Introduction to Monte Carlo Methods

Sarath Seneviratne

(B.Sc (Hons), M.Sc, PG.Dip.Comp, M.Med.Phy)

Contact email: Sarath@alumi.rpi.edu

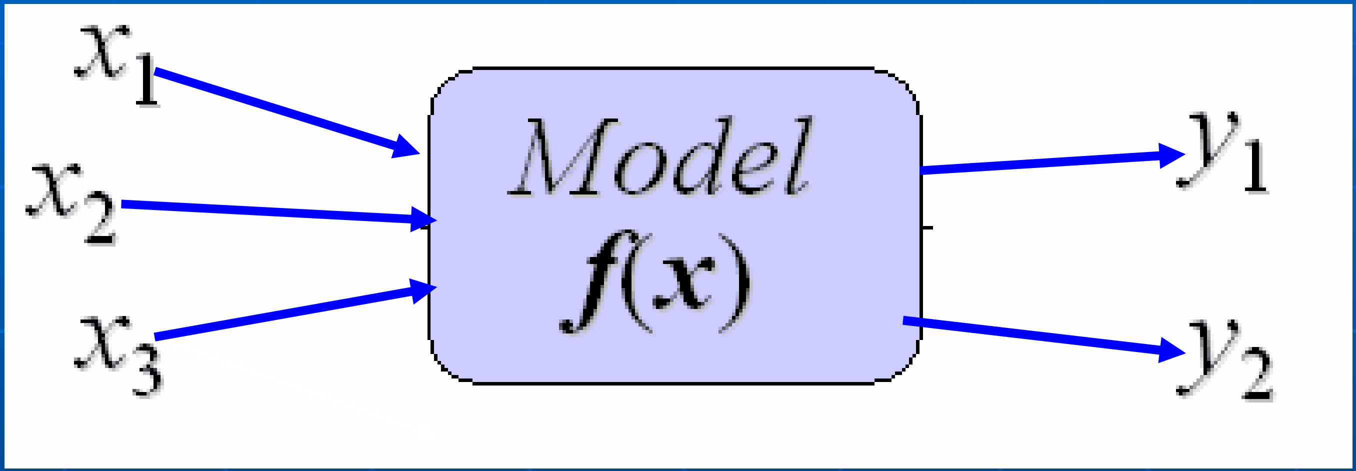
Monte Carlo Method

- Monte Carlo Method: any technique of statistical sampling employed to approximate solutions to quantitative problems
- Stanislaw Ulam (Polish-American mathematician) who worked with John Neumann and Nicholas Metropolis on the "Manhattan Project" had been recognised as the first scientist used the computers to transform non-random problems into random forms that would facilitate their solution via statistical sampling.
- Nicholas Metropolis and Stanislaw Ulam published first paper on the Monte Carlo method in 1949

- MC Simulation plays a major role in various phases of experimental physics projects
  - design of the experimental set-up
  - evaluation and definition of the potential physics output of the project
  - evaluation of potential risks to the project
  - assessment of the performance of the experiment
  - development, test and optimization of reconstruction and physics analysis and data acquisition software
  - contribution to the calculation and validation of physics results
  - Estimation of non-measurable events

Introduction

- Deterministic Model



Produce the same output for a give stating conditions

- For example, return on an investment

$$F = P\left[1 + \frac{r}{m}\right]^{y}$$

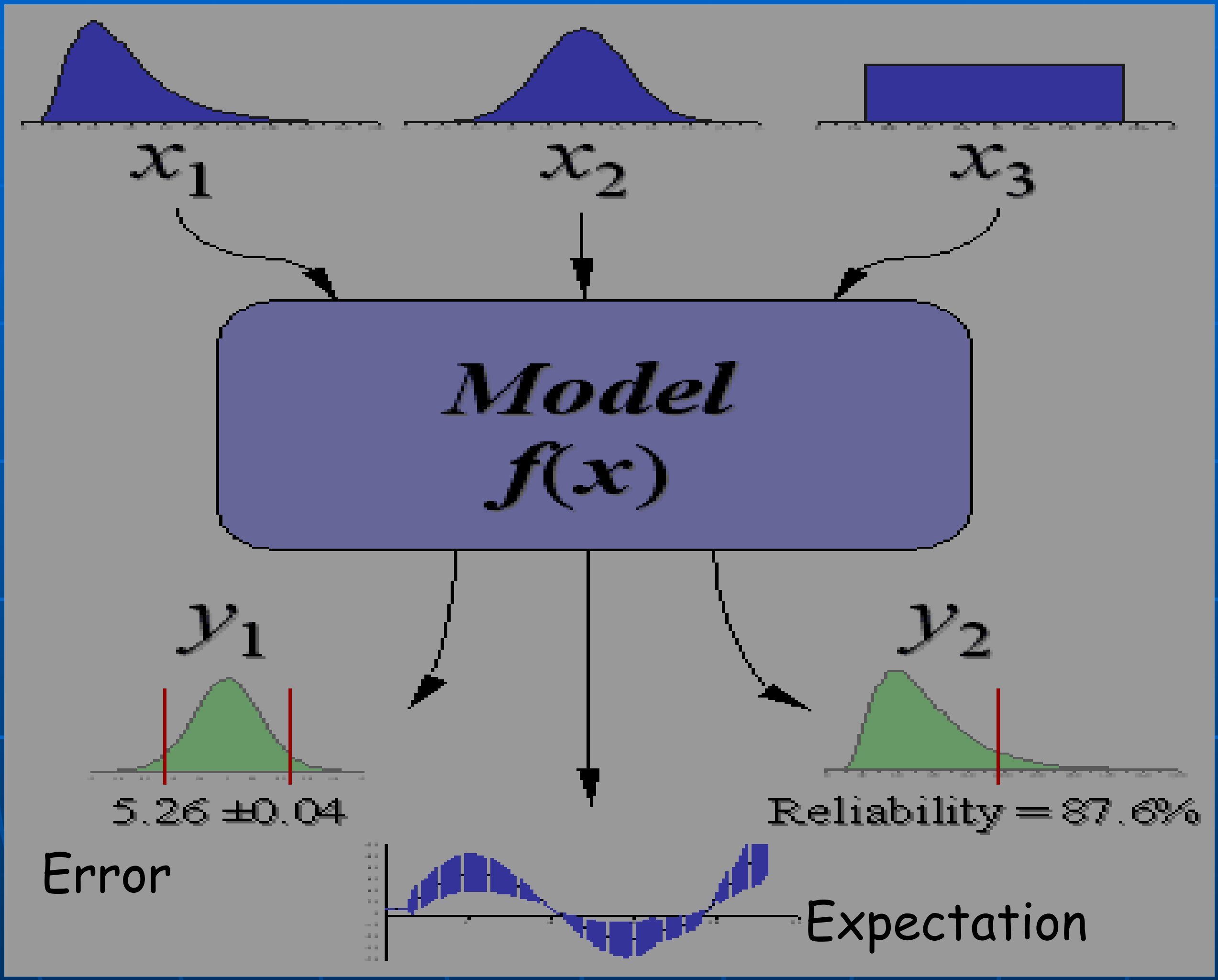
P – Initial investment

r – annual interest rate

m – compounding period

y – number of years

- Stochastic Model



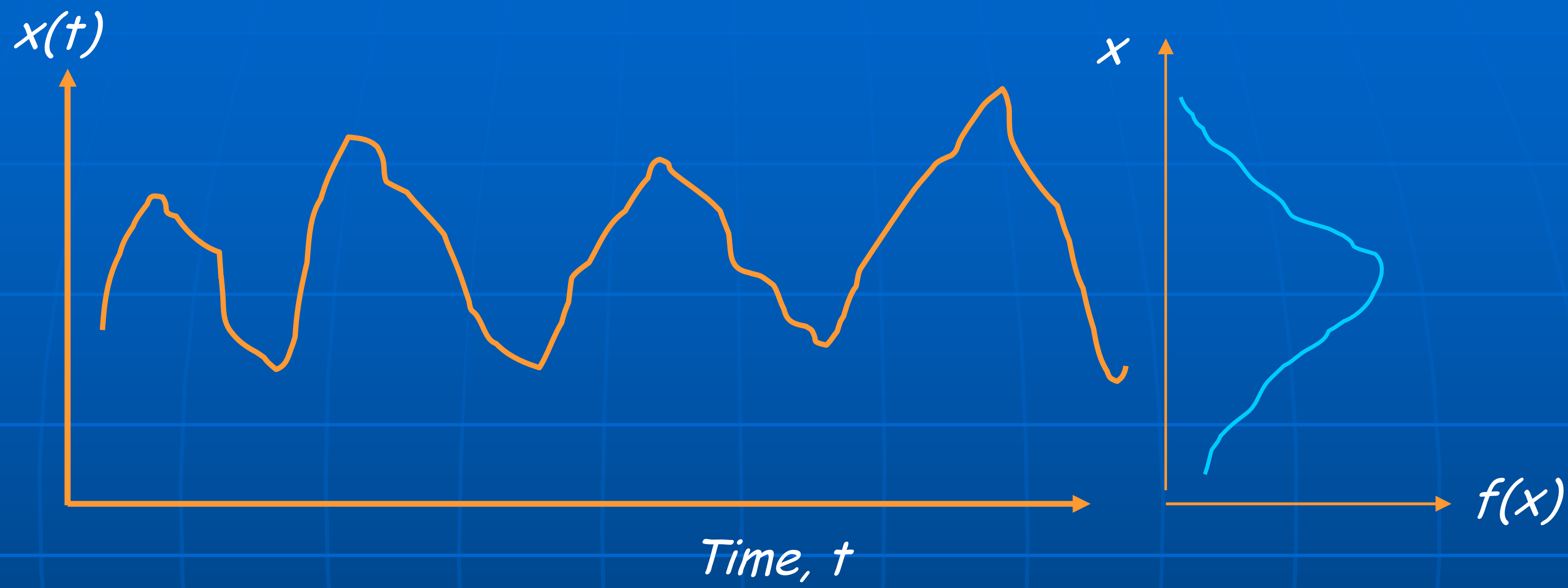
Examples:

- Stock market
- Particle interaction with matter

Random Process

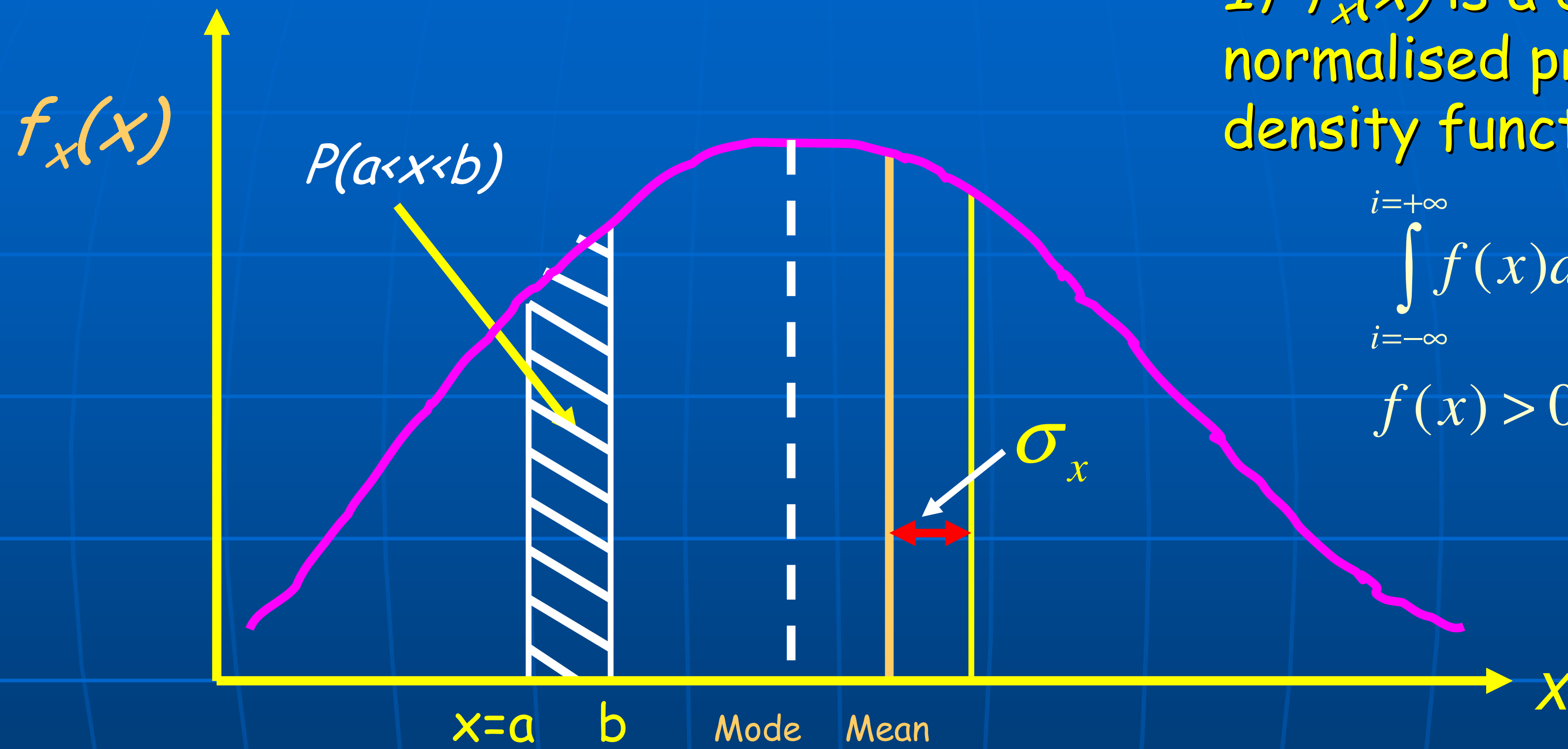
- Described in probabilistic terms
  All the properties are given as averages
  - Ensemble averaging:
    Properties of the process are obtained by averaging over a collection or 'ensemble' of sample records using values at corresponding times
  - Time averaging:
    Properties are obtained by averaging over a single record in time
- Basic properties of a single random process
  mean, standard deviation, auto-correlation, spectral density
- Joint properties of two or more random processes
  correlation, covariance, cross spectral density, simple input-output relations
- Stationary random process :
  Ensemble averages do not vary with time

Random Process



- The probability density function describes the general distribution of the magnitude of the random process, but it gives no information on the time or frequency content of the process

Probability Density Function



If $f_x(x)$ is a continuous normalised probability density function

$$\int_{i=-\infty}^{i=+\infty} f(x)dx = 1$$

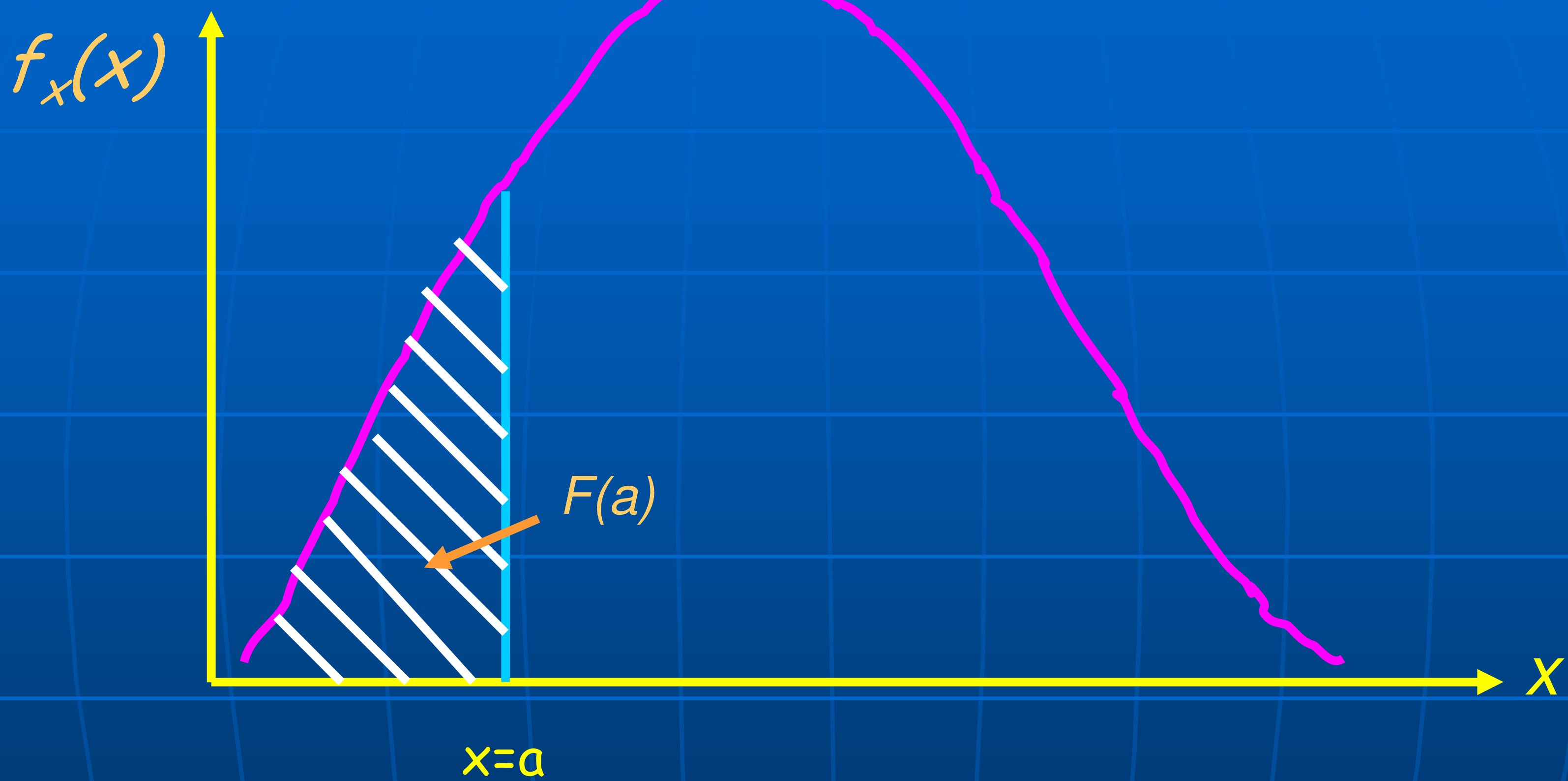
$$f(x) > 0 \quad for\ all\ x$$

$\bar{x}$

Probability that x lies between the values a and b is the area under the graph of $f_x(x)$ defined by x=a and x=b

$$P\{a < x < b\} = \int_a^b f_x(x)dx$$

Cumulative distribution function

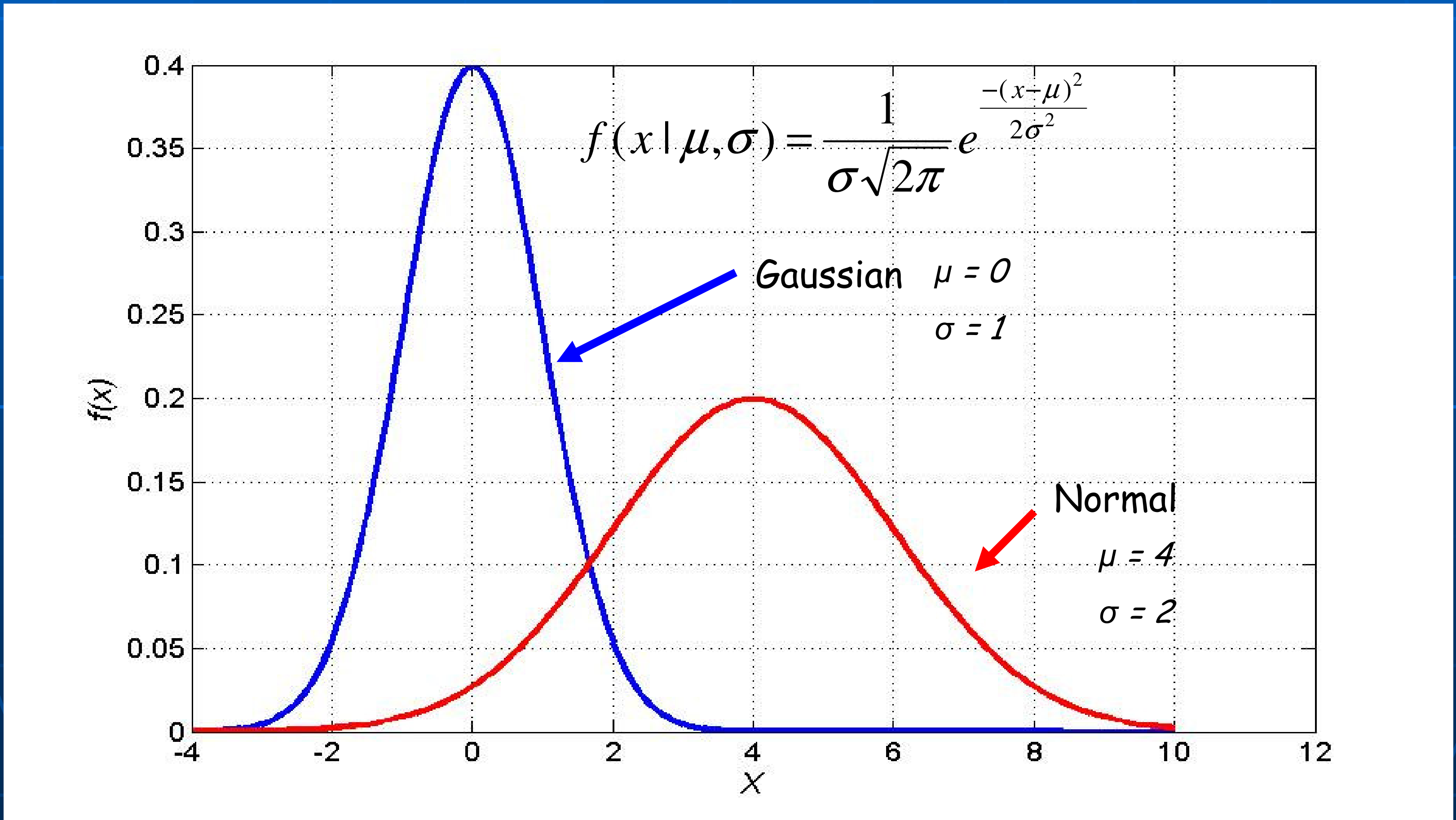


The cumulative distribution function (c.d.f.) is the integral between -∞ and x of $f_X(x)$

$$F_x(x) = \int_{x=-\infty}^{x} f_x(x)dx$$

Central Limit Theorem

- Sample averages from a random process with finite mean and variance always approximate a normal distribution
- Another interpretation, the sum of independent samples from any distribution with finite mean and variance converges to the normal distribution as the sample size goes to infinity.



Mean and Variance

- The expected value (mean) of discrete variables

  If x is a discrete random variable with possible values $x_1, x_2, x_3, \ldots.. x_N$

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Sample Variance $$\sigma_x^{\ 2} \equiv \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

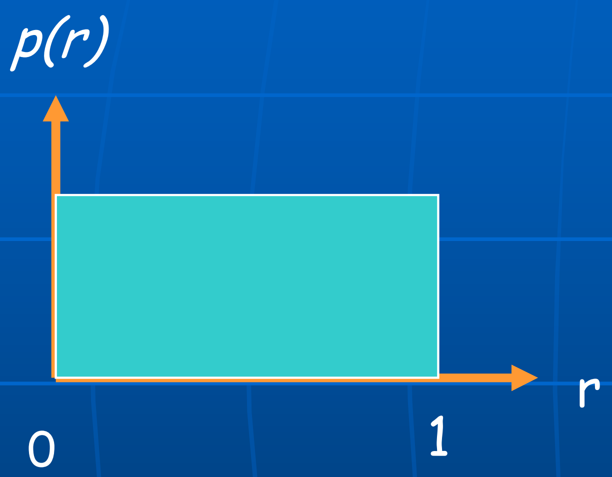
If x is a continuous random variable with probability density function $f(x)$, then the mean value is

$$\bar{x} = \int_{i=-\infty}^{i=+\infty} x f(x) dx$$

Sample Variance $$\sigma_x^{\ 2} = \int_{-\infty}^{\infty} (x - \bar{x})^2 f_x(x) dx$$

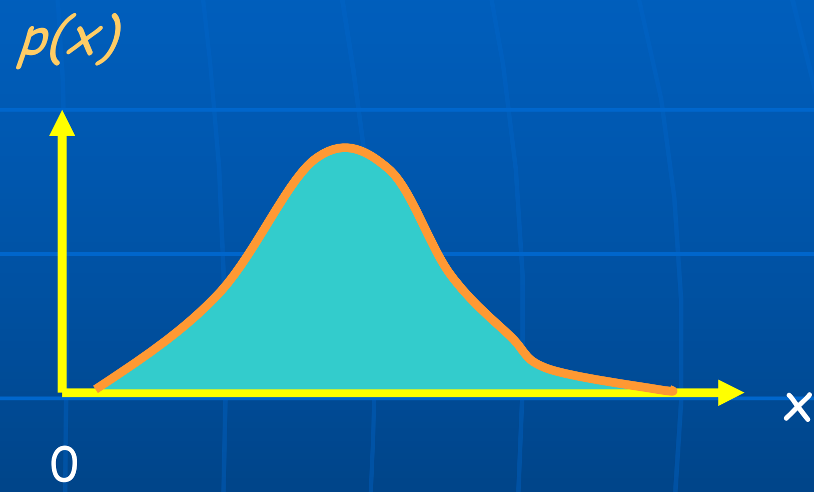
Random Sampling

Uniform distribution in sampling space



$$p(r) = 1 \quad for\, 0 \le r < 1$$
$$= 0 \quad otherwise$$

Probability distribution in which samples are being drawn



$$p(x) \ge 0 \quad for\ all\ x$$

Random Numbers

- Pseudorandom numbers
  - Generated by deterministic computer algorithm
  - Uniform distribution over a predefined range
  - Can repeat a calculation with the same sequence of numbers
    - Easy to debug the simulations
  - Should produce a large number of unique numbers before repeating the cycle.
  - For example, if we are studying the sensitivity of a calculation to variation in a selected parameter, we can reduce the variance of the difference between results calculated with two trail values of the parameter by using the same random number sequence.

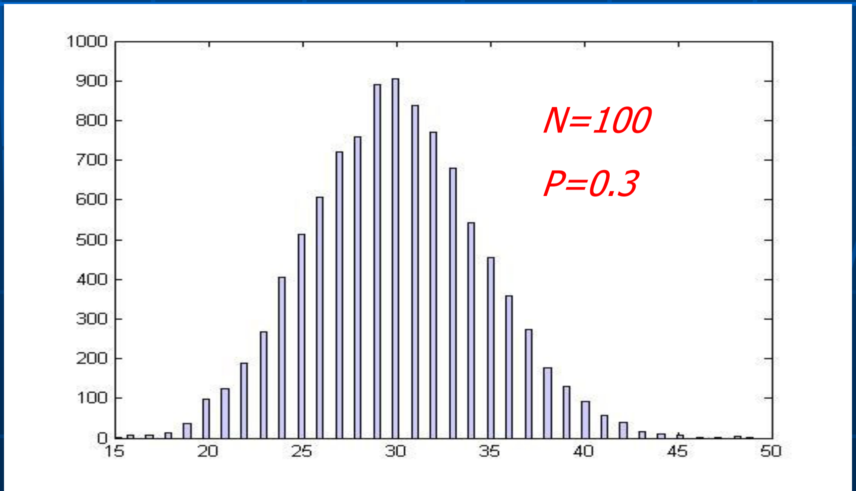
Random Numbers from Probability Distributions

- Direct Method
  - Use the probability distribution directly
  - For example:
    - Binominal random numbers

      Let assume a probability *p* of a head on any single toss of a coin

      If you generate N uniform random numbers on the interval (0,1), the counts the numbers less than p

      - The count is a binomial random numbers with parameters *N* and *p*



Random Numbers from Probability Distributions

- Transformation Method (Inverse transform method)

  Numbers drawn from a specific probability distribution *P(x)*

  - Let us define uniform deviates *p(r)* drawn from a standard probability density distribution that is uniform between *r=0* and *r=1*

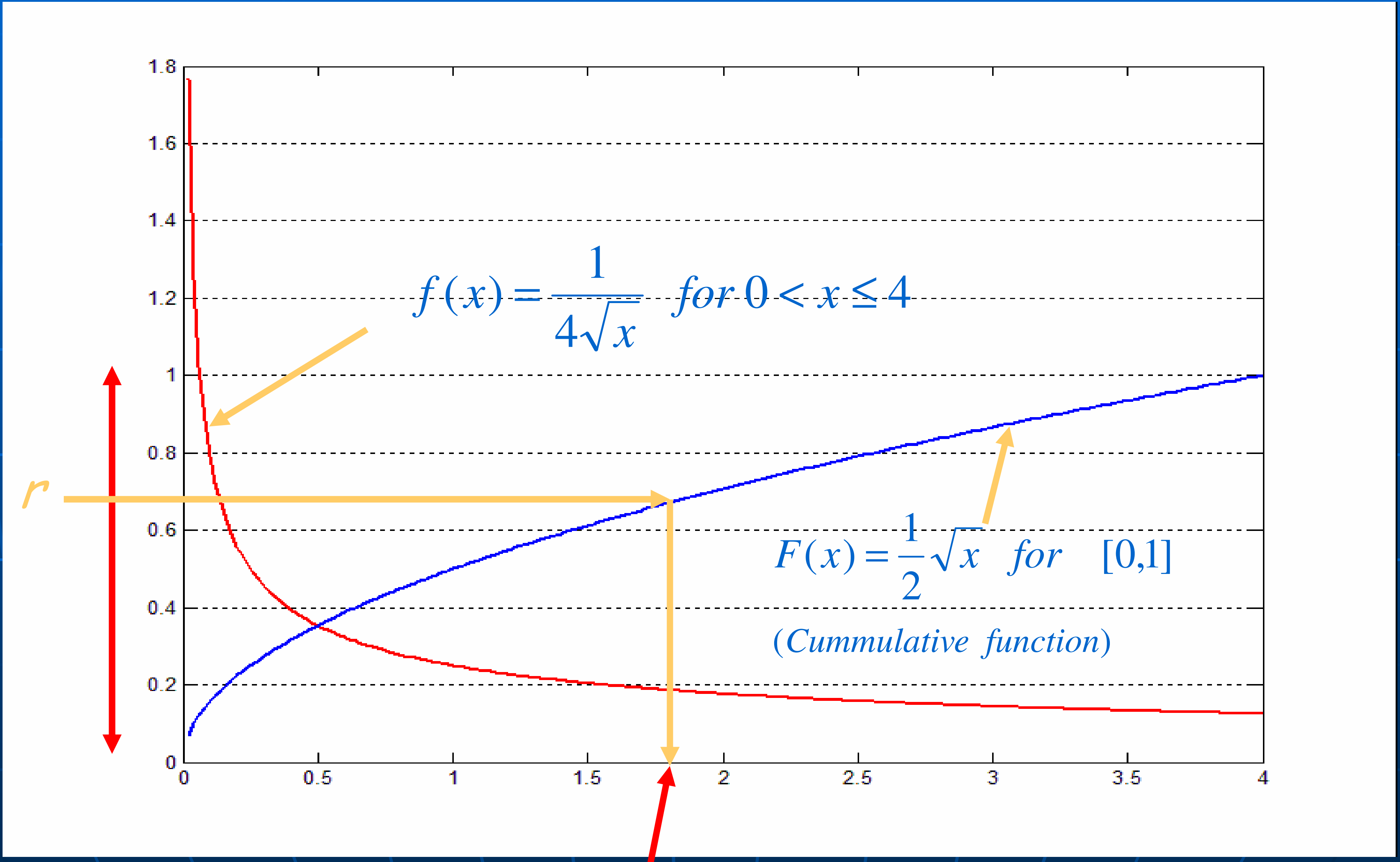
$$p(r) = 1 \qquad for \;\; 0 \le r < 1$$
$$= 0 \qquad otherwise$$

Comulative distribution fuction

$$F_x(x) = \int_{x=-\infty}^{x} P(x)dx$$

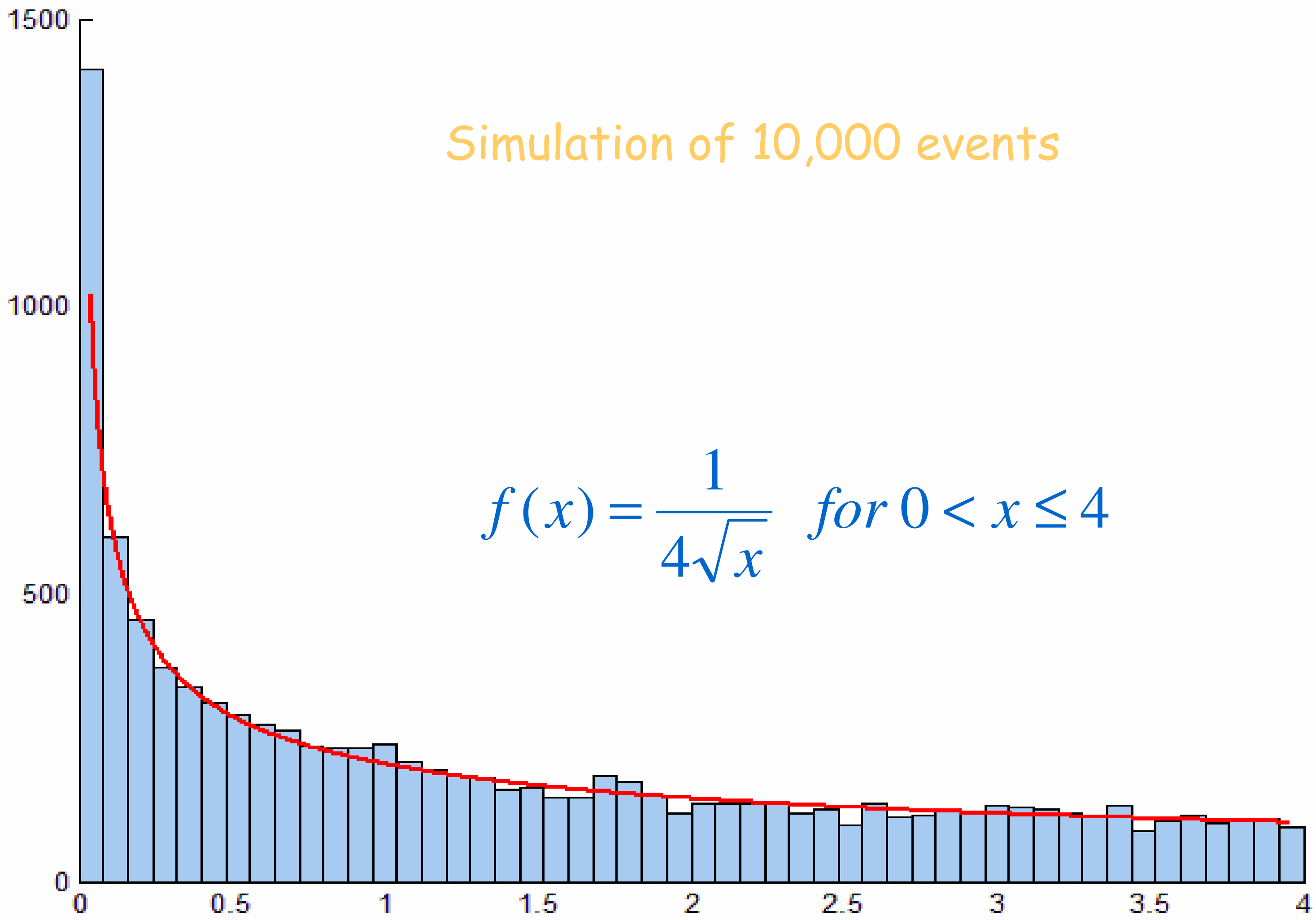
This method is used to generate normal, exponential and other types of distributions where the inverse is well defined and easy to obtain

Transformation Method



$X = F^{-1}(r) = (2*r)^2$

Transformation Method



$$f(x) = \frac{1}{4\sqrt{x}} \quad for\ 0 < x \le 4$$

Random Numbers from Probability Distribution

- Acceptance-Rejection Method
  - Very easy to use
  - Successively generates random values from the uniform distribution
  - Calculate the probability function to estimates the random event
  - If the event falls within the probability distribution function – "Hit", otherwise "reject"
  - In a complex Monte Carlo simulation only a small fraction of the events may survive.

Acceptance-Rejection Method

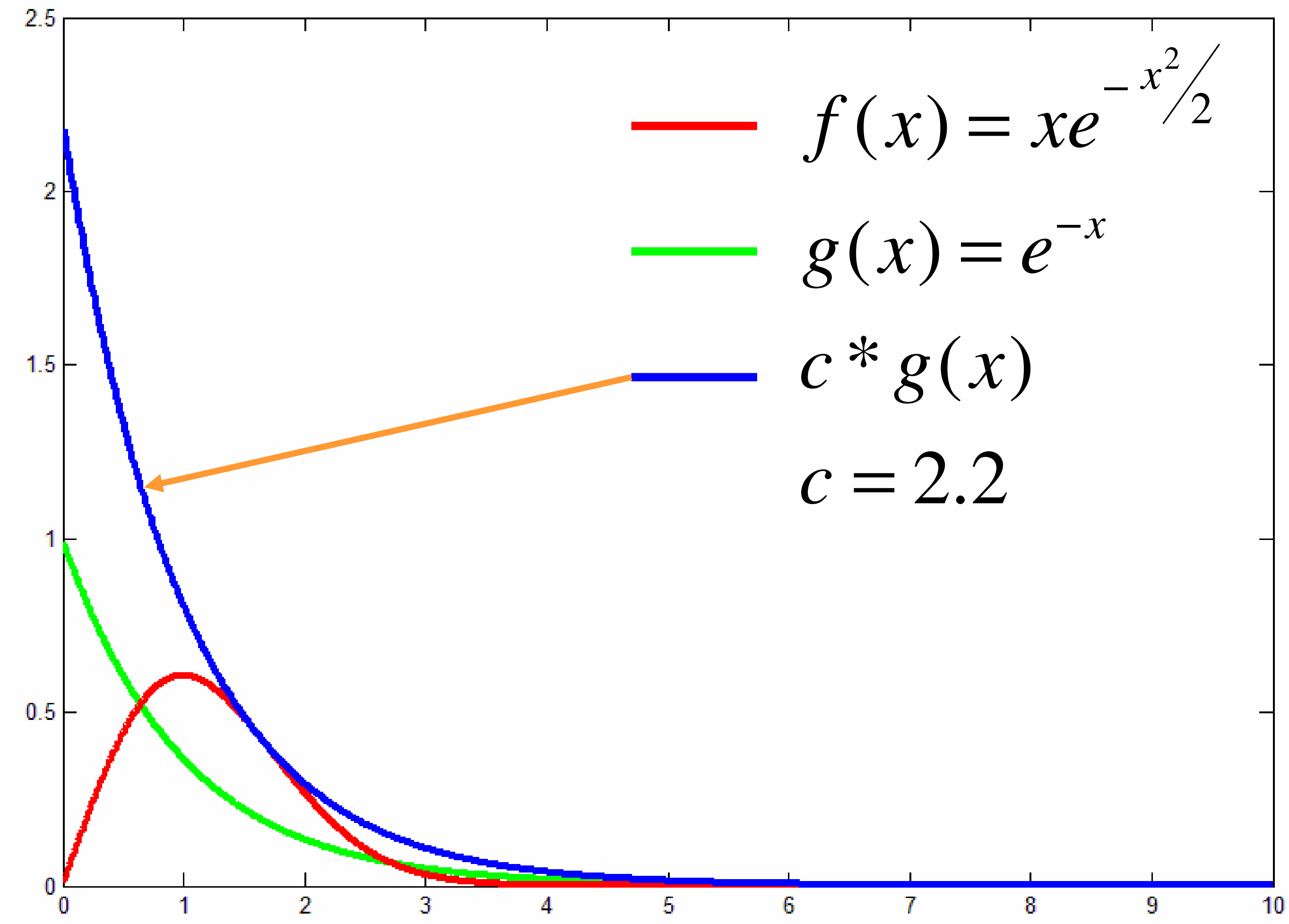
- Aim is to generate a random number from a continuous distribution with PDF *f(x),*

1. Choose a density function *g(x)*
2. Find a constant *c* such that

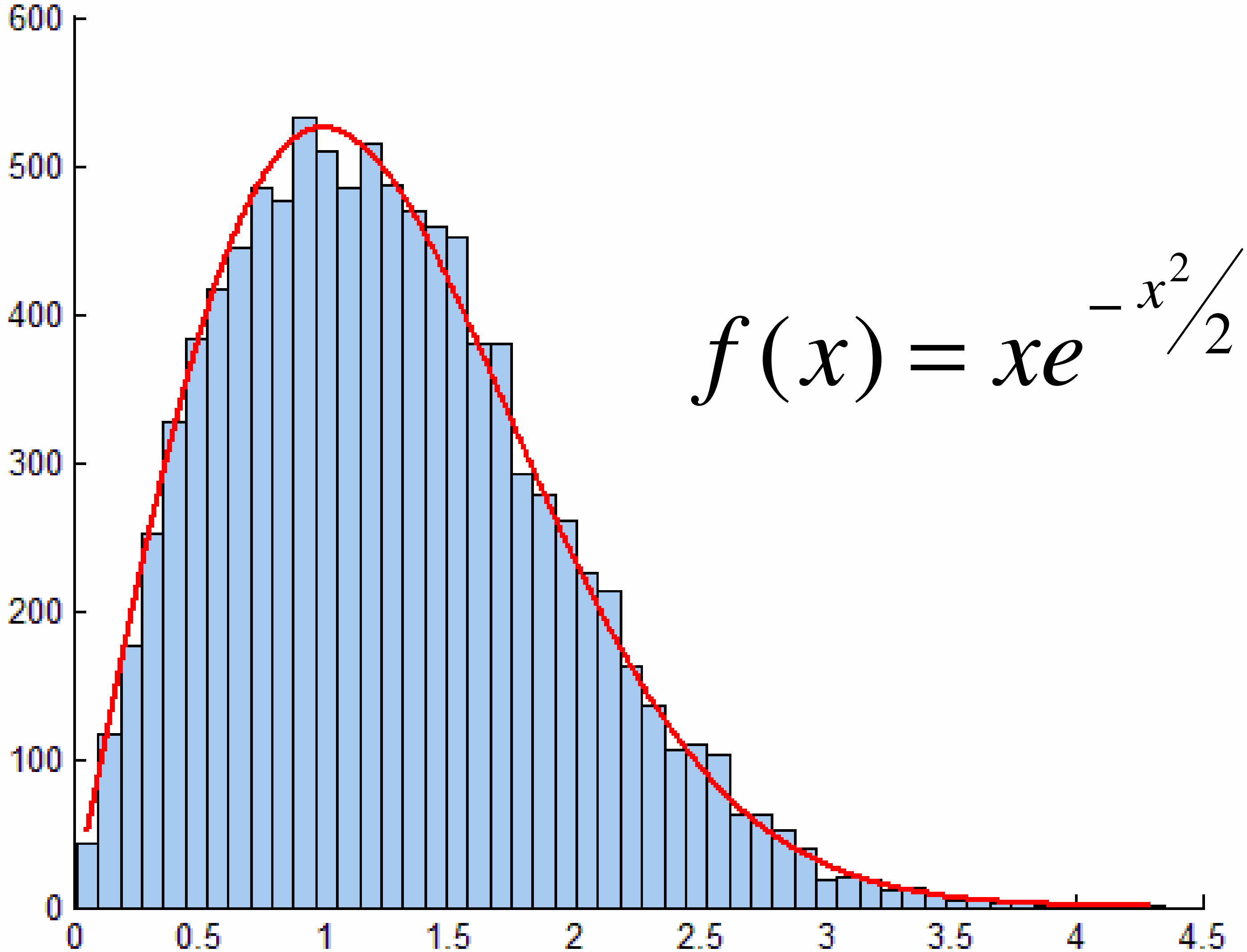
$$f(x) \le cg(x) \quad for\ some\ c\ and\ all\ x$$

3. Generate uniform random number *u* between 0 & 1
4. Generate a random number v from g(x)
5. If c*u ≤ f(v)/g(v), accept and return *v*
6. Otherwise reject and go to step 3

Acceptance-Rejection Method



Acceptance-Rejection Method



$$f(x) = xe^{-x^2/2}$$

Error in Monte Carlo Method

- Variance in Monte Carlo Simulation can be calculated for N number of measurements as

$$s^2 = \left[\frac{1}{N} \sum_{i=1}^{N} x_i^2 - \left(\frac{1}{N} \sum_{i=1}^{N} x_i \right)^2 \right]$$

Standard deviation $\sigma = \sqrt{s^2}$

This is the standard deviation in a single measurement. It will be several orders higher than the actual error

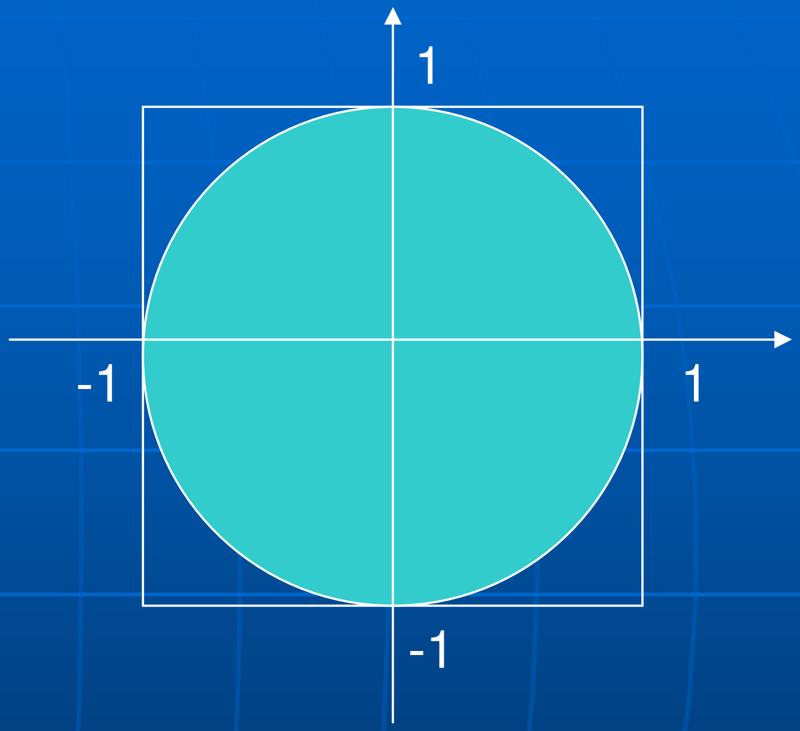
If the simulation divided into *m* sets of measurements each with *n* trials (measurements), then the error will be

$$err = \frac{\sigma}{\sqrt{n - 1}} \approx \frac{\sigma}{\sqrt{n}}$$

Calculate value of π

$$\pi = \frac{4 \times \pi R^2}{(2R)^2} = \frac{4 \times Area\ of\ a\ circle}{Area\ of\ enclosing\ square}$$

$$\pi = 4 * \frac{score\ within\ circle}{Total\ events\ generated}$$



$$\text{Result} = 3.1417 \pm 0.001\ for\ N = 10^7\ events$$

On Gaussian distribution $\pm\sigma$ *and* $\pm 2\sigma$ *fall within* 68.3%, 95% *resply*

Gaussian Distribution

- Almost any MC that simulates experimental measurements will require the generation of deviates drawn from Gaussian distribution
- The normal distribution with mean 0 and the standard deviation 1

$$P(z)dz = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz$$

To scale to different means μ and standard deviation δ by

$$x = \sigma z + \mu$$

Gaussian Distribution

- An interesting method for generating Gaussian deviates is based on the fact that if we reputedly calculate the men of group of numbers drawn randomly from any distribution, the distribution of those means tends to a Gaussian as the number of means increases
    - If we calculate many times the sums of N uniform deviates, drawn from the uniform distribution, we should expect the sums to fall into a truncated Gaussian distribution (bounded by 0 & N) with mean value N/2
    - If we generate N values of r from a uniform distribution

$$p(r) = 1 \quad for\, 0 \le r < 1$$
$$= 0 \quad otherwise$$

$$r_G = \sum_{i=0}^{N} r_i - N/2 \quad where\ \mu = 0, \sigma = \sqrt{N/12}$$

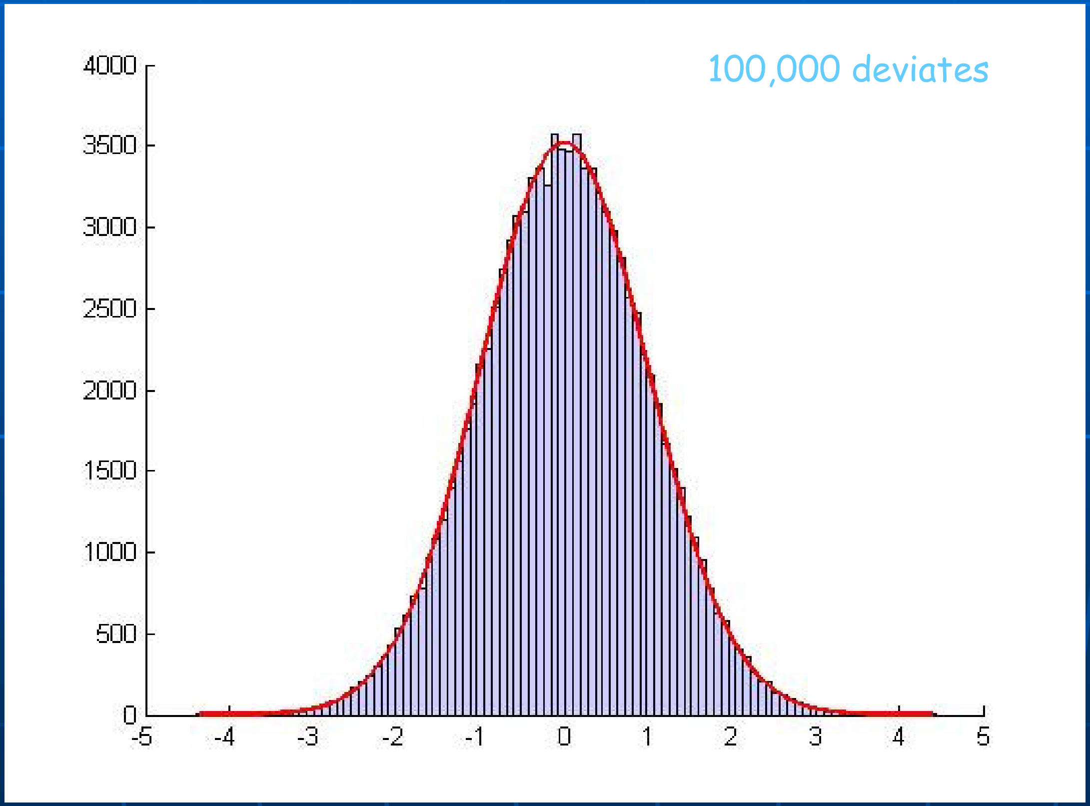
When N=12, the variance and standard deviation becomes 1

Gaussian Distribution

- Box and Muller (1958) suggested to generate two Gaussian deviates z_1 & z_2 from two uniform deviates r_1 & r_2

$$z_1 = \sqrt{-2\ln r_1}\cos 2\pi r_2$$

$$z_2 = \sqrt{-2\ln r_1}\sin 2\pi r_2$$



Poisson Distribution

- Does not have the convenient scaling properties of the Gaussian function
- To find an integer *x* drawn from the Poisson distribution with mean μ,
  - Generate a random variable r from the uniform distribution as

$$r = \sum_{x=0}^{x} \frac{\mu^x}{x!} e^{-\mu}$$

Build the Poisson distribution table and look up from the table for the generated random number.

For example: Geiger counter experiment with an assumed mean counting rate of 8.4 counts per 20s interval. Build a table for Poisson distribution with μ=8.4

At larger values of μ, the Poisson distribution becomes indistinguishable from the Gaussian function

Exponential Distribution

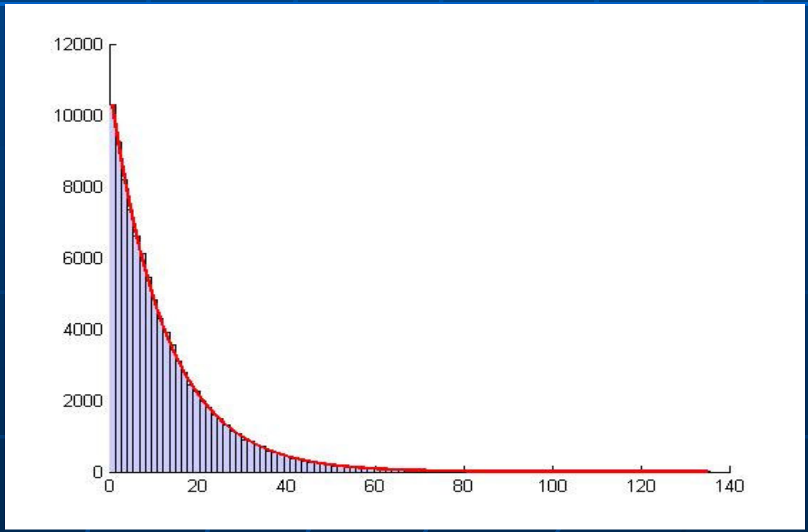
- To simulate unstable state such as particle decay
- The probability density function for obtaining a count at time t from an exponential distribution with mean life τ

$$P(t:\tau) = 0 \quad for\ t < 0$$

$$= \frac{e^{-t/\tau}}{\tau}$$

Random sample t_i from the distribution can be obtained from

$$t_i = -\tau \ln r_i$$



Composition + Rejection

- If the probability density *f(x)* can be written as

$$f(x)=\sum_{i=1}^{n}\alpha_i f_i(x)g_i(x)$$

$f_i(x)$ must be easily sampled & $g_i(x)$ should be not much smaller than unit to keep rejections low

$$where\ 0\le g_i(x)\le 1,\quad \alpha_i>0,\quad f_i(x)\ge 0,\quad \int f_i(x)dx=1$$

First choose an integer from *1* to *n* with probability proportional to α_i of being *i*

A sample x is drawn from the distribution function with frequency function $f_i(x)$ and the value is accepted or rejected by computing the value of $g_i(x)$ and comparing with a random number *r*

if $r > g_i(x)$, reject *x*

$r \le g_i(x)$, accept *x*

So that the probability of accepting or rejection x is $g_i(x)$

For example, Compton scattering

Compton Scattering

- Klein-Nishina DCS

$$\frac{d\sigma}{d\varepsilon} = \pi r_e^2 \frac{m_e c^2}{E_0} Z \left[\frac{1}{\varepsilon} + \varepsilon \right] \left[1 - \frac{\varepsilon \sin^2 \theta}{1 - \varepsilon^2} \right]$$

$r_e - classicl\ electron\ radius$

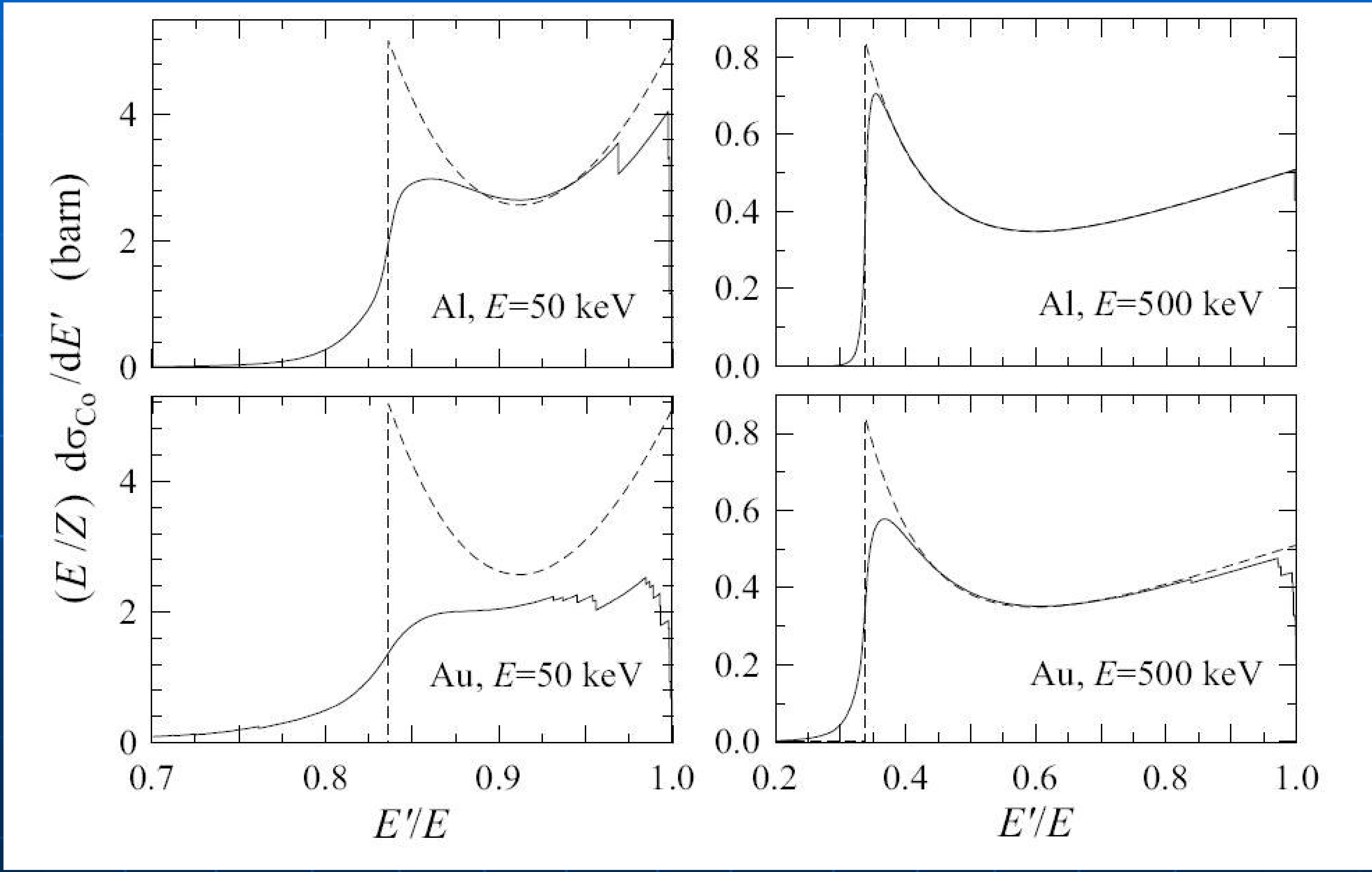
$m_e c^2 - electron\ mass$

$E_0, E_1 - Energy\ of\ incident\ \&\ scattered\ photons$

$$\varepsilon = E_1 / E_0$$

$$E_1 = E_0 \frac{m_e c^2}{m_e c^2 + E_0 (1 - \cos\theta)}$$

Compton Scattering



Compton Scattering

Minimum photon energy - Backward scattering ε_0

$$\varepsilon_0 = \frac{m_e c^2}{m_e c^2 + 2E_0} \qquad \varepsilon \in [\varepsilon_0, 1] \qquad \left[\varepsilon + \frac{1}{\varepsilon} - \sin^2\theta\right]$$

$$\phi(\varepsilon) \cong \left[\frac{1}{\varepsilon} + \varepsilon\right]\left[1 - \frac{\varepsilon \sin^2\theta}{1+\varepsilon^2}\right] = f(\varepsilon) \bullet g(\varepsilon) = [\alpha_1 f_1(\varepsilon) + \alpha_2 f_2(\varepsilon)] \bullet g(\varepsilon)$$

$$where\ \alpha_1 = \ln(1/\varepsilon_0) \qquad f_1(\varepsilon) = 1/(\alpha_1 \varepsilon)$$

$$\alpha_2 = (1-\varepsilon_0^2)/2 \qquad f_2(\varepsilon) = \varepsilon/\alpha_2$$

Rejection function

f_1 and *f_2 are probability density functions $[\varepsilon_0, 1]$*

$$g(\varepsilon) = \left[1 - \frac{\varepsilon}{1+\varepsilon^2}\sin^2\theta\right]$$

$$f(\varepsilon) = \sum_{i=1}^{2} \alpha_i f(\varepsilon) \qquad for\ \varepsilon_0 \geq \varepsilon > 1$$

$$for\ all\ [\varepsilon_0, 1] \Rightarrow 0 < g(\varepsilon) \leq 1$$

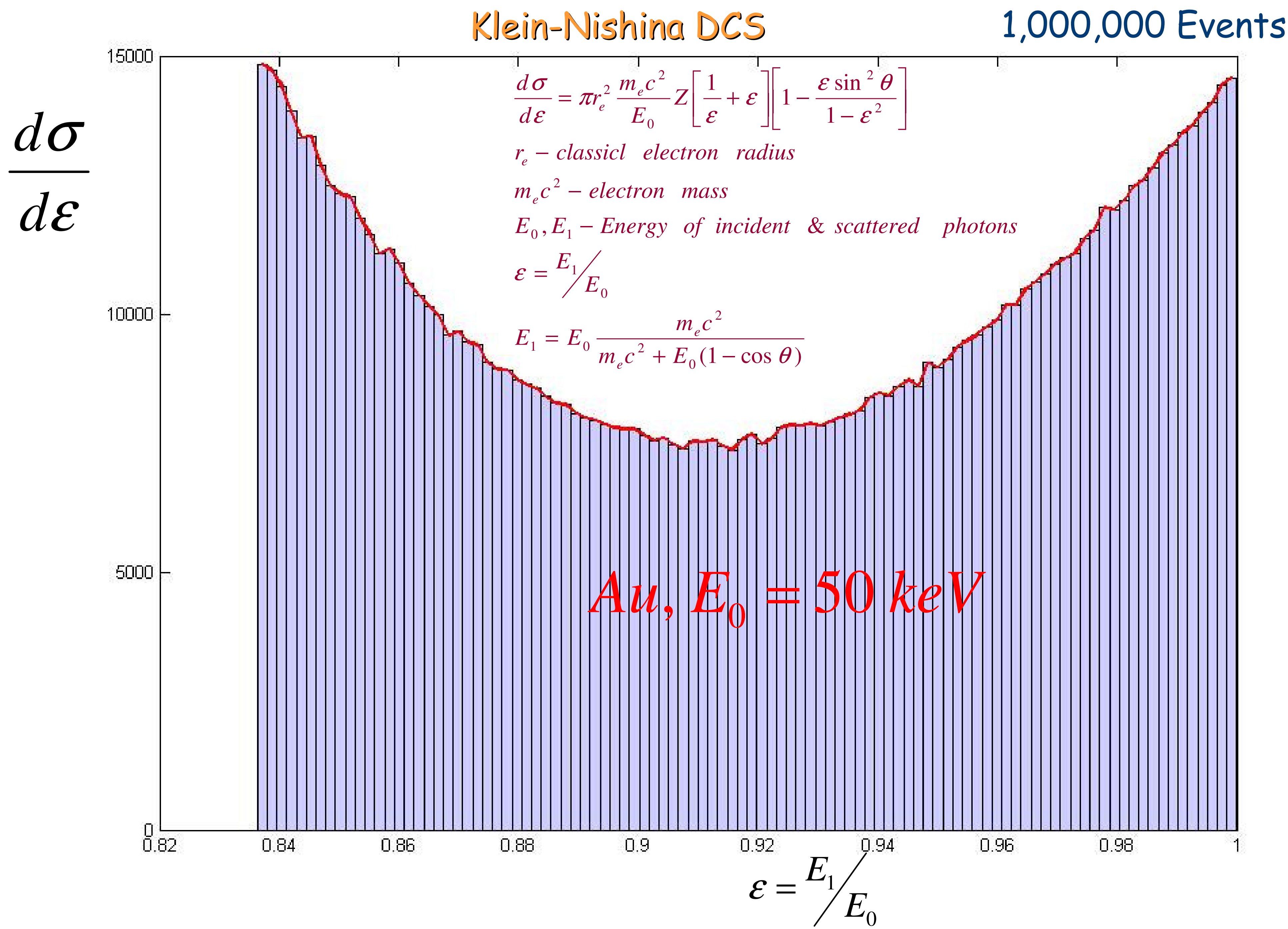
Sampling Photon Energy

f_1 and f_2 - normalised probability density functions [ε_0,1]

Random numbers *(r) are drawn from uniform distribution [0 1]*

- Step 1
  - If $r < \alpha_1/(\alpha_1+\alpha_2)$ select f_1, otherwise f_2
- Step 2
  - Sample ε from the distribution corresponding to f_1 or f_2
    - *For f_1 : $\varepsilon = \exp(-r\alpha_1)$*
    - *For f_2 : $\varepsilon^2 = \varepsilon_0^2 + (1 - \varepsilon_0^2)r$*
- *Step 3*
  - Use the value of ε in stp 2 to calculate $\sin^2\theta = (2-t)$ *and then* $g(\varepsilon)$ Where $t = (1-\cos\theta) = m_e c^2(1-\varepsilon)/(E_0\varepsilon)$
- *Step 4*
  - *If* $r \leq g(\varepsilon)$ accept ε, otherwise *goto* step *1*

Compton Scattering without atomic form factor



Low Energy Compton Scattering in GEANT4

- Hubbel's atomic form factor, SF

  $P(\varepsilon,q)=\Phi(\varepsilon) \times SF(q)$

  Momentum transfer $q = E \sin^2(\theta/2)$

  - Energy distribution of the scattered photon according to the Klein-Nishina formula, multiplied by scattering function F(q) (Hubbel's atomic factor) from EPDL97 data library
  - The effect of scattering function becomes significant at low energies in suppressing forward scattering
  - Angular distribution of the scattered photon and the recoil electron also based on EPDL97