



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



**ESTIMACIÓN MÁXIMO VEROSÍMIL Y MÍNIMOS CUADRADOS EN EL
MODELO ESTADÍSTICO LINEAL MIXTO: METODOLOGÍA Y ANÁLISIS.**

Autor: Lcda. Vianey Méndez

Tutor: Dr. Miguel Balza

**TRABAJO DE GRADO PRESENTADO COMO REQUISITO FINAL PARA
OPTAR AL TÍTULO DE MAGISTER SCIENTIARUM EN ESTADÍSTICA.**

Maracay, noviembre de 2015

DEDICATORIA

*A Dios Todopoderoso por ser mi guía y
ayudador, a ti sea toda la gloria y el honor!*

AGRADECIMIENTOS

Al Dios Altísimo por ser mi refugio en todo momento, porque gracias a Él hoy puedo culminar otra meta más en mi vida, sé que sin ti nada soy, gracias mi Señor.

A mis padres por ser los mejores del mundo, por estar conmigo en todo momento, Uds. después de Dios son mi motor, los amo.

Al Dr. Miguel Balza por toda su colaboración, por brindarme su apoyo en los momentos más difíciles, cuando no conseguía tutor Ud. con mucho gusto se ofreció y por eso le estoy muy agradecida.

A los profesores Luis Pérez y Omar Verde por todo el apoyo brindado, por sus consejos acertados en su justo momento, ha sido un beneplácito haberles conocido y trabajar con Uds.

A los profesores de la Cátedra de Genética de la Facultad de Ciencias Veterinarias gracias por toda la colaboración prestada. A mis compañeros de trabajo profesores de la Cátedra de Bioestadística de la Facultad de Ciencias Veterinarias por sus palabras de aliento en esos momentos tan difíciles, por todo su apoyo mil gracias.

A los profesores de este maravilloso postgrado por impartir tan valiosos conocimientos, son personas a las que admiro mucho y como dicen por allí cuando sea grande quiero ser como ustedes. Al personal administrativo y de control de estudios de este postgrado mil gracias por todo el apoyo brindado.

A mis amigos Marvin Querales y Manuel Maluenga, por compartir sus experiencias, conocimiento y amistad, fue una aventura cursar esta maestría con Uds.

Mi eterno agradecimiento para todos.
Vianzy Méndez.

ÍNDICE GENERAL

	Página
Índice de Cuadros	vi
Índice de Gráficos	vii
Resumen	viii
Introducción	1
1. EL PROBLEMA	3
1.1 Planteamiento y Formulación del Problema	3
1.2 Objetivos de la Investigación.....	4
1.3 Justificación.....	5
2. ANTECEDENTES Y MARCO TEÓRICO	6
3. MARCO METODOLÓGICO	15
3.1 Método de Mínimos Cuadrados.....	15
3.1.1 Estimación de los efectos fijos.....	17
3.1.2 Estimación de los efectos aleatorios.....	17
3.2 Método de Máxima Verosimilitud.....	19
3.2.1 Algoritmo de Newton-Raphson (NR).....	21
3.2.2 Algoritmo de Fisher Scoring (FS).....	22
3.2.3 Algoritmo de Maximización De la Esperanza (EM).....	23
3.2.4 Algoritmo de Información Promedio (AI).....	24
PROC MIXED.....	25
PROC GLM.....	26
Tipos de Sumas de Cuadrados.....	26
Criterio de Selección CME.....	27
Criterios de Información AIC y BIC.....	27
4. ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS	31
5. CONCLUSIONES Y RECOMENDACIONES	57
BIBLIOGRAFÍA	58
ANEXOS	63
ANEXO A	
Instrucciones programa SAS, métodos ML, REML y GLM.....	64
ANEXO B	
Salida del programa SAS, (SC III), datos Balanceados, repetición 3...	65
ANEXO C	
Salida del programa SAS, (SC III), datos Desbalanceados, repetición 3	67
ANEXO D	
Salida del programa SAS, (SC III), datos Balanceados, repetición 5....	69
ANEXO E	
Salida del programa SAS, (SC III), datos Desbalanceados, repetición 5	71

ÍNDICE GENERAL

	Página
ANEXO F	
Salida del programa SAS, (SC III), datos Balanceados, repetición 10..	73
ANEXO G	
Salida del programa SAS, (SC III), datos Desbalanceados, repetición 10	75
ANEXO H	
Salida del programa Statistix (Estadística Descriptiva).....	77
ANEXO I	
Resumen de la salida del SAS, (SC III), (Peso al Nacer).....	78
ANEXO J	
Resumen de la salida del SAS, (SC III), (Peso al Destete).....	79
ANEXO K	
Resumen de la salida del SAS, (SC III), (Peso a los 18 Meses).....	80
ANEXO L	
Resumen de la salida del SAS (Promedios).....	81
ANEXO M	
Resumen de la salida del SAS (Promedios), estimados interacción AN*MN.....	82

ÍNDICE DE CUADROS

	Página
Cuadro 1. Criterios de Información generados por el programa SAS...	28
Cuadro 2. Resultados obtenidos por el programa SAS para datos simulados (Balanceados), (SC III).....	32
Cuadro 3. Resultados obtenidos por el programa SAS datos simulados (Desbalanceados), (SC III).....	35
Cuadro 4. Resultados obtenidos por el programa SAS de la variable Peso al Nacer (PN), (SC III).....	44
Cuadro 5. Análisis de Varianza para Peso al Nacer (PN), (SC III) y Estimación de Parámetros de Covarianza.....	46
Cuadro 6. Resultados obtenidos por el programa SAS usando la variable Peso al destete o a los 205 días (AJUSDE), (SC III).....	47
Cuadro 7. Análisis de Varianza para Peso al Destete (AJUSDE), (SC III) y Estimación de Parámetros de Covarianza.....	48
Cuadro 8. Resultados obtenidos por el programa SAS usando la variable Peso a los 18 meses o 548 días (AJUS18), (SC III).....	49
Cuadro 9. Análisis de Varianza para Peso a los 18 Meses (AJUS18), (SC III) y Estimación de Parámetros de Covarianza.....	50

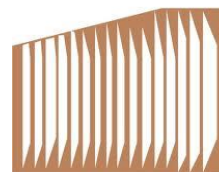
ÍNDICE DE GRÁFICOS.

	Página
Gráfico 1. Distribución de los becerros por año.....	37
Gráfico 2. Edades de las madres al parto de cada becerro evaluado por año.....	38
Gráfico 3. Meses de nacimiento de los becerros por año.....	38
Gráfico 4. Sexo de los becerros por año.....	39
Gráfico 5. Pesos de los Becerros al Nacer.....	40
Gráfico 6. Pesos de los Becerros a los 205 días.....	41
Gráfico 7. Pesos de los Becerros a los 548 días.....	42
Gráfico 8. Promedio del peso al nacer (PN), peso corregido a los 205 días (AJUSDE) y peso corregido a los 548 días (AJUS18) según el año de nacimiento	51
Gráfico 9. Promedio del peso al nacer (PN) y peso corregido a los 205 días (AJUSDE) según la edad de la madre al parto (años).....	52
Gráfico 10. Promedio del peso a los 205 días (AJUSDE) según los meses de nacimiento.....	53
Gráfico 11. Promedio del peso al nacer (PN), peso corregido a los 205 días (AJUSDE) y peso corregido a los 548 días, según el sexo.....	54
Gráfico 12. Promedio del peso al nacer (PN), para la interacción AN*MN.....	54
Gráfico 13. Promedio del peso al destete (AJUSDE), para la interacción AN*MN.....	55
Gráfico 14. Promedio del peso a los 548 días (AJUS18), para la interacción AN*MN.....	55



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

CURSO DE POSTGRADO EN
estadística



ESTIMACIÓN MÁXIMO VEROSÍMIL Y MÍNIMOS CUADRADOS EN EL
MODELO ESTADÍSTICO LINEAL MIXTO: METODOLOGÍA Y ANÁLISIS.

Autor: Lcda. Méndez Vianey

Tutor: Dr. Miguel Balza

Fecha: Noviembre, 2015

RESUMEN

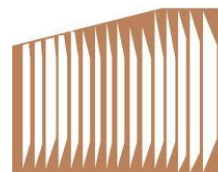
El presente estudio tiene como propósito establecer la metodología de estimación y análisis en el Modelo Lineal Mixto. Estos modelos se utilizan principalmente para describir las relaciones entre una variable respuesta y algunas covariables en los datos que se agrupan de acuerdo a uno o más factores de clasificación, en donde los mismos son de efectos fijos y aleatorios, además brindan la posibilidad de analizar datos dependientes, desbalanceados y con falta de normalidad. Para la presente investigación se describieron los procedimientos teóricos de los métodos de estimación Máximo Verosímil y Máximo Verosímil Restringido (ML y REML), además del Mínimos Cuadrados (GLM). Con el propósito de determinar cuál de las estimaciones del modelo sería la más conveniente, se tomó en consideración el cuadrado medio del residual (CME) y de allí el mejor método será aquel que arroje el menor (CME), ya que este se considera como una herramienta útil para decidir qué tan apropiado es un modelo y determinar la medida en la que el modelo no se ajusta a la información, o si el quitar ciertos términos puede simplificar el modelo de manera favorecedora. El proporciona una forma para elegir el mejor estimador: un CME mínimo a menudo, indica una variación mínima, y por lo tanto se considera un buen estimador. Para contrastar los métodos de estimación antes mencionados, se realizaron dos ejemplos uno para datos simulados (balanceados y desbalanceados) normalmente distribuidos con varianzas homogéneas y otro con datos reales, desbalanceados, enfocados en el área de genética, especialmente en bovinos. Adicionalmente se ajustaron distintos modelos al mismo conjunto de datos reales, por lo tanto, fue necesario seleccionar el más adecuado para éstos, por lo cual se utilizaron los criterios de información como: Criterio de Información de Akaike (AIC) y Criterio de Información de Schwarz o Bayesiano (BIC). Se obtuvo la Máxima Verosimilitud como el método de estimación que arrojó menor cuadrados medios residuales.

Palabras Claves: Modelo Lineal Mixto, Métodos Máximo Verosímil (ML y REML), Método de Mínimos Cuadrados, Cuadrados Medios de los Residuales, AIC y BIC.



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

CURSO DE POSTGRADO EN
estadística



ESTIMATION MAXIMUM LIKELIHOOD AND LEAST SQUARE IN STATISTICAL LINEAR MIXED MODELS: METHODOLOGY AND ANALYSIS.

Author: Lcda. Méndez Vianey

Advisor: Dr. Miguel Balza

Date: November, 2015

SUMMARY

The present study aims to establish the estimation methodology and analysis in the mixed linear model, these models are mainly used to describe the relationship between a response variable and some covariates in the data that are grouped according to one or more factors of classification, where the same are fixed and random effects, in addition provide the possibility of analyzing data no dependent, unbalanced and lacking of normality. For the present investigation described the theoretical procedures of Maximum Likelihood, Restricted Maximum Likelihood (ML and REML) and Least Squares (GLM) estimation methods. In order to determine which of the estimates of the model would be most suitable, was took into consideration the mean square residual (MSE) and then the best method will be that resulting with the lower (MSE), since this is considered as a useful tool to decide what so proper is a model and determine the extent in which the model does not conform to the information, or remove certain terms can be simplified in ways which promote model. It provides a way to choose the best estimator: a minimum mean square residual often indicates a minimum variation, and is therefore considered a good estimator. To contrast the estimation methods previously mentioned, were performed two examples one for simulated data (balanced and unbalanced) normally distributed with homogeneous variances and other with actual unbalanced data focused on the area of genetics, especially in cattle. Additionally different models at the same set of real data were adjusted, so it was necessary to select the most suitable for them, therefore the information criteria were used as: Akaike Information Criterion (AIC) and Schwarz Information Criterion or Bayesian (BIC). It was obtained the maximum likelihood as the method of estimation with the lower square medium residual.

Keywords: Linear Mixed Model, Maximum Credible Methods (ML and REML), Method of Least Squares, Mean Squares of the Residuals, AIC and BIC.

INTRODUCCIÓN

Los diseños de experimentos se basan en modelos estadísticos que indican los factores que modifican la variable respuesta y estos modelos, por lo general son lineales.

Muchos modelos estadísticos comunes pueden ser expresados como modelos lineales que incorporan efectos fijos, que son parámetros asociados con una población entera o con ciertos niveles repetibles de factores experimentales, y efectos aleatorios, que se asocian con las unidades experimentales individuales extraídos al azar de una población (Pinheiro y Bates, 2000).

Un modelo con ambos efectos fijos y efectos aleatorios se conoce como modelo de efectos mixtos, los cuales fueron ampliamente estudiados por Fisher hacia el año 1918, quien los denominó modelos de componentes de varianza, siendo estos de gran utilidad en los estudios de genética cuantitativa y mejoramiento animal. Estos modelos se utilizan principalmente para describir las relaciones entre una variable de respuesta y algunas covariables en los datos que se agrupan de acuerdo a uno o más factores de clasificación, Pu y Niu (2006), Iyit y Genc (2011).

Referido a que existen muchos métodos de estimación de ajustes del modelo, dependiendo de su naturaleza, las estimaciones de estos modelos, en este trabajo, se basarán en los métodos de Máxima Verosimilitud, Máxima Verosimilitud Restringida y Mínimos Cuadrados los cuales serán explicados a profundidad en el desarrollo de esta investigación.

El presente estudio se estructuró en cinco capítulos:

En el Problema, se presenta el planteamiento y formulación del problema, objetivo general, específicos y justificación, siendo esta la que explica el porqué de la

investigación. **En el Marco Teórico**, se hace referencia al marco teórico relacionado con este tipo de modelos, su proceso de estimación por los métodos de Máxima Verosimilitud, Máxima Verosimilitud Restringida y Mínimos Cuadrados, además de antecedentes vinculados con la misma.

En el Marco Metodológico, se explica a profundidad los métodos seleccionados para el análisis de los datos como los métodos de Máxima Verosimilitud y Máxima Verosimilitud Restringida, con una breve descripción de sus algoritmos de estimación, y el método de Mínimos Cuadrados, los diferentes tipos de Sumas de Cuadrados, además de los criterios de selección como el Criterio de Información de Akaike (AIC) y el Criterio de Información de Schwarz o Bayesiano (BIC), y el cuadrado medio del residual los cuales serán utilizados en el desarrollo de los **Resultados**, en donde se hace referencia al análisis e interpretación de los resultados, para los cuales se realizaron dos ejemplos el primero con datos simulados (balanceados y desbalanceados), con el objeto de validación de los métodos, y el segundo con datos reales, desbalanceados, ambos conjuntos de datos se analizaron mediante los métodos de Máxima Verosimilitud, Máxima Verosimilitud Restringida y Mínimos Cuadrados, con la finalidad de hacer la comparación entre los métodos de estimación a través del resultado del cuadrado medio del residual, (CME) , además de la aplicación de los criterios de información AIC y BIC, para la escogencia del mejor modelo de ajuste. **Conclusiones**, es donde se encuentran las conclusiones y recomendaciones referentes al presente estudio.

1. EL PROBLEMA

1.1 Planteamiento y Formulación del Problema

Un modelo estadístico lineal es la expresión de una relación estocástica lineal existente entre una variable que se denota como dependiente y una serie de variables denominadas independientes y generalmente continuas, la linealidad prevalece en los parámetros desconocidos del modelo White y Hodge (1989), Salto (2008). Este modelo es ampliamente utilizado en la experimentación para analizar la variabilidad de observaciones realizadas sobre características de importancia en función de una o más variables predictoras o factores.

Para especificar estos modelos en un experimento es necesario definir si la naturaleza de sus efectos es fija o aleatoria. Según Searle *et al.* (1992) y Salto (2008) si las inferencias están circunscritas a los efectos del modelo, los efectos son considerados fijos, para lo cual destaca que si todos los niveles de clasificación se encuentran en el experimento, o, sólo los niveles de interés del investigador están en el experimento, o, si los niveles del experimento provienen de una población normalmente distribuida pero no fueron elegidos al azar se está en presencia de efectos fijos.

Por su parte Henderson (1984), plantea que un efecto es aleatorio si los niveles del factor en el experimento pueden ser considerados como una muestra aleatoria de alguna población de niveles generados estocásticamente. La decisión de considerar como fijo o aleatorio un conjunto de efectos depende de las inferencias que se extraerán y está ligada al proceso de muestreo White y Hodge (1989), Salto (2008).

Para el análisis estadístico de estos modelos lineales (fijos o aleatorios) se deben cumplir con los supuestos del Análisis de Varianza (ANAVAR), Balzarini *et al.* (2005), como Normalidad de los residuales, Linealidad de los efectos, Homocedasticidad e

Independencia de las observaciones, si existen desviaciones significativas sobre estos requisitos y no se pueden realizar las transformaciones pertinentes según sea el caso, los resultados posteriores pueden ser incorrectos.

Por lo tanto, cuando no se cumplen los supuestos del ANAVAR o también porque se tiene un diseño que no posee todos sus factores fijos o aleatorios, sino más bien una mezcla de estos, se hace difícil utilizar los modelos lineales clásicos del análisis de varianza y regresión. Los modelos lineales mixtos son una solución a estos problemas ya que brindan la posibilidad de analizar datos no independientes, desbalanceados y con falta de normalidad y, en este contexto, es necesario determinar la estimación del modelo y cuál sería la más conveniente. Por ello surge la presente investigación.

1.2 Objetivos de la Investigación

1.2.1 General

Describir la Metodología de Estimación por Máxima Verosimilitud y Mínimos Cuadrados en el Modelo Estadístico Lineal Mixto.

1.2.2 Específicos

Caracterizar el Modelo Estadístico Lineal Mixto.

Describir la Estimación por Mínimos Cuadrados de los parámetros para los Efectos Fijos y Efectos Aleatorios en el Modelo Estadístico Lineal Mixto.

Describir los Métodos de Estimación por Máxima Verosimilitud en el Modelo Estadístico Lineal Mixto.

Analizar el Modelo Estimado por vía de Máxima Verosimilitud y Mínimos Cuadrados.

1.3 Justificación

Una estimación eficiente del vector de parámetros correspondiente a los efectos fijos, y el vector de parámetros correspondiente a los efectos aleatorios, en el modelo lineal mixto, permite descomponer la varianza total en la varianza entre las unidades de muestreo y la varianza dentro de cada unidad. La ventaja que poseen es que, si se escoge una buena estructura de covarianza, se pueden obtener estimaciones más apropiadas de los efectos fijos y de los errores estándar. Si se tiene una estructura jerárquica de los datos, los resultados pueden ser aún más precisos, ya que permite elegir varianzas distintas para diferentes grupos de sujetos. Y a través de estos, se evitan los problemas causados por los datos faltantes cuando se estiman los parámetros.

Los modelos lineales mixtos, son útiles cuando se tienen casos de medidas repetidas, datos longitudinales, curvas de crecimiento, experimentos multiambientales de correlación espacial. Han sido utilizados en agronomía especialmente en mejoramiento genético animal y vegetal para estimar heredabilidades y predecir ganancia genética, Morante (2011).

También se utilizan en ensayos comparativos de rendimiento para estimar componentes de varianzas asociadas a la comparación de efectos de tratamiento conducidos a través de varias localidades y años (diseños espacio – tiempo), asumiendo según Balzarini *et al.* (2005), que los efectos del tratamiento están contenidos en la interacción aleatoria.

Estos modelos constituyen una alternativa al ajuste por regresión mediante mínimos cuadrados ordinarios cuando las observaciones no son independientes entre sí. El ajuste de éste permite obtener estimadores eficientes de los parámetros fijos del modelo y predecir parámetros aleatorios específicos de cada unidad de muestreo que reflejan el patrón de desvío respecto a la media, Calama y Montero (2004).

2. ANTECEDENTES Y MARCO TEÓRICO

López *et al.* (2007), plantean que los modelos de efectos mixtos fueron ampliamente estudiados por Fisher hacia el año 1918, quien los denominó modelos de componentes de varianza, estos modelos fueron de gran utilidad en los estudios de genética cuantitativa y mejoramiento animal. En los estudios de efectos mixtos es fundamental que se tenga en cuenta, la estimación de efectos fijos y aleatorios.

Los modelos mixtos permiten modelar la respuesta de un estudio experimental u observacional como función de factores o covariables, cuyos efectos pueden considerarse tanto como constantes fijas o variables aleatorias, Balzarini *et al.* (2005). Cada modelo estadístico que contiene una media general (μ) y un término de error aleatorio (ϵ) es considerado modelo mixto por definición, ya que posee una constante μ y un efecto aleatorio ϵ , mas sin embargo en la práctica se denomina modelo mixto a aquel modelo que posee efectos fijos distintos de μ y efectos aleatorios diferentes del error ϵ .

Estos modelos según Calama y Montero (2004), Liu *et al.* (2008), incluyen en su formulación parámetros fijos, comunes a toda la población de niveles de factores y parámetros aleatorios específicos a cada unidad de muestreo, estos últimos tienen media cero y varianza que define la componente del error asociada a la unidad de muestreo.

Fox (2002), López *et al.* (2007), Rencher y Schaalje (2008), plantean que el modelo mixto puede ser escrito:

Forma General:

$$y = X\beta + Z_1U_1 + Z_2U_2 + \dots + Z_cU_c + \epsilon$$

Dónde:

- ☉ $\mathbf{y}$ Es el vector aleatorio observable de orden $\mathbf{n} \times \mathbf{1}$.
- ☉ $\mathbf{X}$ Es una matriz conocida de tamaño $\mathbf{n} \times \mathbf{p}$.
- ☉ $\boldsymbol{\beta}$ Es un vector de parámetros desconocidos de dimensión $\mathbf{p} \times \mathbf{1}$.
- ☉ $\mathbf{Z}_i$ Es una matriz conocida de tamaño $\mathbf{n} \times \mathbf{q}_i$, con $i= 1, 2, \dots, c$.
- ☉ $\mathbf{U}_i$ Es un vector aleatorio de dimensión $\mathbf{q}_i \times \mathbf{1}$.
- ☉ $\boldsymbol{\varepsilon}$ Es un vector aleatorio de orden $\mathbf{n} \times \mathbf{1}$.

Para el modelo se asume que $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_c, \boldsymbol{\varepsilon}$ se distribuyen de manera independiente e idénticamente como:

$$\mathbf{U}_i \sim N(\mathbf{0}; \sigma_i^2 \mathbf{I}_{q_i})$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}; \mathbf{R} \sigma_0^2)$$

Esto es, los efectos regresores aleatorios se distribuyen normalmente con media cero y matriz de varianza $\sigma_i^2 \mathbf{I}_{q_i}$; el vector de error se distribuye normalmente con media cero y varianza $\mathbf{R} \sigma_0^2$, donde $\sigma_0^2, \sigma_2^2, \dots, \sigma_i^2$ son constantes desconocidas no negativas a las cuales se les conoce como componentes de varianza y $\mathbf{R}$ e $\mathbf{I}$ son matrices que se asumen conocidas.

La matriz $\mathbf{R}$ es modelada como $\mathbf{R} = \sigma^2 \mathbf{I}$ cuando se considera que los términos del error son independientes y tienen la misma varianza σ^2 . Sea $\mathbf{G}$ una matriz cuadrada de orden $q \times q$, donde q es el número total de parámetros aleatorios, en los modelos de efectos aleatorios, $\mathbf{G}$ es siempre diagonal, cuyos elementos son de la forma $\sigma_{u_i}^2$; $i=1,2,\dots,c$; según Calama y Montero (2004); Rencher y Schaalje (2008), la matriz $\mathbf{G}$ se denota como:

$$\mathbf{G} = \begin{pmatrix} \sigma^2_{u1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \sigma^2_{u2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \sigma^2_{uc} \end{pmatrix} \rightarrow \mathbf{G} = \begin{pmatrix} \sigma^2_1 \mathbf{I}_{n_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \sigma^2_2 \mathbf{I}_{n_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \sigma^2_c \mathbf{I}_{n_c} \end{pmatrix}$$

Además, la matriz $\mathbf{R}$ es:

$$\mathbf{R} = \begin{pmatrix} \sigma^2_\varepsilon & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \sigma^2_\varepsilon & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \sigma^2_\varepsilon \end{pmatrix}$$

Forma Matricial:

$$\mathbf{Y}_{(nx1)} = \mathbf{X}_{(n \times p)} \boldsymbol{\beta}_{(p \times 1)} + \mathbf{Z}_{(n \times q)} \mathbf{U}_{(q \times 1)} + \boldsymbol{\varepsilon}_{(n \times 1)}$$

↓

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} 1 & x_{11} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & x_{np} \end{bmatrix}_{n \times (p+1)} \begin{bmatrix} \mu \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1) \times 1} + \begin{bmatrix} z_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & z_{nq} \end{bmatrix}_{n \times q} \begin{bmatrix} U_1 \\ \vdots \\ U_q \end{bmatrix}_{q \times 1} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n \times 1}$$

Dónde:

- ☞ $\mathbf{Y}$ es el vector respuesta $(n \times 1)$.
- ☞ $\mathbf{X}$ es la matriz diseño $n \times (p + 1)$ para las observaciones de efecto fijo.
- ☞ $\boldsymbol{\beta}$ es el vector $(p + 1) \times 1$ de los coeficientes de efectos fijos.
- ☞ $\mathbf{Z}$ es la matriz diseño $(n \times q)$ para las observaciones de efectos aleatorios.
- ☞ $\mathbf{U}$ es el vector $(q \times 1)$ de los coeficientes de efectos aleatorios.
- ☞ $\boldsymbol{\varepsilon}$ es el vector aleatorio $(n \times 1)$.

Laird y Ware (1982), Pu y Niu (2006), Ho y Hu (2008), establecen que un modelo de efectos mixtos de la forma:

$$Y_i = X_i\beta + Z_iU_i + \varepsilon_i \quad ; i = 1, \dots, m$$

puede ser descrito mediante dos etapas:

Etapla 1: Variación dentro del sujeto: El termino ε_i es el vector de los errores aleatorios dentro del sujeto con $\varepsilon_i \sim N(\mathbf{0}, \mathbf{R})$ condicional en U_i se tiene:

$$E(Y_i|U_i) = X_i\beta + Z_iU_i$$

$$Cov(Y_i|U_i) = \mathbf{R}$$

Etapla 2: Variación intra sujeto: Suponiendo que los efectos aleatorios de U_i provienen de una distribución normal con media cero y matriz de covarianza $\mathbf{G}$ son independientes el uno del otro y del error aleatorio dentro del sujeto. Bajo estos supuestos la media marginal y la matriz de covarianza para Y_i son:

$$E(Y_i) = E[E(Y_i|U_i)] = X_i\beta$$

$$Cov(Y_i) = E[Cov(Y_i|U_i)] + Cov[E(Y_i|U_i)]$$

$$Cov(Y_i) = \mathbf{R} + \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^t = \mathbf{V}$$

Así, bajo los supuestos de normalidad e independencia de U_i y el ε_i incondicionalmente, los U_i normalmente e independientemente distribuidos con media $X_i\beta$ y matriz de covarianza:

$$\mathbf{V} = \mathbf{R} + \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^t$$

típicamente $\mathbf{R} + \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^t$ depende de un vector de parámetros θ con q elementos.

Bajo los supuestos de normalidad en efectos aleatorios U_i y los errores aleatorios ε_i , los coeficientes β pueden ser estimados por el Método de Mínimos Cuadrados Generalizados y la estimación de los componentes de varianza de θ pueden ser basados en el Método de Máxima Verosimilitud (**ML**) o Máxima Verosimilitud Restringida o Residual (**RML**).

En resumen, según Balsells (2009):

$$E(Y) = E(X\beta + ZU + \varepsilon) \rightarrow E(Y) = XE(\beta) + ZE(U) + E(\varepsilon);$$

la esperanza de U y ε se asume igual a cero. Entonces: $E(Y) = X\beta$

La Matriz de Varianzas y Covarianzas de (Y) está dada por:

$$V(Y) = V(X\beta + ZU + \varepsilon) \rightarrow V(Y) = V(X\beta) + V(ZU) + V(\varepsilon)$$

La varianza del vector β , el cual describe los efectos fijos, es cero y Z es una matriz de constantes, por lo tanto:

$$V(Y) = ZV(U)Z^t + V(\varepsilon)$$

Si se determina como G a la varianza de (U) y R a la varianza del error (ε) , ya que los efectos aleatorios y los residuos siguen una distribución normal, entonces: $U \sim N(0, G)$ y $\varepsilon \sim N(0, R)$, se tiene:

$$V(Y) = ZGZ^t + R$$

G es una matriz cuadrada de tamaño q y este es el número total de parámetros aleatorios, ambas matrices G y R son cuadradas, no singulares, positivamente definidas y con elementos que se asumen conocidos, se asume independencia entre los efectos aleatorios y los términos del error, así que:

$$Cov(u, e) = 0.$$

Balsells (2009), Gumedze y Dunne (2011), plantean que algunos métodos para estimar la varianza de los parámetros en los modelos lineales mixtos, como el análisis de varianza para datos balanceados que utiliza el enfoque de los cuadrados medios. Es difícil de aplicar cuando los datos están desbalanceados o cuando se desea modelar la variación en los datos utilizando una estructura de varianza más compleja.

Para datos desbalanceados estos autores coinciden que Rao propuso la norma mínima de estimación cuadrática (**MINIQUE**), que es un método para estimar los

parámetros de varianza, el cual produce estimadores cuadráticos sesgados. Según Searle (1995), Gumedze y Dunne (2011), Rao en el año 1970 propuso un método de estimación que ha llegado a ser conocido como **MINQUE**, “norma mínima de estimación cuadrática imparcial”. No requiere la normalidad, y se desarrolla por estimar una función lineal de componentes de la varianza.

Henderson en el año 1953, citado por Searle (1995), Gumedze y Dunne (2011), propuso tres métodos para la estimación de los parámetros de varianza, conocidos como los métodos de Henderson, Método I usando formas cuadráticas las cuales son análogas a la suma de cuadrados de los diseños balanceados, Método II es una adaptación del Método I y toma en cuenta los efectos fijos en el modelo, Método III también llamado método de ajustes de constantes (**FITCON**), este utiliza las sumas de cuadrados de los modelos mixtos completos como si todos los términos fuesen efectos fijos.

Según Balsells (2009), la estimación de los parámetros del modelo lineal mixto se realiza a través del método de Máxima Verosimilitud (**ML**) el cual maximiza el logaritmo de la función con respecto a los parámetros de la varianza considerando los efectos fijos β como constantes, este método produce estimaciones de los parámetros sesgadas, siendo este mayor cuanto más pequeños sean los grados de libertad.

En los modelos donde las observaciones son independientes, la función de verosimilitud es el producto de las funciones de densidad de cada observación, Balsells (2009), plantea que en los modelos mixtos, las observaciones no son independientes, la función de verosimilitud necesita ser basada en función de densidad multivariante que puede ser definida usando la distribución normal multivariante, como:

$$L = \frac{1}{(2\pi)^{n/2} \sqrt{| \mathbf{ZGZ}^t + \mathbf{R} |}} e^{-1/2(\mathbf{Y}-\mathbf{X}\beta)' \mathbf{V}^{-1}(\mathbf{Y}-\mathbf{X}\beta)}$$

Más, sin embargo en la práctica, Balsells (2009), plantea que se utiliza el logaritmo de la función de verosimilitud, siendo esta:

$$\log(L) = K - \frac{1}{2} [\log|ZGZ^t + R| + (Y - X\beta)^t V^{-1} (Y - X\beta)]$$

dónde: $K = -\frac{1}{2} * n \log(2\pi)$ es una constante que puede ser ignorada en el proceso de maximización y n es el número de observaciones.

Balsells (2009), Gumedze y Dunne (2011), coinciden en que existen otros métodos adicionales al Método de Máxima Verosimilitud para ajustar el modelo lineal mixto, basados en la maximización de la función verosimilitud, tales como:

Máxima Verosimilitud Restringida (REML): Desarrollado por Patterson y Thompson (1971), plantean que el vector de los efectos fijos es eliminado de la función de verosimilitud, esto sirve para estimar los parámetros de covarianza,

$$\log[L(Y; Y - X\hat{\beta})] = K - \frac{1}{2} [\log|ZGZ^t + R| - \log|X^t V^{-1} X|^{-1} + (Y - X\hat{\beta})^t V^{-1} (Y - X\hat{\beta})]$$

$$\text{Siendo } \hat{\beta} = (X^t V^{-1} X)^{-1} X^t V^{-1} Y.$$

Este método maximiza la función de densidad después de ajustar los datos a los efectos fijos.

Diferencias entre **MINQUE** y **REML**: Searle (1995) y Gumedze y Dunne (2011) plantean que **MINQUE** es un procedimiento no iterativo, pero las ecuaciones de **REML** se utilizan de forma iterativa y esto tiene que comenzar con un valor inicial de σ^2 por lo tanto se dice que una estimación

MINQUE = una primera solución iterativa de REML.

Método Iterativo de Mínimos Cuadrados Generalizados (IGLS): Obtiene la estimación de los parámetros de los efectos fijos, minimizando el producto

$$(Y - X\beta)^t V^{-1} (Y - X\beta)$$

y la estimación de la varianza de los parámetros mediante la ecuación:

$$V = (Y - X\hat{\beta})(Y - X\hat{\beta})^t.$$

Este método puede ser usado repetidamente obteniendo el mismo resultado que se tendría usando el método de máxima verosimilitud.

Estimación simultanea de efectos fijos y aleatorios: Son un conjunto de ecuaciones llamadas Ecuaciones de Modelos Mixtos (**EMM**) desarrolladas por Henderson en el año 1984, cuya resolución genera simultáneamente el mejor estimador lineal insesgado BLUE ($X\beta$) y el mejor predictor lineal insesgado BLUP (U), la expresión matricial de EMM es:

$$\begin{bmatrix} \hat{\beta} \\ \hat{U} \end{bmatrix} = \begin{bmatrix} X^T R^{-1} X & X^T R^{-1} Z \\ Z^T R^{-1} X & Z^T R^{-1} Z + G^{-1} \lambda \end{bmatrix}^{-1} \begin{bmatrix} X^T R^{-1} y \\ Z^T R^{-1} y \end{bmatrix}$$

si G y R son conocidas, la solución para los efectos fijos y aleatorios es:

$$\hat{\beta} = (X^T V X)^{-1} X^T V y ; \hat{U} = G Z^T V^{-1} (y - X \hat{\beta})$$

dónde: $\hat{\beta}$ es la solución de los mínimos cuadrados generalizados para β así como su mejor estimador lineal insesgado (BLUE) y $\hat{U}$ es el mejor predictor lineal insesgado (BLUP) de U .

Este último constituye el predictor natural de los efectos aleatorios, presentando propiedades óptimas al momento de minimizar el error cuadrático medio de la predicción dentro del conjunto de los predictores insesgados, el BLUP es una regresión hacia la media general basada en los componentes de varianza y covarianza asociados a los efectos aleatorios del modelo, por su parte el BLUE asocia a la media de los factores de efectos fijos en un promedio realizado sobre todos los niveles del efecto en la población. Henderson (1984), asume que la distribución de (y) es normal

multivariada, las varianzas y covarianzas (**R** y **G**) se asumen conocidas y en la práctica se estiman por el método de máxima verosimilitud restringida (**REML**).

Gumedze y Dunne (2011), establecen diferencias entre el método de máxima verosimilitud (**ML**) y el método de máxima verosimilitud restringida (**REML**). Ellos señalan que en comparación con el **ML**, el **REML** incluye un término adicional para tomar en cuenta los grados de libertad perdidos en la estimación de los efectos fijos, otra diferencia importante entre ellos es que el **ML** es invariante en la reparametrización de los efectos fijos.

Máxima Verosimilitud (**ML**) y Máxima Verosimilitud Restringida (**REML**), son ahora los métodos estándar para la estimación de los parámetros de varianza para los datos balanceados y no balanceados.

Morante (2011), plantea que los modelos de efectos mixtos son herramientas poderosas para el análisis de datos longitudinales, donde se recolectan datos a través del tiempo en puntos o periodos, para hacer inferencias respecto al cambio, sus determinantes y consecuencias. En general los modelos lineales mixtos proveen una herramienta eficaz y flexible para el análisis de una variedad de datos, incluyendo datos agrupados, como datos longitudinales, medidas repetidas, datos bloqueados o multinivel, genéticos, etc.

3. MARCO METODOLÓGICO

La presente investigación comprende una revisión teórica y crítica sobre los modelos estadísticos mixtos, mostrando las propiedades y bondades del uso de estos modelos en la realidad práctica de los modelos mixtos a través de dos métodos de estimación y análisis en los cuales se basa la misma, como son el método de Mínimos Cuadrados y Máxima Verosimilitud.

3.1 Método de Mínimos Cuadrados:

El método clásico de ajuste de modelos lineales es el método de mínimos cuadrados desarrollado por Gauss en 1809 y Markoff en 1900. Los parámetros estimados por esta vía son insesgados y tienen la menor varianza de todos los posibles estimadores lineales; es decir, es el Mejor Estimador Lineal Insesgado (**BLUE**), es mejor porque maximiza la correlación entre el valor verdadero y el valor estimado de los efectos, minimizando la varianza del error, es lineal debido a que los factores para los que se requieren las estimaciones, son funciones lineales de las observaciones y por último se dice que es insesgado porque las estimaciones de los efectos fijos son tales que $E(\hat{\beta}) = \beta$, en donde β es el vector de parámetros desconocidos a estimar. Por lo tanto, el estimador mínimo cuadrado minimiza la suma de cuadrados residual. Salto (2008).

Está basado en el ajuste de variables a una función lineal

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \varepsilon$$

cuando los parámetros son desconocidos, donde el interés se concentra en minimizar la magnitud cuadrática de las diferencias entre el valor observado (y) y $(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)$, por lo tanto se denota:

$$\sum_{i=1}^n \varepsilon^2 = \sum_{i=1}^n (y - \beta_0 - \beta_1 x_1 - \dots - \beta_k x_k)^2$$

Derivando parcialmente con respecto a los parámetros β_k e igualando a cero se obtendrán las $(k+1)$ ecuaciones normales que bajo las restricciones del modelo establecido darán origen a las estimaciones de interés, Wackerly *et al.* (2002).

La estimación de los parámetros en los modelos lineales mixtos se hace más difícil que en el modelo lineal planteado arriba, no solo porque se tiene a β como en el modelo lineal general sino que además se tienen parámetros desconocidos en u , Balsells (2009), esto es:

$$Y_{(nx1)} = X_{(n \times p)} \beta_{(p \times 1)} + Z_{(n \times q)} U_{(q \times 1)} + \varepsilon_{(nx1)}$$

por lo tanto, este nuevo conjunto de parámetros de carácter aleatorio en el modelo, establece nuevas condiciones en el proceso de estimación por mínimos cuadrados, los cuales son necesario evaluar. Es por ello que, para este tipo de modelos, se hace uso del proceso de estimación por Mínimos Cuadrados Generalizados (**GLS**), siendo este más apropiado, ya que permiten obtener la estimación de los parámetros de los efectos fijos β , minimizando el producto:

$$(y - X\beta)^t (ZGZ^t + R)^{-1} (y - X\beta)$$

en donde G , y R ya fueron especificadas en el capítulo anterior, y la estimación de la varianza de los parámetros mediante la solución de las ecuaciones

$$(y - X\beta)(y - X\beta)^t = V$$

Balsells (2009), Gumedze y Dunne (2011), coinciden que a través del Método de Mínimos Cuadrados Generalizados se pueden estimar los efectos fijos y aleatorios en un modelo lineal mixto:

3.1.1 Estimación de los efectos fijos:

Asumiendo $\mathbf{G}$ y $\mathbf{R}$ conocidas por la teoría de Mínimos Cuadrados Generalizados (GLS), Graybill (1976), Medina (2007) y Villa (2008), plantean que el mejor estimador lineal insesgado se obtiene minimizando la expresión

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t(\mathbf{ZGZ}^t + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

De donde el estimador de mínimos cuadrados está dado por:

$$\hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{Y}$$

lo que implica que la $E(\hat{\boldsymbol{\beta}}_{GLS}) = \boldsymbol{\beta}$ y $V(\hat{\boldsymbol{\beta}}_{GLS}) = (\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X})^{-1}$.

3.1.2 Estimación de los efectos aleatorios:

López *et al.* (2007), Gumedze y Dunne (2011), plantean que el mejor predictor lineal insesgado (BLUP) de $\mathbf{U}$ está dado por:

$$\hat{\mathbf{U}}_{GLS} = (\mathbf{GZ}^t\mathbf{V}^{-1})(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

López *et al.* (2007) citando a Henderson (1984), Searle (1987), Gumedze y Dunne (2011), coinciden en que:

$$\mathbf{BLUP}(\mathbf{U}) = \hat{\mathbf{U}} = \mathbf{GZ}^t\mathbf{P}\mathbf{y}$$

Donde a $\hat{\mathbf{U}}$ se le conoce como el mejor predictor para $\mathbf{U}$, con $\mathbf{G} = \mathbf{Var}(\mathbf{U})$, $\mathbf{Z}$ matriz de diseño para las observaciones de efecto aleatorio, $\mathbf{X}$ matriz de diseño para las observaciones de efecto fijo e $\mathbf{y}$ como vector respuesta, además:

$$P = V^{-1} - V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}$$

Siendo en muchas situaciones la matriz de varianza covarianza de (Y),

$$V(Y) = V = (ZGZ^t + R)$$

La cual puede ser factorizada como:

$$V(Y) = \sigma^2 V = \sigma^2 (ZGZ^t + R).$$

Esto se logra con la adecuada modificación en las matrices **G** y **R**, para obtener el parámetro σ^2 , no necesariamente único, se puede realizar la factorización con solo dividir todos los elementos de **G** y **R** por σ^2 , estas versiones reparametrizadas permiten definir apropiadamente **V**, sin perder generalidad, Medina (2007). Un estimador insesgado de este puede ser obtenido por:

$$\hat{\sigma}^2 = \frac{1}{v_2} \hat{e}^t (ZGZ^t + R) \hat{e}$$

Donde:

$$\hat{e} = Y - [X(X^tV^{-1}X)^{-1}](X^tV^{-1}Y) \text{ y } v_2 = n - \text{rango}(X)$$

son los grados de libertad del error, si $X^tV^{-1}X$ no admite inversa, una inversa generalizada podría ser empleada suministrando estimabilidad de la función bajo consideración.

Se supone según Medina (2007), que **U** y **e** poseen una distribución normal multivariada, es decir, $Y \sim N(X\beta, \sigma^2 V)$, los estimadores de β y **U** se obtienen resolviendo las ecuaciones de modelos mixtos, Gumedze y Dunne (2011):

$$\begin{bmatrix} X^T R^{-1} X & X^T R^{-1} Z \\ Z^T R^{-1} X & Z^T R^{-1} Z + G^{-1} \lambda \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{U} \end{bmatrix} = \begin{bmatrix} X^T R^{-1} y \\ Z^T R^{-1} y \end{bmatrix}$$

determinadas por Henderson (1984), Medina (2007), a partir de la función de verosimilitud.

$$L = \frac{1}{(2\pi)^{n/2} \sqrt{|\mathbf{ZGZ}^t + \mathbf{R}|}} e^{-1/2(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-1}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})}$$

estas ecuaciones son una extensión de las ecuaciones normales donde $\hat{\boldsymbol{\beta}}$ y $\hat{\mathbf{U}}$. Son los estimadores de $\boldsymbol{\beta}$ y $\mathbf{U}$ respectivamente. Cuando las matrices $\mathbf{G}$ y/o $\mathbf{R}$ son desconocidas, la estimación de estas matrices puede realizarse mediante el uso de métodos de máxima verosimilitud (**ML**) o máxima verosimilitud restringida o residual (**REML**), bajo el supuesto de distribución normal conjunta multivariada de $\mathbf{U}$ y $\mathbf{e}$.

3.2 Método de Máxima Verosimilitud:

Según Searle (1995), el método de máxima verosimilitud fue desarrollado por Fisher en el año 1920, parece haber sido aplicado primero para la estimación de componentes de varianza. En este y en casi todas las presentaciones posteriores de este tema, la normalidad se asume para los términos del error y todos los efectos aleatorios, la normalidad con media cero, varianza homogénea de todos los efectos aleatorios relacionados con cada factor, y todas las covarianzas cero.

Wackerly *et al.* (2002), plantean que el método de máxima verosimilitud, elige como estimaciones los valores de los parámetros que maximizan la verosimilitud (función de probabilidad conjunta o la función de densidad conjunta) de la muestra observada. Estos estimadores son consistentes y si se les ajusta para que sean insesgados, a menudo proporcionan estimadores insesgados de varianza mínima. Como estos estimadores poseen buenas propiedades, este se hace un método muy común de estimación.

Sean las variables aleatorias $x_1, x_2, \dots, x_n$ las cuales constituyen una muestra aleatoria de una distribución discreta o continua, cuya función de probabilidad o

función de densidad de probabilidad es $f(\mathbf{X}/\boldsymbol{\theta})$, donde el parámetro $\boldsymbol{\theta}$ pertenece a un espacio paramétrico $\boldsymbol{\Omega}$. Así mismo, $\boldsymbol{\theta}$ puede ser un parámetro real o un vector de parámetros. Si para cualquier vector observado $x = (x_1, x_2, \dots, x_n)$ de la muestra, se define la función de probabilidad,

$$L(x_1, x_2, \dots, x_n; \boldsymbol{\theta}) = f(x_1; \boldsymbol{\theta}) \cdot f(x_2; \boldsymbol{\theta}) \dots f(x_n; \boldsymbol{\theta})$$

entonces el estimador Máxima Verosimilitud de $\boldsymbol{\theta}$, denotado por $\hat{\boldsymbol{\theta}}$, basado en la muestra aleatoria $x = (x_1, x_2, \dots, x_n)$ es el valor de $\boldsymbol{\theta}$ que maximiza a $L(x_1, x_2, \dots, x_n; \boldsymbol{\theta})$, donde $L(x_1, x_2, \dots, x_n; \boldsymbol{\theta})$ se le denomina “Función de Verosimilitud”.

Los estimadores de máxima verosimilitud de $\boldsymbol{\theta}$ y de $\mathbf{V}$ ($\mathbf{G}$ y $\mathbf{R}$), según Medina (2007), son obtenidos mediante la maximización de la función normal de verosimilitud

$$l(\boldsymbol{\theta}) = l(\mathbf{G}y\mathbf{R}) = -\frac{1}{2} \ln|\mathbf{V}| - \frac{N}{2} \ln \hat{\boldsymbol{\epsilon}}^t \mathbf{V}^{-1} \hat{\boldsymbol{\epsilon}} - \frac{N}{2} \left[1 + \ln \left(\frac{2\pi}{N} \right) \right]$$

donde:

$$\hat{\boldsymbol{\epsilon}} = \mathbf{Y} - [\mathbf{X}(\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1}] (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{Y})$$

con respecto a estos parámetros.

Lindstrom y Bates (1988), hacen una crítica de los estimadores **ML** para los componentes de varianza en el sentido que ellos son sesgados a la baja, debido a que no toman en cuenta los grados de libertad perdidos de la estimación de $\boldsymbol{\beta}$. El método **REML** corrige esto definiendo las estimaciones de los componentes de varianza como los maximizadores del logaritmo de la función de verosimilitud basándose en N-p contraste de error linealmente independiente, donde N es el número total de

observaciones de todos los individuos. Este logaritmo de verosimilitud, derivado por Harville (1977), es:

$$l_R(\hat{\beta}(\theta), \sigma, (\theta|y)) = -\frac{1}{2} \log|\sigma^{-2} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X}| + l_f(\hat{\beta}(\theta), \sigma, (\theta|y))$$

Donde:

- ☞ Logaritmo de verosimilitud completo (l_f)
- ☞ $l_f(\hat{\beta}(\theta), \sigma, (\theta|y)) = -1/2 \log|\sigma^2 \mathbf{V}| - 1/2 \sigma^2 (\mathbf{y} - \mathbf{X}\beta)^t (\mathbf{Z}\mathbf{G}\mathbf{Z}^t + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{X}\beta)$
- ☞ $\hat{\beta}(\theta) = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{Y}$

Los estimadores de Máxima Verosimilitud restringida o residual pueden considerarse, según Gianola (2002), un intento de tomar en cuenta la “pérdida de grados de libertad” resultante de estimar los efectos fijos, Patterson y Thompson (1971), argumentan que no se pierde información al usar verosimilitud restringida. Estos estimadores son obtenidos mediante la maximización de la función.

$$l_{REML}(\mathbf{G}, \mathbf{R}) = -\frac{1}{2} \ln|\mathbf{V}| - \frac{1}{2} \ln|\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X}| - \frac{N-r}{2} \left[1 + \ln\left(\frac{2\pi}{N-r}\right) \right],$$

donde $r = \text{rango}(\mathbf{X})$.

A continuación se describen cuatro procedimientos que se utilizan para el cálculo de los estimadores de varianza de máxima verosimilitud (**ML**) y máxima verosimilitud restringida (**REML**), Gumedze y Dunne (2011).

3.2.1 Algoritmo de Newton-Raphson (NR):

Este algoritmo asume la concavidad de la función de verosimilitud para obtener la aproximación cuadrática de la función. Cada iteración **NR** requiere el cálculo de la función estimada y su derivada. Tiene como desventaja que es inestable cuando se

encuentra lejos del máximo, sus ventajas son: converge rápidamente y genera errores estándar asintóticos para las estimaciones. El utiliza la expansión de primer orden de la función estimada alrededor de la estimación actual $\phi_{(m)}$ para producir la próxima estimación $\phi_{(m+1)}$.

Considerando la función de logaritmo de verosimilitud $l(\phi)$ para el cual se desea encontrar el máximo de ϕ con $\frac{\partial l(\phi)}{\partial \phi} = 0$.

El procedimiento **NR** se muestra a continuación:

Por la expansión de primer orden se tiene la ecuación vectorial:

$$\frac{\partial l(\phi)}{\partial \phi} = U(\phi) \approx U(\phi_{(0)}) + \frac{\partial^2 l(\phi)}{\partial \phi \partial \phi^t} (\phi - \phi_{(0)}),$$

igualando a cero y resolviendo se tiene:

$$\phi = \phi_{(0)} - \left[\frac{\partial^2 l(\phi)}{\partial \phi \partial \phi^t} \right]^{-1} U(\phi_{(0)}),$$

esta ecuación puede ser utilizada iterativamente para depurar el estimador máximo en la $(m+1)$ -ésima iteración, esto es:

$$\phi_{(m+1)} = \phi_{(m)} - \left[\frac{\partial^2 l(\phi)}{\partial \phi \partial \phi^t} \right]^{-1} U(\phi_{(m)}),$$

comenzando por el valor inicial $\phi_{(0)}$, y,

$$\left[\frac{\partial^2 l(\phi)}{\partial \phi \partial \phi^t} \right]^{-1}$$

es la información observada de la matriz evaluada en $\phi_{(m)}$, (Gumedze y Dunne, 2011).

3.2.2 Algoritmo de Fisher Scoring (FS):

Este algoritmo sustituye a la matriz de información observada por la matriz de información esperada

$$E \left[-\frac{\partial^2 l(\phi)}{\partial \phi \partial \phi^t} \right]$$

en el algoritmo **NR**. Converge fácilmente, genera errores estándar asintóticos para las estimaciones al igual que **NR** pero tiene como desventaja la tendencia a converger a valores fuera del espacio del parámetro, es inestable al igual que el **NR** cuando se encuentra lejos del máximo, Gumedze y Dunne (2011).

Patterson y Thompson (1971), citados por Gilmour *et al.* (1995), sugieren que los componentes de varianza se pueden resolver de forma iterativa utilizando el algoritmo de **FS**.

3.2.3 Algoritmo de Maximización De la Esperanza (EM):

Gumedze y Dunne (2011), plantean que Dempster, Laird, y Rubin en el año 1977 introdujeron el algoritmo **EM** para la estimación de parámetros en modelos con datos desbalanceados. Ellos demostraron cómo el algoritmo **EM** se puede utilizar para obtener la máxima probabilidad de las estimaciones de componentes de varianza en el modelo mixto lineal. Este algoritmo según Pazos *et al.* (2011), aumenta los datos observados de y con datos ocultos de z produciendo así los datos completos de x , elegidos para simplificar el procedimiento de estimación.

La aplicación de este método en este modelo se basa en la visualización de los efectos aleatorios como los datos observados o faltantes. Este algoritmo consiste en esencia de dos pasos: un paso de esperanza (E-paso) y una etapa de maximización (M-paso). Los pasos en la iteración $(m + 1)$ ésimas puede ser descrito como sigue:

E-paso. Usa la estimación m -ésima $\phi_{(m)}$, para evaluar el logaritmo de la verosimilitud de la distribución condicional $(u | y)$ y calcular la esperanza del log-verosimilitud para un nuevo valor de ϕ dada esta distribución condicional. Es decir el E-paso calcula la esperanza de la verosimilitud logarítmica $L(\theta; x)$ de los datos

completos con respecto a la distribución de los datos ocultos dados los datos observados y el parámetro θ en su valor actual $\theta^{(t)}$.

M-paso. Maximiza la esperanza del E-paso con respecto a ϕ para producir $\phi_{(m+1)}$. Este actualiza la estimación del parámetro θ maximizando la función calculada en el paso previo respecto a θ . El procedimiento se completa con **EM** iterando entre los pasos E y M hasta la convergencia.

Las desventajas de este método es que la convergencia es lenta y no da errores estándar para las estimaciones, más sin embargo es numéricamente estable.

Según Gilmour *et al.* (1995), algunos autores sugieren que el algoritmo **EM** como un procedimiento eficaz para la obtención de **REML** para estimar los componentes de varianza. Este utiliza las ecuaciones estimadas y requiere aproximadamente tres veces los cálculos necesarios para la evaluación de la probabilidad **REML** por iteración. Sin embargo, la convergencia aún puede ser lenta, aunque se han realizado varias mejoras al algoritmo **EM** originales.

3.2.4 Algoritmo de Información Promedio (AI):

Este algoritmo fue desarrollado por Gilmour, Thompson y Cullis en el año 1995, para la estimación de los parámetros de varianza en un modelo lineal mixto. Gumedze y Dunne (2011) plantean que este método puede considerarse como un algoritmo de puntuación modificado de Fisher ya que reemplaza la información esperada de la matriz en el algoritmo de **FS** con un promedio de la información de las matrices observada y esperada, llamadas la matriz de información promedio.

Esta matriz de información evita la evaluación de los términos de la traza de la matriz de información observada y esperada por la aproximación de los términos de la traza por sumas de cuadrados con los valores esperados correctos, es decir, el uso de la matriz de información promedio está motivado por la eficiencia computacional ya que

las sumas de los términos cuadrados son más fáciles de calcular que la traza de los términos. Al igual que el **NR** y algoritmos de **FS**, el algoritmo de **AI** se basa en encontrar una solución eficiente de las ecuaciones del modelo mixto. En cada iteración los valores actuales de ϕ se utilizan para resolver ecuaciones del modelo mixto.

Este método tiene la tendencia de converger a valores fuera del espacio del parámetro, tiene la ventaja de que es resistente a los pobres valores iniciales de **NR** y requiere menos iteraciones (más rápido) comparado con **EM**.

De los cuatro procedimientos que se utilizan para el cálculo de los estimadores de varianza descritos anteriormente, se tomará en cuenta para el presente estudio el algoritmo de Newton Raphson, ya que genera errores estándar asintóticos para las estimaciones, y estos van a ser utilizados para el criterio del mejor método de estimación, además el programa seleccionado para realizar el análisis de los datos (SAS) está basado en este.

Los procedimientos de interés de este programa serán el: *PROC MIXED* y *PROC GLM*.

Cadena y Castillo (2000), describen estos de la siguiente manera:

El procedimiento *MIXED* de SAS ajusta modelos lineales mixtos y calcula pruebas basadas en las sumas de cuadrados tipo III para los efectos fijos. Puede proporcionar estimaciones y hacer inferencias tanto de los factores fijos como de los aleatorios basándose en el modelo mixto de Henderson del año 1984. El enunciado **RANDOM** del procedimiento *MIXED* incorpora al modelo los efectos aleatorios que constituyen el vector de éstos en el modelo mixto. El *PROC MIXED* tiene tres métodos de estimación de componentes de varianza: **MIVQUE0**, **ML** y **REML**.

De los cuales para el análisis previsto en la presente investigación se tomarán en cuenta los métodos **ML** y **REML**.

El procedimiento *GLM* de SAS usa el método de mínimos cuadrados para ajustar modelos lineales generales. Puede manejar variables de clasificación y variables continuas, permite la especificación de cualquier grado de interacción y anidamiento de efectos. El *GLM* imprime los cuadrados medios esperados para cada uno de los diferentes tipos de sumas de cuadrados denominados como tipo I, tipo II, tipo III y tipo IV según lo solicite el usuario. Cuando se aplica a un conjunto de datos desbalanceados se tendría que escoger entre las sumas de cuadrados tipos II, III y IV, según el problema, así como los componentes de varianza que de éstas se puedan derivar.

Para Restrepo (2007), la suma de cuadrados se emplea con el fin de efectuar una descomposición de la variabilidad total atribuible a la variable respuesta (y), en los diferentes componentes o factores controlados o manipulados por el investigador, y la adición del error experimental, que constituye la fuente de variación que aglutina a todos los componentes no controlados dentro del modelo de clasificación experimental.

Existen varios tipos de sumas de cuadrados (SC) las cuales Restrepo (2007) define de la siguiente manera:

La suma de cuadrados tipo I, se usa para probar hipótesis sobre medias ponderadas, ajustadas o no, esta suma de cuadrado será igual a las sumas de cuadrados tipo II, III y IV cuando los diseños son balanceados. La suma de cuadrados tipo II se puede generar a partir de la tipo I, en la cual se escoge únicamente las hipótesis asociadas con medias ponderadas, esta suma de cuadrado será igual a la tipo III y IV si no existe interacción en el modelo.

La suma de cuadrados tipo III, se basan sobre medias no ponderadas, es altamente recomendada ante la presencia de no ortogonalidad. En el ANOVA se usa promedio muestral armónico para ajustar el total de la celda, no depende del orden del efecto, o de niveles. Estas SC se recomiendan cuando los datos son desbalanceados. Las sumas de cuadrado tipo III y IV tienen las mismas hipótesis para datos balanceados o no, y trabajan con promedios marginales, donde algunos promedios marginales no

son definidos. La SC tipo IV está referida a hipótesis sobre contrastes entre medias poblacionales de celdas ubicadas en la misma columna o fila.

El conjunto de datos reales que se analizarán en la presente investigación tienen la característica de ser desbalanceados y se asumirá efecto de la interacción, por lo cual el programa SAS y autores como Bancroft (1968) y Restrepo (2007) recomiendan que cuando los datos sean desbalanceados y se requiera una comparación de los efectos principales asumiendo presencia de interacción, la suma de cuadrados apropiada es la SC tipo III.

El criterio para decidir cuál es el mejor método para estimar el modelo lineal mixto será aquel que arroje el menor cuadrado medio del residual (CME). Posada y Rosero (2007), plantean que el CME es una medida que agrupa la variabilidad de aquellos factores que no tiene en cuenta el investigador. Toda vez que el CME corresponde a la varianza residual, los métodos seleccionados por su mayor capacidad de ajuste son aquellos que expresan el menor valor en este criterio. El proporciona una forma para elegir el mejor estimador: un CME mínimo a menudo, indica una variación mínima, y por lo tanto se considera un buen estimador.

Mood *et al.* (1974) plantean que cuando se quieren comparar estimadores examinando sus respectivos CME, naturalmente se preferiría uno con el CME pequeño o más pequeño, el cual se podría definir como mejor estimador.

Adicionalmente, como se ajustarán distintos modelos a un mismo conjunto de datos, será necesario seleccionar el más adecuado para estos, por lo tanto se utilizarán criterios para la comparación de los ajustes, según Pu y Niu (2006), existe en la literatura un considerable número de procedimientos para la selección de modelos entre los cuales destacan el criterio de información de Akaike (AIC) y el criterio de información de Schwarz o Bayesiano (BIC), los cuales serán descritos a continuación:

Gómez *et al.* (2012), plantean que:

Akaike desarrolló un método alternativo para la comparación de modelos, llamado criterio de información de Akaike (AIC). Este método permite determinar con qué eficiencia los modelos se ajustan a una base de datos. El criterio de selección es escoger modelos con valores más bajos de AIC. El modelo que mejor explica los datos con el mínimo número de parámetros es el que presenta más bajo valor de AIC.

El criterio de información de Schwarz (BIC), se denomina bayesiano por basarse en argumentos de la llamada estadística bayesiana. Los métodos bayesianos constituyen una alternativa a la estadística tradicional, que se basa en el contraste de hipótesis. Estos métodos se diferencian en que incorporan información externa al estudio. Con esta información y con los datos observados se estima una distribución de probabilidad para la magnitud efecto que se está investigando. El criterio para elegir el mejor modelo es el mismo que el de Akaike: el que tenga el menor valor de BIC.

Para García *et al.* (2012), los criterios AIC y BIC se utilizan para seleccionar la estructura de covarianza y media del modelo, estos criterios son funciones del logaritmo de verosimilitud y un término de penalidad basado en el número de parámetros del modelo. Jiang y Rao (2003), plantean que un trabajo pionero en los criterios de selección es el algoritmo de información de Akaike en el año 1972.

Cuadro 1. Criterios de Información generados por el programa SAS.

Criterio	Formula	Referencia
AIC	$-2l + 2d$	Akaike (1974)
AICC	$-2l + 2d \frac{n^*}{(n^* - d - 1)}$	Hurvich and Tsai (1989) Burnham and Anderson (1998)
HQIC	$-2l + 2d \log \log n$	Hannan and Quinn (1979)
BIC	$-2l + d \log n$	Schwarz (1978)
CAIC	$-2l + d(\log n + 1)$	Bozdogan (1987)

Fuente: SAS versión 9.1

Es decir:

$$AIC = -2L + 2d = -2L + 2(q + p)$$

$$BIC = -2L + d\log(n) = -2L + (q + p)\log(n)$$

Donde L es el máximo valor de la función de verosimilitud, d es la dimensión del modelo la cual para el estimador de máxima verosimilitud se denota como $(q + p)$; q es el número de parámetros de covarianza estimados y p es el rango de la matriz de diseño X . Bajo las expresiones antes descritas, el mejor modelo resulta ser aquel con menos valor para el indicador.

Para la presente investigación se realizaron dos ejemplos, el primero con datos simulados y el segundo con datos reales, ambos conjuntos de datos se analizaron mediante los métodos de Máxima Verosimilitud, Máxima Verosimilitud Restringida y Mínimos Cuadrados, con la finalidad de hacer la comparación entre los métodos de estimación a través del resultado del cuadrado medio del residual. El objeto del análisis con los datos simulados, es simplemente observar el comportamiento de los métodos de estimación aquí señalados, para efecto de validación de los mismos.

Los datos simulados se conformaron sobre una base de datos que incluyó dos niveles del factor A y dos del factor B, siendo este último aleatorio, con (3, 5 y 10) repeticiones de observación en cada combinación de tratamientos, con un total de diez experimentos por cada repetición, lo que equivale a (120, 200 y 400) observaciones para una variable respuesta, (y) fue producto de una combinación de la media fijada en 100 y los efectos fijos de A, y tres aleatorios conformados por: el error experimental, el efecto aleatorio de B y el efecto aleatorio de la interacción A*B, todos con media cero y varianza 25. Para generar el desbalanceo de estos datos se eliminó al azar el 10% de las observaciones por repetición lo que equivale a (2,2, y 4) respectivamente.

Para el segundo ejemplo se partió de una base de datos reales suministrada por el profesor Omar Verde, profesor jubilado de la Cátedra de Bioestadística y Genética

de la Facultad de Ciencias Veterinarias de la UCV, dicha base de datos estaba conformada por 9181 becerros divididos entre los años (2001-2012) a los cuales se les tomó tres pesos específicos: Peso al Nacer (PN), Peso al Destete (205 días), llamado (AJUSDE), y Peso a los 18 Meses (548 días), que se denotó como (AJUS18), los datos se revisaron detalladamente y se eliminaron aquellos padres que tenían un solo hijo o becerros que no tenían padres registrados, como existían meses de nacimiento con muy pocos datos a consecuencia de la temporada de monta que hace que los becerros no nazcan en todo el año sino en una temporada (no existiendo nacimientos en los meses del 6 al 8), se hicieron agrupaciones en los meses de nacimiento de los becerros, se tomaron en cuenta los meses del 1 al 4, poniendo en 4 los del mes 5, y 10 a 12, poniendo en 10 los del mes 9, de esta forma se asegura que no existan combinaciones sin datos.

Luego de esa exhaustiva revisión quedaron 6354 becerros divididos entre los años (2001-2012) para el estudio, a los cuales se les hizo un análisis estadístico de tipo descriptivo mediante el programa Statistix (Ver: **ANEXO H**).

Con el presente trabajo se busca aportar los elementos necesarios para la estimación del Modelo Lineal Mixto, en forma concreta y detallada de manera que pueda ser de referencia para los estudiantes del Postgrado en Estadística, Docentes y otros, ya que en nuestro entorno hay muy poca información y estudios relacionados con este tema.

4. ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS

El modelo lineal mixto para los datos simulados se define como:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

Dónde:

y_{ijk} = Variable respuesta de la k – ésima observacion, respecto al j
 – ésimo efecto de β e i – esimo efecto de α

μ = Media General

α_i = Es el i – ésimo efecto del factor $A, i = 1, 2$

β_j = Es el j – ésimo efecto aleatorio del factor $B, j = 1, 2$

$(\alpha\beta)_{ij}$ = Es la interacción de primer orden del j
 – ésimo efecto aleatorio del factor B respecto al i
 – ésimo efecto del factor A

ε_{ijk} = Efecto al azar propio de cada observación.

Su expresión matricial sería, como ya se dijo en capítulos anteriores:

$$Y_{(nx1)} = X_{(nxp)}\beta_{(px1)} + Z_{(nxq)}U_{(qx1)} + \varepsilon_{(nx1)}$$

en donde $\beta_{(px1)}$ está conformado por los parámetros μ, α_1 y α_2 , correspondientes a los efectos fijos y $U_{(qx1)}$ está conformado por los parámetros $\beta_1, \beta_2; (\alpha\beta)_{11}, (\alpha\beta)_{12}, (\alpha\beta)_{21}, (\alpha\beta)_{22}$, correspondientes a los efectos aleatorios. Los sub-índices n, p y q corresponderán a los valores indicados para i, j y k .

Los datos antes descritos fueron analizados (Ver: **ANEXO A**) a través del programa SAS versión 9.1 obteniendo los resultados que se muestran a continuación en lo referente al cuadrado medio del residual (CMResidual), criterio de Información de Akaike (AIC) y el criterio de Información Bayesiano de Schwarz (BIC), para cada

repetición, experimento y método de estimación: Máxima Verosimilitud (ML), Máxima Verosimilitud Restringida (REML) y Mínimos Cuadrados (GLM):

Cuadro 2. Resultados obtenidos por el programa SAS para datos simulados Balanceados (SC tipo III).

Exp.		3 Repeticiones			5 Repeticiones			10 Repeticiones		
		ML	REML	GLM	ML	REML	GLM	ML	REML	GLM
1	CMResidual	74.50	74.50	74.50	361.43	361.42	361.43	337.40	337.40	337.40
	AIC	104.1	86.6		185.7	168.7		362.8	345.6	
	BIC	97.5	82.7		180.5	166.1		356.3	341.7	
2	CMResidual	60.33	60.33	60.33	272.78	280.42	280.43	263.39	263.39	263.39
	AIC	99.4	83.4		180.8	166.1		351.1	334.3	
	BIC	92.8	79.5		175.6	162.2		345.9	331.7	
3	CMResidual	108.31	120.33	135.33	282.45	282.45	282.45	398.16	398.16	398.16
	AIC	100.7	85.7		183.2	166.5		368.6	351.5	
	BIC	95.5	83.1		176.7	162.5		362.0	347.6	
4	CMResidual	51.75	51.75	51.75	294.96	305.37	305.38	225.69	225.69	225.69
	AIC	96.1	81.2		181.2	167.0		346.4	330.2	
	BIC	89.6	77.2		176.0	163.1		339.9	326.3	
5	CMResidual	58.81	62.17	62.17	169.58	179.56	188.83	213.36	213.36	213.36
	AIC	96.3	82.9		172.3	156.2		347.3	329.6	
	BIC	91.1	79.0		167.1	153.6		340.7	325.7	
6	CMResidual	149.83	149.83	149.83	552.47	606.90	618.95	387.22	392.05	392.05
	AIC	107.6	91.2		189.0	175.2		362.8	349.3	
	BIC	101.1	87.3		185.1	172.6		357.6	345.4	
7	CMResidual	62.33	62.33	62.33	260.65	275.98	287.18	194.06	194.06	194.06
	AIC	102.4	85.1		177.3	162.1		341.0	324.8	
	BIC	95.9	81.1		172.1	159.5		334.5	320.9	
8	CMResidual	73.83	73.83	73.83	191.85	191.85	191.85	339.81	347.13	347.13
	AIC	99.2	83.1		177.4	160.4		359.6	345.4	
	BIC	94.0	80.5		170.8	156.5		354.3	341.5	
9	CMResidual	116.71	129.68	140.83	124.17	124.18	124.18	340.30	349.50	359.11
	AIC	99.8	85.5		168.3	152.4		358.3	342.9	
	BIC	94.6	82.9		161.8	148.5		353.1	340.2	
10	CMResidual	89.42	89.42	89.42	346.13	346.12	346.13	233.75	235.89	235.89
	AIC	105.5	88.0		185.2	169.0		343.7	330.7	
	BIC	98.9	84.1		178.6	165.1		338.5	326.7	

Análisis:

En el cuadro 2 se observan los resultados para los datos simulados balanceados (Ver: **ANEXOS B, D y F**), incluyendo tres tipos de repeticiones (3, 5 y 10), con 10 experimentos cada una, los cuadrados medios de los residuales obtenidos por los tres métodos de estimación **ML**, **REML**, **GLM**, y los criterios de información **AIC** y **BIC**.

En la repetición 3 se observa que en los experimentos (1, 2, 4, 6, 7, 8 y 10) los cuadrados medios del residual estimados por los tres métodos involucrados en el estudio (Máxima Verosimilitud, Máxima Verosimilitud Restringida y Mínimos Cuadrados) son iguales, sin embargo en los experimentos (3, 5 y 9) se observan diferencias entre los cuadrados medios de los residuales estimados, donde el menor de estos es obtenido a través del método de Máxima Verosimilitud.

En la repetición 5 se observa que los experimentos (1, 3 y 8) los cuadrados medios del residual estimados son iguales, mientras que en los experimentos (2, 4, 5, 6 y 7) se observan diferencias entre los cuadrados medios de los residuales, donde el menor de estos es obtenido a través del método de Máxima Verosimilitud, en el experimento 9 los cuadrados medios de los residuales son iguales entre los métodos de Máxima Verosimilitud Restringida y Mínimos Cuadrados, mientras que el cuadrado medio del método de Máxima Verosimilitud es menor, en el experimento 10 se observa que los cuadrados medios del residual obtenidos por el método de Máxima Verosimilitud y Mínimos Cuadrados son iguales, el generado por el método de Máxima Verosimilitud Restringida es un poco más pequeño que estos.

A través de la repetición 10 se puede observar que los experimentos (1, 2, 3, 4, 5 y 7) los cuadrados medios de los residuales son iguales, por otro lado los experimentos (6, 8 y 10) muestran diferencias entre los cuadrados medios de los residuales, donde el menor de estos fue obtenido a través del método de Máxima Verosimilitud, mientras que los obtenidos por los métodos de Máxima Verosimilitud Restringida y Mínimos Cuadrados son iguales, con el experimento 9 todos los cuadrados medios obtenidos son diferentes y el menor de estos fue generado por el método de Máxima Verosimilitud.

Autores como Cadena y Castillo (2000), coinciden que cuando los datos son balanceados es difícil escoger entre los métodos de máxima verosimilitud y máxima verosimilitud restringida, ya que al aplicarse máxima verosimilitud restringida a datos

balanceados proporciona resultados idénticos a las estimaciones del Análisis de Varianza, pero no proporciona estimaciones de los factores fijos. Por su parte, al aplicarse las estimaciones de máxima verosimilitud a datos balanceados no se obtienen resultados idénticos a los de Análisis de Varianza; pero máxima verosimilitud proporciona estimaciones de máxima verosimilitud de los factores fijos, los que, obviamente, serán BAN (*Best Asymptotically Normal*).

Según Li y Pourahmadi (2013), cuando se desean estimar parámetros de covarianza a menudo se prefiere el método de máxima verosimilitud restringida que el de máxima verosimilitud, especialmente cuando se tienen tamaños de muestras pequeños.

Para estos datos balanceados algunos resultados del cuadrado medio del residual fue menor al aplicar el método de estimación de Máxima Verosimilitud, sin embargo, este grupo de datos cumple con las características mencionadas por Cadena y Castillo (2000), Li y Pourahmadi (2013), en el sentido que las mejores estimaciones son obtenidas a través del método de Máxima Verosimilitud Restringida.

Cuadro 3. Resultados obtenidos por el programa SAS datos simulados Desbalanceados (SC tipo III).

Exp.		3 Repeticiones			5 Repeticiones			10 Repeticiones		
		ML	REML	GLM	ML	REML	GLM	ML	REML	GLM
1	CMResidual	55.49	56.15	56.39	310.09	309.68	309.68	349.39	349.77	349.90
	AIC	87.0	69.6		164.3	148.0		329.1	311.8	
	BIC	80.4	65.7		159.1	145.3		322.6	307.9	
2	CMResidual	71.42	73.23	73.94	281.50	293.84	293.84	271.06	271.33	271.43
	AIC	86.2	70.1		163.7	149.2		320.3	303.4	
	BIC	79.7	66.1		158.5	145.3		313.8	299.4	
3	CMResidual	117.90	134.74	156.97	310.93	310.66	310.66	412.25	412.75	412.92
	AIC	86.2	70.7		167.2	150.4		334.7	317.3	
	BIC	81.0	68.1		160.7	146.5		328.2	313.4	
4	CMResidual	62.48	65.10	65.97	116.35	124.08	132.84	234.93	235.18	235.26
	AIC	81.8	68.5		148.9	133.6		315.0	298.4	
	BIC	76.5	64.5		143.7	131.0		308.5	294.5	
5	CMResidual	59.34	67.55	68.44	187.93	200.43	213.54	221.77	221.90	221.94
	AIC	82.4	68.8		157.9	141.5		315.3	297.7	
	BIC	77.2	64.8		152.7	138.9		308.8	293.7	
6	CMResidual	172.94	180.26	183.53	600.53	675.60	704.35	337.44	347.61	348.08
	AIC	92.7	76.0		172.2	156.1		322.1	309.1	
	BIC	86.2	72.0		168.3	154.7		316.9	305.1	
7	CMResidual	57.72	58.40	58.64	288.29	307.20	315.94	212.13	212.37	212.44
	AIC	87.4	70.0		162.0	146.6		311.1	294.9	
	BIC	80.9			156.8	143.9		304.6	290.9	
8	CMResidual	67.99	69.65	70.31	215.56	215.47	215.47	334.20	334.72	334.90
	AIC	85.9	69.7		162.2	145.4		326.1	309.6	
	BIC	79.4	65.8		155.7	141.4		319.6	305.7	
9	CMResidual	129.20	156.91	166.17	131.75	131.71	131.71	370.60	381.81	393.64
	AIC	83.0	70.5		154.2	137.9		326.0	310.5	
	BIC	79.1	67.9		147.7	134.0		320.8	307.9	
10	CMResidual	99.04	100.79	101.44	257.76	257.65	257.65	238.43	238.95	239.10
	AIC	91.1	73.4		163.5	146.2		311.3	297.8	
	BIC	84.5	69.5		158.3	143.6		306.0	293.9	

Análisis:

En el cuadro 3 se observan los resultados para los datos simulados balanceados (Ver: **ANEXOS C, E y G**), incluyendo tres tipos de repeticiones (3, 5 y 10), con 10 experimentos cada una, los cuadrados medios de los residuales obtenidos por los tres métodos de estimación **ML**, **REML**, **GLM**, y los criterios de información **AIC** y **BIC**.

En los experimentos de la repetición 3 al igual que los de la repetición 10 se puede observar que los cuadrados medios de los residuales estimados por los métodos de Máxima Verosimilitud, Máxima Verosimilitud Restringida y Mínimos Cuadrados

presentaron diferencias entre sí, obteniéndose la menor estimación de estos a través del método de máxima verosimilitud.

En la repetición 5 se observa que los experimentos (2, 4, 5, 6, y 7) los cuadrados medios de los residuales obtenidos a través de los métodos de estimación antes mencionados presentan diferencias entre sí, siendo el menor de estos el obtenido a través del método de Máxima Verosimilitud. En los experimentos (1, 3, 8, 9, y 10) se observan diferencias entre los cuadrados medios generados por los métodos de estimación involucrados, en donde el cuadrado medio del residual obtenido por el método de Máxima Verosimilitud resultó un poco mayor a los restantes, esto pudo ser originado por el desbalanceo de los datos, para dichos experimentos.

En relación a los métodos para seleccionar el mejor modelo de predicción, **AIC**, **BIC**, se colocaron los valores obtenidos, pero no se hizo ningún comentario al respecto, ya que se usó un mismo modelo para efectos de este conjunto de datos simulados, tanto balanceados como desbalanceados.

Cadena y Castillo (2000), recomiendan que cuando los datos son desbalanceados el mejor método de estimación para estos, es el método de Máxima Verosimilitud, coincidiendo con los resultados obtenidos para estos datos desbalanceados simulados.

Ahora, se procede a aplicar los métodos de estimación a los datos reales y sus respectivos análisis, contribuyendo así a analizar desde esta perspectiva un estudio particular que se llevó a cabo en la Facultad de Veterinaria de la UCV.

La metodología estadística de los modelos lineales mixtos, permite estimar los efectos ambientales y predecir los valores genéticos de los animales de forma simultánea. Esto brinda la oportunidad de obtener resultados más confiables de estimadores de varianzas y parámetros genéticos que permitan predecir el valor genético de los animales y como consecuencia de ello, el progreso genético logrado en una determinada población Arias *et al.* (2007).

A continuación se muestra una breve descripción de los datos reales, que serán utilizados para el estudio y análisis de este trabajo:

Los años de nacimiento utilizados para este análisis se encuentran comprendidos entre el 2001 y el 2012. Cabe destacar que un mismo animal (becerro) fue medido en un año, es decir tuvo una sola repetición en el tiempo.

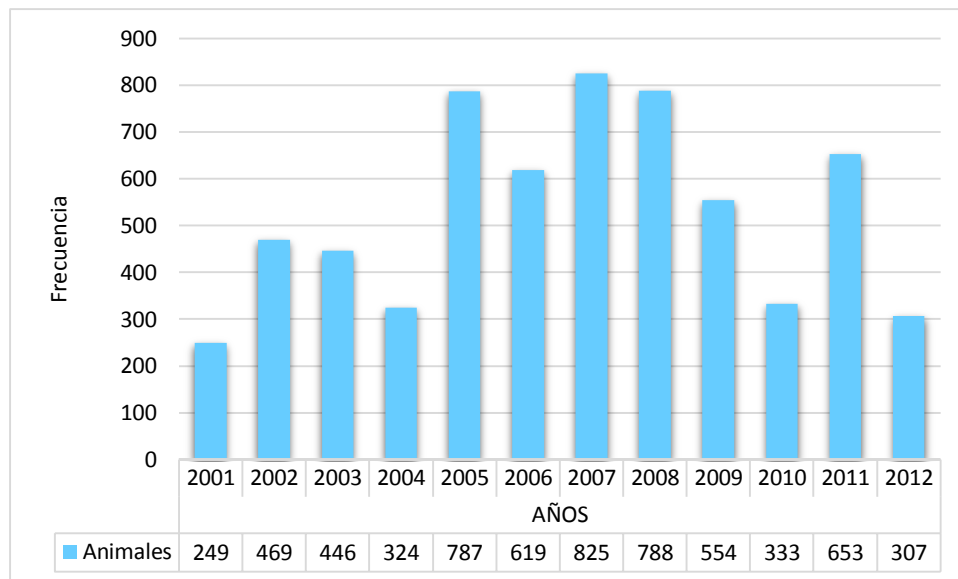


Gráfico 1. Distribución de los becerros por año de nacimiento.

Las edades de las madres al parto de cada becerro evaluado se encuentran comprendidas entre los 3 y 11 años, La edad promedio de estas es de 6.53 años, es decir de 7 años aproximadamente, el 25% y 75% de las madres tienen edades menores o iguales a los 5 años y 8 años, respectivamente. El 50% de estas madres se encuentra por debajo de los 6 años, quedando las edades de las mismas distribuidas de la siguiente manera:

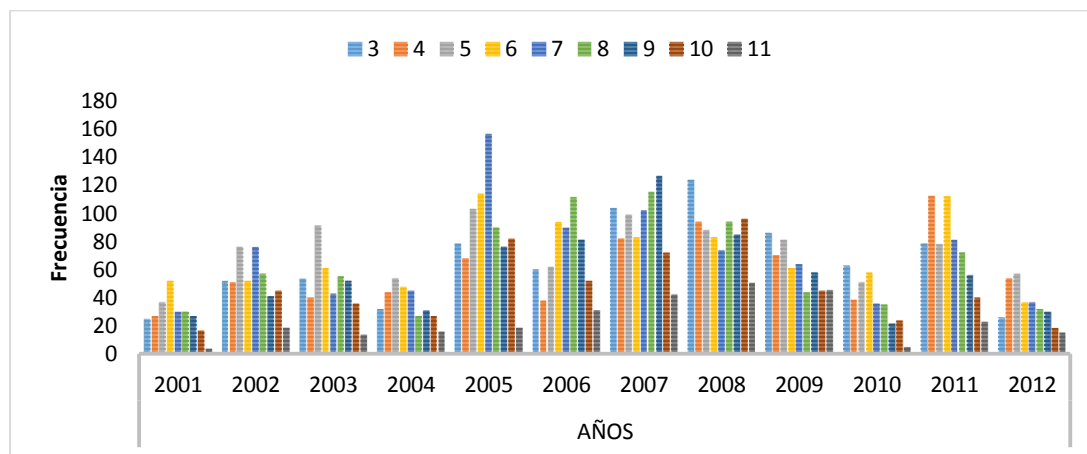


Gráfico 2. Edades de las madres al parto de cada becerro evaluado por año.

Los meses de nacimiento de los becerros seleccionados se encuentran entre los 1 y 4 meses y de los 10 a los 12 respectivamente. Los cuales quedaron distribuidos de la siguiente manera:

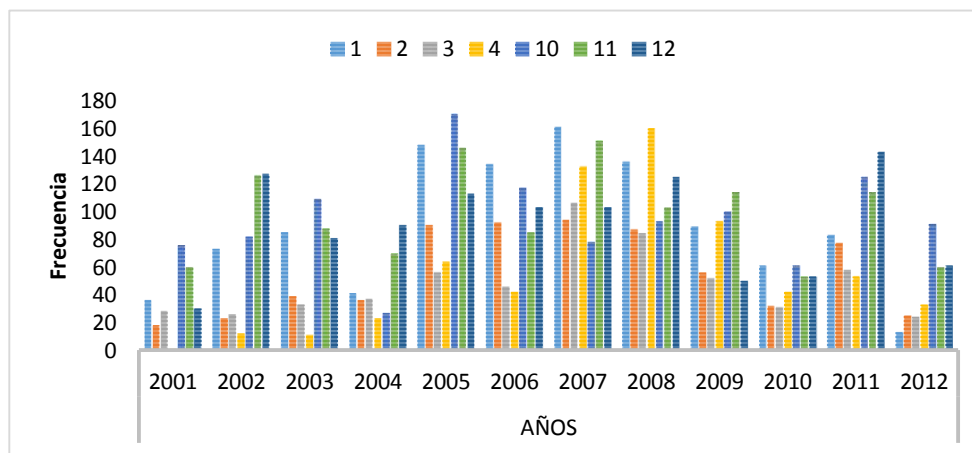


Gráfico 3. Meses de nacimiento de los becerros por año.

Se codificaron los sexos de los becerros quedando de la siguiente manera 1 para hembras y 5 para machos. En total se obtuvieron 2627 hembras y 3727 machos quedando:

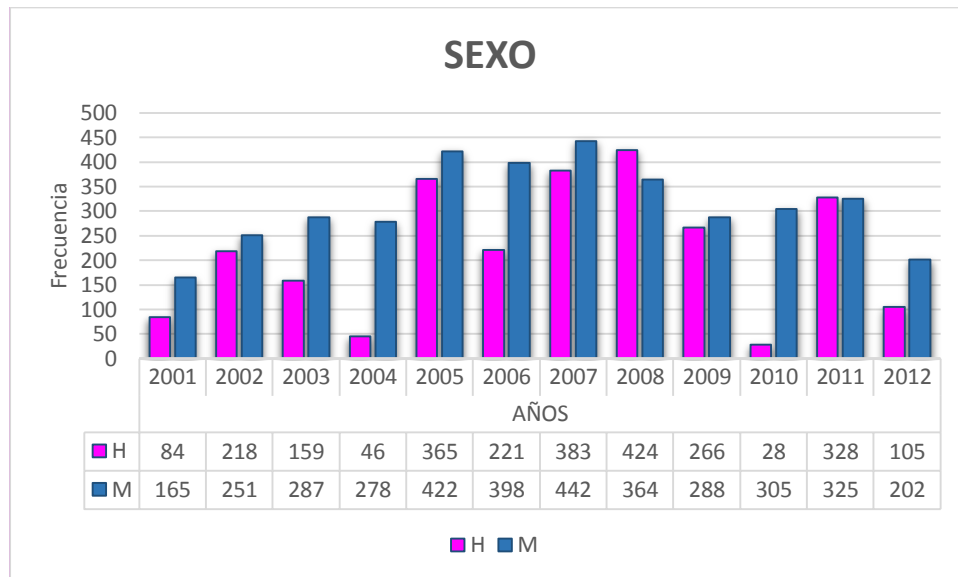


Gráfico 4. Sexo de los becerros por año.

Según Díaz *et al.* (2013), citando a Pariacote y Hhan (1985), el peso al nacer es un valor que indica el nivel de crecimiento de la cría en el período prenatal, su importancia demostrada radica en el hecho que crías con valores de peso adecuados se relacionan con mayores niveles de sobrevivencia durante sus primeras semanas de vida. Pesos al nacer muy bajos están relacionados con alta mortalidad en etapas tempranas de la vida del animal y por el contrario pesos con valores muy altos pueden producir problemas de partos distócicos.

Por lo tanto, el peso al nacer es un indicador que representa un punto crítico en el crecimiento de los animales, que está afectado en gran parte por la capacidad que tenga su madre para permitirle desarrollar su potencial para crecimiento dentro del útero.

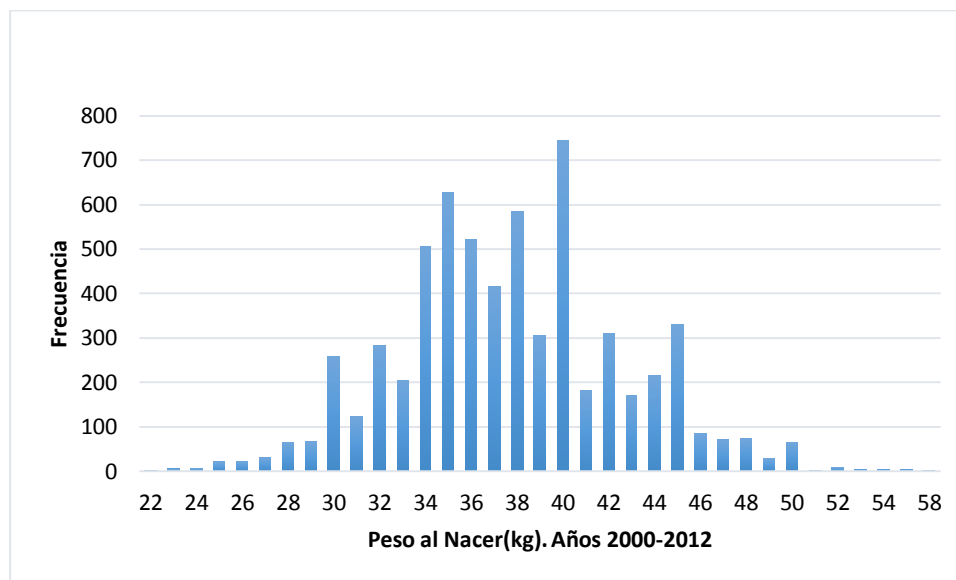


Gráfico 5. Pesos de los Becerros al Nacer.

Para los datos que se manejaron en el presente estudio, los pesos de los becerros al nacer están comprendidos entre los 21 kg y 58 kg. El peso promedio es de 37.524 kg, el 25% y 75% de los mismos son menores o iguales a los 31 kg y 40 kg respectivamente. El 50% de los pesos se encuentra por debajo de los 38 kg.

Por otra parte Díaz *et al.* (2013), plantean que el peso al destete es un valor que indica el crecimiento de la cría desde su nacimiento hasta el destete, momento en el que la alimentación del becerro cambia debido a la separación de su madre y pasa a depender de su capacidad para utilizar el forraje y otro tipo de alimentos, como los concentrados, para llenar sus requerimientos.

La importancia de su conocimiento radica en que permite medir tanto el potencial de crecimiento directo de un animal en particular, así como la habilidad materna de su madre y su nivel de producción lechera.

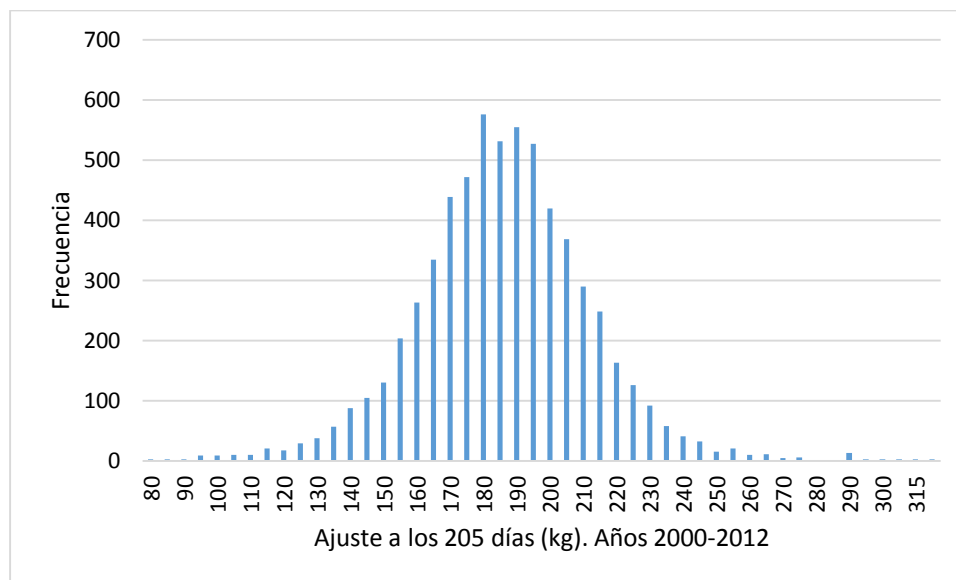


Gráfico 6. Pesos de los Becerros a los 205 días.

El peso de los becerros ajustado a los 205 días o pesos al destete se encuentra comprendido entre los 75 kg y 330 kg. El peso promedio es de 185.85 kg. El 25% y 75% de los mismos son menores o iguales a los 170.75 kg y 330 kg respectivamente. El 50% se encuentra por debajo de 186 kg.

Díaz *et al.* (2013), citando a Plasse (1985), plantean que a partir del destete los animales comienzan una etapa en la que son alimentados básicamente con pasturas y en algunos casos con alimento concentrado o cualquier otro recurso disponible, lo que permite evaluar su capacidad para crecer sin estar bajo el cuidado de sus madres, el peso postdestete se puede medir a diferentes edades, siendo los 18 meses de edad más común en los sistemas de producción de carne en el trópico, edad que permite la evaluación del crecimiento de cada individuo sin la influencia materna, y además permite la selección para su destino, sea como reproductor o animal de ceba y en las hembras para iniciar la etapa reproductiva.

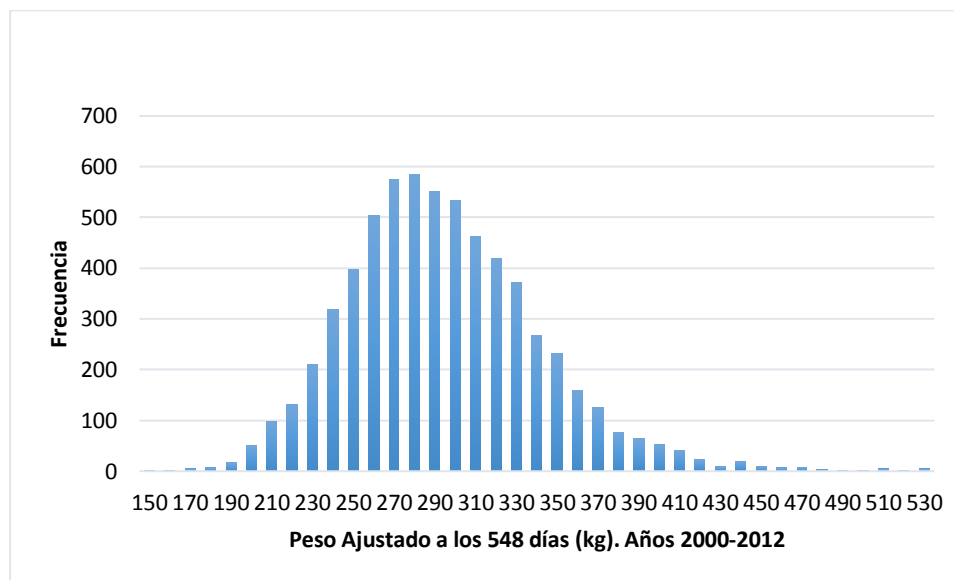


Gráfico 7. Pesos de los Beceros a los 548 días.

El peso de los becerros ajustado a los 548 días o pesos postdestete se encuentra comprendido entre los 148 kg y 658 kg. El peso promedio es de 294.01 kg. El 25% y 75% de los mismos son menores o iguales a los 261.75 kg y 322.00 kg respectivamente. El 50% se encuentra por debajo de 289 kg.

Se observa en los gráficos 5, 6 y 7, cómo dichos histogramas, correspondientes a las frecuencias de la variable peso de los animales, tienen un comportamiento de distribución casi normal. Los efectos fijos que se tomaron en cuenta para la presente investigación son: Año de Nacimiento, Mes de Nacimiento, Edad de la Madre al parto en años y Sexo del becerro, como efecto aleatorio se consideró el Padre (genético).

A través de la experiencia de los profesores de la Cátedra de Genética se seleccionaron modelos que proporcionarán mejores ajustes estadísticos para los efectos fijos y aleatorios que se analizarán mediante los métodos de estimación Máxima Verosimilitud, Máxima Verosimilitud Restringida y Mínimos Cuadrados quedando de la siguiente manera:

Modelo I: Es un modelo simple sin interacción.

$$y_{ijklmn} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \rho_m + \varepsilon_{ijklmn}$$

y_{ijklmn} = Variable Respuesta del n – ésimo becerro del m – ésimo padre, l
 – ésima edad de la madre, del k – ésimo sexo, j
 – ésimo mes de nacimiento, e i – ésimo año de nacimiento.

Donde la variable respuesta será (PN, AJUSDE, AJUS18).

μ = Media General.

α_i = Efecto del i – ésimo año de nacimiento; $i = 2001, \dots, 2012$

β_j = Efecto del j – ésimo mes de nacimiento; $j = 1, \dots, 4; 10, \dots, 12$

γ_k = Efecto del k – ésimo sexo; $k = 1, 5$

δ_l = Efecto de la l – ésima edad de la madre; $l = 3, \dots, 11$

ρ_m = Efecto aleatorio del m – ésimo padre; $m = 10, \dots, 870$ (Codificación de los padres, los mismos se numeraron de 10 en 10).

ε_{ijklmn} = Efecto al azar propio de cada observación.

Modelo II: Al modelo **I** se le adiciona una interacción de tipo fija conformada por el mes de nacimiento y el año, puesto que la misma tiene un impacto ambiental en el peso del becerro, quedando el modelo de la siguiente manera:

$$y_{ijklmn} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \alpha_i * \beta_j + \rho_m + \varepsilon_{ijklmn}$$

$\alpha_i * \beta_j$ = Efecto conjunto del i – ésimo año de nacimiento respecto al j
 – ésimo mes de nacimiento

Modelo III: Adicional a la interacción del tipo fija que se presenta en el modelo **II**, se agrega una de tipo aleatoria conformada por el padre que es de efecto aleatorio y el sexo del becerro que ha sido considerado fijo, quedando el modelo de la siguiente manera:

$$y_{ijklmn} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \alpha_i * \beta_j + \rho_m + \rho_m * \gamma_k + \varepsilon_{ijklmn}$$

$\rho_m * \gamma_k = \text{Efecto conjunto del } m - \text{ésimo padre respecto al } k - \text{ésimo sexo}$

Al igual que los datos simulados, los modelos **I**, **II** y **III**, pueden ser expresados matricialmente como:

$$Y_{(nx1)} = X_{(n \times p)}\beta_{(p \times 1)} + Z_{(n \times q)}U_{(q \times 1)} + \varepsilon_{(nx1)}$$

identificándose en cada caso cada una de sus componentes matriciales.

Cada modelo seleccionado tiene nueve salidas, tres por cada variable respuesta (PN, AJUSDE, AJUS18), una por el procedimiento **ML**, otra por el **REML** y una última por el **GLM**, con la finalidad de hacer la comparación entre los métodos de estimación a través del resultado del cuadrado medio del residual, y para escoger cuál de los tres modelos presentados en los ejemplos resultó el más adecuado, se tomarán en cuenta los resultados de los criterios de información generados por el programa SAS como son el **AIC** (Criterio de Información de Akaike) y el **BIC** (Criterio de Información Bayesiana de Schwarz), de este modo lo que se quiere verificar mediante este estudio es la eficiencia de los métodos de estimación antes mencionados. Los datos antes descritos fueron analizados a través del programa SAS versión 9.1 obteniendo los resultados que se muestran a continuación:

Cuadro 4. Resultados obtenidos por el programa SAS usando la variable Peso al Nacer (PN), SC (III).

	Modelo								
	I			II			III		
Método	ML	REML	GLM	ML	REML	GLM	ML	REML	GLM
CM _{Res}	20.75	20.83	20.83	19.72	20.01	20.01	19.66	19.95	19.95
AIC	37497.1	37476.6		37305.4	37066.1		37305.2	37065.6	
BIC	37567.5	37481.5		37536.4	37071.0		37538.6	37072.9	

Análisis:

En los tres modelos seleccionados (Ver: **ANEXO I**), se observa que el cuadrado medio del residual obtenido por el método de Máxima Verosimilitud es más pequeño

que los conseguidos por los métodos de Máxima Verosimilitud Restringida y Mínimos Cuadrados, estos últimos resultaron a su vez ser iguales entre sí.

Como se puede observar para la variable en estudio el cuadrado medio del residual va bajando en forma continua en los tres modelos para los tres métodos de estimación, lo que induce a que a medida que se introduzcan más componentes al modelo disminuye el cuadrado medio del residual. Cadena y Castillo (2000), citando a (Mood *et al.*, 1975), plantean que desde el punto de vista teórico los estimadores de máxima verosimilitud son los que reúnen las mejores propiedades ya que, si bien son sesgados, tienen la propiedad de ser BAN (*Best Asymptotically Normal*). Para ellos el ML y REML, son formas de estimaciones de máxima verosimilitud y por tanto tienen las mismas propiedades asintóticas para las estimaciones que cada uno de ellos produce.

Para Cadena y Castillo (2000), en los problemas de reproducción animal los datos suelen ser desbalanceados, con muchas observaciones, frecuentemente miles, y con muchos niveles de los factores, en ocasiones cientos; con lo que se fortalecen las propiedades asintóticas de ML, tanto para las estimaciones de los efectos fijos como de los aleatorios.

A través del método de Máxima Verosimilitud en cada uno de los modelos antes mencionados se obtiene el valor más bajo del cuadrado medio del residual, esto indica que existe una mejor estimación de acuerdo al criterio de decisión descrito en el marco metodológico.

Los criterios para seleccionar el mejor modelo AIC y BIC para esta variable respuesta tienen una característica a medida que el AIC disminuye el BIC también lo hace, ambos son criterios independientes el uno del otro, sin embargo en el modelo **III** el AIC sigue disminuyendo pero el BIC incrementa ligeramente, el investigador decidió por el criterio AIC, y para este último, el valor más bajo es el obtenido a través

del modelo **III**, por lo tanto el modelo con interacción fija y aleatoria es el que se adapta mejor para explicar los datos.

Cuadro 5. Análisis de Varianza para Peso al Nacer (PN), SC (III) y Estimadores de Parámetros de Covarianza.

Fuente de Variación	gl/num	gl/denom	F-Valor	Pr> F
AN	11	6095	16.67	<.0001
MN	6	6095	3.48	0.0019
SEXO	1	83	277.35	<.0001
EM	8	6095	11.78	<.0001
AN*MN	66	6095	5.03	<.0001
Componente de Varianza	Estimador	Valor Z	Pr Z	
σ^2 PADRE	1.25	4.64	<.0001	
σ^2 PADRE*SEXO	0.13	1.20	0.1145	

La variable (PN) tuvo un peso promedio ajustado de 36.9987 kg es decir 37 kg aproximadamente. Se puede observar a través del cuadro 5 que las SC tipo III para los efectos fijos arrojan un valor de F basadas en las esperanzas de los cuadrados medios (ECM), para los efectos AN (Año de Nacimiento), EM (Edad de la Madre), AN* MN (Año de Nacimiento* Mes de Nacimiento) el denominador apropiado fue el CMresidual, para el efecto del SEXO su denominador apropiado fue la interacción PADRE*SEXO. Los efectos aleatorios en el PROC MIXED del programa SAS se analizan a través de las estimaciones de los componentes de varianza, es decir, se realiza una estimación del componente de varianza analizado y adicionalmente se construye el intervalo al 95% de confianza correspondiente mediante la metodología de Wald.

Welham y Thompson (1997), el programa SAS versión 9.0 (2005), y Kraemer (2012) coinciden que el IC de Wald está basado en el valor de las estimaciones de los parámetros y su matriz de varianza-covarianza, se distribuye de forma asintótica como una ji-cuadrado, con una $H_0: \sigma^2 = 0$ y una $H_0: \sigma^2 > 0$. El p valor significativo implica que existe el componente de varianza y el efecto aporta información a la variable respuesta, si el p valor es no significativo implica que el IC de Wald presenta valores

muy bajos y eso implica que la varianza puede considerarse cero, es decir, no aporta información significativa a la variable respuesta.

Según el cuadro 5, se observa que el efecto del PADRE aporta información sobre la variable PN, sin embargo la interacción PADRE*SEXO no aporta información sobre la variable en estudio, con esta última se buscaba observar si existían diferencias entre los pesos de los hijos de los padres involucrados, lo que significa es que independientemente del sexo del descendiente, los padres más pesados tenderán a presentar descendencia más pesada, y los padres más livianos, descendencia más liviana, es decir, no hay un cambio de respuesta (peso) ya sea un becerro o una becerra.

Así padres pesados tenderán a tener becerros y becerras más pesados que los padres livianos, quienes tenderán a tener becerros y becerras más livianas.

Cuadro 6. Resultados obtenidos por el programa SAS usando la variable Peso al destete o a los 205 días (AJUSDE), SC (III).

	Modelo								
	I			II			III		
Método	ML	REML	GLM	ML	REML	GLM	ML	REML	GLM
CM _{Res}	567.96	570.16	569.98	527.85	535.44	535.30	527.85	535.43	535.73
AIC	58459.0	58351.0		58113.6	57571.5		58113.6	57571.5	
BIC	58529.5	58355.8		58344.5	57576.4		58344.5	57576.4	

Análisis:

Se observa en los modelos I y II, (Ver: **ANEXO J**), que el cuadrado medio del residual obtenido por el método de Máxima Verosimilitud es más pequeño que el conseguido por el método de Máxima Verosimilitud Restringida y que este último a su vez es algo mayor con respecto al cuadrado medio del residual del método de Mínimos Cuadrados, esto puede deberse a que los efectos involucrados en estos modelos generaron una mayor variabilidad y por ende el cuadrado medio del residual se vio afectado.

Por otra parte en el modelo **III** se puede observar que los cuadrados medios de los residuales obtenidos a través de los tres métodos de estimación seleccionados se encuentran de forma decreciente, donde el menor cuadrado medio del residual se obtuvo a través de la Máxima Verosimilitud.

Los criterios AIC y BIC, para seleccionar el mejor modelo, tienen una característica particular, a medida que el AIC disminuye el BIC también disminuye, esto solo ocurre entre los modelos **I** y **II**, con referencia al modelo **III** ambos criterios se mantienen iguales a los valores obtenidos a través del modelo **II**, es decir que cualquiera de estos modelos (**II** y **III**) pueden ser los más idóneos para explicar mejor los datos, ya que a través de estos se obtuvieron los valores más bajos, pero tomando en consideración que la interacción aleatoria PADRE*SEXO resultó no significativa, según el criterio del IC de Wald, el cual especifica que si el p valor es muy bajo eso implica que la varianza puede considerarse cero, es decir, no aporta información significativa a la variable respuesta, por lo antes expuesto el modelo **II** es el más idóneo para explicar los datos.

Cuadro 7. Análisis de Varianza para Peso al Destete (AJUSDE), SC (III) y Estimadores de Parámetros de Covarianza.

Fuente de Variación	gl/num	gl/denom	F-Valor	Pr> F
AN	11	6178	26.97	<.0001
MN	6	6178	18.78	<.0001
SEXO	1	6178	454.76	<.0001
EM	8	6178	16.73	<.0001
AN*MN	66	6178	7.53	<.0001
Componente de Varianza	Estimador	Valor Z	Pr Z	
σ^2 PADRE	8.39	3.02	0.0012	

La variable (AJUSDE) tuvo un peso promedio ajustado de 183.16 kg, es decir 183 kg aproximadamente. Las SC tipo III del cuadro 7, para los efectos fijos AN (Año de Nacimiento), MN (Mes de Nacimiento), SEXO, EM (Edad de la Madre) y AN* MN

(Año de Nacimiento*Mes de Nacimiento), arrojaron valores de F basados en las (ECM) significativos, siendo el denominador apropiado para cada uno de estos efectos el CMresidual. Se observa un p valor para el efecto del PADRE no significativo lo que implica que el mismo no aporta información sobre la variable respuesta.

Cuadro 8. Resultados obtenidos por el programa SAS usando la variable Peso a los 18 meses o 548 días (AJUS18), SC (III).

Método	Modelo								
	I			II			III		
	ML	REML	GLM	ML	REML	GLM	ML	REML	GLM
CM _{Res}	1455.04	1460.71	1459.41	1389.47	1409.54	1407.98	1374.66	1394.31	1390.84
AIC	64462.8	64328.6		64295.6	63662.1		64280.2	63646.1	
BIC	64533.3	64333.5		64526.5	63667.0		64513.5	63653.4	

Análisis:

En los tres modelos (Ver: **ANEXO K**), se puede observar que existen diferencias entre los cuadrados medios de los residuales estimados a través de los tres métodos, resultando que el cuadrado medio del residual del método de Máxima Verosimilitud Restringida se vio afectado por los tres modelos originando un cuadrado medio de residual más elevado que los obtenidos por el método de Mínimos Cuadrados y Máxima Verosimilitud, siendo el de este último el menor de estos.

En los modelos **II** y **III** al igual que el modelo **I** los cuadrados medios de los residuales obtenidos mediante Máxima Verosimilitud y Mínimos Cuadrados son más pequeños que los generados por la Máxima Verosimilitud Restringida, la diferencia entre estos (REML y GLM) es un poco más grande que la observada en el modelo **I**.

La variable (AJUS18) es la de mayor interés en cuanto al mejoramiento en los rebaños, a partir de este peso depende que animales seguirán en la finca para la reproducción.

Aquí se puede observar que los cuadrados medios de los residuales van disminuyendo progresivamente en cada modelo, independientemente del método utilizado. A través del método de Máxima Verosimilitud en cada uno de los modelos se obtiene el valor más bajo del cuadrado medio del residual.

Los criterios AIC y BIC, para seleccionar el mejor modelo, tienen una característica particular, a medida que el AIC disminuye el BIC también disminuye, por ende si se toma en consideración el valor más bajo de uno u otro criterio, el modelo **III** es el más idóneo para explicar mejor los datos.

Cuadro 9. Análisis de Varianza para Peso a los 18 Meses (AJUS18), SC (III) y Estimadores de Parámetros de Covarianza.

Fuente de Variación	gl/num	gl/denom	F-Valor	Pr> F
AN	11	6095	64.21	<.0001
MN	6	6095	1.45	0.1911
SEXO	1	83	1005.68	<.0001
EM	8	6095	1.76	0.0806
AN*MN	66	6095	4.61	<.0001
Componente de Varianza	Estimador	Valor Z	Pr Z	
σ^2 PADRE	27.49	2.04	0.0206	
σ^2 PADRE*SEXO	34.20	2.60	0.0047	

La variable (AJUS18) tuvo un peso promedio ajustado de 289.53 kg, es decir 290 kg aproximadamente. Se puede observar a través del cuadro 9, el ANAVAR arrojado por las SC tipo III, que para los efectos AN (Año de Nacimiento), MN (Mes de Nacimiento), EM (Edad de la Madre) y AN* MN (Año de Nacimiento* Mes de Nacimiento), el denominador apropiado fue el CMresidual, para el SEXO el denominador apropiado fue la interacción PADRE*SEXO, resultando significativos los efectos AN, SEXO y AN*MN. Se observa que los efectos del PADRE y PADRE*SEXO no significativos lo que implica que no aportan información sobre la variable respuesta, lo que implica que independientemente del sexo del descendiente, los padres más pesados tenderán a presentar descendencia más pesada, y los padres

más livianos, descendencia más liviana, es decir, no hay un cambio de respuesta (peso) ya sea un becerro o becerra. Así padres pesados tenderán a tener becerros y becerras más pesados que los padres livianos, quienes tenderán a tener becerros y becerras más livianas.

Como se pudo observar en los análisis de los datos desbalanceados simulados y reales que fueron seleccionados para la presente investigación, el método de Máxima Verosimilitud tuvo mejor comportamiento en cuanto al criterio de eficiencia seleccionado (Cuadrado Medio del Residual) en comparación con los Métodos de Máxima Verosimilitud Restringida y Mínimos Cuadrados, coincidiendo con Corbeil y Searle (1976) y Hocking and Kutner (1975), quienes en sus estudios resultó ser favorecido el Método de Máxima Verosimilitud en diseños de datos no balanceados.

A continuación se hace una breve descripción de las estimaciones obtenidas a través de los modelos antes descritos (Ver: **ANEXO L**):

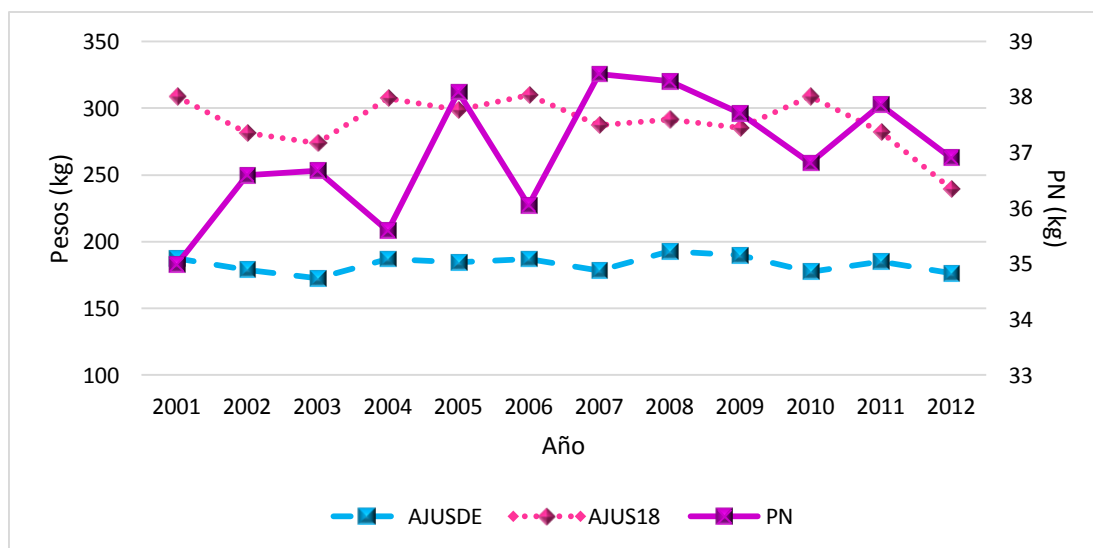


Gráfico 8. Promedio del peso al nacer (PN), peso corregido a los 205 días (AJUSDE) y peso corregido a los 548 días (AJUS18) según el año de nacimiento.

Para Díaz *et al.* (2013), el año de nacimiento (AN) está asociado tanto a factores climáticos como de manejo dentro del hato, que son diferentes en cada año y provocan variabilidad sobre el fenotipo de los animales, este resultado es común encontrarlo para caracteres de crecimiento en distintos sistemas de producción y con diferentes grupos genéticos.

En el gráfico 8, se muestran los comportamientos de los promedios obtenidos para los tres pesos según el año de nacimiento, siendo para la variable PN el año con mayor peso promedio del animal fue el 2007 y con menor peso promedio del animal el año 2000, con valores de 38.41 kg y 34.98 kg respectivamente.

Para la variable AJUSDE el año con mayor peso promedio del animal fue el 2008 y con menor peso promedio del animal el 2003 con valores de 192.94kg y 172.81kg respectivamente. Por su parte para el peso postdestete (AJUS18) se observa que el año con mayor peso promedio del animal fue el 2006 y con menor peso promedio del animal el año 2012 con valores de 309.57 kg y 239.74 kg respectivamente.

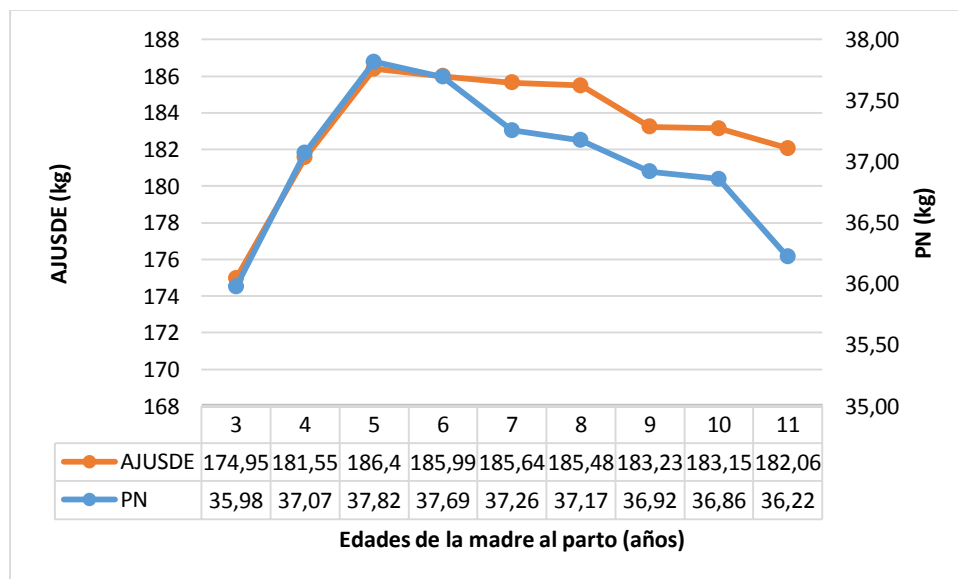


Gráfico 9. Promedio del peso al nacer (PN) y peso corregido a los 205 días (AJUSDE) según la edad de la madre al parto (años).

La edad de la madre resulto significativa en las variables PN y AJUSDE, según Díaz *et al.* (2013), el efecto de la edad de la madre sobre los pesos permite observar como comportamiento regular que madre jóvenes e inexpertas y madres de edad avanzada con deterioro fisiológico tienden a destetar becerros menos pesados, lo que tiene relación directa con los cuidados que ofrece la madre a su cría, su producción de leche y su madurez fisiológica.

En el gráfico 9 se observa que los mayores valores de PN y AJUSDE corresponden a madres de 5 años con 37.82 kg y 186.4 kg, respectivamente, los menores valores se obtuvieron con las madres de 3 años.

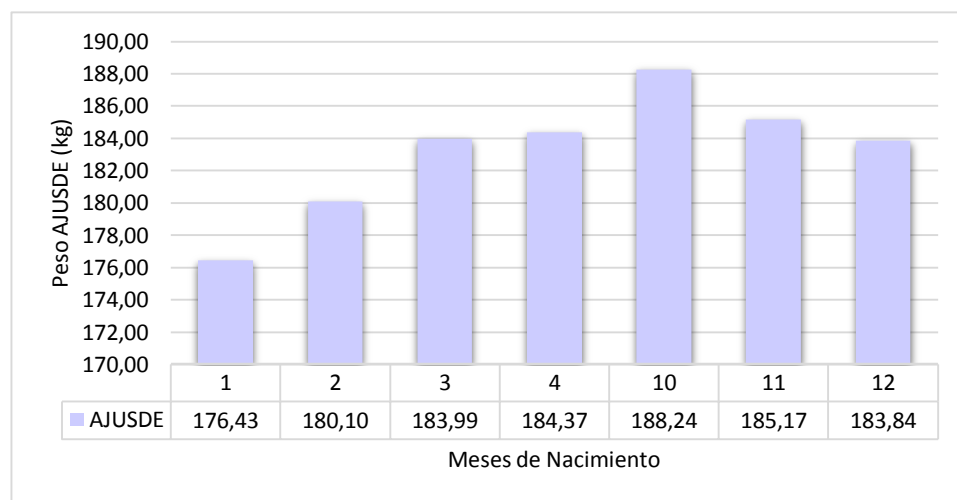


Gráfico 10. Promedio del peso a los 205 días (AJUSDE) según los meses de Nacimiento.

El MN (Mes de Nacimiento) resultó significativo solamente para la variable peso a los 205 días (AJUSDE), donde se puede observar que los meses con mayores promedios de peso fueron el 10 y el 11, es decir los meses de octubre y noviembre con 188.24 kg y 185.17 kg, respectivamente, el mes menos pesado fue enero con 176.43 kg.

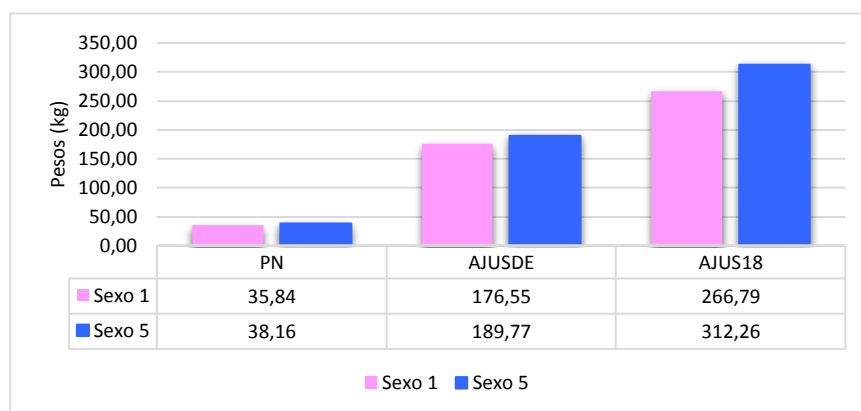


Gráfico 11. Promedio del peso al nacer (PN), peso corregido a los 205 días (AJUSDE) y peso corregido a los 548 días, según el sexo.

A través del gráfico anterior se observa que lo machos son más pesados que las hembras en las tres variables PN, AJUSDE y AJUS18.

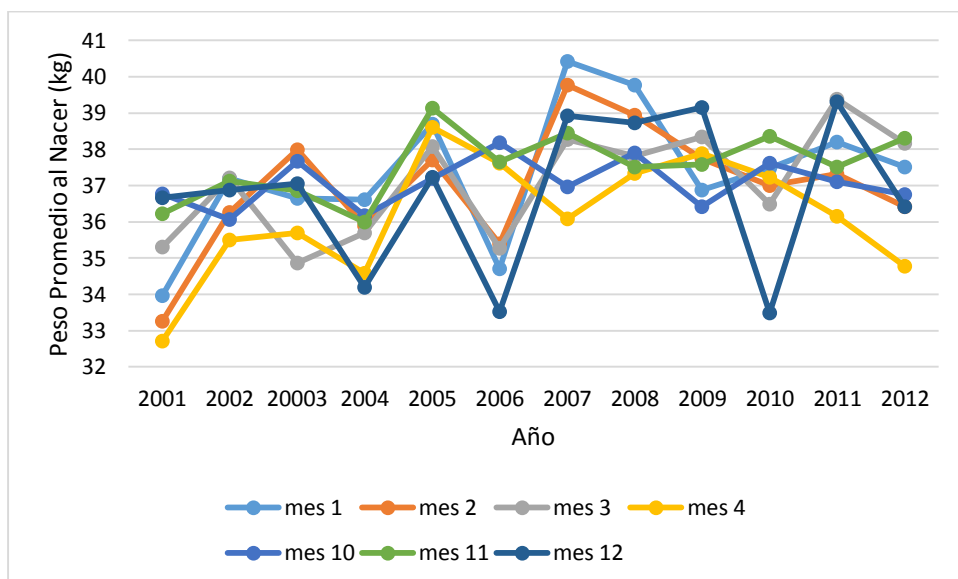


Gráfico 12. Promedio del peso al nacer (PN), para la interacción AN*MN.

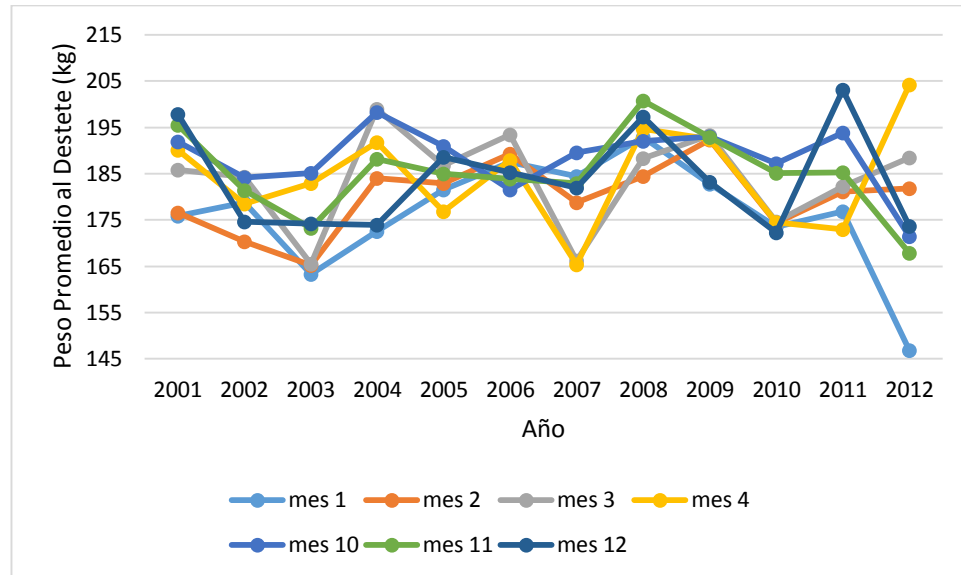


Gráfico 13. Promedio del peso al destete (AJUSDE), para la interacción AN*MN.

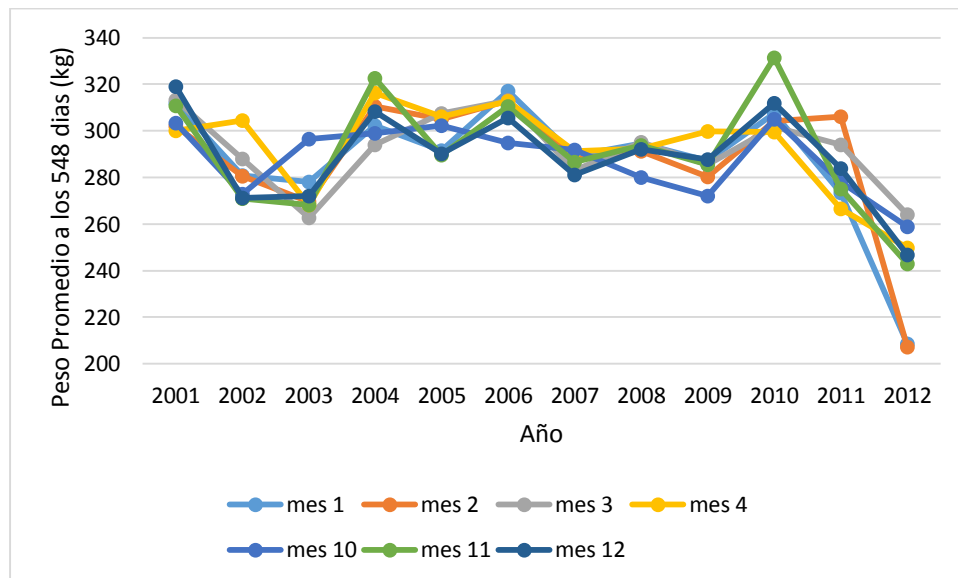


Gráfico 14. Promedio del peso a los 548 días (AJUS18), para la interacción AN*MN.

A través de los gráficos 12, 13 y 14 ver (**ANEXO M**), se observa que el comportamiento de los pesos para los diferentes meses a través de los distintos años no fue homogéneo, es decir hubo años en que un mes que tenía peso bajo presento peso alto y otro año que presento un peso intermedio, este comportamiento vario mucho de un año a otro, en el comportamiento de los meses no se puede esperar que sea el mismo a lo largo de los años, lo que puede ser observado en las tres variables respuestas, PN, AJUDE y AJUS18.

5. CONCLUSIONES Y RECOMENDACIONES

De los 30 experimentos simulados balanceados el 53% de estos generaron estimaciones de los cuadrados medios con igual valor a través de los Métodos Máxima Verosimilitud, Máxima Verosimilitud Restringida y Mínimos Cuadrados, mientras que el 47% restante mostraron que el cuadrado medio del residual estimado por el método de Máxima Verosimilitud fue menor que los obtenidos a través del método de Máxima Verosimilitud Restringida y Mínimos Cuadrados.

Se recomienda que cuando se esté presencia de datos balanceados y tamaños de muestras pequeñas se use el método de Máxima Verosimilitud Restringida, puesto que este genera las mejores estimaciones para este tipo de casos.

Los resultados arrojados por los datos desbalanceados tanto simulados como reales favorecieron al Método de Máxima Verosimilitud.

Se recomienda el uso del método de Máxima Verosimilitud específicamente para datos desbalanceados, específicamente relacionados con reproducción animal, puesto que este método proporcionó mejores estimaciones tanto para los efectos fijos como los aleatorios, en el estudio realizado.

Para este estudio en particular el mejor modelo seleccionado para explicar los datos reales según los criterios **AIC** y **BIC** fue el modelo III que incluía la interacción fija AN*MN y la aleatoria PADRE*SEXO, puesto que a través de este modelo en particular específicamente en las variables respuestas Peso al Nacer y Peso a los 548 días genero los valores más bajos en ambos criterios. Por lo tanto se recomienda que en casos similares a este estudio se utilice un modelo con las características antes descritas.

BIBLIOGRAFÍA.

- Arias, M.; Plasse, D., y Martínez, G. (2007). Variancias genéticas y no genéticas para pesos mensuales hasta los 18 meses en un rebaño Brahman. XXII Cursillo sobre bovinos de carne 18 -19 Octubre 2007. Facultad de Ciencias Veterinarias. UCV. Maracay, Venezuela. pp. 223-334.
- Balsells, M. (2009). Análisis longitudinal de los factores asociados a la pérdida de densidad mineral ósea en pacientes infectados con VIH. Presentado en la Universidad Politécnica De Cataluña, Barcelona, para optar al título: Master de Estadística e Investigación Operativa. Barcelona, España. pp. 23-27.
- Balzarini, M., Macchiavelli, R., y Casanoves, F. (2005). Aplicaciones de Modelos Mixtos en Agricultura y Forestería. Notas de Clase Curso de Posgrado CATIE, Turrialba, Costa Rica. pp. 3-12.
- Bancroft, A. (1968). Topics in Intermediate Statistical Methods. The Iowa State University Press Ames, Iowa, U.S.A. First Edition. pp. 3-30.
- Cadena, J., y Castillo, A. (2000). Una comparación de SAS y Harvey en la estimación de componentes de varianza en modelos mixtos. *Agrociencia*. 34 (1): 57-68. Colegio de Postgraduados Texcoco, México
- Calama, R., y Montero, G. (2004). Aplicación de los modelos mixtos a un caso práctico de modelización del crecimiento y producción de las masas forestales. *Sociedad Española de ciencias Forestales*. 18: 317-321. Grupo Silvicultura Mediterránea. CIFOR-INIA. Madrid, España.
- Corbeil, R., and Searle, R. (1976). A comparision of Variance Component Estimators. *Biometrics Unit, Cornell University, Ithaca, New York* 14853, U.S.A. *Biometrics*. 32: 779-791. New York, USA.
- Díaz, H., Salomón, J., y Romero, R. (2013). Crecimiento pre y postdestete de animales F₁ Holstein-Brahman en un hato de los llanos centrales de Venezuela. XXVIII Cursillo sobre bovinos de carne 24 y 25 de Octubre 2013. Facultad de Ciencias Veterinarias UCV. Maracay, Venezuela. pp. 219-262.
- Fox, J. (2002). Linear Mixed Models Appendix to an R and S-PLUS Companion to Applied Resgression. pp. 1-24.
<https://socserv.socsci.mcmaster.ca/jfox/Books/Companion-1E/appendix-mixed-models.pdf>.
- García, M., Rapelli, C., y Castellana, N. (2012). Usos de Criterios Predictivos para la selección de un Modelo Lineal Mixto. X Congreso Latinoamericano de Sociedades de Estadística Córdoba, Argentina. pp. 1-7.

- Gianola, D. (2002). Statistical Methods in the Genetic Improvement of Farm Animals. Departaments of Animal Sciences, Biostatistics and Medical Informatics, and of Dairy Science, University of Wisconsin-Madison. U.S.A. pp. 7-12.
- Gilmour, A., Thompson, R., and Cullis, B. (1995). Average Information REML: an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics*. 51: 1440–1450.
- Graybill, F. (1976). Theory and application of the linear model. Duxbury Press. Massachusetts, EEUU. pp. 704.
- Gómez, S., Torres, V., García, Y., y Navarro, J. (2012). Procedimientos estadísticos más utilizados en el análisis de medidas repetidas en el tiempo en el sector agropecuario. *Revista Cubana de Ciencia Agrícola*. 46 (1): 1-7.
- Gumedze, F., and Dunne, T. (2011). Parameter estimation and inference in the linear mixed model. *Linear Algebra and its Applications*. Departament of Statistical Sciences, University of Cape Town, South Africa. Elsevier.Inc. 435: 1920-1944.
- Harville, D. (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. *American Statistical Association*. 72 (358): 320-338.
- Henderson, C. (1953). Estimation of variance and covariance components. *Biometrics*. 9: 226–252.
- Henderson, C. (1984). Applications of Linear Models in Animal Breeding. University of Guelph. Ontario, U.S.A. pp. 1-384.
- Ho, R., and Hu, I. (2008). Flexible modelling of random effects in linear Mixed Models a Bayesian approach. *Computational Statistics & Data Analysis*. Departament of Statistics, Chinese University of Hong Kong, Shatin, Hong Kong. Elsevier.Inc. 52: 1347-1361.
- Hocking, R., and Kutner, M. (1975). Some analytical and numerical comparisons of estimators for the mixed A.OV. Model. *Biometrics*. 31: 19-27.
- Iyit, N., and Genc, A. (2011). Constitution of random intercept and slope model (RISM) as a special cas of linear mixed models (LMMs) for repeated measurements data. *Applied Mathematics and Computation* 218: 827-831.
- Jiang, J. and Rao, S. (2003). Consistent Procedures for Mixed Linear Model Selection. *Sankhya: The Indian Journal of Statistics*. 65 (1): 23-42.
- Kraemer, K. (2012). Confidence intervals for variance components and functions of variance components in the random effects model under non-normality.

- Graduate Theses and Dissertations. Paper 12893. Iowa State University. pp: 5-10.
- Laird, N., and, Ware, J. (1982). Random effects models for longitudinal data. *Biometrics*. 38: 963–974.
- Li, E., and Pourahmadi, M. (2013). An alternative REML estimation of covariance matrices in linear mixed models. *Statistics and Probability*. 83: 1071–1077.
- Liu, Q., Rong, J., Liu, Y. (2008). Best linear unbiased prediction for linear combinations in general mixed linear models. *Multivariate Analysis*. 99: 1503-1517.
- Lindstrom, M., and Bates, D. (1988). Newton-Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data. *Journal of the American Statistical Association*, 83 (404): 1014-1022. December 1988.
- López, L., Franco, D., y Barreto, P. (2007). Sobre la construcción del mejor predictor lineal insesgado (BLUP) y restricciones asociadas. *Revista Colombiana de Estadística*. 30 (1): 13-36.
- Medina, D. (2007). La aproximación general mejorada y el método de modelos mixtos en el análisis de experimentos con muestras pequeñas y diseño de medidas parcialmente repetidas. Trabajo de Grado como requisito para optar al grado de Magister Scientiarum en Estadística. Universidad Central de Venezuela. Maracay, Venezuela. pp. 56-62.
- Mood, A., Graybill, F., and Boes, D. (1974). *Introduction to the Theory of Statistics*. McGraw-Hill, USA, pp. 291-292.
- Morante, L. (2011). Métodos de regresión no paramétricos para el análisis de datos longitudinales. Trabajo fin de Master. Master Universitario en Estadística Aplicada. Universidad de Granada. p. 12-18
- Pariacote, F. y Hahn, M. (1985). Estimados parámetros genéticos en una población cruzada destinada a la producción simultanea de carne y leche. Índice de herencia para caracteres de carne en hembras. ALPA X reunión, Acapulco, México (Resumen).
- Patterson, H., and Thompson, R. (1971). Recovery of Inter-Block Information When BlockSizes are Unequal. *Biometrika*. 58: 545-554.
- Pazos, S., Hurtado, M., y Muravchik, C. (2011). Algoritmo con compresión óptima para Modelos Mixtos Lineales Ralos. Primeras Jornadas de Investigación y Tranferencia de la Facultad de Ingeniería UNLP. pp. 1-2.
- Pinheiro, J, y Bates, D. (2000) *Mixed –Effects Models in S and S-PLUS*. Editorial: Springer. Verlag New York, Inc.

- Plasse, D. (1985). La selección en la ganadería de carne. I Selección de toros. I Cursillo de Bovinos de Carne. Universidad Central de Venezuela, Facultad de Ciencias Veterinarias. Maracay, Venezuela. pp. 1-17.
- Posada, S., y Rosero, R. (2007). Comparación de modelos matemáticos: Una aplicación en la evaluación de alimentos para animales. *Revista Colombiana de Ciencias Pecuarias*. 20: 141-148.
- Pu, W., and Niu, F. (2006). Selecting mixed-effects models base don a generalixed information criterion. *Journal of Multivariate analysis*. 97: 733-758.
- Rencher, A., and Schaalje, B. (2008). *Linear Models In Statistics*. Wiley. John Wiley & Sons, Inc. Toronto, Ontario. USA. pp. 479-501.
- Restrepo, L. (2007). Tipos de suma de cuadrados en el análisis de varianza. *Revista Colombiana de Ciencias Pecuarias*. Universidad de Antioquia. Colombia. 20: 209-215.
- SAS Institute Inc, Cary, NC, USA. Copyright 2002-2003. Analytical Software. All Rights Reserved. Version 9.1.
- Salto, S. (2008). Variación Genética en Progenies de Polinización Abierta de *Eucalyptus Tereticornis* SMITH. Trabajo Final de Graduación. Universidad Nacional de Santiago del Estero. Facultad de Ciencias Forestales “Ing. Nestor Rene Ledesma”. pp. 16-22.
- Searle, S. (1995). An overview of variance component estimation. *Metrika* 42: 215-230.
- Searle, S., Casella, G., and McCulloch, C. (1992). *Variance Components*. 2 Ed. Editor J Willey. New York, pp. 501.
- Statistix 8. Copyright 1985-2003. Analytical Software. All Rights Reserved. Version 8.0.
- Villa, A. (2008). Comparación de los Métodos Mínimos Cuadrados Ordinarios y Máxima Verosimilitud para la Estimación de Parámetros en Diseños Factoriales pxq en Modelos Mixtos. Trabajo de Grado como requisito para optar al grado de Magister Scientiarum en Estadística. Universidad Central de Venezuela. Maracay, Venezuela. pp. 11-23.
- Wackerly, D., Mendenhall, W., y Scheaffer, R. (2002). *Estadística Matemática con Aplicaciones*. México. Thomson Learning, Sexta Edición. pp. 448, 537.
- Welham, S., and Thompson, R. (1997). Likelihood Ratio Test for Fixed Model Terms using Residual Maximum Likelihood. *Royal Statistical Society*. 59(3):701-714.

White, F., and, Hodge, G. (1989). Prediction Breeding Values with applications in forest tree improvement. Kluwer Academic Publishers. Dordrech, Netherlands. pp. 367.

ANEXOS

ANEXO A

Instrucciones programa SAS, métodos ML, REML y GLM.

<pre> Data prueba; Input exp A\$ B\$ Rep Y; cards; [INSERTE DATOS AQUÍ] ; proc sort data=prueba; by exp A B; run; proc glm data=prueba; by exp; class A B; model Y=A B A*B; random B A*B; TITLE 'Minimos Cuadrados'; run; proc sort data=prueba; by exp A B; run; proc Mixed data=prueba method=ML; by exp; class A B; model Y=A/SOLUTION; random B A*B/G SOLUTION; TITLE 'Maxima Verosimilitud'; run; proc sort data=prueba; by exp A B; run; proc Mixed data=prueba method=REML; by exp; class A B; model Y=A/SOLUTION; random B A*B/G SOLUTION; TITLE 'Maxima Verosimilitud Restringida'; run; </pre>	<pre> data A; INPUT IDENT AN MN SEXO EM PADRE PN AJUSDE AJUS18; RUN; PROC MIXED METHOD=ML CL COVTEST; CLASS PADRE AN MN SEXO EM; MODEL PN= AN MN SEXO EM; RANDOM PADRE; LSMEANS AN MN SEXO EM; RUN; PROC MIXED METHOD=REML CL COVTEST; CLASS PADRE AN MN SEXO EM; MODEL PN= AN MN SEXO EM; RANDOM PADRE; LSMEANS AN MN SEXO EM; RUN; PROC GLM; CLASS PADRE AN MN SEXO EM; MODEL PN= AN MN SEXO EM PADRE; RANDOM PADRE; LSMEANS AN MN SEXO EM; RUN; </pre>
--	---

ANEXO B

Salida del programa SAS (SC III), datos Balanceados, repetición 3.

		Repetición 3. BALANCEADOS					
Exp.		ML		REML		GLM	
1	σ^2B	117.11		234.22		$R^2=0.823285$	
	σ^2A*B	87.6667		200.17			
	CMResidual	74.500000		74.5000		74.500000	$\bar{Y}=124.6667$
	Verosimilitud -2 Log	94.1		80.6			
	AIC (mejor más pequeño)	104.1		86.6			
	AICC (mejor más pequeño)	114.1		90.6			
	BIC (mejor más pequeño)	97.5		82.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.06	0.8432	0.03	0.8880	0.29	0.6071
	B					27.92	0.0007
2	σ^2B	34.5833		69.1667		$R^2=0.716357$	
	σ^2A*B	44.1111		108.33			
	CMResidual	60.3333		60.333333		60.333333	$\bar{Y}=127.1667$
	Verosimilitud -2 Log	89.4		77.4			
	AIC (mejor más pequeño)	99.4		83.4			
	AICC (mejor más pequeño)	109.4		87.4			
	BIC (mejor más pequeño)	92.8		79.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.17	0.7491	0.09	0.8179	0.55	0.4786
	B					13.27	0.0066
3	σ^2B	43.2804		102.66		$R^2=0.586716$	
	σ^2A*B	0		0			
	CMResidual	108.31		120.33		135.333333	$\bar{Y}=125.1667$
	Verosimilitud -2 Log	92.7		81.7			
	AIC (mejor más pequeño)	100.7		85.7			
	AICC (mejor más pequeño)	106.4		87.4			
	BIC (mejor más pequeño)	95.5		83.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	7.39	0.2244	6.65	0.2355	5.91	0.0411
	B					5.44	0.0480
4	σ^2B	78.1111		156.22		$R^2=0.754691$	
	σ^2A*B	0.7500		18.7500			
	CMResidual	51.7500		51.7500		51.750000	$\bar{Y}=120.8333$
	Verosimilitud -2 Log	86.1		75.2			
	AIC (mejor más pequeño)	96.1		81.2			
	AICC (mejor más pequeño)	106.1		85.2			
	BIC (mejor más pequeño)	89.6		77.2			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.23	0.3758	1.11	0.4828	2.33	0.1658
	B					20.20	0.0020
5	σ^2B	132.20		268.89		$R^2=0.783226$	
	σ^2A*B	0		9.5278			
	CMResidual	58.8084		62.1667		62.166667	$\bar{Y}=120.7500$
	Verosimilitud -2 Log	88.3		76.9			
	AIC (mejor más pequeño)	96.3		82.9			
	AICC (mejor más pequeño)	102.0		86.9			
	BIC (mejor más pequeño)	91.1		79.0			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.04	0.8816	0.02	0.9043	0.03	0.8593
	B					27.41	0.0008
	A*B					1.46	0.2615

ANEXO B

Salida del programa SAS (SC III), datos Balanceados, repetición 3.

		Repetición 3 BALANCEADOS					
Exp.		ML		REML		GLM	
6	σ^2B	55.4167		110.83		$R^2=0.569548$	
	σ^2A*B	26.1111		102.17			
	CMResidual	149.83		149.83		149.833333	$\bar{Y}=125.6667$
	Verosimilitud -2 Log	97.6		85.2			
	AIC (mejor más pequeño)	107.6		91.2			
	AICC (mejor más pequeño)	117.6		95.2			
	BIC (mejor más pequeño)	101.1		87.3			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.04	0.8798	0.02	0.9145	0.06	0.8195
	B					7.48	0.0256
7	σ^2B	58.5000		117.00		$R^2=0.826304$	
	σ^2A*B	115.35		251.47			
	CMResidual	62.3333		62.3333		62.333333	$\bar{Y}=121.0833$
	Verosimilitud -2 Log	92.4		79.1			
	AIC (mejor más pequeño)	102.4		85.1			
	AICC (mejor más pequeño)	112.4		89.1			
	BIC (mejor más pequeño)	95.9		81.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.09	0.8145	0.04	0.8669	0.59	0.4646
	B					24.36	0.0011
8	σ^2B	0		0		$R^2=0.683966$	
	σ^2A*B	73.8882		172.39			
	CMResidual	73.8336		73.8336		73.833333	$\bar{Y}=119.5000$
	Verosimilitud -2 Log	91.2		79.1			
	AIC (mejor más pequeño)	99.2		83.1			
	AICC (mejor más pequeño)	104.9		84.9			
	BIC (mejor más pequeño)	94.0		80.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.33	0.6697	0.16	0.7557	1.30	0.2864
	B					6.87	0.0306
9	σ^2B	7.2320		31.7448		$R^2=0.250720$	
	σ^2A*B	0		0			
	CMResidual	116.71		129.68		140.833333	$\bar{Y}=122.8333$
	Verosimilitud -2 Log	91.8		81.5			
	AIC (mejor más pequeño)	99.8		85.5			
	AICC (mejor más pequeño)	105.5		87.2			
	BIC (mejor más pequeño)	94.6		82.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.14	0.7721	0.13	0.7829	0.12	0.7422
	B					2.27	0.1699
10	σ^2B	84.0278		168.06		$R^2=0.822395$	
	σ^2A*B	98.1944		226.19			
	CMResidual	89.4167		89.4167		89.416667	$\bar{Y}=119.8333$
	Verosimilitud -2 Log	95.5		82.0			
	AIC (mejor más pequeño)	105.5		88.0			
	AICC (mejor más pequeño)	115.5		92.0			
	BIC (mejor más pequeño)	98.9		84.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.00	0.3918	1.00	0.5000	8.59	0.0190
	B					19.87	0.0021
10	σ^2B	84.0278		168.06		$R^2=0.822395$	
	σ^2A*B	98.1944		226.19			
	CMResidual	89.4167		89.4167		89.416667	$\bar{Y}=119.8333$
	Verosimilitud -2 Log	95.5		82.0			
	AIC (mejor más pequeño)	105.5		88.0			
	AICC (mejor más pequeño)	115.5		92.0			
	BIC (mejor más pequeño)	98.9		84.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.00	0.3918	1.00	0.5000	8.59	0.0190
	B					19.87	0.0021
10	σ^2B	84.0278		168.06		$R^2=0.822395$	
	σ^2A*B	98.1944		226.19			
	CMResidual	89.4167		89.4167		89.416667	$\bar{Y}=119.8333$
	Verosimilitud -2 Log	95.5		82.0			
	AIC (mejor más pequeño)	105.5		88.0			
	AICC (mejor más pequeño)	115.5		92.0			
	BIC (mejor más pequeño)	98.9		84.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.00	0.3918	1.00	0.5000	8.59	0.0190
	B					19.87	0.0021

ANEXO C

Salida del programa SAS (SC III), datos Desbalanceados, repetición 3.

		Repetición 3. DESBALANCEADOS					
Exp.		ML		REML		GLM	
1	σ^2B	184.60		360.44		$R^2=0.894019$	
	σ^2A*B	54.6563		128.49			
	CMResidual	55.4936		56.1522		56.388889	$\bar{Y}=126.4000$
	Verosimilitud -2 Log	77.0		63.6			
	AIC (mejor más pequeño)	87.0		69.6			
	AICC (mejor más pequeño)	102.0		75.6			
	BIC (mejor más pequeño)	80.4		65.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.01	0.9281	0.01	0.9485	0.04	0.8434
	B					36.21	0.0010
2	σ^2B	15.4677		29.3886		$R^2=0.729620$	
	σ^2A*B	62.3409		144.43			
	CMResidual	71.4246		73.2294		73.944444	$\bar{Y}=127.1000$
	Verosimilitud -2 Log	76.2		64.1			
	AIC (mejor más pequeño)	86.2		70.1			
	AICC (mejor más pequeño)	101.2		76.1			
	BIC (mejor más pequeño)	79.7		66.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.20	0.7300	0.11	0.7984	0.61	0.4647
	B					7.14	0.0369
3	σ^2B	45.3102		110.81		$R^2=0.631015$	
	σ^2A*B	0		0			
	CMResidual	117.90		134.74		156.972222	$\bar{Y}=125.5000$
	Verosimilitud -2 Log	78.2		66.7			
	AIC (mejor más pequeño)	86.2		70.7			
	AICC (mejor más pequeño)	94.2		73.1			
	BIC (mejor más pequeño)	81.0		68.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	7.81	0.2188	6.83	0.2326	5.86	0.0518
	B					4.29	0.0838
4	σ^2B	81.5888		154.22		$R^2=0.737319$	
	σ^2A*B	1.41E-15		20.0067			
	CMResidual	62.4820		65.1005		65.972222	$\bar{Y}=121.9000$
	Verosimilitud -2 Log	73.8		62.5			
	AIC (mejor más pequeño)	81.8		68.5			
	AICC (mejor más pequeño)	89.8		74.5			
	BIC (mejor más pequeño)	76.5		64.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.06	0.4909	0.58	0.5844	1.00	0.3553
	B					11.90	0.0136
5	σ^2B	146.89		302.78		$R^2=0.801649$	
	σ^2A*B	1.19E-15		0.8606			
	CMResidual	59.3422		67.5478		68.444444	$\bar{Y}=119.6000$
	Verosimilitud -2 Log	74.4		62.8			
	AIC (mejor más pequeño)	82.4		68.8			
	AICC (mejor más pequeño)	90.4		74.8			
	BIC (mejor más pequeño)	77.2		64.8			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.14	0.7751	0.12	0.7911	0.12	0.7431
	B					20.48	0.0040
6	σ^2B	146.89		302.78		$R^2=0.801649$	
	σ^2A*B	1.19E-15		0.8606			
	CMResidual	59.3422		67.5478		68.444444	$\bar{Y}=119.6000$
	Verosimilitud -2 Log	74.4		62.8			
	AIC (mejor más pequeño)	82.4		68.8			
	AICC (mejor más pequeño)	90.4		74.8			
	BIC (mejor más pequeño)	77.2		64.8			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.14	0.7751	0.12	0.7911	0.12	0.7431
	B					20.48	0.0040
7	σ^2B	146.89		302.78		$R^2=0.801649$	
	σ^2A*B	1.19E-15		0.8606			
	CMResidual	59.3422		67.5478		68.444444	$\bar{Y}=119.6000$
	Verosimilitud -2 Log	74.4		62.8			
	AIC (mejor más pequeño)	82.4		68.8			
	AICC (mejor más pequeño)	90.4		74.8			
	BIC (mejor más pequeño)	77.2		64.8			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.14	0.7751	0.12	0.7911	0.12	0.7431
	B					20.48	0.0040

ANEXO C

Salida del programa SAS (SC III), datos Desbalanceados, repetición 3.

		Repetición 3 DESBALANCEADOS					
Exp.		ML		REML		GLM	
6	σ^2B	82.4344		150.98		$R^2=0.595813$	
	σ^2A*B	10.1262		75.4525			
	CMResidual	172.94		180.26		183.527778	$\bar{Y}=126.6000$
	Verosimilitud -2 Log	82.7		70.0			
	AIC (mejor más pequeño)	92.7		76.0			
	AICC (mejor más pequeño)	107.7		82.0			
	BIC (mejor más pequeño)	86.2		72.0			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.24	0.7114	0.13	0.7800	0.26	0.6315
	B					5.27	0.0614
	A*B					1.75	0.2335
7	σ^2B	46.0686		90.0204		$R^2=0.867413$	
	σ^2A*B	132.16		280.81			
	CMResidual	57.7246		58.3970		58.638889	$\bar{Y}=122.8000$
	Verosimilitud -2 Log	77.4		64.0			
	AIC (mejor más pequeño)	87.4		70.0			
	AICC (mejor más pequeño)	102.4		76.0			
	BIC (mejor más pequeño)	80.9					
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.00	0.9958	0.00	0.9970	0.00	0.9871
	B					19.36	0.0046
	A*B					12.18	0.0130
8	σ^2B	42.8930		81.6155		$R^2=0.752285$	
	σ^2A*B	42.1748		105.11			
	CMResidual	67.9847		69.6507		70.305556	$\bar{Y}=120.1000$
	Verosimilitud -2 Log	75.9		63.7			
	AIC (mejor más pequeño)	85.9		69.7			
	AICC (mejor más pequeño)	100.9		75.7			
	BIC (mejor más pequeño)	79.4		65.8			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.44	0.6264	0.23	0.7140	1.06	0.3420
	B					9.58	0.0213
	A*B					4.32	0.0829
9	σ^2B	1.14E-16		7.3363		$R^2=0.250601$	
	σ^2A*B	0		0			
	CMResidual	129.20		156.91		166.166667	$\bar{Y}=122.4000$
	Verosimilitud -2 Log	77.0		66.5			
	AIC (mejor más pequeño)	83.0		70.5			
	AICC (mejor más pequeño)	87.0		72.9			
	BIC (mejor más pequeño)	79.1		67.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.30	0.6822	0.24	0.7075	0.23	0.6477
	B					0.81	0.4021
	A*B					0.61	0.4644
10	σ^2B	32.6288		62.9595		$R^2=0.838930$	
	σ^2A*B	146.66		320.64			
	CMResidual	99.0374		100.79		101.444444	$\bar{Y}=120.9000$
	Verosimilitud -2 Log	81.1		67.4			
	AIC (mejor más pequeño)	91.1		73.4			
	AICC (mejor más pequeño)	106.1		79.4			
	BIC (mejor más pequeño)	84.5		69.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.60	0.4260	0.83	0.5299	7.11	0.0372
	B					11.11	0.0158
	A*B					8.24	0.0284

ANEXO D

Salida del programa SAS (SC III), datos Balanceados, repetición 5.

		Repetición 5. BALANCEADOS					
Exp.		ML		REML		GLM	
1	σ^2B	0		0		$R^2=0.354234$	
	σ^2A*B	86.1996		244.70			
	CMResidual	361.43		361.42		361.425000	$\bar{Y}=124.5500$
	Verosimilitud -2 Log	177.7		164.7			
	AIC (mejor más pequeño)	185.7		168.7			
	AICC (mejor más pequeño)	188.4		169.5			
	BIC (mejor más pequeño)	180.5		166.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.00	0.9646	0.00	0.9750	0.01	0.9354
	B					1.00	0.3323
2	σ^2B	160.39		333.06		$R^2=0.523087$	
	σ^2A*B	0		28.5550			
	CMResidual	272.78		280.42		280.425000	$\bar{Y}=124.0000$
	Verosimilitud -2 Log	172.8		160.1			
	AIC (mejor más pequeño)	180.8		166.1			
	AICC (mejor más pequeño)	183.5		167.9			
	BIC (mejor más pequeño)	175.6		162.2			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.73	0.3466	1.76	0.4113	2.65	0.1228
	B					13.39	0.0021
3	σ^2B	63.2700		126.54		$R^2=0.415898$	
	σ^2A*B	33.2900		123.07			
	CMResidual	282.45		282.45		282.450000	$\bar{Y}=117.5000$
	Verosimilitud -2 Log	173.2		160.5			
	AIC (mejor más pequeño)	183.2		166.5			
	AICC (mejor más pequeño)	187.5		168.2			
	BIC (mejor más pequeño)	176.7		162.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.35	0.6602	0.17	0.7480	0.56	0.4670
	B					7.66	0.0137
4	σ^2B	87.1416		190.96		$R^2=0.388853$	
	σ^2A*B	0		23.5650			
	CMResidual	294.96		305.37		305.375000	$\bar{Y}=123.4000$
	Verosimilitud -2 Log	173.2		161.0			
	AIC (mejor más pequeño)	181.2		167.0			
	AICC (mejor más pequeño)	183.9		168.7			
	BIC (mejor más pequeño)	176.0		163.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.20	0.4715	0.83	0.5289	1.16	0.2984
	B					7.64	0.0138
5	σ^2B	177.64		371.17		$R^2=0.565854$	
	σ^2A*B	0		0			
	CMResidual	169.58		179.56		188.825000	$\bar{Y}=115.4500$
	Verosimilitud -2 Log	164.3		152.2			
	AIC (mejor más pequeño)	172.3		156.2			
	AICC (mejor más pequeño)	175.0		157.0			
	BIC (mejor más pequeño)	167.1		153.6			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.09	0.8192	0.08	0.8240	0.08	0.7856
	B					20.61	0.0003
	σ^2B	0		0		$R^2=0.354234$	
	σ^2A*B	86.1996		244.70			
	CMResidual	361.43		361.42		361.425000	$\bar{Y}=124.5500$
	Verosimilitud -2 Log	177.7		164.7			
	AIC (mejor más pequeño)	185.7		168.7			
	AICC (mejor más pequeño)	188.4		169.5			
	BIC (mejor más pequeño)	180.5		166.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.00	0.9646	0.00	0.9750	0.01	0.9354
	B					1.00	0.3323

ANEXO D

Salida del programa SAS (SC III), datos Balanceados, repetición 5.

		Repetición 5. BALANCEADOS					
Exp.		ML		REML		GLM	
6	σ^2B	0		12.5135		$R^2=0.154335$	
	σ^2A*B	0		0			
	CMResidual	552.47		606.90		618.95000	$\bar{Y}=125.8500$
	Verosimilitud -2 Log	183.0		171.2			
	AIC (mejor más pequeño)	189.0		175.2			
	AICC (mejor más pequeño)	190.5		176.0			
	BIC (mejor más pequeño)	185.1		172.6			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.20	0.4714	1.09	0.4864	1.07	0.3167
	B					1.18	0.2929
7	σ^2B	24.3376		73.2195		$R^2=0.206151$	
	σ^2A*B	0		0			
	CMResidual	260.65		275.98		287.175000	$\bar{Y}=117.0000$
	Verosimilitud -2 Log	169.3		158.1			
	AIC (mejor más pequeño)	177.3		162.1			
	AICC (mejor más pequeño)	180.0		162.9			
	BIC (mejor más pequeño)	172.1		159.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.34	0.6646	0.32	0.6724	0.31	0.5871
	B					3.51	0.0794
8	σ^2B	40.2500		80.5000		$R^2=0.515132$	
	σ^2A*B	77.1500		192.67			
	CMResidual	191.85		191.85		191.850000	$\bar{Y}=125.6000$
	Verosimilitud -2 Log	167.4		154.4			
	AIC (mejor más pequeño)	177.4		160.4			
	AICC (mejor más pequeño)	181.6		162.1			
	BIC (mejor más pequeño)	170.8		156.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.25	0.7036	0.13	0.7827	0.76	0.3962
	B					10.22	0.0056
9	σ^2B	94.7200		189.44		$R^2=0.593785$	
	σ^2A*B	10.4450		45.7250			
	CMResidual	124.17		124.18		124.175000	$\bar{Y}=121.5000$
	Verosimilitud -2 Log	158.3		146.4			
	AIC (mejor más pequeño)	168.3		152.4			
	AICC (mejor más pequeño)	172.6		154.1			
	BIC (mejor más pequeño)	161.8		148.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.72	0.4143	0.86	0.5236	2.45	0.1371
	B					18.10	0.0006
10	σ^2B	19.8800		39.7600		$R^2=0.402383$	
	σ^2A*B	12.6950		94.6150			
	CMResidual	346.13		346.12		346.125000	$\bar{Y}=128.4000$
	Verosimilitud -2 Log	175.2		163.0			
	AIC (mejor más pequeño)	185.2		169.0			
	AICC (mejor más pequeño)	189.4		170.8			
	BIC (mejor más pequeño)	178.6		165.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	4.13	0.2910	2.07	0.3869	4.89	0.0419
	B					3.52	0.0792
	A*B					2.37	0.1435

ANEXO E

Salida del programa SAS (SC III), datos Desbalanceados, repetición 5.

		Repetición 5. DESBALANCEADOS					
Exp.		ML		REML		GLM	
1	σ^2B	0		0		$R^2= 0.322207$	
	σ^2A*B	44.5049		159.93			
	CMResidual	310.09		309.68		309.678571	$\bar{Y}= 120.8333$
	Verosimilitud -2 Log	156.3		144.0			
	AIC (mejor más pequeño)	164.3		148.0			
	AICC (mejor más pequeño)	167.4		148.9			
	BIC (mejor más pequeño)	159.1		145.3			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.06	0.8481	0.04	0.8766	0.16	0.6942
	B					1.78	0.2029
2	σ^2B	116.14		252.75		$R^2=0.438346$	
	σ^2A*B	0		21.3079			
	CMResidual	281.50		293.84		293.842857	$\bar{Y}=123.5556$
	Verosimilitud -2 Log	155.7		143.2			
	AIC (mejor más pequeño)	163.7		149.2			
	AICC (mejor más pequeño)	166.8		151.2			
	BIC (mejor más pequeño)	158.5		145.3			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.16	0.4760	0.90	0.5174	1.27	0.2794
	B					8.97	0.0097
3	σ^2B	38.0213		76.2750		$R^2=0.407466$	
	σ^2A*B	37.7552		146.19			
	CMResidual	310.93		310.66		310.657143	$\bar{Y}=117.3333$
	Verosimilitud -2 Log	157.2		144.4			
	AIC (mejor más pequeño)	167.2		150.4			
	AICC (mejor más pequeño)	172.2		152.4			
	BIC (mejor más pequeño)	160.7		146.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.60	0.5810	0.27	0.6960	0.74	0.4036
	B					5.27	0.0376
4	σ^2B	91.5550		195.65		$R^2=0.515304$	
	σ^2A*B	0		0			
	CMResidual	116.35		124.08		132.839286	$\bar{Y}=119.0556$
	Verosimilitud -2 Log	140.9		129.6			
	AIC (mejor más pequeño)	148.9		133.6			
	AICC (mejor más pequeño)	152.0		134.6			
	BIC (mejor más pequeño)	143.7		131.0			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.87	0.4020	1.82	0.4062	1.77	0.2046
	B					14.03	0.0022
5	σ^2B	183.08		386.35		$R^2=0.550079$	
	σ^2A*B	0		0			
	CMResidual	187.93		200.43		213.535714	$\bar{Y}=115.8333$
	Verosimilitud -2 Log	149.9		137.5			
	AIC (mejor más pequeño)	157.9		141.5			
	AICC (mejor más pequeño)	161.0		142.4			
	BIC (mejor más pequeño)	152.7		138.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.10	0.8036	0.11	0.7963	0.12	0.7343
	B					17.02	0.0010
	σ^2B	0		0		$R^2=0.550079$	
	σ^2A*B	0		0			
	CMResidual	187.93		200.43		213.535714	$\bar{Y}=115.8333$
	Verosimilitud -2 Log	149.9		137.5			
	AIC (mejor más pequeño)	157.9		141.5			
	AICC (mejor más pequeño)	161.0		142.4			
	BIC (mejor más pequeño)	152.7		138.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.10	0.8036	0.11	0.7963	0.12	0.7343
	B					17.02	0.0010
	σ^2B	0		0		$R^2=0.550079$	
	σ^2A*B	0		0			
	CMResidual	187.93		200.43		213.535714	$\bar{Y}=115.8333$
	Verosimilitud -2 Log	149.9		137.5			
	AIC (mejor más pequeño)	157.9		141.5			
	AICC (mejor más pequeño)	161.0		142.4			
	BIC (mejor más pequeño)	152.7		138.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.10	0.8036	0.11	0.7963	0.12	0.7343
	B					17.02	0.0010

ANEXO E

Salida del programa SAS (SC III), datos Desbalanceados, repetición 5.

		Repetición 5. DESBALANCEADOS					
Exp.		ML		REML		GLM	
6	σ^2B	0		1.03E-15		$R^2= 0.125785$	
	σ^2A*B	0		0			
	CMResidual	600.53		675.60		704.35357	$\bar{Y}= 124.8889$
	Verosimilitud -2 Log	166.2		154.1			
	AIC (mejor más pequeño)	172.2		156.1			
	AICC (mejor más pequeño)	174.0		156.3			
	BIC (mejor más pequeño)	168.3		154.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.78	0.5388	0.70	0.5574	0.85	0.3734
	B					0.94	0.3499
7	σ^2B	19.7068		70.5647		$R^2=0.207330$	
	σ^2A*B	0		0			
	CMResidual	288.29		307.20		315.935714	$\bar{Y}=117.6667$
	Verosimilitud -2 Log	154.0		142.6			
	AIC (mejor más pequeño)	162.0		146.6			
	AICC (mejor más pequeño)	165.1		147.5			
	BIC (mejor más pequeño)	156.8		143.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.19	0.7383	0.23	0.7159	0.28	0.6019
	B					2.96	0.1075
8	σ^2B	53.7121		107.57		$R^2=0.483636$	
	σ^2A*B	52.8879		154.58			
	CMResidual	215.56		215.47		215.471429	$\bar{Y}=124.3333$
	Verosimilitud -2 Log	152.2		139.4			
	AIC (mejor más pequeño)	162.2		145.4			
	AICC (mejor más pequeño)	167.2		147.4			
	BIC (mejor más pequeño)	155.7		141.4			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.30	0.6805	0.16	0.7544	0.75	0.3996
	B					8.63	0.0108
9	σ^2B	96.3738		192.90		$R^2=0.602178$	
	σ^2A*B	21.0465		71.8707			
	CMResidual	131.75		131.71		131.710714	$\bar{Y}=121.7778$
	Verosimilitud -2 Log	144.2		131.9			
	AIC (mejor más pequeño)	154.2		137.9			
	AICC (mejor más pequeño)	159.2		139.9			
	BIC (mejor más pequeño)	147.7		134.0			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.39	0.4483	0.72	0.5529	2.54	0.1334
	B					16.44	0.0012
10	σ^2B	0		0		$R^2= 0.571215$	
	σ^2A*B	118.52		295.57			
	CMResidual	257.76		257.65		257.653571	$\bar{Y}=130.5000$
	Verosimilitud -2 Log	155.5		142.2			
	AIC (mejor más pequeño)	163.5		146.2			
	AICC (mejor más pequeño)	166.6		147.2			
	BIC (mejor más pequeño)	158.3		143.6			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.41	0.3645	1.24	0.4658	7.81	0.0143
	B					5.89	0.0293
	A*B					6.31	0.0249

ANEXO F

Salida del programa SAS (SC III), datos Balanceados, repetición 10.

		Repetición 10. BALANCEADOS					
Exp.		ML		REML		GLM	
1	σ^2B	53.0100		106.02		$R^2= 0.374739$	
	σ^2A*B	90.2915		214.32			
	CMResidual	337.40		337.40		337.39722	$\bar{Y}= 124.2750$
	Verosimilitud -2 Log	352.8		339.6			
	AIC (mejor más pequeño)	362.8		345.6			
	AICC (mejor más pequeño)	364.6		346.4			
	BIC (mejor más pequeño)	356.3		341.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.16	0.7580	0.08	0.8247	0.59	0.4486
	B					13.64	0.0007
2	σ^2B	0		0		$R^2=0.416476$	
	σ^2A*B	113.15		252.64			
	CMResidual	263.39		263.39		263.39444	$\bar{Y}=120.4500$
	Verosimilitud -2 Log	343.1		330.3			
	AIC (mejor más pequeño)	351.1		334.3			
	AICC (mejor más pequeño)	352.3		334.7			
	BIC (mejor más pequeño)	345.9		331.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.85	0.5255	0.43	0.6319	4.51	0.0406
	B					10.21	0.0029
3	σ^2B	61.1375		122.28		$R^2=0.392406$	
	σ^2A*B	71.1889		182.19			
	CMResidual	398.16		398.16		398.16111	$\bar{Y}=120.6500$
	Verosimilitud -2 Log	358.6		345.5			
	AIC (mejor más pequeño)	368.6		351.5			
	AICC (mejor más pequeño)	370.3		352.2			
	BIC (mejor más pequeño)	362.0		347.6			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.14	0.3820	1.07	0.4895	5.96	0.0197
	B					11.72	0.0016
4	σ^2B	69.9200		139.84		$R^2=0.459010$	
	σ^2A*B	33.6111		89.7911			
	CMResidual	225.69		225.69		225.68889	$\bar{Y}=121.3000$
	Verosimilitud -2 Log	336.4		324.2			
	AIC (mejor más pequeño)	346.4		330.2			
	AICC (mejor más pequeño)	348.2		330.9			
	BIC (mejor más pequeño)	339.9		326.3			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	3.29	0.3207	1.65	0.4215	8.20	0.0070
	B					17.37	0.0002
5	σ^2B	138.88		277.76		$R^2=0.602729$	
	σ^2A*B	95.7094		212.75			
	CMResidual	213.36		213.36		213.35556	$\bar{Y}=124.9500$
	Verosimilitud -2 Log	337.3		323.6			
	AIC (mejor más pequeño)	347.3		329.6			
	AICC (mejor más pequeño)	349.0		330.3			
	BIC (mejor más pequeño)	340.7		325.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.21	0.4697	0.60	0.5792	6.64	0.0142
	B					37.01	<.0001
6	σ^2B	138.88		277.76		$R^2=0.602729$	
	σ^2A*B	95.7094		212.75			
	CMResidual	213.36		213.36		213.35556	$\bar{Y}=124.9500$
	Verosimilitud -2 Log	337.3		323.6			
	AIC (mejor más pequeño)	347.3		329.6			
	AICC (mejor más pequeño)	349.0		330.3			
	BIC (mejor más pequeño)	340.7		325.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.21	0.4697	0.60	0.5792	6.64	0.0142
	B					37.01	<.0001

ANEXO F

Salida del programa SAS (SC III), datos Balanceados, repetición 10.

		Repetición 10. BALANCEADOS					
Exp.		ML		REML		GLM	
6	σ^2B	65.7376		140.17		$R^2= 0.223007$	
	σ^2A*B	0		20.8572			
	CMResidual	387.22		392.05		392.05278	$\bar{Y}= 120.0750$
	Verosimilitud -2 Log	354.8		343.3			
	AIC (mejor más pequeño)	362.8		349.3			
	AICC (mejor más pequeño)	364.0		350.0			
	BIC (mejor más pequeño)	357.6		345.4			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.12	0.7882	0.08	0.8277	0.12	0.7333
	B					8.68	0.0056
7	σ^2B	77.0000		154.00		$R^2=0.429302$	
	σ^2A*B	34.6739		88.7539			
	CMResidual	194.06		194.06		194.06111	$\bar{Y}=124.7500$
	Verosimilitud -2 Log	331.0		318.8			
	AIC (mejor más pequeño)	341.0		324.8			
	AICC (mejor más pequeño)	342.8		325.5			
	BIC (mejor más pequeño)	334.5		320.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.02	0.9055	0.01	0.9329	0.06	0.8042
	B					21.44	<.0001
8	σ^2B	181.05		375.41		$R^2=0.414235$	
	σ^2A*B	0		6.8894			
	CMResidual	339.81		347.13		347.13056	$\bar{Y}=120.4750$
	Verosimilitud -2 Log	351.6		339.4			
	AIC (mejor más pequeño)	359.6		345.4			
	AICC (mejor más pequeño)	360.7		346.1			
	BIC (mejor más pequeño)	354.3		341.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.46	0.4398	1.19	0.4717	1.43	0.2393
	B					22.83	<.0001
9	σ^2B	87.0246		190.59		$R^2=0.284694$	
	σ^2A*B	0		0			
	CMResidual	340.30		349.50		359.10556	$\bar{Y}=120.6500$
	Verosimilitud -2 Log	350.3		338.9			
	AIC (mejor más pequeño)	358.3		342.9			
	AICC (mejor más pequeño)	359.5		343.2			
	BIC (mejor más pequeño)	353.1		340.2			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.88	0.3390	2.80	0.3427	2.73	0.1072
	B					11.59	0.0016
10	σ^2B	77.1121		158.13		$R^2=0.371295$	
	σ^2A*B	0		15.4739			
	CMResidual	233.75		235.89		235.88611	$\bar{Y}=117.9750$
	Verosimilitud -2 Log	335.7		324.7			
	AIC (mejor más pequeño)	343.7		330.7			
	AICC (mejor más pequeño)	344.9		331.4			
	BIC (mejor más pequeño)	338.5		326.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	4.58	0.2782	2.74	0.3458	4.54	0.0400
	B					15.06	0.0004
	A*B					1.66	0.2064

ANEXO G

Salida del programa SAS (SC III), datos Desbalanceados, repetición 10.

		Repetición 10. DESBALANCEADOS					
Exp.		ML		REML		GLM	
1	σ^2B	81.9781		162.09		$R^2= 0.404320$	
	σ^2A*B	77.9927		192.19			
	CMResidual	349.39		349.77		349.90139	$\bar{Y}= 124.2500$
	Verosimilitud -2 Log	319.1		305.8			
	AIC (mejor más pequeño)	329.1		311.8			
	AICC (mejor más pequeño)	331.1		312.6			
	BIC (mejor más pequeño)	322.6		307.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.16	0.7549	0.08	0.8243	0.46	0.5037
	B					14.03	0.0007
2	σ^2B	2.6854		5.3129		$R^2=0.448095$	
	σ^2A*B	118.08		263.20			
	CMResidual	271.06		271.33		271.42743	$\bar{Y}=120.3056$
	Verosimilitud -2 Log	310.3		297.4			
	AIC (mejor más pequeño)	320.3		303.4			
	AICC (mejor más pequeño)	322.3		304.2			
	BIC (mejor más pequeño)	313.8		299.4			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.97	0.5051	0.48	0.6135	4.60	0.0397
	B					9.91	0.0035
3	σ^2B	34.9385		68.9868		$R^2=0.409912$	
	σ^2A*B	116.18		274.22			
	CMResidual	412.25		412.75		412.91832	$\bar{Y}=121.2222$
	Verosimilitud -2 Log	324.7		311.3			
	AIC (mejor más pequeño)	334.7		317.3			
	AICC (mejor más pequeño)	336.7		318.1			
	BIC (mejor más pequeño)	328.2		313.4			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.15	0.4779	0.57	0.5875	3.92	0.0565
	B					9.80	0.0037
4	σ^2B	61.5249		121.70		$R^2=0.470118$	
	σ^2A*B	52.7528		129.88			
	CMResidual	234.93		235.18		235.26007	$\bar{Y}=121.8889$
	Verosimilitud -2 Log	305.0		292.4			
	AIC (mejor más pequeño)	315.0		298.4			
	AICC (mejor más pequeño)	317.0		299.2			
	BIC (mejor más pequeño)	308.5		294.5			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.82	0.4061	0.91	0.5147	5.36	0.0272
	B					15.01	0.0005
5	σ^2B	151.95		301.98		$R^2=0.616313$	
	σ^2A*B	87.9056		199.18			
	CMResidual	221.77		221.90		221.94262	$\bar{Y}=125.6389$
	Verosimilitud -2 Log	305.3		291.7			
	AIC (mejor más pequeño)	315.3		297.7			
	AICC (mejor más pequeño)	317.3		298.5			
	BIC (mejor más pequeño)	308.8		293.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	1.25	0.4642	0.63	0.5738	5.62	0.0240
	B					33.15	<.0001
	A*B					8.97	0.0053

ANEXO G

Salida del programa SAS (SC III), datos Desbalanceados, repetición 10.

Repetición 10. DESBALANCEADOS							
Exp.		ML		REML		GLM	
6	σ^2B	42.4818		103.32		$R^2= 0.187256$	
	σ^2A*B	0		0.1063			
	CMResidual	337.44		347.61		348.07674	$\bar{Y}= 121.0833$
	Verosimilitud -2 Log	314.1		303.1			
	AIC (mejor más pequeño)	322.1		309.1			
	AICC (mejor más pequeño)	323.4		309.9			
	BIC (mejor más pequeño)	316.9		305.1			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.04	0.8747	0.03	0.8895	0.03	0.8743
	B					6.04	0.0196
7	σ^2B	63.5345		125.63		$R^2=0.405150$	
	σ^2A*B	38.6601		99.4237			
	CMResidual	212.13		212.37		212.44132	$\bar{Y}=124.8611$
	Verosimilitud -2 Log	301.1		288.9			
	AIC (mejor más pequeño)	311.1		294.9			
	AICC (mejor más pequeño)	313.1		295.7			
	BIC (mejor más pequeño)	304.6		290.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	0.00	0.9939	0.00	0.9925	0.00	0.9698
	B					15.57	0.0004
8	σ^2B	134.76		265.07		$R^2=0.415393$	
	σ^2A*B	9.0956		53.9681			
	CMResidual	334.20		334.72		334.89852	$\bar{Y}=119.8889$
	Verosimilitud -2 Log	316.1		303.6			
	AIC (mejor más pequeño)	326.1		309.6			
	AICC (mejor más pequeño)	328.1		310.4			
	BIC (mejor más pequeño)	319.6		305.7			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.07	0.3863	1.04	0.4935	2.51	0.1228
	B					16.31	0.0003
9	σ^2B	66.8863		153.86		$R^2=0.252181$	
	σ^2A*B	0		0			
	CMResidual	370.60		381.81		393.64019	$\bar{Y}=121.1389$
	Verosimilitud -2 Log	318.0		306.5			
	AIC (mejor más pequeño)	326.0		310.5			
	AICC (mejor más pequeño)	327.3		310.9			
	BIC (mejor más pequeño)	320.8		307.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	2.55	0.3561	2.42	0.3638	2.29	0.1402
	B					7.93	0.0082
10	σ^2B	86.3120		169.47		$R^2=0.417849$	
	σ^2A*B	0		25.2448			
	CMResidual	238.43		238.95		239.09991	$\bar{Y}=117.0278$
	Verosimilitud -2 Log	303.3		291.8			
	AIC (mejor más pequeño)	311.3		297.8			
	AICC (mejor más pequeño)	312.6		298.6			
	BIC (mejor más pequeño)	306.0		293.9			
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A	5.22	0.2626	2.65	0.3505	5.11	0.0307
	B					14.32	0.0006
	σ^2B						
	σ^2A*B						
	CMResidual						
	Verosimilitud -2 Log						
	AIC (mejor más pequeño)						
	AICC (mejor más pequeño)						
	BIC (mejor más pequeño)						
	Efecto	F-Valor	Pr> F	F-Valor	Pr> F	F-Valor	Pr> F
	A						
	B						
	A*B					1.90	0.1772

ANEXO H

Salida del programa Statistix (Estadística Descriptiva).

Statistix 8.0					
Descriptive Statistics					
	AN	EM	PN	AJUSDE	AJUS18
N	6354	6354	6354	6354	6354
Missing	0	0	0	0	0
Sum	12750732	41505	239588	1180902	1868135
Lo 95%CI	2006.7	6.4744	37.586	185.22	292.84
Mean	2006.7	6.5321	37.707	185.85	294.01
Up 95%CI	2006.8	6.5898	37.828	186.49	295.18
Minimum	2001.0	3.0000	21.000	75.000	148.00
1 st .Quarti	2005.0	5.0000	34.000	170.75	261.75
Median	2007.0	6.0000	38.000	186.00	289.00
3 rd .Quarti	2009.0	8.0000	40.000	201.00	322.00
Maximum	2012.0	11.000	58.000	330.00	658.00

ANEXO I

Resumen de la salida del SAS, (SC III), (Peso al Nacer).

		ML				REML				GLM	
		Val Z		Pr Z		Val Z		Pr Z			
Modelo I	σ^2 PADRE	1.4096	5.06	<.0001		1.4501	5.05	<.0001		20.8298	$R^2=0.155696$ $PN_{MEDIA}=37.70664$ C.V=12.10387
	CMResidual	20.7491	55.98	<.0001		20.8324	55.87	<.0001			
	Verosimilitud	37439.1				37472.6					
	AIC	37497.1				37476.6					
	AICC	37497.3				37476.6					
	BIC	37567.5				37481.5					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	21.52	<.0001	11	6244	21.41	<.0001	11	6244	20.33	<.0001
	MN	6.24	<.0001	6	6244	6.21	<.0001	6	6244	5.95	<.0001
	SEXO	362.05	<.0001	1	6244	360.58	<.0001	1	6244	359.22	<.0001
EM	11.90	<.0001	8	6244	11.85	<.0001	8	6244	12.21	<.0001	
PADRE									5.93	<.0001	
		Val Z		Pr Z		Val Z		Pr Z			
Modelo II	σ^2 PADRE	1.3203	5.01	<.0001		1.3667	4.99	<.0001		20.0074	$R^2=0.197601$ $PN_{MEDIA}=37.70664$ C.V=11.86253
	CMResidual	19.7218	55.98	<.0001		20.0106	55.58	<.0001			
	Verosimilitud	37115.4				37062.1					
	AIC	37305.4				37066.1					
	AICC	37308.4				37066.2					
	BIC	37536.4				37071.0					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	16.72	<.0001	11	6178	16.47	<.0001	11	6178	15.69	<.0001
	MN	3.49	0.0019	6	6178	3.43	0.0022	6	6178	3.16	0.0043
	SEXO	374.91	<.0001	1	6178	369.49	<.0001	1	6178	369.23	<.0001
EM	11.76	<.0001	8	6178	11.60	<.0001	8	6178	11.92	<.0001	
AN*MN	5.03	<.0001	66	6178	4.96	<.0001	66	6178	4.89	<.0001	
PADRE									5.72	<.0001	
		Val Z		Pr Z		Val Z		Pr Z			
Modelo III	σ^2 PADRE	1.2500	4.64	<.0001		1.2894	4.60	<.0001		19.9527	$R^2= 0.215365$ $PN_{MEDIA}=37.70664$ C.V=11.83146
	σ^2 PADRE*SEXO	0.1295	1.20	0.1145		0.1422	1.27	0.1022			
	CMResidual	19.6633	55.67	<.0001		19.9476	55.28	<.0001			
	Verosimilitud	37113.2				37059.6					
	AIC	37305.2				37065.6					
	AICC	37308.2				37065.6					
	BIC	37538.6				37072.9					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	16.67	<.0001	11	6095	16.41	<.0001	11	6095	15.26	<.0001
	MN	3.48	0.0019	6	6095	3.43	0.0022	6	6095	3.14	0.0045
SEXO	277.35	<.0001	1	83	268.51	<.0001	1	83	245.60	<.0001	
EM	11.78	<.0001	8	6095	11.62	<.0001	8	6095	11.81	<.0001	
AN*MN	5.03	<.0001	66	6095	4.96	<.0001	66	6095	4.83	<.0001	
PADRE									5.36	<.0001	
PADRE*SEXO									1.20	0.1011	

Leyenda: **gl/n** (grados de libertad del numerador) y **gl/d** (grados de libertad del denominador).

ANEXO J

Resumen de la salida del SAS, (SC III), (Peso al Destete).

		ML				REML				GLM	
		Val	Z	Pr	Z	Val	Z	Pr	Z		
Modelo I	σ^2 PADRE	11.8967	3.45	0.0003		12.6022	3.50	0.0002		569.978	$R^2= 0.161704$
	CMResidual	567.96	55.97	<.0001		570.16	55.87	<.0001			AJUSDE _{MEDIA} =
	Verosimilitud	58401.0				58347.0					185.8517
	AIC	58459.0				58351.0					C.V=12.84583
	AICC	58459.3				58351.0					
	BIC	58529.5				58355.8					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	35.69	<.0001	11	6244	35.50	<.0001	11	6244	33.78	<.0001
	MN	12.41	<.0001	6	6244	12.29	<.0001	6	6244	10.13	<.0001
	SEXO	437.98	<.0001	1	6244	436.16	<.0001	1	6244	432.38	<.0001
Modelo II	EM	18.62	<.0001	8	6244	18.55	<.0001	8	6244	18.68	<.0001
	PADRE									2.38	<.0001
		Val	Z	Pr	Z	Val	Z	Pr	Z		
	σ^2 PADRE	8.3905	3.02	0.0012		9.1591	3.11	0.0010		535.296	$R^2= 0.221033$
	CMResidual	527.85	55.96	<.0001		535.44	55.58	<.0001			AJUSDE _{MEDIA} =
	Verosimilitud	57923.6				57567.5					185.8517
	AIC	58113.6				57571.5					C.V=12.44888
	AICC	58116.5				57571.5					
	BIC	58344.5				57576.4					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	26.97	<.0001	11	6178	26.52	<.0001	11	6178	24.79	<.0001
	MN	18.78	<.0001	6	6178	18.43	<.0001	6	6178	16.02	<.0001
	SEXO	454.76	<.0001	1	6178	448.32	<.0001	1	6178	449.08	<.0001
Modelo III	EM	16.73	<.0001	8	6178	16.51	<.0001	8	6178	16.93	<.0001
	AN*MN	7.53	<.0001	66	6178	7.41	<.0001	66	6178	7.13	<.0001
	PADRE									2.07	<.0001
		Val	Z	Pr	Z	Val	Z	Pr	Z		
	σ^2 PADRE	8.3758	3.02	0.0012		9.1697	3.11	0.0010		535.730	$R^2= 0.230876$
	σ^2 PADRE*SEXO	0	.	.		0	.	.			AJUSDE _{MEDIA} =
	CMResidual	527.85	55.96	<.0001		535.43	55.58	<.0001			185.8517
	Verosimilitud	57923.6				57567.5					C.V=12.45393
	AIC	58113.6				57571.5					
	AICC	58116.5				57571.5					
	BIC	58344.5				57576.4					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	26.97	<.0001	11	6095	26.52	<.0001	11	6095	24.25	<.0001
	MN	18.79	<.0001	6	6095	18.43	<.0001	6	6095	16.07	<.0001
	SEXO	454.76	<.0001	1	83	448.33	<.0001	1	83	314.73	<.0001
	EM	16.73	<.0001	8	6095	16.51	<.0001	8	6095	16.58	<.0001
	AN*MN	7.53	<.0001	66	6095	7.41	<.0001	66	6095	6.86	<.0001
	PADRE									1.97	<.0001
	PADRE*SEXO									0.94	0.6342

Leyenda: **gl/n** (grados de libertad del numerador) y **gl/d** (grados de libertad del denominador).

ANEXO K

Resumen de la salida del SAS, (SC III), (Peso a los 18 Meses).

		ML				REML				GLM	
		Val	Z	Pr	Z	Val	Z	Pr	Z		
Modelo I	σ^2 PADRE	51.0519	3.99	<.0001		53.4590	4.02	<.0001		1459.41	$R^2= 0.365061$
	CMResidual	1455.04	55.93	<.0001		1460.71	55.83	<.0001			AJUS18 _{MEDIA} =
	Verosimilitud	64404.8				64324.6					294.0093
	AIC	64462.8				64328.6					C.V=12.99353
	AICC	64463.1				64328.6					
	BIC	64533.3				64333.5					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	74.12	<.0001	11	6244	73.71	<.0001	11	6244	69.38	<.0001
	MN	1.53	0.1643	6	6244	1.53	0.1646	6	6244	1.71	0.1154
	SEXO	1966.85	<.0001	1	6244	1959.20	<.0001	1	6244	1959.40	<.0001
Modelo II	EM	1.90	0.0559	8	6244	1.89	0.0571	8	6244	1.91	0.0543
	PADRE									3.11	<.0001
		Val	Z	Pr	Z	Val	Z	Pr	Z		
	σ^2 PADRE	43.4987	3.73	<.0001		46.2924	3.78	<.0001		1407.98	$R^2= 0.393912$
	CMResidual	1389.47	55.90	<.0001		1409.54	55.52	<.0001			AJUS18 _{MEDIA} =
	Verosimilitud	64105.6				63658.1					294.0093
	AIC	64295.6				63662.1					C.V=12.76253
	AICC	64298.5				63662.1					
	BIC	64526.5				63667.0					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	64.24	<.0001	11	6178	63.21	<.0001	11	6178	59.48	<.0001
	MN	1.38	0.2169	6	6178	1.37	0.2247	6	6178	1.44	0.1967
	SEXO	1959.26	<.0001	1	6178	1931.35	<.0001	1	6178	1933.38	<.0001
Modelo III	EM	1.86	0.0623	8	6178	1.83	0.0670	8	6178	1.83	0.0661
	AN*MN	4.65	<.0001	66	6178	4.58	<.0001	66	6178	4.46	<.0001
	PADRE									2.86	<.0001
		Val	Z	Pr	Z	Val	Z	Pr	Z		
	σ^2 PADRE	27.4888	2.04	0.0206		29.5926	2.10	0.0180		1390.84	$R^2=$
	σ^2 PADRE*SEXO	34.2025	2.60	0.0047		35.6392	2.62	0.0044			0.409332
	CMResidual	1374.66	55.47	<.0001		1394.31	55.10	<.0001			AJUS18 _{MEDIA} =
	Verosimilitud	64088.2				63640.1					294.0093
	AIC	64280.2				63646.1					C.V=12.68462
	AICC	64283.1				63646.1					
	BIC	64513.5				63653.4					
	Efecto	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F	gl/n	gl/d	F-Valor	Pr> F
	AN	64.21	<.0001	11	6095	63.20	<.0001	11	6095	58.96	<.0001
	MN	1.45	0.1911	6	6095	1.43	0.1983	6	6095	1.54	0.1595
Modelo III	SEXO	1005.68	<.0001	1	83	980.27	<.0001	1	83	1474.01	<.0001
	EM	1.76	0.0806	8	6095	1.73	0.0864	8	6095	1.69	0.0946
	AN*MN	4.61	<.0001	66	6095	4.54	<.0001	66	6095	4.37	<.0001
	PADRE									2.89	<.0001
	PADRE*SEXO									1.92	<.0001

Leyenda: **gl/n** (grados de libertad del numerador) y **gl/d** (grados de libertad del denominador).

ANEXO L

Resumen de la salida del SAS (Promedios).

	Año	Promedio	EM	Promedio	MN	Promedio	Sexo	Promedio
Modelo III	2001	34.98	3	35.98	1	37.34	1	35.84
	2002	36.60	4	37.07	2	36.97	5	38.16
	2003	36.68	5	37.82	3	37.07		
	2004	35.59	6	37.69	4	36.18		
	2005	38.09	7	37.26	10	37.07		
	PN	2006	8	37.17	11	37.56		
		2007	9	36.92	12	36.80		
		2008	10	36.86				
		2009	11	36.22				
		2010						
		2011						
		2012						
Modelo II	2001	187.66	1	176.43	3	174.95	1	176.55
	2002	178.85	2	180.1	4	181.55	5	189.77
	2003	172.81	3	183.99	5	186.4		
	2004	186.8	4	184.37	6	185.99		
	2005	184.65	10	188.24	7	185.64		
	AJUSDE	2006	11	185.17	8	185.48		
		2007	12	183.84	9	183.23		
		2008			10	183.15		
		2009			11	182.06		
		2010						
		2011						
		2012						
Modelo III	2001	308.55	1	286.83	3	287.9	1	266.79
	2002	281.24	2	288.08	4	290.04	5	312.26
	2003	273.63	3	291.82	5	293.06		
	2004	307.54	4	292.26	6	291.5		
	2005	298.83	10	287.87	7	289.33		
	AJUS18	2006	11	290.65	8	289.96		
		2007	12	289.14	9	287.52		
		2008			10	288.84		
		2009			11	287.57		
		2010						
		2011						
		2012						

ANEXO M

Resumen de la salida del SAS (Promedios) estimados interacción AN*MN.

		Años					
	Meses	2001	2002	2003	2004	2005	2006
PN	1	33.97	37.19	36.65	36.62	38.69	34.7
	2	33.25	36.25	37.98	35.88	37.7	35.39
	3	35.3	37.22	34.87	35.7	38.07	35.27
	4	32.71	35.5	35.69	34.59	38.61	37.61
	10	36.77	36.06	37.67	36.17	37.19	38.18
	11	36.22	37.13	36.85	35.99	39.13	37.66
	12	36.67	36.87	37.06	34.19	37.23	33.53
AJUSDE	1	175.93	178.71	163.36	172.56	181.52	187.51
	2	176.58	170.35	165.24	184.01	182.97	189.3
	3	185.81	184.3	165.47	198.89	186.83	193.36
	4	190.12	178.57	182.97	191.82	176.79	187.95
	10	191.88	184.1	185.11	198.23	190.89	181.6
	11	195.52	181.36	173.25	188.23	185	183.94
	12	197.79	174.58	174.28	173.89	188.53	185.3
AJUS18	1	311.77	280.77	278.12	302.26	291.5	317.1
	2	301.85	280.56	269.47	310.6	304.88	312.99
	3	313.13	287.9	262.78	294.06	307.41	313.04
	4	300.18	304.32	268.32	316.15	306.02	312.71
	10	303.33	272.89	296.51	298.94	302.12	294.96
	11	310.71	270.95	268.25	322.61	289.63	310.55
	12	318.91	271.28	271.98	308.18	290.23	305.67
		Años					
	Meses	2007	2008	2009	2010	2011	2012
PN	1	40.43	39.77	36.87	37.47	38.2	37.52
	2	39.77	38.95	37.74	37	37.32	36.42
	3	38.27	37.81	38.34	36.49	39.39	38.17
	4	36.08	37.34	37.88	37.23	36.16	34.78
	10	36.97	37.9	36.42	37.62	37.11	36.75
	11	38.44	37.51	37.58	38.36	37.52	38.3
	12	38.93	38.73	39.16	33.49	39.31	36.42
AJUSDE	1	184.48	192.94	182.82	173.52	176.85	146.93
	2	178.77	184.41	192.29	174.35	181.14	181.79
	3	166.22	188.38	193.31	174.66	182.18	188.48
	4	165.41	194.7	192.44	174.47	172.95	204.19
	10	189.51	192.07	192.98	187.19	193.83	171.43
	11	182.92	200.77	192.92	185.14	185.25	167.8
	12	182	197.33	183.17	172.32	203.16	173.72
AJUS18	1	289.44	294.48	287.13	307.48	273.39	208.57
	2	287.71	291.28	280.25	304.11	306.14	207.17
	3	282.94	295.15	285.5	301.78	293.96	264.15
	4	291.35	292.39	299.71	299.55	266.57	249.88
	10	291.78	279.9	272.15	305.07	277.98	258.8
	11	286.49	293.85	285.48	331.32	275.11	242.83
	12	281.05	292.13	287.79	311.89	283.8	246.79