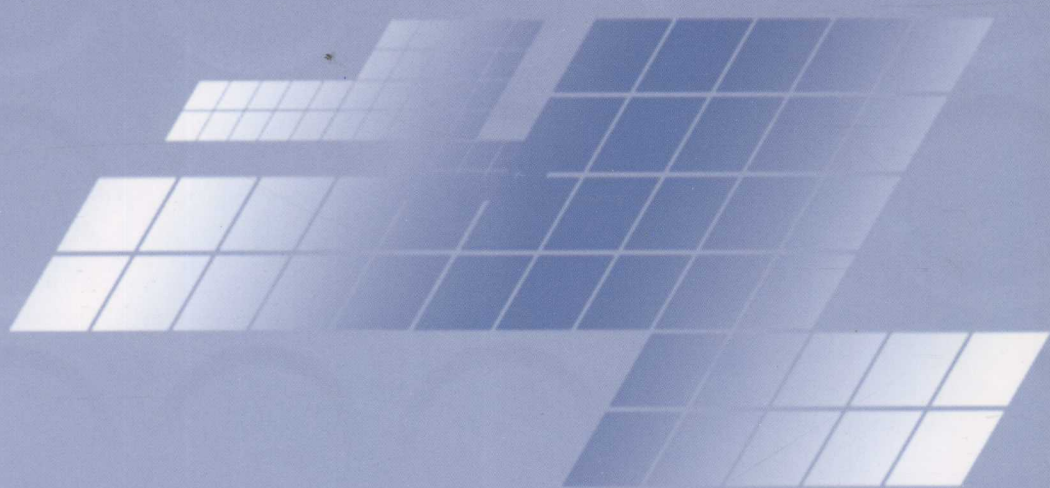




ZHILIANG JIANYAN BUQUEDINGDU
YU YINGYONG SHULI TONGJI

质量检验不确定度 与应用数理统计

何永政 编著



中国计量出版社

CHINA METROLOGY PUBLISHING HOUSE

质量检验

不确定度与应用数理统计

何永政 编著



中国计量出版社

图书在版编目(CIP)数据

质量检验不确定度与应用数理统计/何永政编著. —北京:中国计量出版社,2009.6

ISBN 978 - 7 - 5026 - 3035 - 5

I. 质… II. 何… III. ①测量 - 不确定度②数理统计 - 应用 - 测量 - 不确定度
IV. TB9

中国版本图书馆 CIP 数据核字(2009)第 081362 号

内 容 提 要

本书主要介绍质量检验不确定度与应用数理统计的相关知识。全书共 8 章,主要包括试验误差基本知识、测量不确定度、试验方法精密度、标准样品、随机变量及概率分布、随机抽样与统计推断、方差分析和多重比较以及回归分析等内容。书中以大量的分析实例详细举例说明质量检验中的不确定度和数理统计问题的处理方法。

本书可供从事质量检验与质量管理工作人员学习使用,亦可供工矿企业实验室技术人员及大专院校相关专业师生参考。

中国计量出版社 出版

地 址 北京和平里西街甲 2 号(邮编 100013)
电 话 (010)64275360
网 址 <http://www.zgjl.com.cn>
发 行 新华书店北京发行所发行
印 刷 北京市媛明印刷厂印刷
开 本 787mm × 1092mm 1/16
印 张 19.25
字 数 475 千字
版 次 2009 年 7 月第 1 版 2009 年 7 月第 1 次印刷
印 数 1—2 000
定 价 48.00 元

如有印装质量问题,请与本社联系调换
版权所有 侵权必究

前 言

本书共 8 章,其内容分为质量检验不确定度与应用数理统计两部分。第一章至第四章为第一部分:首先介绍试验误差基本知识,在此基础上介绍测量不确定度及其评定方法,以及试验方法精密度的确定方法及应用,分析了它与测量不确定度的区别与联系;同时介绍了标准样品的特性及应用。这些内容都与测量不确定度有关。第五章至第八章为第二部分:首先介绍随机变量与概率分布的基本知识,在此基础上介绍随机抽样与统计推断,方差分析与多重比较,回归分析等。这些属于质量检验工作常用的数理统计方面的内容。

本书的主要特点:首先是内容比较全面。如试验误差、测量不确定度等都有专门著作,但本书将试验误差、测量不确定度、试验方法精密度、标准样品等,既相互独立又具有内在联系的内容集于一体,不但方便读者的使用,也利于对有关内容内在联系的掌握。其次是深入浅出。测量不确定度、数理统计主要涉及一些比较专门的数学知识,如概率论、广义积分、矩阵运算、特殊函数等内容。

本书既尽量避免一些复杂的数学推导,又以比较通俗的方法分析其原理,说明有关计算公式的来源,明确各种统计方法的使用条件,力求做到知其然亦知其所以然。最后是实例较多。本书以质量检验工作中经常遇到的一些问题,作为实例进行分析,希望能起到依例触类旁通的作用。此外,本书也总结了作者在长期从事质量检验与质量管理工作中的一些经验体会。

本书可供从事质量检验与质量管理工作人员以及工矿企业实验室技术人员学习使用,也可作为技校、大专院校教学的参考读物。

由于作者学识所限,本书缺点错误在所难免,欢迎读者批评指正。

编 者

2009 年 6 月



目 录

第一章 试验误差基本知识	/ 1
第一节 试验误差定义	/ 1
第二节 试验误差的来源	/ 5
第三节 误差的分类与试验结果的评价	/ 6
第四节 异常值的检验与剔除	/ 8
第五节 系统误差和随机误差的检查与处理	/ 11
第六节 有效数字与数值修约规则	/ 12
第七节 提高测试技术水平的几种途径	/ 15
第二章 测量不确定度	/ 19
第一节 概述	/ 19
第二节 直接测量不确定度 A 类评定	/ 21
第三节 间接测量不确定度 A 类评定	/ 25
第四节 B 类不确定度评定	/ 28
第五节 不确定度分量的合成	/ 34
第六节 不确定度评定的程序	/ 35
第三章 试验方法精密度	/ 40
第一节 试验方法精密度及其确定方法	/ 40
第二节 精密度试验统计分析实例	/ 52
第三节 精密度数据的应用	/ 58
第四节 标准测量方法偏倚和实验室偏倚的估计	/ 66
第四章 标准样品	/ 74
第一节 标准样品的发展概况	/ 74
第二节 标准样品及其特性	/ 77
第三节 标准样品的定值	/ 83
第四节 标准样品及其作用	/ 88

第五节 使用标准样品校准试验结果 / 93

第五章 随机变量及概率分布 / 98

第一节 事件与概率初步知识 / 98

第二节 随机变量及其概率分布 / 107

第三节 随机变量统计特征数 / 112

第四节 正态分布 / 115

第六章 随机抽样与统计推断 / 121

第一节 随机抽样 / 121

第二节 样本特性 / 124

第三节 参数估计 / 133

第四节 统计假设检验 / 144

第五节 样本容量的确定 / 162

第七章 方差分析和多重比较 / 170

第一节 单因素试验的方差分析 / 170

第二节 双因素试验的方差分析 / 180

第三节 试验结果的多重比较 / 197

第八章 回归分析 / 217

第一节 一元线性回归 / 217

第二节 一元非线性回归 / 229

第三节 二元线性回归 / 236

第四节 多元线性回归 / 248

第五节 多项式回归 / 253

第六节 Excel 在统计分析计算中的应用 / 258

附表: / 272

附表1 格拉布斯(Grubbs)检验临界值表 / 272

附表2 狄克逊检验临界值表(双侧检验) / 273

附表3 t 分布表 / 273

附表4 相关系数检验表 / 274

附表5 科克伦(Cochran)检验法的临界值表 / 275

附表6 χ^2 分布表 / 277

附表 7	F 分布表	/	278
附表 8	显著性水平为 1% 时曼德尔 h 和 k 统计量的临界值	/	283
附表 9	正态分布表	/	285
附表 10	泊松分布数值表	/	286
附表 11	两总体秩和检验临界值表	/	288
附表 12	参照处理结果为已知值的比较用表—— t 分布分位数表	/	289
附表 13	k 种处理结果之间的两两比较用表—— q 表	/	290
附表 14	参照处理与 k 种处理的两两比较用表—— d 表	/	292
附表 15	若干处理的平均结果之间的比较用表—— s 表	/	294
附表 16	无对比方案的平均数比较试验临界系数 $t_{r,\alpha}$ 表	/	296
参考文献 / 297			

第一章 试验误差基本知识

第一节 试验误差定义

一、引言

经验表明,同一个样品,不同的实验室测试结果不一定相同;同一实验室中,不同的试验员测试结果也不一定相同;甚至,同一试验员重复试验的测试结果也不一定相同。造成这种现象的原因是试验结果存在误差。试验结果存在误差,这是一种普遍现象,正因如此,人们总结出了一条误差公理:测量结果都具有误差,误差自始至终存在于一切科学实验和测量过程之中。

从事质量检验工作离不开各种测试、试验工作,而试验结果总伴随着误差,因此有必要了解一些试验误差的基本知识。

二、绝对误差

所谓误差就是某被测量的给出值(测量值、试验结果、标称值、仪表示值等)与其客观真值之差,这样定义的误差称为绝对误差,简称误差,即

$$(\text{绝对})\text{误差} = \text{给出值} - \text{真值} \quad (1-1)$$

如果给出值是由试验结果得到的,则有

$$\text{试验误差} = \text{试验结果} - \text{真值} \quad (1-2)$$

【例】1—1 用磅秤称物体的质量。若物体的真实质量为 85.0 kg,而称量的结果为 86.0 kg 或 84.5 kg,则称量误差分别为:

$$\text{误差} = \text{给出值} - \text{真值} = 86.0 \text{ kg} - 85.0 \text{ kg} = 1.0 \text{ kg};$$

$$\text{误差} = \text{给出值} - \text{真值} = 84.5 \text{ kg} - 85.0 \text{ kg} = -0.5 \text{ kg}。$$

绝对误差是给出值与真值之差,它不是误差的绝对值,它有正有负。当给出值大于真值时,误差为正,表示给出结果偏大;当给出值小于真值时,误差为负,表示给出结果偏小。

真值有其客观存在,所谓试验误差是试验结果存在的误差,而不是真值存在的误差。

真值往往是未知的,正因为未知才需要测量,才需要做试验。真值未知,由式(1-1)可看出误差也是未知的。但是,在实际工作中,误差可以被估计,给出误差的估计范围,如极限误差等。另外,在某些情况下真值是可知。

(1) 理论真值

如平面三角形三内角之和为 180° ,一个圆的圆心角为 360° 等。

(2) 计量学约定真值

①长度单位——m(米),光在真空中于 $1/299792458$ 秒的时间间隔内行程的长度;

②质量单位——kg(千克),保存在法国巴黎国际计量局的铂-铱合金圆柱体(国际千克

原器)的质量;

③时间单位——s(秒),铯-133原子处于特定的状态(原子基态的两个超精细能级之间的跃迁)时,辐射出9192631770个周期的电磁波所持续的时间;

④电流强度单位——A[安(培)],在真空中两根长而细的平行导线,其间相距1 m,由一恒定的电流通过,如果两导线之间产生的力,在每米长度的导线上为 2×10^{-7} N时,其电流即为1 A。

⑤热力学温度单位——K[开(尔文)],水三相点热力学温度的 $1/273.16$ 为1 K。1 K的温度间隔等于 1°C 的温度间隔。

⑥发光强度单位——cd[坎(德拉)],一个频率为 540×10^{12} Hz的单色辐射光源,如果在给定方向上的辐射强度为 $(1/683) \cdot \text{W/sr}$ (瓦特每球面度)时,则光源在该方向上的发光强度为1 cd。

⑦物质的量单位——mol[摩(尔)],是一系统的物质的量,该系统中所包含的基本单元数与0.012 kg碳-12的原子数目相等时为1 mol。基本单元可以是原子、分子、离子、电子及其他粒子,或是这些粒子的组合。

(3)标准器相对真值

除以上两种真值知道的情况外,在实际工作中,人们经常将标准器相对真值作为真值看待。若高一级标准器的误差与低一级标准器或普通计量仪器的误差相比,为其 $1/5$ (或 $1/3 \sim 1/20$)时,则前者可作为后者的相对真值。比如计量检定的结果就是作为相对真值来使用的。标准样品的标准值也是作为相对真值来使用的。

三、相对误差

一个真值为100 g的砝码,称量结果为101 g,则误差为1 g。一个真值为1000 g的砝码,称量结果为1001 g,误差也是1 g。虽然二者的误差都是1 g,但二者称量的精度是不同的。显然,后者称量的精度高于前者。为了更好地描述测量的精度,我们引入相对误差的定义:

$$\text{相对误差} = \text{误差} \div \text{真值} \quad (1-3)$$

相对误差往往用百分数表示。如上述的例子:

$$\text{相对误差} = 1 \text{ g} / 100 \text{ g} = 0.01 = 1\%; \text{相对误差} = 1 \text{ g} / 1000 \text{ g} = 0.001 = 0.1\%。$$

在实际工作中,我们常用相对误差来表示测试结果的精度。相对误差越小,测试结果的精度越高。

因为真值往往未知,所以也难以用式(1-3)计算相对误差。但当误差较小时有:

$$\text{相对误差} \approx \text{误差} \div \text{给出值} \quad (1-4)$$

【例】1—2 用钢尺测量某一线段的长为80 mm,误差0.5 mm;用量角器测量某一角度为 90° ,误差是 $30'$,试比较两者测量精度的高低。

解:前者

$$\text{相对误差} = \text{误差} \div \text{给出值} = 0.5 \text{ mm} / 80 \text{ mm} = 0.62\%$$

后者

$$\text{相对误差} = \text{误差} \div \text{给出值} = 30' / 90^{\circ} = 0.56\%$$

因此,后者的测量精度高于前者。

本例中,前者的误差是0.5 mm,后者的误差是 $30'$,这是两个不同量纲的量,所以二者的绝对误差无法进行比较。但相对误差表示测试结果的精度,是无量纲的量,所以可以进行相互

比较。

一般说来,绝对误差仅用于描述测试结果误差的大小或真值相近的测试结果的误差的比较。当真值相差很大,或不同量纲的测试结果进行误差比较时,必须用相对误差进行比较。

四、引用误差

一个仪表有若干刻度,每个刻度都有相应的示值误差,根据式(1—4)知道,仪表的示值误差除以仪表的示值就是仪表在该刻度的相对误差。在不同的刻度,仪表的示值不同,示值误差一般也不同。因此,仪表在不同刻度的相对误差彼此不一样,即同一仪表在不同刻度的测试精度不同。那么,如何表示仪表的精度呢?为此,引入引用误差的概念,定义如下:

$$\text{引用误差} = \text{仪表示值误差} \div \text{仪表量程} \quad (1-5)$$

【例】1—3 某电压表刻度为0~10 V(即量程为10 V),在5 V和10 V处,计量检定值分别为4.995 V和10.006 V,求该表在5 V和10 V处的示值误差、示值相对误差和引用误差。

解:①在5 V处

$$\text{示值误差} = \text{示值} - \text{真值} = 5 - 4.995 = 0.005 (\text{V});$$

$$\text{示值相对误差} = \text{示值误差} \div \text{给出值(示值)} = 0.005 \text{ V} / 5 \text{ V} = 0.1\%;$$

$$\text{引用误差} = \text{示值误差} \div \text{仪表量程} = 0.005 \text{ V} / 10 \text{ V} = 0.05\%。$$

②在10 V处

$$\text{示值误差} = \text{示值} - \text{真值} = 10 - 10.006 = -0.006 (\text{V});$$

$$\text{示值相对误差} = \text{示值误差} \div \text{示值} = -0.006 \text{ V} / 10 \text{ V} = -0.06\%;$$

$$\text{引用误差} = \text{示值误差} \div \text{量程} = -0.006 \text{ V} / 10 \text{ V} = -0.06\%。$$

仪表在不同的刻度,引用误差一般不同。其中,引用误差最大的一个称为最大引用误差。根据式(1—5)有:

$$\text{最大引用误差} = \text{仪表最大示值误差} \div \text{仪表量程} \quad (1-6)$$

一个仪表最大引用误差百分数的分子称为该仪表的精度等级。如某仪表的精度为S级,表明该仪表的最大引用误差为S%。若该仪表的量程为 x_m ,根据式(1—6),该仪表的

$$\text{最大示值误差} = x_m \times S\% \quad (1-7)$$

【例】1—4 量程为100Ω的0.2级电阻表,求它的最大示值误差。

解:该电阻表是0.2级表,所以它的最大引用误差0.2%,由式(1—7)知:

$$\text{最大示值误差} = x_m \times S\% = 100\Omega \times 0.2\% = 0.2\Omega。$$

【例】1—5 某0.1级10 A电流表,经检定最大示值误差为8mA,问该仪表是否检定合格。

解:根据式(1—7),该电流表的

$$\text{最大示值误差} = x_m \times S\% = 10 \text{ A} \times 0.1\% = 10 \text{ mA}$$

而检定的最大示值误差=8 mA,它小于10 mA,因此检定合格。

知道仪表的量程 x_m 和最大引用误差S%后,根据式(1—7)知道,在任一测量点 x 处,仪表的示值误差 $\leq x_m \times S\%$,因此,在测量点 x 处的

$$\text{相对误差} \leq \frac{x_m \times S\%}{x} \quad (1-8)$$

因为 $x \leq x_m$,所以 $\frac{x_m}{x} \geq 1$,因此 x 越接近于 x_m , $\frac{x_m}{x}$ 就越小,从式(1—8)看出,此时 x 处的相

对误差也越小,测试精度越高;反之, x 越小于 x_m , $\frac{x_m}{x}$ 就越大,此时 x 测量点对应的相对误差也

越大,测试精度越低。这就是为什么在使用这类仪表时,尽可能在仪表的上限处附近或大于2/3 量程处进行测量的原因所在。

【例】1—6 某待测电流约 1.8 A,现有 0.5 级 20 A 和 1.0 级 2 A 电流表,问用哪个电流表测量较好?

解:根据式(1—8),用 0.5 级 20 A 电流表测量时, $x_m = 20 \text{ A}$, $S = 0.5$,测量的最大相对误差为 $\frac{x_m \cdot S\%}{x} = \frac{20 \times 0.5\%}{1.8} = 5.6\%$;用 1.0 级 2 A 电流表时, $x_m = 2 \text{ A}$, $S = 1.0$,测量的最大相对误差为 $\frac{x_m \cdot S\%}{x} = \frac{2 \times 1.0\%}{1.8} = 1.1\%$ 。因此,用 1.0 级 2A 电流表测量较好。

此例说明,选用测量仪器时,不仅要看仪表的精度等级,还要选择适当的量程。量程选择适当,低精度等级仪表测量结果的精度,可能比量程选择不当的高精度等级仪表测量结果的精度高。

五、应用实例

上面我们介绍了绝对误差、相对误差、引用误差和测试精度的概念,这些概念对于试验误差的分析是很有意义的,下面举几个例子。

【例】1—7 在计量仪器的性能指标中,往往有感量或最小分度值这个指标。根据这个指标,我们可以分析它的测试精度。如用万分之一天平(感量为 0.1 mg)称量 5 mg,10 mg,20 mg 的试样,问其称量精度是多少。

解:我们可近似地将感量(最小分度值)的半个单位看作测试的最大示值误差。根据式(1—4),相对误差分别为:

$0.05 \text{ mg}/5 \text{ mg} = 1\%$, $0.05 \text{ mg}/10 \text{ mg} = 0.5\%$, $0.05 \text{ mg}/20 \text{ mg} = 0.25\%$ 。称量精度分别为 1%,0.5% 和 0.25%。

【例】1—8 要求称量误差小于 1%,用感量为 0.1 mg 的天平称量,试样量至少需要多少?

解:感量 0.1 mg 的天平称量误差 $\leq 0.05 \text{ mg}$,要求称量误差小于 1% 也就是要求 $\frac{\text{称量误差}}{\text{试样量}} \leq 1\%$,即试样量 $\geq \frac{\text{称量误差}}{1\%} = \frac{0.05 \text{ mg}}{0.01} = 5 \text{ mg}$ 。

这就是根据仪器精度和试验精度要求,如何确定试样量的问题。

反过来,如果试样量限定,又要求试验精度时,必须选择仪器的精度。

【例】1—9 如试样量限定为 1 mg,要求称量误差小于 1%,天平的感量应选多大?

解:同样,由题意有 $\frac{\text{称量误差}}{\text{试样量}} \leq 1\%$,即:

$$\text{称量误差} \leq \text{试样量} \times 1\% = 1 \text{ mg} \times 0.01 = 0.01 \text{ mg}$$

因此,要选择感量为 0.01 mg 的天平,此时称量误差 $\leq 0.005 \text{ mg}$,满足称量误差 $\leq 0.01 \text{ mg}$ 的要求。

【例】1—10 若某产品的断裂比强度为 30.0 cN/tex,试验结果要求保留一位小数,问测试精度是多少?

解:试验结果保留一位小数,即要求试验结果的整数位是准确的,它要求试验结果的绝对误差不超过 0.5 cN/tex,相对误差不超过 $0.5/30.0 = 0.017 \approx 2\%$,即测试精度为 2%。

【例】1—11 在废水悬浮固体的测定方法中,使用万分之一分析天平(分度值 0.1 mg),试

验方法要求量取均匀适量水样(使悬浮物大于 2.5 mg)。试分析称量精度。

解:因为水样的悬浮物大于 2.5 mg,而分析天平的分度值 0.1 mg,读数误差 ≤ 0.05 mg,所以,称量相对误差 $\leq \frac{0.05}{2.5} = 0.02 = 2\%$ 。

第二节 试验误差的来源

误差尽管是不可避免的,但随着科学技术水平的提高以及人们的经验、技巧和专门知识的丰富,误差可以被控制得越来越小。我们要求试验结果准确,就是要求测试结果的误差小。为此,我们应该了解误差的来源。

误差的来源主要有以下五个方面:

1. 装置误差。它指的是试验工作中所用的计量器具和辅助设备本身对试验结果所引起的误差。如纤维切断器切断长度偏长或偏短,强力机隔距长度、显微镜放大倍数不准,电测器刻度误差等都会引起试验误差。为克服或减少装置误差,计量器具必须定期检定。

2. 环境误差。由于被测量本身的大小往往随环境条件的变化而变化,又由于测试用的仪器设备的性能往往也与环境条件有关;甚至,环境条件对操作人员的心理、生理也有影响,从而影响其测试技术水平的发挥。由于这一切造成了因环境条件而引起的试验误差。为了减少环境误差对试验结果的影响,试验方法标准一般都对试验的环境条件提出一定的要求,测试时,环境条件必须满足标准提出的要求。

3. 人员误差。这是指由于不同的操作人员在操作手法、固有习惯和生理分辨能力等方面的差异而引起的误差。特别是有的试验方法,手工操作环节比较多,对操作手法要求比较高,这时人员误差往往在总误差中占有较大比重,必须引起足够的注意。为了减少人员误差,操作人员必须经过培训,要持证上岗操作。操作人员要勤学苦练,努力提高自己的测试技术水平。此外,要经常用标准样品来校准自己的测试结果,积极参加全国比对试验等。

4. 方法误差,也称理论误差。这是由于测试方法或计算方法不完善而引起的误差。比如棉花含水率或回潮率的试验,样品到底用箱内称重还是箱外称重,是箱外热称还是箱外冷称,由于称重方法的不一样,就会造成很大的误差。又如使用电测器测试棉花的回潮率,对于回潮率的温差补偿,由于理论研究尚不充分,因此,补偿的结果存在误差。因此,为了减小方法误差,要对试验方法进行研究,使之不断完善,或者用一种新的,精密度更高的方法来取代旧方法。

5. 抽样误差。这实质上是由于试验对象的变动而引起的误差。显然,对抽样单元来说,如果试验对象是绝对均匀的,则不存在抽样误差。但是,有些材料是不可混匀的物质,因此,抽取的样品与样品之间就存在着差异。样品的制备及取样数量的不同,都可能使抽取的样品失之代表性,这会引引起试验结果的较大误差。如何减小抽样误差,对于质检工作来说,是非常重要的问题。为此,除了要严格按照有关标准要求的抽样方法和抽样数量进行抽样和样品制备外,而且往往要求做重复试验,而且重复试验的结果必须符合精密度的要求。

第三节 误差的分类与试验结果的评价

一、误差分类

为了便于对误差进行分析和研究,需要对误差进行分类。按误差的性质,人们通常将误差分为随机误差、系统误差和粗大误差三类。

1. 随机误差

随机误差是指:“测量结果与在重复性条件下,对同一被测量进行无限多次测量所得结果的平均值之差。”即在实际测试条件下,多次测试同一量时,误差的绝对值和符号的变化,时大时小、时正时负,以不可预定的方式变化着的误差。随机误差是由于人们无法控制的随机因素的影响造成的,因此,试验结果的随机误差是不可避免的。这正是误差公理指出的“试验结果都具有误差”的根本原因所在。

2. 系统误差

系统误差是指:“在重复性条件下,对同一被测量进行无限多次测量所得结果的平均值与被测量的真值之差。”即由于某种不可忽视的原因而造成的固定不变的或按某种规律变化着的误差。含有系统误差的试验结果显著偏离被测量的真值,试验结果不正确。因此,在试验工作中,应尽量避免系统误差。一旦发现试验结果含有系统误差,必须予以修正。

3. 粗大误差

粗大误差是指超出在规定条件下预期的误差。它往往是由于读错、抄错数据,拿错样品,计算错误或试验条件的某种意外变故而引起的。粗大误差会明显地歪曲试验结果。含有粗大误差的试验数据称为异常值,要分析情况,予以剔除。

二、试验结果的评价

根据试验误差的情况,可以评价试验结果的好坏。它可以用以下几个量来描述。

1. 精密度

精密度是在规定的条件下,相互独立的测试结果之间的一致程度,如重复试验数据之间的一致程度。精密度仅依赖于试验的随机误差,它反映试验结果中随机误差大小的程度。随机误差越小,试验结果的精密度越好。

2. 正确度

正确度是由大量测试结果得到的平均数与接受参照值之间的一致程度。

所谓接受参照值是指用作比较的经协商同意的标准值,它来自于:①基于科学原理的理论值或确定值;②基于一些国家或国际组织的实验工作的指定值或认证值;③基于科学或工程组织赞助下合作实验工作中的同意值或认证值;④当以上不能获得时,则用被测量的期望,即规定测量总体的均值。因此,接受参照值实际上就是被测量的真值、约定真值、标准值或者是其总体均值(期望)。所以,正确度是指测试结果的平均值与真值或总体均值间的一致程度,正确度是表示试验结果中系统误差大小的程度。系统误差越小,试验结果的正确度越好。

3. 准确度(精确度)

准确度是测试结果与接受参照值间的一致程度。

测试结果可以是几个观测值计算的结果,也可以是观测值本身。所以,准确度这个术语用于一组测试结果时,由随机误差含量和系统误差含量即偏倚分量组成。所以,准确度是试验结果中随机误差和系统误差的综合反映。

准确度是一个定性的概念,不宜将其量化。换言之,我们可以说准确度高、准确度等级或准确度符合 $\times \times$ 标准等,而不宜将准确度与数字直接相连,例如:准确度为 0.25%、16 mg 或 ± 16 mg 等。

由于试验误差的存在,多次重复试验的结果,其数据的分布会呈现下面不同的情况,如图 1—1 所示(图中, + 表示真值位置, · 表示每次试验值)。

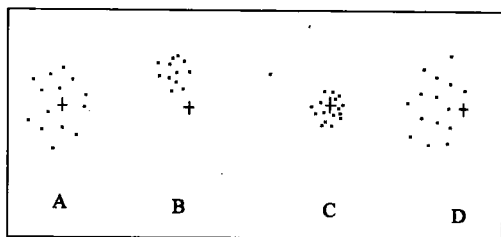


图 1—1 试验数据分布情况

图 1—1 中, A 种情况为重复试验结果的数据较为离散,但它们围绕真值分布,平均值接近真值; B 种情况为重复试验结果的数据分布较为集中,但其平均值偏离真值; C 种情况为重复试验结果的数据较集中,而且其平均值接近真值; D 种情况为重复试验数据较离散,且平均值偏离真值。因此,可以说:

A: 系统误差小, 随机误差大; 正确度好, 精密度差。

B: 系统误差大, 随机误差小; 正确度差, 精密度好。

C: 系统误差和随机误差都小。这种试验结果是最理想的, 准确度好。

D: 随机误差和系统误差都大, 正确度差, 精密度也不好, 准确度差。

三、试验数据处理

我们进行试验时, 试验结果不可避免地含有随机误差, 往往也含有系统误差甚至粗差(粗大误差)。

含有粗差的数据属异常数据, 它会严重歪曲试验结果。因此, 在试验数据的处理中, 必须首先予以剔除。

从试验误差的定义知道:

$$\begin{aligned} \text{误差} &= \text{测量结果} - \text{真值} = (\text{测量结果} - \text{总体均值}) + (\text{总体均值} - \text{真值}) \\ &= \text{随机误差} + \text{系统误差} \end{aligned}$$

另外, 从图 1—1 的 A, B, C, D 四种情况可以看出, 当试验结果含有系统误差时, 试验结果的正确度差, 试验结果不准。因此, 对系统误差的发现与修正是试验工作中应十分重视的问题。而且, 在给出最后的试验结果时, 应消除系统误差成分, 否则, 给出的最后结果不正确。

对一组试验数据, 当剔除了异常值, 消除了系统误差后, 数据还包含有随机误差。这时, 还得检查数据的随机误差是否过大。因为, 当试验数据的随机误差过大, 而重复试验的次数又不多时, 试验结果是不可靠的。也就是说, 这样的结果可信度是不够的, 不能贸然地下试验结论。为此, 试验数据还得进行精密度检查, 只有当试验的精密度符合要求时, 也就是试验的随机误差控制在允许的范围内时, 才能给出试验的最终结果, 否则应增加重复试验的数据个数。

总之, 对一组试验数据, 在进行数据的处理时, 一般应按下述步骤进行: 首先, 检查并剔除

异常数据。然后,检查并修正系统误差。最后,对重复试验数据进行精密度检查,符合精密度要求后,再给出最终试验结果。为此,我们将进一步介绍异常值的检验与剔除,系统误差的发现与修正以及精密度检查等方面的内容。

第四节 异常值的检验与剔除

一、概述

要在试验数据中消除异常值,无非是从两个方面入手:事先防范,事后把关。

事先防范,首先在整个试验过程中要认真。不要拿错样品,不要读错、记错、抄错或算错数据。二要有制度保证,如对试验报告要有三级审核制度等。三是在试验过程中,当试验方法标准中明令要废弃的数据出现时,该数据必须废弃。四是在试验过程中,发现测试条件异常变故的情况,要予以记录,以便供数据分析时参考。

事后把关就是对一组数据首先要进行异常值的检验和剔除。

异常值的检验方法与法则很多,有的还制定了国家标准,如 GB/T 4883《数据的统计处理和解释 正态样本异常值的判断与处理》,对异常值的检验与处理有较详细的论述。下面对正态分布试验数据的异常值检验方法,格拉布斯(Grubbs)检验方法和狄克逊(Dixon)检验方法作一个简单介绍。

二、格拉布斯检验方法

设 n 个测试数据为 $x_1, x_2, \dots, x_n$, 检验步骤如下:

(1) 计算平均值和标准差

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1-9)$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1-10)$$

(2) 排序

将 n 个数据按从小到大的顺序重新排列如下: $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ 。

(3) 计算统计量 G_n

$$G_n = \frac{x_{(n)} - \bar{x}}{s} \text{ 和 } G_n = \frac{\bar{x} - x_{(1)}}{s} \quad (1-11)$$

中的较大者。

(4) 判断

从格拉布斯检验临界值表(见附表 1)中的“一个最大值或一个最小值”栏中,查显著性水平 $\alpha = 0.05$ 和 $\alpha = 0.01$ 时与 n 对应的格拉布斯检验临界值 $G_{0.05}(n)$ 和 $G_{0.01}(n)$, 并进行判断:

若统计量 $G_n \leq G_{0.05}(n)$, 则这 n 个数据无异常值;

$G_{0.05}(n) < G_n \leq G_{0.01}(n)$, 则对应的 $x_{(1)}$ 或 $x_{(n)}$ 为异常值;

$G_n > G_{0.01}(n)$, 则对应的 $x_{(1)}$ 或 $x_{(n)}$ 为高度异常值。

要注意的是,格拉布斯检验方法仅适用于在试验数据中发现一个异常值的情况。

三、狄克逊检验方法

狄克逊检验可用于在试验数据中发现多个异常值的情况。

设 n 个测试数据为 $x_1, x_2, \dots, x_n$, 检验步骤如下:

(1) 排序

将 n 个数据按从小到大的顺序重新排列如下: $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ 。

(2) 计算统计量

根据样本量 n 的大小, 计算相应的狄克逊检验统计量 D , 计算公式见表 1—1。

表 1—1 狄克逊检验统计量公式

样本量 n	统计量 D
3 ~ 7	$D_{10} = \frac{x_{(2)} - x_{(1)}}{x_{(n)} - x_{(1)}}$ 和 $D_{10} = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}}$ 中的较大者
8 ~ 10	$D_{11} = \frac{x_{(2)} - x_{(1)}}{x_{(n-1)} - x_{(1)}}$ 和 $D_{11} = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(2)}}$ 中的较大者
11 ~ 13	$D_{21} = \frac{x_{(3)} - x_{(1)}}{x_{(n-1)} - x_{(1)}}$ 和 $D_{21} = \frac{x_{(n)} - x_{(n-2)}}{x_{(n)} - x_{(2)}}$ 中的较大者
14 ~ 30	$D_{22} = \frac{x_{(3)} - x_{(1)}}{x_{(n-2)} - x_{(1)}}$ 和 $D_{22} = \frac{x_{(n)} - x_{(n-2)}}{x_{(n)} - x_{(3)}}$ 中的较大者

(3) 判断

从狄克逊检验临界值表(见附表 2)中, 查显著性水平 $\alpha = 0.05$ 和 $\alpha = 0.01$ 时与 n 对应的狄克逊检验临界值 $D_{0.05}(n)$ 和 $D_{0.01}(n)$, 并进行判断:

若统计量 $D \leq D_{0.05}(n)$, 则这 n 个数据无异常值;

$D_{0.05}(n) < D \leq D_{0.01}(n)$, 则对应的 $x_{(1)}$ 或 $x_{(n)}$ 为异常值;

$D > D_{0.01}(n)$, 则对应的 $x_{(1)}$ 或 $x_{(n)}$ 为高度异常值。

(4) 对剔除异常值后余下的数据继续进行狄克逊检验

如果数据中的最小值或最大值被剔除后, 应对余下的 $n-1$ 个数据继续进行检验, 直到不能继续检出异常值为止。

四、异常值的处理原则

上面分别介绍了正态样本异常值的格拉布斯检验法和狄克逊检验法。对于检出的异常值如何进行处理, 下面提出几个要点。

(1) 一组数据, 经格拉布斯检验法及狄克逊检验法检验后, 认为是异常值的数据, 首先要从技术上查明原因。若从技术上查明了原因(如确认抄错、算错数据等), 则将此异常值剔除。若不能从技术上查明原因, 则保留异常值, 剔除高度异常值。

(2) 对于一组试验数据, 格拉布斯检验和狄克逊检验都必须分别独立地进行, 两种检验结果分别记录在册。格拉布斯检验只进行一次。两种检验结论不一定相同。如果只

发现一个异常值,取格拉布斯检验所得的结论;如果发现有多于个异常值,则取狄克逊检验所得的结论。

(3)当样本分布的正态性假设不够充分时,有可能导致数据连续地剔除。因此,当出现连续剔除的情况时,对数据的剔除要特别慎重。

【例】1—12 某次试验,6个试样的重复试验结果如下:18.5,17.5,17.2,17.5,17.2,17.3。试检查并剔除异常值。

解:(1)用格拉布斯检验法

①计算平均值和标准差

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = (18.5 + 17.5 + \cdots + 17.3)/6 = 17.53$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{1}{6-1} \times [(18.5 - 17.53)^2 + (17.5 - 17.53)^2 + \cdots + (17.3 - 17.53)^2]} \\ = 0.4926$$

②将数据按从小到大顺序排列为:17.2 ≤ 17.2 ≤ 17.3 ≤ 17.5 ≤ 17.5 ≤ 18.5。

③计算统计量

$$G_n = \frac{x_{(n)} - \bar{x}}{s} = \frac{18.5 - 17.53}{0.4926} = 1.9691 \text{ 和 } G_n = \frac{\bar{x} - x_{(1)}}{s} = \frac{17.53 - 17.2}{0.4926} = 0.6699$$

取 $G_n = 1.9691$ 。

④判断。查附表1格拉布斯检验临界值表, $n=6$ 时, $G_{0.05}(6) = 1.887$, $G_{0.01}(6) = 1.973$ 。

因为, $G_{0.05}(6) = 1.887 < G_n = 1.9691 < G_{0.01}(6) = 1.973$,所以对应的 $x_{(n)} = 18.5$ 为异常值,但不是高度异常值。

(2)用狄克逊检验法

①将数据按从小到大顺序排列为:17.2 ≤ 17.2 ≤ 17.3 ≤ 17.5 ≤ 17.5 ≤ 18.5。

②计算统计量 D 。因为 $n=6$,故用 D_{10} :

$$D_{10} = \frac{x_{(2)} - x_{(1)}}{x_{(n)} - x_{(1)}} = \frac{17.2 - 17.2}{18.5 - 17.2} = 0$$

$$\text{和 } D_{10} = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}} = \frac{18.5 - 17.5}{18.5 - 17.2} = 0.7692$$

取 $D_{10} = 0.7692$ 。

③判断。查附表2狄克逊检验临界值表, $n=6$ 时, $D_{0.05}(6) = 0.628$, $D_{0.01}(6) = 0.740$ 。

因为, $D_{10} = 0.7692 > D_{0.01}(6) = 0.740$,所以对应的 $x_{(n)} = 18.5$ 为高度异常值。

④剔除高度异常值 18.5 后,对余下的 5 个数据继续进行狄克逊检验,结果未能继续检出异常值。

(3)异常值的处理

此组数据,经格拉布斯检验认为 18.5 为异常值,但不是高度异常值;经狄克逊检验认为 18.5 为高度异常值,但余下的数据不能继续检出异常值。

因为从技术上未能查出产生异常值的原因,因此,根据格拉布斯检验法结果,应保留异常值 18.5;根据狄克逊检验法结果,应剔除高度异常值 18.5。但因为只发现一个异常值,所以取格拉布斯检验法的结论,即保留异常值 18.5。