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 03 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**) $e = 2.718282\cdots$ 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 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**)

紫霞满天秋意浓，黄灯暖照映晚风。

学子归途星点散，一路轻歌入梦宫。

8.1 如何证明标准正态分布密度是一个密度函数?

$1^\bullet$ $N(0,1)$ 的密度函数

- 设 $X \sim N(0,1)$, 它的密度函数定义为

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-0.5x^2} > 0, \quad x \in \mathbb{R} = (-\infty, \infty).$$

$2^\bullet$ 证明 $\int_{-\infty}^{\infty}\phi(x)\,\mathrm{d}x = 1$ or $I \triangleq \int_{-\infty}^{\infty} \mathrm{e}^{-0.5x^2}\,\mathrm{d}x = \sqrt{2\pi}$

- 作极坐标变换$x = r\cos(\theta)$, $y = r\sin(\theta)$, 则 $r > 0$, $\theta \in (0, 2\pi)$,

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \det\begin{pmatrix}\cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta)\end{pmatrix} = r.$$

- 我们得到

$$I^2 = \int_{-\infty}^{\infty} \mathrm{e}^{-0.5x^2}\,\mathrm{d}x \cdot \int_{-\infty}^{\infty} \mathrm{e}^{-0.5y^2}\,\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-0.5(x^2+y^2)}\,\mathrm{d}x\mathrm{d}y$$

$$= \int_0^{\infty}\left(\int_0^{2\pi} \mathrm{e}^{-0.5r^2}\cdot r\,\mathrm{d}\theta\right)\mathrm{d}r = 2\pi\int_0^{\infty} r\mathrm{e}^{-0.5r^2}\,\mathrm{d}r$$

$$= 2\pi\left(-\mathrm{e}^{-0.5r^2}\right)\Big|_0^{\infty} = 2\pi(0+1) = 2\pi,$$

因此 $I = \sqrt{2\pi}$。 □

8.2 蛇吞象公式及其应用

3° 蛇吞象公式

- 设 X_1 与 X_2 是两个相依的 (dependent) 随机变量, 蛇吞象公式是指

$$E(X_1) = E\{E(X_1|X_2)\}.$$

4° 蛇吞象公式的应用

- 设 $(X_1, X_2)^\top \sim N_2(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$, 求

$$\begin{aligned} E(X_1X_2^2) &= E\{E(X_1X_2^2|X_2)\} = E\{X_2^2 \cdot E(X_1|X_2)\} \\ &= E\left\{X_2^2 \cdot \left[\mu_1 + \rho\sigma_1\sigma_2^{-1}(X_2 - \mu_2)\right]\right\} \\ &= (\mu_1 - \rho\sigma_1\sigma_2^{-1}\mu_2)E(X_2^2) + \rho\sigma_1\sigma_2^{-1}E(X_2^3), \end{aligned}$$

其中 $E(X_2^2) = \mathrm{Var}(X_2) + [E(X_2)]^2 = \sigma_2^2 + \mu_2^2$, 且

$$0 = E(X_2 - \mu_2)^3 = E(X_2^3) - 3\mu_2E(X_2^2) + 3\mu_2^2E(X_2) - \mu_2^3.$$

8.3 所包含的思政元素

- **扩展思维**: 在**一维直线上**(例如一重积分) 不能解决的问题, 扩展到**二维平面上**去解决 (例如二重积分)。
- **降维思维**: 在**二维平面上**(例如二重积分) 很难解决的问题, 降维到**一维直线上**(例如给定 X_2) 去解决 (例如一重积分)。