# 概率论与数理统计

授课教师：唐宏岩

# 前言

本讲义基于清华大学数学系唐宏岩老师于 2023-2024 学年秋季学期开设的《概率论与数理统计》课程，用于辅助同学们课后复习，助教尽量做到每周课后两天内更新。

提醒大家，由于时间与能力所限，本讲义可能不会出现大段的文字论述（但会包含重要的定义、定理与公式等）。但是，对许多基本概念的深入理解是非常有必要的，同学们可以在浏览时检查自己是否能够回忆起课上的内容，对掌握不够扎实的地方，鼓励大家查阅参考书或在微信群提问以解决问题。

由于此为课程组第一年尝试整理讲义，诸如格式编排、内容完整性方面可能存在许多不足，欢迎大家联系我提出宝贵的意见与建议。

曹子尧

2023 年 9 月

# 目录

# 第一部分

# 初等概率论

# 第一章　事件的概率

## 1.1　概率的发展史

赌博中的 de Méré's Problem：连续掷一个均匀六面骰 4 次，获得至少一次“6”的概率为 $1-(\frac{5}{6})^4 \approx 0.5177$；而连续掷两个均匀六面骰 24 次，获得至少一次“对 6”的概率为 $1-(35/36)^{24} \approx 0.4914$。

Pascal 和 Fermat 的通信中使用初等数学的方法，首创了概率论相当多的数学理论，虽然当时没有总结成通用的定理。

Laplace 创立了采用分析方法的分析概率论。

Kolmogorov 利用测度论方法发展了现代概率理论。

## 1.2　随机试验与事件

**定义 1.1.** 概率论中的随机试验指的是符合下面两个特点的试验：

1. 不能预先确知结果
2. 可以预测所有可能的结果

**定义 1.2.** 样本空间是指一个试验的所有可能结果的集合，常用 $\Omega$ 表示。

**定义 1.3.** 事件是样本空间的一个良定义的子集。

一次随机试验中，一个事件可能发生或不发生。

下面是一些常见的事件：

1. 全事件 $\Omega$（必然事件）
2. 空事件 $\emptyset$（不可能事件）
3. 基本事件 $\{a\}$，其中 $a \in \Omega$，即仅包含单一试验结果的事件

## 1.3 事件的运算

由于事件是集合，因此事件之间可以进行集合之间的运算，如：

1. 余 $A^c = \Omega \setminus A$
2. 和 $A + B = A \cup B = (A^c \cap B^c)^c$
3. 差 $A - B = A \setminus B$
4. 积 $AB = A \cap B = (A^c \cup B^c)^c$

集合的 De Morgan's laws 也适用于事件：$(\bigcup_n A_n)^c = \bigcap_n A_n^c$。

事件的运算像集合的运算一样，可以用 Venn 图来表示。

## 1.4 概率的几种解释

对于概率这一数学概念，人们形成了几种从不同角度出发的解释：

1. 古典解释：基于等可能性的解释
2. 频率解释：基于大量重复试验的解释（频率学派采用的解释）
3. 主观解释：概率是一种对确信程度的度量（Bayes 学派采用的解释）

## 1.5 概率的公理化定义

我们用 $2^\Omega$ 表示 $\Omega$ 的幂集，即 $\Omega$ 的所有子集组成的集合。

**定义 1.4.** 事件集类 $\mathscr{F} \subset 2^\Omega$ 必须满足所谓 $\sigma$-代数的性质：

1. $\Omega \in \mathscr{F}$
2. $A \in \mathscr{F} \Rightarrow A^c \in \mathscr{F}$（对补运算的封闭性）
3. $A_i \in \mathscr{F}, \forall i \in \mathbb{N}^+ \Rightarrow \bigcup_{i=1}^\infty A_i \in \mathscr{F}$（对可列并的封闭性）

**例 1.1.** $\Omega = \{a, b, c, d\}$，以下是一些合法的事件集类：

1. $\mathscr{F}_1 = 2^\Omega$
2. $\mathscr{F}_2 = \{\Omega, \emptyset\}$
3. $\mathscr{F}_3 = \{\Omega, \emptyset, \{a, b\}, \{c, d\}\}$

**定义 1.5.** (Kolmogorov) 概率函数 $P : \mathscr{F} \to \mathbb{R}$ 是满足以下三条公理的映射：

1. $P(A) \geq 0, \forall A \in \mathscr{F}$
2. $P(\Omega) = 1$
3. $A_i \in \mathscr{F}, \forall i \in \mathbb{N}^+, A_iA_j = \emptyset, \forall i \neq j \Rightarrow P(\sum_{i=1}^\infty A_i) = \sum_{i=1}^\infty P(A_i)$（加法公理/可列可加性）

我们称 $(\Omega, \mathscr{F}, P)$ 是一个概率空间。

**命题 1.1.** 关于概率空间，有如下性质：

1. $P(A) \leq 1,\ \forall A \in \mathscr{F}$
2. $P(\emptyset) = 0$
3. $P(A) + P(A^c) = 1$
4. $A_i \in \mathscr{F}, \forall i \in \{1, 2, \cdots, n\}$，$A_iA_j = \emptyset, \forall i \neq j \Rightarrow P(\sum_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$（有限可加性）
5. $A \subset B \Rightarrow P(A) \leq P(B)$（我们称事件 $A$ 蕴含事件 $B$）
6. $P(A_1 + \cdots + A_n) = \sum_{i=1}^n P(A_i) - \sum_{i_1<i_2} P(A_{i_1}A_{i_2}) + \cdots + (-1)^{r+1} \sum_{i_1<i_2<\cdots<i_r} P(A_{i_1}A_{i_2}\cdots A_{i_r}) + \cdots + (-1)^{n+1}P(A_1\cdots A_n)$（容斥公式）

   特别地，$P(A+B) = P(A) + P(B) - P(AB)$。

**例 1.2.** （配对问题）

有 $n$ 个人，每人有一顶帽子。现将所有帽子放到一起，再随机分配给每人一顶，考虑无人拿到自己的帽子的概率。

为此，设事件 $A_i$ 为“第 $i$ 个人拿到自己的帽子”，则 $P(A_i) = 1/n$。

利用容斥公式，至少一人拿到自己帽子的概率为 $P(A_1 + \cdots + A_n) = \sum_{i=1}^n P(A_i) - \sum_{i_1<i_2} P(A_{i_1}A_{i_2}) + \cdots + (-1)^{r+1} \sum_{i_1<i_2<\cdots<i_r} P(A_{i_1}A_{i_2}\cdots A_{i_r}) + \cdots + (-1)^{n+1}P(A_1\cdots A_n)$，

其中 $\sum_{i_1<i_2<\cdots<i_r} P(A_{i_1}A_{i_2}\cdots A_{i_r}) = \frac{(n-r)!}{n!}\binom{n}{r} = \frac{1}{r!}$，即 $P(A_1 + \cdots + A_n) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + (-1)^{r+1}\frac{1}{r!} + \cdots + (-1)^{n+1}\frac{1}{n!}$。

所求概率 $P_n = 1 - P(A_1 + \cdots + A_n) = 1 - (1 - \frac{1}{2!} + \cdots + (-1)^{n+1}\frac{1}{n!}) \to e^{-1} (n \to \infty)$。

思考：恰有 $k$ 个人拿到自己的帽子的概率？

## 1.6　条件概率

**定义 1.6.** 若 $P(B) > 0$，定义条件概率 $P(A|B) = \frac{P(AB)}{P(B)}$。

通常，我们计算条件概率的方法有两种：

1. 在缩小（受限）的样本空间（要求事件 $B$ 发生）上，考虑事件 $A$ 发生的概率
2. 根据定义计算

一种常用的形式是 $P(AB) = P(A|B)P(B) = P(B|A)P(A)$，这可以视作是求解两个事件的积的概率的方法（乘法法则）。

**例 1.3.** 掷一个均匀六面骰，$\Omega = \{1, 2, 3, 4, 5, 6\}, A = \{2, 3, 4, 5\}, B = \{1, 3, 5\}$，则 $P(A) = 4/6, P(B) = 3/6, P(AB) = 2/6, P(A|B) = \frac{P(AB)}{P(B)} = 2/3$。

**例 1.4.** 袋子中有 8 个红球和 4 个白球，无放回地取出两个球，利用组合数可知，两个都是红球的概率为 $\frac{\binom{8}{2}}{\binom{12}{2}}$。
用条件概率可以简化计算：$P(R_1R_2) = P(R_1)P(R_2|R_1) = \frac{8}{12} \times \frac{7}{11}$。

更一般地，我们有 $P(A_1A_2\cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})$，常用于序贯发生的一系列事件的积的概率求解。

**例 1.5.** 回忆上一节的“配对问题”。我们有 $P(A_{i_1}A_{i_2}\cdots A_{i_r}) = P(A_{i_1})P(A_{i_2}|A_{i_1})\cdots P(A_{i_r}|A_{i_1}\cdots A_{i_{r-1}}) = \frac{1}{n} \times \frac{1}{n-1} \times \cdots \times \frac{1}{n-(r-1)} = \frac{(n-r)!}{n!}$。

**命题 1.2.** 对于给定的事件 $B$，$P(\cdot|B) : \mathscr{F} \to \mathbb{R}$ 是概率函数，即 $(\Omega, \mathscr{F}, P(\cdot|B))$ 仍是概率空间。

对于上述命题的证明，只需验证 $P(\cdot|B)$ 满足概率的三条公理即可。

这提示我们，条件概率也是一种概率，如果我们将 $P(A)$ 称为观察到事件 $B$ 之前 $A$ 的“先验概率”，则 $P(A|B)$ 就是相应的“后验概率”。

一个常见的迷思是: 观测到事件 $A$ 已经发生后, 是否可以说事件 $A$ 发生的概率 $P(A) = 1$? 学过条件概率之后，我们知道答案是否定的，实际上是后验概率 $P(A|A) = 1$。

## 1.7　事件的独立性

**定义 1.7.** 若 $P(AB) = P(A)P(B)$，则称事件 $A, B$ 相互独立。

如果 $P(B) > 0$，我们注意到 $A, B$ 独立等价于 $P(A|B) = P(A)$。

**命题 1.3.** 若 $A, B$ 独立，则 $A^c, B$ 独立。

**定义 1.8.** 若 $P(ABC) = P(A)P(B)P(C)$，且 $A, B, C$ 两两独立，则称事件 $A, B, C$ 独立。

注意，仅有 $A, B, C$ 两两独立，不能推出三者独立。

**定义 1.9.** 若对于事件列 $\{A_i\}_{i=1}^{\infty}$，任意取有限个事件 $A_{i_1}, A_{i_2}, \cdots, A_{i_r}$，都有 $P(A_{i_1}A_{i_2}\cdots A_{i_r}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_r})$，则称 $\{A_i\}_{i=1}^{\infty}$ 相互独立。

**例 1.6.** 每周开奖的彩票，各次中奖率均为 $10^{-5}$ 且独立，问连续十年（520 周）不中奖的概率? 令事件 $A_i$ 为第 $i$ 周不中奖，则 $P(A_i) = 1 - 10^{-5}$，故 $P(A_1 \cdots A_{520}) = (1 - 10^{-5})^{520} \approx 0.9948$。

**定义 1.10.** 若事件 $A, B, E$ 满足 $P(AB|E) = P(A|E)P(B|E)$，则我们称 $A, B$ 关于 $E$ 条件独立。

注意，条件独立性和独立性之间没有蕴含关系。

## 1.8 Bayes 公式

**定理 1.1.** （全概率公式）

设 $\{B_i\}$ 是 $\Omega$ 的一个分割，即

1. $\sum_i B_i = \Omega$
2. $B_i B_j = \emptyset, \forall i \neq j$
3. $P(B_i) > 0, \forall i$

则 $P(A) = P(\sum_i (AB_i)) = \sum_i P(AB_i) = \sum_i P(A|B_i)P(B_i)$。

注：$\{B_i\}$ 可以是有限集合，或可数无穷集合。

**例 1.7.** 对于调查问卷中的敏感问题（如“你是否有过某病史”），被调查者可能会有所顾虑而做出虚假的回答。为保护被调查者的隐私，同时取得其信任，考虑引入一个“保护性问题”，即不具有敏感性的问题（如“你是否会游泳”），并让被调查者以抛硬币的方式，随机抽取一个问题回答。这样，抽到敏感问题的、确有过该病史的被调查者在回答“是”时也无须有担心病史暴露之虞。

设人群中，敏感问题答案为“是”的比例为 $p$（未知），保护性问题答案为“是”的比例为 $q$（假设已知），则若收集到 $n$ 个被调查者的结果，其中 $k$ 个为“是”，我们便有 $\frac{1}{2}p + \frac{1}{2}q \approx \frac{k}{n}$，可以据此得到 $p$ 的估计。

**定理 1.2.** （Bayes 公式 / Bayes 准则）

设 $\{B_i\}$ 是 $\Omega$ 的一个分割，则 $P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_j P(B_j)P(A|B_j)}$。

**例 1.8.** （假阳性悖论）

对于一种流行病，$A$ 表示一个人检查呈阳性，$B$ 表示此人确实患病。设 $P(B) = 10^{-4}, P(A|B) = 0.99, P(A|B^c) = 10^{-3}$，则一个检查呈阳性的人真的患病的概率仅为 $P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B)+P(A|B^c)P(B^c)} \approx 9\%$。

如果再次检测仍呈阳性，且两次检测效率不变，结果彼此独立，则此人真的患病的概率为 $P(B|A_1A_2) = \frac{P(A_1A_2|B)P(B)}{P(A_1A_2|B)P(B)+P(A_1A_2|B^c)P(B^c)} = \frac{P(A_1|B)P(A_2|B)P(B)}{P(A_1|B)P(A_2|B)P(B)+P(A_1|B^c)P(A_2|B^c)P(B^c)} \approx 99\%$。