

Part III-B: Probability Theory and Mathematical Statistics

Lecture by 李漫漫

Note by THF

2024 年 9 月 8 日

目录

1 第一章	2
1.1 随机事件	2
1.1.1 现象	2
1.1.2 随机试验	3
1.1.3 样本	3
1.1.4 随机事件	4
1.2 事件关系与运算	4
1.3 事件的概率	5
1.3.1 古典概型	6
1.3.2 几何概型	7

概述

资源

公众号：狗熊会、大数据文摘，好玩的数学

MOOC：爱课程，Coursera，Edx，网易公开课等

教师要求

教材：概率论与数理统计第二版

参考：The Lady Tasting Tea, 程序员数学之概率统计, ...

学习目的：自问自答，自言自语

考核及成绩组成：

期中（10）

作业与考勤（10）

期末（70）

MOOC（10）

课程简介

概率：Probability

统计：Statistics

概率论与数理统计：Probability theory and Mathematical statistics

Notation. 第一章重要但不突出

从概率到概率论：新增时间（随机事件、样本空间变化）

从统计到数理统计：统计最开始为记录性质，后来衍生出预测，通过数学模型引入数理统计

类似的还有政府统计、经济统计等

2000-2015 年间，IT 时代逐渐转换为 DT(Data Technology) 时代，大数据逐渐占时代主体

1 第一章

1.1 随机事件

1.1.1 现象

确定性现象：一定条件下必然发生

随机现象强调统计规律性

Notation. 统计规律性:

1. 每次试验前不能预测结果
2. 结果不止一个
3. 大量试验下有一定规律

Example. 星际旅行时宇航员看到的现象不是随机现象:

对星际旅行的人而言, 无法完成大量试验

宇航员观测到的结果无规律, 只能称为不确定现象 (Uncertain)

Example. 扔一个骰子不能预测结果, 但可以知道结果是 1, 2, 3, 4, 5, 6 的一个, 因此观察扔骰子是随机现象 (Random)

1.1.2 随机试验

随机试验 (E): 研究随机现象时进行的实验或观察等

Notation. 随机试验的特性:

1. 可以在完全相同的条件下重复进行
2. 试验的可能结果在试验前已知
3. 试验的结果不可预测

1.1.3 样本

在随机试验中, 不可再分的最简单结果成为样本点 ω , 全体样本点组成样本空间 Ω

Notation. 随机事件是基本事件的集合

Example. 扔骰子存在 6 个基本事件, 可以产生 2^6 个随机事件, 其中样本空间 $\Omega = \{\omega | \omega \in [1, 6], \omega \in \mathbb{R}\}$

Example. 1. 射击时用 ω_i 表示击中 i 环, 样本空间为:

$$\Omega = \{\omega_0, \omega_1, \omega_2, \dots, \omega_{10}\}.$$

2. 微信用户每天收到信息条数的取值范围是 $[0, +\infty)$, 样本空间为无限集:

$$\Omega = \{N | N \geq 0, N \in \mathbb{R}\}.$$

3. 电视机的寿命样本空间为 $\Omega = \{t | t > 0\}$, 为连续的非负实数集
4. 投掷两枚硬币, 样本空间为 $\Omega = \{(x, y) | x, y = 0, 1\}$, 其中 0, 1 分别代表正面和背面

Notation. 1. 样本点可以不是数

2. 样本空间可以是无限集

1.1.4 随机事件

1.2 事件关系与运算

- 1.1. $A \subset B$: A 发生必然 B 发生
- 1.2. $A = B$: $A \subset B, B \subset A$
2. $A \cup B$: A 和 B 至少有一个发生
- 2.1 $A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i$
3. $A \cap B$: A 和 B 只发生一个
- 4.1. A, B 互斥: 不能同时发生: $AB = \emptyset$
- 4.2. A, B 对立: 非此即彼: $A \cup B = \Omega$
5. $A - B$: $A\bar{B}$ 或 $A(\Omega - B)$, 或 A 发生但 B 不发生

Notation. $A - B = A\bar{B} \subset A$, $B - A = B\bar{A} \subset B$

当 $AB = \emptyset$ 时, $A - B = A, B - A = B$

Notation. $P(\Omega) = 1, P(\emptyset) = 0$, 且 $P(\Omega) + P(\emptyset) = 1$, 即 Ω 与 \emptyset 互斥

6. 结合律: $(A \cup B) \cup C = A \cup (B \cup C)$
7. 分配律: $(AB) \cup C = (A \cup C)(B \cup C)$, $(A \cup B)C = AC \cup BC$
8. 交换律: $A \cup B = B \cup A, AB = BA$

Notation. 德摩根律:

$$\overline{\bigcup_{i=1}^n A_i} = \bigcap_{i=1}^n \bar{A}_i.$$

$$\overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \bar{A}_i.$$

Example.

$$\overline{A \cup B} = \bar{A}\bar{B}.$$

$$\overline{(A \cup B) \cup C} = \overline{A \cup B}\bar{C} = \dots$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}.$$

1.3 事件的概率

概率分类：

$$\left\{ \begin{array}{l} \text{主观概率} \\ \text{统计概率} \\ \text{古典概型} \\ \text{几何概型} \end{array} \right\}.$$

Notation. 德摩根、蒲丰、皮尔逊、维纳均进行过投掷硬币的试验，随着试验次数的增加，出现正面的频率逐渐接近 0.5

大数定律说明，该事件的概率为 0.5

Definition. 统计概率：A 为试验 E 的一个事件，随着重复次数 n 的增加，A 的频率接近于某个常数 p，定义事件 A 的概率为 p，记为 $P(A) = p$

频率的特性：

1. 非负性： $f_n(A) \in [0, 1]$
2. 规范性： $f_n(\Omega) = 1$
3. 有限可加性： A_i 两两互斥，则 $f_n(\sum_{i=1}^n A_i) = \sum_{i=1}^n f_n(A_i)$

Definition. 主观概率：人对某个事件发生与否的可能性的估计

Definition. 完备事件组： A_1, A_2, \dots, A_n 两两互斥，且 $P(\sum_{i=1}^n A_i) = 1$ 或 $\sum_{i=1}^n A_i = \Omega$ ，则称 $A_1 \rightarrow A_n$ 为完备事件组（不重不漏）

Example. A, \bar{A} 是完备事件组

1.3.1 古典概型

古典概型特点：有限等可能性（基本事件数有限，基本事件发生的可能性相等）

Notation. 概率计算：

$$P(A) = \frac{m}{n} = \frac{n(A)}{n(\Omega)}.$$

Example. 某年级有 6 人在 9 月份出生，求 6 个人中没有人同一天过生日的概率

基本事件总数： 30^6

目标事件： $30 \cdot 29 \cdot 28 \cdot 27 \cdot 26 \cdot 25 = P_{30}^6$

概率：

$$P(A) = \frac{P_{30}^6}{30^6}.$$

Example. 有 N 个乒乓球中有 M 个白球、 $N - M$ 个白球，任取 $n (n < N)$ 个球，分有放回和不放回，求取到 m 个黄球的概率

1. 不放回：

基本事件总数： C_N^n

目标事件： $C_M^m C_{N-M}^{n-m}$

概率：

$$P = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}, n = \max\{0, n - (N - M)\}, \dots, \min\{n, M\}.$$

2. 有放回：

$$P = \frac{C_n^m M^m (N - M)^{n-m}}{N^n} = C_n^m \left(\frac{M}{N}\right)^m \left(1 - \frac{M}{N}\right)^{n-m}, m \in [0, n].$$

注意到该概率为伯努利分布 $C_n^m B(n, \frac{M}{N})$

匹配问题：

Example. 麦克斯韦-玻尔兹曼统计问题：

n 个质点随机落入 $N (N > n)$ 个盒子，盒子容量不限，设 A 表示指定的 n 个盒子各有一个质点， B 表示恰好有 n 个盒子装一个质点

基本事件总数: N^n

A 考虑顺序, 即:

$$P(A) = \frac{n!}{N^n}.$$

同理:

$$P(B) = \frac{C_N^n}{N^n}.$$

1.3.2 几何概型

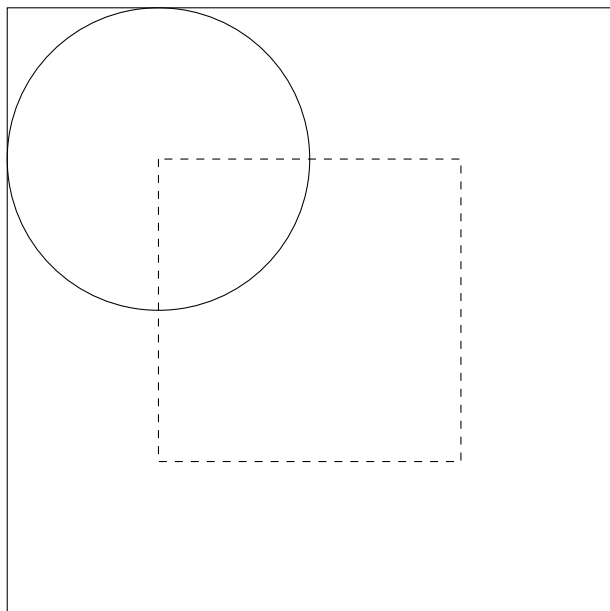
几何概型特点: 使用事件所对应的**几何度量**计算

$$P(A) = \frac{m(A)}{m(\Omega)}.$$

Notation. 度量: 面积、体积、长度等描述几何量大小的测度方式

Example. 地面铺满 2 dm 的地砖, 向地面投掷一个 $r = 0.5$ dm 的光盘, 求光盘不与边线相交的概率

如图:



课后习题: A 组 8 题, B 组 3 题