



概率论笔记

作者：肖程哲

时间：August 14, 2022



苟日新，日日新，又日新

目录

第 1 章 概率基础	1	3.3 估计与预测	29
1.1 概率空间	1	3.3.1 delta 法	29
1.1.1 随机事件	1	3.3.2 预测	29
1.1.2 概率空间	2	3.4 熵与信息	29
1.1.3 概率的性质	3	3.5 其他特征	29
1.2 古典概型与几何概型	6		
1.2.1 古典概型	6	第 4 章 常见分布	30
1.2.2 几何概型	6	4.1 离散分布	30
1.2.3 Bertrand 奇论	8	4.1.1 均匀分布	30
1.3 条件概率	9	4.1.2 伯努利分布	30
1.4 独立性	12	4.1.3 二项分布	31
第 1 章 练习	14	4.1.4 几何分布	32
第 2 章 随机变量	12	4.1.5 负二项分布	33
2.1 随机变量的分布	12	4.1.6 多项分布	34
2.2 多元随机变量	14	4.1.7 泊松分布	35
2.2.1 边际分布	15	4.1.8 超几何分布	37
2.2.2 条件分布	15	4.2 连续分布	38
2.2.3 独立	16	4.2.1 均匀分布	38
2.3 随机变量的函数	16	4.2.2 指数分布	38
2.3.1 分布函数法	17	4.2.3 伽马分布	40
2.3.2 Copula	18	4.2.4 贝塔分布	41
2.3.3 概率密度函数法	19	4.3 正态分布及其导出分布	43
2.3.4 矩母函数法	20	4.3.1 正态分布	43
2.3.5 次序统计量	20	4.3.2 卡方分布	44
第 3 章 随机变量的数值特征	22	4.3.3 F 分布	45
3.1 期望	22	4.3.4 t 分布	46
3.1.1 均值	22	4.3.5 柯西分布	47
3.1.2 方差	23	4.4 各分布间关系	49
3.1.3 协方差	24		
3.1.4 条件期望	25		
3.2 矩母函数与特征函数	26		
3.2.1 矩	26		
3.2.2 矩母函数	27		
3.2.3 联合特征函数	28		
3.2.4 特征函数	28		
第 5 章 概率极限	51		
5.1 收敛	51		
5.2 大数定理	53		
5.3 中央极限定理	54		
第 A 章 测度论基础	56		
A.1 可测空间和可测映射	56		
A.1.1 集合及其运算	56		
A.1.2 集合系	57		

第1章 概率基础

考试重点

□ 概率空间的定义

□ 古典概型与几何概型

□ 条件概率与独立

□ 伯努利试验

1.1 概率空间

1.1.1 随机事件

定义 1.1 (样本空间)

随机试验可能出现的基本结果称为样本点 (sample point), 记为 ω 。样本的全体构成样本空间 (sample space), 记为 $\Omega = \{\omega\}$ 。若样本点有至多可数个, 则称为离散样本空间, 若不可数, 则称为连续样本空间。



定义 1.2 (事件的古典定义)

样本点 ω 的集合称为 (随机) 事件 (event)。



随机现象被抽象为集合。事件间的包含、相等、互不相容关系, 对应成集合间的包含、相等、不相交; 逻辑运算 (且、或、非等) 对应成集合论运算 (交、并、补等)。

例题 1.1 设 A 、 B 、 C 是某个随机现象的三个事件, 则

- 事件 “ A 与 B 发生, C 不发生” 可表示为: $AB\bar{C}$.
- 事件 “ A 、 B 、 C 中至少有一个发生” 可表示为: $A \cup B \cup C$.
- 事件 “ A 、 B 、 C 中至少有两个发生” 可表示为: $AB \cup AC \cup BC$.
- 事件 “ A 、 B 、 C 中恰好有两个发生” 可表示为: $AB\bar{C} \cup A\bar{B}C \cup \bar{A}BC$.
- 事件 “ A 、 B 、 C 同时发生” 可表示为: ABC .
- 事件 “ A 、 B 、 C 都不发生” 可表示为: $\bar{A} \cap \bar{B} \cap \bar{C}$.
- 事件 “ A 、 B 、 C 不全发生” 可表示为: $\bar{A} \cup \bar{B} \cup \bar{C}$.

性质 集合的运算性质:

- 交换律

$$A \cup B = B \cup A, \quad AB = BA \quad (1.1)$$

- 结合律

$$(A \cup B) \cup C = A \cup (B \cup C), \quad (AB)C = A(BC) \quad (1.2)$$

- 分配律

$$(A \cup B) \cup C = A \cup (B \cup C), \quad (1.3)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C). \quad (1.4)$$

- 对偶律 (De Morgan 公式)

$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \quad (1.5)$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}. \quad (1.6)$$

注 并与补是集合中最基本的运算:

- 交的运算可通过并与对立来实现 (对偶律).

- 差的运算可通过对立与交来实现 ($A - B = A\bar{B}$) .

1.1.2 概率空间

为方便概率的定义，避免不可测集的出现，并不把 Ω 的一切子集作为事件。

定义 1.3 (事件域)

事件构成的全体称为事件域 \mathcal{F} ，是 Ω 的子集族 (collection of subsets)，应满足 σ 代数 的要求：

- $\emptyset \in \mathcal{F}$ ，代表无事发生；
- $A \in \mathcal{F} \implies A^c \in \mathcal{F}$ ，即 \mathcal{F} 对补集运算 (逻辑上的非) 封闭；
- $A_1, \dots, A_n, \dots \in \mathcal{F} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ ，即 \mathcal{F} 对可数并运算封闭。



注 可数性是为了在数学上能够恰当地处理无穷的概念，术语中的 σ 指的就是可数并。由对偶原理可得 σ 域同时对可数并运算封闭。即 σ 域对逆，并，交，差的可数次运算封闭。

事件域根据问题的不同要求适当选取。在概率定义没有困难时，应尽量取得大，通常以 Ω 的一切子集作为事件域。当 Ω 给定后，若某些子集必须作为事件处理，能否找到包含他们的 σ 域？

命题 1.1

若给定 Ω 的一个非空集族 \mathcal{G} ，必存在 Ω 上唯一的 σ 域 $m(\mathcal{G})$ ，满足下列性质：

- 包含 \mathcal{G}
- 若其他 σ 域包含 \mathcal{G} ，则必包含 $m(\mathcal{G})$

这个 $m(\mathcal{G})$ 称为包含 \mathcal{G} 的最小 σ 域，或由 \mathcal{G} 扩张而成的 σ 域。



扩张，或者称为延拓，是数学中很重要的一个概念，大抵是将某映射的定义域适当扩大，不改变在初始定义域上的映射取值（注意值域可能是比较抽象的集合，配备了某些操作之后被称为空间），同时在扩大后的定义域上仍然保持某些优良的性质。与此相对的概念是限制，即关心局部上可能更加漂亮的性质，把初始的定义域适当缩小。

证明 由于 Σ 的一切子集构成的集类包含 \mathcal{G} ，所以 m 存在。再取 Σ 上满足此条件的 σ 域之交作为 $m(\mathcal{G})$ 即可。

特别地，实数集 \mathbb{R} 的子集族 $\{(-\infty, x] : x \in \mathbb{R}\}$ 生成的 σ 代数 $\mathcal{B}_{\mathbb{R}}$ 称为 \mathbb{R} 上的 Borel 代数。

定义 1.4 (Borel 集)

设 \mathbb{R}^1 为全集，形为 $[a, b]$ 构成的集类产生的 σ 域称为一维 Borel σ 域，记为 \mathcal{B}_1 ，其中的元素称为一维 Borel 集



若 x, y 为任意实数，由于：

$$\{x\} = \bigcap_{n=1}^{\infty} \left[x, x + \frac{1}{n} \right)$$

$$(x, y) = [x, y] - \{x\}$$

$$[x, y] = [x, y] + \{y\}$$

$$(x, y] = [x, y] + \{y\} - \{x\}$$

因此 \mathcal{B}_1 包含一切开区间，闭区间，单个实数，可列个实数，以及他们经可列次逆，并，交，差运算的集合。

定义 1.5 (概率空间)

设 Ω 为一个样本空间， \mathcal{F} 为定义于其上的一个事件域。定义在事件域(非样本空间) 上的集合函数 P ：
 $\mathcal{F} \rightarrow \mathbb{R}$ 称为概率的条件是：

$$\text{非负性} \quad P(A) \geq 0, \forall A \in \mathcal{F}$$

规范性 $\mathbb{P}(\Omega) = 1$; (如果没有这条就是一般的有限测度)

可列可加性 若 $A_1, \dots, A_n, \dots \in \mathcal{F}$ 两两不交, 即 $A_i \cap A_j = \emptyset, \forall i \neq j$, 则

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

并称 $(\Omega, \mathcal{F}, \mathbb{P})$ 为一个概率空间 (probability space)



1.1.3 概率的性质

性质 概率的性质:

- $P(\emptyset) = 0$;
- $P(A^c) = 1 - P(A)$;
- 若 $A \subseteq B$ 则 $P(A) \leq P(B)$;

例题 1.2 口袋中有编号为 $1, 2, \dots, n$ 的 n 个球, 从中有放回地任取 m 次, 求取出的 m 个球的最大号码为 k 的概率。

解 记事件 A_k 为“取出的 m 个球的最大号码为 k ”。如果直接考虑事件 A_k , 则比较复杂, 因为“最大号码为 k ”可以包括取到 1 次 k 、取到 2 次 k 、...、取到 m 次 k 。为此我们记事件 B_i 为“取出的 m 个球的最大号码小于等于 i ”。则 B_i 发生只需每次从 $1, 2, \dots, i$ 号球中取球即可。由题可知

$$P(B_i) = \frac{i^m}{n^m}, \quad i = 1, 2, \dots, n.$$

又因为 $A_k = B_k - B_{k-1}$, 且 $B_{k-1} \subset B_k$, 所以

$$\begin{aligned} P(A_k) &= P(B_k - B_{k-1}) = P(B_k) - P(B_{k-1}) \\ &= \frac{k^m - (k-1)^m}{n^m}, \quad k = 1, 2, \dots, n. \end{aligned}$$

例题 1.3

$$|P(AB) - P(A)P(B)| \leq \frac{1}{4}$$

解 设 $P(A) \geq P(B)$, 则:

$$P(AB) - P(A)P(B) \leq P(B) - P(B)P(B) = P(B)[1 - P(B)] \leq \frac{1}{4}$$

另一方面有:

$$\begin{aligned} P(A)P(B) - P(AB) &= P(A)[P(AB) + P(\bar{A}B)] - P(AB) \\ &= P(A)P(\bar{A}B) + P(AB)[P(A) - 1] \\ &\leq P(A)P(\bar{A}B) \leq P(A)P(\bar{A}) \\ &= P(A)[1 - P(A)] \leq \frac{1}{4} \end{aligned}$$

命题 1.2 (加法公式)

基础形式:

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

一般形式:

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \\ &\quad \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n). \end{aligned}$$

特别地, 若事件出现个数相同时概率相等, 则可简化为:

$$P\left(\bigcup_{i=1}^n A_i\right) = nP_1 - \binom{n}{2}P_2 + \binom{n}{3}P_3 - \cdots + (-1)^{n-1}P_n$$

推论 1.1

对任意 n 个事件 A_1, \dots, A_n 有

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



例题 1.4 配对问题 在一个有 n 个人参加的晚会上, 每个人带了一件礼物, 且假定各人带的礼物都不相同。晚会期间各人从放在一起的 n 件礼物中随机抽取一件, 问至少有一个人自己抽到自己礼物的概率是多少?

解 以 A_i 记事件“第 i 个人自己抽到自己的礼物”, 则所求概率为 $P(A_1 \cup A_2 \cup \cdots \cup A_n)$ 。因为

$$\begin{aligned} P(A_1) &= P(A_2) = \cdots = P(A_n) = \frac{1}{n} \\ P(A_1 A_2) &= P(A_1 A_3) = \cdots = P(A_{n-1} A_n) = \frac{1}{n(n-1)} \\ P(A_1 A_2 A_3) &= P(A_1 A_2 A_4) = \cdots = P(A_{n-2} A_{n-1} A_n) = \frac{1}{n(n-1)(n-2)} \\ &\vdots \\ P(A_1 A_2 \cdots A_n) &= \frac{1}{n!} \end{aligned}$$

再由概率的加法公式 1.2 得

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + (-1)^{n-1} \frac{1}{n!}$$

显然, 可列可加性可以推出有限可加性。但是一般而言, 由有限可加性不能推出可列可加性。设 $A_i \in \mathcal{F}, i = 1, 2, \dots$ 且两两互不相容, 则由概率的有限可加性只能推出下式成立:

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

这个等式的左边对任意 n 都不超过 1, 因此右边的正项级数收敛。这样应有

$$\lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_i\right) = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(A_i) = \sum_{i=1}^{\infty} P(A_i)$$

即若希望由有限可加性推出可列可加性, 则需要下式成立:

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

可写为:

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

若记

$$S_n = \sum_{i=1}^n A_i$$

则 $S_n \in \mathcal{F}$ 而且 $S_n \subseteq S_{n+1}$, 即 $\{S_n\}$ 是 \mathcal{F} 中一个单调不减的集序列, 这时将上式改写改写为:

$$P(\lim_{n \rightarrow \infty} S_n) = \lim_{n \rightarrow \infty} P(S_n)$$

定义 1.6 (下连续)

对于 \mathcal{F} 上的集合函数 P , 若它对 \mathcal{F} 中任一单调不减的集序列 $\{S_n\}$ 均满足:

$$\lim_{n \rightarrow \infty} P(S_n) = P(\lim_{n \rightarrow \infty} S_n)$$

则称它是下连续的.



定理 1.1

若 P 是 \mathcal{F} 上满足 $P(\Omega) = 1$ 的非负集合函数, 则它具有可列可加性的充要条件为:

- 它是有限可加的;
- 它是下连续的.



证明 由上述推导可知其充分性。对于其必要性, 有限可加也是易知的。设 $\{S_n\}$ 是 \mathcal{F} 中一个单调不减的事件序列, 即

$$\bigcup_{i=1}^{+\infty} S_i = \lim_{n \rightarrow +\infty} S_n$$

若定义 $S_0 = \emptyset$, 则

$$\bigcup_{i=1}^{+\infty} S_i = \bigcup_{i=1}^{+\infty} (S_i - S_{i-1})$$

由于 $S_{i-1} \subset S_i$, 显然诸 $(S_i - S_{i-1})$ 两两不相容, 再由可列可加性得

$$P\left(\bigcup_{i=1}^{+\infty} S_i\right) = \sum_{i=1}^{+\infty} P(S_i - S_{i-1}) = \lim_{n \rightarrow +\infty} \sum_{i=1}^n P(S_i - S_{i-1})$$

由有限可加性得

$$\sum_{i=1}^n P(S_i - S_{i-1}) = P\left(\bigcup_{i=1}^n (S_i - S_{i-1})\right) = P(S_n)$$

所以

$$P(\lim_{n \rightarrow +\infty} S_n) = \lim_{n \rightarrow +\infty} P(S_n)$$

这就证得了 P 的下连续性.

同理定义上连续

定义 1.7 (上连续)

对于 \mathcal{F} 上的集合函数 P , 若它对 \mathcal{F} 中任一单调不增的集序列 $\{B_n\}$ 均满足:

$$\lim_{n \rightarrow \infty} P(B_n) = P(\lim_{n \rightarrow \infty} B_n)$$

则称它是上连续的.



定理 1.2

概率是上连续的



证明 设 $\{B_n\}$ 是单调不增的事件序列，则 $\{\overline{B_n}\}$ 为单调不减的事件序列，由概率的下连续性得

$$\begin{aligned} 1 - \lim_{n \rightarrow +\infty} P(B_n) &= \lim_{n \rightarrow +\infty} [1 - P(B_n)] = \lim_{n \rightarrow +\infty} P(\overline{B_n}) \\ &= P\left(\bigcup_{n=1}^{+\infty} \overline{B_n}\right) = P\left(\overline{\bigcap_{n=1}^{+\infty} B_n}\right) \\ &= 1 - P\left(\bigcap_{n=1}^{+\infty} B_n\right). \end{aligned}$$

由于 $\{B_n\}$ 单调不增，所以

$$\lim_{n \rightarrow +\infty} P(B_n) = P\left(\bigcap_{n=1}^{+\infty} B_n\right) = P\left(\lim_{n \rightarrow +\infty} B_n\right)$$

1.2 古典概型与几何概型

1.2.1 古典概型

古典概型的基本思想：

有限个样本点 所涉及的随机现象只有有限个样本点，譬如为 n 个，且这些事件是两两互不相容的；

等可能性 每个样本点发生的可能性相等

定义 1.8 (古典概型)

若事件 A 含有 k 个样本点，则事件 A 的概率为

$$P(A) = \frac{\text{事件 } A \text{ 所含样本点的个数}}{\Omega \text{ 中所有样本点的个数}} = \frac{k}{n}$$



笔记 古典概型的大部分问题都能形象化地用摸球模型来描述。以后我们经常研究摸球模型，意义即在于此。

例题 1.5 生日问题 n 个人的生日全不相同的概率 p_n 是多少？

解 把 n 个人看成是 n 个球，将一年 365 天看成是 $N = 365$ 个盒子，则“ n 个人的生日全不相同”就相当于“恰好有 n ($n \leq N$) 个盒子各有一球”，所以 n 个人的生日全不相同的概率为：

$$P_n = \frac{365!}{365^n(365-n)!} = \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right)$$

上式可用以下方法作近似计算：

1. 当 n 较小时，原式右边中各因子的第二项之间的乘积 $\frac{i}{365} \times \frac{j}{365}$ 都可以忽略，于是有近似公式

$$p_n \approx 1 - \frac{1+2+\cdots+(n-1)}{365} = 1 - \frac{n(n-1)}{730}$$

2. 当 n 较大时，因为对小的正数 x 有 $\ln(1-x) \approx -x$ ，所以由原式得

$$\ln p_n \approx \frac{1+2+\cdots+(n-1)}{365} = -\frac{n(n-1)}{730}$$

1.2.2 几何概型

几何概型的基本思想：

不可数样本点 如果一个随机现象的样本空间 Ω 充满某个区域，其测度（长度、面积或体积等）大小可用 S_Ω 表示；

测度等可能性 任意一点落在测度相同的子区域内是等可能的。

定义 1.9 (几何概型)

若将“在区域 Ω 中随机地取一点，而该点落在区域 g 中”记为事件 A_g ，则其概率定义为

$$\mathbb{P}(A_g) = \frac{S_g}{S_\Omega}$$



例题 1.6 会商问题 甲乙两人约定在下午 6 时到 7 时之间在某处会面，并约定先到者应等候另一个人 20 min，过时即可离去。求两人能会面的概率。

解 以 x 和 y 分别表示甲、乙两人到达约会地点的时间（以 min 为单位），在平面上建立 xOy 直角坐标系（见图 1.1）。因为甲、乙都是在 0 至 60 min 内等可能地到达，所以由等可能性知这是一个几何概率问题。 (x, y) 的所有可能

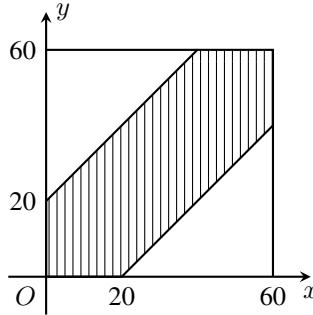


图 1.1：会面问题中的 Ω 与 A 。

取值是边长为 60 的正方形，其面积为 $S_\Omega = 60^2$ 。而事件 A = “两人能够会面”相当于：

$$|(|x - y|) \leq 20$$

即图中的阴影部分，其面积为 $S_A = 60^2 - 40^2$ ，由几何概型计算公式可知

$$P(A) = \frac{S_A}{S_\Omega} = \frac{60^2 - 40^2}{60^2} = 0.5556$$

例题 1.7 蒲丰投针问题 平面上画有间隔为 d ($d > 0$) 的等距平行线，向平面任意投掷一枚长为 l ($l < d$) 的针，求针与任一平行线相交的概率。

解 以 x 表示针的中点与最近一条平行线的距离，又以 φ 表示针与此直线间的交角，见图 1.2。易知样本空间 Ω 满足

$$0 \leq x \leq d/2, \quad 0 \leq \varphi \leq \pi$$

由这两式可以确定 $x - \varphi$ 平面上的一个矩形 Ω ，这就是样本空间，其面积为 $S_\Omega = d\pi/2$ 。这时为了针与平行线相交（记为事件 A ），其充要条件是

$$x \leq \frac{l}{2} \sin \varphi$$

由这个不等式表示的区域是图 1.3 中的阴影部分。

由于针是向平面任意投掷的，所以由等可能性知这是一个几何概率问题。由此得

$$P(A) = \frac{S_A}{S_\Omega} = \frac{\int_0^{\pi} \frac{l}{2} \sin \varphi d\varphi}{\frac{d\pi}{2}} = \frac{2l}{d\pi}.$$

如果 l, d 为已知，则以 π 的值代入上式即可计算得 $P(A)$ 之值。反之，如果已知 $P(A)$ 的值，则也可以利用上式去求 π ，而关于 $P(A)$ 的值，可用从试验中获得的频率去近似它：即投针 N 次，其中针与平行线相交 n 次，则频率 n/N 可作为 $P(A)$ 的估计值，于是由

$$\frac{n}{N} \approx P(A) = \frac{2l}{d\pi}$$

可得

$$\pi \approx \frac{2lN}{dn}$$

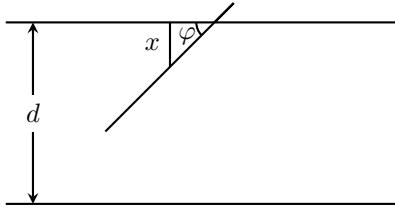
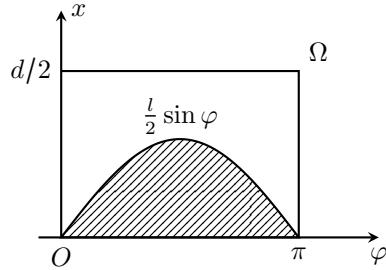


图 1.2: 蒲丰投针问题

图 1.3: 蒲丰投针问题中的 Ω 和 A

例题 1.8 在长度为 a 的线段内任取两点将其分为三段，求它们可以构成一个三角形的概率。

解 由于是将线段任意分成三段，所以由等可能性知这是一个几何概率问题。分别用 x, y 和 $a - x - y$ 表示线段被分成的三段长度，见图1.4。则显然应该有

$$0 < x < a; \quad 0 < y < a; \quad 0 < a - (x + y) < a$$

第三个式子等价于： $0 < x + y < a$ 。所以样本空间为（见图1.5）

$$\Omega = \{(x, y) : 0 < x < a, 0 < y < a, 0 < x + y < a\}$$

Ω 的面积为

$$S_\Omega = \frac{a^2}{2}$$

又根据构成三角形的条件：三角形中任意两边之和大于第三边，得事件 A 所含样本点 (x, y) 必须满足：

$$\begin{aligned} 0 &< a - (x + y) < x + y, \\ 0 &< x < y + (a - x - y), \\ 0 &< y < x + (a - x - y). \end{aligned}$$

整理得

$$\frac{a}{2} < x + y < a; \quad 0 < x < \frac{a}{2}; \quad 0 < y < \frac{a}{2}$$

所以事件 A 可用图 1.6 中的阴影部分表示。事件 A 的面积为

$$S_A = \frac{a^2}{8}.$$

由此得

$$P(A) = \frac{1}{4}.$$

1.2.3 Bertrand 奇论

在一圆内任取一条弦，问其长度超过该圆内接等边三角形的边长的概率是多少？这是一个几何概率问题，它有三种解法，具体如下：

1. 由于对称性，可只考察某指定方向的弦。作一条直径垂直于这个方向。显然，只有交直径于 $1/4$ 与 $3/4$ 之

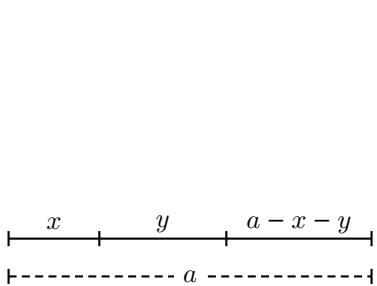
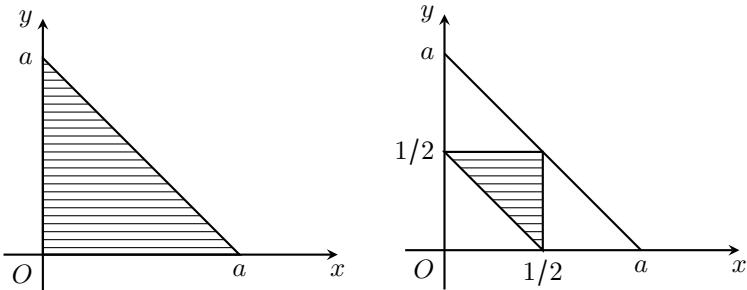
图 1.4: 长度为 a 的线段分成三段.

图 1.6: 构成三角形的条件

间的弦才能超过正三角形的边长（见图），如此，所求概率为 $1/2$ 。

2. 由于对称性，可让弦的一端点固定，让另一端点在圆周上作随机移动。若在固定端点作一切线，则与此切线夹角在 60° 与 120° 之间的弦才能超过正三角形的边长（见图），如此，所求概率为 $1/3$ 。
3. 圆内弦的位置被其中点唯一确定。在圆内作一同心圆，其半径仅为大圆半径的一半，则大圆内弦的中点落在小圆内，此弦长才能超过正三角形的边长（见图），如此，所求概率为 $1/4$ 。

同一问题有三种不同答案，究其原因在圆内“取弦”时规定尚不够具体，不同的“等可能性假定”导致了不同的样本空间。具体如下：其中“均匀分布”应理解为“等可能取点”。

1. 解法一中假定弦的中点在直径上均匀分布，直径上的点组成样本空间
2. 解法二中假定弦的另一活动端点在圆周上均匀分布，圆周上的点组成样本空间
3. 解法三中假定弦的中点在大圆内均匀分布，大圆内的点组成样本空间

可见，上述三个答案是针对三个不同样本空间引起的，它们都是正确的，贝特朗奇论引起人们注意，在定义概率时要事先明确指出样本空间是什么。

1.3 条件概率

定义 1.10 (条件概率)

设 $A, B \in \mathcal{F}$, 且 $P(B) > 0$ 称

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

为基于于 B 的条件概率 (probability conditional on B)。

命题 1.3

条件概率仍是一个概率测度，即若设 $P(B) > 0$ ，则

- $P(A|B) \geq 0, \quad A \in \mathcal{F}$,
- $P(\Omega|B) = 1$,
- 若 \mathcal{F} 中的 $A_1, A_2, \dots, A_n, \dots$ 互不相容，则

$$P\left(\bigcup_{n=1}^{+\infty} A_n | B\right) = \sum_{n=1}^{+\infty} P(A_n | B)$$

证明 从条件概率的定义很容易得出前两点。因为 $A_1, A_2, \dots, A_n, \dots$ 互不相容，所以 $A_1 B, A_2 B, \dots, A_n B, \dots$ 也

互不相容，故

$$\begin{aligned} P\left(\bigcup_{n=1}^{+\infty} A_n | B\right) &= \frac{P\left(\left(\bigcup_{n=1}^{+\infty} A_n\right)B\right)}{P(B)} = \frac{P\left(\bigcup_{n=1}^{+\infty}(A_n B)\right)}{P(B)} \\ &= \sum_{n=1}^{+\infty} \frac{P(A_n B)}{P(B)} = \sum_{n=1}^{+\infty} P(A_n | B). \end{aligned}$$

定理 1.3 (乘法法则)

设 $A, B \in \mathcal{F}$, 且 $P(B) > 0$, 则

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

泛化后有：

$$P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 A_2) \cdots P(A_n | A_1 A_2 \cdots A_{n-1})$$



例题 1.9 罐子模型 设罐中有 b 个黑球、 r 个红球，每次随机取出一个球，取出后将原球放回，还加进 c 个同色球和 d 个异色球。记 B_i 为“第 i 次取出的是黑球”， R_j 为“第 j 次取出的是红球”。若连续从罐中取出三个球，其中有两个红球、一个黑球。则由乘法公式我们可得

$$\begin{aligned} P(B_1 R_2 R_3) &= P(B_1)P(R_2 | B_1)P(R_3 | B_1 R_2) \\ &= \frac{b}{b+r} \frac{r+d}{b+r+c+d} \frac{r+d+c}{b+r+2c+2d}, \\ P(R_1 B_2 R_3) &= P(R_1)P(B_2 | R_1)P(R_3 | R_1 B_2) \\ &= \frac{r}{b+r} \frac{b+d}{b+r+c+d} \frac{r+d+c}{b+r+2c+2d}, \\ P(R_1 R_2 B_3) &= P(R_1)P(R_2 | R_1)P(B_3 | R_1 R_2) \\ &= \frac{r}{b+r} \frac{r+c}{b+r+c+d} \frac{b+2d}{b+r+2c+2d}. \end{aligned}$$

以上概率与黑球在第几次被抽取有关。

罐子模型也称为波利亚 (Polya) 模型，这个模型可以有各种变化，具体见下：

- 当 $c = -1, d = 0$ 时，即为**不返回抽样**。此时前次抽取结果会影响后次抽取结果。但只要抽取的黑球与红球个数确定，则概率不依赖其抽出球的次序，都是一样的。此例中有

$$\begin{aligned} P(B_1 R_2 R_3) &= P(R_1 B_2 R_3) = P(R_1 R_2 B_3) \\ &= \frac{br(r-1)}{(b+r)(b+r-1)(b+r-2)}. \end{aligned}$$

- 当 $c = 0, d = 0$ 时，即为**返回抽样**。此时前次抽取结果不会影响后次抽取结果。故上述三个概率相等，且都等于

$$P(B_1 R_2 R_3) = P(R_1 B_2 R_3) = P(R_1 R_2 B_3) = \frac{br^2}{(b+r)^3}$$

- 当 $c > 0, d = 0$ 时，称为**传染病模型**。此时，每次取出球后会增加下一次取到同色球的概率，或换句话说，每次发现一个传染病患者，以后都会增加再传染的概率。与前面两个一样，以上三个概率都相等，且都等于

$$\begin{aligned} P(B_1 R_2 R_3) &= P(R_1 B_2 R_3) = P(R_1 R_2 B_3) \\ &= \frac{br(r+c)}{(b+r)(b+r+c)(b+r+2c)}. \end{aligned}$$

从以上可以看出：在罐子模型中只要 $d = 0$ ，则以上三个概率都相等。即只要抽取的黑球与红球个数确定，则概率不依赖其抽出球的次序，都是一样的。但当 $d > 0$ 时，就不同了。

- 当 $c = 0, d > 0$ 时，称为**安全模型**。此模型可解释为：每当事故发生 (红球被取出)，安全工作就抓紧一

些，下次再发生事故的概率就会减少；而当事故没有发生时（黑球被取出），安全工作就放松一些，下次再发生事故的概率就会增大。在这种场合，上述三个概率分别为

$$\begin{aligned} P(B_1 R_2 R_3) &= \frac{b}{b+r} \frac{r+d}{b+r+d} \frac{r+d}{b+r+2d}, \\ P(R_1 B_2 R_3) &= \frac{r}{b+r} \frac{b+d}{b+r+d} \frac{r+d}{b+r+2d}, \\ P(R_1 R_2 B_3) &= \frac{r}{b+r} \frac{r}{b+r+d} \frac{b+2d}{b+r+2d}. \end{aligned}$$

定理 1.4 (全概率公式)

设可列个事件 $B_1, B_2, \dots, B_n, \dots$ 互不相容，且包含事件 A 。即 $A \subseteq \bigcup_{i=1}^{\infty} B_i$, $B_i B_j = \emptyset, \forall i \neq j$, 则：

$$P(A) = \sum_{i=1}^{\infty} P(B_i)P(A|B_i), \quad \forall A \in \mathcal{F}$$



笔记 $P(A|B_i)$ 可视为事件 A 在 B_i 上的分量， $P(B_i)$ 则为其权重，加权平均后得到 $P(A)$

证明 因为 $A \subseteq \bigcup_{i=1}^{\infty} B_i$, 所以

$$A = A(\bigcup_{n=1}^{\infty} B_i) = \bigcup_{i=1}^{\infty} (AB_i)$$

且 $AB_1, AB_2, \dots, AB_n, \dots$ 互不相容，所以由可加性得

$$P(A) = P\left(\bigcup_{i=1}^{\infty} (AB_i)\right) = \sum_{i=1}^{\infty} P(AB_i)$$

再将 $P(AB_i) = P(B_i)P(A|B_i)$, 代入上式即得全概率公式。

推论 1.2

$$P(A) = P(B)P(A|B) + P(\bar{B})P(A|\bar{B})$$



例题 1.10 敏感性问题调查 学生阅读黄色书刊和观看黄色影像会严重影响学生身心健康发展。但这属个人隐私行为。现在要设计一个调查方案，从调查数据中估计出学生中阅读黄色书刊和观看黄色影像的比率 p .

敏感性问题的调查方案，关键要使被调查者愿意作出真实回答又能保守个人秘密。经过多年研究和实践，一些心理学家和统计学家设计了一种调查方案，在这个方案中被调查者只需回答以下两个问题中的一个问题，而且只需回答“是”或“否”。

- 你的生日是否在 7 月 1 日之前？
- 你是否看过黄色书刊或影像？

这个调查方来看似简单，但为了消除被调查者的顾虑，使被调查者确信他这次调查不会泄露个人秘密，在操作上有以下关键点：

- 被调查者在没有旁人的情况下，独自一人回答问题。
- 被调查者从只有白球和红球的罐子中随机抽一只球，看过颜色后放回。若抽到白球，则回答问题 A；若抽到红球，则回答问题 B。

如此的调查方法，主要在于旁人无法知道被调查回答的是问题 A 还是问题 B，由此可以极大地消除被调查者的顾虑。现在的问题是如何分析调查的结果。很显然，我们只对问题 B 是感兴趣。

首先假设有 n 张答卷 (n 较大)，其中有 k 张回答“是”。有两个信息我们是预先知道的：

- 在参加人数较多的情况下，任选一人其生日在 7 月 1 日之前的概率为 0.5.
- 罐中红球的比率 π 。

由全概率公式得

$$P(\text{是}) = P(\text{白球})P(\text{是}|\text{白球}) + P(\text{红球})P(\text{是}|\text{红球})$$

将 $P(\text{红球}) = \pi, P(\text{白球}) = 1 - \pi, P(\text{是}|\text{白球}) = 0.5, P(\text{是}|\text{红球}) = p$ 代入上式右边，而上式左边用频率 k/n 代

替概率 $P(\text{是})$, 得

$$\frac{k}{n} = 0.5(1 - \pi) + p \cdot \pi$$

由此得

$$p = \frac{k/n - 0.5(1 - \pi)}{\pi}$$

定理 1.5 (Bayes 定理)

设可列个事件 $B_1, B_2, \dots, B_n, \dots$ 互不相容, 且包含事件 A 。即 $A \subseteq \bigcup_{i=1}^{\infty} B_i$, $B_i B_j = \emptyset, \forall i \neq j$, 则

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^n P(B_j)P(A|B_j)}$$



例题 1.11 某地区居民的肝癌发病率为0.0004, 现用甲胎蛋白法进行普查。医学研究表明, 化验结果是存有错误的。已知患有肝癌的人其化验结果 99 % 呈阳性(有病), 而没患肝癌的人其化验结果 99.9 % 呈阴性(无病)。现某人的检查结果呈阳性, 求其真患肝癌的概率。

解 记 B 为事件“被检查者患有肝癌”, A 为事件“检查结果呈阳性”。由题设知

$$\begin{aligned} P(B) &= 0.0004, \quad P(\bar{B}) = 0.996, \\ P(A|B) &= 0.99, \quad P(A|\bar{B}) = 0.001. \end{aligned}$$

由贝叶斯公式得

$$\begin{aligned} P(B|A) &= \frac{P(B)P(A|B)}{P(B)P(A|B) + P(\bar{B})P(A|\bar{B})} \\ &= \frac{0.0004 \times 0.99}{0.0004 \times 0.99 + 0.9996 \times 0.001} \\ &= 0.284. \end{aligned}$$

条件概率的三公式中: 乘法公式是求事件交的概率; 全概率公式是求一个复杂事件的概率; 而贝叶斯公式是求一个条件概率。在贝叶斯公式中, 称 $P(B_i)$ 为 B_i 的先验概率, 称 $P(B_i|A)$ 为 B_i 的后验概率。贝叶斯公式是专门用于计算后验概率的, 也就是通过 A 的发生这个新信息, 来对 B_i 的概率作出的修正。

1.4 独立性

定义 1.11 (事件的独立性)

如果 $A, B \in \mathcal{F}$ 满足

$$P(A \cap B) = P(A)P(B),$$

则称 A 与 B 独立 (independent), 记为 $A \perp\!\!\!\perp B$.

对于事件集 A_1, A_2, \dots, A_n , 若对于其中任意子集 $A_{i_1}, A_{i_2}, \dots, A_{i_n}$ 有:

$$P(A_{i_1} \cap \dots \cap A_{i_n}) = P(A_{i_1}) \cdots P(A_{i_n})$$

则称此事件集相互独立 (mutually independent)



笔记 当 $P(A) > 0$ 时, 我们有 $P(B|A) = P(B) \iff B \perp\!\!\!\perp A$, 由此可得到 B 独立于 A 的直观理解

性质 独立性是对称的, 即 $A \perp\!\!\!\perp B \iff B \perp\!\!\!\perp A$. 若两事件独立, 则其补集也独立.

$$\begin{array}{c} A \leftrightarrow B \\ \nwarrow \nearrow \\ A^c \leftrightarrow B^c \end{array}$$

定义 1.12 (事件域的独立性)

若 $\mathcal{G} \subset \mathcal{F}$ 与 $\mathcal{H} \subset \mathcal{F}$ 满足

$$A \perp\!\!\!\perp B, \quad \forall A \in \mathcal{G}, B \in \mathcal{H},$$

则称 \mathcal{G} 与 \mathcal{H} 独立, 记为 $\mathcal{G} \perp\!\!\!\perp \mathcal{H}$



测度论告诉我们一个重要结果: 如果 \mathcal{G} 对交集运算封闭, 那么成立 $\mathcal{G} \perp\!\!\!\perp \mathcal{H} \implies \sigma(\mathcal{G}) \perp\!\!\!\perp \mathcal{H}$

例题 1.12 有两名选手比赛射击, 轮流对同一目标进行射击, 甲命中目标的概率为 α , 乙命中目标的概率为 β . 甲先射, 谁先命中谁得胜. 问甲、乙两人获胜的概率各为多少?

解 记事件 A_i 为“第 i 次射击命中目标”, $i = 1, 2$, 因为甲先射, 所以事件“甲获胜”可以表示为

$$A_1 \cup \bar{A}_1 \bar{A}_2 A_3 \cup \bar{A}_1 \bar{A}_2 \bar{A}_3 \bar{A}_4 A_5 \cup \dots,$$

又因为各次射击是独立的, 所以得

$$\begin{aligned} P(\text{甲获胜}) &= \alpha + (1 - \alpha)(1 - \beta)\alpha + (1 - \alpha)^2(1 - \beta)^2\alpha + \dots \\ &= \alpha \sum_{i=0}^{+\infty} (1 - \alpha)^i (1 - \beta)^i \\ &= \frac{\alpha}{1 - (1 - \alpha)(1 - \beta)}. \end{aligned}$$

同理可得

$$\begin{aligned} P(\text{乙获胜}) &= P(\bar{A}_1 A_2 \cup \bar{A}_1 \bar{A}_2 \bar{A}_3 A_4 \cup \dots) \\ &= (1 - \alpha)\beta + (1 - \alpha)(1 - \beta)(1 - \alpha)\beta + \dots \\ &= \beta(1 - \alpha) \sum_{i=0}^{+\infty} (1 - \alpha)^i (1 - \beta)^i \\ &= \frac{\beta(1 - \alpha)}{1 - (1 - \alpha)(1 - \beta)}. \end{aligned}$$

此题在等比级数求和时, 应该有条件: 公比 $|(1 - \alpha)(1 - \beta)| < 1$. 这一点不难从题目的实际意义中得到. 因为对本题而言, α, β 取值为零或 1 均是无意义的.

例题 1.13 系统由多个元件组成, 且所有元件都独立地工作. 设每个元件正常工作的概率都为 $p = 0.9$, 试求以下系统正常工作的概率.

(1) 串联系统 S_1 :

(2) 并联系统 S_2 :

(3) 5 个元件组成的桥式系统 S_3 :

解 设 S_i = “第 i 个系统正常工作”, A_i = “第 i 个元件正常工作”.

(1) 对串联系统而言, “系统正常工作”相当于“所有元件正常工作”, 即 $S_1 = A_1 A_2$, 所以

$$P(S_1) = P(A_1 A_2) = P(A_1)P(A_2) = p^2 = 0.81.$$

这也可看出: 两个正常工作概率为 0.9 的元件组成的串联系统, 其系统正常工作的概率下降为 0.81.

(1) 对并联系统而言, “系统正常工作”相当于“至少一个元件正常工作”, 即 $S_2 = A_1 \cup A_2$, 所以

$$\begin{aligned} P(S_2)P(A_1 \cup A_2) &= P(A_1) + P(A_2) - P(A_1 A_2) \\ &= p + p - p^2 = 0.99. \end{aligned}$$

或

$$\begin{aligned} P(S_2) &= 1 - P(\bar{S}_2) = 1 - P(A_1 \bar{\cup} A_2) = 1 - P(\bar{A}_1 \cap \bar{A}_2) \\ &= 1 - P(\bar{A}_1)P(\bar{A}_2) = 1 - (1-p)^2 = 0.99. \end{aligned}$$

这也可看出：两个正常工作概率为 0.9 的元件组成的并联系统，其系统正常工作的概率提高至 0.99.

(1) 在桥式系统中，第 3 个元件是关键，我们先用全概率公式得

$$P(S_3) = P(A_3)P(S_3|A_3) + P(\bar{A}_3)P(S_3|\bar{A}_3).$$

因为在“第 3 个元件正常工作”的条件下，系统成为先并后串系统（见图 1.7）。所以

$$\begin{aligned} P(S_3|A_3) &= P((A_1 \cup A_4)(A_2 \cup A_5)) = P(A_1 \cup A_4)P(A_2 \cup A_5) \\ &= [1 - (1-p)^2]^2 = 0.9801. \end{aligned}$$

又因为在“第 3 个元件不正常工作”的条件下，系统成为先串后并系统（见图 1.8）。

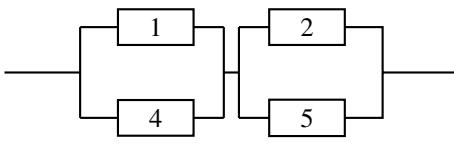


图 1.7: 先并后串系统

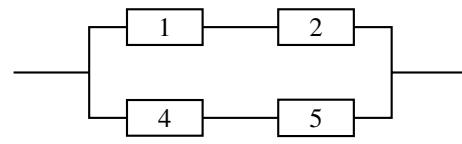


图 1.8: 先串后并系统

所以

$$P(S_3|\bar{A}_3) = P(A_1A_2 \cup A_4A_5) = 1 - (1-p^2)^2 = 0.9639.$$

最后我们得

$$\begin{aligned} P(S_3) &= p[1 - (1-p)^2]^2 + (1-p)[1 - (1-p^2)^2] \\ &= 0.9 \times 0.9801 + 0.1 \times 0.9639 = 0.9785. \end{aligned}$$

错题记录

1. (茆 1.2.4) 从一副 52 张的扑克牌中任取 4 张，求下列事件的概率：
 - (a). 全是黑桃；
 - (b). 同花；
 - (c). 没有两张同一花色；
 - (d). 同色。
2. (茆 1.2.5) 考虑一元二次方程 $x^2 + Bx + C = 0$ ，其中 B, C 分别是将一颗骰子接连掷两次先后出现的点数，求该方程有实根的概率和有重根的概率
3. (茆 1.2.10) 从 n 个数 $1, 2, \dots, n$ 中任取 2 个，问其中一个小于 k ($1 < k < n$)，另一个大于 k 的概率是多少？
4. (茆 1.2.13) 把 10 本书任意地放在书架上，求其中指定的三本书放在一起的概率。
5. (茆 1.2.19) n 个男孩， m 个女孩 ($m \leq n+1$) 随机地排成一排，试求任意两个女孩都不相邻的概率。
6. (茆 1.2.20) 将 3 个球随机地放入 4 个杯子中去，求杯子中球的最大个数分别为 1, 2, 3 的概率各为多少？
7. (茆 1.2.22) 将 n 个完全相同的球 (这时也称球是不可辨的) 随机地放入 N 个盒子中，试求恰好有 m 个空盒的概率
8. (茆 1.2.23) 在区间 $(0, 1)$ 中随机地取两个数，求事件“两数之和小于 $7/5$ ”的概率。
9. (茆 1.3.4) 从 $0, 1, 2, \dots, 9$ 十个数字中任意选出三个不同的数字，试求事件“三个数字中含 0 但不含 5”的概率
10. (茆 1.3.8) 从数字 $1, 2, \dots, 9$ 中可重复地任取 n 次，求 n 次所取数字的乘积能被 10 整除的概率。
11. (茆 1.3.23) 证明： $|P(AB) - P(A)P(B)| \leq \frac{1}{4}$