普通高等教育“十二五”规划教材

概率论与数理统计

崔向照　李　春　主编

科学出版社

普通高等教育“十二五”规划教材

概率论与数理统计

崔向照　李　春　主编

科 学 出 版 社

北　京

内 容 简 介

本书主要介绍概率论与数理统计的基本概念、基本原理和基本方法，注重可读性，突出基本思想，适当淡化技巧，力求简明清晰. 内容包括：随机事件与概率、随机变量及其分布、二维随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、数理统计的基础知识、参数估计、假设检验、方差分析与回归分析，其中标星号的章节可根据实际需要选学.

本书可供高等院校经济类、管理类和其他非数学类专业的学生使用，也可作为相关专业学生的教学参考书.

图书在版编目(CIP)数据

概率论与数理统计/崔向照，李春主编 —北京：科学出版社，2016.6
普通高等教育“十二五”规划教材
ISBN 978-7-03-047979-2

Ⅰ. ①概… Ⅱ. ①崔… ②李… Ⅲ. ①概率论-高等学校-教材②数理统计-高等学校-教材 Ⅳ. ①O21

中国版本图书馆 CIP 数据核字 (2016) 第 064146 号

责任编辑：王胡权 陈曰德 / 责任校对：邹慧卿
责任印制：白 洋 / 封面设计：迷底书装

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

大厂书文印刷有限公司 印刷
科学出版社发行 各地新华书店经销
*
2016 年 6 月第 一 版 开本：720 × 1000 1/16
2016 年 6 月第一次印刷 印张：9
字数：181 000

定价：26.00 元
(如有印装质量问题，我社负责调换)

前　言

概率论与数理统计的很多思想、方法已被广泛地应用到经济、交通、医学、气象等各个与人们生活息息相关的领域，运用抽样数据进行推断已经成为现代社会一种普遍适用并且强有力的思考方式. 但随着我国高等教育从“精英型教育”向“大众化教育”转化，教育规模的迅速扩大，给我国大学教育带来了一系列的问题. 例如，现阶段大学的教育正面临生源录取分数下降、教学课时减少、教学内容增加、对学生实践能力的培养要求提高等一系列似乎矛盾的问题.

本书力求在保持体系完整的前提下，尽可能地弱化理论的推导，强化概率论与数理统计的思维训练和知识的实际应用. 在表述上由浅入深，由易及难，由具体到抽象，使得难点分散，便于教学. 在选材和编排上，充分考虑不同层次学生对课程的学习需求，有较大的灵活性. 全部教学内容按 54 学时设计和安排，其中打上“*”号的内容作为选讲内容，可根据学时安排进行取舍; 在习题安排中，有针对性地安排了 A, B 两组习题，A 组习题是为巩固教材中涉及的知识而安排的，B 组习题是为学有余力的学生准备参加研究生考试而安排的.

本书一共由 9 章组成，第 1~5 章主要安排概率论的基础知识及内容，包含随机事件与概率、随机变量及其分布、二维随机变量及其分布、随机变量的数字特征、大数定理与中心极限定理; 第 6~9 章主要安排了数理统计的基础知识及内容，包含数理统计的基础知识、参数估计、假设检验、方差分析与回归分析，适用于经济类、管理类本科各专业学生，也可供非数学类专业的学生作为学习和参考使用.

本书第 1 章由陈灿编写，第 2, 3 章由崔向照编写，第 4, 5 章及附录由李春编写，第 6, 7 章由赵金娥编写，第 8, 9 章由曾黎编写，全书由崔向照进行统稿.

本书在编写时参考了大量的相关教材和文献资料，并选用了某些内容和习题，在此一并向相关作者表示感谢!

由于编者水平有限，书中难免存在疏漏之处，恳请各位同仁批评指正.

编　者

2015 年 6 月 30 日

目　　录

第 1 章　随机事件与概率

1.1　随机事件及其运算

1.1.1　随机现象

自然界中存在两类现象, 一类是在一定条件下必然出现某一种结果的现象, 称为**确定性现象 (必然现象)**, 例如, 异性电荷会相互吸引; 水在标准大气压下加热到 100°C 沸腾; 在自然状态下, 水从高处往低处流淌等. 另一类是在一定条件下并不总出现相同结果的现象, 称为**随机现象**, 例如, 在抛掷一枚均匀硬币时, 不能事先预知出现正面还是反面; 在股票交易中不能预知下一个交易日某只股票的价格; 在生育前, 不能预知生男孩还是生女孩等.

虽然随机现象在一定条件下可能出现不同的结果, 并且在每次观测前不能预知出现哪个结果, 但经过长期、反复地观测, 人们发现所谓的 "不能预知" 只是对一次或少数几次观测而言的. 当在相同条件下进行重复试验和观测时, 试验结果就会出现某种规律性, 例如, 大量重复地抛掷一枚均匀硬币, 随着抛掷次数的增加, 出现正面的频率会逐渐稳定在 0.5. 随机现象的这种在大量重复试验中呈现出来的规律性称之为**统计规律性**. 概率论与数理统计就是研究和揭示随机现象统计规律的数学学科.

1.1.2　随机试验与样本空间

为了研究随机现象的统计规律, 需要对随机现象在相同条件下进行重复观测, 我们称这种观测为**随机试验**, 记为 E. 下面是一些试验的例子.

(1) 向一个靶射击, 对射击环数进行观测;

(2) 抛一枚硬币, 对出现正面和反面情况进行观测;

(3) 掷一枚骰子, 对出现点数的情况进行观测;

(4) 考察一批炮弹的质量, 对不爆炸的炮弹数进行观测;

(5) 观测某一种品牌手机的使用寿命.

从上面的试验可以发现随机试验具有以下特点:

(1) 试验在相同条件下可以重复进行;

(2) 每次试验的结果不止一个, 但事先明确知道试验的所有可能结果;

(3) 在试验前不能确定哪个结果会出现.

随机试验的每一个可能结果称为**样本点**, 记为 ω; 所有样本点构成的集合称为**样本空间**, 记为 $\Omega = \{\omega\}$, 样本空间通常分为两类. 若样本空间中样本点的个数为有限个或可列个, 则称为**离散样本空间**; 若样本空间中样本点的个数为无限不可列个, 则称为**连续样本空间**.

例如, 在上述试验中, 相对应的样本空间分别为 $\Omega = \{0, 1, 2, \cdots, 10\}$, $\Omega_2 = \{$正, 反$\}$, $\Omega_3 = \{1, 2, 3, 4, 5, 6\}$, $\Omega_4 = \{0, 1, 2, 3, \cdots, n\}$, $\Omega_5 = \{t | \ t \geqslant 0\}$, 其中 Ω_1, Ω_2, Ω_3, Ω_4 是离散样本空间, Ω_5 为连续样本空间.

1.1.3 随机事件

样本空间 Ω 中的某些样本点组成的集合称为**随机事件**, 常用 $A, B, C, \cdots$ 表示, 它是 Ω 的子集. 若某随机事件中的样本点只有一个, 则称该事件为**基本事件**; 若某随机事件在一次试验中一定发生, 则称该事件为**必然事件**, 记为 Ω; 若某随机事件在一次试验中一定不发生, 则称该事件为**不可能事件**, 记为 $\varnothing$.

例如, 在抛掷一枚骰子的试验中, 其样本空间为$\Omega=\{1, 2, 3, 4, 5, 6\}$, 而 $A=\{1,3,5\}$ 是一个事件, $B = \{1\}$ 是一个基本事件, Ω 是一个必然事件, $\varnothing$ 是一个不可能事件.

若事件 A 中的某个样本点出现了, 则说事件 A 发生了.

例如, 在抛掷一枚骰子的试验中, 事件 $A = \{2, 4, 6\}$ 表示出现偶数点, 当抛掷结果出现了 2, 4, 6 中的任何一个点数时, 我们都认为出现偶数点这一事件即 A 发生了.

1.1.4 事件的关系及运算

由于事件 A 是样本空间 Ω 的一个子集, 因此事件的关系及运算可按集合间的关系及运算来处理.

1. 包含关系

若属于 A 的样本点必属于 B, 则称A **包含于** B或称B **包含** A, 记为 $A \subset B$ 或 $B \supset A$, 用概率论语言描述为: 事件 A 发生必然导致事件 B 发生.

例如, 在抛掷一枚骰子的试验中, $A = \{4\}$, $B = \{2, 4, 6\}$, 则有 $A \subset B$, 对任意事件 A, 一定有 $\varnothing \subset A \subset \Omega$.

2. 相等关系

若事件 A 与 B 相互包含, 则称事件 A 与 B **相等**, 记为 $A = B$.

例如, 在抛掷一枚骰子的试验中, $A = \{1, 3, 5\}$, $B = \{$奇数$\}$, 则 $A = B$.

3. 互不相容

若 A 与 B 没有相同的样本点, 则称 A 与 B **互不相容**. 用概率论语言描述为:

事件 A 与 B 不可能同时发生.

例如, 在抛掷一枚骰子的试验中, $A=\{3,5\}$, $B=\{2,4,6\}$, 则 A 与 B 互不相容.

4. 事件的并

由属于 A 或属于 B 的样本点构成的集合, 称为事件 A 与 B 的**并**, 记为 $A\cup B$. 用概率论语言描述为: 事件 A 与 B 至少有一个发生.

例如, 在抛掷一枚骰子的试验中, $A=\{1,3,5\}$, $B=\{1,2,3\}$, 则 $A\cup B=\{1,2,3,5\}$.

类似地, n 个事件 $A_1, A_2, \cdots, A_n$ 中至少有一个发生, 称为 $A_1, A_2, \cdots, A_n$ 的**并**, 记为 $\bigcup\limits_{i=1}^{n} A_i$; 可列个事件 $A_1, A_2, \cdots$ 中至少有一个发生, 则称为 $A_1, A_2, \cdots$ 的**并**, 记为 $\bigcup\limits_{i=1}^{+\infty} A_i$.

5. 事件的交

由既属于 A 又属于 B 的样本点构成的集合, 称为事件 A 与 B 的**交**, 记为 AB. 用概率论语言描述为: 事件 A 与 B 同时发生.

例如, 在抛掷一枚骰子的试验中, $A=\{1,3,5\}$, $B=\{1,2,3\}$, 则 $AB=\{1,3\}$.

类似地, n 个事件 $A_1, A_2, \cdots, A_n$ 同时发生, 称为 $A_1, A_2, \cdots, A_n$ 的**交**, 记为 $\bigcap\limits_{i=1}^{n} A_i$; 可列个事件 $A_1, A_2, \cdots$ 同时发生, 则称为 $A_1, A_2, \cdots$ 的**交**, 记为 $\bigcap\limits_{i=1}^{+\infty} A_i$.

6. 事件的差

由属于 A 但不属于 B 的样本点构成的集合, 称为 A 与 B 的**差**, 记为 $A-B$. 用概率论语言描述为: 事件 A 发生而 B 不发生.

例如, 在抛掷一枚骰子的试验中, $A=\{1,2,3\}$, $B=\{2,4,6\}$, 则 $A-B=\{1,3\}$.

7. 对立事件

事件 A 不发生称为事件 A 的**对立事件**, 记为 $\bar{A}$.

例如, 在抛掷一枚骰子的试验中, $A=\{3,6\}$, 则 $\bar{A}=\{1,2,4,5\}$.

两个事件 A 与 B 互为对立事件的充要条件为 $A\cup B=\Omega$, $AB=\varnothing$.

显然有 $A-B=A\bar{B}$.

概率论描述与集合论描述对照表如下.

符号	概率论描述	集合论描述
$A \subset B$	事件 A 发生必然导致事件 B 发生	集合 A 是 B 的子集
$A = B$	事件 A 与事件 B 相等	集合 A 与 B 相等
$A \cup B$	事件 A 与 B 至少有一个发生	集合 A 与 B 的并集
AB	事件 A 与 B 同时发生	集合 A 与 B 的交集
$A - B$	事件 A 发生而 B 不发生	集合 A 与 B 的差集
$AB = \varnothing$	事件 A 与 B 不可能同时发生	集合 A 与 B 没有公共元素
$\bar{A}$	事件 A 的对立事件	集合 A 的补集

8. 事件的运算性质

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

(2) 结合律: $(A \cup B) \cup C = A \cup (B \cup C)$, $(AB)C = A(BC)$;

(3) 分配律: $(A \cup B)C = (AC) \cup (BC)$, $(AB) \cup C = (A \cup C)(B \cup C)$;

(4) (de Morgan 公式): $\overline{A \cup B} = \bar{A}\bar{B}$, $\overline{AB} = \bar{A} \cup \bar{B}$, 可推广为 $\overline{\bigcup\limits_{i=1}^{+\infty} A_i} = \bigcap\limits_{i=1}^{+\infty} \bar{A}_i$, $\overline{\bigcap\limits_{i=1}^{+\infty} A_i} = \bigcup\limits_{i=1}^{+\infty} \bar{A}_i$.

例 1.1　设 A, B, C 是三个事件, 则

(1) A 与 B 发生而 C 不发生可表示为 $AB\bar{C}$;

(2) A 发生而 B 与 C 都不发生可表示为 $A\bar{B}\bar{C}$;

(3) 三个事件恰好有一个发生可表示为 $A\bar{B}\bar{C} \cup \bar{A}B\bar{C} \cup \bar{A}\bar{B}C$;

(4) 三个事件恰好有两个发生可表示为 $AB\bar{C} \cup A\bar{B}C \cup \bar{A}BC$;

(5) 三个事件都发生可表示为 ABC;

(6) 三个事件都不发生可表示为 $\bar{A}\bar{B}\bar{C}$;

(7) 三个事件中至少有一个不发生可表示为 $\bar{A} \cup \bar{B} \cup \bar{C}$ 或 $\overline{ABC}$.

1.2　概率的定义及确定方法

随机事件在一次试验中可能发生也可能不发生, 人们希望知道的是随机事件在一次试验中发生的可能性有多大、如何度量? 在概率论与数理统计中, 人们用概率来度量随机事件发生的可能性大小.

在概率论的发展史上, 曾出现过概率的统计定义、古典定义、几何定义和公理化定义, 下面分别介绍它们.

1.2.1　概率的统计定义

定义 1.1　在相同条件下进行 n 次试验, 事件 A 发生的次数 n_A 称为 A 发生的**频数**. 比值 $\dfrac{n_A}{n}$ 称为 A 发生的**频率**, 记为 $f_n(A)$. 易见频率具有以下性质.

(1) **非负性** $0 \leqslant f_n(A)$;

(2) **规范性** $f_n(\Omega)=1$;

(3) **有限可加性** 若事件 $A_1, A_2, \cdots, A_m$ 两两互不相容, 则

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

人们长期的实践表明, 随着试验次数 n 的增加, 事件 A 发生的频率 $f_n(A)$ 会逐步稳定到一个常数. 如下表所示, 历史上著名的投硬币试验.

实验者	投硬币次数	出现正面次数	频率
De Morgan	2048	1061	0.5181
Buffon	4040	2048	0.5069
Feller	10000	4979	0.4979
Pearson	12000	6019	0.5016
Pearson	24000	12012	0.5005

随着试验次数的增加, 正面出现的频率逐步稳定到常数 0.5.

定义 1.2 事件 A 发生的频率 $f_n(A)$ 随着试验次数 n 的增加逐步稳定到某个常数 p, 称这个常数 p 为事件 A 发生的**概率**, 记为 $p(A)$.

1.2.2 概率的古典定义

若随机试验满足:

(1) 试验结果只有有限个;

(2) 每个结果发生的可能性相等,

则称这种随机试验为**古典概型**.

定义 1.3 在古典概型中, 若样本空间包含 n 个样本点, 事件 A 包含 k 个样本点, 则事件 A 发生的概率为

$$p(A)=\frac{k}{n}.$$

上述确定概率的方法称为**古典方法**. 用古典方法确定概率的关键是计算 A 包含的样本点数和样本空间包含的样本点数, 因而常常要用到排列组合知识.

例 1.2 n 个人围一圆桌坐, 求甲、乙两人相邻而坐的概率?

解 考虑甲先坐好, 则乙有 $n-1$ 个位置可坐, 而 "甲乙相邻" 只有两种情况, 从而, 所求概率为

$$p(A)=\frac{2}{n-1}.$$

例 1.3 n 个人坐成一排, 求甲、乙两人相邻而坐的概率?

解 (1) 先考虑样本空间的样本点数, 甲先坐、乙后坐, 则共有 $n(n-1)$ 种可能;

(2) 甲在两端, 则乙与甲相邻共有 2 种可能;

(3) 甲在中间 $n-2$ 个位置上, 则乙左右都可坐, 于是共有 $2(n-2)$ 种可能.

因此, 所求概率为

$$p(A)=\frac{2+2(n-2)}{n(n-1)}=\frac{2}{n}.$$

1.2.3 概率的几何定义

若随机试验满足:

(1) 试验结果充满某个可度量的区域;

(2) 每个结果发生的可能性相等,

则称这种随机试验为**几何概型**.

定义 1.4 在几何概型中, 若样本空间的度量为 S_Ω, 事件 A 的度量为 S_A, 则事件 A 发生的概率为

$$p(A)=\frac{S_A}{S_\Omega}.$$

上述确定概率的方法称为**几何方法**. 用几何方法确定概率的关键是将样本空间 Ω 和所求事件 A 对应到一个可度量的空间区域上.

例 1.4 设公共车站每 5 分钟开过一趟车, 而乘客到达车站的时间是任意的, 求每个乘客到车站后等车的时间不超过 3 分钟的概率?

解 记 $A=\{$每个乘客到达车站后等车不超过 3 分钟$\}$. 由于乘客可能在两趟车之间的任一时刻到达, 因而其到达时间落在区间 $[0,5]$ 内; 要使乘客到车站后等车的时间不超过 3 分钟, 其到达时间必须落在区间 $[2,5]$ 内, 于是有样本空间的度量 $S_\Omega=5$, 事件 A 的度量 $S_A=3$, 所以

$$p(A)=\frac{S_A}{S_\Omega}=\frac{3}{5}.$$

例 1.5 甲乙二人相约在 2 点到 3 点之间在某地会面, 约定先到者等候另一人 20 分钟, 过时即可离去. 如果每个人可在指定的一小时内的任意时刻到达, 试求二人能够会面的概率?

解 设 x, y 分别为甲乙二人到达会面地点的时间, 在平面上建立直角坐标系, 如图 1.1 所示, 则样本空间为

$$\{(x,y)|0 \leqslant x \leqslant 60, 0 \leqslant y \leqslant 60\}.$$

记 $A = \{$两人能会面$\}$, 则有

$$A = \{(x,y)|(x,y) \in \Omega \text{且} |x-y| \leqslant 20\},$$

$$\Rightarrow p(A) = \frac{S_A}{S_\Omega} = \frac{60^2 - 40^2}{60^2} = \frac{5}{9}.$$

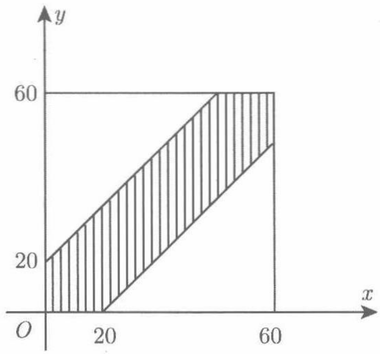

图 1.1 例 1.5 用图

1.2.4 概率的公理化定义

前面讨论的概率的统计定义、古典定义和几何定义都带有一定局限性, 为此, 人们引进概率的公理化定义, 使概率的定义更具有普遍性和严密性.

定义 1.5 设 Ω 是随机试验的样本空间, 若对任意事件 A, 均有唯一的一个实数 $p(A)$ 与之对应, 且 $p(A)$ 满足下列条件:

(1) **非负性** 对任意事件 A, 有 $p(A) \geqslant 0$;

(2) **规范性** $p(\Omega) = 1$;

(3) **可列可加性** 若 $A_1, A_2, \cdots$ 是两两互不相容的一列事件, 则有

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

则称 $p(A)$ 为事件 A 发生的**概率**.

由概率的公理化定义出发, 可得到概率具有以下性质:

性质 1.1 $p(\varnothing) = 0$.

性质 1.2 若 $A_1, A_2, \cdots, A_n$ 是两两互不相容的事件, 则有

$$p\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} p(A_i).$$

性质 1.3 对任意事件 A, 有 $p\left(\bar{A}\right) = 1 - p(A)$.

性质 1.4 对任意两个事件 A, B, 有 $p(A-B) = p(A) - p(AB)$.

特别地, 若 $B \subset A$, 则有 $p(A-B) = p(A) - p(B)$ 且 $p(A) \geqslant p(B)$.

性质 1.5 对任意事件 A, 有 $0 \leqslant p(A) \leqslant 1$.

性质 1.6 对任意两个事件 A, B, 有 $p(A \cup B) = p(A) + p(B) - p(AB)$.

例 1.6 设 $p(A) = 0.5, p(B) = 0.3, p(AB) = 0.2$, 试求下列概率:

(1) $p\left(\bar{A} \cup \bar{B}\right)$;

(2) $p(A\bar{B})$;

(3) $p(A \cup \bar{B})$.

解　(1) $p(\bar{A} \cup \bar{B}) = p(\overline{AB}) = 1 - p(AB) = 1 - 0.2 = 0.8$;

(2) $p(A\bar{B}) = p(A - AB) = p(A) - p(AB) = 0.5 - 0.2 = 0.3$;

(3) $p(A \cup \bar{B}) = p(A) + p(\bar{B}) - p(A\bar{B}) = 0.5 + 0.7 - 0.3 = 0.9$.

例 1.7　已知 $p(B) = 0.3$, $p(A \cup B) = 0.6$, 求 $p(A\bar{B})$?

解　由于 $p(B) = 0.3$ 和 $p(A \cup B) = p(A) + p(B) - p(AB) = 0.6$, 因而

$$\begin{aligned} & p(A) - p(AB) = 0.3 \\ \Rightarrow & p(A\bar{B}) = p(A - B) = p(A) - p(AB) = 0.3. \end{aligned}$$

1.3 条件概率

1.3.1 条件概率

在实际问题中, 常常要考虑在事件 A 已发生的条件下, 事件 B 发生的概率, 称为**在事件 A 发生的条件下事件 B 发生的条件概率**, 简称**条件概率**, 记为 $p(B|A)$.

例 1.8　考察有两个小孩的家庭, 其样本空间 $\Omega = \{$(男男), (男女), (女男), (女女)$\}$, 在 Ω 中 4 个样本点等可能的情况下, 记 $A = \{$第一个是女孩$\}$, $B = \{$第二个是男孩$\}$, 讨论下列事件的概率.

(1) $p(A)$;　(2) $p(AB)$;　(3) $p(B|A)$.

解　(1) $p(A) = \dfrac{2}{4}$;

(2) $p(AB) = \dfrac{1}{4}$;

(3) $p(B|A) = \dfrac{1}{2} = \dfrac{p(AB)}{p(A)}$.

此关系具有一般性, 于是可得到条件概率的定义.

定义 1.6　设 A, B 是 Ω 中的两个事件, 若 $p(A) > 0$, 则称

$$p(B|A) = \frac{p(AB)}{p(A)}$$

为**在事件 A 发生的条件下事件 B 发生的条件概率**, 简称**条件概率**.

不难验证条件概率 $p(B|A)$ 满足概率的三个条件.

例 1.9　一盒子装有 5 件产品, 其中一等品 3 件、二等品 2 件, 从中不放回地任取两次, 每次取一件, 设 $A = \{$第一次取到一等品$\}$, $B = \{$第二次取到一等品$\}$, 求 $p(B|A)$?

解 方法 1: 样本空间的样本点总数为 $\mathrm{P}_5^2=20$, 事件 A 包含的样本点数为 $3\times 4=12$ 种取法, AB 包含的样本点数为 $\mathrm{P}_3^2=6$ 种取法, 于是

$$p(A)=\frac{12}{20},\quad p(AB)=\frac{6}{20}\Rightarrow p(B|A)=\frac{p(AB)}{p(A)}=\frac{6/20}{12/20}=\frac{1}{2}.$$

方法 2: 直接在缩小的样本空间上考虑

$$p(B|A)=\frac{2}{4}=\frac{1}{2}.$$

1.3.2 乘法公式

由条件概率的定义可直接得到下述结果.

定理 1.1 对于任意事件 A, B, 若 $p(A)>0$, 则有 $p(AB)=p(A)p(B|A)$; 若 $p(B)>0$, 则有 $p(AB)=p(B)p(A|B)$.

上述公式称为概率的**乘法公式**. 该公式可推广到有限多个事件的情形, 例如, 对于三个事件 A_1, A_2, A_3, 若 $p(A_1A_2A_3)>0$, 则有

$$p(A_1A_2A_3)=p(A_1)p(A_2|A_1)p(A_3|(A_1A_2)).$$

例 1.10 10 个产品中有 7 个正品、3 个次品, 从中无放回地抽取两个, 已知第一个取到次品, 求第二个又取到次品的概率?

解 设 $A=\{$第一个取到次品$\}$, $B=\{$第二个取到次品$\}$, 则

$$p(A)=\frac{3}{10},\quad p(AB)=\frac{1}{15}.$$

于是, 所求概率为

$$p(B|A)=\frac{p(AB)}{p(A)}=\frac{1/15}{3/10}=\frac{2}{9}.$$

例 1.11 一批零件有 100 个, 其中有 10 个不合格品, 从中一个一个取出, 求第三次才取得不合格品的概率?

解 记 $A_i=\{$第 i 次取出的是不合格品$\}(i=1,2,3)$, 则所求的概率为

$$\begin{aligned}p(\bar{A}_1\bar{A}_2A_3)=&p(\bar{A}_1)p(\bar{A}_2|\bar{A}_1)p(A_3|(\bar{A}_1\bar{A}_2))\\=&\frac{90}{100}\times\frac{89}{99}\times\frac{10}{98}=0.826.\end{aligned}$$

1.3.3 全概率公式

在概率的计算中, 人们常常希望通过已知简单事件的概率去求未知的复杂事件的概率, 为此, 需将复杂事件分解为若干个简单事件, 计算出每个简单事件的概率, 再利用概率的性质计算出复杂事件的概率. 下面介绍一种基本的分解方法.

定义 1.7 若事件 $A_1, A_2, \cdots, A_n$ 满足

(1) $A_iA_j = \varnothing(i, j = 1, 2, \cdots, n$ 且 $i \neq j)$;

(2) $\bigcup\limits_{i=1}^{n} A_i = \Omega$,

则称 $A_1, A_2, \cdots, A_n$ 为 Ω 的一个**完备事件组**.

定理 1.2 (全概率公式) 若 $A_1, A_2, \cdots, A_n$ 为 Ω 的一个完备事件组, 则对任意的事件 B, 有

$$p(B) = \sum_{i=1}^{n} p(A_i)p(B|A_i).$$

全概率公式最简单的形式为

$$p(A) = p(B)p(A|B) + p(\bar{B})p(A|\bar{B}).$$

例 1.12 设 10 件产品中有 3 件不合格品, 从中无放回地取两次, 每次一件, 求取出的第二件为不合格品的概率?

解 设 $A =$\{第一次取得不合格品\}, $B =$\{第二次取得不合格品\}, 则由全概率公式得

$$\begin{aligned} p(B) =& p(A)p(B|A) + p(\bar{A})p(B|\bar{A}) \\ =& \frac{3}{10} \times \frac{2}{9} + \frac{7}{10} \times \frac{3}{9} = \frac{3}{10}. \end{aligned}$$

例 1.13 某工厂有甲、乙、丙三个车间生产同一批产品, 三个车间的产量分别占总产量的 40%, 30%, 30%, 其产品的合格率分别为 0.96, 0.95, 0.90, 求从这批产品中任取一件是不合格品的概率?

解 用 A_1, A_2, A_3 分别表示产品由甲、乙、丙三个车间生产, B 表示取到的产品是不合格品, 则 A_1, A_2, A_3 构成完备事件组, 从而

$$\begin{aligned} p(B) =& p(A_1)p(B|A_1) + p(A_2)p(B|A_2) + p(A_3)p(B|A_3) \\ =& 0.4 \times 0.04 + 0.3 \times 0.05 + 0.3 \times 0.1 = 0.061. \end{aligned}$$

1.3.4 贝叶斯公式

全概率公式解决的问题是借助于一个完备事件组 $A_1, A_2, \cdots, A_n$ 来计算某事件 B 发生的概率, 而下面的公式则刚好相反, 是在已知某个事件 B 发生的条件下, 求完备事件组中某个事件 A_i 发生的条件概率.

定理 1.3 (贝叶斯公式) 设 $A_1, A_2, \cdots, A_n$ 为一完备事件组, 对任意事件 B, 若 $p(B) > 0$, 则有

$$p(A_i|B) = \frac{p(A_i)p(B|A_i)}{\sum\limits_{j=1}^{n} p(A_j)p(B|A_j)}.$$

若把 A_i 看作是事件 B 发生的各个原因 (或条件), 则该公式描述的是: 已知出现了事件 B, 要求出事件 B 产生的各个原因 (或条件)A_i 的概率.

例 1.14 在例 1.13 中, 若已知抽到的一件产品是不合格品, 问这件不合格品是由哪个车间生产的可能性大?

解 在例 1.13 的记号下, 有

$$p(A_1|B)=\frac{p(A_1)p(B|A_1)}{\sum\limits_{j=1}^{3}p(A_j)p(B|A_j)}=\frac{0.4\times0.04}{0.4\times0.04+0.3\times0.05+0.3\times0.1}=0.2623,$$

$$p(A_2|B)=\frac{p(A_2)p(B|A_2)}{\sum\limits_{j=1}^{3}p(A_j)p(B|A_j)}=\frac{0.3\times0.05}{0.4\times0.04+0.3\times0.05+0.3\times0.1}=0.2459,$$

$$p(A_3|B)=\frac{p(A_3)p(B|A_3)}{\sum\limits_{j=1}^{3}p(A_j)p(B|A_j)}=\frac{0.3\times0.1}{0.4\times0.04+0.3\times0.05+0.3\times0.1}=0.4918.$$

由于 $p(A_3|B)$ 最大, 故认为这件不合格品是由第三车间生产的可能性大.

例 1.15 甲胎蛋白免疫检测法被普遍应用于肝癌的普查和诊断. 记 $A=\{$肝癌患者$\}$, $B=\{$检测反应为阳性$\}$, 已知 $p(B|A)=0.94$, $p(B|\bar{A})=0.04$, 由统计数据知, 人群中肝癌的发病率一般为 $p(A)=0.0004$, 现有一人检测结果为阳性, 问该人患肝癌的概率是多大?

解 该人患肝癌的概率为

$$\begin{aligned}p(A|B)&=\frac{p(A)p(B|A)}{p(A)p(B|A)+p(\bar{A})p(B|\bar{A})}\\&=\frac{0.0004\times0.94}{0.0004\times0.94+0.9996\times0.04}=0.0093.\end{aligned}$$

即该人患肝癌的概率只有 0.93%, 连 1% 都不到.

$p(A|B)$ 不大的原因是人群中肝癌的发病率非常低, 仅有 0.0004, 因而在对稀有病例的普查中, 一次检测为阳性者, 其实际患此病的概率并不大.

1.4 事件的独立性

1.4.1 事件的独立性

一般情况下 $p(B|A)\neq p(B)$, 这表明事件 A 的发生对事件 B 发生的概率产生了影响. 但在许多实际问题中, 常常会遇到两个事件的发生互相不影响的情况, 即