# Chapter 5: Law of Large Numbers and Limit Theorem

1. The Law of Large Numbers (大数定律)

2. The Central Limit Theorem (中心极限定理)
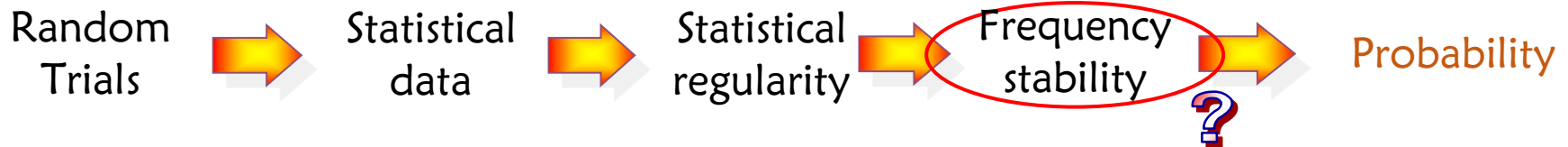
## Chapter 5: Law of Large Numbers and Limit Theorem

1.　The Law of Large Numbers (大数定律)
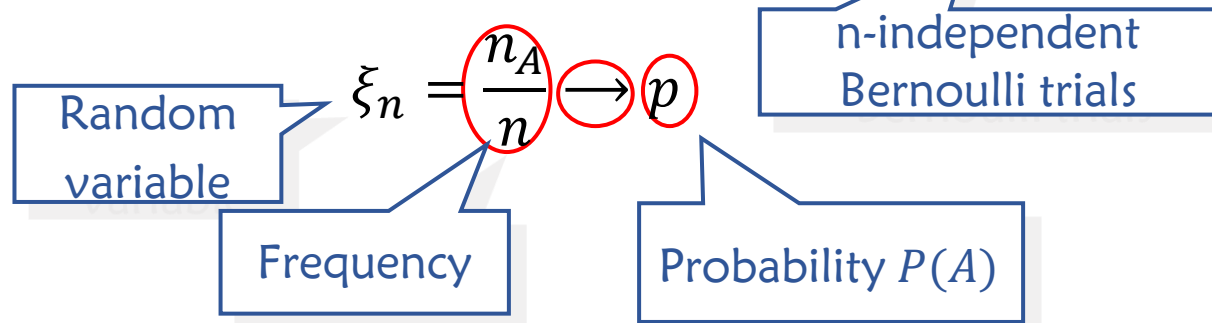
2.　The Central Limit Theorem (中心极限定理)

# Background of The Law of Large Numbers

⬤ Question: How does the concept of "probability" evolve ❓

Random Trials ➡ Statistical data ➡ Statistical regularity ➡ (Frequency stability) ➡ Probability ❓

**Frequency stability (频率稳定性):**

Assume that the number of times event A happens during n repeated independent trials is $n_A$. When $n \to \infty$,

$$\xi_n = \frac{n_A}{n} \longrightarrow p$$

Random variable

Frequency

Probability $P(A)$

n-independent Bernoulli trials

n repeated independent

⬤ Question: How to define the limit (极限) $\lim\limits_{n\to\infty} \xi_n = p$ ❓

⬤ Question: What is the strict mathematical description of "Frequency stability" ❓

(频率稳定性)

**Example:** Coin tossing trials  Toss a coin n times, let

$$A = \{\text{Head}\}$$

$$n_A = \text{The total number of A in n trials}$$

The frequency of $A$ is

$$\xi_n = \frac{n_A}{n} \ (n = 1, 2, \cdots)$$

### 18th-19th century famous coin tossing tests

| Tester | $n$ | $n_A$ | $\xi_n$ |
|--------|------|-------|---------|
| Buffon | 4048 | 2048 | 0.5069 |
| De Morgan | 2048 | 1061 | 0.5181 |
| Pearson | 12000 | 6019 | 0.5016 |
| Pearson | 24000 | 12012 | 0.5005 |

**Frequency stability (频率稳定性):**  $\xi_n \to 0.5 \ (n \to \infty)$

**Example:** Coin tossing trials   Toss a coin n times, let

$$A = \{\text{Head}\}$$

$$n_A = \text{The total number of } A \text{ in n trials}$$

The frequency of $A$ is

$$\xi_n = \frac{n_A}{n} \ (n = 1, 2, \cdots)$$

① $\{\xi_n\}_{n=1}^{\infty}$ is a list of random variables (随机变量列).
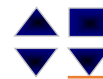
**Simulation trials of n=4048**

Trial results: ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● …… ●

$\xi_n = \frac{n_A}{n}:$  $\frac{1}{1}$  $\frac{1}{2}$  $\frac{1}{3}$  $\frac{2}{4}$  $\frac{3}{5}$  $\frac{3}{6}$  $\frac{3}{7}$  $\frac{3}{8}$  $\frac{4}{9}$  $\frac{4}{10}$  $\frac{5}{11}$  $\frac{6}{12}$  $\frac{6}{13}$  $\frac{7}{14}$  $\frac{7}{15}$  $\frac{7}{16}$  ……  $\frac{2021}{4048}$

**Is it possible to have the following results?**

● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● …… ●

● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● …… ●

② $\{\xi_n = \xi_n(\omega)\}_{n=1}^{\infty}$ is a list of functions (函数列) defined on the sample space $\Omega$.

**Example:** Coin tossing trials  Toss a coin n times, let

$$A = \{\text{Head}\}$$

$$n_A = \text{The total number of A in n trials}$$

The frequency of $A$ is

$$\xi_n = \frac{n_A}{n} \ (n = 1, 2, \cdots)$$

① $\{\xi_n\}_{n=1}^{\infty}$ is a list of random variables (随机变量列).

② $\{\xi_n = \xi_n(\omega)\}_{n=1}^{\infty}$ is a list of functions (函数列) defined on the sample space $\Omega$.

## Review the convergence (收敛性) of a function list in calculus

If functions $f(x), f_n(x), (n = 1, 2, \dots)$ are defined on the interval $(a, b)$.

Then $f_n(x)$ converges to $f(x)$ means that for $\forall x \in (a, b)$:

$$\lim_{n \to \infty} f_n(x) = f(x)$$

Pointwise convergence
(逐点收敛)

For a list of random variables (随机变量列), do we have

$$\lim_{n \to \infty} \xi_n(\omega) = p \ (\forall \omega \in \Omega)$$

?

**Definition:** Assume that $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ is a list of random variables. If for $\forall \varepsilon > 0$:
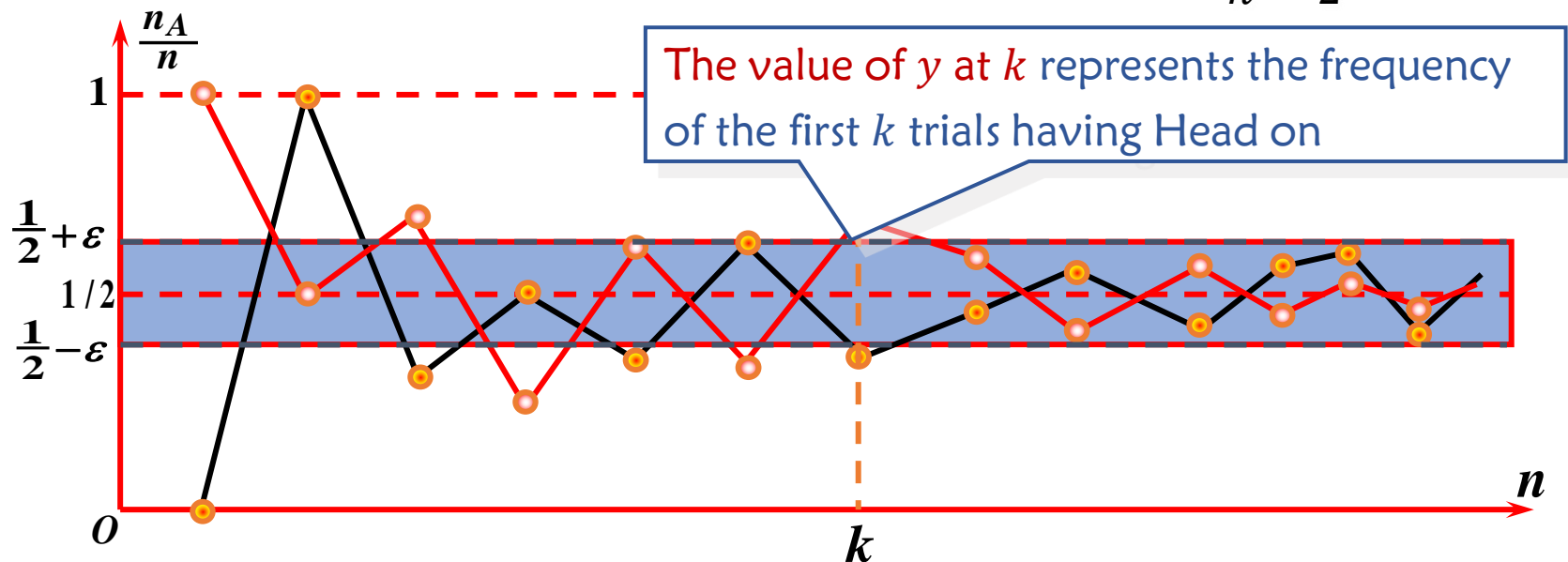
$$\lim_{n \to \infty} P\{|\xi_n - \xi| \geq \varepsilon\} = 0$$

Then $\{\xi_n\}$ is said to converge in probability (依概率收敛) to $\xi$, denoted as $\xi_n \overset{P}{\to} \xi$.

① $\lim_{n \to \infty} P\{|\xi_n - \xi| \geq \varepsilon\} = 0 \iff \lim_{n \to \infty} P\{|\xi_n - \xi| < \varepsilon\} = 1$

② The meaning of $\xi_n \overset{P}{\to} \xi$ : as $n$ increases, the probability that the absolute error (绝对误差) $|\xi_n - \xi|$ being large will become smaller.

③ The frequency stability of a coin tossing trials is $\xi_n = \dfrac{n_A}{n} \overset{P}{\to} \dfrac{1}{2}$.



The value of $y$ at $k$ represents the frequency of the first $k$ trials having Head on

## Theorem 1: Bernoulli's law of large numbers (伯努利大数定律)

Assume that $n_A$ is the number of times event A happens in n trials, and $P(A) = p$. Then for $\forall \varepsilon > 0$,

$$\lim_{n \to \infty} P\left\{ \left| \frac{n_A}{n} - p \right| \geq \varepsilon \right\} = 0.$$

How to prove ?

**Analysis** Let

$$X_i = \begin{cases} 1, & A \text{ happened at the } i\text{th trial} \\ 0, & A \text{ didn't happend at the } i\text{th trail} \end{cases}, i = 1, 2, \cdots$$

Then ① $X_1, X_2, \ldots, X_n$ are independent.

② $E(X_i) = p, D(X_i) = p(1-p), (i = 1, 2, \ldots)$

③ $\dfrac{n_A}{n} = \dfrac{1}{n} \sum_{i=1}^{n} X_i \ (n = 1, 2, \cdots)$

$\{X_n\}$ is a list of independent random variables (随机变量列); $\{X_i\}$ has the same mathematical expectation and variance.

**Therefore**

$$\lim_{n \to \infty} P\left\{ \left| \frac{n_A}{n} - p \right| \geq \varepsilon \right\} = 0. \iff \lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{i=1}^{n} X_i - p \right| \geq \varepsilon \right\} = 0.$$

## Theorem 1: Bernoulli's law of large numbers (伯努利大数定律)

Assume that $n_A$ is the number of times event A happens in n trials, and $P(A) = p$. Then for $\forall \varepsilon > 0$,

$$\lim_{n \to \infty} P\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0.$$

## Theorem 2: Chebyshev's law of large numbers (切比雪夫大数定律)

Assume that $\{X_n\}$ is a list of independent random variables and they have the same mathematical expectation and variance:

$$E(X_i) = \mu, D(X_i) = \sigma^2, i = 1, 2, \cdots$$

Then for $\forall \varepsilon > 0$:

$$\lim_{n \to \infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right\} = 0.$$

### Recall the Chebyshev's inequality

Assume that the variance of the random variable $\xi$ exists, then for $\forall \varepsilon > 0$:

$$P\{|\xi - E(\xi)| \geq \varepsilon\} \leq \frac{D(\xi)}{\varepsilon^2}$$

## Theorem 1: Bernoulli's law of large numbers (伯努利大数定律)

Assume that $n_A$ is the number of times event A happens in n trials, and $P(A) = p$. Then for $\forall \varepsilon > 0$,

$$\lim_{n \to \infty} P\left\{ \left| \frac{n_A}{n} - p \right| \geq \varepsilon \right\} = 0.$$

## Theorem 2: Chebyshev's law of large numbers (切比雪夫大数定律)

Assume that $\{X_n\}$ is a list of independent random variables and they have the same mathematical expectation and variance:

$$E(X_i) = \mu, D(X_i) = \sigma^2, i = 1, 2, \cdots$$

Then for $\forall \varepsilon > 0$:

$$\lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| \geq \varepsilon \right\} = 0.$$

Note   Both Bernoulli's LLN and Chebyshev's LLN require the **same variance** of a random variable list. It can be replaced by "identically distributed".

## Theorem 1: Bernoulli's law of large numbers (伯努利大数定律)

Assume that $n_A$ is the number of times event A happens in n trials, and $P(A) = p$. Then for $\forall \varepsilon > 0$,

$$\lim_{n \to \infty} P\left\{ \left| \frac{n_A}{n} - p \right| \geq \varepsilon \right\} = 0.$$

## Theorem 2: Chebyshev's law of large numbers (切比雪夫大数定律)

Assume that $\{X_n\}$ is a list of independent random variables and they have the same mathematical expectation and variance:

$$E(X_i) = \mu, D(X_i) = \sigma^2, i = 1, 2, \cdots$$

Then for $\forall \varepsilon > 0$:

$$\lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| \geq \varepsilon \right\} = 0.$$

## Theorem 3: Khinchine's law of large numbers (辛钦大数定律)

Assume that $\{X_n\}$ is an independent and identically distributed r.v. list, and $E(X_1) \triangleq \mu$ exists. Then $\{X_n\}$ follows the law of large numbers. That is, for $\forall \varepsilon > 0$:

$$\lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| < \varepsilon \right\} = 1 \quad \text{or} \quad \lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| \geq \varepsilon \right\} = 0.$$

## The significance of the Law of Large Numbers

① It gives the strict mathematical explanation for "frequency stability" (频率稳定性)

$$\lim_{n\to\infty} P\{|\xi_n - \xi| \geq \varepsilon\} = 0 \Longleftrightarrow \lim_{n\to\infty} P\{|\xi_n - \xi| < \varepsilon\} = 1$$

② It provides a method to find out the probability of an event by using trials .

③ It is one of the major theoretical basis for parameter estimation in mathematical statistics.

④ It is the major mathematical theory basis of the Monte Carlo method (蒙特卡洛法).

# Application of The Law of Large Numbers
## --The Monte Carlo Method
### (蒙特卡洛法)

**The Monte Carlo Method (**or random computational simulation method, computational simulation method**)** is one of the most important tools for science and engineering.

The principle of the Monte Carlo Method is the Law of Large numbers.

**Example:** There is a rectangle G. Assume that its area is 1. Draw a closed curve L within the rectangle. What is the area |D| in the enclosed area by L?
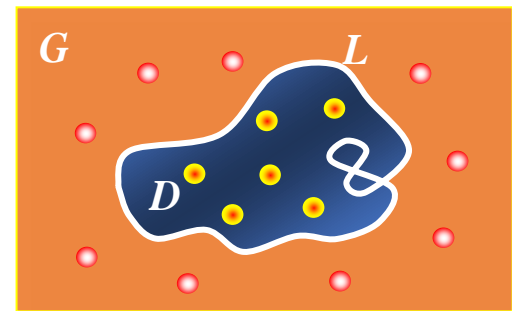
Not possible?!
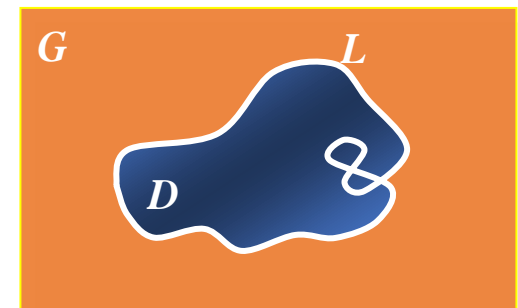
**Example:** There is a rectangle G. Assume that its area is 1. Draw a closed curve L within the rectangle. What is the area |D| in the enclosed area by L?

**Answer:** Use a computer to generate random points (random variables) which are independent, and each random variable follows the uniform distribution on G:

$$(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$$



Denote event $A$ as $A = \{$a randomly generated point falls into $D\}$ and let $n_A$ be the number of times event $A$ happens among $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$.

According to the Bernoulli's Law of Large Numbers

$$\frac{n_A}{n} \xrightarrow{P} P(A) = \frac{\text{The area of } D}{\text{The area of } G} = |D|$$

Thus, if n is large enough, the area of D is $|D| \approx \dfrac{n_A}{n}$.

This is called the Monte Carlo method (蒙特卡洛法)

# Why is called "The Law (定律) of Large Numbers" not "The Theorem (定理) of Large Numbers" ❓

At the early age of **Probability Theory**, the mathematical definition of probability was not clear. Therefore, there was no theoretical basis to understand the concept of **convergence in probability**, the phenomenon that frequency "tending to" probability was considered as the results from **a large number of trials**, just like the laws in physics.

Since an axiomatic system (公理化体系) of probability was established and the law of large numbers can be strictly proved as a **theorem**, then the Law of Large Numbers has been also named as the Theorem of Large Numbers.

谢谢大家