Module 11: Lesson 2 Lab

Callum Arnold

7/19/2021

Contents

Background

Suppose that we have data on incubation periods $y = (y_1, ..., y_n)$. We assume that the data are *independent* draws from a Gamma distribution with shape α and rate β, i.e.

$$y_i \sim \Gamma(\alpha, \beta) \quad \text{i.i.d.}, i = 1, ..., n$$

where the Gamma distribution has probability density function:

$$f(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), \quad \alpha > 0, \beta > 0.$$

The goal is to make (sampling-based) Bayesian inference for the parameters and .

Exercise 1

> Assume that and are *a priori* independent and assign Gamma distributions with parameters $\lambda\alpha$ and $\nu\alpha$, and $\lambda\beta$ and $\nu\beta$, respectively. Write down (i.e. express in mathematical terms) the posterior density of interest, i.e.
>
> $$\pi(\alpha, \beta|y) \propto \pi(y|\alpha, \beta)\pi(\alpha)\pi(\beta)$$
>
> where $\pi(y|\alpha, \beta)$ denotes the likelihood function.

$$
\begin{aligned}
y_i &\sim \Gamma(\alpha, \beta) \\
f(x|\alpha,\beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\exp(-\beta x) \\
\alpha &\sim \Gamma(\lambda_\alpha, \nu_\alpha) \\
\beta &\sim \Gamma(\lambda_\beta, \nu_\beta)
\end{aligned}
$$

$$
\begin{aligned}
\pi(\alpha,\beta|y) &\propto \pi(y|\alpha,\beta)\pi(\alpha)\pi(\beta) \\
\pi(y|\alpha,\beta) &= \frac{\beta^{n\alpha}\prod_k y_k^{\alpha-1}\exp\left(-\beta\sum y_k\right)}{\{\Gamma(\alpha)\}^n} \\
\pi(\alpha,\beta|y) &\propto \frac{\beta^{n\alpha}\prod_k y_k^{\alpha-1}\exp\left(-\beta\sum y_k\right)}{\{\Gamma(\alpha)\}^n} \times \pi(\alpha)\pi(\beta) \\
&\propto \frac{\beta^{n\alpha}\prod_k y_k^{\alpha-1}\exp\left(-\beta\sum y_k\right)}{\{\Gamma(\alpha)\}^n} \times \left(\frac{\nu_\alpha^{\lambda_\alpha}}{\Gamma(\lambda_\alpha)}\alpha^{\lambda_\alpha-1}\exp(-\nu_\alpha\alpha)\right) \times \left(\frac{\nu_\beta^{\lambda_\beta}}{\Gamma(\lambda_\beta)}\beta^{\lambda_\beta-1}\exp(-\nu_\beta\beta)\right) \\
&\propto \left(\frac{\beta^{n\alpha}\prod_k y_k^{\alpha-1}\exp\left(-\beta\sum y_k\right)}{\{\Gamma(\alpha)\}^n}\right) \times \left(\alpha^{\lambda_\alpha-1}\exp(-\nu_\alpha\alpha)\right) \times \left(\beta^{\lambda_\beta-1}\exp(-\nu_\beta\beta)\right)
\end{aligned}
$$

Exercise 2

Having obtained the joint posterior density in Exercise 1, first derive the densities of the posterior distribution of the parameters   and   up to proportionality, i.e. (| ,y) and (| ,y).

Is any of these densities of a known form (i.e. the density of a standard/well-known distribution, e.g. Gamma, Normal etc)?

Full conditional density of α

$$
\pi(\alpha|\beta,y) \propto \left(\frac{\beta^{n\alpha}\prod_k y_k^{\alpha-1}}{\{\Gamma(\alpha)\}^n}\right) \times \left(\alpha^{\lambda_\alpha-1}\exp(-\nu_\alpha\alpha)\right)
$$

This doesn't look like a standard well-known distribution, so we'll need to use M-H algorithm

Full conditional density of β

$$
\begin{aligned}
\pi(\alpha,\beta|y) &\propto \left(\beta^{n\alpha}\prod_k \exp\left(-\beta\sum y_k\right)\right) \times \left(\beta^{\lambda_\beta-1}\exp(-\nu_\beta\beta)\right) \\
&\propto \left(\beta^{n\alpha}\exp\left(-\beta\sum_{k=1}^n y_k\right)\right) \times \left(\beta^{\lambda_\beta-1}\exp(-\nu_\beta\beta)\right) \\
&\propto \beta^{n\alpha+\lambda_\beta-1}\exp\left[\beta\left(\sum_{k=1}^n y_k + \nu_\beta\right)\right] \\
\beta|\alpha,y &\sim \Gamma\left(n\alpha+\lambda_\beta, \sum_{k=1}^n y_k + \nu_\beta\right)
\end{aligned}
$$

We can see that this looks like a Gamma distribution with the associated shape parameters, therefore, for β we can sample directly from its conditional distribution.

Exercise 3

Write a R function to implement an MCMC algorithm (from scratch!) which samples from the joint posterior distribution $\pi(\alpha, \beta|y)$.

First, think about what the input (i.e. arguments) the function should have, e.g. a vector corresponding to data **y**, hyper-parameter values $\lambda\alpha$, etc. Then also think about what the output should be; surely the function should output the posterior samples of α and β. Anything else?

To get you started, a pseudo-code of this MCMC algorithm is given below:

1. Choose initial values for α and β;
2. Update α by sampling from $\pi(\alpha|\beta, y)$ using a Gaussian random-walk;
3. Update β by sampling from $\pi(\beta|\alpha, y)$ directly.
4. Go to to Step (b)

Look at the trace plots and convince yourself that the chain has reached stationarity and is mixing well.

Exercise 4

Test your algorithm is producing sensible results.

One way to do this is to first simulate a dataset from a Gamma distribution with some specific values for α and β, eg. `y <- rgamma(100, 4, 2)` wil simulate a vector of 100 draws from a Ga(4,2).

Then use your MCMC algorithm (in Exercise 3) to sample from the joint posterior distribution of and assuming vague priors, e.g. $\lambda\alpha = \lambda\beta = 1$ and $\nu\alpha = \nu\beta = 10 - 3$.

In principle, your marginal posterior distributions ($\pi(\alpha|y)$ and $\pi(\beta|y)$) should be centered around the values you have chosen to simulate the data from (e.g. 4 and 2 in this case).

Exercise 5

If you have convinced yourself that your MCMC algorithm is doing what is supposed to be doing, then use it to fit the model to the Campylobacter data (24 observations in total) from Evans et al. 1996) by `y <- c(rep(2,2), rep(3, 6), rep(4, 11), rep(5, 3), rep(7,7))`

Does the model offer a good fit? How would you go assessing that?

Exercise 6* [Optional]

Now that you have got a working algorithm, can you improve its mixing?

One way to do this is to update both α and β at the same time (in a block). This can be done, for example, and as discussed in the lecture, by proposing values drawn from some (bivariate distribution) with density $q(\alpha, \beta)$ and accept/reject according to the Metropolis-Hastings ratio. There are many choices here, but not all of them will lead to an efficient sampler.

Here are a few choices to try:

- Gaussian random-walk in a block, i.e. propose values for α and β drawn from a bivariate Normal distribution with mean the current values of α and β and some (2×2) variance-covariance matrix, and then accept/reject using the Metropolis-Hastings ratio.

- Gaussian approximation to the posterior distribution $\pi(\alpha, \beta|y)$. (Hint: use the command `optim` to find the values of α and β that maximise the posterior density).