

# Homework 6, STA 360/602

Shenyang Huang

2. Researchers are studying the length of life (lifetime) following a particular medical intervention, such as a new surgical treatment for heart disease, where the study consists of 12 patients. Specifically, the number of years before death for each is

$$3.4, 2.9, 1.2+, 1.4, 3.2, 1.8, 4.6, 1.7+, 2.0+, 1.4+, 2.8, 0.6+$$

where the + indicates that the patient was alive after  $x$  years, but the researchers lost contact with the patient after that point in time.

One way we can model this data is in the following way:

$$X_i = \begin{cases} Z_i & \text{if } Z_i \leq c_i \\ c_i & \text{if } Z_i > c_i \end{cases} \quad (1)$$

$$Z_1, \dots, Z_n | \theta \stackrel{iid}{\sim} \text{Gamma}(r, \theta) \quad (2)$$

$$\theta \sim \text{Gamma}(a, b) \quad (3)$$

where  $a$ ,  $b$ , and  $r$  are known. In addition, we know:

- $c_i$  is the censoring time for patient  $i$ , which is fixed, but known only if censoring occurs.
- $X_i$  is the observation
  - if the lifetime is less than  $c_i$  then we get to observe it ( $X_i = Z_i$ ),
  - otherwise all we know is the lifetime is greater than  $c_i$  ( $X_i = c_i$ ).
- $\theta$  is the parameter of interest—the rate parameter for the lifetime distribution.
- $Z_i$  is the lifetime for patient  $i$ , however, this is not directly observed.

The probability density function (pdf) associated consists of two point masses: one at  $Z_i$  and one at  $c_i$ . The formula is

$$p(x_i | z_i) = \mathbf{1}(x_i = z_i) \mathbf{1}(z_i \leq c_i) + \mathbf{1}(x_i = c_i) \mathbf{1}(z_i > c_i).$$

.

Now we can easily find the full conditionals (derived in class and reproduced below). Notice that  $z_i$  is conditionally independent of  $z_j$  given  $\theta$  for  $i \neq j$ . This implies that  $x_i$  is conditionally independent of  $x_j$  given  $z_i$  for  $i \neq j$ . Now we have

$$\begin{aligned} p(z_i | z_{-i}, x_{1:n}, \theta) &= p(z_i | x_i, \theta) \\ &\propto_{z_i} p(z_i, x_i, \theta) \\ &= p(\theta) p(z_i | \theta) p(x_i | z_i, \theta) \\ &\propto_{z_i} p(z_i | \theta) p(x_i | z_i, \theta) \\ &= p(z_i | \theta) p(x_i | z_i). \end{aligned}$$

There are now two cases to consider. If  $x_i \neq c_i$ , then  $p(z_i|\theta)p(x_i|z_i)$  is only non-zero when  $z_i = x_i$ . The density devolves to a point mass at  $x_i$ . This corresponds to the case where  $z_i$  is observed, so  $x_i$  is the observed value and we should always sample this value. Practically speaking, we do not sample this value when running the Gibbs sampler.

If  $x_i = c_i$ , then the density becomes  $p(x_i|z_i) = \mathbf{1}(z_i > c_i)$ , so

$$p(z_i|\dots) \propto p(z_i|\theta)\mathbf{1}(z_i > c_i),$$

which is a truncated Gamma.

For the Gibbs sampler, we will use the current value of  $\theta$  to impute the censored data. We will sample from the truncated gamma using a modified version of the iverse CDF trick. For the censored values of  $Z_i$  we know  $c_i$ . If we know  $\theta$  (which we will in a Gibbs' sampler), we know the distribution of  $Z_i|\theta \sim \text{Gamma}(r, \theta)$ . Let  $F$  be the CDF of this distribution. Suppose we truncate this distribution to  $(c, \infty)$ . The new CDF is

$$P(Z_i < z) = \frac{F(z) - F(c)}{1 - F(c)}.$$

Therefore  $Y$  is a sample from the truncated Gamma, as desired.

In the actual code for the Gibbs' sampler we do not sample the observed values. We simply impute the censored values using the method above.

You will find code below (that is also taken from class ) that will help you with the remainder of the problem.

1. (5 points) Write code to produce trace plots and running average plots for the censored values for 200 iterations. Do these diagnostic plots suggest that you have run the sampler long enough? Explain.
2. (5 points) Now run the chain for 10,000 iterations and update your diagnostic plots (traceplots and running average plots). Report your findings for both traceplots and the running average plots for  $\theta$  and the censored values. Do these diagnostic plots suggest that you have run the sampler long enough? Explain.
3. (5 points) Give plots of the estimated density of  $\theta | \dots$  and  $z_9 | \dots$ . Be sure to give brief explanations of your results and findings. (Present plots for 10,000 iterations).
4. (5 points) Finally, let's suppose that  $r = 10, a = 1, b = 100$ . Do the posterior densities in part (c) change for  $\theta | \dots$  and  $z_9 | \dots$ ? Do the associated posterior densities change when  $r = 10, a = 100, b = 1$ ? Please provide plots and an explanation to back up your answer. (Use 10,000 iterations for the Gibbs sampler).