Temporal Point Process

- A temporal point process is a mathematical model of temporally discrete events such as earthquakes, financial transactions, communication in a social network, user activity at a web site, and so on.

- Conditional Intensity function characterizes the temporal point process:
  \[ \lambda(t|H_t) = \frac{P(\text{one event occurs in } [t, t+\Delta t]|H_t)}{\Delta t} \]

- Parametric models of the conditional intensity function:
  - Poisson process: \( \lambda(t|H_t) = \mu \)
  - Hawkes process: \( \lambda(t|H_t) = \mu + \sum_{s \in H_t} \phi(t-s) \)
  - Self-correcting process: \( \lambda(t|H_t) = \exp\left[ \sum_{s \in H_t} \phi(t-s) \right] \)

RNN based Approach to Temporal Point Processes

- The RNN is used to encode the event history. Then the conditional intensity function is formulated via the hazard function as \( \lambda(t|H_t) = \phi(t - t_i|H_t) \).

- The log-likelihood function of the model is given as
  \[ \log L \left( \{ t_i \} \right) = \sum \log \phi(t_{i+1} - t_i|H_t) - \int_0^{t_{i+1}} \phi(t|H_t) dt \]

- [Problem] The hazard function is usually modeled by a specific parametric function. However, such assumption can limit the expressive power of the model.

- (Exponential hazard function [2]) \( \phi(t|H_t) = \exp(w + \theta_1 + \theta_2) \)
  - Constant hazard function [3] \( \phi(t|H_t) = \exp(\theta_1 + \theta_2) \)

- [Our study] We propose a novel RNN based model.
  - Our model can flexibly model the time evolution of the hazard function.
  - The log-likelihood function of our model can be exactly evaluated without any numerical approximations, so the model can be efficiently trained.

Proposed model

- [Problem] If we use a complex model for the hazard function, the log-likelihood function cannot be exactly evaluated because it contains the integral of the hazard function. It is intractable to evaluate the integral using numerical methods.

- [Our approach] Instead of the hazard function, we model the cumulative hazard function,
  \[ \Phi(t|H_t) = \int_0^t \phi(s|H_s) ds \]

- The hazard function is given by differentiating the cumulative hazard function,
  \[ \phi(t|H_t) = \frac{\partial}{\partial t} \Phi(t|H_t) \]

- The log-likelihood function of the model is reformulated as
  \[ \log L \left( \{ t_i \} \right) = \sum \log \left[ \phi(t_{i+1} - t_i|H_t) + \Phi(t_{i+1} - t_i|H_t) \right] \]

- This can be exactly evaluated even for a complex model of the cumulative hazard function!

- The feedforward neural network model of the cumulative hazard function

  The cumulative hazard function is a monotonically increasing function of \( t \), which can be reproduced by constraining the particular network connections to be positive [1].

### References