概率统计

（第四版）

同济大学概率统计教研组　编著

同济大学出版社
TONGJI UNIVERSITY PRESS

概率统计

（第四版）

同　济　大　学
概率统计教研组　编著

同濟大學出版社
TONGJI UNIVERSITY PRESS

内 容 提 要

本书在原《概率统计(工程数学)》第三版的基础上，根据作者多年的教学改革实践修订而成，内容包括随机事件与概率、离散型随机变量及其分布、连续型随机变量及其分布、随机变量的数字特征、随机变量序列的极限、现代概率论基础简介、数理统计的基本概念、参数估计、假设检验、回归分析与方差分析. 书中各章附有相当数量的习题，书末附有习题的参考答案，供读者查阅. 本书在教育部制定的教学大纲的基础上，紧扣硕士研究生入学考试大纲，并以此规范概率统计中的术语与记号.

本书以提高读者解题能力与解决实际问题能力为基本出发点，从实例引入抽象的基本概念，从抽象的数学定理又回到具体的应用问题，有助于读者较快地掌握近代的概率统计知识. 本书可作为高等院校本科生(包括理工类与经济类)概率论与数理统计课程的教材或参考书，也可作为广大概率统计应用人员的工具性参考书.

图书在版编目(CIP)数据

概率统计/同济大学概率统计教研组编著. 何迎晖执笔. 4版. —上海：同济大学出版社，2009.6
ISBN 978-7-5608-4020-8

Ⅰ. 概… Ⅱ. ①同…②何… 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(2009)第 064303 号

概率统计(第四版)
同济大学概率统计教研组 编著
责任编辑 曹 建 责任校对 徐春莲 封面设计 潘向蓁

出版发行 同济大学出版社 www.tongjipress.com.cn
(地址：上海市四平路 1239 号 邮编：200092 电话：021－65985622)
经 销 全国各地新华书店
印 刷 同济大学印刷厂
开 本 787mm×960mm 1/16
印 张 18
印 数 1—5100
字 数 360000
版 次 2009 年 6 月第 4 版 2009 年 6 月第 1 次印刷
书 号 ISBN 978-7-5608-4020-8

定 价 25.00 元

前　言

随着我国高等院校规模的扩大,学生数目的膨胀,在大中城市里,高等教育正走在普及化的道路上,需要在本科阶段学习概率统计知识的学生越来越多.然而,教学大纲与相应的考研要求并没有因此而有所降低.

在新的形势下,具体教学实践中产生了一个新问题.一方面,相当一部分学生出于今后攻读硕士、博士研究生的考虑,不满足简单基础知识的学习,他们中间的优秀者自然合理地希望通过本科阶段的努力学习能顺利地达到硕士研究生入学考试的要求.另一方面,也有为数不少的学生迟迟不能从传统数学方法的束缚中解脱出来以适应随机数学的思维方式,他们总觉得概率统计这门随机数学课程很难理解,因而只满足于达到教学大纲的最低要求.从理论上说,解决这个问题最简单的答案是“因材施教”.然而,在同一课堂上,这恰是“知易行难”的.

在第四版的修订工作中,我们试图从教材的角度对解决这个令人困惑的问题作一些探索.这个思想实际上已经渗透在第二版和第三版的编写之中.单纯满足一部分学生要求的教材是不可取的,我们力求在二者之间寻找一个适当的平衡点.当然,效果如何还有待于教学实践来检验.为此,真诚地希望使用本书的老师与学生提出宝贵的意见.

第四版的修订工作由何迎晖执笔.

编　者

2009 年清明

第三版前言

跨入新世纪以来，随着高等院校招生规模的扩大，我国高等教育正经历着从精英教育向大众教育的转变过程．社会各界对提高教育质量的要求日益高涨．教材质量直接影响到高等学校教学水平的提高．为了使近代的概率统计知识成为广大本科生（包括理工类与经济类）和概率统计应用人员手中的一个重要数学工具，本书再次作了修订．

在保持原书风格与体系的基础上，第三版遵循以下思路进行修订：（1）突出概率论是本科段教学的重点；（2）明确随机变量是概率论的研究对象；（3）强调分布是随机变量的核心；（4）深化随机变量相互独立性在多个随机变量中的作用；（5）强化随机变量数字特征的应用；（6）增加前、后知识点的联系．愿读者通过阅读本书能学以致用，不仅提高解题能力，而且提高解决实际问题的能力．

书中不妥之处，敬请读者指正．

第三版的修订工作由何迎晖执笔．

编　者

2004年元旦

第二版前言

1997年原国家教委把概率统计课程纳入全国硕士研究生入学统一考试内容，此举极大地促进了概率统计的教学工作，同时也推动了本书的修订工作.

第二版在保持原书特色的基础上力图贯彻下列思想：(1) 突出概率论中随机变量分布这一核心内容；(2) 在不脱离现有本科生实际数学水平的基础上，以离散型随机变量为对象，严格叙述并论证概率论中的基本理论；(3) 在内容安排上，适当注意概率统计在经济类与管理类专业中的应用；(4) 与硕士研究生入学统一考试大纲相衔接，并以此规范概率统计中的术语与记号；(5) 与硕士研究生的后续课程——“随机过程”与“应用统计”相衔接；(6) 为了使读者今后深入学习概率统计知识，增加第六章介绍现代概率论基础的一些常识，供课外选读.

根据我们教学实践估计，如果使用本书作教材，只进行概率论部分(前五章)教学约需32学时；按大部分专业的考研要求，进行概率统计(前五章及第七、八、九章)教学约需60学时.

这些年来，承蒙使用本书师生的厚爱，收到了各种修改意见. 我们在此谨表谢意，同时也欢迎读者今后继续赐教.

第二版完成之际正值新旧世纪之交. 若本书在客观上成为21世纪的教材，则纯属巧合，而非编者主观努力的成果.

第二版的修订工作由何迎晖执笔.

编　者

2000年元旦

第一版前言

概率论与数理统计是随机数学的两个分支，它们在各个领域中都有极其广泛的应用．目前，概率论与数理统计已经成为绝大多数工科专业大学生的一门重要的基础理论课程．

为了便于读者掌握概率论与数理统计的基本概念与基本理论，初步掌握处理随机现象的基本思想与方法，提高运用概率统计方法分析和解决实际问题的能力，本书在内容处理上作了如下安排：(1) 以实际例子引进概率统计的基本概念；(2) 强调概率统计在工程技术上的应用；(3) 以离散型随机变量为对象讲清概率论中的基本概念与基本性质，以连续型随机变量为背景强调概率论的计算与应用；(4) 对随机变量的分布强调值域这个容易被忽视的概念；(5)按国家标准采用规范的概率统计术语．

本书是我们教研组全体任课老师历年教学工作的一个总结．油印讲义在教学中曾使用了多次．现根据 1992 年国家教委下达的“高等工业学校概率论与数理统计课程教学基本要求”作了进一步的修订．本书可作为高等工业学校概率论与数理统计课程的教材或教学参考书．

本书由王金宝(第一、二章)，钱志坚(第三、四章)，钱伟民(第五、六章)与庄勇荣(第七、八章)编写初稿，由何迎晖统稿并写第二稿，最后由潘承毅修改、润色、定稿．在编写过程中，闵华玲、蒋凤瑛等老师提供了许多宝贵的意见与帮助，在此向他们致谢．

限于编者的水平，本书的不足之处，衷心希望读者批评指正．

编　者

1992 年 10 月

目　　录

第十章　回归分析与方差分析

附　表

第一章 随机事件与概率

概率论是研究随机现象的统计规律性的一个数学分支. 本章介绍概率论中的基本概念——随机事件与随机事件的概率，并进一步讨论随机事件的关系与运算，以及概率的性质与初等计算方法.

§1.1 随机事件

在自然界与人类的社会活动中常常会出现各种各样的现象. 例如，一枚硬币向上抛起后必然会落地；在相同的大气压与温度下，气罐内的分子对罐壁的压力是个常数. 这类现象的共同特点是，在确定的试验条件下，它们必然会发生，称这类现象为**确定性现象**. 另一类现象则不然. 例如，将一枚硬币上抛，着地时究竟是正面向上还是反面向上，这在上抛前是无法断言的. 但是，人们从长期实践中知道，多次重复上抛同一枚硬币，出现正面向上的可能性占一半左右，这类在个别试验中呈现不确定的结果而在大量重复试验中结果呈现某种规律性的现象称为**随机现象**，这种规律性称为**统计规律性**. 为研究随机现象的统计规律性作准备，本节介绍随机试验、样本空间与随机事件等概念.

一、随机试验

在客观世界中，随机现象是极为普遍的. 例如，某地的年降雨量，河流某处的年最高水位，相同条件下生产的电子元件的寿命，某交通道口中午 1h 内汽车流量，等等. 为了对随机现象的统计规律性进行研究，有时要做一些试验. 这里所说的试验，必须具有以下 3 个特点：

(ⅰ) 试验可以在相同的条件下重复地进行；

(ⅱ) 试验的所有可能结果在试验前已经明确，并且不止 1 个；

(ⅲ) 试验前不能确定试验后会出现哪一个结果.

在概率论中，称具有上述 3 个特点的试验为**随机试验**，简称为**试验**.

下面给出一些随机试验的例子：

例 1.1 上抛 1 枚硬币并观察硬币着地时向上的面，这是一个试验.

例 1.2 观察某交通道口中午 1h 内汽车流量(单位：辆)，这是一个试验. 可能出现的试验结果可以是非负整数中的任意一个，但试验前无法确定究竟会出现哪一个非负整数.

例 1.3 从某厂生产的相同型号的灯泡中抽取 1 只，测试它的寿命(即正常工作

的小时数). 这是一个试验. 可能出现的试验结果可以是非负实数中的任意一个,但试验前无法确定究竟会出现哪一个非负实数.

在实际生活中还存在许多随机试验的例子. 例如,彩票的开奖,质检部门对产品的质量检查,等等.

二、样本空间

要研究一个随机试验,首先要弄清楚这个试验所有可能的结果. 每一个可能出现的结果称为**样本点**,记作 ω(必要时带有下标或上标). 全体样本点组成的集合称为**样本空间**,记作 Ω. 换句话说,样本空间是试验的所有可能结果所组成的集合. 这个集合中的元素就是样本点.

例 1.1(续) 在例 1.1 中,样本空间 $\Omega=\{$正面,反面$\}$,它由两个样本点组成.

例 1.2(续) 在例 1.2 中,样本空间 $\Omega=\{0,1,2,\cdots\}$,它是 1 个数集,由可列无限个[1]样本点组成.

例 1.3(续) 在例 1.3 中,样本空间 $\Omega=[0,\infty)$,它是 1 个数集,由不可列无限个样本点组成.

从这 3 个例子中可以看到,样本空间可以是数集,也可以不是数集;样本空间可以是有限集,也可以是无限集.

有时候,为了数学上处理的方便,可以把样本空间作适当的扩大. 例如,在例 1.3 中,灯泡寿命实际上不会超过某个足够大的正数,但我们仍取样本空间为 $[0,\infty)$,必要时甚至还可以取样本空间为 $(-\infty,\infty)$. 在例 1.2 中也作了类似的扩大.

三、随机事件

当我们通过随机试验来研究随机现象时,常常不是关心某一个样本点在试验后是否出现,而是关心满足某些条件的样本点在试验后是否出现. 例如,在例 1.2 中,要通过对该道口汽车流量的观察来决定是否需要扩建道口. 假定超过 600 辆便认为需要扩建. 这时,我们关心的便是试验结果是否大于 600. 满足这一条件的样本点组成了样本空间的 1 个子集. 称 1 个随机试验的样本空间的子集为**随机事件**,简称为**事件**. 随机事件通常用大写字母 $A,B,C,\cdots$ 表示. 仅含 1 个样本点的随机事件称为**基本事件**. 在例 1.1 中,有 2 个基本事件$\{$正面$\}$、$\{$反面$\}$;在例 1.2 与例 1.3 中,分别有无限多个基本事件.

在试验后,如果出现随机事件 A 中所包含的某个样本点,那么,称**事件 A 发生**;否则,就称**事件 A 不发生**. 在例 1.2 中,设 A 表示"流量大于 600",在试验后事件 A 可能发生,也可能不发生. 如果试验结果是 689,那么,便认为事件 A 发生.

样本空间 Ω 是其自身的 1 个子集,因而也是 1 个事件. 由于样本空间 Ω 包含所有

---

① 可列无限个的含义是:这无限个元素可以按某种次序排成一列. 例如,自然数有可列无限个.

的样本点，因此，每次试验后，必定有 Ω 中的 1 个样本点出现，即 Ω 必然发生. 称 Ω 为**必然事件**. 空集 $\varnothing$ 永远是样本空间的 1 个子集，因而也是 1 个事件. 由于空集 $\varnothing$ 不包含任何一个样本点，因此每次试验后 $\varnothing$ 必定不发生. 称 $\varnothing$ 为**不可能事件**. 必然事件 Ω 与不可能事件 $\varnothing$ 是两个特殊的随机事件.

四、随机事件之间的关系与运算

在一个样本空间中，可以有许多随机事件，希望通过对较简单的事件的了解去掌握较复杂的事件. 为此，需要研究事件之间的关系与事件之间的运算.

由于事件是一个集合，因此，事件之间的关系与事件之间的运算应该按照集合论中集合之间的关系与集合之间的运算来规定.

给定一个随机试验，Ω 是它的样本空间，下列事件 A,B,C 与 $A_i(i=1,2,\cdots)$ 都是 Ω 的子集.

(1) 如果 $A \subset B$(或 $B \supset A$)，那么，称事件 B **包含**事件 A. 它的含义是：事件 A 发生必定导致事件 B 发生. 图 1.1 给出了这种包含关系的一个几何表示.

在例 1.3 中，事件 A 表示“灯泡寿命不超过 200h”，事件 B 表示“灯泡寿命不超过 300h”. 于是，$A \subset B$(或 $B \supset A$).

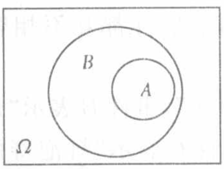

图 1.1　$A \subset B$

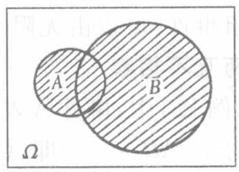

图 1.2　$A \cup B$

(2) 如果 $A \subset B$ 且 $B \subset A$，即 $A = B$，那么称事件 A 与事件 B **相等**.

(3) 事件 $A \cup B = \{\omega : \omega \in A \text{ 或 } \omega \in B\}$ 称为事件 A 与事件 B 的**和事件**(或**并事件**). 它的含义是：当且仅当事件 A 与事件 B 中至少有 1 个发生时，事件 $A \cup B$ 发生. 图 1.2 给出了这种运算的一个几何表示.

例如，在某建筑工地上，事件 A 表示“缺少水泥”，事件 B 表示“缺少黄砂”. 于是，和事件 $A \cup B$ 表示“缺少水泥或黄砂”.

一般地，用 $\bigcup\limits_{i=1}^{n} A_i$ 表示 n 个事件 $A_1,\cdots,A_n$ 的和事件；用 $\bigcup\limits_{i=1}^{\infty} A_i$ 表示可列无限个事件 $A_1,A_2,\cdots$ 的和事件.

(4) 事件 $A \cap B = \{\omega : \omega \in A \text{ 且 } \omega \in B\}$ 称为事件 A 与事件 B 的**积事件**(或**交事件**). 它的含义是：当且仅当事件 A 与事件 B 同时发生时，事件 $A \cap B$ 发生. 积事件也可以记作 AB. 图 1.3 给出了这种运算的一个几何表示.

例如，某输油管长 10km. 事件 A 表示“前 5km 油管正常工作”，事件 B 表示“后 5km 油管正常工作”. 于是，积事件 $A \cap B$ 表示“整个输油管正常工作”.

一般地，用 $\bigcap\limits_{i=1}^{n} A_i$ 表示 n 个事件 $A_1,\cdots,A_n$ 的积事件；用 $\bigcap\limits_{i=1}^{\infty} A_i$ 表示可列无限个事件 $A_1,A_2,\cdots$ 的积事件.

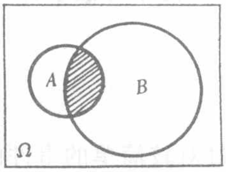

图 1.3　$A \cap B$

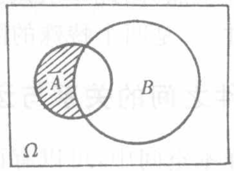

图 1.4　$A - B$

(5) 事件 $A - B = \{\omega : \omega \in A \text{ 且 } \omega \notin B\}$ 称为事件 A 与事件 B 的**差事件**. 它的含义是：当且仅当事件 A 发生且事件 B 不发生时，事件 $A - B$ 发生. 图 1.4 给出了这种运算的一个几何表示.

例如，某种圆柱形零件的长度与外径都合格时才算合格. 事件 A 表示“长度合格”，事件 B 表示“外径合格”. 于是，差事件 $A - B$ 表示“长度合格但外径不合格”.

(6) 如果 $A \cap B = \varnothing$，那么称事件 A 与事件 B **互不相容**（或**互斥**）. 它的含义是：事件 A 与事件 B 在 1 次试验后不会同时发生. 图 1.5 给出了这种运算的一个几何表示.

如果一组事件（可以由无限个事件组成）中任意两个事件都互不相容，那么，称这组事件**两两互不相容**.

例如，在例 1.3 中，事件 A 表示“灯泡寿命不超过 200h”，事件 B 表示“灯泡寿命至少为 300h”. 于是 $AB = \varnothing$，即 A 与 B 互不相容. 如果事件 C 表示“灯泡寿命在(200h，300h) 内”，那么，A,B,C 构成一个两两互不相容的事件组. 又如，任意 2 个基本事件总是互不相容；任意一组基本事件总是两两互不相容.

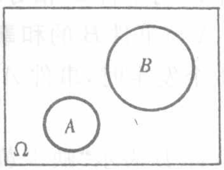

图 1.5　$A \cap B = \varnothing$

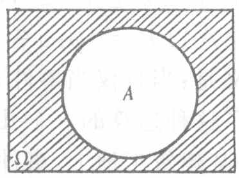

图 1.6　$\overline{A}$

(7) 事件 $\Omega - A$ 称为事件 A 的**对立事件**（或**逆事件**，或**余事件**），记作 $\overline{A} = \Omega - A$. 它的含义是：当且仅当事件 A 不发生时，事件 $\overline{A}$ 发生. 于是 $A \cap \overline{A} = \varnothing$，$A \cup \overline{A} = \Omega$. 由于 A 也是 $\overline{A}$ 的对立事件，因此称事件 A 与 $\overline{A}$ **互逆**（或**互余**）. 图 1.6 给出了这种运算的一个几何表示.

例如，某建筑物在经历一场地震后，事件 A 表示“建筑物倒塌”. 于是，事件 $\overline{A}$ 表示“建筑物幸存”.

按差事件与对立事件的定义，差事件也可以表示成 $A - B = A\overline{B}$.

与集合论中集合的运算一样，事件之间的运算满足下述定律：

(ⅰ) **交换律** $A \cup B = B \cup A, A \cap B = B \cap A$；

(ⅱ) **结合律** $A \cup (B \cup C) = (A \cup B) \cup C$，

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

(ⅲ) **分配律** $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$，

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

(ⅳ) **德·摩根**(De Morgan) **法则** $\overline{A \cup B} = \overline{A} \cap \overline{B}, \overline{A \cap B} = \overline{A} \cup \overline{B}$.

这些定律都可以推广到任意多个事件.

例 1.4 某城市的供水系统由甲、乙两个水源与三部分管道 1,2,3 组成(图 1.7). 每个水源都足以供应城市的用水. 设事件 A_i 表示"第 i 号管道正常工作"，$i = 1, 2, 3$. 于是，"城市能正常供水"可表示为

$$(A_1 \cup A_2) \cap A_3；$$

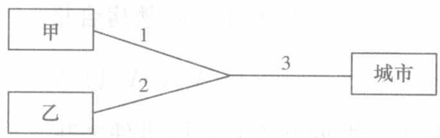

图 1.7 供水系统示意图

由德·摩根法则可知，"城市断水"可表示为

$$\overline{(A_1 \cup A_2) \cap A_3} = \overline{(A_1 \cup A_2)} \cup \overline{A_3} = (\overline{A_1} \cap \overline{A_2}) \cup \overline{A_3}.$$

例 1.5 某工程队承包建造了 3 幢楼房，设事件 A_i 表示"第 i 幢楼房经验收合格"，$i = 1, 2, 3$. 试用 A_1, A_2, A_3 表示下列事件：

(1) 第 1 幢楼房合格；

(2) 只有第 1 幢楼房合格；

(3) 恰有 1 幢楼房合格；

(4) 至少有 1 幢楼房合格；

(5) 至多有 1 幢楼房合格.

解 事件 $\overline{A_i}$ 表示"第 i 幢楼房经验收不合格"，$i = 1, 2, 3$.

(1) A_1. 这时，第 2、第 3 幢楼房可能合格，也可能不合格.

(2)"只有第 1 幢楼房合格"包含了"第 2、第 3 幢楼房不合格"的意思，因此，这个事件可以表示成 $A_1\overline{A_2}\,\overline{A_3}$.

(3)"恰有 1 幢楼房合格"没有指明究竟哪一幢楼房合格，因此，这个事件可以表示成

$$A_1\overline{A_2}\,\overline{A_3} \cup \overline{A_1}A_2\overline{A_3} \cup \overline{A_1}\,\overline{A_2}A_3.$$

$A_1\overline{A_2}\,\overline{A_3}, \overline{A_1}A_2\overline{A_3}, \overline{A_1}\,\overline{A_2}A_3$ 这 3 个事件构成 1 个两两互不相容的事件组.

(4)"至少有 1 幢楼房合格"可以看成 A_1, A_2, A_3 这 3 个事件中至少有 1 个发生，因此这个事件可以表示成 $A_1 \cup A_2 \cup A_3$. A_1, A_2, A_3 这 3 个事件不构成 1 个两两互不相容的事件组. 另一方面，"至少有 1 幢楼房合格"的对立事件是"3 幢楼房全不合格"，因此，所求事件也可以表示成 $\overline{\overline{A_1}\,\overline{A_2}\,\overline{A_3}}$. 由德·摩根法则知道

$$\overline{\overline{A}_1\overline{A}_2\overline{A}_3} = \overline{\overline{A}}_1 \cup \overline{\overline{A}}_2 \cup \overline{\overline{A}}_3 = A_1 \cup A_2 \cup A_3.$$

(5)“至多有 1 幢楼房合格”是下列 2 个互不相容的事件的和事件:“恰有 1 幢楼房合格”与“3 幢楼房全不合格”,因此,所求事件可以表示成

$$A_1\overline{A}_2\overline{A}_3 \cup \overline{A}_1A_2\overline{A}_3 \cup \overline{A}_1\overline{A}_2A_3 \cup \overline{A}_1\overline{A}_2\overline{A}_3.$$

由例 1.5(4) 看出,事件的表达一般不唯一.“至少有 1 幢楼房合格”还可以表示成

$$\bigcup_{i=1}^{3}\{\text{恰有 } i \text{ 幢楼房合格}\} = A_1\overline{A}_2\overline{A}_3 \cup \overline{A}_1A_2\overline{A}_3 \cup \overline{A}_1\overline{A}_2A_3$$
$$\cup\, \overline{A}_1A_2A_3 \cup A_1\overline{A}_2A_3 \cup A_1A_2\overline{A}_3 \cup A_1A_2A_3.$$

这是七个两两互不相容的事件之并.

§1.2　等可能概型

在一次试验后,随机事件 A 可能发生,也可能不发生. 随机事件发生的可能性的大小用区间$[0,1]$中的一个数来刻画,这个数称为**概率**. 事件 $A,B,C,\cdots$ 的概率分别记作 $P(A),P(B),P(C),\cdots$. 作为事件的 2 个特殊情况:必然事件 Ω 与不可能事件 $\varnothing$,自然应该合理地规定

$$P(\Omega) = 1,\quad P(\varnothing) = 0.$$

如何计算概率?这是本章以下内容讨论的主题. 本节讨论最简单的情形 —— **等可能概型**,即样本空间中的每个样本点在一次试验后以相等的可能性出现.

一、古典型概率

上抛 1 枚硬币,观察硬币着地时向上的面. 假定这枚硬币质地均匀,因此“出现正面”与“出现反面”的可能性是相等的. 从常识上知道,这两个事件的概率都应该是$\dfrac{1}{2}$.

1 个口袋中装有 5 只外形相同的球,分别编有号码 $1,\cdots,5$. 现在从这个口袋中任取 1 只,取到偶数号码的球的概率有多大?由于样本空间 $\Omega = \{\omega_1,\cdots,\omega_5\}$,这 5 个样本点在 1 次试验后出现的可能性都相等.“取到偶数号码的球”这一事件 $A = \{\omega_2,\omega_4\}$. 从常识上知道,$P(A) = \dfrac{2}{5}$.

一般地,称具有下列 2 个特征的随机试验的数学模型为**古典概型**:

(i)试验的样本空间 Ω 是个有限集,不妨记作 $\Omega = \{\omega_1,\cdots,\omega_n\}$;

(ii)每个样本点在 1 次试验后以相等的可能性出现,即