

Part III-B: Probability Theory and Mathematical Statistics

Lecture by 李漫漫

Note by THF

2024 年 9 月 13 日

目录

1 第一章	3
1.1 随机事件	3
1.1.1 现象	3
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
1.4 公理化	9
1.5 条件概率与乘法公式	12
1.6 全概率公式	12
1.7 贝叶斯公式	12
1.8 独立性	14

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 $\Omega = \{x|x \in [1, 6], x \in \mathbb{R}\}$

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:

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

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

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

$$P\left(\sum_{i=1}^n A_i\right) = 1.$$

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、 $N - M$ 个黄球，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 $C_n^m B\left(n, \frac{M}{N}\right)$

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

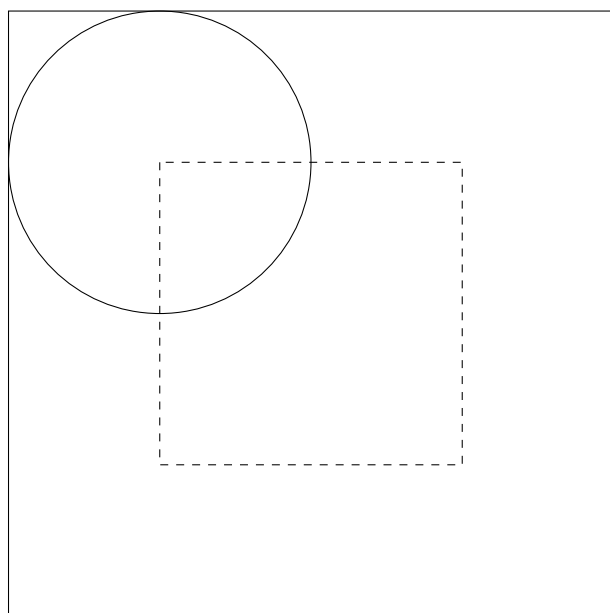
几何概型特点: 使用事件所对应的**几何度量**计算

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

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:

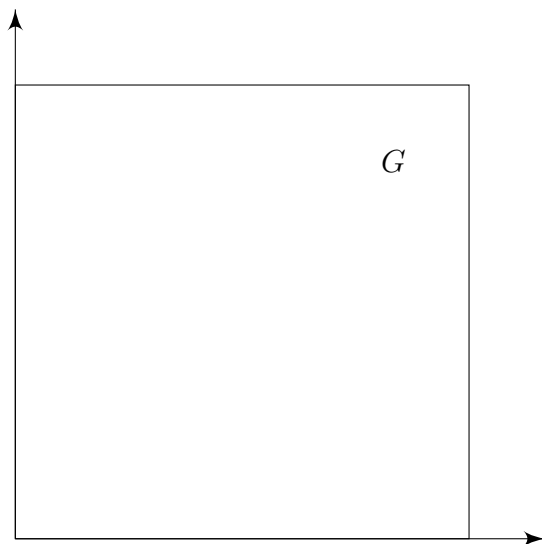


课后习题：A 组 8 题，B 组 3 题

Example. 两人相约 8-9 点间在某地相见，先到的人等待 20 分钟后离去，求二人会面的概率

设 (x, y) 分别表示两人到达的时刻

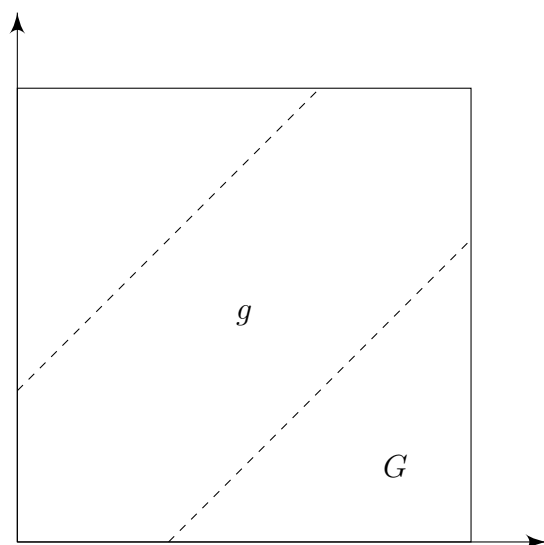
设 G 为样本空间，绘制样本空间：



由题：两人到达的时间之差的绝对值小于 20 分钟 ($\frac{1}{3}$ 小时)，即：

$$|x - y| \leq \frac{1}{3}.$$

将事件绘制：



$$P(g) = \frac{m(g)}{m(G)} = \frac{S(g)}{S(G)} = \frac{1 - \left(\frac{2}{3}\right)^2}{1} = \frac{4}{9}.$$

Notation. 几何概型的特点：

1. 非负性：

$$P(A) \in [0, 1].$$

2. 规范性：

$$P(\Omega) = 1.$$

3. 可列可加性：

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$

1.4 公理化

$$(\Omega, \mathcal{F}, p).$$

Definition. Ω ：随机试验所产生的所有样本点的集合

\mathcal{F} ：集合内所有子集为元素的集合

$P(X)$ ：概率函数

Axiom. 非负性:

$$P(A) \geq 0, A \in \mathcal{F}.$$

Axiom. 规范性:

$$P(\Omega) = 1.$$

Axiom. 可列可加性: 对两两互斥的事件 A_1, A_2, \dots ,

$$P\left(\sum_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} P(A_i).$$

从三条公理得出的性质:

Notation. 1. $P(\emptyset) = 0$

2. 有限可加性:

$$\sum_{i=1}^n P(A_i) = P\left(\sum_{i=1}^n A_i\right).$$

3. $P(\bar{A}) = 1 - P(A)$

4. $A \subset B \implies P(B - A) = P(B) - P(A)$

5. $A \subset B \implies P(A) \leq P(B)$

6. $P(A \cup B) = P(A) + P(B) - P(AB)$

Notation. 6.1.

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC).$$

6.2.

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^n P(A_i A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n P(A_i A_j A_k) \\ &\quad \dots + (-1)^{n-1} P\left(\prod_{i=1}^n A_i\right). \end{aligned}$$

Example. 从有号码 $1, 2, \dots, n$ 的 n 个球中有放回地取 m 个球, 求取出的 m 个球中最大号码为 k 的概率

$$P\{k=1\} = \left(\frac{1}{n}\right)^m.$$

逐个列举计算较复杂, 记事件 B_k 为取出的 m 个球最大号码不超过 k , 只需保证每次摸出的球都不超过 k 即可:

$$P(B_k) = \frac{k^m}{n^m}.$$

又有 $P(A_k) = P(B_k) - P(B_{k-1})$, 且 $B_{k-1} \subset B_k$

所以:

$$P(A_k) = \frac{k^m}{n^m} - \frac{(k-1)^m}{n^m}.$$

Example. 匹配问题: n 个学生各带有一个礼品, 随机分配礼品, 设第 i 个人抽到自己的礼品称为一个配对, 求至少有一个配对的概率

设 A_i 是第 i 个人抽到自己的礼品, A 为目标事件, 则:

$$A = \bigcup_{i=1}^n A_i.$$

$$P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

$$P(A_i A_j) = \frac{(n-2)!}{n!} = \frac{1}{P_n^2}.$$

$$P(A_i A_j A_k) = \frac{1}{P_n^3}.$$

.....

$$P\left(\prod_{i=1}^n A_i\right) = \frac{1}{n!}.$$

$$P(A) = P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^n P(A_i A_j) + \dots$$

1.5 条件概率与乘法公式

Definition.

$$P(A) > 0, P(B|A) = \frac{P(AB)}{P(A)}.$$

即：在 A 发生的条件下，B 发生的概率

Definition. 乘法公式：

$$P(AB) = P(A)P(B|A) = P(B)P(A|B).$$

Notation. A, B 独立： $P(AB) = P(A)P(B)$

结合乘法公式：

$$P(B) = P(B|A).$$

$$P(A) = P(A|B).$$

1.6 全概率公式

Corollary. 事件 A_1, A_2, \dots, A_n 为完备事件组，事件 $B \subset \Omega = \bigcup_{i=1}^n A_i$ ，则：

$$P(B) = \sum_{i=1}^n P(A_i)P(B|A_i).$$

Notation. 此时完备事件组的情况应该已知，通过完备事件组 A 的辅助可以求得较复杂事件 B 的概率

1.7 贝叶斯公式

Corollary.

$$P(A_k|B) = \frac{P(A_k B)}{P(B)} = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^n P(A_i)P(B|A_i)}.$$

贝叶斯公式被称为“逆概率公式/后验公式”，其中事件 B 更可能是事件的结果，将事件组 A 看作结果出现的原因，则贝叶斯公式是一个从“结果”推“原因”的可能性的公式

Notation. 对比一般公式：事件 A 导致 B，求 B 发生的概率

贝叶斯公式：事件 A 导致 B，A 中的一个事件 A_i 导致 B 发生的概率

Axiom. 条件概率的公理:

1. 非负性: $P(A) \in [0, 1]$
2. 规范性: $P(\Omega|A) = 1$
3. 可列可加性:

$$P\left(\sum_{i=1}^{\infty} B_i|A\right) = \sum_{i=1}^{\infty} P(B_i|A).$$

Corollary.

$$P(\bar{B}|A) = P(\Omega - B|A) = P(\Omega|A) - P(B|A).$$

Corollary.

$$\begin{aligned} P(B_1 \cup B_2) &= P(B_1) + P(B_2) - P(B_1 B_2). \\ \implies P(B_1 \cup B_2|A) &= P(B_1|A) + P(B_2|A) - P(B_1 B_2|A). \end{aligned}$$

Corollary. 乘法公式:

$$P(ABC) = P(A(BC)) = P(A)P(BC|A) = P(A)P(B|A)P(C|AB).$$

Example. 8 个红球 2 个白球, 求前三次结果是“红红白”的概率:

1. 不放回取 3 个 (和一次取三个球相同)

所有可能性: $10 \times 9 \times 8$

目标事件: $8 \times 7 \times 2$

或使用乘法公式: 设 A_i 为第 i 次取到红球, 目标事件可表示为 $A_1 A_2 \bar{A}_3$
概率:

$$P(A_1 A_2 \bar{A}_3) = P(A_1)P(A_2|A_1)P(\bar{A}_3|A_1 A_2) = \frac{8}{10} \times \frac{7}{9} \times \frac{2}{8} = \frac{7}{45}.$$

2. 每次取后放回, 并加入两个同色的球, 取 3 次 (不能使用古典概型)
概率:

$$P(A_1 A_2 \bar{A}_3) = \frac{8}{10} \times \frac{8}{12} \times \frac{2}{14} = \frac{8}{105}.$$

Example. 某疾病的发病率为 0.0004, 患病检测呈阳性的概率为 0.99, 误诊为阴性的概率为 0.01, 误诊为阳性的概率为 0.05, 不患病检测呈阴性概率为 0.95, 一个人检测呈阳性, 求其患病的概率

设阳性为 A , 患病为 B

则:

$$P(A|B) = 0.99, P(A|\bar{B}) = 0.05, P(B) = 0.0004.$$

要求: $P(B|A)$

使用贝叶斯公式:

$$\begin{aligned} P(B|A) &= \frac{P(AB)}{P(A)} = \frac{P(B)P(A|B)}{P(AB) + P(A\bar{B})} \\ &= \frac{P(B)P(A|B)}{P(B)P(A|B) + P(A|\bar{B})P(\bar{B})} = 0.0079. \end{aligned}$$

1.8 独立性

Definition. A, B 独立, 则: $P(A|B) = P(A)$

Notation. 证明独立性:

1. $P(A)P(B) = P(AB)$

Notation. 独立事件的特点:

1. A, B 独立有: A, B 所有的组合 (包含补集) 均独立
2. A, B 独立的充要条件: $P(A|B) = P(A)$ or $P(B|A) = P(B)$
3. \emptyset 与任何随机事件独立, Ω 与任何随机事件独立

对于三个事件相互独立:

$$\begin{cases} P(AB) = P(A)P(B) \\ P(AC) = P(A)P(C) \\ P(BC) = P(C)P(C) \\ P(ABC) = P(A)P(B)P(C) \end{cases}.$$

对比乘法公式: $P(ABC) = P(A)P(B|A)P(C|AB)$

Definition. 相互独立:

有 A_1, A_2, \dots, A_n 事件组, 对 $\forall s \in [2, n]$ 个事件 $A_{k_1}, A_{k_2}, \dots, A_{k_s}$ 均有:

$$P\left(\prod_{n=1}^s A_{k_n}\right) = \prod_{n=1}^s P(A_{k_n}).$$

称事件 A_1, A_2, \dots, A_n 相互独立

Definition. 两两独立: 对事件 A_1, A_2, \dots, A_n , 若任意两个事件独立, 则称为两两独立

Notation. 相互独立一定两两独立, 反之不一定

Notation. 相互独立事件组的性质:

1. 事件 A_1, A_2, \dots, A_n 相互独立, 将其中任意部分改为对立事件, 事件组仍为相互独立
2. 事件相互独立, 将事件组任意分为两组 (或多组), 对组内事件进行“并、交、差、补”操作后, 事件间依然相互独立

独立重复实验

Definition. E_1, E_2 中一个试验的任何结果和另一个试验的任何结果相互独立, 则试验相互独立; 若 n 个独立试验相互独立且试验相同, 称 E_1, E_2, \dots, E_n 为 n 次独立重复实验, 或 n 重独立试验

Example. 扔硬币和掷骰子为独立试验, 其中扔硬币为伯努利试验 (只有两个结果)

Definition. n 重独立试验 E 中, 每次试验都是伯努利试验 (可能结果只有两个), 称 E 为 n 重伯努利试验

1. 二项概率公式: 成功 k 次的概率记为 $P_n(k)$, 假定前 k 次成功, 后 $n-k$ 次失败, 则

$$P_i = p^k (1-p)^{n-k}.$$

指定事件 A 发生的位置有 C_n^k 种, 则:

$$P_n(k) = C_n^k p^k (1-p)^{n-k}.$$

称为二项概率公式

2. 几何概率公式: 首次成功恰好发生在第 k 次的概率记为 $G(k)$, 设前 $k-1$ 次失败, 则:

$$G(k) = q^{k-1}p.$$

可以验证: $\sum G(k) = 1$

3. 负二项概率: 需要成功 r 次, 第 r 次成功恰好发生在第 k 次的概率记为 $G_r(k)$, 设前 $k-1$ 次试验有 $r-1$ 次成功, 则:

$$G_r(k) = C_{k-1}^{r-1} p^r q^{k-r}.$$

同样有: $\sum G_r(k) = 1$