Frailty model approach for the clustered interval-censored data with informative censoring

Jinheum Kim\(^1\)
joint work with Youn Nam Kim\(^2\)

\(^1\)Department of Applied Statistics, University of Suwon
\(^2\)Clinical Trials Center Severance Hospital, Younsei University Health System

August 9, 2015
Outline

- Introduction: interval-censored data, informative censoring
- Model specification
- Likelihood construction
- Parameter estimation procedure
- Simulations
- Illustrative real example
- Concluding remarks
What is interval-censored survival data?

- In clinical trials with periodic follow-up, each subject is observed through several examinations. However, a subject may skip one or more pre-appointed visits and then return with the failure already occurred. In these situations, the true event time of interest lies in an interval of the form

\[(L, R],\]

where \(L\) is the last time seen without disease, and \(R\) is the first time the subject appeared with disease.

- So, a subject with \(R = \infty\) is right-censored at \(L\). On the contrary, a subject with \(L = 0\) is left-censored at \(R\).
In what situations may informative censoring occur?

- Most existing methodologies with regression analysis were developed under the assumption of ‘non-informative censoring’ mechanism (Zhang et al., 2005). The failure time and visiting times of subjects are frequently assumed to be independent.

- However, in some situations, this assumption does not hold. For instance, when failure occurs, a patient could experience some symptoms prior to or together with failure. This makes the patient tend to visit the doctor earlier than scheduled (Zhang et al., 2007; Wang et al., 2010).
How do we deal with informative censorung?

- However, it is virtually impossible to observe both the failure time and the censoring times simultaneously. Subsequently, it is not possible to test the dependence or independence assumption of the censoring mechanism.
- One remedy to circumvent these difficulties is to impose extra assumptions or modelling.
Review on related works

- Huang & Wolfe (2002) have dealt with the clustered right-censored data assuming the dependence between the failure time and the censoring time.
- Zhang et al. (2005, 2007) and Wang et al. (2010) have utilized frailty models to explain a dependence structure between the failure time and the censoring times for the interval-censored data with informative censoring.
- Kim & Kim (2012) proposed an estimating procedure using the Cox PH model with a shared frailty for the clustered interval-censored data under the non-informative censoring assumption.
- In this talk, we extend the arguments of Huang & Wolfe (2002) and Kim & Kim (2014) to the clustered interval-censored data in the presence of informative censoring.
Notation

- $T_{ij}$: the failure time for the $j^{th}$ subject within the $i^{th}$ cluster ($i = 1, \ldots, n; j = 1, \ldots, n_i$)
- $U_{ij}, V_{ij}$: two observation times with $U_{ij} \leq V_{ij}$
  - Although we cannot observe the exact failure time $T_{ij}$, it is only less than or equal to $U_{ij}$, between $U_{ij}$ and $V_{ij}$, or greater than $V_{ij}$
- $W_{ij}$: the gap time defined as $W_{ij} = V_{ij} - U_{ij}$ if $V_{ij}$ is available; otherwise $W_{ij} = \infty$
- $\delta_{1ij} = I(T_{ij} \leq U_{ij})$ and $\delta_{2ij} = I(U_{ij} < T_{ij} \leq V_{ij})$
- $x_{ij}$: a $p \times 1$ vector of covariates
- So, the observed data for the $j^{th}$ subject within the $i^{th}$ cluster have the form of
  $$o_{ij} = (U_{ij}, V_{ij}, \delta_{1ij}, \delta_{2ij}, x_{ij}').$$

Subsequently, $o = (o'_1, \ldots, o'_n)'$, where $o_i = (o'_{i1}, \ldots, o'_{in_i})'$
Proposed models

- Assume that $T_{ij}'$'s within the $i$th cluster share an unobservable frailty $r_i$ and conditional on $x_{ij}$ and $r_i$, they are independent
  - $r_i$ : a normal frailty with a mean of 0 and variance $\theta$
- To incorporate the informative censoring, consider Cox PH models with a shared frailty for $T_{ij}$, $U_{ij}$, and $W_{ij}$, respectively:
  
  \[
  \lambda_t(t|x_{ij}, r_i) = \lambda_{0t}(t)\exp\{\beta'_t x_{ij} + r_i\},
  \]  
  \[
  \lambda_u(t|x_{ij}, r_i) = \lambda_{0u}(t)\exp\{\beta'_u x_{ij} + \alpha_u r_i\},
  \]  
  \[
  \lambda_w(t|x_{ij}, r_i) = \lambda_{0w}(t)\exp\{\beta'_w x_{ij} + \alpha_w r_i\},
  \]

where $\beta_t$, $\beta_u$, and $\beta_w$ are the regression coefficients, $\lambda_{0t}(\cdot)$, $\lambda_{0u}(\cdot)$, and $\lambda_{0w}(\cdot)$ are the baseline hazard functions for $T_{ij}$, $U_{ij}$, and $W_{ij}$, respectively, and $\alpha_u$ and $\alpha_w$ are unknown parameters representing the degree of dependency between $T_{ij}$ and $U_{ij}$ and between $T_{ij}$ and $W_{ij}$, respectively.

- Assume that $T_{ij}$, $U_{ij}$, and $W_{ij}$ are conditionally independent given $x_{ij}$ and $r_i$.
Schematic diagram

Starting time = 0 → End of study

1. U = \infty, W = \infty, T = no information
2. U = \infty, W = \infty, T = no information
3. U = \infty, W = \infty, T = no information
4. U = u, W = \infty, T < u, left-censored
5. U = u, W = \infty, T > u, right-censored
6. U = u, W = \infty, T > u, right-censored
7. U = u, W = \infty, (ignored)
8. U = u, W = v - u, T > v, interval-censored
9. U = u, W = \infty, T > v, right-censored
Likelihood construction

- Given $x_{ij}$ and $r_i$, the likelihood function $L_{ij}$ for the $j^{th}$ subject within the $i^{th}$ cluster can be expressed as follows:
  - when $T_{ij}$ is left-censored at $u_{ij}$ and $W_{ij}$ is right-censored at 0,
    \[
    L_{ij} = P(U_{ij} = u_{ij}|x_{ij}, r_i)P(T_{ij} \in (0, u_{ij}]|x_{ij}, r_i);
    \]
  - when $T_{ij}$ is interval-censored in $(u_{ij}, v_{ij}]$ but $W_{ij}$ is exactly observed as $(v_{ij} - u_{ij})$,
    \[
    L_{ij} = P(U_{ij} = u_{ij}|x_{ij}, r_i)P(T_{ij} \in (u_{ij}, v_{ij}]|x_{ij}, r_i)P(W_{ij} = v_{ij} - u_{ij}|x_{ij}, r_i);
    \]
  - when both $T_{ij}$ and $W_{ij}$ are right-censored at $u_{ij}$ and 0, respectively,
    \[
    L_{ij} = P(U_{ij} = u_{ij}|x_{ij}, r_i)P(T_{ij} \in (u_{ij}, \infty)|x_{ij}, r_i),
    \]
  - when $T_{ij}$ is right-censored at $v_{ij}$ but $W_{ij}$ is exactly observed as $(v_{ij} - u_{ij})$,
    \[
    L_{ij} = P(U_{ij} = u_{ij}|x_{ij}, r_i)P(T_{ij} \in (v_{ij}, \infty)|x_{ij}, r_i)P(W_{ij} = v_{ij} - u_{ij}|x_{ij}, r_i).\]
Thus, given $x_{ij}$ and $r_i$, the conditional likelihood for the $j^{th}$ subject within the $i^{th}$ cluster can be written as

$$L_{ij} = P(T_{ij} \in (0, u_{ij}]|x_{ij}, r_i) \delta_{1ij} P(T_{ij} \in (u_{ij}, v_{ij}]|x_{ij}, r_i) \delta_{2ij}$$

$$\times P(T_{ij} \in (u_{ij}, \infty)|x_{ij}, r_i) \delta_{3ij}(1-\psi_{ij}) P(T_{ij} \in (v_{ij}, \infty)|x_{ij}, r_i) \delta_{3ij}$$

$$\times P(U_{ij} = u_{ij}|x_{ij}, r_i) P(W_{ij} = v_{ij} - u_{ij}|x_{ij}, r_i) \psi_{ij},$$

(4)

where $\delta_{1i} = 1 - \delta_{1i} - \delta_{2i}$ and $\psi_{ij} = I(W_{ij} < \infty)$;
Define \((l_{ij}, r_{ij}]\) as
\[
(l_{ij}, r_{ij}] = \begin{cases} 
(0, u_{ij}], & \delta_{1ij} = 1, \\
(u_{ij}, v_{ij}], & \delta_{2ij} = 1, \\
(u_{ij}, \infty), & \delta_{3ij} = 1 \text{ and } \psi_{ij} = 0, \\
(v_{ij}, \infty), & \delta_{3ij} = 1 \text{ and } \psi_{ij} = 1
\end{cases}
\]

Consider an equivalence class of points,

\[
0 = s_0 < s_1 < \cdots < s_m < s_{m+1} = \infty,
\]

of the set \(S = \{ (l_{ij}, r_{ij}]; i = 1, \ldots, n; j = 1, \ldots, n_i \}\)
Parameterization

- Define

\[
\Lambda_{0t}(s_k) = \sum_{q=0}^{k} \exp\{\gamma_q\} \text{ for } k = 0, \ldots, m + 1,
\]

where \(\gamma_0 = -\infty\) and \(\gamma_{m+1} = \infty\), and also for \(k = 1, \ldots, m + 1\),

\[
\phi_{ijk} = I((s_{k-1}, s_k) \in (l_{ij}, r_{ij}]),
\]

\[
S_t(s_k|x_{ij}, r_i) = \exp\{-\Lambda_{0t}(s_k)\exp(\beta'_t x_{ij} + r_i)\},
\]

and

\[
g_{ijk} = S_t(s_{k-1}|x_{ij}, r_i) - S_t(s_k|x_{ij}, r_i)
\]

- So,

\[
P(T_{ij} \in (l_{ij}, r_{ij}] | x_{ij}, r_i) = \sum_{k=1}^{m+1} \phi_{ijk} g_{ijk}
\]
Parameterization

- Let $0 < \xi_1 < \cdots < \xi_a < \infty$ and $0 < \zeta_1 < \cdots < \zeta_b < \infty$ be distinct realizations of $u_{ij}$ and $w_{ij}$, respectively.
- Let $\lambda_{0u} = (\lambda_{0u}(\xi_1), \ldots, \lambda_{0u}(\xi_a))^\prime$ and $\lambda_{0w} = (\lambda_{0w}(\zeta_1), \ldots, \lambda_{0w}(\zeta_b))^\prime$ be the vectors of discrete baseline hazard functions of $U_{ij}$'s and $W_{ij}$'s, respectively.
- Let $\eta = (\beta'_t, \beta'_u, \beta'_w, \alpha_u, \alpha_w, \gamma', \lambda'_{0u}, \lambda'_{0w})^\prime$ denote the vector of the parameters, where $\gamma = (\gamma_1, \ldots, \gamma_m)^\prime$. 
Then the conditional likelihood in (4) can be expressed as

\[ L_{ij}(\eta) = \left\{ \sum_{k=1}^{m+1} \phi_{ijk}g_{ijk} \right\} \]

\[ \times \lambda_{0u}(u_{ij}) \exp\{\beta'_u x_{ij} + \alpha_u r_i\} \exp\{-\Lambda_{0u}(u_{ij}) \exp(\beta'_u x_{ij} + \alpha_u r_i)\} \]

\[ \times [\lambda_{0w}(w_{ij}) \exp\{\beta'_w x_{ij} + \alpha_w r_i\} \exp\{-\Lambda_{0w}(w_{ij}) \exp(\beta'_w x_{ij} + \alpha_w r_i)\}]^{\psi_{ij}} \]

The full likelihood of the \( i^{th} \) cluster based on complete data is defined by

\[ L_i^c(\eta, \theta) = \left\{ \prod_{j=1}^{n_i} L_{ij}(\eta) \right\} f(r_i; \theta), \]

where \( f(r_i; \theta) \) is pdf of the normal frailty with mean 0 and variance \( \theta \).
Therefore, the log-likelihood based on complete data can be written as

\[
l^c(\eta, \theta) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \log \left\{ \sum_{k=1}^{m+1} \phi_{ijk} \gamma_{ijk} \right\} \\
+ \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{p=1}^{a} I(U_{ij} = \xi_p) \{ \log \lambda_{0u}(\xi_p) + \beta'_u x_{ij} + \alpha_u r_i \\
- \Lambda_{0u}(\xi_p) \exp(\beta'_u x_{ij} + \alpha_u r_i) \} \\
+ \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{q=1}^{b} I(W_{ij} = \zeta_q) \psi_{ij} \{ \log \lambda_{0w}(\zeta_q) + \beta'_w x_{ij} + \alpha_w r_i \\
- \Lambda_{0w}(\zeta_q) \exp(\beta'_w x_{ij} + \alpha_w r_i) \} \\
- \frac{1}{2} \sum_{i=1}^{n} \left\{ \log(2\pi \theta) + \frac{r_i^2}{\theta} \right\}
\]  

(5)
Since \( r_i \) is not observable, we employ the EM algorithm for parameter estimation.

We need to replace the terms involving \( r_i \) by the conditional expectations.

For any frailty function \( g(r_i) \), the conditional expectation can be written as:

\[
E[g(r_i)|o_i, \eta, \theta] = \frac{\int_{-\infty}^{\infty} g(r_i) L^c_i(\eta, \theta) dr_i}{\int_{-\infty}^{\infty} L^c_i(\eta, \theta) dr_i}
\]

Using Gauss-Hermite method (Abramowitz and Stegun, 1970), the conditional expectation can be approximated as:

\[
\hat{E}[g(r_i)|o_i, \eta, \theta] = \frac{\sum_{s=1}^{M} g(r_i(x_s)) L^c_i(\eta, \theta; o_i, r_i(x_s)) w_s}{\sum_{s=1}^{M} L^c_i(\eta, \theta; o_i, r_i(x_s)) w_s}
\]

\( x_s \): a horizontal abscissa with weight \( w_s \) for \( s = 1, \ldots, M \), where \( M \) is a pre-specified value.
Parameter estimation

- Since the expectation of (5) is decomposed into four terms, the M-step is proceeded by estimating term by term.
- For estimating \((\beta_t', \gamma')'\), we employed one-step Newton-Raphson algorithm.
  - \((\beta_t^{(l)}', \gamma^{(l)}')'\) : the \(l^{th}\) (\(l = 0, \ldots\)) iterative solution of \((\beta_t', \gamma')'\)
  - The \((l + 1)^{th}\) solution can be derived from the following equation:

\[
\begin{pmatrix}
\beta_t^{(l+1)} \\
\gamma^{(l+1)}
\end{pmatrix} = \begin{pmatrix}
\beta_t^{(l)} \\
\gamma^{(l)}
\end{pmatrix} - \hat{E}[H_t|o, \beta_t^{(l)}, \gamma^{(l)}]^{-1} \hat{E}[U_t|o, \beta_t^{(l)}, \gamma^{(l)}],
\]

where

\[
U_t(\beta_t, \gamma) = \begin{pmatrix}
\frac{\partial l^c(\eta, \theta)}{\partial \gamma'} \\
\frac{\partial l^c(\eta, \theta)}{\partial \beta_t'}
\end{pmatrix} \quad \text{and} \quad H_t(\beta_t, \gamma) = \begin{bmatrix}
\frac{\partial^2 l^c(\eta, \theta)}{\partial \gamma \partial \gamma'} & \frac{\partial^2 l^c(\eta, \theta)}{\partial \gamma \partial \beta_t'} \\
\frac{\partial^2 l^c(\eta, \theta)}{\partial \beta_t \partial \gamma'} & \frac{\partial^2 l^c(\eta, \theta)}{\partial \beta_t \partial \beta_t'}
\end{bmatrix}
\]
Parameter estimation

- For estimating \((\beta_u', \alpha_u)\)', we first determine the Breslow-type estimate for \(\lambda_{0u}\) (Klein & Moeschberger, 2003). Then we use this estimate to derive one-step estimate of \((\beta_u', \alpha_u)\)'

- Similarly, we obtain one-step estimates of \((\beta_w', \alpha_w)\)' after obtaining the Breslow-type estimate for \(\lambda_{0w}\)

- Now let \(\hat{\eta}\) denote the maximum likelihood estimate of \(\eta\)
By the method of Louis (1982) the estimated variance-covariance matrix of \( \hat{\eta} \) can be defined as the inverse of the observed matrix \( I(\hat{\eta}) \), where

\[
I(\hat{\eta}) = E \left[ - \frac{\partial^2 l^c(\eta, \theta)}{\partial \eta \partial \eta'} \left| o, \hat{\eta}, \hat{\theta} \right. \right] - E \left[ \frac{\partial l^c(\eta, \theta)}{\partial \eta} \frac{\partial l^c(\eta, \theta)}{\partial \eta'} \left| o, \hat{\eta}, \hat{\theta} \right. \right]
\]

Thus, the inference regarding \( \beta_t, \beta_u, \) and \( \beta_w \) can be done using the sub-matrix of \( I(\hat{\eta})^{-1} \) in conjunction with these parameters. Similarly, we use the elements of \( I(\hat{\eta})^{-1} \) for the inference of \( \alpha_u \) and \( \alpha_w \)
Setup

- The frailty $r_i$ is generated from a normal distribution with a mean of zero and variance $\theta$
  - $\theta = 1$ and $2$
- The values of $t_{ij}$, $u_{ij}$, and $w_{ij}$ are generated from the models (1)-(3) accordingly
  - A binary covariate $z_{ij}$ is generated from a Bernoulli trial with a success probability of 0.5
  - Set $\beta_u = \beta_w = \beta_t = \alpha_u = \alpha_w = 0.5$, and $\lambda_u(t) = 4$, $\lambda_w(t) = 8$, and $\lambda_t(t) = 16$ if $t > 0$
  - We control the last observation to be 10 so that the values of $t_{ij}$, $u_{ij}$, and $v_{ij}$ cannot be observed
- We use different number of clusters ($n = 25$, $100$, and $200$) with the same $n_i$ being $2$, $3$, or $5$ for each cluster and $n_i$ being generated from the discrete uniform distribution $\{1, \ldots, 5\}$
Table 1: Simulation results of the mean of bias (Bias), the standard deviation (SE), the mean of standard error (SEM) of parameter estimates, and the coverage probability (CP) based on 1,000 replications when \( n_i \)'s are equal for each cluster.

<table>
<thead>
<tr>
<th>( n )</th>
<th>Parameter</th>
<th>( \theta = 1 )</th>
<th>( \theta = 2 )</th>
<th>( \theta = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>( \beta_u )</td>
<td>0.017 0.326 0.310 94.7</td>
<td>0.026 0.267 0.248 92.8</td>
<td>0.014 0.196 0.189 94.1</td>
</tr>
<tr>
<td></td>
<td>( \alpha_u )</td>
<td>0.014 0.185 0.174 93.3</td>
<td>0.016 0.142 0.139 94.9</td>
<td>0.011 0.111 0.106 94.7</td>
</tr>
<tr>
<td></td>
<td>( \beta_w )</td>
<td>0.037 0.383 0.357 94.2</td>
<td>0.020 0.291 0.283 94.6</td>
<td>0.017 0.224 0.215 94.6</td>
</tr>
<tr>
<td></td>
<td>( \alpha_w )</td>
<td>0.032 0.221 0.205 94.4</td>
<td>0.014 0.169 0.162 93.8</td>
<td>0.015 0.129 0.123 93.4</td>
</tr>
<tr>
<td></td>
<td>( \beta_t )</td>
<td>0.076 0.537 0.527 95.2</td>
<td>0.052 0.437 0.415 94.6</td>
<td>0.076 0.314 0.306 94.2</td>
</tr>
<tr>
<td></td>
<td>( \theta )</td>
<td>0.070 0.829 0.643 87.2</td>
<td>0.096 0.696 0.551 88.8</td>
<td>0.087 0.510 0.453 91.2</td>
</tr>
<tr>
<td>100</td>
<td>( \beta_u )</td>
<td>0.004 0.148 0.148 94.6</td>
<td>0.005 0.117 0.120 95.7</td>
<td>0.002 0.093 0.092 94.6</td>
</tr>
<tr>
<td></td>
<td>( \alpha_u )</td>
<td>0.008 0.080 0.080 95.7</td>
<td>0.003 0.067 0.065 95.1</td>
<td>0.003 0.052 0.050 94.3</td>
</tr>
<tr>
<td></td>
<td>( \beta_w )</td>
<td>0.016 0.169 0.167 94.7</td>
<td>-0.006 0.141 0.135 94.1</td>
<td>0.012 0.101 0.104 96.0</td>
</tr>
<tr>
<td></td>
<td>( \alpha_w )</td>
<td>0.004 0.095 0.092 95.2</td>
<td>0.006 0.074 0.075 95.3</td>
<td>0.003 0.057 0.058 95.5</td>
</tr>
<tr>
<td></td>
<td>( \beta_t )</td>
<td>0.021 0.254 0.253 95.1</td>
<td>0.046 0.199 0.202 94.0</td>
<td>0.055 0.152 0.151 94.0</td>
</tr>
<tr>
<td></td>
<td>( \theta )</td>
<td>0.077 0.388 0.291 81.9</td>
<td>-0.005 0.313 0.250 87.7</td>
<td>0.041 0.241 0.213 92.1</td>
</tr>
<tr>
<td>200</td>
<td>( \beta_u )</td>
<td>0.004 0.105 0.104 95.7</td>
<td>0.001 0.084 0.084 96.0</td>
<td>-0.002 0.066 0.065 95.1</td>
</tr>
<tr>
<td></td>
<td>( \alpha_u )</td>
<td>0.003 0.057 0.056 94.0</td>
<td>0.003 0.045 0.046 95.1</td>
<td>0.000 0.035 0.035 95.6</td>
</tr>
<tr>
<td></td>
<td>( \beta_w )</td>
<td>-0.001 0.119 0.117 94.7</td>
<td>-0.007 0.096 0.095 94.8</td>
<td>0.004 0.076 0.073 93.2</td>
</tr>
<tr>
<td></td>
<td>( \alpha_w )</td>
<td>0.004 0.066 0.064 95.2</td>
<td>0.001 0.052 0.052 94.5</td>
<td>0.003 0.041 0.040 94.7</td>
</tr>
<tr>
<td></td>
<td>( \beta_t )</td>
<td>0.020 0.179 0.178 95.5</td>
<td>0.033 0.138 0.142 95.0</td>
<td>0.061 0.110 0.107 94.7</td>
</tr>
<tr>
<td></td>
<td>( \theta )</td>
<td>0.010 0.263 0.201 80.5</td>
<td>0.022 0.214 0.174 87.5</td>
<td>0.044 0.174 0.150 91.0</td>
</tr>
<tr>
<td>100</td>
<td>( \beta_u )</td>
<td>0.004 0.147 0.148 94.3</td>
<td>0.005 0.117 0.120 95.7</td>
<td>0.003 0.093 0.092 94.4</td>
</tr>
<tr>
<td></td>
<td>( \alpha_u )</td>
<td>0.007 0.061 0.061 95.1</td>
<td>0.003 0.050 0.049 94.7</td>
<td>0.003 0.039 0.038 94.8</td>
</tr>
<tr>
<td></td>
<td>( \beta_w )</td>
<td>0.007 0.170 0.169 94.6</td>
<td>-0.007 0.141 0.136 94.2</td>
<td>0.012 0.101 0.105 96.4</td>
</tr>
<tr>
<td></td>
<td>( \alpha_w )</td>
<td>0.005 0.072 0.071 95.6</td>
<td>0.006 0.058 0.058 94.8</td>
<td>0.003 0.044 0.044 94.6</td>
</tr>
<tr>
<td></td>
<td>( \beta_t )</td>
<td>-0.010 0.263 0.268 95.5</td>
<td>0.032 0.209 0.210 95.9</td>
<td>0.052 0.151 0.154 94.1</td>
</tr>
<tr>
<td></td>
<td>( \theta )</td>
<td>-0.347 0.527 0.402 72.0</td>
<td>-0.134 0.468 0.379 82.9</td>
<td>0.023 0.399 0.356 91.7</td>
</tr>
<tr>
<td>200</td>
<td>( \beta_u )</td>
<td>0.004 0.105 0.104 95.5</td>
<td>0.001 0.084 0.084 95.9</td>
<td>-0.002 0.066 0.065 95.1</td>
</tr>
<tr>
<td></td>
<td>( \alpha_u )</td>
<td>0.003 0.043 0.043 94.4</td>
<td>0.003 0.034 0.035 95.4</td>
<td>0.000 0.026 0.027 95.5</td>
</tr>
<tr>
<td></td>
<td>( \beta_w )</td>
<td>-0.002 0.119 0.118 94.4</td>
<td>-0.007 0.097 0.096 94.7</td>
<td>0.004 0.076 0.074 93.6</td>
</tr>
<tr>
<td></td>
<td>( \alpha_w )</td>
<td>0.003 0.051 0.050 94.9</td>
<td>0.002 0.040 0.040 94.6</td>
<td>0.002 0.032 0.031 94.5</td>
</tr>
<tr>
<td></td>
<td>( \beta_t )</td>
<td>-0.009 0.189 0.188 95.9</td>
<td>0.024 0.146 0.148 95.2</td>
<td>0.060 0.111 0.108 94.6</td>
</tr>
<tr>
<td></td>
<td>( \theta )</td>
<td>-0.381 0.364 0.278 61.4</td>
<td>-0.150 0.330 0.264 81.4</td>
<td>0.015 0.290 0.250 91.3</td>
</tr>
</tbody>
</table>

4.2. Lymphatic filariasis data

We analyze data related to lymphatic filariasis (LF) disease which is a parasite helminthiasis often infected by mosquitoes. This dataset was analyzed earlier by several authors including Williamson et al. (2008), Zhang & Sun (2010), and Kim & Kim (2014). Note that this dataset is composed of unequal cluster sizes. In this analysis we investigated...
# Simulation results: when \(n_i\)'s are unequal

Table 2: Simulation results of the mean of bias (Bias), the standard deviation (SE), the mean of standard error (SEM) of parameter estimates, and the coverage probability (CP) based on 1,000 replications when \(n_i\)'s are unequal

<table>
<thead>
<tr>
<th>(n)</th>
<th>Parameter</th>
<th>1 (\theta)</th>
<th></th>
<th></th>
<th>2 (\theta)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bias</td>
<td>SE</td>
<td>SEM</td>
<td>CP</td>
<td>Bias</td>
</tr>
<tr>
<td>25</td>
<td>(\beta_u)</td>
<td>0.023</td>
<td>0.267</td>
<td>0.254</td>
<td>94.2</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(\alpha_u)</td>
<td>0.020</td>
<td>0.149</td>
<td>0.146</td>
<td>94.5</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(\beta_w)</td>
<td>0.003</td>
<td>0.296</td>
<td>0.289</td>
<td>94.1</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(\alpha_w)</td>
<td>0.018</td>
<td>0.181</td>
<td>0.170</td>
<td>95.1</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(\beta_t)</td>
<td>0.051</td>
<td>0.429</td>
<td>0.420</td>
<td>94.8</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>(\theta)</td>
<td>0.078</td>
<td>0.711</td>
<td>0.565</td>
<td>85.7</td>
<td>-0.020</td>
</tr>
<tr>
<td>100</td>
<td>(\beta_u)</td>
<td>0.007</td>
<td>0.123</td>
<td>0.122</td>
<td>95.6</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(\alpha_u)</td>
<td>0.000</td>
<td>0.069</td>
<td>0.066</td>
<td>94.5</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(\beta_w)</td>
<td>0.010</td>
<td>0.141</td>
<td>0.138</td>
<td>94.5</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(\alpha_w)</td>
<td>0.009</td>
<td>0.078</td>
<td>0.077</td>
<td>95.1</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(\beta_t)</td>
<td>0.039</td>
<td>0.210</td>
<td>0.205</td>
<td>94.9</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(\theta)</td>
<td>-0.013</td>
<td>0.331</td>
<td>0.258</td>
<td>85.6</td>
<td>-0.137</td>
</tr>
<tr>
<td>200</td>
<td>(\beta_u)</td>
<td>0.002</td>
<td>0.087</td>
<td>0.086</td>
<td>94.9</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(\alpha_u)</td>
<td>0.004</td>
<td>0.043</td>
<td>0.046</td>
<td>97.2</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(\beta_w)</td>
<td>0.005</td>
<td>0.097</td>
<td>0.097</td>
<td>94.4</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(\alpha_w)</td>
<td>0.001</td>
<td>0.054</td>
<td>0.053</td>
<td>94.6</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(\beta_t)</td>
<td>0.036</td>
<td>0.146</td>
<td>0.144</td>
<td>93.2</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(\theta)</td>
<td>-0.014</td>
<td>0.221</td>
<td>0.181</td>
<td>87.2</td>
<td>-0.123</td>
</tr>
</tbody>
</table>
Simulation results: four different frailty distributions when \( n \) is equal to 100 and \( n_i \)'s are unequal

Table 3: Simulation results of the mean of bias (Bias), the standard deviation (SD), the mean of standard error (SEM) of parameter estimates, and the coverage probability (CP) based on 1,000 replications under four different frailty distributions when \( n \) is equal to 100 and \( n_i \)'s are unequal

<table>
<thead>
<tr>
<th>Frailty distribution</th>
<th>( N(0, 2.43) )</th>
<th>( U(-2.70, 2.70) )</th>
<th>( t(3.40) )</th>
<th>( G(1.46, 0.68) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_u )</td>
<td>0.006 0.123 0.122 95.4</td>
<td>0.005 0.124 0.122 95.0</td>
<td>0.001 0.118 0.122 95.9</td>
<td>-0.005 0.123 0.122 95.1</td>
</tr>
<tr>
<td>( \alpha_u )</td>
<td>0.000 0.049 0.047 93.6</td>
<td>0.002 0.045 0.047 95.4</td>
<td>0.005 0.053 0.051 94.2</td>
<td>0.002 0.049 0.048 94.1</td>
</tr>
<tr>
<td>( \beta_w )</td>
<td>0.010 0.142 0.139 94.7</td>
<td>-0.003 0.144 0.139 94.4</td>
<td>-0.001 0.142 0.139 95.1</td>
<td>0.008 0.137 0.134 94.1</td>
</tr>
<tr>
<td>( \alpha_w )</td>
<td>0.008 0.057 0.056 94.9</td>
<td>0.004 0.057 0.054 94.2</td>
<td>0.004 0.062 0.062 95.4</td>
<td>0.001 0.050 0.052 96.9</td>
</tr>
<tr>
<td>( \beta_t )</td>
<td>0.025 0.216 0.215 95.1</td>
<td>0.009 0.221 0.215 95.0</td>
<td>0.027 0.215 0.213 94.2</td>
<td>-0.055 0.226 0.230 94.5</td>
</tr>
<tr>
<td>( \theta )</td>
<td>-0.184 0.548 0.451 84.1</td>
<td>-0.181 0.506 0.451 87.0</td>
<td>-0.611 0.576 0.386 54.2</td>
<td>-0.836 0.418 0.380 39.7</td>
</tr>
</tbody>
</table>
Mastitis data

- A observation study was conducted to estimate the incidence of different organisms causing mastitis in the dairy cattle population in Flanders.
- A total of 100 cows was monitored at the udder-quarter level for bacterial infections from the time of parturition, at which the cow was included in the cohort and observed infection-free, until the end of the lactation period.
- The four udder quarters are obviously clustered within a cow and udder quarters that experience an event are interval-censored because of periodic follow-up.
- We want to investigate the effect of the covariates that change within cow (e.g. front and rear udder quarters) and covariates that change between cows (e.g. number of calving).
## Results

Table 3: Parameter estimates, their standard errors, and $p$-values of mastitis data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_u$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>location: rear</td>
<td>0.175</td>
<td>0.101</td>
<td>0.082</td>
</tr>
<tr>
<td>calving: 2-4</td>
<td>-0.110</td>
<td>0.109</td>
<td>0.317</td>
</tr>
<tr>
<td>calving: &gt;4</td>
<td>0.640</td>
<td>0.168</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\alpha_u$</td>
<td>0.819</td>
<td>0.053</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\beta_w$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>location: rear</td>
<td>0.086</td>
<td>0.104</td>
<td>0.405</td>
</tr>
<tr>
<td>calving: 2-4</td>
<td>-0.154</td>
<td>0.112</td>
<td>0.170</td>
</tr>
<tr>
<td>calving: &gt;4</td>
<td>0.122</td>
<td>0.169</td>
<td>0.469</td>
</tr>
<tr>
<td>$\alpha_w$</td>
<td>0.200</td>
<td>0.047</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\beta_t$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>location: rear</td>
<td>0.162</td>
<td>0.119</td>
<td>0.172</td>
</tr>
<tr>
<td>calving: 2-4</td>
<td>-0.011</td>
<td>0.334</td>
<td>0.975</td>
</tr>
<tr>
<td>calving: &gt;4</td>
<td>1.618</td>
<td>0.473</td>
<td>0.001</td>
</tr>
<tr>
<td>$\theta$</td>
<td>2.001</td>
<td>0.341</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Summary

- We proposed the Cox PH models with a shared frailty effect incorporated with clustered interval-censored data for which there exits a dependence between the failure time and the censoring times.
- After constructing the likelihood function based on complete data, we employed the EM algorithm for parameter estimation.
- Simulation results showed that when a cluster size is fixed, both Bias and SEM of parameter estimates except for $\theta$ decrease as the number of the clusters increases, and the CPs are close to the nominal level of 0.95.
- The overall trends were similar regardless of whether the number of observations within the same cluster is equal or unequal.
- Moreover, our proposed method was robust to misspecified frailty distribution such as the uniform, $t$, and gamma distributions.
Thank you!