

MODELANDO EL DÍA DE CONTROL: NUEVA TÉCNICA ESTADÍSTICA EN EVALUACIÓN GENÉTICA DE GANADO BOVINO LECHERO¹

Test day model: A new statistical tool for genetic evaluation of dairy cattle

Héctor A. Uribe²

ABSTRACT

Molecular genetics is making enormous improvement in identifying major effect genes or quantitative trait *loci*. However, quantitative genetics still plays an important role on modern animal breeding. Animal models using a single record per lactation assume that environmental effects do not vary through the entire lactation. A test day model permits environmental changes through the lactation. Test day models were developed in Canada and are being used for genetic evaluation in at least two countries. Chile, as a dairy frozen semen importer, uses genetic evaluations done abroad. Although domestic application of such evaluations is limited, it is necessary a minimal understanding of the methodology used in the estimation of breeding values. The objective of this work is to discuss basic aspects of the statistical theory used in test day models. A random regression test day model example is presented. A test day random regression model has been implemented in Canada for dairy cattle genetic evaluation, this allows better modeling of the lactation curve. It is concluded that breeding value estimation, using test day models, is better as compared to single record animal models, however, computer requirements are increased. Persistence, as a byproduct of random regression test day models, may be an important production trait.

Key words: breeding values, lineal models, random regression.

RESUMEN

Aunque la genética molecular está haciendo importantes avances en la identificación de genes o grupos de genes con acción sobre características productivas, la genética cuantitativa aún presenta un rol importante en el mejoramiento animal moderno. Los modelos animales tradicionales, usando un registro de producción por lactancia, asumen que los efectos ambientales que afectan una lactancia no varían a lo largo de ésta. Un modelo de día de control permite variaciones del ambiente a través de una lactancia. Ésta es una herramienta estadística que está siendo utilizada en evaluación de ganado de leche en al menos dos países. Chile, como un importador de semen congelado bovino, usa evaluaciones genéticas realizadas en el extranjero. Aunque la aplicación de dichas evaluaciones es muy limitada en condiciones locales, es necesario un mínimo conocimiento de la metodología usada para estimar valores genéticos. El objetivo de este trabajo es entregar algunos aspectos básicos de la teoría estadística usada en modelos de día de control. Se presenta un ejemplo usando regresores aleatorios, similar al sistema canadiense de evaluación de ganado lechero, esto modela la curva de producción de leche con mayor exactitud ya que permite

¹Recepción de originales: 23 de noviembre de 1999.

²Instituto de Investigaciones Agropecuarias, Centro Regional de Investigación Remehue, Casilla 24-0, Osorno, Chile.
E-mail: huribe@remehue.inia.cl

desviaciones individuales de cada vaca. Se concluye que la estimación de valores genéticos usando un modelo de día de control es más exacta, pero la demanda computacional es mucho mayor. La persistencia de producción, como un subproducto de un modelo de día de control usando regresiones aleatorias, puede ser una característica de interés.

Palabras clave: valores genéticos, modelos lineales, regresiones aleatorias.

INTRODUCCIÓN

El desarrollo y aplicación de la genética molecular avanza muy rápidamente, identificando genes o grupos de ellos y sus acciones sobre características productivas, sin embargo, la mayor parte del avance genético en ganado de leche aún debe sus logros al desarrollo e implementación de herramientas estadísticas aplicadas a genética cuantitativa. Estas técnicas estadísticas, que a primera vista parecen abstractas, han probado su efectividad y gran potencial a nivel práctico. El avance genético en producción de leche de ganado Holstein en países desarrollados ha sido de aproximadamente 100 kg de leche por año (Wiggans, 1997), es decir en promedio las terneras nacidas en este año tienen un potencial genético de producir 100 kg más de leche en la primera lactancia, que las nacidas hace un año atrás. Estos logros incentivan a los investigadores quienes, financiados por productores y asociaciones de criadores, prueban nuevas teorías que puedan ayudarlos a separar el componente ambiental del componente genético, basados en registros de producción y genealógicos.

Una nueva tecnología está siendo usada en evaluación genética en algunos países desarrollados; se trata de los modelos para producción del día de control, entendiéndose como tal a los modelos que usan cada control lechero como una observación separada y particular de cada animal (Ali y Schaeffer, 1987; Ptak y Schaeffer, 1993; Reents *et al.*, 1994; Jara, 1999). Este tipo de modelos ya está en uso en forma comercial en el sistema de evaluación genética de Alemania desde agosto de 1998. Aunque gran parte de la tecnología estadística para los modelos de producción del día de control fue desarrollada en Canadá

(Ptak y Schaeffer, 1993), este es el segundo país en adoptar su uso en forma comercial. Desde la evaluación genética de febrero de 1999, Canadá está expresando los valores genéticos de sus animales de leche basados en resultados de modelos para producción del día de control. España está trabajando en la implementación de un modelo del día de control y pareciera ser el tercer país que usará en forma comercial este tipo de modelos (María J. Carabaño, 1999, Depto. de Mejora Genética y Biotecnología, INIA, España, comunicación personal). En Chile, el uso de un modelo para el día de control usando regresores fijos y datos de producción de leche nacional fue reportado por primera vez por Jara (1999).

En Chile, aproximadamente el 70% del semen congelado usado en ganado de leche es importado, esto indica que los valores genéticos publicados en los catálogos de comercialización serán un producto de esta nueva forma de evaluación. El conocimiento básico de la metodología utilizada para evaluar los reproductores ingresados al país debe ser la mínima aspiración de los agentes nacionales involucrados en la importación de germoplasma de leche.

En ganado de leche, los modelos estadísticos tradicionales de evaluación genética, usan un registro de producción proveniente de una combinación de 4 a 10 controles lecheros realizados durante la lactancia del animal (Misztal *et al.*, 1992). Estos controles individuales son acumulados utilizando diferentes estrategias (Swalve, 1998), para obtener un registro por lactancia. El registro acumulado de la lactancia es generalmente corregido por otros factores ambientales (época de parto, número de lactancia) y estandarizado a una lactancia de 305 días de un animal

adulto, usando factores de corrección aditivos y/o multiplicativos, que muchas veces no han sido generados con información local. A las vacas cuyas lactancias están en progreso se les extrapola su producción asumiendo una curva de lactancia estándar (Wood, 1967); entre otras cosas, este procedimiento asume que los factores no genéticos que afectan todos los controles lecheros dentro de una misma lactancia son los mismos. Al analizar el manejo de un rebaño en producción de leche es claro que el supuesto anterior no se cumple. Sin embargo, con todas sus limitaciones, este sistema de evaluación tradicional ha realizado buenos aportes, considerando el avance genético obtenido en las dos últimas décadas en la producción de leche, grasa y proteína (Wiggans, 1997). Sin embargo, el sistema tradicional es débil y poco flexible al no considerar que las condiciones ambientales cambian desde el inicio al final de la lactancia.

La nueva metodología para analizar la producción del día de control considera el hecho de que cada vaca esté sometida a un manejo (alimentación, clima, etc.) que cambia a través de su lactancia; considera además que sus contemporáneas no son necesariamente las mismas a través de una misma lactancia. La implementación de modelos de producción del día de control no requiere controles de leche adicionales, es decir, no involucra nuevas inversiones o trabajo adicional para las empresas de control lechero. Esta metodología hace un mejor uso de los datos de control lechero existentes. Los registros de vacas que no han terminado su lactancia no necesitan ser extendidos y pueden ser usados directamente aportando información que puede ayudar a reducir el intervalo entre generaciones (Swalve, 1998).

El objetivo de este trabajo es entregar antecedentes de la teoría estadística que se aplica en la estimación de los valores genéticos que aparecen en los catálogos de reproductores de bovinos lecheros de Alemania y Canadá.

METODOLOGÍA

El modelo de evaluación genética canadiense usa regresiones fijas y aleatorias (Henderson, 1984; Jamrozik y Schaeffer, 1997; Uribe, 1998), esto permite obtener valores genéticos para cada día de lactancia. Un subproducto de esta metodología es una evaluación de persistencia de cada animal. Persistencia es una característica presentada como una función de la pendiente de la curva de producción de leche entre los días 60 y 280 de lactancia. Refleja la producción de leche, grasa o proteína en el día 280 en proporción al día 60 (máxima producción). En otras palabras, describe la disminución de leche, grasa o proteína después de alcanzar el máximo de producción, el cual ocurre aproximadamente a los 60 días de lactancia. El modelo usado en Alemania no tiene regresores aleatorios por lo tanto no es posible obtener una medida de persistencia.

Para ilustrar el método se presenta un ejemplo numérico. Los datos del Cuadro 1 son datos ficticios de producción de leche del día de control de 4 vacas en dos rebaños, también se conoce el número de días de lactación al momento de cada control. En total son 10 registros, ya que las vacas números 1, 3 y 4 tienen más de un control de producción de leche. En los datos del ejemplo, la vaca número 4 tiene 4 controles de producción, dos fueron hechos en el segundo rebaño y los últimos fueron hechos en el primer rebaño. El Cuadro 2 muestra la estructura del pedigrí de los animales. En este caso la madre de la vaca 1 es desconocida.

Por simplicidad se puede asumir que la curva de producción de leche puede ser descrita por el efecto de días en leche más el efecto del logaritmo natural de 305 dividido por días en leche (Schaeffer y Dekkers, 1994). Existen varias funciones que describen la curva de producción de leche en ganado de leche (Jamrozik *et al.*, 1997), quizás la más conocida sea la presentada por Wood (1967).

Cuadro 1. Registros de producción de cuatro vacas

Table 1. Production records of four cows

Número vaca	Días de ordeño	Rebaño	Producción (kg d ⁻¹)
1	35	1	15
1	64	1	20
1	98	1	17
2	60	1	23
3	60	2	25
3	95	2	24
4	35	2	16
4	65	2	18
4	95	1	22
4	124	1	20

Cuadro 2. Estructura del pedigrí de los animales

Table 2. Pedigree structure of animals

Vaca	Padre	Madre
1	5	¿?
2	6	1
3	6	2
4	7	2

El modelo estadístico para explicar cada una de los controles de producción es el siguiente:

$$y_{ijk} = R_i + \beta_1 x_1 + \beta_2 x_2 + a_j + \alpha_{1j} x_1 + \alpha_{2j} x_2 + e_{ijk}$$

Donde:

- y_{ijk} = es un registro de producción diaria hecho en el i-ésimo rebaño (i = 1,2) por la j-ésima vaca (j = 1,2,3,4).
- R_i = es el efecto fijo del i-ésimo rebaño.
- x_1 = son los días de lactancia.
- x_2 = $\ln(305/x_1)$.
- β_1 y β_2 = son los coeficientes de regresión fijos, comunes a las cuatro vacas del ejemplo, de producción en un

día de control sobre x_1 y x_2 , respectivamente.

a_j, α_{1j} y α_{2j} = son los coeficientes de regresión aleatorios, específicos para cada vaca, relacionando un control lechero con un intercepto, x_1 y x_2 , respectivamente. En forma conjunta representan el efecto genético aditivo aleatorio de la j-ésima vaca.

e_{ijk} = es el efecto residual aleatorio.

La estructura de (co)varianzas entre los efectos aleatorios del modelo es:

$$V \begin{pmatrix} a \\ \alpha_1 \\ \alpha_2 \\ e \end{pmatrix} = \begin{pmatrix} Ag_{aa} & Ag_{a\alpha_1} & Ag_{a\alpha_2} & 0 \\ Ag_{\alpha_1 a} & Ag_{\alpha_1 \alpha_1} & Ag_{\alpha_1 \alpha_2} & 0 \\ Ag_{\alpha_2 a} & Ag_{\alpha_2 \alpha_1} & Ag_{\alpha_2 \alpha_2} & 0 \\ 0 & 0 & 0 & D\sigma_e^2 \end{pmatrix} = \begin{pmatrix} A * G & 0 \\ 0 & D\sigma_e^2 \end{pmatrix}$$

donde:

- A = es la matriz de parentesco genético aditivo entre los animales involucrados en el análisis (son siete animales ya que es un modelo animal y se incluyen todos los animales del archivo de pedigrí, Cuadro 2).
- G = es una matriz que contiene las (co)varianzas entre los regresores aleatorios incluidos en el modelo (g_{ij}).
- D = es una matriz diagonal, la cual puede tener diferentes valores, permitiendo que la varianza residual sea diferente a medida que la lactancia progresa.

σ_e^2 = es la varianza residual.

De acuerdo al modelo, los regresores aleatorios, los cuales representan el efecto genético aditivo, no son independientes y existen covarianzas entre ellos, las cuales están ponderadas según el grado de parentesco genético aditivo entre los animales. Además, el modelo asume que los efectos residuales son independientes de los efectos genéticos aditivos.

Una de las complicaciones de este tipo de modelos es la estimación de los elementos de la matriz G y la varianza residual. En la actualidad estos parámetros genéticos se estiman usando principios derivados del teorema de Bayes mediante el muestreo de Gibbs (Jamrozik y Schaeffer, 1997; Jamrozik *et al.*, 1997; Uribe, 1997, 1998; Jara, 1999).

En notación matricial el modelo estadístico es:

$$y = Xb + Z_a a + Z_1 \alpha_1 + Z_2 \alpha_2 + e$$

donde:

- y = es el vector de observaciones de producciones del día de control.
- X = es una matriz relacionando cada observación con el rebaño donde se hizo la observación; además, esta matriz contiene los elementos de x_1 y x_2 .
- b = es un vector desconocido que contiene las estimaciones para los dos rebaños y los coeficientes de regresión β_1 y β_2 .
- Z_a = es una matriz de diseño, conteniendo ceros y unos, que relaciona cada observación con el animal que la hizo
- a = es un vector desconocido que contiene el intercepto del efecto genético aditivo para cada animal
- Z_1 = es una matriz relacionando cada observación con los días de lactancia (x_1).
- Z_2 = es una matriz relacionando cada observación con x_2 .

- α_1 y α_2 = son dos vectores desconocidos que contienen los regresores aleatorios, para cada animal, de producción del día de control sobre x_1 y x_2 respectivamente.
- e = es un vector de efectos residuales no explicados por los efectos incluidos en el modelo.

Con los datos del ejemplo la matrices X, Z_a , Z_1 y Z_2 son:

$$X = \begin{pmatrix} 1 & 1 & 0 & 35 & 2,16 \\ 1 & 1 & 0 & 64 & 1,56 \\ 1 & 1 & 0 & 98 & 1,14 \\ 1 & 1 & 0 & 60 & 1,63 \\ 1 & 0 & 1 & 60 & 1,63 \\ 1 & 0 & 1 & 95 & 1,17 \\ 1 & 0 & 1 & 35 & 2,17 \\ 1 & 0 & 1 & 65 & 1,55 \\ 1 & 1 & 0 & 95 & 1,17 \\ 1 & 1 & 0 & 124 & 0,9 \end{pmatrix} \quad Z_a = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$Z_1 = \begin{pmatrix} 0 & 0 & 0 & 35 & 0 & 0 & 0 \\ 0 & 0 & 0 & 64 & 0 & 0 & 0 \\ 0 & 0 & 0 & 98 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 60 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 60 & 0 \\ 0 & 0 & 0 & 0 & 0 & 95 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 35 \\ 0 & 0 & 0 & 0 & 0 & 0 & 65 \\ 0 & 0 & 0 & 0 & 0 & 0 & 95 \\ 0 & 0 & 0 & 0 & 0 & 0 & 124 \end{pmatrix} \quad Z_2 = \begin{pmatrix} 0 & 0 & 0 & 2,16 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1,56 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1,14 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1,63 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1,63 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1,17 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2,17 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1,55 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1,17 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0,90 \end{pmatrix}$$

El vector y contiene los 10 controles lecheros hechos por las 4 vacas:

$$y' = (15 \ 20 \ 17 \ 23 \ 25 \ 24 \ 16 \ 18 \ 22 \ 20)$$

Las ecuaciones de un modelo mixto para obtener soluciones de los vectores desconocidos (Henderson, 1950; 1984) son:

$$\begin{pmatrix} X'X & X'Z_a & X'Z_1 & X'Z_2 \\ Z_a'X & Z_a'Z_a + A^{-1}k_{aa} & Z_a'Z_1 + A^{-1}k_{a1} & Z_a'Z_2 + A^{-1}k_{a2} \\ Z_1'X & Z_1'Z_a + A^{-1}k_{1a} & Z_1'Z_1 + A^{-1}k_{11} & Z_1'Z_2 + A^{-1}k_{12} \\ Z_2'X & Z_2'Z_a + A^{-1}k_{2a} & Z_2'Z_1 + A^{-1}k_{21} & Z_2'Z_2 + A^{-1}k_{22} \end{pmatrix} \begin{pmatrix} b \\ a \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} X'y \\ Z_a'y \\ Z_1'y \\ Z_2'y \end{pmatrix}$$

donde:

A^{-1} = es la inversa de la matriz de parentesco genético aditivo.

k_{ij} = es el ij -ésimo elemento de la matriz G invertida multiplicado por σ_e^2 .

La dimensión (número de filas y columnas) de la matriz del lado izquierdo será de 26, esto es 7 animales cada uno con 3 regresores aleatorios (21), una columna para x_1 y para x_2 , dos columnas para el efecto rebaño y una columna para el intercepto general. De la misma forma el vector de elementos desconocidos tendrá 26 elementos que corresponden a las estimaciones de los efectos indicados anteriormente.

Asumiendo que existen estimaciones de los parámetros del modelo:

$\sigma_e^2 = 2$ (varianza residual) y que la estructura de la matriz G es:

$$G = \begin{bmatrix} 0,22395 & -0,00066 & 0,00175 \\ -0,00066 & 0,000366 & -0,000052 \\ 0,00175 & -0,000052 & 0,005338 \end{bmatrix}$$

entonces:

$$G^{-1} \sigma_e^2 = \begin{bmatrix} 9 & 16 & -2,8 \\ 16 & 5500 & 48 \\ -2,8 & 48 & 376 \end{bmatrix}$$

Las soluciones de los vectores desconocidos se obtienen al invertir la matriz del lado izquierdo multiplicándola por el vector del lado derecho:

$$\begin{pmatrix} \hat{b} \\ \hat{a} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{pmatrix} = \begin{bmatrix} X'X & X'Z_a & X'Z_1 & X'Z_2 \\ Z'_a X & Z'_a Z_a + A^{-1}k_{aa} & Z'_a Z_1 + A^{-1}k_{a1} & Z'_a Z_2 + A^{-1}k_{a2} \\ Z'_1 X & Z'_1 Z_a + A^{-1}k_{1a} & Z'_1 Z_1 + A^{-1}k_{11} & Z'_1 Z_2 + A^{-1}k_{12} \\ Z'_2 X & Z'_2 Z_a + A^{-1}k_{2a} & Z'_2 Z_1 + A^{-1}k_{21} & Z'_2 Z_2 + A^{-1}k_{22} \end{bmatrix}^{-1} \begin{pmatrix} X'y \\ Z'_a y \\ Z'_1 y \\ Z'_2 y \end{pmatrix}$$

Las soluciones para los efectos fijos y los efectos genéticos aleatorios incluidos en el modelo se presentan en los Cuadros 3 y 4, respectivamente.

Cuadro 3. Soluciones de los efectos fijos del modelo

Table 3. Solutions for the fixed effects included in the model

Efecto	Solución
Media general	47,16
Rebaño 1	23,975
Rebaño 2	23,18
Coef. de regresión para días en leche (b_1)	-0,2686
Coef. de regresión Ln (305/días en leche) (b_2)	-20,848

Cuadro 4. Soluciones de los efectos genéticos aleatorios

Table 4. Solutions for the random genetic effects

Animal	a	α_1	α_2
1	-0,123276	-0,022471	-0,003328
2	0,1074068	0,0124298	0,0041555
3	0,2683191	0,0326985	0,0100812
4	-0,049128	0,0023101	-0,003298
5	-0,061638	-0,011236	-0,001664
6	0,1918302	0,0250745	0,0069114
7	-0,068554	-0,002603	-0,003584

a = es un vector desconocido que contiene el intercepto del efecto genético aditivo para cada animal.

α_1 y α_2 = son dos vectores desconocidos que contienen los regresores aleatorios, para cada animal, de producción del día de control sobre días de lactancia y días de lactancia al cuadrado respectivamente.

RESULTADOS Y DISCUSIÓN

En un modelo animal los animales sin registros productivos (5, 6 y 7) también tienen soluciones y es posible calcular su valor genético, esto es posible al incluir la matriz de parentesco genético aditivo (A). Cada animal tiene tres soluciones (a , α_1 y α_2) con lo que es posible calcular el valor genético aditivo de cada uno de ellos para cada día de la lactancia si esto fuera requerido.

Si es de interés obtener el valor genético (V. G.) del j-ésimo animal para el i-ésimo día, éste se calcula de la siguiente manera: $V.G. = k'_i \lambda_j$

donde:

$$k'_i = (1 \quad i \quad \ln(305/i))$$

$$\lambda_j = (a \quad \alpha_1 \quad \alpha_2)_j$$

Por ejemplo, la obtención del valor genético del animal 1 en el día 60 de lactancia se hace de la siguiente manera:

$$V.G._{(1)(60)} = (1 \quad 60 \quad \ln(305/60)) \begin{bmatrix} -0,123 \\ -0,022 \\ -0,003 \end{bmatrix} = -1,47 \text{ kg}$$

Esto significa que la vaca número 1 tiene un valor genético negativo de casi 1,5 kg de leche bajo el promedio poblacional en el día 60 de lactancia. Una producción completa se puede obtener sumando la producción desde el día 1 al 305, de esta manera:

$$k_{305} = \sum_{i=1}^{305} k_i = (305 \quad 46665 \quad 301,220)$$

y el valor genético para el j-ésimo animal en toda la lactancia es:

$$V.G._{(305)} = k'_{305} \lambda_j$$

El valor genético de la vaca número 1 para una lactancia de 305 días es:

$$V.G._{(1)(305)} = (305 \quad 46665 \quad 301,220) \begin{bmatrix} -0,123 \\ -0,022 \\ -0,003 \end{bmatrix} = -1.087,211$$

Esto significa que en una lactancia de 305 días, el animal número 1 tiene un potencial genético de producir 1087,211 kg de leche, por debajo de la media poblacional.

De la misma forma, el valor genético estimado para una lactancia de 305 días del animal 4 es:

$$V.G._{(4)(305)} = (305 \quad 46665 \quad 301,220) \begin{bmatrix} -0,049 \\ 0,022 \\ -0,003 \end{bmatrix} = 91,823 \text{ kg}$$

Siguiendo la misma metodología, los valores genéticos estimados para una lactancia de 305 días de todos los animales del ejemplo se presentan en el Cuadro 5.

$$\% \text{ Persistencia} = \frac{\text{Producción día 280}}{\text{Producción día 60}} \times 100$$

Cuadro 5. Valores genéticos y diferencias predichas para una lactancia de 305 días
Table 5. Breeding values and predicted differences for a 305 days lactation

Animal	Valor genético (kg)	Diferencia predicha (kg)
1	-1087,211	-543,600
2	614,047	307,024
3	1610,745	805,373
4	91,823	45,912
5	-543,629	-271,815
6	1230,692	615,346
7	-143,458	-71,729

Persistencia se definió como el potencial productivo del día 280 de lactancia en relación a la producción del día 60.

Con los resultados del ejemplo y para evitar extrapolar los resultados se usará la producción del día 124 en relación al día 60. La producción del día 60 del animal 1 se puede estimar usando las soluciones fijas y aleatorias del modelo (Cuadros 3 y 4) de la siguiente forma:

$$\text{Prod.}_{(1)(60)} = (1 \quad 1 \quad 0 \quad 60 \quad 1,62 \quad 1 \quad 60 \quad 1,62) \begin{pmatrix} 47,160 \\ 23,970 \\ 23,180 \\ -0,268 \\ -20,840 \\ -0,123 \\ -0,022 \\ -0,003 \end{pmatrix} = 19,64 \text{ kg}$$

De la misma forma la producción del día 124 del animal 1 es:

$$\text{Prod.}_{(1)(124)} = (1 \quad 1 \quad 0 \quad 124 \quad 0,9 \quad 1 \quad 124 \quad 0,9) \begin{pmatrix} 47,160 \\ 23,970 \\ 23,180 \\ -0,268 \\ -20,840 \\ -0,123 \\ -0,022 \\ -0,003 \end{pmatrix} = 16,15 \text{ kg}$$

De esta manera la persistencia de producción para el animal 1 es:

$$\% \text{ Persistencia} = \frac{16,15}{19,64} \times 100 = 82,23\%$$

Esto significa que el día 124 de lactancia este animal puede producir un 82,23% de lo que es capaz de producir el día 60 de la lactancia.

Haciendo el mismo ejercicio para el animal 4 y asumiendo que este producirá en el rebaño 1, resulta en una persistencia de 90,60% ($18,50/20,40 = 90,60\%$). Esto indica que el animal 4 tiene una declinación de producción mucho menos pronunciada que el animal 1.

En este ejemplo el animal 4 tiene mejor persistencia y mejor valor genético que el animal 1. Al comparar dos animales puede existir el caso que

uno tenga mayor valor genético para una lactancia de 305 días pero una menor persistencia que otro con menor valor genético. Algunos productores afirman que un animal con mayor persistencia en la primera lactancia es aquel que no alcanza volúmenes espectaculares en el pico de producción, pero mantiene una producción pareja a través de la lactancia. Una vaca más persistente está sometida a un menor estrés y presenta una mejor condición corporal que aquella que produce demasiado al momento del pico de producción. Este tipo de animal puede alcanzar un mayor número de lactancias, por lo tanto su producción vitalicia puede ser mayor. Esta es una de las razones por la cual genetistas canadienses y españoles (Rekaya *et al.*, 1998) han empezado a estudiar esta característica que, aunque su relación con producción no es del todo clara, puede tener un gran potencial en el futuro para aumentar la longevidad productiva del rebaño. En los catálogos de sementales canadienses la persistencia de un toro se expresa como el promedio de persistencia de sus hijas en un rebaño promedio. Para la raza Holstein se considera una persistencia de 63% como el promedio racial (CDN, 1999).

La persistencia también es importante en aquellas lactancias que por algún motivo se extienden a

más de 305 días. La Figura 1 muestra las curvas de producción de dos vacas con lactancias de más de 305 días, donde la vaca A tiene una mayor persistencia que la vaca B. El sistema tradicional de evaluación de leche no considera producción de leche más allá de los 305 días, de esta manera la vaca A puede ser subevaluada en consecuencia que es un animal de mayor mérito genético que el animal B.

Las siguientes son algunas de las ventajas atribuibles a modelos de día de control con y sin regresores aleatorios:

- No se necesita una proyección de las lactancias que no han terminado. De esta manera disminuye la posibilidad de una estimación errada de la producción real de leche, grasa o proteína.
- Los registros de vacas que aún no han terminado su lactancia pueden ser analizados, de esta manera se obtienen estimaciones genéticas de toros jóvenes en un menor tiempo.
- Una mayor seguridad de estimación del valor genético de un animal (Jamrozik y Schaeffer, 1997; Swalve, 1998). La seguridad de estimación, la cual es la correlación entre el valor

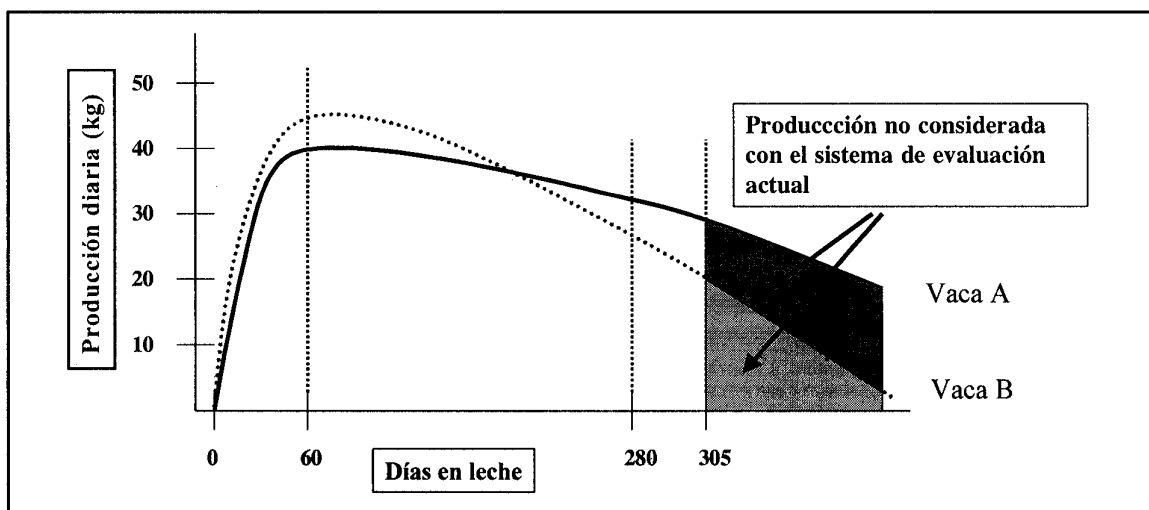


Figura 1. Curvas de producción de dos vacas con lactancias de más de 305 días.

Figure 1. Lactation curve of two cows with more than 305 days in lactation.

genético estimado y el verdadero, depende, entre otras cosas, del número de registros por animal. En contraste al análisis donde se usa un registro por lactación por vaca, un modelo para producción del día de control usa todos los controles (registros) mensuales disponibles de una vaca, esto aumenta la seguridad de estimación. La seguridad de estimación de los valores genéticos también está relacionada a la heredabilidad de la característica. Algunas investigaciones reportan una mayor heredabilidad de producción de leche al usar modelos del día de control con regresores aleatorios (Jamrozik y Schaeffer, 1997) lo que indica claramente una ventaja con respecto a seguridad de estimación (Swalve, 1998). Al aumentar la seguridad de estimación aumenta el cambio o progreso genético.

- Al existir más de una entidad de control lechero es más fácil uniformar las diferencias que existan entre ellas como diferencias en los intervalos de control, horario de control lechero, método de extensión y estandarización de la lactancia, etc.

En el caso de un modelo de día de control con regresores aleatorios existe la ventaja adicional que cada animal tenga su propia curva de lactancia (altura y forma), dentro de lo permisible por la función usada para describir la curva en un modelo particular. Además, al ser posible calcular valores genéticos en cualquier etapa de la lactancia es posible tener estimaciones de persistencia.

Un modelo de día de control requiere mayores recursos de computación. El número de ecuaciones a resolver aumenta a medida que la función que explica la curva de lactancia aumenta sus parámetros. Las funciones usadas para modelar la curva de producción de leche en Canadá tienen 5 parámetros (Ali y Schaeffer, 1987; Jamrozik y Schaeffer, 1997), esto indica que la matriz G tiene 25 elementos. Por simplicidad, en el ejemplo presentado se consideró una función con 3 parámetros, la cual puede no reflejar lo

que es una curva de lactancia. La estimación de los elementos de la matriz G también representa un esfuerzo considerable de programación y optimización de tiempo computacional. La obtención de estos parámetros ha sido posible usando el teorema de Bayes (Box y Tiao, 1992). En estadística bayesiana la distribución posterior de una variable aleatoria está dada por una función de densidad *a priori*, la cual se actualiza con la información entregada por los datos. La distribución posterior conjunta de los parámetros contiene toda la información necesaria para hacer la inferencia (Sorensen *et al.*, 1994).

Sería posible obtener la distribución posterior marginal de un determinado parámetro desde la distribución conjunta por integración de los parámetros contenidos en esa distribución. La integración analítica de una distribución posterior con varios parámetros es casi imposible de obtener (Wang *et al.*, 1994), lo cual impedía la implementación de análisis bayesianos cuando la distribución posterior conjunta presentaba algún grado de complejidad, éste es el caso de un modelo de día de control donde hay varios parámetros a estimar. Existen algunas aproximaciones como es el muestreo de Gibbs (Casella y George, 1992; Wang *et al.*, 1994). El muestreo de Gibbs, el cual es una integración numérica, fue presentado por Geman y Geman (1984), éste genera vectores aleatorios de la distribución posterior marginal, de los parámetros de interés al muestrear, de la distribución posterior conjunta.

Parámetros genéticos como varianzas y heredabilidad también pueden ser estimados para cada día de la lactancia. Efectos ambientales permanentes (aquellos que afectan a todos los registros de producción pero no son de origen genético) también se modelan con regresores aleatorios. La varianza residual se asumió como constante en el ejemplo, en la realidad ésta varía a través de la lactancia y debe modelarse de acuerdo a esta característica.

De todos estos análisis la información que llega al productor de leche es lo que se publica en un

catálogo de comercialización de semen congelado, esto es la mitad del valor genético estimado ($V.G_{.305}$) de una lactancia completa (Cuadro 5), lo que se conoce como Diferencia Predicha. El valor genético se divide por dos, ya que un determinado reproductor sólo transmite a su progenie una muestra aleatoria de la mitad de sus genes, la otra mitad que formará al nuevo individuo se asume que viene también en forma aleatoria de la población.

CONCLUSIONES

Al aumentar la seguridad de predicción, los modelos para producción en el día de control permiten una mejor aproximación al verdadero valor genético del animal. Aunque la comparación directa con modelos tradicionales no es directa,

ya que estos modelan características que pueden ser diferentes.

La demanda computacional de este tipo de modelos es considerablemente mayor que un modelo animal univariado.

La persistencia de producción de leche es un subproducto de esta metodología de evaluación, la cual puede tener bastante importancia en el futuro cuando exista mayor conocimiento de la relación entre persistencia y producción vitalicia.

Desde su desarrollo teórico los modelos de día de control han sido implementados comercialmente en ganado lechero en Canadá (incluyendo regresores aleatorios) y en Alemania.

LITERATURA CITADA

- Ali, T.E., and L.R. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67:637-644.
- Box, G.E.P., and G.C. Tiao. 1992. *Bayesian Inference in Statistical Analysis*. 608 p. John Wiley & Sons, New York, USA.
- CDN. 1999. Canadian Dairy Network. Interpretation and Use of Canadian Bull Proof for Lactation Persistency. Available on: <http://www.cdn.ca/pages/news/persistency.html>. Accessed on February 1999.
- Casella, G., and E.I. George. 1992. Explaining the Gibbs Sampler. *Am. Stat.* 46:167-174.
- Geman, S., and D. Geman. 1994. Stochastic Relaxation, Gibbs Distributions, and the Bayesian restoration of images. *IEEE Transactions of Pattern Analysis and Machine Intelligence* 6:721-741.
- Henderson, C.R. 1984. *Applications of Linear Models in Animal Breeding*. 423 p. University of Guelph, Guelph, Ontario, Canada.
- Henderson, C.R. 1950. Estimation of genetic parameters. *Ann. Math. Stat.* 21:309. (Abstr.).
- Jamrozik, J., and L.R. Schaeffer. 1997. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. *J. Dairy Sci.* 80:762-770.
- Jamrozik, J., G.J. Kistemaker, J.C.M. Dekkers, and L.R. Schaeffer. 1997. Comparison of possible covariates for use in a random regression model for analyses of test day yields. *J. Dairy Sci.* 80:2550-2556.
- Jara, A.A. 1999. Impacto de la interacción genético ambiental sobre producción y calidad láctea en una población multiracial de vacas lecheras. 226 p. Tesis Médico Veterinario. Universidad de Chile, Facultad de Ciencias Veterinarias y Pecuarias, Santiago, Chile.
- Misztal, I., T.J. Lawlor, T.H. Short, and P.M. Van Raden. 1992. Multiple-trait estimation of variance components of yield and type traits using an animal model. *J. Dairy Sci.* 75:544-551.

- Ptak, E., and L.R. Schaeffer. 1993. Use of test day yields for genetic evaluation of dairy sires and cows. *Livest. Prod. Sci.* 34:23-34.
- Reents, R., J.C. M. Dekkers, and L.R. Schaeffer. 1994. Analysis of test day for somatic cell from dairy cows. *J. Anim. Sci.* 72:267 (Abst.).
- Rekaya, R., M.J. Carabaño, and M.A. Toro. 1998. Assessment of heterogeneity of residual variances using changepoint techniques. 49th Annual meeting of European Association of Animal Production, Warsaw, Poland. 24-27 August. Session 6.4. Free Communications/Ideas Box.
- Schaeffer, L.R., and J.C.M. Dekkers. 1994. Random regressions in animal models for test-day production in dairy cattle. Vol. XVIII. p. 443-446. *In Proceedings 5th World Congress Genet. Appl. Livest. Prod.*, University of Guelph, Guelph, Ontario, Canada.
- Schaeffer, L.R. 1997. Random Regression Models. p. 1-15. *In* Van der Werf, J., and L.R. Schaeffer (ed.). *Random Regression in Animal Breeding. Course Notes.* University of Guelph, Department of Animal and Poultry Science, Guelph, Ontario, Canada.
- Sorensen, D.A., C.S. Wang, J. Jensen, and D. Gianola. 1994. Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genet. Sel. Evol.* 26:333-360.
- Swalve, H.H. 1998. Use of test day records for genetic evaluation. Vol. 23:295-302. *In Proceedings 6th World Congress Genet. Appl. Livest. Prod.* 11-16 January. Armidale, Australia.
- Uribe, H.A. 1997. Estimation of genetic merit for conformation traits using random regression in Holstein. 221 p. Doctor of Philosophy Thesis. University of Guelph, Department of Animal and Poultry Science, Guelph, Ontario, Canada.
- Uribe, H.A. 1998. Uso de regresiones aleatorias en predicción de mérito genético en características de producción de leche. p. 207-208. *In XXIII Reunión Anual Sociedad Chilena de Producción Animal A.G.*, 21-23 octubre. Chillán, Chile.
- Wang, C.S., J.J. Rutledge, and D. Gianola. 1994. Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. *Genet. Sel. Evol.* 26:91-115.
- Wiggans, G.R. 1997. Genetic evaluation systems in the United States. *In 32nd International Symposium, Animal Production: Advances in Technology, Accuracy and Management*, September 29 – October 1. Milan, Italy. Available on: <http://www.aipl.arsusda.gov/memos/html/overview.html>. Accessed: September 1998.
- Wood, P.D.P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164-165.