质量管理的七种工具

武汉钢铁公司教育委员会办公室 编

科 学 出 版 社

TQC

内容简介

本书概括地叙述了基本数理统计知识，讲述了抽样和数据整理的方法，重点介绍了质量管理中常用的七种工具，用直接而易使读者了解的方式举了一些实例。它的特点是取材广泛且紧密联系实际问题，具有较高的实用价值。

本书可供工矿企业工人、技术人员及干部参考学习。工人可把重点放在应用方面，数理统计基础理论部分供进一步学习使用。

质量管理的七种工具

武汉钢铁公司教育委员会办公室 编

责任编辑 董明 吴锦文

科学出版社出版发行

北京东黄城根北街16号

邮政编码：100707

仙桃市新华印刷厂印刷

*

1990年2月武汉第一版 开本：787×1092 1/32

1990年2月第一次印刷 印张：2 3/4

印数：0001—60 000 字数：59 000

ISBN 7-03-001860/F·39

定价：1.20元

前　言

全面质量管理在武钢推行已经10年了，实践证明，它对提高武钢的现代化管理水平，提高产品质量，降低消耗，挖掘企业内部潜力起到了积极作用。党的十一届三中全会以来，武钢坚持四项基本原则，坚持改革开放，把工作重点迅速转移到深化企业改革上来。近几年，公司明确提出了"突出质量是武钢生产经营指导思想的中心，突出全面质量管理是推行企业管理现代化的中心"的指导方针，建立起以优质产品为龙头的质量保证体系，采取层层实行质量承包等一系列措施，取得了较好的社会效益和经济效益。经过10年的努力，武钢的产量、产值、劳动生产率提高了一倍，利税翻了两番，多项产品获国家质量奖金、银牌，走出了一条质量效益型发展企业的路子。武钢曾先后被评为部、省、市先进企业，荣获国家二级企业称号，并通过了国家一级企业预评和国家质量管理奖预评。

全面质量管理是群众性的质量管理活动，群体质量意识的建立是至关重要的。因此，在推行全面质量管理过程中，要始终把质量教育工作放在首位。在搞好普及教育的同时，学习内容要逐步深化，进一步强化职工质量意识，不断提高职工队伍的质量管理素质，把群众性的质量管理小组活动在全公司扎扎实实地开展下去，力争取得更丰硕的成果。

全面质量管理是一门管理科学，它不是一种行政手段，更不是表面形式。在学习有关全面质量管理理论的同时，还必须掌握一些数理统计知识和科学管理方法。此书人手一册，其目的是使广大职工更好地掌握质量管理方法，并便于

安排电化教学，让大家更好地、自觉地把有关知识和方法运用到工作与生产实践中去，提高实际运用能力，把我们的质量管理工作推向一个新水平。

愿这本书为提高职工队伍的质量管理素质，起到积极推动作用。

编 者

1989.12

目　　录

第一章　数理统计基本知识

§1.1　数理统计中的数据

在生产过程中，与产品质量有关的各种数据帮助人们了解产品质量的特性，认识产品质量的变化规律，发现质量问题，分析造成的原因，从而为保证和提高产品质量采取措施提供依据。推行全面质量管理，应用科学的质量管理方法，就是要系统地收集数据，通过有关的数理统计方法，对数据加以整理分析，形成一整套数据表格和图形，然后，根据科学的质量管理方法，对产品质量起到预测、管理和控制的作用，从而达到保证和提高产品质量的目的。

数据的种类

在我们生产与工作中，几乎天天都与数据打交道。数据不外乎有两大类：即连续的计量数据和不连续的计数数据.

连续的计量数据：如重量、长度、温度、速度等都可用一定的尺度来衡量，而且所得的数据可以是整数，也可以是小数，它们的数据值可以是连续的。

不连续的计数数据：如钢材的合格与不合格数，零件表面质量是否有缺陷等，它们不用计量仪器的尺度来衡量，直接以合格或不合格的数字来计算，只能是零或正整数（0、1，2，3，…等），不可能出现介乎两个相邻正整数间的小数。这种性质的计数数据，具有一种不连续的特性。

计数数据还可以进一步分为计件数据和计点数据。计件数据是检查一批钢板中有多少张是合格品，有多少张是待处理和废品（以一张为整体单位来衡量）；计点数据针对产品的某一质量特性，只能计点数，如钢板上的结疤数、铸件上的砂眼点数、耐火砖上的熔洞数，等等。另外，如废品率、事故率、出勤率等，以百分数为计数依据。

数据的收集

这是一项重要的基础工作。在收集数据时，首先应明确收集数据的目的。目的不同，所需要的数据也就不同，而且收集方法也可能不一样。收集数据一定要避免那种“有用无用一起收”的做法。为了确保数据的有用性，应当根据收集数据的目的而确定收集方法、要求以及要注意的事项。

收集数据的真实性、可靠性非常重要，只有可靠的数据才能提供真实的信息，而虚假的数据和随意制造的数据，则有百害无一利，只能使上级领导产生错觉，作出错误的判断和决定。收集数据时，应如实记录，不得随意修改和剔除。

收集数据的目的如果是为了工序控制或者工艺验证，就应以工序作为总体，从工序中抽出一批半成品或成品进行检验，从而得到数据，然后整理数据分析与判断生产工序的质量状况（图 1.1）。

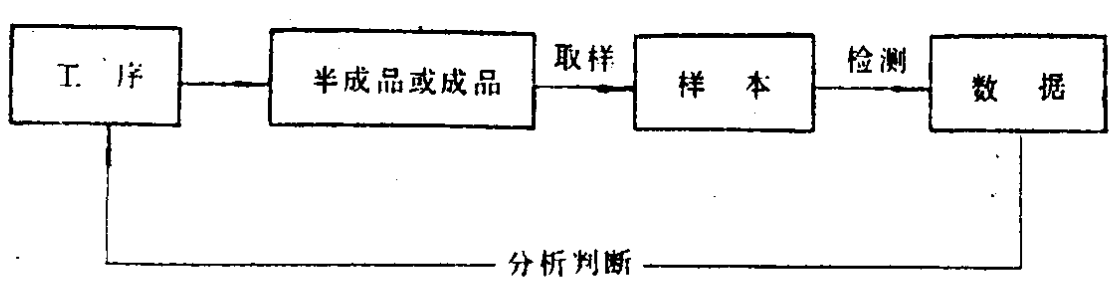

图 1.1

收集数据如果是要对一批产品进行评定，确定合格与否，总体就是一批产品，从中抽取部分产品作为样本，检测后得到数据，然后用这种样本数据来分析确定该批产品的质量水平（图 1.2)。

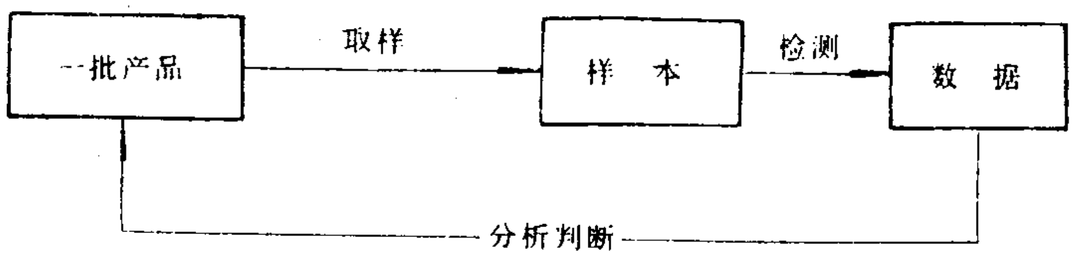

图 1.2

抽样方法

抽样，就是从总体中抽取样本。在工厂中，抽样方式可分为两种：随机抽样和定时定量抽样。

随机抽样：在一批产品或零件中，所有产品或零件都能够以均等的机会被抽取，换句话说，“对总体的任何部分都以同样的概率进行取样”。取样的方法可以有单纯随机取样、系统取样、分层取样、等分取样等。

定时定量抽样：按工艺程序，每隔一定时间持续抽取若干产品作为样本，这样的抽样方式称为定时定量抽样。

这两种抽样方式比较，前者因为每件产品都有均等机会被抽取，抽出的样品有代表性，后者只能代表一定时间的质量水平，而不能代表整批产品的质量状况。

随机抽样，多用于产品的验收检查；而定时定量抽样，多用于生产工序的质量控制。

随机因素与系统因素

要使产品质量能保持稳定，就要设法找出影响产品质量的各种因素。须知，影响质量的因素是很复杂的。从总的方面来看，影响产品质量的因素分为两类：一类是随机因素，又称非系统因素；另一类是系统因素，也称非随机因素。

什么因素属于随机因素或系统因素呢？比如，轧制某一种规格的工字钢，抽查其反映机械性能的屈服强度，我们可以假定：

(1) 用同一钢种的钢来轧制；

(2) 使用同样的轧钢机；

(3) 使这一批工字钢的生产环境前后一致；

(4) 做机械性能试验时用同样的拉力试验机来测试屈服强度；

(5) 做机械性能试验时由同等技术水平的工人操作轧钢机和作拉力试验测试屈服强度。

尽管以上五个条件固定不变，但是当我们抽查这批工字钢并测试其屈服强度时，结果仍各不相同。这是因为影响工字钢机械性能的因素很多，除上面所列五条外，尚有许多因素是不易一一确定的。这些因素不仅不易通过分析研究加以确定，而且即使能确定也难以控制，例如厂房内温度的微小变化，测试产品性能时仪器产生的误差，同一技术工人在八小时工作内由于精力集中程度的不同给产品性能带来微小的差异，等等。这许多微小不可测的因素，它们之间相互独立，但对产品质量起总的影响。我们称这类因素为随机因素。而那些可以列出的、易于控制的因素，如条件 (1)—(5)，则称为影响产品质量的系统因素。

在生产过程中，我们进行抽样检验，其目的就是通过对得到的数据进行加工分析，判断影响产品质量的因素。一个生产过程如果仅因随机因素产品质量产生波动，我们就说生产过程处于正常状态，否则，就说生产过程处于非正常状态。此时，就需要立即找出原因，或调整设备，或检查原材料有无异常。务必要消除系统因素对生产质量的影响。

怎样才能知道生产过程是否处于正常状态呢？即在正常状态下，随机因素应该呈现出什么样的规律性？我们从理论上可以证明，如果生产过程处于正常状态，这时的产品质量指标将形成正态分布。为掌握正态分布的有关知识，我们必须懂得概率及分布。

§1.2 数理统计基础知识

随机事件

在一定条件下可能发生也可能不发生的事件称为随机事件（也称偶然事件）。如，产品抽样检查时，从不合格品率为10%的一批产品中任意抽取5件“都是合格品”，这就是一个随机事件；“5件中有1件不合格品”，又是一个随机事件。

在一定条件下必然要发生的事件叫做必然事件。例如，当一批产品不合格品率为零时，任意从中抽出一件来检验，“必然是合格品”，这就是一个必然事件。

在一定条件下必然不发生的事件称为不可能事件。例如，有50件同类产品，其中只有两件次品，任意抽取3件，“3件都是次品”，是不可能发生的事件。

在自然界中，偶然性是永远伴随着一切事物的，但在偶然中又蕴藏着必然的规律。

什么是概率

概率就是表示某个事件发生的可能性。在日常生活中，概率较通俗的解释是某件事情确实发生的程度或发生可能性的大小。

概率的定义：假设某事情进行了 n 次重复试验，其中事件 A 出现了 m 次，如果试验次数 n 很大，事件 A 出现的频率 m/n 就稳定在某个数值 p 的附近摆动，那么我们定义事件 A 的概率为

$$P(\mathrm{A})=p \tag{1.1}$$

在一般情况下，固定数值 p 是不可能精确地得到的，因此，通常在 n 充分大时，以事件 A 的频率作为事件 A 的概率 p 的近似值，即

$$P(\mathrm{A})\approx\frac{m}{n} \tag{1.2}$$

这样通过频率计算的概率称为统计概率。

例如，我们检查一批在相同条件下生产出来的产品，按产品质量标准规定，当厚度介于 5.6mm 至 6.4 mm 时为合格，否则为次品。我们分别抽取 5 张、10 张、60 张、150 张、600 张、900 张、1200 张、1800 张来检查，其结果列于表 1.1 。

由表 1.1 中看出，随着试验次数（检查的总数）增多，合格品的频率越来越接近于一个稳定的值 0.9。

必然事件的频率为 1，不可能事件的频率为 0，由于频率

m/n 总介于 0 与 1 之间，因而由概率的定义可知，对任何随机事件 A 有 $0\leqslant P(A)\leqslant 1$。

表 1.1 合格品频率表

检查张数	5	10	60	150	600	900	1200	1800
合格品数	5	7	53	131	548	820	1091	1631
合格品频率	1	0.7	0.883	0.873	0.913	0.911	0.909	0.906

概率的计算

1. 互补定理

假设掷一次骰子出四点的概率为 1/6，而从 1 至 6 点出任何点的概率为 1。那么，出四点以外其他点的概率等于从 1 减去 1/6，即 5/6。从这个计算我们可以看到，某件事情发生的概率和不发生的概率之和必须为 1。

设某事件发生的概率为 P，则不发生的概率为 $1-P$。

2. 加法定理

掷骰子出三点概率是 1/6，出四点的概率也是 1/6，出三点和四点都可以的概率则应该是 1/6 + 1/6 等于 2/6。

在概率计算时，互相独立的事件出现其中任何事件的概率是那些事件各自发生的概率之和。如从 52 张（除掉两个王牌）扑克牌中，任意抽出一张，出现红桃 A 的概率为 1/52，而出现任何种类 K 的概率为 1/13，那么出现红桃 A 和任何种类 K 的概率则为 1/52+1/13 = 5/52。

3. 乘法定理

如果将两颗骰子一齐掷出（或一先一后掷也可以），两颗骰子都出三点的概率是 1/6 乘 1/6 等于 1/36。

在概率计算时,相互独立的事件同时发生的概率是这些事件各自发生的概率的乘积。

排列与组合

在研究概率或数理统计时,须掌握一些排列与组合的知识。

1. 排列

让我们先看看下面的问题。

某车间准备从三个懂技术、精通业务的工人师傅赵、钱、孙当中选出正、副班长各一名,则有下列 6 种不同的选法:

(班长, 副班长)① ={ (赵,钱), (赵,孙), (钱,赵), (钱,孙), (孙,赵), (孙,钱) }

从 m 个元素中,每次取出 n 个元素按照一定的顺序摆成一排,叫做从 m 个元素里每次取出 n 个元素的排列。其排列的种数以符号 A_m^n 表示:

$$A_m^n = m(m-1)(m-2)\dots(m-n+1) \tag{1.3}$$

公式 (1.3) 表明, 从 m 个元素里, 每次取 n 个元素, 所有的排列的种数等于 n 个连续自然数的积, 其中最大的一个是 m。

当 $n=m$ 时, 从 m 个元素里每次取出 m 个元素的排列叫全排列。m 个元素所有的全排列的种数,通常用符号 P_m 表示:

① 全体可能结果, 可用集合表示。

$$P_m = m(m-1)(m-2)\cdots 3 \cdot 2 \cdot 1 \tag{1.4}$$

公式（1.4）说明，m 个元素所有的全排列的种数等于自然数 1 到 m 的积，通常用 $m!$（读作“m 的阶乘”）表示。所以公式（1.4）可以写成下面的形式：

$$P_m = m!$$

例 1.1　用 1，2，3，4 四个数字可以组成多少个没有重复数字的四位数？

解：所求的四位数的个数等于 4 个元素的全排列的种数：

$$P_4 = 4! = 24$$

答：可以组成 24 个没有重复数字的四位数。

2. 组合

我们分析一下下面的例子。

从不在一条直线上的三个点 A，B，C 里面，每次取出两个点连接成一条直线，可以得出下面的三条直线：

$$AB，AC，BC$$

从 m 个元素里，每次取出 n 个元素，不管怎样的顺序并成一组，叫做从 m 个元素里每次取出 n 个元素的组合。

从 m 个元素里每次取出 n 个元素所有不同的组合的种数，通常用符号 C_m^n 表示：

$$C_m^n = \frac{A_m^n}{P_n} = \frac{m(m-1)(m-2)\cdots(m-n+1)}{1 \cdot 2 \cdot 3 \cdots n} \tag{1.5}$$

上面公式(1.5)表明，要计算 C_m^n 的值，只需要写出一个公式，使它的分子是 n 个连续自然数的积，其中最大的一

个是 m，分母是从 1 开始的 n 个连续自然数的积。

例 1.2 求 C_7^3，和 C_{100}^3 的值。

解：

$$C_7^3 = \frac{7 \times 6 \times 5}{1 \times 2 \times 3} = 35$$

$$C_{100}^3 = \frac{100 \times 99 \times 98}{1 \times 2 \times 3} = 161700$$

分布及其特征值

通常人们都有这样一种概念，在一个轧钢车间里，同一个生产工人，在同一台轧机上操作，用同一种原材料，按照同一工艺规程连续轧制同一种钢板，如果对这位生产工人某一天所生产的钢板进行逐张测量厚度，就会发现测量结果并不完全相同，而是有一定的差异。钢板厚度的轧制公差就是允许钢板实际尺寸差异的变动范围。

将钢板实际尺寸的变化用某种图形客观地表示出来，就称为钢板实际尺寸的分布。

数理统计学中，用来表示分布及其某些特性的特征值可以分为两类。一类是表示数据的集中位置的，其常用的特征值，有算术平均数、中位数等；另一类是表示数据的分散程度的，它的特征值有标准偏差及极差（或称范围）。

1. 算术平均数

算术平均数是表示数据集中位置最有用的代数值。算术平均数的求法也比较简单，它是把一组数值相加后再用数值的个数去除而得到的平均数通常用 $\bar{x}$ 表示。

例 1.3 如测量同一批钢板的厚度，测得一批数据为 x_1、x_2、x_3，…，x_n，求算术平均数 $\bar{x}$。

$$\overline{x}=\frac{x_1+x_2+x_3+\cdots+x_n}{n} \tag{1.6}$$

或者写成

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \tag{1.7}$$

上式中记号“$\sum$”表示求和的意思，$\sum_{i=1}^{n}x_i$表示从x_1加到x_n。

2. 中位数

在一组数据中，按其数值的大小排列，恰好在中央的数就称为中位数。比中位数大的数据的个数与比中位数小的数据的个数是相等的。

例如，有一组数据为 5，6，6，7，8，9，9，共七个数据，其中数据 7 为中位数，比 7 大的数据有 8，9，9 三个，比 7 小的数据有 5，6，6 三个。

上面一组数据的个数为奇数，中位数就是按数据的大小排列居于中央的那个数据。当一组数据的个数为偶数时，在中央的数据不是一个而是两个，这时取这两个数的算术平均值为中位数。

3. 极差

极差又称范围。在一组数据中，数值存在的区间称为分布的范围。也有时把极差称作变异的幅度，简称变异幅。计算极差的方法，是从$x_1,\cdots,x_n$中的最大值减去最小值，即$x_{i\max}-x_{i\min}$称为$x_1,\cdots,x_n$的极差，通常用R表示：

$$R=x_{i\max}-x_{i\min} \qquad (1\leqslant i\leqslant n) \tag{1.8}$$

例 1.4　有一组数据，由 5，7，8，9，10 五个数据组

成，其最大值为 10，最小值为 5，极差的求法如下：

$$R=10-5=5$$

极差的计算十分简单，只要知道一组数据的最大值与最小值，极差就可以求出，与一组数据中其他数据没有关系。

4. 标准偏差

上面我们介绍了极差，因为它没有充分利用数据提供的资料，因此反映实际分布情况不够准确，于是数学家又给我们提供了另一个表示数据分布离散特性的数值，称为标准偏差。

标准偏差又称标准离差或标准差。通常用 σ 表示总体的标准偏差；用 S 表示样本的标准偏差。

总体标准偏差用下述公式表示：

$$\sigma=\sqrt{\frac{1}{n}[(x_1-\mu)^2+(x_2-\mu)^2+\cdots+(x_n-\mu)^2]}$$

$$=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2} \tag{1.9}$$

式中　x_i—总体数据的各个数值；

μ—数据的算术平均数；

n—数据的个数。

总体的标准偏差，真正要求得是很困难的，特别是破坏性检验，不可能把所有产品都检验。所以，一般我们是通过样本的标准偏差来作估计的。

样本的标准偏差的公式为：

$$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2} \tag{1.10}$$