STATISTICS 225 – Winter 2003

Exam – Wednesday March 5, 2003

1. Let $y_t, t = 1, \ldots, T$ be a sequence of random variables with

$$y_t \sim \begin{cases} \text{Poisson}(\lambda_1), & t = 1, \ldots, k \\ \text{Poisson}(\lambda_2), & t = k+1, \ldots, T \end{cases}$$

where the Poisson density (mass) function is $\Pr(Y = y) = \lambda^y e^{-\lambda}/y!$ for $y = 0, 1, \ldots$. The y_t are independent given λ_1 and λ_2. This is an example of a changepoint model; k is a known timepoint $(1 \leq k \leq T - 1)$ at which the mean of the process changes.

(a) Suppose that $\lambda_1 \sim \text{Gamma}(a_1, b_1)$ and $\lambda_2 \sim \text{Gamma}(a_2, b_2)$ are chosen as prior distributions where a_1, b_1, a_2, b_2 are specified constants. Recall that the density function for a random variable λ having a gamma distribution with parameters a, b is given by $p(\lambda|a, b) = b^a \lambda^{a-1} e^{-\lambda b}/\Gamma(a)$.

i. Show that λ_1 and λ_2 are independent in their joint posterior distribution.

ii. Identify the posterior distributions of λ_1 and λ_2.

(b) Suppose that we now consider k to be unknown in the changepoint model. We take the prior distribution to be uniform over the possible values $(1, \ldots, T - 1)$.

i. Derive the joint posterior distribution of the parameters of the new model $(\lambda_1, \lambda_2, k)$ up to a normalizing constant.

ii. Describe one approach for simulating from this joint posterior distribution. Give enough details for someone to implement the approach. You don't have to write a computer program but it is not sufficient to say that you would "use MCMC" or "use simulation".

(c) One model checking approach (not discussed in class) is to consider the predictive distribution for y_t given all of the other data. Let Y_{-t} denote all of the data other than y_t so that $Y_{-t} = (y_1, \ldots, y_{t-1}, y_{t+1}, \ldots, y_T)$. For this part use the model with **known** changepoint.

i. Show that for $1 \leq t \leq k$

$$p(y_t|Y_{-t}) = \int p(y_t|\lambda_1) p(\lambda_1|Y_{-t}) d\lambda_1.$$

ii. Assume for the moment that we can compute $p(y_t|Y_{-t})$. Explain how this quantity could be used for model checking.

2. (see reverse)

2. The exponential distribution can be used to model the waiting time for a job to be completed. We consider the situation in which there are I distinct types of jobs. Let y_{ij} be the amount of time that is required to complete the jth job of type i. The time required to complete a job depends on the type of job and on the number of available workers. Let m_{ij} be the number of workers available when the jth job of type i is to be carried out. We assume that

$$y_{ij}|\theta_i \sim \text{Exp}(\theta_i m_{ij}), \quad i = 1, \ldots, I; \; j = 1, \ldots, n_i$$

where $\text{Exp}(\lambda)$ denotes the exponential distribution with parameter λ having probability density function $f(y|\lambda) = \lambda e^{-\lambda y}$ for $\lambda > 0$ and $y > 0$. (This exponential distribution has mean $1/\lambda$.) We assume that the y_{ij}'s are independent random variables given the θ_i's. For this problem we treat the m_{ij}'s as known constants and thus don't include them as variables in our notation.

(a) Are the random variables $(y_{i1}, y_{i2}, \ldots, y_{in_i})$ exchangeable? If you answer YES, then explain why; If you answer NO, then identify conditions on the m_{ij}'s under which the random variables $(y_{i1}, y_{i2}, \ldots, y_{in_i})$ would be exchangeable.

(b) Note that each type of job has its own associated parameter θ_i. Suppose that the parameters θ_i are assumed to be iid random variables with a $\text{Gamma}(\alpha, \beta)$ prior distribution (the density of which is provided in problem 1). Let $p(\alpha, \beta)$ denote the prior distribution on the hyperparameters α, β. Obtain the (unnormalized) joint posterior distribution of all of the unknown parameters conditional on the observed data.

(c) Identify the conditional posterior distribution of $\theta = (\theta_1, \ldots, \theta_I)$ given α, β and y.

(d) Explain how the conditional posterior distribution of θ_i demonstrates the Bayesian principle of combining information from the data and the population model.

(e) Describe one approach for simulating from the joint posterior distribution in part (b). Give enough details for someone to implement the approach. You don't have to write a computer program but it is not sufficient to say that you would "use MCMC" or "use simulation".

(f) Describe how you would obtain a predictive distribution for the time to complete a job of type i given that two workers are available.

(g) Describe how you would obtain a predictive distribution for the time required to complete a job of a new type (not seen in the existing data set) given that two workers are available.