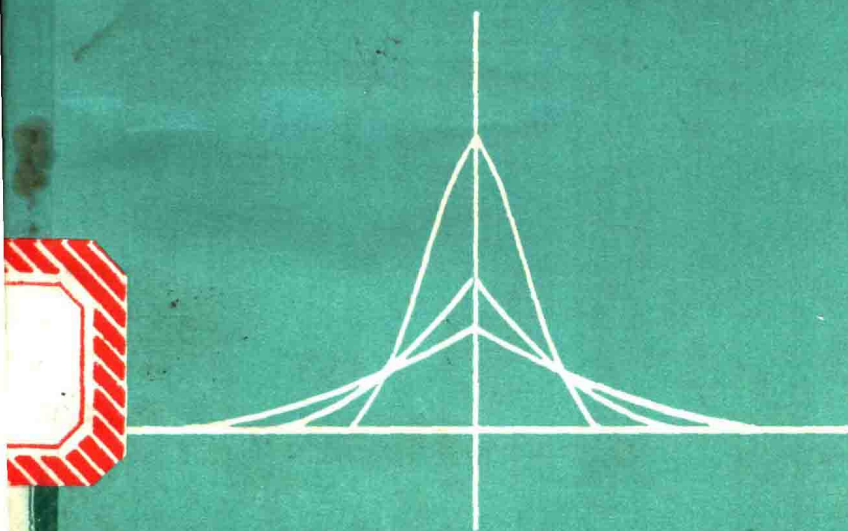


缪征明 编著

四川科学技术出版社

数理统计 在 分析化学中的应用



数 理 统 计
在 分 析 化 学 中 的 应 用

缪 征 明 编 著

四川科学技术出版社

一九八七年一月 成都

责任编辑：李世勋

封面设计：李 勤

技术设计：李世勋

数 理 统 计 在分析化学中的应用

缪征明 编著

四川科学技术出版社出版发行
(成都盐道街三号)

苍溪县装潢印刷厂 印刷

国际书号：ISBN 7—5364—0188—4/O·8

统一书号： 298·126

1987年9月第一版 开本787×1092毫米1/32

1987年9月第一次印刷 字数155千

印数 4500 册 印张9.25插页

定 价：1.95元

前 言

分析化学是一门实践性和应用性很强的学科，无论是在它本身的研究和发展过程中，还是在它为科学研究和国民经济各部门服务的过程中，都必须进行实验。实验的数据必须经过合理地、科学地处理和评价，才能得出正确的结论。因此，可以毫不夸张地说，如果没有正确地数据处理，即使是分析实验很出色，也不能得出合乎实际的结论。对一个分析化学工作者来说，如果他缺乏数据处理的知识能力，那么他就不能很好地完成任务，也就不能算是一个很合格的分析工作者。正因为这样，目前国内外一些《分析化学》课程的教材中，把误差和数据处理常专设一章，篇幅也相当多，虽然如此，但仍感到难以讲解清楚，对于分析化学专业的学生和专门从事分析化验工作的技术人员来说，更是不能满足需要。因此在许多大专和中专学校的分析化学专业的教学计划中，开设了有关数据处理的课程或选修课程，许多工矿化验室中从事分析化验工作的同志，也要求举办有关数据处理的学习班和提供学习资料，本书就是为了满足这种需要而编写的。

分析化学的实验数据中，必然包含有偶然误差，因此必须用概率论和数理统计的方法进行处理。所以，分析化学中的数据处理也就是数理统计在分析化学中的具体应用。对一个分析化学工作者来说，对于大量的概率论和数理统计术语

和繁难的公式，往往是感到生疏和困难的，特别是对公式的推导。然而，作为分析化学工作者，主要要求能够应用数理统计的方法正确地处理和评价实验数据，而不必一定要求熟悉全部术语和能够推导公式，这样就不至于使初学者望而生畏，同时可以收到学以致用，事半功倍的效果。基于这种认识，所以本书不是作为数理统计方面的专业著作编写的，而是从分析化验工作的实际出发，深入浅出的介绍数理统计在分析化学中的具体应用。在编写中，力求作到理论结合实际，简明扼要，通俗易懂，便于自学，学以致用。每章都附有复习题和练习题，以供学习者巩固和应用所学知识。

限于作者水平，错误和不妥之处在所难免，欢迎读者指正。

缪 征 明

于四川教育学院

一九八六年六月三十日

本书所用符号

在有关的数理统计书籍中，所用的符号不尽一致。符号的混乱，往往会给学习造成困难，本书在全书中统一符号的概念，但不一定与其它书籍完全一致，这点希读者注意。

下面将本书所用符号按英文和希腊文及其字母顺序列表如下，以供查检。

A ——用极差计算置信界限的系数

A_2 ——平均值的质量评价因子

$AD(\bar{D} \text{ 或 } \bar{d})$ ——样本平均偏差

a ——变差

$\hat{a}$ ——回归方程的截距的估量值

$\hat{b}$ ——回归方程的斜率的估量值

CI ——置信区间

CL ——置信界限（置信限）

$CL_{\text{上限}}$ ——极差的质量控制图的控制上限

$CL_{\text{下限}}$ ——极差的质量控制图的控制下限

C ——用极差换算标准偏差的无偏估量系数

$D(d)$ ——单次测量的偏差

$\bar{D}$ ——同 AD

D_3 ——极差质量评价下限因子

D_4 ——极差质量评价上限因子

d ——同 D

$\bar{d}$ ——同 AD

d_i ——成对数据的差值

$\bar{d}_i$ ——成对数据的差值的平均值

df 或 f ——自由度

E ——绝对误差

$E\%$ ——相对误差

$E_{\bar{x}}$ ——平均值的绝对误差

F —— F 检验法的统计量

G_1 和 G_n ——格拉布斯 (Grubbs) 检验法的统计量

H_0 ——否定假设

H_1 ——备择假设 (消解假设)

h ——精密度指数

k ——组数

L_t ——代用 t 检验法的统计量 (用于成对试验)

M ——代用 t 检验法的统计量 (用于两个平均值的比较)

$M_{\text{间}}$ 、 $M_{\text{内}}$ 、 M_E ——组间、组内、偶然误差的均方 (方差估计值)

n ——测量次数

P ——概率

R ——极差

$\bar{R}$ ——极差的平均值 (平均极差)

R_d ——成对试验数据差值的极差

r_{10} 、 r_{11} 、 r_{12} 、 r_{22} ——改良 Q 检验法统计量

S ——样本标准偏差

$S_{\text{合}}$ ——合并标准偏差

S^2 ——同V

$S_{\text{合}}^2$ ——合并方差

S_o ——剩余标准偏差（观测标准偏差）

$S(b)$ ——回归方程斜率的标准偏差

S_d ——成对试验数据差值的标准偏差

$S_{\bar{x}}$ ——平均值的标准偏差

$S(x_j)$ ——回归方程中自变量的估量值的标准偏差

$S(\bar{y})$ ——回归方程中因变量的平均值的标准偏差

$S(y_j)$ ——回归方程中因变量的估量值的标准偏差

t ——斯图登 t （“Student's t ”）或学生 t

U ——同 $V_{\text{回}}$

V ——样本方差，偏差平方和

$V_{\text{总}}(V_T)$ ——总平方和

$V_{\text{回}}(U)$ ——回归平方和

$V_{\text{余}}$ ——剩余平方和

V_E, V_A, V_B ——偶然因素平方和，A因素平方和，

B因素平方和

$\bar{X}(\bar{X})$ ——样本平均值

$\bar{\bar{X}}$ ——总体平均值

y ——回归方程中因变量的估量值

α ——危险率，即拒绝好的结果的概率

β ——置信水平，即接受好的结果的概率

γ ——相关系数

δ ——总体平均偏差

μ ——真值

σ ——总体标准偏差

σ^2 ——总体方差

目 录

第一章	基础知识	(1)
§ 1 · 1	数据处理与数理统计方法	(1)
§ 1 · 2	有效数字及其运算规则	(3)
§ 1 · 3	有效数字及其运算规则在分析化学 中的重要作用	(8)
§ 1 · 4	分析化学实验中的误差及其分类	(11)
§ 1 · 5	几个基本数语的涵义	(16)
	复习题	(43)
	练习题	(45)
第二章	偶然误差的理论基础	(47)
§ 2 · 1	偶然误差的基本性质	(47)
§ 2 · 2	偶然误差的正态分布	(49)
§ 2 · 3	偶然误差的概率	(62)
§ 2 · 4	t分布	(67)
	复习题	(70)
	练习题	(70)
第三章	平均值的可靠性及分析结果的 合理表示方法	(72)
§ 3 · 1	最小二乘法原理与平均值	(72)
§ 3 · 2	平均值的误差及其影响因素	(74)

§ 3 · 3	平均值的置信界限.....	(76)
§ 3 · 4	分析结果的正确表示方法.....	(85)
复习题.....		(87)
练习题.....		(87)
第四章	分析数据的评价.....	(89)
§ 4 · 1	无效测量的判断——可疑值的舍弃 (89)	
§ 4 · 2	显著性检验.....	(107)
§ 4 · 3	平均值与标准值的比较.....	(110)
§ 4 · 4	两个平均值的比较.....	(114)
§ 4 · 5	多个平均值的比较.....	(126)
§ 4 · 6	配对比较试验数据的判断.....	(129)
复习题.....		(133)
练习题.....		(135)
第五章	方差分析.....	(138)
§ 5 · 1	方差分析的基础.....	(138)
§ 5 · 2	单因素方差分析.....	(139)
§ 5 · 3	双因素方差分析.....	(145)
§ 5 · 4	其它方差分析概要.....	(154)
复习题.....		(155)
练习题.....		(155)
第六章	一元线性回归——标准曲线的绘制及其精密度的判断.....	(157)
§ 6 · 1	回归分析.....	(158)
§ 6 · 2	一元线性回归方程的建立.....	(162)
§ 6 · 3	一元线性回归分析的应用.....	(173)

§ 6 · 4. 回归方程的显著性检验·····	(180)
§ 6 · 5 测定结果的标准偏差和置信界限·····	(190)
复习题·····	(203)
练习题·····	(204)
第七章 正交试验法 ·····	(206)
§ 7 · 1 试验设计的意义和作用·····	(206)
§ 7 · 2 正交试验法的特点和正交表·····	(209)
§ 7 · 3 正交试验法的方法和步骤·····	(214)
复习题·····	(224)
练习题·····	(225)
第八章 化验工作的质量管理 ·····	(226)
§ 8 · 1 化验工作的质量管理的内容和作用·····	(226)
§ 8 · 2 质量控制(评价)图·····	(228)
§ 8 · 3 平均值的质量控制图及其应用·····	(229)
§ 8 · 4 极差的质量控制图及其应用·····	(235)
复习题·····	(238)
练习题·····	(238)
参考文献 ·····	(240)
附 录 ·····	(242)
附录一 从极差计算标准偏差的系数C表·····	(242)
附录二 正态分布的双侧分位数表·····	(243)
附录三 t分布的双侧分位数表·····	(247)
附录四 由极差计算平均值的置信限的	

	A值表.....	(251)
附录五	极差检验法舍弃无效测量的 t_1 的 临界值表.....	(253)
附录六	格拉布斯 (Grubbs) 检验法G的 临界值表.....	(254)
附录七	改良t检验法L的临界值表.....	(255)
附录八	F检验法F的临界值表.....	(256)
附录九	代用t检验法M的临界值表.....	(266)
附录十	柯奇拉 (Cochrane) 检验法 G_{max} 的 临界值表.....	(269)
附录十一	常用正交表.....	(271)
附录十二	相关系数检验的临界值表.....	(283)
附录十三	平均值质量评价因子 A_2 值表.....	(283)
附录十四	极差质量评价下限因子 D_3 值表...	(284)
附录十五	极差质量评价上限因子 D_4 值表...	(284)

第一章 基础知识

§ 1 · 1 数据处理与数理统计方法

分析化学既是一门理论性的学科，又是一种应用科学技术。作为应用科学技术，分析化学的用途十分广泛，在工业、农业、地质、冶金、环境监测等部门，以及许多科学研究工作中，都起着重要的、甚至是必不可少的作用，被誉为“眼睛”。分析化学工作者既有研究分析化学理论和改进分析方法的任務，又有从事测定物质的化学组成的实际工作。在完成上述任务和工作的过程中，都必须进行实验，因此也都必须进行数据处理工作。在分析化学中，通过实验只能得到一些数据（例如容量分析得到的是毫升数，重量分析法得到的是克数，吸光光度法和原子吸收光谱分析法得到的是吸光度 A 的数值，离子选择性电极法得到的是毫伏值或 PX 值等数据），应用这些数据进行适当的数学运算才能得到分析结果。为了获得准确可靠的分析结果，或要判断测定结果的可靠性，或者对分析方法进行考核和评价，就必须进行多次实验，然后应用这些实验的数据进行适当的数学运算才能得出切合实际的结果或结论。从实验数据出发，到获得分析结果或结论的数学运算过程，称为“数据处理”。

数据处理必需应用概率论与数理统计方法来完成。

概率论是数学学科中的一个分支，它是研究随机现象中

有关事件的规律性的学科。数理统计方法（简称数理统计）是概率论的具体应用，它的任务之一是根据实际观测到的随机试验结果（在分析化学中，实验测得的数据都是属于随机事件），对有关事件的概率或随机变量的分布、数字特征做出估计或推测。所以概率论和数理统计是密切联系相关的，常统称为“概率论和数理统计法”或“概率统计学”。概率统计学已经具有系统的理论，并广泛地应用于科学研究，工程技术，政治和经济领域中。

为什么分析化学试验（测量）数据必需应用数理统计方法来处理呢？这是因为分析化学实验所得到的数据都不是绝对准确的，都含有随机性质的偶然误差。例如，即使是同一个工作者，用同一种分析方法，在相同的条件下，对同一种样品进行多次重复测定，测定结果也不会是完全相同的。偶然误差（关于它的性质和产生原因，在下面讨论）是由于多种随机因素所引起的随机事件，其数值是不可测的。虽然偶然误差是随机性的、不确定的，但是它仍然是有规律性的。而概率论和数理统计方法正是研究和处理随机事件的属性和规律的科学方法，所以分析化学实验的数据，应该和必需运用数理统计方法来进行处理。只有这样，我们才能“透过”偶然误差，从实验数据中获得可靠的结果和结论，才能收到多方面的效果：

- （1）判断应舍弃的错误数据或结果；
- （2）得到最精确的分析结果；
- （3）判断分析结果的可靠程度；
- （4）评价分析方法、分析工作者的质量；
- （5）评价影响分析结果的各种因素的影响；

(6) 从测定数据中得出最多的结论,

(7) 对测定工作进行质量管理,

(8) 使实验工作最经济。

但是, 也应该指出, 数理统计方法是不能改变实验或测定工作本身的质量的, 只能是在已有的测定数据的基础上, 发挥上述作用。

由此可见, 对一个分析化学家或分析化学工作者来说, 如果没有掌握必要的概率论知识和数理统计方法, 不会正确地进行数据处理, 那么即使他的实验工作做得很出色, 也不能得到全面、正确的结果和结论。相反, 正确的数据处理工作, 往往可以发现测定工作中的缺陷或不足之处, 指出改进的途径。

§ 1 · 2 有效数字及其运算规则

在分析化学中, 进行数据处理, 自始至终都会遇到有效数字及其运算。

在日常生活和科学实验中, 我们不可避免地经常地要用到数字。在我们应用数字的时候, 可能有两种情况: 一种情况是数字是表示不随时间、地点和其他条件而改变的一个“绝对数值”, 例如查点物体的件数, 数一个字的笔划, 或者数一句话的字数等。例如“振兴中华”这句话, 无论谁来数、什么时候数、在什么地方和怎样数, 都只能得到同一结论——它是由四个字组成的。这种情况下, 数字是表示绝对数值的, 也就是通常所说的数目。另一种情况, 数字所表示的不是绝对数值, 而表示的是包含着不可靠部分或可疑值的

“非绝对数值”。一切表示测量结果的数字都是非绝对数值，因为无论测量的精确度如何高，所得到的结果都不可能是“绝对”准确的。在分析化学中，天平称量的结果、滴定管的读数、吸光度值读数、电位毫伏值的读数等等，一切表示测量结果的数字都是包含有可疑值的非绝对数值。例如，某次滴定管读数为20.15毫升，那么最后一位数字5是可疑的。因为一般滴定管的分度为0.1毫升，我们读为20.15毫升，是在20.1毫升与20.2毫升两个数字之间内插了一个0.05毫升，这是因为我们认为读数更接近20.15毫升而不是更接近20.14毫升或20.16毫升。显然这种“更接近”并非绝对正确，所以最后这位数字5是可疑值，非绝对数值的数字都是带有误差的，误差的大小与许多因素有关，这在后面要作讨论。

非绝对数值的数字，虽然是包含有可疑值的和有误差的，但这不是说这种数字是完全不可靠的、可以任意确定的，或者它的误差是可以任意大的。正因为如此，所以对于这类数字的值的确定，以及对这类数字的运算都有不同于绝对数值的特殊规定和规则，这就是有效数字及其运算规则。

有效数字是指一个数中的任何一个表示数的大小的数字，而不是用来指示小数点位置的数字。例如前面讲的滴定管读数是20.15毫升，它的意义为十位数上为2，个位数上为零，十分位上为1，百分位上为5，所以这四个数字都是有效数字，从而20.15毫升的有效数字的位数为4。又例如用分析天平称量的结果为1.3020克，这五个数字都是有效数字，它是五位有效数。值得注意的是数字“0”，它可以是有效数字，也可以不是有效数字。在上面两个例子中的0都