

通识教育平台数学课程系列教材

Gailülun

yuShulitongji

概率论与数理统计

罗 汉 彭国强 主编



科学出版社

www.sciencep.com

· 通识教育平台数学课程系列教材 ·

概率论与数理统计

罗 汉 彭国强 主编

科学出版社

北 京

内 容 简 介

本书是为高等本科院校非数学专业学生编写的通识教育平台数学课程系列教材之一。全书共十一章,第一至五章为概率论部分,包括概率论基本概念,随机变量、随机向量及其函数的概率分布,数字特征,大数定理及中心极限定理;第六至十章为数理统计部分,包括抽样分布,参数估计,假设检验,回归分析与方差分析;第十一章讲述了正交试验法的应用。

本书可供高等院校非数学专业学生使用,也可供各类需要提高数学素质和能力的人员使用。

图书在版编目(CIP)数据

概率论与数理统计/罗汉,彭国强主编. —北京:科学出版社,2007
(通识教育平台数学课程系列教材)
ISBN 978-7-03-019718-4

I.概… II.①罗…②彭… III.①概率论—高等学校—教材②数理统计—高等学校—教材 IV.O21

中国版本图书馆CIP数据核字(2007)第129395号

责任编辑:张颖兵 吉正霞/责任校对:梅莹
责任印制:高嵘/封面设计:宝典

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>

武汉市新华印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2007年8月第一版 开本:B5(720×1000)

2007年8月第一次印刷 印张:16 1/2

印数:1-6 000 字数:316 000

定价:26.00元

(如有印装质量问题,我社负责调换)

前 言

教材是体现教学内容和教学方法的知识载体,是进行教育教学的基本工具,也是高等学校深化教育教学改革、全面推进素质教育、培养新世纪创新人才的重要保证.教育部《关于进一步加强高等学校本科教学工作的若干意见》(教高[2005]1号)指出:“加强教材建设,确保高质量教材进课堂.要大力锻炼精品教材,并把精品教材作为教材选用的主要目标.对发展迅速和应用性强的课程,要不断更新教材内容,积极开发新教材,并使高质量的新版教材成为教材选用的主体.”

数学作为科学和技术基础,在决定国家各级人才的素质方面正起着日益重要的作用.高等学校作为培育人才的摇篮,数学课程的开设也就具有特别重要的意义.高等数学是高等教育中涉及学生多、专业门类广、对学生影响深远的基础课程之一,其教材建设工作受到广大教育工作者的普遍关注和重视.

湖南大学历来十分重视高等数学课程建设,其高等数学课程已被评定为国家精品课程.由湖南大学数学与计量经济学院组织编写的《大学数学》系列教材(供非数学专业理工科学生公共数学基础课程使用)被列为普通高等教育“十五”国家级规划教材,其修订版已被列为普通高等教育“十一五”国家级规划教材.为了在教材建设中引进竞争机制,进一步打造精品教材,学院经广泛征求任课教师意见及教学指导委员会研究讨论,决定再组织编写一套高质量的高等数学课程教材,在不同年级学生的教学中两套教材交替使用,并修改、完善.本套教材适合于本科院校非数学专业学生作为数学公共课教材或参考书使用,也可供各类需要提高数学素质和能力的人员使用.

《概率论与数理统计》是本套教材的第四册,由罗汉、彭国强主编,参加本册编写的人员还有王利平、刘先霞等,他们都是长期从事非数学专业本科生数学公共课教学的教师,本书包括随机事件及其概率、随机变量及其概率分布、随机向量、数字特征、大数定律与中心极限定理、样本与抽样分布、参数估计、假设检验、回归分析、方差分析、正交试验法,其中标注“*”号的内容可根据学时多少进行选讲.本书结构严谨,概念、定理及理论叙述准确、精炼,符号使用标准、规范,知识点突出,难点分散,证明和计算过程严谨,例题、习题等均经过精选,具有代表性和启发性.

湖南大学数学与计量经济学院退休教师刘楚中教授在本套教材编写的前期组织中做了大量工作,在此表示衷心感谢.

由于编写时间有限,本教材难免存在不妥之处,希望使用本教材的教师和学生提出宝贵意见.

湖南大学数学与计量经济学院
“概率论与数理统计”教材编写组

2007年5月

目 录

第一章 随机事件及其概率	1
第一节 随机事件	1
一、随机试验和样本空间	1
二、随机事件	2
三、随机事件的关系与运算	2
第二节 随机事件的概率	4
一、古典概型	4
二、随机事件的频率	6
三、概率的公理化定义	7
第三节 概率的加法公式与乘法公式	9
一、概率的加法公式	9
二、条件概率	11
三、概率的乘法公式	12
第四节 全概率公式与贝叶斯公式	14
一、全概率公式	14
二、贝叶斯公式	16
第五节 相互独立性	17
一、事件的相互独立性	17
二、伯努利概型	20
第二章 随机变量及其概率分布	22
第一节 随机变量	22
第二节 离散型随机变量	23
一、离散型随机变量的概率分布	23
二、常见的离散型分布	24
第三节 连续型随机变量	26
一、连续型随机变量的概率密度函数	26
二、连续型随机变量常用的分布	28
第四节 随机变量的分布函数	31
第五节 随机变量函数的分布	38
一、离散型的情形	38
二、连续型的情形	39

第三章 随机向量	43
第一节 随机向量的概率分布	43
一、二维离散型随机向量	43
二、二维连续型随机向量	44
三、随机向量的分布函数	46
第二节 随机向量的边缘分布与随机变量的相互独立性	49
一、随机向量的边缘分布	49
二、随机变量的相互独立性	53
* 第三节 条件概率分布与条件概率密度	58
一、 X, Y 为离散型随机变量	58
二、 X, Y 为连续型随机变量	59
第四节 两个随机变量函数的分布	59
一、 (X, Y) 为离散型	60
二、 (X, Y) 为连续型	61
第四章 数字特征	69
第一节 随机变量的数学期望	69
一、离散型随机变量的数学期望	69
二、连续型随机变量的数学期望	71
三、随机变量函数的数学期望	72
四、数学期望的性质	75
第二节 随机变量的方差	76
一、方差的概念	76
二、常见分布的方差	78
三、方差的性质	81
第三节 随机向量的数字特征	82
一、随机向量函数的数学期望	82
二、数学期望与方差的运算性质	83
三、协方差与相关系数	85
四、随机向量的数学期望与协方差矩阵	88
第五章 大数定律与中心极限定理	90
第一节 大数定律	90
一、切比雪夫不等式	90
二、大数定律	90
第二节 中心极限定理	92
第六章 样本与抽样分布	97
第一节 总体、样本与统计量	97

一、总体	97
二、样本	97
三、统计量	98
第二节 抽样分布	100
一、 U 统计量及其分布	100
二、 χ^2 分布	102
三、 t 分布	104
四、 F 分布	106
第三节 经验分布函数与频率直方图	107
一、经验分布函数	107
二、频率直方图	108
第七章 参数估计	111
第一节 点估计	111
一、矩估计法	111
二、极大似然估计法	114
第二节 估计量的基本评价标准	119
一、无偏性	119
二、有效性	121
三、一致性	123
第三节 正态总体均值的区间估计	125
一、区间估计概述	125
二、单正态总体均值的置信区间	127
三、双正态总体均值差的区间估计	131
第四节 正态总体方差的区间估计	133
一、单正态总体方差的置信区间	133
二、两个正态总体方差比 σ_1^2/σ_2^2 的置信区间	135
第八章 假设检验	139
第一节 假设检验概述	139
一、假设检验的基本思想	139
二、两类错误	140
三、假设检验的基本步骤	141
第二节 正态总体均值的显著性检验	142
一、 U 检验法	142
二、 t 检验法	144

第三节 正态总体方差的显著性检验	148
一、 χ^2 检验法(单个正态总体方差的检验)	148
二、 F 检验法(两个正态总体方差比的检验)	150
* 第四节 总体分布函数的显著性检验	152
一、秩和检验	152
二、 χ^2 检验法	154
第九章 回归分析	158
第一节 一元线性回归分析	158
一、回归分析概述	158
二、一元线性回归模型	160
三、参数的 a, b 和 σ^2 的点估计	161
四、一元线性回归模型的假设检验	164
五、一元线性回归模型的预测与控制	167
第二节 一元非线性回归与多元线性回归	171
一、一元非线性回归的线性化	171
二、曲线回归方程的比较	174
三、多元线性回归分析简介	176
第十章 方差分析	182
第一节 单因素方差分析	182
一、方差分析的数学模型	182
二、方差分析的方法	184
三、未知参数估计	189
第二节 双因素方差分析	190
一、非重复试验的双因素方差分析	190
* 二、等重复试验的双因素方差分析	195
第十一章 正交试验法	201
第一节 不考虑交互作用的正交试验	201
一、正交表	201
二、确定试验方案	202
三、试验结果的直观分析	203
四、正交试验的数学模型	205
五、正交试验的方差分析	205
第二节 水平数不等的正交试验	208
一、利用混合水平的正交表	209
二、拟水平法	210

第三节 考虑交互作用的正交试验·····	212
习题参考答案·····	216
附录 常用概率统计表·····	226
附表一 泊松分布表·····	226
附表二 标准正态分布表·····	228
附表三 χ^2 分布表 ·····	229
附表四 t 分布表 ·····	231
附表五 F 分布表·····	232
附表六 相关系数检验表·····	241
附表七 常用正交表·····	242
附表八 两子样秩和检验的临界值表·····	251

第一章 随机事件及其概率

第一节 随机事件

一、随机试验和样本空间

人们在客观世界中所观察到的现象大致可分为两类：一类现象在一定条件下必然发生，如一个标准大气压下，纯水加热到 100°C 会沸腾，异性的电荷之间一定相互吸引，水往低处流等，这类现象是可事前预言的，其结果是确定的，称为确定性现象或必然现象；另一类现象在一定条件下可能发生也可能不发生，如抛掷一枚质地均匀的对称硬币，落下后可能正面朝上，也可能反面朝上，某厂同一工艺同一生产线生产的同一型号的灯泡，其寿命有长有短，公共汽车站每天某一时段的候车人数有多有少等，这类现象在观察之前无法预知它的准确结果，称为随机现象。

随机现象表面上呈现着偶然性，但实质上却存在着内在的必然规律，通过大量的观察和试验去发现和研究随机现象的这些规律性，是概率统计学科的任务。

通常把观察某种现象和进行各种科学试验统称为试验，凡具有下列特性的试验称为随机试验（简称试验）：

- (1) 试验可以在相同的条件下重复进行；
- (2) 试验前能知道所有可能出现的试验结果，并且结果不止一个；
- (3) 每次试验只出现可能结果中的一个，但在试验前不能预知哪一个结果会出现。

随机试验的每一个可能结果称为基本事件，也称为样本点，常用 ω 表示；全体基本事件组成的集合称为样本空间，常用 Ω 表示。

例 1 抛掷一枚硬币，观察出现正、反面的情况（约定某一面为正面），这是一随机试验，试验的可能结果有两个：出现正面和出现反面。用“正”表示出现正面，“反”表示出现反面，其样本空间 $\Omega = \{\text{正}, \text{反}\}$ 。

例 2 将两枚硬币先后各抛一次，观察出现正反面的情况，这是一随机试验，试验的可能结果有四个：(正, 正)，(正, 反)，(反, 正)，(反, 反)，其样本空间 $\Omega = \{(\text{正}, \text{正}), (\text{正}, \text{反}), (\text{反}, \text{正}), (\text{反}, \text{反})\}$ 。

例 3 掷一颗骰子观察出现的点数是一随机试验，试验的可能结果有六个，用“ i ”表示“出现 i 点”($i = 1, 2, 3, 4, 5, 6$)，其样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$ 。

例 4 记录某电话交换机在 24 小时内的电话转接次数是一个随机试验。每一

个非负整数都是一个可能的试验结果,其样本空间 $\Omega = \{0, 1, 2, 3, \dots\}$.

例5 某灯泡厂生产了一批同型号的灯泡,从中任取一只检测其寿命 x (单位:小时) 是一个随机试验,其样本空间 $\Omega = \{x \mid 0 \leq x < +\infty\}$.

二、随机事件

在随机试验中,有时关心的是带有某些特征的事件是否发生. 如在例3中,我们可以研究

$$A = \{\text{出现6点}\},$$

$$B = \{\text{出现偶数点}\},$$

$$C = \{\text{点数大于3}\}$$

这些结果是否发生?其中 A 是一个基本事件,而 B 与 C 则是由多个基本事件所组成,相对于基本事件,它们是复杂事件. 无论是基本事件还是复杂事件,它们在试验中发生与否,却带有随机性,都称为随机事件或简称为事件,通常用大写字母 $A, B, C, \dots$ 表示.

在每次试验中都一定会发生的事件称为必然事件,用 Ω 表示;在每次试验中一定不会发生的事件称为不可能事件,用 $\emptyset$ 表示. 如在例3中,“点数大于0且小于7”是必然事件,而“点数大于8”是不可能事件.

必然事件与不可能事件本质上没有不确定性,但是为了方便,仍把它们看作随机事件.

三、随机事件的关系与运算

在后面的讨论中均假定随机试验的样本空间 Ω 已确定,而且 A, B, C 及 $A_i (i = 1, 2, \dots)$ 等均是该试验中的事件.

1. 事件的包含与相等

如果事件 A 发生必然导致事件 B 发生,则称事件 B 包含事件 A , 或称 A 是 B 的子事件,记为 $A \subset B$ 或 $B \supset A$.

如果事件 A 包含事件 B , 同时事件 B 也包含事件 A , 即 $A \subset B, B \subset A$ 同时成立,则称事件 A 与 B 相等,记为 $A = B$.

对任意事件 $A, A \subset \Omega$ 总成立;作为特殊情形,还规定 $\emptyset \subset A$.

2. 事件的和(并)

“事件 A 与 B 中至少有一个发生”这一事件称为事件 A 与 B 的和(或并),记为 $A \cup B$. 若有事件 $A_1, A_2, \dots, A_n$, 则“事件 $A_1, A_2, \dots, A_n$ 中至少有一个发生”称为事件 $A_1, A_2, \dots, A_n$ 的和(或并),记为

$$A_1 \cup A_2 \cup \dots \cup A_n \text{ 或 } \bigcup_{i=1}^n A_i.$$

“事件 $A_1, A_2, \dots, A_n, \dots$ 中至少有一个发生”称为可列个事件 $A_1, A_2, \dots, A_n, \dots$ 的

和(或并),记为 $A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots$ 或 $\bigcup_{i=1}^{\infty} A_i$.

3. 事件的积(交)

“事件 A 与 B 同时发生”这一事件称为事件 A 与 B 的积(或交),记为 $A \cap B$ 或 AB . “事件 $A_1, A_2, \cdots, A_n$ 同时发生”称为事件 $A_1, A_2, \cdots, A_n$ 的积(或交),记为 $A_1 \cap A_2 \cap \cdots \cap A_n$ 或 $A_1 A_2 \cdots A_n$ 或 $\bigcap_{i=1}^n A_i$. “事件 $A_1, A_2, \cdots, A_n, \cdots$ 同时发生”称为可列个事件 $A_1, A_2, \cdots, A_n, \cdots$ 的积(或交),记为 $A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots$ 或 $A_1 A_2 \cdots A_n \cdots$ 或 $\bigcap_{i=1}^{\infty} A_i$.

4. 互不相容事件与对立事件

若事件 A 与 B 不能同时发生,或者说 $A \cap B$ 是一个不可能事件,即 $A \cap B = \emptyset$,则称事件 A 与 B 互不相容(或互斥).

如果事件 A, B 满足 $A \cup B = \Omega$,且 $A \cap B = \emptyset$,则称事件 B 为 A 的对立事件,记为 $B = \bar{A}$. 当然这时 A 也是 B 的对立事件,即有 $A = \bar{B}$,因此事件 A, B 互为对立事件. 显然有 $\bar{\bar{A}} = A$.

5. 事件的差

“事件 A 发生而 B 不发生”这一事件称为事件 A 与 B 的差,记为 $A - B$. 易知 $A - B = A \cap \bar{B}$.

6. 事件的运算规则

(1) 交换律:

$$A \cup B = B \cup A; \quad A \cap B = B \cap A.$$

(2) 结合律:

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

(3) 分配律:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C);$$

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

(4) 对偶律(德摩根律):

$$\overline{\bigcup_{i=1}^n A_i} = \bigcap_{i=1}^n \bar{A}_i; \quad \overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \bar{A}_i.$$

例6 设 A, B, C 为三个事件,用 A, B, C 的运算关系表示下列各事件:

(1) A 与 B 都发生而 C 不发生;

(2) 恰有一个事件发生;

(3) 至少有一个事件发生;

(4) 恰有两个事件发生;

(5) 至少有两个事件发生.

解 (1) ABC ;

(2) $\bar{A}\bar{B}C \cup \bar{A}B\bar{C} \cup A\bar{B}\bar{C}$;

- (3) $A \cup B \cup C$ 或 $\overline{A}\overline{B}\overline{C} \cup \overline{A}\overline{B}C \cup \overline{A}B\overline{C} \cup \overline{A}BC \cup A\overline{B}\overline{C} \cup A\overline{B}C \cup AB\overline{C} \cup ABC$;
 (4) $\overline{A}\overline{B}\overline{C} \cup \overline{A}\overline{B}C \cup \overline{A}BC$;
 (5) $AB \cup AC \cup BC$ 或 $AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC \cup ABC$.

习 题 1-1

- 一个口袋中有三只球(分别记为 a, b, c), 试写出下列试验的样本空间:
 - 从袋中任取两球;
 - 从袋中任取一球, 取后不放回, 连续取两次;
 - 从袋中任取一球, 取后放回, 连续取两次.
- 将一枚质地均匀的硬币, 连续抛掷三次, 观察正反面出现的情况, 写出这个试验的样本空间, 并用基本事件表示下列事件:
 - $A = \{\text{恰有一次出现正面}\}$;
 - $B = \{\text{正面出现两次}\}$.
- 同时掷两枚骰子, 记录两枚骰子点数之和,
 - 求样本空间 Ω ;
 - 设事件 A 表示“点数之和为偶数”, 事件 B 表示“点数之和大于 7”, 事件 C 表示“点数之和为小于 5 的偶数”, 试用样本空间 Ω 的子事件表示下列事件:
 $\overline{B}, \overline{C}, A \cup B, A - B, A \cup \overline{B} \cup \overline{C}, ABC$.
- 设 A, B, C 是三个事件, 试用 A, B, C 的运算关系表示下列事件:
 - A 发生而 B 与 C 不发生;
 - A, B, C 中恰有一个不发生;
 - A, B, C 中至少有两个不发生.

第二节 随机事件的概率

随机事件在每次试验中可能发生也可能不发生, 但在相同条件下的大量重复试验中, 人们发现它具有一定的规律性. 刻画事件发生可能性大小的数量指标叫做事件的概率, 事件 A 的概率用 $P(A)$ 表示.

一、古典概型

先讨论一类最简单的随机试验, 它具有如下两个特征:

- 试验的可能结果的个数有限, 即只有有限个基本事件;
- 每个基本事件出现的可能性是相等的.

具有这两个特征的试验称为古典概型试验, 其数学模型称为古典概型. 例如, 抛掷一枚质地均匀的硬币, 观察出现正反面的试验中, 试验结果只有两个, 即出现正面和出现反面. 由于硬币质地均匀, 出现正面和出现反面的可能性相同, 故这个

试验属于古典概型.

对古典概型,事件 A 的概率定义如下:

定义 1 设有一古典概型试验,它的样本空间 Ω 包含 n 个基本事件,并且各个基本事件发生的可能性相等.若事件 A 包含 k 个基本事件,则事件 A 的概率 $P(A)$ 定义为

$$P(A) = \frac{\text{事件 } A \text{ 包含的基本事件数}}{\text{基本事件总数}} = \frac{k}{n}. \quad (1)$$

从定义可以得到下列性质:

(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.设事件 $A_1, A_2, \dots, A_m$ ($m \leq n$) 两两互不相容,则

$$P\left(\bigcup_{i=1}^m A_i\right) = \sum_{i=1}^m P(A_i). \quad (2)$$

证 (1) 因为任一事件 A 所含的基本事件数 k 总满足 $0 \leq k \leq n$, 则有

$$0 \leq \frac{k}{n} \leq 1,$$

即

$$0 \leq P(A) \leq 1.$$

(2) 由于必然事件由全部 n 个基本事件所组成,即必然事件 Ω 所包含的基本事件数为 n , 则

$$P(\Omega) = \frac{n}{n} = 1.$$

(3) 设 $A_i = \{w_1^{(i)}, \dots, w_{k_i}^{(i)}\}$ 是由 k_i ($i \leq n$) ($i = 1, 2, \dots, m$) 个不同的基本事件所组成, 则 $P(A_i) = \frac{k_i}{n}$. 由于 A_i 互不相容, 则

$$\bigcup_{i=1}^m A_i = \{w_1^{(1)}, \dots, w_{k_1}^{(1)}, \dots, w_1^{(m)}, \dots, w_{k_m}^{(m)}\}$$

包含 $\sum_{i=1}^m k_i$ 个不同的基本事件, 故

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

例 1 设有一批产品共 100 件, 其中 5 件次品, 现从中任取 50 件, 问: 无次品的概率是多少?

解 首先从 100 件产品中任取 50 件, 则有 C_{100}^{50} 个不同的结果, 每一个结果就是一个基本事件, 设 A 表示“任取 50 件无次品的产品”, 这 50 件必须是从那 95 件正品中取来的, 可见这种无次品的取法共有 C_{95}^{50} 种 (即事件 A 包含 C_{95}^{50} 个基本事件), 则

$$P(A) = \frac{C_{95}^{50}}{C_{100}^{50}} \approx 0.028.$$

例 2 袋中有 n 只球,一只红的, $n-1$ 只白的, n 个人依次从袋中任取一球,每人取球后不再放回袋中,求第 k 个人($1 \leq k \leq n$) 取出的球是红球的概率.

解 设 A 表示“第 k 个人取出的球为红球”,把白球看作是互不相同的个体(例如,给它们进行编号:1,2, $\dots$, $n-1$).若把各人取出的球依次放在排列成直线的 n 个位置上,则可能的排列方式相当于把 n 个元素进行全排列,故基本事件总数为 $n!$.

当第 k 个人取得红球,考虑在直线上第 k 个位置放红球,其他 $n-1$ 个位置放白球,这时 $n-1$ 个白球可以任意地放在其他 $n-1$ 个位置上,它对应 $n-1$ 次取球,故除第 k 次取球外,其他 $n-1$ 次取球相当 $n-1$ 个元素进行全排列,故事件 A 包含的基本事件数为 $(n-1)!$,故

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

可见,所求概率与 k 无关.它说明无论取球的先后次序如何,取得红球的概率都相等.这与生活中的经验是一致的,例如对不放回的抽签,并不是先抽签的人中签的可能性就大些.

例 3 设有 n 个人,每个人都等可能地被分配到 N 个房间中的任意一间去住($n \leq N$),求下列事件的概率:

(1) 指定的 n 个房间各有一个入住;

(2) 恰好有 n 个房间,其中各住一个人.

解 因为每一个人有 N 个房间可供选择,所以 n 个人住的方式共有 N^n 种,它们是等可能的.

(1) 指定的 n 个房间各有一个入住,其可能总数为 n 个人的全排列 $n!$,于是

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

(2) n 个房间可以在 N 个房间中任意选取,其总数有 C_N^n 个,对选定的 n 个房间有 $n!$ 种分配方式,所以恰有 n 个房间其中各住一个人的概率为

$$P_2 = \frac{C_N^n n!}{N^n} = \frac{N!}{N^n (N-n)!}.$$

二、随机事件的频率

古典概型是以等可能性为基础的,但一般的随机试验不都具有等可能性,对于任意随机试验中的随机事件,表示它发生可能性大小的数值是否存在呢?经验表明,通过大量重复的试验,上述问题的回答是肯定的.为此,我们引进事件频率的概念.

定义 2 设 A 是一随机试验的事件,如果在 n 次重复试验中,事件 A 发生 m 次,则称比值 $\frac{m}{n}$ 为这 n 次重复试验中事件 A 发生的频率,记为 $f_n(A)$,即

$$f_n(A) = \frac{m}{n}, \quad (3)$$

其中 m 称为事件 A 在这 n 次重复试验中发生的频数.

历史上不少人为研究事件的概率曾做过许多抛掷硬币的试验,并统计了试验中出现正面的次数和频率,其结果如下:

实验者	抛币次数 n	正面出现次数 m	频率 $\frac{m}{n}$
德摩根(De Morgan)	2048	1061	0.5181
蒲丰(Buffon)	4040	2048	0.5069
费勒(Feller)	10000	4979	0.4979
皮尔逊(Pearson)	24000	12012	0.5005

实验表明,当试验大量重复时,事件 A 发生的频率具有一种稳定性,即当试验次数充分大时, A 发生的频率会集中在某个常数附近摆动.一般地,随着试验次数增多,这种摆动的幅度越来越小.例如上述抛掷硬币的试验,正面出现的频率便集中在 0.5 附近摆动.

频率的稳定性和稳定值是事件本身固有的一种自然属性,这种属性便成为可以对事件发生的可能性大小进行度量的客观基础.

将任一随机试验重复进行 n 次,设事件 A 发生的次数为 m ,当 n 充分大时,就可以用频率 $\frac{m}{n}$ 作为 $P(A)$ 的近似值.

三、概率的公理化定义

因古典概型适用范围有限,而用频率去确定概率,也存在明显的缺点,故对一般随机现象寻求概率和其他基本概念的确切定义,在历史上一度成为一个突出的问题,直到 1933 年前苏联数学家柯尔莫哥洛夫(Kolmogorov)提出了概率的公理化定义,才使概率论有了快速的发展.

下面给出概率的公理化定义:

公理 1(非负性) 对于任一随机事件 A ,有

$$0 \leq P(A) \leq 1.$$

公理 2(规范性) $P(\Omega) = 1$.

公理 3(完全可加性) 对于两两互不相容的随机事件 $A_1, A_2, \dots, A_n, \dots (A_i A_j = \emptyset, i \neq j)$ 有

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

定义 3 设随机试验的样本空间为 Ω . 如果对于每一事件 A 有唯一实数 $P(A)$ 与之对应,且满足公理 1,2,3,则称 $P(A)$ 为事件 A 的概率.

由定义 3 可以推出概率的一些性质.

性质1 不可能事件的概率为0,即 $P(\emptyset) = 0$. (4)

证 因为 $\Omega = \Omega \cup \emptyset \cup \emptyset \cup \dots$,

即有 $P(\Omega) = P(\Omega) + P(\emptyset) + P(\emptyset) + \dots$,

所以 $P(\emptyset) = 0$.

性质2(有限可加性) 设事件 $A_1, A_2, \dots, A_n$ 满足 $A_i A_j = \emptyset$ ($i \neq j; i, j = 1, 2, \dots, n$), 则

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i). \quad (5)$$

证 因为

$$\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup \dots \cup A_n \cup \emptyset \cup \emptyset \cup \dots,$$

由完全可加性及 $P(\emptyset) = 0$ 知

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

性质3 对任一事件 A , 有

$$P(\bar{A}) = 1 - P(A). \quad (6)$$

证 由 $A \cup \bar{A} = \Omega, A\bar{A} = \emptyset$ 得

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

则

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

性质4 $P(A - B) = P(A) - P(AB)$. (7)

证 因

$$A = A\Omega = A(B \cup \bar{B}) = AB \cup A\bar{B},$$

又

$$AB \cap A\bar{B} = \emptyset,$$

故

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

而

$$P(A - B) = P(A\bar{B}),$$

于是

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

特别地, 若 $A \supset B$, 则 $P(A - B) = P(A) - P(B)$, 并且易知若 $A \supset B$, 则 $P(A) \geq P(B)$.

例4 设有一批产品共100件, 其中5件是次品, 任取3件, 求其中至少有一件是次品的概率.

解 方法一. 设 A 表示“任取3件至少有一件次品”, B_i 表示“任取3件恰好有 i 件次品” ($i = 1, 2, 3$), 则 $A = B_1 \cup B_2 \cup B_3$, 且 B_1, B_2, B_3 两两互不相容. 于是由性质2得

$$\begin{aligned} P(A) &= P(B_1 \cup B_2 \cup B_3) = P(B_1) + P(B_2) + P(B_3) \\ &= \frac{C_5^1 C_{95}^2}{C_{100}^3} + \frac{C_5^2 C_{95}^1}{C_{100}^3} + \frac{C_5^3}{C_{100}^3} \approx 0.1440. \end{aligned}$$

方法二. 显然有 $\bar{A}$ 为“任取3件全是正品”, 而