CSE 4713 (Simulation and Modeling)
Department of Computer Science and Engineering
DUET, Gazipur-1700

Md. Abu Bakkar Siddique (ABS)
Assistant Professor, Department of CSE, DUET
Room #7024, New Academic Building
Mobile: +8801944275646 Ext: 3281
Email: absiddique@duet.ac.bd
absduet@gmail.com

Simulation: Simulation is the process of creating a computer-based or mathematical model that mimics the behavior of a real-world system. It involves running experiments or scenarios to understand how the system functions, make predictions, and analyze its performance under different conditions.

Modeling: Modeling is the process of constructing a simplified representation of a real-world system using mathematical, logical, or computational techniques. It involves identifying the relevant variables, relationships, and assumptions to describe the system's behavior and simulate its dynamics. Models can be used for analysis, prediction, optimization, and decision-making.

Basic Probability and Statistics

Why?

The use of probability and statistics is such an integral part of a simulation study. Probability and statistics are needed to understand how to-

- model a probabilistic system
- validate the simulation model
- choose the input probability distributions
- generate random samples from these distributions
- perform statistical analyses of the simulation output data
- and design the simulation experiments

RANDOM VARIABLES AND THEIR PROPERTIES

Sample Space

An experiment is a process whose outcome is not known with certainty.
The set of all possible outcomes of an experiment is called the sample space and is denoted by S.
The outcomes themselves are called the sample points in the sample space.

EXAMPLE:
If the experiment consists of flipping a coin then
$S = \{H, T\}$
where the symbol { } means the "set consisting of," and "H" and "T" mean that the outcome is a head and a tail, respectively.

RANDOM VARIABLES AND THEIR PROPERTIES

Random Variable

A random variable is a function (or rule) that assigns a real number (any number greater than $-\infty$ and less than ∞) to each point in the sample space S.

EXAMPLE

Consider the experiment of rolling a pair of dice. Then

S = {(1, 1), (1, 2), . . . , (6, 6)}

where (i, j) means that i and j appeared on the first and second die, respectively.

If X is the random variable corresponding to the sum of the two dice, then X assigns the value 7 to the outcome (4, 3).

RANDOM VARIABLES AND THEIR PROPERTIES

Discrete Random Variables

- Discrete random variables are variables that can only take on a countable number of distinct values. They represent outcomes or events that can be counted or enumerated.

- The probabilities associated with each possible value are defined by a probability mass function (PMF).

- Examples include the number of heads obtained in coin flips or the number of children in a family.

RANDOM VARIABLES AND THEIR PROPERTIES

The probability mass function

- For a discrete random variable X, we define the probability mass function p(a) of X by

$$p(a) = P\{X = a\}$$

- The probability mass function p(a) is positive for at most a countable number of values of a. That is, if X must assume one of the values x1, x2, . . . , then

$$p(xi) > 0, i = 1, 2, \ldots$$

$$p(x) = 0, \textit{all other values of } x$$

- Since X must take on one of the values x_i, we have

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

RANDOM VARIABLES AND THEIR PROPERTIES

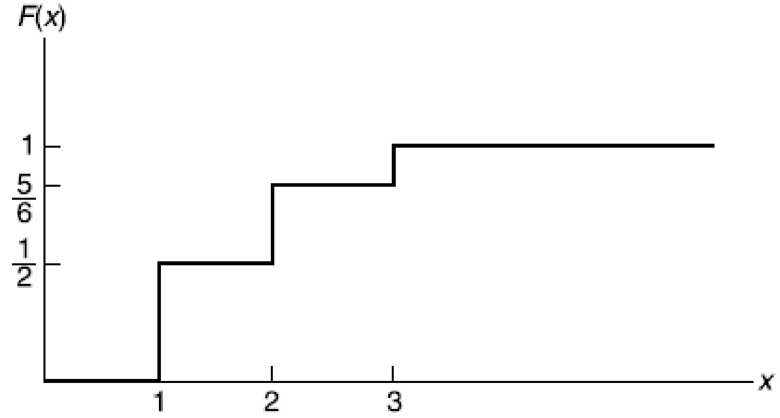
The cumulative distribution function

- The cumulative distribution function F can be expressed in terms of $p(a)$ by

$$F(a) = \sum_{all\ x_i \le a} p(x_i)$$

- For instance, suppose X has a probability mass function given by

$$p(1) = 1/2, \mathrm{p}(2) = 1/3, \mathrm{p}(3) = 1/6$$



RANDOM VARIABLES AND THEIR PROPERTIES

Continuous Random Variable

- X is a *continuous* random variable if there exists a nonnegative function $f(x)$, defined for all real $x \in (-\infty, \infty)$, having the property that for any set B of real numbers

$$P\{X \in B\} = \int_B f(x)dx$$

- The function $f(x)$ is called the *probability density function* of the random variable X.
- *This means* X will be in B may be obtained by integrating the probability density function over the set B. Since X must assume some value, $f(x)$ must satisfy

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx$$

RANDOM VARIABLES AND THEIR PROPERTIES

Continuous Random Variable

- X is a *continuous* random variable if there exists a nonnegative function $f(x)$, defined for all real $x \in (-\infty, \infty)$, having the property that for any set B of real numbers

$$P\{X \in B\} = \int_B f(x)dx$$

- The function $f(x)$ is called the *probability density function* of the random variable X.
- *This means* X will be in B may be obtained by integrating the probability density function over the set B. Since X must assume some value, $f(x)$ must satisfy

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx$$

RANDOM VARIABLES AND THEIR PROPERTIES

The distribution function

- The distribution function $F(x)$ for a continuous random variable X is given by

$$F(x) = P(X \epsilon [(-\infty, x]) = \int_{-\infty}^{x} f(y)\, dy \; for\ all -\infty < x < \infty$$

- if $I = [a, b]$ for any real numbers a and b such that $a < b$, then

$$P(X \in I) = \int_{a}^{b} f(y) dy = F(b) - F(a)$$

Statistical Models in Simulation

1. *Discrete random variables.* Let X be a random variable. If the number of possible values of X is finite, or countably infinite, X is called a discrete random variable. The possible values of X may be listed as. $x_1, x_2, \ldots$ In the finite case, the list terminates; in the countably infinite case, the list continues indefinitely.

Example 5.1

The number of jobs arriving each week at a job shop is observed. The random variable of interest is X, where

$$X = \text{number of jobs arriving each week}$$

The possible values of X are given by the range space of X, which is denoted by R_X. Here $R_X = \{0, 1, 2, \ldots\}$.

Let X be a discrete random variable. With each possible outcome x_i in R_X, a number $p(x_i) = P(X = x_i)$ gives the probability that the random variable equals the value of x_i. The numbers $p(x_i)$, $i = 1, 2, \ldots$, must satisfy the following two conditions:

1. $p(x_i) \geq 0$, for all i
2. $\sum_{i=1}^{\infty} p(x_i) = 1$

The collection of pairs $(x_i, p(x_i))$, $i = 1, 2, \ldots$ is called the probability distribution of X, and $p(x_i)$ is called the probability mass function (pmf) of X.

Discrete random variables.

Example 5.2

Consider the experiment of tossing a single die. Define X as the number of spots on the up face of the die after a toss. Then $R_X = \{1, 2, 3, 4, 5, 6\}$. Assume the die is loaded so that the probability that a given face lands up is proportional to the number of spots showing. The discrete probability distribution for this random experiment is given by

x_i	1	2	3	4	5	6
$p(x_i)$	1/21	2/21	3/21	4/21	5/21	6/21

The conditions stated earlier are satisfied—that is, $p(x_i) \geq 0$ for $i = 1, 2, \ldots, 6$ and $\sum_{i=1}^{\infty} p(x_i) = 1/21 + \cdots + 6/21 = 1$. The distribution is shown graphically in Figure 5.1.

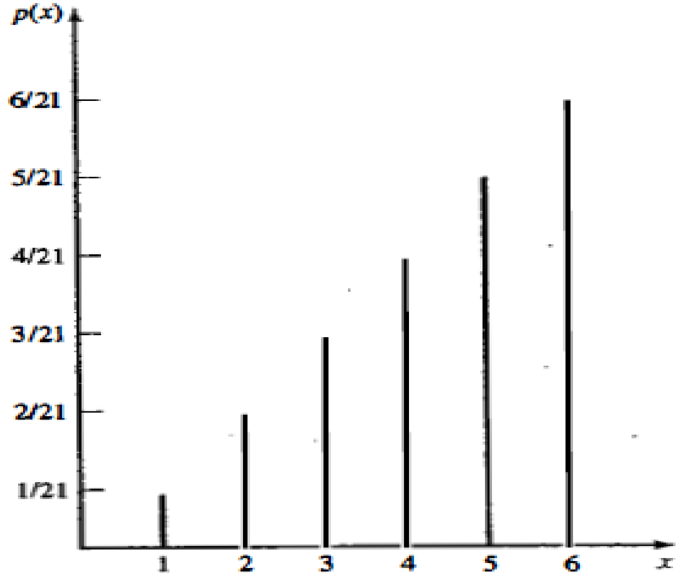


Figure 5.1 Probability mass function for loaded-die example.

Continuous random variables

2. *Continuous random variables.* If the range space R_X of the random variable X is an interval or a collection of intervals, X is called a continuous random variable. For a continuous random variable X, the probability that X lies in the interval $[a, b]$ is given by

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx \tag{5.1}$$

The function $f(x)$ is called the probability density function (pdf) of the random variable X. The pdf satisfies the following conditions:

a. $f(x) \geq 0$ for all x in R_X

b. $\int_{R_X} f(x)dx = 1$

c. $f(x) = 0$ if x is not in R_X

As a result of Equation (5.1), for any specified value x_0, $P(X = x_0) = 0$, because

$$\int_{x_0}^{x_0} f(x)\,dx = 0$$

Continuous random variables.

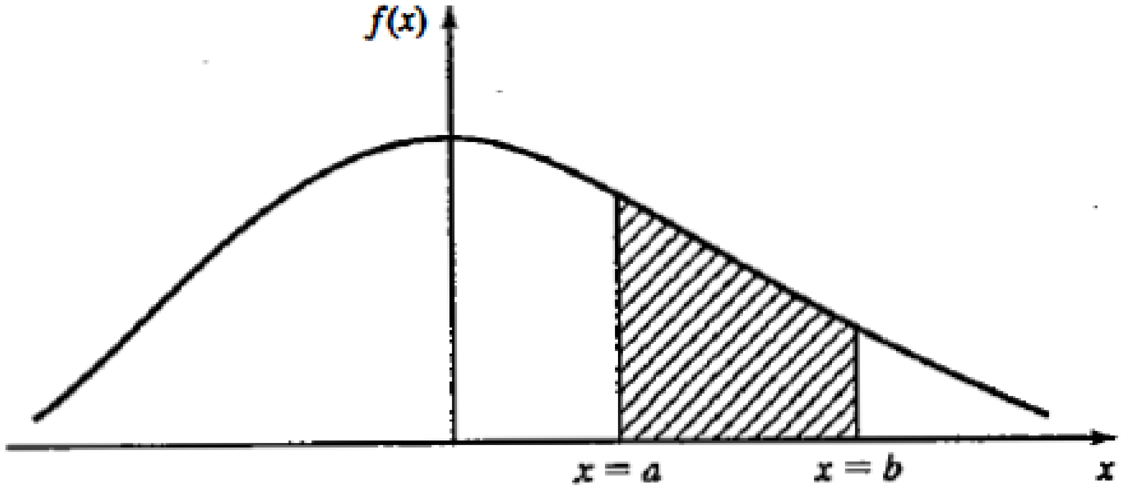


Figure 5.2 Graphical interpretation of $P(a < X < b)$.

$P(X = x_0) = 0$ also means that the following equations hold:

$$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b) \tag{5.2}$$

The graphical interpretation of Equation (5.1) is shown in Figure 5.2. The shaded area represents the probability that X lies in the interval $[a, b]$.

Continuous random variables.

Example 5.3

The life of a device used to inspect cracks in aircraft wings is given by X, a continuous random variable assuming all values in the range $x \geq 0$. The pdf of the lifetime, in years, is as follows:

$$f(x) = \begin{cases} \frac{1}{2} e^{-x/2}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

The probability that the life of the device is between 2 and 3 years is calculated as

$$P(2 \leq X \leq 3) = \frac{1}{2} \int_2^3 e^{-x/2} dx$$

$$= -e^{-3/2} + e^{-1} = -0.223 + 0.368 = 0.145$$

Cumulative distribution function

3. *Cumulative distribution function.* The cumulative distribution function (cdf), denoted by $F(x)$, measures the probability that the random variable X assumes a value less than or equal to x, that is, $F(x) = P(X \le x)$.

If X is discrete, then

$$F(x) = \sum_{\substack{\text{all} \\ x_i \le x}} p(x_i) \tag{5.3}$$

If X is continuous, then

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \tag{5.4}$$

Some properties of the cdf are listed here:

a. F is a nondecreasing function. If $a < b$, then $F(a) \le F(b)$.
b. $\lim_{x \to \infty} F(x) = 1$
c. $\lim_{x \to -\infty} F(x) = 0$

All probability questions about X can be answered in terms of the cdf. For example,

$$P(a < X \le b) = F(b) - F(a) \quad \text{for all } a < b \tag{5.5}$$

Cumulative distribution function

Example 5.5

The cdf for the device described in Example 5.3 is given by

$$F(x) = \frac{1}{2}\int_0^x e^{-t/2} dt = 1 - e^{-x/2}$$

The probability that the device will last for less than 2 years is given by

$$P(0 \leq X \leq 2) = F(2) - F(0) = F(2) = 1 - e^{-1} = 0.632$$

The probability that the life of the device is between 2 and 3 years is calculated as

$$P(2 \leq X \leq 3) = F(3) - F(2) = (1 - e^{-3/2}) - (1 - e^{-1})$$
$$= -e^{-3/2} + e^{-1} = -0.223 + 0.368 = 0.145$$

as found in Example 5.3.

Expectation

4. *Expectation.* An important concept in probability theory is that of the expectation of a random variable. If X is a random variable, the expected value of X, denoted by $E(X)$, for discrete and continuous variables is defined as follows:

$$E(X) = \sum_{\text{all } i} x_i p(x_i) \qquad \text{if } X \text{ is discrete} \tag{5.6}$$

and

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \qquad \text{if } X \text{ is continuous} \tag{5.7}$$

The expected value $E(X)$ of a random variable X is also referred to as the mean, μ, or the first moment of X.

Example 5.7

The mean of the life of the device described in Example 5.3 are computed as follows:

$$E(X) = \frac{1}{2}\int_0^{\infty} xe^{-x/2} dx = -xe^{-x/2}\Big|_0^{\infty} + \int_0^{\infty} e^{-x/2} dx$$

$$= 0 + \frac{1}{1/2} e^{-x/2}\Big|_0^{\infty} = 2 \text{ years}$$

The variance and the standard deviation

The variance of a random variable, X, denoted by $V(X)$ or var(X) or σ^2, is defined by

$$V(X) = E[(X - E[X])^2]$$

A useful identity in computing $V(X)$ is given by

$$V(X) = E(X^2) - [E(X)]^2 \tag{5.10}$$

The mean $E(X)$ is a measure of the central tendency of a random variable. The variance of X measures the expected value of the squared difference between the random variable and its mean. Thus, the variance, $V(X)$, is a measure of the spread or variation of the possible values of X around the mean $E(X)$. The standard deviation, σ, is defined to be the square root of the variance, σ^2. The mean, $E(X)$, and the standard deviation, $\sigma = \sqrt{V(X)}$, are expressed in the same units.

The variance and the standard deviation

Example 5.7

The variance of the life of the device described in Example 5.3 are computed as follows:

To compute $V(X)$ from Equation (5.10), first compute $E(X^2)$ from Equation (5.9) as follows:

$$E(X^2) = \frac{1}{2}\int_0^{\infty} x^2 e^{-x/2} dx$$

∫xe^(-x) dx = uv - ∫v du

= x * (-e^(-x)) - ∫(-e^(-x)) dx

= -xe^(-x) + ∫e^(-x) dx

= -xe^(-x) - e^(-x) + C,

Thus,

$$E(X^2) = -x^2 e^{-x/2}\Big|_0^{\infty} + 2\int_0^{\infty} x e^{-x/2} dx = 8$$

giving

$$V(X) = 8 - 2^2 = 4 \text{ years}^2$$

and

$$\sigma = \sqrt{V(X)} = 2 \text{ years}$$

With a mean life of 2 years and a standard deviation of 2 years, most analysts would conclude that actual lifetimes, X, have a fairly large variability.

Any Question?

Thank You All