STA 290 — Homework 6

Please include all of your C code in an appendix.

The data for this homework, as well as the background of the data, come from Rice (1988) (p. 263, although what I'm asking you to do with the data is quite different, I just want to fully acknowledge the source).

The Poisson distribution has been used by traffic engineers as a model for light traffic, based on the rationale that if the rate is approximately constant and the traffic is light (so the individual cars move independently of each other), the distribution of counts of cars in a given time interval or space area should be nearly Poisson (Gerlough and Schuhl, 1955). The file `turns.dat` (available on the course web page of datasets) shows the number of right turns during 300 3-minute intervals at a specific intersection. We will fit two different models to this data. Note that the data are presented as summary data, for example, the first line indicates that in 14 of the intervals there were no turns observed. Thus the 300 cases are not listed separately. You don't need them to be, but you need to be aware of the format of the data when you are doing your analysis, since the standard way of writing distributions is in terms of the individual cases.

Let y_i be the number of right turns observed in time interval i, $i \in \{1, \ldots, 300\}$. We model each y_i as *iid* Poisson with rate parameter λ (i.e., $E[y_i] = \lambda$). What prior we put on λ varies with the two parts below.

1. Put a prior on λ which is the (re-scaled) positive part of the standard normal distribution, $N^+(0,1)$, i.e., $P(\lambda) = \sqrt{\frac{2}{\pi}} e^{-\lambda^2/2} I_{\{\lambda>0\}}$.

   (a) Use Metropolis-Hastings (with your choice of proposal, symmetric or not) to generate a sample from the posterior for λ. Plot the MCMC run.

   (b) What does this posterior distribution look like? Remove a sufficient number of burn-in runs. Create a histogram. Make Q–Q plots against a normal distribution and an appropriate gamma distribution. Comment on your findings.

2. Use a hierarchical prior for λ, with first an exponential prior with rate θ (i.e., mean $1/\theta$), and then an exponential prior with rate 1 for θ, i.e.

$$
\begin{aligned}
y_i|\lambda &\overset{iid}{\sim} Pois(\lambda) \\
\lambda|\theta &\sim Exp(\theta) \\
\theta &\sim Exp(1)
\end{aligned}
$$

   (a) Find the complete conditional distributions for λ and θ and recognize them as members of a standard family (specifying the parameters of this family).

   (b) Write a Gibbs sampler (for both λ and θ) and generate a sample from the joint posterior. Plot the MCMC run.

   (c) Remove the burn-in runs, then make a Q–Q plot against your posterior sample for λ from the other prior (if you have removed a different number of burn-in runs, make the two runs the same length by removing more samples from the beginning of the longer run). Comment on your findings.