实用材料试验数据处理

赵从宝　编

经济管理出版社

实用材料试验数据处理

赵从宝　　编

经济管理出版社

(京)新登字 029 号

责任编辑：张汉亚

封面设计：赵利军

实用材料试验数据处理

赵从宝　编

出版：经济管理出版社

(北京西城区新街口红园胡同 8 号　邮政编码 100035)

发行：新华书店北京发行所总发行　各地新华书店经销

印刷：4229 印刷厂印刷

850×1168 毫米　32 开　　9 印张　220 千字

1992 年 4 月北京第一版　　1992 年 4 月 北京第一次印刷

印数：1～3000 册

IABN 7-80025-566-2／F·440

定价：4.80 元

内 容 提 要

本书主要介绍了材料试验数据处理及产品质量分析的方法。内容包括：数据整理，可疑值的取舍，正态分布的经验确定，威布尔分布的经验确定，经验公式的建立，内插与外推，数值的近似计算等。文字通俗，说理清楚，联系实际，举例丰富，方法步骤明确，并给出了用 BASIC 语言编制的计算机程序，便于自学和应用。

本书可作为冶金、航天、航空、机械、电力、交通、建筑、化工、石油等部门从事材料试验研究及产品质量分析等工作人员的工具书，也可作为大专院校有关材料、力学、机械及机械设计等专业师生的教学参考书。

前　言

在材料试验研究及产品质量分析中，数据是分析问题和解决问题所不可少的基本依据。但是，在大多数情况下，一堆未经加工的试验数据，往往是什么问题也说明不了的，必须运用数学的方法对试验数据加以整理、分析和研究，以尽可能简明的方式把试验数据中所包含的有用信息提炼出来，才能最大限度地发挥它们的作用。如果试验数据不能得到很好的利用，试验就会功亏一篑，造成人力、物力的浪费。上述对数据的加工，就是通常说的“数据筛”，一般需用到的数学包括概率论、数理统计、微积分、高等代数，可靠性数学等，这对许多数学根底较差的初学者来说，会觉得有一定的困难。针对这种情况，本书从实用出发，尽量本着通俗易懂的原则，紧密结合材料试验研究和产品质量分析中的实际问题，去介绍试验数据的处理方法，并给出了部分用 BASIC 语言编制的计算机程序，这在一定程度上减少了数学方法的抽象性，增加了实际感，便于初学者自学和入门，应用起来也较为方便。

如果本书的出版，能对材料有关部门的材料试验研究及产品质量分析工作起到一些积极的作用，也算是我对祖国的四化事业尽了微薄之力，这也是我所衷心期望的。

本书是我在 1986 年编写的《实验数据处理》讲义的基础上重新整理、修改充实而成的。由于本人学识浅薄，书中肯定存在不少缺点和错误，恳请读者批评指正。

赵从宝

1991 年 10 月于北京冶金部钢铁研究总院

目　　录

第一章　数据整理

试验是科学研究中的一种重要方法，由于人力、物力等原因的限制，研究工作者不可能对研究对象进行全面的分析，只能通过对部分试验结果进行研究，去作出结论。由于测量方法，测量工具、取样等原因都会使试验结果不十分精确，所以，我们必须对试验结果进行去伪存真，去粗取精，作出科学的整理。

第一节　基本概念

为了今后叙述的方便，我们先介绍一下概率论与数理统计中经常用到的基本概念和术语。

一、确定性现象和不确定性现象

自然界和人类社会中发生的现象是多种多样的。其中有一类现象，在一定的条件下，必然发生（或必然不发生）。如：一定的外力作用在试样上，必然出现变形，当试样所承受的载荷超过一定限值时，必然断裂。在标准大气压下，把水加热到 100℃，它就会沸腾。两物体带同性电荷不可能相互吸引，等等。这类现象称为确定性现象。也称必然现象。然而，在自然界和人类社会中也存在着另一类现象，它们在一定条件下可能发生，也可能不发生。如：对某种材料取 50 个试样在相同条件下进行高温持久

寿命试验，其寿命值有多种可能，具体是多少，事先不能确定。又如：在相同条件下，抛同一硬币，落地结果可能是国徽面朝上，也可能是数字面朝上，不论怎样控制抛的条件，在每次抛之前总是无法肯定抛的结果是什么。这类现象称为不确定性现象。人们经过长期实践并深入研究之后，发现不确定性现象在大量重复试验中，具有统计规律性，所以，又将之称为随机现象。概率论和数理统计就是研究和揭示随机现象统计规律性的一门数学学科。

二、随机试验与随机事件

我们遇到过许多试验，如：抛硬币时，观察国徽面和数字面出现的情况；对某一批金属材料取样测试它的抗拉强度，等等。这些试验具有以下共同的特性：

1. 可以在相同的条件下重复地进行。

2. 每次试验的可能结果不止一个，并且能事先明确试验的所有可能结果。

3. 进行一次试验之前不能确定哪一个结果会出现。

一般把具有上述三个特性的试验称为随机试验。

在随机试验中，对在一次试验中可能出现也可能不出现，而在大量重复试验中却以某种规律性出现的事情，称为随机试验的随机事件。换句话说，我们将随机试验的结果称作事件。随机事件属于不确定性现象。在试验中必然会发生的事情叫做必然事件，不可能发生的事情叫做不可能事件。不可能事件和必然事件属于确定性现象。但也可作为随机事件的特殊情况。

我们就是通过研究随机试验研究随机现象的。

三、频率与概率

随机事件在一次试验中是否发生，固然不能事先知道，但当

进行大量重复的试验时，可以发现事件的发生具有一定规律性。为了说明这种规律性是客观存在的，先介绍频率的概念。

1. 频率的定义

定义：在一定的条件下进行 n 次重复试验，如事件 A 出现了 m 次，则称

$$f(A) = \frac{m}{n} \tag{1-1}$$

为事件 A 在 n 次试验中出现的频率。通常称 m 为频数，n 为试验次数。历史上有许多人对抛硬币做过很多试验，如表（1-1）所示。从表中可以看出，不管什么人做此试验，当试验次数增加时，事件 A 出现的频率在常数 0.5 附近摆动，而且逐渐稳定于这个常数值。同样，常数 0.5 反映了事件 A 发生的规律性。

表 1-1

试验人	蒲　丰	皮尔逊	维　尼
m（正面向上次数）	2048	6019	14994
n（上抛次数）	4040	12000	30000
f（A）	0.5069	0.5016	0.4998

2. 概率的定义

概率是度量某一事件发生的可能性大小的量，它是随机事件的函数。在某些教科书中，概率的定义是：当试验条件可以重复，试验次数无限增加时，如果事件 A 发生的频率趋于一稳定的数值，则称这个数值为事件 A 的概率，记为 $P(A)$，简记为 p。

由概率的定义知：必然事件的概率为 1，不可能发生事件的概率为 0，一般随机事件其概率在 0 与 1 之间。若把某一事件发生的概率记为 p，则不发生的概率为 1-p。

四、随机变量

用来描述随机事件的量叫随机变量。随机变量具有两个特点：（1）试验之前只知道它取值的范围而不能预言它具体取什么值。（2）由于随机事件每一个试验结果的出现有一定的概率，所以，随机变量的取值也有一定的概率。如：材料的机械性能是随机变量。σ_s、σ_b、σ_{-1}、t_r 等，即使是同一钢号，同一炉号，在同一根坯料上相同部位取样，且热处理条件和试验条件相同的情况下，作多个试样的试验所得结果，也不会完全一样，数据会因材料和试验方法而表现出不同的分散程度。但是把这些试验数据进行概率统计后，可以从中发现它们具有一定的统计规律性。

随机变量是研究随机事件的有效工具，通过随机变量将各个事件联系起来就可以研究随机事件的全部结果。

五、母体、个体、子样和抽样

母体是指研究对象的全体。把组成母体的每个基本单元称为个体。例如：一车砖，一大批钢筋都是母体。而每块砖，每根钢筋等都是个体。子样是由母体的一部分个体组成。如：从一车砖中抽取 20 块检查，那么这 20 块砖就成为一个子样。又如：对某种材料在相同条件下进行持久寿命试验，测得一组 12 个试样的持久寿命如下：43、45、39、35、42、41、48.5、37、40.5、38.7、43.6、40（单位为：小时）。那么这 12 个试样的持久寿命就成为一个子样，而其中每一个试样的持久寿命则为一个个体。子样中所包含个体的数目称为子样大小，或叫子样容量。当以母体为研究对象时，往往由于母体所包含的个体很多，甚至无穷，以致不可能一一加以考察，怎么办呢？只好进行抽样观察，即从母体中抽取一部分个体来推断母体的性质。换句话说，就是通过对子样的研究来推断母体的性质。

抽样的方法很多，最常用的是简单随机抽样，这种方法要求

满足以下两条：一是在抽取子样时，母体中的每个个体都具有同等概率被抽取到。二是在抽取子样的每一个个体时，母体中的个体成分并不改变。

由此可见，这种抽样方法，实际是在做独立重复试验。由于子样是随机抽取的，所以可以认为子样是随机变量。

第二节　测量值与有效数字

在采用抽样的办法进行试验时，试验的结果表现为测量值。对于母体来说，这个测量值不仅受到抽样的影响，而且还受到测量者的感官和所用仪器的缺陷的影响，不可能得到绝对精确的数值。因此，有必要介绍一下试验中测量值的读取方法。

一、测量值

试验的测量值与通常数学上所说的数值在概念上是有区别的，如数据 2.3、2.30、2.300 等，在数学上被看作是同一个数值，但在表示试验的测量值时就不一样了。现在举个例子来说明这个问题。例如：测定某种材料的拉伸应力松弛曲线，材料在某一位移下承受的负荷随时间的变化是从试验机上的力值表盘上读取的。当力值表盘上的指针所指处无刻度时（如图 1-1 所示），此时，我们自然地将它读成离指针最近的那个数。即 4250N，而实际上，精确的读数值 a 并不正好等于 4250N，而只是满足下列不等式

$$4245<a<4255$$

也就是说，上述的 a 也可能是 4247、4248……，也可能是 4251、4252……等等，此时，我们说，试验的测量值 4250N 直

到最后一位都是有意义的，最后一位数 0 是经过舍或入之后得来的，至于其后面的数字就不清楚了。由此可见，所谓测量值，是在进行试验时，用测量仪器对某一物理量进行测量，从记录装置上读得的数字。测量值含有一定的误差，而数学上的数值是不考虑这一点的，这就是这两者之间的区别之处。

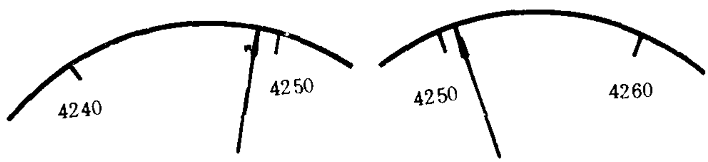

图 1-1　测量值的读取

二、有效数字

由上面讨论可知，任何测量值，总是表现为数字，而这些数值的最后一位为不精确数值，前面的数字都是精确的，也就是说，测量值只代表了测量的近似值。那么，测量值究竟取多少位数合适呢？有人认为，在一个数值中小数点之后的位数越多，便越精确，其实不然，因为小数点的位置仅与所用单位大小有关，再说，还与测量者的感官和所用仪器、测量方法、测量环境等有关。如：测量某试样的直径为 8mm 与 0.008m，其精确度完全一样。又如在用电位差计测量某一试验温度所对应的电势值为 32.24mv，而用高精度的数字多用表测量可得 32.2437mv，可见，由于两种测量仪器的精度不同，所得结果的精确度也不一样。因此，对测量值必须要求给出有效数字的位数。有效数字的定义是：表示一个数中的任何一个有意义的数字。现举例说明如下：

例 1-1：已知某测量值为 3563，则意味着其精确的读数 a 满足

$$3562.5 \leqslant a \leqslant 3563.5$$

也就是说，测量值 3563 的有效数字有四位。

例 1-2：已知某测量值为 34.5，则意味着其精确的读数值 a 应满足

$$34.45 \leqslant a \leqslant 34.55$$

也就是说，测量值 34.5 的有效数字有三位。

关于数字零，需要特别提一下，它可以是有效数字，也可以不是有效数字，主要由其在数据中的位置决定。

例 1-3：已知某测量值为 3.20，则意味着其精确的读数值 a 应满足

$$3.195 \leqslant a \leqslant 3.205$$

也就是说，测量值 3.20 的有效数字有三位。

例 1-4：已知某测量值为 0.00072，由于测量值中的前四个零只与所取的单位有关，而与测量的精确度无关。故前面的四个零均非有效数字。可见，这个测量值可写成 7.2×10^{-4}，则意味着其精确的读数值 a 应满足

$$7.15 \times 10^{-4} \leqslant a \leqslant 7.25 \times 10^{-4}$$

也就是说，测量值 0.00072 的有效数字有二位。

由上面的例子可以看出，有效数字是用来表示数的大小的，而不是用来指示小数点位置的。所以，在记录数值时应该记录的位数中，只允许最末一位数字是可疑的，其余各位数字必须都是精确的，而且读数应该反映出仪器的精度。若测量某一物理量为 10.90mm，若误记为 10.9mm，则所记测量值低于实际所能达到的精确度，这是不能允许的。

第三节　数据的修约与计算规则

我们知道，由试验测量得到的数值都表现为近似数，因此，在计算时也只能用近似数表示，这些近似数的正确取得，对保证测量结果的精确度有很大关系。本节将介绍在试验数据处理时对测量值的舍入和计算规则。

一、数据的修约规则

在试验数据处理时，首先确定有效数字的位数，当有效数字的位数确定之后，例如为 N 位，然后对 N+1 位后的数字进行修约，即舍入。

过去习惯用“四舍五入法”，其实，这样做有一个缺点，因为“四舍五入”遇 5 就进位，很容易使所得的数据系统偏高，无法消除由 5 本身引起的误差，如果有的 5 入，有的 5 舍，就可以使 5 本身引起的正负误差相消。“四舍六入五单双”的数字修约规则就解决了这个问题。“四舍六入五单双法”的具体内容是：四舍六入五考虑；五后非零则进一；五后皆零视奇偶，五前为奇则进一，五前为偶应舍去。

例 1−5：要求将下列数据均舍入成三位有效数字：

345.601→346　　（六入）

0.52343→0.523　（四舍）

9.37501→9.38　（五后非零则进一）

12.4500→12.4　（五前为偶应舍去）

12.7500→12.8　（五前为奇则进一）

二、近似数的计算规则