

Parámetros genéticos de la variabilidad del peso al nacimiento de conejos y ratones usando GSEVM v5.1

Milagros Garcia, Ingrid David, Christelle Robert-Granié, Hervé Garreau, Loys Bodin *

INRA, UR631 Station d'amélioration génétique des animaux, F-31320 Castanet-Tolosan, France

Resumen: El propósito de este artículo es explorar la estimación de los parámetros genéticos asociados con la variabilidad ambiental utilizando diferentes funciones escedásticas del programa GSEVM. El estudio trata de datos del peso al nacimiento recogidos en experimentos de selección de conejos y de ratones. Los resultados obtenidos en conejos muestran una mejor adaptación de modelos lineales sobre la varianza o desviación estándar de la varianza que de modelos no lineales. Los resultados conseguidos en ratones sugieren que ese tipo de análisis esta sensible a la estructura genética de los datos. En ese caso una transformación Box Cox previa para normalizar los datos no permite mejorar los resultados.

Palabras clave: canalización, variabilidad, estimación paramétrica, conejos, ratones, peso al nacimiento.

1 Introducción

Un programa dedicado a modelos heteroscedásticos que integran factores ambientales y genéticos en la función escedástica (modelización de la variabilidad ambiental) ha sido escrito sobre una base de (Ibañez-Escriche *et al.*, 2010) y se encuentra en el dominio público. Para chequear la validez de ese programa hemos analizado datos de varios caracteres y de distintas especies. Este artículo resume los principales resultados obtenidos en experimentos sobre la canalización del peso al nacimiento de los conejos y de los ratones.

2 Material y Método

El programa GSEVM v5.1

Después de una estandarización específica el modelo lineal clásico utilizado para estimar los parámetros y valores genéticos se puede escribir:

$$y_{ij} = \mu + u_i + \sigma_i \varepsilon_{ij}$$

En el cual

y_{ij} corresponde a la observación j del individuo i ,

μ es la media general de las observaciones y sin pérdida de generalidad puede representar también los distintos factores ambientales que afecta la media de las observaciones,

u_i es el valor genético del individuo i , procedente de una distribución $N(0, \sigma_u^2)$

σ_i es la desviación estándar de las observaciones de ese individuo

ε_{ij} es una variable aleatoria residual de distribución $N(0,1)$

En los modelos para analizar la canalización, la modelización de la variabilidad residual individual σ_i se hace mediante una función escedástica $f(\eta + v_i)$

$$y_{ij} = \mu + u_i + f(\eta + v_i)\varepsilon_{ij}$$

en la cual:

f es una función exponencial (EXP) en el modelo propuesto por (SanCristobal-Gaudy *et al.*, 1998); raíz cuadrada (SQR) en el modelo de (Hill *et al.*, 2007), lo que supone el ajuste de un modelo lineal sobre la varianza residual, e identidad (LIN) en el de (García *et al.*, 2009), lo que supone el ajuste de un modelo lineal sobre las desviación típica residual.

η es un término medio, y sin pérdida de generalidad puede representar también los distintos factores ambientales que afectan la variabilidad de las observaciones,

v_i es el valor genético del individuo i para la variabilidad, procedente de una distribución $N(0, \sigma_v^2)$

El programa utiliza un algoritmo MCMC con un Gibbs sampling para los efectos sistemáticos afectando la media, y un Metropolis Hasting para el resto de los efectos. Los parámetros del Metropolis provienen de una auto-parametrización de manera que el ratio de aceptación se encuentre alrededor del valor 0.23. Las distribuciones a priori asumidas para los diferentes parámetros son: para las varianzas de los efectos genéticos y no genéticos distribuciones Chi2 inversas y para los efectos sistemáticos μ y η así como para el correlación genética ρ entre u y v una distribución Uniforme.

Métodos alternativos

Con el objetivos de tener una aproximación de los valores de los parámetros en casos de resultados no fiables y corroborar los resultados, se ha utilizado un proceso en dos etapas realizadas con el programa ASREML, tal como lo ha presentado (Gutierrez *et al.*, 2006) o (Garreau *et al.*, 2008). La primera etapa consiste en ajustar un modelo lineal clásico sobre el valor de los datos. Mientras que en la segunda etapa se ajusta un modelo lineal sobre el logaritmo del cuadrado de los residuos de la primera etapa. La segunda etapa puede implicar un modelo bivariable que permite obtener estimaciones de la correlación genética entre u y v .

Transformación Box Cox

Para tener en cuenta la relación entre la disimetría de la distribución marginal de los datos según el modelo 2 y la correlación entre los efectos genéticos que afectan la media y la varianza, (Yang *et al.*, 2009) proponen de usar una transformación Box Cox de los datos. Esa transformación no esta implementada en la versión actual del programa GSEVM y fue aplicada previamente sobre los datos utilizando la función boxcox de R

Experimento de canalización en conejo

Los datos proceden de un experimento de selección divergente emprendido al INRA desde 2003, sobre la variabilidad intra camada del peso al nacimiento de los gazapos (Garreau *et al.*, 2008). En ese experimento los pesos al nacimiento fueron considerados como un carácter de la madre de tal manera que cada parto permitía estimar directamente su variabilidad. Los resultados presentados aquí corresponden a los datos acumulados durante 10 generaciones. Conciernen el peso al nacimiento de 20175 gazapos nacidos de 2176 camadas de 1078 hembras.

Análisis

Los parámetros genéticos para el peso al nacimiento medio y su variabilidad fueron estimados utilizando las tres funciones escedásticas implementada en el programa GSEVM (500000 iteraciones, burn-in=100000, lag=100). La duración de gestación, el estatuto fisiológico de la hembra pariendo y el tamaño de la camada fueron considerados como efectos ambientales en los modelos para la media y la variabilidad; un efecto permanente relacionado con la hembra entró también en la modelización de la media. La respuesta directa a la selección fue estimada a través de los cambios de valores genéticos según las generaciones. Las repuestas sobre el peso al nacimiento medio y la desviación típica intra camada fueron estudiadas comparando los valores fenotípicos de esas variables entre las líneas divergentes.

Resultados

El cuadro 1 da los valores de los parámetros estimados según las distintas funciones escedásticas. Las varianzas de los efectos genético y permanente que afectan la media son muy parecidas cualquier sea la función escedástica. Son también bastante cercanos de los valores estimadas en la primera etapa (resultados no mostrados); lo que significa que el grado global de heteroscedasticidad es ligero.

Las varianzas genéticas que controlan la variabilidad son evidentemente diferentes según las funciones escedásticas pero son siempre mucho más pequeñas que la varianza genética de la media. De manera interesante, el valor de la correlación entre la media y la variabilidad del peso al nacimiento es nula o muy cerca de cero. El valor del DIC indica que un modelo lineal sobre la varianza ambiental (SQR) es el modelo más pertinente.

	LIN $\eta_{ijk} + v_i$		EXP $e^{\frac{\eta_{ijk} + v_i}{2}}$		SQR $\sqrt{\eta_{ijk} + v_i}$	
σ_u^2	23.9	(2.9)	21.7	(2.8)	24.6	(1.3)
σ_p^2	35.22	(1.7)	36.23	(1.9)	36.23	(1.04)
σ_v^2	3.28	(0.3)	0.22	(2.2 10 ⁻²)	608.8	(1.3)
ρ	8.5 10 ⁻² (6.6 10 ⁻²)		7.7 10 ⁻² (8.0 10 ⁻²)		-8.8 10 ⁻⁵ (3.8 10 ⁻⁷)	
Dic	106421		106515		104140	

Cuadro 1 : Media (desviación estándar) de la distribución a posteriori de los parámetros estimados según las tres funciones escedásticas y Dic de los respectivos modelos.

Experimento de canalización en ratones

Los datos proceden del experimento conducido a la Facultad Veterinaria de Madrid y cuya originalidad es realizar las pruebas sobre descendencia de los machos sobre una línea consanguínea de hembras. El carácter elegido para esa selección divergente es la variabilidad del peso al nacimiento asignado al animal mismo y no a su madre. El diseño y los resultados de las primeras generaciones están descritos por (Cervantes *et al.*, 2010). Aquí presentamos análisis en las cuales el carácter se atribuye a la madre.

Análisis

Para las estimaciones se han considerado los mismos factores de variación y la misma función escedástica EXP que en (Cervantes *et al.*, 2010). Se analizaron también los datos después de una transformación Box Cox con la finalidad de normalizar los valores marginales. Finalmente, intentamos comprobar el rango de los resultados con el proceso en dos etapas.

Resultados

El cuadro 1 muestra los valores de los parámetros estimados según los distintos métodos.

	GSEVM $e^{\left(\frac{\eta_{jk}+v_i}{2}\right)}$	GSEVM (Box Cox)	ASREML
σ_u^2	4.89 10 ⁻² (3.4 10 ⁻²)	4.43 10 ⁻² (2.4 10 ⁻²)	0.22 10 ⁻²
σ_p^2	2.44 10 ⁻² (4.5 10 ⁻³)	1.95 10 ⁻² (4.0 10 ⁻³)	2.35 10 ⁻²
σ_v^2	0.21 (7.8 10 ⁻²)	0.20 (7.6 10 ⁻²)	0.19 10 ⁻¹
$\sigma_{\bar{p}}^2$	0.32 (4.8 10 ⁻²)	0.37 (5.2 10 ⁻²)	0.21
ρ	0.44 (0.43)	0.11 (0.53)	0.77
Dic	-10067	-10979	

Cuadro 2 : Media (desviación estándar) de la distribución a posteriori de los parámetros estimados según la función escedástica EXP sin o con transformación previa Box Cox de los datos y estimaciones de ASREML.

Sobre esos datos, los parámetros (varianza genética y efecto permanente) de la variabilidad del peso al nacimiento dados por GSEVM están muy elevados. Esos resultados son bastante diferentes de las dos etapas de ASREML. Sin embargo se ha visto que cuando los efectos sobre la variabilidad son importantes, el proceso de dos etapas da resultados muy sesgados (Garcia *et al.*, 2010). En particular, una subestimación para la varianza genética y una sobreestimación para la varianza no genética, en cuanto a la correlación genética entre u y v, este valor es siempre subestimado cuando se utiliza el modelo bivariado.

En ese caso, la transformación previa de los datos no cambia los parámetros, excepto la correlación entre los efectos genéticos que afectan la media y la variabilidad. Su distribución a posteriori tiene una media más próxima de cero, pero una variabilidad más elevada todavía.

3 Discusión y conclusion

Sobre los datos del experimento de selección en conejo, un modelo lineal sobre la varianza ambiental (SQR) o sobre la desviación estándar ambiental (LIN) parece adaptarse mejor que sobre el logaritmo de la varianza ambiental. La correlación entre los valores genéticos para la media y para la variabilidad es nula con cualquier método. Los datos de ratones dan resultados distintos. La correlación entre los dos valores genéticos es importante, aunque menos que cuando se atribuye el peso al nacimiento a cada animal ($\rho = -0.96$ resultados no mostrados). A parte de la diferencia entre especies, el contraste entre los resultados de esos dos experimentos puede provenir del origen genético de los efectos que controlan el peso al nacimiento. Se ha demostrado un componente maternal importante en la selección canalizante de los conejos (Bolet *et al.*, 2007) que no puede expresarse en el experimento de selección de los ratones que utiliza hembras homocigotes y por lo tanto permite seleccionar únicamente el efecto directo.

El programa GSEVM v5.1 permite analizar datos de canalización con una gran flexibilidad. Tres funciones escedásticas son disponibles y el valor del DIC puede ayudar a seleccionar la función más pertinente. Sobre esos datos, la transformación Box Cox previa de los marginales no permite mejorar la estimación de la correlación entre los valores genéticos para la media y para la variabilidad.

Agradecimientos

Los autores agradecen a Juan Pablo Gutiérrez, Isabel Cervantes, Blanca Nieto y Concha Salgado por el acceso a sus datos del experimento de selección en ratones.

Referencias bibliograficas

- Bolet G., Garreau H., Joly T., Theau-Clement M., Falières J., Hurtaud J., and Bodin L. 2007. Genetic homogenisation of birth weight in rabbits: Indirect selection response for uterine horn characteristics. *Livestock Science*. **111**(1-2): 28-32.
- Cervantes I., Nieto B., Salgado C., Pérez-Cabal M.A., Ibanez-Escriche N., and Gutierrez J.P. 2010. Resultados de las dos primeras generaciones de selección divergente para variabilidad del peso al nacimiento en ratones. *XV Reunión Nacional de Mejora Genética Animal*, Vigo, Spain.
- Garcia M., David I., Garreau H., Ibañez-Escriche N., Mallard J., Masson J.P., Pommeret D., Robert-Granié C., and Bodin L. 2009. Comparisons of three models for canalising selection or genetic robustness. *60th Annual Meeting of the European Association for Animal Production*, Barcelona, Spain.
- Garcia M., David I., Robert-Granié C., and Bodin L. 2010. Comparison of two steps versus single step estimation of genetic parameters for canalisation. *Submitted*.
- Garreau H., Bolet G., Larzul C., Robert-Granié C., Saleil G., SanCristobal M., and Bodin L. 2008. Results of four generations of a canalising selection for rabbit birth weight. *Livestock Science*. **119**(1-3): 55-62.

- Gutierrez J.P., Nieto B., Piqueras P., Ibañez N., and Salgado C. 2006. Genetic parameters for canalisation analysis of litter size and litter weight traits at birth in mice. *Genetics Selection Evolution*. **38**: 445-462.
- Hill W.G., Mulder H.A., and Zhang X.S. 2007. The quantitative genetics of phenotypic variation in animals. *Acta Agriculturae Scandinavica Section a-Animal Science*. **57**(4): 175-182.
- Ibañez-Escriche N., Garcia M., and Sorensen D. 2010. GSEVM v.2: MCMC software to analyze genetically structured environmental variance models. *Journal of Animal Breeding and Genetics. Short Communication*. (On line).
- SanCristobal-Gaudy M., Elsen J.M., Bodin L., and Chevalet C. 1998. Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. *Genetics Selection Evolution*. **30**(5): 423-451.
- Yang Y., Christensen O.F., and Sorensen D. 2009. Analysis of genetically structured variance heterogeneity and the Box-Cox transformation. *60th Annual Meeting of the European Association for Animal Production*, Barcelona, Spain. 24-27 august.