MTH 511a - 2020: Lecture 5

Instructor: Dootika Vats

1 The Composition method

We have now learned two algorithms for sampling from a discrete distribution: the inverse transform method and the accept-reject algorithm. The inverse transform method can be used for *any* distribution and the accept-reject can be efficient if used properly.

For certain special distributions, it is easier to use a *composition method* for sampling.

Suppose we have an efficient way of simulating random variables from two pmfs $\{p_j^{(1)}\}$ and $\{p_j^{(2)}\}$, and we want to simulate from

$$\Pr(X = j) = \alpha p_j^{(1)} + (1-\alpha)p_j^{(2)} \quad j \geq 0 \text{ where } 0 < \alpha < 1\,.$$

First you should note that the above *composition pmf* is a valid pmf since $\sum_j \Pr(X = j) = 1$. How would we sample in such a situation?

Let $X_1 \sim P^{(1)}$ and $X_2 \sim P^{(2)}$. Set

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

Algorithm 1 Composition method

1: Draw $U \sim U[0,1]$

2: **if** $U < \alpha$ **then** simulate $X_1 \sim P^{(1)}$ **else** simulate X_2 and stop

Proof. We will show that $\Pr(X = j)$ is what is desired. Consider

$$\begin{aligned}
&\Pr(X = j)\\
&= \Pr(X = j, U < \alpha) + \Pr(X = j, \alpha \le U < 1)\\
&= \Pr(X_1 = j, U < \alpha) + \Pr(X_2 = j, \alpha \le U < 1) \qquad \text{(by law of total probability)}\\
&= \Pr(X_1 = j)\Pr(U < \alpha) + \Pr(X_2 = j)\Pr(\alpha \le U < 1) \qquad \text{(by independence of } U \text{ and } X_1, X_2\text{)}\\
&= \alpha p_j^{(1)} + (1-\alpha)p_j^{(2)}\,.
\end{aligned}$$

□

We can set this up more generally for n different distributions. In general, $F_i, i = 1, \dots, n$ are distribution functions, and α_i are such that $0 < \alpha_i < 1$ for all i and $\sum_i \alpha_i = 1$. The composition (or mixture) distribution is

$$F(x) = \sum_{i=1}^{n} \alpha_i F_i(x)\,.$$

Let $X_i \sim F_i$. To simulate from the composition F, set

$$X = \begin{cases} X_1 & \text{with probability } \alpha_1 \\ X_2 & \text{with probability } \alpha_2 \\ \vdots & \\ X_n & \text{with probability } \alpha_n \end{cases}.$$

Example 1 (Zero inflated Poisson distribution)*.* A Poisson(λ) distribution usually has a small mass at 0. But sometimes, we need a counting distribution with large mass at 0. For example, consider the random variable X being the number of COVID-19 patients tested positive every hour. Many hours of the day this number may be 0, and then this number can be quite high for some hours.

In such a case, we may use the *zero inflated Poisson distribution* (ZIP). Recall that if $X \sim$ Poisson(λ)

$$\Pr(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} \quad k = 0, 1, \dots\,.$$

If $X \sim$ ZIP(π, λ)

$$\Pr(X = k) = \begin{cases} \pi + (1-\pi)e^{-\lambda} & \text{if } k = 0 \\ (1-\pi)e^{-\lambda}\dfrac{\lambda^k}{k!} & \text{if } k = \{1, 2, \dots\} \end{cases}.$$

Note that the mean of a ZIP is $(1-\pi)\lambda < \lambda$ since more mass is given at 0.

We will use the composition method to sample from the ZIP distribution. To sample from a ZIP, first $p_j^{(1)}$ be defined as

$$\Pr(X_1 = 0) = 1 \quad \text{and} \quad \Pr(X_1 \neq 0) = 0\,,$$

and let $X_2 \sim \text{Poisson}(\lambda)$. Define the pmf:

$$\Pr(X = k) = \pi p_k^{(1)} + (1-\pi) p_k^{(2)}\,.$$

Then $X \sim \text{ZIP}(\pi, \lambda)$. To see this, plug in $k = 0$ and $k = 1, 2, \ldots$ above:

Algorithm 2 Zero inflated Poisson distribution

1: Draw $U \sim U[0, 1]$

2: **if** $U < \pi$ **then** $X = 0$ **else** simulate $X \sim \text{Poisson}(\lambda)$

2 Bernoulli factories

We have learned how to sample from a Bernoulli distribution. In this section, we learn some tools to draw from a Bernoulli($f(p)$) for some specific function f using only Bernoulli(p) draws.

Suppose you have $X_1, X_2, \ldots \overset{iid}{\sim} \text{Bern}(p)$. Now suppose we wish to construct a Bernoulli random variable with a parameter that is a function of p, $f(p)$. That is, we want to simulate $Y \sim \text{Bern}(f(p))$. This process is called a Bernoulli factory.

There is no universal algorithm for all $f(p)$, but we can construct one on a case by case basis.

Example 2. Suppose we can simulate $X \sim \text{Bern}(p)$. Can we simulate a Bernoulli random variable with success probability

$$f(p) = p^2(1-p)\,?$$

We are free to draw as many samples as we want from Bern(p).

So if we draw three indepenent samples from Bern(p) and look at the event: $\{X_1 = 1, X_2 = 1, X_3 = 0\}$. Then

$$\Pr\left(X_1 = 1, X_2 = 1, X_3 = 0\right) = \Pr(X_1 = 1)\Pr(X_2 = 1)\Pr(X_1 = 0) = p^2(1-p)\,.$$

Thus, the following algorithm returns 1 with probability $f(p) = p^2(1-p)$.

Algorithm 3 Bernoulli factory for $f(p) = p^2(1-p)$

Draw $X_1 \sim \text{Bern}(p)$
if $X_1 = 0$ **then**
    set $X = 0$, stop
Simulate $X_2 \sim \text{Bern}(p)$.
if $X_2 = 0$ **then**
    set $X = 0$, stop
Simulate $X_3 \sim \text{Bern}(p)$.
if $X_3 = 1$ **then**
    $X = 0$, stop
Set $X = 1$.

The above returns $X = 1$ with probability $p^2(1-p)$. There is another even simpler method.

Note that for $X_1, X_2, X_3 \overset{iid}{\sim} \text{Bern}(p)$. Consider $X = X_1X_2(1-X_3)$, then

$$\Pr[X_1X_2(1-X_3) = 1] = \Pr[X_1 = 1]\Pr[X_2 = 1]\Pr[1 - X_3 = 1] = p^2(1-p)\,,$$

where the decomposition is because the only way $X_1X_2(1-X_3) = 1$ is if all three terms are equal to 1.

Algorithm 4 Another Bernoulli factory for $f(p) = p^2(1-p)$

Draw $X_1, X_2, X_3 \overset{iid}{\sim} \text{Bern}(p)$
Return $X = X_1X_2(1-X_3)$

Similarly other polynomials of p can be considered.

3 Questions to think about

1. Can you construct a similar zero inflated Binomial distribution? How would you sample from it?
2. Try setting up a Bernoulli factory for $p^5(1-p)^2$.
3. I claim Algorithm 3 is better than Algorithm 4. Why?