भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Programming for AI

Sampling Techniques, Inverse Transform Technique

Karthik P. N.
Assistant Professor, Department of AI
Email: pnkarthik@ai.iith.ac.in

07 April 2025

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Sampling from a Given Distribution

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Sampling from a Given Distribution

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

Remarks:

- The CDF F may or may not be continuous

The basic building block: Uniform $(0, 1)$ sample(s)

Sampling from a Given Distribution

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

Remarks:

- The CDF F may or may not be continuous
- If F is continuous and differentiable, it admits a probability density function f

The basic building block: Uniform $(0, 1)$ sample(s)

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Sampling from a Given Distribution

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

Remarks:

- The CDF F may or may not be continuous
- If F is continuous and differentiable, it admits a probability density function f
- Instead of the CDF F, we may be given a target PMF or PDF from which to sample

The basic building block: Uniform $(0, 1)$ sample(s)

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Inverse Transform Technique (ITT)

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Inverse Transform Technique

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

Inverse Transform Technique

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

- Sample $U \sim \mathrm{Unif}(0, 1)$

Inverse Transform Technique

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

- Sample $U \sim \mathrm{Unif}(0, 1)$
- Define the function F^{-1} as

$$F^{-1}(u) := \min\{x \in \mathbb{R} : F(x) \geq u\}$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Inverse Transform Technique

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

- Sample $U \sim \mathrm{Unif}(0, 1)$
- Define the function F^{-1} as

$$F^{-1}(u) := \min\{x \in \mathbb{R} : F(x) \geq u\}$$

- F^{-1} is called the quantile function

Inverse Transform Technique

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

- Sample $U \sim \mathrm{Unif}(0, 1)$
- Define the function F^{-1} as

$$F^{-1}(u) := \min\{x \in \mathbb{R} : F(x) \geq u\}$$

- F^{-1} is called the quantile function
- Set $X = F^{-1}(U)$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Inverse Transform Technique

Objective

Given a cumulative distribution function (CDF) $F : \mathbb{R} \to [0, 1]$, generate a sample $X \sim F$.

- Sample $U \sim \mathrm{Unif}(0, 1)$
- Define the function F^{-1} as

$$F^{-1}(u) := \min\{x \in \mathbb{R} : F(x) \geq u\}$$

- F^{-1} is called the quantile function
- Set $X = F^{-1}(U)$
- **Claim:** The CDF of X is exactly equal to F, i.e., $F_X = F$

Example

- Let X be a discrete random variable with the following PMF:

$$p_X(x) = \begin{cases} 0.1, & x = 10, \\ 0.2, & x = 20, \\ 0.3, & x = 30, \\ 0.4, & x = 40, \\ 0, & \text{otherwise.} \end{cases}$$

Use the inverse transform method to generate a sample from the above distribution.

Example

- **[Generating a Sample from Rayleigh Distribution]**

  The PDF of the Rayleigh distribution is given by

$$f(r) = r\, e^{-r^2/2}, \quad r > 0.$$

  Use the inverse transform method to generate a sample from the above distribution.

Gaussian Samples on Python via ITT

- Python's built-in module

  $$\texttt{numpy.random.normal(loc, scale, size)}$$

  generates n independent samples from $\mathcal{N}(\mu, \sigma^2)$, where

  $$n = \texttt{size}, \qquad \mu = \texttt{loc}, \qquad \sigma = \texttt{scale}.$$

- In principle, the above module uses the inverse transform technique

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Gaussian Samples on Python via ITT

1. Let $U_1, U_2 \overset{\text{i.i.d.}}{\sim} \text{Unif}(0, 1)$
2. Let R and Θ be two random variables defined via
$$R = F_1^{-1}(U_1), \qquad \Theta = 2\pi U_2,$$
where F_1 is the CDF of the Rayleigh distribution
3. Let Y_1 and Y_2 be defined as
$$Y_1 = R\,\cos(\Theta), \qquad Y_2 = R\,\sin(\Theta).$$
4. Then, $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$
5. To get $X \sim \mathcal{N}(\mu, \sigma^2)$, simply discard Y_2, and
$$X = \sigma\, Y_1 + \mu.$$
6. Repeat steps 1-5 a total of n times to get $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$