Modelling Functional Disability with Hawkes Process

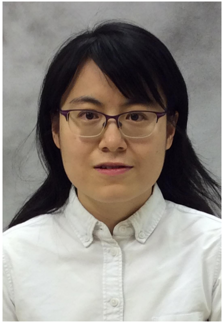
Jiwon Jung

Department of Statistics
Purdue University
jung320@purdue.edu

June 23, 2022

Collaborators

Prof. Kiseop Lee
Department of Statistics
(Data Science in Finance)
Purdue University
kiseop@purdue.edu



Prof. Mengyi Xu
Department of Statistics,
Department of Mathematics
Purdue University
xumy@purdue.edu

Overview

Introduction

Literature review

- Prior literature usually assumes Markov property for modelling health transitions, for which the probabilities of transition at each age depend on the current status only. (see e.g., Fong et al., 2015; Li et al., 2017; Sherris and Wei, 2021)
- Showing that the probabilities of functional status transitions are duration dependent, other literature (Cai et al., 2006; Biessy, 2017) assumes semi-Markov process model to incorporate not only age and the current status but also on the duration in the current state.
- However, the state and duration effect with respect to the past functional disability experience has been less studied.

Motivation

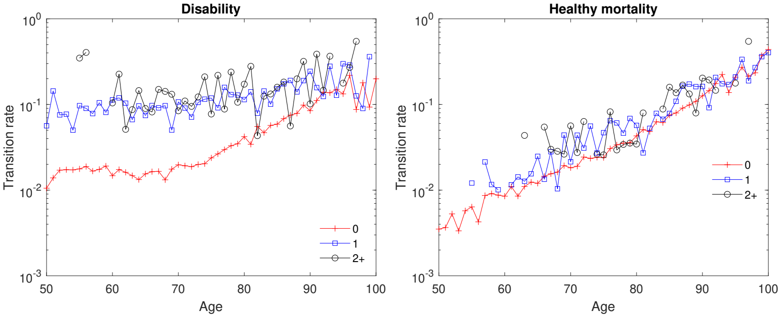


Figure 1. Crude health transition rates of functional disability and healthy mortality. The legend indicates the number of past functional-disabilities during the investigation period.

- Our explanatory data analysis suggest that the elderly who have experienced functional disability have a higher chance of functional disability and mortality than those who were never disabled before.

Hawkes process

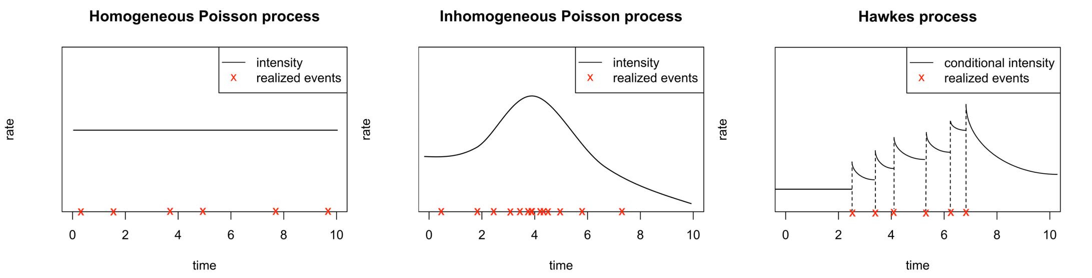


Figure 2. Graphs of a homogenous Poisson process, inhomogeneous Poisson process, and Hawkes process on the real line.

- Hawkes process is a self–exciting point process, in which the occurrence of an event increases the probability of occurrence of another event.

Goal

- Our goal is to estimate the intensity of age and gender-specific transitions by incorporating the impact of past functional disability as well as time spent in the current state using Hawkes process.

Three-state health transition model

Data preparation

- We use the `RAND HRS Data` 1992-2018 from the U.S. Health and Retirement Study (HRS), a nationally representative longitudinal panel survey.[1]
- The HRS is a biennial survey which began in 1992 and follows up with interviews of initially non-institutionalised Americans aged 50 and above.
- The health state is determined by a person's ability to perform activities of daily living (ADLs), such as bathing, toileting, and dressing; An individual needing help in two or more ADLs is functionally disabled.

[1] https://hrs.isr.umich.edu/data-products

Hawkes transition intensity

The transition intensity for individual k of transition type s at time t is given by

$$\lambda_s(t) = \underbrace{\phi_s(t)}_{\text{background intensity}} + \underbrace{\mu_s(t - T_t)}_{\text{exciting function}} \cdot \underbrace{\mathbb{1}_F(t)}_{\text{disability indicator}}, \tag{1}$$

where T_t is the latest transition time before time t.

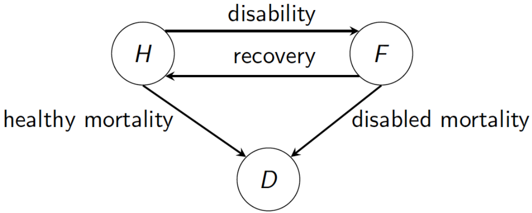


Figure 3. The three-state health transition model. H means healthy; F means functionally disabled; D means dead.

Hawkes transition intensity

- $\phi_s(t)$ is the Gompertz background intensity that captures the impact of observable variates such as (scaled) age $x_k(t)$ and gender indicator F_k at time t

$$\phi_s(t) = \exp(\beta_s^{intercept} + \beta_s^{age} x_k(t) + \beta_s^{female} F_k) \tag{2}$$

- $\mu_s(\cdot)$ is the exciting kernel function that captures the impact of past functional disabilities

$$\mu_s(\tau) = \alpha_s \exp\left(-\delta_s \tau\right) \qquad \text{(exponential decay)} \tag{3}$$

- $\mathbb{1}_F(t)$ is the indicator of the past experience of functional disability

$$\mathbb{1}_F(t) = \begin{cases} 1 & \text{if functionally disabled at least once before time t} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Maximum likelihood estimation

Estimation under left truncation and censoring

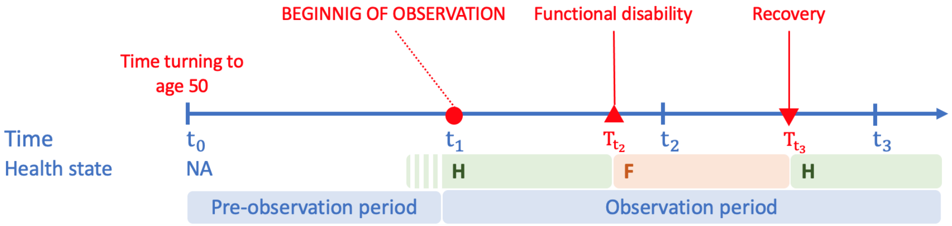


Figure 4. Example of an individual's health transition history. t_j is the time for j^{th} interview.

When an individual was first observed after the age of 50 and he/she was not in the functionally disabled state, we have

1. unknown $\mathbb{1}_F(t_1)$: presence of past functional disability
2. unobserved T_{t_1}: the latest transition time before the first interview (if any)

Complete likelihood function

Suppose there are a total of K individuals, S transition types, and J interview waves. The complete log likelihood function is given by

$$l(\boldsymbol{\theta}) = \sum_{k=1}^{K} \sum_{s=1}^{S} \sum_{j=1}^{J-1} l_{k,s,j}(\boldsymbol{\theta}), \tag{5}$$

where $\boldsymbol{\theta}$ denotes the set of parameters to be estimated, and

$$\begin{aligned} l_{k,s,j}(\boldsymbol{\theta}) = Y_{k,s,j} \ln \lambda_{k,s}(\hat{t}_{k,j}) &- R_{k,s}(t_{k,j}) \int_{t_{k,j}}^{\min\{\hat{t}_{k,j}, t_{k,j+1}\}} \lambda_{k,s}(u)\mathrm{d}u \\ &- R_{k,s}(\hat{t}_{k,j}) \int_{\min\{\hat{t}_{k,j}, t_{k,j+1}\}}^{t_{k,j+1}} \lambda_{k,s}(u)\mathrm{d}u, \end{aligned}$$

Here, we introduce two indicator variables: (1) $Y_{k,s,j} = 1$ if transition type s is observed between the j^{th} and $(j+1)^{\text{th}}$ interviews, and (2) $R_{k,s}(t) = 1$ if the individual is exposed to the risk of transition type s at time t.

EM-algorithm

We use the EM algorithm to perform the estimation.

EM-algorithm

1. Initialize $\theta^{(1)}$ assuming no truncation
2. For $i = 1, 2, 3, \ldots,$ Iterate E-step and M-step until convergence[a]
   2.1 **E-step:** Using 10,000 simulations from $\boldsymbol{\theta}^{(i)}$, approximate the expectation

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) = \mathbb{E}_{\mathbb{1}_F, \tau_{trunc}|\boldsymbol{\theta}^{(i)}}\left[l(\boldsymbol{\theta})\right] = \mathbb{E}_{\mathbb{1}_F|\boldsymbol{\theta}^{(i)}}\left[\mathbb{E}_{\tau_{trunc}|\mathbb{1}_F, \boldsymbol{\theta}^{(i)}}\left[l(\boldsymbol{\theta})\right]\right] \tag{6}$$

   2.2 **M-step:** Update $\boldsymbol{\theta}$

$$\boldsymbol{\theta}^{(i+1)} = \text{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) \tag{7}$$

[a] repeat until the difference between the current and previous Q value is less than10^{-1}

Results

Estimation results

Model	p^*	LRT statistic†	AIC	BIC
Non-Hawkes	9	-	172,128.8	172,201.0
Single Hawkes: disability	11	1,974.0***	170,158.8	170,247.0
Single Hawkes: recovery	11	216.9***	171,915.9	172,004.1
Single Hawkes: healthy mortality	11	620.6***	171,512.3	171,600.5
Single Hawkes: disabled mortality	11	2,172.0***	169,960.8	170,049.0
Full Hawkes: all four-transition	17	**2,234.3*** **	**167,738.6**	**167,874.9**

* number of parameters $\quad$ † Non-Hawkes v. Single Hawkes; Single Hawkes v. Full Hawkes

Table 1. Likelihood ratio test, AIC, and BIC statistics of health transition models.

Estimation results

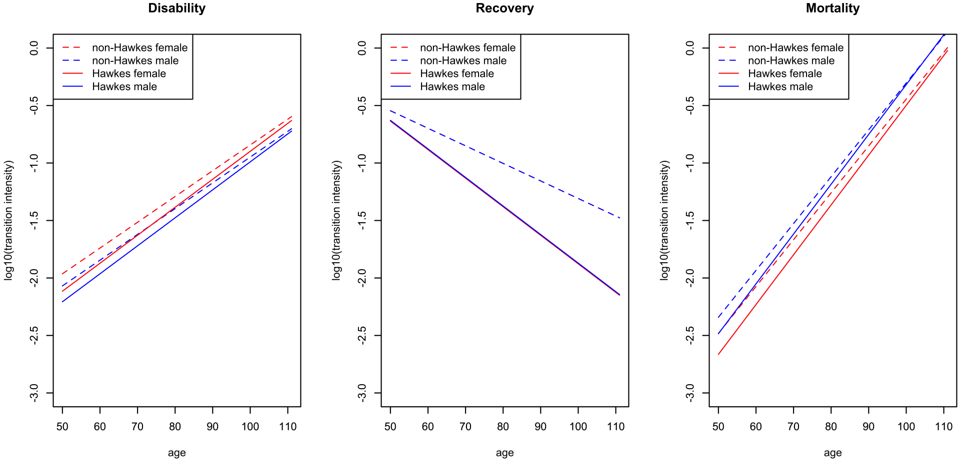


Figure 5. EM estimated background intensity regression parameters.

Estimation results

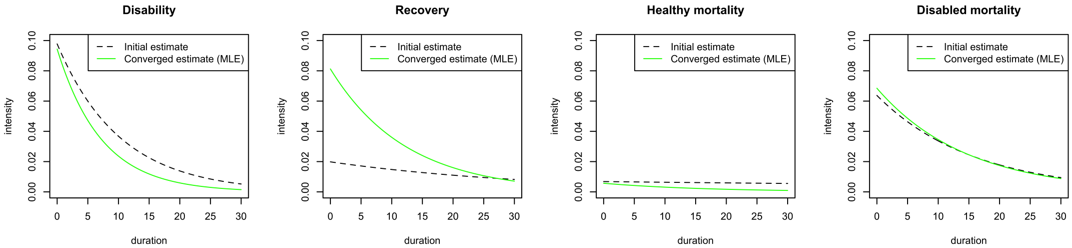


Figure 6. EM estimated exponential decay parameters.

EM convergence: background intensity parameters

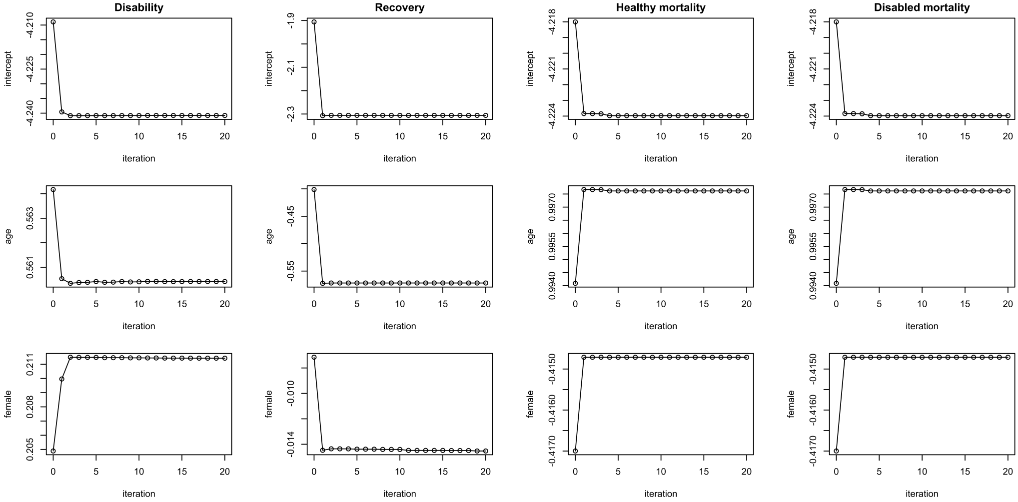


Figure 7. Background intensity coefficients updated by EM algorithm.

EM convergence: Hawkes kernel parameters

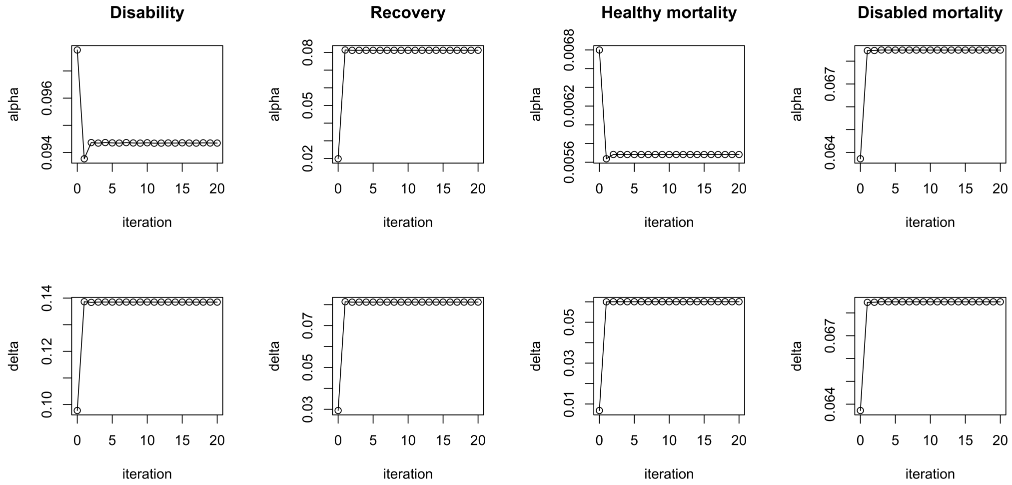


Figure 8. Hawkes kernel coefficients updated by EM algorithm.

Conclusion

Conclusion

- We proposed a multi-state health transition model with Hawkes process.
- The estimation results suggest that the functional disability and mortality intensities significantly increase on the onset of the disability and decay as the duration since the latest transition gets longer.

Reference

- Biessy, G. (2017) Continuous-time semi-markov inference of biometric laws associated with a long-term care insurance portfolio. *ASTIN Bulletin: The Journal of the IAA*, **47**(2), 527–561.
- Cai, L., Schenker, N. and Lubitz, J. (2006) Analysis of functional status transitions by using a semi-markov process model in the presence of left-censored spells. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **55**(4), 477–491.
- Fong, J. H., Shao, A. W. and Sherris, M. (2015) Multistate actuarial models of functional disability. *North American Actuarial Journal*, **19**(1), 41–59.
- Li, Z., Shao, A. W. and Sherris, M. (2017) The impact of systematic trend and uncertainty on mortality and disability in a multistate latent factor model for transition rates. *North American Actuarial Journal*, **21**(4), 594–610.
- Sherris, M. and Wei, P. (2021) A multi-state model of functional disability and health status in the presence of systematic trend and uncertainty. *North American Actuarial Journal*, **25**(1), 17–39.

Appendix

Conditional intensity

Definition

(Conditional intensity function) Consider a counting process $N(t)$ with associated histories $\mathcal{F}^N(t), t \geq 0$. If a non-negative function $\lambda(t)$ exists such that

$$\lambda(t) = \lim_{h \to 0} \frac{\mathbb{E}\left[N(t+h) - N(t) | \mathcal{F}^N(t)\right]}{h}, \tag{8}$$

then it is called the conditional intensity function of $N(t)$. Note that this function is also called the hazard function.

Hawkes process

Definition

(Hawkes process) The one-dimensional Hawkes process is a point process $N(t)$ which is characterized by its conditional intensity $\lambda(t)$ with respect to its natural filtration:

$$\lambda(t) = \phi(t) + \int_0^t \mu(t-s)\mathrm{d}N(s), \tag{9}$$

where $\phi(t)$ is the background intensity function, and the $\mu(t)$ is the excitation function satisfying $\int_0^\infty \mu(s)\mathrm{d}s < 1$.

EM-algorithm

E-step (evaluation):

In the presence of left-truncation, we can split the log likelihood terms into two terms with/without truncation;

$$\log L\left(\boldsymbol{\theta} | \{\mathbb{1}_f, \tau_{trunc}\}\right) = \sum_{k=1}^{K} \sum_{s=1}^{S} \left(\sum_{j:t_j < \inf_i \hat{t}_i} l_{k,s,j}\left(\boldsymbol{\theta}\right) \sum_{j:t_j \geq \inf_i \hat{t}_i} l_{k,s,j}\left(\boldsymbol{\theta}\right) \right) \tag{10}$$

Recall that $\inf_i \hat{t}_i$ denote the first transition time.

EM-algorithm

Now, we can separately calculate the conditional expectation to obtain $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)})$.

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) = Q_{trunc}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) + Q_{full}(\boldsymbol{\theta}) \tag{11}$$

i. Q_{trunc}**:**

$$\begin{aligned} Q_{trunc}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)}) &= \sum_{k=1}^{K}\sum_{s=1}^{S}\sum_{j:t_j<\inf_i \hat{t}_i} \mathbb{E}_{\mathbb{1}_f|\boldsymbol{\theta}^{(i)}}\left[\mathbb{E}_{\tau_{trunc}|\mathbb{1}_f,\boldsymbol{\theta}^{(i)}}\left[l_{k,s,j}(\boldsymbol{\theta})\right]\right] \\ &= \sum_{k=1}^{K}\sum_{s=1}^{S}\sum_{j:t_j<\inf_i \hat{t}_i}\sum_{m=0}^{1} P(\mathbb{1}_f(t_j)=m|\boldsymbol{\theta}^{(i)})\mathbb{E}\left[l_{k,s,j}(\boldsymbol{\theta})|\tau_{t_j}\geq\tilde{\tau}_{t_j},\mathbb{1}_f(t_j)=m,\boldsymbol{\theta}^{(i)}\right] \end{aligned}$$

EM-algorithm

Probability of being functionally disabled at least once before t

Note that

$$P(\mathbb{1}_f(t) = 0 | Z_t, \boldsymbol{\theta}^{(i)}) = \begin{cases} \exp\left(-\int_0^t \phi_1(u) + \phi_3(u) du\right) & \text{, if } Z_t = H \\ 0 & \text{, if } Z_t = F \end{cases}$$

Accordingly, $P(\mathbb{1}_f(t) = 1 | Z_t, \boldsymbol{\theta}^{(i)}) = 1 - P(\mathbb{1}_f(t) = 0 | Z_t, \boldsymbol{\theta}^{(i)})$

EM-algorithm

Expectation under truncation

$$\mathbb{E}\left[l_{k,s,j}(\boldsymbol{\theta})|\tau_{t_j} \geq \tilde{\tau}_{t_j}, \mathbb{1}_f(t_j) = 0, \boldsymbol{\theta}^{(i)}\right] = Y_{k,s,j} \log \phi_{k,s}(\hat{t}_j) - R_{k,s}(t_j) \int_{t_j}^{\min\{\hat{t}_j, t_{j+1}\}} \phi_{k,s}(u)\mathrm{d}u$$
$$- R_{k,s}(\hat{t}_j) \int_{\min\{\hat{t}_j, t_{j+1}\}}^{t_{j+1}} \phi_{k,s}(u)\mathrm{d}u,$$

$$\begin{aligned}
&\mathbb{E}[l_{k,s,j}(\boldsymbol{\theta})|\tau_{t_1} \geq \tilde{\tau}_{t_1}, \mathbb{1}_f(t_j) = 1, \boldsymbol{\theta}^{(i)}] \\
&\quad = \mathbb{E}\Bigg[Y_{k,s,j} \log \left(\phi_{k,s}(\hat{t}_j) + \mu_s(\tau_{\hat{t}_j})\right) - R_{k,s}(t_j) \int_{t_j}^{\min\{\hat{t}_j, t_{j+1}\}} \phi_{k,s}(u) + \mu_s(\tau_u)\mathrm{d}u \\
&\qquad - R_{k,s}(\hat{t}_j) \int_{\min\{\hat{t}_j, t_{j+1}\}}^{t_{j+1}} \phi_{k,s}(u) + \mu_s(\tau_u)\mathrm{d}u|\tau_{t_j} \geq \tilde{\tau}_{t_j}, \mathbb{1}_f(t_j) = 1, \boldsymbol{\theta}^{(i)}\Bigg],
\end{aligned} \tag{12}$$

EM-algorithm

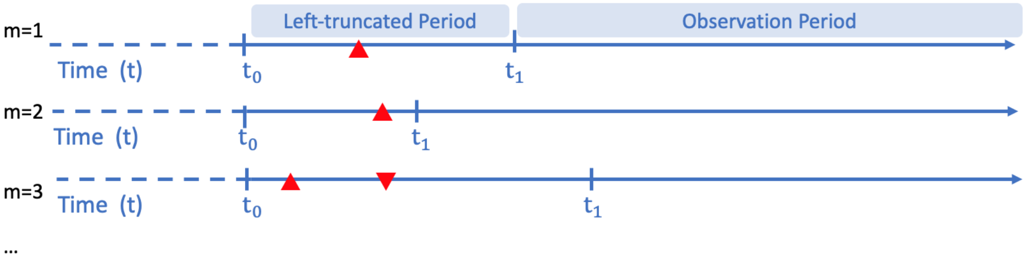


Figure 9. An example of an individual's simulated health transition history.

EM-algorithm

To approximate (12), we use Monte Carlo simulation.

1. Using time-rescaling theorem and Ogata's thinning method, obtain random samples from $T_{t_1}^{(n)} | \mathbb{1}_f(t_1) = 1, \boldsymbol{\theta}^{(i)}$ until we have N number of samples such that $T_{t_1}^{(n)} \leq \tilde{T}_{t_1}$. Note that $\tau_{t_1}^{(n)} = t - T_{t_1}^{(n)}$.

2. Approximate the conditional expectation under truncation

$$(12) \approx \frac{\sum_{n=1}^{N} Y_{k,s,j} \log\left(\phi_{k,s}(\hat{t}_j) + \mu_s(\tau_{\hat{t}_j}^{(n)})\right) - R_{k,s}(t_j) \int_{t_j}^{\min\{\hat{t}_j, t_{j+1}\}} \phi_{k,s}(u) + \mu_s(\tau_u^{(n)}) \mathrm{d}u}{N}$$

EM-algorithm

ii. Q_{full}:

$$Q_{full}(\boldsymbol{\theta}) = \sum_{k=1}^{K} \sum_{s=1}^{S} \sum_{j: t_j \geq \inf_i \hat{t}_i} l_{k,s,j}(\boldsymbol{\theta}) \tag{13}$$

This is the case after the first disability/recovery time (if any). Note that after the first disability/recovery transition, we can remove the expectation since $\mathbb{1}_f(t) = 1$ and $\tau_{t_j} = \tilde{\tau}_{t_j}$ with probability 1.