MATH 640: Bayesian Statistics

Homework 3, due Sunday, February 17

Please submit a PDF or .doc version of your homework to Canvas by 11:59pm on the due date. Please type *all* responses. You are encouraged to use `R` for all calculations.

Theoretical Exercises

1. Multi-parameter distributions often lack convenient conjugate priors (if they have one at all). One such case is when Y_i are iid $Gamma(\alpha, \beta)$ where *both* α and β are unknown. The conjugate prior, while proper, is not a named density. Show that the joint prior

$$p(\alpha, \beta) \propto \frac{\beta^{\alpha s}}{\Gamma(\alpha)^r} p^{\alpha - 1} e^{-\beta q}$$

is actually a conjugate prior for the Gamma distribution with unknown α and β. That is, show that when this joint prior is used, the resulting posterior has the same parametric form. Be sure to determine the parameters. (Hint: this prior is parameterized by p, q, r, and s thus the posterior should have four parameters as well.)

2. Let $\mathbf{y} = [\ y_1 \quad y_2 \quad \cdots \quad y_n\ ]'$ be an $n \times 1$ vector of regression outcomes. Further let X denote an $n \times p$ matrix of covariates and $\boldsymbol{\beta}$ be a $p \times 1$ vector of coefficients. Assume $\mathbf{y}$ is normally distributed of the form

$$\mathbf{y} \sim MVN\left(X\boldsymbol{\beta}, \lambda^{-1} I_{n \times n}\right).$$

That is, the standard regression assumption where we've parametrized the model in terms of the precision, λ. Using the joint prior $\pi(\boldsymbol{\beta}, \lambda) \propto \lambda^{-1}$, find the marginal distribution posterior of $\lambda | \mathbf{y}, X$ and the conditional posterior distribution of $\boldsymbol{\beta} | \lambda, \mathbf{y}, X$.

3. Let $W_i \sim N(\mu, \tau^2)$ for $i = 1, \ldots, n$ where both μ and τ^2 are unknown. Determine the form of normal approximation to the joint posterior of μ and τ^2 when using the non-informative joint prior, i.e. $\pi\left(\mu, \tau^2\right) \propto \left(\tau^2\right)^{-1}$. (Hint: this will require find the posterior modes for both μ and τ^2 as well as the information matrix, i.e. the negative of the Hessian matrix.)

Analysis Exercises

1. The age distribution of the incidence of cancer can be modeled using the Erlang distribution which has as PDF

$$f_X(x; k, \lambda) = \frac{1}{(k-1)!} \lambda^k x^{k-1} e^{-\lambda x}$$

where $x \in [0, \infty)$, $k \in \mathbb{Z}^+$, and $\lambda \in (0, \infty)$. Here the parameter k can be interpreted as the number of carcinogenic events needed for a cancer to develop while $1/\lambda$ is the average time to developing cancer. The data file `incidenceUK.txt` contains age specific incidence of all cancers in both males and females in the United Kingdom for the years 2013 to 2015. Using an Erlang distribution with $k = 22$[1], fixed, find the posterior distribution of the average time to developing cancer in males and females, separately, using the normal approximation to the posterior density. Use Jeffreys' prior for λ. Generate posterior summaries and compare between males and females. Draw a conclusion in context. Use $B = 10000$ samples for each model and set the seed to 2020.

2. The dataset `coup1980.txt` contains the coup risk in the month of June from 1980 for 166 different countries. Using your result from Theoretical Exercise 2, build a linear regression model to predict `logCoup` risk using the covariates `democracy` (1 = yes, 0 = no), `age` (the leader's age in years), and `tenure` (the leader's tenure in months). Conduct relevant inference to determine significant predictors and describe how each variable impacted coup risk during June of 1980. Use $B = 10000$ samples and set the seed to 1980. (Hint: your description of the impact can be an interpretation, in context, of the coefficients.)

[1] Note: 22 is roughly the average number of carcinogenic events needed from the 20 most common cancers.