

UN MODELO BINOMIAL RECURSIVO PARA LA MORTALIDAD DE LOS LECHONES

Varona, L.¹ y Sorensen, D. A.²

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

2 Department of Genetics and Biotechnology. Faculty of Agricultural Sciences. University of Aarhus, DK-8830 Tjele, Dinamarca.

lvarona@unizar.es

INTRODUCCIÓN

Los caracteres de interés económico en producción animal presentan en algunas ocasiones distribuciones muy alejadas de la normalidad. Este hecho dificulta la aplicación de modelos lineales para la evaluación genética y la estimación de componentes de varianza. Pese a esta dificultad, los procedimientos basados en inferencia bayesiana permiten utilizar modelos explicativos de los datos que utilizan distribuciones alternativas para la descripción de la distribución condicional de los datos dados los parámetros.

Un claro ejemplo de un carácter alejado de la normalidad es el número de lechones muertos como carácter atribuible a la madre. En este sentido, Varona y Sorensen (2010) estudiaron el grado de ajuste de varias distribuciones candidatas (Poisson, Binomial y Binomial Negativa), tanto en su versión habitual como en la modalidad de cero inflado. El resultado fue que el modelo más apropiado fue la distribución binomial.

Sin embargo, un limitante para la utilización de estas distribuciones no gaussianas radica en la dificultad para modelizar relaciones entre caracteres. En este sentido, los modelos recursivos (Gianola y Sorensen, 2004; Varona et al., 2007) pueden permitir una interesante alternativa para considerar relaciones entre caracteres con distribuciones no gaussianas. El objetivo de este trabajo es extender el análisis del carácter número de lechones nacidos muertos con la inclusión de recursividad entre el carácter número total de lechones y número de lechones muertos, ambos atribuidos la madre.

MATERIAL Y MÉTODOS

Para este trabajo se utilizaron datos de poblaciones porcinas danesas registrados entre mayo de 2002 y diciembre de 2004. La base de datos consistió en información procedente de 5178 y 3938 camadas de las poblaciones Landrace y Yorkshire, respectivamente. En cada camada, se registró el número de lechones nacidos totales y muertos. Además, se utilizaron dos ficheros genealógicos de 8800 individuos en la población Landrace y de 7143 en la población Yorkshire.

Los datos se analizaron mediante inferencia bayesiana utilizando métodos de MCMC. Para ello se definió la distribución posterior conjunta de los parámetros desconocidos dados los datos.

$$\begin{aligned} & f(\mathbf{b}_\varphi, \mathbf{b}_t, \mathbf{u}_\varphi, \mathbf{u}_t, \mathbf{p}_\varphi, \mathbf{p}_t, \sigma_{u\varphi}^2, \sigma_{ut}^2, \sigma_{p\varphi}^2, \sigma_{pt}^2, \sigma_{e\varphi}^2, \lambda | \mathbf{y}, \mathbf{t}) \\ & \propto f(\mathbf{y} | \varphi, \mathbf{t}) f(\varphi | \mathbf{b}_\varphi, \mathbf{u}_\varphi, \mathbf{p}_\varphi, \lambda, \mathbf{t}) f(\mathbf{t} | \mathbf{b}_t, \mathbf{u}_t, \mathbf{p}_t, \sigma_{e\varphi}^2) f(\mathbf{u}_\varphi | \sigma_{u\varphi}^2) f(\mathbf{u}_t | \sigma_{ut}^2) \\ & f(\mathbf{p}_\varphi | \sigma_{p\varphi}^2) f(\mathbf{p}_t | \sigma_{pt}^2) f(\mathbf{b}_\varphi) f(\mathbf{b}_t) f(\sigma_{u\varphi}^2) f(\sigma_{ut}^2) f(\sigma_{p\varphi}^2) f(\sigma_{pt}^2) f(\sigma_{e\varphi}^2) f(\lambda) \end{aligned}$$

Donde \mathbf{y} es el vector de datos de lechones nacidos muertos por camada, \mathbf{t} es el vector de datos de lechones nacidos totales por camada, φ es el vector de probabilidades de nacer vivo o muerto en cada camada, \mathbf{b} , \mathbf{u} y \mathbf{p} son los vectores de efectos sistemáticos, genéticos aditivos y ambientales permanentes, tanto para la probabilidad de nacer muerto (φ) como para el número total de lechones nacidos (\mathbf{t}). Además σ_u^2 , σ_p^2 son los componentes de varianza genética aditiva y ambiental permanente para φ y \mathbf{t} . Finalmente σ_e^2 es la varianza residual para el número total de nacidos y λ es el parámetro de recursividad entre el número total de nacidos vivos y la transformación logit de la probabilidad de nacer muerto.

La probabilidad de los datos dados los parámetros fue

$$f(\mathbf{y}|\mathbf{t}, \boldsymbol{\varphi}_i) = \prod_{i=1}^N \binom{t_i}{y_i} \varphi_i^{y_i} (1 - \varphi_i)^{t_i - y_i}$$

donde N es la número total de datos, $\mathbf{t} = \{t_i\}$ y $\boldsymbol{\varphi} = \{\varphi_i\}$.

En una jerarquía previa se asume que:

$$\text{logit}(\boldsymbol{\varphi}) = \mathbf{X}\mathbf{b}_\varphi + \mathbf{Z}\mathbf{u}_\varphi + \mathbf{W}\mathbf{p}_\varphi + \lambda\mathbf{t}$$

$$\mathbf{t} = \mathbf{X}\mathbf{b}_t + \mathbf{Z}\mathbf{u}_t + \mathbf{W}\mathbf{p}_t + \mathbf{e}_t$$

Además, se definieron las siguientes distribuciones a priori para los valores genéticos, efectos permanentes y residuales:

$$\mathbf{u}_\varphi | \mathbf{A}, \sigma_{u_\varphi}^2 \sim N(\mathbf{0}, \mathbf{A}\sigma_{u_\varphi}^2) \quad \mathbf{u}_t | \mathbf{A}, \sigma_{u_t}^2 \sim N(\mathbf{0}, \mathbf{A}\sigma_{u_t}^2)$$

$$\mathbf{p}_\varphi | \sigma_{p_\varphi}^2 \sim N(\mathbf{0}, \mathbf{I}\sigma_{p_\varphi}^2) \quad \mathbf{p}_t | \sigma_{p_t}^2 \sim N(\mathbf{0}, \mathbf{I}\sigma_{p_t}^2)$$

$$\mathbf{e}_t | \sigma_{e_t}^2 \sim N(\mathbf{0}, \mathbf{I}\sigma_{e_t}^2)$$

Finalmente, las distribuciones a priori para los componentes de varianza y el parámetro de recursividad se asumieron uniformas en un rango de valores posibles.

Los resultados de este modelo se han comparado con los obtenidos en Varona y Sorensen (2010) mediante la suma de los logaritmos de la ordenada de la distribución condicional predictiva (CPO) que se aproxima mediante:

$$\hat{p}(Y_i = y_i | y_{-i}, M_k) = n \times \left[\sum_{j=1}^{N_s} \frac{1}{p(Y_i = y_i | \boldsymbol{\theta}_i^{(j)}, M_k)} \right]$$

donde $\boldsymbol{\theta}_i^{(j)}$ es la *j*th muestra de los parámetros en el proceso de MCMC y N_s es el número de muestras, 500,000 en este caso, después de descartar las primeras 100,000.

RESULTADOS Y DISCUSIÓN

Las medias y desviaciones típicas posteriores estimadas para los componentes de varianza y el parámetro de recursividad se presentan en las Tabla 1.

Los resultados de componentes de varianza para la transformación logit de la probabilidad de nacer muerto son similares a los obtenidos por Varona y Sorensen (2010), pero cuando se compararon los modelos mediante el logaritmo de la CPO, el modelo con recursión obtuvo una logCPO considerablemente superior, -9320 vs. 10162 en Landrace y -6185 vs. -6358 en Yorkshire. Estos resultados indican que el modelo que incluye recursividad entre el número de lechones totales y la probabilidad de nacer muerto mejora el grado de ajuste del modelo e incluye la relación entre ambos caracteres.

Este resultado muestra que los modelos recursivos permiten incluir relaciones entre caracteres aun cuando la distribución de los datos condicionada a los parámetros sea diferente entre ellos, y además puede generalizarse hacia modelos no lineales de relación como los descritos por López de Maturana et al. (2009) e Ibáñez-Escriche et al. (2010).

REFERENCIAS BIBLIOGRÁFICAS

- Ibáñez-Escriche, N., López de Maturana, E., Noguera, J. L. & Varona, L. 2010. J. Anim. Sci. 88: 3493-3503.
- Gianola, D & Sorensen, D. 2004. Genetics 167:1407-1424.
- Lopez de Maturana, E., Wu. X., Gianola. D, Weigel, K. A. & Rosa G. J. M. 2009. Genetics 181: 277-287.
- Varona, L., Sorensen, D. & Thompson R. 2007. Genetics 177:1791-1799.
- Varona, L. & Sorensen, D. 2010. Genetics 184: 277-284.

Tabla 1. Medias y desviaciones típicas posteriores de los componentes de varianza, heredabilidad y parámetro de recursividad en las poblaciones Landrace y Yorkshire danesas.

	Landrace		Yorkshire	
	Media posterior	Desviación típica posterior	Media posterior	Desviación típica posterior
$\sigma_{u\phi}^2$	0.162	0.032	0.165	0.046
σ_{ut}^2	0.907	0.216	0.800	0.235
$\sigma_{p\phi}^2$	0.281	0.028	0.494	0.049
σ_{pt}^2	0.757	0.320	0.738	0.339
σ_{et}^2	10.015	0.355	9.103	0.378
h_t^2	0.078	0.018	0.075	0.022
λ	0.094	0.005	0.121	0.007

A RECURSIVE BINOMIAL MODEL FOR PIGLET MORTALITY

ABSTRACT: Some interesting traits animal production does not follow a Gaussian distribution. Bayesian analysis allows to model successfully alternative distributions of data, although is not clear how to model relationship between traits. Recursive models provide a useful alternative that it is here used for piglet mortality and its relationship with litter size. The main result indicates a positive effect of litter size in the probability of born dead of a particular piglet. The adequacy of the model was tested by the logarithm of the conditional predictive ordinate of data, showing that the model improve the predictive ability over the model that does not include the recursive relationship with litter size.

Keywords: Bayesian analysis, recursive models, piglet mortality, Markov Chain Monte Carlo