



Igor Kortchemski – ETH Zürich

Probability Theory Autumn 2023





# I. A DREAM

Igor Kortchemski Predicting the Unpredictable

1 / 672



I was on the sand, a thousand miles from any human habitation. I was more isolated than a shipwrecked sailor on a raft in the middle of the ocean.











"If you please — draw me a function at random!"





"If you please — draw me a function at random!"

"What!"





"If you please — draw me a function at random!"

"What!"

"Draw me a function at random!"







"But — what are you doing here?"





"But — what are you doing here?"

And in answer he repeated, very slowly, as if he were speaking of a matter of great consequence:





"But — what are you doing here?"

And in answer he repeated, very slowly, as if he were speaking of a matter of great consequence:

"If you please — draw me a function at random..."







But then I remembered how my studies at ETH Zürich had been concentrated on geography, history, arithmetic, and grammar,



But then I remembered how my studies at ETH Zürich had been concentrated on geography, history, arithmetic, and grammar, but not Probability theory,



But then I remembered how my studies at ETH Zürich had been concentrated on geography, history, arithmetic, and grammar, but not Probability theory, and I told the little chap (a little crossly, too) that I did not know how to draw at random.



But then I remembered how my studies at ETH Zürich had been concentrated on geography, history, arithmetic, and grammar, but not Probability theory, and I told the little chap (a little crossly, too) that I did not know how to draw at random. He answered me:



But then I remembered how my studies at ETH Zürich had been concentrated on geography, history, arithmetic, and grammar, but not Probability theory, and I told the little chap (a little crossly, too) that I did not know how to draw at random. He answered me:

"That doesn't matter.





But then I remembered how my studies at ETH Zürich had been concentrated on geography, history, arithmetic, and grammar, but not Probability theory, and I told the little chap (a little crossly, too) that I did not know how to draw at random. He answered me:

"That doesn't matter. Draw me a function at random..."



But I had never drawn a random. So I drew for him one of the two pictures I had drawn so often.













"No, no, no! I do not want a function that doesn't wiggle.





"No, no, no! I do not want a function that doesn't wiggle. I need a random function.

5 / 672





"No, no, no! I do not want a function that doesn't wiggle. I need a random function. Draw me a function at random."















He looked at it carefully, then he said:







He looked at it carefully, then he said: "No. This one is predictable.







He looked at it carefully, then he said: "No. This one is predictable. Make me another"

For an integer  $n \ge 1$ , define

 $\mathfrak{X}_{n} = \{(c_{0}, c_{1}, \ldots, c_{n}) \in \mathbb{Z}^{n+1} : c_{0} = 0 \text{ and } |c_{i+1} - c_{i}| = 1 \text{ for } 0 \leqslant i \leqslant n-1\}.$ 



Figure: The path (0, 1, 0, -1, -2, -1, -2).

For an integer  $n \ge 1$ , define  $\mathfrak{X}_n = \{(c_0, c_1, \dots, c_n) \in \mathbb{Z}^{n+1} : c_0 = 0 \text{ and } |c_{i+1} - c_i| = 1 \text{ for } 0 \leqslant i \leqslant n-1\}.$ 

And I threw out an explanation with it:

"This is a finite set, with  $2^n$  elements. Choose a function uniformly at random inside. This is the function you asked for."



For an integer  $n \ge 1$ , define 
$$\begin{split} & \chi_n = \{(c_0, c_1, \dots, c_n) \in \mathbb{Z}^{n+1} : c_0 = 0 \text{ and } |c_{i+1} - c_i| = 1 \text{ for } 0 \leqslant i \leqslant n-1 \}. \end{split}$$

And I threw out an explanation with it:

"This is a finite set, with  $2^n$  elements. Choose a function uniformly at random inside. This is the function you asked for."

I was very surprised to see a light break over the face of my young judge:

"That is exactly the way I wanted it!"

For n=1000, I chose a path uniformy at random in the set  $\chi_{1000}$  and I drew it. What did I get?

For n=1000, I chose a path uniformy at random in the set  $\chi_{1000}$  and I drew it. What did I get?



#### Figure: Drawing 1.






Figure: Drawing 1.

Figure: Drawing 2.





Figure: Drawing 1.



Figure: Drawing 3.

Figure: Drawing 2.







Figure: Drawing 3.



Figure: Drawing 2.



Figure: Drawing 4.











Figure: Drawing 2.



Figure: Drawing 4.

**/ →** wooclap.com | ; code **probability**.



Figure: Drawing 3.

Figure: Drawing 4.

**∧→** wooclap.com | ; code **probability**.







Probability theory is more concerned with the realm of the probable, rather than the realm of the possible.

8 / 672

Then, abruptly, without anything to lead up to it, and as if the question had been born of long and silent meditation on his problem, he demanded:

Then, abruptly, without anything to lead up to it, and as if the question had been born of long and silent meditation on his problem, he demanded:

"What happens when  $n \to \infty$ ?"



So, I told him:



So, I told him:

Let E be a Polish space and denote by  $\mathcal{M}_1(E)$  the space of all probability measures on E equipped with its Borel  $\sigma$ -field.

Consider the weak-\* topology  $\mathcal{M}_1(E)$  for which the convergence  $\mu_n \to \mu$  is given by

 $\forall f \in \mathcal{C}_b(E), \qquad \mu_n(f) \to \mu(f).$ 









Goal of the course:

 $\mathcal{N}$  Study stochastic processes in discrete time, that is finite sequences  $X_0, \ldots, X_n$  of random variables, from a mathematical point of view (with measure theory).



Goal of the course:

 $\bigwedge$  Study stochastic processes in discrete time, that is finite sequences  $X_0, \ldots, X_n$  of random variables, from a mathematical point of view (with measure theory). Main focus of interest: as  $n \to \infty$ , asymptotic behavior of  $X_n$ , or functions of  $(X_1, \ldots, X_n)$ .



Goal of the course:

 $\bigwedge$  Study stochastic processes in discrete time, that is finite sequences  $X_0, \ldots, X_n$  of random variables, from a mathematical point of view (with measure theory). Main focus of interest: as  $n \to \infty$ , asymptotic behavior of  $X_n$ , or functions of  $(X_1, \ldots, X_n)$ .

This course is designed for an audience with a variety of interests:



Goal of the course:

 $\bigwedge$  Study stochastic processes in discrete time, that is finite sequences  $X_0, \ldots, X_n$  of random variables, from a mathematical point of view (with measure theory). Main focus of interest: as  $n \to \infty$ , asymptotic behavior of  $X_n$ , or functions of  $(X_1, \ldots, X_n)$ .

This course is designed for an audience with a variety of interests:

 $\wedge$  for students wishing to deepen their mathematical knowledge of probability theory;



Goal of the course:

 $\bigwedge$  Study stochastic processes in discrete time, that is finite sequences  $X_0, \ldots, X_n$  of random variables, from a mathematical point of view (with measure theory). Main focus of interest: as  $n \to \infty$ , asymptotic behavior of  $X_n$ , or functions of  $(X_1, \ldots, X_n)$ .

This course is designed for an audience with a variety of interests:

 $\wedge$  for students wishing to deepen their mathematical knowledge of probability theory;

 $\wedge \rightarrow$  for students who intend to use it in business applications (a good understanding of probability theory is essential in order to be able to orient oneself in the world of applications and to innovate there).



#### I. A DREAM

#### **II.** LAW OF LARGE NUMBERS



#### Imagine that you are given a coin. How can you find out if it is rigged or not?



#### Imagine that you are given a coin. How can you find out if it is rigged or not?



Throw n times in a row the coin.



A bit more formally, for  $i \ge 1$  set  $X_i = 1$  if the i-th throw is heads (happens with a certain probability p) and 0 if it is tails (happens with probability 1 - p). Set  $S_n = X_1 + \dots + X_n$ .

A bit more formally, for  $i \ge 1$  set  $X_i = 1$  if the i-th throw is heads (happens with a certain probability p) and 0 if it is tails (happens with probability 1 - p). Set  $S_n = X_1 + \cdots + X_n$ .

 $\longrightarrow$  What is the behavior of  $\frac{S_n}{n}$  as  $n \to \infty$ ?



A bit more formally, for  $i \ge 1$  set  $X_i = 1$  if the i-th throw is heads (happens with a certain probability p) and 0 if it is tails (happens with probability 1 - p). Set  $S_n = X_1 + \cdots + X_n$ .

 $\bigwedge$  What is the behavior of  $\frac{S_n}{n}$  as  $n \to \infty$ ?



Figure: Simulation of  $\left(\frac{S_n}{n}: 1 \leq n \leq 10\right)$ .

A bit more formally, for  $i \ge 1$  set  $X_i = 1$  if the i-th throw is heads (happens with a certain probability p) and 0 if it is tails (happens with probability 1 - p). Set  $S_n = X_1 + \cdots + X_n$ .

 $\bigwedge$  What is the behavior of  $\frac{S_n}{n}$  as  $n \to \infty$ ?



Figure: Simulation of  $\left(\frac{S_n}{n}: 1 \le n \le 100\right)$ .

A bit more formally, for  $i \ge 1$  set  $X_i = 1$  if the i-th throw is heads (happens with a certain probability p) and 0 if it is tails (happens with probability 1 - p). Set  $S_n = X_1 + \cdots + X_n$ .

 $\longrightarrow$  What is the behavior of  $\frac{S_n}{n}$  as  $n \to \infty$ ?



Figure: Simulation of  $\left(\frac{S_n}{n}: 1 \leq n \leq 100\right)$ .

A bit more formally, for  $i \ge 1$  set  $X_i = 1$  if the i-th throw is heads (happens with a certain probability p) and 0 if it is tails (happens with probability 1 - p). Set  $S_n = X_1 + \cdots + X_n$ .



Figure: Simulation of  $\left(\frac{S_n}{n}: 1 \le n \le 10000\right)$ 

A bit more formally, for  $i \ge 1$  set  $X_i = 1$  if the i-th throw is heads (happens with a certain probability p) and 0 if it is tails (happens with probability 1 - p). Set  $S_n = X_1 + \cdots + X_n$ .

 $\bigwedge$  What is the behavior of  $\frac{S_n}{n}$  as  $n \to \infty$ ?



Figure: 10 simulations of  $\left(\frac{S_n}{n}: 1 \leq n \leq 1000\right)$ .

#### I. A DREAM

#### **II.** LAW OF LARGE NUMBERS

#### **III.** CENTRAL LIMIT THEOREM



A bit more formally, for  $i \ge 1$  set  $X_i = 1$  if the i-th throw is heads (happens with a certain probability p) and 0 if it is tails (happens with probability 1 - p). Set  $S_n = X_1 + \dots + X_n$ .

 $\longrightarrow$  What is the behavior of  $\frac{S_n}{n}$  as  $n \to \infty$ ?



Figure: 10 simulations of  $\left(\frac{S_n}{n}: 1 \le n \le 1000\right)$ .

 $\bigwedge \rightarrow Law$  of large numbers:  $\frac{S_n}{n}$  converges **almost surely** towards p as  $n \rightarrow \infty$ .



Figure: 10 simulations of  $\left(\frac{S_n}{n} - p : 1 \le n \le 1000\right)$  for p = 0.6.





Figure: 10 simulations of  $\left(\frac{S_n}{n} - p : 1 \le n \le 1000\right)$  for p = 0.6.

 $\wedge$  Can we "zoom in"?





Figure: 10 simulations of  $\left(\frac{S_n}{n} - p : 1 \le n \le 1000\right)$  for p = 0.6.

 $\wedge$  Can we "zoom in"?

 $\bigwedge$  Is there a function f(n) such that  $f(n)\left(\frac{S_n}{n}-p\right)$  has a nice behavior for n large?



The speed of convergence is  $\frac{1}{\sqrt{n}}$ . This means that  $\sqrt{n}(\frac{S_n}{n} - p)$  has a non-degenerate behavior as  $n \to \infty$ .



The speed of convergence is  $\frac{1}{\sqrt{n}}$ . This means that  $\sqrt{n}(\frac{S_n}{n} - p)$  has a non-degenerate behavior as  $n \to \infty$ .



Figure: Simulation of  $\left(\sqrt{n} \cdot \left(\frac{S_n}{n} - p\right) : 1 \le n \le 10000\right)$  for p = 0.6.



The speed of convergence is  $\frac{1}{\sqrt{n}}$ . This means that  $\sqrt{n}(\frac{S_n}{n} - p)$  has a non-degenerate behavior as  $n \to \infty$ .



Figure: Another simulation of  $\left(\sqrt{n} \cdot \left(\frac{s_n}{n} - p\right) : 1 \le n \le 10000\right)$  for p = 0.6.


# The Central Limit Theorem





# The Central Limit Theorem

The speed of convergence is  $\frac{1}{\sqrt{n}}$ . This means that  $\sqrt{n}(\frac{S_n}{n} - p)$  has a non-degenerate behavior as  $n \to \infty$ .





# The Central Limit Theorem

The speed of convergence is  $\frac{1}{\sqrt{n}}$ . This means that  $\sqrt{n}(\frac{S_n}{n} - p)$  has a non-degenerate behavior as  $n \to \infty$ .









There is structure in randomness! Look at the "endpoints"  $\sqrt{n}(\frac{S_n}{n} - p)$ 



18 / 672

# Structure in randomness

There is structure in randomness! Look at the "endpoints"  $\sqrt{n}(\frac{S_n}{n} - p)$  and draw the empirical histogram.



There is structure in randomness! Look at the "endpoints"  $\sqrt{n}(\frac{S_n}{n} - p)$  and draw the empirical histogram.





Figure: Empirical histograms of 10000 simulations of  $\sqrt{n} \cdot \left(\frac{S_n}{n} - p\right)$  for n = 10000. Left: p = 0.6; Right:p = 0.4.





Figure: Empirical histograms of 10000 simulations of  $\sqrt{n} \cdot \left(\frac{S_n}{n} - p\right)$  for n = 10000. Left: p = 0.6; Right:p = 0.4.











Figure: Plot of the function  $x \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ .

**Théorème** (Central limit theorem – De Moivre Laplace theorem).

Let  $S_n$  be the sum of n independent Bernoulli random variables of parameter  $p \in (0, 1)$ . Then, for every a < b:

$$\mathbb{P}\left(a \leqslant \frac{\sqrt{n}}{\sqrt{p(1-p)}} \left(\frac{S_n}{n} - p\right) \leqslant b\right) \quad \underset{n \to \infty}{\longrightarrow} \quad \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$



Figure: Plot of the function  $x \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ .

**Théorème** (Central limit theorem – De Moivre Laplace theorem).

Let  $S_n$  be the sum of n independent Bernoulli random variables of parameter  $p \in (0, 1)$ . Then, for every a < b:

$$\mathbb{P}\left(a \leqslant \frac{\sqrt{n}}{\sqrt{p(1-p)}} \left(\frac{S_n}{n} - p\right) \leqslant b\right) \quad \underset{n \to \infty}{\longrightarrow} \quad \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

We say that  $\frac{\sqrt{n}}{\sqrt{p(1-p)}} \left(\frac{S_n}{n} - p\right)$  converges *in distribution* to a Gaussian random variable.

# The central limit theorem



-\̈́\_-

We cannot predict where the endpoint will be, but we know the probability that it belongs to a given region.

### I. A DREAM

#### **II.** LAW OF LARGE NUMBERS

#### **III.** CENTRAL LIMIT THEOREM

### IV. PRACTICAL ORGANIZATION





Lecturer:

∧→ Igor Kortchemski, igor.kortchemski@math.ethz.ch



### Lecturer:

→ Igor Kortchemski, igor.kortchemski@math.ethz.ch

A→ Lectures on Thursdays and Fridays, 8:15-10:00 am (14 weeks, last lecture: Friday 22nd December)

#### Lecturer:

✓→ Igor Kortchemski, igor.kortchemski@math.ethz.ch
✓→ Lectures on Thursdays and Fridays, 8:15-10:00 am (14 weeks, last lecture: Friday 22nd December)

Course website:

∧→ Metaphor website (coordinator: Robert Crowell)



### Lecturer:

✓→ Igor Kortchemski, igor.kortchemski@math.ethz.ch
✓→ Lectures on Thursdays and Fridays, 8:15-10:00 am (14 weeks, last lecture: Friday 22nd December)

### Course website:

∧→ Metaphor website (coordinator: Robert Crowell)

### Material:

 $\wedge \rightarrow$  Lectures will be recorded, lecture notes available.



#### Lecturer:

✓→ Igor Kortchemski, igor.kortchemski@math.ethz.ch
✓→ Lectures on Thursdays and Fridays, 8:15-10:00 am (14 weeks, last lecture: Friday 22nd December)

### Course website:

∧→ Metaphor website (coordinator: Robert Crowell)

### Material:

 $\wedge$  Lectures will be recorded, lecture notes available.

Exercise classes : ∧→ Tuesdays (14:00-15:00 or 15:00-16:00, TA's: Robert Crowell, Nicolas Hotton, Lyu Zifan)

#### Lecturer:

✓→ Igor Kortchemski, igor.kortchemski@math.ethz.ch
✓→ Lectures on Thursdays and Fridays, 8:15-10:00 am (14 weeks, last lecture: Friday 22nd December)

### Course website:

∧→ Metaphor website (coordinator: Robert Crowell)

### Material:

 $\wedge \rightarrow$  Lectures will be recorded, lecture notes available.

Exercise classes : ∧→ Tuesdays (14:00-15:00 or 15:00-16:00, TA's: Robert Crowell, Nicolas Hotton, Lyu Zifan) ∧→ Submission of exercises is optional (yet recommended!)

# Exam: $\Lambda \rightarrow$ Written.

#### Lecturer:

✓→ Igor Kortchemski, igor.kortchemski@math.ethz.ch
✓→ Lectures on Thursdays and Fridays, 8:15-10:00 am (14 weeks, last lecture: Friday 22nd December)

### Course website:

∧→ Metaphor website (coordinator: Robert Crowell)

### Material:

 $\wedge$  Lectures will be recorded, lecture notes available.

Exercise classes : ∧→ Tuesdays (14:00-15:00 or 15:00-16:00, TA's: Robert Crowell, Nicolas Hotton, Lyu Zifan) ∧→ Submission of exercises is optional (yet recommended!)

### Exam:

 $\wedge$  Written.A mock exam will be made available at the middle of the semester.

#### Lecturer:

✓→ Igor Kortchemski, igor.kortchemski@math.ethz.ch
✓→ Lectures on Thursdays and Fridays, 8:15-10:00 am (14 weeks, last lecture: Friday 22nd December)

### Course website:

∧→ Metaphor website (coordinator: Robert Crowell)

### Material:

 $\wedge \rightarrow$  Lectures will be recorded, lecture notes available.

Exercise classes : ∧→ Tuesdays (14:00-15:00 or 15:00-16:00, TA's: Robert Crowell, Nicolas Hotton, Lyu Zifan) ∧→ Submission of exercises is optional (yet recommended!)

### Exam:

 $\wedge \rightarrow$  Written.A mock exam will be made available at the middle of the semester. A forum will be set up towards the end of the semester.  $\rightarrow$  wooclap.com ; code **probability** (questions on your background).

