Random Variate Generation (Part 3)

Dr.Çağatay ÜNDEĞER

Öğretim Görevlisi
Bilkent Üniversitesi Bilgisayar Mühendisliği Bölümü
&
...

e-mail : cagatay@undeger.com
cagatay@cs.bilkent.edu.tr

Random Variate Generation (Outline)

- Random numbers
- Random number generators
- Random variate generation
  - Factors to be considered
  - General principles
    - Inverse Transform Method
    - Acceptance-Rejection Method
    - Composition Method
    - Relocate and Rescale Method
  - Specific distributions

Random Numbers

- A random sequence of numbers obtained from a stochastic process.
- In real world, random numbers may be generated using a dice or a roulette wheel.

Random Numbers

- Random numbers are very important for a simulation,
- Since all the randomness required by the model is simulated by a random number generator,
  - Whose output is assumed to be a sequence of independent and identically (uniformly) distributed random numbers between 0 and 1.
- Then these random numbers are transformed into required probability distributions.

Random Number Generator

- A computational or physical device designed to generate a sequence of numbers that lack any pattern (i.e. appear random).

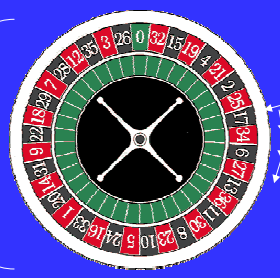
Random Number Generators

- Computer-based generators are simple deterministic programs trying to fool the user by producing a deterministic squence that looks random (pseudo random numbers).
- Therefore, they should meet some statistical tests for randomness intended to ensure that they do not have any easily discernible patterns.

Random Number Seed

- Computer-based generators use random number seeds for seting the starting point of the random number squence.
- These seeds are often initialized using a computer's real time clock in order to have some external noise.

Pseudo random number squence



Seed

Random Variate Generation

- Refers to the generation of variates whose probability distribution is different from the uniform distribution on the interval (0,1).

Factors to Be Considered

- Exactness:
  - Exact if distribution of variates generated has the exact form desired.
- Speed:
  - Related with computing time required to generate variate. Contributions to time are:
    - Setup time,
    - Variable generation time.

Factors to Be Considered

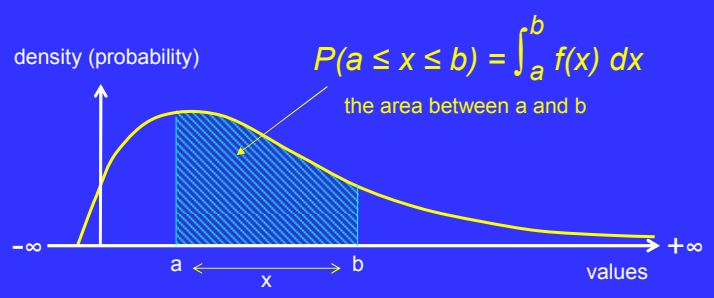
- Space:
  - Computer memory required for the generator.
- Simplicity:
  - Refers to both algorithmic simplicity and implementation simplicity.

General Principles

- We assume that a pseudo random number generator *RN(0,1)* producing a sequence of independent values between 0 and 1 is available.
- General methods:
  - Inverse transform method
  - Acceptance-rejection method
  - Composite method
  - Translations and other simple transforms

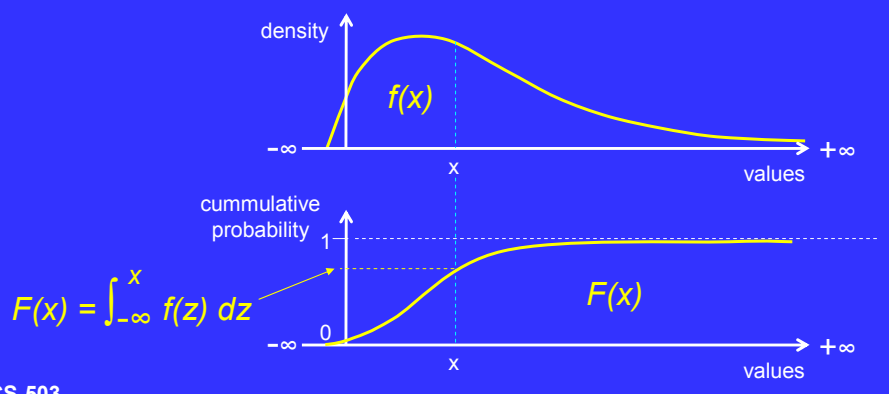
Inverse Transform Method

- A continuous distribution can be defined by its density function *f(x)*.
- The probability that a value x lies between a and b, where a>b, is the integral of function between a and b.



Inverse Transform Method

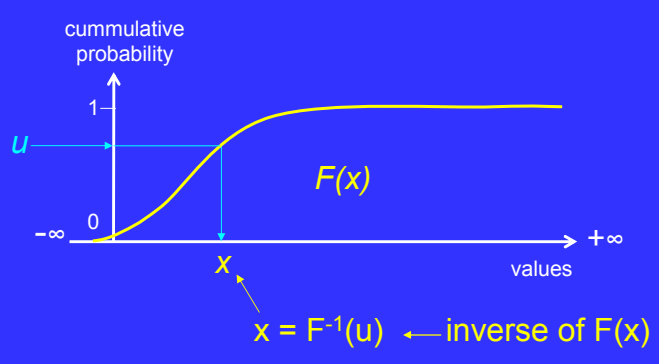
- The cummulative distribution function can be generated for all x by computing the integral of function between -infinite and x.



$$F(x) = \int_{-\infty}^{x} f(z)\, dz$$

Inverse Transform Method

- Function is strictly increasing and contiuous.
- $F(x) = u$ gives a unique x.

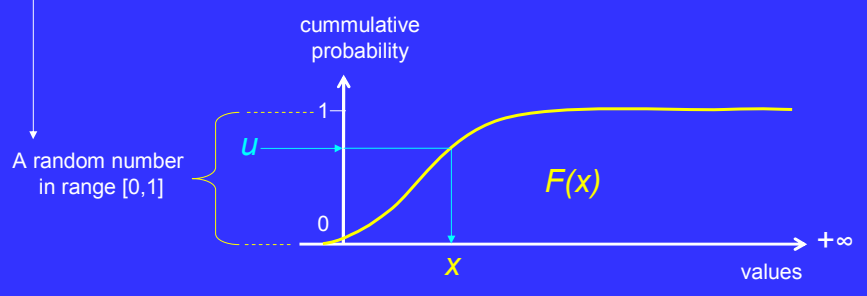


Inverse Transform Method

- For generating random variates with f(x):

  Let u = RN(0,1)

  Return x = $F^{-1}(u)$



Inverse Transform Method (Sample)

- For exponential distribution:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0, \\ 0 & , x < 0. \end{cases}$$ ← Probability density function

$$F(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0, \\ 0 & , x < 0. \end{cases}$$ ← Cummulative distribution function

↓

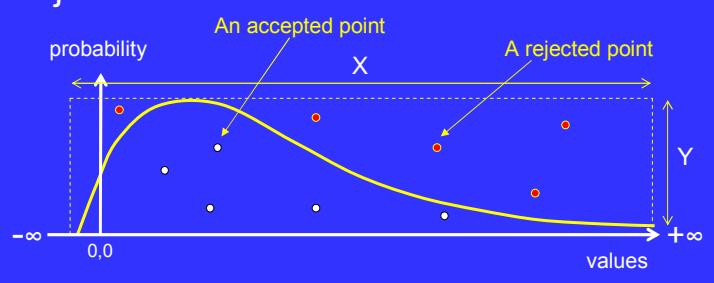
$x = F^{-1}(x,\lambda) = -\ln(1-u)/\lambda$

Acceptance-Rejection Method

- Suppose that we need to sample from a distribution whose inverse function is hard to solve.
- In that case, acceptance-rejection method can be used.

Acceptance-Rejection Method

- Generate a random point (X,Y) on the graph.
- If (X,Y) lies under the graph of f(X) then
  Accept X
- Otherwise
  Reject X

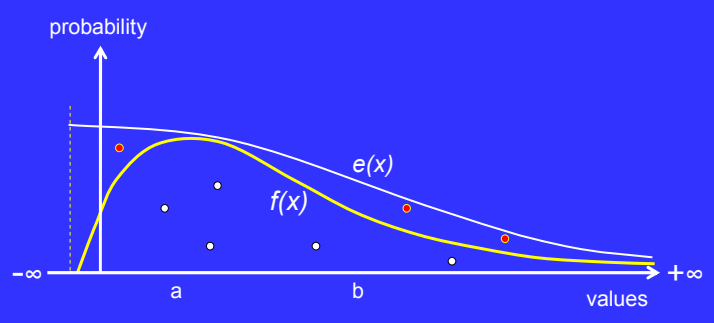


Acceptance-Rejection Method (Drawback)

- Trials ratio: Average number of points (X,Y) needed to produce one accepted X.
- We need to make trial ratio close to 1.
- Else generator may not be efficient enough because of wasted computating effort.

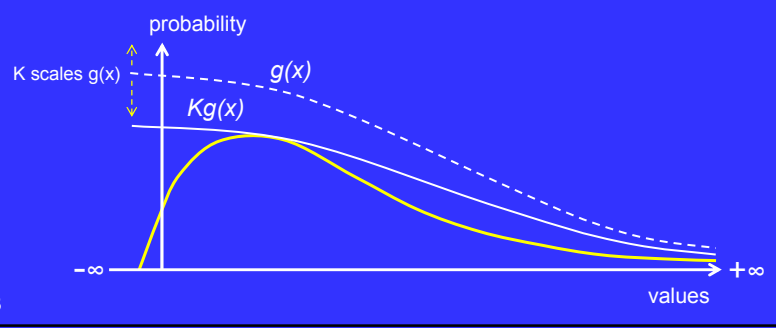
Acceptance-Rejection Method (Making More Efficient)

- One way to make generator efficient is;
  - To generate points uniformly scattered under a function *e(x)*, where area between the graph of *f* and *e* be small.



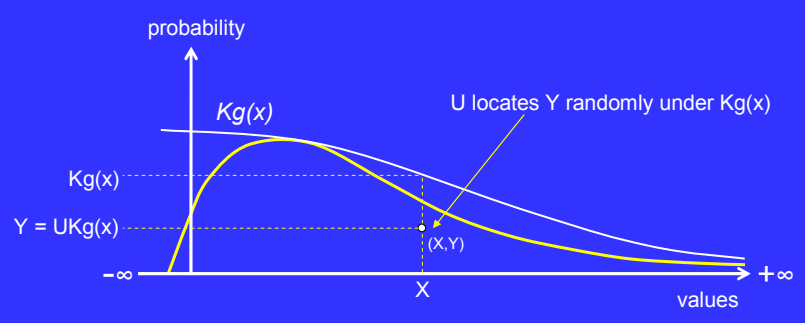
Acceptance-Rejection Method (Constructing e(x))

- Take *e(x) = Kg(x)*
- *g(x) = density function of a distribution for which an easy way of generating variates already exists.*
- *K = scale factor*



Acceptance-Rejection Method (Producing (X,Y))

- Let X = a variate produced from Kg(x)
- Let U = *RN(0,1)*
- (X,Y) = (X, UKg(X))



Composition Method

- Soppose that density f is written as weighted sum of r other density functions (f_i),
  - Where weights (p_i) satisfy:
    - *All* $p_i > 0$
    - Sum of all $p_i = 1$

A mixture or a compound density

$$f(x) = \sum_{i=1}^{r} p_i f_i(x)$$

Composition Method (Producing Variates)

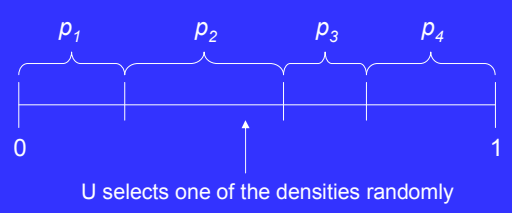
Let U = *RN(0,1)*

Set i = 1

While ($F_i < U$), where $F_i = \sum_{k=1}^{i} p_k$

$i = i+1$

Return X = a variate produced from f_i

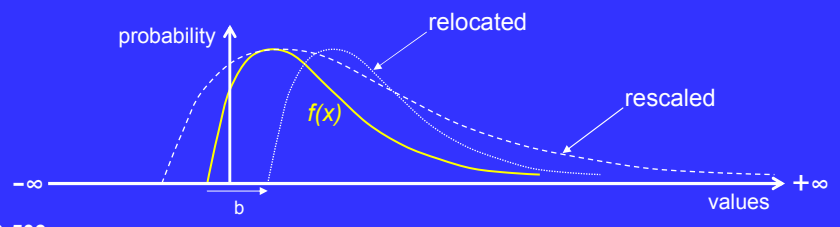


Relocate and Rescale Method

- If X is a random variate drawn from a continuous distribution *f(x)*,
  - We can reposition and rescale the distribution to get a random variate Y.

$$Y = aX + b$$

rescale relocate



Bernoulli Distributions

- Density:
  - Value 1 with success probability *p*, and
  - Value 0 with failure probability $q = 1 - p$.
- Random variate production:

Let U = *RN(0,1)*

If ($U \leq p$)

  Return $X = 1$

Else

  Return $X = 0$

Uniform Distributions

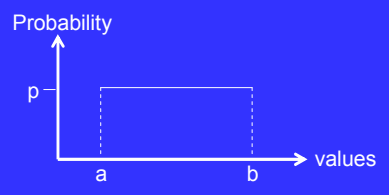
- Density:
  - Values between a and b, where a>b, and probabilities of all the values are equal.
- Production (relocate & rescale):

  Let U = *RN(0,1)*

  Return $X = a + (b-a)U$

for discrete values:

$a + min(int[(b-a+1)U] , b-a)$



Triangular Distributions

- Density:
  - A distribution with lower limit *a*, mode *c*, and upper limit *b*.
- Production (inverse transform, relocate & rescale):

  Let $\beta = (c-a)/(b-a)$

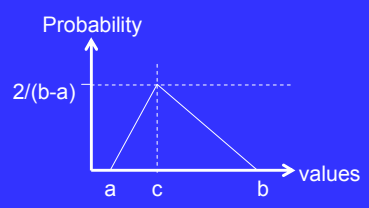
  Let $U_1 = RN(0,1)$, $U_2 = RN(0,1)$

  If $(U_1 \leq \beta)$

  Return $X = a + (c-a)\sqrt{U_2}$

  Else

  Return $X = b - (b-c)\sqrt{U_2}$



Symmetric Triangular Distributions

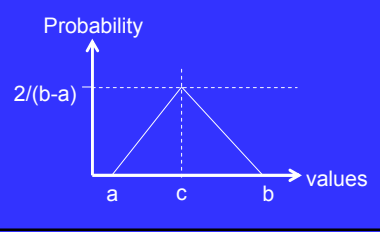
- Density:
  - A triangular distribution with c located at the center of a and b.
- Production (probability theory):
  - If U_1 and U_2 are uniformly distributed between 0 and 1 then
    - $(U_1+U_2)/2$ has a symmetric triangular distribution between 0 and 1.

Symmetric Triangular Distributions

- Density:
  - A triangular distribution with c located at the center of a and b.
- Production (probability t., relocate & rescale):

  Let $U_1 = RN(0,1)$, $U_2 = RN(0,1)$

  Return $X = a + (b-a)(U_1+U_2)/2$



Normal (Gaussian) Distributions

- Density:
  - A bell-like shaped function with μ and σ.
- Production (special transform):

  While (true)

  Let $U_1 = RN(0,1)$, $U_2 = RN(0,1)$

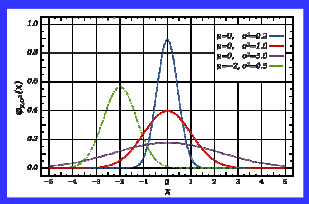
  Let $V_1 = 2U_1-1$, $V_2 = 2U_2-1$

  Let $W = V_1^2+V_2^2$

  If (W<1)

  Let $Y = \sqrt{(-2lnW)/W}$

  Return $X = \mu + \sigma V_1 Y$



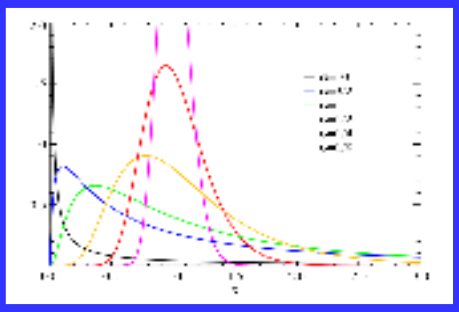
Log-Normal Distributions

- Density:
  - Single-tailed distribution of any random variable whose logarithm is normally distributed.
- Production (special transform):
  - Has no simple closed form.
  - Use the property that;
    - If Y is normal(μ, σ^2) then

      $X = e^Y$ is log-normal(μ, σ^2)

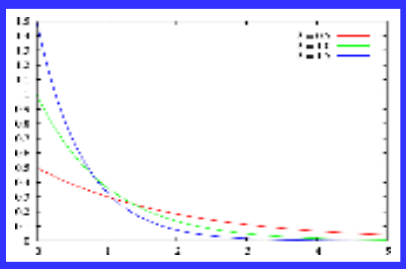
Log-Normal Distributions

- Production (special transform):
  Let Y = Normal(μ, σ^2)
  Return $X = e^Y$



Exponential Distributions

- Density:
  – *D*escribes the times between events in a Poisson process with rate parameter, λ.
- Production (inverse transform, rescale):
  Let U = *RN(0,1)*
  Return $X = -1/\lambda \ln(1-U)$

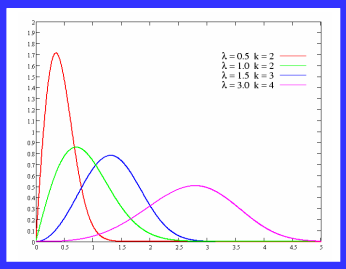


Weibull Distributions

- Density:
  - Describe the size distribution of particles with shape, k, and scale, λ.
- Production (inverse transform, rescale):

  Let U = *RN(0,1)*

  Return $X = \lambda\,[-\ln(1-U)]^{1/k}$



Poisson Distributions

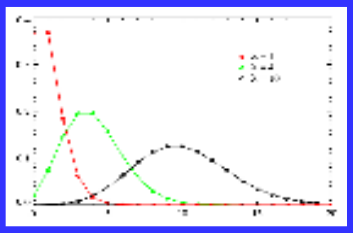
- Density:
  - *P*robability of a number of events occurring in a fixed period of time with;

    k : number of occurrences of an event

    $\lambda > 0$: expected number of occurrences in the fixed interval.

The probability that there are exactly *k* occurrences of event is equal to

$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!},$$



Poisson Distributions

- Production (counting events):

  Let $a = e^{-\lambda}$

  Let $p = 1$

  Let $k = -1$

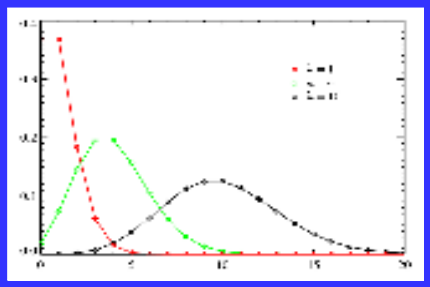
  While ($p > a$)

    Let $U = RN(0,1)$

    Let $p = pU$

    Let $k = k+1$

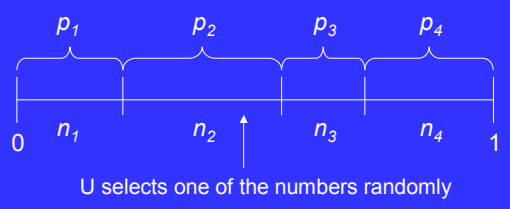
  Return k



Discrete Integer-Valued Random Numbers

- Assume we have n integer numbers with probability of being selected p_i for each

- How can we write a random variate generator?

Discrete Integer-Valued Random Numbers

- Build a table for cummulative probabilities (P_i) of n numbers.
- Use $U = RN(0,1)$ to select n_i from the table with $P_{i-1} < U \leq P_i$ (let $P_0 < 0$).