Exercise 1: (chapter 6.2) Let $\{Y_i\}_{i=1}^n$ be independent and identically distributed random variables that follow a Bernoulli distribution with parameter $0 \le \theta \le 1$. The probability mass function of the Bernoulli distribution is

$$p_Y(y) = \Pr(Y = y) = \theta^y (1-\theta)^{1-y} \qquad y \in \{0, 1\}\,.$$

Let the prior on θ be improper with density $p(\theta) \propto \theta^{-1}(1-\theta)^{-1}$.

1. Find the posterior $p(\theta \mid y)$ and the corresponding normal approximation at its mode.
2. Show that the improper prior on θ is equivalent to a uniform prior on the logit $\beta = \log\{\theta/(1-\theta)\}$.
3. Find the posterior $p(\beta \mid y)$ and the corresponding normal approximation at its mode.
4. Is it more sensible to derive a normal approximation on the probability or logit scale?

Exercise 2 (chapter 6.2): Let $\{Y_i\}_{i=1}^n$ be independent and identically distributed random variables that follow a Poisson distribution with rate parameter $\lambda > 0$. The probability mass function of the Poisson distribution is

$$p_Y(y) = \Pr(Y = y) = \frac{\lambda^y}{y!} \exp\{-\lambda\} \qquad y = 0, 1, \ldots$$

Assume that $\mathbb{E}[\lambda] = 2$ and $\Pr(\lambda > 3) = 0.01$.

1. Describe the prior on λ by a normal distribution and find the posterior $p(\lambda \mid y)$.
2. Derive a normal approximation to the posterior $p(\lambda \mid y)$ at its mode using 100 Poisson observations

y_i	0	1	2	3	4	5	≥ 6
#	18	32	27	15	6	2	0

   and compute the posterior probability $\Pr(\lambda > 2 \mid y)$.

3. Although $\lambda > 0$, the support of the normal prior on λ is unconstrained. Which reparameterization under the bijection $\theta = h(\lambda) \Leftrightarrow \lambda = g(\theta)$ would yield an unconstrained parameter? Describe the prior on θ by a normal distribution using $\mathbb{E}[\theta] = \log 2$ and $\Pr(\theta > \log 3) = 0.01$ and find the posterior $p(\theta \mid y)$.
4. Derive a normal approximation to the posterior $p(\theta \mid y)$ at its mode using same data as above and compute the posterior probability $\Pr(\lambda > 2 \mid y)$ by translating back to the original parameter space (you may use R to find the mode and observed Fisher information).

Exercise 3 (chapter 6.2): Let $\{Y_i\}_{i=1}^n$ be independent and identically distributed random variables that follow an Exponential distribution with rate parameter $\lambda > 0$. The density of the Exponential distribution is

$$f_Y(y) = \lambda \exp\{-\lambda y\} \qquad y > 0\,.$$

Assume that the prior on λ can be described by the following density

$$p(\lambda) \propto \exp\Big\{-20(\lambda - 0.25)^2\Big\} \qquad \lambda > 0\,.$$

1. Find the posterior $p(\lambda \,|\, y)$ and an expression for the normalizing constant.
2. Derive a normal approximation to the posterior at its mode using $n = 10$ and $\bar{y} = 0.5$. Plot the normal approximation together with the true posterior density.

Exercise 4: Let $\{Y_i\}_{i=1}^n$ be independent and identically distributed random variables that follow an Normal distribution with location μ and precision parameter $\tau > 0$. The density of the Normal distribution with precision parameter τ is

$$f_Y(y) = \sqrt{\frac{\tau}{2\pi}} \exp\Big\{-\frac{\tau}{2}(y - \mu)^2\Big\}\,.$$

Assume that $\mu \,|\, \tau \sim \text{Normal}\big(0, \tau^{-1}\big)$ and $\tau \sim \text{Gamma}(1, 1)$.

1. Derive the variational densities $q^\star(\mu \,|\, y) = \exp\{\mathbb{E}_\tau[\ln p(\mu, \tau, y)] - \ln c_\mu\}$ and $q^\star(\tau \,|\, y)$ under the mean–field assumption.
2. Implement a variational algorithm that refines the parameters of the variational distribution until convergence occurs.
3. Compare the variational algorithm to Gibbs sampling with respect to bias and speed using the following simulated data: `set.seed( 50 ) ; y <- rnorm( 100 )`