UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE AGRONOMÍA
COMISIÓN DE ESTUDIOS DE POSTGRADO
POSTGRADO EN ESTADÍSTICA

Métodos de Máxima Verosimilitud para la Estimación de Parámetros en Diseños Factoriales Mixtos pxq y sus Aplicaciones en la Agroindustria

Autor: Ing. José L. Romero
Tutor: Dr. Manuel Milla
Consejeros: **Dr. Miguel Balza**
Dr. Franklin Chacín

Trabajo de Grado presentado ante la Comisión de Estudios de Postgrado de la Facultad de Agronomía de la Universidad Central de Venezuela, como requisito final para optar al título de **Magister Scientiatum en Estadística.**

Maracay, Julio 2015

Agradecimientos

A Dios sobre todas las cosas.

A mis padres, Jaime Rafael y Luisa Elena, por inculcarme la actitud de la constancia y paciencia ante la vida.

Al Dr. Manuel Milla por ser mi tutor y formador en estadística y al Dr. Omar Verde por su valiosa colaboración y aportes orientadores en la conducción del presente trabajo de investigación.

A todos los profesores del Postgrado de Estadística de la Facultad de Agronomía UCV-Maracay. Todo ellos, de una otra manera contribuyeron y contribuyen en mi formación académica.

A mi hija Marianny Yurubi y mi novia Luz Marina por llegar en esta etapa de mi vida.

Agradecido, a todos.

Dedicatoria

“A la firmeza y perseverancia”

Tabla de Contenido

Índice de Cuadros.

Resumen

El presente estudio describe los procedimientos teóricos de los Métodos de Estimación por Máxima Verosimilitud Clásica, Máxima Verosimilitud Restringida y Máxima Verosimilitud Ponderada, con el propósito de contrastar estos métodos se considera las propiedades deseada de estimador en cuanto al insesgamiento, consistencia, eficiencia, suficiencia y completitud, en un escenario de diseños factoriales (pxq), con grupos de datos balanceados o desbalanceados, en condiciones de varianzas homogéneas o heterogéneas con tamaño de muestra n=10 en cada combinación de tratamientos.

Para contratar los métodos de estimación mencionada con antelación, se ilustrara usando un conjunto de datos ficticios, se estudia la eficiencia, obteniendo la Máxima Verosimilitud Ponderada como el método de estimación que arrojó menor cuadrados medios residuales inclusive en condiciones extremas como en datos desbalanceados y violación del supuesto de homogeneidad. En contraste con los métodos de estimación por Máxima Verosimilitud Clásica, Restringida y Mínimos Cuadros Ordinarios, mostrando estos ser sensibles a las condiciones extremas.

Se encontró que la Máxima Verosimilitud Ponderada sea el Método de Estimación en los modelos fijos y mixtos conducido bajo la estructura considerada en el presente trabajo, se sugiere continuar trabajando en línea de investigación, que permita aportar mayor información en la experimentación agroindustrial y evalúen alternativas de eficiencias de métodos.

Palabras claves: Estimación de Parámetros, Diseños Factorial Mixtos, Máxima Verosimilitud Clásica, Máxima Verosimilitud Restringida, Máxima Verosimilitud Ponderada.

Abstract

The present study describes the theoretical procedures Methods maximum likelihood estimation Classic, Restricted Maximum Likelihood and Maximum Weighted Likelihood, in order to contrast these methods are considered the desired properties of an estimator in terms of unbiasedness, consistency, efficiency, adequacy and completeness in a factorial design (pxq) stage, with a groups balanced or unbalanced data, under homogeneous or heterogeneous variance with sample size n = 10 in each treatment combination.

To hire estimation methods mentioned in advance, it will be illustrate using a set of dummy data, efficiency is studied, obtaining the maximum likelihood weighted as the estimation method showed less residual squares means even in extreme conditions and unbalanced data and rape the assumption of homogeneity. In contrast to the methods of maximum likelihood estimation Classic, Restricted and Ordinary least squares showing these extreme conditions sensitive to be.

It was found that the maximum likelihood Weighted is the estimation method in fixed and mixed models conducted under the structure considered in this work, it is suggested to continue working online research, allowing provide more information in the agribusiness experimentation and evaluate alternatives methods efficiencies.

Key Words: Parameter Estimation, Mixed Factorial Designs, Classical Maximum Likelihood, Restricted Maximum Likelihood, Weighted Maximum Likelihood.

1. Introducción

En investigaciones científicas en el ámbito Agroindustrial, se requiere hacer estimaciones de parámetros que permitan tomar decisiones acertadas en el marco de la inferencia estadística. Esta necesidad se presenta con frecuencia al conducir experimentos que obedecen a modelos lineales aditivos con efectos fijos, con efectos aleatorios o con efectos mixtos. La estimación de parámetros es un proceso inferencial mediante la cual se extraen muestras de forma aleatoria de una población de interés, con el propósito de obtener una aproximación o estimar el verdadero valor del parámetro Θ, o un conjunto de valores posibles de parámetro Θ, a partir de los datos muestrales.

Para esto es indispensable la utilización de Métodos de Estimaciones que orienten a determinar los posibles valores de los parámetros. Existen distintos procedimientos para estimar parámetros de una función de distribución de probabilidad, entre los cuales se puede mencionar: Estimación de los Momentos, Estimación Bayesiana, Estimación por Mínimos Cuadrados y Estimación de Máxima Verosimilitud. Este último se desarrollará en el presente estudio y se abordarán los aspectos teóricos de máxima verosimilitud clásica, máxima verosimilitud restringida y máxima verosimilitud ponderada.

Bajo la consideración anterior, la estimación por máxima verosimilitud clásica desarrollada por Fisher (1922), citado por Gallego (2008), consiste en encontrar los valores del parámetro que maximiza la probabilidad de obtener la muestra observada.

Por otro lado, la estimación por máxima verosimilitud restringida, inicialmente desarrollado por Anderson y Bancroft (1952) y Thompson (1962), citados por Searle et al., (1992), consiste en la combinación lineal de los elementos del vector de datos, de tal manera que las combinaciones no contienen efectos fijos.

Así mismo, se presentará el método de estimación por máxima verosimilitud ponderada, el cual se basa en combinar la información de conjunto de datos independientes, cuando en dos o más conjuntos de datos se obtienen variables con medias similares en diferentes muestras obtenidas en condiciones de diferentes grados de precisión y sesgo,(Wang 2001).

De esta manera, el presente trabajo va dirigido a contrastar los diferentes métodos de estimación por máxima verosimilitud mencionados con antelación, bajo un Diseño Factorial (pxq), con grupos de datos balanceados o desbalanceados y en presencia de varianzas homogéneas o heterogéneas, en modelos de efectos fijos y mixtos sometiendo a estudio las bondades de cada uno de los métodos de estimación por máxima verosimilitud ya tendiendo a las propiedades deseadas de un estimador, como son: insesgamiento, consistencia, eficiencia, suficiencia y completitud, y a los criterios de estructura matricial de varianza y covarianza, nivel de significancia y potencia de prueba, que aporten validez a las estimaciones de los parámetros sometidos a estudio.

Para dar mayor entendimiento, se ilustrará con datos provenientes del sector agroindustrial, y se mostrará cuál de los métodos de estimación por verosimilitud es más adecuado al aproximarse al posible valor del parámetro.

Objetivo General

Presentar los Métodos de Máxima Verosimilitud Clásica, Restringida y Ponderada como alternativa para la estimación de parámetros en diseños factoriales (p x q) en modelos de efecto fijos y de efectos mixtos.

Objetivos Específicos

- Describir el procedimiento teórico del método de estimación por Máxima Verosimilitud Clásica, casos: modelos de efectos fijos y modelos de efectos mixtos.

- Describir el procedimiento teórico del método de estimación por Máxima Verosimilitud Restringida, casos: modelos de efectos fijos y modelos de efectos mixtos.

- Describir el procedimiento teórico del método de estimación por Máxima Verosimilitud Ponderada, casos: modelos de efectos fijos y modelos de efectos mixtos.

- Contrastar los métodos de estimación por Máxima Verosimilitud, en base a las propiedades de los estimadores: Insesgamiento, Consistencia, Eficiencia, Suficiencia y Completitud.

- Ilustrar los métodos de estimación por Máxima Verosimilitud usando un conjunto de datos ficticios.

2. Revisión de Literatura

2.1 Propiedades Deseadas en un Estimador

Villa (2008), plantea que diversos autores como Batista et al., (2001), García (2004), Melia (2003), Morris (1988), Ross (2000) mencionan las propiedades de un buen estimador de la manera siguiente:

A.- Insesgamiento: Se dice que un estimador es insesgado si el valor esperado es igual al verdadero valor del parámetro que se está estimando, esto es:

$$\mathrm{E}[Estimador] = Parametro$$

Considerando que la esperanza matemática se refiere al valor promedio, esta propiedad es cerrada, es decir, se concentra en que sea solo el valor promedio el que sea igual al parámetro y, por lo tanto, no permite que sea la mayor parte o cualquiera de los valores del estimador que esté aceptablemente próximo al verdadero valor del parámetro.

B.- Eficiencia: Se dice que un estimador es eficiente si posee la “menor varianza posible”. Se considera al estimador como “bueno” cuando además de tener una varianza mínima, también es insesgado. La unión de estas dos condiciones en un estimador es conocida como “el mejor estimador insesgado”, donde el termino hace referencia a la varianza mínima.

Para evaluar la eficiencia de un estimador, se hace uso de la matriz de información de Fisher, la cual se define como:

$$\mathrm{I}(\theta) = \mathrm{E}\left[\lambda`(X/\theta)^{2}\right] = -\mathrm{E}[\lambda''(X/\theta)]$$

Donde $[\lambda(X/\theta)] = \log f(X/\theta)$, X es una variable aleatoria con función de probabilidad $f(X/\theta)$ que depende de un parámetro desconocido $\theta \in \Omega$. Por lo tanto, $[\lambda''(X/\theta)] = \dfrac{\partial^2 \log f(X/\theta)}{\partial \theta^2}$. De aquí se obtiene que la matriz de varianza θ es la inversa de la información de Fisher, esto es $V(\hat{\theta}) = [-E(\lambda''(\theta))]^{-1}$. Cuando θ corresponde a un vector del parámetro, entonces la matriz de varianza–covarianza está dada como la función de las segundas derivadas, es decir, la matriz Hessiana.

El siguiente teorema nos da otra expresión para $I(\theta)$, más sencillo de determinar.

Teorema # 1. Si $\mathrm{I}(\theta)$ es la cantidad (matriz de información) de Fisher dado por X sobre el parámetro (θ) entonces $I(\theta) = nI(\theta)$, Lacourly. N (2008).

Por otro lado, la desigualdad de Rao-Cramer, para un estimador de θ, establece que la varianzas para todos los posibles estimadores de θ tiene por límite el cociente:

$$V(\hat{\theta}) \geq \frac{\left[\mathrm{E}'\left(\hat{\theta}\right)\right]^2}{nI\left(\theta\right)}$$

De modo que los estimadores que alcanzan dicha desigualdad son considerados los estimadores "más eficientes". Siguiendo este orden y a fin de verificar entre un conjunto de estimadores cual es el que posee la menor varianza, matemáticamente se calcula la razón de los dos estimadores a comparar, lo que es conocido como eficiencia relativa de Varianza (Estimador 2) con respecto al Varianza (Estimador 1) según Montgomery y Runger (2003)y se expresa por:

$$\textbf{Eficiencia Relativa} = \frac{Varianza(Estimador\ 1)}{Varianza(Estimador\ 2)}$$

Montgomery y Runger (2003), si el resultado de la eficiencia relativa es menor que uno (1), se considera que (Estimador 1), es un estimador más eficiente que el parámetro del (Estimador 2), en el sentido de que posee un cuadro medio del error menor.

C.- Suficiencia: Sea $X_1, X_2,......, X_n$ una muestra aleatoria de una distribución continua o discreta, cuya *fdp* o *fd* está dada por $f(x/\theta)$, donde θ es el parámetro desconocido tal que $\theta \in \Omega$. Entonces $\delta(X_1, X_2,......, X_K)$ es un estadístico suficiente si y solo si, la *fdp* conjunta o *fd* conjunta $f(x/\theta)$ de $X_1, X_2,......, X_n$ se puede factorizar para todos los valores de $x=(x_1, x_2,......, x_n) \in R^n$ y todos los valores de $\theta \in \Omega$ de la siguiente manera:

$$f_n(x/\theta) = u(x)v[\delta(x), \theta]$$

Donde la función u y v, son no negativos, la función u solo depende de (x) la función de v solo depende de θ y el valor observado de (x) únicamente a través de los valores $\delta(x)$, lo cual se puede extender a (k) estadísticos, es decir, bajo las condiciones establecidas, se dice que los estadísticos $\delta_1,.........,\delta_k$ son conjunto suficiente si la *fdp* conjunta o *fd* conjunta $f(x/\theta)$ de $X_1, X_2,......, X_n$ se puede factorizar como sigue, para todos los valores de $x=(x_1, x_2,......, x_n) \in R^n$ y todos los valores de $\theta \in \Omega$:

$$f_n(x/\theta) = u(x)v[\delta_1(x),......\delta_k(x), \Theta]$$

Donde u conserva sus características, v ahora dependerá de Θ y (x) únicamente a través de las $\delta_1(x),........,\delta_k(x)$. Finalmente, la suficiencia es una propiedad necesaria para la eficiencia, ya que un estimador es llamado suficiente cuando incluye toda la información necesaria que la muestra puede proporcionar acerca del parámetro.

D.- Consistencia: La consistencia es considerada como una propiedad asintótica de los estimadores, esto es, se considera el comportamiento (en variabilidad) de convergencia de los estimadores con respecto al parámetro que se está estimando cuando $n \rightarrow \infty$. La palabrea "convergencia" hace alusión a que el valor esperado debe acercarse al verdadero valor del parámetro y que la varianza del estimador debe hacerse despreciable.

Al respecto, Meyer (1973), citado por Villa (2008), expone las siguientes definiciones.

Definición 1 Estimación Insesgada de Varianza Mínima.

Sea $\hat{\theta} = \delta(X_1, X_2, \ldots\ldots, X_K)$ una estimación insesgada del parámetro desconocido θ asociada con la distribución de la variable aleatoria (X). Entonces, $\hat{\theta}$ es una estimación insesgada de varianza mínima de θ si para todos los estimadores θ^* tales que $\mathrm{E}[\theta^*] = \theta$, se cumple que $V(\theta) \leq V(\theta^*)$, es decir entre todas las estimaciones insesgadas de θ, $\hat{\theta}$ tiene la varianza más pequeña.

Definición 2 Convergencia de una Estimación.

Sea $\hat{\theta}$ una estimación (basada en una muestra $X_1, X_2, \ldots, X_n$ del parámetro θ). Se dice que $\hat{\theta}$. Se una estimación convergente de θ, si

$$\lim_{n \to \infty} P_\theta\left(\left|\theta - \hat{\theta}\right| \geq \varepsilon\right) = 0, \quad \forall_\varepsilon > 0$$

Teorema#1 Sea $\hat{\theta}$ una estimación de θ en una muestra de tamaño n .Si $\lim\limits_{n\to\infty}$ $E(\hat{\theta})=\theta$ y si $\lim\limits_{n\to\infty} V(\hat{\theta})=0$, entonces $\hat{\theta}$ es un estimador convergente de θ.

Definición 3 Mejor Estimador Lineal Insesgado (MELI).

Se dice que $\hat{\theta}$ es "El mejor estimador lineal insesgado" si $E(\hat{\theta})=\theta$, donde $\hat{\theta}=\sum_{i=1}^{n} a_i x_i$, esto es, $\hat{\theta}$ es una función lineal de la muestra. $V(\hat{\theta})\leq V(\theta^*)$, en donde θ^* es cualquier estimador de θ que satisface las dos condiciones anteriores.

Definición 4 Estimadores Cuadráticos Uniforme Insesgados de Mínima Varianza.

Sea $(Y_1, Y_2, \ldots\ldots, Y_n)$ variables aleatorias con *fdp* acumulada $F_Y(.:\gamma)$, donde γ pertenece a un espacio parametral Ω. Se define a $\tau = r(\gamma)$ como una función de parámetro desconocidos y $\hat{\tau} = q(Y)$ como un estimador de (γ), donde $q(Y)$ es la forma cuadrática del vector aleatorio de (Y). Si en $q(Y)$, $\hat{\tau}$ es insesgado y con varianza mínima, entonces $\hat{\tau}$ se define "El mejor estimador cuadrático insesgado" para todos los valores de $\hat{\tau}$ en Ω.

2.2 Métodos para Encontrar Estimadores

Como es bien sabido, existen diferentes métodos para estimar los valores de los parámetros de una función de distribución de probabilidad de manera metódica, tales como: Estimación de los Momentos, Estimadores de Bayes, Estimación Invariante y Estimación de Máxima Verosimilitud (Casella y Berger, 1990).

2.2.1.- Máxima Verosimilitud Clásica (ML)

Generalidades

Fisher (1922) citado en Yañez (2000), introduce su teoría de estimación de máxima verosimilitud, que consiste en encontrar los valores del parámetro que maximizan la probabilidad de observar los datos que realmente fueron observados. Por su parte, Hartleyy Rao, (1967), utilizaron la metodología para resolver problemas para datos desbalanceados, en presencia de la normalidad de los datos en modelos mixtos.

Por otro lado, Ghosh et al (1980), expresan que, en condiciones regularmente normal y con probabilidad tendiente a 1, la estimación de máxima verosimilitud es de 100(1-α) %, dentro de un intervalo de confianza de ($0 < \alpha < 1$) determinado por la familia de localidades para las pruebas robustas insesgadas de H_o: ($\theta = \theta_{0)}$ vs (H_1: $\theta \neq \theta_0$).

Sen (1996), estudia versiones de estimadores de máxima verosimilitud en pruebas preliminares y la reducción de estimadores de máxima verosimilitud en condiciones de normalidad asintótica que apuntan a reducir significativamente el riesgo del estimador.

Draper y Smith (1998), indican que si los errores son independientes y ε_i ~N $(0,\sigma^2)$, $b = (X'X)^{-1}X'Y$ es un estimador de máxima verosimilitud de β (en términos de vectores se puede escribir como ε ~ **N (0,** $I\sigma^2$), lo que significa que ε sigue una distribución normal multivariada n- dimensional con **E (** ε **) = 0** (donde **0** es un vector compuesto enteramente de ceros, de longitud ε) y $V(\varepsilon) = I\sigma^2$, donde, ε tiene una matriz de varianza covarianza con elementos en la diagonal, $V(\varepsilon_i),\ i = 1,2,...,n$ que son todas σ^2 y aquellos elementos fuera de la diagonal, las covarianzas $(\varepsilon_i,\varepsilon_j), i \neq j = 1,...,n.$ son todos ceros. La función verosimilitud para una muestra $Y_1, Y_2,.., Y_n$ está definida en este caso como el producto.

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} \exp^{\frac{-\varepsilon_i^2}{2\sigma^2}} = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp^{-\varepsilon'\varepsilon/2\sigma^2}.$$

Para valores fijos de σ , la función de verosimilitud maximizada es equivalente a minimizar la cantidad $\varepsilon'\varepsilon.$

En ese mismo sentido, Lerwick (1965), señala que el modelo clásico de regresión lineal se amplía para incluir el caso en que los residuos que dependen de la matriz de covarianza Σ = a S (1) dónde $S_{ij} = r^{|i-j|}$, |r| < 1. En un sistema de ecuaciones de estimadores de máxima verosimilitud para regresión, los coeficientes de a, y r se obtienen por un procedimiento iterativo desarrollado con el fin de resolver el sistema de ecuaciones generadas.

Harter y Moore (1968), indica que el uso y las relaciones entre la función exponencial y la de distribución asintótica de tipo II, permite obtener estimadores de máxima verosimilitud, a partir de muestras censuradas de forma separadas, de los parámetros de escala.

Brown y Rutemiller (1973), comparan dos estimadores puntuales de la fracción defectuosa de la distribución normal cuando dos parámetros, el estimador de mínima varianza insesgado F (x), y el estimador de máxima verosimilitud $\hat{F}(x)$.utilizando el criterio de mínimos cuadrados residuales, demuestra que la elección del estimador depende del verdadero valor de F(x) y el tamaño de la muestra. En el ámbito de .0005 $\leq$ F (x) $\leq$ 0.50, por lo que el estimador de máxima verosimilitud (M.L) es generalmente superior, incluso para los tamaños de muestras pequeños, a excepción de F(x) menos de alrededor de 0,01 o superior a 0,25.

Por otra parte, Gokhale (1973), propone un método iterativo para la obtención de estimaciones de máxima verosimilitud de distribuciones de probabilidades discretas con intervalo finito, bajo conjunto de restricciones lineales. Plantean diferentes hipótesis en una muestra y a diferentes muestras, de esta manera reduciendo los problemas.

Por otro lado, Olkin y Vaeth (1981), realizan análisis de varianza, en modelo de dos vías con errores correlacionados. Asumiendo que las mediciones obtenidas en cada fila tiene una matriz de covarianza general. Las estimaciones de máxima verosimilitud de los parámetros por filas y columnas, como también la matriz de covarianza general se obtienen y es analizada por su distribución asintótica. Finalmente, se utiliza la razón de verosimilitud entre las pruebas estadísticas para las hipótesis planteadas en un diseño de dos vías.

En otro orden de ideas, Grimshaw., (1993), indica que la distribución generalizada de Pareto (GPD) es una familia de dos parámetros de las distribuciones que se pueden utilizar para más de un modelo de superación del umbral. Los estimadores de máxima verosimilitud de los parámetros son preferibles, ya que son asintóticamente normal y asintóticamente eficientes en muchos casos. Los métodos numéricos son necesarios para maximizar el logaritmo de la verosimilitud.

Galán et al, (2003), expresan en su trabajo que los estimadores de los componentes de varianza y covarianza obtenidos por máxima verosimilitud tiene propiedades asintóticas, cuando el tamaño de la muestra es considerablemente grande. Son consistentes, normalmente asintóticos, eficientes y no requieren de conjunto de datos balanceados. Además, la matriz asintótica de varianza y covarianza de los estimadores es conocida, lo que permite el establecimiento de intervalos de confianza y prueba de hipótesis acerca de los componentes de varianza y covarianza.

Rodríguez., (2003), consideró en su trabajo el método de máxima verosimilitud en campos aleatorios gaussianos para realizar inferencia y predicción de procesos espaciales, al estudiar las propiedades de los estimadores puntuales y por intervalos, en términos de sesgo, varianza, error cuadrático medio, y la cobertura de los intervalos de confianza obtenidos por este método.

Por su parte, Villa. (2008), hace referencia a los métodos de máxima verosimilitud apoyados en el algoritmo de la función máximo verosímil, lo cual garantiza estimaciones insesgadas y solo valores positivos de la varianza. Además, requiere del supuesto de normalidad en los efectos aleatorios y del error. Si se presenta la no linealidad de las expresiones, es necesaria la utilización de algoritmos computacionales.

El Halimi. ,(2005) y Galan et al,. (2003), señalan que en las ecuaciones obtenidas de la estructura de la matriz de varianza-covarianza totales para tamaños de muestras pequeñas, los estimadores de máxima verosimilitud son habitualmente sesgados. Este se debe que al estimar los efectos aleatorios, no se toma en cuenta la perdida de grados de libertad asociada a las estimaciones de los efectos fijos, lo que conduce a subestimar la varianza del error (Searle et al, 1992).

2.2.2.- Máxima Verosimilitud Restringida (REML)

Generalidades

Inicialmente desarrollada por primera vez para ciertas situaciones de datos balanceados por Anderson y Bancroft (1952) y Ruseell y Bradley (1958),citados por Searle et al,. (1992).Para luego ser extendido por Thompson (1962) para datos balaceados en general, formalizada y extendida a diseños de bloques o datos desbalanceados con distribución normal por Patterson y Thompson (1971), (1974), generalmente para modelos mixtos, citados por Searle et al,. (1992).Consiste en factorizar la verosimilitud completa en dos partes independientes, una de las cuales no contiene la media, asumiendo que al usar esta parte de la verosimilitud no se pierde información con respecto al usar la verosimilitud completa, citada porLeón (2004).

Esto es importante cuando el número de parámetros fijos y el número de observaciones es de tamaño moderado o pequeño. La verosimilitud restringida, corresponde en realidad con la verosimilitud asociada a una combinación lineal de las observaciones, cuya media es nula y cumple las condiciones de ser un factor independiente del otro con el que se produce la verosimilitud completa y no suponer perdida de información con respecto al uso de los datos originales, citado por Galán et al,.(2003).

En este sentido, O` Neill., (2010), plantea que consiste en dividir la verosimilitud en dos componentes, el primer componente es una verosimilitud de uno o más estadísticos que involucran a todos los parámetros fijos como (μ) (y puede involucrar parámetro de varianza). El segundo componente es una verosimilitud restringida y que solo involucra los parámetros de la varianza de los efectos aleatorios, maximizando de esta manera cada componente de forma separada. Es por esto que las estimaciones de parámetro de varianza son conocidas como REML. Entonces, en presencia de una población normal, el primer componente resulta ser la verosimilitud de la media de la muestra y la segunda verosimilitud son para las variables asociadas con la varianza muestral.

$$\log L = \left[-\frac{\ln\left(\frac{2\pi\sigma^2}{n}\right)}{2} - \frac{1}{2}\left(\frac{\bar{y}-\mu}{\sigma/\sqrt{n}}\right)^2 \right] + \left[-\frac{n-1}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2}\ln(n) - \frac{1}{2}\sum\left(\frac{y_i-\bar{y}}{\sigma}\right)^2 \right]$$

Involucra μ (y es de poco interés σ)

Involucra σ μ (y es de poco interés μ)

Es de agregarque El Halimi. (2005), escribe en su trabajo que el contraste de estimación de máxima verosimilitud es la misma máxima verosimilitud restringida (REML). Consiste en maximizar la verosimilitud de un vector de combinaciones de $X\delta$. Suponga que L es la matriz que produce al vector de combinaciones lineales. Entonces, $Ly = LX\delta + LZ + Le$, es invariante a $X\delta$ si y sólo si $LX = 0$. Pero $LX = 0$ si y sólo si $L = TM$ para $M = I - X(X'X)^{-1}X'$ y alguna T. Para obtener combinaciones lineales que eliminan a los efectos fijos, L debe ser de rango completo. Harville (1977) llamó a los elementos L*y* "contrastes de error" porque tienen un valor esperado cero, E (Ly)= LE (Xδ +Zη +e)=L Xδ =0.

En este mismo orden de ideas,Galán et al,. (2003), realizaron la extensión a partir de modelo mixtos univariados para estimar componentes de varianza y covarianza, lo cual conduce a modelos multivariados, permitiendo estimar la máxima verosimilitud restringida de la covarianza entre características para los diferentes factores, en los modelos de diseños I y II de Carolina del Norte. Los estimadores de componentes de varianza y covarianza de característica múltiples, poseen todas las propiedades teóricas de los estimadores basados en la máxima verosimilitud, que son consistencia, eficiencia, normalidad asintótica y matriz de dispersión asintótica, contribuyendo a formar parte de las bases teóricas en implementar la estimación de REML de características múltiples en el mejoramiento genético.

Corbeil y Searle,. (1976), señala que los estimadores de (REML) poseen propiedad de invarianza bajo la versión de los estimadores de máxima verosimilitud (ML). Los estimadores REML tienen la propiedad adicional de reducir en el análisis de la varianza (ANOVA) muchos estimadores, si no casi en todos, en los casos de datos balanceados (igual número de observaciones). De igual manera que el método de (ML), requiere de un algoritmo de cálculo que fue desarrollado y adaptado a la transformación de Hemmerl y Hartley, lo que reduce los cálculos requeridos al tratar con matrices de igual orden de dimensión del espacio de parámetro, en lugar del espacio muestral. Estas matrices también producen varianzas asintóticas de estimaciones de muestreo.

Rodríguez. (2003), señala que el interés se centra en estudiar el método de máxima verosimilitud restringida para campos aleatorios gaussianos al realizar inferencia y predicción de procesos espacial, centrándose en las propiedades estadística del estimadores y en los intervalos de confianza que son poco conocidos, aplicando técnicas analíticas y de simulación, para investigar las propiedades de sesgo, varianza, error cuadrático medio y de los intervalos de confianza..

Leon et al,.(2009),consideraron una serie de trabajos para la evaluación genética en un centro genético porcino usando metodología BLUP. Se realizó una estimación de componentes de varianza de una población de cerdos. Se empleó para el análisis un modelo animal mixto multivariado, que incluía los efectos: año, época, sexo, edad final (como variable regresora) y animal; para los rasgos: Peso por edad (PPE) y Espesor de la grasa dorsal corregido a 100kg de peso vivo (EGD). El método de Máxima Verosimilitud Restringida (REML), con información promedio y matriz esparcida para análisis multivariado, del paquete AIREML fue el utilizado en el análisis de la información. Se concluye que estas estimaciones son lo suficientemente precisas y confiables para la predicción de valor genético con metodología BLUP.

2.2.3 Máxima Verosimilitud Ponderada (WML)

Generalidades

Wang. (2001), señala que el método de máxima verosimilitud ponderada "se basa en combinar todos los posibles datos de diferentes fuentes, cuando dos o más datos se obtienen de variables con similar medias en diferentes muestras, en cuanto a grado de precisión, sesgo y tamaños de muestras pequeñas, esto con el fin de mejorar la inferencia estadística". En este mismo orden de ideas, Plante (2007), indica que consiste en hacer inferencia sobre una población donde se ha obtenido una muestra, los datos adicionales están disponibles a partir de poblaciones, con distribución similar a la muestra inicial, pero no necesariamente idéntica. Permitiendo el uso de la información principales contenidas en los datos y su incorporación a la inferencia.

Wang. (2001)**,** Markatouy Lindsay (1997 y 1998), señalan que el modelo es fundamentado en ecuaciones de verosimilitud ponderada en el contenido de la estimación robusta, que se describe de la siguiente manera

Sea $(X_1, X_2, ..., X_n)$una muestra aleatoria con distribución de $f(x_i, \theta)$ se define la función de probabilidad ponderada como:

$$\sum_{i=1}^{n} \omega\left(x_i - \hat{F}\right) \frac{\partial}{\partial \theta} \log f\left(x_i; \theta\right).$$

Donde $(\hat{F})$ es la función de distribución empírica acumulada. La función de peso $\omega\left(X_i, \hat{F}\right)$, es seleccionada de tal manera que los valores estén muy cercano a 1, si no hay evidencia de violación del modelo en (x) desde la función de distribución empírica. La función de peso va a resultar cercano a cero (0) ò exactamente cero (0)en los (X_i). Si la función de probabilidad acumulada empírica indica falta de ajuste en cero (0) en(X_i).

Rao. (1991), introduce su definición de máxima verosimilitud ponderada para hacer frente a las irregularidades de los tiempos de observación en los estudios longitudinales de la tasa de crecimiento para ser más específicos. Se define la verosimilitud ponderada como:

$$L_n(\theta) = \prod_{i=1}^{K_n} f\left(x_i, n, t_i, n\theta / \{x_j, n, t_j n : 1 \leq j \leq i-1\}\right)^{\lambda(t_{in}) - \lambda(t_{i-1,n})}$$

Dónde $\lambda(.)$ es una función conocida no decreciente en $[a,b] t_{i,n}$ es el tiempo de observación que tal $a = t_{0,n} < t_{1,n} .... < t_{K_n,n} = b$.

El término de verosimilitud ponderada también fue utilizado por Warm. (1989) en el contexto de la teoría de respuesta de conjunto. Su propuesta describe como maximizan de $w(\theta)L(x/\theta)$ en lugar de la función de verosimilitud tradicional $L(x/\theta)$. La Función de peso, $w(\theta)$ ha sido cuidadosamente elegida de tal manera que el nuevo estimador sean imparcial a la Orden de n^{-1}.

Según Markatou. (2000), señala que los problemas relacionados con el análisis de los datos de distribución mixta incluyen la presencia de valores atípicos en la muestra. El hecho de que un componente no puede ser representado en los datos, además del problema de sesgo que se genera cuando el modelo no está descrito. Lo que indica el estudio de verosimilitud ponderado al brindar mejor desempeño bajo este contexto. El método produce estimaciones con sesgo y cuadrado medios del error bajos, una de sus utilidades es exponer datos de subestructuras en forma de raíces múltiples. Esto a su vez indica la viabilidad en el modelo mixto, debido a la presencia de más componentes de lo que se especifica originalmente en el modelo. Para calcular las estimaciones de verosimilitud ponderadas, se utiliza como punto de partida los valores del método de estimaciones de los momentos.

Wanget al,.(2004)**,** señalan que el método de verosimilitud ponderada fue presentado por Hu y Zidek (1995). Formalmente este abarca una variedad de procedimientos estadísticos para aumentar sesgo, mejorando la precisión de los estimadores. Su enfoque combina toda la información pertinente a través del uso de una versión de la función de verosimilitud ponderada, presentan las propiedades asintóticas de una clase de estimadores de máxima verosimilitud ponderada que contiene consideraciones hechas por Hu y Zidek. (2001).

Imoto., (2013), propone la utilización del método de máxima verosimilitud ponderada para determinar el hipocentro en sismología, cuando se está en presencia de datos contaminados por los errores sistemáticos, ya que se obtienen varianza de error muy próximo al valor asumido en la simulación.

2. 3 Modelos Lineales

Generalidades

Un Modelo Lineal consiste en expresar una variable Y como función de p variables $X_1, X_2,..., X_P$ de la forma:

$$Y = \beta_1 X_1 + \beta_2 X_2,..., \beta_P X_P + \varepsilon \qquad (1)$$

donde, Y es aleatoria debido a que ε, conocido como el termino de error, es considerado aleatorio, $X_1, X_2,..., X_P$ no aleatoriosy $\beta_1 X_1, \beta_2,..., \beta_P$ son parámetros desconocidos. Se pueden detallar estas consideraciones en la siguiente definición.

Definición 2.3.1
Considere las n ecuaciones:

$$Y_j = \sum_{j=1}^{n} X_{ij}\beta_j + \varepsilon_i \quad i = 1,2,..n \qquad (2)$$

$E(\varepsilon_i) = 0$

Donde;

1. Y_i son variables aleatorias observables;
2. X_{ij} ,$1 \leq j \leq p$ son variables no aleatorias, observables en un dominio d;
3. β_j son parámetros desconocidos definidos en un espacio Ω;
4. ε_i son variables aleatorias no observables tal que $E(\varepsilon_i) = 0$, $V(\varepsilon_i) = \sigma^2 I_d$ y $\text{cov}[\varepsilon_i, \varepsilon_j] = 0$;$i \neq j$, es decir, son independientes los errores.

Se pretende en este trabajo de investigación realizarla la estimación del vector de parámetros β_j, por vía de mínimos cuadrados ordinarios, máxima verosimilitud clásica, máxima verosimilitud restringida y máxima verosimilitud ponderada.

2.4.-Diseños Factoriales

Generalidades

Se refiere a la constitución de los tratamientos o combinación de tratamientos que tiene interés el investigador para someter a comparación. Para esto es necesario realizar replicas completas del experimento. Esto permite construir un diseño de tratamientos que refiere a la selección de factores donde se desean estudiar todas las combinaciones posibles de los niveles de los factores.

En necesario mencionar que el diseño de tratamiento es independiente del diseño experimental. Este último se refiere a la manera en que los tratamientos se disponen aleatoriamente en las unidades experimentales y la manera como se controla la variabilidad natural de las unidades experimentales. Es importante destacar que el diseño experimental puede ser conducido por medio de los diseños clásicos tales como; diseño completamente aleatorizado, diseño de bloque al azar y diseño cuadrado latino. Para cada uno de estos diseños, se puede tener un experimento factorial.

Clasificación de los diseños de tratamientos factoriales

- FACTORIALES
  - Factoriales Completos
    - Puros
      - 2^K
      - 3^k
      - $4^k,...,p^k$
    - Mixtos
      - p x q
      - p x q x r
      - p x q x r x s...
  - Factoriales Incompletos
    - Fraccionados.
    - Superficies Respuesta.
  - Pseudo factoriales

Los diseños factoriales producen experimentos más eficientes, pues cada observación proporciona información sobre todos los efectos de todos los factores sometidos a estudio y es factible ver las respuestas de un factor en diferentes niveles de otro factor en el mismo experimento. Kuehl (2003). Es por esto que el efecto de un factor puede considerarse como el cambio en la respuesta producido por un cambio en el nivel del factor. Con frecuencia se le llama efecto principal porque refiere a los factores de interés primarios en el experimento. Montgomery (2005).

Por otro lado en algunos experimentos puede encontrarse que la diferencia de respuesta entre los niveles de un factor no es la misma para todos los niveles de los otros factores. Cuando esto ocurre, existe una interacción entre los factores. Montgomery (2005).

En este sentido Montgomery y Runger. (2003) recomiendan efectuar primero la prueba de interacción y luego evaluar los efectos principales. Si la interacción no es significativa de las pruebas sobre los efectos principales es directa, ahora cuando la interacción es significativa, los efectos principales de los factores que participan en la interacción no tiene mucho valor interpretativo de manera práctica, por lo que se centra es en la interacción más que en los efectos principales

2.4.1.- Diseños Factoriales (pxq)

Generalidades

Dentro de los diseños factoriales se encuentran los diseños factoriales de dos factores (pxq). Se describirse un modelo lineal con interacción:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \qquad (1)$$

Sea y_{ijk} la respuestas observada cuando el factor τ tiene el nivelel i - esimo (i =1,2,..,a) y el factor β tiene el nivel del j -esimo (j =1,2,..,b) en la réplica k -esimo (k =1,2,..,n), donde μ es el efecto promedio global, τ_i es el efecto del nivel i - esimo del factor τ , β_j es el efecto del nivel j -esimo del factor β, $(\tau\beta)_{ij}$ es el efecto de la interacción entre τ_i y β_j y ε_{ijk} es un componente del error aleatorio.

Los modelos de los diseños de experimentos, según sean los factores, pueden clasificarse en modelos de efectos fijos, modelo de efecto aleatorios y modelos de efectos mixtos. Se considera un modelo de efecto fijo cuando el factor o los factores en consideración utilizan niveles que han sido seleccionados por el investigador. Es apropiado cuando el interés se centra en comparar el efecto sobre la respuesta de esos niveles específicos. Un modelo de efecto aleatorio es un factor o varios del que solo se incluyen en el experimento una muestra aleatoria simple de todos los posibles niveles del mismo.

Se utiliza este criterio cuando tienen un número muy grande de niveles y no es razonable usarlos todos ellos y en los modelos mixtos en un modelo en el que hay factores de efectos fijos y factores de efectos aleatorios

2.4.2.- Diseños Factoriales (pxq) en Modelos Fijos

Definición.

Tountenbur y Shalabh (2009) y Mongomery (2005), expresan el modelo de efecto fijo de la siguiente manera; suponga que (a) niveles del factor A y (b) niveles del factor B, para cada combinación (i, j), se realizan (n) replicas y el diseño es completamente aleatorizado. Por lo tanto, el número de observaciones viene dado por (nab) y el modelo de efecto fijo es:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \quad ,(1)$$

Donde

(y_{ijk}) Es la repuesta para el (i^{th}) nivel del factor A y el (j^{th}) nivel del factor B en la (k^{th}) replica.

(μ) Es la media general

(τ_i) Es el efecto del (i^{th}) nivel del factor de A

(β_j) Es el efecto del (j^{th}) nivel del factor de B

$(\tau\beta)_{ij}$ Es el efecto de la interacción de la combinación (i, j)

(ε_i) Es el componente del error aleatorio.

El siguiente supuesto $\varepsilon' = (\varepsilon_{111}, \ldots \varepsilon_{nab})$ entonces $\varepsilon \sim N(0, \sigma^2 I)$. (2)

En presencia de efectos fijos se tiene las siguientes consideraciones; para los efectos de tratamientos se define como las desviaciones de la media general, por lo tanto:

($\sum_{i=1}^{a} \tau_i = 0$) (3)

Luego $\sum_{j=1}^{b} \beta_j = 0$ (4)

de manera similar, los efecto de las interacciones son fijos, se define de tal modo

$\sum_{i=1}^{a} (\tau\beta)_{ij} = \sum_{j=1}^{b} (\tau\beta)_{ij} = 0$

(5)

ya que hay (n)réplicas del experimento, hay (nab) observaciones totales.

Se puede considerar el modelo de efecto fijo en el análisis de varianza, se considera representar de manera matricial como lo indica. Searle al et., (1992) lo expresa:

$$y = X\beta + \varepsilon \qquad (6)$$

Donde;

(y) Es un vector de datos ($Nx1$).

(β) En un vector de ($px1$) de parámetros de efecto fijo presentes en los datos.

(X) Es una matriz de coeficientes conocida (NxP)

(ε) Es un vector de errores definido como ($\varepsilon = y - E(y) = y - X\beta$) y así se tiene ($E(\varepsilon) = 0$), generalmente se atribuye la matriz de dispersión $Var(\varepsilon) = \sigma_{\varepsilon}^{2} I_N$.

Por lo general la matriz (X) es la matriz de incidencia o la matriz del modelo.

2.4.3.- Diseños Factoriales (pxq) en Modelos Mixtos

Definición.

Diversos experimentos se diseñan para estudiar los efectos de un factor sobre la media de la población y los efectos de otro sobre la varianza de la misma. Estos experimentos tienen una mezcla de factores fijos y aleatorios. Se conocen como modelos mixtos y se componen de dos partes; una sería el factor del efecto aleatorio donde se aplica a la variación en una población del total de niveles, mientras que la inferencias para los factores de efectos fijos se restringen a los niveles específicos usado por el investigador.

La estimación de los componentes de varianza en un modelo mixto, es de trascendental importancia para el investigador (Montgomery 2005). Es por esta razón que Galán et al,. (2003), señalan que el objetivo del análisis de los modelos mixtos es el estimar la varianza de cada uno de los modelos de efectos aleatorios así como su covarianzas, conocidas como componentes de varianzas (Corbeil y Searle, 1976; Harville, 1977).

El modelo lineal para experimentos factoriales (pxq) en un modelo mixto pueden ser especificado como $y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$ la situación τ_i es un factor de efecto fijo $\sum_i \tau_i = 0$ y si β_j es un efecto aleatorio entonces $\beta_j \sim NID(0, \sigma_\beta^2)$ con $\varepsilon \sim N(0, \sigma^2)$ .Se supone que la interacción $(\tau\beta)_{ij}$, es de efecto aleatorio, ya que cuando uno de los factores involucrado en el análisis es de efecto aleatorio se tiene $(\tau\beta)_{ij} \sim N(0, \sigma^2_{\tau\beta})$.

El modelo mixto en el análisis de varianza puede ser representado

$$y = X\alpha + U\beta + e$$

Para el modelo básico de y, un vector de n observaciones es especificado por Hartley y Rao (1967), Szatrowski y Miller (1980), Searle et al., (1992). En términos matriciales el modelo lineal mixto se presenta como:

$$y = X\beta + Zu + \varepsilon \qquad (1)$$

Donde y es el vector $nx1$ de observaciones.

β es un vector (px1) de parámetros desconocidos de efectos fijos.

X es una matriz, nxp de constante conocida como matriz incidencia de ceros y unos (matriz de diseño para efectos fijos).

(u) es un vector (qx1) de efectos aleatorios de la forma $u = [u'_1, u'_2, ..., u'_r]'$ particionado es (r) sub-vectores, además

$\mathrm{var}(u_i) = \sigma_i^2 I_{qi}$ para todo i, $\mathrm{cov}(u_i, u'_j) = 0$ para todo i $\neq$j $u \sim (0, \sigma_i^2 I_{qi})$

$i = 1,2,..,r$

q_i es el número de elementos de u_i seria el número de nivele de factor correspondiente u_i. La estructura de varianza de u es:

$$D = \mathrm{var}(u) = \begin{bmatrix} \sigma_1{}^2 I_{q1} & & \\ & \sigma_2{}^2 I_{q2} & \\ & & \sigma_r{}^2 I_r \end{bmatrix} = \{_d\, \sigma_i{}^2 I_{qi}\}$$

Z es una matriz de nxq de incidencia (matriz diseño asociada con efectos aleatorios e interacción), esta matriz se forma una partición de manera adecuada para el producto Zu_i de modo que $Z = [Z_1, Z_2, .., Z_r]$.Con Z_i de nxq_i dimensión.

Se tiene $$v = \mathrm{var}(y) = ZDZ' + \sigma_\varepsilon{}^2 I_N$$

ε es un vector $nx1$ de errores aleatorios que $\mathrm{E}(\varepsilon) = 0$ $\mathrm{var}(\varepsilon) = \sigma_\varepsilon^2 I_N$

3 Metodología

3.1 Método de Mínimos Cuadrados Ordinarios (MCO)

3.1.1Generalidades y Supuestos Básicos

Este método posiblemente sea el que se utiliza con mayor frecuencia para la estimación de parámetros. Consiste en estimar los β_i de una función lineal:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik} + \varepsilon_i \qquad (1)$$

Si x_{ij} denota nivel $i-esimo$ de la variable x_j, el interés se centra en minimizar la suma de los cuadrados de los errores entre el valor observado y_i y el valor estimado

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + ... + \hat{\beta}_k x_{ik} + \varepsilon_i \qquad (2).$$

Se puede denotar el error del modelo el valor observado y_i y el valor estimado $\hat{y}_i$:

$$\varepsilon_i = [y_i - \hat{y}_i] \qquad (3)$$

$$\varepsilon_i = \left[y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + ... + \hat{\beta}_k x_{ik}\right)\right] \qquad (4)$$

Para esto se necesita una función diferenciable tal que el error total de la estimación sea mínima posible, (Montgomery, 2005).La función de mínimos cuadrados es:

$$L = \sum_{i=1}^{n} \varepsilon_i^2 \,. \quad (5)$$

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{n=1}^{n} \left[y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + ... + \hat{\beta}_k x_{ik}\right)\right]^2 \quad (6)$$

Se deriva de manera parcial la función L con respecto a cada uno de los β_k parámetros y se iguala a cero (0), obteniendo $(k+1)$ ecuaciones normales.

Las ecuaciones normales se pueden resolver de manera más sencilla si se expresa de forma matricial $y = X\beta + \varepsilon$. Lo que se quiere es que el vector de los estimadores de mínimos cuadrados de la ecuación normal (4), es decir los $\hat{\beta}$, se obtiene minimizando la suma de los cuadraros de error. Esto se obtiene por la multiplicación de su transpuesta por sí mismo. A través de la función de mínimos cuadrados se denota:

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon'\varepsilon = (y - X\beta)'(y - X\beta) \qquad (7)$$

donde L se puede expresar $L = y'y - 2\beta'X'y + \beta'X'X\beta$ (8)

entonces, los estimadores de mínimos cuadrados que deben satisfacer:

$$\left.\frac{\partial L}{\partial \beta}\right|_{\hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0 \Rightarrow \hat{\beta} = (X'X)^{-1}(X'y) \qquad (9)$$

Por lo general, se requiere determinar una estimación de σ^2. El procedimiento establecido con antelación no permite tal estimación. Sin embargo, partiendo de la estimación basada en el estimador de mínimos cuadrados $\hat{\beta}$ de la ecuación (1), σ^2 es el mismo valor para cada y_i, expresando a través del siguiente teorema planteado por Rencher y Schaalje. (2007).

Teorema # 3. Si $E(\varepsilon) = 0$ y $\operatorname{cov}(\varepsilon) = \sigma^2 I$, para el modelo lineal $y = X\beta + \varepsilon$, donde el estimador insesgado σ^2 viene dado por $\hat{\sigma}^2 = \dfrac{y'y - \hat{\beta}'X'y}{n-k-1}$ con $n-k-1$ grados de libertad.

3.2 Método de Máxima Verosimilitud Clásica (ML).

3.2.1 Generalidades y Supuestos Básicos

Canavos (1998), señala que la esencia del método de máxima verosimilitud es seleccionar como estimador aquel valor del parámetro que tiene la propiedad de maximizar el valor de la probabilidad de la muestra aleatoria observada, es decir, consiste en encontrar el valor del parámetro que maximiza la función de verosimilitud.

La máxima verosimilitud fue introducido por Fisher (1922), intuitivamente pretende obtener el estimativo de un parámetro seleccionado aquel que maximiza la probabilidad de observar los datos que realmente fueron observados. Citado por Yáñez (2004).

Definición 1. Sea X una variable aleatoria discreta o continua cuya función de probabilidad $f(x)$ o densidad $f(x)$ respectivamente, depende solo de un parámetro θ. Supóngase que se efectúa n veces el experimento correspondiente, con la que se obtendrá una muestra de n números: $x_1, x_2, \ldots, x_n$.

Se supone independencia de los n ensayos, como antes, en el caso discreto, la probabilidad que una muestra de tamaño n conste precisamente de estos n valores está dada por el producto;

$$l = f(x_1) f(x_2) \ldots f(x_n). \quad (1)$$

En el caso continuo la probabilidad de que la muestra conste de valores en los pequeños intervalos

$x_1 \le x \le x_1 + \Delta x,\ x_2 \le x \le x_2 + \Delta x, ...,\ \ x_n \le x \le x_n + \Delta x$ está dada por la expresión

$$f(x_1)\Delta x\ f(x_2)\Delta x ... f(x_n)\Delta x. = l(\Delta x)^n. \qquad (2)$$

Los valores de $f(x_1),..., f(x_n)$ dependen del parámetro θ. Se tiene que l depende de $x_1, x_2,..., x_n$ y θ. Si los valores $x_1, x_2,..., x_n$, están dados y fijos, entonces l es una función de θ, que se llama función de verosimilitud. Kreyszig (1974).

Graybill (1972), indica que el caso más importante que considera es aquel en los que $x_1, x_2,..., x_n$ es una muestra aleatoria de densidad $f(x\ ;\theta)$, de modo que la función de probabilidad es:

$$L(\theta) = f(x_1;\theta)\ f(x_2;\theta),..., f(x_n;\theta) \qquad (3)$$

Muchas funciones de verosimilitud satisfacen las condiciones de regularidades, de manera que el estimador de máxima verosimilitud es solución de la ecuación

$$\frac{dL}{d\theta} = 0\ . \qquad (4)$$

Kreyszig (1974), plantea que la ecuación (4) que depende de $x_1, x_2,..., x_n$ se llama estimación de máxima verosimilitud para el parámetro de θ. En este sentido Graybill (1972), expresa que al aplicar logaritmo $L(\theta)$ y $\log[L(\theta)]$ o $\ln[L(\theta)]$ se logra encontrar el máximo de θ .

Si la función verosimilitud contiene k parámetros desconocidos $\theta_1, \theta_2,..., \theta_n$, en lugar de tener solo la ecuación (3) tendremos k ecuaciones decir, si

$$L(\theta_1, \theta_2, \ldots \theta_k) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \ldots, \theta_k) \qquad (5)$$

Entonces, los estimadores de máxima verosimilitud de los parámetros $\theta_1, \theta_2, \ldots, \theta_n$, son las variables aleatorias $\hat{\theta}_1 = d_1(x_1, \ldots, x_n)$; $\hat{\theta}_2 = d_2(x_1, \ldots, x_n)$;…; $\hat{\theta}_k = d_k(x_1, \ldots, x_n)$ donde $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ son evaluados en un espacio vectorial Ω que maximiza $L(\theta_1, \theta_2, \ldots \theta_k)$

Por lo que Kreyszig (1974), presenta las siguientes derivadas parciales

$$\frac{\partial L}{\partial \theta_1} = 0, \ldots\ldots \frac{\partial L}{\partial \theta_k} = 0, \qquad (6).$$

A partir de las cuales se pueden obtener estimaciones para los parámetros. Puesto que $f(x)$ es no negativa, un valor máximo de L será positivo. Como el logaritmo $\ln L$ es una función monotonicamente creciente L, esta tiene un máximo precisamente en el punto en la que L tiene un máximo, por lo que se puede usar $\ln L$ en lugar de L y reemplazar (4) por $\frac{\partial \ln(L)}{d\theta} = 0$.

Además, Graybill (1972), indica que si cumple con ciertas condiciones de regularidad, el punto donde la verosimilitud es un máximo allí es una solución de las k ecuaciones

Entonces la expresión (4) vendrá dada por $\frac{\partial \ln(\theta_1, \ldots, \theta_k)}{d\theta_1} = 0$

$$\frac{\partial \ln(\theta_1, \ldots, \theta_k)}{d\theta_2} = 0$$

$$\ldots\ldots\ldots\ldots$$

$$\frac{\partial \ln(\theta_1, \ldots, \theta_k)}{d\theta_k} = 0$$

Rencher y Schaalje,. (2007), consideran obtener estimadores de máxima verosimilitud, en presencia del supuesto de normalidad. Entonces la función de verosimilitud es la densidad de Y se describe por $L(\beta,\sigma^2)$, que se pueden buscar los valores desconocidos de (β)como a (σ^2) que maximiza $L(\beta,\sigma^2)$ para los valores de y y x_i evaluados en la muestra.

En el caso de la función densidad normal, es posible encontrar estimadores de máxima verosimilitud $\hat{\beta}$ y $\hat{\sigma}^2$por diferenciación. Debido a que la densidad normal involucra un producto y un exponencial, es más sencillo trabajar con $\ln L(\beta,\sigma^2)$, de esta manera alcanzando su máximo para los mismos valores de β y σ^2 como bien lo hace $L(\beta,\sigma^2)$. Los estimadores de máxima verosimilitud para β y σ^2 se obtiene en base al siguiente teorema.

Teorema # 1. Si Y es $N_n(X\beta,\sigma^2 I)$ donde X es $n\ x\ (k+1)$ de rango $(K+1)<n$, los estimadores de máxima verosimilitud de β y σ^2 son

$$\hat{\beta}=(X'X)' X'Y \tag{1}$$

$$\hat{\sigma}^2=\frac{1}{n}\left(Y-X\hat{\beta}\right)'\left(Y-X\hat{\beta}\right) \tag{2}$$

Se tiene

$$L(\beta,\sigma^2)= f(Y;\beta,\sigma^2)=\frac{1}{2\pi^{\frac{n}{2}}\left|\sigma^2 I\right|^{\frac{1}{2}}}\exp^{\frac{-(Y-X\beta)'\left(\sigma^2 I\right)^{-1}(Y-X\beta)}{2}}$$

$$L(\beta,\sigma^2)= f(Y;\beta,\sigma^2)=\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}}\exp^{\frac{-(Y-X\beta)'(Y-X\beta)}{2\sigma^2}} \tag{3}$$

Entonces $\ln L(\beta,\sigma^2)$ se convierte en;

$$\ln L(\beta,\sigma^2)=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta) \tag{4}$$

Tomando las derivadas parciales de $\ln L(\beta,\sigma^2)$ con respecto a β y σ^2, el resultado obtenido se iguala a cero, y se obtendrán las ecuaciones (1) y (2).

Propiedades de $\hat{\beta}$ y σ^2

Si se considera algunas propiedades de $\hat{\beta}$ y σ^2 bajo el modelo normal, las distribuciones de $\hat{\beta}$ y σ^2 son obtenidas en el siguiente teorema.

Teorema #2. Si Y es $N_n(X\beta,\sigma^2 I)$, donde X es $n \ x \ (k+1)$ de rango $(K+1)<n$, y $\beta=(\beta_0,\beta_1,....,\beta_k)'$, entonces los estimadores de máxima verosimilitud, $\hat{\beta}$ y $\hat{\sigma}^2$, dado en el teorema anterior tiene las siguientes propiedades distributivas:

(i) $\hat{\beta}$ es $N_{n+1}\left[\beta,\sigma^2(X'X)^{-1}\right]$.

(ii) $\frac{n\hat{\sigma}^2}{\sigma^2}$ es $\chi^2(n-k-1)$, o es equivalente $(n-k-1)\frac{s^2}{\sigma^2}$ es $\chi^2(n-k-1)$.

(iii) $\hat{\beta}$ y $\hat{\sigma}^2$ son independientes.

Otras propiedades de $\hat{\beta}$ y $\hat{\sigma}^2$ en condiciones normales son estadísticos suficientes.

Teorema #3. Si Y es $N_n(X\beta,\sigma^2 I)$, entonces $\hat{\beta}$ y $\hat{\sigma}^2$ son conjuntamente suficientes para β y σ^2.Si β y s^2 son suficientes, indica no hay otro estimador que pueda mejorar la información extraída de la muestra para estimar β y σ^2. Por lo tanto, no es de extrañar que $\hat{\beta}$ y s^2, sean estimadores insesgados de mínima varianza.

Teorema #4. Si Y es $N_n(X\beta,\sigma^2 I)$, entonces $\hat{\beta}$ y s^2 tiene mínima varianza entre todos los estimadores insesgados.

3.3 Método de Máxima Verosimilitud Restringida (REML)

3.3.1 Generalidades y Supuestos Básicos

Searle et al., (1992), señala en su texto que el REML fue desarrollado por primera vez para ciertas situaciones de datos balanceados por Anderson y Bancroft (1952) y Rusell y Bradley (1958), luego extendido por Thompson (1962) para datos balanceados de manera general, formalizada y extendido a diseño de bloques o datos desbalanceados por Patterson y Thompson (1971) generalizado a modelos de efectos mixtos con distribución normal.

En este sentido, Searle et al., (1992) señalan que los estimadores provenientes de máxima verosimilitud por lo general son sesgados. Esto es debido a que al estimar los efectos aleatorios no toma en cuenta los grados de libertad que se obtienen de la estimación de los efectos fijos. En este sentido, Galán et al,. (2003), señala que REML es una modificación de máxima verosimilitud, para obtener estimadores con menor sesgo. Para corregir esta situación, REML se basa en utilizar directamente el vector de observaciones de (y), es decir, se fundamenta en las combinaciones lineales de los elementos de las observaciones de (y), elegido de tal manera que las combinaciones no contienen los efectos fijos, sin importar cuál es su valor. Estas combinaciones lineales resultan equivalentes a los residuos obtenidos después de haber ajustado los efectos fijos, a partir de la siguiente ecuación:

$$y = X\beta + Zu + \varepsilon \qquad (1)$$

En este sentido Cadena y Castillo (2004) consideran que REML maximiza la verosimilitud de un vector de combinaciones lineales de las observaciones que son invariantes a $X\beta$.

Supóngase que K' es la matriz que produce al vector de combinaciones lineales a partir de un conjunto de valores $(K'y)$ donde los vectores de (K') se eligen de manera que de la ecuación (1) no contiene ningún término de β, se tiene:

$$K'y = K'X\beta + K'ZU + K'\varepsilon \qquad (2)$$

Es decir qué $(K'X\beta) = 0$ Para toda β (3)

Se tieneque $X\beta$ es invariante si y solo si ($K'X = 0$), conservando que $K'X = 0$ sí y solo si $K' = TM$, para M es una matriz de la forma $M = I - X(X'X)^{-}X'$ y alguna T, y así generar combinaciones lineales que no contengan efectos fijos, donde K' y T son de rango completo por filas de $(N-r)$.

Para K' en condiciones de normalidad la función de verosimilitud de $K'y$ es invariante a K' (Searle et al 1992). No importa que funciones de $K'y$ se utilice siempre que K' son de rango completo por filas de $(N-r)$.

En este sentido Harville (1977), citado por Searle el at (1992) llamo a los elemento $K'y$ para K' "contraste de error" porqué su valor esperado es cero entonces

$$E(K'y) = K'X\beta = 0$$

El número de contraste de error es linealmente independiente depende de X : con X de orden de (NxP) y de rango r, proveniente de la ecuación (3), esto satisface solo cuando se tiene $(N-r)$ valores linealmente independiente de K'.

Con las condiciones anteriores de K' las ecuaciones resultantes para estimar REML de σ^2 es:

$$\left\{_c\, tr(\dot{P}Z_i Z_i')\right\}_{i=0}^{r} = \left\{_c\, y'\dot{P}Z_i Z_i'\dot{P}y\right\}^{r}_{i=0} \text{, Con } i = 1,2,..r.$$

3.4 Método de Máxima Verosimilitud Ponderada (WML).

3.4.1 Generalidades y Supuestos Básicos

Wang (2001), señala que la verosimilitud ponderada fue desarrollada para reducir la varianza de los estimadores, esto se logra al reducir el cuadrado medio del error, mejorando la precisión de los estimadores, WML se basa en la combinación de toda la información apropiada a través de una versión ponderada de la función de verosimilitud, el interés se centra en analizar una sola población, un parámetro o vector de parámetro θ_1, es de interés inferencial.

La información de otras poblaciones, es decir una segunda población o tercera población y sucesivamente, están disponibles junto con la información directa de la población 1. Sea m el número total de poblaciones cuyas distribuciones se cree que es "similar" a la población 1.

Sea $n_1,..., n_m$ es el número de observaciones obtenidas de cada población antes mencionadas. Sea las observación variables aleatorias e independientes del vector $X_1, X_2,..X_m$ con funciones de densidad de probabilidad $f_1(.;\theta_1)$, $f_2(.;\theta_2)$, ..., $f_m(.;\theta_m)$, donde $X_i = \left(X_{i1}, X_{i2},.., X_{in_i}\right)^t$.

Se supone que, además, será solo la población con parámetro θ_1 es el vector de parámetro desconocido de interés inferencial, entonces la verosimilitud clásica seria:

$$L_1 = \left(x_1, \theta_1\right) = \prod_{j=1}^{n_1} f\left(x_{1j}; \theta_1\right)$$

Se supone que está interesado en la función de densidad de probabilidad $f_1(.;\theta_1):\theta_1 \in \Theta$ de una variable de estudio o vector de variables X, θ_1 es un parámetro o vector de parámetros desconocido. Al menos en cierto sentido cualitativo, $f_2(.;\theta_2),.....,f_m(.;\theta_m)$ se cree que son "similares" a $f_1(.;\theta_1)$.

Asumiendo que el resto de los parámetros tiene valores muy cercanoa θ_1 esto nos lleva a sugerir verosimilitud ponderada (WL), se define como:

$$WL(x;\theta) = \prod_{i=1}^{m}\prod_{j=1}^{n_i} f_i\left(x_{ij};\theta_1\right)^{\lambda_i}$$

Para valor fijo de $X = x$, donde $\lambda = (\lambda_1, \lambda_2, .., \lambda_{m_i})^t$ es el "vector ponderado", cuyo valor deben ser especificados previamente. Es importante aclarar que los parámetros $\theta_2, ..., \theta_{n_i}$ no aparecen en la verosimilitud ponderada definida con antelación, ya que el interés recae solo en θ_1 es el parámetro de la población 1. En lugar de ellos, la muestras generadas a partir de todas las poblaciones y se incorpora a la verosimilitud ponderada:
Se deduce que:

$$\log WL(x;\theta_1) = \sum_{I=1}^{m}\sum_{j=1}^{n_i}\lambda_i \log f_1(x_{ij},\theta_1).$$

Entonces se deduce que $\tilde{\theta}_1$ es un Estimador de Máxima Verosimilitud Ponderada (WML) de θ_1 si:

$$\tilde{\theta}_1 = \arg\sup_{\theta_1 \in \Theta} WL\left(x;\theta_1\right)$$

en muchos casos la (WME) puede obtenerse resolviendo la ecuación de estimación:

$$\left(\partial \big/ \partial\theta_1\right) \log WL(\theta_1) = 0$$

3.5 Criterios y Contrastes

A continuación se presenta un cuadro donde se observa la manera de contrastes a realizar a los métodos de estimación de parámetros descritos con antelación.

Cuadro # 1 Métodos de estimación, criterios y contraste a realizar.

Método de Estimación	**Criterios**	**Contrastar**
Mínimos Cuadrados Ordinarios. Máxima Verosimilitud Clásica. Máxima Verosimilitud Restringida. Máxima Verosimilitud Ponderada.	Datos balanceados Homogéneos. Datos desbalanceados Homogéneos. Datos balanceados Heterogéneo. Datos desbalanceados Heterogéneos. Estructura Matricial. Nivel de Significancia.	Tamaño de muestras para n=10 por tratamiento. Tamaño de muestras diferentes por tratamientos. Varianza-Covarianza.

3.6 El Modelo y las Hipótesis Planteadas

El modelo mixto pueden ser especificado como $y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$ en un diseño factorial con $k = 1,2,..,K$ $i = 1,2,..,I$ $j = 1,2,..J$ bajo un diseño clásico completamente aleatorizado en condiciones balaceadas o desbalanceadas, donde:

y_{ijk} es la k-esima observación en el i-esimo nivel del factor τ y el j-esimo nivel del factor β

τ_i es el i-esimo nivel del factor de efectos A.

β_j es el j-esimo nivel del factor de efectos B.

$(\tau\beta)_{ij}$ es la interacción del i- esimo factor A con el j-esimo factor B

ε_{ijk} es el error aleatorio.

Es necesario establecer los siguientes supuestos:

Modalidad I: ambos factores (A y B) son de efectos fijos:

$$\sum_{i=1}^{a} \tau_i = 0 \quad \sum_{j=1}^{b} \beta_j = 0 \, , \sum_{i=1}^{a}\sum_{j=1}^{b} (\tau\beta)_{ij} = 0 \, , \quad \varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$$

Modalidad II: el factor A es de efecto aleatorio y el factor B es efectos fijos:

$$\tau_i \sim N(0, \sigma_\tau^2) \, , \quad \sum_{j=1}^{b} \beta_j = 0 \, , \quad (\tau\beta)_{ij} \sim N(0, \sigma_{\tau\beta}^2) \, , \quad \varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$$

Por lo tanto, los componentes de varianza en cualquier observación vendrá dado por

$$V(y_{ijk}) = \sigma_\tau^2 + \sigma_{\tau\beta}^2 + \sigma_\varepsilon{}^2$$

3.7 Materiales y Métodos.

Se procedió a estructurar inicialmente una base de datos que incluyese cuatro (4) niveles del factor (A) y tres (3) del factor (B), con diez (10) repeticiones de observación en cada combinación de tratamientos, para obtener un total de 120 observaciones para una variable respuesta (y), donde se cumpliese el supuesto de homogeneidad de varianza en los doce (12) grupos de tratamientos balanceados homogéneos.

Tomando como referencia esa base de datos inicialmente estructurada, se realizaron cambios deliberados en los valores de la respuesta, de tal manera que se presentase heterogeneidad en las varianzas, generándose el archivo balanceado heterogéneo. Igualmente, y partiendo de la base de datos inicial, se procedió a la eliminación de observaciones dentro de los grupos, generando una base de datos desbalanceada presentando cumplimiento del supuesto de homogeneidad de varianza, obteniendo el archivo desbalanceado homogéneo. Finalmente, a partir de esta última base de datos se realizan cambios en la variable respuesta para generar heterogeneidad en la varianzas de los grupos y crear el archivo desbalanceado heterogéneo.

Estos conjuntos de datos fueron analizados por cuatro (4) metodologías de estimación de parámetros: Mínimos Cuadrados Ordinarios, Máxima Verosimilitud Clásica, Máxima Verosimilitud Restringida y Máxima Verosimilitud Ponderada. Es importante destacar que la Máxima Verosimilitud Ponderada se procedió a crear una variable de ponderación que permita equiparar las varianzas entre los grupos de tratamientos. Esta ponderación es obtenida de la relación entre la varianza menor y las otras varianzas, obteniendo una variable de ponderación igual o inferior a la unidad.

Adicional a las metodologías, se consideraron dos modalidades en los efectos incluidos en cada análisis. La primera modalidad consistió en considerar los factores A y B como modelo de efectos fijos, mientras en una segunda modalidad se consideró el factor A como modelo de efecto aleatorio. En ambos caso se incluye la interacción AxB. Los análisis estadísticos se realizaron utilizando los programas Harvey (1987) y SAS® (Freund littell 1981).

Con estas evaluaciones, se podrá realizar, una revisión técnica de la diferencia en las estimaciones por las diferentes metodologías y, con base a los resultados obtenidos, llevar a cabo la discusión de los resultados solo, se ahondará en el cuadrado medio residual. De esta manera será insumo para contrastar los diferentes métodos de estimación sometidos a estudio, determinar las magnitudes de ellas y establecer criterios que permitan producir recomendaciones prácticas para futuros análisis.

4 Análisis y Discusión de los Resultados.

Aplicación de los métodos de MCO, ML, REML, WMLE

El modelo $y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$ en un diseño factorial con $k = 1,2,..,120$ $i = 1,2,3,4$ $j = 1,2,3$ bajo un diseño clásico completamente aleatorizado en condiciones balaceados o desbalanceados donde:

y_{ijk} es la k-esima observación en el i-esimo nivel del factor τ y el j-esimo nivel del factor β

τ_i es el i-esimo nivel del factor de efectos A. $i = 1,2,3,4$

β_j es el j-esimo nivel del factor de efectosB. $j = 1,2,3$

$(\tau\beta)_{ij}$ es la interacción del i- esimo factor A con el j-esimo factor B

ε_{ijk} es el error aleatorio.

Y como se señaló con anterioridad, se tiene los siguientes supuestos:

Modalidad I: ambos factores (A y B) son modelos de efectos fijos:

$\sum_{i=1}^{4}\tau_i = 0$, $\sum_{j=1}^{3}\beta_j = 0$, $\sum_{i=1}^{4}\sum_{j=1}^{3}(\tau\beta)_{ij} = 0$, $\varepsilon_{ijk} \sim N(0,\sigma_\varepsilon^2)$

Modalidad II: el factor A ese modelo aleatorio y el factor B es fijo:

$\tau_i \sim N(0,\sigma_\tau^2)$, $\sum_{j=1}^{3}\beta_j = 0$, $(\tau\beta)_{ij} \sim N(0,\sigma_{\tau\beta}^2)$, , $\varepsilon_{ijk} \sim N(0,\sigma_\varepsilon^2)$

Por lo tanto, los componentes varianza en cualquier observación viene dado por

$$V(y_{ijk}) = \sigma_\tau^2 + \sigma_{\tau\beta}^2 + \sigma_\varepsilon{}^2$$

Consideraciones iníciales.

A continuación se presentan los cuadros # 1, 2, 3, 4 de los datos incluyendo criterios de cuatro (4) conjuntos de datos generados.

Cuadro #2: Datos balanceados y varianzas homogéneas.

A	B	Rep	Y		A	B	Rep	Y		A	B	Rep	Y		A	B	Rep	Y	
1	1	1	**20**		2	1	1	**15**		3	1	1	**11**		4	1	1	**4**	
1	1	2	**19**		2	1	2	**14**		3	1	2	**7**		4	1	2	**3**	
1	1	3	**15**		2	1	3	**11**		3	1	3	**10**		4	1	3	**3**	
1	1	4	**20**		2	1	4	**15**		3	1	4	**11**		4	1	4	**4**	
1	1	5	**20**		2	1	5	**15**		3	1	5	**11**		4	1	5	**1**	
1	1	6	**19**		2	1	6	**14**		3	1	6	**10**		4	1	6	**3**	
1	1	7	**24**		2	1	7	**15**		3	1	7	**11**		4	1	7	**4**	
1	1	8	**21**		2	1	8	**19**		3	1	8	**15**		4	1	8	**5**	
1	1	9	**21**		2	1	9	**16**		3	1	9	**12**		4	1	9	**8**	
1	1	10	**21**		2	1	10	**16**		3	1	10	**12**		4	1	10	**5**	
1	2	1	**22**		2	2	1	**19**		3	2	1	**10**		4	2	1	**5**	
1	2	2	**17**		2	2	2	**18**		3	2	2	**7**		4	2	2	**4**	
1	2	3	**21**		2	2	3	**18**		3	2	3	**9**		4	2	3	**2**	
1	2	4	**22**		2	2	4	**19**		3	2	4	**10**		4	2	4	**5**	
1	2	5	**22**		2	2	5	**19**		3	2	5	**10**		4	2	5	**5**	
1	2	6	**21**		2	2	6	**14**		3	2	6	**9**		4	2	6	**4**	
1	2	7	**22**		2	2	7	**19**		3	2	7	**10**		4	2	7	**5**	
1	2	8	**27**		2	2	8	**20**		3	2	8	**11**		4	2	8	**8**	
1	2	9	**23**		2	2	9	**24**		3	2	9	**11**		4	2	9	**6**	
1	2	10	**23**		2	2	10	**20**		3	2	10	**13**		4	2	10	**6**	
1	3	1	**23**		2	3	1	**14**		3	3	1	**10**		4	3	1	**9**	
1	3	2	**22**		2	3	2	**10**		3	3	2	**9**		4	3	2	**8**	
1	3	3	**18**		2	3	3	**13**		3	3	3	**5**		4	3	3	**5**	
1	3	4	**23**		2	3	4	**14**		3	3	4	**10**		4	3	4	**9**	
1	3	5	**27**		2	3	5	**14**		3	3	5	**10**		4	3	5	**9**	
1	3	6	**22**		2	3	6	**13**		3	3	6	**9**		4	3	6	**8**	
1	3	7	**23**		2	3	7	**14**		3	3	7	**14**		4	3	7	**9**	
1	3	8	**24**		2	3	8	**18**		3	3	8	**11**		4	3	8	**13**	
1	3	9	**24**		2	3	9	**15**		3	3	9	**11**		4	3	9	**10**	
1	3	10	**24**		2	3	10	**15**		3	3	10	**11**		4	3	10	**10**	

Cuadro # 3: Datos balanceados y varianzas heterogéneas.

A	B	Rep	Y		A	B	Rep	Y		A	B	Rep	Y		A	B	Rep	Y	
1	1	1	**20**		2	1	1	**15**		3	1	1	**11**		4	1	1	**4**	
1	1	2	**19**		2	1	2	**14**		3	1	2	**7**		4	1	2	**3**	
1	1	3	**15**		2	1	3	**5**		3	1	3	**10**		4	1	3	**3**	
1	1	4	**20**		2	1	4	**15**		3	1	4	**11**		4	1	4	**4**	
1	1	5	**20**		2	1	5	**15**		3	1	5	**11**		4	1	5	**1**	
1	1	6	**19**		2	1	6	**9**		3	1	6	**10**		4	1	6	**3**	
1	1	7	**24**		2	1	7	**20**		3	1	7	**11**		4	1	7	**4**	
1	1	8	**21**		2	1	8	**25**		3	1	8	**15**		4	1	8	**5**	
1	1	9	**21**		2	1	9	**16**		3	1	9	**12**		4	1	9	**8**	
1	1	10	**21**		2	1	10	**16**		3	1	10	**12**		4	1	10	**5**	
2	1	1	**15**		2	2	1	**19**		3	2	1	**10**		4	2	1	**5**	
2	1	2	**14**		2	2	2	**18**		3	2	2	**3**		4	2	2	**1**	
2	1	3	**5**		2	2	3	**18**		3	2	3	**4**		4	2	3	**2**	
2	1	4	**15**		2	2	4	**19**		3	2	4	**10**		4	2	4	**2**	
2	1	5	**15**		2	2	5	**19**		3	2	5	**10**		4	2	5	**5**	
2	1	6	**9**		2	2	6	**14**		3	2	6	**9**		4	2	6	**4**	
2	1	7	**20**		2	2	7	**19**		3	2	7	**10**		4	2	7	**5**	
2	1	8	**25**		2	2	8	**20**		3	2	8	**11**		4	2	8	**11**	
2	1	9	**16**		2	2	9	**24**		3	2	9	**16**		4	2	9	**9**	
2	1	10	**16**		2	2	10	**20**		3	2	10	**17**		4	2	10	**6**	
3	1	1	**11**		3	2	1	**10**		3	3	1	**10**		4	3	1	**9**	
3	1	2	**7**		3	2	2	**3**		3	3	2	**3**		4	3	2	**8**	
3	1	3	**10**		3	2	3	**4**		3	3	3	**5**		4	3	3	**5**	
3	1	4	**11**		3	2	4	**10**		3	3	4	**10**		4	3	4	**9**	
3	1	5	**11**		3	2	5	**10**		3	3	5	**10**		4	3	5	**9**	
3	1	6	**10**		3	2	6	**9**		3	3	6	**9**		4	3	6	**8**	
3	1	7	**11**		3	2	7	**10**		3	3	7	**20**		4	3	7	**9**	
3	1	8	**15**		3	2	8	**11**		3	3	8	**11**		4	3	8	**13**	
3	1	9	**12**		3	2	9	**16**		3	3	9	**11**		4	3	9	**10**	
3	1	10	**12**		3	2	10	**17**		3	3	10	**11**		4	3	10	**10**	

Cuadro # 4: Datos desbalanceados y varianzas homogéneas.

A	B	Rep	Y		A	B	Rep	Y		A	B	Rep	Y		A	B	Rep	Y
1	1	1	**20**		2	1	1	**15**		3	1	1	**11**		4	1	5	**1**
1	1	2	**19**		2	1	2	**14**		3	1	2	**7**		4	1	6	**3**
1	1	3	**15**		2	1	3	**11**		3	1	3	**10**		4	1	7	**4**
1	1	4	**20**		2	1	7	**15**		3	1	5	**11**		4	1	8	**5**
1	1	6	**19**		2	1	8	**19**		3	1	6	**10**		4	1	9	**8**
1	1	7	**24**		2	1	9	**16**		3	1	7	**11**		4	1	10	**5**
1	1	8	**21**		2	1	10	**16**		3	1	8	**15**					
1	1	9	**21**							3	1	9	**12**					
1	1	10	**21**							3	1	10	**12**					
1	2	1	**22**		2	2	1	**19**		3	2	1	**10**		4	2	1	**5**
1	2	2	**17**		2	2	2	**18**		3	2	2	**7**		4	2	2	**4**
1	2	3	**21**		2	2	3	**18**		3	2	3	**9**		4	2	3	**2**
1	2	4	**22**		2	2	6	**14**		3	2	4	**10**		4	2	7	**5**
1	2	5	**22**		2	2	7	**19**		3	2	5	**10**		4	2	8	**8**
1	2	6	**21**		2	2	8	**20**		3	2	9	**11**		4	2	9	**6**
1	2	7	**22**		2	2	9	**24**		3	2	10	**13**		4	2	10	**6**
1	2	8	**27**		2	2	10	**20**										
1	2	9	**23**															
1	2	10	**23**															
1	3	1	**23**		2	3	1	**14**		3	3	1	**10**		4	3	1	**9**
1	3	2	**22**		2	3	2	**10**		3	3	2	**9**		4	3	2	**8**
1	3	3	**18**		2	3	3	**13**		3	3	3	**5**		4	3	3	**5**
1	3	4	**23**		2	3	6	**13**		3	3	4	**10**		4	3	4	**9**
1	3	5	**27**		2	3	7	**14**		3	3	5	**10**		4	3	6	**8**
1	3	6	**22**		2	3	8	**18**		3	3	6	**9**		4	3	7	**9**
1	3	7	**23**		2	3	9	**15**		3	3	7	**14**		4	3	8	**13**
1	3	8	**24**		2	3	10	**15**		3	3	8	**11**		4	3	9	**10**
1	3	9	**24**							3	3	9	**11**		4	3	10	**10**
1	3	10	**24**							3	3	10	**11**					

Cuadro # 5: Datos desbalanceados y varianzas heterogéneas.

A	B	Rep	Y	A	B	Rep	Y	A	B	Rep	Y	A	B	Rep	Y
1	1	1	**20**	2	1	1	**15**	3	1	1	**11**	4	1	1	**4**
1	1	2	**19**	2	1	3	**5**	3	1	2	**7**	4	1	2	**3**
1	1	3	**15**	2	1	6	**9**	3	1	3	**10**	4	1	5	**1**
1	1	4	**20**	2	1	7	**20**	3	1	5	**11**	4	1	6	**3**
1	1	6	**19**	2	1	8	**25**	3	1	6	**10**	4	1	9	**8**
1	1	7	**24**	2	1	9	**16**	3	1	7	**11**	4	1	10	**5**
1	1	8	**21**	2	1	10	**16**	3	1	8	**15**				
1	1	9	**21**					3	1	9	**12**				
1	1	10	**21**					3	1	10	**12**				
1	2	1	**22**	2	2	1	**19**	3	2	2	**3**	4	2	2	**1**
1	2	2	**17**	2	2	2	**18**	3	2	3	**4**	4	2	3	**2**
1	2	3	**21**	2	2	5	**19**	3	2	4	**10**	4	2	4	**2**
1	2	4	**22**	2	2	6	**14**	3	2	7	**10**	4	2	7	**5**
1	2	5	**22**	2	2	7	**19**	3	2	8	**11**	4	2	8	**11**
1	2	6	**21**	2	2	8	**20**	3	2	9	**16**	4	2	9	**9**
1	2	7	**22**	2	2	9	**24**	3	2	10	**17**	4	2	10	**6**
1	2	8	**27**	2	2	10	**20**								
1	2	9	**23**												
1	2	10	**23**												
1	3	1	**23**	2	3	1	**14**	3	3	1	**10**	4	3	1	**9**
1	3	2	**22**	2	3	2	**10**	3	3	2	**3**	4	3	2	**8**
1	3	3	**18**	2	3	3	**13**	3	3	3	**5**	4	3	3	**5**
1	3	4	**23**	2	3	6	**13**	3	3	4	**10**	4	3	4	**9**
1	3	5	**27**	2	3	7	**14**	3	3	5	**10**	4	3	6	**8**
1	3	6	**22**	2	3	8	**18**	3	3	6	**9**	4	3	7	**9**
1	3	7	**23**	2	3	9	**15**	3	3	7	**20**	4	3	8	**13**
1	3	8	**24**	2	3	10	**15**	3	3	8	**11**	4	3	9	**10**
1	3	9	**24**					3	3	9	**11**	4	3	10	**10**
1	3	10	**24**					3	3	10	**11**				

Cuadro # 6: Varianza para grupos de datos.

Grupos de Datos				
Combinación Tratamientos	Balanceados		Desbalanceados	
	Homogéneas	Heterogéneas	Homogéneas	Heterogéneas
11	5.11	5.11	5.75	5.75
12	6.00	6.00	6.00	6.00
13	5.11	29.33	5.11	5.11
21	4.00	4.00	5.81	43.81
22	6.00	6.00	7.71	7.55
23	4.00	4.00	5.14	5.14
31	4.00	4.00	4.50	4.50
32	2.44	19.11	3.33	28.48
33	5.11	19.78	5.11	19.78
41	3.33	3.33	5.47	5.60
42	2.44	9.78	3.48	14.48
43	4.00	4.00	4.50	4.50
$\chi 2$ (Barlett)	4.14	29.20	1.81	26.60
p aceptar (Ho)	.9656	.0021	.9990	.0052

Cuadro # 7: Grupos de datos, ponderación de varianzas.

Grupos de Datos								
Combinación Tratamientos	Balanceados				Desbalanceados			
	homogéneas	ponderación	heterogéneas	ponderación	homogéneas	ponderación	heterogéneas	ponderación
11	5.11	0,48	5.11	0,65	5.75	0,58	5.75	0,78
12	6.00	0,41	6.00	0,56	6.00	0,56	6.00	0,75
13	5.11	0,48	29.33	0,11	5.11	0,65	5.11	0,88
21	4.00	0,61	4.00	0,83	5.81	0,57	43.81	0,10
22	6.00	0,41	6.00	0,56	7.71	0,43	7.55	0,60
23	4.00	0,61	4.00	0,83	5.14	0,65	5.14	0,88
31	4.00	0,61	4.00	0,83	4.50	0,74	4.50	1
32	2.44	1	19.11	0,17	3.33	1	28.48	0,16
33	5.11	0,48	19.78	0,17	5.11	0,65	19.78	0,23
41	3.33	0,74	3.33	1	5.47	0,61	5.60	0,80
42	2.44	1	9.78	0,34	3.48	0,96	14.48	0,31
43	4.00	0,61	4.00	0,83	4.50	0,74	4.50	1

Análisis de los Resultados

A partir de estos momentos se realizará discusiones que permita recomendar la utilidad del ò los métodos de estimación de parámetros, sometidos a estudio tomando las siguientes consideraciones se condujo en un diseño factorial a diferentes niveles para el factor (A) está presente en cuatro (4) niveles y el factor (B) a tres (3) niveles, se llevó a cabo en diseño clásico conocido como diseño completamente aleatorizado, bajo situación de datos balanceados o desbalanceados con presencia de varianzas homogéneas o heterogéneas y tamaño de muestra $n = 10$ por tratamientos. Se utilizará el cuadrado medio del residual ya que se considera elemento importante para contrastar los estimadores obtenidos en base de la propiedad de eficiencia.

En los siguientes cuadros se presentan los resultados obtenidos en condiciones mencionadas con antelación para la obtención de la eficiencia de los MCO, ML, REML, WML.

Casos de Efectos Fijos

Cuadro # 8: Modelo de efectos fijos para datos balanceados con varianzas homogéneas.

	CMO	**ML**	**REML**	**WML**
CM residual	**4,2963**	**3,8667**	**4,2963**	**2,2047**
F Efecto A	324,05	360,06	324,05	359,71
F Efecto B	6,98	7,76	6,98	7,94
F Interacción	10,6	11,78	10,6	11,82
- 2 Log	---	502,8	491,6	498

En el caso del modelo de efectos fijos en condición de datos balanceados y con varianzas homogéneas, se puede destacar, que el estimador por WML muestra un menor estimador del menor cuadrado medio residual. Este aspecto está documentado en Wang (2001), Wang et al (2004), Markatau y Lindsay (1997, 1998) y Markatau (2000), De acuerdo con los resultados de este trabajo, la estimación ML tiene el segundo valor menor el estimador del cuadrado medio residual frente a los estimaciones CMO y REML que tienen el mismo valor de estimador del cuadrado medio residual.

Cuadro # 9: Modelo de efectos fijos para datos balanceados con varianzas heterogéneas.

	CMO	**ML**	**REML**	**WML**
CM residual	**9,6296**	**8,6667**	**9,6296**	**4,3635**
F Efecto A	144,58	160,64	144,58	125,31
F Efecto B	3,12	3,46	3,12	3,39
F Interacción	4,73	5,26	4,73	6,31
- 2 Log	---	599,7	578,7	609

Para el caso del modelo de efectos fijos en situación de datos balanceados y el impacto que toleran debido a las condición de violación del supuesto de homogeneidad, note como los métodos de estimaciones CMO y REML evidencia los valores mayores, además de la similitud del estimador del cuadrado medio residual, en relación al método de estimación ML, tiene el segundo valor mayor del estimador del cuadrado medio residual, estadísticamente no sean los más eficientes. Por otro lado el método de estimación WML muestra un menor valor, estadísticamente sea el método de estimación que produce estimador más eficiente al tener el estimador de varianza mínima.

Cuadro # 10: Modelo de efectos fijos para datos desbalanceados con varianzas homogéneas.

	CMO	ML	REML	WML
CM residual	**5,1937**	**4,5705**	**5,1937**	**2,9318**
F Efecto A	221,63	251,85	221,63	258,53
F Efecto B	3,92	4,46	3,92	4,38
F Interacción	6,93	7,87	6,93	7,77
- 2 Log	**---**	435,7	420	433,5

Como se evidencia con el modelo de efectos fijos para los datos desbalanceados y el cumplimiento de supuesto de homogeneidad, el método de estimación por WML presenta menor magnitud en cuanto al estimador del cuadrado medio residual indicando ser un método de estimación que produce un estimador eficiente a pesar del desbalance de los datos, por su parte las estimaciones por CMO y REML tienen los valores mayor, además de la similitud de magnitud en cuanto al estimador del cuadrado medio residual, esto estadísticamente afecta la eficiencia de los estimadores, en este mismo sentido el método de estimación por ML produce un valor que, medianamente afecta la eficiencia del método.

Cuadro # 11: Modelo de efectos fijos para datos desbalanceados con varianzas heterogéneas.

	CMO	ML	REML	WML
CM residual	**11,7437**	**10,3345**	**11,7437**	**3,9615**
F Efecto A	98,94	112,4	98,94	174,57
F Efecto B	2,04	2,31	2,04	1,87
F Interacción	3,18	3,62	3,18	6,07
- 2 Log	---	517,3	491,8	486,1

Como se muestra en el cuadro anterior para el caso del modelo de efectos fijos, desbalanceados y la violación del supuesto de homogeneidad, como los métodos de estimaciones por MCO y REML reaccionan de la misma manera al presentar un valor considerablemente elevados en el estimador del cuadrado medio residual son sensibles condiciones a que son sometidos, de esta manera el estimador ML se presenta un valor moderado en la estimador del cuadrado medio residual, esto indica estadísticamente que los métodos de estimación no producen estimadores de varianzas mínima afecto la eficiencia del estimador, otro lado el método de estimación de WML presenta un valor consideradamente bajo en la estimación de los cuadrados medios residual, no se ve afectada por las condiciones del modelo descrito con anterioridad produciendo estimadores más eficientes.

Cuadro # 12: Resumen del estimador de cuadrado medio residual por el método de estimación WML, en modelo de efectos fijos para datos balanceados y desbalanceados con varianzas homogéneas y varianzas heterogéneas

Balanceados		Desbalanceados	
varianzas homogéneas	varianzas heterogéneas	varianzas homogéneas	varianzas heterogéneas
2,2047	**4,3635**	**2,9318**	**3,9318**

El cuadro # 12 muestra el desempeño del método de estimación por WML en modelo de efectos fijos para datos balanceados en presencia del no cumplimiento del supuesto de homogeneidades sensiblemente afectado el estimador del cuadrado medio residual cuyo valor es superior, en contraste con los datos balanceado y reúne el cumplimiento del supuesto de homogeneidad, el estimador del cuadrado medio residual es menos afectada bajo estas condiciones el estimador es más eficiente. Por otro lado en modelo de efectos fijos para datos desbalanceados en condición de violación del supuesto de homogeneidad es afectado el estimador del cuadrado medio residual, con relación a datos desbalanceado y en presencia del cumplimiento del supuesto de homogeneidad es menos afectada la eficiencia de estimador.

Casos de Efectos Mixtos

Cuadro # 13: Modelo de efectos mixtos para datos balanceados con varianzas homogéneas

	CMO	ML	REML	WML
CM residual	**4,2963**	**4,2963**	**4,2963**	**2,4498**
F Efecto A	324,05			
F Efecto B	0,66	0,88	0,66	0,85
F Interacción	10,6			
σ^2 A	44,8889	33,6667	44,8889	33,7272
σ^2 AB	4,1259	2,987	4,1259	2,98
- 2 Log	--	554	545,2	549,7

Los cuatro métodos de estimación de parámetros con modelo de efectos mixtos para datos balanceados con varianzas homogénea, el método de estimación que muestra valor menor del estimador del cuadrados medios residual es WML en base al modelo planteado generando un estimador más eficiente, este criterio está documentado en Markatou (2000), contrario los métodos de estimaciones MCO, ML, REML tienen el mismo valor de los estimadores de cuadrados medios residuales en base al modelo planteado, de acuerdo a este resultados obtenidos, no se ha logrado encontrar nada de en las referencias consultadas sobre tal efectos.

Cuadro # 14: Modelo de efectos mixtos para datos balanceados con varianzas heterogéneas.

	CMO	ML	REML	WML
CM residual	**9,6296**	**9,6296**	**9,6296**	**4,8233**
F Efecto A	144,58	---	---	---
F Efecto B	0,66	0,88	0,66	0,9
F Interacción	4,73	---	---	---
σ^2 A	44,8889	33,6667	44,8889	31,7641
σ^2 AB	3,5926	2,4537	3,5926	2,7937
- 2 Log	---	641,2	632,3	650,3

Los cuatro métodos de estimación de parámetros con modelo de efectos mixtos para datos balanceados con incumplimiento del supuesto homogeneidad, en base a los resultados obtenidos y tomando en cuenta el menor valor del estimador cuadrado medio residual, el método de estimación más conveniente es WML, en contraposición los métodos de estimación CMO, ML y REML presentan mayores valores en los estimadores de los cuadros medios residuales, además poseen los mismos los estimadores de cuadrados medios residuales, sin embargo no se tiene nada dicho en la literatura examinadas.

Cuadro # 15: Modelo de efectos mixtos para datos desbalanceados con varianzas homogéneas.

	CMO	**ML**	**REML**	**WML**
CM residual	**5,1937**	**5,1947**	**5,1941**	**3,3348**
F Efecto A	218,81			
F Efecto B	0,56	0,78	0,59	0,72
F Interacción	6,93			
σ^2 A	44,2793	33,0054	44,0632	33,0238
σ^2 AB	3,739	2,6434	3,7405	2,5961
- 2 Log	---	482,1	473,3	480,1

El estimador por WML en modelo de efectos mixtos con datos desbalanceados y el cumplimiento del supuesto de homogeneidad el estimador del cuadrados medios residuales es sensiblemente menor, por lo que bajo estas condiciones a que es sometido es más eficiente ante los otros métodos de estimación, además estos se comportan de la misma maneras afectando el estimador del cuadros medios residuales.

Cuadro # 16: Modelo de efectos mixtos para datos desbalanceados con varianzas heterogéneas.

	CMO	ML	REML	WML
CM residual	**11,7437**	**11,7495**	**11,7461**	**4,4655**
F Efecto A	97,51	---	---	---
F Efecto B	0,62	0,87	0,65	0,94
F Interacción	3,18	---	---	---
σ^2 A	44,5795	33,1193	44,2899	32,0061
σ^2 AB	3,1124	1,96	3,1025	2,8679
- 2 Log	---	554,2	545,4	527,3

Se puede notar que los métodos de estimación MCO, ML y REML sus cuadrados medios residuales son diferentes bajo estas condiciones; la utilización de estos métodos no siempre se obtienen los efectos esperados. Caso contrario sucede con WML a pesar de las mismas condiciones de datos desbalanceados heterogéneo con efectos mixtos el estimador de los cuadrados medios residuales es menor.

Cuadro # 17: Resumen del estimador de cuadrado medio residual por el método de estimación WML en presencia de datos balanceados y desbalanceados con varianzas homogéneas y varianzas heterogéneas en modelo de efectos mixtos

Balanceados		**Desbalanceados**	
Varianzas Homogéneas	**Varianzas Heterogéneas**	**Varianzas Homogéneas**	**Varianzas Heterogéneas**
2,4498	4,8233	3,3348	4,4655

Como se puede notar en el cuadro # 17, en términos de eficiencia bajo la influencia de datos balanceados en modelos mixtos, con cumplimiento o no cumplimiento del supuesto de homogeneidad, se presenta el método de estimación WML vulnerado, ante la violación del supuesto de homogeneidad, evidenciando que el método de estimación WML es más eficiente bajo el cumplimiento del supuesto de homogeneidad. Por otro lado en condiciones de datos desbalanceados en modelos mixtos, con cumplimiento o no cumplimiento del supuesto de homogeneidad, el método de estimación WML sigue estando afectado por el no cumplimiento del supuesto de homogeneidad, es claro que el método de estimación WML es eficiente cuando se cumple el supuesto de homogeneidad.

5. Conclusiones y Recomendaciones.

Para los ejemplos aquí presente, se pudo constatar en cuanto al método de estimación WML en presencia de datos balanceados o desbalanceados con varianzas homogéneas o heterogéneas en modelos de efectos fijos y mixtos presento mejor comportamiento en términos de la propiedad de eficiencia al tener menor estimador de cuadrados medios residuales para diseños factoriales (pxq), en contraste con los métodos de estimación MCO, ML y REML fueron sensible a estas condiciones a que fueron sometidos.

No se pretender generalizar en este trabajo, que los métodos de estimación MCO, ML, REML no funcionan bajo estas condiciones; sino que su utilización no siempre se obtiene los efectos esperados.

Ante los hechos, se recomienda el uso del método de máxima verosimilitud ponderada (WML), ya que mostró ser más eficiente al obtener menores cuadrados medios residuales.

En virtud de los diferentes casos planteados, se sugiere continuar trabajando en línea de investigación, que permita aportar mayor información en la experimentación agroindustrial y evalúen alternativas de eficiencias de métodos.

6. Bibliografía

Brown, G and Rutemiller, H., (1973). **The Efficiencies of Maximum Likelihood and Minimum Variance unbiased Estimators of Fraction Defective in the Normal Case**. Thechnometrics vol. 15 No 4.

Cadena, J y Castillo, A., (2004). **Comparación de Diferentes Métodos para la Estimación de Componentes de Varianzas**.
Disponible: www.http://www.redalyc.org/pdf/302/30236609.pdf

Canavos, G,. (1998). **Probabilidad y Estadística. Aplicaciones y Métodos.** Primera Edición, Editorial McGraw-Hill/Interamericana de México, S.A

Casella, G. and Berger.R., (1990).**Statistical Inference.** Wadsworth & Brooks/GoleAdvacedBook & Sofware Pacific Grove, California.

Corbeil, R. and Searle, R,. (1976). **Restricted Maximum Likelihood (REML) Estimation of Variance Components in the Mixed Model.** Technometrics. Vol. 18, No 1 pp.31-37.

Draper, N. and Smith, H., (1998).**Applied Regression Analysis.**3era EditionAwiley-Interscience Publication. JONH WILEY & SONS. INC, pp. 136-137.

El Halimi, R., (2005).**Nonlinear Mixed-effectsModels and Nonparametric Inference.** Memoria Presentada per a Optar al grau de Doctor en Matemàtiques Departament d'Estadística de la Universitat de Barcelona.

Freund, R y Littell, R., (1981). **S.A.S For Linear Models**. A Guide to the Anova and GLM Procedures. Editorial S.A.S Institute. INC pp. 231

Galán, J, Jiménez. P y Cervantes. C., (2003). **Estimación por Máxima Verosimilitud Restringida de Componente de Varianza de Múltiples Características Bajo los Diseños I y II de Carolina del Norte.** Rev. Fitotecnia. México .Vol. 26 (1): 53-66.

Gallego, J., (2008). **Econometría**. Universidad de Cantabria.

Gokhale, D., (1973).**Iterative Maximum Likelihood Estimation for Discrete Distributions.** Sankhya: The Indian Journal Of Statistics, Volumen 35, serie B, Pts 3, pp. 293-298.

Ghosh.J, Sinha B. and Joshi, S., (1980).**A Property of Maximum Likelihood Estimator**. Sankhya: The Indian Journal Of Statistics, Volume 42, series B, Pts 3 and 4, pp. 143-152.

Graybill, F., (1972).**Theory and Application of Linear Model**. Duxbury Press, Massachusetts.

Grimshaw, S., (1993).**Computing Maximum Likelihood Estimates for the Generalized Pareto Distribution.** Technometrics, Vol 35, N° 2.

Hartley and Rao, J., (1967).**Maximum Likelihood Estimation for the Mixed Analysis of Variance Model.** Biometrika, 54, 1 and 2, pp. 93-108.

Harter, L. and Moore, A.,(1968).**Conditional Maximum- Likelihood Estimator, from Singly Censored Samples, of the Scale Parameters of Type II Extreme-Value Distributions.** Technometrics, Vol. 10, N° 2.

Harville, D., (1977).**Maximum Likelihood Approaches to Variance component Estimation and to Related Problems.** J. Amer. Statist. Ass.73, 320-338.
Harvey, W., (1987). User`s Guide for. LSMLMW. PC-1 Version. pp. 59

Hu, F., Zidek, J., (1995). **Incorporating Relevant Sample Information Using the Likelihood.** Technical Report No. 161, Department of Statistics, The University of British Columbia, Vancouver, BC, Canada.

Hu, F. and Zidek, J., (2001).**The Relevance Weighted Likelihood with Applications**. In: Ahmed, S.E., Reid, N. (Eds.), Empirical Bayes and Likelihood Inference. Springer, New York, pp. 211–235.

Imoto, M., (2013). **Application of a Weighted Likelihood Method to Hypocenter Determination.** National Research Institute for Earth Science and Disaster Prevention 3-1 ten-nodri, Tskuba, Ibakari 305-306,Japòn.

Kuelhl, R. (2003). **Diseño de Experimentos: Principios Estadísticos para el Diseño y Análisis de la Investigación**. Editorial Thomson Learning. 2da Edición.

Kreyszig, E., (1974). **Introducción a la Estadística Matemática. Principios y Métodos**. Editorial Limusa.

Lacourly, N,. (2008). **Métodos Estadísticos Predictivo.** Facultad de Ciencias Físicas y Matemáticas. Departamento de matemáticas, Universidad de Chile.
www.ucursos.cl/ingeneria/2008/1/MA44D/1/material_docente/bajar%Fid_material%3D158267

León, E. (2004). **Métodos de Estimación de Componente de Varianza en Poblaciones**. Una Reseña Histórica. Revista computarizada del Instituto de Investigaciones porcinas. Vol: 11 Nº1- 2004, gaveta postal, punta Brava La Habana Cuba.

León. E, Guerra. D, Santana. I, Dieguez. F, Brache. F., (2009).**Estimación de Parámetros Genéticos vía REML en un Modelo Animal.**
Disponible: www.sian.info.ve/porcinos/publicaciones/rccpn/rev62/rcpp62art3.htm.

Lerwick, T., (1965). **Maximum Likelihood Estimators of Regression Coefficients for the case of Autocorrelated Residuals**. Technometrics, vol 7.N°1.

Markatou, B. and Lindsay B., (1997).**Weighted likelihood estimating equations: The discrete case with applications to logistic regression**. Journal of Statical Planning and Inference vol (57), pp. 215-232.

Markatou, B and Lindsay. B., (1998).**Weighted likelihood estimating equations with a bootstrap root search.** Journal of the American Statiscal Planning Association vol (93), pp. 740-759.

Markatou, M., (2000). **Mixture Models, Robustness and the Weighted Likelihood methodology.** Biometricsvol (56), pp. 483-486

Montgomery, D., (2005). **Diseño y Análisis de Experimentos**. 2da Edición, Limusa Wiley. p p. 512-549.

Montgomery, D. y Runger.G., (2003).**Probabilidad y Estadística: Aplicadas a la Ingeniera**. 1^{era} Edición, Mc Graw Hill. pp. 701.

O`Neill. M., (2010). **ANOVA and REML.**A guide to linear mixed models in an experimental design context. Statisticaladvisory and training servicePty. LTD.

Olkin, I and Vaeth, M., (1981).**Maximum Likelihood Estimation in a Two- Way Analysis of Variance With Correlated Error in One Classification**. Biometrika Vol. 68, 3; pp 653-660.

Patterson. H, Thompson R.,. (1971). **Recovery of Inter-Block Information when Block Sizes are Unequal.** Biometrika, 58: 545-554.

Plante, J., (2007).**Adaptive Likelihood Weight and Mixtures of Empirical Distributions.** unpublished doctoral dissertation, Department of Statistics, The University of British Columbia.

Rao, B., (1991).**Asymptotic Theory of Weighted Maximum Likelihood for Growth Models In**: Statical Inference InStachastic Processes 183-208 edited by prabhu, N.U. and Basawa, I.V. Marcel Dekker, Inc., New York.

Rencher, A and Schaalje, B., (2007).**Linear Models In Statistics**. 2da Edition. Wiley-Interscience A Johnwiley & Sonsnc. Department of statistics, Brigham Young University, Provo, Utah.

Rodríguez, L., (2003).**Estimaciones en Campos Aleatorios Gaussianos, Máxima Verosimilitud y Máxima Verosimilitud Restringida.** Trabajo de Asenso. Universidad Centroccidental "Lisandro Alvarado".

Searle. S, Casella. G and McCulloch, C., (1992). **Variance Components**. John Wily ans Sons, INC,. Publication, New Nork.

Sen, Pranab k (1996).**On the Asymptotic Distributional Risks of shrinkage and Preliminary Test Versions of maximum likelihood Estimators.** Sankhya: The Indian Journal of Statistcsvol 40, serie B, Pts 3-4, pp. 192-196-

Szatrowski. T and Miller, J., (1980). **Explicit Maximum Likelihood From Balance data in the Mixed Model of the Analysis of Variance**. The Annals of Statistics. Vol 8 nº4.

Thompson, W., (1962).**The Problem of Negative Estimates of Variance Components**. Annals of Mathematical Statistics, 33, 273-289.

Toutenbur. H and Shalabh., (2009). **Statistical Analysis of Designed Experiments**.3era Edicion Springer. New York Dordrecht Heidelberg London.

Villa, A., (2008). **Comparación de los Métodos de Mínimos Cuadrados Ordinarios y Máxima Verosimilitud para la Estimación de Parámetros en Diseños Factoriales (pxq) en Modelos Mixtos.** Trabajo de Grado para Optar al Título de Magister Scientiatum en Estadística. Postgrado Estadística UCV.

Wang, X, Van C., Zidek, J., (2004). **Asymptotic Properties of Maximum Weighted Likelihood Estimators.** Journal of Statistical Planning and Inference.

Wang, S., (2001).**Maximum Weighted Likelihood Estimation.** Ph.D. Thesis. Unpublished Doctoral Dissertation Department of Statistics, the University of British Columbia, Vancouver, BC, Canada.

Warm T.A., (1989)."**Weighted Likelihood Estimation of Ability in Item Response Theory." Psychometrika,** 54, 427-450.

Yáñez, S., (2004). **La Estadística en la Ciencia del Siglo XX. R.A. Fisher. El Genio**. Revista Universidad Nacional de Colombia, Sede Medellín Matemática y Física. Vol. I, N 2, pp 1-14

ANEXOS

CUADRADOS MÍNIMOS MODELO FIJO

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	324.052	.0000
Efecto B	6.983	.0014
Interacción	10.603	.0000

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 4.2963

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	144.577	.0000
Efecto B	3.115	.0484
Interacción	4.731	.0003

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 9.6296

DESBALANCEADO HOMOGÉNEO (n =100)

	Valor de F	Prob.
Efecto A	221.627	.0000
Efecto B	3.924	.0233
Interacción	6.926	.0000

Promedios de B1 = 12.619
B2 = 14.036
B3 = 14.000

Cuadrado medio del residual = 5.1937

DESBALANCEADO HETEROGÉNEO (n =100)

	Valor de F	Prob.
Efecto A	98.944	.0000
Efecto B	2.037	.1365
Interacción	3.182	.0071

Promedios de B1 = 12.536
B2 = 14.103 B3 = 14.000

Cuadrado medio del residual = 11.7437

CUADRADOS MÍNIMOS MODELO MIXTO

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	324.052	.0000
Efecto B	0.659	.5514
Interacción	10.603	.0000

Componente de varianza AB = 4.1259
A = 44.8889

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0
Cuadrado medio del residual = 4.2963

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	144.577	.0000
Efecto B	0.659	.5514
Interacción	4.731	.0003

Componente de varianza AB = 3.5926
A = 44.8889

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0
Cuadrado medio del residual = 9.6296

DESBALANCEADO HOMOGÉNEO (n =100)

	Valor de F	Prob.
Efecto A	218.806	.0000
Efecto B	0.558	.5996
Interacción	6.926	.0000

Componente de varianza AB = 3.7390
A = 44.2793

Promedios de B1 = 13.174
B2 = 14.564
B3 = 14.529
Cuadrado medio del residual = 5.1937

DESBALANCEADO HETEROGÉNEO (n =100)

	Valor de F	Prob.
Efecto A	97.514	.0000
Efecto B	0.619	.5698
Interacción	3.182	.0071

Componente de varianza AB = 3.1124
A = 44.5795
Promedios de B1 = 13.106
B2 = 14.627
B3 = 14.531

Cuadrado medio del residual = 11.7437

MÁXIMA VEROSIMILITUD MODELO FIJO

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	360.06	.0000
Efecto B	7.76	.0014
Interacción	11.78	.0000

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0
Cuadrado medio del residual = 3.8687

Verosimilitud -2 Log	502.8
AIC (mejor más pequeño)	528.8
AICC (mejor mas pequeño)	532.3
BIC (mejor más pequeño)	565.1

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	160.64	.0000
Efecto B	3.46	.0349
Interacción	5.26	.0000

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 8.6667

Verosimilitud -2 Log	599.7
AIC (mejor más pequeño)	625.7
AICC (mejor mas pequeño)	629.1
BIC (mejor más pequeño)	661.9

DESBALANCEADO HOMOGÉNEO (n =100)

	Valor de F	Prob.
Efecto A	251.85	.0000
Efecto B	4.46	.0143
Interacción	7.87	.0000

Promedios de B1 = 12.619
B2 = 14.036
B3 = 14.000
Cuadrado medio del residual = 4.5705

Verosimilitud -2 Log	435.7
AIC (mejor más pequeño)	461.7
AICC (mejor mas pequeño)	466.0
BIC (mejor más pequeño)	495.6

DESBALANCEADO HETEROGÉNEO (n =100)

	Valor de F	Prob.
Efecto A	112.44	.0000
Efecto B	2.31	.1048
Interacción	3.62	.0030

Promedios de B1 = 12.536
B2 = 14.103
B3 = 14.000

Cuadrado medio del residual = 10.3345

Verosimilitud -2 Log	517.3
AIC (mejor más pequeño)	543.3
AICC (mejor mas pequeño)	547.6
BIC (mejor más pequeño)	577.2

MÁXIMA VEROSIMILITUD MODELO MIXTO

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto B	0.88	.4629

Componente de varianza AB = 2.9870
A = 33.6667

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 4.2963

Verosimilitud -2 Log 554.0
AIC (mejor más pequeño) 566.0
AICC (mejor mas pequeño) 566.8
BIC (mejor más pequeño) 562.4

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto B	0.88	.4629

Componente de varianza AB = 2.4537
A = 33.6667
Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 9.6296

Verosimilitud -2 Log 641.2
AIC (mejor más pequeño) 653.2
AICC (mejor mas pequeño) 653.9
BIC (mejor más pequeño) 649.5

DESBALANCEADO HOMOGÉNEO (n =100)

	Valor de F	Prob.
Efecto B	0.78	.4989

Componente de varianza AB = 2.6434
A = 33.0054

Promedios de B1 = 12.6423
B2 = 14.0554
B3 = 14.0143

Cuadrado medio del residual = 5.1947

Verosimilitud -2 Log 482.1
AIC (mejor más pequeño) 494.1
AICC (mejor mas pequeño) 495.0
BIC (mejor más pequeño) 490.4

DESBALANCEADO HETEROGÉNEO (n =100)

	Valor de F	Prob.
Efecto B	0.87	.4673

Componente de varianza AB = 1.9600
A = 33.1193

Promedios de B1 = 12.5917
B2 = 14.1428
B3 = 14.0311

Cuadrado medio del residual = 11.7495

Verosimilitud -2 Log	554.2
AIC (mejor más pequeño)	566.2
AICC (mejor mas pequeño)	567.1
BIC (mejor más pequeño)	562.5

MÁXIMA VEROSIMILITUD RESTRINGIDA MODELO FIJO

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	324.05	.0000
Efecto B	6.98	.0014
Interacción	10.60	.0000

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 4.2963

Verosimilitud -2 Log	491.6
AIC (mejor más pequeño)	493.6
AICC (mejor mas pequeño)	493.6
BIC (mejor más pequeño)	496.2

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	144.58	.0000
Efecto B	3.12	.0484
Interacción	4.73	.0003

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 9.6296

Verosimilitud -2 Log	578.7
AIC (mejor más pequeño)	580.7
AICC (mejor mas pequeño)	580.8
BIC (mejor más pequeño)	583.4

DESBALANCEADO HOMOGÉNEO (n =100)

	Valor de F	Prob.
Efecto A	221.63	.0000
Efecto B	3.92	.0233
Interacción	6.93	.0000

Promedios de B1 = 12.619
B2 = 14.036
B3 = 14.000

Cuadrado medio del residual = 5.1937

Verosimilitud -2 Log	420.0
AIC (mejor más pequeño)	422.0
AICC (mejor mas pequeño)	422.0
BIC (mejor más pequeño)	424.5

DESBALANCEADO HETEROGÉNEO (n =100)

	Valor de F	Prob.
Efecto A	98.94	.0000
Efecto B	2.04	.1365
Interacción	3.18	.0071

Promedios de B1 = 12.536
B2 = 14.103
B3 = 14.000

Cuadrado medio del residual = 11.7437

Verosimilitud -2 Log	491.8
AIC (mejor más pequeño)	493.8
AICC (mejor mas pequeño)	493.8
BIC (mejor más pequeño)	496.3

MÁXIMA VEROSIMILITUD RESTRINGIDA MODELO MIXTO

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto B	0.66	.5514

Componente de varianza AB = 4.1259
A = 44.8889

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 4.2963

Verosimilitud -2 Log	545.2
AIC (mejor más pequeño)	551.2
AICC (mejor mas pequeño)	551.4
BIC (mejor más pequeño)	549.3

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto B	0.66	.5514

Componente de varianza AB = 3.5926
A = 44.8889

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 9.6296

Verosimilitud -2 Log	632.3
AIC (mejor más pequeño)	638.3
AICC (mejor mas pequeño)	638.5
BIC (mejor más pequeño)	636.5

DESBALANCEADO HOMOGÉNEO (n =100)

	Valor de F	Prob.
Efecto B	0.59	.5842

Componente de varianza AB = 3.7405
A = 44.0632

Promedios de B1 = 12.6365
B2 = 14.0505
B3 = 14.0108

Cuadrado medio del residual = 5.1941

Verosimilitud -2 Log	473.3
AIC (mejor más pequeño)	479.3
AICC (mejor mas pequeño)	479.6
BIC (mejor más pequeño)	477.5

DESBALANCEADO HETEROGÉNEO (n =100)

	Valor de F	Prob.
Efecto B	0.65	.5532

Componente de varianza AB = 3.1025

A = 44.2889

Promedios de B1 = 12.5779
B2 = 14.1329
B3 = 14.0233

Cuadrado medio del residual = 11.7461

Verosimilitud -2 Log	545.4
AIC (mejor más pequeño)	551.4
AICC (mejor mas pequeño)	551.6
BIC (mejor más pequeño)	549.5

RESUMEN SALIDAS DE PONDERADOS

MÁXIMA VEROSIMILITUD PONDERADO MODELO DE EFECTOS FIJOS

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	359.71	.0000
Efecto B	7.94	.0006
Interacción	11.82	.0000

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 2.2047

Verosimilitud -2 Log	498.0
AIC (mejor más pequeño)	524.0
AICC (mejor mas pequeño)	527.5
BIC (mejor más pequeño)	560.3

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	125.31	.0000
Efecto B	3.39	.0375
Interacción	6.31	.0000

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 4.3635

Verosimilitud -2 Log	609.0
AIC (mejor más pequeño)	635.0
AICC (mejor mas pequeño)	638.4
BIC (mejor más pequeño)	671.2

<u>DESBALANCEADO HOMOGÉNEO</u> (n =100)

	Valor de F	Prob.
Efecto A	258.53	.0000
Efecto B	4.38	.0153
Interacción	7.77	.0000

Promedios de B1 = 12.619
B2 = 14.036
B3 = 14.000

Cuadrado medio del residual = 2.9318

Verosimilitud -2 Log	433.5
AIC (mejor más pequeño)	459.5
AICC (mejor mas pequeño)	463.7
BIC (mejor más pequeño)	493.3

<u>DESBALANCEADO HETEROGÉNEO</u> (n =100)

	Valor de F	Prob.
Efecto A	174.57	.0000
Efecto B	1.87	.1605
Interacción	6.07	.0000

Promedios de B1 = 12.536
B2 = 14.103
B3 = 14.000

Cuadrado medio del residual = 3.9615

Verosimilitud -2 Log	486.1
AIC (mejor más pequeño)	512.1
AICC (mejor mas pequeño)	516.4
BIC (mejor más pequeño)	546.0

MÁXIMA VEROSIMILITUD PONDERADOS MODELO DE EFECTOS MIXTOS

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto B	0.85	.4743

Componente de varianza AB = 2.980
A = 33.7272

Promedios de B1 = 12.4967
B2 = 13.9304
B3 = 13.9999

Cuadrado medio del residual = 2.4498

Verosimilitud -2 Log	549.7
AIC (mejor más pequeño)	561.7
AICC (mejor mas pequeño)	562.5
BIC (mejor más pequeño)	558.0

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto B	0.90	.4539

Componente de varianza AB = 2.7937
A = 31.7641

Promedios de B1 = 12.4910
B2 = 14.1933
B3 = 13.9201

Cuadrado medio del residual = 4.8233

Verosimilitud -2 Log	650.3
AIC (mejor más pequeño)	662.3
AICC (mejor mas pequeño)	663.0
BIC (mejor más pequeño)	658.6

DESBALANCEADO HOMOGÉNEO (n =100)

	Valor de F	Prob.
Efecto B	0.72	.5237

Componente de varianza AB = 2.5961
A = 33.0238

Promedios de B1 = 12.6567
B2 = 13.9430
B3 = 14.0255

Cuadrado medio del residual = 3.3348

Verosimilitud -2 Log	480.1
AIC (mejor más pequeño)	492.1
AICC (mejor mas pequeño)	493.0
BIC (mejor más pequeño)	488.4

DESBALANCEADO HETEROGÉNEO (n =100)

	Valor de F	Prob.
Efecto B	0.94	.4410

Componente de varianza AB = 2.8679
A = 32.0061

Promedios de B1 = 12.5057
B2 = 14.3812
B3 = 14.1195

Cuadrado medio del residual = 4.4655

Verosimilitud -2 Log	527.3
AIC (mejor más pequeño)	539.3
AICC (mejor mas pequeño)	540.2
BIC (mejor más pequeño)	535.6

MÁXIMA VEROSIMILITUD RESTRINGIDA MODELOS EFECTOS FIJOS

BALANCEADO HOMOGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	323.74	.0000
Efecto B	7.14	.0012
Interacción	10.64	.0000

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 2.4496

Verosimilitud -2 Log	487.3
AIC (mejor más pequeño)	489.3
AICC (mejor mas pequeño)	489.3
BIC (mejor más pequeño)	491.9

BALANCEADO HETEROGÉNEO (n =120)

	Valor de F	Prob.
Efecto A	112.78	.0000
Efecto B	3.05	.0516
Interacción	5.68	.0000

Promedios de B1 = 12.5
B2 = 14.0
B3 = 14.0

Cuadrado medio del residual = 4.8483

Verosimilitud -2 Log	587.1
AIC (mejor más pequeño)	589.1
AICC (mejor mas pequeño)	589.1
BIC (mejor más pequeño)	591.8

<u>DESBALANCEADO HOMOGÉNEO</u> (n =100)

	Valor de F	Prob.
Efecto A	227.51	.0000
Efecto B	3.86	.0248
Interacción	6.84	.0000

Promedios de B1 = 12.619
B2 = 14.036
B3 = 14.000
Cuadrado medio del residual = 3.3316

Verosimilitud -2 Log	418.1
AIC (mejor más pequeño)	420.1
AICC (mejor mas pequeño)	420.1
BIC (mejor más pequeño)	422.6

<u>DESBALANCEADO HETEROGÉNEO</u> (n =100)

	Valor de F	Prob.
Efecto A	153.62	.0000
Efecto B	1.64	.1991
Interacción	5.34	.0000

Promedios de B1 = 12.536
B2 = 14.103
B3 = 14.000

Cuadrado medio del residual = 4.5017

Verosimilitud -2 Log	463.8
AIC (mejor más pequeño)	465.8
AICC (mejor mas pequeño)	465.8
BIC (mejor más pequeño)	468.3

<u>MÁXIMA VEROSIMILITUD RESTRINGIDA MODELO DE EFECTOS MIXTOS</u>

<u>BALANCEADO HOMOGÉNEO</u> (n =120)

	Valor de F	Prob.
Efecto B	0.64	.5596

Componente de varianza AB = 4.1199

A = 44.9491

Promedios de B1 = 12.4975
B2 = 13.9479
B3 = 13.9992

Cuadrado medio del residual = 2.4497

Verosimilitud -2 Log	540.9
AIC (mejor más pequeño)	546.9
AICC (mejor mas pequeño)	547.1
BIC (mejor más pequeño)	545.0

<u>BALANCEADO HETEROGÉNEO</u> (n =120)

	Valor de F	Prob.
Efecto B	0.66	.5506

Componente de varianza AB = 3.9724
A = 42.8654

Promedios de B1 = 12.4940
B2 = 14.1472
B3 = 13.9309

Cuadrado medio del residual = 4.8314

Verosimilitud -2 Log	641.1
AIC (mejor más pequeño)	647.1
AICC (mejor mas pequeño)	647.3
BIC (mejor más pequeño)	645.3

<u>DESBALANCEADO HOMOGÉNEO</u> (n =100)

	Valor de F	Prob.
Efecto B	0.55	.6015

Componente de varianza AB = 3.6917
A = 44.0830

Promedios de B1 = 12.6472

B2 = 13.9660
B3 = 14.0192

Cuadrado medio del residual = 3.3333

Verosimilitud -2 Log	471.4
AIC (mejor más pequeño)	477.4
AICC (mejor mas pequeño)	477.6
BIC (mejor más pequeño)	475.5

<u>DESBALANCEADO HETEROGÉNEO</u> (n =100)

	Valor de F	Prob.
Efecto B	0.68	.5425

Componente de varianza AB = 4.1045
A = 44.1557

Promedios de B1 = 12.5139
B2 = 14.3184
B3 = 14.0891

Cuadrado medio del residual = 4.4748

Verosimilitud -2 Log	518.1
AIC (mejor más pequeño)	524.1
AICC (mejor mas pequeño)	524.3
BIC (mejor más pequeño)	522.2