Bayesian Inference on Parametric Mixture Cure Model with Spatial Autocorrelation in Random Effects

Bomin Kim

June 30, 2017

Abstract

In this paper, we implement Bayesian inference on parametric mixture cure model, one of the most popular models to estimate the cure rate of treatment and the survival rate of uncured patients at the same time, with spatial autocorrelation in random effects. This includes two parametric survival models, Exponential and Weibull.

Review on Parametric Mixture Cure Model

Likelihood function

We consider a model for the duration t which splits the sample into two groups, one of which will eventually experience the event of interest (i.e., "fail") and the other which will not. We define the latent variable Y such that $Y_i = 1$ for those who will eventually fail and $Y_i = 0$ for those who will not; define $Pr(Y_i = 1) = \delta_i$ such that the probability δ_i is modeled as a logit by including explanatory variables:

$$\delta_i = \frac{\exp(Z_i\gamma)}{1 + \exp(Z_i\gamma)}. \tag{1}$$

The likelihood function derived by Box-Steffensmeier and Zorn (1999) is:

$$L = \prod_{i=1}^{N}[\delta_i g(t_i|X_i, \beta)]^{C_i}[1 - \delta_i + \delta_i G(t_i|X_i, \beta)]^{1-C_i}, \tag{2}$$

where $g(t_i|X_i, \beta) = f(t_i * Y_i = 1|X_i, \beta)$ and $G(t_i|X_i, \beta) = S(t_i * Y_i = 1|X_i, \beta)$. Here, we follow the common definition in survival model liturate such that $f(t_i|X_i, \beta)$ is the density function and $S(t_i|X_i, \beta) = Pr(T_i > t_i|X_i, \beta)$, with β and γ are the parameter vectors to be estimated.

Then the corresponding log-likelihood function is

$$\ln L = \sum_{i=1}^{N} C_i[\ln \delta_i + \ln g(t_i|X_i, \beta)] + (1 - C_i)\ln[1 - \delta_i + \delta_i G(t_i|X_i, \beta)]. \tag{3}$$

We will expand this mixture cure model to be the spatial-cure model, by incorporating spatial autocorrelation as a random effect.

Derivation of Parametric Spatial Mixture Cure Model

Including a random effect or "frailty" term, we define the proportional hazards $h(t_i|X_i, \beta)$ as:

$$\begin{aligned}
h(t_i|X_i, \beta) &= h_0(t_i)\omega_i \exp(X_i\beta) \\
&= h_0(t_i) \exp(X_i\beta + W_i),
\end{aligned} \tag{4}$$

where $h_0(t_i)$ is the baseline hazard and $W_i \equiv \ln \omega_i$ is the individual frailty term, designed to capture differences among the individuals.

For cure model, we can also add the random effect term to $\Pr(Y_i = 1)$ such that

$$\delta_i = \frac{\exp(Z_i\gamma + W_i)}{1 + \exp(Z_i\gamma + W_i)}. \tag{5}$$

Should we use shared random effect W_i? or define different random effect V_i? (as we did for γ)

Then, we can define the likelihood functions for Exponential and Weibull models:

1. Exponential

$$\begin{aligned}
f(t_i|X_i, \beta, W_i) &= \exp(X_i\beta + W_i)\exp(-\exp(X_i\beta + W_i)t_i), \\
S(t_i|X_i, \beta, W_i) &= \exp(-\exp(X_i\boldsymbol{\beta} + W_i)t_i),
\end{aligned} \tag{6}$$

which establish the likelihood and log-likelihood function respectively as

$$\begin{aligned}
L(\beta, \gamma, \mathbf{W}) = \prod_{i=1}^{N} &[\delta_i\exp(X_i\beta + W_i)\exp(-\exp(X_i\beta + W_i)t_i)]^{C_i} \\
&\times [1 - \delta_i + \delta_i\exp(-\exp(X_i\boldsymbol{\beta} + W_i)t_i)]^{1-C_i},
\end{aligned} \tag{7}$$

$$\begin{aligned}
\ln L(\beta, \gamma, \mathbf{W}) = \sum_{i=1}^{N} &C_i[\ln \delta_i + (X_i\beta + W_i) - \exp(X_i\beta + W_i)t_i] \\
&+ (1 - C_i)\ln[1 - \delta_i + \delta_i\exp(-\exp(X_i\boldsymbol{\beta} + W_i)t_i)].
\end{aligned} \tag{8}$$

2. Weibull

$$f(t_i|\lambda, X_i, \beta, W_i) = \lambda(\exp(X_i\beta + W_i))(\exp(X_i\beta + W_i)t_i)^{\lambda-1}\exp(-(\exp(X_i\beta + W_i)t_i)^{\lambda})$$

$$S(t_i|\lambda, X_i, \beta, W_i) = \exp(-(\exp(X_i\beta + W_i)t_i)^{\lambda}),$$

(9)

which establish the likelihood and log-likelihood function respectively as

$$L(\lambda, \beta, \gamma, \mathbf{W}) = \prod_{i=1}^{N}[\delta_i\lambda(\exp(X_i\beta + W_i))(\exp(X_i\beta + W_i)t_i)^{\lambda-1}\exp(-(\exp(X_i\beta + W_i)t_i)^{\lambda})]^{C_i}$$

$$\times [1 - \delta_i + \delta_i\exp(-(\exp(X_i\beta + W_i)t_i)^{\lambda})]^{1-C_i},$$

(10)

$$\ln L(\lambda, \beta, \gamma, \mathbf{W}) = \sum_{i=1}^{N} C_i[\ln \delta_i + \ln \lambda + (X_i\beta + W_i) + (\lambda - 1)(X_i\beta + W_i + \ln t_i)$$

$$- \lambda\exp(X_i\beta + W_i)t_i] + (1 - C_i)\ln[1 - \delta_i + \delta_i\exp(-(\exp(X_i\beta + W_i)t_i)^{\lambda})].$$

(11)

For Bayesian inference, we use simplest specifications for the priors, for example,:

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

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

$$\lambda \sim \text{Gamma}(a_\lambda, b_\lambda) \text{ for Weibull.}$$

(12)

For the hyperparameters, we can follow the common approach and use

$$\boldsymbol{\mu}_\beta = \mathbf{0}, \Sigma_\beta \sim \text{Inverse-Wishart}(p_1\mathbf{I}_{p_1}, p_1),$$

$$\boldsymbol{\mu}_\gamma = \mathbf{0}, \Sigma_\gamma \sim \text{Inverse-Wishart}(p_2\mathbf{I}_{p_2}, p_2),$$

$$a_\lambda = b_\lambda = 0.001,$$

(13)

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 $\Sigma_\beta = \Sigma_\gamma = 10^4 \times \mathbf{I}$ (very slow mixing).

Priors using Conditionally Autoregressive (CAR) model

Now, we need to define the priors for the random effect $\mathbf{W} = \{W_1, ..., W_N\}$ to assign spatial correlation across the individuals. We follow the canonical approach and apply Conditionally Autoregressive (CAR) model.