

Evaluación genética para caracteres de valoración subjetiva en la raza Pirenaica

L. Varona, C. Moreno, J. Altarriba

Unidad de Genética Cuantitativa y Mejora Animal. Universidad de Zaragoza. 50013 Zaragoza

E-mail: lvarona@unizar.es

Resumen

Los esquemas de selección en las especies ganaderas utilizan una amplia variedad de caracteres. En algunos casos, los registros fenotípicos se obtienen a partir de una valoración subjetiva por parte de evaluadores expertos. Esta valoración implica una clasificación en una escala arbitraria, y, por este motivo, puede diferir considerablemente de la distribución Normal. Por otra parte, cada evaluador puede utilizar criterios de clasificación específicos, y diferentes de los otros evaluadores. En este trabajo se propone un modelo multi-umbral para el análisis de datos procedentes de valoraciones subjetivas. El modelo asume una escala observable diferente para cada evaluador o grupo de evaluadores, y una escala subyacente común. El modelo propuesto se ha aplicado a datos de conformación de la canal de la Raza Bovina Pirenaica procedentes del sistema de valoración SEUROP en 12 mataderos del País Vasco y Navarra.

Palabras clave: Modelo Umbral, Valoración Subjetiva, SEUROP, Conformación, Análisis Bayesiano

Summary

Genetic evaluation for subjective traits in the Pirenaica Breed

Selection programs in livestock populations made use of a wide variety of traits. Among them, phenotypic records for some traits are obtained by a subjective evaluation from a set of experts, like sensory, type, carcass or fat score traits. Data from subjective evaluation usually involves a classification under an arbitrary predefined scale. The output of this process can lead to strong departures from the Gaussian distribution. Moreover, different criteria can be achieved for each expert. In this study, we propose a Slaughterhouse Specific Ordered Category Threshold Model, that assumes a specific observable scale for each specialist, and a common subjacent scale. The procedure is applied to SEUROP conformation score data from the Pirenaica Beef Cattle Breed evaluated at 12 different slaughterhouses from the Basque Country and Navarre.

Key words: Threshold model, Subjective Evaluation, SEUROP, Conformation, Bayesian Analysis

Introducción

Los programas de mejora genética de las especies ganaderas utilizan una amplia variedad de caracteres. En algunos casos, los registros fenotípicos se obtienen a partir de una valoración subjetiva, como la valoración

lineal de los caracteres de tipo en vacuno lechero (Gengler *et al.*, 1999), las valoraciones de conformación de la canal o de cobertura grasa en vacuno de carne (Hickey *et al.*, 2007; Altarriba *et al.*, 2008), o en fenotipos procedentes de análisis sensorial (Kararnichou *et al.*, 2007).

Estas valoraciones subjetivas implican una categorización bajo una escala predefinida y arbitraria, y, como consecuencia, los datos procedentes de este proceso pueden presentar una fuerte divergencia frente a la distribución Normal. La alternativa habitual para el análisis de datos categóricos es el modelo umbral (Wright, 1934; Gianola, 1982), que asume una variable subyacente continua, bien Normal – modelo probit– o bien logística –modelo logit–. El modelo umbral asume que la transformación de la escala observable a la escala real se realiza a través de una serie de umbrales que determinan la asignación de cada valor de la escala subyacente a las distintas categorías de la escala real.

En el caso de las valoraciones subjetivas, cada evaluador puede utilizar diferentes regiones de la escala de valoración, o utilizar un rango más o menos amplio de valores. Por lo tanto, el proceso de categorización utilizado por cada evaluador puede ser específico. En este sentido, Varona y Hernandez (2006) propusieron un modelo multi-umbral para el tratamiento de datos de análisis sensorial, y mostraron que cada panelista utilizaba un patrón de valoración específico. El objetivo de este trabajo es desarrollar este modelo multi-umbral en el contexto de un modelo animal sobre datos de conformación de la canal en la Raza Bovina Pirenaica, obtenidos a partir del sistema SEUROP de valoración de canales.

Material y métodos

Datos

La base de datos utilizada en este análisis procede del sistema nacional de registro de canales (SIMOGAN). Se utilizaron 14031 datos de conformación de la canal proce-

dentos de 477 rebaños de la raza Pirenaica. La valoración de conformación de la canal se realiza de manera subjetiva y describe el desarrollo de algunos componentes esenciales en la canal. Bajo el sistema SEUROP de valoración se definen 6 clases, representadas con las letras S, E, U, R, O, P, en una escala que se incrementa desde P, la peor, hasta S, la mejor conformación (CEE 2930/21/1981). La regulación europea permite dividir cada una de las clases en tres categorías, así la categoría S se divide en S+, S y S-, en orden decreciente. Los animales fueron sacrificados entre los años 1999 y 2006 en 12 mataderos localizados en el País Vasco y Navarra. La descripción de la base de datos entre mataderos y categorías se presenta en la tabla 1. Además, para la ejecución del análisis se utilizó una genealogía de 34843 individuos proporcionados por la Confederación Nacional de Asociaciones de Ganado Pirenaico (CONASPI).

Análisis estadístico

El modelo de análisis fue un modelo umbral, pero se definió un grupo de umbrales particular para cada matadero ($t_{i1}, t_{i2}, t_{i3}, \dots, t_{ir}$), donde t_{ij} es el j -ésimo umbral para el i -ésimo matadero y r es el número de umbrales. La distribución condicional de los datos observados dada la variable subyacente y los umbrales es:

$$f(\mathbf{y}|\mathbf{L}, \mathbf{t}) = \prod_{i=1}^{NM} \prod_{j=1}^{n_i} f(y_{ij}|t_{ij}, \mathbf{t}_i)$$

$$f(y_{ij}|t_{ij}, \mathbf{t}_i) = 1(t_{ij} < t_{i1})1(y_{ij} = 1) + 1(t_{i1} < t_{ij} < t_{i2})$$

$$1(y_{ij} = 2) + \dots + 1(t_{ir} < t_{ij})1(y_{ij} = r + 1)$$

Donde NM es el número de mataderos, n_i es el número de datos en el i -ésimo matadero, y \mathbf{t}_i es el vector de umbrales en el i th matadero. El modelo asumido para el vector de variables subyacentes (\mathbf{L}) fue:

$$\mathbf{L} = \mathbf{Xb} + \mathbf{Wp} + \mathbf{Zu} + \mathbf{e}$$

Tabla 1. Distribución de las observaciones del sistema SEUROP en los 12 mataderos analizados
 Table 1. Distribution of SEUROP conformation data on the 12 slaughterhouses

Val	Matadero												Total	
	1	2	3	4	5	6	7	8	9	10	11	12		
S	+	0	0	0	0	0	12	0	0	0	0	0	0	12
	=	0	0	0	0	0	0	0	0	0	0	0	0	0
	-	0	0	0	0	0	0	0	0	4	0	0	0	4
E	+	0	7	0	22	0	5	0	3	46	8	2	0	93
	=	0	0	0	1	1	0	0	0	53	0	0	0	55
	-	0	2	7	15	3	1	0	2	709	3	16	1	759
U	+	21	158	67	262	13	28	69	16	3089	162	229	20	4134
	=	0	4	52	55	4	5	0	1	3294	9	178	5	3607
	-	0	45	110	173	7	24	5	0	1911	26	459	19	2779
R	+	29	117	165	266	214	143	161	15	325	150	640	175	2400
	=	0	1	4	26	0	8	0	0	8	4	40	2	93
	-	0	0	0	0	0	0	0	0	1	0	0	0	1
O	+	0	9	0	6	1	0	0	0	2	5	42	17	82
	=	0	0	0	0	0	0	0	0	0	0	0	0	0
	-	0	0	0	0	0	0	0	0	0	0	0	0	0
P	+	0	0	0	0	1	1	0	0	1	0	0	9	12
	=	0	0	0	0	0	0	0	0	0	0	0	0	0
	-	0	0	0	0	0	0	0	0	0	0	0	0	0
TOTAL		50	343	405	826	244	227	235	37	9443	367	1606	248	14031

donde \mathbf{b} es el vector de efecto sexo (2 niveles), \mathbf{p} es el vector de efectos rebaño-año-estación (477 niveles) y \mathbf{e} es un vector de residuos normales fijado arbitrariamente de media cero y varianza uno. Como distribuciones a-priori, se asumieron distribuciones multivariantes para \mathbf{p} and \mathbf{u} , y distribuciones uniformes acotadas para los efectos sistemáticos y los componentes de la varianza.

Las distribuciones marginales se calcularon utilizando un muestreo de Gibbs (Gelfand and Smith, 1990), con un paso de aumento de datos (Albert and Chib, 1993; Sorensen et al., 1995) para muestrear de las distribuciones predictivas de la variable subyacente. Se utili-

zó una única cadena de muestreo de Gibbs de 500000 iteraciones, tras descartar las primeras 25000 como periodo de quemado.

Resultados y discusión

Los resultados de las medias y desviaciones típicas de las distribuciones posteriores de los umbrales específicos para cada matadero se presentan en la tabla 2. La magnitud de todos los umbrales y de los efectos sistemáticos está condicionada a la definición arbitraria de la varianza residual en 1, y a la asunción del valor 0 para el umbral entre las

Tabla 2. Media y desviación típica posterior de los umbrales
 Table 2. Posterior mean and standard deviation of the threshold values

UMB	Matadero											
	1	2	3	4	5	6	7	8	9	10	11	12
S+/S	-	-	-	-	-	4.18 (0.24)	-	-	-	-	-	-
S/S-	-	-	-	-	-	4.14 (0.24)	-	-	-	-	-	-
S-/E+	-	-	-	-	-	4.06 (0.24)	-	-	4.79 (0.23)	-	-	-
E+/E	-	4.01 (0.22)	-	3.90 (0.23)	-	3.81 (0.24)	-	3.27 (0.44)	3.91 (0.07)	3.76 (0.18)	5.44 (0.38)	-
E/E-	-	3.93 (0.22)	-	3.85 (0.23)	5.29 (0.55)	3.76 (0.25)	-	3.05 (0.42)	3.61 (0.06)	3.71 (0.17)	5.24 (0.36)	-
E-/U+	-	3.74 (0.21)	4.12 (0.31)	3.55 (0.23)	4.36 (0.35)	3.67 (0.25)	-	2.54 (0.36)	2.49 (0.04)	3.50 (0.16)	4.18 (0.25)	4.98 (0.52)
U+/U	2.62 (0.36)	1.37 (0.16)	2.55 (0.26)	1.91 (0.22)	3.33 (0.25)	2.85 (0.23)	2.22 (0.20)	1.13 (0.25)	1.04 (0.02)	1.23 (0.07)	2.59 (0.23)	3.13 (0.27)
U/U-	2.56 (0.35)	1.27 (0.16)	2.01 (0.26)	1.61 (0.22)	3.13 (0.24)	2.44 (0.23)	2.20 (0.20)	0.81 (0.25)	0	1.08 (0.07)	2.08 (0.24)	2.94 (0.26)
U-/R+	2.49 (0.36)	0.84 (0.16)	1.10 (0.26)	0.93 (0.22)	2.84 (0.23)	2.04 (0.22)	2.10 (0.20)	0.72 (0.25)	-1.34 (0.04)	0.84 (0.07)	1.12 (0.24)	2.45 (0.23)
R+/R	-	-1.02 (0.21)	-1.68 (0.34)	-0.89 (0.24)	-1.71 (0.30)	-0.70 (0.28)	-	-	-2.83 (0.10)	-1.34 (0.16)	-0.75 (0.24)	-0.30 (0.24)
R/R-	-	-1.11 (0.22)	-	-1.73 (0.28)	-1.82 (0.32)	-1.44 (0.35)	-	-	-3.20 (0.15)	-1.64 (0.18)	-1.10 (0.24)	-0.37 (0.24)
R-/O+	-	-1.17 (0.23)	-	-1.80 (0.29)	-1.95 (0.34)	-1.59 (0.37)	-	-	-3.34 (0.18)	-1.74 (0.20)	-1.12 (0.24)	-0.40 (0.24)
O+/O	-	-	-	-	-2.29 (0.39)	-1.79 (0.41)	-	-	-3.67 (0.25)	-	-	-1.06 (0.26)
O/O-	-	-	-	-	-2.52 (0.45)	-2.04 (0.46)	-	-	-3.87 (0.30)	-	-	-1.12 (0.27)
O-/P+	-	-	-	-	-2.82 (0.52)	-2.33 (0.52)	-	-	-4.09 (0.37)	-	-	-1.19 (0.27)

categorías U y U- en el matadero 9. Los resultados muestran la especificidad en la asignación de categorías en cada uno de los mataderos. En particular, este fenómeno se observa claramente entre las categorías U- y R+, donde, para el matadero 9, el umbral se sitúa en -1.34 (0.04), mientras para todos los demás mataderos se sitúa en valores posi-

vos. En general, los resultados de los umbrales son consistentes con las observaciones fenotípicas presentadas en la tabla 1, y, en este sentido, la heterogeneidad de criterio en la asignación en categorías es considerada por el modelo, mientras que los procedimientos habituales desprecian este fenómeno. Este hecho se corrobora con los resulta-

dos de comparación de modelos frente a un modelo lineal y a un modelo de datos categóricos con un único grupo de umbrales (resultados no mostrados), que definió claramente al modelo propuesto como el más plausible, tanto mediante el Deviance Information Criterion –DIC– (Spiegelhalter et al, 2002), como mediante el PseudoBayes Factor (Newton and Raftery, 1994).

El modelo proporciona también información acerca de los efectos sistemáticos. En la población analizada, la media posterior (desviación típica) del efecto atribuido a los machos fue de 1.16 (0.09) y el de las hembras fue 0.70 (0.10). El modelo no incluyó ningún efecto sistemático asignado a los mataderos, ya que las diferencias entre estos quedan absorbidas por la diferente localización de los umbrales. En este sentido, se observa cómo el matadero 9 proporciona valores inferiores al resto de los mataderos, de manera consistente con los datos fenotípicos presentados en la tabla 1.

La media posterior del porcentaje de varianza atribuida al efecto rebaño-año-estación fue de 0.27 con una desviación típica de 0.02. En cuanto a la heredabilidad la media posterior fue de 0.18, con una desviación típica posterior de 0.02. Los resultados de heredabilidad fueron inferiores a los obtenidos mediante un modelo lineal en la misma población (Altarriba et al, 2008).

El modelo propuesto considera la conformación como un carácter continuo que se expresa igualmente en todos los mataderos, y, por lo tanto, la variable subyacente es la misma en todos ellos, pero su transformación a la escala observable es dependiente del matadero. Algunas otras alternativas se pueden considerar en el futuro, como el análisis multicaracter, así como la consideración de variables difusas en la definición de las valoraciones subjetivas.

Agradecimientos

Los autores agradecen a la Confederación Nacional de Asociaciones de Ganado Pirenaica (CONASPI) por provisión de datos para este trabajo.

Bibliografía

- Albert JH, Chib S, 1993. Bayesian-analysis of binary and polychotomous response data. *J Am. Stat. Soc.*, 88, 669-79.
- Altarriba J, Yagüe G, Moreno C, Varona L, 2008. Exploring possibilities of the traceability data for genetic improvement (en preparación).
- CEE nº 2930/81. 1981. Community scale for the classification of carcass of adult bovine animals. Official publications of the European communities, L-2985 Luxemburg.
- Gelfand AE, Smith AFM, 1990. Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Soc.*, 85, 398-409.
- Gengler N, Wiggans GR, Wright JR, 1999. Animal Model Genetic Evaluation of Type Traits for Five Dairy Cattle Breeds. *J. Dairy Sci.*, 82, 1350.
- Gianola D, 1982. Theory and analysis of threshold characters. *J Anim Sci*, 54, 1079-96.
- Hickey JM, Keane MG, Kenny DA, Cromie AR, Veerkamp RF, 2007. Genetic parameters for EUROP carcass traits within different groups of cattle in Ireland. *J. Anim. Sci.*, 85, 314-321.
- Kararnichou E, Richardson RI, Nute GR, Wood JD, Bishop SC, 2007. Genetic analyses of sensory characteristics and relationships with fatty acid composition in the meat from Scottish Blackface lambs. *Animal*, 10, 1524-1531.
- Newton MA, Raftery AE, 1994. Approximate Bayesian inference by the weighted likelihood bootstrap (with discussion). *J. Royal Stat. Soc. B.*, 56, 3-48.
- Sorensen DA, Andersen S, GianolaD, Korsgaard I, 1995. Bayesian Inference in threshold models

- using gibbs sampling. *Genet. Sel. Evol.*, 27, 229-49.
- Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A, 2002. Bayesian measures of model complexity and fit (with discussion). *J. Royal Stat. Soc. B.*, 64, 583-639.
- Varona L, Hernandez P, 2006. A multithreshold model for sensory analysis. *J. Food Sci.*, 71, 333-336.
- Wright S, 1934. An analysis of variability in number of digits in an inbred strain of guinea pigs. *Genetics*, 19, 506-36.

(Aceptado para publicación el 28 de abril de 2008)