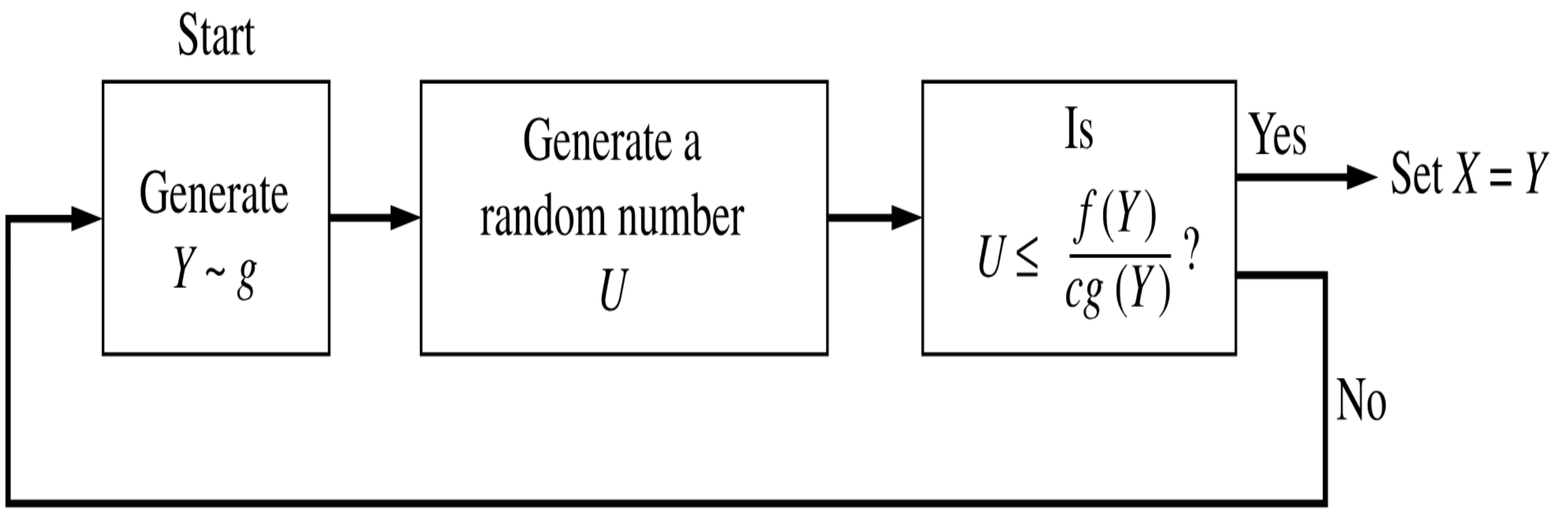
RECAP

- Inverse transform method
- $F_X(X)$ is Uniform!
- Importance Sampling

Accept Reject method

- Suppose you want to generate samples from X with pmf $p(\cdot)$ using samples from Y with pmf $q(\cdot)$.
- Suppose that $\frac{p(y)}{q(y)} \leq c$ for all y.
- The accept reject method is as follows:
- Step 1: Generate a sample $y \sim q(\cdot)$.
- Step 2: Generate $u \sim \mathcal{U}(0, 1)$.
- Step 3: If $u \leq \frac{p(y)}{cq(y)}$, accept y as a sample from X.
- Step 4: If not, reject y and go back to Step 1.

Accept Reject method



- Why does the method work ?
- What is $P(y/accept)$? is it $p(y)$?

Proof of Accept-Reject Method

- To prove that the method produces samples from $p(\cdot)$, we will compute the probability of accepting a sample y from $q(\cdot)$.
- The probability of accepting y is given by:

  $$P(\text{accept} \mid y) = P\left(u \leq \frac{p(y)}{cq(y)}\right) = \frac{p(y)}{cq(y)}$$

  since $u \sim \mathcal{U}(0, 1)$.
- Thus, the joint probability of sampling $y \sim q(\cdot)$ and accepting it is:

  $$P(\text{sample } y \text{ and accept}) = q(y) \cdot \frac{p(y)}{cq(y)} = \frac{p(y)}{c}$$

Proof (cont'd)

- The marginal probability of accepting any sample (i.e., normalizing constant) is:

$$P(\text{accept}) = \sum_y P(\text{sample } y \text{ and accept}) = \sum_y \frac{p(y)}{c} = \frac{1}{c}$$

- The conditional probability of accepting a particular sample y given that the sample was accepted is:

$$P(y \mid \text{accept}) = \frac{P(\text{sample } y \text{ and accept})}{P(\text{accept})} = \frac{\frac{p(y)}{c}}{\frac{1}{c}} = p(y)$$

- Therefore, the accepted samples are distributed according to $p(\cdot)$, proving that the method works.

Stochastic Simulation

- This was a brief introduction to this area of stochastic simulation.
- Refer the book Simulation by Sheldon Ross!
- Some popular techniques in simulation are:
-
  - The inverse transform method
  - Accept-Reject method (rejection sampling)
  - Importance sampling
  - Markov Chain Monte Carlo (MCMC) methods
    - Hasting-Metropolis algorithm
    - Gibbs sampling
    - Slice sampling

Convergence of Random Variables

Pointwise Convergence

- When do we say that $\{x_n\}$ converges to $x \in \mathbb{R}$?

> We say that $\{x_n\}$ converges to $x \in \mathbb{R}$ (denoted by $x_n \to x$) if for every $\epsilon > 0$, we can find an $N(\epsilon) \in \mathbb{N}$ such that for $|x_n - x| < \epsilon$ for $n > N(\epsilon)$.

- What about convergence of functions?
- When do we say that a sequence of functions $F_n(\cdot)$ converge to $F(\cdot)$ on the domain $\mathbb{R}$?

> We say that the sequence of function $F_n(\cdot)$ converge to $F(\cdot)$ pointwise if the sequence $\{F_n(x)\}$ converges to $F(x)$ $(F_n(x) \to F(x))$ for all $x \in \mathbb{R}$.

Uniform Convergence

We say that the sequence of function $F_n(\cdot)$ converge to $F(\cdot)$ pointwise if the sequence $\{F_n(x)\}$ converges to $F(x)$ ($F_n(x) \to F(x)$) for all $x \in \mathbb{R}$.

- For every x, the sequence $\{F_n(x)\}$ coverges to $F(x)$.
- For every ϵ, there exists $N(\epsilon, x)$ which can depend on x.
- Only those $F_n(x)$ are ϵ close to $F(x)$ for which $n > N(\epsilon, x)$.

If $N(\epsilon, x) = N(\epsilon)$ (i.e., independent of x) for every $x \in \mathbb{R}$, then such convergence of $F_n(\cdot)$ to $F(\cdot)$ is called as uniform convergence.

Convergence of Sequence of random variables

- We will now be interested in the convergence properties of an infinite sequence of random variables $\{X_n\}$ to some limiting random variable X.
- What does the convergence $X_n \to X$ even mean ?
- When you perform the random experiment once, you get a sequence of realizations $\{x_n\}$ and x.
- If you are 'lucky', maybe $x_n \to x$.
- But if you were to perform the experiment again, you may not be so 'lucky' and get a different sequence $\{x'_n\}$ which may not converge to x'.
- We will come up with notions of convergence that depend on how often you see the sequence of realizations converging.

Convergence of Sequence of random variables

- Convergence of $X_n \to X$
- Here X could even be a deterministic number.
- $X_n's$ could be dependent on each other.
- Each random variable X_n could have a different law (pmf/pdf).

Modes of Convergence ($X_n \to X$)

Pointwise or Sure convergence

> $\{X_n, n \geq 0\}$ converges to X pointwise or surely if for all $\omega \in \Omega$ we have $\lim_{n\to\infty} X_n(\omega) = X(\omega)$

- Consider $\Omega = \{H, T\}$.
- Further, $X_n = \begin{cases} \frac{1}{n} & \text{if } \omega = H \\ 1 + \frac{1}{n} & \text{if } \omega = T. \end{cases}$ and $X = \begin{cases} 0 & \text{if } \omega = H \\ 1 & \text{if } \omega = T. \end{cases}$