

Discussion of heritability of survival traits

Korsgaard¹, I.R., Andersen², A.H., and Jensen¹, J.

¹Department of Animal Breeding and Genetics, Danish Institute of Agricultural Sciences,
DK-8830 Tjele

²Department of Theoretical Statistics, University of Aarhus, DK-8000 Aarhus-C

Abstract

Different definitions of heritability of survival traits are available in the literature - on different scales and in different models. In the Weibull frailty model, heritability has been defined on the log time scale and transformed to the original time scale using Taylor series expansions. In the semiparametric Cox frailty model, heritability has been defined on the log frailty scale. In this paper, having presented the result that heritabilities on any (linear or approximated linear) time scales are identical, these measures of heritability are discussed. In the semiparametric Cox frailty model, there is a direct link between heritability defined on frailty scale and heritability on time scale. In the Weibull frailty model, the heritability that has been reported on the original time scale appears dubious.

1. Introduction

In the Weibull frailty model, heritability has been defined on the log time scale (e.g. Ducrocq and Casella, 1996) and transformed to the original scale (Ducrocq (1997) personal communication, cited from Vollema (1998)) using Taylor series expansions. In the semiparametric Cox frailty

model, heritability has been defined on the log frailty scale (Korsgaard et al., 1996, 1998). The purpose of this paper is to discuss the relevance of, and links between, these definitions. This is carried out in frailty models without stratification and time-dependent covariates.

2. The model

Let T_i be the random variable representing survival time of animal i . In the frailty model, it is

assumed that conditional on log frailty, $Z_i = z_i$, (or frailty) the hazard function of animal i , is given by

$$\lambda_1(t|z_i) = \lambda_0(t) \exp\{\mathbf{x}_i' \beta + z_i\} \quad (1)$$

where $\lambda_0(t)$ is a common baseline hazard function, \mathbf{x}_i is a vector of time-independent covariates of animal i , and β is the corresponding vector of

regression parameters. In the model specified by (1), the conditional integrated hazard function is

$$\Lambda_i(t|z_i) = \Lambda_0(t) \exp\{\mathbf{x}_i' \beta + z_i\} \quad (2)$$

and the conditional survival function is

$$S_i(t|z_i) = \exp\{-\Lambda_i(t|z_i)\}$$

Let F_i denote the distribution function of T_i in the conditional distribution given $Z_i = z_i$. Because $F_i(T_i|z_i) = 1 - \exp\{-\Lambda_i(T_i|z_i)\}$ is uniformly distributed on the interval $(0;1)$, the transformed random

variable, $Y_i = \Lambda_i(T_i|Z_i)$, conditional on $Z_i = z_i$, is exponentially distributed with parameter 1. In turn, ε_i , the logarithm of Y_i , given by

$$\varepsilon_i = \log(Y_i) = \log(\Lambda_i(T_i|Z_i)) = \log(\Lambda_0(T_i)) + \mathbf{x}_i' \beta + Z_i \quad (3)$$

conditional on $Z_i = z_i$, follows an extreme value distribution. By rearranging terms in (3) it follows,

that the model in (3), is equivalent to a linear model on the $\log(\Lambda_0(\cdot))$ scale:

$$\log(\Lambda_0(T_i)) = -\mathbf{x}_i' \beta - Z_i + \varepsilon_i \quad (4)$$

The unconditional mean and variance of $\log(\Lambda_0(T_i))$ are $-\mathbf{x}_i' \beta - E(Z_i) - \gamma_E$ and $\text{Var}(Z_i) + [\pi^2/6]$,

respectively, where γ_E is Euler's constant.

2.1 Sire model and heritability

Assuming that log frailty

$Z_i = s_j + \sum_{k=1}^K w_{k,k(i)}$, where s_j is the sire of animal i and $k(i)$ is the level of k 'th random effect that animal i has. Furthermore, it is assumed that the vector of sire effects, \mathbf{s} , is independent of the

other random effects, $(\mathbf{w}_1, \dots, \mathbf{w}_K)$. It follows that a natural definition of heritability on the $\log(\Lambda_0(\cdot))$ scale is

$$h_{s, \log(\Lambda_0(T_i))}^2 = \frac{4\sigma_s^2}{\text{Var}(\log(\Lambda_0(T_i)))} = \frac{4\sigma_s^2}{\sigma_s^2 + \text{Var}(\sum_{k=1}^K w_{k,k(i)}) + (\pi^2/6)} \quad (5)$$

2.2 Animal model and heritability

In the animal model, it is assumed that $Z_i = \mathbf{a}_i + \sum_{k=1}^K w_{k,k(i)}$. Assuming, that the vector of animal effects, \mathbf{a} , is independent of $(\mathbf{w}_1, \dots, \mathbf{w}_K)$, a

natural definition of heritability on the $\log(\Lambda_0(\cdot))$ scale is

$$h_{a, \log(\Lambda_0(T_i))}^2 = \frac{\sigma_a^2}{\text{Var}(\log(\Lambda_0(T_i)))} = \frac{\sigma_a^2}{\sigma_a^2 + \text{Var}(\sum_{k=1}^K w_{k,k(i)}) + (\pi^2/6)} \quad (6)$$

2.3 Heritability on any transformed time scale

Let g denote an increasing, one-one and differentiable function. Using a first order Taylor series expansion of $g(\log(\Lambda_0(T_i)))$ around the mean

of $\log(\Lambda_0(T_i))$, it follows that

$$\begin{aligned} g(\log(\Lambda_0(T_i))) &\approx g(-\mathbf{x}_i' \beta - E(Z_i) - \gamma_E) \\ &+ g'(-\mathbf{x}_i' \beta - E(Z_i) - \gamma_E) (\log(\Lambda_0(T_i)) + \mathbf{x}_i' \beta + E(Z_i) + \gamma_E) \\ &= g(-\mathbf{x}_i' \beta - E(Z_i) - \gamma_E) \\ &+ g'(-\mathbf{x}_i' \beta - E(Z_i) - \gamma_E) (-Z_i + \varepsilon_i + E(Z_i) + \gamma_E) \end{aligned} \quad (7)$$

which implies that heritability in the linear model (4) is identical to the heritability obtained from the approximated linear model (linear in random

effects) for any transformation, g , of (4). In particular, note that this holds when transforming to the original scale. The result is valid whether the

baseline hazard is parameterised (e.g., Weibull frailty model) or completely arbitrary (semiparametric frailty model). Furthermore it is also valid irrespective of distributional assumptions

about the random effects (e.g., random effects could be log gamma, normally distributed or a combination of the two).

3. Discussion

In this section, existing definitions of heritability are discussed in relation to the result in section 2.3.

3.1 The semiparametric Cox log normal frailty model

In (the simplest case) the semiparametric Cox log normal frailty model described by Korsgaard et

al. (1996, 1998), the hazard function for animal i , conditional on random effects, a_i and e_i , is given by

$$\lambda_i(t|a_i, e_i) = \lambda_0(t) \exp\{\mathbf{x}_i' \beta + a_i + e_i\} \quad (8)$$

where $\lambda_0(t)$ is an arbitrary baseline hazard function, the random effect, a_i , is an animal effect, and e_i is a random error term in log frailty.

It is assumed, that the frailty of animal i , $\exp(Z_i) = \exp\{a_i + e_i\}$, is log normally distributed, such that

$$(Z_1, \dots, Z_n) \sim N_n(\mathbf{0}, \mathbf{A}\sigma_a^2 + \mathbf{I}\sigma_e^2)$$

In this model, heritability was defined on the log frailty scale by

$$h_{a, Z_i}^2 = \frac{\sigma_a^2}{\text{Var}(Z_i)} \quad (9)$$

where $\text{Var}(Z_i) = \sigma_a^2 + \sigma_e^2$ is the total variance of log frailty.

From the results of section 2, we observe that heritability on the log ($\Lambda_0(\cdot)$) scale (or any

transformed time scale) can be obtained from the heritability on the log frailty scale by

$$h_{a, \log(\Lambda_0(T_i))}^2 = \frac{\sigma_a^2}{\text{Var}(Z_i) + \pi^2/6}$$

Consequently, the heritability on the linear or any approximating linear scale is always lower than the heritability defined on the log frailty scale. Furthermore, we observe that the sire model, where

the hazard function of animal i , conditional on random effects, s_j and \tilde{e}_i , is given by

$$\lambda_i(t|s_j, \tilde{e}_i) = \lambda_0(t) \exp\{\mathbf{x}_i' \beta + s_j + \tilde{e}_i\} \quad (10)$$

is equivalent to the animal model in (8), if only relationships among sires are taken into account and if $\sigma_a^2 = 4\sigma_s^2$ and $\sigma_e^2 = \sigma_e^2 + 3\sigma_s^2$.

In the sire model, (10), heritability on the log frailty scale is given by

$$h_{s, Z_i}^2 = \frac{4\sigma_s^2}{\text{Var}(Z_i)} \text{ and heritability on the log } (\Lambda_0(\cdot))$$

scale (or any transformed time scale) is obtained by

$$h_{s, \log(\Lambda_0(T_i))}^2 = \frac{4\sigma_s^2}{\text{Var}(Z_i) + \pi^2/6}.$$

The heritabilities on time scales are then identical in sire and animal models, similarly for heritabilities on the log frailty scale.

3.2 The Weibull log normal frailty model

In (the simplest case) the Weibull log normal frailty model described by Ducrocq and Casella

$$\lambda_i(t|s_j) = \gamma p(\gamma t)^{p-1} \exp\{\mathbf{x}_i' \boldsymbol{\beta} + s_j\} \quad (11)$$

where the baseline hazard function, $\gamma p(\gamma t)^{p-1}$, is that of a Weibull distribution. It is assumed that

$$\mathbf{s} \sim N_q(\mathbf{0}, \mathbf{A}\sigma_s^2)$$

where q is the number of sires. For this model, it was argued that a natural definition of heritability of log survival time is

$$h_{s, \log(T_i)}^2 = \frac{4\sigma_s^2}{\sigma_s^2 + (\pi^2/6)}$$

This definition is in agreement with the result

$$\lambda_i(t|a_i) = \gamma p(\gamma t)^{p-1} \exp\{\mathbf{x}_i' \boldsymbol{\beta} + a_i\} \quad (12)$$

where it is assumed that $\mathbf{a} \sim N_n(\mathbf{0}, \mathbf{A}\sigma_a^2)$, and where n is the number of animals with uncensored and censored records of the survival trait. Proceeding exactly as for the sire model, it can be argued that a natural definition of heritability of log survival time in the animal model is

$$h_{a, \log(T_i)}^2 = \frac{\sigma_a^2}{\sigma_a^2 + (\pi^2/6)}$$

In several papers, heritability is reported on the log and on the original scales for the Weibull log normal frailty model. The heritabilities reported on the original scale were obtained using the variance of a first order Taylor series expansion of $\exp(s_j)$ ($\exp(a_i)$) in the numerator (multiplied by 4 in the sire model). In the denominator was the variance of a first order Taylor series expansion of $\exp(s_j + \varepsilon_i)$ ($\exp(a_i + \varepsilon_i)$). This led to a reported heritability on the original scale given by

$$h_{s, T_i}^2 = \frac{h_{s, \log(T_i)}^2}{\left[\exp\left\{ -\frac{1}{p} \gamma E \right\} \right]^2} \text{ in the sire model, (11), and}$$

$$h_{a, T_i}^2 = \frac{h_{a, \log(T_i)}^2}{\left[\exp\left\{ -\frac{1}{p} \gamma E \right\} \right]^2} \text{ in the animal model, (12).}$$

According to the results of section 2, the heritabilities reported on the original scale are

(1996), the hazard function for animal i conditional on the random sire effect, s_j , is given by

of section 2 because in the Weibull frailty model, the $\log(\cdot)$ scale is a linear transformation of the $\log(\Lambda_0(\cdot))$ scale (i.e., $\log(\Lambda_0(T_i)) = p \log(T_i) + p \log(\gamma)$).

If the random sire effect, s_j , in (11) is substituted by an animal effect, a_i , we have the (simplest) animal model for survival data, described in Ducrocq and Casella (1996):

dubious.

4. Conclusion

In this paper, we have presented the result, that heritabilities on any time scale (linear or approximated linear) are identical. The result is valid for frailty models without stratification and time-dependent covariates. In Korsgaard et al. (1996, 1998) heritability was reported on an unobservable log frailty scale. In this paper, it is linked to heritability on a time scale. In the Weibull frailty model, heritability was reported on the log scale (Ducrocq and Casella, 1996) and transformed to the original scale (Ducrocq (1997) personal communication, cited from Vollema (1998)). The latter appears dubious.

5. References

- Ducrocq, V., Casella, G. (1996). A Bayesian analysis of mixed survival models. *Genet. Sel. Evol.* 28, 505-529.
- Korsgaard, I.R., Madsen, P., Jensen, J. (1996). Genetic analysis of disease resistance in bulls performance tested for beef traits - using an animal model for survival data. *Book of Abstracts of 47th Annual Meeting of the*

European Association for Animal Production.
(Abstract and Handout).

Korsgaard, I.R., Madsen, P., Jensen, J. (1998).
Bayesian inference in the semiparametric log
normal frailty model using Gibbs sampling.

Genet. Sel. Evol. 30, 241-256.

Vollema, A.R. (1998). Selection for longevity in
dairy cattle. Doctoral thesis, Wageningen
Agricultural University.