Linear Mixed Model for Doubly-Interval-Censored Data with Semiparametrically Specified Distributions

Arnošt Komárek and Emmanuel Lesaffre

Biostatistical Centre, Katholieke Universiteit Leuven, Belgium

arnost.komarek@med.kuleuven.be

Problem

• Regression model for the time to caries $T_{i,l}$ on permanent first molars (teeth 16, 26, 36, 46)

$i = 1, \ldots, N$ (*children*), $l = 1, \ldots, 4$ (*teeth*)

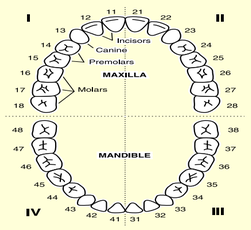
Frequency of brushing

Plaque accumulation

Presence of sealants

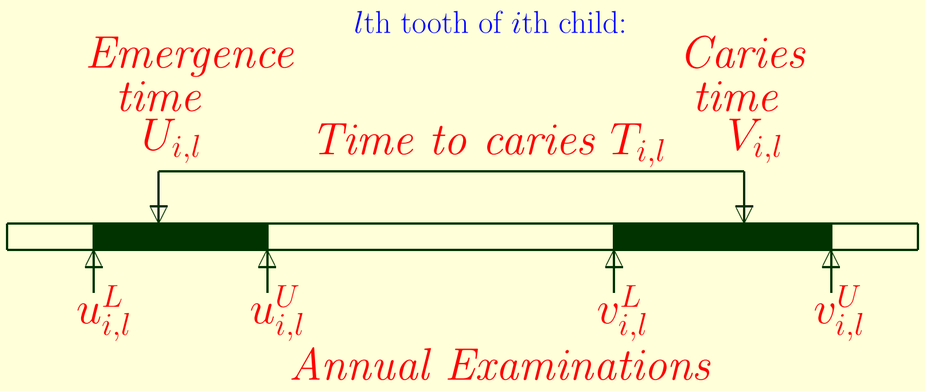
Caries on primary teeth

$\Longleftrightarrow$ CARIES



$\Longrightarrow$ **Clustered data**

$\Longrightarrow$ **Doubly interval censoring**



$$T_{i,l} = V_{i,l} - U_{i,l}$$

only known $u^L_{i,l} < U_{i,l} \leq u^U_{i,l}$ and $v^L_{i,l} < V_{i,l} \leq v^U_{i,l}$

Model

Emergence time:

$$\log(U_{i,l}) = \boldsymbol{\delta}' \boldsymbol{x}^u_{i,l} + d_i + \zeta_{i,l}$$

Caries time:

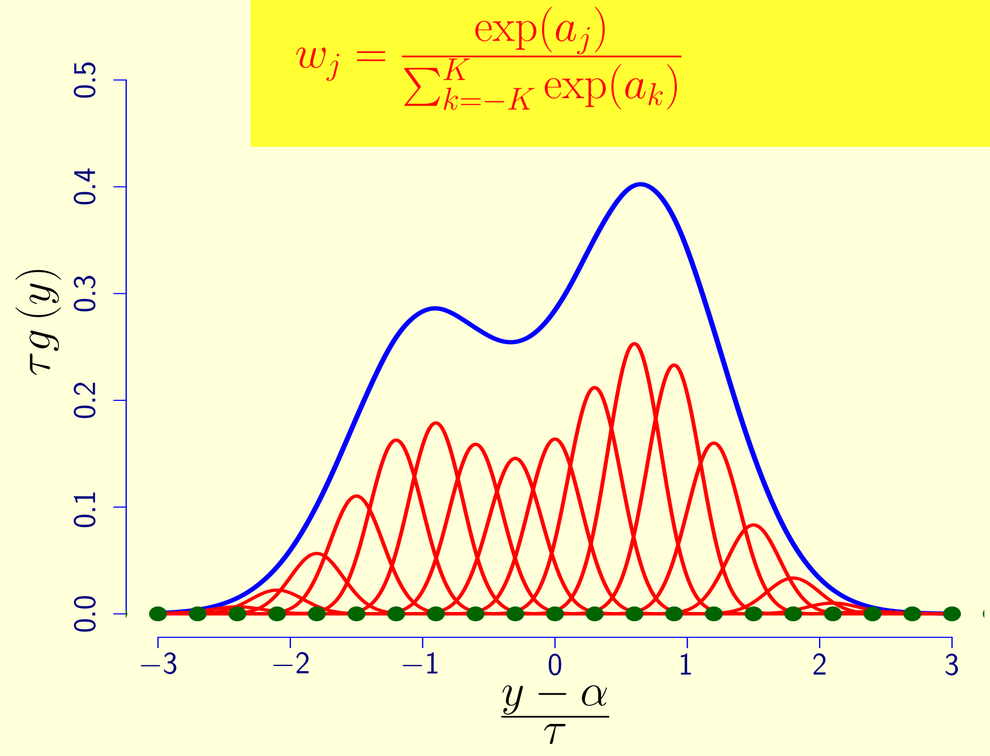
$$\log(T_{i,l}) = \log(V_{i,l} - U_{i,l}) = \boldsymbol{\beta}' \boldsymbol{x}^t_{i,l} + b_i + \varepsilon_{i,l}$$

- $\boldsymbol{x}^u_{i,l}$, $\boldsymbol{x}^t_{i,l}$: covariates
- $\boldsymbol{\delta}$, $\boldsymbol{\beta}$: regression parameters
- $\zeta_{i,l}$: emergence error terms $\overset{i.i.d.}{\sim} g_\zeta(\zeta)$
- d_i: child-specific random effects for emerg. $\overset{i.i.d.}{\sim} g_d(d)$
- $\varepsilon_{i,l}$: caries error terms $\overset{i.i.d.}{\sim} g_\varepsilon(\varepsilon)$
- b_i: child-specific random effects for caries $\overset{i.i.d.}{\sim} g_b(b)$

Spline-like model for densities g_ζ, g_d, g_ε, g_b

$$g(y) \sim \alpha + \tau \sum_{j=-K}^{K} w_j \mathcal{N}(\mu_j, \sigma^2)$$

$$w_j = \frac{\exp(a_j)}{\sum_{k=-K}^{K} \exp(a_k)}$$



Density related model parameters

Fixed	Estimated
• Knots $\{\mu_j\}$	• (Transformed) weights $\{a_j\}$
• Basis variance σ^2	• Intercept α
	• Scale τ

Estimation

• Bayesian using MCMC methodology

• Vague prior distributions for all parameters but the transformed mixture weights $\{a_j\}$

• Markov random field prior for $\{a_j\}$ $\Longleftrightarrow$ penalty

▶ R package bayesSurv (http://www.R-project.org) ◀

Curious to know details and see results of the dental analysis?

Komárek, A. (2006). *Accelerated Failure Time Models for Multivariate Interval-Censored Data with Flexible Distributional Assumptions*, Chapter 9. PhD. Thesis. Katholieke Universiteit Leuven, Faculteit Wetenschappen.