

ULRR

A non-PH accelerated hazard model for analyzing clinical trial data

Item Type	Meetings and Proceedings
Authors	Al-tawarah, Yasin;Mackenzie, Gilbert
Citation	Proceedings of the 18th International Workshop on Statistical Modelling;
Publisher	IWSM
Download date	2026-06-08 17:07:02
Item License	https://creativecommons.org/licenses/by-nc-sa/1.0/
Link to Item	https://hdl.handle.net/10344/2779

A non-PH Accelerated Hazard Model for Analyzing Clinical Trial Data

Yasin Al-Tawarah¹ & Gilbert MacKenzie¹

¹ Centre for Medical Statistics, Keele University, Keele, Staffordshire ST5 5BG, UK. E-mail: g.mackenzie@keele.ac.uk

Abstract

In longitudinal studies with a set of continuous or ordinal repeated response variables it may be convenient to summarize the outcome as a threshold event. Then, the time to this event becomes of interest. In this paper we obtain the general likelihood for the unknown parameters when the underlying survival model is parametric and the survival times are interval-censored. We investigate the use of a member of the Generalized Time Dependent Logistic family of survival distributions (MacKenzie, 1996) which is a non-PH Accelerated Hazard Model and has a logistic baseline hazard function. We use simulation to investigate how inference on the treatment parameter is compromised by using the mis-specified likelihood, which treats the interval-censored survival times as if they were exact.

Keywords: interval censoring, logistic survival, non-PH model, accelerated hazard, mis-specified likelihood

1 Introduction

In classical survival analysis, the exact time to event is usually known. However, in longitudinal clinical trials where outcome is a continuous or ordinal variable measured repeatedly at scheduled follow-up times, the exact time-to-event may be unknown. Such situations arise when the outcome is classified according to threshold of clinical interest. Then scientific interest is focused on the time at which the threshold is crossed. In these studies recruitment is staggered in time and, increasingly, survival-type methods (Kaplan Meier, 1958; Peto & Peto, 1972 and Cox, 1972) are being pressed into service.

These methods are appropriate for right censored 'time to event data' when the exact time of occurrence is known, but strictly inappropriate when the 'time to event' is known only to lie in an interval. Application of conventional methods to interval 'end' or 'mid' points can lead to bias (Lindsey

and Ryan, 1998) and optimistic precision (MacKenzie, 1999). Here, we develop the parametric accelerated life (AL) logistic model (MacKenzie, 1996) in which the baseline hazard follows the time-dependent logistic (TDL) survival model. We compare inference from the correct model with that from the mis-specified model which uses follow-up times as if they were exact.

2 Likelihood Formulation

Suppose there are $m + 1$ scheduled inspection times, $t_0^+, t_1^+, \dots, t_m^+$ at which continuous or ordinal responses Y_0, Y_1, \dots, Y_m are measured. Let T be a non negative variable denoting the time to some outcome of interest defined on the Y s. Let $S(t; \theta)$ and $\lambda(t; \theta)$ be the corresponding survival and hazards functions, respectively, depending on the unknown vector parameter $\theta \in \Theta$, where $\theta = (\alpha', \gamma', \beta)'$. Then for a sample of n independent subjects it may be shown that the true censored likelihood for the unknown parameters is:

$$L_1(\theta) = \prod_{i=1}^n \left\{ S(t_{i_{(k-1)}}; \theta) [1 - S(t_{i_{(k-1)}}, t_{i_k}; \theta)] \right\}^{\delta_i} [S(t_i^*; \theta)]^{1-\delta_i} \quad (1)$$

where typically n_k patients fail between scheduled examination times $t_{(k-1)}^+$ and t_k^+ for $k = 1, \dots, m$ and n_c patients are censored or withdrawn at specific times, t_i^* , such that $n_c + \sum_{k=1}^m n_k = n$. Here, $\delta_i = 1$ denotes an event and $\delta_i = 0$ denotes a censored observation. We may compare (1) with the mis-specified censored likelihood resulting from treating the observed inspection times as if they were exact:

$$L_2(\theta) = \prod_{i=1}^n [\lambda(t_{i_k}; \theta) S(t_{i_k}; \theta)]^{\delta_i} [S(t_{i_k}; \theta)]^{1-\delta_i} \quad (2)$$

Equations (1) and (2) enable us to investigate the effect of mis-specification for any survival model where the function takes closed form. Notice the use of observed inspection times rather than the scheduled times in equations (1) and (2).

3 Model Formulation

MacKenzie's (1996) AL logistic survival model is defined by the hazard function:

$$\lambda(t; x) = \frac{\lambda \exp(tx' \beta + \gamma)}{1 + \exp(tx' \beta + \gamma)} \quad (3)$$

a form which we have modified to make it a strictly AL logistic model in the hazard rather than in the survival function and where we have suppressed

the dependence on θ .

$$\lambda(t; x) = \frac{\lambda \exp(t\alpha\phi)}{1 + \exp(t\alpha\phi)} \quad (4)$$

where $\phi = \exp(x'\beta)$. The corresponding baseline hazard function is:

$$\lambda_0(t) = \frac{\lambda \exp(t\alpha)}{1 + \exp(t\alpha)} \quad (5)$$

and on integrating we obtain the baseline survival function

$$S_0(t) = \left[\frac{1 + \exp(t\alpha)}{2} \right]^{-\lambda/\alpha} \quad (6)$$

We may regard this form as a general survivor function and accelerate it in the classical way (Lawless, 1982), namely:

$$S(t|x) = \left[\frac{1 + \exp(t\alpha\phi)}{2} \right]^{-\lambda/\alpha} \quad (7)$$

whence the corresponding hazard function is:

$$\lambda(t|x) = \lambda\phi \frac{\exp(t\alpha\phi)}{1 + \exp(t\alpha\phi)} \quad (8)$$

a form which is different from the accelerated hazard logistic (AH) model given by (4).

4 Simulation Study

The object of the simulation study is to quantify the degree to which inference about the parameters in the AH & AL models, especially β , is compromised by the use of the mis-specified likelihood. We investigate the 2-sample case, mimicking a RCT in which scientific interest is focused on estimating the treatment effect and its associated standard error. The simulation parameters include: sample size, percentage censored, patterns of follow-up examination is regularly and irregularly spaced, the model parameters (θ). The maximum likelihood estimates will be calculated using the correct and the mis-specified likelihoods.

5 Results

First we compared models (4) and (8) using lung cancer data, and present the conditional fits obtained by each regression model and the marginal

fit of the Kaplan Meier estimator. The (AH) model shows a better fit compared with the (AL) model (Figures 1,2).

Second, we report a subset of the complete simulation using mid-points in the mis-specified likelihood. Tables (1 & 2) show the MLE's for the three parameters using a regular visit schedule. Note that we report $\phi^* = \log_e(\lambda)$ in the tables. Overall, the true likelihood provided consistently better estimates with superiority for the AH model compared with the AL model, when we allowed for drop-out and using a regular schedule. The mis-specified likelihood also produced standard errors which were artificially precise.

6 Summary

The idea of an accelerated hazard model is new. To our knowledge this is the first time that they have been described, and compared empirically with classical accelerated life models, albeit in the context of a single family of survival models - the GTDL (MacKenzie, 1996). The results of the numerical analysis favour the AH model suggesting that the model may be useful in practice. The advantages of these parametric models stem from the closed forms taken by survivor functions and the fact that when $\beta = 0$ the underlying survival functions have testable parametric forms. We have demonstrated by simulation, the use of these two models in the analysis of interval censored survival data arising in longitudinal randomized controlled trials.

7 Key References

- Kaplan Meier (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53,457-481.
- Cox DR (1972). Regression models and life tables (with discussion) *JRSS B*. 34, 187-220.
- MacKenzie G (1996). Regression models for survival data: the generalized time-dependent logistic family. *The Statistician*, 45,No.1,pp.21-34.
- MacKenzie G (1999). Survival analysis for longitudinal data. 14th International Workshop on Statistical Modelling. Graz,Austria. 259-264.

Table 1: Comparison of Mis-specified and True Models Estimators
 AL Model, Mid-point
 Regular follow up (3,6,9,12,15,18,21,24)

		Mis-specified			True		
	n	$\hat{\phi}^*$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\phi}^*$	$\hat{\alpha}$	$\hat{\beta}$
$\phi^* = -0.6, \alpha = 0.2, \beta = -0.5, \% \text{ within censoring} = 0$							
Mean	100	-1.007	1.901	-0.451	-0.601	0.251	-0.511
(se.)		(0.160)	(0.250)	(0.161)	(0.215)	(0.201)	(0.178)
Mean	500	-1.052	1.976	-0.438	-0.626	0.201	-0.501
(se.)		(0.053)	(0.120)	(0.071)	(0.131)	(0.135)	(0.078)
$\phi^* = -0.6, \alpha = 0.2, \beta = -0.5, \% \text{ within censoring} = 30$							
Mean	100	-1.346	1.882	-0.443	-1.060	0.371	-0.479
(se.)		(0.175)	(0.294)	(0.188)	(0.207)	(0.306)	(0.205)
Mean	500	-1.403	1.962	-0.442	-1.127	0.531	-0.481
(se.)		(0.061)	(0.168)	(0.089)	(0.151)	(0.385)	(0.091)

Table 2: Comparison of Mis-specified and True Models Estimators
 AH Model, Mid-point
 Regular follow up (3,6,9,12,15,18,21,24)

		Mis-specified			True		
	n	$\hat{\phi}^*$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\phi}^*$	$\hat{\alpha}$	$\hat{\beta}$
$\phi^* = -0.6, \alpha = 0.2, \beta = -0.5, \% \text{ within censoring} = 0$							
Mean	100	-1.020	1.618	-0.034	-0.624	0.253	-0.552
(se.)		(0.094)	(0.203)	(0.221)	(0.180)	(0.211)	(1.112)
Mean	500	-1.053	1.742	-0.045	-0.635	0.225	-0.538
(se.)		(0.040)	(0.103)	(0.104)	(0.129)	(0.125)	(0.588)
$\phi^* = -0.6, \alpha = 0.2, \beta = -0.5, \% \text{ within censoring} = 30$							
Mean	100	-1.024	1.688	-0.024	-0.632	0.248	-0.393
(se.)		(0.088)	(0.211)	(0.218)	(0.197)	(0.199)	(0.1.206)
Mean	500	-1.055	1.744	-0.045	-0.634	0.212	-0.530
(se.)		(0.040)	(0.108)	(0.099)	(0.125)	(0.109)	(0.584)

FIGURE 1. Predicted AL v KM

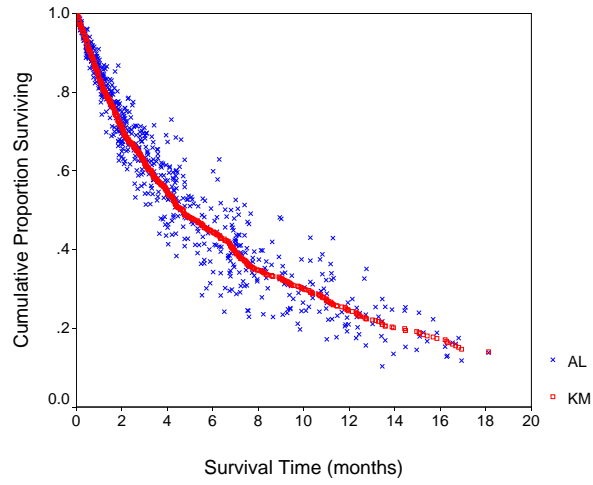


FIGURE 2 Predicted AH v KM

