1. (a) Let $X$ have the standard Laplace distribution. Use importance sampling based on 100000 draws from the standard normal as the proposal density to estimate $E(X)$, $Var(X)$ and $Pr(X > 2)$.

(b) Plot the logarithms of the importance weights as a histogram. Is the distribution of these log-importance-weights symmetric? Do you occasionally get extremely large importance weights? Extremely small ones? Which type of outliers is more worrisome?

(c) Compare your Monte Carlo estimates with the true values. Are there biases in the Monte Carlo estimates? How large are the Monte Carlo standard deviations? (Is there a theoretical formula for the standard deviation of your Monte Carlo estimator?)

(d) Let $Y$ have the standard normal distribution. Use importance sampling based on 100000 draws from the standard Laplace as the proposal density to estimate $E(Y)$, $var(Y)$ and $Pr(Y > 2)$. Repeat parts (b) and (c).

2. Continued from Problem 1 in Homework 2. Consider a Bayesian solution to this Poisson regression problem, using constant priors on $\alpha, \beta$ subject to the natural constraint that $\alpha + \beta X_i \geq 0$ for all $i$.

(a) Choose a suitable fine grid, and evaluate the posterior means of $\alpha$ and $\beta$ by writing these as suitable integrals and then approximating such integrals with finite sums (i.e., the grid method). Note that the joint posterior is not normalized.

(b) Using the MLE and observed Fisher information matrix you calculated for homework 2, construct a normal approximation to the posterior distribution. Report the posterior means and 95 percent credible intervals for $\alpha$ and $\beta$ based on this normal approximation. Compare the posterior mean estimates with those in part (a).

(c) The normal has thin tails. Consider a bivariate t distribution with four degrees of freedom to approximate the posterior. Generate 10000 draws from this bivariate t distribution and report posterior means and 95 percent credible intervals based on these draws. Compare with (b).

(d) We can refine the approximation in part (c) by using importance sampling. Describe how importance sampling works (using the bivariate $t$ as the proposal density) in this context. Compare the estimated posterior means and 95% intervals with those in (b) and (c).

3. Consider the following Markov chain taking values on the nonnegative integers. The one-step transition probabilities are given by $(i, j \geq 0)\\n\Pr(X_{t+1} = j | X_t = i) = p_{ij} = \begin{cases} \frac{1}{2j} & j = i + 1, \ i \geq 0 \\
\frac{j}{2(j+1)} & j = i \geq 1 \\
\frac{1}{2} & j = i = 0 \\
\frac{1}{2} & j = i - 1, \ i \geq 1 \\
0 & |j-i| > 1 \end{cases}$
(a) Verify that these transition probabilities are legitimate.
(b) Argue that this Markov chain is irreducible.
(c) Is this chain aperiodic? Why or why not?
(d) Verify that the Poisson(1) distribution is a stationary distribution.
(e) Starting with $X_0 = 1$, simulate this Markov chain for 10,000 iterations. Compute the mean and variance of these 10,000 draws and compare with the theoretical mean and variance of the stationary distribution.

4. Consider the following model

$$
\phi_1 | (\mu, \sigma^2) \sim N(\mu, \sigma^2/(1 - \rho^2)), \\
\phi_{j+1} | (\phi_1, \ldots, \phi_j, \mu, \sigma^2) \sim N(\mu + \rho(\phi_j - \mu), \sigma^2), \quad j = 1, \ldots, J - 1, \\
y_j | (\phi_1, \ldots, \phi_J, \tau^2) \sim N(\phi_j, \tau^2), \quad j = 1, \ldots, J,
$$

where $\mu, \sigma^2, \tau^2$ are unknown parameters of interest, $\rho$ is a known constant between $-1$ and $1$, $\phi = (\phi_1, \ldots, \phi_J)$ is a sequence of latent variables, and $y = (y_1, \ldots, y_J)$ is the observed data. (This is known as a Gaussian state-space model.)

(a) Assume the prior $p(\mu, \sigma^2, \tau^2) \propto \tau^{-4}e^{-1/(2\tau^2)}$. That is, the prior on $(\mu, \sigma^2)$ is a non-informative constant prior, and independently, the prior on $1/\tau^2$ is an exponential distribution (check this). Write down the joint posterior of all parameters and latent variables given the observed data $y$.

(b) Derive the full conditionals. That is, compute the distribution of each of $\mu, \sigma^2, \tau^2, \phi_1, \ldots, \phi_J$, given all others and the observed data $y$.

(c) Implement a Gibbs sampler for simulating from the joint posterior.

(d) Simulate a data set with $J = 100$, $\rho = 0.8$, $\mu = 0$, $\tau^2 = \sigma^2 = 1$. Use your Gibbs sampler to simulate from the posterior. Comment on its convergence after examining the trajectories of the draws of $\mu, \sigma^2, \tau^2$ and the autocorrelations.

(e) Based on the Gibbs draws, plot the marginal posteriors of $\mu$, $\sigma^2$ and $\tau^2$ and compare with the true values. Do the true values lie within or outside of their respective 95 percent credible intervals?