Sesgo al estimar la varianza de efectos directos y totales de tratamientos en diseños Cuadrados Latinos Balanceados

Bias to estimate the variance of direct and residual effects of treatments on Balanced Latin Square designs

RESUMEN

El sesgo que puede producirse al estimar la varianza de efectos directos y residuales de tratamientos en diseños Cuadrados Latinos Balanceados, fue estudiado mediante la utilización de las producciones diarias de leche de 72 vacas Holstein de mediano potencial, las cuales estaban sometidas a un régimen de manejo y alimentación similares. Los resultados obtenidos demuestran que para evitar estos sesgos, las varianzas de efectos directos y residuales de tratamientos en este tipo de diseño, deben ser calculadas partiendo de la determinación de las varianzas de los contrastes lineal, cuadrático y de mayor orden entre vacas.

Palabras clave: Sesgo, Efectos directos y residuales, Cuadrados Latinos Balanceados, Contrastes.

Dr. Andrés Venereo Bravo PhD

Licenciado en Matemática
Doctor en Ciencias Agrícolas
Universidad Laica Eloy Alfaro de Manabí

ABSTRACT

The bias, which may occur when estimating the variance of direct and residual effects of treatments with Balanced Latin Square design, was studied by using the daily milk production of 72 Holstein cows in mid-potential, which were placed under a similar management and feeding regimen. The results demonstrate that to avoid this bias, the variances on direct and residual effects of treatments in this type of design, should be calculated based on the determination of the variances of linear contrast, quadratic and of higher order among cows.

Key words: Bias, Direct and residual effects, Balanced Latin Squares, Contrasts.

Recibido: 25 de noviembre, 2013
Aceptado: 11 de noviembre, 2014

1. INTRODUCCIÓN

Los diseños Cuadrados Latinos Balanceados han sido ampliamente utilizados desde su creación para el desarrollo de experimentos en producción de leche, en los que la aplicación de tratamientos en secuencia al mismo animal, puede producir efectos residuales entre un período experimental y otro.

La gran popularidad alcanzada por este tipo de diseño está dada por el satisfactorio control que permite sobre el error experimental, y por el reducido número de animales requeridos para el desarrollo de experimentos.

De igual forma, el método propuesto por Cochran, Autrey y Cannon (1941) para el análisis estadístico de tales experimentos, ha sido utilizado de manera general para sacar conclusiones a partir de los resultados obtenidos con la aplicación de estos diseños.

Sin embargo, en la aplicación práctica de este método, una seria deficiencia del mismo debe ser tomada en consideración.

En el método sugerido por estos autores, se parte de la suposición de que la variación de los componentes lineal, cuadrático y de mayor orden de vaca a vaca, se mantienen iguales dentro de cada experimento.

Patterson (1950) ha señalado que en experimentos en producción de leche, en los que la media de períodos exhibe una tendencia notable, esta suposición puede no ser cierta, ya que si la misma se analiza en sus componentes, la variación lineal de vaca a vaca, puede ser considerablemente mayor que las de mayor orden.

Cuando esto sucede, el método propuesto por Cochran *et al.* (1941) no ofrece una estimación insesgada del error y, en consecuencia, el error estándar para las medias de efectos directos y totales de tratamientos está sujeto a sesgos que deben ser eliminados con la aplicación de un modelo más real (Cochran y Cox, 1957).

El presente trabajo tiene como objetivo desarrollar un estudio sobre el sesgo que puede producirse al estimar la varianza de efectos directos y totales de tratamientos en diseños Cuadrados Latinos Balanceados, cuando se aplica el método propuesto por Cochran *et al.* (1941).

2. MATERIALES Y MÉTODOS

Para el presente estudio fueron utilizadas las producciones diarias de leche entre los 6 y 250 días de lactancia de 72 vacas Holstein de mediano potencial, de las cuales 36 eran de primera lactancia y 36 de segunda, las que estaban sometidas a un régimen de manejo y alimentación similares.

Con estas producciones de leche fueron simulados seis experimentos Cuadrados Latinos Balanceados con tres tratamientos y 12 vacas cada uno.

Para la conformación de cada experimento se utilizaron períodos de adaptación y colección de 7 y 21 días de duración, respectivamente, iniciándose los mismos cuando las vacas tuvieron los días de lactancia requeridos según la duración de cada experimento.

La estimación de la varianza experimental y de las varianzas de las diferencias entre efectos directos y residuales de los tratamientos, así como la covarianza entre los mismos, fueron calculadas según dos vías de estimación diferentes, a saber:

1.- Mediante el análisis de varianza de los datos según el modelo lineal:

$$\mathbf{Y_{ijklm} = \mu + t_i + r_j + b_k + f_l + (fb)_{kl} + c_{m(k)} + e_{ijklm}}$$

donde:

$\mathbf{Y_{ijklm}}$= observación correspondiente al efecto directo i- ésimo y residual j-ésimo en la fila l-ésima y columna m-ésima del cuadrado k-ésimo.

μ= constante común a todas las observaciones.
$\mathbf{t_i}$= efecto directo del tratamiento i-ésimo.
$\mathbf{r_j}$= efecto residual del tratamiento j-ésimo.
$\mathbf{b_k}$= efecto del cuadrado k-ésimo.
$\mathbf{f_l}$= efecto de la fila l-ésima.
$\mathbf{(fb)_{kl}}$= efecto de la interacción fila por cuadrado kl-ésimo.
$\mathbf{c_{m(k)}}$= efecto de la columna m-ésima dentro del cuadrado k-ésimo.
$\mathbf{e_{ijklm}}$= variable aleatoria asociada a la ijklm – ésima observación, normal e independientemente distribuida con media cero y varianza σ^2, según el cual, las estimaciones de las varianzas anteriormente señaladas vienen dadas por las expresiones:

$$\hat{\sigma}_e^2 = CM_e$$
$$\hat{V}(d_i - d_j) = \frac{2\sigma_e^2(n^2 - n - 1)}{mn(n^2 - n - 2)} = 0.208333\sigma_e^2$$
$$\hat{V}(r_i - r_j) = \frac{2n\sigma_e^2}{n(n^2 - n - 2)} = 0.375\sigma_e^2$$
$$\hat{Cov}(d_i - d_j, r_i - r_j) = \frac{2n\sigma_e^2}{m(n^2 - n - 2)} = 0.125\sigma_e^2$$

donde:

CM_e = cuadrado medio del error en el análisis de varianza.
n = número de tratamientos (n=3).
m = número de cuadrados (m=4).
d_i , r_i = efecto directo y residual i-ésimo.

2.- Mediante la minimización de la suma de cuadrados ponderada:

$$\sum \frac{W_i \sum_k \left[\sum_j \varepsilon_{ij}(y_j - Y_j) \right]^2}{\sum_j \varepsilon_{ij}^2}$$

donde ε_{ij} es el j-ésimo coeficiente en el polinomio ortogonal de grado i (Fischer y Yates, 1948) para $1 \leq j \leq n$ y $1 \leq i \leq n-1$ (n = 3 = número de tratamientos).

Según este método, las estimaciones de la varianza experimental y de las varianzas de las diferencias entre efectos directos y residuales de tratamientos, así como la covarianza entre los mismos, vienen dadas por las expresiones:

$$\hat{\sigma}_e^2 = \frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_2^2$$
$$\hat{V}(d_i - d_j) = \frac{2(n^2 - n - 1)}{mn(n^2 - n - 2)}\left(\frac{7}{10}\sigma_1^2 + \frac{3}{10}\sigma_2^2\right) = 0.208333\left(\frac{7}{10}\sigma_1^2 + \frac{3}{10}\sigma_2^2\right)$$
$$\hat{V}(r_i - r_j) = \frac{2n}{m(n^2 - n - 2)}\left(\frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_2^2\right) = 0.375\left(\frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_2^2\right)$$
$$\hat{Cov}(d_i - d_j, r_i - r_j) = \frac{2\sigma_1^2}{m(n^2 - n - 2)} = 0.125\sigma_1^2$$

en las cuales:

σ_1^2 = varianza del contraste lineal entre vacas.
σ_2^2 = varianza del contraste cuadrático entre vacas.
n = número de tratamientos (n = 3).
m = número de cuadrados (m = 4).
y:

$$\sigma_i^2 = \frac{1}{W_i} = \frac{E\left[\sum_j \varepsilon_{ij}\left(y_j - Y_j\right)\right]^2}{\sum_j \varepsilon_{ij}^2}$$

Como puede apreciarse, cuando $\sigma_1^2 = \sigma_2^2 = \sigma_e^2$ ambos métodos de estimación son coincidentes, y las estimaciones de las varianzas señaladas también lo serán.

Sin embargo, cuando $\sigma_1^2 = k\ \sigma_2^2$ (K>1), el método (1) produce sistemáticamente un sesgo sobre la varianza de la diferencia entre efectos directos y totales de tratamientos.

El sesgo de la estimación fue calculado mediante la relación:

$$S = \frac{V\ (\textit{método 1})\ \text{-}\ V\ (\textit{método 2})}{V\ (\textit{método 2})}\ x\ 100$$

3. RESULTADOS

Los resultados de la estimación de las varianzas según el método (1), para cada uno de los seis experimentos considerados, se muestran en la TABLA 1: en la cual σ_e^2 representa una estimación insesgada de $\frac{\sigma_1^2 + \sigma_2^2}{2}$, es decir, de la semi – suma de las varianzas de los contrastes lineal y cuadrático entre vacas, las cuales se suponen iguales.

Los resultados del cálculo de estas varianzas así como el de las reportadas en la TABLA 1, aplicando el método (2), se aprecian en la TABLA 2 y TABLA 3:

Tabla 1. Estimación de la varianza experimental y de las varianzas de la diferencia entre efectos directos y residuales según método (1).

Experimento	$\hat{\sigma}_e^2$	$\hat{V}(d_i - d_j)$	$\hat{V}(r_i - r_j)$	$\hat{Cov}(d_i - d_j, r_i - r_j)$
1	2.7844	0.5801	1.0441	0.3480
2	3.7222	0.7754	1.3958	0.4652
3	1.8550	0.3864	0.6956	0.2318
4	3.0646	0.6384	1.1492	0.3831
5	2.3748	0.4948	0.8906	0.2968
6	2.8582	0.5954	1.0718	0.3572

Tabla 2. Estimación de las varianzas del contraste lineal y cuadrático entre vacas.

Experimento	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\frac{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}{2}$
1	2.4602	1.9888	2.2247
2	6.2492	1.3210	3.7851
3	1.7565	0.5733	1.1649
4	3.0684	1.9762	2.5224
5	7.3636	1.1508	4.2572
6	3.9748	0.9934	2.4841

Tabla 3. Estimación de las varianzas de la diferencia entre efectos directos y residuales según el método (2).

Experimento	$\hat{V}(d_i - d_j)$	$\hat{V}(r_i - r_j)$	$\hat{Cov}(d_i - d_j, r_i - r_j)$
1	0.4831	0.8342	0.3076
2	0.9939	1.4194	0.7812
3	0.2920	0.4368	0.2196
4	0.5710	0.9438	0.3836
5	1.1458	1.5964	0.9204
6	0.6417	0.9315	0.4968

DISCUSIÓN

Los resultados de la TABLA 2 muestra como en los seis experimentos estudiados, las varianzas del contraste lineal entre vacas resultó muy superior a las obtenidas para el contraste cuadrático.

Este resultado indica que la estimación de la varianza experimental según el método (1) se encuentra sesgada al no cumplirse la condición $\sigma_1^2 = \sigma_2^2$, y en consecuencia, estarán sesgadas también las varianzas de la diferencia entre efectos directos y residuales de tratamientos y su respectiva covarianza. El cálculo de este sesgo se muestra en la TABLA 4.

CONCLUSIONES

Los resultados demuestran que el método de análisis estadístico propuesto por Cochran *et al.* (1941), no debe ser utilizado para la estimación de las varianzas de efectos directos y residuales de tratamientos en experimentos con vacas lecheras, dado el sesgo que puede producir el mismo.

En su lugar, tales varianzas deben ser obtenidas partiendo de la determinación de las varianzas de los contrastes lineal, cuadrático y de mayor orden entre vacas los cuales pueden ser calculados haciendo uso de los polinomios ortogonales (Fischer y Yates, 1948).

Las conclusiones aquí reportadas mantienen vigencia para los casos de experimentos con más de tres tratamientos, ya que en tales situaciones, al ser mayor el número de períodos experimentales, se produce una diferencia aún mayor entre la componente lineal y las componentes de orden superior.

Tabla 4. Sesgo (%) en la estimación de las varianzas de la diferencia entre efectos directos y residuales obtenidas con el método (1).

Experimento	$\hat{V}(d_i - d_j)$	$\hat{V}(r_i - r_j)$	$\hat{Cov}(d_i - d_j r_i - r_j)$
1	20.08	25.16	13.13
2	-21.98	-1.66	-40.45
3	32.32	59.24	5.56
4	11.80	21.51	-0.13
5	-56.82	-44.21	-67.75
6	-7.22	15.06	-28.10

REFERENCIAS BIBLIOGRÁFICAS

Cochran, W.G., Autrey, K.M. and Cannon, C.Y. 1941. A double change-over designs for dairy cattle feeding experiments. J. Dairy Sci., 24:937.

Cochran, W.G. and Cox, G.M. 1957. Experimental Designs. John Wiley and Sons, Inc., New York. Sec. ed.

Fischer, R.A. and Yates, F. 1948. Statistical Tables for biological, agricultural and medical research. Oliver and Boyd, Edimburgo, 3a. ed.

Patterson, H.D. 1950. The analysis of changeover trials. J. Agric. Sci., 40:375.