A BAYESIAN ANALYSIS OF CORRELATED INTERVAL-CENSORED DATA

Khangelani Zuma

Biostatistics & Epidemiology Unit

Social Aspects of HIV/AIDS & Health

Human Sciences Research Council

PRETORIA

South Africa

Introduction

- Interval-censored data arise where the exact event time is not observed directly.
- For example, HIV seroconversion time.
- HIV diagnosis is delayed compared to HIV transmission.
- Infection can happen before the first exam time or might not have happened at last examination (left, right censoring/truncation).
- In such data, we only know the time interval within which infection occurred but not the event time

Interval Censored Data

- Methods of analysing interval-censored data stem from the Cox proportional hazards model (Cox, 1972).
- Finkelstein (1986) generalised the Cox proportional hazards model.
- Interval-censored data already complicated.
- Dependency presents even more challenges.
- Arise from sampling method used (multistage sampling design).
- In cluster sampling, individuals belonging to the same cluster share the same unobserved cluster-specific effect (frailty) and thus making them positively correlated.

Notation

- Let y_{ij} $(i = 1, \cdots, I; j = 1, \cdots, J_i)$ denote the event time for the jth individual in cluster i.
- Interval-censoring means the exact event time t_{ij} is unobserved.
- Only clinical examination endpoints $v_{ij} = \{v_{ij,1}; v_{ij,2}\}$ encompassing interval-censored event time t_{ij} are observed.
- This means $y_{ij} = t_{ij}$ if event occurred and $y_{ij} = v_{ij,1}$ if the observation is right-censored.
- Let censoring indicator $\delta_{ij} = 1$ if event occurred and 0 otherwise.
- The ith cluster specific frailty is denoted by w_i.

Model Formulation

- Multiplicative frailty model with Weibull baseline hazards is assumed, $\lambda_0(y_{ij}) = \gamma\alpha y_{ij}^{\alpha-1}$.

- Conditional on w_i, survival times are mutually independent with conditional hazards distribution given by

$$h(y_{ij}|x_{ij}, w_i) \quad = \quad \alpha\gamma w_i y_{ij}^{\alpha-1}\theta_{ij}$$

  where $\theta_{ij} = \exp(x'_{ij}\beta)$.

- Those who experienced an event contribute to the likelihood the product of their conditional hazards and conditional survival function whilst those who were right-censored contribute only the conditional survival function.

Complete-data Likelihood

- Using conditional independence between y_i given w_i the complete-data likelihood contribution of cluster i is

$$
\begin{aligned}
& L_i(w_i, v_i, t_i|\psi) \\
= \quad & f(w_i|\eta) \prod_{j=1}^{J_i} [S(t_{ij}|x_{ij}, w_i) h(t_{ij}|x_{ij}, w_i)]^{\delta_{ij}} [S(v_{ij,1}|x_{ij}, w_i)]^{1-\delta_{ij}}
\end{aligned}
$$

where $f(w_i|\eta)$ is a gamma density specified by the shape parameter η and scale parameter η^{-1} and $\psi = \{\eta, \alpha, \gamma, \beta\}$.

Full Bayesian Inference

- Specify priors for hyperparameter η, baseline hazard parameters α and γ, and fixed effects β. Prior for β is assumed MVN($\mathbf{d_0} = \mathbf{0}$, $\mathbf{\Sigma_0} = \upsilon_0 \mathbf{I}$), where υ_0 is a suitably chosen large number.

- A Gamma (ρ_1, ρ_2) prior for η where ρ_1 and ρ_2 are suitably chosen constants is specified. A priori, it is also common to specify Gamma (a_1, a_2) prior and Gamma (b_1, b_2) prior for α and γ, respectively.

- The joint distribution, $f(\text{data}, \beta, \alpha, \gamma, t_{ij}, w_i, \eta)$, of all parameters, hyperparameters and the data for our model is given by

$$f(\beta) f(\alpha) f(\gamma) f(\eta) \times \left\{ \prod_{i=1}^{I} f(w_i|\eta) \prod_{j=1}^{J_i} L(y_{ij}|\beta, \alpha, \gamma, w_i) \right\}.$$

Full Bayesian Inference

- Requires joint posterior distribution of all parameters given the observed data.
- Difficult or impractical to obtain analytically.
- To avoid evaluating high dimensional integrals we use MCMC methods, specifically Gibbs sampler to generate samples from the posterior.
- The algorithm finds Markov chain that has the joint posterior as its long-run distribution.

Gibbs Conditional Distributions

- Cluster specific random effect: $f(w_i|\text{data}, \beta, \alpha, \gamma, t_{ij}, \eta)$

$$\propto \quad w_i^{\sum_{j=1}^{J_i} \delta_{ij} + \eta - 1} e^{-w_i [\eta + \sum_{j=1}^{J_i} \delta_{ij} \Lambda(t_{ij}|x_{ij}) + \sum_{j=1}^{J_i} (1-\delta_{ij}) \Lambda(v_{ij,1}|x_{ij})]}$$

kernel of gamma with shape $\sum_{j=1}^{J_i} \delta_{ij} + \eta$ and inverse scale parameters $\eta + \sum_{j=1}^{J_i} \delta_{ij} \Lambda(t_{ij}|x_{ij}) + \sum_{j=1}^{J_i} (1 - \delta_{ij}) \Lambda(v_{ij,1}|x_{ij})$.

- Random effect inverse variance: $f(\eta|\text{data}, \beta, \alpha, \gamma, w_i)$

$$\propto \quad \eta^{\rho_1 - 1} \times \left(\frac{\eta^\eta}{\Gamma(\eta)} \right)^I \left(\prod_{i=1}^{I} w_i \right)^{\eta - 1} e^{-\eta \left[\rho_2 + \sum_{i=1}^{I} w_i\right]}.$$

This is a simple *log-concave* distribution in η and can be sampled efficiently using the *adaptive-rejection sampling* scheme

Baseline Scale and Shape Parameters

- The scale parameter of the baseline hazard: $f(\gamma|\text{data}, \beta, \alpha, \eta, t_{ij}, w_i)$

$$\propto \quad \gamma^{\sum_{i=1}^{I}\sum_{j=1}^{J_i}\delta_{ij}+b_1-1} e^{-\gamma[b_2+\sum_{i=1}^{I}\sum_{j=1}^{J_i}\delta_{ij}\phi(t_{ij})+\sum_{i=1}^{I}\sum_{j=1}^{J_i}(1-\delta_{ij})\phi(v_{ij,1})]}$$

  where $\phi(t_{ij}) = w_i t_{ij}^{\alpha}\theta_{ij}$ and $\phi(v_{ij,1}) = w_i v_{ij,1}^{\alpha}\theta_{ij}$. This conditional is also recognized as the *kernel* of a gamma distribution.

- The shape parameter of the baseline hazard: $f(\alpha|\text{data}, \beta, \gamma, \eta, t_{ij}, w_i)$

$$\begin{aligned}\propto \quad & \alpha^{\sum_{i=1}^{I}\sum_{j=1}^{J_i}\delta_{ij}+a_1-1} \times e^{-[\alpha a_2+\sum_{i=1}^{I}\sum_{j=1}^{J_i}\delta_{ij}H(t_{ij}|x_{ij},w_i)]} \\ \times \quad & +e^{[\sum_{i=1}^{I}\sum_{j=1}^{J_i}(1-\delta_{ij})H(v_{ij,1}|x_{ij},w_i)]} \times (\prod_{i=1}^{I}\prod_{j=1}^{J_i} t_{ij}^{\delta_{ij}})^{\alpha-1}\end{aligned}$$

  which does not simplify to any standard distribution.

Gibbs Conditional for Interval Censored Event Time

- The unobserved interval-censored event time:

$$f(t_{ij}|v_{ij,1} < t_{ij} \leq v_{ij,2}, \text{data}, \beta, \alpha, \gamma, \eta, w_i)$$

$$\begin{aligned} &= \frac{f(t_{ij}|\text{data}, \beta, \alpha, \gamma, \eta, w_i)}{\int_{v_{ij,1}}^{v_{ij,2}} f(t_{ij}|\text{data}, \beta, \alpha, \gamma, \eta, w_i) dt_{ij}} \\ &= \frac{\exp(-w_i \gamma t_{ij}^{\alpha} \theta_{ij}) \times w_i \alpha \gamma t_{ij}^{\alpha-1} \theta_{ij}}{S(v_{ij,1}|x_{ij}, w_i) - S(v_{ij,2}|x_{ij}, w_i)} \\ &\propto \exp(-w_i \gamma t_{ij}^{\alpha} \theta_{ij}) \times w_i \alpha \gamma t_{ij}^{\alpha-1} \theta_{ij} \end{aligned}$$

which we recognize as the *kernel* of a Weibull distribution with shape parameter α and scale parameter $\gamma w_i \theta_{ij}$. The node can also be sampled directly on condition that the sampled value $t_{ij} \in (v_{ij,1}, v_{ij,2}]$.

Fixed Effects Gibbs Conditional

- $f(\beta|$,data$, \alpha, \gamma, \eta, t_{ij}, w_i)$, does not simplify.
- If the improper prior $f(\beta) \propto 1$ is assumed, the posterior mode will be proportional to the maximum likelihood (ML) estimate for β.
- A modified EM algorithm can be used to compute the ML estimates (Klein, 1992) with event times and frailties fixed and known.
- The normal approximation, with mean and covariance matrix equal to the mode and inverse of the Fisher information matrix obtained from the ML estimation, can be used as the conditional density.
- The β's are then generated by inserting a Metropolis step using the normal approximation as the candidate generating density.

Model Implementation

- The model was implemented on the data from migrant men and their female sexual partners, and non-migrant men and their female sexual partners (Lurie, *et al* 2003a; 2003b).
- Acceptance rate for candidate β's was about 54%, which was well within 30% and 70%, the recommended acceptance rate (Raftery and Lewis, 1996).
- Fixed effects sampling scheme involves an EM estimation of ML estimates and calculation of Fisher information for the proposal density in the Metropolis step.
- This can be computationally intensive.

Recommendation

- It is recommended to modify the iteration scheme to iterate through $w_i, t_{ij}, \alpha, \gamma$ and η several times for each draw of β.
- The modification improves the efficiency.
- Equal sampling of these nodes leads to successive simulated values of $w_i, t_{ij}, \eta, \gamma$ and α that are highly autocorrelated while those for β are nearly nearly independent.

Conclusion

- Advents of MCMC make possible the use of flexible Bayesian models to analyse correlated interval-censored data.

- Simplify to two iterative steps involving *imputation step* which draws $w_i^{(r)}$ and $t_{ij}^{(r)}$ from their respective conditional predictive distributions $f(w_i|\text{data}, \psi, t_{ij}^{(r-1)})$ and $f(t_{ij}|v_{ij,1} < t_{ij} \leq v_{ij,2}, \text{data}, \psi, w_i^{(r)})$, and a

- *posterior step* which draws $\psi^{(r)}$ from conditional posterior distribution $f(\psi|\text{data}, w_i^{(r)}, t_{ij}^{(r)})$.

- The two iterative steps can be viewed as the stochastic counterparts to the E-step and M-step of the EM algorithm.