SQC-9

统计质量控制

STATISTICAL QUALITY CONTROL

统计方法应用演示50例

50 examples of application on statistical method

陈国铭　编

国防工业出版社

SQC-9

统计质量控制

STATISTICAL QUALITY CONTROL

统计方法应用演示50例

50 *examples of application on statistical method*

陈国铭 编

中国标准出版社

（京）新登字 048 号

内 容 提 要

本册是《统计质量控制》前八册内容的应用实例演示。实例中的 30%来源于石化企业生产实践，70%来源于国内外流程型行业应用实践。为使读者易于理解，对实例进行了加工整理。对每一例题中使用的方法和公式都做了简要回顾，目的在于使读者既掌握各种方法的要点，同时又收到复习的功效。在例题的编排上则与本丛书的顺序基本一致，使读者易于达到贯通全书。

本书可供企业质量管理干部、工程技术人员及有关院校师生使用。

SQC-9
统计质量控制
STATISTICAL QUALITY CONTROL
统计方法应用演示 50 例
50 examples of application on statistical method
陈国铭 编
*
中国石化出版社出版发行
（北京朝阳区太阳宫路甲 1 号 邮政编码：100029）
煤炭工业出版社印刷厂排版
中国纺织出版社印刷厂印刷
新华书店北京发行所经销
*
850×1168 毫米 大 32 开本 5½ 印张 144 千字 印 1－6400
1995 年 3 月北京第 1 版 1995 年 3 月北京第 1 次印刷
ISBN 7-80043-562-8/O・027 定价：6.50 元

中国石油化工总公司质量管理协会组织编写

生产技术顾问：张德义
统计技术审核：王经涛

主　编：陈国铭
副主编：张祖荫　郭耀曾
编　委（按姓氏笔划）：李世英　陈国铭　杨丽春
张祖荫　饶上建　郭耀曾　崔廷铨
其他编辑校核人员：万　涛　刘秋萍　吕巧云
邱以玲　田从金

序　言

为了适应国际贸易往来和经济合作的要求，国际标准化组织经过十多年的努力，于1986年和1987年相继正式发布ISO8402《质量——术语》标准和ISO9000质量管理和质量保证系列标准，将世界多年质量管理的经验进行了标准化。ISO9000系列标准的基本点是要求企业在生产过程中建立有效的质量保证体系，并对质量体系中相互关联、相互作用的若干要素进行有效的控制。在过程质量控制中，科学、有效方法之一就是数理统计方法。因此在ISO9000系列标准的各个模式中以及质量管理和质量体系要素指南中都要求在市场分析、产品设计、工序控制、性能评定、数据分析等方面广泛使用统计技术，其范围包括实验设计、方差分析、显著性检验、累积和控制图、抽样检验等技术。因此，研究学习统计质量控制技术对于贯彻ISO9000质量保证系列标准，提高科学管理水平是非常必要的。

回顾世界质量管理的发展史，可以看出，数理统计技术在质量管理中发挥了重要作用。从19世纪末到现在，质量管理在历史上经过了检验质量管理、统计质量控制和全面质量管理三个阶段。单纯检验质量管理的严重缺点：一是只能从产品中发现和挑出废品，事前预防功能不强，二是由于检验人员的差错，即使全数检验也可能漏检或错检；三是至关重要的破坏性试验不可能全数进行。产品是生产出来的，单靠检验是不能防止产生废品的。1924年美国贝尔研究所的休哈特（W. A. Shewhart）运用数理统计的原理提出了控制生产过程中的“6σ”方法，即后来发展的质量控制图和预防缺陷的概念。与此同时，同属贝尔研究所的道奇（H. E. Dodge）和罗米格（H. G. Romig）联合提出了在破坏性试验情况下采用的“抽样检验表”。二次大战初期，美国大批民用品转入军

工生产，由于事先无法控制废品而不能满足交货期要求，又由于军工生产多属破坏性试验，全数检验不可能也不允许。美国国防部为了解决这一难题，邀集休哈特、道奇、罗米格以及美国材料与试验协会、美国标准协会、美国机械工程师协会等有关人员研究，于 1941～1942 年先后公布一系列“美国战时质量管理标准”，要求各公司普遍实行统计质量控制方法，结果半年内取得显著成效。后来统计质量控制取得了很大发展。

我国自从 1978 年从日本引进全面质量管理十多年取得了显著成效。纵观我国的质量管理发展历史，是由检验质量管理直跃全面质量管理，对数理统计方法的运用远不是像当年美国那样深入广泛，不少决策、设计、科研、生产、销售、服务部门在提出问题、解决问题、检查结果时有些人还不习惯于进行科学的数理解析。

为了普及数理统计基本知识并在生产实际中发挥作用，我们组织石化行业中具有实践经验的质量管理专家编写了这套《统计质量控制》系列丛书。本书共分十册，第一册是数据收集和整理，第二册概率和数理统计基础，第三册估计和检验，第四册控制图，第五册方差分析，第六册实验设计，第七册相关和回归分析，第八册抽样检验，第九册统计方法应用演示 50 例，第十册数表。

数理统计方法就是通过对生产实践中大量数据的收集、整理、解析，研究生产实际中的内在规律的数学方法。和目前国内其它有关数理统计的书籍相比，本系列丛书的显著特点：一是它不同于一般的数学教科书，特别突出了实际应用，因此在编写中尽量减少不必要的公式推导，是一本实用性较强的书籍；第二个特点是书中列举了大量社会和生产（特别是石油化工生产）实例，文章从实例引出理论，又从理论回到实例，便于读者理解和应用，适合于工业企业特别是石油化工等流程型行业设计、研究、生产、销售、辅助等系统技术人员和管理干部学习参考；第三是语言既通俗易懂，又有一定深度和广度，既可用于中等水平人员学习应用，又可适用于高等水平技术人员研究参考。

为了更好应用本书，建议学习中注意几点：一是随着计算机的高度发展，许多数理统计方法可完全不需用手工计算，即可很快得出结果，已经掌握了统计方法的人可直接借助计算机，但对于初学之人，还是先用手算为好，防止知其然而不知其所以然，不利于在实践中灵活运用；二是对于现场技术人员，不要去深究公式推导，只要求会实际灵活运用；三是统计方法只提供解决问题的手段，必须和固有技术相结合才能解决问题，因此要使读者学会用数学的思维考虑专门技术问题；四是质量管理所用的方法不限于数理统计方法，还包括许多其它方法，如价值分析（VA）、生产工学（IE）、操作研究（OR）、价值工程（VE）、可靠性工程（RE）等，本书这次没有列入，读者可根据需要深入研究，灵活运用。

徐英宝

1995 年 1 月

目　录

1　数据收集和整理

例1　某油田为调查输出原油含水情况，在其外输原油中，连续抽取了10个罐次，其含水分别是0.11，0.10，0.09，0.15，0.14，0.12，0.11，0.13，0.10，0.10（单位是%）。求其均值，并标出各个数出现的频数。

解：

将数字按大小排序排列，并标出各个数出现的频数。

x_i（含水%）	0.09	0.10	0.11	0.12	0.13	0.14	0.15
f_i（频数）	1	3	2	1	1	1	1

（1）求均值 $\bar{x}=\dfrac{\sum f_i x_i}{n}$

其中　$n=10$

$$\sum f_i x_i=0.09\times1+0.10\times3+0.11\times2+0.12\times1+0.13\times1+0.14\times1+0.15\times1=1.15$$

$\therefore\quad \bar{x}=\dfrac{1.15}{10}=0.12$（%）

（2）求中位数 $\tilde{x}=\dfrac{0.11+0.11}{2}=0.11$（%）

（因10个数按大小顺序排列，第5、6个数都是0.11）

（3）众数为0.10，因出现0.10的频数最大。

例2　某厂生产的SD级内燃机油的Ca、Mg、Zn含量测定值如表2-1所示，计算其均值和标准偏差。

解：

为了理解计算原理和步骤，现用手工计算。

（1）将数据变换，简化计算方法：

将数据按下式简化：　$X_i=(x_i-x_0)\times h_i$

设 x_1，x_2，x_3 分别表示Ca、Mg、Zn含量值，

表 2-1　SD 内燃机油金属含量测定值

金属元素 \ 测定次数	1	2	3	4	5	6	7	8
Ca　%	0.345	0.319	0.320	0.318	0.321	0.313	0.323	0.324
Mg　ppm	7.4	5.9	10.9	7.1	7.8	5.0	8.3	8.4
Zn　%	0.104	0.102	0.106	0.106	0.105	0.107	0.110	0.109

令　$x_{10}=0.320$　　$h_1=1000$

$x_{20}=8.0$　　$h_2=10$

$x_{30}=0.100$　　$h_3=1000$　则：

Ca 含量值：$X_1=(x_{1i}-x_{10})\times h_1=(x_{1i}-0.320)\times 1000$

Mg 含量值：$X_2=(x_{2i}-x_{20})\times h_2=(x_{2i}-8.0)\times 10$

Zn 含量值：$X_3=(x_{3i}-x_{30}\times h_3=(x_{3i}-0.100)\times 1000$

数据变换后变为表 2-2：

表 2-2　X_i 值

	1	2	3	4	5	6	7	8	$\sum$
Ca X_1	25	−1	0	−2	1	−7	3	4	23
Mg X_2	−6	−21	29	−9	−2	−30	3	4	−32
Zn X_3	4	2	6	6	5	7	10	9	49

将 X_i 平方后得表 2-3：

表 2-3　X_i^2 值

	1	2	3	4	5	6	7	8	$\sum$
Ca X_1^2	625	1	0	4	1	49	9	16	705
Mg X_2^2	36	441	841	81	4	900	9	16	2328
Zn X_3^2	16	4	36	36	25	49	100	81	347

$$X_1\text{ 平方和 }S_{X_1}=\sum X_{1r}^2-\frac{(\sum X_{1r})^2}{n}=705-\frac{23^2}{8}=638.9$$

$$X_2\text{ 平方和 }S_{X_2}=\sum X_{2r}^2-\frac{(\sum X_{2r})^2}{n}=2328-128=2200$$

$$X_3\text{ 平方和 }S_{X_3}=\sum X_{3r}^2-\frac{(\sum X_{3r})^2}{n}=347-300=47$$

（2）求无偏方差 V：

$$V_{X_1}=\frac{S_{X_1}}{n-1}=\frac{638.9}{7}=91.27$$

$$V_{X_2}=\frac{S_{X_2}}{n-1}=\frac{2200}{7}=314$$

$$V_{X_3}=\frac{S_{X_3}}{n-1}=\frac{47}{7}=6.71$$

（3）求平均值 $\bar{x}$：

$$\bar{x}_i=x_0+\frac{\sum X_i}{n}\times\frac{1}{h}$$

n 为数据个数。

$$\therefore\quad \bar{x}_{1r}=0.320+\frac{23}{8}\times\frac{1}{1000}=0.3228$$

$$x_{2r}=8+\frac{-32}{8}\times\frac{1}{10}=7.60$$

$$\bar{x}_{3r}=0.100+\frac{49}{8}\times\frac{1}{1000}=0.1061$$

（4）求原始数据的方差：

将数据变回原来的状态

$$V_{x_1}=S_{X_1}^2\times\frac{1}{10^{3\times2}}=0.00009127$$

$$V_{x_2}=S_{X_2}^2\times\frac{1}{10^2}=3.14$$

$$V_{x_3}=S_{X_3}^2\times\frac{1}{10^6}=0.00000671$$

（5）求标准偏差 $\sqrt{V}$：

$$\sqrt{V_{x_1}}=\sqrt{0.00009127}=0.0095$$

$$\sqrt{V_{x_2}}=\sqrt{3.14}=1.77$$

$$\sqrt{V_{\hat{x}}}=\sqrt{0.00000671}=0.00259$$

（6）说明：

①求平均值，当 $n>5$ 时取值比测定值多1位，n 在70以上时可比测定值多取2位。

②x 标准偏差取位数和均值一样，实用中已足够。

③用计算器计算过程中可不考虑位数，其最后结果取到上述足够位数。

④平均值和标准偏差计算较复杂，会产生修约数字的误差，如平方和是根据 $S=\sum(x_i-\bar{x})^2$ 定义。如按此式计算修约误差会产生 n 次，若用 $S=\sum x_i^2-\frac{(\sum x_i)^2}{n}$ 计算，用 n 去除只有一次，修约误差小，将 $\frac{(\sum x_i)^2}{n}$ 称为修正项。

⑤数字个数 $n\geqslant100$ 时，求 $\bar{x}$ 和 $\sqrt{V}$ 可用直方图法。

⑥由于这里求出的值 $\bar{x}$、V、S 均为统计量，是从同一总体中抽取出样本测得的数据，因此是一变量，为区别于总体平均值和标准偏差，称为样本均值，样本标准偏差。总体的平均值 μ、方差 σ^2 和标准偏差 σ，可分别用样本的值估计，总体均值、方差和标准偏差的估计值分别用 $\hat{\mu}$、$\hat{\sigma}^2$、$\hat{\sigma}$ 表示时，

$$\hat{\mu}=\bar{x}$$

$$\hat{\sigma}^2=V$$

$$\hat{\sigma}=\frac{\sqrt{V}}{C_2}$$

$n=8$　$C_2=0.9650$（可查数表）

⑦S/n 称为样本方差，其平方根称为样本标准偏差，在统计书中及电子计算器经常看到，请注意。

⑧n 值小，μ 的估计值精度低，方差和标准偏差的精度也低，精度是指估计值的偏差程度。在这里估计的值是点估计，此外还可进行区间估计，在以后例题经常碰到。

⑨不偏方差用 $S/n-1$ 计算，式中 $n-1$ 称为自由度，用 ν 表

示，也可用 f、φ 字母表示。

例 3 影响污水处理 COD 去除率的因素分析。某 QC 小组成员用小组讲座的方式从 $4M1E$ 五因素入手找出了一系列影响 COD 去除率的原因。为了找出影响去除率的主要原因，以便针对主要问题制定对策，解决问题。试用排列图解析（兰化梁兵）。

表 3-1 COD 去除率低原因统计表（讨论方式）

序 号	项 目	频 数	累积数	累积%
1	DO 低	62	62	53.9
2	COD 波动	33	95	82.6
3	SS 波动	12	107	93.0
4	CN 高	5	112	97.4
5	操作失误	1	113	98.3
6	其它	2	115	100.0

作排列图如图 3-1。

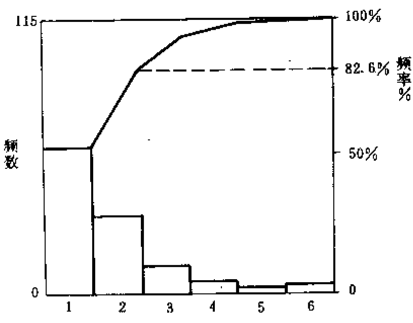

图 3-1 COD 去除率低原因排列图

从排列图中把影响 COD 去除率的原因分为了三类，即 A 类、

B 类、C 类，A 类因素是造成 COD 去除率低的主要原因，为 DO（溶解氧）低和 COD 波动，占原因的 80%以上，B 类因素是次要原因，为 SS 波动大，其它均为 C 类就是一般因素。所以，我们应该重点针对 DO 低和 COD 波动的问题，制定出切实可行的对策措施，加以实施，从而提高 COD 的去除率，这个问题解决了，就等于问题解决了 80%。

例 4 求下列数据的平均值 $\bar{x}$、中位数 $\tilde{x}$、极差 R、偏差平方和 S、无偏方差 V 和标准偏差 $\sqrt{V}$。

数据：21.5，22.1，20.0，23.6，21.1，20.7

解：

这本是非常简单地问题，但为了理解其意义及计算方法，我们按步骤解析。

（1）求中位数 $\tilde{x}$：

将数据按大小次序排列：$n=6$

20.0，20.7，21.1，21.5，22.1，23.6

$$\tilde{x}=\frac{21.1+21.5}{2}=21.3$$

（2）求平均值 $\bar{x}$：

为了解数据变换对结果的影响，将数据变换简化后，再进行计算。数据少的情况下似乎毫无意义，但数据多了就显得简单多了。

令 $X_i=(x_i-20)\times 10$

则							
x_i	21.5	22.1	20.0	23.6	21.1	20.7	$\sum$
X_i	15	21	0	36	11	7	90

$$\therefore\quad \bar{X}=\frac{90}{6}=15$$

$$\bar{x}=20+\bar{X}\times\frac{1}{10}=20+15\times\frac{1}{10}=21.5$$

（3）求偏差平方和 S：

按定义，$S=\sum(x_i-\bar{x})^2=(21.5-21.5)^2$
$+(22.1-21.5)^2+(23.6-21.5)^2$

$$+(21.1-21.5)^2+(20.7-21.5)^2$$

用简化计算式为：

$$S=\sum x_i^2-\frac{(\sum x_i)^2}{n}$$

$\sum x_i^2=21.5^2+22.1^2+20.0^2+23.6^2+21.1^2+20.7^2=$ 2781.32

$$\frac{(\sum x_i)^2}{n}=\frac{(21.5+22.1+20.0+23.6+21.1+20.7)^2}{6}$$

$$=\frac{129^2}{6}=2773.5$$

$\therefore\quad S=2781.32-2773.5=7.82$

以上计算比较复杂，如用变换后数字计算，

$$S=\left(\sum X_i^2-\frac{(\sum X_i)^2}{n}\right)\cdot\frac{1}{10^2}=\left(2132-\frac{90^2}{6}\right)\frac{1}{100}=7.82$$

比上述计算方法简单多了。

(4) 求无偏方差 V：

$$V=\frac{S}{n-1}=\frac{7.82}{5}=1.564$$

(5) 求标准偏差 $\sqrt{V}$：

$$\sqrt{V}=\sqrt{1.564}=1.25$$

(6) 求极差 R：

$$R=x_{max}-x_{min}=23.6-20.0=3.6$$

以上均为手算方式，实际上目前计算器很普遍，只要用带有统计功能的计算器，将计算器置于统计功能下，把 6 个数据一一输入，则可立即求出上述各数据。

例 5 在某恒温装置入厂检查中，其测定精度是十分重要的，由于精度问题经常产生和出厂合格证不同的结果。为减少争议，要进行测定误差解析。对同一产品重复测定 3 次，计算出 $\bar{x}$ 和 R 值列于表 5-1，然后作成图 5-1 控制图，试解析。

表 5-1 $\bar{x}$—R 值 单位:℃

时间 日	1	2	3	4	5	6	7	8	9	10
$\bar{x}$	92	110	95	122	111	105	104	79	139	104
R	15	14	22	25	32	10	17	8	41	41
时间 日	11	12	13	14	15	16	17	18	19	20
$\bar{x}$	130	98	108	96	97	96	110	94	90	112
R	40	23	13	13	23	21	15	16	15	17
时间 日	21	22	23	24	25	26	27	28	29	30
$\bar{x}$	80	90	90	112	86	80	78	108	103	122
R	1	14	22	25	17	8	8	8	23	20

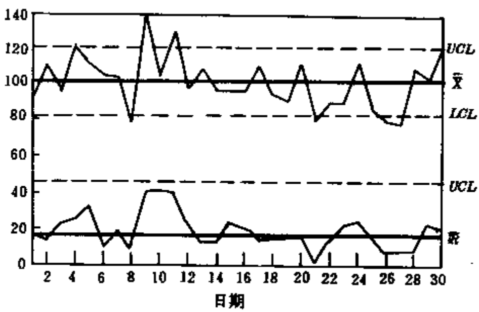

图 5-1 测定误差的 $\bar{x}$—R 控制图

解:

从控制图中可以看出,R 控制图没有发现控制限以外的点,可以认为测定误差在可控状态下。

$\overline{R}$ 为极差均值,从图 5-1 中可知 $\overline{R}=18.9$,$n=3$,查数表 38 可知 $d_2=1.693$,所以操作温度测定误差:

$$\hat{\sigma}_M=\frac{\overline{R}}{d_2}=\frac{18.9}{1.693}=11.164$$

从 R 的上控制线为 $D_2\hat{\sigma}_M=4.358\times11.164=48.653$(查数表 $D_2=4.358$),$\bar{x}$ 控制图中的 $\bar{x}$ 均值,显示了一个个样品值的变化,

$\bar{x}$ 大幅度变动显示了工序发生异常，从考核测定精度来说这是好现象，如果分组后 $\bar{x}$ 控制图处于管理状态时，说明测定误差过大，意味着不能发现工序的异常变化，因此有必要改善缩小测定误差。

x 控制图控制界限为 $\bar{x} \pm A_2\bar{R} = 101.4 \pm 1.023 \times 18.9 = 120.6，82.1$

（A_2 从数表 38 可查到）。

〔说明〕

我们日常所取得的数据，其中包含了测定误差，如将 σ_p 定为工序变动，σ_M 定为测定误差，测定值的方差应是：$\sigma_X^2 = \sigma_p^2 + \sigma_M^2$。欲求测定误差的大小，必须如图 5-1 所示，用同一样品重复测定 2～5 次，将这些数据为一组，作 $n=2\sim5$ 的 $\bar{x}$--R 控制图，R 控制图在管理状态下即可控状态下，可判定测定误差在较好地可控状态下，其测定误差估计值可用下式求出：

$$\hat{\sigma}_M = \frac{\bar{R}}{d_2}$$

d_2 根据 n 的大小查数表 38 可得。

一般，测定误差 σ_M 必须是工序偏差 σ_p 的 1/3 以下，因此如 $\bar{x}$ 控制图中有 k 个点，应有一半以上的点（$n=2$ 时为 $0.49k$ 个，$n=3$ 时 $0.57k$ 个，$n=4$ 时为 $0.62k$ 个，$n=5$ 时为 0.66 个）在 $\bar{x}$ 控制限以外，本例 $n=3$，在 $\bar{x}$ 控制图中，$0.57k = 0.57 \times 30 = 17.1$，即应有 18 个点在界限外较好，但图 5-1 中的界外点只有 8 个，可认为测定误差过大。

如果 R 控制图在非可控状态下，表示测定不正常，必须将测定仪器和方法进行标准化。

例 6 某厂生产一种设备，每天生产 200 台，这种设备分 Ⅰ 型和 Ⅱ 型 2 种，经过两个月的生产，组装工程中的不合格品的修理仍没有减少。为了减少不合格品的修理，最近一个月，调查了不合格品内容和缺陷项目，如表 6-1 所示。根据调查内容，到底应向什么方向努力才能减少不合格品？请用排列图、控制图解析。