$$h(x) = \frac{f(x)}{Mg(x)} = \frac{cx^{\alpha-1}(1-x)^{\alpha-1}}{(1/2)^{\alpha-1}cx^{\alpha-1}} = (2(1-x))^{\alpha-1}.$$

As predicted, c drops out and its value is not needed.

The algorithm for $\mathrm{Be}(\alpha, \alpha)$ is then:

1. generate $U \sim U(0,1)$ and put $Y = (1/2)U^{1/\alpha}$;
2. as the side experiment, if $R \sim U(0,1) < h(Y)$ then accept Y and go to step 3; else return to step 1;
3. again sample $U \sim U(0,1)$; if $U < 1/2$ return $X = Y$; otherwise return $X = 1 - Y$.

2.8 Composition Example: The Gamma distribution

Suppose we have n independent random variables X_i with probability density functions f_i together with n probabilities $p_i > 0$, $\sum_i p_i = 1$. Construct the random variable X as follows:

$$\begin{aligned}&\text{with probability } p_i\text{, select index } i \in \{1, 2, \ldots, n\}\\&\text{return a sample } X_i \text{ from density } f_i.\end{aligned} \tag{2.24}$$

We refer to f as the *composition* of the f_i.

By letting the roulette wheel random variable M denote the index selection, and the function $\mathbb{1}_k(M)$ be equal to 1 if $M = k$ and 0 otherwise, we may write X as

$$X = \sum_{i=1}^{n} X_i \mathbb{1}_i(M). \tag{2.25}$$

The pdf for f is

$$f(x) = p_1 f_1(x) + p_2 f_2(x) + \cdots + p_n f_n(x). \tag{2.26}$$

This is seen as follows, for an event A

$$\begin{aligned}\int_A f(y)\,dy = \Pr(X \in A) &= \sum_{i=1}^{n} \Pr(X \in A \,|\, i \text{ is selected})\Pr(i \text{ is selected})\\ &= \sum_{i=1}^{n} p_i \int_A f_i(y)\,dy = \int_A \left(\sum_{i=1}^{n} p_i f_i(y)\right) dy.\end{aligned}$$

For the composite mean we have

$$\begin{aligned} E(X) &= \int y f(y)\,dy = \int y \sum_i p_i f_i(y)\,dy = \sum_i p_i \int y f_i(y)\,dy \\ &= \sum_i p_i \mu_i. \end{aligned} \tag{2.27}$$

The formula for the variance is a little more complicated,

$$\operatorname{var}(X) = \sum p_i \operatorname{var}(X_i) + \sum_{i \neq j} p_i p_j \Big(\mu_i - \mu_j\Big)^2. \tag{2.28}$$

First we show $E(X^2) = \sum p_i E(X_i^2)$;

$$\begin{aligned} E(X^2) &= \int y^2 f(y)\,dy = \int y^2 \Big(\sum p_i f_i(y)\Big)\,dy \\ &= \sum p_i \int y^2 f_i(y)\,dy = \sum p_i E(X_i^2). \end{aligned} \tag{2.29}$$

To finish we do the $n = 2$ case, the full derivation is given in Fig. 2.8,

$$\begin{aligned} \operatorname{var}(X) &= E(X^2) - E(X)^2 = \sum p_i E(X_i^2) - \Big(\sum p_i \mu_i\Big)^2 \\ &= \sum p_i \Big(\operatorname{var}(X_i) + E(X_i)^2\Big) - \Big(\sum p_i \mu_i\Big)^2 \\ &= \sum p_i \operatorname{var}(X_i) + p_1 \mu_1^2 + p_2 \mu_2^2 - p_1^2 \mu_1^2 - 2 p_1 p_2 \mu_1 \mu_2 - p_2^2 \mu_2^2 \\ &= \sum p_i \operatorname{var}(X_i) + p_1 (1 - p_1) \mu_1^2 - 2 p_1 p_2 \mu_1 \mu_2 + p_2 (1 - p_2) \mu_2^2 \\ &= \sum p_i \operatorname{var}(X_i) + p_1 p_2 (\mu_1^2 - 2 \mu_1 \mu_2 + \mu_2^2) \\ &= p_1 \operatorname{var}(X_1) + p_2 \operatorname{var}(X_2) + p_1 p_2 (\mu_1 - \mu_2)^2. \end{aligned} \tag{2.30}$$

Sampling from a composite density goes just as described in (2.24). An index may be selected, for example, via discrete cdf inversion or the alias method. Having chosen i, sample from f_i by an appropriate method for that density and return the sampled value.

Composition can give rise to a great variety of probability distributions. One class of examples is the sum of piecewise constant functions such as a histogram. In this case the f_i are just uniform densities and are easily sampled. Any density can be approximated in this way. A better approximation is to use piecewise trapezoidal regions. In this case the f_i are appropriately scaled linear functions and these too are easily sampled.

Another way in which composition can be exploited for sampling purposes is by tailoring a piecewise envelope or proposal density g

From (2.29) we can write

$$\begin{aligned}
\text{var}(X) &= E(X^2) - E(X)^2 = \sum p_i E(X_i^2) - \left(\sum p_i \mu_i\right)^2 \\
&= \sum p_i \text{var}(X_i) + \sum_i p_i E(X_i)^2 - \sum_i p_i^2 \mu_i^2 - 2\sum_{i \neq j} p_i p_j \mu_i \mu_j \\
&= \sum p_i \text{var}(X_i) + \sum_i p_i (1 - p_i)\mu_i^2 - 2\sum_{i \neq j} p_i p_j \mu_i \mu_j \\
&= \sum p_i \text{var}(X_i) + \sum_i \sum_{j \neq i} p_i p_j \mu_i^2 - 2\sum_{i \neq j} p_i p_j \mu_i \mu_j \\
&= \sum p_i \text{var}(X_i) + \sum_i \sum_{j > i} p_i p_j \mu_i^2 + \sum_j \sum_{j < i} p_i p_j \mu_j^2 - 2\sum_{i \neq j} p_i p_j \mu_i \mu_j \\
&= \sum p_i \text{var}(X_i) + \sum_{i \neq j} p_i p_j \left(\mu_i^2 + \mu_j^2\right) - 2\sum_{i \neq j} p_i p_j \mu_i \mu_j \\
&= \sum p_i \text{var}(X_i) + \sum_{i \neq j} p_i p_j \left(\mu_i - \mu_j\right)^2 . \qquad (2.31)
\end{aligned}$$

for use in the rejection method. This gives rise to an exact sampling method. The combination of composition and rejection provides a powerful tool for many distributions that cannot be treated in any other way.

We already saw a restricted example of this approach in the last section, restricted in the sense that by symmetry the two components were essentially the same distribution. Here we show how the composition/rejection technique can provide a solution for treating the gamma distribution. This is another important distribution that arises in conjunction with the Poisson process.

2.8.1 The Gamma Distribution

In the Poisson process with event rate λ we saw that the exponential distribution is the waiting time until the first event. By way of generalization, the *gamma distribution* with parameter α is the waiting time W until the αth event occurs. Using (2.9) we obtain the gamma cdf as follows:

$$\begin{aligned}
F(w) &= \Pr(W \leq w) = 1 - \Pr(W > w) \\
&= 1 - \Pr(\text{fewer than } \alpha \text{ events occur in } [0, w])
\end{aligned}$$