Revista Colombiana de Estadística - Applied Statistics
January 2024, volume 47, issue 1, pp. 87–109
http://doi.org/10.15446/rce.v47n1.108065

AovBay: An R Package for Application and Visualization of Parametric, Non-parametric and Bayesian ANOVA

AovBay: un paquete de R para la aplicación y visualización de análisis de varianza paramétrico, no paramétrico y bayesiano

Johny J. Pambabay-Calero[1,a], Mauricio J. Rojas-Campuzano[1,b], Sergio A. Bauz-Olvera[1,c], Omar H. Ruiz-Barzola[2,d]

[1]Faculty of Natural Sciences and Mathematics - ESPOL, Polytechnic University, Guayaquil, Ecuador

[2]Centro de Estudios e Investigaciones Estadísticas - ESPOL, Faculty of Life Sciences, Polytechnic University, Guayaquil, Ecuador

Abstract

The analysis of variance is a statistical technique widely used in the design of experiments and different research areas. It has been modeled using a classical or frequentist approach. With the computational power that is currently available, the Bayesian approach is an essential statistical tool related to hypothesis testing. However, conformity with classical techniques, ignorance of Bayesian statistics, and lack of easy-to-use software are obstacles to its frequent application. In this work, the use of a reproducible statistical package in R is proposed. It presents options to perform an analysis of variance in a classical (frequentist) and Bayesian way when the assumptions of the frequentist approach are not met or when a level of more specific inference such as quantifying the evidence provided by a data set for a given hypothesis, with the possibility of contributing to the understanding of the rejection or not of the statistical hypotheses raised, through interactive graphics presented in an emerging Shiny panel.

***Key words*:** ANOVA; Bayesian; Non-parametric; Parametric; R Software.

[a]Ph.D. E-mail: jpambaba@espol.edu.ec
[b]Ph.D. E-mail: maujroja@espol.edu.ec
[c]Ph.D. E-mail: serabauz@espol.edu.ec
[d]Ph.D. E-mail: oruiz@espol.edu.ec

Resumen

El análisis de varianza es una técnica estadística ampliamente utilizada en el diseño de experimentos y diferentes áreas de investigación. Ha sido modelado utilizando un enfoque clásico o frecuentista. Con el poder computacional que se tiene actualmente, el enfoque bayesiano es una herramienta estadística esencial relacionada con las pruebas de hipótesis. Sin embargo, la conformidad con técnicas clásicas, el desconocimiento de la estadística bayesiana y la falta de software fácil de usar son obstáculos para su aplicación frecuente. En este trabajo, se propone el uso de un paquete estadístico reproducible en R. Presenta opciones para realizar un análisis de varianza de manera clásica (frecuentista) y bayesiana cuando no se cumplen los supuestos del enfoque frecuentista o cuando se requiere un nivel de inferencia más específico, como cuantificar la evidencia proporcionada por un conjunto de datos para una hipótesis dada, con la posibilidad de contribuir a la comprensión del rechazo o no de las hipótesis estadísticas planteadas, a través de gráficos interactivos presentados en un panel Shiny emergente.

***Palabras clave*:** ANOVA; Bayesiano; No paramétrico; Paramétrico; Software R.

1. Introduction

In the realm of statistical methodologies, numerous techniques, each with distinct specifications, assumptions, advantages, and disadvantages, have been scrutinized based on their applicability to specific data frames; notably, the analysis of variance (ANOVA) stands out, maintaining its position as a prevailing mode (Kozak & Piepho, 2018). It is important to remember the variability that exists between the profiles of the users of the model since, just as researchers or statisticians can use it, there are many cases in which people, who are not used to studying statistical theory, use it as well. This is a potential reason for ANOVA to be subject to multiple violations in its assumptions. It is generally assumed that the residuals are normally distributed with mean 0 and variance σ^2 to estimate the unknown parameters in the model (Çelik & Şenoğlu, 2018). This is probably the best-known assumption, yet its validation is often mishandled. It is common to see normality tests applied to raw data and not to residuals (Kozak & Piepho, 2018). Similarly, there are cases in which the assumption of homoscedasticity is not evaluated in the best way. Other scenarios to consider are potential outliers (Barnett et al., 1994) and missing values (Rubin, 1987). Once the test is shown to apply to the data frame, reliable inferences can be made by validating the assumptions.

If the assumptions are not met, especially the assumption of normality, the immediate path that researchers usually take is to choose one of the non-parametric tests (Flores et al., 2017). In the case of ANOVA, Kruskal-Wallis is used for the same purposes. It does not assume normality in the residuals (Ostertagova et al., 2014). However, these non-parametric tests, although it is said that they are not linked to specific conditions, have a fundamental assumption, symmetry, which is often not considered (Duller & Vorhauer, 2020). If the mean, median, and mode

have the same value, it can be asserted that the statistical distribution of the data is symmetric, which speeds up statistical inferences. Thus, applying tests with medians and ranges becomes a plausible choice. If this is not true, it is good practice to look for other methods that supply more reliable inferences. In addition to this, it is known that non-parametric tests, as they belong to classical statistics, use the p-value as a potential conclusion point. Increasing the sample size may suggest statistical significance due to some differences, although they may be small (Jiménez-Paneque, 2016).

While traditional frequentist statistics pose certain challenges, especially in complex experimental designs, there are many other ways to approach analysis of variance use cases. The Bayesian approach to variance analysis not only addresses these inherent difficulties but also provides a unified framework for understanding data variability and uncertainty. As highlighted by Wedel & Dong (2020), this method offers a holistic perspective, bridging gaps that are often encountered in classical statistical methodologies

Furthermore, Bayesian inference provides more opportunities for users allowing knowledge to be updated based on data and previous knowledge about a specific case. You can even monitor the evidence until the result is convincing enough or the available resources have been exhausted (Wagenmakers et al., 2018). Nevertheless, using Bayesian inference techniques is less frequent than using classical inference (Pambabay-Calero et al., 2020). It may be due to three causes:

- That the data meet the necessary assumptions for the correct use of classic ANOVA
- Lack of knowledge of Bayesian models.
- Absence of an easy-to-use computational complement.

These three points are not exclusive because the probability that someone will feel satisfied with the classical methods can increase if ignorance of the Bayesian approach methods or if no theoretical knowledge of software channels is added. It allows it to be put into practice with relative ease. It is known that the Bayesian analysis of variance is better addresses the difficulties caused by the lack of validity in the assumptions of classical statistics. The first cause presented above can already be discarded; more advantages are added to this. While the classic ANOVA concludes considering the p-value and the level of significance determined, having as options *Reject the null hypothesis* or *there is no statistical evidence to reject the null hypothesis*. The Bayesian approach allows quantifying how much evidence a data set provides for a hypothesis (Wedel & Dong, 2020). The Bayes factor quantifies the relative predictive performance of rival models. In addition, there is no need to limit the value of the Bayes factor. Finally, unlike the p-value, with the Bayes factor, there is no need to establish a critical point or rejection zone to perform a hypothesis test (van den Bergh et al., 2020). However, it is necessary to include values that help the interpretation of the Bayes factor.

A significant impediment to the widespread adoption of Bayesian tests is possibly the lack of easy- to-use software for mainstream statistical problems

(Wagenmakers et al., 2018). Currently, many researchers or people seeking the application of analysis of variance, with sufficient theoretical knowledge or not, have some tools at their disposal, the *aov* function of the **stats** package in R (Collyer & Adams, 2018), the analysis of variance module of STATA (StataCorp., 2019), the ANOVA procedure in SAS.(SAS Institute Inc., 1999), among others. For this article, we will use the statistical software R (R Core Team, 2012). The idea of acquiring other statistical programs with Bayesian inference implements, such as JASP (van den Bergh et al., 2020), is not unreasonable. However, what would happen if an interactive statistical complement that can be reproduced from R for the Bayesian ANOVA test is presented? The likelihood that the person reading this article will be completely satisfied with the classical methods may be reduced. Until recently, these tests had not been implemented in any software, much less user-friendly software (Wagenmakers et al., 2018). The question stops being in a utopian sense since that complement now exists, the interactivity that it possesses makes it friendly, and the Bayesian procedures make it reliable.

An example is presented with the database of an experiment where the capacity of the vetiver plant (*Chrysopogon zizanioides*) was analyzed to remove five pharmaceutical products classified as contaminants from the water (Checa-Artos et al., 2021), For this example, only the contaminating substance *SULFAMETHOXAZOLE* will be considered. The diagnosis of the statistical tests of these data is presented in Table 1. For the decision, an α (Type 1 error) $= 0.05$ is considered.

Table 1: Statistical diagnosis for the different tests

Assumptions	Test	p-value	Value	Decision
Normality of residuals	Kolmogorov-Smirnov	0.3832		Is met
	Shapiro-Wilk	0.3865		
Homoscedasticity of the residuals	Bartlett Test	$< 2.2x10^{-16}$		Is not met
Symmetry of the residuals			0.1290*	Is not met
Independence of the residuals	Durbin Watson	0.006		Is not met

* Asymmetry coefficient

To be able to apply classic ANOVA, the assumptions in Table 1 must be met. However, this is not the case since the only assumption met is normality, while the assumptions of homoscedasticity, symmetry, and independence are not met. Therefore, to decide which path to take, given the inconvenience of not meeting the assumptions, a decision diagram was designed, systematically facilitating the statistical technique selection, depending on the compliance or non-compliance of assumptions.

Figure 1 shows a flowchart, where the assumptions are mentioned and the available ways according to their meeting are established, ending with the technique applicable to the data frame.

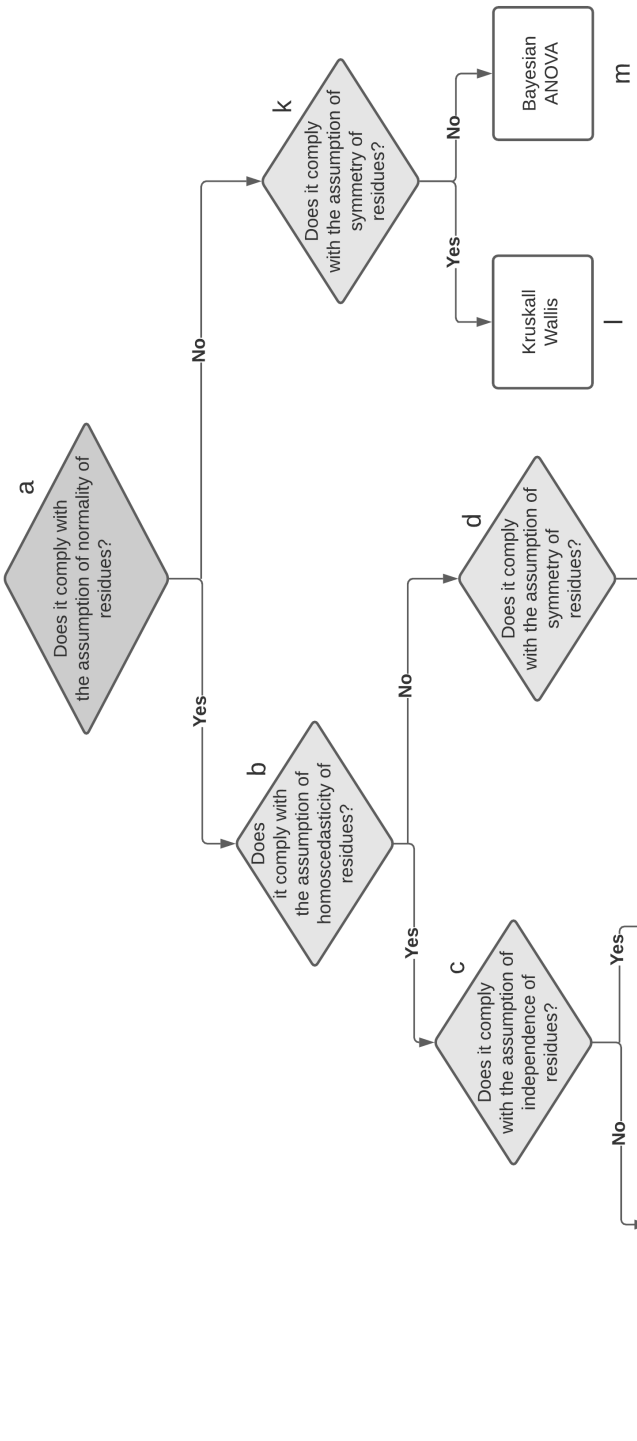

FIGURE 1: Flowchart: Assumptions and statistical techniques

The journey begins at the decision diamond **a**, where the assumption of normality of the residuals is questioned. If met, we proceed to diamond **b** where the assumption of homoscedasticity of the residuals is asked if met. If this assumption is validated, we move on to diamond **c**, where the assumption of independence of the residuals is questioned. If the assumption is favorable, it is good practice to apply the classic ANOVA model. However, if the assumption of independence, homoscedasticity, or normality, mentioned in the decision diamonds **c**, **b** and **a** respectively, are not met, we evaluate the assumption of symmetry of the residuals (diamonds **d**, **e** and **k**), If the assumption is met, the recommended technique is Kruskal Wallis; if not met, an analysis of variance from the Bayesian approach is advised. It is important to mention that the process can start directly with Bayesian ANOVA to avoid evaluating assumptions.

The most appropriate technique to be used will be decided using Table 1 and the flow chart, reviewed in Figure 1, subjecting the data set systematically to the assumptions suggested in the graph. We begin by evaluating the assumption of normality of the residual, the diagnosis of the data frame, observed in Table 1, indicates that it is met. This answer leads to the next question, which deals with homoscedasticity. Table 1 shows that this assumption is not met. From this point, the classic ANOVA application is dropped. Following the path suggested by the diagram is the symmetry assumption, which is not met either. Therefore the non-parametric equivalent, Kruskal Wallis, should not be considered. Finally, the path suggested by the diagram is applying the analysis of variance with a Bayesian approach.

In the following, Section 2, "Materials and Methods", examines the classical ANOVA model, the Bayesian ANOVA model, and lays the foundation for an illustrative example. Subsequently, in Section 3, the package, its availability, and functions are introduced. In Section 4, the main results obtained from analyzing the data set of the illustrative example using **AovBay** are presented. Finally, Section 5 provides a discussion, followed by an annex in Section Appendix B.

2. Materials and Methods

2.1. One-Way ANOVA

When we only have one factor, the ANOVA model can be denoted as follows:

$$Y_{ij} = \mu_i + \epsilon_{ij} \quad ; \quad i = 1, \ldots, k \quad , \quad j = 1, \ldots, n_i \tag{1}$$

where

- i represents the ith treatment of the factor, and the total number of treatments is denoted by k.
- j represents observation j, the maximum number of j is n_i, meaning the total number of observations in treatment i.

- Y_{ij} is the value taken by the response variable in the jth observation of the ith treatment.
- μ_i is the mean of the ith treatment.
- ϵ_{ij} is the random error, independent and identically distributed normally with mean 0 and variance σ^2.
- $n = n_1 + n_2 + \cdots + n_k$.

The classical analysis of variance focuses on testing the following hypotheses:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k = \mu \qquad \text{versus} \qquad H_1 : \mu_i \neq \mu \text{ for some } i$$

The alternate hypothesis indicates that at least one μ_i (Treatment mean i) is different from the others.

The analysis of variance model compares the variations between groups (treatments) and within groups (Ostertagova et al., 2013). If there is no effect of the factor in the response variable, these variations will not be different. Otherwise, the difference between them will be reflected in the F statistic, and the null hypothesis will be rejected.

To attain the F, statistic, it is necessary to obtain the variation sources and the degrees of freedom. The square sum of $SSTR$ treatments with $k-1$ degrees of freedom, the square sum of the SSE residuals with $n-k$ degrees of freedom, and the total square sum $SSTO$ with $n-1$ degrees of freedom. The F statistic is as follows:

$$F = \frac{SSTR/(k-1)}{SSE/(n-k)} \tag{2}$$

It is known that under the null hypothesis H_0 the test statistic F follows an $F_{k-1,n-k}$, distribution and under the alternate hypothesis as a non-central $F_{k-1,n-k,\lambda}$, is distributed where λ is the decentralization parameter (Solari et al., 2008).

2.2. Bayesian ANOVA

The Bayesian paradigm, rooted in Bayes' theorem, provides a statistical approach that facilitates the updating of probabilities as new evidence emerges. Contrary to frequentist statistics, which focuses on observed frequencies, Bayesian statistics interprets probability as a degree of belief or certainty. This perspective allows for the integration of prior knowledge through prior distributions and their subsequent updating with observed data to derive posterior distributions. While it offers advantages in terms of flexibility and quantifying uncertainty, it also poses challenges, such as the requirement for advanced computational techniques and the inherent subjectivity in selecting prior distributions (Rendón-Macías et al., 2018).

Bayesian inference's main objective is to update knowledge through observations, which can be called *learning* (van den Bergh et al., 2020). In the one-way analysis of variance, the null hypothesis is initially proposed. This hypothesis

indicates that there is no effect on the mean of the distribution of the response variable or on the alternative hypothesis, which considers the scenarios in which this does not happen. With the Bayesian ANOVA, it is possible to determine which hypothesis is most probable by comparing the subsequent probabilities of each candidate model. If the scope is limited to one-way models, the candidate models are the null model Θ_0, which considers a single constant mean, and the model that considers the effect of the treatments on the dependent variable Θ_1. To reach the subsequent model comparison, one must start from the prior knowledge and include the plausibility of the data.

2.2.1. Prior

Once the hypotheses are identified, their previous or prior probability $\pi(\Theta_i)$ is determined. Having two hypotheses, the ratio between them is usually calculated:

$$\frac{\pi(\Theta_1)}{\pi(\Theta_0)} \tag{3}$$

so that $\pi(\Theta_1)$ and $\pi(\Theta_0)$ are the prior probabilities of the alternative and null hypotheses, respectively. Having this ratio, there are three scenarios (Rendón-Macías et al., 2018).

1. The ratio is equal to 1; this implies that the two hypotheses are equally probable.

2. The ratio is greater than 1; previous information indicates that the alternative hypothesis is more likely.

3. The ratio is less than 1; it is previously known that the null hypothesis is more likely.

2.2.2. Likelihood

Once the prior probabilities are in place, to update the knowledge, the information provided by the data (y), the sampling distribution or distribution of the data $p(y|\Theta_i)$ are required. The Bayesian inference indicates that two probability models $p(y|\Theta)$ with the same likelihood function produce the same inference for any hypothesis, under the assumption of equal prior probabilities.

2.2.3. Posterior

If we condition the known value of the data, using the conditional probability theory or Bayes rule, the posterior density $p(\Theta|y)$ can be obtained (Gelman et al., 2013).

$$p(\Theta|y) \propto \pi(\Theta)p(y|\Theta) \tag{4}$$

where, $\pi(\Theta)$ is the prior probability and $p(y|\Theta)$ the likelihood.

2.2.4. Bayes Factor

The Bayes factor (FB), analogous to the likelihood ratio, uses prior knowledge to transform them into inferences, that is, to determine that a conclusion is correct or incorrect based on probabilities (Goodman, 1999). Being a ratio, it can be expressed as follows:

$$FB_{01} = \frac{p(y|\Theta_0)}{p(y|\Theta_1)} \tag{5}$$

Where $p(y|\Theta_0)$ and $p(y|\Theta_1)$ are the data distributions under the null and alternate hypotheses, respectively.

If FB_{10}, is required, just get the reciprocal of FB_{10}, $FB_{01} = \frac{1}{FB_{10}}$ (Rouder et al., 2012).

This likelihood ratio can be used to find a ratio that involves further knowledge by looking at the equation 4.

$$\frac{p(\Theta_0|y)}{p(\Theta_1|y)} = \frac{p(y|\Theta_0)}{p(y|\Theta_1)}\frac{\pi(\Theta_0)}{\pi(\Theta_1)} = FB_{01}\frac{\pi(\Theta_0)}{\pi(\Theta_1)} \tag{6}$$

where $\frac{\pi(\Theta_0)}{\pi(\Theta_1)}$ is the ratio stated in equation 3.

In this way, it can be concluded with the subsequent knowledge. The Bayes factor is responsible for the change in beliefs, reflected in the ratio of the posterior probabilities (Rouder et al., 2012).

The Bayesian Factor, pivotal in Bayesian statistics, facilitates a quantitative juxtaposition of evidence between competing hypotheses, capitalizing on its ability to assimilate prior information and its detachment from p-values. Nonetheless, its deployment can be computationally demanding, and the subjective selection of prior distributions, coupled with the imperative for adept interpretation and the potential for inconclusive outcomes, mandates that scholars approach this instrument with discernment and profound acumen (Ramos-Vera, 2021; Guillen & Chaparro, 2021; Lintusaari et al., 2017)

2.3. Illustrative Example

We used a study in which five pharmaceutical products, aqueous solutions were prepared using ultrapure methanol at four concentrations: 3 mg/L, 6 mg/L, 9 mg/L, 12 mg/L and four sampling times. Later the vetiver species was introduced (*Chrysopogon zizanioides*), the experiment was carried out under specific conditions. The response variable is the percentage of removal of these pollutants. If the reader wants to delve into this investigation, carried out with response surfaces, see Checa-Artos et al. (2021).

In this example, the polluting substance *SULFAMETOXAZOLE* is used. The concentration of the pollutant (in ppm) is delimited as a factor.

Table 2 shows that it is not a balanced design since the number of experimental units is not the same at each factor level. Therefore, to perform the complete analysis of this data set, the **AovBay**, package will be used, described below.

Table 2: Illustrative example data description

Treatments	Experimental units
Concentration 3	4
Concentration 6	2
Concentration 9	2
Concentration 12	4

3. The Package

3.1. Availability

The package **AovBay** (Rojas-Campuzano & Pambabay-Calero, 2021) is available from the Comprehensive R Archive Network at `https://cran.r-project.org/web/packages/AovBay/`. It can be downloaded as follows:

```
install.packages('AovBay')
```

3.2. Functions and Arguments

It includes the *aovbayes()*, function, which only has the argument **data**, corresponding to the data set. The function can be called as follows:

```
library(AovBay)
data("PollutionData")
aovbayes(PollutionData)
```

The output of this statement is a shiny panel, where a window dedicated to loading or updating the data set is displayed. This window can be seen in Figure 2.

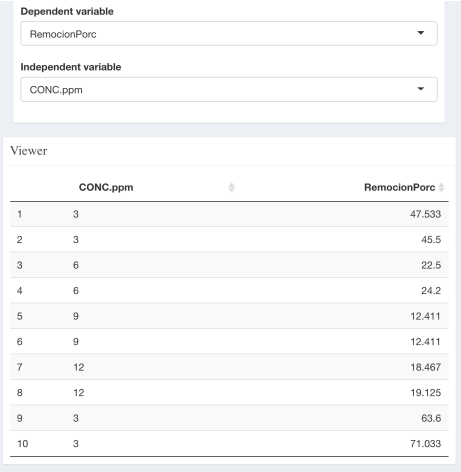

Figure 2: Partial data upload window, upload from the function

It is even possible to access the shiny panel, calling the function without specifying any argument, as follows:

```
library(AovBay)
aovbayes()
```

The user can enter the database from the panel without entering it in R. The output can be seen in Figure 3.

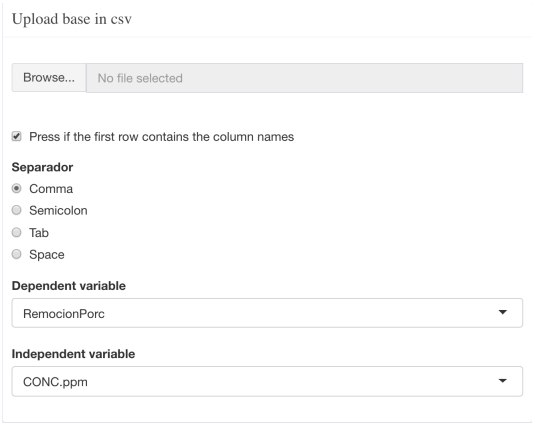

FIGURE 3: Data upload window, upload from the panel

The function does not contain more arguments; however, this does not mean that the user receives a fixed, unchangeable output. It is possible to change arguments in the statistical tests or simulations within the pop-up application in a friendly and reactive way. The latter implies that if any argument is changed, the output will be refreshed and presented based on the new argument without the need to rerun the function.

A menu is presented, where the assumptions, a window with the output of the classic ANOVA, another with Kruskal Wallis, and another with Bayesian ANOVA are made available to the user, as seen in Figure 4.

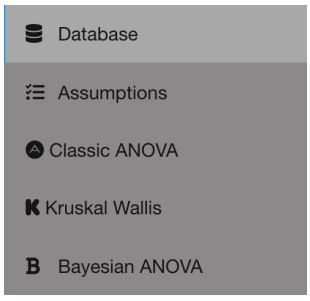

FIGURE 4: Emergent shiny panel menu

4. Results

For the description of the results, the diagram of techniques is used according to the compliance of the assumptions reviewed in Figure 1.

4.1. Assumptions

The package allows the evaluation of assumptions graphically and through statistical tests. First, the conclusion is presented, based on the p-value of the applied test and the α (type 1 error) with which the user wishes to carry out the contrasts. If this value is not to be changed, $\alpha = 0.05$ remains. Subsequently, an account of the validated assumptions is made, and the appropriate technique is suggested.

4.1.1. Normality of Residuals

To evaluate the assumption of normality of the residual, the Kolmogórov-Smirnov test is used with the Lilliefors correction. The Shapiro-Wilk test was used because the sample size is less than 30 (Gandica de Roa, 2020). The *aovbayes()* function evaluates this condition and returns the appropriate residual normality test. Finally, based on the p-value and the α (type 1 error) selected by the user, the decision is presented. Additionally, there is the QQ plot, which contributes to the verification of the inference.

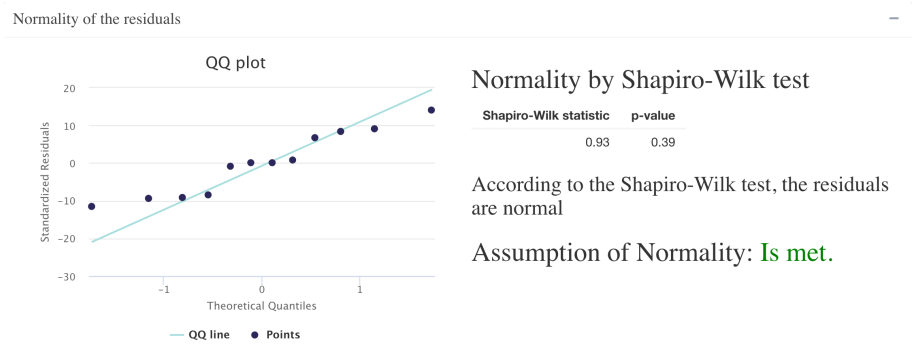

FIGURE 5: Normality of the residual, assumption of the classical approach, package output

Figure 5 presents the output provided by the package when evaluating the normality of the residuals with the data from the illustrative example. The Shapiro-Wilk test suggests that the residuals can come from a normal distribution. The QQ plot shows this inference.

Following the route proposed in Figure 1, the next step is to verify if the assumption of homoscedasticity of the residuals of the ANOVA model, shown in equation 1, is met.

4.1.2. Homoscedasticity of the Residuals

The assumption of homoscedasticity of the residuals is verified with the Bartlett test, which aims to prove that the variances in the k populations, determined by the treatments, are the same (Aslam, 2020). The graph of residuals versus Adjusted Values also identifies possible sections in which the variance is not the same.

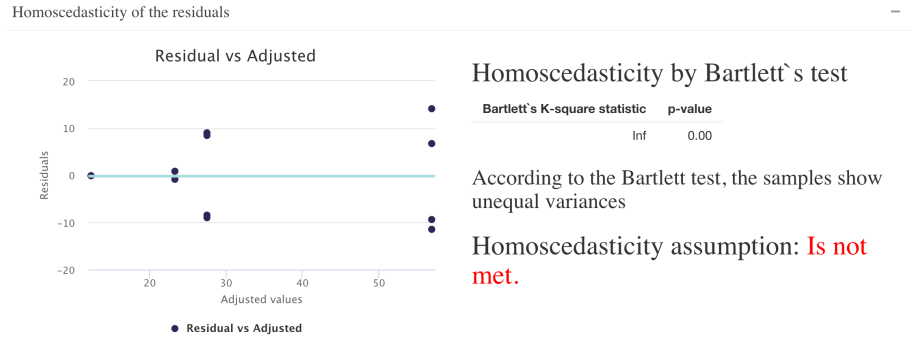

FIGURE 6: Homoscedasticity of the residual, assumption of the classical approach, package output

The result of the Bartlett test is observed in Figure 6. The p-value is very close to zero, concluding that the homoscedasticity assumption is not met. This is confirmed with the graph of residuals versus adjusted values, where the increase in variance is appreciated as one moves to the right, showing heteroscedasticity.

By not meeting this assumption, the diagram presented in Figure 1 suggests testing the assumption of symmetry. However, to finish the assumptions related to parametric statistics, the assumption of independence of the residuals is presented.

4.1.3. Independence of the Residuals

The assumption of independence is verified with the Durbin Watson test. The presence of autocorrelation in the residuals of the linear model is contrasted. This happens when the sequential residuals are correlated so that if the residuals are plotted against time, graphic patterns are obtained. This behavior is called *first-order autoregressive*.

The Durbin Watson test is observed in Figure 7. The p-value is very close to zero; it can be inferred that it is highly probable that there are autocorrelations, that is, that the independence assumption is not met.

4.1.4. Symmetry of the Residuals

The assumption of symmetry is tested with the coefficient of skewness. A symmetric distribution has an asymmetry coefficient equal to 0. Therefore, the cases in which this is not met are considered responsible for the non-validation of

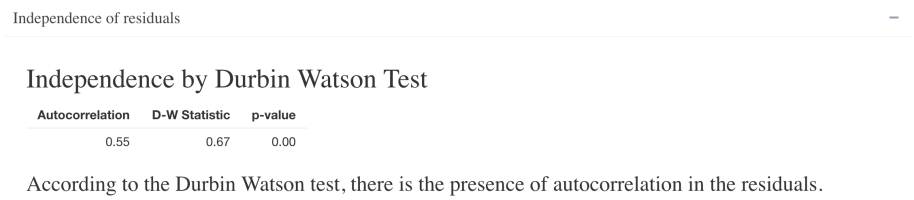

FIGURE 7: Independence of the residual, assumption of the classical approach, package output

the assumption. In addition, the histogram of the residuals is included to verify that the symmetry is present immediately.

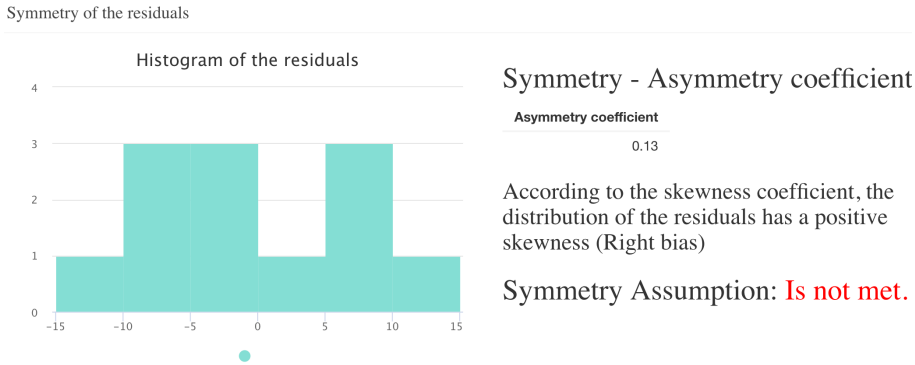

FIGURE 8: Symmetry of the residual, assumption of the non-parametric approach, package output

Figure 8 shows the coefficient of skewness is positive; there is a bias in the right tail. Therefore, the symmetry assumption is not met, and consequently, the use of non-parametric statistics is discarded.

4.2. Appropriate Technique Selection

Once the assumptions have been analyzed, the next step is to select the appropriate technique. To this end, the package provides a suggestion to the user.

FIGURE 9: Recommended technique based on compliance with assumptions, package output

Depending on the compliance with the assumptions, the available techniques, and the most appropriate ones are presented. Figure 9 shows the package suggestion of the most appropriate technique based on the assumptions. The red color indicates that the test is not feasible due to non-compliance with the statistical assumptions. While the green color indicates that making use of this test is feasible.

4.3. Bayesian Analysis of Variance

The packages *BayesFactor* (Morey & Rouder, 2018) and *rstan* (Stan Development Team, 2020). were used to obtain the results of the analysis of variance with a Bayesian approach. The first step is specifying the prior of the alternative model. A scenario is proposed in which it is not known if the factor affects the response variable. Therefore, the prior probability will remain with a 0.5 value. The number of iterations for Monte Carlo[1] , simulations is initially set to 1000 as the minimum reference number (Flegal et al., 2008). However, if the user wants to increase the value of this argument, they have the option of doing so. Similarly, it is possible to choose the number of Markov[2], chains. Increasing this number provides the user with a multiplication effect of the sample size, decreasing the sampling error and the autocorrelation. In addition, this reduces the probability that the chain will focus on a local optimum.

The Bayesian analysis of variance table, visible in Figure 10, presents the prior probabilities of the alternative model and the null model; subsequently, it shows the Bayes factor, $BF_{10} = 18.359$, This means that the alternative model considering the effect of the concentration of SULFAMETOXAZOLE on the percentage of removal of the pollutant in the presence of the vetiver species (*Chrysopogon zizanioides*), is 18,359 times more likely than the null model, which only considers the large mean. Therefore, according to the scale of interpretation values of the Bayes factor, revised in Gelman et al. (2013), there is strong evidence in favor of the rejection of the null hypothesis.

Additionally, in Figure 10, there is a Table showing the posterior mean *Mean*, the standard error of the mean of the *SE Mean*, parameter, the standard deviation of the *SD* parameter, the 2.5, 25, 50, 75, 97.5 percentiles, the estimated effective sample size *n eff*[3] and the potential downscaling statistic *Rhat*[4], for the subsequent distributions of the overall mean, the deviation between groups, deviation within groups, and each treatment.

As part of the method, the Markov chain method is used, which, through a Monte Carlo simulation, presents the subsequent distribution of the mean, variance, and treatments. This can be seen in Figure 12.

[1] Markov chain is a popular tool in the random systems environment, where the simulation of each value is tied only to the immediately preceding value (Li et al., 2020)

[2] The Monte-Carlo Markov chain (MCMC) is a widely used method for estimating subsequent distributions in Bayesian inference (Pambabay-Calero et al., 2021)

[3] The generated Markov chain has approximately the same precision that would be obtained from an independent sample of n eff size (Ruppert, 2011). The higher this value, the better

[4] When Rhat is close enough to 1, the diagnosis declares convergence (Vats & Knudson, 2021)

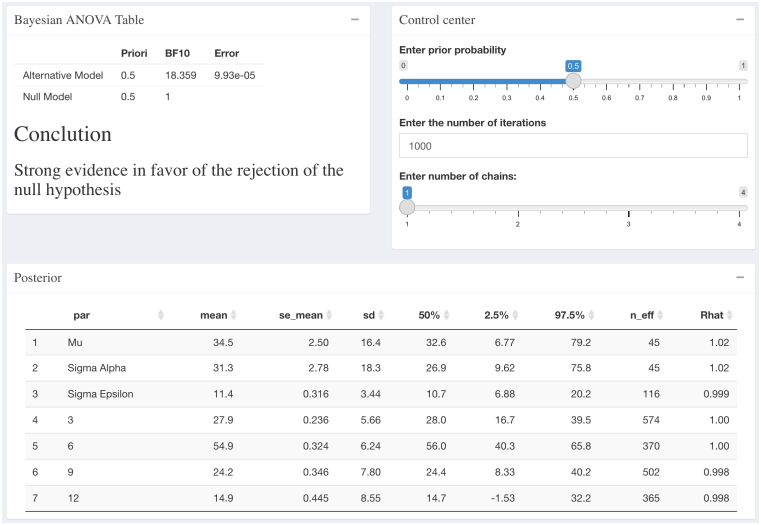

FIGURE 10: Bayes factor and subsequent summary. Bayesian analysis of variance, package output

Abbreviations: BF, Bayes Factor; par, Parameter; se_mean, Mean Squared Error; sd, Standard Desviation; n_eff, Effective sample size; Rhat, potential scale reduction factor on split chains: Mu, Mean.

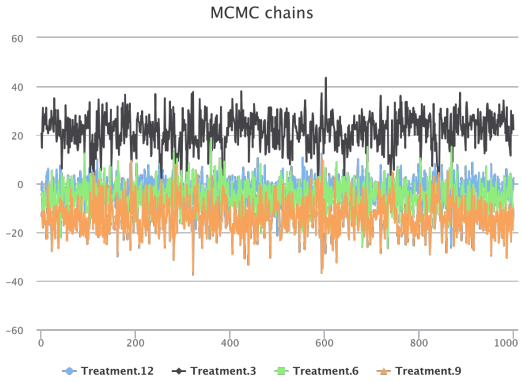

FIGURE 11: Markov Chains with 1000 iterations

In Figure 11 you can see the Markov chains and the statistical distributions generated by the MCMC method. It is observed that, with the treatment that considers a concentration of 3 mg/L, there is a different behavior, observable in the density curves. There is a shift to the right, while the distributions of the treatments with 6 mg/L, 9 mg/L, and 12 mg/L are close together, even almost entirely overlapping. This can also be evidenced in the Markov chains, shown in Figure 11. The blue line, corresponding to the treatment with 3 mg/L of pollutant concentration, is higher than the others, corresponding to the other treatments.

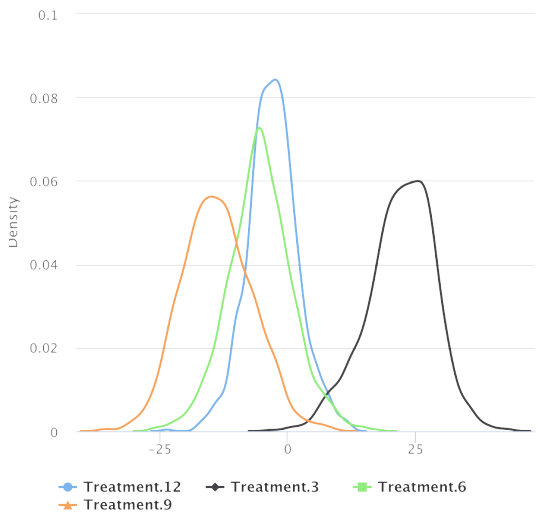

FIGURE 12: Posterior marginal distributions

Work is underway on the new version of the AovBay package, in which the Evans relative belief ratio will be incorporated.

5. Discussion

This article emphasizes the advantages of using the Bayes factor before the*p*-value. However, the test of statistical assumptions, parametric and non-parametric contrasts, being part of classical statistics, continue to depend on this value. Even the decision that is presented in the shiny panel, when evaluating the assumptions, is the result of the interpretation of the test using the*p*-value. The Bayes factor is invariant before linear transformations and is more computationally stable because it has algorithms defined for sampling and obtaining the subsequent distributions. However, when it comes to statistical inference, the*p*-value is more used in most statistical contrasts. According to Rouder et al. (2012), this may be since it is not usual to find formulas or computational tools based on the Bayes factor that is easy to use. The **AovBay** package breaks this paradigm, demonstrating that it is possible to incorporate Bayesian inference, R statistical software, and interactivity, responsible for the easy use of the tool.

Table 2, which describes the data for the illustrative example, indicates that there are few experimental units in each treatment in addition to not being a balanced design. It isn't easy to describe the behavior of a population when there are only two observations. When performing the Bayesian analysis of variance, this is no longer an inconvenience since, when the Markov Monte-Carlo chains are generated, the subsequent distribution of these populations, belonging to the treatments, is obtained with the number of iterations that the user wants to per-

form. It is easier and more convenient to infer on a population with 1000 (or more) observations than with 2, as are the treatments with concentrations 6 mg/L and 9 mg/L (Table 2). Furthermore, when comparing the subsequent distributions with the density curves, it is no longer necessary to make multiple comparisons of means, as in the case of classical analysis of variance.

The results of the analysis of variance with a Bayesian approach, obtained with the proposed tool, reflect that there are significant differences in the concentration levels of the pollutant. If the subsequent density curves presented in Figure 12, are analyzed, the treatment with 3 mg/L presents a shift to the right, away from the other curves, a behavior that is also seen in the Markov chains, also shown in Figure 11. In this case, the line that corresponds to the simulations of the distribution of the population from the treatment with 3 mg/L is higher than the others. With this outcome, it can be inferred that the treatment with 3 mg/L is responsible for a higher percentage of pollutant removal.

The results obtained by Checa-Artos et al. (2021),using response surfaces, indicate that it would be necessary to find the optimum removal percentage to reduce the concentrations in the case of the sulfamethoxazole pollutant. Still, due to the conditions of the experiment, the 3 mg/L concentration is considered the best option to obtain a higher percentage of pollutant removal. Therefore, the conclusion obtained with the results of the **AovBay** package is the same as that obtained in the study of response surfaces.

If the user wishes to compare the result obtained with the Bayesian approach and the one obtained with classical statistics. In that case, the emerging shiny panel also includes the classical analysis of variance, with post hoc test and the non-parametric equivalent of Kruskal Wallis, visible in Figures B1, B2 and B3, respectively.

In the case of analysis of variance, the proposed tool to decide which technique to apply provides researchers with ease in interpreting the statistical results through detailed data analysis and the interactivity achieved with the *highcharter* packages and *shiny* within RStudio. Although initially the most appropriate tool is recommended, depending on the compliance of assumptions, the shiny panel presents the three approaches to the analysis of variance so that data analysts and users of this tool can assess it and make appropriate decisions.

Due to those mentioned above, the analysis of univariate variance aims to test linear hypotheses about the influence of the different levels of one or more factors on the behavior of a variable (one-dimensional). Thus, future work intends to carry out an interactive tool for the multivariate analysis of variance (MANOVA), which considers a vector (multidimensional) of variables.

[Received: March 2023 — Accepted: November 2023]

References

Aslam, M. (2020), 'Design of the bartlett and hartley tests for homogeneity of variances under indeterminacy environment', *Journal of Taibah University for*

Science **14**(1), 6–10. https://doi.org/10.1080/16583655.2019.1700675

Barnett, V., Lewis, T. et al. (1994), *Outliers in statistical data*, Vol. 3, Wiley New York.

Checa-Artos, M., del Castillo, D. S., Eulalia, V. M., Ruiz-Barzola, O. & Barcos-Arias, M. (2021), 'Remoción de cinco productos farmacéuticos catalogados como contaminantes emergentes en medio acuoso utilizando la especie vetiver (Chrysopogon zizanioides)', *Bionatura* **6**(1). http://dx.doi.org/10.21931/RB/2021.06.01.7

Collyer, M. L. & Adams, D. C. (2018), 'RRPP: An R package for fitting linear models to high-dimensional data using residual randomization', *Methods in Ecology and Evolution* **9**(7), 1772–1779.

Duller, C. & Vorhauer, D. (2020), Comparison of four nonparametric tests of symmetry, *in* 'Actas de la conferencia AIP', Vol. 2293, AIP Publishing LLC, p. 420110.

Flegal, J. M., Haran, M. & Jones, G. L. (2008), 'Markov chain Monte Carlo: Can we trust the third significant figure?', *Statistical Science* **23**(2), 250–260. http://www.jstor.org/stable/27645897

Flores, E., Miranda, M. G. & Ángel Villasís, M. (2017), 'El protocolo de investigación VI: cómo elegir la prueba estadística adecuada. estadística inferencial', **64**(3), 364–370. https://doi.org/10.29262/ram.v64i3.304

Gandica de Roa, E. M. (2020), 'Potencia y robustez en pruebas de normalidad con simulación montecarlo', *Revista Scientific* **5**(18), 108–119.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. (2013), *Bayesian Data Analysis*, Chapman and Hall/CRC. http://www.stat.columbia.edu/ gelman/book/

Goodman, S. N. (1999), 'Toward evidence-based medical statistics. 2: The bayes factor', *Annals of Internal Medicine* **130**(12), 1005–1013. PMID: 10383350.

Guillen, F. G. F. & Chaparro, J. B. (2021), 'Análisis bayesiano del compromiso académico en estudiantes de psicología: diferencias según sexo y edad', *Revista de investigación en psicología* **24**(1), 5–18.

Jiménez-Paneque, R. (2016), 'El valor de p en entredicho: significación estadística, clínica y práctica', *Medwave* **16**(08).

Kozak, M. & Piepho, H.-P. (2018), 'What's normal anyway? residual plots are more telling than significance tests when checking anova assumptions', *Journal of Agronomy and Crop Science* **204**(1), 86–98. https://onlinelibrary.wiley.com/doi/abs/10.1111/jac.12220

Li, Z., Huang, M., Meng, X. & Ge, X. (2020), 'The limit theorems for function of markov chains in the environment of single infinite markovian systems', *Mathematical Problems in Engineering* **2020**. https://doi.org/10.1155/2020/8175723

Lintusaari, J., Gutmann, M. U., Dutta, R., Kaski, S. & Corander, J. (2017), 'Fundamentals and recent developments in approximate bayesian computation', *Systematic biology* **66**(1), e66–e82.

Morey, R. D. & Rouder, J. N. (2018), *BayesFactor: Computation of Bayes Factors for Common Designs*. R package version 0.9.12-4.2. https://CRAN.R-project.org/package=BayesFactor

Ostertagova, E., Ostertag, O. & Kováč, J. (2014), 'Methodology and application of the kruskal-wallis test', *Applied Mechanics and Materials* **611**, 115–120.

Ostertagova, Eva, Ostertag & Oskar (2013), 'Methodology and application of one-way ANOVA', *American Journal of Mechanical Engineering* **1**, 256–261.

Pambabay-Calero, J., Bauz-Olvera, S., Nieto-Librero, A. B., Galindo-Villardón, M. P. & Sánchez-García, A. B. (2020), 'A tutorial for meta-analysis of diagnostic tests for low-prevalence diseases: Bayesian models and software', *Methodology* **16**(3), 258–277.

Pambabay-Calero, J., Bauz-Olvera, S., Nieto-Librero, A., Sánchez-García, A. & Galindo-Villardón, P. (2021), 'Hierarchical modeling for diagnostic test accuracy using multivariate probability distribution functions', *Mathematics* **9**(11). https://www.mdpi.com/2227-7390/9/11/1310

R Core Team (2012), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org/

Ramos-Vera, C. A. (2021), 'Un aporte del análisis bayesiano en la investigación psicológica', *CES Psicología* **14**(3), 191–194.

Rendón-Macías, M., Riojas-Garza, A., Contreras-Estrada, D. & Martínez-Ezquerro, J. (2018), 'Análisis bayesiano. conceptos básicos y prácticos para su interpretación y uso', *Revista alergia México* **65**, 285–298.

Rojas-Campuzano, M. & Pambabay-Calero, J. (2021), *AovBay: Classic, Nonparametric and Bayesian One-Way Analysis of Variance Panel*. R package version 0.1.0. https://CRAN.R-project.org/package=AovBay

Rouder, J. N., Morey, R. D., Speckman, P. L. & Province, J. M. (2012), 'Default bayes factors for anova designs', *Journal of Mathematical Psychology* **56**(5), 356–374. https://www.sciencedirect.com/science/article/pii/S0022249612000806

Rubin, D. B. (1987), *Multiple Imputation for Nonresponse in Surveys*, Wiley Series in Probability and Statistics.

Ruppert, D. (2011), *Bayesian Data Analysis and MCMC*, Springer New York, pp. 531–578.

SAS Institute Inc. (1999), *Guía del usuario de SAS / STAT, versión 8.*, SAS Institute, Cary, NC.

Solari, F., Liseo, B. & Sun, D. (2008), 'Some remarks on bayesian inference for one-way anova models', *Annals of the Institute of Statistical Mathematics* **60**, 483–498.

Stan Development Team (2020), 'RStan: the R interface to Stan'. R package version 2.21.2. http://mc-stan.org/

StataCorp. (2019), *STATA USER'S GUIDE RELEASE 16*, StataCorp LLC., College Station, TX.

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E.-J., Derks, K., Dablander, F., Gronau, Q. F., Šimon Kucharský, Gupta, A. R. K. N., Sarafoglou, A., Voelkel, J. G., Stefan, A., Ly, A., Hinne, M., Matzke, D. & Wagenmakers, E.-J. (2020), 'A tutorial on conducting and interpreting a bayesian anova in jasp', *L'Année psychologique* **120**. https://www.cairn.info/revue-l-annee-psychologique-2020-1-page-73.htm

Vats, D. & Knudson, C. (2021), 'Revisiting the gelman–rubin diagnostic', *Statistical Science* **36**(4), 518–529.

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Smíra, M., Et, S. E. A., Matzke, D., de Jong, T., van den Bergh, D., Sarafoglou, A., Steingroever, H., Derks, K., Rouder, J. N. & Morey, R. D. (2018), 'Bayesian inference for psychology. part ii: Example applications with jasp', *Psychonomic Bulletin & Review* **25**. https://doi.org/10.3758/s13423-017-1323-7

Wedel, M. & Dong, C. (2020), 'BANOVA: Bayesian analysis of experiments in consumer psychology', *Journal of Consumer Psychology* **30**(1), 3–23. https://onlinelibrary.wiley.com/doi/abs/10.1002/jcpy.1111

Çelik, N. & Şenoğlu, B. (2018), 'Robust estimation and testing in one-way anova for type ii censored samples: skew normal error terms', *Journal of Statistical Computation and Simulation* **88**(7), 1382–1393. https://doi.org/10.1080/00949655.2018.1433670

Appendix A. Relevant Formulas

Appendix A.1. One-way ANOVA

$$Y_{ij} = \mu_i + \epsilon_{ij} \quad ; \quad i = 1, \ldots, k \quad , \quad j = 1, \ldots, n_i$$

where

- i represents the ith treatment of the factor, and the total number of treatments is denoted by k.

- j represents observation j, the maximum number of j is n_i, meaning the total number of observations in treatment i.
- Y_{ij} is the value taken by the response variable in the jth observation of the ith treatment.
- μ_i is the mean of the ith treatment.
- ϵ_{ij} is the random error, independent and identically distributed normally with mean 0 and variance σ^2.
- $n = n_1 + n_2 + \cdots + n_k$.

Appendix A.2. F statistic

$$F = \frac{SSTR/\left(k-1\right)}{SSE/\left(n-k\right))}$$

Appendix A.3. Prior

The ratio between two priori probabilities

$$\frac{\pi(\Theta_1)}{\pi(\Theta_0)}$$

so that $\pi(\Theta_1)$ and $\pi(\Theta_0)$ are the prior probabilities of the alternative and null hypotheses, respectively.

Appendix A.4. Posterior

The posterior density $p(\Theta|y)$ can be obtained.

$$p(\Theta|y) \propto \pi(\Theta)p(y|\Theta)$$

where, $\pi(\Theta)$ is the prior probability and $p(y|\Theta)$ the likelihood.

Appendix A.5. Bayes Factor

The Bayes factor (FB) can be expressed as follows:

$$FB_{01} = \frac{p(y|\Theta_0)}{p(y|\Theta_1)}$$

Where $p(y|\Theta_0)$ and $p(y|\Theta_1)$ are the data distributions under the null and alternate hypotheses, respectively.

Appendix B. Annexes

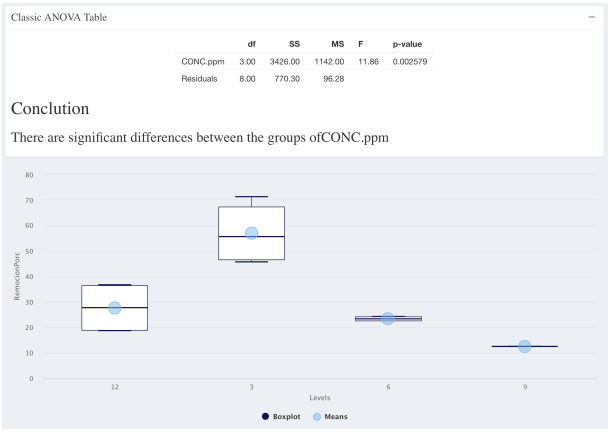

FIGURE B1: Analysis of variance from the classical approach, package output

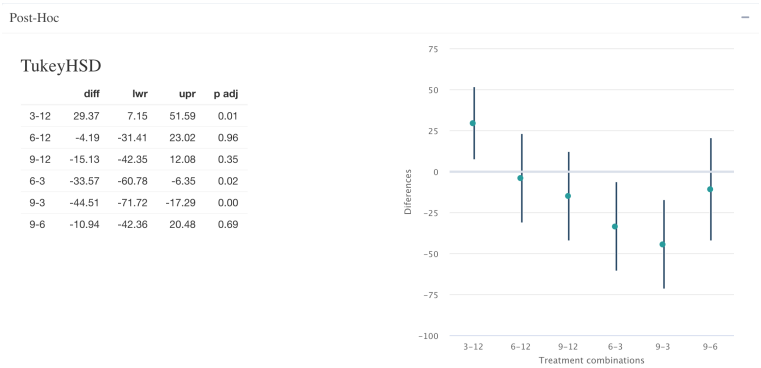

FIGURE B2: Post Hoc Tests (Classic Anova), Package output

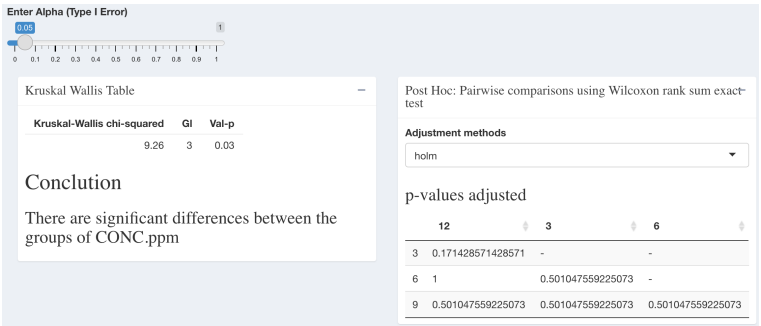

FIGURE B3: Kruskal Wallis (Non-parametric analysis of variance), package output