

# Mean Pattern Detecting in Hypnogram Data Using Discrete Semi-Markov Process

Haoqun Cao

Nov 17, 2023

## Abstract

Hypnogram is a plot describing the sleep stages during the sleep process. Hypnogram plays an important role in sleep study. In this work, we propose to model hypnogram data with **nonhomogeneous discrete semi-markov process**. We give a formal definition of such model and prove its likelihood can be decomposed into each transition type through proper approximation. Then we use a novel parametrization for nonhomogeneity which can separate the baseline transition intensity and the effect of nonhomogeneity meanwhile can be efficiently estimated. We propose a method of deriving a mean pattern which is representative and simple compared to the rather noisy data. Our empirical results are consistent with existing medical sleep studies.

**Keywords.** semi-markov process, competing risk model

## 1 Introduction

Hypnogram is a plot describing the sleep stages during the sleep process. It can be perceived as a categorical time series. Hypnogram plays an important role in sleep study (Swihart et al., 2008). When given large samples of hypnograms, an efficient way to explore its routine in a statistical way is urgently needed. We also hope to summarise the heterogeneous hypnogram observations into a simple but representative mean pattern.

From the perspective of categorical time series, statistical modeling methods can be summarised into 4 methods. The Multi-State model, the link function approach (Fahrmeir and Kaufmann, 1987), the likelihood-based approach (Fokianos and Kedem, 1998) and spectral envelope approach (Stoffer et al., 1993). In this report, we majorly follow the research line of the Multi-State model. We extend the existing method by proposing weaker assumptions.

The Multi-State Model is defined as a model for a continuous or discrete time stochastic process allowing individuals to move between a finite number of states. It can be generally specified by the transition kernel. However, parametrization and estimation of transition kernel is inefficient. The major way to reduce computational cost is by introducing assumptions. The most widely used assumption is the homogeneous Markov assumption where sojourn times follow a exponential distribution (Andersen and Keiding, 2002). However, the homogeneous Markov assumption can be too restrictive. Weaker assumptions are proposed including the semi-markov assumption where sojourn time can take any distribution.

Existing literature on semi-markov process (SMP) modeling majorly focuses on the homogeneous and continuous case. Kang and Lagakos (2007) proposes to model homogeneous SMP using transition intensity (equivalently cause-specific hazard). The estimation of the transition function is

done by jointly optimizing the likelihood function, which is costly. [Asanjarani et al. \(2021\)](#) proposes to decompose the likelihood of continuous homogeneous SMP into each transition type and use the standard survival analysis method to separately estimate transition intensity. Their method is easy to implement, however, homogeneous assumption is still too strong for real-world applications.

Therefore, we introduce the nonhomogeneous discrete semi-markov process to model our hypnogram data. We manage to extend the method of likelihood decomposition to the discrete non-homogeneous situation through an approximation. Then we explicitly model nonhomogeneity with a time-varying covariate in the proportional hazard model. Through that, we manage to separately specify the baseline transition intensity and the change of intensity caused by nonhomogeneity. Compared to existing method on modeling nonhomogeneity as shown in [Fahrmeir and Klinger \(1998\)](#), our method is easier to implement meanwhile can separate the effect of baseline intensity of effect of nonhomogeneity, therefore more interpretable.

After property modeling our data, we summarise the estimation result by predicting a **mean pattern**. The mean pattern can be perceived as a *mean* of an infinite-dimensional mixed random vector. We propose a way to derive the mean pattern that can capture the major mode of sleep process as shown in the experiments.

Our report covers the following parts in the following order. In section ??, we introduce the notations and basic properties of the probabilistic model we use. In section ??, we introduce our methods of parametrization for the model and estimation procedure. In section 2.4, we introduce how we derive the mean pattern after requiring the estimation. In section 3, we show our result on the oura ring pregnant women's data.

## 2 Methodology

### 2.1 Discrete Nonhomogeneous Semi-Markov Process

Let  $\mathcal{S}$  denote a discrete state set. Map its element to integers in  $[\mathcal{S}] = [m] := \{1, 2, \dots, m\}$ . Denote a sequence of r.v. take value in  $[m]$  as  $\{X_t, t \in \mathbb{N}\}$  representing the state of system at time point  $t$ . Denote  $T_0 = 0$ ,  $T_n = \min\{t, X_t \neq X_{T_{n-1}}\}, n \geq 1$ . and  $J_n = X_{T_n}, \tau_n = T_n - T_{n-1}$ . We call  $\{X_t, t \in \mathbb{N}\}$  a discrete-time multi-state process. And  $J_n$  and  $T_n$  are  $n$ th state and corresponding transition time respectively. In the context of hypnogram modeling, we always assume there is a deterministic initial state  $J_0 = s_0$ . We also assume there is an absorbing state denoted as  $s_a$ .

The semi-markov property holds if and only if,

$$Pr\{J_n = j, \tau_n = t | J_{n-1} = i, T_{n-1}, \dots, T_1, J_0\} = Pr\{J_n = j, \tau_n = t | J_{n-1} = i, T_{n-1}\}, \forall n. \quad (2.1)$$

Further, if the joint conditional probability of  $J_n, \tau_n$  given  $J_{n-1}, T_{n-1}$  is the same for any  $n$ , we call  $\{X_t, t \in \mathbb{N}\}$  a **discrete nonhomogeneous semi-markov process(dnsmp)**.

Due to the semi-markov property, it's straightforward to verify that  $\{X_t, t \in \mathbb{N}\}$  can to specified by the conditional joint distribution of  $J_n, \tau_n$  given  $J_{n-1}, T_{n-1}$  which is denoted as,

$$p(j, t | i, T_{n-1}) := Pr(J_n = j, \tau_n = t | J_{n-1} = i, T_{n-1}). \quad (2.2)$$

An alternative to specify the process is by transition intensity function defined as,

$$\lambda_{ij}(t | T_{n-1}) := Pr(\tau_n = t, J_n = j | J_{n-1} = i, \tau_n \geq t, T_{n-1}).$$

It's straightforward to verify that the transition intensity function can derive  $p(j, t | i, T_{n-1})$ . Define the conditional survival function of  $\tau_n$  as

$$S_i(t | T_{n-1}) = Pr(\tau_n \geq t | j_{n-1} = i, T_{n-1}).$$

Then we have  $p(j, t | i, T_{n-1}) = \lambda_{ij}(t | T_{n-1})S_i(t | T_{n-1})$ .

## 2.2 Multinomial Representation

### 2.2.1 Multinomial Reponse as Member of Multivariate Exponential Family and VGM

Suppose  $\tilde{\mathbf{y}}^T = (y_1, y_2, \dots, y_m) \sim M(1, \lambda_1, \lambda_2, \dots, 1 - \sum_{k=1}^{m-1} \lambda_k)$ . Denote  $\lambda_m = 1 - \sum_{k=1}^{m-1} \lambda_k$ . The distribution function of  $\tilde{\mathbf{y}}$  is,

$$\begin{aligned} f(\tilde{\mathbf{y}}) &= \prod_{k=1}^m (\lambda_k)^{y_k} \\ &= \exp\left[\sum_{k=1}^{m-1} y_k \log(\lambda_k) + (1 - \sum_{k=1}^{m-1} y_k) \log(\lambda_m)\right] \\ &= \exp\left[\sum_{k=1}^{m-1} y_k \log\left(\frac{\lambda_k}{\lambda_m}\right) + \log(\lambda_m)\right] \end{aligned}$$

Define  $\mathbf{y}^T = (y_1, \dots, y_{m-1})$ , it belongs to the exponential family since,

$$\begin{aligned} f(\mathbf{y}, \boldsymbol{\theta}) &= \exp[\mathbf{y}^T \boldsymbol{\theta} + c(\boldsymbol{\theta})], \\ \text{where } \boldsymbol{\theta}^T &= \left(\log\left(\frac{\lambda_1}{\lambda_m}\right), \dots, \log\left(\frac{\lambda_{m-1}}{\lambda_m}\right)\right), \end{aligned}$$

here  $\boldsymbol{\theta}$  is called natural parameter, and  $c$  is a function of natural parameter in the context of exponential family. The conditional mean of  $\mathbf{y}$  is  $\boldsymbol{\mu} = (\lambda_1, \dots, \lambda_{m-1})^T$ . As long as we define the predictor part  $\boldsymbol{\eta} = (\eta_1(\mathbf{x}), \dots, \eta_{m-1}(\mathbf{x}))$ , and find a link function  $g(\boldsymbol{\mu}) = \boldsymbol{\eta}$ , we complete the parametrization of the model.

We can see that it differs from the GLM and GAM in that  $\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\mu}$  are all multivariate. We generally call it **Vector Generalized Model(VGM)** for we cannot decide yet whether the predictor part  $\boldsymbol{\eta}$  is linear w.r.t  $\mathbf{x}$ .

### 2.2.2 Multinomial Representation of Discrete SMP

Consider **one day's** sleep trajectory  $\{J_n, T_n\}_{n=0}^N$ . Define  $\tau_n = T_n - T_{n-1}$ . The likelihood is,

$$\begin{aligned} \mathcal{L} &= \prod_{n=1}^N p(J_n, \tau_n | T_{n-1}, J_{n-1}) = \prod_{n=1}^N \lambda_{J_{n-1}J_n}(\tau_n | T_{n-1}) S_{J_{n-1}}(\tau_n | T_{n-1}) \\ &= \prod_{n=1}^N \lambda_{J_{n-1}J_n}(\tau_n | T_{n-1}) \prod_{t \leq \tau_n - 1} \left[1 - \sum_{k \neq J_{n-1}} \lambda_{J_{n-1}k}(t | T_{n-1})\right]. \end{aligned}$$

The second equation holds due to the semi-markov property. And the third equation holds due to the relationship between survival and hazard function.

Now define  $Y_{nt}$  as,

$$Y_{nt} = \begin{cases} J_n & \text{if } \tau_{n+1} > t, \\ k & \text{if } \tau_{n+1} = t \text{ and } J_{n+1} = k. \end{cases}$$

Define  $\mathbf{y}_{nt}$  to be the one-hot encoder of  $Y_{nt}$  where,

$$\mathbf{y}_{nt}^T = (y_{nt0}, y_{nt1}, \dots, y_{ntk}, \dots, y_{ntm}) = (0, \dots, 1, \dots, 0),$$

where  $y_{ntk} = 1$  for  $Y_{nt} = k$ .

Using  $\mathbf{y}_{nt}$ , we can rewrite the likelihood as

$$\mathcal{L} = \prod_{n=1}^N \prod_{t=1}^{\tau_n} \left\{ \left[ \prod_{\substack{k=1 \\ k \neq J_{n-1}}}^m \lambda_{J_{n-1}k}(t|T_{n-1})^{y_{n-1tk}} \right] \left[ 1 - \sum_{\substack{k=1 \\ k \neq J_{n-1}}}^m \lambda_{J_{n-1}k}(t|T_n)^{y_{n-1t0}} \right] \right\}$$

The final step is to swap the order of multiplication. We introduce indicator  $\mathbf{1}_{J_n=i}$  and  $I_i$  represents the index set of all  $n \in \{0, 1, \dots, N-1\}$  which satisfies that  $J_n = i$ . Then write likelihood as  $\mathcal{L} = \prod_{i \in [m]} \mathcal{L}_i$ , where

$$\begin{aligned} \mathcal{L}_i &= \prod_{n=1}^N \left\{ \prod_{t=1}^{\tau_n} \left[ \prod_{\substack{k=1 \\ k \neq J_{n-1}}}^m \lambda_{ik}(t|T_{n-1})^{y_{n-1tk}} \right] \left[ 1 - \sum_{\substack{k=1 \\ k \neq J_{n-1}}}^m \lambda_{J_{n-1}k}(t|T_{n-1})^{y_{n-1t0}} \right] \right\}^{\mathbf{1}_{J_{n-1}=i}} \\ &= \prod_{n \in I_i} \mathcal{L}_n^i \\ \mathcal{L}_n^i &= \prod_{t=1}^{\tau_{n+1}} \left\{ \left[ \prod_{\substack{k=1 \\ k \neq i}}^m \lambda_{ik}(t|T_n)^{y_{ntk}} \right] \left[ 1 - \sum_{\substack{k=1 \\ k \neq i}}^m \lambda_{ik}(t|T_n)^{y_{nti}} \right] \right\}. \end{aligned} \quad (2.3)$$

Firstly, since the likelihood can be decomposed into different  $\mathbf{L}_i$ , and each  $\mathbf{L}_i$  will not contain joint parameters as we illustrate later. We can deal with them separately.

Furthermore, We can see that for each  $n \in I$ , equation 2.3 is the same as the likelihood for the  $\tau_{n+1}$  observations  $\mathbf{y}_{n1}, \dots, \mathbf{y}_{n\tau_{n+1}}$  of a multinomial response model. The indicator  $y_{ntk}$  variables actually represent the distributions given that a specific epoch( $t$ ) is reached. Given that after system reaches  $J_n = i$  and exactly  $t$  epoch has passed, then the response is multinomially distributed with  $\mathbf{y}_{nt}^T = (y_{nt1}, \dots, y_{nti}, \dots, y_{ntm}) \sim M(1, \lambda_{i1}(t|T_n), \dots, 1 - \sum_{\substack{k=1 \\ k \neq i}}^m \lambda_{ik}(t|T_n), \dots, \lambda_{im}(t|T_n))$ .

In conclusion, now for each  $i$  we have independent response variable  $\{Y_{nt}\}_{n \in I, t \in [\tau_{n+1}]}$  or equally  $\{\mathbf{y}_{nt}\}_{n \in I, t \in [\tau_{n+1}]}$ . And its corresponding covariate(or predictor) is  $t$  and  $T_n$  as illustrated in equation 2.3. We use Vector Generalized Model in section 2.2.1 to model the regression relationship between  $Y_{nt}$  and  $(t, T_n)$ .

### 2.2.3 Vector Generalized Semi-parametric Parametrization

To simplify notation, we only show how we parameterize samples from  $I_m$ . Now we parameterize the model by designing the predictor part  $\boldsymbol{\eta}_{nt}^m = (\eta_1^m(t, T_n), \dots, \eta_{m-1}^m(t, T_n))^T$ . The most generalized way is to assume  $\eta_k^m$  is an unspecified function of  $t$  and  $T_n$ . However, it's redundant, especially for  $t$  to perceive it as a continuous variable thus incorporating nonlinear component w.r.t.  $t$ . Notice in reality,  $t$  represents the number of episodes system stays after it jumps to a certain state, therefore  $t$  only takes a finite number of values. We regulate the model and assume support of  $\tau_n$  is finite  $\{1, 2, \dots, r\}$ . Therefore we can perceive  $t$  as a categorical variable and further one-hot-encode it as  $\mathbf{t} \in \{0, 1\}^{r-1}$ .

For  $T_n$ , there's no need to assume linearity for  $\eta_k^m$  w.r.t.  $T_n$ , and we incorporate nonlinear part into it. Finally, we parameterize the predictor part as,

$$\eta_k^m(t, T_n) := \mathbf{t}^T \boldsymbol{\beta}_k^m + f_k^m(T_n), k = 1, \dots, m-1$$

The predictor is partially linear w.r.t. the covariate  $\mathbf{x}_{nt} = (\mathbf{t}^T, T_n)^T$ , and is semi-parametric. So we classify our model in the literature of **Vector Generalized Semi-parametric(or partially linear) model**.

Finally, corresponding the standard VGM modeling introduced in section 2.2.1, we only need to specify the link function which is the canonical link,

$$\begin{aligned} g &= h^{-1} \\ h(\boldsymbol{\eta}) &= (h_1(\boldsymbol{\eta}), \dots, h_{m-1}(\boldsymbol{\eta}))^T \\ h_k(\boldsymbol{\eta}) &= \frac{\exp(\eta_k)}{1 + \sum_{j=1}^{m-1} \exp(\eta_j)} \end{aligned}$$

We summarise our model as follows, for all samples  $\{Y_{nt}, \mathbf{x}_{nt}\}_{n \in I_i, t \in [\tau_{n+1}]}$ , we have,

$$\begin{aligned} Pr(Y_{nt} = k) &= \begin{cases} \lambda_{ik}(t|T_n), & \text{if } k \neq i \\ 1 - \sum_{j=1}^m \lambda_{ij}(t|T_n), & k = i \end{cases}, \text{ if } t = 1, 2, \dots, \tau_n \\ \lambda_{ik}(t|T_n) &= h_k(\boldsymbol{\eta}_{nt}^i), \\ \boldsymbol{\eta}_{nt}^i &= (\eta_1^i(\mathbf{t}, T_n), \dots, \eta_{i-1}^i(\mathbf{t}, T_n), \eta_{i+1}^i(\mathbf{t}, T_n), \dots, \eta_{m-1}^i(\mathbf{t}, T_n))^T \\ \eta_k^i(\mathbf{t}, T_n) &= \mathbf{t}^T \boldsymbol{\beta}_k^i + f_k^i(T_n), k \neq i \\ h_k(\boldsymbol{\eta}_{nt}^i) &= \frac{\exp[\eta_k^i(\mathbf{t}, T_n)]}{1 + \sum_{j=1, j \neq i}^m \exp[(\eta_j^i(\mathbf{t}, T_n))]} \end{aligned} \quad (2.4)$$

The log-likelihood of our model for each starting state  $i$  is,

$$l_i = \sum_{n \in I_i} \sum_{t=1}^{\tau_{n+1}} \left\{ \sum_{\substack{k=1 \\ k \neq i}}^m y_{ntk} [\log[h_k(\boldsymbol{\eta}_{nt}^i)]] + y_{nti} \log\left(1 - \sum_{\substack{j=1 \\ j \neq i}}^m h_j(\boldsymbol{\eta}_{nt}^i)\right) \right\}.$$

## 2.2.4 Discussion

We categorize our model into **Vector Generalized Semi-parametric Model**. In statistical literature, generalized linear model is a widely used model for regression of non-gaussian response variables. Soon the model is extended to the setting where covariates have nonlinear effect and generalized additive model(GAM) is brought up as a simple way to deal with smoothing multiple predictors. However, **our model is between them for we both have linear and nonlinear component**. In statistical literature, it is mostly referred to generalized partially nonlinear model or generalized semi-parametric model.

Another thing to notice is that most of the GLM or GAM or GPLM literature considers the case when the response variable is univariate, however, a multinomial distribution is a special case in exponential family because its response is multivariate. Therefore, our model **also differs in that the response is a vector**.

## 2.3 Estimation

Since we introduce nonlinear(or nonparametric) part in equation 2.4, we have to consider how to estimate  $\mathbf{f}_k^i$  (here  $i$  is the subscript for different starting state and  $k$  is the state system jumps to). When estimating functions, it involves literature of **smoothing**. From my knowledge, there are

Here, we talk about two broad categories of smoothers:

- Regression or series smoothers(regression spline, polynomial series).
- Smoothing spline.

### 2.3.1 Regression Spline

In the first category, we show how to estimate the model through regression spline, which is also the simplest way in three categories.

To the best of our knowledge, regression spline is simply approximating  $f_k^i$  with a linear combination of spline bases  $\sum \gamma_s B_s()$ . We simply evaluate  $B_s()$  in given points and incorporate them into design matrix. In conclusion, estimation procedure using regression is the same as Vector GLM.

For vector GLM, the estimation of parameters is through Iteratively Reweighted Least Square(IRLS)

### 2.3.2 Smoothing Spline

Different from regression spline, smoothing spline add smoothing penalty to the loss function and estimate the coefficient.

## 2.4 Mean Pattern Derivation

After we estimate all the parameters and have the whole probability model in hand, our next job is to derive a mean pattern to summarise the noisy data.

We define the mean pattern to be a predicted hypnogram denoted with  $\{\hat{J}_h, \hat{T}_h\}_{h=0}^{\hat{H}}$ . However, since  $J_h$  is a discrete random variable that does not take value in  $\mathbb{R}$ , it's not trivial to define its mean value. Therefore, we combine the idea of maximum probability and conventional mean to derive the mean pattern. The mean pattern is computed by first setting  $\hat{J}_0 = s_0$  and  $\hat{T}_0 = 0$ , and then compute,

$$\hat{J}_h = \underset{j}{\operatorname{argmax}} \operatorname{Pr}(J_h = j | J_{h-1} = \hat{J}_{h-1}, T_{h-1} = \hat{T}_{h-1}). \quad (2.5)$$

Then compute

$$\hat{\tau}_h = \mathbb{E}[\tau_h | J_h = \hat{J}_h, J_{h-1} = \hat{J}_{h-1}, T_{h-1} = \hat{T}_{h-1}]. \quad (2.6)$$

Then  $\hat{T}_h = \hat{T}_{h-1} + \hat{\tau}_h$ . And the mean pattern is ended when the predicted trajectory reaches absorbing state  $s_a$ .

## 3 Real Data Studies

Our data is collected from 17 women wearing oura ring during their pregnancy. Our data consists of consecutive sleep observations from observation of up to 12 months.

We fit the model using the data we gathered. In implementation,  $\beta_{ij}(\cdot)$  is estimated using B-spline functions. After acquiring estimation, we adopt the method in section 2.4 and derive the mean pattern. The result is shown in figure 1. When drawing this plot, we did not include covariate into the model.

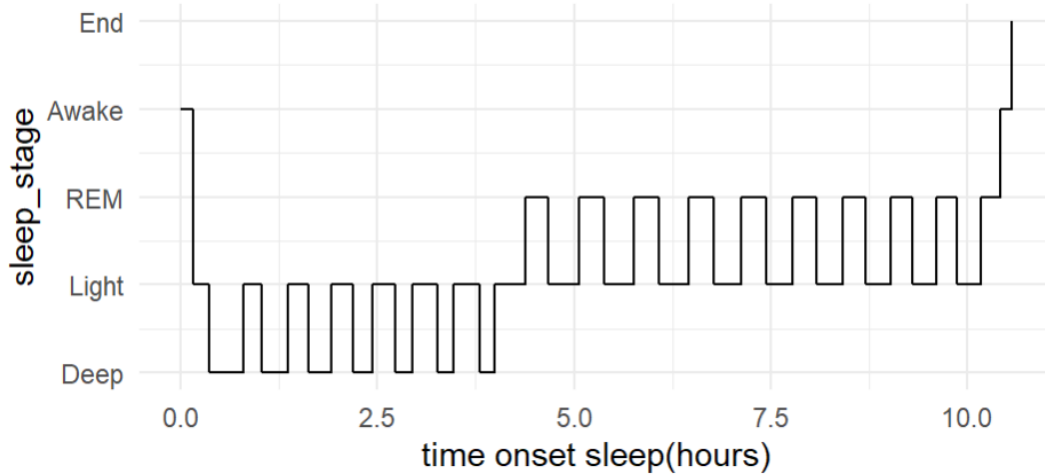


Figure 1: Predicted Mean Pattern

The mean pattern is simple and representative compared to the rather noisy original data. We draw three conclusions from the fitted mean pattern. 1. At early sleep process, one alternates between light and deep sleep for several times. 2. During later half of the sleep process, one alternates between light and REM stage for several times. 3. The duration of each sleep stage undergoes gradual and monotonous change. The duration of deep sleep is increasing while the sleep duration of light sleep first increase then decrease.

We also fit the mean pattern for different phases of pregnancy. We use data from 10-19, 20-29, 30-39 months of pregnancy separately to fit model and to derive mean pattern. The results are shown in figure 2.

We observe that as time passes, pregnant women’s sleep quality is deteriorating as they have less deep sleep. They also have trouble getting into deep sleep as observed from the mean pattern of 30-39 weeks of pregnancy.

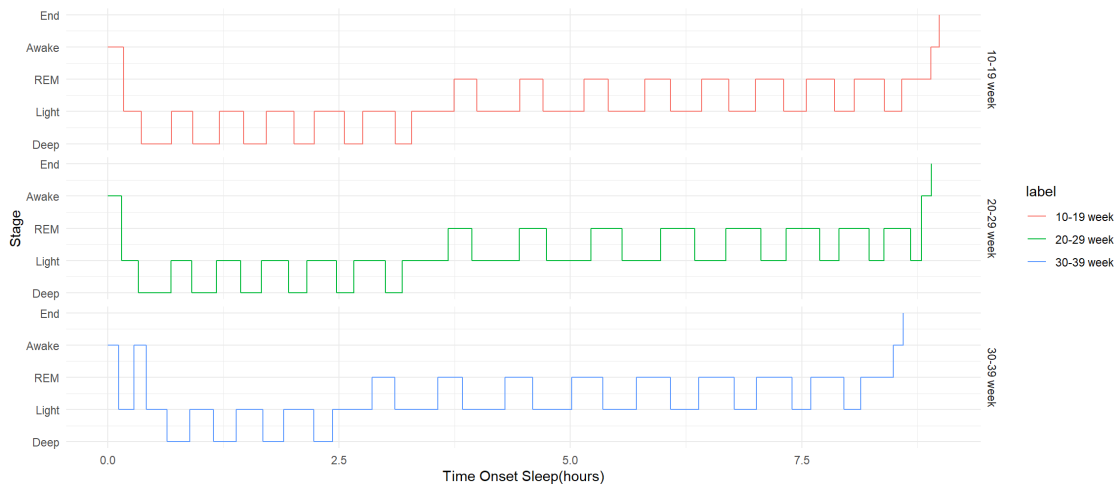


Figure 2: Change of Mean Pattern During Pregnancy

Finally, we fit mean pattern of different individuals. It turns out that there is strong heterogeneity from different individual’s sleep, however they all follow the overall mean pattern which is

first deep and light sleep, and then light and REM sleep. Serious deviations from this pattern may indicate abnormal health conditions. For example, as figure 3 has shown, individual 19 has abnormal mean pattern, which might be explained by the fact that she is over 30 years old and is severely overweight.

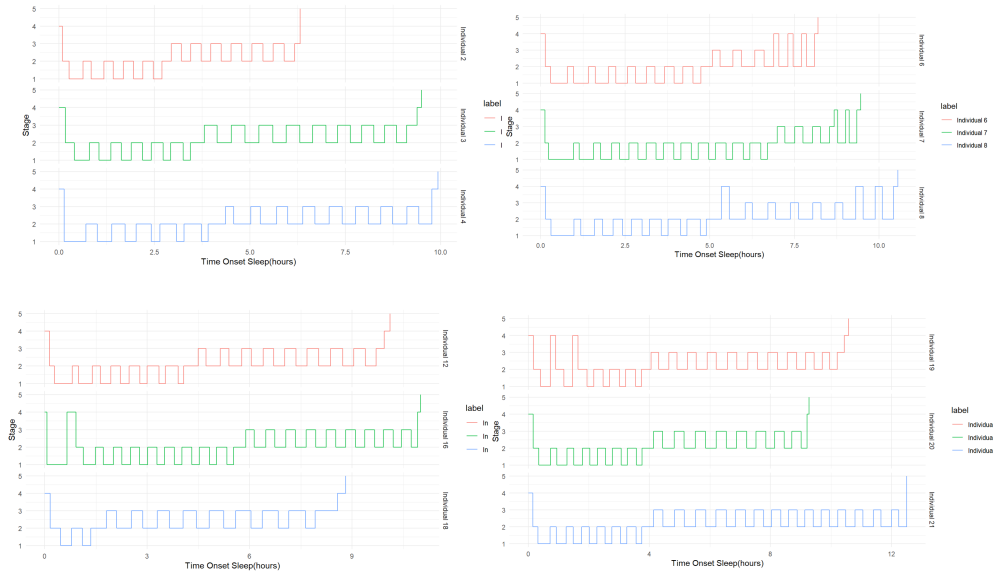


Figure 3: Individual Mean Patterns

## 4 Conclusion and Discussions

The conclusion and contribution of our work are as follows. Firstly, we give a simple definition to nonhomogeneous discrete semi-markov process. We prove its likelihood can be decomposed into binary regression through proper approximation. Secondly, we use a novel parametrization for nonhomogeneity which can separate the baseline transition intensity and the effect of nonhomogeneity meanwhile can be efficiently estimated. Third, we propose a method of deriving mean pattern and our empirical results are consistent with existing medical sleep studies.

Our drawbacks and future work are as follows: the estimation’s consistency and inference have not been discussed in this report. Though Lee et al. (2018) gave a consistency theory in their paper, which demonstrates the estimation is asymptotic normal and unbiased, and they use a sandwiched estimator for the variance. However, their proof directly neglects the approximation error of the likelihood therefore is not the true error under the proposed model. Our future work can include deriving the comprehensive consistency theory for such estimation.

## References

- ANDERSEN, P. K. and KEIDING, N. (2002). Multi-state models for event history analysis. *Statistical methods in medical research* **11** 91–115.
- ASANJARANI, A., LIQUET, B. and NAZARATHY, Y. (2021). Estimation of semi-markov multi-state models: a comparison of the sojourn times and transition intensities approaches. *The international journal of biostatistics* **18** 243–262.



- FAHRMEIR, L. and KAUFMANN, H. (1987). Regression models for non-stationary categorical time series. *Journal of time series Analysis* **8** 147–160.
- FAHRMEIR, L. and KLINGER, A. (1998). A nonparametric multiplicative hazard model for event history analysis. *Biometrika* 581–592.
- FOKIANOS, K. and KEDEM, B. (1998). Prediction and classification of non-stationary categorical time series. *Journal of multivariate analysis* **67** 277–296.
- KANG, M. and LAGAKOS, S. W. (2007). Statistical methods for panel data from a semi-markov process, with application to hpv. *Biostatistics* **8** 252–264.
- LEE, M., FEUER, E. J. and FINE, J. P. (2018). On the analysis of discrete time competing risks data. *Biometrics* **74** 1468–1481.
- STOFFER, D. S., TYLER, D. E. and MCDUGALL, A. J. (1993). Spectral analysis for categorical time series: Scaling and the spectral envelope. *Biometrika* **80** 611–622.
- SWIHART, B. J., CAFFO, B., BANDEEN-ROCHE, K. and PUNJABI, N. M. (2008). Characterizing sleep structure using the hypnogram. *Journal of Clinical Sleep Medicine* **4** 349–355.