Inference for Mean Residual Life Function Under Nonparametric Mixture Modeling of Survival Distribution

Valerie Poynor and Athanasios Kottas vpoynor@soe.ucsc.edu University of California, Santa Cruz

Baskin Engineering UC SANTA CRUZ

Introduction

The mean residual life function provides the expected remaining life given that the subject has survived (i.e., is event-free) up to a particular time. The mean residual life function characterizes the survival distribution, and thus it can be used in fitting a model to the data.

We review the key properties of the mean residual life function and investigate its form for some common distributions. We next develop Bayesian nonparametric inference for mean residual life functions obtained from a flexible model for the corresponding survival distribution, using Dirichlet process mixtures of lognormal or Weibull distributions.

We compare with an exponentiated Weibull model, a parametric survival distribution that allows various shapes for the mean residual life function. The approach is illustrated with two data examples, one involving comparison of lifetimes of experimental subjects under different diets, and one on right censored survival times of liver metastases patients.

Mean Residual Life Function

The Survival function S(x) defines the probability of survival beyond time x:

$$S(x) \quad = \quad Pr(X > x) = 1 - F(x)$$

with F(x) the distribution function. The Hazard function gives the probability of failure in the next instant given survival up to time x:

$$h(x) \quad = \quad \lim_{\Delta x \to 0} \frac{Pr[x < X \le x + \Delta x | X > x]}{\Delta x} \overset{\text{(for continuous X)}}{=} \frac{f(x)}{S(x)}$$

where f(x) is the probability density function. The Mean Residual Life (MRL) function computes the expected remaining survival time of a subject given survival up to time x. The MRL function with finite mean (μ) is defined as:

$$m(x) \quad = \quad E(X - x | X > x) \overset{\text{(for continuous X)}}{=} \frac{\int_x^\infty S(t)dt}{S(x)}$$

For a continuous random variable X with finite mean (μ) the Survival function is defined through the MRL (Inversion Formula):

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

When the Hazard function isdecreasing (DCR), the MRL function is increasing (INC) and vice versa. For an Bathtub (BT) shape Hazard function, if $h(0) > 1/\mu$ then the MRL function has a Upside Down Bathtub (UBT) shape, and if $h(0) \le 1/\mu$ then the MRL function is decreasing (DCR). An UBT MRL function is commonly seen in biological situations where the subject has a lower MRL during infancy and elderly age and a higher MRL during the middle ages. Generally, the MRL function does not have a closed form expression and must be obtained numerically.

Common Survival Distributions

The distributions in the table below may be restrictive in the form of the MRL function. These distributions would require knowledge of the shape the MRL function for the data to ensure the proper model is fit. Moreover, many of these distributions have monotonic MRL functions which is not characteristic for most practical situations.

Distribution	Hazard	Mean Residual Life
Gamma(α, λ) shape parameter $\alpha > 0$ scale parameter $\lambda > 0$	$\alpha < 1$ DCR $\alpha = 1$ constant $(1/\lambda)$ $\alpha > 1$ INC	$\alpha < 1$ INC $\alpha = 1$ constant(λ) $\alpha > 1$ DCR
Gompertz(α, λ) shape parameter $\alpha > 0$ scale parameter $\lambda > 0$	$\forall \alpha$ INC	$\forall \alpha$ DCR
Loglogistic(α, λ) shape parameter $\alpha > 0$ scale parameter $\lambda > 0$	$\alpha \le 1$ DCR $\alpha > 1$ UBT	$\alpha \le 1$ undefined $\alpha > 1$ BT
Weibull(α, λ) shape parameter $\alpha > 0$ scale parameter $\lambda > 0$	$\alpha < 1$ DCR $\alpha = 1$ constant $(1/\lambda)$ $\alpha > 1$ INC	$\alpha < 1$ INC $\alpha = 1$ constant(λ) $\alpha > 1$ DCR

Model

The Exponentiated Weibull Distribution has an MRL function that can be INC, DCR, UBT, BT, or constant. For $t, \alpha, \theta, \sigma > 0$, the probability density function is defined as:

$$\frac{\alpha\theta}{\sigma} \left[1 - exp\left(- \left(\frac{t}{\sigma} \right)^\alpha \right) \right]^{\theta - 1} exp\left(- \left(\frac{t}{\sigma} \right)^\alpha \right) \left(\frac{t}{\sigma} \right)^{\alpha - 1}$$

α and θ are shape parameters while σ is a scale parameter The form of the MRL function depends only on the values of α and θ The Dirichlet Process (DP) Mixture Model:

$$x_i | G \overset{ind}{\sim} \int K(x_i; \boldsymbol{\theta}) dG(\boldsymbol{\theta}) = \sum_{l=1}^{L} p_l K(x_i; \boldsymbol{\theta}_l) \text{ for } i = 1:n$$

$$G \sim DP(\alpha, G_o)$$

such that $G = \sum_{l=1}^{\infty} w_l \delta_{\boldsymbol{\theta}_l}$ is truncated by $G_L = \sum_{l=1}^{L} p_l \delta_{\boldsymbol{\theta}_l}$, and where p_l are the weights, obtained, via DP Stick-Breaking (SB) construction, corresponding to the component θ_l, and L is the total number of components specified in the model.

→The Lognormal (LN) Dirichlet Process Mixture Model is fitted to the survival times of the rats under different diets. In this model, the kernel distribution is the Lognormal distribution and the mixing is performed on the location and scale parameters, $\boldsymbol{\theta} = (\mu, \sigma^2)$, respectively.

→The Weibull Dirichlet Process Mixture Model is fitted to the survival times of the liver metastasis patients. Here the kernel is the Weibull distribution and the mixing is performed on the shape and scale parameters, $\boldsymbol{\theta} = (\gamma, \sigma)$, respectively.

→ A mixture of Weibull distributions is preferred over a mixture of Lognormal distributions in the presence of censoring due to the closed form of the Weibull survival distribution.

Results

We fit an Exponentiated Weibull Model and LN DP Mixture Model to two data sets in an experiment that studied the lifetimes of rats under different diets. The Restricted diet group consisted of 106 rats while the Ad Libitum (free-eating) group consisted of 108 rats.

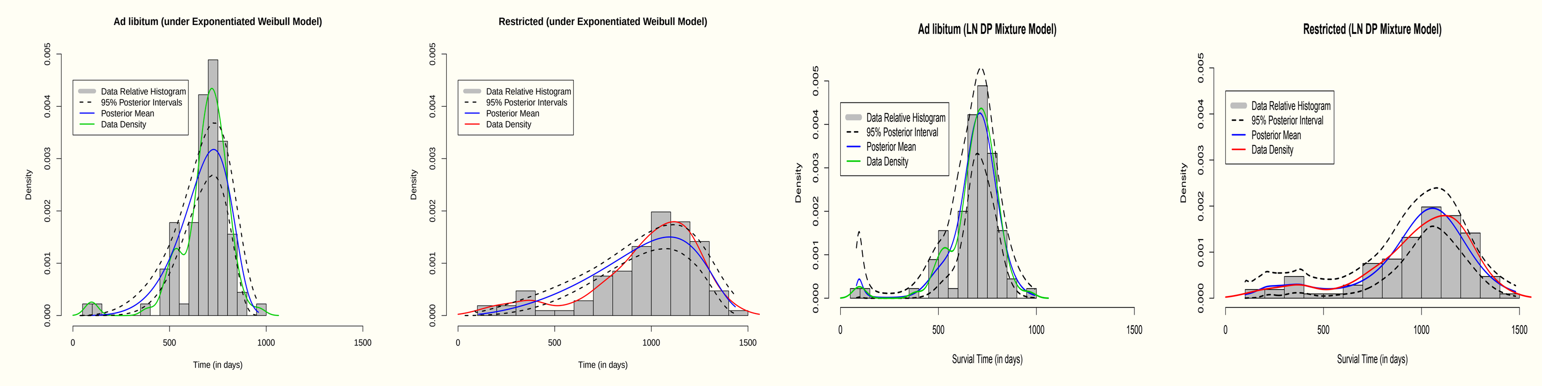

Figure 1

Figure 1 indicates that the nonparametric model does a much better job at capturing the unique characteristics of the data sets. Using our posterior samples from the LN DP Mixture Model we simulated point and interval estimates for the Densities, Survival, Hazard, and MRL functions for the two experimental groups (Figure 2). The MRL function of the Restricted diet is significantly higher than the Ad libitum diet.

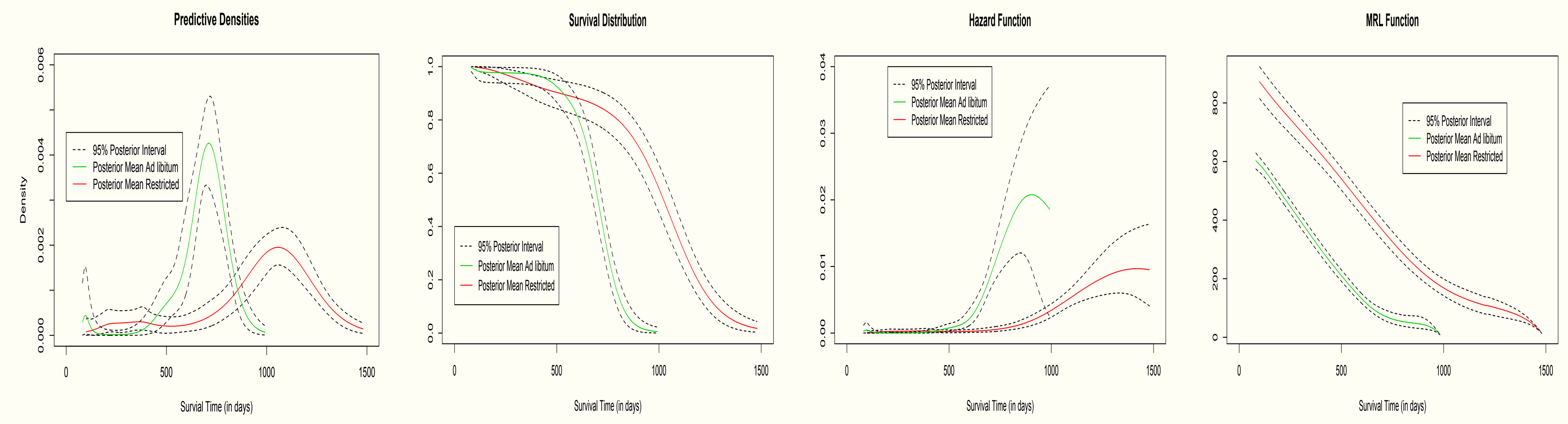

Figure 2

We fit a Weibull DP Mixture Model to the survival times of 360 patients with liver metastasis. There were 259 right censored times. Posterior point and interval estimates for the Density, Survival, Hazard, and MRL functions are shown in Figure 3.

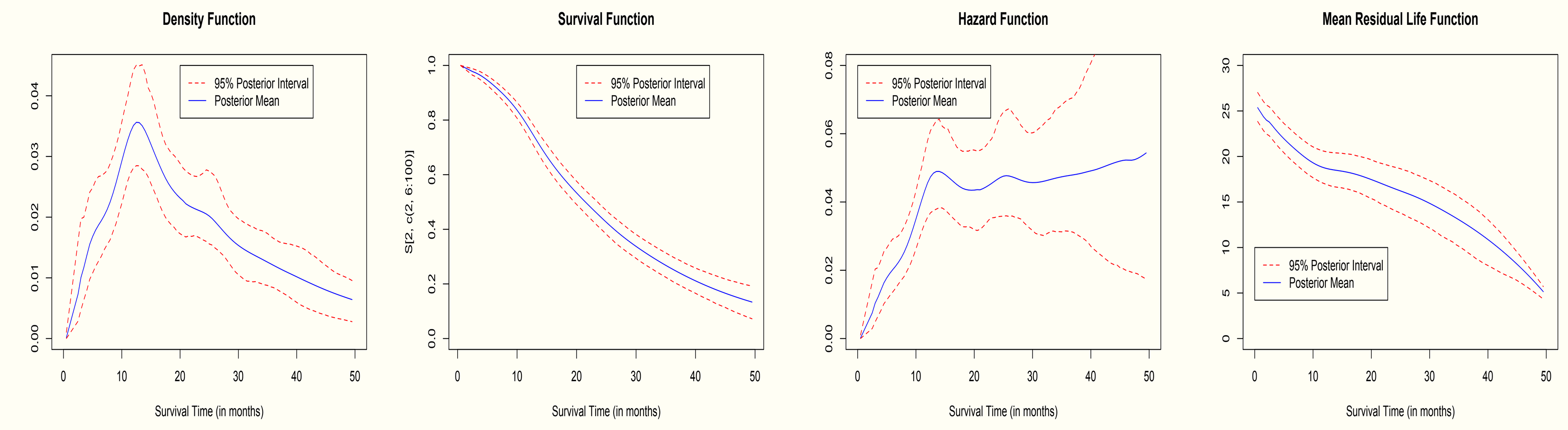

Figure 3

The Hazard function in Figure 3 exhibits multi-model behavior. The MRL function shows a sharp decline in the remaining life expectancy towards the beginning age, and continues to decrease with points of inflection around the corresponding ages of the modes in the Hazard function.

Future Work

→ Extensions to regression modeling with censored responses.
→ Models that develop priors directly for the MRL function.