

质量改进统计方法

ZHILIANG GAIJIN TONGJI FANGFA

孙顺利 编著



冶金工业出版社
Metallurgical Industry Press

质量改进统计方法

孙顺利 编著

北 京

冶金工业出版社

2015

内 容 提 要

本书以现场统计数据为基准,系统讲解分层法、调查表法、排列图法、因果分析图法、散布图法、直方图法和控制图法等七种统计方法的特点、目的、适用时机、注意事项和应用案例,为学生掌握相关知识提供参考,为生产现场的质量改进活动提供简捷而有效的支持。

本书可供工业工程等专业的师生使用,也可供从事相关专业的工程技术人员参考。

图书在版编目(CIP)数据

质量改进统计方法/孙顺利编著. —北京:冶金工业出版社, 2015. 10

ISBN 978-7-5024-7065-4

I. ①质… II. ①孙… III. ①生产统计—统计方法
IV. ①F273

中国版本图书馆 CIP 数据核字(2015)第 241399 号

出 版 人 谭学余

地 址 北京市东城区嵩祝院北巷 39 号 邮编 100009 电话 (010)64027926

网 址 www.cnmp.com.cn 电子信箱 yjcbcs@cnmp.com.cn

责任编辑 郭冬艳 美术编辑 彭子赫 版式设计 孙跃红

责任校对 郑 娟 责任印制 李玉山

ISBN 978-7-5024-7065-4

冶金工业出版社出版发行;各地新华书店经销;固安华明印业有限公司印刷

2015 年 10 月第 1 版, 2015 年 10 月第 1 次印刷

169mm×239mm; 11.25 印张; 218 千字; 170 页

34.00 元

冶金工业出版社 投稿电话 (010)64027932 投稿信箱 tougao@cnmp.com.cn

冶金工业出版社营销中心 电话 (010)64044283 传真 (010)64027893

冶金书店 地址 北京市东四西大街 46 号(100010) 电话 (010)65289081(兼传真)

冶金工业出版社天猫旗舰店 yjgycbs.tmall.com

(本书如有印装质量问题,本社营销中心负责退换)

前 言

基于数据的质量改进是当代质量管理领域最活跃的理论研究和实践应用类方法和工具之一。随着全面质量管理在全球的开展,国际标准化组织制定了质量管理与质量保证的国际标准,即 ISO9000 系列标准。ISO 将统计方法从其标准 ISO9001:1994 (4.20 统计技术) 中的一个要素提升为 ISO9001:2000 (2.10 统计技术的作用) 质量管理体系基础之一。这一变化反映了质量管理的先进水平。我国也举办了质量工程师资格考试。在企业层面,统计方法在质量管理中的应用愈来愈受到重视,不少企业开展了六西格玛质量改进活动,对统计方法的要求愈来愈迫切。

由于我国大多数组织的质量管理跨越了统计质量管理阶段,在推行全面质量管理中又未能抓住时机补上一课,所以在许多组织中统计技术的应用一直是一个普遍存在的薄弱环节,在一定程度上制约了我国企业质量管理水平的提高。ISO9001:2000 标准中十分明确指出了“统计技术的作用”。应用统计技术可帮助组织了解变异,进而对这类变异进行测量、描述、分析、解释和建立模型。这种数据的统计分析能对更好地理解变异的性质、程度和原因提供帮助。从而有助于解决,甚至防止由变异引起的问题,并促进持续改进。

本书以现场统计数据为基础,在内容安排上既力求科学、系统、完整,又有所侧重,系统讲解分层法、调查表法、排列图法、因果分析图法、散布图法、直方图法和控制图法,这七种统计方法的特点、目的、适用时机、注意事项和应用案例,为生产现场的质量改进活动提供简捷而有效的支持。

本书有几个显著的特点:一是强调基础性、可操作性和实用性,

对通常的质量管理的七种统计方法作了十分详细的讲解,力图使读者掌握、会用且能够在现场条件下有效地去应用;二是浅显易懂,将统计知识和现场工程问题紧密结合,略去不必要的理论证明,可读性强;三是结合质量统计软件 MINITAB,在给出每种统计方法软件操作步骤的同时,对输出结果也进行了明确的统计解释,适合没有系统学习过统计学的现场技术人员使用。

本书可以作为工业工程专业统计学、质量管理等课程的实验教学材料,也可以作为各类制造企业、服务企业等各类企业进行现场质量改进的指南或员工培训教材。

本书是河南理工大学 2013 年教育教学改革研究项目 (2013JG058) 和 2014 年教育教学改革研究项目 (2014JG065) 的阶段成果。在编写过程中参考了国内出版的许多质量管理和统计学方面的著作,得到了河南理工大学能源学院以及河南理工大学教务处有关领导和同仁的关心、鼓励和支持,在此表示衷心感谢。

由于作者水平和能力有限,书中难免存在疏漏,诚望读者和专家不吝指教。

作 者
2015 年 6 月

目 录

1 统计技术基础	1
1.1 引言	1
1.2 数据收集	1
1.2.1 总体与样本	2
1.2.2 数据的分类	3
1.2.3 数据收集的目的	4
1.3 质量变异的描述	4
1.3.1 系统因素与非系统因素	4
1.3.2 产品质量波动	6
1.4 质量管理中的常用分布	7
1.4.1 二项分布	7
1.4.2 泊松分布	10
1.4.3 超几何分布	12
1.4.4 正态分布	13
2 调查表与分层	16
2.1 调查表	16
2.2 调查表的用途和种类	16
2.3 调查表的应用步骤	16
2.4 注意事项	17
2.5 应用实例	18
2.5.1 缺陷位置调查表	18
2.5.2 不良项目调查表	19
2.5.3 不良原因调查表	20
2.5.4 质量分布调查表	20
2.6 分层	22
2.7 应用实例	22
3 排列图	25
3.1 排列图	25

3.2 应用场合	25
3.3 排列图的应用步骤	25
3.4 注意事项	27
3.5 排列图的应用	30
3.5.1 分析主要缺陷形式	30
3.5.2 分析不合格的原因	31
4 因果图	35
4.1 因果图	35
4.2 因果图的制作步骤	36
4.3 绘制因果图的注意事项	36
4.4 如何确定主要原因	37
4.4.1 逐条识别	37
4.4.2 确定程序	38
4.4.3 确定方法	38
4.5 确定主要原因不正确的方法	39
4.6 因果图的类型	40
5 相关与回归分析	42
5.1 相关与回归概念	42
5.1.1 确定性关系与相关关系	42
5.1.2 相关分析与回归分析	43
5.1.3 相关分析与回归分析的区别与联系	43
5.1.4 相关分析应注意的问题	44
5.2 相关分析	45
5.2.1 散布图	45
5.2.2 散布图的意义和作用	45
5.2.3 散布图的分析 and 判断	46
5.2.4 相关系数判别法	47
5.2.5 散布图的绘制	49
5.2.6 应用示例	50
5.3 一元线性回归	52
5.3.1 模型	52
5.3.2 回归系数的最小二乘估计及其性质	52
5.3.3 一元线性回归方程的显著性检验	55

5.3.4	一元线性回归方程总效果的度量	59
5.3.5	回归系数的显著性检验	61
5.3.6	回归方程的失拟性检验	63
5.3.7	回归方程残差诊断	65
5.3.8	利用回归方程进行预测	69
5.3.9	利用回归方程进行控制	75
6	直方图	76
6.1	数据分布	76
6.2	数据的波动性和规律性	76
6.3	直方图的原理	77
6.4	直方图的作用	77
6.5	直方图的作图步骤	78
6.6	直方图的应用示例	78
6.7	直方图的观察与分析	80
6.7.1	分布形状分析与判断	80
6.7.2	直方图与规格(标准)相比较	82
6.8	直方图的分层处理	85
7	控制图	88
7.1	统计过程控制简介	88
7.1.1	统计过程控制技术的发展	88
7.1.2	统计过程控制的理论基础	89
7.1.3	统计控制图的分类	90
7.2	统计控制图的原理	93
7.2.1	统计控制图的原理与构造	93
7.2.2	统计控制图的两类错误	94
7.2.3	判断过程正常的准则	95
7.2.4	判断过程异常的准则	96
7.2.5	常规控制图判断准则的使用	101
7.3	常规控制图作图方法	102
7.3.1	建立控制图的预备工作	102
7.3.2	计量值控制图应用步骤	104
7.4	应用统计控制图应注意的问题	104
8	常规计量值控制图	107
8.1	均值-极差控制图	107

8.1.1 $\bar{X}-R$ 控制图, 标准值给定	107
8.1.2 $\bar{X}-R$ 控制图, 标准值未给定	111
8.2 均值-标准差控制图 ($\bar{X}-s$ 控制图)	114
8.3 单值-移动极差控制图 ($\bar{X}-R_s$ 控制图)	117
8.4 中位数-极差控制图 ($Me-R$ 控制图)	119
9 常规计数值控制图	121
9.1 不合格品率控制图 (p 图)	121
9.2 不合格品数控制图 (np 图)	127
9.3 缺陷数控制图 (c 图)	129
9.3.1 c 控制图的原理	130
9.3.2 c 控制图的绘制	130
9.4 单位缺陷数控制图 (u 图)	132
9.4.1 u 控制图的原理	132
9.4.2 u 控制图的绘制	133
10 过程能力分析	136
10.1 过程能力和过程能力指数	136
10.1.1 过程能力	136
10.1.2 过程能力指数	138
10.1.3 过程不合格品率的计算	142
10.2 过程能力的评价	146
10.2.1 过程能力的判定	146
10.2.2 提高过程能力的对策	146
10.2.3 提高过程能力指数的措施	147
10.3 过程能力调查	148
10.3.1 过程能力调查流程	148
10.3.2 过程能力调查的方法	149
10.3.3 计量数据的过程能力分析	150
10.3.4 计点值数据的过程能力分析	154
10.3.5 计件值数据的过程能力分析	158
附 录	162
参考文献	170

1 统计技术基础

1.1 引言

任何企业，只要进行生产就必定存在产品质量问题。产品质量的优劣，不但直接关系到企业的信誉、经济收益和前途，并且在一定程度上也反映了该企业的技术水平和管理水平。当今，质量的概念已经渗透到各个领域。例如武器、弹药的质量关系到国防的防卫实力；药物、食品的质量关系到人们的身体健康；汽车、房屋、电器……的质量无不与人们的生产、生活紧密相关。质量是人们对完善、适用、精美、真实、可靠等含意的综合表述。

质量离不开数量。产品质量的提高，要用数量来表示；不合格品率的降低，也要用数量来表示；产品销售量、产品的规格、控制生产所用的管理标准、……，都要用数量来表示。现代工业生产的专业化方向早已到来。专业化的一个重要特点就是要求成批、大量的生产。企业生产量往往很多，一般以万、十万、百万来计算，……。面对产量巨大的生产，如何才能保证它的质量呢？旧有的办法是用大量的检验人员，对产品进行全数检查，这显然是不经济的十分落后的办法。在产品生产过程中，能否及时发现不良征兆，发出警报以确保生产出来的产品都是合格品呢？如果影响产品质量的因素很多，能否从中找出主要因素并及时进行控制呢？根据生产实际所积累的数据，能否对设计参数、工艺参数进行优选，以保证达到预期的目的呢？

实践证明，质量管理统计方法能够提高产品质量、降低成本、优选设计参数及工艺参数。所谓统计质量管理，就是由数据来说话，用统计方法处理过的统计数字来表明管理结果。事实上，产品质量反映在数据的波动上，现代工业企业管理的技巧，主要在于数据的运用。

1.2 数据收集

通过有目的的收集数据，运用数理统计的方法处理所得的原始数据，提炼出有关产品质量、生产过程的信息，再分析具体情况，作出决策，从而达到提高产品质量的目的。这就是通常所说的质量管理。

初期的质量管理，往往只是进行“事后把关”。全数检查即将出厂的产品，测试有关的质量特征指标，从中挑出不合格品，出厂合格品。这种方法，一方面

不能预防生产过程中不合格品的产生；另一方面，如果要知道每一产品或工序被破坏时的指标，全数检查的方法显然是行不通的。

随着生产的发展和产量的大幅度提高，迫切地要求用经济、可靠的方法解决企业产品的质量检查问题，这就导致把全数检查发展成抽样检验。这种检验不仅要最终产品进行抽样检验，以便对不合格品采取一定的措施，减少可能引起的损失，而且要在产品生产过程中及时做抽样检验，预防不合格品的出现，从而使生产过程保持稳定的状态。

什么是抽样检验呢？简而言之，抽样检验就是通过检验一部分产品而对一批产品的质量作出估计。只有合理地抽取样本，正确地运用统计方法处理数据，才能得到可靠的估计。关于抽样方法的具体内容，请参阅相关教材。我们在这里仅仅借助于抽样问题说明数理统计的基本概念和一些运算。

1.2.1 总体与样本

1.2.1.1 总体

“总体”一词，是统计学中常用的一个术语。一批产品、一台设备或在某段时间内生产的同类产品的全体等，都可以叫做一个总体。例如，某灯泡厂三月份生产的全部灯泡，某纺织厂生产的布匹，或一道工序加工后的半成品，制造产品的原材料等，都能成为一个总体。

构成总体的基本单位，叫做个体。这个基本单位又可叫做单位产品。单位产品有时可以很自然地划分出来，如果总体是一批电灯泡，那么其中的每只灯泡都可以看作一个个体。有些单位产品却不能自然地进行划分。比如，如果把一匹布作为一批产品，即叫做一个总体，那么这总体中的每个个体可以是一米布、十米布或者二十米布，在这种情况下，个体的划分需要根据具体问题而定。

所谓抽样检验，指的是从总体中抽取一部分个体，并测试被抽到的每个个体的有关质量特征的数量指标，得到一组数据，再对这些数据进行处理，然后对总体作出估计和判断。每个单位产品都有一个或一组表现其质量的数量指标与之对应。如电阻元件与阻值；电灯泡与其使用寿命；一米棉布与其上的疵点数。

总体中的每一个个体肯定要与某一个（组）数相对应，这是确定无疑的事实。这个数就是每个个体的质量表征，由于这个数的具体取值因个体的不同而异，因此通常称它为随机变量，并记作 X 。

综上所述，可以说总体是问题所涉及的全体对象，总体就是随机变量 X 的全体取值，也可以说总体就是随机变量 X 。

总体可以是有限的，也可以是无限的。一个工厂生产的电子管的数目是有限的，但这些电子管的寿命作为一个总体来看就是无限的。

1.2.1.2 样本

从一批产品中抽取一部分进行检验，被抽取的这一部分单位产品的全体，就此为试读，需要完整PDF请访问：www.ertongbook.com，

叫做一个样本。换句话说，样本就是从总体中抽取的一部分个体的全体。

例如，我们要了解即将出厂的一批螺钉的长度是否合格，今从中抽取 n 个个体 $X_1, X_2, \dots, X_n$ 。我们就称 $\{X_1, X_2, \dots, X_n\}$ 为总体的一个样本。这里的 n 通常称为样本容量或样本大小。

$X_1, X_2, \dots, X_n$ 分别表示 n 个螺钉的长度， n 个螺钉被抽取后，客观上就有 n 个长度与其对应，当长度未被测定时，则都用随机变量表示。待具体测定它们的长度后，所得值记作 $\{x_1, x_2, \dots, x_n\}$ ，称为样本值。

样本中的每一个个体叫做一个样品。今后，如果我们说抽取了样本大小为 n 的一个样本，实际上就是抽取了 n 个样品，也就是抽取了 n 个单位产品。如果在有限总体中，包含个体的总数为 N ，则常称 N 为总体的批量。

1.2.2 数据的分类

抽取到样本以后，我们就要测试每个单位产品的质量特征指标，并把测得的原始数据记录在表格内。单位产品的质量可以用不同的方法来衡量。有的用“计数”方法衡量，有的用“计量”的方法衡量。

1.2.2.1 计数值

描述一块布或一个铸件表面的质量，需要一块棉布上的疵点数、一个铸件表面的气孔数或砂眼数表达。这就是所谓“计点”值描述产品质量的方法。

存在着一些单位产品的质量特征不能用定量的方法去度量。如洗衣机外表面的漆层有剥落，就可以认为其外观不合格。有许多产品，其外观不能用定量的方法来衡量，只能定性地分为好的与坏的，合格的与不合格的，这叫做用“计件”的方法表示产品的质量特征。

我们把“计件的”和“计点的”方法，统一称为计数的方法。计数的方法有一个重要的特点，即它的数据是不连续的，也就是说数据是以离散的状态出现的。

1.2.2.2 计量值

许多产品的质量要用连续数量来表示。如棉纱的强度、钢的化学成分、电灯泡的寿命等，都是衡量产品质量的连续变量。使用连续变量来定量地度量产品质量的方法，称为计量方法。

计量值与计数值的不同，在于计量方法得出的数据是一组连续变量，而计数值是一组离散数据。如何区分离散数据和连续型变量呢？所谓离散数据又称为间断数据。它的特点是在同一变数之下，任何两个不同变数值之间不能插入无穷多个数值，而连续型变量则可以。

例如，记录机器每天发生故障的次数，属于计数值。记录得出来的数据是离散的。我们在 3 和 7 之间，插入 4.56 是无意义的。因为机器发生故障的次数不可能

取 4.56 次。在测试电灯泡寿命的一组数据里,取任意两个不同的数值,如 1230h 与 2000h,在其中插入 1457.8h 是有意义的。因此,电灯泡的寿命属于计量值。

1.2.3 数据收集的目的

数据是我们行动的向导。足够、翔实的数据是正确进行统计分析的前提条件。从数据中我们可以得知相关的事实(包括过程是否受控和产品质量的实际状况),并根据这些数据做出相应的判断。在收集数据前,重要的是确定自己要达到什么目的。

在质量控制中,收集数据的目的通常有:

- (1) 监视、测量和控制生产过程;
- (2) 分析不合格产品的现状及其原因;
- (3) 质量验收(例如抽样检验)。

数据的收集要能为实现预期的目的所用,要注意防止“事与愿违”的情况出现。比如:在一个机械加工车间,检查员对一种急于投放市场的产品进行转入下一道工序前的抽样检验。为追求生产进度,一批本应该拒收的产品被接收了,但他们并没有对这批本应被拒收的产品做任何处理(比如:进行筛选,挑出不合格品)。这就是说,合格品与不合格品混在一起,全都进入了下一道工序。抽检数据本来是用来确定产品的合格率的,确定是否可以放行的,此时却没有起到它应有的作用。

实践中,一旦确定了收集数据的目的,也就确定了要收集的数据的类型(计量型、计数型)和频次。例如:假设这里有一个涉及产品质量变差的问题,如果每天只收集一组数据,那么要在一天之内就了解和确定这种变差是不可能的。或者,你想知道两个操作者都是怎么生产出不合格品的,就需要分别收集两个人的数据,以便对其操作做出比较。如果将两组数据进行比较,能够显示出两个人有明显的差别,那么就有可能制定出能够消除两位操作者之间的差别的纠正措施,同时也会减少生产过程中产生的变差。

根据一定的原则把一组数据分成几个小组的方法称作“分层处理”。这一点非常重要,让人们在相对比较复杂的情况下,养成应用分层处理的习惯很有必要。

如果你了解两个变量(如某种金属零件的硬度与其材料成分)之间的相关关系,在这种情况下,数据应该是成对收集的。对成对数据的分析可参阅本书后面介绍的“散布图”。

1.3 质量变异的描述

1.3.1 系统因素与非系统因素

要使产品的质量保持稳定,就要设法找出影响产品质量的各种因素。我们知

道,影响质量的因素往往是十分复杂的。从总的方面来看,影响产品质量的因素可分为两类:一类是随机因素,又可称为非系统因素;另一类是非随机因素,也可称为系统因素。

什么因素属于随机因素,什么因素又属于非随机因素呢?比如,我们要轧制某一种规格的工字钢,并抽查其反映力学性能的屈服点。我们假定:

- (1) 用同一型号的钢来轧制;
- (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) 做力学性能试验时用同样的拉力试验机,测试屈服点;
- (5) 做力学性能试验时由同等技术水平的工人操作轧钢机做拉力试验,测试屈服点。

尽管以上五条固定不变,但是当我们抽查这批工字钢并测试它们的屈服点时,仍各不相同。这是因为影响工字钢力学性能的因素很多,除上面所列出的五条以外,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,而且即使能够确定也难加以控制。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的微小差异,等等。这里列举的各种微小的不可测的因素,它们之间相互独立,对产品质量起总的影响,我们称这类因素为随机因素,或叫作非系统因素。对于可以列出的,易于控制的因素,如条件(1)~条件(5),则称为影响产品质量的系统因素,简称为系统因素,或非随机因素。

影响产品质量的因素,总括起来,就是这两种类型。这两类因素的根本区别是:随机因素是属于不易确定或者根本不能确定的各因素作用的总和。它对产品质量的影响是随机的,具有偶然性,因而也是难以控制的。由于随机因素对于产品质量的影响是偶然的,故当生产过程仅受到随机因素的影响时,我们就认为生产过程处于正常状态。相反,系统因素是可以确定的,为数不多的几种因素的总和,它们的存在对产品质量的影响是系统的,非偶然的,而且直接关系到产品质量的波动。

我们仍然以轧制工字钢为例,在条件(1)~条件(5)中,仅改变条件(1),即使用另一种型号的钢作为轧制工字钢的原材料,经测试如两种钢材轧出的工字钢,其屈服点的平均值产生的差异,我们称这种差异为系统差异。这种差异,可以采取的措施,如通过调整条件(1)~条件(5)使屈服点的平均值达到我们原定的要求。

在生产过程中,我们进行抽样检验,其目的就是通过抽取的数据,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使产品质量产生波动,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。

态。此时,就需要立即找出原因,或调整机器,或检查原材料有无异常,等等,务必消除系统因素对生产的影响。

1.3.2 产品质量波动

波动是客观存在的,它存在于任何事物中。产品质量具有波动性和规律性。在生产实践中,即便操作者、机器、原材料、加工方法、测试手段、生产环境等条件相同,但生产出来的一批产品的质量特性数据却并不完全相同,总是存在着差异,这就是产品质量的波动性。因此,产品质量波动具有普遍性和永恒性。当生产过程处于统计控制状态时,生产出来的产品的质量特性数据,其波动又服从一定的分布规律,这就是产品质量的规律性。

从统计学的角度来看,可以把产品质量波动分成正常波动和异常波动两类。

1.3.2.1 正常波动

正常波动是由随机因素(也称一般原因或偶然原因)引起的产品质量波动,是过程固有的波动。这些随机因素在生产过程中大量存在,对产品质量经常产生影响,但它所造成的质量特性值波动往往比较小。例如,原材料的成分和性能上的微小差异;机器设备的轻微振动器;温度、湿度的微小变化;操作方面、测量方法、检测仪器的微小差异;等等。对这些波动的随机因素的消除,在技术上难以达到,在经济上的代价也很大。因此,一般情况下这些质量波动在生产过程中是允许存在的,所以称为正常波动。公差就是承认这种波动的产物。我们把仅有正常波动的生产过程称为过程处于统计控制状态,简称受控状态或稳定状态。

1.3.2.2 异常波动

异常波动是由系统原因(也称特殊原因或可查明的原因)引起的产品质量波动。这些系统因素在生产过程中并不大量存在,也不经常对产品质量产生影响,而一旦存在,它对产品质量的影响就比较显著。比如,原材料的质量不符合规定要求;机器设备带病运转;操作者违反操作规程;测量工具带系统性误差;等等。由于这些原因引起的质量波动大小和作用方向一般具有一定的周期性或倾向性,因此比较容易查明,容易预防和消除。又由于异常波动对质量特性值的影响较大,一般说来,在生产过程中是不允许存在的。我们把有异常波动的生产过程称为过程处于非统计控制状态,简称失控状态或不稳定状态。

质量管理的一项重要工作,就是要找出产品质量波动的规律,把正常波动控制在合理的范围内,消除因系统原因引起的异常波动。

从微观角度看,引起产品质量波动的原因主要来自以下六个方面:

- (1) 人(man): 操作者的质量意识、技术水平、文化素养、熟练程度和身体素质等。
- (2) 机器(machine): 机器设备、工具夹的精度和维护保养状况等。

- (3) 材料 (material): 材料的化学成分、物理性能和外观质量等。
- (4) 方法 (method): 加工工艺、操作规范和作业指导书的正确程度等。
- (5) 测量 (measure): 测量设备、试验手段和测试方法等。
- (6) 环境 (environment): 工作地的温度、湿度、含尘度、照明、噪声和震动等。

通常把上述因素称为引起产品质量波动的六大因素或简称“5M1E”因素、质量波动源。

1.4 质量管理中的常用分布

概率分布是将变量在总体中的取值与其发生的概率两者相联系的数学模型。概率分布有离散概率分布与连续概率分布两种类型。在质量管理中, 常见的离散概率分布有二项分布与泊松分布, 常见的连续概率分布有正态分布。

1.4.1 二项分布

二项分布产生的背景是:

(1) 重复进行 n 次随机试验。例如, 把一枚硬币连续抛 n 次、检验 n 个产品的质量、对一个目标连续射击 n 次等。

(2) n 次试验间相互独立, 即每一次试验结果不对其他次试验结果产生影响。

(3) 每次试验仅有两个可能的结果。例如, 正面与反面、合格与不合格、命中与不命中、具有某种特性与不具有某种特性, 以下统称为“成功”与“失败”。

(4) 每次试验成功的概率均为 p , 失败的概率均为 $1-p$ 。

在上述四个条件下, 设 X 表示 n 次独立重复试验中成功出现的次数, 显然 X 是可以取 $0, 1, \dots, n$ 等 $n+1$ 个值的离散随机变量, 且 $X=x$ 的概率为:

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x=0, 1, 2, \dots, n \quad (1-1)$$

这一分布称为二项分布, 记为 $b(n, p)$, 其中 $\binom{n}{x}$ 是从 n 个不同元素中取出 x 个的组合数, 它的计算公式为:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

若不合格品率为 p 的一批产品批量很大, 现从中随机地抽取 n 个产品相对较少, 观察抽出的产品是合格品还是不合格品, 那么:

(1) 由于一批产品的批量很大, 以至于抽取一个产品不管是否为不合格品,

不会影响下次抽到不合格品的概率, 因此可以认为各次抽取结果是相互独立的。

(2) 每次抽取只有两个结果: 合格品与不合格品, 且抽到不合格品的概率皆为 p , 抽到合格品的概率皆为 $1-p$ 。

若记 n 个产品中不合格品数为 X , 则 X 服从二项分布 $b(n, p)$ 。

二项分布 $b(n, p)$ 的均值、方差与标准差分别为:

$$E(X) = np$$

$$\text{Var}(X) = np(1-p)$$

$$\sigma(X) = \sqrt{np(1-p)}$$

特例: $n=1$ 的二项分布称为二点分布。它的概率函数 (即分布列) 为:

$$P(X=x) = p^x(1-p)^{1-x} \quad x=0,1$$

或如表 1-1 所示。

表 1-1 (0-1) 分布的分布律

X	0	1
P	$1-p$	p

它的均值、方差与标准差分别为:

$$E(X) = p \quad \text{Var}(X) = p(1-p) \quad \sigma(X) = \sqrt{p(1-p)}$$

[例 1-1] 已知一批晶体管中, 一级品率为 20%, 现从中任意抽取 10 只, 问:

- (1) 10 只中恰有 d 只 ($d=0, 1, \dots, 10$) 一级品的概率;
- (2) 不超过 3 个一级品的概率;
- (3) 二项分布 $b(10, 0.2)$ 的均值、方差与标准差。

解: (1) 题目未指明晶体管的批量, 就可以看作批量 N 很大, 而且抽查的数量 $n=10$ 相对于 N 来说又很少, 因此可直接用二项分布计算结果。这里 $n=10$, $p=0.2$, 由式 (1-1) 可立即得到:

$$\begin{aligned} P\{X=d\} &= b(10, 0.2) = C_{10}^d (0.2)^d (1-0.2)^{10-d} \\ &= \frac{10!}{d!(10-d)!} (0.2)^d (0.8)^{10-d} \quad (d=0, 1, 2, \dots, 10) \end{aligned}$$

上述结果, 见表 1-2。

表 1-2 d 只一级品的概率

X	0	1	2	...	9	10
P	0.107	0.268	0.301	...	0.000004	0.0000001

从表 1-2 中可以看到, 当 X 的取值 d 增加时, 概率 $P(X=d)$ 先是由小增大, 直至达到最大值, 随后又逐渐减少。对于任意给定的 n 及 p , 二项分布皆具