Homework 1

Show your steps for each problem. Type up answers in LaTeX and turn in by October 14.

1. **Poisson-Gamma**. Datum y is distributed Poisson(λ), where λ is the mean of the Poisson distribution. A priori, $\lambda \sim \text{Gamma}(a, b)$ where $a > 0$ and $b > 0$ are known scalars.

   (a) What is the posterior distribution of λ given y?

   (b) Calculate the normalizing constant $f(y) = \int_0^\infty f(y|\lambda)f(\lambda)d\lambda$.

   (c) Thought of as a distribution in y given a and b, what distribution with what parameters is $f(y|a, b)$? (If you need a hint: Wikipedia Gamma-Poisson mixture).

   (d) Plot the prior and the posterior for λ on the same plot for $y = a = b = 1$. Repeat for $y = 4, a = 8, b = 2$. Label appropriately.

2. **Poisson-Gamma, cont'd**. Data $Y = (y_1, \ldots, y_n)'$ is an n-vector of observations that are independently distributed as Poisson given λ, $y_i|\lambda \sim \text{Poisson}(\lambda)$. A priori, $\lambda \sim \text{Gamma}(a, b)$ where $a > 0$ and $b > 0$ are known scalars.

   (a) What is the posterior distribution of λ given Y?

3. **Half-normal distribution**. The half-normal distribution given parameter τ has density

$$f(y|\tau) = \left(\frac{2}{\pi\tau}\right)^{1/2} \exp\left(-\frac{y^2}{2\tau}\right) \mathbf{1}\{0 < y\}$$

   and let us denote this distribution as

$$y|\tau \sim \text{HN}(y|\tau).$$

   Let $y_i|\tau \sim \text{HN}(y_i|\tau)$ for $i = 1, \ldots, n$ with $Y = (y_1, \ldots, y_n)'$.

   (a) What are the mean, median, mode and variance of the half-normal distribution?

   (b) What is the sufficient statistic?

(c) The inverse gamma distribution $\tau \sim \text{InverseGamma}(a/2, b/2)$ is a convenient and conjugate prior for τ. Derive the posterior of $\tau|Y$.

4. **Half-normal distribution**, cont'd.

   (a) What kind of data might you model using the half-normal distribution? That is, what characteristics/features of the data would lead you to use a half-normal distribution as a sampling model? (Answer with a short list of a few items.)

   This requires you think about what characteristics of a data set are important and that should be reflected in choosing a sampling density. For example, would minimum daily temperatures in Fahrenheit in Fairbanks, Alaska be appropriately modeled by a half-normal density? (yes/no)

   (b) Give a specific example of data that might be modeled by the half-normal density. (Like the Fairbanks temp data of previous question.)

   (c) What competitor distributions might you use in place of the half-normal distribution to model non-negative data?

5. **Half-normal distribution**, cont'd. Suppose you parameterize the halfnormal distribution in terms of the unknown standard deviation, $\sigma = \tau^{1/2}$.

   (a) Write the sampling density of $y_i|\sigma$.

   (b) Is an inverse gamma prior for σ conjugate when $y_i|\sigma \sim \text{HN}(y_i|\sigma^2)$? [Yes/no.] Please explain [1 sentence].

   (c) (Not to be turned in. Advanced problem.) Is there a conjugate prior distribution for σ? What is it?

6. **Power Distribution**. We define the power distribution for $\theta > 0$ and for $0 < y < 1$ to have density

$$f(y|\theta) \propto y^{\theta-1}\mathbf{1}\{0 < y < 1\}.$$

   (a) Find the normalizing constant for $f(y|\theta)$ so that the right hand side is a density.

(b) Give the likelihood for a sample y_i of size n, $i = 1, \ldots, n$, with $Y = (y_1, \ldots, y_n)'$. Find the sufficient statistic. Is there a particularly convenient (i.e. interpretable) form of the sufficient statistic? What is the interpretable form?