Probability & Statistics

Fundamentals

Rahul Singh
rsingh@arrsingh.com

Experiments & Events

Experiment (aka Trial)

Any procedure that can be repeated infinitely and has a well defined set of outcomes

Random Experiment (aka Trial)

If the experiment can have more than one possible outcome, then its a random experiment

Event

An event is the subset of the outcomes of an experiment

Example

Experiment: Flipping a fair coin (this is a random event)

Outcomes: This experiment has two outcomes: Heads or Tails

Event: The coin lands with heads on top

Random Variable

A numeric quantity associated with random events.

Random variable maps the set of possible outcomes (the domain) to a numeric quantity (the range)

Discrete Random Variable

A random variable with a finite, countable, distinct set of outcomes

Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{cases}$$

Two Possible Outcomes: Heads or Tails. Heads is mapped to 1 and Tails is mapped to 0

Random Variable

A numeric quantity associated with random events.

Random variable maps the set of possible outcomes (the domain) to a numeric quantity (the range)

Discrete Random Variable

A random variable with a finite, countable, distinct set of outcomes

Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\,\alpha = Heads \\ 0, & if\,\alpha = Tails \end{cases}$$

Two Possible Outcomes: Heads or Tails. Heads is mapped to 1 and Tails is mapped to 0

X is a discrete random variable

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{Bmatrix} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{Bmatrix}$$

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{Bmatrix} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{Bmatrix}$$

$$P(X = 1) = \frac{1}{2} = 0.5$$

Probability that the coin lands on Heads ($X = 1$)

$$P(X = 1) = \frac{Number\ of\ Desired\ Outcomes}{Total\ Number\ of\ Outcomes} = \frac{1}{2} = 0.5$$

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{cases}$$

$$P(X = 1) = \frac{1}{2} = 0.5$$

$$P(X = 0) = \frac{1}{2} = 0.5$$

Probability that the coin lands on Heads ($X = 1$)

$$P(X = 1) = \frac{Number\ of\ Desired\ Outcomes}{Total\ Number\ of\ Outcomes} = \frac{1}{2} = 0.5$$

Probability that the coin lands on Tails ($X = 0$)

$$P(X = 0) = \frac{Number\ of\ Desired\ Outcomes}{Total\ Number\ of\ Outcomes} = \frac{1}{2} = 0.5$$

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{cases}$$

$$P(X = 1) = \frac{1}{2} = 0.5$$

$$P(X = 0) = \frac{1}{2} = 0.5$$

We can plot these probabilities on a graph

Probability that the coin lands on Heads ($X = 1$)

$$P(X = 1) = \frac{Number\ of\ Desired\ Outcomes}{Total\ Number\ of\ Outcomes} = \frac{1}{2} = 0.5$$

Probability that the coin lands on Tails ($X = 0$)

$$P(X = 0) = \frac{Number\ of\ Desired\ Outcomes}{Total\ Number\ of\ Outcomes} = \frac{1}{2} = 0.5$$

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

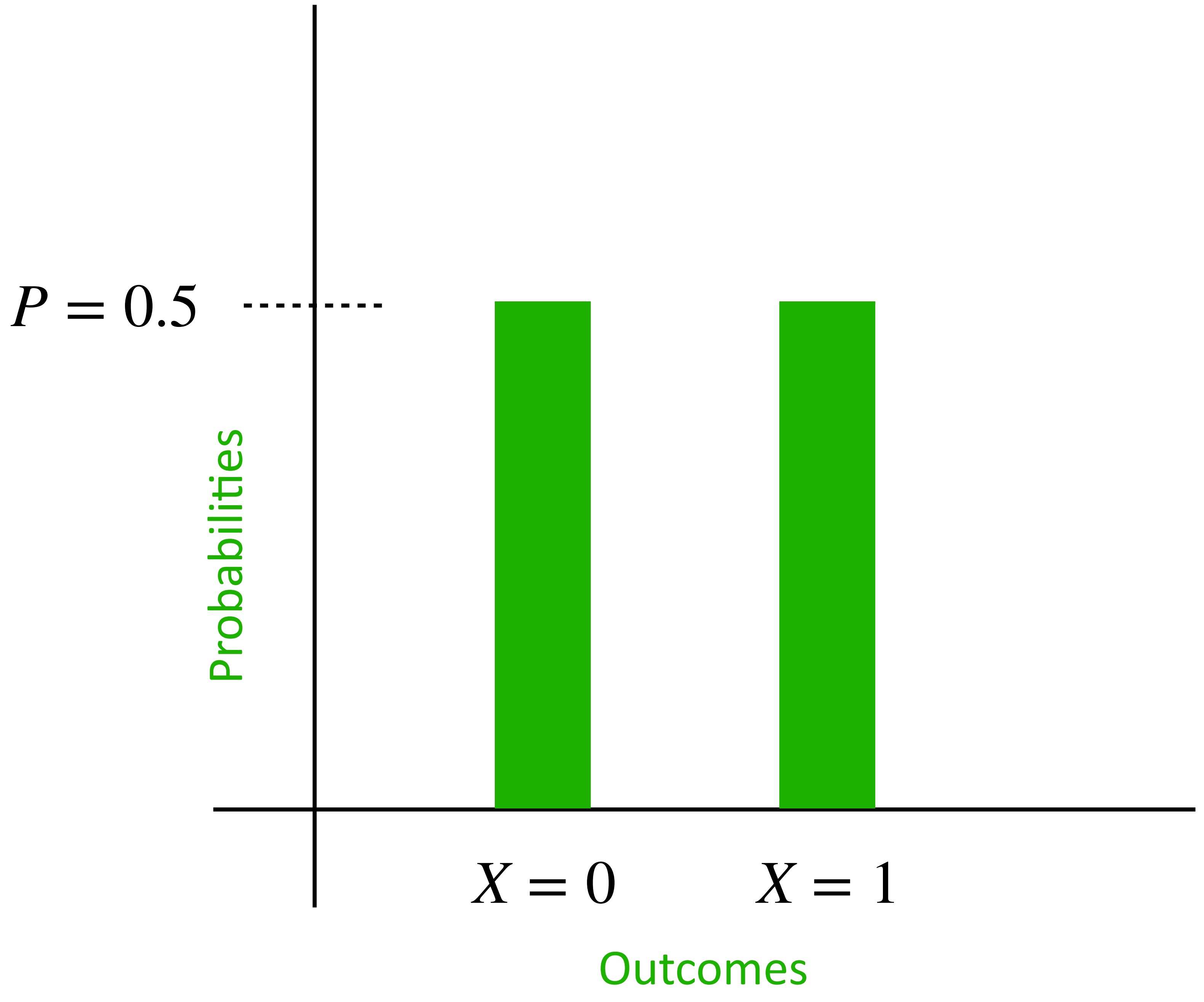
Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{cases}$$

$$P(X = 1) = \frac{1}{2} = 0.5$$

$$P(X = 0) = \frac{1}{2} = 0.5$$



We can plot these probabilities on a graph

Probability Mass Function

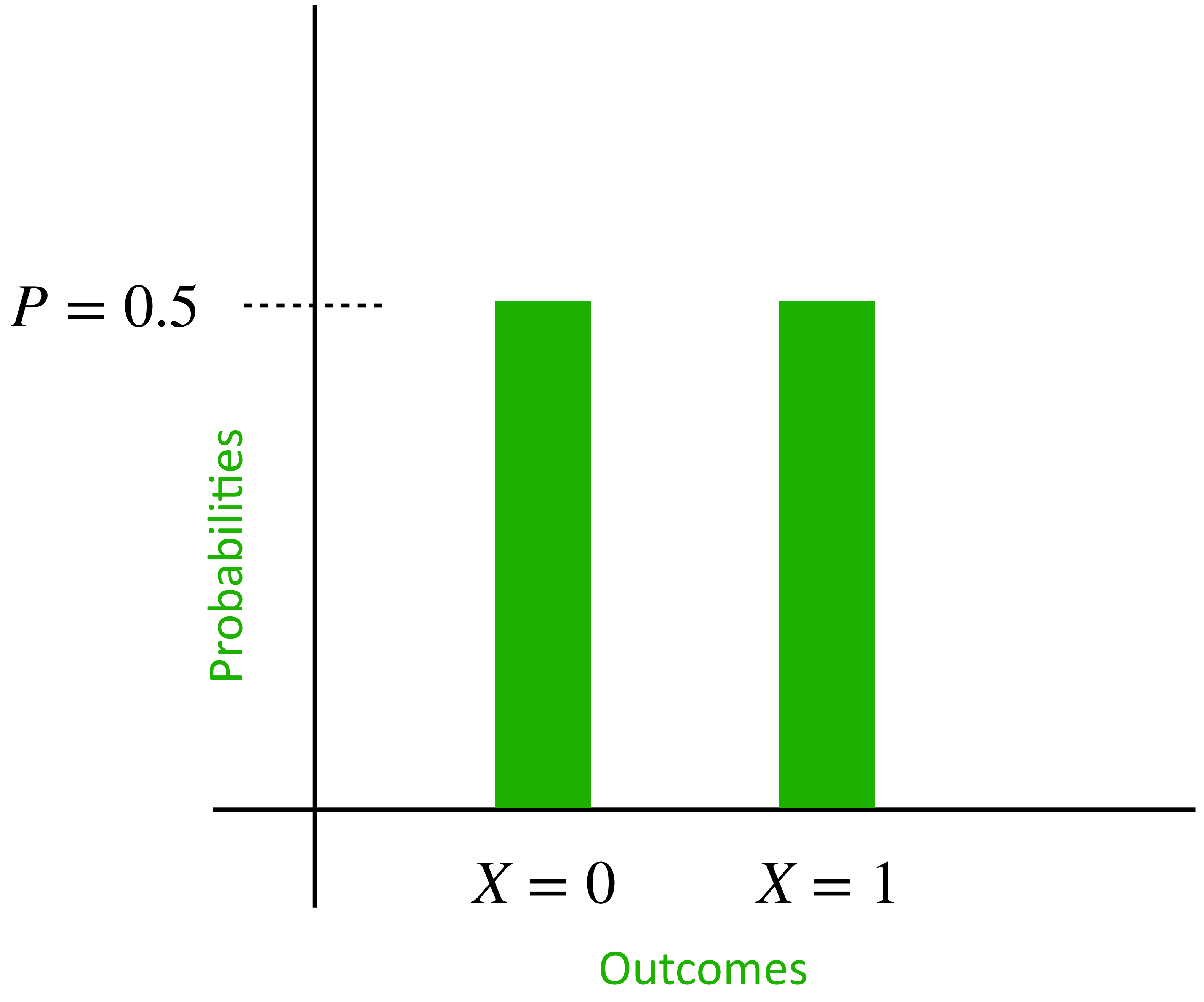
Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{Bmatrix} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{Bmatrix}$$



Probability Mass Function

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

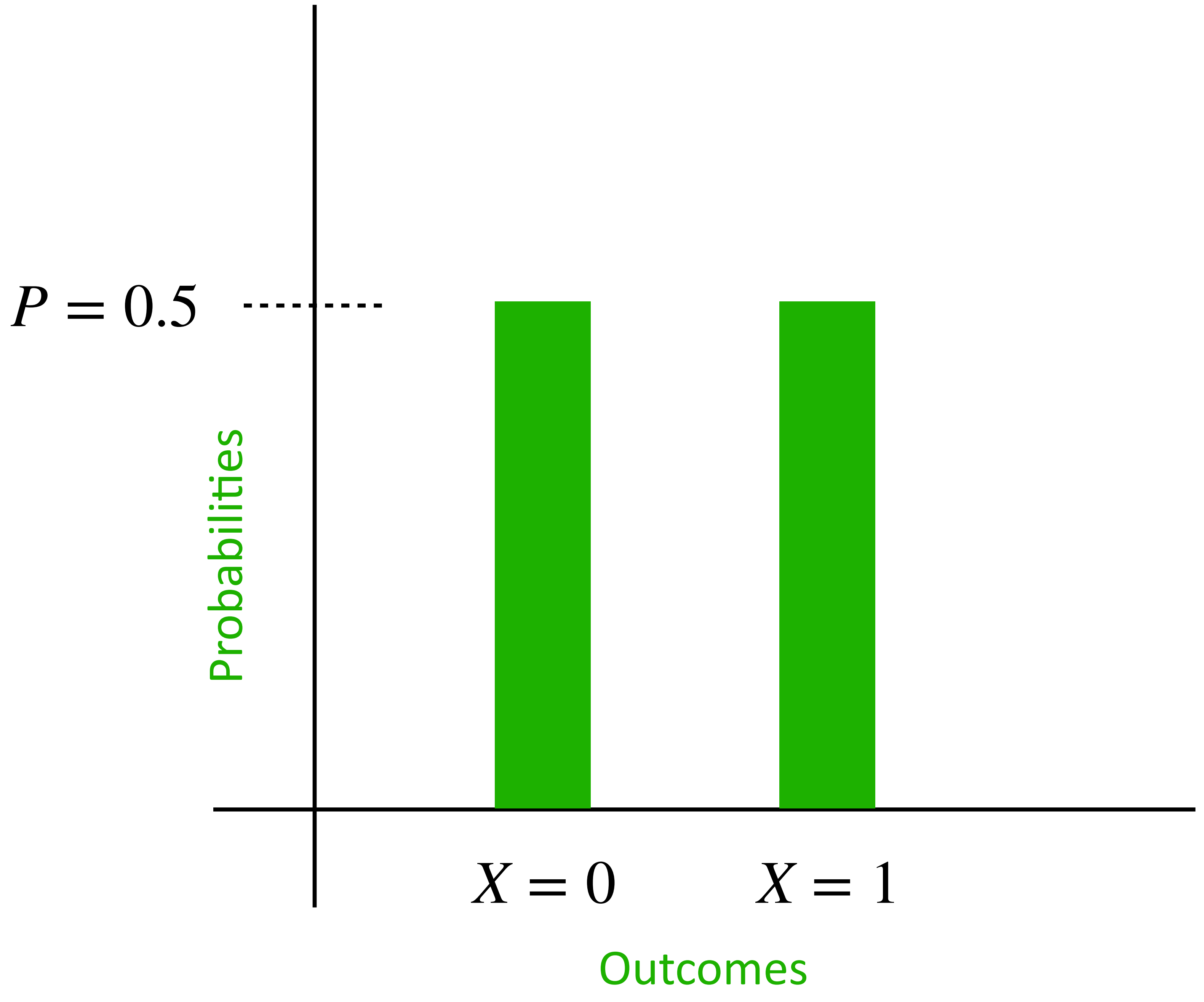
Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\,\alpha = Heads \\ 0, & if\,\alpha = Tails \end{cases}$$

$$P(X = 1) = \frac{1}{2} = 0.5$$

$$P(X = 0) = \frac{1}{2} = 0.5$$



Probability Mass Function

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

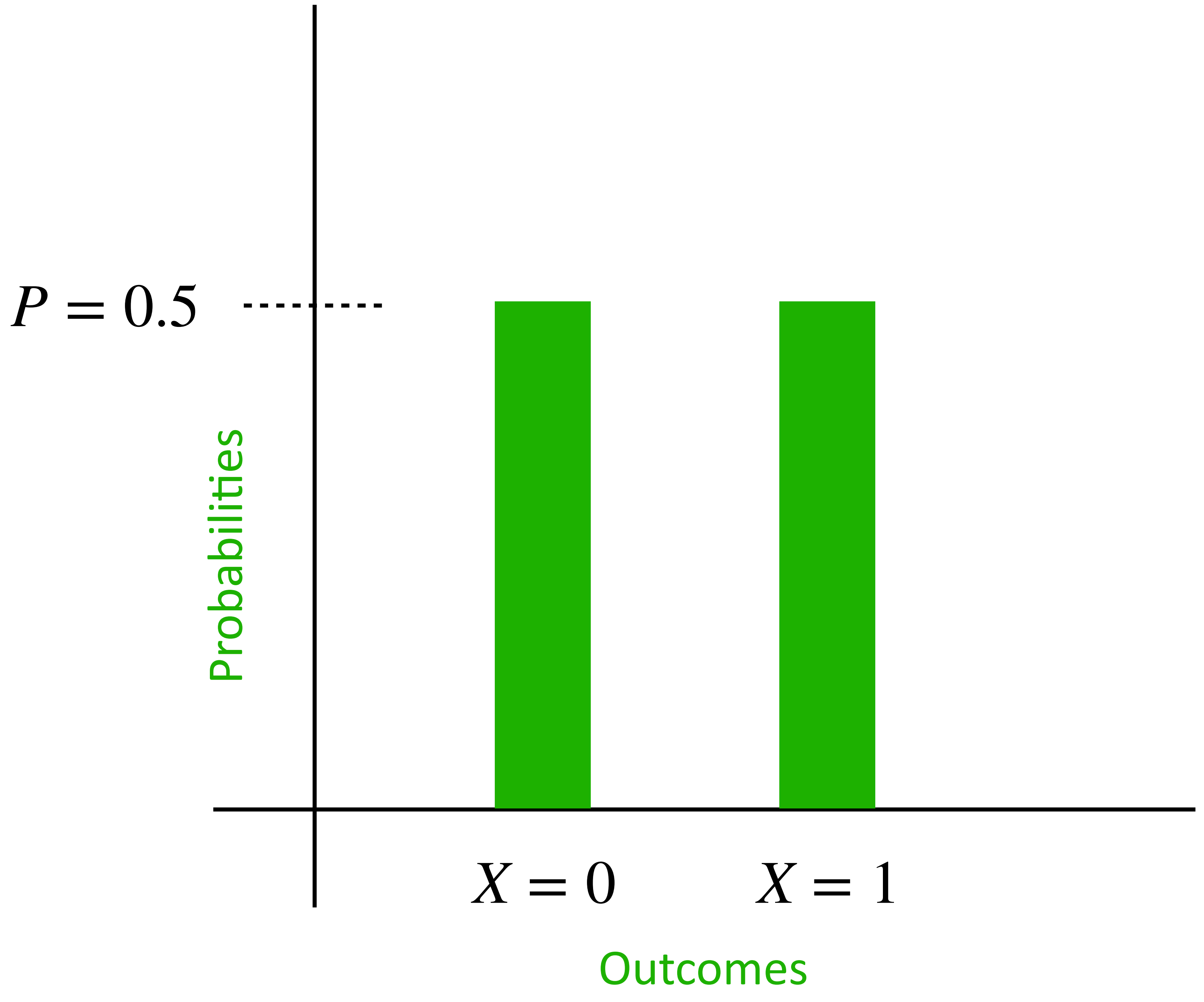
Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{cases}$$

$$P(X = 1) = \frac{1}{2} = 0.5$$

$$P(X = 0) = \frac{1}{2} = 0.5$$



Probability Mass Function

Flipping a fair coin

$$f(x) = \begin{cases} 0.5, & if\, x = 1 \\ 0.5, & if\, x = 0 \end{cases}$$

Probability Mass Function

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{cases}$$

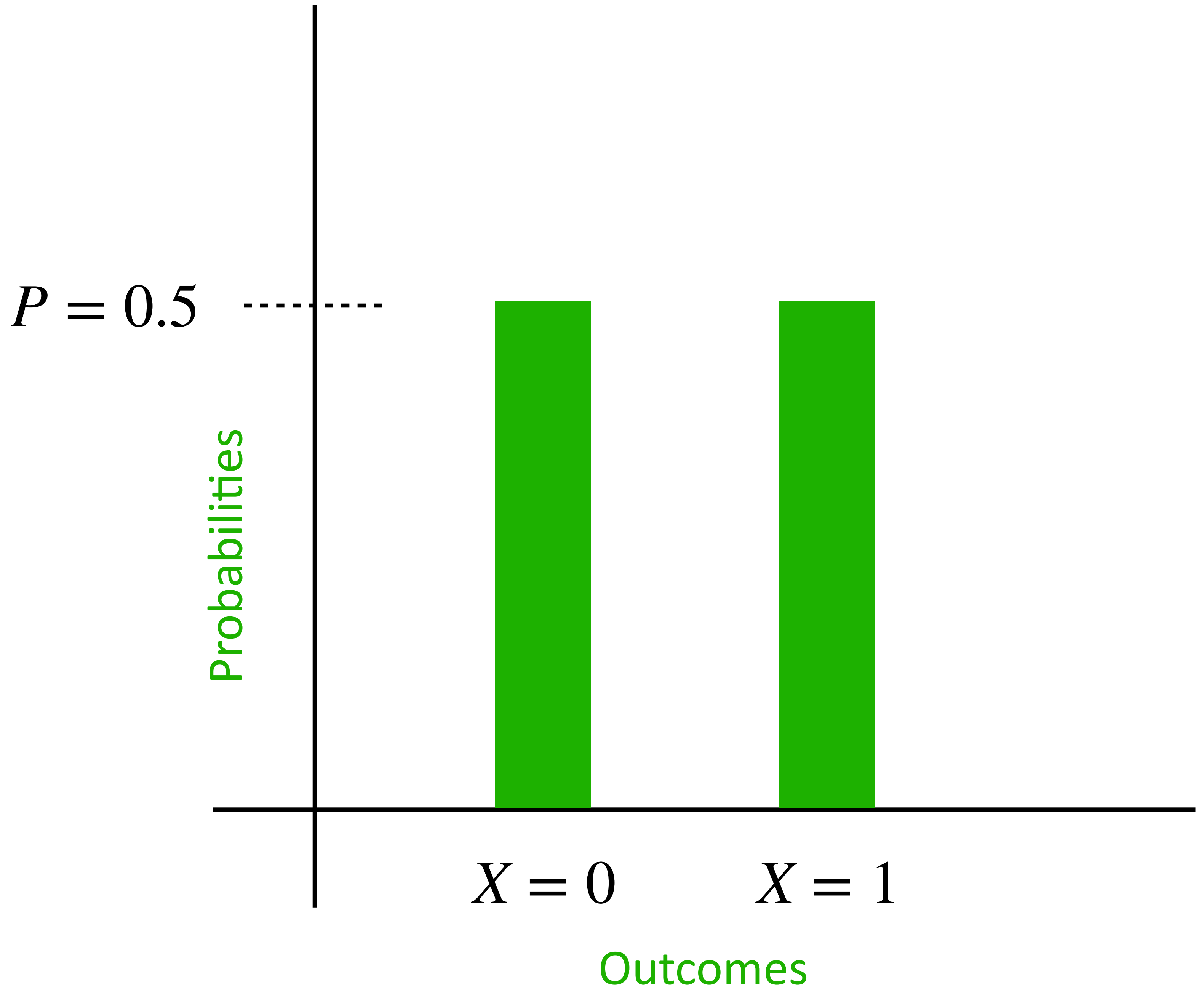
$$P(X = 1) = \frac{1}{2} = 0.5$$

$$P(X = 0) = \frac{1}{2} = 0.5$$

Probability Mass Function

Flipping a fair coin

$$f(x) = \begin{cases} 0.5, & if\, x = 1 \\ 0.5, & if\, x = 0 \end{cases}$$



This is the Probability Mass Function

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

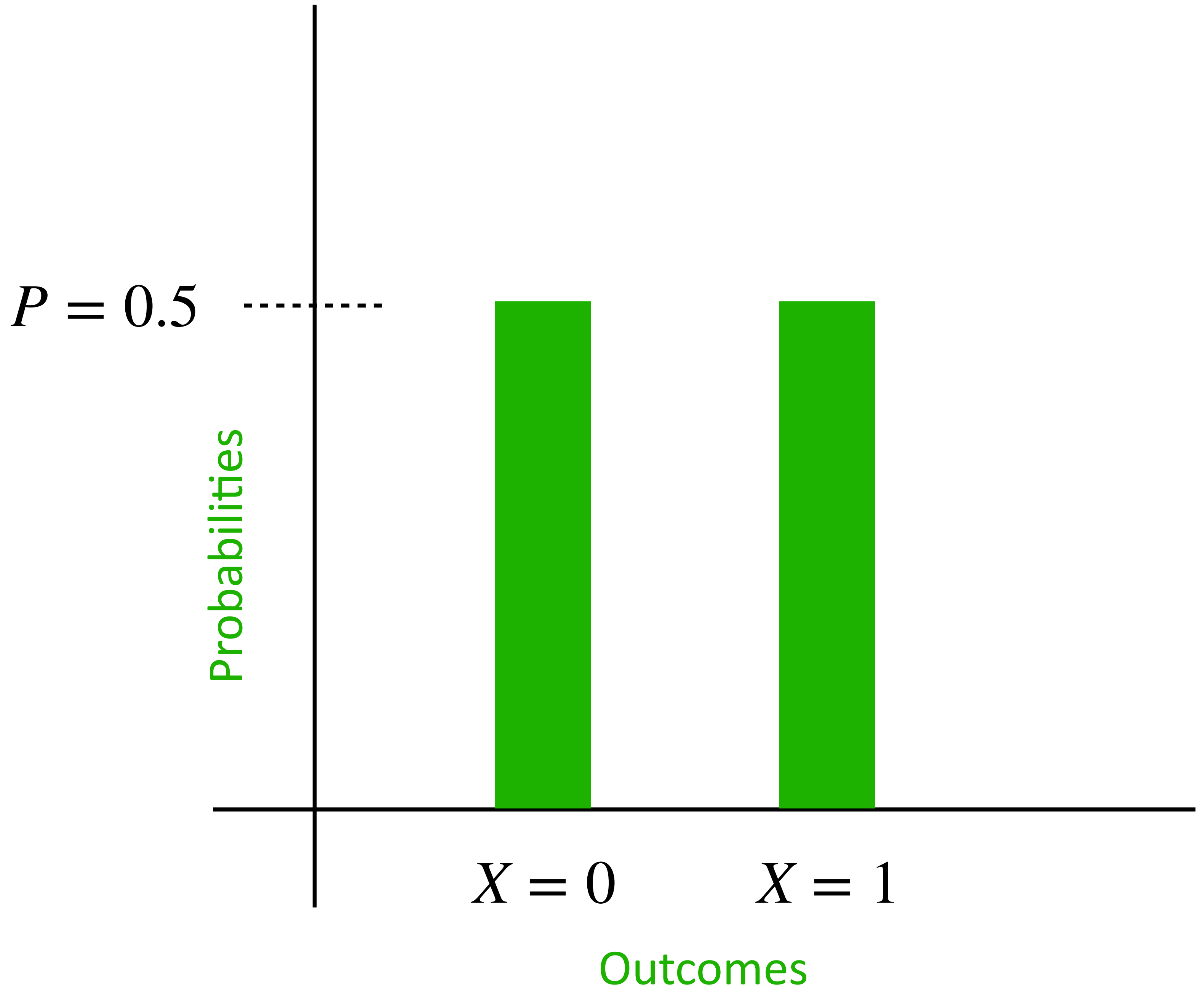
Mathematical Representation

Flipping a fair coin

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = Heads \\ 0, & if\, \alpha = Tails \end{cases}$$

$$P(X=1) = \frac{1}{2} = 0.5$$

$$P(X=0) = \frac{1}{2} = 0.5$$



Probability Mass Function

Flipping a fair coin

$$f(x) = \begin{cases} 0.5, & if\, x = 1 \\ 0.5, & if\, x = 0 \end{cases}$$

The Probability Mass Function gives the probability that a discrete random variable is equal to some value

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Rolling a six sided fair die

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Rolling a six sided fair die

$$X(\alpha) = \begin{Bmatrix} 1, & if\,\alpha = One \\ 2, & if\,\alpha = Two \\ 3, & if\,\alpha = Three \\ 4, & if\,\alpha = Four \\ 5, & if\,\alpha = Five \\ 6, & if\,\alpha = Six \end{Bmatrix}$$

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Rolling a six sided fair die

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = One \\ 2, & if\, \alpha = Two \\ 3, & if\, \alpha = Three \\ 4, & if\, \alpha = Four \\ 5, & if\, \alpha = Five \\ 6, & if\, \alpha = Six \end{cases}$$

$$P(X = 1) = \frac{1}{6} = 0.1667$$

$$P(X = 2) = \frac{1}{6} = 0.1667$$

$$P(X = 3) = \frac{1}{6} = 0.1667$$

$$P(X = 4) = \frac{1}{6} = 0.1667$$

$$P(X = 5) = \frac{1}{6} = 0.1667$$

$$P(X = 6) = \frac{1}{6} = 0.1667$$

Probability of an Event Occurring

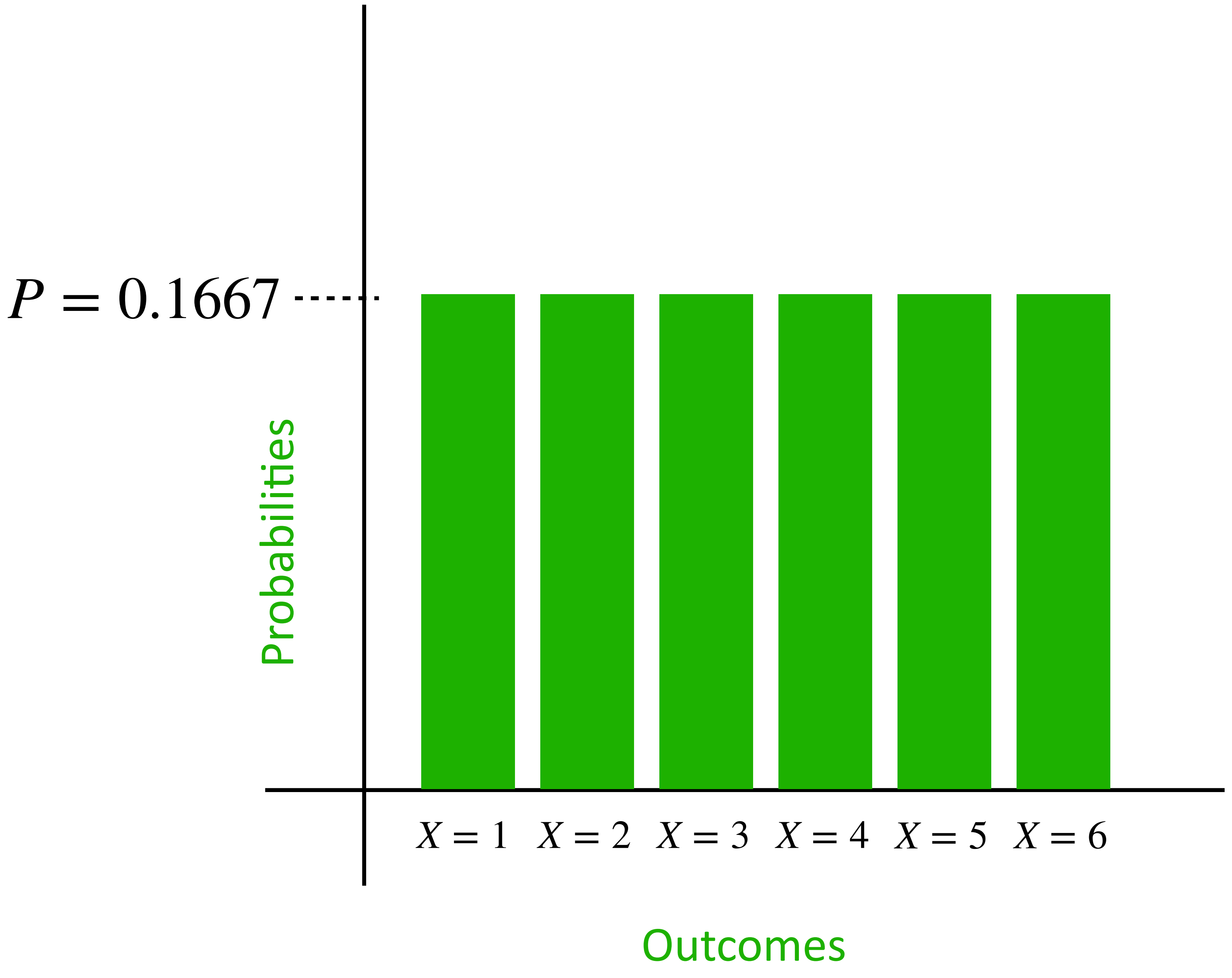
The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Rolling a six sided fair die

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = One \\ 2, & if\, \alpha = Two \\ 3, & if\, \alpha = Three \\ 4, & if\, \alpha = Four \\ 5, & if\, \alpha = Five \\ 6, & if\, \alpha = Six \end{cases}$$

$$P(X = 1) = \frac{1}{6} = 0.1667$$
$$P(X = 2) = \frac{1}{6} = 0.1667$$
$$P(X = 3) = \frac{1}{6} = 0.1667$$
$$P(X = 4) = \frac{1}{6} = 0.1667$$
$$P(X = 5) = \frac{1}{6} = 0.1667$$
$$P(X = 6) = \frac{1}{6} = 0.1667$$



Probability Mass Function

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

Rolling a six sided fair die

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = One \\ 2, & if\, \alpha = Two \\ 3, & if\, \alpha = Three \\ 4, & if\, \alpha = Four \\ 5, & if\, \alpha = Five \\ 6, & if\, \alpha = Six \end{cases}$$

$$P(X = 1) = \frac{1}{6} = 0.1667$$

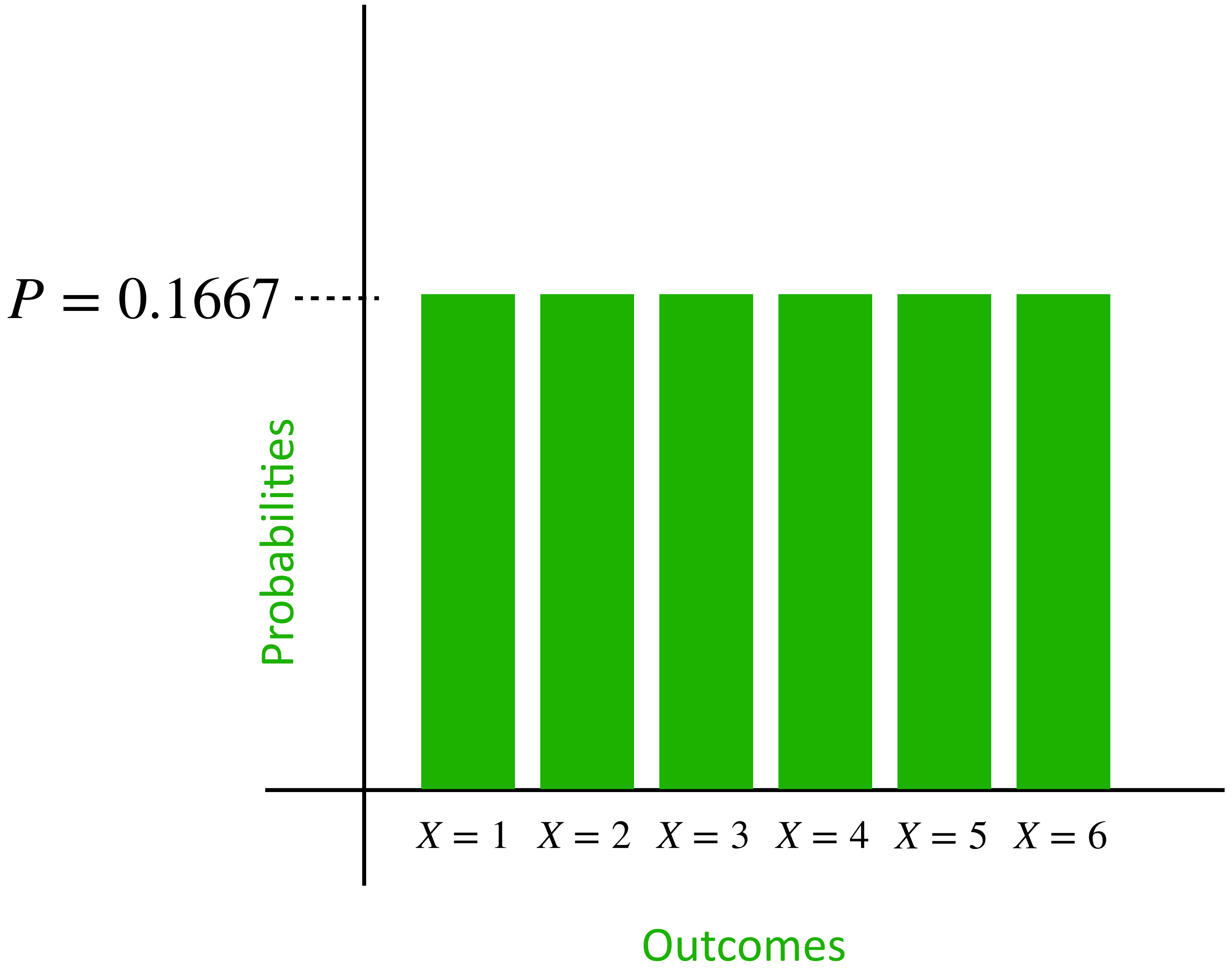
$$P(X = 2) = \frac{1}{6} = 0.1667$$

$$P(X = 3) = \frac{1}{6} = 0.1667$$

$$P(X = 4) = \frac{1}{6} = 0.1667$$

$$P(X = 5) = \frac{1}{6} = 0.1667$$

$$P(X = 6) = \frac{1}{6} = 0.1667$$



The Probability Mass Function gives the probability that a discrete random variable is equal to some value

Probability Mass Function

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

Mathematical Representation

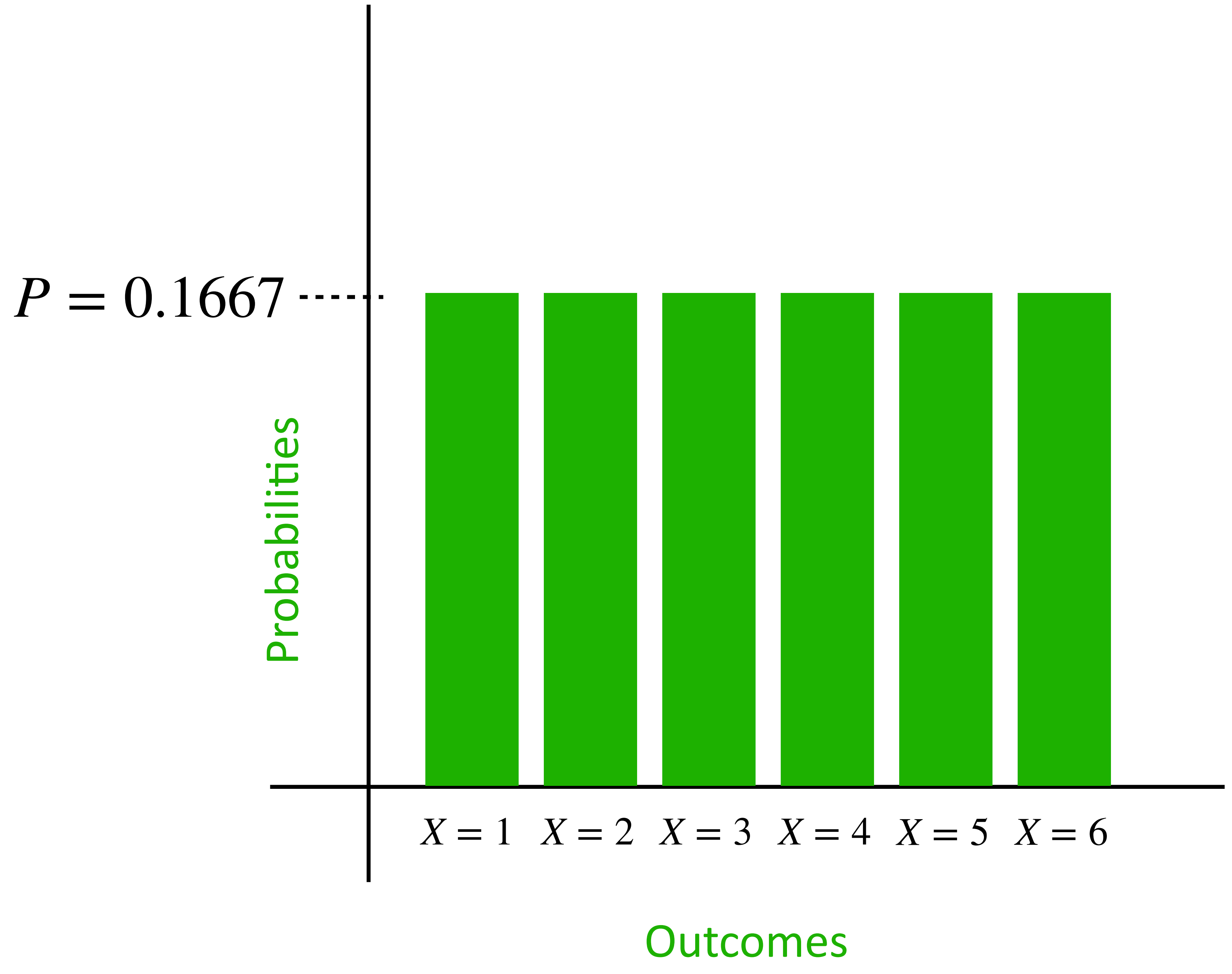
Rolling a six sided fair die

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = One \\ 2, & if\, \alpha = Two \\ 3, & if\, \alpha = Three \\ 4, & if\, \alpha = Four \\ 5, & if\, \alpha = Five \\ 6, & if\, \alpha = Six \end{cases}$$

Probability Mass Function

Rolling a six sided fair die

$$f(x) = \begin{cases} 0.1667, & if\, x = 1 \\ 0.1667, & if\, x = 2 \\ 0.1667, & if\, x = 3 \\ 0.1667, & if\, x = 4 \\ 0.1667, & if\, x = 5 \\ 0.1667, & if\, x = 6 \end{cases}$$



The Probability Mass Function gives the probability that a discrete random variable is equal to some value

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

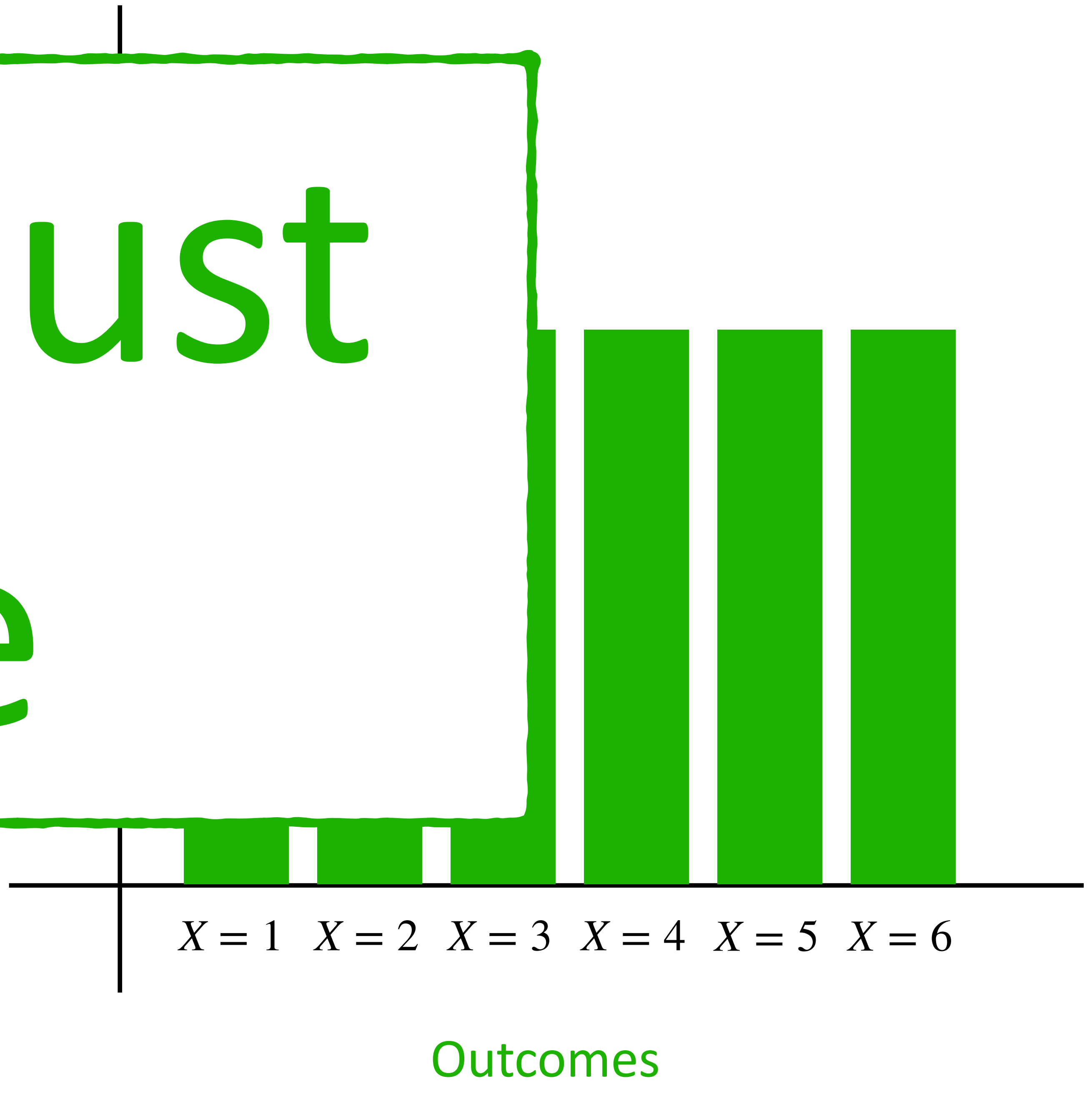
Mathematical Representation

Rolling a six

$$X(\alpha) = \begin{cases} 1, & \\ 2, & \\ 3, & \\ 4, & \\ 5, & if\, \alpha = Five \\ 6, & if\, \alpha = Six \end{cases}$$

Probability Mass Function

$$\begin{cases} 0.1667, & if\, x = 5 \\ 0.1667, & if\, x = 6 \end{cases}$$



The Probabilities must all add up to one

The Probability Mass Function gives the probability that a discrete random variable is equal to some value

Cumulative Distribution Function

Mathematical Representation

Rolling a six sided fair die

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = One \\ 2, & if\, \alpha = Two \\ 3, & if\, \alpha = Three \\ 4, & if\, \alpha = Four \\ 5, & if\, \alpha = Five \\ 6, & if\, \alpha = Six \end{cases}$$

Probability Mass Function

Rolling a six sided fair die

$$f(x) = \begin{cases} 0.1667, & if\, x = 1 \\ 0.1667, & if\, x = 2 \\ 0.1667, & if\, x = 3 \\ 0.1667, & if\, x = 4 \\ 0.1667, & if\, x = 5 \\ 0.1667, & if\, x = 6 \end{cases}$$

Cumulative Distribution Function

Mathematical Representation

Rolling a six sided fair die

$$X(\alpha) = \begin{cases} 1, & if\, \alpha = One \\ 2, & if\, \alpha = Two \\ 3, & if\, \alpha = Three \\ 4, & if\, \alpha = Four \\ 5, & if\, \alpha = Five \\ 6, & if\, \alpha = Six \end{cases}$$

Probability Mass Function

Rolling a six sided fair die

$$f(x) = \begin{cases} 0.1667, & if\, x = 1 \\ 0.1667, & if\, x = 2 \\ 0.1667, & if\, x = 3 \\ 0.1667, & if\, x = 4 \\ 0.1667, & if\, x = 5 \\ 0.1667, & if\, x = 6 \end{cases}$$

Probability of rolling a 2 or less:

$$P(X = 1) + P(X = 2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} = 0.3333$$

Cumulative Distribution Function

Mathematical Representation

Rolling a six sided fair die

$$X(\alpha) = \begin{Bmatrix} 1, & if\, \alpha = One \\ 2, & if\, \alpha = Two \\ 3, & if\, \alpha = Three \\ 4, & if\, \alpha = Four \\ 5, & if\, \alpha = Five \\ 6, & if\, \alpha = Six \end{Bmatrix}$$

Probability Mass Function

Rolling a six sided fair die

$$f(x) = \begin{Bmatrix} 0.1667, & if\, x = 1 \\ 0.1667, & if\, x = 2 \\ 0.1667, & if\, x = 3 \\ 0.1667, & if\, x = 4 \\ 0.1667, & if\, x = 5 \\ 0.1667, & if\, x = 6 \end{Bmatrix}$$

Probability of rolling a 2 or less:

$$P(X=1) + P(X=2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} = 0.3333$$

Probability of rolling a 6 or less:

$$P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6} = 1$$

Cumulative Distribution Function

Mathematical Representation

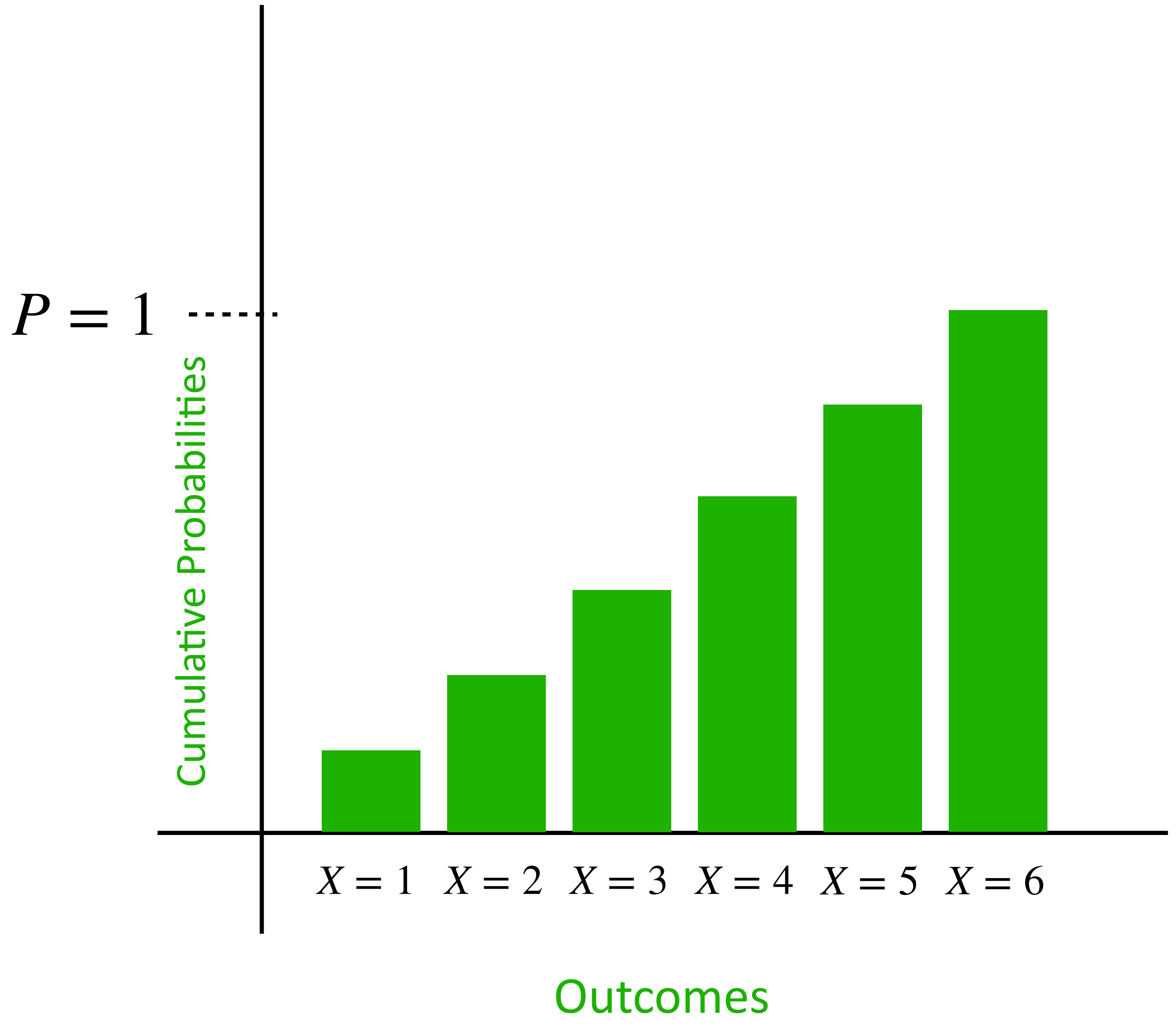
Rolling a six sided fair die

$$X(\alpha) = \begin{Bmatrix} 1, & if\,\alpha = One \\ 2, & if\,\alpha = Two \\ 3, & if\,\alpha = Three \\ 4, & if\,\alpha = Four \\ 5, & if\,\alpha = Five \\ 6, & if\,\alpha = Six \end{Bmatrix}$$

Probability Mass Function

Rolling a six sided fair die

$$f(x) = \begin{Bmatrix} 0.1667, & if\,x = 1 \\ 0.1667, & if\,x = 2 \\ 0.1667, & if\,x = 3 \\ 0.1667, & if\,x = 4 \\ 0.1667, & if\,x = 5 \\ 0.1667, & if\,x = 6 \end{Bmatrix}$$



Probability of rolling a 2 or less:

$$P(X = 1) + P(X = 2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} = 0.3333$$

Probability of rolling a 6 or less:

$$P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6} = 1$$

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Infinite Outcomes: Height can be any value from 0 to infinity.

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Infinite Outcomes: Height can be any value from 0 to infinity.

A Continuous Random Variable doesn't have a Probability Mass Function because there are infinite outcomes.

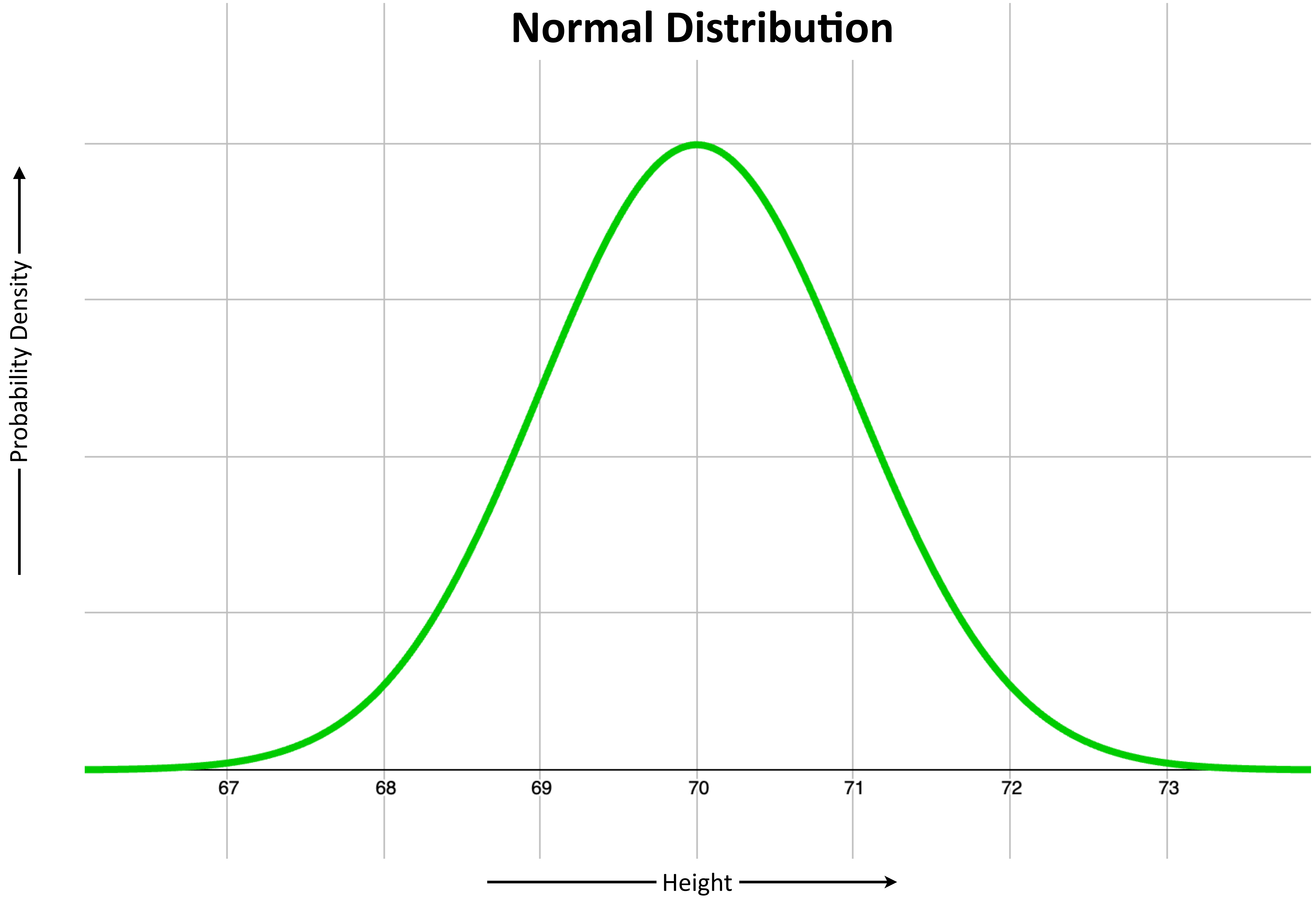
Instead, a continuous random variable has a Probability Density Function (PDF)

Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria



The Probability Density Function returns the probability of an outcome within a certain range

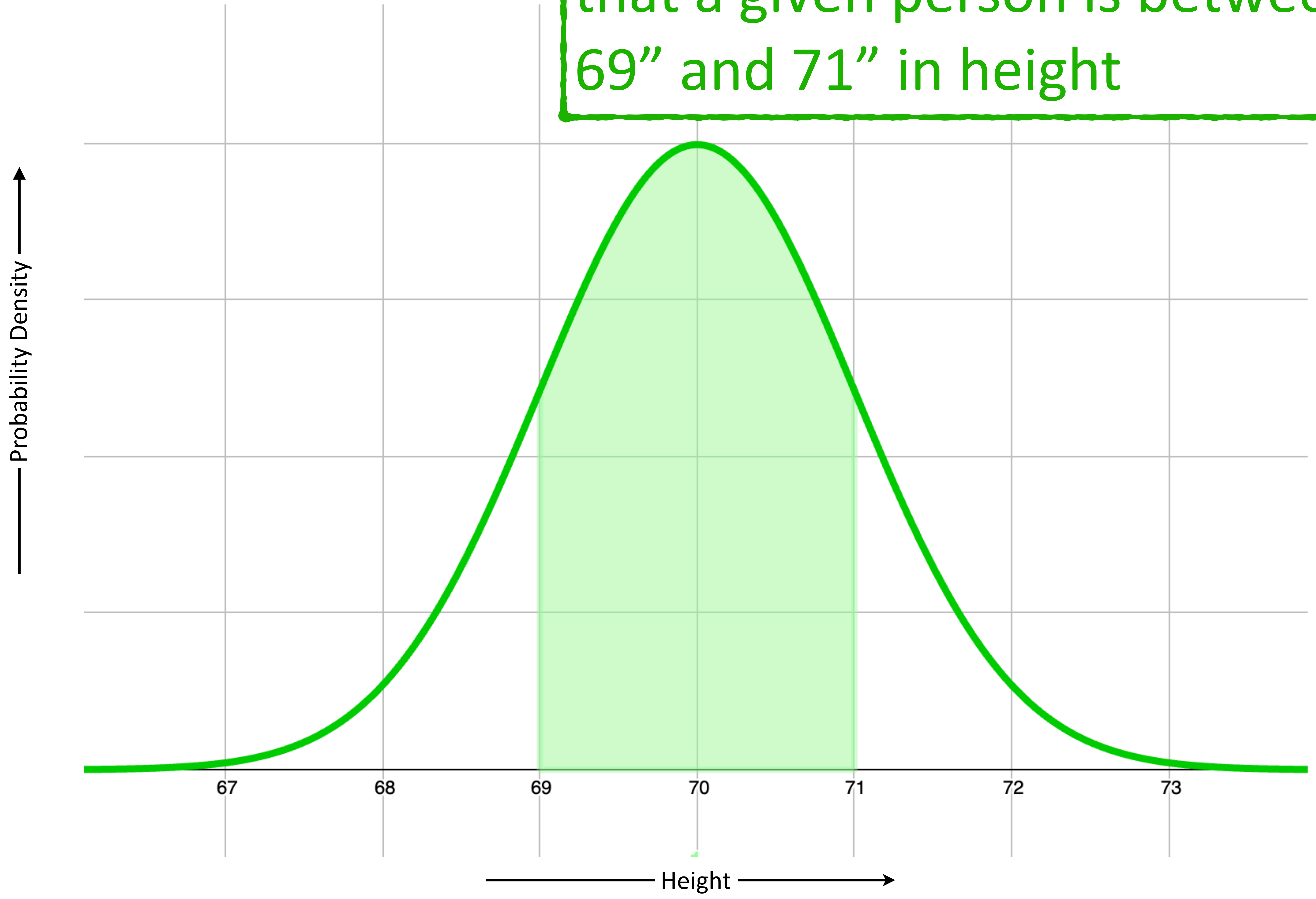
Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Area under the curve: Probability that a given person is between 69” and 71” in height



The Probability Density Function returns the probability of an outcome within a certain range

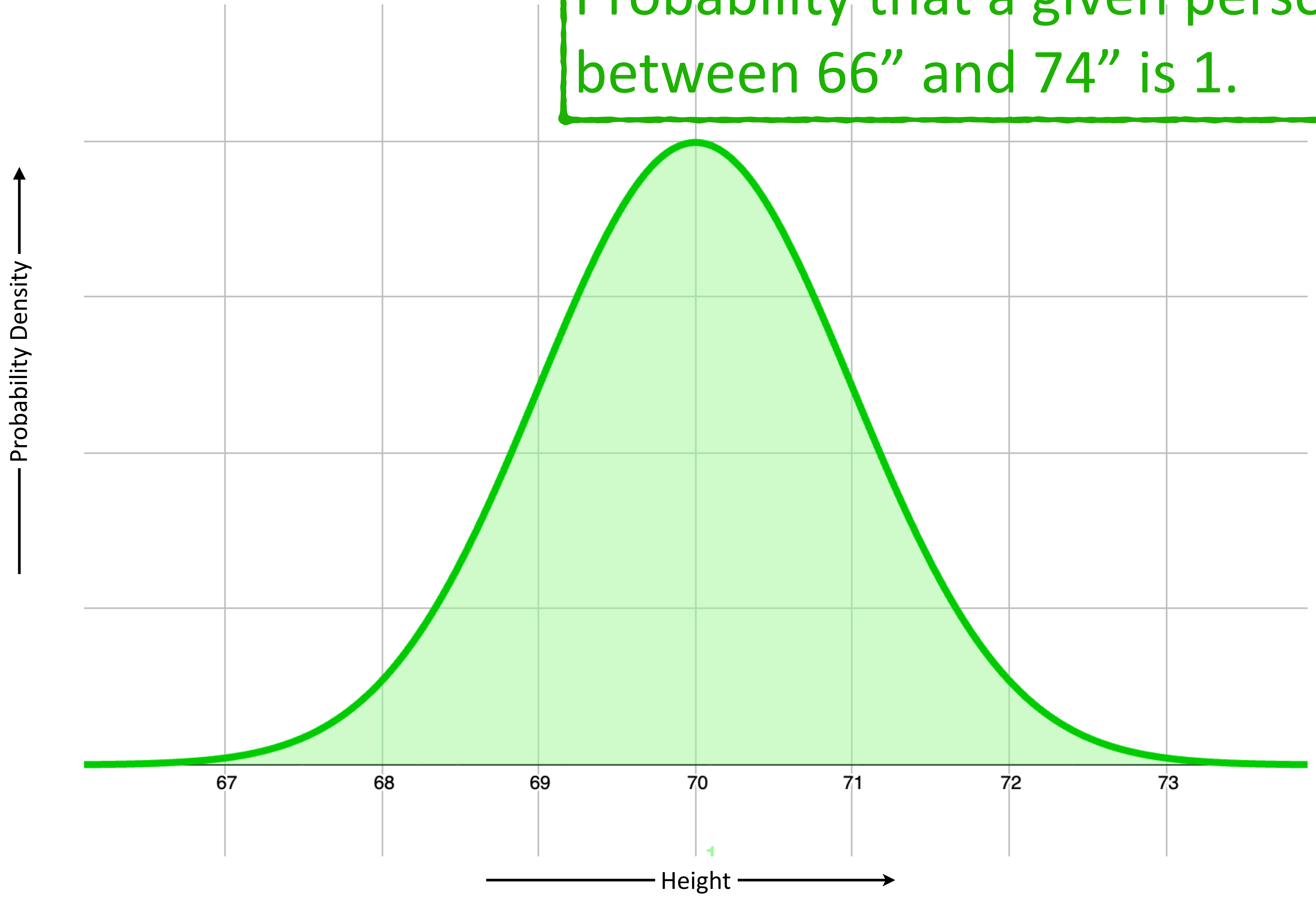
Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Total Area under the curve:
Probability that a given person is between 66” and 74” is 1.



The Probability Density Function returns the probability of an outcome within a certain range

Probability Density Function

Continuous Random Variable

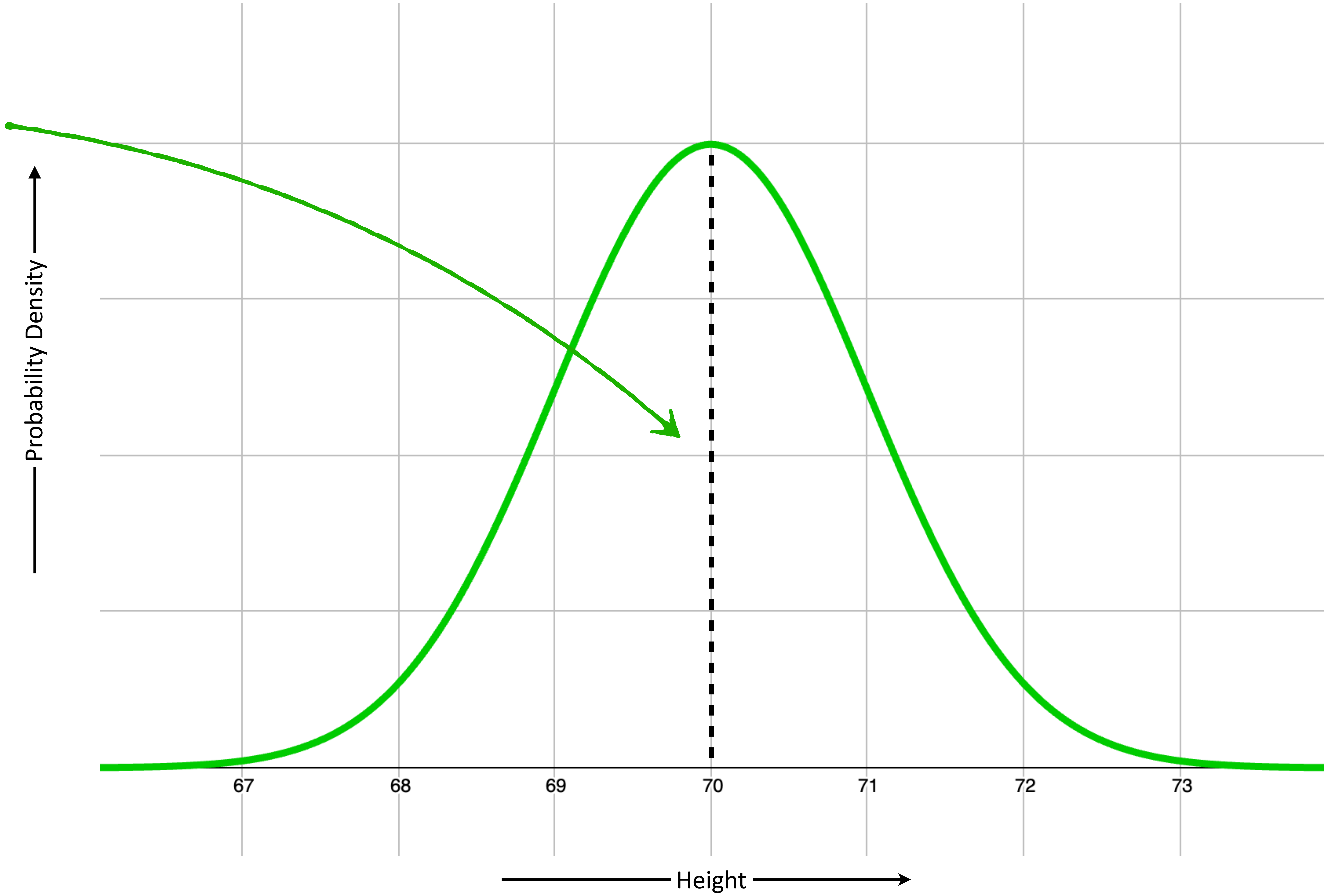
A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Mean (μ) of the distribution

The Average Height of a citizen of Bulgaria is 70”

The Probability Density Function returns the probability of an outcome within a certain range

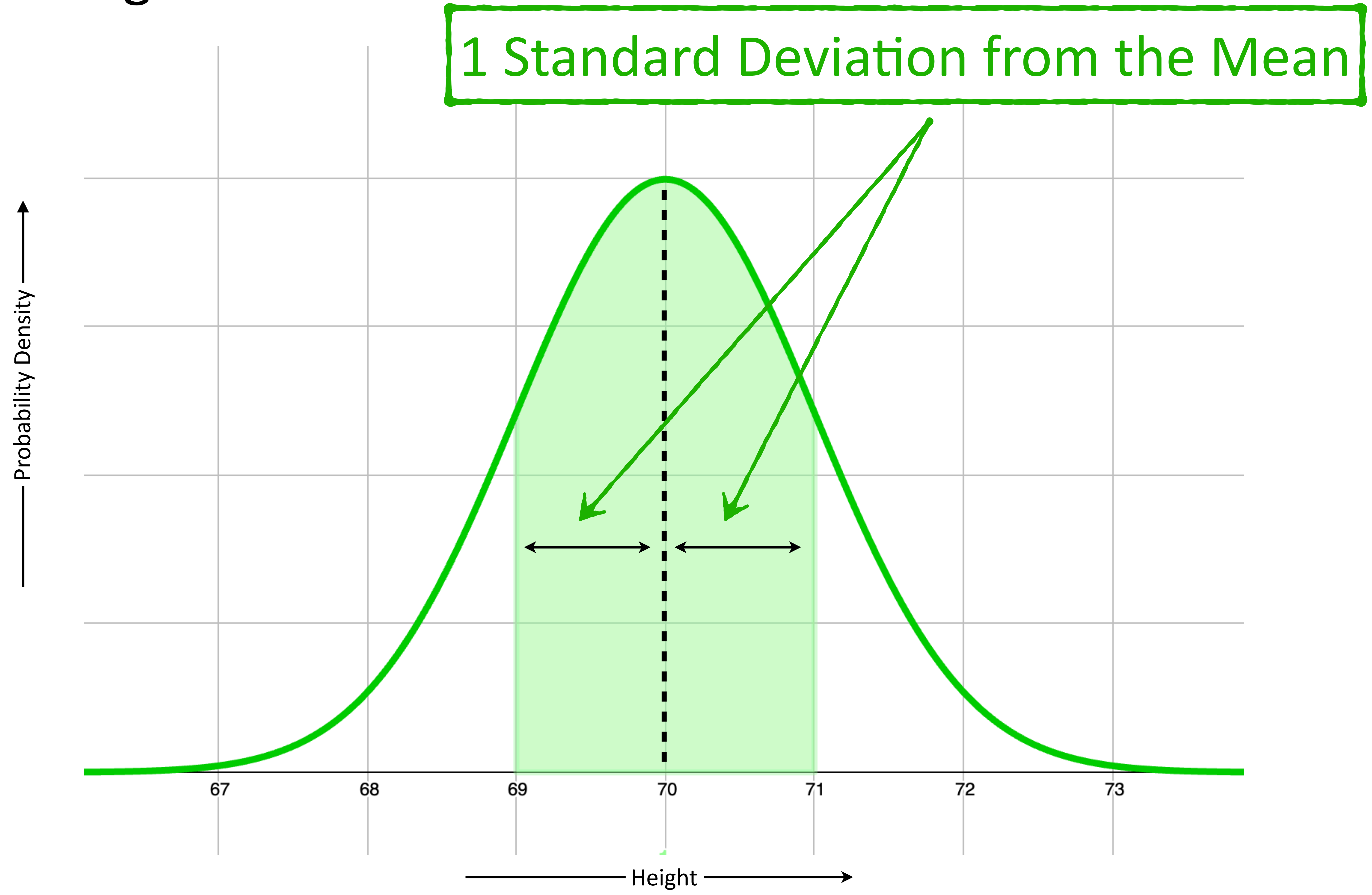


Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria



The Probability Density Function returns the probability of an outcome within a certain range

Probability Density Function

Continuous Random Variable

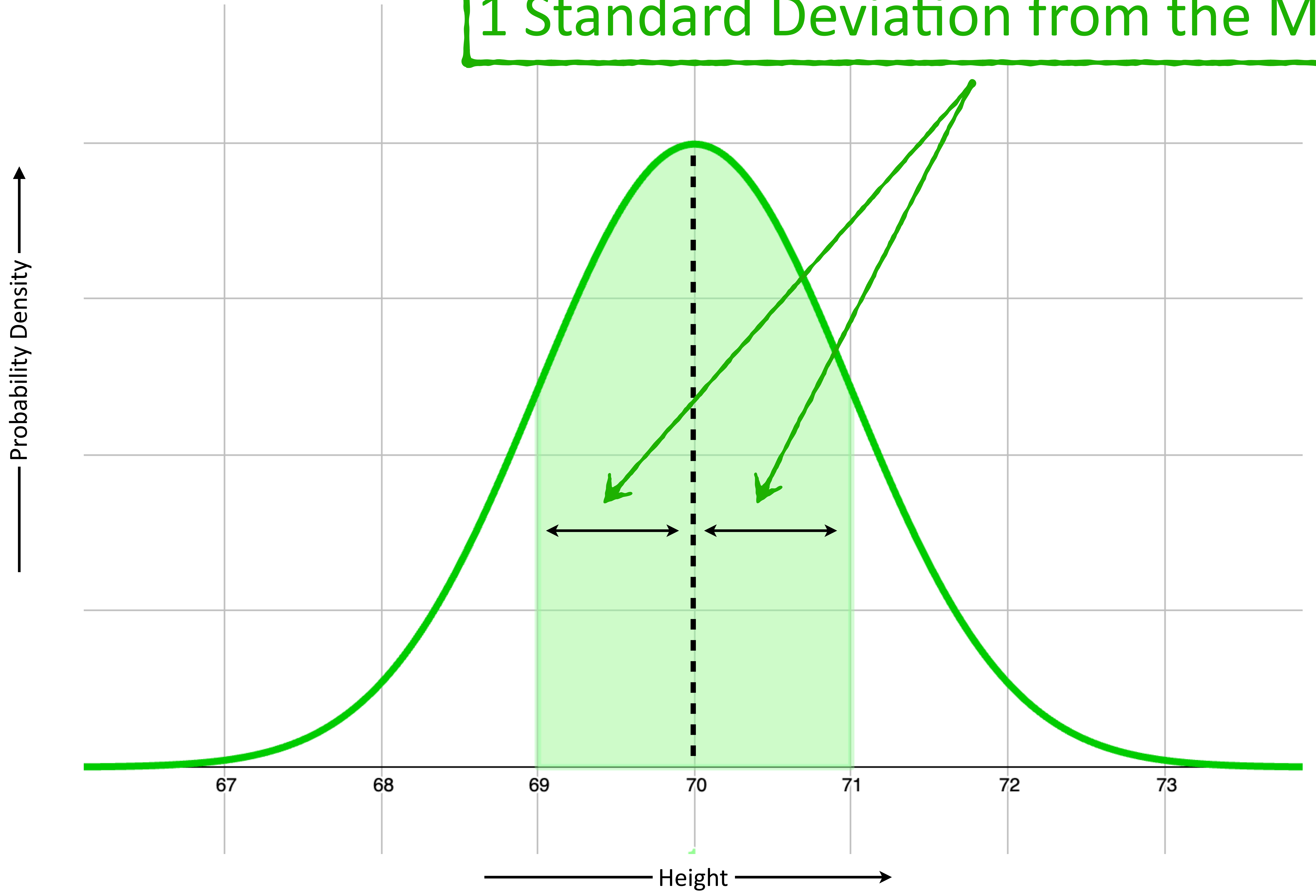
A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

1 Standard Deviation from the Mean

Standard Deviation (σ): A measure of distance of the data from the mean.

The Probability Density Function returns the probability of an outcome within a certain range



Probability Density Function

Continuous Random Variable

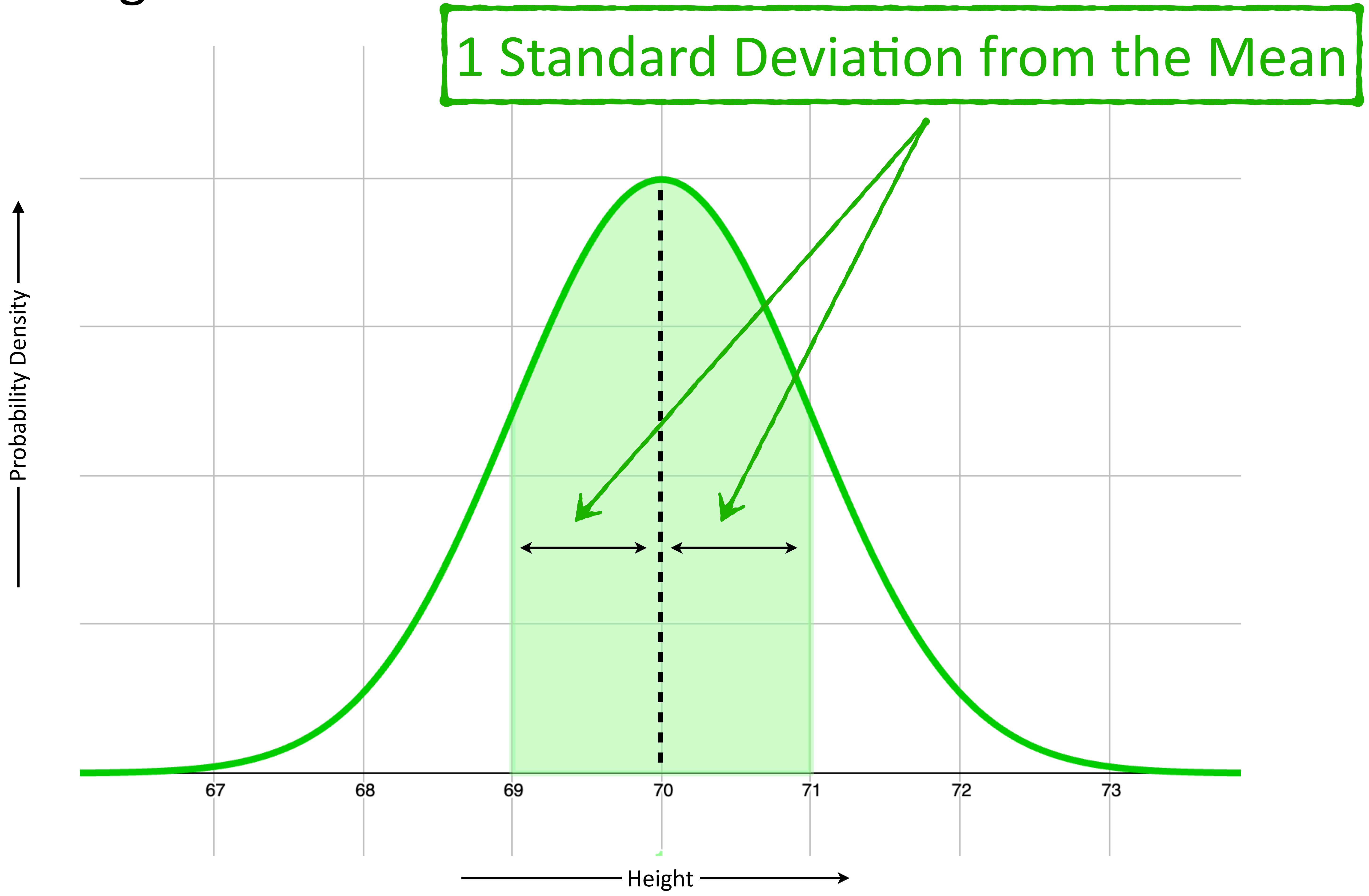
A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Variance (σ^2) is the square of the Standard Deviation (σ)

Standard Deviation (σ): A measure of distance of the data from the mean.

The Probability Density Function returns the probability of an outcome within a certain range



Probability Density Function

Continuous Random Variable

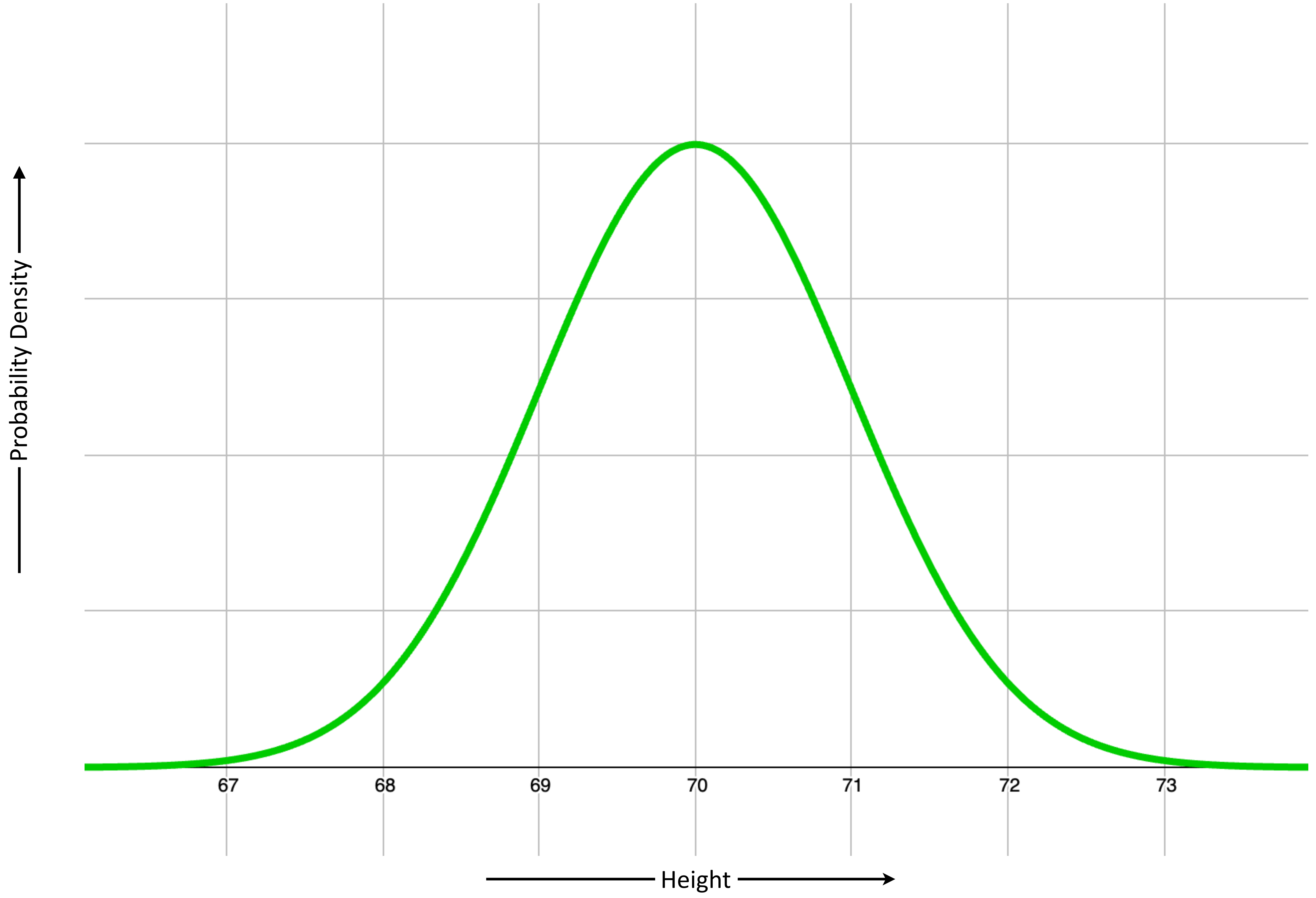
A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Mean = μ

Variance = σ^2

Mean & Variance are the two parameters that determine the shape and position the normal distribution



Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

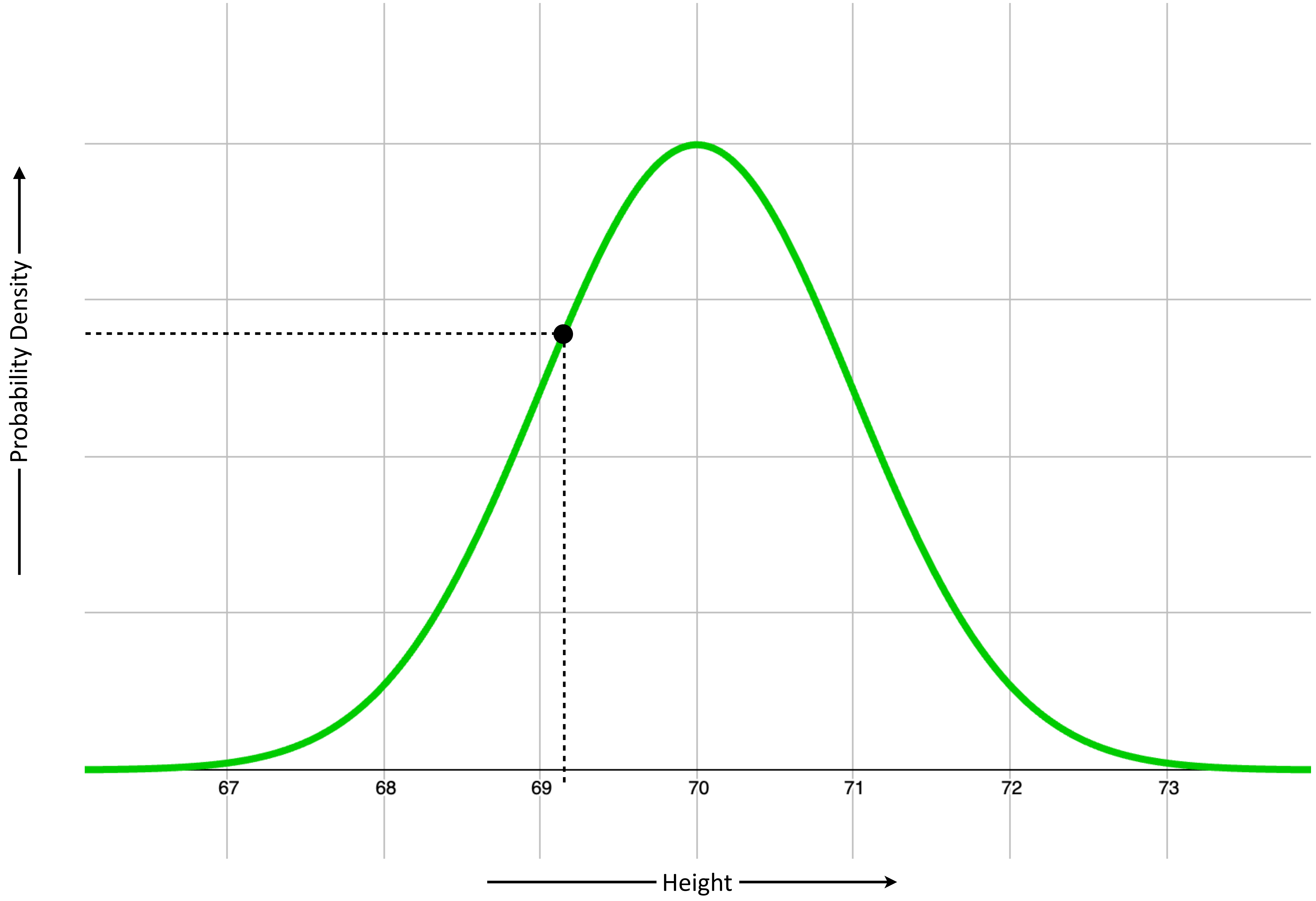
Example: X is the height of a citizen of Bulgaria

Mean = μ

Variance = σ^2

Likelihood is not Probability

A single point on the curve:
Likelihood of observing a person of that specific height for this specific distribution ($\mu = 70, \sigma^2 = 1$)



Probability Density Function

Continuous Random Variable

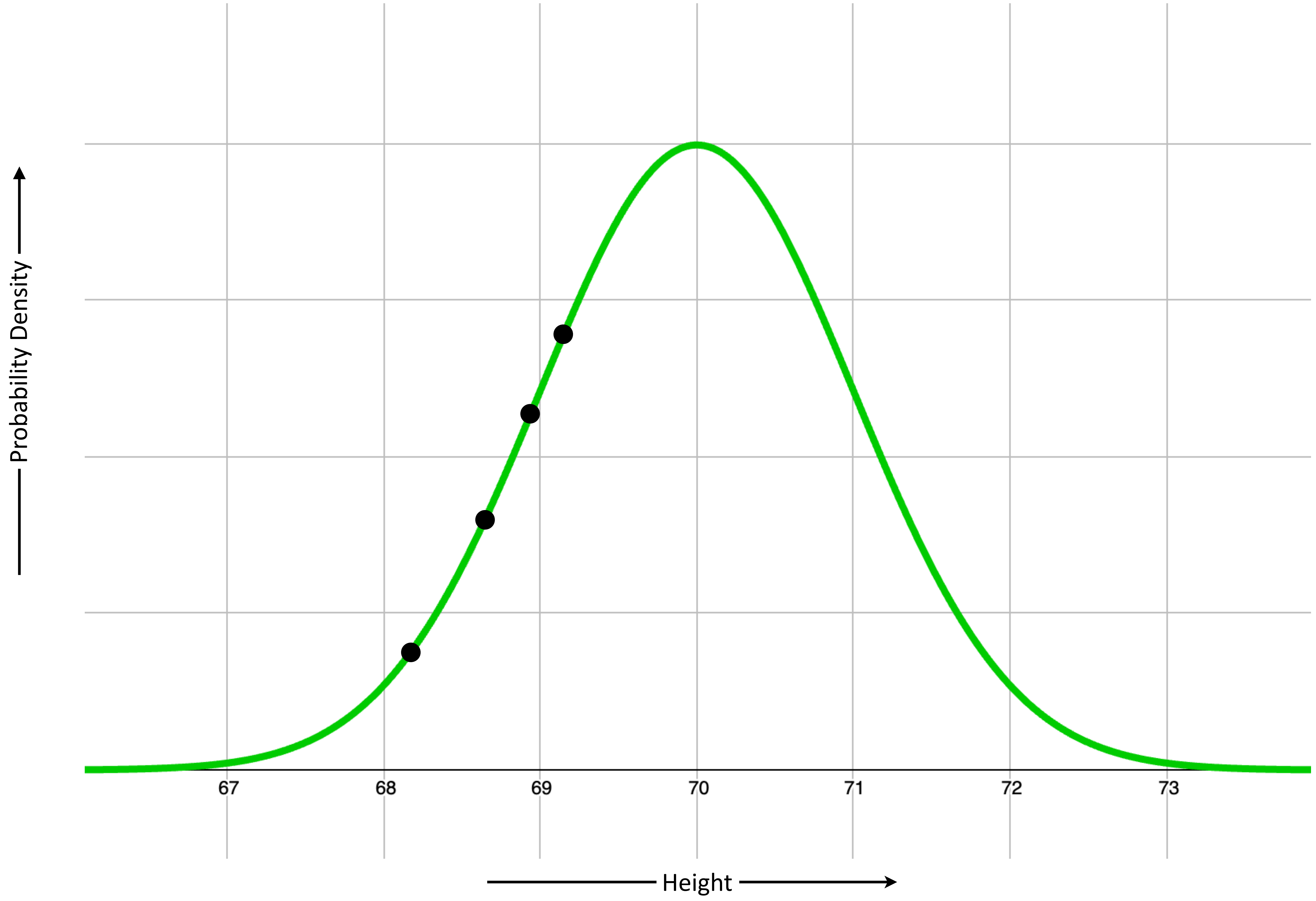
A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Mean = μ

Variance = σ^2

Multiple points on the curve:
Multiply the likelihoods of every point



Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

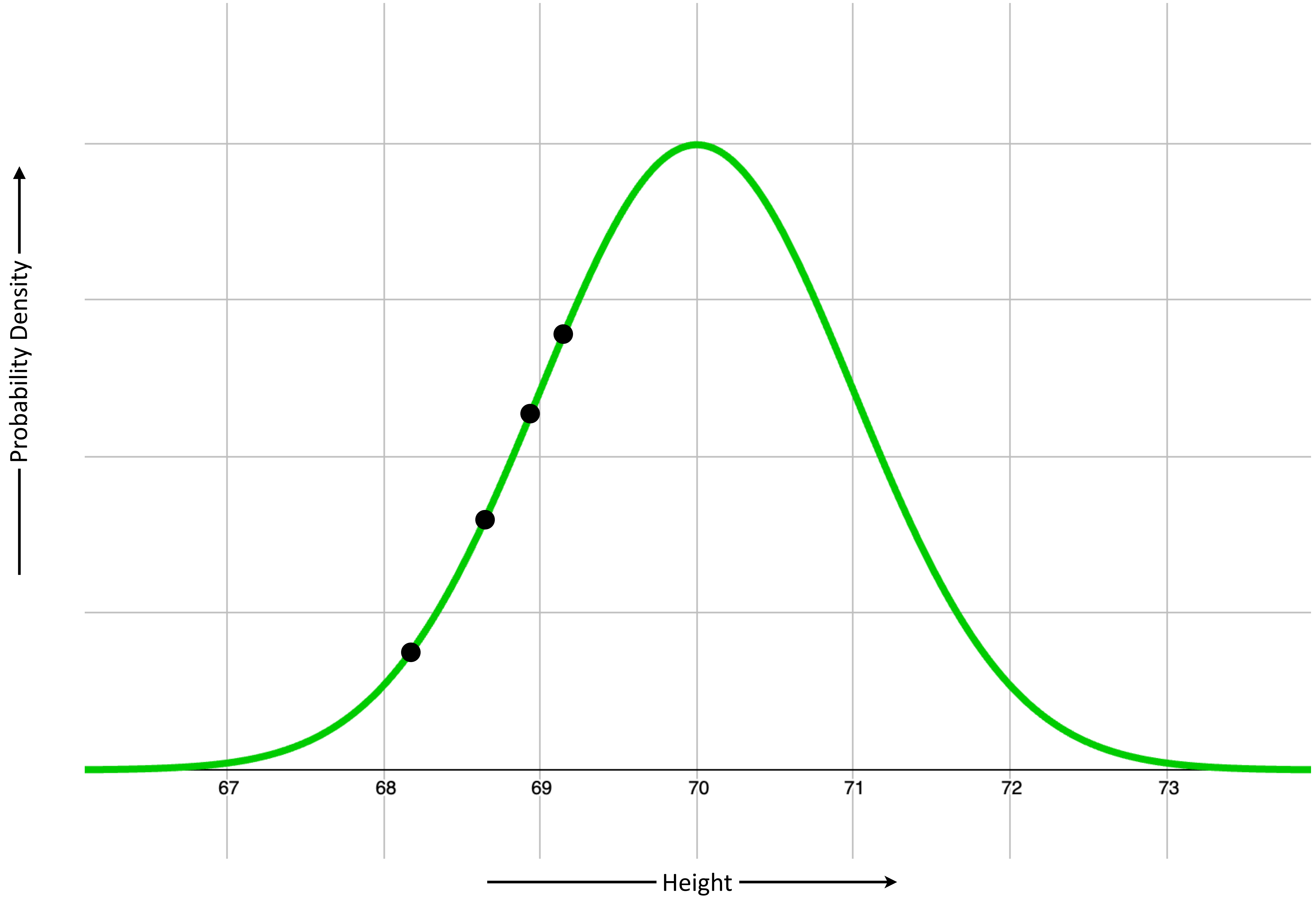
Example: X is the height of a citizen of Bulgaria

Mean = μ

Variance = σ^2

Log Likelihood

Multiple points on the curve:
Sum of the log of the likelihoods of every point



Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Mean = μ

Variance = σ^2

Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Mean = μ

Variance = σ^2

Probability Density function: Function that returns the probability of x within a range given fixed parameters μ and σ

Probability Density Function

$f(x \mid \mu, \sigma^2)$

Probability Density Function

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

Mean = μ

Variance = σ^2

Probability Density Function

$f(x \mid \mu, \sigma^2)$

Likelihood Function

$\mathscr{L}(\mu, \sigma^2 \mid x)$

Probability Density function: Function that returns the probability of x within a range given fixed parameters μ and σ

Likelihood function: Function that calculates the plausibility of x taking a specific value for parameters μ and σ

Maximum Likelihood Estimation

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

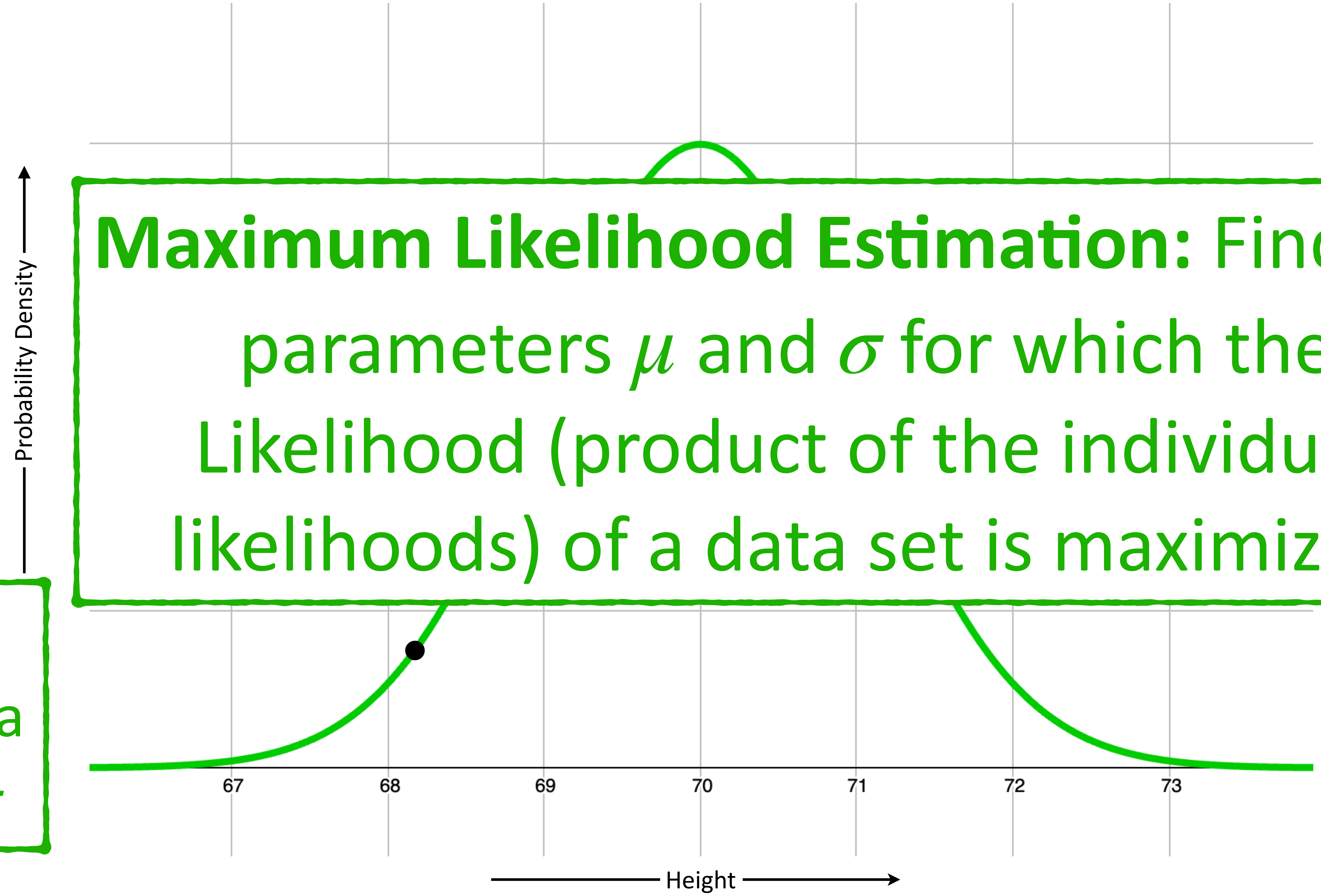
Mean = μ

Variance = σ^2

Likelihood Function

$\mathscr{L}(\mu, \sigma^2 \mid x)$

Likelihood function: Function that calculates the plausibility of x taking a specific value for parameters μ and σ



Maximum Likelihood Estimation: Find the parameters μ and σ for which the Likelihood (product of the individual likelihoods) of a data set is maximized

Continuous Random Variable

A random variable with an infinite, uncountable set of outcomes

Example: X is the height of a citizen of Bulgaria

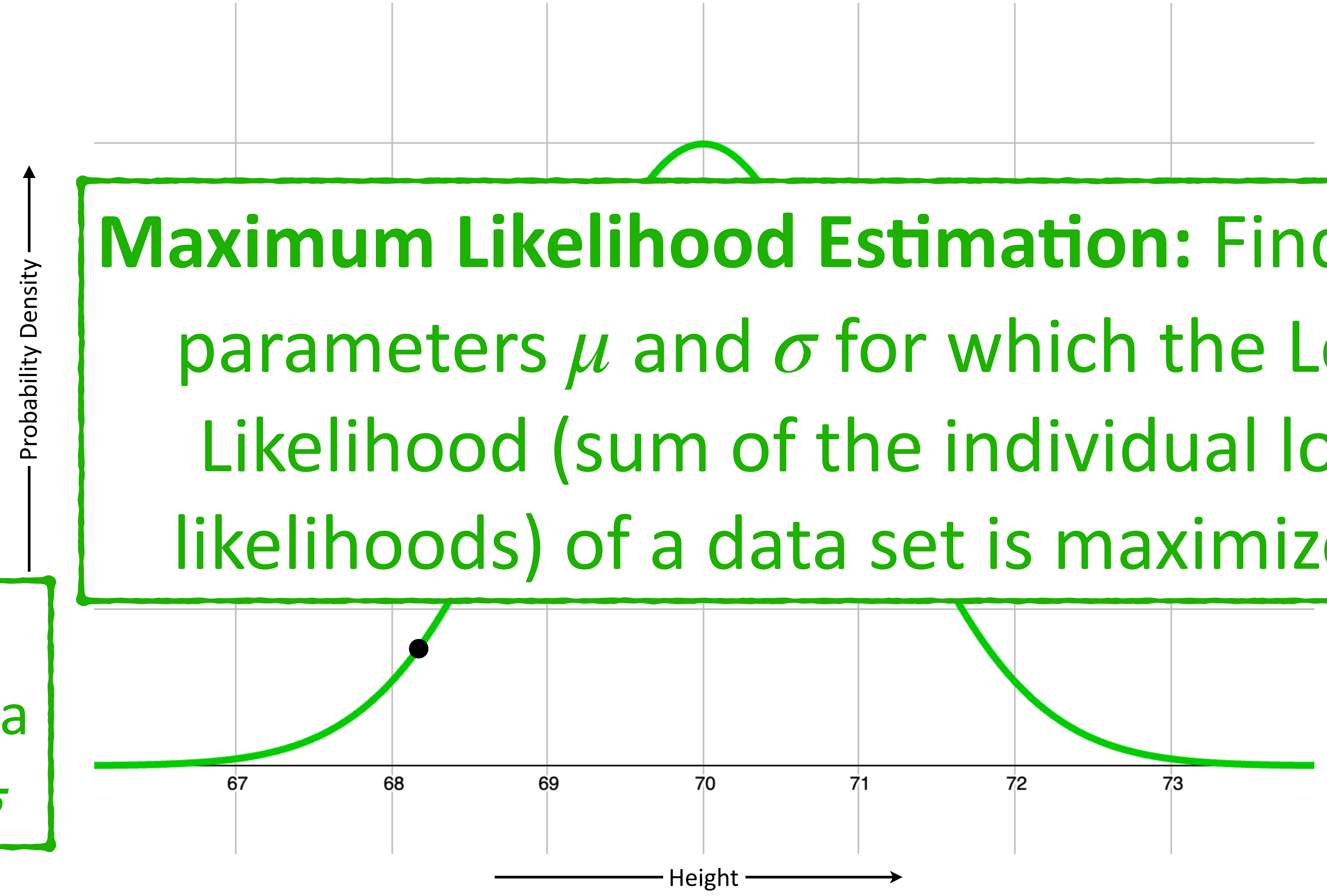
Mean = μ

Variance = σ^2

Likelihood Function

$\mathscr{L}(\mu, \sigma^2 \mid x)$

Likelihood function: Function that calculates the plausibility of x taking a specific value for parameters μ and σ



Maximum Likelihood Estimation: Find the parameters μ and σ for which the Log Likelihood (sum of the individual log likelihoods) of a data set is maximized

Lets review some well known probability distributions

Normal Distribution

Normal Distribution: The normal distribution is symmetric around the mean - outcomes near the mean occur more frequently than those further away from the mean. Also known as the Gaussian Distribution and characterized by the "Bell Curve"

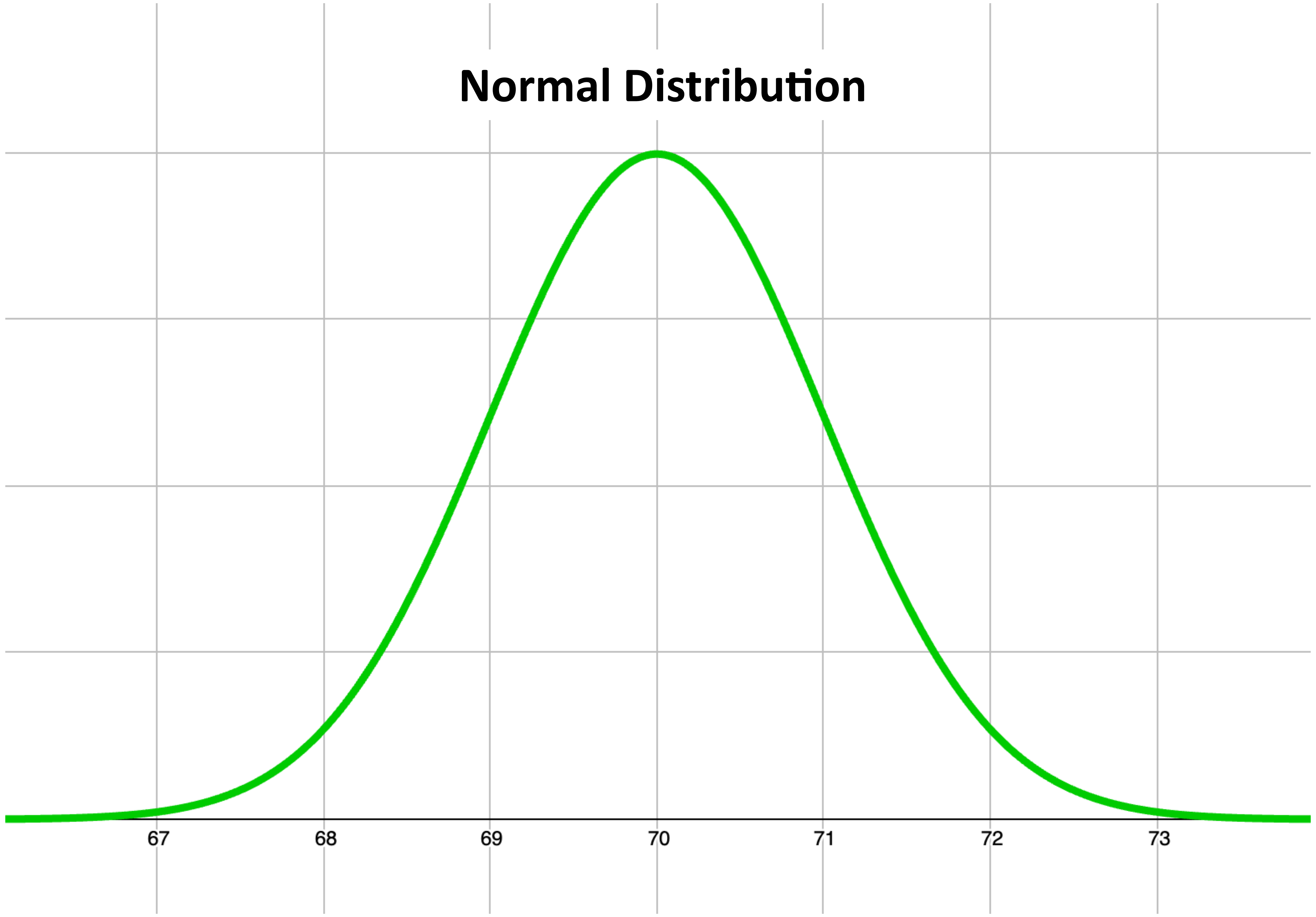
Parameters:

Mean = μ

Variance = σ^2

Probability Density Function (PDF)

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Binomial Distribution

Binomial Distribution: A Discrete Probability distribution of the number of successful outcomes in a sequence of n independent trials where each trial has binary outcome - true with probability p and false with a probability $1 - p$

Parameters:

Number of trials = n

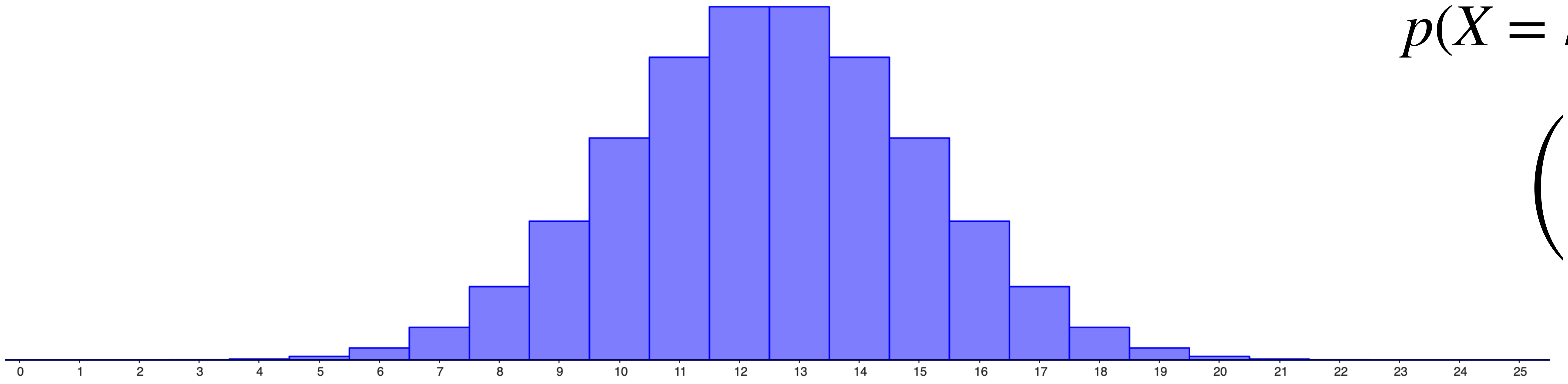
Probability of success of each trial = p

Probability Mass Function (PMF)

Probability of getting k successes in n independent Bernoulli trials

$$p(X = k) = \binom{n}{k} p^k (1 - p)^{(n-k)}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$



Binomial Distribution

n = 25, p = 0.5

Bernoulli Distribution: A Discrete Probability distribution of a random variable that has an outcome of 1 with a probability p and an outcome of 0 with probability $1 - p$. Its a special case of the binomial distribution with $n = 1$ (a single trial)

Parameters:

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

Probability Mass Function (PMF)

$$f(x) = \begin{Bmatrix} p, & if\, x = 1 \\ 1 - p, & if\, x = 0 \end{Bmatrix}$$

$$f(x) = p^k (1 - p)^{1-k}$$

Bernoulli Distribution
n = 1, p = 0.2

Probability of an Event Occurring

The number of desired outcomes divided by the total number of outcomes

$$P(X) = \frac{Number\ of\ Desired\ Outcomes}{Total\ Number\ of\ Outcomes}$$

Odds of an Event Occurring

The Probability that the event will occur divided by the probability that the event will not occur

$$O(X) = \frac{p}{1-p}$$

Related Tutorials & Textbooks

Logistic Regression

An introduction to Logistic Regression. A Logistic Regression model use used to predict a binary value (the dependent variable) for one or more independent variables using a threshold to classify a probability.

Multiple Regression

Multiple regression extends the two dimensional linear model introduced in Simple Linear Regression to k + 1 dimensions with one dependent variable, k independent variables and k+1 parameters.

Cost Function & Gradient Descent for Logistic Regression

An introduction to the Cost function for Logistic Regression long with its partial derivative (the gradient vector). The model parameters (B & W) are then optimized using Maximum Likelihood Estimation and Gradient Descent.

For a complete list of tutorials see:
https://arrsingh.com/ai-tutorials