A score test for the agronomical overlap effect in a two-way classification model

Test score para el efecto del solapamiento agronómico en un modelo de clasificación de dos vías

Aquiles Darghan[1], Sinha Surendra[2], and Julio Monroy[3]

ABSTRACT

In some agricultural research, a treatment applied to an experimental unit may affect the response in the neighboring experimental units. This phenomenon is known as overlap. In this article, a test to evaluate this effect in the Draper and Guttman model was developed by imposing side conditions on the parameters of a two-way classification model to obtain a re-parameterized model which can be used in different neighboring patterns of experimental units, usually plants within a crop, whenever the nearest neighbor is considered a directly affected experimental unit and the two-way model is used. Three methods, namely maximum likelihood, least squares with side conditions and generalized inverse, were used to estimate the parameters of the original model in order to calculate the value of the test statistics for the null hypothesis associated with the absence of the overlapping effect. The three alternatives were invariant with respect to the use of test. The proposed test is simple to adopt and can be implemented in agronomy since its asymptotic nature is in agreement with the large number of experimental units which generally exist in this type of research, where each plant represents the experimental unit being assessed.

Key words: perpendicular projection operator, side conditions, experimental design.

RESUMEN

En algunas investigaciones agrícolas, un tratamiento aplicado sobre una unidad experimental puede afectar la respuesta de unidades experimentales vecinas. Este fenómeno es conocido como solapamiento. En este artículo se desarrolló un test para evaluar este efecto, sobre el modelo de Draper y Guttman, mediante la imposición de condiciones laterales sobre los parámetros del modelo de clasificación de dos vías para obtener un modelo reparametrizado, el cual puede usarse bajo diferentes patrones de vecindad de las unidades experimentales, usualmente plantas dentro de un cultivo, siempre y cuando sea considerado el vecino más cercano como la unidad experimental directamente afectada y el modelo sea de dos vías. Fueron usados tres métodos para estimar los parámetros del modelo original, a saber, el método de máxima verosimilitud, el método de mínimos cuadrados con imposición de condiciones laterales y el uso de una inversa generalizada para calcular el valor del estadístico de prueba para la hipótesis nula asociada a la ausencia del efecto de solapamiento. Las tres alternativas resultaron invariantes con respecto al uso del test. La prueba propuesta es sencilla de adoptar y se puede implementar en el campo de la agronomía, ya que su naturaleza asintótica está de acuerdo con el gran número de unidades experimentales que generalmente existen en este tipo de investigaciones, donde cada planta representa la unidad experimental evaluada.

Palabras clave: operador de proyección perpendicular, condiciones laterales, diseño experimental.

Introduction

In some agricultural tests, a treatment that is applied to an experimental unit may affect the response in the neighboring experimental units. This phenomenon is known as overlap. For example, in variety testing, the effect of a neighbor can be attributed to differences in height between plants, strength of roots, and date of germination, among others. Hide and Read (1990) discussed this situation in potato cultivation. The treatments applied to crops for fertilization plans, irrigation and pesticide applications can be dispersed to other plots or adjacent experimental units, which can affect the response in the neighboring units. An example of this situation for an irrigation experiment can be found with Bhalli *et al*. (1964). This overlapping phenomenon has been modeled by several researches; for example, Pearce (1957) considered a model in which each treatment had a direct influence on the plot in which they were applied and an overlap effect on the neighboring plots. Draper and Guttman (1980) also considered such a

Received for publication: 9 October, 2014. Accepted for publication: 27 November, 2014. Doi: 10.15446/agron.colomb.v32n3.46086

[1] Department of Agronomy, Faculty Agricultural Sciences, Universidad Nacional de Colombia. Bogota (Colombia). aqedarghanco@unal.edu.co
[2] Institute for Applied Statistics and Computers, Universidad de Los Andes. Merida (Venezuela).
[3] Department of Mathematics, Faculty of Science, Universidad de Pamplona. Pamplona (Colombia).

model, discussing some approximate testing methods as well as a confidence interval for the overlap effect. Draper and Guttman (1980) used a nonlinear model in which the overlapping was attributed to all the effects considered in the model. Such a model is written as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \alpha\mathbf{W}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where the random vector $\mathbf{Y}$ denotes the response of $n$experimental units, $\mathbf{X}$ is a known design matrix of dimension $n \times p$ with rank $q > p$, $\boldsymbol{\beta}$ is a p-dimensional vector of unknown parameters consisting of the effects of blocks and treatments, α is the overlapping coefficient, and $\mathbf{W} = (w_{ij})$ is a matrix of known weights of dimension $n \times n$, where w_{ij} denotes the effect of unit j on unit i; such that $\sum_{j=1}^{n} w_{ij} = 1$ for all i; and $w_{ii} = 0$, $w_{ij} \geq 0$ for all i and j. The fact that the $\mathbf{X}$ matrix is of an incomplete rank generates the possibility of imposing side conditions on the parameters so that an estimator for the $\boldsymbol{\beta}$ vector can be obtained. Finally, it is assumed that the distribution of the error vector $\boldsymbol{\varepsilon}$ is normal and independent with an expected value of zero ($E(\boldsymbol{\varepsilon}) = 0$) and variance of $\sigma^2 I$, where $\mathbf{I}$ is an identity matrix. Shukla and Subrahmanyan (1999) considered a generalization of the first model represented by (1), which only included a subset of all the direct effects influencing the neighboring plots. In the following sections of the article, some important results which allowed the construction of Rao's score test for α of model (1) are presented; in the methods section, we explain the side condition technique, the obtainment of several matrices associated with the model of Draper and Guttman (1980) and, finally, we present Rao's score test under the imposition of side conditions.

Methods

Side conditions and associated matrices in the Draper and Guttman model

The technique to impose side conditions is very well-known in the area of experimental design since it provides the necessary (linear) restrictions which can assure that the estimation of the parameters is unique. Another use of side conditions is the imposition of arbitrary restrictions on the estimators so that the normal equations can be simplified. In this case, the estimators have exactly the same behavior when a generalized inverse of $\mathbf{X}^T\mathbf{X}$ is used to obtain the vector of parameters. Let q be the rank of the design matrix $\mathbf{X}$ where $q<p<n$ in (1), then $\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}$ represents a set of p estimable functions of $\boldsymbol{\beta}$. The side conditions must be non-estimable functions of $\boldsymbol{\beta}$. Since the rank of $\mathbf{X}$ is q and, hence, the deficiency of this rank is p-q; conditions such as $\mathbf{L}\boldsymbol{\beta} = 0$ or $\mathbf{L}\hat{\boldsymbol{\beta}} = 0$ to obtain a unique solution for $\hat{\boldsymbol{\beta}}$ should be defined where $\mathbf{L}$ is a matrix of dimension $(p\text{-}q) \times p$ with a rank of $\mathbf{L}$ being p-q such that $\mathbf{L}\boldsymbol{\beta}$ is a set of non-estimable functions. The (1) model under the null hypothesis $\alpha = \alpha_0$ along with the side conditions is written as:

$$\begin{bmatrix}\mathbf{Y}_{n\times1}\\ \mathbf{0}_{l\times1}\end{bmatrix} = \begin{bmatrix}\mathbf{X}_{n\times p}\\ \mathbf{L}_{l\times p}\end{bmatrix}\boldsymbol{\beta}_{p\times1} + \alpha_0 \begin{bmatrix}\mathbf{W}_{n\times n} & \mathbf{0}_{n\times l}\\ \mathbf{0}_{l\times n} & \mathbf{0}_{l\times l}\end{bmatrix}\begin{bmatrix}\mathbf{X}_{n\times p}\\ \mathbf{L}_{l\times p}\end{bmatrix}\boldsymbol{\beta}_{p\times1} + \begin{bmatrix}\boldsymbol{\varepsilon}_{n\times1}\\ \mathbf{0}_{l\times1}\end{bmatrix}$$

or

$$\mathbf{Y}_A = \mathbf{U}\boldsymbol{\beta} + \alpha_0\mathbf{W}_A\mathbf{U}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_A \tag{2}$$

where $\mathbf{U}^T = [\mathbf{X}^T\ \mathbf{L}^T]$ has rank p, $\mathbf{W}_A$ is a matrix of weights with l additional rows, and where l represents the number of imposed linear restrictions. The $\mathbf{U}^T\mathbf{U}$ matrix is of a $p \times p$ dimension and rank p, which generate the system of normal equations $\mathbf{U}^T\mathbf{U}\hat{\boldsymbol{\beta}} = \mathbf{U}^T\mathbf{Y}_A$, where $\mathbf{Y}_A$ is the vector of responses expanded with zeros and has a unique solution for $\hat{\boldsymbol{\beta}}$ (Graybill, 1976). In order to illustrate the above described procedure, let us suppose that we have a two-way classification model (treatments and blocks) which will be written as $Y_{ij} = \mu + \tau_i + \delta_j + \varepsilon_{ij}$; ($i$ = 1, 2,..., t >1); (j = 1, 2,..., b >1) with τ_i as the effect of the ith treatment and δ_j as the effect of the jth block, which in matrix form is written as $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. In this model, there is a deficiency in rank p-q = 2, with $p = t + b + 1$ and $q = t + b - 1$. In order to solve this problem of rank deficiency in this design model, the usual non-estimable functions $\sum_{i=1}^{t}\tau_i = 0$ and $\sum_{j=1}^{b}\delta_j = 0$ can be used, which can be expressed in the matrix form as the following set of non-estimable functions:

$$\mathbf{L}\boldsymbol{\beta} = \begin{bmatrix}0 & \mathbf{J}_t^T & 0\\ 0 & 0 & \mathbf{J}_b^T\end{bmatrix}\boldsymbol{\beta} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

denoting $\mathbf{J}_t$ and $\mathbf{J}_b$ as the vectors of ones of the t and b elements, respectively. The matrices $\mathbf{U}^T\mathbf{U}$, $(\mathbf{U}^T\mathbf{U})^{-1}$, $(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{X}^T$ and $\mathbf{X}(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{X}^T$ are needed in the estimation of the vector of parameters $\boldsymbol{\beta}$, error variance and, therefore, for obtaining Rao's score test for overlapping. However, here, we only present the final expression (Darghan, 2010); it is the comprehensive development of related matrix calculations), namely:

$$\mathbf{X}\left(\mathbf{U}^T\mathbf{U}\right)^{-1}\mathbf{X}^T = b^{-1}\mathbf{X}_\tau\mathbf{X}_\tau^T + t^{-1}\mathbf{X}_\delta\mathbf{X}_\delta^T - n^{-1}\mathbf{J}\mathbf{J}^T = \mathbf{M}_\tau + \mathbf{M}_\delta - \mathbf{M}_J \tag{3}$$

with $n=bt$, $\mathbf{X} = (\mathbf{J}\ \mathbf{X}_\tau\ \mathbf{X}_\delta)$, where $\mathbf{X}_\tau$ and $\mathbf{X}_\delta$ are the submatrices of the design matrix associated with the effects of the treatments and blocks respectively, $\mathbf{M}_\tau = b^{-1}\mathbf{X}_\tau\mathbf{X}_\tau^T$, $\mathbf{M}_\delta = t^{-1}\mathbf{X}_\delta\mathbf{X}_\delta^T$ and $\mathbf{M}_J = n^{-1}\mathbf{J}\mathbf{J}^T$, which are all perpendicular projection matrices onto the column spaces of $\mathbf{X}_\tau$, $\mathbf{X}_\delta$

$\hat{\beta}$

and **J**, respectively, denoted with $C(X_\tau)$, $C(\mathbf{X}_\delta)$ and $C(\mathbf{J})$. This last statement is shown in a later theorem. The space spanned by the columns of a vector or matrix called a column space is written as $C(\quad)$, placing in the brackets the vector or matrix of interest. Another interesting use of side conditions for obtaining the re-parameterization of a model can be described as follows: Let $\mathbf{L\beta} = 0$ define the side conditions in the model (1). If $\mathbf{L\beta} = 0$ holds, then $\boldsymbol{\beta}$ belongs to the orthogonal ($\perp$) complement of $C(L^T)$, namely $\boldsymbol{\beta} \in C(\mathbf{L}^T)^\perp$. Let $\mathbf{Z}$ be such that $C(\mathbf{L}^T) = C(\mathbf{Z})$, then $\boldsymbol{\beta} = \mathbf{Z}\gamma$ for some γ; then, by substituting $\boldsymbol{\beta}$ with $\mathbf{Z}\gamma$ and defining $\mathbf{X}_0 = \mathbf{XZ}$ in model (1), we get

$$\mathrm{Y} = \mathrm{X}_0\boldsymbol{\gamma} + \alpha \mathrm{W}\mathrm{X}_0\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{4}$$

Besides, if $L\beta$ is not estimable, then $C(X) = C(X_0)$ and the third model represented by (4) is a re-parameterization of model (1) (Christensen, 2011).

Rao's score test

This section summarizes the theory of likelihood for the score test. Further information about score tests can be found in (Rao, 1973) and (Cox and Hinkley, 2000). The article published by (Rao, 1948) introduced the fundamental principle of a test based on the score function as an alternative method to the likelihood ratio test and to Wald's method. Several authors have described the attractive properties of the method; among them, Chandra and Joshi (1985), Bera and Mckenzie (1986) and Bera and Bilias (2001). For a better understanding of the test, we introduce its notation, which is maintained until the development of the test for α. Suppose there are n independent observations $Y_1,..., Y_n$ identically distributed with a density function of $f(y; \theta)$ that satisfies the conditions of a regularity given by (Rao, 1973), where θ is a vector of parameters of dimension $p \times 1$, with $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \Re^p$. The log-likelihood function, the score function and the expected information matrix are defined respectively as $l(\boldsymbol{\theta}) = \log\left(\prod_{i=1}^{n} f(y_i;\boldsymbol{\theta})\right)$, $s(\boldsymbol{\theta}) = \partial l(\boldsymbol{\theta}) / \partial\boldsymbol{\theta}$, $F(\boldsymbol{\theta}) = -E\left(\partial^2 l(\boldsymbol{\theta}) / \partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T\right)$.

The hypothesis to test is H_0: $h(\boldsymbol{\theta}) = c$, where $h(\boldsymbol{\theta})$ is a r-dimensional vector function of $\boldsymbol{\theta}$ with $r \leq p$ and c is a vector of known constants. In addition, it is assumed that $H(\boldsymbol{\theta}) = \partial h(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ is a full column rank matrix. In the one-dimensional case of c, when $p = 1$ with H_0: $\theta = \theta_0$ and using the Neyman-Pearson lemma, (Rao and Poty, 1946) proved that the most powerful local test for H_0 is $ks(\theta_0) > \lambda$, where λ is determined such that the size of the test is equal to a pre-assigned value of the significance level, with k as +1 or -1, respectively for alternatives $\theta > \theta_0$ or $\theta < \theta_0$ (Wald, 1941), and as the result of which under H_0, $s(\theta_0)$ has asymptotically normal distribution with a mean of zero and a variance of $F(\theta)$. This result led (Rao, 1948) to suggest a test based on $s^2(\theta_0) / F(\theta)^{-1}$ as a χ^2_1 variable when n is large. The generalization of the test when $p \geq 2$ was developed by (Rao, 1948), which led him to the statistical test based on $s(\theta_0)^T F(\theta)^{-1} s(\theta_0)$, which has a χ^2_g distribution with g degrees of freedom.

Results

Here, the results obtained by applying Rao's test score to the overlap model are presented as well as some theorems generated as a result of the partition of the design matrix to solve the problem of incomplete rank in the estimation process and the statistic test to evaluate the effect of overlapping.

Suppose we want to test the hypothesis H_0: $\alpha=0$ against H_0: $\alpha \neq 0$ in model (1), under H_0, this model is the usual linear model $Y = X\beta+\varepsilon$; however, X is not of a full column rank; in this sense, the p parameters in β are not unique. We then ascertained whether β could be estimated. Using least-squares in (1), we obtained the normal equations $X^TX\hat{\beta} = X^TY$. However, since X was not of a full column rank, X^TX had no inverse; as a consequence, the normal equations did not have a unique solution, despite the fact that this system was consistent if and only if $X^TX(X^TX)^-X^TY(X^TY)$, where $(X^TX)^-$ is a generalized inverse of (X^TX). Since the normal equations were consistent, a solution is:

$$\hat{\beta} = (X^TX)^-X^TY \tag{5}$$

For a particular generalized inverse $(X^TX)^-$, the expected value of $\hat{\beta}$ in (1) under H_0 is $E(\hat{\beta}) = (X^TX)^-X^TY\beta$, thus $\hat{\beta}$ is not an unbiased estimator of β; furthermore, the expression $(X^TX)^-X^TY\beta$ is not invariant to the choice of $(X^TX)^-$, thus, $\hat{\beta}$ in (5) does not estimate β. With respect to the estimation of σ^2, in model (1), we define:

$$\hat{\sigma}^2 = \mathbf{Y}^T\left[\mathbf{I}_n - \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^-\mathbf{X}^T\right]\mathbf{Y} \,/\left(n-q\right) \tag{6}$$

where n is the number of rows of $\mathbf{X}$ which has rank q. The estimator in (6) is unbiased for σ^2 and invariant to the choice of $\hat{\boldsymbol{\beta}}$ and to the choice of generalized inverse $(\mathbf{X}^T\mathbf{X})^-$. For the non-full column rank model (1), under H_0, we now assume that $\boldsymbol{\varepsilon}$ is distributed as $N_n(0; \sigma^2\mathbf{I})$. With the normality assumption we can obtain a maximum likelihood estimator for (1) under H_0. In this case, we have the same estimator for $\boldsymbol{\beta}$ which is given by the least square estimation but a biased estimator for σ^2 is obtained, which can be written as $\hat{\sigma}^2 = [n^{-1}\mathbf{Y}^T\mathbf{I}_n - \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T]\mathbf{Y}$. The lack of uniqueness in the estimate for $\boldsymbol{\beta}$ using two different estimation methods in (1) leads us to the use of (2). In this model, the problem of the rank of $\mathbf{X}$ is solved with the imposition of side conditions to proceed with the estimation of the parameters using least squares. Using model (2) restricted by H_0 we obtain $\hat{\beta} = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{Y}_A = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{Y}$ and $\hat{\sigma}^2 = \mathbf{Y}^T[\mathbf{I}_n - \mathbf{X}(\mathbf{U}^T\mathbf{U})^{-}\mathbf{X}^T]$ Y / $(n - q)$, which are unique and unbiased. Also the error vector has two interesting features: it has singular normal distribution and, therefore, the log-likelihood function doesn't exist and contains the matrix given in (3), which is very important in the development of the theorems presented below and can be extended in the balanced two-way classification model with more than one observation per cell by just modifying the design and weight matrices in order to adjust for the replications of each treatment.

Theorem 1. Let $\mathrm{M}_\tau = b^{-1}\mathbf{X}_\tau\mathbf{X}^T_\tau$, $\mathbf{M}_\delta = t^{-1}\mathbf{X}_\delta\mathbf{X}^T_d$ and $\mathbf{M}_J = n^{-1}\mathbf{J}\mathbf{J}^T$ be perpendicular projection matrices onto the respective column spaces of $\mathbf{X}_\tau$, $\mathbf{X}_\delta$ and $\mathbf{J}$, such that $\mathrm{C}(\mathbf{M}_\tau) = \mathrm{C}(\mathbf{X}_\tau)$, $\mathrm{C}(\mathbf{M}_\delta) = \mathrm{C}(\mathbf{X}_\delta)$, and $\mathrm{C}(\mathbf{M}_J) = \mathrm{C}(\mathbf{J})$.

Proof. An inspection of the matrices $\mathbf{M}_\tau$, $\mathbf{M}_\delta$ and $\mathbf{M}_J$ allows for a verification of the symmetry and idempotent properties, which is enough to prove that these matrices are all perpendicular projection matrices onto the respective column spaces of $\mathbf{X}_\tau$, $\mathbf{X}_\delta$ and $\mathbf{J}$. To prove that $\mathrm{C}(\mathbf{M}_\tau) = \mathrm{C}(\mathbf{X}_\tau)$, we can refer to the B.51 proposition in (Christensen, 2011). In the cases where $\mathrm{C}(\mathbf{M}_\delta) = \mathrm{C}(\mathbf{X}_\delta)$ and $\mathrm{C}(\mathbf{M}_J) = \mathrm{C}(\mathbf{J})$, the same procedure can be followed.

Theorem 2. For matrices $\mathbf{M}_\tau$, $\mathbf{M}_\delta$ and $\mathbf{M}_J$, it holds that if $\mathbf{M}_\tau\mathbf{M}_\delta = \mathbf{M}_J$, $\mathbf{M}_\tau\mathbf{M}_J = \mathbf{M}_J$ and $\mathbf{M}_\delta\mathbf{M}_J = \mathbf{M}_J$, then $\mathrm{C}(\mathbf{I} - \mathbf{X}(\mathbf{U}^T\mathbf{U})^{-}\mathbf{X}^T) = \mathrm{C}(\mathbf{I} - \dot{\mathbf{M}}) \subset \mathrm{C}(\mathbf{M}_J)^{\perp}$.

Proof. The proof is obvious by the very definition of the matrices M_τ, M_δ, and M_J in the two-way classification design without interaction. To prove that $\mathrm{C}(\mathbf{I} - \dot{\mathbf{M}}) \subset \mathrm{C}(\mathbf{M}_J)^{\perp}$, it is sufficient to multiply $\mathrm{I} - \dot{\mathrm{M}}$ with M_J and observe that their product is zero. In addition, the rank of $\mathrm{C}(\mathbf{M}_J)^{\perp}$ as $n - 1$ is greater than the rank of $\mathrm{C}(\mathbf{I} - \dot{\mathbf{M}})$, whose value is $n - q$, whenever $q > 1$ in the two-way classification model without interaction.

Theorem 3. $\dot{\mathrm{M}}$ is the perpendicular projection matrix on $\mathrm{C}(\mathbf{X})$, where $\dot{\mathbf{M}} = \mathbf{X}(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{X}^T$.

Proof. In order to prove that $\dot{\mathrm{M}}$ is the perpendicular projection matrix on $\mathrm{C}(\mathbf{X})$, it is necessary to verify that $\dot{\mathrm{M}}$ is symmetric and idempotent. The symmetry obviously results from equation (3) when verifying that $\dot{\mathrm{M}} = \dot{\mathrm{M}}^T$. So we only have to prove idempotence ($\dot{\mathbf{M}} = \dot{\mathbf{M}}^2$). Since $\dot{\mathbf{M}} = \mathbf{M}_\tau + \mathbf{M}_\delta - \mathbf{M}_J$, therefore $\dot{\mathbf{M}}^2 = (\mathbf{M}_\tau + \mathbf{M}_\delta - \mathbf{M}_J)^2$, and using the previous theorems, we have $\dot{\mathbf{M}} = \dot{\mathbf{M}}^2$.

Theorem 4. The matrix $(\mathbf{U}^T\mathbf{U})^{-1} = (\mathbf{X}^T\mathbf{X} + \mathbf{L}^T\mathbf{L})^{-1}$ is a g-inverse of $\mathbf{X}^T\mathbf{X}$.

Proof. A generalized inverse of $\mathrm{X}^T\mathrm{X}$ is any G matrix such that $\mathbf{X}^T\mathbf{X}\mathbf{G}\mathbf{X}^{Tx}\mathbf{X} = \mathbf{X}^T\mathbf{X}$. Let $\mathbf{G} = (\mathbf{X}^T\mathbf{X} + \mathbf{L}^T\mathbf{L})^{-1}$, then $\mathbf{X}^T\mathbf{X}\mathbf{G}\mathbf{X}^T\mathbf{X} = \mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X} + \mathbf{L}^T\mathbf{L})^{-1}$, but replacing $\dot{\mathrm{M}}$ with theorem 3, we have $\mathbf{X}^T\mathbf{X}\mathbf{G}\mathbf{X}^T\mathbf{X} = \mathbf{X}^T\dot{\mathbf{M}}\mathbf{X}$ and as $\dot{\mathbf{M}}\mathbf{X} = \mathbf{X}$ for the same theorem, we have $\mathbf{X}^T\mathbf{X}\mathbf{G}\mathbf{X}^T\mathbf{X} = \mathbf{X}^T\mathbf{X}$.

Theorem 5. The matrices $\mathbf{M}_\alpha = \mathbf{M}_\tau - \mathbf{M}_J$ and $\mathbf{M}_\eta = \mathbf{M}_\delta - \mathbf{M}_J$ are perpendicular projection operators; additionally, $\mathrm{C}(\mathbf{M}_\alpha) \perp \mathrm{C}(\mathbf{M}_\eta)$, $\mathrm{C}(\mathbf{M}_\alpha) \perp \mathrm{C}(\mathbf{M}_\delta)$ and $\mathrm{C}(\mathbf{M}_\eta) \perp \mathrm{C}(\mathbf{M}_\tau)$.

Proof. To start, we will prove that $\mathbf{M}_\alpha$ and $\mathbf{M}_\eta$ are idempotent matrices, which we obtain by verifying that $(\mathbf{M}_\tau - \mathbf{M}_J)^2 = \mathbf{M}^2_\tau + \mathbf{M}^2_J - \mathbf{M}_\tau\mathbf{M}_J - \mathbf{M}_J\mathbf{M}_\tau = \mathbf{M}_\tau - \mathbf{M}_J$. The symmetry of $\mathbf{M}_\tau - \mathrm{M}_J$, obviously results when we substitute M_τ with $b^{-1}\mathbf{X}_\tau\mathbf{X}^T_J$ and $\mathbf{M}_J$ with $n^{-1}\mathbf{J}\mathbf{J}^T$, from which we obtain $\mathbf{M}^T_\alpha = \mathbf{M}_\alpha$. The same procedure may be used in the case of $\mathbf{M}_\eta$. In order to prove that $\mathrm{C}(\mathbf{M}_\alpha) \perp \mathrm{C}(\mathbf{M}_\delta)$ or, in other words, to prove that $\mathrm{C}(\mathbf{M}_\alpha)$ is contained in the orthogonal complement of $\mathrm{C}(\mathbf{M}_\eta)$, it should be sufficient to verify that the product of the matrices associated with these spaces is null. In the first case, we have $(\mathbf{M}_\tau - \mathbf{M}_J)(\mathbf{M}_\delta - \mathbf{M}_J) = \mathbf{M}_\tau\mathbf{M}_\delta - \mathbf{M}_\tau\mathbf{M}_J - \mathbf{M}_J\mathbf{M}_\delta + \mathbf{M}^2_J = 0$; for the cases $\mathrm{C}(\mathbf{M}_\alpha) \perp \mathrm{C}(\mathbf{M}_\delta)$ and $\mathrm{C}(\mathbf{M}_\eta) \perp \mathrm{C}(\mathbf{M}_\tau)$ the same procedure is followed.

To develop Rao's score test for overlapping, we need the expected information matrix which is constructed from the log-likelihood function. In the (1) model, the log-likelihood is singular due to the deficiency in the rank of $\mathbf{X}$. Thus, models (1) and (2) are not used in the construction of the test, but we still use their estimates for $\boldsymbol{\beta}$ and σ^2. The log-likelihood function for (4) is:

$$l\left(\alpha,\gamma,\sigma^2\right) = -n\log\left(2\pi\right)/2 - n\log\left(\sigma^2\right)/2 - \left\|\left(\mathbf{Y} - \mathbf{X}_0\gamma + \alpha\mathbf{W}\mathbf{X}_0\gamma\right)/\sigma\right\|^2/2$$

from which, under H_0, we obtain: $\hat{\boldsymbol{\gamma}} = \left(\mathbf{X}_0^T\mathbf{X}_0\right)^{-1}\mathbf{X}_0^T\mathbf{Y}$ and $\hat{\sigma}^2 = n^{-1}\mathbf{Y}^T\left[\mathbf{I}_n - \mathbf{M}_0\right]\mathbf{Y}$ as the perpendicular projection matrix on $\mathrm{C}(\mathrm{X}_0)$. The score vector is:

$$s\left(\alpha,\gamma,\sigma^2\right) = \frac{1}{\sigma^2}\begin{bmatrix} \left(\mathbf{W}\mathbf{X}_0\gamma\right)^T\left(\mathbf{Y} - \mathbf{Q}\gamma\right) \\ \mathbf{X}_0^T\left(\mathbf{Y} - \mathbf{Q}\gamma\right) \\ -n/2 + \left\|\left(\mathbf{Y} - \mathbf{Q}\gamma\right)/\sigma\right\|^2/2 \end{bmatrix}$$

with $\mathbf{Q} = (\mathbf{I}_n + \alpha\mathbf{W})\mathbf{X}_0$. If we denote $s_\alpha(\alpha, \gamma, \sigma^2)$ the component of the score vector corresponding to the parameter of interest (α), the score statistic to test H_0: $\alpha = 0$ is:

$$S_{ovp} = s_\alpha\left(0,\hat{\gamma},\hat{\sigma}^2\right)^T \hat{\mathcal{F}}^{11}\left(0,\hat{\gamma},\hat{\sigma}^2\right)s_\alpha\left(0,\hat{\gamma},\hat{\sigma}^2\right)$$

where $\hat{\boldsymbol{\gamma}}$ and $\hat{\sigma}^2$ are the maximum likelihood estimators of $\boldsymbol{\gamma}$, σ^2 in (4) and $\hat{\mathcal{F}}^{11}$ $(0, \hat{\boldsymbol{\gamma}}, \hat{\sigma}^2)$ is the component in the expected information matrix evaluated at $\alpha = 0$. The expected information inverse matrix is:

$$\hat{\mathcal{F}}^{-1}\left(0,\hat{\gamma},\hat{\sigma}^2\right) = \hat{\sigma}^2 \begin{bmatrix} \left\|\hat{\mathbf{H}}\right\|^{-2} + \left\|\hat{\mathbf{H}}\right\|^{-2} \hat{\mathbf{H}}^T\mathbf{X}_0\hat{\mathbf{G}}\mathbf{X}_0^T\hat{\mathbf{H}}\left\|\hat{\mathbf{H}}\right\|^{-2} & -\left\|\hat{\mathbf{H}}\right\|^{-2}\hat{\mathbf{H}}^T\mathbf{X}_0\hat{\mathbf{G}} & 0 \\ -\left\|\hat{\mathbf{H}}\right\|^{-2}\hat{\mathbf{G}}\mathbf{X}_0^T\hat{\mathbf{H}} & \hat{\mathbf{G}} & \mathbf{0} \\ 0 & \mathbf{0} & 2\hat{\sigma}^2/n \end{bmatrix}$$

where $\hat{\mathbf{H}} = \mathbf{W}\mathbf{X}_0\boldsymbol{\gamma}$ and $\hat{\mathbf{G}} = (\mathbf{X}_0^T(\mathbf{I}_n - \mathbf{M}_{\hat{H}})\mathbf{X}_0)^{-1}$ with $\mathbf{M}_{\hat{H}} = \hat{\mathbf{H}}\|\hat{\mathbf{H}}\|^{-2}\hat{\mathbf{H}}^T$ ($\mathbf{M}_{\hat{H}}$ as the perpendicular projection operator on C($\hat{\mathbf{H}}$)). Now, when evaluating $s_\alpha(\alpha, \boldsymbol{\gamma}, \sigma^2)$ in $s_\alpha(0, \hat{\boldsymbol{\gamma}}, \hat{\sigma}^2)$ and grouping terms, we obtain:

$$S_{ovp} = \hat{\sigma}^2\left(\hat{\mathbf{H}}^T\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^T\left(\mathbf{I}_n + \mathbf{X}_0\hat{\mathbf{G}}\mathbf{X}_0^T\right)\hat{\mathbf{H}}\left\|\mathbf{M}_{\hat{H}}\left(\mathbf{I}_n - \mathbf{M}_0\right)\mathbf{Y}\right\|^2 \quad (7)$$

The score statistic test for overlap has a χ^2 distribution with a g =1 degree of freedom, which tests α in the Draper and Guttman model. The statistic in (7) depends on $\hat{\boldsymbol{\beta}}$ only through $\hat{\mathbf{H}}$ (and hence $\hat{\mathbf{G}}$). The estimates for $\boldsymbol{\beta}$ and σ^2 in models (1) and (4) can be substituted in (7), verifying that the same value of the statistic is obtained. Such a result is not surprising given that models (1) and (4) have the same estimation space C($\mathbf{X}$). Darghan (2010) showed the invariance property of S_{ovp} using $(\mathbf{X}^T\mathbf{X} + \mathbf{L}^T\mathbf{L})^{-1}$ or $(\mathbf{X}^T\mathbf{X})^-$.

Discussion

A statistical test based on Rao's score test has been developed for the overlapping effect in a re-parameterization of the Draper and Guttman model (Draper and Guttman, 1980). Although the phenomenon of overlap is not new in agricultural sciences, evaluation by means of a model is relatively new. Few authors have modeled this phenomenon. In 1999, Shukla and Subrahmanyam (1999) proposed an exact test and confidence intervals to assess the overlap coefficient using the model of Draper and Gutman, but using the Likelihood ratio rest. Despite the side conditions imposed in the original model to resolve the no-full-rank restriction in the design matrix and obtain unique and unbiased estimated parameters in (2), the error vector introduced the singular normal distribution; so, in this case, it is not defined as a log-likelihood function and, hence, it was impossible to obtain the estimators by maximum likelihood as it was not possible to apply the methodology of Rao to build the hypothesis testing. So far, we can say that the original model and the extended model by side conditions presented the respective restrictions of not being able to obtain the inverse of the expected Fisher information matrix or singular normal distribution. Thus, the test was developed by re-parameterization of the original model. The natural maximum likelihood estimators were obtained for the parameters γ and σ^2 under H_0 and, with these same estimators, the score test for overlapping was evaluated. The dependence of the statistical test on the parameters allows for replacing the two sets of estimators obtained by least square estimation: one for β and another for σ^2 in each studied model. The statistical test was invariant, which was expected because models (1) and (4) were equivalent and model (2) provided an estimate for β, belonging to the same estimation space of models 1 and 4. The statistical test for overlapping should be of wide application in agronomical research and has extensive practical value because the application of a variance analysis on a data set that was collected in the field has experienced the effect of overlap, generating spurious results as the effect of the treatments is confused since the same experimental unit could be receiving more than one treatment at a time. This test is easy to adopt as long as the layout of the used experimental design involves a model similar to that of Draper and Guttman whenever the nearest neighbor is considered a possible source of overlapping. Although the used example involved a test of overlapping by means of only the effects included in the two-way model, it can also be tested for any number of factors and their interactions as well as for a subgroup of effects and not only the use of the nearest neighbor by just modifying the design matrix and the matrix of the weights in the re-parameterized model. Statistics have been tested in various agricultural applications (Darghan *et al*., 2012) as well as in the area of education (Darghan *et al*., 2014) not only in design models, but also in the context of response surface modeling (Darghan *et al*., 2011), obtaining similar results as observed in the field (Darghan, 2010).

Conclusions and recommendations

The developed overlap test can be used in the field of agronomy where it is increasingly suspected that applied treatments can move from one experimental unit to that of the nearest neighbor and that the presence of overlap may invalidate the comparison of treatments when using the analysis of variance associated with a linear model, in this case, a two-way classification model. The asymptotic nature of the test requires a large amount of experimental units for it to be valid. The results obtained in applications using linear models have been similar to those observed in field results. Once the overlap coefficient has been estimated, the analysis of variance can be corrected by the overlap effect. Monte Carlo simulation studies as well as agricultural applications in the field of information technology and communication will complement the properties of the test, which can be extended to more complex experimental designs using neighboring patterns that cover all of the experimental units that are being studied.

Literature cited

Bera, A.K. and Y. Bilias. 2001. Rao's score, Neyman's C(α) and Silvey's LM tests: an essay on historical developments and some new results. J. Stat. Plan. Infer. 97, 9-44. Doi: 10.1016/S0378-3758(00)00343-8

Bera, A.K. and C.R. Mckenzie. 1986. Alternative forms and properties of thescore test. J. Appl. Stat. 13, 13-25. Doi: 10.1080/02664768600000002

Bhalli, M.A., A.D. Day, H. Tucker, R.K. Thompson, and G.D. Massey. 1964. Endborder effects in irrigated barley yield trials. J. Appl. Stat. 56, 346-348. doi:10.2134/agronj1964.00021962005600030028x

Chandra, T.K. and S.N. Joshi. 1985. Comparison of the likelihood ratio, Rao's and Wald's tests and a conjecture of C.R. Rao. Shankhyã: the Indian J. Stat. Ser. A 45, 226-246.

Christensen, R. 2011. Plane answers to complex questions. the theory of linear models. 4th ed. Springer-Verlag, New York, NY.

Cox, D.R. and D.V. Hinkley. 2000. Theoretical statistics. Chapman and Hall; CRC, London.

Darghan, A.E. 2010. Test score de Rao para el modelo de solapamiento de Draper y Guttman. PhD thesis. Faculty of Sciences, Universidad de los Andes, Merida, Venezuela.

Darghan, A.E., P. Sinha, and A. Goitia. 2011. Score test para el coeficiente de solapamiento en modelos de superficies de respuesta de primer y segundo orden. Dyna 165, 234-245.

Darghan C., A.E. P. Sinha S., and A. Goitia A. 2012. Una aplicación del test de solapamiento en un modelo de clasificación de dos vías reparametrizado. Rev. F. Agron. (LUZ) 29, 124-137.

Darghan, A.E., C. Monroy, and G. Montañez. 2014. Evaluación del solapamiento de información en las investigaciones que comparan modalidades de enseñanza asociados a las tecnologías de información y comunicación. Bistua: Rev. Fac. Cienc. Básicas 12, 3-15.

Draper, N.R. and I. Guttman. 1980. Incorporating overlap effects from neighbouring units into response surface models. Appl. Statist. 29, 128-134.

Graybill, F.A. 1976. Theory and application of the linear model. Duxbury Press, North Scituate, MA.

Hide, G.A. and P.J. Read. 1990. Effect of neighbouring plants on the yield of potatoes from seed tubers affected with gangrene (*Phoma foveata*) or from plants affected with stem canker (*Rhizoctonia solani*). Ann. Appl. Biol. 116, 233-243. Doi: 10.1111/j.1744-7348.1990.tb06603.x

Pearce, S.C. 1957. Experimenting with organisms as blocks. Biometrika 44, 141-149.

Rao, C.R. 1948. Large sample tests of statistical hypothesis concerning several parameters with applications to problems of estimation. Math. Proc. Camb. Phil. Soc. 44, 50-57. Doi: 10.1017/S0305004100023987

Rao, C.R. 1973. Linear statistical inference and its applications. 2nd ed. Wiley, New York, NY. Doi: 10.1002/9780470316436

Rao, C.R. and S.J. Poty. 1946. On locally most powerful tests when alternative are one sided. Shankhya 7, 439.

Shukla, G.K. and G.S.V. Subrahmanyan. 1999. A note on a test and confidence interval for competition and overlap effects. Biometrics 55, 273-276. Doi: 10.1111/j.0006-341X.1999.00273.x

Wald, A. 1941. Some examples of asymptotically most powerful tests. Ann. Math. Statist. 12, 396-408. Doi: 10.1214/aoms/1177731678