# Modern Statistics

## Lecture 1: Introduction

Xiangyu Chang(常象宇)

Email: xiangyuchang@xjtu.edu.cn
Center for Intelligent Decision-Making and Machine Learning
School of Management
Xi'an Jiaotong University

# Introduction


## Professor

Center for Intelligent Decision-Making and Machine Learning
Department of Information Systems and Intelligent Business
School of Management

## Address

Room #750, Main Building of Management School
Xi'an Jiaotong University (XJTU)
China

## Email

xiangyuchang AT xjtu DOT edu DOT cn


## Employment Experience

- 2018.12-2019.12, Visiting Associate Professor
  Department of Industrial & Systems Engineering, University of Washington, Seattle.
- 2016.01-2016.02, Research Associate
  Department of Mathematics, the City University of Hong Kong
- 2015.04-2016.04, Visiting Assistant Professor
  Department of Business Statistics and Econometrics, Guanghua School of Management, Peking University

## Education

- 2013-2014, Ph.D., Applied Math, XJTU, China
- 2010-2012, Visiting Ph.D. Student, Statistics, UC Berkeley, U.S.
- 2009-2010, Ph.D. Candidate, Applied Math, XJTU, China
- 2008-2009, M.S., Applied Math, XJTU, China
- 2003-2007, B.S., Applied Math, XJTU, China

## Research Interests

- Statistics & Machine Learning
- Business Data Science

**http://xiangyuchang.github.io/**

SOM-XJTU

**Modern Statistics**

Spring 2025

HOME SCHEDULE LECTURES ASSIGNMENTS PROJECT MATERIALS

# Modern Statistics / Spring 2025

## Updates

- New course in Spring 2025 is comming soon.
- New Lecture is up: Lecture 1 Review [notes]

## Course Description

This course is to present the basic concepts and theories in modern statistics for undergraduate students at School of Management, Xi'an Jiaotong University.

**Location:** Main Building D-404

**Date**: 3rd&4th class, Mondays and 3rd&4th class, Thursdays, 02/17 - 04/17, 2025.

**https://som-course.github.io/stats-spring-2025/**

## Course Description

This course is to present the basic concepts and theories in modern statistics for undergraduate students at School of Management, Xi'an Jiaotong University.

**Location:** Main Building D-404

**Date**: 3rd&4th class, Mondays and 3rd&4th class, Thursdays, 02/17 - 04/17, 2025.

## Instructors

Xiangyu Chang

## Teaching Assistants

Zhihong Liu

# Course Introduction

- 32hrs, 1-8 week
- To learn the basic concept and theory of modern statistics
- To learn LaTeX for the scientific writing (HW)
- To practice Python or R for handling simulation problems

# Outline

## Part I: Probability Theory

- Probability
- Distribution
- Random Vector
- CLT and LLN

## Part II: Statistical Inference

- Parametric Inference
- Hypothesis Testing
- Linear Model
- Non-parametric Inference

## Part III: Advance Topics

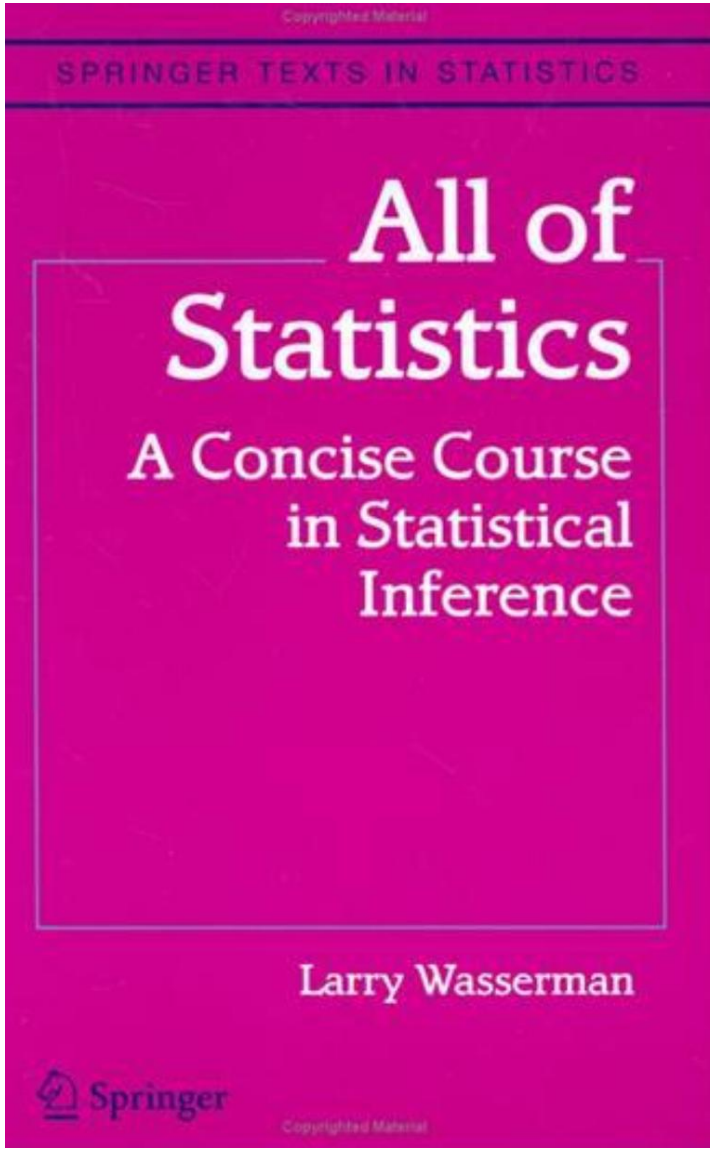
- Bootstrop
- Stochastic Processing
- Cause Inference

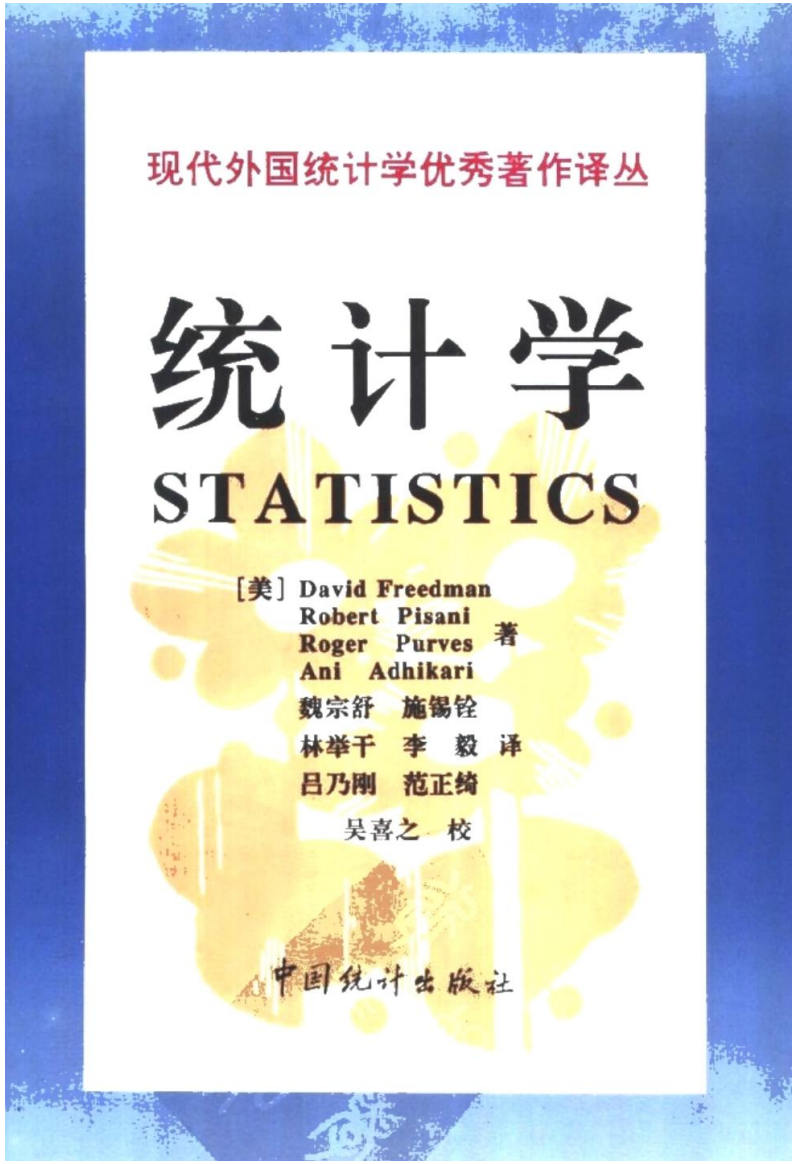
# Evaluation

- **10%Lecture + 40%HW + 50%Final= 100%**

- **Edit your Homeworks and Lectures by LaTeX!**

# Text Books


Copyrighted Material
SPRINGER TEXTS IN STATISTICS
All of Statistics
A Concise Course in Statistical Inference
Larry Wasserman
Springer
Copyrighted Material


现代外国统计学优秀著作译丛
统计学
STATISTICS
[美] David Freedman
Robert Pisani
Roger Purves 著
Ani Adhikari
魏宗舒 施锡铨
林举干 李 毅 译
吕乃刚 范正绮
吴喜之 校
中国统计出版社

# Materials

## Suggested Text Book

- All of Statistics by Larry Wasserman
- Statistical Inference by George Casella and Roger Berger
- Statistics by David Freedman, Robert Pisani, Roger Purves and Ani Adhikari. 译：魏宗舒，施锡铨，林举干，李毅，吕乃刚，范正绮。
- Veridical Data Science by Bin Yu and Rebecca L. Barter

SOM-XJTU

# Modern Statistics

Spring 2025

HOME | SCHEDULE | LECTURES | ASSIGNMENTS | PROJECT | MATERIALS

## Lectures

You can download the lectures here. We will try to upload lectures prior to their corresponding classes.


**Lecture 1 Introduction**

**tl;dr:** introduction lecture

**[slides] [notes]**

**Suggested Readings:**

- LaTeX Materials
- Chapter 1 of AOS
- Chapter 13-15 of STATS

课后复习用，每次课结束后公布

课前预习用，每次课之前公布

# Assignments

You can download the assignments here. Also check out each assignment page for any additional info.

**HW1: Master LaTeX Template**

# HW1: Master LaTeX Template

Released on Monday 02/17/2025

**Due Date: 2025-02-20 10:00**
**Download [problems] [attachment]**

## Late Policy

- Late homework will not be accepted.

This is a simple assignment.

# Schedule

| EVENT | DATE | DESCRIPTION | COURSE MATERIAL |
|---|---|---|---|
| Lecture | 02/17/2025<br>Monday | Lecture 1 Introduction<br>[slides] [notes] | **Suggested Readings:**<br>○ LaTeX Materials<br>○ Chapter 1 of AOS<br>○ Chapter 13-15 of STATS |
| **Assignment** | 02/17/2025<br>Monday | **HW1: Master LaTeX Template released!** | [HW1: Master LaTeX Template] |
| **Due** | **02/20/2025 10:00<br>Thursday** | **Master LaTeX Template for Course Notes.** | |

# LaTeX vs. Word

$$B_{t,\alpha} = \begin{cases} 1 - i\beta \mathrm{sgn}(t) \tan(\frac{\pi\alpha}{2}) & if \quad \alpha \neq 1 \\ 1 + i\beta \mathrm{sgn}(t) \frac{2}{\pi} \log|t| & if \quad \alpha = 1 \end{cases} \qquad (2)$$

Where $-\infty < \delta < +\infty$, $\gamma > 0$, $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$. $\alpha(0 < \alpha \leq 2)$ is the characteristic exponent and sets the degree of impulsiveness of the distribution. The smaller the value of

$$B_{t,\alpha} = \begin{cases} 1 - i\beta \,\mathrm{sgn}(t) \tan(\dfrac{\pi\alpha}{2}) & if \quad \alpha \neq 1 \\ 1 + i\beta \,\mathrm{sgn}(t) \dfrac{2}{\pi} \log|t| & if \quad \alpha = 1 \end{cases} \qquad (2)$$

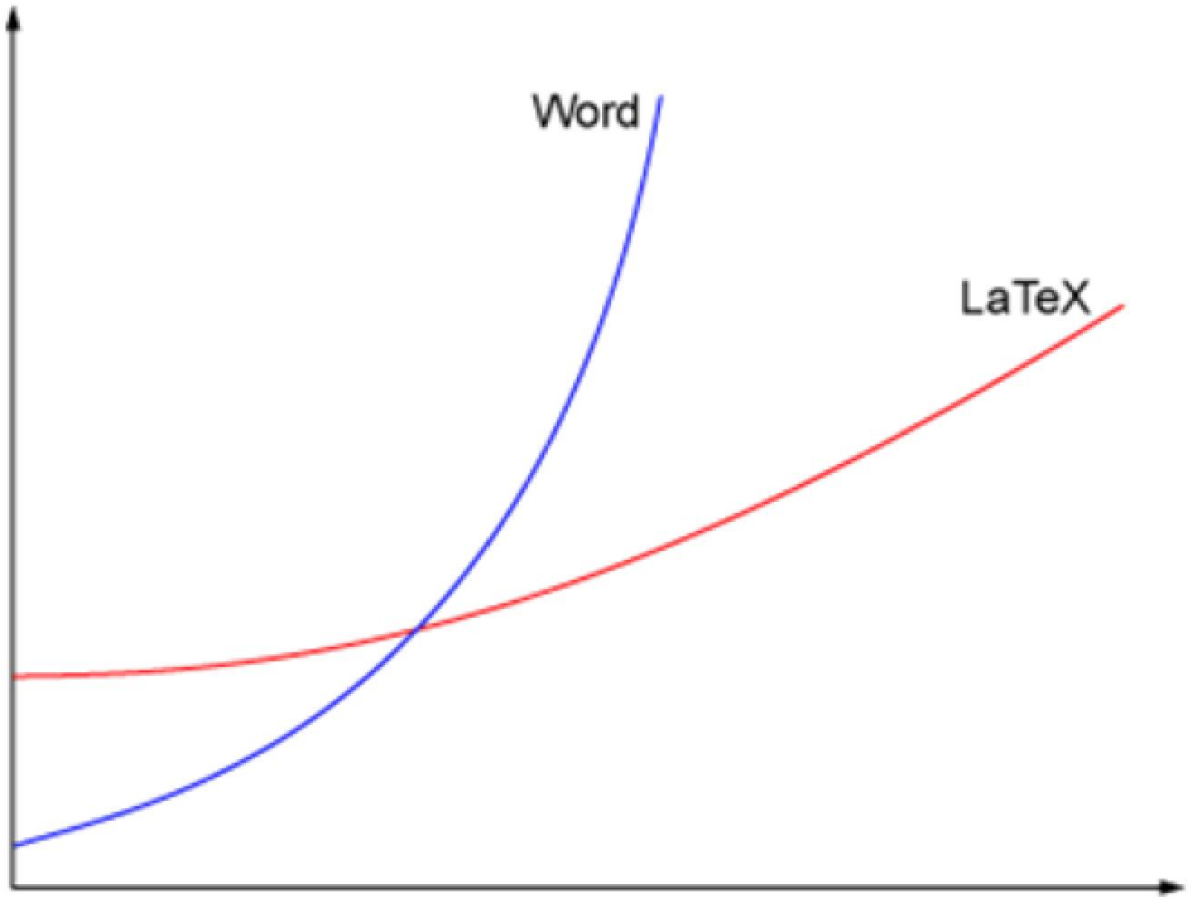
Where $-\infty < \delta < +\infty$ , $\gamma > 0$ , $0 < \alpha \leq 2$ , $-1 \leq \beta \leq 1$ .

$\alpha$ ( $0 < \alpha \leq 2$ ) is the characteristic exponent and sets the degree of impulsiveness of the distribution. The smaller the value of $\alpha$ ,

Latex文档格式均匀，黑白和行距均匀，数学格式排版美观

Slides from Prof. Yao Wang

# LaTeX vs. Word




Slides from Prof. Yao Wang

# LaTeX vs. Word

- 如果只是排版一些简单的公文, word就够用了
- 如果需要排版高质量的包含大量数学公式的科技文档，那就需要用LaTex
- 用LaTeX时间越长，就越会发现word的不足，从而更爱LaTeX
- LaTeX开源，可免费获得(这一条理由就足够)

Slides from Prof. Yao Wang

# LaTeX: Install

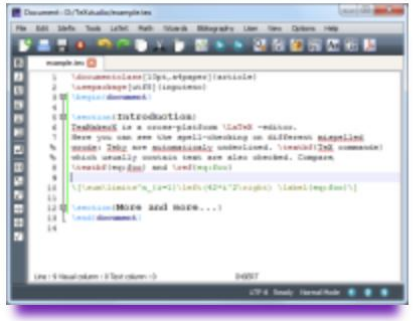

# LaTeX: Suggested IDE

TeXstudio
LaTeX made comfortable

- Home
- News
- Features
- Download
- Participate
- About

## Welcome to TeXstudio

TeXstudio is an integrated writing environment for creating LaTeX documents. Our goal is to make writing LaTeX as easy and comfortable as possible. Therefore TeXstudio has numerous features like syntax-highlighting, integrated viewer, reference checking, and various assistants. For more details see the features.



TeXstudio is open-source and is available for all major operating systems.

**Download now**
TeXstudio 3.1.1 (OSX dmg)
for macOS 10.14 or higher

## News

2021-02-22 A new bugfix release **TeXstudio 3.1.1** is available.Mainly it fixes crashes loading included files.
see changelog

2021-02-17 A new bugfix release **TeXstudio 3.1.0** is available.Mainly changing the GUI language has been fixed. Furthermore a global TOC is available besides the structure view. It is work in progress and not (yet) automatically updated. Please give feedback.
see changelog

# LaTeX: How to Learn

## Learn LaTeX in 30 minutes

In this guide, we hope to give you your first introduction to LaTeX. The guide does not require you to have any prior knowledge of LaTeX, but by the time you are finished, you will have written your first LaTeX document, and hopefully will have a good knowledge of some of the basic functions provided by LaTeX.

## Contents

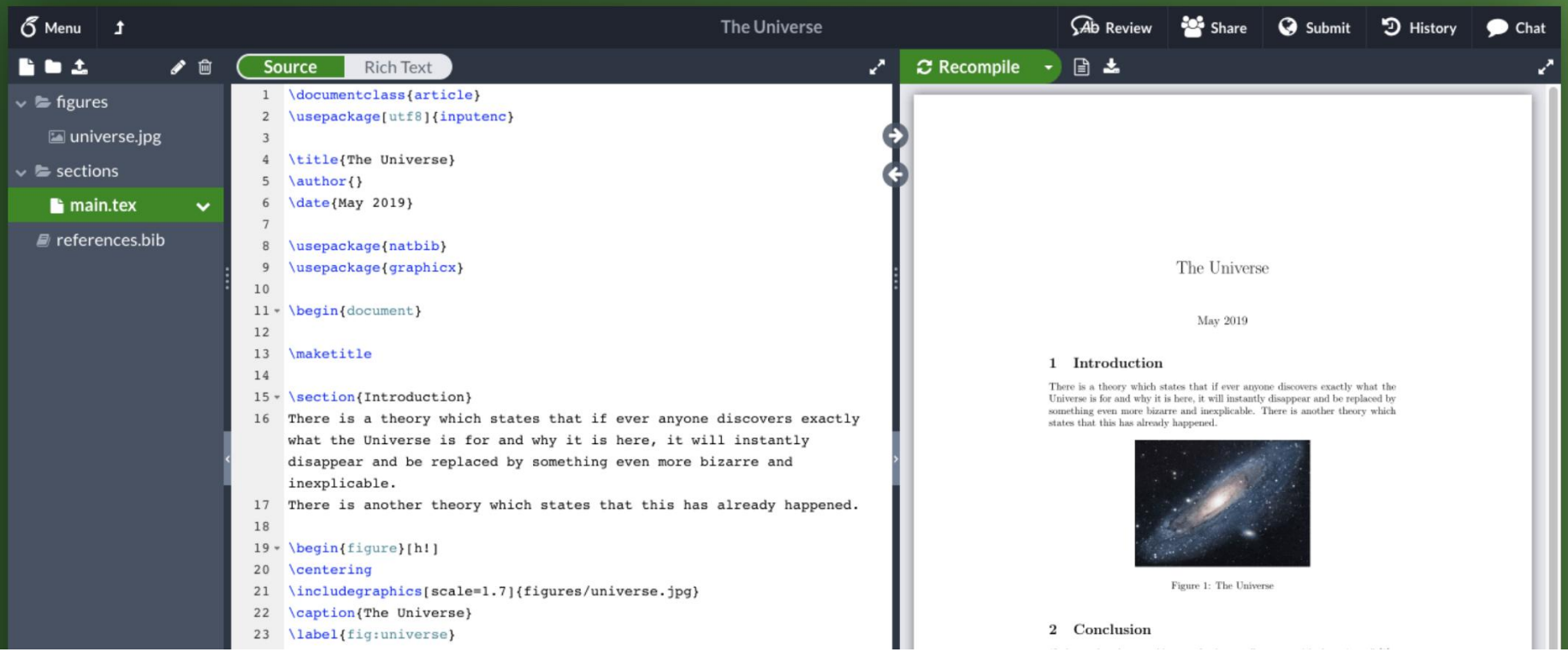
# Use Overleaf!!!

Overleaf

Features & Benefits ▾ Templates Plans & Pricing Help ▾ Register

## LaTeX, Evolved

The easy to use, online, collaborative LaTeX editor



## R

- R
- RStudio: Suggested IDE
- R Cheatsheet
- An introduction course of R by Xiangyu Chang: Lecture 1, Lecture 2, Lecture 3, Lecture 4, Lecture 5, Lecture 6

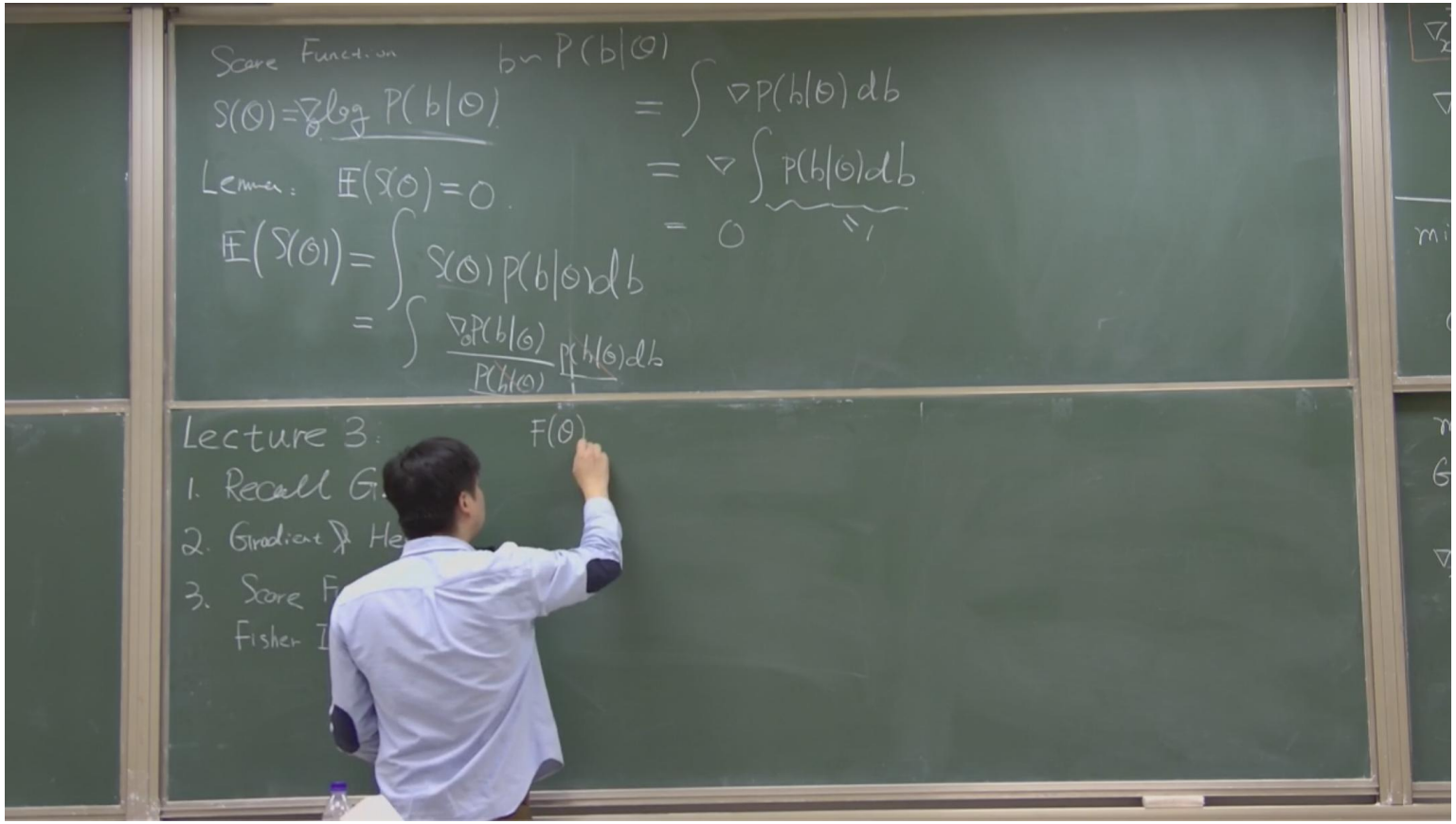
## Python

- Python for Beginners
- Python Cheatsheet
- Python 数据科学实践 by 常象宇，曾智亿，李春艳，程茜著
- SciPy Tutorial
- NumPy Quickstart Tutorial

## LaTeX

- MacTeX
- CTeX
- Learn LaTeX in 30 minutes
- Use Overleaf
- TeXStudio: Suggested IDE

# Teaching Style




- **Writing on the blackboard in English for main lectures, Speaking in Chinese**

- **Demostraing codes or visual examples by sildes**

# Why?

手机人民网 m.people.cn | 观点 > 人民网评

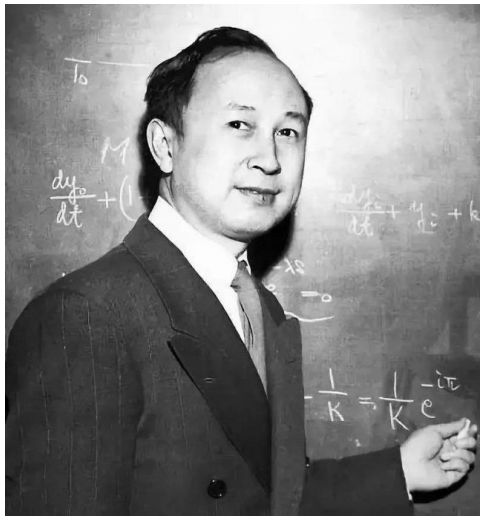
人民网评：一流大学建设要坚持社会主义办学方向

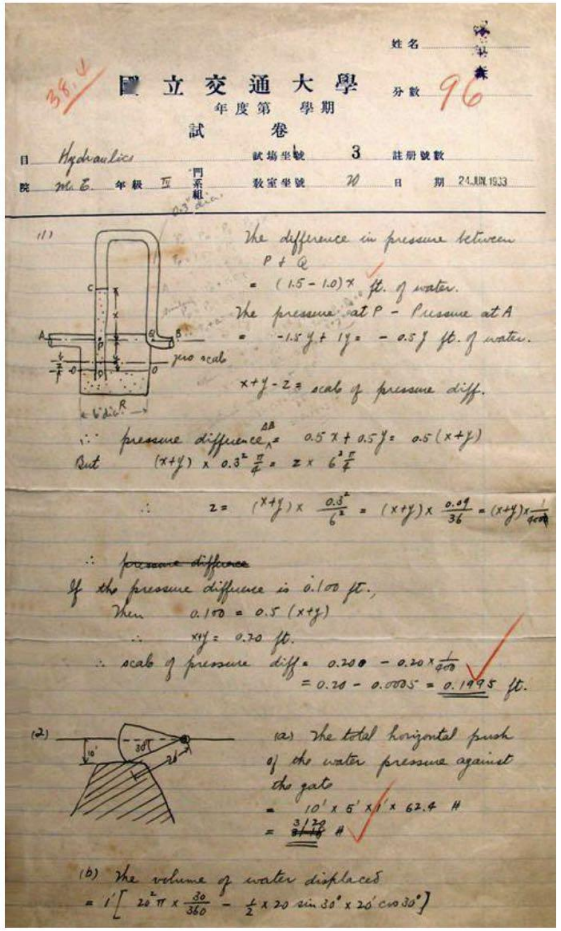
2021-04-22 08:07 人民网-观点频道 作者：岚山


文章指出，教育兴则国家兴，教育强则国家强。在国家发展进程中**办好高等教育，办出世界一流大学，首先要在体现中国特色上下功夫。**

交通大学上海徐家汇校区老校门（1920年代）


1956年交通大学主体西迁西安兴庆宫





姓名

國立交通大學

年度第 學期

分數 96

試卷

Hydraulics 試場座號 3 註冊號數

系組 年級 五 教室座號 N 日期 24.JUN.1933

(1) The difference in pressure between P & Q
= (1.5 − 1.0) × ft. of water.
The pressure at P − Pressure at A
= −1.5 y + 1y = −0.5 y ft. of water.
x + y − z = scale of pressure diff.
∴ pressure difference = 0.5 x + 0.5 y = 0.5 (x + y)
But (x+y) × 0.3² × π/4 = z × 6² × π/4
∴ z = (x+y) × 0.3²/6² = (x+y) × 0.09/36 = (x+y) × 1/400
∴ pressure difference
If the pressure difference is 0.100 ft.,
Then 0.100 = 0.5 (x+y)
∴ x+y = 0.20 ft.
∴ scale of pressure diff. = 0.200 − 0.20 × 1/400
= 0.20 − 0.0005 = 0.1995 ft.

(2) (a) The total horizontal push of the water pressure against the gate
= 10' × 5' × 1/2 × 62.4 #
= 3120 #

(b) The volume of water displaced
= 1' [ 20²π × 30/360 − 1/2 × 20 sin 30° × 20 cos 30° ]

# Questions

# Grouping by DeepSeek

# What is Statistics?

Statistics is the science concerned with developing and studying methods for collecting, analyzing, interpreting and presenting empirical data. Statistics is a highly interdisciplinary field; research in statistics finds applicability in virtually all scientific fields and research questions in the various scientific fields motivate the development of new statistical methods and theory. In developing methods and studying the theory that underlies the methods statisticians draw on a variety of mathematical and computational tools.

Two fundamental ideas in the field of statistics are uncertainty and variation. There are many situations that we encounter in science (or more generally in life) in which the outcome is uncertain. In some cases the uncertainty is because the outcome in question is not determined yet (e.g., we may not know whether it will rain tomorrow) while in other cases the uncertainty is because although the outcome has been determined already we are not aware of it (e.g., we may not know whether we passed a particular exam).

Probability is a mathematical language used to discuss uncertain events and probability plays a key role in statistics. Any measurement or data collection effort is subject to a number of sources of variation. By this we mean that if the same measurement were repeated, then the answer would likely change. Statisticians attempt to understand and control (where possible) the sources of variation in any situation.

We encourage you to continue exploring our website to learn more about statistics, our academic programs, our students and faculty, as well as the cutting-edge research we are doing in the field.

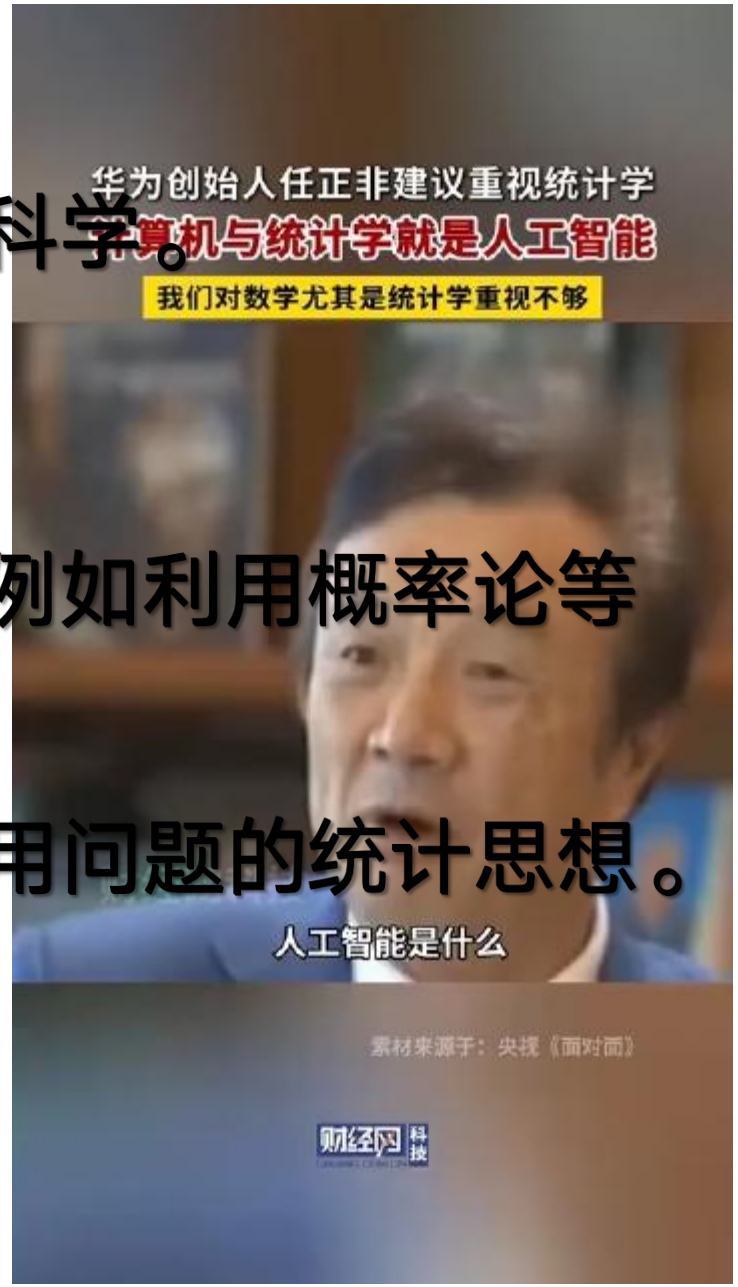
统计是研究数据的收集，整理，分析和解释的科学。

统计是研究不确定性的科学手段。

数理统计是描述统计学的严谨的科学语言——例如利用概率论等语言严格描述。

数据科学——现代应用统计，用于解决实际应用问题的统计思想。

任正非：人工智能的本质是统计。




RUSSIAN


SPANISH


FRENCH


VIETNAMESE

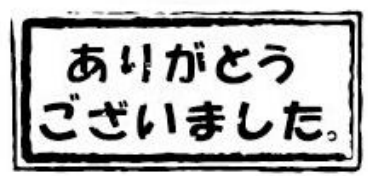

CHINESE


ENGLISH


GREEK


ARABIC


KOREAN




JAPANESE