

● 王 清 王琪延 郝成义 等编
刘利兰 郭 英
于秀林 审

应用数理统计学

● 中国统计出版社

应用数理统计学

YING YONG SHU LI TONG JI XUE

王 清 王琪延 郝成义 刘利兰 郭 英 编 著

※

中国统计出版社出版
顺义县太平印刷厂印刷

※

787×1092毫米 32开本 12.625印张 26万字

1991年2月第1版 1991年2月北京第1次印刷

印数 1—1500

ISBN 7—5037—0539—6/0.9

定价：9.50元

前 言

在经济领域中，迫切需要应用定量分析的技术方法解决大量的理论和实践问题，本书正是为了适应这方面的需要而编写的。

本书介绍了数理统计的基本方法及其在经济管理实践中的应用，在编写中，既注意保持了该学科的系统性和完整性，又突出了经济管理的需要。一方面为读者提供了必要的基础理论，并对其主要理论依据加以扼要的介绍，但未作繁琐的数学推导。另一方面，着重从经济应用角度，介绍了具有普遍性的若干应用数理统计方法，也介绍了一些近代该学科研究的新成就。书中列举了比较丰富的经济实例，以便读者对各种方法的应用更为直观了解。

全书共十三章，参加各章编写的是：第一、十二章王琪延，第二、三、四章王清，第五、十三章刘利兰，第六章张宗乔，第七、十一章郝成义，第八、九章郭英，第十章方燕，全书由王清、王琪延、郝成义同志负责总纂。

于秀林同志审阅了全部手稿，并提出许多宝贵的意见，在此，表示衷心的感谢。

本书既可作为财经院校的教材，也可供从事经济研究和管理工作的人员自学参考。

由于编者水平有限，书中会有许多不妥之处，恳请广大读者不吝批评指正。

作者

1990年12月

目 录

第一章 绪论

第二章 随机事件和概率

第一节 随机事件及其计算…………… (5)

第二节 频率的稳定性与概率的统计定义… (10)

第三节 古典概率…………… (16)

第四节 条件概率…………… (23)

第五节 全概率公式与贝叶斯公式…………… (27)

第三章 随机变量及其概率分布

第一节 随机变量…………… (33)

第二节 随机变量的概率分布…………… (36)

第三节 基本随机变量的分布…………… (43)

第四章 随机变量的数字特征

第一节 随机变量的数学期望…………… (67)

第二节 随机变量的方差…………… (74)

第三节 主要分布的数学期望和方差…………… (79)

第五章 极限定理

第一节 大数定律…………… (87)

第二节 中心极限定理…………… (91)

第六章 抽样分布与抽样调查

第一节 随机样本与统计量…………… (95)

第二节 抽样分布…………… (97)

第三节 抽样调查的意义与应用…………… (104)

第四节 总体和样本的关系…………… (107)

第五节	抽样误差与推断·····	(112)
第六节	各种抽样组织方式的误差和抽样单位 数目的确定·····	(129)
第七章 参数估计		
第一节	参数估计的基本问题·····	(151)
第二节	确定估计量的方法·····	(154)
第三节	评价估计量优劣的标准·····	(162)
第四节	平均数的区间估计·····	(170)
第五节	比例和方差的区间估计·····	(191)
第八章 参数假设检验		
第一节	参数假设检验的概念和基本问题··	(209)
第二节	总体平均数的假设检验·····	(231)
第三节	总体比例的假设检验·····	(241)
第四节	总体方差的假设检验·····	(247)
第九章 方差分析		
第一节	方差分析的基本原理·····	(258)
第二节	单因素方差分析·····	(262)
第三节	双因素无交互作用的方差分析····	(274)
第四节	双因素有交互作用的方差分析····	(282)
第十章 回归与相关分析		
第一节	回归分析与相关分析的基本概念··	(295)
第二节	一元线性回归的分析·····	(302)
第三节	回归方程的评价·····	(308)
第四节	相关分析·····	(311)
第五节	多元线性回归与相关分析简介····	(325)
第六节	非线性回归分析简介·····	(337)

第十一章	非参数方法简介	
第一节	非参数统计的基本问题·····	(341)
第二节	关于单个样本的非参数方法·····	(345)
第三节	关于两个独立样本的非参数方法···	(364)
第四节	关于两个相关样本的非参数方法···	(376)
第五节	相关程度的度量及显著性检验·····	(390)
第十二章	统计质量控制	
第一节	产品质量与统计质量控制·····	(403)
第二节	产品质量统计控制图——变量控制图·····	(408)
第三节	产品质量统计控制图——品质控制图·····	(420)
第四节	工序能力指数·····	(425)
第五节	产品质量影响因素的分析方法·····	(432)
第六节	产品质量的统计控制与经济效益···	(440)
第十三章	统计预测与决策	
第一节	统计预测的基本问题·····	(445)
第二节	因果回归分析法·····	(447)
第三节	季节变动预测法·····	(479)
第四节	统计决策·····	(494)
第五节	几种常用的统计决策·····	(497)

第一章 绪论

随着社会经济的发展，信息的数量以及人们对信息的需要在不断地迅速扩大，为了对信息进行系统和有效的利用，在各个领域的科学研究中正在愈来愈广泛地使用各种各样的定量分析方法，数量统计学就是进行定量分析的主要方法之一。

数理统计学的应用范围极其广泛，它的理论和方法不仅愈来愈深入地渗透到自然学科的研究中，而且愈益普遍地应用到社会经济领域的研究之中，特别是在经济管理领域中，如生产经营决策、市场预测、产品质量控制、经济活动影响因素分析等方面正在逐步引入数理统计方法。

一、数理统计学的研究对象

数理统计学是研究大量随机现象的一门科学。在现实生活中，存在着这一种现象，在相同条件下，它可能发生，也可能不发生，这种现象称之为随机现象。随机现象不确定性产生的原因是什么呢？我们知道任何现象都是受多因素影响的，在诸多因素中，有的是确定性因素，有的是偶然性因素。例如产品质量，它受工人的技术水平、设备状况、原材料性质、管理水平等因素的影响，这些因素是确定性因素；同时，它也受温度，工人工作过程中的精神状况等因素的影响，这此因素是偶然性因素。确定性因素和偶然性因素综合作用的结果，表现出来就是产品质量的不稳定性——有的是

合格品，有的是不合格品。这也就是不确定性。又如，在经营管理决策中，决策的正确与否，也受很多因素的影响，象决策者对该行业情况的熟悉程度、决策者知识水平、社会提供的信息量等因素是确定性因素；而决策者决策时的情绪、国家政治经济形势的变动、国家经济政策的变动等都属于偶然性因素。这些因素综合作用，决定着决策的成败。

不确定现象常带有偶然性，但是在这些不确定的背后隐藏着必然性，即规律性。例如，一个质量均匀的硬币，抛掷后正面或反面朝上的可能性是一样的。如果只抛一次，可能出现正面，也可能出现反面，当抛掷次数很多时，正面和反面朝上的次数就各接近于一半，即出现正面或反面向上的机会各接近于 $1/2$ 。这就是不确定现象所呈现出来的一种规律性。

可见，对于不确定的现象通过一次观察，它的发生与否是带有偶然性的，但是通过大量观察，就会透过这种偶然性发现其隐藏在背后的必然性。也就是说，统计学不是以个别事例来研究问题，而是对这些客观事物中不确定的问题，通过大量的观察，寻找其规律性。

二、应用数理统计学及其研究内容

数理统计学有理论数理统计学和应用数理统计学之分，前者侧重于研究一般的原理和方法，从数学角度去探讨；后者主要根据某一领域的特点，探讨适合在该领域中理论数理统计的方法如何加以应用问题。因此，应用数理统计学不以数学推论为重点，而是着重阐明应用时所必须满足的数学假定和应用方法。但是，为了便于系统介绍，本书对于一些简单的数学原理和预备知识也给予了一定的阐述。

本书分为三部分内容：①预备知识——随机事件与概率；②用于描述现象的统计方法——随机现象的概率分布与数学特征、极限定理；③用于分析、推断、估计、预测和决策的统计方法——估计理论、假设检验、方差分析、非参数检验、回归分析、相关分析、抽样调查、质量控制、预测和决策理论等等。在上述三部分内容中，第二部分称为描述统计方法，第三部分称为推论统计方法。在统计实践中，描述统计是推论统计的基础，推论统计是描述统计的目的。应用数理统计学主要是根据样本资料，找出样本的数学特征及其分布，对现象的总体进行分析、判断，以便预测和决策。

应用数理统计的方法来源于理论数理统计学，但是又不同于理论数理统计学，应用数理统计学不仅要阐述数理统计的一般方法，还要说明数理统计方法在各研究领域中的应用的条件、应用的过程及注意事项，并且，也要探讨应用过程中出现的一些问题。例如在方法的应用条件不能满足时，如何根据现有情况，进行方法的修正，使之适合客观情况，以达到应用的目的。

综上所述，数理统计学是人们认识自然、认识社会、进行数量分析的一个重要手段，正在各个领域中显示出它的威力。特别是电子计算机的应用，大大地提高了数理统计方法处理问题的速度，降低了使用这种方法的难度，使过去难以应用的统计方法在今天得以应用。

第二章 随机事件和概率

第一节 随机事件及其计算

一、必然现象与随机现象

在社会经济活动中，人们观察到的现象大体可归纳为两种类型。一类是可以事前预言的，即在准确地重复某些条件下，它的结果总是肯定的；或是根据它过去的状态，在相同条件下完全可以预言将来的发展。我们把这类现象称之为确定性现象或必然现象。另一类现象是事前不可预言的，即在同一条件下重复进行试验，每次结果未必一样；或是知道它过去的状况，在相同条件下，未来的发展事前却不能完全肯定。这类现象我们称之为偶然性现象或随机现象。随机现象在社会经济领域中是广泛存在、不胜枚举的。例如，每位顾客到商店购买服装所需的型号是不同的，售货员无法在顾客购货之前准确地预料到各种型号服装的销售数量；再如，新出生的婴儿可能是男孩，也可能是女孩，即表现为一种随机现象；在工业生产中，从同一生产线上用同一种工艺生产出来的产品的质量是有差异的；在农业生产中，相同条件下生长的农产品的产量仍然会出现一定程度的变动。这些都是由于客观上存在着大量的系统的随机因素的结果，正因为如此，在社会经济发展过程中，我们不仅要研究确定性现象，也要研究不确定性现象，把握各种随机现象的变动规律，以

促进国家经济建设和生产的发展。

随机现象的具体内容虽然是千变万化的，但随机现象本身并非杂乱无章。通过大量的观察和试验，可以揭示出这种现象的规律性。例如，通过观察服装商店以往销售服装的统计资料，可以了解顾客的消费特点和变化规律，逐步掌握顾客需要不同型号服装的比例，并据以安排商店的进货计划。再如，根据各个国家各个时期的人口统计资料可以了解到，新生婴儿中男孩和女孩的比例大约总是保持在 1 : 1。这种在相同条件下进行大量观察时，偶然现象所呈现出的某种规律性，称为随机现象的统计规律性。数理统计学正是以现实世界中随机现象的规律性作为自己的研究对象。

二、随机试验与随机事件

对随机现象的研究必然要联系到对客观事物进行“观察”或“试验”。为了叙述方便，我们把对经济现象进行观察或进行一次科学试验统称为一个试验。如果这个试验在相同条件下可以重复进行，而且每次试验的结果事前不可预言，我们就称它为一个随机试验。

在随机试验中，我们感兴趣的主要是试验的结果。例如，对某工厂生产的某种产品进行质量检验，这里每检验一件产品就相当于进行一次试验，我们最关心的是出现合格品，还是出现不合格品，这是两个可能出现的结果。试验的每一个可能结果一般称为随机事件，简称为事件，常用字母 $A, B, C, \dots$ 表示。在一次试验中必然发生的事件称为必然事件，用字母 U 表示，在每一次试验中都不会发生的事件称为不可能事件，用字母 V 表示。为了便于讨论本书把必然事件和不可能事件都算作随机事件。

三、事件间的关系与运算

在实际问题中，往往需要考虑同一个试验之下的多个事件，而这些事件又是相互关联的。例如，从一批产品中连抽两件进行检验，可能出现：“有一件合格品、一件次品”；“都是合格品”；“都是次品”；“第一件是次品、第二件是合格品”；“至少有一件合格品”；“至少有一件次品”等，以上每一种可能结果都可以看作是一个事件。显然，这些事件彼此之间又有一定的联系。下面我们引进事件之间的一些重要关系和运算，这将有利于今后对事件和它的概率的叙述与研究。

（一）事件的包含与相等

设有事件 A 与 B ，如果事件 A 发生必然导致事件 B 发生，则称事件 B 包含事件 A 。记作

$$A \subset B$$

例如，从一批产品中连取两件，令 A 表示“有一件合格品”， B 表示“至少有一件合格品”，显然有 $A \subset B$ 。

如果事件 A 包含事件 B ，同时事件 B 也包含事件 A ，则称事件 A 与 B 相等。记作

$$A = B$$

在上例中，若以 A 表示“至少抽到一件合格品”， B 表示“抽到了合格品”。则有 $A = B$ 存在。

（二）事件的和

用 U 表示“两事件 A 与 B 至少有一件发生”这一事件叫做事件 A 与 B 的和。记作

$$U = A + B \text{ 或 } U = A \cup B$$

$A + B$ 实质上包括三种情况：

(1) A 发生 B 不发生.

(2) B 发生 A 不发生.

(3) A 、 B 同时发生.

例如, 在对某产品进行检验时规定, 必须是体积和重量同时合格才能称为合格产品. 若令 A 表示“体积不合格”; B 表示“重量不合格”; C 表示“产品不合格”. 则产品不合格就意味着体积不合格, 或重量不合格, 或两者都不合格, 则事件 C 便是事件 A 与 B 的和的事件. 即 $C=A+B$.

(三) 事件的积

“两事件 A 与 B 同时发生”这一事件叫做事件 A 与 B 的积. 记作

$$AB \text{ 或 } A \cap B$$

上列中, 若以 A 表示“体积合格”; B 表示“重量合格”, 则 AB 表示“产品合格”.

注意: 事件的和与事件的积都可以推广到有限多个事件. 请读者自行完成, 这里不再赘述.

(四) 事件的差

“事件 A 发生而事件 B 不发生”这一事件叫做事件 A 与 B 之差. 记作

$$A-B$$

例如, 对产品进行一、二、三级划分, 事件 A 表示“抽出一件产品为一级或二级产品”; 事件 B 表示“抽出一件产品为二级或三级产品”, 则 $A-B$ 表示“抽出一件产品为一级产品”.

(五) 两事件的互斥 (或不相容)

如果用 C 表示事件 A 与事件 B 不可能同时发生, 即记为:

$$AB=V$$

那么，称 A 与 B 两事件是互斥的（或不相容的）。

在上例中，若以 A 表示“抽到一件产品为一级品”， B 表示“抽到一件产品为二级品”，显然事件 A 与 B 不可能同时发生，则 A 与 B 就是互斥事件。

（六）对立事件（或逆事件）

设 A 是一个发生事件，那么“ A 不发生”这一事件就叫做 A 的对立事件。记作 $\bar{A}$

例如， A 表示“抽出一件产品为合格品”， $\bar{A}$ 表示“抽出一件产品为不合格品”，则事件 A 与 $\bar{A}$ 就是对立事件。

在一次试验中，显然有

$$(1) A\bar{A}=V \quad (A \text{ 和 } \bar{A} \text{ 不能同时发生})$$

$$(2) A+\bar{A}=U \quad (A \text{ 和 } \bar{A} \text{ 必有一个发生})$$

$$(3) (\bar{\bar{A}})=A \quad (A \text{ 是 } \bar{A} \text{ 的对立事件})$$

（七）事件的运算规律

事件的运算遵循以下规律

$$(1) A+B=B+A \quad (\text{加法交换律})$$

$$(2) A+(B+C)=(A+B)+C \quad (\text{加法结合律})$$

$$(3) A+\bar{A}=U$$

$$(4) A \cdot B=B \cdot A \quad (\text{乘法交换律})$$

$$(5) (AB)C=A(BC) \quad (\text{乘法结合律})$$

$$(6) A \cdot \bar{A}=V$$

$$(7) A(B+C)=AB+AC \quad (\text{乘法分配律})$$

$$(8) \overline{A+B}=\bar{A} \cdot \bar{B}$$

$$(9) \overline{A \cdot B}=\bar{A}+\bar{B}$$

（八）事件的几何图形表示

对于事件之间的关系及其运算，如果用几何图形表示，则较直观且易于理解。

假定以平面上的某一矩形表示样本空间，矩形内的每一点表示样本点，则事件的运算可通过平面上的几何图形表示。如果用两个小圆形表示事件 A 和 B ， U 表示样本空间（如图2—1所示），阴影部分表示事件 A 与 B 的各种关系及运算。

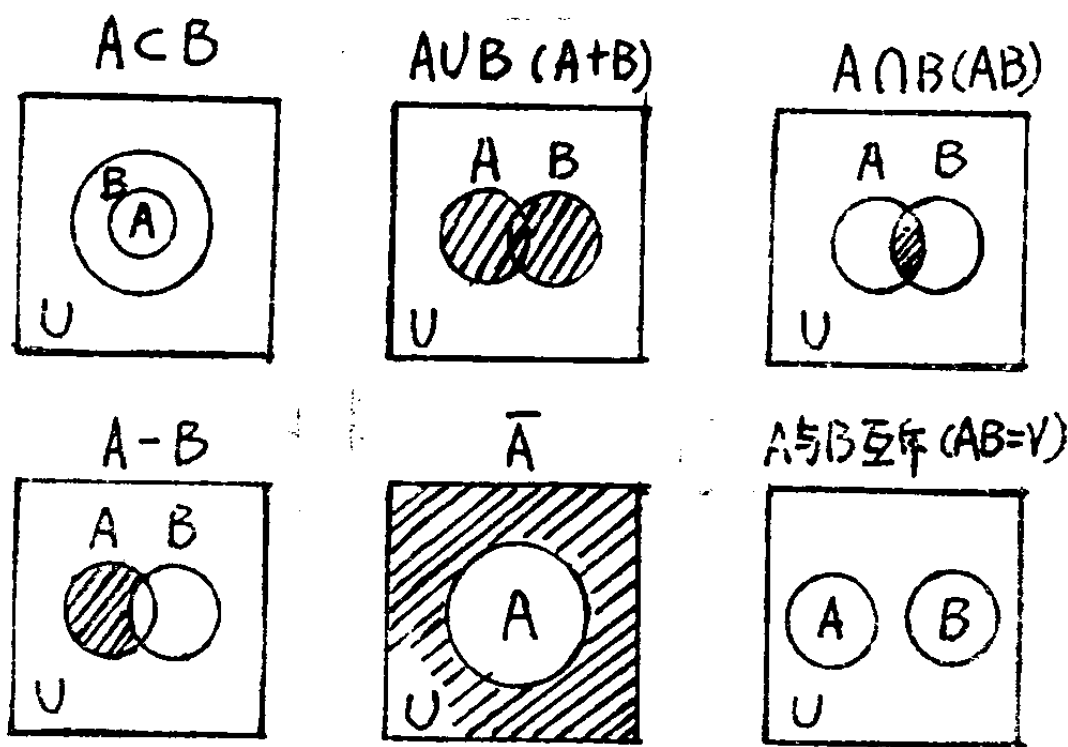


图 2—1 事件的各种关系及运算

第二节 频率的稳定性与概率的统计定义

一、频率的稳定性

对于随机事件，我们知道它在条件具备时可能发生，也可能不发生。但仅仅知道这一点是不够的，更重要的是要研

究各种事件出现的可能性大小，揭示出这些事件内在的统计规律。为了能定量地研究这种规律性，首先给出频率的概念。

在相同条件下进行 N 次重复试验，设随机事件 A 在这 N 次试验中出现了 M 次，称 M 为事件 A 出现的频数。则比值 $\frac{M}{N}$ 为随机事件 A 在 N 次试验中出现的频率。记作

$$F_N(A) = \frac{M}{N} \tag{2.1}$$

人们经过长期的实践发现，虽然个别随机事件在某次试验或观察中，可能出现，也可能不出现，但在大量试验中它却呈现出明显的规律性——频率的稳定性。

例如，在掷一枚硬币时，即可能出现正面，也可能出现反面，预先作出准确的判断是不可能的。但是，假如硬币质地均匀，直观上讲，出现正面与出现反面的机会应该相等。即在大量试验中，出现正面与反面的频率都应接近于50%。为了验证这点，历史上曾有不少人做过这个试验，其结果如表2.1所示。

表 2 . 1 历史上几个数学家作过的掷硬币试验

实验者	投 掷 次 数 N	出现正面的次数 M	频率 f_N
德 · 摩根	2048	1061	0.5180
蒲 丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005