

Existing non-parametric alternatives to two-way ANOVA

Non-parametric two-way ANOVA is a statistical method used to analyze data with two categorical independent variables (factors) and one continuous dependent variable, where the assumptions of normality and homogeneity of variance are not met.

1. Scheirer-Ray-Hare test

The Scheirer-Ray-Hare test is a nonparametric test used for a two way factorial design investigating that there are no interactions between the two factors. The test is named after James Scheirer, William Ray and Nathan Hare, who published it in 1976. This is the two factor version of the Kruskal-Wallis test. It has been suggested that the observations should be balanced and that each cell in the interaction should have at least five observations. Note that for unbalanced designs, the `scheirerRayHare` function uses a type-II sum-of-squares approach by default. There is an option to use type-I sum-of-squares.

Note: Type I SS is calculated sequentially, meaning the factors are added to the model one by one in a specific order. The sum of squares for each factor is computed after the previous factors have already been included. Type II SS calculates the effect of each factor while adjusting for the other factors in the model. It doesn't depend on the order in which the factors are added to the model.

Appropriate data

- Two-way data arranged in a factorial design
- Dependent variable is ordinal, interval, or ratio
- There are two treatment or group independent variables. Each is a factor with two or more levels
- Observations are independent. That is, they are not paired or repeated measures

Mathematical foundation of the SRH process

1. Rank the Data

Combine all the observations from the entire dataset into a single list.

Assign ranks to the observations, with the smallest value receiving rank 1. If there are ties, assign the average rank to the tied values.

Let Y_{ijk} be the observed value for the k -th replicate in the i -th level of factor A and the j -th level of factor B . The ranks R_{ijk} are computed for all observations.

- ##### 2. Compute the Sum of Ranks for Each Group
- For each combination of factor levels (i.e., each cell in the design), compute the sum of ranks:

$$SS_{ij} = \sum_{k=1}^{n_{ij}} R_{ijk}$$

where n_{ij} is the number of observations in the i -th level of factor A and the j -th level of factor B .

3. Total Sum of squares and sum of squares of main effects and interaction

$$SST = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ij}} R_{ijk}^2 - \frac{N(N+1)^2}{4}$$

$$SSA = \sum_{i=1}^a \frac{(\sum_{j=1}^b S S_{ij})^2}{b \cdot n_{i.}} - \frac{N(N+1)^2}{4}$$

$$SSB = \sum_{j=1}^b \frac{(\sum_{i=1}^a S S_{ij})^2}{a \cdot n_{.j}} - \frac{N(N+1)^2}{4}$$

$$H_{AB} = \frac{12}{N(N+1)} \sum_{i=1}^a \sum_{j=1}^b \frac{R_{ij}^2}{n_{ij}} - 3(N+1)$$

where A and B are the main factors

N is the total no of observations , a- no of levels in factor A , b- no of levels in factor B .

In R, the Scheirer–Ray–Hare (SRH) test can be performed using the rcompanion package with the scheirerRayHare() function. The test has low statistical power which may not be best approach for detecting interactions. This is because it applies the Kruskal-Wallis test structure to factorial designs. Works best with balanced designs

ref: https://rcompanion.org/handbook/F_14.html

practical example:

```
### Assemble the data

Location = c(rep("Olympia" , 6), rep("Ventura", 6),
             rep("Northampton", 6), rep("Burlington", 6))
Tribe = c(rep(c("Jedi", "Sith"), 12))
Midichlorians = c(10, 4, 12, 5, 15, 4, 15, 9, 15, 11, 18, 12,
                  8, 13, 8, 15, 10, 17, 22, 22, 20, 22, 20, 25)
Data = data.frame(Tribe, Location, Midichlorians)

### Scheirer-Ray-Hare test
#install.packages("rcompanion")
library(rcompanion)

Warning: package 'rcompanion' was built under R version 4.4.3
```

```
scheirerRayHare(Midichlorians ~ Tribe + Location, data = Data)
```

DV: Midichlorians

Observations: 24

D: 0.9917391

MS total: 50

	Df	Sum Sq	H	p.value
Tribe	1	8.17	0.1647	0.68487
Location	3	746.58	15.0560	0.00177
Tribe:Location	3	315.58	6.3642	0.09517
Residuals	16	70.17		

2. Rank Transform Approach (RT) in factorial designs.

The rank transform procedure as proposed by Iman and Conover (1976) is carried out by replacing original observations with their respective ranks, computing parametric tests on these ranks (typically an analysis of variance), and referring the values of the test statistics so obtained to the usual table of critical values. That is the major steps of the RT process.

The authors found that rank transform tests perform similarly to normal theory tests when data follows a normal distribution. However, rank transform tests have a significant power advantage when the data is non-normal.

Limitations of RT approach

In certain cases, the validity of the rank transform procedure applied to factorial designs with interactions may be questioned.

when applying non-linear transformations such as rank transform are made on set of data, main effects and interactions effects in original data may or may not exist in transformed data.

The rank transform provides misleading results in testing significance of interactions when there are large main effects involving both factors but no true interactions.

ref: <https://pdfs.semanticscholar.org/43c3/b80ef17d98180bcaa6c95339064901242eb9.pdf>

practical approach :

```
library(dplyr)
```

```
Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:stats':
```

```
filter, lag
```

```
The following objects are masked from 'package:base':
```

```
intersect, setdiff, setequal, union
```

```
# the dataset
Location <- c(rep("Olympia", 6), rep("Ventura", 6),
              rep("Northampton", 6), rep("Burlington", 6))
Tribe <- c(rep(c("Jedi", "Sith"), 12))
Midichlorians <- c(10, 4, 12, 5, 15, 4, 15, 9, 15, 11, 18, 12,
                  8, 13, 8, 15, 10, 17, 22, 22, 20, 22, 20, 25)
data <- data.frame(Tribe, Location, Midichlorians)
data$Tribe <- as.factor(data$Tribe)
data$Location <- as.factor(data$Location)
# Rank the responses
data <- data %>%
  mutate(Ranked_Midichlorians = rank(Midichlorians))

# Perform ANOVA on the ranked data
rt_anova <- aov(Ranked_Midichlorians ~ Tribe * Location, data = data)
summary(rt_anova)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Tribe	1	8.2	8.17	1.862	0.191
Location	3	746.6	248.86	56.747	9.52e-09 ***
Tribe:Location	3	315.6	105.19	23.987	3.66e-06 ***
Residuals	16	70.2	4.39		

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

3. Aligned Rank Transform (ART) ANOVA

The **Aligned Rank Transformation (ART) ANOVA** is a non-parametric procedure used when the assumptions of parametric ANOVA (e.g., normality and homogeneity of variance) are violated. It allows for factorial designs with interactions by aligning and ranking data before applying standard ANOVA.

ART ANOVA assumptions

- Independence of observations
- No strict normality requirement
- Adequate sample size

The Process of ART

For two way layout, the general steps to do aligned rank tests are as follows:

1. Estimate the effects on which the aligning is to be done. Subtract the estimated effects from the original scores (align the scores).

For a two-way ANOVA with factors A and B , the observed data Y_{ijk} (where i is the level of factor A , j is the level of factor B , and k is the replicate) is aligned to remove the main effects of A and B .

The aligned data Y_{ijk}^* is computed as:

$$Y_{ijk}^* = Y_{ijk} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$$

where: $\bar{Y}_{i..}$ is the mean for level i of factor A , $\bar{Y}_{.j.}$ is the mean for level j of factor B , $\bar{Y}_{...}$ is the overall mean.

2. Rank the aligned scores in the total sample.

The aligned data Y_{ijk}^* is then ranked across all observations. Let R_{ijk} be the rank of Y_{ijk}^* .

3. A traditional two-way ANOVA is performed on the ranks R_{ijk} . The F-tests from this ANOVA are used to assess the significance of the main effects and interaction.
 - The ARTool package in R makes it easy to implement ART for two-way ANOVA designs.

```
# Load necessary libraries
library(ARTool)
# Create the dataset
Location <- c(rep("Olympia", 6), rep("Ventura", 6),
              rep("Northampton", 6), rep("Burlington", 6))
Tribe <- c(rep(c("Jedi", "Sith"), 12))
Midichlorians <- c(10, 4, 12, 5, 15, 4, 15, 9, 15, 11, 18, 12,
                  8, 13, 8, 15, 10, 17, 22, 22, 20, 22, 20, 25)

data <- data.frame(Tribe, Location, Midichlorians)
# Convert categorical variables to factors
data$Tribe <- as.factor(data$Tribe)
data$Location <- as.factor(data$Location)

# Apply Aligned Rank Transform (ART)
art_model <- art(Midichlorians ~ Tribe * Location, data = data)
anova_results <- anova(art_model)
print(anova_results)
```

Analysis of Variance of Aligned Rank Transformed Data

Table Type: Anova Table (Type III tests)

Model: No Repeated Measures (lm)

Response: art(Midichlorians)

	Df	Df.res	F value	Pr(>F)
1 Tribe	1	16	3.0606	0.099364 .
2 Location	3	16	34.6201	3.1598e-07 ***
3 Tribe:Location	3	16	29.9354	8.4929e-07 ***

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

Limitations of ART approaches

ART allows for nonparametric testing of interactions and repeated measures without relying on distributional assumptions. When data has many ties, ART replaces them with tied ranks, which may not fully capture the data's structure. For data with extreme skew (e.g., power-law distributions), ART reduces the skew, which could be undesirable if the skew is meaningful. ART works best with completely randomized designs; while it can be applied to other designs, it may not fully account for all effects. The aligned rank transform has been found to perform better than the rank transform, but it still has problems with elevated Type I error rates and with nonnormal error terms.

Ref :

<https://www.jstor.org/stable/1165296> -Non parametric competitors to the two way ANOVA

https://www.researchgate.net/publication/384698597_Nonparametric_tests_for_interaction_in_two-way_ANOVA_with_balanced_replications

4. Permutation tests

A permutation test (also called re-randomization test or shuffle test) is an exact statistical hypothesis test making use of the proof by contradiction. A permutation test involves two or more samples. The null hypothesis is that all samples come from the same distribution. Under the null hypothesis, the distribution of the test statistic is obtained by calculating all possible values of the test statistic under possible rearrangements of the observed data. Permutation tests are therefore a form of re sampling.

Permutation tests can be understood as surrogate data testing where the surrogate data under the null hypothesis are obtained through permutations of the original data.

In other words, the method by which treatments are allocated to subjects in an experimental design is mirrored in the analysis of that design. If the labels are exchangeable under the null hypothesis, then the resulting tests yield exact significance levels; see also changability. Confidence intervals can then be derived from the tests. The theory has evolved from the works of [Ronald Fisher](#) and [E. J. G. Pitman](#) in the 1930s.

Permutation tests exist in many situations where parametric tests do not (e.g., when deriving an optimal test when losses are proportional to the size of an error rather than its square). All simple and many relatively complex parametric tests have a corresponding permutation test version that is defined by using the same test statistic as the parametric test, but obtains the p-value from the sample-specific permutation distribution of that statistic, rather than from the theoretical distribution derived from the parametric assumption.

The major drawbacks to permutation tests are that they

- Can be computationally intensive and may require “custom” code for difficult-to-calculate statistics. This must be rewritten for every case.

- Are primarily used to provide a p-value. The inversion of the test to get confidence regions/intervals requires even more computation.

Based on the concept of synchronized permutations, we introduce an exact permutation solution (Pesarin, 2001; Salmaso, 2003; Basso et al., 2007) for testing for fixed effects in replicated two-way factorial designs with continuous responses. This permutation solution, since it is conditional on a set of sufficient statistics, is a distribution-free nonparametric test. It is worth noting that asymptotically distribution-free (but not exact) tests could also be developed using the approach by Draper (1988) or the recent development of a generalization of the Kruskal-Wallis approach to two- and three-way layouts given in Brunner and Puri (2001). Among the exact tests, we can differentiate synchronized permutation tests from those inspired by the two-way ANOVA F test, such as the tests proposed by Edgington (1995), Maritz (1995), and Sprent (1998).

Let us remember that permutation tests are conditional procedures in which conditioning is with respect to a set of joint sufficient statistics under the null hypothesis. Hence, the permutation approach for a two-way layout should be based on such a set of joint sufficient statistics. We will present the theory in the case of a balanced two-factor design where factor A has I levels and factor B has J levels.

The linear model is;

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk} \text{ where } i=1,\dots,I; j=1,\dots,J; k=1,\dots,n$$

where Y_{ijk} are the experiment response, μ is the general mean, α_i and β_j are the main factors, $\alpha\beta_{ij}$ are interaction effects, ϵ_{ijk} are exchangeable experimental error, with zero mean, from an unknown continuous distribution P and n is the number of replicates in each cell. The usual side conditions are

$$\sum_i \alpha_i = 0, \sum_j \beta_j = 0, \sum_i \alpha\beta_{ij} = 0 \forall j, \sum_j \alpha\beta_{ij} = 0 \forall i$$

Usually, the experimenter's major interest is testing separately for two main effects and interactions. Hence, there are three sub null hypotheses of interest — $H_{0A} : \alpha_i = 0 \forall i$, $H_{0B} : \beta_j = 0 \forall j$, $H_{0AB} : \alpha\beta_{ij} = 0 \forall i, j$ – and the emphasis is on finding three separate and possibly uncorrelated tests. What experimenters are generally looking for is, for instance, to test H_{0A} against $H_{1A} : \text{there exist } i \text{ such that } \alpha_i \text{ is not equal to } 0$, irrespective of whether H_{0B} or H_{0AB} is true or not, etc.

In order to attain this goal within a permutation framework, we must find the proper set of jointly sufficient statistics for all three testing sub problems: H_{0A} irrespective of $H_{0B} \cup H_{0AB}$ is true or not, H_{0B} irrespective of $H_{0A} \cup H_{0AB}$ is true or not, and H_{0AB} irrespective of $H_{0A} \cup H_{0B}$ is true or not. In this framework, this set is $y = [y_{11}, y_{12}, \dots, y_{IJ}]'$, the vector of the observed response partitioned into $I \times J$ blocks. This is due to the definition of sufficient statistic: in fact, two points of the sample space, y and y' , lay in the same orbit of a sufficient statistic if the likelihood ratio

$$\frac{L(\alpha, \beta, \alpha\beta; \mathbf{y})}{L(\alpha, \beta, \alpha\beta; \mathbf{y}')} = \frac{\prod_{ijk} L(\alpha_i, \beta_j, \alpha\beta_{ij}; y_{ijk})}{\prod_{ijk} L(\alpha_i, \beta_j, \alpha\beta_{ij}; y'_{ijk})} = h(\mathbf{y}, \mathbf{y}')$$

does not depend on the parameters $\alpha_i, \beta_j, \alpha\beta_{ij}$. This only occurs if \mathbf{y}' is a permutation of the units within the blocks of \mathbf{y} . From a naive point of view, we are only allowed to permute data within each block, but this is useless since any permutation within blocks gives the same value of any suitable test.

- The function `aovp()` is part of the R package `lmPerm` and stands for “Analysis of Variance with Permutation.” It’s specifically designed for conducting permutation-based analysis of variance tests.

Advantages

- No distributional assumptions
- Valid for small sample sizes
- Robust against outliers
- Maintains Type I error control

Limitations

- Computationally intensive, especially with large datasets
- Requires exchangeability of observations
- Works best with balanced designs
- Testing specific effects requires specialized permutation schemes
- Interpretation differs from parametric ANOVA
- Limited software implementation and standardization
- Post-hoc testing is challenging
- May have limited precision with very small samples

The report explores non-parametric alternatives to two-way ANOVA, such as the Scheirer-Ray-Hare test, Rank Transform Approach, Aligned Rank Transform (ART) ANOVA, and Permutation Tests. These methods are useful when data violates assumptions of normality or equal variances. While they offer flexibility and robustness, each has limitations, such as reduced power for detecting interactions, computational intensity, or challenges with unbalanced data. The choice of method depends on the data structure and research goals, but these alternatives provide valuable tools for analyzing complex datasets without relying on strict parametric assumptions.