

ST2132 Introduction

Semester 1 2022/2023

If printing, do DOUBLE-SIDED, each side TWO slides.

- ▶ BSc Mathematics 1996 NUS. MSc Applied Mathematics 1997 NUS. PhD Statistics 2002 University of California.
- ▶ Evolution of DNA; applied statistics, especially in biology and medicine; teaching statistics and mathematics to undergraduate and pre-university students.
- ▶ International Association of Statistics Educators, Singapore Mathematical Society, Nature Society (Singapore).

A box of tickets

A box has 10,000 tickets; K are white.

100 tickets are drawn without replacement; 78 are white.

Mathe: K must be between 78 and 9978.

Stata: **If the draws are random**, K is around 7800, with a likely error of 400.

Mathe: You just use some formulae to get your numbers, right?

Stata: Yes, mathematics is used, but the assumption is crucial.

“Random” does not have a precise mathematical definition. The statistician has to judge whether the draws are random.

Statistics is a **mathematical science** concerned with answering numerical questions arising from the study of nature.

What are our learning goals?

1. To solve estimation problems such as the above, by formulating a statistical model.
2. To assess the goodness-of-fit of a model, and more generally to conduct hypothesis tests.
3. To appreciate the importance of statistical assumptions (randomness, model specification) in statistical inference.
4. To appreciate the role of mathematics, statistics, and science in making sense of our world.

In learning statistical procedures, pay attention to the intentions first, and then the techniques.

Important topics beyond this module

1. Descriptive statistics: summary and visualisation of data
2. Intensive coding and software packages
3. Prediction: machine learning, etc.

Where applicable, connections to them will be made, for you to pursue later.

Outline of topics

Reference: *Mathematical Statistics and Data Analysis* by John A. Rice.

1. Review (chapters 1-6)
2. Parameter estimation (chapters 7, 8)
3. Testing hypotheses (chapter 9)

Besides probability, need basic knowledge in linear algebra, calculus, R computing, and a scientific/skeptical attitude.

I will help you consolidate relevant background knowledge, though you may need to review past modules.

- ▶ Probability will be defined on random experiments that can be repeated many times.
- ▶ All random variables have expectation and variance. Practical problems hardly ever forces a random variable that has an infinite variance, or which do not even have an expectation.
- ▶ Similarly, theorems may not be stated in the most general form. But mathematical arguments that are important to applications will be scrutinised formally and intuitively.
- ▶ Notation will be kept consistent. The unfortunate few instances where it is not will be pointed out.

Assessment (tentative weights)

- ▶ 20%: the best 20 in-lecture quizzes
- ▶ 10%: attending at least 8 tutorials
- ▶ 20%: two assignments
- ▶ 50%: final exam Wednesday 30 November 9 am – 11 am

No cheat sheet allowed in the final. The paper will have a formula sheet.

Get the most out of this module

- ▶ Come to every lecture and tutorial. Lecture recordings should only be for revision or enforced absence.

If absent with MC, please send a copy to me (lecture) or your tutor (tutorial).

- ▶ Form a study group.
- ▶ Ask questions on Canvas, or email your tutor, or me.

I want to help you become a better statistician, and also a better mathematician, scientist, etc. Best of luck with this semester.