Stats 545: Homework 7

Due before class on Tuesday, Nov 5.
All plots should have labelled axes and titles.

Important:`Rcode`, tables and figures should be part of a single .pdf or .html files from `R Markdown` and `knitr`. See the class reading lists for a short tutorial.Any derivations can also be in Markdown, in Latex or neatly written on paper which you can give to me.

1 Problem 1: Generating gamma random variables via rejection sampling [80]

We are going to write some functions to generate random variables in `R`. The only in-built `R` function we will access to is `runif()` that generates numbers uniformly between 0 and 1. For each question, you can also use functions defined in earlier questions.

We will first generate an exponential random variable.

1. Let U be a Uniform$(0, 1)$ random variable. How will you transform this to generate an exponential with rate λ? Use the inverse-cdf method. [4 pts]

2. Write an `R` function `my_exp` that takes two inputs `n` and `lambda` and returns n independent exponentials with rate `lambda`. Sample 1000 variables for `lambda = 5` and plot the histogram/density. [4 pts]

3. Let Y be a random variable on $[0, 1)$, with density $p(Y = y) \propto 1/y^a$, where $a \in (0, 1)$. This is the truncated Pareto distribution. What is it's normalization constant? Describe the inverse-cdf method to generate Y by transforming U. [7 pts]

4. Write a function `my_pareto` that takes two inputs `n` and `a`, and returns n random variables distributed as a. Plot the histogram/density for 1000 such variables for $a = 0.5$. [8 pts]

5. Write down the density of a Gamma random variable with shape parameter α and rate parameter β. Call this $p_{\mathrm{Gamma}}(x|\alpha, \beta)$. Note that calculating its cdf is not easy. [5pts]

6. Note that the sum of two exponentials with rate 1 is distributed as Gamma$(2, 1)$. Write a function `my_gamma_int` that takes two inputs, `n` and `alpha_int`, and returns n independent Gamma variables with parameters (`alpha_int`, 1), where `alpha_int` must be an integer. [8pts]

Next, we will generate a Gamma random variable, with `alpha` < 1

7. Note that for $\alpha < 1$, $p_{\mathrm{Gamma}}(x|\alpha, 1) \leq C_1 x^{\alpha-1}$ for $x < 1$, and $p_{\mathrm{Gamma}}(x) \leq C_2 \exp(-x)$ for $x \geq 1$. What are C_1 and C_2? [5pts]

8. Use this information to write down a probability density $q(x|\alpha)$ and a scalar M such that $p_{\mathrm{Gamma}}(x|\alpha, 1) \leq M \cdot q(x|\alpha)$. [8pts]

9. Now provide pseudocode for a rejection sampler to sample from $p_{\mathrm{Gamma}}(x|\alpha, 1)$. Write down the acceptance probability. [7pts]

10. Write a function `my_gamma_l1` that takes two inputs, `n` and `alpha_l1`, and returns n independent Gamma variables with parameters `(alpha_l1,1)`, where `alpha_l1` must be less than or equal to one. [8pts]

11. How would you use this to generate a Gamma with arbitrary shape parameter, and rate equal to one. How about arbitrary shape and rate parameters? [8pts]

12. Write a function `my_gamma` to do this. Generate 1000 independent variables with parameter $(2.5, 2)$. Plot the histogram/density and compare with `R`'s `rgamma` function. [8pts]

2 Problem 2: Gibbs sampling for a mixture of Poissons [20]

In a hypothetical experiment, a neuron emits a burst of x spikes over a short interval in response a stimulus. The neuroscientist models x as Poisson-distributed, with mean λ, where λ is the voltage input to the neuron (in appropriate units). λ is unknown, but assumed constant throughout the experiment and the neuroscientist measures a sequence of counts $X = (x_1, x_2, \ldots, x_N)$.

1. What is the maximum likelihood estimate of λ? [2 pts]

Now, even though λ for this neuron is unknown, the neuroscientist knows that across all neurons in the brain, λ has mean 10 and standard-deviation 5. She decides to place a conjugate Gamma(a, b) prior on λ.

2. What are the parameters (a, b) of this conjugate prior? [2 pts]

3. What are the parameters of the posterior given observations X? [2 pts]

4. What is the MAP solution (corresponding to the mode of the posterior)? Wikipedia has an expression for this. [2 pts]

Now assume that the electrode picks up signal not just from one but K neurons. Assume $K = 5$. Let λ_k be the mean of neuron k (these are assumed independent). Each count x_i now belongs to one of the 5 neurons, let c_i be its cluster assignment. Let all neurons have the sample probability i.e. $p(c_i = k) = 1/5 = \pi_k$. The overall model is then:

$$\begin{aligned}
\lambda_k &\sim \text{Gamma}(a, b), & k &= 1, \ldots, 5 \\
c_i &\sim \pi & i &= 1, \ldots, N \\
x_i &\sim \text{Poisson}(\lambda_{c_i})
\end{aligned}$$

We observe only the x_i's with the c_i's missing. The EM-algorithm would allow us to calculate point estimates of $\vec{\lambda}$ (the MLE or even the MAP solution). However, we now want the full posterior distribution over $\vec{\lambda}$ given X. We will sample from $p(C, \vec{\lambda}|X)$ using a Gibbs sampler that alternately samples the C given $\vec{\lambda}, X$ and then $\vec{\lambda}|C, X$.

5. For any observation i, write down $P(c_i|X, \vec{\lambda}, C^{\neg i})$. [4 pts]

6. For neuron k, write down $P(\lambda_k|X, \vec{\lambda}^{\neg k}, C)$. [4 pts]

7. Write a few lines of psuedocode describing your overall Gibbs sampler. [2 pts]

8. How would you use these samples to approximate the posterior mean and variance? [2 pts]