Variabilidad genética y selección de la prolificidad en porcino

Leopoldo Alfonso Ruiz

I S B N: 84-89727-64-3
Depósito Legal: S. 54-98

Servei de Publicacions
Universitat de Lleida

TESITEX, S.L.
c/. Melchor Cano, 15
Télf. 923 - 25 51 15
Fax 923 - 25 87 03
37007 SALAMANCA

ÍNDICE GENERAL

- 1.- INTRODUCCIÓN GENERAL
- 2.- OBJETIVOS
- 3.- ESTIMACIÓN DE COMPONENTES DE VARIANZA MEDIANTE UN MODELO CON REPETIBILIDAD
    - 3.1.- INTRODUCCIÓN
    - 3.2.- MATERIAL Y MÉTODOS
        - 3.2.1.- Material.
            - 3.2.1.1.- Material animal.
                - 3.2.1.2.- Poblaciones simuladas.
                - 3.2.1.2.1.- Descripción general.
                - 3.2.1.2.2.- Poblaciones simuladas con selección en la población base.
        - 3.2.2.- Métodos.
            - 3.2.2.1.- Descripción general.
            - 3.2.2.2.- Modelo utilizado.
            - 3.2.2.3.- Análisis de los datos de campo.
                - 3.2.2.3.1.- Estimación de componentes de varianza.
                - 3.2.2.3.2.- Comparación de conjuntos de datos.
                - 3.2.2.3.3.- Análisis de la selección de los animales de la población base.
            - 3.2.2.4.- Análisis de las poblaciones simuladas.
    - 3.3.- RESULTADOS Y DISCUSIÓN.
        - 3.3.1.- Estimación de la variabilidad genética mediante un modelo con repetibilidad.
        - 3.3.2.- Análisis de la existencia de selección en la información de los animales base.
            - 3.3.2.1.- Efecto de la selección de los datos de la población base por truncamiento.
            - 3.3.2.2.- Comparación de conjuntos de datos de las poblaciones analizadas.
            - 3.3.2.3.- Efecto de la reducción de la variabilidad genética en los animales de la población base.
            - 3.3.2.4.- Asunción de distintos grados de selección de los animales base de las poblaciones analizadas.
- 4.- ANÁLISIS DE LA EXISTENCIA DE EFECTOS MATERNOS
    - 4.1.- INTRODUCCIÓN
    - 4.2.- MATERIAL Y MÉTODOS
        - 4.2.1.- Material.
            - 4.2.1.1.- Material animal.
            - 4.2.1.2.- Poblaciones simuladas.
        - 4.2.2.- Métodos.
            - 4.2.2.1.- Modelos utilizados.
            - 4.2.2.2.- Contraste entre modelos.
    - 4.3.- RESULTADOS Y DISCUSIÓN
        - 4.3.1.- Estimaciones en los datos de granja.

- 4.3.2.- Estimaciones en poblaciones simuladas.

- 5.- ANÁLISIS DE LA EXISTENCIA DE HETEROGENEIDAD GENÉTICA ENTRE PARTOS
  - 5.1.- INTRODUCCIÓN
  - 5.2.- MATERIAL Y MÉTODOS
    - 5.2.1.- Material animal.
    - 5.2.2.- Métodos.
  - 5.3.- RESULTADOS Y DISCUSIÓN
    - 5.3.1.- Análisis unicarácter.
    - 5.3.2.- Análisis multicarácter.
- 6.- IMPLICACIONES DE LA HETEROGENEIDAD GENÉTICA ENTRE PARTOS
  - 6.1.- INTRODUCCIÓN
  - 6.2.- MATERIAL Y MÉTODOS
    - 6.2.1.- Material.
      - 6.2.1.1.- Material animal.
      - 6.2.1.2.- Poblaciones simuladas.
    - 6.2.2.- Métodos.
      - 6.2.2.1.- Revisión de la hipótesis de identidad genética entre partos mediante índices de selección con información familiar.
      - 6.2.2.2.- Simulación de un proceso de selección hiperprolífica en la población A.
      - 6.2.2.3.- Simulación de un proceso de selección hiperprolífica en una población simulada con la misma estructura que la población A.
        - 6.2.2.3.1.- Simulación del proceso de selección.
        - 6.2.2.3.2.- Decisión del modelo de evaluación.
  - 6.3.- RESULTADOS Y DISCUSIÓN
    - 6.3.1.- Revisión de la hipótesis de identidad genética entre partos mediante índices de selección con información familiar.
    - 6.3.2.- Simulación de un proceso de selección hiperprolífica en la población A.
    - 6.3.3.- Simulación de un proceso de selección hiperprolífica en una población simulada con la misma estructura que la población A.
- 7.- DISCUSIÓN GENERAL
- 8.- CONCLUSIONES
- 9.- BIBLIOGRAFÍA
- ANEJO

Agradecimientos:

A todos aquellos que de una u otra forma han hecho posible la realización de este trabajo.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

RESUMEN

A partir de la información de dos poblaciones porcinas (A y B), de raza Landrace y especialización materna, se han estimado los componentes de varianza y covarianza de la prolificidad, tanto número de lechones nacidos vivos como totales, bajo distintos supuestos genéticos. El objetivo fue determinar la existencia de efectos maternos, la existencia de heterogeneidad genética entre partos de una misma cerda, y el interés de seguir utilizando el modelo con repetibilidad habitualmente empleado en la selección por prolificidad de esas poblaciones, haciendo especial referencia al esquema de selección hiperprolífica realizado en la población A (Noguera *et al.*, 1994).

Inicialmente se consideró un modelo animal con repetibilidad. En ambas poblaciones los valores estimados de heredabilidad oscilaron entre 0.05 y 0.07, y los de repetibilidad entre 0.13 y 0.15. Los valores de heredabilidad fueron más bajos que los habitualmente referidos, pero pueden estar infraestimados por desconocer una parte importante de las relaciones de parentesco entre los animales. También se ha analizado la existencia de selección en la información disponible y su efecto sobre las estimaciones de heredabilidad, viendo que en la población A la información de prolificidad de los animales más antiguos era una muestra seleccionada que condujo a una importante infraestimación (50%). Esa infraestimación se evitó al despreciar la información seleccionada, obteniendo un conjunto de datos que presentaba mejores propiedades para seleccionar por prolificidad esa población. Mediante simulación se concluyó que una selección de las observaciones de los animales base puede producir infraestimación de la heredabilidad cuando se estima mediante un modelo mixto que agrupa la información seleccionada dentro de un mismo nivel de un factor fijo. Este resultado permitió explicar el bajo valor estimado en la población A al considerar toda la información disponible, dado que en el modelo de estimación consideraba un efecto fijo granja-año-estación que agrupaba la información más antigua en unos mismos niveles. Por otra parte, se revisó como modelizar, en el modelo animal, la existencia de selección en la población base para obtener estimaciones de heredabilidad no sesgadas por la presencia de esa selección.

En segundo lugar se consideró la existencia de efectos maternos sobre la prolificidad. En ninguna de las dos poblaciones se estimaron efectos maternos significativos, pese a haber desconsiderado la posible existencia de un efecto ambiental de camada de lactación. En poblaciones simuladas se comprobó que ignorar la existencia de ese efecto ambiental cuando realmente existe implica sobrestimar los efectos maternos de origen genético. Esos resultados indicaron la conveniencia de considerar un modelo sin efectos maternos en la selección por prolificidad de las poblaciones A y B, posiblemente debido a una suficiente estandarización del tamaño de camada durante el periodo de lactación.

Los componentes de (co)varianza también se estimaron bajo la hipótesis de heterogeneidad genética entre la prolificidad de distintos partos de una misma hembra. Para ello se realizó toda una serie de análisis bicarácter para los cinco primeros partos, cuyos resultados mostraron que las heredabilidades estimadas no eran distintas entre sí, y tampoco de las estimadas bajo el modelo de repetibilidad. Respecto a las correlaciones genéticas entre partos, exceptuando la correspondiente al primer-tercer parto en la población B, no resultaron distintas de la unidad. Pese a que los resultados de ese conjunto de análisis bicarácter pueden estar afectados por la existencia de selección de las cerdas en base a la información de partos previos, ofrecieron más credibilidad que los obtenidos mediante un análisis multicarácter de todos los partos, análisis que, por otra parte, sólo permitió obtener soluciones en la población A.

Finalmente, se analizó la repercusión de seleccionar mediante un modelo con

repetibilidad cuando existe heterogeneidad genética entre la prolificidad de los cinco primeros partos. Considerar el modelo de repetibilidad condujo a pérdidas de hasta un 7% al seleccionar utilizando índices con información familiar, pérdidas que justificarían el uso de un modelo multicarácter. No obstante, al simular el proceso de selección hiperprolífica que se está realizando en la población A, se comprobó que la pérdida de respuesta esperada podía ser mayor al utilizar inapropiadamente el modelo multicarácter que el de repetibilidad. En este sentido, utilizando elementos de la teoría de la decisión y simulando el mismo proceso de selección realizado en poblaciones también simuladas con la estructura de la población A, se concluyó que el modelo por el que en general se optaría para realizar un proceso de selección hiperprolífica era el modelo con repetibilidad. La simulaciones realizadas indicaron que la respuesta genética esperada únicamente vía hembra es de casi 0.35 lechones por parto.

Los resultados alcanzados mostraron que el modelo con repetibilidad, que se ha venido utilizando durante los tres últimos años para seleccionar por prolificidad las dos poblaciones estudiadas, parece más adecuado que un modelo que asuma heterogeneidad genética entre partos o la existencia de efectos maternos. Por esta razón, la respuesta genética esperada en el experimento de selección hiperprolífica que se está llevando a cabo en la población A, es poco probable que se vea afectada por haber desconsiderado ambas hipótesis genéticas.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

SUMMARY

Variance and covariance components of prolificacy (number of piglets born alive and total number of piglets born) were estimated under different genetic hypotheses in two maternal Landrace pig populations (A and B). The objective was to study the existence of maternal effects, the existence of genetic heterogeneity between parities, and the interest of using the repeatability model usually employed to select prolificacy of these two populations, with specially reference of hyperprolific scheme carried out in population A (Noguera *et al.*, 1994).

A repeatability animal model was initially considered. Estimates of heritability ranged between 0.05 and 0.07, and estimates of repeatability between 0.13 and 0.15, in both populations. These values of heritability are smaller than the values usually referred in the literature, but could be underestimated because an important number of relationships was unknown. The existence of selected information and its effect over the heritability estimations were also analyzed; the information of older animals of population A was a selected sample that implied a relevant underestimation of heritability (50%). This underestimation could be avoided by ignoring the selected information, which results in a data set with better properties to select animals of this population. In simulated populations it was concluded that a phenotypic selection of data of base animals could produce an underestimation of heritability using a mixed model with all selected records grouped into a level of a fixed factor. This result allowed us to explain the value obtained in population A when all known information was used, because the estimation model included a fixed effect farm-year-season that grouped older information into the same levels. Moreover, it was revised how to consider selection of base population under the animal model in order to obtain heritability estimations not biased by this selection.

The existence of maternal effects over litter size was also studied. Despite of ignoring the possible existence of a common litter environment, no population presented significant maternal effects. The genetic component of maternal effects was overestimated when the environment component was ignored in populations simulated. These results showed the interest of considering a model without maternal effects to select for prolificacy populations A and B, maybe because litter size standardization during suckling period is sufficient.

After, (co)variance components were estimated assuming the hypothesis of genetic heterogeneity between parities of a sow. A set of bivariate analyses was realized for the first five parities, which showed that heritabilities in each parity were not different from each other, and neither of the value estimated with a repeatability model. Genetic correlations were not different of unity, except the value of first-third parities of population B. Although these results could be affected by the existence of sows selection in base of previous parities, they were more credible than those obtained in a general multivariate analysis of all parities; in addition, this analysis produced solutions just in population A.

Finally, implications of the existence of genetic heterogeneity between parities when a repeatability model is used were analyzed. The use of the repeatability model produced loss of response up to 7% when a family index information was considered, which could justify the use of a multivariate model. However, the simulation of a hyperprolific scheme, as the applied in population A, showed that loss of expected response could be bigger when using an unsuitable multivariate model than when using a repeatability model. In this sense, using decision theory elements and simulating the same hyperprolific scheme, it was concluded that in general the model that should be chosen to carry out selection is a repeatability model. In this scheme simulation showed that expected response only by selection of females is near of

0.35 piglets/litter.

In conclusion, the repeatability model used during the last three years to select for prolificacy the two populations analyzed, is more appropriate than a model assuming genetic heterogeneity between parities and(or) the existence of maternal effects. It is hardly possible that the expected genetic response of the hyperprolific scheme carried out in population A could be affected by the assumption of a repeatability model without maternal effects.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

1.- Introducción general

La eficiencia en la producción de carne magra es el principal objetivo de la mejora genética porcina (Brascamp y De Vries, 1992). De entre los caracteres que determinan la eficiencia de esa producción - la velocidad de crecimiento, el porcentaje de magro de la canal, la eficacia en la conversión del alimento y el número de lechones criados por cerda y año - hasta hace pocos años sólo los tres primeros formaban parte del objetivo de selección en la mayoría de las poblaciones europeas (Webb y Bampton, 1989). No obstante, la situación ha ido cambiando; el número de lechones criados por cerda y año ha ido adquiriendo mayor importancia en la rentabilidad de las explotaciones porcinas y, para determinados productos, se han ido alcanzado niveles óptimos de magro en las canales (Hill y Webb, 1982).

Una medida del número de lechones criados por cerda y año es la productividad numérica (Legault, 1978), entre cuyos componentes, el criterio de selección que permite alcanzar un mayor progreso genético es la prolificidad (Bichard y David, 1985; Bolet y Legault, 1982; Gianola, 1988; Legault, 1987). Pese a que el objetivo de selección puede variar si se establece para un programa colectivo que integra el estrato comercial de animales o para programas individuales que deben competir en un mercado libre de venta de reproductores (Brascamp y De Vries, 1992; De Vries y Sorensen, 1990), la mejora de la prolificidad aumenta la rentabilidad económica tanto en el estrato comercial (Skjervold, 1979) como en el propio núcleo de selección (Houska *et al.*, 1994). Por otra parte, aunque los programas individuales de mejora puedan ser reticentes a aceptarlo por razones económicas, la selección en base a líneas especializadas es, desde un punto de vista global, más eficiente (De Vries y Sorensen, 1990) o al menos nunca menos eficiente (Webb y Bampton, 1989) que la selección en base a una única línea general. Así, el interés que se ha mostrado por esta especialización, utilizando líneas paternas fuertemente seleccionadas por el contenido magro de las canales (Webb y Bampton, 1989), ha potenciado el interés de considerar la prolificidad dentro del objetivo de selección de las líneas maternas (Bichard y Seidel, 1982).

El progreso genético esperado en prolificidad es relativamente discreto aplicando los métodos tradicionales, dado que la heredabilidad de este carácter es baja: 0.09 y 0.07 para el número de lechones nacidos vivos por parto, según las revisiones de Haley *et al.* , (1988) y Lamberson (1990), respectivamente. No obstante, se ha demostrado teóricamente (Avalos, 1985; Avalos y Smith, 1987) que mediante la aplicación de métodos de selección que permiten aumentar la precisión con que se selecciona, la mejora puede ser rápida y efectiva. En realidad, la inclusión de un carácter en el objetivo de selección no depende sólo de su importancia económica, sino también de la precisión con que se puede seleccionar (De Vries y Sorensen, 1990), de modo que la disponibilidad de metodologías de selección más precisas (Bampton, 1992) ha favorecido el interés de seleccionar por prolificidad.

En definitiva y por las razones apuntadas, gran parte de las organizaciones y empresas de mejora porcina seleccionan actualmente líneas maternas considerando la prolificidad dentro del objetivo de selección (Bampton, 1992; Ducos *et al.*,1992; Estany *et al.*, 1993; Frey *et al.*, 1994; Knap *et al.*, 1993; Luján, 1994; Sehested y Ianssen, 1993).

Los primeros intentos realizados para mejorar la prolificidad no parecían avalar el interés teórico de su selección. De los escasos experimentos de selección directa dentro de rebaño realizadas en porcino, en pocas se había obtenido respuesta. Posiblemente la más representativa es la referida por Ollivier y Bolet (1981): tras 10 generaciones de selección, la respuesta no fue significativa, fundamentalmente debido a la imposibilidad de alcanzar la intensidad de selección prevista. Los resultados publicados posteriormente (Bolet *et al.*, 1989) sobre 11 generaciones de selección, ratifican esos resultados. Vangen (1981) tampoco obtuvo respuesta tras 6 generaciones de selección en dos poblaciones noruegas.

Estos experimentos pusieron en evidencia la necesidad de conseguir una mayor intensidad de selección para obtener respuestas significativas, pese a que pueda haber otras causas responsables de no alcanzar las expectativas de respuesta iniciales: desde el tamaño finito de las poblaciones (Wang *et al.*, 1994) hasta la inadecuación de los modelos empleados (Bichard y David, 1985; Haley *et al.*, 1988; Bidanel y Ducos, 1994).

Legault y Gruand (1976) propusieron la creación de una línea hiperprolífica mediante la aplicación de una elevada intensidad de selección sobre las madres de los verracos, para posteriormente retrocruzar a las hijas con una nueva generación de verracos seleccionados. Mediante un esquema clásico de selección basado en un 33% de hembras seleccionadas sobre dos partos registrados, el progreso genético esperado es de 0.25 lechones por camada (Ollivier, 1982), mientras que en un esquema hiperprolífico basado en un máximo de un 2% de hembras seleccionadas sobre la base de 4 camadas registradas, se lograría una respuesta de 1.15 lechones por camada (Haley *et al.*, 1988). Además, si se forma un núcleo hiperprolífico, seleccionando tanto machos como hembras, se evita la necesidad de retrocruzar (Avalos, 1985), permitiendo obtener progresos superiores incluso con presiones de selección más débiles (Bichard y Seidel, 1982).

Los resultados alcanzados mediante una línea hiperprolífica creada en base a la población francesa de Large White, fueron positivos (Legault, 1981; Legault, 1987; Le Roy *et al.*, 1987). Recientemente, Bidanel *et al.* (1994) estiman en 0.7 lechones nacidos totales por parto la superioridad de las hijas de los verracos hiperprolíficos frente a las cerdas contempóraneas de la población francesa.

Experimentos posteriores también revelan la posibilidad de obtener respuesta en prolificidad. Lamberson *et al.*, (1991) obtienen una respuesta de más de un lechón en el primer parto de una línea previamente seleccionada por tasa de ovulación, y que tras varias generaciones de reposición al azar, se sometió 9 generaciones de selección directa por prolificidad. Sorensen y Vernessen (1991) llevan a cabo en Dinamarca un experimento de selección hiperprolífica en el que, a diferencia del experimento francés, se seleccionaron tanto machos como hembras para fundar un núcleo experimental. Aunque no se han publicado los resultados definitivos, los primeros resultados indican que la respuesta obtenida, pese a ser menor a la esperada, era positiva.

En la actualidad, otro experimento de selección hiperprolífica se está llevando a cabo en España en una línea materna Landrace (Noguera *et al.*, 1994). Pese a que el experimento aún no ha finalizado se observa, por el momento, una importante respuesta positiva en los animales seleccionados (J.L.Noguera comunicación personal).

En esos dos últimos experimentos, los animales se escogieron según las predicciones BLUP (*Best Linear Unbiased Prediction*; Henderson, 1984) de su valor genético, calculadas a través de un modelo animal con repetibilidad. Esas predicciones se realizaron considerando la prolificidad independiente del resto de caracteres que formaban parte del objetivo de mejora (Sorensen y Vernessen, 1991; Estany *et al.*, 1993), puesto que los caracteres productivos, de composición corporal y de calidad de carne presentan, en general, débiles correlaciones genéticas con el tamaño de la camada (Legault, 1970; Hill y Webb, 1982; Bolet *et al.*, 1984; Rydhmer, 1993; Short *et al.*, 1994).

El uso de esa metodología, en general referida como metodología del modelo mixto, se ha extendido notablemente en porcino desde su primera aplicación en caracteres productivos en Canadá (Hudson y Kennedy, 1985), utilizándose actualmente en la mayoría de los programas porcinos de selección (Brandt, 1992; Ducos *et al.*, 1992; Luján, 1994).

La metodología del modelo mixto permite obtener predicciones, insesgadas y de varianza mínima (Henderson, 1984), del valor genético de los animales candidatos a la

selección, pero para ello el valor considerado de las varianzas y covarianzas genéticas ha de coincidir con su valor verdadero en el conjunto de los animales base de la población a la que pertenecen. El conocimiento de esos componentes de (co)varianza no sólo es necesario para construir predictores del valor genético de los animales candidatos a ser seleccionados, sino también para predecir las consecuencias de la selección, diseñar experiencias eficientes y estimar diferencias entre poblaciones (Gianola, 1989). Dado que los valores de esos componentes en la población base no son conocido, debemos forzosamente estimarlos a partir de las observaciones fenotípicas disponibles, para poder así sustituirlos por sus estimadores. Desgraciadamente, en muy pocos casos esas observaciones proceden de un diseño experimental (Hill, 1979), sino de granjas comerciales, y por tanto son datos desequilibrados, recogidos en distintos ambientes y probablemente seleccionados.

Para datos desequilibrados y con un gran número de niveles de efectos fijos, el método REML (*Restricted,* o *residual, Maximum Likelihood* ; Patterson y Thompson, 1971) es el que presenta mejores propiedades desde un punto de vista de la estadística clásica (Robinson, 1991), incluso cuando los datos han sufrido ciertos tipos de selección (Thompson, 1973). No obstante, el método REML no siempre puede evitar hallar estimas de los parámetros genéticos fuertemente desviadas de los valores verdaderos, en particular cuando los animales de la población base no son una muestra aleatoria de los que, conceptualmente, forman la población base de referencia (Van der Werf y Boer, 1990; Van der Werf, 1992; Van der Werf y Thompson, 1992).

Prácticamente en todos los experimentos y programas de selección en que se considera la prolificidad, el modelo utilizado es un modelo animal con repetibilidad (Quaas *et al.*, 1984). Desde un punto de vista genético es un modelo muy sencillo: se considera una única variable como responsable de la variabilidad genética existente entre todas las observaciones del carácter. Es decir, ha sido habitual ignorar la existencia de efectos maternos sobre la prolificidad, y considerar que todos los partos de una cerda tienen la misma determinación genética (Ducos *et al.*, 1992; Bidanel y Ducos, 1994).

Son varios los trabajos en donde se estudia la influencia de los efectos maternos sobre el tamaño de camada en porcino, pero los resultados no siempre han sido coincidentes (Nelson y Robinson, 1976; Alsing *et al.*, 1980; Van der Steen, 1983; Rutledge, 1980a; Kirkpatrick y Rutledge, 1988; Stewart y Diekman, 1989; Southwood y Kennedy, 1990). En general se acepta que puedan estar negativamente correlacionados con la prolificidad, estimándose que su influencia conduce a una reducción de la respuesta a la selección de entre un 5 y un 10% de su posible rango de valores (Van der Steen, 1983). Aunque es difícil concluir si los efectos maternos son económicamente tan importantes como para ser considerados (Bichard y David, 1985), el desarrollo de un programa de selección eficiente para la productividad de las cerdas ha de considerar su posible existencia (Robinson, 1981; Southwood y Kennedy, 1990).

La asunción de homogeneidad genética entre partos implica considerar que las correlaciones genéticas entre ellos son iguales a la unidad y que la heredabilidad es constante para todos los partos. No obstante, existen evidencias, tanto en porcino como en otras especies polítocas, de que no siempre se cumplen ambas consideraciones (Tartar, 1981; Le Roy *et al.*, 1987; Bolet *et al.*, 1984; Avalos, 1985; Haley *et al.*, 1988; Knap *et al.*, 1993; Rydhmer, 1993; Irgang *et al.*, 1994; Gómez, 1994). Si realmente la prolificidad de los sucesivos partos producidos por una hembra es la expresión de distintos caracteres, tenerlo en cuenta en el proceso de selección aumenta su eficacia y evita sobrestimar la respuesta genética esperada (Ollivier y Bolet, 1981; Tartar y Bolet, 1984; Gómez, 1994), pero hace necesario estimar las correlaciones genéticas entre partos y los coeficientes de ponderación

económica que forman parte del objetivo global a mejorar.

Aun existiendo evidencias de heterogeneidad genética entre partos, la decisión de emplear un modelo genético más complejo que el habitualmente considerado no es fácil de tomar. Así, por ejemplo, errores en la asignación de las correlaciones genéticas pueden producir varianzas de los errores de predicción del valor genético de los animales superiores a las de un modelo simplificado que las considere iguales a la unidad (Schaeffer, 1984). Adicionalmente, errores superiores a un 50% en la estimación de los coeficientes de ponderación económica podrían conducir a pérdidas significativas de respuesta teórica a la selección (Vandepitte y Hazel,1977). La teoría de la decisión (p.ej. Romero, 1993), pese a haber sido poco utilizada en mejora animal (Meuwissen, 1991; Wyler *et al.*, 1994), ofrece herramientas interesantes para tomar decisiones entre distintas alternativas, pues permitir considerar tanto aspectos objetivos como subjetivos de la decisión.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

2.- Objetivos

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

En la Introducción se ha revisado el interés de seleccionar por prolificidad poblaciones porcinas de especialización materna, la necesidad de estimar los componentes de varianza de la prolificidad para llevar a cabo un programa de selección, y la posibilidad de que se incumplan ciertas asunciones habitualmente realizadas en la modelización de este carácter. Por ese motivo, centrándonos en el análisis de la prolificidad de dos poblaciones porcinas Landrace de especialización materna, se han establecido los siguientes objetivos:

1º.- Estimar los componentes de varianza de la prolificidad mediante un modelo con repetibilidad y sin efectos maternos, y discutir el efecto de la selección de la población base sobre esas estimaciones.

2º.- Analizar la existencia de efectos maternos sobre la prolificidad.

3º.- Analizar la existencia de heterogeneidad genética entre la prolificidad de distintos partos de una misma cerda.

4º.- Evaluar las implicaciones de la existencia de heterogeneidad genética entre partos en la evaluación genética de los animales candidatos a ser padres, y decidir, haciendo referencia a un experimento de selección hiperprolífica, qué modelo de evaluación genética utilizar.

Con el objeto de facilitar la presentación del trabajo, se tratarán por separado cada uno de los cuatro objetivos.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

3.- Estimación de componentes de varianza mediante un modelo con repetibilidad

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

3.1.- Introducción

En la Introducción General se comentó que para obtener un predictor del valor genético de los animales candidatos a la selección, insesgado y de mínima varianza, se debía conocer el valor de los parámetros genéticos en su población base. También se comentó que conocer esos parámetros implica estimarlos a partir de observaciones habitualmente desequilibradas, recogidas en distintos ambientes, y seleccionadas.

Se han dedicado importantes esfuerzos para poder disponer de métodos de estimación de componentes de varianza que, adaptándose a las características de los datos, ofreciesen estimadores con propiedades deseables[1]. Fundamentalmente interesa disponer de estimadores suficientes, que sean aplicables a modelos mixtos desequilibrados, y tengan en cuenta la selección de los datos. De los distintos métodos de estimación de componentes de varianza desarrollados en estadística clásica - ANOVA[2] y derivado[3]; método de máxima verosimilitud (ML[4]; métodos de norma mínima (MINQUE) o varianza mínima (MIVQUE)[5]; método de máxima verosimilitud restringida (REML)[6]-, el método REML es el más utilizado en la actualidad en mejora animal. Eso es así porque los estimadores REML son aplicables a modelos mixtos desequilibrados, permiten utilizar fácilmente toda la información que contienen los datos, no subestiman la varianza residual y, como veremos más adelante, no están afectados por la existencia de determinados tipos de selección o, si lo están, lo están menos que otros estimadores (Meyer, 1990). Sin embargo, hay que tener en cuenta que los estimadores basados en la verosimilitud sólo son asintóticamente insesgados, y que además

---

[1] Insesgados; suficientes; con expresión explícita de la varianza de muestreo; de varianza mínima; no afectados por el desequilibrio de los datos; con estimas dentro del espacio paramétrico; aplicables a cualquier tipo de modelo (fijo, aleatorio, mixto, con interacciones,...); sin necesidad de asunción de distribución de muestreo; no afectados por la selección de los datos; etc... Existen numerosos trabajos que revisan los distintos métodos de estimación de componentes de varianza existentes, por ejemplo, los de Searle (1989), Baselga y Carabaño (1993), Foulley (1990), San Cristóbal (1992), García-Cortés (1992), Searle et al. (1992), etc...

[2] En el caso de datos equilibrados y bajo un modelo aleatorio, ofrece un estimador óptimo (insesgado y de varianza mínima) pero no ocurre así en el caso de datos desequilibrados (Searle, 1989).

[3] Métodos Henderson I, II y III, que permiten analizar datos desequilibrados mediante un modelo aleatorio, un modelo mixto sin interacciones entre efectos fijos y aleatorios, y todo tipo de modelo mixto respectivamente. No obstante, entre otras limitaciones, presentan la de no ser adecuados en presencia de selección (Searle, 1989).

[4] Basado en maximizar la verosimilitud de la muestra, ofrece estimadores con propiedades más útiles que los mínimo cuadráticos: consistencia, normalidad asintótica y matriz de dispersión muestral asintótica conocida. Sin embargo, existe una pérdida de grados de libertad, asociada a la estimación de los efectos fijos, que conduce a subestimar la varianza del error (Searle et al., 1992).

[5] Presentan propiedades óptimas aún en el caso de datos desequilibrados, pero si las poblaciones han sido seleccionadas y los valores de las componentes de varianza considerados a priori, en el proceso iterativo de resolución, no son las verdaderas, los valores que se obtienen están sesgados (Sorensen y Kennedy, 1984a).

[6] Se caracteriza por trabajar con funciones residuales de los datos libres de los efectos fijos, y por tanto, a diferencia del método ML, no subestima la varianza residual (Searle, 1989).

los REML generalmente pierden esa propiedad por restringir su espacio paramétrico (Searle, 1989; Foulley, 1993).

Existen otros inconvenientes al utilizar estimadores REML. Excepto en algunos casos equilibrados, no se pueden obtener analíticamente, dado que el sistema de ecuaciones a resolver no es lineal respecto a los componentes de varianza a estimar. Ello obliga a utilizar métodos iterativos basados en localizar el máximo de la función de verosimilitud (Meyer, 1989), es decir, a usar algoritmos de maximización. Es necesario diferenciar entre estimadores y soluciones REML, puesto que aunque los estimadores REML fuesen insesgados, las soluciones obtenidas mediante un método iterativo no necesariamente lo serían (Searle, 1989). Se consideran soluciones REML las obtenidas al alcanzar la convergencencia en un proceso iterativo; en determinadas situaciones un método iterativo difícilmente puede garantizar la convergencia en el máximo global de la función de verosimilitud, siendo posible hallar varias soluciones REML para un mismo conjunto de datos (Groeneveld y Kovac, 1990).

Hay, entre otros, dos importantes algoritmos iterativos de cálculo basados en los elementos de las ecuaciones del modelo mixto de Henderson (Henderson, 1984): un algoritmo de desarrollo teórico, el EM (*Expectation and Maximization*, Dempster *et al.*(1977)), y un algoritmo de desarrollo empírico, el DF (*Derivative Free*, Graser *et al.* (1987)). El algoritmo DF presenta la ventaja de ser mucho menos demandante computacionalmente pero mucho más oscuro en cuanto a conocer los verdaderos mecanismos que determinan las estimaciones (Thompson y Atkins, 1990). El desarrollo de este algoritmo (Meyer, 1989), su implementación informática aprovechando distintas estrategias para aumentar su eficiencia (Boldman y Van Vleck, 1991), y el importante desarrollo informático experimentado en la última década, han permitido el uso generalizado de la metodología REML en mejora animal.

Hasta ese momento, las estimaciones de heredabilidad de la prolificidad obtenidas en porcino mediante otros métodos indicaban como valores medios 0.10 para el número de nacidos totales por parto y 0.08 para el número de nacidos vivos por parto[7]. No obstante, se argumentaba que algunas estimas podían estar afectadas por criterios como la elección de las cerdas según registros tempranos del carácter (Haley *et al.*, 1988). Vangen (1986) halló valores superiores de heredabilidad cuando sólo consideraba los datos de las hembras que habían alcanzado un determinado parto. Ese resultado indujo a Haley *et al.* (1988) a sugerir que la selección de las cerdas en base a la información de partos previos posiblemente sobrestimase el valor de la heredabilidad, si bien indicaban lo incierto de la dirección de esos sesgos.

La base teórica para conocer el efecto de la selección sobre los parámetros genéticos parte de considerar que, en una población de tamaño infinito y bajo un modelo aditivo infinitesimal (Fisher, 1918), el genotipo de un individuo se modeliza como la semisuma de los genotipos de los padres más una componente de recombinación[8]. Pese a que la varianza aditiva de la descendencia se altera por la selección de los padres, produciéndose un cambio temporal que desaparece progresivamente al cesar la selección, la varianza de la componente

[7] Valores medios de las revisiones de Haley et al. (1988) y Lamberson (1990).

[8] Partiendo del concepto de "path coeficients", como medida de la porción de la desviación típica debida a un solo parental (Wrigth, 1921), se puede deducir como en una población en equilibrio Hardy-Weinberg y sin consanguinidad, esa componente de recombinación, conocida habitualmente como efecto de muestreo mendeliano, se distribuye según una normal de media cero y varianza igual a la mitad de la varianza genética.

de recombinación no se ve afectada, pues es independiente del valor de los parentales (Bulmer, 1971). De esta forma, bajo ciertas formas de selección, se pueden predecir los cambios de la varianza genética aditiva y, por lo tanto, conocer su incidencia sobre distintos métodos de estimación[9].

Respecto al método REML, los resultados de algunos trabajos de simulación (p.ej. Sorensen y Kennedy, 1984a; Thompson y Atkins, 1990; Van der Werf y Boer, 1990; Visscher y Thompson, 1990; Van der Werf, 1992; Van der Werf y Thompson, 1992) y el análisis teórico basado en funciones de verosimiltud (Gianola *et al.* , 1989; Im *et al.*, 1989), indican que ofrece estimadores que no se ven afectados por la selección de los reproductores siempre que se conozcan todas las relaciones de parentesco, y toda la información utilizada en el proceso de selección, o al menos una muestra aleatoria que no afecte su distribución. Eso significa que las estimas obtenidas en base a datos de poblaciones sometidas a selección, de las que desconocemos algunas relaciones de parentesco, estarán afectadas por la selección, más cuanta menos información familiar se conozca[10] (Van der Werf, 1992; Van der Werf y Thompson, 1992).

Los estimadores REML tampoco se ven afectados por la existencia de selección en los animales de la población base, si no hay selección posterior y tratamos esos animales como efectos fijos, tal como propusieron Graser *et al*. (1987). Tratar los animales base como efectos fijos implica ignorar la información que aportan sobre la variabilidad genética de la población[11]. En una población de tamaño infinito, esa información se puede despreciar cuando no existe selección posterior, pues en ese caso la varianza genética no se ve afectada generación a generación. Por contra, cuando existe selección, la varianza genética sí se ve afectada generación a generación, produciéndose cambios que son función de la varianza genética de la población base.

Cuando la población base no es una muestra tomada al azar, es difícil obtener estimaciones que no estén afectadas por la selección si su descendencia también es seleccionada; tratar los animales base como efectos fijos implica evitar una fuente de error, pero a la vez generar otra que puede ser de mayor importancia (Van der Werf y Thompson, 1992).

En los últimos años, el método REML se ha utilizado en el análisis de la prolificidad de numerosas poblaciones porcinas (p.ej. Gu *et al.*, 1989; Mercer y Crump, 1990; Southwood y Kennedy, 1990; Kaplon *et al.*, 1991; etc...), incluidas algunas poblaciones españolas (Pérez-Enciso y Gianola (1992) y Rodríguez *et al.*, (1993) en cerdo ibérico, y Alfonso *et al*. (1993) en razas de producción intensiva[12]), pero no siempre se puede asegurar que esas estimas estén

---

[9] Por ejemplo, asumiendo que la selección se practique por un solo carácter conocido, de forma que todos los animales que superen cierto valor se utilicen como padres, Robertson (1977) muestra cómo la regresión padre-hijo no está afectada por la selección de los padres, dado que la covarianza entre los hijos de los padres seleccionados y la varianza entre los padres seleccionados están afectadas por el mismo término; y también que la correlación entre medios hermanos cuando uno de los sexos parentales está seleccionado, está afectada por esa selección, dado que mientras la varianza dentro de los grupos de medios hermanos no está afectada, sí lo está la varianza entre grupos.

[10] En ausencia de selección y en poblaciones finitas las estimaciones también se verían afectadas por no poder tener en cuenta la deriva genética.

[11] Considerar que la varianza entre los animales base es infinita equivale a no tener información.

[12] Respecto a estas últimas, los valores oscilan entre 0.03 y 0.11, observándose diferencias

poco afectadas por el desconocimiento de determinadas relaciones de parentesco o por la existencia de selección.

Cuando no existen sistemas de recogida de información fiables resulta difícil asegurar que no existan determinados tipos de selección en los datos registrados de las poblaciones animales. En España y en referencia al porcino, nos encontramos en esa situación, pues hasta hace pocos años no han existido sistemas de información de granjas de selección públicos (Estany *et al.*, 1992) y los privados presentaban, en general, un grado de desarrollo incompleto.

Por las razones expuestas, podemos concretar dos objetivos en este capítulo. En primer lugar, estimar los componentes de varianza de la prolificidad de dos poblaciones porcinas a partir de un modelo animal con repetibilidad. En segundo lugar, analizar el efecto de la selección de la información de los animales base sobre las estimaciones de la heredabilidad.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

entre razas, semejantes a las observadas por otros autores (Lamberson, 1990): las poblaciones de animales Duroc y Large White presentan mayores h^2 que las de animales Landrace.

3.2.- Material y métodos

3.2.1.- Material.

3.2.1.1.- Material animal.

Se consideraron los datos de prolificidad (tanto número de lechones nacidos vivos por parto como número total de lechones nacidos por parto) de dos poblaciones porcinas de selección (que identificaremos como A y B), de raza Landrace, comerciales y de especialización materna. Ambas poblaciones están integradas en el Sistema GTEP-IRTA desarrollado en el Area de Producción Animal del Centro UdL-IRTA de Lleida (Noguera *et al.*, 1992), y son genéticamente evaluadas de forma regular a través de su sistema de evaluación (Estany *et al.*, 1992).

Toda la información, tanto productiva como genealógica, se filtró según los criterios de dicho Sistema (Estany *et al.*, 1992). Dado que las bases de datos del Sistema GTEP-IRTA se actualizan trimestralmente para incorporar la nueva información producida, no siempre fue posible considerar la misma cantidad de información en todos los análisis. No obstante, las comparaciones siempre se realizaron dentro de una misma base de datos, y además se intentó que las variaciones fuesen mínimas.

En la Tabla 3-1 se indica la información disponible de las poblaciones analizadas. Los datos correspondientes a la población A pertenecían a dos núcleos de selección y tres multiplicadoras, y los de la población B a dos núcleos y seis multiplicadoras.

	Población A	**Población B**
Nº total de animales	15208	6512
Nº de animales con partos	10612	4405
Nº de animales sin ascendientes conocidos	5419	1190
Nº de padres	131	293
Nº de madres	1673	1334
Coeficiente de consanguinidad medio	0.6%	0.5%
Nº de partos	38047	17111
Periodo analizado	1982-1992	1983-1992

Tabla 3-1.- Descripción de la información disponible de las poblaciones A y B.

La población B se seleccionó hasta 1991 fundamentalmente por caracteres productivos (ganancia media diaria de peso y espesor de grasa dorsal), mediante un índice muy sencillo que no tenía en cuenta la existencia de efectos ambientales, ni las diferencias en la cantidad de información entre animales[13]. La población A fue seleccionada de forma similar durante todo el periodo analizado. En ninguna de ambas poblaciones la prolificidad

[13] La selección se hacía intralote, mediante el índice: I=15 + GMD/15 - EGD, siendo GMD la ganancia media de peso en gr/día, y EGD el espesor graso dorsal en mm.

fue el objetivo de selección de mayor peso económico[14]. A partir de 1991 la población B se empezó a seleccionar de acuerdo a un índice tipificado que combina el valor genético de la ganancia media diaria de peso, el espesor de grasa dorsal y la prolificidad (Estany *et al.*, 1993). Los valores genéticos integrados en dicho índice se calculan mediante un BLUP multicarácter en el que las correlaciones genéticas entre la prolificidad y los otros dos caracteres se consideran nulas.

En 1993 se inició un experimento de selección por prolificidad en la población A (Noguera *et al.*, 1994). Se seleccionaron las mejores 175 hembras, de las aproximadamente 3500 disponibles, y 1 macho de cada una de las 25 mejores camadas, con el objetivo de comparar 150 hijas, una de cada hembra, con las hijas de una población control de 200 hembras y 15 machos tomados al azar.

En ambas poblaciones la información disponible es el resultado de una reciente informatización de la toma de datos de granja. El proceso utilizado al inicio de la informatización fue distinto en cada una de las poblaciones, pero a partir de 1990 ambas poblaciones pasaron a utilizar un mismo sistema de información, el sistema GTEP-IRTA. Por ese motivo, la información antigua es menos homogénea entre empresas que la información producida a partir de 1990.

No se conocen exactamente las características de la informatización de las bases de datos históricas, pero se sabe que, fundamentalmente en la población A, se eliminó toda la información productiva de las hembras que aproximadamente sobre 1988 no estaban en activo. Es decir, no se conoce toda la información producida en los primeros años registrados (Figura 3-1 y Figura 3-2), sino sólo la de las hembras que años después seguían estando en activo. Eso significa que la información más antigua es una muestra seleccionada de la producida, seleccionada por criterios difíciles de conocer (¿longevidad, fertilidad, ganancia media diaria,...,?), pero algunos posiblemente relacionados con la prolificidad.

En cada población, y para conocer si tenía implicaciones despreciar los datos más antiguos, se tomaron todos los datos disponibles (datos que identificaremos como *total*), y dos subconjuntos de ellos. Un primer subconjunto fue el formado por los partos producidos a partir de 1988 (datos que identificaremos como *1988*). Un segundo subconjunto (que identificaremos como *granja*) fue el formado por los partos producidos en cada granja a partir del momento en que su número se estabilizaba, tal como recogen las Figura 3-3 y Figura 3-4. Una justificación más detallada de la formación de esos subconjuntos de datos se puede encontrar en Ruiz (1993). La cantidad de información considerada en los distintos conjuntos de datos, así como sus medias y desviaciones típicas, se recoge en la Tabla 3-2. En la comparación de distintos conjuntos de datos se tuvo en cuenta toda la información genealógica disponible, y sólo se analizó el carácter nacidos vivos por parto, al juzgar redundante realizar el análisis para ambos caracteres.

[14] El criterio de selección aplicado tenía básicamente en cuenta aspectos zootécnicos: la selección de las cerdas de reposición se realizaba en base al valor esperado de su madre, valor que se calculaba en base a la productividad numérica media de los dos o tres últimos ciclos, corregida por un coeficiente fijado en función del orden de parto.

		Población A	Población B
total	**NA**	10612	4405
	n	38047	17111
	μ	8.59	9.77
	σ_ρ	2.45	2.42
granja	**NA**	9468	2613
	n	29330	9778
	μ	8.62	9.74
	σ_ρ	2.59	2.43
1988	**NA**	8728	3797
	n	26222	12945
	μ	8.60	9.78
	σ_ρ	2.59	2.59

Tabla 3-2.- Número de animales con datos (NA), número de registros (n), media (μ) y desviación típica fenotípica (σ_ρ) del número de nacidos vivos por parto, para toda la información disponible (total) y los subconjuntos granja(*) y 1988(**) de las poblaciones A y B.
(*) formado por los partos producidos a partir del momento en que su número se estabiliza en cada granja.
(**) formado por los partos producidos a partir de 1988.

3.2.1.2.- Poblaciones simuladas.

3.2.1.2.1.- Descripción general.

Se simularon varias poblaciones animales que, a efectos de estructura, se podrían considerar organizadas en un núcleo de selección y un estrato de multiplicación (Figura 3-5).

Se partió de una población base de 10 machos y 200 hembras. Tras aparear al azar cada macho con 20 hembras, se escogió, también al azar, una hija de cada camada. El resto de hembras (nh1-1) se consideró que pasaban al estrato de multiplicación. Los hijos se eligieron tomando uno de cada macho, en algunos casos al azar y en otros según el valor fenotípico de su madre. Formada la primera generación se repitió el proceso para formar una segunda generación de hembras (nh2).

Los valores genéticos de los animales de la población base se simularon considerando que la población estaba en equilibrio Hardy-Weinberg y en equilibrio de ligamiento, y que la varianza genética era el resultado de la contribución de un número infinito de loci no ligados bajo un modelo completamente aditivo, con una distribución $N(0,I.\sigma^2_a)$ (p. ej., Sorensen y Kennedy, 1984b). Los valores genéticos de las siguientes generaciones se obtuvieron mediante la expresión:

$$a_i = \frac{1}{2}(a_p + a_m) + \phi_i$$

donde: a_i: valor genético del individuo i
a_p: valor genético de su padre
a_m: valor genético de su madre

ϕ_i: efecto del muestreo mendeliano, distribuido según una $N(0,½.I.\sigma^2_a)$

Los valores fenotípicos se calcularon para un único parto según:

$$y_i=\mu+ a_i+ e_i$$

donde: y_i: valor fenotípico del individuo i

μ: media del carácter

e_i: valor residual de la observación del individuo i, distribuido según una $N(0,I.\sigma^2_e)$

Todas las poblaciones se simularon asumiendo que la media y la varianza del carácter eran 9 y 7.12 respectivamente.

3.2.1.2.2.- Poblaciones simuladas con selección en la población base.

Se simularon poblaciones, bajo la estructura y en la forma anteriormente vistas (con nh1=1 y nh2=2), considerando la existencia de dos tipos de selección:

1- [**Sω**]: Selección por truncamiento de los datos de las hembras de la población base. Se consideraron conocidos los registros de las hembras de la población base cuyo valor fenotípico fuese superior o inferior a un determinado valor (ω). Se asumieron dos valores de h^2: 0.051 y 0.5. Se simularon cuatro casos:

* ω> 9 (la media del carácter)
* ω> 12
* ω< 6
* ω< 6 y ω> 12

2- [**Sδ**]: Reducción de la variabilidad genética de los animales de la población base. Se consideró que la varianza genética de los animales de la población base era menor (con un coeficiente de reducción δ) que el doble de la varianza del muestreo mendeliano. Se asumió un único valor de h^2, 0.051, y las siguientes situaciones:

* selección de machos y hembras de la población base:
  ** δ= 0.5
  ** δ= 0.9
* selección únicamente de los machos de la población base:
  ** δ= 0.5
  ** δ= 0.9

En este caso, también se tuvo en cuenta que los machos de la primera generación podían estar seleccionados o tomados al azar, por lo que en total fueron 8 el número de casos simulados.

3.2.2.- Métodos.

3.2.2.1.- Descripción general.

Los análisis se basaron en la metodología del modelo mixto. En la estimación de los parámetros genéticos se utilizó la metodología REML a partir de un algoritmo libre de derivadas (Graser *et al.*, 1987). La estimación de los errores estándar de esos parámetros se realizó, cuando fue posible, por aproximación cuadrática a la matriz de información esperada (Smith y Graser, 1986). En la evaluación genética de los animales se utilizó la metodología BLUP bajo un modelo animal (Kennedy, 1989).

Al margen del *software* propio del Área de Producción Animal UdL-IRTA y del específicamente desarrollado, se utilizaron los programas PEST (Groeneveld *et al.*, 1990), DFREML (Meyer, 1989, 1991b)[15] y SAS (SAS Inc.,1987). En la realización de los análisis se

[15] En algunos análisis se utilizó una versión del programa DFREML con las modificaciones de Boldman y Van Vleck (1991) basadas en el programa SPARSPAK (Chu et al., 1984).

utilizaron, fundamentalmente, un MicroVax II y una VAX-Station 6040, ambos bajo sistema operativo VMS (Área de Producción Animal UdL-IRTA).

3.2.2.2.- Modelo utilizado.

El modelo utilizado fue un Modelo Animal con repetibilidad en el que se consideró toda la información familiar dispo, tal como se viene realizando para la evaluación de las poblaciones porcinas controladas por el Sistema GTEP-IRTA (Estany *et al.*, 1992):
Figura 3-6

Los factores fijos incluidos en el modelo fueron los actualmente considerados en la evaluación genética de ambas poblaciones:

- la granja-año-estación en que tuvo lugar el parto; agrupaciones trimestrales para cada año y granja: diciembre-febrero, marzo-mayo, junio-agosto, septiembre-noviembre. En total se consideraron 122 niveles en la población A y 118 en la B.
- el orden de parto; 4 niveles: parto 1º; parto 2º; partos 3º, 4º y 5º; 6º parto y posteriores (Alfonso *et al.*, 1991).

3.2.2.3.- Análisis de los datos de campo.

3.2.2.3.1.- Estimación de componentes de varianza.

Tras analizar la evolución de las medias y desviaciones típicas fenotípicas en el tiempo, se estimaron los componentes de varianza del carácter nacidos vivos por parto en los tres conjuntos de datos que se habían formado (*total, granja, 1988*) en ambas poblaciones.

El criterio de finalización del proceso iterativo se estableció en obtener una varianza del valor $-2\log_e(L)$ menor a 10^{-8} (Meyer, 1989)[16], siendo L el valor de la verosimilitud.

3.2.2.3.2.- Comparación de conjuntos de datos.

Los conjuntos de datos *total* y *granja* se compararon de acuerdo con los siguientes criterios:

1- Correlación entre los valores genéticos BLUP de los animales activos, obtenidos a partir de ambos conjuntos de datos.

2- Pérdida de respuesta esperada que implica considerar un conjunto de datos frente a otro. Considerando distintas presiones de selección y para un mismo conjunto de datos, la pérdida de respuesta se calculó como la diferencia entre la media de los valores genéticos de las hembras que se seleccionaban con un conjunto de datos y la media de los valores genéticos de las seleccionadas con el otro conjunto. Posteriormente, esa pérdida se expresó como porcentaje sobre la respuesta esperada, calculada ésta como la diferencia entre el valor genético medio predicho para las hembras seleccionadas y la media del total de hembras activas.

3- Capacidad predictiva de cada conjunto de datos (*cross-validation*). Tal como recoge la Figura 3-7, se realizaron una serie de particiones de ambos conjuntos de datos: primero se dividió cada conjunto de datos en dos partes (*total1* y *total2*; *granja1* y *granja2*) de igual tamaño, con el mismo número de animales y una distribución similar de registros dentro de cada nivel de efectos fijos; luego se tomó la parte común de *total1* con *granja1*, y *total2* con *granja2* (*unión1* y *unión2* respectivamente). Realizadas esas particiones, se predijeron los datos de cada unión en base a los datos que no contenían, es decir, se predijo *unión1* en base a *total2* y *granja2,* y *unión2* en base a *total1* y *granja1*, haciendo un total de cuatro predicciones.

Cada predicción se realizó a partir de las soluciones BLUE y BLUP obtenidas

[16] Ese mismo criterio se utilizó en todos los análisis unicarácter realizados posteriormente.

previamente[17] y operando según:
Figura 3-8

Para cada predicción se calculó el coeficiente de correlación (ρ) entre los valores predichos y los realizados, y el error cuadrático medio según la expresión:
Figura 3-9

El ECM así calculado se expresó dividido por la varianza de las observaciones ($\sigma^2_{y^{(p)}_{ijk}}$), para así hacerlo comparable entre particiones con distinta variabilidad fenotípica.

Para conocer la bondad de predicción de los conjuntos *total* y *granja*, se calculó el valor medio de y ECM de las dos predicciones realizadas en base a los datos de cada uno de ellos.

Este tipo de predicciones, si bien sólo nos permite conocer cual es el conjunto de datos que mejor predice los valores fenotípicos, ha sido utilizado por otros autores como una buena aproximación para inferir sobre la predicción de valores genéticos (Pérez-Enciso *et al.*, 1993; Estany y Sorensen, 1995).

3.2.2.3.3.- Análisis de la selección de los animales de la población base.

Tal como recoge el Anejo 1, partiendo de los trabajos de Graser *et al.* (1987) y Van der Werf y Thompson (1992), se puede hallar una expresión sencilla de las ecuaciones del modelo mixto (MME) del modelo animal que permite hacer distintas hipótesis de selección de los animales de la población base. De forma análoga, para un modelo con repetibilidad la expresión sería:
Figura 3-10

[1][18]

Esas ecuaciones son fáciles de construir. La construcción de A^{*-1} es muy sencilla, pues todos sus elementos son los mismos que los de la inversa de la matriz de parentesco A^{-1}, exceptuando los de la diagonal. Así, basta con atribuir a los elementos del vector **d** de las reglas de Henderson valor 0 en vez de 1 cuando no conocemos ni padre ni madre, es decir cuando se trata de un animal de la población base. Por otra parte, la matriz H^{*-1} de la expresión [1] es de la forma:
Figura 3-11
en donde H_b es una matriz que permite recoger la existencia de selección en la población que actúa como base. La matriz de varianzas y covarianzas genéticas de los animales de la población base puede expresarse como:
$V(a_b)=H_b.\ \sigma^2_a$

De esta forma, considerando distintas matrices H_b, la expresión [1] permite hacer múltiples hipótesis sobre los animales de la población base: todos los animales seleccionados, los machos seleccionados, varianzas heterogéneas entre animales de la población base,...,. Así, por ejemplo, se puede ver que tratar los animales de la población base como fijos no es más que un caso particular; el de asumir que la varianza en la población base es infinita, lo que equivale a que los elementos de H_b tengan valor, y por lo tanto H^{*-1} sea una matriz de ceros[19].

La expresión [1] se utilizó para la obtención de estimas REML de las componentes de

[17] Considerando los componentes de varianza estimados en los conjuntos total y granja según el caso.
[18] En el Anejo 1 también se puede ver como bajo las hiótesis habitualmente asumidas respecto a la población base (no selección, no parentesco y no consaguinidad), esta expresión conduce a la expresión habitual de las MME del modelo animal.
[19] Coincidiendo la expresión [1] con la expresión de Graser et al. (1987).

varianza en la población A, teniendo en cuenta distintos grados de selección de los animales base (machos y hembras).

Para ello, llamando:

Figura 3-12

y definiendo δ como:

$\delta = \sigma^2_a$(población considerada base) / σ^2_a (verdadera población base),

se consideraron distintos valores de δ correspondientes a distintos grados de reducción de la varianza aditiva en la población base conocida; concretamente 0.9, 0.5 y 0.1.

Las soluciones REML se obtuvieron modificando la parte del programa DFREML que construye la inversa de la matriz de parentesco de acuerdo a las reglas indicadas, de forma que los animales de la población base en vez de tener 1 como elemento de la diagonal de esa inversa, tuvieron 1/δ.

3.2.2.4.- Análisis de las poblaciones simuladas.

En los dos tipos de poblaciones simuladas (**S*ω*** y **Sδ**) se realizaron los siguientes análisis:

1- En las poblaciones **S*ω*** (con selección de los datos de las hembras de la población base) se estimó la heredabilidad mediante dos modelos:

* un modelo aleatorio (el mismo que el utilizado en la simulación de los datos)

* un modelo mixto, en el que se consideró un efecto fijo generación (con dos niveles: generación base y resto de generaciones).

2- En las poblaciones **Sδ** (con selección de los animales de la población base) se estimó la heredabilidad mediante otros dos modelos:

* un modelo aleatorio que no tiene en cuenta la selección simulada (con la expresión habitual de las MME).

* un modelo aleatorio que sí la tiene en cuenta (con la expresión de las MME expuesta en el apartado anterior (3.2.2.3.3))

En todos los casos el proceso se repitió un mínimo de 100 veces, tras comprobar que no existiesen problemas de autocorrelación entre las distintas réplicas (Morgan, 1984; pág. 200).

3.3.- Resultados y discusión.

3.3.1.- Estimación de la variabilidad genética mediante un modelo con repetibilidad.

Los valores de h^2 estimados en ambas poblaciones considerando toda la información disponible son muy bajos (0.03 y 0.05), especialmente en la población A (Tabla 3-3). Pese a ser superiores a los obtenidos por otros autores (p.ej., Haley y Lee, 1992; Ferraz y Johnson, 1993), son inferiores a los habitualmente referidos en la bibliografía (Haley *et al.*, 1988; Lamberson, 1991) y los obtenidos en otras poblaciones españolas (Alfonso *et al.*, 1993). Sin embargo, la proporción de varianza fenotípica debida al efecto permanente (c^2) es mayor de lo esperado, de forma que la repetibilidad es sólo ligeramente inferior a 0.15, valor medio tradicionalmente considerado (Haley *et al.*, 1988).

Población	h^2	c^2	σ^2_a	σ^2_c
A	0.03 ± 0.01	0.09 ±0.01	0.19	0.55
B	0.05 ± 0.01	0.08 ± 0.01	0.27	0.43

Tabla 3-3.- Estimaciones de la heredabilidad, la proporción de varianza fenotípica explicada por el efecto permanente (±s.e.), y las varianzas aditiva y del efecto permanente, del número de nacidos vivos en las poblaciones A y B.

Es posible que no exista una diferencia significativa entre las estimaciones de h^2 obtenidas en ambas poblaciones, pero dado el bajo valor que tienen, el hecho de considerar una u otra condiciona notablemente las expectativas de progreso genético. Bajo selección individual e igualdad en el resto de parámetros de los que depende la esperanza de respuesta genética, en la población A sólo se esperaría alcanzar el 60% de la respuesta de la población B. Las escasas expectativas de progreso genético que se derivan de una h^2 tan baja, contrastan con los resultados provisionales del experimento de selección por prolificidad realizado a partir de la población A, y comentado al describir el Material animal (apartado 3.2.1.1.). Con los primeros partos realizados en aproximadamente un 30% de las hembras de la primera generación de la línea seleccionada, existe una diferencia fenotípica entre esa línea y la control de cerca de 1 lechón (J.L. Noguera, comunicación personal).

Es posible que el valor de h^2 estimado en la población A, 0.03, esté infraestimado. Una primera causa es desconocer una parte importante de las relaciones de parentesco entre los animales (Van der Werf y Boer, 1990). En la población A se desconocen los padres del 36% de los animales. En la población B, en la que el valor estimado de h^2 es el doble que en la población A, se desconocen los padres del 18% de los animales. Es decir, en ambas poblaciones el valor de h^2 obtenido puede estar infraestimado, fundamentalmente en la A. Al utilizar un modelo con repetibilidad con la matriz de parentesco incompleta una parte de la variabilidad genética puede pasar a ser explicada por el efecto permanente, de forma que la repetibilidad se vea poco afectada. De hecho, ya hemos comentado que los valores de repetibilidad estimados se separan poco del valor bibliográfico de referencia.

La existencia de determinados tipos de selección podría ser otra causa que condujese a valores infraestimados de h^2. La información que tenemos sobre ambas poblaciones indica que es muy poco probable que haya existido una selección por prolificidad de los padres lo suficientemente importante como para que afecte a las estimas obtenidas. Por contra, existen indicios para sospechar que la información disponible no es una muestra aleatoria de toda la producida.

La Figura 3-13 recoge, para las dos poblaciones, la evolución de la media y la desviación típica fenotípicas del número de nacidos vivos respecto al año de nacimiento de las cerdas. En la población A la desviación típica aumenta significativamente a medida que se avanza en el tiempo, permaneciendo constante la media, mientras que en la población B media y desviación son estables en el tiempo (Tabla 3-4). El comportamiento de la varianza fenotípica en la población A parece confirmar la existencia de algún tipo de selección, selección que podría estar relacionada con el bajo valor de h^2 estimado.

		a	**b**
Población A	μ	8.635	-0.006
	σ_p	1.720	0.094 **
Población B	μ	9.799	-0.009

	σ_ρ	2.434	0.005

Tabla 3-4.- Resullados de la regresión (y = a + bx) de la media (μ) y la desviación típica (σρ) fenotípicas del número de nacidos vivos sobre el aflo de nacimiento de las cerdas en las poblaciones A y B. ** (α=0.01)

Tal como se comentó en el apartado 3.2.1.1. la información procedía, fundamentalmente en la población A, de una base de datos que sólo recogió la información reproductiva histórica de los animales que en ese momento estaban en activo. Por ese motivo, de los años más antiguos sólo se conocen los datos de las hembras más longevas, y no de los de sus contemporáneas. Por otra parte, su población base se caracteriza por estar formada por un elevado número de hembras, un 45%, de las que se conocen algunos de sus partos, habiendo producido más del 50% de los partos conocidos de los años más antiguos de cada granja.

3.3.2.- Análisis de la existencia de selección en la información de los animales base.

Cuando se hace inferencia sobre algún parámetro, dado un conjunto de datos, se puede pensar en la existencia de varios tipos de selección: selección basada en los datos (***y***); selección basada en ***y*** y otros factores desconocidos, pero independientes de los parámetros a estimar; y selección basada en ***y*** y otros factores desconocidos de los que dependen los parámetros sobre los cuales se desea hacer inferencia (Gianola *et al.*, 1989; Im *et al.*, 1989). El tercer tipo es el más general aunque imposible de tener en cuenta, pues no se dispone de toda la información sobre cómo se ha realizado la selección.

Por su parte, Henderson (1975, 1984) diferencia tres casos de selección, que se pueden enmarcar dentro de los casos comentados: selección basada directamente sobre ***y***, selección basada sobre los valores genéticos (***u***), y selección basada sobre los residuos (***e***). Pensar en una selección directa sobre ***u*** o ***e*** es un tanto artificial, pues ni ***u*** ni ***e*** se pueden conocer con exactitud (Thompson, comentario en Im *et al.*, 1989). En realidad, sólo se puede realizar selección sobre ***y***. Esa selección afectará tanto a la distribución de ***u*** como a la de ***e***; en función del valor de la h^2 esa selección afectará más a la distribución de ***u*** que a la de ***e***, o la inversa. Por ese motivo, si sólo conocemos un subconjunto seleccionado de ***y*** para determinados grupos de animales, podemos pensar en tener en cuenta su efecto tanto sobre la distribución de ***u*** como sobre la de ***e.***

Asumiendo normalidad, los principales efectos de esa selección son sobre la media y sobre la varianza de ***e*** y ***u.*** Si la selección afectase sólo a la media de ***e*** se podría tener en cuenta incluyendo un efecto fijo para los distintos grupos de observaciones afectadas. Si afectase sólo a la media de ***u*** también se podría tener en cuenta a través de otro efecto fijo, en este caso asignando grupos genéticos (Quaas, 1988). No obstante, no sólo afectará a ambas medias conjuntamente, sino que también afectará a las varianzas de ***u*** y ***e*** , por lo que incluir ambos efectos tratándolos como fijos no bastaría. Si afectase únicamente a la varianza de ***e***, se debería considerar la existencia de varianzas residuales heterogéneas entre grupos de observaciones afectadas por distintos efectos fijos (p.ej. Henderson, 1984). Finalmente, si la selección sólo afectase a la varianza de ***u***, se habría de considerar la existencia de varianzas genéticas heterogéneas entre grupos de animales (p.ej. Visscher y Thompson, 1990; Lo *et al.*, 1993).

Una alternativa a la complejidad que supone considerar dos efectos adicionales y heterogeneidad de varianzas asociada a cada uno de ellos, es despreciar la información seleccionada, asumiendo que no se ha utilizado para escoger los animales reproductores de

cada generación, pues en caso contrario los valores de h^2 que se estimasen estarían en mayor o menor medida sesgados.

Como se vió en el apartado anterior, hay suficientes indicios para pensar en una selección de las observaciones de las hembras base de la población A. Además se sabe que no se ha seleccionado directamente por prolificidad. Por esos motivos, despreciar la información seleccionada, considerando toda la genealogía conocida puede permitir obtener estimaciones de la h^2 no sesgadas por la selección. El primer inconveniente que se presenta es conocer exactamente cual es la información seleccionada, y qué conjunto de datos está realmente libre de ella. Por otra parte, supone despreciar una gran cantidad de información: en el mejor de los casos, casi un 23% de los partos registrados, es decir cerca de 9000 (Tabla 3-2). Si bien en esta población no es un serio inconveniente por disponer de un gran volumen de datos, en otras situaciones puede ser muy limitante proceder de ese modo.

Si no se quisiese o pudiese despreciar la información seleccionada, se deberían tener en cuenta todos los efectos que la selección puede tener sobre las distribuciones de $\boldsymbol{u}$ y $\boldsymbol{e}$, o al menos los más importantes. En un carácter de baja heredabilidad como la prolificidad, una selección de los datos de la población base consistente en conocer sólo los mejores registros afectaría fundamentalmente a la media de $\boldsymbol{e}$, por lo que parece razonable tenerla en cuenta incluyendo un efecto fijo como el que se incluyó en las poblaciones analizadas: granja-año-estación. Además, cabe pensar que el efecto de la selección sobre la varianza de $\boldsymbol{u}$, aunque pequeño puede afectar de forma importante a la inferencia que hagamos sobre el valor de la h^2.

Una selección que implique cambios en la varianza genética de la población base, se puede tener en cuenta tal como se indicó en el apartado 3.2.2.3.3. Ello supone estimar más parámetros en base a los mismos datos, por lo que el proceso de estimación se hace más complejo e impreciso. Una posible alternativa es asumir conocidas las diferencias entre las varianzas genéticas de distintos grupos de animales y posteriormente maximizar la verosimilitud condicional (tal como la definen Gianola *et al.*, 1989).

Considerar exclusivamente un factor fijo 'ambiental' y una reducción de la varianza genética de los animales base, sólo permite tener en cuenta parcialmente una selección como la presumida en la población A, pero es una aproximación relativamente sencilla que recoge, además, otras situaciones habituales de selección de los animales base. Por eso, a continuación se analizarán por separado los siguientes aspectos:

- el efecto, sobre las estimaciones de la h^2, de considerar un factor fijo de tipo 'ambiental' en poblaciones simuladas con selección por truncamiento de la información de la población base,
- el interés de despreciar, según los resultados anteriores, la información presuntamente seleccionada de las hembras más antiguas de la población A, dado que si se desprecia la información seleccionada no existen cambios en $\boldsymbol{e}$,
- el efecto de la reducción de la variabilidad genética de los animales base de poblaciones simuladas, sobre las estimaciones de la h^2,
- y las implicaciones de considerar una reducción de ese tipo en la población A.

3.3.2.1.- Efecto de la selección de los datos de la población base por truncamiento.

Los resultados de las Tablas 3-5 y 3-6 corresponden a poblaciones simuladas (Sω, apartado 3.2.1.2.2.) en las que se ignoraron ciertos registros de las hembras base, y en las que se estimó la h^2 bajo un modelo aleatorio y un modelo mixto que consideraba como efecto fijo la generación (dos niveles, uno para la generación base y otro para el resto de generaciones).

Las estimaciones se realizaron considerando la matriz de parentesco simulada.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

Selección (ω)	Población con h^2=0.051	Población con h^2=0.5
>9	0.189 ±0.004	0.533 ±0.003
>12	0.156 ±0.007	0.596 ±0.004
<6	0.158 ±0.006	0.601 ±0.004
<6 y >12	0.078 ±0.005	0.608 ±0.004

Tabla 3-5.- Heredabilidades (±s.e.) estimadas con un modelo aleatorio en una población en la que se seleccionaron los datos de las hembras de la población base por truncamiento sobre sus valores fenotípicos (ω).

ω	h^2=0.051	h^2=0.5
>9	0.043 ±0.003	0.409 ±0.003
>12	0.053 ±0.003	0.457 ±0.004
<6	0.050 ±0.003	0.465 ±0.004
<6 y >12	0.078 ±0.005	0.591 ±0.004

Tabla 3-6.- Heredabilidades (±s.e.) estimadas con la inclusión de un efecto generación en una población en la que se seleccionaron los datos de las hembras de la población base por truncamiento sobre sus valores fenotípicos (ω).

Si nos ceñimos al modelo aleatorio (Tabla 3-5) podemos observar que en esas poblaciones siempre se sobrestima la h^2. Esto ocurre tanto cuando se conocen sólo los mejores datos, como los peores, o ambos. El grado de sobrestimación depende tanto del punto de truncamiento como del valor de h^2. Para una h^2 de 0.051 la sobrestimación (con un valor de hasta un 278%) es siempre mayor que para una h^2 de 0.5 (con un valor máximo de un 22%). Cuando el punto de truncamiento se aleja de la media, la sobrestimación disminuye para el valor de h^2 de 0.051, pero aumenta cuando ésta es de 0.5. Lo mismo sucede cuando se seleccionan los dos extremos de las observaciones, los peores y los mejores registros. Estas sobrestimaciones se deben al hecho de estar atribuyendo diferencias de ambiente a causas genéticas.

Esos resultados no permiten explicar los obtenidos en la población A, pero ayudan a entender algunos resultados referidos en la bibliografía. Vangen (1986) analizando sólo la información de las cerdas que habían alcanzado un determinado parto, halló algunos valores de h^2 superiores a los obtenidos al utilizar la información de todas las cerdas. Así, por ejemplo, cuando analizaba la h^2 del segundo parto de las hembras que habían alcanzado el quinto (pasando de considerar 17713 partos a considerar 5096) obtenía una h^2 de 0.14 frente a 0.09. Es una situación con cierta similitud a la que hemos analizado puesto que los 5096 segundos partos de las hembras seleccionadas cabe esperar que sean los mejores de los 17713 producidos.

En la Tabla 3-6 podemos ver cómo afecta el incluir un factor fijo de generación (semejante al factor granja-año-estación considerado en el análisis de los datos de campo), para ajustar las observaciones conocidas por los valores ausentes, es decir, para corregir las diferencias ambientales entre observaciones con medias distintas. Para una h^2 de 0.05 se

produce una infraestimación (0.043) cuando el punto de truncamiento se sitúa en la media, pero las estimaciones no parecen verse afectadas si la cantidad de información seleccionada es escasa, es decir, si sólo conocemos los partos por encima de 12 o por debajo de 6. No ocurre lo mismo para una h^2 de 0.5, pues siempre que existe un único punto de truncamiento, se produce infraestimación.

La infraestimación de h^2 que se observa bajo el modelo mixto debe atribuirse a la confusión que se produce entre ***u*** y ***e,*** pues tal como hemos indicado anteriormente una selección sobre ***y*** como la simulada afecta a la distribución de ***u***, y su magnitud depende del parámetro que queremos estimar. Así, por ejemplo, el valor de 0.043, obtenido para una h^2 de 0.051 cuando sólo se consideran los datos de los animales base que superan la media, corresponde a una varianza genética estimada de 0.29, cuando la simulada fue de 0.36.

Al comparar ambos modelos se observa que para una h^2 de 0.5 el sesgo que se produce bajo el modelo mixto puede ser mayor que el producido bajo el modelo aleatorio. Para una h^2 de 0.051 no sucede lo mismo, de forma que el efecto fijo considerado siempre permite tener parcialmente en cuenta una selección como la simulada, excepto cuando la selección se basa en dos puntos de truncamiento; en ese caso no aporta nada tenerlo en cuenta. Estos resultados son explicables si pensamos que se ha producido una selección tanto sobre ***u*** como sobre ***e***, que sólo tenemos en cuenta la selección sobre ***e***, y que cuando la h^2 es baja la selección sobre ***u*** es muy pequeña.

En ese sentido, si se comparan los resultados de Johansson y Kennedy (1985) con los que obtuvieron previamente en la misma población sin considerar un efecto año-rebaño, se observa como se pasó de unos valores medios iniciales de h^2 de 0.16 a valores de 0.07, estando posiblemente el primer resultado sobrestimado y el segundo infraestimado.

Otra situación que se ha recogido en la Tabla 3-5 y Tabla 3-6 es la selección basada en dos puntos de truncamiento, tal que haga que sólo conozcamos los datos por debajo y por encima de dos determinados valores. Pese a que es una situación que difícilmente podemos encontrar en poblaciones comerciales, permite ver, como ya se ha apuntado, que incorporar un efecto fijo al modelo de estimación no aporta nada, pues siempre conduce a valores sobrestimados de h^2. Esto es así, por que en realidad existen tres grupos de observaciones y no sólamente los dos considerados, de modo que no se recogen los cambios en las medias de ***e***.

3.3.2.2.- Comparación de conjuntos de datos de las poblaciones analizadas.

Como se comentó, el modelo utilizado en el análisis de los datos de campo contenía un efecto fijo granja-año-estación con niveles trimestrales para cada año y granja, es decir un efecto fijo que podría tener parcialmente en cuenta la existencia de selección en los datos agrupados en sus primeros niveles. No obstante, de acuerdo a los resultados obtenidos en las poblaciones simuladas se podría pensar que conduce a valores de h^2 infraestimados.

Al estimar la h^2 en los subconjuntos de datos *granja* y *1988* de la población A, en los que no se consideraron los datos más antiguos, se obtuvo un valor del doble (0.06) del obtenido al considerar toda la información (Tabla 3-7).

Población	Datos (*)	h^2	c^2	σ^2_a	σ^2_c
A	*granja*	0.06 ± 0.01	0.08 ±0.01	0.37	0.52
	1988	0.06 ± 0.02	0.08 ±0.01	0.36	0.50
B	*granja*	0.06 ± 0.02	0.09 ±0.02	0.36	0.59
	1988	0.05 ± 0.01	0.09 ±0.01	0.27	0.48

Tabla 3-7.- Estimaciones de la heredabilidad, la proporción de varianza fenotípica debida al efecto permanente (±s.e.), y las varianzas aditiva y del efecto permanente, del número de nacidos vivos por parto en dos subconjuntos de datos de las poblaciones A y B.
(*) *granja* : formado por los partos producidos a partir del momento en que su número se estabiliza en cada *granja*; *1988* formado por los partos producidos a partir de 1988.

En la Tabla 3-7 se puede ver que los subconjuntos granja y 1988 de la población A conducen al mismo resultado al considerar la misma información genealógica pero distinta cantidad de información reproductiva (el 77 y el 69% del total disponible respectivamente). Lo mismo sucede en la población B, en la que no existen indicios de selección; incluso despreciando el 43% de información (subconjunto granja) se obtiene prácticamente el mismo valor (0.06±0.02 frente a 0.05±0.01).

También en la Tabla 3-7 se ve como el valor estimado de la proporción de varianza fenotípica debida al efecto permanente es equivalente al estimado al considerar todos los datos, de forma que la variación observada en la h^2 de la población A no se puede vincular a un distinto reparto de la variabilidad debida al animal, entre el valor aditivo y el efecto permanente.

Mientras en la población B los conjuntos total y granja conducen a las mismas decisiones de selección, no sucede lo mismo en la población A. El coeficiente de correlación entre los valores genéticos BLUP de los animales activos de la población A (1549), predichos utilizando toda la información disponible o parte de esa información (granja), fue de sólo 0.89. Por contra, en la población B, al realizar el mismo análisis (857 animales activos), el coeficiente de correlación fue de 0.99.

En la población A, de los dos subconjuntos formados, granja y 1988, parece más adecuado considerar granja dado que conduciendo a los mismos resultados considera mayor cantidad de información.

La Tabla 3-8 recoge los porcentajes de pérdida de respuesta esperada al seleccionar distintos porcentajes de animales de la población A en base a una evaluación BLUP que considera uno de los dos conjuntos de datos (total o granja), frente a la evaluación que considera el otro. En ella se puede ver cómo para intensidades muy elevadas (presiones de un 1%), características de esquemas de selección hiperprolíficos como el que se realiza en la población A, utilizar toda la información disponible implica hasta un 35% de pérdida respecto a utilizar la información parcial, mientras que proceder de forma inversa supone a lo sumo un 8%.

% animales Seleccionados	Evaluación *total* (*)	Evaluación *granja* (**)
1	35%	8%
5	15%	14%
10	7%	11%
20	9%	7%

Tabla 3-8.- Porcentaje de pérdida de respuesta esperada en la población A al seleccionar según una evaluación BLUP con un conjunto de datos en relación a la obtenida utilizando otro.
(*) conjunto de todos los datos disponibles; (**) conjunto formado por los partos producidos a partir del momento en que su número se estabiliza en cada granja.

Long *et al*., (1991) en el caso extremo de ignorar el 21% de los peores datos de prolificidad de la población porcina que analizan sólo hallan una pérdida del 5% en respuesta genética. Los resultados de la Tabla 3-8, indican pérdidas superiores desconsiderando un porcentaje similar de información (23%), pero a diferencia del trabajo referido, los valores que se señalan se obtuvieron teniendo en cuenta el valor de h^2 estimado en cada conjunto de datos, y no un valor constante de h^2.

El mismo análisis realizado en la población B, dio como resultado una pérdida nula en todos los casos, confirmando que las soluciones obtenidas con uno u otro conjunto de datos eran equivalentes.

Finalmente, la Tabla 3-9 recoge los resultados del análisis de la capacidad predictiva que tenían los conjuntos total y granja, viéndose que el conjunto granja es el que presenta mayor capacidad, dado que el cociente entre el error cuadrático medio y la varianza de las observaciones utilizadas es menor (un 1%) y el coeficiente de correlación mayor (un 4%). Las diferencias entre ambos conjuntos de datos son, pese a todo, escasas, pero es difícil esperar grandes variaciones en modelos en los que se explica una pequeña fracción de la variabilidad total, tal como se puede ver en resultados de otros autores como Pérez-Enciso et al. (1993) o Estany y Sorensen (1995).

Conjunto de datos	ecm/σ^2_y	*r*
total (*)	0.94	0.26
granja (**)	0.93	0.27

Tabla 3-9.- Capacidad predictiva de dos conjuntos de datos de la población A en té~ de error cuadrático medio tipificado (ecm/σ^2_y) y correlación entre los valores observados y los predichos (r).
(*) conjunto de todos los datos disponibles; *(**)* conjunto formado por los partos producidos a partir del, momento en que su número se estabiliza en cada granja.

De todos estos resultados, se desprende el interés de no considerar todos los datos disponibles de la población A, sino sólo los partos de aquellos años-estación en los que la

información no es una muestra seleccionada de la que se produjo, es decir los contenidos en el conjunto *granja*. Ya se comentó que despreciar información seleccionada es posible cuando se dispone de mucha información, pero no en situaciones en las que la cantidad de información es limitante.

3.3.2.3.- Efecto de la reducción de la variabilidad genética en los animales de la población base.

Cuando no se puede despreciar información y existe selección en la población base, se debe intentar modelizarla para evitar obtener estimaciones de h^2 sesgadas por esa selección. Analizando lo que sucede en poblaciones simuladas (Sδ, apartado 3.2.1.2.2.), los resultados de la Tabla 3-10 muestran que la selección de la población base, consistente en una reducción de su varianza genética, infraestima la h^2 si no se tiene en cuenta, siendo mayor el sesgo cuanto más intensa es la selección. También se puede ver que cuando se seleccionan machos y hembras, es decir, cuanto mayor es el porcentaje de animales base seleccionados respecto al total de animales, también es mayor el sesgo que se produce. Así, por ejemplo, cuando se seleccionan los machos de la primera generación, una reducción de la variabilidad genética aditiva de los animales base de un 50%, infraestima la h^2 un 29% cuando machos y hembras están seleccionados, y un 20% si sólo lo están los machos; si la variabilidad es un 90% de la total, el efecto pasa a ser de un 10 y un 6% respectivamente. El sentido de los sesgos que se producen coincide, en general, con el observado en otros trabajos (Van der Werf y Thompson, 1992; Pieramati y Van Vleck, 1993). En esos trabajos se concluye que el efecto de la selección de la población base depende de la existencia o no de selección posterior. En nuestras poblaciones no se observa con claridad esa dependencia, posiblemente por el bajo valor de h^2 asumido (0.051) y la débil selección realizada sobre los machos de la primera generación.

Animales base seleccionados	**δ simulado**	**δ considerado**	**NS**	**S**
Machos y hembras	**0.90**	**1**	0.053±0.003	0.046±0.003
	0.50	**1**	0.038±0.002	0.036±0.002
	0.90	**0.90**	0.050±0.003	0.051±0.003
	0.50	**0.50**	0.052±0.004	0.052±0.004
Sólo machos	**0.90**	**1**	0.050±0.003	0.048±0.003
	0.50	**1**	0.040±0.003	0.041±0.003

Tabla 3-10. Heredabilidades estimadas en poblaciones simuladas con una h^2=0.051 y menor varianza genética en algunos animales base (reducida en un coeficiente δ), al considerar o no esa reducción y en ausencia (NS) o presencia (S) en los machos de la primera generación.

La forma que habitualmente se ha empleado para evitar la infraestimación que hemos observado ha sido tratar los animales de la población base como fijos (Graser *et al.*, 1987). No obstante, ya se indicó que ese tratamiento no es adecuado cuando se seleccionan los animales de las siguientes generaciones (Van der Werf, 1992), siendo preferible tener en

cuenta la selección de la población base en la construcción de las MME, tal como se desarrolla en el Anejo y aplican Van der Werf y Thompson (1992). Los resultados de la Tabla 3-10 muestran que proceder de ese modo es una forma efectiva de tener en cuenta la selección de los animales base, dado que existe coincidencia entre los valores de h^2 simulados y los estimados, tanto en ausencia como en presencia de selección posterior. El inconveniente de esta forma de proceder es asumir que se conoce la selección realizada, condición que difícilmente se producirá al analizar poblaciones animales. Pese a todo, si la población base está seleccionada y además existe selección posterior puede ser preferible asumir cierto grado de selección de los animales base, que tratarlos como fijos, o al menos hacer varios supuestos de selección y ver si la estimación de h^2 es sensible a ellos.

En el apartado 3.3.2.1. se discutió que realizar una selección fenotípica de las observaciones de los animales base implicaba una infraestimación de la h^2 al incluir un efecto fijo generación. La inclusión de ese efecto permitía corregir parcialmente las diferencias entre las medias de los dos grupos de observaciones que producía la selección, pero no tenía en cuenta los cambios en sus varianzas. Se asumió que existía la misma varianza genética entre los animales de la población base que entre los animales de las siguientes generaciones, cuando en realidad no era así; la varianza entre los animales base está afectada por la selección realizada, pero no la de los animales no base.

El cambio que se produce en la varianza de las observaciones de los animales base se puede predecir fácilmente cuando la selección es por truncamiento. Como señala Robertson (1977), la varianza de una variable y normalmente distribuida y seleccionada por truncamiento en un valor ω, de forma que seleccionemos una proporción p, se reduce en una proporción $K=(f(\omega) / p) [(f(\omega) / p) - \omega]$, siendo $f(\omega)$ el valor de la distribución en el eje de ordenadas. Si se conoce ω y la h^2, la reducción de la varianza genética se puede predecir, dado que es Kh^2. De esta forma, podemos tener en cuenta la reducción de la varianza genética en las estimaciones de la h^2 de la Tabla 3-6. Los resultados se recogen en la Tabla 3-11, para los puntos de truncamiento 9 y 12.

ω	δ	h^2 =0.05		h^2=0.5
>9	0.97	0.051 ±0.003	0.68	0.512 ±0.006
>12	0.96	0.051 ±0.003	0.59	0.550 ±0.006

Tabla 3-11.- Heredabilidades (±s.e.) estimadas con la inclusión de un efecto generación en una población en la que se seleccionaron los datos de la población base por truncamiento sobre sus valores fenotípicos (ω), y teniendo en cuenta el efecto de esa selección sobre la varianza genética de la población base (δ=1-K).

Al comparar los resultados de la Tabla 3-6 con los de la Tabla 3-11 se observa, para una h^2=0.051, que al considerar la reducción de varianza genética en la población base, se obtienen estimaciones insesgadas. Para el valor de h^2 de 0.5, cuando el truncamiento se realiza en la media fenotípica (9) el sesgo se reduce desde un valor máximo de un 18% a un valor de sólo un 2%, pero si se realiza en un punto extremo (12) se comete prácticamente el mismo sesgo, aunque en sentido contrario.

Considerar un efecto fijo y reducción de la varianza genética no es más que una aproximación a la obtención de estimas libres de los sesgos producidos por selección por truncamiento; como se indicó, para tenerlos completamente en cuenta se debería considerar

adicionalmente las diferencias entre las medias genéticas y las varianzas residuales que provocan la selección. No obstante, para bajas h^2 un tratamiento de este tipo puede ser suficiente, posiblemente debido a una escasa repercusión de la existencia de varianzas residuales heterogéneas y la pequeña diferencia entre las medias genéticas de los animales con información seleccionada y el resto de animales.

3.3.2.4.- Asunción de distintos grados de selección de los animales base de las poblaciones analizadas.

Como acabamos de comentar, una aproximación suficiente al problema de disponer de información seleccionada de los animales base de la población A, que no obligue a despreciar la información presuntamente seleccionada, puede ser considerar el efecto fijo granja-año-estación habitualmente incluido, y una reducción de la varianza genética en los animales base. Esa reducción no se conoce, pero se puede ver como afecta a la estimación de la h^2 considerando distintos grados de reducción. En la Tabla 3-12 se recogen los valores de h^2 así estimados al considerar toda la información disponible (conjunto *total*)[20]. A medida que nos separamos de la hipótesis de no selección de los animales de la población base, considerando mayores grados de reducción en su varianza genética, la h^2 va aumentando hasta alcanzar un valor de 0.06 cuando se considera que la varianza genética es sólo un 10% de la varianza genética de la población.

	Conjunto *total*		***Conjunto granja***	
δ	**h^2**	**c^2**	**h^2**	**c^2**
1	*0.02±10.01*	*0.10±0.01*	*0.06±0.01*	*0.08±0.01*
0.9	0.02±0.01	0.09±0.01	0.06±0.01	0.08±0.01
0.5	0.04±0.01	0. 9±0.01	0.05±0.01	0.09±0.01
0.1	0.06:10.01	0.09±0.01	0.04±0.01	0.10±0.01

Tabla 3-12.- Heredabilidad y proporción de varianza fenotípica explicada por el efecto permanente (±s.e.) estimadas al considerar distintos grados de selección de la población base (δ), usando los conjuntos total y granja de la población A.

Podemos ver que la h^2 se mueve en el intervalo 0 - 0.06, dentro del cual se encuentra el valor obtenido al despreciar la información más antigua. De este modo, la selección de la información más antigua de la población A parece tener un efecto equivalente a una reducción de la variabilidad genética de los animales base, de más de un 50%. Aunque parece difícil pensar en una selección tan fuerte, en la Figura 3-13 se pudo ver que la varianza fenotípica de las observaciones de los animales nacidos en 1980 era inferior en más de un 50% a la de los animales nacidos en 1990.

Tal como cabía esperar, al realizar el mismo análisis con el conjunto *granja* en el que se puede asumir que se ha evitado la información seleccionada, no se observan diferencias

[20] La diferencia entre el valor 0.02 y el valor 0.03 de la Tabla 3-5 debe atribuirse a las variaciones causadas en la información por la actualización trimestral que sufren las bases de datos.

tan grandes como las observadas al utilizar toda la información disponible. Es más, tal como recoge la Tabla 3-12, no se puede considerar que existan diferencias significativas entre los valores estimados. Pese a todo, el valor de h^2 parece tender a disminuir al aumentar el grado de selección de los animales base, mostrando una tendencia opuesta a los resultados obtenidos por Van der Werf y Thompson (1992) al trabajar con poblaciones simuladas. Esa oposición se puede deber tanto al grado de conocimiento de las relaciones de parentesco, como a la existencia de selección de los machos de las siguientes generaciones y a la propia estructura poblacional; mientras en la población A se desconoce una parte importante de los parentescos y no existe selección posterior, en las poblaciones simuladas por Van der Werf y Thompson (1992) se conocen todas las relaciones de parentesco, existe selección posterior, y la estructura familiar es muy sencilla, implicando sólo 102 animales.

El desconocimiento de un importante número de relaciones de parentesco se puede traducir en cierta confusión entre el valor genético y otros efectos del modelo, básicamente el efecto permanente; si reducimos artificialmente la varianza genética, puede ser más difícil separar el efecto animal del efecto permanente, tal como parece indicar el hecho de que la repetibilidad sea constante para todos los grados de reducción de la varianza genética considerados (Tabla 3-12).

Estos resultados parecen ratificar el valor de h^2 estimado al despreciar la información más antigua, 0.06, como una estimación adecuada del valor de la h^2 de NV de la población A.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

4.- Análisis de la existencia de efectos maternos

4.1.- Introducción

Se conoce como efecto materno la influencia que una madre ejerce sobre un carácter expresado en un descendiente y que no se debe al efecto directo de los genes que le ha transmitido. Es un efecto con un componente genético respecto a la madre pero que actúa exclusivamente como ambiental sobre su descendencia (Van Vleck, 1993). Se pueden definir dos periodos de influencia materna: prenatal, cuyo componente principal es la influencia uterina, y postnatal, en el que actúa principalmente la influencia de la lactación (Willham, 1963). La influencia del periodo postnatal sobre la prolificidad puede eliminarse considerablemente al estandarizar los tamaños de camada durante la lactación (Ollivier, 1982; Haley *et al.*, 1988). Por contra, la influencia del periodo prenatal sólo puede corregirse, en la práctica, estadísticamente (Kirkpatrick y Rutledge, 1988).

Se han planteado diversos modelos para tener en cuenta la existencia de efectos maternos, como son los de Willham (1963) y Falconer (1965). El modelo de Willham (1963) trata un carácter con influencia materna como el resultado de un efecto genotípico y ambiental relativo a ese carácter y al individuo en el que se expresa (efecto directo), más un efecto genotípico y ambiental relativo a ese u otro carácter y a su madre (efecto indirecto de origen materno[21]. El modelo de Falconer (1965) también considera esos caracteres como el resultado de efectos genotípicos, ambientales y maternos, pero definiendo el componente de efectos maternos como una función lineal del valor fenotípico de la madre. Esa función lineal viene dada por un coeficiente de regresión parcial del valor de un individuo sobre el de su madre, referido a la media de la población y en ausencia de variación genética entre las madres[22]. Cuando se considera la existencia de un componente ambiental común entre animales criados por una misma hembra, bajo el modelo de Falconer este componente no sólo contiene los efectos ambientales comunes entre esos animales, sino también los efectos maternos no relacionados con el carácter (Van der Steen, 1983)[23].

Partiendo del modelo de Willham, el modelo animal se puede adaptar fácilmente para considerar la existencia de efectos maternos. Se trata de añadir dos nuevas variables al

---

[21] Px = Gox + Eox + Gmw + Emw, siendo, Px : valor fenotípico del individuo x;G: valor genotípico; E: valor ambiental; o y m: origen directo y materno, respectivamente, de los efectos; w: madre del individuo x. Posteriormente Willham (1972) incorpora al modelo un efecto abuela, el de la abuela materna, como primera generalización a otro tipo de antecesores. Por otra parte, en este modelo el componente ambiental directo se puede dividir en un componente ambiental común (C) entre animales criados por una misma hembra, y en un componente debido a otras causas ambientales (Eox*), de forma que: Eox = C + Eox* (Van der Steen, 1983).

[22] P = G + mP' + E, siendo, P : valor fenotípico; G : efecto genotípico; mP': función lineal sobre el valor fenotípico de la madre; E : efecto ambiental (que al igual que en el modelo de Willham se puede descomponer en un componente ambiental común entre animales criados por una misma hembra (C) y un componente debido a otras causas ambientales (E*), de forma que E = C + E*).

[23] Van der Steen (1983), divide los efectos maternos del modelo de Willham, en efectos maternos relacionados con el tamaño de camada (1) y en efectos maternos no relacionados (n), pudiendo comparar los modelos de Willham (W) y Falconer (F):

W: $P_x = (G_{ox}) + (G_{mlw}+E_{mlw}) + (C+G_{mnw}+E_{mnw}) + (E_{ox})$

F: P = (G) + (mP') + (C) + (E*).

modelo tradicional, una que recoja los efectos no genéticos que influyen sobre los fenotipos de todos los hijos de una hembra, y otra que recoja el efecto genético de la madre[24] (Quaas *et al.*, 1984).

En determinadas especies y carácteres, como en vacuno de carne y peso al destete, existen evidencias para afirmar no sólo que existen efectos maternos, sino también que estos están, generalmente, correlacionados de forma negativa con el efecto genético directo (ver por ejemplo la revisión de Cantet (1990)). No obstante, respecto al tamaño de camada en porcino (tanto nacidos vivos, como totales por parto) la existencia de estos efectos no es evidente.

Han sido varios los trabajos que han analizado la influencia de los efectos maternos sobre la prolificidad en cerdos, pero los resultados no siempre han sido coincidentes. Algunos resultados indican que las cerdas criadas en camadas grandes paren un menor número de lechones que las criadas en camadas de tamaño menor (Revelle y Robinson, 1973; Nelson y Robinson, 1976; Rutledge, 1980a), y que, posiblemente, esa relación no es lineal (Urban *et al.*, 1966). La existencia de un efecto negativo de ese tipo debería implicar una reducción del valor de heredabilidad cuando se estima por regresión entre el parto en que nace una cerda y los que ella produce (Alsing *et al.*, 1980). Sin embargo, los valores de heredabilidad así estimados no siempre son significativamente distintos de los estimados utilizando otros registros u otras relaciones de parentesco (Strang y Smith, 1979; Vangen, 1980; Hill y Webb, 1982; Van der Steen, 1983; Avalos, 1985).

Los resultados obtenidos más recientemente, utilizando el modelo animal con efectos maternos y métodos basados en la verosimilitud, indican la inexistencia de efectos maternos significativos sobre la prolificidad (Tabla 4-1). Resultados semejantes se obtienen, de forma análoga, en otras especies polítocas como es el caso de la cunícola (Gómez, 1994). Pese a que en la Tabla 4-1 se recogen algunos resultados significativos que indican una correlación negativa entre los efectos directo y materno, éstas no son siempre elevadas, y además van asociadas a la inexistencia de variabilidad genética directa, por lo que la significación debe considerarse poco relevante.

---

[24] Podemos expresar el modelo animal con efectos maternos como:

$Y = X \cdot b + Z_a \cdot a_a + Z_m \cdot a_m + T \cdot l + e$

siendo, Y, b, a_d, a_m, l, e : vectores de observaciones, efectos fijos, valores genéticos directos, valores genéticos maternos, efectos de ambiente común, y residuos, respectivamente; y X, Z_d, Z_m, T = matrices que relacionan Y con b, a_d, a_my l, respectivamente. De esta forma la relación entre este modelo y el de Willham viene dada por: $a_d = G_{ox}$;$a_m = G_{mlw}$; $l = E_{mlw} + C + G_{mnw} + E_{mnw}$; $e = E_{ox}^{*}$.

Referencia	Raza	Car.	Parto	Parámetros				
				h_a^2	h_m^2	r_{am}	c^2	l^2
Mercer y Crump (1990)	**L**	**NV**	**Pl**	0.11	0.00	0.00		0.00
			VP	0.09	0.01	0.00	--	0.02
		NT	**Pl**	0.11	0.0	0.00		0.00
			VP	0.10	0.02	0.00	--	0.03
Southwood y Kennedy (1990)	**L**	**NV**	**Pl**	0.18 (0.06)	0.08 (0.03)	-0.93 (0.11)		
		NT		0.25 (0.05)	0.05 (0.03)	-0.92 (0.14)		
	Y	**NV**		0.10 (0.06)	0.04 (0.04)	-0.23 (0.53)	--	--
		NT		0.13 (0.06)	0.06 (0.04)	-0.45 (0.3 5)		--
Haley y Lee (1992)	**LW**	**NV**	**Pl**	0.01 (NS)	0.00 (NS)	-0.98 (NS)		0.13(NS)
Pérez-Enciso y Gianola (1992)	**IB**	**NV**	**VP**	0.04 (0.03)	0.00 (NS)	0.00 (NS)	0.07 (0.03)	--
		NT		0.06 (0.03)	0.00 (NS)	0.00 (NS)	0.07 (0.03)	
Ferraz y Johnson (1993)	**I-LW**	**NV**	**VP**	0.00	0.01(**)	-0.07(NS)	0.17(**)	
Rodriguez *et al.* (1994)	**I**	**NV**		0.06(0.03)	0.00 (NS)	--	0.06(0.02)	
***Irgang et al.* (l 994)**	**D**	**NV**	**Pl**	0.16(0.11)	0.04 (NS)	-0.72 (NS)	--	0.06 (NS)
			P2	0.15(0.10)	0.09(NS)	-0.88(NS)		0.00 (NS)
				0.27(0.18)	0.07(NS)	-1.00(NS)		0.00 (NS)
	L	**NV**	**Pl**	0.15(0.05)	0.00(NS)	-0.50(NS)	--	0.00 (NS)
			P2	0.27 (0.07)	0.08 (NS)	-0.63 (NS)		0.00 (NS)
			P3	0.03 (0.08)	0.01 (NS)	-1.00 (NS)		0.00 (NS)
	LW	**NV**	**Pl**	0. 11 (0.06)	0.00 (NS)	-0.95 (NS)	--	0.00 (NS)
				0.01 (0.09)	0.12 (S)	-0.42(S)		1 (NS)
				0.25(0.09)	0.02(NS)	-1.00(NS)		0 (NS)

Tabla 4-1.- Estimaciones REML de los componentes de varíanza (nivel de significación o por delecto, errores estándar), obtenidas utilizando un modelo animal con efectos maternos, en nacidos vivos (NV) y totales (NT), y distintas razas porcinas.

Abreviaciones: h_a^2: heredabilidad directa; h_m^2 : heredabilidad efectos maternos; r_{am} :correlación genética entre el efecto directo y materno; c^2: proporción de varianza fenotípica explicada por el efecto permanente; l^2: proporción de varianza fenotípica explicada por el efecto ambiental de camada; L Landrace; LW: Large White; Y: Yorkshire; I: Ibérico; D: Duroc; P 1: primer parto; P2: segundo parto; P3: tercer parto; VP: varios partos; NS: no significativo; (**): P<0.0 1; S: significativo (P no indicada).) Sólo se han considerado los errores estándar calculados por aproximación cuadrática a la matriz de información de Fisher y las significaciones LRT.

No obstante, se sabe que la estima de efectos maternos encierra determinados

problemas ligados a la confusión existente entre efectos genéticos directos y maternos (Willham, 1980), y es difícil hallar valores significativos mediante el modelo animal con determinadas estructuras de
datos (Gerstmayr, 1991)[25] o si los efectos son poco importantes aún trabajando con diseños específicos y gran cantidad de información (Meyer, 1992).

Si existiesen efectos maternos, el haberlos ignorado permitiría explicar parcialmente por qué no se han alcanzado las expectativas teóricas de progreso genético en algunos experimentos de selección *(Robinson, 1972; Rutledge, 1980b; Alsing et al.*, 1980; Le Roy *et al.*, 1987). Algunos trabajos indican que, si existe un efecto negativo sobre la prolificidad, cabe esperar una reducción de la respuesta a la selección (Van der Steen, 1983), incluso si sólo se da sobre los primeros partos (Vangen, 1980). De todas formas, asumiendo que exista reducción de respuesta, ésta puede evitarse en gran parte si se estandariza el tamaño de camada durante la lactación (Van der Steen, 1983; Avalos, 1985).

Ante este contexto, no sólo es difícil concluir si la influencia de los efectos maternos sobre la prolificidad es significativa, y tiene o no un efecto negativo, sino que también lo es saber si puede ser económicamente tan importante como para tenerla en cuenta (Bichard y David, 1985). No obstante, se considera que el desarrollo de un programa de selección eficiente para la productividad de las cerdas debe tener en cuenta la posible existencia de efectos maternos (Robinson, 1981), puesto que ignorarlos conduciría a predicciones de los valores genéticos y estimaciones de la respuesta a la selección sesgadas (Roehe y Kennedy, 1993). El valor absoluto de este sesgo depende fundamentalmente del valor de la covarianza entre los efectos genéticos maternos y directos (Southwood y Kennedy, 1990), pero si su signo fuese negativo, la respuesta genética real sería siempre menor a la esperada (Gómez, 1994).

El objetivo de este capítulo es analizar la existencia de efectos maternos sobre los caracteres número de nacidos vivos y nacidos totales por parto, en las dos poblaciones consideradas en este trabajo.

4.2.- Material y métodos

4.2.1.- Material.

4.2.1.1.- Material animal.

Se consideraron los datos de nacidos vivos (NV) y nacidos totales (NT) por parto de los animales con madre conocida, pues de esta forma se puede utilizar la misma matriz de parentesco para los efectos genéticos directos y maternos, simplificándose la construcción de las ecuaciones del modelo mixto.

Las variables NV y NT se analizaron separadamente para el primero y todos los órdenes de parto. La información utilizada se describe en la Tabla 4-2, procediendo la de la población A del conjunto de datos *granja* descrito en el capítulo anterior, y la de la población B del conjunto de todos los datos disponibles.

[25] La estructura más desfavorable es no conocer ningún registro de las madres y un sólo hijo por madre, en la que no es estimable el componente de varianza asociado a los efectos maternos.

Población	Parto		NT	NV	A	B	D
A	Primero	μ	8.79	8.31	0.67	0.58	8.26
	(*n*=7269)	σ	2.34	2.34	1.38	1.13	1.50
	Todos	μ	9.33	8.81	0.65	0.70	8.58
	(*n*=24886)	σ	2.49	2.42	1.39	1.27	1.48
B	Primero	μ	9.89	8.98	0.48	0.43	8.75
	(*n*=3647)	σ	2.44	2.33	1.22	0.94	1.92
	Todos	μ	10.45	9.46	0.45	0.41	9.15
	(*n*=16309)	σ	2.60	2.40	1.18	0.90	1.86

Tabla 4-2.- Descripción de la información utilizada en el análisis de los efectos matemos sobre nacidos vivos y totales por parto en las poblaciones A y B. (NT: nacidos totales, NV: nacidos vivos, A: adoptados, B: bajas en lactación, D: destetados).
Símbolos: *n* = número de datos; μ = media ; σ = desviación típica.

4.2.1.2.- Poblaciones simuladas.

Se simularon varias poblaciones de la forma descrita en el apartado 3.2.2.1. del capítulo anterior, considerando que no existe ningún tipo de selección, y tomando de cada parto 4 hijas con registros, según la estructura que se vió en la Figura 3-5 (nh1=nh2= 4).

Las simulaciones se realizaron (200 réplicas) considerando un único parto por hembra y bajo dos modelos distintos (Tabla 4-3):

1 - [MS1]: un modelo en el que se consideró un efecto directo de tipo genético, un efecto de ambiente común y un efecto residual. Esos efectos se simularon considerando una σ_p= 2.39, una heredabilidad (h^2_a) de 0.10 y una fracción de varianzas debida al efecto ambiental (l^2) de 0.10. La simulación se realizó tal como se indicó en el apartado 3.2.1.2. El efecto de ambiente común se consideró igual para todos los animales de una misma camada, de forma que al no realizar ningún tipo de adopciones, se debe entender como un efecto de ambiente común entre hermanos.

2 - [MS2]: un modelo como el anterior en el que además se tuvo en cuenta un efecto materno de tipo genético, correlacionado con el efecto directo. Ambos efectos se simularon como se describió en el apartado 3.2.1.2, utilizando las varianzas de la forma canónica de la matriz de (co)varianzas entre ambos (Searle, 1982). Los valores de esa matriz correspondieron a los asumidos en el modelo anterior, una heredabilidad del efecto materno (h^2_m) de 0.05 y una correlación genética entre ambas variables (r_{am}) de -0.5. Posteriormente, ambas variables se transformaron utilizando las propiedades de la transformación canónica (Lin, 1988). Se siguió el mismo procedimiento para simular los valores de los componentes de recombinación de ambos efectos[26] .Figura 4-1.

[26] Si para un individuo i,f_{di} es el componente de recombinación de los efectos directos y f_{mi} el componente de recombinación de los efectos maternos, se cumple (Cantet et al., 1992):
Ver Figura 4-1.

	Parámetros			
Modelos	h^2_m	h^2_m	r_{am}	l^2
MS1	√			√
MS2	√	√	√	√

Tabla 4-3.- Modelos empleados en la simulación de poblaciones, en función de los parámetros considerador (símbolod utilizados en el texto).

4.2.2.- Métodos.

Se estimaron los componentes de varianza para la prolificidad de las poblaciones descritas utilizando el programa DFREML (Meyer, 1991b). Los errores estándar de los componentes de varianza se calcularon de la forma indicada en el apartado 3.2.2.1. Posteriormente se contrastaron los valores estimados de los componentes de (co)varianza asociados a los efectos maternos, mediante un contraste de razón de verosimilitudes (LRT - *Likelihood Ratio Test*-, Kendall y Stuart (1979)).

4.2.2.1.- Modelos utilizados.

Se utilizaron diversos modelos basados en un modelo animal con repetibilidad, un efecto materno y un efecto de ambiente común entre las hembras criadas por una misma cerda:

Figura 4-2

En la Tabla 4-4 se recogen los diversos modelos considerados a partir de ese modelo general. En todos los casos se tuvo en cuenta toda la información familiar disponible; en el análisis de los datos de campo se consideraron los mismos factores fijos que en los análisis del capítulo anterior: la granja-año-estación y el orden de parto.

Parámetros						
Datos	Modelos	h^2_a	h^2_m	r_m	c^2	l^2
De campo: primeros partos	ME l	√				
	ME2	√	√			
	ME3	√	√	√		
De campo: todos los partos	ME4	√			√	
	ME5	√	√		√	
	ME6	√	√	√	√	
Poblaciones simuladas	MEI	√				
	ME2	√	√			
	ME3	√	√	√		
	ME7	√				√
	ME8	√	√			√
	ME9	√	√	√		√

Tabla 4-4.- Relación de modelos empleados en el análisis de efectos maternos en distintos conjuntos de datos, en función de los parámetros considerados (símbolos utilizados en el texto).

4.2.2.2.- Contraste entre modelos.

Los valores estimados de los componentes de (co)varianza asociados a los efectos maternos se contrastaron a través del contraste de la razón de verosimilitudes (LRT), que se basa en la distribución X^2 que presenta $2\log_e\lambda$, cuando la hipótesis nula (H_0) es cierta, siendo λ el cociente entre la verosimilitud (L) de una muestra bajo H_0 y la verosimilitud de la misma muestra bajo la hipótesis alternativa (H_1). De acuerdo a esta definición, $2\log_e\lambda$ se puede expresar como la diferencia entre las funciones soporte de la verosimilitud, $l = -2 \log_e L$, bajo H_1 y H_0.

Asumiendo que bajo H_0 estimamos un vector de parámetros $\hat{\vartheta}_0$, que pertenece a un espacio Ω_0 , en el que maximizamos la verosimilitud de la muestra, y bajo H_1 hacemos lo propio para un vector $\hat{\vartheta}_1$, que pertenece a un espacio Ω_1, el contraste de razón de verosimilitudes será:

Figura 4-3

En el caso en que H_0 sea cierta, $2\log_e\lambda$, se distribuirá siguiendo una distribución X^2 con $[\dim(\Omega_1)-\dim(\Omega_0)]$ grados de libertad.

Este contraste se puede aplicar aunque se trabaje con verosimilitudes "restringidas" (REML), siempre y cuando se consideren los mismos efectos fijos y el mismo número de niveles en todos los modelos de análisis.

4.3.- Resultados y discusión

4.3.1.- Estimaciones en los datos de granja.

Al analizar la significación de los componentes de varianza de los distintos modelos considerados, mediante un contraste de razón de verosimilitudes, sólo resultó significativo, para todos los caracteres y poblaciones, el valor de la varianza del efecto genético directo ($P \leq 0.05$). Ni los valores de la varianza materna ni los de la covarianza entre los efectos directo y materno resultaron significativos (Tabla 4-5), excepto el de ese término de covarianza para NV de la población A cuando se analizaron todos los partos conjuntamente. En ese caso el valor de la varianza materna no era significativo, y tampoco lo fue el contraste conjunto de los valores de varianza y covarianza, de forma que no se puede afirmar que exista variación genética de origen materno. Estos resultados son completamente coincidentes con los hallados por otros autores en porcino (Mercer y Crump, 1990; Southwood y Kennedy, 1990; Haley y Lee, 1992; Pérez-Enciso y Gianola, 1992; Ferraz y Johnson, 1993; Rodríguez *et al.*, 1994; Irgang *et al.*, 1994) tal como refleja la Tabla 4-1 de la Introducción.

BIBLIOTECA VIRTUAL

Componente contrastado (grados de libertad)					
Población	**Carácter**	σ^2_m	σ_{am}	$\sigma^2_m + \sigma_{am}$	**logL (·)**
		(1)	**(1)**	**(2)**	
A	**NV1**	1.48	0	1.08	-17240.88
	NV	0	62.06	0.78	-41316.66
	NT1	3.08	1.48	4.56	-17190.37
	NT	0	0	0	-41662.23
B	**NV1**	0	2.40	2.40	-9053.56
	NV	0.77	1.16	1.94	-26455.85
	NT1	0	0.48	0.40	-9101.66
	NT	0	0	0	-27171.63

Tabla 4-5.- Valores y significación de la razón de verosimilitud (2logλ) de distintos contrastes asociados a los componentes de (co)varianza de los efectos maternos para nacidos vivos y totales en primer parto (NV1 y NT1 respectivamente) y en todos los partos (NV y NT respectivamente).
* (PO.05)
(•) logaritmo, natural de la verosimilitud del modelo sin efectos maternos: σ^2_p $\sigma^2_a+\sigma^2_e$ para NV1 y NT1; σ^2_p $\sigma^2_a+ +\sigma^2_c+\sigma^2_e$ para NV y NT.

Para el carácter NV, las estimaciones oscilaron entre 0.17 y 0.48 para la varianza directa, entre 0 y 0.15 para la materna y entre -0.12 y -0.02 para la covarianza entre ambos efectos. Para NT los valores de varianzas fueron semejantes, entre 0.13 y 0.36 para la varianza directa y entre 0 y 0.06 para la materna, pero los valores de la covarianza entre

ambos efectos no siempre fueron negativos (oscilaron entre -0.03 y 0.06). En conformidad con las significaciones halladas, la inclusión de los efectos maternos, ya sea tratándolos correlacionada o incorrelacionadamente con los efectos directos, no supuso una reducción importante de la varianza residual.

Los parámetros genéticos estimados bajo los distintos modelos empleados, tanto para NV y NT en los primeros partos como para el conjunto de todos los partos de las poblaciones A y B, se recogen en la Tabla 4-6. El mayor valor de heredabilidad materna supera escasamente el 2%, mientras que la heredabilidad directa se mueve en un rango del 3-8%, prácticamente igual que en el Capítulo anterior. La correlación, como se desprendía de las covarianzas obtenidas fue siempre negativa, excepto en un único caso (NT del primer parto en la Población A), en el que además su valor fue prácticamente del 90%. El rango de variación de la correlación genética indica un comportamiento errático, pues cubre prácticamente todo su dominio.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

Pob.	Car.	Modelo	h^2_a	h^2_m	r_{am}	c^2	σ_p
A	**NV1**	**ME1**	0.033 (0.019)	--			2.485
		ME2	0.027	0.006	--		2.484
		ME3	0.031	0.007	-0.189	--	2.484
	NV	**ME4**	0.045	--	--	0.078	2.465
		ME5	0.058 (0.023)	0.023 (0.019)	--	0.088 (0.016)	2.501
		ME6	0.054	0.000	-0.999	0.074	2.466
	NT1	**ME1**	0.040 (0.035)	--	--	--	2.469
		ME2	0.028	0.011	--		2.466
		ME3	0.021 (0.014)	0.006 (0.007)	0.897		2.468
	NT	**ME4**	0.054 (0.010)	--		0.080 (0.008)	2.507
		ME5	0.055	0.000		0.079	2.507
		ME6	0.060	0.000	-0.999	0.077	2.508
B	**NVI**	**ME1**	0.053 (0.022)	--	--	--	2.394
		ME2	0.053 (0.028)	0.000 (0.020)	--		2.394
		ME3	0.084 (0.045)	0.005 (0.031)	-1.000	--	2.396
	NV	**ME4**	0.043 (0.010)	--	--	0.116 (0.011)	2.431
		ME5	0.040 (0.011)	0.003 (0.004)	--	0.116 (0.011)	2.431
		ME6	0.048	0.006	-0.395	0.112	2.431
	NT1	**ME1**	0.038	--	--	--	2.422
		ME2	0.038	0.000	--		2.426
		ME3	0.053	0.000	-0.853		2.425
	NT	**ME4**	0.055 (0.014)			0.089 (0.000)	2.536
		ME5	0.055	0.000	--	0.088	2.536
		ME6	0.051 (o.006)	0.000 (0.000)	-1.000	0. 109 0.005)	2.540

Tabla 4-6.- Valores estímados de los parámetros genéticos (·) y desviación típica fenotípica (σ_p) para nacidos vivos y totales en los primeros (NVI y NT1) y todos los partos (NV y NT), en distintos modelos. (Entre paréntesis se ofrecen los errores estándar cuando fue posible estimarlos).
(·) h^2_a= heredabilidad directa; h^2_m= heredabilidad materna; r_{am}= correlación genética entre los efectos directos y maternos; c^2= fracción de variabilidad correspondiente a los efectos ambientales permanentes.

En el supuesto que los efectos maternos de origen prenatal fuesen poco importantes, y los de origen postnatal estuviesen directamente relacionados con el número de individuos que una cerda amamanta, la estandarización de las camadas tras el nacimiento implicaría una

reducción de la varianza asociada al componente de efectos maternos, y eventualmente su eliminación. Esto permitiría explicar por qué los efectos maternos son tan poco relevantes en las poblaciones analizadas. De hecho, en ambas poblaciones, existe un importante intercambio de lechones entre camadas, de entre 0.45 y 0.67 lechones adoptados por parto (Tabla 4-2), valores superiores a los referidos por otros autores, como por ejemplo Souhtwood y Kennedy (1990) (0.44 lechones/parto). Este intercambio de lechones implica una estandarización que se ve reflejada en la disminución de la variabilidad numérica del tamaño de camada entre el parto y el destete. En la Tabla 4-2 se puede ver como se pasa de valores de desviación típica de NV de entre 2.33 y 2.42, a valores de entre 1.48 y 1.92, para el número de destetados (ND), valores que como cabía esperar son inferiores a los mostrados por Souhtwood y Kennedy (1990) (2.58). La mortalidad durante la lactación también puede tender a estandarizar el tamaño de camada al destete, pero en ninguna de las dos poblaciones la mortalidad fue especialmente alta en ese periodo. Sus valores, de entre un 4.8 y un 7.9% según se desprende de la Tabla 4-2, fueron inferiores a los valores medios referido por el Sistema GTEP-IRTA (tal como se recoge en García (1994)), pero semejantes a los referidos por otros autores como Kaplon *et al.*, (1990) e Irgang *et al.*, (1994).

Se puede considerar que la estandarización observada en el tamaño de camada durante la lactación, es suficiente, el 90% de las camadas tienen entre 6 y 12 lechones. No obstante, es posible que existan efectos ambientales comunes entre las producciones de cerdas pertenecientes a una misma camada de lactación que no sean directamente atribuibles al tamaño de camada durante la lactación. Si nos fijamos en el modelo de Willham, podemos ver que dentro de ese efecto podemos recoger influencias como la aptitud maternal de la cerda de lactación, el ambiente de la camada de nacimiento de la madre, etc... Si en realidad fuese así, las estimaciones obtenidas estarían sesgadas, por lo que sería necesario tener en cuenta esos efectos de ambiente común. No obstante, los datos disponibles no lo permitieron, puesto que sólo se conocía la camada de lactación de menos del 10% de los animales (814 y 518 en las poblaciones A y B respectivamente). En las simulaciones realizadas en el siguiente apartado se analiza cómo puede afectar desconsiderar la existencia de tales efectos.

4.3.2.- Estimaciones en poblaciones simuladas.

Las Tablas 4-7 y 4-8 recogen los resultados de los análisis realizados en poblaciones simuladas con el objeto de conocer, por una parte, hasta que punto son detectables influencias maternas de pequeño efecto y, por otra, la repercusión de ignorar la influencia de efectos de ambiente común entre animales criados en una misma camada, tal como se hizo con los datos de campo.

Modelo simulación	Modelo	estimación	h^2_a	h^2_m	r_{am}	l^2
Con efectos maternos (MS1)	Sin efecto camada	ME1	0.19			
		ME2	0.06	0.09	--	
		ME3	0.09	0.12	-0.41	
	Con efecto camada	ME7	0.09	--		0.10
		ME8	0.08	0.02		0.09
		ME9	0.09	0.03	-0.33	0.09
Sin efectos maternos (MS2)	Con efecto camada	ME7	0.10			0.09
		ME8	0.09	0.02	--	0.08
			0.10	0.03	-0.28	0.09

Tabla 4-7.- Estimaciones de los parámetros genéticos bajo tres modelos de análisis, en una población simulada con efectos maternos [MSI:h^2_a=0.1, h^2_m=0.05, r_{am}=-0.5, l^2=0. 1] y otra sin efectos maternos [MS2: h^2_a= 0 1, l^2= 0.1] (error estándar <0.005)

Componente contrastado (grados de libertad)					
Modelo simulación	Modelo estimación	σ^2_m (1)	σ_{am} (1)	$\sigma^2_m + \sigma_{am}$ (2)	$\log_e$ L (·)
Con efectos maternos	Sin efecto camada	40.46 ± 6.02 **	-3.62 ± 6.75	33.22 ± 6.96 **	-3414.92 ±2.21
(MS1)	Con efecto camada	-4.26 ± 5.74	2.18 ± 5.74	0.10 ± 5.72	-3390.84 ±2.02
Sin efectos maternos (MS2)	Con efecto camada	5.44 ± 5.80	-3.12 ± 6.23	-0.40 ± 5.58	-3393.74 ±1.80

Tabla 4-8.- Valores y significación de la razón de verosimililud ($2\log_e\lambda$) de distintos contrastes asociados a los componentes de (co)varianza de los efectos maternos, cuando estos fueron simulados y estimados bajo modelos alternativos. ** (P=0.05)
(•) logaritmo natural de la verosimilitud del modelo sin efectos maternos: $\sigma^2_p + \sigma^2_a + \sigma^2_e$

En la Tabla 4-7 se puede ver como, cuando se incluyen los efectos de ambiente común, los valores estimados para los distintos parámetros (h^2_a; h^2_m; r_{am}; l^2) son prácticamente iguales, con independencia de que la población se haya simulado con o sin efectos maternos (p.ej. MS1-ME9 frente a MS2-ME9). Esos valores son bastante cercanos a

los simulados, pero no son significativos (Tabla 4-8), es decir que ni aun existiendo efectos maternos fue posible detectarlos. En poblaciones simuladas a partir de h² mayores, pero con matrices de parentesco incompletas, Quintanilla *et al.* (1995) tampoco logran estimar los valores simulados. Si bien en nuestras poblaciones simuladas se conocía completamente la matriz de parentesco, hay que tener en cuenta que sólo se conocían dos generaciones con datos.

Por otra parte, al ignorar el efecto ambiental de camada cuando en realidad existe se sobrestima el valor de la h² de los efectos maternos (MS1-ME3 frente a MS1-ME9, (Tabla 4-7). Esa sobrestimación está asociada a la significación de la varianza de los efectos maternos, pero no a la del término de covarianza, pues en ningún caso este parámetro es significativo (Tabla 4-8). Cuando tampoco se tienen en cuenta los efectos genéticos maternos, es el valor de h² de los efectos directos el que se ve sobrestimado (MS1-ME1 frente a MS1-ME7, (Tabla 4-7).

Estos resultados indican que en los datos de campo, al haber ignorado un posible efecto camada de ambiente común, las estimaciones de los efectos maternos pueden estar sobrestimadas, aunque aun así recordemos que no eran significativas. Por otra parte, ayudan a explicar valores especialmente altos de efectos maternos hallados por otros autores al desconsiderar la existencia de un efecto ambiental entre cerdas criadas en una misma camada (Southwood y Kennedy, 1990; Ferraz y Johnson, 1993).

No obstante, en los datos de campo la relación existente entre efecto materno y efecto camada es menor, pues en las poblaciones simuladas cada hembra producía un único parto, es decir no existían medias hermanas de madre. Además, la estandarización del tamaño de camada durante la lactación, y las adopciones de lactantes que se realizan durante ese periodo, son prácticas habituales de manejo. Por esos motivos, cabe pensar que las sobrestimaciones que se puedan producir sean menores.

La dificultad de detectar efectos maternos significativos se debe, en parte, a que la estructura de las poblaciones utilizadas no es la óptima para estimar este tipo de parámetros, y en parte, a que la eficiencia de los procedimientos de búsqueda del máximo de la función de verosimilitud, utilizados en los algoritmos libres de derivadas, depende fuertemente de la estructura de covarianzas entre los parámetros que se quieren estimar, siendo menor cuando estos están correlacionados (Meyer, 1992). Por estos motivos, el enfoque bayesiano puede ser particularmente útil, al proveer de mayor información sobre el grado de credibilidad que podemos dar a las estimaciones de los componentes de (co)varianza de los efectos maternos (Cantet, 1990; Janss *et al.*,1994; Jensen *et al.*, 1994).

5.- Análisis de la existencia de heterogeneidad genética entre partos

5.1.- Introducción

Tradicionalmente se ha considerado que el determinismo genético de la prolificidad era el mismo en todos los partos de una cerda. Bajo esa hipótesis, el modelo de evaluación genética más apropiado es un modelo con repetibilidad, que es el habitualmente utilizado en las líneas porcinas seleccionadas por prolificidad (Ducos *et al.*, 1992). Este modelo se ha estudiado con anterioridad en los Capítulos 3 y 4.

No obstante, al igual que en su día distintos trabajos apuntaron la hipótesis de correlaciones genéticas entre lactaciones de una misma vaca, distintas a la unidad, lo mismo sucedió para la prolificidad en porcino (Tartar, 1981; Tartar y Bolet, 1984).

Bichard y David (1985) presentaron una revisión de las correlaciones genéticas estimadas entre distintos órdenes de parto. Señalando la considerable variación existente entre ellas, concluyeron que en general son próximas a la unidad, aunque menores entre órdenes de parto más separados en el tiempo; no cabía, por lo tanto, esperar una gran reducción en el progreso genético por asumir identidad genética entre partos.

Posteriormente, Vangen (1986) revisó otros resultados, así como las estimaciones de heredabilidad (h^2) obtenidas en distintos órdenes de parto. Observó que había poca variación entre las h^2 de los distintos órdenes de parto, pero las correlaciones genéticas (r_g) se separaban bastante de la unidad, y al igual que señalaron Bichard y David (1985) eran menores cuanto más separados estaban los partos entre sí, correspondiendo las más bajas a las obtenidas entre el primero y el resto de partos. Esos resultados le condujeron a concluir que el tamaño de camada del primer parto está parcialmente determinado por otros factores genéticos que el resto de partos, y además, que en los programas de selección se debería considerar cada parto como un carácter diferente.

Haley *et al.* (1988) recogiendo los resultados publicados para los primeros cuatro partos, y ponderándolos en función de su error de estimación, también concluyeron que existe poca variación entre la h^2 de los distintos partos, y que las correlaciones genéticas son altas entre partos consecutivos, siendo sólo la r_g entre 1º y 4º parto, con un valor de 0.3, menor de 0.7. Respecto a esas correlaciones señalan que pueden estar afectadas, entre otras causas, por la selección de las cerdas en base a sus registros previos, por lo que sería interesante obtener estimaciones REML, menos afectadas por ese tipo de selección que las hasta entonces obtenidas por otros métodos. Avalos (1985) ya indicaba que el escaso número de registros existentes para órdenes de parto tardíos, la progresiva selección de las hembras en sus primeros partos, y la baja heredabilidad del carácter, hacían difícil obtener buenas estimaciones de las correlaciones genéticas.

Pese al desarrollo informático alcanzado en los últimos años, la obtención de estimaciones REML multicarácter sigue siendo onerosa, fundamentalmente cuando el número de caracteres que se considera es elevado. Posiblemente por esa razón se han publicado muy pocas estimaciones REML de las correlaciones genéticas entre distintos órdenes de parto: Knap *et al.* (1993), Rydhmer (1993) e Irgang *et al.* (1994) en porcino, y Gómez (1994) en conejos.

Knap *et al.* (1993), utilizando un modelo bivariante en tres líneas distintas, y considerando el primero y el resto de partos como caracteres diferentes, no obtienen diferencias en los valores de h^2. De las tres líneas analizadas, las de especialización materna (Landrace y Yorkshire) presentaron una r_g menor (0.86 y 0.77 respectivamente) que la línea Yorkshire de machos terminales (0.95).

Rydhmer (1993), también mediante un modelo bivariante, pero considerando sólo los primeros y segundos partos de la población de Yorkshire sueco, obtiene valores semejantes

de h² y una r_g de 0.67.

Finalmente, Irgang *et al.* (1994) analizan, mediante un modelo que considera cada parto como carácter distinto, los registros de los tres primeros partos de poblaciones Duroc, Landrace y Large White de Brasil. Las correlaciones genéticas que obtienen no muestran concordancia con los resultados recopilados por Vangen (1986) y Haley *et al.* (1988), pues son bajas entre primer y segundo parto (entre 0.37 y 0.48), pero elevadas entre primero y tercero y segundo y tercero (entre 0.77 y 1). Pese a que no ofrecen los valores de h² obtenidos mediante el modelo multicarácter, los que obtienen analizando separadamente cada parto indican que, excepto en un caso, no existen diferencias entre ellos.

Teniendo en cuenta el conjunto de todos los resultados publicados, se podría concluir que las h² de distintos órdenes de parto no son significativamente distintas entre sí, que es posible que el primero y el resto de partos no sean exactamente el mismo carácter, y que la r_g entre segundo y siguientes órdenes de parto está próxima a la unidad.

Gómez (1994) analiza datos de dos líneas de otra especie prolífica como son los conejos. Mediante varios modelos bicarácter, y considerando los tres primeros partos como caracteres distintos, obtiene resultados diferentes en ambas líneas; mientras en una de ellas el modelo de repetibilidad parece justificado, en la otra se observa heterogeneidad de h² y correlaciones genéticas sensiblemente inferiores a la unidad.

En definitiva, es difícil concluir si el determinismo genético de los distintos órdenes de parto es o no es el mismo en todos ellos. El objetivo de este capítulo es estimar las h² y las r_g entre distintos órdenes de parto en las dos poblaciones consideradas, y comparar los resultados con los obtenidos bajo el modelo con repetibilidad tradicionalmente empleado.

5.2.- Material y métodos

5.2.1.- Material animal.

Se utilizaron los datos de las poblaciones A y B para el número de nacidos vivos y totales por parto, descritas en el Capítulo [27].

Para analizar la variabilidad genética de cada orden de parto por separado, se consideraron todos los que tuviesen más de 1000 observaciones (Tabla 5-1). Para realizar un análisis conjunto de varios órdenes de parto, se tomaron los cinco primeros, por ser los de mayor interés económico. La Tabla 5-2 recoge la cantidad de información disponible en las distintas combinaciones de los cinco primeros partos en las dos poblaciones analizadas.

Población A						**Población B**				
		NV		**NT**			**NV**		**NT**	
Parto	**n**	μ	σ	μ	σ	**n**	μ	σ	μ	σ
1	9277	8.09	2.53	8.59	2.50	3730	9.18	2.40	10.11	2.45
2	6233	8.78	2.54	9.20	2.58	3158	9.58	2.45	10.35	2.58
3	4823	9.08	2.49	9.53	2.54	2766	10.13	2.38	11.03	2.54

[27] La información considerada de la población A fue la correspondiente al conjunto granja, y la de la B toda la disponible.

4	3744	9.17	2.43	9.63	2.48	2212	10.19	2.38	11.25	2.59
5	2978	9.11	2.45	9.65	2.51	1726	10.19	2.40	11.26	2.56
6	2317	9.11	2.45	9.70	2.54	1320	10.11	2.34	11.25	2.51
7	1748	8.83	2.53	9.49	2.62	--	--		--	
8	1246	8.53	2.63	9.21	2.73	--	--		--	
Todos	34417	8.68	2.55	9.19	2.59	16750	9.74	2.44	10.73	2.59

Tabla 5-1.- Número de registros (n), media (μ) y desviación típica (σ) del número de nacidos vivos (NV) y totales (NT) en las poblaciones A y B, para los órdenes de parto con más de 1000 registros, y de todos los partos conocidos.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

Población	A				B			
Partos	**2**	**3**	**4**	**5**	**2**	**3**	**4**	**5**
1	5828	4269	3166	2401	2935	2385	1759	1247
2		4536	3372	2559		2590	1939	1409
3			3551	2701			2104	1558
4				2802				1653

Tabla 5-2. - Número de registros conocidos para las distintas combinaciones de los cinco primeros partos, en las poblaciones A y B.

5.2.2.- Métodos.

Se realizaron tres tipos de análisis en la estimación de componentes de varianza: unicarácter, bicarácter y multicarácter (5 partos). El modelo empleado en el análisis multicarácter, y del cual se derivan los empleados en los análisis uni y bicarácter, fue el siguiente: Figura 5-1

De esta forma, el número de componentes de varianza a estimar fue 2 en los análisis unicarácter, 6 en los bicarácter y 30 en los multicarácter. En todos los análisis se utilizó toda la información familiar disponible en la construcción de la matriz de parentesco, y como factor fijo, únicamente la granja-año-estación en que tuvieron lugar los partos.

También se aplicó el modelo con repetibilidad utilizado en el Capítulo 3 (apartado 3.2.2.2.) para compararlo con los modelos anteriores, dado que se consideró un mayor número de partos por haber actualizado las bases de datos.

La estimación de los componentes de varianza se realizó mediante el programa DFREML (Meyer, 1991a, 1991b). Se consideró alcanzada la convergencia cuando los cambios entre cada uno de los componentes estimados en dos iteraciones consecutivas eran menores a 10^{-4}. No se realizó ninguna aproximación a los errores estándar de las estimaciones de los parámetros genéticos (excepto en los análisis unicarácter).

El proceso de estimación se interrumpía, por sí mismo, cuando se obtenían matrices singulares de (co)varianzas genéticas o residuales. Cuando esto sucedía, se reiniciaba el proceso a partir de las matrices correspondientes al punto de máxima verosimiltud previamente alcanzado.

En los análisis bicarácter se realizaron contrastes de la razón de verosimilitudes (LRT), con el fin de probar si los valores de correlación genética estimados por debajo de 0.7 eran estadísticamente distintos a la unidad. Al no ser posible calcular la verosimilitud cuando la correlación genética es la unidad, pues la matriz de (co)varianzas es singular, se tomó como hipótesis que la correlación estimada fuese distinta a un valor cercano a la unidad (0.97). Al trabajar estimando covarianzas, no se contrastaron directamente correlaciones, sino que se contrastó si a igualdad de varianzas (genéticas y residuales), la covarianza genética estimada era igual o distinta a la correspondiente a una correlación de 0.97.

Al margen del equipo informático señalado en el Capítulo 3, para realizar los análisis multicarácter considerando conjuntamente los cinco primeros partos se utilizó un CRAY YMP-232 bajo sistema operativo UNIX (CESCA, Generalitat de Catalunya).

5.3.- Resultados y discusión

5.3.1.- Análisis unicarácter.

En la Tabla 5-3 se recogen los resultados de los análisis realizados mediante el modelo con repetibilidad, al considerar tanto el número de nacidos vivos (NV) como totales (NT). Esos resultados coinciden con los obtenidos previamente en el Capítulo 3 para el carácter NV. Para NT los valores de h², sin que se puedan considerar significativamente distintos, son mayores que los estimados para NV. Lamberson (1990), revisando algunas estimaciones de h² publicadas, señala que el valor medio para NT es superior al de NV (0.10 frente a 0.07), pese a indicar que es posible que la diferencia se deba simplemente a haber analizado distintas poblaciones para cada carácter. Al analizar cada parto por separado (Tabla 5-4), también se obtienen valores de h² algo mayores para NT que para NV, pese a que en este caso las diferencias son aún menos significativas.

	NV				NT			
Población	h^2	c^2	σ^2_a	σ^2_c	h^2	c^2	σ^2_a	σ^2_c
A	0.06±0.01	0.07±0.01	0.40	0.43	0.07±0.01	0.07±0.01	0.42	0.47
B	0.05±0.01	0.09±0.01	0.30	0.53	0.06±0.01	0.09±0.01	0.37	0.58

Tabla 5-3.- Estimaciones de la heredabilidad, la proporción de varianza fenotípica explicada por el efecto permanente (±s.e.) y las varianzas aditiva y del efecto permanente, del número de nacidos vivos (NV) y totales (NT), según un modelo con repetibilidad.

	Población A				Población B			
	NV		NT		NV		NT	
Parto	h^2	σ^2_a	h^2	σ^2_a	h^2	σ^2_a	h^2	σ^2_a
1	0.04±0.01	0.24	0.05±0.02	0.29	0.05±0.02	0.28	0.04±0.02	0.21
2	0.05±0.02	0.30	0.05±0.02	0.37	0.01±0.02	0.08	0.02±0.02	0.16
3	0.05±0.03	0.29	0.07±0.03	0.30	0.09±0.03	0.50	0.10±0.03	0.64
4	0.06±0.03	0.30	0.06±0.03	0.41	0.09+0.02	0.12	0.10±0.03	0.66
5	0.04±0.03	0.32	0.06± 0.04	0.37	0.04±0.03	0.21	0.04±0.03	0.26
6	0.06±0.05	0.24	0.07±0.05	0.39	0.06±0.04	0.34	0.07±0.04	0.39
7	0.06±0.05	0.38	0.07±0.05	0.50				
8		--	0.03±0.05	0.20				

Tabla 5-4. - Estimaciones de la heredabilidad (±s.e.) y la varianza aditiva del número de nacidos vivos (NV) y totales (NT) de cada parto, según un análisis unicarácter.
n.e.: no se alcanzaron soluciones dentro del espacio paramétrico.

Si no existiese ningún tipo de selección y asumieramos que en todos ellos existe la misma determinación genética, los valores estimados bajo el modelo con repetibilidad deberían coincidir con los estimados en cada parto, al analizarlos separadamente. Si bien

parece cumplirse en la población A, en la población B se observan variaciones en los valores de h^2 estimados; se obtienen, respectivamente para NV y NT, valores de 0.01 y 0.02 en el segundo parto, y 0.09 y 0.10 en el tercero, frente a valores de 0.05 y 0.06 en el conjunto de todos los partos.

En otros trabajos también se ha observado esa disminución del valor de la h^2 en el segundo parto y su posterior aumento en el tercero. Legaut (1970), por ejemplo, estimó los valores de h^2 en los cuatro primeros partos, en 0.08, 0.03, 0.15 y 0.22 respectivamente. Avalos (1985) halló valores similares: 0.08, 0.08, 0.13 y 0.17, para los cuatro primeros partos, respectivamente. Sin embargo, Haley *et al.* (1988) concluyen, REVISANDO valores publicados, que pese a existir variaciones entre las estimaciones de h^2 en distintos partos, no pueden considerarse significativas. Una causa que permite justificar esas variaciones es el mayor error que se comete en órdenes de parto superiores, debido a disponer de menos información. Otra causa, posiblemente de mayor importancia, puede ser la eliminación de las cerdas en base a la prolificidad del último parto realizado, o en base a una serie de factores productivos genéticamente relacionados con ella, como puede ser fertilidad, facilidad al parto, aptitud maternal, etc. Si se realiza una selección de ese tipo, se verían afectadas las estimaciones de h^2 obtenidas al analizar, independientemente, segundo y posteriores partos. En este sentido, Irgang *et al.*, (1994), en un análisis similar al nuestro, en el que hallan una importante reducción de la h^2 del tercer parto de la población Landrace que analizan, argumentan que esa reducción posiblemente se deba a la eliminación de las cerdas que en el segundo parto mostraron una baja prolificidad.

De la misma forma se podrían explicar algunas de las diferencias observadas en las h^2 estimadas en la población B, como son las bajas h^2 del segundo parto para NV y NT, y del cuarto parto para NV. Es más difícil encontrar una explicación satisfactoria para las altas h^2 de NT en tercer y cuarto parto, y de NV en tercer parto, que responda a razones de manejo. Se puede pensar en la existencia de heterogeneidad genética entre la prolificidad de distintos partos de una misma cerda. Un análisis multicarácter del conjunto de varios partos, cuyos resultados se recogen en el siguiente apartado, es mucho más informativo sobre el origen de las diferencias observadas.

5.3.2.- Análisis multicarácter.

Los valores de heredabilidad obtenidos en los análisis realizados en las combinaciones parto a parto, entre los cinco primeros (un total de 10), se sitúan entre 0.03 y 0.11 para ambos caracteres y ambas poblaciones (Tabla 5-5), tal como sucedía en los análisis unicarácter.

		NV					NT				
	Parto	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Población A	**1**	**0.04 (34)**	0.92	0.91	0.70	0.82	**0.04 (4-5)**	0.90	0.93	0.67	0.71
	2	*0.14*	**0.05 (4-6)**	1.00	0.97	0.98	*0.14*	**0.06 (5-7)**	0.89	0.92	0.79
	3	*0.11*	*0.12*	**0.05 (4-6)**	0.92	0.95	*0.12*	*0.12*	**0.05 (4-6)**	1.00	0.92
	4	*0.07*	*0.12*	*0.15*	0.05 (4-5)	1.00	*0.08*	*0.13*	*0.17*	**0.06 (4-7)**	0.95
	5	*0.08*	*0.11*	*0.16*	*0.18*	**0.06 (6-6)**	*0.08*	*0.16*	*0.17*	*0.18*	**0.06 (6-6)**
Población B	**1**	**0.06** ***(3-9)***	*0.87*	*0.47*	*0.92*	*0.98*	***0.04*** ***(4-5)***	*0.95*	*0.53*	*0.75*	*0.90*
	2	*0.12*	**0.05 (2-7)**	0.83	0.97	n.e.	*0.12*	**0.04 (2-6)**	0.9o	0.86	n.e
	3	*0.13*	*0.13*	**0.09 (7-10)**	0.86	0.44	*0.13*	*0.13*	**0.10 (9-11)**	0.75	0.47
	4	*0.12*	*0.19*	*0.19*	**0.05 (3-6)**	0.93	*0.12*	*0.19*	*0.19*	**0.09** ***(6-11)***	0.89
	5	*0.15*	*n.e.*	*0.12*	*0.17*	**0.05 (3-6)**	*0.15*	*n.e.*	*0.12*	*0.17*	**0.06 (4-7)**

Tabla 5-5.- Estimaciones medias de la heredabilidad (y valores mínimos y máximos en %) -en la diagonal-, correlaciones genéticas -sobre la diagonal-, y correlaciones residuales -bajo la diagonal-, del número de nacidos vivos (NV) y totales (NT) en los cinco primeros partos, según un modelo bicarácter.
n.e.: no se alcanzaron soluciones dentro del espacio paramétrico.

Las diferencias que existían entre los resultados unicarácter de los distintos órdenes de parto en la población B, se mantienen pero de forma atenuada. Los valores más bajos que se obtenían (2º y 4º parto en NV y 2º en NT) son en los análisis bicarácter mayores, y próximos al valor del modelo con repetibilidad. De hecho, si existe selección de las cerdas en base únicamente a la información de su parto anterior, las estimaciones univariantes estarían teóricamente infraestimadas, y su valor aumentaría al realizar análisis bicarácter. Si atendemos a los errores estándar aproximados de los análisis univariantes, no se puede concluir que las h^2 estimadas en uno y otro análisis sean distintas entre sí.

De todas formas, lo más interesante de estos análisis son las correlaciones genéticas estimadas, pues de ellas depende fundamentalmente la bondad de la hipótesis de homogeneidad genética entre partos. Los valores estimados son muy elevados, sobre todo en la población A, en la que tanto para NV como NT, todos los valores se sitúan por encima de 0.7 excepto para los partos 1º-4º. En la población B, existen más valores por debajo de 0.7: las combinaciones entre el 1º-3º y 3º-5º, con un valor en torno a 0.5, y prácticamente iguales para NV que para NT. No obstante, tras realizar un contraste LRT para ver si esas 6 correlaciones por debajo de 0.7 equivalían a una correlación próxima a la unidad (0.97), se obtuvo que sólo las correlaciones entre los partos 1º-3º se podían considerar, para ambos carácteres, distintas a uno. En la población B no siempre fue posible estimar los parámetros genéticos, posiblemente por un valor de correlación genética entre partos muy próximo a la unidad.

Si en realidad existiese selección de las cerdas en base a sus registros previos, las estimaciones de los parámetros genéticos obtenidas mediante análisis bicarácter podrían estar sesgadas (Sorensen y Kennedy, 1984a), de forma que se debería realizar un análisis multicarácter considerando los cinco partos conjuntamente. En la Tabla 5-6 se ofrecen los valores de las componentes de (co)varianza estimadas según ese modelo para NV y NT en la población A. En la población B se alcanzaron, repetidamente, valores fuera del espacio paramétrico, por lo que pese a iniciar el proceso iterativo con distintos valores iniciales, no se pudieron obtener soluciones que se pudiesen considerar apropiadas.

	NV					NT				
Parto	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
1	**0.02** **2.50**	0.74	0.55	0.34	0.57	**0.02** **2.46**	0.72	0.56	0.67	0.16
2	*0.14*	**0.02** **2.50**	0.55	0.40	0.57	*0.14*	**0.03** **2.53**	0.54	0.35	0.20
3	*0.11*	*0.14*	**0.02** **2.46**	0.22	0.58	*0.12*	*0.13*	**0.02** **2.50**	0.57	0.35
4	*0.07*	*0.13*	*0.14*	**0.04** **2.42**	0.31	*0.09*	*0.15*	*0.19*	**0.03** **2.47**	0.23
5	*0.13*	*0.14*	*0.17*	*0.17*	**0.04** **2.44**	*0.08*	*0.12*	*0.17*	*0.15*	**0.06** **2.47**

Tabla 5-6.- Estimaciones de la heredabilidad y desviaciones típicas fenotípicas -en la diagonal-, correlaciones genéticas -sobre la diagonal-, y correlaciones residuales -bajo la diagonal-, del número de nacidos vivos (NV) y totales (NT) en los cinco primeros partos de la población A, según un modelo multicarácter.

Los valores medios estimados de h^2 (0.03 para NV y NT) son más bajos que los estimados mediante los análisis unicarácter (0.05 para NV y 0.06 para NT), o bicarácter (0.05 para NV y NT); lo mismo sucedió respecto a la rg cuyo valor medio es de 0.48 y 0.44 para NV y NT frente a los valores medios de 0.92 y 0.87 para NV y NT, en los análisis bicarácter. Por contra, los valores de correlación fenotípica que se deducen de los valores de la tabla, se sitúan entre 0.07 y 0.18, valores del mismo orden que los hallados en los análisis bicarácter, y concordantes con la repetibilidad estimada en el análisis unicarácter (0.14).

La Figura 5-2 recoge la evolución de logaritmo de la verosimilitud ($\log_e L$). En la resolución del modelo se partió de unos valores medios iniciales de h^2 y rg de 0.14 y 0.09 respectivamente. Al final del proceso de iteración la curva del $\log_e$ L era prácticamente plana. Esa evolución, aunque mucho más lenta, es semejante a la del $\log_e L$ del modelo de repetibilidad (Figura 5-3), siendo indicativo de que las estimaciones obtenidas pueden corresponder al máximo de la función. No obstante, al observar el valor del $\log_e$ L correspondiente a los distintos valores de h^2 o rg evaluados, no es evidente que se haya alcanzado el valor máximo de verosimilitud. Tomando como ejemplo la h^2 del tercer parto para NV (Figura 5-4) y comparando su perfil de verosimilitud con el de la h^2 del modelo de repetibilidad (Figura 5-5) se observa que no existen puntos de menor $\log_e L$ a ambos lados del punto de mayor verosimilitud. Si en vez de fijarnos en la estima de h^2 lo hacemos en la de rg , vemos como se produce la misma situación (Figura 5-6). No se tiene la seguridad de haber obtenido la estimación máximo-verosímil.

Spilke y Groeneveld (1994) señalan varios inconvenientes sobre el tipo de análisis bicarácter que se ha realizado, al compararlo con un análisis multicarácter. De entre ellos cabe destacar el ya mencionado de poder obtener estimaciones que no tienen en cuenta la selección realizada en base a registros de un tercer carácter, el hecho de tener que realizar

muchos análisis bicarácter frente a un único análisis multicarácter (en nuestro caso 10 frente a 1) y finalmente, la posibilidad de obtener matrices de (co)varianzas no definidas positivas, debiendo recurrir a técnicas de *'bending'*, basadas en concentrar los valores propios de una matriz en torno a su valor medio (ver p.ej. Ducos, 1994). En nuestro caso, todas las matrices de (co)varianzas estimadas para los cinco carácteres fueron definidas positivas, por lo que, si fuese necesario, se podrían utilizar directamente en un proceso de evaluación genética multicarácter.

Por contra, realizar varios análisis bicarácter puede ser recomendable para ciertos modelos. Meyer (1991b) ya indica que la aplicación del programa DFREML a un modelo con cinco caracteres no ha sido contrastada. Spilke y Groeneveld (1994) al analizar conjuntamente, mediante ese programa, cinco carácteres de gallinas de puesta con una única componente aleatoria (animal), no logran estimar las 30 componentes de varianza implicadas, por lo que se ven obligados a realizar 10 análisis bicarácter.

Ante estos inconvenientes, y la falta de seguridad de haber alcanzado una estima máximo-verosímil, se debería haber repetido el proceso partiendo de distintos puntos iniciales. El alto coste que suponía el proceso (más de 10 horas de CPU en un CRAY YMP-232) impidió actuar de ese modo, por lo que se optó por comprobar si existía una fuerte dependencia entre los valores iniciales y las soluciones alcanzadas en una población simulada. Esa población se simuló con la misma estructura que las simuladas en los Capítulos anteriores, considerando cinco caracteres sin correlación genética ni residual entre ellos, y en ausencia de selección (Figura 3-5, con una población base de 50 machos y 1000 hembras, y nh1=1 y nh2=2). Se estimaron los componentes de varianza usando un modelo multicarácter partiendo de tres conjuntos distintos de valores iniciales, y se consideró el mismo criterio de parada que el usado en el análisis de los datos de la población A. En ninguno de los tres casos se obtuvo como solución el conjunto de valores con que se simuló, y en cada uno, las soluciones obtenidas fueron distintas y dependientes de los valores iniciales (con una correlación de 0.995). No pasó lo mismo al estimar de igual forma los componentes bajo un modelo bicarácter, pues independientemente al punto de partida se alcanzaban siempre, como solución, los valores simulados.

Los actuales algoritmos basados en la función de verosimilitud no aseguran alcanzar soluciones precisas en modelos multicarácter, fundamentalmente si es grande el número de componentes a estimar (Groeneveld y Kovac, 1990). Misztal (1994) señala como causas principales la precisión de cálculo de la función de verosimilitud y las propiedades de los métodos utilizados en su maximización. La precisión en el cálculo de la función de verosimilitud disminuye al aumentar el tamaño de las ecuaciones del modelo mixto y la dependencia lineal de los caracteres implicados. Los métodos utilizados en la maximización numérica de la función de verosimilitud son lentos, pero los libres de derivadas son poco precisos, pues la función de verosimilitud es más plana cuanto mayor es el número de parámetros a estimar. En futuros análisis multicarácter, que implicasen estimar numerosos componentes de varianza, se podría pensar en utilizar algoritmos basados en las primeras derivadas, como por ejemplo el EM o el AI-REML (Madsen *et al.*, 1994), este último basado en las primeras derivadas y una aproximación a las segundas derivadas, realizada calculando la media de las matrices de información observada y esperada.

Al margen de los posibles problemas de las estrategias de estimación, nos interesa conocer hasta que punto los resultados obtenidos en los distintos análisis son coherentes entre sí. Van Vleck y Gregory (1992) muestran como la h^2 bajo un modelo de repetibilidad se infraestima si el modelo de repetibilidad es incorrecto, concretamente por asumir correlaciones genéticas de 1 cuando en realidad son inferiores. En ese caso, la h^2 del modelo

de repetibilidad es aproximadamente el producto de la h² media y la rg media de las estimaciones obtenidas bajo un modelo multicarácter, valor que podemos llamar h² ponderada (h^2_w)[28].

En nuestro caso, el modelo de repetibilidad ofrecía estimaciones de h² de 0.06 y 0.07 para NV y NT en la población A y 0.05 y 0.06 en la población B. Los valores medios de h² obtenidos en los análisis bicarácter no son mayores en la población A (0.05 para NV y NT) por lo que se puede pensar que la rg media es uno, cumpliéndose la hipótesis de identidad genética. Los resultados del análisis multicarácter conjunto indican unas h² ponderadas (0.01 para NV y NT) que se alejan más de los valores de h² estimados mediante el modelo con repetibilidad, que según los modelos bicarácter.

En la población B, los valores medios de h² y rg en los análisis bicarácter son para NV y NT respectivamente 0.06 y 0.81, y 0.07 y 0.78. Es decir, los valores de h² del modelo de repetibilidad son ligeramente inferiores y coinciden con los valores de h² ponderada del modelo de heterogeneidad genética (0.05 y 0.06). Se podría pensar que el modelo de repetibilidad no se ajusta en la población B, aunque como se vió sólo una de las 10 correlaciones estimadas puede considerarse que es distinta a la unidad, la correspondiente al primer y tercer parto. Es difícil encontrar una explicación de por qué esa correlación se separa de la unidad. Irgang *et al.* , (1994) argumentan que la existencia de efectos maternos puede ser una causa del bajo valor de correlación genética que obtienen entre el primer y segundo parto de una de las poblaciones que analizan. En nuestra población no se puede pensar en una explicación de ese tipo pues, como se vió en el capítulo anterior, los efectos maternos no parecen influir sobre el carácter. Parece más lógico pensar en problemas asociados a la estimación o a los datos disponibles que en razones de tipo genético. Gómez (1994) obtiene, en una de las dos poblaciones de conejos que analiza, unos resultados semejantes a los de la población B; las correlaciones genéticas entre los tres primeros partos son prácticamente la unidad a excepción de la correlación entre primer y tercer parto. Sin embargo, pese a esos resultados concluye que el modelo de repetibilidad parece más adecuado, dado que no observa heterogeneidad de heredabilidades entre partos.

[28] $h^2_w = h^2 \bullet r_g$

6.- Implicaciones de la heterogeneidad genética entre partos

6.1.- Introducción

El interés de conocer si existe heterogeneidad genética entre la prolificidad de diversos partos de un mismo animal viene fundamentalmente determinado por las repercusiones que puede tener aplicar incorrectamente un modelo con repetibilidad en un proceso de selección. Esas repercusiones se han analizado en varios trabajos bajo condiciones de población infinita o infinitas repeticiones de un proceso de selección.

Tartar (1981) y Tartar y Bolet (1984) analizan, mediante un índice de selección individual, las implicaciones de utilizar un modelo de repetibilidad si las correlaciones genéticas entre partos no son la unidad. Considerando los cinco primeros partos, una h^2 de 0.1 y el caso extremo de correlaciones nulas, la pérdida de respuesta sería de un 9%, mientras que la esperanza de respuesta se vería sobrestimada en un 300%. García *et al.*, (1989) considerando dos partos y distintos valores de correlación genética evalúan los riesgos de sobrestimación de las tasas de progreso genético esperado debidos a la aceptación de la hipótesis de identidad genética. Utilizando un índice de selección que considera la madre, 2 hermanas y 12 medias hermanas de la madre, y 3 hermanas y 12 medias hermanas del padre, obtienen sobrestimaciones de más de un 30% cuando la correlación genética es 0.4. Gómez (1994) simulando 10 generaciones de selección mediante BLUP, y considerando 3 partos, una h^2 de 0.1 y una r^g de 0.5, halla una respuesta un 30% inferior a la que se obtendría si la r^g fuese igual a 1. No parece tener tanta importancia la heterogeneidad de h^2 entre partos, pues para coeficientes de variación del 40% en tres partos con r^g=1, en ese mismo trabajo la respuesta sólo disminuye en un 6%.

Estos resultados indican que la pérdida de respuesta puede ser lo suficientemente importante como para rechazar el modelo de repetibilidad en la evaluación genética por prolificidad, si en realidad no existe homogeneidad genética entre partos. No obstante, seleccionar mediante un modelo multicarácter plantea otra serie de problemas, pues es necesario estimar más parámetros genéticos, emplear más recursos computacionales en la evaluación genética, y establecer un agregado económico que es posible que sólo sea válido a corto o medio plazo, pues pueden aparecer nuevos objetivos de selección que lo modifiquen.

Vandepitte y Hazel (1977) analizan el efecto de los errores en la asignación de los pesos económicos sobre la precisión de los índices de selección. Al analizar por separado el efecto de los errores en un único peso económico hallan que errores de hasta un 50% conllevan una pérdida de respuesta menor al 1%, pese a que la respuesta teórica esperada se puede ver considerablemente sobrestimada (hasta cerca de un 30%). Al analizar conjuntamente el efecto de errores cometidos al azar en todos los pesos económicos del agregado, obtienen menores sobrestimaciones de la respuesta, pero valores similares de pérdida, que incluso para errores del 100% no supera el 15%. Posteriormente, Weller (1994) confirma el escaso efecto que tienen pequeños cambios en los pesos económicos sobre la eficacia de la selección, concluyendo que son necesarios errores muy grandes para que afecten de forma importante.

Ducrocq (1994) señala el uso de parámetros genéticos posibles pero incorrectos como uno de los problemas importantes asociados a las predicciones genéticas a través de modelos multicarácter. Schaeffer (1984) analizó este caso, a través del efecto que tenía utilizar estimaciones incorrectas de los parámetros genéticos sobre la Varianza del Error de Predicción (PEV) de las predicciones obtenidas mediante un modelo mixto bicarácter. Obtuvo que el aumento de la PEV estaba casi directamente relacionado con el valor absoluto de la diferencia entre los errores cometidos en las correlaciones genéticas y ambientales Figura 6-1, y que valores absolutos de esa diferencia menores de 0.5, implicaban un aumento

de la PEV menor del 5%.

En la actualidad existen programas informáticos que permiten obtener estimaciones REML en modelos multicarácter con numerosas componentes de varianza (Misztal, 1994), pero tal como se vio en el Capítulo anterior, no aseguran cometer errores lo suficientemente pequeños como para ignorarlos en la decisión de qué modelo de evaluación emplear. Visscher (1994) indica que no parece recomendable incluir en un índice de selección caracteres cuyos parámetros genéticos estén sujetos a grandes errores. Van der Werf *et al.* (1992) señalan que los caracteres con baja h^2 son más sensibles a este tipo de errores, aunque también indican que el Modelo Animal puede atenuar su efecto si se dispone de suficiente cantidad de información.

Clásicamente, la genética cuantitativa ha ignorado la existencia de subjetividad en la mayor parte de decisiones que se deben de tomar en mejora animal, pero es difícil asumir que nuestras decisiones están exentas de subjetividad. Por ejemplo, respecto al modelo de evaluación, hemos visto que la aplicación del modelo de repetibilidad puede implicar pérdidas considerables de respuesta si se incumple la hipótesis de identidad genética entre partos; que los errores en los pesos económicos, aunque de forma poco importante, afectan a la respuesta; que es difícil estimar de forma precisa parámetros genéticos bajo un modelo multicarácter; que los errores en esos parámetros afectan la respuesta especialmente para caracteres poco heredables; que los pesos económicos pueden variar a largo plazo; etc... Esa subjetividad se puede desconsiderar a medida que nos acercamos a condiciones de población infinita o infinitas repeticiones de un proceso de selección, pues el riesgo que encierra no tomar la decisión adecuada tiende a desaparecer (Smith y Hammond, 1987). También se puede desconsiderar cuando para todas las decisiones el riesgo es el mismo. Sin embargo, nuestro interés suele estar centrado en una única repetición de un proceso de selección, realizado en base a una población de tamaño finito y con distinto grado de riesgo en cada decisión. En estos casos no parece justificado tomar la decisión sobre qué modelo de evaluación genética utilizar, exclusivamente en base a la tasa esperada de progreso genético, y es útil emplear elementos de la teoría de la decisión.

Teoría de la decisión.

El proceso de toma de decisiones se puede modelizar, desde un enfoque clásico, en base a una serie de elementos que permiten buscar de entre todas las decisiones posibles, la que ofrece la consecuencia óptima. Estos elementos son (Prieto, 1973) :

1- El conjunto de los posibles estados de la naturaleza ($\boldsymbol{\theta}$)**.** Es la realidad y por tanto desconocida. Existe incertidumbre que se intenta expresar en términos de probabilidad ($p(\boldsymbol{\theta})$).

2- El conjunto de las decisiones posibles ($\boldsymbol{\delta}$)**.** Sobre ellas siempre existe una actitud previa, una preferencia que se expresa asignándoles un orden.

3- El conjunto de las consecuencias ($\boldsymbol{\omega}$) de las combinaciones de $\boldsymbol{\theta}$ y $\boldsymbol{\delta}$ para un determinado criterio. Esas consecuencias se han medido tradicionalmente en términos de 'valor'.

4- La relación entre $\boldsymbol{\omega}$, $\boldsymbol{\theta}$ y $\boldsymbol{\delta}$ [$\boldsymbol{\omega} = \boldsymbol{\varphi}(\boldsymbol{\theta}, \boldsymbol{\delta})$], función que establece el valor de las consecuencias de cada decisión para cada posible estado de la naturaleza[29] tabla 6-1..

[29] Es habitual representar estos elementos matricialmente, a través de una matriz de decisión: Ver tabla 6-1.

	$p(\theta_1)$	$p(\theta_2)$	.....	$p(\theta_n)$	← conocimiento existente sobre la realidad
	θ_1	θ_2	.....	θ_n	←estados de la realidad
δ_1	$\omega_{11}=\varphi(\theta_1\ \delta_1)$	$\omega_{12}=\varphi(\theta_2,\delta_1)$	.....	$\omega_{1n}=\varphi(\theta_n\ ,\delta_1)$	←consecuencias de cada decisión según cada estado de la realidad
δ_2	$\omega_{21}=\varphi(\theta_1\ \delta_2)$	$\omega_{22}=\varphi(\theta_2\ \delta_2)$		$\omega_{2n}=\varphi(\theta_n\ \delta_2)$	
:	:	:	:	:	
δ_m	$\omega_{m1}=\varphi(\theta_1,\delta n)$	$\omega_{m2}=\varphi(\theta_2,\delta_n)$	.....	$\omega_{mn}=\varphi(\theta_n\ \delta_n)$	

decisiones posibles

BIBLIOTECA VIRTUAL

Tabla 6-1

A partir de esos cuatro elementos asigna un valor a las consecuencias de cada decisión [υ], que depende de la probabilidad de ocurrencia de cada estado de la naturaleza [$p(\boldsymbol{\theta})$]. En función del conocimiento que se tenga de esas probabilidades se habla de *ambiente determinista*, *ambiente de riesgo* y *ambiente de incertidumbre*.

Se habla de ambiente determinista cuando existe un único estado de la naturaleza posible [$p(\boldsymbol{\theta})=1$], por lo que no existe incertidumbre. Es el proceso de toma de decisiones más sencillo, pues la decisión se basa directamente en $\boldsymbol{\omega}$, pero es el menos frecuente.

Los ambientes de riesgo hacen referencia a la existencia de cierto grado de incertidumbre. Existen varios estados de la naturaleza posibles pero se conoce algo sobre su ley de probabilidades [$p(\boldsymbol{\theta})$], de forma que somos capaces de calcular $\boldsymbol{\upsilon}$. A diferencia de lo que ocurre en ambiente de certidumbre $\boldsymbol{\omega}$ no es único; se pueden asumir dos situaciones, que $\boldsymbol{\omega}$ sea una variable discreta[30] o continua[31]. Tomarla como discreta implica asumir que los estados de la naturaleza son finitos y obliga a explicitar la ley de probabilidades, mientras que considerarla como continua implica la existencia de infinitos estados de la naturaleza, obligando a asumir su tipo de distribución.

Los ambientes de incertidumbre se dan cuando, exitiendo también varios estados posibles de la naturaleza, no se conoce nada sobre la probabilidad de que sucedan, de forma que no se puede conocer $\boldsymbol{\upsilon}$.

Tanto $\boldsymbol{\omega}$ como $\boldsymbol{\upsilon}$ se pueden cuantificar como ganancias, pérdidas, ganancias relativas, etc... Asociada a cada caso hay matriz de decisión distinta, conocida como matriz de ganancias, matriz de pérdidas, matriz de ganancias relativas, etc... Es la medida más sencilla e

[30] $\boldsymbol{\upsilon_i}= p(\theta_1)\ \omega_{i1}+ p(\theta_2)\ \omega_{i2}+...,.. + p(\theta_n)\ \omega_{in}\quad 1\leq i\leq m$

[31] $\boldsymbol{\upsilon_i} = \int^{+\infty}_{-\infty} f(\omega_i)\ d\omega_i$,siendo $f(\omega_i)$ la función de densidad de la variable aleatoria $\omega_i=(\omega_{i1}, \omega_{i2},...,.., \omega_i)\quad 1\leq i\leq m$

intuitiva, pero presenta la desventaja de no tener en cuenta las características del decisor, por lo que es preferible cuantificar las consecuencias de cada decisión en términos de utilidad.

La utilidad es el criterio que se considera que los decisores toman para evaluar distintas alternativas, escogiendo aquella que maximiza la utilidad esperada (Kaylen *et al.*, 1987). Las bases de la teoría de la utilidad se pueden resumir en buscar una función de utilidad (U) que incorporando elementos subjetivos del decisor, permita pasar de una matriz de decisión de, por ejemplo, ganancias, a una matriz de utilidades esperadas sobre la cual calcular la utilidad esperada de cada decisión.

Para que una función se pueda considerar función de utilidad se deben cumplir una serie de supuestos (Romero, 1993), como por ejemplo, que sea monótona creciente, de forma que a medida que aumente el valor de una consecuencia (p.ej. ganancia monetaria) también aumente su utilidad (Mallo *et al.*, 1987). Berger (1985) formaliza los pasos a realizar en el proceso de construcción de las funciones de utilidad para cada individuo, pero dada la complejidad que supone tratar separadamente a cada individuo, se suele trabajar con formas analíticas que permitan expresar rasgos comunes en la mentalidad de los decisores.

Se dice que existen tres actitudes ante la decisión: de aversión al riesgo, de preferencia por el riesgo y de indiferencia ante el riesgo. La Figura 6-2 es la ilustración clásica de las tres actitudes cuando se relaciona "valor" y "utilidad",y en ella se puede ver como las funciones de aversión al riesgo son siempre cóncavas, las de preferencia por el riesgo convexas y las de indiferencia no son más que funciones lineales. La actitud habitual de los decisores es de aversión al riesgo, y es esta actitud la que se conoce como "comportamiento racional ante la decisión".

Existen dos formas fundamentales de funciones analíticas de utilidad, los polinomios de grado n y las funciones exponenciales (Kaylen *et al.*, 1987; Mallo *et al.*, 1987).

Dentro de los polinomios de grado n[32] existe un caso particular que se conoce como función de Markovitz (Mallo *et al.*, 1987) o función de utilidad cuadrática (habitualmente expresada como $U(\omega) = \omega - \lambda\, \omega^2$)[33]. La primera y segunda derivada de $U(\omega)$ respecto a [34] indican que el valor máximo de λ es $\frac{1}{2}\, \omega^*$, siendo $\omega^* = \{\text{máx } \omega\}$. Si $\lambda > 0$, la función es cóncava por lo que expresa aversión al riesgo. Si $\lambda < 0$ la función es convexa, expresando afición al riesgo. Asumiendo que el comportamiento racional es de aversión al riesgo, el rango de valores de λ es $[0, \frac{1}{2}\, \omega^*]$.

De las muchas funciones que pertenencen a la familia de las exponenciales, podemos tomar la función de Freund [$U(\omega) = 1 - e^{-\lambda\omega}$] (Mallo *et al.*, 1987) como modelo de función de utilidad. La adición de constantes no afecta este tipo de funciones, por lo que se pueden encontrar expresiones como la utilizada por Wyler *et al.* (1994)[35] en el contexto de la mejora animal. Primera y segunda derivada de la función de Freund[36], permiten ver que la función es

---

[32] $U(\omega) = \lambda_n\, \omega^n + \lambda_{n-1}\, \omega^{n-1} + \ldots\ldots + \lambda_0$

[33] La esperanza de esta función es: $E[U(\omega)] = E[\omega - \lambda\omega^2] = E[\omega] - \lambda\, E[\omega^2] = E[\omega] - \lambda\, E[\omega]^2 - \lambda\, V[\omega]$

[34] $U'(\omega) = 1 - 2\lambda\omega$; $U''(\omega) = -2\lambda$

[35] $U(\omega) = -\, e^{-\lambda\omega}$, siendo su esperanza: $E[U(\omega)] = E[\omega] - (\lambda/2)\, V[\omega]$ (Wyler et al., 1994).

[36] $U'(\omega) = \lambda e^{-\lambda\omega}$; $U''(\omega) = -\, \omega\lambda^2 e^{-\lambda\omega}$

siempre cóncava y que λ puede tomar cualquier valor, a diferencia de la función de Markovitz en la que λ presenta un valor máximo.

En mejora animal, Meuwissen (1990, 1991) aplica funciones cuadráticas para decidir entre distintos diseños de esquemas de selección en vacuno de leche. Sin embargo, es posible que el uso de funciones cuadráticas no sea justificable. Smith y Hammond (1987), recogiendo la opinión de otros autores, indican que las funciones de utilidad cuadrática explican comportamientos de decisión poco realistas. Posteriormente Wyler *et al.* (1994) se inclinan por las funciones exponenciales basándose fundamentalmente en que en las funciones cuadráticas los coeficientes de riesgo no pueden tomar cualquier valor, a diferencia de las exponenciales, lo cual puede conducir a situaciones en que su dominio sea lo suficientemente estrecho como para no poder describir las actitudes de los mejoradores. Esta ventajosa característica es a su vez un inconveniente de las funciones exponenciales, pues hace que su determinación no sea tan obvia como en las funciones cuadráticas.

La aplicación de la teoría de la utilidad se ve limitada a los ambientes de decisión de riesgo, pues en ambientes de incertidumbre, al desconocer las probabilidades de ocurrencia de los posibles estados de la naturaleza, no se puede calcular la utilidad de cada decisión. Hablar de decisión en ambiente de incertidumbre sólo tiene sentido desde el análisis frecuentista. En análisis bayesiano si bien no se conocen las probabilidades de ocurrencia de los estados de la naturaleza, sí es posible atribuir probabilidades subjetivas - *a priori*-, de forma que seguiríamos en un ambiente de riesgo (Smith, 1988).

No obstante, si no se quieren establecer probabilidades subjetivas, por creer que estas no tienen sentido, por existir poca información previa o simplemente por criticismo de las probabilidades *a priori*, existen distintos criterios por los que se puede optar. De entre ellos, los más clásicos son (Mallo *et al.*, 1987; Ballesteros, 1983): el criterio de Wald o del pesimismo; el criterio de Savage o de pérdida de oportunidad; el criterio de Agrawal-Heady o del beneficio; el criterio maximax; el criterio de Laplace, de la equiprobabilidad o de la razón insuficiente; y el criterio de Hurwicz o de juego contra la naturaleza. Los tres primeros parten del principio de que una vez tomada una decisión la naturaleza siempre se opondrá a ella, pero ante ese hecho, adoptan distintas posturas, lamentándose por lo que se pierde (Wald), lamentándose por lo que se deja de ganar (Savage) y alegrándose por lo que se deja de perder (Agrawal-Heady). El criterio maximax es el único que parte del principio de que una vez se ha decidido, la naturaleza siempre favorece esa decisión, mostrándose absolutamente optimista y alegrándose por lo que se gana. En una situación intermedia se sitúa el criterio de Hurwicz, pues permite asumir un grado de optimismo/pesimismo variable. El único criterio que muestra indiferencia es el de Laplace, pues se enfrenta a la incertidumbre otorgando equiprobabilidad de ocurrencia a todos los posibles estados de la naturaleza.

Pese a que el proceso de toma de decisiones que hemos planteado se caracteriza por una gran solidez lógica, Romero (1993) indica que presenta serias debilidades empíricas, dado que los procesos reales de toma de decisiones rara vez se realizan en base a un único criterio. Para superar esas debilidades empíricas, a mediados de este siglo aparecieron las primeras ideas de un nuevo enfoque de los procesos de decisión: la teoría de la decisión multicriterio, que considera la existencia de varios criterios a tener en cuenta en la toma de decisiones (Romero, 1993).

No obstante, parece adecuado hacer una primera aproximación a la toma de decisiones en mejora animal tomando como único criterio el progreso genético, tal como lo plantean, por ejemplo, Smith y Hammond (1987), Rogers (1990), Meuwisen (1991) o Wyler

et al., (1994).

El primer objetivo de este capítulo será analizar el efecto que tiene, sobre la respuesta genética, integrar diversas fuentes de información familiar cuando se utiliza un modelo con repetibilidad y existe heterogeneidad genética entre partos. Un segundo objetivo será evaluar, desde un punto de vista genético, el interés de aplicar un modelo multicarácter en un experimento de selección hiperprolífica llevado a cabo en una de las poblaciones analizadas (Noguera*et al.*, 1994). Finalmente, y como tercer objetivo, se analizará ese mismo proceso de selección mediante simulación estocástica, usando distintos criterios y ambientes de decisión.

6.2.- Material y métodos

6.2.1.- Material.

6.2.1.1.- Material animal.

El experimento de selección hiperprolífica referido al establecer los objetivos de este Capítulo en la Introducción (Noguera *et al.*, 1994), está basado en los animales de la población A. Recordando este experimento, tal como se comentó en el Capítulo 3, en 1993 se seleccionaron las mejores 175 hembras, de las aproximadamente 3500 disponibles, y 1 macho de cada una de las 25 mejores camadas de la población A. El objetivo es comparar 150 hijas de la población seleccionada, una de cada hembra, con las hijas de una población control de 200 hembras y 15 machos tomados al azar; inicialmente se seleccionaron 175 hembras con la finalidad de asegurar las 150 hijas previstas en la comparación.

Por este motivo, para alcanzar los objetivos previstos en este Capítulo se utilizaron los datos de los cinco primeros partos de las hembras de la población A (Tabla 5-1, Capítulo 5), así como la estructura familiar conocida de toda la población (Tabla 3-1, Capítulo 3).

6.2.1.2.- Poblaciones simuladas.

Teniendo en cuenta la estructura familiar de la población A se simularon los datos de prolificidad de acuerdo a tres hipótesis genéticas:

- [HO]: hipótesis de homogeneidad genética entre partos, con los parámetros estimados en el Capítulo 3;
- [HE1]: hipótesis de heterogeneidad genética, con los parámetros estimados en el Capítulo 5 al analizar todos los partos conjuntamente;
- [HE2]: hipótesis de heterogeneidad genética, con unos parámetros arbitrarios que cumpliesen que la heredabilidad media ponderada de los cinco partos, definida en el apartado 5.3.2. (Van Vleck y Gregory, 1992), equivalía a la heredabilidad estimada bajo la hipótesis de homogeneidad.

La Tabla 6-2 recoge los valores utilizados en la simulación para cada una de las tres hipótesis.

HO	Homogeneidad genética (modelo animal de repetibilidad) a^2= 0.37, e^2 = 5.62 y c^2 = 0.52 (h^2=0.06)									
HE1	Heterogeneidad genética (modelo animal multícarácter)									
	G ao = 0. 10	0.08	0.06	0.05	0.08	Ro = 6.17	0.76	0.62	0.40	0.73
		0.12	0.07	0.07	0.09		6.15	0.78	0.79	0.77
			0.14	0.04	0.10			5.90	0.81	0.92
				0.23	0.07				5.62	0.99
					0.22					5.73
	H^2media = 0.03 , r g media = 0.5 (h^2 ponderada = 0.02)									
HE2	Heterogeneidad genética (modelo animal multicarácter)									
	Gao = 0.46	0.41	0.37	0.32	0.28	Ro = 6.05	0.57	0.61	0.66	0.70
		0.46	0.41	0.37	0.32		6.05	0.57	0.61	0.66
			0.46	0.41	0.37			6.05	0.57	0.61
				0.46	0.41				6.05	0.57
					0.46					6.05
	H^2 media = 0.07, r g media = 0.8 (h^2 ponderada de = 0.06)									

Tabla 6-2.- Hipótesis genéticas consideradas en la simulación de una población con una estructura equivalente a la población A.

Para la hipótesis HO, los valores genéticos de los animales base se simularon extrayendo valores de una distribución $N(0,I.\sigma^2_a)$, y los del resto de animales, dado que la consanguinidad media no alcanza ni el 1% (Tabla 3-1, Capítulo 3), se simularon de acuerdo a la expresión:

$$a_i = \frac{1}{2}(a_p + a_m) + \varnothing_i$$

donde:

a_i: valor genético del individuo i
a_p: valor genético de su padre
a_m: valor genético de su madre

$\varnothing_i$: efecto del muestreo mendeliano, distribuido según una $N(0,\frac{1}{2}.I.\sigma^2_a)$

Para las hipótesis de heterogeneidad genética se hizo uso de las propiedades de la transformación canónica, procediendo de forma análoga a la simulación de los efectos maternos que se comentó en el apartado 4.2.1.2. Así, los valores genéticos de los animales base se simularon extrayendo valores de una distribución $N(\underline{\mathbf{0}},I\otimes G_{ao})$ (siendo G_{ao} la matriz de (co)varianzas genéticas entre los cinco partos), y los del resto de animales se simularon de

acuerdo a la expresión:

$a_{ij}= ½ (a_{pj}+ a_{mj}) + ø_{ij}$

donde:

a_{ij}: valor genético del individuo i para el carácter j

a_{pj}: valor genético de su padre para el carácter j

a_{mj}: valor genético de su madre para el carácter j

$ø_{ij}$: efecto del muestreo mendeliano ø,distribuido según una N(**0**, I⊗G_{ao}), para el carácter j del individuo i

Teniendo en cuenta la distribución real de los registros dentro de los efectos fijos orden de parto y granja-año-estación, y tomando para esos efectos los valores estimados en la población mediante el modelo con repetibilidad descrito en el Capítulo 3, se simularon los valores fenotípicos, según las siguientes expresiones:

- para la hipótesis H0:

$y_{ijk}= \mu+ op_i+ gae_j+ a_k+ c_k+ e_{ijk}$

donde:

y_{ijk}: prolificidad correspondiente a la hembra k, el orden de parto i y la granja-año-estación j

μ: media del carácter

op_i: efecto del orden de parto (i=1,2,3,4,5)

gae_j: efecto granja-año-estación (j=1,...,123)

a_k: valor genético de la hembra k, previamente simulado.

c_k: efecto permanente en la hembra k; simulado extrayendo valores de una N(0,I.σ^2_c)

e_{ijk}: valor residual de la prolificidad correspondiente a la hembra k, el orden de parto i, y la granja-año-estación j; simulado extrayendo valores de una N(0,I. σ^2_e)

- para las hipótesis HE1 y HE2:

$y_{ijk}=_k+ gae_{ik}+ a_{jk}+ e_{ijk}$

donde:

y_{ijk}: prolificidad correspondiente al parto k de la hembra j y la granja-año-estación i

μ_k: media del parto k

gae_{ik}: efecto granja-año-estación (i=1,...,123), para el parto k

a_{jk}: valor genético de la hembra j, para el carácter k previamente simulado

e_{ijk}: valor residual de la prolificidad correspondiente al parto k de la hembra j y la granja-año-estación i; simulado extrayendo valores de una N(**0**,I⊗Ro) (siendo Ro la matriz de (co)varianzas residuales entre los cinco partos) y utilizando las propiedades de la transformación canónica.

6.2.2.- Métodos.

6.2.2.1.- Revisión de la hipótesis de identidad genética entre partos mediante índices de selección con información familiar.

Una primera comparación entre los modelos de repetibilidad y multicarácter se realizó en el contexto de la teoría general de los índices de selección (Hazel, 1943). Esta aproximación permite conocer las implicaciones de considerar incorrectamente el modelo de repetibilidad en una situación de tamaño poblacional infinito y ausencia de efectos fijos.

Concretamente, se amplió el trabajo de Tartar y Bolet (1984) a índices de selección con información familiar, evaluando la sobrestimación y la pérdida de respuesta genética esperada cuando se asume que la hipótesis de identidad genética es correcta, y en realidad no lo es. Para ello, se utilizaron distintos niveles de información familiar (Tabla 6-3) y distintas

hipótesis sobre el carácter (Tabla 6-4). En total se consideraron 16 casos para cada índice de la Tabla 6-3, casos que se recogen en la Tabla 6-5 (para algunos índices el número de casos es en realidad menor, pues no se ven afectados por ciertos supuestos). En todos ellos se calculó la respuesta exclusivamente vía hembra,
a partir de un diseño jerárquico en el que cada macho se apareaba con cuatro (o diez) hembras, y por simplicidad, sólo se consideraron los cuatro primeros partos de una cerda en el objetivo de selección.

índice	**Infonnación familiar considerada.**
1	P_i
2	P_i+P_d
3	$P_i+P_d+P_{hs}$
4	$P_i+P_d+P_{fs}$
5	$P_i+P_d+P_{fs}+P_{hs}$
6	P_d+P_{hs}

Tabla 6-3.- Indices con distinto grado de información familiar considerados en la comparación de los modelos de repetibilidad y multicarácter.
siendo: P_i: registros del candidato; P_d: registros de su madre; P_{fs}: registros de sus hermanas; P_{hs}: registros de sus hermanas de padre.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Parámetros	**Valores**
h^2 (igual para todos los partos)	0.09 (*)
Desviación típica fenotípica	a) 1º parto 2.7 ; resto 2.9 b) todos 2.85(**)
Correlación residual entre partos de una rnisma cerda	0.08(*)
Correlación residual entre partos de una cerda con los de su madre (r^{em}) y con los de sus hermanas (r^{eh}).	a)r^{em}=0 ;r^{eh}=0 b) r^{em} . = -0.1 (***); r^{eh} =0 c) r^{em} = 0 ;r^{eh} 0.05 (**)
Correlación genética entre partos (*) 1º y 2º parto 1º y 3º parto 1º y 4º parto 2º y 3º parto 2º y 4º parto 3º y 4º parto	 0.89 0.72 0.33 0.78 0.87 1.00
Nº hermanas, de cada cerda, con Partos	2
Nº hermanas de padre, de cada cerda, con partos	a) 9 b) 27
Nº partos de cada cerda.	a) 2 partos. b) 4 partos.

Tabla 6-4.- Valores asumidos de los distintos parámetros que intervienen en el cálculo de los índices de selección utilizados en el análisis de los modelos de repetibilidad y multicarácter.
(*) valores tomados de Haley *et* al.(1988); (**) valores tomados de Avalos y Smith (1987);(***) valores tomados de Alsing *et* al. (1980).

Casos	σ^{p1} (*)	σ^{pr}	Nº partos por cerda	Nº de medios hermanos	r^{em} (**)	r^{eh}
caso 1	2.7	2.9	2	9	00	
caso 2	2.7	2.9	2	9	-0.1	0.05
caso 3	2.7	2.9	2	27	0	0
caso 4	2.7	2.9	2	27	-0.1	0.05
caso 5	2.7	2.9	4	9	0	0
caso 6	2.7	2.9	4	9	0	0.05
caso 7	2.7	2.9	4	27	0	0
caso 8	2.7	2.9	4	27	0	0.05
caso 9	2.85	2.85	2	9	0	0
caso 10	2.85	2.85	2	9	-0.1	0.05
caso 11	2.85	2.85	2	27	0	0
caso 12	2.85	2.85	2	27	-0.1	0.05
caso 13	2.85	2.85	4	9	0	0
caso 14	2.85	2.85	4	9	0	0.05
caso 15	2.85	2.85	4	27	0	0
caso 16	2.85	2.85	4	27	0	0.05

Tabla 6-5- Relación de los casos analizados en la comparación de los modelos de repetibilidad y multicarácter, en función de los parámetros considerados.
(*)σ^{p1} desviación típica fenotípica del primer parto; σ^{pr} desviación típica fenotípica del resto de partos.
(**) r^{em} correlación residual entre partos de una cerda con los de su madre; r^{eh} correlación residual entre partos de una cerda col los de sus hermanas.

De forma análoga se procedió considerando una de las situaciones que Avalos y Smith (1987) analizaron para evaluar las ventajas de incorporar información familiar en la evaluación genética por prolificidad. Concretamente, se consideró un esquema basado en aparear 1 macho con 10 hembras, y conocer dos partos de 3 hembras de cada apareamiento. De las mejores camadas de primer parto se escogieron tres hembras y un macho como futuros reproductores[37]. Las camadas se seleccionaron mediante un índice que consideraba la información de la madre, la de la abuela paterna, la de las hermanas de la madre y del padre y la de las medias hermanas de la madre y el padre. También se consideraron los parámetros de

[37] Un esquema de estas características supone aplicar una intensidad de selección de 1.10 en las hembras y 1.78 en los machos. Para una h^2 de 0.10, una desviación típica de la prolificidad de 2.85 y considerando homogeneidad genética entre partos, la respuesta genética esperada es de 0.47 lechones por cerda y año (Avalos y Smith, 1987).

la Tabla 6-4 a excepción del valor de h², para el que se tomó 0.10 con el objeto de que los resultados fuesen perfectamente comparables con los obtenidos por Avalos y Smith (1987).

Se consideraron dos agregados económicos (H) a partir del valor aditivo (a) de cada uno de los cuatro primeros partos, y un vector de pesos económicos (α):

1) $H^1 = \alpha^1.\mathbf{a} = a^1 + 0.88\ a^2 + 0.72\ a^3 + 0.58\ a^4$. Los pesos económicos se derivaron de la pirámide de edades de las cerdas controladas por el programa de gestión técnica de explotaciones porcinas GTEP-IRTA, al ponderar cada parto según su frecuencia (Tartar y Bolet, 1984)[38].

2) $H^2 = \alpha^2.\mathbf{a} = a^1 + a^2 + a^3 + a^4$. En este caso, los pesos económicos se establecieron en función de la ganancia marginal que resulta al aumentar una unidad el tamaño de camada de cada parto[39].(Figura 6-3) (Avalos, 1985).

El valor del agregado económico H^i, para un individuo cualquiera, se estimó mediante las siguientes expresiones:

- bajo la hipótesis unicarácter:

$$I^u_i = [\Sigma^4_{j=1}(\alpha_{ij}/4)].\ \mathbf{C}(a,\mathbf{p'}).\mathbf{V}(\mathbf{p})^{-1}.\mathbf{p} \qquad i=1,2;\ j=1,..,4.$$

siendo:

p: vector de valores fenotípicos [dim(nx1); n=número de observaciones].

a : valor genético aditivo.

C(a,**p'**) : matriz de covarianzas entre el valor genético y los valores fenotípicos [dim(1xn)].

V(**p**)$^{-1}$: inversa de la matriz de (co)varianzas entre los valores fenotípicos [dim (nxn)].

α_i: vector de pesos económicos del agregado H^i (con elementos α_{ij}; j=1,..,4).

- y bajo la hipótesis multicarácter:
-

$$\mathbf{1}^m_i = \boldsymbol{\alpha_i}'.\ \mathbf{C}(\mathbf{a},\mathbf{p'}).\mathbf{V}(\mathbf{p})^{-1}.\mathbf{p}$$

siendo :

a: vector de valores genéticos aditivos del individuo [dim(4x1)];

C(**a,p'**) : matriz de covarianzas entre los valores genéticos y los valores fenotípicos [dim(4xn)].

Esas expresiones permitieron calcular para cada agregado la diferencia en respuesta esperada (R) y la pérdida de respuesta esperada (P) al aplicar el índice unicarácter (I_u), cuando el índice correcto era el multicarácter (I_m). Asumiendo una intensidad de selección constante, R y P vienen dadas, expresadas en porcentaje, por:

$$R_i\ (\%) = [(\sigma_{1ui} - \sigma_{1mi}) / \sigma_{1mi}] \bullet 100$$

$$P_i\ (\%) = [(r_{(Hi\ 1mi)} - r_{(Hi\ 1ui)}) / r_{(Hi\ 1mi)}] \bullet 100$$

siendo: $r_{(Hi\ 1ui)} = Cov\ (H_i, I_{ui}) / \sigma H_i.\ \Sigma\sigma_{Iui}$

$$r_{(Hi\ Imi)} = \sigma_{Imi} / \Sigma\sigma_{Hmi}$$

6.2.2.2.- Simulación de un proceso de selección hiperprolífica en la población A.

---

[38] De esta forma se tiene en cuenta que a los partos más avanzados llega menor número de cerdas. Sin embargo, no se tiene en cuenta que el tamaño de camada es menor en el primer parto que en el resto, y mayor en tercero y cuarto que en segundo (diferencias de 0.3 y 0.9 lechones nacidos vivos en los segundos y terceros-cuartos partos respecto al primero (Alfonso et al., 1991)).

[39] El peso económico del carácter j (α_{2j}) se calculó asumiendo que los costes de cada parto son independientes del tamaño de camada y del orden de parto, y que la h² era la misma para todos los partos, según:

Ver Figura 6-3

Se realizaron dos evaluaciones BLUP de la población A para nacidos vivos, considerando dos modelos genéticos: un modelo de repetibilidad con los parámetros genéticos estimados en el Capítulo 3, y un modelo multicarácter, tomando los valores que más se separaban de la hipótesis de identidad estimados en el Capítulo 5.

Una vez realizadas las evaluaciones, se simuló un proceso de selección basado en seleccionar el 1 o el 10% de las mejores hembras activas según esas evaluaciones. La comparación entre ambos modelos consistió en ver las diferencias en las pérdidas de respuesta esperada al seleccionar mediante un modelo cuando se considera el otro como correcto.

La respuesta esperada se calculó como la mitad de la diferencia entre el valor genético medio predicho para las hembras seleccionadas y la media del total de hembras activas; bajo el modelo de repetibilidad, se calculó considerando sólo los cinco primeros partos, y bajo el modelo multicarácter, considerando un agregado que otorgaba igual importancia a cada uno de esos cinco partos (agregado H_2 del apartado anterior). Este agregado también se comparó con el agregado obtenido al ponderar cada parto en función de su probabilidad de ocurrencia en el estrato comercial de la población ((1, 0.71, 0.61, 0.48, 0.42); agregado H_1 del apartado anterior, calculado con los datos de la población A).

La pérdida de respuesta se calculó como la diferencia entre la media de los valores genéticos de las hembras que se seleccionaban mediante un modelo y la media de los valores genéticos de las seleccionadas con el otro modelo. Posteriormente, esa pérdida se expresó como porcentaje sobre la respuesta esperada previamente calculada.

6.2.2.3.- Simulación de un proceso de selección hiperprolífica en una población simulada con la misma estructura que la población A.

6.2.2.3.1.- Simulación del proceso de selección.

Tras simular los datos de la población A como se vio en el apartado 6.2.1.2. se asignó el estado activo a aproximadamente 3500 hembras tomadas al azar de entre todas las hembras conocidas, para posteriormente seleccionar a las 150 con mejor valor genético predicho.

Se realizaron predicciones BLUP del valor genético de las hembras considerando toda la información genealógica disponible y utilizando tres modelos distintos:

* un Modelo Animal de repetibilidad (MR), idéntico al usado en la simulación bajo la hipótesis HO.

* un Modelo Animal multicarácter (MM1), idéntico al usado en la simulación bajo la hipótesis HE1.

* un Modelo Animal multicarácter (MM2), idéntico al usado en la simulación bajo la hipótesis HE2.

Bajo la hipótesis de heterogeneidad genética, se tomó como agregado económico el considerado en el apartado anterior que otorgaba la misma importancia económica a cada uno de los cinco primeros partos (H_2).

La selección de las hembras se realizó bajo la restricción de conocer padre y madre de la hembra seleccionada y no tomar más de 10 hembras hijas de un mismo macho, simulando el proceso llevado a cabo en la realidad.

Se simularon nueve procesos de selección de acuerdo a los tres modelos de evaluación considerados (MR, MM1 y MM2) y las tres hipótesis genéticas asumidas (HO, HE1 y HE2). En cada uno de ellos se calculó, para los valores simulados, la media de las hembras activas, la media de las hembras seleccionadas, y su diferencia. Cada proceso se simuló 10 veces, de forma que en total se simularon 90 procesos de selección, correspondiendo 30 a cada modelo de evaluación.

De esta forma, se tenían tres hipótesis genéticas (HO, HE1, HE2), tres decisiones de evaluación (modelos MR, MM1, MM2) y nueve valores de respuesta genética esperada, combinación de las tres hipótesis genéticas y los tres modelos de evaluación.

6.2.2.3.2.- Decisión del modelo de evaluación.

Una primera toma de decisión se realizó en ambiente de incertidumbre, asumiendo absoluta ignorancia sobre la probabilidad de ocurrencia de cada una de las tres hipótesis genéticas consideradas. En este caso se trabajó sobre la matriz de ganancias y se aplicaron los siguientes criterios para decidir que modelo de evaluación aplicar en el proceso de selección:

- Criterio de Laplace. Consiste en otorgar igual probabilidad a cada estado de la naturaleza y tomar la decisión que maximice el valor esperado.
- Criterio de Wald. Consiste en maximizar la mínima ganancia, dado que parte del principio de que una vez se haya tomado la decisión, la naturaleza se opondrá a ella.
- Criterio Maximax. Consiste en maximizar la máxima ganancia, dado que parte del principio de que una vez se haya tomado la decisión, la naturaleza siempre la favorecerá.
- Criterio de Savage. Consiste en minimizar la máxima pérdida en relación a la decisión de máxima ganancia.
- Criterio de Agrawal-Heady. Consiste en maximizar el mínimo beneficio en relación a la decisión de máxima ganancia.
- Criterio de Hurwicz. Consiste en tomar una decisión intermedia entre la consecuencias extremas estableciendo, previamente, un coeficiente de optimismo (α). Si llamamos Cm a la consecuencia más favorable y Cp a la más desfavorable, la decisión a tomar es la que hace máxima la cantidad: $Cm.\alpha + Cp.(1-\alpha)$

Una segunda toma de decisión se realizó en ambiente de riesgo; se asumió que no se conocía la verdadera hipótesis genética, pero sí se podían atribuir probabilidades. Para ello se trataron las hipótesis como si fuesen conocidas o desconocidas.

En el primer caso, en el que se consideró que se conocían las hipótesis genéticas, se atribuyeron las siguientes probabilidades subjetivas: HO: 0.45; HE1: 0.1; HE2: 0.45. Estas probabilidades expresaban nuestra credibilidad sobre las distintas hipótesis. A la hipótesis HE1, basada en unas estimaciones de parámetros genéticos que ofrecían muy poca credibilidad (como se discutió en el Capítulo 5), se le otorgó una baja probabilidad (0.1), atribuyendo igual probabilidad al resto de hipótesis (0.45).

En el segundo caso, en el que se consideró que no se conocían las hipótesis genéticas, se asumió que la respuesta, en los 30 procesos de selección simulados para cada modelo de evaluación, se distribuía siguiendo una normal. Esta es la forma habitual de proceder cuando la realidad es desconocida y no simulada, recurriendo a algún contraste para determinar si es aceptable o no asumir normalidad. En nuestro caso, para asegurar la normalidad de la distribución de la respuesta, se utilizó el contraste de Shapiro-Wilk (SAS Inc., 1987).

Para ambos casos, el criterio de decisión fue considerar como modelo de evaluación óptimo, aquel que presentase mayor utilidad. Para ello se calculó la utilidad de cada modelo mediante dos funciones (siendo R respuesta y V(R) su varianza):

1- una función cuadrática: $U= R - \lambda R^2$, determinada por un valor del coeficiente de aversión al riesgo calculado como la mitad del valor máximo, pues de esta forma nos situamos en una situación intermedia de aversión al riesgo. El valor máximo viene dado por la siguiente expresión (Meuwissen, 1991):

λ (máx.) = 1/[2.R(máx) + 4.{V(R(máx))}½]

2- una función exponencial: $U=-e^{-\lambda R}$, determinada por un valor de λ que viene dado por la utilidad de las respuestas extremas, R1 y R2, con varianzas V(R1) y V(R2), y tales que R1 > R2, según la siguiente expresión (Wyler *et al.*, 1994):

$\lambda = 2\ \{[R1*U(R2) - R2*U(R1)]\ /[V(R1)*U(R2) - V(R2)*U(R1)]\}$

El inconveniente de las funciones exponenciales es que la utilidad de las respuestas extremas no se conoce, por lo que se deben fijar arbitrariamente. En nuestro caso se consideró un cociente entre las utilidades de las respuestas extremas arbitrariamente menor que el cociente entre esas respuestas. Esta forma de proceder es justificable cuando se cumple que V(R1)>V(R2), tal como sucedió en nuestra situación.

Los valores esperados de ambas funciones vienen dados, como se vió en la Introducción, por:

- función cuadrática: $E[U]=E[R] - \lambda E[R]^2 - V[R]$
- función exponencial: $E[U]=E[R] - \frac{1}{2}(\lambda V[R])$

6.3.- Resultados y discusión

6.3.1.- Revisión de la hipótesis de identidad genética entre partos mediante índices de selección con información familiar.

En la Tabla 6-6 se presentan los resultados más relevantes, obtenidos al evaluar la pérdida y la sobrestimación de respuesta genética esperada, que implica asumir identidad genética entre partos cuando no se cumple (en la Tabla 6-7 se recogen todos los resultados). En ella se puede ver cómo aumenta la pérdida de respuesta esperada a medida que se incorpora más información familiar a los índices de selección. Para el conjunto de los índices construidos la pérdida de respuesta esperada que conlleva considerar la hipótesis de identidad genética entre partos, frente a la hipótesis multicarácter, oscila entre un 0 y un 7%. Las pérdidas son prácticamente nulas al utilizar un índice de selección individual; el valor máximo del 7% corresponde al índice que considera más fuentes de información familiar: el individuo, la madre, las hermanas y las hermanas de padre.

Información familiar considerada

	Nº partos	**Individuo**	**Individuo y madre**	**Individuo, madre y hermanas**	**Individuo, madre, hermanas y . heninanas de padre**
Pérdida (%)	**2**	1	1	1	2
	4	3	4	4	7
Sobrestimación (%)	**2**	19	18	18	17
	4	19	18	18	16

Tabla 6-6.- Porcentaje de pérdida y sobrestimación de respuesta genéfica es~ al aplicar distintos índices de selección bajo la hipótesis de homogeneidad genética entre partos, cuando ésta no es cierta.

Índice	**1**		**2**		**3**		**4**		**5**		**6**	
Agregado	**H_1**	**H_2**	**H_1**	**H_2**	**H_1**	**H_2**	**H_1**	**H_2**	**H_1**	**H_2**	**H_1**	**H_2**
Caso 1	1.10	1.42	1.76	1.30	0.67	1.16	1.52	2.01	1.42	1.89	2.02	2.44
	18.86	*22.59*	*18.67*	*22.26*	*18.28*	*21.92*	*17.11*	*21.82*	*17.93*	*21.41*	*18.24*	*21.81*
Caso 2			0.70	0.97	1.08	1.33	0.66	1.13	0.61	1.27	1.75	2.17
			18.54	*22.23*	*18.29*	*21.91*	*18.21*	*21.88*	*18.02*	*21.55*	*18.35*	*21.90*
Caso 3					2.04	2.52			1.34	1.97	2.11	2.45
					17.74	*21.22*			*17.48*	*20.89*	*17.34*	*20.86*
Caso 4					1.98	2.45			0.96	1.58	2.17	2.80
					17.78	*21.31*			*17.57*	*21.01*	*17.57*	*20.86*
Caso 5	1.32	1.21	2.44	2.44	4.64	4.25	4.85	4.83	5.64	5.79	3.62	3.89
	18.73	*18.20*	*18.16*	*17.57*	*16.97*	*16.46*	*1664*	*1617*	*15.94*	*15.49*	*17.01*	*1633*
Caso 6					4.70	4.87	4.71	4.88	2.95	2.94	2.23	2.33
					17.20	*16.0*	*17.64*	*17.03*	*16.85*	*16.28*	*17.36*	*1696*
Caso 7					4.84	4.83			4.92	4.91	4.76	4.75
					15.74	*15.29*			*15.08*	*14.68*	*14.94*	*14.46*
Caso 8					5.61	5.59			5.56	5.71	4.95	4.93
					16.05	*15.57*			*15,82*	*15.36*	*15.53*	*15.03*
Caso 9	0.55	0.85	1.01	1.30	2.02	2.54	0.65	0.00	2.43	2.92	2.68	3.11
	18.49	*22.25*	*18.33*	*21.95*	*17.97*	*21.65*	*17.82*	*21.47*	*17.64*	*21.17*	*17.98*	*21.54*
Caso 10			0.00	0.00	1.93	2.21	0.87	1.12	1.22	1.68	2.09	2.51
			18.22	*21.95*	*17.98*	*21.65*	*17.90*	*21.62*	*17.71*	*21.22*	*18.08*	*21.62*
Caso 11					1.22	1.67			2.10	2.55	2.89	3.24
					17.46	*20.97*			*17.13*	*20.66*	*17.13*	*20.54*
Caso 12					1.18	1.63			1.72	2.37	2.17	2.79
					17.59	*21.07*			*17.30*	*20.70*	*17.34*	*20.65*
Caso 13	3.08	3.06	3.67	3.66	4.08	4.07	3.77	3.76	3.94	3.92	4.18	4.17
	18.95	*18.45*	*18.36*	*17.80*	*17.23*	*16.74*	*16.90*	*1637*	*16.17*	*15.71*	*17.06*	*1655*
Caso 14					3.20	3.18	3.73	3.91	3.87	3.85	3.80	3.79
					17.47	*1697*	*17.83*	*17.25*	*17.03*	*1649*	*17.68*	*17.18*
Caso 15					5.02	5.00			4.60	4.42	5.44	5.20
					15.98	*15.48*			*15.30*	*14.86*	*15.18*	*14.71*
Caso 16					5.09	4.90			6.77	6.75	5.42	5.40
					16.29	*15.83*			*1606*	*15.61*	*15.77*	*15.29*

Tabla 6-7.- Porcentaje de pérdida y sobrestimación (en cursiva) de la respuesta genética esperada, al aplicar distintos índices bajo la hipótesis de homogeneidad genética entre partos cuando no es cierta.
Nota: $H1=(a^1 + 0.88\ a^2 + 0.72\ a^3 + 0.58\ a^4)$; $H2=(a^1+a^2+a^3+a^4)$;
Casos descritos en la Tabla 6-5.

Tal como hallan Tartar y Bolet (1984), en un índice individual la pérdida de respuesta es mayor cuando mayor es el número de partos que se consideran, y como cabía esperar ocurre lo mismo en índices que incorporan información familiar (Tabla 6-6).

La pérdida de respuesta esperada fue prácticamente la misma bajo los dos vectores de pesos económicos considerados (Tabla 6-7). La diferencia media entre ellos era de un 20%, por lo que según los resultados de Vandepitte y Hazel (1977) cabía esperar pocas diferencias entre ambos. Tampoco existen diferencias al considerar la misma variabilidad en los registros de todos los órdenes de parto o considerar una menor variabilidad en el primer parto, ni considerar las correlaciones ambientales entre partos de animales emparentados, iguales o distintas a cero (Tabla 6-6 y Tabla 6-7).

Pérdidas de un 7% son lo suficientemente importantes como para tenerlas en cuenta. Así, como ejemplo, se puede pensar que la aplicación del BLUP en caracteres reproductivos de especies prolíficas, puede suponer incrementos semejantes en progreso genético (Estany, 1987), si bien, existen trabajos que muestran incrementos superiores (Long *et al.*, 1991). Además, cabe pensar que las implicaciones de la heterogeneidad genética puedan ser mayores cuando se realizan predicciones BLUP, en las que se utiliza toda la información familiar disponible.

Analizando un esquema de selección como el de Avalos y Smith (1987), la pérdida de progreso genético esperado que conlleva considerar homogeneidad genética cuando no existe, no es muy elevada, entre un 2-3% (Tabla 6-8), lo que supone aproximadamente 0.01 lechones/año/cerda. Por contra, las expectativas de progreso se sobrestimarían considerablemente (sobre un 20% -Tabla 6-8-), lo que equivaldría a esperar un progreso de aproximadamente 0.08 lechones/cerda/año más de los que realmente se pueden alcanzar. La pequeña tasa de pérdida de respuesta, 2-3%, se justifica al tener en cuenta que en este esquema sólo se dispone de la información de dos partos, situación en la que la Tabla 6-6 también muestra que la pérdida de respuesta es poco relevante.

Casos (*)		**Respuesta genética esperada (echones/cerda/año)**		**Pérdida (%)**		**Sobrestimación (%)**	
		H_1	H_2	H_1	H_2	H_1	H_2
$\sigma^2p_1 \neq \sigma^2p_2$	$r_e=0$	0.386	0.396	2	3	17	21
$\sigma^2p_1 \neq \sigma^2p_2$	$r_e \neq 0$	0.368	0.380	2	3	19	21
$\sigma^2p_1 = \sigma^2p_2$	$r_e=0$	0.389	0.398	2	2	17	20
$\sigma^2p_1 = \sigma^2p_2$	$r_e \neq 0$	0.370	0.380	2	3	17	21

Tabla 6-8.- Respuesta genética esperada en el esquema descrito por Avalos y Smith (1987) cuando existe heterogeneidad genética entre partos (considerando los agregados H_1 y H_2), y pérdida y sobrestimación que supone considerar la hipótesis de identidad.

(*)$\sigma^2p_1 \neq \sigma^2p_2$: varianza fenotípica del primer parto=7.29, y del segundo=8.41
$\sigma^2 p_1 = \sigma^2p_2$: varianza fenotípica de primer y segundo parto=8.13
$r_e=0$: no existe correlación residual entre partos de distintas cerdas.

$r_e \neq 0$: correlación residual entre partos de una cerda con su madre = -0. 1, y con sus hermanas = 0.05.
Nota: $Hl=(a_1 + 0.88\ a_2 + 0.72\ a_3 + 0.58\ a_4)$; $F12=(a_1+a_2+a_3+a_4)$

Tartar y Bolet (1984) demostraron que asumir correlaciones genéticas mayores de las verdaderas suponía sobrestimar la respuesta genética esperada, llegando a ser del orden del 300% en el caso extremo de que las correlaciones genéticas fuesen nulas y se considerasen igual a la unidad. Cuando se consideran índices de selección con información familiar, la sobrestimación que se produce es ligeramente menor, tal como muestran los resultados de la Tabla 6-6. En todos los casos analizados (Tabla 6-7) su valor osciló entre un máximo del 23 %, correspondiente a un índice individual, y un mínimo del 15 %, correspondiente a un índice que considera información familiar; estos valores están en conformidad con los obtenidos por García *et al.*, (1989) tanto de forma teórica como por simulación.

La selección en base a un índice con información familiar es, en términos de pérdida de progreso genético, más sensible al incumplimiento del modelo de homogeneidad genética (más cuantos más partos se consideren) que la selección individual. Sin embargo, y pese a afectar de forma importante, afecta ligeramente menos a la valoración de las expectativas de progreso. Según estos resultados, utilizar erróneamente un modelo animal con repetibilidad considerando todas las relaciones de parentesco conocidas, puede llevar a pérdidas de respuesta no despreciables.

BIBLIOTECA VIRTUAL

6.3.2.- Simulación de un proceso de selección hiperprolífica en la población A.

Pese a que los resultados del apartado anterior pueden aconsejar utilizar un modelo multicarácter, la Tabla 6-9 muestra que considerar la hipótesis de heterogeneidad genética cuando no es cierta, puede conducir a pérdidas de respuesta mayores, que considerar un modelo de repetibilidad, aun siendo también incorrecto. Si bien para intensidades moderadas sería irrelevante considerar uno u otro modelo, pues las pérdidas son prácticamente iguales (14 o 15%), para intensidades elevadas (1%) la pérdida pasa de ser un 8%, cuando el modelo de repetibilidad es erróneo, a un 19% cuando lo es el multicarácter.

		Hipótesis genética			
		Heterogeneidad		**Homogeneidad**	
Modelo	**p**	**R**	**P (%)**	**R**	**P(%)**
Multicarácter	1	0.12	-	0.22	19
	10	0.09	-	0.15	15
Repetibifidad	1	0.11	8	0.27	-
	10	0.07	14	0.18	- -

Tabla 6-9.- Respuesta esperada (R, en lechones nacidos vivos por parto) y pérdida de respuesta esperada (P, en %) al seleccionar, para distintas presiones (p, en %), la población A mediante los modelos multicarácter o de repetibilidad, y considerar la hipótesis de heterogeneidad u homogeneidad genética entre partos alternativamente correctas.

Por tanto, en un proceso de selección, parece más prudente asumir el riesgo que implica aceptar la hipótesis de identidad que utilizar un modelo multicarácter, fundamentalmente cuando se apliquen de forma puntual intensidades de selección muy

elevadas.

No obstante, predecir la respuesta genética al seleccionar un 1 o un 10% de las mejores cerdas de la población A mediante un modelo de repetibilidad, puede conducir a sobrestimarla hasta más de un 150%, mientras que utilizar un modelo multicarácter implicaría infraestimarla menos de un 50%. Por lo tanto, también parece prudente considerar la hipótesis de heterogeneidad genética cuando se deba evaluar el interés de un proceso de selección antes de tomar la decisión de llevarlo a cabo.

La Tabla 6-10 nos ratifica la escasa importancia que tienen pequeños errores en el valor de los pesos económicos sobre la tasa de respuesta genética, tal como se encontró trabajando con índices de selección. Errores con un coeficiente de variación del 35% respecto al verdadero objetivo de selección no implican pérdida de respuesta esperada.

		Objetivo			
		H1		**H2**	
Criterio	**P**	**R**	**P(%)**	**R**	**P(%)**
H1	1	0.12	-	0.11	0
	10	0.09	-	0.08	0
H2	1	0.12	0	0.11	-
	10	0.09	0	0.08	-

Tabla 6-10.- Respuesta esperada (R, en lechones nacidos vivos por parto) y pérdida de respuesta esperada (P, en %) al seleccionar, para distintas presiones (p, en %), la población A mediante un modelo multicarácter y dos agregados económicos considerados alternativamente correctos.
Nota: $Hl=(a_1 + 0.71\ a_2 + 0.61\ a_3 + 0.48\ a_4 + 0.42\ a_5)$; $H2=(a_1+a_2+a_3+a_4+a_5)$

6.3.3.- Simulación de un proceso de selección hiperprolífica en una población simulada con la misma estructura que la población A.

En la Tabla 6-11, se recogen los resultados de la simulación de un proceso de selección hiperprolífica en una población simulada con la misma estructura que la población A y tres hipótesis genéticas alternativas (HO, HE1, HE2), al considerar tres modelos de evaluación (MR, MM1, MM2, correspondientes respectivamente a las hipótesis anteriores). Se observa que si bien existen fuertes diferencias en la esperanza de la respuesta, la variabilidad de la respuesta es lo suficientemente baja como para que *a priori* se pueda pensar que no van a condicionar demasiado nuestras decisiones. Los resultados mostrados en esa tabla conducen a las mismas conclusiones a las que se llegaba en el apartado anterior, en el que se comparaban los modelos de evaluación en base a los datos reales de la población, teniendo en cuenta que ahora la respuesta corresponde únicamente a las hembras seleccionadas. En este caso, la pérdida que implica utilizar el modelo de repetibilidad incorrectamente es del 7%, y de hasta un 9%, en el caso de utilizar también inapropiadamente el modelo multicarácter, frente al de repetibilidad. La diferencia entre modelos no llega a ser tan importante a la observada en el apartado anterior al aplicar una presión de selección del 1%, hecho que se debe relacionar a que en este caso, la presión aplicada fue de aproximadamente un 4%.

Hipótesis genética

Modelo	HO	HE1	HE2
MR	0.664 (0.008)	0.176(0.001)	0.564(0.006)
MM1	0.606 (0.008)	0.189(0.002)	0.528(0.007)
MM2	0.649 (0.007)	0. 167 (0.001)	0.569 (0.007)

Tabla 6-11.- Respuesta media (y varianza) de las nueve combinaciones entre modelos (MR: repetibilidad, MMI: multicarácter con h^2 ponderada 0.02; MM2: h^2 ponderada 0.06) e hipótesis (HO=MR, HEl=MM1, HE2=MM2) considerados en la simulación del esquema hiperprolífico de la población A.

La respuesta genética esperada al simular y evaluar en base a la hipótesis de identidad genética, seleccionando sólo las hembras de la población fundadora del esquema hiperprolífico, es de aproximadamente 0.33 lechones por parto. Este resultado coincide con simulaciones previas del experimento realizadas en base a la misma hipótesis genética (Noguera *et al.*, 1994), en las que la respuesta esperada, al seleccionar tanto hembras como machos fundadores de la línea hiperprolífica, fue de 0.41±0.04 lechones/parto. La escasa diferencia observada al seleccionar o no los machos se debe a que la diferencia entre los valores genéticos medios predichos para los machos de la línea seleccionada y la control fue de sólo 0.22 lechones/parto.

Si nos situamos en un ambiente de incertidumbre, considerando que no tenemos ninguna información sobre el modelo genético, podemos basar la decisión de qué modelo de evaluación emplear, en las matrices de respuesta esperada, absoluta y relativa, y de pérdida relativa que se recogen en la Tabla 6-12.

Modelo		Hipótesis genética HO	HE1	HE2
MR	**RA**	0.664	0.176	0.564
	RR	0.058	0.009	0.036
	PR	0	0.013	0.005
MM1	**RA**	0.606	0.189	0.528
	RR	0	0.022	0
	PR	0.058	0	0.041
MM2	**RA**	0.649	0.167	0.569
	RR	0.043	0	0.039
	PR	0.015	0.022	0

Tabla 6-12.- Matrices de ganancia absoluta (RA), de ganancia relativa (RR)(*) y de pérdida relativa de respuesta genética esperada (PR)(*), en función de las distintas hipótesis genéticas y modelos de evaluación considerados.
(*) Siendo i la hipótesis genética y j el modelo de evaluación, se tiene que:
$RR_{ij}= RA_{jj} - \text{mín}\{RA_{ij}\}j=_{MR,MMI,MM2}$
$PR_{ij}= \text{máx}\ (RA_{ij}\}=_{MR,\ MM1,\ MM2} - RA_{ij}$
Abreviaturas: Modelos: MR: de repetibilidad, MM1: multicarácter con h^2 ponderada 0.02; MM2: multicarácter con h^2 ponderada 0.06 ; Hipótesis: HO corresponiente al modelo MR; HE1 correspondiente al modelo MM1; HE2 correspondiente al modelo MM2.

Todos los criterios, excepto el de Wald, indican que deberíamos decidirnos por el modelo de repetibilidad (Tabla 6-13). El criterio de Wald se basa en que una vez se ha decidido, la naturaleza siempre se opondrá a nuestra decisión. Por eso muestra preferencia por el modelo multicarácter con las correlaciones genéticas más separadas de la unidad y una menor h^2 ponderada, es decir por la situación más desfavorable. Es un criterio absolutamente pesimista que conduce a adoptar siempre la postura más conservadora.

	DECISIÓN
•*Criterio de Wald:* máx [(mín) RA] = máx (0. 176,0.189,0.167) = 0. 189	**MM1**
•*Criterio maximax:* máx [(máx) RA] = máx (0.664,0.606,0.649) 0.664	**MR**
•*Criterio de Agrawal-Heady:* máx [(mín) RR] = máx (0.009,0,0) 0.009	**MR**
•*Criterio de Savage:* mín [(máx) PR] = mín (0.013,0.058,0.022) 0.013	**MR**
•Criterio de *Laplace* MR: (0.664+0.176+0.564)/3 = 0.468 MM1: (0.606+0.189+0.528)/3 = 0.441 MM2: (0.649+0.167+0.569)/3 = 0.462 máx [0.468,0.441,0.462] = 0.468	**MR**

Tabla 6-13.- Decisiones sobre el modelo de evaluación en un contexto de incertidumbre.
Abreviaturas: Modelos: MR: de repetibilidad, MM1: multicarácter con h^2 ponderada 0.02; MM2: multicarácter con h^2 ponderada 0.06.

Los criterios de Savage y Agrawal-Heady también son absolutamente pesimistas, pues parten del mismo principio de oposición de la naturaleza, pero como vemos conducen a decisiones distintas que el de Wald. Las bases psicológicas de los tres criterios no coinciden: el de Wald se lamenta por lo que se pierde, el de Savage por lo que se deja de ganar y el de Agrawal-Heady se alegra por lo que se deja de perder. Ballesteros (1983) califica las tres posturas como inconformista, timorata y conformista respectivamente.

El criterio maximax no es un criterio habitual de decisión. Se basa en una postura poco racional, pues asume que la naturaleza siempre favorecerá la decisión que tomemos, de forma que encierra una predisposición total por el riesgo. En nuestro caso coincide con la mayoría de criterios, lo cual es un indicativo de que en realidad existe poco riesgo en la decisión del modelo de evaluación a emplear. Eso también se puede ver en la Tabla 6-14: el criterio de Hurwicz indica que debemos escoger el modelo de repetibilidad, excepto si nos aproximamos mucho al criterio de Wald [que no es más que I = 0 Máx (R) + 1 Mín (R)], dado que en este caso nos estamos mostrando casi absolutamente pesimistas.

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MR	0.225	0.274	0.322	0.371	0.420	0.469	0.518	0.566	0.615
MM1	0.231	0.272	0.314	0.356	0.398	0.439	0.481	0.523	0.564
MM2	0.215	0.263	0.312	0.360	0.408	0.456	0.504	0.553	0.601
Decisión	MM1	MR	MR	MR	MR	MR	MR	MR	MR

Tabla 6-14.- Valores del índice utilizado por el criterio de Hurwicz (I=a.Máx R+(I-(x). Mín R) y decisiones a las que conducen, en función de¡ grado de optimismo que se exprese (Cc).
Abreviaturas: Modelos: MR: de repetibilidad, MMI: multicarácter con h^2 ponderada 0.02; MM2: multicarácter con h^2 ponderada 0.06.

La ventaja del criterio de Hurwicz frente al resto de criterios es que permite considerar situaciones intermedias entre el optimismo y el pesismismo absoluto. Coincide con el de Laplace cuando sólo existen dos posibles estados de la realidad y se asume el mismo grado de optimismo que de pesimismo. Cuando existen más estados de la realidad posibles, presenta la desventaja de despreciar la información de las consecuencias intermedias.

Tal como se comentó en el apartado 6.2.2.3.2., en un ambiente de riesgo, los resultados del proceso de simulación para cada una de las alternativas de decisión que tenemos, se pueden considerar como el resultado de finitos o infinitos estados de la realidad.

Los coeficientes de riesgo utilizados al asumir que las tres hipótesis consideradas son las únicas que pueden explicar el carácter, fueron 0.295 para la función de utilidad cuadrática[40] y 65,2 para la exponencial[41]. Esos coeficientes se derivaron de los resultados de la Tabla 6-11, tal como se expuso en el apartado anteriormente citado, y permitieron calcular la utilidad esperada para cada decisión aplicando la ley de probabilidades asumida [prob(HE1) = 0.1; prob(HO) = prob(HE2) = 0.45].

Los resultados que se recogen en la Tabla 6-15 indican que, independientemente a la función de utilidad empleada, la mejor decisión es la de utilizar el modelo de repetibilidad como modelo de evaluación. Eso es así por ser el que maximiza la utilidad esperada, aunque cabe destacar que existen muy pocas diferencias con el modelo multicarácter MM2.

Utilidad esperada		
Modelo	**E[UC]**	**E[UE]**
MR	0.468	0.361
MM1	0.442	0.303
MM2	0.465	0.356

Tabla 6-15.- Uilidad esperada para cada modelo de evaluación MR, MM1, MM2) al asumir la existencia de tres hipótesis genéticas y una ley de probabilidades conocida, para las funciones cuadrática (UC) y exponencial (UE). Abreviaturas: Modelos: MR: de repetibilidad, MM1: multicarácter con h^2 ponderada 0.02; MM2: multicarácter con h ponderada 0.06.

Si se considera que las tres hipótesis genéticas son una muestra extraida al azar de un infinito número de hipótesis, para cada decisión tenemos un conjunto de 30 procesos de simulación con una media y una variabilidad de respuesta que se muestran en la Tabla 6-16. En este caso los coeficientes de riesgo de las funciones de utilidad a aplicar, fueron 0.27 para

[40] $1/(2[2.0.664 + 4.(0.008)^{1/2}])=0.295$

[41] Se consideró que la utilidad de la respuesta máxima era 3 veces la de la respuesta mínima pues el cociente entre ambas respuestas era 4, pero la respuesta máxima tenía mayor varianza (2.{[0.664*1-0.167*3]/[0.008*1-0.001*3]})=65.2

la función de utilidad cuadrática[42] y 5.08 para la función de utilidad exponencia[43].

Modelo	R (σ^2)	Prob
MR	0.486 (0.048)	0.997
MM1	0.457 (0.038)	0.987
MM2	0.479 (0.049)	0.998

Tabla 6-16.- Matriz de respuestas esperadas (R) y probabilidad de distribución de los datos según una normal (contraste de Shapiro-Wilk), para los tres modelos de evaluación considerados (MR, MMI, MM2). Abreviaturas: σ^2: varianza de la respuesta esperada Modelos: MR: de repetibilidad, MM1: multicarácter con h^2 ponderada 0.02; MM2: multicarácter con h^2 ponderada 0.06.

Las utilidades esperadas para cada decisión, mostradas en la Tabla 6-17, conducen a la misma decisión que al tratar de forma discreta la realidad: el modelo de repetibilidad es el más adecuado. La probabilidad de ocurrencia de HE1 es baja al asumir una distribución normal, por ser el caso más alejado de la respuesta esperada media. No obstante, es lo suficientemente grande como para que el modelo MM1 se sitúe en el segundo lugar del orden de preferencias con la función de utilidad exponencial, aunque no con la función cuadrática. Esto es posiblemente debido a que el estrecho dominio de los valores de los coeficientes de aversión de riesgo en la función cuadrática hace que esta sea mucho menos sensible a variaciones de utilidad (Wyler *et al.,*1994).

	Utilidad esperada	
Modelo	**E[UC]**	**E[UE]**
MR	0.409	0.364
MM1	0.390	0.360
MM2	0.404	0.355

Tabla 6-17.- Utilidad esperada para cada modelo de evaluación MR, MM1, MM2) al asumir la existencia de infinitas hipótesis genéticas, para las funciones cuadrática (UC) y exponencial (UE).
Abreviaturas: Modelos: MR: de repetibilidad, MM1: multicarácter con h^2 ponderada 0.02; MM2: multicarácter con h^2 ponderada 0.06.

En términos de respuesta nuestra decisión parece clara: la utilización de un modelo de evaluación multicarácter no está justificada en la fundación de un núcleo de selección basada en la aplicación de una elevada intensidad de selección, por lo que optaríamos por un modelo de repetibilidad. Sólo un individuo ultraconservador y absolutamente pesimista, actitud poco

[42] $1/(2[2.0.486 + 4.(0.048)^{1/2}])=0.27$

[43] Se consideró que la utilidad de la respuesta máxima era 1.01 veces la de la respuesta mínima pues el cociente entre ambas respuestas era 1.06, pero la respuesta máxima también tenía mayor varianza (2{[0.486*1-0.457*1.01] / [0.048*1-0.038*1.01]}))= 5.08

racional, optaría por un modelo multicarácter. Esta conclusión corrobora los resultados obtenidos en el apartado anterior, por simple comparación de pérdidas de respuestas esperadas al utilizar los partos registrados de la población A.

La aplicación de elementos de la teoría de la decisión no permite afirmar cual es el modelo genético verdadero, pero permite aclarar cual es el modelo más útil: es posible que las correlaciones genéticas entre partos no sean la unidad, pero desde un punto de vista práctico, para seleccionar una población animal no parece tener mucha importancia. Sí puede tener importancia en la evaluación, tanto *a priori* como *a posteriori* de un proceso de selección.

Esos elementos también permiten conocer mejor la mentalidad que existe detrás de una decisión de este tipo. Los criterios de ambiente de incertidumbre nos permiten ver que algunas decisiones lógicas que se toman en mejora animal no son absolutamente pesimistas. Con la información que existía previamente al análisis de decisión habría pocas personas que adoptasen el modelo multicarácter más desfavorable (MM1), algunas personas estarían dispuestas a adoptar el otro modelo multicarácter (MM2), y la mayor parte adoptaría el modelo de repetibilidad.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

7.- Discusión general

Considerar la prolificidad como un objetivo de selección en líneas porcinas de especialización materna implica conocer los distintos componentes de varianza y covarianza que la determinan. Su desconocimiento impide establecer un criterio óptimo de selección de los animales como padres de futuras generaciones, y puede condicionar la eficiencia de un programa de mejora. El conocimiento que se puede alcanzar sobre los componentes de (co)varianza de la prolificidad de una determinada población depende, fundamentalmente, de la información disponible, del modelo genético y estadístico utilizado, y del método de estimación aplicado.

Se han realizado numerosos esfuerzos por disponer de métodos de estimación que presenten propiedades óptimas frente a las situaciones que habitualmente se dan en las poblaciones animales, básicamente en las que son objeto de programas de mejora animal. Lo mismo podríamos decir sobre los esfuerzos realizados durante los últimos diez años entorno a los algoritmos y los desarrollos informáticos que permiten aplicar esos métodos de estimación. La traducción de esos esfuerzos es disponer de métodos de estimación como el REML (Patterson y Thompson, 1971) y distintos programas informáticos (Misztal, 1994) fácilmente aplicables al análisis de la variabilidad de cualquier población animal.

No obstante, esa facilidad no rompe la dependencia que existe entre las estimaciones que podemos obtener, las características de la información disponible y el modelo genético y estadístico que consideramos. Por el contrario, dado que los estimadores REML sólo asintóticamente presentan propiedades óptimas, crece la necesidad de disponer de gran cantidad de información aumentando, frecuentemente, la dificultad de conocer sus características y de asegurar su calidad.

Además, la complejidad de los modelos ideales que se pueden plantear, hace que éstos no siempre sean operativos. En este sentido, en el Capítulo 4 se discute que la estimación de efectos maternos presenta problemas asociados a la información disponible en las poblaciones animales, que se traducen en una insuficiente eficiencia de los algoritmos basados en la función de verosimilitud, y en el Capítulo 5 se discuten los problemas que se presentan en la obtención de estimadores REML al utilizar modelos multicarácter que contienen un gran número de parámetros a estimar.

El interés del presente trabajo se ha centrado, básicamente, en analizar las características de la información disponible, concretamente las que hacen referencia a la población base, y las asunciones habitualmente realizadas en el modelo genético y estadístico utilizado en el análisis de la prolificidad.

Importancia de la definición de la población base de referencia.

De la misma forma que se han realizado esfuerzos para disponer de mejores métodos de estimación de componentes de (co)varianza, también se han dedicado recursos para disponer de más y mejores herramientas informáticas de gestión en granja, así como de Sistemas de Información públicos (p.ej. Noguera *et al.*, 1992) que permitiesen conocer mejor las poblaciones animales y asegurar la calidad de la información captada en ellas.

Es necesario conocer si existe selección de la información disponible, dado que esa selección puede afectar a nuestras inferencias sobre la heredabilidad. Frecuentemente, para apoyar la validez de los valores de heredabilidad estimados en poblaciones animales se asume que no se ha producido selección, o en caso contrario, que los estimadores REML tienen en cuenta esa selección por conocer toda la información que se ha utilizado para realizarla. Difícilmente se dan esas circustancias (siempre se desconocen algunas observaciones, relaciones de parentesco, etc...), pero aun dándose, no son suficientes si la

población base no está formada por una muestra de animales tomados al azar. Determinados trabajos como los de Van der Werf y Boer (1990), Van der Werf (1992), y Van der Werf y Thompson (1992), permiten ver la importancia que tiene la selección de los animales de la población base que se toma de referencia en la estimación de parámetros genéticos; aun conociendo toda la información utilizada en un proceso de selección, los estimadores REML están sesgados si la población base está seleccionada, y tratar los animales de esa población como efectos fijos (Graser *et al.*, 1987) no permite evitarlos.

En el Capítulo 3 se vio que considerar indiscriminadamente toda la información disponible de la población A conducía a infraestimar el valor de la heredabilidad de la prolificidad. Esto era debido a disponer de una muestra seleccionada de la información producida en los primeros años registrados, aquellos en los que no se disponía de herramientas eficientes de recogida de información. Por esta misma razón, se puede pensar que en otras poblaciones animales, la información más antigua también sea la más sensible a la existencia de selección; en definitiva, un problema importante que puede presentar la información captada en las poblaciones animales es el asociado a la definición de la población base.

En el mismo Capítulo 3, y trabajando con poblaciones simuladas, se comprueba que es posible obtener valores sobrestimados de la heredabilidad si existe selección por truncamiento de las observaciones de los animales de la población base. Para evitarlo se puede tener en cuenta el efecto que esa selección tiene sobre la distribución de las observaciones, es decir sobre la distribución de los valores genéticos y residuales. Por una parte, se debería tener en cuenta el efecto sobre la media de ambas variables; sobre la de los residuos, incluyendo un efecto fijo para los distintos grupos de animales afectados, y sobre la de los valores genéticos, a través de otro efecto fijo, en este caso asignando grupos genéticos (Quaas, 1988). Por otra parte, se debería considerar el efecto sobre la varianza tanto de los residuos como de los valores genéticos; es decir, considerar la existencia de varianzas residuales heterogéneas entre grupos de observaciones afectadas por distintos niveles de un efecto fijo (p.ej. Henderson, 1984), y la existencia de varianzas genéticas heterogéneas entre grupos de animales (p.ej. Visscher y Thompson, 1990; Lo *et al.*, 1993). La heterogeneidad de varianzas genéticas entre grupos de animales se puede modelizar fácilmente, como se deduce del Anejo, cuando esos grupos se asocian a la definición de la población base.

Estimar la heterogeneidad de varianzas entre animales base y no base puede ser difícil mediante métodos basados en la maximización de funciones de verosimilitud. En poblaciones simuladas con los animales base seleccionados, se intentó estimar, mediante un algoritmo EM, el cociente entre la varianza de la población base y el doble de la varianza de recombinación, pero los resultados iniciales indicaban que no se podía diferenciar entre ambos valores. Pese a la cautela con que se deben interpretar esos resultados (razón fundamental por la que no se han incluido en este trabajo), es posible que la información simulada fuese insuficiente para distinguir ambos valores, aunque es difícil pensar en situaciones reales en las que se disponga de un elevado número de hermanos completos, tal que permita diferenciar mejor entre ambas varianzas.

Otra forma de evitar el efecto que causa analizar información seleccionada es, siempre y cuando se pueda detectar, despreciarla. Proceder de este modo no es recomendable cuando se dispone de poca información, o cuando la información pertenece a una población sometida a selección (Van der Werf y Boer, 1990), pero puede ser apropiado en caso contrario, como era el caso de la población A, facilitando realizar otro tipo de inferencias sobre el carácter en esa población, como por ejemplo podría ser la predicción de valores genéticos, la estimación de tendencias genéticas y ambientales, etc...,

Pese a todo, las heredabilidades estimadas en ambas poblaciones (0.05-0.07) siguen siendo más bajas de lo esperado (Haley *et al.*, 1988; Lamberson, 1990), de modo que las expectativas de respuesta a la selección son menores de las que en principio se podían tener. Esos valores contrastan con los resultados provisionales de un experimento de selección hiperprolífica que se está realizando en la población A (Noguera *et al.*, 1994). La respuesta esperada al simular el experimento con esos valores de heredabilidad está entorno a los 0.3 - 0.4 lechones/parto, mientras que los resultados provisionales indican una respuesta de cerca de 1 lechón/parto (J.L.Noguera, comunicación personal). Cabe pensar que los valores estimados de heredabilidad están infraestimados. Esa infraestimación es justificable por no conocer, en ambas poblaciones, gran parte de las relaciones de parentesco entre sus animales.

La mejor forma de garantizar que las inferencias que podemos hacer a partir de un conjunto de datos son acertadas, es disponer de un sistema fiable de recogida de la información de las poblaciones animales. Si no es así, difícilmente se pueden cumplir todas las hipótesis de partida necesarias para obtener estimaciones insesgadas, de los parámetros genéticos asociados a un determinado carácter.

En poblaciones relativamente grandes como las consideradas en este trabajo puede ser intrascendente considerar valores ligeramente infraestimados de heredabilidad, pero en poblaciones pequeñas y cerradas, no se puede rechazar el seleccionar considerando un valor de heredabilidad superior al estimado. Seleccionar mediante predicciones BLUP frente a realizar selección masal permite aumentar considerablemente la respuesta genética, sobre todo para caracteres poco heredables, pero ese aumento es proporcionalmente menor que el aumento de la tasa de consanguinidad de la población (Wray, 1989). Integrar la información existente *a priori* (resultados experimentales, desconocimiento de relaciones de parentesco, etc...) y trabajar con predictores de tipo bayesiano empírico (Gianola *et al.*, 1986), es posiblemente la opción más adecuada cuando no se puede asegurar el uso de predictores insesgados (Gianola *et al.*, 1988).

Importancia de las asunciones establecidas en los modelos genéticos y estadísticos.

Al inicio de esta discusión comentábamos que el conocimiento que se puede alcanzar sobre los componentes de (co)varianza de la prolificidad también depende del modelo genético y estadístico que utilizamos. La existencia de efectos maternos y de heterogeneidad genética entre partos son dos razones que se han apuntado como posible causa de infraestimación de la heredabilidad (Le Roy *et al.*, 1987); no obstante, habitualmente se asume que no se da ninguna de ambas circunstancias.

En las poblaciones analizadas parece apropiado asumir que no existe influencia materna sobre la prolificidad. Esos resultados se obtuvieron ignorando la posible existencia de efectos de ambiente común entre animales de una misma camada de lactación, por no disponer de la información de esas camadas. Dado que en ambas poblaciones había cierto grado de estandarización del tamaño de las camadas durante la lactación, es difícil pensar en la existencia de un importante efecto ambiental de camada.

Tal como se comprobó en poblaciones simuladas, ignorar efectos significativos de ambiente común entre animales de una misma camada de lactación, cuando en realidad existen, conduce a sobrestimar el valor de los efectos maternos de tipo genético. Esto permitiría explicar por qué en algunas poblaciones, como las analizadas por Southwood y Kennedy (1990), en las que se ignoran los efectos de ambiente común de camada de lactación, los efectos maternos sí parecen tener influencia. También nos permite afirmar que, en nuestras poblaciones, los efectos maternos de origen genético pueden ser incluso menores que los estimados, por lo que se puede cuestionar que en realidad existan. No obstante, el

modelo genético que se emplea para explicar la influencia materna es posiblemente demasiado complejo para la información que, en general, se puede llegar a conocer, de modo que pueden existir efectos maternos que seamos incapaces de estimar.

Por otra parte, en las actuales poblaciones porcinas comerciales cabe pensar que la principal influencia materna que existe sobre la prolificidad es de tipo prenatal, básicamente uterino, dado que la estandarización de camadas es una práctica habitual de manejo. Ese efecto prenatal, de existir, parece ser lo suficientemente pequeño como para no mostrarse significativo y no implicar importantes pérdidas de respuesta a la selección. Para bajas heredabilidades maternas (0.02) e incluso correlaciones genéticas negativas elevadas, Gómez (1994) encuentra pérdidas de menos del 3% al simular un proceso de selección BLUP de 10 generaciones. No obstante, tal como señalan Roehe y Kennedy (1993), aun cuando la estandarización del tamaño de camada durante la lactación implicase la eliminación de la influencia materna, las predicciones de los valores genéticos directos podrían estar sesgadas debido a los cambios producidos en los valores genéticos maternos.

Respecto a la hipótesis genética entre distintos partos de una misma cerda, la que parece más apropiada para explicar la prolificidad es la hipótesis de homogeneidad. En las poblaciones estudiadas, no existen diferencias substanciales entre las estimas de heredabilidad obtenidas, para cada carácter y población, bajo las hipótesis de homogeneidad o heterogeneidad genética. Bajo esta última hipótesis tampoco existen diferencias entre las heredabilidades de distintos órdenes de parto. Además, las correlaciones genéticas entre la prolificidad de distintos partos, aunque son menores de la unidad, no se puede afirmar que lo sean significativamente. Sólo la correlación genética entre la prolificidad de primer y tercer parto de una de las dos poblaciones, podría considerarse distinta de la unidad (con un valor próximo a 0.5).

Los resultados alcanzados son análogos a los de otros autores, tanto en porcino (Knap *et al.*, 1993; Rydhmer, 1993; Irgang *et al.*, 1994), como en conejos (Gómez, 1994), pero sigue siendo difícil concluir sobre el modelo genético que realmente rige la prolificidad. Intelectualmente es más fácil aceptar el modelo de heterogeneidad pero no existen resultados sólidos que lo apoyen.

De todos modos, desde un punto de vista práctico de la mejora animal es posiblemente más importante conocer las implicaciones del modelo estadístico que aplicamos, que determinar el verdadero modelo genético de la prolificidad. Si las correlaciones genéticas no son la unidad, utilizar un modelo de repetibilidad puede implicar, teóricamente, una considerable pérdida de respuesta genética esperada. Esta pérdida es mayor si se selecciona utilizando información familiar, que si se realiza selección individual. No obstante, en la práctica, si no se tiene ningún conocimiento sobre el verdadero modelo genético, o se considera que lo más probable es que las correlaciones genéticas no difieran mucho de la unidad, la decisión sobre qué modelo de evaluación genética emplear es la de utilizar el modelo con repetibilidad. Sólo desde una actitud absolutamente pesimista y asumiendo que las correlaciones genéticas pueden alejarse considerablemente de la unidad, tendría sentido utilizar un modelo multicarácter.

Finalmente, los resultados de simulación del Capítulo 6 indican que aun considerando bajas heredabilidades, los esquemas de selección hiperprolífica pueden ser una interesante vía para mejorar la prolificidad de las cerdas; la respuesta genética esperada en la población A, considerando los parámetros genéticos estimados bajo la hipótesis de identidad genética entre partos, es de más de 0.3 lechones por parto en las cerdas procedentes de los apareamientos de hembras seleccionadas y machos tomados al azar.

8.- Conclusiones

A partir de los resultados alcanzados en los capítulos anteriores vamos a exponer las conclusiones que, a nuestro juicio, se derivan sobre los objetivos iniciales.

1 - Los valores de heredabilidad estimados para la prolificidad de las dos poblaciones estudiadas, bajo las hipótesis de inexistencia de efectos maternos y homogeneidad genética entre distintos partos de una misma cerda, oscilan entre 0.05 y 0.07, y los de repetibilidad entre 0.13 y 0.15.

2 - En la estimación de los parámetros genéticos de un carácter se deben analizar previamente las características de la población base de referencia; la existencia de selección en la población base conduce a estimaciones sesgadas si no se evita o se considera en el modelo de estimación. La información de los animales más antiguos de la población A es una muestra seleccionada y produce una infraestimación del valor de la heredabilidad de la prolificidad de aproximadamente un 50%; conviene despreciar esa información para seleccionar la población por ese carácter. En la población B no se detecta ningún tipo de selección de los animales base.

3 - En ninguna de las dos poblaciones se han estimado efectos genéticos maternos significativos sobre la prolificidad, probablemente por haberse producido una suficiente estandarización del tamaño de camada durante la lactación. Ignorar efectos ambientales comunes entre hembras criadas en una misma camada de lactación cuando en realidad existen, puede sobrestimar el valor de los efectos genéticos maternos.

4 - No se puede afirmar que existan diferencias entre los valores de heredabilidad estimados para la prolificidad de los cinco primeros partos, y entre éstos y el valor estimado bajo un modelo con repetibilidad. Los valores estimados de correlación genética entre la prolificidad de los cinco primeros partos corroboran la hipótesis de identidad genética; sólo el valor de la correlación genética entre primer y tercer parto de la población B puede considerarse distinto a la unidad.

5 - Seleccionar por prolificidad mediante un modelo con repetibilidad, cuando existe heterogeneidad genética entre la prolificidad de distintos partos, implica pérdidas no despreciables de respuesta genética esperada, mayores a medida que se considera más información familiar. No obstante, para realizar procesos de selección como el llevado a cabo en una de las poblaciones analizadas (población A), es preferible utilizar un modelo con repetibilidad; las pérdidas que se pueden producir se ven compensadas por el mayor riesgo que supone seleccionar mediante un modelo multicarácter cuando no se conocen los verdaderos valores de correlación genética entre partos.

6 - El modelo animal con repetibilidad que se ha venido utilizando durante los tres últimos años para seleccionar por prolificidad las dos poblaciones porcinas analizadas, puede considerarse como un modelo operativo más adecuado que un modelo multicarácter o un modelo con efectos maternos. Es poco probable que la respuesta genética esperada en el experimento de selección hiperprolífica que se está llevando a cabo en la población A, se vea afectada por haber desconsiderado ambas hipótesis genéticas en la fundación de la línea seleccionada. Las simulaciones realizadas indican una respuesta, sólo por selección de las hembras fundadoras, de más de 0.3 lechones por parto.

9.- Bibliografía

Alfonso, L., Babot, D., Estany, J., Ruiz, J., Noguera, J.L. (1993) Parámetros genéticos de la prolificidad en varias poblaciones porcinas. *V Jor. Prod. Animal. AIDA.* Zaragoza, *ITEA.* Vol.extra **12(I) :** 266-268.

Alfonso, L., Noguera, J.L., Pomar, J., Babot, D., Estany, J. (1991) Un modelo de evaluación genética de la prolificidad en porcino. *IV Jor. Prod. Animal. AIDA.* Zaragoza, *ITEA.* Vol.extra **11(I):** 547-549.

Alsing, I., Krippl, J., Pirchner, F. (1980). Maternal effects on the heritability of litter traits of pigs. *Z. Tierzüchtg. Züchtgsbiol.* **97** : 241-249.

Avalos, E. (1985) *Estimation of genetic parameters and reponses in selection for litter size in pigs.* Ph.D. Thesis, Edinburgh University.

Avalos, E., Smith, C. (1987) Genetic improvement of litter size in pigs. *Anim. Prod.* , **44**: 153-164.

Ballesteros, E. (1983). *Principios de economía de la empresa.* Alianza Editorial, Madrid.

Bampton, P.R. (1992) Best linear unbiased prediction for pigs - the commercial experience. *Pig News and Information.* **Vol.13 No.3** : 125N-129N.

Baselga, M., Carabaño, M.J. (1993). Estimación de componentes de varianza. *V Reunión Nacional Mejora Genética Animal.* Córdoba, 1990. (En: Colec.Congresos y Jornadas nº 27/92. Junta Andalucía).

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis* . Springer-Verlag, New York.

Bichard, M., David, P.J. (1985) Effectiveness of genetic selection for prolificacy in pigs. *J. Reprod. Fert. Suppl.*, **33** : 127-138.

Bichard, M., Seidel, C.M. (1982) Selection for reproductive performance in maternal lines of pigs. *Proc. of 2nd. World Congress on Genetics Applied to Livestock Production*, Madrid, **VIII** , pp. 565-569.

Bidanel, J.P., Ducos, A. (1994). Utilisation du BLUP Modèle Animal pour l'évaluation génétique des porcs de race Large White et Landrace Français sur la prolificité. *Journées Recherche Porcine en France*, **26**: 321-326.

Bidanel, J.P., Gruand, J., Legault, C. (1994). An overview of twenty years of selection for litter size in pigs using "hyperprolific" schemes. *Proc. 5th World Congress on Genetics Applied Livestock Production, Guelph.* **XVII**:512-515.

Boldman, K.G., Van Vleck, L. D. (1991) Derivative-Free Restricted Maximum Likelihood Estimation in Animal Models with a Sparse Matrix Solver. *J. Dairy Sci.* **74**: 4337-4343.

Bolet, G., Legault, C. (1982) New aspects of genetic improvement of prolificacy in pigs. *Proc. of 2nd World Congress on Genetics Applied to Livestock Production*, Madrid, **V**: 548-567.

Bolet, G., Ollivier, L., Dando, P. (1989) Sélection sur la prolificité chez le porc. I. Résultats d'une expérience de sélection sur onze générations. *Gen. Sel. Evol*. **21**:93-106.

Bolet, G., Tartar, M., Laloe, D., Felgines, C. (1984) Efficacité théorique de la sélection sur la prolificité chez le porc. Possibilités de modifier l'indice actuel. *Journées de la Recherche Porcine en France*, **16**: 475-480.

Brandt, H. (1992) Estimation of Breeding values in pigs. New developments in mixed model methodology. *43rd Annual Meeting of EAAP*. Madrid.

Brascamp, E.W., De Vries, A.G. (1992) Defining the breeding goals for pig improvement. *Pig News and Information.* **Vol. 13 No. 1**: 21N-26N

Bulmer, M.G. (1971) The effect of selection on genetic variability. *Am.Nat.* **105**: 201-209. (In: Quantitative Genetics, Part 2. Ed.W.G.Hill. Van Nostrand Reinhold Company Inc. New York).

Cantet, R.J.C. (1990) *Estimation and prediction problems in mixed linear models for maternal genetic effects.* PhD. Thesis. University of Illinois.

Cantet, R.J.C., Fernando, R.L., Gianola, D., Misztal, I. (1992) Genetic grouping for direct and maternal effects with differential assignement of groups. *Gen. Sel. Evol.* **24**: 211-223.

Chu, E., George, A., Liu J., Ng, **E.** (1984). *SPARSPAK: Waterloo Sparse Matrix Package . User's Guide for SPARSPAK-A.* Dep. of Computer Science. Univ. of Waterloo. Waterloo. Ontario. Canada.

De Vries, A.G., Sorensen, D.A. (1990). Optimization of present pig breeding programmes. *Proc. of 4th. World Congress on Genetics Applied to Livestock Production*, Edinburgh, **XV**: 395-404.

Dempster, A.P., Laird, N.M., Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J.Royal Stat.Soc.* Series B. **39**:1-38.

Ducos, A. (1994) *Evaluation génétique des porcs controlés dans les stations publiques a l'aide d'un modèle animal multicaractère.* Thèse. Institut National Agronomique de Paris-Grignon - Institut National de la Recherche Agronomique. Jouy-en-Josas.

Ducos, A., Bidanel, J.P., Ducrocq, V. (1992) Bilan sur l'utilisation du BLUP pour l'evaluation génétique des reproducteurs porcins dans le monde. *Journées Rech. Porcine en France*, **24**: 39-46.

Ducrocq, V. (1994). Multiple trait prediction: principles and problems. *Proc. of 5th World Congress on Genetics Applied Livestock Production,* Guelph. Vol 18: 445-462.

Estany, J. (1987*). Metodología de la selección en líneas de reproducción y crecimiento de conejo de carne*. Tesis Doctoral. Universidad Politécnica, Valencia.

Estany, J., Alfonso, L., Babot, D., Noguera, J.L. (1993). Genetic evaluation of production and reproduction traits from field pig data in Spain. *Application of mixed linear models in the prediction of genetic merit in pigs.* E. Groeneveld (Ed.) Commission of the European Communities. 1995. pp. 26-31.

Estany, J., Pomar, J., Alfonso, L., Babot, D., Noguera, J.L., (1992) A recording data system for efficient genetic evaluation in pig breeding. *Proc. Satellite symposium on pig management information systems*, Madrid. Monografías INIA, **68** : 221-228.

Estany, J., Sorensen, D. (1995) Estimation of genetic parameters for litter size in Danish Landrace and Yorkshire pigs. *Anim. Science.* **60(2)**:315-324.

Falconer, D.S. (1965). Maternal effects and selection response. *Genetics Today. Proc. XIth Int. Cong. on Genetics,* **3**:763-774.

Ferraz, J.B.S., Johnson, R.K. (1993). Animal model estimation of genetic parameters and response to selection for litter size and weight, growth and backfat in closed seedstock populations of Large-White and Landrace swine. *J. Anim. Sci.* **71**: 850-858.

Fisher, R.A. (1918) The correlation between relatives on the supposition of mendelian inheritance. *R.Soc. (Edinburgh)* **52**:399-433. (En: Quantitative Genetics. Part 1. Ed. W.G.Hill. Van Nostrand Reinhold Company Inc. New York).

Foulley, J.L. (1990). *Estimation des composantes de la variance en modèle linéaire.* Diplôme d'Etudes Approfondis. Génétique Quantitative et des Populations. Université d'Orsay. Paris.

Foulley, J.L. (1993) A simple argument showing how to derive restricted maximum likelihood. *J.Dairy Sci.* **76:** 2320-2324.

Frey, M., Hofer, A., Künzi, N. (1994) Precision of comparisons of estimated breeding values across herds in the genetic evaluation for litter size in pigs. *45th Annual Meeting of EAAP*. Edinburgh.

García, A., Silió, L., Toro, M.A. (1989). Selección para prolificidad: efecto de la correlación genética entre partos. *II Congreso Monográfico Internacional SEPOR/89*. Lorca.

García, J. (1994) *Influència dels efectes materns sobre la prolificitat en porcí*. Proyecto Fin Carrera. Universidad de Lleida.

García-Cortés, L.A. (1992) *Estimación de componentes de varianza y covarianza: contraste e incremento de su eficacia en modelos de genética cuantitativa*. Tesis Doctoral. Universidad de Zaragoza.

Gerstmayr, S. (1991). Data structure for the estimation of maternal effects in an animal model. *42nd Annual Meet. of the European Assoc. of Animal Production.* Berlin.

Gianola, D. (1988) Aspectos metodológicos de la evaluación genética por la prolificidad en el ganado porcino. *I Congreso Monográfico Internacional SEPOR/88*. Lorca.

Gianola, D. (1989) Estimación de componentes de varianza heterogéneos: dos enfoques bayesianos. *II Congreso Monográfico Internacional SEPOR/89*. Lorca.

Gianola, D., Fernando, R.L., Im, S., Foulley, J.L. (1989) Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. *Genome*, **31**: 768-777.

Gianola, D., Foulley, J.L., Fernando, R.L. (1986). Prediction of breeding values when variances are not know. *Gen. Sel. Evol.* **18(4):** 485-498.

Gianola, D., Im, S., Fernando, R.L. (1988). Prediction of breeding value under Henderson's selection model: A revisitation. *J. Dairy Sci.* **71**: 2790-2798.

Gómez, E. (1994). *La selección del tamaño de camada en el conejo de carne: Influencia de los efectos maternos y de la heterogeneidad genética entre partos.* Tesis Doctoral. Universidad Politécnica, Valencia.

Graser, H.-U., Smith, S.P., Tier, B . (1987) A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood. *J.Anim.Sci.* **64** :1362-137.

Groeneveld, E., Kovac, M. (1990). A note on multiple solutions in multivariate restricted maximum likelihood covariance component estimation. *J.Dairy Sci.* **73**: 2221-2229.

Groneveld, E., Kovac, M., Wang, T. (1990). PEST, a general purpose BLUP package for multivariate prediction and estimation. *Proc. of 4th. World Congress on Genetics Applied to Livestock Production*, Edinburgh, **XII**:488-491.

Gu, Y., Haley, C.S., Thompson, R. (1989) Estimates of genetic and phenotypic parameters of litter traits from closed lines of pigs. *Anim. Prod.*, **49**: 477-482.

Haley, C.S., Avalos, E., Smith, C. (1988) Selection for litter size in the pig. *Animal Breeding Abstracts*., **56**: 317-332.

Haley, C.S., Lee, G.J. (1992). Genetic factors contributing to variation in litter size in British Large White gilts. *Livest. Prod. Sci.*, **30**: 99-113.

Hazel, L.N. (1943). The genetic basis for constructing selection indexes. *Genetics*, **28:** 476-490. (En: Quantitative Genetics. Part 2. Ed. W.G.Hill. Van Nostrand Reinhold Company Inc. New York).

Henderson, C. (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics.* **31**: 423-447.

Henderson, C. (1984) *Applications of linear models in animal breeding*. Guelph University Press, Guelph.

Henderson, C. (1988) A simple method to account for selected base population. *J.Dairy Sci.*

71: 3399-3404.

Hill, W.G. (1979) *Design of quantitative genetic selection experiments. (En: Selection experiments in Laboratory and Domestic Animals. Ed: A.Robertson)*

Hill, W.G., Webb, A.J. (1982) Genetics of reproduction in the pig. *Control of Pig Reproduction. (Eds. D.J.A. Cole y G.R. Foxcroft)*: pp. 541-564. Butterworths, London.

Houska, L., Fiedler, J., Pulkrábek, J. (1994) An evaluation of production efficiency of pig nucleus herd by means of bioeconomic model. *45th. Annual Meeting of European Assoc. of Animal Production.* Edinburgh.

Hudson, G.F.S., Kennedy, B.W. (1985). Genetic trend of growth rate and backfat thickness of swine in Ontario. *J. Anim. Sci.* **61(1)**: 92-97.

Im, S., Fernando, R.L., Gianola, D. (1989) Likelihood inferences in animal breeding under selection:a missing-data theory view point. *Gen. Sel. Evol.* **21**:399-414.

Irgang, R., Fávero, J.A., Kennedy, B.W. (1994). Genetic parameters for litter size of different parities in Duroc, Landrace and Large White sows. *J. Anim. Sci.* **72**:2237-2246.

Janss, L.L.G., Van Arendonk, J.A.M., Van der Waaij, E.H., Visscher, A.H. (1994) Estimation of genetic parameters using Gibbs sampling with an application to direct and maternal genetic variance for weaning weight in sheep. *45th Annual Meeting of European Assoc. of Animal Production*, Edinburgh.

Jensen, J., Wang, C.S., Sorensen, D.A., Gianola, D. (1994). Bayesian inference on variance and covariance components for traits influenced by maternal and direct genetic effects, using the Gibbs sampler. *Acta. Agric. Scand.* **44**: 193-201.

Johansson K., Kennedy, B.W. (1985) Estimation of genetic parameters for reproductive traits in pigs. *Acta Agric. Scand.* **35**: 421-431.

Kaplon, M.J., Rothschild, M.F., Berger, P.J., Healey, M. (1991) Population parameter estimates for performance and reproductive traits in Polish Large White nucleus herds. *J. Anim. Sci.* **69**: 91-98.

Kaylen, M.S., Preckel, P.V., Loehman, E.T. (1987) Risk modeling via direct utility maximization using numerical quadrature. *Amer. J. Agr. Econ.* **69:** 701-706.

Kendall, M., Stuart, A. (1979). *The advanced theory of statistics. Vol. 2. Inference and Relationship*. Charles Griffin and Company Limited, High Wycombe.

Kennedy, B.W., (1989). *Animal model BLUP. Erasmus intensive graduate course.* Trinity College, Dublin.

Kirkpatrick, B.W., Rutledge, J.J. (1988) Influence of prenatal and postnatal fraternity size on reproduction in swine. *J. Anim. Sci.* , **66** : 2530-2537..

Knap, P.W., van Alst, G.J.M., Versteeg, J.G., Kanis, E. (1993) Realized genetic improvement of litter size in Dutch Pig Herdbook breeding. *Pig News and Information,* **14** (3): 119N-121N.

Lamberson, W.R. (1990) Genetic parameters of quantitative characters. *Genetics of swine.* L.D. Young (Ed.), Nebraska. pp.70-106.

Lamberson, W.R., Johnson, R.K., Zimmerman, D.R., Long, T.E. (1991) Direct responses to selection for increased litter size, decreased age at puberty, or random selection following selection for ovulation rate in swine. *J. Anim. Sci* . **69** :3129-3143.

Le Roy, P., Legault, C., Gruand, J., Ollivier, L. (1987) Héritabilité réalisée pour la taille de la portée dans la sélection des truies dites "hyperprolifiques". *Gen. Sel. Evol.* , **19 (3):** 351-364.

Legault, C. (1970). Étude statistique et génétique des performances d'élevage des truies de

race Large White. *Ann. Génét. Sél. Anim.* , **2**: 209-227.

Legault, C. (1978) Génetique et réproduction chez le porc. *Journées de la Recherche Porcine en France* . **13** :43-60.

Legault, C. (1987). Selección y mejora de caracteres de reproducción en el porcino. *Proc. IX Symposium Anaporc: 50-57, Barcelona.*

Legault, C., Gruand, J. (1976) Amélioration de la prolificité des truies par la création d'une lignee "hiperprolifique " et l'usage de l'insemination artificielle: principe et résultats experimentaux préliminaires. *Journées de la Recherche Porcine en France*, **8**: 201-206.

Legault, C., Gruand, J., Bolet, G. (1981) Résultats de l'utilisation en race pure et en croisement de la lignée dite "hyperprolifique". *Journées de la Recherche Porcine en France*, **13**: 255-260.

Lin, C.Y. (1988) Four equivalent sets of mixed-model equations with relationship matrix foe estimation of genetic parameters. *J.Anim.Sci.* **66**: 1627-1635.

Lo, L.L., Fernando, R.L., Grossman, M. (1993) Covariance between relatives in multibreed populations: additive model. *Theor. Appl. Genet.* **87**: 423-430.

Long, T.E., Johnson, R.K., Keele, J.W. (1991) Effects of selection of data on estimates of breeding values by three methods for litter size, backfat, and average daily gain in swine. *J.Anim.Sci.* **69**: 2787-2794.

Luján, F.J. (1994). *Estructuras actuales en granjas de selección de porcino.* Trabajo Final Carrera. Universidad de Lleida.

Madsen, P., Jensen, J., Thompson, R. (1994). Estimation of (co)variance components by REML in multivariate mixed linear models using average of observed and expected information. *Proc. 5th World Congress on Genetics Applied Livestock Production,* Guelph. **XXII**: 19-22.

Mallo, F., Mures, M.J., Huerga, C. (1987). *Decisión empresarial.* Universidad de León. Serv.Public.

Mercer, J.T., Crump, R.E. (1990) Genetic parameter estimates for reproduction traits in purebred Landrace pigs. *Proc. 4th. World Congress on Genetics Applied to Livestock Production*, Edinburgh, **XV**:489-492.

Meuwissen, T.H.E. (1990). *Optimization of dairy cattle breeding plans with increased female reproductive rates.* Ph.D. Thesis. Agricultural University, Wageningen.

Meuwissen, T.H.E. (1991). Expectation and variance of genetic gain in open and closed nucleus and progeny testing schemes. *Anim. Prod.* **53:** 133-141.

Meyer, K. (1989) Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. *Gen. Sel. Evol.* **21**: 317-340.

Meyer, K. (1990) Present status of knowledgement about statistical procedures and algorithms to estimate variance and covariance components. *Proc. 4th World Congress on Genetics Applied Livestock Production,* Edinburgh, **XIII:** 407-418.

Meyer, K. (1991a) Estimating variances and covariances for multivariate animal models by restricted maximum likelihood. *Genet. Sel. Evol.* **23**: 67-83.

Meyer, K. (1991b) DFREML. Programs to Estimate Variance Components by Restricted Maximum Likelihood Using a Derivative-Free Algorithm. User Notes v.2.0.

Meyer, K. (1992) Bias and sampling covariances of estimates of variance components due to maternal effects. *Genet. Sel. Evol.* **24**: 487-509.

Misztal, I. (1994). Comparison of software packages in animal breeding. *Proc. 5th World Congress on Genetics Applied Livestock Production, Guelph.* **XXII**: 3-10.

Morgan, B.J.T. (1984). *Elements of simulation.* Champman and Hall, London.

Nelson, R.E., Robison, O.W., 1976. Effects of postnatal maternal environment on reproduction of gilts. *J. Anim. Sci.* **43(1)**: 71-77.

Noguera, J.L., Estany, J., Babot, D., Alfonso, L., Galofré, R., Pomar, J. (1992). GTEP-IRTA: a pig information system for technical and selection management and research. *Proc. Satellite symposium on pig management information systems,* Madrid. Monografías INIA, **68** : 49-57.

Noguera, J.L., Pérez-Enciso, M., Alfonso L., Babot, D., Estany, J. (1994) A selection experiment for increasing litter size of Landrace pigs in Spain. *Proc. 5th World Congress on Genetics Applied Livestock Production, Guelph.* **XVII**: 343-346.

Ollivier, L. (1982) Selection for prolificacy in the pig. *Pig News and Information*, **3**: 383-388.

Ollivier, L., Bolet, G. (1981) La sélection sur la prolificité chez le porc: résultats d'une expérience de sélection sur dix générations. *Journées de la Recherche Porcine en France*, **13**: 261-268.

Patterson, H.D., Thompson, R. (1971) Recovery of inter-block information when blok size are unequal. *Biometrika.* **58**: 545-554.

Pérez-Enciso, M., Gianola, D. (1992). Estimates of genetic parameters for litter size in six strains of Iberian pigs. *Livest. Prod. Sci.*, **32**: 283-293.

Pérez-Enciso, M., Tempelman, R., Gianola, D. (1993). A comparison between linear and Poisson mixed models for litter size in Iberian pigs. *Livest. Prod. Sci.*, **35**: 303-316.

Pieramati, C., Van Vleck, L.D. (1993). Effect of genetic groups on estimates of additive genetic variance. *J.Anim Sci.* **71:** 66-70.

Prieto, E. (1973). *Teoría de la inversión.* Ediciones ICE. Madrid.

Quaas, R.L. (1988). Additive genetic model with groups and relationships. *J. Dairy Sci.* **71:** 1338-1345.

Quaas, R.L., Anderson R.D., Gilmour, A.R. (1984). *BLUP school handbook.Use of mixed models for prediction and estimation of (co)variance components*. University of New England.

Quintanilla, R., Pujol, M.R., Piedrafita, J. (1995). Estimación de componentes de varianza para los effectos directos y maternos sobre el peso al destete: análisis en la población bovina Bruna dels Pirineus. *VI Jor. Prod. Animal, AIDA,* Zaragoza. *ITEA,* **16:** 230-232.

Revelle, T.J., Robinson, O.W. (1973) An explanation for the low heritability of litter size in swine. *J. Anim. Sci.* **37(3)**: 668-675.

Robertson, A. (1977) The effect of selection on the estimation of genetic parameters. *Z. Tierzüchtg. Züchtgsbiol,* **94**: 131-135.

Robinson, G.K. (1991). That BLUP is a good thing: the estimation of random effects. *Statistical Science.***6** (1): 15-51

Robinson, O.W. (1972). The role of maternal effects in animal breeding: V. Maternal effects in swine. *J. Anim Sci.,* **35**: 1303-1315.

Robinson, O.W. (1981). The inluence of maternal effects on selection: a review. *Livest. Prod. Sci.* **8**: 121-137.

Rodríguez, C., Béjar, F., Rodrigañez, J., Silió, L. (1993) Componentes de varianza, heterosis y depresión consanguínea en el tamaño de camada de cerdos ibéricos. *Invest.Agr.: Prod. Sanid. Anim.* **8** (1): 43-53.

Rodríguez, C., Rodrigañez, J., Silió., L. (1994). Genetic analysis of maternal ability in Iberian pigs. *J. Anim. Breed. Genet.,* **111:** 220-227.

Roehe, R., Kennedy, B.W. (1993). The influence of maternal effects on accuracy of evaluation of litter size in swine. *J. Anim. Sci.* **71**: 2353-2364.

Rogers, G.W. (1990) A utility function for ranking sires that considers production, linear type traits, semen cost, and risk. *J. Dairy Sci.* **72:** 532-538.

Romero, C.(1993). *Teoría de la decisión multicriterio: Conceptos, técnicas y aplicaciones.* Alianza Editorial, Madrid.

Ruiz, J. (1993). *Heredabilidad del tamaño de la camada en poblaciones porcinas españolas.* Proyecto Fin Carrera. Universidad de Lleida.

Rutledge, J.J. (1980a) Fraternity size and swine reproduction. I. Effect on fecundity of gilts. *J.Anim.Sci.* **51(4)**: 868-870.

Rutledge, J.J. (1980b) Fraternity size and swine reproduction. I. Genetical consequences. *J.Anim.Sci.* **51(4)**: 871-874.

Rydhmer, L. (1993). *Pig reproductive genetics and correlation between reproduction and production traits.*Thesis. Swedish University of Agricultural Sciences, Uppsala.

San Cristóbal, M. (1992). *Méthodes d'inférence statistique en modélisation de la variance. Application en génétique quantitative*. These de Doctorat d'Université. Université Paul Sabatier, Toulouse. **SAS Inc.** (1987). *SAS User's guide: Statistics.* SAS Institute Inc., Cary, NC.

Schaeffer, L.R. (1984). Sire and cow evaluation under multiple traits model. *J.Dairy Sci.* **67(7)**: 1567-1580.

Searle, S.R. (1982). *Matrix algebra useful for statistics.* John Wiley & Sons. New York.

Searle, S.R. (1989) Variance components - some history and a summary account of estimation methods. *J. Anim. Breed. Genet.,* **106:** 1-29.

Searle, S.R., Casella, G., McCulloch, C.E. (1992). *Variance components.* John Wiley & Sons, New York.

Sehested, E., Ianssen, K. (1993). Multivariate genetic evaluation in Norwegian pig breeding. *Application of mixed linear models in the prediction of genetic merit in pigs.* E. Groeneveld (Ed.) Commission of the European Communities. 1995. pp. 47-54.

Short, T.H., Wilson, E.R., McLaren, D.G. (1994). Relationships between growth and litter traits in pig dam lines. *Proc. 5th World Congress on Genetics Applied Livestock Production,* Guelph. **XVII**: 413-416.

Sjkervold, H. (1979). What about the genetic improvement of litter size?. *Acta Agr. Scand.* **21** (Suppl.): 176-184.

Smith, J.Q. (1988). *Decision Analysis. A Bayesian approach.* Chapman and Hall, London.

Smith, S.P., Graser, H.-U. (1986) Estimating variance components in a class of mixed models by restricted maximum likelihood. *J. Dairy Sci.*, **69**: 1156-1165.

Smith, S.P., Hammond, K. (1987). Portfolio theory, utility theory and mate selection. *Gen. Sel. Evol.* **19**: 321-326.

Sorensen, D.A., Kennedy, B.W. (1984a) Estimation of genetic variances from unselected and selected populations. *J. Anim. Sci.* **59**: 1213-1223.

Sorensen, D.A., Kennedy, B.W. (1984b) Estimation of response to selection using least-squares and mixed model methodology. *J. Anim. Sci.* **58**: 1097-1106.

Sorensen, D.A., Vernersen, A.H. (1991) Large escale selection for number of born piglets using an animal model. *42nd Annual Meeting of the European Association of Animal Production.* Edinburgh.

Southwood, O.I., Kennedy, B.W. (1990) Estimation of direct and maternal genetic variance for litter size in Canadian Yorkshire and Landrace swine using an animal model. *J. Anim. Sci.* **68**: 1841-1847.

Spilke, J., Groeneveld, E. (1994). Comparison of four multivariate REML (co)variance component estimation packages. *Proc. 5th World Congress on Genetics Applied Livestock Production, Guelph.* **XXII**: 11-14.

Stewart, T.S., Diekman, M.A. (1989). Effect of birth and fraternal litter size and cross-fostering on growth and reproduction in swine. *J. Anim. Sci.*, **67**: 635-640.

Strang, G.S., Smith, C. (1979). A note on the heritability of litter traits in pigs. *Anim. Prod.*, **28**: 403-406.

Tartar, M. (1981). *Analyse critique des possibilités d'estimation du potenciel génétique des truies reproductrices*. Diplôme d'Etudes Approfondis. Génétique quantitative et appliquée. Université Paris XI. Paris.

Tartar, M., Bolet, G. (1984). Application de la théorie des indices de sélection à des caractères répétés. Example de la sélection sur la prolificité chez le porc. *Gen. Sel. Evol.*, **16**: 319-334.

Thompson, R. (1973) The estimation of variance and covariance components with an application when records are subject to culling. *Biometrics*, **29**: 527-550.

Thompson, R., Atkins, K.D. (1990). Estimation of heritability in selection experiments. *Proc. 4th World Congress on Genetics Applied Livestock Production*, **XV**: 257-260.

Urban, W.E., Shelby, C.E., Chapman, A.B., Whatley, J.A., Garwood, V.A. (1966). Genetic and environmental aspects of litter size in swine. *J.Anim. Sci.*, **25**: 1148-1153.

Van der Steen, H.A.M. (1983*) Maternal and genetic influences on production and reproduction traits in pigs.* Ph.D. Thesis. Agricultural University, Wageningen.

Van der Werf, J. (1992). Restricted maximum likelihood estimation of additive genetic variance when selected base animals are considered fixed. *J. Anim. Sci.*, **70**: 1068-1076.

Van der Werf, J., de Boer, I.J.M. (1990) Estimation of additive genetic variance when base populations are selected. *J. Anim. Sci*. **68**: 3124-3132.

Van der Werf, J., Thompson, R. (1992). Variance decomposition in the estimation of genetic variance with selected data. *J. Anim. Sci.*, **70**: 2975-2985.

Van der Werf, J.H.J., van Arendonk, J.A.M., De Vries, A.G. (1992). Improving selection of pigs using correlated characters. *43th Annual Meeting of the European Assoc. of Animal Production,* Madrid.

Van Vleck, L.D. (1993). *Selection index and introduction to mixed model methods*. CRC Press, Boca Raton.

Van Vleck, L.D., Gregory, K.E. (1992). Differences in heritability estimates from multiple-trait and repeated-records model. *J. Anim. Sci.,* **70:** 2994-2998.

Vandepitte, W., Hazel, L. (1977) The effect of errors in the economic weights. *Ann. Génét. Sél. Anim.*, **9 (I):** 87-103.

Vangen, O. (1980). Studies on a two trait selection experiment in pigs. VI. Heritability estimates of reproductive traits. Influence of maternal effects. *Acta Agr. Scand.* **30**: 320-326.

Vangen, O. (1981). Problems and possibilities for selection for fecundity in multiparous species. *Pig News and Information* **2(3)**: 257-263.

Vangen, O. (1986). Genetic control of reproduction in pigs: from parturition to puberty. *Proc. 3th World Congress on Genetics Applied Livestock Production,* Nebraska. **XI:** 168 179.

Visscher, P.M. (1994). Bias in genetic R^2 from halfsibs designs. *Proc. 5th. World Congress on Genetics Applied Livestock Production,* Guelph, **XVIII:** 394-397.

Visscher, P.M., Thompson, R. (1990). REML estimates of parameters for fat yield in

pedigree herds in the U.K. using an individual animal model; male and female heritability estimates. *Proc. 4th. World Congress on Genetics Applied Livestock Production,* Edinburgh, **XIV:** 233-236.

Wang, C.S., Gianola, D., Sorensen, D.A., Jensen, J., Christensen, A., Rutledge, J.J. (1994). Response to selection for litter size in Danish Landrace pigs: a Bayesian analysis. *Theor. Appl. Genet.* **88:** 220-230.

Webb, A.J., Bampton, P.R. (1989). Choice of selection objectives in specialized sire and dam lines for commercial crossbreeding. *39th Annual Meeting of European Assoc. of Animal Production.* Lisboa.

Weller, J.I. (1994) *Economic aspects of animal breeding.* Chapman and Hall. London.

Willham, R.L. (1963) The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics*. **19** :18-27. (En: Quantitative Genetics. Part 1. Ed. W.G.Hill. Van Nostrand Reinhold Company Inc. New York).

Willham, R.L. (1972) The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. *J. Anim. Sci.* **35(6)**: 1288-1293.

Willham, R.L. (1980). Problems in estimating maternal effects. *Livest. Prod. Sci.* **7**: 405-418.

Wray, N.R. (1989). *Consequences of selection in finite populations with particular reference to closed nucleus herds of pig.* PhD. University of Edinburgh.

Wrigth, S. (1921). Systems of mating. I. The biometrical relations between parent and offspring. *Genetics* **6**: 111-123.(In: Quatitative Genetics: Part 1. Ed. W.G.Hill. Van Nostrand Reinhold Company Inc. New York).

Wyler, A., Hagger, C. Stranzinger, G. (1994). Different risk-utility functions in animal breeding. *Proc. 5th World Congress on Genetics Applied Livestock Production, Guelph.* **XVIII**: 244-247.

Anejo

Deducción de una expresión general de las ecuaciones del Modelo Animal cuando la población base está seleccionada.

A partir de la expresión del Modelo Animal y diferenciando entre animales de la población base (b) y el resto de animales (r) (Graser *et al.,* 1987) tenemos:

Figura 10-1 [1]

Bajo un modelo infinitesimal (Fisher, 1918) la expresión del valor genético de un individuo (a_i) en base a los valores genéticos de padre (a_p) y madre (a_m), y el muestreo mendeliano (s_i) es:

$$a_i = ½\, a_m + ½\, a_p + s_i$$

Esta expresión nos permite escribir $a_r$de [1] como :

Figura 10-2 [2]

donde:

- P_1 y P_2 son matrices que relacionan cada individuo con sus padres, por lo que tendrán como elementos coeficientes 0 o ½.
- s representa el valor aleatorio de segregación o muestreo mendeliano.

A partir de la expresión [2], y llamando:

1) $\mathbf{Q = (I - P_2)^{-1} \bullet P_1}$

que no es más que la matriz que relaciona los animales de la población base con sus descendientes, y que corresponde a la utilizada por Quaas (1988) para tratar la existencia de grupos genéticos diferenciados en una población,

y 2) $\mathbf{s^* = (I - P_2)^{-1} \bullet s}$

podemos encontrar una nueva expresión de a_r:

Figura 10-3 [3]

A partir de [1] y [3] podemos expresar el Modelo Animal como (Graser *et al.* 1987):

Figura 10-4 [4]

Esta expresión es útil en la descomposición de la varianza fenotípica, al poder asumir distintas hipótesis sobre las características de los animales conocidos de la población base. La varianza de **y** se puede expresar según [4] en:

Figura 10-5 [5]

en donde:

1) Figura 10-6 [6]

2) Figura 10-7 [7]

siendo Figura 10-8 [8]

3) Figura 10-9 [9]

Asumiendo que no existe consanguinidad[44], nos podemos referir a H_0 y H_b como:

- H_0 una matriz de elementos ½ en la diagonal y 0 fuera de la diagonal (Wrigth, 1921).
- H_b una matriz que recoge la existencia de selección y parentesco en los animales de la población base conocida.

Llamando ahora:

Figura 10-10

y considerando la expresión [8] podemos escribir [7] como:

Figura 10-11 [10]

Usando [6], [9] y [10] podemos expresar la varianza de las observaciones [5] como:

[44] Con el objeto de simplificar las expresiones que se obtienen.

Figura 10-12 [11]

que nos permite ver que depende tanto de la varianza (δ^2_{a0}) como de las características (H_b) de la población base conocida, es decir de la existencia, en ella, de selección y/o parentesco.

Atendiendo a las ecuaciones del modelo mixto (MME) podemos escribir, en relación a la expresión [4]:

Figura 10-13

se puede llegar a una expresión equivalente a la de Van der Werf y Thompson (1992):

Figura 10-14

Dado que $Z'_i Z_j=0$, esta expresión se puede simplificar a:

Figura 10-15 [12]

A partir de esa expresión [12], haciendo operaciones algebraicas, se puede hallar una expresión simplificada; llamando M a la matriz de coeficientes, b al vector de incógnitas, y c al vector de la parte derecha, es decir:

$$Mb=c$$

y aplicando la transformación utilizada por Graser *et al.* (1987):

$$TMb=Tc$$

o lo que es lo mismo,

$$TMT'[(T')^{-1}b]=Tc$$

donde T es:

Figura 10-16

se puede expresar [12] com[45]Ver Figura 10-17(Figura 10-17):

Figura 10-18 [13]

De las expresiones anteriormente citadas de S y Q se puede deducir:

$$S^{-1}= 2 (I- P_2)' (I - P_2)$$
$$-S^{-1} Q = 2P_1 (I - P_2)$$
$$Q' S^{-1} Q= 2P_1'P_1$$

con lo que [13] también se puede expresar como:

Figura 10-19

o lo que es lo mismo,

Figura 10-20 [14]

De esta expresión [14], si nos fijamos en la submatriz:

Figura 10-21 [15]

podemos ver su similitud con la inversa de la matriz de parentesco. Siguiendo a Henderson (1984), la matriz de parentesco se define, cuando la población base es una muestra tomada al azar, como:

Figura 10-22

Haciendo uso de las matrices P_1, P_2 y Q, la podemos expresar como:

Figura 10-23

cuya inversa es:

Figura 10-24

Es decir, llamando:

1) Figura 10-25 [16]

2) Figura 10-26

podemos expresar la inversa de la matriz de parentesco cuando la población base es una muestra tomada al azar como:

---

[45] Siguiendo los siguientes pasos:

$$A^{-1} = A^{*-1} + H^{*-1}$$

La construcción directa de A^{*-1} es muy sencilla: sus elementos son los mismos que los de la inversa de la matriz de parentesco A^{-1}, exceptuando los de la diagonal. Así, basta con atribuir valor 0 en vez de 1 a los animales de los que no conocemos ni padre ni madre, es decir cuando se trata de un animal de la población base.

Ahora retomando la expresión [15] podemos ver que la podemos construir directamente utilizando las reglas de Henderson (Henderson, 1984) con la modificación comentada, y que además nos permite considerar la existencia de selección y parentesco en la población base, pues la matriz H_b^{-1} no tiene por que ser una matriz identidad, sólo se debe cumplir que H_b no sea singular para que exista.

Llamando:

1) $Z = [Z_1\ Z_2]$

2) Figura 10-27

y usando la expresión [16], podemos expresar [14] como:

Figura 10-28 [17]

que es la expresión general de las ecuaciones del Modelo Animal para cualquier tipo de población base.

Así, de esa expresión [17] se derivan distintos casos particulares:

- Cuando la población base es una muestra tomada al azar de animales no emparentados, $H_b^{-1} = I$, y la expresión [17] no es más que la expresión habitual de las MME.

- Cuando la selección de los animales base se puede describir mediante una función lineal del tipo:

Figura 10-29

la expresión [17] es equivalente a la planteada por Henderson (1988).

- Cuando los animales de la población base se tratan como efectos fijos, $H_b^{-1} = 0$, de modo que la expresión [17] no es más que la derivada por Graser *et al.* (1987).

- Cuando se considera un modelo con grupos genéticos se puede representar cada grupo mediante un 'padre fantasma' siguiendo el modelo de Westell-Robinson. En ese caso, si tomamos $H_b^{-1} = 0$ para todos esos 'padres fantasmas', la expresión [17] equivale a tratar los grupos genéticos como efectos fijos, siendo la expresión derivada por Quaas (1988).

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES