Bayesian statistics and pomp

Qianying (Ruby) Lin Spencer J. Fox

Lecture outline

1. Motivating Bayesian statistics
2. Short introduction to Bayesian statistics and theory
3. Introduction to MCMC
4. Introduction to PMCMC
5. Simple influenza case study

What we have covered so far

$$p(y|\theta)$$

- y can be thought of as your data or observations
- θ can be thought of as the model or parameter values
- Called the "Likelihood"

Issues with maximum likelihood estimation (MLE)

- Assumes results occur with some given "frequency" over period time or replicates/repeated experiments
  - If we had the same outbreak hundreds of time, what proportion of them would provide confidence intervals that contain the true value for the R_0
- Some difficulties in constraining parameter values based on outside data, information, or expert opinion
- Just not really intuitive…
  - We typically want to say something about the parameters based on the data, $p(\theta|y)$

Bayesian statistics

- Bayes theorem provides an intuitive framework to update parameter estimates based on both prior knowledge and experimental data
- End result is a posterior distribution, $p(\theta|y)$, directly describing the parameter and model of interest
- Easy to communicate results
  - "The reproduction number is estimated to be x, with a 95% credible interval from y to z"
- Issues
  - Computationally expensive
  - Without enough data, prior can bias posterior distribution, but this is what you want!

Bayes theorem 1

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

- $p(\theta|y)$ is the posterior distribution
- $p(y|\theta)$ is the likelihood
- $p(\theta)$ is the prior distribution
- $p(y)$ is the marginal distribution (sometimes called a normalizing constant as it doesn't depend on the parameters)

Bayes theorem 2

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

- $p(y) = \int p(y|\theta)p(\theta)d\theta$
- Probability of observing y marginal over all possible values of θ
- Typically is very difficult to calculate
- The good news is that $p(y)$ is a constant

Bayes theorem 3

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

- Since $p(y)$ is a constant, the posterior distribution is proportional to the likelihood times the prior
- If we can solve this we can get the posterior distribution because $\int p(\theta|y)d\theta = 1$
- Intuitively our parameter estimates are based on a combination of our observations $p(y|\theta)$ and our prior beliefs $p(\theta)$
- Only need to sample from the likelihood and prior distribution to get the posterior

How do we do so?

- Many ways to do so (and many software packages), but we're only going to talk about one...
- Markov chain Monte Carlo (MCMC) is a class of algorithms used to draw samples from a probability distribution
- Will not cover the theoretical details, but will attempt to motivate

Assume you have an unknown probability distribution (hill) to explore... How would you do so?

- You don't know where it is in parameter space
- You don't know it's shape