

Homework 2 – AMS 276

Mickey Warner

M1 – Weibull baseline

The likelihood for the proportionl hazards model with a Weibull as the baseline hazard function is given by

$$L(\mathbf{y}|\mathbf{x}, \boldsymbol{\nu}, \alpha, \gamma, \beta) = \prod_{\{i:\nu_i=1\}} \alpha \gamma y_i^{\alpha-1} \times \prod_{i=1}^n \exp \left[-\gamma y_i^\alpha e^{x_i^\top \beta} \right]$$

where for subject i , y_i is the length of time in the study, ν_i is an indicator of whether a failure was observed, and x_i is an indicator for the aneuploid group.

I place independent gamma priors on α, γ and a normal prior on β :

$$\begin{aligned} \alpha &\sim \text{Gamma}(1, 0.1) \\ \gamma &\sim \text{Gamma}(1, 0.1) \\ \beta &\sim \text{Normal}(0, 10^2) \end{aligned}$$

The full conditional for the joint parameter vector (α, γ, β) is proportional to the likelihood times each of the priors:

$$\pi(\alpha, \gamma, \beta|\mathbf{y}, \mathbf{x}, \boldsymbol{\nu}) = \prod_{\{i:\nu_i=1\}} \alpha \gamma y_i^{\alpha-1} \times \prod_{i=1}^n \exp \left[-\gamma y_i^\alpha e^{x_i^\top \beta} \right] \times e^{-0.1\alpha} e^{-0.1\gamma} \exp \left(-\frac{\beta^2}{2 \cdot 10^2} \right)$$

The parameters are updated jointly using Metropolis-Hastings updates. Plots of their marginal posterior distributions and an estimate of the survival functions for the two groups (Aneuploid in blue and Diploid in red) are given at the end.

For this model, the survival function is computed by

$$S(t|x, \alpha, \gamma, \beta) = \exp \left[-\gamma t^\alpha \exp(x\beta) \right].$$

M2 – Piece-wise constant

We define a partition $0 = s_0 < s_1 < \dots < s_J$ by the quantiles of the failure times. We let $J = 20$, with s_j being the $(5 \times j)$ th quantile. Define $\delta_{ij} = 1$ if $y_i \in (s_{j-1}, s_j]$ and 0 otherwise.

The likelihood for the piece-wise constant hazard model is given by

$$L(\mathbf{y}|\mathbf{x}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \beta) = \prod_{i=1}^n \prod_{j=1}^J (\lambda_j e^{x_i^\top \beta})^{\delta_{ij} \nu_i} \exp \left(-\delta_{ij} \left[\lambda_j (y_i - s_{j-1}) + \sum_{g=1}^{j-1} \lambda_g (s_g - s_{g-1}) \right] e^{x_i^\top \beta} \right)$$

If $j = 1$, then the summation in the exponent (which is for calculating the cumulative hazard function) becomes 0.

In this model we include an intercept term (described why later). For priors, we have

$$\begin{aligned}\lambda_k &\stackrel{iid}{\sim} \text{Gamma}(1, 0.1), & k = 1, \dots, J \\ \beta_l &\stackrel{iid}{\sim} \text{Normal}(0, 10^2), & l = 0, 1\end{aligned}$$

The full posterior distribution is given by

$$\pi(\boldsymbol{\lambda}, \boldsymbol{\beta} | \mathbf{y}, \mathbf{x}, \boldsymbol{\nu}) \propto L(\mathbf{y} | \mathbf{x}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\beta}) \prod_{k=1}^J p(\lambda_k) \prod_{l=0}^1 p(\beta_l)$$

We again update the parameters jointly, and I believe it is for this reason why it was necessary for me to include an intercept term. I don't fully understand it, but I would not get a correct answer if I exclude the intercept.

At time s_j , the survival function is computed by

$$S(s_j | \mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\beta}) = \exp \left[- \exp(\mathbf{x}^\top \boldsymbol{\beta}) \sum_{g=1}^j \lambda_g (s_g - s_{g-1}) \right].$$

M3 – Gamma process

We use the same partition over the failure times as used in M2. We further define

$$\begin{aligned}R_j &= \{i : s_{j-1} < y_i\} \\ D_j &= \{i : s_{j-1} < y_i \leq s_j, \nu_i = 1\} \\ R_j - D_j &= R_j \cap D_j^c\end{aligned}$$

We assume the baseline cumulative hazard function follows a gamma process, i.e. $H_0 \sim GP(c_0 H^*, c_0)$, where $c_0 > 0$ and $H^*(t) > 0$ is an increasing function. This implies the following prior for the jumps in the cumulative hazard

$$h_j = H_0(s_j) - H_0(s_{j-1}) \stackrel{ind}{\sim} \text{Gamma}(c_0(H^*(s_j) - H^*(s_{j-1})), c_0)$$

We use a grouped-data likelihood:

$$\begin{aligned}L(\mathbf{y} | \mathbf{x}, \boldsymbol{\nu}, \mathbf{h}, \boldsymbol{\beta}) &= \prod_{j=1}^J G_j \\ &= \prod_{j=1}^J \left\{ \exp \left(-h_j \sum_{R_j - D_j} e^{\mathbf{x}_i^\top \boldsymbol{\beta}} \right) \prod_{D_j} \left[1 - \exp(-h_j e^{\mathbf{x}_i^\top \boldsymbol{\beta}}) \right] \right\}\end{aligned}$$

In the case of $D_j = \emptyset$ or $R_j - D_j = \emptyset$, the respective terms in the likelihood give no contribution. As in M2, we again include an intercept term.

H^* is chosen to be Weibull, so that $H^*(t) = \gamma t^\alpha$. The priors we select are

$$\begin{aligned}\alpha &\sim \text{Gamma}(1, 1) \\ \gamma &\sim \text{Gamma}(1, 1) \\ \beta_0 &\sim \text{Normal}(0, 10^2) \\ \beta_1 &\sim \text{Normal}(0, 10^2) \\ c_0 &\sim \text{Gamma}(1, 1)\end{aligned}$$

From M1, we don't expect (α, γ) in this case to be much different than around 1. Placing priors on H^* and c_0 should provide us with additional flexibility.

The full posterior is

$$\pi(\mathbf{h}, \boldsymbol{\beta}, \alpha, \gamma, c_0 | \mathbf{y}, \mathbf{x}, \boldsymbol{\nu}) \prod_{j=1}^J [G_j \times p(h_j | \alpha, \gamma, c_0)] p(\alpha) p(\gamma) p(c_0) p(\beta_0) p(\beta_1)$$

We obtain posterior samples by alternating updates between two groups of parameters: $(\alpha, \gamma, \beta_0, \beta_1, c_0)$ and \mathbf{h} . This is solely for mixing purposes; we use $\pi(\mathbf{h}, \boldsymbol{\beta}, \alpha, \gamma, c_0 | \cdot)$ in both cases, as opposed to the full conditionals for these joint parameter vectors.

At time s_j the survival function is calculated by

$$S(s_j | \mathbf{x}, \mathbf{h}, \boldsymbol{\beta}) = \exp \left[- \exp(\mathbf{x}^\top \boldsymbol{\beta}) \sum_{g=1}^j h_g \right].$$

Remarks about the posteriors and figures

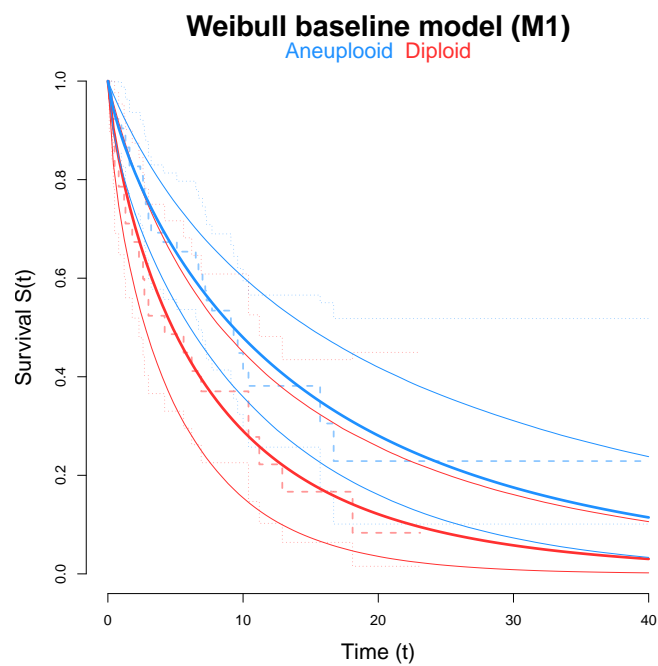
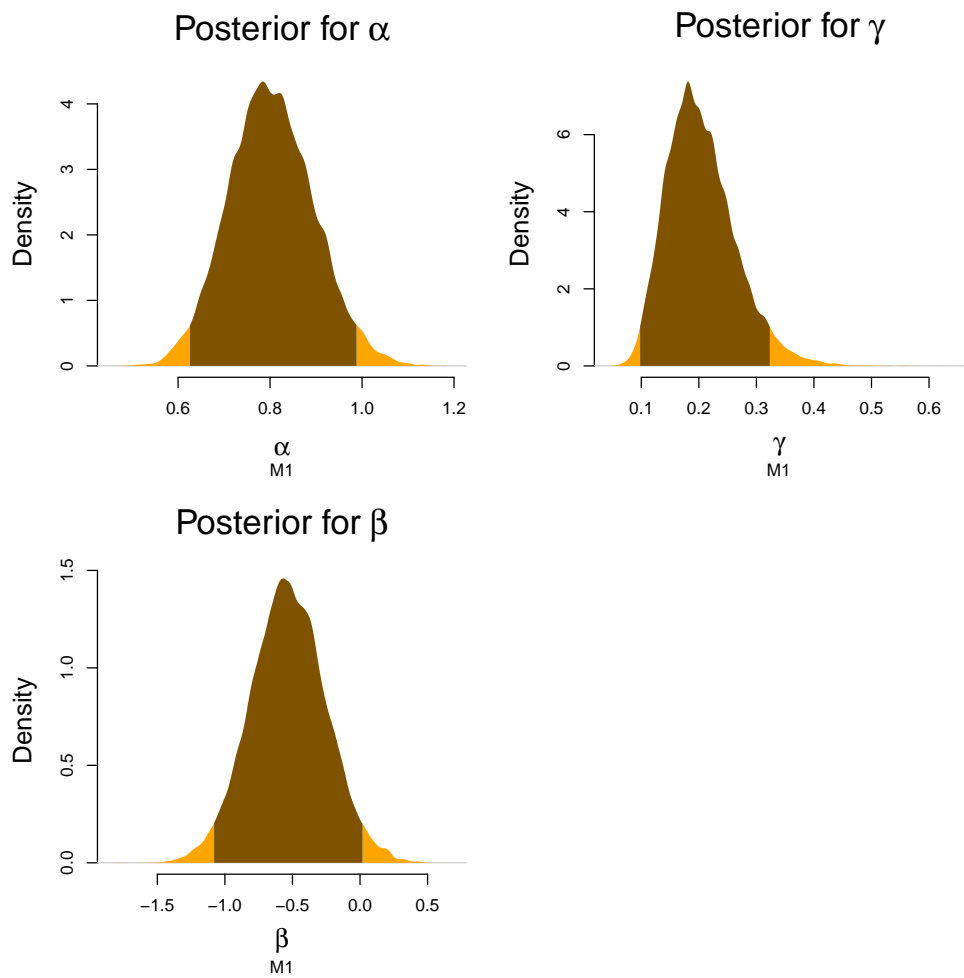
For the most part, there did not seem to be any issue with convergence in the posteriors. When such issues did arise, it was sufficient to increase the number of post-burn-in samples.

For each model, the posterior survival function for the two groups is given. The dashed lines are the same across the models, being the Kaplan-Meier estimates of S . The solid lines are from the posterior samples. The thicker lines are the means and the thinner are 95% credible bands. In every case, our posterior estimates for the survival function seem to agree nicely with K-M. This may lead toward favoring M1, as it is the simplest.

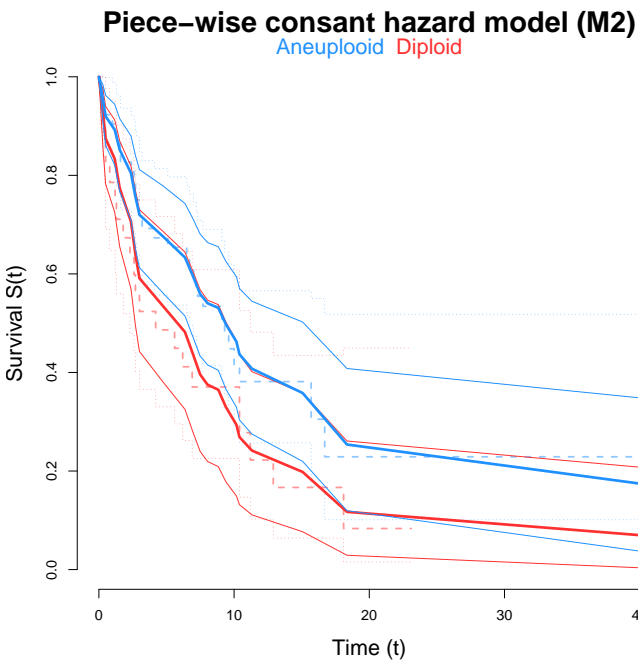
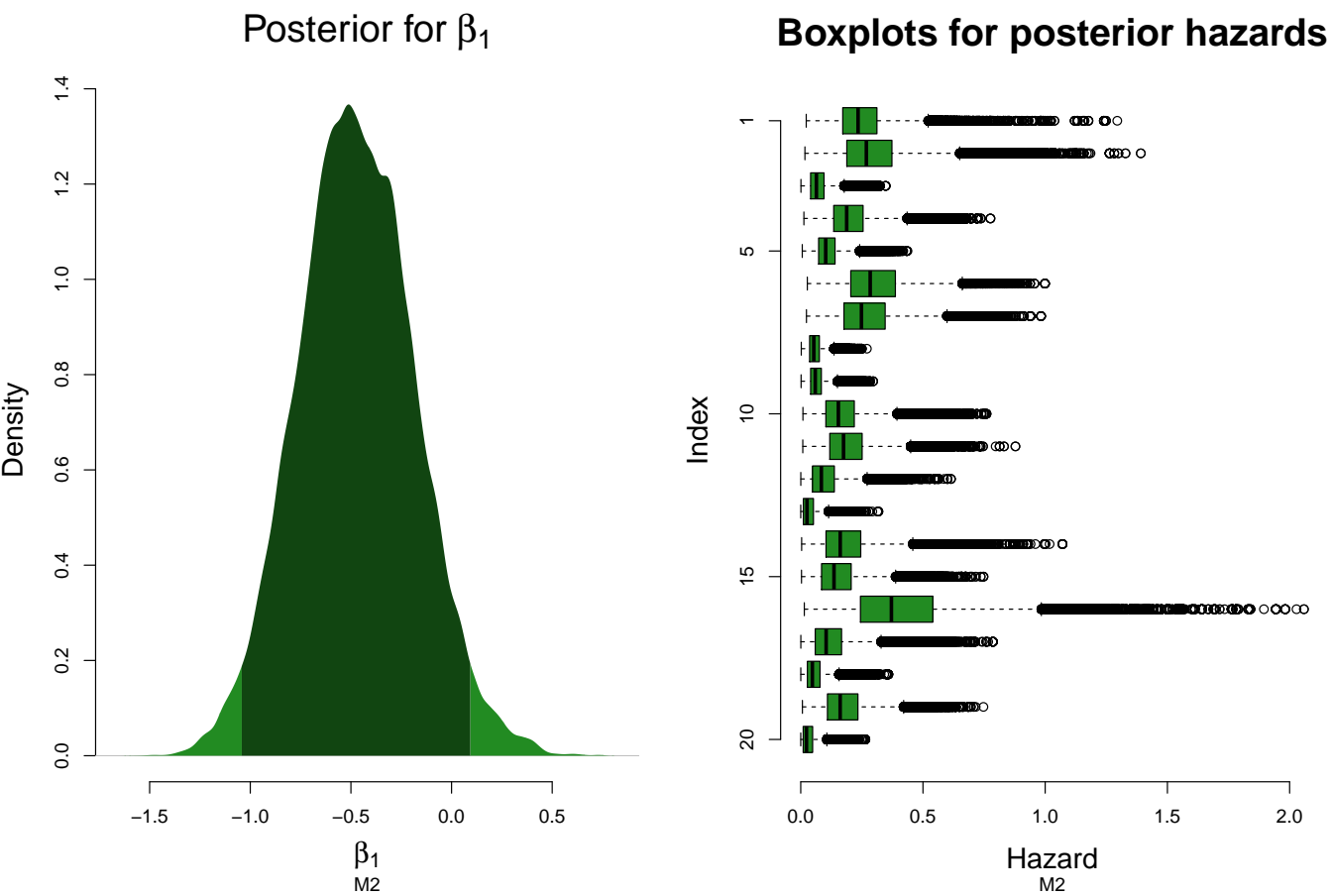
For M1, density estimates and 95% credible intervals are given for each of the three parameters. For M2, the distribution for β_1 is given (comparable to M1's β). Boxplots of the $J = 20$ posterior hazards are shown. Since we had an intercept term, the parameters shown in the boxplot are not the λ_j 's, but rather $\lambda_j \exp(\beta_0)$, with the latter being the height of the hazard function at within $(s_{j-1}, s_j]$. The figures for M3 are comparable to those in M2.

Overall, each model seems appropriate. M1 was far easier to implement than M2 and M3, but is not as robust. My main issue has been with the necessary inclusion of the intercept term β_0 for M2 and M3. I suspect it is because I did not update the parameters one at a time with their full conditionals, but I did not think that should have been an issue.

Figures for M1



Figures for M2



Figures for M3

