ST 740: Markov Chain Monte Carlo

Alyson Wilson

Department of Statistics
North Carolina State University

October 16, 2013

Flour Beetle Mortality

Data from Bliss (1935), Model from Carlin and Louis (2000)

Dosage w_i	Number Killed y_i	Number Exposed n_i
1.6907	6	59
1.7242	13	60
1.7552	18	62
1.7842	28	56
1.8113	52	63
1.8369	53	59
1.8610	61	62
1.8839	60	60

Sampling Distribution

The number of insects killed y_i after five hours of exposure to dose w_i of gaseous carbon disulphide is Binomial(n_i, p_i), where the probability of death is modeled as

$$p_i = \left(\frac{\exp(x_i)}{1 + \exp(x_i)} \right)^{m_1}$$

where

$$x_i = \frac{w_i - \mu}{\sigma}$$

This is called a *generalized logit model*.

The unknown parameters in the model are μ, σ, and m_1.

Prior Distribution

Assume the prior distributions for μ, σ, and m_1 are independent. We will use improper priors for μ and σ, with

$$\begin{aligned} \pi(\mu) &\propto 1, \ -\infty < \mu < \infty \\ \pi(\sigma) &\propto \frac{1}{\sigma}, \ 0 < \sigma < \infty \end{aligned}$$

We assign a gamma prior to m_1 with parameters $a_0 = 0.25$ and $b_0 = 0.25$.

$$\begin{aligned} \pi(m_1) &\propto \exp(-b_0 m_1) m_1^{a_0 - 1} \\ &\propto \exp(-0.25 m_1) m_1^{0.25-1} \end{aligned}$$

The joint prior distribution is

$$\pi(\mu, \sigma, m_1) \propto \frac{\exp(-0.25 m_1) m_1^{-0.75}}{\sigma}$$

Posterior Distribution

The posterior distribution is

$$\pi(\mu, \sigma, m_1 \mid y, n, w) \propto \left[\prod_{i=1}^{8} p_i^{y_i} (1 - p_i)^{n_i - y_i} \right] \frac{\exp(-0.25 m_1) m_1^{-0.75}}{\sigma}$$

where

$$p_i = \left(\frac{\exp(x_i)}{1 + \exp(x_i)} \right)^{m_1}$$

and

$$x_i = \frac{w_i - \mu}{\sigma}$$

Questions of Interest

What questions are we interested in answering about this data?

- We would like to plot the "dose-response" curve, or a curve that relates the dose of gaseous carbon disulphide to the probability of death for a flour beetle.
- We would also like to estimate "LD50," or the dose of gaseous carbon disulphide that kills 50% of the flour beetles.

Computational Considerations

As a general computational strategy, we will compute using the log posterior after transforming all parameters to the real line. In particular, let $\theta_2 = \log(\sigma)$ and $\theta_3 = \log(m_1)$.

The joint posterior distribution is not a common distributional form, and neither are the full conditional distributions. We'll use the Metropolis-Hastings algorithm directly on the posterior distribution to get a random sample from the posterior that we will then use to address our questions. Since we are using the Metropolis algorithm, we only need to know our posterior distribution up to a normalizing constant.

Computational Considerations

$$\log(\pi(\mu, \sigma, m_1 | y, n, w)) = \log\left(\left[\prod_{i=1}^{8} p_i^{y_i}(1 - p_i)^{n_i - y_i}\right] \frac{\exp(-b_0 m_1) m_1^{a_0 - 1}}{\sigma}\right)$$

$$= \left[\sum_{i=1}^{8} y_i \log(p_i) + (n_i - y_i)\log(1 - p_i)\right] - b_0 m_1 + (a_0 - 1)\log(m_1) - \log(\sigma)$$

$$\log(\pi(\mu, \theta_2, \theta_3 | y, n, w)) = \left[\sum_{i=1}^{8} y_i \log(p_i) + (n_i - y_i)\log(1 - p_i)\right] - b_0 \exp(\theta_3) + a_0\theta_3$$

Normal Approximation to Posterior

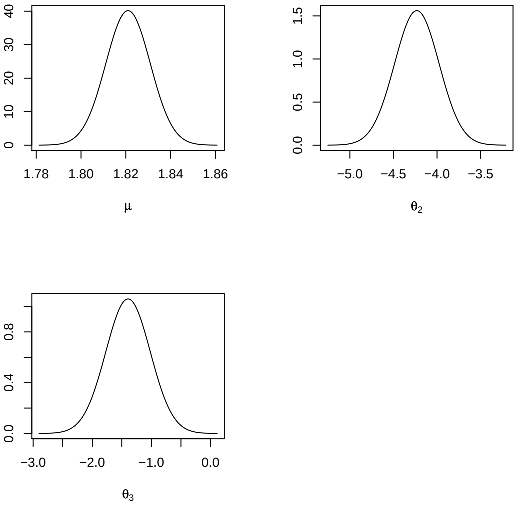
We find that a normal approximation to the posterior has mean $(1.82, -4.23, -1.39)$ and variance-covariance matrix

$$\begin{pmatrix} 0.000099 & -0.0020 & -0.0034 \\ -0.0020 & 0.065 & 0.090 \\ -0.0034 & 0.090 & 0.14 \end{pmatrix}$$

The correlations among the parameters based on the normal approximation are $\text{Corr}(\mu, \theta_2) = -0.77$, $\text{Corr}(\mu, \theta_3) = -0.91$, and $\text{Corr}(\theta_2, \theta_3) = 0.93$. The large correlations suggest that we may need to take correlation into account when we set up our proposal density.

Normal Approximation to Posterior

Approximate marginal posterior distributions



Proposal Density

We'll use a multivariate normal proposal density. Our mean will be at the previous draw. Our proposal variance matrix will use the variances from our normal approximation.

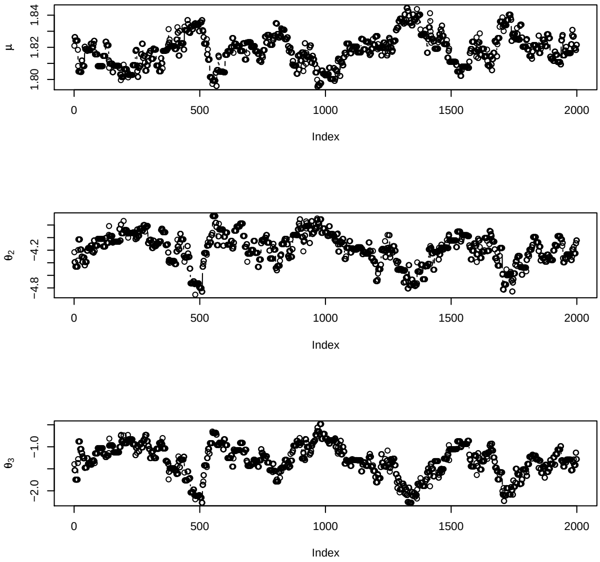
$$\begin{pmatrix} 0.000099 & 0.0 & 0.0 \\ 0.0 & 0.065 & 0.0 \\ 0.0 & 0.0 & 0.14 \end{pmatrix}$$

We'll run 2000 observations from our Metropolis algorithm starting at the posterior mode.

Proposal Density

When we run 2000 observations with this proposal density, the acceptance ratio is about 10%. This is too low, so we want to “decrease” our proposal variance. Through some experimentation, we decide to multiply our current proposal variance matrix by 0.4. This gives us an acceptance ratio of about 23%

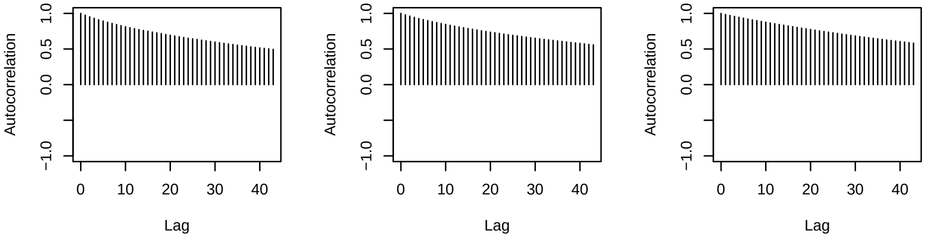
Trace Plots


μ
1.80
1.82
1.84
0
500
1000
1500
2000
Index
θ2
−4.8
−4.2
0
500
1000
1500
2000
Index
θ3
−2.0
−1.0
0
500
1000
1500
2000
Index

Trace Plots

There's a problem with these trace plots. The chains are mixing poorly. Notice that they seem to be "wandering" through the space.

Poor Mixing



How do you know this isn't going to wander off into another part of the sample space after you run longer? You are trying to have good coverage of the entire support of the posterior distribution.

New Proposal Density

Now we'll run 2000 observations with our proposal density variance set to the normal approximation variance-covariance matrix

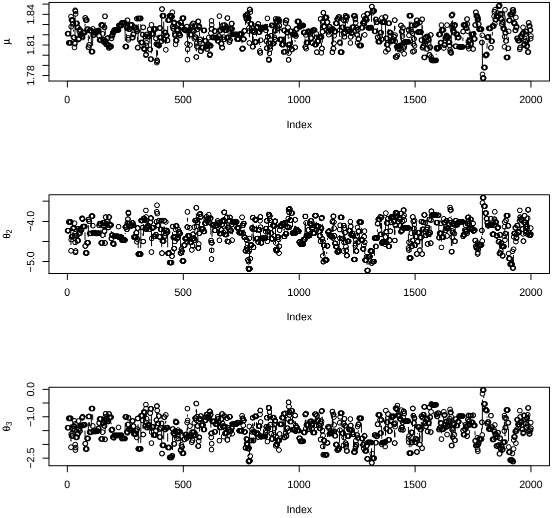
$$\begin{pmatrix} 0.000099 & -0.0020 & -0.0034 \\ -0.0020 & 0.065 & 0.090 \\ -0.0034 & 0.090 & 0.14 \end{pmatrix}$$

This has an acceptance ratio of about 45%.

This is too high, so we want to "increase" our proposal variance. Through some experimentation, we decide to multiply our current proposal variance matrix by 2.0. This gives us an acceptance ratio of about 30%.

Trace Plots

These chains are mixing much better.



How Many Iterations

Using our chain of length 2000, we run the Raftery diagnostic. It tells us that we need at least 3746 runs with the default parameter settings. So we run our chain 5000 iterations and find that the suggested number of iterations are 57182 (μ), 45588 (θ_2), and 30,348 (θ_3). We need to go with the largest number, so we'll run 60,000 observations.

The maximum number of burn-ins suggested is 53. We'll come back to this once we've run multiple chains.

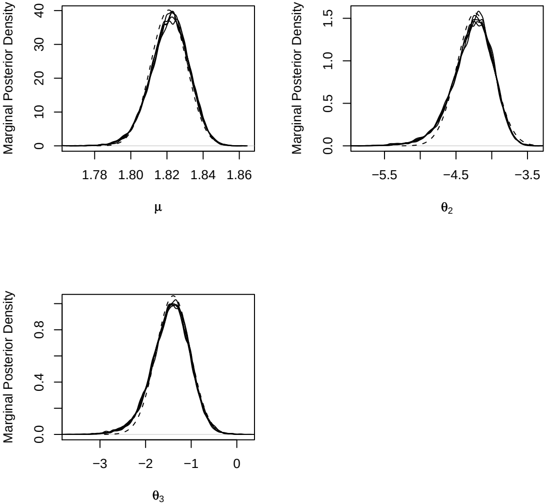
Multiple Chains

Now, we'll run 5 chains of 60,000 runs each starting from values simulated from a multivariate t distribution with 4 degrees of freedom and the mean and variance structure suggested by our normal approximation.

Looking at the 5 chains, it is hard to choose a number of burn-ins. We'll go with 100 (just to be safe).

Marginal Posterior Distributions

We plot the marginal posterior distributions for the three parameters and overlay the marginal approximated using the normal.



Questions of Interest

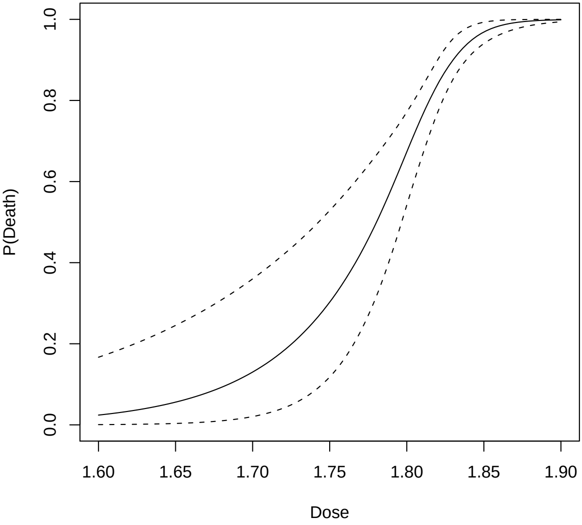
Now that we have our posterior draws, let's remember what questions we interested in answering about this data.

- We would like to plot the "dose-response" curve, or a curve that relates the dose of gaseous carbon disulphide to the probability of death for a flour beetle.
- We would also like to estimate "LD50," or the dose of gaseous carbon disulphide that kills 50% of the flour beetles.

Dose-Response Curve

We want to plot the posterior distribution of p at many values of possible dose. We'll choose a range of 1.60 to 1.90 for the doses, grid 100 points, use our 299,500 posterior draws at each of the 100 doses to calculate p (which is a function of our parameters), and plot the posterior mean, 5th quantile, and 95th quantile of p at each dose.

Dose-Response Curve with 90% Credible Interval



LD50

We want to calculate the dose w_{50} so that the probability of response is 0.5.

$$0.5 = \left(\frac{\exp(x)}{1+\exp(x)}\right)^{m_1}$$

After some algebra, we have

$$w_{50} = \mu + \sigma \log\left(\frac{.5^{1/m_1}}{1-0.5^{1/m_1}}\right)$$

Again, w_{50} is a function of our three parameters, so we calculate posterior draws of w_{50} using the posterior samples from the three parameters.

We find the central 95% credible interval for LD50 is $(1.74, 1.80)$.

Posterior Distribution of LD50