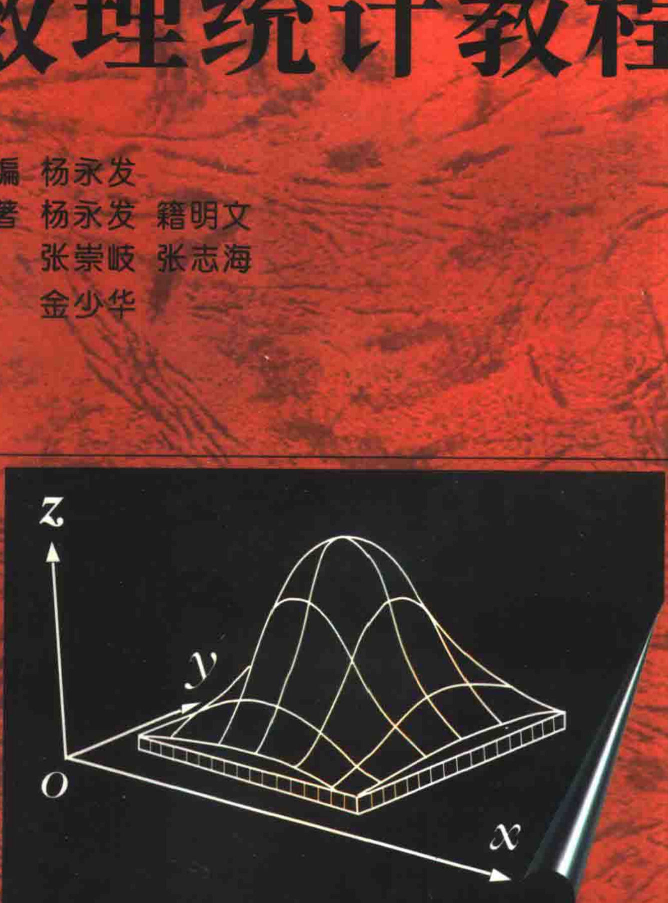
概率论与数理统计教程

主编 杨永发
编著 杨永发 籍明文
张崇岐 张志海
金少华



南开大学出版社

概率论与数理统计教程

主编　杨永发
编著　杨永发　籍明文　张崇岐
　　　张志海　金少华

南开大学出版社

图书在版编目(CIP)数据

概率论与数理统计教程 / 杨永发主编 .—天津:南开大学出版社，2000
ISBN 7-310-01446-4

Ⅰ.概... Ⅱ.杨... Ⅲ.①概率论-教材②数理统计-教材 Ⅳ.021

中国版本图书馆 CIP 数据核字(2000)第 68649 号

出版发行 南开大学出版社
地址:天津市南开区卫津路 94 号
邮编:300071 电话:(022)23508542

出版人 肖占鹏
承　印 天津宝坻第二印刷厂印刷
经　销 全国各地新华书店
版　次 2000 年 11 月第 1 版
印　次 2000 年 11 月第 1 次印刷
开　本 850mm×1168mm 1/32
印　张 12.875
字　数 319 千字
印　数 1— 6000
定　价 18.00 元

内 容 简 介

本书是为工科大学“概率论与数理统计”课程而编写的教材。全书分概率论与数理统计两部分。前4章为概率论部分，内容包括：概率论的基本概念，随机变量及其分布，随机变量的数字特征，大数定律和中心极限定理。后5章为数理统计部分，内容包括：数理统计的基本概念，参数估计，假设检验，方差分析和一元线性回归分析。书末附有名词索引、常用概率分布表及练习题、习题参考答案等。

本书结合工科教学实际，注意理论与实际的结合，选材适当，论述严谨，条理清楚，便于教学及学生自学，可作为高等工科院校本、专科“概率论与数理统计”课程的教材，也适合非数学类理科及管理各专业使用。

前　言

概率论与数理统计是工科大学的一门重要基础课，在培养学生运用概率统计独特的思维方式分析和解决实际问题的能力方面具有重要的作用，并为后继课程和未来的工作实践提供必备的随机数学基础。随着教学改革的深入，数学教学向着加强基础、拓宽应用的方向推进，该课程已经被列为工科大学理、工科及管理类绝大多数专业的必修课，并被列入工学、经济学硕士研究生入学考试的必考科目。我们编写本书，就是为了适应教学改革的发展，希望能为本课程提供一本合适的教学用书。

本书分概率论和数理统计两部分。前 4 章为概率论部分，后 5 章是数理统计部分，概率论是数理统计的基础。本书内容包括该课程教学基本要求规定的全部内容，并可满足硕士研究生入学考试中"数学一"考试的基本要求。

本书结合工科数学教学实际，在选材上注意理论与实际的统一，结合概率与数理统计的基本理论，适当介绍了在社会人文科学、医学、经济管理、化工等领域中的一些应用实例，可以拓宽学生视野、激发学生的兴趣。在叙述上力求论述严谨，条理清楚，便于阅读。每节中穿插有若干练习题，作为基本训练之用。章末附有习题，书后附有练习题及习题的参考答案或简要解答，供教师和学生参考。

作为教材使用，概率论和数理统计各需约 32 学时；若删掉一些内容（如第八、九章，第二章部分内容及一些应用实例），也可以在 48 学时内完成主要部分的教学。

本书第一章由杨永发编写，第二章由籍明文编写，第三、四章由金少华编写，第五、六、七章由张崇岐编写，第八、九章由张志海编写。杨永发统稿并最后定稿。本书的编写，得到了蒋春澜教授的支持和鼓励，于慎根教授、刘国欣教授先后审阅了全书并提出许多很好的建议，南开大学出版社李正明编审对本书的出版倾注了大量的心血，作者对此一并致以由衷的感谢。

我们也感谢“参考书目”中所列著作的诸位作者。本书从他们的著作中吸收了许多优秀思想和若干富有启发性的例、习题及一些生动有趣的应用实例，极大地丰富了本书的内容。

由于水平所限，书中难免有错、漏及欠妥之处，作者热忱地希望使用本书的教师和同学指正，以便再版时修订。

编　者

2000 **年** 4 **月**

目　录

第一章　概率论的基本概念 …………………………… (1)
§1.1　随机现象与随机事件 …………………………… (1)
1.1.1　随机现象 …………………………… (1)
1.1.2　随机试验和样本空间 …………………………… (2)
1.1.3　随机事件 …………………………… (5)
§1.2　概率的定义 …………………………… (11)
1.2.1　随机事件的频率 …………………………… (11)
1.2.2　概率的定义 …………………………… (13)
1.2.3　概率的性质 …………………………… (14)
§1.3　等可能概型 …………………………… (18)
1.3.1　古典概型 …………………………… (18)
*1.3.2　几何概型 …………………………… (24)
§1.4　条件概率 …………………………… (26)
1.4.1　条件概率的概念 …………………………… (26)
*1.4.2　条件概率的性质 …………………………… (30)
1.4.3　全概率公式和贝叶斯公式 …………………………… (32)
§1.5　随机事件的独立性 …………………………… (37)
1.5.1　两个随机事件的独立性 …………………………… (37)
1.5.2　多个随机事件的独立性 …………………………… (39)
1.5.3　n重伯努利试验 …………………………… (42)
习题1 …………………………… (46)
第二章　随机变量及其概率分布 …………………………… (49)
§2.1　随机变量与分布函数 …………………………… (49)

2.1.1 随机变量…………………………………………………… (49)
2.1.2 分布函数…………………………………………………… (51)
§2.2 离散型随机变量………………………………………… (57)
2.2.1 定义与基本概念 ……………………………………… (57)
2.2.2 几种常见的离散型随机变量 ………………………… (60)
§2.3 连续型随机变量………………………………………… (66)
2.3.1 定义与基本概念 ……………………………………… (66)
2.3.2 几种常见的连续型随机变量 ………………………… (70)
§2.4 二维随机向量…………………………………………… (79)
2.4.1 随机向量及其分布函数 ……………………………… (79)
2.4.2 离散型随机向量 ……………………………………… (83)
2.4.3 连续型随机向量 ……………………………………… (87)
2.4.4 二维均匀分布和二维正态分布 ……………………… (92)
*2.4.5 n 维正态分布 ……………………………………… (96)
§2.5 条件分布与随机变量的独立性………………………… (97)
2.5.1 条件分布…………………………………………………… (97)
2.5.2 随机变量的独立性……………………………………… (101)
§2.6 随机变量的函数的概率分布 ………………………… (109)
2.6.1 一个随机变量的函数 ………………………………… (109)
2.6.2 两个随机变量的函数 ………………………………… (115)
习题 2 ……………………………………………………………… (127)
第三章 随机变量的数字特征………………………………… (130)
§3.1 数学期望 ……………………………………………… (130)
3.1.1 离散型随机变量的数学期望 ………………………… (130)
3.1.2 连续型随机变量的数学期望 ………………………… (137)
3.1.3 随机变量的函数的数学期望 ………………………… (141)
3.1.4 数学期望的性质 ……………………………………… (147)
§3.2 方差 …………………………………………………… (150)

§3.3 协方差与相关系数 …………………………… (159)
§3.4 矩和协方差矩阵 ……………………………… (166)
3.4.1 矩 ………………………………………… (166)
3.4.2 协方差矩阵 ……………………………… (167)
习题 3 ………………………………………………… (169)
第四章 大数定律和中心极限定理 ……………… (172)
§4.1 大数定律 ……………………………………… (173)
§4.2 中心极限定理 ………………………………… (177)
习题 4 ………………………………………………… (183)
第五章 数理统计的基本概念 …………………… (185)
§5.1 数理统计的基本问题 ………………………… (185)
§5.2 总体与样本 …………………………………… (186)
5.2.1 总体与样本 ……………………………… (186)
5.2.2 经验分布函数与直方图 ………………… (189)
§5.3 统计量与抽样分布 …………………………… (193)
5.3.1 统计量 …………………………………… (193)
5.3.2 抽样分布 ………………………………… (196)
5.3.3 抽样分布定理 …………………………… (201)
习题 5 ………………………………………………… (206)
第六章 参数估计 ……………………………… (209)
§6.1 点估计 ………………………………………… (209)
6.1.1 矩估计法 ………………………………… (209)
6.1.2 最大似然估计法 ………………………… (212)
6.1.3 估计量的评选标准 ……………………… (219)
§6.2 区间估计 ……………………………………… (223)
6.2.1 置信区间 ………………………………… (223)
6.2.2 正态总体均值的区间估计 ……………… (226)
6.2.3 正态总体方差的区间估计 ……………… (228)

6.2.4 两正态总体均值差的区间估计 …………………… (229)
6.2.5 两正态总体方差比的区间估计 …………………… (233)
§6.3 非正态总体参数的区间估计 ………………… (234)
6.3.1 总体均值的区间估计 ……………………………… (235)
6.3.2 两总体均值差的区间估计 ………………………… (236)
§6.4 单侧置信区间 ………………………………………… (238)
习题 6 …………………………………………………………… (241)
第七章 假设检验 ………………………………………… (244)
§7.1 假设检验的基本概念 ……………………………… (244)
§7.2 单个正态总体参数的假设检验 ………………… (247)
7.2.1 检验假设 $H_0:\mu=\mu_0$ ………………………………… (247)
7.2.2 检验假设 $H_0:\sigma^2=\sigma_0^2$ ………………………………… (249)
§7.3 两个正态总体参数的假设检验 ………………… (251)
7.3.1 检验假设 $H_0:\mu_1=\mu_2$ ………………………………… (251)
7.3.2 检验假设 $H_0:\sigma_1^2=\sigma_2^2$ ……………………………… (256)
§7.4 单边假设检验 ………………………………………… (259)
§7.5 非正态总体参数的假设检验 …………………… (264)
*§7.6 假设检验中的两类错误分析 …………………… (268)
§7.7 分布假设检验 ………………………………………… (271)
习题 7 …………………………………………………………… (278)
第八章 方差分析 ………………………………………… (281)
§8.1 单因素方差分析 ……………………………………… (281)
8.1.1 单因素方差分析的统计意义 …………………… (281)
8.1.2 单因素方差分析的基本方法 …………………… (283)
8.1.3 单因素方差分析的计算程序和实例 ………………… (286)
§8.2 双因素方差分析 ……………………………………… (289)
8.2.1 双因素无重复试验的方差分析 ………………… (289)
8.2.2 双因素等重复试验的方差分析 ………………… (297)

习题 8 ………………………………………………………… (305)
第九章 一元线性回归分析 ………………………………… (309)
§ 9.1 一元线性回归 ………………………………………… (310)
9.1.1 一元线性回归的原理和方法 ……………………… (310)
9.1.2 回归方程的显著性检验 ………………………… (317)
§ 9.2 一元线性回归方程的应用 ………………………… (323)
9.2.1 预测 ………………………………………………… (323)
9.2.2 控制 ………………………………………………… (327)
§ 9.3 可线性化的非线性回归 …………………………… (331)
9.3.1 非线性回归问题的线性化方法 ………………… (331)
9.3.2 常用曲线的图形 …………………………………… (335)
习题 9 ………………………………………………………… (338)
练习题、习题参考答案 ………………………………………… (342)
参考书目……………………………………………………… (379)
名词索引……………………………………………………… (380)
附表 I 泊松分布表……………………………………………… (385)
附表 II 标准正态分布表………………………………………… (386)
附表 III t 分布上侧分位数表 ……………………………… (388)
附表 IV χ^2 分布上侧分位数表 ………………………… (389)
附表 V F 分布上侧分位数表………………………………… (391)

第一章　概率论的基本概念

概率论(Probability Theory)是研究随机现象的统计规律性的一门数学分支.

§1.1　随机现象与随机事件

1.1.1　随机现象

自然界中的现象,从它们发生的必然性的角度区分,可以分为两类:一类是确定性现象,一类是不确定性现象.确定性现象是指这样一类现象:当一定的条件实现时,它一定发生,而当条件不满足时,它一定不发生.例如,带同种电荷的两个小球相互排斥,而带异种电荷的两个小球必相互吸引;硬币在空中自由降落,必然落向地面;函数 $f(x)$ 在区间$[a,b]$上连续,它在$[a,b]$上的积分一定存在,等等,都是确定性现象的例子.不确定性现象,是指当一定的条件实现时,这种现象可能出现,也可能不出现.例如,当我们抛出的硬币落到地面时,它可能正面(我们规定硬币的某一面为正面,另一面为反面)向上,也可能反面向上;随意射出的一发子弹,可能击中目标,也可能偏离目标;一个口袋,装有红、白两种球,任意摸出一个,假设摸到每个球的可能性相同,那么我们可能摸出红球,也可能摸出白球.在一定的条件下,对不确定性现象,重复同一个试验,出现的结果常常是不同的.但是,不确定性现象表现出的偶

然性又是受事物内部的固有规律性支配的，因此，在大量的重复试验中，某现象的出现，又会呈现出一定的规律性. 例如，当向同一目标发射多发炮弹时，将会看到弹着点的分布具有一定的规律性，它大致关于目标成中心对称，且越靠近目标，弹着点越密；远离目标的区域，弹着点则较稀疏. 弹着点落入目标周围任意一个区域内的频率，是基本稳定的. 再如，大量重复抛硬币这一试验，将会发现正面向上的次数约占 1/2. 一般说来，试验的次数越多，这种规律性就越明显. 在大量重复试验中不确定性现象表现出的这种规律性叫做**统计规律性**(statistical regularity). 这种在个别试验中出现的结果具有偶然性，而在大量重复试验中表现出统计规律性的现象称为**随机现象**(random phenomenon).

1.1.2 随机试验和样本空间

研究随机现象时，我们称条件实现一次为一次**试验**(trial). 概率论中的"试验"是一个含义广泛的术语，它既包含科学研究和工程技术中的各种实验，也包括对一些事物的某一特征的观测. 一般说来，我们希望试验具有如下一些特征：

(1) 试验可以在相同的条件下重复进行；

(2) 试验可能出现的所有结果是事先预知的；

(3) 每次试验有且只有其中的一个结果出现，但在每次试验结束之前，不知道哪一个结果会出现.

对试验的这些要求是合理的. 我们称具有上述三个特征的试验为**随机试验**(random trial). 以后我们常用字母 E 表示随机试验. 本书中以后提到的试验都是指随机试验.

对随机试验，我们首先关心的是它可能出现的结果有哪些. 随机试验的每一个可能出现的结果叫做一个**样本点**(sample point). 用字母 ω 表示，而把试验 E 的所有可能结果的集合叫做 E 的**样本空间**(sample space)，并用字母 Ω 表示. 样本点的全体构成样本空

间．例如我们规定硬币的一面为正面，另一面为反面，则抛一枚硬币，有两个可能的结果：正面向上或反面向上．正面向上，记为 ω，反面向上，记为 $\bar{\omega}$，则 ω 和 $\bar{\omega}$ 就是这个试验的两个样本点．样本空间就是

$$\Omega = \{\omega, \bar{\omega}\}.$$

根据随机试验的特征(3)，样本点 ω 应是随机试验的最简单的、不可以再分的结果．当随机试验确定之后，样本空间 Ω 就是已知的．

下面我们考察几个简单的例子．

［**例 1.1**］ 设 E_1 表示抛一枚硬币 2 次，观察正、反面出现的情况，则 E_1 的样本空间为

$$\Omega = \{\omega\omega, \omega\bar{\omega}, \bar{\omega}\omega, \bar{\omega}\bar{\omega}\}.$$

其中表示每一个样本点的两个字母中，第一个表示第一次抛出时得到的结果，第二个表示第二次抛出得到的结果．例如 $\omega\omega$ 表示两次都出现正面，$\omega\bar{\omega}$ 表示第一次出现正面而第二次出现反面，余类推．

［**例 1.2**］ 仍考虑抛一枚硬币 2 次，观察正面出现的次数，这个试验记作 E_2，E_2 可能出现的结果有 0 次，1 次，2 次．每种结果是一个样本点，故样本空间为 $\Omega = \{0,1,2\}$．

［**例 1.3**］ 设 E_3：记录在时间$[0,t]$内进入某一商店的顾客人数，其样本点为非负整数，且无一固定的上界，故样本空间为

$$\Omega = \{0,1,2,\cdots\}.$$

［**例 1.4**］ 设 E_4：检验某个灯泡的寿命，以 t（小时）记其寿命，则 E_4 的样本空间为

$$\Omega = \{t \mid 0 \leqslant t < +\infty\}.$$

［**例 1.5**］ 设 E_5：记录本地区一昼夜间的最低气温 x 和最高气温 y，分别以 T_0 和 T_1 表示本地区的最低和最高气温限，则 E_5 的样本空间为

$$\Omega = \{(x,y) \mid T_0 \leqslant x < y \leqslant T_1\}.$$

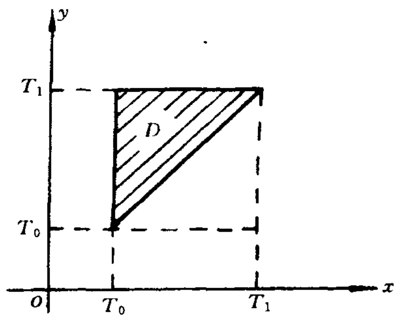


图 1—1

可以看到，样本空间 Ω 中包含样本点的数目，可以是有限个（如 E_1,E_2），也可以是无限多个（如 E_3,E_4,E_5）. 在 E_3 中，样本点的个数和自然数可以一一对应，故样本点是“可数”的，这时我们称样本点的个数是可数无限多个，或可列无限多个. 在 E_4 和 E_5 中，样本点的个数分别可以和半直线上的点一一对应或和三角形区域 $D: T_0 \leqslant x < T_1, x < y \leqslant T_1$（图 1—1）中的点一一对应，这时称样本点的个数为不可数（不可列）无限多个.

练习 1　在例 1.4 中，假设我们关心的是灯泡是否为合格品，并规定灯泡的寿命在 1 000 小时以上为合格品（记为“1”），否则为不合格品（记为“0”），写出这个试验的样本空间. 这个试验和 E_4 是否相同？

练习 2　从包含两件正品 a_1,a_2 和一件次品 b 的 3 件产品中，依次取出 2 件，写出这个试验的样本空间. 如果将题目中的“依次”改为“同时”，是否仍为同一个试验？写出后一个试验的样本空间.

练习 3　将 r 个球随机地放入 n 个盒子中，每一种放法是一个样本点，这个试验共有多少个样本点？考虑将两个球 a,b 放入 3 个盒子中的情形，例如，用

a	b	

表示球 a 放入第一个盒子、球 b 放入第二个盒子的情形. 写出这个试验的所有样本点.

1.1.3 随机事件

对随机试验,引入样本空间后,人们常常关心的是试验的结果是否包含在样本空间中满足一定条件的某一个子集中. 例如,若规定灯泡的寿命在 1 000 小时以上为合格品,则对一个灯泡,我们关心的是它的寿命 t 是否在 1 000 小时以上. 这里集合 $A=\{t \mid t \geqslant 1\,000\}$ 就构成 $\Omega=\{t \mid 0 \leqslant t<+\infty\}$ 的一个子集. 我们称样本空间的每一个子集为一个**随机事件**(random event),简称为事件. 事件常以大写英文字母 $A,B,\cdots$ 表示.

设某次试验的结果 $\omega \in A$,则称在这次试验中事件 A 发生,否则称在这次试验中 A 没有发生. 根据定义,空集 Φ 和整个样本空间 Ω 也是随机事件. Φ 在每次试验中都不出现,称为**不可能事件**,而 Ω 在每次试验中一定出现,称为**必然事件**. 每个样本点 ω 可作为 Ω 的一个单子集,因而也是随机事件,称为随机试验的**基本事件**(elementary event). 样本空间 Ω 也称为**基本事件空间**.

下面进一步讨论前面给出的例.

[**例 1.1(续)**] 对试验 E_1,$A=\{\omega\omega,\omega\overline{\omega}\}$ 表示事件:第一次出现正面;$B=\{\overline{\omega}\omega,\omega\omega\}$ 表示事件:第二次出现正面;$C=\{\omega\omega,\omega\overline{\omega},\overline{\omega}\omega\}$ 表示事件:至少出现一次正面;$D=\{\overline{\omega}\overline{\omega}\}$ 表示事件:两次都出现反面. 其中 D 是 E_1 的一个基本事件.

[**例 1.2(续)**] 对试验 E_2,设 A 表示事件:正面出现不少于 1 次,则 $A=\{1,2\}$;设 B 表示事件:出现正面不多于 2 次,则 $B=\{0,1,2\}$;记 C 表示事件:出现 3 次正面,则 $C=\Phi$. 其中 $B=\Omega$ 是必然事件,C 是不可能事件.

[**例 1.3(续)**] 设 A 表示事件:进入商店的人数不超过 3 个.

则 $A=\{0,1,2,3\}$；设 B 表示事件：进入商店的人数不少于 5 个，则 $B=\{5,6,7,8,\cdots\}$；设 C 表示事件：无人进入商店，则 $C=\{0\}$. 这是 E_3 的一个基本事件.

［**例 1.4(续)**］ 设事件 A：灯泡寿命小于 200 小时，则 $A=\{t\mid 0\leqslant t<200\}$；设事件 B：灯泡寿命在 200 到 1 000 小时之间，则 $B=\{t\mid 200\leqslant t\leqslant 1\ 000\}$；$C=\{t=1\ 035\}$ 表示灯泡寿命恰为 1 035 小时，这是 E_4 的一个基本事件.

［**例 1.5(续)**］ 设事件 A：最低气温不低于 10℃，则 $A=\{(x,y)\mid T_0\leqslant 10℃\leqslant x<y\leqslant T_1\}$；设事件 B：最高气温与最低气温之差不超过 10℃，则 $B=\{(x,y)\mid y-x\leqslant 10℃, T_0\leqslant x<y\leqslant T_1\}$. A,B 可由图 1—2 表示.

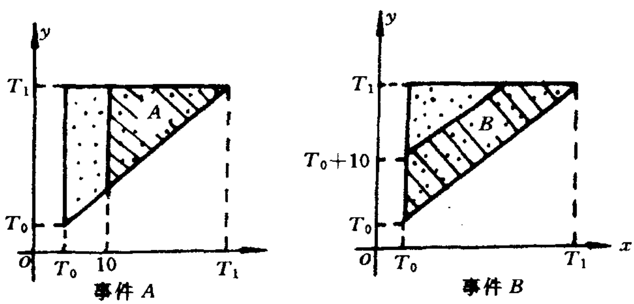


图 1—2

样本点、样本空间与随机事件之间的关系等同于集合论中元素、集合与子集之间的关系，因此我们可以仿照集合论中子集的关系与运算来定义随机事件之间的关系和运算.

在下面的叙述中，设 Ω 是随机试验 E 的样本空间，ω 表示其样本点，$A,B,C,A_k,B_k,\cdots$ 均表示 E 的事件. 为了直观起见，我们用平面上的一个矩形区域表示样本空间 Ω，ω 即为该矩形域中的一点，用矩形域中的两个圆分别表示事件 A 和事件 B.

(1) 事件的包含关系