H36.6 (ihbcab@bahnhof.se)

Survival analysis of the length of competition life of Standardbred trotters in Sweden

T.Árnason[*], *International Horse Breeding Consultant AB, Knubbo, S-74494 Morgongava, Sweden*

Introduction

The breeding goal for the Standardbred trotter in Sweden is a fast, sound, sustainable, endurable, well tempered, talented and good gaited racing trotter of international standard, capable both as an young and as an older trotter. Current genetic evaluations involve racing performance traits based on the racing results accomplished as 3- to 5-year-old. The heritabilty of the racing traits have been shown to drop with increasing age and therefore the genetic evaluations for the traits of early performance have mainly been used as criteria for selection (Árnason, *et al.*, 1989). The genetic evaluations are annually computed by MT-AM-BLUP for the following six traits: Number of races, Racing status (0 or 1), Percentage placed 1-3, Earnings/races, Earnings and Best racing time (Árnason, 1999). It may be argued that all these traits are to some extent expressing soundness and endurance of the horse, which is the prerequisite for the horse's success in races. However, none of these variables are directly reflecting the true endurance or the longevity as a racing horse (stayability). Number of races in the age interval of 3 to 5 years is a trait that is meant to indicate to a certain degree the ability of the horse to withstand many races. The problem is that both superior and inferior racing trotters are expected to have fewer races than an average performer. The outstanding racing horses usually do not start very often each year, but on the other hand they start in races where they meet tough competition and compete for big prize money. Presumably a better measure of stayability of the racing trotter would be to count the number of years which the horse participates in races. However, since the trait "number of years in races" is right censored (many horses have still not completed their racing career) it is not a suitable variable for analysis by the traditional mixed linear model methodology. If the genetic evaluations of this trait would be computed by BLUP then stallions with many young progenies in the beginning of their racing career would be underestimated compared with stallions with many progenies which have completed their racing career. Ducrocq (1987) showed that survival analysis could be effective for analysis of length of productive life of livestock due to their ability to handle right censored production records and to account properly for the non-linear characteristics of such data.

In Sweden the trotters are allowed to race up to the age of twelve years. Within this restriction stallions and geldings generally compute as long they are competitive and sound, while good mares may in some cases go to breeding earlier. Ricard and Fournet-Hanocq (1997) used the number of years in competition for analysing the stayability in French jumping horses. They used survival analysis which were based on Cox's proportional hazard model. The objective of this study was to analyse the length of competition life of Standardbred trotters in Sweden with similar methodology and to investigate the use of survival analysis as a tool for genetic evaluations for stayability in trotters.

Material and methods

The material used in this study was extracted from the file which was used for the routine genetic evaluations of Swedish standardbred trotters in December 2005. Only racing records on horses born in Sweden 1976-2001 were utilized and only progenies of sires with at least 20 offspring with racing records were considered in the analysis. Data set I involved records on 63,967 offspring of both sexes sired by 514 stallions. Data set II contained only records on stallions and geldings in total 32,799 offspring of 393 sires.

Since the time observations of years in competitions are expressed on a discrete scale (min=1; max=10) the traditional Cox or Weibull models are not suitable for the survival analysis. Instead the grouped data model of Prentice and Gloeckler (1978) had to be used (Ducrocq, 1999). The proportional hazard model can then be expressed as:

$$\lambda(t;\, \mathbf{x}(t), \mathbf{z}(\mathrm{t})) = \lambda^{*}_{0}(t)\ \exp\ \{\mathbf{x}(t)'\mathbf{b} + \mathbf{z}(t)'\mathbf{s}\}$$

where, $\lambda(t;\, \mathbf{x}(t), \mathbf{z}(\mathrm{t}))$ denotes the hazard function at time t (instantaneous rate of failure at time t, conditional upon survival up to t), $\lambda^{*}_{0}(t)$ is the piecewise constant baseline hazard as estimated at every discrete failure time interval (Prentice and Gloeckler, 1978), $\mathbf{x}(t)$ is an incidence vector of fixed explanatory variables with a corresponding **b** vector of regression coefficients, $\mathbf{z}(t)$ is the incidence vector for the random sire effects of the model and **s** is the corresponding vector of sire effects.

The Survival Kit V3.1 programs were used perform the analyses (Ducrocq and Sölkner, 1999). Several models were employed where the fixed effects **b** were as follows:

Data set I:
b includes sex/birth-year subclasses (52) and the (time dependent) discrete time units (10 classes).
b includes sex/birth-year subclasses (52), a continuous covariate of inbreeding coefficient (F%) and the discrete time units (10).
b includes sex/birth-year subclasses (52), a continuous covariate of inbreeding coefficient (F%), a continuous covariate of phenotypic racing performance index and the discrete time units (10).
b includes birth-year subclasses (52), a continuous covariate of inbreeding coefficient (F%), a continuous covariate of BLUP total-index and the discrete time units (10).

Data set II:
b includes birth-year subclasses (26) and the discrete time units (10 classes).
b includes birth-year subclasses (26), a continuous covariate of inbreeding coefficient (F%) and the discrete time units (10).
b includes birth-year subclasses (26), a continuous covariate of inbreeding coefficient (F%), a continuous covariate of phenotypic racing performance index and the discrete time units (10).
b includes birth-year subclasses (26), a continuous covariate of inbreeding coefficient (F%), a continuous covariate of BLUP total-index and the discrete time units (10).

The phenotypic racing performance index was based on racing results as 3-5-year-olds. The racing variables were expressed as standardized deviations from corresponding sex/birth-year subclasses. The sire variance components were estimated as the mean of the marginal posterior distribution of the normal distribution using Gauss-Hermite integration with five points in each iteration (Ducrocq and Cacella, 1996, Ducrocq, 1999). The estimates of the sire variance σ^2_s were used to compute the "equivalent" heritability as defined by Yazdi *et al.* (2002) as:

$$h^2_{equ} = 4\sigma^2_s / (\sigma^2_s + 1/p)$$

where p is the proportion of uncensored records.

The estimated hazard of each sire i was transformed into genetic standard deviation unit as: $g_i = -s_i / \sigma_s$ which was then converted to an index value as: Stayability Index$_i$ $= 100 + 10(g_i - \mu_g)/\sigma_g$. The accuracy of the index was calculated as $r_{TI} = \sqrt{(n/(n + k))}$, where n is the number of uncensored progeny observations and $k = (4 - h^2_{equ})/ h^2_{equ}$ (Ducrocq, 1999).

Pearson's correlation coefficients and Spearman's rank-order correlation coefficients were calculated between the index values obtained in the different analyses and between the BLUP index and the survival analysis stayability index.

Results

In data set I (both sexes) the proportion of uncensored records was p=0.87 and in data set II p=0.85. Table 1 shows the estimated sire variance components and the corresponding "equivalent" heritability from the different analyses.

Table 1. Sire variance components σ_s and "equivalent" heritabilty h^2_{equ}.

Analysis (run)	σ_s	h^2_{equ}
a) Both sexes; sex/birth-year	0.0138	0.047
b) a) + F%	0.0143	0.049
c) b) + phenotypic performance	0.0151	0.052
d) b) + BLUP	0.0488	0.163
e) Males; year	0.0210	0.070
f) e) + F%	0.0216	0.072
g) f) + phenotypic performance	0.0122	0.041
h) f) + BLUP	0.0350	0.116

Analyses a), b), e), and f) are survival analyses which are not adjusted for performance, while analyses c) and g) are adjusted for early phenotypic performance (as 3-5-year-old), and analyses d) and h) are adjusted for estimated breeding values for the early performance traits (BLUP based on 3-5-year-old records). For the unadjusted models the results indicate slightly higher estimates of the sire variance component when only males were included in the analysis. For the performance adjusted analysis the results are the opposite. The estimated sire variance components are slightly higher when both sexes were included in the analysis.

Inbreeding had significant effect on the hazard function, where the hazard increased with higher inbreeding coefficient of the offspring. Inclusion of F in the model had only marginal effect on the estimated sire variance, though. In general the low heritability estimates are rather disencouraging. The survivor function S(t) and the conditional survivor function S(t|T>t(k-1)) from run f) are shown in Figure 1. The average length of racing career calculated from uncensored records were 3,4 years (3,8 for males only).

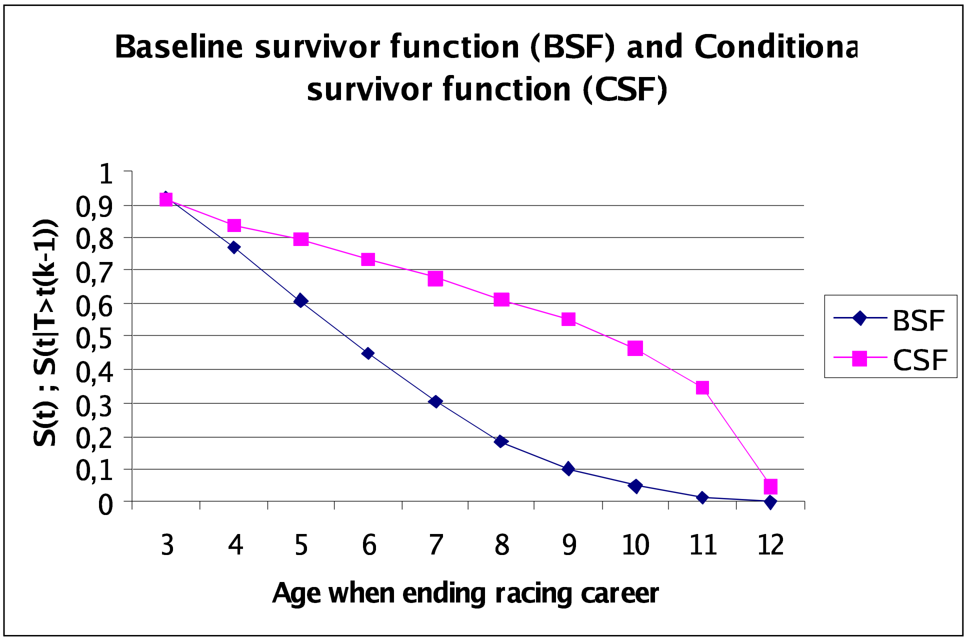

Figure 1

The correlations between the stayability indices (survival index) resulting from the different runs (data, model) are shown in Table 2. The correlations were computed between paired solutions for sires which occur in both data sets.

Table 2. Person's correlations between stayability indices from different runs above the diagonal, Spearman's rank correlations below the diagonal.

Run	a	b	c	d	e	f	g	h
a	-	0.99	0.77	0.39	0.83	0.83	0.70	0.30
b	0.99	-	0.80	0.42	0.81	0.81	0.72	0.33
c	0.74	0.76	-	0.79	0.48	0.48	0.77	0.66
d	0.37	0.39	0.78	-	0.09	0.09	0.50	0.75
e	0.81	0.79	0.45	0.09	-	1.00	0.77	0.31
f	0.81	0.79	0.45	0.09	1.00	-	0.77	0.30
g	0.68	0.69	0.72	0.49	0.76	0.76	-	0.68
h	0.29	0.30	0.62	0.73	0.32	0.31	0.68	-

Summarising Table 2 the correlation between the unadjusted stayability index and the phenotypic racing performance (production) adjusted stayability index was estimated to *r*=0.8. The corresponding correlation between the unadjusted stayability index and the BLUP adjusted stayability index was estimated to *r*=0.4. The correlation between the adjusted stayability indices was *r*=0.7.

The correlations between the stayability indices and the BLUP indices for performance traits as 3-5-year-old are shown in Table 3.

Table 3. Correlations (Pearson's) between stayability indices and BLUP indices for racing performance traits

BLUP I	Run :	b	c	d	f	g	h
Total index		0.35	-0.06	-0.47	0.46	0.06	-0.38
Number of starts		0.50	0.10	-0.29	0.57	0.18	-0.25
Racing performance		0.34	-0.08	-0.48	0.45	0.04	-0.38
Start status		0.35	-0.02	-0.41	0.44	0.08	-0.33

In data set I the average number of uncensored progeny records *per* sire were 128.5 and the average accuracy of selection were r_{TT} =0,72 in run b). In data set II (only males) the average number of uncensored progeny records *per* sire were 70,9 and the corresponding mean accuracy were r_{TT} =0,65 in run f). The increased estimated hazard of mares compared to stallions and geldings (males) are shown in Figure 2.

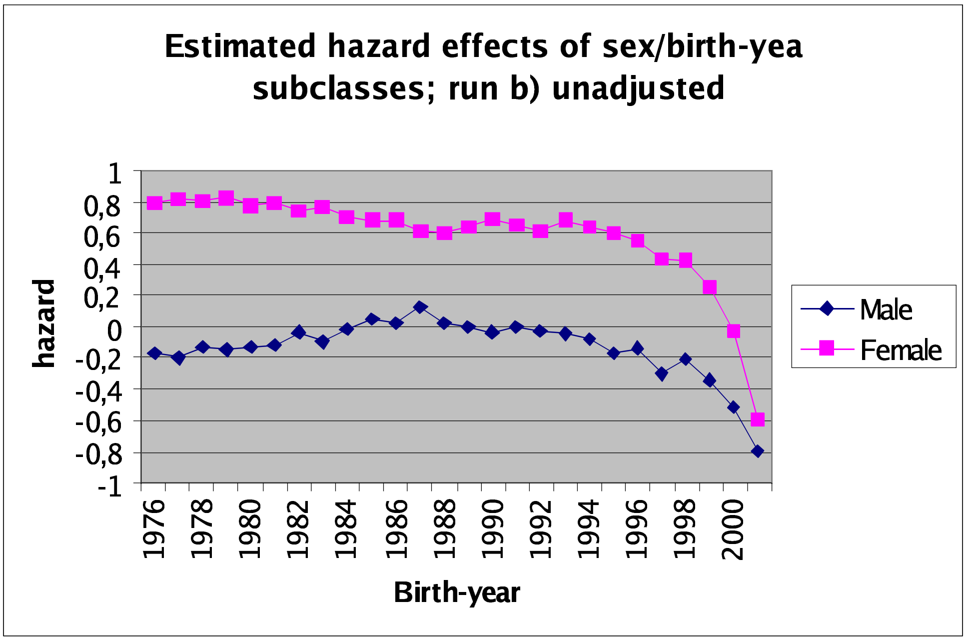

Figure 2

Discussion

Survival analysis seems to be the right tool to compute stayability index in right censored data on the length of competitive life of racing horses measured as their number of years in racing. Unfortunately the length of the racing career in standardbred trotters seems to be very much dependent on environmental factors as indicated by the relatively low heritability estimates. However the heritability estimates of the survivor variables are significantly higher than zero and certain response to genetic selection should be expected. Due to the low heritability the stayability index will be most valuable for evaluating stallions with relatively many raced offspring.

Inbreeding had significantly negative effects on the length of competitive life of the racing trotters. Stallions and geldings have somewhat longer competitive life than mares and the estimated hazard of mares is considerable higher than those of males. Obviously some good racing mares quit racing and become covered for breeding and this may introduce slight bias. However, the increased number of records which become available by the use of the data on mares suggest that records on both sexes should be used for the genetic evaluations of sires for stayability due to the increase in accuracy and in the number of sires with reliable index values. In the remaining part of the discussion only results on data set I (both sexes, including inbreeding; runs: b-d) will be considered.

The unadjusted stayability index of the 506 stallions is positively correlated with the genetic evaluations (BLUP) of the stallions for early racing performance. This is an important finding which shows that there is a general medium high positive correlation between racing capacity and the length of the competitive life. The stayability index adjusted for phenotypic early racing performance is almost uncorrelated with the sires' BLUP indices of early racing performance. As to be expected the sires' stayability index adjusted for the progenies BLUP indices is negatively correlated with their BLUP index.

Which stayability indices should be used in practice by the breeders? Those which are adjusted for racing ability or not? My preliminary answer would be that all should be used simultaneously together with the BLUP index. However, there might be a difficult pedagogic problem to get breeders to interpret the adjusted indices, especially the one adjusted for the genetic racing ability (BLUP). The length of the competition life is obviously dependent both on the horse's racing performance ability and the stayability in form of soundness and durability of the horse. The "time of failure" for any horse is either depending on lack of racing capacity compared with the contemporaries or it suffers from injuries or disease (which also consequently suppresses their observed racing capacity). Any adjustment of the survivor function for racing performance will automatically benefit horses with less racing capacity. They will not strain their locomotion and respiratory organs to the same extent as good racing horses and may stay healthier for longer time. Therefore the different indices should be read simultaneously. Stallions with high BLUP for racing performance together with high stayability indices are likely to give offspring with slightly better lastingness in racing. On the other hand, stallions with high stayability indices are not interesting as candidates for selection if they lack genetic ability for racing performance. Perhaps the stayability index adjusted for the progenies' BLUP indices favours too much the inferior racing trotters, which are not exerting to the utmost in races, but lasting relatively long. Therefore this index will certainly be questioned by breeders.

The results of this study are preliminary and the material will be studied further with newer software and possibly more complete models (Damgaard, 2006) before this procedure will be applied for genetic evaluations of Swedish trotter stallions for soundness and stayability.

Conclusions

Survival analysis are promising tools for genetic analysis and evaluation of breeding values for length of competitive life of racing horses. The low genetic variance obtained for the survival variables studied were rather disappointing (heritability about 5%). Inbreeding had negative effect on the length of competive life of the standardbred trotters in Sweden. Stallions and geldings have longer competition life than mares. Estimated breeding values for racing performance and unadjusted stayability index and stayability index adjusted for racing performance should be read and interpreted simultaneously.

References

Árnason, T., 1999. Genetic evaluation of Swedish standard-bred trotters for racing performance and racing status. *J. Anim. Breed. Genet. 116:387-398.*
Árnason, T., Bendroth, M., Philipsson, J., Henriksson, K. and Darenius, A., 1989. Genetic evaluations of Swedish trotters. In *State of Breeding Evaluation in Trotters.* EAAP Publication No. 42. Pudoc, Wageningen, pp. 106-130.
Damgaard, L.H., 2006. Joint quantitative genetic analysis of survival, linear Gaussian, and ordered categorical traits. *8th World Congress on Genetics Applied to Livestock Production, August 13-18, 2006*, Belo Horizonte, MG, Brazil.
Ducrocq, V., 1987. An analysis of length of productive life in dairy cattle. Ph.D. Dissertation. Cornell University, Ithaca, NY, USA.
Ducrocq, V., 1999. Survival Analysis Applied to Animal Breeding and Epidemiology. INRA, Jouy-en-Josas, France.
Ducrocq, V. and Cacella, G., 1996. A Bayesian analysis of mixed survival models. *Gen. Sel. Evol. 28:505-529.*
Ducroq, V. and Sölkner, J., 1999. The Survival Kit V3.1 User's Manual. INRA, Jouy-en-Josas, France.
Prentice, R. and Gloeckler, L., 1978. Regression analysis of grouped survival data with application on breast cancer data. *Biometrics 34:57-67.*
Ricard, A. and Fournet-Hanocq, F., 1997. Analysis of factors affecting length of competition life of jumping horses. *Genet. Sel. Evol. 29:251-267.*
Yazdi, M.H., Visscher, P.M., Ducrocq, V. and Thompson, R., 2002. Heritability, reliability of genetic evaluatons and response to selection in proportional hazard models. *J. Dairy Sci. 85:1563-1577.*