Thinking like a Frequentist

Student

January 2026

This page is intentionally left blank.

Preface

When I was learning Probability and Statistics (ProbStat for short), I didn't really understand why things were distributed *normally*. Were they *normal* because they were, or were *we just assuming "Everything is normal for simplicity"*? Sometimes, I felt we were overusing this term and in many situations, we were separating our theory from the actual data due to the intital assumption of *normal* distribution.
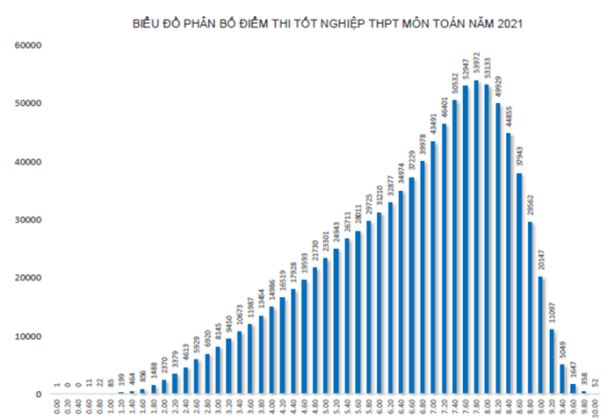

Figure 1: Distribution of Math scores in the 2021 National Entrance Exam

As you can see here, this graph **doesn't really resemble a bell-shaped curve**; and clearly, this data is not distributed normally. In fact, not much real-world statistical data fits a bell-shaped curve. But why are we still using it?

In my opinion (as a junior CE student), I think the main reason is due to **Central Limit Theorem** (CLT for short), every sample in the same population (regradless of how the data is distributed) has a **mean** that converges to a *normal distribution*, or **bell-shaped curve**. This topic will be explained in detail in **Chapter 8: Fundamental of Statistics**.

This book approaches the fundamental of probability and statistics in a very mathematically rigourous way to ensure maximum accuracy. I encourage readers to prove all of theorems, corollaries; and you should create your own examples for each theorem to gain a deeper understading of their origins.

—Student—

Acknowledgement

First of all, I want to thank you for reading this book. Although I am not a native English speaker, but I truly enjoy writing in English. I have tried my best to express my ideas in English, but minor grammatical or spelling errors are unavoidable. If you find any of them, I would be very happy to receive your feedback via my email address: `tinvu1309@gmail.com`

Secondly, I am extremly grateful and would like to express my thanks to the authors of the textbook "Probability and Statistics for Engineers and Scientists", 9th edition. This book truly saved my student life.

Finally, the idea of writing this book was inspired by Prof. Steve Brunton lectures on YouTube. You should check out his videos too!

—Student—

Contents

List of Figures

List of notations

Since much of my work is handwritten, so I've modified some commonly used notations for probability distribution funcions using curved script for convenience. You should notice that my conventions are not the international standards.

1. $\mathscr{I}$: Bernoulli distribution

2. $\mathscr{B}$: Binomial distribution

3. $\mathscr{B}^*$: Negative binomial distribution

4. $\mathscr{H}$: Hypergeometric distribution

5. $\mathscr{G}^*$: Geometric distribution

6. $\mathscr{P}$: Poisson distribution

7. $\mathscr{N}$: Normal distribution

8. $\mathscr{G}$: Gamma distribution

9. $\mathscr{E}$: Exponential distribution

10. $\mathscr{C}$: Chi-squared distributon

11. $\mathscr{T}$: t-distribution (or Student distribution)

Chapter 1

Introduction to Probability and Statistics

1.1 Tossing a coin

Imagine you have a coin, like this one. It's an ordinary coin that can be found everywhere.

Figure 1.1: Tossing a coin

Everyone knows that the chance of head (H) appearing after each toss is 50%, no one even doubts that. So, the **probability** of an **event** that head appearing is 50%. But because you're a very curious person, you don't easily accept the truth like others, so you will find a way to test it. The fact "The probability of head appearing is 50%" will be tested, and it can be called a **hypothesis**.

The simplest way to test your hypothesis is tossing a coin many times. You might toss a coin $N = 100$ times and see heads appear $n = 40$ times, so your actual probability is:

$$p = \frac{n}{N} = 0.4$$

But your previous assumtion (or hypothesis) states that the ideal probability is:

$$\hat{p} = 0.5$$

Is there anything wrong here? Dissatisfied, you continue your experiment. This time, you toss a coin $N = 500$ times and see heads appear $n = 270$ times, so the new actual probability is:

$$p = \frac{n}{N} = 0.54$$

This time the result is closer with intial hypothesis $\hat{p}$, but your hand now must be very tired after tossing a coin 600 times. So do you think you have **enough** evidence to conclude that if $N \to +\infty$, then $p \to \hat{p}$? If so, your assumption is correct because *can't be refuted*; and if not, our common sense might be wrong because the evidence *strongly refutes it.*

Another strategy to toss a coin is fixing and dividing N. Instead of tossing $N = 600$ times and only getting the final p value, you can divide large number of coin tosses $N = 600$ to smaller tosses $N_1 = N_2 = \cdots = N_6 = 100$ (but don't too small, at least each $N_i > 30$), and get 6 values of p_i. Suppose that you would get:

p_1	p_2	p_3	p_4	p_5	p_6
0.43	0.46	0.56	0.53	0.49	0.51

You might see that the value $\hat{p} = 0.5$ too idealistic to occur in our experiment, and p constantly changes. Therefore instead of determining the **exact value** of p, we **predict** that the p value lies inside a **closed interval** with a **certainty** of $(1-\alpha)100\%$. For example, we can confidently conclude that the value of p lies within the range (0.48,0.51) with 95% accuracy.

1.2   Weather forecasting

You don't need to know anything about geography to predict what the weather will be tomorrow. Everything you need is just knowledge about Poisson, exponential distribution and the **average number** of rainy days per month where you are living.
For example, in December there are **average** 5 rainy days. If today is the 10th and you haven't seen any rain since the beginning of the month, what are the probabilities of:

1. There will be 3 rainy days this month.

2. Tomorrow will be a rainy day.

This problem will be covered in detail in **Chapter 5: Discrete Random Variables** and **Chapter 6: Continuous Random Variables**, but if you do have experience with random variables, you can try to solve it!
The answer for the first question is:

$$P(X = 3) = \frac{e^{-5}5^3}{3!} \approx 14.03\%$$

The answer for the second question is:

$$P(X < 11 | X > 10) = 1 - P(X > 11 | X > 10) = 1 - e^{\frac{-5}{31}} \approx 14.89\%$$

Since 14.89% is quite low, so tomorrow you don't have to bring an umbrella.

1.3   Relationship between Probability and Statistics

Chapter 2

H-test and T-test

Thanks to my Prof Ha, Do Thi Thu for saving my life. We can efine contour integral: aas

Theorem 2.0.1. *Let f be a function whose derivative exists in every point, then f is a continuous function.*

Theorem 2.0.2 (Pythagorean theorem)**.** *This is a theorem about right triangles and can be summarised in the next equation*
$$x^2 + y^2 = z^2$$

And a consequence of theorem 2.0.2 is the statement in the next corollary.

Corollary 2.0.2.1. *There's no right rectangle whose sides measure 3cm, 4cm, and 6cm.*

You can reference theorems such as 2.0.2 when a label is assigned.

Lemma 2.0.3. *Given two line segments whose lengths are a and b respectively there is a real number r such that $b = ra$.*

Definition 2.0.1 (Fibration)**.** *A fibration is a mapping between two topological spaces that has the homotopy lifting property for every space X.*

Proof. Let z be some element of $xH \cap yH$. Then $z = xa$ for some $a \in H$, and $z = yb$ for some $b \in H$. If h is any element of H then $ah \in H$ and $a^{-1}h \in H$, since H is a subgroup of G. But $zh = x(ah)$ and $xh = z(a^{-1}h)$ for all $h \in H$. Therefore $zH \subset xH$ and $xH \subset zH$, and thus $xH = zH$. Similarly $yH = zH$, and thus $xH = yH$, as required. $\square$