NE 155
Introduction to Numerical Simulations in Radiation Transport
Random Sampling

R. N. Slaybaugh

April 20, 2021

MAJOR COMPONENTS OF MC ALGORITHM

- *PDFs: the physical/mathematical system must be described by a set of pdfs.*
- **Random number generator**: a source of random #s uniformly distributed on the unit interval.
- *Sampling rule: prescription for sampling the pdf (given having random #s)*
- **Scoring**: the outcomes must be accumulated/<u>tallied</u> for quantities of interest
- **Error estimation**: an estimate of the statistical error (<u>variance</u>) of the solution
- **Variance Reduction**: methods for reducing the variance and computation time simultaneously
- **Parallelization**: efficient use of computers

OUTLINE

1. Physics as Probability
2. Definitions: PDF & CDF
3. Motivation & Goal of Random Sampling
4. Basic Random Sampling Techniques
   - Direct Discrete Sampling
   - Direct Continuous Sampling
   - Rejection Sampling

Notes derived from Jasmina Vujic and Paul Wilson

LEARNING OBJECTIVES

1. Provide examples of probabilistic representations of physics
2. Distinguish between a PDF and CDF
3. Distinguish between a *discrete* PDF/CDF and a *continuous* PDF/CDF
4. Describe the goal of random sampling
5. Identify and implement the best random sampling technique for a given distribution

PHYSICS AS PROBABILITY

Various physical phenomena can be represented by probability distributions

- Photon emission energy
  - Each possible energy has a different probability (intensity)

PHYSICS AS PROBABILITY

Various physical phenomena can be represented by probability distributions

- Photon emission energy
  - Each possible energy has a different probability (intensity)

- Scattering cross-sections
  - Each possible scattering angle has a different probability as a function of the energy

PHYSICS AS PROBABILITY

Various physical phenomena can be represented by probability distributions

- Photon emission energy
  - Each possible energy has a different probability (intensity)

- Scattering cross-sections
  - Each possible scattering angle has a different probability as a function of the energy

- Transmission through a medium
  - Probability of reaching a particular position depends on the cross-section

PROBABILITY DENSITY FUNCTIONS

All variables, x, have a <u>Probability Density Function (PDF)</u>, $p(x)$, with the following characteristics:

<div align="center">

Continuous   <u>Discrete</u>

</div>

$$p\{a \leq x \leq b\} = \int_a^b p(x)dx$$

$$p(x = x_k) = p_k \equiv p(x_k)$$
$$k = 1, \ldots, N$$

$$p(x) \geq 0$$

$$p_k \geq 0$$

$$\int_{-\infty}^{\infty} p(x)dx = 1$$

$$\sum_{k=1}^{N} p_k = 1$$

CUMULATIVE DISTRIBUTION FUNCTIONS

All PDFs, $p(x)$, have an associated
Cumulative Distribution Function (CDF), $P(x)$, with the following
properties:

Continuous

Discrete

$$P\{x' \le x\} = P(x) = \int_{-\infty}^{x} p(x')dx' \qquad P\{x' \le x\} = P_k \equiv P(x_k) = \sum_{j=1}^{k} p_j$$

$$P(-\infty) = 0 , \quad P(\infty) = 1$$

$$0 \le P(x) \le 1$$

$$\frac{dP(x)}{dx} \ge 0$$

$$k = 1, \ldots, N$$
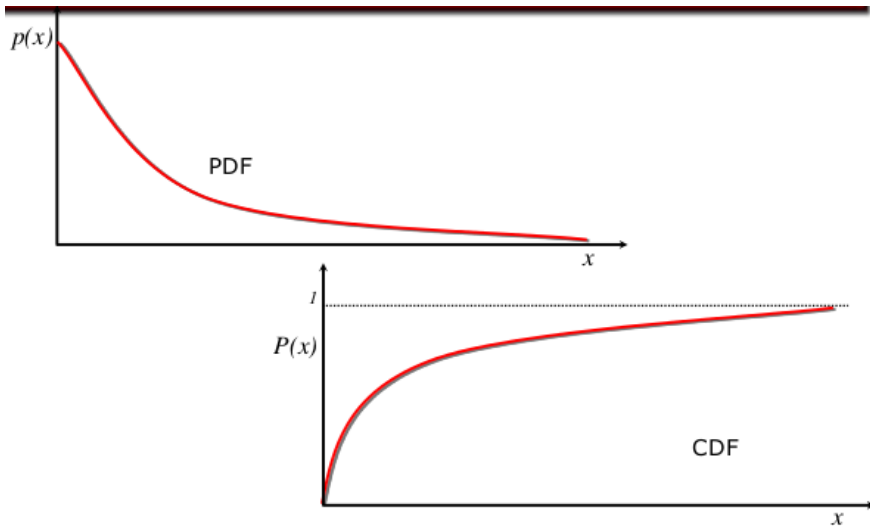
$$P_0 = 0 , \quad P_N = 1$$
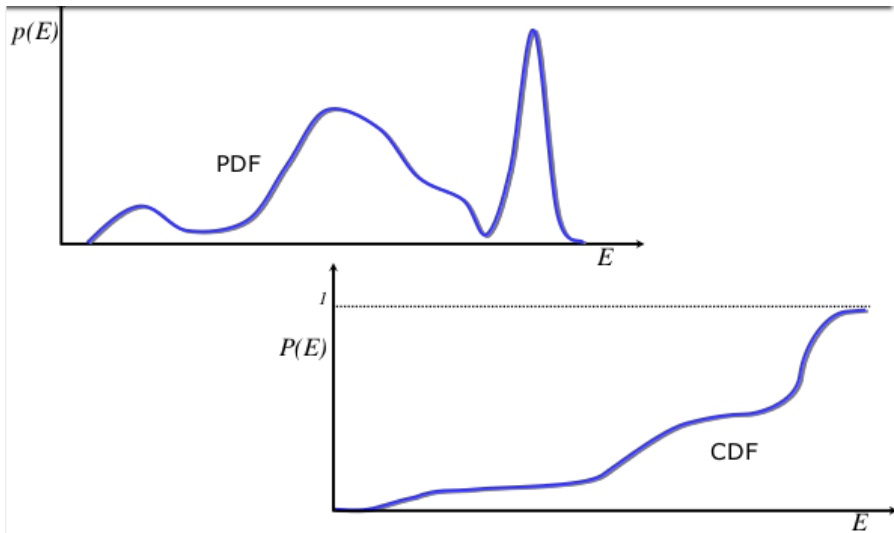
$$0 \le P_k \le 1$$

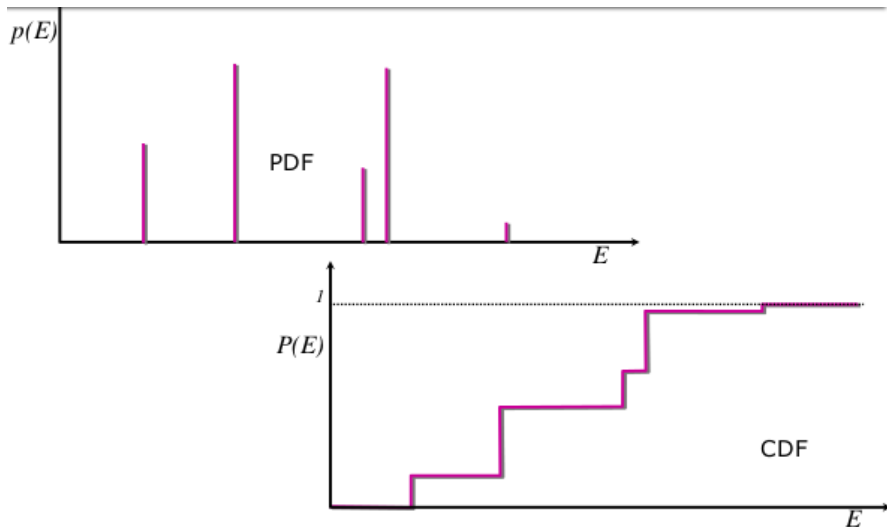$$P_k \ge P_{k-1}$$

RANDOM SAMPLING BASICS
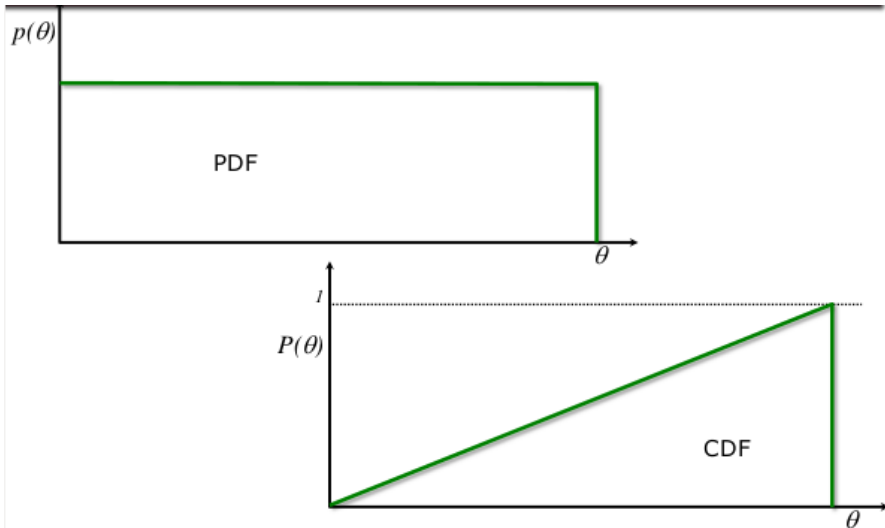
RANDOM SAMPLING BASICS

RANDOM SAMPLING BASICS

RANDOM SAMPLING BASICS

WHY RANDOM SAMPLING

Various physical phenomena can be represented by probabilistic distributions

- The known probability distribution represents the *collective* behavior

- We need to know the behavior at *each* single event

- We need to <u>recreate</u> the collective behavior after <u>many</u> events

RANDOM SAMPLING PURPOSE

Use a random process to select a single value with the following
requirements

- Each sample should be independent from other samples

- The PDF formed from a large number of samples should converge
  to the initial PDF

- Recover the full resolution of the initial PDF

SAMPLING TECHNIQUES

Random sampling uses uniformly distributed random variables to choose a value for a variable according to its probability density function

- *Basic* sampling techniques
  - Direct discrete sampling
  - Continuous discrete sampling
  - Rejection sampling

SAMPLING TECHNIQUES

Random sampling uses uniformly distributed random variables to choose a value for a variable according to its probability density function

- *Basic* sampling techniques
  - Direct discrete sampling
  - Continuous discrete sampling
  - Rejection sampling

- *Advanced* sampling techniques
  - Histogram
  - Piecewise linear
  - Alias sampling
  - Advanced continuous PDFs

UNIFORMLY-DISTRIBUTED RANDOM VARIABLE

- Standard notation
  - Single random variable: ξ
  - Pair of random variables: (ξ, η)
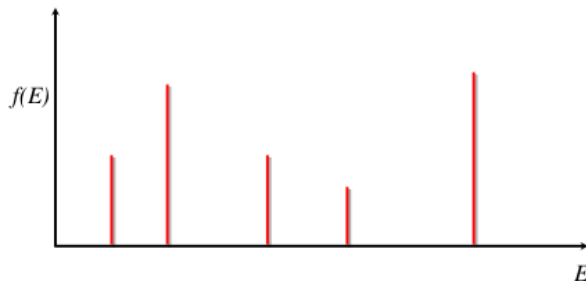
- PDF for random variables:

$$p(\xi) = \begin{cases} 1 & 0 \leq \xi < 1 \\ 0 & \text{otherwise} \end{cases}$$

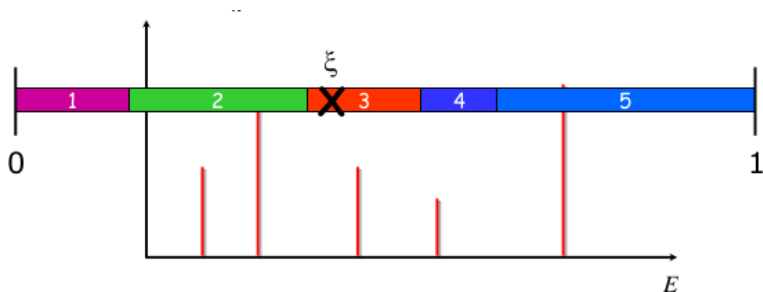# DIRECT DISCRETE SAMPLING

Sampling Procedure

- Generate ξ
- Determine k such that $P_{k-1} \leq \xi \leq P_k$
- Return $x = x_k$
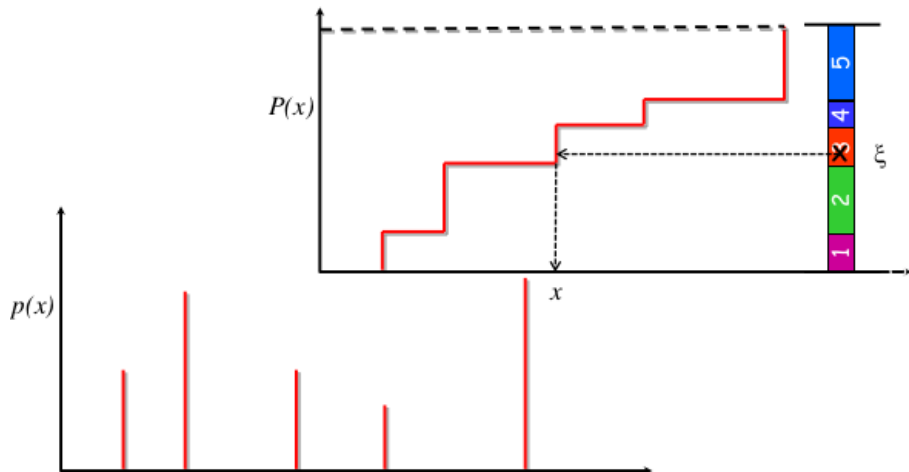
DIRECT DISCRETE SAMPLING

Sampling Procedure

- Generate ξ
- Determine k such that $P_{k-1} \leq \xi \leq P_k$
- Return $x = x_k$

DIRECT DISCRETE SAMPLING

Consider the CDF

DIRECT DISCRETE SAMPLING

- Requires a table search on P_k
  - Linear search requires $O(N)$ time
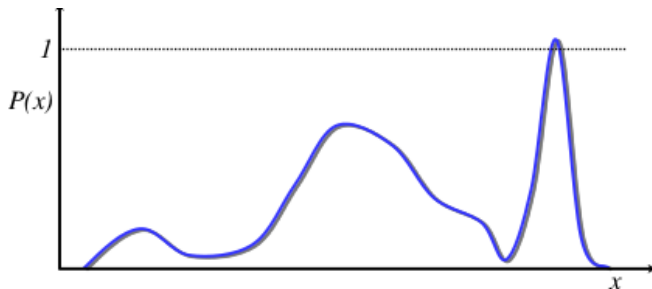 - Binary search requires $O(\log_2 N)$ time

DIRECT DISCRETE SAMPLING

- Requires a table search on P_k
  - Linear search requires $O(N)$ time
  - Binary search requires $O(\log_2 N)$ time

- Special case: Uniform discrete PDF
  - $p_k = 1/N$
  - $P_k = k/N$
  - $k = \lfloor 1 + N\xi \rfloor$ (floor function)

DIRECT CONTINUOUS SAMPLING

- Can only be used if CDF can be inverted
- Direct solution of $P(x) = \xi$
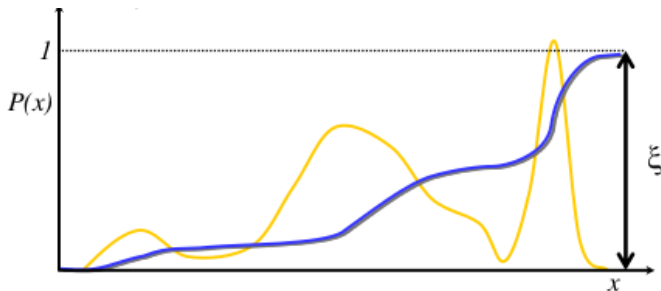- Sampling Procedure:

  Generate ξ,   Determine $x = P^{-1}(\xi)$

DIRECT CONTINUOUS SAMPLING

- Can only be used if CDF can be inverted
- Direct solution of $P(x) = \xi$
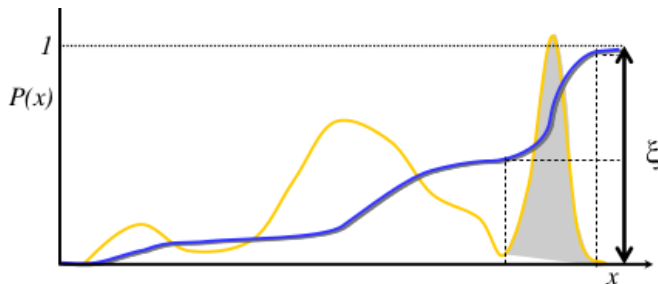- Sampling Procedure:

    Generate ξ,   Determine $x = P^{-1}(\xi)$

DIRECT CONTINUOUS SAMPLING

- Can only be used if CDF can be inverted
- Direct solution of $P(x) = \xi$
- Sampling Procedure:

  Generate ξ,   Determine $x = P^{-1}(\xi)$

DIRECT CONTINUOUS SAMPLING

- Advantages:
  - Straightforward math & coding

- Disadvantages:
  - Can involve computationally slow functions
  - Not always possible to invert $P(x)$

NORMALIZATION

- Random sampling depends on **shape** and not on ~~magnitude~~
- Normalization for formal definition of PDF/CDF required

$$g(t)dt = e^{-\lambda t}dt \,, \quad t > 0$$

NORMALIZATION

- Random sampling depends on **shape** and not on ~~magnitude~~
- Normalization for formal definition of PDF/CDF required

$$g(t)dt = e^{-\lambda t}dt, \quad t > 0$$

$$G(t) = \int_{-\infty}^{t} g(t')dt' = \int_{0}^{t} g(t')dt' = \left[-\frac{e^{-\lambda t'}}{\lambda}\right]_{0}^{t} = \frac{1}{\lambda}\left(1 - e^{-\lambda t}\right)$$

$$G(\infty) = \frac{1}{\lambda}$$

NORMALIZATION

- Random sampling depends on **shape** and not on ~~magnitude~~
- Normalization for formal definition of PDF/CDF required

$$g(t)dt = e^{-\lambda t}dt, \quad t > 0$$

$$G(t) = \int_{-\infty}^{t} g(t')dt' = \int_{0}^{t} g(t')dt' = \left[-\frac{e^{-\lambda t'}}{\lambda}\right]_{0}^{t} = \frac{1}{\lambda}\left(1 - e^{-\lambda t}\right)$$

$$G(\infty) = \frac{1}{\lambda}$$

$$p(t) = \lambda g(t) = \lambda e^{-\lambda t}dt, \quad t > 0$$

$$P(t) = \int_{-\infty}^{t} p(t')dt' = \int_{0}^{t} \lambda f(t')dt' = \left[e^{-\lambda t'}\right]_{0}^{t} = 1 - e^{-\lambda t}$$

$$P(\infty) = 1$$

SHIFTED UNIFORM

$$g(x)dx = C \quad a \leq x < b$$

SHIFTED UNIFORM

$$g(x)dx = C \quad a \le x < b$$

$$G(x) = \int_{-\infty}^{x} g(x')dx' = C \int_{a}^{x} dx' = C\big[x'\big]_{a}^{x} = C(x - a)$$

$$G(\infty) = G(b) = C(b - a)$$

SHIFTED UNIFORM

$$g(x)dx = C \quad a \leq x < b$$

$$G(x) = \int_{-\infty}^{x} g(x')dx' = C \int_{a}^{x} dx' = C\big[x'\big]_{a}^{x} = C(x - a)$$

$$G(\infty) = G(b) = C(b - a)$$

$$p(x) = \frac{g(x)}{G(\infty)} = \frac{C}{C(b-a)} = \frac{1}{b-a} \quad a \leq x < b$$

SHIFTED UNIFORM

$$g(x)dx = C \quad a \le x < b$$

$$G(x) = \int_{-\infty}^{x} g(x')dx' = C \int_{a}^{x} dx' = C[x']_{a}^{x} = C(x - a)$$

$$G(\infty) = G(b) = C(b - a)$$

$$p(x) = \frac{g(x)}{G(\infty)} = \frac{C}{C(b - a)} = \frac{1}{b - a} \quad a \le x < b$$

$$P(x) = \int_{-\infty}^{x} p(x')dx' = \frac{1}{b - a} \int_{a}^{x} dx' = \frac{x - a}{b - a}$$

$$x = P^{-1}(\xi) = \xi(b - a) + a$$

SIMPLE LINE, SLOPE = m

$$g(x)dx = mx \qquad 0 \leq x < 1$$

SIMPLE LINE, SLOPE = m

$$g(x)dx = mx \qquad 0 \leq x < 1$$

$$G(x) = \int_{-\infty}^{x} g(x')dx' = \int_{0}^{x} mx'dx' = \frac{m}{2}\left[x'^2\right]_{0}^{x} = \frac{m}{2}x^2$$

$$G(\infty) = G(1) = \frac{m}{2}$$

SIMPLE LINE, SLOPE = m

$$g(x)dx = mx \qquad 0 \leq x < 1$$

$$G(x) = \int_{-\infty}^{x} g(x')dx' = \int_{0}^{x} mx'dx' = \frac{m}{2}\left[x'^2\right]_0^x = \frac{m}{2}x^2$$

$$G(\infty) = G(1) = \frac{m}{2}$$

$$p(x) = \frac{mx}{\frac{m}{2}} = 2x \qquad 0 \leq x < 1$$

SIMPLE LINE, SLOPE = m

$$g(x)dx = mx \qquad 0 \le x < 1$$

$$G(x) = \int_{-\infty}^{x} g(x')dx' = \int_{0}^{x} mx'dx' = \frac{m}{2}\left[x'^2\right]_0^x = \frac{m}{2}x^2$$

$$G(\infty) = G(1) = \frac{m}{2}$$

$$p(x) = \frac{mx}{\frac{m}{2}} = 2x \qquad 0 \le x < 1$$

$$P(x) = \int_{-\infty}^{x} p(x')dx' = \int_{0}^{x} 2x'dx' = \left[x'^2\right]_0^x = x^2$$

$$x = P^{-1}(\xi) = \sqrt{\xi} \qquad \text{Independent of } m$$

SHIFTED LINE

$$g(x)dx = m(x - a) \qquad a \leq x < b$$

$$G(x) = \int_{-\infty}^{x} g(x')dx' = \int_{a}^{x} m(x' - a)dx' = \frac{m}{2}\left[(x' - a)^2\right]_0^x = \frac{m}{2}(x - a)^2$$

$$G(\infty) = G(1) = \frac{m}{2}(b - a)^2$$

$$p(x) = \frac{m(x - a)}{\frac{m}{2}(b - a)^2} = 2\frac{x - a}{(b - a)^2} \qquad a \leq x < b$$

$$P(x) = \int_{-\infty}^{x} p(x')dx' = \frac{1}{(b - a)^2}\int_{a}^{x} 2(x' - a)dx' = \frac{(x - a)^2}{(b - a)^2}$$

$$x = P^{-1}(\xi) = \sqrt{\xi}(b - a) + a \qquad \text{Independent of } m$$

REJECTION SAMPLING

- Many CDFs cannot be inverted
  - e.g. Klien-Nishina cross-section

REJECTION SAMPLING

- Many CDFs cannot be inverted
  - e.g. Klien-Nishina cross-section

- Use an approach that is more graphical
  - Select a point in a 2-D domain
  - Determine whether that point is above or below the PDF
 - Keep those that are below
  - Start over if above

REJECTION SAMPLING

- Select a bounding function, g(x), such that
  - $g(x) \geq p(x)$ for all x
  - $g(x)$ is easy to sample
- Simplest choice is $g(x) = C$
- May not be best choice

REJECTION SAMPLING

- Select a bounding function, g(x), such that
  - $g(x) \geq p(x)$ for all x
  - $g(x)$ is easy to sample
- Simplest choice is $g(x) = C$
- May not be best choice

- Generate pair of random variables, (ξ, η)
  - $x' = G^{-1}(\xi)$
  - If $\eta < p(x')/g(x')$, accept x'
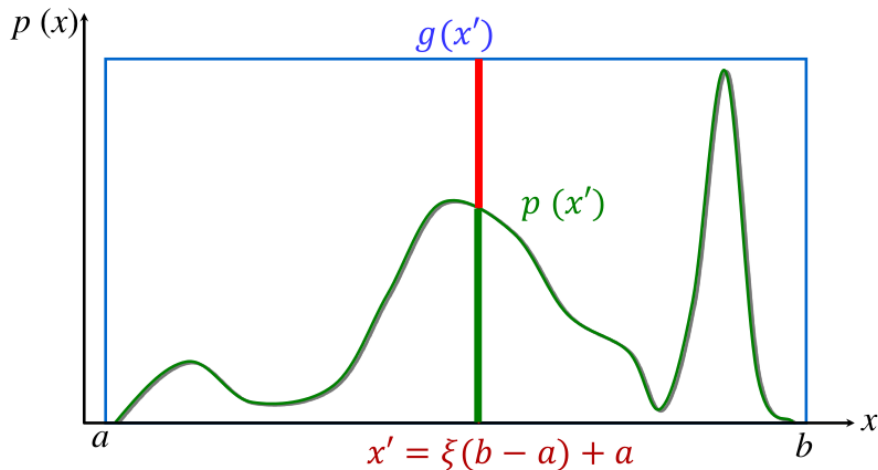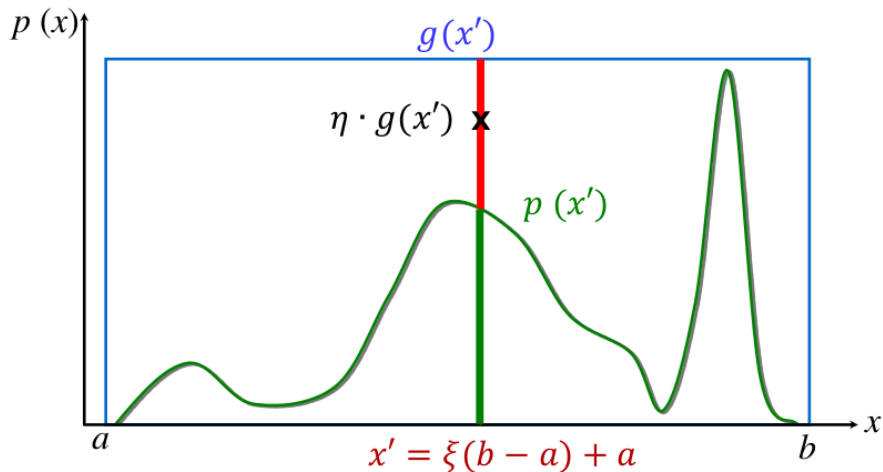 - Else, reject x'

REJECTION SAMPLING
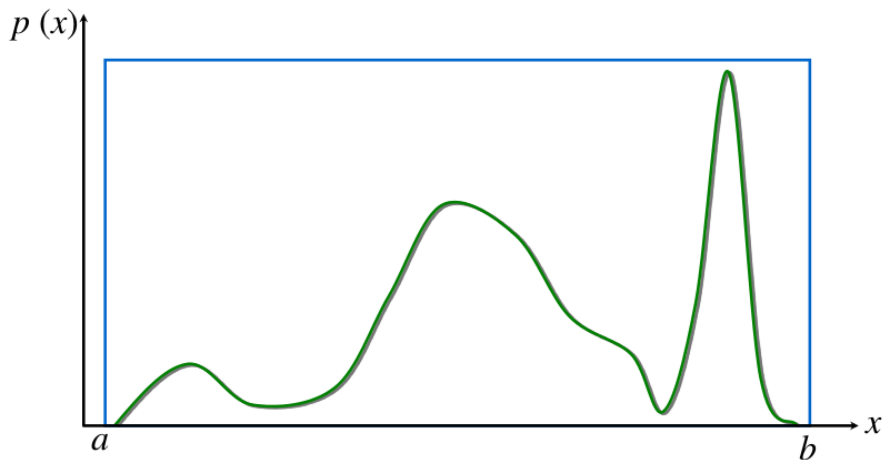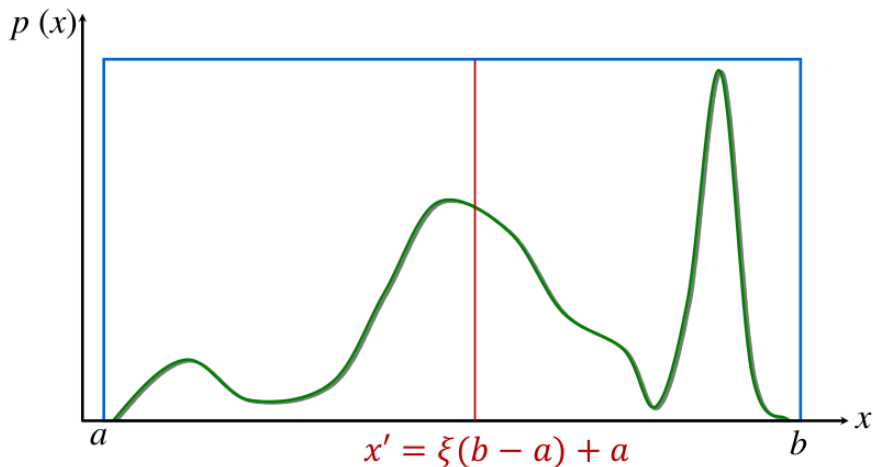
REJECTION SAMPLING

REJECTION SAMPLING

REJECTION SAMPLING

REJECTION SAMPLING

- Advantages
  - Computationally simple
  - Always works

REJECTION SAMPLING

- Advantages
  - Computationally simple
  - Always works

- Disadvantages
  - Will be inefficient if shapes of $g(x)$ and $p(x)$ are not similar

$$\text{Efficiency} = \frac{\int p(x)dx}{\int g(x)dx}$$

RANDOM SAMPLING SUMMARY

- Physics can be represented *probabilistically*

RANDOM SAMPLING SUMMARY

- Physics can be represented *probabilistically*

- We can create PDFs and from those generate CDFs

RANDOM SAMPLING SUMMARY

- Physics can be represented *probabilistically*

- We can create PDFs and from those generate CDFs

- These can be either continuous or discrete

RANDOM SAMPLING SUMMARY

- Physics can be represented *probabilistically*

- We can create PDFs and from those generate CDFs

- These can be either continuous or discrete

- We learned some basic ways to use random numbers to *sample* from these distributions to simulate physics