Problem 1

a) Using $g(x) = e^{-x}$ as proposal,

$$\frac{f(x)}{cg(x)} = 2(1 - e^{-x})/c \leq 1$$

for all $x \geq 0$ implies that we must have $c \geq 2$, with $c = 2$ being the optimal choice.

repeat
  Generate $x \sim g$
  Generate $u \sim \text{unif}(0, 1)$
until $u < 1 - e^{-x}$
return x

In this case and in general, the overall acceptance rate is

$$EP(U < \frac{f(X)}{cg(X)}|X) = \int \frac{f(x)}{cg(x)} g(x)dx = \frac{1}{c} = 1/2.$$

b) The cdf of X is

$$F(x) = \int_0^x 2e^{-t} - 2e^{-2t}dt = \left[-2e^{-t} + e^{-2t}\right]_0^x = e^{-2x} - 2e^{-x} + 1$$

Equating this to u gives

$$u = F(x) = e^{-2x} - 2e^{-1} + 1 = (1 - e^{-x})^2$$

which solved for x gives

$$x = F^{-1}(u) = -\ln(1 - \sqrt{u}).$$

We can thus generate X using the algorithm

Generate $u \sim \text{unif}(0, 1)$
$x \leftarrow -\ln(1 - \sqrt{u})$
return x.

c) If $V \sim \text{Exp}(1)$ and $W \sim \text{Exp}(2)$, the joint density of $X = V + W$ and $Y = W$ becomes

$$\begin{aligned} f_{X,Y}(x, y) &= f_{V,W}(v, w) \left| \frac{\partial(v, w)}{\partial(x, y)} \right| \\ &= 2e^{-v-2w} \cdot 1 \\ &= 2e^{-(v+w)-w} \\ &= 2e^{-x-y} \end{aligned}$$

for $y > 0$ and $x > y$ implying that the marginal density of X is

$$f_X(x) = \int_0^x f_{X,Y}(x,y)dy = 2e^{-x}\int_0^x e^{-y}dy = 2e^{-x}(1-e^{-x})$$

equal to $f(x)$ given in the problem statement.

We can thus generate X also using the algorithm

Generate $v \sim \exp(1)$
Generate $w \sim \exp(2)$
$x \leftarrow v + w$
return x.

Problem 2

a) Since both $\ln x$ and $\sqrt{x}$ are concave functions we see that

$$\ln f^*(x) = (a-1)\ln x - bx + c\sqrt{x}$$

is convex or linear at least when both $c \leq 0$ and $a \leq 1$. It is thus not always log-concave. The transformed variable $Y = \sqrt{X}$, however, has (unnormalised) density

$$\begin{aligned} f_Y^*(y) &= f^*(x)\left|\frac{dx}{dy}\right| \\ &= (y^2)^{a-1}e^{-by^2+cy}2y \\ &\propto y^{2a-1}e^{-by^2+cy} \end{aligned}$$

and log-density

$$\ln f_Y^*(y) = (2a-1)\ln y - by^2 + cy$$

which is concave since $-by^2$ is concave given that b is strictly positive, $(2a-1)\ln y$ is always concave since a is strictly larger than $1/2$, and cy is linear.

An efficient method for generating Y is therefore adaptive rejection sampling. Briefly, this is a form of rejection sampling where we use a piecewise log-linear proposal density $g(y)$ (a mixture of truncated exponential distributions) such that $cg(y)$ envelopes the target density $f^*(y)$. This proposal is adaptively refined by adding futher subdivisions each time a proposed sample is rejected. In this particular case, since Y has a lower bound of zero, a single exponential distribution can be used as the initial proposal. Having simulated Y, X can be generated by the inverse transformation $X = Y^2$.

Problem 3

a) The joint density of all random variables in the model is

$$
\begin{aligned}
f(\tau,\alpha,\beta,\mathbf{y}) &\propto I(\tau>0)\frac{1}{\tau}e^{-\frac{(\ln\tau-\mu_0)^2}{2\sigma_0^2}}\prod_{i=1}^{n}\sqrt{\frac{\tau}{2\pi}}e^{-\frac{1}{2}\tau(y_i-\alpha-\beta x_i)^2} \\
&\propto I(\tau>0)\frac{1}{\tau}e^{-\frac{(\ln\tau-\mu_0)^2}{2\sigma_0^2}}\tau^{n/2}e^{-\frac{\tau}{2}\sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2}
\end{aligned}
$$

b) Using the fact

$$\pi(\beta|\alpha,\tau,\mathbf{y})\propto f(\tau,\alpha,\beta,\mathbf{y})$$

it is clear that the full conditional of β is a Guassian density since the full conditional is proportional to the exponential function of a quadratic in β. Completing the square in β, we find that

$$
\begin{aligned}
\sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2 &= \beta^2\sum_{i=1}^{n}x_i^2-2\beta\sum_{i=1}^{n}x_i(y_i-\alpha)+\sum_{i=1}^{n}(y_i-\alpha)^2 && (1)\\
&= \left(\sum_{i=1}^{n}x_i^2\right)\left(\beta-\frac{\sum_{i=1}^{n}x_i(y_i-\alpha)}{\sum_{i=1}^{n}x_i^2}\right)^2+C && (2)
\end{aligned}
$$

where C is a constant not involving β. Thus the full conditional of β is a Gaussian density with mean

$$\frac{\sum_{i=1}^{n}x_i(y_i-\alpha)}{\sum_{i=1}^{n}x_i^2}$$

and variance

$$\left(\tau\sum_{i=1}^{n}x_i^2\right)^{-1}$$

c) Similarly, the full conditional of τ simplifies to

$$\pi(\tau|\alpha,\beta,\mathbf{y})\propto I(\tau>0)\frac{1}{\tau}e^{-\frac{(\ln\tau-\mu_0)^2}{2\sigma_0^2}}\tau^{n/2}e^{-\frac{\tau}{2}\sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2}$$

Assuming that the part coming from the likelihood dominates the part coming from the log-normal prior, a good choice of proposal in a Metropolis-within-Gibbs step is the density

$$Q(\tau'|\tau,\alpha,\beta,\mathbf{y})\propto\tau'^{n/2}e^{-\frac{\tau'}{2}\sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2}$$

that is, a Gamma distribution with shape parameter $n/2+1$ and rate parameter $\frac{1}{2}\sum_{i=1}^{n}(y_i - \alpha - \beta x_i)^2$. This leads to

$$\frac{\pi(\tau'|\dots)Q(\tau|\tau',\dots)}{\pi(\tau|\dots)Q(\tau'|\tau,\dots)} = \frac{\tau}{\tau'}e^{-\frac{(\ln\tau'-\mu_0)^2-(\ln\tau-\mu_0)^2}{2\sigma_0^2}}$$

and a log acceptance probability of

$$\log\alpha = \min\left(0, \ln\tau - \ln\tau' - \frac{(\ln\tau'-\mu_0)^2-(\ln\tau-\mu_0)^2}{2\sigma_0^2}\right)$$

Problem 4

a) If we assume that $X|Z=z \sim N(0,\sigma_z^2)$ and $Z \sim \text{Bernoulli}(p)$ so that $f_Z(z) = p^z(1-p)^{1-z}$, then by the law of total probability, the marginal density of X is

$$f_X(x) = \sum_{z=0}^{1} f_X|Z(x|z)f_Z(z) = (1-p)\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{x_i^2}{2\sigma_0^2}} + p\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{x_i^2}{2\sigma_1^2}}$$

If we observed both $\mathbf{x}$ and $\mathbf{z}$, the likelihood would be

$$\begin{aligned} f(\mathbf{x},\mathbf{z};p,\sigma_0^2,\sigma_1^2) &= \prod_{i=1}^{n} f_{X,Z}(x_i,z_i) \\ &= \prod_{i=1}^{n} f_{X|Z}(x_i|z_i)f_Z(z_i) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_{z_i}}\exp\left\{-\frac{x_i^2}{2\sigma_{z_i}^2}\right\}p^{z_i}(1-p)^{1-z_i} \end{aligned}$$

and the log-likelihood, omitting unnecessary constants,

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

b) Conditional on $\mathbf{x}$ and current parameter values, each z_i is Bernoulli distributed. Using Bayes theorem,

$$\begin{aligned} w_i^{(t)} &= P(z_i = 1|\mathbf{x},\sigma_0^{(t)},\sigma_1^{(t)},p^{(t)}) \\ &= \frac{f_{X|Z}(x_i|1)f_Z(1)}{\sum_{z=0}^{1} f_{X|Z}(x_i|z)f_Z(z)} \\ &= \frac{p^{(t)}\frac{1}{\sigma_1^{(t)}}e^{-\frac{x_i^2}{2\sigma_1^{2(t)}}}}{p^{(t)}\frac{1}{\sigma_1^{(t)}}e^{-\frac{x_i^2}{2\sigma_1^{2(t)}}} + (1-p^{(t)})\frac{1}{\sigma_0^{(t)}}e^{-\frac{x_i^2}{2\sigma_0^{2(t)}}}} \end{aligned}$$

Taking conditional expectation of the complete data log likelihood

$$
\begin{aligned}
&Q(\sigma_0^2, \sigma_1^2, p | \sigma_0^{(t)}, \sigma_1^{(t)}, p^{(t)}) \\
=&E(\ln f(\mathbf{x}, \mathbf{z}; p, \sigma_0^2, \sigma_1^2) | \mathbf{x}, \sigma_0^{(t)}, \sigma_1^{(t)}, p^{(t)}) \\
=&\sum_{i=1}^{n} \left(w_i^{(t)} \Big(-\frac{1}{2}\ln(\sigma_1^2) - \frac{x_i^2}{2\sigma_1^2} + \ln(p) \Big) + (1 - w_i^{(t)}) \Big(-\frac{1}{2}\ln(\sigma_1^2) - \frac{x_i^2}{2\sigma_1^2} + \ln(1-p) \Big) \right)
\end{aligned}
$$

c) At its maximum $\partial Q / \partial p = 0$, $\partial Q / \partial(\sigma_0^2) = 0$, $\partial Q / \partial(\sigma_1^2) = 0$ which after some algebra yields

$$
p^{(t+1)} = \frac{1}{n} \sum_{i=1}^{n} w_i^{(t)}, \quad \sigma_1^{2(t+1)} = \frac{\sum_{i=1}^{n} w_i^{(t)} x_i^2}{\sum_{i=1}^{n} w_i^{(t)}}, \quad \sigma_0^{2(t+1)} = \frac{\sum_{i=1}^{n} (1 - w_i^{(t)}) x_i^2}{\sum_{i=1}^{n} (1 - w_i^{(t)})}
$$