

- 中图分类法:: [O212.1 一般数理统计](#)
- 叙述词法:: [#placeholder/index/thesaurus](#)
- 主题:: [#placeholder/index/topic](#)

统计至简 - Chap 4 离散随机变量 (Review)

Intro

- 文件: Book5Ch04 离散随机变量 统计至简 鸢尾花书 __ 从加减乘除到机器学习.pdf
- 版本: GitHub commit [10dc2e6](#)
- Bilibili (尚未发布 2023-07-14): <https://space.bilibili.com/513194466/video>

修改意见

目录 (导图)

- typo: 思维导图的 二元 - 分布函数 分支下 `joint PMF` 前缺少空格

4.1 随机变量

无

4.2 期望

无

4.2 方差

勘误建议: 本节编号与上一节重复, 都是 `4.2` 需要修改

- Page 15:
 - 建议
 - 原文引述
 - 统计中, 样本的方差计算方法类似上式, 不过要将分母中的 n 换成 $n-1$ 。
 - 修改意见
 - 是否需要解释一句为什么?
 - 本书后面也许会涉猎 「无偏估计」 这个概念? (我还没有看后面的章节)
- page 16
 - 勘误建议
 - 引述 「请大家尝试使用 (25) 计算 (20) 的方差。」
 - 意见 「请大家尝试使用 (23) 计算 (20) 的方差。」
 - 勘误建议
 - 引述 「下面我们聊聊 (25) 的几何含义」
 - 意见 「下面我们聊聊 (23) 的几何含义」
 - 建议 (于几何意义)

- 我认为 与解释下面的等式相比

$$\text{var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

, 解释 "移项之后的加法" 更为直观:

$$\mathbb{E}(X^2) = \text{var}(X) + \mathbb{E}(X)^2$$

这个式子的解释是: 散点相对于原点的离散程度 = 散点相对于质心的离散程度 + 质心相对于原点的偏离

- 我认为这个解释更好的原因有二
 - 一方面, 我认为 思考 "量的累加" 比 思考 "量的减少" 更加直观.
 - 另一方面
 - 正如 期望 对应 力学中刚体的 质心 (刚体平动的惯性量)
 - 此处的 方差 对应 力学中刚体的 转动惯量 (刚体旋转的惯性量)
 - 刚体相对于原点 A_O 的转动惯量 = 刚体相对于过质心的轴 A_C 的转动惯量 + 质心关于 A_O 的转动惯量

4.3 累积分布函数

- 问题
 - 我无法理解 CDF 的图像 (图11)
- 建议
 - 关于公式 (33), 我认为最后一项写作如下形式易读性更高:

$$\Pr(\{X = x\} \cap \{Y = y\})$$

4.4 二元离散随机变量

无

4.5 协方差

- Page 21
 - 建议
 - 原文引述: X 的取值为 $x^{(i)}$ ($i = 1, 2, \dots, n$)
 - 修改意见: X 在 $\{x^{(i)} \mid i = 1, 2, \dots, n\}$ 中取值
 - 原因: 我认为这样表意更清晰

4.6 边缘概率

无

4.7 条件概率

- 原文引述: 也就是说, 对于 $p_{X|Y}(x|y)$, $\{Y = y\}$ 是新的样本空间。
- 修改意见: [...], $\{Y = y\}$ 定义了一个新的样本空间。

Page 27

- 原文: $p_Y(2)$ 为一定值
- 更爱建议: 「 $p_Y(2)$ 是一个定值」 或 「 $p_Y(2)$ 为定值」
 - 由于中文分词问题, 此句可能误读作 「为 + 一定 + 值」

4.8 独立性

无

4.9 以鸢尾花数据为例: 不考虑分类标签

无

4.10 以鸢尾花数据为例: 考虑分类标签

无

4.10 再谈概率 1: 展开、折叠

勘误建议: 本节编号与上一节重复, 都是 4.10, 需要更新编号