

# DISEÑOS EXPERIMENTALES CON TESTIGOS REPETIDOS

## EXPERIMENTAL DESIGNS WITH REPEATED CONTROLS

Juan A. BURGUEÑO-Ferreira<sup>1</sup>, Ángel Martínez-Garza<sup>†</sup>, José Crossa<sup>2</sup> y Ángel Mastache-Lagunas<sup>3</sup>

<sup>1</sup>Estadística. Campus Montecillo. Colegio de Postgraduados. 56230. Montecillo, Estado de México. (juanbu@yahoo.com) (angel@colpos.mx). <sup>2</sup>CIMMYT. Apartado Postal 6-641. 06600. México, D. F. (jcrossa@cgiar.org). <sup>3</sup>Universidad de Guerrero. Javier Méndez Aponte #1, Colonia Servidor Agrario. 39070. Chilpancingo, Guerrero. (gamma32@starmedia.com)

### RESUMEN

En algunas fases del mejoramiento de ciertos cultivos se tienen conjuntos muy numerosos de nuevas variedades, de las cuales los mejoradores tienen que seleccionar las más prometedoras. La selección se realiza por comparación directa de las nuevas variedades que se ensayan en una sola unidad experimental con variedades testigo que se intercalan sistemáticamente entre ellas, repitiéndose sobre un número grande de parcelas. A menudo no se entiende que, bajo reglas sencillas de diseño, podrían ser diseños experimentales susceptibles de un análisis estadístico exacto y correcto. En este trabajo se discute este tema con precisión, estableciendo reglas de diseño y la técnica de análisis estadístico adecuada para los diseños experimentales completamente aleatorizado o en bloques completos al azar con bloques de igual tamaño.

**Palabras clave:** Diseño experimental, diseño no repetido, ensayos varietales.

### INTRODUCCIÓN

En las investigaciones sobre mejoramiento de plantas, en las primeras etapas del proceso de mejora genética, el elevado número de genotipos, la escasez de semilla y la baja disponibilidad de recursos, no permiten plantear un diseño experimental que cumpla con el requisito de tener repeticiones de todos los tratamientos. Una práctica frecuente es intercalar testigos entre las líneas a evaluar y, por comparación directa con éstos, los investigadores seleccionan tan rigurosamente como les es posible sus variedades prometedoras. La selección entonces es subjetiva; pero es posible, siguiendo ciertas reglas de diseño, obtener un diseño experimental susceptible de una interpretación estadística válida.

Se han propuesto distintas soluciones de diseño para examinar situaciones donde algunos tratamientos no se repiten, en tanto que los repetidos se emplean para obtener una estimación del error experimental. Se han realizado numerosos trabajos cuando los factores a analizar

### ABSTRACT

In some stages of the breeding of crops, breeders must select the most promising lines from a very large number of sets of new varieties. Selection is conducted via direct comparison of new varieties that are tested within a single experimental unit, with control varieties systematically intercropped among them and replicated over a large number of plots. Researchers often do not realize that, by following some simple design rules, they might be susceptible of a precise and accurate statistical analysis. This paper discusses the subject with precision, and establishes rules of design and a statistical analysis technique appropriate for completely randomized experimental designs or randomized complete blocks with blocks of the same size.

**Key words:** Experimental design, unreplicated design, variety trials

### INTRODUCTION

In plant breeding trials, during the early stages of the improvement process, it is not possible to use an experimental design that satisfies the requirement of replicating all the treatments because of the large number of genotypes involved, the small amount of seed and the low availability of resources. A common practice is to intercrop check varieties among test lines. By comparing the materials directly, they are able to select the most promising lines as rigorously as is possible. The selection is, then, subjective; nonetheless, by following certain design rules it is possible to develop an experimental design susceptible to a valid statistical interpretation.

Different design solutions have been proposed to study situations in which some of the treatments are unreplicated, and the replicated ones are used to estimate the experimental error. Many studies have focused on analyzing quantitative factors. Box and Wilson (1951) included only replications of the central point in the composite central designs used in industrial research. However, there are few studies focusing on qualitative problems. Federer (1961) introduced augmented designs and developed a methodology for analyzing one replication per block of controls. Other designs, and their

---

Recibido: Diciembre, 2004. Aprobado: Agosto, 2005.

Publicado como ENSAYO en *Agrociencia* 39: 679-691. 2005.

son cuantitativos. Box y Wilson (1951) consideran exclusivamente repeticiones del punto central, en los diseños compuestos centrales usados en investigaciones industriales. Son pocos los trabajos enfocados a problemas de carácter cualitativo. Federer (1961) introduce los diseños aumentados, y genera la metodología de análisis para una repetición por bloque de los testigos. Otros diseños, y sus análisis, han sido propuestos por Lin y Poushinsky (1983). Diseños de tipo sistemático se han sugerido para este propósito (Besag y Kempton, 1986; Méndez, 1976).

Hasta hoy no parece existir un examen sistemático de los métodos de diseño y análisis estadístico de tales arreglos experimentales. En Estados Unidos y México es común el uso de las técnicas del vecino cercano (Papadakis, 1984), mientras que en Australia y Gran Bretaña se usa la modelación de los errores. Federer *et al.* (2000) han promovido el uso de modelos polinomiales para modelar la variabilidad espacial. Sin embargo, en la mayoría de los casos, no se considera la relación que debe haber entre el diseño experimental y la metodología de análisis.

El propósito de este trabajo fue considerar una extensión de los diseños aumentados de Federer (1961), que incluyen testigos repetidos dentro de los bloques. Se examina el caso más sencillo (diseño completamente al azar), y luego la situación más usual y balanceada de este tipo de investigaciones en bloques, con el propósito de establecer reglas claras de diseño y el método de análisis estadístico.

## DISEÑO Y ANÁLISIS

### Diseño completamente al azar

El caso más balanceado, y posiblemente el más frecuente, ocurre cuando cada una de  $q$  variedades testigo, repetidas  $r$  veces cada una, se intercala al azar sobre  $v$  nuevas variedades. El proceso de aleatorización ocurriría en dos pasos: 1: las  $r$  repeticiones de cada uno de los  $q$  testigos se asignan al azar a las  $rq+v$  parcelas disponibles; 2: las  $v$  variedades nuevas se asignan al azar a las  $v$  unidades experimentales restantes.

El modelo lineal para interpretar los resultados del experimento es:

$$y_{k^0j} = \mu + \tau_{k^0} + e_{k^0j} \quad k=1,2,\dots,q, \quad j=1,2,\dots,r \quad (1a)$$

para los tratamientos testigo, y

$$y_{ij} = \mu + \tau_i + e_{ij} \quad i=1,2,\dots,v, \quad j=1, \quad (1b)$$

analyses, have been proposed by Lin and Poushinsky (1983). Systematic designs have also been suggested for this purpose (Besag and Kempton, 1986; Méndez, 1976).

To date, there does not seem to be a systematic study of design and statistical analysis methods for such experimental arrangements. In the USA and México the use of the nearest neighbor techniques is common (Papadakis, 1984), while in Australia and Great Britain error modeling is most commonly used. Federer *et al.* (2000) have promoted the use of polynomial models to model spatial variability. However, in most cases, the relationship that should exist between the experimental design and the analysis methodology is not taken into account.

The purpose of this paper was to consider an extension of Federer's augmented designs (1961) that would include replicated controls within blocks. We analyze the simplest design (a completely randomized design), and then the most common and balanced example of a randomized complete block design, to establish clear rules of design and the method for statistical analysis.

## DESIGN AND ANALYSIS

### Completely randomized design

The most balanced and perhaps most common case, is when each one of  $q$  control varieties, replicated  $r$  times each, is randomly intercropped among  $v$  new varieties. The randomization process is conducted in two steps: 1:  $r$  replications of each of the  $q$  controls are randomly assigned to the  $rq+v$  plots available; 2:  $v$  new varieties are randomly assigned to the  $v$  remaining experimental units.

The linear model for interpreting the results of the experiment is:

$$y_{k^0j} = \mu + \tau_{k^0} + e_{k^0j} \quad k=1,2,\dots,q, \quad j=1,2,\dots,r \quad (1a)$$

for the control treatments, and

$$y_{ij} = \mu + \tau_i + e_{ij} \quad i=1,2,\dots,v, \quad j=1, \quad (1b)$$

for the new treatments,  $y_{ij} (y_{k^0j})$  being the value of the trait being studied,  $\mu$  an effect common to all observations,  $\tau$  the treatment effects, and  $e_{ij} (e_{k^0j})$ , errors with normal distribution and with the usual properties (mean 0, unknown variance  $\sigma^2$  and uncorrelated).

Applying least squares to model (1) leads, with the usual statistical notation, to the system of normal equations

para los nuevos tratamientos, siendo  $y_{ij}(y_{k^0_j})$  el valor de la característica en estudio,  $\mu$  un efecto común a todas las observaciones,  $\tau$  los efectos de tratamientos, y los  $e_{ij}(e_{k^0_j})$ , errores con distribución normal y con las propiedades usuales (media 0, varianza desconocida  $\sigma^2$  y no correlacionados).

La aplicación de mínimos cuadrados al modelo (1) conduce, con la notación estadística usual, al sistema de ecuaciones normales

$$\mu : (rq + v)\hat{\mu} + \sum_{i=1}^v \hat{\tau}_i + r \sum_{k=1}^q \hat{\tau}_{k^0} = Y_{..} = G,$$

$\tau_i : \hat{\mu} + \hat{\tau}_i = y_{i1} = T_i \quad i=1,2,\dots,v$ , para un tratamiento nuevo, o

$\tau_{k^0} : r\hat{\mu} + r\hat{\tau}_{k^0} = Y_{k^0} = T_{k^0}, \quad k=1,2,\dots,q$ , para un tratamiento testigo, las cuales conducen al análisis de varianza del Cuadro 1. Nótese en este Cuadro que

$$SCE = \sum_{k=1}^q \sum_{j=1}^r y_{k^0_j}^2 - \frac{\sum_{k=1}^q T_{k^0}^2}{r}, \text{ es decir, que el error se}$$

estima exclusivamente a partir de los tratamientos testigo.

Por otra parte se tiene (Searle, 1971)

$$MELI(\tau_{k^0} - \tau_{l^0}) = \bar{y}_{k^0} - \bar{y}_{l^0},$$

$$VAR(\bar{y}_{k^0} - \bar{y}_{l^0}) = \frac{2\sigma^2}{r}, \quad 1 \leq k, l \leq q$$

$$MELI(\tau_{k^0} - \tau_i) = \bar{y}_{k^0} - \bar{y}_i, \quad VAR(\bar{y}_{k^0} - \bar{y}_i)$$

$$= \left(1 + \frac{1}{r}\right)\sigma^2, \quad k = 1, 2, \dots, q, i = 1, 2, \dots, v,$$

$$MELI(\tau_i - \tau_m) = \bar{y}_i - \bar{y}_m; \quad VAR(\bar{y}_i - \bar{y}_m) = 2\sigma^2,$$

$$1 \leq i, m \leq v,$$

donde  $\bar{y}_i = T_i / 1 \quad i=1,2,\dots,v \quad \bar{y}_{k^0} = T_{k^0} / r \quad k=1,2,\dots,q$

### Diseño en bloques

Considérese ahora la situación en que el material experimental presenta un gradiente de variabilidad, y se requiere controlarlo con bloques de tamaño razonable,

$$\mu : (rq + v)\hat{\mu} + \sum_{i=1}^v \hat{\tau}_i + r \sum_{k=1}^q \hat{\tau}_{k^0} = Y_{..} = G,$$

$\tau_i : \hat{\mu} + \hat{\tau}_i = y_{i1} = T_i \quad i=1,2,\dots,v$ , for a new treatment, or

$\tau_{k^0} : r\hat{\mu} + r\hat{\tau}_{k^0} = Y_{k^0} = T_{k^0}, \quad k=1,2,\dots,q$ , for a control treatment, which in turn lead to the analysis of variance in Table 1. Note that in Table 1

$$SCE = \sum_{k=1}^q \sum_{j=1}^r y_{k^0_j}^2 - \frac{\sum_{k=1}^q T_{k^0}^2}{r}, \text{ that is, the error is}$$

estimated based on the control treatments only.

We also have (Searle, 1971):

$$BLUE(\tau_{k^0} - \tau_{l^0}) = \bar{y}_{k^0} - \bar{y}_{l^0},$$

$$VAR(\bar{y}_{k^0} - \bar{y}_{l^0}) = \frac{2\sigma^2}{r}, \quad 1 \leq k, l \leq q$$

$$BLUE(\tau_{k^0} - \tau_i) = \bar{y}_{k^0} - \bar{y}_i, \quad VAR(\bar{y}_{k^0} - \bar{y}_i)$$

$$= \left(1 + \frac{1}{r}\right)\sigma^2, \quad k = 1, 2, \dots, q, i = 1, 2, \dots, v,$$

$$BLUE(\tau_i - \tau_m) = \bar{y}_i - \bar{y}_m; \quad VAR(\bar{y}_i - \bar{y}_m) = 2\sigma^2,$$

$$1 \leq i, m \leq v,$$

where  $\bar{y}_i = T_i / 1 \quad i=1,2,\dots,v \quad \bar{y}_{k^0} = T_{k^0} / r \quad k=1,2,\dots,q$

### Block design

Let us now consider an example where the experimental material present a variability gradient that needs to be controlled by blocks of reasonable size, but constant for all blocks, to achieve a design more or less balanced. In the most balanced design, all the blocks will have, say,  $k$  experimental units, with each block containing  $q$  controls replicated  $c$  times within a block. If the experiment has  $v$  new treatments replicated once, and a total of  $b$  blocks, it obviously consists of  $bqc+v$  experimental units, and each block contains  $k-cq=v/b$  new treatments. Again, the randomization process would be conducted in two steps; 1:  $c$  repetitions of each of the  $q$  control treatments would be assigned completely at random, independently within each block in the  $k$  available plots; 2: in the  $b(k-cq)=v$  experimental units of the  $b$

**Cuadro 1. Análisis de varianza de un experimento con testigos repetidos en un diseño completamente aleatorio.**  
**Table 1. Analysis of variance for an experiment with repeated controls, in a completely randomized design.**

Fuentes de variación	Grados de libertad	Sumas de cuadrados	Cuadrados medios	F calculada
Tratamientos	$q+v-1$	$SCT = \sum_{i=1}^v \frac{T_i^2}{1} + \frac{\sum_{k=1}^q T_k^2}{r} - \frac{G^2}{N}$	$CMT = \frac{SCT}{q+v-1}$	$\frac{CMT}{s^2}$
Error	$q(r-1)$	SCE (por diferencia)	$CME = \frac{SCE}{q(r-1)} = s^2$	
Total	$N-1$	$\sum_{i=1}^v y_{i1}^2 + \sum_{k=1}^q \sum_{j=1}^r y_{k^0j}^2 - \frac{G^2}{N}$		

En este Cuadro  $N = rq + v$  y  $G = Y_{..}$  = Gran total.

pero constante para todos los bloques, para lograr un diseño más o menos balanceado. El caso más balanceado ocurrirá cuando todos los bloques tengan, por ejemplo,  $k$  unidades experimentales, incluyendo cada bloque  $q$  testigos repetidos  $c$  veces dentro de un bloque. Si el experimento tiene  $v$  nuevos tratamientos repetidos una vez, y  $b$  bloques en total, es fácil ver que éste consta de  $bqc+v$  unidades experimentales, y que cada bloque incluye  $k-cq = v/b$  nuevos tratamientos. Nuevamente, el proceso de aleatorización se haría en dos pasos; 1: las  $c$  repeticiones de cada uno de los  $q$  tratamientos testigo se asignarían completamente al azar, en forma independiente dentro de cada bloque en las  $k$  parcelas disponibles; 2: en las  $b(k-cq) = v$  unidades experimentales de los  $b$  bloques que no recibieron tratamiento, se asignarían al azar los  $v$  nuevos tratamientos.

Otra situación experimental que puede considerarse es cuando el número de nuevas variedades no es un múltiplo del número de bloques. En este caso los bloques deberán ser de distinto tamaño para mantener el balance de los testigos. La solución presentada abarca ambas situaciones experimentales.

El modelo lineal apropiado para interpretar los resultados del experimento es

$$y_{k^0jl} = \mu + \beta_1 + \tau_{k^0} + e_{k^0jl} \quad \begin{matrix} k=1,2,\dots,q, \\ j=1,2,\dots,c \\ l=1,2,\dots,b \end{matrix}$$

para los tratamientos testigo, y:

$$y_{ijl} = \mu + \beta_1 + \tau_i + e_{ijl} \quad i=1,2,\dots,v; \quad j=1; \quad l=1,2,\dots,b$$

para los nuevos tratamientos. Por consiguiente, el sistema de ecuaciones normales reducidas para tratamientos, eliminando bloques, cuya forma general Martínez (1988), se simplifica como sigue:

blocks that received no treatment, the  $v$  new treatments would be assigned at random.

Another experimental situation worth considering is when the number of new varieties is not a multiple of the number of blocks. In this case the blocks must be of different sizes to keep the balance of the controls. The solution here presented includes both experimental situations.

The appropriate linear model to interpret trial results is

$$y_{k^0jl} = \mu + \beta_1 + \tau_{k^0} + e_{k^0jl} \quad \begin{matrix} k=1,2,\dots,q, \\ j=1,2,\dots,c \\ l=1,2,\dots,b \end{matrix}$$

for the control treatments, and:

$$y_{ijl} = \mu + \beta_1 + \tau_i + e_{ijl} \quad i=1,2,\dots,v; \quad j=1; \quad l=1,2,\dots,b$$

for the new treatments. Therefore, the reduced system of normal equations for treatments eliminating blocks, whose general form is described in Martínez (1988), is simplified as follows:

**Equations for a control treatment**

$$\left( bc - \frac{bc}{k} \right) \bar{\tau}_{k^0} - \frac{c}{k} \sum_{i=1}^v \bar{\tau}_i - \frac{bc^2}{k} \sum_{\substack{i=1 \\ i \neq m^0}}^q \bar{\tau}_{i^0} = Q_{k^0}, \quad (2)$$

$$k^0=1,2,\dots,q,$$

where  $Q_{k^0} = Y_{k^0..} - \sum_{l=1}^b \frac{n_{kl} Y_{..l}}{k} = T_{k^0} - \frac{1}{k} \sum_{l=1}^b n_{kl} B_l$ , being  $Y_{k^0..} = T_{k^0}$  the total of the  $k^{\text{th}}$  control treatment,  $B_l$  the total of the  $l^{\text{th}}$  block,  $\sum_{l=1}^b B_l = Y_{..} = G$  the grand total;  $n_{kl}$

**Ecuaciones para un tratamiento testigo**

$$\left( bc - \frac{bc}{k} \right) \hat{\tau}_{k^0} - \frac{c}{k} \sum_{i=1}^v \hat{\tau}_i - \frac{bc^2}{k} \sum_{\substack{i=1 \\ i \neq m^0}}^q \hat{\tau}_{i^0} = Q_{k^0},$$

$$k^0=1,2,\dots,q, \tag{2}$$

donde  $Q_{k^0} = Y_{k^0} \dots - \frac{1}{k} \sum_{l=1}^b \frac{n_{kl} Y_{..l}}{k} = T_{k^0} - \frac{1}{k} \sum_{l=1}^b n_{kl} B_l$ , siendo  $Y_{k^0} \dots = T_{k^0}$  el total del  $k$ -ésimo tratamiento testigo,  $B_l$  el total del  $l$ -ésimo bloque,  $\sum_{l=1}^b B_l = Y_{\dots} = G$  el gran total;  $n_{kl}$  una variable indicadora que cuenta el número de veces que el  $k$ -ésimo tratamiento ocurre en el bloque  $l$ . Como las  $n_{kl}$  son todas iguales a  $c$ , finalmente  $Q_{k^0} = T_{k^0} - cG/k$ . Sumando y restando  $(bc^2/k)\hat{\tau}_{k^0}$  al lado izquierdo de la ecuación (2), después de hacer algunas simplificaciones, se obtiene

$$\left( bc - \frac{bc^2}{k} + \frac{bc^2}{k} \right) \hat{\tau}_{k^0} - \frac{c}{k} \left( \sum_{i=1}^v \hat{\tau}_i + bc \sum_{i=1}^q \hat{\tau}_{i^0} \right) = Q_{k^0},$$

y de aquí, introduciendo la restricción  $\sum_{i=1}^v \hat{\tau}_i + bc \sum_{i=1}^q \hat{\tau}_{i^0} = 0$

$$\hat{\tau}_{k^0} = \frac{1}{bc} Q_{k^0}, \quad k^0=1,2,\dots,q \tag{3}$$

**Ecuaciones para un tratamiento nuevo**

$$\hat{\tau}_i - \frac{1}{k} \sum_{m=1}^v \delta_{ij} \hat{\tau}_m - \frac{1}{k} \sum_{k=1}^q \hat{\tau}_{k^0} = Q_i, \quad i=1,2,\dots,v$$

donde  $Q_i = y_{i1} - \frac{1}{k} \sum_{l=1}^b n_{il} Y_{..l}$ , y  $\delta_{ij} = \sum_{l=1}^b n_{il} n_{ml} = 1$  si los nuevos tratamientos  $i$ -ésimo y  $m$ -ésimo ocurren juntos en el mismo bloque, e igual a 0 de otra manera. Obsérvese que  $\sum_{m=1}^v \delta_{ij} \hat{\tau}_m$ , es la suma de los  $k-qc-l$  efectos de los nuevos tratamientos que ocurren en el mismo bloque que el  $i$ -ésimo tratamiento nuevo, más  $\hat{\tau}_i$ . Asimismo,  $\sum_{l=1}^b n_{il} Y_{..l}$  es el total del bloque donde ocurre el  $i$ -ésimo tratamiento nuevo. De aquí, si se considera por ejemplo

an indicator variable that counts the number of times the  $k^{\text{th}}$  treatment occurs in the  $l$  block. Since the  $n_{kl}$  variables are all equal to  $c$ , finally  $Q_{k^0} = T_{k^0} - cG/k$ . By adding and subtracting  $(bc^2/k)\hat{\tau}_{k^0}$  to the left-hand side of equation (2), and then doing some simplifications, one obtains

$$\left( bc - \frac{bc^2}{k} + \frac{bc^2}{k} \right) \hat{\tau}_{k^0} - \frac{c}{k} \left( \sum_{i=1}^v \hat{\tau}_i + bc \sum_{i=1}^q \hat{\tau}_{i^0} \right) = Q_{k^0},$$

from which, by introducing the restriction  $\sum_{i=1}^v \hat{\tau}_i + bc \sum_{i=1}^q \hat{\tau}_{i^0} = 0$

$$\hat{\tau}_{k^0} = \frac{1}{bc} Q_{k^0}, \quad k^0=1,2,\dots,q \tag{3}$$

**Equations for a new treatment**

$$\hat{\tau}_i - \frac{1}{k} \sum_{m=1}^v \delta_{ij} \hat{\tau}_m - \frac{1}{k} \sum_{k=1}^q \hat{\tau}_{k^0} = Q_i, \quad i=1,2,\dots,v$$

where  $Q_i = y_{i1} - \frac{1}{k} \sum_{l=1}^b n_{il} Y_{..l}$ , and  $\delta_{ij} = \sum_{l=1}^b n_{il} n_{ml} = 1$  if the  $i^{\text{th}}$  and  $m^{\text{th}}$  new treatments occur together in the same block, but otherwise are equal to 0. It should be noted that  $\sum_{m=1}^v \delta_{ij} \hat{\tau}_m$ , is the sum of  $k-qc-l$  effects of the new treatments occurring in the same block as the  $i^{\text{th}}$  new treatment, plus  $\hat{\tau}_i$ . Similarly,  $\sum_{l=1}^b n_{il} Y_{..l}$  is the total for the block where the  $i^{\text{th}}$  new treatment occurs. Thus, if we consider, for example, the first block of the trial, the normal equations reduced for the new treatments may be written using the matrix notation as follows

$$\begin{bmatrix} 1 - \frac{1}{k} & -\frac{1}{k} & \dots & -\frac{1}{k} \\ -\frac{1}{k} & 1 - \frac{1}{k} & \dots & -\frac{1}{k} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{k} & -\frac{1}{k} & \dots & 1 - \frac{1}{k} \end{bmatrix} \begin{bmatrix} \hat{\tau}_1 \\ \hat{\tau}_2 \\ \vdots \\ \hat{\tau}_{k-qc} \end{bmatrix} = \begin{bmatrix} Q_1 + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \\ Q_2 + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \\ \vdots \\ Q_{k-qc} + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \end{bmatrix} \tag{4}$$

el primer bloque del experimento, las ecuaciones normales reducidas para los tratamientos nuevos pueden escribirse en notación matricial, como sigue

$$\begin{bmatrix} 1-\frac{1}{k} & -\frac{1}{k} & \dots & -\frac{1}{k} \\ -\frac{1}{k} & 1-\frac{1}{k} & \dots & -\frac{1}{k} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{k} & -\frac{1}{k} & \dots & 1-\frac{1}{k} \end{bmatrix} \begin{bmatrix} \hat{\tau}_1 \\ \hat{\tau}_2 \\ \vdots \\ \hat{\tau}_{k-qc} \end{bmatrix} = \begin{bmatrix} Q_1 + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \\ Q_2 + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \\ \vdots \\ Q_{k-qc} + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \end{bmatrix} \quad (4)$$

La matriz en el lado izquierdo de la expresión (4), de dimensiones  $(k-qc) \times (k-qc)$ , es de la forma  $aI+bJ$ , donde  $I$  es una matriz idéntica de dimensiones  $(k-qc) \times (k-qc)$ ,  $J$  es una matriz de unos de las mismas dimensiones,  $a=1$  y  $b=-1/k$ ; y tiene una inversa de estructura similar:  $dI+eJ$ , con  $d=1$  y  $e=1/(qc)$ . Por tanto, la solución a la ecuación matricial (4), es

$$\begin{bmatrix} \hat{\tau}_1 \\ \hat{\tau}_2 \\ \vdots \\ \hat{\tau}_{k-qc} \end{bmatrix} = \begin{bmatrix} 1+\frac{1}{qc} & \frac{1}{qc} & \dots & \frac{1}{qc} \\ \frac{1}{qc} & 1+\frac{1}{qc} & \dots & \frac{1}{qc} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{qc} & \frac{1}{qc} & \dots & 1+\frac{1}{qc} \end{bmatrix}$$

$$\begin{bmatrix} Q_1 + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \\ Q_2 + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \\ \vdots \\ Q_{k-qc} + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \end{bmatrix}$$

o bien, después de realizar las operaciones indicadas

$$\hat{\tau}_i = Q_i + \frac{1}{qbc} \sum_{k=1}^q Q_{k^0} + \frac{1}{qc} \sum_{m=1}^v \left( \sum_{l=1}^b n_{il} n_{ml} \right) Q_m, \quad i=1,2,\dots,k-qc \quad (5)$$

The matrix on the left-hand side of equation (4), with  $(k-qc) \times (k-qc)$  dimensions, has the form  $aI+bJ$ , where  $I$  is an identity matrix with  $(k-qc) \times (k-qc)$  dimensions,  $J$  is a matrix of 1s of the same dimensions,  $a=1$  and  $b=-1/k$ , and has an inverse of similar structure:  $dI+eJ$ , with  $d=1$  and  $e=1/(qc)$ . Therefore, the solution to matrix equation (4) is

$$\begin{bmatrix} \hat{\tau}_1 \\ \hat{\tau}_2 \\ \vdots \\ \hat{\tau}_{k-qc} \end{bmatrix} = \begin{bmatrix} 1+\frac{1}{qc} & \frac{1}{qc} & \dots & \frac{1}{qc} \\ \frac{1}{qc} & 1+\frac{1}{qc} & \dots & \frac{1}{qc} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{qc} & \frac{1}{qc} & \dots & 1+\frac{1}{qc} \end{bmatrix}$$

$$\begin{bmatrix} Q_1 + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \\ Q_2 + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \\ \vdots \\ Q_{k-qc} + \frac{1}{bk} \sum_{x=1}^q Q_{k^0} \end{bmatrix}$$

or, after performing the operations indicated above

$$\hat{\tau}_i = Q_i + \frac{1}{qbc} \sum_{k=1}^q Q_{k^0} + \frac{1}{qc} \sum_{m=1}^v \left( \sum_{l=1}^b n_{il} n_{ml} \right) Q_m, \quad i=1,2,\dots,k-qc \quad (5)$$

$\sum_{m=1}^v \left( \sum_{l=1}^b n_{il} n_{jm} \right) Q_m$  is the sum of the  $Q_m$  corresponding to those new treatments occurring in the same block as the  $i^{th}$  new treatment, plus  $Q_i \cdot \frac{1}{qbc} \sum_{k=1}^q Q_{k^0} = \frac{1}{q} \sum_{k=1}^q \hat{\tau}_{k^0}$  is the average of the control treatment effects.

Similar solutions for the remaining  $\hat{\tau}_i$  are obviously obtained in the other blocks.

If in equation (5) we substitute the specific forms for  $Q_m$ , we obtain

$\sum_{m=1}^v \left( \sum_{l=1}^b n_{il} n_{jm} \right) Q_m$  es la suma de las  $Q_m$  correspondientes

a aquellos tratamientos nuevos que ocurren en el mismo bloque que el  $i$ -ésimo tratamiento nuevo más  $Q_i$ .

$\frac{1}{qbc} \sum_{k=1}^q Q_{k^0} = \frac{1}{q} \sum_{k=1}^q \hat{\tau}_{k^0}$  es el promedio de los efectos de

los tratamiento testigo.

Obviamente, soluciones similares para las  $\hat{\tau}_i$  restantes se obtienen en los demás bloques.

Si se sustituyen en (5) las formas expresas de las  $Q_m$  se tiene

$$\hat{\tau}_i = y_{i1} - \frac{1}{k} \sum_{l=1}^b n_{il} Y_{..l} + \frac{1}{q} \sum_{k=1}^q \hat{\tau}_{k^0} + \frac{1}{qc} \sum_{m=1}^v \sum_{l=1}^b (n_{il} n_{ml}) \left( y_m - \frac{1}{k} \sum_{l=1}^b n_{il} Y_{..l} \right)$$

y de ahí

$$\hat{\tau}_i = y_{i1} - \left[ \frac{1}{qc} \sum_{l=1}^b n_{il} Y_{..l} - \frac{1}{q} \sum_{k=1}^q \hat{\tau}_{k^0} - \frac{1}{qc} \sum_{m=1}^v \left( \sum_{l=1}^b n_{il} n_{ml} \right) y_m \right]$$

donde el segundo término entre paréntesis es la media del bloque donde se encuentra el  $i$ -ésimo tratamiento nuevo considerando únicamente las parcelas con testigos, menos el efecto de éstos. Así, una expresión más sencilla para el efecto de un nuevo tratamiento es

$$\hat{\tau}_{i1l} = y_{i1l} - (\bar{Y}_{..l}^* - \bar{Y}_{...}^*) - \bar{Y}_{...}$$

es decir que la estimación del efecto de un nuevo tratamiento es el valor observado menos el efecto del bloque en el que se encuentra, estimado únicamente con los tratamientos testigo menos la media general.

Las observaciones de los nuevos tratamientos, que no están repetidos, no dan información sobre los efectos de bloque ni sobre el error experimental, por tanto, el efecto de bloque y el error experimental se estima con las observaciones de los testigos repetidos. Si se incluyera un mayor número de nuevas variedades, las estimaciones y varianzas de las ya incluidas serían las mismas.

Cuando se tienen bloques de distinto tamaño, con balance de los testigos, se obtienen las mismas soluciones, ya que tenemos un diseño de bloques completos con repeticiones para los testigos donde estimar los efectos de

$$\hat{\tau}_i = y_{i1} - \frac{1}{k} \sum_{l=1}^b n_{il} Y_{..l} + \frac{1}{q} \sum_{k=1}^q \hat{\tau}_{k^0}$$

$$+ \frac{1}{qc} \sum_{m=1}^v \sum_{l=1}^b (n_{il} n_{ml}) \left( y_m - \frac{1}{k} \sum_{l=1}^b n_{il} Y_{..l} \right)$$

from which

$$\hat{\tau}_i = y_{i1} - \left[ \frac{1}{qc} \sum_{l=1}^b n_{il} Y_{..l} - \frac{1}{q} \sum_{k=1}^q \hat{\tau}_{k^0} - \frac{1}{qc} \sum_{m=1}^v \left( \sum_{l=1}^b n_{il} n_{ml} \right) y_m \right]$$

where the second term within the parentheses is the mean of the block where the  $i$ <sup>th</sup> new treatment is found, taking into account only the plots with controls, minus the effect of those. Thus, a simpler equation for the effect of a new treatment is:

$$\hat{\tau}_{i1l} = y_{i1l} - (\bar{Y}_{..l}^* - \bar{Y}_{...}^*) - \bar{Y}_{...}$$

that is, the estimated effect of a new treatment is the observed value minus the effect of the block where it is found, estimated based only on the control treatments minus the general mean.

Observations on new, unreplicated treatments give no information on the block effects or experimental error; therefore, the block effects and the experimental error are estimated based on observations made on replicated controls. If a higher number of new varieties were to be included, the estimates and variances of the ones already included would be the same.

When there are blocks of different sizes, and balance on controls we arrive at the same solutions, since we have a complete block design with replicated control, based on which we can estimate the block and experimental error effects. It is clear that the blocks may be of different sizes, as long as they can be considered a homogeneous unit.

The analysis of variance in Table 2 was constructed based on the above results, where  $B_i$  are the block totals,

$G = Y_{...} = \sum_{l=1}^b B_l$  is the grand total, whereas  $\hat{\tau}_i$  and  $\hat{\tau}_{k^0}$  are given by equations (5) and (3), respectively. Similarly,

$$Q_i = y_{i1} - \frac{1}{k} \sum_{l=1}^b n_{il} B_l \quad i=1,2,\dots,v$$

and:

$$Q_{k^0} = T_{k^0} - \frac{cG}{k}, \quad k^0=1,2,\dots,q$$

bloque y el error experimental. Es claro que los bloques pueden ser distintos en tamaño, siempre que puedan considerarse como una unidad homogénea.

Con los resultados anteriores se construye el análisis de varianza del Cuadro 2, donde las  $B_l$  son totales de bloques,  $G = Y_{...} = \sum_{l=1}^b B_l$  es el gran total, en tanto que las  $\hat{\tau}_i$  y las  $\hat{\tau}_{k^0}$  están dadas, por las relaciones (5) y (3); asimismo

$$Q_i = y_{i1} - \frac{1}{k} \sum_{l=1}^b n_{il} B_l \quad i=1,2,\dots,v$$

y:

$$Q_{k^0} = T_{k^0} - \frac{cG}{k}, \quad k^0=1,2,\dots,q$$

Nótese que en la expresión para las  $Q_i, \sum_{i=1}^v n_{il} B_l$  es exactamente igual al total del bloque donde ocurre el nuevo tratamiento.

Ahora bien, como se tienen los estimadores

$$\hat{\tau}_{k^0} - \hat{\tau}_i = \frac{Q_{k^0}}{bc} - Q_i - \frac{1}{qbc} \sum_{m=1}^q Q_{m^0} - \frac{1}{qc} \sum_{s=1}^v \left( \sum_{l=1}^b n_{il} n_{sl} \right) Q_s$$

$$\hat{\tau}_{k^0} - \hat{\tau}_{m^0} = \frac{Q_{k^0}}{bc} - \frac{Q_{m^0}}{bc};$$

para dos nuevos tratamientos en el mismo bloque,

$$\hat{\tau}_i - \hat{\tau}_s = Q_i - Q_s;$$

Note that in the equation for  $Q_i, \sum_{i=1}^v n_{il} B_l$  is exactly the same as the total for the block where the new treatment occurs.

Since we have the estimators

$$\hat{\tau}_{k^0} - \hat{\tau}_i = \frac{Q_{k^0}}{bc} - Q_i - \frac{1}{qbc} \sum_{m=1}^q Q_{m^0} - \frac{1}{qc} \sum_{s=1}^v \left( \sum_{l=1}^b n_{il} n_{sl} \right) Q_s$$

$$\hat{\tau}_{k^0} - \hat{\tau}_{m^0} = \frac{Q_{k^0}}{bc} - \frac{Q_{m^0}}{bc};$$

for two new treatments in the same block,

$$\hat{\tau}_i - \hat{\tau}_s = Q_i - Q_s;$$

for two new treatments in different blocks,

$$\hat{\tau}_i - \hat{\tau}_s = Q_i + \frac{1}{qc} \sum_{m=1}^v \left( \sum_{l=1}^b n_{l=1} n_{ml} \right)$$

$$Q_m - Q_s - \frac{1}{qc} \sum_{m=1}^v \left( \sum_{l=1}^b n_{sl} n_{ml} \right) Q_m$$

The following results are also obtained

$$VAR(\hat{\tau}_{k^0} - \hat{\tau}_i) = \left( \frac{1}{bc} - \frac{1}{qbc} + \frac{1}{qc} + 1 \right) \sigma^2,$$

$$VAR(\hat{\tau}_{k^0} - \hat{\tau}_{m^0}) = \frac{2\sigma^2}{bc},$$

**Cuadro 2. Análisis de varianza de un experimento con testigos repetidos en un diseño en bloques.**

**Table 2. Analysis of variance for an experiment with repeated controls in a complete blocks design.**

Fuentes de variación	Grados de libertad	Sumas de cuadrados	Cuadrados medios	F calculada
Bloques (ignorando tratamientos)	$b-1$	$SC(BIT) = \frac{\sum_{l=1}^b B_l^2}{k} - \frac{G^2}{bk}$	CM(BIT)	
Tratamientos (eliminando bloques)	$v+q-1$	$SC(TEB) = \sum_{i=1}^v \hat{\tau}_i Q_i + \sum_{k^0=1}^q \hat{\tau}_{k^0} Q_{k^0}$	CM(TEB)	CM(TEB)/ $s^2$
Error	$b(qc-1)-q+1$	SCE (por diferencia)	CME= $s^2$	
Total	$bqc+v-1$	$\sum y^2 - \frac{G^2}{bk}$		



para dos nuevos tratamientos en distintos bloques,

$$\hat{\tau}_i - \hat{\tau}_s = Q_i + \frac{1}{qc} \sum_{m=1}^v \left( \sum_{l=1}^b n_{l=1} n_{ml} \right)$$

$$Q_m - Q_s - \frac{1}{qc} \sum_{m=1}^v \left( \sum_{l=1}^b n_{sl} n_{ml} \right) Q_m$$

Se obtienen además los siguientes resultados

$$VAR(\hat{\tau}_{k^0} - \hat{\tau}_i) = \left( \frac{1}{bc} - \frac{1}{qbc} + \frac{1}{qc} + 1 \right) \sigma^2,$$

$$VAR(\hat{\tau}_{k^0} - \hat{\tau}_{m^0}) = \frac{2\sigma^2}{bc},$$

$$VAR(\hat{\tau}_i - \hat{\tau}_s) = 2\sigma^2,$$

para dos nuevos tratamientos en el mismo bloque; y

$$VAR(\hat{\tau}_i - \hat{\tau}_s) = 2 \left( 1 + \frac{1}{qc} \right) \sigma^2,$$

para dos nuevos tratamientos en distintos bloques.

### UN EJEMPLO NUMÉRICO

Considérense los datos del experimento de la Figura 1, generados a partir de números aleatorios.

Por inspección es claro que  $b=3, k=7, q=2, c=2, N=21$ . Después de realizar la aritmética correspondiente, se obtienen las siguientes cantidades:  $B_1=344, B_2=303, B_3=428; T=313, T=310, T_1=49, T_2=49, T_3=63, T_4=53, T_5=12, T_6=60, T_7=70, T_8=17, T_9=79; G=1075; Q=5.857143, Q=2.857143, Q_1=-0.1428571, Q_2=-0.1428571, Q_3=13.857143, Q_4=9.714286, Q_5=-31.285714, Q_6=16.714286, Q_7=8.857143, Q_8=-44.142857, Q_9=17.857143$ . De aquí, usando las ecuaciones (3) y (5), resulta:  $\hat{\tau}_{1^0}=0.9761905, \hat{\tau}_{2^0}=0.4761905, \hat{\tau}_1=3.976190, \hat{\tau}_2=3.976190, \hat{\tau}_3=17.97619, \hat{\tau}_4=9.22619, \hat{\tau}_5=-31.77381, \hat{\tau}_6=16.22619, \hat{\tau}_7=5.22619, \hat{\tau}_8=-47.77381$  y  $\hat{\tau}_9=14.22619$ . Puede

comprobarse que  $\sum_{k=1}^2 Q_{k^0} + \sum_{i=1}^9 Q_i = 0$ , en tanto que  $qc \sum_{k=1}^2 \hat{\tau}_{k^0} + \sum_{i=1}^9 \hat{\tau}_i = 0$ , excepto por errores de redondeo.

$$VAR(\hat{\tau}_i - \hat{\tau}_s) = 2\sigma^2,$$

for two new treatments in the same block; and

$$VAR(\hat{\tau}_i - \hat{\tau}_s) = 2 \left( 1 + \frac{1}{qc} \right) \sigma^2,$$

for two new treatments in different blocks.

### A NUMERICAL EXAMPLE

Let us consider the data of the trial illustrated in Figure 1, which were generated based on random numbers.

Upon inspection it is clear that  $b=3, k=7, q=2, c=2, N=21$ . After doing the relevant arithmetical computations, we obtained the following:  $B_1=344, B_2=303, B_3=428; T=313, T=310, T_1=49, T_2=49, T_3=63, T_4=53, T_5=12, T_6=60, T_7=70, T_8=17, T_9=79; G=1075; Q=5.857143, Q=2.857143, Q_1=-0.1428571, Q_2=-0.1428571, Q_3=13.857143, Q_4=9.714286, Q_5=-31.285714, Q_6=16.714286, Q_7=8.857143, Q_8=-44.142857, Q_9=17.857143$ . From which, by using equations (3) and (5), we obtained:  $\hat{\tau}_{1^0}=0.9761905, \hat{\tau}_{2^0}=0.4761905, \hat{\tau}_1=3.976190, \hat{\tau}_2=3.976190, \hat{\tau}_3=17.97619, \hat{\tau}_4=9.22619, \hat{\tau}_5=-31.77381, \hat{\tau}_6=16.22619, \hat{\tau}_7=5.22619, \hat{\tau}_8=-47.77381$  y  $\hat{\tau}_9=14.22619$ . We can

prove that  $\sum_{k=1}^2 Q_{k^0} + \sum_{i=1}^9 Q_i = 0$ , while  $qc \sum_{k=1}^2 \hat{\tau}_{k^0} + \sum_{i=1}^9 \hat{\tau}_i = 0$ , except for rounding errors.

To obtain mean values of the varieties, we added to the effect of each treatment the estimated general mean

Bloque I		Bloque II		Bloque III	
81	2 <sup>0</sup>	53	4	50	1 <sup>0</sup>
3	1 <sup>0</sup>	19	1 <sup>0</sup>	70	7
49	1	46	2 <sup>0</sup>	39	2 <sup>0</sup>
67	1 <sup>0</sup>	12	5	91	2 <sup>0</sup>
49	2	92	1 <sup>0</sup>	17	8
32	2 <sup>0</sup>	60	6	82	1 <sup>0</sup>
63	3	21	2 <sup>0</sup>	79	9

Figura 1. Experimento con dos testigos y nueve nuevas variedades.  
Figure 1. Experiment with two controls and nine new varieties.

Para obtener valores medios de las variedades se suma al efecto de cada tratamiento la estimación de la media general,

$$\hat{\mu} = \frac{Y_{...}}{bqc + v} = 1075 / (3 \cdot 2 \cdot 2 + 9) = 51.2 \text{ lo que da:}$$

$$\begin{aligned} \hat{\mu}_{1^0} &= 52.17, & \hat{\mu}_{2^0} &= 51.67, & \hat{\mu}_1 &= 55.17, & \hat{\mu}_2 &= 55.17, \\ \hat{\mu}_3 &= 69.17, & \hat{\mu}_4 &= 60.42, & \hat{\mu}_5 &= 19.42, & \hat{\mu}_6 &= 67.42, \\ \hat{\mu}_7 &= 56.42, & \hat{\mu}_8 &= 3.42, & \hat{\mu}_9 &= 65.42. \end{aligned}$$

Completando la aritmética se obtiene el análisis de varianza del Cuadro 3. Se considerará enseguida el cálculo de algunas estimaciones y sus varianzas, sólo para ejemplificar:

$$\begin{aligned} \hat{\tau}_{1^0} - \hat{\tau}_1 &= 0.9761905 - 3.9761905 = -3.0, \\ \hat{\tau}_{1^0} - \hat{\tau}_{2^0} &= 0.9761905 - 0.4761905 = 0.5, \\ \hat{\tau}_1 - \hat{\tau}_3 &= 3.976190 - 17.976190 = -14, \\ \hat{\tau}_1 - \hat{\tau}_4 &= 3.976190 - 9.226190 = -5.25; \end{aligned}$$

$$\begin{aligned} \text{VAR}(\hat{\tau}_{1^0} - \hat{\tau}_1) &= \left[ \frac{1}{bc} - \frac{1}{qbc} + \frac{1}{qc} + 1 \right] \sigma^2 \\ &= \left[ \frac{1}{2 \times 3} - \frac{1}{2 \times 3 \times 2} + \frac{1}{2 \times 2} + 1 \right] \sigma^2 = \frac{4}{3} \sigma^2 \end{aligned}$$

$$\text{VAR}(\hat{\tau}_{1^0} - \hat{\tau}_{2^0}) = \frac{2\sigma^2}{bc} = \frac{2}{3 \times 2} \sigma^2 = \frac{1}{3} \sigma^2,$$

$$\text{VAR}(\hat{\tau}_1 - \hat{\tau}_3) = 2\sigma^2$$

$$\text{VAR}(\hat{\tau}_1 - \hat{\tau}_4) = 2 \left( 1 + \frac{1}{qc} \right) \sigma^2 = 2 \left( 1 + \frac{1}{2 \times 2} \right) = \frac{5}{2} \sigma^2$$

### Uso de SAS para el análisis de los experimentos

Cuando los experimentos se analizan por métodos computacionales, el procedimiento GLM de SAS produce el análisis de la varianza básico para la interpretación de los resultados, en tanto que el comando LSMEANS

$$\hat{\mu} = \frac{Y_{...}}{bqc + v} = 1075 / (3 \cdot 2 \cdot 2 + 9) = 51.2, \text{ which gives:}$$

$$\begin{aligned} \hat{\mu}_{1^0} &= 52.17, & \hat{\mu}_{2^0} &= 51.67, & \hat{\mu}_1 &= 55.17, & \hat{\mu}_2 &= 55.17, \\ \hat{\mu}_3 &= 69.17, & \hat{\mu}_4 &= 60.42, & \hat{\mu}_5 &= 19.42, & \hat{\mu}_6 &= 67.42, \\ \hat{\mu}_7 &= 56.42, & \hat{\mu}_8 &= 3.42, & \hat{\mu}_9 &= 65.42. \end{aligned}$$

We finished the arithmetical calculations to obtain the analysis of variance in Table 3. Next, let us calculate some estimates and their variances, just as examples:

$$\begin{aligned} \hat{\tau}_{1^0} - \hat{\tau}_1 &= 0.9761905 - 3.9761905 = -3.0, \\ \hat{\tau}_{1^0} - \hat{\tau}_{2^0} &= 0.9761905 - 0.4761905 = 0.5, \\ \hat{\tau}_1 - \hat{\tau}_3 &= 3.976190 - 17.976190 = -14, \\ \hat{\tau}_1 - \hat{\tau}_4 &= 3.976190 - 9.226190 = -5.25; \end{aligned}$$

$$\begin{aligned} \text{VAR}(\hat{\tau}_{1^0} - \hat{\tau}_1) &= \left[ \frac{1}{bc} - \frac{1}{qbc} + \frac{1}{qc} + 1 \right] \sigma^2 \\ &= \left[ \frac{1}{2 \times 3} - \frac{1}{2 \times 3 \times 2} + \frac{1}{2 \times 2} + 1 \right] \sigma^2 = \frac{4}{3} \sigma^2 \end{aligned}$$

$$\text{VAR}(\hat{\tau}_{1^0} - \hat{\tau}_{2^0}) = \frac{2\sigma^2}{bc} = \frac{2}{3 \times 2} \sigma^2 = \frac{1}{3} \sigma^2,$$

$$\text{VAR}(\hat{\tau}_1 - \hat{\tau}_3) = 2\sigma^2$$

$$\text{VAR}(\hat{\tau}_1 - \hat{\tau}_4) = 2 \left( 1 + \frac{1}{qc} \right) \sigma^2 = 2 \left( 1 + \frac{1}{2 \times 2} \right) = \frac{5}{2} \sigma^2$$

### Using SAS for trial analysis

When analyzing trials using computational methods, the SAS GLM procedure produces a basic analysis of variance for interpreting the results, while the LSMEANS command requires printing the treatment means adjusted for block effects. In the sample trial, the SAS program in Figure 2 gives the desired results.

**Cuadro 3. Análisis de varianza del ejemplo ilustrativo.**  
**Table 3. Analysis of variance for the illustrative example.**

Fuentes de variación	Grados de libertad	Sumas de cuadrados	Cuadrados medios	F calculada
Bloques (ignorando tratamientos)	2	1160.0952	580.0476	
Tratamientos (eliminando bloques)	10	4019.1429	401.9143	0.36
Error	8	9036.0000	1129.5000	
Total	20	14215.2381		

requiere la impresión de las medias de tratamientos corregidas por los efectos de bloques. Para los resultados del experimento ilustrativo, el programa SAS de la Figura 2 conduce a los resultados deseados.

Los resultados que genera SAS se presentan en el Cuadro 4. Éstos coinciden con los resultados previos, excepto por las estimaciones de efectos, ya que SAS utiliza una restricción distinta para resolver las ecuaciones normales (toma como cero el efecto del último tratamiento). Sin embargo, por resultados de la teoría de la estimación lineal (Searle, 1971), independientemente de la solución mínimo-cuadrática, el estimador de cualquier contraste debe ser el mismo, ya que estas funciones son estimables, y su estimación única para cualquier restricción que se imponga las ecuaciones normales. Esto es fácil de comprobar con la solución sugerida, puesto que si se estiman los contrastes propuestos en líneas anteriores usando las medias mínimo-cuadráticas, se obtiene

$$\hat{\tau}_{1^0} - \hat{\tau}_1 = 52.1667 - 55.1667 = -3.0,$$

$$\hat{\tau}_{1^0} - \hat{\tau}_{2^0} = 52.1667 - 51.6667 = 0.5,$$

$$\hat{\tau}_1 - \hat{\tau}_3 = 55.1667 - 69.1667 = -14.0,$$

$$\hat{\tau}_1 - \hat{\tau}_4 = 55.1667 - 60.4167 = -5.25.$$

Estimaciones idénticas a las previamente encontradas.

### CONSIDERACIONES FINALES

El punto más importante para realizar este tipo de diseños, tal vez sea no tener repeticiones de los tratamientos de mayor interés. Esto supone que la medida obtenida para estos tratamientos no se afecta por el error experimental, y que no hay interacción con la parcela. Al no tener repeticiones, la precisión está determinada únicamente por el error experimental, y si este es grande no se podrán diferenciar estadísticamente los tratamientos de interés. La ejecución de estos diseños requiere de un buen manejo experimental. Para eliminar esta restricción se han propuesto distintas metodologías de análisis que permiten corregir el resultado individual de una parcela, por las parcelas vecinas; estas metodologías incluyen la confección de bloques muy pequeños, la corrección a través de vecinos cercanos, ya sea por diferenciación o uso de covariables, el ajuste de polinomios de tendencia y más recientemente la definición de estructuras especiales para la matriz de varianzas y covarianzas de los errores. Si los bloques fueran de tamaño 2 se obtendrían resultados similares a los obtenidos con las otras metodologías propuestas, pues entonces se elimina cualquier tipo de relación entre los errores experimentales. La matriz de varianzas y

Results produced by SAS are shown in Table 4. They agree with previous results, except for the estimated effects, since SAS uses a different restriction to solve normal equations (it considers the effect of the last treatment to be zero). However, based on the results of the linear estimation theory (Searle, 1971), and independently of the least squares solution, the estimator of any contrast has to be the same, given that these functions are estimatable and their estimation is unique for any restriction imposed to the normal equations. This is easily proven using the suggested solution, given that, if the contrasts proposed above are estimated using the least squares means, we obtain

$$\hat{\tau}_{1^0} - \hat{\tau}_1 = 52.1667 - 55.1667 = -3.0,$$

$$\hat{\tau}_{1^0} - \hat{\tau}_{2^0} = 52.1667 - 51.6667 = 0.5,$$

$$\hat{\tau}_1 - \hat{\tau}_3 = 55.1667 - 69.1667 = -14.0,$$

$$\hat{\tau}_1 - \hat{\tau}_4 = 55.1667 - 60.4167 = -5.25.$$

Estimates identical to the ones previously found.

### FINAL CONSIDERATIONS

Perhaps the most important thing to remember when developing these designs is the fact that the most relevant treatments are unreplicated. This assumes that the measure of these treatments is not greatly affected by experimental error and there is no plot interaction. Since there are no replications, the precision is given only by the experimental error, so that if the error is big, the treatments of interest cannot be statistically differentiated. Applying these designs, thus, requires good trial management. Various analysis methodologies that allow adjusting individual plot results based on neighboring plots have been proposed to eliminate this restriction. These methodologies include setting up very small blocks, adjusting them based on neighboring plots, either through differentiation or the use of covariables, adjusting the polynomial trends and, more recently, defining special structures for the error variance-covariance matrix. If the blocks were of size 2, results would be similar to those obtained using other proposed methodologies, because any relationship between the experimental errors is eliminated. The error variance-covariance matrix is of the

```
DATA UNO; INPUT BLOQUE TRAT $ Y @@; CARDS;
1 2° 81 1 1° 3 1 1 49 1 1° 67 1 2 49 1 2° 32 1 3 63
2 4 53 2 1° 19 2 2° 46 2 5 12 2 1° 92 2 6 60 2 2° 21
3 1° 50 3 7 70 3 2° 39 3 2° 91 3 8 17 3 1° 82 3 9 79
PROC GLM; CLASSES BLOQUE TRAT;
MODEL Y=BLOQUE TRAT; LSMEANS TRAT; RUN;
```

Figura 2. Programa SAS para el análisis del ejemplo ilustrativo.  
Figure 2. SAS program for the analysis of the illustrative example.

**Cuadro 4. Resultados de SAS.**  
**Table 4. SAS results.**

General Linear Models Procedure						
Dependent Variable: Y						
Source	DF	Sum of squares	Mean square	F value	Pr > F	
Model	12	5179.2	431.6	0.38	0.9352	
Error	8	9036.0	1129.5			
Corrected Total	20	14215.2				
	R-Square	C.V.	Root MSE		Y Mean	
	0.364344	65.65291	33.6		51.2	
Source	DF	Type I SS	Mean Square	F Value	Pr > F	
BLOQUE	2	1160.1	580.0	0.51	0.6168	
TRAT	10	4019.1	401.9	0.36	0.9357	
Parameter	Estimate		Standard error	t Value	Pr >  t	
Intercept	79.00000000 B		33.60803475	2.35	0.0466	
BLOQUE 1	-19.75000000 B		23.76446928	-0.83	0.4300	
BLOQUE 2	-21.00000000 B		23.76446928	-0.88	0.4026	
BLOQUE 3	0.00000000 B		.	.	.	
TRAT 1	-10.25000000 B		53.13896875	-0.19	0.8519	
TRAT 1°	-13.25000000 B		38.80721582	-0.34	0.7416	
TRAT 2	-10.25000000 B		53.13896875	-0.19	0.8519	
TRAT 2°	-13.75000000 B		38.80721582	-0.35	0.7323	
TRAT 3	3.75000000 B		53.13896875	0.07	0.9455	
TRAT 4	-5.00000000 B		53.13896875	-0.09	0.9273	
TRAT 5	-46.00000000 B		53.13896875	-0.87	0.4119	
TRAT 6	2.00000000 B		53.13896875	0.04	0.9709	
TRAT 7	-9.00000000 B		47.52893855	-0.19	0.8545	
TRAT 8	-62.00000000 B		47.52893855	-1.30	0.2283	
TRAT 9	0.00000000 B		.	.	.	

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

Least Squares Means	
TRAT	Y
	LSMEAN
1	55.1666667
1°	52.1666667
2	55.1666667
2°	51.6666667
3	69.1666667
4	60.4166667
5	19.4166667
6	67.4166667
7	56.4166667
8	3.4166667
9	65.4166667

covarianzas de los errores es de la forma  $\lambda \mathbf{I}$  y los ajustes por vecinos se realizan al comparar dentro de bloque. Un diseño con bloques de tamaño 2 implicaría el uso de excesivo material experimental, por lo que sería recomendable cuando se quiere comparar un número pequeño de nuevas variedades. La inclusión de un mayor número de testigos y más repeticiones de éstos permite mejorar la confiabilidad y precisión de las estimaciones, con la limitante de tener un experimento extremadamente grande. En general se recomienda el uso de 2 a 4 testigos, que ocupen entre 10 y 20% de las unidades experimentales.

form  $\lambda \mathbf{I}$  and adjustments by neighboring plots are performed when making within-block comparisons. A design with size 2 blocks would imply the use of many test lines and is recommended for comparing a small number of new varieties. Including a greater number of controls as well as more replications of those checks makes it possible to improve the reliability and precision of the estimates, but results in an extremely large trial. In general we recommend using 2-4 controls that would occupy 10-20% of the test plots.

—End of the English version—

**AGRADECIMIENTOS**

Se agradece al Editor los comentarios realizados, que contribuyeron a la presentación de un trabajo mejor y más completo.

**LITERATURA CITADA**

- Besag, J. E., and R. A. Kempton. 1986. Statistical analysis of field experiments using neighbouring plots. *Biometrics*, 42: 231-251.
- Box, G. E. P., and K. B. Wilson. 1951. On the experimental attainment of optimum conditions. *J. R. Stat. Soc. B* 13.1-45.
- Federer, W. T. 1961. Augmented designs with one-way elimination of heterogeneity. *Biometrics*, 17: 447-473.
- Federer, W. T., M. Singh, and R. D. Wolfinger. 2000. SAS/GLM and SAS/MIXED for trend analysis using fourier and polynomial regression for centered and noncentered varieties. Technical Report Series. Department of Biometrics. Cornell University. BU-1498-M. 17 p.
- Lin, C. S., and G. Poushinsky. 1983. A modified augmented design for an early stage of plant selection involving a large number of test lines without replication. *Biometrics*, 39: 553-561
- Martínez Garza, A. 1988. Diseños experimentales. Ed. Trillas. México. pp: 95.
- Méndez, I. 1976. Métodos alternativos para comparación de un número grande de variedades usando funciones de tendencia. *Agrociencia*. 26: 153-165.
- Papadakis, J. S. 1984. Advances in the Analysis of Field Experiments. *Proc. Acad. Athens*. 59: 326-342.
- Searle, S. R. 1971. Linear models. J. Wiley & Sons. New York. pp: 88, 180-188.