

ESTIMA DE COMPONENTES DE VARIANZA Y DE RESPUESTA A LA SELECCION EN LA RAZA OVINA MANCHEGA MEDIANTE MUESTREO DE GIBBS

Jurado, J.J.*; Serrano, M.*; Montoro, V.** y Pérez-Guzmán, M.D.**

* Area de Mejora Genética Animal del CIT-INIA de Madrid. España.

** CERSYRA de Valdepeñas. Ciudad Real. España.

INTRODUCCION

El éxito de un programa de mejora genética se puede cuantificar por el cambio del valor genético medio que se va produciendo a lo largo de los años en la población seleccionada. Dicho éxito puede ser atribuido al grado de organización del esquema de selección, a la calidad de la valoración genética de los reproductores y al adecuado uso que de ellos se haga. La valoración del mérito genético de machos y hembras depende en parte de la utilización de unos parámetros genéticos correctos. La respuesta a la selección está, asimismo, muy influida por los componentes de varianza (**CV**) que se utilicen para estimarla (SORENSEN, 1986). El procedimiento clásico es estimar los **CV** a partir de la información acumulada, utilizarlos para hacer una valoración genética y a partir de esta estimar la respuesta a la selección, lo cual no asegura si la respuesta estimada es distinta de cero o de otro valor arbitrario, ya que se desconoce la distribución y la varianza del estimador. También es posible utilizar una metodología Bayesiana que permite estimar los parámetros genéticos y la respuesta simultáneamente a partir de las distribuciones posteriores de estos. El objeto de este trabajo es presentar los primeros resultados de una aplicación Bayesiana mediante muestreo de GIBBS (**MG**) para estimar los **CV** y la tendencia genética en la raza ovina Manchega, y compararla con la metodología más clásica (**MC**).

MATERIAL Y METODOS

Se utilizó la base de datos acumulada en el esquema de selección de la raza ovina Manchega desde el año 1986 hasta el año 1992. Esta incluye 28099 lactaciones correspondientes a 18622 hembras, así como información acerca de 201 machos de inseminación artificial. Los caracteres estudiados fueron leche normalizada al 6% de grasa (**LN**), leche total (**LT**), porcentaje de grasa (**PG**) y porcentaje de proteína (**PP**), todos ellos estandarizados a 120 días de lactación.

El modelo utilizado fue el descrito por SERRANO et al. (1994):

$$y = \mu + \text{rem} + \text{tp} + \text{nped} + \text{int} + u + c + e$$

donde:

-y es el valor fenotípico del carácter; μ es la media general; **rem** es el efecto conjunto de rebaño-año-época móvil (731 niv.); **tp** es el tipo de parto (4 niv.); **nped** es el efecto conjunto número de parto-edad oveja (161 niv.); **int** es el intervalo parto-primer control (3 niv.: <44, 45-52, >52 días); **u** es el valor genético de la oveja ($n_u=18622$ niv.); **c** es el efecto permanente ($n_c=9477$ niv.) y **e** es el residuo ($n_e = 28099$).

En notación matricial este modelo sería:

$$y = X\beta + Z_1 u + Z_2 c + e$$

donde: X, Z₁ y Z₂ son las matrices de incidencia que relacionan β , u y c con los valores de y.

Se cumple que $u \sim N(0, A\sigma_u^2)$, $c \sim N(0, I_c\sigma_c^2)$ y $e \sim N(0, I_e\sigma_e^2)$, siendo σ_u^2 , σ_c^2 y σ_e^2 los **CV** y A la matriz de relaciones genéticas entre animales. u, c y e se suponen independientes. El número total de niveles de efectos fijos es de 899.

-Método clásico (MC).

Se calcularon las estimas de los **CV** mediante el método **REML**, via algoritmo libre de derivadas, utilizando los programas **DFREML** (MEYER, 1993) para cada uno de los caracteres (unicarácter). Estas estimas se utilizaron en las ecuaciones del modelo

mixto para obtener las predicciones **BLUP** de los genotipos de los animales. Por último se estimó la tendencia genética mediante el coeficiente de regresión de la media genética de los animales nacidos en un año sobre dicho año, para los años 90, 91 y 92.

-Análisis Bayesiano mediante muestreo de Gibbs (MG).

Se utilizó la metodología descrita por WANG et al., (1994) y SORENSEN et al. (1994). Las inferencias sobre las estimas de efectos fijos, predicciones genéticas y **CV** están basadas en sus distribuciones marginales "a posteriori", calculándose la marginalización mediante **MG**.

Para realizar las inferencias bayesianas se utilizaron dos tipos de "a priors" para los **CV**: PLANOS (**G1**), consistentes en distribuciones uniformes ($p(\sigma^2_i) \propto Cte; i=u,c,e$) e INFORMATIVOS (**G2**), que suponen distribuciones χ^2 invertidas (independientes para u,c y e) según la siguiente expresión:

$$p(\sigma^2_i / v_i, S^2_i) \propto (\sigma^2_i)^{-((v_i/2)+1)} \exp^{-v_i S^2_i / 2\sigma^2_i} \quad (i = u, c, e)$$

donde v_i es el grado de credibilidad de la información "a priori" y S^2_i es un valor "a priori" para la varianza (con $v_i = -2$ y $S^2_i = 0$, la distribución es uniforme). En el caso **G2** los valores de v_i fueron tales que: $v_i/(v_i + n_i) = 0.01$ (esto implica un peso relativo de la información "a priori" de solo un 1%), y los de S^2_i las estimas **DFREML** de los **CV**. Tanto para **G1** como para **G2** se utilizaron como "a priors" de β y de u distribuciones uniformes ($p(\beta), p(u) \propto Cte.$).

En ambos casos, **G1** y **G2**, se corrió una de cadena de una longitud de 200000 muestras para cada uno de los caracteres estudiados. Las primeras 3000 muestras fueron descartadas y a partir de estas, se tomó una de cada 10. El total de muestras utilizadas fue de 19700.

Para el proceso de cálculo se utilizaron programas propios escrito en FORTRAN77 ejecutados en una estación de trabajo SUN-10. Los números aleatorios fueron generados usando subrutinas descritas en NUMERICAL RECIPES (1990).

RESULTADOS Y DISCUSION

Las estimas **DFREML** de los **CV** presentaron valores significativamente diferentes según el valor inicial empleado en el proceso de búsqueda de los máximos de la función de verosimilitud, lo cual no permite considerarlas como definitivas. Fueron las estimas de σ^2_c (varianza del efecto permanente) las que mostraron una mayor variación y error típico. En la Tabla 1 se presenta, para cada uno de los caracteres estudiados, el valor medio de varias ejecuciones para la heredabilidad (h^2) y la repetibilidad (r).

En el contexto Bayesiano, la utilización de distintos tipos de información "a priori" nos condujo a inferencias sobre los **CV** distintas en cada caso. En la Tabla 1 se presentan los valores modales obtenidos para las distribuciones posteriores de h^2 y r en los casos **G1** y **G2**. Los distintos resultados obtenidos bajo **G1** y **G2** permiten concluir que los datos son poco informativos para la estima de los parámetros genéticos, ya que su valor depende del tipo de información "a priori" utilizada. La técnica Bayesiana nos permite profundizar en la razón de la baja calidad de las estimas **DFREML** de los **CV**.

Las estimas de la tendencia genética (b) y de la respuesta total (rs) (media año92-media año 90) calculadas mediante el **MC** y **MG** (**G1** y **G2**), se presentan en la Tabla 2. El **MC**, utilizando las estimas **DFREML**, da valores de b inferiores a los obtenidos mediante **MG**, siendo estos en el primer caso no significativamente distintos de cero y si en el segundo. La rs es la misma cuando se compara el **MC** con el caso **G1** y **G2** del **MG**. El **MG** presenta valores modales inferiores a los valores medios en **G2**, pero iguales en **G1**. Estos

resultados parecen mostrar, que aunque no hay valores definitivos para **b** y **rs**, si hay una cierta evidencia de su existencia.

TABLA 1: Estimaciones de heredabilidad (h^2) y repetibilidad (r) para los caracteres leche normalizada (LN), leche total (LT), porcentaje de grasa (PG) y porcentaje de proteína (PP) obtenidas mediante DFREML y Muestreo de Gibbs con distintos "a priori" (G1 y G2).

CARAC	DFREML		MUESTREO DE GIBBS			
	$h^2 \pm ET$	r	MODA G1		MODA G2	
			h^2	r	h^2	r
LN	0.29 ± .024	0.48	0.34	0.34	0.23	0.37
LT	0.33 ± .024	0.50	0.30	0.38	0.27	0.40
PG	0.23 ± .005	0.36	0.24	0.24	0.19	0.28
PP	0.35 ± .020	0.47	0.27	0.40	0.29	0.41

TABLA 2: Tendencia genética (**b**) y respuesta total (**rs**) para los caracteres leche normalizada (LN) y leche total (LT) obtenidas mediante la metodología clásica (MC) y Muestreo de Gibbs (casos G1 y G2).

Param	MC	MUESTREO DE GIBBS					
		MEDIA G1	MEDIA G2	MODAG1	MODAG2	ASIMG1	ASIMG2
b(LN)	0.842 ± 0.62	1.239 ± 0.59	1.008 ± 0.52	1.232	0.864	-0.005	-0.015
rs(LN)	2.37	2.479 ± 1.19	2.016 ± 1.03	2.464	1.728	-0.005	-0.015
b(LT)	0.805 ± 0.71	1.138 ± 0.46	1.076 ± 0.44	1.063	0.936	0.005	-0.016
rs(LT)	2.50	2.276 ± 0.93	2.153 ± 0.88	2.127	1.873	0.005	0.016

ASIM = Asimetría.

CONCLUSIONES

- Las diferencias de las estimas DFREML de los componentes de varianza según el valor inicial empleado, fué un primer indicio de que la función de verosimilitud no estaba bien definida, ya que se alcanzan máximos de la función diferentes en cada ejecución.
- El estudio Bayesiano revela que los datos son poco informativos en lo que respecta a la estima de los parámetros genéticos, lo cual se manifiesta en que el tipo de información "a priori" utilizada es determinante en el valor que toman las estimas de los componentes de varianza.
- Las estimas de la tendencia genética y de la respuesta total no son concluyentes, pero existen fuertes indicios de que pudieran tomar valores distintos de cero.
- El Muestreo de Gibbs es fácil de programar a partir de paquetes informáticos para predicción BLUP, pero es muy costoso en cuanto a tiempo de ejecución.

BIBIOGRAFIA

- MEYER, K.(1993). Programa DFREML versión 2.1.08.
- NUMERICAL RECIPES, (1990). Fortran version. Cambridge University Press.
- SERRANO, M. et al. (1994). 45th Annual Meeting of the European Association of Animal Production. Edimburg, 5-8 September, 1994.
- SORENSEN, D.(1986). J. Anim. Sci. 63: 245-258.
- SORENSEN, D. et al. (1994). Genet. Sel. Evol (in press).
- WANG, C.S. et al. (1994). Theor. and Appl. Genet. 88:220-230.