

教书育人、不忘初心,《数理统计》本科课程 — 课程思政 20 讲

田国梁 统计学教授

南方科技大学 ● 统计与数据科学系

Email: tiangl@sustech.edu.cn

中国 ● 广东 ● 深圳

2024 年 04 月 15 日

提纲 (Outline) Part I

第 1 讲 **South**与**Southern**之区别

第 2 讲 用**14 年创新编写**《数理统计》英文教材

第 3 讲 **Bayes**如何译成中文名? 英文名如何读?

第 4 讲 全概率公式 (**Law of Total Probability**) 和 Bayes 公式

第 5 讲 自然常数 (**Natural Constant**) $e = 2.718282 \dots$ 的起源

第 6 讲 从矩母函数与密度函数的关系出发, 深度理解**国王函数** e^x

第 7 讲 从对数似然函数出发, 深度理解**王后函数** $\log(x)$

第 8 讲 **标准正态分布密度**和**蛇吞象公式**

第 9 讲 从函数的一阶泰勒展开式到**线段中/外**任何一点之数学表达式

第 10 讲 函数的一阶泰勒展开之**四种形式**

提纲 (Outline) Part II

- 第 11 讲 指数分布与几何分布的无记忆性
- 第 12 讲 二项分布的生存函数与贝塔分布的累积分布函数之恒等式
- 第 13 讲 深度理解中心极限定理
- 第 14 讲 二项分布的正态近似和泊松近似
- 第 15 讲 从矩估计量到 Monte Carlo 积分
- 第 16 讲 从 KL 散度的角度来理解极大似然估计之定义
- 第 17 讲 从 Laplace 提出问题到 Gauss 解决问题: 正态分布的发现过程
- 第 18 讲 度量点估计量好坏的指标: 均方误差
- 第 19 讲 克拉默-拉奥 (Cramér-Rao) 不等式
- 第 20 讲 建立参数的置信区间过程中的枢轴量(Pivotal Quantity)

第 13 讲 深度理解中心极限定理



商院楼内微风起，六颜五色四季春

13.1 中心极限定理 (Central Limit Theorem) 的内容

1* 独立同正态分布同参数的均值之分布

- 设 $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$, 则对任何样本大小 n , 均有

$$\bar{X}_n \triangleq \frac{1}{n} \sum_{i=1}^n X_i \sim N(\mu, \sigma^2/n) \quad \text{或} \quad \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0, 1). \quad (13.1)$$

2* 独立同其它分布同参数的均值之分布 (CLT)

- 设 $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} F(\cdot)$, 且 $E(X_1) = \mu$ and $\text{Var}(X_1) = \sigma^2 < \infty$, 则

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{L} N(0, 1) \quad \text{as } n \rightarrow \infty, \quad (13.2)$$

其中 \bar{X}_n 由 (13.1) 定义。

13.2 其它中心极限定理

3* 独立同贝努利分布同参数 (De Moivre CLT, 1733)

- 设 $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Bernoulli}(\theta)$, 定义 $Z_n \triangleq \sum_{i=1}^n X_i$, 则

$$\frac{Z_n - n\theta}{\sqrt{n\theta(1-\theta)}} \xrightarrow{L} N(0, 1) \quad \text{as } n \rightarrow \infty. \quad (13.3)$$

4* 独立不同分布不同参数 (李雅普诺夫, Liapounov CLT)

- 设 $\{X_i\}_{i=1}^{\infty}$ 独立, $X_i \sim F_i(\cdot)$, 且 $E(X_i) = \mu_i$ and $\text{Var}(X_i) = \sigma_i^2 < \infty$. 令 $s_n^2 = \sum_{i=1}^n \sigma_i^2$. 如果对于某个 $\delta > 0$, 下述 Liapounov 条件成立:

$$\frac{1}{s_n^{2+\delta}} \sum_{i=1}^n E|X_i - \mu_i|^{2+\delta} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

则

$$\frac{\sum_{i=1}^n (X_i - \mu_i)}{s_n} \xrightarrow{L} N(0, 1) \quad \text{as } n \rightarrow \infty. \quad (13.4)$$

13.3 所包含的思政元素

- (13.1) 表明, 当母体为**正态分布**(富裕家庭; 西方国家) 时, **不需要** $n \rightarrow \infty$ (不需要刻苦努力; 周末不需要工作), \bar{X}_n **就服从正态分布**(物质生活很好; 住洋楼坐豪车)。
- (13.2) 表明, 当母体为**其它分布**(一般家庭; 一般状态; 中国经济; 中国) 时, **需要** $n \rightarrow \infty$ (需要刻苦努力; 奋斗拚搏; 14 亿人艰苦奋斗; 14 亿人团结一致), \bar{X}_n **也能服从正态分布**(富裕家庭; 颠峰状态; 经济总量世界第 2; 中华民族伟大复兴)。
- 从**量变到质变**的哲学观。