

概率论 与数理统计

PROBABILITY AND MATHEMATICAL
STATISTICS

傅冬生 赵进 谢兆茹 刘荣丽 编 ©



科学出版社

概率论与数理统计

傅冬生 赵 进 谢兆茹 刘荣丽 编

科学出版社

北 京

内 容 简 介

本书内容包括随机事件与概率, 随机变量及其概率分布, 随机向量及其分布, 极限理论, 统计量与抽样分布, 参数估计, 假设检验, 方差分析, 回归分析等。本书不仅重视基础知识的完整性与易懂性, 有丰富的例题解释定理与理论, 而且还重视理论与应用的结合, 注意应用性例题的选择, 引导学生注重概率统计在本专业的应用。

本书可作为高等院校(非数学专业)概率论与数理统计课程的教材或参考书, 也可作为具有高等数学知识的实际工作者的自学参考书。

图书在版编目(CIP)数据

概率论与数理统计/傅冬生等编. —北京: 科学出版社, 2014.6

ISBN 978-7-03-041182-2

I. ①概… II. ①傅… III. ①概率论-高等学校-教材 ②数理统计-高等学校-教材 IV. ①O21

中国版本图书馆 CIP 数据核字 (2014) 第 129062 号

责任编辑: 于盼盼 顾晋饴 / 责任校对: 胡小洁

责任印制: 肖 兴 / 封面设计: 许 瑞

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

新科印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2014 年 6 月第 一 版 开本: 787×1092 1/16

2014 年 7 月第二次印刷 印张: 15 1/4

字数: 360 000

定价: 38.00 元

(如有印装质量问题, 我社负责调换)

前 言

概率论与数理统计作为高等数学教育中的一门重要基础课,在自然科学、社会科学等诸多领域有着广泛的应用。尤其是近年来,随着计算机技术的迅速发展,概率统计在经济、金融、保险、农业、医学、气象、环境保护等方面均起到了重要的作用。本书是我们在南京大学多年教学实践的基础上编写的,本书可以作为高等学校理工、农医、经济、管理等专业概率论与数理统计课程的教材或参考书。

本书在写作与取材上,一方面着重数学理论的严谨性,对一些重要的基本概念,特别是统计概念的由来与实质,采取较多的篇幅加以说明;另外在内容及符号的书写上,参照了硕士生招生考试中概率统计部分的要求,以求适合部分考研读者的需要;最后,我们还注意例题的多样性,书中给出了很多具有实际背景的例题。

本书第一、二章由傅冬生编写,第三~五章由刘荣丽编写,第六~八章由赵进编写,第九、十章由谢兆茹编写。本书的编写与出版得到了南京大学数学系领导朱晓胜老师和黄卫华老师的大力支持,在此表示衷心的感谢。

由于编者水平有限,错谬之处在所难免,恳请广大读者批评指正。

编 者

2014 年 3 月

目 录

前言

第一章 随机事件与概率	1
第一节 随机事件及其运算	1
一、随机试验与随机事件	1
二、事件间的关系及运算	2
第二节 事件的概率及性质	4
一、频率与概率	4
二、概率的定义及性质	5
第三节 等可能概型 (古典概型)	7
第四节 几何概率	9
第五节 条件概率	12
一、条件概率	12
二、乘法公式	14
三、全概率公式与贝叶斯公式	15
第六节 独立性	18
一、独立性定义	18
二、多个事件的独立性	18
三、可靠性分析	20
第七节 独立重复试验概型	21
习题一	24
第二章 随机变量及其概率分布	27
第一节 随机变量及其分布函数	27
一、随机变量	27
二、随机变量的分布函数	28
第二节 离散型随机变量及其分布	31
一、离散型随机变量	31
二、常见离散型随机变量	32
第三节 连续型随机变量及其分布	36
一、连续型随机变量	36
二、常见连续型随机变量	38
第四节 随机变量函数的分布	43
一、离散型随机变量的函数	43
二、连续型随机变量的函数	45

习题二	48
第三章 随机向量及其分布	51
第一节 二维随机向量及其分布函数	51
一、二维离散型随机向量	51
二、二维连续型随机向量	55
三、 n 维随机向量及其分布	59
第二节 条件分布	60
一、离散型随机向量的条件概率分布	60
二、连续型随机向量的条件概率	62
第三节 随机变量的独立性	65
第四节 二维随机向量函数的分布	68
一、二维离散型随机向量函数的分布	68
二、二维连续型随机向量函数的分布	70
习题三	76
第四章 随机变量的数字特征	79
第一节 数学期望	79
一、数学期望的定义	79
二、常见分布的数学期望	82
三、随机变量函数的数学期望	84
四、数学期望的性质	86
第二节 方差与矩	88
一、方差的定义	88
二、常见分布的方差	89
三、方差的性质	91
四、矩	94
第三节 协方差与相关系数	95
一、协方差	95
二、相关系数	98
*第四节 条件数学期望简介	100
习题四	101
第五章 极限理论	104
第一节 大数定律	104
第二节 中心极限定理	106
习题五	108
第六章 统计量与抽样分布	110
第一节 总体与样本	110
一、总体与个体	110
二、样本	111

第二节 统计量与抽样分布	111
第三节 正态总体	113
一、 χ^2 分布	113
二、 t 分布	114
三、 F 分布	115
四、上 α 分位点	116
五、正态总体的样本均值与样本方差的分布	117
习题六	118
第七章 参数估计	120
第一节 矩估计	120
第二节 极大似然估计	122
第三节 估计量的评价标准	126
一、无偏性	126
二、均方误差准则	127
三、一致性	128
第四节 区间估计	129
一、基本概念与枢轴变量法	129
二、正态总体 $N(\mu, \sigma^2)$ 中均值 μ 的置信区间	131
三、正态总体 $N(\mu, \sigma^2)$ 中方差 σ^2 的置信区间	132
四、两个正态总体 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ 的均值差 $\mu_1 - \mu_2$ 的置信区间	132
五、两个正态总体 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ 的方差比 σ_1^2/σ_2^2 的置信区间	133
六、非正态总体均值的区间估计 (大样本法)	134
习题七	135
第八章 假设检验	137
第一节 假设检验的基本概念	137
一、假设检验问题的提出	137
二、假设检验的步骤	138
三、假设检验的两类错误	138
四、 p 值检验法	139
第二节 正态总体均值的假设检验	140
一、单个正态总体 $N(\mu, \sigma^2)$ 均值 μ 的假设检验	140
二、两个正态总体 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ 的均值差的检验	142
三、基于成对数据的假设检验	143
第三节 正态总体方差的假设检验	144
一、单个正态总体 $N(\mu, \sigma^2)$ 方差 σ^2 的假设检验	144
二、两个正态总体 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ 的方差比 σ_1^2/σ_2^2 的假设检验	145
第四节 拟合优度检验	146
第五节 独立性检验	149

习题八	151
第九章 方差分析	154
第一节 单因素实验的方差分析	154
一、单因素试验	154
二、平方和分解	156
三、 S_E, S_A 的统计特性	157
四、假设检验问题的拒绝域	159
五、未知参数的估计	160
第二节 双因素试验的方差分析	162
一、双因素等重复实验的方差分析	162
二、双因素无重复试验的方差分析	168
习题九	171
第十章 回归分析	173
第一节 一元线性回归	173
一、一元回归模型	174
二、 a, b 的最小二乘 (LS) 估计	175
三、参数的极大似然估计	176
四、线性假设的显著性检验	180
五、系数 b 的置信区间	186
六、回归函数 $\mu(x) = a + bx$ 函数值的点估计和置信区间	186
七、 Y 的观察值的点预测和区间预测	187
八、可化为一元线性回归的例子	188
第二节 多元线性回归	190
一、 $b_0, b_1, \dots, b_p$ 的最小二乘估计	191
二、对多元线性回归的各种统计分析	194
习题十	199
习题答案	202
附录	208
附表 1 几种常用的概率分布	208
附表 2 泊松分布表	211
附表 3 标准正态分布表	217
附表 4 t 分布表	218
附表 5 χ^2 分布表	220
附表 6 F 分布表	223
附表 7 检验相关系数的临界值表	233

第一章 随机事件与概率

第一节 随机事件及其运算

一、随机试验与随机事件

自然、社会中千变万化的各种现象可以分为两类：确定性现象与非确定性现象。确定性现象只有一种结果，如太阳每天从东边升起，西边落下；正常情况下，水在 $100\text{ }^{\circ}\text{C}$ 会沸腾。非确定性现象有多种结果，如抛一枚硬币，可能正面向上，也可能反面向上；观察未来某天的天气，可能有晴天、阴天、雨天等多种情况。

非确定性现象也称随机现象。随机现象可以通过随机试验来研究。所谓随机试验是指对随机现象的一次观测。把随机试验记为 E ，随机试验具有如下特点：① 在相同的条件下试验可重复进行；② 每次试验具有多种结果，但试验之前可知道试验的所有可能结果；③ 每次试验会出现这些可能结果之一，但试验前不能确定哪一个结果会出现。

在随机试验 E 中，把所有可能结果的集合称为样本空间，记为 Ω 。样本空间的元素，即 E 的每个可能结果称为基本事件或称样本点，记为 e 。例如：

E_1 ：某班级共 n 人，观察某天上课迟到的人数，

$\Omega_1 = \{0, 1, 2, \dots, n\}$ ，样本点共 $n+1$ 个；

E_2 ：抛两颗骰子，观察先后出现的点数，

$\Omega_2 = \{(1, 1), \dots, (1, 6), (2, 1), \dots, (2, 6), \dots, (6, 1), \dots, (6, 6)\}$ ，样本点共 36 个；

E_3 ：某人在一小时内随机到达车站，记录到达时刻（实数，单位：分钟），

$\Omega_3 = \{x | 0 \leq x \leq 60\}$ ，样本点无穷多个。

随机事件是可能发生也可能不发生的事件。从随机试验的角度，可以把随机事件定义为样本空间 Ω 的子集合。随机事件常记为 A, B, C 等。随机事件可以简称为事件。

例如：袋中有 1~5 号球，任取一只，观察球号。 $\Omega = \{1, 2, \dots, 5\}$ 。

定义随机事件： A ：取得偶数号球； B ：取得奇数号球； C ：球号不超过 3。 A, B, C 所包含的样本点分别为： $A = \{2, 4\}$ ； $B = \{1, 3, 5\}$ ； $C = \{1, 2, 3\}$ 。从集合的角度看，随机事件 A, B, C 均为样本空间 Ω 的子集合。

随机事件的发生：该随机事件所包含的某个样本点在随机试验 E 中出现，则称该随机事件发生。如袋中取球的例子，若取到 2 号球，因为样本点 2 属于事件 A ，也属于事件 C ，所以说事件 A 发生了，事件 C 也发生了。

两个特殊的事件：

必然事件：样本空间 Ω 。因为每次试验必有 Ω 中的样本点出现，因此 Ω 必然发生。

不可能事件：空集 Φ 。因为 Φ 不含任何样本点，每次试验都没有 Φ 中的样本点出现，因此 Φ 不可能发生。

二、事件间的关系及运算

同一样本空间上可以定义多个随机事件, 为了进一步研究这些随机事件间的联系以及用简单的随机事件来表示复杂的随机事件, 我们需要研究事件的关系与运算。

以下设 Ω 是给定的样本空间, 此样本空间上的随机事件为 $A, B, C, A_k (k=1, 2, \dots)$ 等。

由于随机事件定义为样本空间 Ω 的子集合, 因此事件的关系与运算和集合的关系与运算是一致的。

常用 Venn 图来表示事件的关系与运算。即用矩形表示样本空间 Ω , 用 Ω 内的几何图形表示随机事件。

(一) 事件间的关系

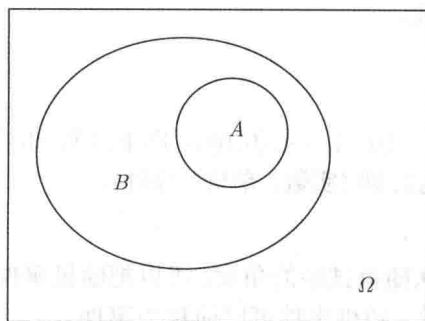
1. 包含关系

如果事件 A 发生必导致事件 B 发生, 则称事件 B 包含事件 A , 记 $B \supset A$ 。

例如, 袋中有 1~10 号球, 任取一个, 记录号码。事件 A : 号码大于 7, 事件 B : 号码大于 5。 A 发生, 号码大于 7, 则一定大于 5, 所以导致 B 发生, $B \supset A$ 。从样本点看, $A = \{8, 9, 10\}$, $B = \{6, 7, \dots, 10\}$ 与集合的包含关系一致 (图 1-1)。

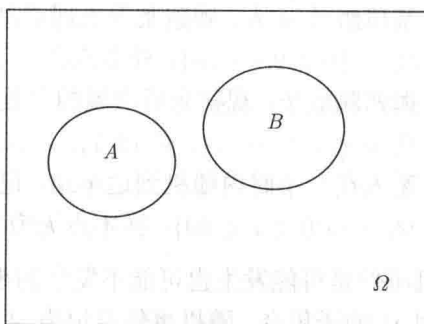
2. 互不相容关系

若事件 A 与 B 不可能同时发生, 则称 A, B 互不相容 (或互斥)。此时, A 与 B 没有共同的样本点 (图 1-2)。



$B \supset A$

图 1-1



A, B 互不相容

图 1-2

3. 相等关系

若 $B \supset A$, 且 $A \supset B$, 则称 A 与 B 相等, 记为 $A = B$ 。 A, B 相等时, 两事件的样本点完全相同。

(二) 事件的运算

1. 事件的并

事件 A 与 B 至少发生一个所构成的事件称为 A 与 B 的并, 记为 $A \cup B$ 。 $A \cup B$ 的样本点由 A 与 B 的样本点合并而成 (图 1-3)。

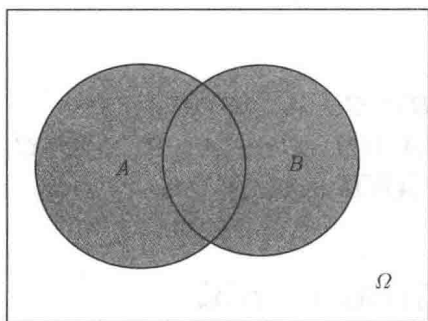
例如: 甲乙射击, A : 甲命中, B : 乙命中。则 $A \cup B$ 表示至少一人命中。

推广：事件 $A_1, A_2, \dots, A_n$ 至少发生一个的事件称为 $A_1, A_2, \dots, A_n$ 的并，记为： $A_1 \cup A_2 \cup \dots \cup A_n$ 或 $\bigcup_{i=1}^n A_i$ 。

2. 事件的交

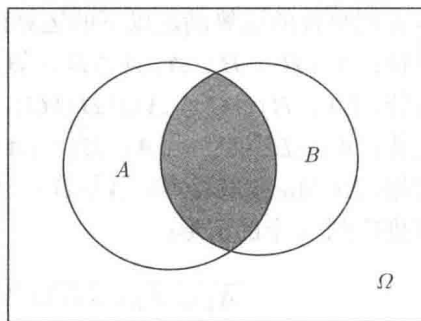
事件 A 与 B 同时发生的事件称为 A 与 B 的交，记为 $A \cap B$ 或 AB 。 $A \cap B$ 的样本点由 A 与 B 共同的样本点组成 (图 1-4)。上例中， $A \cap B$ 表示甲乙同时命中。

推广：事件 $A_1, A_2, \dots, A_n$ 同时发生的事件称为 $A_1, A_2, \dots, A_n$ 的交，记为： $A_1 \cap A_2 \cap \dots \cap A_n$ 或 $\bigcap_{i=1}^n A_i$ 或 $A_1 A_2 \dots A_n$ 。



$A \cup B$

图 1-3



$A \cap B$

图 1-4

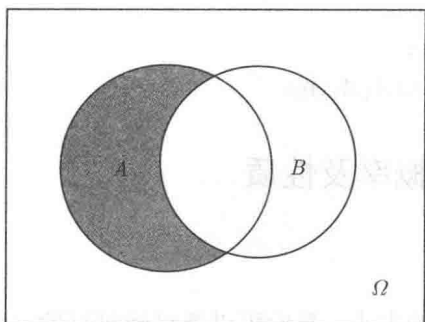
3. 事件的差

事件 A 发生而 B 不发生的事件称为 A 与 B 的差，记为 $A - B$ 。 $A - B$ 的样本点由 A 中除去 B 的样本点组成 (图 1-5)。

上例中， $A - B$ 表示甲命中而乙未命中。

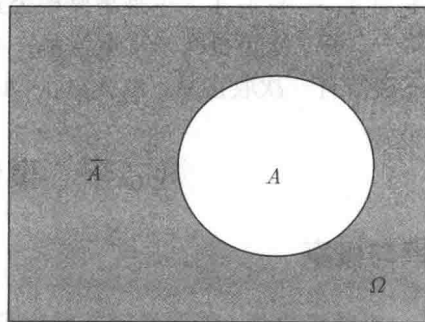
4. 对立事件

A 不发生的事件称为 A 的对立事件，记为 $\bar{A}$ 。 $\bar{A}$ 的样本点由 Ω 中除 A 以外的样本点组成 (图 1-6)。



$A - B$

图 1-5



$\bar{A}$

图 1-6

注意如下等式：

$$(1) A = (A - B) \cup AB.$$

(2) $A \cup B = A \cup (B - A) = (A - B) \cup AB \cup (B - A)$, 第二个等式表示 $A \cup B$ 可以分为互不相容的三个部分。

$$(3) A - B = A - AB = A\bar{B}, \text{ 第二个等式表示事件的差可以用事件的交表示。}$$

$$(4) \bar{A} = \Omega - A, A \cup \bar{A} = \Omega.$$

例 1.1 设事件 $A = \{\text{甲来听课}\}$, $B = \{\text{乙来听课}\}$, 则:

$$A \cup B = \{\text{甲乙至少有一人来}\}, A \cap B = \{\text{甲乙同时来}\}$$

$$\overline{A \cup B} = \bar{A}\bar{B} = \{\text{甲乙都不来}\}, \bar{A} \cup \bar{B} = \overline{A \cap B} = \{\text{甲乙至少有一人不来}\}$$

可以证明事件的运算满足以下的运算律:

$$\text{交换律: } A \cup B = B \cup A, A \cap B = B \cap A$$

$$\text{结合律: } (A \cup B) \cup C = A \cup (B \cup C), (A \cap B) \cap C = A \cap (B \cap C)$$

$$\text{分配律: } A \cup (B \cap C) = (A \cup B) \cap (A \cup C), A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$\text{德摩根 (De Morgan) 定律: } \overline{A \cup B} = \bar{A} \cap \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}$$

可以推广到 n 个的情形:

$$\overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n$$

$$\overline{A_1 \cap A_2 \cap \cdots \cap A_n} = \bar{A}_1 \cup \bar{A}_2 \cup \cdots \cup \bar{A}_n$$

例 1.2 一批产品有正品与次品, 从中抽取 3 个, 用 A_i 表示第 i 个取次品 ($i=1,2,3$), 试用 A_i 的运算表示下列运算。

(1) 三次都取正品;

(2) 三次至少有一次取正品;

(3) 第一、第二次至少有一次取次品;

(4) 三次恰有一次取正品。

解 (1) 三次都取正品: $\bar{A}_1 \bar{A}_2 \bar{A}_3$;

(2) 三次至少有一次取正品: $\bar{A}_1 \cup \bar{A}_2 \cup \bar{A}_3$, 或表示为三次都是次品的对立事件: $\overline{A_1 A_2 A_3}$ 。

而 $\bar{A}_1 \cup \bar{A}_2 \cup \bar{A}_3 = \overline{A_1 A_2 A_3}$, 正是德摩根公式;

(3) 第一、第二次至少有一次取次品: $A_1 \cup A_2$;

(4) 三次恰有一次取正品: $\bar{A}_1 A_2 A_3 \cup A_1 \bar{A}_2 A_3 \cup A_1 A_2 \bar{A}_3$ 。

第二节 事件的概率及性质

一、频率与概率

对于一个随机事件, 希望知道其发生可能性的大小。我们可以通过随机试验 E 来了解这一可能性。设随机事件 A 在 n 次重复试验中共出现 n_A 次, 将其比值 n_A/n 定义为 A 发生的频率, 记为 $f_n(A)$ 。即

$$f_n(A) = \frac{n_A}{n}.$$

例如抛骰子 100 次, 出现六点 16 次, 则在 100 次试验中, “出现六点” 的频率为 $16/100$ 。从直观上看, 频率可以在一定程度上反映事件发生可能性的大小。

事件 A 频率具有如下性质:

- (1) $0 \leq f_n(A) \leq 1$;
- (2) $f_n(\Omega) = 1$;
- (3) 若 $A_1, A_2, \dots, A_n$ 互不相容, 则

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

频率的稳定性 在 n 次试验中, 随着试验次数 n 的增加, 事件的频率将在某个数值 p 附近稳定地摆动, 一般来说 n 越大, 摆动的幅度越小。此数值可以定义为 A 发生的概率, 称为统计概率。

历史上有多个学者做过抛投硬币的试验, 观察正面向上的次数 (表 1-1)。

表 1-1

试验者	投币次数	正面向上次数	频率
蒲丰	4040	2048	0.5096
皮尔逊	24 000	12 012	0.5005
维尼	30 000	14 994	0.4998

试验表明, 随着 n 的增加, 这几次试验中频率逐渐稳定靠近常数 0.5。

图 1-7 为某 120 次抛投硬币试验中, 频率的变化曲线。

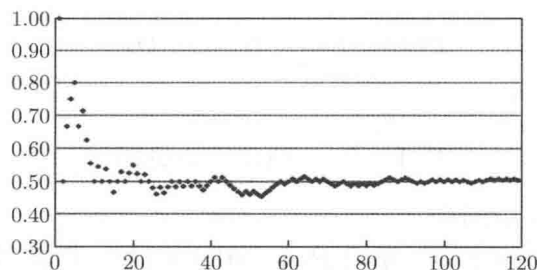


图 1-7

二、概率的定义及性质

1933 年前苏联数学家柯尔莫哥洛夫提出了概率的公理化定义。即通过规定概率应具有的基本性质来定义概率。柯尔莫哥洛夫提出的公理很简洁, 但在此基础上建立起了概率论的理论体系。

定义 1.1 设 E 是随机试验, Ω 为其样本空间。对于 E 的每一个随机事件 A , 对应一个实数 $P(A)$, 如果集合函数 $P(\cdot)$ 满足:

- (1) 非负性 $P(A) \geq 0$;
- (2) 规范性 $P(\Omega) = 1$;

(3) 可列可加性 若 $A_1, A_2, \dots, A_n, \dots$ 两两互不相容, 则

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

则称 $P(A)$ 为事件 A 的概率。

利用概率的公理化定义, 可以得到概率的一些性质。

性质 1 $P(\Phi) = 0$ 。

证明 $\Omega = \Omega \cup \Phi \cup \Phi \cup \Phi \dots$, 由可列可加性:

$$P(\Omega) = P(\Omega) + P(\Phi) + P(\Phi) + P(\Phi) + \dots$$

从而有 $P(\Phi) = 0$ 。

性质 2 若 $A_1, A_2, \dots, A_n$ 为 n 个两两互不相容事件, 则有

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

证明 在可列可加性中, 取 $A_{n+1} = A_{n+2} = \dots = \Phi$, 因为 $P(\Phi) = 0$, 得证。

性质 3 对任意两事件 A, B 有

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

证明 因为 $A = (A - B) \cup AB$, 而 $A - B$ 与 AB 互不相容, 由性质 2,

$$P(A) = P(A - B) + P(AB),$$

从而

$$P(A - B) = P(A) - P(AB)。$$

同理 $P(B - A) = P(B) - P(AB)$ 。

当 $A \supset B$ 时, $P(AB) = P(B)$, 从而 $P(A - B) = P(A) - P(B)$ 且 $P(A) \geq P(B)$ 。

性质 4 对任意两事件 A, B 有加法定理

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

证明 因为 $A \cup B = A \cup (B - A)$, A 与 $B - A$ 互不相容, 再由性质 3, 得到

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

利用归纳法, 可推广到 n 个事件的加法定理

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

性质 5 对任意事件 A , $P(\bar{A}) = 1 - P(A)$ 。

证明 $A \cup \bar{A} = \Omega$, $A \cap \bar{A} = \Phi$, 由性质 2 得

$$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}),$$

从而 $P(\bar{A}) = 1 - P(A)$ 。

以上概率性质中的公式, 也可在 Venn 图得到反映。只要定义 Ω 的面积为 1, 某事件在 Venn 图的面积对应其概率。

如 $P(A \cup B) = P(A) + P(B) - P(AB)$, 在 Venn 图中, $P(A \cup B)$ 的面积正是 A 的面积加上 B 的面积减去 AB 的面积。

第三节 等可能概型 (古典概型)

有一类随机试验, 具有如下特点:

- (1) 样本空间 Ω 的元素个数为有限个;
- (2) 样本空间中的每个基本事件 (样本点) 发生的可能性相同。

把这类随机试验称为等可能概型, 由于它是概率论发展初期的主要研究对象, 又称为古典概型。

设随机试验 E 为等可能概型, 若其样本空间 Ω 含有 n 个样本点, 而事件 A 含有 m 个样本点, 则事件 A 的概率为

$$P(A) = \frac{m}{n} = \frac{A \text{ 包含的样本点数}}{\text{样本点总数}}$$

按这样的定义是合理的。设 $\Omega = \{e_1, e_2, \dots, e_n\}$, 由于等可能, $P(e_i) = 1/n$ 。设 A 含有 m 个基本事件, $A = \{e_{i_1}, e_{i_2}, \dots, e_{i_m}\}$, 则

$$P(A) = P(e_{i_1}) + P(e_{i_2}) + \dots + P(e_{i_m}) = \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n} = \frac{m}{n}$$

例 1.3 一盒元件装有 45 个正品, 5 个次品, 从中任取 5 个。求以下概率:

- (1) A : 恰好取得 3 个正品;
- (2) B : 至多取到 2 个次品。

解 样本点总数为从总共 50 个产品中取出 5 个所有的取法种数 $n = C_{50}^5$ 。

(1) 恰好取得 3 个正品, 另两个为次品, 因此 A 含的样本点数为 $C_{45}^3 C_5^2$, 则

$$P(A) = \frac{C_{45}^3 C_5^2}{C_{50}^5} = 0.0670$$

(2) 至多取到 2 个次品可能有 5 个正品、4 个正品、1 个次品, 3 个正品、2 个次品三种情况, B 含有基本事件数为 $C_{45}^5 + C_{45}^4 C_5^1 + C_{45}^3 C_5^2$, 则

$$P(B) = \frac{C_{45}^5 + C_{45}^4 C_5^1 + C_{45}^3 C_5^2}{C_{50}^5} = 0.9952。$$

例 1.4 袋中有 a 只白球 b 只红球。从中将球依次取出, 问第 k 次取出的球是红球的概率是多少?

解 样本点总数为 $a+b$ 只球依次排列的所有种数 $(a+b)!$ 。设 A : 第 k 个为红球, 则第 k 个红球有 b 种选择, 其余 $a+b-1$ 个位置共有 $(a+b-1)!$ 种排法, 因此, A 包含样本点数为 $b(a+b-1)!$ 。得到

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

注意此结果与 k 无关。由此引申出“抽签原理”。即袋中有 a 只白球, b 只红球。若干人依次从中各取一球, 取后不放回, 则任一人取得红球的概率相等, 都是 $\frac{b}{a+b}$ 。

例 1.5 N 件产品, 含 M 件次品, 其余为正品。现从中任意取出 n 件, 按不放回抽样和有放回抽样两种情况, 分别求其中恰有 k ($k \leq M$) 件次品的概率。

解 **不放回抽样**, 每次取出不放回, 继续抽下一件。样本点总数为 N 件产品取出 n 件的所有取法 C_N^n 。恰有 k 件次品, 则同时取 $n-k$ 件正品, 取次品 C_M^k 种取法, 取正品 C_{N-M}^{n-k} 种取法。因此所求概率为

$$p = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}。$$

有放回抽样, 每次取出放回, 再任取下一件。样本点总数为从 N 件产品中有放回地抽取 n 件的排列种数 N^n 。恰有 k 件次品的取法, 取出的 n 个排成一行, 在 n 个位置选 k 个位置放次品 C_n^k 种, 每个次品有 M 种取法, 取次品共: $C_n^k M^k$ 种; 取 $n-k$ 个正品, $(N-M)^{n-k}$ 种。因此所求概率为

$$p = \frac{C_n^k M^k (N-M)^{n-k}}{N^n} = C_n^k \left(\frac{M}{N}\right)^k \left(1 - \frac{M}{N}\right)^{n-k}。$$

注意, 第二个等号后的写法正好对应后面所讲的独立重复试验概型。

例 1.6 某厂家称一批数量为 1000 件的产品中次品数不超过 50 件。现从该批产品中抽取了 30 件, 经检验发现有次品 5 件, 问该厂家是否谎报了次品数?

解 按最多 50 件次品计, 求出抽取 30 件含 5 件次品的概率。由上例不放回抽样, $N=1000$, $M=50$, $n=30$, $k=5$

$$p = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = \frac{C_{50}^5 C_{950}^{25}}{C_{1000}^{30}} = 0.011$$

这是很小的概率, 即抽取 30 件含 5 件次品的可能性很小。根据“小概率事件原理”, 即概率很小的事件在一次试验中几乎是不发生的。因此几乎可以推断厂家谎报了次品率。

例 1.7 将 n 只小球随机放入 N 个盒子 ($N \geq n$), 求以下概率: (1) 事件 A : 恰有 n 个盒子每盒 1 球; (2) 事件 B : 某指定的 n 个盒子中各有一球。

解 样本点总数: n 只球放入 N 个盒子, 共有 $N \times N \times \cdots \times N = N^n$ 种放法。

(1) n 个盒子是可选的, 共有 C_N^n 种, n 个球放入 n 个盒子 $n!$ 种, 因此 A 含有 $C_N^n n!$ 个样本点, 则 $P(A) = \frac{C_N^n \cdot n!}{N^n}$ 。

(2) 盒子指定, 只需放入 n 个球, $n!$ 种放法, B 含 $n!$ 个样本点, 因此

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

本例的概率模型有很多应用, 例如统计物理中的麦克斯韦-玻尔兹曼统计, 玻色-爱因斯坦统计。以下的“生日问题”同样可以用这一模型解决。

例 1.8 某班级有 n 个同学, 求至少两人生日相同的概率。(假设每人的生日在 365 天中任一天等可能)。

解 设 A : 至少 2 人生日相同, A 是一个复杂的事件。考虑 A 的对立事件: n 个同学的生日都不同。这相当于将 n 球放入 365 个盒子, 使得恰有 n 个盒子每盒 1 球。按例 1.3(1) 的情形, 此概率为

$$P(\bar{A}) = \frac{C_N^n \cdot n!}{N^n}$$

因此

$$P(A) = 1 - \frac{C_N^n \cdot n!}{N^n}$$

可以求得, $n=23$ 时, $P(A)=0.507$, 已经超过一半; $n=60$ 时, $P(A)=0.994$ 。这个概率比直观感觉要大。

例 1.9 从 6 双不同的鞋子中任取 4 只, (1) 求恰有 2 只鞋子配成对的概率; (2) 求至少有 2 只鞋子配成对的概率。

解 样本点总数: 从 12 只鞋子中取出 4 只的所有取法, 共 C_{12}^4 种。

(1) 设 A_1 : 恰有 2 只鞋子配成 1 对。先从 6 双中任取 1 双, C_6^1 种; 余下 5 双中任取 2 双, C_5^2 种; 每双再取一只, 2^2 种。因此 A_1 包含 $C_6^1 C_5^2 2^2$ 个样本点。

$$P(A_1) = \frac{C_6^1 C_5^2 2^2}{C_{12}^4} = \frac{16}{33}$$

(2) B : 至少有 2 只鞋子配成对; 设 A_2 : 恰有 4 只鞋子配成 2 对, 则 $B = A_1 \cup A_2$ 。

$$P(A_2) = \frac{C_6^2}{C_{12}^4} = \frac{1}{33}$$

$$P(B) = P(A_1 \cup A_2) = P(A_1) + P(A_2) = \frac{16}{33} + \frac{1}{33} = \frac{17}{33}.$$

第四节 几何概率

古典概型中要求样本空间中的样本点总数有限。而很多实际问题中, 样本空间中的样本点无限, 但各样本点仍具有等可能的特点。这种情况下, 可以用几何方法计算概率。

定义 1.2 若随机试验的样本空间 Ω 对应一个度量有限的几何区域 S , 每一基本事件与 S 内任意点一一对应, 则任一随机事件 A 对应 Ω 中的某一子区域 D 。若事件 A 的概率只和 A 对应的区域 D 的度量成正比, 与 D 的形状及 D 在 S 中的位置无关。 A 发生的概率定义为