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 25 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**) $\mathrm{e} = 2.718282\cdots$ 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 15 讲 从矩估计量到 Monte Carlo 积分

青山绿水长岭陂，云淡碧波南科大

15.1 矩估计量 (Moment Estimator)

1° 矩估计量的定义

- 设母体随机变量 $X \sim g(x;\theta)$, 其中 $g(x;\theta)$ 是一个具有单参数 θ 的密度函数。设 $X_1,\ldots,X_n \overset{\text{iid}}{\sim} g(x;\theta)$, 则用样本均值估计母体均值

$$\frac{1}{n}\sum_{i=1}^{n} X_i = E(X) = \int x\,g(x;\theta)\,\mathrm{d}x, \tag{15.1}$$

就可以得到 θ 的矩估计量, **只要 n 充分大**。

2° 指数分布均值参数的矩估计量

- 设 $X_1,\ldots,X_n \overset{\text{iid}}{\sim}$ **Exponential**(θ), 其密度函数 $g(x;\theta) = (1/\theta)\mathrm{e}^{-x/\theta}$, $\theta > 0$, 则用样本均值估计母体均值

$$\bar{X} = E(X_1) = \theta,$$

就得到 θ 的矩估计量为 $\bar{X}$。

$3^\bullet$ Monte Carlo 积分公式

- 设$X_1, \ldots, X_n, X \overset{\text{iid}}{\sim} g(x;\theta)$, 其中 $g(x;\theta)$ 是一个密度函数。对任意的一个函数 $h(x)$, 假设它的定义域 $\mathbb{X}$ 与 $g(x;\theta)$ 的定义域相同, 则

$$E\{h(X)\} = \int_{\mathbb{X}} h(x) \cdot g(x;\theta)\,\mathrm{d}x = \frac{1}{n}\sum_{i=1}^{n} h(X_i), \quad n \to \infty, \qquad (15.2)$$

称为 Monte Carlo 积分公式, 它是 (15.1) 的推广。

$4^\bullet$ 证明公式 (15.2)

- 定义 $Y_i = h(X_i)$, $i = 1, \ldots, n$, 且 $Y = h(X)$, 当 $n \to \infty$, 则

$$\frac{1}{n}\sum_{i=1}^{n} h(X_i) = \frac{1}{n}\sum_{i=1}^{n} Y_i \overset{(15.1)}{\to} E(Y) = E\{h(X)\}.$$

15.3 深度理解公式 (15.2)

- Monte Carlo 积分的**本质**：是用**样本均值估计母体均值**。
- 在 Monte Carlo 积分公式 (15.2) 的右手边中，并未出现密度函数 $g(x;\theta)$，它的隐形贡献在于：它生产了 $X_1,\ldots,X_n$，即 $X_1,\ldots,X_n \overset{\text{iid}}{\sim} g(x;\theta)$。
- 公式 (15.2) 从**左到右**容易理解/记住，从**右到左**就不容易记记住与应用。

15.4 所包含的思政元素 (Part I)

- **对立的两面可以相互转换**的哲学思想。(15.1) 是 (15.2) 的特例 (令 $h(x)=x$), (15.2) 是 (15.1) 的推广，在另一个维度 (即 Y 维度，其中 $Y=h(X)$), (15.2) 就是 (15.1)。从 (15.2) 的证明过程来看：**一般公式 $\Longleftrightarrow$ 特殊公式**。

- 将积分转化为求和，从哲学角度看，则是将 "**求面积很困难**" 的问题向 " **求面积很容易**" 的问题，在一定条件 ($n\to\infty$) 下，之**矛盾转化**。

- 积分是连续的，求和是离散的。体现了**连续与离散**之间的相互转化。

- 将无限到有限的转化中，对立的两个方面 (有限个长方形面积之和与有限个曲边梯形面积之和) 得到了统一 (曲边梯形面积)，体现了**对立统一规律**。

- **中华民族之伟大复兴**($\int_{\mathbb{X}} h(x) \cdot g(x;\theta)\,\mathrm{d}x$), 等于全国 14 亿人 ($n = 14$ **亿**) 的贡献 ($h(X_i) \times (1/n)$ **即长方形面积**) 之和; 而我们的先辈 ($g(x;\theta)$) 的隐形贡献: 在于创造了中华民族的优秀文化, 滋养了 14 亿中华儿女 ($X_1, \ldots, X_n \overset{\text{iid}}{\sim} g(x;\theta)$)。

- **南科大的成功**($\int_{\mathbb{X}} h(x) \cdot g(x;\theta)\,\mathrm{d}x$), 等于每位南科人的贡献 ($h(X_i) \times (1/n)$) 之和; 而党、政府、深圳市人民 ($g(x;\theta)$) 的隐形贡献: 在于引领正确的方向, 并向南科人提供了良好的工作、学习与生活环境 ($X_1, \ldots, X_n \overset{\text{iid}}{\sim} g(x;\theta)$)。

15.5 所包含的思政元素 (Part III)

- **南科大学生的成功**($\int_{\mathbb{X}} h(x) \cdot g(x; \theta)\,\mathrm{d}x$), 等于该学生在各个维度的成功 ($h(X_i) \times (1/n)$) 之和; 而南科大 ($g(x; \theta)$) 的隐形贡献: 在于向学生们在各个方面的正能量成长提供了良好的学习与发展条件 ($X_1, \ldots, X_n \overset{\text{iid}}{\sim} g(x; \theta)$)。

- **学生的总 GPA**($\int_{\mathbb{X}} h(x) \cdot g(x; \theta)\,\mathrm{d}x$), 等于该学生在大学期间所学每门课程的 GPA($h(X_i)$) 之平均值; 而南科大 ($g(x; \theta)$) 的隐形贡献: 在于向学生们提供了逐步优化的课程设置、优秀的上课老师、装备良好的教室、环境优美的校园设施 ($X_1, \ldots, X_n \overset{\text{iid}}{\sim} g(x; \theta)$)。