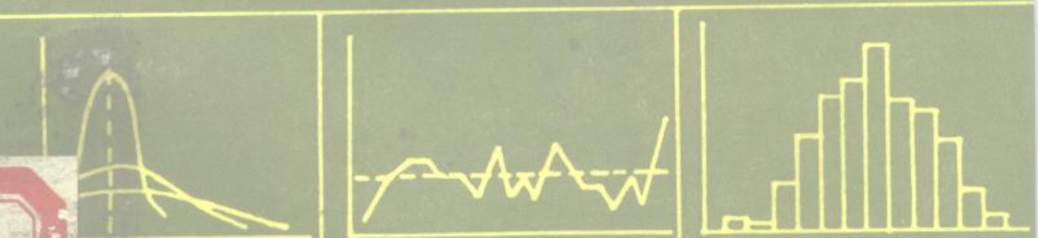
建筑施工工程师技术丛书

应用数理统计学

沈荣芳 编著



中国建筑工业出版社

建筑施工工程师技术丛书

应用数理统计学

沈荣芳 编著

中国建筑工业出版社

本书以数理统计学为内容，着重于在建筑业中的应用。全书共十章，分成三个部分。第一部分主要介绍数理统计学的基础知识。第二部分是本书的核心，结合大样本介绍了统计推断的基本原理，同时介绍了小样本情况下的统计推断理论和方法。第三部分主要介绍数理统计学的应用。分别讨论了方差分析、回归和相关分析、抽样检验和质量控制等方法及其应用。

本书末附有数理统计名词索引，供读者查阅。

本书可供工程技术人员，高等工科院校各专业，特别是土建类专业、管理工程类专业本科生、研究生作为教学参考书。

* * *

责任编辑：林婉华

建筑施工工程师技术丛书
应用数理统计学
沈荣芳 编著
*
中国建筑工业出版社出版（北京西郊百万庄）
新华书店北京发行所发行 各地新华书店经售
中国建筑工业出版社印刷厂印刷（北京阜外南礼士路）
*
开本：850×1168毫米 1/32 印张 7½ 字数：202 千字
1987年4月第一版 1987年4月第一次印刷
印数：1—6,270册 定价：1.75元
统一书号：15040·5165

出 版 说 明

当前，新技术革命浪潮冲击着一切经济部门，建筑业也不例外。许多现代化的科学技术方法和管理手段正逐步地应用在建筑业中，取得了越来越大的经济效益。党的十一届三中全会以来，我国的建筑事业得到了蓬勃发展，各种现代化的建筑如雨后春笋，逐年增多。常年奔波在施工生产第一线的建筑施工工程师们，担负着繁重而复杂的施工任务。他们渴望学习新技术，提高业务水平；渴望更新自己的知识以适应现代化的要求。从科学技术的发展和四化建设的需要考虑，对在职科技人员进行继续教育的重要性和迫切性也日益突出。为此，我们组织出版了这套丛书，希望这套书能对他们有所裨益，并在工程实践中广泛应用新技术，建造出更多优良的工程，取得更佳的经济效益。

城乡建设环境保护部曾委托同济大学、重庆建筑工程学院、哈尔滨建筑工程学院从一九八一年开始举办建筑施工工程师进修班。这套丛书就是根据这些班的教学内容，结合当前施工技术的发展，将施工新技术、新材料、新结构的课题适当加多，以同济大学的老师为主组织编写的。可作为工程师进修班的教材，也可作为建筑施工工程师和有关人员自学参考。计划列题十余种，三年左右出齐。成书时尽量做到内容完整系统，文字叙述深入浅出，以便于现场施工工程师和技术员自学。当然，书中的内容选材是否适当，能否满足读者的要求，还希望广大读者提出意见，以便我们改进。谢谢！

城乡建设环境保护部干部局
中国建筑工业出版社
1986年6月

前　言

长期以来，工程技术和管理的问题中，常常把问题看作是确定性的，应用得比较多的是确定性的数学工具。近几十年来，随着科学和实验技术的发展，在工程技术和管理科学中，往往需要考虑问题的不确定方面，随机性数学模型越来越多，所以概率论和数理统计学的应用日益广泛。可以预期，随着数理统计学知识的普及，其应用范围还将更加扩大，更加普遍。

本书内容以数理统计学为核心，包括了部分作为预备知识的基础概率，以及几种基本的应用方法及其原理。

在翟立林教授倡导下，同济大学建筑管理工程专业，于1960年首次开设了数理统计学这门课程，经过多次教学逐步形成了本书。从最近的发展看，在实际中应用较多的一些方法，它们的原理本书基本上都已提到。

本书中所运用的数学理论是以一般工科院校本科生所学的高等数学为基础的，不需其他更高深的数学内容。书中在说明方法的同时，还适当地介绍了方法的数学原理及其理论推导。如果读者越过这些数学推导，直接学习方法的应用，那么，对问题的理解也不会有太多的影响。部分牵涉数学理论较多、较深的定理，则不加证明地只叙述它们的结论。

本书出版前虽经多次使用修改，这次付印前又作了必要的补充与删减，但是由于编者水平有限，缺点及错误在所难免，衷心希望读者不吝指正。在本书出版过程中，得到了赵志缙、王永安、杨道生等同志的协助，在此谨致谢忱。

沈荣芳

1986年元月

目　　录

绪　论

在自然界和社会中广泛地存在着一类所谓随机现象。例如，我们在进行某种测量时，由于种种偶然因素的影响（如测量仪器受大气的影响，观测者生理上或心理上的变化等），不可避免地会产生测量的误差。又如，生产工人每天的劳动生产率，由于种种原因，每天都会有一定的变化。此外，象单位工程量的材料消耗数，它也不是固定不变的。它们都围绕着某个确定的数值（例如，定额）作随机摆动。

实践证明，对同类随机现象作大量的观察或试验后，通常总能发现一种确定的规律性，这是随机现象所特有的一种规律性。例如，在射击中，当射击次数不多时，靶上命中点似乎是杂乱无章的，没有规律；但是，当射击次数增多时，命中点的分布就开始呈现出一定的规律性；射击次数越多，规律性就越明显。因此可以知道，尽管个别随机现象是没有规律的，但是，大量性质相同的随机现象却有其确定的统计规律性。概率论和数理统计学就是一门研究随机现象数量规律的科学。

概率论和数理统计学中的一切定律（例如，大数定律），都是宇宙中客观存在的大量随机现象统计规律性的反映。必须指出，随机现象同确定现象之间，偶然性同必然性之间并没有不可逾越的鸿沟。恩格斯曾经指出："那些被断定为必然的东西，是由种种纯粹的偶然所构成的，而被认为是偶然的东西，则是一种有必然性隐藏在里面的形式"❶。事实上，宇宙中几乎没有一件

❶ 恩格斯：《费尔巴哈与德国古典哲学的总结》，人民出版社，1959年，第34页。

实际现象不带有某种程度的随机成分，可以说，任何确定现象都不可避免地伴有随机偏差。在许多实际问题中，为了处理问题的方便，在所要求的精确度的范围内，常常略去那些造成随机偏差的次要因素，而只考虑起主要和基本作用的因素。这是研究自然现象时，常采用的做法。也就是先找出那些对现象起决定作用的、最主要和最基本的因素，然后用数学工具对这些起决定作用的因素进行分析，并从中得出结论。例如，在考虑枪弹的轨道时，略去大气对它的影响；又如，在测定挖土机挖土的定额时，忽略土中含水量细小变化的影响。但是，这种做法是以降低精确度为代价的。在某些精确度要求较高的问题中，就必须把这些随机因素也考虑在内。因此，除了需要那些研究确定性现象的数学工具之外，还需要采用概率论和数理统计学的方法。

本书的主要内容——数理统计学——同概率论一样，是研究随机现象统计规律性的科学，但在使用的方法上有其特点。数理统计学主要是以实际观测的资料为基础，来研究随机现象的某些数量规律的。

显然，如果我们对带有随机性的现象作大量的观测，那么，就能看出这些随机现象所呈现的规律性。但是，在客观上往往只允许我们对随机现象作次数不多的观测。这是我们面临的问题。实际上，只要能充分利用观测得到的数据，以及局部同整体的内在联系，并用它们来进行分析和推断，就仍然能认识到这些规律性。数理统计学的中心任务就是要研究如何合理地搜集资料，并利用这些观测得到的资料来对随机现象的某些数量规律进行估计、分析和推断。

由于随机现象存在的普遍性，随着我国社会主义建设事业的发展，对科学技术的要求越来越高，因此，上述从局部观测去推断整体的方法越来越显得重要，数理统计学的应用也日益广泛。它同各种具体的研究对象相结合，能解决许多实际中存在的问题。例如，工程设计中安全度的统计分析；工业生产中产品质量控制和抽样检查；以及科学实验中如何从试验数据中找出客观存

在的规律等等，数理统计学都得到了广泛的应用。

近二、三十年来，数理统计学在我国的应用，得到了很大的发展。在许许多多的领域里，诸如，经济、管理；工业、农业、商业，医学，等等，都有广泛的应用。我们相信，随着我国建设事业的发展，数理统计学的应用必将有更大的发展。

本书将侧重于应用，特别是结合建筑业中的问题，介绍数理统计学的方法及应用。

第一章　概率的基本概念及其运算

从第一章到第四章是数理统计的基础部分，许多数理统计的概念都是在它们的基础上发展起来的。在这一章里，首先讨论有关概率的基本概念和它的运算。

1-1　基　本　概　念

本节将介绍概率论和数理统计学中的三个基本概念。

1-1-1　随机实验

对于某些事物的试验来说，在同一组条件实现之下，就必然得到同样的试验结果。例如，水在标准大气压下，加热到摄氏100度的时候，必然会沸腾；又如，将一枚硬币向上抛后，它必然会受地心的引力而下落；等等。但是，也确实存在着另外一类试验，对于这种试验，在同一组条件实现之下，不一定得到相同的试验结果；例如，旋转一枚硬币，等它停下来后，就既可能正面向上（简称为出正），也可能反面向上（简称为出反）；每一个可能的试验结果都有一定的出现机会，或者说有一定的可能性。对于后一类试验，就称为**随机实验**。在这本书中，将要讨论的试验一般都是指随机实验而言的。

1-1-2　事件

随机实验的结果叫做**事件**，例如，掷一个硬币，“出正面”就是一件事件。

事件可以是定性的，例如，上面提到的掷一个硬币出正或出反，出生婴孩的性别，等等；事件也可以是定量的，例如，投一骰子出“2”，出生的婴孩体重在3～4kg之间，等等。

事件可以是单一性质的，也可以是多重性质的。前者，例如

只考虑出生婴儿的性别，譬如说，“生男孩”这一事件；后者，例如要同时考虑婴儿的性别和体重，譬如说，“生女孩、体重在3～4kg之间”这一事件，等等。

事件又可以是简单的或者是复合的。前者，指不能再行分拆的事件；后者，指由简单事件复合而成的事件。例如，当我们观测某服务机构前的排队人数时，“有4个人在排队”是一简单事件；而“有不多于4个人在排队”则为一复合事件，因为这一事件是由“有0、1、2、3、4个人排队”这样五个简单事件的复合。

由此可见，更确切地说，事件是随机实验的所有可能结果里面的一个集合。

事件是概率论和数理统计学中最基础的概念。这个概念是同一个随机实验以及它的所有可能的结果联系在一起的。只有明确了随机实验和它的全部可能结果，才能确定事件可能性的大小（以后我们说确定事件的概率）。

1-1-3 随机变量

表示随机实验结果的一个数量叫做**随机变量**。随机变量取什么值，在试验之前是无法确切地知道的，它决定于试验的结果。例如 在一定的生产技术条件下生产出来的砖，试压它的强度所得到的值，就是一个随机变量；在某一小时内，电话总机所接到的呼唤次数也是一个随机变量；等等。如果试验的结果不是直接由测量或计数而得，那么，随机变量所取的值可以按照适当的惯例而确定。例如，在产品的质量检查中，记废品为1、成品为0。这样，随机变量取什么值，就可根据检查结果是废品还是成品来决定了。

不言而喻，随机变量取什么值，是有一定的机会或可能性的，而每次试验，事实上，就是对随机变量的一次观测。

1-2 概率的定义

如上所述，事件有它不确定的一面，即可能发生也可能不发

生。每一个事件的出现，都有一定的可能性。概率就是对这种可能性的数量描绘。它通常有三种定义法，即：概率的**古典定义**、概率的**统计定义**和概率的**公理化定义**。前两者，将在下面的两小节内分别予以介绍；后者，由于牵涉到的数学内容较多，在本书中将不去提它了。

1-2-1 概率的古典定义

假如相对于一个试验，有并且也只有 n 个可能出现的结果，同时，每个情况都是等可能的，其中恰恰有 m 个结果具有性质 A，那么，就称 A 的概率是 $\left(\frac{m}{n}\right)$，并记为

$$P\{A\}=\left(\frac{m}{n}\right) \tag{1-1}$$

上面，就是概率的**古典定义**。

【例 1-1】 由均匀物质构成的正六面体骰子，六面分别写上 1、2、3、4、5、6六个数字，试求掷一骰子出现偶数的概率。

【解】 关于这个试验，有也只有六个可能发生的情况，即：出 1 或出 2 或出3、4、5、6；对于均匀的正六面体的骰子而言，出现每一面的可能性可以认为是相同的；在六个可能发生的情况中，具有偶数性质的情况，恰有三个，即：出 2 、 4 和 6 。因此可知，出现偶数的概率是 $\left(\frac{3}{6}\right)=\frac{1}{2}$。

1-2-2 概率的统计定义

在上一小节里介绍了根据古典定义确定某一偶然事件出现的概率的方法。但是，从古典定义的内容来看，它的适用范围是有一定的限度的。因为它只适用于试验结果能归结为有限多个等可能的简单事件的情况，在考虑自然科学和工程技术性质的问题时，这种情况虽然不少，但是确实有非常多的试验，它们的结果不能归结为有限多个等可能的简单事件的情况❶，因而也就不能根据古典定义来计算概率。对于这种不能用古典定义来确定概

❶ 例如，至少在目前，我们还不能用判断事件等可能性所依据的对称想法，决定出生婴儿性别的概率。

率的事件，可用统计定义来找到这种事件的概率的近似值。

我们知道，在试验之前一般不能预期试验结果会出现什么事件，不经研究也很难说出出现某种事件的可能性有多大。但是，在多次重复的试验的观察中，人们还是可以发现其中存在着的某种规律性。下面，先用例子来说明这种规律性。

【例 1-2】 现在要检查一大批在相同条件下生产出来的产品，产品质量标准中规定，当产品长度介于13.60cm到13.90cm之间时为合格，否则就列为次品。现在从中分别抽取5件、10件、60件、150件、600件、900件、1200件、1800件来检查，检查结果列于表1-1和图1-1中。

合格品频率表 表 1-1

检查件数	5	10	60	150	600	900	1200	1800
合格品数	5	7	53	131	548	820	1091	1631
合格品频率①	1	0.7	0.883	0.873	0.913	0.911	0.909	0.906

①合格品频率$=\dfrac{\text{合格品数}}{\text{检查件数}}$。

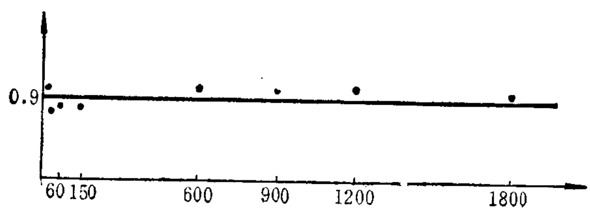


图 1-1 合格品频率图

从例1-2可以看出，虽然在抽出的产品中，次品数是随机的，然而随着试验次数（抽查的总数）的增多，频率越来越清楚地显现出稳定性来，也就是说频率越来越接近于某个固定的数值 p，例中为 $p=0.9$，只是偶而产生较大的偏差。

另一方面，对于概率的古典定义所能运用的那些情况中，有许多例子，证明上述频率的波动的确是围绕着事件的概率 p 在摆动的，例1-3就是其中的一个。

【例 1-3】 蒲丰(Бюффон)和K·皮尔逊(Pearson)两人曾就掷硬币进行了大量的试验，结果列于1-2表中。

掷硬币试验的频率表　　　　表 1-2

实验者	掷硬币次数	掷出国徽次数	频率
蒲丰	4048	2048	0.5059
K·皮尔逊	12000	6019	0.5016
K·皮尔逊	24000	12012	0.5005

例1-3这一类例子说明了一条普遍的规律，就是在同一组条件下进行试验，只要试验的次数足够多,那么,频率通常是同概率接近的；并且随着试验次数的增多，频率将更接近于概率；这也就是所谓的**频率稳定性**。由于存在这种规律性，因而有理由认为：在一般场合下，也存在一个“常数”，而频率就在它的左右摆动。例如，在例1-2中，合格品的频率就围绕着常数0.9在波动。

将上述常数（频率围绕它波动的数）作为所研究的随机事件的概率的近似值，就是概率的**统计定义**的主要意思。例如，在例1-2中，根据概率的统计定义，可以认为产品合格的概率是0.9。

最后必须说明，当要通过大量的实验观测来探求事件的概率时，必须保证每次试验在相同的条件下进行。因为条件不同时，事件发生的可能性也将发生变化。

1-3 概率的属性

我们把出现或发生具有性质A的情况称为事件A。对于事件A的概率有下列一些性质。

（1）事件A的概率是非负数，即

$$P\{A\}\geqslant 0 \qquad (1\text{-}2)$$

如果相对于某一试验，事件A不可能发生，那么，就称A为**不可能事件**，同时，记它的概率为零，即

$$P\{A\} = 0$$

（2）事件A的概率不大于1，即

$$P\{A\} \leqslant 1 \tag{1-3}$$

如果对于一个试验，事件A必然发生，那么，就称A为**必然事件**，同时，记它的概率为

$$P\{A\} = 1$$

（3）记不具有性质A的事件为$\bar{A}$，并称$\bar{A}$为A的**对立事件**，对于A和$\bar{A}$恒有

$$P\{A\} + P\{\bar{A}\} = 1 \tag{1-4}$$

显然，A同$\bar{A}$互为对立事件，$\bar{A}$当且仅当A不发生时才发生。

上列属性，都能很容易地从概率的古典定义而导出，对概率的统计定义而言，也是正确的。

1-4 概率的加法和乘法

在这一节里将讨论概率的基本运算。为了简单起见，讨论将在概率古典意义的基础上进行。如果利用概率的统计定义，那么，也可以参照下面的说明进行运算。

下面先谈一下关于事件的记号。

事件和 对于任意事件A_1、A_2、…、A_n，我们用总和

$$A_1 + A_2 + \cdots + A_n = \sum_{i=1}^{n} A_i$$

表示至少发生事件A_1、A_2、…、A_n之一，读作“事件A_1或A_2…或A_n”。这个总和叫做事件A_1、A_2、…、A_n的事件和。许多概率统计书中，上式中的+号写成$\cup$号，其意义相同。

事件和本身也是一个事件，它可看作为一个具有n个事件性质中任一个或一个以上的性质的事件。