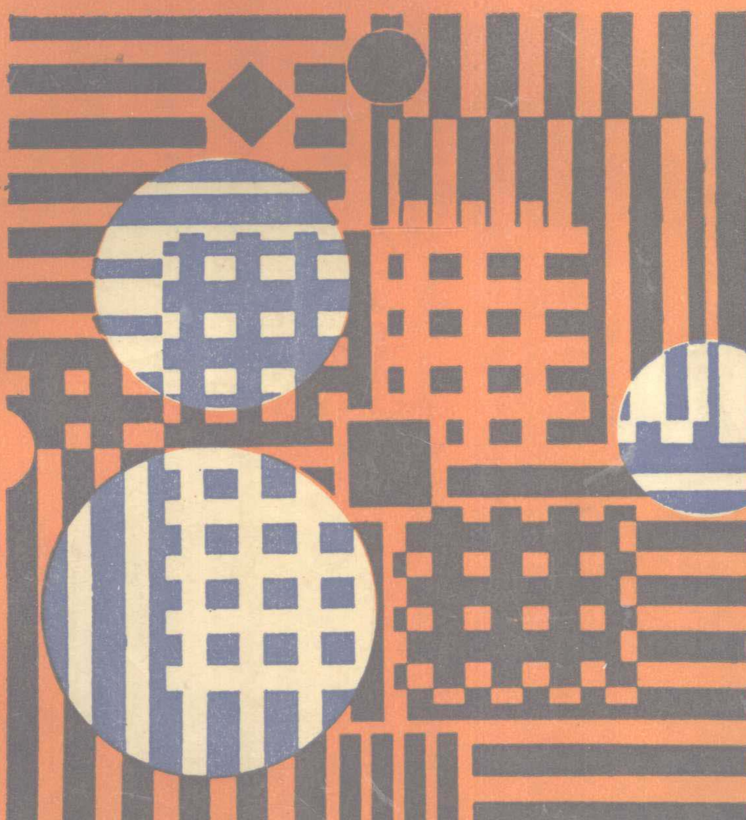
统计学系列教材

数理统计学

周兆麟
余望之 主编

暴奉贤 主审

暨南大学出版社



粤新登字13号

数理统计学

周兆麟
余望之 主编

暴奉贤 主审

•

暨南大学出版社出版

（广州·石牌）

广东省新华书店经销

广东省封开县人民印刷厂印刷

•

开本：850×1168 1/32 印张：12 字数：311千字

1991年12月第1版 1991年12月第1次印刷

印数：1——5000册

ISBN 7—81029—109—2/C·3

定价：5.50元

前　言

数理统计方法在经济领域中和其他领域一样有着广泛的应用。掌握该学科的理论和方法，对于统计、会计、经济管理等方面从事定量分析的工作人员来说是十分重要的。为了解决广东地区高等院校经济类专业开设数理统计学课程的教材问题，暨南大学、广东商学院、中山大学、广州外贸学院、佛山大学等高等院校商定编写出版本教材。

本教材可供高等财经院校或其他院校本科、电大、函授、夜大学、自修大学、专修科教学选用，并可根据专业要求、学时安排等具体情况对内容作适当增减。对于广大经济管理干部和其他领域的统计工作者、研究工作者来说，本书在充实其统计学基本知识和方法上也大有裨益。

本书由周兆麟（广东商学院）、余望之、（中山大学）主编。参加编写的有郑宗成、关伟德（中山大学）、李焯平（佛山大学）、张春汉、梁江浩、李纬(暨南大学)、戴念昆、何福祖、朱伟雄（广东商学院）、董黎明（广州对外贸易学院）。暴奉贤、谢启南、马俊林、熊达棣、杨慎辉、刘兆洪、谭中元及有关同志为本书的完成和出版提供了帮助，谨表示衷心的感谢。

由于作者水平有限，书中缺点错误在所难免，敬请读者批评指正。

编　者

1991年4月

导　　言

开展科学管理离不开定量分析，数理统计是定量分析的工具之一，对于管理工作者来说，研究和运用它是很重要的。当前，统计方法在许多领域有着极为广泛和深入的应用，几乎没有一门自然科学和技术科学不在某种形式下应用统计的方法。在经济管理领域里，质量控制与管理、经济预测与决策、公用事业服务、信息探索、因素影响与最优化研究等方面，统计方法的运用也日益获致显著的效果。

数理统计学科作为一门学科，它的研究对象和内容是什么呢？它的研究方法和程序又是怎样的呢？这里就这些问题作些介绍。

现实生活中有这样一类现象，在相同的条件下它可能发生，也可能不发生，表现出不确定性。或者说，它的发生呈偶然性。例如，为了了解产品质量，从一批产品中任意抽出一件来检验，它可能是正品，也可能是次品，事先无法肯定。又如要掌握生产的某种零件厚度尺寸数据的变化，了解到工人加工该零件时，其尺寸标准是 2 mm，但不见得正好每件都是 2 mm，即使在同一时间、地点，由同一名工人在同一台设备上就同一批原材料等条件下加工，每个零件的尺寸也不可能正好是 2 mm，很可能取 1.98mm、1.99mm、2.01mm等数值，这些零件有不同程度的差别，即表现出不确定性。又例如，从全厂职工上班时刻统计分析中看到，若上午 8：00时为标准上班时刻，在正常情况下每个职工实际进厂时刻却可能不一致，有的稍早，有的略晚，又是带不确定性的。再如，从经济分析中了解到，产量和单位成本指标之间

有密切关系。如某种产品第一月的产量为50件，单位成本是1050元，第二个月的产量为85件，单位成本是980元，在条件相同的情况下，第三个月仍拟生产85件，能否断言单位成本也会是980元呢？那不见得，实际上总会有出入，也呈现出不确定性。实际生活中，类似这些不确定现象的例子是很多的，我们也常和它们打交道，如明天会不会下雨？每粒种子种下去会不会发芽？掷一枚硬币，正面是否会朝上？得了感冒，几天可以好？等等。

这类现象为什么会呈现不确定性呢？我们知道，影响一个事物的因素是很多的，有些是基本因素，如零件加工，会受到工人技术水平、设备状况、原材料性能、加工条件等基本因素的影响。还有大量的偶然因素，诸如室温、场外微小振动、工人加工时精神状态等，这些大量的、作用相当微小的、彼此无关的、偶然的因素同样影响着零件的加工。上述因素作用的结果，综合地反映在数据的不稳定上，从而表现出不确定性。

不确定现象带偶然性，似乎不好捉摸，他们的变化是不是没有什么内在规律可循呢？从最简单的掷硬币这事来看，一个质量匀称的硬币，抛掷后正面朝上或背面朝上的机会是一样的。如果只抛一次，那就很难说正面一定朝上，但如果抛掷多次，计算正面朝上的次数m对抛掷总数n之比的频率指标m/n，我们会看到，随着抛掷次数的增加，m/n越会来越接近$\frac{1}{2}$这个常数。也就是说，在大量观察中，硬币正面朝上的机会大体上是一半，这和我们实际的想象也是相吻合的。频率在随着n越来越大而逐渐稳定在一个常数附近这个事实，在其他场合也会遇到。这些确定的常数描述了该现象发生可能性的大小，这是不确定性现象变化的一种规律的反映。又如，观察零件尺寸取值、职工上班时刻统计，也会看到，从个体上观察（就抽取的一个零件，登记一名职工的上班时刻来观察），它们的一次取值是偶然的，不确定的，但若通过大量观察，统计成百上千的零件尺寸，登记成百上千职工下班时刻，将它们系统

整理后就会发现：就某工人加工的大量零件寸尺看，尽管它们之间各不一样，但大部分零件的尺寸是在 2 mm左右，偏离 2 mm越大的零件越来越小，呈“两头小，中间大”大的山峰形状的曲线。就职工到厂的时刻而言，尽管各人来厂时刻不同，但从大量观察可看到，大部分职工是在8时以前某个时刻（比如7时55分）左右进厂，随着时间差距的拉大，提前很早到厂和迟到很晚的人，会越来越少，也大致呈“两头小，中间大”的情形。所有这些，和我们的通常想象也是相符合的，只要这些事物仅仅受大量偶然因素的影响，情况就会是如此。这反映出某种规律性的东西，即正是大量的偶然性中内含着某种必然性。统计的特点是作大量的观察，我们常把这种通过大量观察概括得到的规律性，称为统计的规律性。

不确定现象又称随机现象。数理统计学的研究对象就是这种不确定性现象，它利用一定的数学模型来描述随机现象的统计规律性，以之作为分析研究与指导解决大量实际问题的基础。

其实，用一定的数学模式来描述实践中一些变量的规律性，在其它场合我们也见过。比如说，几何上的圆，即一动点与一定点等距离运动所得的轨迹，这个几何模式的出现和对它的研究，也是通过大量实践的启发：太阳象“圆”，车轮、月亮、锅盖等具体事物都是“圆”的，把这个内在的共同规律性归纳和抽象出来，用一定的数学模式描述它，并取名为“圆”。有了圆这个理论上的模式，就可指导千千万万的实际事物，开展各种研究。数理统计也正是运用了这种抽象归纳的方法。

当然，就个别某天或某批去统计，职工上班时刻或零件尺寸的实际分布情况，不见得正好和理论上的光滑曲线相吻合，而会有一些出入。实际分布曲线和理论曲线不完全符合，这是正常的。如果就大量的天数或次数来统计，只要各天都只受大量偶然因素的影响，平均说起来就会与理论的情况十分吻合。理论上和实践中的若干出入，并不妨碍这种理论上曲线规律对实践的指导作

理。正好比我们在现实生活中观察到的各种各样的圆形事物，严格说来，世界上并不存在和理论上的几何圆完全吻合的形体，多少也会有点出入，但这并不妨碍几何圆对实践的指导作用。如设计一个圆盘，就得依靠圆面积的计算去考虑用料，等等。

我们所研究的一定的不确定性现象客体，统计上常称为总体。由于总体中包含众多直至无穷多个个体，无法一一研究，实际上只能从总体中抽取一个有代表性的局部，这一部分个体常称为样本（又称子样），先对样本作观察和研究，然后据以对总体的情况作各种推断。从研究内容上看，数理统计经常是通过对样本信息的研究来推断总体的有关结论；它还通过比较事物间的差异，分析影响事物变化的因素等来研究现象间相互关系；它也研究观察实验设计的科学方法等等。象数据整理和样本统计量研究，统计推断(包括估计和假设检验两大方面)、方差分析、相关分析和回归分析、抽样理论、质量控制和试验设计等，都是数理统计学研究的主要内容，因此，数理统计的内容是十分丰富的。

开展数理统计研究时，首先需要对所研究的事物赋予一定的数量进行测度，并加以收集整理。以一定的计量尺度对事物进行测度，根据事物的特性和刻划信息的深浅程度不同可分为这样几种：

①列名尺度。这是计量中最弱的一种尺度，它仅用事物的名称或给予一定的编号码表示事物间的区别。例如抽检产品时，为了定量分析方便，当抽到正品时赋以代码“0”，抽到废品时赋以代码“1”，这就是列名尺度的例子。

②顺序尺度。这也是一种弱计量尺度，当把某些事物按一定的特征排成由小到大（或由大到小）的顺序等级时，就是顺序尺度。例如产品分为优等品、一等品、二等品、等外品等等；企业按超额完成计划、完成计划、未完成计划分成各个等级。这种尺度不仅能说明事物间的区别，还能说明它们的优劣、大小、多少等，但不能说明究竟大多少、小多少。

目　录

第一章　随机事件与概率

第一节　随机事件及其运算

一、随机事件

在自然界中，人们观察到的现象可概括为两大类型：一为事前可以预测到的，即在某些相同的条件、环境之下，它所出现的结果总是肯定的，这一种现象称为**必然现象**。譬如说，水在100℃时会沸腾，重物从高处总是垂直落地等等；另一种现象是事前不可预知的，即在相同条件下重复试验，得出的结果未必相同。换言之，虽然知道只有几种不同的结果，但发展下去，结果如何事前都不能肯定，这种现象称为**随机现象**。例如抛掷一枚分币，结果可能是正面（有印值的一面）朝上或背面（有国徽的一面）朝上。

是不是随机现象没有什么规律可寻呢？事实上并非如此。人们经过长期反复观察和实践，逐渐发现所谓不可预测，只是对一次或少数次观察或实践而言。如在相同的条件下，进行大量观察时，随机现象就会呈现出某种规律性，因此它也是可以预测的。例如前面所说的抛掷分币的试验，如果分币是匀称的，放在手心上用一定的动作向上抛，让分币自由落在具有弹性的桌面上，重复地抛掷多次，正面朝上的次数比例大约是１／２，而抛掷的次数越多，频率就越接近这个比值。因此，现实世界里存在着具有如下特性的现象：在一定的条件下，有可能出现多种结果，事前人们并不能预言将出现哪种结果，但大量重复观察时，所得出的结果却呈现出某种规律性，这种规律性称之为**随机现象的统计规律**

性。**概率论**就是研究随机现象统计规律性的一门数学学科。

人们对自然现象进行观察或进行一次科学实验，称为一个**试验**。如果这些试验在相同条件下可以重复进行，而且每次试验的可能结果不止一个，并能事先明确所有可能结果，但每次试验的结果却事前不可预知，我们就称它为一个**随机试验**。在随机试验中，每一个可能结果一般称为此 随 机 试验的**随机事件**，简称**事件**。我们把不可能再分的事件称为**基本事件**。由若干基本事件组合而成的事件，称为**复合事件**。

〔例1.1〕 投掷分币一次的试验中，“正面朝上”和“反面朝上”，都是基本事件。

〔例1.2〕 投掷一颗骰子的试验中，“出现 1 点”、“出现 2 点”、……“出现 6 点”都是基本事件。“出现偶数点”也是一个随机事件，它是由“出现 2 点”、“出现 4 点”和“出现 6 点”这三个基本事件所组成，故为复合事件。

在试验中必然会发生的事件称作**必然事件**；不可能发生的事件称作**不可 能 事件**。在例1.2中，“点 数 不 大 于 6 ”是必然事件；“点数大于 6 ”是不可能事件。必然事件和不可能事件本来没有不确定性，但为今后讨论问题方便起见，我们把它们当作一种特殊的随机事件。

为了运用集合的概念来研究随机事件，我们将随机试验中的所有基本事件组成的集合叫做该随机试验的 **样本空间**，记为U，U中的元素就是试验中的基 本事件。因此，基本事件也称为**样本点**，我们用W表示样本点。

〔例1.3〕 写出下列随机试验的 样本空间U，并指出样本空间U中元素的个数。

（1）一个口袋中有三个球，分别标上号码1*、2*、3*，从口袋中任取一个球后不放回袋中，再从袋中任取一球，记录二次取球的结果；

（2）将（1）中的取球方式改为第一次取球后放回袋中，

再作第二次取球，记录两次取球的结果；

（3）将（1）中的取球方式改为一次从口袋中任取两球，记录取球的结果；

（4）将（1）中的取球方式改为不放回地从袋中一个按一个地取球，直到取到1#号球为止，记录取球的结果。

对于（1）因为取球有先后顺序，假设第一次取到的是1#号球，第二次取到的球应当是除1#以外的某个球，不妨设为2#号球，我们将这个基本事件记为（1，2），余此类推，我们可以得到样本空间如下：

$$U=\{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\}$$

样本空间U中元素的个数为

$$A_3^2=3\times2=6$$

对于（2）由于第一次取出的球放回后再取第二个球，所以可以取到两个相同的号码的球，例如第一次取到1#号球，第二次又取到1#号球，记作（1，1），此时样本空间为

$$U=\{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)\}$$

样本空间U中的元素的个数是3个球任取二个可以重复的排列数，即$3^2=9$。

对于（3），它与（1）的区别在于取出的两个球没有先后顺序，例如（1，2）与（2，1）是没有区别的，所以样本空间为

$$U=\{(1,2),(1,3),(2,3)\}$$

U中的元素的个数为

$$C_3^2=3$$

对于（4），其样本空间U为

$$U=\{(1),(2,1),(3,1),(2,3,1),(3,2,1)\}$$

U中的元素个数为

$$1+A_2^1+A_2^2=1+2+2=5$$