

Modelos recursivos para el análisis simultáneo de caracteres productivos y de supervivencia

J.C. Martínez Ávila¹ L.A. García Cortés^{1,2}

1. Instituto Nacional de Investigación y Tecnología Agraria (INIA). Departamento de Mejora Genética Animal. Ctra de la Coruña km 7.5 MADRID 28040. jcarlos@inia.es

2. Escuela Técnica Superior de Ingenieros Agrónomos (UPM). Departamento de Producción Animal Avda. Complutense s/n MADRID

INTRODUCCIÓN

La importancia de los caracteres de supervivencia, tales como la longevidad, se basa fundamentalmente en la relación inversa entre los costes y el tiempo de vida del animal.

Una evaluación genética conjunta es preferible cuando varios caracteres relacionados genéticamente están incluidos en los objetivos de selección. También es útil para evitar el sesgo debido a la selección de los otros caracteres (Damgaard y Korsgaard, 2006) y para mejorar la precisión de los valores mejorantes. Esto es especialmente relevante en caracteres con baja heredabilidad y poco informativos como suelen ser los caracteres de supervivencia (Tarrés, Piedrafita y Ducroq, 2006). El análisis de supervivencia normalmente utiliza una alta proporción de datos censurados que proporcionan poca información.

Analizar simultáneamente caracteres de supervivencia y lineales no es trivial dado el problema de definir la distribución conjunta. No está muy claro como la distribución de valor extremo que siguen los residuos de supervivencia pueden ser covariados con los residuos gaussianos de los caracteres lineales. Por ejemplo Tarrés (Tarrés, Piedrafita y Ducroq, 2006) siguiendo el método de los pseudoregistros, dividen la parte residual del carácter lineal en dos partes, una de ellas correlacionada con la distribución de valor extremo. Por otro lado Damgaard (Damgaard y Korsgaard, 2006) incluyen a un residuo gaussiano dentro de la distribución de Weibull pero obviando cualquier correlación residual con la distribución de valor extremo.

Una forma de expresar la relación entre dos caracteres es la propuesta recientemente por Sorensen y Gianola (Gianola y Sorensen, 2004), usando modelos recursivos donde una de las fuentes de variación viene dada por la inclusión del fenotipo de un carácter en la descripción del otro. Pero este tipo de modelos recursivos necesitan restricciones en los parámetros a estimar para asegurar la identificabilidad tal y como describen Varona y Sorensen (Varona y Sorensen, In press).

Los modelos multivariantes que incluyen caracteres de supervivencia y producción no son equivalentes cuando se asume recursividad, como en el caso bivariado con caracteres gaussianos. Considerando el carácter productivo como dependiente del fenotipo del carácter de supervivencia, este mantendrá su distribución gaussiana.

El objetivo del presente estudio es presentar un nuevo acercamiento al problema de la evaluación genética simultánea. Se usan modelos recursivos para analizar el caso bivariado producción-supervivencia. Se lleva a cabo un estudio de simulación para comprobar la bondad del método propuesto.

MATERIAL Y MÉTODOS

La verosimilitud en el modelo bivariado habitual $f(y, t|\theta)$ puede ser dividida en dos partes.

$$p(\theta|y, t) \propto f(y|t, \theta) f(t|\theta) p(\theta), \quad (1)$$

Con θ el vector de las variables a estimar y t e y , son los fenotipos de supervivencia y producción respectivamente.

Sin pérdida de generalidad se asume que t sigue una distribución de Weibull tal que:

$$t|\theta \sim Weibull(\rho, \lambda) \quad (2)$$

donde ρ y λ son el parámetro de forma y el parámetro de escala respectivamente. El parámetro de escala es explicado por medio de efectos ambientales y genéticos.

$$\lambda = \mathbf{X}\mathbf{b}_t + \mathbf{Z}\mathbf{u}_t, \quad (3)$$

El carácter productivo es descrito condicionalmente como:

$$\mathbf{y}|\mathbf{t}, \theta \sim N(\phi(\mathbf{t}) + \mathbf{X}\mathbf{b}_y + \mathbf{Z}\mathbf{u}_y + \mathbf{e}, \mathbf{I}\sigma_e^2), \quad (4)$$

donde la única diferencia con el modelo mixto es $\phi(t)$, que es una función recursiva tal que $\phi(t) = \beta_1 t + \beta_2 t^2$ permitiendo así una mayor flexibilidad entre el carácter productivo y de supervivencia.

Se simuló una población de 30.000 animales con 10 generaciones no solapadas y 3.000 animales por generación. Siete casos fueron analizados donde la única diferencia fue la correlación genética entre el carácter productivo y de supervivencia, y cuyos valores fueron, 0,9, 0,5, 0,3, 0, -0,3, -0,5 y -0,9. De acuerdo con (Ducroq, 2005) el parámetro de forma se tomó como $\rho = 2$.

Los valores verdaderos de las varianzas genéticas del carácter productivo g_{yy} y del carácter de supervivencia g_{tt} fueron 10 y 3, respectivamente.

Se implementó un muestro de Gibbs para hacer inferencias en los parámetros de interés, se usó también un *adaptive rejection sampling* y las cadenas obtenidas después del *burn-in* fueron analizadas con el programa BOA (Smith, 2005).

RESULTADOS Y DISCUSIÓN

Los siete casos propuestos para mostrar la influencia de la covarianza genética en el modelo propuesto están recogidos en el cuadro 1. La varianza genética aditiva g_{yy} muestra el mismo comportamiento a lo largo de los nueve escenarios propuestos y las estimas son muy similares, el valor estimado se encuentra dentro del intervalo de alta densidad posterior. Las medias posteriores de g_{tt} tienden al valor verdadero dentro de los intervalos de alta densidad posterior, pero su precisión esta claramente influida por la correlación genética. Esta precisión es mayor en los casos de mayor correlación genética debido a la información proveniente del carácter productivo

El cuadro 2 muestra la precisión de las distribuciones marginales posteriores de 30.000 valores genéticos aditivos para el carácter de supervivencia. Esta claro que los casos donde la correlación genética es mayor, la precisión en los criterios de selección para supervivencia aumenta. El caso con covarianza genética nula es equivalente a un análisis univariante y tomándolo como referencia se aprecia como con la inclusión del carácter productivo mejora la precisión de las estimas.

Cuadro 1: Media posterior(PM), desviación estandar posterior(PSD) e intervalo de alta densidad posterior(HPD), para los componentes de (co)varianza en los siete casos

CASO*	$g_{yy} \mathbf{y}, \mathbf{t}$			$g_{tt} \mathbf{y}, \mathbf{t}$			$g_{yt} \mathbf{y}, \mathbf{t}$		
	PM	PSD	HPD[95 %]	PM	PSD	HPD[95 %]	PM	PSD	HPD[95 %]
0,9(4,92)	9,73	0,22	(9,26, 10,21)	2,58	0,36	(2,04, 3,45)	4,72	0,26	(4,20, 5,22)
0,5(2,73)	9,75	0,25	(9,25, 10,24)	2,55	0,62	(1,43, 3,60)	2,73	0,34	(2,03, 3,36)
0,3(1,64)	9,74	0,24	(9,25, 10,24)	2,75	0,61	(1,55, 3,93)	1,65	0,37	(0,89, 2,35)
0(0)	9,73	0,25	(9,21, 10,21)	2,84	0,61	(2,03, 3,99)	-0,07	0,31	(-0,63, 0,62)
-0,3(-1,64)	9,72	0,25	(9,23, 10,22)	2,93	0,66	(1,62, 3,99)	-1,74	0,37	(-2,47, -1,00)
-0,5(-2,73)	9,71	0,25	(9,20, 10,19)	2,93	0,50	(1,97, 3,85)	-2,78	0,35	(-3,47, -2,09)
-0,9(-4,92)	9,68	0,23	(9,22, 10,20)	2,76	0,31	(2,25, 3,59)	-4,96	0,27	(-5,55, -4,39)

* correlación genética aditiva verdadera y su correspondiente covarianza entre paréntesis

Cuadro 2: Precisión en el componente genético aditivo de supervivencia: Varianza marginal(MV) promediada sobre la población total, varianza entre estimadores(VBP) y coeficiente de correlación de Pearson con los verdaderos valores aditivos

Caso	MV	VBP	Correlación
0,9	1,305	1,453	0,718
0,5	1,945	0,546	0,438
0,3	2,349	0,286	0,323
0,0*	2,703	0,167	0,241
-0,3	2,709	0,336	0,329
-0,5	2,255	0,619	0,443
-0,9	1,386	1,622	0,719

* Equivalente a un análisis unicaracter de supervivencia

El modelo recursivo presentado, lo es en una vía, pero pudiera serlo en dos vías. En general los modelos recursivos pueden ser usados para describir situaciones donde una o mas distribuciones no gaussianas están implicadas.

Independientemente de las similitudes con otras propuestas al mismo problema, el objetivo de este trabajo no fue mejorar la eficiencia de dichos métodos. Creemos que el presente trabajo es una respuesta elegante y flexible al problema de la evaluación genética simultanea de caracteres de supervivencia y producción.

Los códigos escritos en Fortran, que ademas de los presentado aqui, incluyen datos censurados de supervivencia y estimación de ρ estan disponibles solicitandolos a los autores

REFERENCIAS BIBLIOGRÁFICAS

- Damgaard L.H., Korsgaard I.R., Gen. Sel. Evol. (2006) 38.45-64
 Ducrocq V., Anim Sci. (2005) 80.249-246
 Gianola D., Sorensen D., Genetics (2004) 167.1407-1424
 Tarres J., Piedrafita J., Ducrocq V., Genet. Sel. Evol. (2006) 38.65-83
 Smith B.J., University of Iowa. (2005) Iowa City
 Varona L., Sorensen D., Genetics (En prensa)