SQC-2

统计质量控制

STATISTICAL QUALITY CONTROL

概率和数理统计基础

Basic of Probability and Mathematical statistics

陈国铭 主编
郭耀曾 编

中国石化出版社

SQC-2

统计质量控制

STATISTICAL QUALITY CONTROL

概率和数理统计基础

Basic of Probability and Mathematical statistics

陈国铭　主编
郭耀曾　　编

中国石化出版社

(京) 新登字 048 号

Z619/07

内 容 提 要

本册作为《统计质量控制》丛书的数学基础，主要介绍了概率论和数理统计的基本知识和线性方程组的解法。

前一部分介绍了随机事件的概率及常用的计算公式，随机变量的分布及常用的离散分布和连续分布，随机变量的数字特征，以统计量为主的随机变量函数的分布等内容。后一部分则介绍了线性方程组的行列式及矩阵解法，还介绍了表算法。这一部分是线性回归理论的数学基础。

本书主要突出实际应用，尽量避免复杂的数学推导，并将行列式和矩阵以及统计量分布中要用到的 Γ 函数和 B 函数，作为附录提供给读者，以便需要时查阅。

本书可供质量管理干部和企业工程技术人员使用。

SQC-2

统计质量控制

STATISTICAL QUALITY CONTROL

概率和数理统计基础

Basic of Probability and Mathematical Statistics

陈国铭 主编

郭耀曾 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 6 印张 157 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 7-80043-555-5/O·020 定价：7.00 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ISO900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. S hewhart）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”。二次大战时期，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，特别突出了实际实用，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。

1995 年 1 月

目　　录

1　预　备　知　识

1.1　加法原理和乘法原理

1.1.1　加法原理

例 1-1　从兰州到北京可以乘坐火车或飞机，火车一天有 3 班，飞机一天有 2 班。问某人自兰赴京可以有几种不同的走法？

解：若乘火车有 3 种走法（3 个不同的火车班次），若乘飞机有 2 种走法（2 个不同的航班），不论乘火车或飞机均可到达目的地，故一共可以有

$$3+2=5$$

种不同的走法。

这类问题在现实生活中经常遇到，人们归纳出解决这类问题的加法原理（*principle of addition*）。

加法原理　完成某一事件有 m 类办法，在第一类办法中有 n_1 种方法，在第二类办法中有 n_2 种方法，……，在第 m 类办法中有 n_m 种方法。任选一种方法都可以完成这一事件，那么完成该事件共有

$$M=n_1+n_2+\cdots\cdots+n_m$$

种不同的方法。

1.1.2　乘法原理

例 1-2　欲从北京经兰州联系工作后赴长庆油田。从北京到兰州可乘火车或飞机，从兰州到长庆油田可乘火车、汽车或飞机，若不考虑火车或飞机的班次，问有几种不同方法的走法？

解：由于不考虑班次，故乘同一种交通工具算一种方式。

从北京到兰州可以有两种方式，即乘火车和乘飞机。不论采用哪种方式到兰州后，再去长庆油田均有三种方式，即乘汽车、火车或飞机，故从北京到长庆油田有

$$2\times3=6$$

种不同的方式，即

火车——火车；　　飞机——火车；

火车——飞机；　　飞机——飞机；

火车——汽车；　　飞机——汽车；

对这类常见的问题，人们归纳出乘法原理（*principle of multiplication*）：

乘法原理　完成某一事件有 m 个步骤（阶段），完成第一步（阶段）有 n_1 种方法，完成第二步（阶段）有 n_2 种方法，……，完成第 m 步（阶段）有 n_m 种方法，必须完成这 m 个步骤（阶段）才能完成这一事件。那么，完成这一事件一共可能有

$$M=n_1\times n_2\times\cdots\cdots\times n_m$$

种不同的方法。

1.1.3　$\sum$ 和 Π

为了使加法原理和乘法原理的表达式加以简化，引进两个记号，它们在高等数学中经常使用，即 $\sum$ 和 Π。

1) $\sum$　$\sum$ 称为总和号，它是若干个数求和时的简单表示方法。使用该符号时，往往呈下列形式：

$$\sum_{i=1}^{n} n_i$$

其中，n_i 表示第 i 个加数；

$i=1$ 表示从第一个加数开始求和；

n 表示求和的加数到第 n 个为止。

在意义明确不致发生混淆时，也简写为"$\sum$"。

例 1-3　有一列数 x_1，x_2，……x_n，问 $\sum_{i=1}^{n}x_i$，$\sum_{i=1}^{18}x_i$，$\sum_{i=5}^{15}x_i$ 各表示什么意义？

解：$\sum_{i=1}^{n}x_i=x_1+x_2+\cdots\cdots+x_n$，即一列数全相加；

$\sum_{i=1}^{18}x_i=x_1+x_2+\cdots\cdots x_{18}$，即从第一个数加到第 18 个数

的和；

$\sum_{i=5}^{15} x_i = x_5 + x_6 + \cdots\cdots x_{15}$，即从第 5 个数开始，加到第 15 个数为止的 11 个数的和。

2）Π　Π 称为连乘积符号，它是若干个数连续相乘时的简单表示方法。使用该符号时，往往呈下列形式：

$$\prod_{i=1}^{n} n_i$$

其中，n_i 表示第 i 个因子；

$i=1$ 表示从第一个因子开始求积；

n 表示求积时乘到第 n 个因子为止。

例 1-4　有一列数 $x_1, x_2, \cdots\cdots, x_n$ 问 $\prod_{i=1}^{n} x_i$，$\prod_{i=1}^{18} x_i$ 和 $\prod_{i=5}^{15} x_i$ 各表示什么意义？

解：$\prod_{i=1}^{n} x_i = x_1 + x_2 + \cdots\cdots + x_n$，即一列 n 个数的连乘积；

$\prod_{i=1}^{18} x_i = x_1 \times x_2 \times \cdots\cdots \times x_{18}$，即从第 1 个因子开始，连续 18 个因子的乘积；

$\prod_{i=5}^{15} x_i = x_5 \times x_6 \times \cdots\cdots \times x_{15}$，即从第 5 个因子开始到第 15 个因子为止，连续 11 个数的乘积。

例 1-5　（综合运用加法原理和乘法原理的例），某厂在调整产品结构适应市场需求的过程中准备上一项技措，该厂从设计到施工均有能力完成；若考虑外委设计，有 3 家设计院愿意接受，若考虑外委施工，有 4 家建筑安装单位准备投标。问完成该项技措一共可能有多少种不同的设计、施工方案？

解：完成此项技措共有 3 类不同的办法：自行承担设计和施工；自行承担和外委相结合；完全外委。

第一类办法，即该厂自行设计，自行施工，只有这样 1 种方法；

第二类办法，即自行承担和外委相结合，又可分为自行设计

外委施工和自行施工外委设计两种情况，前者有

$$1\times4=4$$

种方法，后者有

$$3\times1=3$$

种方法，故第二类办法共有

$$3+4=7$$

种方法。

第三类办法，即完全外委，全部技措分两个阶段，第一阶段是设计，可以有 3 种方法，第二阶段是施工，可以有 4 种方法，故共有

$$3\times4=12$$

种不同的方法。

综合三类办法，完成此技措总共可能有

$$1+7+12=20$$

种不同的方法。

1.2 排列

1.2.1 排列的概念和类型

排列（*permutation*）系指 n 个元素，将其中一部分或是全体按不同的顺序排起队来，一共有多少种不同排法的问题。这是在日常生活中经常遇到的问题。

按其元素是否全部参加排列分为全排列（*perfect permutation*）和选排列（*partial permutation*）。前者是 n 个元素全部拿来排除，而后者是将 n 个元素中的 k（$k<n$）个元素拿来进行排队。按其元素是否允许重复出现分为重复排列（*permutation with repetition*）和无重复排列（*permutation without repetition*），重复排列的情况下一个元素允许出现多次，而无重复排列的情况下一个元素只允许出现一次。

1.2.2 无重复全排列的计算

例 1-6 某化工厂有 8 套装置，一质量检查员欲对这 8 套装置

的控制质量进行巡回检查，每套装置只去1次，不计起点，问有多少种可能的检查路线。

解：由于不计起点，则该检查员第一个检查的装置有8套可供选择，即有8种可能；由于每套装置只去1次，故该检查员第二套检查的装置就剩下7种可能了（还有7装置没去），欲检查的第三套装置有6种可能，第四套装置有5种可能，……，欲检查的第八套装置只有1种可能。根据乘法原理，该检查员的检查路线共有

$$8\times7\times6\times\cdots\cdots\times2\times1=40320\ (\text{种})$$

如果把8套装置看作8个元素，这些元素全部参加排列，且每个元素（每套装置）只出现1次，这实质上是$n=8$的无重复全排列的问题。按照与本例题同样的考虑，可知：n个元素的无重复全排列一共有

$$n\times(n-1)\times(n-2)\times\cdots\cdots3\times2\times1$$

种不同的排列方法。上式是从自然数n开始，每次减1，n个连续自然数的乘积，我们称它为n的阶乘，记作，

$$n!=\prod_{k=1}^{n}k=n\times(n-1)\times(n-2)\times\cdots\cdots\times3\times2\times1$$

若用P_n表示n个元素无重复全排列的排列数，则有

$$P_n=n! \tag{1-1}$$

为了运算的需要，规定$0!=1$，且有

$$P_1=1!=1$$

$$P_2=2!=2$$

$$P_3=3!=6$$

$$P_4=4!=24$$

$$P_5=5!=120$$

$$P_6=6!=720$$

$$P_7=7!=5040$$

$$P_8=8!=40320$$

$$P_9=9!=362880$$

$$P_{10}=10\,!=3628800$$

从上述数据可知，随着 n 的增大，$n!$ 迅速变大，它的增加速度很快，因此计算起来相当麻烦。为了计算的方便，给出阶乘的近似计算公式，即斯特林公式（Stirling）：

$$n!\sim\sqrt{2n\pi}\left(\frac{n}{e}\right)^{n} \tag{1-2}$$

式中“～”表示“等价于”，即当 $n\longrightarrow\infty$ 时，符号两侧数量的比值 $\longrightarrow 1$，π 和 e 是两个最重要的无理数，π 为圆周率，其取值为 3.14159……，e 为自然对数的底，取值为 2.71828……。

用斯特林公式计算 $n!$，随着 n 的变大，其绝对误差可以变得很大，而相对误差却可以变得很小。例如 $n=10$，用斯特林公式计算

$$10\,!=\sqrt{2\times10\times3.14159}\left(\frac{10}{2.71828}\right)^{10}\doteq3598600$$

直接计算为：

$$10\,!=3628800$$

二者的绝对误差为 30200，相对误差为 0.8%；当 n=100 时，两种方法计算结果的绝对误差为：

$$7.8\times10^{154}$$

而相对误差却只有 0.084%。

1.2.3 无重复选排列的计算

例 1-7 某化工厂有 8 套装置，一检查员欲对其中任意 4 套装置进行抽检，不计起点，问有多少种可能的抽检路线？

解：由抽检的任意性可知，8 套装置中任何一套均可作为抽检的第一站，故有 8 种可能；抽检的第二站是剩余 7 套装置中的任何一套，故有 7 种可能；第三站有 6 种可能；第四站有 5 种可能。根据乘法定理知，总共可能有

$$8\times7\times6\times5=1680$$

条不同的抽检路线。

本例属于无重复的选排列问题。这类问题的一般提法是：从

n 个不同的元素中任意选出 r 个元素进行排列，任何元素在每个排列中都只允许出现一次（即不允许重复出现），有多少种不同的排法。例 1-7 即为从 8 套装置中选出 4 套进行选排列的问题。

无重复的选排列问题之解法可如下讨论：

当 $r=1$ 时，即只选一个元素进行排列。这时选出的元素只有 1 种排列方法，而可供选择的元素共有 n 个，故 $r=1$ 时一共有 n 种不同的排列方法。

当 $r=2$ 时，即从 n 个元素中选出 2 个元素进行排列。这时取第一个元素时有 n 种选法，取第 2 个元素时有 $n-1$ 种选法（因为不允许重复，选过的元素不能再选了），根据乘法定理，这时一共有

$$n\times(n-1)=n\times[(n-2)+1]$$

种不同的排列方法。

当 $r=3$ 时，根据类似的讨论知，总共有

$$n\times(n-1)\times(n-2)=n\times(n-1)\times[(n-3)+1]$$

种不同的排列方法。

………

当 $r=r$ 时，不同的排列数即为

$$n\times(n-1)\times(n-2)\times\cdots\cdots\times[(n-r)+1]$$
$$=n(n-1)(n-2)\cdots\cdots(n-r+1)$$

于是，从 n 个不同元素中选取 r 个元素无重复选排列的排列数为：

$$P_n^r=n(n-1)(n-2)\cdots\cdots(n-r+1)$$
$$=\frac{n(n-1)\cdots\cdots(n-r+1)(n-r)(n-r-1)\cdots\cdots\times3\times2\times1}{(n-r)(n-r-1)\cdots\cdots\times3\times2\times1}$$
$$=\frac{n!}{(n-r)!} \tag{1-3}$$

选排列的符号有时也用 A_n^r。

特别地，当 $r=n$ 时，即从 n 个不同的元素中选出 n 个元素进行排列，亦即 n 个不同元素的全排列，这时由式（1-3）有：

$$P_n^n=\frac{n!}{(n-n)!}=\frac{n!}{0!}=n!\ =P_n$$

即为全排列的计算公式（1-1）。

用式（1-3）计算例1-7，得

$$P_8^4=\frac{8\,!}{(8-4)\,!}=8\times7\times6\times5=1680$$

条不同的抽检路线。

1.2.4 重复排列

例1-8 北京市的电话号码是7位数，问一共可以安排多少部不重号的7位号码的电话？

解：本例与例1-6和例1-7均不同。那两个例子都属于无重复的排列，即在某个特定的排列中，任何元素都至多允许出现1次（也可以不出现），而本例在某个特定的排列中允许元素多次出现，例如中国石化总公司的电话号码为4225533，其中数字（元素）2，3，5都出现了两次，数字排列顺序不同则代表不同的电话所持有的号码。

电话号码的第一位数字可从0，1，2，……9共10个数字中任选1个，即有10种可能；由于允许元素重复，故第二位数字也有10种可能；……第七位即最后一位数字同样有10种可能。根据乘法原理，总共可以有

$$10\times10\times10\times\cdots\cdots\times10=10^7$$

个不同的号码，即可容纳$10^7=1000$万部电话。

本例属于允许重复的选排列的类型。这种类型排列问题的一般提法是：

从n个不同的元素中选择r个元素进行排列，允许同一元素在某一个特定的排列中重复出现，有多少不同的排列方法。

这时，任何一个位置均有n种可能，故共有

$$\underbrace{n\cdot n\cdot n\cdots\cdots n}_{r\text{个}}=n^r \qquad (1-4)$$

不同的排法。式（1-4）即为允许重复选排列的排列数之计算公式。

特别有，$r=n$时，即允许重复全排列的排列数为n^n。

1.2.5 元素不尽相异的全排列