ANALYSIS OF LENGTH OF PRODUCTIVE LIFE OF MEXICAN HOLSTEIN COWS USING A FRAILTY MODEL.

F. J. Ruiz and B. Trejo-Valdivia
Centro Nacional de Investigaciones en Fisiología y Mejoramiento Animal-INIFAP-SARH
and Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas-UNAM.
MEXICO.

SUMMARY

A previous work reported the importance of a sire effect in the modeling of length of productive life of Mexican Holstein cows with a Weibull proportional hazards model. Another approach to this problem was taken in this paper which consisted of using a frailty model, assuming a distribution for the random frailty term, estimating the distribution parameters, estimating individual frailties through their expectations and comparing these to the sire effects of the Weibull model. The results showed a good agreement between individual frailty terms and sire effects and further reinforces the possibility of genetic improvement of length of productive life in this population.

INTRODUCTION

The modeling of the instantaneous probability of culling (called hazard) has been used for the analysis of length of productive life of dairy cows (LPL), most recently by Ducrocq et al. (1988a,b) and Ruiz et al. (1994). Hazard models with an unobserved random term (called frailty models and frailty respectively) were proposed by Vaupel et al. (1979) and were recently used by Ruiz and Trejo (1993) to examine the presence of a frailty term in LPL of Mexican Holstein cows. These authors found evidence of the need of an extra term in the model and concluded that the cow's sire effect could reproduce the missing term.

If a distribution is assumed for the frailty term in the model, this term can be estimated, and later analyzed to investigate if there is a relationship with the sire's term.

MATERIALS AND METHODS

A frailty model with a Weibull baseline hazard was fitted to describe the length of productive life of 36477 Mexican dairy cows. The model was of the form:

$$h(t|\mathbf{z}) = h_0(t)\, e^{(\beta' \mathbf{z})}\, \xi$$

where h(t|z) represents the hazard at time t given a set of covariates z, β is a vector of unknown parameters measuring the log-linear effect of z, $h_0(t)$ represents the Weibull baseline hazard function, and ξ is the random term modelling the effect of relevant covariates (independent of z) out of the model but affecting the survival time. Covariates in the model were sire's country of origin (United States, Canada and Mexico), within herd-year standardized level of production (9 levels) and herd-year effect (770 herd-years). In order to associate the missing effects with the sire effects, we fitted a Weibull proportional hazards model with the previous terms plus a sire effect and used these results to analyze the frailty model in more detail. The model was then, of the form:

$$h(t|z,\ sire\ i) = h_0(t)e^{(\beta' z)}e^{\gamma_i}$$

where β and $h_0(t)$ are as before and e^{γ_i} represents the i-th sire effect.

A gamma distribution was assumed for the frailty term with unit mean and variance Θ. The generalized residuals from the frailty model were then used to calculate the maximum likelihood estimate of the frailty variance. This was done using the procedure suggested by Trejo (1991) in which a third order Taylor's expansion around zero is considered.

The gamma family is closed under the follow-up analysis, this implies that the conditional frailty distribution among the cullings at time t and covariates z is also gamma with parameters,

$$\theta^{-1} + 1 \quad and, \quad \theta^{-1} + r(t,z)$$

with $r(t,z) = \Lambda_0(t)\, e^{\beta' z}$ the generalized residual.

Also, the conditional frailty distribution among the survivors at time t and covariates z is gamma with parameters,

$$\theta^{-1} \quad and, \quad \theta^{-1} + r(t,z)$$

These two distributional results, allow us to estimate the individual frailty. Two candidates might be considered, the mean or, as proposed by McGilchrist and Aisbett (1991), the mode. However, a general analysis of the resulting estimates showed that

the mean gives a better representation of the situation. Therefore, the estimated frailty for an individual with covariates z and censoring indicator δ is given by,

$$\hat{\xi} = E\left[\xi \mid \hat{r}(t,z),\delta\right] = \frac{\hat{\theta}^{-1} + \delta}{\hat{\theta}^{-1} + \hat{r}(t,z)}$$

with $\hat{r}(t,z) = \hat{\Lambda}_0(t)\, e^{\beta' z}$.

Differences of individual frailties from sire effects can be used to evaluate the similarity between both effects. The variation within sire can also be seen as the degree of homogeneity among daughters of a sire.

RESULTS AND DISCUSSION

The estimate of the variance of the frailty terms (Θ) was .260 for these data, and it was later used to estimate the individual frailty terms as described before.

Expected frailty terms ranged from 0.338 to 1.260 with an average of 1.025 result that was in agreement with the fact that these terms come from a gamma distribution.

The correlation between individual frailties and their observed LPL was large and negative (-.709), result that was also expected because larger frailties imply greater hazards and shorter LPL's. The average overall difference of the estimated individual frailty and their respective sire regression coefficient was close to 0 (-.007, s.e. .002), finding that reinforces the conclusion of correspondence of frailty and sire effects.

To study the effect within sire, the average individual frailty within sire was calculated for all sires and later compared with the sire effect of the proportional hazards model. The mean difference between the average of individual frailties and their sire effect was -0.042 (s.e. .015). The standard deviation within sires for the above mentioned differences ranged from .004 to .052.

In general these findings further support the importance of a sire effect in the description of LPL of Mexican cows, and with this, the possibility of genetic improvement. Ducrocq et al. (1988b) used a somewhat similar model on cows from the North East of the United States. In their work, they considered z as a step function of time and the sire effect as a random effect gamma distributed. They estimated the gamma parameter equal to .018 which is smaller than the .260 estimated in this work.

Additional studies will be needed to asses if this discrepancy is due to differences in the models used or differences in the populations.

REFERENCES

DUCROCQ, V., QUAAS, R.L., POLLACK, E.J., and CASELLA, G. (1988a) J. Dairy Sci. 71:3061.

DUCROCQ, V., QUAAS, R.L., POLLACK, E.J., and CASELLA, G. (1988b) J. Dairy Sci. 71:3071.

McGILCHRIST, C.A. and AISBETT, C.W. (1991) Biometrics 47:461.

RUIZ, F.J. and TREJO-VALDIVIA, B. (1993) Proceedings of the 1993 Annual Meeting of the European Association of Animal Production. Aarhus, Denmark. (Abs.)

RUIZ, F.J., OLTENACU, P.A. and BLAKE, R.W. (1994). Técnica Pecuaria en México. (Accepted).

TREJO-VALDIVIA, G.M.B. (1991) Proportional hazards models with heterogeneous treatment effects. Ph.D. Thesis. University of London.

VAUPEL, J.W., MANTON, L.G., and STALLARD, E. (1979) Demography, 16, 439.