Ankara Üniv Vet Fak Derg, **54**, 219-222, 2007

Kısa Bilimsel Çalışma / *Short Communication*

Survival analysis on calving interval and gestation length in Simmental x South Anatolian Red F_1XB_1 crossbred cows

İsmail Safa GÜRCAN[1], Aytaç AKÇAY[1]

[1]Ankara Üniversitesi Veteriner Fakültesi Biyoistatistik Anabilim Dalı, Dışkapı, Ankara.

Summary: The objective of this study was to compare the survival analysis for calving interval and gestation length in Simmental x South Anatolian Red crossbred F_1XB_1 cows. The data set consisted of 1775 lactation records from January 1987 to October 1991, of which 3.10% were censored. The mean calving interval was 382.26 days (15.39 S.D) for young cow group while it was 386.27 days (8.90 S.D) for old cow group. The mean calving interval in young cow group is shorter than old cow group ($p<0.01$). The mean gestation length was 276.04 days (10.55 S.D) for young cow group while it was 278.89 days (9.63 S.D) for old cow group. The mean gestation length in young cow group was 2 days shorter than old cow group ($p<0.001$). The distribution of longevity data is extremely skewed. The methods based on assumptions of normality may give biased parameter estimates. Survival analysis to estimate breeding values for longevity should be able to cope with censored data and time dependent variables.

Key words: Calving interval, censored data, gestation length, survival analysis,

Simental x Güneydoğu Anadolu Kırmızısı F_1XB_1 melezi ineklerin buzağılama aralığı ve gebelik süresinin yaşam analizi ile incelenmesi

Özet: Bu çalışmada, Simmental x Güneydoğu Anadolu F_1XB_1 melezi genç ve ergin çağdaki ineklerde, buzağılama aralığı ve gebelik süreleri, yaşam analizi metodu ile incelenmiştir. Çalışmanın verileri, Ceylanpınar tarım işletmesinde Ocak 1987 ile Ekim 1999 tarihleri arasında yer alan 1775 laktasyon kayıdından elde edilmiştir. Verilerin %3.1'i sensörlüdür. Genç ineklerin ortalama buzağılama aralığının, (382.26 ± 15.39 Gün) ergin yaştakilerden (386.27 ± 8.9 Gün) daha kısa olduğu ($p<0.01$), ergin yaştakilerin gebelik süresinin (278.89 ± 9.63 S) gençlerden (276.04 ±10.55 S) 2 gün daha fazla olduğu görülmüştür ($p<0.001$). Yaşam verilerine ait damızlık değerlerinin tahmin edildiği çalışmalarda normallik varsayımı üzerine kurulu metotlar yanlı tahminler vermektedir. Zamana bağımlı ve sansürlü verilerin analizinde yaşam analizi metodu kullanılmalıdır.

Anahtar sözcükler: Buzağılama aralığı, gebelik süresi, sansürlü veri, yaşam analizi.

Variables used to denote the fertility of a dairy cow are calving interval (CI), days open, gestation length (GL), and number of inseminations per pregnancy. CI defined as the number of days between the two consecutive calvings. The CI is the sum of two components, the interval from the last calving date to the date of conception (a) and the length of gestation (b). Thus CI= a+b.

Gestation length appears to have an indirect influence on calving problems. As GL increases, the birth weights of the calves increase as much as half a kg per day. GL is a trait that can be selected for. This means that a potential exists to select cattle for shorter length and subsequently lighter birth weight. The consequence of decrease in fertility include additional insemination and increased veterinary costs (1).

Famula (9) was the first researcher who proposed survival analysis as a method to analyze length of productive life in dairy cattle. Smith (14) and Smith and Quaas (15) used survival analysis techniques to estimate breeding values of sires based on the length of productive life of their daughters. The techniques were further developed and adjusted for large scale applications by Ducrocq (7). By now, routine genetic evaluation of sires has been implemented in several European countries.

A number of economically important traits measure the time until an event occurs. These traits pose a number of challenges including non-number distributions and censoring. A specific feature of survival analysis is that it can accommodate censoring and by this way the use of information available on animals that are still alive.

A censored observation is defined as an observation with incomplete information. When an observation is censored it means that the information is incomplete because the subject did not have an event during the time that the subject was part of the study(1).

Miller (11), described the statistical techniques pertinent to the analysis of survival data with censoring and discussed the theoretical principles underlying those methods. In survival analysis, it is important to compare the survival experiences among two or more groups. Gehan (10), Breslow (2) and Cox (5) proposed some statistical comparisons based on mortality from any cause of death, whether or not related to the disease under study. Brown (3) carried out another statistical procedure for comparing the survival of two or more groups of patients adjusted for calculating relative survival. Buckley (4) discussed additive and multiplicative models for relative survival rates.

The objective of this study was to compare survival probabilities of calving interval and gestation length of young and mature Simmental x South Anatolian Red crossbred cows.

Lactation data of herdbook registered cows raised at Ceylanpınar State Farm from January 1987 to October 1991 were used. The data set consisted of 1775 lactation records were obtained from 138 Simmental x South Anatolian Red F_1XB_1 crossbred cows. The cows gave their first calves younger than 630 days were discarded from the analyses. Records of cows still present in the herd were right censored. The cows, at the first and the second lactation, were classified as young cows while the cows at the third and higher lactations were classified as mature cows.

The goal of survival analysis is to analyze positive measures describing the "width" of the interval between an origin point and an end point. The end point, called failure, corresponded to death or culling of the animal. But the end point may also correspond to the occurrence of any type of event, e.g, recovery from a disease (8).

There are three basic goals of survival analysis: to estimate and interpret survivor and hazard function from survival data; to compare survivor and hazard functions; to assess the relationship of explanatory variables to survival time.

The survivor function, S(t), gives the probability that an animal survives longer than some specified time t. The hazard function, h(t), gives the instantaneous potential per unit of time for the event to occur, given that the animal has survived up to time t.

The survivor and hazard functions can be mathematically expressed, respectively, as

$$S(t) = \Pr(T > t)$$
$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr\left(t \le T < t + \Delta t \middle| T \ge t\right)}{\Delta t} \qquad [1]$$

where T is the survival time variable. To estimate of the survival probabilities of animals under study, Kaplan – Meier method is a commonly used method in survival analysis (11).

$$\hat{S}_{KM}(t) = \prod_{k|T_{[k]}<t} \left(\frac{n_k - d_k}{n_k} \right) \qquad [2]$$

where $\hat{S}_{KM}(t)$ is the value of the survivor function at time t, $T_{[k]}$ represents the ordered failure times, from the first occurrence of the failure to the last one, n_k is the number of animals at risk at $T_{[k]}$, and d_k is the number of animals that actually died at. $T_{[k]}$. $\hat{S}_{KM}(t)$ is called the product limit or Kaplan – Meier estimate of the survivor function.

The survival analysis was performed using SPSS version 11.5 (16).

A total 1775 lactation records were analyzed, of which 3.10% were censored. The mean CI was 382.26 days (15.39 S.D) for young cow group while it was 386.27 days (8.90 S.D) for old cow group. The calving interval of young cow group is one day shorter than old cow group ($p<0.01$). The mean GL for young cow group was 276.04 days (10.55 S.D) while it was 278.89 days (9.63 S.D) for old cow group. However it was slightly skewed to the right. The mean GL of young cow group was 2 days shorter than old cow group ($p<0.001$).

The results according to The Kaplan – Meier analysis and the survival functions are shown in Table1, Figure 1 and Figure 2.

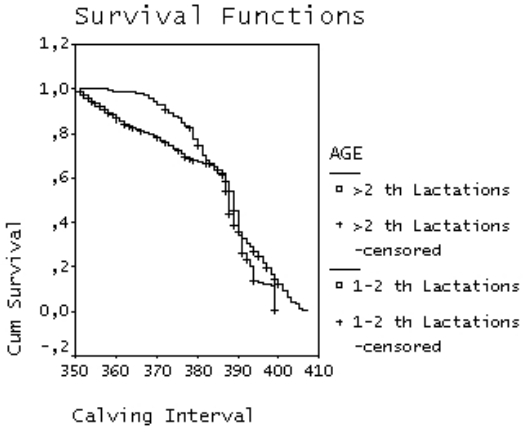

Figure 1. Plot of Calving Interval Survivor functions of the cows

Şekil 1. İneklerde buzağılama aralığının yaşam fonksiyon grafiği

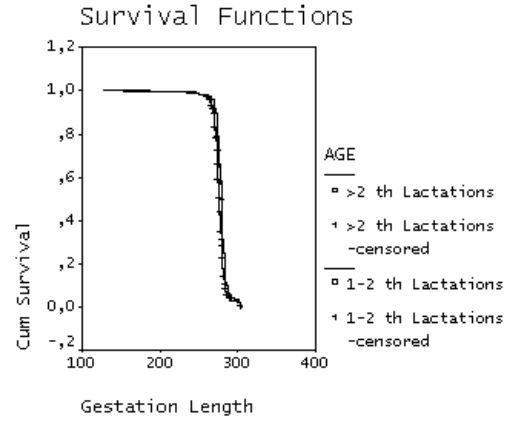

Figure 2. Plot of Gestation Lengths Survivor functions of the cows

Şekil 2. İneklerde gebelik süresinin yaşam fonksiyon grafiği

Table 1. CI and GL survivor functions of the cows
Tablo 1. İneklerde buzağılama aralığı ve gebelik süresi yaşam fonksiyonları

			Calving interval		Gestation length	
	n	Number of censored %	Median	St.Error	Median	St.Error
1-2 th. Lactation	1285	43 (3.35%)	388	0.16	277	0.19
> 2 th. Lactation	490	12 (2.45%)	389	0.30	279	0.23
Total	1775	55 (3.10%)	Log Rank=7.75 df=1 p<0.01		Log Rank=50.89 df=1 p<0.001	

Data on gestation length could be used to exploit the relationship between gestation length and birth weight and more importantly the relationship with calving difficulty. One might speculate that selection for shorter gestation lengths might reduce calving problems with less negative impact on growth rate than might be expected, for example, from selection for lower birth weights.

A reseach was made by Kumlu et al.(12) on milk yield and reproductive traits of Holstein Fresian Breeding herds in Turkey, the mean calving interval at 1-2th lactations were 406 ± 0.9 days standard error of mean (sem) while it was 398 ± 1.0 days sem at 2-3th lactations and 394 ± 1.4 days. All of the parameters estimations were calculated using GLM methods. Another research was made for investigating fertility and milk yield of Simmental cows, the general linear model was used (6). The mean gestation length was 282.5 ± 0.63 days (sem) for young cow group while it was 283.5 ± 0.66 days (sem) at calving age 3 and it was 283.9 ± 0.66 days (sem) at calving age 4. It was indicated that young group gestation length significantly shorter than other age groups. The mean calving interval was 381.4 ± 3.62 days (sem) for young cow group while it was 369.4 ± 3.90 days (sem) at calving age 3 and it was 375.8 ± 4.20 days (sem) at calving age 4. Gestation length results show parallel results with our findings. However, there were no significantly different between age groups on Calving interval parameter estimation was reported in the research whereas significantly different at age groups at our findings.

Evaluation of traits, which are measured in days, months, or years, poses a number of challenges. These traits consist of the length of time between two events. For example, a breeder may be interested in the length of productive life. The trait would than be the length of time an animal is productive. The breeder is then faced with the following challenges. First, the endpoints of the interval must be defined. Second, how will a record be treated if the animal leaves the herd for a factor unrelated to production? Third, how will a record be treated if the animal is still productive when the evaluation takes place? Forth, the distribution is heavily skewed. Survival analysis is an approach to analyzing traits such as these.

Health traits are important parts of total economic merit in dairy cattle and are expected to be an important element in decision on culling of individuals cows. It is therefore expected that there are strong relations between health and longevity.

Longevity is one of the most important components of dairy cow profitability. Calving interval and gestation length are frequently used values for the measurement of longevity. The distribution of longevity data is extremely skewed. The methods based on assumptions of normality may give biased parameter estimates. Methodology to estimate breeding values for longevity should be able to cope with censored data and time dependent variables.

Because animal health and welfare, sustainable breeding and more balanced selection objectives are progressively playing a more central role in animal breeding, there is definite need for better statistical tools to access the genetic components of fitness in domestic animals. Therefore, there is no doubt that survival analysis will take a more prominent place among animal breeders, not only cattle breeding but also in other species.

It is concluded that, survival analysis appears to be an appropriate method to handle this type of data in dairy cow farming.

In this search, The calving interval of young cow group was estimated one day shorter than old cow group and the mean GL of young cow group was 2 days shorter than old cow group (p<0.001).

References

1. **Boichard D** (1990): *Estimation of the economic value of conception rate in dairy cattle*. Livestock Prod Sci, **24**, 187 – 204.
2. **Breslow NE** (1970): *A generalized Kruskal – Wallis test for comparing k samples subject to unequal patterns of censorship.* Biometrika, **57**, 579 – 594.
3. **Brown CC** (1983): *The statistical comparison of relative survival rates.* Biometrics, **39**, 941 -948.
4. **Buckley JD** (1984): *Additive and multiplicative models for relative survival rates.* Biometrics, **40**, 51 – 62.
5. **Cox DR** (1975): *Partial likelihood*. Biometrika, **62**, 269 – 276.
6. **Çilek S, Tekin ME** (2005): *Environmental factors affecting milk yield and fertility traits of Simmental cows*

raised at the Kozova state farm and phenotypic correlations between these traits. Turk J Vet Anim Sci, **29**, 987-993.

7. **Ducrocq V** (1997): *Survival analysis, a statistical tool for longevity data.* 48 th. Annual Meeting of the EAAP, Vien, Austria.
8. **Ducrocq V** (1994): *Statistical analysis of length of productive life for dairy cows of the Normande breed.* J Dairy Sci, **77**, 855-866.
9. **Famula TR** (1981): *Exponential stayability model with censoring and covariates.* J Dairy Sci, **64**, 538.
10. **Gehan EA** (1965): *A generalized Wilcoxon test for comparing arbitrarily singly censored samples.* Biometrika, **52**, 203 – 223.
11. **Kaplan EL, Meier P** (1958): *Nonparametric estimation from incomplete observations*. J Amer Stat Assoc, **53**, 457 – 481.
12. **Kumlu S, Akman N** (1999): *Türkiye damızlık siyah alaca sürülerinde süt ve döl verimi.* Lalahan Hay Araşt Enst Derg, **39**, 1-15.
13. **Miller RG** (1981): *Survial Analysis*. John Wiley and Sons, NY.
14. **Smith SP** (1983): *The Extension of Failure Time Analysis to Problems of Animal Breeding.* PhD Diss, Cornell Univ, Ithaca, NY.
15. **Smith SP, Quass RL** (1984): *Productive lifespan of bull progeny groups: failure time analysis.* J Dairy Sci, **67**, 2999-3007.
16. **SPSS ver. 11.5**, Statistical Packet for Social Sciences, SN: 902 4147, Chicago, IL.

Geliş tarihi: 01.03.2006 / Kabul tarihi: 21.12.2006

Yazışma adresi:
Arş. Gör. Dr. İ.Safa Gürcan
Ankara Üniversitesi
Veteriner Fakültesi
Biyoistatistik Anabilim Dalı
Dışkapı/Ankara/Türkiye
e-mail:sgurcan@veterinary.ankara.edu.tr