

## BIBLIOGRAFIA

- 1) BOGART, R. (1975).—Biological indicators of inherited differences in rate and efficiency of gains in beef cattle. In *Genetics lectures*. Vol. 4, Oregon State University Press: 139-175.
- 2) CIMA, M.; SÁNCHEZ GARCÍA, L., y RINCÓN, C. (1977).—Papel de las agrupaciones bovinas autóctonas Asturiana, Gallega y Pirenaica, como razas paternas para la producción de carne. *Zootechnia*, XXVIII (4, 5, 6): 165-180.
- 3) GASPARI, J., y STEVENS, R. W. C. (1968).—Bovine serum amylase isozymes in several breeds of domestic cattle. *Can. J. Genet. Cytol.* (10): 148.
- 4) KADIEV, A. K. (1974).—Genetic diversity of transferrin and amylase in relation to growth indices in young cattle. *Tsitologiya i Genetika*, 8 (1): 13-19.
- 5) MASON, I. L. (1975).—Razas españolas de ganado e investigación en grupos sanguíneos. *Informe al C.S.I.C. en relación con las actividades del G.E.I.S.I.M.A.*
- 6) MEDVEDEVA, N. V.; MEDVEDEV, G. F., y KUZUR, E. P. (1975).—Blood characters of Kostroma heifers and their relationships with age, growth intensity, sexual maturity and fertility. *Sbornik Nauchnykh Trudov. Belorusskaya sel'skokhozyais tvennaye Akedemiya* (136): 38-43.
- 7) POULINK, M. D. (1957).—Starch gel electrophoresis in a discontinuous system of buffers. *Nature* (180): 1.147.
- 8) ROWLANDS, G. J.; STARK, A. J.; MANSTON, R.; LEWIS, W. H. E., y SAUNDERS, R. W. (1977).—The blood composition of different breeds of bulls undergoing beef performance test. *Research in Veterinary Science*, 23 (3): 348-350.
- 9) SÁNCHEZ GARCÍA, L. (1978).—Raza Rubia gallega. Evolución, situación actual y perspectivas zootécnicas. *Asociación Nacional de Criadores de ganado vacuno selecto de la raza Rubia gallega*. 319 p.p.
- 10) VALLEJO, M. (1971).—Estudio de la conformación, rendimientos y calidades cárnicas de siete razas bovinas españolas. *Anal. Fac. Vet. Zaragoza*, VI, (6): 263-329.
- 11) VALLEJO, M.; MONGE, E.; RODERO, A.; ZARAZAGA, I.; GARZÓN, R., y LAMUELA, J. M. (1977).—Polimorfismos bioquímicos en razas vacunas españolas. I. Rubia gallega, pirenaica, retinta y morenas del noroeste. *Trab. Cient. Univ. Córdoba* (23): 34 pp.
- 12) VALLEJO, M.; MONGE, E.; ZARAZAGA, I.; LASIERRA, J. M., y LAMUELA, J. M. (1978).—Segregación de los polimorfismos bioquímicos Tf y Am en ganado vacuno. *Zootechnia*, XXVII (1, 2, 3): 15-30.
- 13) YAMAMOTO, Y. (1975).—Transferrin types and performance in Japanese Black cattle. *Jap. J. Zoot. Sci.*, 46 (11): 656-657.

CATEDRA DE GENETICA  
(Prof. Dr. M. VALLEJO VICENTE)

## ESTUDIO GENETICO DE LOS FACTORES QUE INFLUYEN EN LA PRODUCCION LACTEA DEL GANADO OVINO. IV. COMPARACION ENTRE LOS METODOS DE ESTIMACION DE MAXIMA VEROSIMILITUD (ML) Y DE MINIMOS CUADRADOS (LS) PARA EL MODELO MIXTO

Por J. A. Carriedo  
F. San Primitivo

### INTRODUCCION

La influencia de los factores de variación genéticos y ambientales, sobre los caracteres productivos de los animales domésticos, se estudia generalmente mediante la utilización de análisis de varianza.

Cuando para estos estudios se utilizan datos productivos, procedentes de registros de campo, los análisis de varianza suelen presentarse con subclases desiguales y fuertemente desequilibrados; incluso, en determinados casos, pueden manifestar una cierta desconexión entre subclases<sup>10</sup>. En estos casos, adquiere una especial importancia la metodología estadística empleada.

Los análisis de varianza que se utilizan para este tipo de estudios, generalmente siguen un modelo mixto, en el que los factores de tipo genético se consideran «aleatorios» y los ambientales «fijos». Entre los métodos de estimación paramétrica que con mayor frecuencia se han utilizado para el modelo mixto, figuran el de máxima verosimilitud (ML)<sup>9, 11, 12</sup> y el de mínimos cuadrados (LS)<sup>3, 5</sup>.

El análisis de varianza puede utilizarse, bien con la finalidad de cuantificar la importancia relativa de los factores de variación, en términos de proporción de varianza explicada por cada factor significativo, bien con objeto de estimar los efectos de los factores, representados por los términos que figuran en el modelo matemático.

Un factor de variación presenta, en cada uno de sus niveles o subclases, un efecto sobre la variable en estudio, que se corresponde con su desviación respecto a la media de la población, en esa subclase. Los métodos LS y ML permiten estimar estos efectos correspondientes a cada factor de variación.

El método ML proporciona estimaciones con propiedades estadísticamente muy

*An. Fac. Vet. León*, 1982, 28, 95-102.

deseables<sup>1, 7, 12</sup>. Mediante el algoritmo de cálculo iterativo indicado por Henderson<sup>9</sup> y Schaeffer<sup>11</sup>, se obtienen estimaciones por el método de máxima verosimilitud para el modelo mixto. En el estudio de los efectos de los factores, este método proporciona estimaciones lineales de mínima varianza (BLUP), muy adecuadas desde el punto de vista estadístico<sup>13</sup>.

El método LS para el modelo mixto<sup>3, 5</sup> proporciona estimaciones de los efectos con propiedades estadísticamente menos deseables, en comparación con el método ML, siendo este aspecto particularmente aplicable al sesgo de las estimaciones<sup>7</sup> y a su varianza<sup>4</sup>. En contraposición, el método ML es computacionalmente más complejo que el método LS.

Cuando todos los factores del modelo son «fijos», el método LS es muy adecuado y proporciona estimaciones BLUP de los efectos<sup>1</sup>.

El presente trabajo es el cuarto de una serie dedicada al estudio de los factores de variación genéticos y ambientales de la producción láctea en el ganado ovino.

El objetivo general de este estudio es la discusión de los resultados obtenidos mediante la utilización de los métodos ML y LS, en la estimación de los efectos de los factores, a partir de un análisis de varianza con el modelo mixto.

En sentido estricto, el análisis comparativo se basa en las propiedades sesgo y varianza de las estimaciones. Para el modelo mixto, admitimos «a priori» que el método ML es estadísticamente más adecuado que el método LS. Por esta razón, tomando como referencia las estimaciones de los efectos obtenidos por el método ML, trataremos de establecer el grado de validez estadística de las estimaciones mínimo-cuadráticas.

Incluimos también un modelo sencillo, con todos los factores como «fijos», para el que se sigue, únicamente, el método LS.

## METODOLOGIA ESTADISTICA

Se han utilizado 1.125 datos productivos de primera lactación y 830 de la segunda, normalizados a 150 días, así como los correspondientes datos genealógicos, comprobados mediante los polimorfismos bioquímicos Hemoglobina y Transferrina.

Los análisis de varianza se han realizado para cada una de las dos primeras lactaciones por separado. El número de elementos dentro de cada subclase siempre fue superior a cinco.

El modelo mixto utilizado ha sido el siguiente:

$$y_{ijklmn} = \mu + R_i + T_j + E_k + A_{l(i)} + S_{m(i)} + RE_{ik} + e_{ijklmn}$$

donde  $R_i$ ,  $T_j$  y  $E_k$  representan los efectos de los factores de variación rebaño, tipo de parto y estación de parto, respectivamente.  $A_{l(i)}$  y  $S_{m(i)}$  corresponden a los efectos de

los factores año de parto y semental, respectivamente, ambos subordinados al factor rebaño.  $RE_{ik}$  representa los efectos del factor interacción rebaño por estación.  $y_{ijklmn}$  representa la producción láctea de la oveja «n», en la lactación considerada (primera o segunda), perteneciente al rebaño «i», descendiente del semental «m», cuyo parto tipo «j» tuvo lugar en la estación «k» y en el año «l». El factor semental se admite en el modelo como «aleatorio», y los restantes como «fijos».

En nuestro caso específico, debido al tipo de distribución no aleatoria de los sementales a lo largo de los años, creemos que la discusión de la metodología se debe centrar en el factor «aleatorio» semental o en el «fijo» año de parto, ya que el tipo de metodología empleado incide sobre los resultados que se obtengan para ambos factores<sup>6</sup>.

Henderson<sup>8</sup> ya ha realizado un estudio crítico de estos dos métodos (ML y LS), para el modelo mixto, basándose en los factores aleatorios.

En consecuencia, se ha elegido como base de estudio el factor «fijo» año de parto, que, por otra parte, se ha mostrado como el factor más influyente sobre la producción láctea y con un mayor número de niveles o subclases (ver tabla n.º 1).

La discusión de los dos métodos de estimación paramétrica, LS y ML para el modelo mixto, se han efectuado en base a las propiedades sesgo y varianza de las estimaciones de los efectos del factor año de parto, bajo el criterio de que estas estimaciones deben ser insesgadas y de mínima varianza.

Con objeto de analizar el sesgo, las estimaciones mínimo cuadráticas de los efectos del año de parto se han comparado con las realizadas mediante el método de máxima verosimilitud, suponiendo estas últimas como correctas «a priori». El grado de similitud se ha determinado mediante la correlación entre los dos tipos de estimaciones.

La varianza de los estimadores de los efectos se ha determinado a partir de los errores típicos de cada uno de los efectos del año de parto. A partir de estos

**TABLA 1**  
**Resultados de los análisis de varianza siguiendo los métodos LS y ML para el modelo mixto**

Factor de variación (modelo mixto)	Primera lactación			Segunda lactación		
	Grados de libertad	F y significación		Grados de libertad	F y significación	
		LS	ML		LS	ML
Rebaño (R)	2	0,0 NS	0,4 NS	2	1,6 NS	2,3 NS
Estación (E)	3	0,2 NS	0,5 NS	3	2,8 *	3,0 *
Tipo parto (T)	1	1,2 NS	1,7 NS	1	13,6 ***	14,0 ***
Año de parto (A)	18	7,7 ***	8,1 ***	16	6,3 ***	7,7 ***
Interacción (RE)	6	3,6 ***	3,6 ***	6	2,9 **	3,2 **
Semental (S)	36	2,1 ***	—	28	2,1 ***	—
Residual (e)	1.058	—	—	812	—	—

Nivel de significación: \*  $p \leq 0,05$ ; \*\*  $p \leq 0,01$ ; \*\*\*  $p \leq 0,001$ ; NS  $p < 0,05$

resultados, se han obtenido los valores promedio de estos errores típicos, para ambos métodos ML y LS.

Además de considerarse la precisión estadística de las estimaciones de los efectos del año de parto, se han tenido en cuenta los errores típicos de las estimaciones de las medias de este factor. En cada nivel o subclase del factor año de parto, la variable en estudio presenta una «media», que corresponde a la suma del respectivo efecto más la media poblacional. Para ambos métodos, ML y LS, se han obtenido los valores promedio de los errores típicos de las estimaciones de las «medias» debidas al factor año de parto.

En el estudio estadístico se ha incluido, además, un modelo matemático que contempla todos los factores ambientales incluidos en el modelo mixto descrito anteriormente. En este modelo no se incluye el factor «aleatorio» semental y todos los restantes se tratan como «fijos». Por esta razón, únicamente se ha utilizado para su estudio el método de mínimos cuadrados<sup>5</sup>.

En el análisis del sesgo que presentan las estimaciones realizadas por los tres métodos, se han determinado los coeficientes de correlación lineal y de Spearman<sup>14</sup>.

Los análisis de varianza y las estimaciones de los efectos de los factores, con sus correspondientes errores típicos, se han efectuado utilizando los programas LSMLGP y LSML76 de Harvey<sup>2</sup>, 4.

## RESULTADOS Y DISCUSION

En la tabla n.º 1 figuran los resultados de los análisis de varianza, referentes a la primera y segunda lactación, efectuados siguiendo los métodos ML y LS para el modelo mixto. En ella se indica, para cada factor de variación, los grados de libertad y el valor del estadístico «F» del Snedecor con su significación estadística.

En el método de máxima verosimilitud, el factor semental, considerado como «aleatorio» en el modelo, se ha «absorbido» en el proceso de computación<sup>4, 12</sup>, por lo que no figuran resultados referentes a su significación estadística. Cuando se ha seguido el método LS, el factor semental se mostró altamente significativo, siendo la proporción de varianza explicada por él, 4,7 % para la primera y 4,5 % para la segunda lactación. Considerando la influencia altamente significativa del factor semental sobre la producción láctea, se ha creído adecuado contemplar, en nuestra discusión, un modelo mixto en el que figurase este factor «aleatorio» semental.

El año de parto se manifestó como el factor de variación más importante, siendo la proporción de varianza por él explicada, para la primera y segunda lactación: 20 % y 15 %, respectivamente. La importante influencia del año de parto sobre la producción láctea, ha sido uno de los motivos, como ya indicamos, por lo que se ha centrado nuestra discusión en las estimaciones referentes a este factor.

En la tabla n.º 2 figura, para cada uno de los dos métodos LS y ML, así como para las dos primeras lactaciones, el valor promedio de los errores típicos de los efectos

estimados del factor año de parto, el error típico de la media poblacional estimada y el valor promedio de los errores típicos de las «medias» estimadas del año de parto.

**TABLA 2**  
Valores promedio de los errores típicos de los efectos y «medias» estimados por LS y ML (modelo mixto)

N.º de lactación	Método de estimación	Tipo de estimación		
		Factor año de parto Efecto	«Media»	Media poblacional
1	ML	5,69	6,68	2,41
	LS	8,85	11,21	3,96
2	ML	5,84	7,02	2,53
	LS	8,34	10,50	3,40

Tanto para las «medias» del año de parto, como para los efectos y la media poblacional, los errores típicos de las estimaciones mínimo cuadráticas fueron considerablemente superiores, en relación a los obtenidos por el método ML, siendo este aumento del orden del 50 %.

Aunque «a priori» se esperaba que el método ML proporcionase estimaciones cuya varianza no fuera inferior a la de las estimaciones mínimo cuadráticas<sup>4</sup>, se desconocía si ambos métodos, en nuestro tipo de estudios, proporcionarían estimaciones cuya precisión estadística fuese apreciablemente diferente.

En este sentido, nuestros resultados nos permiten admitir que, para el modelo mixto seguido en nuestro estudio, el método ML proporciona estimaciones de los efectos y de las «medias» del año de parto, con una notoria mayor precisión, en relación al método LS.

Se han determinado los coeficientes de correlación entre las estimaciones puntuales de los efectos del factor año de parto, obtenidas por los siguientes métodos estadísticos:

- 1.º Modelo con factores fijos y método de mínimos cuadrados (LSf).
- 2.º Modelo mixto y método de mínimos cuadrados (LSm).
- 3.º Modelo mixto y método de máxima verosimilitud (MLm).

En la tabla n.º 3 figuran, para cada pareja de estos tres métodos, las estimaciones del coeficiente de correlación lineal (r) y de Spearman (rs), obtenidos. Por encima de la diagonal se incluyen los coeficientes referentes a la primera lactación, y por debajo, los relativos a la segunda.

Para los métodos MLm y LSf, el coeficiente de correlación obtenido, 0,97, fue lo suficientemente alto para permitirnos suponer que ambos métodos proporcionan estimaciones puntuales de los efectos del año de parto, relativamente semejantes. Estos resultados están a favor de admitir que, para el tipo de material que se ha

**TABLA 3**  
**Coefficientes de correlación entre las estimaciones de los efectos del año de parto efectuadas a partir de los tres métodos**

	Método de estimación		
	LSr	MLm	LSm
LSr	— —	r = 0,97 rs = 0,97	r = 0,91 rs = 0,92
MLm	r = 0,98 rs = 0,98	— —	r = 0,94 rs = 0,94
LSm	r = 0,82 rs = 0,84	r = 0,87 rs = 0,88	— —

estudiado, la no inclusión en el modelo del factor «aleatorio» semental, no implica la introducción de un sesgo en las estimaciones mínimo cuadráticas de los efectos ambientales «fijos». Esta indicación concuerda con la discusión realizada por Henderson<sup>8</sup>.

Para los métodos LS<sub>m</sub> y ML<sub>m</sub>, los coeficientes de correlación que se han obtenido, 0,91, 0,92, 0,82 y 0,84, creemos que, en conjunto, no fueron lo suficientemente altos como para poder considerar que ambos métodos suministraron estimaciones similares. Para el modelo mixto se ha admitido «a priori», que el método de máxima verosimilitud proporciona estimaciones de los efectos insesgadas<sup>7</sup>, por lo que los resultados que se han obtenido nos hacen suponer que, en nuestro caso, las estimaciones mínimo cuadráticas de los efectos tienen una menor validez estadística.

En conclusión, los resultados referentes al sesgo y varianza de las estimaciones de los efectos, nos induce a admitir que, en el tipo de estudios considerado, para realizar estas estimaciones a partir del modelo mixto, el método de mínimos cuadrados presenta una validez estadística apreciablemente inferior en relación al método de máxima verosimilitud. En nuestro caso, para realizar estimaciones mínimo cuadráticas de los efectos ambientales «fijos», es más adecuado estadísticamente no incluir en el modelo el factor «aleatorio» semental.

#### AGRADECIMIENTOS

A D. Rafael Díez, del Departamento de Genética Cuantitativa y Mejora Animal del CRIDA 06, por su cooperación en el tratamiento informático llevado a cabo en el Centro de Cálculo del INIA (Madrid).

#### RESUMEN

Se discute la utilización del método de máxima verosimilitud (ML) y de mínimos cuadrados (LS) para el modelo mixto, con el factor genético progenitor o semental

como «aleatorio» y los ambientales «fijos». Para ello se han considerado en los análisis 1.125 y 840 registros de producción láctea ovina, de la primera y segunda lactación, respectivamente.

Los resultados referentes al sesgo y varianza de las estimaciones de los efectos de los factores, nos inducen a admitir que, en nuestro tipo de estudios, para realizar estas estimaciones a partir del modelo mixto, el método LS tiene una validez estadística apreciablemente inferior al método ML. Si se opta por utilizar el método LS para estimar los efectos ambientales, es estadísticamente más adecuado no incluir en el modelo el factor «aleatorio» semental.

### GENETIC STUDY OF SOME FACTORS INFLUENCING THE MILK PRODUCTION OF DAIRY EWES. IV. COMPARATION BETWEEN THE MAXIMUM LIKELIHOOD AND THE LEAST SQUARES METHODS FOR THE MIXED MODEL

#### SUMMARY

The validity of the maximum likelihood (ML) and the least squares (LS) methods for the mixed model (with the random sire factor and the fixed environmental factors) are discussed. For this, we have considered in the analysis 1125 and 840 records of milk production from the first and second lactation, respectively.

The results relating to bias and variance of the effect estimations would indicate that in order to estimate the effects of the factors for the mixed model, the LS has a lower statistic validity than the ML method. If we choose the LS method to estimate the environmental effects, it is more suitable, statistically, not to include the sire random factor in the model.

#### BIBLIOGRAFIA

- 1) ARNAIZ, G. (1978).—*Introducción a la Estadística Teórica*. Ed. lex Nova, Madrid: pp. 734-746.
- 2) HARVEY, W. R. (1968).—User's guide for LSMLGP. The Ohio State University. Columbus. Mimeografía.
- 3) HARVEY, W. R. (1970).—Estimation of variance and covariance components in the mixed model. *Biometrics*, **26**: 485-496.
- 4) HARVEY, W. R. (1977).—User's guide for LSML76. The Ohio State University. Columbus. Mimeografía.
- 5) HARVEY, W. R. (1979).—Least-Squares Analysis of Data. Mimeografía.
- 6) HENDERSON, C. R. (1953).—Estimation of variance and covariance components. *Biometrics*, **9**: 226-252.
- 7) HENDERSON, C. R. (1974).—General flexibility of linear model techniques for sire evaluation. *J. Dairy Sci.*, **57**: 963-972.
- 8) HENDERSON, C. R. (1975).—Comparison of alternative sire evaluation methods. *J. Anim. Sci.*, **41**: 760-770.

- 9) HENDERSON, C. R. (1977).—Prediction of future records. Proceedings of the International Conference on Quantitative Genetics. Iowa State University Press, Ames, Iowa.
- 10) SCHAEFFER, L. R. (1975).—Disconnectedness and variance component estimation. *Biometrics*, **31**: 969-977.
- 11) SCHAEFFER, L. R. (1976).—Maximum likelihood estimation of variance components in dairy cattle breeding research. *J. Dairy Sci.*, **59**: 2.146-2.151.
- 12) SEARLE, S. R. (1971).—*Linear Models*. Ed. John Wiley & Sons, New York: pp. 226-327 y 458-470.
- 13) RIOS, S. (1977).—*Métodos Estadísticos*. Ediciones del Castillo. Madrid: pp. 331-334.
- 14) SNEDECOR, G. W., y COCHRAN, W. G. (1975).—*Métodos Estadísticos*. Ed. Compañía Editorial Continental, S. A., México: pp. 244-246.

CATEDRA DE MATEMÁTICAS Y ESTADÍSTICA  
(Prof. Dr. A. ALVAREZ PRIETO)

## DISTRIBUCION DE WEIBULL, CARACTERIZACION Y ALGUNAS APLICACIONES

Por A. Alvarez Prieto  
M. Díaz Gabela

### INTRODUCCION

La distribución Weibull es una distribución triparamétrica, definida a partir de la distribución exponencial.

Su utilización se extiende, de forma general, a los casos en que no se cumplen las condiciones de aleatoriedad suficientes para aplicar la distribución exponencial; sin embargo, no existen razonamientos matemáticos lo suficientemente claros que indiquen cuándo esta distribución deba ser usada.

Nuestro trabajo, que parte de un estudio de caracterización de la distribución Weibull, se plantea como objetivo fundamental la búsqueda de situaciones, procesos y variables aleatorias que lleven asociada la distribución Weibull, es decir, nos proponemos el estudio de circunstancias concretas, en las que dicha distribución pueda aplicarse.

### MATERIAL Y METODOS

Partimos de una serie de propiedades, métodos y conocimientos necesarios para el desarrollo de este trabajo.

#### 1. Caracterización de la distribución Weibull

a) Función de densidad.—Una variable aleatoria "X" se dice que sigue una distribución Weibull si existen parámetros  $\alpha$ ,  $\delta$ ,  $\xi$ , tal que  $\alpha > 0$ ,  $\delta > 0$ ,  $\xi \geq 0$ , que hacen que la variable aleatoria  $Y = \left(\frac{X-\xi}{\alpha}\right)^\delta$  siga una distribución exponencial con función de densidad:  $f(y) = e^{-y}$ ;  $y > 0$ , por lo que la función de densidad de la v.a. "X" será:

$$f(X) = \frac{\delta}{\alpha} \left(\frac{X-\xi}{\alpha}\right)^{\delta-1} e^{-\left(\frac{X-\xi}{\alpha}\right)^\delta} ; X-\xi \geq 0$$

An. Fac. Vet. León, 1982, 28, 103-115.