

Genetic Analyses of Days Open Using a Random Regression Model

A.S. Leaflet R2304

Gustavo A. Gutierrez, graduate research assistant;
Mary H. Healey, program assistant II;
P. Jeff Berger, professor of animal science

Summary and Implications

The purpose of this study was to investigate the feasibility of using a linear random regression model for the analysis of female fertility as expressed by days open. The analysis provided insight about the kind and amount of variation over the trajectory for age at calving from 24 to 72 mo. Correlations between permanent environmental effects changed from positive to negative values with greater distance between days open in young cows and days open in older cows; $r(24,36)=0.88$, $r(24,72)=-.65$. The random regression animal model proved to be a better alternative to a repeatability animal model for prediction of breeding values for days open. Random regression enables breeders to pick the optimum age at calving to select female replacements with enhanced fertility or fewer days open per lactation. Further research is needed to evaluate the advantages of using random regression model on the genetic analysis of days open.

Introduction

Genetic progress for milk yield has increased since 1960 due to the higher selection intensity for sires and higher reliability of sire's Predicted Transmitting Ability (PTA) from modern progeny testing schemes. Female fertility, however, has declined in Holsteins (Lucy et al., 2001). Therefore, there has been a concerted effort among dairy geneticists to find appropriate models and statistical techniques for genetic analysis of common measures of reproductive performance, e.g., days open (DO) (Oseni et al., 2004, Gonzales-Recio et al., 2006), number of services (Chang et al. 2006), or non-return rate (Heringstad et al., 2006). Genetic evaluation of sires for daughter pregnancy rate was implemented in 2004 at the national level (Van Raden et al, 2004). These genetic evaluations are published annually for sires available through the artificial insemination industry in the United States.

Female fertility data in dairy cattle is characterized by having repeated measurements per cow and for being highly unbalanced (i.e., variable number of records per cow), because a great proportion of cows leave their herds of origin following their first, second, third or later parity due to reproductive failure. Recently, random regression models (RR) have been applied in the analysis of repeated measurement data where the trajectory of time was either

days in milk or parity (e.g., daily milk yield, Schaeffer, 2004; number of services, Nishida et al 2006). Therefore, the objectives of this research were to implement techniques for using RR to analyze repeated measurements of DO across the trajectory of time when each measurement is expressed, e.g., parity (1, 2, 3 or later). RR modeling enables researchers to estimate mathematical functions (formally called covariance function (CF)) of the (co)variance between all possible points on the trajectory of time when each measurement is taken. The CF is a mathematical function to give estimates of all possible combinations of (co)variance measures at points on the trajectory (Van der Werf, 2001). The CF is estimated from the (co)variance matrix of the RR coefficients (Meyer and Hill, 1997). The objective of this study was to estimate genetic parameters (heritability (h^2), additive genetic variance, permanent environmental variance (PE)) for DO by using a covariance function (CF)-random regression model.

Materials and Methods

Data in this study were from cows enrolled in a long term sire selection project designed to explain the genetic changes in progeny of sires selected for either high or average PTA-Fat+Protein. There were two concurrent lines of cattle and the overall long-term objective of selection was to show the difference between progeny of high and average sires. Three new sires were selected for each line each year. Data were collected at the Ankeny dairy research farm at Iowa State University from 1986 to 2004. A total of 3830 records from Holstein cows ($n=766$) were used in the analysis. Data were restricted to less than 6 parities. All cows were required to have a first parity record, but not all cows had an opportunity to complete all parities. Data were analyzed using two models: 1) a repeatability animal model, fixed effects for line of sire selection (high or average PTA-Fat+Protein), year-season, parity, and linear-Legendre polynomial for age at calving (mean=39, min=20, max= 90, SD=16 mo), random effects for animal and permanent environment; and 2) RR model, fixed effects as defined above for model 1, random linear Legendre polynomial coefficients for additive genetic and permanent environment effects. Analyses were implemented by using ASREML v1.1. Due to convergence problems the genetic correlation between records at different ages was fixed at 0.99. The covariance matrix of random regression coefficients was used to define the CF for the additive genetic and permanent environmental (co)variances across the trajectory of age at calving from 20 to 90 mo.

Results and Discussion

Table 1 gives the number of cows and percentage of cows with missing records by parity. By definition, heifers (parity=0) do not have a value for DO. Thus only proven fertile heifers become lactating cows where DO becomes available. This may be a criticism of DO as a measure of fertility early in the reproductive life of dairy cows, because it does not address the dual nature of fertility, i.e., conception and the ability to maintain pregnancy. The percentage of cows with missing records tended to increase with parity due to greater involuntary culling among older aged cows.

Table 2 gives the mean DO by parity for the two lines. There was an increasing trend for DO in cows that had records for all 5 parities in both selection lines, but previous statistical analysis showed no significant ($P \leq 0.05$) difference between selection lines for DO.

Table 3 compares estimates of parameters at fixed ages. In model 1, estimated values of heritability and repeatability

for DO were 0.08 and 0.09, respectively. In model 2, the additive genetic variance tended to increase with age at calving (Figure 1) and the heritability estimates increased slightly from 24 to 72 mo. The percentage of phenotypic variance explained by permanent environmental variance varied from 3 to 19% across the 5 fixed ages. Repeatability ranged from 0.17 to 0.29. Permanent environmental correlations were highly variable. For example, the PE correlations between DO at 24 mo and 36, 48, 60, and 72 mo were: 0.88, 0.06, -0.48, and -0.65, respectively

Table 1. Number of cows and percentage of missing records per parity.

	Parity					
	0	1	2	3	4	5
Cows(n)	1049	766	528	322	169	67
% Missing records	-	27	50	69	84	94

Table 2. Means and standard deviation for days open from cows that have all 5 parity records.

Parity	AFP (n=27) ¹					HFP(n=40) ²				
	1	2	3	4	5	1	2	3	4	5
DO	116(77)	105(52)	116(63)	150(78)	174(97)	126(66)	117(56)	126(58)	130(60)	157(85)
Age at calving (mo)	25(1)	38(3)	50(3)	63(3)	77(4)	25(2)	39(3)	52(4)	65(5)	78(5)

¹ AFP = mean DO for progeny of average PTA fat + protein sires

² HFP = mean DO for progeny of high PTA fat + protein sires

Table 3. Heritability (h^2 , (SD)) for days open.

Age (mo)	Model I ¹ h^2	Model II ²		
		h^2	% Var (PE)	Repeat-ability
24	0.08 (0.03)	0.07	11.	0.17
36	0.08 (0.03)	0.08	4.	0.12
48	0.08 (0.03)	0.09	3.	0.13
60	0.08 (0.03)	0.10	9.	0.19
72	0.08 (0.03)	0.10	19.	0.29

¹ Simple repeatability animal model.

² Random regression animal model.

Acknowledgments

Financial support was provided by the Iowa Agriculture Experiment Station, Ames, IA, Project Number 6623, and by Hatch Act and State Iowa funds. Special thanks to Joe Detrick for managing the Ankeny dairy farm and for his assistance collecting data. G. Gutierrez's graduate studies were partially sponsored by the Fulbright Commission and Universidad Nacional Agraria La Molina, Peru.

Figure 1. Additive and permanent environment variances for days open.

