

概率论基础

Probability Theory

目录

- 第一章 事件与概率
- 第二章 条件概率与统计独立性
- 第三章 随机变量与分布函数
- 第四章 数字特征与特征函数
- 第五章 极限定理

概率论简史

- **研究和揭示随机现象数量规律的学科.**
- 亦称赌博法, 机遇论, 猜测艺术等, 它的思想可追溯自公元前220年以前的中国的一些文献. 不过真正的历史却只有三百来年而已.
- 如今, 但凡要进行信息处理、决策制定、实验设计等等, 只要涉及数据, 必用概率统计的模型和方法. 例如, 在经济、管理、工程、技术、物理、化学、生物、环境、天文、地理、卫生、教育、语言、国防等领域都有非常重要的应用.

概率论简史

- **萌芽时期: 1653年之前**
 - 内容: 赌博和占卜中的一些问题;
 - 工具: 计数;

概率论简史

- **诞生: 1654年7月29日**
 - 这一天, 法国的职业赌徒De Mere向Pascal提出了“**分赌注问题**”: 甲、乙两赌徒下了赌注后对赌, 其技巧相当, 事先约定谁先赢得s局便算赢家而赢得所有赌注. 但在一人赢m局, 另一人赢n局时因故中止了赌局 ($m, n < s$), 那么赌注应该如何分配才公平?
 - 为了解决这一难题, Pascal与Fermat通过书信进行讨论, 深入细致地研究赌博中的数学问题, 从而导致概率论的诞生!

概率论简史

- **古典概率时期: 1654-1811年**
- 工具: 排列组合、代数分析方法;
- 内容: 离散型随机变量;
- 特征: 直观具体, 逻辑基础不严格;
- 主要工作:
 - Pascal与Fermat的7封通信, 1654年7-10月;
 - Huygens, 《论赌博中的计算》, 1657年;
 - Bernoulli, 《猜度术》, 1713年;
 - de Moivre, 《机会学说》, 1718年;
 - Thomas Bayes, 逆概率思想;

概率论简史

● 分析概率时期：1812-1932年

- 工具：特征函数、微分方程、差分方程；
- 内容：连续型随机变量；
- 主要工作：
 - Laplace——《分析概率论》，1812年，实现了由组合技巧向分析方法的过渡；
 - Poisson——泊松分布，泊松定理，泊松大数定律；
 - 圣彼得堡数学学派：Chebyshev, Markov, Liapunov——对大数定律和中心极限定理的发展；

概率论简史

● 现代概率时期：1933年-至今

- 工具：集合论和测度论
- 标志：Kolmogorov 《概率论基础》
- 意义：借助20世纪初完成的Lebesgue测度和积分理论以及抽象测度和积分理论，建立了一套严密的概率公理体系，成为现代概率论的基础，使概率论成为严谨的数学分支。

概率论简史

- 蓬勃发展：自1933年以来，在公理化的基础上，现代概率论不仅在理论上取得了一系列突破，在应用上也取得了巨大的成就，其应用几乎遍及所有的科学领域。
- 理论研究：极限理论，独立增量过程，马氏过程，平稳过程和时间序列，鞅和随机微分方程等
- 应用领域：天气预报、地震预报、产品的抽样调查、经济最优决策、金融保险、通讯工程、服务系统、生物医学等。

第1章 事件与概率

- 1.1 随机现象与统计规律
- 1.2 样本空间与事件
- 1.3 古典概型
- 1.4 几何概型
- 1.5 概率空间

1.1 随机现象与统计规律性

1.1.1 随机现象

自然界和人类社会中存在着两类现象：

确定现象(必然现象)：在一定条件下必然发生或不发生的现象。

不确定现象(随机现象)：事先无法准确预知结果的现象。

例：

- ◆ 向上抛出的物体会掉落地上 —— 确定
- ◆ 明天天气状况 —— 不确定
- ◆ 彩票中奖情况 —— 不确定
- ◆ 投掷硬币出现正反面情况 —— 不确定

随机现象的特点：不能事先预知结果，即使条件不变，在重复试验中出现结果也不一定相同。

虽然随机现象在个别的试验或观察中它的结果具有不确定性，但是在大量的重复试验下它的结果又具有某种规律，称为**统计规律性**。

概率论是研究与揭示随机现象的统计规律性的一门数学学科。

数理统计学则应用概率论的方法，研究如何收集与分析处理数据、并作出科学的决策。

1.1.2 随机试验与随机事件

对随机现象的观察、记录、试验统称为**随机试验**。它具有以下特性：

1. 可以在相同条件下重复进行
2. 事先知道可能出现的结果
3. 进行试验前并不知道哪个结果会发生

随机试验或者观察的结果称**随机事件**，常用大写字母 A, B, C 等表示，它是试验中可能发生、也可能不发生的情况。

随机试验的例子：

- 例：抛一个均匀硬币三次，观察正面(H)、反面(T)出现的情况。
- 例：记录一天内一家商场的顾客数量。
- 例：从一批电子元件里任意抽取一只测试寿命。
- 例：从一副去掉大小王的扑克牌中任意抽出一张，观察它的花色和点数。
- 例：任意找出一名同学，测量身高、体重。

随机事件的例子：

- ◆明天天气晴朗
- ◆买彩票中了奖
- ◆投掷硬币出现正面

1.1.3 频率的稳定性

1. 频率的定义

在相同的条件下进行了 n 次重复试验，记 n_A 是随机事件 A 发生的次数(称为频数)，则定义随机事件 A 发生的频率为

$$F_n(A) = \frac{n_A}{n}。$$

频率描述了一个随机事件发生的频繁程度。

大量的随机试验表明：

- (1) 频率具有**随机波动性**，即对于同一个随机事件来说，在相同的试验次数下，得到的频率也不一定会相同。
- (2) 频率还具有**稳定性**，总是在某一个具体数值附近波动，随着试验次数的不断增加，频率的波动会越来越小，逐渐稳定在这个数值。

频率的稳定性表明随机现象也具有规律性，称为是统计规律(大量试验下体现出的规律)。

例：抛硬币出现的正面的频率

试验序号	$n=5$		$n=50$		$n=500$	
	n_H	$f_n(H)$	n_H	$f_n(H)$	n_H	$f_n(H)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.50	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1.0	25	0.50	253	0.506
5	1	0.2	24	0.48	251	0.502
6	2	0.4	21	0.42	246	0.492
7	4	0.8	18	0.36	244	0.488
8	2	0.4	24	0.48	258	0.516
9	3	0.6	27	0.54	262	0.524
10	3	0.6	31	0.62	247	0.494

例：历史上著名的投掷硬币试验。

表 1.1 历史上投掷硬币试验的记录

试验者	投掷次数 (n)	正面次数 (r_n)	正面频率 ($\frac{r_n}{n}$)
De Morgan	2 048	1 061	0.5181
Buffon	4 040	2 048	0.5069
Pearson K	12 000	6 019	0.5016
Pearson K	24 000	12 012	0.5005

例：高尔顿钉板试验

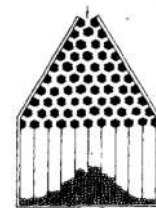


图 1 高尔顿板

2. 概率的描述性定义:

频率的稳定性说明: 随机事件发生的可能性大小是随机事件本身固有的、不随人们意志改变的一种客观属性, 因此可以对它进行度量。

随机事件A发生的可能性大小的度量, 称为A发生的**概率** (probability), 记作 $P(A)$ 。

概率 $\xrightarrow{\text{表现}}$ 频率

概率的频率定义

自然地, 可以采用一个随机事件的频率的稳定值去描述它在一次试验中发生的可能性大小, 即用频率的极限来作为概率的定义, 称为**概率的统计定义**。

统计概率的特性:

1. 直观, 易于理解, 生活中比比皆是;
2. 大量重复试验的局限性, 只能得到近似值;
3. 用现象定义本质, 未抓住概率本质。

3. 频率的性质

(1) (非负有界) $0 \leq F_n(A) \leq 1$;

(2) (规范性) $F_n(\Omega) = 1$;

(3) (有限可加)

如果 A_1, A_2, \dots, A_m 两两互不相容, 则:

$$\begin{aligned} F_n(A_1 + A_2 + \dots + A_m) \\ = F_n(A_1) + F_n(A_2) + \dots + F_n(A_m) \end{aligned}$$

1.2 样本空间与事件

1.2.1 样本空间

随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**(sample space), 用符号 Ω 来表示。

样本空间的元素, 即 E 的每一个可能的结果称为 E 的**样本点**(sample point), 用符号 ω 来表示。

E_1 : 抛一个均匀硬币三次, 观察正面(H), 反面(T)出现的情况

$\Omega_1 = \{ HHH, HHT, HTH, HTT, THH, THT, TTH, TTT \}$ 有限个点

E_2 : 记录一天内一家商场的顾客数量

$\Omega_2 = \{ 0, 1, 2, 3, \dots \}$ 可数个点

E_3 : 从一批电子元件里任意抽取一只测试寿命

$\Omega_3 = \{ t \mid t \geq 0 \}$ 不可数点构成的区间

E_4 : 从一副去掉大小王的扑克牌中任意抽出一张, 观察它的花色和点数

$\Omega_4 = \{ (x, y) \mid x \text{ 表示花色, 有 4 种可能; } y \text{ 是点数, } 1 \leq y \leq 13 \}$

E_5 : 任意找出一名同学, 测量身高、体重

$\Omega_5 = \{ (x, y) \mid a \leq x \leq b, c \leq y \leq d \}$

可见, 样本空间可以是有限或无限多个离散的点; 或者是有限 (或无限) 的区间; 甚至还可以是二维以及任意维数的集合。

Remark:

同一随机试验可构造出不同的样本空间, 如:
 E_1 : 抛均匀硬币三次, 以正面 H 出现的次数来构造样本空间 $\Omega_1 = \{0, 1, 2, 3\}$

1.2.2 随机事件(random event)

随机事件是样本点的集合, 即随机试验 E 的样本空间 Ω 的子集。一般用 A, B, \dots 表示。

显然, 样本点是不可分割的最基本随机事件。

称一个**随机事件发生**当且仅当其中的某一个样本点出现; 而一个样本点出现, 则所有包含这个样本点的随机事件都要发生。

◆ 两个特殊的随机事件:

样本空间 Ω (包含全部样本点)称为**必然事件**;

空集 Φ (不包含任何样本点)称为**不可能事件**。

例1.2.1 抛掷均匀硬币三次, 考虑随机事件 $A = \{\text{第一次出现的是正面H}\}$, 则
 $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$, $A = \{HHH, HHT, HTH, HTT\}$.

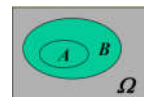
如果 A 发生, 说明 A 中某一个样本点出现了;

反之, 如果样本点 $\{HHT\}$ 出现, 则不仅表明随机事件 $A = \{\text{第一次是正面}\}$ 发生了, 同时随机事件 $B = \{\text{第二次是正面}\}$ 、 $C = \{\text{第三次是反面}\}$ 、 $D = \{\text{正面比反面多出现一次}\}$ 、 $E = \{\text{正面反面都出现}\}$ 、 $F = \{\text{正面比反面先出现}\}$ 等也都发生。

1.2.3 事件的关系与运算

1. 事件的包含、等价关系

如果 A 发生必然导致 B 的发生, 则称 A **包含于** B 中, 记为 $A \subset B$. 即 A 的每个样本点也都属于 B .



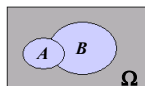
例. $A = \{HHH\}$, 三次都是正面, $B = \{H**\}$, 第一次是正面, 则 $A \subset B$.

特别地, 对任意 $A, \Phi \subset A \subset \Omega$.

如果 $A \subset B$ 与 $B \subset A$ 同时成立, 则称 A 与 B **等价**, 记为 $A = B$. 等价事件同时发生, 二者样本点完全相同。

2. 事件的并运算

- ✓ A 与 B 的**并事件**, 记为 $A \cup B$, 由属于 A 或 B 的所有样本点组成, 即
 $A \cup B = \{x | x \in A \text{ 或 } x \in B\}$
 $\Leftrightarrow A$ 与 B 至少有一发生。



例. $A = \{HHH\}$, $B = \{TTT\}$, 则
 $A \cup B = \{HHH, TTT\}$, 三次都是同一面。

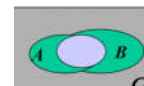
特别地, 对任意的随机事件 A ,

$$A \cup A = A, A \cup \Phi = A, A \cup \Omega = \Omega$$

当 A, B 不相容时, 称它们的并为**和**, 并记作 $A+B$.

3. 事件的交运算

- ✓ A 与 B 的**交事件**, 记为 $A \cap B$ 或 AB , 由属于 A 及 B 的样本点组成, 即
 $A \cap B = \{x | x \in A \text{ 且 } x \in B\}$
 $\Leftrightarrow A$ 与 B 同时发生。



例. $A = \{H**\}$, $B = \{\square H \square\}$, 则
 $AB = \{HH*\}$, 前两次都是正面。

特别地, 对任意的随机事件 A ,

$$A \cap A = A, A \cap \Phi = \Phi, A \cap \Omega = A.$$

事件的并与交运算可推广到可列个事件的情形：

$$\bigcup_{k=1}^{\infty} A_k = \lim_{n \rightarrow \infty} \bigcup_{k=1}^n A_k :$$

$A_1, A_2, \dots, A_n, \dots$ 中至少有一个发生

$$\bigcap_{k=1}^{\infty} A_k = \lim_{n \rightarrow \infty} \bigcap_{k=1}^n A_k :$$

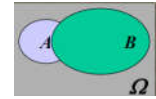
$A_1, A_2, \dots, A_n, \dots$ 同时发生

4. 事件的差运算

✓ A与B的**差事件**, 记为 $A-B$, 由属于A但不属于B的样本点组成, 即

$$A-B = \{x | x \in A \text{ 且 } x \notin B\}$$

$\Leftrightarrow A$ 发生而 B 不发生。



例. $A = \{HH*\}, B = \{\square\square T\}$, 则

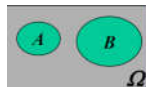
$$A-B = \{HHH\}, \text{ 三次都是正面.}$$

特别地, 对任意的随机事件 A ,

$$A-\Phi = A, A-\Omega = \Phi, A-A = \Phi.$$

5. 互不相容事件

若 $A \cap B = \Phi$, 即它们没有公共样本点, 则表示 A 与 B 不会同时发生, 称 A 与 B **互不相容**.

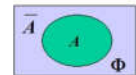


例. $A = \{HHH\}$, 三次都是正面,
 $B = \{TTT\}$, 三次都是反面,
则 A 与 B 互不相容。

特别地, Φ 与任意一个事件 A 互不相容。

6. 对立事件

✓ 由不属于A的所有样本点组成的事件称为A的**逆事件**或**对立事件**, 记为 \bar{A} , 表示A不发生。



例. 抛掷均匀硬币三次, $A = \{HHH, TTT\}$, 则 A 的对立事件的样本点是 $\{HHT, HTH, HTT, THH, THT, TTH\}$, 即三次出现的结果不全相同。

性质: $\bar{\bar{A}} = A, A \cup \bar{A} = \Omega, A \cap \bar{A} = \Phi.$

附:

事件与集合的关系及其运算对照表

符号	集合论含义	概率论含义
Ω	空间或全集	样本空间或必然事件
Φ	空集	不可能事件
ω	元素	样本点
A	子集	随机事件
$\omega \in A$	ω 是 A 的元素	事件 A 包含样本点 ω

符号	集合论含义	概率论含义
$A \subset B$	A 是 B 的子集	A 发生将导致 B 发生
$AB = \Phi$	A, B 不相交	A, B 不可能同时发生
$A \cup B$	并集	A, B 至少有一个发生
$A \cap B$	交集	A, B 同时发生
$A-B$	差集	A 发生而 B 不发生
\bar{A}	补集(余集)	A 不发生

1.2.4 随机事件的运算规则

类似于集合运算, 事件的运算满足:

(1) 交换律 $A \cup B = B \cup A, AB = BA$;

(2) 结合律 $(A \cup B) \cup C = A \cup (B \cup C),$
 $(A \cap B) \cap C = A \cap (B \cap C);$

(3) 分配律 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$
 $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)。$

(4) De Morgan(德摩根) 公式

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \overline{A \cap B} = \overline{A} \cup \overline{B}$$

例 A, B, C 是三个随机事件, 利用事件的关系与运算表示出下面新的随机事件。

(1) A 发生而 B 与 C 都不发生:

$$A - B - C \text{ 或 } A - (B \cup C) \text{ 或 } A\overline{B}\overline{C}$$

(2) A, B 都发生而 C 不发生:

$$AB - C \text{ 或 } AB - ABC \text{ 或 } ABC$$

(3) A, B, C 都发生: ABC

(4) 恰好只有一个发生: $A\overline{B}\overline{C} + \overline{A}B\overline{C} + \overline{A}\overline{B}C$

(5) 恰好只有两个发生: $AB\overline{C} + A\overline{B}C + \overline{A}BC$

(6) 至少有一个发生:

$$A \cup B \cup C \text{ 或}$$

$$A\overline{B}\overline{C} + \overline{A}B\overline{C} + \overline{A}\overline{B}C + AB\overline{C} + A\overline{B}C + \overline{A}BC + ABC$$

1.2.5 有限样本空间

若样本空间 Ω 只包含有限个样本点, 对每个样本点赋予一个非负实数: $\omega_k \rightarrow p_k$, 则这些 p_k 称为**样本点 ω_k 的概率**, 如果全部这些 p_k 的和为1.

定义: 对任何随机事件 $A \subset \Omega$, **A 发生的概率**就是它包含的样本点概率之和。

1.3 古典概率模型

1.3.1 模型与计算公式

如果试验 E 满足

(1) 样本空间只含有限多个样本点, 即

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$$

(2) 且每种结果发生的可能性相同, 即

$$P(\{\omega_i\}) = \frac{1}{n}, i = 1, \dots, n,$$

则称这样的试验模型为等可能概率模型或古典概率模型. 简称为等可能型或**古典概型**.

古典概率的计算公式:

$$P(A) = \frac{\text{随机事件 } A \text{ 包含的样本点个数}}{\text{样本空间 } \Omega \text{ 包含的样本点总数}}$$

Remark: 古典概率问题中构造样本空间时必须保证每个样本点是等可能发生的。

例: 抛均匀硬币三次, 计算 $P(A) = P\{\text{恰好出现一次正面}\}$.

$\Omega_1 = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$, 因此 $P(A) = 3/8$;

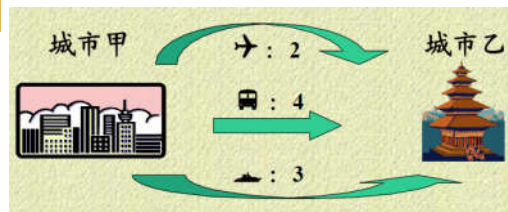
$\Omega_2 = \{0, 1, 2, 3\}$, 因此 $P(A) = 1/4$?

1.3.2 组合分析有关知识

1. 加法原理与乘法原理

加法原理: 假设做一件事情可采用 A 或 B 两类不同方式, B 方式有 m 种不同的方法可以完成这件事, A 方式有 n 种不同的方法可以完成这件事, 则完成这件事情一共有 $n+m$ 种不同的方法。

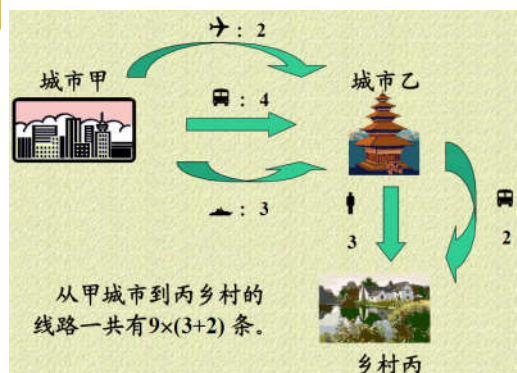
类似地, 如果有若干类方式, 就把所有方式的各种方法全部相加。



则从甲城市到乙城市一共有
 $2+4+3=9$ 条线路

乘法原理: 假设做一件事情必须经过 A 与 B 两个不同步骤, 步骤 A 包含了 n 种不同的方法, 步骤 B 包含 m 种不同的方法, 完成这件事情一共有 $n \times m$ 种不同方法。

类似地, 如果有若干个步骤, 就把所有步骤的各种方法全部相乘。



从甲城市到丙乡村的
线路一共有 $9 \times (3+2)$ 条。

2. 基本排列组合公式

(1) 不可重复的排列

从 n 个不同的元素中无放回地任意取出 r 个 ($1 \leq r \leq n$) 排成有顺序的一列, 称为 n 取 r 的**不可重复排列** (又称为**选排列**), 其不同的排列方法一共有:

$$A_n^r = n(n-1)\dots(n-r+1) = \frac{n!}{(n-r)!}$$

特别当 $r=n$ 时, 上述选排列称**全排列**. 全排列数为

$$P_n = A_n^n = n!$$

(2) 可以重复的排列

从 n 个不同元素中允许**放回**任意取 r 个出来排成有顺序的一列 (即取出的这些元素可以相同)。所有不同的排列方式一共有

$$\underbrace{n \times n \times \dots \times n}_r = n^r$$

例如一个城市的电话号码是 8 位数字, 那么理论上这个城市可以容纳 10^8 , 即一亿门电话。

(3) 二项式组合

从 n 个不同元素中不允许放回任意取 r 个 ($r \leq n$) 构成一个集合, 称为 n 取 r 的**组合**.
构成这个集合的不同组合方法一共有

$$C_n^r = \binom{n}{r} = \frac{A_n^r}{r!} = \frac{n!}{(n-r)!r!} \quad (r \leq n)$$

称 $\binom{n}{r}$ 为**二项系数**, 是下列二项展开式的系数:

$$(a+b)^n = \sum_{r=0}^n \binom{n}{r} a^r b^{n-r}$$

(4) 多项式组合

把 n 个不同元素分成 k 个部分, 各个部分包含的元素个数分别是: r_1, r_2, \dots, r_k ;
则全部不同的分配方式一共有:

$$\left[\begin{matrix} n \\ r_1 \ r_2 \ \dots \ r_k \end{matrix} \right] = \frac{n!}{r_1! \times r_2! \times \dots \times r_k!}$$

称上式中的数为**多项系数**, 是下列多项式展开式的系数:

$$(x_1 + x_2 + \dots + x_k)^n = \sum_{r_1, r_2, \dots, r_k} \left[\begin{matrix} n \\ r_1 \ r_2 \ \dots \ r_k \end{matrix} \right] x_1^{r_1} x_2^{r_2} \dots x_k^{r_k}$$

例如把 52 张扑克牌平均分给 4 个人, 每人 13 张, 则所有不同的分配方案有:

$$\frac{52!}{13! \times 13! \times 13! \times 13!}$$

(5) 关于二项系数的一些公式

设 n, r 为自然数, x 为任意实数.

将排列组合公式中自然数 n 推广至可以取成任意实数 x , 定义:

$$A_x^r = x(x-1)\dots(x-r+1);$$

$$C_x^r = \binom{x}{r} = \frac{A_x^r}{r!} = \frac{x(x-1)\dots(x-r+1)}{r!};$$

$$\text{约定: } 0! = 1, \quad \binom{x}{0} = 1.$$

设 n, r 为自然数, a, b, x 为任意实数, 则下列公式成立:

$$(1) \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n} = 2^n.$$

$$(2) \binom{-x}{r} = (-1)^r \binom{x+r-1}{r}, \text{ 或 } \binom{x}{r} = (-1)^r \binom{-x+r-1}{r};$$

$$(3) \binom{a+b}{n} = \sum_{i=0}^n \binom{a}{i} \binom{b}{n-i}, \text{ 特别地, } \binom{2n}{n} = \sum_{i=0}^n \binom{n}{i}^2.$$

证: (1) 在二项式展开式 $(1+x)^n = \sum_{k=0}^n \binom{n}{k} x^k$, 令 $x=1$ 即得.

(2) 由定义, 易见

$$\binom{-x}{r} = \frac{-x(-x-1)\dots(-x-r+1)}{r!} = (-1)^r \binom{x+r-1}{r};$$

(3) 由 Taylor 公式, 当 $-1 < x < 1$ 时, 有

$$(1+x)^a = \sum_{k=0}^{\infty} \binom{a}{k} x^k, \text{ 称为 } \textbf{牛顿二项式}.$$

$$\text{类似有, } (1+x)^b = \sum_{r=0}^{\infty} \binom{b}{r} x^r, (1+x)^{a+b} = \sum_{n=0}^{\infty} \binom{a+b}{n} x^n.$$

利用幂级数的乘法，计算 $(1+x)^a \cdot (1+x)^b$ 的幂级数展开式中 x^n 项前面的系数知：

$$\binom{a}{0}\binom{b}{n} + \binom{a}{1}\binom{b}{n-1} + \cdots + \binom{a}{n}\binom{b}{0} = \binom{a+b}{n}$$

特别地 $\binom{n}{0}\binom{n}{n} + \binom{n}{1}\binom{n}{n-1} + \cdots + \binom{n}{n}\binom{n}{0} = \binom{2n}{n}$

或

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n}$$

1.3.3 古典概率的例子

- 古典概率的计算一般是先计算样本空间里的样本点总数，再从中挑选出随机事件包含的样本点个数，即“有利场合数”。

古典模型的基本特征：

样本点总数有限，每个样本点等可能发生。

古典模型中的几个经典问题：

- 1、生日问题
- 2、抽签问题
- 3、产品抽样

1. 生日问题(又称分房问题)

[例4] (投球入格)设有 n 个球,每个球都能以同样的概率 $\frac{1}{N}$ 落到 N 个格子 ($N \geq n$) 的每一个格子中,试求:

- (1) 某指定的 n 个格子中各有一个球的概率;
- (2) 任何 n 个格子中各有一个球的概率.

[解] 这是一个古典模型问题,由于每个球可落入 N 个格子中的任一个,所以 n 个球在 N 个格子中的分配相当于从 N 个元素中选取 n 个进行有重复的排列,故共有 N^n 种可能分配.

在第一个问题中,有利场合相当于 n 个球在那指定的 n 个格子中全排列,总数为 $n!$,因而所求概率为

$$P_1 = \frac{n!}{N^n}$$

(2) 将 n 个球放入 N 个盒子,

$$P = \frac{N(N-1)\cdots(N-n+1)}{N^n} = \frac{A_N^n}{N^n} = \frac{N!}{N^n(N-n)!} \quad \text{直接放球}$$

或

$$P = \frac{\binom{N}{n} n!}{N^n} = \frac{A_N^n}{N^n} = \frac{N!}{N^n(N-n)!} \quad \text{先选好格子,再放球}$$

例 (生日问题) 设每个人的生日在一年365天中的任一天是等可能的,即都等于 $\frac{1}{365}$,那么随机选取 n ($n \leq 365$) 人。

- (1) 他们的生日各不相同的概率为多少?
- (2) n 个人中至少有两个人生日相同的概率为多少?

解 (1) 设 $A =$ “ n 个人的生日各不相同”

$$P(A) = \frac{A_{365}^n}{365^n}$$

(2) 设 $B =$ “ n 个人中至少有两个人生日相同”

$$P(B) = 1 - P(A) = 1 - \frac{A_{365}^n}{365^n}$$

n	20	23	30	40	50	64	80	100
p	0.411	0.507	0.706	0.891	0.970	0.997	0.999	0.9999997

2、抽签问题

[例5] 口袋中有 a 只黑球, b 只白球,它们除颜色不同外,其他方面没有差别,现在把球随机地一只只摸出来,求第 k 次摸出的一只球是黑球的概率 ($1 \leq k \leq a+b$)。

解法一: 把小球编号,按 $(a+b)$ 次取球情况构造样本空间,则样本点总数为 $(a+b)!$;第 k 次取到黑球,有 a 种,其余的顺序可以任意排列,因此

$$p_k = \frac{a(a+b-1)!}{(a+b)!} = \frac{a}{a+b}$$

解法二: 把小球编号,按前 k 次取球情况构造样本空间

$$p_k = \frac{a A_{a+b-1}^{k-1}}{A_{a+b}^k} = \frac{a}{a+b}$$

3、产品抽样(二项分布与超几何分布)

[例6] 如果某批产品中有 a 件次品 b 件合格品,我们采用有放回及不放回抽样方式从中抽 n 件产品,问正好有 k 件是次品的概率各是多少?

【放回抽样】

把 $a+b$ 件产品进行编号,有放回的抽 n 次,把可能的重复排列全体作为样本空间。

n 次取出的产品中
究竟哪 k 次是次品
 C_n^k

a 件次品中
取 k 个次品
 a^k

b 件合格品
取出 $n-k$ 件
 b^{n-k}

$$b_k = \frac{\binom{n}{k} a^k b^{n-k}}{(a+b)^n} = \binom{n}{k} \left(\frac{a}{a+b}\right)^k \left(\frac{b}{a+b}\right)^{n-k}$$

b_k 是二项式 $\left(\frac{a}{a+b} + \frac{b}{a+b}\right)^n$ 展开式的一般项。

$\{b_k, k=0,1,\dots,n\}$ 称为 **二项分布**——次品数的概率分布。

【不放回抽样】

从 $a+b$ 件产品中取出 n 件产品的可能组合全体作为样本空间。

$$h_k = \frac{\binom{a}{k} \binom{b}{n-k}}{\binom{a+b}{n}}$$

$\{h_k, k=0,1,\dots,n\}$ 称为 **超几何分布**。

注意: 当产品总数很大而抽样数不大时,采用有放回抽样与采用不放回抽样,差别不大。

1.3.4 古典概率的性质

(1) **非负性:** 对任一事件 A , 有 $0 \leq P(A) \leq 1$ 。

(2) **规范性:** 对必然事件 Ω , 有 $P(\Omega)=1$

(3) **有限可加性:** 若事件 A_1, A_2, \dots, A_n 两两互不相容, 则

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$$

推论: $P(\bar{A}) = 1 - P(A)$

1.3.5 对立事件的概率

[例8] (德·梅尔问题) 一颗骰子投4次至少得到一个六点与两颗骰子投24次至少得到一个双六, 这两件事中哪一件有更多的机会遇到?

解: 以 A 表示一颗骰子投4次至少得到一个六点这一事件, 则 \bar{A} 表示投一颗骰子4次没有出现六点, 故

$$P(\bar{A}) = \frac{5^4}{6^4}$$

因而

$$P_1 = 1 - P(\bar{A}) = 1 - \frac{5^4}{6^4} = 0.5177$$

同理, 若以 B 表示两颗骰子投24次至少得到一个双六, 则

$$P_2 = 1 - P(\bar{B}) = 1 - \frac{35^{24}}{36^{24}} = 0.4914$$

因而, 这两件事情中, 前面一件事情更容易遇到。

这个问题在概率论发展史上颇有名气, 因为它是德梅尔向巴斯卡提出的问题之一。正是这些问题导致了巴斯卡的研究和他与费马的著名通信。他们的研究标志着概率论的诞生。

1.4 几何概型

有时, 试验的可能结果是某区域中的一个点, 这个区域可以是一维的, 也可以是二维的, 还可以是 n 维的, 这时不管是可能结果全体, 还是我们感兴趣的结果都是无限的。

此时, 等可能性可以通过下列方式来赋予意义: 落在某区域 g 的概率与区域的“几何度量” $m(g)$ (长度、面积、体积等) 成正比并且与其位置和形状无关。这种区域的度量统称为“勒贝格 (Lebesgue) 测度”。

即: 若以 A_g 记“在区域 Ω 中随机的取一点, 而该点落在区域 g 中”这一事件, 则其概率定义为:

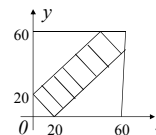
$$P(A_g) = \frac{m(g)}{m(\Omega)}.$$

例1 (会面问题) 两人相约7点到8点在某地会面, 先到者等候另一人20分钟, 这时就可离去, 试求这两人能会面的概率。

解 以 x, y 分别表示两人到达的时刻, 则会面的充要条件为 $|x - y| \leq 20$ 。

可能的结果全体是边长为60的正方形中的点, 能会面的点的区域用阴影标出, 故所求的概率为

$$p = \frac{60^2 - 40^2}{60^2} = \frac{5}{9}$$

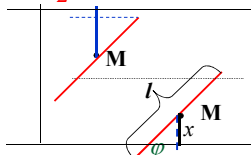


实际上, 我们假定了两人到达的时间在7点到8点之间的机会均等且互不影响。

例2 (蒲丰Buffon投针问题) 平面上有一族平行线. 其中任何相邻的两线距离都是 a ($a > 0$). 向平面任意投一长为 l ($l \leq a$) 的针, 试求针与一条平行线相交的概率。

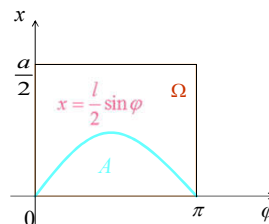
解: 设 x 是针的中点 M 到最近的平行线的距离, φ 是针与此平行线的交角, 投针问题就相当于向平面区域 Ω 取点的几何概型。

$$\Omega = \{(\varphi, x) \mid 0 \leq \varphi \leq \pi, 0 \leq x \leq \frac{a}{2}\}$$



$$A = \{(\varphi, x) \mid 0 \leq \varphi \leq \pi, 0 \leq x \leq \frac{l}{2} \sin \varphi\}$$

$$p = \frac{A \text{ 的面积}}{D \text{ 的面积}} = \frac{\int_0^\pi \frac{l}{2} \sin \varphi d\varphi}{\frac{a}{2} \pi} = \frac{2l}{\pi a}. \quad (1.4.2)$$



◆ 模拟计算 π

由于最后的答案与 π 有关, 因此蒲丰设想利用它来计算 π 的数值, 其方法是投针 N 次, 计算针与线相交的次数 n , 再以频率值 $\frac{n}{N}$ 作为概率 p 之值代入 (1.4.2), 求得

$$\pi = \frac{2lN}{an} \quad (1.4.3)$$

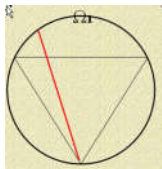
试验者	时间	针长	投掷次数	相交次数	π 的近似值
Wolf	1850	0.8	5000	2532	3.1596
Smith	1855	0.6	3204	1218	3.1554
De Morgan	1860	1.0	600	382	3.137
Fox	1884	0.75	1030	489	3.1595
Lazzerini	1901	0.83	3408	1808	3.1415929
Reina	1925	0.5419	2520	859	3.1795

几何概型在现代概率概念的发展中曾经起过重大作用。19世纪时, 不少人相信, 只要找到适当的等可能性描述, 就可以给概率问题以唯一的解答, 然而有人却构造出这样的例子, 它包含着几种似乎都同样有理却相互矛盾的答案。下面就是一个著名的例子。

贝特朗(Bertrand)奇论 在半径为1的圆内随机地取一弦, 求其长超过该圆内接等边三角形的边长 $\sqrt{3}$ 的概率。

【解法一】

任何弦交圆周两点，不失一般性，先固定其中一点于圆周上，以此点为顶点作等边三角形，显然只有落入此三角形内的弦才满足要求，这种弦的弧长为整个圆周的 $1/3$ ，故所求的概率为 $1/3$ 。

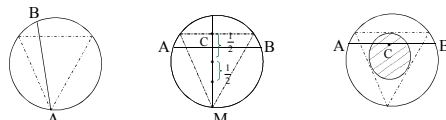


【解法二】

弦长只跟它与圆心的距离有关，而与方向无关，因此可以假定它垂直于某一直径，当且仅当它与圆心的距离小于 $1/2$ 时，其长才大于 $\sqrt{3}$ ，因此所求的概率为 $1/2$ 。

【解法三】

弦被其中点唯一确定，当且仅当其中点属于半径为 $1/2$ 的同心圆时，弦长才大于 $\sqrt{3}$ ，此小圆面积为大圆面积的 $1/4$ ，故所求的概率为 $1/4$ 。



同一问题有三种不同的答案，细究其原因，发现是在取弦时采用了不同的等可能性假定。在第一种解法中，假定端点在圆周上均匀分布，在第二种解法中，假定弦的中点在直径上均匀分布，而在第三种解法中，又假定弦的中点在圆内均匀分布。这三种答案针对三种不同的随机试验，对于各自的随机试验而言，它们都是正确的。

因此在使用术语“随机”、“等可能”、“均匀分布”等时，应明确指明其含义，这又因试验而异。

◆ 几何概率的性质：

(1) **非负性**：对任一事件 A ，有 $0 \leq P(A) \leq 1$ 。

(2) **规范性**：对必然事件 Ω ，有 $P(\Omega)=1$ 。

(3) **可列可加性**：若事件 A_1, A_2, \dots ，两两互不相容，则

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

1.5 概率空间

采用等可能性来定义概率适用范围有限，对一般的随机现象如何明确地定义概率及其他基本概念？

前苏联数学家Kolmogorov于1933年在《概率论基础》一书中，用公理化的方法与集合论的观点成功地解决了这一问题，他在定义概率这一基本概念时只指明概率应具有的基本性质，而把具体概率的给定放在一边，这样做的好处是能针对不同的随机试验给定适当的概率。

Kolmogorov提出的公理为数很少且极为简单，但在此基础上建立起了概率论的宏伟大厦。

一、事件域

- **随机试验**：一个试验(或观察)，若它的结果预先无法确定，则称之为随机试验，简称试验；
- **样本空间**：所有试验的可能结果组成的集合，称为样本空间，记作 Ω ；
- **样本点**： Ω 中的元素称为样本点，用 ω 表示；
- **事件**：事件 A 定义为 Ω 的一个子集，它包含若干样本点，事件 A 发生当且仅当 A 所包含的样本点中有一个发生。常用大写字母 A, B, C 等表示；
- **集类**：由 Ω 中的若干子集构成的集合称为集类，用花写字母 \mathcal{B}, \mathcal{F} 等表示；

问题: 针对哪些事件给出概率?

- 记 \mathcal{F} 为研究的所有事件的全体。
- \mathcal{F} 一般不包括所有的事件, 即样本空间 Ω 的一切子集, 因为这将给给定概率带来困难。
- \mathcal{F} 必须把感兴趣的事件包含进来。
- 因为事件的交、并、逆等也应该为事件, 也应该有相应的概率, 需要把它们亦包括进来。
- 当然, Ω 和 Φ 必不可少。

定义 (σ 域) 满足下列条件的集类 \mathcal{F} 称为 σ 域或 σ 代数:

- (i) $\Omega \in \mathcal{F}$; (包含全集)
- (ii) 如果 $A \in \mathcal{F}$, 那么 $\bar{A} \in \mathcal{F}$; (对逆运算封闭)
- (iii) 如果 $A_n \in \mathcal{F}$, $n=1, 2, \dots$, 那么 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$; (对可列并运算封闭)

若 \mathcal{F} 是由样本空间 Ω 的一些子集构成的一个 σ 域, 则称它为事件域. \mathcal{F} 中的元素称为事件, Ω 称为必然事件, Φ 称为不可能事件.

很显然, 根据定义, 必然事件和不可能事件都在事件域中, 事件的有限及可列交、并也都在事件域中.

例1: $\mathcal{F} = \{\emptyset, \Omega\}$ 为一 σ 域。

例2: $\mathcal{F} = \{\emptyset, A, \bar{A}, \Omega\}$ 为一 σ 域。

例3: $\Omega = \{\omega_1, \dots, \omega_n\}$, \mathcal{F} 是由 Ω 的一切子集构成。

这时, \mathcal{F} 是一个有限的集合。共有元素 2^n 个。

\mathcal{F} 为一 σ 域。

对于一般的 Ω , 若 \mathcal{F} 由 Ω 的一切子集构成。

可以验证 \mathcal{F} 为一 σ 域。

注: 事件域可以很简单, 也可以十分复杂, 要根据问题的不同要求来选择适当的事件域。

命题 给定 Ω 的一个非空集类 \mathcal{A} , 必然存在唯一的一个 Ω 上的 σ 域 $m(\mathcal{A})$, 满足: (1) 包含 \mathcal{A} , (2) 若有其它 σ 域包含 \mathcal{A} , 则必包含 $m(\mathcal{A})$ 。

$m(\mathcal{A})$ 称为包含 \mathcal{A} 的最小 σ 域, 或由 \mathcal{A} 产生的 σ 域。

\mathcal{A} 包含 Ω 的某些子集, 从它的这些最基本子集出发反复进行“最多可列次并、交、逆运算(Borel运算)”, 不断添加子集从而得到所需要的 σ 代数。或者把包含这些基本子集的全部 σ 代数做交运算。

■ 博雷尔(Borel) 集类

以后, 用 \mathbf{R}^1 记数直线或实数全体, 用 \mathbf{R}^n 记 n 维欧几里得 (Euclid) 空间。

一维博雷尔(Borel)点集

由一切形为 $[a, b)$ 的有界左闭右开区间构成的集类所产生的 σ 域称为一维博雷尔 σ 域, 记为 \mathcal{B}_1 , \mathcal{B}_1 中的集合称为一维博雷尔点集。

n 维博雷尔点集:

由一切 n 维矩形产生的 n 维博雷尔 σ 域。

若 x, y 表示任意实数, 由于

$$\begin{aligned} \{x\} &= \bigcap_{n=1}^{\infty} [x, x + \frac{1}{n}) & (x, y) &= [x, y) - \{x\} \\ [x, y] &= [x, y) + \{y\} & (x, y] &= [x, y) + \{y\} - \{x\} \end{aligned}$$

因此, \mathcal{B}_1 中包含一切开区间, 闭区间, 单个实数, 可列个实数, 以及由它们经可列次并、交运算而得出的集合。这是一个相当大的集合, 足够把实际问题中感兴趣的点集都包括在内。

同样, \mathcal{B}_n 也是一个相当大的集合, 足够把实际问题中感兴趣的点集都包括在内。

二、概率的公理化定义

定义：设 Ω 为样本空间， \mathcal{F} 为 Ω 上的事件域，称定义上的集合函数 $P(\cdot)$ 为 \mathcal{F} 上的一个**概率测度**，若它满足：

- (i) **非负性**： $\forall A \in \mathcal{F}, P(A) \geq 0$
- (ii) **规范性**： $P(\Omega) = 1$;
- (iii) **可列可加性**：若 $A_n \in \mathcal{F}, n = 1, 2, \dots$ 且两两互不相容，则

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

其中，对任意具体事件 A ，函数值 $P(A)$ 称为事件 A 的**概率**。三元组 (Ω, \mathcal{F}, P) 称为**概率空间**。

■ 概率测度的其他性质：

$$1^0 \quad P(\emptyset) = 0$$

$$2^0 \quad \text{有限可加性：即若 } A_i A_j = \emptyset, \text{ 则}$$

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

$$3^0 \quad P(\bar{A}) = 1 - P(A)$$

$$4^0 \quad \text{若 } A \subset B, \text{ 则 } P(B - A) = P(B) - P(A)$$

$$\textcircled{1} \quad \text{如果 } A \subset B, \text{ 则 } P(A) \leq P(B)$$

$$\textcircled{2} \quad \forall A \subset \Omega, \text{ 有 } 0 \leq P(A) \leq 1$$

$$5^0 \quad \text{概率的加法公式：} P(A \cup B) = P(A) + P(B) - P(AB)$$

$$(1) \quad \text{布尔不等式：} P(A \cup B) \leq P(A) + P(B)$$

$$(2) \quad \text{Bonferroni不等式：} P(AB) \geq P(A) + P(B) - 1$$

6⁰ 加法公式的推广：

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1, \dots, n} P(A_i) - \sum_{\substack{i < j \\ i, j = 1, \dots, n}} P(A_i A_j) \\ + \sum_{\substack{i < j < k \\ i, j, k = 1, \dots, n}} P(A_i A_j A_k) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n)$$

利用上面的公式来作概率的计算，常能使解题思路清晰，计算便捷。

例5 (最大车牌号) 某区域有 N 部卡车，车牌号从1到 N ，有一外地人到该区域去，把遇到的 n 部卡车的车牌号抄下来（可以重复），以 A 表示“抄到的最大号码正好为 k ($1 \leq k \leq N$)”，求 A 的概率。

解：记 $A_k = \{\text{记到的最大号码为 } k\}$ ， $B_k = \{\text{记到的最大号码不超过 } k\}$ ，则 $A_k = B_k - B_{k-1}$ ，

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

$$P(A_k) = P(B_k) - P(B_{k-1}) = \frac{k^n}{N^n} - \frac{(k-1)^n}{N^n}$$

例6 (匹配问题)

某人写好 n 封信，又写好 n 只信封，然后在黑暗中把每封信放入一只信封中，试求至少有一封信放对的概率。

解：若以 A_i 记第 i 封信与信封符合，则所求的事件为 $A_1 \cup A_2 \cup \dots \cup A_n$ ，则

$$P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}, \quad P(A_i A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)},$$

$$P(A_i A_j A_k) = \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)}, \quad \dots, P(A_1 A_2 \dots A_n) = \frac{1}{n!}$$

因此，

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \binom{n}{1} \frac{1}{n} - \binom{n}{2} \frac{1}{n(n-1)}$$

$$+ \binom{n}{3} \frac{1}{n(n-1)(n-2)} + \dots + (-1)^{n-1} \frac{1}{n!}$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{n-1} \frac{1}{n!}$$

$n \rightarrow \infty, P = ?$

三、概率的可列可加性与连续性

定义1: 若 $S_n \in \mathcal{F}, n=1, 2, \dots$ 且 $S_n \subset S_{n+1}$, 则 S_n 是 \mathcal{F} 中的一个**单调不减的集序列**。

若 $S_n \in \mathcal{F}, n=1, 2, \dots$ 且 $S_n \supset S_{n+1}$, 则 S_n 是 \mathcal{F} 中的一个**单调不减的集序列**。

定义2: 对于 \mathcal{F} 上的集合函数 $P(\cdot)$, 若它对 \mathcal{F} 中任何一个**单调不减的集序列** $\{S_n\}$ 均有:

$$\lim_{n \rightarrow \infty} P(S_n) = P(\lim_{n \rightarrow \infty} S_n) \quad (1)$$

成立, 则我们称它是**下连续的**。

若(1)式对 \mathcal{F} 中任何一个**单调不减的集序列** 均成立, 则我们称它是**上连续的**。

定理 若 P 为 \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:

- (i) 它是有限可加的;
- (ii) 它是下连续的。

分析: 即要证明:

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) (*) \Leftrightarrow \begin{cases} P\left(\sum_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) & (1) \\ P(\lim_{n \rightarrow \infty} S_n) = \lim_{n \rightarrow \infty} P(S_n) & (2) \end{cases}$$

其中 A_i 互不相容, $\{S_n\}$ 为单调不减的集序列, 即 $S_n \uparrow$ 。

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) (*) \Leftrightarrow \begin{cases} P\left(\sum_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) & (1) \\ P(\lim_{n \rightarrow \infty} S_n) = \lim_{n \rightarrow \infty} P(S_n) & (2) \end{cases}$$

其中 A_i 互不相容, $\{S_n\}$ 为单调不减的集序列, 即 $S_n \uparrow$ 。

证明: " \Rightarrow " (1) 已证明, 下面证明 (2)。

$$\begin{aligned} P\left(\lim_{n \rightarrow \infty} S_n\right) &= P\left(\bigcup_{i=1}^{\infty} S_i\right) = P\left(\sum_{i=1}^{\infty} (S_i - S_{i-1})\right) \quad S_0 = \emptyset \\ &\stackrel{(*)}{=} \sum_{i=1}^{\infty} P(S_i - S_{i-1}) = \sum_{i=1}^{\infty} [P(S_i) - P(S_{i-1})] \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n [P(S_i) - P(S_{i-1})] = \lim_{n \rightarrow \infty} P(S_n). \quad (2) \text{ 得证。} \end{aligned}$$

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) (*) \Leftrightarrow \begin{cases} P\left(\sum_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) & (1) \\ P(\lim_{n \rightarrow \infty} S_n) = \lim_{n \rightarrow \infty} P(S_n) & (2) \end{cases}$$

其中 A_i 互不相容, $\{S_n\}$ 为单调不减的集序列, 即 $S_n \uparrow$ 。

" \Leftarrow "

$$\begin{aligned} P\left(\sum_{i=1}^{\infty} A_i\right) &= P\left(\lim_{n \rightarrow \infty} \left(\sum_{i=1}^n A_i\right)\right) \stackrel{(2)}{=} \lim_{n \rightarrow \infty} P\left(\sum_{i=1}^n A_i\right) \\ &\stackrel{(1)}{=} \lim_{n \rightarrow \infty} \sum_{i=1}^n P(A_i) = \sum_{i=1}^{\infty} P(A_i) \end{aligned}$$

这样, 我们便证得 (*) 式。

推论1 概率是下连续的。

推论2 概率是上连续的。

证明 设 $A_n \downarrow$, 则 $\bar{A}_n \uparrow$, 这样, 由推论1可知:

$$\lim_{n \rightarrow \infty} P(\bar{A}_n) = P(\lim_{n \rightarrow \infty} \bar{A}_n) = P\left(\bigcup_{n=1}^{\infty} \bar{A}_n\right) = P\left(\overline{\bigcap_{n=1}^{\infty} A_n}\right)$$

因而

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

即

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

$$\text{推论 3. } P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i).$$

四、概率空间的实际例子

在科尔莫戈罗夫的概率论公理化结构中, 称三元组 (Ω, \mathcal{F}, P) 为概率空间, 以此为出发点讨论种种问题。至于实际问题中, 如何选定 Ω , 怎构造 \mathcal{F} , 怎样给定 P , 要视具体情况而定。

例7 Bernoulli概率空间

取 $\mathcal{F} = \{\emptyset, A, \bar{A}, \Omega\}$, 其中 A 为 Ω 的非空真子集. 任取两个正数 p 与 q ($p+q=1$), 令

$$P(\emptyset) = 0, P(A) = p, P(\bar{A}) = q, P(\Omega) = 1$$

易证, 此 P 是一个概率测度, 从而 (Ω, \mathcal{F}, P) 是一个概率空间。它是描述 Bernoulli 试验的概率空间。

例8 有限概率空间

样本空间为有限集 $\Omega = \{\omega_1, \dots, \omega_n\}$, 事件域 \mathcal{F} 取为 Ω 的一切子集 (共 2^n 个) 组成的集类. 取 n 个非负实数

$$p_1, p_2, \dots, p_n, \quad \text{使} \quad p_1 + p_2 + \dots + p_n = 1.$$

最后, 对 Ω 的每一个子集 A , 令

$$P(A) = \sum_{\omega_i \in A} p_i \quad (4)$$

易证, 此 P 是一个概率测度, 从而 (Ω, \mathcal{F}, P) 是一个概率空间。

特别取 $p_i = \frac{1}{n}, i = 1, 2, \dots, n$, 就是古典概型空间。

例9 离散概率空间

样本空间为可列集 $\Omega = \{\omega_1, \omega_2, \dots\}$, 事件域 \mathcal{F} 仍取为 Ω 的一切子集组成的集类. 取非负实数列 $\{p_n\}$ 使 $\sum_{n=1}^{\infty} p_n = 1$.

再按(4)式定义概率 $P(A)$, 则 (Ω, \mathcal{F}, P) 是一概率空间, 称为离散概率空间。

例10 一维几何概率空间

样本空间 Ω 为 $(-\infty, +\infty)$ 中的博雷尔点集, 具有正的有限的勒贝格测度 $m(\cdot)$ 。事件域 \mathcal{F} 取作 Ω 中的博雷尔集类。

对每个事件 A , 取 $P(A) = \frac{m(A)}{m(\Omega)}$, 则它为一概率测度。

于是得到几何概型的概率空间。

从上面的例子可以看到下面两点:

(1) 选定了 (Ω, \mathcal{F}) 之后, 对于事件概率的给定还有相当大的灵活性。因为只有这样, 才能用概率空间来描述不同的随机现象。

(2) 事件 A 的概率不能任意给定, 即在事件域中, 各事件的概率有一定的关系, 给定概率必须满足这些关系。