NON-IGNORABLE MISSING COVARIATE DATA IN PARAMETRIC SURVIVAL MODELS

K.L.Boyd[†], J.L.Hutton

University of Warwick, UK

[†] E-mail: *k.l.boyd@warwick.ac.uk*

Within medical settings, it is common to face incomplete covariate data. In 1976, Rubin [1] established a framework for dealing with this missing data based upon the mechanism acting upon it. Many methods for dealing with missing data require the slightly restrictive assumption of MAR. Our motivation is a data set from Bristol looking at the survival of individuals with cerebral palsy. We question the validity of making this MAR assumption in this case. We therefore propose a fully parametric joint model for survival and the missing data mechanism based upon a latent variable which indicates the observation, or lack of observation, of the covariate of interest. This model framework is based upon methods used by Copas and Shi [2] to consider publication bias in meta-analysis and can incorporate any of the commonly used survival distributions. Application to the Bristol data is discussed and the inclusion of left-truncated survival times investigated. Results suggest that naive analysis of this data produces biased survival estimates. Simulation studies show that this method is consistent with basic methods applicable when data is MCAR. We also consider the model's performance over various levels of censoring and missing data.

[1] Rubin, D.B., 'Inference and missing data (with discussion)', *Biometrika*, 1976, 63, 581-592.
[2] Copas, J.B. and Shi, J.Q., 'A sensitivity analysis for publication bias in systematic reviews', *Statistical Methods in Medical Research*, 2001, 10, 251-265.