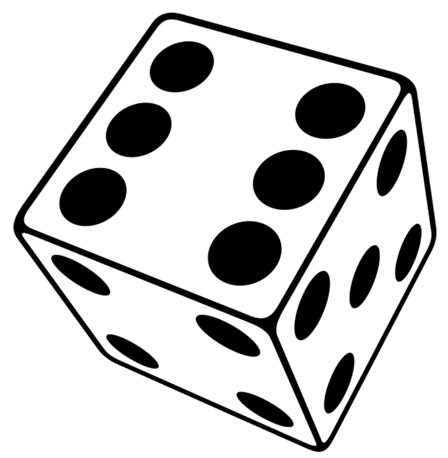


Probability:
Theory and practice

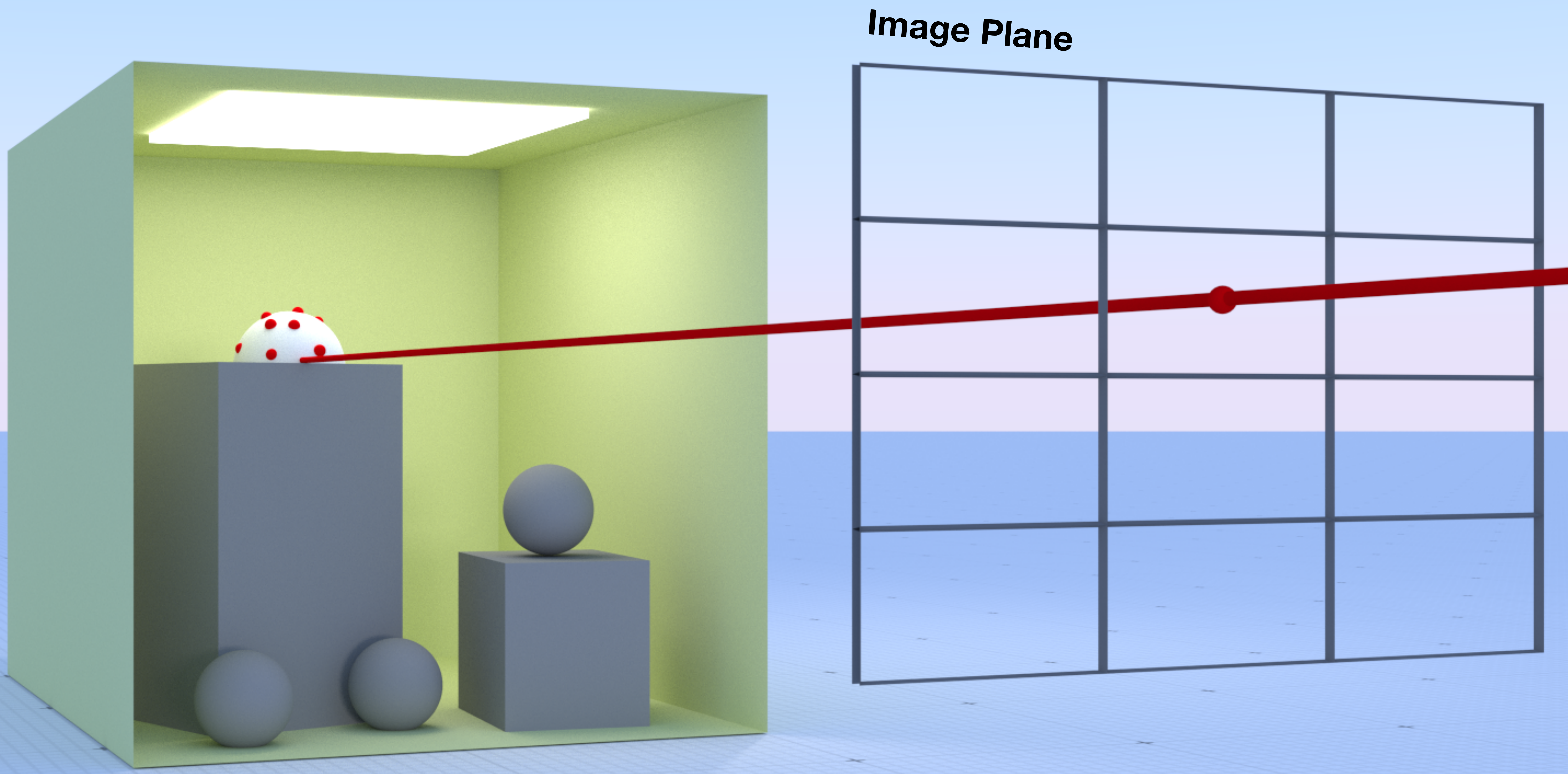
Philipp Slusallek   *Karol Myszkowski*
Gurprit Singh

A la Carte

- σ- algebra and measure
- Random Variables
- Probability distribution functions (PDFs and PMFs)
- Conditional and Marginal PDFs
- Expected value and Variance of a random variable

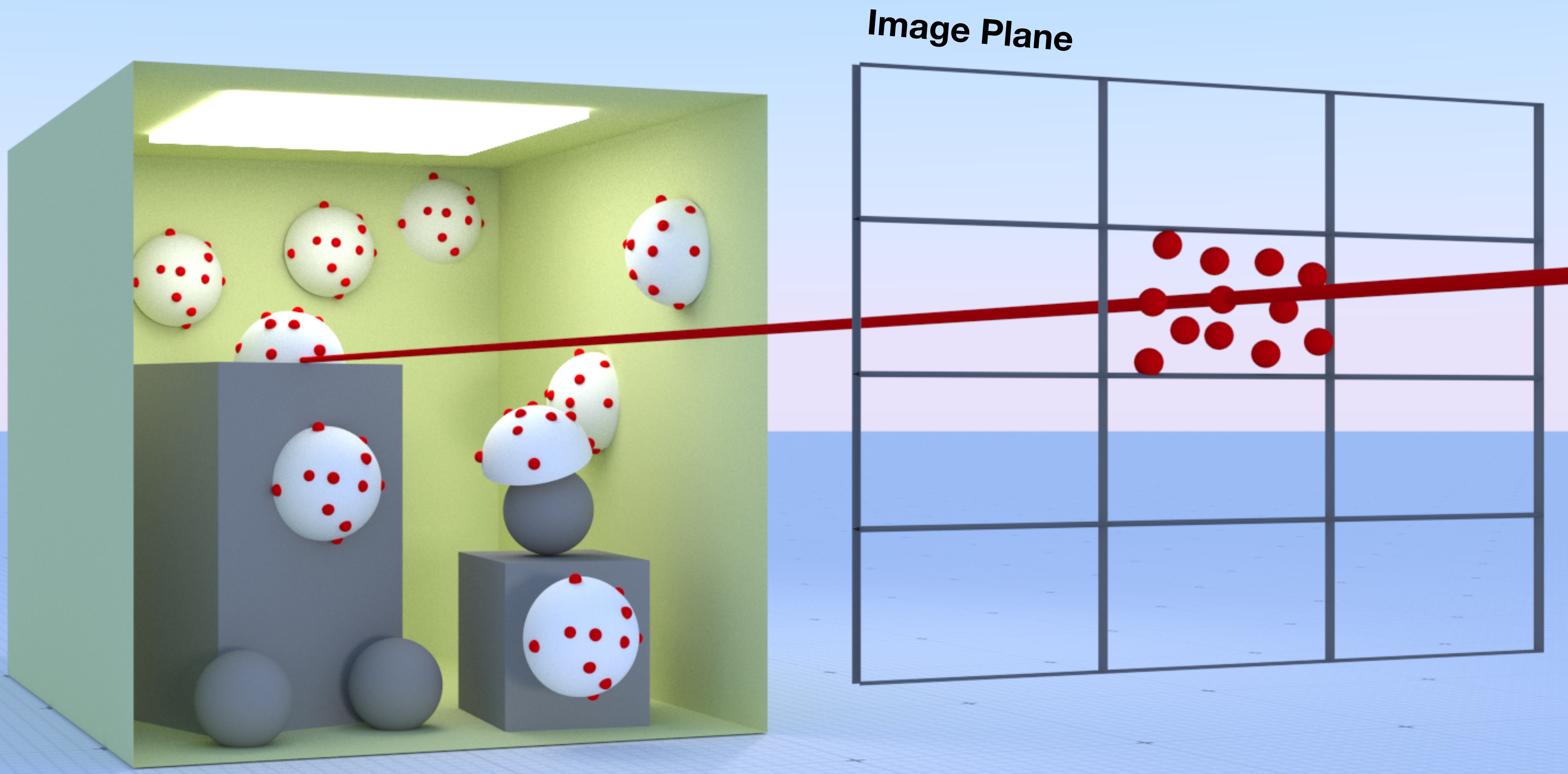
Motivation: Ray Tracing

Ray Tracing



Scene designed by David Coeurjolly

Ray Tracing



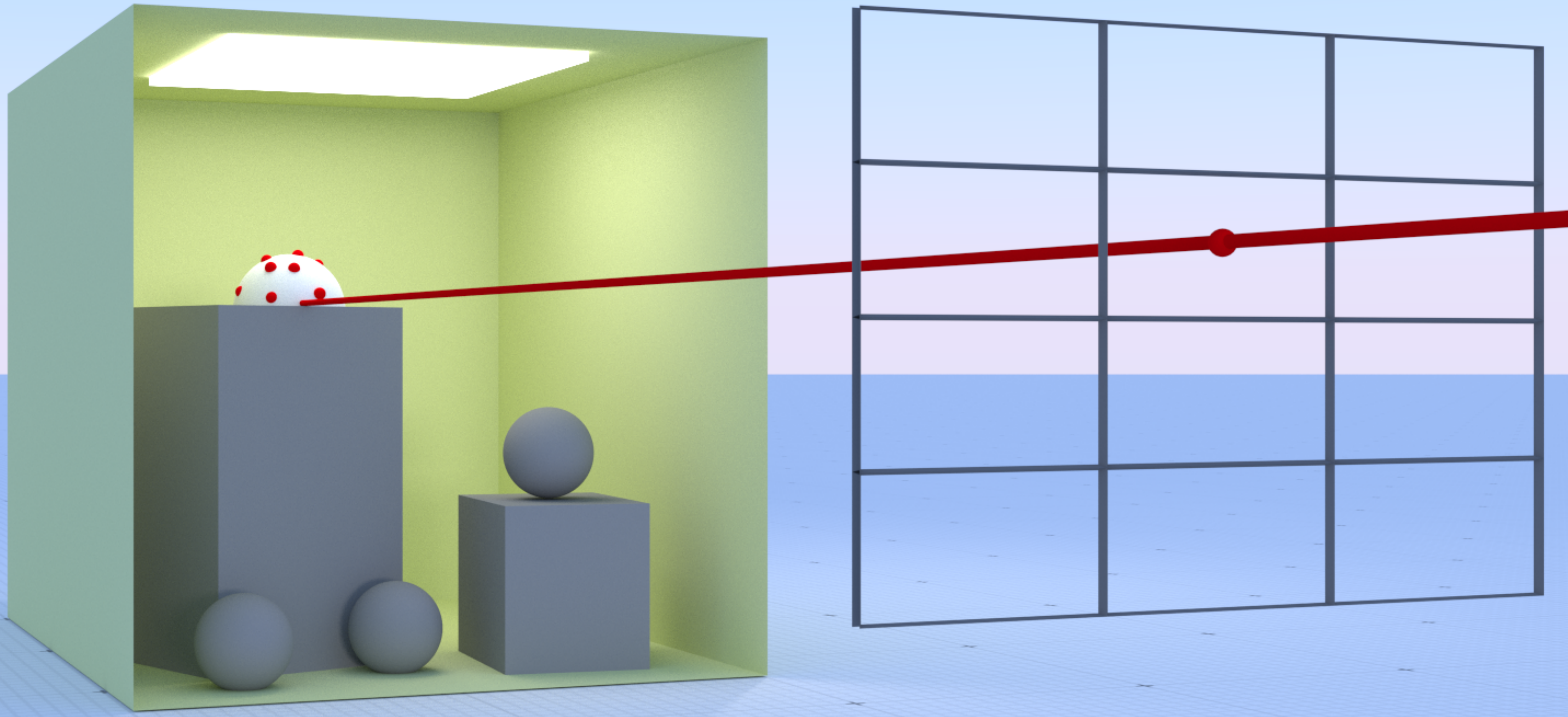
Scene designed by David Coeurjolly

Direct Illumination

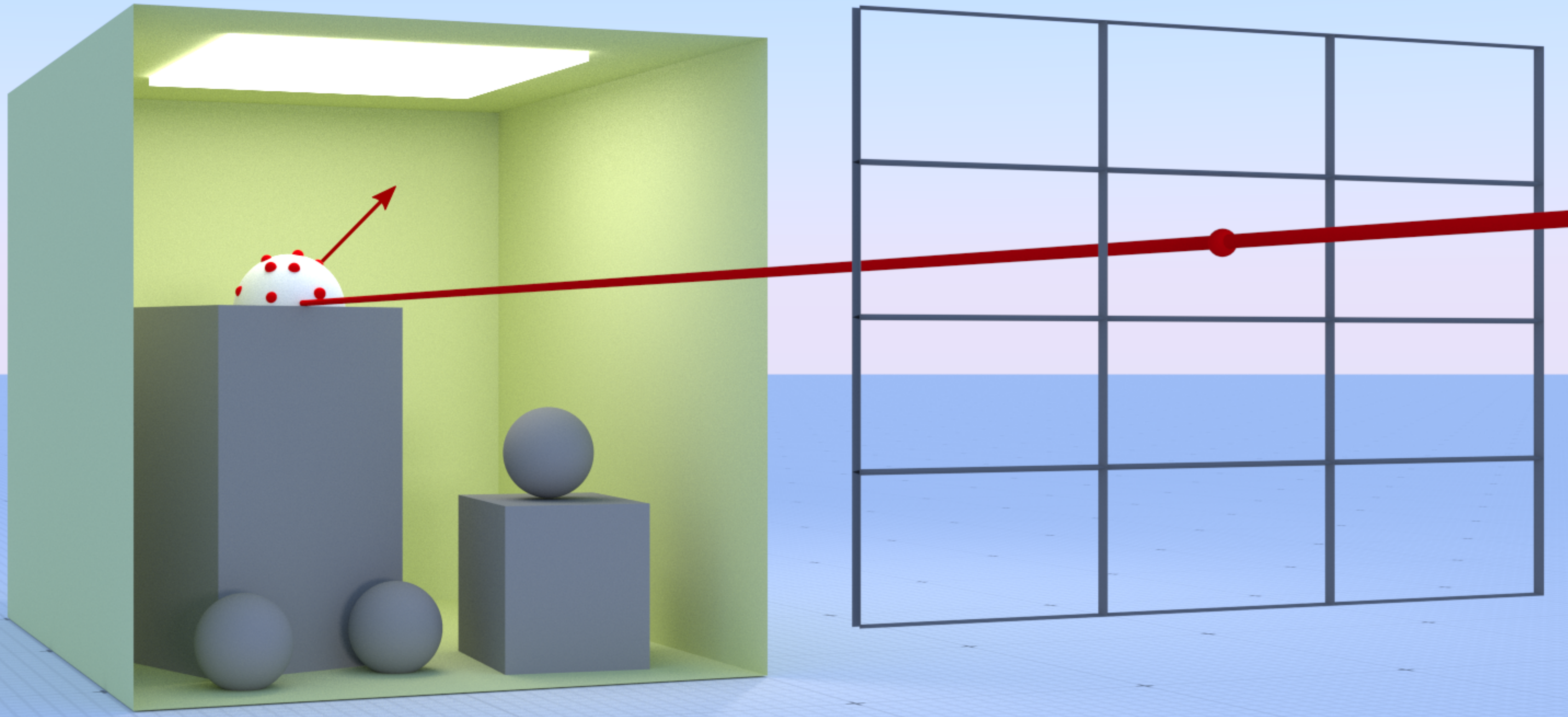
Direct Illumination

Direct and Indirect Illumination

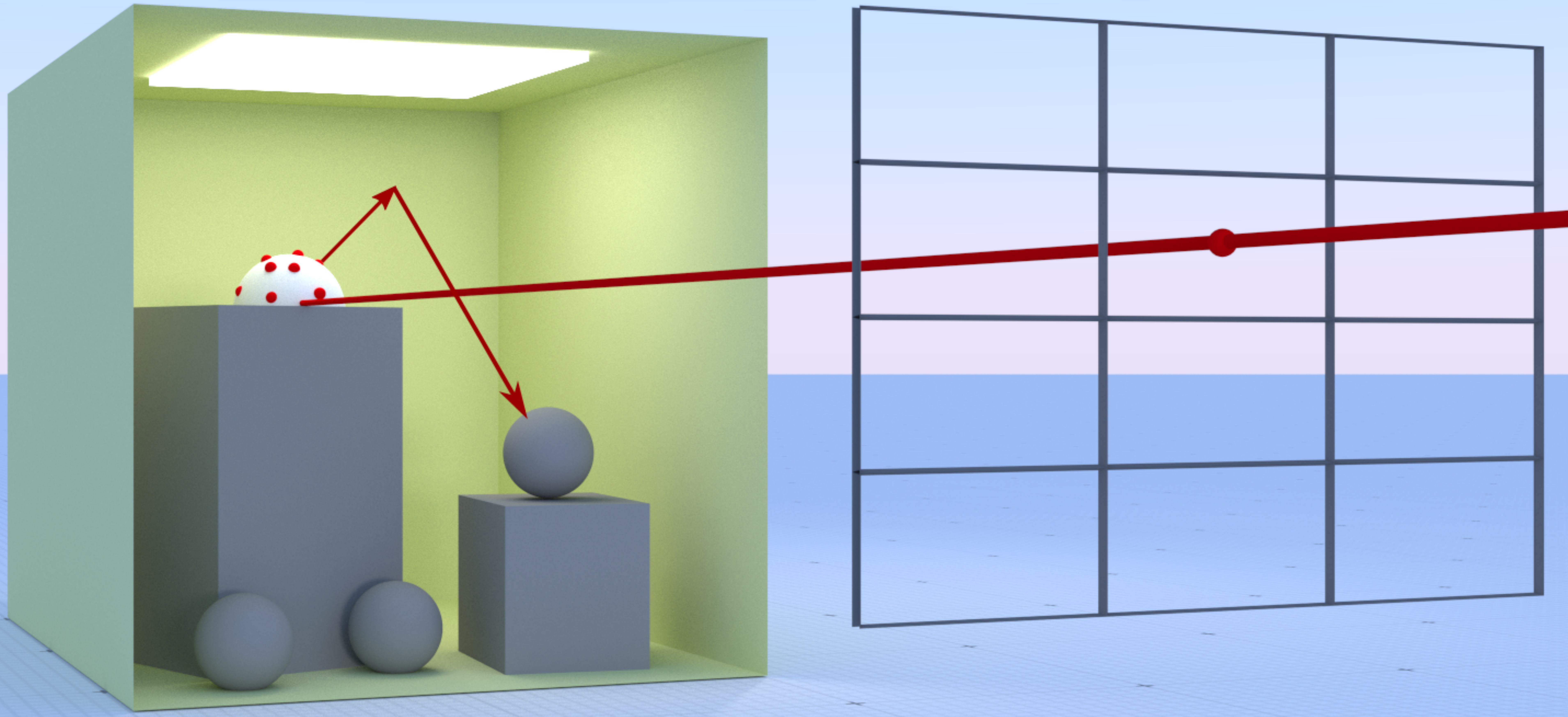
Path Tracing



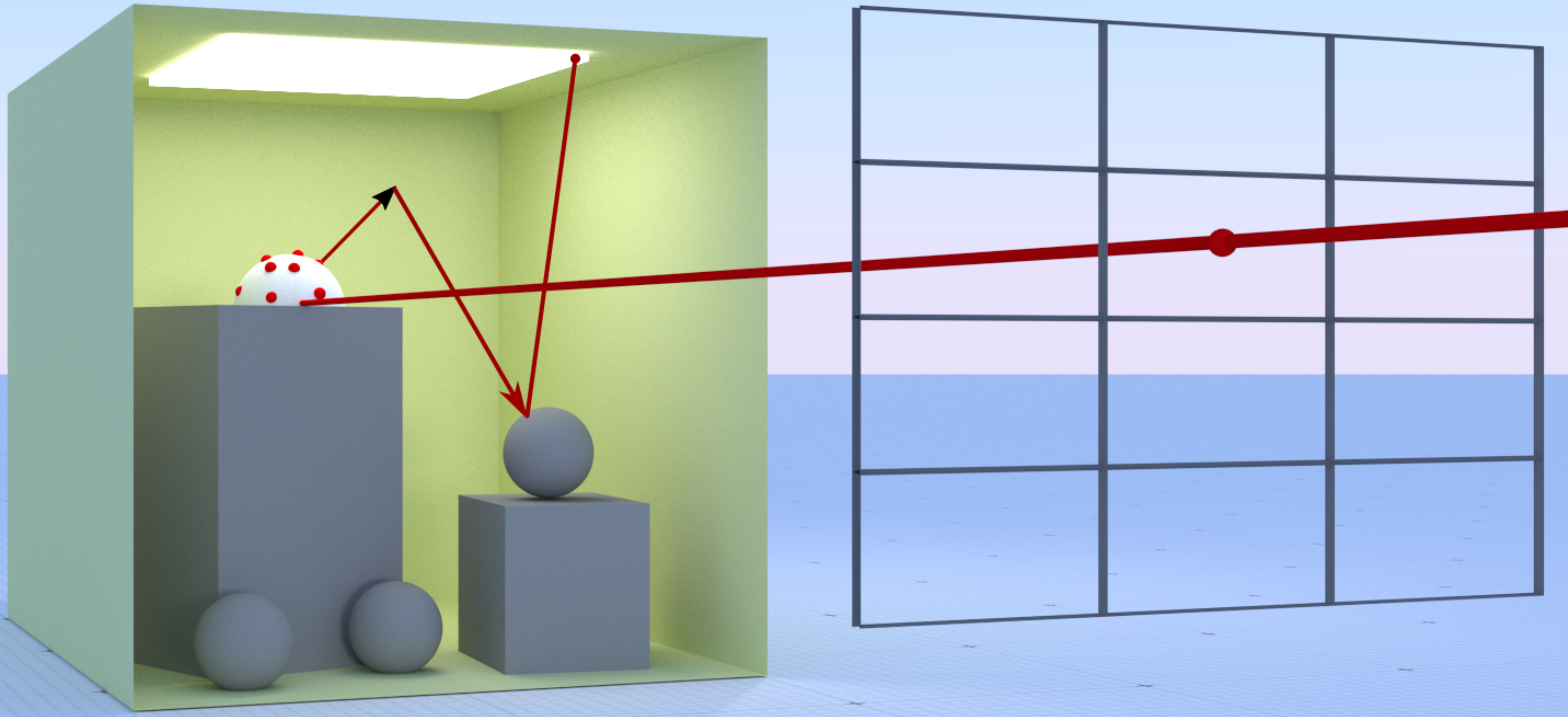
Path Tracing



Path Tracing



Path Tracing



Direct and Indirect Illumination

How can we analyze the noise present in the images ?

Probability Theory
and/or
Number Theory

Probability Theory

- Discrete Probability Space
- Continuous Probability Space

Rolling a fair dice

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- Finite outcomes: **discrete** random experiment
- Can ask the outcome is a number: 1 or 6
- Can ask the outcome is a subset, e.g. all prime numbers:

Rolling a fair dice

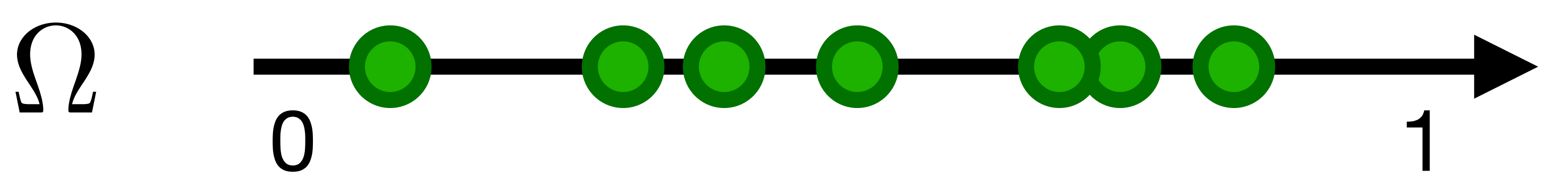
$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- **R1**: Apart from elementary values, the focus lies on subsets of Ω
- **R2**: A probability assigns each element or each subset of a positive real value

The first requirement leads to the concept of σ-algebra

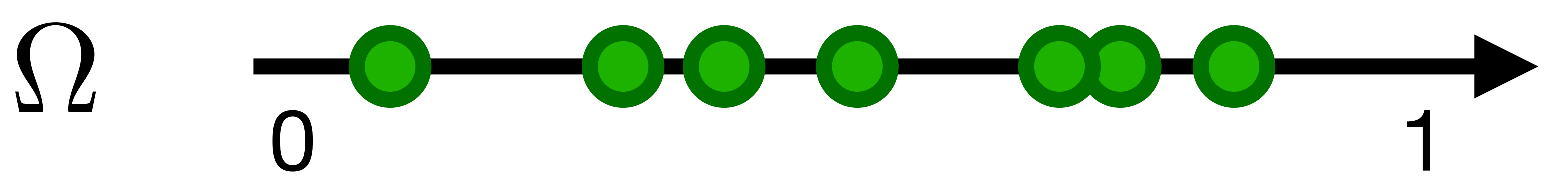
The second to the mathematical construct of a measure

Random number in [0,1]



- Uncountably infinite outcomes: **continuous** random experiment
- Does not make sense to ask for one number as output, e.g. 0.245
- We need to ask for the probability of a region, e.g. [0.2,0.4] or [0.36,0.89]

Random number in [0,1]



- **R1**: As in discrete case, focus lies on subsets of Ω, also called events
- **R2**: A probability assigns each subset of a positive real value.

The first requirement leads to the concept of Borel σ-algebra

The second to the mathematical construct of a Lebesgue measure

σ-Algebra

- Mathematical construct used in probability and measure theory

  1. Take on the role of system of events in probability theory

- Simply spoken: Collection of subsets of a given set

  A. A non-empty collection of subsets c hat is **closed** under the set theoretical operations of: countable unions, countable intersections, and complement

σ-Algebra

- For discrete set Ω :

    1. The sigma-algebra corresponds to the power set of omega (set of all subsets)

σ-Algebra

- For discrete set Ω :

1. The sigma-algebra corresponds to the power set of omega (set of all subsets)

$$\Omega = \{0, 1\}$$

$$\Sigma = \{\{\phi\}, \{0\}, \{1\}, \{0, 1\}\}$$

σ-Algebra

- For discrete set Ω :

1. The sigma-algebra corresponds to the power set of omega (set of all subsets)

$$\Omega = \{0, 1\}$$

$$\Sigma = \{\{\phi\}, \{0\}, \{1\}, \{0, 1\}\}$$

$$\Omega = \{a, b, c, d\}$$

$$\Sigma = \{\{\phi\}, \{a, b\}, \{c, d\}, \{a, b, c, d\}\}$$

σ-Algebra

- For continuous set Ω:

A. The associated sigma algebras are the Borel sets ove i.e., the collection of all open sets over omega that can be generated via countable unions, countable intersections, and complement of open sets

σ-Algebra

- For continuous set Ω:

A. The associated sigma algebras are the Borel sets over Ω, i.e., the collection of all open sets over omega that can be generated via countable unions, countable intersections, and complement of open sets

$I = [p, q), p, q \in \mathbb{R}$ Fixed half-interval

σ-Algebra

- For continuous set Ω:

A. The associated sigma algebras are the Borel sets over Ω, i.e., the collection of all open sets over omega that can be generated via countable unions, countable intersections, and complement of open sets

$I = [p, q), p, q \in \mathbb{R}$ Fixed half-interval

$\mathbb{T} = [\alpha, \beta) \subseteq [p, q)$ Collection of all half-intervals

σ-Algebra

- For continuous set Ω:

A. The associated sigma algebras are the Borel sets over Ω, i.e., the collection of all open sets over omega that can be generated via countable unions, countable intersections, and complement of open sets

$I = [p, q), p, q \in \mathbb{R}$ Fixed half-interval

$\mathbb{T} = [\alpha, \beta) \subseteq [p, q)$ Collection of all half-intervals

Here, $\mathbb{T}$ is not a σ-algebra because, generally speaking, neither the union nor the difference of two half-intervals is a half-interval.

σ-Algebra

It is the mathematical construct that allows defining a measure

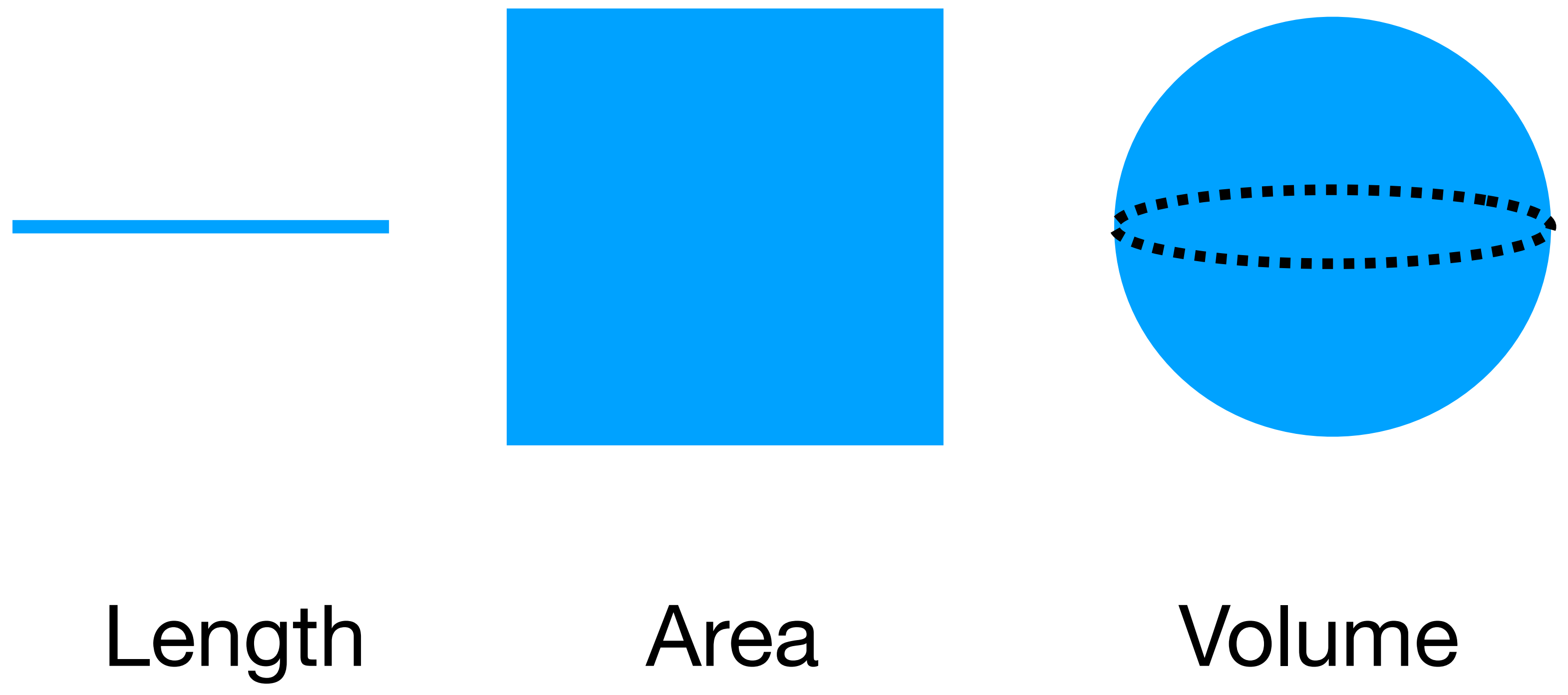
Measure

- In probability theory, it plays the role of a probability distribution
- A real-valued set function defined on a sigma-algebra that assigns each subset of a sigma-algebra a non-negative real number.
- A sigma-additive set function: i.e., the measure of the union of disjoint sets is equal to the sum of the measures of the individual sets

Lebesgue Measure

- Standard way of assigning measure to subsets of n-dimensional Euclidean space.

- For n = 1,2 or 3, it coincides with the standard measure of length, area or volume, respectively.



Random Variable

- Central concept in probability theory

- Enables to construct a simpler probability space from a rather complex one

- Correspond to a measurable function defined on a gebra that assigns each element to a real number

Random Variable

- A random variable X is a value chosen by some random process
- Random variables are always drawn from a domain: discrete (e.g., a fixed set of probabilities) or continuous (e.g., real numbers)
- Applying a func random var ults in a new random variable

Discrete Probability Space

Discrete Random Variable

- Random variable (RV):

$$X : \Omega \rightarrow E \qquad \Omega = \{x_1, x_2, \ldots, x_n\}$$

- Probabilities:

$$\{p_1, p_2, \ldots, p_n\}$$

$$\sum_{i=1}^{N} p_i = 1$$

Discrete Random Variable

- Example: Rolling a Die

$$x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5, x_6 = 6$$

- Probability of each event:

$p_i = 1/6 \quad \text{for} \; i = 1, \ldots, 6$ $\qquad P(X = i) = \frac{1}{6}$

UNIVERSITÄT DES SAARLANDES

Discrete Random Variable

$$P(2 \leq X \leq 4) = \sum_{i=2}^{4} P(X = i)$$

$$= \sum_{i=2}^{4} \frac{1}{6} = \frac{1}{2}$$

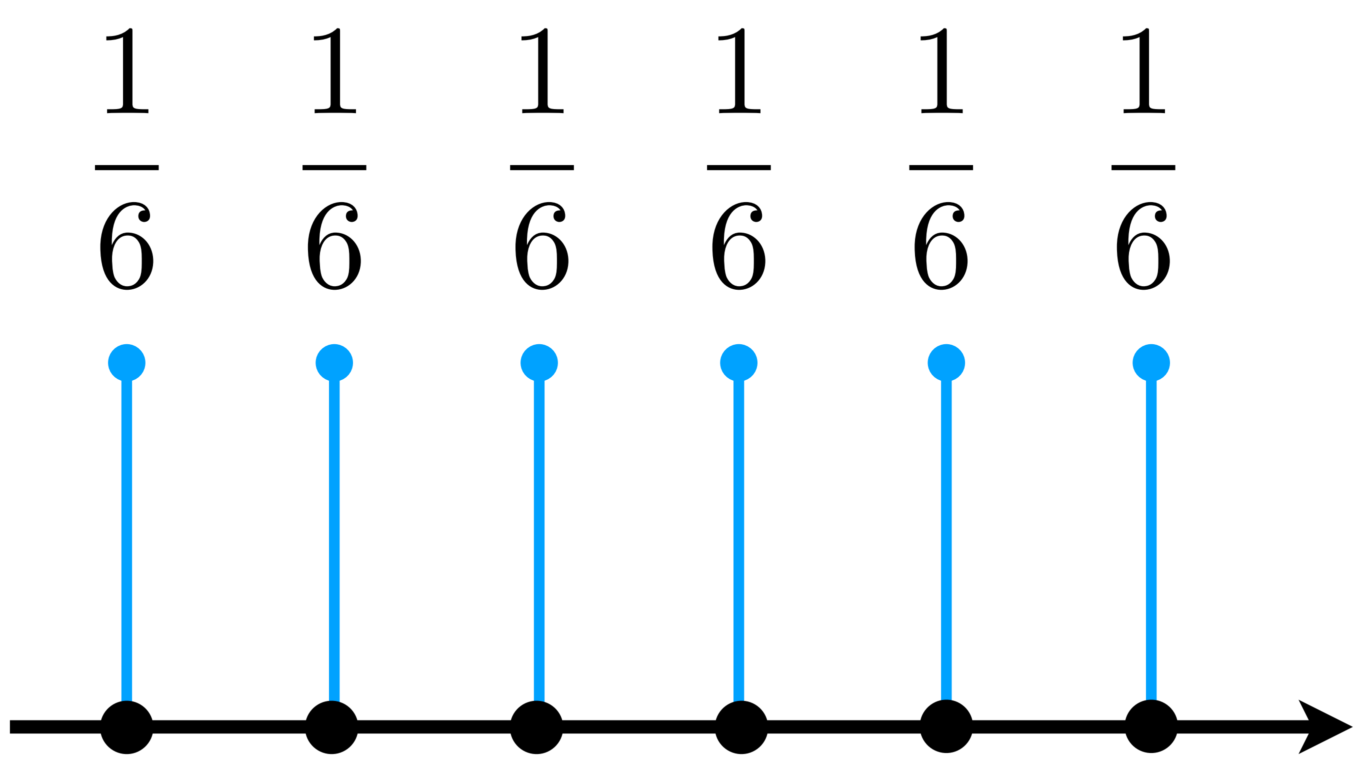
Probability mass function

- PMF is a function that gives the probability that a discrete RV is exactly equal to some value.

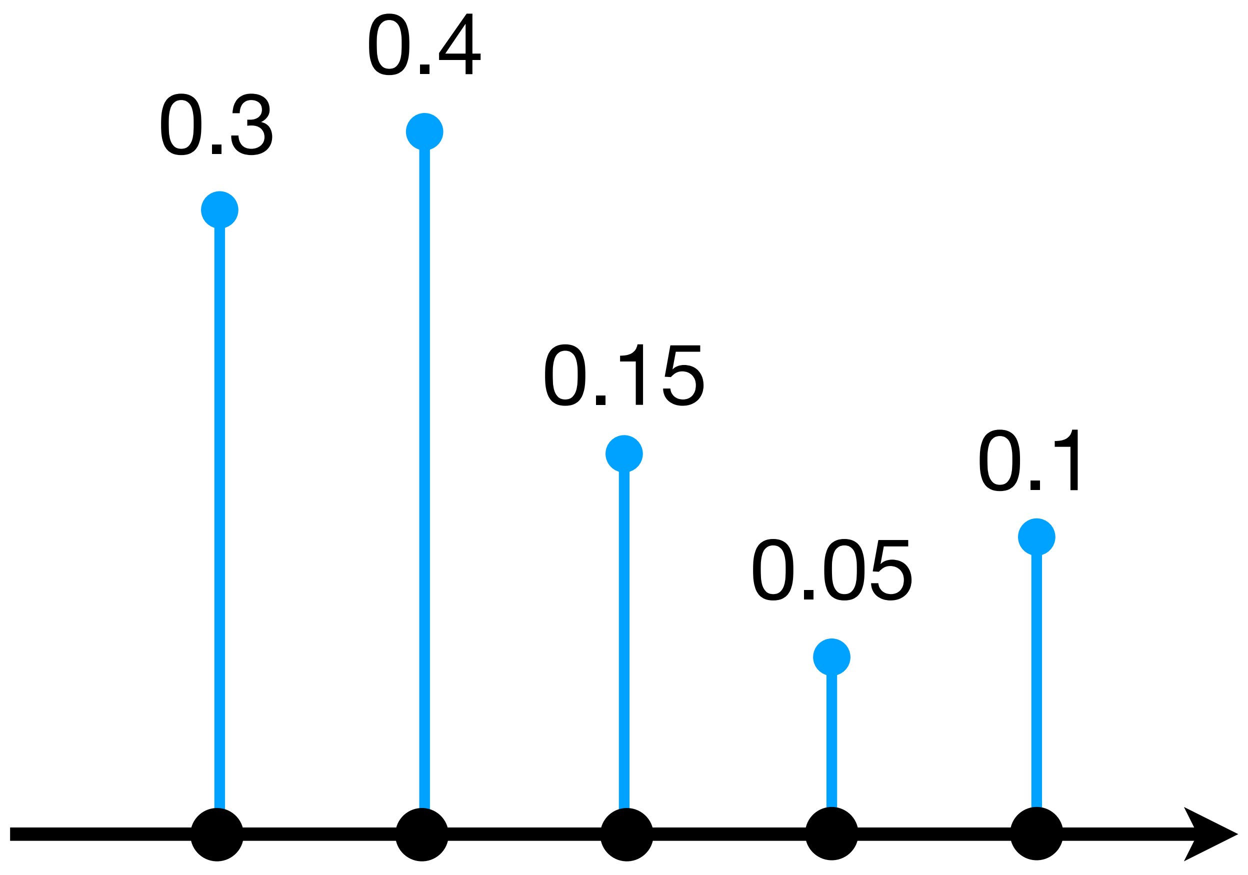
- PMF is different from PDF (probability density function) which is for continuous RVs.

Probability mass function

Constant PMF



Non-uniform PMF

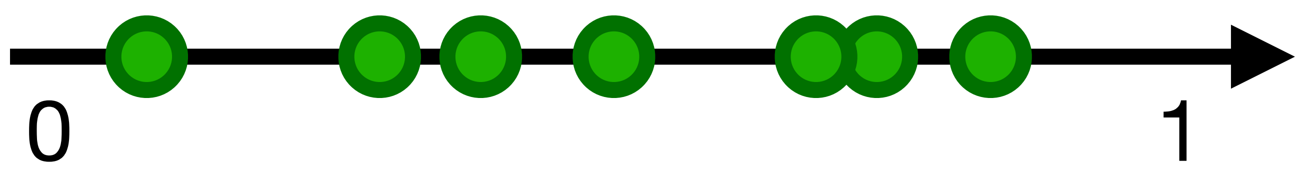


Continuous Probability Space

Continuous Random Variable

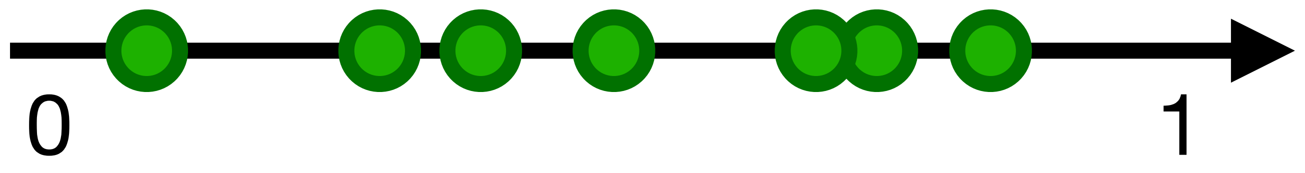
- In rendering, discrete random variables are less common than continuous random variables
- Continuous random variables take on values that ranges of continuous domains (e.g. real numbers or directions on the unit sphere)
- A particularly important random variable is the *canonical uniform random variable, which we write a*

Continuous Random Variable



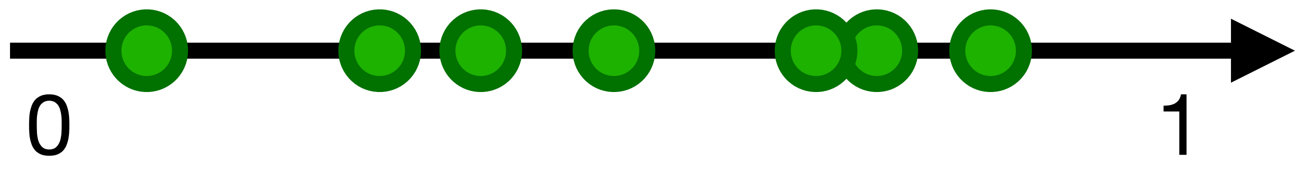
$$\xi \in [0, 1)$$

Continuous Random Variable



- We can take a continuous, uniformly distributed random variable $\xi \in [0, 1)$ and map to a discrete random variable, choosing X_i if:

Continuous Random Variable



- We can take a continuous, uniformly distributed random variable $\xi \in [0, 1)$ and map to a discrete random variable, choosing X_i if:

$$\sum_{j=1}^{i-1} p_j < \xi \leq \sum_{j=1}^{i} p_j$$

$$X_i = \{1, 2, 3, 4, 5, 6\}$$

Visual Break

Image rendered using PBRT

Visual Break

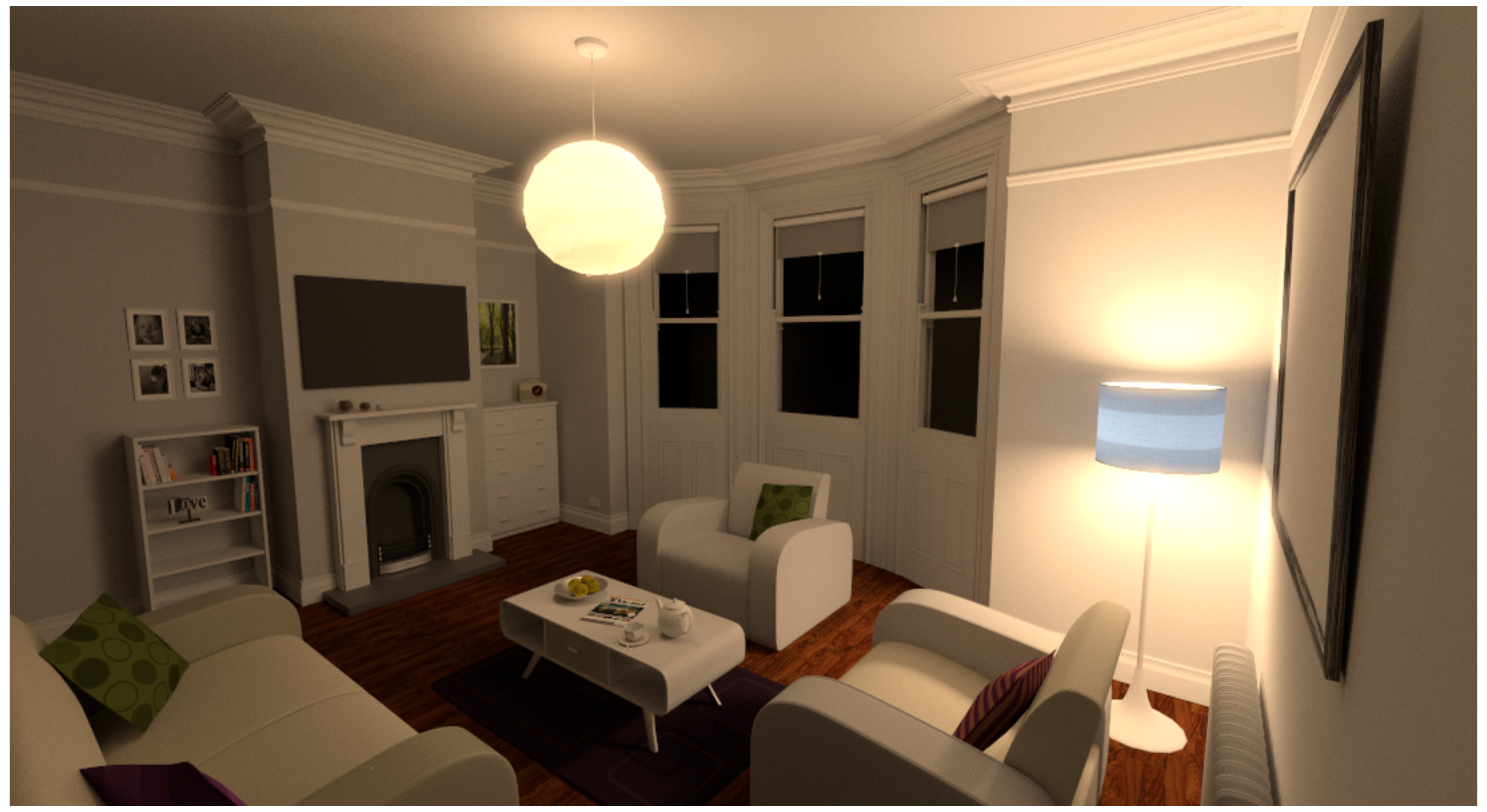
Image rendered using PBRT

Continuous Random Variable

- For lighting application, we might want to define probability of sampling illumination from each light source in the scene based on its power Φ_i

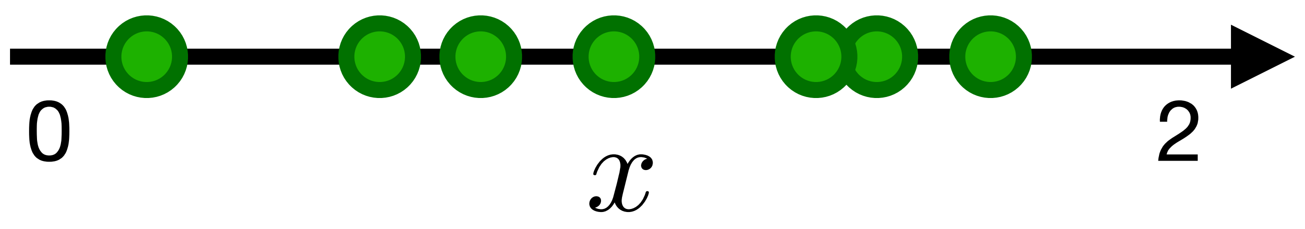
$$p_i = \frac{\Phi_i}{\sum_j \Phi_j}$$

Here, the probability is relative to the total power



Probability Density Functions

Probability density function



- Consider a continuous RV that ranges over real numbers: $[0, 2)$, where the probability of taking on any particular value x is **proportional** to the value $2 - x$

- It is twice as likely for this random variable to take on a value around 0 as it is to take around 1, and so forth.

Probability density function

- The probability density function (PDF) formalizes this idea: it describes the relative probability of a RV taking on a particular value.
- Unlike PMF, the values of the PDFs are not the probabilities as such: a PDF must be integrated over an interval to yield a probability

Probability density function

For uniform random variables:

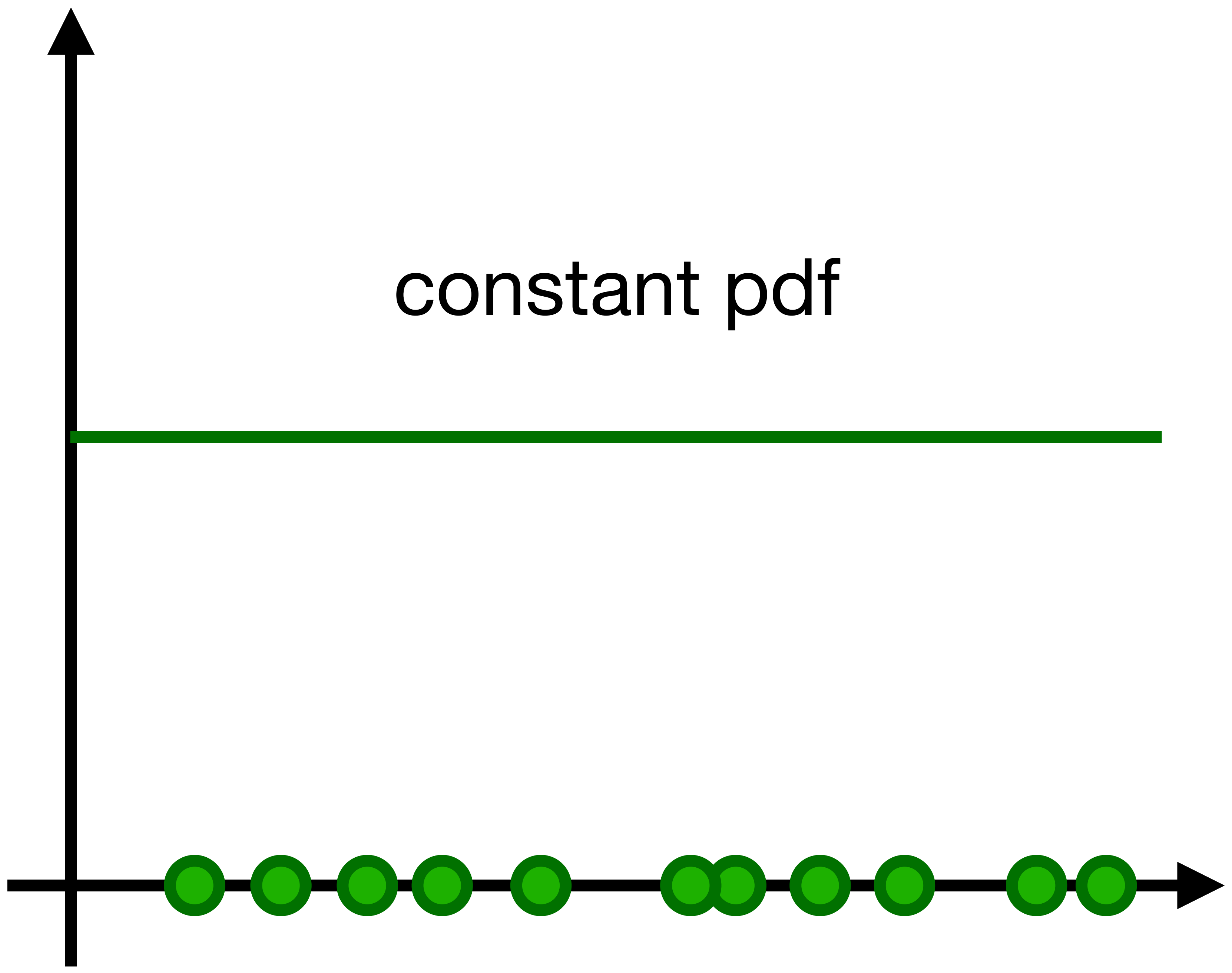
$$p(x) = \begin{cases} 1 & x \in [0,1) \\ 0 & \text{otherwise} \end{cases}$$

For non-uniform random variables:

$p(x)$ could be any function

Probability density function

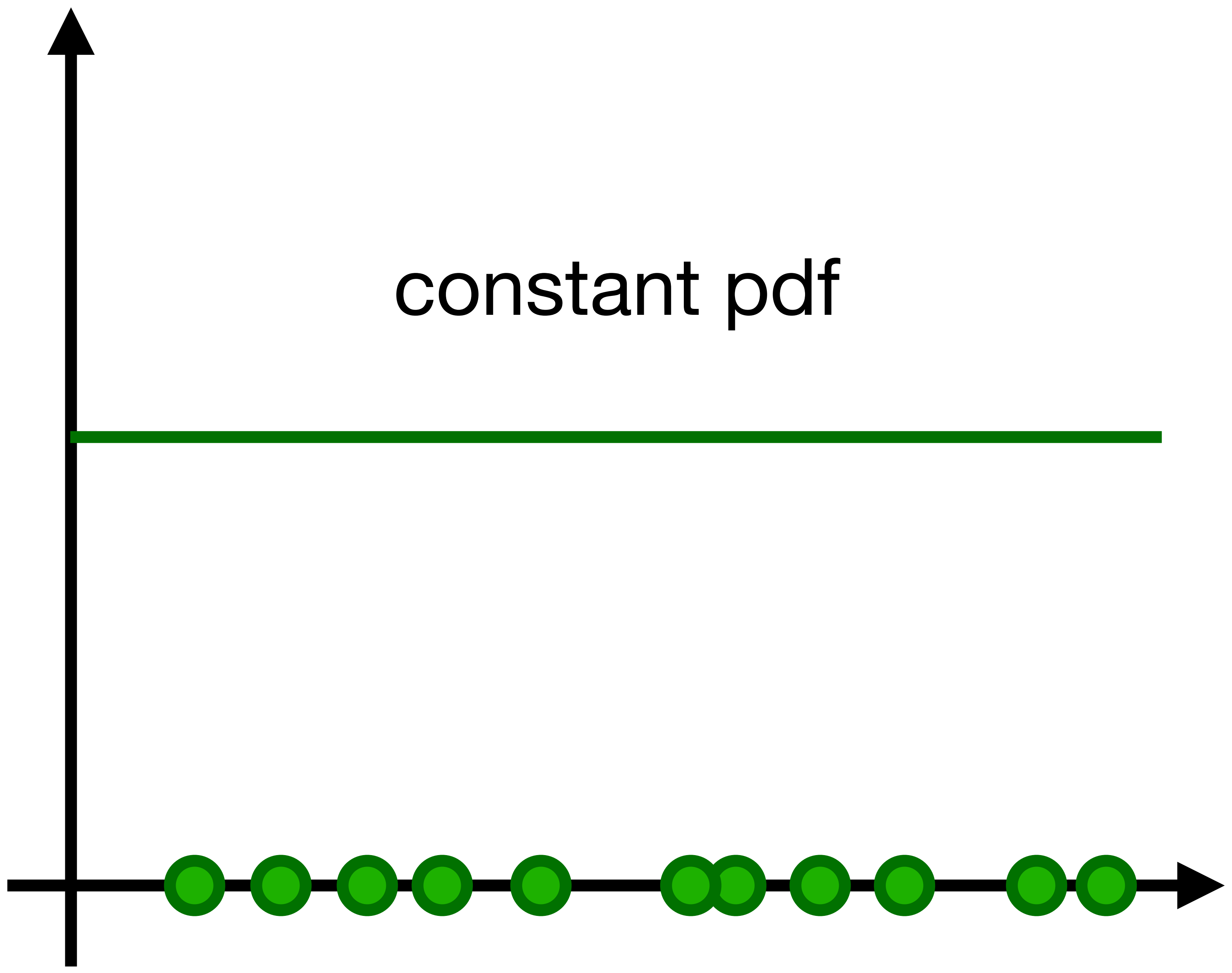
Uniform distribution



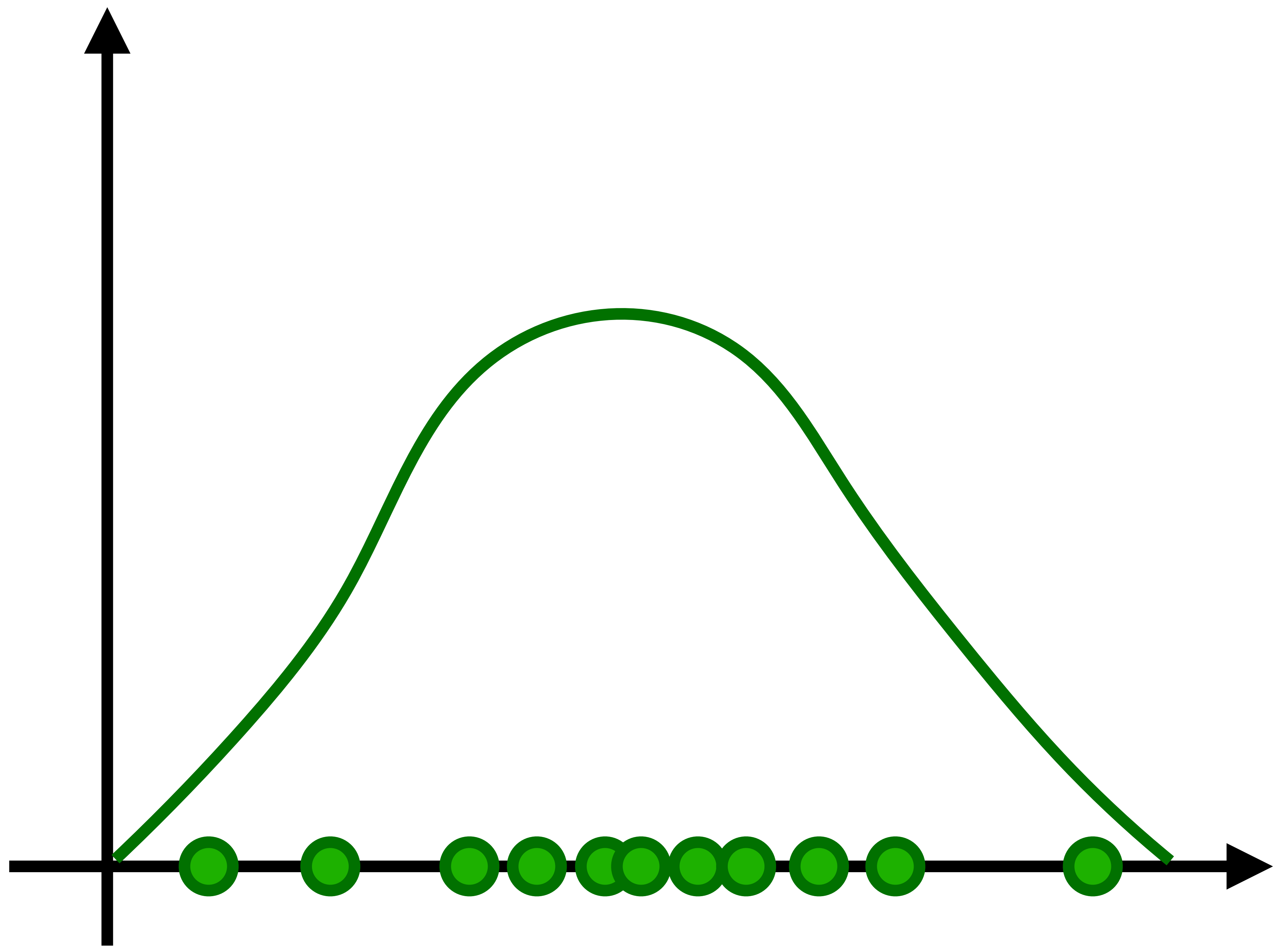
Non-uniform distribution

Probability density function

Uniform distribution



Non-uniform distribution



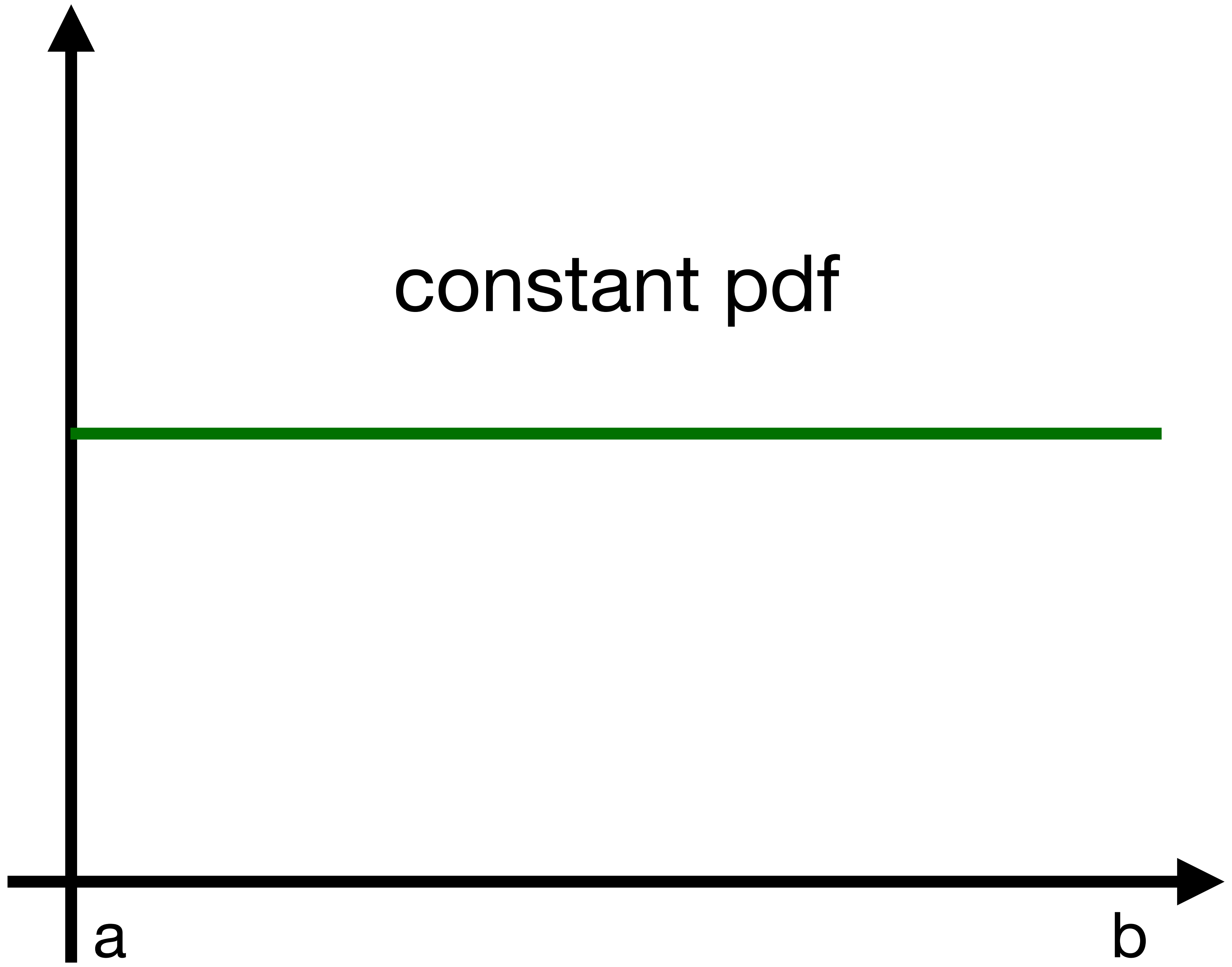
Probability density function

Some properties of PDFs:

$$p(x) > 0$$

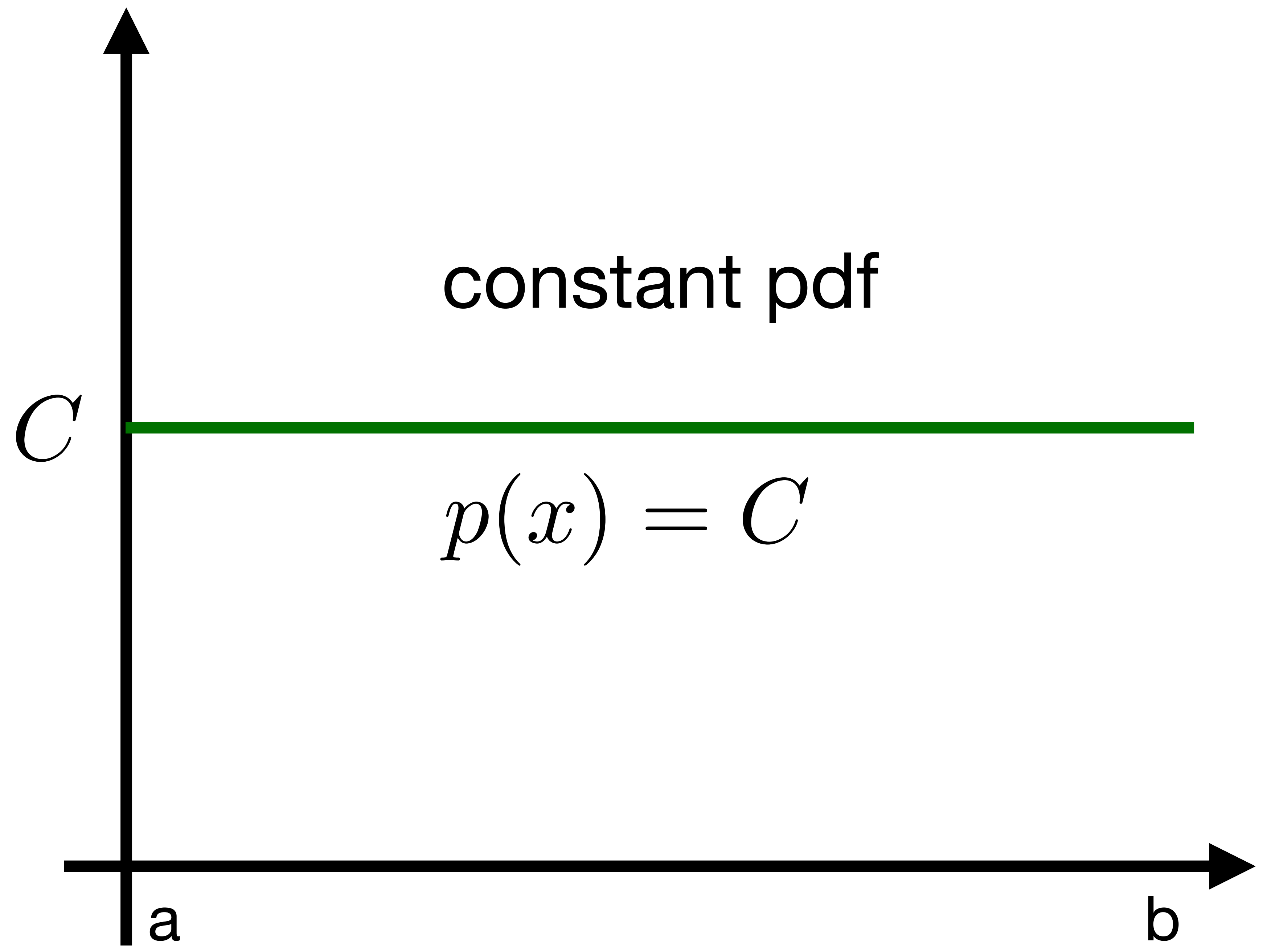
$$\int_{-\infty}^{\infty} p(x)dx = 1$$

Probability density function



$$\int_a^b p(x)dx = 1 \quad x \in [a, b)$$

Probability density function



$$\int_a^b p(x)dx = 1 \quad x \in [a,b)$$

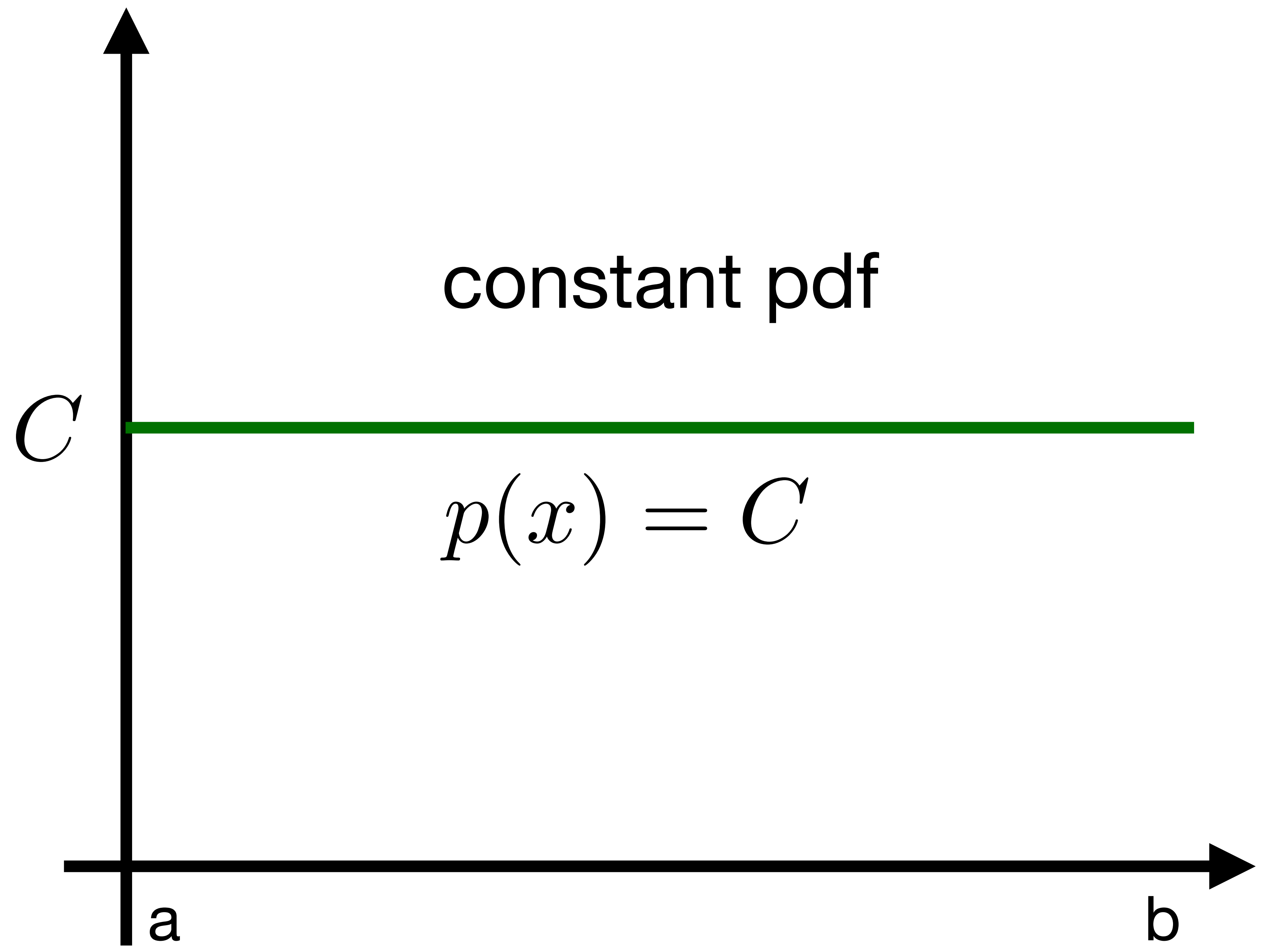
$$\int_a^b C\ dx = 1$$

$$C\int_a^b dx = 1$$

$$C(b-a) = 1$$

$$C = \frac{1}{b-a}$$

Probability density function



$$\int_a^b p(x)dx = 1 \quad x \in [a, b)$$

$$\int_a^b C\ dx = 1$$

$$C \int_a^b dx = 1$$

$$C(b-a) = 1$$

$$C = \frac{1}{b-a}$$

$$p(x) = \frac{1}{b-a}$$

Cumulative distribution function

- The PDF $p(x)$ is the derivative of the random variable's CDF:

UNIVERSITÄT DES SAARLANDES

Cumulative distribution function

- The PDF $p(x)$ is the derivative of the random variable's CDF:

$$p(x) = \frac{dP(x)}{dx}$$

$P(x)$: cumulative distribution function (CDF) ,
also called cumulative density function

Cumulative distribution function

- The PDF $p(x)$ is the derivative of the random variable's CDF:

$$p(x) = \frac{dP(x)}{dx}$$

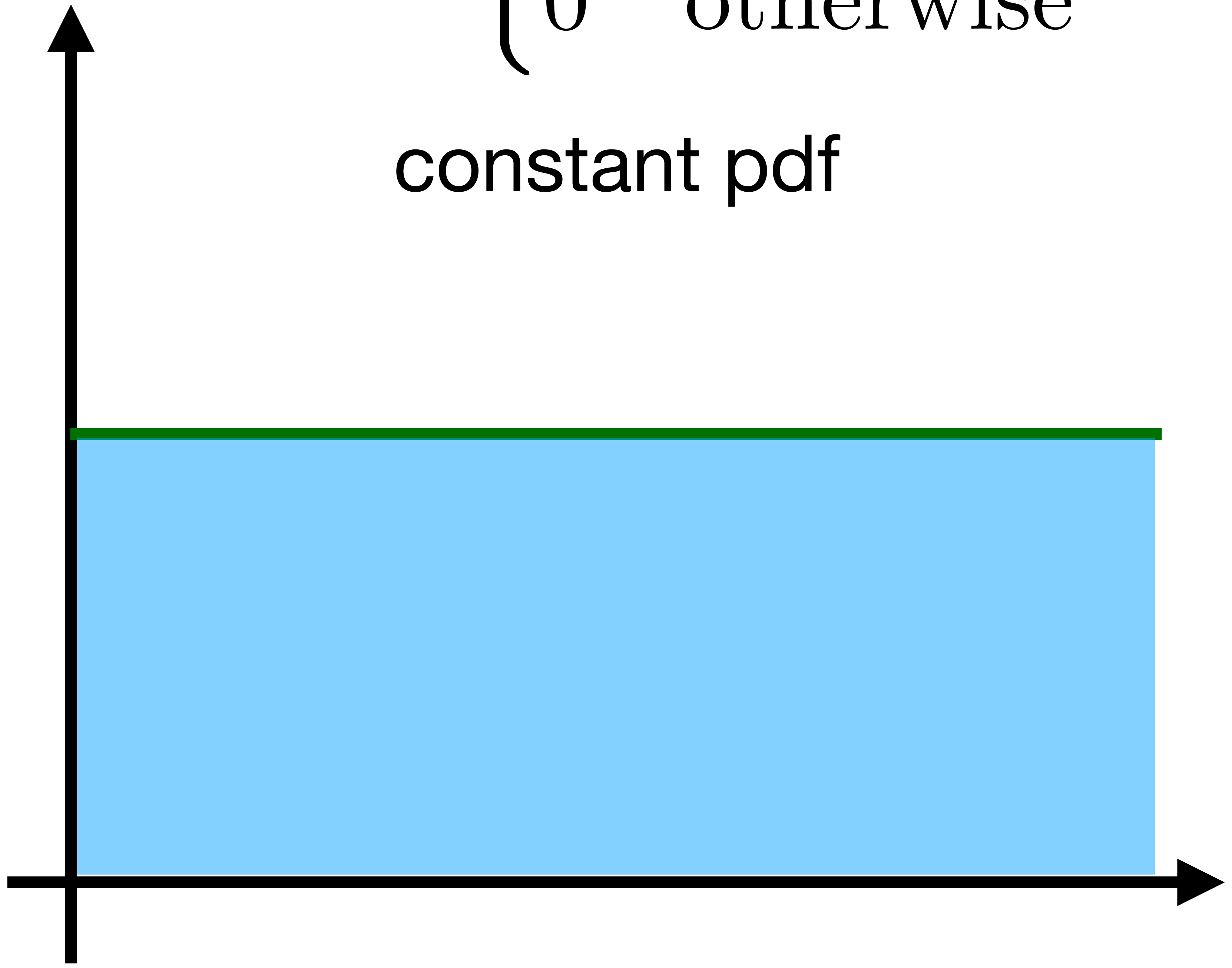
$$P(x) = \int_{-\infty}^{x} p(x)dx$$

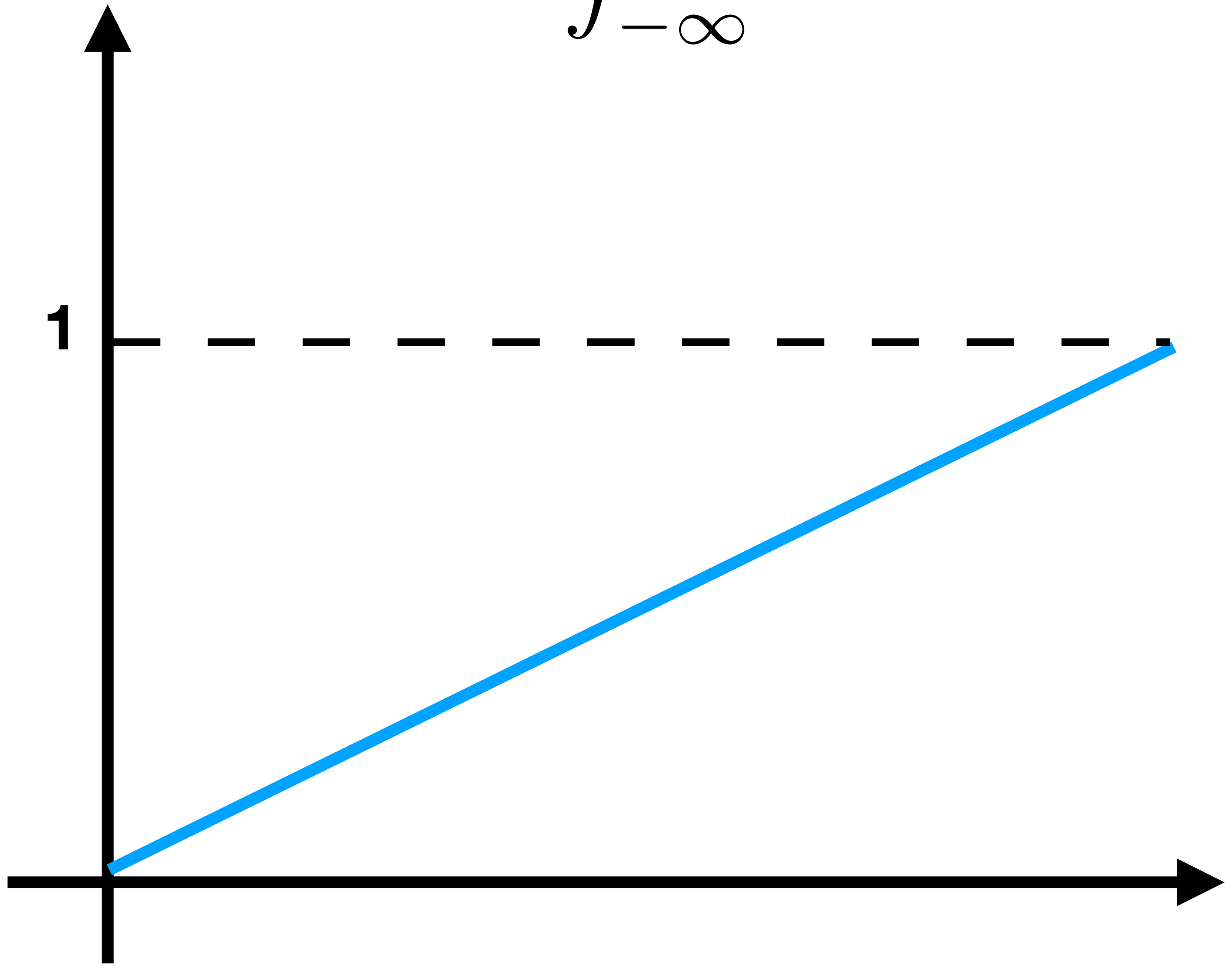
$P(x)$: cumulative distribution function (CDF) ,
also called cumulative density function

Cumulative distribution function

$$p(x) = \begin{cases} 1 & x \in [0,1) \\ 0 & \text{otherwise} \end{cases}$$

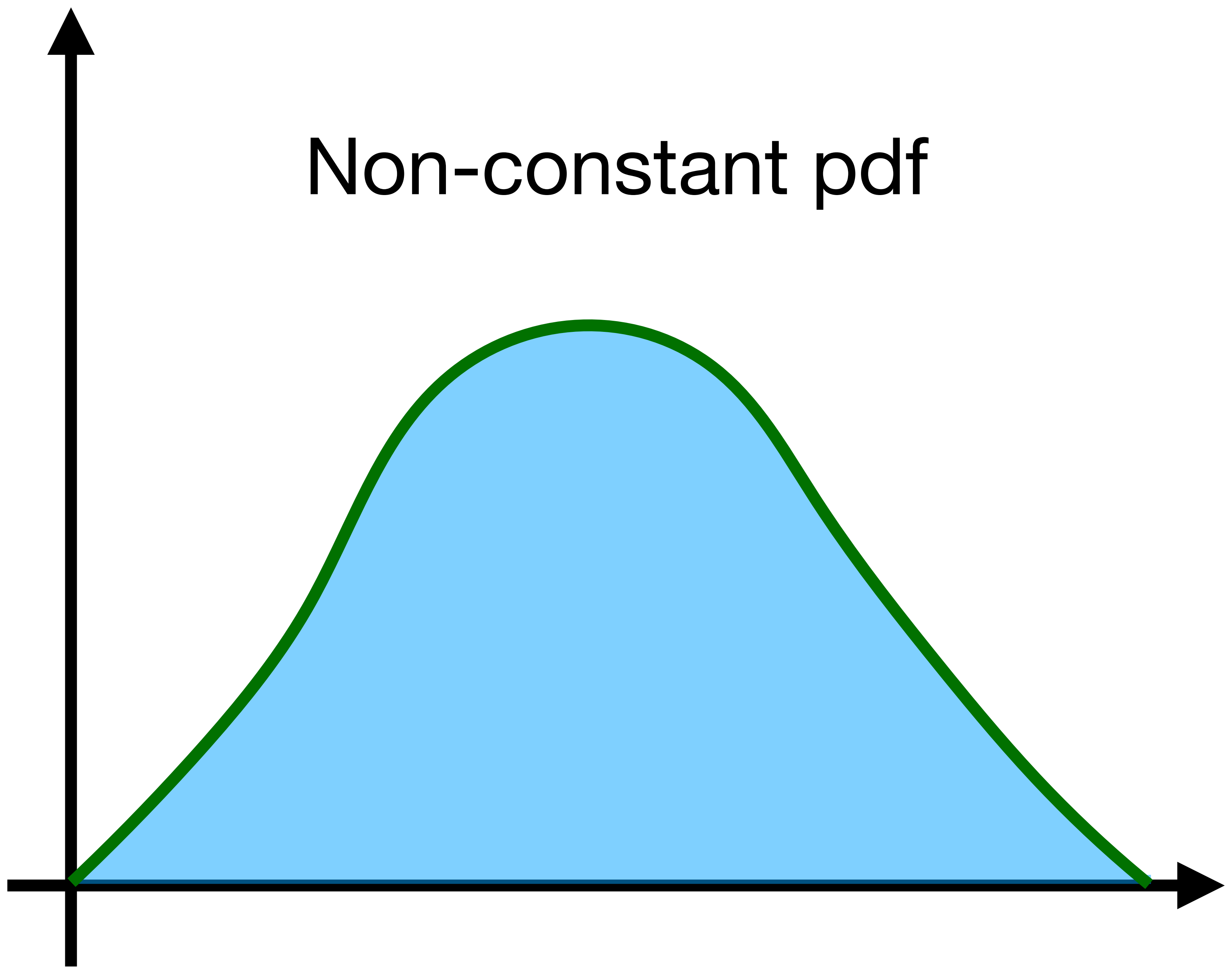
$$P(x) = \int_{-\infty}^{x} p(x)dx$$



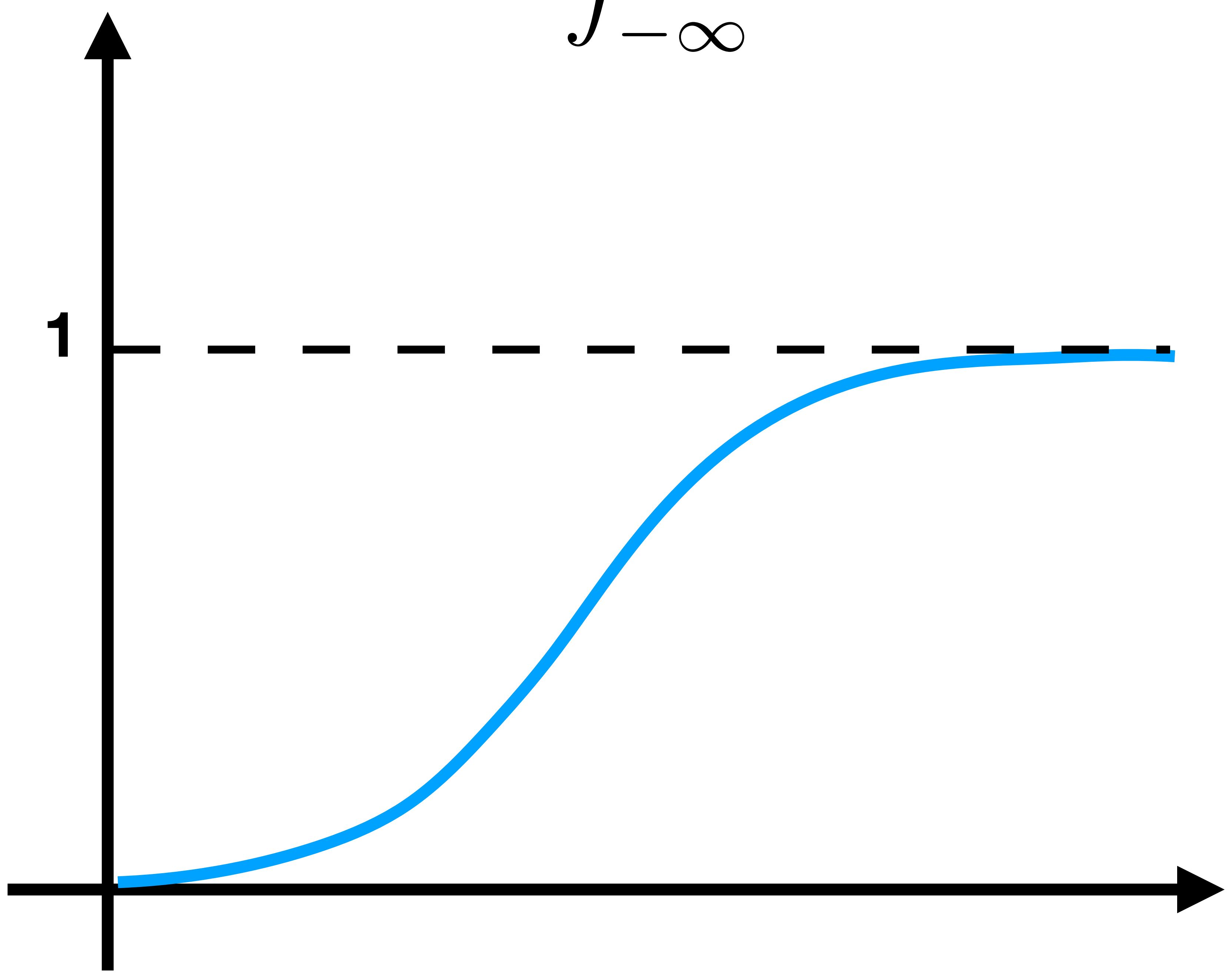


Cumulative distribution function

$$p(x)$$



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



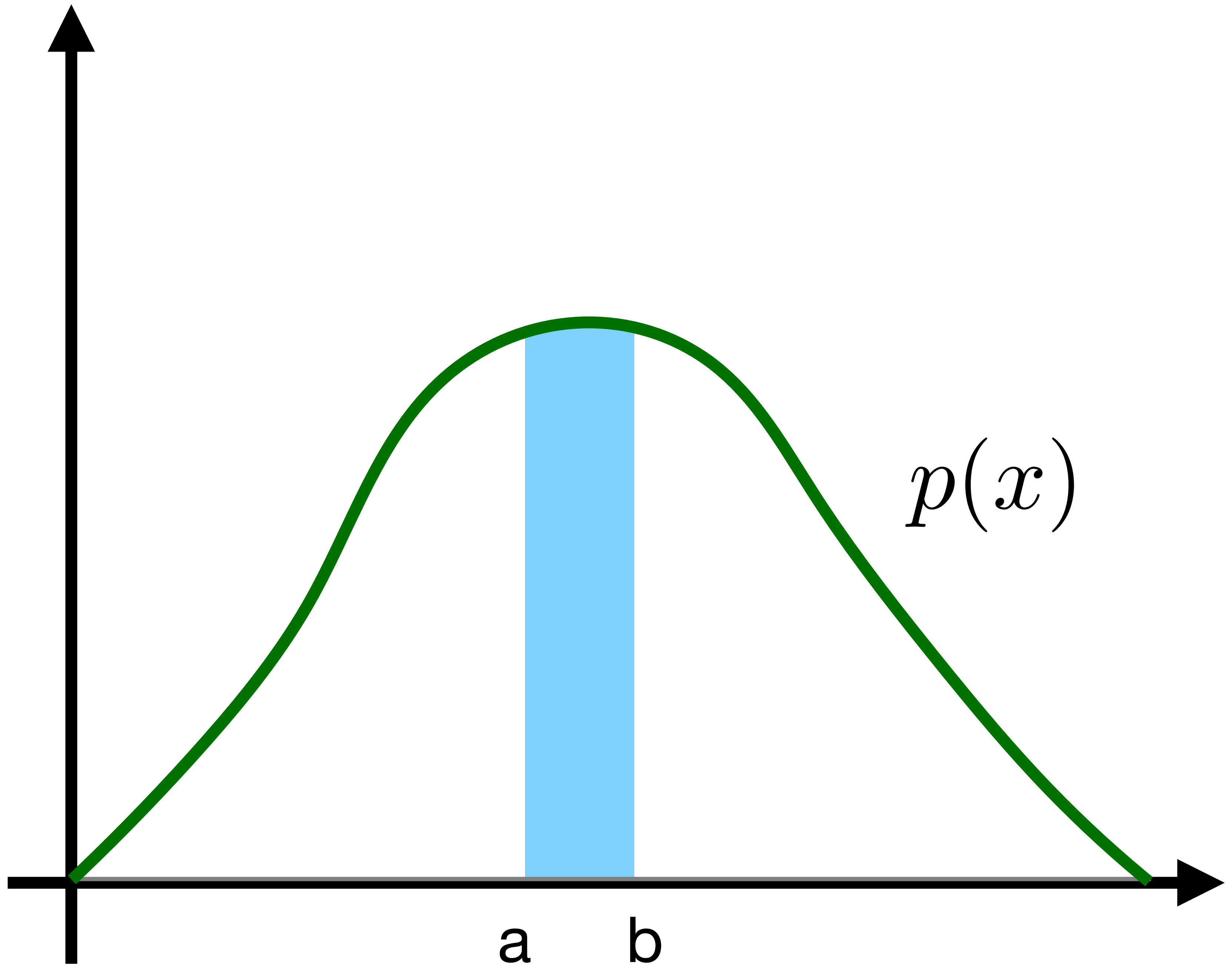
Visual Break

Image rendered using PBRT

Probability: Integral of PDF

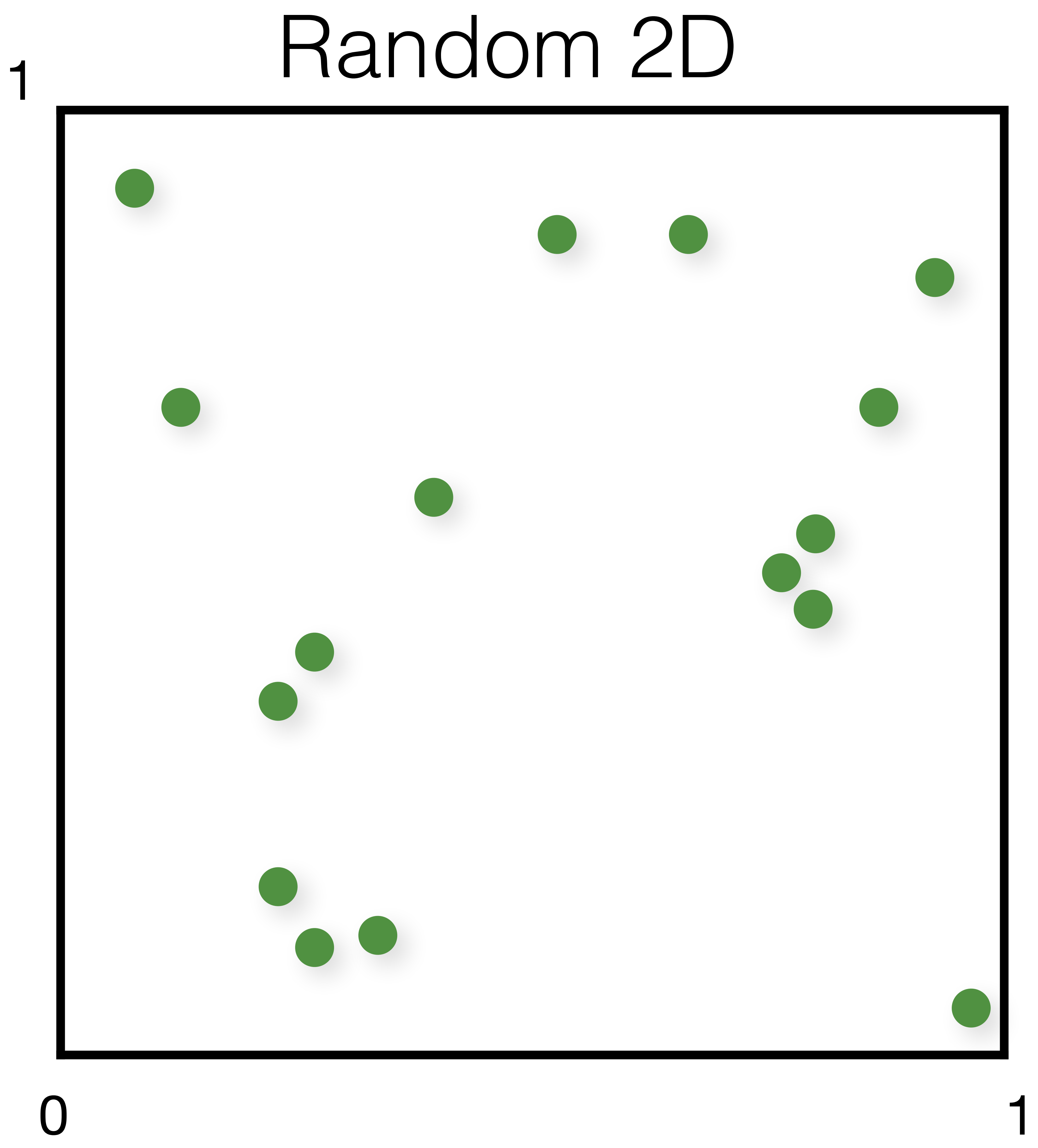
- Given the arbitrary interval $[a, b]$ in the domain, integrating the PDF gives the probability that a RV lies inside that interval:

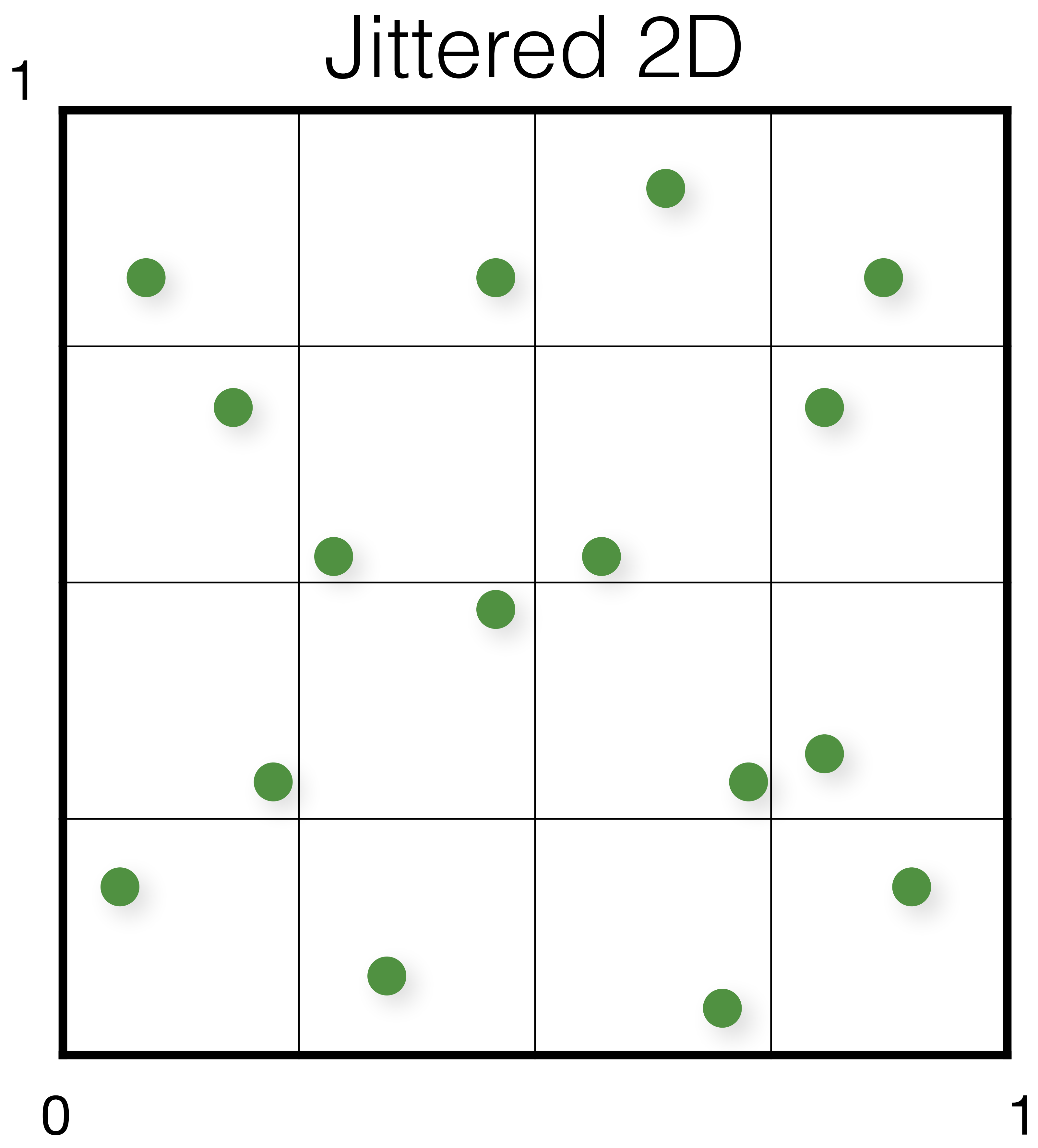
$$P(x \in [a, b]) = \int_a^b p(x)dx$$



Examples: Sampling PDFs

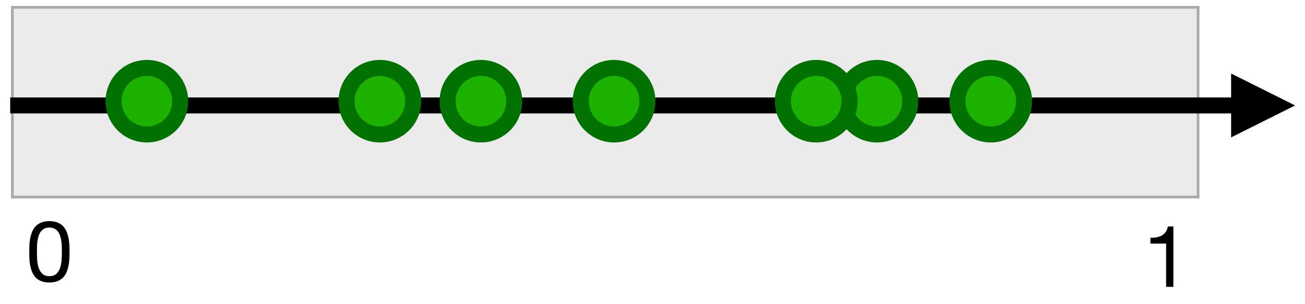
Constant Sampling PDFs





Constant Sampling PDFs

Random 1D

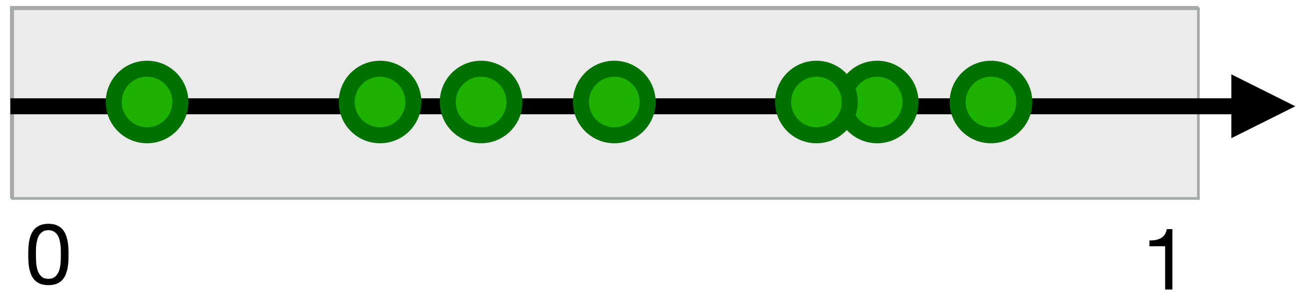


$$\xi \in [0, 1)$$

Sampling a unit domain with uniform random samples

Constant Sampling PDFs

Random 1D

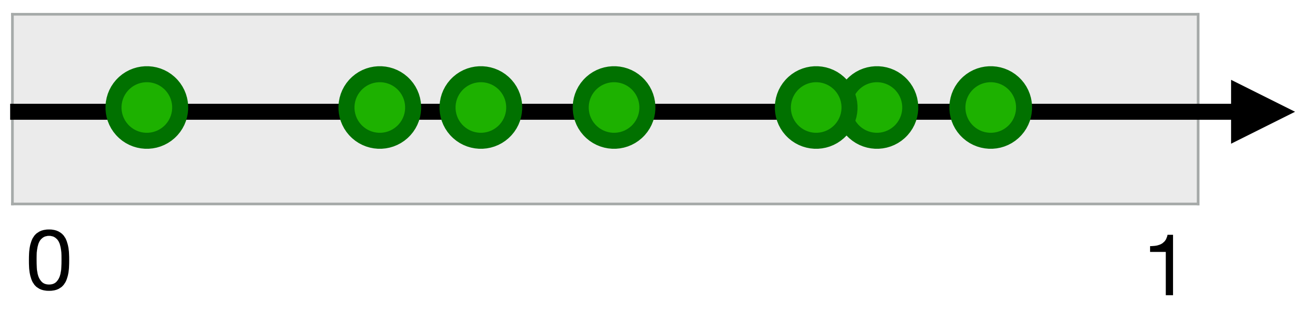


$$\xi \in [0, 1)$$

Sampling a unit domain with uniform random samples

Constant Sampling PDFs

Random 1D



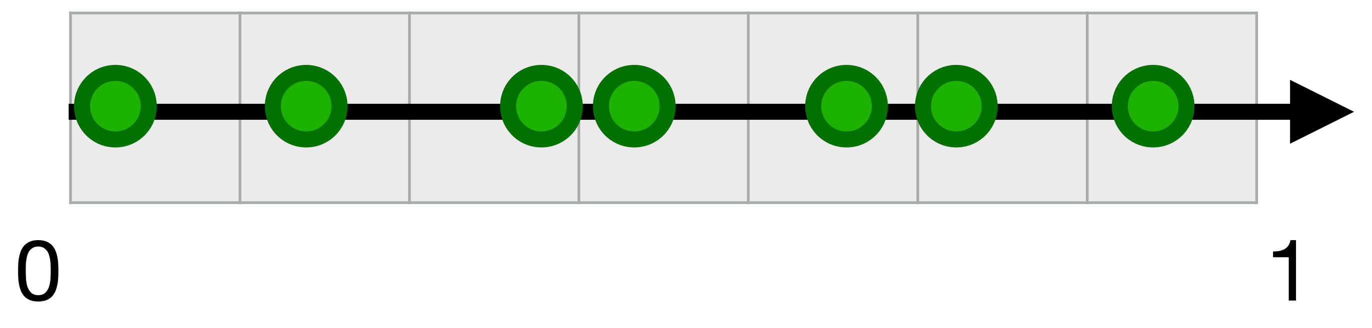
$$\xi \in [0, 1)$$

$$p(x) = \begin{cases} C & x \in [0, 1) \\ 0 & \text{otherwise} \end{cases}$$

Sampling a unit domain with uniform random samples

Constant Sampling PDFs

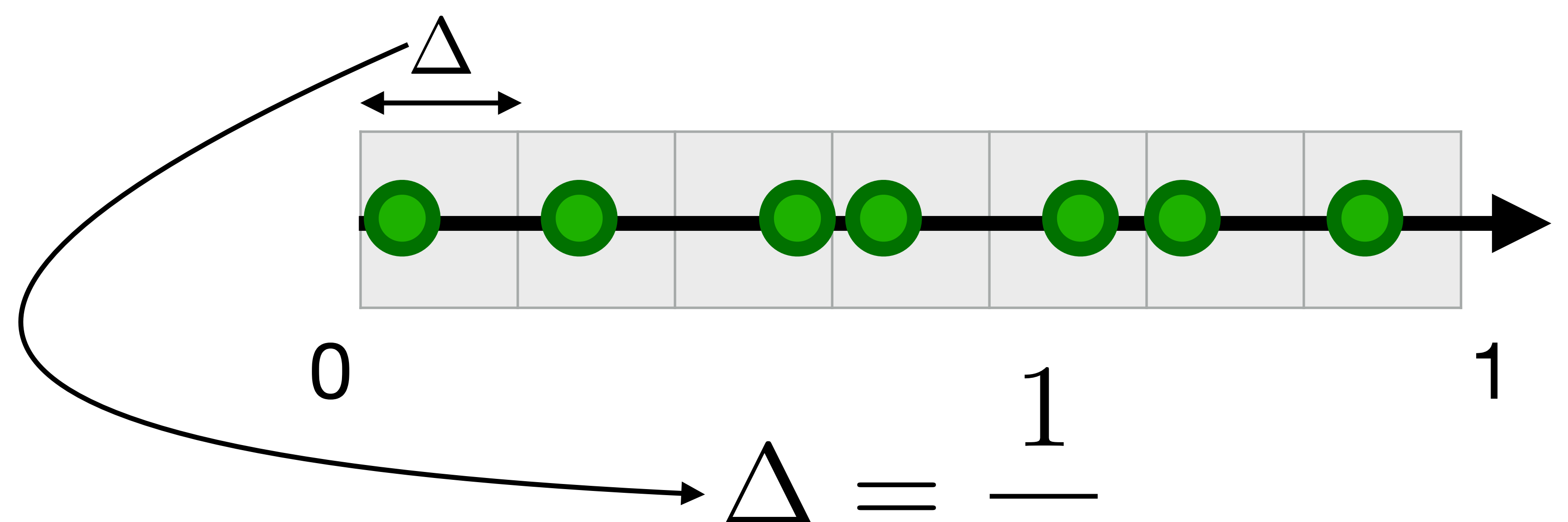
Jittered 1D



Sampling each stratum with uniform random samples

Constant Sampling PDFs

Jittered 1D

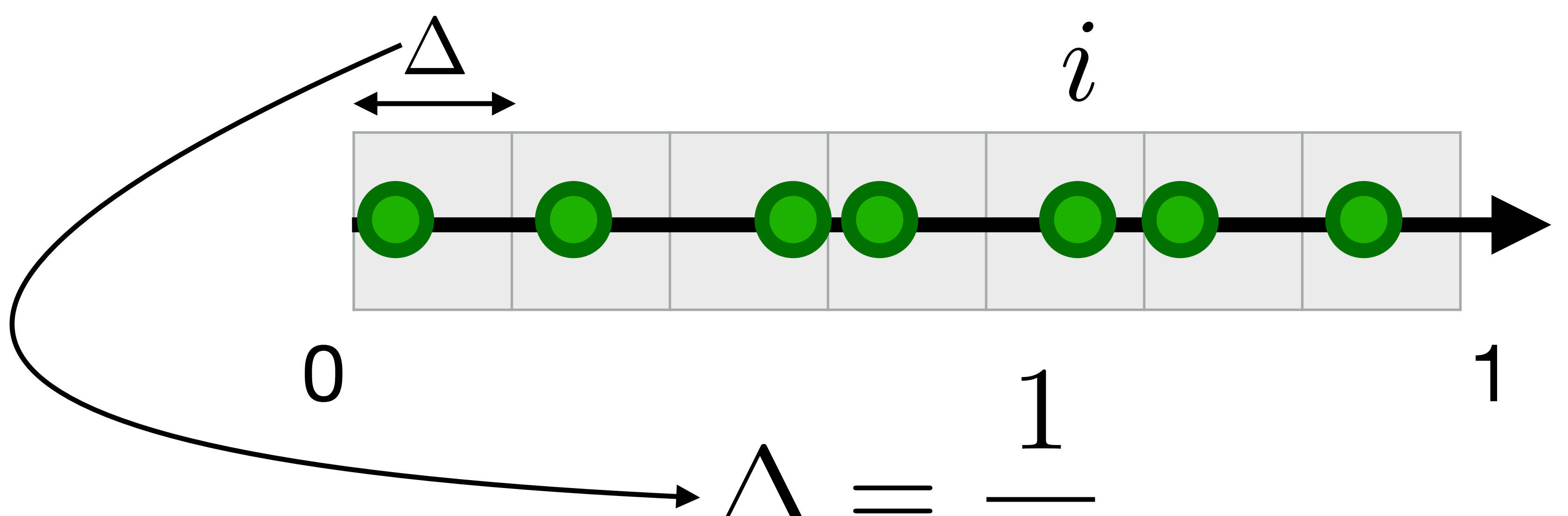


$$\Delta = \frac{1}{N}$$

Sampling each stratum with uniform random samples

Constant Sampling PDFs

Jittered 1D



$$\Delta = \frac{1}{N}$$

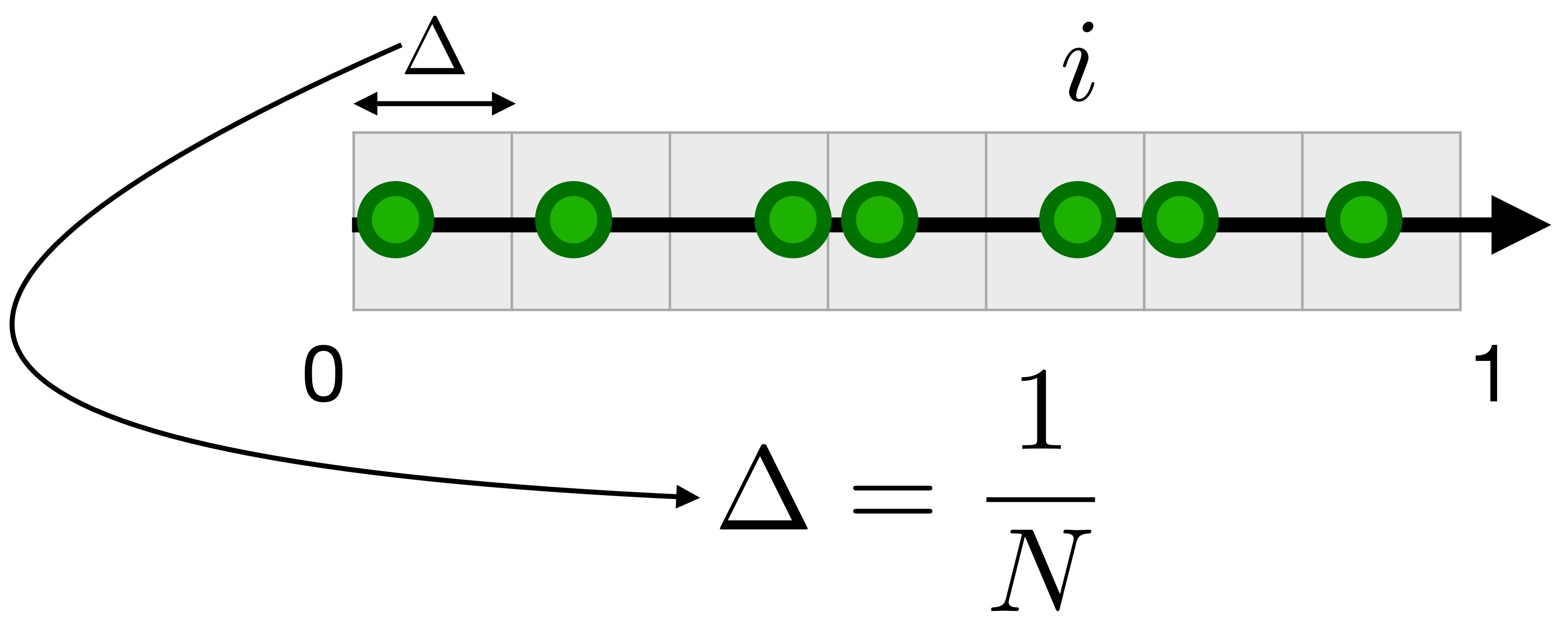
Probability density of generating a sample in an i-th stratum is given by:

$$p(x_i) = ???$$

Sampling each stratum with uniform random samples

Constant Sampling PDFs

Jittered 1D



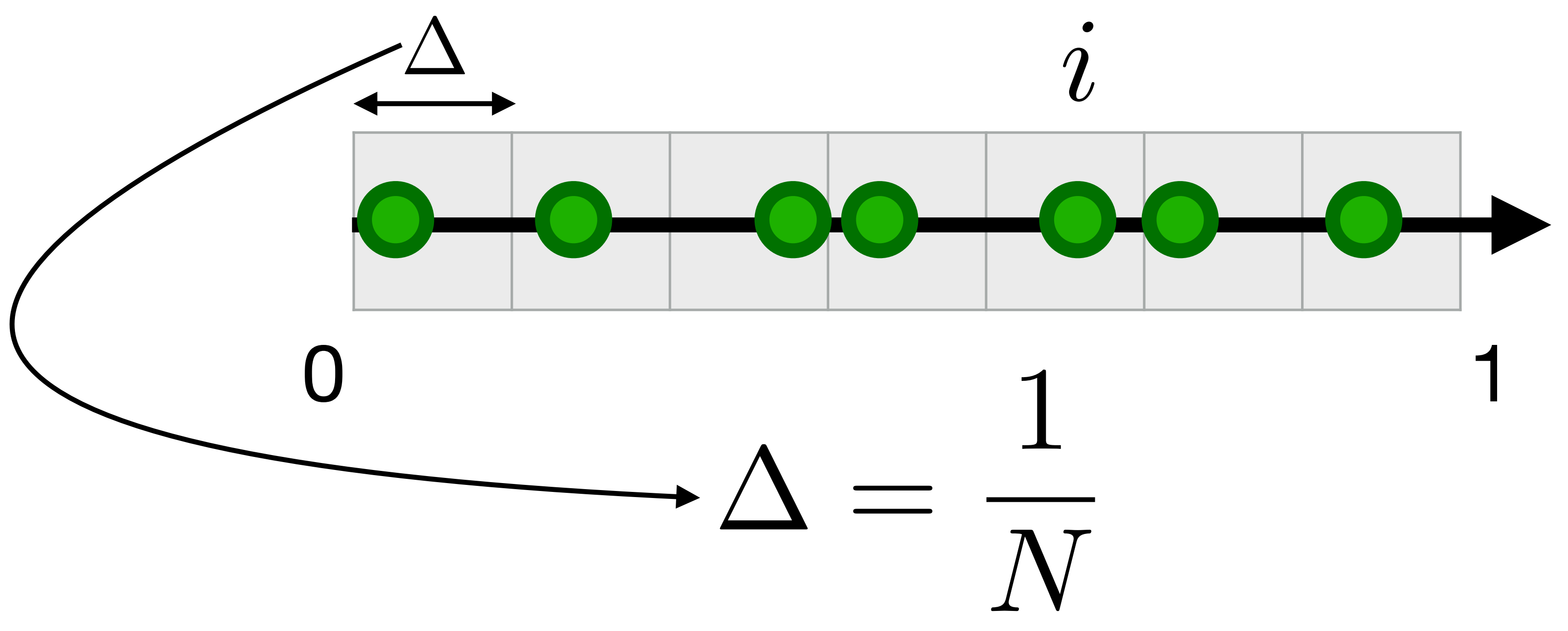
Probability density of generating a sample in an i-th stratum is given by:

$$p(x_i) = \begin{cases} N & x \in [\frac{i}{N}, \frac{i+1}{N}) \\ 0 & \text{otherwise} \end{cases}$$

Sampling each stratum with uniform random samples

Joint PDFs

Jittered 1D



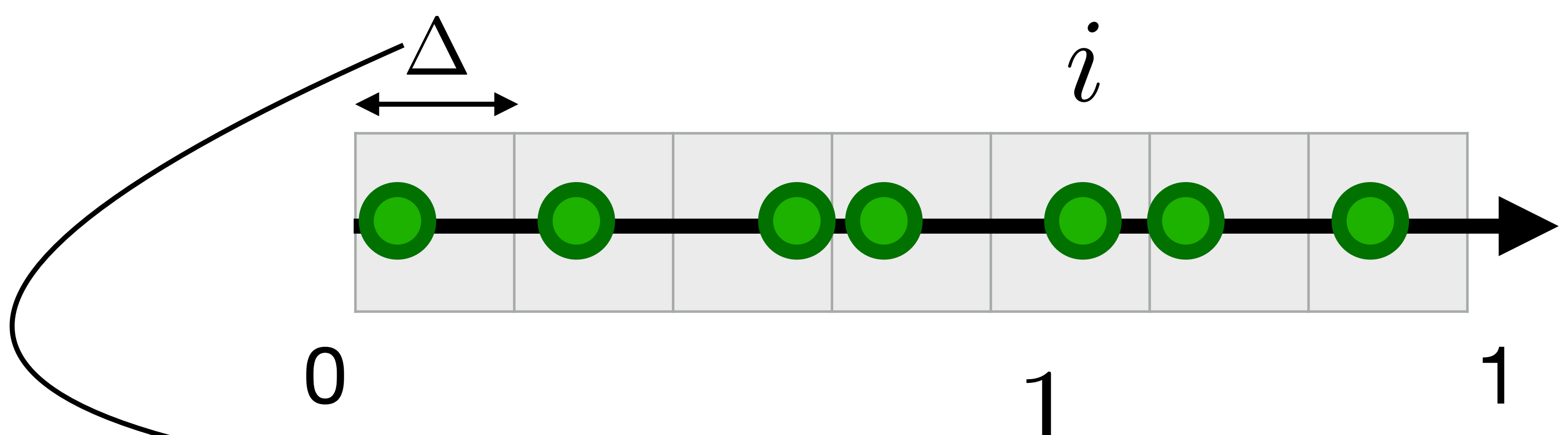
First, we divide the domain into equal strata.

Second, we sample the domain.

This implies that two samples are correlated to each other.

Joint PDFs

Jittered 1D



$$\Delta = \frac{1}{N}$$

First, we divide the domain into equal strata.

Second, we sample the domain.

This implies that two samples are correlated to each other.

For two different strata i and j, what is the joint PDF for jittered sampling ?

$$p(x_i, x_j) = ???$$

Conditional and Marginal PDFs

Joint PDF

For two random variables X_1 and X_2, the joint PDF $p(x_1, x_2)$ is given by:

Joint PDF

For two random variables X_1 and X_2, the joint PDF $p(x_1, x_2)$ is given by:

$$p(x_1, x_2) = p(x_2|x_1)p(x_1)$$

Joint PDF

For two random variables X_1 and X_2, the joint PDF $p(x_1, x_2)$ is given by:

$$p(x_1, x_2) = p(x_2|x_1)p(x_1)$$

where,

$X_1 = x_1$ $\quad p(x_2|x_1)$: conditional density function

$X_2 = x_2$ $\quad p(x_1)$: marginal density function

UNIVERSITÄT DES SAARLANDES

Joint PDF

For two random variables X_1 and X_2, the joint PDF $p(x_1, x_2)$ is given by:

$$p(x_1, x_2) = p(x_2|x_1)p(x_1)$$

where,

$X_1 = x_1$ $\quad p(x_2|x_1)$: conditional density function

$X_2 = x_2$ $\quad p(x_1)$: marginal density function

Joint PDF

For two random variables X_1 and X_2, the joint PDF $p(x_1, x_2)$ is given by:

$$p(x_1, x_2) = \underline{p(x_1|x_2)}\,\underline{p(x_2)}$$

where,

$X_1 = x_1$ $\quad p(x_1|x_2)$: conditional density function

$X_2 = x_2$ $\quad p(x_2)$: marginal density function

Marginal PDF

$$p(x_1) = \int_{\mathbb{R}} p(x_1, x_2) dx_2$$

$$p(x_2) = \int_{\mathbb{R}} p(x_1, x_2) dx_1$$

We integrate out one of the variable.

Conditional PDF

$$p(x_1|x_2) = \frac{p(x_1, x_2)}{p(x_2)}$$

$$p(x_2|x_1) = \frac{p(x_1, x_2)}{p(x_1)}$$

The conditional density function is the density function for x_i given that some particular x_j has been chosen.

Conditional PDF

If both x_1 and x_2 are independent then:

$$p(x_1|x_2) = p(x_1)$$

$$p(x_2|x_1) = p(x_2)$$

Conditional PDF

If both x_1 and x_2 are independent then:

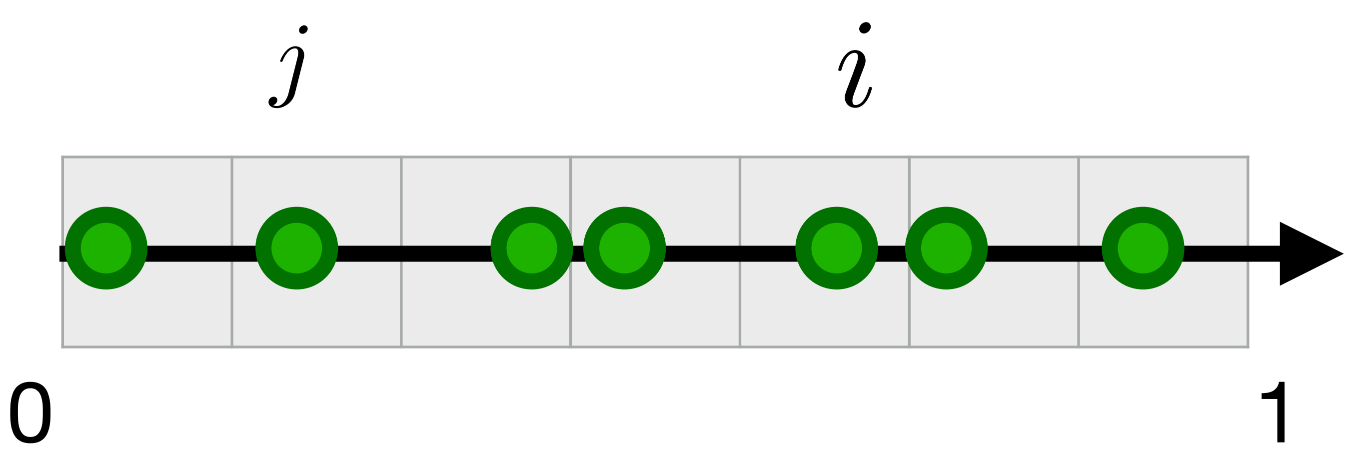
$$p(x_1|x_2) = p(x_1)$$

$$p(x_2|x_1) = p(x_2)$$

That gives:

$$p(x_1, x_2) = p(x_1)p(x_2)$$

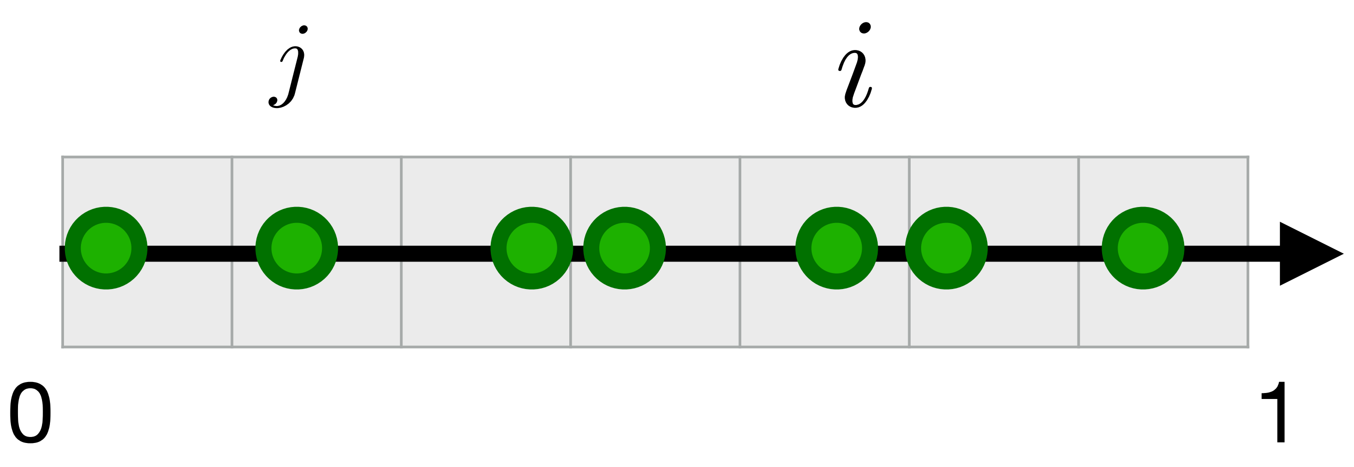
Joint PDF of Jittered 1D Sampling



For two different strata i and j, what is the joint PDF for jittered sampling ?

$$p(x_i, x_j) = ???$$

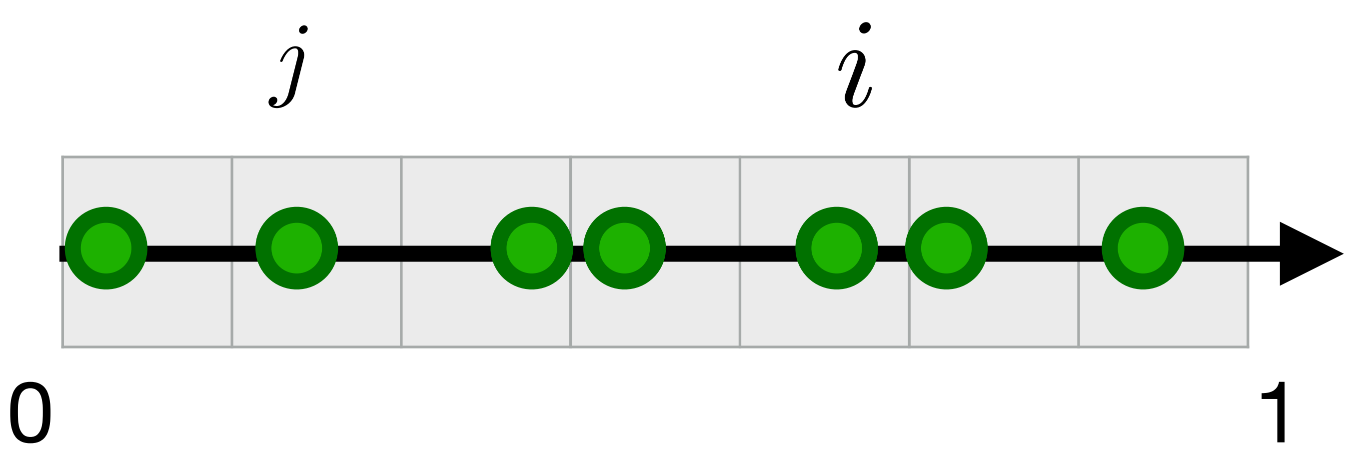
Joint PDF of Jittered 1D Sampling



$$p(x_1, x_2) = p(x_1|x_2)p(x_2)$$

UNIVERSITÄT DES SAARLANDES

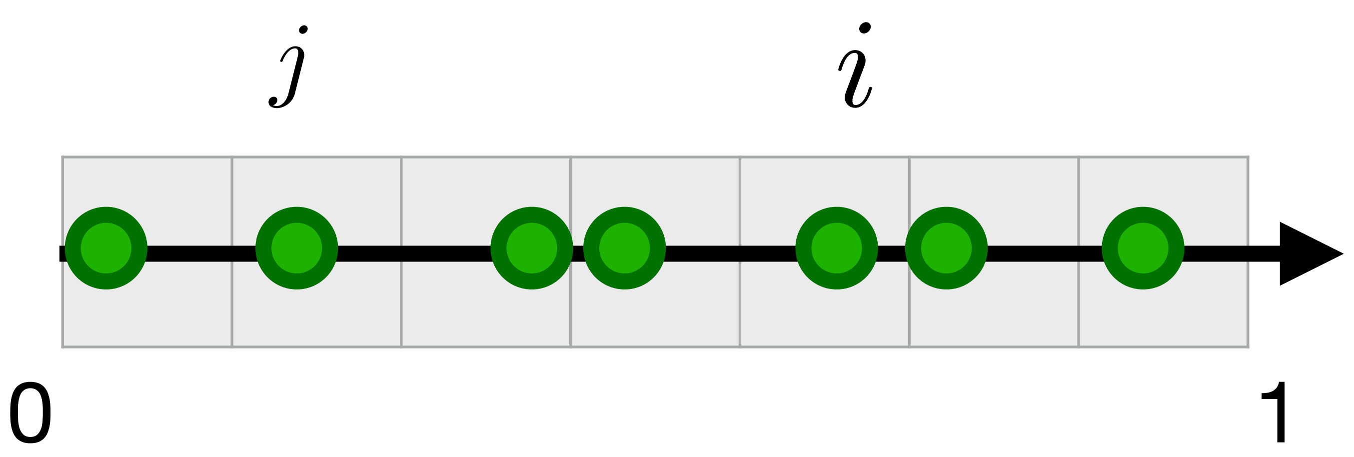
Joint PDF of Jittered 1D Sampling



$$p(x_1, x_2) = p(x_1 | x_2) p(x_2)$$

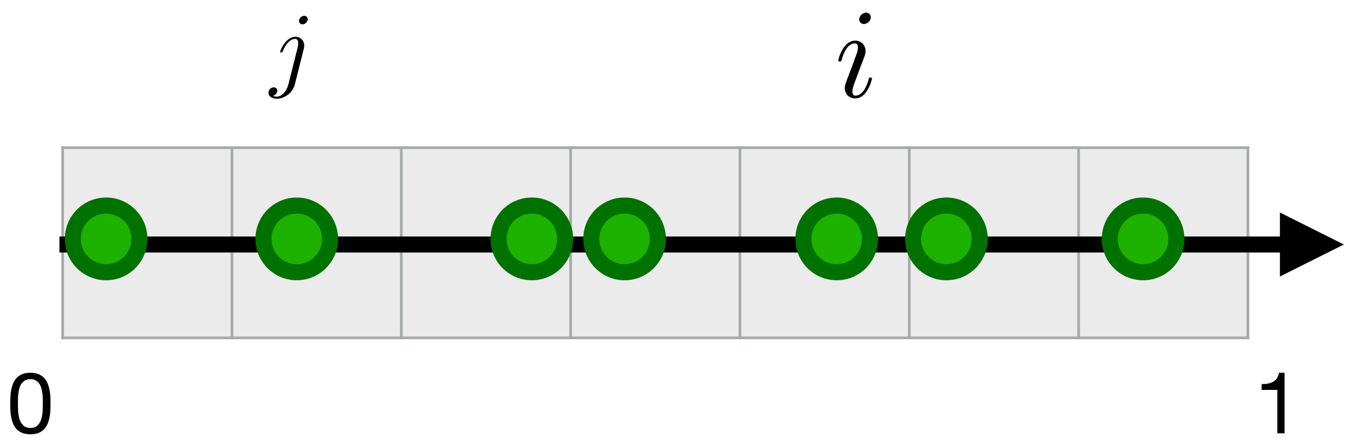
$$p(x_1, x_2) = p(x_1) p(x_2)$$

Joint PDF of Jittered 1D Sampling



$$p(x_i, x_j) = \begin{cases} p(x_i)p(x_j) & i \neq j \\ 0 & otherwise \end{cases}$$

Joint PDF of Jittered 1D Sampling



$$p(x_i, x_j) = \begin{cases} p(x_i)p(x_j) & i \neq j \\ 0 & otherwise \end{cases}$$

$$p(x_i, x_j) = \begin{cases} N^2 & i \neq j \\ 0 & otherwise \end{cases}$$

Since, $p(x_i) = N$

Visual Break

Image rendered using PBRT

Expected Value

Expected value

- Expected value: average value of the variable

$$E[X] = \sum_{i=1}^{N} x_i p_i$$

- example: rolling a die

$$E[X] =$$

UNIVERSITÄT DES SAARLANDES

Expected value

- Expected value: average value of the variable

$$E[X] = \sum_{i=1}^{N} x_i p_i$$

- example: rolling a die

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

Expected value

- Properties:

$$E[X + Y] = E[X] + E[Y]$$

Expected value

- Properties:

$$E[X+Y] = E[X] + E[Y]$$
$$E[X+c] = E[X] + c$$

Expected value

- Properties:

$$E[X+Y] = E[X] + E[Y]$$
$$E[X+c] = E[X] + c$$
$$E[cX] = cE[X]$$

Estimating expected values

- To estimate the expected value of a variable
  - choose a set of random values based on the probability
  - average their results

$$E[X] \approx \frac{1}{N} \sum_{i=1}^{N} x_i$$

- example: rolling a die
  - roll 3 times: $\{3, 1, 6\} \rightarrow E[\mathrm{x}] \approx$

Estimating expected values

- To estimate the expected value of a variable
  - choose a set of random values based on the probability
  - average their results

$$E[X] \approx \frac{1}{N} \sum_{i=1}^{N} x_i$$

- example: rolling a die
  - roll 3 times: $\{3, 1, 6\} \rightarrow E[\mathrm{x}] \approx (3 + 1 + 6)/3 = 3.33$

Estimating expected values

- To estimate the expected value of a variable
  - choose a set of random values based on the probability
  - average their results

$$E[X] \approx \frac{1}{N} \sum_{i=1}^{N} x_i$$

- example: rolling a die
  - roll 3 times: $\{3, 1, 6\} \rightarrow E[\mathrm{x}] \approx (3 + 1 + 6)/3 = 3.33$

Estimating expected values

- To estimate the expected value of a variable
  - choose a set of random values based on the probability
  - average their results

$$E[X] \approx \frac{1}{N}\sum_{i=1}^{N} x_i$$

- example: rolling a die
  - roll 3 times: $\{3, 1, 6\} \rightarrow E[\mathrm{x}] \approx (3 + 1 + 6)/3 = 3.33$
  - roll 9 times: $\{3, 1, 6, 2, 5, 3, 4, 6, 2\} \rightarrow E[\mathrm{x}]$

Estimating expected values

- To estimate the expected value of a variable
  - choose a set of random values based on the probability
  - average their results

$$E[X] \approx \frac{1}{N}\sum_{i=1}^{N} x_i$$

- example: rolling a die
  - roll 3 times: $\{3, 1, 6\} \rightarrow E[\mathrm{x}] \approx (3 + 1 + 6)/3 = 3.33$
  - roll 9 times: $\{3, 1, 6, 2, 5, 3, 4, 6, 2\} \rightarrow E[\mathrm{x}]$

Estimating expected values

- To estimate the expected value of a variable
  - choose a set of random values based on the probability
  - average their results

$$E[X] \approx \frac{1}{N} \sum_{i=1}^{N} x_i$$

- example: rolling a die
  - roll 3 times: $\{3, 1, 6\} \rightarrow E[\mathrm{x}] \approx (3 + 1 + 6)/3 = 3.33$
  - roll 9 times: $\{3, 1, 6, 2, 5, 3, 4, 6, 2\} \rightarrow E[\mathrm{x}] \approx 3.51$

Law of large numbers

- By taking infinitely many samples, the error between the estimate and the expected value is statistically zero
  - the estimate will converge to the right value

$$\text{probability}\left[E[x] = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x_i\right] = 1$$

Variance

Variance

- Variance: how much different from the average

$$\sigma^2[X] = E[(X - E[X])^2]$$

Variance

- Variance: how much different from the average

$$\sigma^2[X] = E[(X - E[X])^2]$$
$$= E[X^2 + E[X]^2 - 2XE[X]]$$

Variance

- Variance: how much different from the average

$$\begin{aligned}\sigma^2[X] &= E[(X - E[X])^2] \\ &= E[X^2 + E[X]^2 - 2XE[X]] \\ &= E[X^2] + E[E[X]^2] - 2E[X]E[E[X]]\end{aligned}$$

Variance

- Variance: how much different from the average

$$
\begin{aligned}
\sigma^2[X] &= E[(X - E[X])^2] \\
&= E[X^2 + E[X]^2 - 2XE[X]] \\
&= E[X^2] + E[E[X]^2] - 2E[X]E[E[X]]] \\
&= E[X^2] + E[X]^2 - 2E[X]^2
\end{aligned}
$$

Variance

- Variance: how much different from the average

$$
\begin{aligned}
\sigma^2[X] &= E[(X - E[X])^2] \\
&= E[X^2 + E[X]^2 - 2XE[X]] \\
&= E[X^2] + E[E[X]^2] - 2E[X]E[E[X]]] \\
&= E[X^2] + E[X]^2 - 2E[X]^2 \\
&= E[X^2] - E[X]^2
\end{aligned}
$$

Variance

- Variance: how much different from the average

$$
\begin{aligned}
\sigma^2[X] &= E[(X - E[X])^2] \\
&= E[X^2 + E[X]^2 - 2XE[X]] \\
&= E[X^2] + E[E[X]^2] - 2E[X]E[E[X]]] \\
&= E[X^2] + E[X]^2 - 2E[X]^2 \\
&= E[X^2] - E[X]^2
\end{aligned}
$$

$$\sigma^2[X] = E[X^2] - E[X]^2$$

Variance

- example: Rolling a die
  - variance:

$$\sigma^2[X] = E[X^2] - E[X]^2$$

$$\sigma^2[X] = \ldots =$$

Variance

- example: Rolling a die
  - variance:

$$\sigma^2[X] = E[X^2] - E[X]^2$$

$$\sigma^2[X] = \ldots =$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

Variance

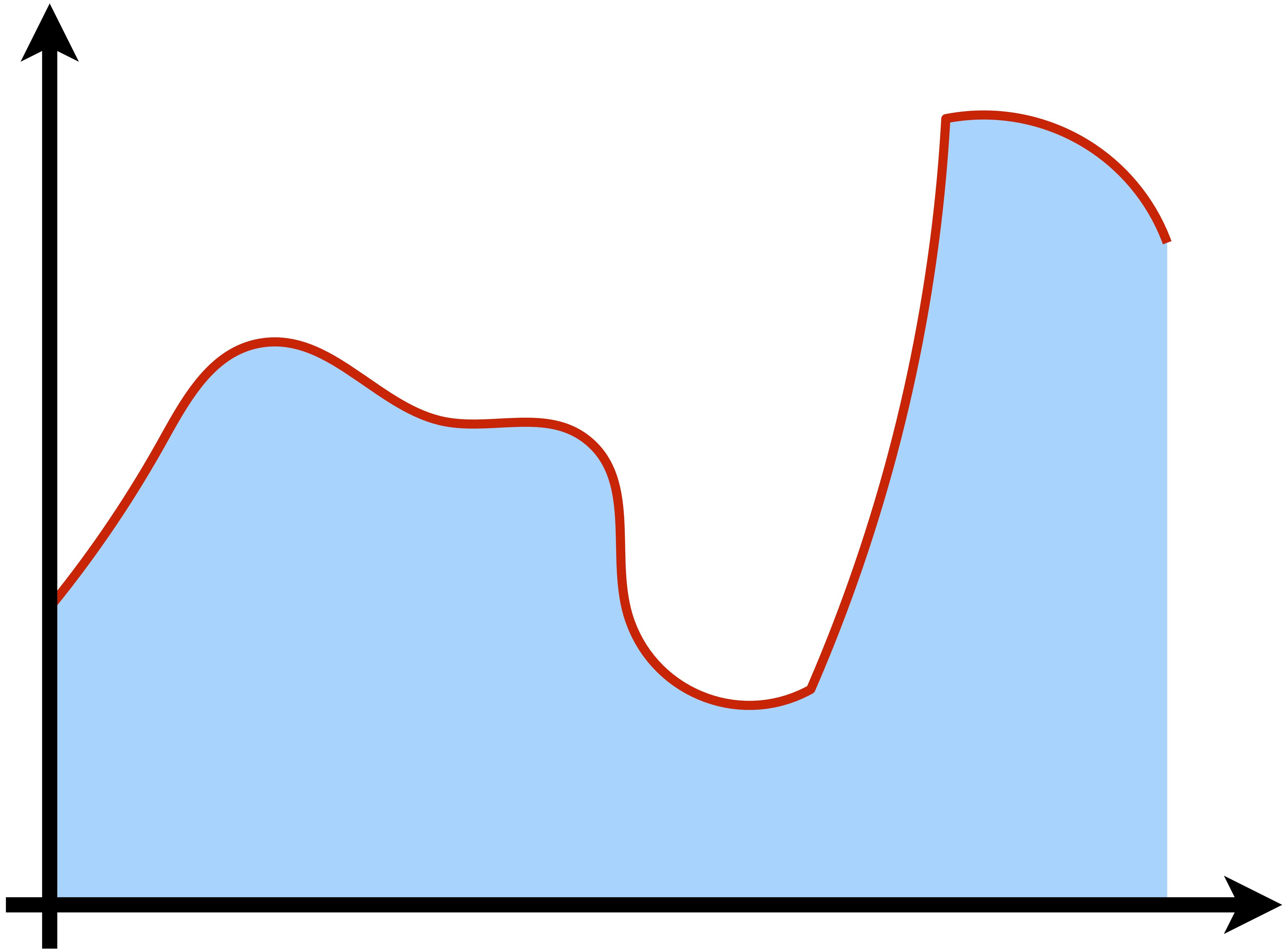
- example: Rolling a die
  - variance:

$$\sigma^2[X] = E[X^2] - E[X]^2$$

$$\sigma^2[X] = \ldots = 2.917$$

Monte Carlo Integration

$$I = \int_D f(x) \, \mathrm{d}x$$



Slide after Wojciech Jarosz


Questions ?
Image rendered using PBRT