Flexible Mixture Modeling Approaches to Renewal Processes

Advancement to Candidacy Presentation
UCSC Statistical Science Program

Zach Horton
July 28, 2022

Presentation Outline

Renewal Process Basics

Renewal Process Basics

What is a renewal process?

- A type of model for recurring events
- Characterized by good-as-new behavior after an event (homogeneous)
- Examples: printer replacement, earthquakes, neuron firing patterns.

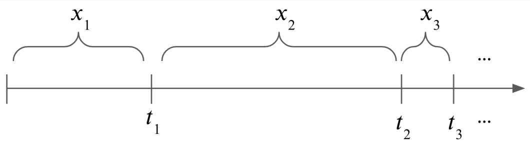


Fig. 1: Relating events and inter-arrival times.

Renewal Process Basics

What is a renewal process?

- A type of model for recurring events
- Characterized by good-as-new behavior after an event (homogeneous)
- Examples: printer replacement, earthquakes, neuron firing patterns.

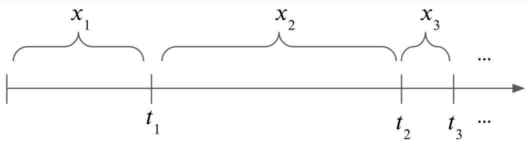


Fig. 1: Relating events and inter-arrival times.

Notation:

- $t \in (0, T)$: Observation time window.
- $N(t)$: Number of events up to time t.
- $N(T) = n$: Total number of events.
- $\boldsymbol{t} = (t_1, \ldots, t_n)$: Event times.
- $\boldsymbol{x} = (x_1, \ldots, x_n)$: Elapsed times or *inter-arrival times*
- Notice $x_i = t_i - t_{i-1}$ and $t_i = \sum_{\ell=1}^{i} x_\ell$, with $t_0 = 0$

Renewal Process Basics

NOAA Significant Earthquakes Data Set:

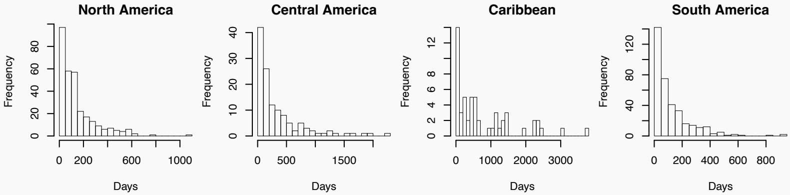


Fig. 2: Inter-arrival times in days for NOAA earthquake data

- Contains earthquakes occurring between Jan. 1, 1900 and Dec. 31, 2018
- Subset only includes earthquakes occurring in regions of the Americas
- "Significant": At least \$1M in damages, magnitude 7.5 or higher, 10 deaths or more, Modified Mercalli Intensity X, or generated a tsunami.
- Longest inter-arrival: Puerto Rico, Nov. 12 1918 to coast of Venezuela, Jan. 17, 1929.

Renewal Process Basics

When studying renewal process data, 3 distributional functions are often of interest:

1. Density Function $f(x)$
2. Hazard Function $h(x)$
3. Renewal Function $M(t)$ and/or K-function $K(t)$

Renewal Process Basics

The inter-arrival density or PDF $f(x), x \in \mathbb{R}^+$:

- Describes the inter-arrival histogram shape
- CDF $F(x) = Pr(X \leq x) = \int_0^x f(u)du$
- Survival function $S(x) = Pr(X > x) = 1 - F(x)$
- Finite mean restriction: $\int_0^\infty xf(x)dx < \infty$

Some examples:

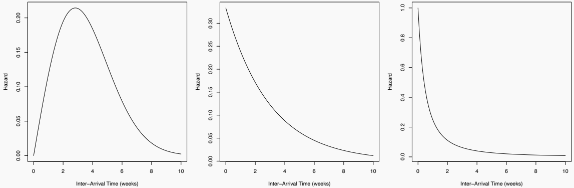


Fig. 3: Inter-arrival densities: Weibull (left), Exp (mid), Lomax (right). Note the Weibull case displays shape > 1. Using shape < 1 leads to decreasing density.

Renewal Process Basics

The inter-arrival hazard $h(x)$

- Describes instantaneous risk of an event happening.
- *Similarly shaped densities can have differently shaped hazards!*
- $h(x) = f(x)/S(x)$
- Cumulative hazard $H(x) = \int_0^x h(u)du$

Some examples:

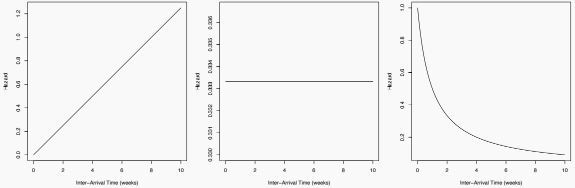


Fig. 4: Inter-arrival hazards: Weibull (left), Exp (mid), Lomax (right)

Renewal Process Basics

The renewal function $M(t)$

- Measures expected number of events: $M(t) = E[N(t)]$
- $\mathcal{L}_M(s) = \frac{\mathcal{L}_f(s)}{s(1-\mathcal{L}_f(s))}$, where $\mathcal{L}_g(s) = \int_0^\infty e^{-xs} g(x)dx$
- K-function is standardized: $K(t) = M(t) \times E[X]$
- $K(t) - t$ describes (de)clustering relative to exponential

Some $K(t) - t$ examples:

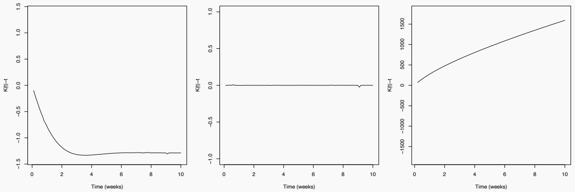
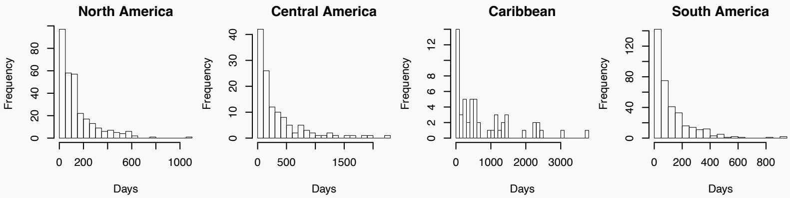


Fig. 5: Inter-arrival K-function: Weibull (left), Exp (mid), Lomax (right)

Research Objectives



Objectives:

(1) Develop a flexible model framework that can adapt to whatever functional shapes are supported by the data.

(2) Build a more structured model suitable for decreasing shape-constrained applications such as earthquakes.

(3) Incorporate model extensions which smoothly evolve over time.

Approach is two-pronged: build mixture models for both the density and hazard.

Mixture Models for the Inter-Arrival Density

Mixture Models for the Inter-Arrival Density

Basic idea: Dirichlet Process Mixture Model (DPMM) framework,

$$f(x|\boldsymbol{\omega}, \boldsymbol{\theta}) = \sum_{j=1}^{\infty} \omega_j k(x|\theta_j)$$

$$\theta_j \overset{iid}{\sim} G_0$$

$$\omega_j = \nu_j \prod_{\ell=1}^{j-1} (1 - \nu_\ell)$$

$$\nu_\ell \overset{iid}{\sim} Beta(1, \alpha)$$

What should we use for kernel $k(x|\theta)$ and prior G_0?

Mixture Models for the Inter-Arrival Density

Kernel $k(x|\theta)$ and Prior G_0 considerations:

- Finite mean conditions
  1. $E_k(\theta) = \int_0^\infty x k(x|\theta) dx < \infty$ for all $\theta \in \Theta$
  2. $\int_\Theta E_k(\theta) dG_0(\theta) < \infty$
- Mixture shape flexibility
- CDF available in closed form (likelihood and hazard calculations)
- Laplace transform available in closed form (K-function evaluation)
- Conditional conjugacy (convenient sampling)

Mixture Models for the Inter-Arrival Density

Kernel $k(x|\theta)$ and Prior G_0 considerations:

- Finite mean conditions
  1. $E_k(\theta) = \int_0^\infty x k(x|\theta) dx < \infty$ for all $\theta \in \Theta$
  2. $\int_\Theta E_k(\theta) dG_0(\theta) < \infty$
- Mixture shape flexibility
- CDF available in closed form (likelihood and hazard calculations)
- Laplace transform available in closed form (K-function evaluation)
- Conditional conjugacy (convenient sampling)

The gamma distribution is a promising choice for modeling unrestricted, continuous density shapes.

We went a different direction, focusing on *decreasing* densities using the uniform distribution kernel.

- Meets requirements
- Closed form expressions
- Empirical support in earthquake literature
- Unexplored for renewal models

Mixture Models for the Inter-Arrival Density

Hierarchical model with membership variable augmentation:

$$
\begin{aligned}
& x_i | z_i, \boldsymbol{\theta}, N(T) \overset{iid}{\sim} U(0, \theta_{z_i}) && i = (1, \ldots, n) \\
& P(N(T) = n | z_{n+1}, \boldsymbol{\theta}) = 1 - F_u(T - t_n | \theta_{z_{n+1}}) && \\
& P(z_i = j | \boldsymbol{\omega}) = \omega_j && i = (1, \ldots, n+1); j = (1, \ldots, J) \\
& \quad \boldsymbol{\omega} | \alpha \sim GDir_J((1, \ldots, 1), (\alpha, \ldots, \alpha)) && \\
& \quad \theta_j | b_\theta \overset{i.i.d.}{\sim} IG(a_\theta, b_\theta) && \\
& \quad b_\theta \sim Ga(a_0, b_0); \quad \alpha \sim Ga(a_\alpha, b_\alpha) &&
\end{aligned}
$$

where $a_\theta > 1, a_\alpha, b_\alpha, a_0, b_0$ are all fixed hyperparameters and the stick-breaking prior is reduced to $GDir$ after truncating to J components.

Mixture Models for the Inter-Arrival Density

Gibbs sampling outline:

- α and b_θ are conditionally conjugate gamma
- $\boldsymbol{\omega}$ is conditionally conjugate generalized Dirichlet
- z_i is discrete with unnormalized weights $\omega_j U(x_i|0, \theta_j)$
- θ_j for $z_{n+1} \neq j$ is truncated inverse-gamma
- θ_j for $z_{n+1} = j$ is truncated inverse-gamma times the $T - t_n$ term. A slice sampling step can be used to draw θ_j conditionally from a new truncated inverse-gamma.

The main difference from the standard blocked-Gibbs DPMM is the slice sampling step used to handle z_{n+1}.

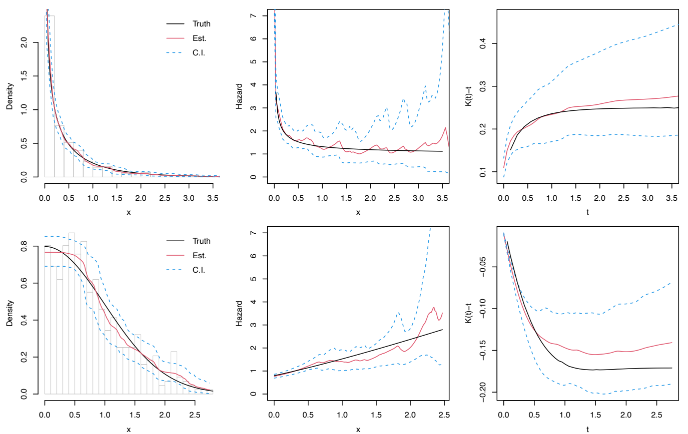


Fig. 6: Uniform mixture model fit to simulated gamma inter-arrival data (top row) and half-normal inter-arrival data (bottom row). Both true and fitted versions of the density (left), hazard (middle), and K-function (right) are displayed along with 95% posterior intervals.

Mixture Models for the Inter-Arrival Density: Data Analysis

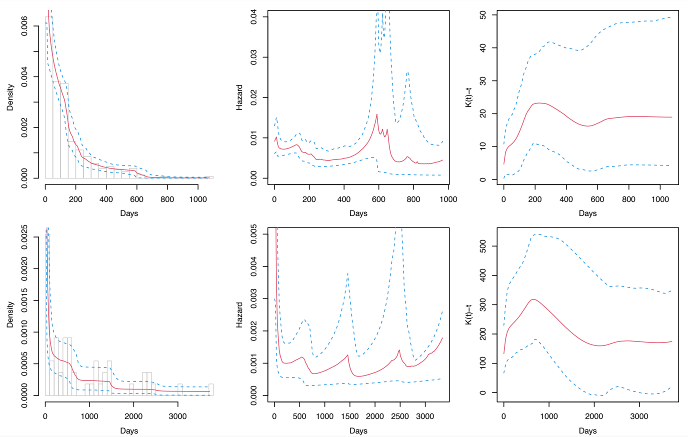


Fig. 7: Uniform mixture model fit to NOAA earthquake occurrence data in the North American region (top row) and Caribbean region (bottom row). Displayed are the fitted density (left), hazard (middle), and K-function (right) along with 95% posterior intervals. Compare to Xiao et al. (2021)

Mixture Models for the Inter-Arrival Density

Uniform mixture model satisfies our objectives:

- ✓ Flexible within the decreasing shape constraint
- ✓ Computationally manageable

Mixture Models for the Inter-Arrival Density

Uniform mixture model satisfies our objectives:

- ✓ Flexible within the decreasing shape constraint
- ✓ Computationally manageable
- ✓ Can be extended for time-dependence

Rather than being characterized by a single inter-arrival distribution $f(x) = f(t - \mathcal{H}(t))$ where $\mathcal{H}(t) = t_{N(t)}$ is the time of the most recent event, an inhomogeneous renewal process is characterized by the collection $\{f_{\mathcal{H}(t)} : \mathcal{H}(t) \in [0, T])\}$.

Mixture Models for the Inter-Arrival Density

We propose incorporating time-dependence into the weights:

$$\omega_j(\mathcal{H}(t)) = \nu_j(\mathcal{H}(t)) \prod_{\ell=1}^{j-1} (1 - \nu_\ell(\mathcal{H}(t))$$

$$\log\left(\frac{\nu_j(\mathcal{H}(t))}{1 - \nu_j(\mathcal{H}(t))}\right) = \beta_{0j} + \beta_{1j}\mathcal{H}(t)$$

Features:

- Preserves computational structure for the θ_j parameters.
- Corresponds to stagnant components with dynamic membership.
- Pólya-Gamma augmentation can be used to easily sample β parameters
- Linear form w.r.t. $\mathcal{H}(t)$ can be extended through polynomials or other covariates
- Contains homogeneous renewal model as special case

Basis Mixture Models for the Inter-Arrival Hazard

Basis Mixture Models for the Inter-Arrival Hazard

Mixture model placed directly on the hazard function:

$$h(x|\boldsymbol{\omega}, \theta, \boldsymbol{c}) = \sum_{j=1}^{J} \omega_j b_{c_j}(x|\theta)$$

More similar to linear combinations than mixtures:

- b = known basis function
- ω_j = positive coefficient
- c_j = knot location
- θ = common dispersion parameter

What should we use for basis functions? How do we model the weights?

Basis Mixture Models for the Inter-Arrival Hazard

Desirable properties for a basis set:

- Non-negative with infinite integral
- Basis mixture shape flexibility

For given knot location c and dispersion θ, we define the basis:

$$b_c(x|\theta) = \frac{\theta}{c} h\left(c, \frac{c}{\theta}\right) = \frac{1}{x}\left(\frac{1}{1 + \left(\frac{c}{x}\right)^{\frac{c}{\theta}}}\right)$$

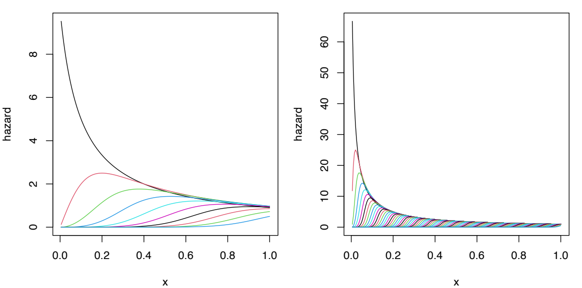


Fig. 8: Log-Logistic basis for $J = 10$ (left) and $J = 100$ (right) with $\theta = \frac{1}{J}$ and evenly spaced knots $c_j = \frac{j}{J}$.

where $h(x|a,b)$ is the log-logistic hazard function: $\frac{b}{a}\left(\frac{x}{a}\right)^{b-1}\left(1 + \left(\frac{x}{a}\right)^b\right)^{-1}$ with corresponding cumulative hazard $H(x|a,b) = \log\left(1 + (x/a)^b\right)$.

Basis Mixture Models for the Inter-Arrival Hazard

Features of the log-logistic basis:

- Satisfies infinite integral requirement
- Primarily unimodal shape (similar to B-splines and Erlang mixtures)
- Closed form expressions
- A convergence result:

$$h(x) = \lim_{\theta \to 0^+} \lim_{J \to \infty} \sum_{j=1}^{J} \frac{R}{J} \phi^*(c_j) b_{c_j}(x|\theta)$$

$$R = \text{upper bound for } x$$

$$c_j = \frac{j}{J} R; \quad \omega_j = \frac{R}{J} \phi^*(c_j); \quad \phi^*(x) = \frac{d}{dx} x h(x)$$

This suggests modeling the weights as evaluations along a function $\omega_j = \phi(c_j) = \frac{R}{J} \phi^*(c_j)$.

Basis Mixture Models for the Inter-Arrival Hazard

Prior model for the weights:

$$\log \phi(x) \sim GP(\mu(x), k(x, x'))$$
$$\mu(x) = \beta_0 + \beta_1 \log(x) \quad \Longleftarrow \text{ centered on Weibull}$$
$$k(x, x') = \sigma^2 \exp\left(-\frac{(x - x')^2}{\gamma}\right)$$

Additional details:

- Set $R = \max(T - t_n, x_1, \dots, x_n)$.
- We use $c_j = \frac{jR}{J}$, although other grids can be used.
- Fix $\theta = \frac{R}{J}$, which simplifies computation and may be justified theoretically.
- A sufficient condition for finite mean is $\sum \omega_j > 1$, which is frequently satisfied but difficult to enforce directly.
- Truncating $\beta_1 > -1$ is consistent with Weibull centering

Basis Mixture Models for the Inter-Arrival Hazard

Hierarchical model with membership variable augmentation:

$$p(\boldsymbol{x}|\boldsymbol{\omega}, \boldsymbol{z}) \propto \exp\left(-\sum_{j=1}^{J} \omega_j B_j^*\right) \prod_{i=1}^{n} \Omega b_{c_{z_i}}(x_i|\theta)$$

$$P(z_i = j|\boldsymbol{\omega}) = \omega_j/\Omega$$

$$\boldsymbol{\omega} \sim MLN(\boldsymbol{\mu} = \boldsymbol{C}\boldsymbol{\beta}, \Sigma)$$

$$\boldsymbol{\beta} \sim N(\boldsymbol{m}, \tau^2 I); \;\; \sigma^2 \sim IG(2, b_\sigma); \;\; \gamma \sim Exp(R/J)$$

- Abbreviated terms: $\Omega = \sum \omega_j$, $B_j^* = \sum_{i=1}^{n+1} B_{c_j}(x_i|\theta)$, $\boldsymbol{C} = [\mathbf{1}; \log \boldsymbol{c}]$
- $\boldsymbol{\mu}$ and Σ come from evaluating $\mu(x)$ and $k(x, x')$ along the knot vector $\boldsymbol{c}$.
- $\boldsymbol{m}, \tau^2, b_\sigma$ are fixed hyperparameters.
- The GP prior structure induces the MLN prior on the weights.

Basis Mixture Models for the Inter-Arrival Hazard

Gibbs sampling scheme:

- $\boldsymbol{\beta}$ is conditionally conjugate normal.
- σ^2 is conditionally conjugate inverse-gamma.
- γ can be sampled using M-H (as is common for GP models).
- z_i is discrete with unnormalized weights $\omega_j b_{c_j}(x_i|\theta)$.
- $p(\boldsymbol{\omega}|-) \propto MLN(\boldsymbol{\omega}|\boldsymbol{\mu}, \Sigma) \prod_{j=1}^{J} Ga(\omega_j|n_j + 1, B_j^*)$ where $n_j = |\{i : z_i = j\}|$
  - Introduce latent uniform u_j such that $p(\boldsymbol{\omega}, \boldsymbol{u}|-) \propto MLN(\boldsymbol{\omega}|\boldsymbol{\mu}, \Sigma) \times \prod_{j=1}^{J} I\{u_j < Ga(\omega_j|n_j + 1, B_j^*)\}$
  - Sample $\boldsymbol{u}$ and $\boldsymbol{\omega}$ with a Gibbs step. Each u_j is sampled from a uniform distribution.
  - The $\boldsymbol{\omega_j}$ arise from a truncated MLN: $p(\boldsymbol{\omega}|\boldsymbol{u}, -) \propto MLN(\boldsymbol{\omega}|\boldsymbol{\mu}, \Sigma) \times \prod_{j=1}^{J} I(L_j < \omega_j < U_j)$ and we sample from this using minimax-tilting algorithm from Botev (2017).

Basis Mixture Models for the Inter-Arrival Hazard: Simulations

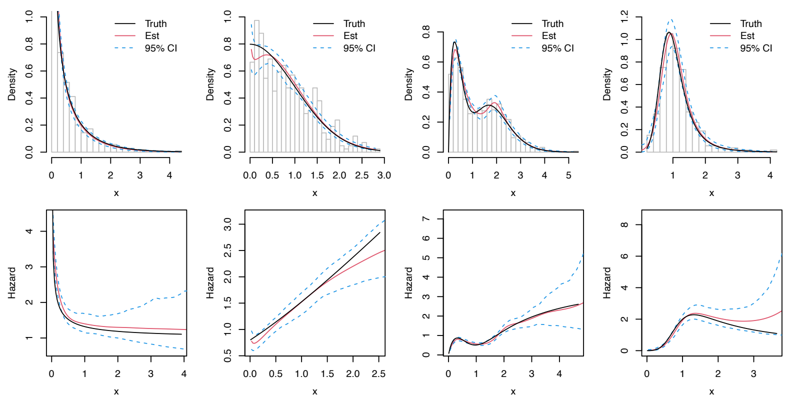


Fig. 9: Log-logistic hazard basis mixture fit to simulated inter-arrival times. Both true and fitted versions of the density (top) and hazard (bottom) are displayed along with 95% posterior intervals. Distributions from left to right: gamma, half-normal, gamma mixture, log-logistic.

Basis Mixture Models for the Inter-Arrival Hazard: Simulations

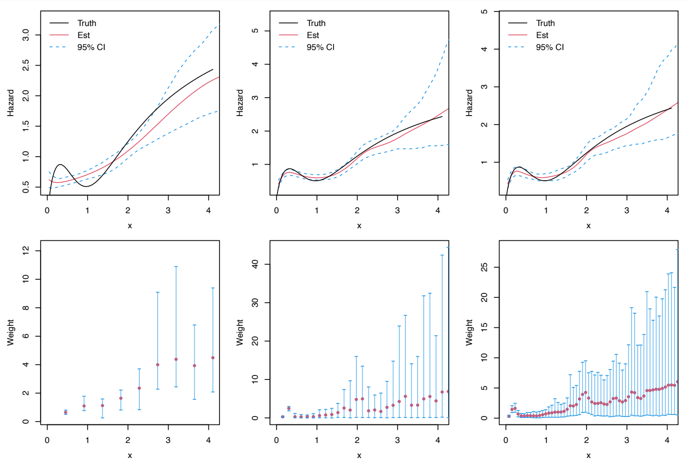


Fig. 10: Estimated hazard function (top) and mixture weights (bottom) obtained from fitting the basis mixture model to simulated gamma mixture data. Weights are placed at corresponding knot locations. From left to right the values of J are 10, 30, and 60.

Basis Mixture Models for the Inter-Arrival Hazard: Data Analysis

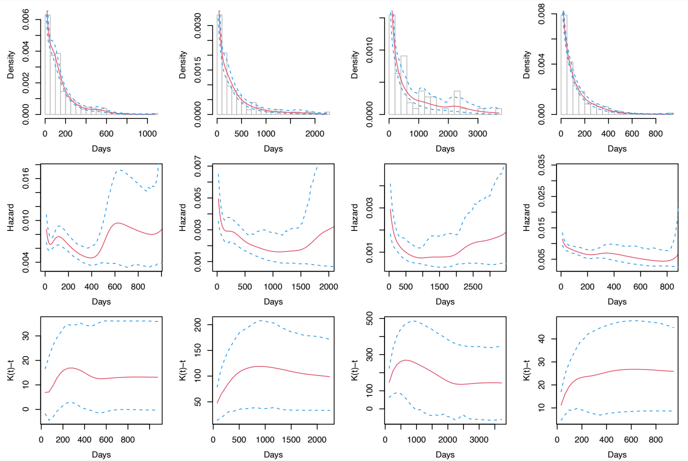


Fig. 11: Log-logistic basis mixture fit to NOAA earthquake occurrence data. Displayed are the fitted density (top), hazard (middle), and K-function (bottom) with 95% posterior intervals. Regions from left to right: North Am., Central Am., Caribbean, South Am. Compare to Xiao et al. (2021)

Basis Mixture Models for the Inter-Arrival Hazard

Features of the hazard basis mixture model:

- ✓ Scales well to larger data sets
- ✓ Flexible yet mitigates overfit
- ✓ Computationally manageable
- ✓ Applicable to more general survival analysis

Basis Mixture Models for the Inter-Arrival Hazard

Features of the hazard basis mixture model:

- ✓ Scales well to larger data sets
- ✓ Flexible yet mitigates overfit
- ✓ Computationally manageable
- ✓ Applicable to more general survival analysis

Time-dependence can be approached through a modulated renewal process:

$$\lambda(t|\mathcal{H}(t)) \equiv \lambda(t, t - \mathcal{H}(t)) = \lambda^*(t)h(t - \mathcal{H}(t))$$

where λ^* is some modulating function and h is the inter-arrival hazard.

Basis Mixture Models for the Inter-Arrival Hazard

Modulated renewal process likelihood:

$$P(t_1, \ldots, t_n | h, \lambda^*) = \prod_{i=1}^{n} \lambda^*(t_i) h(t_i - t_{i-1}) \prod_{i=1}^{n+1} \exp\left(- \int_{t_{i-1}}^{t_i} \lambda^*(t) h(t - t_i) dt \right)$$

Possible mixture model framework:

- Model the hazard using the basis mixture:

$$h(t - \mathcal{H}(t)) = \sum_{j=1}^{J} \omega_j b_{c_j}(t - \mathcal{H}(t) | \theta)$$

- For the modulating function, we use the Erlang mixture of Kim and Kottas (2022):

$$\lambda^*(t) = \sum_{\ell=1}^{L} \omega^*_\ell Ga(t | \ell, \theta^*)$$

- Employ two independent sets of latent membership variables.
- Handle the exponential via numerical integration or uniformization

Appendix

Likelihood Censoring Term

The likelihood function is often used to estimate inter-arrival distribution parameters from observed data. The homogeneous renewal process likelihood can be expressed as:

$$S(T - t_n) \prod_{i=1}^{n} f(x_i) = e^{H(T-t_n) - \sum_{i=1}^{n} H(x_i)} \prod_{i=1}^{n} h(x_i)$$

The bit coming from $T - t_n$ represents partial information about the next event t_{n+1}. Its presence usually complicates what otherwise would allow for a simple likelihood procedure.

The presented models handle this term in different ways. The uniform mixture contains an extra membership variable z_{n+1} and a slice sampling step. The hazard basis mixture adds an extra term in B_j^*.

K-Function Numerical Laplace Transform

The Laplace transform $\mathcal{L}_f(s)$ allows for complex s, that is $s = a + ib$ where a, b are real and $i = \sqrt{-1}$. When numerically constructing the Laplace transform, the following expression represents complex support in terms of real integrals:

$$\begin{aligned}
\mathcal{L}_f(s) &= \int_0^\infty e^{-xs} f(x)dx = \int_0^\infty e^{-x(a+ib)} f(x)dx \\
&= \int_0^\infty e^{-xa} e^{-ixb} f(x)dx = \int_0^\infty e^{-xa}[\cos(-xb) + i\sin(-xb)]e^{-ixb} f(x)dx \\
&= \int_0^\infty \cos(xb)e^{-xa} f(x)dx - i\int_0^\infty \sin(xb)e^{-xa} f(x)dx
\end{aligned}$$

which can be quickly evaluated using Gauss-Laguerre quadrature.

Convergence Result Proof Sketch

Prove $h(x) = \lim_{\theta \to 0^+} \lim_{J \to \infty} \sum_{j=1}^{J} \frac{R}{J} \phi(c_j) b_{c_j}(x|\theta)$:

1. Show that $h(x) = \int_0^\infty \frac{1}{x} I(x > c) \phi(c) dc$ where $\phi(c) = \frac{d}{dc} ch(c)$.
2. Show that $\lim_{\theta \to 0^+} b_c(x|\theta) = \frac{1}{x} I(x > c)$.
3. Plug 2. into 1. and assume some arbitrary upper bound R.
4. Use dominated convergence to bring limit outside of the integral.
5. Use Reimann sum with evenly spaced grid to arrive at result.

Note that uniform convergence is not possible through this proof structure, only point-wise. Also note that any knot grid system supported by general Riemann sum theory will also suffice for this result.

Comments on Fixing θ

The parameter θ controls basis dispersion, but it also influences how many bases are unimodal vs decreasing. There will be exactly $m = \lfloor \frac{\theta J}{R} \rfloor$ purely decreasing basis functions. Ensuring only 1 by fixing $\theta = \frac{R}{J}$ is consistent with both B-splines and Erlang mixtures.

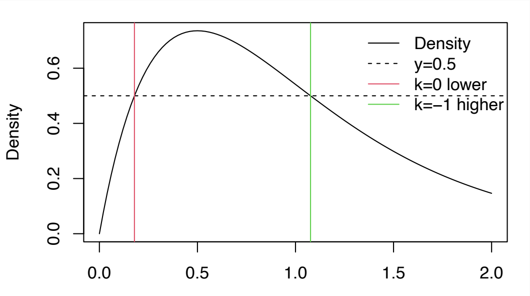
Additionally, there may be theoretical justification. The convergence result is an iterated limit. Showing that an iterated limit is equivalent to a simultaneous limit (in which setting $\theta = \frac{R}{J}$ is valid) requires some added conditions. Given the absolute continuity of the limiting argument, we may be able to prove those are met.

Inverting the Gamma Density

Inverting the gamma density to compute the bounds involves Lambert's W-function. In short, to solve for x in the equation $y = Ga(x|a, b)$, then the following expression provides the solutions:

$$x = -\frac{a-1}{b} W_k \left(\frac{-1}{a-1} \left(\frac{\Gamma(a) y}{b} \right)^{\frac{1}{a-1}} \right)$$

where $k = 0$ for lower bounds and $k = -1$ for upper bounds. To demonstrate, consider solving for $y = 0.5$ of the gamma density with parameters $a = 2$, $b = 2$, displayed below:



Hazard Mixture Simulation Details

Distribution	Notation/Parameters	T	n	R
Gamma	$Ga(0.5, 1)$	250	523	4.36
Half-Normal	$\lvert N(0, 1)\rvert$	350	421	2.80
Gamma Mixture	$\frac{1}{2}Ga(2, 4) + \frac{1}{2}Ga(8, 4)$	1000	781	5.20
Log-Logistic	$LL(4, 1)$	300	266	4.04

Table 1: Simulation distributions used to generate inter-arrival times.

Markov Renewal Model Applied to Uniform Mixture (1/2)

Markov renewal processes are (the only?) method of incorporating marks, such as magnitude, into a renewal process. Supposed marks are categorical $\mathcal{M} = (1, \ldots, M)$ and inter-arrival times and marks are observed in pairs (x_i, m_i). The joint distribution is:

$$P\left(m_{i+1} = j, x_{i+1} < x | m_i = k, x_i\right) = p_{jk} F_{jk}(x)$$

The sequence $(m_i)_{i>0}$ forms a Markov chain with transition probabilities $P(m_{i+1} = j | m_i = k) = p_{jk}$ for $j, k \in \mathcal{M}$. Conditioned on $(m_i)_{i>0}$, the inter-arrival times $(x_i)_{i>0}$ form a sequence of independent variables with CDF $F_{m_{i-1}, m_i}(x)$.

Note that the censored observation x_{n+1} has unobserved mark, thus $P(x_{n+1} < x) = \sum_{m=1}^{M} p_{m_n, m} F_{m_n, m}(x)$. Thus the likelihood is:

$$L = \left(\sum_{m=1}^{M} p_{m_n, m} f_{m_n, m}(x) \right) \times \prod_{i=1}^{n} p_{m_{i-1}, m_i} f_{m_{i-1}, m_i}(x_i)$$

Markov Renewal Model Applied to Uniform Mixture (2/2)

Consider the state pair $(m_{i-1}, m_i) = (j, k)$. and let N_{jk} denote the number of times this pair is visited in the Markov chain. Also let $\boldsymbol{x}_{jk}$ denote the corresponding set of inter-arrival times. Finally, introduce a latent m_{n+1} with probabilities $p_{m_n,k}$ for all k and update N_{jk} accordingly. The augmented likelihood can be written as:

$$L = S_{m_n,m_{n+1}}(T - t_n) \times \prod_{j,k} p_{jk}^{N_{jk}} \times \prod_{j,k} \prod_{x_i \in \boldsymbol{x}_{jk}} f_{jk}(x_i)$$

We can incorporate the uniform mixture into this likelihood. Ultimately, we will be independently sampling weight sets $\boldsymbol{\omega}^{jk}$ and $\boldsymbol{\theta}^{jk}$.

- The sampling scheme of $\boldsymbol{\omega}^{jk}$ and $\boldsymbol{\theta}^{jk}$ for (j, k) subset will be standard blocked Gibbs for DPMM, except for the set containing m_{n+1} which will use presented scheme with added slice step.
- The value of m_{n+1} will be drawn discretely, affecting $S_{m_n,m_{n+1}}(T - t_n)$ as well as each N_{jk}.
- The transition matrix rows $p_{ij}, j \in \mathcal{M}$ will be drawn from a Dirichlet distribution.