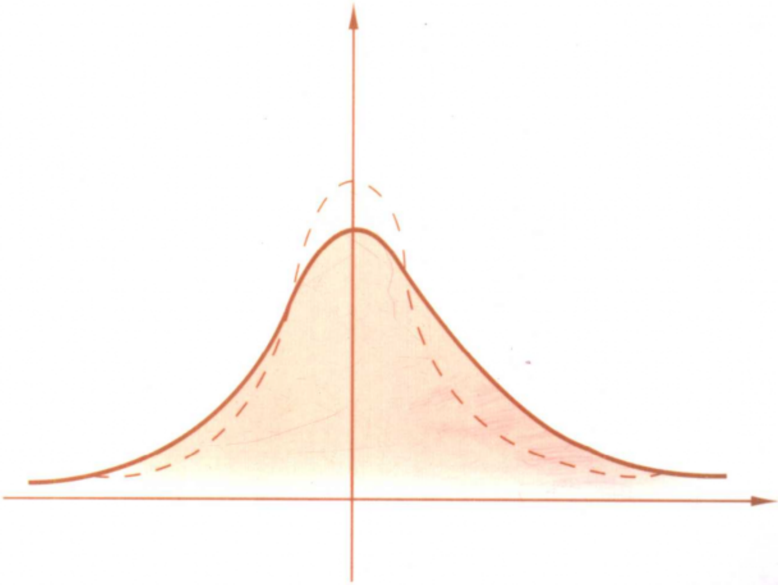
高等教材

全国高等农林院校教材

数理统计（第4版）

贾乃光 张 青 李永慈 编著



中国林业出版社

中国林业出版社教材建设与出版管理中心
责任编辑／杜建玲　　封面设计／忠信工作室

ISBN 7-5038-3668-7

9 787503 836688 >

ISBN 7-5038-3668-7
定价：25.00元

全国高等农林院校教材

数理统计

（第4版）

贾乃光　张　青　李永慈 编著

中国林业出版社

图书在版编目（CIP）数据

数理统计/贾乃光，张青，李永慈编著．—4 版．—北京：中国林业出版社，2006．7

（全国高等农林院校教材）

ISBN 7-5038-3668-7

Ⅰ．数…　Ⅱ．①贾…②张…③李…　Ⅲ．数理统计－高等学校-教材　Ⅳ．O212

中国版本图书馆 CIP 数据核字（2006）第 068123 号

中国林业出版社·教材建设与出版管理中心

策划编辑：牛玉莲　　**责任编辑**：杜建玲

电话：66188720　66170109　　**传真**：66170109

出版发行　中国林业出版社（100009　北京市西城区德内大街刘海胡同 7 号）
　　　　　E-mail：cfphz@public.bta.net.cn　电话：（010）66184477
　　　　　网　址：http：//www.cfph.com.cn

经　　销　新华书店

印　　刷　北京市昌平百善印刷厂

版　　次　1980 年 10 月第 1 版
　　　　　2006 年 7 月第 4 版

印　　次　2006 年 7 月第 19 次印刷

开　　本　850mm×1168mm　1/16

印　　张　18.5

字　　数　391 千字

定　　价　25.00 元

凡本书出现缺页、倒页、脱页等质量问题，请向出版社图书营销中心调换。

前言

什么是数理统计？一方面它以观测（或搜集）到的样本数据（data）为基础；另一方面以经验和其他各方面的知识提出一个模型为基础，对总体（population）做描述、参数估计及检验（决策）（description、parameter estimation & test——decision）和预报（prediction）。

本课程是以它前行课概率论（probability）为基础的，后续课程有多元统计（multivariate statistical analysis）、统计决策论（statistical decision theory）、随机过程（random process）等。

抽样比全面观测当然省时、省力、省经费，然而由于哪些单元被抽中而成为样本是有偶然性（随机性）的，这就产生抽样误差（sampling error）。那么，抽样误差的大小与哪些因素有关，如何减少它，它对统计结论产生什么影响等问题就成为本课程贯彻始终的内容。

作为教材必须考虑哪些是最基本的原理和方法，结合学时数慎重取舍。根据因材施教和对应用本书内容较多的专业，对此课更感兴趣的学生设立带星号的内容、注和附录，使所授内容有一定的弹性（灵活性 flexibility）。

数理统计的内容近来有长足的发展。我们目录中的每一章差不多都发展为一个分支，如《抽样技术》、《实验设计》、《正交设计》、《近代回归分析》、《非参数统计方法》等。本教材的内容，可以说是入门，在实际应用中肯定还应参考相应的专著。

在我们统计方法的背后，有坚实的数学理论基础。我们涉及的所有的统计量所服从的分布，都有严格的数学证明。这里，在应用中有一个问题：我们说某一变量 X 是正态变量，当 Z 代表的是现实中的某一变量时，它就不大可能是严格正态的，只能是近似正态的。那么，$(Z-EZ)/\sigma(Z)$ 是标准正态也是近似的，而且 EZ 和 $\sigma(Z)$ 又往往是不知道的，用 $\bar{Z}$ 及 S 代替，这就又增加了一层近似，从而其平方为 $\chi^2(1)$ 分布就包含了这些近似。所以，只有严格的理论基础还是不够的，还要能在现实中使许多的近似尽量做得好、近似得好。

当然，在抽样中又出现了抽样误差，在计算中还有计算误差。所有这些近似和误差都加在一起之后形成怎样的情形，这里实践经验及在每一个步骤中细节的慎重就是很重要的了。所以，在应用中，应该注重反复的实践，和在实践中的思考、经验的总结。

美国的一些教材是值得参考的，它们更注重操作方法和对理论的直观理解，而且淡化理论的数学证明和充满数学特色的陈述。比如，连续型随机变量 X 的

定义，大多数中文教材都表述为 X 的分布函数 $F(x)$ 可表述为

$$F(x) = \int_{-\infty}^{x} \varphi(t)\,\mathrm{d}t$$

可积函数 $\varphi(x)$ 被称为 X 的概率密度。而许多美国非数学、非统计专业的书籍都避免以上充满数学美的定义，而更看重于让学生更易于接受，顺利和自然地理解。

本书无星号的内容是最基本、简要的部分，适于约 40 学时的教学内容；加带“*”部分适于约 70 学时的教学；带“**”部分适于约 100 个学时。带“#”号的内容为编者个人的理解，仅供参考。每章中的注，多为数学性较强的内容，供有兴趣的读者参考；每章的附录属于内容上的扩张，大多安排一段外文教材中的内容，便于与中文内容对比及专业英语的学习。

编者学识浅陋，错误、不妥之处望读者及同行们多多指教。

编 者

2005 年 11 月

目　录

前　言

第 1 章　概率论

【本章提要】　正如几何学是测地学的理论基础，概率论是统计学的理论基础。统计方法和结论的依据是样本、是抽样，于是就有抽样误差，它的描述、计算和控制方法即要研究概率、随机变量及其分布、常用的一些分布、分布的数字特征、大数定律及中心极限定理。

§1.1　随机事件与集合

1.1.1　随机事件

今后我们常举两个例：掷骰子和打靶。前者只可能出现 6 种可能，即出现的点数可能为 $\{1,2,3,4,5,6\}$，它是全集、有限集，所有可能出现的事件都是全集的子集。如‘出现奇数点’这一事件即 $\{1,3,5\}$；‘出现点数小于等于 2’这一事件即集合 $\{1,2\}$ 等等。在掷一颗骰子这一随机试验中，有可能出现的事件被称为随机事件。

定义　在某一随机试验中有可能出现、也可能不出现的事件被称为随机事件，或简称为事件。用 $A,B,C,\cdots$ 等表示。

如果把试验可能出现的基本事件（即不能再分解得更小的事件）组成的集合称为全集 Ω，则事件就是 Ω 的子集。

空集 $\varnothing$ 也是 Ω 的子集，与 $\varnothing$ 相应的事件我们称为不可能事件，即试验之前我们已经知道此事件不可能出现。如掷骰子中出现‘点数大于 7’的事件、‘点数小于 1’的事件、‘点数为 1.25’的事件等都是不可能事件。不可能事件用 $\varnothing$ 表示。

全集 Ω 自己也是 Ω 的子集，与 Ω 相应的事件我们称为必然事件，即试验之前我们已肯定知道此事件必然出现。如掷骰子中出现‘点数为整数’的事件；‘点数在区间 $[0,6]$ 内’的事件；‘点数为 1 或 2 或 3 或 4 或 5 或 6’的事件等等皆为必然事件。必然事件用 Ω 表示，有时也用 U 表示。

我们常举的第 2 个例子是打靶。如图 1.1 所示。假设打中的是图中正方形的一个点。此时全集 U 为无穷集，图中的 A、B 皆为 U 的子集，是随机事件（此例不考虑脱靶）。按图 1.1 则‘既打中 A 又打中 B’的事件为一个不可能事件。

这里，我们将事件与集合建立了对应的关系：

事件$\longleftrightarrow$集合　记为 $A,B,C,\cdots$

必然事件⟷全集 记为 U 或 Ω

不可能事件⟷空集 记为 $\varnothing$

于是有:任意的事件 A,则

$$\varnothing \subset A \subset \Omega \tag{1.1}$$

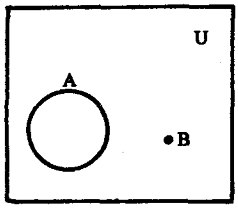


图 1.1

1.1.2 事件之间的关系与运算

1. 包含关系

事件 A 包含事件 B,记为 $A \supset B$;或事件 B 被事件 A 包含,记为 $B \subset A$。

$A \supset B$ 或 $B \subset A$ 均表示若有元素 $\omega \in B$,则必有 $\omega \in A$。如图 1.2 所示。

或者说,若事件 B 发生,则事件 A 一定发生。如在掷一骰子中,若 B 为‘出现3点’的事件,A 表示‘出现奇数点’的事件,则 B 发生肯定也有 A 发生,所以 $B \subset A$。

注意,我们有

$$A \subset A \tag{1.2}$$

即包含关系是自返的。

在公式(1.1)中,$\varnothing \subset A$,可以用反证法证明(请读者自己练习并作为习题)。

2. 事件的相等 $A = B$

若 $A \subset B$ 且又有 $B \subset A$,则称 A、B 相等,记为 $A = B$。即事件 A 若发生,必有事件 B 发生。反之亦然。

3. 事件的和(或并)$A + B$

事件 A、B 中至少一个发生的事件被称为事件 A、B 的和,记为 $A + B$。如图 1.3 阴影部分所示。

同样事件 $A_1, A_2, \cdots, A_m$ 中至少一个发生的事件被称为 $A_1, A_2, \cdots, A_m$ 的和,记为 $A_1 + A_2 + \cdots + A_m$。显然有:

$$A + B = B + A;(A + B) + C = A + (B + C) \tag{1.3}$$

$$A + A = A;A + \varnothing = A;A + U = U \tag{1.4}$$

$$\sum_{i=1}^{m} A_i = A_1 + A_2 + \cdots + A_m$$

例如图 1.4 表示电路的两开关并联时,至少一个开通电路的事件。

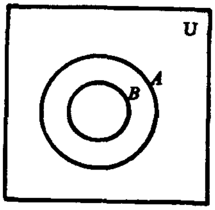


图 1.2

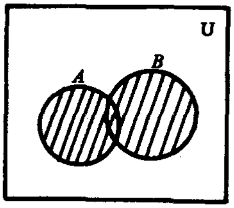


图 1.3

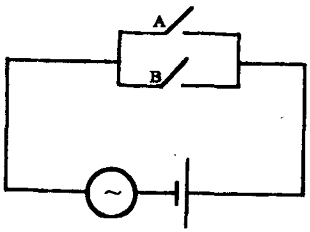


图 1.4

又如,在掷骰子时,若 A 表示'出现 1,3,4'的事件;B 表示'出现 2,4'的事件,则 $A+B$ 为'出现 1,2,3,4'的事件。

4. 事件的积(或交)$A \cdot B$

事件 A、B 同时发生的事件被称为 A、B 的积,记为 $A \cdot B$。如图 1.5 中的阴影部分所示。

同样,$A_1, A_2, \cdots, A_m$ 的积为同时发生的事件,记为 $A_1 \cdot A_2 \cdot \cdots \cdot A_m$。显然有

$$A \cdot B = B \cdot A;\ (A \cdot B) \cdot C = A \cdot (B \cdot C) \tag{1.5}$$

$$A \cdot A = A;\ A \cdot \varnothing = \varnothing;\ A \cdot U = A \tag{1.6}$$

$$\prod_{i=1}^{m} A_i = A_1 \cdot A_2 \cdot \cdots \cdot A_m$$

在和与积之间还容易证明有以下的分配律:

$$(A+B) \cdot C = A \cdot C + B \cdot C \tag{1.7}$$

$$A \cdot (B+C) = A \cdot B + A \cdot C \tag{1.8}$$

另外还有(证明留给读者):

$$A \subset A+B;\ B \subset A+B \tag{1.9}$$

$$A \cdot B \subset A;\ A \cdot B \subset B \tag{1.10}$$

例如,图 1.6 表示电路两开关 A、B 串联时,只有 A、B 同时开通,线路才能通的事件。

图 1.7 中,3 个事件 A、B、C 的交的事件为图中阴影部分。

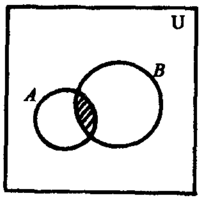


图 1.5

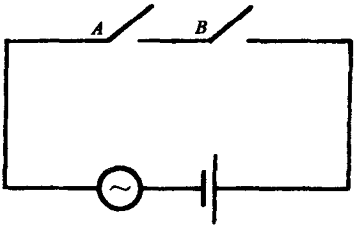


图 1.6

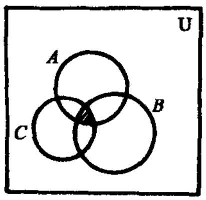


图 1.7

5. 事件的差 $A-B$

事件 A 发生但事件 B 不发生的事件被称为 $A-B$,如图 1.8 阴影部分所示。

如在掷一骰子的试验中,设 A 为'出现 1,2,3 点'的事件,B 为'出现 3,4 点'的事件,则 $A-B$ 为'出现 1,2 点'的事件。即将事件 A 中同时属于 B 的部分挖去。显然有

$$A-B \neq B-A;A-A = \varnothing;A-\varnothing = A \tag{1.11}$$

$$A-B = A-A \cdot B;A-U = \varnothing \tag{1.12}$$

$$\left.\begin{aligned}(A-B)+B &\neq A(\text{左端} = A+B)\\ (A+B)-B &\neq A(\text{左端} = A-B)\end{aligned}\right. \tag{1.13}$$

$$A-(B+C) = (A-B)-C \tag{1.14}$$

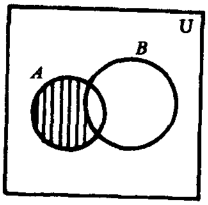


图 1.8

6. 事件的补(或逆)$\overline{A}$

事件 A 未发生也是一个事件,被称为 A 的补或逆,记为 $\overline{A}$,如图 1.9 阴影部分所示。

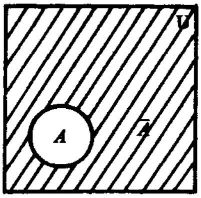


图 1.9

显然

$$\overline{A}=U-A;A-B=A\cdot\overline{B} \tag{1.15}$$

$$\overline{\varnothing}=U;\overline{U}=\varnothing;\overline{\overline{A}}=A \tag{1.16}$$

式(1.17)被称为对偶律(或摩尔律):

$$\overline{A+B}=\overline{A}\cdot\overline{B};\overline{A\cdot B}=\overline{A}+\overline{B} \tag{1.17}$$

例如,事件 A 为某一班内会英语的同学,事件 B 为会日语的同学,则 $A+B$ 为英、日语至少会一门的同学,$\overline{A+B}$则为两门都不会的同学,即 $\overline{A}\cdot\overline{B}$。

又如,在掷一骰子中,设 A 由'1,2,3 点'组成,B 由'3,4 点'组成,则 $A+B$ 由'1,2,3,4 点'组成,于是$\overline{A+B}$由'5,6 点'组成。而 $\overline{A}$ 由'4,5,6 点'组成,$\overline{B}$ 由'1,2,5,6 点'组成,故 $\overline{A}\cdot\overline{B}$ 也由'5,6 点'组成。

7. 事件的互斥(或互不相容)

若 $A\cdot B=\varnothing$,则称 A、B 互斥或互不相容。如图 1.10 所示。

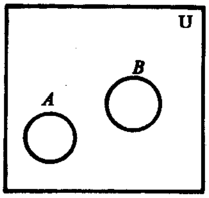


图 1.10

A 与 $\overline{A}$ 被称为对立的事件。则有:

(1)两事件对立,则必互斥。但反之未必。

这是因为,若 A、B 对立,即 $B=\overline{A}$,有 $A+B=U$,且 $A\cdot B=\varnothing$,后者说明 A、B 互斥。反之 A、B 互斥,有 $A\cdot B=\varnothing$,但未必有 $A+B=U$,所以未必对立。

(2)若 A、B 互斥,则 $\overline{A}$、$\overline{B}$ 未必互斥。举一反例:如在掷一骰子中,设 A 为'出现 1 或 2 点';B 为'出现 5 或 6 点',显然 A、B 互斥。而 $\overline{A}\cdot\overline{B}$ 为'出现 3 或 4 点'并非不可能事件,所以 $\overline{A}$、$\overline{B}$ 不互斥,是相容的。

(3)若 A、B 互斥,则 $A\subset\overline{B}$,$B\subset\overline{A}$。

证明 ①若 $A=\varnothing$,$\varnothing$属任何事件,故$\varnothing$属 $\overline{B}$,即 $A\subset\overline{B}$ 成立。

②若 $A\neq\varnothing$,任 $\omega\in A\to\omega\notin B\to\omega\in\overline{B}$,所以 $A\subset\overline{B}$ 成立。

(4)若 A、B 互斥,则 $\overline{A}+\overline{B}=U$。

此题之证明留给读者作为思考题。

8. 互斥事件完备群

若 $A_1,A_2,\cdots,A_m$ 这 m 个事件满足以下两条件,则称这 m 个事件为互斥事件完备群。如图 1.11 所示。

$$\begin{cases}A_1+A_2+\cdots+A_m=U\\ A_i\cdot A_j=\varnothing\ (i\neq j;i,j=1,2,3,\cdots,m)\end{cases} \tag{1.18}$$

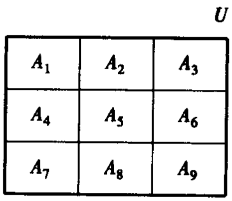


图 1.11

【例 1.1】 设在某一图书馆,令 A 表示数学类书,B 表示中文书,且已知非数学类书都是中文的。问是否外文书都是数学类书?

解 非数学类书为 $\overline{A}$,都是中文的,即

$$\overline{A} \subset B(\text{由于若} A \subset B \text{可推出} \overline{B} \subset \overline{A})$$

所以

$$\overline{B} \subset \overline{\overline{A}} = A$$

而 $\overline{B}$ 为外文书,上式说明外文书都是数字类书。

【例 1.2】 是否有 $A-(B-C)=(A-B)+C$ 或 $A-(B-C)=A-B-C$

解 以上等式都不成立。可举反例如下:在掷一骰子中,令 $A=\{1,2\}$,$B=\{2,3,4\}$,$C=\{1,4,5\}$。

上式左端 $A-(B-C)=\{1\}$,而右端为 $\{1,4,5\}$,两者不等;第二个式子的右端为 $\varnothing$,也不等于 $\{1\}$。

§1.2 概率的定义

某事件 A 的频率:设同一试验被重复进行 n 次,其中有 m 次 A 出现,$n-m$ 次 A 未出现,则称 m/n 为 A 的频率。

定义 以上当 n 充分大时,一般地有:A 的频率愈来愈稳定,其稳定中心被称为 A 的概率,记为 $P(A)$。

以上若 n 充分大时,A 的频率并不趋于稳定,则 A 并无概率可言(如混沌状态即是)。

这里的 $P(A)$ 是事件 A 发生的客观的可能性的大小,与试验无关;而频率是与具体的试验有关的。一个与试验有关的量 m/n,当 n 充分大之后其稳定中心变得与试验无关了,这充分显示了 n 充分大所发挥的作用。当 n 小时,两组 n 次试验所得的频率可以相差很大。以上所说正是概率论中的一个本质特征,即我们在本章的最后要介绍的"大数定律"。

虽然 n 大时频率可以摆脱具体试验的偶然性而逐渐显现概率值大小的客观真值,但它的缺点是,依此要做概率的计算变得几乎不可能。古典概型的计算方法可以帮助我们解决这个问题。

§1.3 古典概型

一个随机试验所可能出现的事件若不能再分解为更简单的事件时,这样的事件被称为基本事件,记为 $\omega_1,\omega_2,\cdots$。所有基本事件组成的集合被称为样本空间,记为 Ω 或 U,也就是我们在上一节所讲的全集。

若 Ω 是有限的,由 $\omega_1,\omega_2,\cdots,\omega_n$ 组成,且每个 ω_i 的出现具有等可能性,则

$$P(A)=\frac{A\text{ 中所包含的基本事件的个数}}{n} \tag{1.19}$$

例如,掷一骰子,当骰子质地均匀且充分旋转时,即为古典概型,其中 $\omega_1=\{1\}$,$\omega_2=\{2\}$,$\cdots$,$\omega_6=\{6\}$。若 A 为出现奇数点,则

$$P(A) = 3/6 = 1/2$$

那么,公式(1.19)与我们所定义的概率是频率的稳定中心是否一致呢?不少科学家对掷钱币做了多次试验,证明两者是一致的。他们的试验结果为表1.1。

表1.1 掷钱币试验

试验者	所掷次数	出现正面的次数	出现正面的频率
摩　根[德]	2 048	1 061	0.518 1
蒲　丰[法]	4 040	2 048	0.506 9
皮尔逊[英]	12 000	6 019	0.501 6
皮尔逊[英]	24 000	12 012	0.500 5

古典概型的计算方法公式(1.19)的缺点是要求样本空间Ω是有限集,且要求每一基本事件$\omega_i(i=1,2,3,\cdots,n)$是等可能的。而这两点在我们现实生活中是常常得以满足的。在第2章中我们将知道,我们要研究的总体即Ω往往都是有限的,而我们要求抽样必须是随机的,即保证每一个总体单元被抽中都是等可能的。另外,在概率的频率稳定中心的定义中,试验必须是可重复进行的,而且要能够无限次地重复进行。那么,不能重复进行的试验、不能重复观测的事件就没有概率可言了。而古典概型的概率计算方法公式(1.19)式则避免了这一点,只要理论上认定各ω_i是等可能性的,即可按公式(1.19)计算概率。

【例1.3】 在打靶的例子中,由于靶中的点无穷,样本空间Ω是无限集,故不能按古典概型应用公式(1.19),但若假设靶中每一点被击中的可能性是相同的(即不瞄准打枪,随机地打),则可以如下定义概率(图1.12):

$$P(B) = \frac{B\text{的面积}}{U\text{的面积}} \qquad (1.20)$$

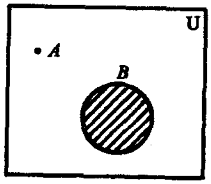


图1.12

若B收缩到一个点成为A,由于A的面积为0,故$P(A)=0$

$P(A)=0$即A为概率为零的事件,但A不是绝对不可能发生的事件,只是一个几乎不可能发生的事件。这说明:不可能事件$\varnothing$的概率为0,但概率为0的事件未必是不可能事件。

【例1.4】 设同一宿舍中有4个同学,4人的生日可能为星期1,2,…,6,日共7个可能,假定它们是等可能的。问此4人恰好在同一个星期几的概率为多少?

解 先求总的可能数。每人在7个可能中任选一个,有7种可能,4个人的搭配数为7^4个可能。

所求的事件设为A,为同一个星期几,在7个可能中只能选一个。按公式(1.19)有

$$P(A) = 7/7^4 = 1/7^3$$

【例1.5】 设有4人同乘一电梯,他们可能在第2,3,…,8层下。问4人恰好在同一层下这一事件A的概率?

解 由例1.4易知,$P(A)=1/7^3$。

通过以上两例让我们看到,完全不同领域的问题归结到数学上变成同一个问题,这就启发我们采用模型的方法将现实问题分类。

§1.4 概率的性质

1. $0 \leqslant P(A) \leqslant 1$ (1.21)

其中 A 为任意的一个事件。因为频率 m/n 中有 $0 \leqslant m \leqslant n$,故频率必在$[0,1]$内,其稳定值 $P(A)$ 也在$[0,1]$内。

2. $P(\varnothing)=0, P(U)=1$ (1.22)

因为每一实验$\varnothing$都不会出现,其频率 m/n 中之 $m=0$,频率为 0,其稳定值 $P(\varnothing)$ 也为 0。类似知 $P(U)=1$。

3. 若 $A \cdot B=\varnothing$,则

$$P(A+B) = P(A) + P(B) \tag{1.23}$$

例如,按图 1.13 知

$$P(A+B)=\frac{A\text{ 的面积}+B\text{ 的面积}}{U\text{ 的面积}}=\frac{A\text{ 的面积}}{U\text{ 的面积}}+\frac{B\text{ 的面积}}{U\text{ 的面积}}$$

$$=P(A)+P(B)$$

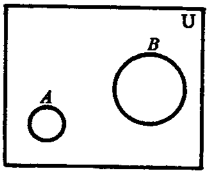


图 1.13

公式(1.23)又被称为概率的可加性。

* 以上 3 个性质可以作为概率的公理定义方法中的 3 个公理。概率的公理定义方法是前苏联柯尔莫廓洛夫提出的。

4. $P(\bar{A})=1-P(A)$ (1.24)

因为 $A+\bar{A}=U$,A、$\bar{A}$ 互斥。由公式(1.23),$1=P(U)=P(A+\bar{A})=P(A)+P(\bar{A})$,再移项即得公式(1.24)。

5. 若 $A \subset B$,则

$$P(B-A) = P(B) - P(A) \tag{1.25}$$

因为,$B=AB+(B-AB)$,而 $AB=A$

所以 $B=AB+(B-A)$

而 AB 与 $B-A$ 互斥,由(1.23)知

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

移项即得公式(1.25)。

6. 加法公式

$$P(A+B) = P(A) + P(B) - P(A \cdot B) \tag{1.26}$$

因为,$A+B=A+(B-AB)$,而 A 与$(B-AB)$互斥(因为 $A \cdot (B-AB)=AB-AB=\varnothing$),由公式(1.23)有

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

再由公式(1.25)得

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

代入上式即得所求。

加法公式如图 1.14 所示。从面积的意义来讲，$A \cdot B$ 的面积被 A 的面积及 B 的面积重复地计算，故相加时应将重复计算部分去除。

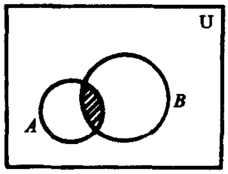


图 1.14

【例 1.6】 一副扑克(52 张)中任抽 13 张。问 13 张中至少有 1 张黑桃的概率？

解 设 A 为 13 张中至少有 1 张黑桃的事件；A_i 为 13 张中恰有 i 张黑桃的事件。于是

$$A = A_1 + A_2 + \cdots + A_{13}$$

上式右端各事件显然两两互斥，多次应用公式(1.23)得

$$P(A) = \sum_{i=1}^{13} P(A_i)$$

而

$$P(A_i) = \frac{C_{13}^{i} C_{39}^{13-i}}{C_{52}^{13}}$$

但此问题计算 $\overline{A}$ 则更简单。$\overline{A}$ 为 13 张中没有黑桃的事件，于是

$$P(\overline{A}) = C_{39}^{13} / C_{52}^{13}$$

再由公式(1.24)，$P(A) = 1 - P(\overline{A})$，即得所求。

§1.5 条件概率及概率乘法公式

例如，设肺癌发病率在成年人中为 3/1 000，在抽烟的成年人中为 1/100。这里，全集 U 为成年人，A 为成年人中之肺癌发病者；B 为成年人中之吸烟者。以上假设可写为

$$P(A) = 3/1\,000; P(A \mid B) = 1/100$$

其中 $A|B$ 表示 B 条件下的 A。以上 3/1 000 为 A 的无条件概率，后者为在 B 条件下 A 的条件概率，显然两者是不同的。

1. 条件概率

定义 若 $P(B) \neq 0$，记 $A|B$ 为在已知 B 已发生的条件下 A 发生的事件，则定义：

$$P(A \mid B) = P(A \cdot B) / P(B) \tag{1.27}$$

2. 概率乘法公式

从公式(1.27)得

$$P(A \cdot B) = P(B) \cdot P(A \mid B) \tag{1.28}$$

若 $P(A) \neq 0$，显然也有

$$P(A \cdot B) = P(A) \cdot P(B \mid A) \tag{1.29}$$

公式(1.28)及公式(1.29)被称为概率的乘法公式。它们还可以推广到多个，如

$$P(A \cdot B \cdot C) = P(A) P(B \mid A) P(C \mid AB) \tag{1.30}$$

因为，上式左端 $= P[(A \cdot B) \cdot C] = P(A \cdot B) P(C|A \cdot B)$，再应用公式(1.29)即得公式(1.30)。