

ESTIMA DE PARÁMETROS GENÉTICOS PARA PROLIFICIDAD EN GANADO OVINO MEDIANTE MUESTREO DE GIBBS. RESULTADOS PRELIMINARES

M^a. J. Espinosa Pascual

J.J. Jurado García

Area de Mejora Genética y Biotecnología
CIT-INIA

Crta de La Coruña, Km. 7,00
28040 Madrid

España

RESUMEN

Se presentan los primeros estudios llevados a cabo sobre la información recogida en el esquema de selección de Carne-Aragón S.C.L. en raza Rasa-Aragonesa. El objetivo de selección es el incremento de la productividad numérica y el criterio de selección es la prolificidad de las ovejas por parto.

La prolificidad presenta distribución fenotípica discontinua, con un reducido número de clases y claramente asimétrica dudándose se pueda aplicar la metodología de modelos mixtos a estos caracteres. A los efectos de valoración genética se puede optar, bien por ignorar su naturaleza discontinua (asumiendo que los modelos mixtos son suficientemente robustos para soportar tales clases de anomalías), o bien por asumir modelos más complejos como los no lineales o el modelo umbral (con su doble escala visible y subyacente). En este último caso y dado que la escala subyacente es más heredable parecería lógico que los reproductores se eligieran según su valoración en esta escala. La forma propuesta por SORENSEN (1995) podría ser una forma de abordar este problema ya que obtiene valoraciones y parámetros genéticos en la escala subyacente, mediante metodología Bayesiana y muestreo de Gibbs.

El modelo utilizado para la valoración genética es el modelo animal con medidas repetidas e incluye los efectos **Rebaño-Año**, estación de parto, numero de parto, intervalo entre partos, modo de cubrición, valor genético y el efecto permanente. La base de datos incluye 21.605 ovejas con 71.130 partos.

Las estimas de la heredabilidad y repetibilidad en la escala visible fueron de $0,049 \pm 0,004$ y de $0,106 \pm 0,002$ y de $0,101 \pm 0,008$ y de $0,164 \pm 0,004$ en la escala subyacente. Estos valores confirman lo citado en la literatura acerca de la mayor heredabilidad del carácter en la escala subyacente.

Palabras Claves: Ovino, Prolificidad, Parámetros genéticos. Muestreo de Gibbs.

SUMMARY

GENETIC PARAMETER ESTIMATION FOR PROLIFICACY IN SHEEPS USING GIBBS SAMPLING: PRELIMINARY RESULTS

First studies based on the results obtained from Carne-Aragón S.C.L. selection scheme on Rasa-Aragonesa sheep are presented. Selection objective is to increase numeric productivity and the selection criterion was the ewes prolificacy.

Prolificacy presents a discontinuous and asymmetrical phenotypic distribution with a small number of classes which is not the ideal situation to apply mixed models methodology. Animal breeding values can be estimated ignoring discontinuous nature of this trait (trusting in mixed models method robustness) or using more complex models as non linear or threshold (with two scales: visible and subjacent) ones. Under the threshold model the heritability for prolificacy in the subjacent scale is higher than that in the visible scale. Therefore animals should be selected in the basis of their breeding values estimated in the subjacent scale. Sorensen (1995) proposed a Bayesian approach using Gibbs Sampling to estimate breeding values and genetic parameters in the subjacent scale under a threshold model.

In this case a threshold animal model with repeated measures including herd-year, lambing season, lambing number, period between lambings and covering mode as fixed effects was applied. Data consisted on 21605 ewes with 71130 lambings.

Heritability and repeatability estimates were 0.049 ± 0.004 and 0.106 ± 0.002 in the visible scale and 0.101 ± 0.008 and 0.164 ± 0.004 in the subjacent one, respectively. These values corroborate literature estimates in the sense of higher heritability values obtained in the subjacent scale.

Key words: Meat ewes, Prolificacy, Genetic parameters, Gibbs Sampling.

Introducción

La cooperativa Carne-Aragón S.C.L. está llevando a cabo un programa de selección genética en ganado ovino de raza Rasa Aragonesa en colaboración con la DGA de Aragón y el INIA de Madrid, cuyo objetivo de selección es incrementar la productividad numérica, utilizando como criterio de selección la prolificidad de las ovejas. Este programa se puso en marcha a finales de 1995.

La metodología BLUP es generalmente, aceptada como la forma más eficiente y versátil de obtener una clasificación correcta de los animales en función de su valor genético. En condiciones adecuadas las predicciones genéticas son insesgadas y de mínima varianza. No obstante hay considerables reservas de tipo teórico acerca de su

posible aplicación a caracteres de tipo discreto.

La prolificidad es un carácter que presenta distribución fenotípica discontinua, con un reducido número de clases y claramente asimétrica. Estas características ponen en cuestión que su análisis pueda llevarse a cabo mediante la técnica de los modelos lineales mixtos. No obstante dicha naturaleza discontinua, en ocasiones, se suele ignorar aduciendo que los modelos mixtos es una metodología robusta que soporta anomalías en la distribución de los caracteres y que permite una aceptable ordenación de los animales en función de su valor genético.

La definición del modelo umbral por WRIGHT (1934) y su posterior desarrollo por DEMSPETER y LERNER (1950), en el caso de un solo umbral, puso en evidencia que las heredabilidades en la escala subyacente

y en la escala visible pueden ser diferentes y de mayor magnitud en la primera que en la última, estando ambas relacionadas en función de la incidencia del carácter y del valor de la ordenada en el valor umbral. GIANOLA (1982) generaliza estos resultados a un mayor número de umbrales.

GIANOLA y FOULLEY (1983) y HARVILLE y MEE (1984) proponen métodos con modelos no lineales con el propósito de obtener predicciones genéticas, pero presentan grandes problemas computacionales cuando se pretende aplicar un modelo animal con medidas repetidas sobre los datos de una población de gran tamaño. FOULLEY y col. (1987) proponen la utilización de un modelo Poisson con estos mismos propósitos, no encontrando ventajas respecto a otros modelos.

Desde un punto de vista teórico, y puesto que la escala subyacente es más heredable, sería deseable que la elección de los reproductores se basara en los valores genéticos predichos sobre la escala subyacente más que en los predichos sobre la visible, lo que implica acceder de algún modo a dicha escala. SORENSEN y col. (1995) proponen un método para efectuar inferencias bayesianas en modelos umbrales mediante muestreo de Gibbs, basado en el muestreo de funciones de densidad truncadas que conducen a distribuciones posteriores de los parámetros genéticos y de las predicciones genéticas de los animales en la escala subyacente.

El propósito de este trabajo es presentar los primeros resultados de la aplicación de la metodología descrita por SORENSEN y col. (1995) mediante un modelo animal con medidas repetidas, al esquema de selección genética de Carne-Aragón, en especial a la obtención de estimas de heredabilidad y

repetibilidad de la prolificidad, obtenidas de formas diferentes y en las dos escalas.

Material y métodos

Los datos utilizados fueron parte de los acumulados en la base de datos del programa de selección y que se emplearon en la última valoración de reproductores en Abril de 1998. El número total de animales en dicha base de datos fue de 36.159 ovejas en 44 rebaños que sumaban un total de 107.921 partos. Hasta la fecha hay datos de 27 machos con 366 hijas controladas.

Los datos utilizados fueron los correspondientes a los rebaños conectados, entendiéndose por tales los que tenían al menos 8 hijas de machos de referencia. El número de ovejas en estas circunstancias fue de 21.605 con un total de 71.130 partos y 24 machos. El modelo utilizado fue el siguiente:

$$Y_{ijkl} = \mu + RA_i + Est_j + Np_k + Itp_l + Mc_m + U_n + Ep_n + \varepsilon_{p(ijklmn)}$$

En donde:

Y_{ijkl} es la prolificidad de la oveja en el parto.

μ es la media general de la población.

RA_i es el grupo de comparación rebaño-año de parto (108 niveles).

Est_j es el efecto de la estación de parto (12 niveles).

Np_k es el efecto del número de parto (10 niveles).

Itp_l es el efecto del intervalo parto y siguiente cubrición (6 niveles: ovejas de primer parto, <90 días, 90-160 días, >160,

anomalías: intervalo entre parto < 150 días, > 600 días).

Mc_m es el efecto del modo de cubrición (3 niveles, sincronización de celos sin IA, sincronización mas IA, monta natural).

U_n es el efecto genético del animal que produce el dato.

Ep_n es el efecto ambiental permanente

$\epsilon_{p(ijklmn)}$ es el residuo del modelo

Para estimar los parámetros genéticos se utilizaron diversos métodos con el fin de comparar resultados.

Sobre la escala visible se utilizó el programa VCE de GROENEVELD (1998) que implementa la metodología REML usando un método de gradiente analítico para maximizar la función de verosimilitud.

También se utilizó sobre la escala visible el muestreo de Gibbs según WANG (1994), realizándose un muestreo de los efectos fijos, efectos aleatorios y las varianzas genéticas, del efecto permanente y residual. Para los efectos fijos se asumieron "priors" uniformes de modo que todo los valores tuvieran la misma probabilidad y para los componentes de varianza se asumieron distribuciones χ^2 invertidas con bajos grados de credibilidad ($v_c=4$). Se limitó el campo de variación de las varianzas entre límites amplios.

Para trabajar sobre la escala subyacente se utilizó la metodología propuesta por SORENSEN (1995), mediante la técnica conocida como de "aumento de datos" (data augmentation) citada por el mismo autor. De acuerdo con el modelo umbral se asume la existencia de una escala subyacente sobre la que se definen los valores fenotípicos de las ovejas (y_s) correspondientes a los de la escala visible. Se asume asimismo un modelo lineal igual al usado en la escala

visible en el que la variable es y_s . Se asume también la existencia de unos umbrales (t_1 , t_2 , y t_3) que determinan unos intervalos en la distribución subyacente. En cada iteración el valor y_s de cada parto se obtiene mediante la suma de los correspondientes efectos fijos, genéticos y permanentes calculados en la iteración anterior (y_{sc}), mas el efecto del residuo (ϵ_s). Este último se deduce a partir de una distribución normal de media y_{sc} y varianza, la obtenida como residual en la iteración anterior. Si la oveja es de parto simple el valor ϵ_s se toma al azar de la distribución entre $-\infty$ y t_1 , si es de parto doble se toma entre t_1 y t_2 , si es de parto triple se toma entre t_2 y t_3 y si es cuádruple o más se toma entre t_3 y $+\infty$. Una vez muestreados los y_s , se muestrean los efectos fijos, los efectos aleatorios y las componentes de varianza de forma semejante a la citada anteriormente. La estructura de "priors" utilizada fue la misma que la usada para el caso de la escala visible.

Los umbrales elegidos fueron los deducidos a partir de las frecuencias fenotípicas en la escala visible. La aplicación del método conduce a obtener estimas de los parámetros genéticos y de los efectos fijos y predicciones de los valores genéticos de todos los animales como es propio del modelo animal. Puesto que los umbrales se asumen fijos, el proceso de muestreo conduce a que las varianzas tomen valores según los umbrales elegidos, pero los parámetros genéticos obtenidos a partir de las mismas son constantes con independencia de los umbrales fijados. Las cadenas de muestreo utilizadas tuvieron longitudes variables. Se descartaron las primeras iteraciones y a partir de una determinada se tomaron periódicamente algunas de ellas.

También se utilizó la técnica de cadenas pareadas (coupling method) aplicada por GARCIA-CORTÉS y col. (1997) al modelo

animal, con el fin de estimar el número de iteraciones que deberían ser descartadas previamente para conseguir buenas inferencias marginales. El periodo de calentamiento se consideró terminado cuando las diferencia entre los valores de los parámetros en ambas cadenas pareadas fue inferior a un valor prefijado.

Con el fin de estudiar la autocorrelación entre los puntos que integran la distribución marginal se calcularon las estimas de parámetros usando una única cadena de gran longitud con un periodo de calentamiento variable, así como también fue variable el intervalo entre las iteraciones elegidas. En todos estos casos se usan cadenas pareadas para vigilar la convergencia del proceso. La autocorrelación y el tamaño efectivo del conjunto de puntos utilizados para formar la distribución marginal se calculó por la técnica descrita por COX y MILLER (1965) y utilizada por SORENSEN (1995).

Resultados

En el cuadro 1 se presenta una síntesis de los resultados obtenidos hasta este momento, por lo que éstos deben ser considerados como preliminares.

Los valores obtenidos en la escala visible para la heredabilidad (h^2) de $0,059\pm 0,006$ y para la repetibilidad (r_{ep}) de $0,102\pm 0,006$ mediante REML están dentro del orden de valores obtenidos por otros autores para esta raza. OCCON (1988), obtiene valores de h^2 para el primer parto que oscilan entre 0,07 y 0,21. La r_{ep} oscila entre 0,11 y 0,16.

Cuando se utiliza el muestreo de Gibbs en la escala visible se usó una única cadena con 300.000 iteraciones, un periodo previo de 10.000 y se eligió una de cada 70 iteraciones. En total se obtuvieron 4.142 puntos que integran las distribuciones marginales. Los valores de los parámetros son muy se-

CUADRO 1
ESTIMAS DE LA HEREDABILIDAD Y REPETIBILIDAD OBTENIDAS POR REML Y MUESTREO DE GIBBS EN LAS ESCALA VISIBLE Y SUBYACENTE, ASI COMO AUTOCORRELACIÓN Y TAMAÑO EFECTIVO

Metod.	Escala	Longitud Cadena	Periodo Previo	Int.	Parámetros Genéticos	Autocorr. Tam. Efec.	
Reml	Visible				$h^2= 0,059\pm 0,006$ $r_{ep}= 0,102\pm 0,006$		
Gibbs	Visible	300.000	10.000	70	$h^2= 0,049\pm 0,004$ $r_{ep}=0,106\pm 0,002$ $n= 4.142$	0,77 0,106	491,5 2.890,2
Gibbs	Subyac	30.100	3.000	50	$h^2= 0,101\pm 0,008$ $rep= 0,164\pm 0,004$ $n= 542$	0,74 0,136	40,02 425,5
Gibbs	Subyac	50.000	6.000	200	$h^2=0,094\pm 0,008$ $r_{ep}= 0,168\pm 0,004$ $n= 220$	0,541 -0,015	52,13 190,7

mejantes a los obtenidos mediante REML. ($h^2 = 0,049 \pm 0,004$ y $r_{ep} = 0,106 \pm 0,002$). En este caso se observa en la h^2 una alta autocorrelación (0,77) lo que supone un tamaño efectivo de 491,5. En cambio en la r_{ep} esta autocorrelación es de 0,105 con un tamaño efectivo de 2.890,18, lo cual indicaría que los puntos que integran la distribución marginal de la r_{ep} son más independientes que los de la h^2 , debido seguramente a que la r_{ep} implica dos varianzas muestreadas frente a solo una de la h^2 .

En la aplicación del muestreo de Gibbs a la escala subyacente, las cadenas pareadas tuvieron 30.100 iteraciones con un periodo previo de 3.000 y se tomó una de cada 50 iteraciones. En cada una de las cadenas los valores iniciales de los parámetros genéticos fueron diferentes, pero no así los efectos fijos. Se observó que en la iteración 3.000, la diferencia máxima entre ambas cadenas para las soluciones del modelo mixto era del orden de 0,01. En el caso de las varianzas genéticas y permanentes dicha diferencia entre cadenas fue de 0,0004 lo que supone 1% de la diferencia inicial y en el caso de la varianza residual fue de 0,0001, un 0,005% de la inicial. Esto implica que el periodo de calentamiento fue suficiente para asegurar que las cadenas habían convergido y se estaba ya muestreando de la distribución conjunta (figuras 1 y 2). El valor obtenido de h^2 fue de $0,101 \pm 0,008$ y el de la r_{ep} de $0,164 \pm 0,004$, que son claramente superiores a los obtenidos trabajando sobre la escala visible. Este mayor valor de h^2 en la escala subyacente que en la visible era esperable según la teoría y demuestra la bondad del método propuesto por SORENSSEN (1985). Las predicciones genéticas obtenidas de esta forma tienen la ventaja de estar corregidas por los efectos ambientales según un modelo animal con medidas repetidas (como es propio de la metodología

BLUP) y, además, están basadas en unas medidas del carácter mucho más heredables como es la escala subyacente. El número de puntos que en este caso integran las distribuciones marginales fue de 542. La autocorrelación para h^2 fue de 0,747 lo que supone un tamaño efectivo de 40,02. Para la r_{ep} dicha autocorrelación fue de 0,136 con un tamaño efectivo de 425,5. Nuevamente se presenta la misma circunstancia que anteriormente, los puntos que integran la distribución marginal de la h^2 son menos independientes que los que integran la r_{ep} .

Con el propósito de estudiar si la alta autocorrelación afectaba de alguna forma a las estimas de los parámetros genéticos, se corrió otra cadena semejante a la anterior de 50.000 iteraciones, pero con un periodo previo más largo (6.000) y tomando una de cada 200 iteraciones. En total se obtuvieron 220 puntos. En este caso los valores alcanzados para h^2 ($0,094 \pm 0,008$) y r_{ep} ($0,168 \pm 0,004$) fueron del mismo orden que en el caso anterior con una autocorrelación mas baja (0,53) y un tamaño efectivo de 52,13 para la h^2 y de -0,051 y 190,69 para r_{ep} respectivamente. Esto podría ser un indicio de que la autocorrelación no afecta de forma notable a las estimas de los parámetros genéticos.

Conclusiones

Las conclusiones a las que podemos llegar en función de los resultados obtenidos serían las siguientes.

a) Los valores estimados de la heredabilidad y repetibilidad en la escala visible en esta población de ovino de raza Rasa-Aragonesa fueron de $0,049 \pm 0,004$ y $0,107 \pm 0,002$ respectivamente.

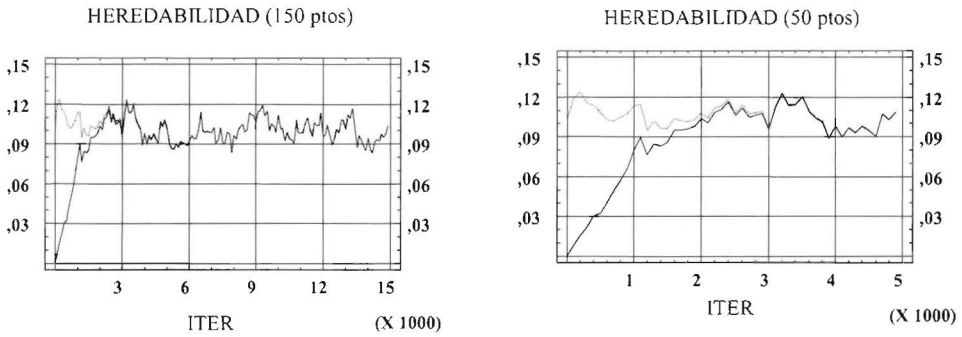


Figura 1. 150 y 50 primeros puntos de las cadenas pareadas para heredabilidad

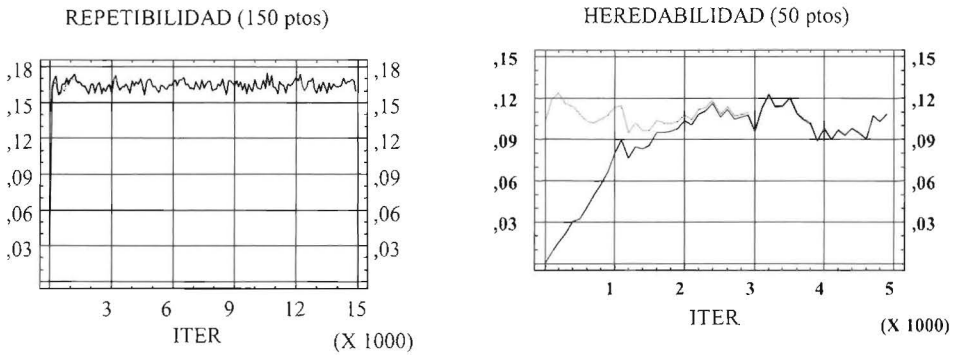


Figura 2. 150 y 50 primeros puntos de las cadenas pareadas para repetibilidad

b) En la escala subyacente el valor de dichos parámetros fueron: para la heredabilidad de $0,101 \pm 0,008$ y para la repetibilidad de $0,164 \pm 0,004$, considerablemente mayores que en la escala visible.

c) Las predicciones genéticas de los animales de la población para el carácter prolificidad en la escala subyacente es posible obtenerlas mediante la técnica del aumento de datos y muestreo de Gibbs.

d) Aunque se presenta autocorrelación entre los puntos integrantes de las distribuciones marginales, esto no parece afectar de forma notable a las estimas de los parámetros genéticos.

e) Parece razonable que las predicciones genéticas de los animales en la escala subyacente deberían ser el criterio de selección de reproductores en razón de su mayor heredabilidad.

Bibliografía

- COX D.R., MILLER H.D., 1965. The theory of Stochastic Process. Chapman and Hall, London. U.K.
- DEMPSTER R., LERNER I.M., 1950. Heritability of threshold characters. *Genetics*, 35, :212.
- FOULLEY J.L., GIANOLA D., IM S., 1987. Genetic evaluation of traits distributed as Poisson-binomial with reference to reproductive traits. *Theor. Appl. Genet.* 73, 870.
- GARCÍA-CORTÉS L.A., RICO M., GROENEVELD E., 1997. Coupling the gibbs Sampler to asses convergence in animal Models. Fotocopia.
- GIANOLA D., 1980. A method of sire evaluation for dichotomies. *J. Anim. Sci.* 51, 1.266-1.271.
- GIANOLA D., 1982. Theory and analysis of threshold characters. *J. Anim. Sci.* 54, 1.079-1.096.
- GIANOLA D., FOULLEY J.L., 1983. Sire evaluation for ordered categorical data with a threshold model. *Gen. Sel. Evol.* 15(2), 201-224.
- GRPEMEVELD, E., 1998.- Programa VCE 4, versión 4.2.
- HARVILLE D.A., MEE R.W., 1984. A mixed-model procedure for analysing ordered categorical data. *Biometrics*, 40, 393-408.
- HOESCHELE I., TIER B., 1995. Estimation of variance components of treshold characters by marginal posterior modes and means via Gibbs sampling. *Genet. Sel. Evol.* 27, 519-540.
- OCCON A., 1988. Repetibilidad y heredabilidad de la prolificidad en la raza rasa Aragonesa. Optimización de los progresos genéticos en los procesos de selección. Tesis doctoral.
- SORENSEN D., ANDERSEN J., GIANOLA D., KORSGAARD I., 1995. Bayesian inference in treshold models using Gibbs sampling. *Genet. Sel. Evol.* 27, 229-249.
- WANG C.S., RUTLEDGE J.J., GIANOLA D., 1994.- Bayesian analysis of mixed models via gibbs sampling with application to litter size in Iberian pigs. *Genet. Sel. Evol.* 26, 91-115.
- WRIGHT., 1934. An analysis in number of digits in an inbred strain of guinea pigs. *Genetics*, 19, 506.