

APLICACIÓN DEL MÉTODO REML EN MODELOS ESTADÍSTICOS LINEALES EN MEJORAMIENTO ANIMAL



<https://doi.org/10.22533/at.ed.811112520035>

Data de aceite: 05/08/2025

José Raúl Pérez González

Universidad politécnica territorial de
Maracaibo (UPTMA)
<https://orcid.org/0009-0007-2442-486X>

Ernesto José Jiménez Quintero

Asociación Venezolana de criadores de
ganado Carora (ASOCRICA)
Barinas-Barinas
<https://orcid.org/0009-0000-8783-1701>

Gilberto Antonio Pérez Quintero

Universidad Centroccidental “Lisandro
Alvarado” (UCLA)
Barquisimeto-Lara. Venezuela

Ricaurte Lopera-Vásquez

Universidad Cooperativa de Colombia.
Grupo de Investigación IMPRONTA
Ibagué-Tolima. Colombia
<https://orcid.org/0000-0001-6792-1961>

El método de máxima verosimilitud restringida (REML) es una técnica estadística que se utiliza en modelos estadísticos para estimar parámetros, fue desarrollado por Patterson and Thompson (1971) para ser usado en diseños de bloques incompletos balanceados, pero ha sido ampliado para ser usado en

diferentes áreas, con especial énfasis en el mejoramiento genético animal.

El REML puede ser utilizado para probar hipótesis de comparación de medias en modelos de efectos fijos, pero su aplicación principal es en la estimación de componentes de varianza bajo modelos aleatorios o mixtos, especialmente en el campo del mejoramiento genético animal, donde se busca estimar las variaciones de los factores aleatorios, que dan lugar a los parámetros genéticos y ambientales.

MODELO DE EFECTOS FIJOS

Los modelos fijos o modelos de efectos fijos son aquellos en los que se asume que los efectos de los factores considerados son constantes y no varían entre las unidades de observación. En este tipo de modelos, los niveles de los factores se seleccionan intencionalmente y se consideran representativos del conjunto completo. Por ejemplo, si se está estudiando el efecto de diferentes tratamientos en un experimento, cada tratamiento se considera fijo y se analiza su efecto específico (Castejón-Sandoval, 2011).

En un modelo fijo, se estima un parámetro para cada nivel del factor, lo que permite hacer inferencias sobre esos niveles específicos. Un ejemplo clásico de un modelo fijo es el caso, donde se comparan las medias de diferentes grupos para determinar si existen diferencias significativas entre ellas. El modelo estadístico lineal general en algebra de matrices para el modelo fijo es (Blasco, 2017):

$$y = Xb + e$$

Donde y es un vector de datos, X es una matriz de incidencia que relaciona los datos con los efectos fijos, b es un vector de parámetros desconocidos para los efectos fijos del modelo y e es un vector de residuales. Bajo cualquier modelo fijo el único componente aleatorio será el residual e , y los fijos pueden ser las covariables (variables cuantitativas) o variables de clasificación (variables cualitativas). La solución del Vector b para el caso general es un estimador de máxima verosimilitud, el cual tiene la siguiente expresión matemática (Gutiérrez, 2010):

$$b = (X'V^{-1}X)^{-1}(X'V^{-1}y)$$

Donde V es la matriz de varianzas y covarianzas fenotípicas de los efectos aleatorios del modelo, la cual al contemplar como único efecto aleatorio el residual, tiene la siguiente expresión matemática:

$$V = R$$

Donde $R = I\sigma_e^2$ es la matriz de varianzas y covarianzas residuales, siendo I una matriz identidad de tamaño $n \times n$ y σ_e^2 es la varianza residual. En el modelo fijo como $V=R$ las estimaciones de máxima verosimilitud para los parámetros desconocidos son:

$$b = (X'R^{-1}X)^{-1}(X'R^{-1}y) = (X'X)^{-1}(X'y)$$

Y como R es una matriz diagonal su inversa es inmediata. Para calcular la varianza residual σ_e^2 se debe maximizar una función de los datos dados los parámetros, vista como una función de los parámetros llamada función de verosimilitud, en el caso de REML esta función es la función de verosimilitud restringida o L_r , en la práctica, se maximiza el logaritmo natural de debido a que es más fácil, por lo tanto, la forma matemática de $Ln(L_r)$ es (Pérez González, 2025):

$$Ln(L_r) = -0.5Ln|V| - 0.5Ln|X'V^{-1}X| - 0.5 y'Py - \frac{n-p}{2} \ln(2\pi)$$

Pero como en el modelo fijo la función queda definida como:

$$Ln(L_r) = -0.5Ln|R| - 0.5Ln|X'R^{-1}X| - 0.5 y'Py - \frac{n-p}{2} \ln(2\pi)$$

Donde $-\frac{n-p}{2} \ln(2\pi)$ es una constante que desaparece al derivar y $y'Py$ es la suma de cuadrados residual generalizada, por lo tanto, toda la función queda completamente definida por los valores de σ_e^2 para el modelo fijo. La primera derivada de $Ln(L)$ con

respecto σ_e^2 a para el modelo fijo da lugar a la siguiente estimación para la varianza residual (Pérez, 2024):

$$\sigma_{e(REML)}^2 = CMe$$

Donde **CMe** es el cuadrado medio residual del modelo, el cual viene definido como (Montgomery & St, 2022):

$$CMe = \frac{SCe}{GLE} = \frac{SCtotal - SCresto}{GLtotal - GLresto}$$

Donde **SCe** : suma de cuadrados residual del modelo, **SCtotal** es la suma de cuadrados total, **SCresto** es la sumatoria de todas las sumas de cuadrados de los factores del modelo, **GLtotal** : son los grados de libertad total, **GLresto** es la sumatoria de todos los grados de libertad de los factores del modelo y **GLE** son los grados de libertad residual del modelo estadístico. Para estimar si existe diferencias estadísticamente significativas entre los niveles de un factor de interés, se utiliza la prueba F, la cual tiene la siguiente fórmula matemática (Castejón-Sandoval, 2011):

$$F = \frac{CM_{factor}}{\sigma_{e(REML)}^2}$$

Donde F es la prueba F y **CM_{factor}** es el cuadrado medio del factor de interés. Para conocer el efecto de los factores fijos, se utiliza una prueba F y el criterio del valor P, donde valores de $P < 0.05$ (5%) indican diferencias significativas entre los niveles de un factor.

En general, para el modelo fijo se necesita calcular los **CM** del residuo y de los factores de interés, para esto es necesario estimar los grados de libertad y las sumas de cuadrados, estos procedimientos pueden variar si el modelo usa una data balanceada o desbalanceada (un modelo con data balanceada es aquel en el que cada nivel de los factores tiene el mismo número de datos, mientras que uno con data desbalanceada es aquel en el que los niveles de los factores no tienen el mismo número datos).

Para el caso balanceado, en un modelo que no posea interacción, los grados de libertad quedan definidos de la siguiente manera:

$$GLtotal = n - 1$$

$$GL_{factor} = n_{factor} - 1$$

Donde **GL_{factor}** son los grados de libertad del factor de interés, n es el número de datos y n_{factor} es el número de niveles del factor de interés. Para las sumas de cuadrados, las fórmulas (Searle et al., 1992) son:

Donde **SC_{factor}** es la suma de cuadrados para el factor de interés, k es el número de réplicas y $\sum y_{factor}^2$ es la suma de los datos de cada nivel del factor al cuadrado.

$$SC_{total} = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$SC_{factor} = \frac{\sum y_{factor}^2}{k} - \frac{(\sum y)^2}{n}$$

MODELO CON EFECTOS MIXTOS

Los modelos mixtos son aquellos en los que se asume que los efectos de alguno de los factores son variables y están sujetos a variabilidad aleatoria y otros factores son fijos. En este enfoque, los niveles de los factores aleatorios se consideran como una muestra aleatoria de una población más grande y los fijos de una población finita, por lo tanto, en lugar de estimar un parámetro para cada nivel de un factor aleatorio, se estima una varianza que representa la variabilidad (Castejón-Sandoval, 2011).

El modelo estadístico lineal para un modelo con efectos mixtos es (Gianola y Sorensen, 2002):

$$y = Xb + Za + e$$

Donde Z es una matriz de incidencia que relaciona los datos con los efectos aleatorios y a es un vector de predicciones desconocidas para el efecto aleatorio (Littell et al., 2006).

En el modelo mixto existen mínimo dos factores aleatorios en el modelo, por lo tanto, la matriz de varianza y covarianza fenotípica viene dado por (Mrode, 2014):

$$V = ZGZ' + R$$

Donde $G = I\sigma_i^2$ es la matriz de varianza y covarianzas para el factor aleatorio y σ_i^2 es la varianza del factor aleatorio del modelo. Interesa estimar las varianzas σ_i^2 y σ_e^2 del modelo, por lo tanto, se tiene que maximizar $Ln(L_r)$ en función de los componentes de varianza. En particular para el caso balanceado, las estimas REML son idénticas a las de ANOVA (análisis de la varianza) siempre y cuando se cumpla la desigualdad $CM_{factor} > CMe$ en caso contrario ($CM_{factor} < CMe$) la varianza del factor $\sigma_i^2 = 0$. La fórmula para los componentes de varianza de los factores de interés, derivadas de la maximización de $Ln(L_r)$ en data balanceada y si solo existe un efecto aleatorio (diferente del residual) (Montgomery & St, 2022) son:

$$\sigma_{i(REML)}^2 = \frac{CM_{factor} - CMe}{k}$$

$$\sigma_{e(REML)}^2 = CMe$$

Para el caso desbalanceado, la primera derivada de $Ln(L_r)$ genera ecuaciones no lineales que tienen que resolverse por iteración numérica. Otra alternativa es aplicar algún algoritmo numérico a la verosimilitud logarítmica (Meyer, 1989) para minimizar $-2Ln(L_r)$, lo cual es equivalente a maximizar, en algebra de matrices el $-2Ln(L_r)$ es:

$$-2Ln(L) = \frac{n-p}{2} \ln(2\pi) + Ln|C| + \ln|G| + \ln|R| + y'Py$$

Donde C es la matriz de coeficientes de las ecuaciones normales de Henderson (Henderson, 1984). Algunos algoritmos que son usados para minimizar $-2Ln(L_r)$ son el SIMPLEX y el POWELL (Boldman y col, 1995) los cuales se caracterizan por no usar derivadas, otro muy usado es el AI-REML (Johnson y Thompson, 1995).

MODELO PADRE

En genética cuantitativa, es necesario estimar parámetros genéticos para predecir los BLUP (mejor estimador lineal insesgado) con el objetivo de poder mejorar los procesos de selección genética. La heredabilidad, es el parámetro genético más importante, ya que explica cuanto de las diferencias observadas en el fenotipo de un grupo de animales se puede heredar, matemáticamente la fórmula es:

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2}$$

Donde h^2 es la heredabilidad, σ_p^2 es la varianza fenotípica y σ_a^2 es la varianza genética aditiva.

Para estimar los componentes de varianza σ_a^2 y σ_e^2 se propusieron varios procedimientos, entre los cuales está utilizar la información de los hijos de n reproductores, es decir, la información de medios hermanos paternos. Para realizar este procedimiento se utiliza el modelo padre, que en algebra de matrices viene dado por (Cerón Muñoz et al, 2012):

$$y = Xb + Zs + e$$

Donde s es un vector de predicciones desconocidas para cada reproductor. Bajo este modelo la varianza la matriz de varianza y covarianzas para el factor aleatorio viene dada por:

$$G = I\sigma_s^2$$

Donde σ_s^2 es la varianza entre toros, la cual representan $\frac{1}{4}\sigma_a^2$ y σ_e^2 representa $\frac{3}{4}\sigma_a^2$ y toda la varianza ambiental (σ_{ev}^2), por último, $\sigma_p^2 = \sigma_s^2 + \sigma_e^2$ por lo tanto, la heredabilidad puede calcularse como (Pérez, 2024):

$$h^2 = \frac{4\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$$

Los BLUP se calculan resolviendo el siguiente sistema de ecuaciones lineales llamado, ecuaciones normales de Henderson (Román y Aranguren, 2014) :

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + I\alpha \end{bmatrix} \begin{bmatrix} b_i \\ s_i \end{bmatrix} = \begin{pmatrix} X'y \\ Z'y \end{pmatrix}$$

Donde a es un escalar que introduce el valor de la heredabilidad en las ecuaciones, el cual viene dado por:

$$\alpha = \frac{4 - h^2}{h^2}$$

Si se posee la genealogía entre padres, la matriz G ya no es una matriz diagonal, ya que se utiliza toda la información genealógica para evaluar un reproductor, por lo tanto, la matriz de varianzas y covarianzas es (Gutiérrez, 2010):

$$G = A\sigma_s^2$$

Donde A es la matriz de parentesco. Las ecuaciones normales de Henderson para $G = A\sigma_s^2$ son (Elzo, 2012):

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\alpha \end{bmatrix} \begin{bmatrix} b_i \\ s_i \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

La correlación entre el predictor de S_i y su valor real viene dado por:

$$r_{ss} = \sqrt{1 - C_{ii}\alpha}$$

Donde r_{ss} es la correlación y C_{ii} es el elemento inverso de la diagonal de la matriz de coeficiente para el bloque de animales del modelo mixto. El error estándar de s_i viene dado por:

$$EE(s) = \sqrt{C_{ii}\sigma_e^2}$$

Donde $EE(s)$ es el error estándar.

Ejemplo 1.

En la tabla 1 se presenta la siguiente base de datos de peso vivo en ganado vacuno, donde la primera columna son los toros (padres), luego el sexo de sus hijos y por último el peso al destete de los hijos. Esta base de datos se usó para mostrar el procedimiento de estimación de \hat{h}^2 , b y los BLUP en data balanceada y no balanceada bajo un modelo padre:

Animal	Sexo (hijo)	Peso destete
1	M	205
1	M	260
1	H	160
1	H	171
2	M	260
2	M	210
2	H	170
2	H	144
3	M	200

3	M	205
3	H	130
3	H	150

Tabla 1. Base de datos ejemplo 1

Los modelos usados se presentan en la tabla 2:

Modelos	Formula algebra escalar	Formula matricial	Modelo en palabras
Fijo	$y_{ij} = \mu + b_i + e_{ij}$	$y = Xb + e$	PD = media + sexo + error
Mixto	$y_{ijk} = \mu + b_i + s_j + e_{ijk}$	$y = Xb + Zs + e$	PD = media + sexo + DEP+ error

Tabla 2. Modelos utilizados

Para la resolución del modelo fijo se utilizó el programa SAS y para el modelo mixto el programa WOMBAT (Meyer, 2007) y SAS.

CASO BALANCEADO MODELO FIJO.

Para el modelo de efectos fijos, los cálculos de cuadrados medios son:

Cuadrado medio para el factor sexo:

$$CM_{\text{sexo}} = \frac{\frac{(205 + 260 + 260 + 210 + 200 + 205)^2}{6} + \frac{(160 + 171 + 170 + 144 + 130 + 150)^2}{6} - \frac{2265^2}{12}}{2 - 1} = 14352.08333$$

Suma de cuadrados total:

$$SC_{\text{total}} = 205^2 + \dots + 150^2 - \frac{2265^2}{12} = 19708.25$$

Cuadrado medio para el residuo:

$$CMe = \frac{19708.25 - 14352.08333}{11 - 1} = \frac{5356.16667}{10} = 535.616667$$

Componente de varianza residual:

$$\sigma_{e(REML)}^2 = CMe = 535.616667$$

Prueba F:

$$F = \frac{CM_{\text{sexo}}}{\sigma_{e(\text{REML})}^2} = \frac{14352.08333}{535.616667} = 26.79543$$

Estimador de los efectos fijos:

$$b = \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix}^{-1} \begin{pmatrix} 1340 \\ 925 \end{pmatrix} = \begin{pmatrix} 1340/6 \\ 925/6 \end{pmatrix} = \begin{pmatrix} 223.3333 \\ 154.16666 \end{pmatrix}$$

Código SAS PROC GLIMMIX para el modelo fijo:

```
data trabajo;
input animal sexo$ pd @@;
datalines;
1 m 205 2 m 260 3 m 200
1 m 260 2 m 210 3 m 205
1 h 160 2 h 170 3 h 130
1 h 171 2 h 144 3 h 150;
;
proc glimmix;
class animal sexo;
model pd = sexo / solution;
run;
```

En la figura 1 se presentan los resultados del SAS para el modelo fijo:

Procedimiento GLIMMIX				
Información del modelo				
Conjunto de datos	WORK.TRABAJO			
Variable de respuesta	pd			
Distribución de respuesta	Gaussian			
Función de vínculo	Identidad			
Función de varianza	Predeterminado			
Matriz de varianza	Diagonal			
Técnica de estimación	Verosimilitud máxima restringida			
Método de grados de libertad	Residual			

Test de tipo III de efectos fijos				
Efecto	DF Num	DF Den	Valor F	Pr > F
sexo	1	10	26.80	0.0004

Figura 1. Salida del SAS para el modelo fijo.

Resultados idénticos a los encontrados de manera manual. Como el valor $P < 0.05$ (0.0004) se concluye que las medias entre machos y hembras difieren estadísticamente.

Para obtener soluciones de b iguales a las plasmadas manualmente con PROC GLIMMIX se debe aplicar la sentencia NOINT para que el SAS ignore el intercepto, (model pd = sexo / solution noint;).

Código SAS PROC GLIMMIX para obtener b de machos y hembras:

```
proc glimmix;  
class animal sexo;  
model pd = sexo / solution noint;  
run;
```

En la figura 2 se presentan los resultados SAS para los estimadores del factor sexo:

Estimaciones de parámetro						
Efecto	sexo	Estimación	Error estándar	DF	t valor	Pr > t
sexo	h	154.17	9.4483	10	16.32	<.0001
sexo	m	223.33	9.4483	10	23.64	<.0001
Escala		535.62	239.54	.	.	.

Figura 2. Salida SAS para los estimadores del factor sexo.

Resultados idénticos a los expuestos de manera manual. Se debe resaltar que en este caso la prueba F cambia usando la sentencia NOINT.

CASO BALANCEADO MODELO MIXTO

Para el modelo mixto, los cálculos son:
Cuadrado medio para el factor padre:

$$CMs = \frac{\frac{(205 + 260 + 160 + 171)^2}{4} + \frac{(260 + 210 + 170 + 144)^2}{4} + \frac{(200 + 205 + 130 + 150)^2}{4} - \frac{2265^2}{12}}{3 - 1} = 927.75$$

Cuadrado medio residual:

$$CMe = \frac{19708.25 - (1855.5 + 14352.08333)}{11 - (2 + 1)} = \frac{3500.6667}{8} = 437.58333$$

Componente de varianza residual:

$$\sigma_{e(REML)}^2 = CMe = 437.58333$$

Componente de varianza entre padres:

$$\sigma_{s(REML)}^2 = \frac{CMS - CMe}{k} = \frac{927.75 - 437.58333}{4} = \frac{490.16667}{4} = 122.5416675$$

Componente de varianza aditivo:

$$\sigma_{a(REML)}^2 = 4(122.5416675) = 490.16667$$

Componente de varianza fenotípico:

$$\sigma_{p(REML)}^2 = 437.58333 + 122.5416675 = 560.12997$$

Heredabilidad:

$$h^2 = \frac{490.16667}{560.12997} = 0.87094$$

Escalar α :

$$\alpha = \frac{4 - 0.87094}{0.87094} = \frac{3.12906}{0.87094} = 3.59273$$

Ecuaciones normales de Henderson:

$$\begin{bmatrix} 6.00000 & 0.00000 & 2.00000 & 2.00000 & 2.00000 \\ 0.00000 & 6.00000 & 2.00000 & 2.00000 & 2.00000 \\ 2.00000 & 2.00000 & 7.59273 & 0.00000 & 0.00000 \\ 2.00000 & 2.00000 & 0.00000 & 7.59273 & 0.00000 \\ 2.00000 & 2.00000 & 0.00000 & 0.00000 & 7.59273 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1340 \\ 925 \\ 796 \\ 784 \\ 685 \end{bmatrix}$$

$$b_1 2 + b_2 2 + 7.59273 s_1 = 796$$

$$b_1 2 + b_2 2 + 7.59273 s_2 = 784$$

$$b_1 2 + b_2 2 + 7.59273 s_3 = 685$$

Y las soluciones son:

$$s_1 = \frac{796 - 223.3333(2) - 154.16666(2)}{7.59273} = 5.39991$$

$$s_2 = \frac{784 - 223.3333(2) - 154.16666(2)}{7.59273} = 3.81945$$

$$s_3 = \frac{685 - 223.3333(2) - 154.16666(2)}{7.59273} = -9.21933$$

Y el error estándar y la correlación para los 3 animales (por que tienen el mismo número de datos) es:

$$r_{ss} = \sqrt{1 - 0.18058(3.59273)} = 0.59264$$
$$EE(s) = \sqrt{0.18058(437.58333)} = 8.88925$$

Código SAS PROC GLIMMIX para el modelo mixto:

```
proc glimmix;  
class animal sexo;  
model pd = sexo / solution ;  
random animal/ solution;  
run;
```

En la figura 3 se presentan los resultados SAS para los componentes de varianza:



Figura 3. Resultados SAS para los componentes de varianza.

Resultados idénticos a los calculados de manera manual. En la figura 4 se muestra los resultados SAS para las predicciones del BLUP de los tres toros:

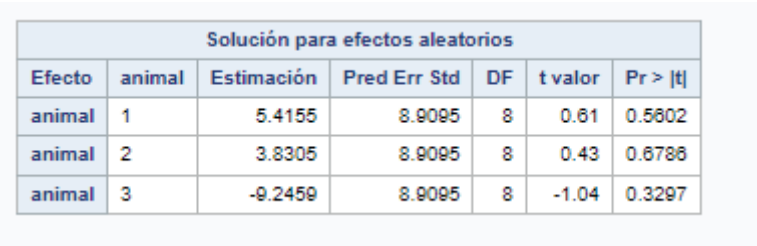


Figura 4. Predicciones de los BLUP para los 3 toros.

Soluciones casi idénticas a las encontradas de manera manual (las diferencias son por el número de decimales usados).

En la figura 5 se muestra el archivo de parámetros, de datos y genealógico en WOMBAT para las obtener los componentes de varianza:

```

com sire model
PEDS ped.ped SIRE
DAT sire.dat

animal
sire
dam
sexo 2
pd
end

anal uni
model
ran animal nrm
fix sexo
trait pd
end mod

var animal 1
103
var error 1
410

```

1	0	0	1	205
1	0	0	1	260
1	0	0	2	160
1	0	0	2	171
2	0	0	1	260
2	0	0	1	210
2	0	0	2	170
2	0	0	2	144
3	0	0	1	200
3	0	0	1	205
3	0	0	2	130
3	0	0	2	150

1	0	0
2	0	0
3	0	0

Figura 5. Archivo de parámetros, de datos y genealógico en WOMBAT para los componentes de varianza

En la figura 6 se muestra la salida de los componentes de varianza con WOMBAT:

```

**** Estimates of residual covariances ****
Order of fit          =      1
Covariance matrix
1      437.58
Matrix of correlations and variance ratios
1      0.7812
Covariances & correlations with approximate sampling errors
1 COVS Z 1 1      437.580      218.789      vrat      0.781      0.362

**** Estimates for RE 1 "animal" ****
No. of levels          =      3
Covariance structure    = NRM
Order of fit           =      1
Covariance matrix
1      122.53
Matrix of correlations and variance ratios
1      0.2188
Covariances & correlations with approximate sampling errors
2 COVS A 1 1      122.529      238.280      vrat      0.219      0.362

**** Estimates of phenotypic covariances ****
Covariance matrix
1      560.11
Covariances & correlations with approximate sampling errors
3 COVS T 1 1      560.109      284.098

```

Figura 6. Salida WOMBAT para los componentes de varianza

Resultados similares a los encontrados anteriormente. Luego de estimados los componentes de varianza, pueden usarse las estimaciones para generar los BLUP, los errores estándares y las correlaciones, usando la siguiente secuencia --BLUP, esto se muestra en la figura 7:

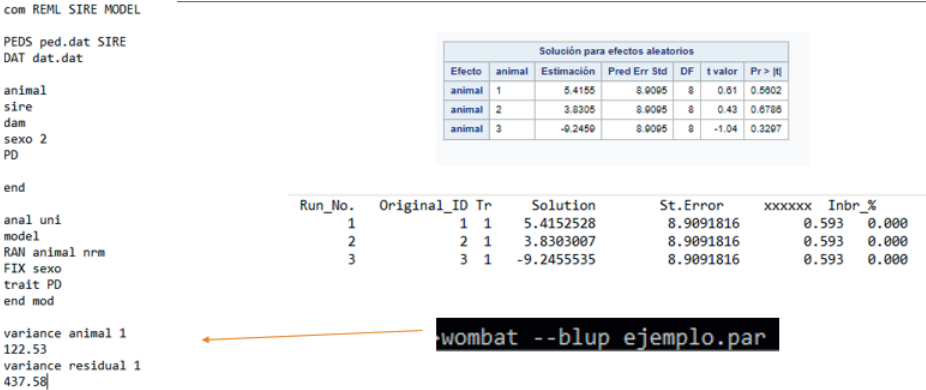


Figura 7. BLUP, errores estándares y correlaciones vía WOMBAT.

Resultados idénticos a los encontrados de manera manual y con SAS.

CASO NO BALANCEADO MODELO FIJO

Como el modelo es a una vía de clasificación, los cálculos son similares, pero cambia la forma de estimar el CMsexo (en este caso se asume el último dato perdido (150)):

Cuadrado medio para el factor sexo:

$$CM_{\text{sexo}} = \frac{\frac{(205 + 260 + 260 + 210 + 200 + 205)^2}{6} + \frac{(160 + 171 + 170 + 144 + 130)^2}{5}}{2 - 1} - \frac{2115^2}{11}$$

$$= 12734.84848$$

Suma de cuadrado total:

$$SC_{\text{total}} = 18070.18182$$

Cuadrado medio residual:

$$CMe = \frac{18070.18182 - 12734.84848}{10 - 1} = \frac{5335.3334}{9} = 592.81481$$

Componente de varianza residual:

Prueba F:

$$F = \frac{CM_{sexo}}{\sigma^2_{e(REML)}} = \frac{12734.84848}{592.81481} = 21.48200$$

Código SAS PROC GLIMMIX para el modelo fijo no balanceado:

```
data trabajo;  
input animal sexo$ pd @@;  
missing x;  
datalines;  
1 m 205 2 m 260 3 m 200  
1 m 260 2 m 210 3 m 205  
1 h 160 2 h 170 3 h 130  
1 h 171 2 h 144 3 h x  
;  
proc glimmix;  
class animal sexo;  
model pd = sexo / solution;  
run;
```

En la figura 8 se presentan la prueba F para el caso no balanceado del modelo fijo:

Test de tipo III de efectos fijos				
Efecto	DF Num	DF Den	Valor F	Pr > F
sexo	1	9	21.48	0.0012

Figura 8. Prueba F para el modelo fijo en el caso no balanceado

Resultados idénticos a los encontrados manualmente. Como el valor $P < 0.05$ ($P = 0.0012$) se concluye nuevamente que existen diferencias estadísticamente significativas entre los niveles del factor sexo.

CASO NO BALANCEADO MODELO MIXTO

En la figura 9 se presentan las iteraciones de $-2Ln(L_r)$ y los componentes de varianza con SAS:

Estimaciones del parámetro de covarianza		
Parm Cov	Estimación	Error estándar
animal	165.58	325.07
Residual	468.83	254.64

Historial de iteración					
Iteración	Reinicio	Evaluaciones	Función objetivo	Cambio	Gradiente máx
0	0	4	85.956473735	.	0.509623
1	0	3	85.936537574	0.01993616	0.172049
2	0	2	85.931648187	0.00488939	0.052782
3	0	2	85.931225076	0.00042311	0.008983
4	0	2	85.931213296	0.00001178	0.000383
5	0	2	85.931213275	0.00000002	2.836E-6

Figura 9. Iteraciones de $-2Ln(L_r)$ y componentes de varianza con SAS modelo mixto no balanceado.

En este caso la heredabilidad da un número mayor de 1, lo cual se debe a una sobre estimación, debido a un aumento de la varianza entre toros.

MODELO ANIMAL

El modelo animal es un modelo estadístico lineal, que permite predecir los BLUP de todos los animales del rebaño, usando los registros y genealogías disponibles en las bases de datos. El modelo animal, fue propuesto por el Dr. Charles Roy Henderson, como un método de evaluación genética en estados unidos, y posteriormente fue adoptado a nivel mundial.

El modelo más simple, que permite predecir el VG de los animales, para una característica es:

$$y_{ij} = \mu + a_i + e_{ij}$$

Donde y_{ij} es la variable respuesta, μ es el intercepto o media del rebaño, a_i es el VG y e_{ij} es el error residual. Bajo este modelo, los parámetros a_i y e_{ij} son aleatorios, mientras que μ es fijo. Las matrices de varianza y covarianza para los parámetros aleatorios son (Gutiérrez, 2010):

$$G = A\sigma_a^2 \text{ y } R = I\sigma_e^2$$

Donde G es la matriz de varianzas-covarianzas genéticas aditivas y R es una matriz de residuales. Si no se dispone de información genealógica de los animales en la base de datos, la matriz A= I por lo tanto, $G = I\sigma_a^2$ con lo cual su inversa es inmediata al ser diagonal.

Usando el modelo $y_{ij} = \mu + a_i + e_{ij}$ y asumiendo A= I, la ecuación que permite predecir el BLUP para cada animal es (Pérez, 2024):

$$a_i = h^2(y_{ij} - \mu)$$

Mientras que las soluciones para μ y e_{ij} son:

$$\mu = \frac{\sum y}{n}$$

$$e_{ij} = y_{ij} - (a_i + \mu) = y_{ij} - \hat{y}_{ij}$$

Por lo tanto, la suma de cuadrados residual es (Montgomery, 2022):

$$SC_e = \sum (y_{ij} - \hat{y}_{ij})^2$$

Y la correlación entre el BLUP real y el estimado es (Pérez, 2025):

$$r_{\hat{a}a} = \sqrt{h^2}$$

Y el error estándar de los BLUP para este modelo es (Mrode, 2005):

$$EE(a) = \sqrt{(1 - r_{\hat{a}a}^2)\sigma_a^2}$$

Ejemplo 2.

Se requiere estimar el BLUP de 5 animales para peso al destete, se asume $h^2=0.5$ ($\sigma_a^2 = 1$ y $\sigma_e^2 = 1$). Los datos se presentan en la tabla 3:

animal	y
1	205
2	200
3	195
4	188
5	170

Tabla 3. datos para PD del ejemplo 2

La media para estos cinco datos viene dada por:

$$\mu = \frac{205 + 200 + 195 + 188 + 170}{5} = 191.6$$

Por lo tanto, los BLUP, los residuales, la correlación y los EE para los 5 animales se presentan en la tabla 4:

animal	VG	Residual	Correlación	EE
1	$0.5(205-191.6)=6.7$	$205-(6.7+191.6)=6.7$	$\sqrt{0.5} = 0.707$	$\frac{\sqrt{(1 - 0.707^2)}1}{= 0.707}$
2	$0.5(200-191.6)=4.2$	$200-(4.2+191.6)=4.2$	$\sqrt{0.5} = 0.707$	$\frac{\sqrt{(1 - 0.707^2)}1}{= 0.707}$

3	$0.5(195-191.6)=1.7$	$195-(1.7+191.6)=1.7$	$\sqrt{0.5} = 0.707$	$\frac{\sqrt{(1-0.707^2)}1}{= 0.707}$
4	$0.5(188-191.6)=-1.8$	$188-(-1.8+191.6)=-1.8$	$\sqrt{0.5} = 0.707$	$\frac{\sqrt{(1-0.707^2)}1}{= 0.707}$
5	$0.5(170-191.6)=-10.8$	$170-(-10.8+191.6)=-10.8$	$\sqrt{0.5} = 0.707$	$\frac{\sqrt{(1-0.707^2)}1}{= 0.707}$

Tabla 4. BLUP, residual, correlación y EE para el ejemplo 2

Y la suma de cuadrados residual es:

$$SC_e = 6.7^2 + 4.2^2 + 1.7^2 + (-1.8)^2 + (-10.8)^2 = 185.3$$

Estos resultados indican que el animal 1 esta 6.7kg por encima de la media, por lo tanto, es el mejor genéticamente, mientras el animal 5 esta 10.8kg por debajo de la media, con lo cual queda calificado como el peor. La correlación entre el BLUP real y el estimado es de 70.7%, con lo cual entra en el rango de buena.

Uso del WOMBAT-REML

En la figura 10, se presenta el archivo genealógico y de datos para el ejemplo 2:

ejemplo: Bloc de notas						ejemplo: Bloc de notas					
Archivo	Edición	Formato	Ver	Ayuda		Archivo	Edición	Formato	Ver	Ayuda	
1	0	0				1	0	0	205		
2	0	0				2	0	0	200		
3	0	0				3	0	0	195		
4	0	0				4	0	0	188		
5	0	0				5	0	0	170		

Figura 10. Archivo genealógico y de datos (WOMBAT)para el ejemplo 2.

En la figura 11, se presenta el archivo de parametros para el ejemplo 2:

```

ejemplo: Bloc de notas
Archivo Edición Formato Ver Ayuda
COM MA A=I

PEDS ejemplo.ped
DAT ejemplo.dat

animal
sire
dam
PD

END

ANAL UNI
MODEL
RAN animal NRM
TRAIT PD
END MOD

var animal 1
1
var residual 1
1

```

Figura 11. Archivo de parámetros para el ejemplo 2.

La opción usada en wombat para calcular las predicciones genéticas fue –BLUP. La Salida del WOMBAT para BLUP, EE y las correlaciones del ejemplo se presenta en figura 12:

RnSoln_animal: Bloc de notas

Run_No.	Original_ID	Tr	Solution	St.Error	xxxxxx	Inbr_%
1	1	1	6.7000000	0.70710678	0.707	0.000
2	2	1	4.2000000	0.70710678	0.707	0.000
3	3	1	1.7000000	0.70710678	0.707	0.000
4	4	1	-1.8000000	0.70710678	0.707	0.000
5	5	1	-10.800000	0.70710678	0.707	0.000

Figura 12. Salida WOMBAT para VG, EE y correlaciones del ejemplo 2.

Estos resultados son idénticos a los encontrados manualmente. Los residuales y la suma de cuadrados residual calculados por WOMBAT se presentan en la figura 13:

Residuals: Bloc de notas					SumEstimates: Bloc de notas				
Archivo	Edición	Formato	Ver	Ayuda	Archivo	Edición	Formato	Ver	Ayuda
Residual			Y-hat	Observation	Adj.pheno.	SS residuals	185.300		
6.70000			198.300	205.000	205.000				
4.20000			195.800	200.000	200.000				
1.70000			193.300	195.000	195.000				
-1.80000			189.800	188.000	188.000				
-10.8000			180.800	170.000	170.000				

Figura 13. Salida WOMBAT para los residuales y SSe del ejemplo 2

Resultados idénticos a los encontrados de manera manual

En el ejemplo 2, el único factor fijo es μ , pero es bien sabido que la variación fenotípica es atribuible a varios factores, por ejemplo, el sexo del animal (por lo general los machos son más pesados que las hembras), por lo tanto, el nuevo modelo que introduce el efecto del sexo es:

$$y_{ijk} = \mu + a_i + b_j + e_{ijk}$$

Donde b_j es el efecto del sexo del animal. En el caso de múltiples efectos fijos, el modelo se puede expresar en algebra de matrices de la siguiente forma (Mrode, 2014):

$$y = Xb + Za + e$$

Donde a es un vector de predicciones desconocidas para los BLUP de los animales. Bajo este modelo matricial el vector de residuales es:

$$e = y - (Xb + Za)$$

Y la suma de cuadrados de los residuales es:

$$SCe = (y - Xb - Za)'(y - Xb - Za)$$

Y el BLUP viene dado por (Blasco, 2021):

$$a_i = GZ'V^{-1}(y - Xb)$$

Donde $V = ZGZ' + R$ es la matriz de varianzas-covarianzas fenotípica del modelo animal.

Si asumimos que $A \neq I$ (es decir, que utilizamos información de parentesco) entonces G no es una matriz diagonal, por lo tanto, la inversa de V no es inmediata y demanda grandes esfuerzos calcularla. Bajo este problema el Dr C. Henderson propuso usar un sistema de ecuaciones donde no se tuvieran que utilizar inversas de matrices, por lo que se sustituye el cálculo de G y R por la inversa de A la cual es calculada sin necesidad de inversión matricial. Estas ecuaciones llamadas ecuaciones normales de Henderson son (Martínez y col, 2012):

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\alpha \end{bmatrix} \begin{bmatrix} \hat{b}_i \\ \hat{a}_i \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Donde a es un escalar que introduce la heredabilidad en las ecuaciones, el cual, tiene la siguiente expresión matemática en el modelo animal:

$$\alpha = \frac{1 - h^2}{h^2}$$

Y la solución de las ecuaciones normales de Henderson viene dada por:

$$\begin{bmatrix} \hat{b}_i \\ \hat{a}_i \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\alpha \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Donde la solución del vector de parámetros y predicciones desconocidas es el producto de la inversa de la matriz de coeficientes con el vector del lado derecho, con lo cual se puede pensar en una solución dificultosa debido a la inversa inmediata, sin embargo, al ser este un sistema de ecuaciones lineales, puede ignorarse la inversión directa y calcular las soluciones con cualquier método numérico iterativo.

De forma general, para calcular la correlación entre el BLUP real y las predicciones se utiliza la siguiente ecuación (Aranguren y Román, 2014):

$$r_{\hat{a}a} = \sqrt{1 - C_{ij}\alpha}$$

Donde C_{ij} es el elemento inverso de la diagonal del bloque de animales en la matriz de coeficientes de las ecuaciones normales de Henderson.

Ejemplo 3.

Se utilizan la misma base de datos que en el ejemplo 2, con la diferencia que se identificó el sexo de cada animal y la genealogía, los datos se presentan en la tabla 5:

animal	Padre	Sexo	y
1	0	M	205
2	0	M	200
3	0	H	195
4	1	H	188
5	2	H	170

Tabla 5. Datos ejemplo 3.

Para construir las ecuaciones normales de Henderson tenemos que calcular los productos matriciales de la matriz de coeficientes y el vector del lado derecho, los cuales son contruidos igual que en las ecuaciones normales de mínimos cuadrados:

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu & b_1 & b_2 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}$$

$$Z'X = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu & b_1 & b_2 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

$$X'Z = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Para construir la inversa de es necesario aplicar las reglas de Henderson (Henderson, 1976), para este caso es:

$$A^{-1} = \begin{bmatrix} 1 + \frac{1}{3} & 0 & 0 & -2/3 & 0 \\ 0 & 1 + \frac{1}{3} & 0 & 0 & -2/3 \\ 0 & 0 & 1 & 0 & 0 \\ -2/3 & 0 & 0 & 4/3 & 0 \\ 0 & -2/3 & 0 & 0 & 4/3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} & 0 & 0 & -2/3 & 0 \\ 0 & \frac{4}{3} & 0 & 0 & -2/3 \\ 0 & 0 & 1 & 0 & 0 \\ -2/3 & 0 & 0 & 4/3 & 0 \\ 0 & -2/3 & 0 & 0 & 4/3 \end{bmatrix}$$

Y el escalar a es:

$$\alpha = \frac{1 - 0.5}{0.5} = 1$$

Por lo tanto $Z'Z + A^{-1}\alpha$ es:

$$Z'Z + A^{-1}\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \frac{4}{3} & 0 & 0 & -\frac{2}{3} & 0 \\ 0 & \frac{4}{3} & 0 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & 0 & 0 \\ -\frac{2}{3} & 0 & 0 & \frac{4}{3} & 0 \\ 0 & -\frac{2}{3} & 0 & 0 & \frac{4}{3} \end{bmatrix} \quad (1)$$

$$= \begin{bmatrix} 7/3 & 0 & 0 & -2/3 & 0 \\ 0 & 7/3 & 0 & 0 & -2/3 \\ 0 & 0 & 2 & 0 & 0 \\ -2/3 & 0 & 0 & 7/3 & 0 \\ 0 & -2/3 & 0 & 0 & 7/3 \end{bmatrix}$$

y el vector de lado derecho viene dado por:

$$(X'y) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{pmatrix} 205 \\ 200 \\ 195 \\ 188 \\ 170 \end{pmatrix} = \begin{pmatrix} 958 \\ 405 \\ 553 \end{pmatrix}$$

$$(Z'y) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 205 \\ 200 \\ 195 \\ 188 \\ 170 \end{pmatrix} = \begin{pmatrix} 205 \\ 200 \\ 195 \\ 188 \\ 170 \end{pmatrix}$$

Por lo tanto, las ecuaciones quedan definidas de la siguiente forma:

$$\begin{bmatrix} 5 & 2 & 3 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 7/3 & 0 & 0 & -2/3 & 0 \\ 1 & 1 & 0 & 0 & 7/3 & 0 & 0 & -2/3 \\ 1 & 0 & 1 & 0 & 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & -2/3 & 0 & 0 & 7/3 & 0 \\ 1 & 0 & 1 & 0 & -2/3 & 0 & 0 & 7/3 \end{bmatrix} \begin{bmatrix} \mu \\ b_1 \\ b_2 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} = \begin{pmatrix} 958 \\ 405 \\ 553 \\ 205 \\ 200 \\ 195 \\ 188 \\ 170 \end{pmatrix}$$

Debido a la introducción de factores fijos diferentes de la media, el determinante de la matriz de coeficiente es 0, por lo tanto, existe una dependencia lineal en el sistema de ecuaciones, por lo que se debe eliminar un efecto, en este caso se utilizó $\mu = 0$, por lo tanto, se elimina la fila y la columna para μ y en el vector de lado derecho el primer valor, por lo que las ecuaciones se reducen y quedan determinadas de la siguiente forma:

$$\begin{bmatrix} 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 7/3 & 0 & 0 & -2/3 & 0 \\ 1 & 0 & 0 & 7/3 & 0 & 0 & -2/3 \\ 0 & 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & -2/3 & 0 & 0 & 7/3 & 0 \\ 0 & 1 & 0 & -2/3 & 0 & 0 & 7/3 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} = \begin{pmatrix} 405 \\ 553 \\ 205 \\ 200 \\ 195 \\ 188 \\ 170 \end{pmatrix}$$

Y la solución de este sistema es:

$$\begin{bmatrix} 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 7/3 & 0 & 0 & -2/3 & 0 \\ 1 & 0 & 0 & 7/3 & 0 & 0 & -2/3 \\ 0 & 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & -2/3 & 0 & 0 & 7/3 & 0 \\ 0 & 1 & 0 & -2/3 & 0 & 0 & 7/3 \end{bmatrix}^{-1} \begin{pmatrix} 405 \\ 553 \\ 205 \\ 200 \\ 195 \\ 188 \\ 170 \end{pmatrix} = \begin{pmatrix} 203.83333 \\ 184.33333 \\ 1.03333 \\ -3.7 \\ 5.33333 \\ 1.86666 \\ -7.2 \end{pmatrix}$$

Y los residuales son:

$$e = \begin{pmatrix} 205 \\ 200 \\ 195 \\ 188 \\ 170 \end{pmatrix} - \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} 203.83333 \\ 184.33333 \end{pmatrix} - \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1.03333 \\ -3.7 \\ 5.33333 \\ 1.86666 \\ -7.2 \end{pmatrix} = \begin{pmatrix} 0.133334 \\ -0.133334 \\ 5.33334 \\ 1.80001 \\ -7.133333 \end{pmatrix}$$

Por lo tanto, la suma de cuadrados de los residuos es:

$$SCe = \begin{pmatrix} 0.133334 \\ -0.133334 \\ 5.33334 \\ 1.80001 \\ -7.133333 \end{pmatrix}' \begin{pmatrix} 0.133334 \\ -0.133334 \\ 5.33334 \\ 1.80001 \\ -7.133333 \end{pmatrix} = 82.6045$$

Y la confiabilidad del BLUP para los 5 animales es:

$$r_{\hat{a}_1 a_1} = \sqrt{1 - 0.7125(1)} = 0.53$$

$$r_{\hat{a}_2 a_2} = \sqrt{1 - 0.7125(1)} = 0.53$$

$$r_{\hat{a}_3 a_3} = \sqrt{1 - 0.66666(1)} = 0.57$$

$$r_{\hat{a}_4 a_4} = \sqrt{1 - 0.65(1)} = 0.59$$

$$r_{\hat{a}_5 a_5} = \sqrt{1 - 0.65(1)} = 0.59$$

El animal 3 al no conocerle genealogía y tener dato propio, puede calcularse su BLUP despejando de las ecuaciones a_3 :

$$1b_2 + 2a_3 = 195 \rightarrow a_3 = \frac{195 - 184.3333}{2} = \frac{10.6667}{2} = 5.33335$$

Es decir que la predicción viene dada por (Pérez, 2025):

$$a_3 = \frac{y - b_2}{1 + \alpha}$$

Si descomponemos esta ecuación:

$$a_3 = (1 - b_2)(1 + \alpha)^{-1} = (1 - b_2) \left(1 + \frac{1 - h^2}{h^2}\right)^{-1} = (1 - b_2) \left(\frac{h^2 + 1 - h^2}{h^2}\right)^{-1}$$

La expresión $(1 - b_2) \left(\frac{h^2 + 1 - h^2}{h^2}\right)^{-1}$ puede organizarse de manera que:

$$a_3 = (1 - b_2) \left(\frac{h^2 + 1 - h^2}{h^2}\right)^{-1} = (1 - b_2) \left(\frac{1}{h^2}\right)^{-1} = (1 - b_2)h^2$$

Por lo tanto:

$$a_3 = 0.5(195 - 184.33333) = 5.33335$$

Aunque esta fórmula final parece sencilla, implica conocer el valor de b_2 por lo tanto, no es muy usada.

Si se despeja el BLUP de un animal con genealogía, por ejemplo, el 4, se tiene:

$$\begin{aligned} b_2 - \frac{2}{3}(a_1) + \frac{7}{3}(a_4) &= 188 \rightarrow a_4 = \frac{188 - b_2 + \frac{2}{3}(a_1)}{\frac{7}{3}} \\ &= \frac{188 - 184.3333 + \frac{2}{3}(1.03333)}{2.3333} = 1.866 \end{aligned}$$

Este valor para el BLUP del animal 4, es idéntico a los encontrados usando las ecuaciones normales de Henderson, pero implica conocer y el BLUP del padre.

Uso del WOMBAT

En la figura 14 se presenta el archivo genealógico y de datos:

ejemplo: Bloc de notas						ejemplo: Bloc de notas					
Archivo	Edición	Formato	Ver	Ayuda		Archivo	Edición	Formato	Ver	Ayuda	
1	0	0				1	0	0	1	205	
2	0	0				2	0	0	1	200	
3	0	0				3	0	0	2	195	
4	1	0				4	1	0	2	188	
5	2	0				5	2	0	2	170	

Figura 14. Archivo genealógico y de datos del ejemplo 3.

En la figura 15, se presenta el archivo de parámetros del ejemplo 3:

```
ejemplo: Bloc de notas
Archivo Edición Formato Ver Ayuda
COM ejemplo modelo animal 2

PED ejemplo.ped
DAT ejemplo.dat

animal 5
sire
dam
sexo 2
PESO

END

ANAL UNI
MODEL
RAI animal NRM
FIX sexo
TRAIT PESO
END MOD

VAR animal 1
2
VAR error 1
2
```

Figura 15. Archivo de parámetros ejemplo 3.

Y la opción usada en wombat para calcular las predicciones genéticas fue –BLUP. La Salida del WOMBAT para BLUP, EE y las correlaciones del ejemplo 3 se presentan en la figura 16:

RnSoln_animal: Bloc de notas

Archivo	Edición	Formato	Ver	Ayuda			
Run_No.	Original_ID	Tr	Solution	St.Error	xxxxxx	Inbr_%	
1	1	1	1.0333333	1.1937336	0.536	0.000	
2	2	1	-3.7000000	1.1937336	0.536	0.000	
3	3	1	5.3333333	1.1547005	0.577	0.000	
4	4	1	1.8666667	1.1401754	0.592	0.000	
5	5	1	-7.2000000	1.1401754	0.592	0.000	

Figura 16. Salida del WOMBAT para los BLUP, EE y las correlaciones del ejemplo 3.

Resultados idénticos a los encontrados manualmente. Finalmente, los residuales calculados por Wombat son:

Residuals: Bloc de notas					SumEstimates: Bloc de notas				
Archivo	Edición	Formato	Ver	Ayuda	Archivo	Edición	Formato	Ver	Ayuda
Residual		Y-hat	Observation	Adj.pheno.	SS residuals		82.6044		
0.133333		204.867	205.000	192.767					
-0.133333		200.133	200.000	187.767					
5.333333		189.667	195.000	202.267					
1.800000		186.200	188.000	195.267					
-7.133333		177.133	170.000	177.267					

Figura 17. Salida del WOMBAT para los residuales y la SSe figura 18

Encontrando nuevamente resultados idénticos a los calculados manualmente.

ESTIMACIÓN DE COMPONENTES USANDO UN MODELO ANIMAL

Para estimar los componentes de varianza usando un modelo animal, es necesario utilizar métodos numéricos iterativos para encontrar los valores de los parámetros que maximicen la verosimilitud logarítmica restringida, una explicación detallada del caso puede encontrarse en Searle et al (1992) y Sorensen (2023).

BIBLIOGRAFÍA

Aranguren, A y Román, R (2014). El modelo animal simple: una metodología para los genetistas. Logros & Desafíos de la Ganadería Doble Propósito. GIRAZ. 120-136.

Blasco, A. (2017). Correction to: Bayesian Data Analysis for Animal Scientists. In Bayesian Data Analysis for Animal Scientists: The Basics (pp. E1-E2). Springer International Publishing. https://doi.org/10.1007/978-3-319-54274-4_11

Blasco, A. (2021). Mejora genética animal. Editorial síntesis. De: <https://www.sintesis.com/libro/mejora-genetica-animal>

Boldman, K. G., L. A. Kriese, L. D. Van Vleck, C. P. Van Tassell and S. D. Kachman. (1995). A Manual for Use of MTDFREML. A Set of Programs To Obtain Estimates of Variances and Covariances [DRAFT]. U.S. Department of Agriculture, Agricultural Research Service.

Castejón-Sandoval, O. (2011). Diseño y análisis de experimentos con statistix. <https://libros.fondoeditorial.uru.edu/index.php/lb/catalog/book/castejonsandoval2011>

Cerón Muñoz, M. F., Elzo, M. A., Vergara G, O., Arboleda, E., & Hurtado, N. (2012). Modelación aplicada a las ciencias animales: II. Evaluaciones genéticas Fondo editorial Biogénesis. <https://revistas.udea.edu.co/index.php/biogenesis/article/view/326021?articlesBySameAuthorPage=1>

Gutiérrez, P. (2010). Iniciación a la valoración genética animal. Editorial Complutense. <https://docta.ucm.es/entities/publication/8d1fdea5-0f8f-40cd-8626-dfb069f69fa8>

Henderson CR. (1976). A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics; 32:69–83. <https://doi.org/10.2307/2529339>

Henderson CR. (1984) Applications of linear models in animal breeding. 1sted. Canada: University of Guelph. <https://www.ancp.org.br/wp/wp-content/uploads/2011/02/Livro-Applications-of-Linear-Models-in-Animal-Breeding.pdf>

Johnson, D. and Thompson, R. (1995). Restricted Maximum Likelihood Estimation of Variance Components for Univariate Animal Models Using Sparse Matrix Techniques and Average Information. J. Dairy Sci. 78: 449-456. [https://doi.org/10.3168/jds.S0022-0302\(95\)76654-1](https://doi.org/10.3168/jds.S0022-0302(95)76654-1)

Littell, R., Milliken, G., Stroup, W., Wolfinger, R., & Schabenberger, O. (2006). SAS for mixed model. (2nd ed. ed.). SAS press. <https://support.sas.com/documentation/onlinedoc/stat/141/mixed.pdf>

Martínez, C. Manrique, C. Elzo, M. (2012). La evaluación genética de vacunos: una percepción histórica. Rev Colomb Cienc Pecu. 5:293-311 <https://doi.org/10.17533/udea.rccp.324760>

Meyer, K. (1989). Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. *Genetics Selection Evolution*, 21(3), 317. <https://doi.org/10.1186/1297-9686-21-3-317>

Meyer, K. (2007). WOMBAT: a tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *J Zhejiang Univ Sci B*, 8(11), 815-821. <https://doi.org/10.1631/jzus.2007.B0815>

Montgomery, D., & St, C. (2022). *Design and Analysis of Experiments*, 9th Edition.

Mrode, R. (2014). *Linear models for the prediction of animal breeding values*.

Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3), 545-554. <https://doi.org/10.1093/biomet/58.3.545>

Peréz González, J. (2025). Guía rápida de análisis genéticos bajo modelos mixtos con WOMBAT. *Revista Multidisciplinar Epistemología de las Ciencias*, 2, 19-39. <https://doi.org/10.71112/pw47g958>

Pérez, J. (2024). *Estadística Aplicada al Mejoramiento Genético Animal*. <https://libros.fondeditorial.uru.edu/index.php/lb/catalog/book/perez2024>

Román, R., & Aranguren, A. (2014). *Evaluación genética de reproductores: logros y desafíos*. Fundación GIRARZ.

Searle, S., Casella, G., & McCulloch, C. (1992). *Variance Components*. John Wiley & Sons, Inc. . <https://doi.org/10.1002/9780470316856>

Sorensen, D. (2023). *Statistical Learning in Genetics*. Springer Nature. <https://link.springer.com/book/10.1007/978-3-031-35851-7>

Sorensen D, Gianola D. *Métodos de verossimilhança, bayesianos e MCMC em genética quantitativa*. Nova York: Springer. 2002.