Non-PH Survival Modelling with Frailty and Structured Dispersion

Joseph Lynch 1, Gilbert MacKenzie 1

1 Centre for Biostatistics, University of Limerick, Ireland

Abstract

In the Cox proportional hazard (PH) regression model, the hazard function is given by \( \lambda(t; x) = \lambda_0(t) \exp(x'\beta) \), where the unspecified form of the baseline hazard, \( \lambda_0(t) \), means that the data essentially determine the shape of the function. However, in parametric models the baseline hazard is specified and we focus on the Generalised Time-Dependent logistic distribution (GTDL) which is a non-PH model in which the hazard function given by

\[
\lambda(t; x) = \lambda_0 \frac{\exp(t\alpha + x'\beta)}{1 + \exp(t\alpha + x'\beta)}
\]

where \( \alpha \) is a time parameter, \( \beta \) is a \( p \times 1 \) vector of regression parameters associated with fixed covariates, \( x' = (x_1, \ldots, x_p) \) and \( \lambda > 0 \) is a scalar. In this model the \( \beta \)s measure the linear influence of the covariates on a generalised log odds scale, rather than on the log-linear scale, as in Cox’s PH model. The time dependent relative risk, \( \rho(t) \), the ratio of hazard rates for two subjects with different covariate vectors, \( x_1 \) and \( x_2 \), is given by

\[
\rho(t; x_1, x_2) = \lambda(t; x_1)/\lambda(t; x_2) = \exp\{(x_1 - x_2)'\beta\} \psi(t; x_1, x_2),
\]

(2)

The leading term on the right hand side of (2) is Cox’s constant of proportionality (the RR in a PH model) and in the GTDL model this constant is moderated by \( \psi(t; \cdot) \), a function of both time and covariates, thus generalising the basic Cox model.

Further generalisations via frailty are currently being developed. Firstly, we extend the GTDL model to include a univariate multiplicative frailty component \( u_i \sim \Gamma(1, \sigma^2) \). Next we develop a structured dispersion version of the model (Lee & Nelder, 2006) in which the frailty variance \( \sigma^2 = \exp(\omega_i) \) is not constant, but is governed by baseline covariates according to another regression model, \( \omega_i = x_i'\beta^* \) where the covariates may or may not be the same as in the hazard regression model. Finally, for the purposes of comparison, we develop a PH structured dispersion model based on the Weibull.

We illustrate our various findings using a dataset of over 15,000 women diagnosed with breast cancer between 1991 and 1995 in the West Midlands of the U.K., comparing and contrasting survival functions in the various models fitted.

References

