

How to Perform a Repeated Measures ANOVA in R?

Authored by
stats writer

December 20, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Perform a Repeated Measures ANOVA in R?*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=108062>

The Repeated Measures ANOVA (RM-ANOVA) is a powerful statistical test essential for analyzing experimental designs where the same subjects are measured under multiple conditions or across various time points. This technique is often preferred over traditional independent-samples ANOVA when dealing with within-subjects factors, as it efficiently controls for individual differences, thereby increasing statistical power. In the context of the R programming language, implementing RM-ANOVA requires careful structuring of the input data and appropriate use of the built-in `aoV()` function to correctly specify the error term associated with the repeated measurements.

Unlike a standard one-way ANOVA, which assumes complete independence between groups, the RM-ANOVA accounts for the correlation between measurements taken from the same individuals. This correlation structure is critical because it reduces the noise (unexplained variance) in the analysis, making it easier to detect a true treatment effect if one exists. Successfully performing this analysis in R involves several key stages: meticulous data preparation into a long-format data frame, model specification using the `aoV()` function with the appropriate error structure, and finally, detailed interpretation of the resulting summary statistics, often followed by post-hoc testing.

This comprehensive guide details the necessary steps for conducting a one-way repeated measures ANOVA in R, ensuring that the syntax correctly handles the within-subject design. We will walk through data entry, model fitting, and the interpretation of the output, focusing on how R separates the variance components attributable to the subjects themselves versus the variance attributable to the experimental treatments.

A repeated measures ANOVA is specifically designed for experiments where every participant experiences all levels of the independent variable. This design minimizes subject-to-subject variability, which is a major source of error in between-subjects designs. The fundamental goal remains to determine if there is a statistically significant difference among the population means of three or more related groups.

The subsequent sections will provide a practical, step-by-step methodology for executing a one-way repeated measures ANOVA using the R statistical environment, focusing on clarity and reproducibility.

Understanding the Prerequisites and Assumptions

Before proceeding with the analysis, it is vital to ensure that the data meets the necessary assumptions of the RM-ANOVA. While some assumptions are shared with standard ANOVA (like normality of residuals), a unique and crucial assumption for repeated measures designs is **sphericity**. Sphericity refers to the condition where the variances of the differences between all pairs of within-subject conditions (levels of the independent variable) are equal. Violation of this assumption can lead to an inflated F-statistic and an increased risk of Type I error.

If the sphericity assumption is severely violated, standard RM-ANOVA results become unreliable. In such cases, adjustments like the Greenhouse-Geisser or Huynh-Feldt corrections should be applied, or alternative methods like multivariate ANOVA (MANOVA) might be considered. However, in R's basic `anova()` implementation for balanced designs, the output often assumes sphericity; therefore, it is good practice to explicitly test this assumption using specialized packages for rigorous analysis, especially when dealing with more complex designs or larger datasets.

The core assumptions for a valid RM-ANOVA include: 1) The dependent variable must be measured at the interval or ratio level. 2) The data must be sampled independently across participants (though not across conditions). 3) The distribution of the dependent variable should be approximately normal for each condition. 4) The assumption of sphericity must be met, particularly when there are three or more levels of the within-subjects factor.

Setting Up the Research Scenario

To illustrate the application of RM-ANOVA, let us consider a typical experimental design. A team of researchers is investigating the efficacy of four distinct pharmaceutical formulations--referred to here simply as Drug 1, Drug 2, Drug 3, and Drug 4--on the reaction time of human subjects. The primary objective is to determine if the mean reaction times are significantly different depending on which drug is administered. Since reaction time is the outcome variable measured on a continuous scale, it serves as our dependent variable.

Crucially, this experiment employs a **within-subjects design**: the same five patients participate in all four experimental conditions. Each patient receives all four drugs, typically separated by a sufficient washout period to prevent carryover effects. Since each patient is measured on each of the four drugs, the measurements are dependent, necessitating the use of the repeated measures ANOVA. The total sample size for the analysis, therefore, is 20 observations (5 patients \times 4 drugs).

The statistical hypothesis framework for this experiment is established as follows. We are testing against the baseline assumption that all drugs have an equivalent effect: the Null Hypothesis (H₀) states that the population mean reaction times are equal across all four drugs ($\mu_1 = \mu_2 = \mu_3 = \mu_4$). The Alternative Hypothesis (**H_a**) posits that at least one drug yields a significantly different mean reaction time from the others. Our analysis in R will determine whether we have sufficient evidence to reject **H₀**.

Step 1: Preparing and Entering Data in R

The first critical step in R is correctly structuring the data. For repeated measures analysis using the standard `anova()` function, the data must be in the "long" format. This means that each row represents a single measurement, and we need separate columns to identify the subject (patient),

the condition (drug), and the measured response time. This structure is essential for R to correctly distinguish between the variance due to the treatment (Drug) and the variance due to the individual subject (Patient).

We use the `data.frame()` function to construct our data frame, ensuring that the `patient` variable is correctly repeated to correspond to the four drug measurements taken from each individual, and the `drug` variable cycles through the four levels for each patient. The `response` variable holds the outcome measurement (reaction time). The resulting data object, named `df`, contains 20 rows and three columns.

Note the use of the `rep()` function in R, which is key to generating the structure efficiently. We repeat the patient identifiers (1 through 5) four times *each* to indicate the four measurements per patient, and we repeat the drug levels (1 through 4) five times overall (once for each patient) using the `times` argument. This coding ensures the data matrix accurately reflects the dependency inherent in the repeated measures design.

#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

Step 2: Executing the RM-ANOVA Model (using aov())

Once the data is correctly formatted, we proceed to fit the RM-ANOVA model using R's base function, `aov()`. The syntax for a repeated measures model is slightly specialized because we must explicitly instruct R to partition the error variance into two components: the variance associated with individual differences (between-subjects error) and the remaining variance used to test the within-subjects factor (drug effect).

The core of the model specification is the `Error()` term. In our model, `aov(response ~ factor(drug) + Error(factor(patient)))`, the term `factor(drug)` represents the within-subjects independent variable whose effect we are testing. The critical part is `Error(factor(patient))`. This instructs R to treat `patient` as a blocking factor. It tells the function to remove the variability that can be attributed solely to differences between patients before calculating the F-ratio for the drug effect.

It is important that both the independent variable (`drug`) and the subject identifier (`patient`) are treated as factors using the `factor()` function. If they are left as numeric integers, R might incorrectly treat them as continuous variables, leading to an invalid model calculation. By isolating the individual subject variance, the residual error used for testing the drug effect is smaller, enhancing the power of the test.

The code below demonstrates how to fit and display the summary of the repeated measures ANOVA model:

```
#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 Output and Hypothesis Testing

The output generated by `summary(model)` is partitioned into two distinct sections, corresponding to the error terms defined in the model syntax. The first section, labeled "Error: factor(patient)," represents the between-subjects variance. Since this variance component is entirely removed from the test of the treatment effect, no F-statistic or P-value is calculated for the patient factor itself; it is only used to establish the degrees of freedom for the model.

The second and most critical section is labeled "Error: Within." This portion contains the results for the within-subjects factor, `factor(drug)`. It shows how the remaining variance (after accounting for patient differences) is partitioned between the drug effect and the true residual error. The analysis relies on comparing the variability explained by the treatment (Mean Sq for drug) against the unexplained, remaining error variance (Mean Sq for Residuals).

In our example, the key statistics for the treatment effect of the drug are the Degrees of Freedom ($Df = 3$), the Mean Square (Mean Sq = 232.7), the F-statistic (F value = 24.76), and the associated P-value ($Pr(>F) = 1.99e-05$). To make a decision regarding the Null Hypothesis (H₀), we compare the P-value to our predetermined significance level, typically $\alpha = 0.05$.

With an observed P-value of 1.99e-05, this value is substantially smaller than 0.05. Therefore, we firmly reject the Null Hypothesis and conclude that there is a statistically significant difference in mean reaction times attributable to the type of drug administered. The extremely high F-statistic (24.76) relative to the small residual mean square confirms a strong effect.

Step 4: Reporting Statistical Findings

Reporting the results of a Repeated Measures ANOVA requires adherence to standard statistical reporting guidelines, typically following the format used in psychological or scientific journals (e.g., APA style). The report must include the type of test conducted, the independent and dependent variables, the calculated F-statistic, the degrees of freedom for both the numerator and the denominator, and the exact P-value.

When reporting, the degrees of freedom are presented in parentheses as $F(\text{df error numerator}, \text{df error denominator})$. In our output, the drug effect has 3 degrees of freedom, and the corresponding residuals (Error: Within) have 12 degrees of freedom. The resulting F-ratio is 24.76. Since the P-value is very small, we typically report it as **$p < 0.001$** rather than the highly precise scientific notation.

A clear, concise reporting statement summarizing the findings is essential for communicating the impact of the treatment. This statement confirms the significance and contextualizes the statistical outcome back into the research question, providing the critical link between the mathematical result and the pharmacological reality being studied.

Here is the standard exemplary format for reporting the findings:

A one-way repeated measures ANOVA was conducted on five individuals to examine the effect that four different drugs had on response time.

Results showed that the type of drug used lead to statistically significant differences in response time ($F(3, 12) = 24.76, p < 0.001$).

Conducting Post-Hoc Analysis for Specific Comparisons

While the RM-ANOVA confirms that a difference exists among the four drug means, it does not specify which pairs of drugs are significantly different from each other. To pinpoint these specific differences, we must perform a **post-hoc test**. Common post-hoc procedures include Tukey's Honest Significant Difference (HSD) test or the Bonferroni correction for multiple comparisons.

For repeated measures designs, performing post-hoc comparisons using the standard Tukey HSD on the overall model summary can be misleading because it often does not correctly use the within-subjects error term. A more robust approach involves using the `pairwise.t.test()` function in R, which allows the use of the `p.adjust.method` argument (e.g., "bonferroni") to control the family-wise error rate, or using specialized packages designed for within-subjects contrasts, such as `emmeans` or `rstatix`.

The goal of the post-hoc test is to compare every possible pair of drug means (Drug 1 vs. Drug 2, Drug 1 vs. Drug 3, etc.) while maintaining an acceptable level of error control. By performing these targeted comparisons, researchers can determine exactly which drugs produce the fastest or slowest reaction times and thus derive clinically relevant conclusions from the overall significant ANOVA result.

Related Statistical Resources

Further exploration of Repeated Measures ANOVA and related methods is 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](#)

ARABPSYCHOLOGY.COM