

Fall 16 – AMS276 Homework 1

Due: Tuesday October 18.

1. (ICS Ex 2.1) Suppose y_1, \dots, y_n is a random sample from an exponential distribution with mean $1/\lambda$. Further suppose that y_1, \dots, y_k are fully observed, and y_{k+1}, \dots, y_n are right censored.
 - (a) Derive Jeffreys's prior for λ . Note: Jeffreys's prior is defined to be any prior which is proportional to the square root of the Fisher information in λ
 - (b) Derive the posterior distribution of λ using the prior in part (a)
 - (c) Using part (b), derive the predictive distribution of a future observation z , where z is assumed to be fully observed, i.e., not censored.
 - (d) Consider including a single covariate x_i in the exponential model, so that $\lambda_i = \exp(x_i\beta)$. Let $\theta = \exp(\beta)$ and suppose that $\theta \sim \text{Gamma}(\alpha_0, \lambda_0)$ a priori. Derive a closed form expression for the posterior distribution of θ when $x_i = 1$ or when x_i is a binary covariate.
2. (ICS Ex 2.3) Suppose n pairs of random variables (y_i, ν_i) are recorded. Assume y_i are iid from a Weibull distribution, $W(\alpha, \gamma)$ with pdf

$$f(y \mid \alpha, \gamma) = \alpha \gamma y^{\alpha-1} \exp(-\gamma y^\alpha).$$

Thus, we can write down the likelihood function of (α, γ) as

$$\mathcal{L}(\alpha, \gamma) = (\alpha \gamma)^{\sum_{i=1}^n \nu_i} \exp \left\{ \sum_{i=1}^n (\nu_i (\alpha - 1) \log(y_i) - \gamma y_i^\alpha) \right\}.$$

- (a) Suppose α is known. Derive the posterior distribution of γ using $\pi(\gamma) \propto \gamma^{-1}$.
 - (b) Using part (a), derive the predictive distribution of a future observation z , where z is assumed to be fully observed, i.e., not censored.
3. (KM Ex 12.1) In Section 1.11, a study of the effects of ploidy on survival for patients with cancer of the tongue was described. In the study patients were classified as having either an aneuploid or diploid DNA profile. The data is presented in Table 1.6. (check their R package for the data and a short description of the data). Use the R commands below to load the data into R;

```
library(KMsurv) # To get the datasets in K-M
library(survival) # R functions
```

```
data(tongue)
tongue$time <- tongue$time/10 ## to obtain computational stability
```

- (a) For both the aneuploid and diploid groups fit a Bayesian Weibull model $W(\alpha, \lambda)$ (one for each, separately) to the data. Discuss your prior distributions. Summarize your posterior and write your conclusions. Do a sensitivity analysis for the prior on α and λ . Comment on α for each group.

- (b) Fit a Weibull AFT model to this data with a single covariate, x , that is equal to 1 if the patient had an aneuploid DNA profile and 0 otherwise. Discuss your prior distributions. Summarize your posterior and write your conclusions. Provide an interpretation of the coefficient.
- (c) Fit log-normal and log-logistic models. For each model, (i) Discuss your prior distributions (ii) Summarize your posterior and write your conclusions.
- (d) Compare the three AFT models based on DIC or some other model comparison criterion. What is your conclusion?