Purpose & Overview

- Develop understanding of generating samples from a specified distribution as input to a simulation model.

- Illustrate some widely-used techniques for generating random variates.
  - Inverse-transform technique
  - Acceptance-rejection technique
  - Special properties
Inverse-transform Technique

- **The concept:**
  - For cdf function: \( r = F(x) \)
  - Generate \( r \) from uniform (0,1)
  - Find \( x \):

\[
x = F^{-1}(r)
\]

Exponential Distribution

- **Exponential Distribution:**
  - Exponential cdf:
    \[
    r = F(x) = 1 - e^{-\lambda x} \quad \text{for} \ x \geq 0
    \]
  - To generate \( X_1, X_2, X_3 \ldots \):

\[
X_i = F^{-1}(R) = -\left(\frac{1}{\lambda}\right) \ln(1-R) \quad \text{[Eq'n 8.3]}
\]

Figure: Inverse-transform technique for \( \exp(\lambda = 1) \)
Exponential Distribution

Example: Generate 200 variates $X_i$ with distribution $\exp(\lambda = 1)$

- Generate 200 Rs with U(0,1) and utilize eq’n 8.3, the histogram of $X$s become:

- Check: Does the random variable $X_1$ have the desired distribution?

\[ P(X_1 \leq x_0) = P(R \leq F(x_0)) = F(x_0) \]

Other Distributions

Examples of other distributions for which inverse cdf works are:

- Uniform distribution
- Weibull distribution
- Triangular distribution
Empirical Continuous Dist’n  [Inverse-transform]

- When theoretical distribution is not applicable
- To collect empirical data:
  - Resample the observed data
  - Interpolate between observed data points to fill in the gaps
- For a small sample set (size n):
  - Arrange the data from smallest to largest
    
    $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$
  - Assign the probability $1/n$ to each interval $x_{(i-1)} \leq x \leq x_{(i)}$

\[
X = \hat{F}^{-1}(R) = x_{(i-1)} + a_i \left( \frac{R - (i-1)}{n} \right)
\]

where
\[
a_i = \frac{x_{(i)} - x_{(i-1)}}{1/n - (i-1)/n} = \frac{x_{(i)} - x_{(i-1)}}{1/n}
\]

Example: Suppose the data collected for 100 broken-widget repair times are:

<table>
<thead>
<tr>
<th>Interval (Hours)</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Frequency, $c_i$</th>
<th>Slope, $a_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.81</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0.10</td>
<td>0.41</td>
<td>5.0</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>0.25</td>
<td>0.66</td>
<td>2.0</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
<td>0.34</td>
<td>1.00</td>
<td>1.47</td>
</tr>
</tbody>
</table>

Consider $R_1 = 0.83$:

$c_3 = 0.66 < R_1 < c_4 = 1.00$

$X_1 = x_{(3)} + a_3 (R_1 - c_3) = 1.5 + 1.47(0.83-0.66) = 1.75$
Discrete Distribution

- All discrete distributions can be generated via inverse-transform technique
- Method: numerically, table-lookup procedure, algebraically, or a formula
- Examples of application:
  - Empirical
  - Discrete uniform
  - Gamma

Example: Suppose the number of shipments, \( x \), on the loading dock of IHW company is either 0, 1, or 2

- Data - Probability distribution:

<table>
<thead>
<tr>
<th>( x )</th>
<th>( p(x) )</th>
<th>( F(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>1</td>
<td>0.30</td>
<td>0.80</td>
</tr>
<tr>
<td>2</td>
<td>0.20</td>
<td>1.00</td>
</tr>
</tbody>
</table>

- Method - Given \( R \), the generation scheme becomes:

\[
\begin{align*}
x & = \begin{cases} 
0, & 0 \leq R \leq 0.5 \\
1, & 0.5 \leq R \leq 0.8 \\
2, & 0.8 \leq R \leq 1.0
\end{cases}
\end{align*}
\]

Consider \( R_1 = 0.73 \):
- \( F(x_0) < R_1 \leq F(x_1) \)
- \( x_1 = 1 \)

Hence, \( x_1 = 1 \)
Acceptance-Rejection technique

- Useful particularly when inverse cdf does not exist in closed form, a.k.a. thinning
- Illustration: To generate random variates, $X \sim U(1/4, 1)$

Procedures:
- Step 1. Generate $R \sim U[0,1]$
- Step 2a. If $R \geq 1/4$, accept $X=R$.
- Step 2b. If $R < 1/4$, reject $R$, return to Step 1

- $R$ does not have the desired distribution, but $R$ conditioned ($R'$) on the event \{R $\geq$ 1/4\} does.
- Efficiency: Depends heavily on the ability to minimize the number of rejections.

NSPP

Non-stationary Poisson Process (NSPP): a Poisson arrival process with an arrival rate that varies with time

Idea behind thinning:
- Generate a stationary Poisson arrival process at the fastest rate, $\lambda^* = \max \lambda(t)$
- But "accept" only a portion of arrivals, thinning out just enough to get the desired time-varying rate
Example: Generate a random variate for a NSPP

**Data: Arrival Rates**

<table>
<thead>
<tr>
<th>t (min)</th>
<th>Mean Time Between Arrivals</th>
<th>Arrival Rate λ(t) (#/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>15</td>
<td>1/15</td>
</tr>
<tr>
<td>60</td>
<td>12</td>
<td>1/12</td>
</tr>
<tr>
<td>120</td>
<td>7</td>
<td>1/7</td>
</tr>
<tr>
<td>180</td>
<td>5</td>
<td>1/5</td>
</tr>
<tr>
<td>240</td>
<td>8</td>
<td>1/8</td>
</tr>
<tr>
<td>300</td>
<td>10</td>
<td>1/10</td>
</tr>
<tr>
<td>360</td>
<td>15</td>
<td>1/15</td>
</tr>
<tr>
<td>420</td>
<td>20</td>
<td>1/20</td>
</tr>
<tr>
<td>480</td>
<td>20</td>
<td>1/20</td>
</tr>
</tbody>
</table>

**Procedures:**

1. \( \lambda^* = \max \lambda(t) = 1/5 \), \( t = 0 \) and \( i = 1 \).
2. For random number \( R = 0.2130 \), \( E = -5\ln(0.213) = 13.13 \)
   \( t = 13.13 \)
3. Generate \( R = 0.8830 \)
   \( \lambda(13.13)/\lambda^* = (1/15)/(1/5) = 1/3 \)
   Since \( R > 1/3 \), do not generate the arrival.
4. For random number \( R = 0.5530 \), \( E = -5\ln(0.553) = 2.96 \)
   \( t = 13.13 + 2.96 = 16.09 \)
5. Generate \( R = 0.0240 \)
   \( \lambda(16.09)/\lambda^* = (1/15)/(1/5) = 1/3 \)
   Since \( R < 1/3 \), \( T_1 = t = 16.09 \), and \( i = i + 1 = 2 \)

**Special Properties**

- Based on features of particular family of probability distributions
- For example:
  - Direct Transformation for normal and lognormal distributions
  - Convolution
  - Beta distribution (from gamma distribution)
Direct Transformation [Special Properties]

- **Approach for normal(0, 1):**
  - Consider two standard normal random variables, $Z_1$ and $Z_2$, plotted as a point in the plane:
    
    In polar coordinates:
    
    $Z_1 = B \cos \phi$
    $Z_2 = B \sin \phi$

  - $B^2 = Z_1^2 + Z_2^2 \sim \text{chi-square distribution with 2 degrees of freedom} = \text{Exp}(\lambda = 2)$. Hence, $B = (-2 \ln R)^{1/2}$
  - The radius $B$ and angle $\phi$ are mutually independent.
    
    $Z_1 = (-2 \ln R)^{1/2} \cos(2\pi R_1)$
    $Z_2 = (-2 \ln R)^{1/2} \sin(2\pi R_2)$

Direct Transformation [Special Properties]

- **Approach for normal($\mu$, $\sigma^2$):**
  - Generate $Z_i \sim N(0, 1)$
    
    $X_i = \mu + \sigma Z_i$

- **Approach for lognormal($\mu$, $\sigma^2$):**
  - Generate $X \sim N((\mu, \sigma^2)$
    
    $Y_i = e^{X_i}$
Summary

- Principles of random-variate generate via
  - Inverse-transform technique
  - Acceptance-rejection technique
  - Special properties
- Important for generating continuous and discrete distributions