MAT7035: Computational Statistics

Suggested Solutions to Assignment 1

1.5 Solution: (a) The cdf of $X \sim \text{Logistic}(\mu, \sigma^2)$ with density

$$f(x) = \frac{\exp(-\frac{x-\mu}{\sigma})}{\sigma\{1+\exp(-\frac{x-\mu}{\sigma})\}^2}, \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma > 0,$$

is given by

$$F(x) \;=\; \left[1 + \exp\left(-\frac{x-\mu}{\sigma}\right)\right]^{-1}.$$

Thus, $F(X) \stackrel{\text{d}}{=} U \sim U(0,1)$ implies

$$X \stackrel{\text{d}}{=} F^{-1}(U) = \mu + \sigma \log\left(\frac{U}{1-U}\right).$$

The inverse method is as follows:

— Step 1: Draw $U = u \sim U(0,1)$;

— Step 2: Return $x = \mu + \sigma \log[u/(1-u)]$.

(b) The cdf of the Rayleigh distribution with density

$$f(x) = \sigma^{-2} x \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x > 0, \quad \sigma > 0,$$

is given by

$$\begin{aligned} F(x) &= \int_0^x \sigma^{-2} y \exp\left(-\frac{y^2}{2\sigma^2}\right) \,\text{d}y \\ &= -\exp\left(-\frac{y^2}{2\sigma^2}\right)\Bigg|_0^x = 1 - \exp\left(-\frac{x^2}{2\sigma^2}\right). \end{aligned}$$

Thus, $F(X) \stackrel{\mathrm{d}}{=} U \sim U(0,1)$ implies

$$X \stackrel{\mathrm{d}}{=} F^{-1}(U) = \sigma\sqrt{-2\log(1-U)} \stackrel{\mathrm{d}}{=} \sigma\sqrt{-2\log(U)}\ .$$

The inverse method is as follows:

— Step 1: Draw $U = u \sim U(0,1)$;

— Step 2: Return $x = \sigma\sqrt{-2\log(u)}$.

(c) The cdf of the triangular distribution with density

$$f(x) = \frac{2}{a}\left(1 - \frac{x}{a}\right), \quad 0 \leqslant x < a, \quad a > 0,$$

is given by

$$\begin{aligned} F(x) &= \int_0^x \frac{2}{a}\left(1 - \frac{y}{a}\right) \mathrm{d}y \\ &= \frac{1}{a}\left(2y - \frac{y^2}{a}\right)\Bigg|_0^x = \frac{1}{a}\left(2x - \frac{x^2}{a}\right). \end{aligned}$$

Thus, $F(X) \stackrel{\mathrm{d}}{=} U \sim U(0,1)$ implies

$$X \stackrel{\mathrm{d}}{=} F^{-1}(U) = a\Big(1 - \sqrt{1-U}\Big) \stackrel{\mathrm{d}}{=} a\Big(1 - \sqrt{U}\Big).$$

The inverse method is as follows:

— Step 1: Draw $U = u \sim U(0,1)$;

— Step 2: Return $x = a(1 - \sqrt{u}\,)$.

(d) The cdf of the Pareto distribution with density

$$f(x) = ab^a/x^{a+1}, \quad x \geqslant b > 0, \quad a > 0,$$

is given by

$$F(x) = \int_b^x \frac{ab^a}{y^{a+1}}\,\mathrm{d}y = -(b/y)^a\Big|_b^x = 1 - (b/x)^a.$$

Thus, $F(X) \overset{\mathrm{d}}{=} U \sim U(0,1)$ implies

$$X \overset{\mathrm{d}}{=} F^{-1}(U) = \frac{b}{(1-U)^{1/a}} \overset{\mathrm{d}}{=} \frac{b}{U^{1/a}}.$$

The inverse method is as follows:

— Step 1: Draw $U = u \sim U(0,1)$;

— Step 2: Return $x = b/u^{1/a}$.

(e) The cdf of the Gumbel–minimum distribution with density

$$f(x) = \frac{1}{\sigma}\,\mathrm{e}^{\frac{x-\mu}{\sigma}} \exp(-\,\mathrm{e}^{\frac{x-\mu}{\sigma}}), \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma > 0,$$

is given by

$$F(x) \;=\; 1 - \exp\left(-\,\mathrm{e}^{\frac{x-\mu}{\sigma}}\right).$$

Thus, $F(X) \overset{\mathrm{d}}{=} U \sim U(0,1)$ implies

$$X \overset{\mathrm{d}}{=} F^{-1}(U) = \mu + \sigma \log\{-\log(1-U)\} \overset{\mathrm{d}}{=} \mu + \sigma \log(-\log U).$$

The inverse method is as follows:

— Step 1: Draw $U = u \sim U(0,1)$;

— Step 2: Return $x = \mu + \sigma \log(-\log u)$.

(f) The cdf of the Gumbel–maximum distribution with density

$$f(x) = \frac{1}{\sigma}\,\mathrm{e}^{-\frac{x-\mu}{\sigma}} \exp(-\,\mathrm{e}^{-\frac{x-\mu}{\sigma}}), \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma > 0,$$

is given by

$$F(x) \;=\; \exp\left(-\,\mathrm{e}^{-\frac{x-\mu}{\sigma}}\right).$$

Thus, $F(X) \overset{\mathrm{d}}{=} U \sim U(0,1)$ implies

$$X \overset{\mathrm{d}}{=} F^{-1}(U) = \mu - \sigma \log(-\log U).$$

The inverse method is as follows:

— Step 1: Draw $U = u \sim U(0, 1)$;

— Step 2: Return $x = \mu - \sigma \log(-\log u)$.

(g1) Let $U \sim U(0, 1)$. The cdf of $X_{(1)} = \min(X_1, \dots, X_n)$ is

$$\begin{aligned} G_1(x) &= \Pr(X_{(1)} \leqslant x) = 1 - \Pr\{\min(X_1, \dots, X_n) > x\} \\ &= 1 - \Pr(X_1 > x, \dots, X_n > x) = 1 - [1 - F(x)]^n. \end{aligned}$$

We have

$$\begin{aligned} & U \stackrel{\mathrm{d}}{=} G_1(X_{(1)}) = 1 - [1 - F(X_{(1)})]^n, \\ \Rightarrow\ & 1 - F(X_{(1)}) \stackrel{\mathrm{d}}{=} (1 - U)^{1/n} \stackrel{\mathrm{d}}{=} U^{1/n}, \\ \Rightarrow\ & X_{(1)} \stackrel{\mathrm{d}}{=} F^{-1}(1 - U^{1/n}). \end{aligned}$$

The inverse method for generating $X_{(1)}$ is as follows:

— Step 1: Draw $U = u \sim U(0, 1)$;

— Step 2: Return $x = F^{-1}(1 - u^{1/n})$.

(g2) Let $U \sim U(0, 1)$. The cdf of $X_{(n)} = \max(X_1, \dots, X_n)$ is

$$\begin{aligned} G_n(x) &= \Pr(X_{(n)} \leqslant x) = \Pr\{\max(X_1, \dots, X_n) \leqslant x\} \\ &= \Pr(X_1 \leqslant x, \dots, X_n \leqslant x) = [F(x)]^n. \end{aligned}$$

We have $U \stackrel{\mathrm{d}}{=} G_n(X_{(n)}) = [F(X_{(n)})]^n$, i.e., $F(X_{(n)}) \stackrel{\mathrm{d}}{=} U^{1/n}$, Thus, $X_{(n)} \stackrel{\mathrm{d}}{=} F^{-1}(U^{1/n})$. The inverse method for generating $X_{(n)}$ is as follows:

— Step 1: Draw $U = u \sim U(0, 1)$;

— Step 2: Return $x = F^{-1}(u^{1/n})$.

1.6 Proof: (a) Let $Y = c\,U f(\mathbf{x})$ and we need to prove that $\binom{\mathbf{x}}{Y} \sim U(\mathbb{A})$.

Since $\mathbf{x} \sim f(\boldsymbol{x})$, $U \sim U(0,1)$ and $\mathbf{x} \perp\!\!\!\perp U$, the joint density of $\binom{\mathbf{x}}{U}$ is

$$h(\boldsymbol{x}, u) = f(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d, \quad 0 < u < 1.$$

Making the following transformation

$$\begin{cases} \boldsymbol{x} &= \boldsymbol{x}, & \boldsymbol{x} \in \mathbb{R}^d, \\ y &= c\,u f(\boldsymbol{x}), & 0 \leqslant y \leqslant c\,f(\boldsymbol{x}), \end{cases}$$

the joint density of $\binom{\mathbf{x}}{Y}$ is given by

$$\begin{aligned} g(\boldsymbol{x}, y) &= h(\boldsymbol{x}, u) \times \left|\frac{\partial(\boldsymbol{x}, u)}{\partial(\boldsymbol{x}, y)}\right| = h(\boldsymbol{x}, u) \Big/ \left|\frac{\partial(\boldsymbol{x}, y)}{\partial(\boldsymbol{x}, u)}\right| \\ &= f(\boldsymbol{x}) \Big/ \left|\det\begin{pmatrix} \boldsymbol{I}_d & \mathbf{0}_d \\ * & cf(\boldsymbol{x}) \end{pmatrix}\right| = 1/c, \end{aligned}$$

which is an constant. In other words, $\binom{\mathbf{x}}{Y} \sim U(\mathbb{A})$ and the volume of $\mathbb{A}$ is c.

(b) If $\binom{\mathbf{z}}{W} \sim U(\mathbb{A})$, then $\boldsymbol{z} \in \mathbb{R}^d$, $0 \leqslant w \leqslant cf(z)$, and their joint density is

$$h(\boldsymbol{z}, w) = \frac{1}{v(\mathbb{A})} = \frac{1}{c}.$$

Thus, the marginal density of $\mathbf{z}$ is given by

$$\int_0^{cf(\boldsymbol{z})} h(\boldsymbol{z}, w)\,\mathrm{d}w = \frac{1}{c}\int_0^{cf(\boldsymbol{z})} \mathrm{d}w = f(\boldsymbol{z}).$$

1.7 Solution: Use the SIR method, we consider the logistic distribution $g(x)$ as the importance sampling density. Thus, the importance ratio $\mathrm{w}(x) = f(x)/g(x)$.

THE SIR METHOD:

Step 1. Generate $X^{(1)}, \ldots, X^{(J)} \stackrel{\text{iid}}{\sim} g(\cdot)$;[1]

Step 2. Select a subset $\{X^{(k_i)}\}_{i=1}^m$ from $\{X^{(j)}\}_{j=1}^J$ via resampling without replacement from the discrete distribution on $\{X^{(j)}\}$ with probabilities $w_j = \mathrm{w}(X^{(j)}) / \sum_{i=1}^J \mathrm{w}(X^{(i)})$, $j = 1, \ldots, J$.

For example, we run the SIR algorithm by setting $\theta_0 = 1/2$, $J = 200,000$ and $m = 10,000$. Figure 1.1(b) shows that the histogram entirely recovers the target density function $f(x)$ very well.

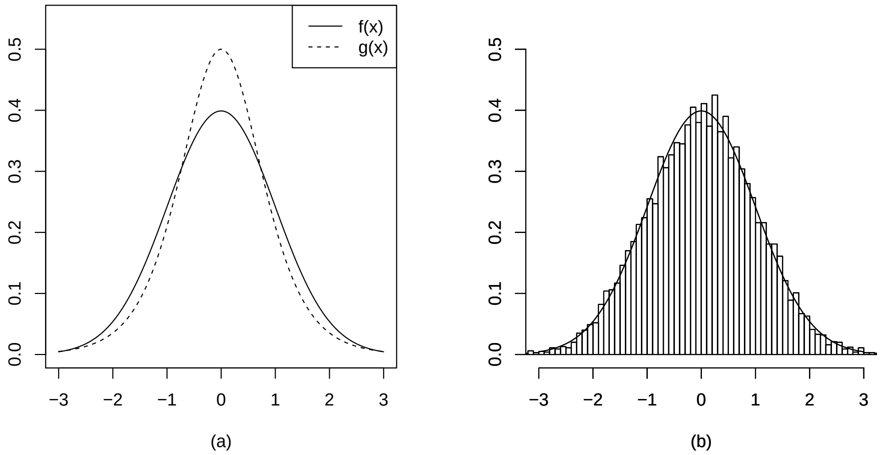

Figure 1.1: The SIR method.

1.8 Proof: Method 1: The Inversion Method. The cdf of X is given by

$$\begin{aligned} F(x) &= \int_0^x f(y)\,\mathrm{d}y = \int_0^x \frac{a}{2\sinh(a)} \sin(y) \exp\{a\cos(y)\}\,\mathrm{d}y \\ &= \frac{-1}{2\sinh(a)} \exp\{a\cos(y)\}\Big|_0^x = \frac{\mathrm{e}^a - \mathrm{e}^{a\cos(x)}}{\mathrm{e}^a - \mathrm{e}^{-a}}, \quad 0 < x < \pi. \end{aligned}$$

Thus, $F(X) \stackrel{\mathrm{d}}{=} U \sim U(0,1)$ implies

$$X \stackrel{\mathrm{d}}{=} F^{-1}(U) = \arccos\left[a^{-1} \log\left\{(1-U)\,\mathrm{e}^a + U\,\mathrm{e}^{-a}\right\}\right].$$

[1]See the solution to Exercise 1.5 (a).

<u>Method 2: The Stochastic Representation Method</u>. Let

$$Y = \arccos\left[a^{-1}\log\left\{(1-U)\,\mathrm{e}^{a} + U\,\mathrm{e}^{-a}\right\}\right], \tag{SA.2}$$

where $U \sim U(0,1)$, we only need to prove that the density of Y is

$$f(y) = \frac{a}{2\sinh(a)}\sin(y)\exp\{a\cos(y)\}, \quad 0 < y < \pi.$$

We argue as follows: From (SA.2), we obtain

$$a\cos(y) = \log[(1-u)\,\mathrm{e}^{a} + u\,\mathrm{e}^{-a}] = \log[\,\mathrm{e}^{a} - u(\,\mathrm{e}^{a} - \,\mathrm{e}^{-a})], \tag{SA.3}$$

yielding

$$\mathrm{e}^{a} - u(\,\mathrm{e}^{a} - \,\mathrm{e}^{-a}) = \,\mathrm{e}^{a\cos(y)}. \tag{SA.4}$$

On the other hand, differentiating both sides of (SA.3) with respect to u, we have

$$a[-\sin(y)]\frac{\mathrm{d}y}{\mathrm{d}u} = \frac{-(\,\mathrm{e}^{a} - \,\mathrm{e}^{-a})}{\mathrm{e}^{a} - u(\,\mathrm{e}^{a} - \,\mathrm{e}^{-a})},$$

which results in

$$\begin{aligned}\frac{\mathrm{d}u}{\mathrm{d}y} &= \frac{a\sin(y)[\,\mathrm{e}^{a} - u(\,\mathrm{e}^{a} - \,\mathrm{e}^{-a})]}{\mathrm{e}^{a} - \,\mathrm{e}^{-a}} \qquad \text{by (SA.4)}\\ &= \frac{a\sin(y)\,\mathrm{e}^{a\cos(y)}}{2\sinh(a)}.\end{aligned}$$

Therefore, the density of Y is

$$f(y) = h(u)\left|\frac{\mathrm{d}u}{\mathrm{d}y}\right| = \frac{a\sin(y)\,\mathrm{e}^{a\cos(y)}}{2\sinh(a)}.$$

In addition, when $u=0$, from (SA.3), we have $a\cos(y)=a$, i.e., $y=0$. When $u=1$, from (SA.3), we have $a\cos(y)=-a$, i.e., $y=\pi$. Thus, $0<y<\pi$.

1.9 Proof. Note that when $a<y<b$, the cdf of Y is

$$G(y) = \frac{F(y)-F(a)}{F(b)-F(a)}.$$

Thus, $G(Y) \stackrel{\mathrm{d}}{=} U \sim U(0,1)$ implies

$$F(Y) \stackrel{\mathrm{d}}{=} F(a) + U[F(b) - F(a)],$$

i.e.,

$$Y \stackrel{\mathrm{d}}{=} F^{-1}(F(a) + U[F(b) - F(a)]).$$

1.10 Solution: To derive the marginal distribution of $(X_1, \ldots, X_i)$, we need to use the following integral identity:

$$\int_{\mathbb{R}^m} h\left(\sum_{i=1}^m x_i^2\right) \mathrm{d}x_1 \cdots \mathrm{d}x_m = \frac{\pi^{m/2}}{\Gamma(m/2)} \int_0^\infty y^{m/2-1} h(y)\, \mathrm{d}y, \quad \text{(SA.5)}$$

where $h(\cdot)$ is an arbitrary non-negative measurable function. Let

$$h(y) = 1/(\Delta_i + y)^b, \quad \text{where} \quad \Delta_i \,\hat{=}\, 1 + x_1^2 + \cdots + x_i^2,$$

then the marginal density of $(X_1, \ldots, X_i)$ is given by

$$\begin{aligned}
& f(x_1, \ldots, x_i) \\
= & \int_{\mathbb{R}^{d-i}} f(x_1, \ldots, x_i, x_{i+1}, \ldots, x_d)\, \mathrm{d}x_{i+1} \cdots \mathrm{d}x_d \\
= & \frac{\Gamma(b)}{\pi^b} \int_{\mathbb{R}^{d-i}} \frac{1}{(\Delta_i + \sum_{j=i+1}^d x_j^2)^b}\, \mathrm{d}x_{i+1} \cdots \mathrm{d}x_d \\
\stackrel{\text{(SA.5)}}{=} & \frac{\Gamma(b)}{\pi^b} \cdot \frac{\pi^{(d-i)/2}}{\Gamma((d-i)/2)} \int_0^\infty y^{(d-i)/2-1} h(y)\, \mathrm{d}y \\
= & c \int_0^\infty \frac{y^{(d-i)/2-1}}{(\Delta_i + y)^b}\, \mathrm{d}y \\
= & \frac{c}{\Delta_i^b} \int_0^\infty \frac{y^{(d-i)/2-1}}{(1 + y/\Delta_i)^b}\, \mathrm{d}y \qquad [\text{let } z = y/\Delta_i] \\
= & \frac{c}{\Delta_i^{b-(d-i)/2}} \int_0^\infty \frac{z^{(d-i)/2-1}}{(1+z)^b}\, \mathrm{d}z \qquad [\text{let } w = 1/(1+z)]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{c}{\Delta_i^{b-(d-i)/2}} \int_0^1 w^{b-(d-i)/2-1}(1-w)^{(d-i)/2-1}\,\mathrm{d}w \\
&= \frac{c}{\Delta_i^{b-(d-i)/2}} \cdot B(b-(d-i)/2, (d-i)/2) \\
&= \frac{\Gamma(\frac{i+1}{2})}{[\pi(1+x_1^2+\cdots+x_i^2)]^{\frac{i+1}{2}}}, \qquad \text{(SA.6)}
\end{aligned}
$$

i.e., $(X_1, \ldots, X_i)$ follow the i-dimensional Cauchy distribution. Especially, X_1 follows the standard Cauchy distribution, denoted by $X_1 \sim \text{Cauchy}(1)$. By symmetry, we have $X_i \sim \text{Cauchy}(1)$ for $i = 1, \ldots, d$.

The conditional distribution of X_i given $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}$ is then given by

$$
\begin{aligned}
f_i(x_i|x_1, \ldots, x_{i-1}) &= \frac{f(x_1, \ldots, x_i)}{f(x_1, \ldots, x_{i-1})} \\
&\overset{\text{(SA.6)}}{=} \frac{\dfrac{\Gamma(\frac{i+1}{2})}{[\pi(1+x_1^2+\cdots+x_i^2)]^{\frac{i+1}{2}}}}{\dfrac{\Gamma(\frac{i}{2})}{[\pi(1+x_1^2+\cdots+x_{i-1}^2)]^{\frac{i}{2}}}} \\
&= \frac{\Gamma(\frac{i+1}{2})}{\Gamma(\frac{i}{2})\sqrt{\pi}} \cdot \frac{\Delta_{i-1}^{i/2}}{(\Delta_{i-1}+x_i^2)^{(i+1)/2}}.
\end{aligned}
$$

Making transformation $y_i = x_i\sqrt{i/\Delta_{i-1}}$, then

$$
f_i(y_i|x_1, \ldots, x_{i-1}) = \frac{\Gamma(\frac{i+1}{2})}{\Gamma(\frac{i}{2})\sqrt{i\pi}} \cdot \left(1 + \frac{y_i^2}{i}\right)^{-(i+1)/2}, \qquad \text{(SA.7)}
$$

which is the density of the t-distribution with i degrees of freedom. Let $T_i \sim t(i)$, then (SA.7) implies

$$
X_i\sqrt{i/\Delta_{i-1}}\Big|(x_1, \ldots, x_{i-1}) \overset{\mathrm{d}}{=} T_i,
$$

or

$$
X_i|(X_1, \ldots, X_{i-1}) \overset{\mathrm{d}}{=} T_i\sqrt{(1+\textstyle\sum_{j=1}^{i-1} X_j^2)/i}, \quad i = 2, \ldots, d.
$$

Therefore, the conditional sampling method can be used to generate a d-dimensional Cauchy distribution as follows:

- Draw $X_1 \sim \text{Cauchy}(1)$, which is a special case of Example 1.4 with $\mu = 0$ and $\sigma = 1$;
- Draw $T_i \sim t(i)$ and set $X_i = T_i\sqrt{(1 + \sum_{j=1}^{i-1} X_j^2)/i}, \quad i = 2, \ldots, d.$

1.11 Solution: We make the following transformation

$$y_i = \frac{z_i}{z}, \quad i = 1, \ldots, d-1, \quad \text{and} \quad z = z_1 + \cdots + z_d.$$

Its inverse transformation is given by

$$z_i = y_i z, \quad i = 1, \ldots, d-1, \quad \text{and} \quad z_d = (1 - y_1 - \cdots - y_{d-1})z.$$

The Jacobian determinant is

$$\begin{aligned}
J(\boldsymbol{z} \to \boldsymbol{y}_{-d}, z) &= \frac{\partial(z_1, \ldots, z_{d-1}, z_d)}{\partial(y_1, \ldots, y_{d-1}, z)} \\
&= \det \begin{pmatrix} z & 0 & \cdots & 0 & y_1 \\ 0 & z & \cdots & 0 & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & z & y_{d-1} \\ -z & -z & \cdots & -z & y_d \end{pmatrix} \\
&= z^{d-1}.
\end{aligned}$$

Therefore, the joint density of $(Y_1, \ldots, Y_{d-1}, Z)^\top$ is

$$\begin{aligned}
& f(y_1, \ldots, y_{d-1}, z) \\
= \;& f(z_1, \ldots, z_{d-1}, z_d) \cdot |J(\boldsymbol{z} \to \boldsymbol{y}_{-d}, z)| = \left\{ \prod_{i=1}^{d} \mathrm{e}^{-z_i} \right\} \cdot z^{d-1} \\
= \;& \Gamma(d) \cdot \frac{1}{\Gamma(d)} z^{d-1}\, \mathrm{e}^{-z} = f(y_1, \ldots, y_{d-1}) \cdot f(z),
\end{aligned}$$

which implies that

$$(Y_1, \ldots, Y_{d-1})^{\top} \sim U(\mathbb{V}_{d-1}), \quad Z \sim \text{Gamma}(d, 1),$$

and $(Y_1, \ldots, Y_{d-1})^{\top} \perp\!\!\!\perp Z$, where $\mathbb{V}_{d-1}$ is defined by (1.24). In addition, the volume of $\mathbb{V}_{d-1}$ is $1/\Gamma(d) = 1/(d-1)!$. Finally, $(Y_1, \ldots, Y_{d-1})^{\top} \sim U(\mathbb{V}_{d-1})$ if and only if $(Y_1, \ldots, Y_d)^{\top} \sim U(\mathbb{T}_d)$, where $Y_d = 1 - \sum_{i=1}^{d-1} Y_i$.

1.12 Solution: (a) We only need to prove that the distribution of $Y = ZX$ is given by (1.36), where $Z \sim \text{Bernoulli}\,(1-\phi)$, $X \sim \text{Poisson}\,(\lambda)$ and $Z \perp\!\!\!\perp X$. In fact, we have

$$\Pr(Y=0) = \Pr(Z=0) + \Pr(Z=1, X=0) = \phi + (1-\phi)\,\mathrm{e}^{-\lambda}$$

by independence and

$$\Pr(Y=y) = \Pr(Z=1, X=y) = (1-\phi)\,\frac{\mathrm{e}^{-\lambda}\lambda^y}{y!}, \quad y > 0.$$

(b) Let $X \sim \text{Poisson}(\lambda)$ and its pmf be denoted by $\text{Poisson}(x|\lambda)$. When $\lambda = 0$, we obtain $E(X) = \text{Var}(X) = \lambda = 0$, so that $X \sim \text{Degenerate}(0)$. In other words, $\text{Poisson}(0) = \text{Degenerate}(0)$ or

$$\text{Poisson}(x|0) = \text{Poisson}(0|0) = \Pr(X=0) = 1.$$

The joint distribution of W and Y is

$$\Pr(W=w, Y=y) = \Pr(W=w) \cdot \Pr(Y=y|W=w).$$

Therefore, the marginal distribution of Y is

$$\begin{aligned}
\Pr(Y=y) &= \sum_{w=0}^{1} \Pr(W=w, Y=y) \\
&= \sum_{w=0}^{1} \Pr(W=w) \cdot \Pr(Y=y|W=w)
\end{aligned}$$

$$\begin{aligned} &= \phi \Pr(Y=y|W=0) + (1-\phi)\Pr(Y=y|W=1) \\ &= \phi \cdot \text{Poisson}(y|0) + (1-\phi)\text{Poisson}(y|\lambda) \\ &= \phi \cdot I(y=0) + (1-\phi)\text{Poisson}(y|\lambda), \end{aligned}$$

which is the same as (1.36).

1.13 Solution: (a) From (1.38), i.e.,

$$X = \begin{cases} X_1, & \text{with probability } \phi, \\ X_2, & \text{with probability } 1-\phi, \end{cases}$$

we know that the density of X is given by

$$f_X(x) = \phi f_{X_1}(x) + (1-\phi) f_{X_2}(x). \tag{SA.8}$$

On the other hand, from (1.38), we have

$$X - X_1 = \begin{cases} 0, & \text{with probability } \phi, \\ X_2 - X_1, & \text{with probability } 1-\phi. \end{cases} \tag{SA.9}$$

By comparing (SA.9) with (1.37), we have

$$X - X_1 \stackrel{\mathrm{d}}{=} Z(X_2 - X_1), \tag{SA.10}$$

where $Z \sim \text{Bernoulli}\,(1-\phi)$ and $Z \perp\!\!\!\perp (X_2 - X_1)$. In general, we cannot add X_1 on both sides of (SA.10). But, we guess

$$X \stackrel{\mathrm{d}}{=} X_1 + Z(X_2 - X_1) = (1-Z)X_1 + ZX_2, \tag{SA.11}$$

where $Z \sim \text{Bernoulli}\,(1-\phi)$ and $Z \perp\!\!\!\perp \{X_1, X_2\}$. To verify the correctness of (SA.11), we need to show that the pdf of X defined by (SA.11) is identical to (SA.8). In fact, the cdf of X is

$$\Pr(X \leqslant x) = \Pr\{(1-Z)X_1 + ZX_2 \leqslant x\}$$

$$
\begin{aligned}
&= \sum_{z=0}^{1} \Pr(Z=z)\cdot \Pr\{(1-Z)X_1 + ZX_2 \leqslant x | Z=z\} \\
&= \Pr(Z=0)\cdot \Pr(X_1 \leqslant x | Z=0) \\
&\quad + \Pr(Z=1)\cdot \Pr(X_2 \leqslant x | Z=1) \\
&= \phi \Pr(X_1 \leqslant x) + (1-\phi)\Pr(X_2 \leqslant x),
\end{aligned}
$$

so that the pdf of X is

$$f_X(x) = \phi f_{X_1}(x) + (1-\phi) f_{X_2}(x),$$

which is identical to (SA.8).

(b) An SR of X defined by (1.39) is

$$X \stackrel{\mathrm{d}}{=} Z_1 X_1 + \cdots + Z_n X_n, \tag{SA.12}$$

where $\mathbf{z} = (Z_1, \ldots, Z_n)^\top \sim \text{Multinomial}(1; \phi_1, \ldots, \phi_n)$, and

$$\mathbf{z} \perp\!\!\!\perp \{X_1, \ldots, X_n\}.$$

The corresponding pdf of X is

$$f_X(x) = \phi_1 f_{X_1}(x) + \cdots + \phi_n f_{X_n}(x). \tag{SA.13}$$

1.14 Solution: Method I: From mixture representation to SR. For $0 \leqslant x \leqslant 2$, we can write

$$
\begin{aligned}
f_X(x) &= \frac{5}{12}\Big[1 + (x-1)^4\Big] = \frac{5}{6}\times\frac{1}{2} + \frac{1}{6}\times\frac{5}{2}(x-1)^4 \\
&= \phi f_{X_1}(x) + (1-\phi) f_{X_2}(x),
\end{aligned}
$$

where $\phi = 5/6$, $X_1 \sim U[0,2]$ with pdf

$$f_{X_1}(x) = \frac{1}{2}\cdot I(0 \leqslant x \leqslant 2) \quad \text{and} \quad f_{X_2}(x) = \frac{5}{2}(x-1)^4 \cdot I(0 \leqslant x \leqslant 2).$$

On the one hand, we have $X_1/2 \stackrel{\mathrm{d}}{=} U_1 \sim U[0,1]$, so that $X_1 \stackrel{\mathrm{d}}{=} 2U_1$. On the other hand, the cdf of X_2 is

$$F(x_2) = \int_0^{x_2} \frac{5}{2}(x-1)^4 \mathrm{d}x = 0.5[(x_2-1)^5+1], \quad x_2 \in [0,2].$$

Since $F(X_2) \stackrel{\mathrm{d}}{=} U_2 \sim U[0,1]$, we have $X_2 \stackrel{\mathrm{d}}{=} (2U_2-1)^{1/5}+1$.

From (SA.11), we have

$$X \stackrel{\mathrm{d}}{=} (1-Z)X_1 + ZX_2,$$

where $Z \sim \text{Bernoulli}\,(1-\phi)$ and $Z \perp\!\!\!\perp \{X_1, X_2\}$. We summarize the algorithm as follows.

— Step 1: Draw $U_1 = u_1, U_2 = u_2 \stackrel{\text{iid}}{\sim} U[0,1]$ and independently draw $Z = z \sim \text{Bernoulli}\,(1/6)$;

— Step 2: Let $x_1 = 2u_1$ and $x_2 = (2u_2-1)^{1/5}+1$;

— Step 3: Return $x = (1-z)x_1 + zx_2$.

<u>Method II: The grid method</u>. To generate X, we first select a set of appropriate grid points $\{x_i\}_{i=1}^d$ with $x_i = 2i/d$ and $d = 100$, that cover the support $[0,2]$, and then approximate the pdf $f_X(x)$ by a discrete distribution at $\{x_i\}_{i=1}^d$ with probabilities

$$p_i = \frac{f_X(x_i)}{\sum_{j=1}^d f_X(x_j)}, \quad i = 1, \ldots, d.$$

In other words, we have $X \dot{\sim} \text{FDiscrete}_d(\{x_i\}, \{p_i\})$.

1.15 Solution: (i) The joint density of X and Y is

$$\begin{aligned} f_{(X,Y)}(x,y) &= ny^{-n}\,\mathrm{e}^{-xy} \cdot I(x>0, y \geqslant 1) \\ &= \frac{nI(y \geqslant 1)}{y^{n+1}} \cdot y\,\mathrm{e}^{-yx} I(x>0) \\ &= f_Y(y) \cdot f_{(X|Y)}(x|y), \end{aligned}$$

so that $Y \sim \text{Pareto}(n, 1)$ with pdf

$$f_Y(y) = \frac{nI(y \geqslant 1)}{y^{n+1}},$$

and $X|(Y = y) \sim \text{Exponential}(y) = \text{Gamma}(1, y)$. We have

$$yX|(Y = y) \sim \text{Gamma}(1, 1) = \text{Exponential}(1),$$

which is independent of y, so $YX \stackrel{\text{d}}{=} W \sim \text{Exponential}(1)$ and $W \perp\!\!\!\perp Y$. We obtain $X \stackrel{\text{d}}{=} Y^{-1}W$.

(ii) The mixture representation method for generating $X \sim f_X(x)$ is as follows:

Step 1: Draw $Y = y \sim \text{Pareto}(n, 1)$ and independently draw $W = w \sim \text{Exponential}(1)$;

Step 2: Return $x = y^{-1} \cdot w$.

From Exercise 1.5(d) and Example 1.1 in Lecture Notes, we know that Step 1 is equivalent to

Step 1′: Draw $U_1 = u_1, U_2 = u_2 \stackrel{\text{iid}}{\sim} U(0, 1)$, and set $y = u_1^{-1/n}$ and $w = -\log(u_2)$.