

产品验收中的 数理统计方法

张福昌

中国对外经济贸易出版社

产品验收中的数理统计方法

张福昌 编著

中国对外经济贸易出版社

2573/25

产品验收中的数理统计方法

张福昌 编著

※

中国对外经济贸易出版社出版
(北京安定门外大街东后巷28号)

新华书店北京发行所发行

海淀印刷厂印刷

※

787×1092 毫米 32 开本17.875印张385千字

1987年4月第1版 1987年4月第1次印刷

印数: 1—5500

书号: 4222·68

定价: 3.65元

编者的话

在我国的经济改革中,加强对产品质量的检验工作,有着特殊重要的意义。我国已经有了一支庞大的从事商品检验、产品验收、质量控制的专职人员队伍、如何用统计理论来指导检验工作,提高检验方法的科学性,是当前十分迫切的任务。

本书所介绍的产品验收中的数理统计方法,都是一些经常用到的统计方法。有关产品验收实践以及科学研究中的许多问题都可从中求得解决。书中汇集了国内外一些最新方法和资料,可供读者参考。

数理统计本身同其他数学方法一样有着严格的理论依据,但本书的重点放在方法的具体应用上:主要是通过大量实例说明对各种类型的问题应该怎样进行统计处理,并介绍具体的计算与分析步骤,以及怎样科学地解释检测的结果。

书中采用的例子大部分是编者从近几年所接触到的实际例子中选取的,在此中间得到国家商检局的大力支持。本书成稿以后经中国煤炭科学院高级工程师罗颖都同志作了认真审阅和修改,谨在此表示深切的谢意。

由于水平所限,书中内容难免有错误和不妥之处,诚恳希望读者批评指正。

目 录

第一章 引论	(1)
一、产品验收与数理统计	(1)
二、检测数据的不确定性	(4)
三、总体、个体和样本	(5)
四、检测数据的分类及处理步骤	(7)
五、概率	(9)
六、统计量与参数	(9)
第二章 检测数据的频数分布	(11)
一、频数分布表	(11)
二、频数分布直方图	(15)
三、频率分布表与频率分布图	(17)
四、累积频数与累积频率	(18)
五、离散变量的频数分布	(21)
第三章 检测数据的统计特征数	(25)
一、算术平均值 $\bar{x}$	(26)
二、算术平均值 $\bar{x}_w$	(30)
三、几何平均值 G	(32)
四、调和平均值 H	(34)
五、中位数 (中值、中序数) Me	(35)
六、众数 M_o	(37)
七、极差 R	(42)
八、平均差 δ	(43)
九、方差 σ^2, S^2	(46)

十、标准偏差 σ , S	(47)
十一、变异系数 $C.V$	(54)
第四章 概率及随机变量分布理论	(59)
一、概率的统计定义.....	(60)
二、随机变量与分布函数.....	(64)
三、正态分布.....	(68)
(一) 标准正态分布.....	(72)
(二) 正态概率纸.....	(78)
(三) 偏度 (Degree of Skewness)	(80)
(四) 对数正态分布.....	(82)
四、二项分布.....	(84)
(一) 二项分布的计算.....	(86)
(二) 二项分布的特征.....	(87)
(三) 二项分布的另一种表达式.....	(88)
(四) 负二项分布.....	(91)
五、普哇松 (Poisson) 分布 (波松分布).....	(93)
六、超几何分布.....	(98)
(一) 超几何分布的几个重要特征.....	(101)
(二) 超几何概率的近似计算.....	(102)
七、几种特征数分布.....	(106)
(一) 样本平均值 $\bar{x}$ 的分布.....	(107)
(二) 中位数的分布.....	(109)
(三) 样本方差 S^2 的分布.....	(109)
(四) 极差 R 的分布.....	(111)
第五章 复合统计量的分布	(113)
一、 t 分布.....	(113)

(一) t 分布曲线	(113)
(二) t 分布一定区间的面积	(117)
二、 F 分布	(119)
三、 χ^2 分布 (卡方)	(124)
第六章 对检测总体的统计估算	(132)
一、对总体参数的定值估计 (点估计)	(132)
(一) 对总体均值 μ 的定值估计	(133)
(二) 对总体标准偏差 σ 的定值估计	(135)
二、区间估计	(140)
(一) 总体均值 μ 的区间估计	(140)
(二) 两总体均值差的区间估计	(145)
(三) 总体比率的区间估计	(148)
(四) 总体比率差的区间估计	(149)
(五) 总体方差的区间估计	(151)
(六) 区间估计所需样本数 n 的确定	(153)
第七章 显著性检验	(157)
一、统计结论的两类错误	(157)
二、关于差异的显著性	(161)
三、置信限与显著性检验的关系	(162)
四、平均值的显著性检验	(164)
五、平均值差异的显著性检验	(170)
六、三组以上平均值差异的显著性检验	(177)
七、关于比率 (不合格品率) 的显著性检验 (计 数值数据)	(180)
(一) 对于总体比率 (不合格率) 的显著性检 验	(180)

(二) 关于两总体比率相等的显著性检验·····	(182)
(三) 三个以上总体比率的显著性检验·····	(184)
八、关于总体方差的显著性检验·····	(185)
(一) 总体方差等于 σ_0^2 的显著性检验·····	(185)
(二) 两总体方差相等的显著性检验·····	(187)
(三) 两个以上总体方差相等的显著性检验···	(189)
九、关于分布形式的检验·····	(192)
第八章 有效数字的应用及检测数据的取舍和插补	
·····	(206)
一、有效数字的应用·····	(206)
(一) 有效数字·····	(206)
(二) 有效数字的计算法则·····	(207)
二、异常数据的发现与取舍·····	(210)
三、数据插补·····	(220)
第九章 误差原理及计算 ·····	(225)
一、误差的定义·····	(225)
二、误差的分类·····	(227)
三、检测误差的表达方法·····	(232)
四、准确度和精密度·····	(235)
第十章 方差分析 ·····	(240)
一、离差平方和的加和性与方差分析·····	(241)
二、单因素方差分析·····	(245)
(一) 总离差平方和的分解·····	(246)
(二) 方差的估算·····	(250)
(三) 多重比较·····	(252)
(四) 单因素方差分析举例·····	(253)

(五) 单因素试验的极差分析·····	(258)
三、两个因素的方差分析·····	(260)
四、有交互作用的两个因素的影响问题·····	(267)
五、系统分组的方差分析·····	(277)
第十一章 回归分析 ·····	(282)
一、回归分析·····	(282)
二、散点图与回归直线·····	(284)
三、回归直线求法与最小二乘法原理·····	(286)
四、回归直线的精度与置信区间·····	(291)
(一) 回归直线的方差分析·····	(291)
(二) 回归直线精度与置信区间·····	(295)
(三) 回归方程的稳定性·····	(297)
五、相关系数与相关关系显著性检验·····	(299)
(一) 相关系数的含义·····	(299)
(二) 相关系数的分布·····	(303)
(三) 相关关系的显著性检验·····	(305)
六、回归直线的简便求法·····	(311)
七、两条回归直线的比较·····	(313)
八、一元非线性回归·····	(318)
九、回归分析的其他一些应用·····	(324)
第十二章 正交试验法 ·····	(327)
一、利用正交表安排试验·····	(329)
(一) 正交拉丁方·····	(329)
(二) 正交试验设计的几何意义·····	(332)
(三) 正交表·····	(333)
(四) 利用正交表安排试验·····	(335)

(五) 试验结果的直观分析.....	(337)
(六) 试验结果的方差分析.....	(340)
二、对因素水平数不同的试验安排.....	(342)
(一) 利用混合型正交表.....	(342)
(二) 拟水平法.....	(346)
(三) 并列表.....	(349)
(四) 部分追加法.....	(354)
三、有交互作用的试验设计.....	(357)
(一) 两列间的交互作用表.....	(358)
(二) 有交互作用的试验安排.....	(360)
(三) 试验结果的分析.....	(363)
第十三章 产品验收检测方法的几项统计指标.....	(367)
一、准确度.....	(367)
(一) 准确度的量度及表达.....	(369)
(二) 提高检测分析准确度的措施.....	(384)
二、精密度 (Precision)	(386)
(一) 精密度的表达和计算.....	(387)
(二) 安排平行试验测定方法精密度.....	(390)
(三) 不进行重复测定来确定方法精密度.....	(393)
三、允许差.....	(394)
(一) 允许差分类.....	(395)
(二) 制定允许差的协同试验安排.....	(396)
(三) 试验数据的统计分析与允许差的计算... (398)	
(四) 允许差在测定农药产品有效成分规格中 的应用.....	(416)
第十四章 抽样检查方法.....	(431)

一、基本概念	(432)
(一) 目前产品验收的主要方法	(432)
(二) 单位产品和批	(435)
(三) 随机抽样法	(441)
(四) 抽样检查方法的分类	(444)
二、计数抽样检查的原理和方法	(448)
(一) 计数标准型抽检方案	(465)
(二) 计数调整型抽检方案	(468)
三、计量抽样检查的原理及方法	(470)
(一) 以不合格品率衡量批的质量	(472)
(二) 以平均值衡量批的质量	(478)
四、进口袋装化肥重量的抽样检查方法	(484)
五、对散装物品进行化学分析时的抽样检查方法	(496)
(一) 分析样品抽取的初步考虑	(497)
(二) 散装物料的抽样	(498)
(三) 偏集物料 (Segregated Material) 的抽 样	(502)
(四) 离散单元物料的样品抽取	(503)
(五) 颗粒混合物取样的粒度问题	(504)
附 录	
主要参考文献	(555)

第一章 引 论

一、产品验收与数理统计

在产品验收、质量控制工作中人们越来越广泛地应用数理统计方法，即根据产品质量波动的统计规律，通过抽样检查来达到评估、预测产品质量的目的，同时也促使生产部门提高质量减少次品。因而从现状及发展趋势来看，数理统计方法作为产品验收、质量控制以及科研工作不可缺少的工具越来越受到人们的重视。

数理统计扎根于概率论。这是因为作为数理统计的一个主要任务是要进行“从部分推断整体”的统计推断，这就注定了，在作出这种推断的出发点（问题的提法）、所用方法和结论的表述方面，都极大地依靠以研究随机现象为任务的概率论。因此，概率论是数理统计学的基础，而数理统计学是概率论的应用。

有关数理统计的定义，至今仍没有统一的论述。要对数理统计下一个严格的数学定义，是很困难的，这需要比较高深的数学知识和大量的数学符号。对于应用数理统计作为工具的广大科学工作者来说，并没有必要去探究这些。就本文所涉及的内容来看，可作这样的理解：数理统计是以概率论为基础，是一门关于收集、整理和分析受随机性影响的数据，以便从中得出有关数据所来自总体的统计性结论的数学学科。更进一步讲，是研究从一定母体中所抽子样的某些特征

数字所表现的概率规律性的科学。

例如，我们要验收一批产品，其质量规格按合同规定应符合某一数量值。现在从这批产品（总体）中抽出 n 个，进行检验测定，取得数据为 $x_1, x_2, \dots, x_n$ ，就要从这些数据去估计整批产品的质量情况。在此，为收集数据 $x_1, x_2, \dots, x_n$ ，应设计一种抽样方法，以保证抽出的 n 个产品有尽可能强的代表性。至于随机性影响，一是由于抽样本身的误差，二是由于组成产品批的各产品个体之间本来就不绝对一致，另一方面在于测量本身也存在着误差，因此作出的结论是统计性的，即其正确性不是百分之百而只能用一定的概率来刻划。例如，结论可以是这样：“以95%的概率说明这批产品的不合格率在3~5%之间”。其所以不能作出百分之百的正确结论，理由很明显：因为我们所观察到的只是总体的一部分，而需要作结论的却是总体，何况检测过程中尚有随机误差的干扰。

数据的收集有两种方式：一是通过抽样检查，二是通过安排特定的试验。

抽样检查是在这种样一种场合进行，即所研究的总体是由一些有形的个体所构成的。例如，为要研究某地区稻米中残留农药的含量，这时需要决定抽多大数量的个体作为样本，以及应采取怎样的抽法可以使抽出的稻米在本地区有尽可能强的代表性。诸如此类的问题便形成了数理统计学中的一个专门分支——抽样技术。

安排特定的试验适用于这样的情况，即总体中的“个体”并非现实而有形地存在，必须通过试验去“实现”出来，并且试验条件在一定限度内是人所能控制的。例如，一个检验

方法与试剂、仪器、温度、电压等诸因素的水平有关。在这种情况下，总体应被理解为“在种种条件下所能检验的全部结果”。特定试验的内容就是选择若干组条件来进行检验测试。但如何挑选这些条件呢？这就需要考虑到这些条件要有代表性，它们能尽量消除系统误差，能够使试验次数恰当，能够使所得到的试验数据有便于进行统计分析的结构，等等。这些问题的讨论便构成了数理统计学的另一个重要分支——试验设计。

观察和检测所得的全部数据是对总体进行“统计推断”的依据。在推断之前，需要对一大堆杂乱无章的数据作整理加工工作，如绘制图表，计算某些平均数和其他特征数字等。这些工作还不是推断，因为它们还停留在数据本身，没有与总体联系起来。拿前面某地区稻米中残留农药含量研究的例子来说，从抽出的稻米中算出农药含量平均值为 0.51ppm ，这就不是统计推断，但是如果使用统计方法，作出诸如“可以用95%的概率断言，该地区稻米中残留农药的含量在 $0.48\sim 0.53\text{ppm}$ 之间”的结论，这就是统计推断了。当然统计研究的最终目的，是要达到预测和控制，这里就不多谈了。

总之，数理统计对于产品验收、商品检验、工业生产、质量控制等方面都有广泛的指导作用。但把统计方法用于实际问题，不是一件简单地套公式的工作，除了应对方法有透彻的理解之外，还要掌握所研究问题的专业知识、经验以及判断力等等。

二、检测数据的不确定性

由于随机性的影响贯穿于产品验收过程的始终，因而任何检测都不可能得到与实际情况完全相符的测定值。例如，上节调查某一地区水稻中残留农药含量一例，尽管采用尽量合理的抽样方法，但仍会存在有抽样误差；加上测定方法本身系统误差以及随机误差的影响，因此，采用统一的方法、仪器，由同一人员多次测定或由不同人员测定，由于实验室不同，以及由于不同的温度、湿度、气压等环境因素的影响，都会使每次检测的数值不尽相同。多次测定所得结果不可避免地一定范围内出现数值上的波动，这种现象叫做检测数据的不确定性。

数据是人们用来了解和研究客观事物的性质、运动规律，以及事物间相互关系的基本素材，虽然由于随机性的影响，使得检测数据都具有不确定性，但是这些参差不齐的测定值既然是被测对象某种性质的客观反映，就不会是“杂乱无章”的，其中必寓有反映这种性质的某种客观规律。就是说，这些数据具有两重性，一是波动性，二是规律性。

为了更好地发挥数据的效能，例如，正确估计所测结果的可靠程度，在数据中提取更多需要的信息，对结果提出更合理和准确的解释等，就要对数据进行一系列数学上的加工处理，这叫做检测数据的数学处理，简称数据处理。

数理统计就是从有波动的数据中找出其规律性的一种数学方法。

具体说，数理统计方法使人们能从最少的一组样本观察

结果中以最大的可靠程度（即以相当大的概率）作出关于对总体的某一判断，同时，算出发生错误判断的可能性。因此，数理统计学为科学研究、产品验收、商品检验、产品质量控制的精确性、可靠性和经济性提供了科学方法。

在这方面数理统计经常处理的三类问题是：

- （1）反映样本特征的统计指标的计算；
- （2）总体特征值的估计；
- （3）统计假设的检验。

三、总体、个体和样本

总体和个体 总体也称母体，是指在进行某次统计分析工作中，所要研究的对象的全体，而个体则是所要研究的全体对象中的一个单位。例如，我们研究在正常生产条件下维尼纶的纤度，那么，凡是正常生产条件下生产的纤维，其纤度的全体就是一个总体，而每一个纤度则是一个个体。当研究对象改变时，总体和个体也随之改变。在整理数据之前，必须把它们弄清楚。因为总体、个体并不是一成不变的，而要根据每一次研究的任务而定。如果我们现在要研究某钢厂5月15日这一天所产钢材的拉伸强度，那么，这个厂5月15日这一天所生产的所有钢材的拉伸强度便构成我们研究的全部对象，也就是构成我们研究的总体，而5月15日这天该厂生产的每一根钢材的拉伸强度则是我们研究的个体。可是，如果我们现在要研究这个钢厂最近三个月（即九十天）中每天所产钢材平均拉伸强度逐日变化的情况，那么，这家钢厂最近三个月（即九十天）中每天所产钢材的平均拉伸强

度的全体便成为研究的对象，也就是我们要研究的总体，而某天所产钢材的平均拉伸强度则为我们研究的个体。

样本 样本也称子样，是指总体的一部分，也就是从总体中被抽出来的个体的集合体。例如，为研究正常生产条件下维尼纶的纤度，我们从中抽取 100 个个体进行研究，这 100 个个体的集合就叫样本。

样本中所含个体的数目，叫做样本的大小或样本容量。

总体的性质由其中的各个个体的性质而定，所以要了解总体的性质，就必须测定各个个体的性质。但在实际检测过程中，要把总体中所有个体都加以测定会有种种困难而无法实现。这主要是指：第一，在很多情况下，总体中个体数目甚多，甚至近似无限多，事实上不可能把总体中所有个体都加以测定。例如，测定某地区水源的农药污染情况，就属于这种类型。第二，特别是那种由无形个体组成总体的情况，就属于这种类型。在这种情况下，人们无法对其所谓个体一一进行测定。比如，对一个特定试验方法的评定分析就属于此类。第三，也有不少这样的情况：总体中的个体数目并不多，但对各个体的某种性质的测定是一种具有破坏性的测定，如灯泡的寿命试验、自行车架的冲击试验等。试验一个也就破坏一个，因此，尽管总体所含个体的数目不是很多，也不允许全部加以考察。我们只能通过抽取一定数量有代表性的个体组成的样本来了解总体。

数理统计方法就是要解决如何从样本来研究总体的问题。