STATISTICS 225 – Fall 2018

Exam – Tuesday November 20, 2018

1. The Pareto distribution is sometimes used in economics to model the distribution of incomes above a specified threshold M. The form of the density function is

$$p(y|\theta, M) = \theta M^{\theta}/y^{\theta+1} \qquad \text{for } y \geq M,$$

where $\theta > 0$ is an unknown parameter. The threshold M is assumed known throughout this problem.

(a) Sketch the distribution of y.

(b) Derive Jeffrey's prior distribution for θ.

Assume that we observe a sample of independent Pareto observations from I subpopulations (e.g., these may be states or regions). Let $y_i = (y_{i1}, \ldots, y_{in_i})$ denote the observations from subpopulation i (for $i = 1, \ldots, I$) and assume that the elements of y_i are independent Pareto observations conditional on the parameter θ_i. Further suppose that the θ_i's are independent draws from a Gamma(α, β) prior distribution (recall that the gamma density is $p(\theta) = \beta^{\alpha}\theta^{\alpha-1}e^{-\beta\theta}/\Gamma(\alpha)$ and the distribution has mean $\frac{\alpha}{\beta}$). Finally, assume that we use the vague hyperprior distribution $p(\alpha, \beta) \propto \beta^{-5/2}$.

(c) Identify the joint posterior distribution of all the unknown parameters up to a normalizing constant.

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

(e) Identify the marginal posterior distribution $p(\alpha, \beta|y)$ (again up to a normalizing constant is sufficient).

(f) Describe a computational approach for obtaining draws from the joint posterior distribution of θ, α, β. You do not need to write any computer code but you should provide enough information about your approach so that someone else could. (Note that there is more than one correct answer here.)

2. Consider the normal linear regression model with no intercept for analyzing bivariate observations $(x_i, y_i), i = 1, \ldots, n$:

$$y_i \mid \beta, \sigma^2, x_i \sim N(\beta x_i, \sigma^2) \qquad i = 1, \ldots, n.$$

In this problem we treat the x_i's as known constants and assume σ^2 is known too. The y_i's are assumed independent given β (and the other known quantities). For the moment we take the prior distribution on β to be $p(\beta)$.

(a) The y_i's are assumed independent conditional on β (and the other known quantities) but they are not independent random variables in their marginal distribution. Explain.

(b) Now assume the prior distribution on β is $N(\beta_o, \tau^2)$ for specified values of β_o and τ^2. Derive the posterior distribution of β.

(c) The traditional least squares approach to this model yields $\hat{\beta} = \sum_i x_i y_i / \sum_i x_i^2$ with variance $\sigma^2/\sum_i x_i^2$. Show that the posterior mean of β is a linear combination of the prior mean and the least squares estimate.

(d) We are interested in making a prediction for a case with predictor value x^*. Identify the posterior predictive distribution for y corresponding to a case with predictor value x^*.

(e) You are concerned that the linear model is not appropriate for the data. Assume that you have obtained a set of simulations from the posterior distribution $(\beta^{(s)}, s = 1, \ldots, S)$. Describe one Bayesian approach for checking the fit of the linear model. Provide enough detail so that someone could carry out the model check.