Week 11:
Ramp- and triangular distributions, Acceptance/rejection methods and generating normal variates

- In this section, we present and discuss several procedures for generating random variates from continuous distributions.
- 2 ways of doing it :

1) Acceptance–Rejection Method
2) Inverse Transformation Method (triangular)

Direct method and Acceptance/

- This method is generally used for distributions whose domains are defined over finite(so from point a to b that IS NOT infinite) intervals.
- defined over the interval a <= x <= b.

The acceptance/rejection method supposes a probability density function is defined over a finite interval, thus $f(t)$ is defined over interval $[a, b]$. To generate a sequence random variable values $\bar{t} = \{t_1, t_2, \ldots, t_n\}$, where n is the required number of random variates, follow this procedure:

1. Select a constant M such that $M = \max\{f(t) | t \in [a, b]\}$
2. Initialize $i \leftarrow 1$ and $\bar{t} \leftarrow \emptyset$
3. While $i < n$ continue, else exit.
4. Generate two random numbers r_1 and r_2.
5. Compute $t^* \leftarrow a + (b - a)r_1$.
6. Evaluate $f(t^*)$.
7. If $r_2 \leq \frac{f(t^*)}{M}$ then $t_i \leftarrow t^*$ and $i \leftarrow i + 1$
8. Return to step 3.

I will follow these steps with Lingo the example :

```
#Part 2
#The only code required to carry out the simulation follows below
mm=2^31-1
aa=7^5
cc=100100100
x=123456789
```

<- Code used to generate random numbers

```
#Create a function to evaluate "t"-values in the PDF
f=function(t)return((-3/4000)*t^2+(9/200)*t-3/5)
```

<-This is the given function

```
num=1000

#Initialize

i=1
```

2 Initialize $i \leftarrow 1$ and $\bar{t} \leftarrow \emptyset$

```
a=20
b=40
```

<- Given a and b value, M is the highest point in the function

```
M=f(30)
```

1 Select a constant M such that $M = \max\{f(t) | t \in [a, b]\}$

M can either be given or seen by looking at a graph of the function ,or it can be caculated to get the highest f(x) value;

```
tbar=c()
```

2 Initialize $i \leftarrow 1$ and $\bar{t} \leftarrow \emptyset$

...

```
while(i<num){
```

8 Return to step 3.

```
  x=(aa*x+cc)%%mm
  r1=x/mm
```

4 Generate two random numbers r_1 and r_2.

```
  x=(aa*x+cc)%%mm
  r2=x/mm

  tstar=a+(b-a)*r1
```

5 Compute $t^* \leftarrow a + (b - a)r_1$.

```
  ftstar=f(tstar)
```

6 Evaluate $f(t^*)$.

```
  if(r2<=(ftstar/M)){
```

7 If $r_2 \leq \frac{f(t^*)}{M}$ then $t_i \leftarrow t^*$ and $i \leftarrow i + 1$

```
    tbar[i]=tstar

    i=i+1

  }

}

#hist(tbar)
```

EXAMPLE 5 The Acceptance–Rejection Method

Use the acceptance–rejection method to generate random variates from a triangular distribution whose pdf is given by

$$f(x) = \begin{cases} -\frac{1}{6} + \frac{x}{12} & 2 \le x \le 6 \\ \frac{4}{3} - \frac{x}{6} & 6 \le x \le 8 \end{cases}$$

Solution For simplicity, we redefine $f(x) = -\frac{1}{6} + \frac{x}{12}$ as $f_1(x)$, and $f(x) = \frac{4}{3} - \frac{x}{6}$ as $f_2(x)$. This distribution is represented graphically in Figure 12.

Since this distribution is defined over two intervals, we must modify steps 4 and 5 of the acceptance–rejection method to account for these ranges. The first three steps of the algorithm, however, stay the same as before. That is, step 1 determines M, step 2 generates r_1 and r_2, and step 3 transforms r_1 into a value x^* of $\mathbf{X}$.

From the graph of the pdf in Figure 12, it is clear that $M = \frac{1}{3}$. This distribution has the endpoints [2, 8], which implies that $a = 2$ and $b = 8$. If we now substitute these endpoints in step 3, the x^* values are generated by the equation $x^* = 2 + 6r_1$. Then we see that if r_1 is between 0 and $\frac{2}{3}$, x^* will lie in the range from 2 to 6. If $r_1 > \frac{2}{3}$, x^* will lie in the interval [6, 8]. To account for this, we make our first modification in step 4. If x^* lies between 2 and 6, then in step 4, we use the function $f_1(x)$ to evaluate $f(x^*)$. Otherwise, we use $f_2(x)$ to compute $f(x^*)$. Step 4 now can be summarized as follows: If $2 \le x^* \le 6$,

$$\begin{aligned} f(x^*) &= f_1(x^*) \\ &= -\frac{1}{6} + \frac{x^*}{12} \\ &= \frac{r_1}{2} \end{aligned}$$

If $6 \le x^* \le 8$,

$$\begin{aligned} f(x^*) &= f_2(x^*) \\ &= \frac{4}{3} - \frac{x^*}{6} \\ &= 1 - r_1 \end{aligned}$$

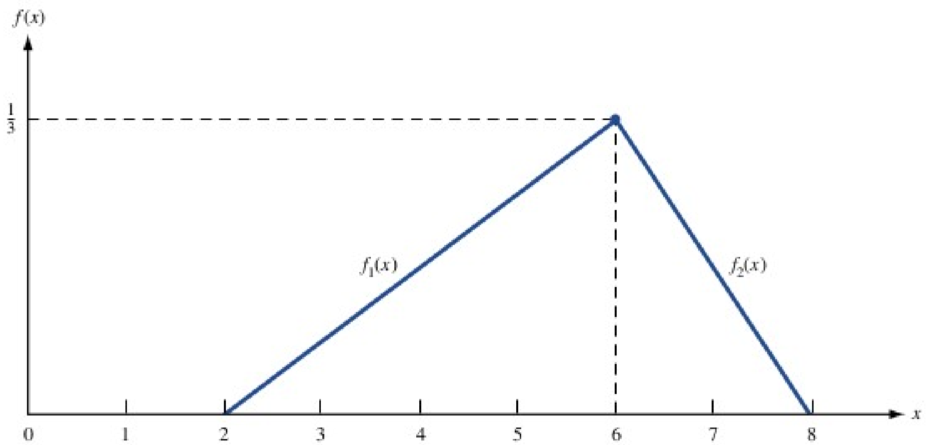

FIGURE 12 The pdf of Triangular Distribution of Example 5

The next step in the algorithm is either to accept or to reject the current value of x^*. We accept x^* if the condition $r_2 \le \frac{f(x^*)}{M}$ is satisfied. However, following step 4, we need to evaluate this condition over the two intervals by substituting the appropriate function, $f(x^*)$, into the relation. In other words, step 5 for this distribution will now be as follows: For $2 \le x^* \le 6$, we accept x^* if $r_2 \le \frac{f_1(x^*)}{M}$—that is, if $r_2 \le \frac{3r_1}{2}$. For $6 \le x^* \le 8$, we accept x^* if $r_2 \le \frac{f_2(x^*)}{M}$—that is, if $r_2 \le 3(1 - r_1)$. If x^* is rejected, we go back to step 2 and repeat the process.

As before, some of the x^* values will be rejected. In this case, also, the probability of accepting a random variate is .5. That is, one half of all random variates generated in step 3 will, in the long run, be rejected in step 5.

Direct method for Normal distribution :

In the direct method for client i, first two standard normal distributed variables z_{i1} and z_{i2} are calculated from which two normal distributed variables, t_{i1} and t_{i2} are obtained. During the simulation, either the stream of t_{i1} or the stream of t_{i2} will be used to provide values to the required t_i's, but not both. The simulator decides in advance which stream will be used.

Regardless of whether t_{i1} or t_{i2} is going to be used to form the t_i's stream, always generate two random numbers R_{i1} and R_{i2} and insert into in any one of

$$z_{i1} = \sqrt{-2 \ln R_{i1}} \sin 2\pi R_{i2} \quad \text{of/or} \quad z_{i2} = \sqrt{-2 \ln R_{i1}} \cos 2\pi R_{i2}$$

from which

$$t_{i1} = \mu + \sigma z_{i1} \quad \text{of/or} \quad t_{i2} = \mu + \sigma z_{i2},$$

is calculated.

As mentioned earlier, use either t_{i1} or t_{i2} to add to the required stream of t_i's according to the prior decision.

#Direct method to generate normal variates

#Say we are required to generate 200 t-values
#that are from a normal distribution with a
#mean of 180 and a standard deviation of 25.

num=200
mu=180
sigma=25

m=2^31-1
a=7^5
c=101010101
x=987654321

<---Used to generate R_1 and R_2

tbar=c()

<-tbar will hold the values

for(i in 1:num){

x=(a*x+c)%%m
r1=x/m

<-generate R_1 and R_2

x=(a*x+c)%%m
r2=x/m

Z=sqrt(-2*log(r1))*sin(2*pi*r2)

$z_{i1} = \sqrt{-2 \ln R_{i1}} \sin 2\pi R_{i2}$

tbar[i]=mu+Z*sigma

$t_{i1} = \mu + \sigma z_{i1}$

}

hist(tbar)

triangular distribution:
(we were told to go through this example)

EXAMPLE 4 **The Triangular Distribution**

Consider a random variable **X** whose pdf is given by

$$f(x) = \begin{cases} \frac{1}{2}(x - 2) & 2 \le x \le 3 \\ \frac{1}{2}(2 - \frac{x}{3}) & 3 \le x \le 6 \\ 0 & \text{otherwise} \end{cases}$$

Use the ITM to generate observations from the distribution. This distribution, called a *triangular* distribution, is represented graphically in Figure 9. It has the endpoints [2, 6], and its mode is at 3. We can see that 25% of the area under the curve lies in the range of x from 2 to 3, and the other 75% lies in the range from 3 to 6. In other words, 25% of the values of the random variable **X** lie between 2 and 3, and the other 75% fall between 3 and 6. The triangular distribution has important applications in simulation. It is often used to represent activities for which there are few or no data. (For a detailed account of this distribution, see Banks and Carson (1984) or Law and Kelton (1991).)

Solution The cdf of this triangular distribution is given by the function

$$F(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{4}(x - 2)^2 & 2 \le x \le 3 \\ -\frac{1}{12}(x^2 - 12x + 24) & 3 \le x \le 6 \\ 1 & \text{otherwise} \end{cases}$$

For simplicity, we redefine $F(x) = (\frac{1}{4})(x - 2)^2$, for $2 \le x \le 3$, as $F_1(x)$, and $F(x) = (-\frac{1}{12})(x^2 - 12x + 24)$, for $3 \le x \le 6$, as $F_2(x)$.

This cdf can be represented graphically as shown in Figure 10. Note that at $x = 3$, $F(3) = 0.25$. This implies that the function $F_1(x)$ covers the first 25% of the range of the cdf, and $F_2(x)$ applies over the remaining 75% of the range. Since we now have two separate functions representing the cdf, the ITM has to be modified to account for these two functions, their ranges, and the distribution of the ranges. As far as the ITM goes, the distribution of the ranges is the most important. This distribution is achieved by using the random number from step 2. In other words, if $r < 0.25$, we use the function $F_1(x) = (\frac{1}{4})(x - 2)^2$ in step 3. Otherwise, we use $F_2(x) = (-\frac{1}{12})(x^2 - 12x + 24)$. Since $r < 0.25$ for 25% of the time and $r \ge 0.25$ for the other 75%, we achieve the desired distribution. In

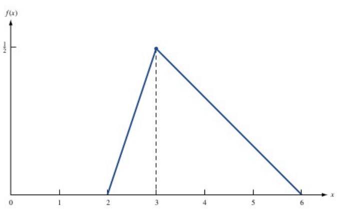

FIGURE 9
Density Function for a Triangular Distribution

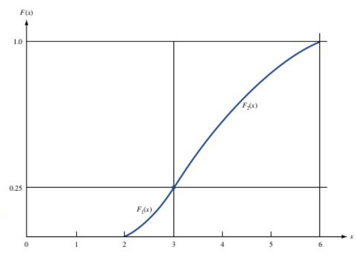

FIGURE 10
The cdf of a Triangular Distribution

either case, we set the function $F_1(x)$ or $F_2(x)$ equal to r and solve for x. That is, we solve one of the following equations:

$$\begin{aligned} (\tfrac{1}{4})(x - 2)^2 &= r && \text{for } 0 \le r < 0.25 \\ (-\tfrac{1}{12})(x^2 - 12x + 24) &= r && \text{for } 0.25 \le r \le 1.0 \end{aligned}$$

x will then be our random variable of interest.

As the graph in Figure 10 shows, a random number between 0 and 0.25 will be transformed into a value of x between 2 and 3. Similarly, if $r \ge 0.25$, it will be transformed into a value of x between 3 and 6.

To solve the first equation, $(\frac{1}{4})(x - 2)^2 = r$, we multiply the equation by 4 and then take the square root of both sides. This gives us

$$\begin{aligned} x - 2 &= \pm\sqrt{4r} \\ x &= 2 \pm 2\sqrt{r} \end{aligned}$$

Since x is defined only for values greater than 2, $x = 2 - 2\sqrt{r}$ is infeasible, leaving

$$x = 2 + 2\sqrt{r}$$

as the process generator for a random number in the range from 0 to 0.25. Note that when $r = 0$, $x = 2$, the smallest possible value for this range. Similarly, when we generate $r = 0.25$, it will be transformed to $x = 3$.

To solve the second equation, $(-\frac{1}{12})(x^2 - 12x + 24) = r$, we can use one of two methods: (1) employing the quadratic formula or (2) completing the square. (See Banks and Carson (1984) for details of the quadratic formula method.) Here, we use the method of completing the square. We multiply the equation by -12 and rearrange the terms to get

$$x^2 - 12x = -24 - 12r$$

To complete the square, we first divide the x term's coefficient by 2. This gives us -6. Next, we square this value to get 36. Finally, we add this resultant to both sides of the equation. This leaves us with the equation

$$\begin{aligned} x^2 - 12x + 36 &= 12 - 12r \\ (x - 6)^2 &= 12 - 12r \end{aligned}$$

Writing the equation in this form enables us to take the square root of both sides. That is,

$$\begin{aligned} x - 6 &= \pm\sqrt{12 - 12r} \\ x &= 6 \pm 2\sqrt{3 - 3r} \end{aligned}$$

As before, part of the solution is infeasible. In this case, x is feasible only for values less than 6. Thus, we use only the equation $x = 6 - 2\sqrt{3 - 3r}$ as our process generator. Note that when $r = 0.25$, our random variate is equal to 3. Similarly, when $r = 1$, we generate a random variate equal to 6.