Another Gibbs Example (Normal Mixture)

**Example 4** (Monkey Eye Data): $X_1, \ldots, X_{48}$ are a random sample of peak sensitivity wavelength measurements from a monkey’s eyes (Bowmaker et al., 1985)

- The data are assumed to come from a mixture of two normal distributions, i.e.,

  $$X_i \overset{\text{indep}}{\sim} N(\lambda_{T_i}, \tau) \text{ and } T_i \sim \text{Bernoulli}(p)$$

  where $T_i (= 1 \text{ or } 2)$ indicates the true group the $i$th observation came from.

- $\lambda_1 =$ mean of group 1, $\lambda_2 =$ mean of group 2, $\tau =$ common **precision** parameter (reciprocal of variance)

- For computational reasons, we let $\lambda_1 < \lambda_2$ and define the “mean shift” $\theta = \lambda_2 - \lambda_1$, $\theta > 0$. 
We use the following independent noninformative priors on $\lambda_1$, $\theta$, $\tau$, and $p$:

- $p \sim \text{beta}(1, 1)$
- $\theta \sim \mathcal{N}(0, \tau = 10^{-6})I_{[\theta > 0]}$ ($\Rightarrow \sigma^2 = 10^6$)
- $\lambda_1 \sim \mathcal{N}(0, \tau = 10^{-6})$
- $\tau \sim \text{gamma}(0.001, 0.001)$

Do example in WinBUGS with 1000-draw burn-in and then 10000 further draws.

See convergence diagnostics in WinBUGS.
When the full conditionals for each parameter cannot be obtained easily, another option for sampling from the posterior is the Metropolis-Hastings (M-H) algorithm.

The M-H algorithm also produces a **Markov chain** whose values approximate a sample from the posterior distribution.

For this algorithm, we need the form (except for a normalizing constant) of the posterior $\pi(\cdot)$ for $\theta$, the parameter(s) of interest.

We also need a **proposal** (or **instrumental**) distribution $q(\cdot|\cdot)$ that is easy to sample from.
The M-H algorithm first specifies an initial value for $\theta$, say $\theta^{[0]}$. Then:

- After iteration $t$, suppose the most recently drawn value is $\theta^{[t]}$.
- Then sample a candidate value $\theta^*$ from the proposal density.
- Let the $(t + 1)$-st value in the chain be

$$
\theta^{[t+1]} = \begin{cases} 
\theta^* & \text{with probability } \min\{a(\theta^*, \theta^{[t]}), 1\} \\
\theta^{[t]} & \text{with probability } 1 - \min\{a(\theta^*, \theta^{[t]}), 1\}
\end{cases}
$$

where

$$
a(\theta^*, \theta^{[t]}) = \frac{\pi(\theta^*)}{\pi(\theta^{[t]})} \frac{q(\theta^{[t]}|\theta^*)}{q(\theta^*|\theta^{[t]})}
$$

is the “acceptance ratio.”
In practice we would accomplish this by sampling $U[t] \sim U(0, 1)$ and choosing $\theta^{[t+1]} = \theta^*$ if $a(\theta^*, \theta^{[t]}) > u[t]$; otherwise choose $\theta^{[t+1]} = \theta^{[t]}$.

Note that if the proposal density $q(\cdot | \cdot)$ is symmetric such that $q(\theta^{[t]} | \theta^*) = q(\theta^* | \theta^{[t]})$, then the acceptance ratio is simply

$$\frac{\pi(\theta^*)}{\pi(\theta^{[t]})}.$$