Stats 579 – Intermediate Bayesian Modeling

Assignment # 3

1. An hypothetical study considers the lifespan of fluorescent light bulbs. Let $y_1, ..., y_n$ be the duration (in years) it takes for each of n light bulbs to fail. Assume that all tests are performed under laboratory conditions and observations are *iid*. Researchers are interested in whether bulb lifespan is better modeled with model M_1, an $\text{Exp}(\lambda)$ distribution, or with model M_2, a $\text{Weibull}(3, \lambda)$ distribution. For both models, the researchers assume $p(\lambda) = e^{-\lambda}$.

   For this problem, please use the Exponential and Weibull parameterizations from your textbook, which give

$$
\begin{aligned}
y_E &\sim \text{Exp}(\lambda) \\
f(y_E \mid \lambda) &= \lambda \exp(-\lambda y) \, I_{(0,\infty)}(y) \\
y_W &\sim \text{Weibull}(\alpha, \lambda) \\
f(y_W \mid \alpha, \lambda) &= \lambda \alpha y^{\alpha-1} \exp(-\lambda y^{\alpha}) \, I_{(0,\infty)}(y)
\end{aligned}
$$

   (a) Obtain the marginal density for these data under each model. (HINT: Take advantage of conjugacy.)

   (b) Obtain an expression for the Bayes factor comparing M_1 to M_2.

   (c) Evaluate the Bayes factor when the data are: $\{8.05, 6.56, 3.20, 6.85, 5.67\}$.

   (d) Explain which model seems preferable based on the Bayes factor. Explain which model would be preferable if you had a prior belief that M_1 were nine times more likely to be correct than M_2.

2. Using the same set-up as in Problem 1, the researchers want to compare these models in terms of AIC.

   (a) Find the MLE for λ under M_1 and M_2. Note that although both models have a parameter named λ, these parameters are not the same and may maximize at different values for each model. It may be helpful to write the MLEs as $\hat{\lambda}_1$ and $\hat{\lambda}_2$ to help distinguish them.

   (b) Calculate AIC for each model using the data provided above. Which model seems preferable based on AIC?

3. Let $y_1, ..., y_n$ be independent conditional on some model parameter θ. Let $y_i \sim f(y_i \mid \theta)$ and let θ have prior $p(\theta)$. Consider the conditional predictive ordinate for an observation y_j,

$$
CPO_j = f\left(y_j \mid y_{(j)}\right),
$$

   where $y_{(j)}$ denotes the set $\{y_1, ..., y_{j-1}, y_{j+1}, ..., y_n\}$.

   (a) Show that

$$
CPO_j = \frac{\int \prod_{i=1}^{n} f(y_i \mid \theta) p(\theta)}{\int \prod_{i \neq j} f(y_i \mid \theta) p(\theta)}.
$$

   (b) Now show that

$$
CPO_j^{-1} = \int \left[\frac{1}{f(y_j \mid \theta)} \right] p(\theta \mid y) d\theta.
$$