Agrociencia

ANALYSIS FOR MULTIPLE RESPONSES IN A COMPLETELY RANDOMIZED EXPERIMENTAL DESIGN

Gloria Selene **Herrera-Reyes**[1], Miguel Ángel **Martínez-Reyes**[1], Paulino **Pérez-Rodríguez**[1], Juan Manuel **Romero-Padilla**[1], Ignacio **Luna-Espinoza**[1], Javier **Suárez-Espinosa**[1*]

[1]Colegio de Postgraduados Campus Montecillos. Carretera México-Texcoco km 36.5, Montecillo, Texcoco, State of Mexico, Mexico. C. P. 56264.
* Author for correspondence: sjavier@colpos.mx

ABSTRACT

Multiple responses are often generated in agricultural and forestry research. For example, the moisture content, fatty acids, carbohydrates, size, diameter, length, shape, and hardness, among other characteristics, are measured in cottonseeds. Multivariate analysis of variance (MANOVA) can be useful for multiple response analysis when differences in treatment effects are to be determined. However, the performance of current *post hoc* tests in this context is not satisfactory due to the limitations of the available methods or because they are difficult to use for non-statistician researchers. Furthermore, this methodology requires the assumptions of multivariate normality and homogeneity of variance and covariance matrices, assumptions that are difficult to verify if the sample size is small. This research proposes an alternative analysis to test the hypothesis of equality of effects between treatments and *post hoc* tests in the case of multiple responses. An asymptotic result is demonstrated for the random variable generated in the proposal for the case of uncorrelated normal variables, and the case for correlated normal random variables is left open. A simulation study shows that the performance of the proposal with small samples is satisfactory in terms of power and that it has advantages compared to MANOVA. Furthermore, the methodological approach allows for *post hoc* testing in the case of multiple responses in the completely randomized experimental design.

Keywords: ANOVA, MANOVA, assumptions, data transformation, Euclidean norm.

INTRODUCTION

The generation of multiple responses is common in research in a variety of areas. For example, Pérez-López *et al*. (2014) presented a study of fava bean cultivars where the following responses were recorded: plant height, number of branches, number of flower nodes, number of pods per plant, pod weight per plant, number of seeds per pod and per plant, total seed weight per plant, number of clean seeds per plant, and weight of clean seed per plant of 100 seeds and of spotted seed per plant. In other research, measurements of weight, color, texture, protein, fat, and vitamin content were obtained from a portion of chicken meat (Sosnówka-Czajka *et al*., 2023); the number of bacteria, pH, and fiber and vitamin content were obtained from cactus

Citation: Herrera-Reyes GS, Martínez-Reyes MA, Pérez-Rodríguez P, Romero-Padilla JM, Luna-Espinoza I, Suárez-Espinosa J. 2024. Analysis for multiple responses in a completely randomized experimental design. **Agrociencia.** https://doi.org/ 10.47163/agrociencia.v58i2.3164

Editor in Chief:
Dr. Fernando C. Gómez Merino

Received: February 07, 2024.
Approved: March 03, 2024.
Published in Agrociencia:
March 20, 2024.

(El-Mostafa *et al*., 2014; Hernández-Anguiano *et al*., 2016); and moisture content, fatty acids, carbohydrates, size, diameter, length, shape, and hardness were measured from cottonseed (Anitha *et al*., 2022).

Multivariate analysis of variance (MANOVA) is one methodology used to evaluate the hypothesis of equality of effects between two or more treatments when there are many responses. However, when MANOVA rejects the hypothesis of equal treatment effects, there are no satisfactory alternatives for *post hoc* testing. The methodology developed by Seo *et al.* (1994) was tested for a limited number of treatments and variables. This methodology requires the assumptions of multivariate normality, homogeneity of covariance matrices. Unfortunately, these assumptions are difficult to verify if the sample size is small (Hair, 1999; Dattalo, 2013). Moreover, such a methodology is difficult to implement for non-statistician researchers. Warne *et al*. (2012) found that 5 out of 62 articles that used MANOVA in educational psychology journals had correctly applied *post hoc* procedures. Furthermore, Warne (2014) screened the top three psychology journals and found that, in 58 articles, researchers used MANOVA between 2009 and 2013; however, none of these articles used *post hoc* procedures.

Much of the statistical methodology proposed for comparing multiple mean vectors is based on the T^2_{max} statistic (Seo, 2002; Nishiyama *et al*., 2014). However, as argued by Nishiyama and Seo (2013) and Nishiyama *et al*. (2014), finding the distribution of the test statistic is difficult, even in the simplest cases of pairwise comparison of vector means, assuming normality. Hence, the upper quantiles of the statistic T^2_{max} have been determined only for particular cases. For example, assuming normality in the data, Nishiyama and Seo (2013) determined the 0.9, 0.95, and 0.99 quantiles of the distribution of the T^2_{max} statistic as part of their proposed methodology for testing four vectors of correlated means.

In this context, the present research paper proposes an alternative analysis for the determination of between-treatment effects and *post hoc* tests for the case of multiple responses.

MATERIALS AND METHODS

Multiple response data generated in a completely randomized experimental design (CRD) can be modeled by the following: $\boldsymbol{Y}_{ij} = \boldsymbol{\mu} + \boldsymbol{\tau}_i + \boldsymbol{e}_{ij}$, where each component of the model is a p-dimensional vector: $\boldsymbol{Y}_{ij} = (Y_{ij1}, ..., Y_{ijp})^t$ is the random vector of response variables for the j-th repetition of the i-th treatment $i = 1,..., t$, $j = 1, ... r$, whose element Y_{ijk} corresponds to the k-th random variable, $k = 1, ..., p$; $\boldsymbol{e}_{ij} = (e_{ij1}, ..., e_{ijp})^t$ is the vector of random errors; $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)^t$ is the vector of overall means; and $\boldsymbol{\tau}_i = (\tau_{i1}, ..., \tau_{ip})^t$ is the vector of effects of the i-th treatment. In practice, $\boldsymbol{\mu}$ and $\boldsymbol{\tau}_i$ are unknown parameters (Rencher and Christensen, 2012).

The set of hypotheses used in the MANOVA is as follows:

H_0: $\tau_1 = \tau_2 = .. = \tau_t$ vs. = H_a: $\tau_i \neq \tau_j$, for at least a $i \neq j$.

Johnson and Wichern (2007) stated that, in order for the data to meet the basic assumptions for the MANOVA result to be reliable, the observations must be random samples of size r of treatment i, the random samples of the treatments must be independent, and each treatment must have a multivariate normal distribution with a common variance and covariance matrix for all treatments, i.e: $\boldsymbol{Y}_{ij} \sim N_p\ (\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, where $\mu_i = (\mu_1 + \tau_{i1}, \ldots, \mu_p + \tau_{ip})^t$.

If H_0 is rejected, it is necessary to identify which treatments have different effects from each other; for this purpose, vector mean comparison methods are used. A very popular alternative is the Hotelling T^2 method by applying the Bonferroni correction. However, this method is very conservative (Dattalo, 2013). Another is the generalized Tukey conjecture, developed by Seo *et al*. (1994) and Seo and Nishiyama (2008), which is a generalization of the univariate Tukey-Kramer methodology. In the procedure, confidence intervals are generated for the differences by pairs of mean vectors. This proposal has the limitation that it is only used for a maximum of four treatments, vectors with five variables, and 60 degrees of freedom (with $\upsilon = N - p - 1$, degrees of freedom).

ANOVA is another methodology to determine treatment effects with a simpler model. For example, the CRD model, $Y_{ij} = \mu + \tau_i + e_{ij}$, which is similar to MANOVA; only the components are scalar, and its basic assumptions are independence, normality, and homoscedasticity (Montgomery, 2004). When any of the assumptions are not met, methods can be used to transform the data, such as the Box and Cox (1964) methodology, although Driscoll (1996) and Salkind (2010) agree that ANOVA is robust to non-normality of the data.

When the hypothesis of equality of treatments is rejected in an ANOVA, it is necessary to identify which treatments cause the difference. For this purpose, comparisons of means are carried out. Montgomery (2004) and Hinkelmann and Kempthorne (2005) mention that the main *post hoc* methods for such comparisons are the Fisher's least significant difference (LSD), Tukey's honest significant difference (HSD), Dunnett's least significant difference (LSD), Duncan's multiple range, and the Student-Newman-Keuls (SNK) test. It should be noted that ANOVA has fewer limitations than MANOVA, as well as the development of several tests for comparison of means; however, ANOVA is not designed to analyze data with multiple responses. As a result, an alternate methodology is proposed for data from three or more treatments with various responses gathered through experimental designs.

Methodological proposal

The proposal is to reduce each vector of response variables to a scalar in order to obtain data that can be analyzed by means of an ANOVA and, subsequently, by means of a *post hoc* method to make a comparison of means.

It is proposed that each variable of the p-vector Y_{ij} be transformed with the quadratic function of the Euclidean norm as follows:

$$X_{ij} = \boldsymbol{Y}_{ij}^t \boldsymbol{Y}_{ij} = \sum_{k=1}^{p} Y_{ijk}^2 \qquad (1)$$

where Y_{ijk} is the k-th random variable ($k = 1, ..., p$) from vector $\boldsymbol{Y}_{ij}$ (Table 1).

Table 1. Use of the Euclidean norm in the data set.

Treatment	Variable 1	...	Variable p		Use of the square of the norm
	y_{111}	...	y_{11p}		$y_{t21}{}^2 + ... + y_{11p}{}^2 = \|y_{11}\|^2 = x_{11}$
1	y_{121}	...	y_{12p}		$y_{121}{}^2 + ... + y_{12p}{}^2 = \|y_{12}\|^2 = x_{12}$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	y_{1r1}		y_{1rp}		$y_{1r1}{}^2 + ... + y_{1rp}{}^2 = \|y_{1r}\|^2 = x_{1r}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\rightarrow$	$\vdots$
	y_{t11}	...	y_{t1p}		$y_{t11}{}^2 + ... + y_{t1p}{}^2 = \|y_{11}\|^2 = x_{t1}$
t	y_{t21}	...	y_{t2p}		$y_{t21}{}^2 + ... + y_{t2p}{}^2 = \|y_{12}\|^2 = x_{t2}$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	y_{tr1}		y_{trp}		$y_{tr1}{}^2 + ... + y_{trp}{}^2 = \|y_{1r}\|^2 = x_{tr}$

By transforming the data with multiple respBy transforming the data with multiple responses to a scalar value, a sequence of independent random variables X_{ij}, is generated, which can be analyzed by means of an ANOVA. Each of the p characteristics is obtained from the same object, so they may correlate with each other. However, the sequence of variables X_{ij} (Equation 1) and referring to the treatment i in its repetition j, can be considered independent because, *a priori*, the researcher must ensure the independence of them by randomization.
Now, note that X_{ij} it is a sum of random variables, so the following central limit theorem for the sum of random variables can be applied:

Theorem 1: Let $W_1, W_2, ..., W_n$ be a sample of n independent random variables with distribution functions $F_1, F_2, ..., F_n$, respectively, such that $E(W_i) = \mu_i$ and $Var(W_i) = \sigma_i^2$ for $i = 1, ..., n$, and $s_n^2 = \sum_{i=1}^{n} \sigma_i^2$, then:

$$S_n^* = s_n^{-1} \sum_{i=1}^{n} (W_i - \mu_i) \xrightarrow{d} Z,$$

where $Z \sim N(0, 1)$ provided that the F is absolutely continuous with density function f_i, such that the following, known as the Lindeberg condition, is satisfied:

$$\lim_{n\to\infty} s_n^{-2} \sum_{i=1}^{n} \int_{|w-\mu_i|>\epsilon s_n} (w-\mu_i)^2 f_i(w)dw = 0,$$

As the Lindeberg condition is met, the vector size (p) is sufficiently large, the variances of X_{ij} are homogeneous for all i, and the conclusions obtained from the ANOVA will be valid. The expression "large enough" is controversial and should be taken with caution, because whether certain sample sizes are considered "large enough" depends on the shape of the original distribution (Correa-Londoño and Castillo-Morales, 2000). Although there are potentially many multivariate distributions, such that the vector p-variate Y_{ij} under the transformation (Equation 1) can meet the above conditions, the most typical case will be explored.

Case 1. Uncorrelated normal variables (theoretical result)

Since MANOVA works under the assumptions of multivariate normality and homogeneity of variances and covariances, these assumptions will be used as a starting point to apply the methodological proposal.

Assuming that the vector $\boldsymbol{Y}_{ij}$ has $N_p(\boldsymbol{\mu}_i, \sigma_i^2\ \boldsymbol{I}_p)$ distribution, the random variable generated from the squared function of the norm $X_{ij} = Y_{ij}^t\, Y_{ij} = \sum_{k=1}^{p} Y_{ijk}^2$ has a non-central chi-squared distribution, with mean $p + \lambda$ and variance $2(p + 2\lambda)$, for $p > 0$ which specifies the degrees of freedom and $\lambda \geq 0$ which is the non-centrality parameter (Casella, 2008):

$$\boldsymbol{Y}_{ij}^t \boldsymbol{Y}_{ij} \sigma^{-2} \sim \chi_p^2(\lambda), \lambda = 0.5 \boldsymbol{\mu}_i^t \boldsymbol{\mu_i}\, \sigma^{-2}.$$

Even if the same $\boldsymbol{Y}_{ij}$ variance is assumed for all i, X_{ij} have different variances because they depend on the mean of each treatment. Under this scenario, the random variables X_{ij} do not follow a normal distribution and do not have homogeneous variances. Therefore, if X_{ij} are used, the ANOVA results will not be valid. However, if p is sufficiently large, X_{ij} may converge to the normal distribution, so it would be feasible to use the proposed methodology in this case.

Convergence demonstration for the case of uncorrelated variables

Let $X_{ij} = \sum_{k=1}^{p} Y_{ijk}^2$ be a random a random variable with $E(Y_{ijk}^2) = \mu_{ik}$ and $Var(Y_{ijk}^2) = a_{ik}^2$ it can be assumed that there exists a constant a, such that: $|a_{ik}| \leq a$, since a_{ik} depends on μ_{ik} and σ^{-2}, so this assumption is reasonable for Case 1. On the other hand, $\sum_{k=1}^{p} a_{ik}^2 \to \infty, p \to \infty$, since the variances are always positive.

Under these considerations it can be seen that

$$s_p^{-2}\sum_{k=1}^{p}\int_{\left|Y_{ijk}^2\right|>\varepsilon s_p}\left(Y_{ijk}^2-\mu_{ik}\right)^2 f_Y\left(y_{ijk}^2\right)dy \le a^2 s_p^{-2}\sum_{k=1}^{p}P\left(\left|Y_{ijk}^2-\mu_{ik}\right|>\varepsilon s_p\right),$$

where $\mu_{ik} = E\left(Y_{ijk}^2\right)$.

Applying Chebyshev's inequality:

$$\le a^2 s_p^{-2}\sum_{i=1}^{p}Var\left(Y_{ijk}^2\right)\varepsilon^{-2}s_p^{-2}$$

$$\le a^2\varepsilon^{-2}s_p^{-2} \to 0, p \to \infty$$

Therefore, the Lindeberg condition is hold, and so

$$X_{ij}\dot{\sim}N\left(\sum_{i=1}^{t}\left(p+0.5\,\mu_i^t\mu_i\sigma^{-2}\right), 2\sum_{i=1}^{t}\left(p+\mu_i^t\mu_i\sigma^{-2}\right)\right),$$

if p es is large enough.

Furthermore, under the null hypothesis, the variances are homogeneous and, therefore, p is sufficiently large and X_{ij} meet the assumptions of ANOVA.

Case 2. Correlated variables

If the vector $\boldsymbol{Y}_{ij}$ has a distribution N_p ($\boldsymbol{\mu}_i$, $\boldsymbol{\Sigma}$), the distribution of the random variable is not known $X_{ij} = \boldsymbol{Y}_{ij}^t\,\boldsymbol{Y}_{ij} = \sum_{k=1}^{p} Y_{ijk}^2$ and remains open. However, in this case, a simulation study was conducted to examine the performance of the proposal in terms of MANOVA.

Performance assessment of the proposal for Case 1

Sometimes, the sample size does not need to be so large to obtain satisfactory convergence results, so a simulation study is presented to evaluate the proposal with sample sizes that usually appear in practice in the case of a CRD. To assess the power of the ANOVA using the transformed data, a Monte Carlo simulation study was performed, using 2000 replicates (B) and a significance level of 0.05, under the assumption that the multiple responses come from a multivariate normal distribution. The simulation study was carried out with R software version 4.3.2 (R Core Team, 2023). The parameters to be set in the simulation of multivariate normal distribution data were: 1) mean vectors for each treatment, where the main vector will bee μ_1 and from which the differences between mean vectors were generated; 2) the variance matrix, for which uncorrelated variable matrices were considered ($\Sigma = I_p$); 3) number of treatments, t = 3, 5, 7; 4) number of variables: p = 3, 5, 7; and 5) number of replicates per treatment, r = 4, 8, 12, 16.

The performance of the proposed methodology through power estimation was evaluated as follows: 1) select p, t y r; 2) generate a random sample with multivariate normal distribution for each treatment, with vector of means μ_i and a common variance matrix Σ in all treatments; 3) obtain the transformed variables X_{ij} from the sample; 4) perform an ANOVA on the X_{ij} and obtain the degrees of freedom of the treatments (*glTrat*), the degrees of freedom of the error (*glTrat*) and the mean square of the error (CM_E); 5) calculate the means of each treatment as: $p + \hat{\mu}_i^t \hat{\mu}_i$; 6) calculate the average of the treatment means $(\hat{\bar{\mu}})$; 7) calculate the estimator of the non-centrality parameter, $\hat{\lambda} = r \sum_{i=1}^{t} (\hat{\mu}_i - \hat{\bar{\mu}})^2 \; CM_E^{-1}$; 8) obtain pf as the cumulative distribution function of the non-central F $\left(F_{glError}^{glTrat}(\hat{\lambda})\right)$; 9) obtain $F_{critical}$ as the quantile 1 - pf of the central F distribution $F_{glError}^{glTrat}$; 10) estimate the power of the ANOVA as assessed 1 - pf in the $F_{critical}$; 11) repeat B times steps 2–10; and 12) estimate the power as the proportion of times it was rejected in the H_0 simulation.

Comparison of MANOVA results with the proposal using ANOVA

To compare the proposed methodology to the MANOVA, data with multivariate normal distribution were simulated by varying the following parameters of interest:

Vectors of means (μ´*s*)

To calculate the distance between the means of the elements in μ_1, random numbers drawn from the uniform distribution using the *runif* function in R (R Core Team, 2023) were considered, so that it μ_1 was contained in one of the following intervals: [1, 10] or [50, 100]. For example, in the case of $\mu_1 \in [1, 10]$ whit p = 3, μ_1 = [2.9, 8.3, 5.7].

Between the mean vectors, two differences were considered between the μ_i´*s*. The small differences consist of differences between 10 and 20 % compared to μ_1. e.g. $\mu_2 = \mu_1$ x 1.1 (10 % difference to μ_1). The large differences include very high percentages between the differences between the mean vectors, ranging from 50 to 300 % compared to μ_1, e.g. $\mu_2 = \mu_1$ x 3 (300 % difference compared to μ_1).

Covariance Matrices (Σ´s)

For uncorrelated variables (Case 1), $\Sigma = \sigma^2 I_p$, considering σ^2 = 1 or 10. In the case of correlated variables (Case 2), the *genPositiveDefMat* function of the *clusterGeneration* library (Qiu and Joe, 2023) was used to generate positive definite random variance and covariance matrices. With this function, two matrices were generated: one with variances in the range [1, 2] (together with their respective covariances) and another with variances in the range [8, 12] to emulate the variances in the matrices studied in the

case of uncorrelated variables, with the following conditions: number of treatments: $t = 3,7$; number of variables: $p = 3,7$; and number of replicates per treatment: $r = 4,16$.
The data were analyzed with MANOVA and the value of the approximation to the Pillai Trace statistic $F\ (\upsilon_H, \upsilon_E)$ (Pillai and Samson, 1959) and the p - *value*. In cases of three treatments with three variables and four replicates, confidence intervals were obtained using Tukey's generalized conjecture (Seo and Fujikoshi, 1994). In other cases, this determination could not be made due to the limited degrees of freedom.
The data were then transformed using the proposed methodology and analyzed using ANOVA for the CRD model. In this case, the assumptions of normality and homoscedasticity were checked for compliance with the Shapiro-Wilk (SW) (Shapiro and Wilk, 1965) and Levene (L) (Levene, 1960) tests. If the transformed data did not meet any of the assumptions, the Box-Cox transformation (Box and Cox, 1964) was performed. If ANOVA rejects the null hypothesis of equal treatment effects, the comparison of means was performed using Tukey's test.

RESULTS AND DISCUSSION

The proposal to transform the vectors generated in a CRD into scalars and test the hypothesis of equal treatment effects with ANOVA performs satisfactorily in terms of power. There is a good performance of the test power with increasing sample size (Figures 1, 2, and 3), which is expected according to statistical theory (Casella, 2008; Hinkelmann and Kempthorne, 2005). As the number of variables (p) increases, the power of the test also increases, suggesting that the asymptotic result of convergence from the X_{ij} normal distribution works with medium sample sizes. Finally, the results show that, as the maximum differences between the vectors increase, the power of the test increases, which is also the expected behavior (Casella, 2008).
When there are four repetitions to achieve a power greater than 0.8, the difference between the mean vectors is required to be at least 40 %. For the cases of 8, 12, and 16 repetitions, this power is obtained when there is a 20 to 30 % difference between the mean vectors. It should be noted that this simulation study was much more extensive; however, only a small sample is presented to show the relevant aspects..

Comparison of MANOVA performance against the proposed methodology

The results of the performance comparison between the MANOVA and the methodology proposed in the test of the hypothesis of equal treatment effects are also satisfactory in the case of a CRD with multiple responses. The methodological approach in most cases detects smaller differences than the MANOVA (Table 2). In some cases, MANOVA cannot be used to analyze the data due to the limitations of the methodology (Table 3), because when r-1>p, the residual matrix W is not of full rank, and hence the test statistic $\Lambda^* = |W| \ / \ |B + W|$ used in MANOVA is not useful because $|W| = 0$ (Strang, 2006).
While the methodological proposal does not present such a problem, the multivariate case was transformed into the univariate case, and therefore the analysis in this case

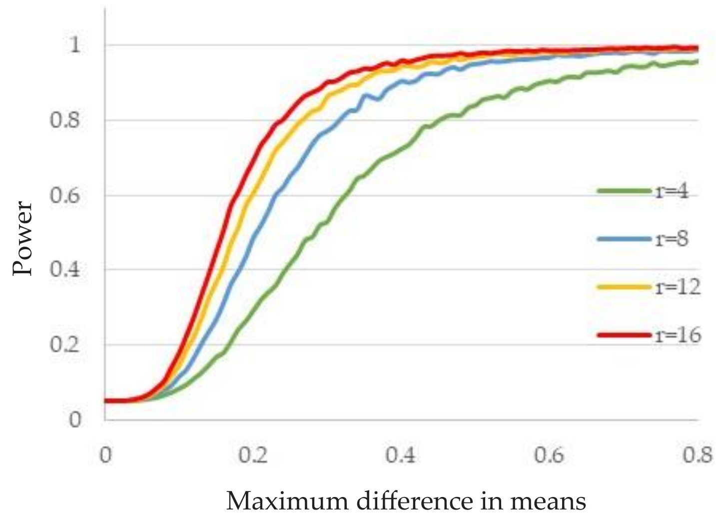

Figure 1. Estimated powers for $p = 3$ (number of variables) in $t = 3$ (number of treatments) with r replicates per treatment: $\mu_1 = [\mu_1, \mu_2, \mu_3]^t$, $\mu_1 \in [1, 10]$, $\mu_2 = (1 + escal\,[h]) \times \mu_1$, $\mu_3 = (1 + escal\,[h]) \times \mu_1$; scal=[0, 0.01, 0.02, …,0.8], for h=1, 2,…; $\Sigma = I_p$; number of Monte Carlo samples $B = 2000$; level of significance used: $\alpha = 0.05$.

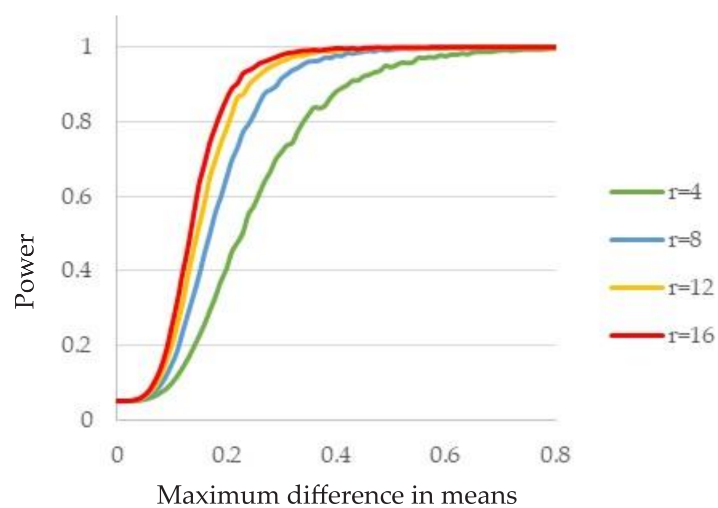

Figure 2. Estimated powers for $p = 5$ (number of variables) in $t = 3$ (number of treatments) with r replicates per treatment: $\mu_1 = [\mu_1, \mu_2, \mu_3, \mu_4, \mu_5]^t$, $\mu_1 \in [1, 10]$, $\mu_2 = (1 + escal\,[h]) \times \mu_1$, $\mu_3 = (1 + escal\,[h]) \times \mu_1$; scal=[0, 0.01, 0.02, …,0.8], for h=1, 2,…; $\Sigma = I_p$; number of Monte Carlo samples $B = 2000$; level of significance used: $\alpha = 0.05$.

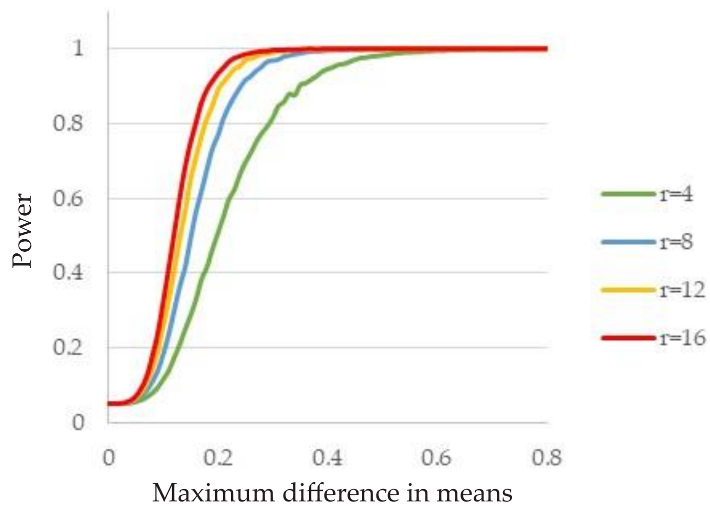

Figure 3. Estimated powers for p = 7 (number of variables) in t = 3 (number of treatments) with r replicates per treatment: $\mu_1 = [\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6, \mu_7]^t$, $\mu_1 \in [1, 10]$, $\mu_2 = (1 + escal\,[h]) \times \mu_1$, $\mu_3 = (1 + escal\,[h]) \times \mu_1$; scal=[0, 0.01, 0.02, …,0.8], for h=1, 2,…; $\Sigma = I_p$; number of Monte Carlo samples B = 2000; level of significance used: α = 0.05.

can be carried out. A satisfactory performance of the methodological approach can be observed when correlated observations are available (Table 4). When the maximum differences between the mean vectors are greater than 100 %, the MANOVA and the methodological approach reject the hypothesis of equal treatment effects, although the significance level is higher in the MANOVA. It was also observed that, in some cases, the transformed data did not meet the assumption of normality and generally met the assumption of homoscedasticity.

Often, the problem of non-compliance with the assumptions was solved with the Box-Cox transformation. No changes in the significance of the MANOVA and ANOVA are observed when changing the variance and covariance matrix. In all cases studied, Tukey's HSD methodology presented results in accordance with the simulation parameters. This research consisted of studying the parameters established in the methodology and generating results with simulation in 46 tables, although only a sample is presented here in order to show the relevant aspects.

The results of using the proposed methodology to test the hypothesis of equality of treatments with multiple responses generated by a CRD with correlated normal variables are satisfactory. This methodology could be applied to the case of other experimental designs after investigating their performance through simulation.

Table 2. Comparisons for case t = 3 (number of treatments), p = 3 (number of variables), r = 4 (number of replicates per treatment), and small differences between μ'^{s}_{i}.

μ	Σ	MANOVA $\sim F_{\upsilon H, \upsilon E}$ / Pr (> F)	Tukey's generalised conjecture	Transformed data p - value SW	L	ANOVA $\sim F_{\upsilon 1, \upsilon 2}$ / Pr (> F)	HSD Tukey
μ_1 = [2.9, 8.3, 5.7] $\mu_2 = \mu_1 \times 1.1$ $\mu_3 = \mu_1 \times 1.2$	I_3	1.638 0.201	T1 vs. T2: [-10.8, 7.4] T1 vs. T3: [-12.5, 5.7] T2 vs. T3: [-10.8, 7.4]	0.0201	0.960	5 0.035	T3 662.025 a T2 519.346 ab T1 398.076 b
μ_1 = [90.3, 57.6, 82.3] $\mu_2 = \mu_1 \times 1.1$ $\mu_3 = \mu_1 \times 1.2$	I_3	2.663 0.055	T1 vs. T2: [-32.1, 13.8] T1 vs. T3: [-55.1, -36.8] T2 vs. T3: [- 32.1, 13.8]	0.0318	0.954	597.2 2.7×10^{-10}	T3 33 811.123 a T2 23 828.921 b T1 16 239.583 c
μ_1 = [2.9, 8.3, 5.7] $\mu_2 = \mu_1 \times 1.1$ $\mu_3 = \mu_1 \times 1.2$	10 * I_3	0.366 0.890	T1 vs. T2: [-30.6, 27.2] T1 vs. T3: [-32.3, 25.5] T2 vs. T3: [- 30.6, 27.2]	0.0654	0.909	0.705 0.519	No significant difference
μ_1 = [90.3, 57.6, 82.3] $\mu_2 = \mu_1 \times 1.1$ $\mu_3 = \mu_1 \times 1.2$	10 * I_3	2.633 0.057	T1 vs. T2: [-51.9, 5.9] T1 vs. T3: [-74.9, -17.0] T2 vs. T3: [- 51.9, 5.9]	0.0351	0.9517	61.29 5.73×10^{-6}	T3 32 397.475 a T2 22 749.481 b T1 15 437.039 c

Table 3. Comparisons for case t = 3 (number of treatment), p = 7 (number of variables), r = 4 (number of replicates per treatment), small differences between μ'^{s}_{i}.

μ	Σ	MANOVA $\sim F_{\upsilon H, \upsilon E}$ / Pr (> F)	Transformed data p - value SW	L	ANOVA $\sim F_{\upsilon 1, \upsilon 2}$ / Pr (> F)	HSD Tukey
μ_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5] $\mu_2 = \mu_1 \times 1.1$ $\mu_3 = \mu_1 \times 1.2$	I_7	Residuals have rank 3 < 7	0.056	0.971	28.14 0.0001	T3 411.386 a T2 345.129 b T1 284.782 c
μ_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7] $\mu_2 = \mu_1 \times 1.1$ $\mu_3 = \mu_1 \times 1.2$	I_7	Residuals have rank 3 < 7	0.2872	0.9494	413.5 1.39×10^{-9}	T3 2332.021 a T2 1954.931 b T1 1611.166 c
μ_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5] $\mu_2 = \mu_1 \times 1.1$ $\mu_3 = \mu_1 \times 1.2$	10 * I_7	Residuals have rank 3 < 7	0.112	0.973	2.111 0.177	No significant difference
μ_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7] $\mu_2 = \mu_1 \times 1.1$ $\mu_3 = \mu_1 \times 1.2$	10 * I_7	Residuals have rank 3 < 7	0.050	0.905	32.34 7.78×10^{-5}	T3 2229.853 a T2 1866.041 b T1 1535.553 c

Table 4. Comparisons for case t = 3 (number of treatments), p = 3 (number of variables), r = 4 (number of replicates), large differences between μ'^{s}_{i}, correlated variables.

μ	Σ	MANOVA $\sim F_{vH,\, vE}$ / $Pr\ (> F)$	Tukey's generalised conjecture	Transformed data *p - value* SW	L	ANOVA $\sim F_{v1,\, v2}$ / $Pr\ (> F)$	HSD Tukey
μ_1 = [2.9, 8.3, 5.7] $\mu_2 = \mu_1 \times 3$ $\mu_3 = \mu_1 \times 2$	*A*	2.666 0.0548	T1 vs. T2: [-38.9, -28.6] T1 vs. T3: [-22.0, -11.7] T2 vs. T3: [11.1, 22.6]	0.498	0.295	183.4 5.09×10^{-8}	T2 1010.855 a T3 456.426 b T1 121.577 c
μ_1 = [90.3, 57.6, 82.3] $\mu_2 = \mu_1 \times 3$ $\mu_3 = \mu_1 \times 2$	*A*	2.667 0.055	T1 vs. T2: [-465.5, -455.2] T1 vs. T3: [-235.3, -225.0] T2 vs. T3: [224.4, 235.9]	0.439	0.542	654865 $< 2 \times 10^{-16}$	T2 164033.520 a T3 7286.500 b T1 18191.760 c
μ_1 = [2.9, 8.3, 5.7] $\mu_2 = \mu_1 \times 3$ $\mu_3 = \mu_1 \times 2$	*B*	2.659 0.055	T1 vs. T2: [-47.9, -19.6] T1 vs. T3: [-31.0, -2.7] T2 vs. T3: [2.7, 31.0]	0.731	0.245	166.9 7.71×10^{-8}	T2 1010.454 a T3 467.568 b T1 144.262 c
μ_1 = [90.3, 57.6, 82.3] $\mu_2 = \mu_1 \times 3$ $\mu_3 = \mu_1 \times 2$	*B*	2.667 0.055	T1 vs. T2: [-474.5, -446.2] T1 vs. T3: [-244.3, -216.0] T2 vs. T3: [214.4, 24.9]	0.331	0.451	32511 $<2 \times 10^{-6}$	T2 164116.560 a T3 72934.270 b T1 18242.260 c

$$A = \begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix};\ B = \begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$$

CONCLUSIONS

In this research, an alternative analysis was proposed to test the hypothesis of equality of effects between treatments and *post hoc* tests in the case of multiple responses. The simulation study shows that the performance of the proposal with small samples is satisfactory in terms of power and that it has advantages compared to MANOVA. Furthermore, the methodological approach allows for *post hoc* testing in the case of multiple responses in the completely randomized experimental design. The transformed data, from the proposed methodology, have problems holding the normality assumption when the number of variables (p) is relatively small, which is usually solved by the Box-Cox transformation.

REFERENCES

Anitha G, Ramireddy KVS, Kumar AA, Edukondalu L, Srinivasa Rao V. 2022. Physical and mechanical properties of cotton seeds. The Pharma Innovation Journal 11 (10): 1331–1333.

Box GEP, Cox DR. 1964. An analysis of transformations. Journal of the Royal Statistical Society, Series B (Methodological) 26 (2): 211–252.

Casella G. 2008. Statistical design. Springer Texts in Statistics: Gainesville, FL, USA. 330 p.

Correa-Lodoño GA, Castillo-Morales A. 2000. Tamaño de muestra para aproximación de un estadístico a la distribución normal. Agrociencia 34 (4): 467–484.

Dattalo P. 2013. Analysis of multiple dependent variables. Oxford University Press: New York, NY, USA. 190 p.

Driscoll WC. 1996. Robustness of the ANOVA and Tukey-Kramer statistical tests. Computers and Industrial Engineering 31 (1–2): 265–268. https://doi.org/10.1016/0360-8352(96)00127-1

El-Mostafa K, El-Kharrassi Y, Badreddine A, Andreoletti P, Vamecq J, El-Kebbaj M, Latruffe N, Lizard G, Nasser B, Cherkaoui-Malki M. 2014. Nopal cactus (*Opuntia ficus-indica*) as a source of bioactive compounds for nutrition, health and disease. Molecules 19 (9): 14879–14901. https://doi.org/10.3390/molecules190914879

Hair JF, Anderson RE, Tatham RL, Black WC. 1999. Análisis multivariante de datos (Quinta edición). Pearson Prentice Hall: Madrid, Spain. 814 p.

Hinkelmann K, Kempthorne O. 2005. Design and analysis of experiments. Wiley Interscience: Hoboken, NJ, USA. 808 p. https://doi.org/10.1002/0471709948

Hernández-Anguiano AM, Landa-Salgado P, Eslava-Campos CA, Vargas-Hernández M, Patel J. 2016. Microbiological quality of fresh nopal juice. Microorganisms 10 (4): 46. https://doi.org/10.3390/microorganisms4040046

Johnson RA, Wichern DW. 2007. Applied multivariate statistical analysis (Sixth edition). Prentice Hall: Upper Saddle River, NJ, USA. 773 p.

Levene H. 1960. Robust tests for equality of variances. *In* Olkin I. (ed.), Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press: Palo Alto, CA, USA, pp: 278–292.

Montgomery DC. 2004. Diseño y análisis de experimentos (Segunda edición). Limusa Wiley: Ciudad de México, México. 686 p.

Nishiyama T, Seo T. 2013. The multivariate Tukey-Kramer multiple comparison procedure among four correlated mean vectors. American Journal of Mathematical and Management Sciences 28 (1–2): 115–130. https://doi.org/10.1080/01966324.2008.10737720

Nishiyama T, Hyodo M, Seo T. 2014. Recent developments of multivariate multiple comparisons among mean vectors. SUT Journal of Mathematics 50 (2): 247–270. https://doi.org/10.55937/sut/1424793883

Pérez-López D de J, González-Huerta A, Franco-Mora O, Rubí-Arriaga M, Ramírez-Dávila JF, Castañeda-Vildózola A, Aquino-Martínez JG. 2014. Aplicación de métodos multivariados para identificar cultivares sobresalientes de haba para el Estado de México, México. Revista Mexicana de Ciencias Agrícolas 5 (2): 265–278. https://doi.org/10.29312/remexca.v5i2.965

Pillai KCS, Samson P. 1959. On Hotelling's generalization of T^2. Biometrika 46 (1–2): 160–168. https://doi.org/10.2307/2332818

Qiu W, Joe H. 2023. ClusterGeneration: random cluster generation (with specified degree of separation). R package version 1.3.8. https://CRAN.R-project.org/package=clusterGeneration

R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Rencher AC, Christensen WF. 2012. Methods of multivariate analysis (Third edition). Wiley: Hoboken, NY, USA. 727 p.

Salkind NJ. 2010. Encyclopedia of research design, volume 3. Sage Publications: New York, NY, USA. 1644 p. https://doi.org/10.4135/9781412961288

Seo T, Nishiyama T. 2008. On the conservative simultaneous confidence procedures for multiple comparisons among mean vectors. Journal of Statistical Planning and Inference 138 (11): 3448–3456. https://doi.org/10.1016/j.jspi.2007.03.061

Seo T. 2002. The effect of nonnormality on the upper percentiles of T^2_{max} statistic in elliptical distributions. The Japan Statistical Society 32 (1): 57–76. https://doi.org/10.14490/jjss.32.57

Seo T, Mano S, Fujikoshi Y. 1994. A generalized Tukey conjecture for multiple comparisons among mean vectors. Journal of the American Statistical Association 89 (426): 676–679. https://doi.org/10.1080/01621459.1994.10476792

Shapiro SS, Wilk MB. 1965. An analysis of variance test for normality (complete samples), Biometrika 52 (3–4): 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Sosnówka-Czajka E, Skomorucha I, Obremski K, Wojtacha P. 2023. Performance and meat quality of broiler chickens fed with the addition of dried fruit pomace. Poultry Science 102 (6): 102631. https://doi.org/10.1016/j.psj.2023.102631

Strang G. 2006. Linear algebra and its applications (Fourth edition). Cengage Learning: Boston, MA, USA. 496 p.

Warne RT, Lazo M, Ramos T, Ritter N. 2012. Statistical methods used in gifted education journals, 2006-2010. Gifted Child Quarterly 56 (3): 134–149. https://doi.org/10.1177/0016986212444122

Warne RT. 2014. A primer on multivariate analysis of variance (MANOVA) for behavioral scientists. Practical Assessment, Research, and Evaluation 19 (1): 17. https://doi.org/10.7275/sm63-7h70