Bayesian nonparametric modeling and inference for mean residual life functions

Athanasios Kottas

Department of Applied Mathematics and Statistics, University of California, Santa Cruz

Joint work with Valerie Poynor, California State University, Fullerton

2016 Joint Statistical Meetings
Chicago, Illinois
August 1, 2016
Let $T$ be an $\mathbb{R}^+$-valued random variable representing survival time.

- The **survival function** defines the probability of survival beyond time $t$, $S(t) = \Pr(T > t) = 1 - F(t)$, where $F(t)$ is the distribution function.

- The **hazard rate function** computes the probability of a failure in the next instant given survival up to time $t$,

$$h(t) = \lim_{\Delta t \to 0} \Pr[t < T \leq t + \Delta t \mid T > t]/(\Delta t)$$

For continuous $T$, $h(t) = f(t)/S(t)$, where $f(t)$ is the density function.
The mean residual life (MRL) function computes the expected remaining survival time of a subject given survival up to time $t$.

- Suppose $F(0) = 0$ and $\mu \equiv E(T) = \int_0^\infty S(t) \, dt < \infty$
- Then, the MRL function for continuous $T$ is defined as:

$$m(t) = E(T - t \mid T > t) = \frac{\int_t^\infty (u - t)f(u) \, du}{S(t)} = \frac{\int_t^\infty S(u) \, du}{S(t)}$$

and $m(t) \equiv 0$ whenever $S(t) = 0$.
The **mean residual life (MRL) function** computes the expected remaining survival time of a subject given survival up to time $t$.

- Suppose $F(0) = 0$ and $\mu \equiv E(T) = \int_0^\infty S(t) \, dt < \infty$.
- Then, the **MRL function** for continuous $T$ is defined as:

\[
m(t) = E(T - t \mid T > t) = \frac{\int_t^\infty (u - t)f(u) \, du}{S(t)} = \frac{\int_t^\infty S(u) \, du}{S(t)}
\]

and $m(t) \equiv 0$ whenever $S(t) = 0$.

\[
\text{Pr}(T \leq t)
\]

- expectation of the residual lifetime distribution at time $t$ ("residual survival function" at $t$, $S_t(u) = S(u)/S(t)$, for $u \geq t$)
Properties of MRL functions

→ The MRL function is of particular interest in survival and reliability analysis.

→ It characterizes the survival distribution through the Inversion Formula:

\[ S(t) = \frac{m(0)}{m(t)} \exp \left[ - \int_0^t \frac{1}{m(u)} du \right] \]

(key advantage compared to percentile residual life functions).
Properties of MRL functions

The MRL function is of particular interest in survival and reliability analysis.

It characterizes the survival distribution through the Inversion Formula:

\[
S(t) = \frac{m(0)}{m(t)} \exp \left[ - \int_0^t \frac{1}{m(u)} \, du \right]
\]

(key advantage compared to percentile residual life functions).

Characterization theorem for MRL functions (Hall & Wellner, 1981)

key properties: right-continuity for \( m(t) \), non-decreasing trend for function \( m(t) + t \)
Properties of MRL functions

→ The MRL function is of particular interest in survival and reliability analysis.

→ It characterizes the survival distribution through the Inversion Formula:

\[ S(t) = \frac{m(0)}{m(t)} \exp \left[-\int_0^t \frac{1}{m(u)} \, du \right] \]

(key advantage compared to percentile residual life functions).

→ Characterization theorem for MRL functions (Hall & Wellner, 1981)
key properties: right-continuity for \( m(t) \), non-decreasing trend for function \( m(t) + t \)

→ MRL function shape typically limited to be monotonic for standard parametric distributions; more flexible parametric distributions (mainly extensions of the Weibull) that achieve UBT and BT shapes have been developed.
Inference for MRL functions: literature review

- Classical (semiparametric) estimation:
  - Nonparametric estimators (Yang, 1978; Kochar et al., 2000).
  - A class of distributions with linear MRL functions (Hall & Wellner, 1984), extended in a semiparametric fashion to a family having proportional MRL functions (Oakes & Dasu, 1990).
  - Regression setting, $m(t; z) = \exp(\psi z)m_0(t)$ (Maguluri & Zhang, 1994; Chen & Cheng, 2005).

Very little attention in the Bayesian (nonparametric) literature:
- Empirical Bayes estimators based on a Dirichlet process (DP) prior for the distribution function (Lahiri & Park, 1991).
- Bayesian estimation under a specific form of censored/grouped data, using a DP prior for the corresponding survival distribution (Johnson, 1999).
Introduction

Inference for MRL functions: literature review

- Classical (semiparametric) estimation:
  - Nonparametric estimators (Yang, 1978; Kochar et al., 2000).
  - A class of distributions with linear MRL functions (Hall & Wellner, 1984), extended in a semiparametric fashion to a family having proportional MRL functions (Oakes & Dasu, 1990).
  - Regression setting, \( m(t; z) = \exp(\psi z) m_0(t) \) (Maguluri & Zhang, 1994; Chen & Cheng, 2005).

- Very little attention in the Bayesian (nonparametrics) literature:
  - Empirical Bayes estimators based on a Dirichlet process (DP) prior for the distribution function (Lahiri & Park, 1991).
  - Bayesian estimation under a specific form of censored/grouped data, using a DP prior for the corresponding survival distribution (Johnson, 1999).
Objectives

- Modeling the **MRL function** directly?
  - possible to define nonparametric priors on the space of mrl functions
  - however, obtaining the likelihood from the *inversion formula* is difficult due to the integration over the reciprocal of the MRL function.
Objectives

- Modeling the **MRL function** directly?
  - possible to define nonparametric priors on the space of mrl functions
  - however, obtaining the likelihood from the **inversion formula** is difficult due to the integration over the reciprocal of the MRL function.

- Instead, explore MRL function inference under flexible mixture modeling for the survival distribution:
  - interpretable mixture form for the implied MRL function
  - incorporation of covariates
  - more structured modeling for ordered MRL functions.
Objectives

- Modeling the **MRL function** directly?
  - possible to define nonparametric priors on the space of mrl functions
  - however, obtaining the likelihood from the inversion formula is difficult due to the integration over the reciprocal of the MRL function.

- Instead, explore MRL function inference under flexible mixture modeling for the survival distribution:
  - interpretable mixture form for the implied MRL function
  - incorporation of covariates
  - more structured modeling for ordered MRL functions.

- **Research objective:** develop a set of flexible inferential tools for MRL functions, using methods from the world of Bayesian nonparametrics.
Model formulation

We use a nonparametric mixture model for the density of the survival distribution:

$$f(t \mid G) = \int_{\Theta} k(t \mid \theta) \, dG(\theta); \quad G \sim \text{DP}(\alpha, G_0)$$

→ A Dirichlet process (DP) prior (Ferguson, 1973) is placed on the mixing distribution, $G$, mixing over the parameters of the kernel density $k(t \mid \theta)$.
Model formulation

We use a nonparametric mixture model for the density of the survival distribution:

\[
f(t \mid G) = \int_\Theta k(t \mid \theta) \, dG(\theta); \quad G \sim DP(\alpha, G_0)
\]

→ A Dirichlet process (DP) prior (Ferguson, 1973) is placed on the mixing distribution, \( G \), mixing over the parameters of the kernel density \( k(t \mid \theta) \)

→ Choice of the NPB prior for the mixing distribution not that critical (and the DP is still “the most beautiful of all”!)
Model formulation

We use a nonparametric mixture model for the density of the survival distribution:

\[ f(t \mid G) = \int_{\Theta} k(t \mid \theta) \, dG(\theta); \quad G \sim \text{DP}(\alpha, G_0) \]

→ A Dirichlet process (DP) prior (Ferguson, 1973) is placed on the mixing distribution, \( G \), mixing over the parameters of the kernel density \( k(t \mid \theta) \).

→ Choice of the NPB prior for the mixing distribution not that critical (and the DP is still “the most beautiful of all”!)

→ Choice of kernel distribution more important to ensure desirable properties for the implied MRL function of the mixture.
Recall the stick-breaking constructive definition of the DP (Sethuraman, 1994).

→ Let \( \{v_r : r = 1, 2, \ldots\} \) and \( \{\theta_l : l = 1, 2, \ldots\} \) be independent sequences of random variables:

- \( v_r \overset{iid}{\sim} \text{Beta}(1, \alpha) \), for \( r = 1, 2, \ldots \) (where \( \alpha \) is the precision parameter)
- \( \theta_l \overset{iid}{\sim} G_0 \), for \( l = 1, 2, \ldots \) (where \( G_0 \) is the baseline distribution)

→ Define \( \omega_1 = v_1 \) and \( \omega_l = v_l \prod_{r=1}^{l-1}(1 - v_r) \), for \( l = 2, 3, \ldots \)

→ Then, a realization, \( G \), from the DP(\( \alpha, G_0 \)) is almost surely of the form

\[
G = \sum_{l=1}^{\infty} \omega_l \delta_{\theta_l}
\]
We use a truncated version of the DP stick-breaking construction:

* \( G_N = \sum_{\ell=1}^{N} p_\ell \delta_{\theta_\ell} \), where \( \theta_\ell \sim G_0 \), for \( \ell = 1, \ldots, N \)

* \( p_1 = v_1; \ p_\ell = v_\ell \prod_{r=1}^{\ell-1} (1 - v_r) \), for \( \ell = 2, 3, \ldots N - 1 \), with \( p_N = 1 - \sum_{\ell=1}^{N-1} p_\ell \), where \( v_r \sim \text{Beta}(1, \alpha) \), for \( r = 1, \ldots, N - 1 \)

* truncation level \( N \) can be specified using standard DP properties.

The mixture model for the survival density becomes:

\[
 f(t \mid G_N) = \int_{\Theta} k(t \mid \theta) \, dG_N(\theta) = \sum_{\ell=1}^{N} p_\ell k(t \mid \theta_\ell)
\]
Mixture modeling for mean residual life function inference

We use a truncated version of the DP stick-breaking construction:

\[ G_N = \sum_{\ell=1}^{N} p_\ell \delta_{\theta_\ell}, \text{ where } \theta_\ell \overset{iid}{\sim} G_0, \text{ for } \ell = 1, \ldots, N \]

\[ p_1 = v_1; \ p_\ell = v_\ell \prod_{r=1}^{\ell-1} (1 - v_r), \text{ for } \ell = 2, 3, \ldots N - 1, \text{ with } p_N = 1 - \sum_{\ell=1}^{N-1} p_\ell, \]

where \( v_r \overset{iid}{\sim} \text{Beta}(1, \alpha) \), for \( r = 1, \ldots, N - 1 \)

* truncation level \( N \) can be specified using standard DP properties.

The mixture model for the survival density becomes:

\[
 f(t \mid G_N) = \int_{\Theta} k(t \mid \theta) \, dG_N(\theta) = \sum_{\ell=1}^{N} p_\ell k(t \mid \theta_\ell)
\]

Alternatively, truncation can be applied only when inference for \( G \) is needed (using marginal or slice samplers for MCMC posterior simulation).
Implied form for the MRL function

$S(t \mid \theta)$ and $m(t \mid \theta)$: survival and MRL function of the kernel distribution.

The MRL function of the mixture can be expressed as

$$m(t \mid G_N) = \frac{\int_{t}^{\infty} \int_{\Theta} S(u \mid \theta) dG_N(\theta) \, du}{S(t \mid G_N)}$$

$$= \frac{\sum_{\ell=1}^{N} p_{\ell} (\int_{t}^{\infty} S(u \mid \theta_{\ell}) \, du)}{\sum_{\ell=1}^{N} p_{\ell} S(t \mid \theta_{\ell})}$$

$$= \sum_{\ell=1}^{N} q_{\ell}(t) m(t \mid \theta_{\ell})$$

where $q_{\ell}(t) = p_{\ell} S(t \mid \theta_{\ell}) / \{\sum_{r=1}^{N} p_{r} S(t \mid \theta_{r})\}$ are normalized weights.
Implied form for the MRL function

- $S(t \mid \theta)$ and $m(t \mid \theta)$: survival and MRL function of the kernel distribution.

- The MRL function of the mixture can be expressed as

$$m(t \mid G_N) = \frac{\int_t^\infty \int_{\Theta} S(u \mid \theta) dG_N(\theta) du}{S(t \mid G_N)}$$

$$= \frac{\sum_{\ell=1}^N p_\ell \left( \int_t^\infty S(u \mid \theta_\ell) du \right)}{\sum_{\ell=1}^N p_\ell S(t \mid \theta_\ell)}$$

$$= \sum_{\ell=1}^N q_\ell(t) m(t \mid \theta_\ell)$$

where $q_\ell(t) = p_\ell S(t \mid \theta_\ell)/\{\sum_{r=1}^N p_r S(t \mid \theta_r)\}$ are normalized weights.

- Implied prior structure for the MRL function: mixture of parametric kernel MRL functions with time-dependent weights defined through the DP stick-breaking probabilities and the kernel survival function.
Choice of the mixture kernel

→ For the MRL function of the mixture to be well-defined, finite mean for the mixture survival distribution is required

* if the kernel mean, \( E(T \mid \theta) \), is finite, and \( \int_{\Theta} E(T \mid \theta) \, dG_0(\theta) < \infty \), then
\[
E(T \mid G) = \int_0^\infty S(t \mid G) \, dt < \infty \text{ (a.s.)}
\]
Choice of the mixture kernel

→ For the MRL function of the mixture to be well-defined, finite mean for the mixture survival distribution is required

* if the kernel mean, $E(T \mid \theta)$, is finite, and $\int_{\Theta} E(T \mid \theta) \, dG_0(\theta) < \infty$, then

$$E(T \mid G) = \int_0^\infty S(t \mid G) \, dt < \infty \, (a.s.)$$

→ Tail behavior of the MRL function for the mixture distribution:

* $\lim_{t \to \infty} m(t \mid \theta) = 0(\infty)$, $\forall \theta \in \Theta \Rightarrow \lim_{t \to \infty} m(t \mid G_N) = 0(\infty)$
Choice of the mixture kernel

→ For the MRL function of the mixture to be well-defined, finite mean for the mixture survival distribution is required

* if the kernel mean, $E(T \mid \theta)$, is finite, and $\int_{\Theta} E(T \mid \theta) \, dG_0(\theta) < \infty$, then $E(T \mid G) = \int_0^\infty S(t \mid G) \, dt < \infty$ (a.s.)

→ Tail behavior of the MRL function for the mixture distribution:

* $\lim_{t \to \infty} m(t \mid \theta) = 0 (\infty), \forall \theta \in \Theta \Rightarrow \lim_{t \to \infty} m(t \mid G_N) = 0 (\infty)$

→ Kernel distributions that allow both increasing and decreasing MRL function shapes: gamma and Weibull distributions.

→ Much easier to achieve the finite mean restriction under the gamma distribution,

$$k(t \mid \theta) = \Gamma(t \mid \exp(\theta), \exp(\phi)), \quad \theta = (\theta, \phi) \in \mathbb{R}^2$$
Prior specification and posterior inference

→ Dependent baseline distribution, $G_0(\theta, \phi) = N_2((\theta, \phi)' | \mu, \Sigma)$ to facilitate learning for the location and dispersion of the mixture components.

→ Hyperpriors: normal for $\mu$; inverse-Wishart for $\Sigma$; gamma for $\alpha$
Prior specification and posterior inference

→ Dependent baseline distribution, $G_0(\theta, \phi) = N_2((\theta, \phi)' | \mu, \Sigma)$ to facilitate learning for the location and dispersion of the mixture components.

→ Hyperpriors: normal for $\mu$; inverse-Wishart for $\Sigma$; gamma for $\alpha$

→ MCMC posterior simulation using blocked Gibbs sampling (for data sets that may include different types of censoring).

→ The posterior samples for $G_N \equiv \{(p_\ell, \theta_\ell, \phi_\ell) : \ell = 1, ..., N\}$ can be used to obtain inference for the density, survival, and hazard functions at any time point $t$, by directly evaluating the expressions for these functions under the gamma DP mixture model.
For the MRL function:

\[
m(t) = \frac{\int_t^\infty S(u) \, du}{S(t)} = \frac{\int_0^\infty S(u) \, du - \int_0^t S(u) \, du}{S(t)} = \mu - \int_0^t S(u) \, du \]

where \( \mu = E(T \mid G_N) = \sum_{\ell=1}^N p_\ell \exp(\theta_\ell - \phi_\ell) \)

Computing over a grid of survival times, \( t_{0,j} \) for \( j = 1, \ldots, J \)

We evaluate the MRL function at the first grid point by

\[
m(t_{0,1} \mid G_N) = \left[ E(T \mid G_N) - 0.5(t_{0,1}(1 + S(t_{0,1} \mid G_N))) \right] / S(t_{0,1} \mid G_N)
\]

and use the following expression for \( j = 2, \ldots, J \):

\[
m(t_{0,j} \mid G_N) = \frac{E(T \mid G_N) - \frac{1}{2} \left( t_{0,1}(1 + S(t_{0,1} \mid G_N)) + \sum_{i=2}^j (t_{0,j} - t_{0,j-1})(S(t_{0,j} \mid G_N) + S(t_{0,j-1} \mid G_N)) \right)}{S(t_{0,j} \mid G_N)}
\]
Simulation example 1

Data set of 200 realizations from $0.35\Gamma(10, 0.5) + 0.4\Gamma(20, 1) + 0.15\Gamma(30, 5) + 0.1\Gamma(40, 8)$
Simulation example 2

Data set of 100 realizations from $0.3 \Gamma(15, 0.2) + 0.25 \Gamma(12, 0.5) + 0.35 \Gamma(8, 2) + 0.1 \Gamma(3, 6)$
Survival times of patients with small cell lung cancer

Study involving two treatments for small cell lung cancer (Ying et al., 1988): survival times (in days) for 121 patients (23 right censored) randomly assigned to one of two treatments – Arm A, under which 62 patients received cisplatin (P) followed by etoposide (E), and Arm B, where 59 patients received (E) followed by (P).
Survival times of patients with small cell lung cancer

Study involving two treatments for small cell lung cancer (Ying et al., 1988): survival times (in days) for 121 patients (23 right censored) randomly assigned to one of two treatments – Arm A, under which 62 patients received cisplatin (P) followed by etoposide (E), and Arm B, where 59 patients received (E) followed by (P).
Figure: Left: posterior mean estimates for the MRL function of Arm A (blue) and Arm B (green). Right: Pr($m_A(t) > m_B(t)$) (red dashed) and Pr($m_A(t) > m_B(t) \mid \text{data}$) (black solid) as a function of time.
Density regression for survival responses

→ The density regression approach has been explored in the NPB literature for real-valued and ordinal responses, using primarily multivariate normal kernels (Müller et. al., 1996; Taddy & Kottas, 2010; DeYoreo & Kottas, 2015).

→ Benefits for survival regression:
  * non-standard response distributions and non-linear regression relationships
  * survival analysis applications typically involve a small to moderate number of (random) covariates.
Density regression for survival responses

The density regression approach has been explored in the NPB literature for real-valued and ordinal responses, using primarily multivariate normal kernels (Müller et. al., 1996; Taddy & Kottas, 2010; DeYoreo & Kottas, 2015).

Benefits for survival regression:
- non-standard response distributions and non-linear regression relationships
- survival analysis applications typically involve a small to moderate number of (random) covariates.

DP mixture model for the joint response-covariate density:

$$f(t, x \mid G) = \int_{\Theta} k(t, x \mid \theta) \, dG(\theta) \approx \sum_{\ell=1}^{N} p_{\ell} k(t, x \mid \theta_{\ell})$$

where $x$ is a vector of random covariates.
Useful interpretation for regression functionals.

Mean regression:

\[ E(T \mid x_0, G_N) = \sum_{\ell=1}^{N} q_\ell(x_0) E(T \mid x_0, \theta_\ell) \]

where \( q_\ell(x_0) = \frac{p_\ell k(x_0 \mid \theta_\ell)}{\sum_{r=1}^{N} p_r k(x_0 \mid \theta_r)} \) are covariate-dependent weights.
Useful interpretation for regression functionals.

Mean regression:

\[ E(T \mid x_0, G_N) = \sum_{\ell=1}^{N} q_\ell(x_0) E(T \mid x_0, \theta_\ell) \]

where \( q_\ell(x_0) = \frac{p_\ell k(x_0 \mid \theta_\ell)}{\left\{ \sum_{r=1}^{N} p_r k(x_0 \mid \theta_r) \right\}} \) are covariate-dependent weights.

Mean residual life regression:

\[ m(t \mid x_0, G_N) = \sum_{\ell=1}^{N} q_\ell(t, x_0) m(t \mid x_0, \theta_\ell) \]

where \( q_\ell(t, x_0) = \frac{p_\ell k(x_0 \mid \theta_\ell) S(t \mid x_0, \theta_\ell)}{\left\{ \sum_{r=1}^{N} p_r k(x_0 \mid \theta_r) S(t \mid x_0, \theta_r) \right\}} \) are covariate-dependent and time-dependent weights.
Specification for the joint kernel?
Specification for the joint kernel?

The condition that ensures the finiteness for the mean can be extended: if \( E_{G_0}[E(T \mid x_0, \theta)] < \infty \), then \( E(T \mid x_0, G) < \infty \)
Specification for the joint kernel?

The condition that ensures the finiteness for the mean can be extended:
if $E_{G_0}[E(T \mid x_0, \theta)] < \infty$, then $E(T \mid x_0, G) < \infty$

Earlier considerations again favor a gamma kernel component for the survival response variable

* product kernel, $k(t, x) = k(t)k(x)$ (with a gamma distribution taken for $k(t)$)
* incorporate dependency between the covariates and the survival responses within the kernel, e.g., consider an appropriate marginal $k(x)$ and

$$k(t \mid x) = \Gamma(t \mid \exp(\theta), \exp(x^T\beta)),$$

such that $E(T \mid x, \theta, \beta) = \exp(\theta - x^T\beta)$
Interest often lies in modeling survival times for treatment and control groups. 

→ Benefits in modeling dependency across groups.

→ Let $s \in S$ represent in general the index of dependence – we consider $S = \{ T, C \}$ where $T$ and $C$ are the treatment and control groups, respectively.
Dependent mixture model for treatment/control settings

Interest often lies in modeling survival times for treatment and control groups.

→ Benefits in modeling dependency across groups.

→ Let $s \in S$ represent in general the index of dependence – we consider $S = \{T, C\}$ where $T$ and $C$ are the treatment and control groups, respectively.

→ DP mixture regression model:

$$f(t, x \mid G_s) = \int_{\Theta} k(t, x \mid \theta) dG_s(\theta), \text{ for } s \in S$$

where we seek to model the pair of dependent random mixing distributions $(G_C, G_T)$. 
Interest often lies in modeling survival times for treatment and control groups.

→ Benefits in modeling dependency across groups.

→ Let $s \in S$ represent in general the index of dependence – we consider $S = \{T, C\}$ where $T$ and $C$ are the treatment and control groups, respectively.

→ DP mixture regression model:

$$f(t, x \mid G_s) = \int_{\Theta} k(t, x \mid \theta) dG_s(\theta), \text{ for } s \in S$$

where we seek to model the pair of dependent random mixing distributions $(G_C, G_T)$.

→ General DDP prior structure (MacEachern, 2000), $G_s = \sum_{l=1}^{\infty} \omega_{ls} \delta_{\theta_{ls}}$, where marginally, $G_s \sim DP(\alpha_s, G_{0s})$, for each $s \in S$. 

Athanasios Kottas (UCSC)

NPB inference for MRL functions

August 1, 2016 21 / 27
We might expect the two groups to be comprised of similar components, but possibly having varying prevalence, motivating modeling dependency only through the weights.
We might expect the two groups to be comprised of similar components, but possibly having varying prevalence, motivating modeling dependency only through the weights.

We use mixing distribution, \( G_s = \sum_{l=1}^{\infty} \omega_l \delta_{\theta_l} \), with a bivariate beta distribution defining the dependent stick-breaking weights (thus retaining the DP marginally).

With the truncated version of \( G_s \approx \sum_{\ell=1}^{N} p_{\ell s} \delta_{\theta_{\ell}} \), the model

\[
f(t, x \mid G_s) = \int_{\Theta} k(t, x \mid \theta) \, dG_s(\theta) \approx \sum_{\ell=1}^{N} p_{\ell s} k(t, x \mid \theta_{\ell}), \text{ for } s \in \{T, C\}
\]
We might expect the two groups to be comprised of similar components, but possibly having varying prevalence, motivating modeling dependency only through the weights.

We use mixing distribution, $G_s = \sum_{l=1}^{\infty} \omega_l \delta_{\theta_l}$, with a bivariate beta distribution defining the dependent stick-breaking weights (thus retaining the DP marginally).

With the truncated version of $G_s \approx \sum_{\ell=1}^{N} p_\ell s \delta_{\theta_\ell}$, the model

$$f(t, x \mid G_s) = \int_\Theta k(t, x \mid \theta) \, dG_s(\theta) \approx \sum_{\ell=1}^{N} p_\ell s k(t, x \mid \theta_\ell), \text{ for } s \in \{T, C\}$$

Practical benefit: modeling dependency only through the weights is not affected by the dimensionality of the mixture kernel.
Small cell lung cancer example (revisited)

Results under the DDP mixture model with a gamma kernel, applied to the data set comprising responses from both treatments (Arm A and B).
**Small cell lung cancer example (revisited)**

Results under the DDP mixture model with a gamma kernel, applied to the data set comprising responses from both treatments (Arm A and B).
Results under the DDP mixture model with a product gamma/lognormal kernel, applied to the full data set which includes also the patient's age (in years) at entry in the study.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Expected survival time (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>400</td>
</tr>
<tr>
<td>50</td>
<td>600</td>
</tr>
<tr>
<td>60</td>
<td>800</td>
</tr>
<tr>
<td>70</td>
<td>1000</td>
</tr>
<tr>
<td>80</td>
<td>1200</td>
</tr>
</tbody>
</table>

Figure: Posterior mean and 80% interval estimates for the mean regression, $E(T | \text{age})$, for Arm A (left) and Arm B (right).
Results under the DDP mixture model with a product gamma/lognormal kernel, applied to the full data set which includes also the patient’s age (in years) at entry in the study.

Figure: Posterior mean and 80% interval estimates for the mean regression, $E(T \mid \text{age})$, for Arm A (left) and Arm B (right).
Figure: Estimates of the MRL function of Arm A (blue) and Arm B (green) for ages 50 (left), 60 (middle), and 78 (right).
Modeling and inference methods for mean residual life (MRL) functions:

Papers and future work

→ Modeling and inference methods for mean residual life (MRL) functions:

→ In some applications, we may wish to impose the restriction that the average remaining lifetime for one population is higher than that of the other population.

→ Nonparametric modeling for two MRL ordered distributions ($T_1 \leq_{mrl} T_2$ if $m_1(t) \leq m_2(t)$, for all $t \in \mathbb{R}^+$), using structured Bernstein polynomial or DP mixture priors.
Contact info:  
    e-mail: thanos@soe.ucsc.edu, web: http://www.ams.ucsc.edu/~thanos

MANY THANKS!