

# 可靠性试验及其统计分析

上 册

戴树森 费鹤良  
王玲玲 苏德清 编著  
白鹤翔 滕怀流

## 内 容 简 介

本书共十七章，分上下两册。上册介绍了概率和数理统计的初步知识、可靠性基本概念、各种寿命分布和常用的概率分布、回归分析、寿命试验(疲劳寿命试验)和加速寿命试验的图分析法以及各种数值分析法；下册介绍了寿命试验(疲劳寿命试验)与加速寿命试验的极大似然估计法、指数分布的各种分析法、抽样验收方案和威布尔分布的抽样验收方案、可靠性筛选试验和加速寿命试验的失效物理分析。对每种方法既介绍了其简单推理，又举例介绍了如何使用该方法，因此，既照顾了初次接触可靠性工作的人员的需要，也一定程度地满足了对可靠性作深入钻研的要求，尤其是对于从事可靠性试验的人员提供了必备的统计分析知识，同时还适用于目前正在全国推行的全面质量管理(TQC)。本书不仅适宜电子、电气、机械、航空、舰船、兵工、纺织、冶金、化工等领域从事可靠性和质量管理的人员学习、使用，对于科研和教学人员也有一定的参考价值。

## 可靠性试验及其统计分析

### 上 册

戴树森 费鹤良  
王玲玲 苏德清 编著  
白鹤翔 滕怀流

\*

国防工业出版社出版

新华书店北京发行所发行 各地新华书店经售  
国防工业出版社印刷厂印装

\*

850×1168<sup>1</sup>/<sub>32</sub> 印张22<sup>1</sup>/<sub>4</sub> 569千字

1983年7月第一版 1983年7月第一次印刷 印数：0,001—7,500册

统一书号：15034·2467 定价：3.25元

## 序

随着社会主义建设事业的发展，电子设备已广泛用于国防和国民经济的各个部门。而且，无论是现代化武器装备、现代化通讯、交通设施、医疗设备、工业自动化仪表还是空间技术，它们所使用的电子设备由于功能多、自动化程度高而愈来愈复杂，所用元器件数量也越来越多。要使这些电子设备长期稳定可靠地工作，一方面对设备本身要开展可靠性工作，另一方面，所使用的元器件必须长寿命、高可靠。由此可见，可靠性问题是电子工业产品的一个十分重要的问题。一些工业发达国家，对产品的可靠性问题十分重视，他们在可靠性研究上不惜巨资，投入了大量的人力、物力和财力。以美国为例，他们在许多重要工程，如《民兵Ⅰ导弹》、《阿波罗》登月飞行、《空中实验室》计划等工程中都施行了规模庞大的可靠性保证计划，并取得了很好的效果。不少产品的可靠性技术指标有很大的提高，许多基础产品的平均使用失效率能够达到 $1 \times 10^{-10} \sim 1 \times 10^{-12}$  (1/小时)数量级的水平。

我国电子工业的可靠性研究工作虽然在五十年代即已开始，但只是在近十年来才获得较大的发展。目前，电子元器件的可靠性试验已比较普遍地展开，高可靠生产线正在逐步建立，随着全面质量管理的迅速推广，工业管理技术的不断改进，电子元器件产品的可靠性水平必将有大幅度的提高。与此同时，电子设备、系统工程的可靠性设计和可靠性试验也逐步开展起来，某些电子设备如电视机的可靠性试验已经收到显著的效果。

可靠性研究工作，不仅在电子工业中有所成效，而且在机械制造、航空、造船、轻工、橡胶、医药卫生等部门中也有很大进

展。现在各行各业都对可靠性技术提出了进一步的要求。

一般说来，一个产品的可靠性水平主要决定于产品的设计、制造和所选择材料的性能（称为固有可靠性），同时也与使用的条件和状态有关（称为使用可靠性）。为了定量地确定产品的可靠性指标，我们需要对产品进行可靠性试验，并从试验所获得的信息中，采用适当的数学方法进行数据处理，以求得到我们所需要的各种可靠性特征量，如可靠寿命、平均寿命、可靠度、失效率等。但应当注意，这种试验室内的模拟试验终究还不能完全反映使用的实际情况，因此还应当进行现场使用试验，并积累现场数据，进行质量反馈，以改进产品的设计，改进可靠性预测和试验室的可靠性试验。这种设计、制造、试验和使用之间的相互循环，就使产品的质量 and 可靠性水平获得不断的提高。在这种循环中，无论设计时的可靠性预计、生产过程中的可靠性试验还是使用过程中的现场数据收集，都要涉及到数据处理和统计分析问题。人们通过对模拟试验和现场使用数据的积累和分析来寻求产品可靠性的客观规律。因此，在可靠性研究中，如何正确地进行可靠性试验和处理数据就是一个很重要的问题。

为了适应可靠性研究工作发展的需要，本书则着重于介绍可靠性试验的数据处理和统计分析方法。为了方便读者对可靠性技术的深究，本书还提供了有关参考资料目录，以便读者检索。在编写过程中，我们不仅收集分析了国外大量的可靠性资料和数据，而且听取了各方面的意见，尽可能做到由浅入深，通俗易懂，理论与实际相结合，以方便读者实际应用。在第一、二、三章中，扼要介绍了概率论的初步知识和有关可靠性的基本内容；第四、五、六章系统地阐述了各种概率分布、参数估计及回归分析等统计分析方法，特别是对在可靠性研究中经常遇到的某些概率分布作了较详细的介绍；第七至十二章分别介绍了寿命试验和加速寿命试验的设计和试验数据的几种重要的处理方法，如图估计法、最好线性无偏估计法、最好线性不变估计法以及极大似然估计法

等。对于极大似然估计，书中还给出了 ALGOL 算法语言编制的计算程序。指数分布由于使用方便，因而在可靠性研究中占有重要的地位，目前许多国家的技术标准中，绝大多数都以指数分布为基础对电子设备和元器件产品的可靠性指标进行鉴定。我国制订的国家标准《电子元器件失效率试验方法》(GB1772-79)和《电子设备可靠性验证试验》也都是以指数分布为理论基础的。本书第十三章专门论述了指数分布的参数估计和假设检验，这对于理解产品可靠性指标的鉴定方法和标准会有较大帮助，并且提供了理论依据；第十四章介绍了假设检验的基本概念和各种有关的假设检验方法；第十五章叙述了可靠性筛选的数学方法；第十六章论述了指数分布和威布尔分布的抽样原理和具体的抽样方法；第十七章则向读者介绍了阿伦尼斯方程的物理化学原理。

书中论及的内容，不仅适用于电子产品的可靠性研究，而且也适用于其他各个行业的可靠性研究工作，如机械制造业的可靠性疲劳试验、橡胶工业中的老化试验以及其他行业中对产品质量的统计分析等。同时，书中所述的统计学知识和数学方法，如概率分布、参数估计、回归分析、假设检验等，对当前在全国推行的全面质量管理 (TQC) 也是适用的。

本书在编写过程中，不仅考虑到初次接触可靠性工作的同志学习可靠性知识的需要，而且也考虑到对可靠性研究工作进一步深入和熟悉的要求。作者力图用各种可靠性试验的实例来加以说明，这对于只要求学会实际使用方法的读者来说，只要通过阅读各章的实例便可掌握各种试验的分析方法；而那些需要进一步寻求理论依据的读者，则可以阅读本书的全部内容。因此，本书不仅可以作为从事可靠性工作的工程技术人员的阅读资料 and 工具书，而且也可以作为高等院校有关专业的参考书。

参加本书编写工作的有中国科学院系统科学研究所戴树森同志，上海师范学院费鹤良同志，华东师范大学王玲玲同志，四机

部标准化研究所苏德清同志，天津南开大学白鹤翔同志和上海复旦大学滕怀流同志等。

由于时间仓促，水平所限，资料不可能收集得很齐全，书中也难免出现错误和不妥之处，诚恳地欢迎读者对本书提出批评和指正意见。

**编著者**

一九八一年

## 目 录

第一章 数理统计的基本概念 .....	1
1.1 母体与子样 .....	1
1.2 频数分布和频率分布 .....	5
1.3 累积频率分布 .....	12
1.4 频率分布的统计特征量 .....	15
参考资料 .....	20
第二章 概率初步 .....	21
2.1 随机事件和子样空间 .....	21
2.2 随机事件的运算规则和事件之间的关系 .....	24
2.3 事件的概率定义 .....	29
2.4 概率的运算 .....	37
2.5 全概率公式和贝叶斯公式 .....	48
2.6 随机变量 .....	51
2.7 离散随机变量的概率分布 .....	54
2.8 离散的累积分布函数 .....	56
2.9 离散随机变量的期望值 .....	59
2.10 离散随机变量的方差 .....	62
2.11 两个离散型随机变量的联合分布 .....	65
2.12 两个离散型随机变量的协方差和相关系数 .....	72
2.13 离散的随机变量 $X$ 和 $Y$ 相互独立 .....	81
2.14 连续型随机变量的概率分布 .....	82
2.15 连续型随机变量的特征量 .....	85
2.16 关于概率密度函数的变换和随机变量 $X$ 的函数的 期望值 .....	88
2.17 母体的方差和标准离差 .....	90
2.18 连续型的两个随机变量的联合分布 .....	91
2.19 连续型的两个随机变量函数的期望值 .....	93

2.20	连续型随机变量的方差、协方差和相关系数	95
2.21	连续型的随机变量 $X$ 和 $Y$ 相互独立	97
2.22	连续型的条件密度函数	99
2.23	多个连续型随机变量的情形	103
2.24	多个随机变量函数的密度函数	104
参考资料		105
<b>第三章 可靠性的基本概念和特征量</b>		106
3.1	引言	106
3.2	可靠性的定义	109
3.3	可靠性的数量特征	110
参考资料		136
<b>第四章 各种概率分布</b>		137
4.1	引言	137
4.2	离散型失效概率分布	140
4.3	某些连续型失效概率分布	168
4.4	某些推导出来的概率分布	228
附表		257
参考资料		276
<b>第五章 各种子样和估计的概念</b>		281
5.1	引言	281
5.2	随机抽样和各种子样	282
5.3	统计量及其分布	287
5.4	某些极限定理	308
5.5	点估计量及其优劣的准则	311
5.6	置信区间估计	318
5.7	寻求点估计量的计算方法	325
参考资料		326
<b>第六章 回归分析</b>		327
6.1	回归分析的用途	327
6.2	单变量线性回归分析	329
6.3	多个自变量的回归分析	413
附录		424



参考资料 .....	446
<b>第七章 寿命试验和加速寿命试验 .....</b>	<b>449</b>
7.1 寿命试验和加速寿命试验的类型 .....	449
7.2 寿命试验的用途 .....	451
7.3 安排寿命试验时的某些实际考虑 .....	451
7.4 加速寿命试验的三种类型 .....	455
7.5 恒定应力加速寿命试验的直观背景 .....	457
7.6 恒定应力加速寿命试验时的某些实际考虑 .....	461
参考资料 .....	464
<b>第八章 图分析法 .....</b>	<b>465</b>
8.1 引言 .....	465
8.2 正态概率图分析法 .....	465
8.3 对数正态概率图分析法 .....	473
8.4 威布尔概率图分析法 .....	473
8.5 短命试验的图分析法 .....	479
8.6 对数正态分布的恒定温度加速寿命试验的图分析法 .....	483
8.7 对数正态分布的恒定电压(或压力等)加速寿命试验的 图分析法 .....	491
8.8 威布尔分布的恒定电压(或压力等)加速寿命试验的 图分析法 .....	494
8.9 威布尔分布的恒定温度加速寿命试验的图分析法 .....	501
8.10 $F(t_i)$ 置信区间的证明 .....	504
附录 .....	506
参考资料 .....	520
<b>第九章 最好线性无偏估计法 .....</b>	<b>521</b>
9.1 最好线性无偏估计量 .....	521
9.2 寿命分布中位置参数和尺度参数的最好线性无偏估计 .....	522
9.3 极值分布中位置参数和尺度参数的最好线性无偏估计 .....	530
9.4 威布尔分布参数 $m$ 和 $\eta$ 的最好线性无偏估计 .....	536
9.5 正态分布参数 $\mu$ 和 $\sigma$ 的最好线性无偏估计 .....	548
9.6 对数正态分布参数 $\mu$ 和 $\sigma$ 的最好线性无偏估计 .....	552
9.7 恒定应力加速寿命试验中参数的最好线性无偏估计 .....	558
9.8 寿命分布为威布尔分布时的恒定应力加速寿命试验的 最好线性无偏估计 .....	565

9.9	寿命分布为对数正态分布时的恒定应力加速寿命试验的最好线性无偏估计 .....	578
9.10	两个应力下恒定应力加速寿命试验中参数的最好线性无偏估计 .....	589
附录	.....	599
参考资料	.....	615
<b>第十章</b>	<b>最好线性不变估计法</b> .....	<b>617</b>
10.1	最好线性不变估计量 .....	617
10.2	寿命分布中位置参数和尺度参数的最好线性不变估计 .....	618
10.3	极值分布中位置参数和尺度参数的最好线性不变估计 .....	625
10.4	威布尔分布参数 $m$ 和 $\eta$ 的最好线性不变估计 .....	628
10.5	正态分布参数 $\mu$ 和 $\sigma$ 的最好线性不变估计 .....	636
10.6	对数正态分布参数 $\mu$ 和 $\sigma$ 的最好线性不变估计 .....	638
10.7	恒定应力加速寿命试验中参数的最好线性不变估计 .....	640
10.8	寿命分布为威布尔分布时的恒定应力加速寿命试验的最好线性不变估计 .....	648
10.9	寿命分布为对数正态分布时的恒定应力加速寿命试验的最好线性不变估计 .....	657
10.10	两个应力下恒定应力加速寿命试验中参数的最好线性不变估计 .....	663
参考资料	.....	672
<b>第十一章</b>	<b>简单线性无偏估计和简单线性不变估计</b> .....	<b>673</b>
11.1	简单线性无偏估计法 .....	673
11.2	极值分布参数的简单线性无偏估计 .....	678
11.3	威布尔分布参数的简单线性无偏估计 .....	685
11.4	正态分布和对数正态分布参数的简单线性无偏估计 .....	692
11.5	恒定应力加速寿命试验中参数的简单线性无偏估计 .....	693
11.6	极值分布参数的简单线性不变估计 .....	694
11.7	威布尔分布参数的简单线性不变估计 .....	698
11.8	正态分布和对数正态分布参数的简单线性不变估计 .....	699
11.9	恒定应力加速寿命试验中参数的简单线性不变估计 .....	700
参考资料	.....	700

# 第一章 数理统计的基本概念

## 1.1 母体与子样

在生产中，我们经常想知道某一批产品的质量情况。譬如说某批晶体管的电流放大倍数  $h_{FE}$ ，某批继电器的工作次数，某批电子管的废品率，某织布机的断头率，等等。对于社会或自然现象，有时我们想了解某些规律，如想了解某地区成年人的身高，黄河近百年来每年最高的水位，等等。在这些问题的研究中，往往会联想到利用数理统计的分析方法来帮助解决。数理统计这门科学就是研究如何从大量的偶然现象中，找出其中内在的必然规律，从而作出比较科学的推断，以针对不同的问题，给予不同的解决办法。数理统计的方法虽然很多，千差万别，但是有一个概念是共同的，这就是在数理统计学中经常提到的名词——“母体”，或者称作“总体”。这个概念在研究具体问题时，必须加以明确，否则将无从下手，或者将会作出错误的结论，造成前功尽弃。

数理统计学中的“母体”，用简单的话来讲，就是我们所要研究的对象的全体。例如，要研究某生产批的晶体管的电流放大系数  $h_{FE}$  时，将该生产批的所有晶体管的  $h_{FE}$  参数作为母体，而不是指该批晶体管的全体，也不是指该批晶体管其他电参数的全体。如果我们考虑的范围不限于某一生产批的晶体管，而是一周内所有生产批的晶体管的  $h_{FE}$  参数，那么这个母体就扩大为一周内所有晶体管的  $h_{FE}$  参数的全体。如果我们考虑的问题是在一周内各生产批的晶体管的  $h_{FE}$  参数的平均值，而一周内共有 18 个生产批，那么这个母体就是这 18 个生产批的所有晶体管  $h_{FE}$  参数的平均值。

1111451

从上例可以看出，母体是由每个个体组成的。因此，母体又有大小之分。母体中个体数量较大的可以看作无限大，这样的母体称作无穷母体；如果个体数量较少，这样的母体就称作有限母体。当然，个体数量较大的母体，也可以作为有限母体，但是在数理统计方法的研究中，若将较大的母体当作无穷母体，处理问题时，往往有其方便之处，不但造成的误差小，而且也更加合理些。

我们对母体进行研究时，需要通过对其个体的观察来作统计分析。但是，这并不是说对母体中的所有个体都要进行观察，而后再进行统计分析，因为这样不仅工作量很大，而且有时甚至是办不到的。举例来说，我们要研究某生产批的晶体管的放大系数  $k_{FE}$  参数，这个生产批如果数量较小，逐个进行人工测试还是可能的，如果该批管子的数量很大，有数百万只，那么人工测试就不能胜任了。况且，由于观测人员的疲劳，还将会造成人为的误差。有的单位曾经做过全验后是否会有错漏检的实验，即由工作人员将某批管子都逐个进行检测，剔除废品，然后将检测过的管子再重复逐个检测一次，结果又剔出了一些废品。这就说明，全验也会出现错漏检的现象。虽然我们可以利用自动测试设备和计算机，但有时从经济效果来看，也不一定是最科学、最合理的，况且有的试验是破坏性的，例如寿命试验、机械疲劳试验等，都只能抽取部分产品而不允许将全部产品投入试验。那么，为了知道母体的情况，是否可以通过对部分产品的观测来推断母体的情况呢？回答是：可以！这就是从母体中抽取一部分个体进行观察，取得数据，然后利用数理统计的分析方法，对母体作出科学的合理的结论。这里所说的“一部分个体”，是数理统计学中又一个经常用到的“子样”的概念。本章只引入“简单子样”的概念，简称“子样”，又称“样本”。在说明“子样”之前，先引入“随机抽取”的概念。所谓“随机抽取”，是一种在母体中抽取个体的方法，这种方法要求保证使得母体中的每个个体都有同等被抽取到的机

会。为了能做到这一点，通常可利用“随机数表”法<sup>[2]</sup>或抽签法。这种随机抽取的方法可以避免人为因素的掺杂，使所取得的部分个体具有对母体的代表性。

如果我们利用随机抽取方法，在母体中抽取  $n$  个个体，进行观测，得到  $n$  个观测值，并以  $x_1, x_2, \dots, x_n$  表示，这  $n$  个观测值  $x_1, x_2, \dots, x_n$ ，就称为一个子样（样本），或叫做简单子样。

随机抽取，对无穷母体和有限母体有不同的理解。有限母体的随机抽取是这样进行的，抽取第一个个体，进行观测，得到观测值  $x_1$ ，然后将这个被取出的个体放回母体混匀，再抽取第二个个体观测，然后再放回母体混匀，一直进行到第  $n$  个个体抽出时为止。为什么抽取第二个个体时要把第一个个体放回母体中去呢？因为有限母体的个体数量较少，如果抽取第一个个体后不再放回母体，那么母体中的个体总数改变了，减少了一个。如果在减少了一个个体的母体中抽取第二个个体，这第二个个体被取到的机会与第一个个体被取到机会就不相等了，这样的抽取方法就不是随机抽取的方法了。而对于无穷母体，则可以不必再放回去，因为母体中的个体数量很大，抽取少量的个体并不改变母体本身的性质。

本书所论及的内容，主要是针对无穷母体情况的。

**例1.1** 从某一批电阻器中，随机地抽取 100 只，测量其阻值，将测得的阻值列入表 1-1 中，这 100 个观测值就是一个子样。利用上述记号， $n = 100$ ， $x_1 = 197$ ， $x_2 = 201$ ， $x_3 = 219$ ， $\dots$ ， $x_{99} = 207$ ， $x_{100} = 199$ 。子样中的观测值的个数  $n$  称作“子样容量”或“子样大小”。前面所说的“子样”就是子样容量为 100 的子样，其中的数值 197，201， $\dots$ ，又称“子样观测值”，简称“观测值”。

为什么要采取随机抽取的方式呢？这是与我们的目的有关系的。我们的目的是为了研究和弄清母体的性质，但是由于某些条件的限制，不得不只抽取母体中的一部分个体来进行观测，以便

表1-1 100只电阻器的阻值 (单位: 欧姆)

197	187	184	213	203	191	195	185	204	200
201	212	191	196	186	202	209	199	196	184
219*	198	200	205	189	211	197	207	215	213
189	205	211	195	196	193	209	203	192	202
199	192	198	201	185	196	201	185	202	190
207	200	207	194	202	195	215	199	217	206
200	191	197	217	192	214	190	210	194	195
190	202	218	205	181*	207	204	206	201	212
195	193	205	198	208	200	205	196	194	207
210	197	200	186	196	192	202	209	183	199

从这些观测的数据中对整个母体的性质作出估计和推断。因此就必须要求这些被用作观测的个体，对母体的性质具有代表性，也就是说，这一部分个体应当是整个母体的缩影，它们的内在规律应当是一致的（实际上只能做到近似）。而随机抽取的方式就能够保证子样具有这种代表性。这种抽取方式另一个特点是，在随机抽取每个个体时，它们之间互不影响、互不关联，都是“独立”进行的。

以上所讨论的都是简单子样，这种子样在数理统计中是最基本的，也是最简单的一种，这在许多情况下也是容易做到的，即随机抽取  $n$  个个体，进行  $n$  次观测，得到  $n$  个观测值，组成一个子样  $(x_1, x_2, \dots, x_n)$ 。但是，在有些情况下，不容易得到  $n$  个观测值，而只能得到其中的一部分观测值。譬如说，在可靠性试验中，无论是疲劳试验或是寿命试验，在低应力情况下，往往做不到全部失效，也就是得不到  $n$  个失效的时间，而只能得到  $r$  ( $r < n$ ) 个最先失效的失效时间。这  $r$  个观测值就不同于简单子样的观测值，这样的观测值所组成的是“截尾子样”。在可靠性研究中，还有其他各种类型的截尾子样，以后再作具体讨论。而如何根据这样的截尾子样来进行推断和估计，则是本书很重要的一部分内容。

## 1.2 频数分布和频率分布

我们仍以例 1.1 中的实际数据为例。在表 1-1 中所列出的 100 个电阻值，尽管都是从某一批规格相同的电阻器中随机抽取的，而且这批电阻器的生产过程是稳定的，原材料、设备、工艺、环境等都尽量保持一致，但是仍难免有各种各样的随机因素对电阻器的生产发生影响。我们说原材料是“一致的”，只是相对而言，实际上它并不完全一致，而总是包含着不一致性。因为在生产原材料的过程中，同样会受到各种不可控制的随机因素的影响。再说生产设备和环境，实际上也不是理想中的稳定和一致，也是不断地受到许多不可控制因素的影响而带有随机波动性。所有这些不一致性，都会对电阻器的生产发生作用，因此，同一规格电阻器的阻值和其它质量指标，就会有不同程度的差异。表 1-1 的数据就明显地说明了这一点，电阻器的阻值散布在 181 欧姆到 219 欧姆之间。表 1-1 中的数据是按测量的次序排列的，表面上看来似乎是杂乱无章的，但只要我们对这些数据稍加适当的处理，就会发现这些数据包含着规律性。这是一种内在的规律性，是与其母体的性质密切相关的，表现出母体规律性的特点。子样中呈现的这种规律性，正是我们所要重视的一个方面。

对于数据的处理可按以下步骤进行。先将原始数据以大小来分组，按散布在各组的数据个数作成分布图，这种图形称作频数分布图。具体作图步骤如下：

(1) 在原始数据中找出最大值和最小值，计算两者之差(通常称作极差，用  $R$  表示)。

$$R = \text{最大值} - \text{最小值}$$

(2) 计算分组数目  $k$

分组数目可按下式计算：

$$k = 1 + 3.31 \lg n \quad (1-1)$$

式中  $k$  —— 分组数目；

$n$ ——数据总个数。

(1-1) 式的提出可参阅资料〔1〕，也可从图 1-1 中查出  $k$  的数值。

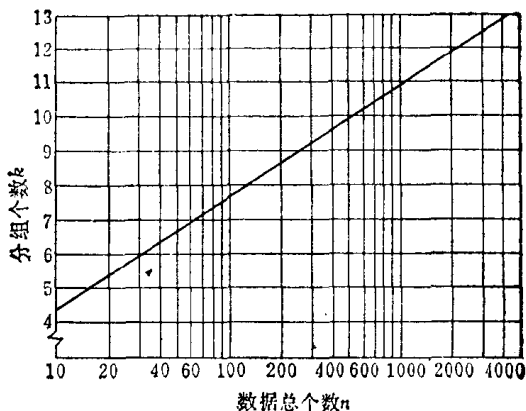


图1-1 分组数目图

分组数目  $k$  的确定不必很严格，上述的公式和图只供作参考。一般分组数目  $k$  在 5~15 之间为宜。

### (3) 计算区间宽度

$$\text{区间宽度} = \frac{R}{k}$$

将整个数据的范围分成  $k$  个区间。(上式计算的区间宽度，也是供参考。)

(4) 将原始数据以其数值的大小，归属于  $k$  个区间中的一个区间内，累加各个区间的数据个数，作数据个数的分布方块图。该图即称作频数分布图，又称累方图或直方图。

下面以例 1.1 的数据为例按上述步骤作图如下：

(1) 极差  $R = 219 - 181 = 38$

(2) 根据  $n = 100$ ，由图 1-1 查得  $k \approx 8$

(3) 计算区间宽度  $= 38 / 8 = 4.75 \approx 5$

现将区间分段和各区间的频数(即数据个数)列入表1-2中。



表1-2 频数表

区间宽度	组中值	记号	频数
180.5~185.5	183		7
185.5~190.5	188		8
190.5~195.5	193		17
195.5~200.5	198		23
200.5~205.5	203		19
205.5~210.5	208		13
210.5~215.5	213		9
215.5~220.5	218		4

在确定区间分点时，为了避免数据落在分点上，可将分点的数值比数据的数值多取一位小数，表1-2中的分点就是这样确定的。表中的“组中值”就是各区间的中点值。

(4) 作频数分布图 (见图1-2)

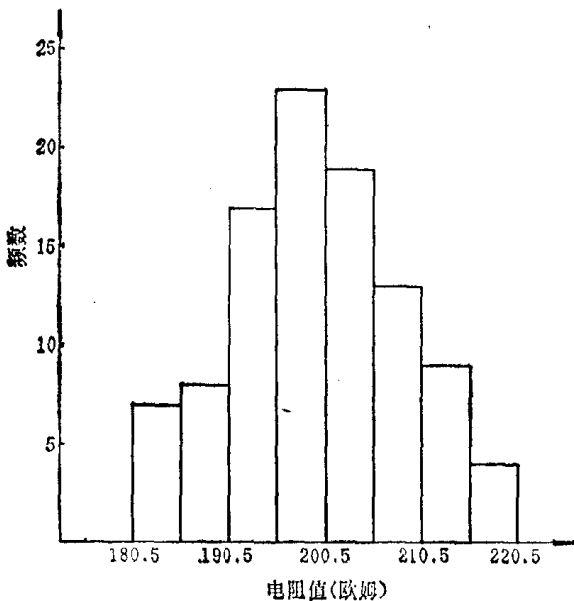


图1-2 频数分布图