1. Suppose $y$ is a random variable with cdf $F(y) = 1 - e^{-\lambda y^\alpha}$ for $y > 0$, $\alpha > 0$. We say $y \sim \text{Weibull}(\lambda, \alpha)$. Explain how to simulate $y$ from Uniform $(0, 1)$ random variables. Note that $y$ is a simple transformation of an Exp $(\lambda)$ random variable. What is the transformation?

We begin by observing the nature of the transformation. Let $x \sim \text{Exp}(\lambda)$. Then $f_x(x) = \lambda e^{-\lambda x}$. Consider the transformation $y = g(x) = |x^{1/\alpha}|$ with inverse transformation $x = g^{-1}(y) = y^\alpha$. By the change of variables formula, this means the pdf of $y$ must be

$$f_y(y) = f_x \left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right|$$

$$= f_x \left(y^\alpha\right) |\alpha y^{\alpha-1}|$$

$$= \lambda \alpha y^{\alpha-1} \exp(-\lambda y^{\alpha}),$$

which is the pdf for a Weibull $(\lambda, \alpha)$ random variable. With this information, and knowing how to simulate an exponential random variable from a Uniform $(0, 1)$ random variable, simulating a Weibull $(\lambda, \alpha)$ random variable is straightforward.

Let $u \sim \text{Uniform}(0, 1)$. Let $F_x(x) = (1 - u)$ where $F_x$ is the Exp $(\lambda)$ cdf—or in other words, let

$$x = F_x^{-1}(1 - u) = \frac{-\log u}{\lambda}. $$

This gives us $x \sim \text{Exp}(\lambda)$. Then, given what we’ve found above regarding the transformation, let $y = |x^{1/\alpha}|$. This gives us $y \sim \text{Weibull}(\lambda, \alpha)$.
2. If $Y_i \overset{iid}{\sim} \Gamma(a_i, b)$ for $i \in \{1, ..., k\}$, we have shown that

$$(Z_1, ..., Z_k) = \left( \frac{Y_1}{S}, ..., \frac{Y_k}{S} \right) \sim \text{Dirichlet} (a_1, ..., a_k),$$

where $S = \sum_{i=1}^k Y_i$. Use Proposition B.4 to transform $Y_1, ..., Y_k$ into the random vector $Z_1, ..., Z_{k-1}, S$. Show that $S$ is independent of the other variables by showing that the joint density of the random vector is the product of a Dirichlet $(a_1, ..., a_k)$ density and a $\Gamma\left(\sum_{i=1}^k a_i, b \right)$ density. The general Dirichlet density is an obvious extension of the three-parameter Dirichlet density given in Table 2.1.

First, we observe that the joint density of the $Y_i$’s for $i \in \{1, ..., k\}$ is

$$f(Y_1, ..., Y_k) = \prod_{i=1}^k f(Y_i)$$

$$= \prod_{i=1}^k \frac{b^{a_i}}{\Gamma(a_i)} y_i^{a_i - 1} \exp(-by_i) I(0, \infty)(y_i)$$

$$= \frac{b^{\sum_{i=1}^k a_i}}{\prod_{i=1}^k \Gamma(a_i)} \left( \prod_{i=1}^k y_i^{a_i - 1} \right) \exp \left( -b \sum_{i=1}^k y_i \right)$$

Now, if for every $Z_i$ we define a transformation function

$$\begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{k-1} \\ S \end{bmatrix} = g \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{k-1} \\ Y_k \end{bmatrix} = \begin{bmatrix} Y_1/\sum_{i=1}^k Y_i \\ Y_2/\sum_{i=1}^k Y_i \\ \vdots \\ Y_{k-1}/\sum_{i=1}^k Y_i \\ \sum_{i=1}^k Y_i \end{bmatrix},$$

which admits an inverse

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{k-1} \\ Y_k \end{bmatrix} = g^{-1} \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{k-1} \\ S \end{bmatrix} = \begin{bmatrix} Z_1 S \\ Z_2 S \\ \vdots \\ Z_{k-1} S \\ S \left(1 - \sum_{i=1}^{k-1} Z_i \right) \end{bmatrix}$$
To calculate the distribution of the transformed variables, we need to obtain the Jacobian.

\[
|J| = \begin{vmatrix}
\frac{\partial}{\partial Z_1} Z_1 S & \frac{\partial}{\partial Z_2} Z_1 S & \ldots & \frac{\partial}{\partial Z_{k-1}} Z_1 S & \frac{\partial}{\partial Z_k} Z_1 S \\
\frac{\partial}{\partial Z_1} Z_2 S & \frac{\partial}{\partial Z_2} Z_2 S & \ldots & \frac{\partial}{\partial Z_{k-1}} Z_2 S & \frac{\partial}{\partial Z_k} Z_2 S \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{\partial}{\partial Z_1} Z_{k-1} S & \frac{\partial}{\partial Z_2} Z_{k-1} S & \ldots & \frac{\partial}{\partial Z_{k-1}} Z_{k-1} S & \frac{\partial}{\partial Z_k} Z_{k-1} S \\
\frac{\partial}{\partial Z_1} (1 - \sum Z_i) & \frac{\partial}{\partial Z_2} (1 - \sum Z_i) & \ldots & \frac{\partial}{\partial Z_{k-1}} (1 - \sum Z_i) & \frac{\partial}{\partial Z_k} (1 - \sum Z_i)
\end{vmatrix}
\]

\[
= \begin{vmatrix}
S & 0 & \ldots & 0 & Z_1 \\
0 & S & \ldots & 0 & Z_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & S & Z_{k-1} \\
-S & -S & \ldots & -S & \left(1 - \sum_{i=1}^{k-1} Z_i\right)
\end{vmatrix}
\]

By careful consideration of the minors of the matrix, we are able to determine that the determinant \(|J|\) is equal to \(S^{k-1}\). (This takes too much space for me to want to show how to do it in \LaTeX. The work proceeds by choosing the \(S\) row of the matrix and obtaining the minors for each element of this row; then choosing the \(\partial S\) column within the minor; recognizing that the subminors for each element of this row must either be diagonal or contain a row of zeroes, and calculating the determinant accordingly.)

Then we can obtain the joint density for the random vector \(\{Z_1, \ldots, Z_{k-1}, S\}\). First define \(Z_k = 1 - \sum_{i=1}^{k-1} Z_i\), then the density can be written as follows:

\[
f_{Z,S}(Z_1, \ldots, Z_{k-1}, S) = f_Y \left(g^{-1} \left(\begin{bmatrix} Z_1 & \ldots & Z_{k-1} & S \end{bmatrix}^T\right)\right) \ |J| \\
= \frac{b^{\sum_{i=1}^k a_i}}{\prod_{i=1}^k \Gamma(a_i)} \left(\prod_{i=1}^{k-1} (Z_i S)^{a_i-1}\right) \exp \left(-b \sum_{i=1}^{k-1} (Z_i S)\right) \\
\times \left(S \left(1 - \sum_{i=1}^{k-1} Z_i\right)\right)^{a_k-1} \exp \left(-bS \left(1 - \sum_{i=1}^{k-1} Z_i\right)\right) \ S^{k-1} \\
= \frac{b^{\sum_{i=1}^k a_i}}{\prod_{i=1}^k \Gamma(a_i)} \left(\prod_{i=1}^k Z_i^{a_i-1}\right) S^{(\sum_{i=1}^k a_i)-1} \exp \left(-bS\right) \\
= \frac{\Gamma \left(\sum_{i=1}^k a_i\right)}{\prod_{i=1}^k \Gamma(a_i)} \left(\prod_{i=1}^k Z_i^{a_i-1}\right) \\
\times \frac{b^{\sum_{i=1}^k a_i}}{\Gamma \left(\sum_{i=1}^k a_i\right)} S^{(\sum_{i=1}^k a_i)-1} \exp \left(-bS\right)
\]

Observe that in the final equality: the first line is the density for a Dirichlet \((a_1, \ldots, a_k)\) random vector; the second line is the density for a Gamma \(\left(\sum_{i=1}^k a_i, b\right)\) random variable; and since the two do not share random elements, the Dirichlet vector and the Gamma variable are independent.
3. In acceptance-rejection sampling, consider the general choice of candidate distribution for a log-concave target density as presented in class. Recall that we have \( \tilde{\theta}_1 \) and \( \tilde{\theta}_2 \), points on either side of the mode of \( \ell(\theta) \equiv \log[p_*(\theta)] \); and that we have tangent lines to \( \ell(\theta) \) calculated at these points,

\[
\gamma_i(\theta) = \ell(\tilde{\theta}_i) + \ell'(\tilde{\theta}_i)(\theta - \tilde{\theta}_i),
\]

where \( i \in \{1, 2\} \). Then \( p_*(\theta) = e^{\ell(\theta)} \leq e^{\gamma_i(\theta)} \) for both \( i = 1 \) and \( i = 2 \). We choose

\[
Mq(\theta) = \min \left\{ e^{\gamma_1(\theta)}, e^{\gamma_2(\theta)} \right\}.
\]

(a) Find \( \theta_* \) by setting \( \gamma_1(\theta_*) = \gamma_2(\theta_*) \) and solving.

\[
\gamma_1(\theta_*) = \gamma_2(\theta_*) \\
\ell(\tilde{\theta}_1) + \ell'(\tilde{\theta}_1)(\theta_* - \tilde{\theta}_1) = \ell(\tilde{\theta}_2) + \ell'(\tilde{\theta}_2)(\theta_* - \tilde{\theta}_2) \\
\ell(\tilde{\theta}_1) - \ell(\tilde{\theta}_2) = \ell'(\tilde{\theta}_2)(\theta_* - \tilde{\theta}_2) - \ell'(\tilde{\theta}_1)(\theta_* - \tilde{\theta}_1) \\
= \theta_* \left[ \ell'(\tilde{\theta}_2) - \ell'(\tilde{\theta}_1) \right] - \ell'(\tilde{\theta}_2)\tilde{\theta}_2 + \ell'(\tilde{\theta}_1)\tilde{\theta}_1
\]

(b) Integrate \( Mq(\theta) \) over \(( -\infty, \infty )\) to determine \( M \) as a function of \( \theta_* \), \( \tilde{\theta}_1 \), and \( \tilde{\theta}_2 \).

Observe that by construction, one of the \( \gamma_i \)’s will have a positive slope and one will have a negative slope. Without loss of generality, we assume that \( \ell'(\tilde{\theta}_1) > 0 \) and \( \ell'(\tilde{\theta}_2) < 0 \). Then for all \( \theta < \theta_* \), we will have \( \gamma_1(\theta) < \gamma_2(\theta) \) and \( Mq(\theta) = e^{\gamma_1(\theta)} \). Similarly, for all \( \theta > \theta_* \) we have \( Mq(\theta) = e^{\gamma_2(\theta)} \). Then we can divide the integral into two halves:
\[ \int_{-\infty}^{\infty} Mq(\theta) d\theta = \int_{-\infty}^{\theta_*} Mq(\theta) d\theta + \int_{\theta_*}^{\infty} Mq(\theta) d\theta \]

\[ = \int_{-\infty}^{\theta_*} \exp \left( \ell(\tilde{\theta}_1) + \ell'(\tilde{\theta}_1)(\theta - \tilde{\theta}_1) \right) d\theta + \int_{\theta_*}^{\infty} \exp \left( \ell(\tilde{\theta}_2) + \ell'(\tilde{\theta}_2)(\theta - \tilde{\theta}_2) \right) d\theta \]

\[ = \int_{-\infty}^{\theta_*} e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1} \exp \left( \ell'(\tilde{\theta}_1)\theta \right) d\theta + \int_{\theta_*}^{\infty} e^{\ell(\tilde{\theta}_2) - \ell'(\tilde{\theta}_2)\tilde{\theta}_2} \exp \left( \ell'(\tilde{\theta}_2)\theta \right) d\theta \]

\[ = e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1} \int_{-\infty}^{\theta_*} \exp \left( \ell'(\tilde{\theta}_1)\theta \right) d\theta + e^{\ell(\tilde{\theta}_2) - \ell'(\tilde{\theta}_2)\tilde{\theta}_2} \int_{\theta_*}^{\infty} \exp \left( \ell'(\tilde{\theta}_2)\theta \right) d\theta \]

\[ = \frac{e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1}}{\ell'(\tilde{\theta}_1)} \left[ \exp \left( \ell'(\tilde{\theta}_1)\theta \right) \right]_{-\infty}^{\theta_*} + \frac{e^{\ell(\tilde{\theta}_2) - \ell'(\tilde{\theta}_2)\tilde{\theta}_2}}{\ell'(\tilde{\theta}_2)} \left[ \exp \left( \ell'(\tilde{\theta}_2)\theta \right) \right]_{\theta_*}^{\infty} \]

\[ = \frac{e^{\ell(\tilde{\theta}_2) - \ell'(\tilde{\theta}_2)\tilde{\theta}_2}}{|\ell'(\tilde{\theta}_2)|} \left[ \exp \left( \ell'(\tilde{\theta}_2)\theta \right) \right]_{\theta_*}^{\infty} \]

\[ = M \]

(c) Obtain the cdf based on the density \( q(\theta) \), say

\[ Q(v) = \int_{-\infty}^{v} q(\theta) d\theta. \]

Do this first for \( v \leq \theta_* \) and then for \( v > \theta_* \). Calculate the latter as \( \int_{-\infty}^{\theta_*} q(\theta) d\theta + \int_{\theta_*}^{v} q(\theta) d\theta \).

Observe that after the preceding step, we can now concisely (ha ha) define \( q(\theta) \) as

\[ q(\theta) = \begin{cases} \frac{1}{M} \exp \left( \ell(\tilde{\theta}_1) + \ell'(\tilde{\theta}_1)(\theta - \tilde{\theta}_1) \right) & \theta \leq \theta_* \\ \frac{1}{M} \exp \left( \ell(\tilde{\theta}_2) + \ell'(\tilde{\theta}_2)(\theta - \tilde{\theta}_2) \right) & \theta > \theta_* \end{cases} \]

or even more concisely as

\[ q(\theta) = \begin{cases} \frac{e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1}}{M} \exp \left( \ell'(\tilde{\theta}_1)\theta \right) & \theta \leq \theta_* \\ \frac{e^{\ell(\tilde{\theta}_2) - \ell'(\tilde{\theta}_2)\tilde{\theta}_2}}{M} \exp \left( \ell'(\tilde{\theta}_2)\theta \right) & \theta > \theta_* \end{cases} \]

We consider this latter form more concise, because we can recognize it as the concatenation of a reflected \( \text{Exp} \left( \ell'(\tilde{\theta}_1) \right) \) density and a \( \text{Exp} \left( -\ell'(\tilde{\theta}_2) \right) \) density with appropriate rescaling factors. Because we know how to find the cdf for an exponential density, putting \( q(\theta) \) in this
form helps us see how to proceed.

Then for \( v \leq \theta_* \) we can express \( Q(v) \) as

\[
Q(v) = \int_{-\infty}^{v} \frac{e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1}}{M} \exp \left( \ell'(\tilde{\theta}_1)\theta \right) \, d\theta
\]

\[
= \frac{e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1}}{\ell'(\tilde{\theta}_1)M} \int_{\theta_* - v}^{\infty} \ell'(\tilde{\theta}_1) \exp \left( -\ell'(\tilde{\theta}_1)\theta \right) \, d\theta
\]

\[
= \frac{e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1}}{\ell'(\tilde{\theta}_1)M} \left[ 1 - \int_{0}^{\theta_* - v} \ell'(\tilde{\theta}_1) \exp \left( -\ell'(\tilde{\theta}_1)\theta \right) \, d\theta \right]
\]

\[
= \frac{e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1}}{\ell'(\tilde{\theta}_1)M} \left[ 1 - \left[ 1 - \exp \left( -\ell'(\tilde{\theta}_1)(\theta_* - v) \right) \right] \right]
\]

\[
= \frac{e^{\ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1}}{\ell'(\tilde{\theta}_1)M} \left[ \exp \left( -\ell'(\tilde{\theta}_1)(\theta_* - v) \right) \right]
\]

\[
= \frac{\exp \left( \ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1 \right)}{\ell'(\tilde{\theta}_1)M} \left[ \theta_1 + \theta_* - v \right]
\]

When \( v = \theta_* \), this reduces to

\[
Q(\theta_*) = \frac{\exp \left( \ell(\tilde{\theta}_1) - \ell'(\tilde{\theta}_1)\tilde{\theta}_1 \right)}{\ell'(\tilde{\theta}_1)M}.
\]

Finally, for \( v > \theta_* \) we have

\[
Q(v) = Q(\theta_*) + \int_{\theta_*}^{v} \frac{e^{\ell(\tilde{\theta}_2) - \ell'(\tilde{\theta}_2)\tilde{\theta}_2}}{M} \exp \left( \ell'(\tilde{\theta}_2)\theta \right) \, d\theta
\]

\[
= Q(\theta_*) + \frac{e^{\ell(\tilde{\theta}_2) - \ell'(\tilde{\theta}_2)\tilde{\theta}_2}}{-\ell'(\tilde{\theta}_2)M} \int_{0}^{v - \theta_*} \left[ -\ell'(\tilde{\theta}_2) \right] \exp \left( \ell'(\tilde{\theta}_2)\theta \right) \, d\theta
\]

\[
= Q(\theta_*) + \frac{e^{\ell(\tilde{\theta}_2) - \ell'(\tilde{\theta}_2)\tilde{\theta}_2}}{-\ell'(\tilde{\theta}_2)M} \left[ 1 - \exp \left( \ell'(\tilde{\theta}_2)[v - \theta_*] \right) \right]
\]

(d) Solve \( Q(v) = u \) for \( v \) so that \( v = Q^{-1}(u) \). Thus if we sample \( U \sim \text{Uniform}(0, 1) \), we have \( Q^{-1}(U) \sim q(\cdot) \).

You know what? If you’ve made it this far, just pat yourself on the back and go home. I don’t want to do this bit, and neither do you—but at this point, you should be able to recognize that solving for \( v \) can be done with a stack of algebra, and it is possible to get a general solution to the whole mess in terms of our arbitrarily chosen tangent points.