



普通高等教育“十一五”国家级规划教材

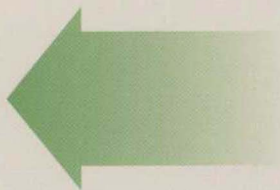
高等学校交通运输专业规划教材

运输统计基础

(第二版)

YUNSHU TONGJI JICHU

王慈光·主编



西南交通大学出版社

[Http://press.swjtu.edu.cn](http://press.swjtu.edu.cn)

运输统计基础

(第二版)

王慈光 主编

西南交通大学出版社

· 成 都 ·

内 容 简 介

统计学是一门应用十分广泛的独立的学科,统计工作是各行各业都不可缺少的重要工作。运输业属于第三产业,无论是日常的运输生产,还是长远的运输规划,都离不开统计。本书从“大统计”的观念出发,系统地介绍概率论、数理统计和一般统计学的基本理论和常用的统计分析方法,并结合各种运输方式的特点,分别介绍铁路、道路、水路、航空、管道运输统计的业务知识和指标计算方法。

本书内容丰富,自成体系,深入浅出,重在应用,可作为高等学校交通运输及相关专业的教材,也可供交通运输部门从事实际工作的生产人员、技术人员和管理人员参考。

图书在版编目(CIP)数据

运输统计基础 / 王慈光主编. —2版. —成都:
西南交通大学出版社, 2010.7
普通高等教育“十一五”国家级规划教材
ISBN 978-7-5643-0701-1

I. ①运… II. ①王… III. ①运输统计—高等学校—
教材 IV. F502

中国版本图书馆CIP数据核字(2010)第113899号

普通高等教育“十一五”国家级规划教材
高等学校交通运输专业规划教材

运输统计基础

(第二版)

王慈光 主编

*

责任编辑 刘婷婷

封面设计 本格设计

西南交通大学出版社出版发行

(成都二环路北一段111号 邮政编码: 610031 发行部电话: 028-87600564)

<http://press.swjtu.edu.cn>

成都蓉军广告印务有限责任公司印刷

*

成品尺寸: 185 mm×260 mm 印张: 21.75

字数: 542千字 印数: 3 001—6 000册

2004年8月第1版 2010年7月第2版 2010年7月第3次印刷

ISBN 978-7-5643-0701-1

定价: 38.00元

图书如有印装质量问题 本社负责退换
版权所有 盗版必究 举报电话: 028-87600562

第二版前言

统计工作是国民经济各行各业都不可缺少的重要工作。统计的基本任务是：对国民经济和社会发展情况进行统计调查、统计分析，提供统计资料和统计咨询意见，实行统计监督。

统计学是用以指导统计工作的一系列原理和方法，是一门综合性很强的方法论科学。从本质上说，统计学是关于数据的科学，它的研究对象是各种现象总体的数量方面，其目的是揭示客观事物的数量特征及其统计规律性，为科学管理和正确决策提供依据。随着科学技术的发展和人类文明的进步，统计学已发展成为由多学科分支组成的学科体系，统计方法也已广泛应用于自然科学和社会科学的众多领域以及国民经济各部门之中。

统计学原理在各领域和各部门的应用逐渐形成一些相对独立的统计学分支，如经济统计学、管理统计学、社会统计学、气象统计学、生物统计学、医疗卫生统计学、商业统计学等。一般统计学原理在交通运输业的具体应用便形成了交通运输统计学，它所研究的是运输业范围内大量社会经济现象的数量特征和数量关系。运输业属于第三产业，是一个特殊的物质生产部门，它生产的产品是旅客和货物（含行包）的位移，这种产品不具有实物形态，不能储存和调拨，产品的生产和消费同时进行；运输业的生产空间广阔，生产环节众多，生产工具经常处于流动状态。以上这些特点决定了交通运输统计学不同于其他部门的统计学，从统计指标的设计、计量单位的确定，到统计资料的收集、整理和分析，都带有运输业的特色。运输业发展到现代，形成了铁路运输、道路运输、水路运输、航空运输和管道运输共五种主要的运输方式。它们各有优势和特长，互相协调和补充，共同构成国家统一的综合运输体系。这五种运输方式具有运输业的共同特性，但又各有其不同的技术经济特征和适用范围，运输生产过程和作业组织方法也不尽相同，因而在统计上也有一定差别。本书以交通运输部门最新统计规章为依据，全面介绍这五种运输方式的统计理论和统计方法。

本书是在第一版的基础上，根据新的教学计划修订而成的。第二版对第一版中社会经济统计学原理、铁路和道路运输统计部分作了全面改写，新增了水路、航空、管道运输统计的内容。为有利于树立“大统计学”的思想观念，第二版仍保留了第一版中概率论与数理统计的基本内容。作为教材，本书定位于高等学校本科生层次，在编写过程中，

注意理论联系实际，突出重点，深入浅出，着重介绍统计基本原理和基本的计算、分析方法，而不拘泥于统计报表的填记和统计规章的条文解释。为便于读者巩固知识和加深理解，各篇均附有适量的习题，书末附有习题答案。

本书共分五篇 14 章，编写分工为：第一篇和第三篇由王慈光执笔，第二篇由左大杰执笔，第四篇由霍娅敏执笔，第五篇由李宗平执笔，全书由王慈光总纂和定稿。

本书可作为高等学校交通运输及相关专业的教材，也可供交通运输部门从事实际工作的生产人员、技术人员和管理人员参考。由于编者水平所限，书中可能存在疏漏和不当之处，欢迎读者批评指正。

编 者

2010 年 8 月 18 日

第一版前言

统计学是一门独立的学科，又是一门综合性很强的学科，它与数学的关系极为密切，又广泛应用于国民经济各部门之中。作为第三产业的交通运输业，无论是日常的运输生产，还是长远的运输规划，都离不开统计。统计学曾经并正在发挥着十分重要的作用。

随着统计学科的发展，当代一些统计学界学者认为建立“大统计学”的思想基本形成，其内容包括概率论、数理统计理论、统计学原理、统计史学、统计法学、管理统计学、社会统计学、国民经济统计学以及应用于各行各业的部门经济统计学等。事实上，数理统计与社会统计学在许多方面是相通的。因此，本书将概率论、数理统计、社会经济统计学的基本原理和陆上运输（含铁路、公路）统计的业务知识汇集在一起，有利于读者建立大统计的概念：既了解统计的有关理论知识，又知晓交通运输业的具体统计内容及计算方法。

本书作为高校专业基础课教材，定位于本科生层次。概括地说，本书具有以下特点：

1. 打破传统分散的教学方式，将概率论、数理统计、社会经济统计学、交通运输统计的内容集中在一本教材中，有利于学生树立“大统计”的观念，系统掌握统计学知识，以便融会贯通，灵活运用。

2. 本教材读者主要是工科专业本科生，既不同于数学专业，又有别于研究生和专科生，在材料的选取、内容的阐述上均考虑到这一点。对概率论、统计学的理论部分，侧重于讲清基本概念、基本思想、基本的计算方法和分析方法，而不纠缠于公式的推导和定理的证明；对运输统计部分，侧重于讲清统计指标的内涵、计算及它们之间的相互关系，而不拘泥于统计报表的填记和统计规章的条文解释。总之，“讲清原理，重在应用”是本书编写过程中遵循的一个重要原则。

3. 统计理论与交通运输专业的有关内容紧密结合，举例和习题多与交通运输领域的实际问题联系起来，使学生了解统计理论在运输专业中的实际应用，激发专业兴趣，为以后的专业课学习打下良好的基础，同时对于交通运输业的实际工作者来说也是极具参考价值的。

4. 每章后附有一定数量的习题，供学生课外（或课堂）练习，以巩固所学知识。

5. 逻辑严谨，叙述简练，由浅入深，清楚易懂，不少内容经过精心编排，组合归纳，更加易于理解和掌握。

本书共分五章。第一章至第四章由王慈光编写，第五章由霍娅敏编写，全书由王慈光统稿。刁明碧教授和熊天文教授仔细审阅了全书，并提出了宝贵的意见，在此向他们表示衷心的感谢。由于编者水平所限，书中不当之处在所难免，欢迎广大读者批评指正。

本书由西南交通大学出版基金资助。

编 者

2004年5月

目 录

第一篇 概率论与数理统计

第一章 概率论基础	1
第一节 随机事件及其概率	1
第二节 随机变量及其分布	11
第三节 随机变量的数字特征	22
第四节 大数定律和中心极限定理	26
第二章 数理统计初步	29
第一节 参数估计	29
第二节 假设检验	37
习 题	43

第二篇 统计学的原理

第三章 概 述	50
第一节 统计学的研究对象	50
第二节 统计学的研究方法	51
第三节 统计学的内容	52
第四章 统计调查与整理	54
第一节 统计调查	54
第二节 统计整理	56
第三节 数据的概括性度量	61
第五章 统计分析	75
第一节 对比分析法	76
第二节 动态分析法	77
第三节 因素分析法	84
第四节 回归分析法	93

第六章 统计预测	101
第一节 移动平均法	101
第二节 速度预测法	102
第三节 指数平滑法	103
第四节 预测误差分析	108
习 题	110

第三篇 铁路运输统计

第七章 铁路客、货运输统计	114
第一节 铁路货运统计调查	114
第二节 铁路货运统计分组	116
第三节 铁路货运统计指标	119
第四节 铁路客运统计调查及分组	128
第五节 铁路客运统计指标	129
第六节 行包运输统计	134
第七节 铁路运输总产品统计	136
第八章 铁路车辆统计	138
第一节 货车现在车统计	138
第二节 装卸车统计	142
第三节 货车停留时间统计	147
第四节 货车运用效率统计	156
第五节 客车运用效率统计	165
第六节 列车正点统计	165
第七节 车辆检修统计	168
第九章 机车统计	170
第一节 机车统计调查及分组	170
第二节 机车运用效率统计	172
第三节 机车燃料消耗统计	182
第四节 机车检修统计	183
习 题	184

第四篇 道路运输统计

第十章 公路运输统计	190
第一节 公路运输统计范围	190
第二节 公路运输统计调查	191
第三节 运输量统计	193
第四节 运输车辆及其他装备统计	201
第五节 车辆运用统计	204
第六节 公路站场生产统计	217
第七节 公路运输业产值统计	222
第十一章 城市公共交通统计	227
第一节 城市公交统计调查	227
第二节 城市公交车辆统计	228
第三节 城市公交线网及线路统计	231
第四节 城市公交站点及场站统计	234
第五节 城市公交客运情况统计	237
第六节 城市公交其他情况统计	240
第七节 城市出租车统计	243
习 题	246

第五篇 水路、航空及管道运输统计

第十二章 水路运输统计	252
第一节 水路运输与统计简述	252
第二节 水路运输量统计	255
第二节 港口运输统计	259
第四节 船舶营运统计	263
第十三章 航空运输统计	272
第一节 民用航空简介	272
第二节 航空运输业务量统计	278
第三节 飞机数量与利用统计	293
第四节 航空公司运营状况统计	298

第十四章 管道运输统计	301
第一节 管道运输与统计简述	301
第二节 管道运输统计内容	304
习 题	309
习题答案	314
附表 1 标准正态分布表	322
附表 2 t 分布表	324
附表 3 χ^2 分布表	326
附表 4 F 分布表	330
附表 5 相关系数检验表	336
参考文献	337

第一篇

概率论与数理统计

第一章 概率论基础

第一节 随机事件及其概率

一、随机事件

1. 随机现象

在自然界和社会生活中存在着两类不同的现象：一类是在一定条件下必然会发生（或必然不发生）某一种结果的现象，叫做必然现象（或确定性现象）。例如，在标准大气压下，水被加热到 100°C 时必沸腾；异性电荷互相吸引等。另一类是在一定条件下具有多种可能的结果，究竟哪种结果出现，事先不能确定，这类现象称为随机现象。例如，抛掷一枚质量均匀的硬币，落下后可能正面朝上，也可能反面朝上；从一批产品中任取一件，取到的可能是正品，也可能是次品等。

对于随机现象，虽然在个别试验（或观察）中呈现出不确定性，但在大量重复试验（或观察）之后，会发现它具有一定的规律性，这种规律性就是通常所说的统计规律性。

仍以掷硬币为例，有人曾作过大量试验，结果如表 1.1 所示。

表 1.1 抛掷硬币的试验结果

抛掷次数 m	正面朝上的次数 n	n/m
2 048	1 061	0.518 1
4 040	2 048	0.506 9
12 000	6 019	0.501 6
24 000	12 012	0.500 5

很明显，正面朝上的次数与总次数之比趋近于 0.5，且随着抛掷次数的增加而趋于稳定。

随机现象的这种统计规律性是普遍存在的。概率论便是研究随机现象统计规律性的一门数学学科，它已广泛应用于自然科学和社会科学的各个领域以及国民经济的各个部门之中。

2. 随机试验

人们是通过随机试验（亦简称试验，通常用 E 表示）来研究随机现象的。所谓随机试验，是指具有下列特性的试验：

- (1) 可以在相同的条件下重复进行；
- (2) 每次试验的可能结果不止一个，且能事先确定试验的所有可能结果；
- (3) 试验之前不能确切预知出现哪种结果。

需要指出的是，“试验”这个词在这里已经泛化，不应狭隘地理解为“做实验”，它还包括观察、测试、记录、抽样等内容。以下所列便是随机试验的几个例子：

- (1) 抛掷一枚硬币，观察正、反面出现的情况；
- (2) 在一批灯泡中任意抽取一只，测试其寿命；
- (3) 掷一枚骰子，观察出现点数的情况；
- (4) 记录长江流经某地的水位；
- (5) 在一批产品中任意抽取 3 件，记录次品的件数；
- (6) 记录某地出生婴儿的情况（性别、体重等）；
- (7) 在交通路口观察、记录车辆通行情况；
- (8) 观察射击的中靶情况；
- (9) 记录技术站车列牵出消耗的时间。

3. 随机事件

在随机试验中，每一个可能出现的结果叫做该试验的基本事件（也叫样本点），用 ω 表示；全体基本事件组成的集合叫做该试验的样本空间，用 Ω 表示。以掷骰子的随机试验为例，基本事件为：出现 1 点、出现 2 点、……、出现 6 点，分别记作 $\omega_1, \omega_2, \dots, \omega_6$ ，于是，样本空间为

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$$

如果在掷骰子的随机试验中，考察“出现奇数点”这样一个事件 A ，那么当出现 1 点、3 点、5 点时，事件发生；而出现 2 点、4 点、6 点时，事件不发生，用集合表示就是： $A = \{\omega_1, \omega_3, \omega_5\}$ 。像这样在随机试验中可能发生也可能不发生的事件叫做随机事件，简称事件，通常用 $A、B、C$ 等大写字母表示。显然，随机事件是样本空间的一个子集，而基本事件是最简单的不可再分解的随机事件。

在随机试验中，必然发生的事件叫做必然事件，记作 Ω ；不可能发生的事件叫做不可能事件，记作 $\emptyset$ 。比如在掷骰子的随机试验中，“点数小于 7”是必然事件，“点数大于 6”是不可能事件。这两类事件本来具有确定性，但为了讨论问题方便，仍把它们视作两类特殊的随机事件。

【例 1.1】 随机试验：甲、乙两位棋手下两局棋，观察比赛的结果。随机事件 A ：甲两局都不败。试验的结果不外乎如下 9 种：

	第一局	第二局		第一局	第二局
ω_1	甲胜	甲胜	ω_6	乙胜	和局
ω_2	甲胜	乙胜	ω_7	和局	甲胜
ω_3	甲胜	和局	ω_8	和局	乙胜
ω_4	乙胜	甲胜	ω_9	和局	和局
ω_5	乙胜	乙胜			

可知样本空间由 9 个基本事件组成: $\Omega = \{\omega_1, \omega_2, \dots, \omega_9\}$, 随机事件 A 由 4 个基本事件组成: $A = \{\omega_1, \omega_3, \omega_7, \omega_9\}$ 。

二、事件的关系与运算

在研究实际问题时, 常常要根据研究的目的考察随机试验中各种不同的事件, 而这些事件往往是相互关联的, 因此, 需要讨论事件之间的关系及其运算。

1. 事件的包含与相等

若事件 A 发生必然导致事件 B 发生, 就说事件 B 包含事件 A , 记作 $B \supset A$; 或说事件 A 包含于事件 B , 记作 $A \subset B$ 。

若事件 B 包含事件 A , 且事件 A 包含事件 B , 就说事件 A 和事件 B 相等, 记作 $A = B$ 。

例如掷一枚骰子, $A = \{\text{出现 2 点}\}$, $B = \{\text{出现 2 点、4 点、6 点}\}$, $C = \{\text{出现偶数点}\}$, 则有 $A \subset B$, $B \subset C$ 。

2. 事件的和、积、差

事件 A 与事件 B 的和是一个事件, 它表示事件 A 与 B 至少有一个发生, 记作 $A \cup B$ 。事件的和也叫做事件的并。

事件 A 与事件 B 的积是一个事件, 它表示事件 A 与 B 同时发生, 记作 $A \cap B$ 或 (AB) 。事件的积也称为事件的交。

事件 A 与事件 B 的差是一个事件, 它表示事件 A 发生而事件 B 不发生, 记作 $A - B$ 。

事件的和、积、差可以用图形直观地表示出来。图 1.1、图 1.2、图 1.3 中的阴影部分分别表示 $A \cup B$ 、 $A \cap B$ 、 $A - B$ 。

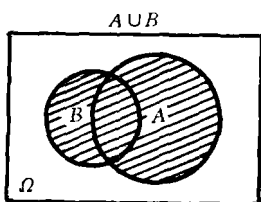


图 1.1

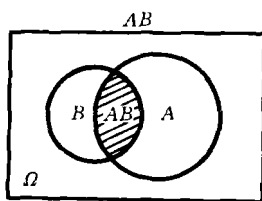


图 1.2

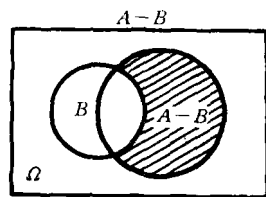


图 1.3

两个事件的和可以推广到多个事件的情况。用 $A_1 \cup A_2 \cup \dots \cup A_n$ 或 $\bigcup_{i=1}^n A_i$ 表示 n 个事件 $A_1, A_2, \dots, A_n$ 的和, 即这 n 个事件中至少有一个发生; 用 $A_1 \cup A_2 \cup \dots \cup A_n \dots$ 或 $\bigcup_{i=1}^{\infty} A_i$ 表示可数个事件, $A_1, A_2, \dots, A_n \dots$ 的和, 即它们之中至少有一个发生。同样, 两个事件的积也可以作类似的推广, n 个事件 $A_1, A_2, \dots, A_n$ 的积表示这 n 个事件同时发生, 记作 $A_1 \cap A_2 \cap \dots \cap A_n$ 或 $\bigcap_{i=1}^n A_i$; 可数个事件 $A_1, A_2, \dots, A_n \dots$ 的积记作 $A_1 \cap A_2 \cap \dots \cap A_n \dots$, 或 $\bigcap_{i=1}^{\infty} A_i$ 。

3. 对立事件

在随机试验中, 若事件 A 和事件 B 必有一个发生, 且仅有一个发生, 即满足条件

$$A \cup B = \Omega, A \cap B = \Phi$$

则称 B 是 A 的对立事件, 也称 A 和 B 互逆, 记为 $B = \bar{A}$ 。

这里, “对立” 具有相互的意义。说 B 是 A 的对立事件, 同时意味着 A 是 B 的对立事件。事实上, $\bar{A}$ 所代表的意义就是: 在试验中, 事件 A 不发生, 除 A 以外的所有事件都发生。所以, 任一事件 A 都有自己的对立事件, 那就是 $\bar{A}$, 如图 1.4 所示。

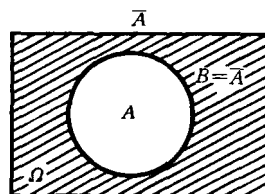


图 1.4

4. 事件的互不相容性

若事件 A 和事件 B 不能同时发生, 即 $A \cap B = \Phi$, 则称 A 和 B 是互不相容事件 (亦称互斥事件), 如图 1.5 所示。显然, 所有的基本事件都是互不相容的。

【例 1.2】 盒中装有 10 个球, 号码为 1, 2, …, 10。今从中任取 1 球, 观察其号码。定义事件 $A_1 \sim A_6$ 如下:

A_1 : 球的号码等于 7,

A_2 : 球的号码小于 8,

A_3 : 球的号码为奇数,

A_4 : 球的号码为偶数,

A_5 : 球的号码大于 5,

A_6 : 球的号码小于 1。

分析这 6 个随机事件, 易知存在下面的关系:

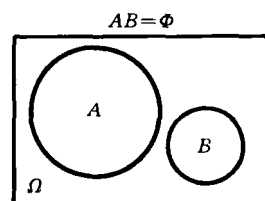


图 1.5

$$A_1 \subset A_2, A_1 \subset A_3, A_1 \subset A_5, A_2 \cap A_3 \cap A_5 = A_1$$

其中, A_3 和 A_4 是对立事件, A_6 是不可能事件, A_1 和 A_4 是互不相容的。

从上面的讨论可以知道, “概率论” 中的随机事件与 “集合论” 中的集合有许多相似之处, 事件之间的关系及运算实质上与集合之间的关系及运算是完全一致的。因此, 可以用集合论的观点来研究随机事件, 利用集合间的运算来分析事件之间的关系。

三、概率的定义和性质

我们知道, 随机事件是随机试验中可能发生也可能不发生的事件。那么, 就某一 (或某些) 事件而言, 它 (们) 发生的可能性究竟有多大呢? 这往往是人们最为关心的问题, 也就是说, 希望用一个数来表示随机事件发生的可能程度。概率就是这样的一个数量指标。

1. 概率的统计定义

把一个随机试验重复 n 次, 如果随机事件 A 共发生 m 次, 则称比值 m/n 为事件 A 在 n 次试验中发生的频率, m 称为 A 发生的频数。无数事例表明, 当 n 充分大时, 频率 m/n 在某一常数左右摆动。随着 n 的继续增大, 频率摆动的幅度越来越小, 逐渐稳定于这个常数, 这就是所谓的频率的稳定性。表 1.1 所示的掷硬币的例子清楚地说明了这一点。下面的例子 (引自文献 [1]) 也反映了频率稳定性这一特性。

从一批产品中任意抽取 n 件, 检查出现废品的件数, 这相当于重复试验 n 次, 观察 “出现废品” 这一事件发生的次数。按 $n=25$ 、 $n=250$ 和 $n=2\,500$ 的三种情况抽样检验, 每种情况抽样 10 次, 得到的结果如表 1.2 所示。

表 1.2 抽取 n 件产品得到的废品件数 m 及频率 m/n

抽样 序号	$n = 25$		$n = 250$		$n = 2\,500$	
	m	m/n	m	m/n	m	m/n
1	1	0.04	12	0.048	157	0.062 8
2	4	0.16	14	0.056	152	0.060 8
3	0	0	17	0.068	157	0.062 8
4	0	0	11	0.044	136	0.054 4
5	1	0.04	22	0.088	152	0.060 8
6	1	0.04	9	0.036	135	0.054 0
7	2	0.08	15	0.06	143	0.057 2
8	0	0	14	0.056	160	0.064 0
9	1	0.04	21	0.084	149	0.059 6
10	1	0.04	8	0.032	153	0.061 2

从表 1.2 可以看出, 随着 n 的增大, 频率 m/n 摆动的幅度的确越来越小。从数值上看, 逐渐稳定的常数可以认为是 0.06, 于是, 这批产品的废品率可以近似地确定为 0.06。

概率的统计定义正是基于频率的稳定性这一客观事实而作出的。

定义 1.1 如果在大量的重复试验中, 随机事件 A 发生的频率稳定地围绕某一个常数上下波动, 且试验的次数越多, 波动的幅度越小, 那么这个常数就称为事件 A 发生的概率, 记作 $P(A)$ 。

从上述定义不难看出, 这个常数的精确程度与试验的次数多少有关, 次数越多越精确, 但不可能进行无限多次的试验, 所以在实际工作中, 当试验次数适当多时, 就用频率作为它的近似值。

2. 概率的数学定义

概率的统计定义比较直观, 容易掌握, 具有很大的实用价值, 但是不要因此而以为欲求某事件的概率, 非进行大量的试验不可。事实上, 事件的概率是该事件的一种固有属性, 是不依赖于试验的客观存在。为了深入研究概率的性质, 数学家柯尔莫哥洛夫提出了概率论的公理化结构, 据此作出的定义就是概率的数学定义, 也称公理化定义。

定义 1.2 设随机试验 E 的样本空间为 Ω , 对于 E 的每一个随机事件 A 赋予一个实数 $P(A)$ 与之对应, 如果它满足下列条件:

- (1) $0 \leq P(A) \leq 1$;
- (2) $P(\Omega) = 1$
- (3) 对于互不相容的事件 $A_i (i=1, 2, \dots)$, 有

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

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

则称 $P(A)$ 为事件 A 的概率。条件 (3) 中的两个式子分别叫做概率的有限可加性和可列可加性。

3. 概率的主要性质

下面不加证明地列出概率的一些主要性质：

(1) 设 $\bar{A}$ 是 A 的对立事件，则 $P(A) = 1 - P(\bar{A})$ 。

(2) $P(\emptyset) = 0$ 。

(3) 设 A 、 B 为两个事件，则

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

这一性质称为概率的加法定理，它可以推广到 n 个事件的情况。例如，当 $n=3$ 时，对 A_1 、 A_2 、 A_3 这 3 个事件，有

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 A_2) - P(A_1 A_3) - P(A_2 A_3) + P(A_1 A_2 A_3)$$

(4) 设 A 、 B 为两个事件，若 $A \subset B$ ，则

$$P(A) \leq P(B), P(B - A) = P(B) - P(A)$$

四、古典概型

古典概型是概率论发展初期被主要研究的概率模型，它适用于具有下列特点的随机试验：

(1) 试验的样本空间只包含有限个基本事件；

(2) 试验中每个基本事件出现的可能性相等（等可能性）。

由于具有第二个特点，古典概型又称为等可能概型。

对于古典概型，随机事件 A 的概率 $P(A)$ 用下式计算：

$$P(A) = \frac{k}{n} \quad (1.1)$$

式中 k —— A 所包含的基本事件数；

n ——样本空间 Ω 包含的基本事件总数。

在计算古典概率时，经常要用到排列组合的有关知识，下面对这部分内容作一简要的复习。

(1) 加法法则。如果完成事件 A 的方法有 a 种，完成事件 B 的方法有 b 种，则完成“事件 A 或事件 B ”的方法共有 $a+b$ 种。

(2) 乘法法则。如果完成事件 A 的方法有 a 种，完成事件 B 的方法有 b 种，则完成“事件 A 与事件 B ”的方法共有 ab 种。

乘法法则也可以叙述为：如果完成一件事需要 m 个步骤，第 1 步有 a_1 种方法，第 2 步有 a_2 种方法，……，第 m 步有 a_m 种方法，则完成该件事总共有 $a_1 a_2 \cdots a_m$ 种方法。

(3) 排列。从 n 个不同的元素中取 r 个 ($r \leq n$) 元素按次序排列, 元素不重复, 其排列种数记作 P_n^r 。当 $r < n$ 时, 称为选排列:

$$P_n^r = \frac{n!}{(n-r)!} = n(n-1)\cdots(n-r+1)$$

当 $r = n$ 时, 称为全排列, 记为 $P_n^n = n!$ 。

(4) 重复排列。从 n 个不同的元素中取 r 个元素按次序排列, 其中元素可以重复出现, 且不要求 $r \leq n$, 其排列种数为 n^r 。

(5) 组合。从 n 个不同的元素中取 r 个元素不计次序地组合在一起, 其组合种数记作 C_n^r 。

$$C_n^r = \frac{n!}{r!(n-r)!}$$

【例 1.3】有 40 件产品, 其中 3 件是次品, 现从中任取两件。设 A_1 : 两件中恰有 1 件次品; A_2 : 两件中无次品; A_3 : 两件中至少有 1 件是次品。求 $P(A_1)$, $P(A_2)$, $P(A_3)$ 。

解 由题意, 样本空间包含的基本事件总数 $n = C_{40}^2 = 780$, A_1 包含的基本事件数为:

$$k_1 = C_3^1 C_{37}^1 = 3 \times 37 = 111$$

故

$$P(A_1) = \frac{k_1}{n} = \frac{111}{780} = 0.1423$$

同理

$$P(A_2) = \frac{k_2}{n} = \frac{C_{37}^2}{C_{40}^2} = \frac{666}{780} = 0.854$$

$$P(A_3) = \frac{k_3}{n} = \frac{C_3^1 C_{37}^1 + C_3^2}{C_{40}^2} = \frac{111+3}{780} = 0.146$$

求 $P(A_3)$ 的另一解法是: 因 A_3 是 A_2 的对立事件, 故

$$P(A_3) = 1 - P(A_2) = 1 - 0.854 = 0.146$$

【例 1.4】某市电话号码由数字 0, 1, ..., 9 中的 7 个数字组成, 问 7 个数字都不相同的概率是多少?

解 由于电话号码中的数字是允许重复的, 属于重复排列, 所以 7 位数中的任何一位都可以从 0~9 这 10 个数字中选取 1 个数字, 有 10 种选法, 故总共有 10^7 种选法, 这就是基本事件的总数。要求 7 个数字都不相同, 就是从 10 个数字中选取 7 个数字不允许重复地排列在一起, 排列总数为 $P_{10}^7 = \frac{10!}{(10-7)!}$ 。由此, 所求概率 $P = \frac{10!}{(10-7)!10^7} = 0.06048$ 。

五、条件概率和事件的独立性

1. 条件概率

设 A, B 是两个随机事件, $P(A) > 0$, 在事件 A 发生的条件下事件 B 发生的概率称为条件概率, 记作 $P(B|A)$ 。