

SI140A: Probability & Statistics for EECS

Ziyu Shao

School of Information Science and Technology
ShanghaiTech University

February 07, 2023

Outline

1 Teaching Team

2 Course Information

3 Probability & Statistics

Outline

1 Teaching Team

2 Course Information

3 Probability & Statistics

The Dark Forest in Random Kingdom



Life is Random



Life was like a box
of chocolates. You
never know what
you're gonna get.



Forrest Gump [1994]
www.geckoandfly.com

Even Great Scientists Failed

- Isaac Newton: “I can calculate the motion of heavenly bodies, but not the madness of people.”



Even Great Scientists Failed

- Albert Einstein: “God does not play dice with the universe.”



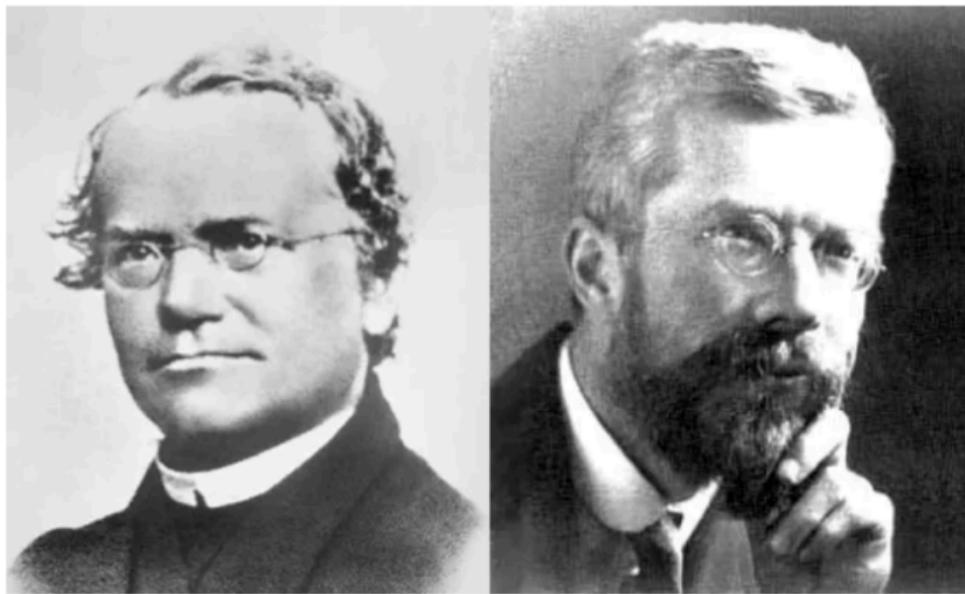
Sometimes Mysterious

- Schrodinger's Cat: “遇事不决，量子力学；解释不通，穿越时空。”



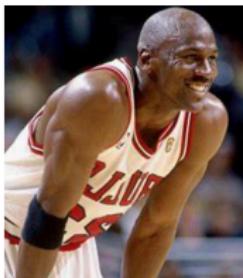
Sometimes Controversial

- Mendel-Fisher Controversy: Cherry Picking?



Sometimes Controversial

Case 3: Who is the greatest NBA player?



VS



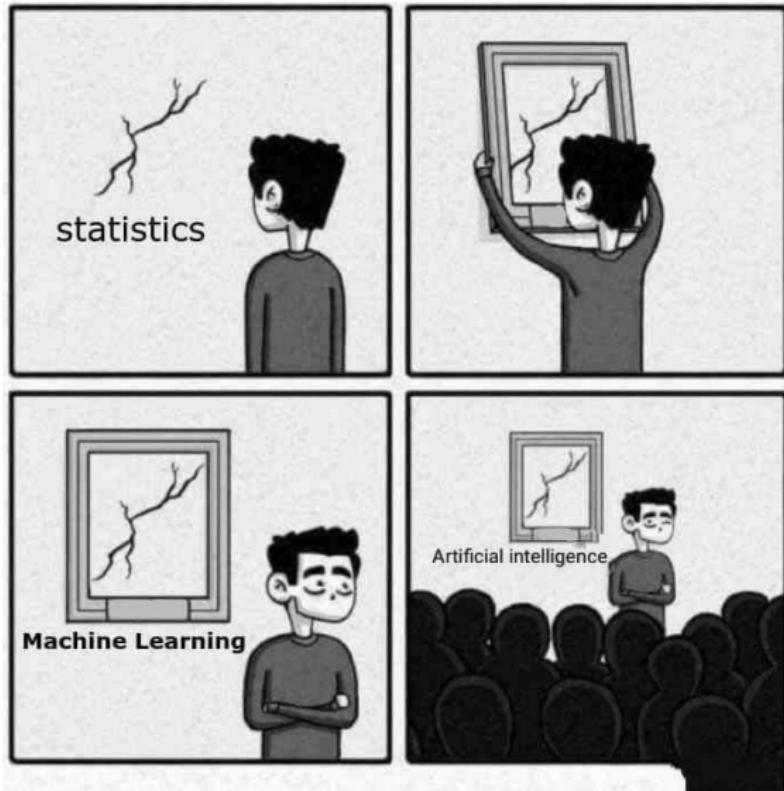
篮球得分	乔丹	詹姆斯
二分球	$12192/24537 = 49.7\%$	$12424/24654 = 50.4\%$
三分球	$581/1778 = 32.7\%$	$1860/5409 = 34.4\%$
总计	$12773/26315 = 48.5\%$	$14284/30063 = 47.5\%$

Sometimes Reinventing the Wheel

Reduction of random noise has many names in different fields.

- Terminology in EECS: signal processing (filter)
- Terminology in Probability & Statistics: estimation
- Terminology in ML & AI: prediction

Sometimes Reinventing the Wheel



You Need Shepherd!



TAs and Professors

- TA Shangshang Wang: 王上上
- TA Yuhang Wang: 王宇翰
- TA Bowen Xu: 徐博文
- Professor Dingzhu Wen: 文鼎柱
- Professor Ziyu Shao: 邵子瑜

Professor Ziyu Shao: 邵子瑜



<https://faculty.sist.shanghaitech.edu.cn/faculty/shaozy>

Outline

1 Teaching Team

2 Course Information

3 Probability & Statistics

Time & Venue

- **Time:** every Tuesday (10:15-11:55am) and Thursday (10:15-11:55am)
- **Location:** Room 302, Teaching Center
- **Course Forum in Piazza:**
`https://piazza.com/shanghaitech.edu.cn/spring2023/si140a`

Course Grade

- *Homework* 20%: 5 or 6 problems per week
- *Behavior* 5%: miscellaneous behaviors
- *Final Project* 15%: due after final exam
- *Midterm Exam* 30%
- *Final Exam* 30%
- *Overall Winners:* lunch
 - ▶ Gauss Award
 - ▶ Markov Award
 - ▶ Bernoulli Award
 - ▶ Laplace Award
 - ▶ Fisher Award

Homework Policy

- Submit homework through Gradescope.
- Write the homework with LaTeX and output it as a PDF file.
- Write the programming assignment with Python and submit source codes.
- **Later** Homework receives no credit.
- You are allowed to discuss with others and use any references, but if you do so please list your collaborators and cite your references for each question.
- Not writing your own solutions or not listing your collaborators or not citing your references may be considered plagiarism.

Project Policy

- Project can be done by a single student or a team with two students.
- Teams with three and more students are NOT allowed.
- Your team is required to use Python for the programming part.
- Your team needs to submit both the Python code and a report (in PDF format) including your simulation results, analysis, discussions, tables, figures, etc.
- **Later** Project receives no credit.

LaTeX Setup

- LaTeX Package (Unix, Windows): TeXLive
(<https://tug.org/texlive/>)
- LaTeX Package (MacOs): MacTeX (<http://tug.org/mactex>)
- Popular LaTeX Editors
 - ▶ Texpad (MacOs): <https://www.texpad.com>
 - ▶ TeXstudio (Windows,MacOs): <http://www.texstudio.org>
 - ▶ LyX(Windows,MacOs): <https://www.lyx.org>
 - ▶ WinEdt (Windows): <http://www.winedt.com/>
- Online LaTeX Editor: Overleaf (<http://overleaf.com/>)

LaTeX Tutorials

- Wiki LaTeX Book: <http://en.wikibooks.org/wiki/LaTeX>
- Collection of links for LaTeX: <https://www.tug.org/begin.html>
- Forum: <https://tex.stackexchange.com/>

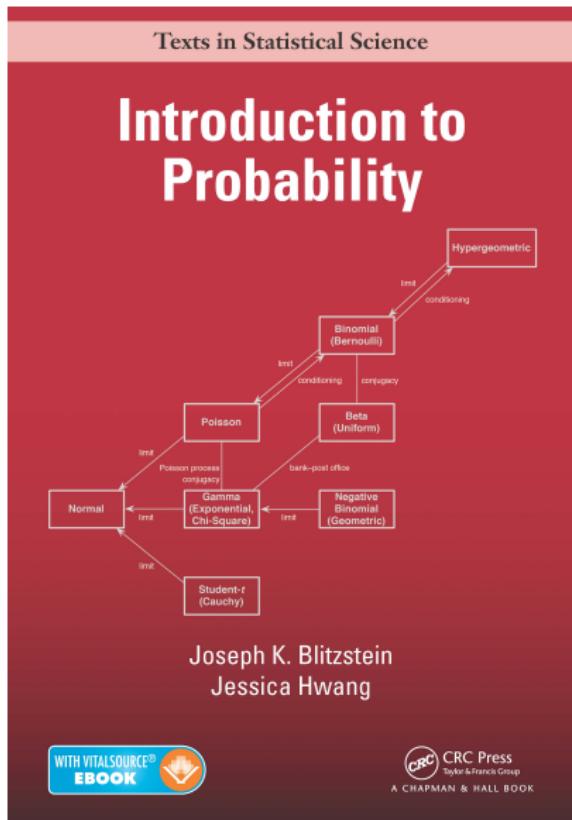
R: Optional

- Setup: <https://www.r-project.org/>
- Editor: <https://www.rstudio.com/>
- Data Visualization in R: <http://ggplot2.org/>
- Reference 1: *R in Action: Data analysis and graphics with R*, Robert I. Kabacoff, Manning Publications, 2015.
- Reference 2: *R for Data Science*, Hadley Wickham & Garrett Grolemund, O'Reilly Media, 2016.

Textbooks: Required

- *Introduction to Probability*, Joseph K. Blitzstein & Jessica Hwang, Chapman & Hall/CRC, 2014. (**BH**)
- *Introduction to Probability (2nd Edition)*, Dimitris P. Bertsekas & John N. Tsitsiklis, Athena Scientific, 2008. (**BT**)

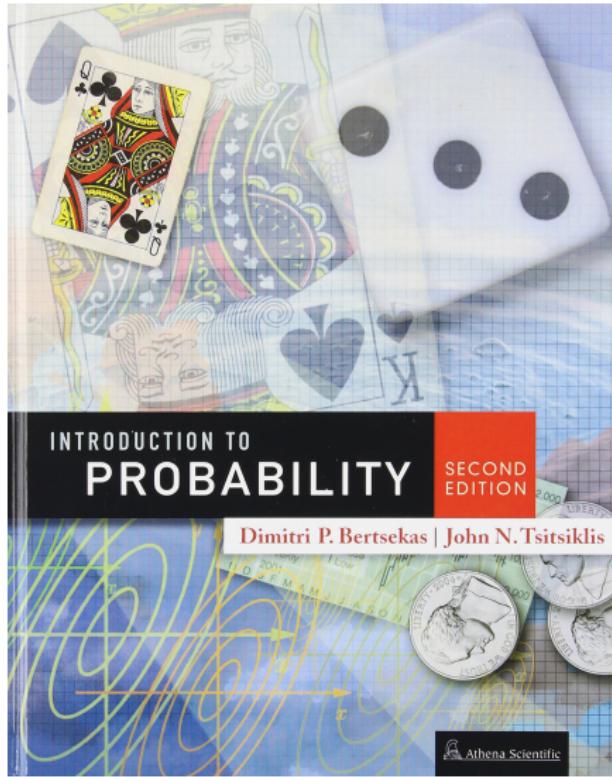
Required Textbook: BH



Joseph K. Blitzstein &
Jessica Hwang

- Introduction to Probability
- Chapman & Hall/CRC, 2014.

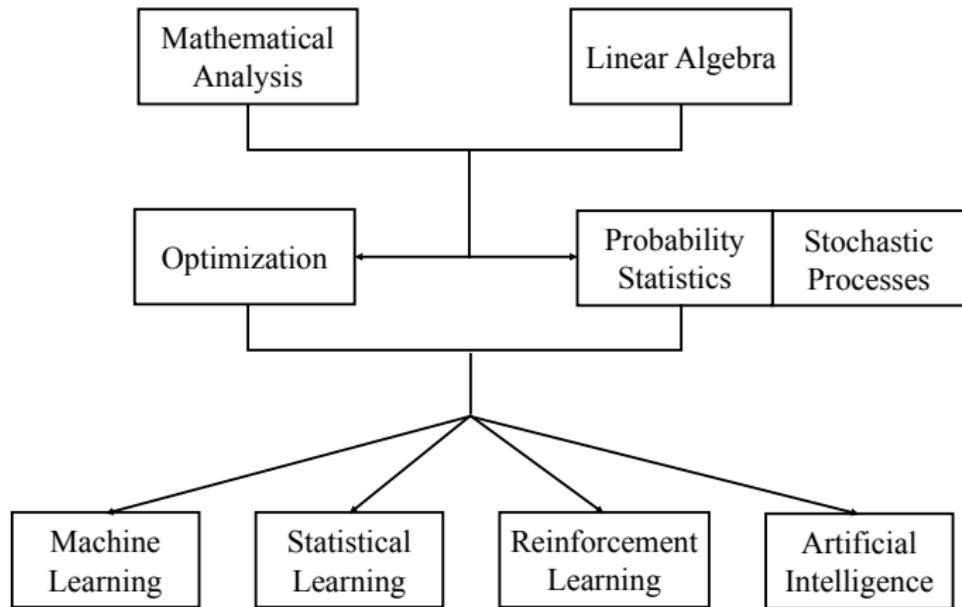
Required Textbook: BT



Dimitris P. Bertsekas &
John N. Tsitsiklis

- Introduction to Probability
(2nd Edition)
- Athena Scientific, 2008.

The Role of This Course



Course Objectives

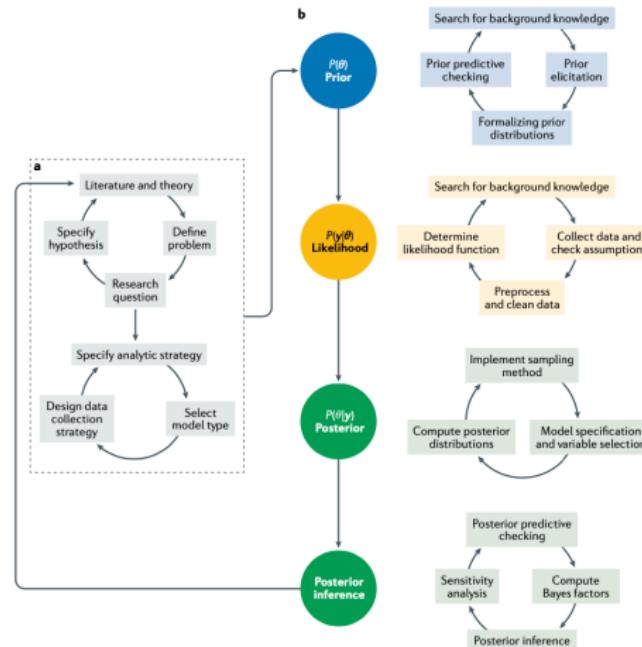
- Integration of Probability, Statistics, Data Science, Random Algorithm
- **Conversation between mathematics and engineering:** theory, model, algorithm, practice
 - ▶ **Solid foundation of probability & statistics:** modeling, analysis, and theoretical proof
 - ▶ **Computational thinking:** algorithmic perspective and programming practice
 - ▶ **Application mentality:** many examples from EECS

Course Focus

- **Motivation:** In the ocean of mathematical definitions, theorems, and equations, why should we spend our time on this particular topic but not another?
- **Intuition:** When going through the derivations, is there a physical intuition beyond those equations?
- **Implication:** After we have learned a topic, what new problems can we solve?

Theme

- One Core: Bayesian Inference
- Two Basic Tools: Conditional Probability & Conditional Expectation



Basic Contents: Traditional Perspective

- Probability and Counting
- Conditional Probability
- Random Variables and Distributions
- Expectation
- Continuous Random Variables
- Joint Distributions
- Transformations
- Conditional Expectation
- Inequalities & Limit Theorems & Monte Carlo
- Classical Statistical Inference
- Bayesian Statistical Inference
- Markov Chain (optional)

Outline

1 Teaching Team

2 Course Information

3 Probability & Statistics

Probability vs. Statistics

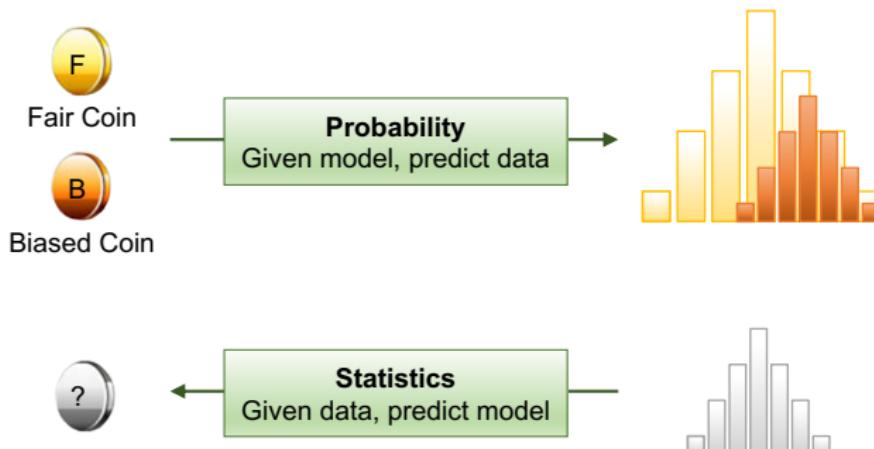


Statistics: Given the information in your hand, what is in the pail?



Probability: Given the information in the pail, what is in your hand?

Probability vs. Statistics



Probability vs. Statistics

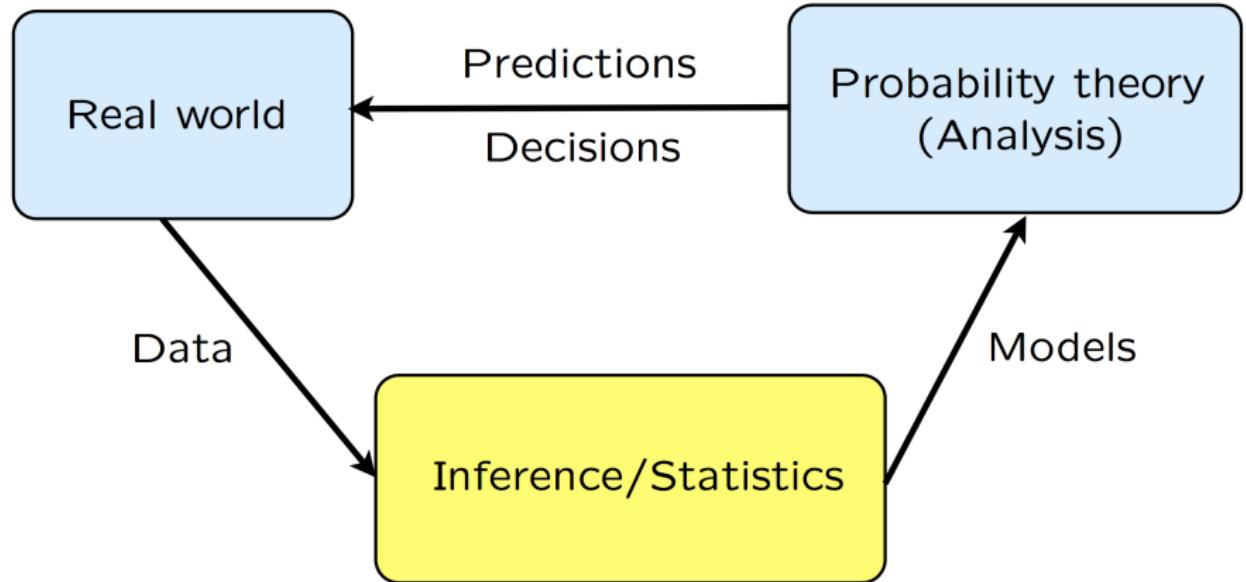
- **Focus of Probability:** given a data generating process (model), what are the properties of the outcomes?
- **Focus of Statistics:** the inverse of probability, i.e., given the outcomes, what can we say about the process that generated the data (model)?

随机 非随意
概率 破玄机
无序 隐有序
统计 解谜离

词以境界为最上。有境界自成高格，
自有名句。有造境，有写境，此理想
与现实二派之所分。

人间词话-王国维

A Framework to Quantify Uncertainty



Let the Adventure Begin!

