

经济数学(三)

PROBABILITY THEORY
AND MATHEMATICAL STATISTICS

概率论 与数理统计

孙激流 沈大庆\主编

 首都经济贸易大学出版社

经济数学(三)

概率论与数理统计

主 编 孙激流 沈大庆
副主编 孙 阳 刘 强 张 琳
 聂 力 陶桂平 梅超群

首都经济贸易大学出版社
·北 京·

图书在版编目(CIP)数据

概率论与数理统计/孙激流,沈大庆主编.—北京:首都经济贸易大学出版社,2005.10
ISBN 7-5638-1324-1

I.概… II.①孙… ②沈… 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(2005)第 090917 号

概率论与数理统计

孙激流 沈大庆 主编

出版发行 首都经济贸易大学出版社
地 址 北京市朝阳区红庙(邮编 100026)
电 话 (010)65976483 65065761 65071505(传真)
E-mail publish@cueb.edu.cn
经 销 全国新华书店
照 排 首都经济贸易大学出版社激光照排服务部
印 刷 北京市通县永乐印刷厂印刷
开 本 787毫米×980毫米 1/16
字 数 390千字
印 张 20.75
版 次 2005年10月第1版 第1次印刷
印 数 1~3 000
书 号 ISBN 7-5638-1324-1/O·29
定 价 29.00元

图书印装若有质量问题,本社负责调换
版权所有 侵权必究

孙激流 沈大庆\主编

责任编辑

薛晓红

平面设计

首都经济贸易大学出版社照排室

前 言

本书是首都经济贸易大学出版的经济数学系列教材的第三部,是参照国家制定的经济管理类概率论与数理统计教学大纲、研究生入学考试大纲和经济管理专业自学考试大纲的基本要求编写的。当然,我们也根据目前教育改革的新形势、学科发展的现状和学生的学习情况,适当地增补或删减了部分内容。参编本书的教师老、中、青结合,遵循系列教材编写的指导思想——注重系统性、严密性和科学性,注重联系实际和兼顾各层次学生的不同需求,注重与数学软件相结合,力求使教材具有自己的特色。

本书的习题安排同经济数学系列教材的第一部。

本书由孙激流、沈大庆任主编,由孙阳、刘强、张琳、聂力、陶桂平、梅超群任副主编。本书分为十一章,第一章由张琳编写,第二章和第五章由孙激流编写,第三章和第十一章由刘强编写,第四章由聂力编写,第六章和第七章由陶桂平编写,第八章由梅超群编写,第九章和第十章由孙阳编写。

本书既适合于深层次的本科生教学,又兼顾到一般层次的专科生和成人教育的教学。由于编者水平有限,错误在所难免,恳请读者批评指正。

编 者

2005年9月1日于首都经济贸易大学

目 录

第一章	随机事件及其概率	(1)
§ 1.1	随机事件	(1)
§ 1.2	事件的概率	(7)
§ 1.3	古典概型与几何概型	(12)
§ 1.4	条件概率与全概公式	(17)
§ 1.5	独立性	(23)
	典型例题分析	(27)
	习题一	(31)
第二章	随机变量及其分布	(38)
§ 2.1	随机变量的概念	(38)
§ 2.2	离散型随机变量	(39)
§ 2.3	连续型随机变量的分布	(48)
§ 2.4	随机变量的分布函数	(54)
§ 2.5	随机变量函数的分布	(57)
	典型例题分析	(61)
	习题二	(62)
第三章	多维随机变量及其分布	(71)
§ 3.1	二维随机变量及其分布	(71)
§ 3.2	边缘分布	(77)

§ 3.3	条件分布	(82)
§ 3.4	随机变量的独立性	(87)
§ 3.5	两个随机变量的函数的分布	(89)
	典型例题分析	(96)
	习题三	(100)
<hr/>		
第四章	随机变量的数字特征	(112)
§ 4.1	数学期望	(112)
§ 4.2	方差	(123)
§ 4.3	协方差及相关系数	(128)
§ 4.4	矩、协方差矩阵	(132)
	典型例题分析	(135)
	习题四	(138)
<hr/>		
第五章	大数定律与中心极限定理	(146)
§ 5.1	切贝雪夫不等式	(146)
§ 5.2	大数定律	(148)
§ 5.3	中心极限定理	(151)
	典型例题分析	(154)
	习题五	(155)
<hr/>		
第六章	样本及抽样分布	(157)
§ 6.1	随机样本	(157)
§ 6.2	抽样分布	(159)
	典型例题分析	(171)
	习题六	(173)

第七章	参数估计	(178)
§ 7.1	点估计	(178)
§ 7.2	估计量的优劣标准	(187)
§ 7.3	区间估计	(190)
§ 7.4	正态总体均值与方差的区间估计	(193)
§ 7.5	单侧置信区间	(199)
	典型例题分析.....	(202)
	习题七	(205)
第八章	假设检验	(212)
§ 8.1	假设检验的原理	(212)
§ 8.2	正态总体参数的假设检验	(215)
§ 8.3	非参数假设检验	(224)
	典型例题分析.....	(232)
	习题八	(237)
第九章	方差分析	(244)
§ 9.1	单因素试验的方差分析	(244)
§ 9.2	双因素试验的方差分析	(249)
	典型例题分析.....	(255)
	习题九.....	(257)
第十章	回归分析	(262)
§ 10.1	一元线性回归分析模型.....	(262)
§ 10.2	一元回归模型线性假设显著性检验	(267)
§ 10.3	一元线性回归模型的应用	(270)

§ 10.4 化曲线回归为线性回归	(273)
§ 10.5 多元线性回归模型	(275)
典型例题分析	(282)
习题十	(284)

第十一章	SPSS 系统及其在统计中的应用	(288)
-------------	-------------------------------	--------------

§ 11.1 SPSS 简介	(288)
§ 11.2 SPSS 在统计中的应用	(291)

第一章 随机事件及其概率

在自然界和人类社会中,我们所遇到的各种现象按其结果能否准确预言来划分,可以分为两大类:一类是必然现象;另一类是随机现象.

在一定条件下,必然出现某一种结果的现象称为必然现象;在一定条件下可能出现这样的结果,也可能出现那样的结果,而且不能预先断言出现哪种结果的现象称为随机现象.让我们先做两个简单的试验.

试验 I:一个盒子中有十个完全相同的白球,搅匀后从中任意摸取一球;

试验 II:一个盒子中有十个相同的球,但 5 个是白色的,另外 5 个是黑色的,搅匀后从中任意摸取一球.

对于试验 I,在球没有取出之前,我们就能确定取出的必定是白球.而对于试验 II 来说,在球没有取出以前,不能确定取出的球是白色的还是黑色的.试验 I 所对应的现象就是必然现象,而试验 II 所对应的现象就是随机现象.骤然一看,随机现象似乎没有什么规律可言,但是实践告诉我们,如果从盒子中反复多次取球(每次取出一球,记录球的颜色后仍把球放回盒子中并且搅匀),那么总可以观察到这样的事实:当试验的次数 n 相当大时,出现白球的次数 n_1 和出现黑球的次数 n_2 是很接近的,比值 $\frac{n_1}{n}$ (或 $\frac{n_2}{n}$) 会逐渐稳定于 $\frac{1}{2}$. 即“大数次”地重复一个试验,试验的结果会遵循某些规律,这种规律称为统计规律性. 概率论与数理统计就是研究和应用随机现象统计规律性的一个数学分支.

§ 1.1 随机事件

一、随机试验

以下把在对随机现象进行研究时,对客观事物进行的“调查”、“观察”或“实验”

统称为随机试验,简称试验.概率论中所说的试验有下列三个特点:

- (1) 试验可以在相同的条件下重复进行;
- (2) 每次试验结果不止一个,而且试验之前就能明确所有可能出现的结果;
- (3) 每次试验总是恰好出现这些可能结果中的一个,但在一次试验之前却不能肯定每次试验会出现哪一个结果.

下面给出几个随机试验的例子.

试验 E_1 : 抛一枚硬币,观察正面 H ,反面 T 出现的情况.

试验 E_2 : 将一枚硬币抛掷三次,观察正面 H ,反面 T 出现的情况.

试验 E_3 : 将一枚硬币抛掷三次,观察出现正面的次数.

试验 E_4 : 掷一颗骰子,观察出现的点数.

试验 E_5 : 记录某一车站某一时间段内候车的人数.

试验 E_6 : 在一批灯泡中任意抽取一只,测试它的寿命.

这些例子中的每个试验,都有前面所说的三个特点.例如,试验 E_1 有两种可能结果,出现 H 或者出现 T .但在抛掷之前不能确定出现 H 还是出现 T .这个试验可以在相同的条件下重复进行.又如试验 E_6 ,我们知道灯泡的寿命(以小时计) $t \geq 0$,但在试验之前不能确定它的寿命有多长.这一试验也可以在相同条件下重复进行.

二、随机事件

我们把随机试验的结果,称为随机事件,简称事件.如掷一颗骰子出现二点,出现三点,出现偶数点,出现奇数点都是事件.如果一个事件不包含其它事件,则称为基本事件,如出现二点.包含某些基本事件的事件则称为复杂事件或复合事件.如出现偶数点这一事件,它包含出现二点,四点,六点三个基本事件.复杂事件是由基本事件构成的集合.因为随机试验的所有结果是明确的,从而所有基本事件也是明确的,它们的全体称作样本空间,通常用字母 Ω 表示. Ω 中的点,即基本事件,有时也称作样本点,常用 ω 表示.而任何一个事件都是样本空间的子集^①,常用大写的英文字母 $A, B, C \dots$ 来表示.

下面写出试验 $E_k (k = 1, 2, \dots, 6)$ 的样本空间 Ω_k 和一些事件.

^① 严格地说,事件是指 Ω 中的满足某些条件的子集.当 Ω 是由有限个或可列无限个元素组成时,每个子集都可作为一个事件.若 Ω 是由不可列无限个元素组成时,某些子集必须排除在外.幸而这种不可允许的子集在实际应用中几乎不会遇到.今后,每当我们讲到一个事件时,都是假定它是允许考虑的那种子集.

$$\Omega_1 = \{H, T\};$$

$$\Omega_2 = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\};$$

$$\Omega_3 = \{0, 1, 2, 3\};$$

$$\Omega_4 = \{1, 2, 3, 4, 5, 6\};$$

$$\Omega_5 = \{0, 1, 2, \dots\};$$

$$\Omega_6 = \{t | t \geq 0\}.$$

在 E_2 中, 若令事件 $A_1 =$ “第一次出现 H ”, 则

$$A_1 = \{HHH, HHT, HTH, HTT\}.$$

在 E_4 中, 若令事件 $A_2 =$ “出现的点数为偶数”, 则

$$A_2 = \{2, 4, 6\}.$$

在 E_6 中, 若令事件 $A_3 =$ “寿命小于 1000 小时”, 则

$$A_3 = \{t | 0 \leq t < 1000\}.$$

在概率论的一次试验中, 若某个事件包含的一个样本点出现, 则称这一事件发生.

因为 Ω 由所有基本事件组成, 而在任一次试验中, 必然要出现 Ω 中的某一个样本点 ω , 也就是在任一次试验中, Ω 一定会发生, 所以称 Ω 为必然事件. 相应地, 空集 Φ 可以看作是 Ω 的子集, Φ 不包含任何样本点, 因此每次试验, 它都一定不会发生, 所以称它为不可能事件. 虽然必然事件与不可能事件都不是随机事件, 但为了今后讨论方便, 我们也把它们当作特殊的随机事件.

注意: 样本点的选取与试验的目的有关. 如掷两颗骰子一次, 如果目的是观察 6 点出现的次数时, 则 $\Omega = \{0, 1, 2\}$; 如果目的是观察可能出现的点数之和时, 则 $\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

三、事件间的关系及事件的运算

因为要讨论的事件可能由一些较简单的事件组成, 所以有必要先了解事件间的关系及事件可以进行怎样的计算.

事件是样本点的集合, 所以事件间的关系与事件的运算类似于集合论中定义的集合之间的关系和集合的运算. 下面根据“事件发生”的含义, 给出这些关系和运算在概率论中的解释.

1. 事件的包含关系

若 B 中的每一个样本点都包含在 A 中, 则记为 $A \supset B$ 或 $B \subset A$, 并称事件 A

包含事件 B , 这时事件 B 发生必导致事件 A 发生. 如图 1-1 所示.

例如, 在掷骰子试验 E_4 中, 若 $A =$ “点数等于 2”, $B =$ “点数是偶数”, 则 $A \subset B$.

显然对任何事件 A , 必有 $A \subset \Omega$, 且规定 $\Phi \subset A$.

2. 事件的相等关系

若 $A \subset B$ 与 $B \subset A$ 同时成立, 则称 A 与 B 等价或称 A 等于 B , 记为 $A = B$. 等价的两个事件同时发生, 也同时不发生.

3. 事件的和

$A \cup B$ 表示至少属于 A 或 B 中的一个的所有样本点的集合, 即 $A \cup B = \{\omega \mid \omega \in A \text{ 或 } \omega \in B\}$. 在概率论中记为 $A + B$, 并称它为 A 与 B 的和事件, 简称为和. 事实上, $A + B$ 表示在试验中“ A, B 中至少有一个发生”这一事件. 如图 1-2 中阴影部分所示.

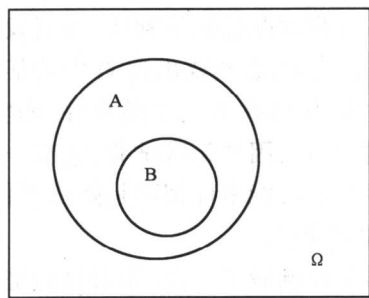


图 1-1

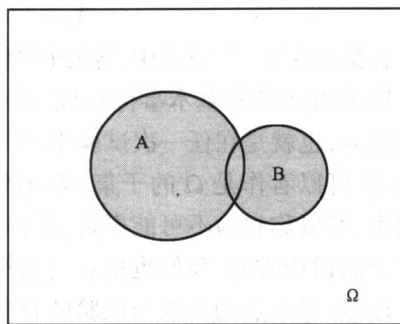


图 1-2

例如, 某产品要经过两道工序加工, 只要有一道工序加工不合格, 产品就不合格, 于是事件“产品不合格”就是事件“第一道工序加工不合格”与“第二道工序加工不合格”的和事件.

类似地, “ $A_i (i = 1, 2, \dots, n)$ 中至少有一个事件发生”这一事件称为 n 个事件 A_1, A_2, \dots, A_n 的和事件, 记为 $\sum_{k=1}^n A_k$. 可列个事件“ $A_i (i = 1, 2, \dots)$ 中至少有一个事件发生”这一事件称为可列个事件 A_1, A_2, \dots 的和事件, 记为 $\sum_{k=1}^{\infty} A_k$.

4. 事件的积

$A \cap B$ 表示所有同时属于 A 及 B 的样本点的集合, 即 $A \cap B = \{\omega \mid \omega \in A \text{ 且}$

$\omega \in B$ }. 在概率中记为 AB , 并称它为 A 与 B 的积事件, 简称为积. 事实上, AB 表示在试验中“ A, B 同时发生”这一事件. 如图 1-3 中阴影部分所示.

在上例中, 事件“产品合格”是事件“第一道工序加工合格”与“第二道工序加工合格”的积事件.

类似地, A_1, A_2, \dots, A_n 同时发生这一事件, 称为 n 个事件 A_1, A_2, \dots, A_n 的积事件, 记为 $\prod_{k=1}^n A_k$. 可列个事件 A_1, A_2, \dots 同时发生这一事件称为可列个事件 A_1, A_2, \dots 的积事件, 记为 $\prod_{k=1}^{\infty} A_k$.

5. 事件的差

用 $A - B$ 表示包含在 A 中而不包含在 B 中的样本点的全体, 即 $A - B = \{\omega \mid \omega \in A \text{ 且 } \omega \notin B\}$, 称它为 A 与 B 的差事件. 事件 $A - B$ 表示 A 发生, 而 B 不发生这一事件. 如图 1-4 中阴影部分所示.

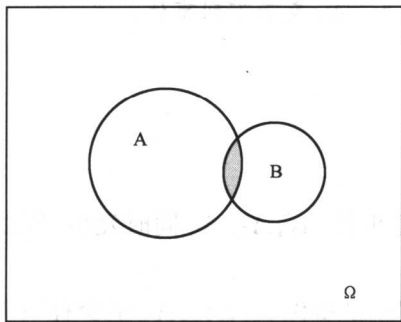


图 1-3

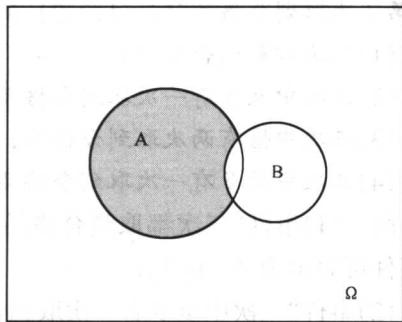


图 1-4

6. 事件的互不相容(或互斥)关系

若 $AB = \Phi$, 则表示事件 A 与事件 B 不可能同时发生, 此时称事件 A 与事件 B 是互不相容的(或互斥的). 如图 1-5 所示.

两个不同的基本事件是互不相容的.

7. 事件的对立(或互逆)关系

若 $A + B = \Omega$ 且 $AB = \Phi$, 则表示事件 A 与事件 B 不可能同时发生, 但又必定有一个发生, 此时称事件 A 与事件 B 是对立的(或互逆的). 显然若事件 A 是事件 B 对立(或互逆)的事件, 则事件 B 也是事件 A 对立(或互逆)的事件. 如图 1-6 所示.

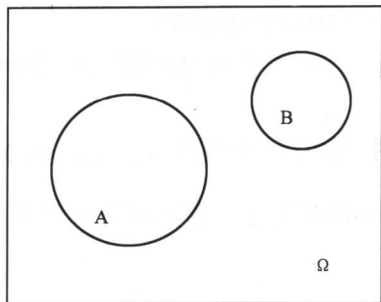


图 1-5

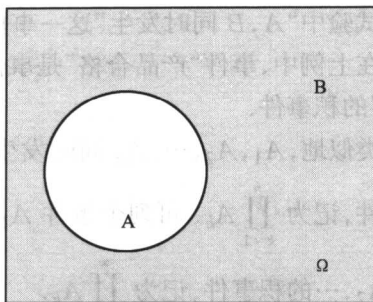


图 1-6

与 A 对立的事件记为 \bar{A} .

显然 $\bar{\bar{A}} = A, \bar{A} = \Omega - A, A - B = A\bar{B}$.

【例 1】 从一批产品中每次取出一个产品进行检验,共取三次,用 A_i 表示事件“第 i 次取到合格品”($i=1,2,3$).试用 A_1, A_2, A_3 表示下列事件:

- (1) 三次都取到合格品;
- (2) 三次中至少有一次取到合格品;
- (3) 三次中恰有两次取到合格品;
- (4) 三次中最多有一次取到合格品.

解 (1) 事件“三次都取到合格品”意味着事件 A_1, A_2, A_3 同时发生,所以这一事件可表示为 $A_1A_2A_3$;

(2) 事件“三次中至少有一次取到合格品”就是事件 A_1, A_2, A_3 至少有一个发生,所以这一事件可表示为 $A_1 + A_2 + A_3$;也可以表示成 $A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3 + A_1A_2\bar{A}_3 + A_1\bar{A}_2A_3 + \bar{A}_1A_2A_3 + A_1A_2A_3$;

(3) 事件“三次中恰有两次取到合格品”意味着两次取到合格品,而另一次取到不合格品,所以这一事件可表示为 $A_1A_2\bar{A}_3 + A_1\bar{A}_2A_3 + \bar{A}_1A_2A_3$;

(4) 事件“三次中最多有一次取到合格品”就是事件 A_1, A_2, A_3 中至少有两个未发生,所以这一事件可表示为 $\bar{A}_1\bar{A}_2 + \bar{A}_2\bar{A}_3 + \bar{A}_1\bar{A}_3$.

既然事件的运算就是集合的运算,那么集合的运算律自然就是事件的运算律.如对于事件 A, B, C 有下述运算律:

交换律: $A + B = B + A, AB = BA$;

结合律: $(A + B) + C = A + (B + C), (AB)C = A(BC)$;

分配律: $(A+B)C = AC + BC, AB + C = (A+C)(B+C)$;

德·摩根律: $\overline{A_1 + A_2} = \overline{A_1} \overline{A_2}, \overline{A_1 A_2} = \overline{A_1} + \overline{A_2}$

对于 n 个事件, 甚至对于可列个事件, 德·摩根律也成立.

【例2】 设 A, B, C 为三个事件, 说明下列各式所表示的事件.

(1) $(A+B)(A+\overline{B})$;

(2) $(A+B)(\overline{A}+B)(A+\overline{B})$.

解 因为 $(A+B)(A+\overline{B}) = A + B\overline{B} = A + \Phi = A$

$$(A+B)(\overline{A}+B)(A+\overline{B}) = (A+B)(A+\overline{B})(\overline{A}+B) = A(\overline{A}+B) = \Phi + AB = AB$$

所以

(1) $(A+B)(A+\overline{B})$ 表示事件 A 发生;

(2) $(A+B)(\overline{A}+B)(A+\overline{B})$ 表示事件 A, B 同时发生.

§ 1.2 事件的概率

在研究随机现象时, 人们除了要了解随机试验的结果, 即事件外, 还需了解事件发生的可能性. 概率就是用来刻画事件在一次试验中发生的可能性大小的数. 人们把概率规定为 0 到 1 之间的实数, 且满足事件发生的可能性越大, 事件的概率也应越大.

在概率论的发展过程中, 人们曾在不同的条件下, 使用不同的方法定义概率, 其中有统计概率, 古典概率, 几何概率等. 不同的概率在各自适合条件下, 刻画了事件发生的可能性, 但这些概率定义的局限性和不统一, 又阻碍了概率论的发展. 于是人们又从各种概率满足的性质中, 概括出三条公理, 给出了概率的公理化定义, 推动了概率论的发展.

一、统计概率

首先给出事件发生的频率的定义.

定义 1.1 如果在相同条件下重复进行了 n 次试验, 事件 A 发生了 m 次, 则比值 $\frac{m}{n}$ 称为事件 A 在这 n 次试验中发生的频率, 记作 $f_n(A)$.

由这个定义, 易知频率具有以下性质: