Random Variate Generation
We've seen how to generate **UNIFORM** random variables:

How about other distributions?

Some simple examples

Let $U$ be a uniform $U(0,1)$ random variable

1. Continuous uniform on $[a,b]$

   $f(t) = \frac{1}{b-a}$ so $F(t) = \int_a^t \frac{1}{b-a} \cdot ds = \frac{t-a}{b-a}$

Now define $X = a + (b-a) \cdot U$

   $\mathbb{P} \{ X < t \} = \mathbb{P} \{ a + (b-a) \cdot U < t \} = \mathbb{P} \{ U < \frac{t-a}{b-a} \}$

   $= F(t)$
2. Discrete uniform on $[a, b]$, $a, b$ integers

$$X = \text{integer}\left((b+1-a) \cdot U\right)$$

Example: $b = 5, a = 1$

```
1 2 3 4 5 6
```

5 values, 5 equi-length regions of number line

$$\begin{align*}
R \{ x \in \mathbb{Z} \mid x = \text{integer}\left((b+1-a) \cdot U\right) = i \} \\
R \{ \frac{i}{b+1-a} \leq U \leq \frac{i+1}{b+1-a} \} & = \frac{1}{b+1-a}
\end{align*}$$
INVERSE TRANSFORM METHOD

- extremely general

Formally $F^{-1}$ exists because $F$ is monotone and covers $[0,1]$.

$\Rightarrow$ for every $p$ there exists unique $x$ with $F'(p) = x$, namely $x$ s.t. $F(x) = p$

$$F^{-1}(p) = \inf_x \{ F(x) \geq p \}$$
**Theorem**

Let $Y$ be uniform $U(0,1)$ and define $X = F^{-1}(Y)$.

Then $X$ has distribution $F$.

**Proof**

\[
R \{ X \leq t \} = R \{ F^{-1}(Y) \leq t \} \\
= R \{ \inf_{x} \{ F(x) \geq Y \} \leq t \} \\
= R \{ Y \leq F(t) \} \\
= F(t)
\]

Any discrete distribution

Note: for $p_0 + p < u \leq p_0 + p + p_2$

\[
\inf_{x} \{ F(x) \geq u \} = 2
\]

i.e. we round up.

Continuous needs no rounding.
Application

a. Exponential

We know \( r = F(x) = 1 - e^{-\lambda x} \)

For what \( x \) is \( F^{-1}(r) = x \)?

Set up and solve \( r = 1 - e^{-\lambda x} \)

\[ 
1 - r = e^{-\lambda x} \\
\ln(1 - r) = -\lambda \ln(e^x) \\
\ln(1 - r) = -\lambda x \\
\Rightarrow -\frac{1}{\lambda} \ln(1 - r) = x 
\]

So sample \( x = -\frac{1}{\lambda} \ln(1 - r) \)

to generate a sample from \( \text{exp}(\lambda) \)
Aside

IF we have two independent exponentials, with rate $\lambda$, and $\lambda_2$ it is always possible when $\lambda_1 << \lambda_2$ to still get a sample from $\text{exp}(\lambda_1)$ that is larger than a sample from $\text{exp}(\lambda_2)$

BUT imagine you want to COMPARE a system with a “slow” exponential service rate with one that is “faster”

**System #1**

$x_{1,i} = \frac{1}{\lambda_1} \ln(1-u_i)$

**System #2**

$x_{2,i} = \frac{1}{\lambda_2} \ln(1-u_i)$

$\uparrow$ \hspace{1cm} same $u_i$ \hspace{1cm} $\uparrow$

$x_{1,i} > x_{2,i}$
3. Weibull \[ F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \]

parameters \( \lambda, k \)

Set up \( u = 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \)

\[ \Rightarrow \quad 1 - u = e^{-\left(\frac{x}{\lambda}\right)^k} \]

\[ \Rightarrow \quad \ln(1-u) = \ln\left(e^{-\left(\frac{x}{\lambda}\right)^k}\right) \]

\[ = -\frac{1}{\lambda^k} \cdot \ln\left(e^{x^k}\right) \]

\[ = -\frac{1}{\lambda^k} x^k \]

\[ x = \left(-\lambda^k \ln(1-u)\right)^{1/k} \]

\[ \Rightarrow \quad x = -\lambda \ln(1-u)^{1/k} \]
4. Geometric 
\[ F(k) = 1 - (1-p)^{k+1} \]

\[ F^{-1}(y) = \min \{ k \mid F(k) \geq y \} = \min \{ k \mid 1 - (1-p)^{k+1} \geq y \} \]

So, given \( u \sim U[0,1] \), sample

\[ F^{-1}(u) = \min \{ k \mid 1 - (1-p)^{k+1} \geq u \} \]

Observe that

\[ 1 - (1-p)^{k+1} \geq u \]
\[ \Rightarrow 1 - u \geq (1-p)^{k+1} \]
\[ \Rightarrow \log_{1-p}(1-u) \geq k+1 \]
\[ \Rightarrow \frac{\ln(1-u)}{\ln(1-p)} - 1 \geq k \]

\[ \Rightarrow \left[ \frac{\ln(1-u)}{\ln(1-p)} \right] - 1 = k \]
Normal N(0,1) distribution

inverse transform method difficult because inverting CDF is difficult

Try a transform

Refresher on polar coordinates

A point can be described as \((r, \theta)\) in polar coordinates or \((x, y)\) in Cartesian coordinates

relationship \( r^2 = x^2 + y^2 \) \( \tan(\theta) = \frac{y}{x} \)

so \( x = r \cos(\theta) \)

\( y = r \sin(\theta) \)
Consider joint density function for independent normals

\[ X \sim N(0,1), \ Y \sim N(0,1) \]

\[ p(x,y) = p(x)p(y) = \left( \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) \left( \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \right) \]

\[ = \left( \frac{1}{2\pi} \right) \left( e^{-\frac{(x^2+y^2)}{2}} \right) \]

\[ = \left( \frac{1}{2\pi} \right) \left( e^{-\frac{r^2}{2}} \right) \text{ where } r^2 = x^2 + y^2 \]

we can think of \( p(x,y) \) as the product of two different density functions
\( \frac{1}{2\pi} \) is density for uniform \((0, 2\pi)\)

and \( e^{-r^2/2} \)?

Consider CDF for \( r^2 \) defined by

\[ R\{ z \leq r^2 \} = 1 - e^{-r^2/2} \], i.e. exponential in \( r^2 \)

then density function is

\[ \frac{d}{dr} \left( 1 - e^{-r^2/2} \right) = (2)(\frac{1}{2}) e^{-r^2/2} \]

\[ = e^{-r^2/2} \]

\[ \Rightarrow \text{ we can view } e^{-r^2/2} \text{ as the exponential pdf of } x^2 + y^2 = r^2 \]

which is \( \exp(\frac{1}{2}) \)
remember we sample \( \exp(\lambda) \) by \(-\frac{1}{\lambda} \ln(1-u)\)

\[ R^2 = -2 \ln(u) \]

So we can sample "angle" component by

\[ \left( u_2 \cdot 2\pi \right) \]

So we're going from sampling \((x, y), x \sim N(0,1), y \sim N(0,1)\)

to sampling \((\Theta, R^2)\)

\[ \Theta \sim U(0, 2\pi) \quad R^2 \sim \exp\left(\frac{1}{2}\right) \]

\[ \Theta = u_2 \cdot 2\pi \quad R^2 \sim -2 \ln(u_1) \]

so

\[ R \sim \sqrt{-2 \cdot \ln(u_1)} \]
Now use polar → rectangular coordinate transformation

\[ x = r \cos(\theta) = (2 \ln(u_1))^{1/2} \cdot \cos(u_2 \cdot 2\pi) \]

\[ y = r \sin(\theta) = (2 \ln(u_1))^{1/2} \cdot \sin(u_2 \cdot 2\pi) \]

So two independent uniform random variables can be used through this transformation to obtain two independent \( N(0,1) \) random variables.
Transformation of $N(0,1) \rightarrow N(\mu, \sigma)$

Given a $N(0,1)$ normal, it is easy to create a Gaussian with mean $\mu$ and standard deviation $\sigma$.

Given $Z \sim N(0,1)$ and parameters $\mu, \sigma$, define

$$X = \sigma Z + \mu$$

and $X$ will be Gaussian with mean $\mu$ and standard deviation $\sigma$.

Conversely, if $X$ is Gaussian with mean $\mu$ and standard deviation $\sigma$, then

$$Y = \frac{X-\mu}{\sigma}$$

is standard normal.
These relationships mean that we can infer the CDF values for a Gaussian from the CDF of a standard normal.

If $X \sim N(\mu, \sigma^2)$ then

$$F_X(t) = \Pr\{X \leq t\}$$

$$= \Pr\{\sigma Z + \mu \leq t\}$$

$$= \Pr\{Z \leq \frac{t - \mu}{\sigma}\}$$

$$= \Phi\left(\frac{t - \mu}{\sigma}\right)$$
REJECTION SAMPLING

BASIC FORM

a) Sample $x$ from an "easy" distribution

b) accept $x$ if it meets some criteria
    reject otherwise

example: sample points within red polygon uniformly

1. sample $(x, y)$,
   $x \sim U(a, b)$, $y \sim U(c, d)$

2. Use geometry to determine whether $(x, y)$ falls within red polygon
    - accept or reject
Let $f, g$ be probability density functions

- $g$ is "easy" to sample from
- $f(x)$ can be computed

Suppose there is $c > 1$ such that for all $x$

\[ f(x) \leq c \cdot g(x) \]

Algorithm

1. Sample $y$ from $g$
2. Sample $z$ from $U[0, c \cdot g(y)]$
3. if $z < f(y)$, accept $y$ else reject
Pr \{ sample is accepted \} = \int_{-\infty}^{\infty} g(x) \cdot \frac{f(x)}{c \cdot g(x)} \cdot dx

so \# \text{ samples to accept is geometric } \left( \frac{1}{c} \right) = \frac{1}{c}

We are interested in the distribution of accepted samples

\[ Pr \{ Y \leq t \mid Y \text{ is accepted} \} = \frac{Pr \{ Y \leq t \& Y \text{ is accepted} \}}{Pr Y \text{ is accepted}} \]

\[ = \int_{x \leq t} g(x) \cdot \frac{f(x)}{c \cdot g(x)} \cdot dx \]

\[ = \frac{1}{c} \]

\[ = \int_{x \leq t} f(x) \cdot dx = F(t) \]
Application to Normal

Find $g(x)$, $C$, such that $\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \leq C \cdot g(x)$ for all $x$

Try $g(x)$ based on exponential

- with probability $\frac{1}{2}$ go positive,

$$g(x) = \frac{1}{2} \cdot e^{-x} \quad \text{(here } \lambda = 1, \frac{1}{2} \text{ for selection of direction)}$$

want $C$ s.t. $C \cdot g(x) \geq f(x)$ for all $x$

$$\Rightarrow C \cdot \frac{1}{2} \cdot e^{-x} \geq \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$$

$$\Rightarrow C \geq \frac{2}{\sqrt{2\pi}} \cdot e^{-x^2/2 - x} \quad \text{for all } x$$

Now $-(x^2/2 - x)$ is maximized at $x = 1$ (use calculus)

$$\Rightarrow C \geq \frac{2}{\sqrt{2\pi}} \cdot e^{1/2} = \sqrt{\frac{4 \cdot e}{2\pi}} = \sqrt{\frac{2 \cdot e}{\pi}} \approx 1.315$$
Creating uniformly random permutation

Problem: Given sequence \( d_1, d_2, d_3, \ldots, d_k \)
generate a permutation which is exactly as likely as any other

Note: there are \( k! \) permutations, so a uniformly random one will have probability \( \frac{1}{k!} \)

BASIC IDEA: Permute "in place"

1 2 3 4 5
[\[\ldots\]\[\[\ldots\]\]

already selected

\[\]

\[\]

\[\]

4 choose some element in not yet selected subarray uniformly at random

\[\]

\[\]
for i in xrange(0, k-1):
    j = random.randint(i, k-1)
    swp = L[i]
    L[i] = L[j]
    L[j] = swp

Proof of correctness:
\[ R\{ s_1, s_2, s_3, \ldots, s_k \} \]
\[ = R\{ s_1 \text{ from } k \} \cdot R\{ s_2 \text{ from } k-1 \} \cdot \cdots \]
\[ = \frac{1}{k} \cdot \frac{1}{k-1} \cdot \frac{1}{k-2} \cdot \cdots \cdot \frac{1}{2} \]
\[ = \frac{1}{k!} \]
CONVOLUTION METHODS

- Some random variates are naturally expressed as a sum of others

- Sum of $N$ Bernoulli r.v.s is a binomial
- Sum of $k$ exponentials is an Erlang-$k$

Some computational issues to consider

memory/computation

example: Binomial
- Constant (small) space, $N$ samples from $U(0,1)$, $O(N)$ time
- Use inverse dist. function
  - $O(N)$ time to create, $O(N)$ space
    - $O(\log N)$ time per subsequent sample

- Alias table: $O(N)$ space, $O(N)$ time to create, $O(1)$ time to sample
Alias Table — Big Picture

- Works for finite discrete distributions
- Given $N$ possible values, create table with $N$ entries
  
  EACH entry has form $(q_i, V_j, V_k)$

  where $q_i$ is a probability, $V_j$ and $V_k$ are possible values

To Sample
1. CHOOSE $i^{th}$ table entry $(q_i, V_j, V_k)$ uniformly at random
2. WITH probability $q_i$ select $V_j$ as sample, else $V_k$

THAT'S ALL!

NOTE — a value $V_j$ may appear in multiple entries
— the magic part is in setting up the entries
BASIC IDEA

Imagine values laid out on $[0,1]$ with length proportional to probability.

\[ \Pr \{ V_1 = \frac{2}{5} \}, \quad \Pr \{ V_2 = \frac{1}{5} \}, \quad \Pr \{ V_3 = \frac{1}{8} \}, \quad \Pr \{ V_4 = \frac{11}{40} \} \]

Key idea is spreading a sample's probability mass among equi-sized bins.

Some care needed though to limit at most 2 samples per bin (notice $V_2$, $V_3$ and $V_4$ have mass in same bin above).
CREATING ALIAS TABLE

1. Partition values into two sets
   \[ G : \text{those with probability } \leq 1/N \]
   \[ H : \text{all others} \]

2. For every \( i = 1 \) to \( N \)
   \[ r_i = \Pr \{ V_i \} \]

3. For \( i = 1 \) to \( N-1 \)
   - Choose any \( V_j \) in \( G \)
   - Set \( q_i = r_j \cdot N \) \hspace{1cm} // assured to be \(<1\)
   - Make \( V_j \) first component of \( i^{th} \) table entry

   - Choose \( V_k \in H \), make 2\(^{nd} \) component

   - Set \( r_k = r_k - (1-q_i) / N \) // subtract allocation to \( i \)
   - Last table entry take single value left in \( H - (1,V_c,-) \)

Note on reset of \( r_k \): subtract off product of selecting entry \((1/N)\)
 times selecting \( V_k \)
**Example Alias Table Construction**

\[ Pr \{ V_1 = \frac{2}{5} \}, \quad Pr \{ V_2 = \frac{1}{5} \}, \quad Pr \{ V_3 = \frac{1}{8} \}, \quad Pr \{ V_4 = \frac{11}{40} \} \]

1. \( G = \{ V_2, V_3 \} \), \( H = \{ V_1, V_4 \} \)
\( r_1 = 2/5, \quad r_2 = 1/5, \quad r_3 = 1/8, \quad r_4 = 11/40 \)

2. **Table construction**

<table>
<thead>
<tr>
<th>From G</th>
<th>From H</th>
<th>entry</th>
<th>residual</th>
<th>action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>( V_2 )</td>
<td>( V_1 )</td>
<td>( \left( \frac{4}{5}, V_2, V_1 \right) )</td>
<td>( \frac{2}{5} - \frac{1}{20} = \frac{7}{20} )</td>
</tr>
<tr>
<td>2.</td>
<td>( V_3 )</td>
<td>( V_1 )</td>
<td>( \left( \frac{1}{2}, V_3, V_1 \right) )</td>
<td>( \frac{7}{20} - \frac{1}{8} = \frac{9}{40} )</td>
</tr>
<tr>
<td>3.</td>
<td>( V_1 )</td>
<td>( V_4 )</td>
<td>( \left( \frac{9}{10}, V_1, V_4 \right) )</td>
<td>( \frac{11}{40} - \frac{1}{40} = \frac{10}{40} )</td>
</tr>
<tr>
<td>4.</td>
<td>( V_4 )</td>
<td>-</td>
<td>( (1.0, V_4, -) )</td>
<td></td>
</tr>
</tbody>
</table>