Supplement: Bayesian Nonparametric Estimation for Dynamic Treatment Regimes with Sequential Transition Times

A: Details of MCMC for Fitting DDP-GP Model

Summary of Model

$$p(y_i^k \mid \boldsymbol{x}_i^k, F^k) = F^k(y_i^k \mid \boldsymbol{x}_i^k)$$
$$F^k \sim \text{DDP-GP}\left\{\{\mu_h^k\}, C^k; \alpha^k, \{\boldsymbol{\beta}_h^k\}, \sigma^k\right\} \tag{1}$$

$$F^k(y \mid \boldsymbol{x}^k) = \sum_{h=0}^{\infty} w_h^k \, N(y;\ \theta_h^k(\boldsymbol{x}^k), \sigma^k). \tag{2}$$

$$\{\theta_h^k(\boldsymbol{x}^k)\} \sim\ GP(\mu_h^k(\boldsymbol{x}^k), C^k(\boldsymbol{x}^k)). \qquad h = 1, 2, ...$$

$$\mu_h^k(\boldsymbol{x}_i^k) = \boldsymbol{x}_i^k \boldsymbol{\beta}_h^k.$$

$k = 1, \ldots, n_{\text{trans}}$. We complete the model construction by assuming $\boldsymbol{\beta}_h^k \sim N(\boldsymbol{\beta}_0^k, \Sigma_0^k)$, $(\sigma^k)^{-2} \overset{\text{i.i.d.}}{\sim} \text{Ga}(\lambda_1, \lambda_2)$ and $\alpha^k \overset{\text{i.i.d.}}{\sim} \text{Ga}(\lambda_3, \lambda_4)$.

Posterior Computation

To evaluate the posterior in a DDP-GP model, we first marginalize (1) analytically with respect to the random probability measures $F^k(\cdot|\boldsymbol{x}^k)$. The form is not obvious from the earlier definition. Temporarily suppress the superscripted transition index k. Consider generating a sample $(Y_1, \boldsymbol{x}_1), \cdots, (Y_n, \boldsymbol{x}_n)$ by first sampling from a covariate distribution, $p(\boldsymbol{x})$, and then from a conditional transition time distribution, $F(\cdot|\boldsymbol{x})$. We rewrite (2) as a hierarchical model with a new latent indicator variable γ_i for the normal mixture summand index h,

$$(Y_i \mid \gamma_i = h,\ \ \boldsymbol{x}_i) \sim N(\theta_h(\boldsymbol{x}_i), \sigma^2) \ \text{ and } \ p(\gamma_i = h) = w_h, \tag{3}$$

for $i = 1, \cdots, n$. Let $\widetilde{\theta}_i(\cdot) = \theta_{\gamma_i}(\cdot)$ denote a realization of the stochastic process selected by γ_i. Next, we re-index the $\theta_h(\cdot)$ such that $\sum_{i=1}^n I(\gamma_i = h) \geq 1$ for $h = 1, \ldots, H$. That is, we let $h = 1, \ldots, H$ index the realizations $\theta_h(\cdot)$ that are selected by some of the γ_i's, so

that $\{\theta_1, \cdots, \theta_H\}$ are the unique values of the n realizations $\{\widetilde{\theta}_i,\ i = 1, \ldots, n\}$. If clusters of patients are defined as $S_h = \{i : \ \widetilde{\theta}_i = \theta_h\}$, then the γ_i's are interpreted as cluster membership indicators. Posterior simulation makes use of these indicators and the vectors $\boldsymbol{\theta}_h = (\theta_h(\boldsymbol{x}_1), \ldots, \theta_h(\boldsymbol{x}_n))$. After marginalization with respect to F_x, we are left with the marginal model for $\{\gamma_i,\ \boldsymbol{\theta}_h(\boldsymbol{x}_i);\ i = 1, \ldots, n,\ h = 1, \ldots, H\}$.

For each transition k, we update parameters using finite DP algorithm as follows. Denote $\#\{i : \gamma_i^k = h\} = n_h^k$.

- Update σ^k

$$(\sigma^k)^2 \mid \cdot \sim \text{Inverse Gamma}(\lambda_1 + \frac{n^k}{2}, \lambda_2 + \frac{\sum_{h=1}^{H}\sum_{\gamma_i^k=h}(y_i^k - \theta_h^k(\boldsymbol{x}_i))^2}{2}) \tag{4}$$

- Update $\boldsymbol{\theta}_h^k$

$$\begin{aligned} p(\boldsymbol{\theta}_h^k \mid \cdot) &\propto p(\boldsymbol{\theta}_h^k) \prod_{i:\gamma_i^k=h} p(y_i^k \mid \theta_h^k(\boldsymbol{x}_i^k)) \\ &\propto \exp\{-\frac{1}{2}(\boldsymbol{\theta}_h^k - \boldsymbol{X}^k\boldsymbol{\beta}_h^k)'(C^k)^{-1}(\boldsymbol{\theta}_h - \boldsymbol{X}^k\boldsymbol{\beta}_h^k)\} \times \exp\{-\frac{\sum_{i:\gamma_i^k=h}(y_i^k - \theta_h^k(\boldsymbol{x}_i))^2}{2(\sigma^k)^2}\} \\ &\sim \text{N}(((C^k)^{-1} + \frac{U'U}{(\sigma^k)^2}I)^{-1}(U'\frac{\boldsymbol{y}_h^k}{(\sigma^k)^2} + (C^k)^{-1}\boldsymbol{X}^k\boldsymbol{\beta}_h^k), ((C^k)^{-1} + \frac{U'U}{(\sigma^k)^2}I_{n^k\times n^k})^{-1}), \end{aligned}$$

where $\boldsymbol{y}_h^k = \{y_i^k, \gamma_i^k = h\}$, $I_{n^k\times n^k}$ is an $n^k \times n^k$ identity matrix, $\boldsymbol{X}^k$ is an $n^k \times M^k$ matrix with the covariates $\boldsymbol{x}_i^k$ of the i-th patient in row i. U is a $n_h^k \times n^k$ matrix: if patient i is the j-th element of $\gamma_i^k = r$, then $U_{ji} = 1$. All other elements are 0.

- Update $\boldsymbol{\beta}_h^k$

$$\begin{aligned} p(\boldsymbol{\beta}_h^k \mid \cdot) &\propto p(\boldsymbol{\beta}_h^k)\exp\{-\frac{1}{2}(\boldsymbol{\theta}_h^k - \boldsymbol{X}^k\boldsymbol{\beta}_h^k)'(C^k)^{-1}(\boldsymbol{\theta}_h^k - \boldsymbol{X}^k\boldsymbol{\beta}_h^k)\} \\ &\sim \text{N}(\Sigma_h^k[((\boldsymbol{X}^k)'(C^k)^{-1}\boldsymbol{\theta}_h^k + \Sigma_0^k\boldsymbol{\beta}_0^k], \Sigma_h^k), \end{aligned}$$

where $\Sigma_h^k = ((\boldsymbol{X}^k)'(C^k)^{-1}\boldsymbol{X}^k + (\Sigma_0^k)^{-1})^{-1}$.

- Update w_h^k

$$v_h^k \sim Beta(1 + n_h^k, \alpha^k + \sum_{j>h} n_j^k),$$

where $n_h^k = \sum_{i=1}^{n_k} I(\gamma_i^k = h)$ is the number of observations such that $\gamma_i^k = h$. Then $w_h^k = v_h^k \prod_{j>h}(1 - v_j^k)$.

- Update γ_i^k

  - If y_i^k is not censored,

$$Pr(\gamma_i^k = h \mid \cdot) \quad \propto \quad w_h^k \int p(y_i^k \mid \theta_h^k(\boldsymbol{x}_i)) p(\theta_h^k(\boldsymbol{x}_i^k) \mid \theta_h^k(\boldsymbol{x}_{-i}^k)) d(\theta_h^k(\boldsymbol{x}_i^k)),$$

    where $\boldsymbol{x}_{-i}^k = \{\boldsymbol{x}_j^k : \gamma_j^k = h, j \neq i\}$.

  - If y_i^k is censored, Let

$$p_h^k(t) = \int p(y_i^k \mid \theta_h^k(\boldsymbol{x}_i)) p(\theta_h^k(\boldsymbol{x}_i^k) \mid \theta_h^k(\boldsymbol{x}_{-i}^k)) d(\theta_h^k(\boldsymbol{x}_i^k)).$$

    Then

$$Pr(\gamma_i^k = h \mid \cdot) \quad \propto \quad \int_{V_i^k}^{\infty} w_h^k p_h^k(t) dt,$$

    where V_i^k is the observed time for patient i in transition k.

- Update α^k

  Using data augmentation, we first sample an m from beta distribution beta$(\alpha^k + 1, n^k)$. Then we sample the new α^k value from

$$\alpha^k \sim \pi \text{Ga}(\lambda_3 + H, \lambda_4 - \log(m)) + (1 - \pi)\text{Ga}(\lambda_3 + H - 1, \lambda_4 - \log(m)),$$

  where $\frac{\pi}{1-\pi} = \frac{\lambda_3 + H - 1}{n^k(\lambda_4 - \log(m))}$.

B: Survival Time Regression Simulation

This simulation was designed to study the DDP-GP regression model by comparing inference for a survival function with the simulation truth. In this study, we did not evaluate a regime effect, but rather focused on inference for the survival curve.

For each subject, we generated T = survival time, the covariates x_1 = tumor size (0=small, 1=large) and x_2 = body weight, and x_3 = a biomarker (0=absent, 1=present). We assumed that small and large tumor sizes each had probability .50. Body weights were computed by sampling from a uniform distribution, Unif$(80, 150)$, with the covariate x_2 defined by shifting and scaling to obtain mean 0 and variance 1. The biomarker was associated with tumor size, as follows. Patients in the large tumor size group were biomarker negative with probability 0.7 and biomarker positive with probability 0.3. Patients with small tumor size were biomarker negative with probability 0.3 and biomarker positive with probability 0.7. Let $Y \sim \text{LN}(m, s)$ denote a lognormal random variable $Y = \log T$ for $T \sim \text{N}(m, s)$. By a slight abuse of notation, we also use LN(m, s) to denote the lognormal p.d.f. Let $\boldsymbol{x}_i = (1, x_{i,1}, x_{i,2}, x_{i,3})$ denote the covariates for patient i, here we include 1 in the covariate to indicate the intercept. We simulated each sample $Y_1, \cdots, Y_n$ of n observations from a mixture of lognormal distributions, $Y_i|\boldsymbol{x}_i \sim 0.4\ \text{LN}(\boldsymbol{x}_i\boldsymbol{\beta}_1, \sigma^2) + 0.6\ \text{LN}(\boldsymbol{x}_i\boldsymbol{\beta}_2, \sigma^2)$, where the true covariate parameters of the mixture components were $\boldsymbol{\beta}_1 = (1, 2, -2, 1)'$ and $\boldsymbol{\beta}_2 = (2, -1, 3, -3)'$, with $\sigma^2 = 0.4$. For comparison, we also fit an AFT regression model, assuming

$$Y_i \ = \ \log(T_i) = \boldsymbol{x}_i'\boldsymbol{\beta} + \sigma\epsilon_i, \quad i = 1, \ldots, n$$

with ϵ_i following an extreme value distribution, so that T_i follows a Weibull distribution.

In this simulation, we considered four scenarios, with $n = 50, 100$, or 200 observations without censoring or $n = 200$ with 23% censoring. For each scenario, $N = 1,000$ trials were simulated. For each simulated data set we fit a DDP-GP survival regression model $F(Y_i \mid \boldsymbol{x}_i)$. For simulation j, let $\overline{S}(t \mid \boldsymbol{x}) = p(T_{n+1} \geq t \mid \boldsymbol{x}_{n+1,j} = \boldsymbol{x}, data)$ denote the posterior expected survival function for a future patient with covariate $\boldsymbol{x}$. Using the empirical distribution

$\frac{1}{n}\sum_{i=1}^{n}\delta_{\boldsymbol{x}_{ij}}$ to marginalize w.r.t. $\boldsymbol{x}_{n+1,j}$ and averaging across simulations, we get

$$\overline{S}(t) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{n}\sum_{i=1}^{n}\overline{S}(t \mid \boldsymbol{x}_{ij}).$$

Figure S1 compares $\overline{S}(\cdot)$ under the DDP-GP model with the simulation truth

$$S_0(t) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{n}\sum_{i=1}^{n}S_0(t \mid \boldsymbol{x}_{ij}),$$

and maximum likelihood estimates (MLE) under Weibull AFT, Lognormal AFT, and Exponential AFT models. In each scenario, the true curve is given as a solid black solid line, the MLE of the survival functions under the AFT regression model assuming Weibull distribution, Lognormal distribution and Exponential distribution as green, blue, magenta solid lines respectively, and the posterior mean survival function under the DDP-GP model as a solid red line with point-wise 90% credible bands as two dotted red lines.

In all four scenarios, the DDP-GP model based estimate reliably recovered the shape of the true survival function and avoided the excessive bias seen with the Weibull, lognormal and exponential MLE. As expected, the three scenarios without censoring show that increasing sample size gives more accurate estimation. With 23% censoring, the DDP-GP estimate becomes less accurate, but it still is much closer to the simulation truth than the AFT regression models with Weibull, lognormal and exponential distributions.

References

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2013). Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika*, page ast014.

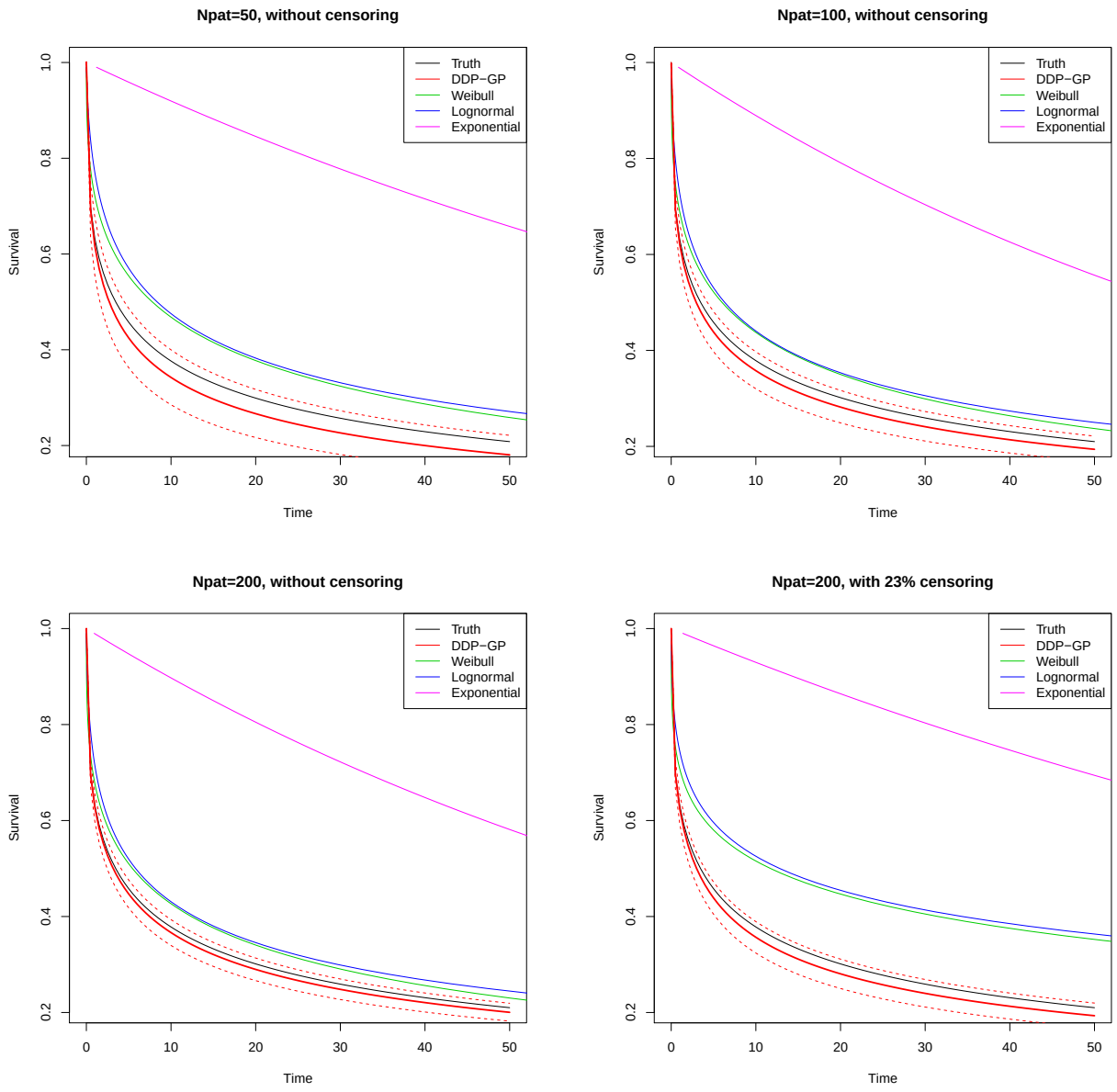

Figure S1: Simulation 1. True mean survival functions (black color) and estimated mean survival functions under the DDP-GP model (red color) for sample sizes $n = 50, 100, 200$ and $n = 200$ with 23% censoring for 1,000 simulations. For comparisons, we also show the MLE under an AFT regression with Weibull distribution (green color), Lognormal distribution (blue color) and Exponential distribution (magenta) . In all cases, the point-wise 90% credible bands are also displayed as the region between two dotted red lines.