Lecture 7: Basic Sampling

Professor Ilias Bilionis

Pseudo-random number generators

PREDICTIVE
SCIENCE LABORATORY

Pseudo-random number generators

- Computers are deterministic machines and therefore they cannot generate completely random numbers?

- Idea: Are there deterministic sequences of numbers that look random?

- Pseudo-random number generators do exactly that.

- We use statistical tests to see how good they are.

Pseudo-random number generators

How do you generate a uniform random number?

John von Neumann. (Los Alamos)

The middle-square method

412523
seed

170175225529
*seed * seed*

175225
output

output becomes seed

The first, but it doesn't pass all statistical tests.

Linear congruential generators

Seed x_0

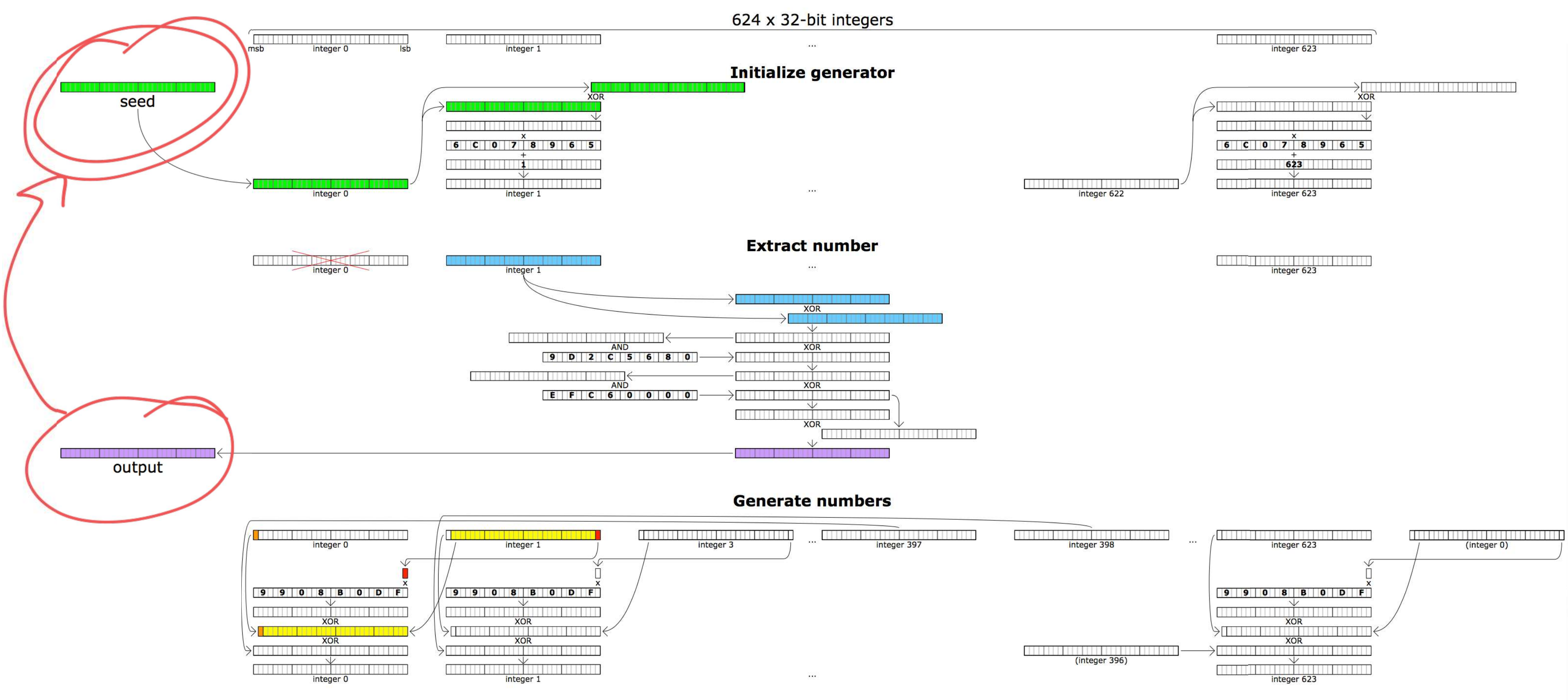
$$x_{i+1} = (a x_i + b) \mod m$$

big number

Mersenne Twister PRNG

- This is what is inside numpy.random.

- Details beyond the scope of this class.



Lecture 7: Basic Sampling

Professor Ilias Bilionis

Sampling the uniform

PRNG to uniform

- PRNG's generate random integers from 0 to m.
- How can we get samples from the uniform?
- Step 1: Sample a random integer d.
- Step 2: Set:

$$x = \frac{d}{m}$$

PRNG to Uniform

How do we know that the samples are indeed uniform?

$X \sim U([0,1])$

$F(x) = P[X \le x] = x$

We can compare the empirical CDF with the ideal CDF.

$\hat{F}_N(x)$

samples

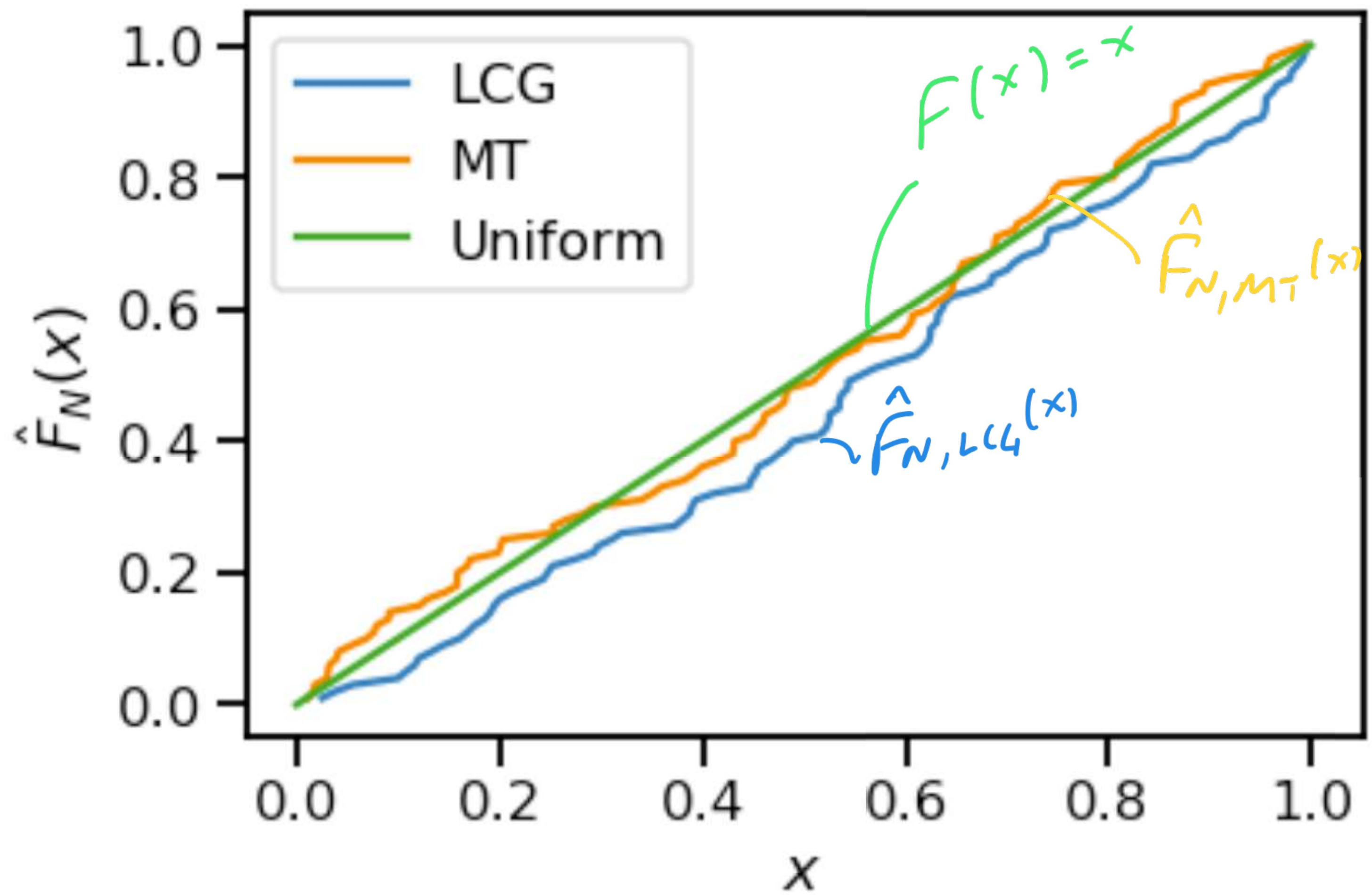
But what is the empirical CDF of a bunch of samples $x_{1:N}$?

It is defined as follows:

$$\boxed{\hat{F}_N(x) = \frac{\text{number of elements in sample} \le x}{N}}$$

PREDICTIVE
SCIENCE LABORATORY

How do we know that the samples are indeed uniform?



PREDICTIVE
SCIENCE LABORATORY

Lecture 7: Basic Sampling

Professor Ilias Bilionis

Sampling the categorical

Example: Sampling from the Bernoulli distribution

$$X \sim \text{Bernoulli}(\theta); \quad X = \begin{cases} 1, & \text{w/pr. } \theta \\ 0, & \text{otherwise} \end{cases}$$

To sample from it, we do the following steps:

$U \sim U([0,1])$

$X = \begin{cases} 1, & \text{if } U \leq \theta \\ 0, & \text{otherwise} \end{cases}$

$p(X=1) = P(U \leq \theta) = F_U(\theta) = \theta$

$p(X=0) = 1-\theta$

- Sample a uniform number u. $\sim U([0,1])$
- If $u \leq \theta$, then set $x = 1$.
- Otherwise, set $x = 0$.

PREDICTIVE SCIENCE LABORATORY

Sampling discrete distributions

K : possibilities

$0, 1, 2, \dots, K-1$

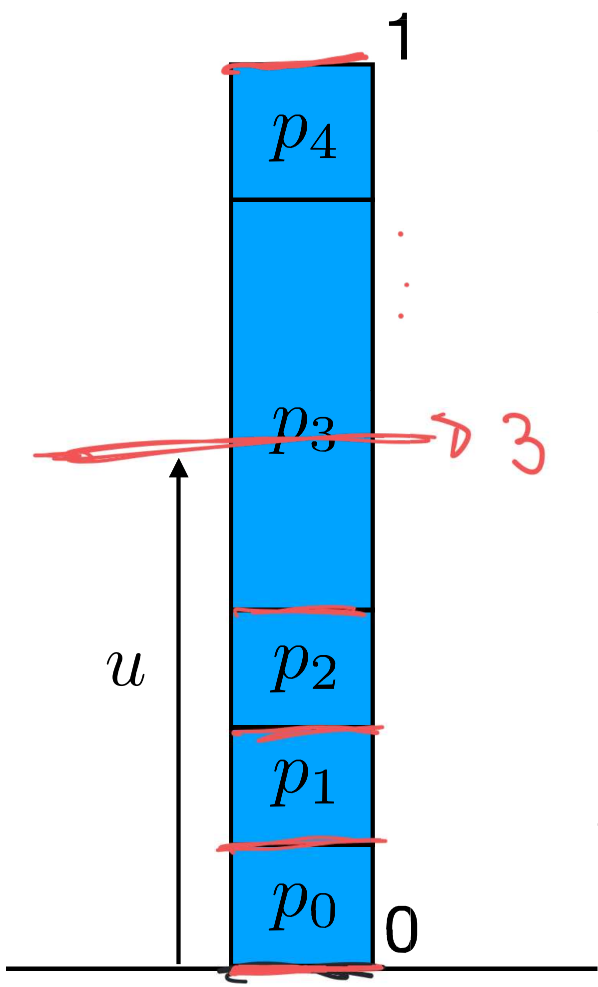
$p_0, p_1, p_2, \dots, p_{K-1}$

- Consider a generic discrete random variable taking different values, with probability:

$$\boxed{p(X = k) = p_k.}$$

$$X \sim \text{Categorical}(p_0, p_1, \dots, p_{K-1}) \; ; \; X = \begin{cases} 0, & \text{w/ pr. } p_0 \\ 1, & \text{w/ pr. } p_1 \\ \vdots & \\ K-1, & \text{w/ pr. } p_{K-1} \end{cases}$$

Sampling Discrete Distributions



- Draw a uniform number u. $\sim U([0,1])$

- Find j such that:

$$\sum_{k=0}^{j-1} p_k \leq u < \sum_{k=0}^{j} p_k$$

- j is your sample

Sampling Discrete Distributions

$$U \sim U([0,1])$$

$$X = \begin{cases} 0, & \text{if } U \leq p_0 \\ 1, & \text{if } p_0 \leq U \leq p_0 + p_1 \\ 2, & \text{if } p_0 + p_1 \leq U \leq p_0 + p_1 + p_2 \\ \vdots & \\ K-1, & \text{otherwise} \end{cases}$$

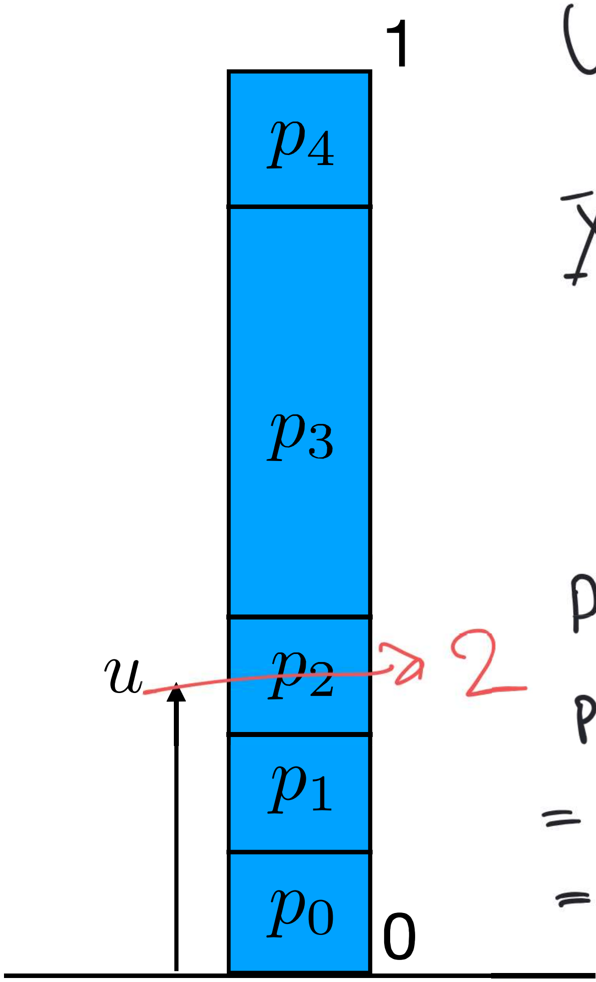
$$P(X=0) = P(U \leq p_0) = p_0$$

$$P(X=1) = P(p_0 \leq U \leq p_0 + p_1)$$

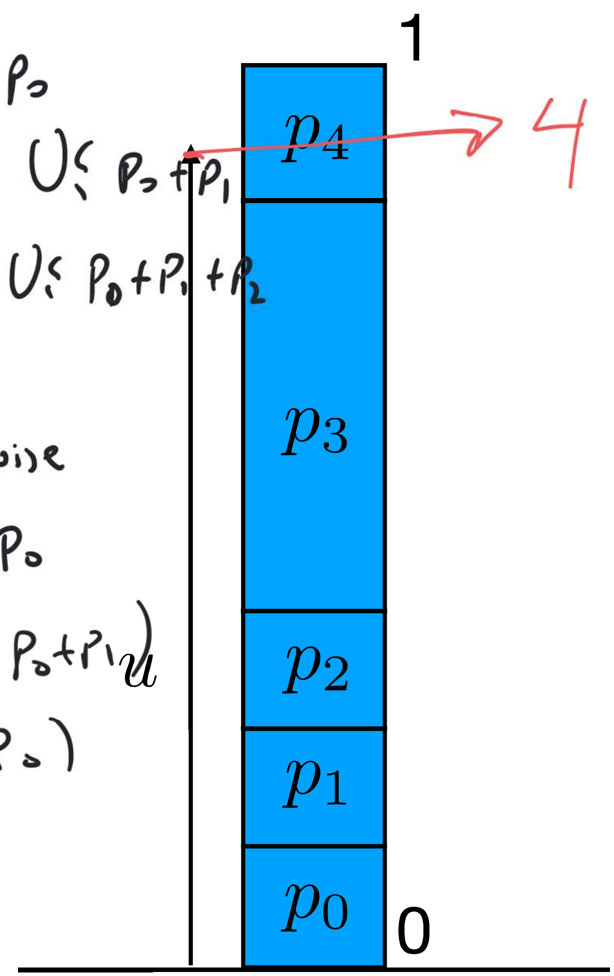
$$= F_U(p_0 + p_1) - F_U(p_0)$$

$$= \cancel{p_0} + p_1 - \cancel{p_0} = p_1$$

$$P(X=2) = \ldots$$



Sample 1



Sample 2

Lecture 7: Basic Sampling

Professor Ilias Bilionis

Inverse sampling

PREDICTIVE
SCIENCE LABORATORY

Inverse Sampling

- Consider an arbitrary univariate continuous random variable X with CDF $F(x)$, How do you sample from it?

- Draw a uniform number u. $\sim U([0,1])$

- Set:

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

inverse of the CDF

- and you get your sample!

Why does inverse sampling work?

- Let $U \sim U([0,1])$ be a uniform random variable.

- For any CDF $F(x)$ define the random variable:

$$X = F^{-1}(U)$$

- The CDF of X is:

$$p(X \leq x) = p\left(F^{-1}(U) \leq x\right) = p\left(F(F^{-1}(U)) \leq F(x)\right)$$
$$= p(U \leq F(x)) = F_U(F(x)) = F(x)$$

PREDICTIVE
SCIENCE LABORATORY

Example: The exponential distribution

- Take an exponential random variable as an example:

$$X \sim \mathrm{Exp}(r)$$

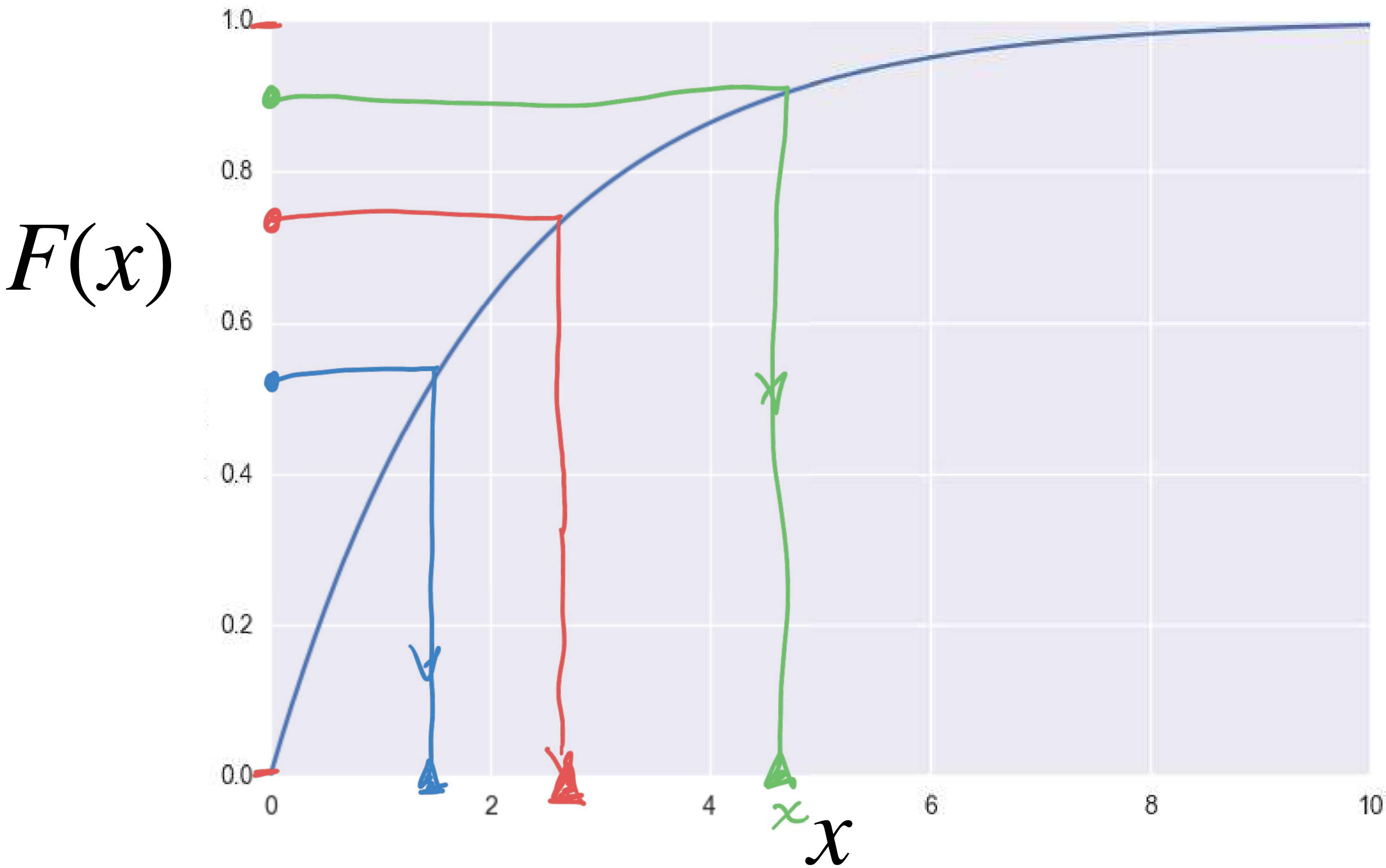
- The CDF is:

$$F(x) = 1 - e^{-rx}$$

- The inverse of the CDF is:

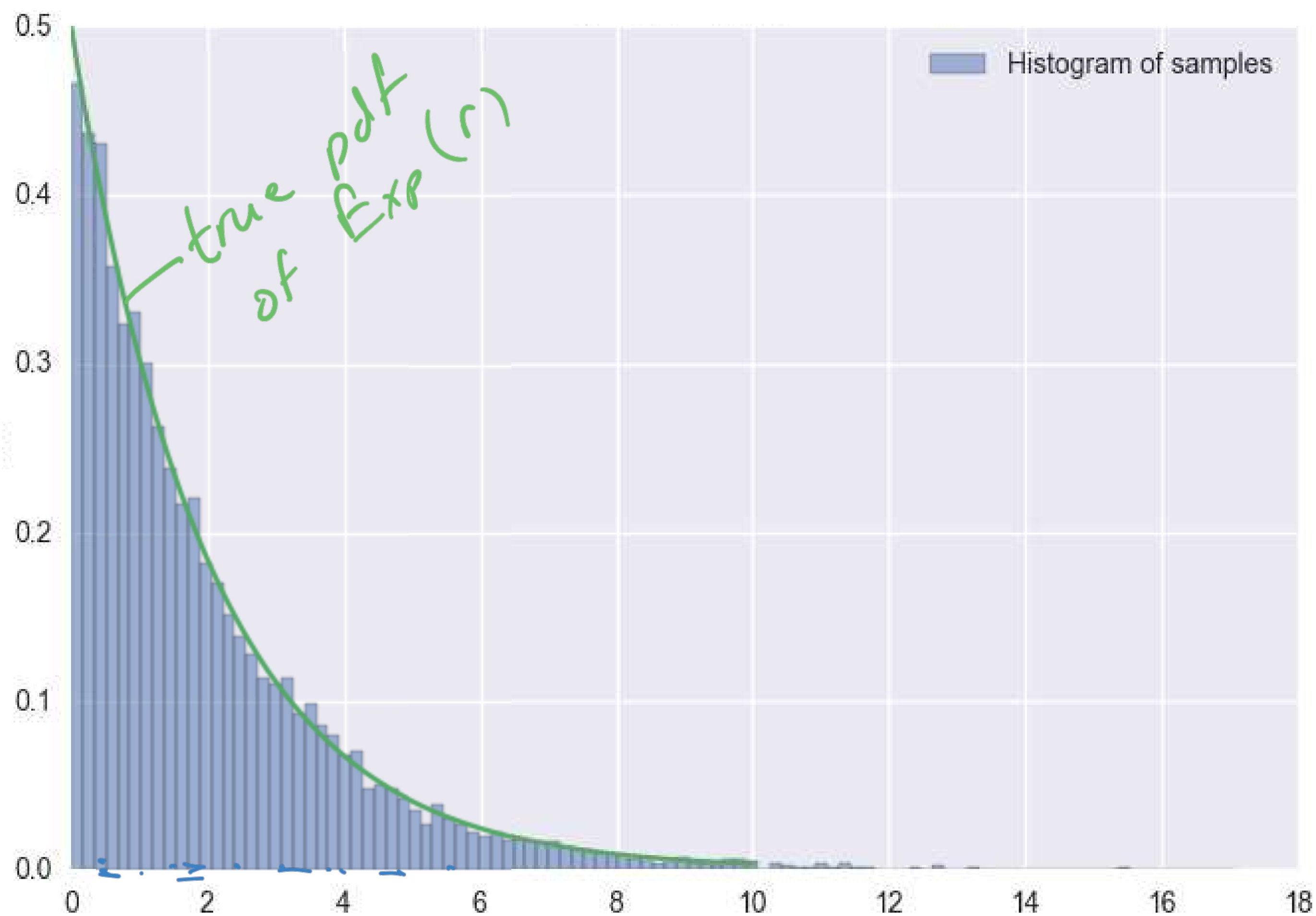
$$F^{-1}(u) = -\frac{\ln(1-u)}{r}$$

The Exponential Distribution



Inverse Sampling for Exponential

$p(x)$



x