Multivariate Survival Mixed Models for Genetic Analysis of Longevity Traits

Rafael Pimentel Maia, Per Madsen and Rodrigo Labouriau [*]
Department of Molecular Biology and Genetics, Aarhus University

March 2013

Abstract

A class of multivariate mixed survival models for continuous and discrete time with a complex covariance structure is introduced in a context of quantitative genetic applications. The methods introduced can be used in many applications in quantitative genetics although the discussion presented concentrates on longevity studies. The framework presented allows to combine models based on continuous time with models based on discrete time in a joint analysis. The continuous time models are approximations of the frailty model in which the hazard function will be assumed to be piece-wise constant. The discrete time models used are multivariate variants of the discrete relative risk models. These models allow for regular parametric likelihood-based inference by exploring a coincidence of their likelihood functions and the likelihood functions of suitably defined multivariate generalized linear mixed models. The models include a dispersion parameter, which is essential for obtaining a decomposition of the variance of the trait of interest as a sum of parcels representing the additive genetic effects, environmental effects and unspecified sources of variability; as required in quantitative genetic applications. The methods presented are implemented in such a way that large and complex quantitative genetic data can be analyzed.

[*] Corresponding author. Email:rodrigo.labouriau@agrsci.dk

Contents

1 Introduction

Longevity is an important trait often considered in animal breeding programs [31, 30, 33, 37, 27, 35, 17, 9, 6, 32, 1]. Even small changes in the longevity of a population under production might have remarkable economic, welfare and ethics consequences [30, 12]. Since the study of longevity involves several types of incomplete observation (*e.g.* censoring, truncation, late entry and competing risks), survival and event-history-analysis techniques are typically used [8, 19, 3]. However, the use of those techniques in the context of quantitative genetics of longevity involves several non-trivial challenges. We will present a model framework here that allows to overcome these challenges. These models will have a structure of means and covariances similar to the gaussian linear mixed models classically used in quantitative genetics. They will allow for a proper representation of quantitative genetic phenomena and for efficient implementations required in practice. We will show that these models extend the class of models currently in use for studying the genetic of longevity.

The first challenge in the use of survival models for characterizing the genetic aspects of longevity is the high complexity of the models. Typically it is necessary to adjust simultaneously for the effects of several explanatory variables, some continuous with linear effects (*e.g.* breed composition and heterosis [21]), some factors with many levels (*e.g.* herd) or even time-dependent effects (*e.g.* year and season). Further complexity is added by the necessity of representing complex genetic effects (several generations deep pedigrees) under different genetic models (*e.g.* sire model, sire-dam-model etc). A typical scenario of applications in quantitative genetics involves a very large number of observations (individuals); indeed often the analyses involve several hundred thousand of observations (we will present two relative small illustrative examples involving 142,133 and 200,084 individuals). This leads to inference problems of the complexity equivalent to solve systems of linear equations with several hundred of thousand, or even several million, simultaneous linear equations (in our examples between 8,248 and 118,432 simultaneous linear equations) when using classic gaussian linear mixed models. Now, the complexity of survival and event-history models is even larger since these models require to keep track of the individuals' history. Moreover, as it will be apparent from the description of the methods available and from our discussion, some of those models might present very flat likelihood functions making the inference problem hard and numerically unstable. This rules out the naive use of the standard methods of survival analysis and makes the use of specially constructed models and approximations mandatory.

A second challenge is the proper representation of the genetic phenomena in play. The interpretation of the gaussian linear mixed models in genetic terms requires a linear decomposition of the so called total variance (*i.e.* the phenotypic variance after having corrected for the effects of a range of non-genetic related aspects with known effects on the trait of interest)[18, 34]. This decomposition of the total variance occurs naturally in the gaussian linear mixed models but not in all the adaptations of survival models specially constructed for genetic evaluations of longevity. We will show that a way to circumvent this problem is to consider models containing a dispersion parameter which plays the role of the residual variance (*i.e.* the part of the total variance that is neither attributed to genetic components nor to identifiable sources of variation) and base the inference on the quasi-likelihood theory [36, 4]. The use of generalised linear mixed models with a dispersion parameter occurs only occasionally in the literature of quantitative genetics [14, 5], and the genetics and the statistics consequences of using a free dispersion parameter has not being systematically explored yet. Moreover, survival models with dispersion parameter were, to the best of our knowledge, never considered in the literature of quantitative genetics.

A third challenge is the presence of incomplete observations following complex patterns. Variants of the so called Cox proportional model [8] were used for univariate traits describing longevity [9, 13, 10]. These techniques succeeded to implement models of approximations that were operational for some of the purposes of animal breeding. However, all those models were of univariate nature (*i.e.* they consider one trait at time) [30, 11], which makes it difficult, if not impossible, to address some key questions related to the issues of competing risks, informative censoring and inhomogeneity of the failure mechanism. A clear example is the study of time to death were the individuals on the study may die from one among several different causes of death. Moreover, the use of multivariate models is necessary to study the time death in a competing risk scenario, by evaluating the risk of death for each specific cause simultaneously. This allows to form the basis to study further aspects as the presence of common genetic determining factors and/or independent or specific genetic factors for each of the death causes. In the framework presented here, it will be possible to treat these issues by using suitable multivariate models.

Two types of models will be considered: models based on continuous time and models based on discrete time. The continuous time models (CTM) will be suitable approximations of the semi-parametric Cox proportional model in which the hazard function will be assumed to be piece-wise constant. These models allow for regular parametric likelihood-based inference by exploring a coincidence of their likelihood functions and the likelihood function of a Poisson model applied to a specially constructed data [2]. When including gaussian random components (log gaussian frailties) the inference can be based on a Laplace approximation to a high dimensional integral [4, 24, 28]. Here, we will include also a dispersion parameter that will play the role of the residual variance which is an essential element for the genetic interpretation of the models. The CTM will depend on an arbitrary choice of time cut points used to define the intervals at which the hazard function is assumed to be constant.

The discrete time models (DTM) used are multivariate variants of the classic discrete relative risk models [19]. These models present advantages in terms of computational and statistical complexity as compared to the CTM: the algorithms run faster relatively to the analogous algorithm for CTMs, are numerically more stable, and the statistical inference is more efficient. The loss of genetic information occurred when switching from the CTM to the DTM is minimal in typical applications in quantitative genetics and this loss is fairly compensated by the avoidance of capturing a large amount of noise. Here, we will also introduce the use of a dispersion parameter, which will be essential for well representing genetic scenarios.

The aim of this article is to introduce, characterize and discuss a class of multivariate mixed survival models for continuous and discrete time with a complex covariance structure in a context of quantitative genetics applications. The methods introduced here can be used in many other applications in quantitative genetics although the discussion presented concentrates on longevity.

The paper is organized as follows. Section 2 describes the basic set-up and the genetic scenario discussed. There, a suitable multivariate version of the proportional hazard model is introduced in general terms. Those models will encompass models for competing risks possibly defined with different types of time scale (continuous and discrete time). The techniques for statistical inference under those models are presented in section 3, including a general discussion on the calculations involved in the likelihood based inference and some connections with multivariate generalized linear mixed models (section

3.1). A Poisson approximation useful for efficiently implementing likelihood based inference for continuous time models is presented in section 3.2. Section 3.3 presents an extension of the models to a situation where there is a stratification variable. A discussion comparing the amount of statistical information between models constructed with discrete time and models based on continuous time is given in section 3.4. The quantitative genetic theory behind the models considered here requires a special decomposition of the phenotypic variance in terms of the variance of the random components and a scale parameter present in the models. This will be necessary for the calculation of the so called heritability, which is crucial in quantitative genetic applications. Section 4 will discuss methods for that and some technical details involving counting processes and martingale theory are presented in the appendix. Two illustrative examples involving longevity of sows and dairy cattle will be presented in section 5.1 and 5.2 respectively. Some discussion will be given in section 6.

2 The basic set-up and genetic scenario

The class of models we will discuss are thought to be applied in the following general scenario. The life spans of a population of n individuals are observed. Genetic information (typically in the form of a reasonably deep pedigree or comprehensive genomic data) and the values of k explanatory variables are available for each individual. The interest lies in modeling the longevity, *i.e.* the length of the life span or the length of the productive life of the individuals, with particular interest on some forms of genetic determination typically used in quantitative genetics.

The longevity can be operationally measured using a continuous time, by determining the time elapsed between two life events in the life of the individuals (*e.g.* the time elapsed from the first parity the to the culling of cows under production) or a discrete time, by counting a number of certain events during the life of the individual (*e.g.* by counting the number of survived parities of cows). We will present two illustrative examples: one involving the longevity of sows and the other studying the longevity of dairy cattle. Many similar examples can be found in the literature with other husbandry animals as ewes [25], salmon [20] among others.

An additional issue that might occur in longevity studies is the presence of competing risks, which arises when the individuals could be culled for one of two or more types of culling. For instance, in the example of dairy cattle longevity, the cows might die or be slaughtered, and the interest is in studying the culling rates for both causes. We assume in this scenario that there are two causes of death or culling, which will be labeled by the index j. It is straightforward to extend the setup described here to the general case with more than two culling causes or to the case with only one culling type.

The longevity will be characterized by the time development of the rate at which the individuals die or are culled. This is measured differently when using continuous or discrete time as we define below. In order to describe the models we have in mind, consider two random variables: T representing the time (continuous or discrete) of culling and J indicating the cause (or type) of culling. In the case of the continuous time, we define the *cause-specific hazard function* [19] for the j^{th} culling cause (for $j = 1, 2$) by

$$\lambda_{[j]}(t) = \lim_{\Delta \downarrow 0} \frac{P(t \leq T < t + \Delta, J = j | T > t)}{\Delta} \quad , \forall t \geq 0 \, .$$

In contrast, when the time is discrete, we define the *cause-specific hazard probability function* [19] for the j^{th} culling cause (for $j = 1, 2$) by

$$\lambda_{[j]}(t) = P(T = t, J = j | T \geq t) \text{ for } t = 0, 1, 2, \dots \, .$$

Although these two characterisations of the time development of the culling rate are of different nature, they will essentially play the same rule in the models discussed and will be generically referred simply as the cause-specific hazard function. Note that in the case where only one cause of culling is studied the cause-specific hazard function and the cause-specific hazard probability function defined above coincide with the hazard function and the hazard probability function used in the literature of survival analysis (see [19]).

The models presented below will assume a scenario where there are two causes of culling (indexed by j). Here, the cause-specific hazard functions (or the hazard probability functions) are specified in terms of k explanatory variables (the fixed effects) and a set of gaussian random components. Without loss of generality, we consider only two random components for each culling cause: $\mathbf{U}_1$ and $\mathbf{U}_2$ that will represent additive genetic effects (usually determined using information on the pedigree of the animals in play) for cause 1 and 2 respectively and $\mathbf{V}_1$ and $\mathbf{V}_2$, representing an environment effect (*e.g.* the effect of herd where an animal is kept) for cause 1 and 2, respectively.

The random components related to genetics (*i.e.* $\mathbf{U}_1$ and $\mathbf{U}_2$) are independent of the components related to environment ($\mathbf{V}_1$ and $\mathbf{V}_2$) by construction. This constraint can be relaxed for describing possible interactions between environmental and genetic effects, but we will not pursue this project here. The extension of the present definition to a situation with a different number of random components is straightforward. According to the general model we describe, the vector of cause specific hazard functions of the i^{th} individual (for $i = 1, \ldots, n$), conditionally on $\mathbf{U}_1 = \mathbf{u}_1$, $\mathbf{U}_2 = \mathbf{u}_2$, $\mathbf{V}_1 = \mathbf{v}_1$ and $\mathbf{V}_2 = \mathbf{v}_2$ (for the realizations $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{v}_1$ and $\mathbf{v}_2$ of $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{V}_2$ and $\mathbf{V}_2$, respectively), is given by

$$\begin{bmatrix} \lambda_{[1],i}(t_1|\mathbf{U}_1, \mathbf{V}_1) \\ \lambda_{[2],i}(t_2|\mathbf{U}_2, \mathbf{V}_2) \end{bmatrix} = \begin{bmatrix} \lambda_1(t_1) \exp\left(\mathbf{X}_{1i}^{'}(t_1)\boldsymbol{\beta}_1 + \mathbf{Z}_{1i}^{'}\mathbf{u}_1 + \mathbf{W}_{1i}^{'}\mathbf{v}_1\right) \\ \lambda_2(t_2) \exp\left(\mathbf{X}_{2i}^{'}(t_2)\boldsymbol{\beta}_2 + \mathbf{Z}_{2i}^{'}\mathbf{u}_2 + \mathbf{W}_{2i}^{'}\mathbf{v}_2\right) \end{bmatrix} . \quad (1)$$

Here the times t_1 and t_2 might be continuous or discrete; moreover, t_1 and t_2 are not necessarily of the same type (as in the example of sows longevity considered in section 5.1). Equation (1) holds for $t_j \geq 0$ when t_j is continuous and for $t_j = 0, 1, 2, \ldots$ if t_j is discrete ($j = 1, 2$). Furthermore, $\lambda_1(\cdot)$ and $\lambda_2(\cdot)$ are baseline hazard functions describing a common time development of the cause specific hazard functions (typically considered as a nuisance parameter), $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{R}^k$ are (finite dimensional) parameters representing the fixed effects (also viewed as a nuisance parameter), $\mathbf{X}_i$, $\mathbf{Z}_{ji}$ and $\mathbf{W}_{ji}$ are incidence matrices for the fixed, genetic and environment effects for the j^{th} cause and the i^{th} individual. When the time is continuous $\lambda_{[j],i}(\cdot)$ in (1) is a hazard function and the marginal model described will correspond to a variant of the Cox proportional hazard model with frailties [19, 3]. If the time is discrete, $\lambda_{[1],i}$ in (1) is a hazard probability function and the marginal model described will be a variant of the discrete relative risk model with frailties [19].

In the case of the models based on continuous time, we assume that the baseline function, $\lambda_j(\cdot)$ is piece-wise constant. That is, we assume that there is a set of disjoint intervals, $I_1, \ldots, I_K$ covering the positive real line, in which $\lambda_j(\cdot)$ is constant. This constraint in the class of possible baseline functions will allow us to connect the models described here with some generalized linear mixed models: Moreover, this will allow us to perform efficient inference with complex data using existent software.

The structure of covariance of the random components is a crucial part of the construction of the models for longevity described since this is the part of the model where the additive genetics and the environment contributions to the cause specific hazard functions are considered. We assume the random components to be multivariate normally distributed, *i.e.* $(\mathbf{U}_1,\ \mathbf{U}_2,\ \mathbf{V}_1,\ \mathbf{V}_2)' \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ where

$$\mathbf{\Sigma} = \left[\begin{array}{cc} \Sigma_1 \otimes \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \otimes \mathbf{I} \end{array} \right] . \tag{2}$$

Here, $\mathbf{I}$ is an identity matrix and $\mathbf{A}$ is the $n \times n$ additive relationship matrix constructed as follows. Each entry of $\mathbf{A}$ represents a pair of individuals in the population. $A_{ii} = 1$ for $i = 1, \ldots, n$, and $A_{ij} = 0$ if the pair of individuals (i, j) are not related. In the case were the pair of individuals (i, j) are related, we consider the degree of relationship r given by 1 for relatives of first-degree, 2 for relatives of second-degree, 3 for relatives of third-degree and so on. In that case $A_{ij} = (1/2)^r$. The additive relationship between two individuals measures the proportion of identical by descent genes expected to be shared by pairs of individuals in the population. The methods introduced here could also be applied to genomic analyses where the relationship matrices are inferred from genomic data.

The covariance structure between the genetic random components is given by

$$\Sigma_1 = \left[\begin{array}{cc} \sigma^2_{g.1} & \sigma_{g.12} \\ \sigma_{g.12} & \sigma^2_{g.2} \end{array} \right] .$$

In particular the genetic correlation between the two culling causes is given by $\rho_g = \frac{\sigma_{g.12}}{\sqrt{\sigma^2_{g.1}\sigma^2_{g.2}}}$. Finally, the covariance structure between the environment random components is given by

$$\Sigma_2 = \left[\begin{array}{cc} \sigma^2_{e.1} & \sigma_{e.12} \\ \sigma_{e.12} & \sigma^2_{e.2} \end{array} \right] .$$

3 Inference

3.1 Representation of the counting process and connections with generalized linear mixed models

The statistical inference for the longevity models can be made by using the coincidence of the conditional likelihood of those models (conditioned on the frailties) with the conditional likelihood of a generalized linear model applied to a specially designed pseudo data representing an underlying counting process as briefly described below. In the case of the continuous time, the pseudo data is constructed by generating one pseudo observation for each time interval (in which the baseline function is assumed to be constant) observed for each individual. In the case of the discrete time, the pseudo data contains one observation per each unit of time observed for each individual. To facilitate the description we use the term "observed times" to refer to the observed discrete times or the time intervals related to the observed continuous time. A pseudo response variable was created for each of the culling causes. These variables are such that they take the value 0 when an individual is at risk of being culled for the current reason but was not culled or take the value 1 if the individual was culled for the current reason in that period. This pseudo data describes the groups of individuals at risk at each observed time.

The statistical inference is performed by using a multivariate generalized linear mixed model (MGLMM) applied to the pseudo data. One dimensional versions of the calculations we describe below were given previously in [2] for models without random components and in [15] for mixed models for continuous time models (piece-wise constant baselines). For one-dimensional discrete times models without random components see [19]. Each of the marginal model of this MGLMM corresponds to a culling reason and is defined with a logarithmic link function. The distribution of each marginal model is Bernoulli (or binomial) when the corresponding time is discrete or Poisson when the time is continuous. Furthermore, the marginal models constructed with continuous time include an offset variable with the logarithm of the length of the corresponding time intervals. Additionally, a discrete explanatory variable counting the order of the time period for each individual should be included in the model in order to represent the baseline function. The model should also specify the covariance structure of the random components given in equation (2). Furthermore, each marginal model includes a dispersion parameter ϕ_j (known in the literature of generalized linear models as an over-dispersion parameter for binomial and Poisson models), which allowed us to better characterizing the genetic scenario.

The MGLMM described above can be adjusted with any software implementing multivariate generalized linear mixed models, which allows the inclusion of over-dispersion parameters in binomial and Poisson models via quasi-likelihood inference [36, 4]. We used the software DMU version 6.0, release 5.1 [23, 24], which implements multivariate generalized linear mixed models efficiently and facilities to specify the required covariance structure of the random components.

3.2 A Poisson approximation for discrete-time models

The inference for the discrete time models presented here is equivalent to making inference for a Bernoulli model where the response variable is a variable indicating whether a culling event occurred. Here, we present a Poisson approximation that will be useful for performing likelihood-based inference for the discrete time models, specially in the case of very large populations with relatively low occurrence of culling. More precisely, for each time t ($t = 0, 1. \ldots, \tau$, where τ is the maximum observed time) and for each individual i ($i = 0, 1, \ldots, n_t$ where n_t is the number of individuals at risk in the time t), we constructed a pseudo observation of a response variable Y_{it} taking the value 1 if the individual i died at time t and 0 otherwise. Exploring a coincidence of the likelihood function, we treated the observations Y_{it}, for $t = 0, 1. \ldots, \tau$ and $i = 0, 1, \ldots, n_t$, as independent Bernoulli random variables with $P(Y_{it} = 1) = p_{it}$, where p_{it} is the hazard probability for the individual i at the time t.

The following condition will ensure that the likelihood function of a suitably defined Poisson model will approximate the likelihood function of the Bernoulli models considered here. Suppose that

$$\max_{1 \leq t \leq \tau} \bar{p}_{t,n_t} \to 0 \ , \tag{3}$$

where $\bar{p}_{t,n_t} = \max_{1 \leq i \leq n_t} p_{it}$, and

$$\sum_{i=1}^{n_t} p_{it} = \gamma_t \ \text{(fixed)} \ \text{ for all } t \text{ as } \min_{1 \leq t \leq \tau} n_t \to \infty \ . \tag{4}$$

The general result on convergence for arrays of probabilities presented in [7] implies that, under the conditions (3) and (4), for each time t and $r = 0, 1, \ldots$,

$$\lim_{n_t \to 0} P(W_{n_t} = r) = \frac{e^{\gamma_t} \gamma_t^r}{r!} \ , \tag{5}$$

where $W_{n_t} = \sum_{i=1}^{n_t} Y_{it}$. Note that the right side of the equation (5) is the density probability function of a Poisson random variable with expected value equal to γ_t. In other words, if the number of individuals at risk in each time t (*i.e.* n_t) is larger enough and the probability of death in each time t is small enough, then the likelihood function of the Bernoulli model is approximated by the likelihood function of a Poisson model.

Note that the conditions given in (3) and (4) are reasonable in the example presented here, since the number of individuals is of the order of 100 thousands at the first observed period and is of the order of 16 thousands at the last observed period. Moreover, for a given individual i, the pseudo random variables Y_{it} $(t = 1, \ldots, t_i)$ is a sequence of zeros until time $t_i - 1$ and takes the value1 at t_i only if this individual was culled, therefore p_{it} will be small in general.

3.3 Continuous models with stratification

The examples of survival models with continuous time presented here required the use of stratification of the baseline function, which is a standard technique of survival analysis. For example, we used the number of days from the first parity to the culling day as one of the characterizations of the longevity of animals (sows or dairy cows), in the two examples presented here. Exploratory analyses of the data used indicated that there were differences in the mortality rates between the different parities and that the form of the baseline functions were not the same for the different parities, *i.e.* the baseline functions for the different parities were not proportional. A way to circumvent this issue is to use the stratification technique defined here. This type of model is referred in the literature of dairy cattle longevity as the "lactation basis model" (see [29]).

Define t_{p_i} as the time of the occurrence of the p^{th} parity of the individual i ($i = 1, \ldots, n$ and $p = 1, \ldots, P_i$ where P_i is the maximum parity observed for the individual i) and t_i^* as the observed time (death or censure) for the i^{th} individual. The hazard function for the i^{th} individual at the p^{th} parity, for $i = 1, 2, \ldots, n$ and for $p = 1, \ldots, P_i$ conditional on the random components

$\mathbf{U} = \mathbf{u}$ and $\mathbf{V} = \mathbf{v}$ is then

$$\lambda_{i,p}(t|\mathbf{U}, \mathbf{V}) = Y_{i,p}(t)\lambda_p(t) \exp\left(\mathbf{X}_{i,p}\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u} + \mathbf{W}_i\mathbf{v}\right),$$

for $t \in [t_{p_i}, t_{(p_i+1)})$ when $p < P_i$ and for $t \in [t_{P_i}, t_i^*)$ when $p = P_i$. Here $\lambda_p(\cdot)$ is the baseline hazard function for the p^{th} parity. The variable $Y_{i,p}(t)$ is equal to 1 if the individual was at risk at $t \in [t_{p_i}, t_{(p_i+1)})$ and 0 otherwise. In this study, the time dependent explanatory variables will be a function of p and then $X_{i,p}(t) = X_{i,p}$ for all $t > 0$.

3.4 Comparison of the statistical information between discrete and continuous time models

When applying the longevity models described above it is often possible to choose between models defined with continuous time and models defined with discrete time. For instance, in the examples studied in section 5 the longevity can be characterized by the time between the first parity and the culling (continuous time) using a stratification by parity or, alternatively, by the number of survived parities (discrete time). The complexity and the statistical stability of these two types of models differ, which is an important aspect to take into account when deciding which type of model to use. We will argue next that the likelihood function of a continuous time model is typically much flatter than the likelihood function of a corresponding discrete time model. This explains certain anomalies in the inference under continuous time models.

In order to simplify the exposition, we assume a scenario where two models are considered: a discrete time model where the number of survived parities is modeled and a continuous time model where the number of survived days is modeled but there is a stratification by parity (as described in section 3.3). Moreover, we restrict the discussion to the case where there is only one cause of culling and there is only one random component, say $\mathbf{U}$, in the model. Here, $\mathbf{U}$ will be a multivariate normal random variable with $E(\mathbf{U}) = \mathbf{0}$, $Var(\mathbf{U}) = \mathbf{A}\sigma^2$, where $\mathbf{A}$ is a known matrix. We denote the density function of the distribution of $\mathbf{U}$ by $\Phi\left(\cdot, \sigma^2\right)$ and represent the multiple integral for integration with respect to this density function by a single integration sign.

The likelihood function of the piece-wise constant hazard model stratified by parity based on the number of survival days is

$$\int \prod_{i,p,k} \exp\left\{-\Delta_{ipk} \exp\left(\eta_{ipk}\right)\right\} \left\{\Delta_{ipk} \exp\left(\eta_{ipk}\right)\right\}^{y_{ipk}} \Phi\left(\mathbf{u}, \sigma^2\right) d\mathbf{u} \ , \tag{6}$$

where the indices i, p and k indexes the individual, the parity and the time intervals in each periods, respectively. Here, y_{ipk} is a indicator variable taking the value 1 if the i^{th} individual died at the p^{th} parity in the k^{th} time interval and 0 otherwise; Δ_{ipk} is the length of the k^{th} time interval at the p^{th} parity; and $\eta_{ipk} = \log(\lambda_{pk}) + \mathbf{X}^{'}_{ipk}\boldsymbol{\beta} + \mathbf{Z}^{'}_{i}\mathbf{u}$ is the linear predictor.

On the other hand, by the Poisson approximation discussed in section 3.3, the likelihood function of the discrete relative risk model based on the number of survived parities is approximated by

$$\int \prod_{i,p} \exp\left\{-\exp\left(\eta_{ip}\right)\right\} \left\{\exp\left(\eta_{ip}\right)\right\}^{y_{ip}} \Phi\left(\mathbf{u}, \sigma^2\right) d\mathbf{u} \ , \tag{7}$$

where i and p indexes the individual and the parity, respectively; y_{ip} is the indicator variable that the i^{th} individual died at the p^{th} parity and $\eta_{ip} = \log(\lambda_p) + \mathbf{X}^{'}_{ip}\boldsymbol{\beta} + \mathbf{Z}^{'}_{i}\mathbf{u}$ is the linear predictor.

Note that, when there are many censured observations and when there are many survived parities or many observed time intervals in each parity, the variables y_{ipk} and y_{ip} will take the value zero in most of the cases. Therefore, under these conditions, the integrand in (6) will be dominated by

$$\prod_{i,p,k} \exp\left\{-\Delta_{ipk} \exp\left(\eta_{ipk}\right)\right\} = \exp\left\{-\sum_{i,p,k} \Delta_{ipk} \exp\left(\eta_{ipk}\right)\right\} \tag{8}$$

and the integrand in (7) will be dominated by

$$\prod_{i,p} \exp\left\{-\exp\left(\eta_{ip}\right)\right\} = \exp\left\{-\sum_{i,p} \exp\left(\eta_{ip}\right)\right\} \ . \tag{9}$$

Note that the sum in the right side of (8) has typically a larger number of parcels as compared to the sum in the right side of (9). Moreover, the parcels in the sum in the right side of (8) are multiplied by the factors Δ_{ipk} which are typically larger than 1. Note, moreover, that if one reduces the size of the intervals used in the piecewise constant hazard model, the factors Δ_{ipk} decrease, but the number of parcels in the sum in the right side of (8) increases; on the other hand, if one decreases the number of intervals, the number of parcels decreases, but the factors Δ_{ipk} might increase substantially. This makes the likelihood function of the piecewise constant hazard models much flatter, since the likelihood function is positive. More precisely, the curvature of the support curve (*i.e.* the graph of the logarithm of the likelihood function) near the maximum likelihood estimate will be much smaller for the piecewise constant hazard models. Therefore, the Fisher information will be smaller, and the maximization of the likelihood function will be a much harder problem, for the piecewise constant hazard models as compared to the corresponding discrete relative risk model.

4 Decomposition of the phenotypic variance and heritability

The notion of heritability is crucial in quantitative genetic applications; it measures the magnitude of the detected additive genetic signal relative to the total variance of a trait of interest and is typically used to quantify the potential response to selection. Heritability (in the narrow sense) is defined operationally by the ratio

$$h^2 = \frac{\sigma_g^2}{\sigma_P^2} \ , \tag{10}$$

where σ_g^2 is the variance of an unobserved random component representing all the additive genetic variation (termed the genetic variance) and σ_p^2 is the variance of the trait of interest (called the phenotypic variance). For details see [22, 18]. In the classic scenario, where gaussian linear mixed models are used, the calculation of the heritability is straightforward because the phenotypic variance can be expressed as the sum of the genetic variance σ_g^2 and the variances of the other random components present in the model. However, the calculation of the heritability is much more complicated in the context of survival analysis considered here and are, therefore, developed in full details below. It will be necessary to introduce some details of the counting process behind the models used and to define the trait of interest carefully. Here, we present the main results informally and expose the precise formal definitions and the technical details using the mathematical machinery of counting processes in appendix A. For simplicity of exposition, we discuss only the one-dimensional case were only one cause of culling is present (referred as death) and assume the same model structure presented in section 3 but with dimension one.

Let T be a non-negative random variable representing the observed survival time and D be an indicator variable taking the value 1 if a death is observed and 0 otherwise (censure). The random variable T will take values in $\mathbb{R}_+$ if the time is continuous or in $\mathbb{Z}_+$ (*i.e.* in $\{0, 1, 2, \ldots\}$) if the time is discrete. We assume that for each time t a vector $X(t)$ of explanatory variables (possibly changing with time) is known. Denote by $\mathbf{X}(t) = \{X(s) : 0 \leq s < t\}$ the trajectory (or path) function until the time just before t (denoted by $t-$). Additionally, we consider two independent gaussian random components, $\mathbf{U}$ and $\mathbf{V}$, representing the additive genetic effects and some environmental effects, respectively. Here, we follow the same convention as in section 2, so $Var(\mathbf{U}) = \mathbf{A}\sigma_g^2$ and $Var(\mathbf{V}) = \mathbf{I}\sigma_e^2$, where $\mathbf{A}$ is the relationship matrix and $\mathbf{I}$ is an identity matrix with suitable dimensions. The hazard function for the i^{th} individual takes then the form

$$\lambda_i(t|\mathbf{U}, \mathbf{V}) = \lambda_0(t) \exp\left(\mathbf{X}_i^{'}(t)\boldsymbol{\beta} + \mathbf{Z}_i^{'}\mathbf{u} + \mathbf{W}_i^{'}\mathbf{v}\right) . \tag{11}$$

In order to describe the models presented in section 3 properly, we need to introduce two stochastic processes: the death (or event) counting process and the at risk process. To define the death process we consider, for each t, the random variable $N(t)$, which indicates whether a death occurred until the time t, *i.e.* $N(t) = \mathbf{1}(T \leq t, D = 1)$. The stochastic process $\mathbf{N} = \{N(t) : t \in \mathbb{R}_+ \text{ or } \mathbb{Z}_+\}$ is the death counting process. It is convenient to introduce the notation $dN(t)$ to denote the increments of the counting process $\mathbf{N}$ at each time t, *i.e.* $dN(t) = N(t) - N(t-)$, where $N(t-) = \lim_{\Delta \downarrow 0} N(t - \Delta)$. Furthermore, the at risk process given by $\mathbf{Y} = \{Y(t) : t \in \mathbb{R}_+ \text{ or } \mathbb{Z}_+\}$, where, for each time t, the random variable $Y(t)$ takes the value 1 when the individual is at risk at time t or 0 otherwise.

We will calculate the phenotypic variance and the heritability for two characteristics of interest: the hazard function evaluated at a given time and the accumulated hazard function up to a given time. These two characteristics are associated to the intensity and the cumulative intensity of the counting process of death $\mathbf{N}$. Due to the structure of the models used here (see equation (1) and (11) for instance), the entire hazard function of an individual is multiplied by the factor $\exp(\mathbf{u})$, where $\mathbf{u}$ is the realization of the genetic random component $\mathbf{U}$ for that individual; so, according to this model, when selecting individuals by the predicted values of $\mathbf{U}$ (as it is the usual practice in animal breeding), the entire hazard curve is affected and therefore, it is indifferent if one measures the potential response to selection (*i.e.* the heritability) of the hazard evaluated at a given time-point, say t, or even of the accumulated hazard up to t.

In order to study the heritability related to the hazard function evaluated at a given time t (in $\mathbb{R}_+$ or in $\mathbb{Z}_+$), we define the conditional random variable

$$\mathcal{Y}(t) = dN(t) \,|\, T \geq t, X(t)\,. \tag{12}$$

Note that $E\left(\mathcal{Y}(t) \,|\, \mathbf{U}=\mathbf{u},\, \mathbf{V}=\mathbf{v}\right) = \lambda\left(t \,|\, \mathbf{U}=\mathbf{u},\, \mathbf{V}=\mathbf{v}\right)$, therefore $\mathcal{Y}(t)$ will be treated as the trait of interest, for which we calculate the phenotypic variance and the heritability. A more precise definition of $\mathcal{Y}(t)$ using counting processes theory is given in appendix A. There we show, taking advantage of the basic theory of martingales and using a suitable Taylor expansion, that in the continuous time case the variance of $\mathcal{Y}(t)$ is approximated by

$$Var\left[\mathcal{Y}(t)\right] \approx Y(t)\left\{\left[\lambda^*(t)\right]^2\left(\sigma_g^2+\sigma_e^2\right)+\phi\lambda^*(t)\right\}\ , \tag{13}$$

where $\lambda^*(t)$ is the hazard function evaluated at t with vanishing random components, *i.e.*

$$\lambda^*(t) = \lambda(t|\mathbf{U}=\mathbf{0},\mathbf{V}=\mathbf{0}) = \lambda_\mathbf{0}(\mathbf{t})\exp\left\{\mathbf{X}_\mathbf{i}^{'}(\mathbf{t})\boldsymbol{\beta}\right\}\ . \tag{14}$$

Note that according to (13) the variance of $\mathcal{Y}(t)$ is additively decomposed in three components: one depending on σ_g^2, one depending on σ_e^2 and one depending on ϕ. This is a situation analogous to the classic gaussian linear mixed models. Here, the component of the variance of $\mathcal{Y}(t)$ associated to σ_g^2, namely $\left[Y(t)\lambda^*(t)\right]^2\sigma_g^2$, plays the role of the genetic variance and therefore the heritability for the hazard evaluated at the time t is given by

$$h^2_{\lambda(t)} \approx Y(t)\frac{\sigma_g^2}{\sigma_g^2+\sigma_e^2+\frac{\phi}{\lambda^*(t)}}\ . \tag{15}$$

Analogous calculations (see appendix A) yield for the discrete time case that the variance of $\mathcal{Y}(t)$ is approximated by

$$Var\left[\mathcal{Y}(t)\right] \approx Y(t)\left\{[\lambda^*(t)]^2\left(\sigma_g^2+\sigma_e^2\right)+\phi\lambda^*(t)[1-\lambda^*(t)]\right\}\ . \tag{16}$$

Here, as in the case of the continuous time, the variance of $\mathcal{Y}(t)$ is additively decomposed in three separate a components depending on σ_g^2, on σ_e^2 and on ϕ, respectively. Therefore, the heritability is approximated by

$$h^2_{\lambda(t)} \approx Y(t)\frac{\sigma_g^2}{\sigma_g^2+\sigma_e^2+\phi\frac{\lambda^*(t)}{1-\lambda^*(t)}}\ . \tag{17}$$

Moreover, when a Poisson approximation is used, the approximated heritability takes the form given in (15).

We turn now to the calculation of the heritability of the accumulated hazard function evaluated at a time point t. In the continuous time case we define the accumulated hazard function, for each $t \in \mathbb{R}_+$, by

$$\Lambda(t) = \int_0^t Y(s)\lambda(s)ds\,, \tag{18}$$

and in the discrete time case the accumulated hazard function is given, for each $t \in \mathbb{Z}_+$, by

$$\Lambda(t) = \sum_{s\le t} Y(s)\lambda(s)\,, \tag{19}$$

where $\lambda(\cdot)$ is the hazard function in play. Furthermore, the conditional random variable

$$\mathcal{Z}(t) = N(t)\,|\,T \ge t, X(t) \tag{20}$$

is such that $E\left(\mathcal{Z}(t)\,|\,\mathbf{U}=\mathbf{u},\,\mathbf{V}=\mathbf{v}\right) = \Lambda\left(t\,|\,\mathbf{U}=\mathbf{u},\,\mathbf{V}=\mathbf{v}\right)$. Therefore, $\mathcal{Z}(t)$ will be treated as the trait of interest, for which we calculate the phenotypic variance and the heritability. A more precise definition of $\mathcal{Z}(t)$ using counting processes theory is given in appendix A. It can be shown (see appendix A) that in the case of the continuous time the variance of $\mathcal{Z}(\cdot)$ evaluated at a given $t \in \mathbb{R}_+$ is approximated by

$$Var\left[\mathcal{Z}(t)\right] \approx [\Lambda^*(t)]^2\left(\sigma_g^2+\sigma_e^2\right)+\phi\Lambda^*(t) \tag{21}$$

where $\Lambda^*(t) = \int_0^t Y(s)\lambda^*(s)ds$. Therefore, the heritability is approximated by

$$h^2_{\Lambda(t)} \approx \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2 + \phi\left[\Lambda^*(t)\right]^{-1}} \ . \tag{22}$$

In the case of the discrete time the variance of $\mathcal{Z}(\cdot)$ evaluated at a given $t \in \mathbb{Z}_+$ is approximated by

$$Var\left[\mathcal{Z}(t)\right] \approx \left[\Lambda^*(t)\right]^2 \left(\sigma_g^2 + \sigma_e^2\right) + \phi\gamma(t), \tag{23}$$

where $\gamma(t) = \sum_{i=0}^{t} \phi Y(t)\lambda^*(t)[1 - \lambda^*(t)]$ and the cumulative hazard is $\Lambda^*(t) = \sum_{s=0}^{t} Y(s)\lambda^*(s)$ .Then, the heritability is approximated by

$$h^2_{\Lambda(t)} \approx \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2 + \phi\gamma(t)[\Lambda^*(t)]^{-2}} \ . \tag{24}$$

When the Poisson approximation is used the heritability is approximated by the right side of (22).

The calculations above are performed for one individual. We propose that given a population of n individuals, the phenotypic variances and the heritabilities for the hazard function or the cumulative hazard (evaluated at a time t) should be calculated using the estimated variance components and an estimated hazard or an estimated cumulative hazard as described below. In the continuous time case, where the PCHM is used, the observed time is split in intervals $I_k = [t_k, t_{k+1})$, for $k = 0, \ldots, K$ where $0 = t_0 < t_1 < \ldots < t_K < t_{K+1} = \tau$, where τ is the largest observed time. Given that $\mathbf{U} = \hat{\mathbf{u}}$ and $\mathbf{V} = \hat{\mathbf{v}}$, where $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ are the BLUP of the random components, the conditional hazard is defined, for each $t \in I_k$, by

$$\lambda_i^*(t) = \lambda_k \exp\left[\mathbf{X}_{i,k}^{'}\boldsymbol{\beta} + \mathbf{Z}_i^{'}\hat{\mathbf{u}} + \mathbf{W}_i^{'}\hat{\mathbf{v}}\right] = \exp\left[\eta_{i,k}(t)\right] \ ,$$

where $\eta_{i,k}(t) = \log(\lambda_k) + \mathbf{X}_{i,k}^{'}\boldsymbol{\beta} + \mathbf{Z}_i^{'}\hat{\mathbf{u}} + \mathbf{W}_i^{'}\hat{\mathbf{v}}$ and $\mathbf{X}_{i,k} = \mathbf{X}_i(t)$ for $t \in I_k$. We obtain $\hat{\eta}_{i,k}(t) = \widehat{\log(\lambda_k)} + \mathbf{X}_{i,k}^{'}\hat{\boldsymbol{\beta}} + \mathbf{Z}_i^{'}\hat{\mathbf{u}} + \mathbf{W}_i^{'}\hat{\mathbf{v}}$ from the adjusted model and then calculate, for each time $t \in I_k$,

$$\tilde{\lambda}^*(t) = \tilde{\lambda}_k^* = \exp\left[\bar{\eta}_k\right] \ ,$$

where $\bar{\eta}_k = n^{-1}\sum_{i=1}^{n}\hat{\eta}_{i,k}$. The cumulative hazard is estimated, for $t \in I_k$, by

$$\tilde{\Lambda}^*(t) = \sum_{j=0}^{k-1}\Delta_j\tilde{\lambda}_j^* + (t - t_k)\tilde{\lambda}_k^* \ ,$$

where $\Delta_j = t_{j+1} - t_j$ is the length of the interval I_j.

In the discrete time case the hazard, conditionally that $\mathbf{U} = \hat{\mathbf{u}}$ and $\mathbf{V} = \hat{\mathbf{v}}$, is given, for each $t \in \mathbb{Z}_+$, by

$$\lambda_i^*(t) = \lambda_t \exp\left[\mathbf{X}_i^{'}(t)\boldsymbol{\beta} + \mathbf{Z}_i^{'}\hat{\mathbf{u}} + \mathbf{W}_i^{'}\hat{\mathbf{v}}\right] = \exp\left[\eta_i(t)\right] \ ,$$

where $\eta_i(t) = \log(\lambda_t) + \mathbf{X}_i^{'}(t)\boldsymbol{\beta} + \mathbf{Z}_i^{'}\hat{\mathbf{u}} + \mathbf{W}_i^{'}\hat{\mathbf{v}}$. We obtain $\hat{\eta}_i(t) = \widehat{\log(\lambda_t)} + \mathbf{X}_i^{'}(t)\hat{\boldsymbol{\beta}} + \mathbf{Z}_i^{'}\hat{\mathbf{u}} + \mathbf{W}_i^{'}\hat{\mathbf{v}}$ from the adjusted model and then calculate for each time t $(t = 0, 1, 2, \ldots,)$

$$\tilde{\lambda}(t) = \exp\left[\bar{\eta}(t)\right] \ ,$$

where $\bar{\eta}(t) = n^{-1}\sum_{i=1}^{n}\hat{\eta}_i(t)$. The cumulative hazard is estimated by $\tilde{\Lambda}^*(t) = \sum_{j=0}^{t}\tilde{\lambda}^*(j)$. The heritabilities will be estimated at $\tilde{\lambda}^*(t_m)$ and $\tilde{\Lambda}^*(t_m)$ where t_m represents the median survival time, *i.e.* $P(T \geq t_m) = 0.5$ in the illustrative examples.

5 Two illustrative examples

The methods presented above will be illustrated by two examples: one involving the longevity of sows (section 5.1) and another characterizing the longevity of dairy cattle (section 5.2). In both examples, we will characterize the longevity of female animals in two different ways: first, by the length of the productive life of the animals (sows or cows) measured by the number of days from the first parity to the culling day (ND); second, by the number of survived parities (NP). Models based on continuous time and models based on discrete time were applied to study ND and NP, respectively.

Incomplete observations are present in both examples because some animals were still alive at the end of the study, some animals were moved, sold or exported during the study and due to intentional right truncation. In the study of sow longevity the proportion of censure was relatively low (9.6%), while in the study of longevity of cows a much higher proportion of censure was observed (23.5%). Furthermore, in the study of longevity of cows it was possible to distinguish two culling causes: slaughter (67.7% of the observations) and death (only 8.8% of the observations), which characterizes a typical scenario of competing risks. Therefore, it is natural to use bivariate models in this example to model these two culling causes simultaneously. Bivariate models will be used in the example of sows longevity to model ND and NP simultaneously and use that to discuss how much the information of these two characterizations of the longevity overlap.

5.1 Longevity of sows

The data in this example was retrieved from the registers of the Danish Pig Center of 200,084 pure Landrace sows that were giving birth in the years between 1999 and 2010. Only the sows that had the first parity in the comprised period were included in the study. All the models proposed for ND and NP included the following explanatory variables: age at the first parity, year and season at the first parity, litter size (total number of piglets born per parity) and herd year size (total number of sows farrowing per herd per year). The litter size and the herd year size were both time dependent variables. In addition two random components were included: a sire component with pedigree, representing the sire additive genetic effects and a herd-year component representing the environment effects. Since a sire model was used, the additive genetic variance was estimated by multiplying the additive sire variance by four.

The Kaplan-Meyer estimate of the median survived NP was 2 parities (CI95% 2;3) and of the median survived ND was 320 days (95% CI, 319 - 322). The median number of days between two consecutive parities for the sows that did not die in the respective period varied from 152 to 155 days. Thus, the median survived time of 320 days from the first parity is approximately the time necessary for the sow, give it had one parity, have the second and survive until a third parity.

Figure 1 displays the logarithm of the cumulative hazard curves for ND stratified by parity. The curves per parity presented similar behavior, however, they are not parallels indicating a non proportional effect of parity. The dashed lines represents the cut-points of the intervals where the baseline hazard function was assumed to be constant.

Table 1 presents the estimates of the variance components and heritabilities from the piece-wise constant hazard model stratified by parity based on the ND and for two discrete relative risk models based on the NP (an exact model and an approximated model via Poisson approximation). The estimates of variance for the sire and the herd-year components were very similar among the three adjusted models (~ 0.065 with SE ~ 0.003 for the sire component and ~ 0.220 with SE ~ 0.009 for the herd-year component). Comparing the execution time between the two adjusted discrete models the approximated models was around 72 times faster compared to the time spent to execute the exact model.

A bivariate model describing the ND and the NP traits (using a Poisson approximation for the discrete marginal models) was fitted (see Table 2).

Table 1: Estimated variance components (with asymptotic standard error in parenthesis) and heritabilities for the number of days (ND) and for the number of parities (NP) at the sows longevity study.

Source	ND-PCHM[a]	NP-DRRM[b]	NP-DRRM$_P^{c}$
Sire	0.071 (0.005)	0.066 (0.002)	0.063 (0.003)
Herd-Year	0.217 (0.010)	0.220 (0.005)	0.219 (0.009)
Dispersion	3.089 (0.002)	0.978 (0.002)	0.653 (0.001)
$h_\lambda^{2\,\mathrm{d}}$	-	0.098	0.096
$h_\Lambda^{2\,\mathrm{e}}$	0.180	0.161	0.165
Execution time[e]	48.5 min	8h 28 min	7 min

[a] PCHM: Piece-wise constant hazard model stratified by parity.

[b] DRRM: Discrete relative risk model.

[c] DRRM$_P$: Discrete relative risk model via Poisson approximation.

[d] Estimated heritability for the hazard (λ) at the median survival time.

[e] Estimated heritability for the cumulative hazard (Λ) at the median survival time.

[f] Execution time (Intel i5 processor).

Table 2: Estimated variance components and correlations from the bivariate model based on the number of days (ND) and number of parity (NP) at the sows longevity study.

Source	Trait	σ^2 (SE)	Cor (SE)
Sire	**ND**	0.022 (0.001)	1.000 (0.004)
	NP	0.081 (0.004)	
Herd-Year	**ND**	0.332 (0.015)	1.000 (0.002)
	NP	0.231 (0.010)	
Dispersion	**ND**	3.130 (0.002)	-
	NP	0.657 (0.001)	

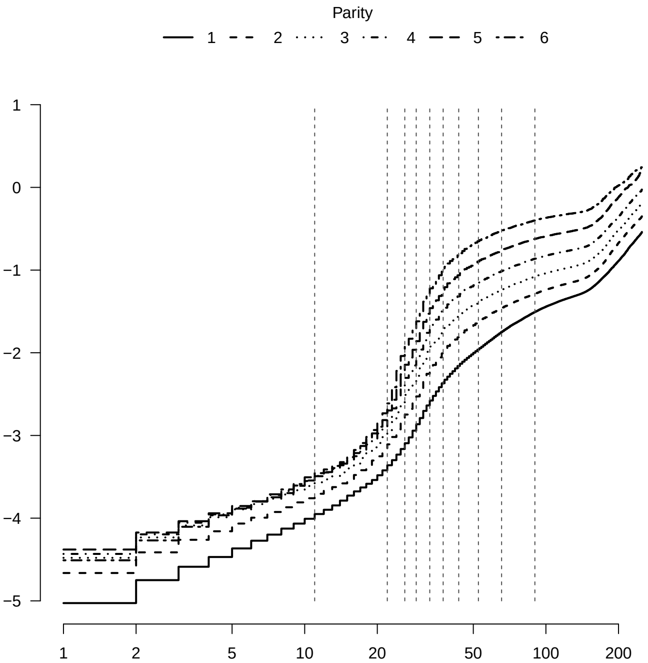

Figure 1: Log of cumulative hazard curves for the number of days trait(ND) stratified by parity at the sow's longevity study. Axis x is presented on the logarithm of days scale. The vertical dashed lines are the cut points for the intervals where the baseline function was assumed to be constant.

5.2 Longevity of dairy cattle

The data analyzed in this example was retrieved from the register of Knowledge Centre for Agriculture of 142,133 Jersey Danish dairy cows that were calving in the period between 1990 to 2006. The explanatory variables (fixed effects) included in the models were: herd year size (number of calving per herd per year), year and season of each parity, as time dependent variables and age at the first parity and a coefficient of general heterosis (defined as a continuous explanatory variables), as time independent variables. Two random components were included in the model, a combination of herd-year-season component representing the environment effects and a sire component with the pedigree representing the genetic effect. In addition, the cows could be culled by one of two possible reasons: death and slaughter.

The Kaplan-Meyer estimate of the overall median NP was 3 parities (95% CI, 3;3) and overall median ND was 826 days (95% CI, 820 - 831). The median number of days between 2 consecutive parities for the cows that did not die in the respective period varied from 361 days to 367 days. Figure 2 displays the stratified cumulative incidence probability curves for ND for the slaughter and the death, respectively. The dashed lines represent the cut points for the intervals where the baseline hazard were assumed to be constant for both specific reasons.

Tables 3 displays the estimates of the variance and covariance components from the competing risk models describing the ND and NP traits. The sire end herd-year-season variances for death rate based on the ND trait were very small, both for the death and the slaughter rate. In contrast, the discrete time model detected significantly larger variances of the random components. The failure of the continuous time models can be explained by the large amount of censure in the data.

Table 3: Estimated variances and covariances components for the number of parities (NP) and the number of days (ND) from a discrete and continuous multivariate competing risk model, respectively, at the dairy cattle longevity study.

		NP		ND	
Source	**Cause**	σ^2 (**SE**)	Cov (**SE**)	σ^2 (**SE**)	Cov (**SE**)
Sire	**Death**	0.102 (0.041)	0.002	0.7×10^{-3} (0.006)	0.002
	Slaughter	0.029 (0.003)	(0.004)	0.008 (0.002)	(0.003)
HYS [a]	**Death**	0.458 (0.015)	-0.055	0.1×10^{-5} (0.079)	0.7×10^{-7}
	Slaughter	0.061 (0.002)	(0.004)	0.3×10^{-6} (0.010)	(0.021)
ϕ_j [b]	**Death**	0.691 (0.002)	-	5.165 (0.005)	-
	Slaughter	0.710 (0.002)		4.148 (0.004)	
h^2_λ [c]	**Death**	0.013		-	
	Slaughter	0.040		-	
h^2_Λ [d]	**Death**	0.044		0.2×10^{-3}	
	Slaughter	0.108		0.012	

[a] Herd-year-season random component.

[b] Marginal dispersion parameters.

[c] Marginal heritabilities for the hazard (λ) at the median overall survival time.

[d] Marginal heritabilities for the cumulative hazard (Λ) at the median overall survival time.

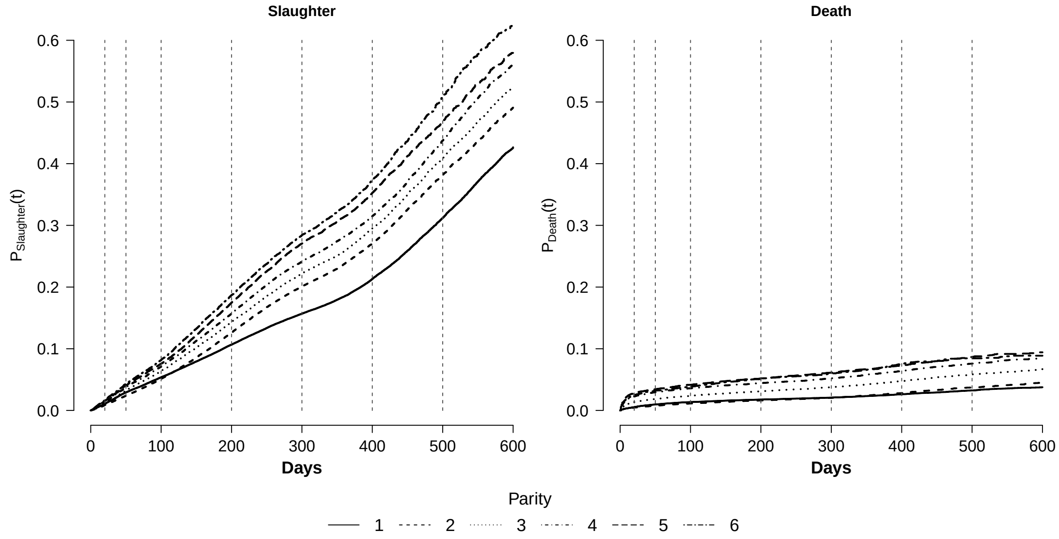

Figure 2: Cumulative incidence probability curves for the ND stratified by parity for the cow's longevity study. The dashed vertical lines are cut points for the intervals were the hazard was assumed constant. $P_J(t) = P(T \leq t, J = j)$ for $J = (Death, Slaughter)$.

6 Discussion

The several variants of models for studying longevity in quantitative genetics can be described using the following general approach. The models are one-dimensional and describe the conditional hazard function, given the random components, as taking the form

$$\lambda_i(t|\mathbf{U} = \mathbf{u}, \mathbf{V} = \mathbf{v}) = \lambda_0(t) \exp\left(\mathbf{X}_i'(t)\boldsymbol{\beta} + \mathbf{Z}_i'\mathbf{u} + \mathbf{W}_i'\mathbf{v}\right) ,$$

for t in $\mathbb{R}_+$ or in $\mathbb{Z}_+$. The models based on continuous time used in the literature differ essentially in the form of the baseline hazard function $\lambda_0(\cdot)$: the classical Cox model (with frailties) assumes the hazard function to be a smooth function (almost everywhere) imposing in this way almost no restriction on $\lambda_0(\cdot)$ (see [19, 26, 28]); the piece-wise Weibull models (see [9]) assume $\lambda_0(\cdot)$ to be a saw function (*i.e.* a piecewise linear function); while the piece-wise constant hazard models (as described here) assume $\lambda_0(\cdot)$ to be a step function (*i.e.* a piece-wise constant function). These models are nested, in the sense that saw functions are smooth almost everywhere, and step functions are particular cases of saw functions. Moreover, smooth functions can be arbitrarily approximated (point-wisely) by saw functions or by step functions. A fourth category of models, which was introduced here, are the piece-wise constant hazard models with free dispersion parameter. These models are an instance of quasi-likelihood based models (see [36, 4]). They contain the traditional piece-wise constant hazard models as a particular case (by setting the dispersion parameter ϕ to be constant equal to 1) and might arbitrarily approximate the piece-wise Weibull models and the Cox frailty models. Although in many practical situations those models yield equivalent representations of the data and produce similar results, this is not always the case for the quantitative genetics applications discussed here.

One of the interesting features of the piece-wise constant hazard model with (free) dispersion parameter presented here is that the phenotypic variance can be approximately decomposed as the sum of parcels representing the additive genetic effects, environmental effects and unspecified sources of variability. This is analogous to the classic decomposition of the phenotypic variance obtained in gaussian linear mixed genetic models. Here, the component of the phenotypic variance involving the dispersion parameter ϕ plays the rule of the residual variance in the classic models. Note that, if we fix the dispersion parameter ϕ, say by setting it equal to 1 as in the piece-wise constant hazard model, it might be that we obtain models that describe the data well and the referred decomposition phenotypic variance would still exists, but this would be equivalent to set the residual variance to be constant, which clearly causes problems in the representation of quantitative genetic phenomena. For instance, in a sire model the genetic variance transmitted by the dam is not captured by the random component representing the sire, therefore, this part of the variability should be represented in the residual variance (*i.e.* should compose part of the residual variance); but this is not possible to occur if the residual variance is set to be constant. This issue occurs also in the piece-wise Weibull models (see [9]). Furthermore, in the most general case where λ_0 is only assumed to be smooth (the Cox frailty model) it might be argued that λ_0 could be point wisely approximated (almost every where) by a sequence of step functions obtained even from distributions in a model with dispersion parameter, which would have a proper decomposition of the phenotypic variance. However, in this case it would not be possible to identify the dispersion parameter ϕ (as essentially argued in [13]); therefore, one would still have the representation issues referred above.

The piece-wise constant hazard model and the piece-wise Weibull model are parametric models, making it possible to implement relatively efficient inference techniques. This explains why they are implementable for complex and large models as in quantitative genetics typical applications. Moreover, when the baseline hazard function $\lambda_0(\cdot)$ is assumed to be piece-wise linear (a saw function), as in the piece-wise Weilbull models, the rate of approximation of $\lambda_0(\cdot)$ to a smooth function (as in the frailty model) is improved (specially if there are regions where the smooth baseline is steep) as well argued in [9]; however, in this case the inference via generalized linear mixed models would not be possible without using profile likelihood techniques (to estimate the shape parameter), which rules out the direct use of extensions containing a dispersion parameter or multivariate extensions already implemented. Finally, a disadvantage of the piece-wise constant hazard models and the piece-wise Weibull models is that both depend on an arbitrary choice of time cutting points.

Models based on discrete time (*i.e.* variants of the discrete relative risk models) has been occasionally used in quantitative genetic previously; however, without incorporating a dispersion parameter. These models do not assume any pre-specified form for the baseline hazard function and also not require any arbitrary definition of time cutting points; therefore, they are similar to the Cox frailty models. Due to the coincidence with the Bernoulli models with random components (with or without dispersion parameter), techniques for parameter inference are at hand, allowing for efficient implementations capable to handle complex large problems as illustrated here. Moreover, these models allow to incorporate time-dependent variables (by splitting data with record for each survived time for each individual) both as fixed effects and as random components. The Poisson approximation presented here might represent a substantial save in computational resources and yields essentially the same results as the exact inference, as illustrated in the example of sows longevity presented in section 5.1.

The continuous time models and the discrete time models yielded similar results in the example of sows longevity. However, the continuous time models failed in the example of dairy cattle longevity (section 5.2), which can be explained by the fact that, in this case, both rare events and high frequencies of censored observations are observed. This is in line with the limitations of the continuous time models discussed in section 3.4 and shows that those limitations do occur in practical situations.

The lack of multivariate techniques for analyzing several traits simultaneously has been a serious limitation in the application of survival models in quantitative genetics [30, 11]. This lack not only makes the study of the relation between longevity and other traits impossible, but also limits very much the possibility of performing time-to-event-analysis in the presence of competing risks. Indeed, if there are two competing culling reasons, one might use standard survival models for modeling the rate of occurrence of one of the culling reasons by considering the observations where the other culling reason occurred as censored. However, this naive approach implies in the implicit use of the assumption that the competing culling reasons are independent. The example of dairy cattle longevity presented here illustrates that this assumption might be not reasonable. Indeed, in this example we detected a significant negative correlation between the herd-year-season random components; moreover, a range of fixed effects were found to be significant for both culling reasons, implying that a naive use of marginal models would violate the assumption of absence of informative censoring.

Acknowledgements

Rafael Pimentel Maia was financed by the project "Svineavl: Developing New Methods for Genetic Selection of Sow Durability" , Ministry of Food,

Agriculture and Fisheries of Denmark.

References

[1] A. Abdelqader, A.A. Yacoub, and M. Gauly, *Factors influencing productive longevity of Awassi and Najdi ewes in intensive production systems at arid regions*, Small Ruminant Research, 104 (2012), 37–44.

[2] N. Aitkin and D. Clayton, *The fitting of exponential and extreme value distributions to complex censored data using GLIMM*, Journal of the Royal Statistical Society. Series C (Applied Statistics), 29(2) (1980), pp. 156–163.

[3] P.K. Andersen, Ø. Borgan, R.D. Gill and N. Keiding *Statistical models based on counting processes* Spring Series in Statistics, Springer, 1997.

[4] N.E. Breslow, and D.G. Clayton, *Approximate inference in generalized linear mixed models*, Journal of the American Statistical Association, 88(421) (1993), pp. 9–25.

[5] D. Butler, B.E. Cullis, A.R. Gilmour, B.J. Gogel, *Analysis of mixed models for S-language Environments:* ***ASReml-R*** *Reference Manual*, Queensland DPI, Brisbane, Australia (2004). Available at http://www.vsni.co.uk/resources/doc/asreml-R.pdf.

[6] D.Z. Caraviello, Weigel, K. and D. Gianola, *Prediction of longevity breeding values for US Holstein sires using survival analysis methodology*, Journal of Dairy Science, 87(10) (2004), pp. 3518-25.

[7] L. H. Y. Chen , *On the convergence of Poisson Binomial to Poisson Distributions.* The Annals of Probability, 2(1) (1974), pp. 178–180.

[8] D.R. Cox, *Regression Models and Life-Tables.* Journal of Statistical Society. Series B (methodologial), 34(2) (1972), pp. 187–220.

[9] V. Ducrocq, R. L. Quaas, E.K. Pollak, and G. Casella, *Length of Productive Life of Dairy Cows. 1. Justification of a Weibull Model*, Journal of Dairy Science, 71(11) (1988), pp. 3061–3070.

[10] V. Ducrocq and V.J. Sölkner, *The Survival Kit, a Fortran Package for the Analysis of Survival Data* proceedings of the $5th$ World Congress on Genetics Applied to Livestock Production, Vol. 21, Univeristy of Guelph ontario, Canada, (1994), pp. 51–52.

[11] V. Ducrocq, *Topics that may deserve further attention in survival analysis applied to dairy cattle breeding: some suggestions*, in: Proceedings of Workshop on genetic improvement of functional traits in cattle, Jouy-en-Josas, Interbull Bulletin 21 (1999), pp. 181189, Swedish University of Agricultural Sciences, Dept. of Anim. Breed. and Genet., Uppsala.

[12] L. Engblom, N. Lundeheim, A. Dalin, and K. Andersson, *Sow removal in Swedish commercial herds*, Livestock Science, 106(1) (2007), 76–86.

[13] S.R. Giolo, and C. G.B. Demétrio, *A frailty modeling approach for parental effects in animal breeding*, Journal of Applied Statistics, 38(3) (2011), 619–629.

[14] J.L.L. Guerra *Statistical models and genetic evaluation of binomial traits*, Ph.D. diss. Program in Animal and Dairy Science, Louisiana State University and Agricultural and Mechanical College, Baton Rouge, (2004).

[15] D. Ha and Y. Lee, *Estimating frailty models via Poisson hierarchical generalized linear models*, Journal of Computational and Graphical Statistics, 12(3) (2003), pp. 663–661.

[16] A. Hald, *Statistical theory with engineering applications.* John Wiley & Sons, Inc., New York, 1952.

[17] M. Hoque and J. Hodges, *Genetic and phenotypic parameters of lifetime production traits in Holstein cows*, Journal of dairy science, 63(11) (1980), pp. 1900–1910.

[18] A. Jacquard, *Heritability : One Word , Three Concepts*, Biometrics, 39(2) (1983), 465–477.

[19] J.D. Kalbfleish and R.L. Prentice, *The Statistical Analysis of Failure Time Data*, 2^{nd} Edition, Wiley series in Prob. and Statisc., John Wiley & Sons, Inc., Hoboken, New Jersey, 2002.

[20] R. Labouriau, M. Wetten and P. Madsen. *Simultaneous analysis of two challenge tests and body weight in fish: an example of multivariate generalized linear mixed model.* In: Book of Abstracts of the 62nd Annual Meeting of the European Federation of Animal Science. Vol. 16 Wageningen Academic Publishers (2011). p. 325–325.

[21] N. López-Villalobos, M. Penasa, R.D. Zotto, M. Cassandro, W. Brade, O. Distl, R. Evans, and A. Cromie, A. *Calculation of a cow culling merit index including specific heterosis in a multibreed dairy population*, Archiv Tierzucht, 53(1) (2010), 9–17.

[22] M. Lynch and B. Walsh, *Genetics and analysis of quantitative traits.* Sinauer, Massachusetts, 1998.

[23] P. Madsen and J. Jensen, *A Users Guide to DMU: A Package for Analyzing Multivariate Mixed Models*, Manual, Faculty of Agricultural Science, University of Aarhus (2010). Available at http://dmu.agrsci.dk/dmuv6_guide.5.0.pdf.

[24] P. Madsen, Su, G., R. Labouriau and O.F. Christensen, *DMU - A Package for Analyzing Multivariate Mixed Models*, In: Proceedings CD - paper 732. Gesellschaft fur Tierzuchtwissenschaften e. V., (2010). Available at http://www.kongressband.de/wcgalp2010/assets/pdf/0732.pdf.

[25] R. P. Maia, L. S. Eikje, I. A. Boman, P. Madsen and R. Labouriau. *Genetic determination of sheep culling rates: a competing risk analysis.* In Preparation (2013).

[26] V.S. Pankratz, M. de Andrade, and T.M. Therneau, *Random-effects Cox proportional hazards model: general variance components methods for time-to-event data*, Genetic Epidemiology, 28(2) (2005), pp. 97–109.

[27] A. Ricard, and C. Blouin, *Genetic analysis of the longevity of French sport horses in jumping competition*, Journal of animal science, 89(10) (2011), 2988–94.

[28] S. Ripatti, and J. Palmgren, *Estimation of multivariate frailty models using penalized partial likelihood*, Biometrics, 56(4) (2000), pp. 1016–22.

[29] A.R. Roxström, V. Ducrocq, and E. Strandberg, *Survival analysis of longevity in dairy cattle on a lactation basis*, Genetics Selection Evolution, 35 (2003), pp. 305–318.

[30] T. Serenius and K.J. Stalder, *Selection for sow longevity*, Journal of Animal Science, 84 (2006), pp. 166–171.

[31] B.R. Southey and K.A. Leymaster, *Discrete time survival analysis of lamb mortality in a terminal sire composite population*, Journal of Animal Science, 81 (2003), pp. 1399–1405.

[32] B.R. Southey, R.V. Knox, J.F. Connor, J. F., Lowe and B. F. Roskamp, *The bioeconomic evaluation of sow longevity and profitability*, Journal of Animal Science, 81 (2003), pp. 2915–2922.

[33] J. Tarrs, J.P. Bidanel, A. Hofer, and V. Ducrocq, *Analysis of longevity and exterior traits on Large White sows in Switzerland*, Journal of animal science, 84(11) (2006), 2914–24.

[34] P.M. Visscher, W.G. Hill, and N.R. Wray *Heritability in the genomics era–concepts and misconceptions* Nature reviews. Genetics, 9(4) (2008), 255–66.

[35] L.D. Van Vleck ,*Stayability Evaluation as a Categorical Trait and by Considering Other Traits*, Journal of Dairy Science, 63(7) (1980), pp. 1172–1180, doi:10.3168/jds.S0022-0302(80)83063-3.

[36] R.W.M. Wedderburn, *Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method*, Biometrika 61(3) (1974), pp. 439–447.

[37] M.H. Yazdi, P.M. Visscher, V. Ducrocq, V. and R. Thompson, *Heritability, reliability of genetic evaluations and response to selection in proportional hazard models*, Journal of dairy science, 85(6) (2002), 1563–77.

A Technical details on the decomposition of the phenotypic variance

Let T be a non-negative random variable representing the observed survival time and D be an indicator variable taking the value 1 if a death is observed and 0 otherwise (censure). Denote by $\mathbf{X}(t) = \{X(s) : 0 \leq s < t\}$ the trajectory function of the known vector of explanatory variables until the time just before t. Additionally, we consider two independent gaussian random components, $\mathbf{U}$ and $\mathbf{V}$, representing the additive genetic effects and some environmental effects, respectively.

We define the death counting process $\mathbf{N} = \{N(t) : t \in \mathbb{R}_+ \text{ or } \mathbb{Z}_+\}$, where, for each time t, $N(t) = \mathbf{1}(T \leq t, D = 1)$. The increment of the death process at the time t is given by $dN(t) = N(t) - N(t-)$, where $N(t-) = \lim_{\Delta \downarrow 0} N(t - \Delta)$. The at risk process is given by $\mathbf{Y} = \{Y(t) : t \in \mathbb{R}_+ \text{ or } \mathbb{Z}_+\}$, where, for each time t, the random variable $Y(t)$ takes the value 1 when the individual is at risk at time t or 0 otherwise.

The death counting process and the at risk process are both adapted to the filtration $\mathbf{F} = \{\mathcal{F}_{t^-} : t \in \mathbb{R}_+ \text{ or } \mathbb{Z}_+\}$, where, for each time t, $\mathcal{F}_{t^-} = \sigma\{N(s), Y(s), X(s), 0 < s < t\}$ is the σ-algebra generated by $N(\cdot), Y(\cdot), X(\cdot)$ up to $t-$ (*i.e.* up to the time just before t) . Here, $\mathcal{F}_{t^-}$ represents what is known up to (but not including) time t.

Conditional on the random components $\mathbf{U}$ and $\mathbf{V}$, under the standard independent censoring assumption (see [3, 19]), we have that

$$P\left[dN(t) = 1 | \mathbf{U} = \mathbf{u}, \mathbf{V} = \mathbf{v}, \mathcal{F}_{t^-}\right] = Y(t)\lambda(t|\mathbf{u}, \mathbf{v}), \tag{25}$$

where $\lambda(t|\mathbf{u}, \mathbf{v}) = \lambda_0(t) \exp\left[X(t)'\beta + Z\mathbf{u} + W\mathbf{v}\right]$ and $Y(t)\lambda(t|\mathbf{u}, \mathbf{v})$ is the intensity associated to the counting process $\mathbf{N}$ at time t. Moreover, the cumulative intensity of $\mathbf{N}$ up to time t is $\Lambda(t|\mathbf{u}, \mathbf{v}) = \int_0^t Y(s)\lambda(s|\mathbf{u}, \mathbf{v})ds$. Note that $\mathbf{\Lambda} = \{\Lambda(t|\mathbf{u}, \mathbf{v}) : t \in \mathbb{R}_+ \text{ or } \mathbb{Z}_+\}$ is predictable with respect to the filtration $\mathbf{F}$. Define, for each time t,

$$M(t) = N(t) - \Lambda(t) \ . \tag{26}$$

The process $\mathbf{M} = \{M(t) : t \in \mathbb{R}_+ \text{ or } \mathbb{Z}_+\}$ is a martingale. Here, (26) is the Doob-Meyer decomposition of the death counting process where $\mathbf{\Lambda}$ is the compensator. Define, for each time t, the increment of the martingale $\mathbf{M}$ by $dM(t) = M(t) - M(t-)$, where $M(t-) = \lim_{\Delta \downarrow 0} M(t - \Delta)$. Since $\mathbf{M}$ is a martingale, $E[dM(t)|\mathcal{F}_{t^-}] = 0$ for each time t, which shows that $\mathbf{M}$ plays the role of the residuals of the model. The predictable variation process of the martingale $\mathbf{M}$ at the time t is

$$\langle M \rangle (t) = \int_0^t Var[dM(s)|\mathcal{F}_{s^-}]$$

and

$$d\langle M \rangle (t) = Var[dM(t)|\mathcal{F}_{t^-}] \ .$$

Now, we study the decomposition of the phenotypic variance related to the hazard function evaluated at a given time t. Define the conditional random variable

$$\mathcal{Y}(t) = dN(t)|\mathcal{F}_{t^-} \ , \tag{27}$$

which is equivalent to $\mathcal{Y}(t)$ informally defined by (12) in section 4.

Given the random components $\mathbf{U}$ and $\mathbf{V}$,

$$E\left[\mathcal{Y}(t)|\mathbf{U},\mathbf{V}\right] = Y(t)\lambda(t|\mathbf{u},\mathbf{v}) \tag{28}$$

and

$$Var\left[\mathcal{Y}(t)|\mathbf{U},\mathbf{V}\right] = \phi Var\left[dM(t)|\mathbf{U},\mathbf{V},\mathcal{F}_{t^-}\right], \tag{29}$$

where ϕ is the dispersion parameter. Using (28) and (29), the variance of $\mathcal{Y}(t)$ is given by

$$\begin{aligned} Var[\mathcal{Y}(t)] &= Var\left\{E\left[\mathcal{Y}(t)|\mathbf{U},\mathbf{V}\right]\right\} + E\left\{Var\left[\mathcal{Y}(t)|\mathbf{U},\mathbf{V}\right]\right\} \qquad (30)\\ &= Var\left\{Y(t)\lambda(t|\mathbf{U},\mathbf{V})\right\} + E\left\{\phi Var[dM(t)|\mathbf{U},\mathbf{V},\mathcal{F}_{t^-}]\right\} \ . \end{aligned}$$

A Taylor expansion argument (see [16], page 118) yields

$$\begin{aligned} E\left\{Var\left[\mathcal{Y}(t)|\mathbf{U},\mathbf{V}\right]\right\} &\approx Var\left[Y(t) \,|\, \mathbf{U}=\mathbf{0},\mathbf{V}=\mathbf{0}\right] \\ &= \phi Var\left[dM(t)|\mathbf{U}=\mathbf{0},\mathbf{V}=\mathbf{0},\mathcal{F}_{t^-}\right] \ . \qquad (31) \end{aligned}$$

Moreover, a second order Taylor expansion (see [16], page 118) yields

$$Var\left\{E\left[\mathcal{Y}(t)|\mathbf{U},\mathbf{V}\right]\right\} \approx \sigma_g^2 \left\{\frac{\partial}{\partial \mathbf{u}} E\left[\mathcal{Y}(t)|\mathbf{U},\mathbf{V}\right]\right\}^2_{(\mathbf{U}=\mathbf{0},\mathbf{V}=\mathbf{0})}$$

$$
\begin{aligned}
& +\sigma_e^2 \left\{ \frac{\partial}{\partial \mathbf{v}} E\left[\mathcal{Y}(t)|\mathbf{U},\mathbf{V}\right] \right\}^2_{(\mathbf{U}=\mathbf{0},\mathbf{V}=\mathbf{0})} \\
= \quad & \left[Y(t)\lambda^*(t)\right]^2 \left(\sigma_g^2 + \sigma_e^2\right) \ , \qquad (32)
\end{aligned}
$$

where $\mathbf{u}$ and $\mathbf{v}$ are realisations of the random components. Moreover, $\lambda^*(t) = \lambda(t|\mathbf{U}=\mathbf{0},\mathbf{V}=\mathbf{0})$.

In the continuous time case, for each $t \in \mathbb{R}_+$,

$$Var\left[dM(t)|\mathbf{U}=\mathbf{0},\mathbf{V}=\mathbf{0},\mathcal{F}_{t^-}\right] = Y(t)\lambda^*(t) \ ,$$

(see [19, 3]) and therefore, using (31) and (32) yields

$$Var\left[\mathcal{Y}(t)\right] \approx Y(t)\left\{\left[\lambda^*(t)\right]^2\left(\sigma_g^2+\sigma_e^2\right) + \phi\lambda^*(t)\right\} \ ,$$

which proves the decomposition of the phenotypic variance given in (13).

In the discrete time case, for each $t \in \mathbb{Z}_+$,

$$Var\left[dM(t)|\mathbf{U}=\mathbf{0},\mathbf{V}=\mathbf{0},\mathcal{F}_{t^-}\right] = Y(t)\lambda^*(t)\left[1-\lambda^*(t)\right] \ ,$$

(see [19, 3]) and therefore, using (31) and (32) yields

$$Var\left[\mathcal{Y}(t)\right] \approx Y(t)\left\{\left[\lambda^*(t)\right]^2\left(\sigma_g^2+\sigma_e^2\right) + \phi\lambda^*(t)\left[1-\lambda^*(t)\right]\right\} \ , \qquad (33)$$

which proves the decomposition of the phenotypic variance given in (16).

For the decomposition of the phenotypic variance related to the cumulative hazard function evaluated at a given time t we define the conditional random variable

$$\mathcal{Z}(t) = N(t)|\mathcal{F}_{t^-} \ .$$

This is equivalent to conditional random variable $\mathcal{Z}(t)$ informally defined in (20). Note that,

$$E\left[\mathcal{Z}(t)|\mathbf{U},\mathbf{V}\right] = \Lambda(t|\mathbf{U},\mathbf{V}) \text{ and } Var\left[\mathcal{Z}(t)|\mathbf{U},\mathbf{V}\right] = \langle M\rangle(t)|\mathbf{U},\mathbf{V} \ .$$

The decomposition of the variance of $\mathcal{Z}(t)$ is obtained by applying an analogous calculation as in the decomposition of the variance of $\mathcal{Y}(t)$. This yields the equations (21) and (23) for the decomposition of the phenotypic variance for the continuous time and the discrete time cases.