Note: This document is a part of the lectures given during the Jan-May 2020 Semester.

Rayleigh Distribution:

We consider the Rayleigh Distribution:

$$F(x) = 1 - e^{-2x(x-b)} \ , \ x \geq b.$$

Solving the equation $F(x) = u \ , \ u \in (0,1)$, results in a quadratic with roots:

$$x = \frac{b}{2} \pm \frac{\sqrt{b^2 - 2\log(1-u)}}{2}.$$

The inverse is given by the larger of the two roots. Replacing U with $(1-U)$ we get,

$$x = \frac{b}{2} + \frac{\sqrt{b^2 - 2\log(U)}}{2}.$$

Note that even if the inverse of F is not known explicitly, the inverse transform method is still applicable through numerical evaluation of F^{-1}. Computing $F^{-1}(u)$ is equivalent to finding a root x of the equation $F(x) - u = 0$. For a distribution F with density f, Newton's method for finding roots produces a sequence of iterates:

$$x_{n+1} = x_n - \frac{F(x_n) - u}{f(x_n)},$$

given a starting point x_0.

Discrete Distribution:

In the case of a discrete distribution, the evaluation of F^{-1} reduces to a table look up. Consider, for example, a discrete random variable whose possible values are $c_1 < c_2 < \cdots < c_n$. Let p_i be the probability attached to $c_i \ , \ i = 1, 2 \ldots, n$ and set $q_0 = 0$. Also,

$$q_i = \sum_{j=1}^{i} p_j \ , \ i = 1, 2, \ldots, n.$$

These are cumulative probabilities associated with the c_i, that is, $q_i = F(c_i) \ , \ i = 1, 2, \ldots, n$. To sample from this distribution:

1. Generate a uniform $U \sim \mathcal{U}[0,1]$.
2. Find $K \in \{1, 2, \ldots, n\}$ such that $q_{K-1} < U \leq q_K$.
3. Set $X = c_K$.

Conditional Distribution:

Suppose X has distribution F and consider the problem of sampling X conditional on $a < X \leq b$ with $F(a) < F(b)$. Using the inverse transform method, this is no more difficult than generating X unconditionally. If $U \sim \mathcal{U}[0,1]$, then the random variable defined by $V = F(a) + [F(b) - F(a)]U$ is uniformly distributed between $F(a)$ and $F(b)$ and $F^{-1}(V)$ has the desired conditional distribution. To see this observe that,

$$\begin{aligned} P(F^{-1}(V) \leq x) &= P(F(a) + [F(b) - F(a)]U \leq F(x)) \\ &= P(U \leq [F(x) - F(a)]/[F(b) - F(a)]) \\ &= \frac{F(x) - F(a)}{F(b) - F(a)}. \end{aligned}$$

This is precisely the distribution of X given $a < X \leq b$.

<u>Acceptance Rejection Method:</u>

Introduced by Von-Neumann, this method is among the most widely applicable mechanism for generating random samples. This method generates samples from a target distribution by first generating candidates from a more convenient distribution and then rejecting a random subset of the generated candidates. The rejection mechanism is designed so that the accepted samples are indeed distributed according to the target distribution. This technique is by no means restricted to univariate distribution.

Suppose we wish to generate samples from a density f defined on some set $\mathcal{X}$. This could be a subset of the real line, of $\mathbb{R}^d$ or a more general set.

Let g be a density on $\mathcal{X}$ from which we know how to generate samples and with the property that:

$$f(x) \leq cg(x) \ , \ \forall\, x \in \mathcal{X}.$$

In the acceptance rejection method, we generate a sample X from g and accept the sample with probability $f(X)/cg(X)$. This can be implemented by sampling U uniformly over $(0, 1)$. If X is rejected, a new candidate is sampled from g and the acceptance test applied again. The process repeats until the acceptance test is passed and the accepted value is returned as a sample from f.

<u>Algorithm:</u>

1. Generate X from distribution g.
2. Generate U from $\mathcal{U}[0, 1]$.
3. If $U \leq f(X)/cg(X)$, return X, otherwise go to step 1.

To verify the validity of the acceptance rejection method, let Y be the sample returned by the algorithm and observe that Y has the distribution of X conditional on $U \leq f(X)/cg(X)$. Thus for any $A \subseteq \mathcal{X}$,

$$\begin{aligned} P(Y \in A) &= P(X \in A | U \leq f(X)/cg(X)) \\ &= \frac{P(X \in A, U \leq f(X)/cg(X))}{P(U \leq f(X)/cg(X))}. \end{aligned}$$

Given X, the probability that $U \leq f(X)/cg(X)$ is simply $f(X)/cg(X)$, because U is uniform and hence the denominator is given by,

$$P(U \leq f(X)/cg(X)) = \int_{\mathcal{X}} \frac{f(x)}{cg(x)} g(x)dx = \frac{1}{c}.$$

Making a substitution we find that,

$$P(Y \in A) = cP(X \in A, U \leq f(X)/cg(X)) = c\int_{A} \frac{f(x)}{cg(x)} g(x)dx = \int_{A} f(x)dx.$$

Since A is arbitrary it verifies that Y has density f.