

ANÁLISIS GENÉTICO CONJUNTO DE LA LONGEVIDAD FUNCIONAL EN TIEMPOS DISCRETOS Y LA PRODUCCIÓN LECHERA. DESARROLLO METODOLÓGICO.

O. González-Recio y R. Alenda

Departamento de Producción Animal. ETSI Agrónomos. Universidad Politécnica de Madrid. Ciudad Universitaria s/n. 28040 Madrid.

oscar.grecio@upm.es

INTRODUCCIÓN

La importancia de la selección por longevidad ha aumentado considerablemente en los últimos años en la mayoría de las especies domésticas de interés ganadero [1, 3, 10]. Hasta el momento, el análisis de supervivencia ha sido la herramienta que mayor aceptación ha tenido a la hora de evaluar la longevidad. Este método presenta numerosas ventajas frente a los métodos tradicionales, ya que permite un tratamiento adecuado de datos censurados, posibilita incluir efectos dependientes del tiempo, y no requiere que los caracteres analizados presenten distribuciones normales.

Este creciente interés de la supervivencia en las especies domésticas ha generado la necesidad de estimar las correlaciones genéticas con otros caracteres de interés productivo, como la producción en vacuno de leche. Recientemente, varios autores han propuesto diferentes métodos para el análisis conjunto de un carácter de supervivencia y otro gaussiano. Tarrés et al. [11] justificaron por simulación una aproximación en dos pasos a través de unos pseudo-valores de la longevidad, que son equivalentes a los datos ajustados por todos los efectos no genéticos en los modelos lineales [5]. Un procedimiento más exacto fue propuesto por Nguti et al. [9] en el que se describe un análisis conjunto de un carácter de supervivencia y otro descrito a través de un modelo lineal longitudinal. Damgaard y Korsgaard [4] propusieron aplicar un análisis bayesiano conjunto de la supervivencia con un carácter gaussiano utilizando métodos MCMC (Markov chain Monte Carlo). Sin embargo, su implementación no es sencilla y el tiempo de computación es aún prohibitivo.

Los caracteres de supervivencia pueden medirse en tiempos discretos (p.e. número de lactaciones en ovino o vacuno de leche, número de partos en porcino o vacuno de carne, etc) y analizarse con metodologías de análisis de supervivencia adaptadas, como el modelo umbral secuencial [1, 6]. Este método describe los procesos fisiológicos y de decisión que ocurren de forma secuencial y permite incorporar efectos dependientes del tiempo y datos censurados. Además se ajusta perfectamente a la mayoría de las especies domésticas donde el desecho suele hacerse al final de cada ciclo productivo, aunque esto no es una condición necesaria para el modelo.

El objetivo de este trabajo fue implementar el método umbral secuencial en un análisis bivariado de la supervivencia en tiempos discretos con un carácter gaussiano, aplicado al número de lactaciones y producción lechera en la población Holstein del País Vasco y Navarra, sin recurrir a una segunda aproximación.

MATERIAL Y MÉTODOS

Para la realización de este estudio se utilizaron datos de supervivencia y producción correspondientes a 96.642 lactaciones de 41.170 vacas procedentes de las comunidades autónomas del País Vasco y Navarra.

Se observó la supervivencia durante las 5 primeras lactaciones junto con sus respectivas producciones estandarizadas a 305 días (KGL305). El método aplicado se basa en que la lactación l se alcanza únicamente tras haber observado $l - 1$ lactaciones previas. La supervivencia en cada lactación (SUPV) se midió como una variable dicotómica indicando si la vaca fue desechada (0) o sobrevivió a la siguiente lactación (1). Dicha supervivencia se relacionó con sus respectivas KGL305. El conjunto de datos observados queda definido por \mathbf{y} , donde $\mathbf{y}_i = (y_{i1}, y_{i2})$ son las observaciones para SUPV (y_{i1}) y KGL305 (y_{i2}) correspondientes al

animal i . Según el modelo umbral secuencial [1], la probabilidad de sobrevivir hasta la lactación l , condicionada a que se haya alcanzado la l -ésima lactación viene dada por:

$$\Pr(y_{i1} = l \mid y_{i1} \geq l-1, \boldsymbol{\gamma}, \mathbf{X}'_{i1} \boldsymbol{\theta}_{i1}) = \Phi[\boldsymbol{\gamma}_i - \mathbf{X}'_{i1} \boldsymbol{\theta}_{i1}],$$

donde $\boldsymbol{\gamma}$ es el vector de 'cutpoints' y $\mathbf{X}'_{i1} \boldsymbol{\theta}_{i1}$ representa los efectos incluidos en el modelo para SUPV. Este modelo puede formularse en términos de una variable subyacente que indica la probabilidad de sobrevivir a la siguiente lactación que permite una implementación y programación sencilla [2]. De esta manera, a cada observación para SUPV ($y_{i1} = L$) le corresponden L variables subyacentes $\{\omega_{il}\}$, observando:

$$y_i = \begin{cases} 1 & \text{if } \omega_{i1} \leq \gamma_1 \\ 2 & \text{if } \omega_{i1} > \gamma_1, \omega_{i2} \leq \gamma_2 \\ 3 & \text{if } \omega_{i1} > \gamma_1, \omega_{i2} > \gamma_2, \omega_{i3} \leq \gamma_3 \\ 4 & \text{if } \omega_{i1} > \gamma_1, \omega_{i2} > \gamma_2, \omega_{i3} > \gamma_3, \omega_{i4} \leq \gamma_4 \\ 5 & \text{if } \omega_{i1} > \gamma_1, \omega_{i2} > \gamma_2, \omega_{i3} > \gamma_3, \omega_{i4} > \gamma_4, \omega_{i5} \leq \gamma_5 \end{cases}$$

Para cada observación $\mathbf{y}_i = (y_{i1}, y_{i2})$ definimos \mathbf{W}_i , siendo $w_i = (\omega_{il}, y_{2l})$ la pareja de datos para el animal i en su lactación l , que indican la probabilidad de sobrevivir a la siguiente lactación (ω_{il}) y la producción de leche estandarizada a 305 días en la presente lactación (y_{2l}).

El modelo propuesto sería $\mathbf{W} \mid \boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{p}, \mathbf{a}, \mathbf{R} \sim N(\boldsymbol{\gamma} + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_p \mathbf{p} + \mathbf{Z}_a \mathbf{a}, \mathbf{R} \otimes \mathbf{I})$, siendo $\boldsymbol{\beta}$ los efectos sistemáticos de número de lactación (para KGL305), nivel de producción (para SUPV), mes de parto, año-estación de parto y rebaño, \mathbf{p} el efecto permanente del animal para KGL305 con varianza σ_p^2 y \mathbf{a} el efecto aditivo genético con distribución $\mathbf{a} \mid \mathbf{G} \sim N(\mathbf{0}, \mathbf{G} \otimes \mathbf{A})$. La matriz de

covarianzas residuales viene dada por $\mathbf{R} = \begin{bmatrix} 1 & \sigma_{e12} \\ \sigma_{e21} & \sigma_{e2}^2 \end{bmatrix}$. Para $\boldsymbol{\theta} = (\boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{p}, \mathbf{a}, \sigma_p^2, \mathbf{G}, \mathbf{R})$, la

distribución posterior quedaría descrita por:

$$p(\boldsymbol{\theta}, \mathbf{W} \mid \mathbf{y}, \boldsymbol{\delta}) \propto p(\mathbf{y}, \boldsymbol{\delta}, \mathbf{W} \mid \boldsymbol{\theta}) p(\mathbf{p} \mid \sigma_p^2) p(\sigma_p^2) p(\mathbf{a} \mid \mathbf{G}) p(\mathbf{G}) p(\mathbf{e} \mid \mathbf{R}) p(\mathbf{R}),$$

donde

$$p(\mathbf{y}, \boldsymbol{\delta}, \mathbf{W} \mid \boldsymbol{\theta}) \propto \prod_{i=1}^N |\mathbf{R}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{W}_i - \mathbf{E}(\mathbf{W}_i))' \mathbf{R}^{-1} (\mathbf{W}_i - \mathbf{E}(\mathbf{W}_i)) \right\}.$$

Las distribuciones posteriores de los parámetros de interés se estimaron usando un algoritmo Gibbs/Metropolis. El análisis consistió en una única cadena de 100.000 iteraciones, descartando las 10.000 primeras como periodo de quemado.

RESULTADOS Y DISCUSIÓN

Las medias posteriores de la heredabilidad para SUPV y KGL305 fueron 0,11 y 0,27 con una desviación posterior estándar de 0,01 para ambas estimas (Tabla 1). La media posterior para la correlación genética fue de -0,53, con un intervalo de confianza al 95% entre -0,62 y -0,46. Este valor indica que las vacas de mayor producción tienen una vida productiva más corta, posiblemente porque presentan mayores problemas metabólicos, de fertilidad y menor resistencia a enfermedades que provocan un mayor desecho involuntario.

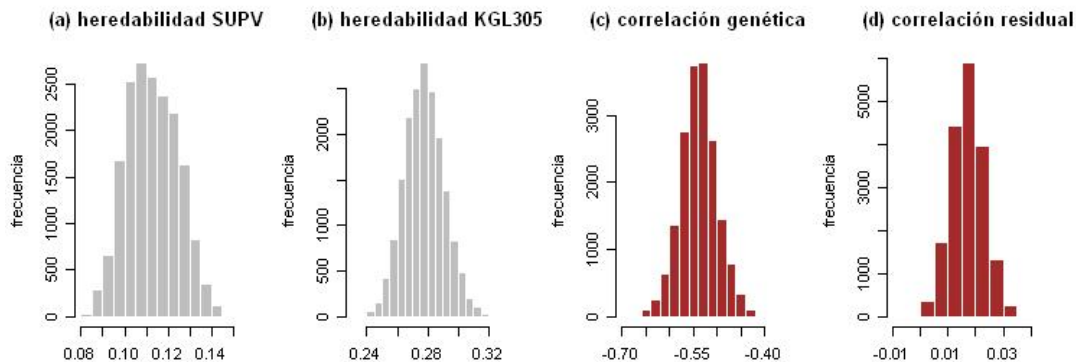
Es necesario recordar que se ha evaluado la longevidad funcional, eliminando el efecto del desecho voluntario del ganadero debido a niveles bajos de producción. Estos resultados son acordes a los encontrados por otros autores [3, 8]. Los histogramas de las distribuciones

posteriores para las heredabilidades y la correlación genética presentaron un aspecto gaussiano (Figura 1). Este método parece tener un comportamiento adecuado, sin problemas de estimación, a la vista de las cadenas obtenidas (resultados no mostrados), aunque podría validarse esta hipótesis con un estudio de simulación. La aplicación del modelo umbral secuencial a la supervivencia medida en tiempos discretos permite realizar un análisis conjunto con otro carácter gaussiano, sin necesidad de segundas aproximaciones. Este método posibilita la estimación de componentes genéticas con otros caracteres de interés productivo, que puedan servir como predictores tempranos de la longevidad en las especies de interés.

Tabla 1: Medias posteriores, desviación estándar (d.e.) y error de Monte-Carlo para las heredabilidades de la supervivencia (SUPV), la producción (kg) de leche estandarizada a 305-días (KGL305), y sus correlaciones genética (ρ_g) y residual (ρ_e).

	SUPV	KGL305	ρ_g	ρ_e
Media	0,11	0,27	-0,53	-0,02
d.e.	0,01	0,01	0,04	0,001
Error de Monte Carlo	0,001	0,001	0,004	0,001

Figura 1: Histogramas de las distribuciones posteriores para las heredabilidades de la supervivencia (SUPV) y la producción de leche a 305-días (KGL305), y de sus respectivas correlaciones genéticas y residuales.



REFERENCIAS BIBLIOGRÁFICAS

- [1] Albert, J.H. y S. Chib. 1997. *Biometrics* 57: 829-836.
- [2] Chang, Y.M., O. González-Recio, K.A. Weigel, D. Gianola. *Genet. Sel. Evol.* (enviado).
- [3] Caraviello, D.Z., K.A. Weigel, D. Gianola. 2004. *J. Dairy Sci.* 87: 3518-3525.
- [4] Damgaard, L.H. y I.R. Korsgaard. 2006. *Genet. Sel. Evol.* 38: 45-64.
- [5] Ducrocq, V. D. Boichard, A. Barbat, H. Larroque. 2001. *Proc. 51st EAAP Meeting*, 2 (abstr.).
- [6] González-Recio, O., Y.M. Chang, D. Gianola, K.A. Weigel. 2005. *J. Dairy Sci.* 88: 3655-3662.
- [7] Guo, S.F., D. Gianola, R. Rekaya and T. Short. 2001. *Livest. Prod. Sci.* 72: 243-252.
- [8] Kadarmideen, H.N., R. Thompson, M.P. Coffey, M.A. Kossaibati. 2003. *Livest. Prod. Sci.* 81: 183-195.
- [9] Nguti, R., T. Burzykowski, J. Rowlands, D. Renard, P. Janssen. 2005. *Genet. Sel. Evol.* 37: 175-197.
- [10] Sánchez, J.P., M. Baselga, R. Peiro, M.A. Silvestre. 2004. *Livest. Prod. Sci.* 90: 227-234.
- [11] Tarrés, J., J. Piedrafita, V. Ducrocq. 2006. *Genet. Sel. Evol.* 38: 65-83.