Bayesian Inference on Parametric Zombie Survival Model

Bomin Kim

May 14, 2017

Abstract

In this paper, we implement Bayesian inference on the new split population survival model, that explicitly models the misclassification probability of failure (vs. right censored) events. This includes two parametric survival models (Exponential and Weibull) and (possibly) Cox proportional hazards regression model.

Review on Parametric Zombie Survival Model

Likelihood function

Recall from Ben's "Parametric Zombie Survival Model" that the probability of misclassification (that is, subset of non-censured failure outcomes that are being misclassified) is

$$\alpha = \Pr(C_i = 1|\widetilde{C}_i = 0). \tag{1}$$

The unconditional density is thus given by the combination of an observation's misclassification probability and its probability of experiencing an actual failure conditional on not being misclassified,

$$\alpha_i + (1 - \alpha_i) * f(t_i) \tag{2}$$

And the unconditional survival function is therefore

$$(1 - \alpha_i) * S(t_i), \tag{3}$$

where

$$\alpha_i = \frac{\exp(\mathbf{Z}\gamma)}{1 + \exp(\mathbf{Z}\gamma)}. \tag{4}$$

The likeihood function of the Parametric Zombie Survival Model is defined as

$$L = \prod_{i=1}^{N} [\alpha_i + (1 - \alpha_i) f(t_i|\mathbf{X}, \boldsymbol{\beta})]^{C_i} [(1 - \alpha_i) S(t_i|\mathbf{X},\boldsymbol{\beta})]^{1-C_i} \tag{5}$$

And the log likelihood is

$$lnL = \sum_{i=1}^{N} \{C_i \ln[\alpha_i + (1 - \alpha_i) f(t_i|\mathbf{X}, \boldsymbol{\beta})] + (1 - C_i) \ln[(1 - \alpha_i) S(t_i|\mathbf{X},\boldsymbol{\beta})]\}. \tag{6}$$

Posterior Distribution of Parametric Zombie Survival Models

Exponential

For exponential survival model, the density function and survival function are

$$\begin{aligned} f(t_i|X_i,\boldsymbol{\beta}) &= \exp(X_i\boldsymbol{\beta})\exp(-\exp(X_i\boldsymbol{\beta})t_i) \\ S(t_i|X_i,\boldsymbol{\beta}) &= \exp(-\exp(X_i\boldsymbol{\beta})t_i). \end{aligned} \tag{7}$$

Then, the likelihood function of Exponential Zombie survival model is

$$L(\boldsymbol{\beta},\boldsymbol{\gamma}) = \prod_{i=1}^{N}[\alpha_i + (1-\alpha_i)\exp(X_i\boldsymbol{\beta})\exp(-\exp(X_i\boldsymbol{\beta})t_i)]^{C_i}[(1-\alpha_i)\exp(-\exp(X_i\boldsymbol{\beta})t_i)]^{1-C_i} \tag{8}$$

where X_i is the i^{th} row of the covariate matrix $\mathbf{X}$.

In the exponential survival model, we assume the prior of $\boldsymbol{\beta} = \{\beta_1, ..., \beta_{p_1}\}$ as

$$\boldsymbol{\beta} \sim \text{Multivariate Normal}_{p_1}(\boldsymbol{\mu}_\beta, \Sigma_\beta), \tag{9}$$

thus the conditional posterior distribution for $\boldsymbol{\beta}$ parameters is given by

$$\pi(\boldsymbol{\beta}|\mathbf{C},\mathbf{X},\mathbf{Z},\mathbf{t},\boldsymbol{\gamma}) \propto L(\boldsymbol{\beta}|\mathbf{C},\mathbf{X},\mathbf{Z},\mathbf{t},\boldsymbol{\gamma}) \times \pi(\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \Sigma_\beta). \tag{10}$$

Moreover, we can also assign mutivariate Normal prior to $\boldsymbol{\gamma} = \{\gamma_1, ..., \gamma_{p_2}\}$,

$$\boldsymbol{\gamma} \sim \text{Multivariate Normal}_{p_2}(\boldsymbol{\mu}_\gamma, \Sigma_\gamma), \tag{11}$$

and the corresponding conditional posterior distribution of $\boldsymbol{\gamma}$ becomes

$$\pi(\boldsymbol{\gamma}|\mathbf{C},\mathbf{X},\mathbf{Z},\mathbf{t},\boldsymbol{\beta}) \propto L(\boldsymbol{\gamma}|\mathbf{C},\mathbf{X},\mathbf{Z},\mathbf{t},\boldsymbol{\beta}) \times \pi(\boldsymbol{\gamma}|\boldsymbol{\mu}_\gamma, \Sigma_\gamma). \tag{12}$$

Weibull

If the survival time t has a Weibull distribution of $W(t|\lambda, X_i\boldsymbol{\beta})$, the density function and survival function are

$$\begin{aligned} f(t_i|\lambda, X_i, \boldsymbol{\beta}) &= \lambda(\exp(X_i\boldsymbol{\beta}))(\exp(X_i\boldsymbol{\beta})t_i)^{\lambda-1}\exp(-(\exp(X_i\boldsymbol{\beta})t_i)^{\lambda}) \\ S(t_i|\lambda, X_i, \boldsymbol{\beta}) &= \exp(-(\exp(X_i\boldsymbol{\beta})t_i)^{\lambda}), \end{aligned} \tag{13}$$

which shows that $\lambda = 1$ reduces to Exponential survival model, which is a well-known property. The likelihood function of Weibull Zombie survival model is

$$\begin{aligned} L(\lambda, \boldsymbol{\beta}, \boldsymbol{\gamma}) = &\prod_{i=1}^{N}[\alpha_i + (1-\alpha_i)\lambda(\exp(X_i\boldsymbol{\beta}))(\exp(X_i\boldsymbol{\beta})t_i)^{\lambda-1}\exp(-(\exp(X_i\boldsymbol{\beta})t_i)^{\lambda})]^{C_i} \\ &\times [(1-\alpha_i)\exp(-(\exp(X_i\boldsymbol{\beta})t_i)^{\lambda})]^{1-C_i}. \end{aligned} \tag{14}$$

For the two parameters λ and $\boldsymbol{\beta} = \{\beta_1, ..., \beta_{p_1}\}$, we assign prior to each parameter as

$$\begin{aligned} &\lambda \sim \text{Gamma}(a_\lambda, b_\lambda) \\ &\boldsymbol{\beta} \sim \text{Multivariate Normal}_{p_1}(\boldsymbol{\mu}_\beta, \Sigma_\beta), \end{aligned} \tag{15}$$

where the conditional distribution for λ and $\boldsymbol{\beta}$ parameters are given by

$$\begin{aligned} &\pi(\lambda|\mathbf{C}, \boldsymbol{\alpha}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \propto L(\lambda|\mathbf{C}, \boldsymbol{\alpha}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \times \pi(\lambda|a_\lambda, b_\lambda) \\ &\pi(\boldsymbol{\beta}|\mathbf{C}, \boldsymbol{\alpha}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\gamma}, \lambda) \propto L(\boldsymbol{\beta}|\mathbf{C}, \boldsymbol{\alpha}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\gamma}, \lambda) \times \pi(\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \Sigma_\beta). \end{aligned} \tag{16}$$

Same as Exponential case, we can assign mutivariate Normal prior to $\boldsymbol{\gamma} = \{\gamma_1, ..., \gamma_{p_2}\}$,

$$\boldsymbol{\gamma} \sim \text{Multivariate Normal}_{p_2}(\boldsymbol{\mu}_\gamma, \Sigma_\gamma), \tag{17}$$

and the corresponding conditional posterior distribution becomes

$$\pi(\boldsymbol{\gamma}|\mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\beta}, \lambda) \propto L(\boldsymbol{\gamma}|\mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\beta}, \lambda) \times \pi(\boldsymbol{\gamma}|\boldsymbol{\mu}_\gamma, \Sigma_\gamma). \tag{18}$$

Markov chain Monte Carlo (MCMC) algorithm

Prior specifications

Since we do not have any informative prior for parametric Zombie survival model, we can follow the common approach and specify our hyperparameters as below:

$$\begin{aligned} &\boldsymbol{\mu}_\beta = \mathbf{0}, \Sigma_\beta \sim \text{Inverse-Wishart}(\mathbf{I}_{p_1}, \nu_1), \\ &\boldsymbol{\mu}_\gamma = \mathbf{0}, \Sigma_\gamma \sim \text{Inverse-Wishart}(\mathbf{I}_{p_2}, \nu_2), \\ &a_\lambda = b_\lambda = 0.001, \end{aligned} \tag{19}$$

where we use hierarchical Bayesian modeling to esitmate Σ_β and Σ_γ using Inverse-Wishart distribution. Note that if this step seems to be unnecessary, we can instead simply fix those such as $\Sigma_\beta = \Sigma_\gamma = 10^4 \times \mathbf{I}$.

Sampling scheme

In the survival regression setting, closed forms for the posterior distribution of $\boldsymbol{\beta}$ (as well as $\boldsymbol{\gamma}$ in our model) are generally not available (which is also our cases: Equation (10), (12), (16), (18)), and therefore one needs to use numerical integration or Markov chain Monte Carlo (MCMC) methods. Here we will use MCMC methods with the following update scheme:

Step 0. Choose an arbitrary starting point $\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0$, (and λ_0 if Weibull) and set $i = 0$.

Step 1. Update $\Sigma_\beta \sim \pi(\Sigma_\beta|\boldsymbol{\beta}_i)$ and $\Sigma_\gamma \sim \pi(\Sigma_\gamma|\boldsymbol{\gamma}_i)$ using Gibbs sampler.

Step 2. Update $\boldsymbol{\beta} \sim \pi(\boldsymbol{\beta}|\mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\gamma}, \lambda, \boldsymbol{\mu}_\beta, \Sigma_\beta)$ and $\boldsymbol{\gamma} \sim \pi(\boldsymbol{\gamma}|\mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\beta}, \lambda, \boldsymbol{\mu}_\gamma, \Sigma_\gamma)$ using slice sampling.

Step 2′. If Weibull, update $\lambda \sim \pi(\lambda|\mathbf{C}, \boldsymbol{\alpha}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\beta}, \boldsymbol{\gamma}, a_\lambda, b_\lambda)$ using slice sampling.

Step 3. Set $i = i + 1$, and go to Step 1.

Step 4. After N iterations, summarize the parameter estimates using all sampled values (e.g. confidence intervals for coefficient estimation)

Gibbs sampling for Σ_β and Σ_γ

The closed form of full conditional distributions of $\pi(\Sigma_\beta|\boldsymbol{\beta}_i)$ and $\pi(\Sigma_\gamma|\boldsymbol{\gamma}_i)$ in Step 1 are derived as below:

1. Σ_β:

$$\begin{aligned}
\pi(\Sigma_\beta|\boldsymbol{\beta}) &\propto \pi(\boldsymbol{\beta}|\boldsymbol{\mu}_\beta=\mathbf{0},\Sigma_\beta)\times\pi(\Sigma_\beta)\\
&\propto |\Sigma_\beta|^{-\frac{1}{2}}\exp\Big\{-\frac{1}{2}\Big(\boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\beta}\Big)\Big\}\times|\Sigma_\beta|^{-\frac{\nu_1+p_1+1}{2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}\Big(\mathbf{I}_{p_1}\Sigma_\beta^{-1}\Big)\Big\}\\
&=|\Sigma_\beta|^{-\frac{1+\nu_1+p_1+1}{2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}\Big(\boldsymbol{\beta}\boldsymbol{\beta}'\Sigma_\beta^{-1}\Big)\Big\}\times\exp\Big\{-\frac{1}{2}\mathrm{tr}\Big(\mathbf{I}_{p_1}\Sigma_\beta^{-1}\Big)\Big\}\\
&=|\Sigma_\beta|^{-\frac{1+\nu_1+p_1+1}{2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}\Big((\boldsymbol{\beta}\boldsymbol{\beta}'+\mathbf{I}_{p_1})\Sigma_\beta^{-1}\Big)\Big\}\\
&\sim \text{Inverse-Wishart}(\boldsymbol{\beta}\boldsymbol{\beta}'+\mathbf{I}_{p_1},\quad 1+\nu_1)
\end{aligned}$$

2. Σ_γ:

$$\begin{aligned}
\pi(\Sigma_\gamma|\boldsymbol{\gamma}) &\propto \pi(\boldsymbol{\gamma}|\boldsymbol{\mu}_\gamma=\mathbf{0},\Sigma_\gamma)\times\pi(\Sigma_\gamma)\\
&\propto |\Sigma_\gamma|^{-\frac{1}{2}}\exp\Big\{-\frac{1}{2}\Big(\boldsymbol{\gamma}'\Sigma_\gamma^{-1}\boldsymbol{\gamma}\Big)\Big\}\times|\Sigma_\gamma|^{-\frac{\nu_1+p_2+1}{2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}\Big(\mathbf{I}_{p_2}\Sigma_\gamma^{-1}\Big)\Big\}\\
&=|\Sigma_\gamma|^{-\frac{1+\nu_2+p_2+1}{2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}\Big(\boldsymbol{\gamma}\boldsymbol{\gamma}'\Sigma_\gamma^{-1}\Big)\Big\}\times\exp\Big\{-\frac{1}{2}\mathrm{tr}\Big(\mathbf{I}_{p_2}\Sigma_\gamma^{-1}\Big)\Big\}\\
&=|\Sigma_\gamma|^{-\frac{1+\nu_2+p_2+1}{2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}\Big((\boldsymbol{\gamma}\boldsymbol{\gamma}'+\mathbf{I}_{p_2})\Sigma_\gamma^{-1}\Big)\Big\}\\
&\sim \text{Inverse-Wishart}(\boldsymbol{\gamma}\boldsymbol{\gamma}'+\mathbf{I}_{p_2},\quad 1+\nu_2)
\end{aligned}$$

Slice Sampling for $\boldsymbol{\beta},\boldsymbol{\gamma}$ and λ

In recent decade, slice sampling has been widely used as an alternative to Metropolis-Hastings algorithm. Following the current practice of Bayesian mixture survival model, we use univariate slice sampler with stepout and shrinkage (Neal, 2003) in Step 2 (and Step 2$'$ if Weibull), where closed form of full conditional distribution is intractable. We also follow the modifications made in 'BayesMixSurv' R package (Mahani, Mansour, and Mahani, 2016). Below are the steps to perform slice sampling for $\boldsymbol{\beta}$, and slice sampling for $\boldsymbol{\gamma}$ and λ could be done in the exactly same manner:

For β_p, $p = 1, ..., P$,

Step 0. Choose an arbitrary starting point β_{p0} and size of the slice w, and set $i = 0$.

Step 1. Draw y from Uniform$(0, f(\beta_{p0}))$ defining the slice $S = \{\beta_p : y < f(\beta_p)\}$, where

$$\begin{aligned} f(\beta_p) &\propto \pi(\beta_p | \boldsymbol{\beta}_{-p}, \mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\gamma}) && \text{if Exponential (Eq. (10))} \\ &\propto \pi(\beta_p | \boldsymbol{\beta}_{-p}, \mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \boldsymbol{\gamma}, \lambda) && \text{if Weibull (Eq. (16))} \end{aligned}$$

Step 2. Find an interva, $I = (L, R)$, around β_{p0} that contains all, or much, of the slice, where the initial interval is determined as

$$u \sim \text{Uniform}(0, w)$$
$$L = \beta_{p0} - u$$
$$R = \beta_{p0} + (w - u)$$

and expand the interval until its ends are outside the slice or until the limit on steps (limit on steps $= m$) is reached (“stepping-out" procedure), by comparing y and $(f(L), f(R))$.

$$J = \text{Floor}(\text{Uniform}(0, m))$$
$$K = (m - 1) - J$$
$$\text{Repeat while } J > 0 \text{ and } y < f(L):$$
$$L = L - w, J = J - 1$$
$$\text{Repeat while } K > 0 \text{ and } y < f(R):$$
$$R = R + w, K = K - 1$$

Step 3. Draw a new point β_{p_1} from the part of the slice within this interval I, and shrink the interval on each rejection ("shrinkage" procedure)

$$
\begin{aligned}
&\text{Repeat}\\
&\beta_{p1} \sim \text{Uniform}(L, R)\\
&\ \text{if } y < f(\beta_{p_1}),\ \text{accept } \beta_{p_1} \text{ then exit loop}\\
&\ \text{if } \beta_{p_1} < \beta_{p_0},\ \text{then } L = \beta_{p_1}\\
&\qquad\qquad\qquad\ \ \text{else } R = \beta_{p_1}
\end{aligned}
$$

Step 4. Set $i = i + 1$, $\beta_{p_0} = \beta_{p_1}$, and go to Step 1.

Step 5. After N iterations, summarize the parameter estimates using all sampled values (e.g. confidence intervals for coefficient estimation)