高等学校教材

应用概率统计方法

（第 2 版）

朱燕堂　赵选民　徐　伟　编

$$P\{x=k\}=C_n^k p^k(1-p)^{n-k}$$

$$k=0,1,\cdots,n.$$

西北工业大学出版社

（陕）新登字 009 号

【内容简介】 本书共分十章，主要内容包括随机事件及其概率、随机变量及其概率分布、随机变量的数字特征及极限定理、参数估计、假设检验、方差分析、回归分析、试验设计、概率论基础、统计计算与统计软件。本书各章均配有适量的习题和复习思考题，书末附有习题答案。

本书可作为高等学校工科各专业概率统计课程的教材，也可作为师范、农林、理科、财经、管理等专业本、专科生的教材或教学参考书，还可供工程技术人员参考使用。

高等学校教材

应用概率统计方法

（第 2 版）

朱燕堂　赵选民　徐伟　编

责任编辑　李　杰

责任校对　耿明丽

*

（710072　西安市友谊西路 127 号　电话 8493844）

陕西省新华书店经销

西安福利彩印厂印装

ISBN 7-5612-0974-6/O·130

*

开本：787×1092 毫米 1/16　　印张：19.25 字数：468 千字

1986 年 4 月第 1 版　　2000 年 1 月第 5 次印刷

印数：19 001－23 000 册　　定价：22.50 元

第1版 前　　言

本书的前身是朱燕堂同志在西北工业大学热加工各专业讲授“应用数学”课程时所编写的讲义，后经多次修改作为该校大部分专业工程数学“概率统计”课程的教材。这次编者又根据几年的教学实践经验和几位同志的教学使用情况进行了较大的修改，并补充了部分内容，由西北工业大学出版社出版。

全书共分十章，前五章主要介绍了随机事件及其概率、随机变量及其分布、数字特征、极限定理概述、参数估计、假设检验等概率统计的基础知识，从第六章至第十章着重介绍了正交试验设计、方差分析、回归分析、正交多项式回归及其在正交设计中的应用，非线性回归的样条函数拟合法等常用的数据处理方法。本书各章均配有适量习题，书末附有习题答案。

在编写此书过程中，编者考虑到初学概率统计的人不易接受这门课程的基本概念和处理问题方法的实际情况，力求做到由浅入深，通俗易懂，便于自学，但也不失对基本理论的要求，为读者进一步学习概率统计更高一级的课程打下必要的基础。另外，编者考虑到现在工科院校各专业对概率统计课需要的教学时数不尽相同，因此，此书的结构为模块式，可供教学时数为40～70学时的各专业选用，读者也可根据自己的情况选学有关章节。为了使读者能够把学到的常用数理统计方法付诸于实际应用，本书的许多例题和习题都是从生产实际的应用中选来的，其中有些例题和习题就是编者亲自参与解决过的。

本书由朱燕堂同志主编，杨敬娟同志修改了第一章至第三章，第四章至第九章由赵选民同志修改，张永曙同志编写了第十章，全书的插图由李时哲同志描绘。在编写本书的过程中，西北工业大学一系和数学教研室许多同志曾给予大力支持和热情帮助，提出了许多宝贵意见，该校先后有五届的学生学习使用了此书，亦提出了许多建设性的意见；本书全稿由北京航空学院韩于羹同志主审，并提出了许多有价值的意见和建议，在此一并致谢。

本书可作为高等工科院校各专业概率统计课的教材，也可作为电大、函授、夜大、专科各专业的教学参考书，亦可供有关工程技术人员和管理工作者参考。

由于编者水平有限，书中错误在所难免，恳切希望读者不吝指正。

编　者1985年6月

第2版 前 言

本书自1986年出版以来，累计发行15 000余册，在我校40多个专业及部分兄弟院校使用。根据10余年的教学实践，参照国家教委“概率论与数理统计课程教学基本要求”，在广泛征求读者及任课教师意见的基础上，本次对原书进行了修订。

修订中对原书结构进行了适当调整，取消了原书中第十章“非线性回归问题的样条函数拟合法”，将有关试验设计和回归分析的内容分别各作为一章进行叙述，并且增写了近代回归分析简介及常用试验设计方法简介等内容，以开阔读者视野。增写了第十章“统计计算与统计软件”，以提高读者应用计算机统计软件进行数学建模和统计分析的能力。

另外考虑到面向21世纪工科大学本科概率统计课程教学改革和实际应用的需要，本次修订中对第一章至第七章的部分内容作了修改与补充，增写了第九章“概率论基础”，从严格化的理论的角度，阐明了工科大学概率统计教材中不容易讲清楚的一些最基本的概念。为了帮助读者抓住学习要点，加强对基本概念的理解和掌握，每章章末增写了“小结”，阐明教学的基本要求；编写了“复习思考题”，以加强读者对基本概念、理论和方法的理解与复习。此外，对原书的习题作了一些调整与补充，增补了一些历年的考研试题，并对全部习题给出了答案；书后列出了主要参考文献。

本书的第一、二、三章由朱燕堂同志修改，第四、五、六、八章由赵选民同志修改，第七、九、十章由徐伟同志修改与编写。在修订过程中，西北工业大学应用数学系的许多同志提出了宝贵的意见与建议，西安交通大学范金诚教授仔细地审阅了原稿，并提出许多宝贵的意见与建议，西北工业大学出版社对本书的修订与出版给予了大力的支持与帮助，在此一并致以衷心的谢意。

本书力求做到适应21世纪概率统计课程教学改革和实际应用的需要，对基础理论与基本方法力求循序渐近，由浅入深，推理严谨，分析透彻，另外尽量讲清它们的实际背景及应用前景。但由于作者水平有限，书中难免存在缺点与错误，敬请读者不吝赐教。

编 者
1996年12月于西安

目　　录

第一章　随机事件及其概率 …… 1

§1　随机事件的概念及其运算 …… 1
§2　随机事件的概率 …… 5
§3　条件概率、全概率公式、贝叶斯公式 …… 13
§4　事件的独立性、独立试验序列概型 …… 18
第一章小结 …… 24
复习思考题一 …… 25
习题一 …… 25

第二章　随机变量及其分布 …… 29

§1　一维随机变量及其分布 …… 29
§2　多维随机变量及其分布 …… 42
§3　随机变量的函数及其分布 …… 53
第二章小结 …… 62
复习思考题二 …… 63
习题二 …… 64

第三章　随机变量的数字特征及极限定理概述 …… 69

§1　随机变量的数字特征 …… 69
§2　极限定理概述 …… 80
第三章小结 …… 85
复习思考题三 …… 85
习题三 …… 86

第四章　参数估计 …… 90

§1　数理统计的基本概念 …… 90
§2　抽样分布 …… 95
§3　求参数的点估计的常用方法及估计量的衡量标准 …… 103
§4　区间估计 …… 113
第四章小结 …… 120
复习思考题四 …… 120
习题四 …… 121

第五章　假设检验……125
§1　假设检验的基本概念……125
§2　正态总体下参数的假设检验……128
§3　非正态总体的假设检验……137
§4　χ^2 拟合优度检验……140
第五章小结……143
复习思考题五……143
习题五……144
第六章　方差分析……148
§1　单因素方差分析……148
§2　双因素方差分析……152
§3　考虑交互作用的双因素方差分析……156
第六章小结……160
复习思考题六……160
习题六……161
第七章　回归分析……163
§1　一元线性回归分析……164
§2　可化为线性回归的非线性回归……171
§3　多元线性回归分析……175
§4　正交多项式回归……182
§5　几种近代回归分析方法简介……191
第七章小结……192
复习思考题七……193
习题七……193
第八章　试验设计……196
§1　正交试验设计的直观分析法……196
§2　有交互作用的试验……204
§3　正交试验设计的方差分析……208
§4　重复试验、重复取样的方差分析……212
§5　正交多项式回归在正交设计中的应用……217
* §6　常用试验设计方法简介……222
第八章小结……227
复习思考题八……228
习题八……228

*第九章　概率论基础 …… 232
§1　测度与概率 …… 232
§2　可测函数与随机变量 …… 237
§3　数学期望和特征函数 …… 242
第九章小结 …… 245
复习思考题九 …… 245
习题九 …… 246
*第十章　统计计算与统计软件简介 …… 248
§1　概率统计计算 …… 248
§2　统计软件 …… 249
§3　SPSS 统计软件简介 …… 250
§4　SAS 软件简介 …… 253
第十章小结 …… 255
复习思考题十 …… 255
*习题十 …… 255
附表 …… 257
附表 1　泊松分布表 …… 257
附表 2　正态分布数值表 …… 260
附表 3　t 分布的上侧分位数表 …… 262
附表 4　χ^2—分布临界值表 …… 264
附表 5　F—分布临界值表(F_α)($\alpha=0.05$) …… 265
附表 6　F—分布临界值表($\alpha=0.10$) …… 269
附表 7　F—分布临界值表($\alpha=0.01$) …… 270
附表 8　F—分布临界值表($\alpha=0.025$) …… 274
附表 9　相关系数临界值表 …… 275
附表 10　常用正交表 …… 277
(1) $L_4(2^3)$ …… 277
(2) $L_8(2^7)$ …… 277
(3) $L_{16}(2^{15})$ …… 278
(4) $L_{32}(2^{31})$ …… 279
(5) $L_{12}(2^{11})$ …… 281
(6) $L_9(3^4)$ …… 281
(7) $L_{27}(3^{13})$ …… 282
(8) $L_{18}(2^1\times 3^7)$ …… 284
(9) $L_{16}(4^5)$ …… 284
(10) $L_{25}(5^6)$ …… 285

附表 11. 常用正交多项式表 …………………………………………………… 285
习题答案……………………………………………………………………………… 287
参考文献……………………………………………………………………………… 299

第一章　随机事件及其概率

§1　随机事件的概念及其运算

一、随机现象和随机事件

1. *必然现象和随机现象*

在自然界和人类社会中，人们观察到的现象大体可分为两种类型，一类是在相同条件下，多次观察，总是有相同的结果。我们把这类现象称为确定性现象或必然现象。例如在 101 325 Pa 压强作用下，水加热到 100℃，不管谁来做试验，每次试验结果都是相同的，“水沸腾了”。在没有外力作用的条件下，作匀速直线运动的物体必然继续作匀速直线运动等。早期的科学就是研究这一类现象的数量规律性，所应用的数学工具是诸如数学分析、几何、代数、微分方程等。但随着科学技术和生产的发展，人们逐渐发现还广泛存在着与必然现象有本质区别的另一类现象，例如：用同一仪器多次测量一物体的重量，所得到的结果总是略有差异；在相同条件下，多次抛掷一枚匀称的硬币，其结果可能是正面(有国徽的一面为正面) 朝上，也可能是正面朝下；远距离射击较小的目标，可能击中，也可能击不中；自动车床加工出来的零件可能是合格品，也可能不是合格品等等。类似的例子还可举出很多。

上面所举的现象的共同特点是在相同条件下，经多次试验或观察会得到一系列不同的结果，而且每次试验之前不知道会出现哪种结果，这类现象称为随机现象。

人们经过长期实践并深入研究之后，发现这类现象虽然就每次试验或观察结果来说，它具有不确定性，但在大量重复试验或观察下它的结果却呈现出某种规律性。例如，多次重复抛掷一枚匀称硬币得到正面朝上大致有半数，多次远距离射击同一小目标的弹着点按照一定规律分布等等。这种在大量重复试验中所呈现出的规律性称为统计规律性。

这样，随机现象可更确切地解释为在个别试验中可能出现这种结果或那种结果，即呈出不确定性，但在大量重复试验中，又具有统计规律性的现象。

概率论与数理统计就是研究随机现象的统计规律性的数学分支。

2. *随机试验*

前面所说的对随机现象进行观察总是在一定条件下进行的。若把一次观察视为一次试验，观察的结果就是试验结果。在概率论中把满足下列三个条件的试验称为随机试验：① 允许在相同的条件下重复地进行；② 每次试验结果不一定相同；③ 试验之前不知道会出现哪种结果。今后我们所指的试验都是随机试验，它是个广泛的名词，包括各种各样的科学实验，甚至对某一事物的某一特征的观察也认为是一种试验。例如：(1) 掷一枚匀称的硬币，观察出现正面或反面的情况；(2) 掷一颗骰子，观察出现的点数；(3) 记录某电话总机在一天内所接到的呼唤次数；(4) 在一批灯泡中任取一只，测试它的寿命等等都是随机试验的例子。

3. *随机事件、基本空间*

定义 1.1 在随机试验中，可能出现、也可能不出现的事情叫做随机事件。例如：

(1) 在掷硬币的试验中，“正面朝上”这一事情，便是随机事件；

(2) 在掷一颗骰子的试验中，“出现偶数点”，“出现一点”，“出现两点”等都是随机事件；

(3) 在一分钟内，一个电话总机“至少接到 10 次呼唤”是随机事件；

(4) 测得“灯泡寿命为 1 000 h”是随机事件。

由上可知，随机事件可以是随机试验的某一个结果(如“正面朝上”，“出现一点”等)也可以是若干个试验结果组合而成的事件(如“出现偶数点”便是由“出现两点”，“出现四点”，“出现六点”组成的)。

在随机试验中，每一个可能出现的结果都是随机事件，它是这个试验的最简单的随机事件，称为基本事件。其他的一些随机事件可由若干个基本事件组成。

例如，掷硬币试验中，“出现正面”，“出现反面”是这个试验的基本事件。掷骰子的试验中，“出现一点”，“出现两点”，……，“出现六点”就是这个试验的基本事件。而“出现偶数点”也是随机事件，它是由“出现两点”，“出现四点”，“出现六点”三个基本事件组成的。

在随机试验中必然会出现的事情叫做必然事件，必然不出现的事情叫做不可能事件。例如，在掷骰子试验中，“点数不大于 6”是必然事件，“点数大于 6”是不可能事件。必然事件和不可能事件是随机事件的两个特例，本来这两类事件没有不确定性，也就是说它们不是随机事件，但为了今后讨论方便起见，把它们当做一种特殊的随机事件。

今后，把随机事件常简称为事件，并用大写字母 $A,B,C\cdots$ 等表示随机事件。用 Ω 表示必然事件，用 Φ 表示不可能事件。

定义 1.2 为了研究随机试验，我们把所有可能的试验结果(或所有基本事件)的全体所构成的集合叫做基本空间(或样本空间)，记为 Ω，而把每一个可能出现的结果(或每一个基本事件)称为一个样本点，记为 ω。

一般地说，基本空间 Ω 可以由有限个基本事件(样本点)所组成，也可以由无限个基本事件(样本点)所组成，甚至是某范围内的全体实数所组成。如：

在例(1)中 $\Omega=\{$出现正面，出现反面$\}$；

在例(2)中 $\Omega=\{1,2,3,4,5,6\}$；

在例(3)中 $\Omega=\{0,1,2,\cdots\}$；

在例(4)中 $\Omega=\{t\mid t\geqslant 0\}$。

由于随机事件是基本事件或由若干个基本事件所组成，因而它是基本空间的子集，事件 A 出现，就是当且仅当 A 所包含的某一样本点出现。特别地，必然事件就是基本空间 Ω，不可能事件就是空集 $\varnothing$。

二、事件间的关系及其运算

在实际问题中，往往不只研究随机试验的一个事件，而要研究很多事件，而这些事件之间又有一定的联系。例如，在检查某些圆柱形产品时要求它的长度及直径都符合规格才算合格。这时要考虑“产品合格”，“产品不合格”，“直径合格”，“直径不合格”，“长度合格”，“长度不合格”，“长度合格但直径不合格”等等事件，显然这些事件之间是有一定联系的。为了表述类似于上述事件之间的联系，下面引进事件之间的几种主要关系以及事件的运算。

(1) 事件的包含和相等。设有事件 A 及 B，如果事件 A 出现必然导致事件 B 出现，则称事

件 B 包含事件 A，或称 A 是 B 的子事件。记做

$$A \subset B \quad 或 \quad B \supset A$$

例如，令 A 表示"直径不合格"，B 表示"产品不合格"，显然，若"直径不合格"这一事件 A 出现，则"产品不合格"这一事件 B 就必然出现。故 $A \subset B$。对任一事件 A，规定 $\varnothing \subset A$，显然对任一事件 A，必有 $\varnothing \subset A \subset \Omega$。

如果事件 A 包含事件 B，同时事件 B 也包含事件 A，即 $B \subset A$ 和 $A \subset B$ 同时成立，则称事件 A 与事件 B 相等，记做 $A = B$。

(2) 事件的和与积。设 A,B 是两个事件，"事件 A 与 B 至少有一个出现"这一事件称为事件 A 与事件 B 的和事件，记做 $A \cup B$。这就是说，$A \cup B$ 出现，它意味着 A,B 中至少有一个出现，也就是表示事件 A 出现或事件 B 出现。因此 $A \cup B$ 也可记为"A 或 B"。例如，若用 A 表示"直径不合格"，B 表示"长度不合格"，则 $A \cup B$ 表示"产品不合格"，即"产品不合格"是"直径不合格"与"长度不合格"两事件的和事件。

一般地，"事件 $A_1,A_2,\cdots,A_n,\cdots$ 中至少有一个出现"这一事件，称为事件 $A_1,A_2,\cdots,A_n,\cdots$ 的和事件，记做 $A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots$ 或 $\bigcup\limits_{n=1}^{\infty} A_n$。

"事件 A 与事件 B 同时出现"这一事件称为事件 A 与事件 B 的积事件，记做 $A \cap B$(或记做 AB)。例如，用 A 表示"直径合格"，B 表示"长度合格"，则 $A \cap B$ 表示"产品合格"，即"产品合格"是"直径合格"与"长度合格"两事件的积事件。

类似地，"事件 $A_1,A_2,\cdots,A_n,\cdots$ 同时出现"这一事件称为事件 $A_1,A_2,\cdots,A_n,\cdots$ 的积事件(或称为交)，记做 $A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots$。

(3) 事件的互斥与互逆。如果事件 A 与事件 B 不能同时出现，即 A,B 的积事件为不可能事件，则称 A,B 为互斥事件(或称为互不相容事件)。记为 $AB = \varnothing$。

对于互斥事件 A,B，可以把和事件 $A \cup B$ 记做 $A + B$。例如在一批包含有正品、次品的产品中，任意抽取三个，"三个都是次品"这一事件记为 A，"三个都是正品"这一事件记为 B，显然 A 与 B 是互斥事件。

如果两事件 A,B 同时满足关系式：$A \cup B = \Omega, AB = \varnothing$(即 A,B 中必出现其一，但 A 与 B 不能同时出现)，则称 A,B 两事件互逆或对立。并称 A 是 B 的逆事件(或对立事件)，或称 B 是 A 的逆事件(或对立事件)。把事件 A 的逆事件记做 $\overline{A}$。例如，掷一枚硬币时，"出现正面"与"出现反面"是互逆事件，检查圆柱形产品时，"直径合格'与"直径不合格"是互逆事件等等。

对任一事件 A，显然有 $\overline{\overline{A}} = A$，即 A 也是 $\overline{A}$ 的对立事件。我们看到：在一次试验中，A 与 $\overline{A}$ 不会同时出现(即它们互斥)，而且 $A,\overline{A}$ 至少有一个出现，就是说 A 和 $\overline{A}$ 满足：$A\overline{A} = \varnothing, A + \overline{A} = \Omega$。

(4) 事件的差。"事件 A 出现而事件 B 不出现"这一事件称为事件 A 与事件 B 的差，记做 $A - B$。例如："直径合格"记为 A，"长度合格"记为 B，"直径合格但长度不合格"记为 C，于是有 $C = A - B$。又如，在一批包含有正品、次品的产品中，任取三个，"至少有一个次品"记为 A_1，"至少有两个次品"记为 A_2，"恰有一个次品"记为 A_3。于是 $A_3 = A_1 - A_2$。

显然有 $A - B = A\overline{B}, \overline{A} = \Omega - A$。

下面把事件间的关系及其运算用图形表示出来更为直观些，这种图称为文(Venn)图。

事件 $\Omega, A, \overline{A}, A \cup B, AB, A - B$ 在图 1-1 中分别以阴影部分表示，不难理解，$B \supset A$ 相

应于 A 的图形完全包含在 B 的图形中；A 和 B 互不相容，则相应于 A 和 B 的图形不相交。

上述图形也可以用打靶的例子来说明，事件 A 代表命中小圆内，事件 B 代表命中大圆内，则 $A\cup B$ 代表命中图 1-1 中 $A\cup B$ 的阴影部分，其他也可类似说明。

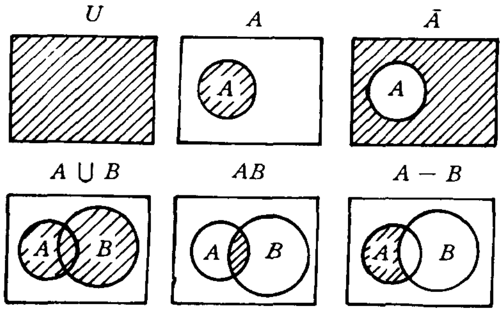

图 1-1

熟悉集合论的读者或许早就发现，事件间的关系及运算与集合论中集合间的关系与运算是完全相似的。若把概率论中的基本事件看做集合论中的元素，由若干个基本事件组成的事件，便可看做包含若干元素的集合。而把由基本事件的全体构成的基本空间看做集合论中的全集或空间。为了便于对照，把它们的术语列表如表 1-1 所示：

表 1-1

符号	概率论	集合论
Ω	基本空间，必然事件	空间，全集
$\varnothing$	不可能事件	空集
ω	基本事件	元素
A	事件	子集
$\overline{A}$	A 的对立事件	集合 A 的余集
$A\subset B$	A 是 B 的子事件	A 是 B 的子集
$A=B$	事件 A 与事件 B 相等	集合 A 与集合 B 相等
$A\cup B$	事件 A 与 B 至少有一个出现	A 与 B 的并集
AB	事件 A 与 B 同时出现	A 与 B 的交集
$A-B$	事件 A 出现而 B 不出现	A 与 B 的差集
$AB=\varnothing$	事件 A 与事件 B 互斥	A 与 B 没有相同元素

事件与集合一样，有如下的运算规律：

(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：$(A\cup B)C=AC\cup BC$，

　　$(AB)\cup C=(A\cup C)(B\cup C)$；

(4) 对偶律：$\overline{A\cup B}=\overline{A}\,\overline{B}$，　$\overline{AB}=\overline{A}\cup\overline{B}$。

一般地，设 $A_1,A_2,\cdots A_n$ 是 n 个随机事件，则

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

这样，我们可以把对事件的分析转化为对集合的分析，利用集合间的运算关系来分析事件

间的关系。不过还要注意学会用概率论的语言来解释这些关系及运算，并且会用这些运算关系来表示一些事件。下面举例说明之。

【例 1-1】 设 A,B,C 为三个事件，则

(1) A 出现而 B 与 C 都不出现，可表示为 $A\bar{B}\bar{C}$（或 $A-B-C$，或 $A-(B\cup C)$）；

(2) A 与 B 都出现而 C 不出现，可表示为 $AB\bar{C}$（或 $AB-C$，或 $AB-ABC$）；

(3) 所有这三个事件都出现，可表示为 ABC；

(4) 这三个事件恰好出现一个，可表示为 $A\bar{B}\bar{C}+\bar{A}\,\bar{B}C+\bar{A}BC$；

(5) 这三个事件恰好出现两个，可表示为 $AB\bar{C}+A\bar{B}C=\bar{A}BC$；

(6) 这三个事件至少出现一个，可表示为 $A\cup B\cup C$，或 $A\bar{B}\,\bar{C}+\bar{A}B\bar{C}+\bar{A}\bar{B}C+AB\bar{C}+A\bar{B}C+\bar{A}BC+ABC$，或 $\Omega-\bar{A}\bar{B}\bar{C}$；

(7) 这三个事件中不多于一个事件出现，可表示为 $\bar{A}\bar{B}\bar{C}+A\bar{B}\bar{C}+\bar{A}B\bar{C}+\bar{A}\,\bar{B}C$，或 $\overline{AB\cup BC\cup AC}$。

【例 1-2】 在两次射击靶子的试验中，令 A 表示第一次射击击中目标的事件，B 表示第二次射击击冲目标的事件，则

(1) $C=A\cup B$ 表示至少有一次射中目标的事件，不管它是第一次射中或第二次射中或两次都射中目标。

(2) $C=AB$ 表示两次射击都射中目标的事件。

(3) $C=\bar{A}\bar{B}$ 表示两次射击都未射中目标的事件。

§2 随机事件的概率

随机事件虽然有其不确定性的一面，即在一次试验中，可能出现，也可能不出现，但在大量重复试验中，人们发现它是有内在规律性的，即它出现的可能性大小有所不同，而且是可以度量的。例如，在一次试验中，掷一枚匀称的硬币出现“正面朝上”比掷二枚匀称的硬币出现“两个都是正面朝上”的可能性大；在多次重复试验中，“射击 10 次而命中靶子”比“射击 2 次而命中靶子”的可能性大。这说明随机事件出现的可能性大小是客观存在的。其次，在许多实际问题中人们也希望了解某些事件出现的可能性有多大。例如，要在某河流上建筑一座防洪水坝，为了确定水坝的高度，就要知道该河流在造水坝地段每年最大洪水达到某高度的可能性的大小。最大洪水达到某一高度是随机事件。这样，很自然地，人们想用一个数 p 来衡量事件出现的可能性的大小，且当事件 A 出现的可能性较大，就用较大的数表示，出现的可能性较小，就用较小的数来表示。这个数 p 我们称为事件 A 出现的概率，通常记为 $P(A)=p$。因此，随机事件 A 的概率就是用来度量随机事件出现的可能性大小的一个概念，它是概率论中最基本的概念之一。

对于已给事件 A，怎样来确定 $P(A)$ 的值呢？我们先从随机事件的频率出发，引进一般情形下确定 $P(A)$ 的值的方法，即概率的统计定义，然后提出在一些简单的随机试验情况下确定 $P(A)$ 的值的方法，即概率的古典定义及几何定义。最后以公理的形式引入概率的一般定义，并由此推出概率的一些常用性质。

一、概率的统计定义

我们先以掷一枚匀称硬币这个随机试验为例为说明确定“出现正面”这一事件的概率的方法,将硬币抛掷 n 次,观察在 n 次试验中“出现正面”的次数 m,m/n 称为出现正面的频率。

历史上有些人做过成千上万次抛掷硬币的试验,表 1-2 列出他们的试验记录:

从表 1-2 看出,不管什么人去投掷,当实验次数逐渐增多时,“正面朝上”的频率越来越明显地稳定并接近于 1/2。这个数能反映出“出现正面”的可能性的大小。所以,我们将用表 1-2 中最后一列中的频率作为投掷硬币“出现正面”的概率的近似值,并认为 $P(A)$ 近似等于 1/2。

表 1-2

实验者	投掷次数 n	出现“正面朝上”的次数 m(即频数)	频率 $=\frac{m}{n}$
德莫根	2 048	1 061	0.518
浦　丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

再来分析一个例子,在一个口袋中装有 6 只球(4 只白球,2 只红球),从袋中任取一球,问取到白球的可能性有多大。设 A 表示取出的球为白球这一事件。现在做许多次试验,观察取到白球的次数,并算出它出现的频率。有人做 600 次试验(从袋中任取一球,观察其颜色后放回袋中,搅匀后再任取一球观察其颜色,这样继续下去),得数据如表 1-3 所示。

表 1-3

试验次数 n	100	200	300	400	500	600
频数 m	69	139	198	261	337	401
频率 m/n	0.690	0.695	0.660	0.653	0.674	0.668

从表 1-3 中可以看出,频率在 $0.66\cdots = 2/3$ 附近摆动,如再继续下去,将逐渐稳定于 2/3。这个数能反映事件 A 出现的可能性的大小。这一事实不因人而异,这就是说在相同条件下,不论谁,只要做大量的重复试验(尽管具体数字有出入),其频率总是稳定于这个常数上。

经验证明,任何随机事件 A,只要试验是相同条件下多次重复进行的,那么事件 A 出现的频率就具有稳定性,就是说,当试验次数充分大时,事件 A 出现的频率总在[0,1]区间上的某个确定的数字 p 附近摆动。因为频率总是介于 0 与 1 之间的一个数,这个常数 p 是客观存在的,这也是我们下面定义事件概率的客观基础。

定义 1.3　在一个随机试验中,如果事件 A 出现的频率 m/n 随着试验次数 n 的增大,它在区间[0,1]上的某个常数 p 附近摆动,那末定义事件 A 的概率为

$$P(A) = p$$

概率的这种定义,称为概率的统计定义。

由概率的统计定义可以推得概率具有下列性质:

(1) 对任一事件 A,有

$$0 \leqslant P(A) \leqslant 1$$

这是因为频率 m/n 总是在区间$[0,1]$上，所以相应的 $p = P(A)$ 也总在区间$[0,1]$上。

(2) $P(\Omega) = 1, P(\varnothing) = 0$。这是因为对于必然事件 Ω，不可能事件 $\varnothing$，频率依次固定为 1，0，所以相应的 $p = P(A)$ 也依次为 1，0。

(3) 对于两两互斥的有限多个随机事件 $A_1, A_2, \cdots, A_n$，有

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

这是由于 $A_1, A_2, \cdots, A_n$ 两两互斥，所以 $A, A_1, A_2, \cdots, A_n$ 的频率 $r/n, r_1/n, r_2/n, \cdots, r_n/n$ 满足等式(其中 $A = A_1 + A_2 + \cdots + A_n$)

$$\frac{r}{n} = \frac{r_1}{n} + \frac{r_2}{n} + \cdots + \frac{r_n}{n}$$

相应的概率应该满足

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

概率的统计定义虽然直观，但据此计算某事件的概率是困难的，仅能以 A 的频率作为 $P(A)$ 的近似值，而且当试验次数 n 越大时越能准确地定出近似值。然而 n 要多大，准确到什么程度，都没有确切的说明，但在实际问题中有许多随机现象满足一定的条件，便能直接计算概率，这就是下面要引进的概率的古典定义。

二、概率的古典定义

若随机现象有下列一些特征：

(1) 基本试验结果的个数是有限的，就是说，该试验的所有可能结果仅有有限个，即仅有有限个基本事件；

(2) 每个基本试验结果出现的可能性相同；

(3) 在任一次试验中，只能出现一个结果，也就是有限个基本事件是两两互斥的。

这类现象在概率论发展初期曾是主要的研究对象，一般把这类现象的数学模型称为古典概型。例如，掷骰子试验，掷硬币试验，袋中摸球，产品质量检查等实际问题都属于古典概型。

对于古典概型的情形，设试验的所有可能结果为 n 个，即有 n 个基本事件，而事件 A 包含有其中的 m 个试验结果，即 A 中有 m 个基本事件，于是定义事件 A 的概率为

$$P(A) = \frac{m}{n} = \frac{A\text{ 中所包含的基本事件数}}{\text{基本事件的总数}} \tag{1-1}$$

概率的这种定义，称为概率的古典定义。由等可能性的假定，便容易理解上述定义确实客观地反映了随机事件出现的可能性的大小。

【例 1-3】 掷一颗骰子，求得奇数点的概率。

【解】 显然，$\Omega = \{1,2,3,4,5,6\}$，即基本事件的总数 $n = 6$，而 $A = \{1,3,5\}$，即事件 A—— 得奇数点所包含的基本事件数 $m = 3$，于是由定义，所求概率为

$$P(A) = \frac{3}{6} = 0.5$$

【例 1-4】 设箱中装有 100 件产品，其中有 3 件次品，为检查产品质量从中任意抽取 5 件，求所取 5 件产品中恰有 1 件次品的概率。

【解】 从 100 件产品中任意抽取 5 件产品，共有 C_{100}^5 种抽取方法，即基本事件总数 $n =$

C_{100}^{5}，设 A 表示“取到的 5 件产品中恰有 1 件次品”这一事件，则 A 中所包含的基本事件数可以这样计算：1 件次品从 3 件次品中取得，共有 C_3^1 种取法，又 4 件正品从 97 件正品中取得，共有 C_{97}^4 种取法。因而事件 A 所包含的基本事件数 $m = C_3^1 \times C_{97}^4$。这样，所求的概率为

$$P(A) = \frac{C_3^2 \times C_{97}^4}{C_{100}^5} \approx 0.138$$

【例 1－5】 续例 1－4，求抽取的 5 件产品中至多有一件次品的概率。

【解】 设 A 表示“所取 5 件产品中至多有一件次品”的事件，则 A 中所包含的基本事件数为：$m = C_{97}^5 + C_3^1 \times C_{97}^4$，因此，所求概率为

$$P(A) = \frac{C_{97}^5 + C_3^1 \times C_{97}^4}{C_{100}^5} \approx 0.994$$

【例 1－6】 设有 0，1，2，…，9 等 10 个数字，在这 10 个数字中顺次有放回地抽取 5 次（即抽了第 1 个数字后放回去，再抽第 2 个数字），求抽到的 5 个数字全不相同的概率。

【解】 设 A 表示“抽到的 5 个数字全不相同”的事件，第一次抽取方法有 10 种；因有放回，所以第 2 次抽取方式也有 10 种；这样，直至第 5 次抽取方式仍有 10 种，所以连续 5 次有放回的抽样方式共有 $10 \times 10 \times 10 \times 10 \times 10 = 10^5$ 种，即基本事件总数 $n = 10^5$，而抽到 5 个数字都不相同的抽取方式有 A_{10}^5 种，即 A 中所包含的基本事件数 $m = A_{10}^5$，这样，所求的概率为

$$P(A) = \frac{A_{10}^5}{10^5} \approx 0.3024$$

【例 1－7】 有 n 个可分辨的球，随机地放入 $N(N \geqslant n)$ 个盒子中的每一个盒，试求：

(1) 指定的 n 个盒中各有一球的概率；

(2) 任何 n 个盒中各有一球的概率；

(3) 某指定的一个盒中恰有 $m(m \leqslant n)$ 个球的概率。

【解】 设球以同样的可能性落入 N 个盒中的每一个，每一个球一共有 N 种放法，n 个球可以有 $N \times N \times \cdots \times N = N^n$ 种放法，这就是基本事件总数

(1) 以 A 表示“指定的 n 个盒中各有一球”的事件，今固定某 n 个盒，第一个球可以落入这 n 个盒中任何一个，有 n 种方法，第二个球可落在余下的 $n-1$ 个盒中的任何一个，有 $n-1$ 种方法，……，第 n 个球落在最后一个盒，只能有一种方法，因此事件 A 共含有 $n!$ 个不同的基本事件数，故所求的概率为

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

(2) 以 B 表示“任何 n 个盒中各有一球”的事件。因为任何 n 个盒可以从 N 个盒中任意选取，共有 C_N^n 种选法，选出这 n 个盒后，再按问题(1) 可知事件 B 共含 $C_N^n \cdot n!$ 个不同的基本事件数，故所求的概率为

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

(3) 以 C 表示“某指定的一个盒中恰有 m 个球”的事件，因为 m 个球可以从 n 个球中任意选出，共有 C_n^m 种选法，其余 $n-m$ 个球可以任意落入其余 $N-1$ 个盒中共有 $(N-1)^{n-m}$ 种落法，因此事件 C 共含 $C_n^m \cdot (N-1)^{n-m}$ 个不同的基本事件，故所求的概率为

$$P(C) = \frac{C_n^m (N-1)^{n-m}}{N^n}$$

由以上各例可以看到，在古典概型的概率计算中常用到一些排列组合的有关知识，熟悉这方面的知识是必要的。但古典概型中的许多问题的求解并非都如此容易，而是富有技巧的。

由概率的古典定义可知，概率也具有下列三个性质：

(1) 对任意事件 A，有 $0 \leqslant P(A) \leqslant 1$；

(2) $P(\Omega) = 1, P(\varnothing) = 0$；

(3) 设事件 $A_1, \cdots, A_n$ 两两互斥，则

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

证 (1) 由定义显然成立，因为包含在 A 中的基本事件数 $0 \leqslant m \leqslant n$，故 $0 \leqslant P(A) = m/n \leqslant 1$；

(2) 因为 Ω 中所含的基本事件数就是总的基本事件数，所以 $P(\Omega) = n/n = 1$，同理 $P(\varnothing) = 0$；

(3) 为简单起见，仅就 $n = 2$ 的情形进行证明。

设 A_1 由 r 个基本事件组成，A_2 由 s 个基本事件组成，由于 A_1 与 A_2 互斥。所以 $A_1 + A_2$ 由 $r + s$ 个基本事件组成，基本事件总数设为 n，于是

$$P(A_1 + A_2) = \frac{r+s}{n} = \frac{r}{n} + \frac{s}{n} = P(A_1) + P(A_2)$$

【例 1-8】 袋中有白球 5 个，黑球 3 个，现从中接连取两个球，第一次取出的球不放回去，求取得两球为一白一黑的概率。

【解】 设 A 为“取得两球为一白一黑”的事件；

A_1 为“第一次取得白球，第二次取得黑球”的事件；

A_2 为“第一次取得黑球，第二次取得白球”的事件。

显然，A_1, A_2 互斥且 $A = A_1 + A_2$，于是所求的概率为

$$P(A) = P(A_1) + P(A_2) = \frac{5 \times 3}{8 \times 7} + \frac{3 \times 5}{8 \times 7} = \frac{15}{28}$$

三、概率的几何定义

根据概率的统计定义求事件的概率，需作大量重复试验，这在实际问题中往往是很麻烦的，甚至是不可能的。在概率的古典定义中，不需要试验可直接根据公式求出事件的概率，这是它最大的优点，但是，它也有局限性，因为它要求试验的全部可能结果的数目是有限的，而且每个试验结果出现的可能性相等。如果试验的全部可能结果是无限的，古典定义就不适用了，必须从另外的角度来考虑这类事件的概率。

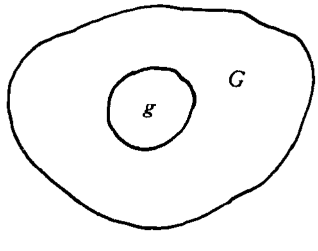

图 1-2

下面看一个试验的全部可能结果数目是无限的，而各个可能结果具有等可能性的一个典型例子。

例如，设有一平面区域 G，其中有一小区域为 g，如图 1-2。今若向区域 G 内任意抛掷一点 M，试问此点落于 g 内的概率为多少？这里，“在区域 G 内任意投掷一点”这句话应理解为：是必定落于 G 中，而且落在区域 G 内的任何部分内的概率只与这部分的面积成正比例，而与其位置和形状无关。于是，在区域 G 内任意投掷一点而该点落在区域 g 内的概率就可定义为