



普通高等教育“十一五”国家级规划教材

P 概率论

robability theory

M 与数理统计

and M athematical statistics

第二版

王明慈 沈恒范 主编



高等教育出版社
Higher Education Press

021

267

2007

普通高等教育“十一五”国家级规划教材

概率论与数理统计

第二版

王明慈 沈恒范 主编

高等教育出版社

内 容 提 要

本书第一版是按工科院校概率论与数理统计课程第Ⅱ类(概率少、统计多)教学基本要求编写的,第二版参照新修订的概率论与数理统计课程教学基本要求进行修订,但仍保留了“概率少、统计多”的特色。前4章是概率论基本内容,为数理统计准备必要的理论基础;后5章是在概率论基础上侧重分析介绍如何用统计方法分析、解决带有随机性的实际问题。两部分内容配合紧密。每章末的综合例题,是全面运用该章理论与方法解决问题的范例。编写特点:全书讲解清楚,文字通顺;内容安排重点突出,难点分散,由浅入深,便于接受;对于用统计方法对随机变量的概率特征作出科学推断的基本思想、推断方法分析透彻,归纳总结方法条理清楚。本书可作为工科院校本科各专业的教材或教学参考书。

图书在版编目(CIP)数据

概率论与数理统计/王明慈,沈恒范主编.—2版.
—北京:高等教育出版社,2007.4
ISBN 978-7-04-020741-5

I. 概… II. ①王…②沈… 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024118号

策划编辑 李 蕊 责任编辑 崔梅萍 封面设计 张 楠 责任绘图 吴文信
版式设计 余 杨 责任校对 王 超 责任印制 朱学忠

出版发行 高等教育出版社
社 址 北京市西城区德外大街4号
邮政编码 100011
总 机 010-58581000

经 销 蓝色畅想图书发行有限公司
印 刷 山东鸿杰印务有限公司

开 本 787×960 1/16
印 张 19
字 数 350 000

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landraco.com>
<http://www.landraco.com.cn>
畅想教育 <http://www.widedu.com>

版 次 1999年6月第1版
2007年4月第2版
印 次 2007年4月第1次印刷
定 价 20.10元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究
物料号 20741-00

第二版序言

本书第一版于1999年6月出版,是当时普通高等学校概率论与数理统计课程唯一一本按照本课程教学基本要求(概率少、统计多)编写的教材,因而受到高等学校教师和学生的好评。

本书主要参照教育部非数学类专业数学基础课程教学指导分委会2005年9月修订的概率论与数理统计课程教学基本要求对第一版进行了修订,增减了部分内容,更换了若干例题与习题。同时,为了普通高等学校本科毕业生报考硕士学位研究生的需要,修订时也参考了教育部考试中心拟定的硕士学位研究生入学考试数学课程的考试大纲。

本书由王明慈和沈恒范主编,参加修订、编写工作的作者、所在学校及修订编写内容如下:

陈鸿建(四川大学)修订第一、四章;

王明慈(四川大学)修订第二、三章;

王丽霞(北京邮电大学)修订第五、六章;

沈恒范(湖北汽车工业学院)修订第七、八章;

马秀兰(天津商学院)修订第一版各章中原有的综合例题;

沈侠(中国人民大学)修订第九章,并为各章增补若干综合例题。

限于作者的水平,修订后的本书难免还存在某些缺点和错误,真诚希望读者批评指正。

王明慈 沈恒范

2006年8月

第一版序言

遵照原国家教育委员会的指示,高等学校工科数学课程教学指导委员会(以下简称工科数学课委会)从1992年至1993年修订了工科数学各门课程的教学基本要求。修订后的《概率论与数理统计课程教学基本要求》分为两种类型:第Ⅰ类型侧重于概率论,即“概率多、统计少”;第Ⅱ类型侧重于数理统计,即“概率少、统计多”。1995年起,我们联合编写了适用于本课程第Ⅱ类型教学基本要求的试用教材,并在各自所在的学校中进行试点讲授。在试点教学过程中曾经多次修改,形成本书的初稿。1997年6月,陶永德教授和齐植兰教授将本书初稿推荐给工科数学课委会。

1997年10月,工科数学课委会聘请汪国强教授、吴翊教授、陶宗英教授、范金城教授、苏化明教授等专家组成教材评审组,对本书初稿进行评审,提出了不少宝贵的意见和有益的建议。根据这些意见和建议,我们对本书初稿作了较多的修改和补充。以后,齐植兰教授对本书进行复审,提出了若干改进的意见,我们再次修改、定稿。以上各位专家、教授的意见和建议对于提高本书在科学性和教学适用性方面的质量都起到了重要的作用。为此,我们并代表本书全体参编者向他们致以诚挚的谢意!

本书内容完全符合《概率论与数理统计课程(第Ⅱ类型)教学基本要求》,我们的教学实践表明,讲授本书全部内容约需50学时,其中概率论部分22学时,数理统计部分28学时。本书有较多的例题和习题,每章有综合例题分析及解答,书末附有习题答案,便于自学本书的读者参考。

本书由王明慈和沈恒范主编,参加编写的作者、所在学校及编写内容如下:

陈鸿建(四川大学)编写第一、四章,

王明慈(四川大学)编写第二、三章,

王丽霞(湖北汽车工业学院)编写第五、六章,

沈恒范(湖北汽车工业学院)编写第七、八、九章,

马秀兰(天津商学院)编写各章的综合例题。

本书编写过程中,曾经得到四川大学、湖北汽车工业学院及天津商学院有关领导同志的支持和帮助,我们也向上述各校的领导同志们致以诚挚的谢意!

限于作者的水平,本书难免还存在若干缺点和错误,恳请读者批评指正。

王明慈 沈恒范

1998年5月

目 录

第一章 随机事件及其概率	1
§ 1.1 样本空间 随机事件	1
§ 1.2 随机事件的频率与概率的定义及性质	7
§ 1.3 古典概型	10
§ 1.4 条件概率 概率乘法公式	14
§ 1.5 随机事件的独立性	18
§ 1.6 伯努利概型	20
§ 1.7 综合例题	22
习题一	26
第二章 随机变量及其分布	30
§ 2.1 随机变量的概念	30
§ 2.2 离散随机变量	31
§ 2.3 超几何分布 二项分布 泊松分布	32
§ 2.4 连续随机变量	38
§ 2.5 均匀分布 指数分布	40
§ 2.6 随机变量的分布函数	41
§ 2.7 多维随机变量及其分布	44
§ 2.8 随机变量的独立性	50
§ 2.9 随机变量函数的分布	52
§ 2.10 综合例题	56
习题二	63
第三章 随机变量的数字特征	67
§ 3.1 数学期望	67
§ 3.2 方差	74
§ 3.3 原点矩与中心矩	78
§ 3.4 协方差与相关系数	79
§ 3.5 切比雪夫不等式与大数定律	83
§ 3.6 综合例题	86
习题三	92
第四章 正态分布	95
§ 4.1 正态分布的概率密度与分布函数	95

§ 4.2	正态分布的数字特征	99
§ 4.3	正态随机变量的线性函数的分布	101
§ 4.4	二维正态分布	103
§ 4.5	中心极限定理	107
§ 4.6	综合例题	109
习题四	116
附表	常用分布及其数学期望与方差	118
第五章	数理统计的基本知识	119
§ 5.1	总体与样本	119
§ 5.2	样本分布函数 直方图	121
§ 5.3	样本函数与统计量	125
§ 5.4	χ^2 分布 t 分布 F 分布	128
§ 5.5	正态总体统计量的分布	132
§ 5.6	综合例题	140
习题五	145
第六章	参数估计	149
§ 6.1	参数的点估计	149
§ 6.2	判别估计量好坏的标准	155
§ 6.3	正态总体参数的区间估计	158
§ 6.4	两个正态总体均值差与方差比的区间估计	163
§ 6.5	非正态总体参数的区间估计举例	168
§ 6.6	单侧置信限	170
§ 6.7	综合例题	171
习题六	180
第七章	假设检验	184
§ 7.1	假设检验的基本概念	184
§ 7.2	单个正态总体参数的假设检验	188
§ 7.3	两个正态总体参数的假设检验	191
§ 7.4	非正态总体参数的假设检验举例	195
§ 7.5	总体分布的拟合检验	197
§ 7.6	综合例题	202
习题七	209
第八章	方差分析	212
§ 8.1	单因素试验的方差分析	212
§ 8.2	双因素无重复试验的方差分析	217

§ 8.3 双因素等重复试验的方差分析	223
§ 8.4 综合例题	228
习题八	234
第九章 回归分析	237
§ 9.1 回归分析的基本概念	237
§ 9.2 线性回归方程	239
§ 9.3 线性相关的显著性检验	241
§ 9.4 利用线性回归方程预测和控制	246
§ 9.5 非线性回归分析	248
§ 9.6 多元线性回归分析	255
§ 9.7 综合例题	260
习题九	268
习题答案	271
附录	283
表 1 函数 $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ 数值表	283
表 2 满足等式 $P(\chi^2 \geq \chi_\alpha^2(k)) = \alpha$ 的 $\chi_\alpha^2(k)$ 数值表	285
表 3 满足等式 $P(t \geq t_\alpha(k)) = \alpha$ 的 $t_\alpha(k)$ 数值表	286
表 4 满足等式 $P(F \geq F_\alpha(k_1, k_2)) = \alpha$ 的 $F_\alpha(k_1, k_2)$ 数值表	287
表 5 满足等式 $P(r \geq r_\alpha(n-2)) = \alpha$ 的 $r_\alpha(n-2)$ 数值表	295

第一章 随机事件及其概率

§ 1.1 样本空间 随机事件

1. 随机试验与随机事件

在自然界,在人们的实践活动中,所遇到的现象一般可以分为两类:

一类现象是,在一定的条件下,必然会出现某种确定的结果.例如,向上抛一枚硬币,由于受到地心引力的作用,硬币上升到某一高度后必定会下落.我们把这类现象称为确定性现象(或必然现象).同样,任何物体没有受到外力作用时,必定保持其原有的静止或等速运动状态;导线通电后,必定会发热;等等也都是确定性现象.

另一类现象则是,在一定的条件下,可能会出现各种不同的结果,也就是说,在完全相同的条件下,进行一系列观测或实验,却未必出现相同的结果.例如,抛掷一枚硬币,当硬币落在地面上时,可能是正面(有国徽的一面)朝上,也可能是反面朝上,在硬币落地前我们不能预知究竟哪一面朝上.我们把这类现象称为随机现象(或偶然现象).同样,自动机床加工制造一个零件,可能是合格品,也可能是不合格品;射击运动员一次射击,可能击中10环,也可能击中9环、8环……甚至脱靶;等等也都是随机现象.

随机现象,从表面上看,由于人们事先不能知道会出现哪一种结果,似乎是不可捉摸的,其实不然.人们通过实践观察到并且证明了,在相同的条件下,对随机现象进行大量的重复试验(观测),其结果总能呈现出某种规律性.例如,多次重复抛一枚硬币,正面朝上和反面朝上的次数几乎相等;对某个靶进行多次射击,虽然各次弹着点不完全相同,但这些点却按一定的规律分布;等等.我们把随机现象的这种规律性称为统计规律性.

正是由于大量重复试验中随机事件的统计规律性是客观存在的,才使得数学家和统计学家对各种随机现象进行深入的研究,并取得了极其丰富的重要成果,形成了若干门研究随机现象的学科,概率论与数理统计只是其中的一门学科,它在自然科学、社会科学、工程技术、国家经济建设等各方面有着广泛的应用.

为了研究随机现象的统计规律性,我们把各种科学试验和对某一事物的观测统称为试验.如果试验具有下述特点:

- (1) 试验可以在相同条件下重复进行;
- (2) 每次试验的所有可能结果都是明确可知的,并且不止一个;
- (3) 每次试验之前不能预知将会出现哪一个结果,

则称这种试验为随机试验(简称试验). 通常用字母 E, E_1, E_2, \dots 表示随机试验.

例 1 试验 E_1 : 抛一枚硬币, 观察出现的结果, 可能是正面朝上或反面朝上.

例 2 试验 E_2 : 从一批产品中任意取 10 个样品, 观测其中的次品数, 可能是 $0, 1, 2, \dots, 10$.

例 3 试验 E_3 : 记录某段时间内电话交换台接到的呼唤次数, 可能是 $0, 1, 2, \dots$.

例 4 试验 E_4 : 测量某个零件的尺寸与规定尺寸的偏差 x (mm), 所有可能的结果是 x 在某个区间 (a, b) 内, 即 $a < x < b$.

我们把试验的结果中发生的现象称为事件. 在每次试验的结果中, 如果某事件一定发生, 则称为必然事件; 相反, 如果某事件一定不发生, 则称为不可能事件.

在试验的结果中, 可能发生、也可能不发生的事件称为随机事件. 通常用字母 A, B, C, \dots 表示随机事件. 我们以上述各项试验为例, 考虑某些随机事件:

例 5 在试验 E_1 中, A_1 ——“正面朝上”, A_2 ——“反面朝上”, 都是随机事件.

例 6 在试验 E_2 中, B ——“取出的 10 个样品中有 1 至 3 个次品”也是随机事件.

例 7 在试验 E_3 中, C ——“在该段时间内电话交换台接到的呼唤次数不超过 8 次”也是随机事件.

例 8 在试验 E_4 中, D ——“测得零件的尺寸与规定尺寸的偏差小于 0.1 mm”(即 $|x| < 0.1$) 也是随机事件(这里假定 $a < -0.1 < 0.1 < b$).

2. 样本空间

为了更好地研究随机试验与随机事件, 我们引进随机试验的样本空间的概念.

随机试验的每一个可能的结果称为样本点, 记作 $\omega_1, \omega_2, \dots$; 随机试验的所有样本点组成的集合称为样本空间, 记作 Ω , 即

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

在例 1 中, 我们有样本空间 $\Omega_1 = \{\omega_1, \omega_2\}$, 其中样本点 ω_1 表示“正面朝上”, ω_2 表示“反面朝上”.

在例 2 中, 我们有样本空间 $\Omega_2 = \{\omega_0, \omega_1, \omega_2, \dots, \omega_{10}\}$, 其中样本点 ω_i 表示“取出的 10 个样品中有 i 个次品”($i = 0, 1, 2, \dots, 10$).

在例 3 中, 我们有样本空间 $\Omega_3 = \{\omega_0, \omega_1, \omega_2, \dots\}$, 其中样本点 ω_i 表示“在该

段时间内电话交换台接到 i 次呼唤” ($i=0,1,2,\dots$).

在例 4 中,我们有样本空间 $\Omega_4 = \{\omega_x | a < x < b\}$, 其中样本点 ω_x 表示“测得该零件的尺寸与规定尺寸的偏差为 x (mm)” ($a < x < b$).

我们指出,任一随机事件 A 都是样本空间 Ω 中的一部分样本点组成的集合;也就是说,任一随机事件 A 都是样本空间 Ω 的一个子集. 在每次试验的所有可能结果中,显然有且仅有一个样本点 ω 发生,如果这个样本点 ω 属于随机事件 A ,那么随机事件 A 也就发生了;反之,如果这个样本点 ω 不属于随机事件 A ,那么随机事件 A 就不发生.

在例 5 中,我们有随机事件 $A_1 = \{\omega_1\}$, $A_2 = \{\omega_2\}$, 显然 A_1, A_2 都是样本空间 Ω_1 的子集.

在例 6 中,我们考虑的随机事件是 $B = \{\omega_1, \omega_2, \omega_3\}$, 这里 B 是样本空间 Ω_2 的子集.

在例 7 中,我们考虑的随机事件是 $C = \{\omega_0, \omega_1, \dots, \omega_8\}$, 这里 C 是样本空间 Ω_3 的子集.

在例 8 中,我们考虑的随机事件是 $D = \{\omega_x | |x| < 0.1\}$, 这里 D 是样本空间 Ω_4 的子集.

因为必然事件在每次试验中一定发生,这就表明,在试验的结果中任一样本点发生时必然事件都发生,所以必然事件是所有样本点组成的集合. 由此可见,必然事件就等于样本空间. 今后我们就把必然事件记作 Ω .

因为不可能事件在每次试验中一定不发生,这就表明,在试验的结果中任一样本点发生时不可能事件都不发生,所以不可能事件是不包含任何样本点的空集. 今后我们就把不可能事件记作 \emptyset .

还应指出,试验的任一样本点 ω 也是随机事件,今后我们把试验的样本点称为试验的基本事件. 显然,基本事件就是仅包含单个样本点的子集.

最后我们强调指出,随机事件的样本点与样本空间是由试验的目的决定的. 例如,设随机试验是将一枚硬币连抛两次,如果观察正面或反面朝上的情况,并用数字 1 表示“正面朝上”,数字 0 表示“反面朝上”,样本点 ω_{ij} 表示“第一次结果为数字 i ($i=0$ 或 1),第二次结果为数字 j ($j=0$ 或 1)”,则样本空间是

$$\Omega' = \{\omega_{00}, \omega_{01}, \omega_{10}, \omega_{11}\};$$

如果我们观察正面朝上的次数,则样本空间是

$$\Omega'' = \{\omega_0, \omega_1, \omega_2\},$$

其中样本点 ω_i 表示正面朝上的次数为 i ($i=0,1,2$).

3. 事件的关系及运算

我们已知,任一事件 A 是样本空间 Ω 的子集,为了叙述简便,不妨就把这个

子集称为集合 A . 于是,事件的关系及运算与集合的关系及运算是完全类似的. 事件的关系及运算对于研究复杂的事件是十分重要的.

(1) 事件的包含

若事件 A 发生必然导致事件 B 发生,则称事件 B 包含事件 A ,或称事件 A 包含于事件 B ,记作

$$B \supset A \quad \text{或} \quad A \subset B.$$

这时,集合 A 中的样本点一定属于集合 B ,但集合 B 中的样本点不一定属于集合 A .

(2) 事件的相等

若事件 A 包含事件 B ,且事件 B 也包含事件 A ,即

$$A \supset B \quad \text{且} \quad B \supset A,$$

则称事件 A 与事件 B 相等,记作

$$A = B.$$

这时,集合 A 与集合 B 中的样本点是相同的.

(3) 事件的并

“两个事件 A 与 B 中至少有一个事件发生”这一新事件称为事件 A 与 B 的并,记作

$$A \cup B.$$

显然,这个事件是由两个集合 A 与 B 中所有的样本点组成的集合(两个集合共同的样本点不重复选取).

事件的并的概念可以推广到有限个或可列无穷多个事件的情形:

“ n 个事件 A_1, A_2, \dots, A_n 中至少有一个事件发生”这一新事件称为这 n 个事件的并,记作

$$A_1 \cup A_2 \cup \dots \cup A_n \quad (\text{简记为 } \bigcup_{i=1}^n A_i).$$

“可列无穷多个事件 $A_1, A_2, \dots, A_n, \dots$ 中至少有一个事件发生”这一新事件称为这些事件的并,记作 $\bigcup_{i=1}^{\infty} A_i$.

(4) 事件的交

“两个事件 A 与 B 同时发生”这一新事件称为事件 A 与 B 的交,记作

$$A \cap B \quad \text{或} \quad AB.$$

显然,这个事件是由两个集合 A 与 B 中共同的样本点组成的集合.

事件的交的概念可以推广到有限个或可列无穷多个事件的情形:

“ n 个事件 A_1, A_2, \dots, A_n 同时发生”这一新事件称为这 n 个事件的交,记作

$$A_1 \cap A_2 \cap \dots \cap A_n \quad \text{或} \quad A_1 A_2 \dots A_n \quad (\text{简记为 } \bigcap_{i=1}^n A_i).$$

“可列无穷多个事件 $A_1, A_2, \dots, A_n, \dots$ 同时发生”这一新事件称为这些事件的交, 记作 $\bigcap_{i=1}^{\infty} A_i$.

(5) 互不相容事件

若两个事件 A 与 B 不可能同时发生, 即

$$AB = \emptyset,$$

则称事件 A 与 B 是互不相容的(或互斥的). 这时, 两个集合 A 与 B 没有共同的样本点.

若 n 个事件 A_1, A_2, \dots, A_n 中任意两个事件不可能同时发生, 即

$$A_i A_j = \emptyset \quad (1 \leq i < j \leq n),$$

则称这 n 个事件是互不相容的(或互斥的).

若可列无穷多个事件 $A_1, A_2, \dots, A_n, \dots$ 中任意两个事件不可能同时发生, 即

$$A_i A_j = \emptyset \quad (1 \leq i < j < \dots),$$

则称这些事件是互不相容的(或互斥的).

今后我们把两个互不相容事件 A 与 B 的并记作

$$A + B.$$

把 n 个互不相容事件 A_1, A_2, \dots, A_n 的并记作

$$A_1 + A_2 + \dots + A_n \quad (\text{简记为 } \sum_{i=1}^n A_i).$$

把可列无穷多个互不相容事件 $A_1, A_2, \dots, A_n, \dots$ 的并记作

$$A_1 + A_2 + \dots + A_n + \dots \quad (\text{简记为 } \sum_{i=1}^{\infty} A_i).$$

(6) 对立事件

若两个互不相容事件 A 与 B 中必有一个事件发生, 即

$$AB = \emptyset \quad \text{且} \quad A + B = \Omega,$$

则称事件 A 与 B 是对立的, 也称事件 B 是事件 A 的对立事件(或逆事件); 同样, 事件 A 也是事件 B 的对立事件, 记作

$$B = \bar{A} \quad \text{或} \quad A = \bar{B}.$$

于是, 我们有

$$\bar{\bar{A}} = A,$$

$$A\bar{A} = \emptyset,$$

$$A + \bar{A} = \Omega.$$

若用平面上某个矩形区域表示样本空间 Ω , 矩形区域内的点表示样本点, 则上述事件的关系及运算可以用集合图形直观地表示出来(图 1.1).

与集合运算的性质类似, 事件的运算具有下面的性质:

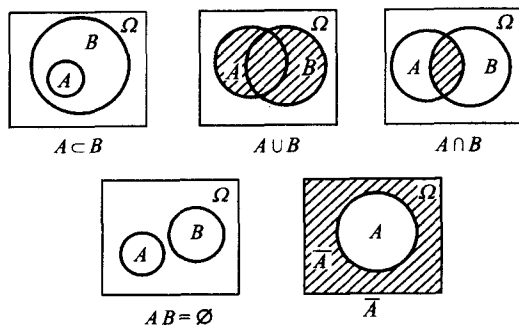


图 1.1

- (1) 交换律: $A \cup B = B \cup A,$
 $AB = BA.$
- (2) 结合律: $(A \cup B) \cup C = A \cup (B \cup C),$
 $(AB)C = A(BC).$
- (3) 分配律: $A(B \cup C) = AB \cup AC,$
 $A \cup (BC) = (A \cup B)(A \cup C).$
- (4) 德摩根(De Morgan)定律:

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

$$\overline{AB} = \bar{A} \cup \bar{B}.$$

对于 n 个事件,有

$$\overline{\bigcup_{i=1}^n A_i} = \prod_{i=1}^n \bar{A}_i,$$

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

例 9 如图 1.2 所示的电路中,设事件 A, B, C 分别表示继电器接点 a, b, c 闭合,事件 D 表示指示灯亮,则因为当且仅当接点 a 闭合,且接点 b 与 c 中至少有一个闭合时,指示灯亮,所以有

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

反之,当且仅当“接点 a 未闭合”与“接点 b 及 c 都未闭合”这两个事件中至少有一个事件发生时,指示灯不亮.所以有

$$\bar{D} = \bar{A} \cup (\bar{B}\bar{C}).$$

显然,这个等式也可以由等式 $D = A(B \cup C)$ 利用上述的德摩根定律得到,事实上,我们有

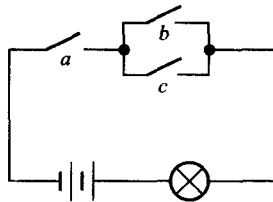


图 1.2

$$\bar{D} = \overline{A(B \cup C)} = \bar{A} \cup (\overline{B \cup C}) = \bar{A} \cup (\bar{B} \bar{C}).$$

§ 1.2 随机事件的频率与概率的定义及性质

1. 随机事件的频率 概率的统计定义

用数字表现大量重复试验中随机现象的统计规律性时,需要用到频率的概念.

定义 1 设随机事件 A 在 n 次试验中发生了 n_A 次,则比值 $\frac{n_A}{n}$ 称为随机事件 A 的频率,记作 $f_n(A)$,即

$$f_n(A) = \frac{n_A}{n}. \quad (1.1)$$

实践证明:在大量重复试验中,随机事件的频率具有稳定性.也就是说,在不同的试验序列中,当试验次数 n 充分大时,随机事件 A 的频率 $f_n(A)$ 常在某个确定的数字附近摆动.

例如,一些著名的统计学家进行过抛硬币的试验,得到如下的结果:

试验者	抛硬币次数 n	正面朝上次数 n_A	频率 $f_n(A)$
Buffon	4 040	2 048	0.506 9
Fisher	10 000	4 979	0.497 9
Pearson	12 000	6 019	0.501 6
Pearson	24 000	12 012	0.500 5

这些结果表明,在抛硬币的试验中,“正面朝上”这一随机事件 A 的频率 $f_n(A)$ 稳定在数字 0.5 的附近.类似的例子还可以举出很多.

由公式(1.1)可知,任一随机事件 A 的频率 $f_n(A)$ 具有下列性质:

(1) 非负性: $f_n(A) \geq 0$.

这是因为,在 n 次试验中随机事件 A 发生的次数 $n_A \geq 0$,所以

$$f_n(A) = \frac{n_A}{n} \geq 0.$$

(2) 规范性: $f_n(\Omega) = 1$.

这是因为,在 n 次试验中事件 Ω 发生的次数 $n_\Omega = n$,所以

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

(3) 有限可加性: 设 l 个事件 A_1, A_2, \dots, A_l 是互不相容的, 则

$$f_n\left(\sum_{i=1}^l A_i\right) = \sum_{i=1}^l f_n(A_i).$$

这是因为, 若在 n 次试验中 l 个互不相容事件 A_1, A_2, \dots, A_l 分别发生了 n_1, n_2, \dots, n_l 次, 则它们的并 $\sum_{i=1}^l A_i$ 就应恰好发生 $\sum_{i=1}^l n_i$ 次, 由此得

$$f_n\left(\sum_{i=1}^l A_i\right) = \frac{1}{n} \sum_{i=1}^l n_i = \sum_{i=1}^l \frac{n_i}{n} = \sum_{i=1}^l f_n(A_i).$$

设在大量重复试验中随机事件 A 的频率 $f_n(A)$ 稳定在某个数字 p ($0 < p < 1$) 的附近. 如果这个数字 p 较大, 则事件 A 的频率 $f_n(A)$ 也相应较大, 这就是说, 事件 A 在 n 次试验中发生的次数较多, 从而表明事件 A 发生的可能性较大; 相反, 如果这个数字 p 较小, 则表明事件 A 发生的可能性较小. 这样一个表明随机事件 A 在试验中发生的可能性大小的数字 p 就是随机事件 A 的概率, 记作 $P(A)$, 即

$$P(A) = p.$$

当试验次数 n 充分大时, 随机事件 A 的频率 $f_n(A)$ 将在概率 $P(A)$ 的附近摆动, 所以有

$$P(A) \approx f_n(A). \quad (1.2)$$

随机事件的概率的这个定义通常称为概率的统计定义.

根据概率的统计定义, 虽然不可能确定任何随机事件的概率, 但是却提供了一种估计概率的方法. 例如: 在人口的抽样调查中, 根据随机抽取的一部分人口的资料可以估计全国人口受教育程度的各种百分比; 在工业生产中, 根据随机抽取的一部分产品的检验结果可以估计整批产品的次品率; 在医学方面, 根据临床试验积累的资料可以估计某种新药对某种疾病的有效率; 等等.

2. 概率的公理化定义

为了使概率论有严谨的数学理论基础, 科尔莫戈罗夫 (Колмогоров) 在 1933 年提出了概率的公理化定义, 叙述如下:

定义 2 设试验的样本空间为 Ω , 对于任一随机事件 $A (A \subset \Omega)$, 都有确定的实值函数 $P(A)$, 满足下列性质:

(1) 非负性: $P(A) \geq 0;$ (1.3)

(2) 规范性: $P(\Omega) = 1;$ (1.4)

(3) 有限可加性: 对于 l 个互不相容事件 A_1, A_2, \dots, A_l , 有

$$P\left(\sum_{i=1}^l A_i\right) = \sum_{i=1}^l P(A_i) \quad (1.5)$$

或者, 更一般地

(3') 可列可加性: 对于可列无穷多个互不相容事件 A_1, A_2, \dots , 有

$$P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i), \quad (1.6)$$

则称 $P(A)$ 为随机事件 A 的概率.

3. 概率的性质

由随机事件的概率的定义, 不难得到概率的某些基本性质:

(1) 不可能事件的概率等于零, 即

$$P(\emptyset) = 0. \quad (1.7)$$

事实上, 因为 $\Omega = \Omega + \emptyset$, 由概率的有限可加性得

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

所以

$$P(\emptyset) = 0.$$

(2) 关于对立事件 A 与 \bar{A} , 有

$$P(A) = 1 - P(\bar{A}). \quad (1.8)$$

事实上, 因为 $A + \bar{A} = \Omega$, 所以有

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

由此即得等式 (1.8).

(3) 若事件 A 包含于事件 B , 即 $A \subset B$, 则

$$P(A) \leq P(B). \quad (1.9)$$

事实上, 若 $A \subset B$, 则有 (如图 1.3, 其中阴影部分表示 $\bar{A}B$)

$$B = A + \bar{A}B.$$

由概率的有限可加性得

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

再由概率的非负性知 $P(\bar{A}B) \geq 0$, 从而不等式 (1.9) 成立.

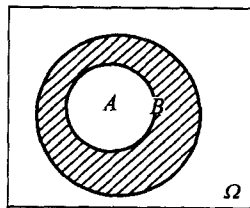


图 1.3

(4) 对于任一随机事件 A , 有

$$P(A) \leq 1. \quad (1.10)$$

事实上, 因为 $A \subset \Omega$, 由性质 (3) 即得

$$P(A) \leq P(\Omega) = 1.$$

(5) 对于任意两个随机事件 A 与 B , 有

$$P(A \cup B) = P(A) + P(B) - P(AB). \quad (1.11)$$

事实上, 因为 (如图 1.4, 其中阴影部分表示 $\bar{A}B$)

$$A \cup B = A + \bar{A}B,$$

由概率的有限可加性得

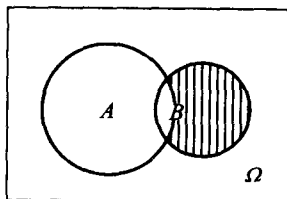


图 1.4