

数理统计浅说

陈永秉 编著

SHU LI TONG JI QIAN SHUO

农业出版社

数 理 一 十 浅 说

陈永秉 编著

农 业 出 版 社

数 理 统 计 浅 说

陈永秉 编著

农业出版社出版 (北京朝内大街 130 号)

新华书店北京发行所发行 农业出版社印刷厂印刷

787×1092 毫米 32 开本 9.75 印张 200 千字

1983 年 7 月第 1 版 1983 年 7 月北京第 1 次印刷

印数 1—17,500 册

统一书号 4144·465 定价 1.25 元

前 言

随着国家社会主义现代化经济建设的发展，数理统计在经济建设中的应用越来越广泛。运用数理统计方法分析工农业生产实践和科学研究中的资料，处理数据，安排试验，对提高生产技术、劳动生产率和产品质量，节约原材料和能源，降低成本，加强企业管理，增加积累等都有重大作用。对发展社会主义经济，加速四化建设具有重要意义。

本书部分章节系为河北省干部学校统计局长班《数理统计》讲座的讲稿，为了适应国家四化建设需要，将其修改补充，增加内容，使之完整。本书特点：以初等数学为基础，力求避免运用高等数学方法，并侧重数理统计方法在社会经济方面的应用。本书可供大、中专财经院校有关专业学生参考，也可供工农业技术人员、科学研究人员以及具有高中水平的在职统计人员和企业管理人员自学之用。

陈永秉

1981年2月

073639

目 录

第一章 绪论	1
第一节 什么叫数理统计	1
第二节 数理统计的作用	2
第三节 数理统计的方法	4
第四节 数理统计的内容	6
第五节 数理统计中常用的几个名词	10
第二章 概率论的基本知识	14
第一节 概率的概念	14
第二节 概率的统计定义	17
第三节 大数定律	20
第四节 随机变量	23
第五节 分布函数	28
第三章 统计特征数	32
第一节 统计特征数的意义	32
第二节 平均数	33
第三节 调和平均数与几何平均数	38
第四节 众数和中位数	46
第五节 标志变动度	49
第四章 统计检验	59
第一节 统计检验的概念	59
第二节 u 检验	62
第三节 χ^2 检验	67

第四节	t 检验	71
第五节	F 检验	76
第五章	方差分析	81
第一节	方差分析的意义	81
第二节	一个因素的方差分析	83
第三节	两个因素的方差分析	102
第四节	两个因素试验有交错作用的方差分析	115
第六章	回归分析	121
第一节	回归分析的意义	121
第二节	一元线性回归	123
第三节	一元非线性回归	140
第四节	多元线性回归	149
第七章	抽样法	159
第一节	抽样法的意义和作用	159
第二节	抽样数目的确定	162
第三节	抽样调查的组织方式及其抽样误差	176
第四节	抽样资料的推断	185
第五节	抽样法实例	187
第八章	质量控制	196
第一节	质量控制的意义	196
第二节	计量控制	198
第三节	计件控制	212
第四节	计点控制	218
第九章	正交试验设计	223
第一节	正交试验设计的意义及方法	223
第二节	正交表的结构	233
第三节	有交互作用的试验	243
第四节	水平数不等的试验	249

第五节	正交试验的方差分析	255
第十章	可靠性统计	259
第一节	可靠性统计的意义	259
第二节	故障率和可靠性尺度	260
第三节	平均寿命的计算	264
第四节	可靠性的抽样检验	272
第五节	可行性研究统计	280
附录		283
附表 I	常态曲线纵座标 $[\phi(t)]$ 与曲线下面积 $[\int_0^t \phi(t) dt]$ 的值	283
附表 II	t 的定性限值	284
附表 III a	F 分布表 ($\alpha = 0.25$)	285
附表 III b	F 分布表 ($\alpha = 0.05$)	287
附表 III c	F 分布表 ($\alpha = 0.01$)	289
附表 IV	χ^2 的定性限值	291
附表 V	相关系数值表	292
附表 VI	部分常用正交表	293

第一章 绪 论

第一节 什么叫数理统计

数理统计是以样本为根据，运用数学模型来推断总体的一门科学。运用数理统计，可以研究大量的自然现象和社会现象的规律性。

数理统计是以概率论为基础，对大量的偶然现象的统计资料进行分析、研究，得出这种现象概率的规律性，给以科学的解释。也就是说，它是运用大量现象的资料，研究母体中所抽取部分子样材料作为计算某些特征数，来推算母体所表现的概率规律性的科学。这样，数理统计就形成一门数理统计学。

例如，一台车床加工的母螺丝是很多的，如果我们从其中抽取若干只来测量它们的口径，则这些母螺丝的平均口径可能大于或小于这种母螺丝所规定的标准口径。这里，就能发现一定的概率规律性。这种概率规律性，就是属于数理统计所要研究的范围。

由此可见，数理统计所要研究的总体有其一定的特点。也就是说，总体中各个体的量或特征的表现是有变异的，并且是不确定性的，是随机性的。这是由于在总体中作用于各

个体的因素是复杂的，有主要因素，也有次要因素，而这些因素对总体的影响，是不确定的。这就要运用数理统计的方法，先在总体中运用随机抽样法抽取样本，而后根据样本计算样本指标，并以样本指标推算总体指标，再将推断总体指标加以检验，以确定其统计量的可信限，即可信的程度，用以说明抽样结果的可靠性。由此可见，根据有限数量的观测数据，进行推断所得的总体结果，总难免有某种不确定性。数理统计就是要运用它自己的特有方法，将其自然现象或社会现象的有限数据，将其归纳性地推断所产生的不确定性可以得到度量。这也就是数理统计的目的。

第二节 数理统计的作用

数理统计的方法，就是用来对大量偶然现象的规律性进行归纳的研究，因此，它的发展和现代化大生产以及科学研究工作的发展密切相联系的，它在最近几十年来得到飞速的发展。在我国，随着生产和科学的发展，特别是四个现代化建设的发展，数理统计的研究也将得到更快地发展，它的应用越来越广泛。数理统计与各种具体的研究对象结合起来，就能解决许多从实践中提出来的具体问题。例如农作物产量预测抽样，工业生产中的产品质量抽样检查和质量控制，气象预测中的天气预报，水文数据所作出的长期定量预报，工程设计中的安全系数的统计分析，以及国民经济的各部门中，数理统计都得到广泛的应用。

因此，数理统计的作用，主要可以归结为以下几个方

面：中銀海海 對善显的系关系因湖同文叠更突帮 四

一、帮助设计观察和试验 在实践中，在一些不必要或不可能用全面调查的方法进行观察的事物，可以运用数理统计的方法进行观察或试验。同时在设计观察或试验时，要考虑以最少的费用，最快的时间，取得及时准确的统计数据资料，或获得足够正确的答案，也只有运用数理统计的方法进行。例如，对农作物产量预测调查，只能采用数理统计中的抽样法调查。因为农作物产量预测，范围广，数量大，而且要求快，才能满足产量预测的要求。在抽样方法上，还要研究采用什么方法，以便既节约人力，又便于组织管理，用最少的费用，获取足够正确的抽样数据。

二、利用样本指标推断总体指标 数理统计利用样本指标推断总体指标，以便提高效率，保证以最少费用获得足够正确的答案。因此要采用抽样法抽取样本，并由样本指标推断总体指标。例如，对工业产品质量的检验，有些产品必须破坏产品来进行检验，如电灯泡的寿命检验。又如商业中的蛋品检验，医学中的肝功能验血等，都只能采用抽样法来抽取样本，由样本来推断总体。

三、分析现象之间的关系 运用数理统计方法，可以分析现象间的关系，如两变量数列之间的关系程度可用乘积矩相关系数来表明。两个变量数列的关系，可用数学模型代表，这种数学模型就是回归方程。

两个品质数列之间的关系问题，可把品质划分为等级，化品质数列为等级数列，用等级相关系数来表明品质数列之间的关系程度。

四、研究现象之间的因果关系的显著性 在试验中，因试验条件不同而引起的试验结果的差异，叫做“条件误差”。除试验条件以外的因素所产生的试验结果的差异，叫做“试验误差”。条件误差是系统性的，试验误差是随机性的。随机性的因素所起的作用是成正态分布的，而非随机性因素作用则不能形成正态分布。这种研究现象间因果关系的显著性，在数理统计中运用方差分析来进行。

此外，数理统计作为一门科学，它与计量经济学也有密切的联系。计量经济学是经济学与数理统计学相互配合相互渗透的一门科学。它的内容是：用数学模型，经过统计检验后，就可作为预测之用或作为制定计划的参考。例如投入—产出分析，就是运用数理统计方法，确定一套联立方程式来表达国民经济各部门之间产品的供给与需求关系。

第三节 数理统计的方法

数理统计是一门科学。与其它科学一样，它的指导思想是马列主义辩证唯物主义。因此，数理统计方法的理论基础是唯物辩证法。

但作为一门科学的数理统计学，还有它自己所特有的方法，它既运用数学的方法，也运用统计学的方法。

数理统计学所应用的数学方法是非常广泛的，特别是概率论，它成为数理统计学的基础。概率论所研究的概率规律性是一种数量规律性，概率规律性就是大量偶然现象或事件的客观存在的集体规律性。因为偶然因素容易发生变化，所

以只研究个别的偶然现象或事件，就不能对事物的进程得出一般的结论。偶然因素只有在大量的集合中，对偶然因素互相抵消，互相平衡，而通过大量观察可以看出事物的一般趋势，找出其中的概率规律性。这就是数理统计学应用数学方法的理论基础。此外，数理统计要应用数学原理和公式去计算数据的特征和进行分析、推断。

数理统计学也应用统计学的方法，它和一般统计学一样，要应用大量观察法、分组法和综合指标法。

一、大量观察法 统计学主要是研究大量现象，而数理统计学则是研究大量的偶然现象（或事件）的某一类客观规律的科学方法。这是由于数理统计学及其基础概率论，都是研究偶然事件，而世界上任何现象都是和其它现象有多种多样的联系，所以动态规律性总是和概率规律性或统计规律性相交织在一起。这些偶然事件，只有通过大量观察法，才能显示规律性。

二、分组法 分组法是统计学上的特有方法，数理统计学上也广为采用。数理统计在整理、计算或分析研究之前，一定要按照事物的特点将所收集的资料加以科学的分组，才能有正确的推断，得出科学的说明。

三、综合指标法 统计学上分析事物现象或发展过程时，常应用综合指标，如总量指标、相对指标和平均指标。数理统计也是如此。数理统计在计算统计特征数，进行分析和统计推断时，都必须运用综合指标法。

第四节 数理统计的内容

数理统计应用的范围非常广泛，正如前面所说的那样，它是涉及各门科学和工业、农业、商业、国防及科学技术等各个领域。同时，它所研究的问题的性质也是多种多样的。当前世界各国科学技术的不断发展和完善，特别是电子计算技术的发展，将促进数理统计的发展。在我国，由于社会主义现代化建设的发展，现代的工业、农业、国防和科学技术的不断发展，新的数理统计问题也随时提出新课题，有待进一步研究解决。同时，数理统计本身及其应用，也都在日新月异地变化着和发展着。

当前，我们要高度重视数理统计的研究，要迅速提高数理统计的理论水平，不断联系实际，解决问题，为我国实现四个现代化做出应有的贡献。

数理统计要研究的内容是很广泛的。本书不可能遍及数理统计的各个部门，只能是涉及在工农业生产中和科学研究中最常用的一些基本原理和几种常用的方法。今将其各主要部分简单介绍如下：

一、统计特征数 在工业生产中工人所生产的产品数，农业生产中的农作物产量，以及这些工农业产品数量之间都存在着千差万别的差异，这种差异的变化也是很大的，是多种多样的。如果要对这些差异进行检查，给以逐个地对比，既不胜其繁，在时间上也不可能，也没有必要，更不可能对检查的总体有一个概括而全面的了解。这样，就只能通过用

少数几个数字来说明总体的情况。那么，怎样用少数几个数字来说明总体的情况呢？这就是数理统计中首先要碰到的而必需解决的问题。例如在工厂中要检查产品的质量，如果逐个检查是不可能的，只能运用数理统计的方法，抽选若干产品加以检验，求得一个平均数，再将这种平均数与标准平均数对比，才能概括地说明该产品总体的概况。这种用少数几个数字来说明整体的情况，在数理统计中叫做统计特征数。统计特征数能概括地说明一件事物的性质。

二、统计检验 在比较二件事物或二批同类的工业产品或农业产品的差异时，往往由于产品质量上存在着差异，测量时又存在着差异，情况比较复杂。数理统计在研究这类问题时，要进行分析，首先要对两件事物或二批同类产品的性质给以某种假设，例如假设两批工业产品所要比较的指标特征数是相同的，而后在每一批产品中抽取若干个体进行测量，如果被测定产品的某一特征的差异超过了一定程度时，即认为所假设为不可信。如果差异不超过一定程度时，那么就认为所作假设是可信的。这在数理统计中叫做可信程度，或叫可信度。这种方法，就叫做统计检验。

三、方差分析 在比较二批同一种类的产品时，两种产品的质量存在着一定差异。引起这种差异的原因是很多的，可能是由于原材料的不同，可能是由于加工处理方法或工艺技术的不同，也可能是由于测量时所引起的测量误差。这些因素，可能是一种因素引起差异，也可能是二种或三种因素纠缠在一起所引起差异。运用数理统计方法，可以对这类问题进行分析研究，作出正确的处理，也可以计算各个因

素所产生的影响程度。因此，这种分析影响事物变化因素的方法，在数理统计中，叫做方差分析。

四、相关分析 在工农业生产中和科学实验中，许多事物相互之间有着密切关系，例如在纺织厂中，棉纤维成熟度与单纱强力的关系，棉纤维成熟度高时，单纱强力也较高。在农业生产中，使用化肥量对农产量的关系，深耕程度、雨量及田间管理等与农产量的关系，都是有一定关系，数理统计的方法，就要分析一件事物的两种性质之间可能存在的相互关系，或者是一事物多种性质之间可能存在的相互关系。对于各种现象间的关系，在数理统计中叫做“相关”，这种分析的方法，叫做相关分析或叫做回归分析。

五、抽样法 要想知道一件事物的性质，我们不可能，也没有必要对被研究事物的全部作个体检查，只要从其中抽取一部分进行检查和分析即可。例如，对工业产品质量检查时，对一批电灯泡的质量检查，不可能把全部电灯泡破坏了来检查；对商品检验中，如对鸡蛋的检查，不可能把全部鸡蛋敲破检查；又如在医药上的检查；对人的血液检查，只能抽取几个毫升的血液加以检验。这些检验的方法，都不可能采用全面的检查，只能抽取一部分来说明全体的方法。在数理统计上，叫做抽样法，或叫抽样检验，或叫抽样推断。

六、质量控制 现代化生产，需要现代化的质量管理，用数理统计方法，对产品质量中存在的散差进行计算分析，运用散差分析统计的规律性，以掌握产品质量的统计分析，再按统计推断进而预测和控制产品的质量。因此，统计质量控制抓住统计规律，可预见全部产品的情况，用不着

去注意单个产品，从杂乱的变动中找出规律性，例如，在金属线验收检查时，要从许多金属线中取出几个试样测定其抗拉强度，而后根据测试结果决定这批金属线是否全部合格。在此，测得的每个数据是否满足规格要求？可不必关心，主要是根据这些数据推断整批产品是否合格。这种方法在数理统计中叫做统计质量管理，或叫质量控制。

七、正交试验设计 正交试验设计，是多因素的实验设计，它的目的在于用较少的人力、物力，在较短时间内获得比较多的成果。它是进行科学研究工作的依据，是保证科学研究成果的重要条件，是处理实验结果的先决条件。当前，科学实验成果的水平，很大程度上取决于试验设计的水平，实验设计错误，会使全部科研失去意义，造成人力、物力的浪费，丢失宝贵时间。同样，如果没有先进的试验设计水平，就不会有先进的科学实验成果。而这种科学实验，就是要靠数理统计中的正交试验设计来解决。

八、可靠性统计 在工业生产设计中，提高可靠性的方法中，要用到数值方法，比较后备设计，采用大安全系数，使用冗余法和过应力试验等，都要用到定量分析，估计和纠正设计缺点以及编制维修计划，也要用到数理统计方法。其它方法，如失效分析和功率老化，也要运用数理统计方法，效果较好。在近代电子工业、航空工业和宇宙空间工业中，对重量和性能，对计算安全系数和了解应力、强度分布以及他们的关系都很重要，在计算中必须运用数理统计方法。这种方法就叫做可靠性统计。

总之，数理统计经常遇到的是上述几类问题，但并不包

括数理统计的全部内容。这里所讲的内容，只是结合工农业生产等国民经济方面的应用，运用通俗地叙述数理统计中最基本的原理和几种常用的方法。并且所运用的数学基础，基本上是限于初等数学的范围，以利于更广泛地供一般统计工作者和企业经济管理工作者参考。

第五节 数理统计中常用的几个名词

在学习数理统计时，必须首先弄清楚几个常用的名词，如母体、个体和子样，抽样总体和全及总体，抽样指标和全及指标等。

一、母体、个体与子样 所谓母体是指某一次统计分析中，所要研究的对象的全体，而个体则是指所要研究的全体中的一个单位。例如某机械制造厂，要了解全部生产的机器，则全部机器是母体，而每一部机器就是个体。如果研究全国整个机械制造业的机器，则对全国的所有机械制造厂为母体，每一个厂则为个体。因而，母体与个体并不是一成不变的，而是根据所要研究的任务而定的。

对于上述机械制造厂所生产的机器或零件是一定的数量，因此一个母体可以包含有限个个体。如果遇到另外一种情况，例如对化工厂某种产品的制造温度，这时测定的数字变化很大，可以是无限的，因此，这样测定的个体是无限个数。这样，在数理统计中把前者叫做有限大母体，后者叫做无限大母体。当有限大母体的个数相当多时，在数理统计中常把它当作无限大母体，但如果母体所含个体不甚多