STATISTICS FOR DATA SCIENCE

Generation of Random Variates

Prof.D. Uma

Computer Science and Engineering

STATISTICS FOR DATA SCIENCE

Generation of Random Variates

Prof.D. Uma
Department of Computer Science and Engineering

Topics to be covered...

- Random Numbers
- Random Variate Generator
- Random Variates
- Techniques for Generating Random Variates

Random Numbers

- Random numbers are very important for a simulation.

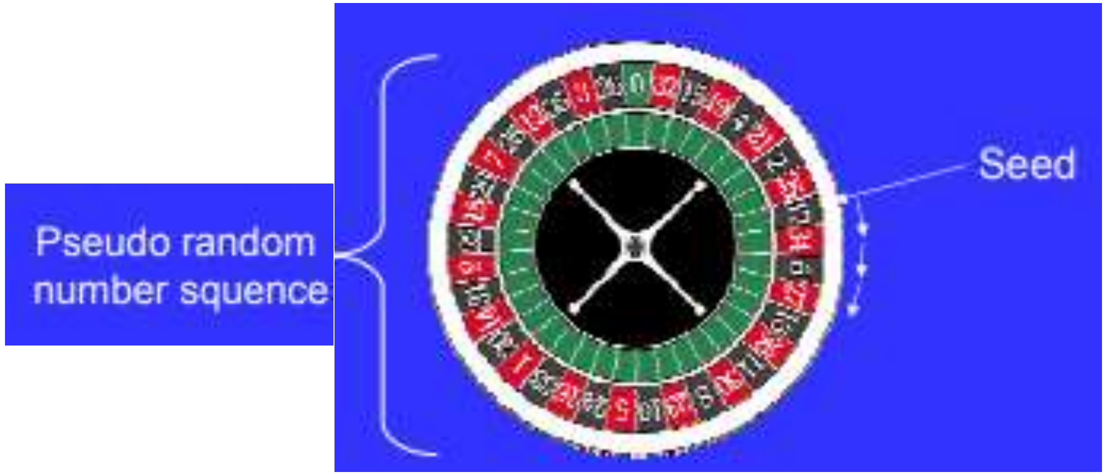
- Random number generator output: Sequence of independent and identically (uniformly) distributed random numbers between 0 and 1.

- These random numbers are transformed into required probability distributions.

Random Number Seed

- Computer-based generators use random number seeds for setting the starting point of the random number sequence.

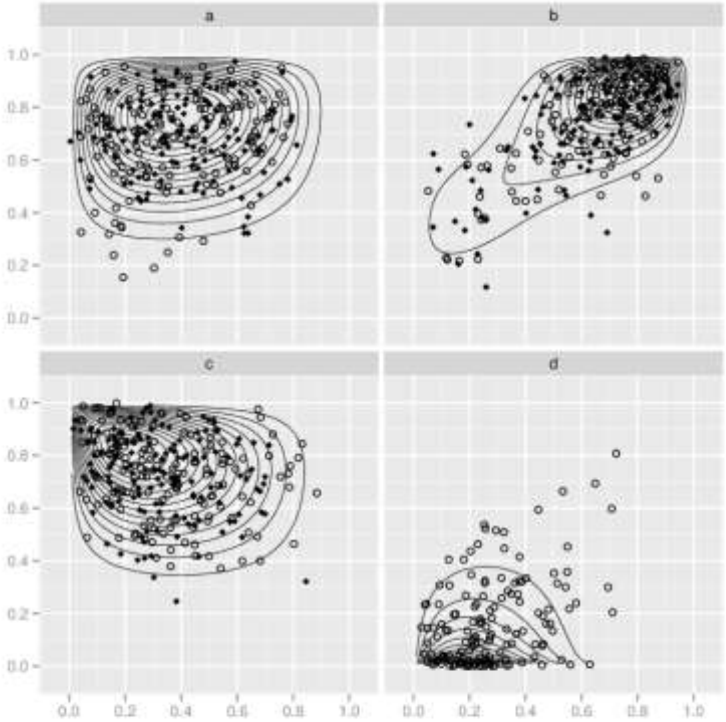
- These seeds are often initialized using a computer's real time clock in order to have some external noise.



Source: http://www.cs.bilkent.edu.tr/~cagatay/cs503/_M&S_04_Random_Variate_Generation.pdf

Random Variate Generation

It is assumed that a distribution is completely specified and we wish to generate samples from this distribution as input to a simulation model.



Source: mathworld.wolfram.com , www.slideshare.net,ResearchGate, Lambert

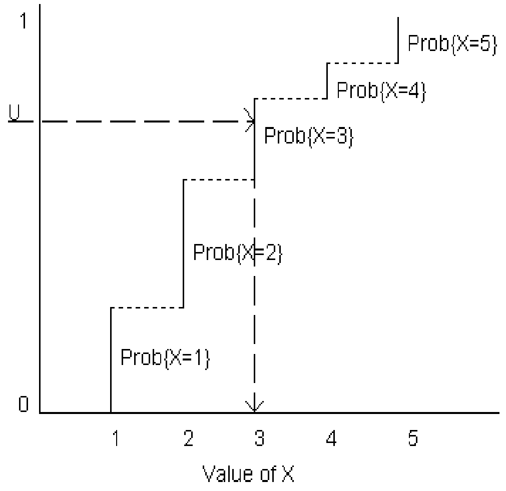
Random Variate Generation

- Process of producing observations that have the distribution of the given random variables.

- This is to develop simulation models for the purpose of analysis and decision making.

- It rely on generating uniformly distributed random number on the interval (0,1).

- Random variate generators use as starting point random numbers distributed U[0,1].



Random Number Generators

- A computational or physical device designed to generate a sequence of numbers that lack any pattern (i.e. appear random).
- Computer-based generators are simple deterministic programs trying to fool the user by producing a deterministic sequence that looks random (pseudo random numbers).
- They should meet some statistical tests for randomness intended to ensure that they do not have any easily discernible patterns.

Source: http://www.cs.bilkent.edu.tr/~cagatay/cs503/_M&S_04_Random_Variate_Generation.pdf

Random Variate

A **random variate** is a variable generated from uniformly distributed pseudorandom numbers.

Depending on how they are generated, a **random variate** can be uniformly or non-uniformly distributed.

Random variates are frequently used as the input to simulation models.

Examples: Inter-arrival time and service time.

Source: mathworld.wolfram.com , www.slideshare.net

Random Variate

RV Generators – Techniques used to generate random variates.

- Inverse transform technique
- Direct transformation for the Normal Distribution
- Convolution Method
- Acceptance and Rejection Technique

Source: mathworld.wolfram.com , www.slideshare.net

Steps in Inverse Transform technique

1. Compute CDF of the desired random variable X.
2. Set F(X)=R on the range of X.
3. Solve the equation F(X)=R for X in terms of R.
4. Generate uniform random numbers R1,R2,R3… and compute the desired random variate by

$$X_i = F^{-1}(R_i)$$

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Steps in Inverse Transform Technique

Inverse transform method – Uniform Distibution Example:

Step 1 – compute *cdf* of the desired random variable X

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x < b \\ 1, & x \ge b \end{cases}$$

Step 2 – Set F(X) = R where R is a random number ~U[0,1)

$$F(x) = R = \frac{x-a}{b-a}$$

Step 3 – Solve F(X) = R for X in terms of R. $X = F^{-1}(R)$.

$$R(b-a) = X - a, \quad X = R(b-a) + a$$

Step 4 – Generate random numbers R_i and compute desired random variates:

$$X_i = R_i(b\text{-}a) + a$$

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Generation of Bernoulli and Binomial Random Variates

Generation of Bernoulli and Binomial Random Variates

Generation of Poisson Random Variate

Generation of Poisson Random Variate

Generation of Normal Random Variate

Generation of Normal Random Variate

Inverse-transform Technique: Other Distributions

Examples of other distributions for which inverse CDF works are:

- Uniform distribution
- Weibull distribution
- Triangular distribution

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Inverse-transform Technique: 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
- Geometric

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Inverse-transform Technique: Continuous Distributions

A number of continuous distributions do not have a closed form expression for their CDF, e.g. Normal, Gamma and Beta.

The presented method does not work for these distributions.

Solution

- Approximate the CDF or numerically integrate the CDF

Problem

- Computationally slow

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Acceptance and Rejection Technique

- Useful particularly when inverse CDF does not exist in closed form

- **Illustration:** To generate random variates, X ~ U(1/4,1)

Procedure:

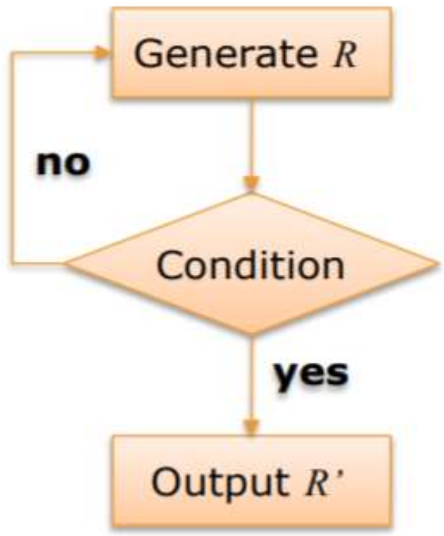
Step 1. Generate $R \sim U(0,1)$

Step 2. If $R \geq 1/4$, accept $X=R$.

Step 3. If $R < 1/4$, reject R, return to Step 1

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Acceptance and Rejection Technique



R does not have the desired distribution, but R conditioned (R’) on the event {R ≥ ¼} does.

- Efficiency: Depends heavily on the ability to minimize the number of rejections.

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Acceptance and Rejection Technique – Poisson Distribution

Procedure of generating a Poisson random variate N is as follows

1. Set n=0, P=1
2. Generate a random number R_{n+1}, and replace P by $P \times R_{n+1}$
3. If $P < \exp(-\alpha)$, then accept $N=n$
   - Otherwise, reject the current n, increase n by one, and return to step 2.

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Acceptance and Rejection Technique – Poisson Distribution

- Example: Generate three Poisson variates with mean α=0.2
  - exp(-0.2) = 0.8187
- Variate 1
  - Step 1: Set $n = 0, P = 1$
  - Step 2: $R1 = 0.4357$, $P = 1 \times 0.4357$
  - Step 3: Since $P = 0.4357 < \exp(-0.2)$, accept $N = 0$
- Variate 2
  - Step 1: Set $n = 0, P = 1$
  - Step 2: $R1 = 0.4146$, $P = 1 \times 0.4146$
  - Step 3: Since $P = 0.4146 < \exp(-0.2)$, accept $N = 0$
- Variate 3
  - Step 1: Set $n = 0, P = 1$
  - Step 2: $R1 = 0.8353$, $P = 1 \times 0.8353$
  - Step 3: Since $P = 0.8353 > \exp(-0.2)$, reject $n = 0$ and return to Step 2 with $n = 1$
  - Step 2: $R2 = 0.9952$, $P = 0.8353 \times 0.9952 = 0.8313$
  - Step 3: Since $P = 0.8313 > \exp(-0.2)$, reject $n = 1$ and return to Step 2 with $n = 2$
  - Step 2: $R3 = 0.8004$, $P = 0.8313 \times 0.8004 = 0.6654$
  - Step 3: Since $P = 0.6654 < \exp(-0.2)$, accept $N = 2$

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Acceptance and Rejection Technique – Poisson Distribution

- It took five random numbers to generate three Poisson variates
- In long run, the generation of Poisson variates requires some overhead!

N	R_{n+1}	P	Accept/Reject		Result
0	0.4357	0.4357	$P < \exp(-\alpha)$	Accept	N=0
0	0.4146	0.4146	$P < \exp(-\alpha)$	Accept	N=0
0	0.8353	0.8353	$P \geq \exp(-\alpha)$	Reject	
1	0.9952	0.8313	$P \geq \exp(-\alpha)$	Reject	
2	0.8004	0.6654	$P < \exp(-\alpha)$	Accept	N=2

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Direct Transformation

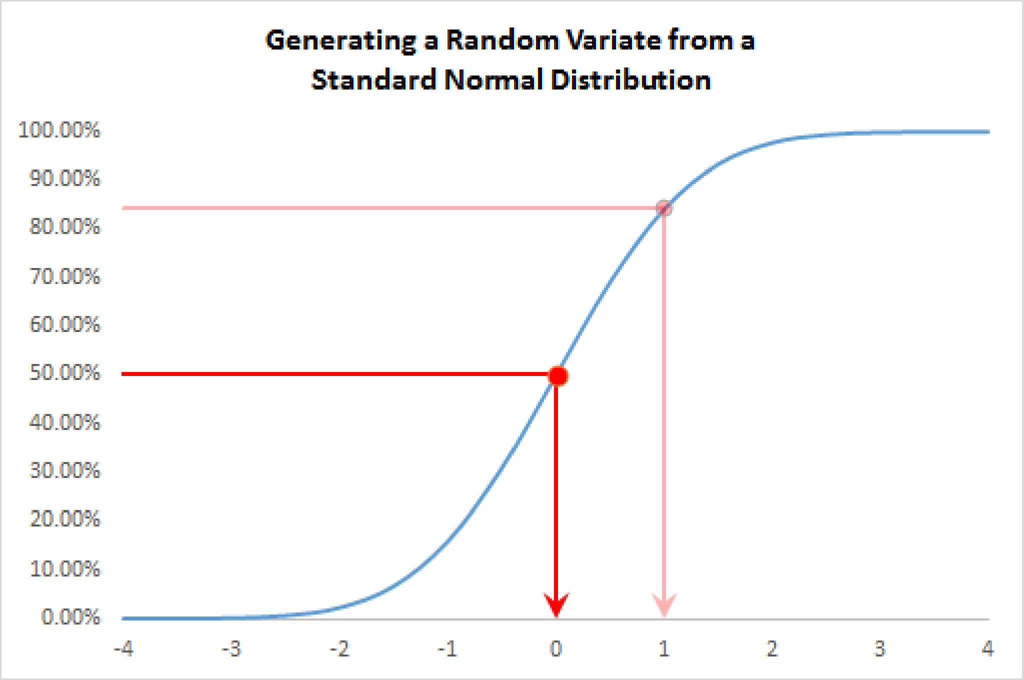
Approach for $N(0,1)$

- PDF

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

- CDF, No closed form available

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$



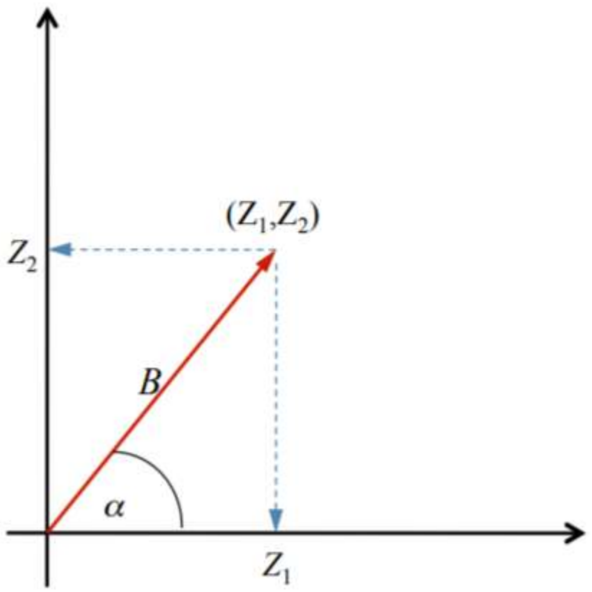
Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf, tvmcalcs.com

Direct Transformation

Approach for N(0,1)

- Consider two standard normal random variables, Z1 and Z2, plotted as a point in the plane:

- In polar coordinates:
  - Z1 = B cos(α)
  - Z2 = B sin(α)



Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Direct Transformation

- Approach for $N(\mu, \sigma^2)$:
  - Generate $Z_i \sim N(0,1)$

$$\boxed{X_i = \mu + \sigma Z_i}$$

- Approach for Lognormal(μ,σ^2):
  - Generate $X \sim N(\mu,\sigma^2)$

$$\boxed{Y_i = e^{X_i}}$$

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Direct Transformation-Example

Let R1 = 0.1758 and R2=0.1489

- Two standard normal random variates are generated as follows:

$$Z_1 = \sqrt{-2\ln(0.1758)}\cos(2\pi 0.1489) = 1.11$$

$$Z_2 = \sqrt{-2\ln(0.1758)}\sin(2\pi 0.1489) = 1.50$$

- To obtain normal variates Xi with mean μ=10 and variance σ^2 = 4

$$X_1 = 10 + 2 \cdot 1.11 = 12.22$$

$$X_2 = 10 + 2 \cdot 1.50 = 13.00$$

Source:https://www.mi.fu-berlin.de/inf/groups/ag-tech/teaching/2012_SS/L_19540_Modeling_and_Performance_Analysis_with_Simulation/07.pdf

Do It Yourself !!!!

Implement Random Variate Generation for **Poisson Distribution.**

Implement Random Variate Generation for **Normal Distribution**.

THANK YOU

D. Uma
Department of Computer Science and Engineering
umaprabha@pes.edu
+91 99 7251 5335