

概率论与数理统计

范大茵 陈永华 编

浙江大学出版社

021

概率论与数理统计

范大茵 陈永华编

浙江大学出版社

内容提要

本书为浙江大学新编的工程数学规范化教材《概率论与数理统计》部分。全书共分十章。前四章为概率论。主要内容有概率论的基本概念、一元与多元随机变量及其分布、数字特征与极限定理。后六章为数理统计部分。主要内容有数理统计的基本概念、参数估计、假设检验、方差分析、回归分析、正交试验设计法。

本书叙述通俗易懂,注重应用,并有较多的例子介绍概率论与数理统计的思想和方法。

本书可作为高等院校工科、理科(非数学专业)概率论与数理统计课程的教材,同时可作为各类专业人员学习概率论与数理统计的参考读物。

概率论与数理统计

范大哲 陈永华 编

责任编辑 贾吉柱

浙江大学出版社出版

(杭州玉古路 20 号 邮政编码 310027)

浙江大学出版社电脑排版中心排版

德清第二印刷厂印刷

浙江省新华书店发行

*

850×1168 32 开 11.75 印张 316 千字

1996 年 10 月第 1 版 1998 年 5 月第 3 次印刷

印数: 6061—10000

ISBN 7-308-01620-X/O · 189 定价: 12.00 元

编写说明

工程数学由多门数学课程组成,它的涉及面和应用性很广,几年来我们曾先后出版过多种工程数学教材,为了适应形势,及时总结经验,不断扩大教学成果,使我校工程数学教学质量能保持持续的稳定与提高,我们重新组织人力,对《线性代数》、《概率论与数理统计》、《复变函数与拉普拉斯变换》、《常微分方程》、《工程技术中的偏微分方程》、《数值计算方法》等六门课程的教材全部作了改编,以适应我校教学规范化的要求。

最近出版的这一套工程数学教学规范化系列教材,具有如下特点:

1. **适应性** 提高和加宽了知识的深度与广度,能较广泛地适用于多种专业、不同层次的要求。

2. **少而精** 虽然新教材充实了不少新的内容,但整个篇幅与学时数都没有增加。

3. **实用性** 学以致用是编写这套教材的一个原则,因此新教材增加了很多实用性内容,例如,在《常微分方程》中增加了建模的实例;在《复变函数与拉普拉斯变换》中增加了用复变函数方法解决工程实际问题的实例;在《工程技术中的偏微分方程》中增加了差分法及有限元法等内容,以便扩大计算机的应用范围;在《概率论与数理统计》中增加了统计部分比重,等等。

4. **便于教学** 新教材融入了我系教师长年积累的经验与资料,用更为直观,更易为学生接受的方式处理疑难内容,起到深入浅出的效果,值得一提的是,《线性代数》、《复变函数与拉普拉斯变换》中还选编了大量的思考题,以帮助同学们复习。

我们希望这套工程数学系列教材能成为深受广大师生欢迎的教材。

浙江大学应用数学系
1994年3月

前 言

概率论与数理统计是研究随机现象数量规律性的一门学科。它是数学中与现实世界联系最密切,应用最广泛的学科之一。随着科学技术的进步,过去用决定性的数学形式描述的一些现象,为了得到更精确的描述,就需要使用概率论与数理统计的方法。

概率论与数理统计和其它的数学分支有着密切的联系,同时又有自己的特点,它与其它的数学分支一样有严格的数学形式,又有它独特的“概率思想”与“统计思想”。随着电子计算机技术的发展,概率论与数理统计的方法已在工农医林等自然科学领域,以及社会、经济的各个部分都得到了卓有成效的应用。数理统计学与其它学科相结合形成了许多边缘性学科,如统计物理学,地质数学,数量经济学,医学统计,生物统计学,数量遗传学等等。这些都充分说明了数理统计学的重要性。现有的概率论与数理统计教材,由于学时的限制,在数理统计部分主要讲述了参数估计和假设检验,没有介绍试验设计方法的内容,对回归分析的内容也介绍得比较少。这些实用性较强的基本内容来不及为学生所掌握,不能不说是一个缺陷。为改变这种状况,我们在不增加篇幅与学时的前提下,增加数理统计部分实用性内容的比重,同时对概率论部分的内容进行提炼,这就是编写本书的出发点。

本书共分十章。概率论部分(第一章至第四章),主要内容为概率论的基本概念,一元与多元随机变量及其分布,数字特征,极限定理。数理统计部分(第五章至第十章)介绍了数理统计中的基本概念,参数估计,假设检验,方差分析,一元及多元线性回归分析,逐步回归分析,正交试验设计等内容。为了帮助读者加深理解基本内容,提高学以致用用的能力,各章均配有习题,并在书末附有答案。

为便于读者理解和运用,本书在选材和叙述上力求将概念写得

清晰易懂,详细介绍概率论与数理统计的思想和方法,注重应用的前提和步骤的说明。书中一部分内容能直接应用于解决实际问题,另一部分内容为读者今后进一步学习有关课程或在实际应用方面奠定一定的基础。

本书可作为各类高等院校工科、理科(非数学专业)概率论与数理统计课程的教材,同时可供各类专业技术人员进行生产和科学研究时参考。

阅读本书的内容,基本上假定读者具有微积分学及线性代数的基础知识。标有“*”号的小段或用小字编排的定理证明等内容,初学时可以跳过,跳过它们并不损害全书的连贯性。本书的后三章具有相对独立性,通过适当的取舍可以适应不同学时数的教学需要。

本书概率论部分(第一、二、三、四章)由范大茵编写,数理统计部分(第五、六、七、八、九、十章)由陈永华编写。华东师范大学茆诗松教授对全书作了认真、负责、细致的审阅,提出了许多宝贵的意见,对本书的最后形成起了重要的作用,在此特向茆诗松教授致以诚挚的感谢。同时,浙江大学概率论与数理统计教研室盛骤教授及许多老师对本书提出了宝贵的意见,对此一并表示衷心的感谢。书中个别例题是编者在解决企业实际课题中遇到的问题,全书也引用了国内外有关书籍中的一些例题和习题,恕不一一指明出处,在此一并向有关人员致谢。

书中的不足之处,尚请读者不吝指正。

编者

1994年12月

目 录

第一章 概率论的基本概念	1
§ 1.1 随机试验	2
§ 1.2 样本空间、随机事件	2
§ 1.3 频率与概率	9
§ 1.4 等可能概型(古典概型)	13
§ 1.5 条件概率	20
§ 1.6 独立性	30
习题 1	34
第二章 随机变量及其分布	38
§ 2.1 随机变量	38
§ 2.2 离散型随机变量	40
§ 2.3 随机变量的分布函数	47
§ 2.4 连续型随机变量	49
§ 2.5 随机变量的函数的分布	62
习题 2	66
第三章 多维随机变量及其分布	70
§ 3.1 二维随机变量	70
§ 3.2 边缘分布	78
§ 3.3 条件分布	85
§ 3.4 随机变量的独立性	89
§ 3.5 随机变量的函数的分布	95
习题 3	100
第四章 随机变量的数字特征、极限定理	105
§ 4.1 数学期望	105
§ 4.2 方差	115
§ 4.3 几种重要分布的数学期望与方差	119
§ 4.4 协方差和相关系数	124

§ 4.5	矩、协方差矩阵	130
§ 4.6	大数定律	131
§ 4.7	中心极限定理	133
习题 4		137
第五章	数理统计的基本概念	142
§ 5.1	总体和样本	142
§ 5.2	概率论和矩阵代数的基础知识	145
§ 5.3	几个常用的分布和抽样分布	149
习题 5		160
第六章	参数估计	162
§ 6.1	参数的点估计	162
§ 6.2	估计量的评选标准	169
§ 6.3	参数的区间估计	174
习题 6		184
第七章	假设检验	187
§ 7.1	假设检验的基本概念	187
§ 7.2	正态总体下参数的假设检验	192
§ 7.3	非参数假设检验	201
习题 7		210
第八章	方差分析	215
§ 8.1	单因素试验的方差分析	216
§ 8.2	双因素试验的方差分析	225
习题 8		240
第九章	回归分析	245
§ 9.1	一元线性回归	245
§ 9.2	多元线性回归	255
§ 9.3	一元与多元非线性回归的例子	266
§ 9.4	逐步回归分析	270
习题 9		276
第十章	正交试验设计法	283
§ 10.1	正交试验设计的基本方法	283

§ 10.2 水平数不同的试验,多指标试验	289
§ 10.3 考虑交互作用的正交试验设计法	294
§ 10.4 正交试验的方差分析	298
习题 10	302
参考文献	307
附录 常用数理统计表	308
附表一 标准正态分布表	308
附表二 χ^2 分布表	309
附表三 t 分布表	311
附表四 F 分布表	312
附表五 符号检验表	324
附表六 秩和检验表	325
附表七 W 检验法的系数表 $a_i(w)$	326
附表八 W 检验法统计量 W 的 p 分位数 Z_p	330
附表九 常用正交表	331
附表十 泊松分布表	344
附表十一 几种常用的概率分布	346
习题答案	350

第一章 概率论的基本概念

概率论和数理统计是研究随机现象的统计规律性的一门学科。人们在实践中经常会遇到各种随机现象，例如丢一枚硬币，可能是正面朝上也可能是反面朝上，事先无法断言；又如观察某机床加工出来的零件，可能是正品也可能是次品，在观察之前无法肯定哪个结果会出现；又如观察某地区一月份的最高温度及最低温度，在观察之前也无法断言最高温度及最低温度一定是多少；又如观察某车站某一时间区间内的候车人数，可能 0 个，可能 1 个，…，事先无法肯定；又如观察某一工厂生产的灯泡的寿命，可能大于 1000 小时，也可能小于 1000 小时，事先无法确定其寿命一定是多少等等这一类现象，尽管我们在每次试验之前无法断言将得到哪一种结果，但是如果进行大量的重复的观察，我们会发现其出现的结果还是有一定的规律可循的。如丢钱币的问题，如钱币确实均匀的话，那么当投掷次数很多时，将会发现大约有一半的次数出现正面。又如测量某物体的长度，如测量多次的话，一般来说测量值会在某一数值附近较为集中，而离该数值较远的数据往往较少。我们将这种规律称为统计规律性。

随着科学技术的不断发展，设计及测量精度的不断提高，随机现象受到人们越来越多的重视，概率论和数理统计的理论与方法得到了越来越广泛的应用。概率论和数理统计的应用几乎遍及所有科学技术领域，工农业生产和国民经济各个部门。例如气象预报、水文、地震预报、元件和系统可靠性评估、产品抽样验收方案的制订、寻求最佳生产条件的试验设计、产品寿命预测等等。概率论与数理统计与其它数学分支也有着密切的联系，概率论与数理统计的方法正向许多基础学科、工程学科渗透。概率论与其它学科相结合发展了许多边缘学科，如生物统计、统计物理、数学地质等。概率论与数理统计也是可靠性理论、信息论、控制论等重要学科的理论基础。

§ 1.1 随机试验

所谓随机试验是指具有以下三个特点的试验：

(1) 试验可以在相同的条件下大量重复进行.

(2) 每次试验的可能结果不止一个,且在试验之前已知试验的所有可能结果.

(3) 在每次试验之前无法断言哪一个结果会出现,但若进行大量重复试验的话,其可能结果出现又有一定的统计规律性.

下面列举一些随机试验的例子.

E_1 : 丢一颗骰子^①,观察所得点数.

E_2 : 从一副扑克牌(52张)中任选13张牌,观察得牌情况.

E_3 : 向平面上某目标射击,观察弹落点的位置.

E_4 : 观察某电话交换台,在某时间区间内接到的呼叫次数.

E_5 : 从一批灯泡中任取一只,观察其寿命.

E_6 : 丢两颗骰子,观察所得点数.

大家也可以自己列举一些随机试验的例子.本书以后提到的试验都指随机试验.

显然,对随机现象所进行的试验或观察就是随机试验.

§ 1.2 样本空间、随机事件

一、样本空间

对于随机试验,虽然在试验之前我们无法断言哪一个结果会出现,但是对于所有可能出现的试验结果我们是预先知道的,我们就将试验 E 所有可能结果组成的集合称为试验 E 的样本空间,记为 S .样

^① 骰子为正六面体,其各面分别标有数字1,2,3,4,5,6.

本空间的每个元素,即试验的每个可能结果,称为样本点.

若记 § 1.1 中随机试验 E_K 的样本空间为 S_K , ($K = 1, 2, 3, 4, 5, 6$), 即么容易得到:

$$S_1 = \{1, 2, 3, 4, 5, 6\}.$$

$S_2 =$ 为 52 张牌中选 13 张牌的各种组合的全体, 共有 $\binom{52}{13}$ 个元素.

$$S_3 = \{(x, y) \mid -\infty < x < +\infty, -\infty < y < +\infty\}.$$

$$S_4 = \{0, 1, 2, 3, \dots\}.$$

$$S_5 = \{t \mid 0 \leq t < +\infty\}.$$

$$S_6 = \{(x, y) \mid x = 1, 2, \dots, 6, y = 1, 2, \dots, 6\}.$$

二、事件

有了样本空间的概念就可以定义事件. 先看如下例子.

例 1 将一颗骰子连掷两次, 依次记录所得点数, 则所有可能出现的结果是:

$$\begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6). \end{array}$$

把这 36 个可能结果作为样本点, 其全体构成了样本空间, 每做一次试验这 36 个样本点中必有一个出现且仅有一个出现. 在许多场合, 我们常常对 S 的某些子集感兴趣, 并将这些子集称之为事件. 如

A: 两次投掷所得点数之和为 8.

B: 两次投掷所得点数相等.

显然事件 A 发生等价于下列样本点之一出现: $(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)$; 记为

$$A = \{(2,6), (3,5), (4,4), (5,3), (6,2)\},$$

A 是样本空间 S 的子集.

易知事件 B 的发生等价于下列样本点之一出现: $(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)$, 记为:

$$B = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\},$$

B 是样本空间的子集.

上例中的 A, B 我们称之为事件.

一般, 我们称随机试验 E 的样本空间的子集为 E 的**随机事件**, 简称**事件**. 在每次试验中, 当且仅当这一子集中的某一个样本点出现时就称这一事件发生.

特别, 仅由一个样本点组成的单点集称为**基本事件**.

样本空间 S 包含所有的样本点, 它是 S 自身的子集, 在每次试验中它总是发生的, 称为**必然事件**. 空集 \emptyset 不包含任何样本点, 它也是 S 的子集, 它在每次试验中总不发生, 称为**不可能事件**.

三、事件的关系和运算

设试验 E 的样本空间是 S , 而 $A, B, A_k, (k = 1, 2, 3, \dots)$ 都是 S 的子集.

1° 若 $A \subset B$, 则称**事件 B 包含事件 A** , 亦称事件 A 是事件 B 的**子事件**, 亦即事件 A 的发生必导致事件 B 的发生.

2° 若 $A \subset B$ 且 $B \subset A$, 则称事件 A 与事件 B **相等**, 记为 $A = B$.

3° 事件 $A \cup B = \{x | x \in A \text{ 或 } x \in B\}$, 称为事件 A 与事件 B 的**和事件**, 当且仅当事件 A 或事件 B 至少有一个发生时, 和事件 $A \cup B$ 发生.

类似地, 对于 n 个事件 A_1, A_2, \dots, A_n , 定义 $\bigcup_{i=1}^n A_i$ 为这 n 个事件的**和事件**, 即当且仅当 A_1, A_2, \dots, A_n 中至少有一个发生时, 和事件发生.

对于事件列 $A_1, A_2, \dots, A_n, \dots$, 可类似定义这一列事件的**和事件**

$$\bigcup_{n=1}^{+\infty} A_n.$$

4° 事件 $A \cap B = \{x | x \in A \text{ 且 } x \in B\}$, 称为 A 事件与 B 事件的积事件, 即当且仅当 A 事件, B 事件都发生时, 积事件 $A \cap B$ 发生. 积事件 $A \cap B$ 也可简记为 AB .

类似地, 可以定义有限个事件 A_1, A_2, \dots, A_n 的积事件 $\bigcap_{i=1}^n A_i$, 以及定义事件列 $A_1, A_2, \dots, A_n, \dots$ 的积事件 $\bigcap_{n=1}^{+\infty} A_n$.

5° $A - B = \{x | x \in A \text{ 且 } x \notin B\}$, 即当且仅当 A 发生而 B 不发生时事件 $A - B$ 发生, 称 $A - B$ 为 A 事件与 B 事件的差事件.

6° 若 $A \cap B = \emptyset$, 则称事件 A 与事件 B 互不相容, 或称为互斥的, 即若 A, B 两事件互不相容, 则在每次试验中事件 A 与事件 B 决不会同时发生.

7° 若 $A \cup B = S$ 且 $A \cap B = \emptyset$, 则称事件 A 与事件 B 互为逆事件. 亦称 A, B 互为对立事件, 记为 $A = \bar{B}$ 或 $B = \bar{A}$. 若 A, B 互为对立事件, 则每做一次试验, 事件 A 或事件 B 中总有一个且仅有一个事件要发生.

用图 1.1 ~ 图 1.7 可较直观地表示事件之间的关系及运算.

例如图 1-3 中, 正方形表示样本空间 S , 大圆表示事件 A , 小圆表示事件 B , 则两圆的公共部分, 即图 1-3 的阴影部分表示积事件 $A \cap B$.

在进行事件运算时经常要用到下面的定律:

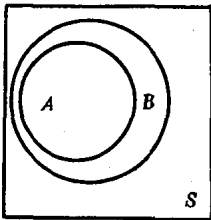


图 1-1

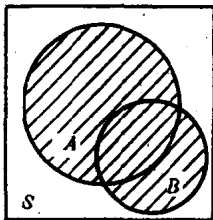


图 1-2

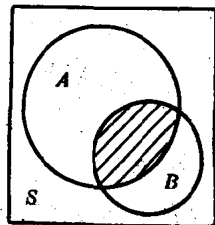


图 1-3

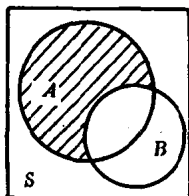


图 1-4

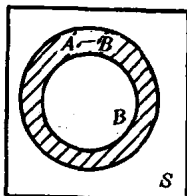


图 1-5

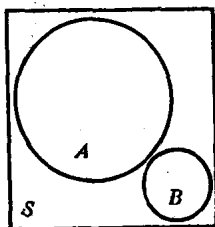


图 1-6

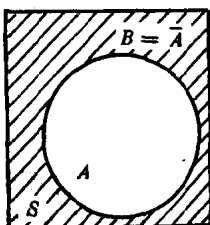


图 1-7

设 A, B, C 是事件, 则有

交换律: $A \cup B = B \cup A, A \cap B = B \cap A;$

结合律: $A \cup (B \cap C) = (A \cup B) \cap C,$

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

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

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

德·摩根律: $\overline{A \cup B} = \bar{A} \cap \bar{B},$

$\overline{A \cap B} = \bar{A} \cup \bar{B}.$

例 2 在某人群中随机地调查一人, 设“被调查者为男性”用事件 A 表示, “被调查者身高大于 1.65 米”用事件 B 表示, “被调查者体重大于 60 公斤”用事件 C 表示, “被调查者具有大专以上学历”用事件 D 表示. 则

(1) “被调查者或是男性或身高大于 1.65 米” = $A \cup B.$

(2)“被调查者身高大于 1.65 米且体重大于 60 公斤” = $B \cap C$.

(3)“被调查者身高小于或等于 1.65 米” = \bar{B} .

(4)“被调查者为男性但体重小于等于 60 公斤” = $A\bar{C}$.

(5)“被调查者是男性或身高大于 1.65 米两条中至少满足一条,且具有大专以上学历” = $(A \cup B)D$.

(6)“被调查者身高小于或等于 1.65 米,或体重小于等于 60 公斤” = $\bar{B} \cup \bar{C}$.

(7)“被调查者是男性但不具有大专以上学历” = $A - D = A\bar{D}$.

(8)“被调查者或是男性或身高大于 1.65 米或体重大于 60 公斤三条中至少满足一条,且具有大专以上学历” = $(A \cup B \cup C)D$.

(9)“被调查者为女性或具有大专以上学历两条中至少满足一条,且身高大于 1.65 米或体重大于 60 公斤两条中至少满足一条” = $(\bar{A} \cup D)(B \cup C)$.

例 3 一颗骰子投掷两次,依次记录所得点数,设 A 为“所得点数之和为 8”的事件, B 为“所得数对中至少有一个为 2 点”的事件, C 为“所得数对中两数相等”的事件, D 为“两数之和小于或等于 3”的事件. 则

$$A = \{(2,6), (3,5), (4,4), (5,3), (6,2)\};$$

$$B = \{(2,1), (2,2), (2,3), (2,4), (2,5), (2,6), \\ (1,2), (3,2), (4,2), (5,2), (6,2)\};$$

$$C = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\};$$

$$D = \{(1,1), (1,2), (2,1)\}.$$

且 $C \cup D = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6), (1,2), (2,1)\} =$ “所得两数相等或两数之和小于等于 3”.

$A \cap B = \{(2,6), (6,2)\} =$ “所得两数之和为 8 且两数中至少有一个为 2”.

$B \cap C = \{(2,2)\} =$ “所得数对中至少有一个为 2 且两数相等.”

$A - B = \{(3,5), (4,4), (5,3)\} =$ “所得数对两数之和为 8 但两数中没有一个是 2”.