Metropolis-Hastings Example

Example 5 (Sparrow data): We gather data on a sample of 52 sparrows:

$$X_i = \text{age of sparrow (to nearest year)}$$
$$Y_i = \text{Number of offspring that season}$$

- We expect that the offspring number rises and then falls with age, so we assume a quadratic trend.
- We model the number of offspring at a given age x as Poisson:

$$Y|x \sim \text{Pois}(\mu_x)$$

Metropolis-Hastings Example

- Since we know μ_x must be positive, we use the model:

$$E[Y|x] = e^{\beta_0 + \beta_1 x + \beta_2 x^2}$$

- This Poisson regression model is a **generalized linear model** (GLM).
- Our parameter of interest is $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$.
- But note that conjugate priors do not exist for **non-normal** GLMs.
- We will use the M-H algorithm to sample from our posterior.

Metropolis-Hastings Example

- Let the prior on $\boldsymbol{\beta}$ be multivariate normal with **independent** components:

$$\boldsymbol{\beta} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}), \text{ where } \boldsymbol{\Sigma} = 100 \times \mathbf{I}_3$$

- We will choose our **proposal** density to be multivariate normal with mean vector $\boldsymbol{\beta}^{[t]}$ (the current value).
- The covariance matrix of the proposal density is sort of a tuning parameter. We will choose

$$\hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1} \text{ where } \hat{\sigma}^2 = \text{var}\{\ln(y_1 + 0.5), \ldots, \ln(y_n + 0.5)\}.$$

- We can adjust this if our **acceptance rate** is too high or too low.
- Usually we like an acceptance rate between 20% and 50%.

Metropolis-Hastings Example

- Since our proposal density is symmetric, our acceptance ratio is simply

$$\frac{\pi(\boldsymbol{\beta}^*)}{\pi(\boldsymbol{\beta}^{[t]})} = \frac{L(\boldsymbol{\beta}^*|\mathbf{X},\mathbf{y})p(\boldsymbol{\beta}^*)}{L(\boldsymbol{\beta}^{[t]}|\mathbf{X},\mathbf{y})p(\boldsymbol{\beta}^{[t]})}$$

$$= \frac{\prod_{i=1}^{n} \text{dpois}(y_i, \exp[\mathbf{x}_i^T\boldsymbol{\beta}^*]) \prod_{j=1}^{3} \text{dnorm}(\beta_j^*, 0, 10)}{\prod_{i=1}^{n} \text{dpois}(y_i, \exp[\mathbf{x}_i^T\boldsymbol{\beta}^{[t]}]) \prod_{j=1}^{3} \text{dnorm}(\beta_j^{[t]}, 0, 10)}$$

  where the Poisson density `dpois` and the normal density `dnorm` can be found easily in `R`.
- See `R` example with real sparrow data.

Other Metropolis-Hastings Issues

- In practice, it is recommended to check the acceptance rate (the proportion of proposed $\boldsymbol{\beta}^*$ values that are "accepted").
- We also check the serial correlation of the $\left\{\beta_j^{[t]}\right\}$ values using a plot of the **autocorrelation function**.
- If the values do not "appear" independent, we can alleviate this by choosing every k th value in the chain as our posterior sample (**thinning**).

CHAPTER 6(b) SLIDES START HERE

Checking Model Adequacy

- Checking the adequacy of a Bayesian model involves:
  1. determining how sensitive the posterior is to the specification of the prior and the likelihood
  2. checking that the values we obtain in our sample fit those we would expect to see, given our posterior knowledge
  3. checking robustness to individual data values

Sensitivity Analysis

- Checking the sensitivity to the specification of the data model/likelihood should be done regularly, but rarely is.
- We might examine the effect on the posterior of choosing related data models (e.g., Poisson vs. negative binomial for count data).
- Far more often, we check the sensitivity of the posterior to the **prior** specification.
- Assume Poisson(θ_1) and Poisson(θ_2) models for the data.
- We might ask: What happens to the posterior when we:
  1. change the functional form of the prior?
  2. keep the same form, but change the parameter(s) of the prior?
- If the posterior is **robust** to such changes in the prior, we may be more comfortable with the posterior inferences we make.

Sensitivity Analysis

Example 1(a): Consider $X_1, \ldots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$ with σ^2 known.

- The conjugate prior for μ is $\mu \sim N(\delta, \tau^2)$.
- A noninformative prior for μ is $p(\mu) = 1$.
- Another choice of prior for μ might be a t-distribution centered at δ.
- How would the posterior change for these 3 prior choices?
- We could examine (1) plots of the posterior in each case, or (2) several posterior quantiles in each case.
- See `WinBUGS` example with Kenya lead data.