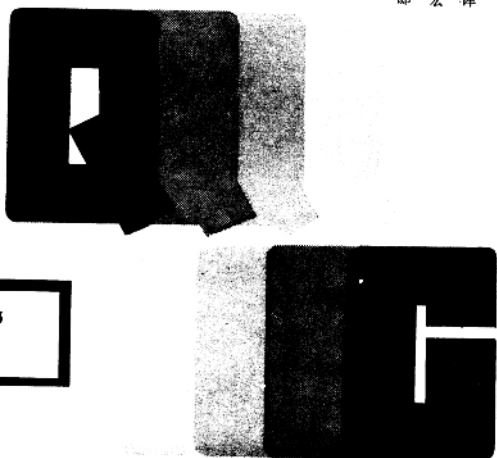


数据的收集方法和应用(二)

〔日〕 藏 健司 编

大滝厚 谷津进 著

邱 宏 译



中国经济出版社

前 言

日本的质量管理有两个最显著的特征。第一个是全公司的质量管理（TQC），另一个是质量管理小组活动。在质量管理活动中，上至管理的领导者，下至从业人员都坚持不懈地应用统计方法，并努力根据质量管理的思想，以数据为基础进行管理和改善。

要分析出隐藏在数据中的事实，最基本的方法就是所谓“现场质量管理的七种工具”，即质量管理方法。这些方法既简单又能被大家理解，所以，在质量管理小组活动中使用得很广泛，在发现问题、解决问题的各个步骤中应用这些方法颇见成效。甚至被称赞为是一种强有力的武器：“只要使用质量管理七种工具，现场管理及改善工作中90%的问题就能迎刃而解”。人们已经承认质量管理小组成员首先要学习这种工具的应用方法，在质量管理的各个阶段中实际应用这种工具解决问题并已大见成效。这些成果本身也从另一个侧面说明了质量管理小组活动并不是光靠精神战术所能奏效的。

全面质量管理中的质量管理小组活动之所以能象今天这样活跃，可以说主要是因为有了“恰当的工具”，特别是统计方法。

本书是本讲座之五的姐妹篇——《数据的收集方法 和应用（一）》的续篇。本书对作为确定重点项目的 方法 而经常使用的排列图（巴雷特图）；对分析数据离散状态的方法——直方图；对了解影响质量的主要原因之间关系的方

法——散布图，为掌握各个数据中潜在的总体性质的数量方法——数据的数量化等的步骤及其应用的要点，通过实例进行了说明。

本书在执笔时，考虑到能使现场监督者、从业人员以及质量管理小组各位成员都能通过自学来理解本书的内容，因此有意地在管理方法章节中举有示例，并按步骤进行解题，说明应用要点。最后，安排了供自学者用的研究课题。另外，本书和它的姐妹篇《数据的收集方法和应用（一）》，在内容上是互相独立的，这种编排方式使读者无论从哪一本开始阅读都很方便。

在质量管理方面，用PDCA循环来进行管理和改善是非常重要的。希望诸位也能通过这本关于方法的学习来最大限度地掌握PDCA循环的思想，努力使自己本身以及小组的管理和改善能力更上一层楼。要巧妙地循环PDCA，就要有“用数据来说话”的态度。要达到此目的，最重要的是读者要有一种把从本书中得到的东西贯彻到实践中去，并为本公司留下业绩的进取精神。并不是说一经学了工具的使用方法，马上就会使用这种工具。所谓“如果不去反复地进行实践，不对工具本身了如指掌，则不可能对工具运用自如”和“趁热打铁”这些话都是这个道理。希望诸位务必在学习之后进行应用，努力使它成为自己身上的血和肉。

最后，向通读我执笔的这本书，并在内容结构等方面向我提出宝贵意见的織健司先生和鼓励我不管进度多慢也要坚持写下去的日本规格协会的石川健先生表示衷心感谢。

著 者

1983年10月

目 录

1. 排列图(巴雷特图)	(1)
1.1 什么是排列图	(1)
1.2 排列图的制作方法	(7)
1.3 排列图的识图及使用方法	(13)
1.4 注意事项及使用要领	(18)
2. 直方图	(21)
2.1 什么是频数分布	(21)
2.2 频数分布表、直方图及累积频数图的制作方法	(24)
2.2.1 频数分布表的制作方法	(24)
2.2.2 直方图及累积频数图的制作方法	(28)
2.2.3 取集数据制作频数表、直方图及 累积频数图的方法	(30)
2.3 直方图的识图方法	(35)
2.3.1 分布形态的识别方法	(35)
2.3.2 与规格的比较	(37)
3. 散布图	(48)
3.1 什么是散布图	(48)
3.2 散布图的制作方法	(49)
3.3 散布图的识图方法	(52)
3.3.1 散布图的识图方法	(52)
3.3.2 分析散布图时的注意事项	(53)
3.4 散布图的应用——相关与回归	(55)
3.4.1 相关的检定	(56)

3.4.2 回归直线的求法	(57)
4. 数据的数量化	(67)
4.1 中心和离散的数量化	(67)
4.1.1 数据中心位置的表示方法	(67)
4.1.2 数据离散(幅度)的表示方法	(69)
4.2 平均值和标准偏差的计算方法	(72)
4.2.1 根据频数分布表计算平均值和标准偏差	(72)
4.2.2 根据变数变换计算平均值和标准偏差	(75)
4.2.3 计算值的修正方法	(77)
4.3 平均值和标准偏差的应用	(79)
4.3.1 分布的一般形态——正态分布	(80)
4.3.2 工序能力指数	(83)
4.3.3 正态分布表	(85)
研究课题答案	(93)
参考文献	(99)

1. 排列图（巴雷特图）

1.1 什么是排列图

在我们的企业中尚存在着各种各样的问题。所谓问题，是指本企业应有的状态、应有的目标同现状水平之间的差距而言的。因此，不但要弄清楚产品有无缺陷的问题，而且还要事先弄清楚各项工作中存在着什么样的问题和这些工作本身所应有的状态。同时要制定出能最大限度地将工作实绩表示出来的指标，并在此基础上设立一个能够掌握问题点的组织机构。那些总认为“没有问题”的人，实际上是一些找不出问题所在的人。

在日常的维持管理和改善活动中，能不断地找出关键问题的人，才是最有效率的人。这类关键问题，就是使自己最头痛的问题和改善后效果最显著的问题。如果过多地去解决那些无关紧要的问题，而对那些重要的应该干的问题却忽略掉了的话，那就会劳而无功！在找出多种问题之后，再进一步弄清真正该解决的重要问题是什么，然后再着手解决。质量管理并不是眉毛胡子一把抓，而是要着重指出重点问题。要在现场准确选定什么是最重要的和该解决的问题，就要使用排列图。

所谓排列图，是把已构成问题的不良以及修理、修补、缺陷、索赔、事故、故障等等问题按其现象和原因等分门别类加以整理，得出数据，然后将整理出来的数据按不良

个数或损失金额多少的顺序排列起来，将其大小用柱状图表示，再配上表示累积的折线，这样形成的图表就叫排列图。

意大利的经济学家巴雷特 (V. Pareto) 在1897年提出收入分布偏于低收入阶层，并发表了这个收入分布的规律。此外，还有美国的经济学家洛伦茨 (M. C. Lorenz)，他在1907年通过被称为洛伦茨曲线的图把收入分布的不均等程度表示出来了。巴雷特的学说也好，洛伦茨的学说也好，都强调了极少数人占有着大部分财富这样一种事实。

用横坐标表示收入金额由高至低按顺序排列起来的人数累积数，用纵坐标表示按各人收入金额和累积收入金额，这样就得到了如图1.1所示的曲线。这曲线图表示从最富有的人到收入为E日元的人占a%，财富的Ec%被占a%的最富有的人所占有。如果所有的人的收入都均等的话，那么累积曲线必然是45°的均等分布线，可是实际上是不均等的。因为少数的高额收入阶层占有了相当的所得，所以图表上显示出一种急剧上升的曲线，而在低收入阶层范围内则成了缓慢的饱和型的曲线。

美国的经营顾问朱兰 (J. M. Juran) 主张把排列图用于质量管理中，此时，横坐标表示按频次顺序排列的不良项目，纵坐标表示不良个数或损失金额和积累百分率 (见图1.2)。

借助这样的排列图，我们就可知道不良个数和损失金额的大部分是被众多项目中的少数几个项目所占有的，就会懂得在改善活动中，选择那些极少数已构成主要问题的项目并解决它，要比选择那些为数众多但损失轻微的项目更有效益。我们称之为巴雷特原则。

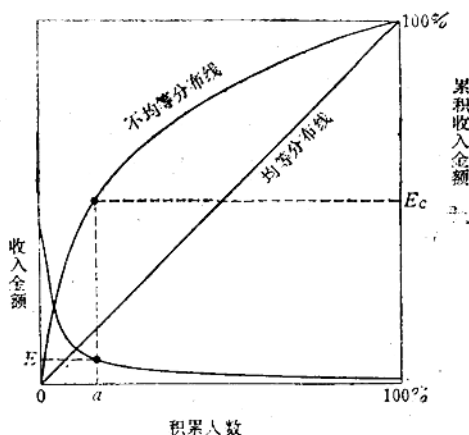


图 1.1 收入的不均等分布曲线

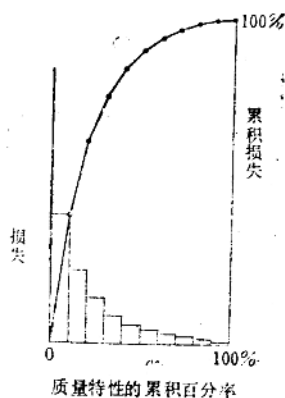


图 1.2 朱兰的排列图

在与此完全相同的想法中，还有一个使用P-Q图表（P代表产品，Q代表数量）和ABC曲线的重点指向问题。

参考

巴雷特法则 年收入金额的分布如图 1.3 所示，明显地倾向低收入阶层一方。也就是说，属于低收入阶层的人数很多。相反，属于高收入阶层的人数则是递减的。因此，如果把横坐标和纵坐标都按对数刻度，从高收入阶层开始把收入者的人数累积值表示在纵坐标上，如图 1.4，这就会形成像巴雷特所证实的那样一条折线。此图表示收入在 E 日元以上的人是全体的 $a\%$ 。

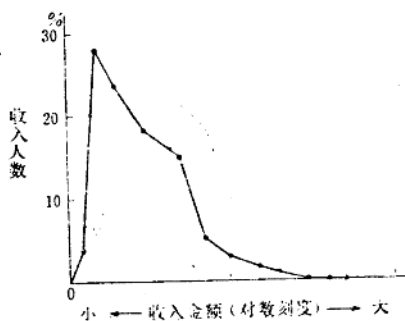


图 1.3 按不同收入阶层的收入人数的分布

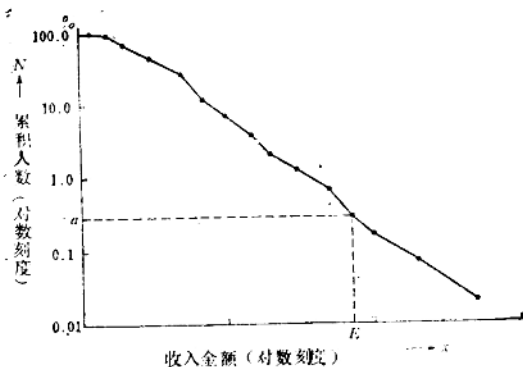


图 1.4 巴雷特折线

设 x 为收入金额， N 为收入 x 以上的人数，如果把图1.4用式子来表示，则为

$$\log N = \log A - \alpha \log x$$

或

$$N = \frac{A}{x^\alpha}$$

这个式子称为巴雷特法则。

为了把这个法则和通常使用的巴雷特图结合起来，把横坐标与纵坐标对换表示，并刻上普通刻度，则成为图1.5。这图说明，在收入等级上属上等的第 a 位的人的收入为 E 日元。并看出 α 值越小，则财富越向高收入阶层集中。这个 α 叫做巴雷特系数。

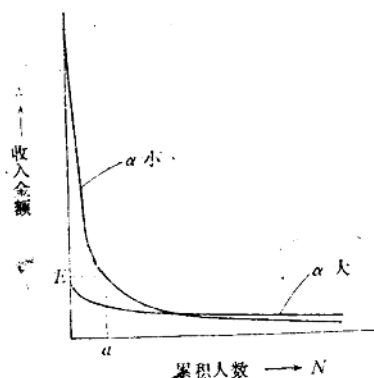


图 1.5 巴雷特法则

洛伦茨曲线 如果把年收入金额适当地分成若干区间，根据各区间的收入人数制成频数图，则成图1.3。如果这是正态分布，并且算出平均值和标准偏差，就能够表现出它的离散状态。但是，对于表示偏向低收入阶层的非对称性曲线则需要下一番功夫研究。

首先，如表1.1所示，根据收入阶层的构成人员的收入从低到高的顺序，计算累积值(巴雷特是从高收入阶层开始累积的)。然后求各阶层的收入总数(相当于频数表的 u 值)。这个值是从低收入阶层开始累积的。如果用横坐标表示收入阶层的累积人数，用纵座标表示累积收入金额，就能得到如图1.6所示的洛伦茨曲线。

表 1.1 收入的频数表

收入阶层 u		f	uf
低 ↑	○ 日元	↓	↓
	○ 日元		
	○ 日元		
↓ 高	○ 日元	↓	↓
	○ 日元		
		横坐标	纵坐标

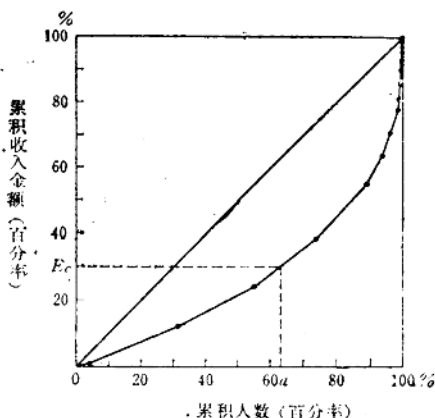


图 1.6 洛伦茨曲线

这个图表示，从低收入阶层数起的 $a\%$ 的人，只占有 $E_c\%$ 的财富。如果收入按人均分配，则 10% 的收入者应各均占 10% 的财富，所以，这个点应落在对角线上。这个 45° 线叫均等线。一般来说，越是偏离这条均等线越是不均等。也就是说，越是靠近高收入阶层就越偏离均等线。它意味着一部分高收入阶层占有着社会的大部分财富。

为了使其与常用的排列图进行比较，累积值最好不要从低收入阶层开始计算，而从高收入阶层开始计算，也就是把图 1.6 的刻度反转过来，旋转一百八十度成图 1.1 的形状。

ABC分析法 在管理项目很多时，如果全部实行同样管理就会感到困难，这就要进行重点管理。为了减少重点管理项目而采用的方法，

就是ABC分析法。以库存管理为例,将多数库存品按消费金额的大小顺序排列,制作它的累积曲线(ABC曲线)。把前位的约10%定为A组,其次的约20%定为B组,其余的定为C组,在多半情况下,A组的合计值占全体的半数以上,所以最好将A组作为重点管理项目(见图1.7)。

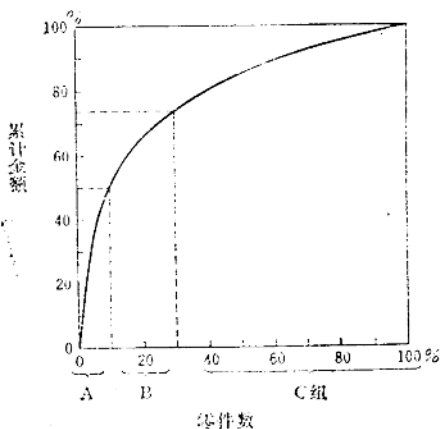


图 1.7 ABC曲线

P—Q图 是表示产品(Product)和数量(Quantity)关系用的一种图表。在横坐标上将产品按品种产量的大小顺序排列,纵坐标上为各产品的生产量。根据累积值制作这种折线图来研究不同品种中产品的重点。此法也可以用于多品种、小批量工场总体设计的计划等方面。

1.2 排列图的制作方法

把在某塑料注塑成形工厂一周内发生的不良品数据收集整理,如表1.2所示。下面,我们以此为例来说明排列图的制作方法。

表 1.2

不良品数据

不良项目	不良品个数
尺寸超差	12
毛刺	73
气孔	17
光洁度	43
接缝	6
其它	20

步骤 1 决定分类项目，收集数据

在决定数据的项目分类、收集时间、收集方法后，再收集数据。关于项目分类，尽量以被认为是原因的项目进行分类，这样比较容易采取对策。一般采用如下的分类方法。

按不同原因分类：如材料、机械、操作者、操作方法等。

按不同内容分类：如不良品和各种费用的项目、场所、工序、时间、不同车间等。

收集数据的时间一般以一周或一个月那样便于划分的一个期间为好。应以取得最近的数据为原则。如果应用检查表，那么数据就更能简单地取得。

步骤 2 把数据按大小顺序排列

把数据按大小顺序排列，在各项目中填写上各自的数据。这时，把若干个数值小的项目合起来作为“其它”项目填写在最后（见表1.3）。

表 1.3

所有项目列入表中

项 目	数 据 值 (不良品个数)	累 积 值	数据值百分率	累积值百分率
毛 刺	73			
光 洁 度	43			
气 孔	17			
尺寸超差	12			
接 缝	6			
其 它	20			
合 计	171			

步骤 3 计算累积值

把多个项目的数据值，依次加起来计算累积数。

例，气孔栏的累积值是以 $73 + 43 + 17 = 133$ 的形式累积，但因为开始的两个项目的合计已算出，已经在光洁度一栏中得到了累积值116，所以也可以以 $116 + 17 = 133$ 进行累积计算（见表1.4）。

表 1.4

累积值的计算

项 目	不良品个数	累积值	不良品个数百分率	累积值百分率
毛 刺	73	73		
光洁度	43	116 (= 73 + 43)		
气 孔	17	133 (= 116 + 17)		
尺寸超差	12	145 (= 133 + 12)		
接 缝	6	151 (= 145 + 6)		
其 它	20	171 (= 151 + 20)		
合 计	171			

步骤 4 计算数据值百分率和累积值百分率

按下式求出各自的百分率，保留小数点后 1 位数。

$$\text{数据值百分率} = \frac{\text{数据值}}{\text{合计}} \times 100\%$$

$$\text{累积值百分率} = \frac{\text{累积值}}{\text{合计}} \times 100\%$$

数据值百分率的合计因受数值四舍五入的影响，所以往往不是100.0（见表1.5）。

表 1.5 数据值百分率和累积值百分率的计算

项 目	数 据 值 (不良品个数)	累 积 值	不良品个数百 分率	累积数百分率
毛 刺	73	73	42.7%	42.7%
光 洁 度	43	116	25.1	67.8
气 孔	17	133	9.9	77.8
尺寸超差	12	145	7.0	84.8
接 缝	6	151	3.5	88.3
其 它	20	171	11.7	100.0
合 计	171	—	99.9	—

步骤 5 画出横座标和纵座标，制成柱形图

以纵座标表示特性，刻上包括数据值合计在内的刻度。

在横座标上，从数据值多的项目依次从左向右注明项目的名称。在确定项目间隔时要尽量使纵座标和横座标的长度接近1比1（完成了的排列图基本上是正方形的）。

把各项目的数据值制成柱形图。柱形要宽度相等，不留间隔地绘制（见图1.8）。

步骤 6 作累积折线

把累积值在柱形图的右上角打点，把这些点用直线连接起来形成的一条折线称为累积折线（见图1.9）。

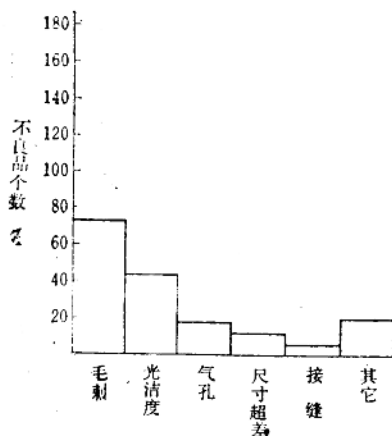


图 1.8 柱形图的作法

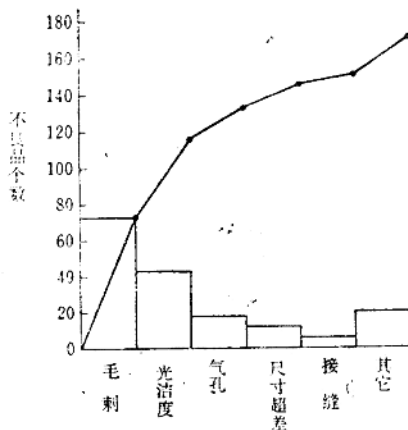


图 1.9 累积折线的作法

步骤 7 在右侧画上纵座标并标明刻度

折线的始点是 0%，终点是 100%，将 0~100% 之间分成 5 个等分，再标上刻度，填上 0、20、40、60、80、100%

的数值。这就是显示累积值百分率的刻度（见图1.10）。

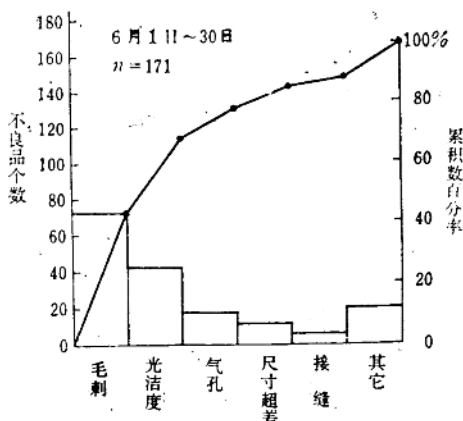


图 1.10 塑料注塑成形不良品的排列图

步骤 8 记录必要事项

在作成的排列图上常需填写标题、取数据的时间、数据总数、工序名称、作图者、作图日期等。为了应用排列图，一般在纵座标和横座标上表示如下内容。

（1）在纵座标上表示如下内容

- ①金额：损失金额，各种经费，零件单价，销售金额等
- ②质量：不良品件数，维修、修好件数，缺陷数，退货数等
- ③时间：作业时间，维修、修好时间，故障时间等
- ④安全：事故件数，灾害件数，故障件数等
- ⑤劳动态度：出勤率，缺勤率，建议次数等

（2）在横座标上表示如下内容

- ①现象：不同的不良项目、缺陷内容、事故内容、位置等