经济应用数学基础

JIANMINGJIAOCHENG

概率统计
简明教程

杜之韩　白淑敏　编著

TONGJI JIANMING JIAOCHENG

BAI SHUMIN　BIANZHU

西南财经大学出版社
SOUTHWESTERN UNIVERSITY OF FINANCE & ECONOMICS PRESS

GAILUTONGJI
JIANMINGJIAOCHENG

概率统计

简明教程

杜之韩　白淑敏　编著

GAILU TONGJI JIANMING JIAOCHENG

DU ZHIHAN　BAI SHUMIN　BIANZHU

西南财经大学出版社
SOUTHWESTERN UNIVERSITY OF FINANCE & ECONOMICS PRESS

图书在版编目(CIP)数据

概率统计简明教程/杜之韩,白淑敏编著 .—成都:西南财经大学出版社,2004 3

ISBN 7-81088-186-8

Ⅰ 概 . Ⅱ.①杜...②白... Ⅲ ①概率论—函授大学—教材②数理统计—函授大学—教材 Ⅳ.021

中国版本图书馆 CIP 数据核字(2003)第 120277 号

概率统计简明教程

杜之韩 白淑敏 编著

责任编辑·韩江

封面设计:大涛视觉传播设计事务所

出版发行·	西南财经大学出版社(四川省成都市光华村街 55 号)
网 址·	http://www xcpress com/
电子邮件:	xcpress@mail sc cninfo net
邮政编码:	610074
电 话.	028-87353785 87352368
印 刷.	西南财经大学印刷厂
开 本	890mm×1240mm 1/32
印 张:	7 75
字 数:	198 千字
版 次:	2004 年 3 月第 1 版
印 次:	2004 年 3 月第 1 次印刷
书 号·	ISBN 7-81088-186-8/F·163
定 价:	15.00 元

前　言

这是一本为高等财经类大学本科学生撰写的《概率论与数理统计》教材，其内容涵盖了教育部颁布的高等财经类专业核心课程《经济数学基础》教学大纲中概率统计部分的全部基本要求（即大纲中不带*号的内容），并略有拓宽，可以满足普通高等财经院校绝大多数专业及成人院校对本课程的要求．

编者编写本教材的立意是：采用最有利于学生接受的内容体系，从较低的教学起点出发，讲清概率论与数理统计课程最低限度要求中的最基本问题．希望这样一种安排能使包括成人院校学生和常为本课程内容的艰深而却步的自考生在内的所有学生们从本书都有所收获．

本书的编写得到西南财经大学经济数学系、成人教育学院领导以及经济数学系长期从事概率论与数理统计课程教学的同事们的大力支持，谨此向他们表示衷心的感谢．虽然作者为本书已竭尽心志，但限于水平，纰漏乃至谬误或难以避免．恳请同仁、读者一经发现及时指教．

编者

2003 年冬于光华园

目 录

第一章 随机事件及其概率…………………………………………（1）

§1.1 随机事件及其关系与运算 ……………………………（2）
§1.2 概率 ……………………………………………………（8）
§1.3 古典概型…………………………………………………（14）
§1.4 条件概率…………………………………………………（19）
§1.5 事件的独立性……………………………………………（27）
习题一 ………………………………………………………（32）

第二章 随机变量的分布及数字特征 …………………………（37）

§2.1 随机变量的概念…………………………………………（37）
§2.2 离散型随机变量…………………………………………（39）
§2.3 几种重要的离散型分布…………………………………（41）
§2.4 分布函数…………………………………………………（48）
§2.5 连续型随机变量及其分布………………………………（52）
§2.6 几种重要的连续型分布…………………………………（56）
§2.7 随机变量的函数的分布…………………………………（62）
§2.8 随机变量的数字特征……………………………………（67）
习题二 ………………………………………………………（79）

第三章 多维随机变量 ……………………………………………（85）

§3.1 多维随机变量及其分布函数……………………………（85）
§3.2 二维离散型随机变量的分布……………………………（87）
§3.3 二维连续型随机变量的分布……………………………（92）

§3.4　两个重要的二维连续型分布……………………（99）
§3.5　二维随机变量的函数的分布 ……………………（104）
§3.6　条件分布 ……………………………………………（110）
§3.7　二维随机变量的数字特征 …………………………（114）
§3.8　大数定律 ……………………………………………（123）
§3.9　中心极限定理 ………………………………………（127）
习题三………………………………………………………（131）

第四章　数理统计的基本概念…………………………………（137）

§4.1　总体、样本与统计量………………………………（137）
§4.2　抽样分布 ……………………………………………（140）
习题四………………………………………………………（151）

第五章　参数估计……………………………………………（153）

§5.1　参数的点估计 ………………………………………（153）
§5.2　点估计量的评价标准 ………………………………（160）
§5.3　区间估计 ……………………………………………（164）
习题五………………………………………………………（170）

第六章　假设检验……………………………………………（174）

§6.1　假设检验的基本概念 ………………………………（174）
§6.2　一个正态总体的假设检验 …………………………（179）
§6.3　两个正态总体的假设检验 …………………………（188）
习题六………………………………………………………（194）

第七章　回归分析初步………………………………………（197）

§7.1　一元线性回归模型 …………………………………（198）
§7.2　一元线性回归的显著性检验 ………………………（202）
§7.3　一元线性回归的预测 ………………………………（206）
习题七………………………………………………………（210）

习题参考答案…………………………………………………… (212)

附表…………………………………………………………… (226)

1　泊松分布表 ………………………………………… (226)
2　正态分布表 ………………………………………… (229)
3　χ^2 分布表 ………………………………………… (230)
4　t 分布表 …………………………………………… (232)
5　F 分布表 …………………………………………… (233)

第一章　随机事件及其概率

在客观世界中,我们发现许多现象在一定的条件下或者一定发生,或者一定不发生．例如,在没有外力作用下,作匀速直线运动的物体必将继续作匀速直线运动;又如,在标准大气压下,水加热到100℃时必然会沸腾,若只加热到80℃,则沸腾现象决不可能发生．这些现象,我们称之为确定性现象．

但是在自然界和社会生活中,也广泛存在着与确定性现象有着本质区别的另一类现象,即所谓随机现象．这种现象的特点是,在一定条件下,既可以发生这样的结果,也可以发生那样的结果．比如,掷一枚硬币,观察哪面向上,结果可能正面向上,也可能反面向上;又如,在抽样检查产品质量时,如果任意从被检产品中抽100件,那么其中的次品可能有5件,可能有10件,也可能只有1件;再如,汽车行驶到马路的交叉路口,可能遇到红灯,也可能遇到绿灯,等等．

随机现象的这种不确定性是否表明随机现象就完全杂乱无章、不可捉摸呢?人们经过长期的实践和研究发现,如果重复地进行大量的观察或试验,任何随机现象都会呈现出一定的规律性,我们称之为统计规律性．

概率论与数理统计就是揭示和研究随机现象的统计规律性的一门学科,它是近代数学的重要组成部分,同时也是我们探讨与研究经济问题的重要数学工具．本章将介绍概率论的一些基本概念,并讨论某些特殊场合下的概率计算问题．

§1.1　随机事件及其关系与运算

一、随机试验与随机事件

对随机现象的研究必然要联系到对客观事物进行观察或试验，若这种观察或试验满足：

(1) 在相同条件下可重复进行；

(2) 全部可能结果不止一个，且事先知道这些可能会出现的结果；

(3) 每次观察或试验都会出现上述结果中的某一个，但具体哪一个结果出现却事先无法预知

则称其为随机试验，简称试验，记为 E.

随机试验的结果称为**随机事件**，简称**事件**，用大写的英文字母 A、B、C 等来表示．

例 1　掷一颗骰子，观察其出现的点数，则

A = "出现 1 点"；B = "出现 2 点"；C = "出现 3 点"；

D = "出现 5 点"；F = "出现奇数点"；G = "点数不超过 2"

都是随机事件．

对于一个随机试验，我们首先关心的是，其所有可能出现的基本结果，它们是试验中最简单的、不可再分解的随机事件，我们称之为基本事件．不是基本事件的事件称作复合事件，它是由基本事件组合而成的．比如在例 1 中，A，B，C，D 就都是基本事件，而 F 则是复合事件，它由 A，C，D 组合而成；G 也是复合事件，它由 A 和 B 组合而成．

此外，我们将试验中必定要发生的事件称为**必然事件**，而将试验中必定不会发生的事件称作**不可能事件**．虽然从本质上说它们已不具有随机性，但为了讨论起来方便，通常把它们视作两个特殊的随机事件，分别用希腊字母 Ω 和 ϕ 表示．在例 1 的试验中，"点数小于 7" 显然是必然事件，而"点数大于 6" 则显然为不可能事件．

二、样本空间

为了使对事件的描述更加准确，也为了使我们将要讨论的事件间的关系和运算比较直观而易于理解，我们引入样本点和样本空间的概念．

称试验 E 的每一个基本事件为一个**样本点**，并以小写希腊字母 ω 记之；称全体样本点所构成的集合为试验 E 的**样本空间**，记之为 Ω.

在例 1 中，令 ω_i = "出现 i 点"，$(i = 1,2,3,4,5,6)$，则有

$$\Omega = \{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$$

或简记为

$$\Omega = \{1,2,3,4,5,6\}$$

于是事件 A 就由一个样本点 ω_1 组成，即 $A = \{\omega_1\}$ 或 $A = \{1\}$. 而因为当且仅当掷出 1,3,5 三种点数的任何一种时，事件 F = "出现奇数点"发生，所以 F 含有 3 个样本点 $\omega_1,\omega_3,\omega_5$，即 $F = \{1,3,5\}$. 说 F 发生了，则意味着出现了 1,3,5 三个点数的某一点数；反之若出现了 1,3,5 中的某一点数，则 F 就发生了．

例 2 将一枚硬币连掷两次，观察所出现的正反面的情况，则该试验有四个样本点，样本空间为

$$\Omega = \{(\text{正},\text{正}),(\text{正},\text{反}),(\text{反},\text{正}),(\text{反},\text{反})\}$$

若以 A 表示"两次掷出相同的一面"这一事件，则 A 包含两个样本点(正,正),(反,反). 而事件 C = "至少掷出一个正面"则包含了三个样本点(正,正),(正,反),(反,正).

例 3 一名射手向某目标射击，直至命中目标为止，观察其命中目标所进行的射击次数(从理论上讲，只要没击中目标，射手就会一直射击下去)，则样本空间为

$$\Omega = \{1,2,\cdots,n,\cdots\}$$

若以 A 表示事件"至少射击 3 次才命中目标"，则 A 发生即意味着射击次数是 3，或 4，或 5，…；而若射击次数为 3，或 4，或 5，…，则意味着事件 A 发生了．这表明事件 A 由样本点"3"，"4"，"5"，… 构成．

例 4　设某公共汽车站每 10 分钟有一辆公共汽车通过，观察某人到达汽车站后的等车时间，则样本空间为

$$\Omega = \{x \mid 0 \leqslant x \leqslant 10\}$$

若令 A 表示“等车时间在 1 分钟到 2 分钟内”这一事件，则

$$A = \{x \mid 1 < x < 2\}$$

A 显然也是 Ω 的一个子集．

例 5　向数轴上投掷一个质点，观察其落点的坐标，由于实数集中任何一个数都是一个样本点，所以

$$\Omega = \{x \mid x \in R\}$$

对上述问题的分析表明，作为试验的结果的事件，其实就是试验的样本空间的子集．

事实上，对于含有有限个或可列个样本点的样本空间（可称离散样本空间），可以将其任意一个子集称作事件．而对于含有不可列个样本点的样本空间而言，由于其子集的复杂性，使得人们不能够将其任意子集都称作事件，而是将满足某种条件的那些子集称作事件．因此，我们一般讲，事件是样本空间的满足某种条件的子集，事件发生当且仅当该子集中某个样本点在试验中出现．

样本空间 Ω 本身也是一个事件，它是一个必然事件，这是因为每次试验总会出现全部基本事件的某一个，即样本点之一总会出现．正基于此，样本空间与必然事件使用了同一个记号 Ω．

三、事件的关系和运算

如前所述，事件即是样本点的集合．因此，集合的包含与相等等概念以及集合的诸种运算也应当同样适用于事件．为此，先要搞清事件的包含、相等及诸种运算的概率论涵义．

1．事件的包含

作为集合，$A \subset B$ 即“A 包含于 B”是说，A 中元素都在 B 中．于是，作为事件，既然 A 的样本点都在 B 中，那么当 A 的样本点出现于试验之中即 A 发生时，B 当然也就发生了．可见 $A \subset B$ 表示“A 的发生必然导致 B 的发生”．

例 6　在掷一颗骰子的试验中，若设 $A=\{1,2,3\}$，$B=\{2,3,4\}$，$C=\{1,2,3,4\}$，则有 $A\subset C$ 且 $B\subset C$.

对任何事件 A，显然总有 $A\subset\Omega$. 又，为了以后讨论的方便，我们约定不可能事件 ϕ 包含于任何事件 A 中，即 $\phi\subset A$. 于是，总有 $\phi\subset A\subset\Omega$.

2. 事件相等

作为集合，$A=B$ 是说"$A\subset B$ 且 $B\subset A$". 于是，作为事件，$A=B$ 就是"A 的发生必导致 B 的发生，同时 B 的发生也必导致 A 的发生". 相等的事件含有相同的样本点 . 事件相等又称事件等价 .

3. 事件的并(和)

作为集合，$A\cup B$ 中的元素或者属于 A，或者属于 B(当然，有的可能同时属于 A,B). 所以事件的并 $A\cup B$ 表示"A 或 B 至少有一个发生".

不难看出，$A\cup B$ 由 A 与 B 的所有样本点构成 .

例如，在例 6 中，$A\cup B=\{1,2,3,4\}=C$

4. 事件的交(或乘积)

作为集合，$A\cap B$ 由 A,B 的公共元素组成 . 所以，事件的交 $A\cap B$ 表示"A,B 同时发生". $A\cap B$ 也简记为 AB.

显然，$A\cap B$ 由 A 与 B 的全部共同样本点构成 .

例如，在例 6 中，$A\cap B=\{2,3\}$

事件的并和交的概念可以推广到有限个和可列个事件的情形

n 个事件 $A_i(i=1,2,\cdots n)$ 的并 $A_1\cup A_2\cup\cdots\cup A_n$ 为一个新事件，简记为 $\bigcup_{i=1}^{n}A_i$，当且仅当"$A_1,A_2,\cdots,A_n$ 中至少有一个发生"时，该事件发生 .

n 个事件 $A_i(i=1,2,\cdots n)$ 的交 $A_1\cap A_2\cap\cdots\cap A_n$ 为一个新事件，简记为 $\bigcap_{i=1}^{n}A_i$，当且仅当"$A_1,A_2,\cdots,A_n$ 都发生"时，该事件发生 .

类似地，可定义可列个事件的并 $\bigcap_{i=1}^{\infty}A_i$ 及可列个事件的交 $\bigcap_{i=1}^{\infty}A_i$.

5. 逆事件(对立事件)

作为集合,A 的余集 $\bar{A}$ 由全集 Ω 中所有不属于 A 的元素组成．于是,作为事件,$\bar{A}$ 当且仅当 A 不发生时发生,称作事件 A 的逆事件．利用上述事件的并和交的运算符号,有

$$A \cup \bar{A} = \Omega \text{ 及 } A\bar{A} = \phi$$

显然,A 与 $\bar{A}$ 互为逆事件．

例如,在掷骰子的试验中,若 $A = \{1,3\}$,则 $\bar{A} = \{2,4,5,6\}$;若 A = "点数小于 3",则 $\bar{A}$ = "点数不小于 3".

6. 事件的差

作为集合,A 与 B 的差集 $A - B$ 由 A 中全部不属于 B 的元素组成．于是,事件 A 与 B 的差 $A - B$ 表示"A 发生而 B 不发生". 显然有 $A - B = A\bar{B}$

7. 互斥事件

集合论中,若 $AB = \phi$,则表明 A,B 没有公共元素．作为事件,若 $AB = \phi$,则表明 A,B 不同时发生,称 A 与 P 互斥(或不相容).

显然,若 A 与 B 互为逆事件,则 A,B 一定互斥,但互斥事件不一定互为逆事件．

例如,在掷骰子的试验中,若令 A = "点数小于 3",B = "点数大于 3",则 A 与 B 互斥,但 A 与 B 并不互逆．

事件的关系与运算,可用集合论中的文(ven)图直观地予以表示(如图 1.1).

与集合的运算一样,事件的运算也满足如下规律:

(1) 交换律:$A \cup B = B \cup A, A \cap B = B \cap A$;

(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:$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,

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

(4) 对偶律:$\overline{A \cup B} = \bar{A} \cap \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}$.

上述运算律也可以推广到任意有限个或可列个事件的情形．例如,对可列个事件 $A_i (i = 1,2,\cdots,n,\cdots)$,有对偶律

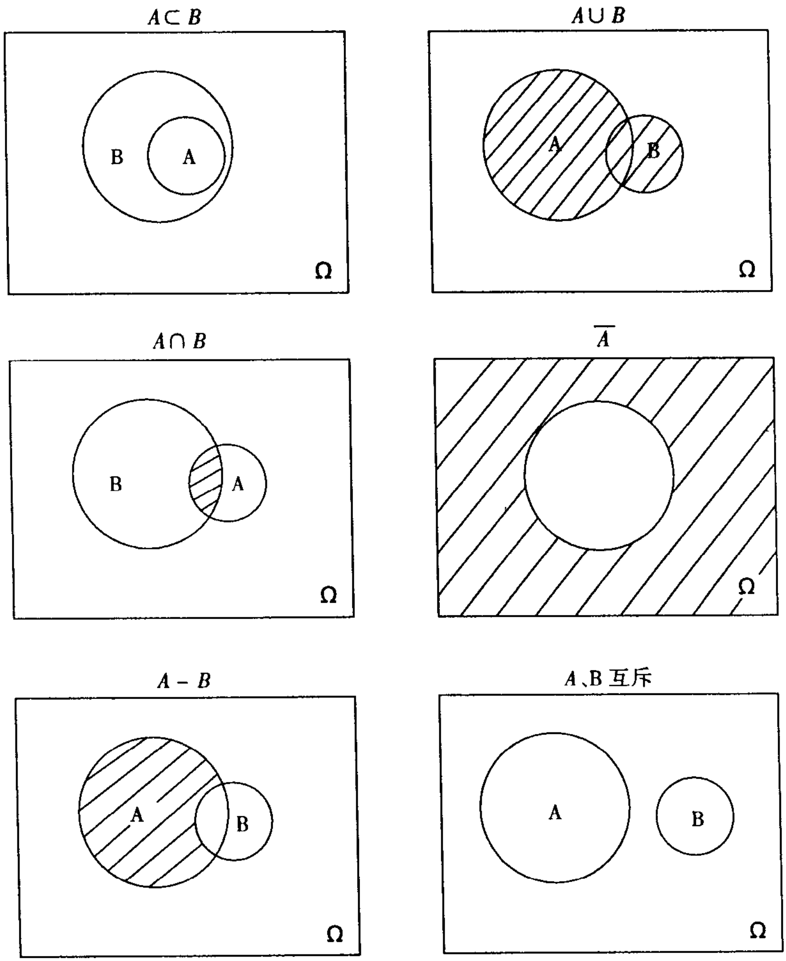

图 1.1

$$\overline{\bigcup_{i=1}^{\infty} A_i} = \bigcap_{i=1}^{\infty} \overline{A_i}, \overline{\bigcap_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \overline{A_i}$$

例 7 从某厂的产品中随机抽取三件产品．设 A 表示“三件中至少有一件是废品”，B 表示“三件中至少有两件是废品”，C 表示“三

件都是正品”,问:$\bar{A},\bar{B},A\cup B,A\cap B,A\cup C,A\cap C,A-B$ 各表示什么事件?

解 $\bar{A}=C$ 表示“三件都是正品”;$\bar{B}$ 表示“三件中至多有一件是废品”;$A\cup B=A$ 表示“至少有一件是废品”;$A\cap B=B$ 表示“至少有两件是废品”;$A\cup C=\Omega$ 为必然事件;$A\cap C=\phi$ 为不可能事件;$A-B$ 表示“恰有一件废品”.

例8 某人连续三次购买彩票,每次一张.令 A,B,C 分别表示其第一,二,三次所买的彩票中奖的事件,试用 A,B,C 及其运算表示下列事件:

(1) 恰有一次中了奖;

(2) 不止一次中奖;

(3) 只有第二次中奖;

(4) 一次奖也未中;

(5) 至多中奖两次.

解(1)$A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$;

(2)$AB\cup BC\cup AC$;

(3)$\bar{A}B\bar{C}$;

(4)$\bar{A}\bar{B}\bar{C}$ 或$\overline{A\cup B\cup C}$;

(5) $\overline{ABC}$ 或$\bar{A}\cup\bar{B}\cup\bar{C}$.

§1.2 概率

一、频率与概率

对于一个随机试验,我们不仅关心它可能出现那些结果,更需要知道某些结果出现的可能性大小.我们希望用一个数字来度量试验中某事件 A 发生的可能性大小,并将这个表征可能性大小的数字称之为事件 A 的概率,记作 $P(A)$.

如何选择这种度量呢?什么样的数能够准确地描述事件发生的可能性大小?

考虑这样一个问题，设 A 是试验 E 的一个可能结果，若在相同条件下将试验 E 连作 100 次，结果事件 A 发生了 90 次，你一定会自然地据此认为，事件 A 发生的可能性是比较大的，因为在总的 100 次试验中，A 发生的次数占了 90%，如果 A 发生的可能性不大，似乎不应该出现这种结果．就是说，人们会认同，在 n 次试验中 A 发生的次数 n_A 与 n 的比值在一定程度上反映了事件 A 发生的可能性的大小．

定义 1.1　称在相同条件下所做的 n 次试验中事件 A 发生的次数 n_A 为 A 发生的频数，并称比值 $\frac{n_A}{n}$ 为 A 发生的频率，记作

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

试验表明，事件 A 的频率不是一个固定的数，它会随着试验次数 n 的变化而变化．即使 n 不变，在不同的两轮 n 次试验中，由于 n_A 可能不同，$f_n(A)$ 也不一定相同．但反映事件 A 发生可能性大小的概率 $P(A)$ 应是客观存在的一个数，它应和所进行试验的次数无关．因此，用 $f_n(A)$ 作为 A 发生的可能性大小的度量是有缺陷的．

但是，人们通过大量的实践观察到，当试验的次数不断增大时，事件 A 的频率会逐步趋于稳定．在概率论的发展历史上，我们的前人曾做了大量的有关频率稳定性的试验．表 1.1 列举了历史上一些著名学者掷硬币试验的记录．记录表明，当掷币次数 n 足够大时，掷出正面的频率稳定地在常数 0.5 的上下摆动，而且随着 n 的不断增大，频率摆动的幅度在逐渐减小．频率的这种稳定性，正是随机现象统计规律性的典型表现．

表 1.1

试验者	掷币次数	频数 n_A	频率 $f_n(A)$
蒲丰	4040	2048	0.5069
费勒	10 000	4979	0.4979
皮尔逊	12 000	6019	0.5016
皮尔逊	24 000	12 012	0.5005

从实践中观察到的频率的稳定性令人们推测，对任何事件 A，都应该有一个由 A 自身所决定的数 p 存在，频率 $f_n(A)$ 正是以这样一个客观存在的数为中心作上下摆动的，而且随着 n 的增大，这种摆动的幅度趋于减小．数 p 的客观存在性以及频率的稳定性有待严格的理论证明，我们将在§3.8中对此问题作进一步讨论．

定义 1.2(概率的统计定义)　在相同条件下所做的 n 次试验中，事件 A 发生的频率 $f_n(A)$ 当 $n\to\infty$ 时稳定在常数 p 的附近，称此常数 p 为事件 A 发生的概率，记作 $P(A)=p$.

定义 1.2 的重要性在于，它提供了一种估计概率的方法，即当 n 充分大时，可以取频率作为概率的估计值．在许多实用问题中，当概率不易求出时，往往就是这样做的．

由频率的定义，可得频率的如下性质：

1°　对任何事件 A，有 $0\leqslant f_n(A)\leqslant 1$;

2°　$f_n(\Omega)=1$，Ω 为必然事件；

3°　若 $A_1,A_2,\cdots,A_m$ 为两两互斥的 m 个事件，则有

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

证明　1°、2°显然成立，现在证明 3°.

设在 n 次试验中，事件 $A_1,A_2,\cdots,A_m$ 发生的频数分别是 n_{A_1}，$n_{A_2},\cdots,n_{A_m}$，由于 $A_1,A_2,\cdots,A_m$ 两两互斥，因此在 n 次试验中事件 $\bigcup\limits_{i=1}^{m}A_i$ 发生的频数等于诸 $A_i(i=1,2,\cdots,m)$ 各自发生的频数之和，即 $n_{A_1}+n_{A_2}+\cdots+n_{A_m}$，从而

$$f_n\left(\bigcup_{i=1}^{m}A_i\right)=\frac{n_{A_1}+n_{A_2}+\cdots+n_{A_m}}{n}=\sum_{i=1}^{m}f_n(A_i)$$

概率的统计定义虽然一般而且直观，但不符合数学的严密性，况且要通过此定义确定出事件 A 发生的概率是困难的．因此有必要寻找更好的定义概率的方法．

下面将给出的概率的公理化定义是大数学家柯尔莫哥洛夫于