Bayesian Inference For Survival Data With a Surviving Fraction

By Debajyoti Sinha * , Ming-Hui Chen † , and Joseph G. Ibrahim ‡

Medical University of South Carolina, University of Connecticut, and University of North Carolina

We develop Bayesian methods for right censored survival time data for populations with a surviving fraction, i.e., with a positive probability of cure. Our model, based on a two-stage process for the development of cancer, has a proportional hazards structure with the covariates depending naturally on the cure rate. We derive several properties of the model and physical interpretation of the model parameters. We also establish mathematical relationships of our model with other existing survival and cure rate models. The new model is computationally appealing, and novel computational Markov chain Monte Carlo (MCMC) methods are developed to sample from the posterior distribution of the parameters. We characterize the propriety of the joint posterior distribution of the parameters using a class of noninformative improper priors. Our model is flexible enough for implementing parametric as well as semiparametric Bayesian inference. We discuss novel nonparametric prior process and very practical prior elicitation for such purposes. A multivariate extension of this model is also possible. A real dataset from a melanoma clinical trial is presented to illustrate our methodology.

1. Introduction. Survival models incorporating a cure fraction, often referred to as *cure rate models*, have been used for modeling time-to-event data for various types of cancers, including breast cancer, non-Hodgkins lymphoma, leukemia, prostate cancer, melanoma, and head and neck cancer, where for these diseases, a significant proportion of patients are "cured". Perhaps the most popular type of cure rate model is the mixture model discussed by Berkson and Gage [1]. In this model, the survivor function for the entire population, denoted by $S_1(t)$, is given by

$$(1) \qquad S_1(t) = \pi + (1 - \pi)S^*(t),$$

where a fraction π of the population are considered "cured", and the remaining $1 - \pi$ are not cured. $S^*(t)$ denotes the survivor function for the non-cured group in the population. Popular choices for $S^*(t)$ are the exponential and Weibull distributions. We shall refer to the model in (1) as the *BG model*. The BG model has been extensively discussed in the statistical literature by several authors, including recently by Taylor [21], Ewell and Ibrahim [7], Stangl and Greenhouse [19], and Sy and Taylor [20]. We refer the reader interested in details about the frequentist methods for the BG model to these articles and to [15] and the references therein.