Survival Models: Correlated Data and Non-parametric Estimation

Chair: Julie Atherton (McGill University)

HENRIK STRYHN, Atlantic Veterinary College, University of Prince Edward Island

Estimating Correlated Frailties in a Cox Model

The inclusion of random effects (frailties) in survival analysis has become a standard approach to account for hierarchical data structures. It is less common to include random regression coefficients ("random slopes"), perhaps due to both the computational complexity and the scarcity of datasets that support such modelling. Using four estimation methods available in standard software and previously published data from veterinary epidemiology, with animals clustered in herds, we demonstrate how the inclusion of herd-level correlated frailty terms may substantially affect fixed effect estimates and inferences. An additional simulation study provided a thought-provoking comparison of the performance of the four methods.

YI NIU, Queen’s University

A Marginal Mixture Cure Model for Clustered Survival Data

Marginal models have been widely used to analyze correlated survival data. Most of the studies assume that the event of interest will eventually occur given sufficiently long follow-up. We consider a marginal mixture cure model for clustered failure time data with a possible surviving fraction, and propose novel generalized estimating equations to incorporate the correlation within clusters in the marginal model. A simulation study demonstrates the substantial efficiency gain over the conventional EM method. The model and the proposed method are applied to a data set of failure times from a cancer study.

HEDY JIANG, Cancer Care Ontario and McMaster University

Survival Models for A Large Cohort Study in Pickering

In population-based cancer incidence studies, the majority of people are not affected by cancer. The cancer-free fraction is of interest and a useful measure to monitor trends in cancer occurrences. In a Pickering cohort study regarding the risk of radiation sensitive cancers in relation to residential proximity to a large nuclear plant, we propose a cure rate model with spatial frailties to estimate the cancer-free fraction and the spatial dependence of census tracts. A Gaussian random field with a Matern correlation function is added to the baseline hazard to model the intracluster correlation in census tracts.

DAVID E. MATTHEWS, University of Waterloo

Exact Nonparametric Likelihood Confidence Bands for the Survivor Function

A method to produce exact simultaneous confidence bands for the empirical cdf that was first described in Owen (1995), and later modified by Jager and Wellner, is the starting point for deriving simultaneous confidence bands for the survivor function, \( F(\cdot) \), of any positive random variable. We invert a nonparametric likelihood test of uniformity to obtain simultaneous lower and upper bands for \( F(\cdot) \) with global confidence level \( 1 - \alpha \). Noé’s recursion provides the computational engine. Various aspects of these exact bands are investigated, using as an illustration survival times for non-Hodgkin’s lymphoma patients with Stage 4, advanced disease.

MOHAMMAD DEHGHAN, Laval University

Nonparametric Estimation of the Conditional Survival Function with Time-Varying Covariate and Interval Censoring

In this talk we propose a nonparametric approach to estimate the conditional survival function of failure time given a time-varying covariate, \( S(t|z(y):0 \leq y \leq t) \), when time is subject to interval censoring. In this case, observations appear as \( \{(L_i, R_i, Z(t_{ij})): j = 1, \ldots, k_i; i = 1, \ldots, n\} \). We assume that \( Z(t_{ij}) = \alpha_i + \beta_i t_{ij} + \varepsilon_{ij}; j = 1, \ldots, k_i \), where \( \alpha_i \) and \( \beta_i \)
are random effects. We treat the estimate of the slope $\beta_i$ of the $i$-th path as a positive covariate. We propose a weighted average estimator based on generalized Turnbull and Kaplan-Meier estimators to estimate $P(T_i > t | z(y_i); 0 \leq y_i \leq t)$, $P(T_i > t | T_i > t_0, z(y_i); 0 \leq y_i \leq t)$ and the $\gamma$-th quantile of the distribution of $\{T_i | z_i(y_i); 0 \leq y_i \leq t\}$, $q_{\gamma i}$.

ZHIHUI (AMY) LIU, McGill University

*Recovering the Raw Data behind a Kaplan-Meier or Nelson-Aalen Survival Curve*

It is preferable but not always possible to obtain from the authors the raw data that were used to create published Kaplan-Meier and Nelson-Aalen survival curves. Using examples, we show that much of the original information can be recovered from the curve itself, by measuring the lengths of the vertical steps – by digitization or accessing the postscript files. These reconstruction techniques are useful for retrieving the raw data when there is no censoring; when either the numbers at risk or the numbers of deaths at various follow-up times are given; and when neither of them are available.