



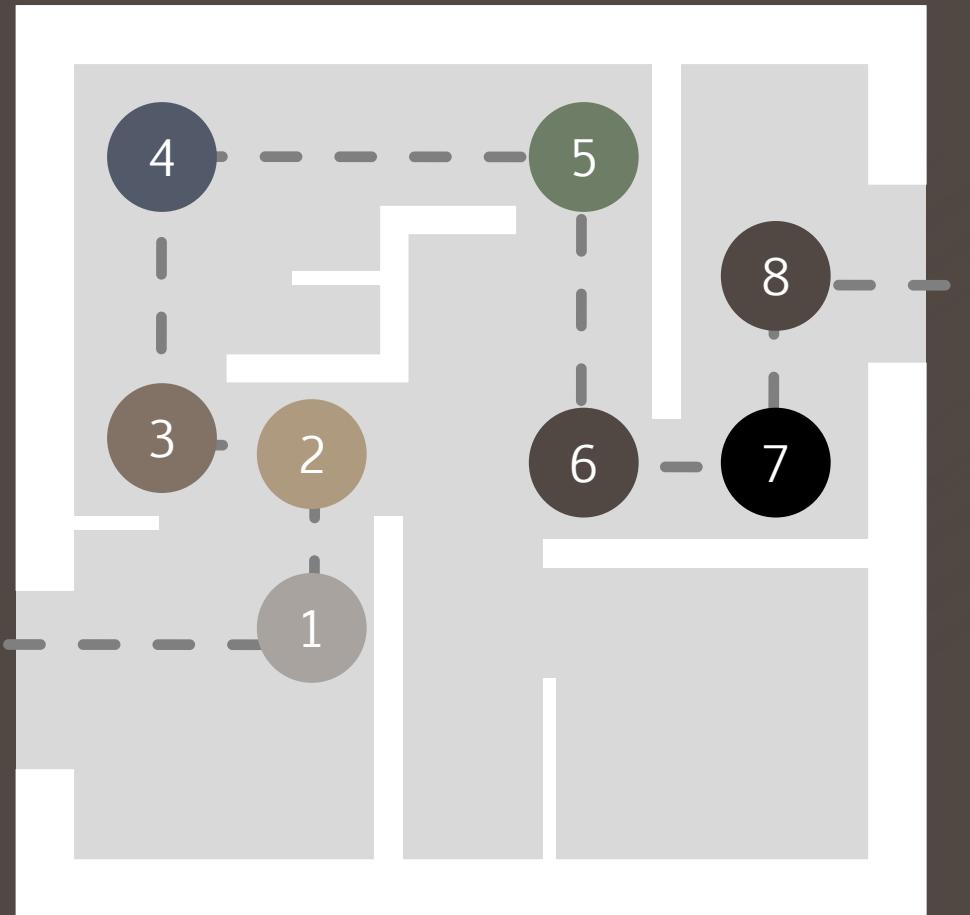
# MATH 20: PROBABILITY

Fundamental Theorems of Probability Theory



# Fundamental Theorems of Probability Theory

Law of Large  
Numbers



Central Limit  
Theorem

## Law of Large Numbers

- frequency as interpretation of probability
  - convergence of the sample mean





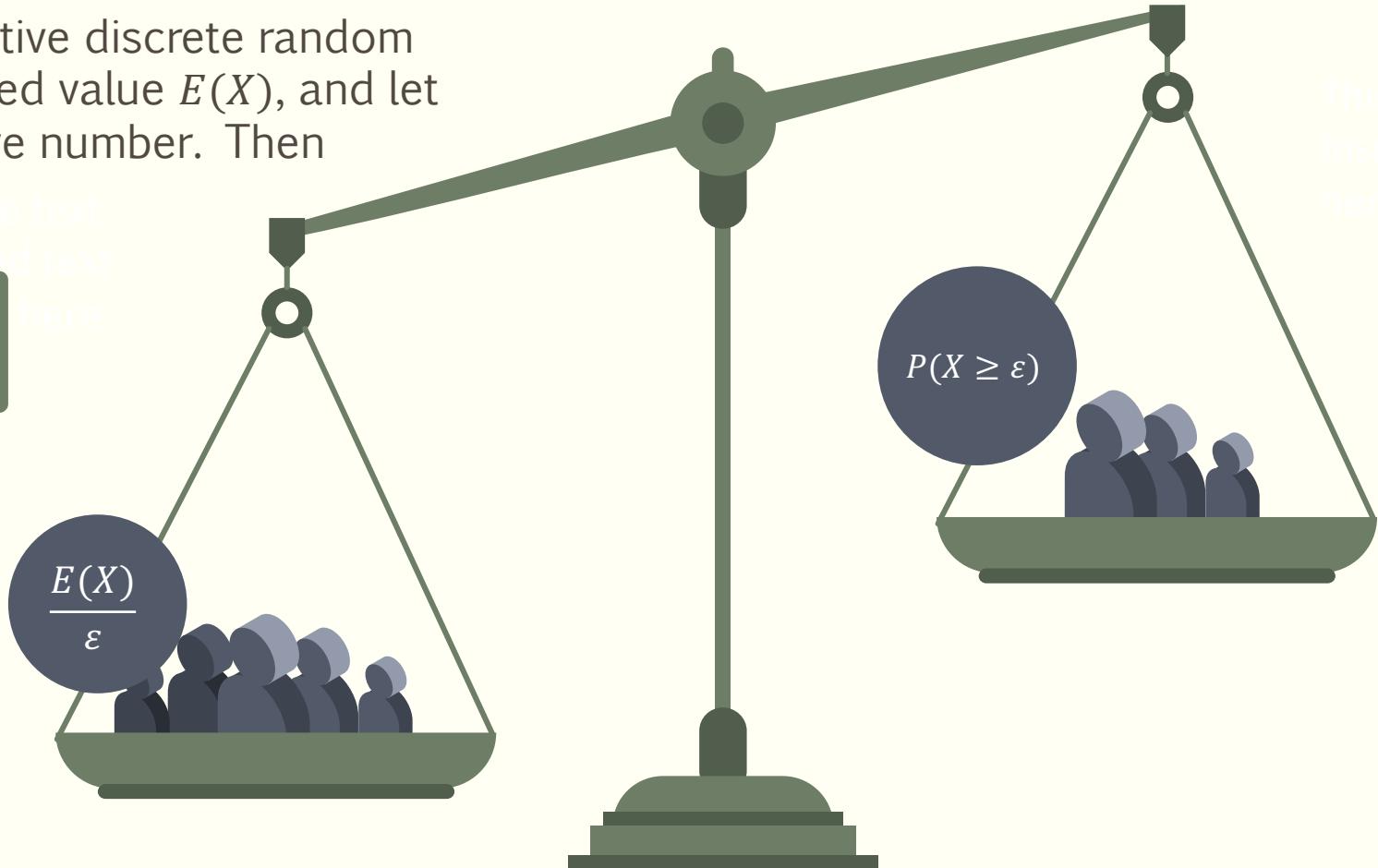
# LAW OF LARGE NUMBERS FOR DISCRETE RANDOM VARIABLES

discrete random variables

Let  $X$  be a nonnegative discrete random variable with expected value  $E(X)$ , and let  $\varepsilon > 0$  be any positive number. Then

This is a sample text.  
Insert your desired text  
here

**$X$ : nonnegative**



This is a sample text.  
Insert your desired text  
here

**Markov** inequality

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}$$

# Proof

Let  $X$  be a **positive** discrete random variable with expected value  $E(X)$ , and let  $\varepsilon > 0$  be any positive number.

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}$$

$$\sum_x m(x) = 1$$

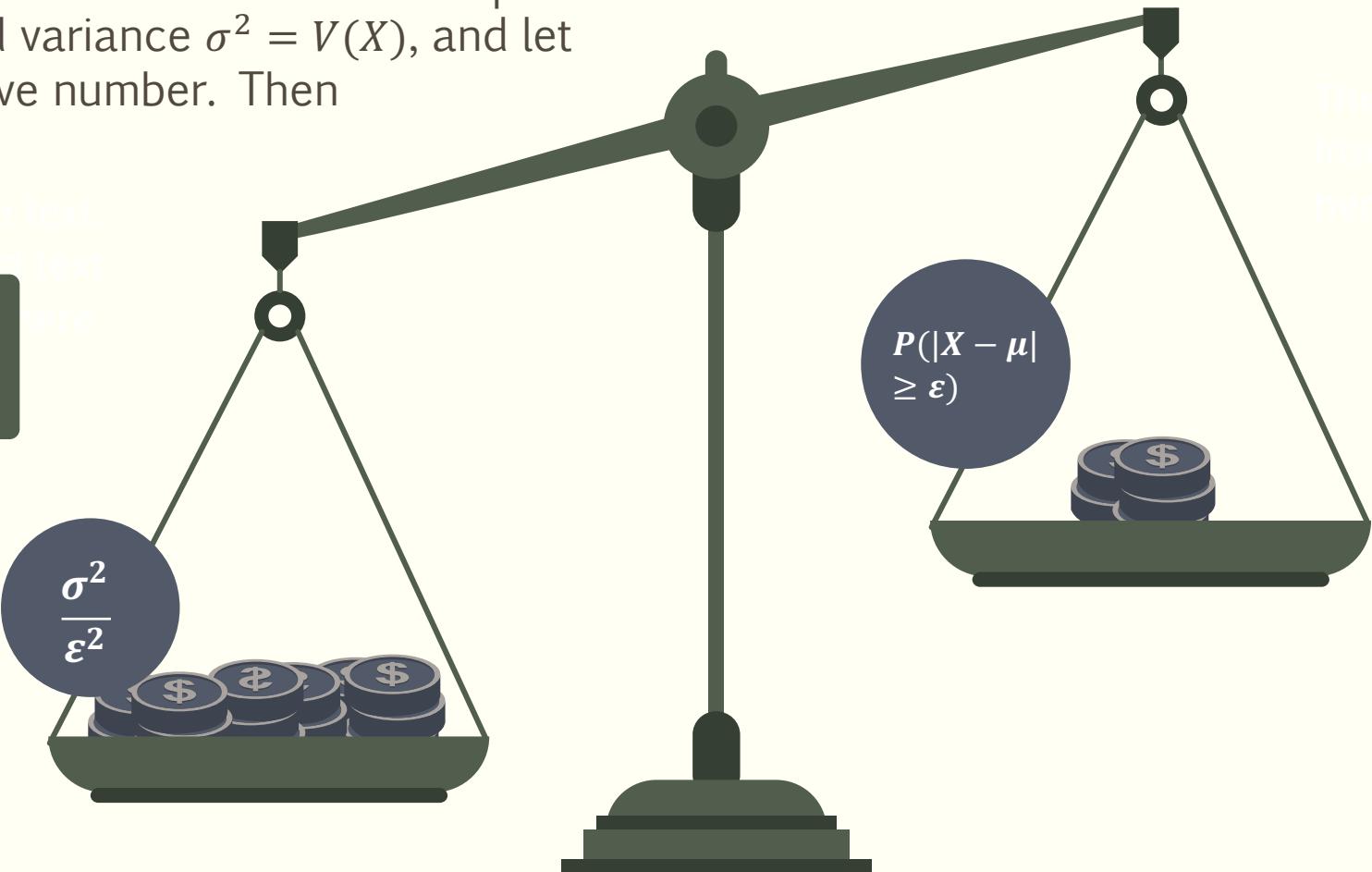
$$\sum_x xm(x) = E(X)$$

$$P(X \geq \varepsilon) = \sum_{x \geq \varepsilon} m(x)$$

$$E(X) = \sum_x xm(x) \geq \sum_{x \geq \varepsilon} xm(x) \geq \varepsilon \sum_{x \geq \varepsilon} m(x) = \varepsilon P(X \geq \varepsilon)$$

Let  $X$  be a discrete random variable with expected value  $\mu = E(X)$  and variance  $\sigma^2 = V(X)$ , and let  $\varepsilon > 0$  be any positive number. Then

X: not necessarily  
nonnegative



Chebyshev  
inequality

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

This is a sample text.  
Insert your desired text  
here.

# Proof

Let  $X$  be a discrete random variable with expected value  $\mu = E(X)$  and variance  $\sigma^2 = V(X)$ , and let  $\varepsilon > 0$  be any positive number.

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

$$\sum_x m(x) = 1$$

$$P(|X - \mu| \geq \varepsilon) = \sum_{|x-\mu| \geq \varepsilon} m(x)$$

$$\sum_x (x - \mu)^2 m(x) = V(X)$$

$$\begin{aligned} \sigma^2 = V(X) &= \sum_x (x - \mu)^2 m(x) \geq \sum_{|x-\mu| \geq \varepsilon} (x - \mu)^2 m(x) \\ &\geq \varepsilon^2 \sum_{|x-\mu| \geq \varepsilon} m(x) = \varepsilon^2 P(|X - \mu| \geq \varepsilon) \end{aligned}$$

## Markov inequality

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}$$

$X$ : nonnegative

$E(X)$ : known

## Chebyshev inequality

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

$X$ : not necessarily  
nonnegative

$E(X)$ : known  
 $V(X)$ : known

## Example 1

---

---

- Let  $X$  be any random variable which takes on values  $0, 1, 2, \dots, n$  and has  $E(X) = V(X) = 1$ . Show that for any positive integer  $k$ ,

$$P(X \geq k + 1) \leq \frac{1}{k^2}.$$

### Markov Inequality

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}$$

### Chebyshev Inequality

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

## Example 1

---

---

- Let  $X$  be any random variable which takes on values  $0, 1, 2, \dots, n$  and has  $E(X) = V(X) = 1$ . Show that for any positive integer  $k$ ,

$$P(X \geq k + 1) \leq \frac{1}{k^2}.$$

### Chebyshev Inequality

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

$$\mu = \sigma^2 = 1, \varepsilon = k$$

$$P(|X - 1| \geq k) = P(X \geq k + 1) \leq \frac{1}{k^2}$$

## Example 2

---

---

- Choose  $X$  with distribution

$$m(x) = \begin{pmatrix} -\varepsilon & \varepsilon \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

$$P(|X| \geq \varepsilon) = 1$$

### Chebyshev Inequality

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

$$\mu = 0, \sigma^2 = \varepsilon^2$$

$$P(|X| \geq \varepsilon) \leq \frac{\varepsilon^2}{\varepsilon^2} = 1$$

# Law of Large Numbers

---

---

- Let  $X_1, X_2, \dots, X_n$  be an independent trials process, with same finite expected value  $\mu = E(X_j)$  and finite variance  $\sigma^2 = V(X_j)$ .
- Let  $S_n = X_1 + X_2 + \dots + X_n$ . Then for any  $\varepsilon > 0$ ,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \rightarrow 0, \text{ as } n \rightarrow +\infty.$$

Large  
Numbers

- Equivalently,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right) = 1.$$

As  $\frac{S_n}{n}$  is an average of the individual outcomes, the LLN is often referred to as the law of averages.

!

# Law of Large Numbers

---

---

- Let  $X_1, X_2, \dots, X_n$  be an independent trials process with  $E(X_j) = \mu$  and  $V(X_j) = \sigma^2$ .
- Let  $S_n = X_1 + X_2 + \dots + X_n$  be the sum, and  $A_n = \frac{S_n}{n}$  be the average. Then

$$E(A_n) = \mu$$



$n \rightarrow +\infty$

$$V(A_n) \rightarrow 0$$

$$V(A_n) = \frac{\sigma^2}{n}, D(A_n) = \frac{\sigma}{\sqrt{n}}$$



$$D(A_n) \rightarrow 0$$

# Proof

Let  $X_1, X_2, \dots, X_n$  be an independent trials process, with same finite expected value  $\mu = E(X_j)$  and finite variance  $\sigma^2 = V(X_j)$ .

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \rightarrow 0, \text{ as } n \rightarrow +\infty.$$

$$E\left(\frac{S_n}{n}\right) = \mu$$

$$V\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n}$$

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

$$n \rightarrow +\infty$$

$$\frac{\sigma^2}{n\varepsilon^2} \rightarrow 0$$

# Law of Large Numbers

convergence of the sample mean

(Strong)  
Law of Large  
Numbers

$$P \left( \lim_{n \rightarrow \infty} \frac{S_n}{n} = \mu \right) = 1$$

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{S_n}{n} - \mu \right| < \varepsilon \right) = 1$$

(Weak)  
Law of Large  
Numbers



## Example 3

- Consider the general Bernoulli trial process.
- As usual, we let  $X = 1$  if the outcome is a success and  $0$  if it is a failure.

### Bernoulli trial

$$m(x) = \begin{cases} p, & X = 1 \\ 1 - p, & X = 0 \end{cases}$$

### Expected value $E(X)$

$$\sum_{x \in \Omega} xm(x) = 1 \times p + 0 \times (1 - p) = p$$

### Variance $V(X)$

$$E(X^2) - \mu^2 = p - p^2 = p(1 - p)$$

## Example 3

- Now consider  $n$  Bernoulli trials.
- Then  $S_n = \sum_{i=1}^n X_i$  is the number of successes in  $n$  trials and  $\mu = E\left(\frac{S_n}{n}\right) = E(X_i) = p$ .

### Law of Large Numbers

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right) = 1$$

$$\mu = p$$

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - p\right| < \varepsilon\right) = 1$$

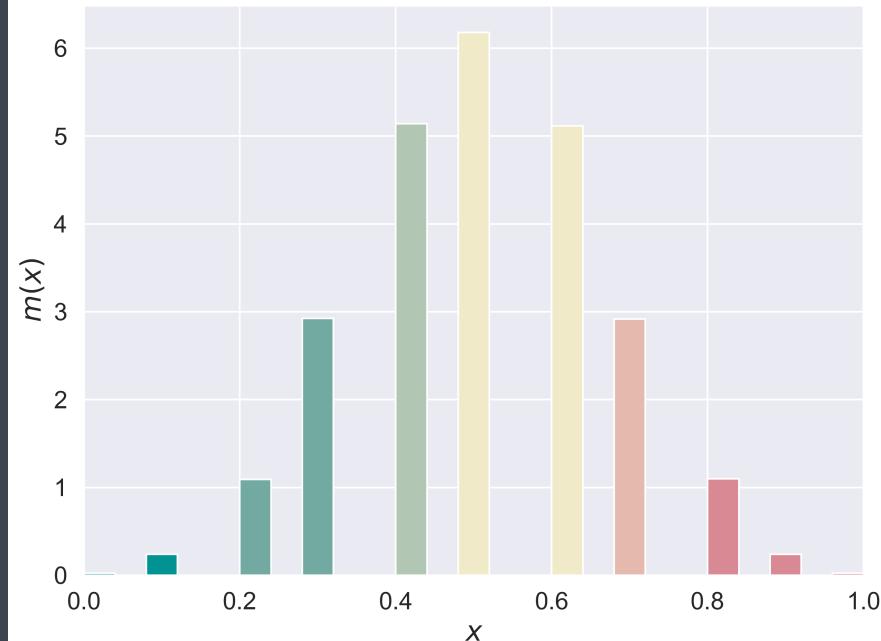
# Coin Tossing

$$n \rightarrow +\infty$$

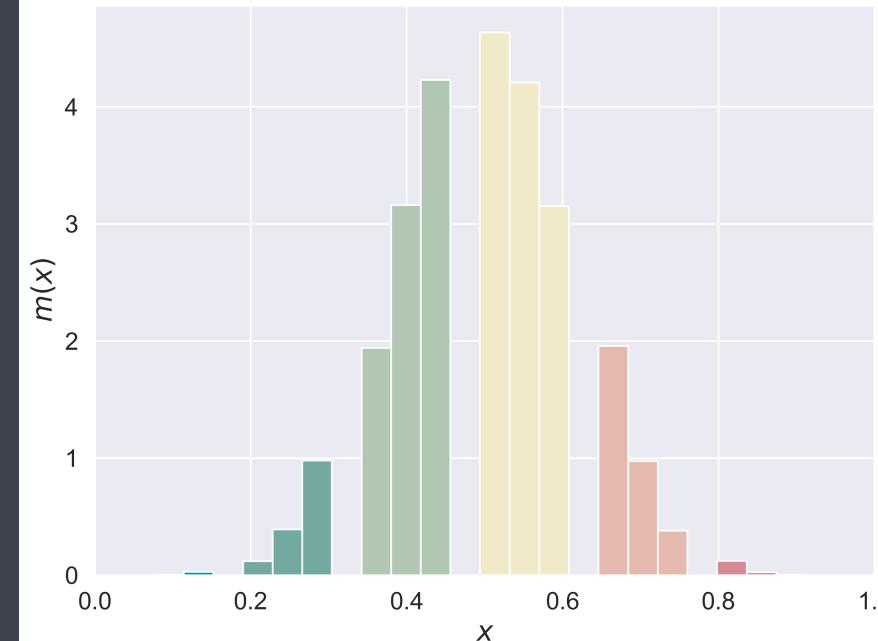
$$\mu = \frac{1}{2}$$

$$\frac{X_1 + X_2 + \dots + X_n}{n}$$

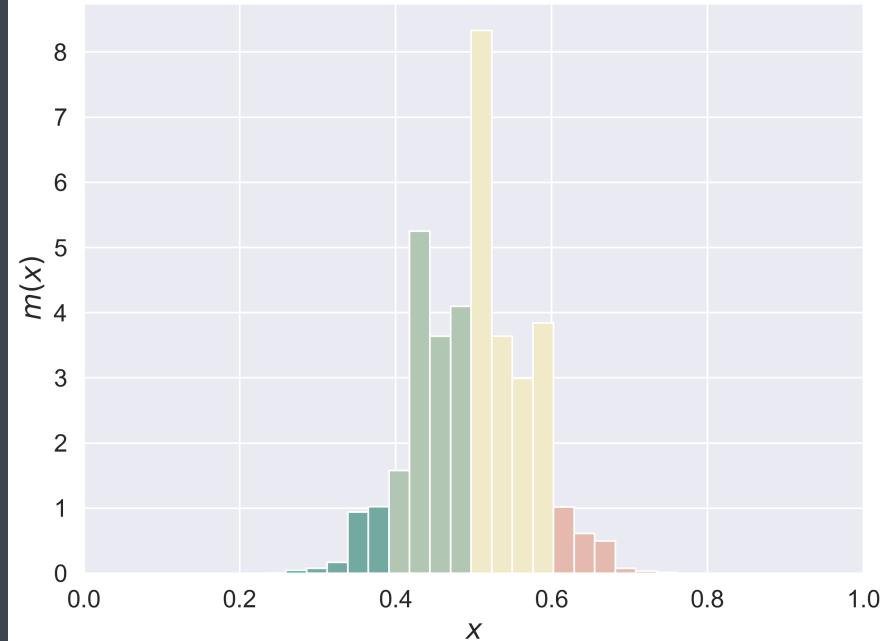
Coin Tossing:  $n = 10$



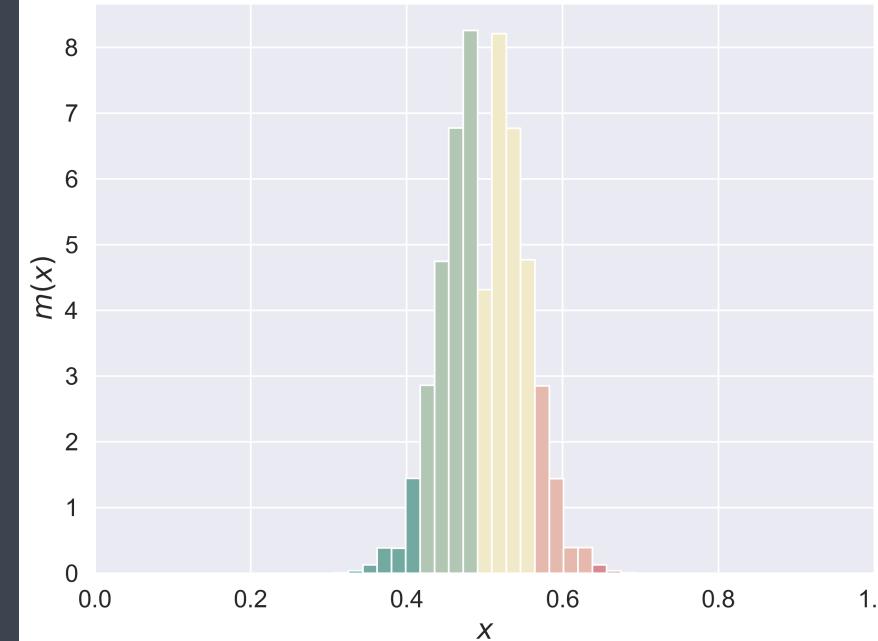
Coin Tossing:  $n = 20$



Coin Tossing:  $n = 50$



Coin Tossing:  $n = 100$



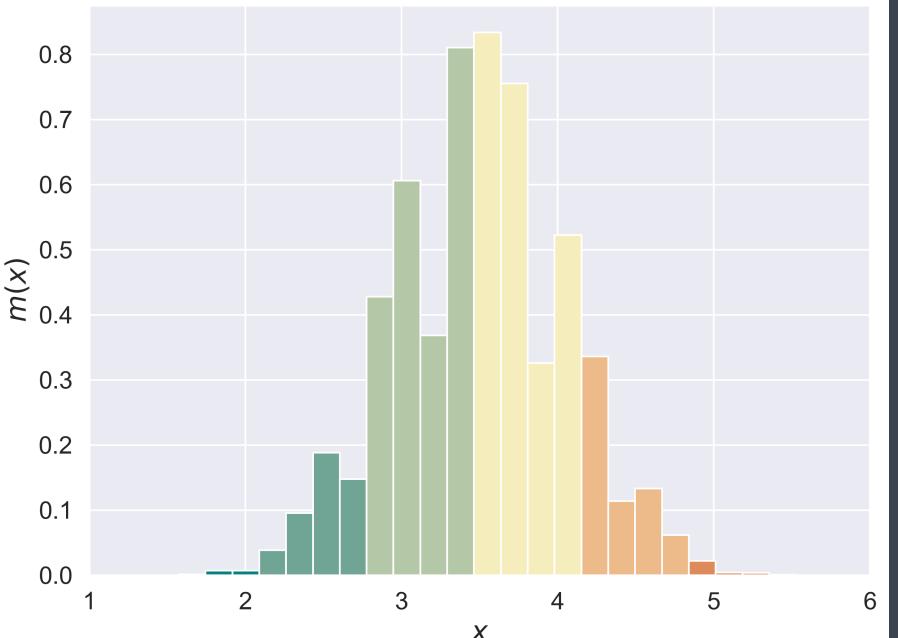
# Dice Rolling

$$n \rightarrow +\infty$$

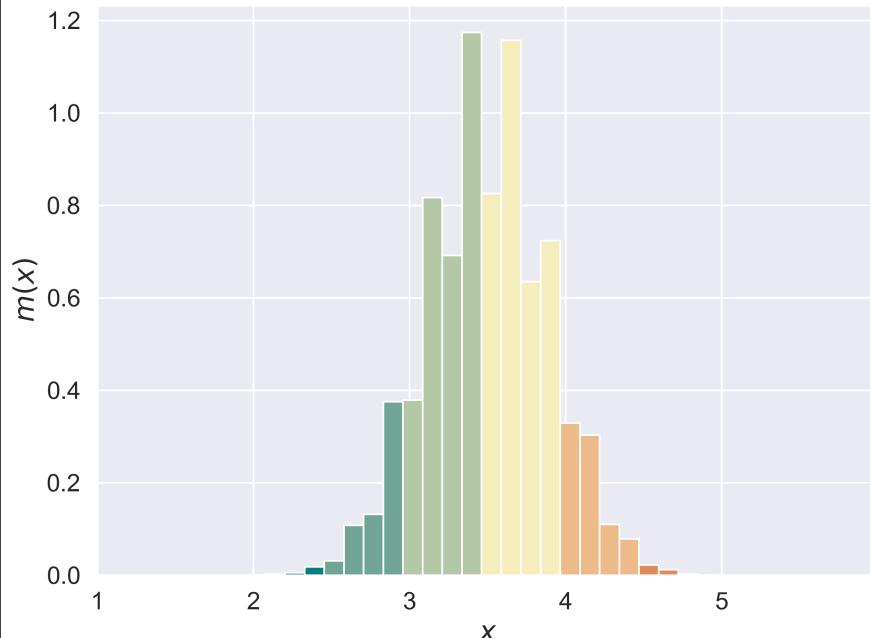
$$\mu = \frac{7}{2}$$

$$\frac{X_1 + X_2 + \dots + X_n}{n}$$

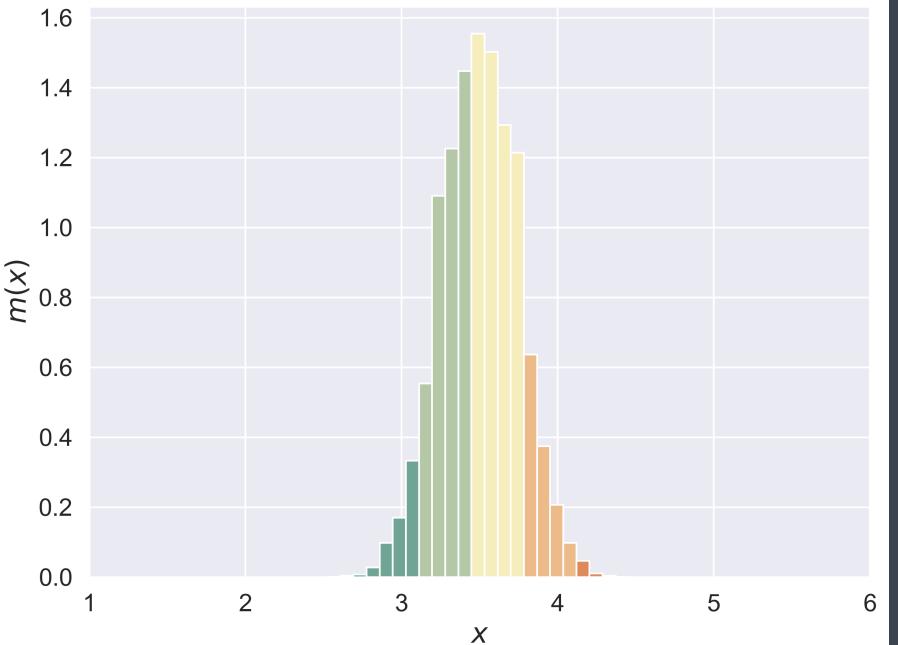
Dice Rolling:  $n = 10$



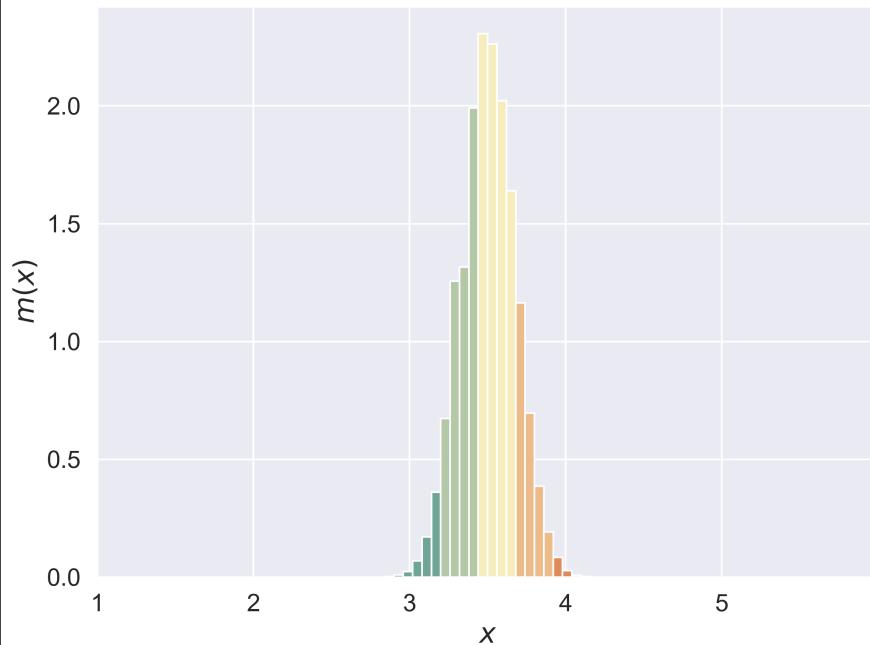
Dice Rolling:  $n = 20$



Dice Rolling:  $n = 50$



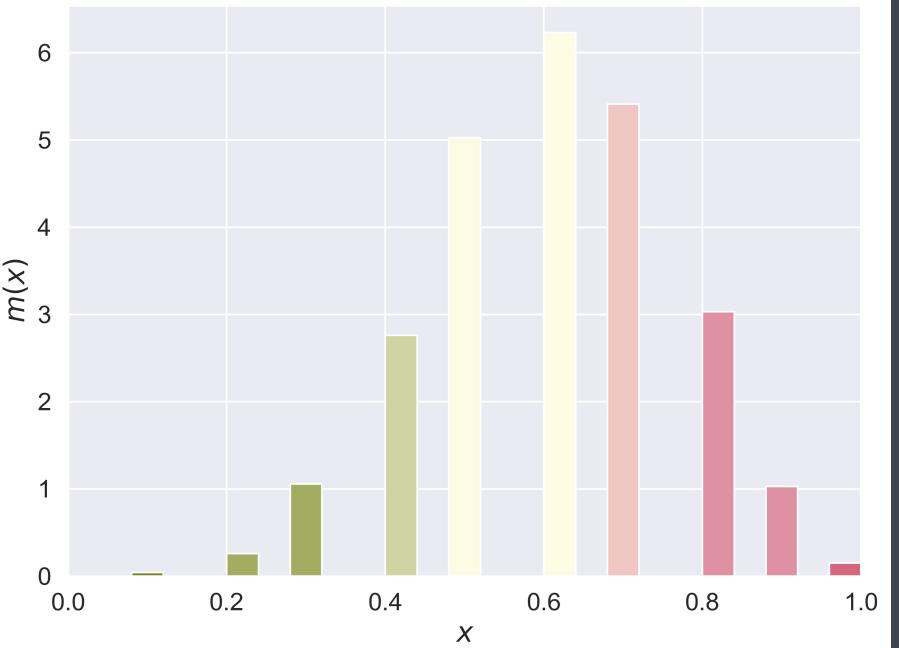
Dice Rolling:  $n = 100$



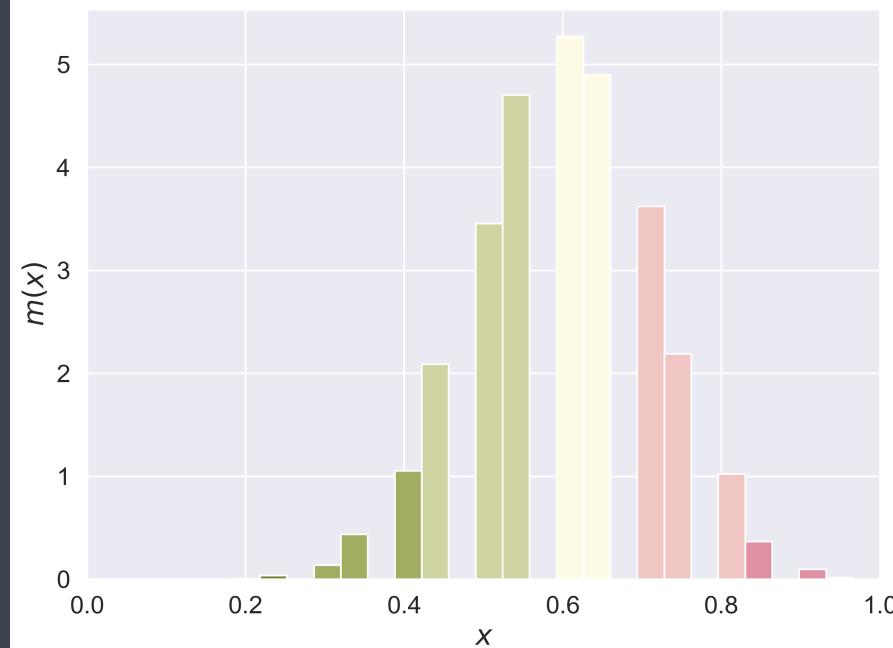
# Bernoulli Trials

We can start with a random experiment about which little can be predicted and, by taking averages, obtain an experiment in which the outcome can be predicted with a high degree of certainty.

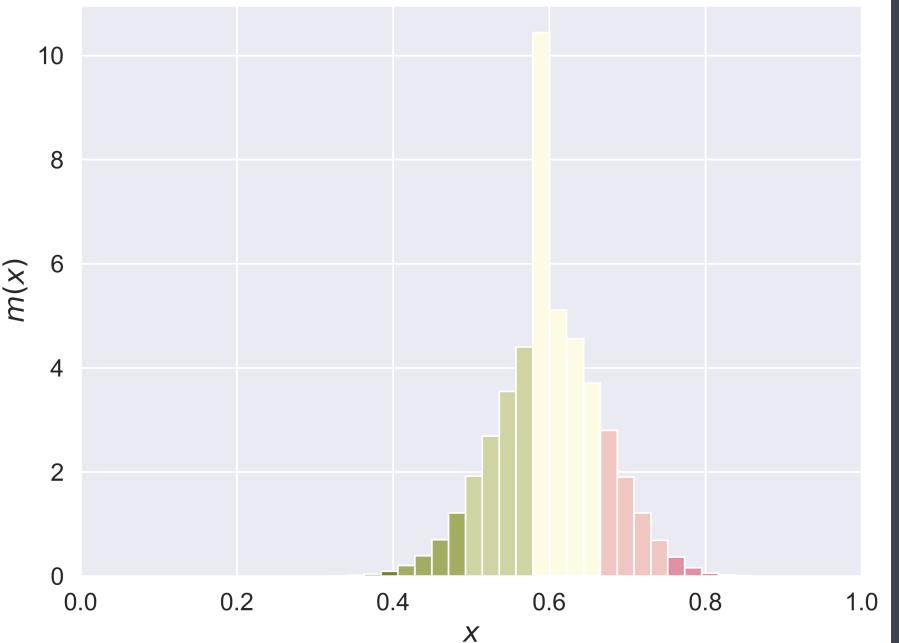
Bernoulli Trials:  $n = 10$



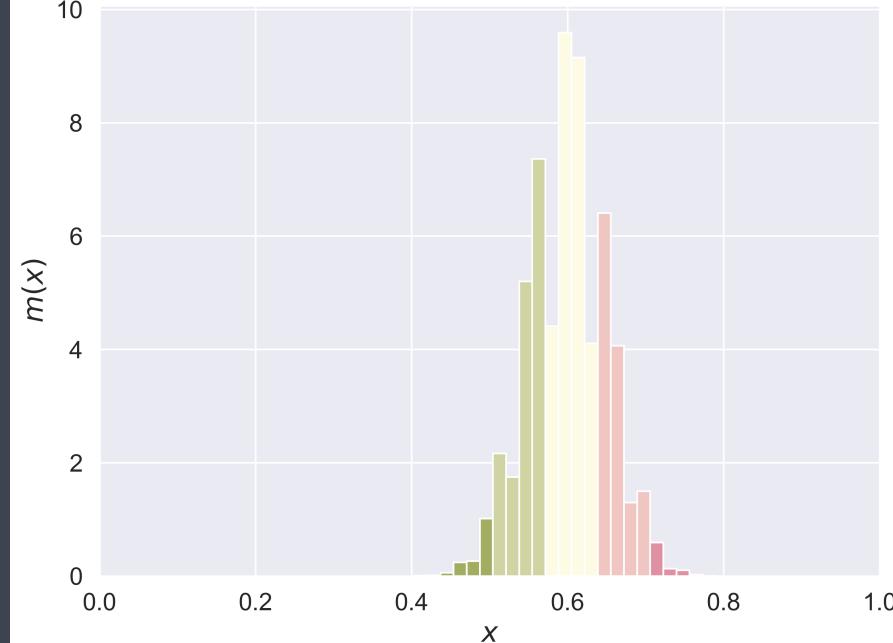
Bernoulli Trials:  $n = 20$



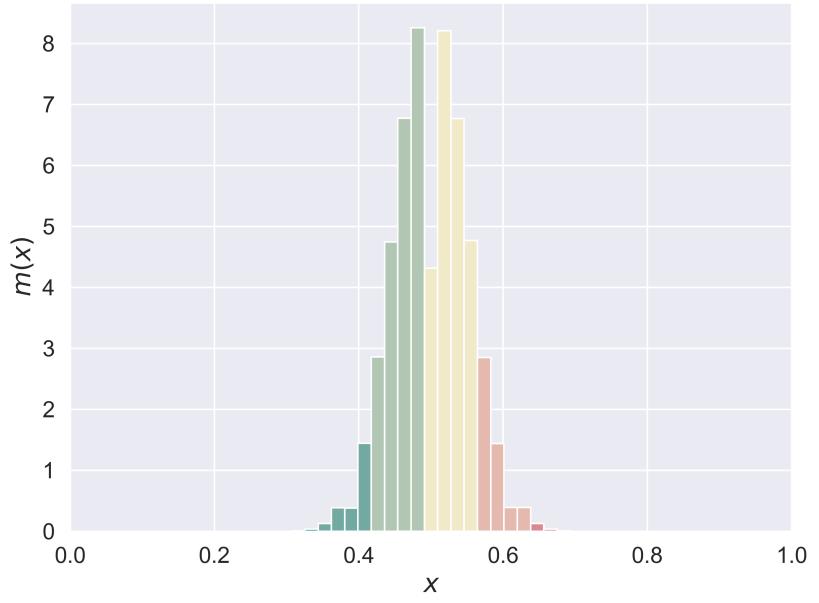
Bernoulli Trials:  $n = 50$



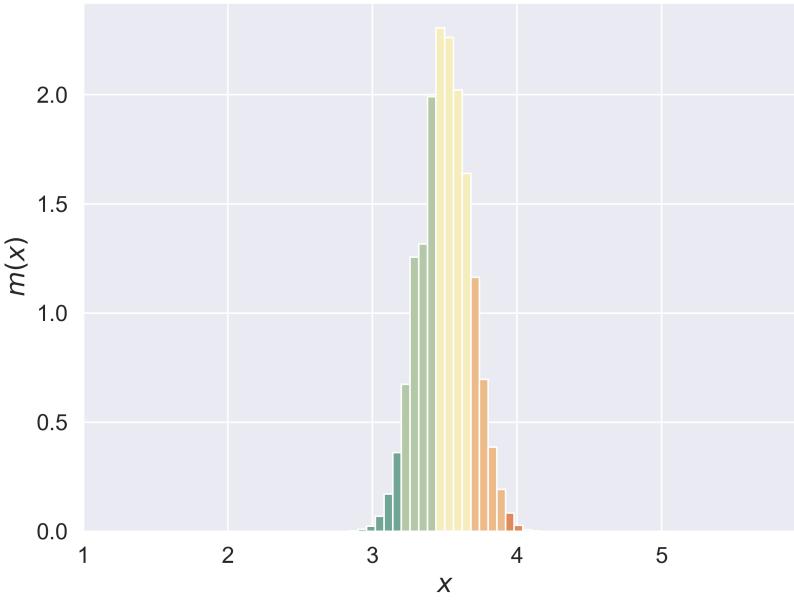
Bernoulli Trials:  $n = 100$



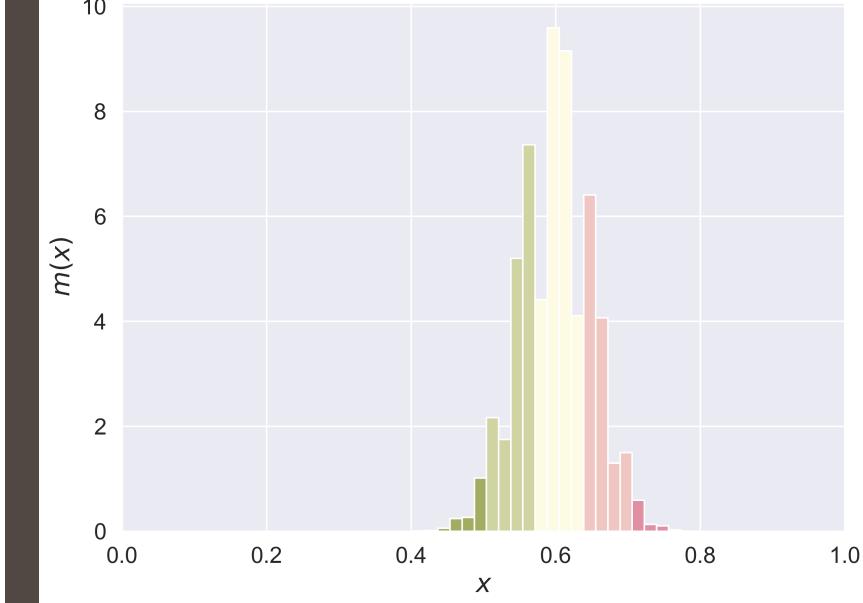
Coin Tossing: n = 100



Dice Rolling: n = 100



Bernoulli Trials: n = 100



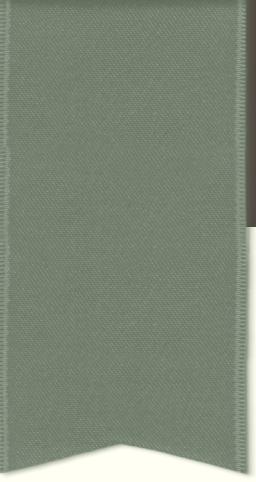
## Law of Large Numbers

- frequency as interpretation of probability
  - convergence of the sample mean

$$S_n = X_1 + X_2 + \cdots + X_n, A_n = \frac{S_n}{n}$$

$$E(A_n) = \mu, V(A_n) = \frac{\sigma^2}{n}$$

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$



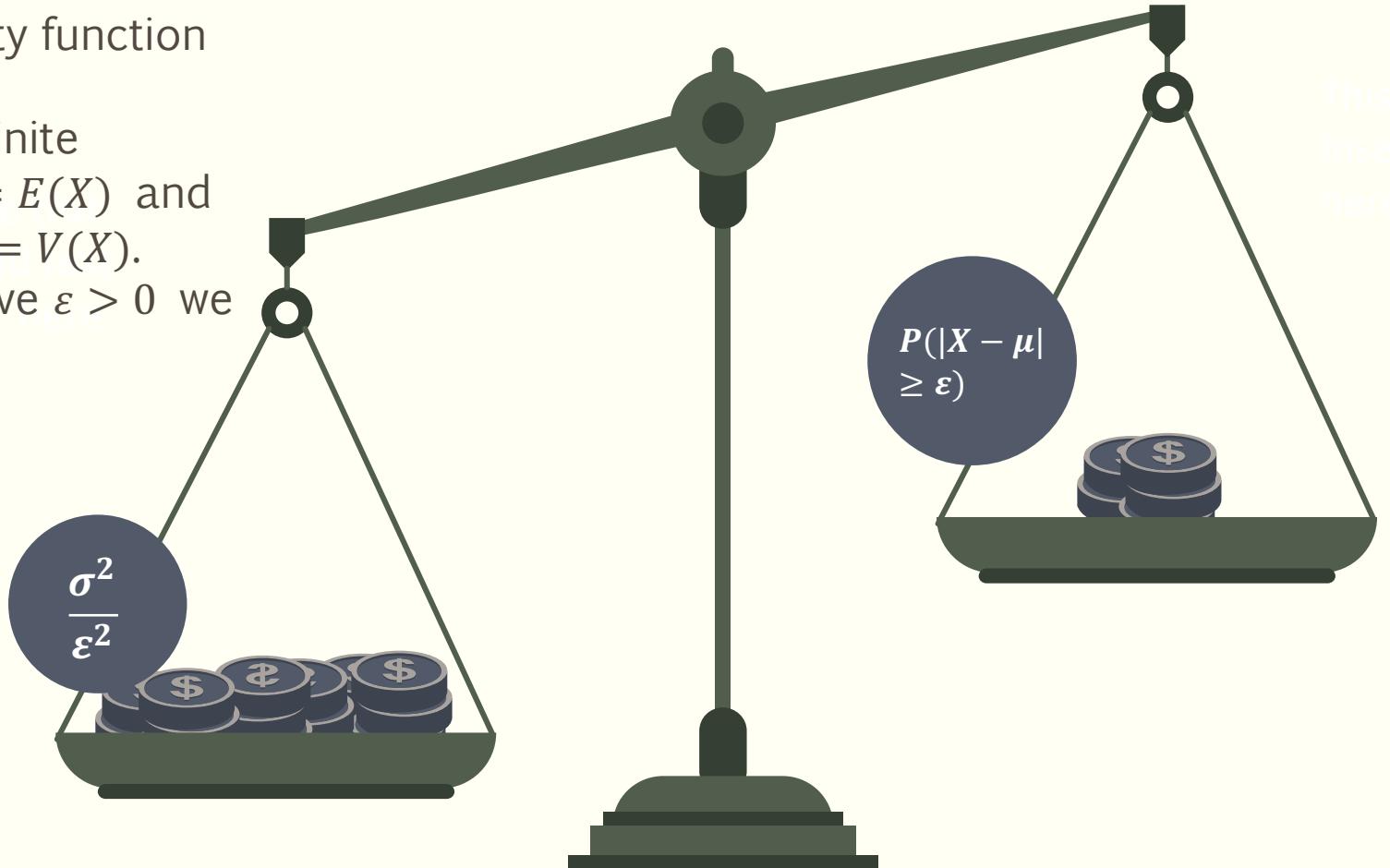
# LAW OF LARGE NUMBERS FOR CONTINUOUS RANDOM VARIABLES

continuous random variables

Let  $X$  be a continuous random variable with density function  $f(x)$ .

Suppose  $X$  has a finite expected value  $\mu = E(X)$  and finite variance  $\sigma^2 = V(X)$ .

Then for any positive  $\varepsilon > 0$  we have



Chebyshev  
inequality

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

# Law of Large Numbers

---

---

- Let  $X_1, X_2, \dots, X_n$  be an independent trials process with a continuous density function  $f$ , **finite** expected value  $\mu$  and **finite** variance  $\sigma^2$ .
- Let  $S_n = X_1 + X_2 + \dots + X_n$ . Then for any  $\varepsilon > 0$ ,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \rightarrow 0, \text{ as } n \rightarrow +\infty.$$

- Equivalently,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right) = 1.$$

As  $\frac{S_n}{n}$  is an average of the individual outcomes, the LLN is often referred to as the **law of averages**.

!

## Example 4

- Suppose we choose at random  $n$  numbers from the interval  $[0, 1]$  with uniform distribution.
- Let  $X_i$  describes the  $i$ th choice.

### Uniform distribution

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$$

$$f(x) = 1, \quad 0 \leq x \leq 1$$

### Expected value $E(X)$

$$E(X) = \frac{1}{2}(a + b) = \frac{1}{2}(0 + 1) = \frac{1}{2}$$

### Variance $V(X)$

$$\begin{aligned} V(X) &= E(X^2) - \mu^2 \\ &= \frac{1}{12}(b - a)^2 = \frac{1}{12}(1 - 0)^2 = \frac{1}{12} \end{aligned}$$

## Example 4

- Suppose we choose at random  $n$  numbers from the interval  $[0, 1]$  with uniform distribution.
- Let  $X_i$  describes the  $i$ th choice and  $S_n = \sum_{i=1}^n X_i$ .

Expected value

$$E\left(\frac{S_n}{n}\right) = \frac{1}{2}$$

Chebyshev Inequality

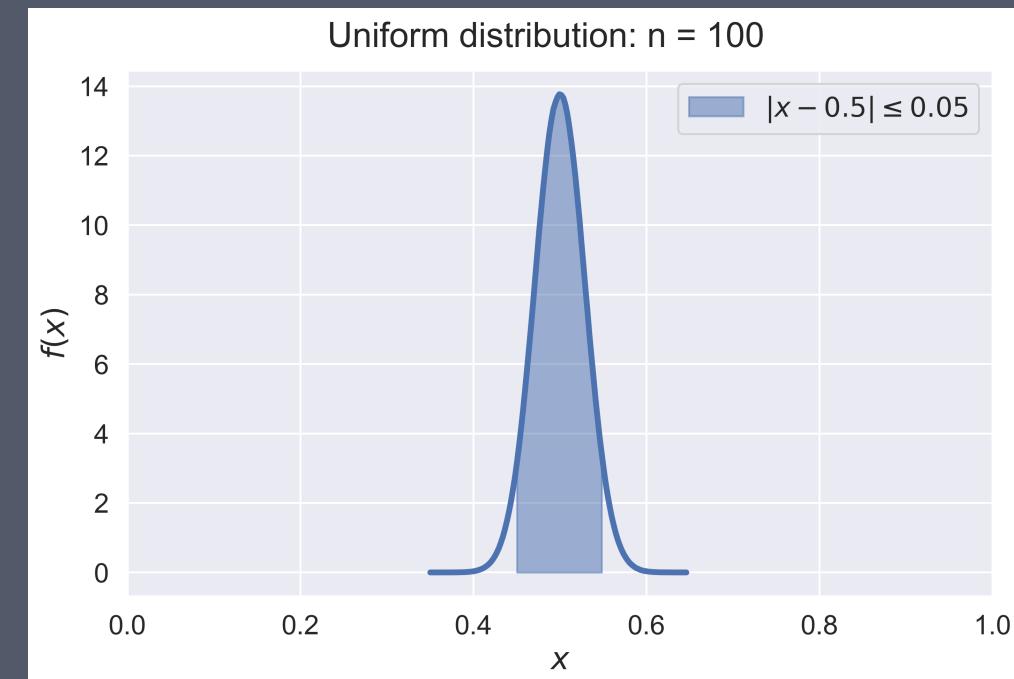
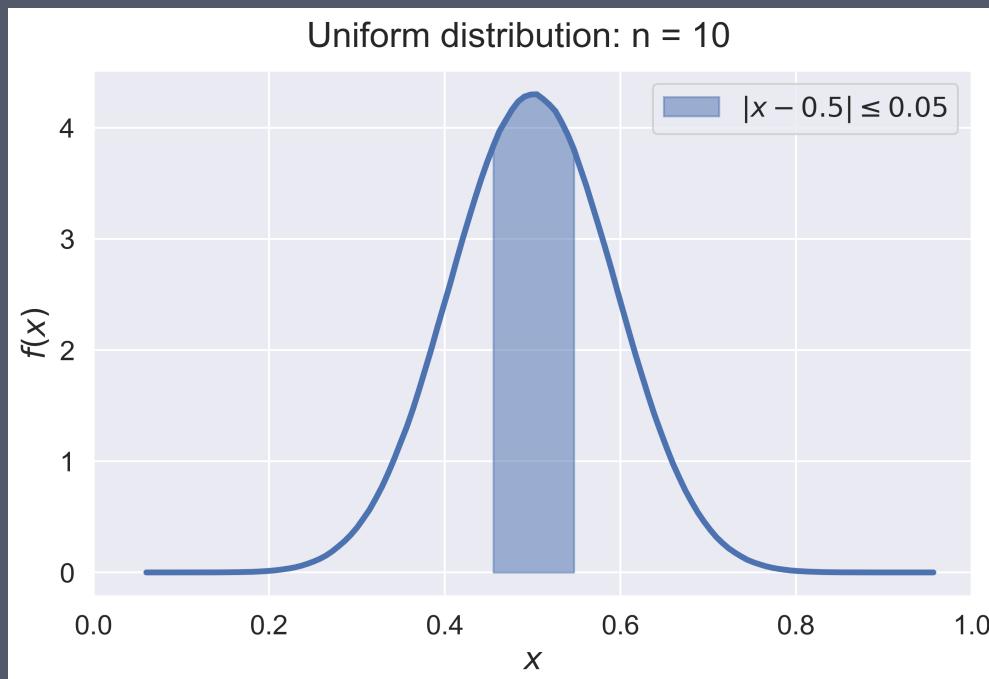
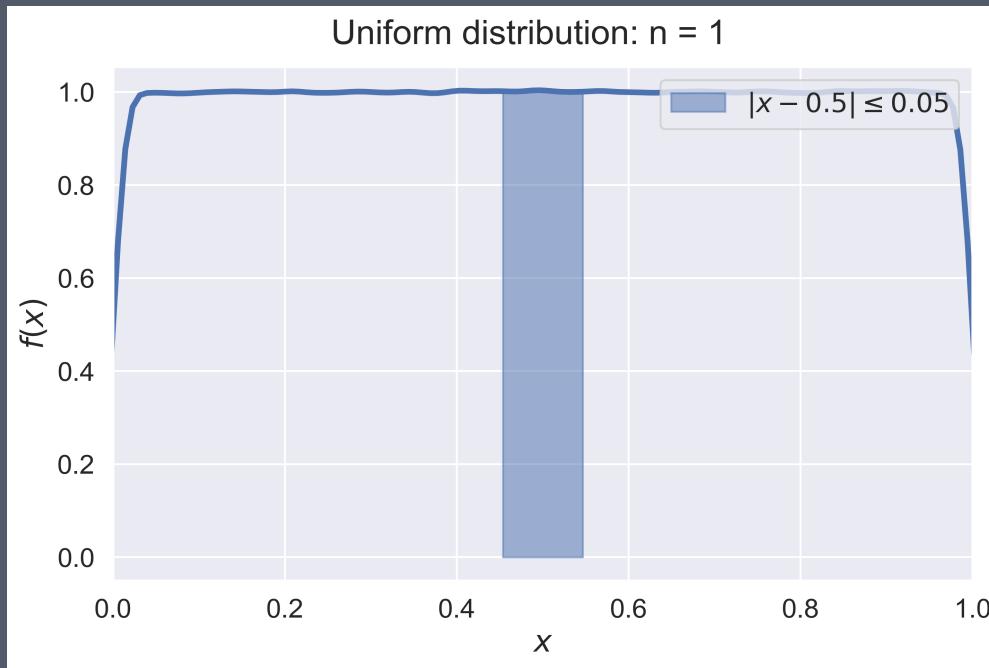
$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

$$\mu = \frac{1}{2}, \sigma^2 = \frac{1}{12}$$

Variance

$$V\left(\frac{S_n}{n}\right) = \frac{1}{12n}$$

$$P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right) \leq \frac{1}{12n\varepsilon^2}$$



## Example 5

- Suppose we choose  $n$  real numbers at random, using a normal distribution with mean 0 and variance 1. Then
  - $\mu = E(X_i) = 0$
  - $\sigma^2 = V(X_i) = 1$

Expected value

$$E\left(\frac{S_n}{n}\right) = 0$$

Chebyshev Inequality

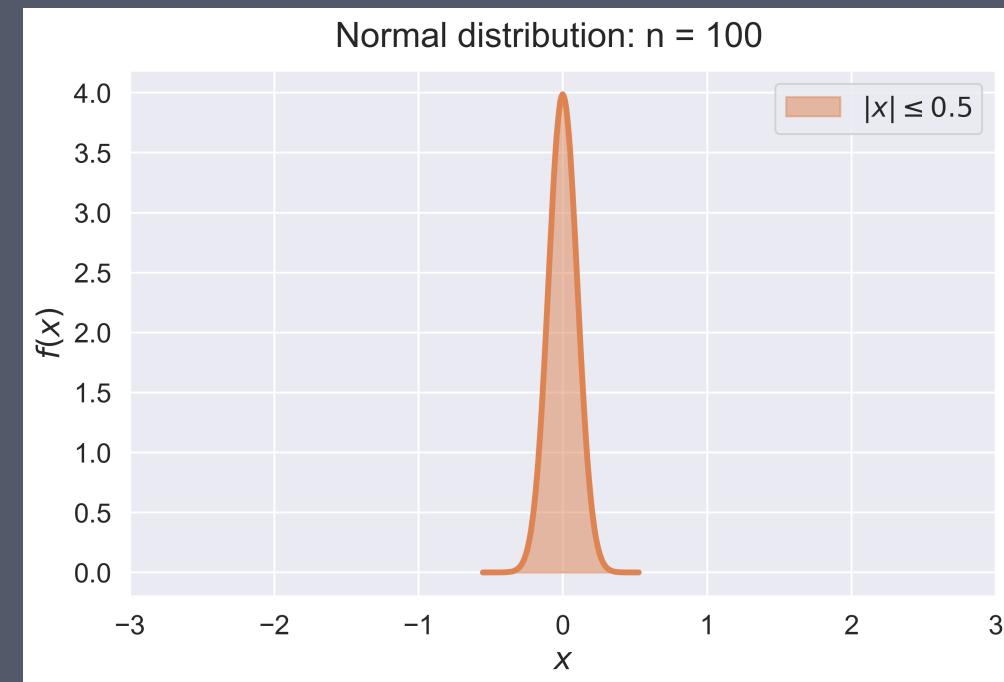
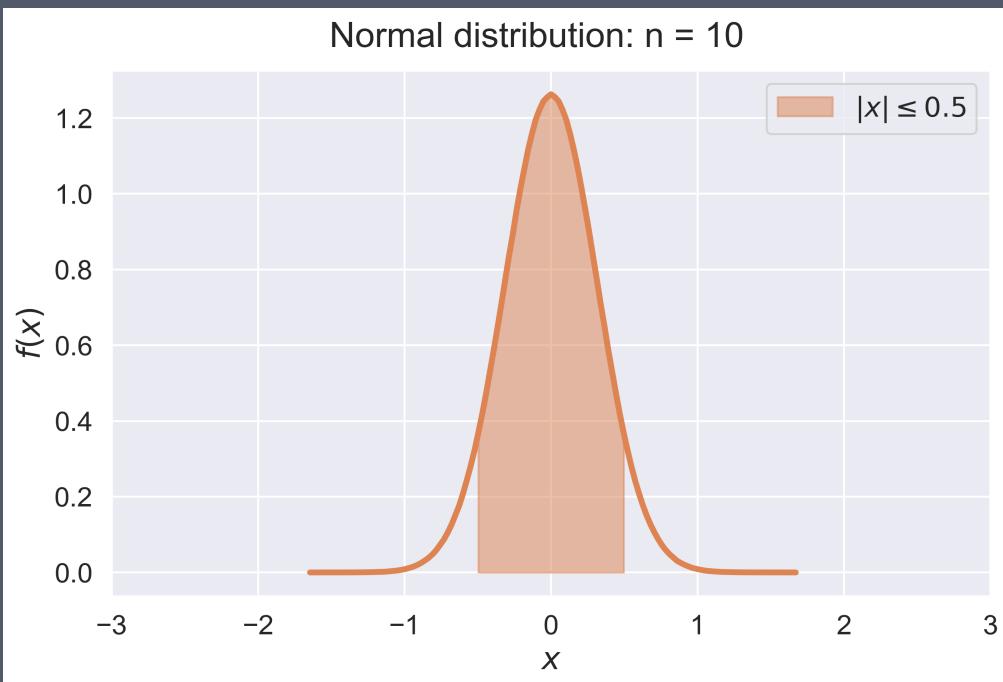
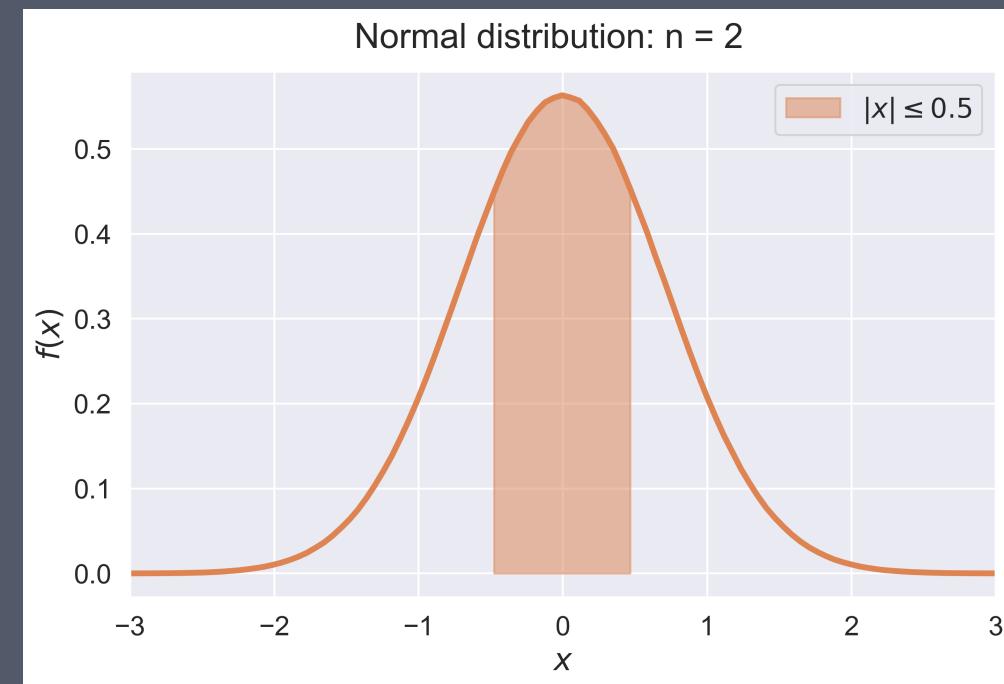
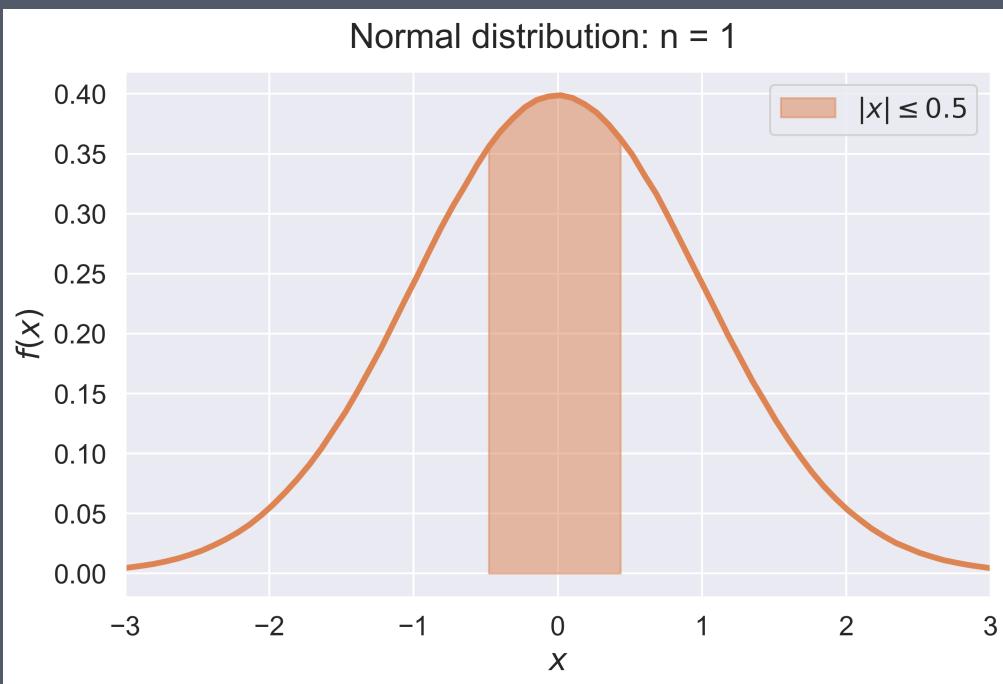
$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

$$\mu = 0, \sigma^2 = \frac{1}{n}$$

Variance

$$V\left(\frac{S_n}{n}\right) = \frac{1}{n}$$

$$P\left(\left|\frac{S_n}{n}\right| \geq \varepsilon\right) \leq \frac{1}{n\varepsilon^2}$$



## Example 6

- Suppose we choose n numbers from  $(-\infty, +\infty)$  with a Cauchy density with parameter  $a = 1$ .
- We know that for the Cauchy density the expected value and variance are undefined.

### Cauchy distribution

$$f(x) = \frac{1}{\pi a(1 + \frac{x}{a})^2}$$

$$f(x) = \frac{1}{\pi(1 + x)^2}$$

Expected value  $E(X)$

undefined

Variance  $V(X)$

undefined

## Example 6

- In this case, the density function for  $A_n = \frac{s_n}{n}$  is given by

Cauchy distribution

$$f(x) = \frac{1}{\pi(1+x)^2}$$

The density function for  $A_n$  is the same for all  $n$ .

!

The Law of Large Numbers does not hold.

!

Why?

## Example 6

- In this case, the density function for  $A_n = \frac{s_n}{n}$  is given by

Cauchy distribution

$$f(x) = \frac{1}{\pi(1+x)^2}$$

The density function for  $A_n$  is the same for all  $n$ .

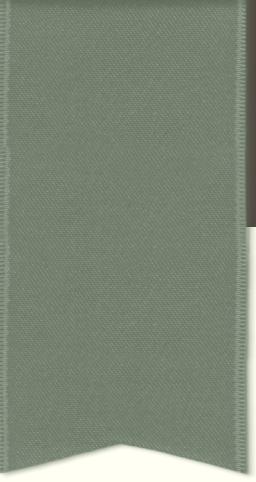
!

The Law of Large Numbers does not hold.

!

Why?

finite expected value and finite variance!

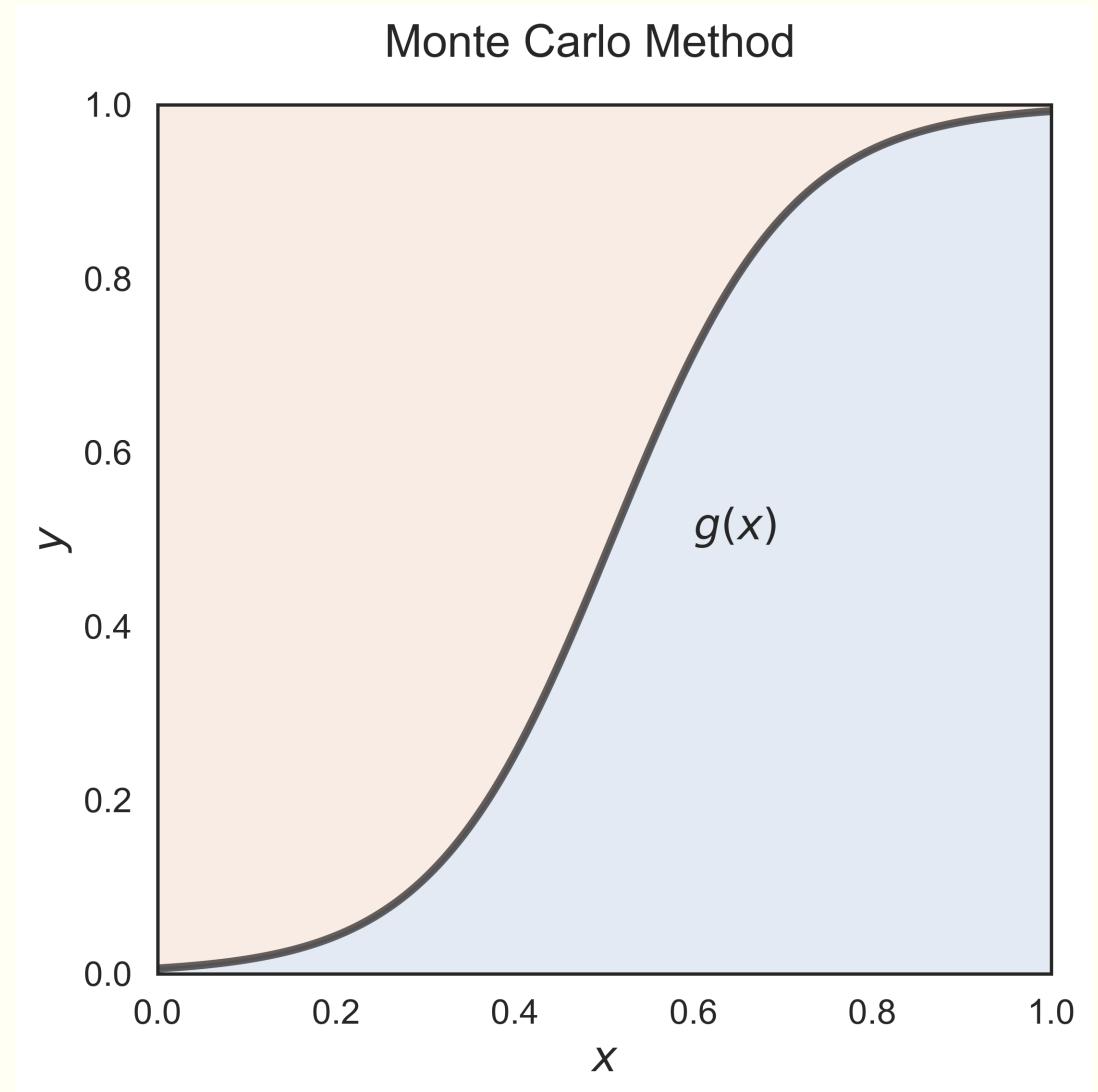


# MONTE CARLO METHOD

estimate the area of the region

# Monte Carlo method

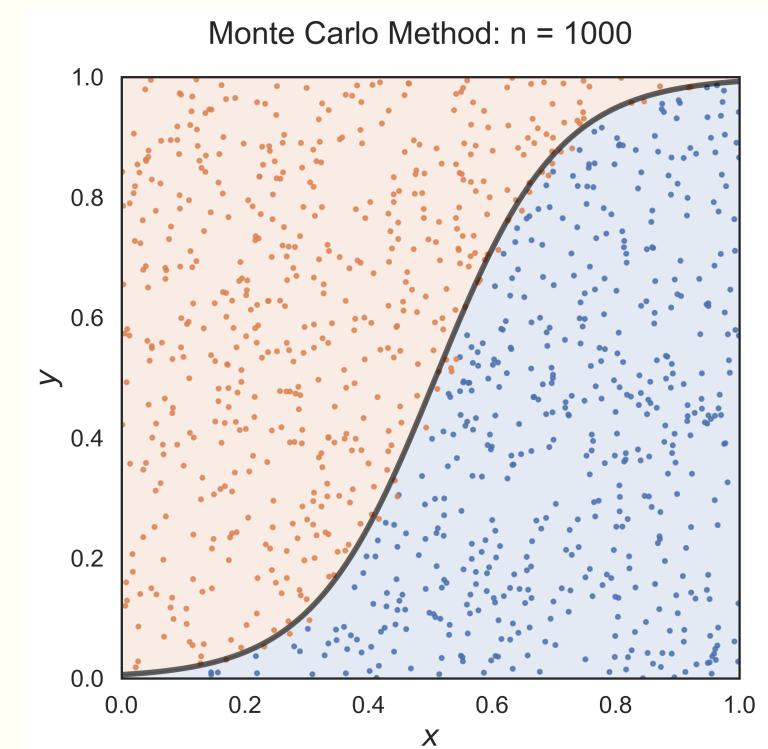
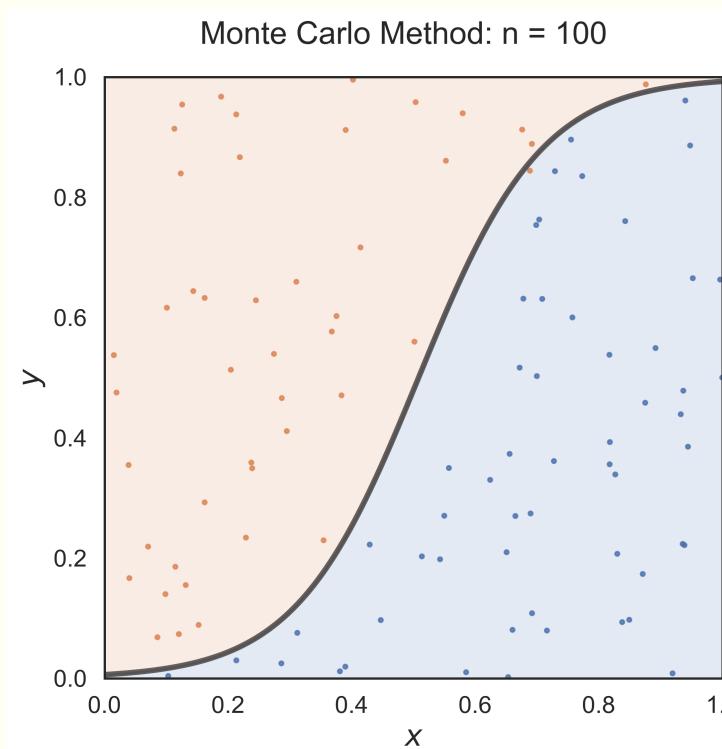
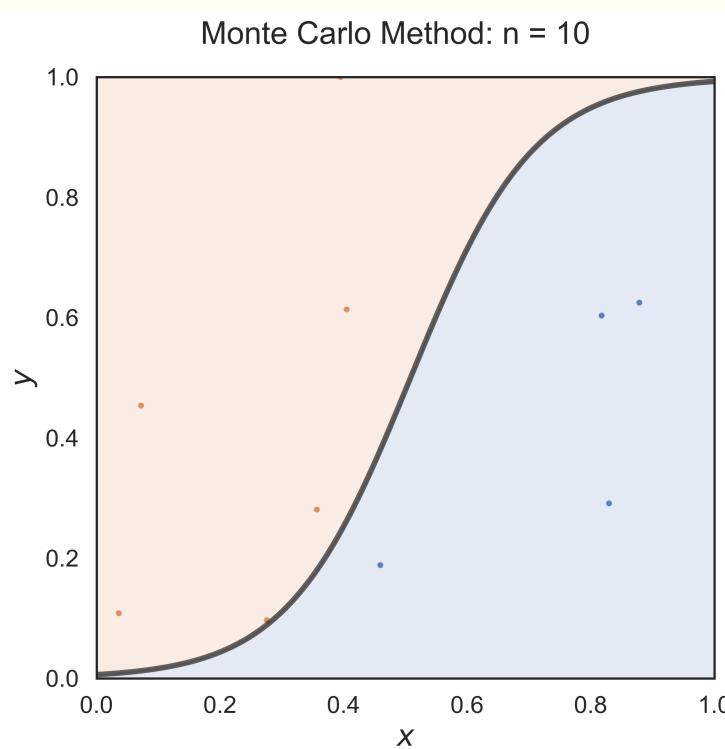
- Let  $g(x)$  be continuous function defined for  $x \in [0, 1]$  with values in  $[0, 1]$ .
- How to estimate the area of the region under the graph of  $g(x)$ ?



# Monte Carlo method: naive

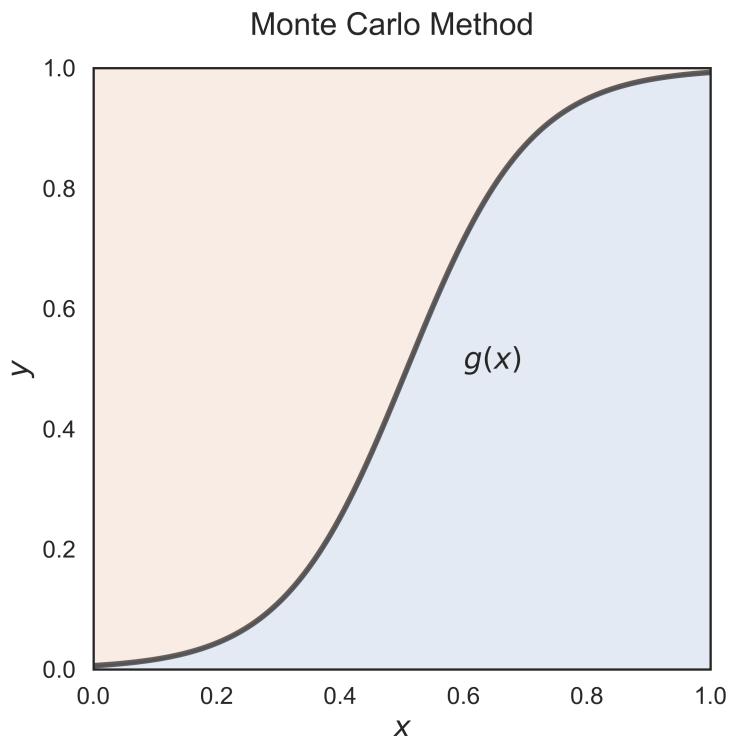
- Let  $g(x)$  be continuous function defined for  $x \in [0, 1]$  with values in  $[0, 1]$ .
- How to estimate the area of the region under the graph of  $g(x)$ ?

Choose a large number of random values for  $x$  and  $y$  with uniform distribution and seeing what fraction of the points  $P(x, y)$  fall inside the region under the graph.



# Monte Carlo method: advanced

- Let  $g(x)$  be continuous function defined for  $x \in [0, 1]$  with values in  $[0, 1]$ .
- How to estimate the area of the region under the graph of  $g(x)$ ?



Choose a large number of independent values  $X_n$  at random from  $[0, 1]$  with uniform density.

Set  $Y_n = g(X_n)$ .

Area:  $A$

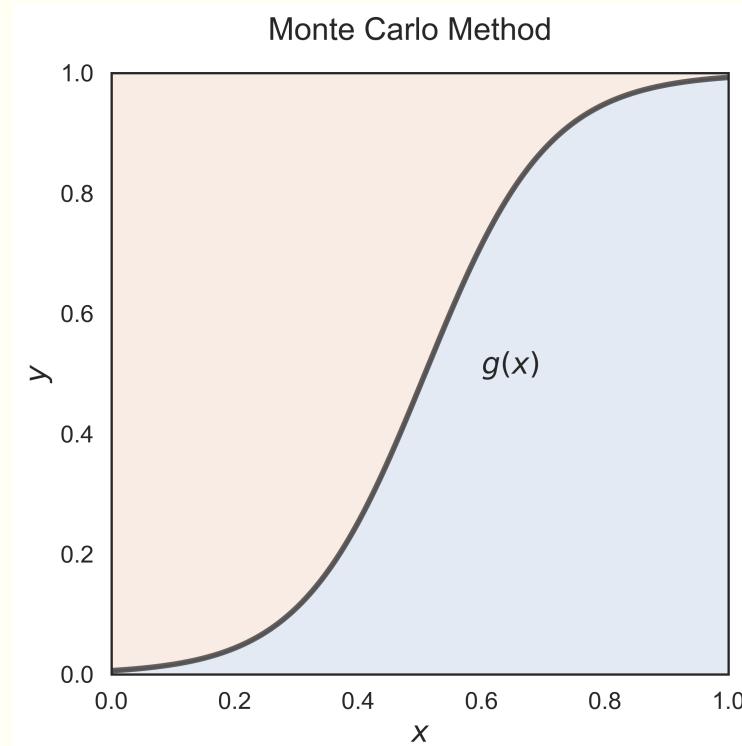
$$\int_0^1 g(x)dx$$

Expected value:  $\mu$

$$\begin{aligned} E(Y_n) &= E(g(X_n)) \\ &= \int_0^1 g(x)f(x)dx \\ &= \int_0^1 g(x)dx \end{aligned}$$

# Monte Carlo method: advanced

- Let  $g(x)$  be continuous function defined for  $x \in [0, 1]$  with values in  $[0, 1]$ .
- How to estimate the area of the region under the graph of  $g(x)$ ?



Choose a large number of independent values  $X_n$  at random from  $[0, 1]$  with uniform density.

Set  $Y_n = g(X_n)$ .

Area:  $A$

Expected value:  $\mu$

$$\int_0^1 g(x)dx$$

$$E(Y_n) = \\ = \int_0^1 g(x)dx$$

$$A = \mu \approx \frac{1}{n} \sum_{i=1}^n Y_n = \frac{1}{n} \sum_{i=1}^n g(X_n)$$

# Monte Carlo method: advanced

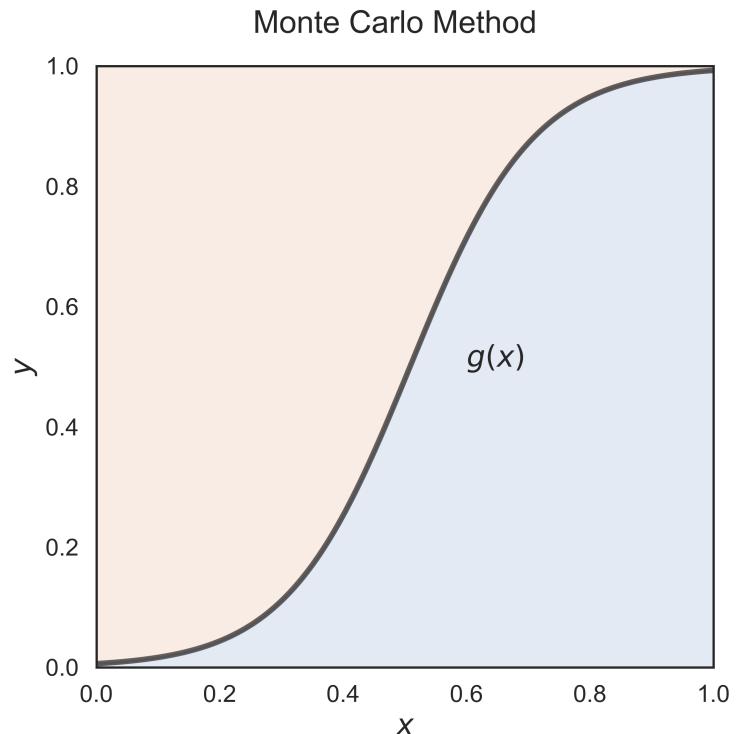
- Let  $g(x)$  be continuous function defined for  $x \in [0, 1]$  with values in  $[0, 1]$ .
- How to estimate the area of the region under the graph of  $g(x)$ ?

Area:  $A$

Expected value:  $\mu$

$$\int_0^1 g(x)dx$$

$$E(Y_n) = \\ = \int_0^1 g(x)dx$$



$$A = \mu \approx \frac{1}{n} \sum_{i=1}^n Y_n = \frac{1}{n} \sum_{i=1}^n g(X_n) = A_n$$

$$\sigma^2 = E((Y_n - \mu)^2) \\ = \int_0^1 (g(x) - \mu)^2 dx \leq 1$$

$$P(|A_n - A| \geq \varepsilon) \\ \leq \frac{\sigma^2}{n\varepsilon^2} \leq \frac{1}{n\varepsilon^2}$$



$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$



## Quiz 12 Question 3

A student's score on a particular probability final is a random variable with values of  $[0, 100]$ , mean 70, and variance 25.

Using Chebyshev's Inequality, find a lower bound for the probability that the student's score will fall between 65 and 75.

$$\begin{aligned} P(65 \leq X \leq 75) &= P(|X - 70| \leq 5) = 1 - P(|X - 70| \geq 5) \\ &\geq 1 - \frac{25}{5^2} = 1 - 1 = 0 \end{aligned}$$

not interesting



$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$



## Quiz 12 Question 4

A student's score on a particular probability final is a random variable with values of  $[0, 100]$ , mean 70, and variance 25.

If 10 students take the final, find a lower bound for the probability that the class average will fall between 65 and 75.

$$\begin{aligned} P(65 \leq A_n \leq 75) &= P(|A_n - 70| \leq 5) = 1 - P(|A_n - 70| \geq 5) \\ &\geq 1 - \frac{25/10}{5^2} = 1 - \frac{1}{10} = \frac{9}{10} \end{aligned}$$