Homework 4
Statistics 201B
Due Nov 22, 9:30am

1. Consider a Bayesian model in which, conditional on unknown parameter λ, $X_1, \ldots, X_n$ are iid with exponential PDF

$$f(x|\lambda) = \frac{1}{\lambda} e^{-x/\lambda}$$

for $x > 0$, and the prior distribution is $InverseGamma(a, b)$, with PDF

$$f(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{-a-1} e^{-b/\lambda}$$

for $\lambda > 0$.

   (a) Find the posterior distribution for λ, conditioning on $X_1, \ldots, X_n$. It is fine to write the family and specify its parameters; you do not need to write out the CDF or PDF.

   (b) Show that the posterior mean can be written as a weighted average of the prior mean and the MLE for λ. You may use the fact that the mean of an Inverse Gamma distribution with parameters a and b is $\frac{b}{a-1}$. What happens as $n \to \infty$?

2. Here is an example for which we can use the model in problem 1. Let λ represent the average time (in units of days) between earthquakes in the Berkeley area. To make this more precise, let's consider only earthquakes with magnitude 3 or greater on the Richter scale, and whose epicenter is within a 10 mile radius of downtown Berkeley.

   (a) Consider using an Inverse Gamma prior for λ. We need to choose the parameters a and b. You may have some prior knowledge about λ, but it may be difficult to translate this into a choice of a and b. To facilitate this, write expressions for a and b in terms of the prior mean m and the prior variance v, using that $m = \frac{b}{a-1}$ and $v = \frac{b^2}{(a-1)^2(a-2)}$ when $a > 2$.

(b) Based on your current knowledge, choose parameters a and b, and make a plot of the prior PDF. You may find it useful here and in the rest of the problem to modify the `R` code in the file `BetaBinomial.R`, which is on bCourse. (There is a `dinvgamma` function in the `R` package `MCMCpack`, which you can install and load using `install.packages("MCMCpack")` and then `library(MCMCpack)`, or you can just code the mathematical form of the prior PDF directly.) Turn in a sentence of explanation with your plot regarding how your prior knowledge (or lack of it) informed your choice of prior distribution.

(c) The file `BerkeleyEarthquakes.RData` on bCourse contains a data frame called `earthquakes` with information about earthquakes within a 10 mile radius of Berkeley, from 1969-2008. Load it in and take a look at the first few lines, then extract the waiting time between each earthquake using

```
load("BerkeleyEarthquakes.RData")
head(earthquakes)
x <- earthquakes$Lag[-1] # First element is NA
```

Using the results you found in problem 1, calculate the posterior distribution for λ, conditional on the observed waiting times. Make a plot comparing your posterior PDF to your prior PDF. Turn in a sentence of explanation with your plot regarding any changes in your knowledge about λ after seeing the data.

3. Consider rejection sampling when the target density $h(\theta) = f(\theta|x^n)$. In class we considered taking the proposal density $g(\theta) = f(\theta)$, i.e., the prior PDF. We set $M = \mathcal{L}(\hat{\theta}_n)$, where $\hat{\theta}_n$ is the MLE for θ. Explain why we should not take it to be any smaller. Note that the candidate θ's are sampled from the proposal density.

4. Consider a Bayesian model in which, conditional on unknown parameter λ, $X_1, \ldots, X_n$ are iid with exponential PDF

$$f(x|\lambda) = \frac{1}{\lambda} e^{-x/\lambda}$$

for $x > 0$.

(a) Find the Jeffreys prior for λ.

(b) Is the Jeffreys prior proper? Why or why not?

5. Suppose $X|p_1 \sim Bin(n, p_1)$ and $Y|p_2 \sim Bin(m, p_2)$, with X and Y independent given p_1 and p_2. Let $H_0 : p_1 = p_2$, and suppose under H_0 we assign prior distribution $p_1 \sim Unif(0, 1)$ (and $p_2 = p_1$) and under $H_1 : p_1 \neq p_2$ we assign independent priors $p_1 \sim Unif(0, 1)$ and $p_2 \sim Unif(0, 1)$.

   (a) Calculate $f(x, y|H_1)$. *Hint: Use the independence assumptions.*

   (b) Calculate $f(x, y|H_0)$.

   (c) Use (a) and (b) to compute the Bayes factor BF_{10} for comparing H_1 to H_0.

   (d) Let p_1 denote the "true" batting average for Albert Pujols and p_2 denote the "true" batting average for Ichiro Suzuki. Consider the data

   Pujols: 5146 at bats; 1717 hits

   Suzuki: 6099 at bats; 2030 hits

   treating each at bat as a Bernoulli trial with probability p_1 or p_2 of getting a hit. Calculate BF_{10} for $H_1 : p_1 \neq p_2$ to $H_0 : p_1 = p_2$. Explain the level of evidence this indicates for H_1 relative to H_0.