STA 5106
Computational Methods in Statistics I

Department of Statistics
Florida State University

Class 17
October 31, 2019

Review: Congruential Generators

- Two steps to simulate a given distribution:
  1. Generate samples from a standard uniform probability on [0, 1],
  2. Transform those samples appropriately to simulate the desired distribution.

- We start with an initial number, $x_0 \in \{1, 2, . . . , m - 1\}$, called the **seed** of the number generator.

- **The remaining sequence is determined by the relation**

$$\boldsymbol{x_{i+1} = (ax_i) \bmod m,}$$

- **The sequence x_i/m, $i = 1, 2, \ldots$ is a simulation of uniform distribution on [0, 1] (or U[0, 1])**.

5.3 Simulating Other Random Variables

Discrete Random Variables

- A discrete random variable is characterized by its probability mass function defined as

$$P(x_1) = p_1, P(x_2) = p_2, \ldots, P(x_n) = p_n, \ldots$$

such that for all i, $0 \leq p_i \leq 1$, and $\Sigma\, p_i = 1$.

- Commonly used discrete random variables are binomial, Poisson, geometric and negative-binomial. For example,

1. **Poisson** random variable with parameter λ:

$$p_i = e^{-\lambda} \frac{\lambda^i}{i!}$$

2. **Binomial** random variable, with parameters (n, p):

$$p_i = \binom{n}{i} p^i (1-p)^{n-i}$$

Definition of Simulation

- **Definition 14** For a given random variable, with a specified probability mass function $\{(x_i, p_i), i = 0, 1, 2, \dots\}$, the process of selecting a value x_i with probability p_i is called **Simulation**.

- If this selection is performed many times, generating a sequence $\{X_j\}$, then
$$\frac{1}{n}\sum_{j=1}^{n} I_{X_j}(\{x_i\}) = \frac{\#\{X_j = x_i\}}{n} \to p_i.$$

- Two standard simulation techniques for discrete random variables: **1. inverse transform method; 2. acceptance/rejection method.**

Inverse Transform Method

- Let X be a discrete random variable with a given probability mass function. We have to utilize numbers generated using U[0, 1] to simulate values of X. The algorithm is as follows:

(a) Generate U according to U[0, 1].

(b) If U < p1

    Set x = x1

else if U < p2 + p1

    Set x = x2

. . .

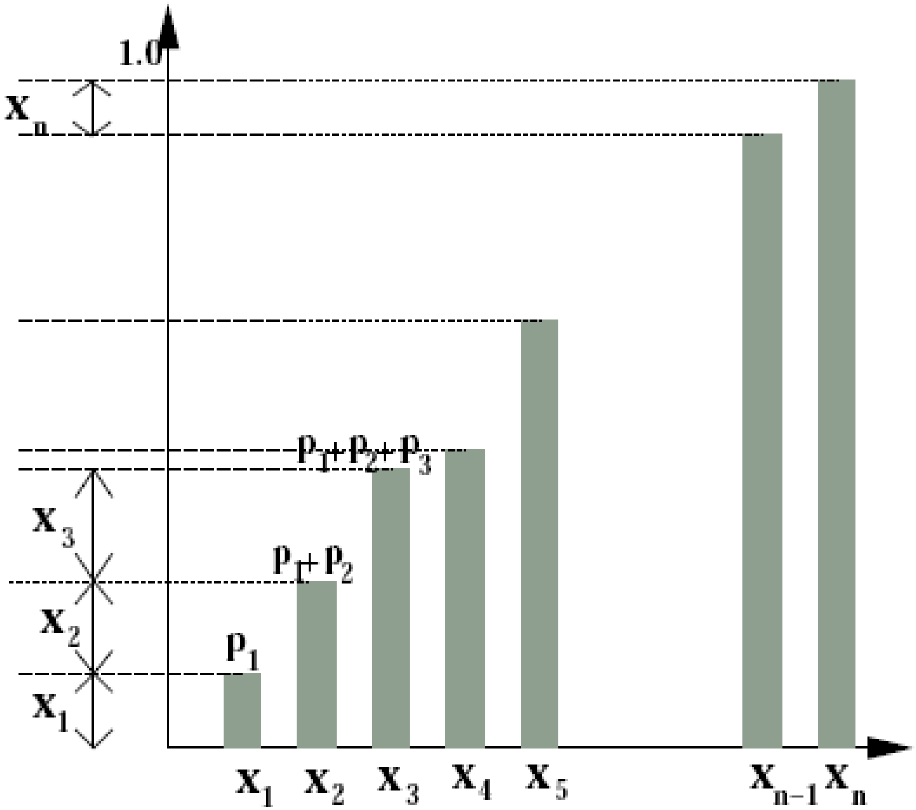
else if U < p1 + p2 + … + pn

    Set x = xn

. ..

Inverse Transform Method

- The random variable U will take value between 0 and 1 on the vertical axis. That value is projected to the right till it meets with a vertical bar and the corresponding value is selected.



Inverse Transform Method

- One can show that the value x_j is selected with probability p_j.

- To verify that X generated according to these steps has the desired probability distribution:

$$P(X = x_j) = P\left(\sum_{i=1}^{j-1} P(x_i) \le U \le \sum_{i=1}^{j} P(x_i)\right)$$

$$= \sum_{i=1}^{j} P(x_i) - \sum_{i=1}^{j-1} P(x_i) = P(x_j) = p_j$$

- This is a general technique for sampling discrete random variables. We take some examples.

Geometric random variable

- A random variable X is called geometric if its probability density function is given by

$$P\{X = i\} = p(1-p)^{i-1},\ i = 1, 2, \ldots .$$

- Consider an experiment involving independent and identical trials, each resulting in either success with probability p or failure with probability $(1-p)$.

- X denotes the number of trials performed to reach the first success. Then

$$P(X \leq j) = \sum_{i=1}^{j} p(1-p)^{i-1} = 1-(1-p)^{j}$$

Geometric random variable

- Let U be distributed according to U[0, 1]. The algorithm sets $X = j$ if

$$P(X \le j-1) \le U < P(X \le j)$$

or

$$1-(1-p)^{j-1} \le U < 1-(1-p)^{j}$$

$$(1-p)^{j-1} \ge 1-U > (1-p)^{j}$$

- In other words, set

$$X = \min\{j : (1-p)^{j} < 1-U\} = \min\{j : j > \log(1-U)/\log(1-p)\}$$

- Since U is the same random variable as $1 - U$,

$$X = ceil\{\log(U)/\log(1-p)\}$$

Poisson Random Variable

- X is a Poisson random variable with parameter λ if its probability mass function is given by

$$P\{X = i\} = \exp(-\lambda)\frac{\lambda^i}{i!}, \quad i = 0,1,2,\ldots$$

- Therefore,

$$\frac{P\{X = i+1\}}{P\{X = i\}} = \frac{\lambda}{i+1}.$$

- Using this results, the cumulative distribution function can be iteratively calculated as

$$P\{X \le i+1\} = P\{X \le i\} + \frac{\lambda}{i+1}P\{X = i\}.$$

Poisson Random Variable

- An algorithm to simulate X using inverse transform method is given by:

i. Generate U according to U[0, 1]

ii. Set $i = 0$, $p = \exp(-\lambda)$, and $F = p$.

iii. If $U < F$, set $X = i$ and stop.

iv. Set $p = p\lambda/(i+1)$, $F = F + p$, $i = i + 1$

v. Go to Step (iii).

Acceptance/Rejection Method

- This method is useful in situations where we want to simulate a random variable X and we have a technique of simulating another random variable Y.

- Let the probability density functions of X and Y be given by

$$P\{X = j\} = p_j, \quad P\{Y = j\} = q_j, \quad j = 0, 1, 2, \ldots .$$

- We will assume that there exists a constant c such that

$$p_j / q_j \leq c, \text{ for all } j \text{ for which } q_j > 0 .$$

Algorithm

- The general algorithm for acceptance/rejection is given by:

  (i) Simulate a value of *Y*.

  (ii) Generate U according to U[0, 1].

  (iii) If $U < p_Y/(cq_Y)$, then set X = Y,

  Else, return to (i)

- We can prove that this procedure simulates *X*. That is, for an *X* generated according to this procedure

$$P\{X = j\} = p_j \,.$$

Proof:
$$P(X = j) = P(Y = j \,|\, U \le p_Y/(cq_Y))$$
$$= P(Y = j)P(U \le p_Y/(cq_Y) \,|\, Y = j) / P(U \le p_Y/(cq_Y))$$
$$= q_j \cdot \frac{p_j}{cq_j} / \sum_j \frac{p_j}{cq_j} \cdot q_j = \frac{p_j}{c} / \frac{1}{c} = p_j.$$