Exercise 1: Let the observed data $\{y_i\}_{i=1}^n$ be binary with each y_i being an independent realization of a Bernoulli random variable Y_i with parameter $\mu_i = \Pr(Y_i = 1)$. Each $\mu_i = \mathbb{E}[Y_i]$ is modeled with a logistic function, that is,

$$\mu_i = \frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}} \quad \Leftrightarrow \quad \log\left\{\frac{\mu_i}{1-\mu_i}\right\} = \beta_0 + \beta_1 x_i\,,$$

where $\boldsymbol{x} = [x_1, \ldots, x_n]^{\mathrm{T}}$ is a vector of predictor values and β_0 and β_1 are parameters.

1. Write down the likelihood $p(\boldsymbol{y} \,|\, \beta_0, \beta_1)$ for the logistic regression model.
2. Find the posterior $p(\beta_0, \beta_1 \,|\, \boldsymbol{y})$ assuming independent Normal priors with $\mu_{\beta_j} = 0$ and variance $\sigma^2_{\beta_j}$, where $j = 1, 2$.
3. Demonstrate that marginalizing $\boldsymbol{u}$ out of the joint posterior

$$p(\boldsymbol{u}, \beta_0, \beta_1 \,|\, \boldsymbol{y}) \propto \prod_{i=1}^n 1\left(u_i < \frac{e^{\beta_0 y_i + \beta_1 x_i y_i}}{1+e^{\beta_0+\beta_1 x_i}}\right) \exp\left\{-\frac{(\beta_0-\mu_{\beta_0})^2}{2\sigma^2_{\beta_0}} - \frac{(\beta_1-\mu_{\beta_1})^2}{2\sigma^2_{\beta_1}}\right\}$$

   yields the posterior $p(\beta_0, \beta_1 \,|\, \boldsymbol{y})$.
4. Implement slice sampling by generating

$$\begin{aligned} u_i \,|\, \boldsymbol{u}_{-i}, \beta_0, \beta_1, \boldsymbol{y} &\sim \mathrm{Unif}\left(0, \frac{e^{\beta_0 y_i + \beta_1 x_i y_i}}{1+e^{\beta_0+\beta_1 x_i}}\right) \\ \beta_0 \,|\, \boldsymbol{u}, \beta_1, \boldsymbol{y} &\sim \mathrm{Normal}(\beta_0 \,|\, \mu_{\beta_0}, \sigma_{\beta_0}) \prod_{i=1}^n 1\left(u_i < \frac{e^{\beta_0 y_i + \beta_1 x_i y_i}}{1+e^{\beta_0+\beta_1 x_i}}\right) \\ \beta_1 \,|\, \boldsymbol{u}, \beta_0, \boldsymbol{y} &\sim \mathrm{Normal}(\beta_1 \,|\, \mu_{\beta_1}, \sigma_{\beta_1}) \prod_{i=1}^n 1\left(u_i < \frac{e^{\beta_0 y_i + \beta_1 x_i y_i}}{1+e^{\beta_0+\beta_1 x_i}}\right) \end{aligned}$$

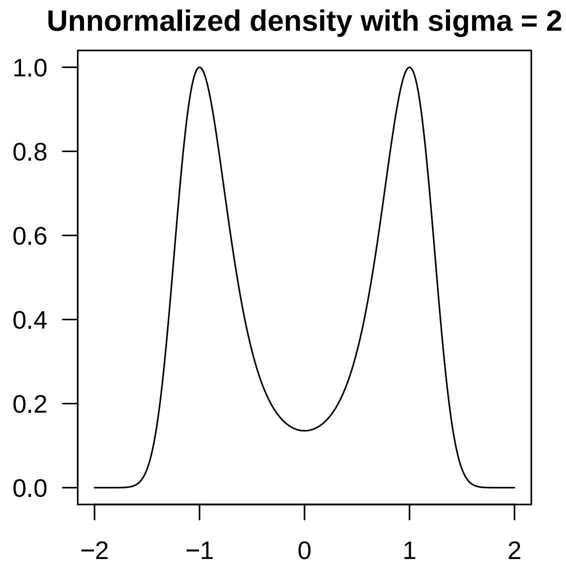
   Note that the full conditionals $p(\beta_0 \,|\, \boldsymbol{u}, \beta_1, \boldsymbol{y})$ and $p(\beta_1 \,|\, \boldsymbol{u}, \beta_0, \boldsymbol{y})$ are truncated Normal distributions.
5. Apply the slice sampler to the following simulated data:

```
n <- 100
beta0 <- 2 ; beta1 <- 0.5
set.seed( 100 ) ; x <- abs( rnorm( n ) )
eta <- beta0 + beta1 * x ; mu <- exp( eta ) / ( 1 + exp( eta ) )
y <- rbinom( n, 1, mu )
```

Exercise 2: Let X be a random variable with unnormalized density

$$p(x \,|\, \sigma) \propto \exp\left\{-\sigma\left(x^2-1\right)^2\right\}.$$

The density of X is bimodal as can be seen from the following figure:



1. Implement a random walk Metropolis–Hastings sampler based on $\text{Normal}(0, \sigma^2)$ noise.
2. Implement a random walk Metropolis–Hastings sampler that runs 4 chains in parallel with different values of σ. That is, the first Markov Chain has the target $p_1(x \mid \sigma_1)$, the second Markov chain has the target $p_2(x \mid \sigma_2)$ and so forth. In each iteration, accept a swap between the states x_i and x_j of two randomly chosen chains i and j with probability

$$\alpha = \min\left\{1, \frac{p_i(x_j \mid \sigma_i) p_j(x_i \mid \sigma_j)}{p_i(x_i \mid \sigma_i) p_j(x_j \mid \sigma_j)}\right\}.$$

3. Run both methods for $\sigma = 1, 2, 4, 8$ and inspect the histograms for $\sigma = 8$ to see if both methods can approximate the bivariate distribution.