Problem 1

Suppose X has a hypoexponential distribution with probability density function

$$f(x) = \begin{cases} 2e^{-x}(1-e^{-x}) & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases}$$

a) Find a method for simulating X via rejection sampling using the proposal density $g(x) = e^{-x}$ for $x \geq 0$. Derive the overall acceptance rate of the algorithm.

b) Find $F(x) = P(X \leq x)$ and its inverse F^{-1}. Based on this, specify another algorithm for simulating X (without using rejection sampling). Hint: You will encounter an equation that is quadratic in e^{-x}.

c) Suppose that V and W are independent exponentially distributed with rate parameters 1 and 2 respectively. Let $X = V + W$ and let Y be another suitably chosen function of V, W. Derive the joint density of X, Y (including its support) and use this to show that the marginal density of X is the above density $f(x)$. Based on this, provide a third method for simulating X.

Problem 2 Suppose that X has density proportional to

$$f^*(x) = \begin{cases} x^{a-1}e^{-bx+c\sqrt{x}} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0, \end{cases}$$

where the parameters $a > 1/2$, $b > 0$ and $-\infty < c < \infty$.

a) Is this density always log-concave? If not, find a suitable tranformation of the form $Y = w(X)$ such that the density of Y becomes log-concave. Hint: A power transformation of the form $Y = X^p$ may perhaps do the trick. Name an efficient method of simulation that is particularly suitable for simulating random variables with densities that are log-concave and briefly explain the main ideas behind the method.

Problem 3

Suppose that we have observations (y_i, x_i) from a linear regression model such that

$$y_i \sim N(\alpha + \beta x_i, 1/\tau)$$

for $i = 1, 2, \ldots, n$ and that we use a non-informative improper prior density on the parameters given by

$$\pi(\alpha, \beta, \tau) \propto \frac{1}{\tau} e^{-\frac{(\ln \tau - \mu_0)^2}{2\sigma_0^2}} I(\tau > 0),$$

that is, we use flat priors on the regression coefficents and a log-normal prior on the precision parameter τ.

a) Write down the joint density of $y_1, y_2, \ldots, y_n, \alpha, \beta, \tau$ up to a normalising constant.

b) Find the full conditional of β, that is, the conditional distribution of β conditional $y_1, y_2, \ldots, y_n, \alpha, \tau$. Hint: You will need to complete a square in β.

c) Similarly, find the full conditional of τ. Assuming that the log-normal prior on the precision τ is vague (σ_0^2 large), suggest a suitable proposal density $Q(\tau'|\tau)$. Derive the resulting log acceptance probability.

Problem 4 Suppose that we observe and iid sample $x_1, x_2, \ldots, x_n$ from the density

$$f_X(x) = (1-p)\frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{x^2}{2\sigma_0^2}} + p\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}},$$

where $0 < p < 1$, $\sigma_0^2 > 0$ and $\sigma_1^2 > 0$ are unknown parameters. We will consider an expectation-maximisation (EM) algorithm for computing the maximum likelihood estimates of these unknown parameters.

a) Show that the distribution of each x_i can be seen as a finite mixture by introducing latent variables $z_i \sim \text{Bernoulli}(p)$, $i = 1, 2, \ldots, n$ and by a suitable choice of $f_{X|Z}(x_i|z_i)$.

Assuming that we observed $z_1, z_2, \ldots, z_n$ in addition to $x_1, x_2, \ldots, x_n$, show that the resulting "complete data" log-likelihood can be written as

$$\ln f(\mathbf{x}, \mathbf{z}; p, \sigma_0^2, \sigma_1^2) = \sum_{i=1}^{n} \left(-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma_{z_i}^2) - \frac{x_i^2}{2\sigma_{z_i}^2} + z_i \ln(p) + (1 - z_i)\ln(1-p) \right).$$

b) What is the distribution of each z_i conditional on the observed data $\mathbf{x}$ and current values of the parameters at the tth iteration of the algorithm? In particular, derive the conditional probabilities

$$w_i^{(t)} = P(z_i = 1|\mathbf{x}, \sigma_0^{(t)}, \sigma_1^{(t)}, p^{(t)}).$$

To complete the E-step of the algorithm, find a closed-form expression for the function Q defined by

$$Q(\sigma_0^2, \sigma_1^2, p|\sigma_0^{(t)}, \sigma_1^{(t)}, p^{(t)}) = E(\ln f(\mathbf{x}, \mathbf{z}; p, \sigma_0^2, \sigma_1^2)|\mathbf{x}, \sigma_0^{(t)}, \sigma_1^{(t)}, p^{(t)}).$$

c) Derive the details of the M-step of the algorithm.