Monte Carlo

Non-Uniform Random Numbers

Francesco Banterle, Ph.D. - July 2021

Non-Uniform Random Numbers

Introduction

- Typically, to draw random numbers in a non-uniform way following a given distribution is not an easy task; and it needs to be crafted for each distribution!
- A solution is to convert uniform random number into a non-uniform one.
- How?
  - All the information that we need about how a random variable X is distributed is inside its CDF:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)dx.$$

Inverting the CDF

Inverting the CDF

Main Idea

- How do we extract this information from the CDF?
- Let's say we generate a random value $u \in \mathbf{U}(0,1)$, and we set $X = F_X^{-1}(U)$, we obtain:

$$P\left(X \leq x\right) = P\left(F_X^{-1}(u) \leq x\right) = P\left(F_X(F_X^{-1}(u)) \leq F_X(x)\right) =$$

$$P(u \leq F_X(x)) = F_X(x).$$

- In this way, we can have X values with F_X as distribution!

Inverting the CDF

Main Idea

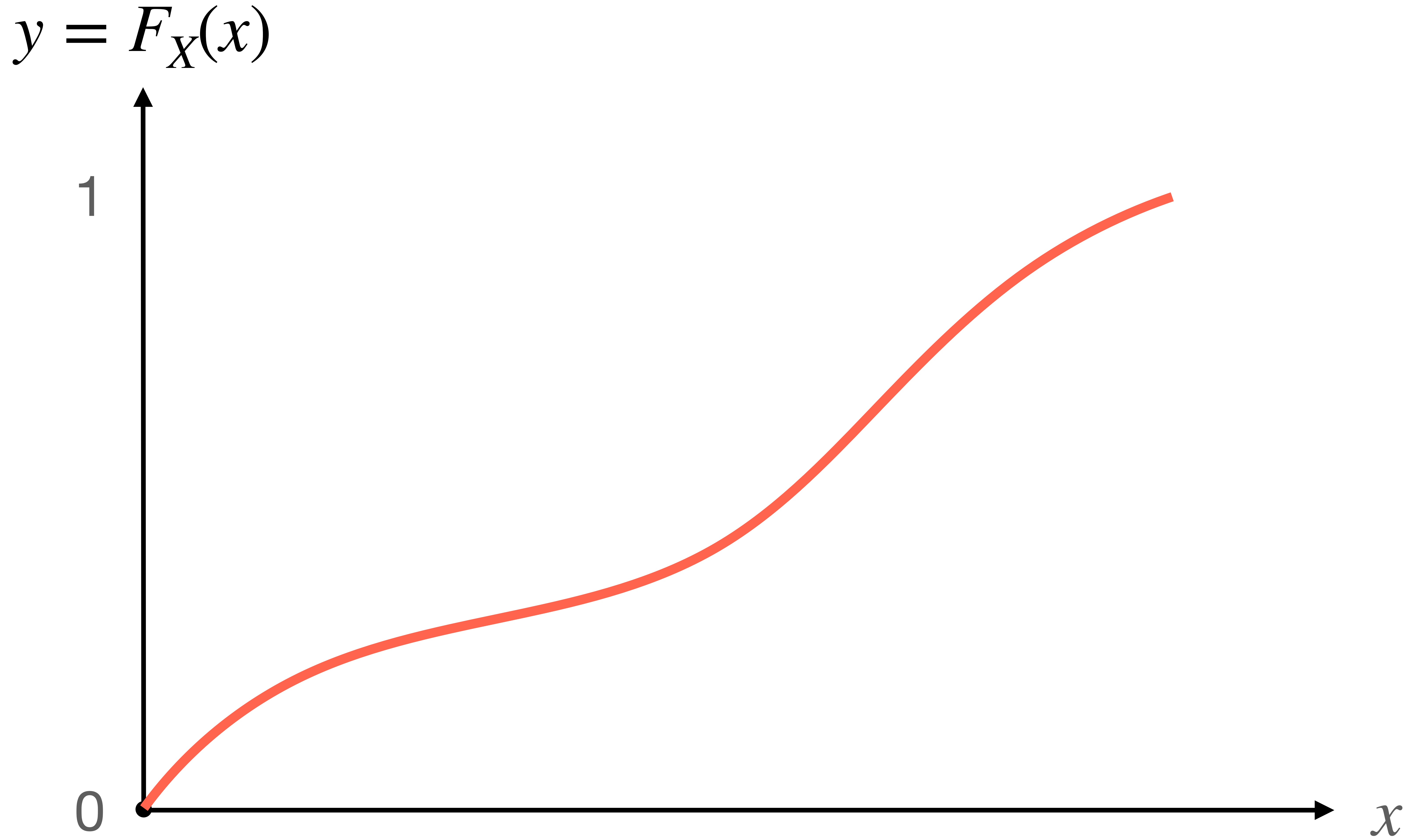
- Given the CDF of a distribution:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} p_X(x)dx.$$

- We generate a non-uniform random numbers as:
  - We first generate a uniform random number, $u \in \mathbf{U}(0,1)$;
  - Then, we compute:

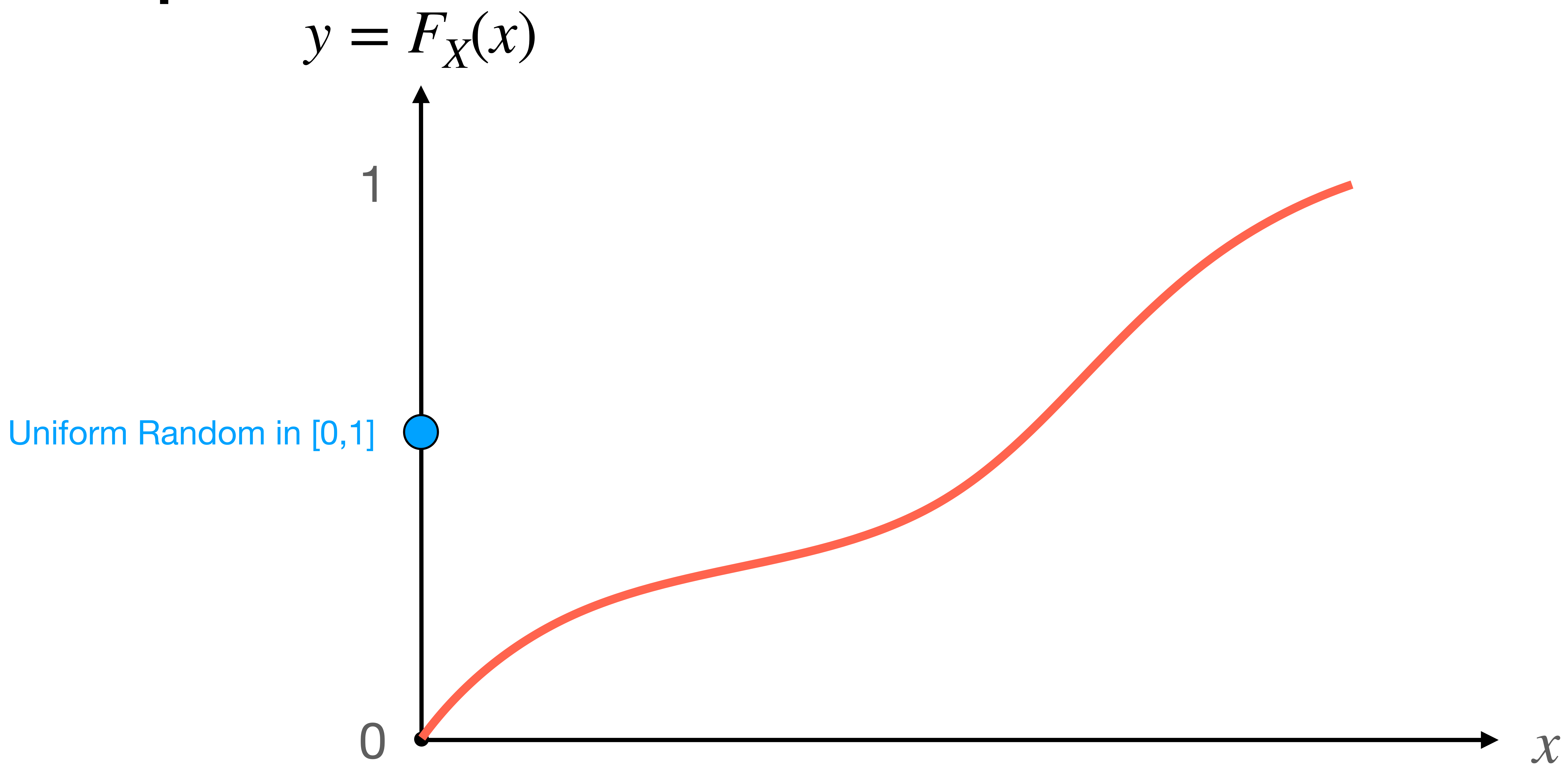
$$u' = F_X^{-1}(u).$$

Inverting the CDF
Example



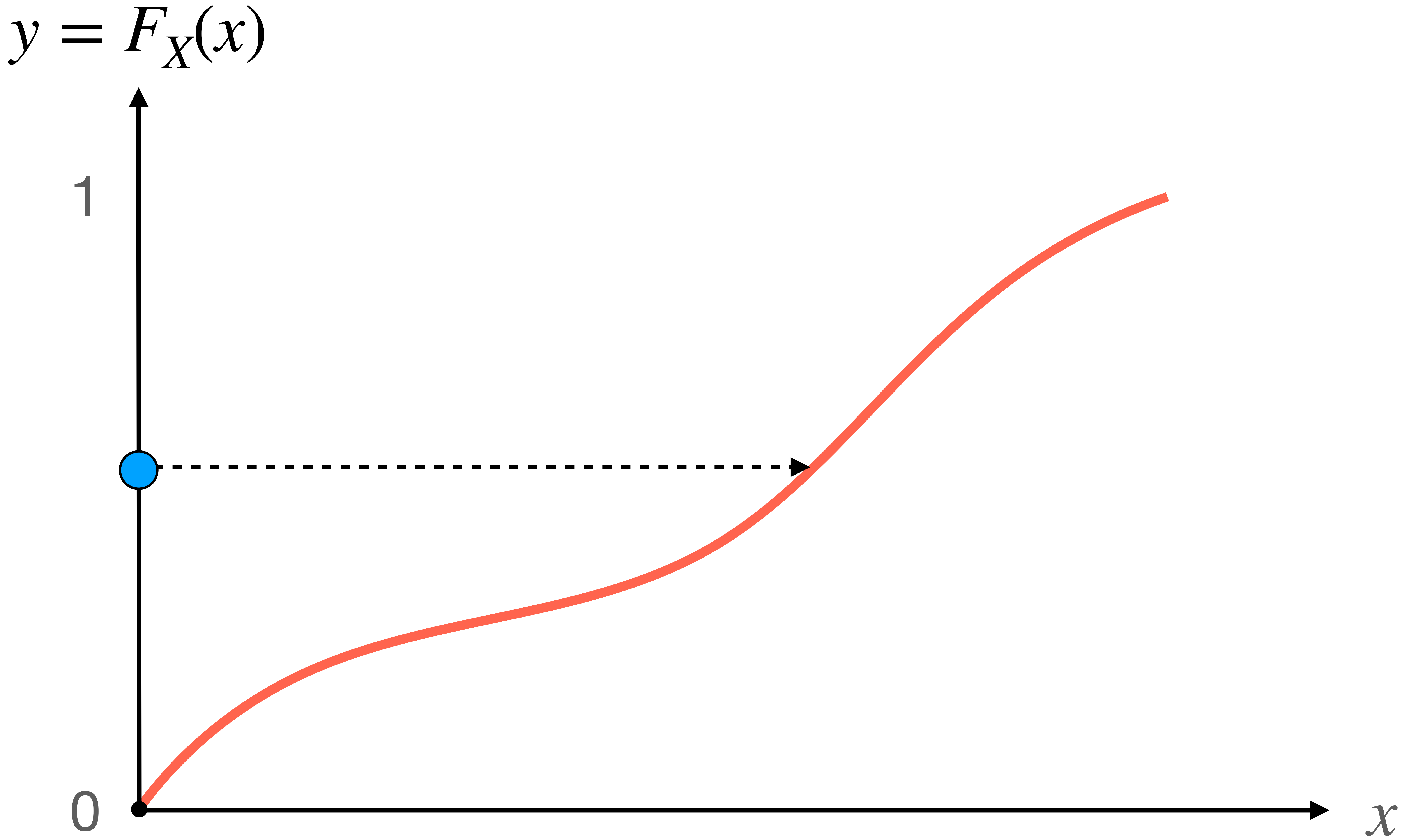
Inverting the CDF

Example



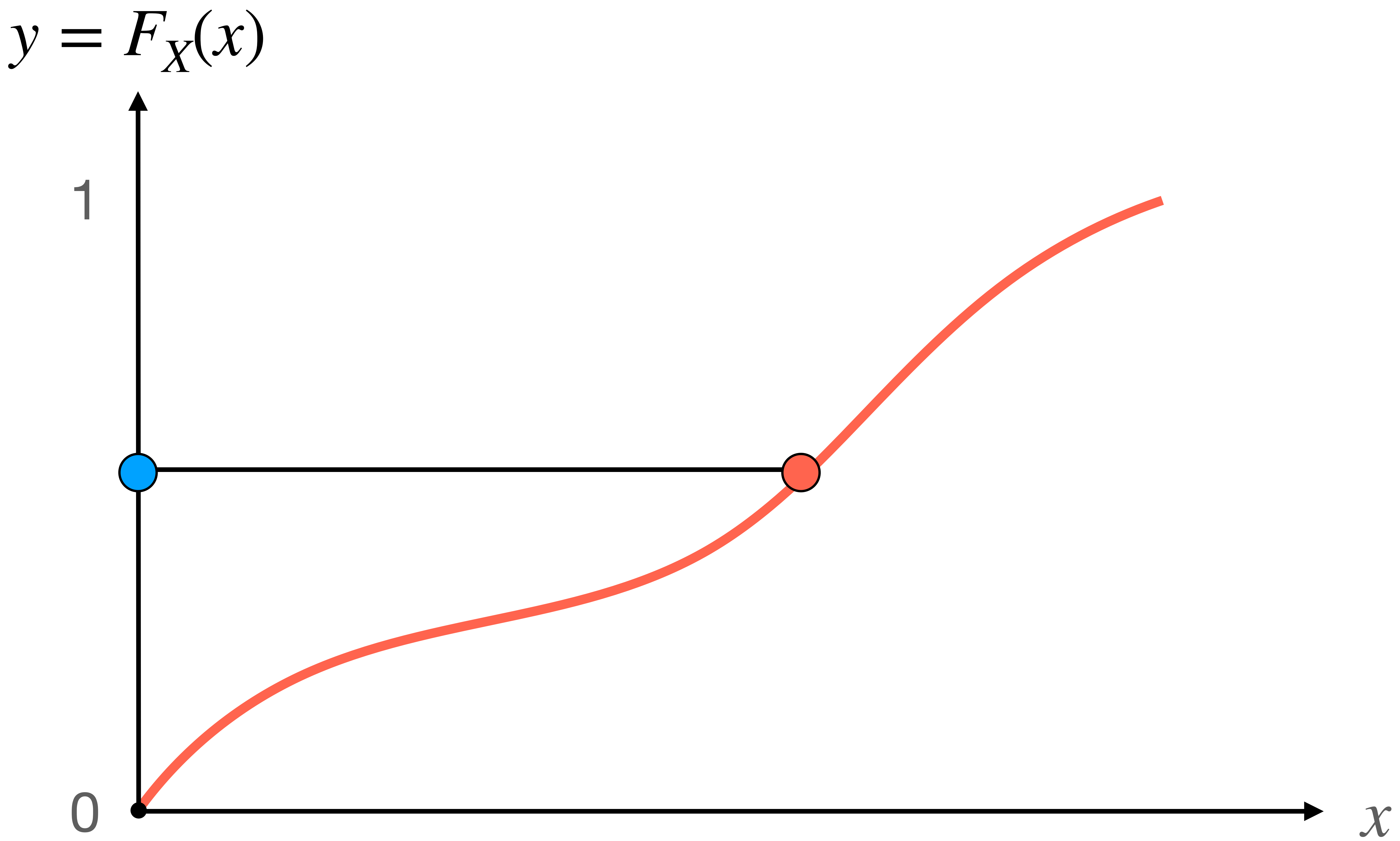
Inverting the CDF

Example



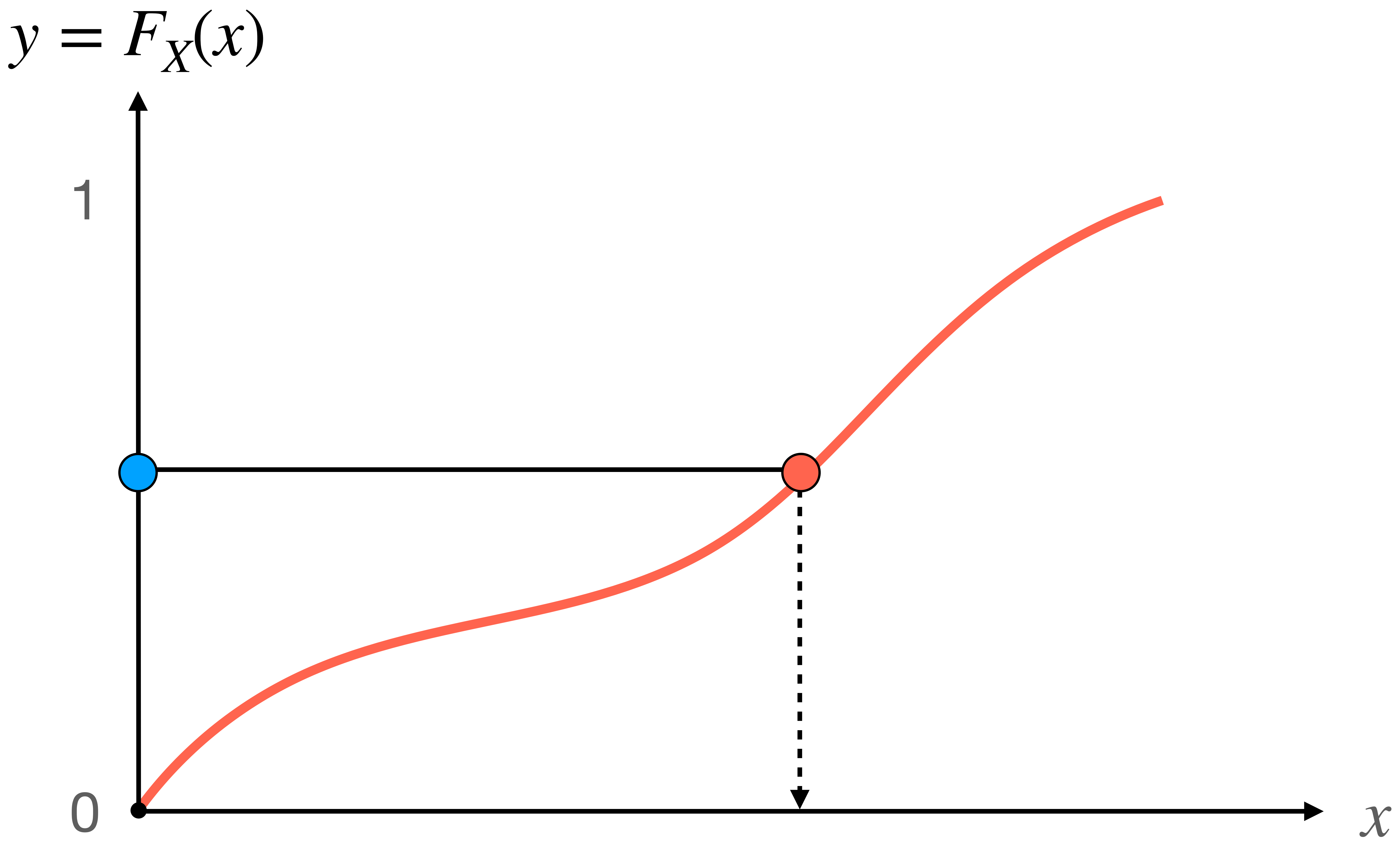
Inverting the CDF

Example



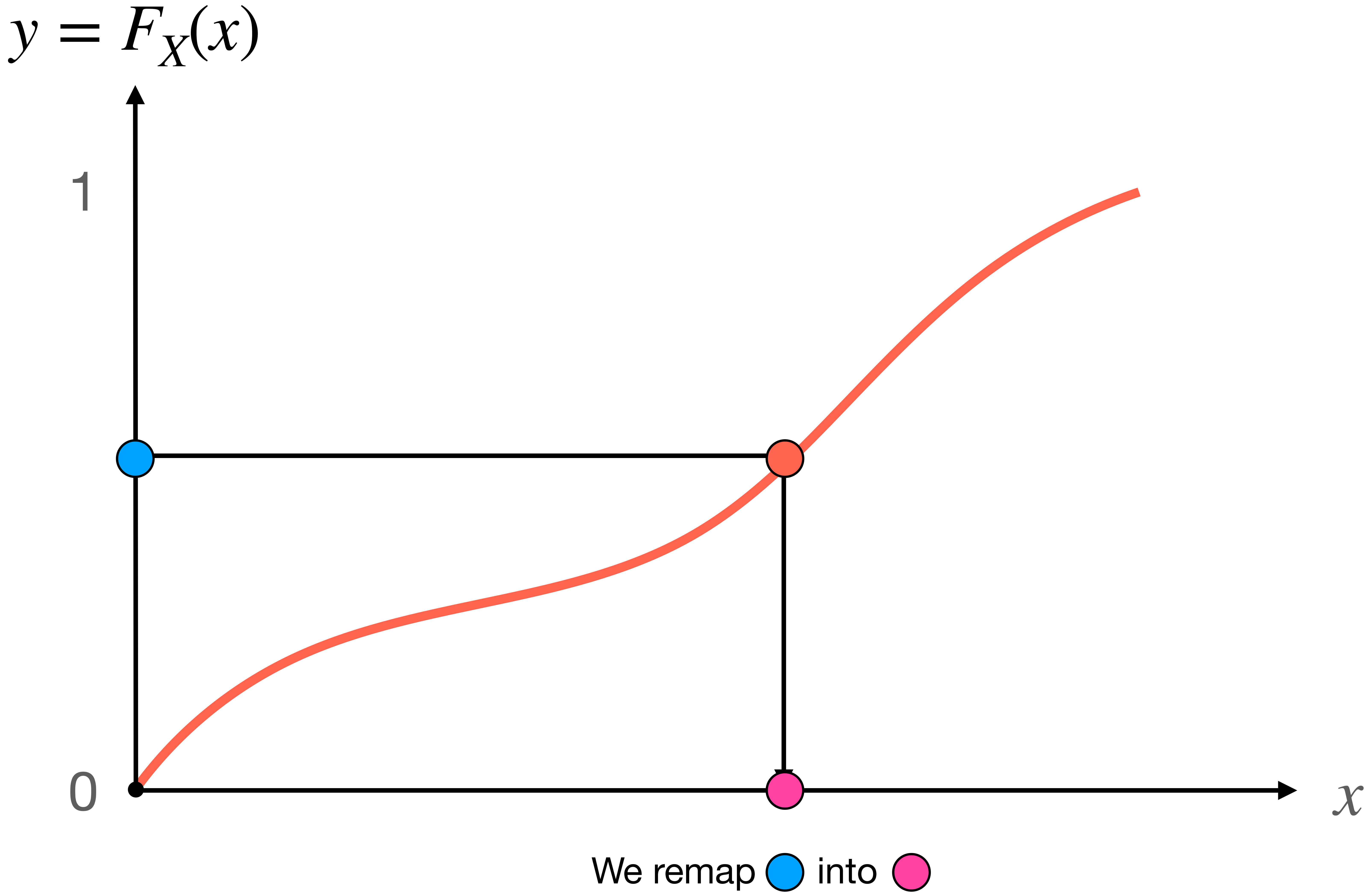
Inverting the CDF

Example



Inverting the CDF

Example



Inverting the CDF

Main Idea

- Note that we draw uniform random numbers $u \in (0,1)$.
- Why?

Inverting the CDF

Main Idea

- Note that we draw uniform random numbers $u \in (0,1)$.
- Why?
  - 0 and 1 may generate some singularities:
    - NaN, +Inf, -Inf

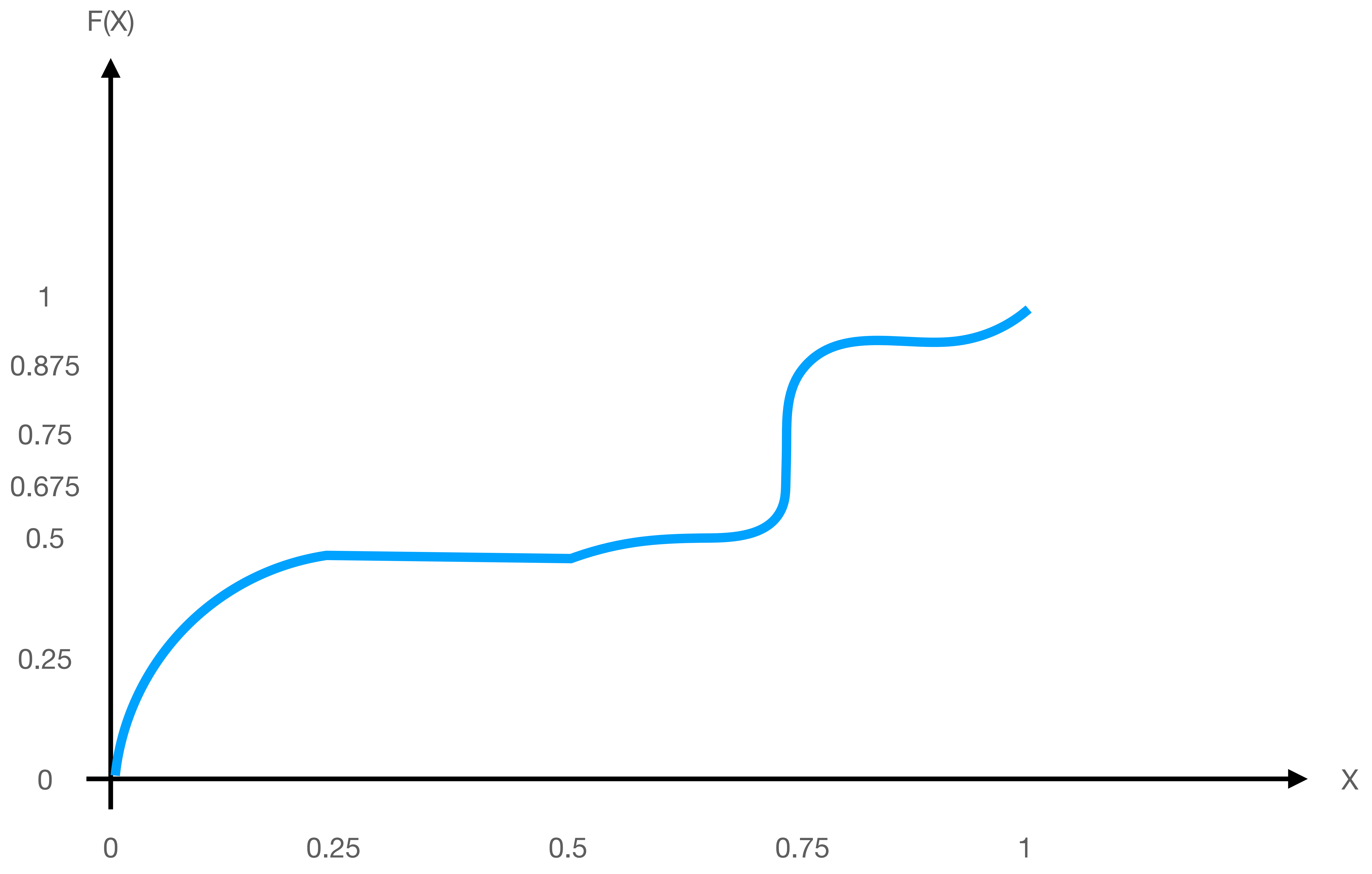
Inverting the CDF

Main Idea

- Note that if $u \sim \mathbf{U}(0,1)$ we have that $1 - u \sim \mathbf{U}(0,1)$.
- This means that $F^{-1}(1-u) \sim F$.
  - In some cases, to compute $F^{-1}(u)$ may be difficult.
  - In these cases the complementary inversion equation may be easier to compute!

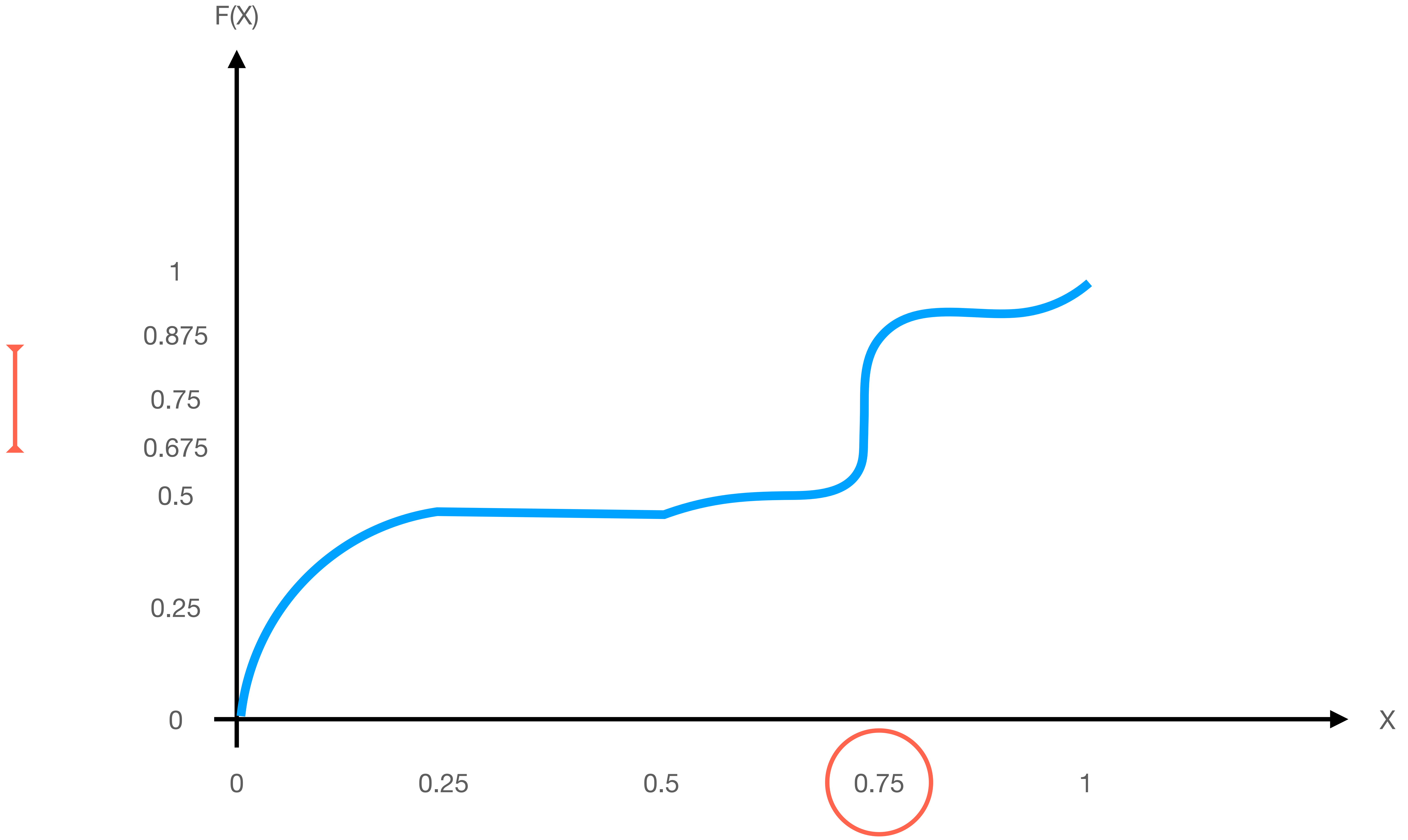
Inverting the CDF

Issues



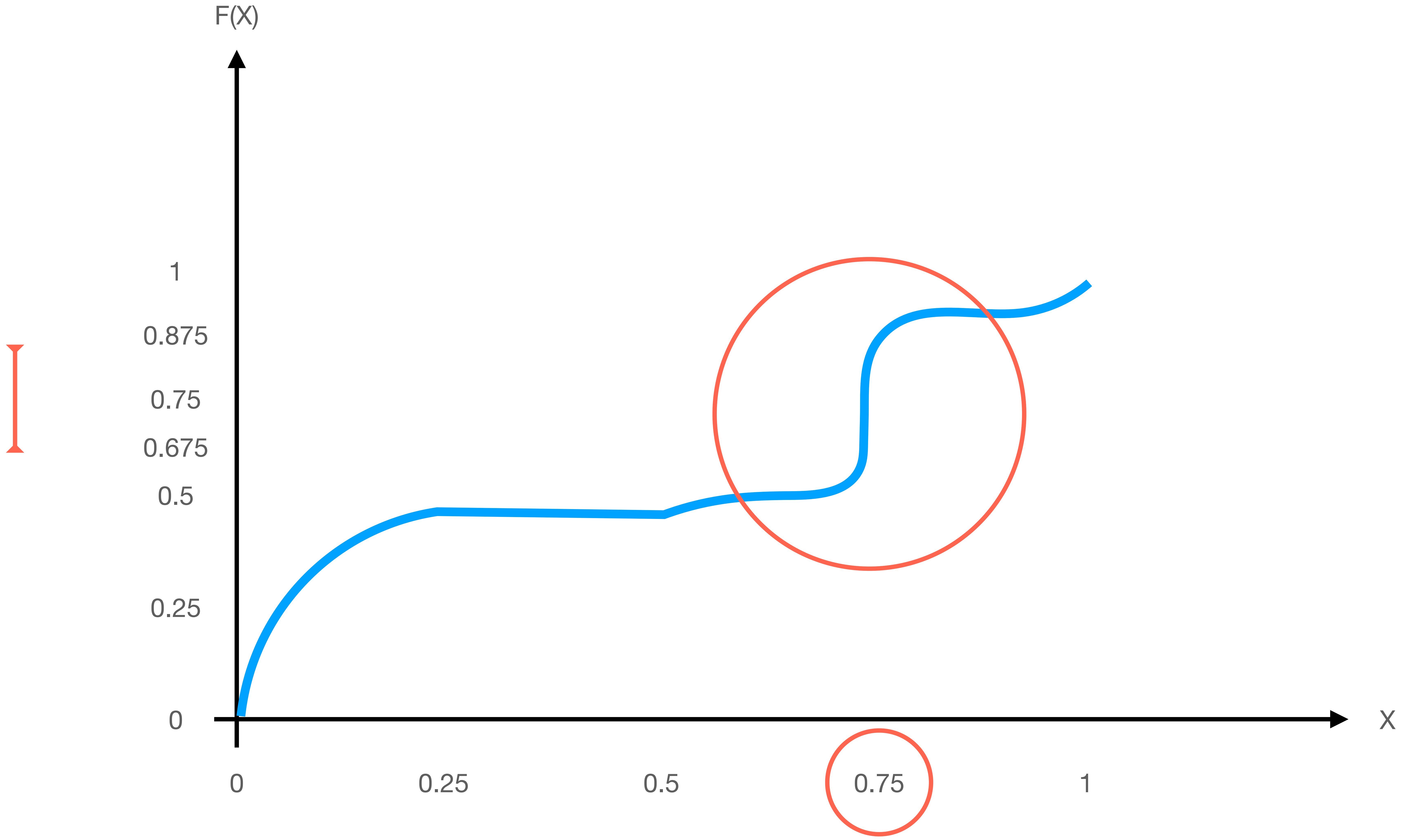
Inverting the CDF

Issues



Inverting the CDF

Issues



Inverting the CDF

Issues

- In such cases, the inverse is not unique, and it can happen for both continuous and discrete distributions!
- A solution to this problem is:

$$F_X^{-1}(u) = \inf\left\{ x \,\middle|\, F_X(u) \geq u \wedge u \in (0,1) \right\}.$$

Inverting the CDF

Example: Uniform Distribution

- The uniform distribution is defined as

$$f(x) = \frac{1}{b-a} \qquad x \in [a, b].$$

- Its CDF is given by:

$$F(x) = \int_{-\infty}^{x} \frac{1}{b-a} dx = \frac{1}{b-a} \int_{-\infty}^{x} dx = \frac{x}{b-a}.$$

- So let's compute its inverse:

$$y = \frac{x}{b-a} \quad \text{multiply both sides by } (b-a)$$

$$x = y(b-a)$$

Inverting the CDF

Example: Exponential Distribution

- Standard exponential distribution is:

$$f(x) = \exp(-x) \qquad x > 0.$$

- Its CDF is given by:

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

- So let's compute its inverse:

$$y = 1 - e^{-x}$$

$y - 1 = -e^{-x}$ add -1 both sides

$1 - y = e^{-x}$ multiply by -1 both sides

$\log(1 - y) = \log(e^{-x})$ apply log to both sides

$x = -\log(1 - y)$ simplify and multiply by -1 both sides

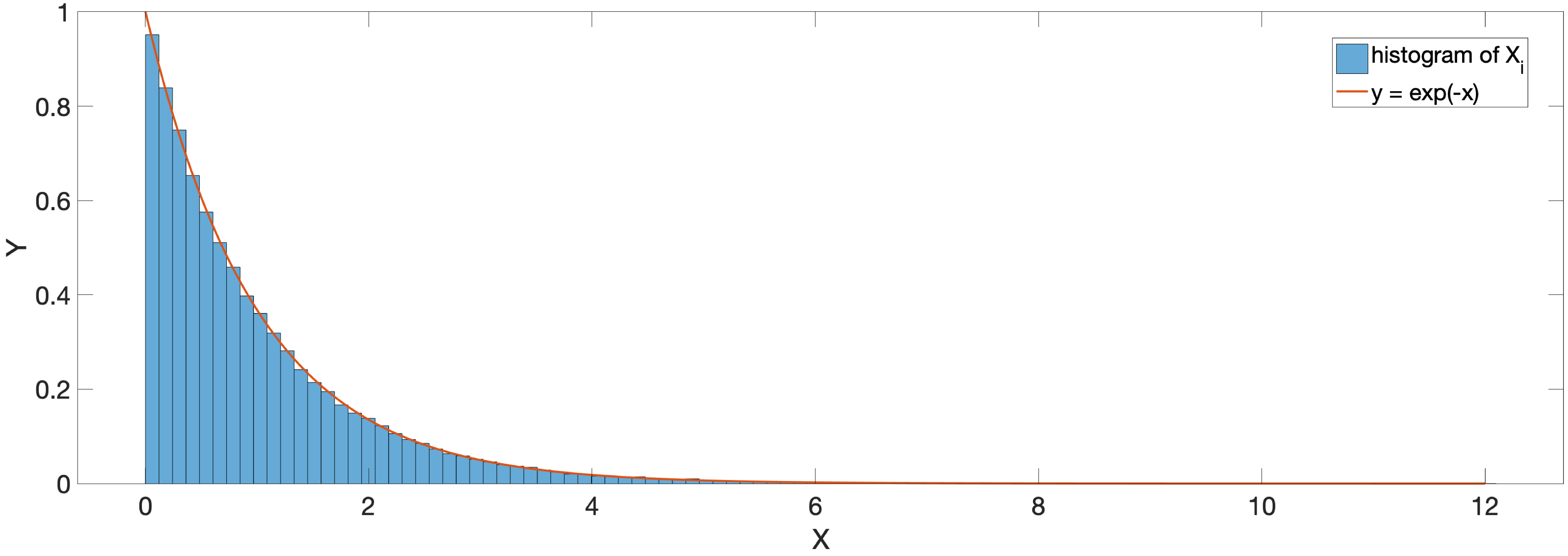
Inverting the CDF

Example: Exponential Distribution

- Now, in order to draw samples exponentially distributed, $X_i \sim \text{Exp}(1)$, we do:
  - $Y_i \in \mathbf{U}(0,1)$;
  - $X_i = -\log(1 - Y_i)$.
- Note that doing the inversion, we have the same distribution and its faster:
  - $Y_i \in \mathbf{U}(0,1)$;
  - $X_i = -\log(Y_i)$.
- In this case it would not be safe to draw 0 and 1 for Y_i because depending on the method it may create a singularity!

Inverting the CDF

Example: Exponential Distribution



Inverting the CDF

Example: Normal Distribution

- Normal distribution $\mathcal{N}(0,1)$:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

- Its CDF is:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{x^2}{2}\right) dx = \Phi(x).$$

- Note that there is not closed form for $\Phi(x)$.

- $\Phi(x)$ is related to the Erf function:

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} \exp\left(-t^2\right) dt \qquad \Phi(x) = \frac{\text{erf}(x/\sqrt{2}) + 1}{2}.$$

Inverting the CDF

Example: Normal Distribution

- In this case, we need to invert $\Phi(x)$ to obtain $\Phi^{-1}(x)$:
  - There is no closed-form for $\Phi^{-1}(x)$.
  - Typically, we have algorithms for erf and its inverse:

$$\Phi^{-1}(x) = 2\pi \text{erf}^{-1}(2x - 1).$$

  - We need to use an approximation such as the AS70:
    - R. E. Odeh and J.O. Evans. “Algorithm AS 70: the percentage points of the normal distribution”. Applied Statistics, 23(1):96-97. 1974.

Inverting the CDF

Transformations: Linear Transformation

- In some cases, if we have a distribution F with mean 0 and variance 1, we may want to shift its mean by μ and scale it to have variance $\sigma^2 > 1$:
  - $X \sim F_X \to Y = \sigma X + \mu$, and Y is our random variable with the desired distribution.
  - To achieve this, we have to:

$$f_Y(y) = \frac{1}{\sigma} f_X\left(\frac{x - \mu}{\sigma}\right).$$

Inverting the CDF

Transformations

- Transformations can be very general. Let's assume:
  - $X \sim F_X$;
  - $Y = \tau(X)$ where τ is an invertible increasing function. This means:

$$P(Y \leq y) = P(\tau(X) \leq y) = P(X \leq \tau^{-1}(y)).$$

  - Therefore, Y has the following PDF:

$$f_Y(y) = \frac{d}{dy} P\big(X \leq \tau^{-1}(y)\big) = f_X(\tau^{-1}(y)) \frac{d}{dy} \tau^{-1}(y).$$

  - Note that:

$$\frac{d}{dx} P(X \leq x) = \frac{d}{dx} \left(\int_{-\infty}^{x} f_X(x) dx \right).$$

Inverting the CDF

Transformations: An Example

- Let's define:

$$\tau(x) = x^p \text{ where } p > 0.$$

  - Let's assume that $X \sim \mathbf{U}(0,1)$:
    - This means: $Y = \tau(X) = X^p$ with PDF:

$$f_Y(y) = \frac{1}{p} y^{\frac{1}{p}-1} \quad y \in (0,1).$$

Inverting the CDF

Numerical Inversion

- It can happen that we may have F, but we cannot invert it.
- In such cases there are other options:
  - We can use bisection algorithms to search x such that $F(x) = u$.
  - Although bisection can get the job done, it is very slow. Another viable option is to Newton's method:

$$x_{i+1} = x_i - \frac{F(x_i) - u}{f(x_i)}.$$

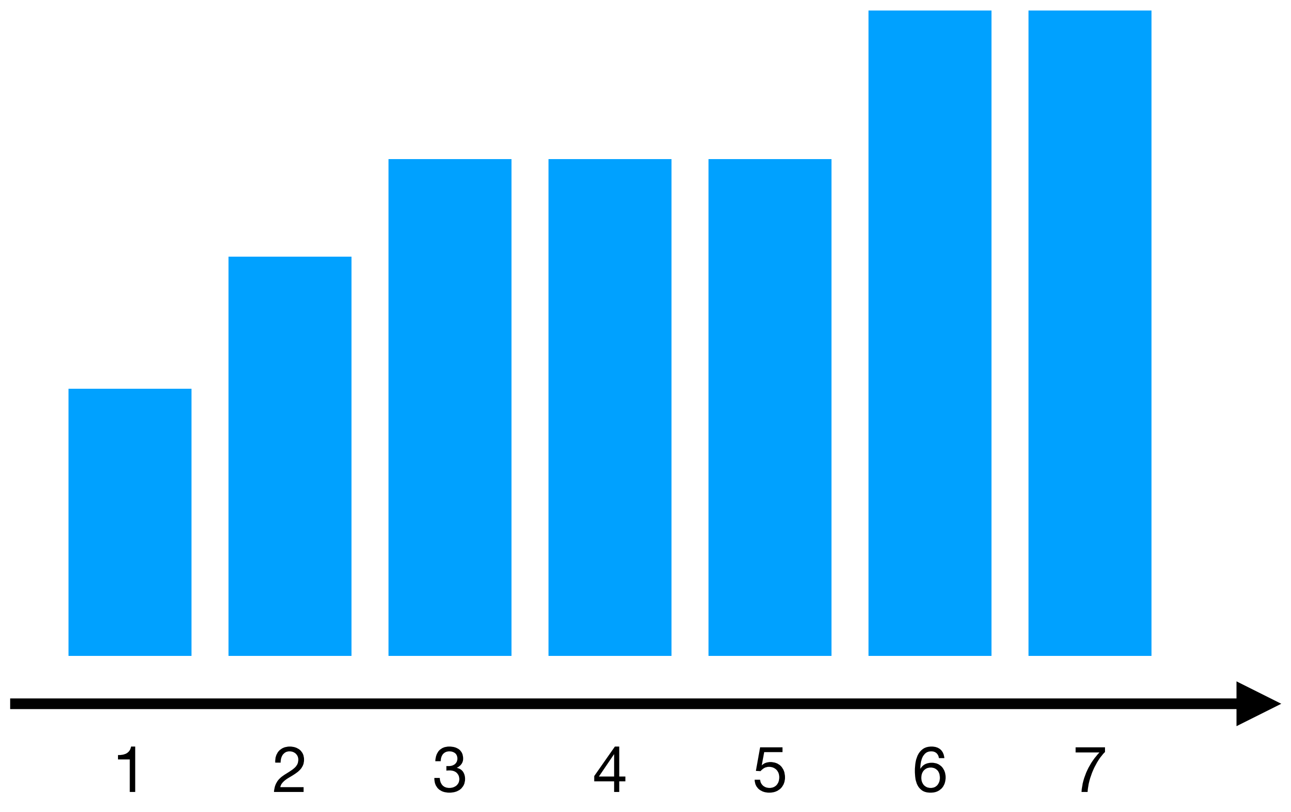
  - The only issue here is that this method may not converge when f is close to 0.

Inverting the CDF

Inversion for Discrete Random Variables

- In many situations, we may face to have discrete distributions; i.e., histograms.

- In a histogram H, we have $1, \ldots, N$ bins and each bin has a frequency number associated to that bin.



- We can convert a histogram into a discrete by normalizing it (i.e., sum of all $H[i]$) obtaining H'.

Inverting the CDF

Inversion for Discrete Random Variables

- At this point, we have can define a random variable X such that

$$P(X = k) = p_k = H'[k] \geq 0.$$

- In this case, the cumulative distribution is defined as:

$$P_k = \sum_{i=1}^{k} p_i \text{ with } P_0 = 0.$$

- In order to compute:

$$F^{-1}(u) = k \quad u \in (P_{k-1}, P_k],$$

we have to run the binary search on the cumulative distribution using $u \sim \mathbf{U}(0,1)$.

Acceptance-Rejection

Acceptance-Rejection

Main Idea

- In some cases, we cannot use the inversion method to get the F distribution that we want.
- When this happens, we can employ another distribution G; key concepts:
  - We reject some values from G;
  - We accept other values from G;
  - In accepting and rejecting, we try to get F.

Acceptance-Rejection

Main Idea

- The first step is to find a distribution G such that its PDF $g(x)$:
  - $f(x) \leq cg(x) \quad c \geq 1$ always holds;
  - We can compute:

$$\frac{f(x)}{g(x)}.$$

Acceptance-Rejection

Main Idea

repeat

$Y \sim g;$

$U \sim \mathbf{U}(0,1);$

until $U \leq f(Y)/(cg(Y))$

$X \leftarrow Y$

return X

Acceptance-Rejection

Main Idea

repeat

$Y \sim g;$

$U \sim \mathbf{U}(0,1);$

until $U \leq f(Y)/(cg(Y))$

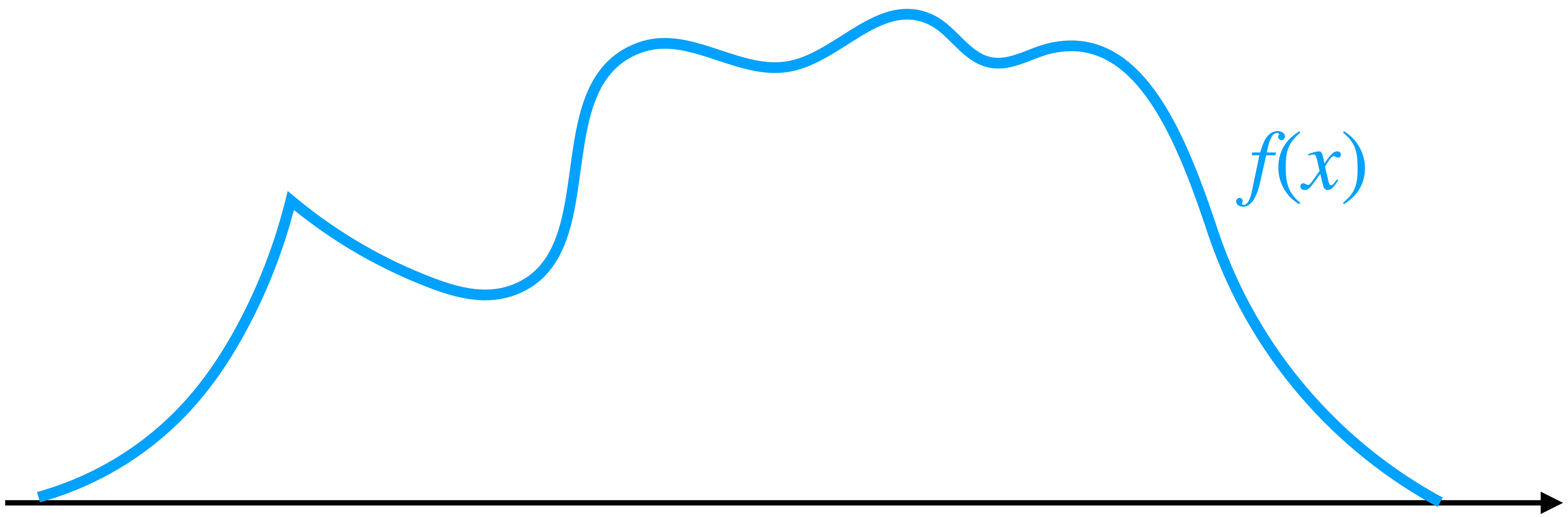
$X \leftarrow Y$

return X

Theorem 4.2 in Owen's book tells us that the generated samples have PDF f.

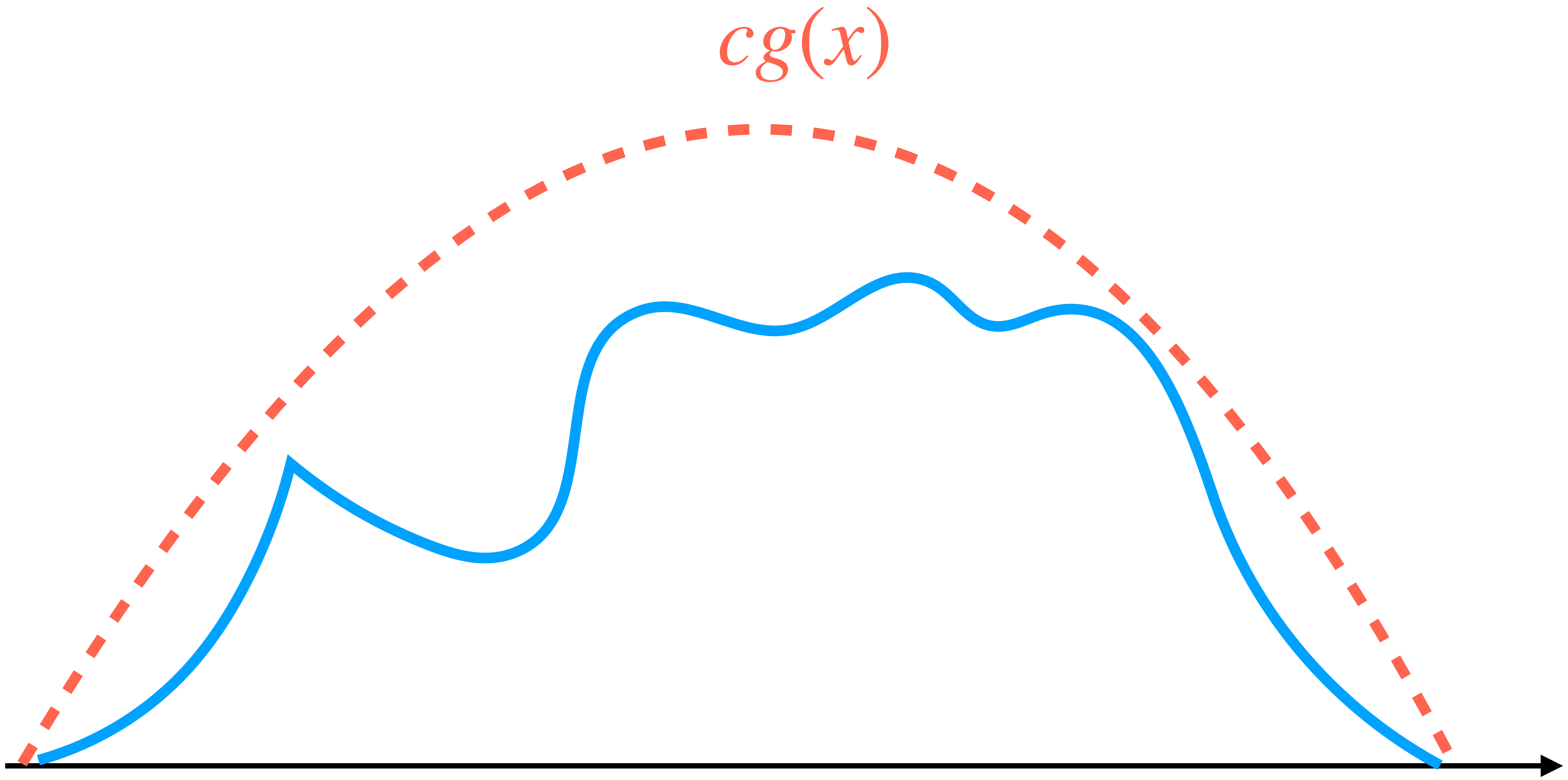
Acceptance-Rejection

Main Idea



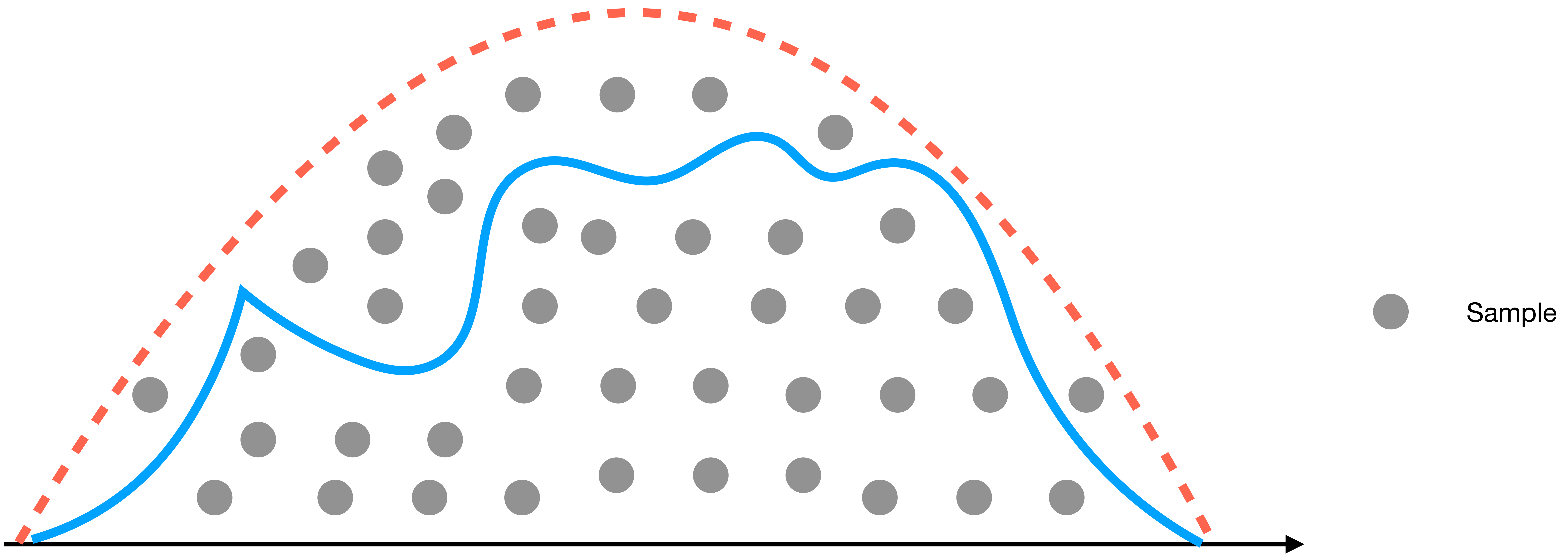
Acceptance-Rejection

Main Idea



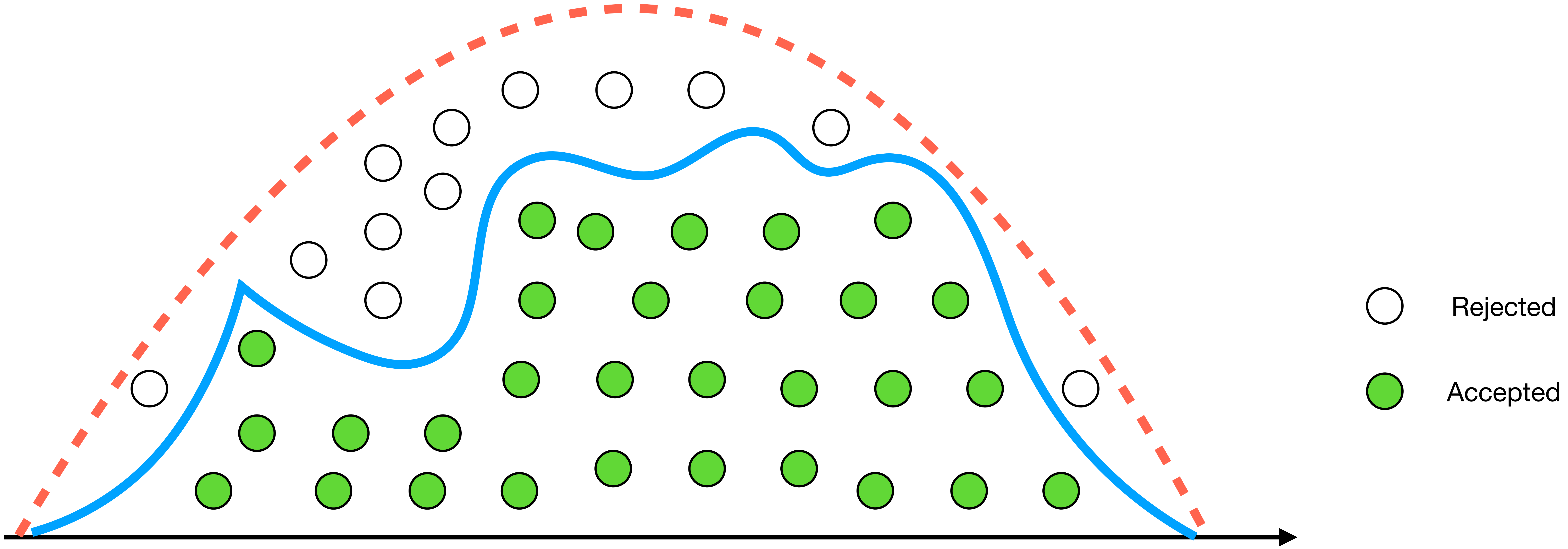
Acceptance-Rejection

Main Idea



Acceptance-Rejection

Main Idea



Acceptance-Rejection

The Ziggurat Algorithm

- The Ziggurat algorithm is an acceptance-rejection method for drawings sampling according to normal distribution (i.e., half).

- The method divides the region below $\mathcal{N}(0,1)$ into k (e.g., 256) horizontal regions that are ideally of similar area; i.e., equiprobable.

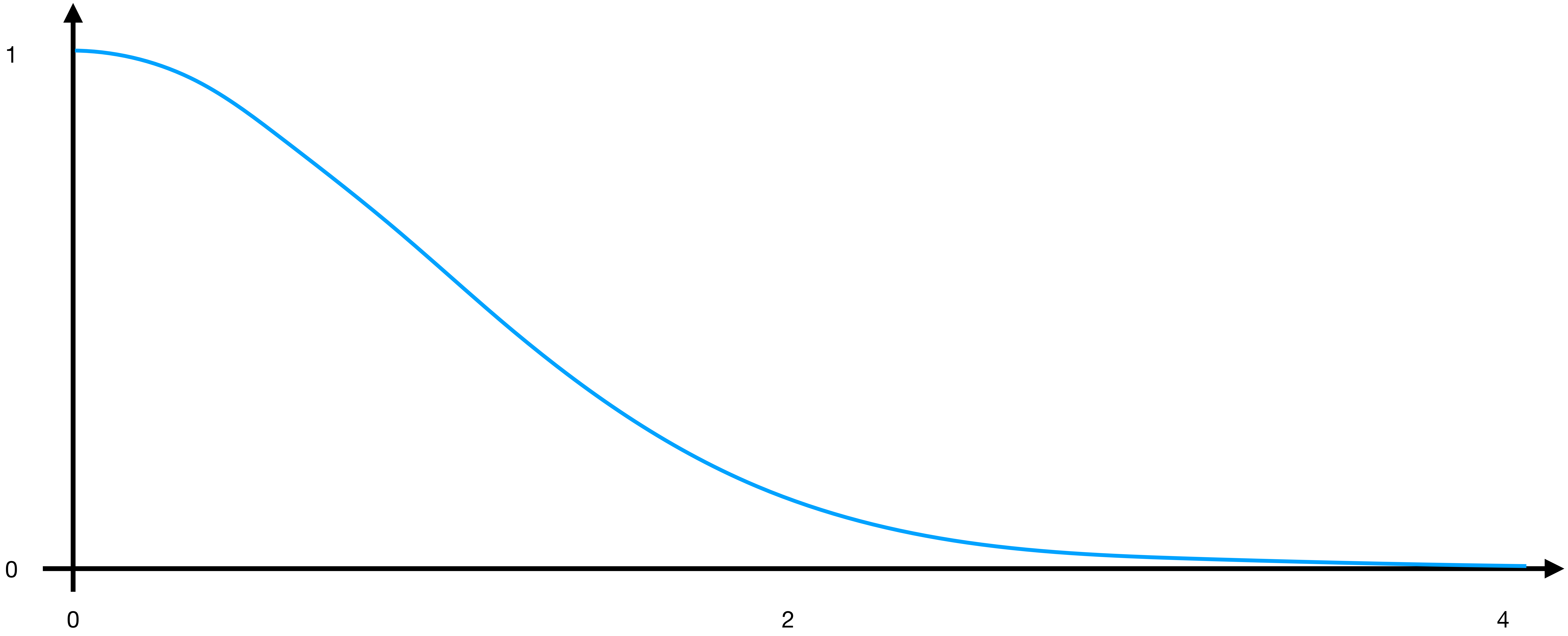
- At this point, the method generate samples points (Z, Y) uniformly distributed in each region such that:

$$\left\{ (z, y) \middle| y \in [0, \exp(-z^2/2); x \in [0,\infty)] \right\}.$$

- Typically, the normalization factor $1/\sqrt{2\pi}$ is ignored for speeding the algorithm up.

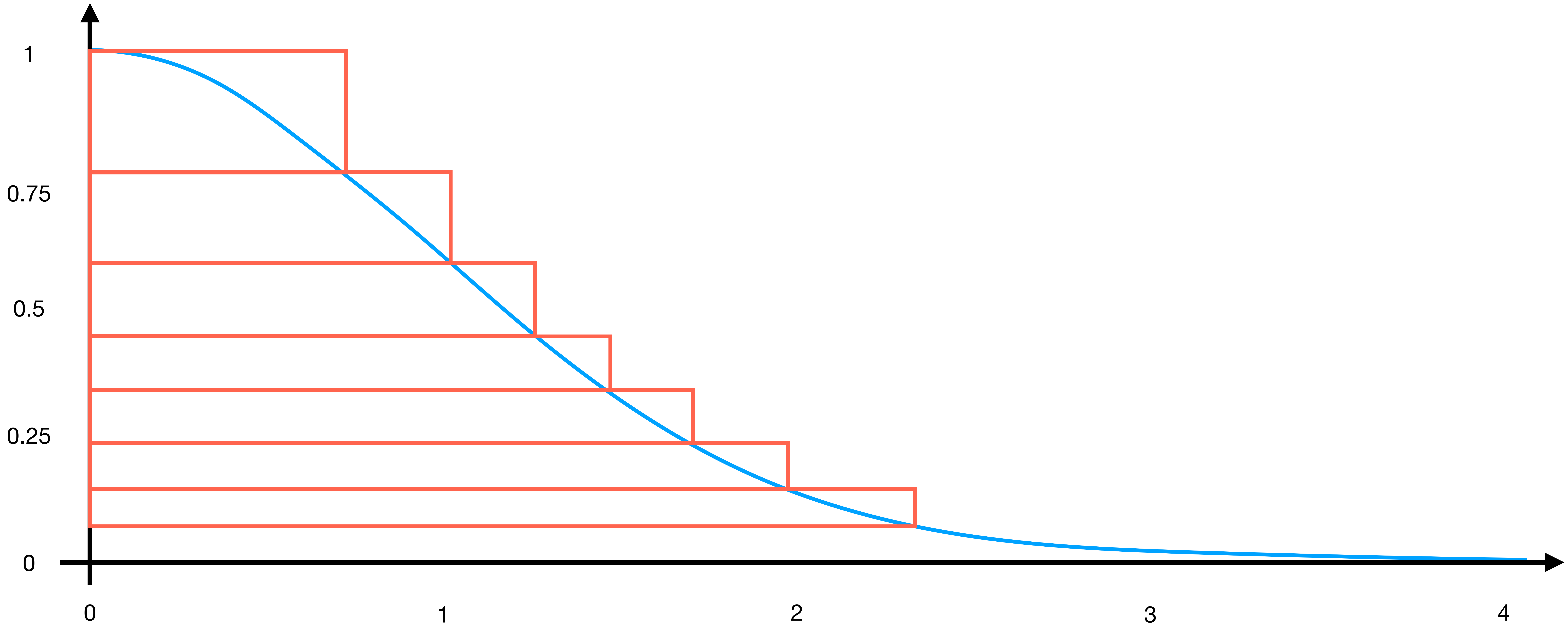
Acceptance-Rejection

The Ziggurat Algorithm



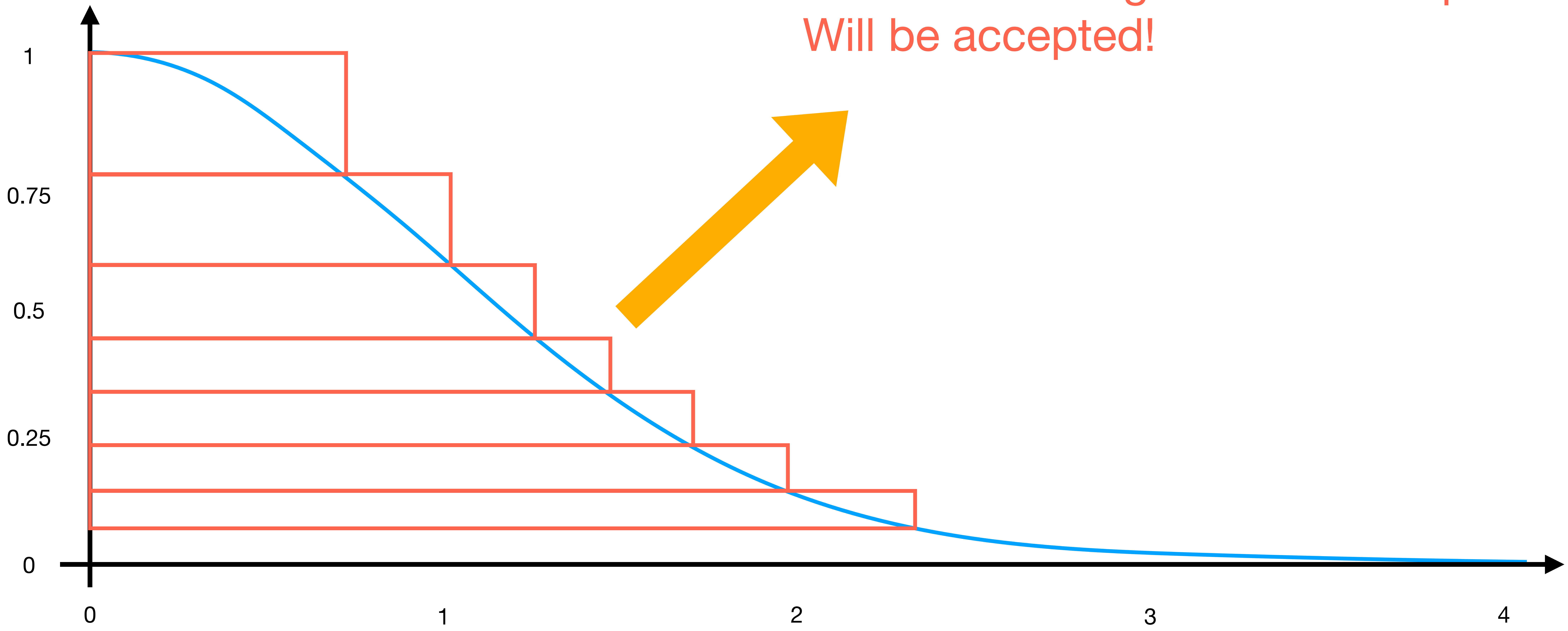
Acceptance-Rejection

The Ziggurat Algorithm



Acceptance-Rejection

The Ziggurat Algorithm



NOTE: most of generated samples Will be accepted!

Random Vectors aka Joint PDFs

Joint PDFs

Main Idea

- Typically, it can happen to have joint probabilities; e.g., sampling shapes such as disks, triangles, etc. So we end up to have:

$$p(x, y).$$

- In such cases, we firstly compute the marginal density $p(x)$ as:

$$p(x) = \int_{\mathcal{D}_x} p(x, y)dy.$$

- Then, we compute the conditional density as:

$$p(y|x) = \frac{p(x, y)}{p(x)}.$$

Joint PDFs
Main Idea

- At this point, we compute the CDF of $p(x)$ and $p(y|x)$ through integration:

$$P(x) = \int_{-\infty}^{x} p(t)dt, \text{ and}$$

$$P(y|x) = \int_{-\infty}^{y} p(t|x)dt.$$

- Finally, we draw samples by inverting these CDFs:

$$n_1 = P^{-1}(u_1) \qquad u_1 \in \mathbf{U}(0,1),$$

$$n_2 = P^{-1}(u_1|u_2) \qquad u_2 \in \mathbf{U}(0,1).$$

Joint PDFs

Main Idea

- The method, we have just seen, is called sequential inversion.
- This process can be extended to d dimension.

Joint PDFs

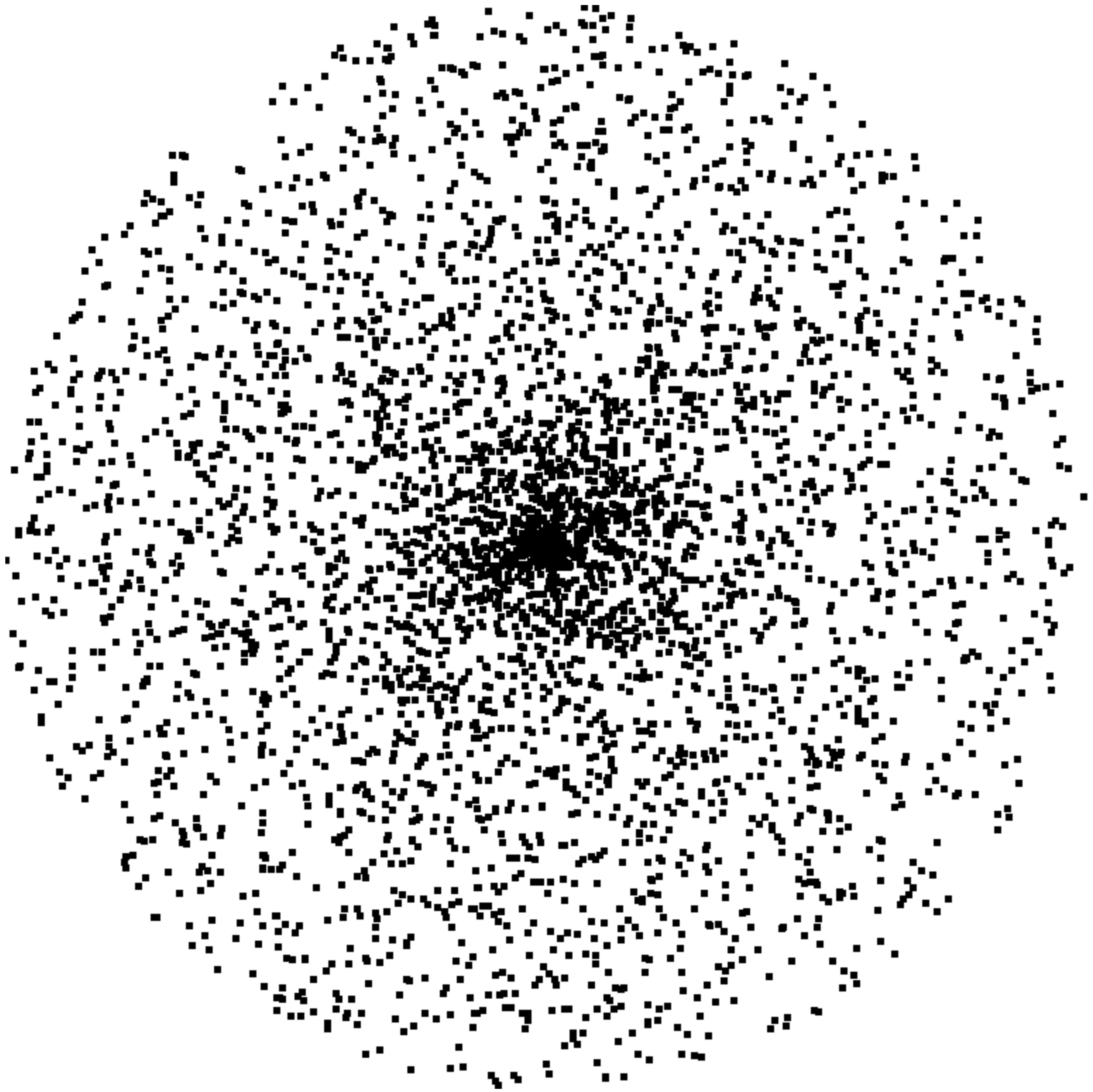
The Unit Disk

- Let's say we want to sample a unit disk in a uniform way.
- The disk looks simple, but it has hidden insidious challenges!
- The wrong approach:
  - $$r = u_1 \qquad \theta = 2\pi u_2 \qquad u_1 \in \mathbf{U}(0,1) \quad u_2 \in \mathbf{U}(0,1).$$
  - Then, we remap into XY coordinates:

$$(x, y) = [\cos(\theta) r, \sin(\theta) r].$$

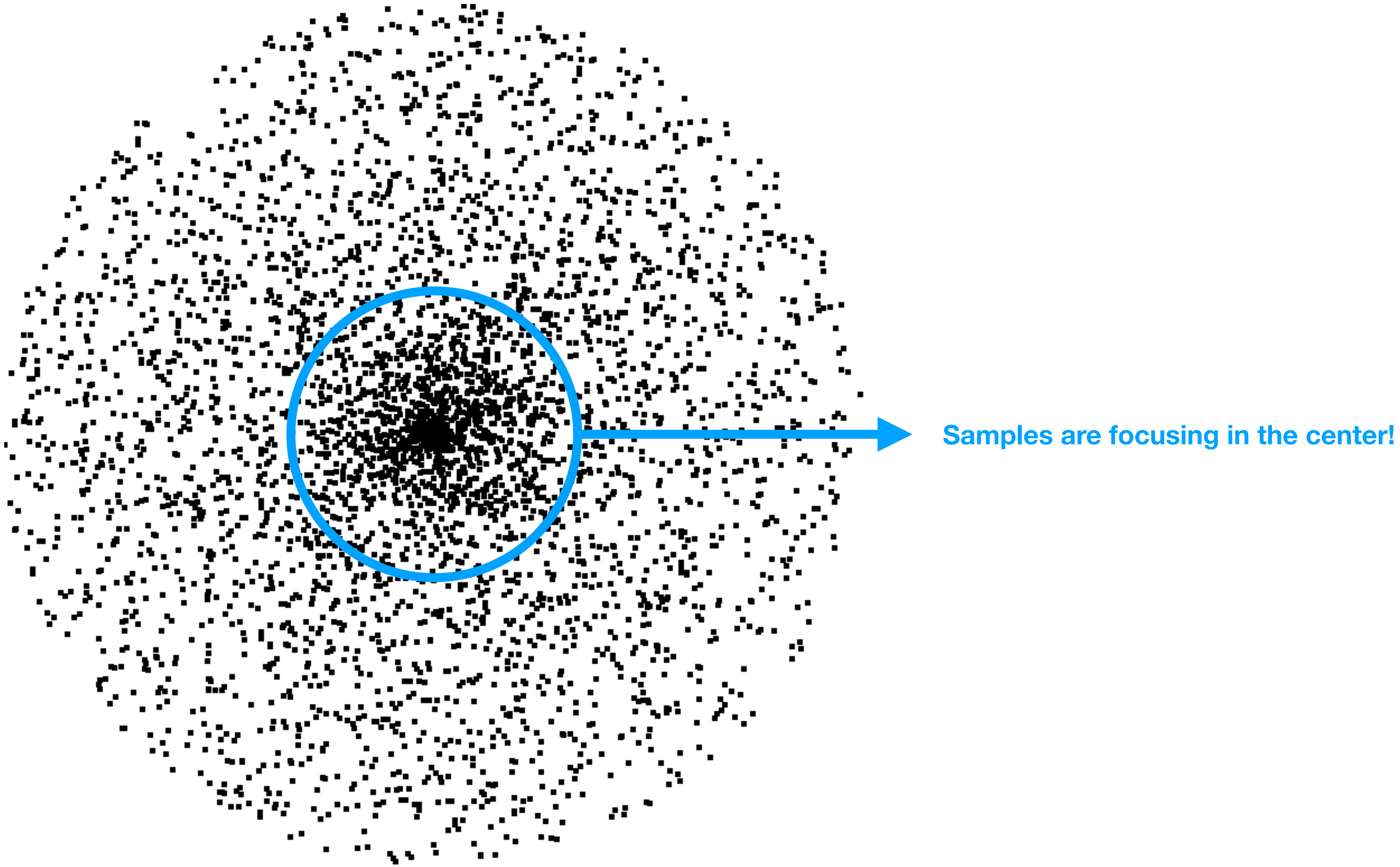
Joint PDFs

The Unit Disk



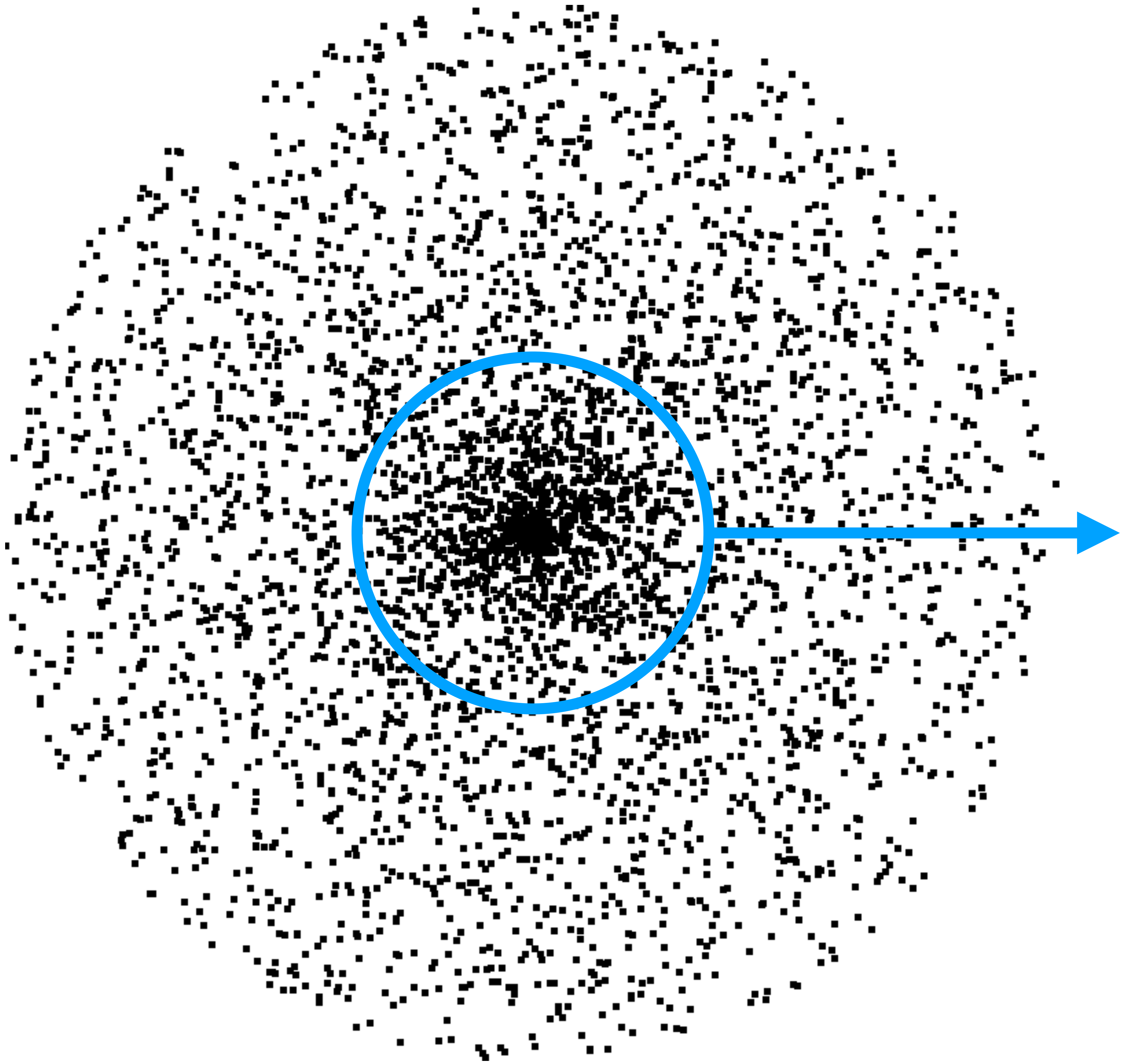
Joint PDFs

The Unit Disk



Joint PDFs

The Unit Disk



BY THE WAY, THAT'S VEERY BAD!

Joint PDFs
The Unit Disk

- The PDF, $p(x, y)$, has to be a constant!
- Assuming a unit disk, this has to be:

$$p(x, y) = \frac{1}{\pi}.$$

- Let's transform it in polar coordinates:

$$p(r, \theta) = \frac{r}{\theta}.$$

Joint PDFs

The Unit Disk

- Let's compute the marginal density:

$$p(r) = \int_0^{2\pi} p(r,\theta)d\theta = \int_0^{2\pi} \frac{r}{\pi} d\theta = \frac{r}{\pi}\int_0^{2\pi} d\theta = 2r.$$

- Now, we can compute the conditional density:

$$p(\theta\,|\,r) = \frac{p(r,\theta)}{p(r)} = \frac{\frac{r}{\pi}}{2r} = \frac{r}{\pi}\frac{1}{2r} = \frac{1}{2\pi}.$$

- We need to invert their CDFs!

Joint PDFs

The Unit Disk

- The first CDF is:

$$P(r) = \int_0^r 2x dx = r^2 \to P^{-1}(x) = \sqrt{x}.$$

- The second CDF is:

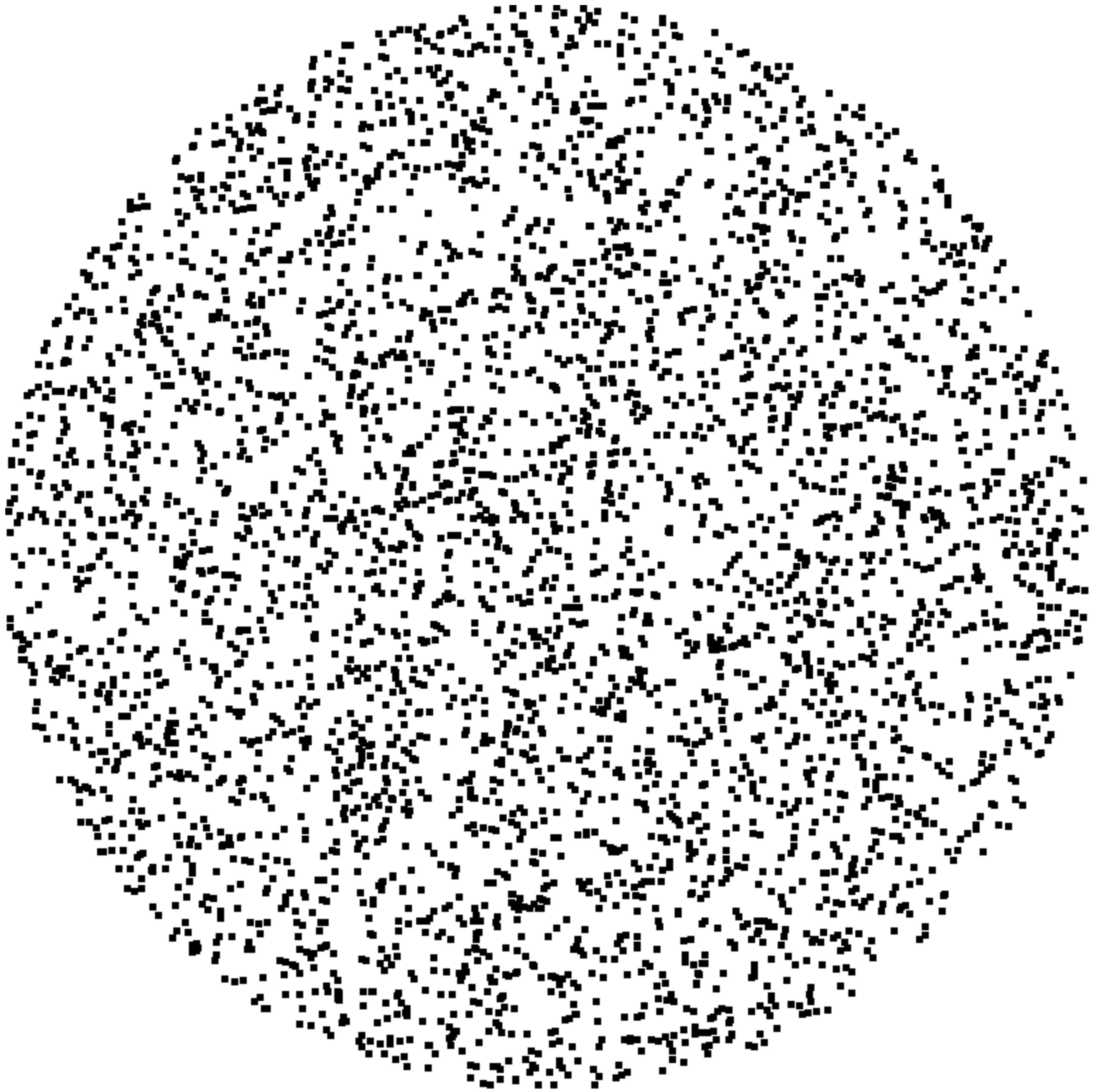
$$P(\theta \mid r) = \int_0^\theta \frac{1}{2\pi} dx \to P^{-1}(x) = 2\pi x.$$

- Now, we have all pieces to generate samples:

$$r = \sqrt{u_1} \qquad \theta = 2\pi u_2 \qquad u_1 \in \mathbf{U}(0,1) \quad u_2 \in \mathbf{U}(0,1).$$

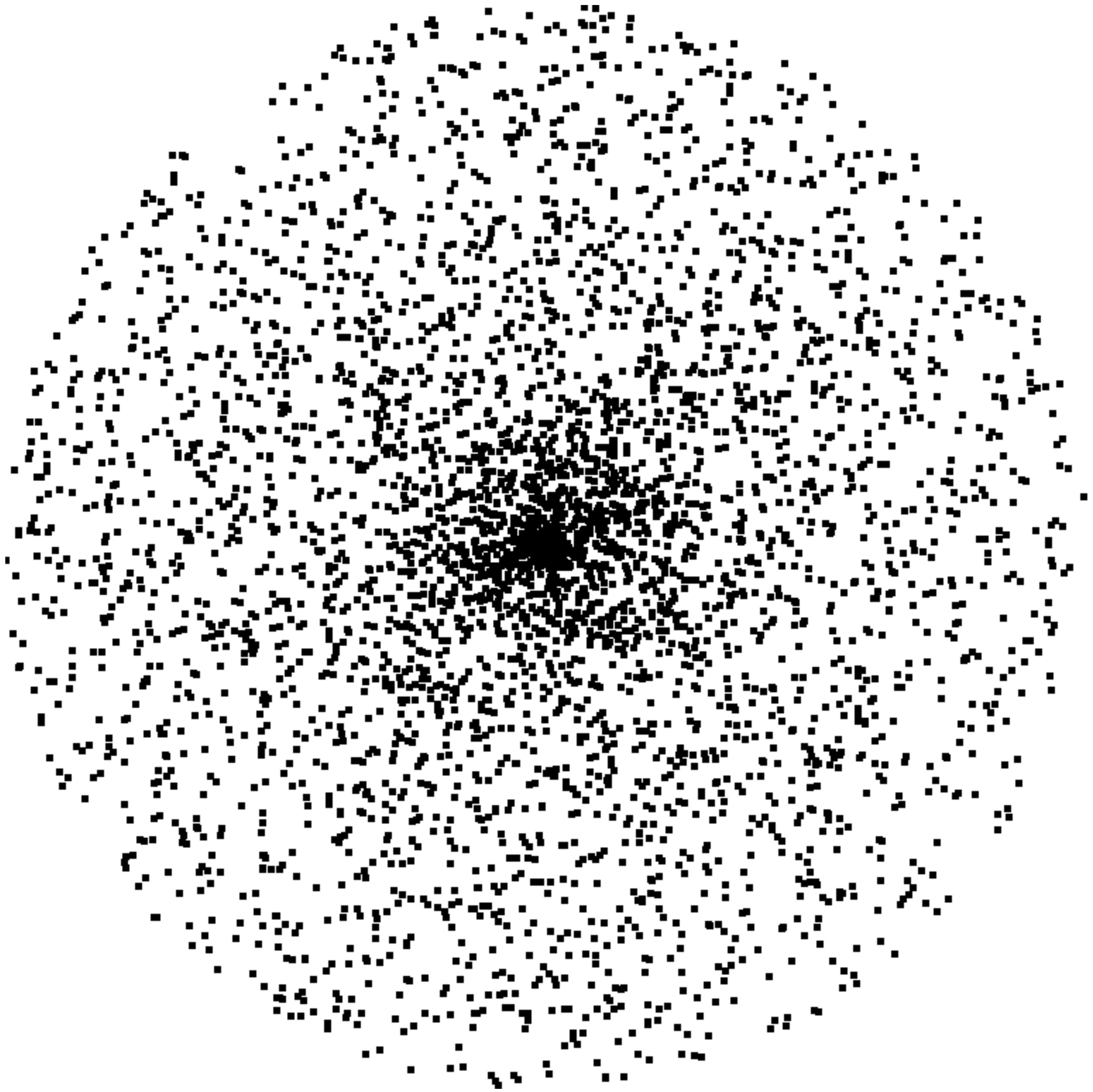
Joint PDFs

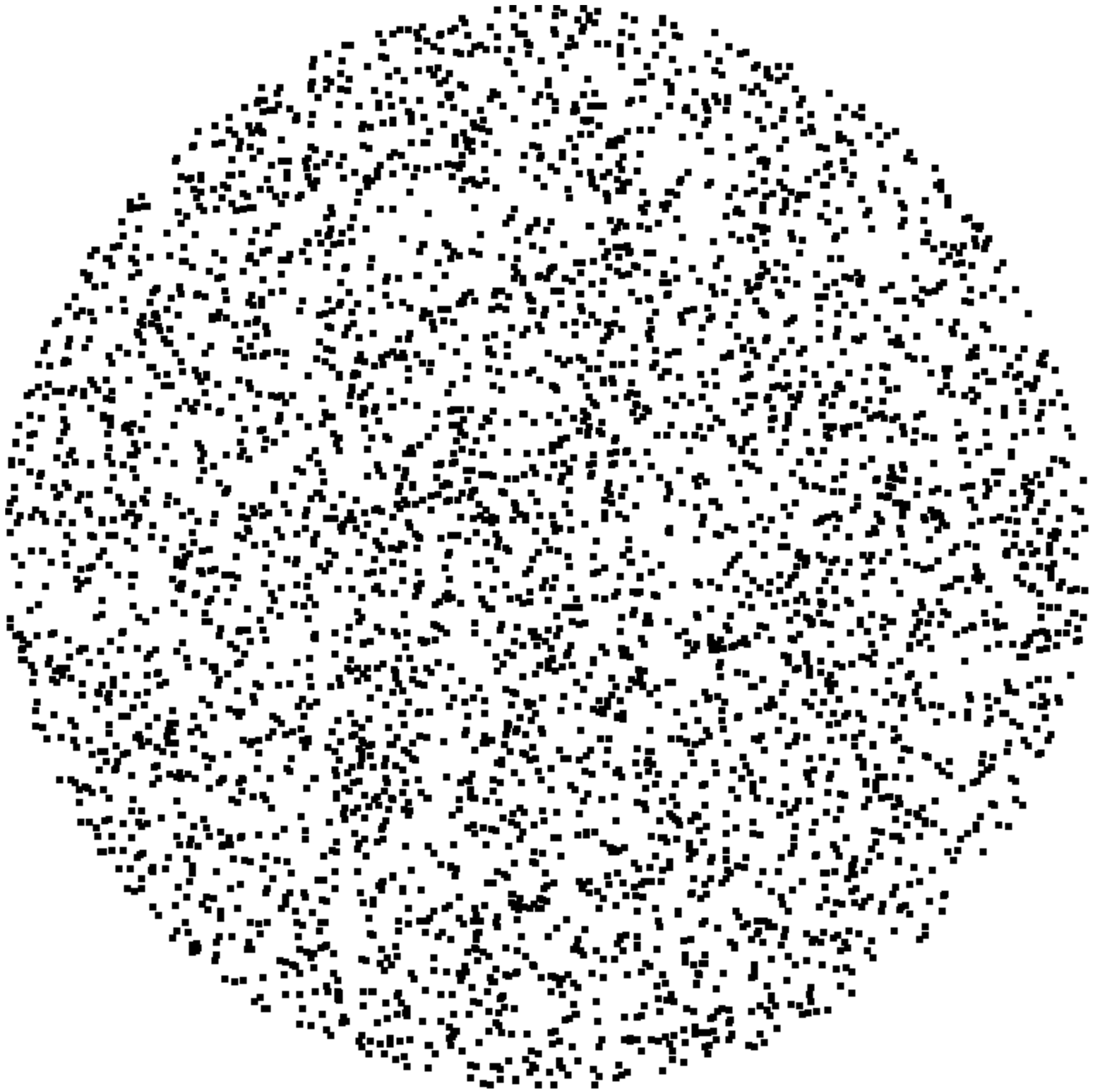
The Unit Disk



Joint PDFs

The Unit Disk





Joint PDFs

Transformations: Box Muller

- An alternative to generate normally distributed random numbers, without inverting Φ, is to use transformations:
  - Box-Muller Method:
    - Let's say, we have two independent variables, X and Y, that have normal distribution.
    - Their joint PDF is:

$$p_{XY}(x, y) = p_X(x)p_Y(y) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}} \cdot \frac{\exp(-y^2/2)}{\sqrt{2\pi}} = \frac{\exp(-(x^2+y^2)/2)}{2\pi}.$$

Joint PDFs

Transformations: Box Muller

- We convert the distribution in coordinate(x, y) in polar coordinates (r, θ) using the Jacobian matrix:

$$J = \frac{\partial(x, y)}{\partial(r, \theta)} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & r\sin(\theta) \\ \sin(\theta) & -r\cos(\theta) \end{bmatrix}.$$

- Knowing that $x^2 + y^2 = r$ and $|\det(J)| = r$, we can define the joint PDF as:

$$f(r, \theta) = \frac{1}{2\pi}\exp(-r^2/2)r \qquad \theta \in [0,2\pi] \quad r \in (0,\infty).$$

- Note that θ and R are independent variables:

$$X = R\cos(\theta) \qquad Y = R\sin(\theta).$$

Joint PDFs
Transformations: Box Muller

- We can compute the PDF of R as:

$$f_R(r) = r\exp(-r^2/2) \quad r \in (0,\infty).$$

- This leads to:
    - $X = \sqrt{-2\log U_1}\cos(2\pi U_2)$,
    - $Y = \sqrt{-2\log U_1}\sin(2\pi U_2)$,

  where $U_1, U_2 \sim \mathbf{U}(0,1)$.

Joint PDFs

Transformations: Box Muller

- We can compute the PDF of R as:

$$f_R(r) = r\exp(-r^2/2) \quad r \in (0,\infty).$$

- This leads to:
  - $X = \sqrt{-2\log U_1}\cos(2\pi U_2)$,
  - $Y = \sqrt{-2\log U_1}\sin(2\pi U_2)$,

  where $U_1, U_2 \sim \mathbf{U}(0,1)$.

Always check $U_1 \in (0,1)$, and better to add: $\sqrt{\max(-2\log U_1, 0)}$

Joint PDFs

Uniform Directions over a Hemisphere

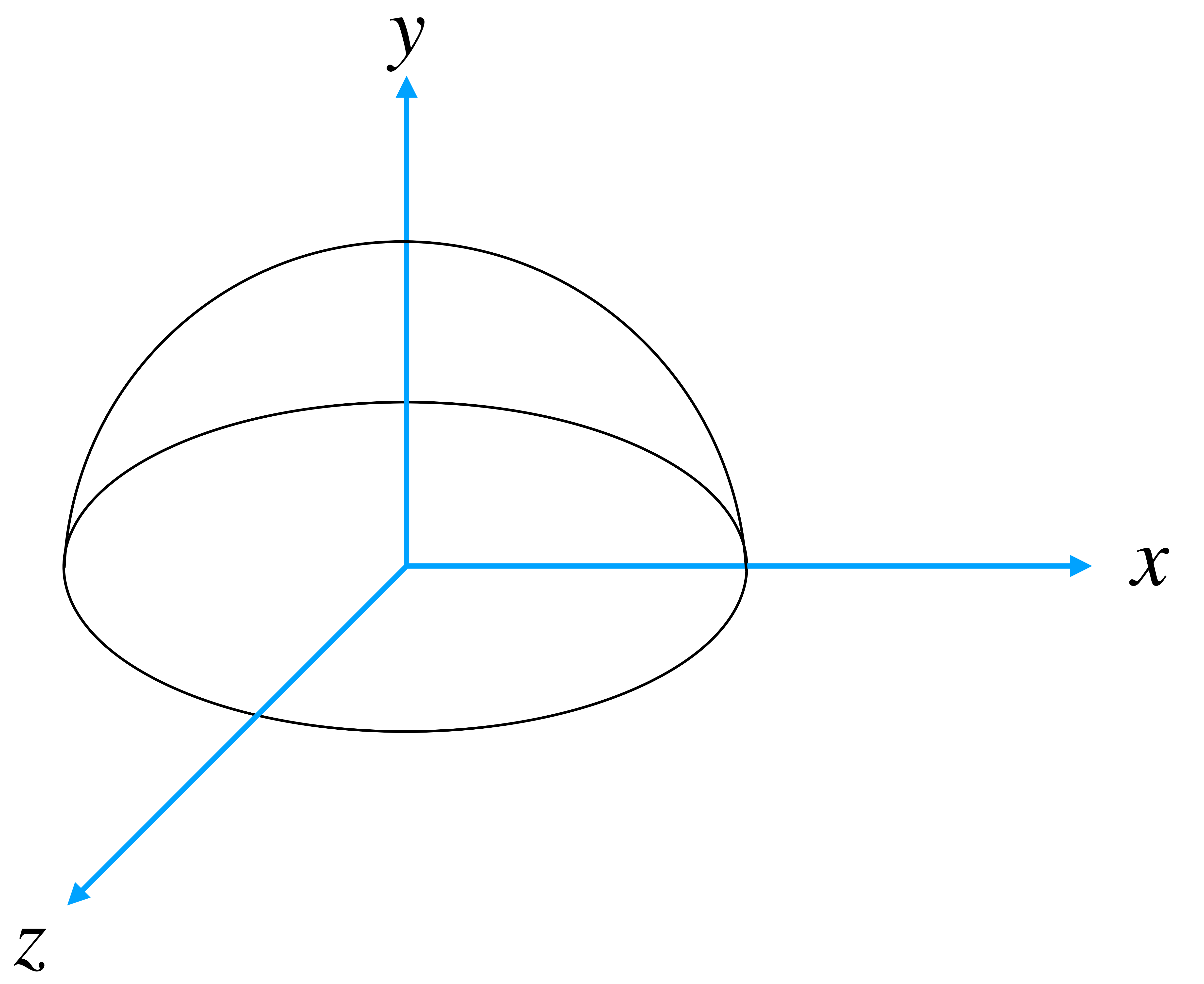
- In this case, we want to generates random vectors, directions, that are normalized; i.e., $||\overrightarrow{\omega_i}|| = 1$.
- This problem is similar to generating points on the surface of the hemisphere, $\mathbf{x}_i^s$, because we can convert them into normal directions as:

$$\overrightarrow{\omega}_i = \frac{\mathbf{x}_i^s - \mathbf{c}}{||\mathbf{x}_i^s - \mathbf{c}||}, \quad \overrightarrow{\omega}_i(\theta, \phi) = \begin{bmatrix} \cos\phi\sin\theta \\ \cos\theta \\ \sin\phi\sin\theta \end{bmatrix},$$

where $\mathbf{c}$ is the center of the hemisphere.

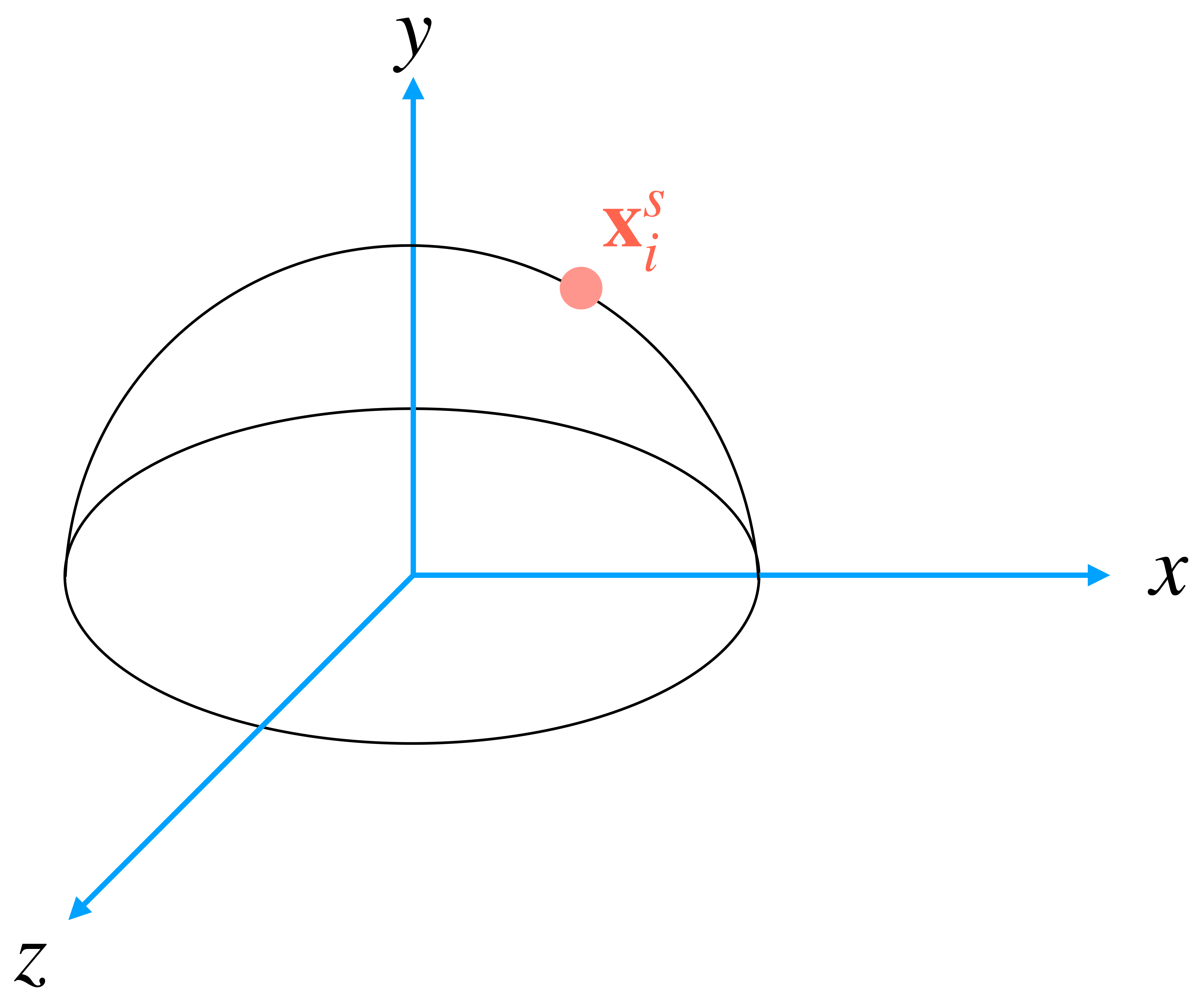
Joint PDFs

Uniform Directions over a Hemisphere



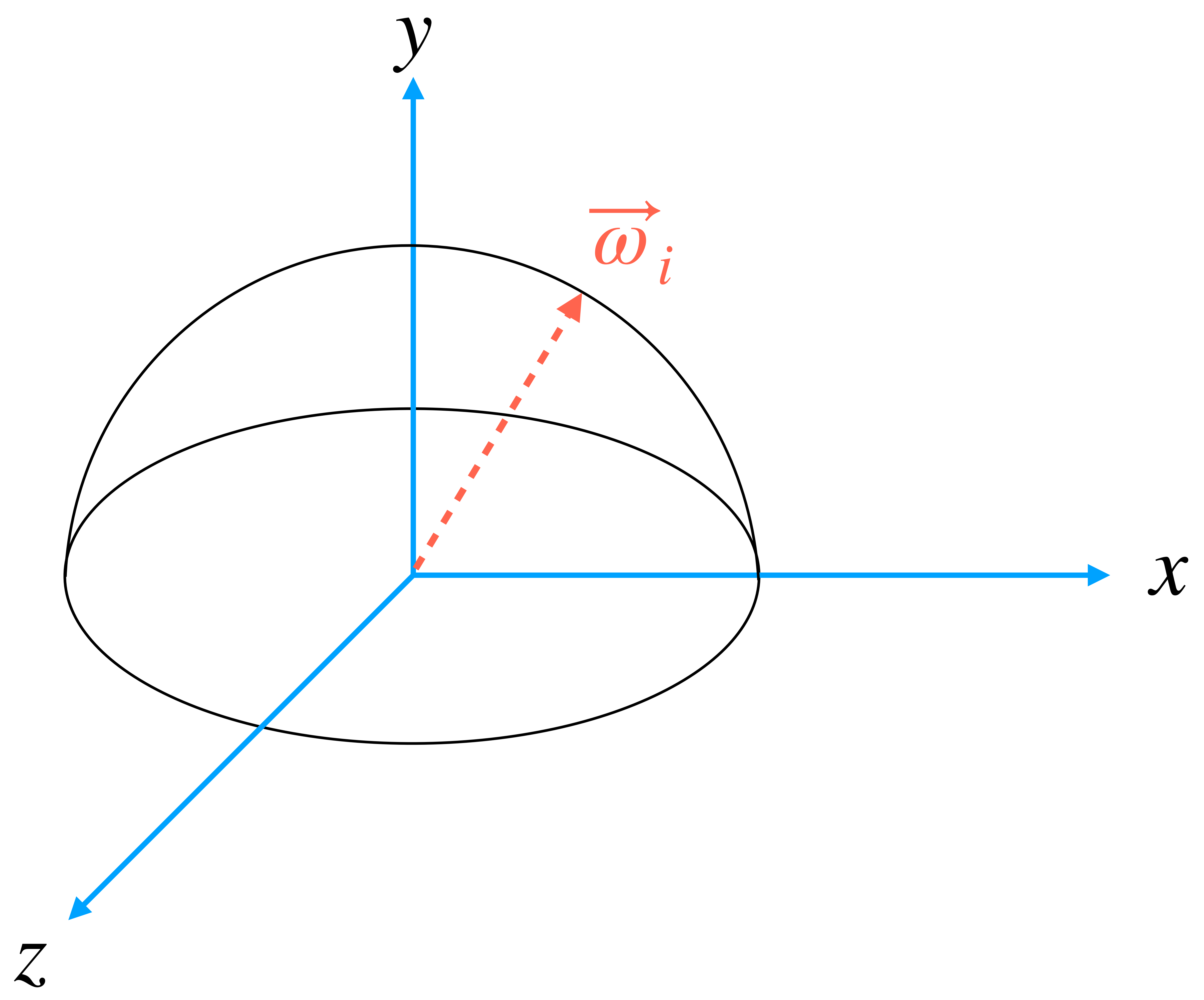
Joint PDFs

Uniform Directions over a Hemisphere



Joint PDFs

Uniform Directions over a Hemisphere



Joint PDFs

Uniform Directions over a Hemisphere

- Let's assume that the sphere has radius 1. Since it is a uniform sampling, the PDF is constant:

$$p(\overrightarrow{\omega}_i) = \frac{1}{2\pi}\text{; i.e., the inverse of the area of half sphere.}$$

- Note that:

$$\omega_x = \sin\theta\cos\phi \quad \omega_y = \cos\theta \quad \omega_x = \sin\theta\sin\phi.$$

- We need to convert from $p(\omega)$ to $p(\theta, \phi)$. Therefore, we need to compute the Jacobian for such transformation:

$$p(\omega) = p(\theta, \phi)\,|J_t| \qquad |J_t| = \sin\theta \rightarrow p(\omega) = p(\theta, \phi)\sin\theta.$$

Joint PDFs

Uniform Directions over a Hemisphere

- At this point, we compute the marginal density:

$$p(\theta) = \int_0^{2\pi} p(\theta, phi) d\phi = \int_0^{2\pi} \frac{1}{2\pi} \sin\theta = \frac{1}{2\pi} \int_0^{2\pi} \sin\theta = \sin\theta.$$

- Then, we compute the conditional density as:

$$p(\phi \,|\, \theta) = \frac{p(\theta, \phi)}{p(\theta)} = \frac{1}{2\pi}.$$

- Finally, we compute the marginal of both these densities, we invert them, and we get:

$$\theta = \cos^{-1} u_1 \qquad \phi = 2\pi u_2 \qquad u_1, u_2 \in \mathbf{U}(0,1).$$

Joint PDFs

Uniform Directions over a Hemisphere

- Practically, we do not compute θ, but we compute directly $\cos\theta$ as:
  - $\cos\theta = u_1 \qquad u_1 \in \mathbf{U}(0,1)$.
  - $\sin\theta = \sqrt{1-(\cos\theta)^2} = \sqrt{1-u_1^2}$.
  - The direction vector is given by:

$$\overrightarrow{\omega} = \begin{bmatrix} \cos\phi\sin\theta \\ \cos^{-1}\theta \\ \sin\phi\sin\theta \end{bmatrix} = \begin{bmatrix} \cos(2\pi u_2)\sqrt{1-u_1^2} \\ u_1 \\ \sin(2\pi u_2)\sqrt{1-u_1^2} \end{bmatrix}.$$

  - Note: we could generate our vector with less math by using rejection sampling, but it would take more time.

Joint PDFs

Uniform Directions over a Hemisphere

- Practically, we do not compute θ, but we compute directly $\cos\theta$ as:
  - $\cos\theta = u_1 \qquad u_1 \in \mathbf{U}(0,1)$.
  - $\sin\theta = \sqrt{1-(\cos\theta)^2} = \sqrt{1-u_1^2}$.
  - The direction vector is given by:

Always check $U_1 \in (0,1)$, and better to add: $\sqrt{\max(1-u_1^2, 0)}$

$$\vec{\omega} = \begin{bmatrix} \cos\phi\sin\theta \\ \cos^{-1}\theta \\ \sin\phi\sin\theta \end{bmatrix} = \begin{bmatrix} \cos(2\pi u_2)\sqrt{1-u_1^2} \\ u_1 \\ \sin(2\pi u_2)\sqrt{1-u_1^2} \end{bmatrix}.$$

- Note: we could generate our vector with less math by using rejection sampling, but it would take more time.

Joint PDFs

From Hemisphere To Sphere

- In this case, $\cos^{-1}\theta = 1 - 2u_1$, so with a few changes:

$$\vec{\omega} = \begin{bmatrix} \cos\phi\sin\theta \\ \cos^{-1}\theta \\ \sin\phi\sin\theta \end{bmatrix} = \begin{bmatrix} \cos(2\pi u_2)2\sqrt{u_1(1-u_1)} \\ 1 - 2u_1 \\ \sin(2\pi u_2)2\sqrt{u_1(1-u_1)} \end{bmatrix}.$$

Joint PDFs

The Multi-Dimensional Sphere

- The d-dimensional sphere is defined:

$$S = \left(\mathbf{x} \middle| \|\mathbf{x}\| = 1 \right).$$

- In order to generate uniform samples over S is to compute:

$$X = \frac{\mathbf{Y}}{\|\mathbf{Y}\|} \qquad Z \sim N(0, I_d) \ .$$

- Where the PDF is:

$$p_Y(\mathbf{y}) = \frac{1}{(2\pi)^{-\frac{d}{2}}} \exp\left(-\frac{\|\mathbf{y}\|^2}{2} \right) \ .$$

One More Thing…

One Last Thing…

Other Random Objects

- Permutations:
  - We may need to generate random permutations in uniformly.
- Matrices:
  - We may need to create random matrices following a given distribution. For example, orthogonal matrices.
- Graphs:
  - To generate a random graphs, $G = (V, E)$, is useful to have models of real-world networks; e.g., a social network.
  - The problem is basically to generate a $n \times n$ binary random matrix; i.e., the graph is defined by its adjacency matrix.

One Last Thing…

Random Objects: Permutations

- A permutation, π, of n elements is defined as:

$$\pi = \begin{pmatrix} 1, & \ldots, & n \\ \pi_1, & \ldots, & \pi_n \end{pmatrix}.$$

- A uniform random permutations can be computed as:

$$\pi = (1, \ldots, n)$$

for $i = n, \ldots, 2$ do

$$j \sim \mathbf{U}(1, i)$$

swap(π_i, π_j)

- This is uniform algorithm has probability $\dfrac{1}{n!}$.

Bibliography

- Art Owen. “Chapter 4: Non-Uniform Random Numbers” from the book “Monte Carlo theory, methods and examples”. 2013.
- Art Owen. “Chapter 5: Random vectors and objects” from the book “Monte Carlo theory, methods and examples”. 2013.
- Matt Pharr, and Greg Humphreys. Chapter 13: “Monte Carlo Integration” from “Physically Based Rendering: From Theory To Implementation” Second Edition. Morgan Kaufmann. 2010.

Thank you for your attention!