

大学数学系列教材

概率论与数理统计

主编 干晓蓉

013067791

大学数学系列教材

021-43
228

概率论与数理统计

Gailülun yu Shuli Tongji

主编 干晓蓉



北航

C1674081



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内容提要

本书内容包含随机事件及其概率、随机变量及其分布、随机变量的数字特征、样本及抽样分布、参数估计、假设检验、方差分析和回归分析共七章。每章都配有一定数量的习题,书末附有习题答案与提示。全书叙述简明扼要,条理清楚,通俗易懂,方便教学。

本书可作为普通高等学校应用型本科各专业的概率论与数理统计教材,也可作为成人教育各专业的概率论与数理统计教材或教学参考书。

图书在版编目(CIP)数据

概率论与数理统计/干晓蓉主编. --北京:高等教育出版社,2013.8

ISBN 978-7-04-038191-7

I. ①概… II. ①干… 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(2013)第 174239 号

策划编辑 张长虹
插图绘制 尹 莉

责任编辑 张长虹
责任校对 杨凤玲

封面设计 姜 磊
责任印制 刘思涵

版式设计 余 杨

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100120
印 刷 山东鸿杰印务集团有限公司
开 本 787 mm × 960 mm 1/16
印 张 12
字 数 220 千字
购书热线 010-58581118

咨询电话 400-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landaco.com>
<http://www.landaco.com.cn>
版 次 2013 年 8 月第 1 版
印 次 2013 年 8 月第 1 次印刷
定 价 18.10 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换
版权所有 侵权必究
物 料 号 38191-00

总 序

在我国改革开放的大潮流中,大学数学教学改革不断深入,本科办学层次亦不断地细化和丰富。适应于各层次各类别的学生学习的大学数学教材相继出版,呈现出百花齐放的可喜态势。近年来,以高职本科和独立学院为代表的应用型本科教育异军突起,招生规模迅猛发展,为我国高等教育的应用型人才培养做出了重要贡献。

以微积分、线性代数、概率论与数理统计为主要内容的大学数学课程,作为理学、工程、农林、经济、管理类专业学生的重要基础课程,担负着向学生传授各自专业所必需的数学基础知识,培养逻辑思维能力,提高数学素养的重要作用。对应用型本科学生而言,更重要的是能够运用这些数学基础知识来理解专业学科的知识,并解决将来在工作中的实际问题,也为后续的专业学习和未来的工作需要打下坚实的基础。

于晓蓉教授主编的“大学数学系列教材”(含《高等数学》、《线性代数》、《概率论与数理统计》),以教育部高等学校数学与统计学教学指导委员会颁布的工科类、经济管理类本科数学基础课程教学基本要求为基础,立足于应用型本科学生的实际情况,结合编写团队多年来在不同层次、不同类别专业上的教学经验,以理工类专业为主,兼顾经济管理类专业。对理论知识的讲解深入浅出,注重基本概念、基本运算、基本应用,尽可能联系实际问题,引导学生学以致用。本书配置了丰富的典型例题和习题,便于学生巩固基本知识,并有一定数量的提高题,为学生考研做准备。

教学内容和课程体系的改革是教学改革的重点和难点。鼓励不同层次、不同类别、不同模式、不同侧重的各种教学改革,是当前高等教育发展的必然趋势。我相信这套有自身鲜明特色教材的出版,将为应用型本科数学的教材建设添砖加瓦,必将为大学数学课程教学质量和教学效果的提高、为学生的数学素养和应用能力的培养做出积极的贡献。

李继彬

2012年7月25日

前 言

本书是为高等学校应用型本科各专业编写的概率论与数理统计教材,也是我们所编写的大学数学系列教材之一。本书内容包含随机事件及其概率、随机变量及其分布、随机变量的数字特征、样本及抽样分布、参数估计、假设检验、方差分析和回归分析共七章。

本书在选材上,着重介绍概率论与数理统计中最基本的基础知识;在内容安排上注意到方便教学、节省教学时间和有利学生的学习;在叙述上尽量做到概念清晰,说理清楚,演算过程详尽,注意利用例题加深学生对概念的理解,加强对所学方法的掌握;在例题和习题中都配有一定数量的应用题,通过这些应用题的学习,可以增加学生的学习兴趣,提高应用能力,对于较难的习题,在书末的答案中还给出一定的提示。所有这一切,都是为了使学生能把本书所介绍的知识学懂、学好,达到本课程的教学要求。书中有些加了*的内容,教学中可以不讲,供学生参阅。

本书由于晓蓉教授任主编,并编写了第一章和第二章内容,王彭德、郭民之、沙春宏任副主编,参加了全书的审稿工作和第六章的编写工作。参加教材编写的教师还有:王跃虹(第三章)、杨斌(第四章)、董艳梅(第五章)、李庶民(第七章),周绍艳、李振鹏老师对书稿进行了认真细致的校对。全书由于晓蓉统稿定稿。

限于作者水平,书中难免会有不妥之处,敬请读者给予批评、指正。

编 者

2013年6月

目 录

第一章 随机事件及其概率	1
第一节 随机事件	1
第二节 随机事件的概率	4
第三节 条件概率	13
第四节 事件的独立性	19
习题 1	22
第二章 随机变量及其分布	26
第一节 随机变量及其分布函数	26
第二节 离散型随机变量	28
第三节 连续型随机变量	34
第四节 二维随机变量	42
第五节 边缘分布及条件分布	46
第六节 随机变量的独立性	54
第七节 随机变量的函数的分布	57
习题 2	64
第三章 随机变量的数字特征	68
第一节 数学期望	68
第二节 方差	73
第三节 协方差与相关系数	77
第四节 大数定律与中心极限定理	81
习题 3	85
第四章 样本及抽样分布	90
第一节 随机样本及统计量	90
第二节 抽样分布	94
习题 4	101
第五章 参数估计	104
第一节 参数的点估计	104
第二节 参数的区间估计	113
习题 5	122

第六章 假设检验	125
第一节 假设检验的基本概念	125
第二节 单个正态总体的均值与方差的假设检验	128
第三节 两个正态总体的均值及方差的比较	134
第四节 分布拟合检验	137
习题 6	140
* 第七章 方差分析和回归分析	144
第一节 单因素试验的方差分析	144
第二节 一元线性回归分析	150
习题 7	159
附表	163
附表 1 标准正态分布表	163
附表 2 泊松分布表	164
附表 3 t 分布表	165
附表 4 χ^2 分布表	166
附表 5 F 分布表	167
部分习题答案与提示	172

第一章 随机事件及其概率

在科学研究和社会生活中,经常要对某些事物进行试验.这里说的试验,含义比较广泛,可以是科学试验,也可以只是对事物的一种观察.在各种试验中,常会看到两类现象.一类是在一定条件下必然发生的现象,称为**确定性现象**.例如,在空地上用手向上抛一块石子,石子必然还会落下来;在标准大气压下将水加热到 100°C ,水必然会沸腾.这些都是确定性现象.另一类现象是在一定条件下,有多种试验结果可能发生,在一次试验中,发生哪一种结果纯属偶然,预先不知道哪一种结果会发生,这类现象称为**随机现象**.例如,抛一枚硬币,落下后可能出现正面(数字面),也可能出现反面(图案面),不能预先知道出现哪一面;掷一颗骰子,出现的点数有6种,不能预先知道会出现哪一种点数;从一大批产品中任取一件产品,取到的可能是正品,也可能是次品,预先不知道会取到哪一种产品.这些都是随机现象.概率论是研究随机现象的数量规律性的数学学科,是数理统计的理论基础,或者说,数理统计是概率论的应用.本书前三章介绍概率论的基本知识,后四章则是数理统计的内容.

随机事件及其概率,是概率论中的两个最基本的概念.本章讨论与这两个基本概念有关的一些概念及计算公式,它们是进一步学习概率论的基础.

第一节 随机事件

一、随机试验与样本空间

满足以下三个条件的试验,称为**随机试验**:

- (i) 试验的所有可能结果是已知的或者可以预先确定的;
- (ii) 每次试验的结果是不能预先知道的;
- (iii) 在相同的条件下,试验是可以重复进行的.

我们常用字母 E 表示随机试验,今后随机试验简称为试验.

随机试验 E 的所有可能的结果所组成的集合,称为 E 的**样本空间**,常用字母 S 表示.试验 E 的每个结果,即样本空间 S 的每个元素,称为**样本点**.

例 1 下面是一些随机试验及其样本空间的例子.

(1) 试验 E_1 : 抛一枚硬币,观察其出现正面或反面的情况;样本空间为 $S_1 = \{\text{正面}, \text{反面}\}$;

(2) 试验 E_2 : 掷一颗骰子,观察其出现的点数;所有可能结果是 $1, 2, 3, 4, 5,$

6, 样本空间为 $S_2 = \{1, 2, 3, 4, 5, 6\}$;

(3) 试验 E_3 : 记录某电话台一天接到呼叫的次数; 样本空间为 $S_3 = \{0, 1, 2, 3, \dots\}$, 就是全体非负整数组成的集合;

(4) 试验 E_4 : 在一批灯泡中任取一只, 测量其寿命 t (小时); 样本空间为 $S_4 = \{t | t \geq 0\} = [0, +\infty)$;

(5) 试验 E_5 : 在圆周 $x^2 + y^2 = R^2$ 内任取一点 M , 记录 M 点的坐标 (x, y) ; 则样本空间为 $S_5 = \{(x, y) | x^2 + y^2 < R^2\}$.

二、随机事件

在观察随机试验的结果时, 我们不仅关心某个个别结果, 更关心的可能是某一类结果. 比如对于例 1 中的试验 E_4 , 在一批灯泡中任取一只灯泡时, 我们关心的不仅是灯泡的寿命 t 是否等于 1000 小时, 更关心的是灯泡的寿命 t 是否有 1000 小时以上. 也就是关心样本空间 S_4 的子集 $A = \{t | t > 1000\}$, A 就称为试验 E_4 的一个随机事件, 即 A 是“灯泡的寿命 t 为 1000 小时以上”这个事件, 而事件 A 发生, 就是试验结果 $t > 1000$, 即 $t \in A$.

一般, 随机试验 E 的样本空间 S 的子集称为 E 的随机事件, 简称事件. 特别地, 由一个样本点组成的单点集称为基本事件. 随机事件常用字母 A, B, C 等来表示. 在一次试验中, 当且仅当事件 A 中有一个样本点出现时, 称事件 A 发生.

S 是 S 的子集, 故 S 是随机事件, 因为每一次试验, S 中的样本点必然出现, 即事件 S 必然发生, 称 S 为必然事件; 空集 $\emptyset$ 也是 S 的子集, 因而也是随机事件, 因为它不含样本点, 每次试验都不可能发生, 称 $\emptyset$ 为不可能事件.

例 2 考察例 1 中的试验 E_2 : 掷一颗骰子, 观察其出现的点数; 样本空间为 $S_2 = \{1, 2, 3, 4, 5, 6\}$, 基本事件为 $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$, 以下是试验 E_2 的一些事件:

$A = \text{“点数为偶数”} = \{2, 4, 6\}$;

$B = \text{“点数为奇数”} = \{1, 3, 5\}$;

$C = \text{“点数大于 3”} = \{4, 5, 6\}$;

“点数不超过 6” $= \{1, 2, 3, 4, 5, 6\} = S_2$ 为必然事件;

“点数大于 6” $= \emptyset$ 为不可能事件.

三、事件间的关系及运算

既然试验 E 的事件是 E 的样本空间 S 的子集, 那么, S 的子集间的关系及运算, 就是事件间的关系及运算. 下面给出集合间的关系及运算在概率论中的称呼, 并根据“事件发生”的含义, 指出它们在概率论中的含义.

设试验 E 的样本空间为 S , 而 $A, B, A_i (i=1, 2, 3, \dots)$ 是 S 的子集, 也就是试验 E 的事件.

1. 若关系 $A \subset B$ 成立, 则称事件 B 包含事件 A , 根据集合的包含关系, $A \subset B \Leftrightarrow$ 若 $x \in A$, 则 $x \in B$, 再根据“事件发生”的含义, $A \subset B$ 在概率论中的含义是事件 A 发生, 必导致事件 B 发生.

若 $A \subset B$ 且 $B \subset A$, 即 $A = B$, 则称事件 A 与事件 B 相等.

2. 事件 $A \cup B = \{x \mid x \in A \text{ 或 } x \in B\}$ 称为事件 A 与事件 B 的和事件, 事件 $A \cup B$ 发生的含义是 A 发生或者 B 发生, 也就是 A 与 B 中至少有一个发生.

类似地, 称 $\bigcup_{i=1}^n A_i$ 为 n 个事件 $A_1, A_2, \dots, A_n$ 的和事件, 其含义是 n 个事件 $A_1, A_2, \dots, A_n$ 中至少有一个发生; 称 $\bigcup_{i=1}^{\infty} A_i$ 为可列个事件 $A_1, A_2, \dots$ 的和事件, 其含义是 $A_1, A_2, \dots$ 中至少有一个发生.

3. 事件 $A \cap B = \{x \mid x \in A \text{ 且 } x \in B\}$ 称为事件 A 与事件 B 的积事件, 事件 $A \cap B$ 发生的含义是 A 发生并且 B 也发生, 也就是事件 A 与事件 B 同时发生.

今后 $A \cap B$ 经常记作 AB .

类似地, 称 $\bigcap_{i=1}^n A_i$ 为 n 个事件 $A_1, A_2, \dots, A_n$ 的积事件, 其含义是 n 个事件 $A_1, A_2, \dots, A_n$ 同时发生; 称 $\bigcap_{i=1}^{\infty} A_i$ 为可列个事件 $A_1, A_2, \dots$ 的积事件, 其含义是 $A_1, A_2, \dots$ 同时发生.

4. 若 $A \cap B = \emptyset$ (或写成 $AB = \emptyset$), 则称事件 A 与事件 B 互斥或互不相容, 其含义是事件 A 与事件 B 不同时发生.

5. 事件 $B - A = \{x \mid x \in B \text{ 且 } x \notin A\}$ 称为事件 B 与事件 A 的差事件, 事件 $B - A$ 发生的含义是 B 发生且 A 不发生.

6. 事件 $\bar{A} = S - A$ 称为 A 的对立事件, 其含义是事件 A 不发生.

A 与 $\bar{A}$ 满足条件: $A \cup \bar{A} = S, A \cap \bar{A} = \emptyset$. 即 A 与 $\bar{A}$ 至少有一个发生, 但不同时发生.

上述事件的关系与运算, 可以用图 1-1 中的 1, 2, 3, 4, 5, 6 的图形来表示, 其中矩形内部表示样本空间 S , 圆圈内部表示事件 A 或事件 B , 有斜线部分表示相应运算的结果.

事件运算满足以下的运算律 (设 $A, B, C, A_i (i=1, 2, \dots, n)$ 为事件):

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

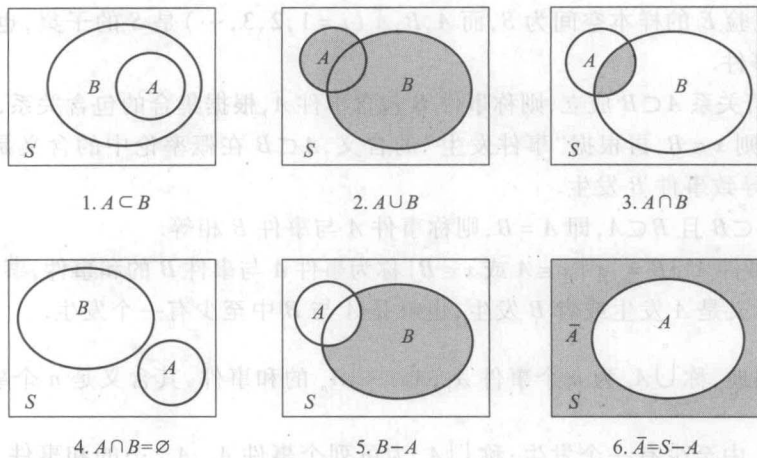


图 1-1

(4) 对偶律(德摩根律): $\overline{A \cup B} = \bar{A} \bar{B}$, $\overline{AB} = \bar{A} \cup \bar{B}$;
 $\overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \bar{A}_1 \bar{A}_2 \cdots \bar{A}_n$, $\overline{A_1 A_2 \cdots A_n} = \bar{A}_1 \cup \bar{A}_2 \cup \cdots \cup \bar{A}_n$;

(5) 其他: $A\bar{B} = A - B = A - AB$, $AB \subset A \subset A \cup B$, $A \cup A = A$, $AA = A$, $A \cup S = S$,
 $AS = A$, $A \cup \emptyset = A$, $A\emptyset = \emptyset$.

例 3 设 A, B, C 是同一个试验的三个事件, 试将下列各题中的事件, 用 A, B, C 的运算表示出来.

- (1) A, B, C 都发生; (2) A, B, C 都不发生;
 (3) A, B, C 中至少有一个发生; (4) A 发生, B, C 都不发生;
 (5) A, B, C 中不多于一个发生; (6) A, B, C 中不多于两个发生;
 (7) A, B, C 中至少有两个发生; (8) A, B, C 中恰有两个发生.

解 (1) ABC ; (2) $\bar{A}\bar{B}\bar{C}$;

(3) $A \cup B \cup C$; (4) $\bar{A}\bar{B}\bar{C}$ 或写成 $A - B - C$;

(5) $\bar{A}\bar{B}\bar{C} \cup \bar{A}\bar{B}C \cup \bar{A}B\bar{C} \cup \bar{A}BC$;

(6) $\overline{ABC}$ (不是都发生) 或写成 $\bar{A} \cup \bar{B} \cup \bar{C}$ (至少有一个不发生);

(7) $AB \cup BC \cup AC$ 或写成 $ABC \cup AB\bar{C} \cup AC\bar{B} \cup BC\bar{A}$;

(8) $AB\bar{C} \cup AC\bar{B} \cup BC\bar{A}$.

第二节 随机事件的概率

在随机试验中, 我们会发现有些事件发生的可能性大一些, 有些事件发生的可能性小一些. 例如, 在掷一枚骰子的试验中, “出现偶数点” 总比 “出现两点” 的
 此为试读, 需要完整PDF请访问: www.ertongbook.com

可能性大一些. 表示事件发生的可能性大小的数量称为事件的概率, 事件 A 的概率记作 $P(A)$.

在历史上, 出现几种概率的定义, 下面分别进行介绍.

一、事件的频率

设在相同条件下进行的 n 次试验中, 事件 A 发生了 n_A 次, n_A 称为事件 A 发生的频数, 而比值 $\frac{n_A}{n}$ 称为事件 A 发生的频率, 记为 $f_n(A)$, 即

$$f_n(A) = \frac{n_A}{n}$$

根据频率的定义, 容易证明, 事件的频率具有以下一些性质:

- (i) $0 \leq f_n(A) \leq 1$;
- (ii) $f_n(S) = 1$ (必然事件的频率等于 1);
- (iii) 若 $A_1, A_2, \dots, A_k$ 是两两互不相容的事件, 即 $A_i A_j = \emptyset, i \neq j, i, j = 1, 2, \dots, k$, 则有

$$f_n(A_1 \cup A_2 \cup \dots \cup A_k) = f_n(A_1) + f_n(A_2) + \dots + f_n(A_k)$$

证 (i), (ii) 很明显, 下面证明性质 (iii). 设在 n 次试验中, 事件 A_1 发生 n_1 次, 事件 A_2 发生 n_2 次, $\dots$, 事件 A_k 发生 n_k 次, 因为 $A_1, A_2, \dots, A_k$ 两两互不相容, $A_1, A_2, \dots, A_k$ 中任何两个事件不可能同时发生. 所以, 事件 $A_1 \cup A_2 \cup \dots \cup A_k$ 发生了 $n_1 + n_2 + \dots + n_k$ 次, 故有

$$\begin{aligned} f_n(A_1 \cup A_2 \cup \dots \cup A_k) &= \frac{n_1 + n_2 + \dots + n_k}{n} \\ &= \frac{n_1}{n} + \frac{n_2}{n} + \dots + \frac{n_k}{n} \\ &= f_n(A_1) + f_n(A_2) + \dots + f_n(A_k) \end{aligned}$$

历史上有许多著名学者做过掷一枚均匀硬币的试验, 观察其出现正面的频率, 下表列出了他们的试验结果.

试验者	试验次数	出现正面的频数	出现正面的频率
德摩根	2048	1061	$\frac{1061}{2048} = 0.5181$
蒲丰	4040	2048	$\frac{2048}{4040} = 0.5069$
费希尔	10000	4979	$\frac{4979}{10000} = 0.4979$
皮尔逊	12000	6019	$\frac{6019}{12000} = 0.5016$
皮尔逊	24000	12012	$\frac{12012}{24000} = 0.5005$

从表中可以看出,出现正面的频率在 0.5 附近摆动,随着试验次数的增大,摆动的幅度越来越小,频率越来越接近于 0.5.

通过各种试验证实,当试验的次数 n 增大时,事件 A 发生的频率,总会稳定在某个常数 p 附近摆动. 这个常数 p 就称为事件 A 的概率,记为 $P(A) = p$. 这就是概率的统计定义. 根据这个定义,上述掷硬币的试验,出现正面的概率等于 0.5.

既然概率是频率的稳定值,由频率的性质,可知概率有以下性质:

- (1) $0 \leq P(A) \leq 1$;
- (2) $P(S) = 1$ (必然事件的概率等于 1);
- (3) 若 $A_1, A_2, \dots, A_k$ 是两两互不相容的事件,则有

$$P(A_1 \cup A_2 \cup \dots \cup A_k) = P(A_1) + P(A_2) + \dots + P(A_k)$$

概率的统计定义中,并没有给出“ n 增大时,事件发生的频率,总会稳定在某个常数 p 附近摆动”的准确含义,因此在理论上是不严密的. 它的实用价值是可以利用频率近似计算概率.

二、古典概型

如果随机试验有以下两个特征:

- (i) 试验的样本空间只含有限个元素,即只有有限个基本事件;
- (ii) 每次试验中,各个基本事件发生的可能性相等,

则这种试验称为**古典概率模型**,简称**古典概型**.

在古典概型中,设试验的样本空间 S 包含有 n 个基本事件,而事件 A 包含其中 k 个基本事件,则定义事件 A 的概率为

$$P(A) = \frac{k}{n} = \frac{A \text{ 包含的基本事件个数}}{S \text{ 包含的基本事件总数}}$$

这样定义的概率,称为**古典概率**.

例 1 将一枚均匀硬币抛掷一次,求出现正面的概率.

解 抛掷一次硬币,可能结果是出现正面或反面,因为硬币均匀,出现正面或反面的可能性相等,故为古典概型. 样本空间为 $S = \{\text{正面}, \text{反面}\}$, S 包含的基本事件数为 2, 设 A 表示“出现正面”,则 A 含 1 个基本事件,故 A 的概率为

$$P(A) = \frac{1}{2} = 0.5$$

例 2 掷两颗均匀的骰子,求两颗骰子点数之和等于 4 的概率.

解 掷两颗骰子,观察其出现的点数,基本事件是两颗骰子点数 i, j 的一个排列 (i, j) , 骰子均匀,出现各种排列的可能性是相同的,故为古典概型. 样本空间为 $S = \{(i, j) \mid i, j = 1, 2, 3, 4, 5, 6\}$, 基本事件总数为 $6 \times 6 = 36$. 设 A 表示“两

骰子点数之和等于4”,则 $A = \{(1,3), (2,2), (3,1)\}$ 包含3个基本事件,故 A 的概率为

$$P(A) = \frac{3}{36} = \frac{1}{12}$$

当样本空间元素较多时,不必把 S 与 A 中的元素一一列出,只要计算出 S 及 A 中所含的基本事件数即可.而在这类计算中,经常会遇到一种最基本的计数原理.比如,设由甲地到乙地有两条路可走,由乙地到丙地有三条路可走,则由甲地经乙地到达丙地,就共有 $2 \times 3 = 6$ 条路可走.把这个例子一般化,就是下面的计数原理.

计数原理:设有 k 个试验,已知做第一个试验有 n_1 个结果,做完第一个试验再做第二个试验时有 n_2 个结果, ..., 做完前 $k-1$ 个试验再做第 k 个试验时,有 n_k 个结果,则依次做这 k 个试验时,共有 $n_1 \times n_2 \times \cdots \times n_k$ 个结果.

例3 设有一批同型号的产品,共有100件,其中有30件为次品.现按以下两种不同抽取方式,先后从这批产品中随机抽取两件产品:(a)有放回抽取,即任意抽取一件产品,查明是正品或次品后放回产品堆中,再从中抽取一件产品;(b)不放回抽取,即任意抽取一件产品,抽后不放回,从剩下的产品中再任意抽取一件产品.试对这两种抽取方式,分别计算

- (1) 两次都抽到次品的概率;
- (2) 第一次抽到次品,第二次抽到正品的概率.

解 试验的基本事件是两次抽取产品为正品或次品的一个排列.易知本题的试验是古典概型.记 $A =$ “两次都取到次品”, $B =$ “第一次抽到次品,第二次抽到正品”.

(a) 有放回抽取

第一次从100件产品中任意抽取一件,有100种抽取的结果,因为抽后放回,故第一次抽取后第二次抽取时,仍有100种抽取结果,由计数原理,先后进行两次抽取,共有 $100 \times 100 = 100^2$ 种结果,即基本事件总数 100^2 ;事件 A 发生就是两次都是从30件次品中抽取一件,因此,事件 A 含有的基本事件数为 $30 \times 30 = 30^2$.根据古典概率的定义,事件 A 的概率为

$$P(A) = \frac{30^2}{100^2} = \frac{9}{100} = 0.09$$

事件 B 发生就是第一次在30件次品中任取一件,第二次在70件正品中任取一件,事件 B 含有 30×70 个基本事件,故 B 的概率为

$$P(B) = \frac{30 \times 70}{100^2} = \frac{21}{100} = 0.21$$

(b) 不放回抽取

第一次从 100 件产品中任意抽取一件,有 100 种抽取的结果,第一次抽取后因为不放回,第二次抽取时只有 99 种抽取结果,由计数原理,两次抽取共有 100×99 种抽取的结果,基本事件总数为 100×99 . 事件 A 发生即第一次在 30 件次品中任取一件,因为不放回,第二次只能在剩下的 29 件中任取一件,所以事件 A 发生有 30×29 个结果, A 含的基本事件数 30×29 ,故 A 的概率为

$$P(A) = \frac{30 \times 29}{100 \times 99} = \frac{29}{330} = 0.088$$

事件 B 发生就是第一次在 30 件次品中任取一件,第二次是在 70 件正品中任取一件,事件 B 发生共有 30×70 种可能结果, B 含的基本事件数为 30×70 ,故 B 的概率为

$$P(B) = \frac{30 \times 70}{100 \times 99} = \frac{7}{33} = 0.212$$

例 4 (抽奖问题) 设某商店进行有奖销售,投放 n 张不同的奖券,其中只有一张有奖,每个顾客可抽一张,求第 k 个顾客得奖的概率($1 \leq k \leq n$).

解 设 $A =$ 第 k 位顾客得奖. 一次试验就是前 k 位顾客从 n 张奖券中各抽取一张奖券,基本事件就是取出的 k 张奖券的一个排列. 因为第一个顾客有 n 个可能的抽取结果,第一个顾客抽取后第二个顾客再抽取时,有 $n-1$ 个可能的抽取结果,……前 $k-1$ 个顾客抽取后,第 k 个顾客抽取时有 $n-(k-1) = n-k+1$ 个可能的抽取结果. 因此,由计数原理,前 k 个顾客从 n 张奖券中各抽取一张奖券的排列总数为 $n(n-1)\cdots(n-k+1)$,这就是基本事件的总数;第 k 个顾客得奖(事件 A),就是前 $k-1$ 个顾客抽到的是无奖的奖券,而第 k 个顾客则抽到有奖的奖券. 即前 $k-1$ 个顾客在 $n-1$ 张无奖的奖券各取一张奖券,类似于前面的计算方法(n 换作 $n-1$, k 换作 $k-1$),共有

$(n-1)(n-2)\cdots[(n-1)-(k-1)+1] = (n-1)(n-2)\cdots(n-k+1)$ 个可能结果. 而第 k 个顾客抽取有奖的奖券,只有一个结果. 因此, A 含有的基本事件总数为 $(n-1)(n-2)\cdots(n-k+1) \cdot 1$. 这样,第 k 个顾客得奖的概率为

$$P(A) = \frac{(n-1)(n-2)\cdots(n-k+1) \cdot 1}{n(n-1)\cdots(n-k+1)} = \frac{1}{n}$$

概率 $P(A)$ 与 $k(1 \leq k \leq n)$ 无关,也就是前 n 个顾客中,每个顾客抽到奖品的概率都是 $\frac{1}{n}$,这说明抽奖方法对每个抽奖者来说都是公平的.

注 如果把 n 张奖券看作 n 个不同元素,把前 k 个顾客从 n 张奖券各抽取的一张奖券,按先后顺序排成一列,就是从 n 个不同元素取 k 个元素所作的一个不重复的排列,而前 k 个顾客从 n 张奖券中各抽取一张奖券的排列总数,就等于从 n 个不同元素取 k 个元素作不重复排列的排列总数,这个排列总数通常记为 A_n^k ,由前面的推导可知,

$$A_n^k = n(n-1)\cdots(n-k+1)$$

当 $k=n$ 时,就是把 n 个不同元素全部进行不重复的 n 元排列,称为 n 元全排列,由 A_n^k 的公式可知, n 元全排列的排列总数为 $A_n^n = n(n-1)\cdots 1 = n!$.

如果从 n 个不同元素取 k 个元素,不考虑其先后次序,组成一组,称为一个组合,从 n 个不同元素取 k 个元素所作的组合总数记为 C_n^k ,如果把其中每一个组合的 k 个元素再作全排列,每个组合可得到 $k!$ 个排列, C_n^k 个组合共可得到 $C_n^k \times k!$ 个排列,这个排列数就是从 n 个不同元素取 k 个元素所作的排列总数 A_n^k ,因此, $C_n^k \times k! = A_n^k$,于是得到 n 个不同元素取 k 个元素组合的组合数 C_n^k 的计算公式:

$$C_n^k = \frac{A_n^k}{k!} = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

用 $(n-k)!$ 乘分子、分母,可得

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

在上式中将 k 换作 $n-k$,得到

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

组合记号 C_n^k 有些书记作 $\binom{n}{k}$.

例 5 设有一批产品,共 100 件,其中有 5 件是次品. 从这批产品中任意取出 3 件产品,求取出的产品中恰有两件是次品的概率.

解 一次试验得到的是 3 件产品的一个组合,100 件产品任意取 3 件的组合总数为 $C_{100}^3 = \frac{100 \times 99 \times 98}{3!} = 100 \times 33 \times 49$; 设 $A =$ “取出的 3 件产品中有两件次品”, A 发生就是从 5 件次品取两件次品,再从 95 件正品中取一件正品所得的组合,由计数原理,组合数为

$$C_5^2 \times C_{95}^1 = \frac{5 \times 4}{2!} \times 95 = 950$$

故所求事件的概率为

$$P(A) = \frac{C_5^2 \times C_{95}^1}{C_{100}^3} = \frac{950}{100 \times 33 \times 49} = \frac{19}{3234}$$

例 6 某接待站在某个星期接待 12 次来访,这 12 次来访都是在星期一和星期三这两天进行的,问是否可以推断接待站的接待时间是有规定的.

解 假设接待时间没有规定,即一星期 7 天都可以来访,则每次来访都有 7 种可能的选择,12 次来访共有 7^{12} 种选择的结果,即基本事件数为 7^{12} . 设 $A =$ “12 次来访都在星期一和星期三那两天进行的”,则每次来访都只有两种选择,12 次来访共有 2^{12} 种可能选择结果,即 A 包含的基本事件数为 2^{12} ,事件 A 发生的概率为

$$P(A) = \frac{2^{12}}{7^{12}} = \left(\frac{2}{7}\right)^{12} < \left(\frac{3}{10}\right)^{12} = \frac{531441}{10^{12}} = 0.000000531441$$

这个概率是一个很小的数,人们长期实践总结出“概率很小的事件,在一次试验中几乎是不发生的”(称为实际推断原理),现在概率很小的事件 A ,竟在一次试验中发生了,因此有理由怀疑假设的正确性,从而推断接待站不是每天都接待来访者,即接待时间是有规定的.

三、几何概型

设试验结果是几何上的一个随机点,它等可能地出现在区域 S 中的每一个点,这种试验模型称为几何概型.这里样本空间 S 可以是直线上的区间、平面区域或空间区域,几何概型的样本空间 S 含有无穷多个样本点,突破了古典概型只含有限个样本点的限制,但仍保留样本点出现为等可能的限制.

设事件 A 是 S 的一个子区域,则定义 A 的概率为

$$P(A) = \frac{|A|}{|S|} = \frac{A \text{ 的度量}}{S \text{ 的度量}}$$

当样本空间为区间、平面区域或空间区域时,相应地 $|A|$ 与 $|S|$ 则表示区间长度、平面区域的面积或空间区域的体积,这样定义的概率,称为几何概率.

例 7 设某公交车站每天早上从 7:00 开始,每隔 15 min 来一趟车,有一乘客在 7:00 到 7:30 之间随机来到该车站.求

(1) 该乘客等车时间不超过 5 min 的概率;

(2) 该乘客等车时间超过 10 min 的概率.

解 设 A = “乘客等车时间不超过 5 min”, B = “乘客等车时间超过 10 min”,乘客到达时间为 7 时 t 分. 样本空间 $S = \{t \mid 0 < t < 30\}$, 因为乘客到达时,第一趟车已开走,只能等第二、三趟车. 要乘客等车时间不超过 5 min, 乘客到达时间 t 只能是 $10 \leq t < 15$ 或 $25 \leq t < 30$, 即事件 $A = [10, 15) \cup [25, 30)$. 同理, $B = (0, 5) \cup [15, 20)$. 故

$|S| = 30$, $|A| = 5 + 5 = 10$, $|B| = 5 + 5 = 10$, 根据几何概率的定义, 有

$$P(A) = P(B) = \frac{10}{30} = \frac{1}{3}$$

例 8(会面问题) 设有甲、乙两人相约在第二天早上 8:00 至 12:00 之间在某地见面, 先到的人等待 1 h 即可离开, 求两人可以见面的概率.

解 设甲、乙到达约会地点的时间分别为 $8+x$ 与 $8+y$ (单位: h), 则 $x, y \in [0, 4]$, 可认为点 (x, y) 等可能地落在如图 1-2 所示的边长等于 4 的正方形(记为 S)

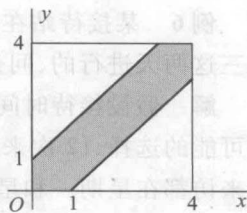


图 1-2