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 01 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 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 10 讲　函数的一阶泰勒展开之四种形式

大沙河畔商理院，华晖云门宝能城

10.1 单变量函数的一阶泰勒展开之四种形式

$1^\bullet$ 精确版本 (exact version of the first–order Taylor expansion)

- 设 $f(x)$ 定义在 $\mathbb{R}$ 上，连续且一阶导数存在。对任意一点 $x_0 \in \mathbb{R}$，则存在一点 $x^* = (1-\alpha)x_0 + \alpha x$, $\alpha \in [0,1]$，使得

$$f(x) = \underbrace{f(x_0) + (x - x_0) f'(x^*)}_{\text{Is this a straight line?}}. \tag{10.1}$$

- **数学追求形式完美。**

$2^\bullet$ 近似版本 (approximate version)

- 用 x_0 取代 (10.1) 中的 x^*，我们得到

$$f(x) \approx \underbrace{f(x_0) + (x - x_0) f'(x_0)}_{\text{Is this a straight line?}}. \tag{10.2}$$

- **统计追求简单实用。**

3° 积分版本 (integral version)

- 将 $f(x_0)$ 从 (10.1) 的左手边移到右手边，我们得到均值定理的两个版本：

$$\underbrace{f(x) - f(x_0)}_{(A)} = \underbrace{(x - x_0)f'(x^*)}_{(B)} = \underbrace{\int_{x_0}^{x} f'(y)\,\mathrm{d}y}_{(C)}. \tag{10.3}$$

— (A)=(B) 为**均值定理**，在 x–$f(x)$ 平面上的几何解释，见图 10.1(a)。

— (B)=(C) 在 x–$f'(x)$ 平面上的几何解释，见图 10.1(b)。

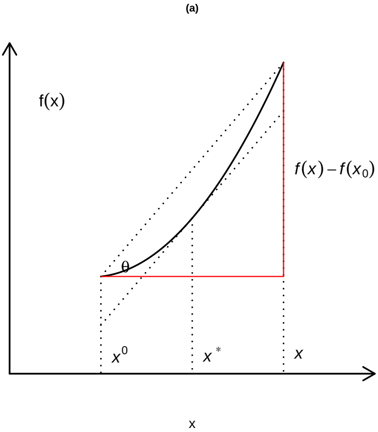
- 因此，一阶泰勒展开的积分版本是

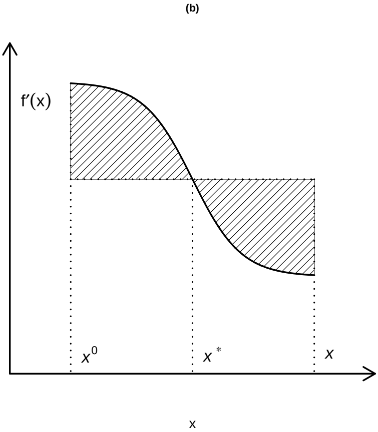
$$f(x) = f(x_0) + \int_{x_0}^{x} f'(y)\,\mathrm{d}y. \tag{10.4}$$

该公式的最大优点是：**不包含难处理的 x^***。

— (A)=(C) 为**积分版本**，可用 cdf 和 pdf 来解释。

图 10.1: (a) 均值定理; (b) $(x - x_0) f'(x^*) = \int_{x_0}^{x} f'(y)\, \mathrm{d}y$





4° 小 o 版本 (small o version)

- 从 (10.1), 我们有

$$f'(x^*) = \frac{f(x) - f(x_0)}{x - x_0}.$$

- 在上式的两边取**极限**, 我们得到

$$f'(x_0) = \lim_{x \to x_0} f'(x^*) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}, \tag{10.5}$$

这就是 $f(x)$ 的**导数**在 $x = x_0$ 处求值**的定义**。

- 我们将 (10.5) 重写为

$$0 = \lim_{x \to x_0} \frac{f(x) - f(x_0) - (x - x_0) f'(x_0)}{x - x_0};$$

即 $f(x) - f(x_0) - (x - x_0) f'(x_0) = o(x - x_0)$。

- 因此, **一阶泰勒展开的小 o 版本**为

$$f(x) = f(x_0) + (x - x_0) f'(x_0) + o(x - x_0). \tag{10.6}$$

10.2 所包含的思政元素

5° 辩证统一的哲学思想

- (10.1) 表明:

— **数学追求形式上的完美**。

— **函数族**在一定条件下 (e.g., 连续且一阶导存在) 可以转换, 困难点在于如何找到 x^*?

— 可引伸为: **学习成绩之好坏 (or 贫富、命运)**在一定条件下 (e.g., 勤奋与刻苦) 可以相互转换, 困难点 在于如何找到适合自己的大学学习方法 (or 机遇、贵人、外部的帮助、情商、综合素养)?

- 公式 (10.2) 表明: 在一定的范围内 (例如在 x_0 的一个很小的邻域内)**曲线**[即 $f(x)$] 可以用通过 $(x_0, f(x_0))$ 这一点的**切线**来近似。

— 说明**统计追求简单实用**。避免了找 x^* 这个困难点。

6° 知识点之间的关联

- 公式 (10.3) 和 (10.4), 将**均值定理、导数、积分、积分版本、cdf、pdf**串联起来。使这些知识形成一个有机的整体。

- 公式 (10.6) 的导出, 将**极限、导数、小 o 版本**关联起来。

7° 一元非线性 (包括积分) 方程之解

- **计算数学基础问题**: 为了求一元非线性方程之根, 存在的**牛顿方法** (Newton's Method) 对初始值之选取是十分敏感的, 它经常不收敛或收敛到一个错误的根。
- **寻根问题**: 给定一个定义于实轴 $\mathbb{R} \triangleq (-\infty, \infty)$ 上的单变量的连续函数 $g(x)$, 需要找一点 $x^* \in \mathbb{R}$, 使得 $g(x^*) = 0$。这些问题通常被称为寻根问题。
- **用统计的语言叙述寻根问题**: 设 $g(\theta)$ 是一个单变量的非线性函数, 并且对如下的方程存在唯一的根 θ^*:

$$g(\theta) = 0, \quad \theta \in \Theta \subseteq \mathbb{R}, \tag{10.7}$$

其中 Θ 是参数空间或实轴 $\mathbb{R}$ 中的一个区间/子集。

$8^\bullet$ 牛顿方法 (Newton's Method)

- 在使用者的计算工具箱中，第一个想到的就是**牛顿方法**。求解方程 (10.7) 之唯一根 θ^* 的牛顿迭代为:

$$\theta^{(t+1)} = \theta^{(t)} - g(\theta^{(t)})/g'(\theta^{(t)}), \tag{10.8}$$

其中 $\theta^{(t)}$ 表示 θ^* 的第 t 次迭代。

- **牛顿方法的缺点**: 它依赖于初始值之选取，它经常不收敛或收敛到一个错误的解。

$9^{\bullet}$ 上穿求解 (Upper-crossing/Solution, 简记为 US) 算法

- **US 算法中一个不失一般性的假设**: 假设直接求解非线性方程 (10.7) 的唯一根非常困难。不失一般性, 我们假设

$$g(\theta) > 0, \quad \forall\, \theta < \theta^* \quad \text{和} \quad g(\theta) < 0, \quad \forall\, \theta > \theta^*. \tag{10.9}$$

当 $\theta < \theta^*$ 时, 如果 $g(\theta) < 0$, 那么我们可以在 (10.7) 的两边同乘以 -1, 获得一个新的方程 $-g(\theta) = 0$, 它满足假设 (10.9) 并具有相同的根 θ^*。

- **一阶导下界 (First–derivative Lower Bound, FLB) 函数方法**: 设 $g'(\cdot)$ 存在, 并且它有一个全局下界函数 $b(\cdot)$, 满足

$$g'(\theta) \geqslant b(\theta), \quad \forall\, \theta \in \Theta, \tag{10.10}$$

其中 $b(\theta)$ 称为**一阶导下界 (FLB)** 函数。我们建立

$$U(\theta|\theta^{(t)}) \triangleq g(\theta^{(t)}) + \int_{\theta^{(t)}}^{\theta} b(z)\,\mathrm{d}z, \quad \forall\, \theta, \theta^{(t)} \in \Theta, \tag{10.11}$$

是 $g(\theta)$ 在 $\theta = \theta^{(t)}$ 处的一个 U–函数。

10° 解 U–方程

- 求方程 (10.7) 的唯一根 θ^* 等价于解如下简单的**U–方程:** $U(\theta|\theta^{(t)}) = 0$, 其解可表示为

$$\theta^{(t+1)} = \textbf{sol}\,\{U(\theta|\theta^{(t)}) = 0,\ \forall\ \theta, \theta^{(t)} \in \Theta\}. \tag{10.12}$$

- 由 (10.12) 给出的解 $\theta^{(t+1)}$, 一般来说, 具有显式表达式。通常, U–方程 $U(\theta|\theta^{(t)}) = 0$ 可以是任何具有解析解的方程, 例如: 二次方程, 甚至是线性方程。