https://repository.uaeh.edu.mx/revistas/index.php./MJMR/issue/archive

Mexican Journal of Medical Research ICSa

Biannual Publication, Vol. 12, No. 23 (2024) 22-26

ISSN: 2007-5235

Step-by-step one-way ANOVA analysis with the Jamovi program

Paso a Paso de análisis ANOVA de una vía con el programa Jamovi

Norma A. Milanes-Baños [a]

Abstract:

ANOVA is the acronym for Analysis of Variance, and is used to establish statistically significant differences between the means of two or more independent groups. Currently there are several programs that perform this analysis, which allows the analyses of statistical data in a simpler way. This manuscript addressees a one-way or one-factor analysis of variance, in order to be able to perform it, there must be a dependent variable or a continuous quantitative type of response and at least one factor or independent variable of a categorical type with more than two levels. In a fictitious situation, with this method it will be observed how the independent variables shifted to the dependent variable, using the Jamovi program, which is a free and easy-to-install advanced calculation program that allows the efficient creation of complex statistical calculations, for manipulation, filtering, selection and combination of data.

Keywords:

Jamovi, Variance, ANOVA

Resumen:

ANOVA es el acrónimo de Análisis de Varianza, y se utiliza para establecer diferencias estadísticamente significativas entre las medias de dos o más grupos independientes. Actualmente existen varios programas que realizan este análisis permitiéndonos analizar datos estadísticos de una forma más sencilla. En este manuscrito abordaremos un análisis de varianza unidireccional o unifactorial, para poder realizarlo debe existir una variable dependiente o respuesta de tipo cuantitativo continuo y al menos un factor o variable independiente de tipo categórico con más de dos niveles. En una situación ficticia, observaremos con este método cómo las variables independientes se desplazan a la variable dependiente, utilizando el programa Jamovi, que es un programa de cálculo avanzado gratuito y fácil de instalar que permite realizar cálculos estadísticos complejos de manera eficiente, para manipulación, filtrado, selección y combinación de datos.

***Palabras Clave**:*

Jamovi, Varianza, ANOVA

INTRODUCTION

ANOVA or analysis of variance is a statistical method that studies how one or more factors (independent variables) affect a dependent variable. In order to be able to perform it, there must be a dependent or response variable of a continuous quantitative type and at least one factor or independent variable of a categorical type with more than two levels. This analysis must have data from populations that have a normal distribution with similar or equal variances between groups.[1–3]

Analysis of variance is a basic technique to study observations that depend on several factors, and it is a fundamental tool in the analysis of Linear Regression and Design of Experiments models.[1,2] The basic idea of the analysis of variance consists of decomposing the total variability observed in some data into a series of terms, associated with the effects of each factor studied and their possible interactions, plus a residual part with which the former will then be compared.[3] Analysis of variance is a very useful statistical tool. There are multiple application examples, which are focused mainly on the comparison of multiple columns of data.[2] There are some variations of this analysis, and they depend mainly on whether the measurements are repeated or not (independent measures ANOVA and repeated measures ANOVA), and on the number of independent variables that are considered in the study design (one-way ANOVA, two-way ANOVA, etc.).[4] This paper deals in a practical way with a problem to be solved by means of the

[a] Corresponding author, Universidad Autónoma del Estado de Hidalgo, https://orcid.org/0000-0002-6699-4920, Email: norma_milanez@uaeh.edu.mx

Received: 08/03/2023, Accepted: 23/05/2023, Postprint: 20/10/2023, Published: 05/01/2024

DOI : https://doi.org/10.29057/mjmr.v12i23.10664

statistical technique of one-way analysis of variance for independent samples using the Jamovi statistical package.[5]

ASSUMPTIONS

To use one-way ANOVA of independent samples satisfactorily, several assumptions must be met, although slight deviations from ideal conditions are acceptable:

1. The dependent variable should be measured on a continuous quantitative scale.
2. The independent variable must consist of two or more independent categorical groups. Generally, an ANOVA is used when it has three or more independent categorical groups, but it can be used for only two groups.
3. There must be independence of observations, meaning that there is no relationship between the observations in each group or between the groups themselves. For example, there must be different participants in each group with no participant being in more than one group. This is more of a study design problem than something that a program can test, but it is an important assumption of ANOVA.
4. There should be no significant outliers. Outliers are simply single data points within your data that do not follow the usual pattern.
5. The dependent variable should have an approximate normal distribution for each independent variable category. To prove this, the Shapiro-Wilk normality test can be run.
6. There must be homogeneity of variances. This assumption can be tested using Levene's test for homogeneity of variances.[2,6-8]

EXAMPLE

Assume the following experimental design: A student investigated how effective is to wash with soap to kill bacteria. To do this, she used four different methods: washing just with water, washing with regular soap, washing with antibacterial soap (ABS), and spraying hands with antibacterial spray (AS, containing 65% ethanol as the active ingredient). The experiment consisted of one experimental factor, the washing method, at four levels. She suspected that the number of bacteria on her hands before washing them might vary considerably from one day to another. To help balance the effects of these changes, she generated random numbers to determine the order of the four treatments. Each morning, she washed her hands according to the randomly chosen treatment. Then, she placed her right hand on a plate of sterile media designed to stimulate bacterial growth. She incubated each plate for 2 days at 36 °C, after which she counted the bacterial colonies. She replicated this procedure 8 times for each of the four treatments.

Considering the experimental design described above, the following can be established:

Dependent variable: Bacterial count (measured on a continuous quantitative scale).

Independent variable: Cleaning method (measured on a categorical scale with 4 levels).

Measurements: independent.

Null hypothesis: There are no statistically significant differences in bacterial counts among the hand cleaning methods.

Considering this experimental design, the indicated analysis to test the hypothesis that one cleaning method is more effective than another in reducing bacterial counts is the one-way analysis of variance of independent measures.

CONFIGURATION IN THE JAMOVI PROGRAM

In Jamovi, the analysis groups were separated by creating a grouping variable called cleaning method (i.e., the independent variable), and it was set up as a quantitative variable. The number of viable bacteria was entered under the variable name antimicrobial count and was set up as a nominal type variable (Figure 1).[9-12]

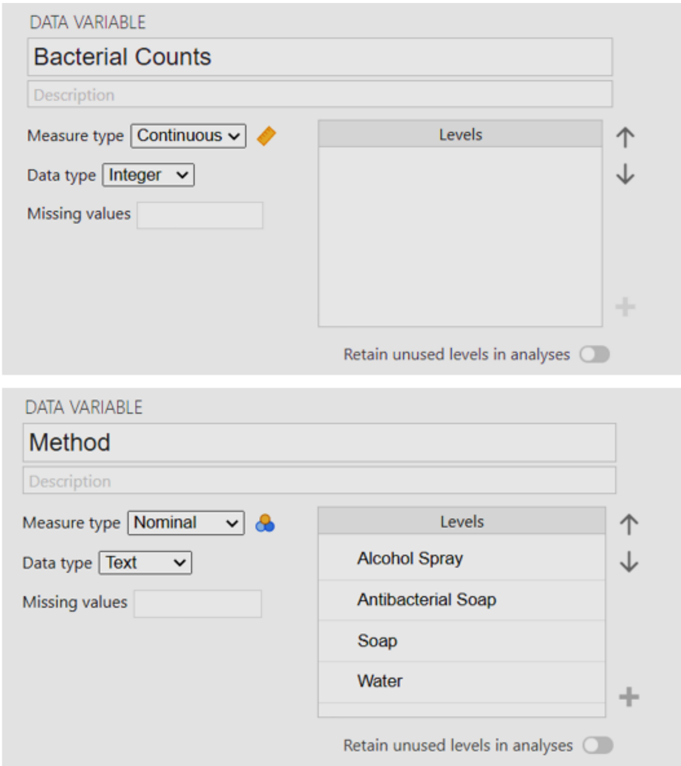

Figure 1. Configuration of the variables.[9,13-15]

Bacterial Counts	Method
74	Water
84	Soap
70	Antibacterial Soap
51	Alcohol Spray
135	Water
51	Soap
164	Antibacterial Soap
5	Alcohol Spray
102	Water
110	Soap
88	Antibacterial Soap
19	Alcohol Spray
124	Water
67	Soap
111	Antibacterial Soap
18	Alcohol Spray
105	Water
119	Soap
73	Antibacterial Soap
58	Alcohol Spray
139	Water
108	Soap
119	Antibacterial Soap
50	Alcohol Spray
170	Water
207	Soap
20	Antibacterial Soap
82	Alcohol Spray
87	Water
102	Soap

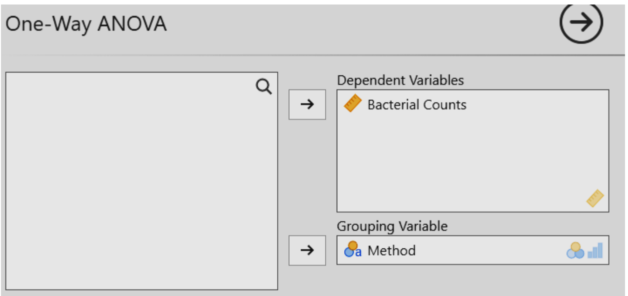

Figure 2. Selection of the variables of the methodological design.[9,13-15]

In Figure 2, once the variables have been set up, the following procedure can be executed: Analysis ⇢ ANOVA ⇢ One-factor ANOVA. A new window will be displayed where the dependent variables (antimicrobial count) and the grouping variable (cleaning method) will have to be chosen.

Other options that should be activated to verify that the data meet the requirements of normality and homogeneity in variance are those highlighted in the "Assumption check" option (Figure 3).[9,13-15]

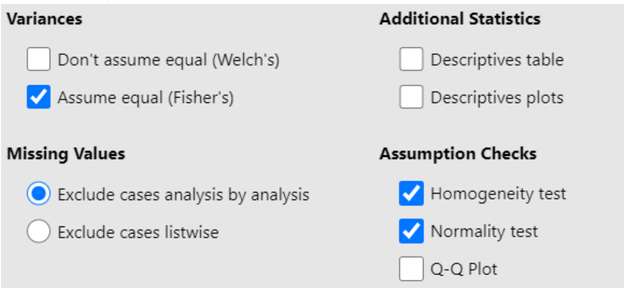

Figure 3. Checking assumptions.[9,13-15]

The verification of the normality and homogeneity requirements of variances is presented in the results window. In the case of the data series analysed for this example, the *p* values obtained suggest that the data meet the requirements to be analysed with an ANOVA test, $p = 0.391$ for normality, and $p = 0.910$ for homogeneity in variance (Figure 4).[9,13-15] In general, the data are assumed as meeting these requirements when the p values for the Shapiro-Wilk and Levene's test are greater than 0.05. When these requirements are not met, a nonparametric test or Welch's analysis of variance should be chosen (Figure 3).[9,13-15]

Assumption Checks

Normality Test (Shapiro-Wilk)

	W	p
Bacterial Counts	0.966	0.391

Note. A low p-value suggests a violation of the assumption of normality

Homogeneity of Variances Test (Levene's)

	F	df1	df2	p
Bacterial Counts	0.179	3	28	0.910

Figure 4. Checking assumptions.[9,13-15]

Once the assumptions have been checked, the table of the one-factor analysis of variance (Figure 5)[9,13-15] can be seen in the results window. This table shows the F-statistic and the p-value obtained. For the data series being analysed in this example, it should be observed the result of the Fisher's test line (since the data meet the requirement of equality of variances). In this case, the *p*-value obtained is 0.001, which leads to reject the null hypothesis, and to state that the cleaning method significantly influences the bacterial count.

One-Way ANOVA

		F	df1	df2	p
Bacterial Counts	Welch's	10.7	3	15.2	< .001
	Fisher's	7.06	3	28	0.001

Figure 5. Analysis of variance.[9,13-15]

If the ANOVA test is significant (as it was the case in this example), it is important to know *a posteriori* which groups can be considered equal and which different. When there are more than two groups, ANOVA only tells us that differences have been found between the groups and that our grouping variable or factor is significant, so we know that differences exist but we do not know where. There are several tests, called post-hoc tests, which allow to know which groups differ from each other. The fact of performing post-hoc tests indicates that there will have to be made multiple comparisons between the groups to be compared (Figure 6). [9,13-15] Some of the best known statistics are LSD (Least Significant Difference), Bonferroni, Scheffé, S-N-k (Student-Neuwman-Keuls), Tukey, Duncan, Dunnet, among others. Many statistical packages incorporate these tests in the analysis of variance. The Jamovi package uses only two (Figure 5), depending on whether or not the requirement of equality in the variances of the groups was met. [9,13-15]

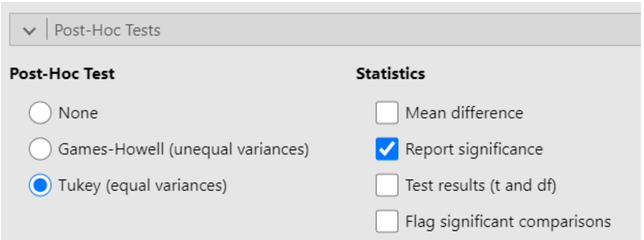

Figure 6. Configuration to perform Tukey's post-hoc analysis.[9,13-15]

In the results window, the result of the post-hoc test can be observed. In order to be able to interpret these results, it is of vital importance to make a table with the descriptive data of the data series being analysed (Figure 7). [9,13-15]

Tukey Post-Hoc Test – Bacterial Counts

		Alcohol Spray	Antibacterial Soap	Soap	Water
Alcohol Spray	p-value	—	0.032	0.006	0.001
Antibacterial Soap	p-value		—	0.889	0.568
Soap	p-value			—	0.936
Water	p-value				—

Figure 7. Post-hoc results.[9,13-15]

For the interpretation of the results of the post-hoc analysis, the following sequence will be used. Reading the comparisons of the first line, the comparisons between "water vs soap" - p = .936 and "water vs antibacterial soap" - p = .568, are not significant, while the comparison "water vs alcohol spray" - p = .001, is significant. Then, the second line is analysed, where it can be observed that the comparison "soap vs antibacterial soap" is not significant (p = .899), while the comparison "soap vs alcohol spray" is significant (p = .006). Finally, when reviewing the third line, it can be verified that the comparison "antibacterial soap vs alcohol spray" is significant (p = .032). With all this, it can be affirmed that the "Alcohol spray" group is different from the rest of the groups, while the "Water", "Soap" and "Antibacterial soap" groups are similar. This can be easily represented in Table 1 as shown below:

Table 1. *Bacterial count (Mean [SD]) as a function of cleaning method.*

Cleaning method	Mean (SD)
Water	117.0 (31.1)[a]
Soap	106.0 (47.0)[a]
Antibacterial Soap	92.5 (42.0)[a]
Alcohol Spray	37.5 (26.6)[b]

Different lowercase letters indicate the presence of statistically significant differences ($p < .05$).[9,13-15]

CONCLUSION

The analysis of statistical data with free and easy to use programs such as Jamovi facilitates statistical work, as in this case the one-way ANOVA, it is of utmost importance to know the use of programs in which statistical work is much easier and accurate, as they make work easier and speed up the results.

REFERENCES

[1] Mishra P, Singh U, Pandey CM, Mishra P, Pandey G. Application of Student's test, analysis of variance and covariance. Ann. Card. Anaesth. 2019; 22(4): 407–11.

[2] Hazra A., Gogtay N. Módulo 3 de la serie Bioestadística: Comparación de grupos: Variables numéricas. Indio J. Dermatol. 2016; 61:251–60.

[3] Hlávka Z, Hlubinka D, Koňasová K. Functional ANOVA based on empirical characteristic functionals. J. Multivar. Anal. 2022;189:12-8.

[4] Potts D, Schmischke M. Interpretable Transformed ANOVA Approximation on the Example of the Prevention of Forest Fires. Front. Appl. Math. Stat. 2022;8:1-8.

[5] Jamovi - open statistical software for the desktop and cloud [Internet]. Jamovi.org. 2022. Available from: https://www.jamovi.org/

[6] Lamboni M. Weak derivative-based expansion of functions: ANOVA and some inequalities. Math. Comput. Simul. 2022;194:691-718.

[7] Zondervan-Zwijnenburg M, van de Schoot R, Hoijtink H. Testing ANOVA Replications by Means of the Prior Predictive p-Value. Meta-Psychology. 2022;6.

[8] Borgonovo E, Li G, Barr J, Plischke E, Rabitz H. Global Sensitivity Analysis with Mixtures: A Generalized Functional ANOVA Approach. Risk Analysis. 2022;42(2): 304-33.

[9] Şahin M, Aybek E. Jamovi: An Easy-to-Use Statistical Software for the Social Scientists. Int. J. Asst. Tools Educ. 2019;4: 670–92.

[10] Acutis M, Tadiello T, Perego A, di Guardo A, Schillaci C, Valkama E. EX-TRACT: An excel tool for the estimation of standard deviations from published articles. Environ. Model. Softw. 2022;147: 105236.

[11] Sánchez Villena A. Uso de programas estadísticos libres para el análisis de datos: Jamovi. Rev. Perspective. 2019;20(1):112-4.

[12] Bax L, Yu LM, Ikeda N, Moons KGM. A systematic comparison of software dedicated to the meta-analysis of causal studies. BMC Med. Res. 2007; 7:40.

[13] Castejón J, Ríos J, Sánchez E, Maurandi A. Didáctica de las matemáticas, software libre y desarrollo de recursos mediante Learnr y Shiny. Educateconciencia. 2021: 29(31), 101-121.

[14] Strunk KK, Mwavita M. Design and analysis in educational research using Jamovi: ANOVA designs. Design and Analysis in Educational Research Using Jamovi: ANOVA Designs. Routledge Londres; 2021. p. 1-32.

[15] Santabárbara Serrano J, Lasheras I. Docencia de Bioestadística en Medicina con software gratuito Jamovi: una ventana de oportunidad. Rev. Esp. Educ. Méd. 2020; 1:9-10.