

# 数理统计基础



深圳市第一职业技术学校

THE FIRST VOCATIONAL TECHNICAL SCHOOL OF SHENZHEN

2021届现代学徒制班

# 数理统计基础

教研组组长 : 尹庆玲、朱建顺、刘会龙

教研组成员 : 廖庆江、蔡哲豪、燕海涛

教务组 : 蔡哲豪、刘艳玉、肖国政、胡燕妮、罗雄新

(排名不分先后)

---

校企合作·现代学徒制学徒手册第一版  
校企合作项目组编

出版机构 : 欣旺达大学

出版日期 : 2019-09-12

页数 : 130

开本 : 210mm\*290mm

版次 : 2019年09月第一版

欣旺达电子股份有限公司

Sunwoda Electronic Co., Ltd.

电话: 86-755-29516888

邮编: 518108

网址: <http://www.sunwoda.com>

地址: 深圳市宝安区石岩街道石龙社区颐和路2号

公司内部资料, 严禁外传。

Sunwoda confidential



# 目录

前言.....	4
<b>第一章 数理统计的概述.....</b>	<b>6</b>
第一节 统计学的背景.....	7
第二节 统计学的基础介绍.....	9
第三节 复习题.....	12
<b>第二章 数理统计基础概念.....</b>	<b>13</b>
第一节 数据的基本概念.....	14
第二节 数据的重要特征.....	18
第三节 复习题.....	26
<b>第三章 数据分析手法.....</b>	<b>28</b>
第一节 QC 七大手法的背景.....	29
第二节 QC 七大手法.....	30
第三节 复习题.....	62
<b>第四章 质量变异的规律性分析.....</b>	<b>64</b>
第一节 随机概率及其分布.....	65
第二节 排列与组合.....	70
第三节 常用离散型分布.....	75
第四节 常用连续分布.....	78
第五节 复习题.....	89
<b>第五章 过程质量控制.....</b>	<b>91</b>
第一节 过程质量.....	92
第二节 过程能力及其指数.....	99
第三节 复习题.....	110
<b>期末测试.....</b>	<b>111</b>
<b>习题解答.....</b>	<b>113</b>
第一章 复习题答案.....	113
第二章 复习题答案.....	114
第三章 复习题答案.....	115



---

第四章 复习题答案 .....	119
第五章 复习题答案 .....	120
期末测试答案 .....	121
附表 1 标准正态分布表 .....	123
附表 2 控制图判异规则 .....	124
附表 3 计量值控制图系数表 .....	125
参考文献 .....	126



## 前言

随着社会的发展不断进步，数理统计学科的重要性日益凸显，在工厂中对这些统计理论知识的人才需求日益增加，全国性的数理统计教材较为高深相对于中职业院校学生理解困难。

为了主动适应中职业学校学生在今后的企业工作中能够更快的进入角色，发挥自己的才能。在企业中，普遍采用老师傅带徒弟的模式，老师傅，跟学生讲解的都是较为直接使用的统计方法，未涉及统计学的基本原理，学生学到了方法却对基本原理没有了解，对此会产生困惑。所以要选择一套合适的教材，对学生进行辅导教学，目前尚不容易。由此，我们在编写《2019年常用办公软件基础教材中职业院校版》时便萌生了编写一套在新形势下对中职生和企业的发展具有意义的关于统计方面的基础教材的想法。对此，我们经过调研和与企业领导的交流，组织有10多年数理统计经验的专家通力合作，这本“数理统计教材中职版”终于完本了。

这本教材具有以下显著特点

1、传统内容与工厂实际应用相结合，自始至终为了培养中职生在企业中运用数理统计方法为企业创造价值，提升学生的自我能力。在本书中加入工厂中对实际遇到的问题数据的分析解决方法。

2、基本内容与信息技术有机融合，将信息技术运用到工厂实际问题中，快速解决问题，同时丰富了本书内容。

3、在本书中化抽象为直观，化枯燥为有趣，化繁难为简易，化理论性为应用性。

在数学教学中，数理统计学作为一门较为深奥的学科，在教学活动中并不如人意，其中它枯燥的公式与理论让学生们的注意力放到了别的地方。让数理统计教学陷入了困难。要解决这些问题，对教学者来说，是具有一定的挑战性的。教学工作者应该以饱满的工作激情参与每日的教学活动，积极解放自己的思想，通过了解学生的需求和与时俱进、不断更新完善教学内容、改进教学的方法，针对不同的层次的学生，开设不同的数理统计教程。本书的特点是根据计算机信息技术的发展，引入了计算机对数据的统计和计算，其中，也使得学生能清晰地计算机上运算复杂的公式，让他们看到统计学的魅力所在，所以在编写过程中，我们力求做到：

1、数理统计的基本内容，为学生在工厂实习中带来较大的帮助。



2、在保证了教学内容的基础上，教学内容力求语言不晦涩，浅显易懂

3、计算机软件的引入，统计计算部分都通过 Excel 软件来实现。

我们诚挚的希望学生们能够在轻松的氛围体会数理统计方法的好处，充分运用所学到的知识，在工厂中解决遇到的专业性的问题。

最后，我们在编写这门课程是以丰富的背景、巧妙的思维和有趣的结论让枯燥的数理统计学科更加的浅显易懂，使学生的学习积极性调动起来，增加对数理统计学科的兴趣。从而掌握数理统计的基本概念以及在学习在工厂中需要运用到数理统计的知识的运用。总之我们已经尽了自己最大的努力，可能会有不尽如人意的地方，只要这门教材能为企业和学生带来微薄的帮助，我们就十分开心了。



# 第一章 数理统计的概述

## [教学目标]

- (1) 掌握数理统计学的产生与发展
- (2) 掌握统计学的基本概念
- (3) 掌握区分总体与总体单位、标志、指标和统计指标体系。

## [教学难点]

- (1) 统计学的基本概念的掌握
- (2) 统计学基本概念的区别与简单运用



## 第一节 统计学的背景

随着社会的不断进步，数理统计技术的发展也逐渐地大众化，在工厂中很多实际应用中会用到数理统计的知识。数理统计的发展历程从古典统计学时期到近代统计学时期然后到现代统计学时期，经过这一段的发展历程，统计学变得更加的完善。

### 一、古典统计学

#### 1. 政治算术学派

17 世纪，英国，代表人物威廉·佩第博士，于 1671—1676 年写成《政治算数》一书。用数字表述，用数字、重量和尺度来计量，并配以朴素的图表，来刻画社会经济现象的特征。现代统计学广为采用。

另一代表人物约翰·格朗特（John. Graunt, 1620—1674）对伦敦人口的出生率和死亡率进行了分类统计，著有《关于死亡表的自然观察与政治观察》一书，他编制了世界上的第一张“死亡率”统计表。但是，该学派“有统计学之实，无统计学之名”。

#### 2. 记述学派（国势学派）

18 世纪，德国，代表人物康令博士，以叙述国家显著事项和国家政策关系为特征。

主要继承人阿亨瓦尔，《近代欧洲各国国势学概论》首次使用统计学代替国势学，认为统计学是关于各国基本制度的学问，是一个国家显著事项的整体。因在拉丁文中“国势”与“统计”词义相同，Status (Statistics) 后正式命名为“统计学”。该学派“有统计学之名，无统计学之实”。

#### 3. 古典概率论学派

古典概率论的研究始于 16 世纪的意大利，直到 17 世纪中叶才得到一般化的解法。并在 18 世纪的法国、瑞士等国得到广泛的发展，最终于 19 世纪初也有法国数学家、统计学家拉普拉斯（1749—1827）在总结前人成果的基础上，出版了名著《概率分析理论》一书，从而形成了完整的应用理论体系。

### 二、近代统计学

1. 19 世纪比利时的凯特勒博士（1796—1874）深受拉普拉斯的影响，著有《社会物理学》和《关于概率论的书信》等书籍，并首次把概率统计引入统计学中，大力提倡采用





统计方法对自然现象和社会现象的规律进行观察。

## 2. 社会统计学派

代表人物：克尼斯、恩格尔(恩格尔定律)、梅尔代表著《人口统计学》。他们继承和发展了“政治算数学派”的理论体系，完善了统计有关描述性的分析方法。如统计指标等。主要方法是大量观察法。

## 三、现代统计学

### 1. 数理统计学派

不断吸取生物学研究中的成果，提出并发展了回归和相关、假设检验、t分布、F分布等理论和方法，逐渐发展成为一门完整的学科。

哥赛特：首创小样本 t-分布理论。

费希尔：提出实验设计等。

尼曼：提出统计假设理论。

### 2. 社会统计学派

费拉斯卡姆波：《一般统计学》

史密斯：《统计学原理》

高野岩山郎：《社会统计史研究》

### 3. 社会经济统计学派：

斯特鲁米林：发展社会经济统计。

廖佐夫：理论与方法。

列昂捷夫：投入产出分析理论。

## 四、中国的统计学发展

大体可以划分为四个阶段：

1. 新中国成立后，中国的统计工作得到了顺利开展。

2. “文革”期间主要参照前苏联模式。

3. 改革开放后，统计工作步入正轨。

4. 加入世界贸易组织后，我国的统计与国际接轨的速度进一步加快。



## 第二节 统计学的基础介绍

### 一、统计学的定义

统计学的定义具有两种，一种是《不列颠百科全书》中的定义：统计学是“用以收集数据、分析数据和由数据得出结论的一组概念、原则和方法”。另外一种是在“普通高等学校本科专业目录和专业设置”中给出的定义为：“统计学是研究如何收集数据，分析数据，以便得出正确认识结论的方法论科学。”

### 二、统计数据的分类

#### 1. 分类的数据：

- (1) 按数据的种类，事物数据进行分类。
- (2) 按数据表现形式，数据为类别，可用文字来表述。
- (3) 按数据性质，例如人口按性别分为男、女两类。

#### 2. 顺序的数据：

是对事物类别顺序的测度；数据表现为类别，用文字来表述例如，产品按质量分为一等品、二等品、三等品、次品等。

#### 3. 数值型数据：

是对事物精确的测度；结果表现为具体的数值；例如：身高为 175cm、168cm、183cm。

#### 4. 观察的数据：

是通过调查或观测而收集到的数据；在没对事物人为控制的条件下而得到的有关社会经济现象的统计数据几乎都是观测数据。

#### 5. 试验的数据：

在试验中控制试验对象而收集到的数据比如，对一种新药疗效的试验，对一种新的农作物品种的试验等；自然科学领域的的数据大多数是试验数据。

#### 6. 截面的数据：

在相同或近似相同的时点上收集的数据描述现象在某一时刻的变化情况；比如，2005 年末我国人口总数。

#### 7. 时序的数据：



是在不同时间上收集到的数据；描述现象随时间变化的具体情况；比如，2000年至2005年国内生产总值数据。

### 三、统计学总体与总体单位

#### 1. 定义

凡是客观存在的，在同一性质基础上结合起来的许多事物（个体）的整体，就是统计总体。组成统计总体的个体称为总体单位，下面举例说明总体和总体单位。

**【例题一】**某工厂 IPQC 对当日生产的电芯进行质量检测，研究电芯的质量是否有问题？什么是统计总体，什么总体单位。

解： 总体是：当日生产的全部电芯。

总体单位是：每一个电芯。

**【例题二】**某高校对男同学的身高进行测量？

解： 总体是：高校所有的男同学。

总体单位是：每一个男同学。

**【例题三】**再如，研究某市的工业生产设备情况？

解： 总体：所有的工业生产设备 。

总体单位：每一台工业生产设备。

#### 2. 统计总体的特点

统计总体的特点分别是同质性、大量性，下面对统计总体的两个特点进行说明。

##### （1）同质性

总体中各个单位具有某种共同的性质。

##### （2）大量性

总体总是包含大量的单位

#### 3. 统计学的类型

统计学的类型按照变量的取值的个数是否可数来进行划分：

有限总体：变量取值的个数是可数的。

例如，某班级学生数量；某线体设备的数量；某工厂一日生产的电芯。

无限总体：变量取值的个数是不可数的。



例如，天上的星星；海边的沙子；草原上的青草；森林里的树木；海洋里的渔业资源等等。



## 第三节 复习题

### 一、单项选择题

1. 下面哪一位是古典统计学时期代表人物 ( )

A. 威廉·佩第博士      B. 拉普拉斯      C. 马云

2. 下列论述中错误的是 ( )

A. 凡是客观存在的，在同一性质基础上结合起来的许多事物（个体）的整体，就是统计总体。

B. 数值型数据：是对事物精确的测度

C. 有限总体的变量取值的个数是不可数的。

### 二、简述题

1. 统计学的定义？

2. 什么是统计总体与总体单位？

3. 现在某工厂要购进吸尘器，要调查某商店销售的吸尘器情况，请指出总体、总体单位是什么？



## 第二章 数理统计基础概念

### [教学目标]

- (1) 掌握数据的基本概念
- (2) 掌握数据重要特征值的计算公式
- (3) 掌握数理统计的相关概念

### [教学难点]

- (1) 重要特征值的计算公式
- (2) 数据的认识及相关计算方法

## 第一节 数据的基本概念

数理统计的基础就是数据，它是事物真实特性的具体反映，数据的收集是实行统计方法的基础工作。

### 一、数据的分类

在工厂、车间、生产现场经常会碰到统计数据，有员工工资单、加班单、考勤汇总单、每日产品产量表、不合格品数、合格品数、成品尺寸、质量、长度、强度、化学成分、时间、电阻等。这些统计数据有些是可以测量出来的，有些是可以数出来的，根据这些数据我们知道数据是统计方法运用的前提，是产品质量特性的具体反映，从统计的角度来看，一般把上述所说的归纳为两大类：即计量数据和计数数据。

#### 1. 计量数据

是指可以连续取值，或者说可以用测量工具测量出小数点以下数值的数据，例如物体长度、体积等，一般来说，计量数据服从正态分布，后文讲解的正态分布数据都是计量数据。

计量型数据见下图 2-1 所示，例如：长度、重量、强度、化学成分、时间、电阻等。

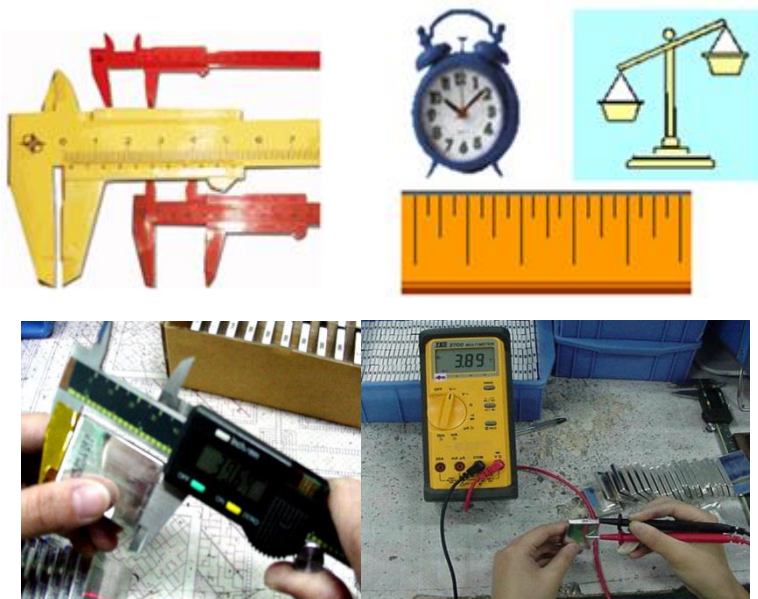


图 2-1 计量型数据

#### 2. 计数型数据

是指不能连续取值，或者说就是用测量工具也得不到小数点以下数值，而只能得到自然数的这类数据，例如计件数据就是按件计数的数据—某生产的不合格品，计点数据就是按缺陷点计数的数据—某生产线体设备缺陷数。

计数型数据见下图 2-2 所示，例如合格/不合格、通过/失败、接受/不接受、出席/缺席、好/坏及缺陷个数等。

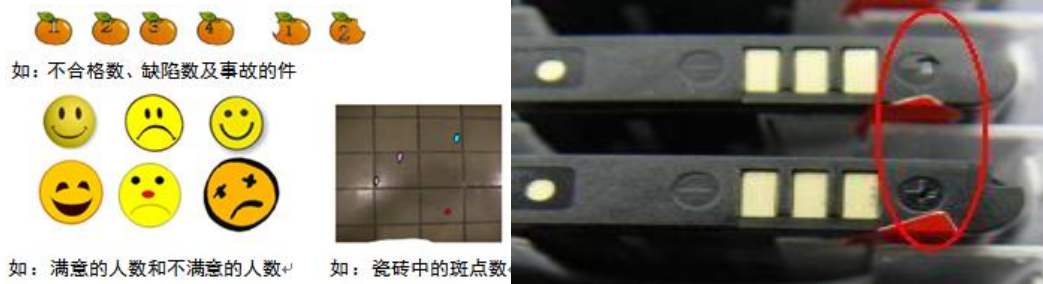


图 2-2 计数型数据

注意事项：当数据用百分率表示时，是计量数据还是计数数据应取决于得到百分率的计算公式的分子，若分子是计量数据，则得到百分率数据是计量数据，反之亦然。

计数数据还可以进一步的细分为：计件数据和记点数据。

(1) 计件数据：计件数据就是如它的名字所说，就是按件计数的数据。在生产中，例如一个不合格的电芯、一台设备、一卷工厂胶纸等。

(2) 计点数据：按照缺陷点（项）计数的数据。例如，在生产中，发现设备中生锈的螺钉上的锈斑数、一个电芯上存在的缺陷数等。

其中，计件数据服从二项分布；计点数据服从泊松分布。

【例题一】：如表 2-1 为某线体头部胶不良数据统计，分析其中不良率是属于计数数据还是计点数据。

表 2-1 某线体头部胶不良数据表

日期	拉线/班次	良品数	头部胶不良	拉线/班次	良品数	头部胶不良	总计	不良率
8/21	A(白班)	2250	24	A(夜班)	3750	25	49	0.81%
8/22	A(白班)	3150	32	A(夜班)	0	53	85	2.63%
8/23	A(白班)	4200	15	A(夜班)	4200	46	61	0.72%
8/24	A(白班)	3300	42	A(夜班)	3450	102	144	2.09%
8/25	A(白班)	2700	60	A(夜班)	3000	70	130	2.23%
8/26	A(白班)	3000	43	A(夜班)	3450	32	75	1.15%
8/27	A(白班)	1500	3	A(夜班)	300	12	15	0.83%
8/28	A(白班)	1500	12	A(夜班)	4200	46	58	1.01%
8/29	A(白班)	4050	30	A(夜班)	3900	14	44	0.55%
8/30	A(白班)	4350	97	A(夜班)	3900	29	126	1.50%
总计	A(白班)	30000	358	A(夜班)	30150	429	787	1.29%



解：由图中数据分析可以知道，不良率=不良数/生产总数，其中，不良数为计数数据，根据前文所说当数据用百分率表示时，是计量数据还是计数数据应取决于得到百分率的计算公式的分子，若分子是计量数据，则得到百分率数据是计量数据，反之分子是计数数据，则得到的百分率数据是计数数据。所以有上表中头部胶的不良率为计数数据。

## 二、数据的要点

数据具有六大性质：统一性、针对性、完整性、准确性、及时性、连续性、统一性。关系图如图 2-3 所示。



图 2-3 数据的要点

### 1. 数据具有针对性

公司部门在收集数据时，应该考虑自身需求，针对性的收集对过程控制和体系起到作用的数据。还需要根据职位的不同来选择收集的数据。数据的收集至少能提供以下四方面的信息，见图 2-4 所示。

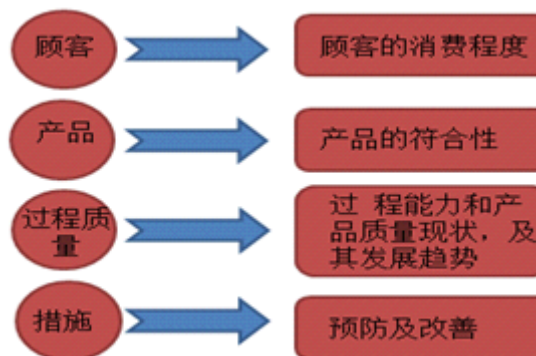


图 2-4 数据的收集信息

### 2. 数据具有完整性

数据的完整性要求有三点，第一点，数据的过程要完整；第二点数据的来源需要记录下来；第三点，数据表格中规定的栏目需要填写完整。



### 3. 数据具有准确性

数据要具有真实性，需要能真实反映事情的实际情况，一个错误的数字，不但不能起到应有的作用，还有可能导致一个错误的结论。

### 4. 数据具有连续性

连续性的数据能让公司更好的发现产品和体系的动态变化规律，不连续型的数据，会导致数据的获得不够多，影响数据的分析结果。

### 5. 数据具有及时性

要做到及时记录、及时传递、及时处理和及时通知。

### 6. 数据具有统一性

数据的位数、数据的修约规则、数据的表示都具有统一性。

## 三、建立数据系统

收集数据主要是为了获得所需的信息，企业应设立一个统一的管理企业信息的管理机构——信息中心，各职能部门或基层单位将收集到的数据(信息)提交信息中心，经中心加工处理后，再向有关领导报告和向有关部门反馈所需信息。这样就有利于提高管理的效率，促进质量管理体系的持续改进和产品质量的不断提高。



## 第二节 数据的重要特征

在现实生活中，在比较两组以上数据分布时，无法定量地表征它们之间的差别。为了解决这个问题，我们常用两类统计量表征数据即数据的位置特征值和数据的离散特征值。下面对这两类数据进行讲解。

### 一、位置特征值

#### 1. 位置特征值

表示位置的特征（集中趋势的度量）有平均值、中位数、众数等。

（1）平均值：也叫均值，是指将一组数据中的所有数据相加后得到的总和除以数据个数总合，平均值记为  $\bar{X}$ 。

$$\text{公式为: } \bar{X} = \frac{\sum_{i=1}^N (X_i)}{N}$$

这里举例题来说明平均值

**【例题一】**：如下表 2-2，某生产线，现在要对设备的节拍时间 CT 进行测试速度，员工针对 21 号机台的节拍时间进行测试。为了减小测试误差，这里测试八组数据，统计情况所示。求机台平均 CT。

表 2-2 CT 测试数据表

某线体21号机台CT测试数据								
测试次数	1	2	3	4	5	6	7	8
测试速度	56	58	55	60	57	56	58	59

解：根据公式计算平均 CT，公式为： $\bar{X} = \frac{\sum_{i=1}^N (X_i)}{N}$ ，计算结果为 57.4。

可得到，机台平均 CT 为 57.4S。

注意：CT 为节拍时间，节拍时间对工厂的产品生产效率极为重要，设备维护人员会对线体的每一个工站的 CT 进行测试，解决瓶颈工站 CT，实现设备综合效率 OEE 的提高。

（2）最大值：在整组数据中最大的数据值；最小值：在整组数据中最小的数据值



**【例题二】** 找出表 2-2 的 CT 的最大值和最小值？

解：将数据从大到小排列得出：60 59 58 57 56 55 可知道最大值即为 60，最小值为 55。

(3) 众数：众数是指在用频数分布表表示测定值时，是数据集中出现频率最高的数据。

例如某服装厂，卖出衣服的尺码之中 30 码最多人购买，那么 30 就是众数。服装厂依据众数可以生产更多的 30 码的衣服以供需求。

**【例题三】**某机械加工工厂，某机械设备在生产的一小时过程中，出现较多异常品，数据统计如下表 2-3 所示，求出其异常类型众数。

表 2-3 异常品数据

组数	功能不良	扫码不良	温度不良	尺寸不良
1	1	1	1	1
2	3	3	2	3
3	3	1	3	2
4	3	4	1	2
5	5	5	2	2
6	1	3	1	2
7	3	2	3	3
8	2	2	1	2
9	4	4	4	4
10	5	3	1	4
11	3	3	3	4
12	3	3	3	3
13	2	2	2	2
14	1	1	1	1
合计	39	37	28	35

解题思路：首先对异常类型的总数进行排序由大到小，即：39 37 35 28，其中可以发现功能不良总数为 39。即功能不良属于众数，应重点关注生产中造成功能不良的原因，提出改善措施。

(4) 中值：在数据中的最大值和最小值之和的平均值。

**【例题四】** 计算表 2-2 的 CT 数据中值为多少？

解：根据中值的计算公式， $M=(X_1+X_2)/2$ ，找出表 2-2 中的最大值与最小值分别为  $X_1=39$ ， $X_2=28$ 。根据公式计算得  $M=33.5$ 。

(5) 中位数：在古代以前将数据从小到大排列，若是单数，则中间值就是中位数；若是偶数，则取中间两个数之和的平均值。

**【例题五】** 下面有甲乙两组数据（见下图 2-5），其中第一组数据为奇数，第二组数据为偶数，分别计算 2 组数据的中位数，它们有什么不同？

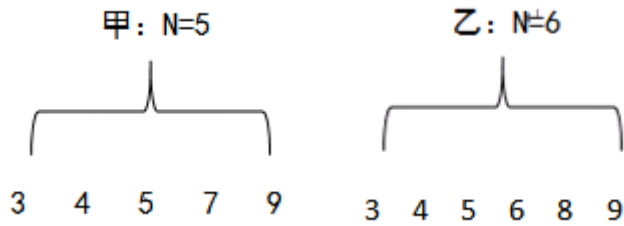


图 2-5

解题思路：计算甲的中位数为  $Me=5$

计算乙的中位数为  $Me=(5+6)/2=5.5$

## 二、离散特征值

实践中已经证明，由于变异性的普遍存在，无论采取什么样的措施都无法加工出两件完全相同的零件，零件不论在尺寸、性能上都会存在或大或小的差异，称之为离散性，即波动，同时这种离散性具有一定的规律。

1. 离散性：在生产产品中由于变异性的普遍存在，不管事前如何严格控制，反映产品质量的数据总是不尽相同，总会产生差别(部分差别微小，但还是会有波动的)。

**【例题一】**：下表是线体焊接点检时做的拉力数据表如表 2-4。试着去分析一下图表中拉力数据的波动性和规律性。

表 2-4 拉力数据表

某线体碰片与极耳焊接拉力数据											
日期	早上			中午			晚上			最大	最小
15	15.3	14.2	15.1	15.1	14.5	14.8	13.9	14.3	14.6	15.3	13.9
16	13.9	14.5	14.2	15.0	14.8	14.0	14.1	14.5	14.3	15.0	13.9
17	15.1	14.4	14.3	15.1	14.7	14.9	14.2	14.8	14.6	15.1	14.2

可以发现这三日的拉力值都是在一定的范围内(13.9~15.3 之间波动)，具有一定的规律。

对数据进行分析：首先，正确并客观地记录生产各个阶段的数据；其次，整理和分析数据；最后，选用恰当的统计方法揭示数据的离散和规律情况并进而运用统计规律，



达到控制产品质量，预防问题发生和质量的改进的目的。

【例题二】下表为某工厂 1、2 线体 5 号机器测试的整形气缸的下压压力数值，比较两组数据的离散程度，如下表 2-5 整形气缸下压力测试数据表。

表 2-5 整形气缸下压力测试数据表

整形气缸下压力测试							
1线	3.1	3.2	3.4	3.6	3.5	3.8	3.3
2线	2.8	2.9	3	4.5	4.3	3.3	3.1

解题思路：

①计算第一组数据的均值

$$X_1 = (3.1 + 3.2 + 3.4 + 3.6 + 3.5 + 3.8 + 3.3) / 7 = 3.4$$

②计算第二组数据的均值

$$X_2 = (2.8 + 2.9 + 3.0 + 4.5 + 4.3 + 3.3 + 3.1) / 7 = 3.4$$

通过计算得出两组数据均值相等，但从图表中数据可以看出第二组的数据要比第一组的波动要大，即离散程度要大。那么我们需要一个反映数据离散程度亦是变异程度的特征值。

常用的特征值有以下几种，如下表 2-6 所示。

表 2-6 常用离散特征值

常用离散特征值		
名称	定义	公式
极差 (R)	数据中的最大值与最小值的差	$R = X_{\max} - X_{\min}$
偏差平方 (S)	各个数据与平均值之差后的平方相加	$S = \sum_{i=1}^n (X_i - \bar{X})^2$
无偏方差	各个测定值的偏差平方和除以 (n-1) 后所得的值称为无偏方差 (简称方差)	$S^2 = \frac{S}{n-1} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
标准偏差 (s)	无偏方差的算术平方根就是标准方差	$s = \sqrt{S^2}$

【例题三】在下列数据表 2-7 中，计算数学期望，若期望相同，则计算其极差，判断波动性，用标准偏差验证其判断准确。若期望不同，则直接用无偏方差计算，得出其波动性。



表 2-7 某线体每日不良数

某线体每日不良数											
日期	7月1日	7月2日	7月3日	7月4日	7月5日	7月6日	7月7日	7月8日	7月9日	7月10日	7月11日
一线	15	17	16	18	13	16	17	15	16	16	17
二线	12	15	10	8	17	16	22	24	25	15	14

解题思路：分别计算出一线和二线的数学期望，对得出的数学期望进行对比，若相同，则采取第一种方案，计算极差，判断波动性，再用标准偏差验证，若不同，则采取二方案，直接采用无偏方差计算，比较大小，得出其波动性

①计算数学期望：

$$\text{一线: } \bar{X} = \frac{\sum_{i=1}^N (X_i)}{N} = (15+16+17+18+13+16+17+15+16+16+17) / 11 = 16$$

$$\text{二线: } \bar{X} = \frac{\sum_{i=1}^N (X_i)}{N} = (12+15+10+8+17+16+22+24+23+15+14) / 11 = 16$$

对比可得数学期望相等，所以采用第一种方案。

②计算极差：

$$\text{一线: } R_1 = 18 - 13 = 5$$

$$\text{二线: } R_2 = 25 - 8 = 17$$

根据极差性质进行对比可得  $R_2 > R_1$ ，所以判断二线波动比一线大。

验证：根据标准偏差原理一越小越稳定，根据计算公式可以得出二线波动较大，一线较为平稳的结果。与极差的判断结果相符合。

$$s_1 = \sqrt{S_1^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} = \sqrt{1.8} = 1.34$$

$$s_2 = \sqrt{S_2^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} = \sqrt{272} = 16.50$$

### 三、数理统计的相关概念

#### 1. 数理统计与统计方法

统计技术主要是指“数理统计”，它是建立在概率论基础上的一门数学分支，统计

技术是研究事物变异性及其规律的科学。统计方法分为两大类：描述型统计技术和推断型统计技术，关系如下图 2-6 所示

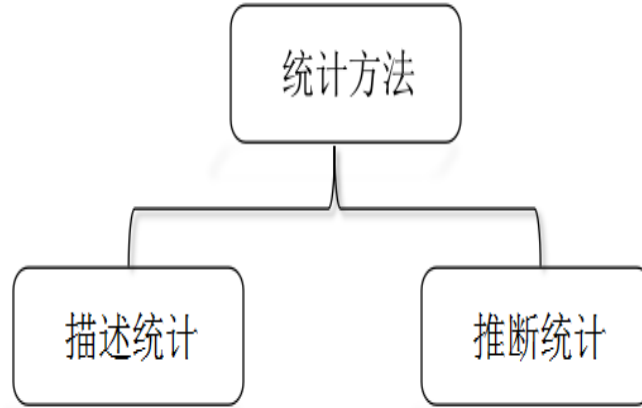


图 2-6 统计方法分类

(1) 描述型统计技术：主要是利用数据的特征或有关图表描述事物。

例如某服装店统计人们购买裤子尺寸数据如下表 2-8 所示

表 2-8 男士裤子尺寸表

男士裤子					
尺寸	29.00	24.00	24.00	25.00	25.00
	24.00	23.00	24.00	26.00	28.00
	28.00	24.00	24.00	27.00	30.00

由上图可知，服装店购买裤子尺寸为 24 码最多，所以根据描述型统计方法可以估计人群中需求 24 码的人最多，可以多购买进 24 码裤子

(2) 推断型统计技术：主要解决从样本如何推断总体。概率论和数理统计研究的对象大多属于此类。推断型统计的方法如下图 2-7 所示。

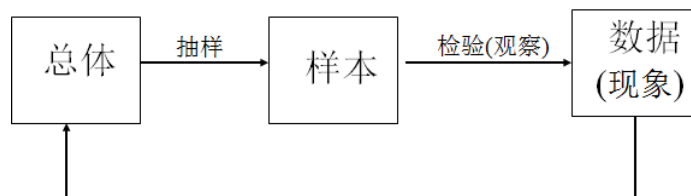


图 2-7 推断性统计技术的方法

例题分析，某车间生产了 5 万个产品，现抽取 1000 个产品检验，发现有 2 个不合格品，合格率为 99.8%，从样本合格率推断出总体产品的合格率大概为 99.8%。



例如，通过参加培训的学员身高数据的收集分析，可以推断出本地区人员的平均身高的范围是：（171.7，174.2）而且，我们对推断结果的正确性有 95% 的信心。

#### 四、总体与样本

统计学是一门应用性非常强的学科，它的历史已经有 300 多年，即使从皮尔逊（K. Pearson, 1857---1936）和费希尔（R. A. Fisher, 1890---1962）的工作算起，统计学也已有 100 多年的历史，并且取得了良好的社会和经济效益。

一般认为，统计学是一门研究如何有效的收集数据和分析数据到随机因素影响数据的学科。经过多年的研究和发展，统计学已经深入到了多个学科中，可以说是一个实际问题涉及一批数据，我们都应该利用统计学方法去分析它，解决它。统计学是唯一的对数据进行整理和分析的学科，随着统计学的发展和完善，其研究的内容已经非常丰富，且形成了多个学科的分支，如抽样调查，试验设计，回归分析，多元统计分析，事件序列分析、非参数统计、贝叶斯方法等等。总体与样本关系图如 2-8 所示。

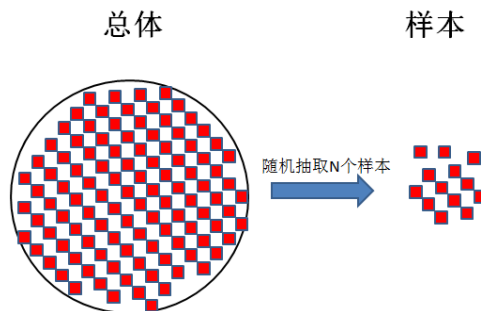


图 2-8 总体与样本关系图

总体：①在一个统计问题中；②研究对象的全体。

个体：构成总体的每一个单位称为个体。

总体以及个体，在生活中会碰到许多有关总体与个体的问题，例如说，统计某高校的学生的体重与身高是否符合标准，那么高校的全体学生就是这一个统计问题中的总体，那么每一个学生就代表着个体。总体就是一个分布，而其数量指标就是分布的随机变量，因为，每个学生所具有的数量指标值身高就是个体，而将所有身高全体看成总体，这样一来，总体就是指的一堆数，数有大有小，因此用一个概率分布来描述和归纳总体是合适的。

样本指的就是我们从总体中随机的抽取  $n$  个个体，记为  $X_1, X_2, X_3 \cdots X_n$  则这  $n$  个



个体称为总体的一个样本。

样本量指的是，将总体随机抽取的  $n$  个个体称为样本的容量或称为样本量。

样品指的是样本中的个体。

样本的性质：

①样本具有随机性，即要求总体中的每一个个体都有同等的机会被选入样本，这便意味着选中的学生都是随机的，不确定的。

②样本具有独立性，即要求样本中每一样品的取值不影响其他样品的取值，这意味着相互独立。

例如，某线体一日生产 5 万个电芯，对电芯进行检验，但不可能全部检验，此时，进行抽样检验，抽出 1000 个电芯进行检验，根据检验结果推断出总体的检验效果，如其中 5 万个中抽取 1000 个电芯中具有 2 个不良，此时样本中的合格率为 99.8%，推断出总体的合格率也为 99.8%。

## 五、生产批与检验批

生产批：过程在受控状态下连续生产的一批产品称为一个生产批。

批量：组成一批产品的单位个数称为批量。

检验批：待检验的一批产品称为检验批。

一般而言，一个生产批，即为一个检验批。但在某些特殊情况下，如批量过大、生产周期过长、可以将一个生产批划分为若干检验批。但要保证生产过程是稳定的，各检验批之间质量均衡，不可人为地任意划分。

注意：为保证检验批的代表性，任何情况下不能将两个生产批合并为一个检验批。



### 第三节 复习题

#### 一、单选题

1. 下列选项中属于计量数据的是（ ）

- A. 不合格产品数                      B. 合格产品数                      C. 电池的长度

2. 下列选项中错误的是（ ）

A. 计量数据指的是指可以连续取值，或者说可以用测量工具测量出小数点以下数值的数据。

B. 计数数据指的是不能连续取值，或者说就是用测量工具也得不到小数点以下数值，而只能得到自然数的这类数据

C. 产品的长度属于计数数据

#### 二、多选题

1. 数据的位置特征值有（ ）

- A. 平均值                              B. 最大值                              C. 最小值

2. 数据的离散特征值有（ ）

- A. 标准差                              B. 平均数                              C. 偏差

#### 三、判断题

1. 在特殊的情况下能将两个生产批合并为一个检验批（ ）

2. 表示位置的特征（集中趋势的度量）有平均值、中位数、众数、方差等（ ）

3. 统计学是唯一的对数据进行整理和分析的学科（ ）

4. 数据的完整性要求有三点，第一点，数据的过程要完整；第二点数据的来源需要记录下来；第三点，数据表格中规定的栏目需要填写完整（ ）

5. 由于变异性的普遍存在，无论采取什么样的措施都无法加工出两件完全相同的零件（ ）

#### 四、简述题



1. 统计方法主要分为哪两大类？并介绍这两大类。
2. 数据的位置特征有哪些？
3. 计数数据与计量数据的区别？
4. 数据的六大性质是什么？
5. 建立数据系统的意义？



## 第三章 数据分析手法

### [教学目标]

- (1) 了解 QC 七工具的起源
- (2) 掌握 QC 七工具的用途
- (3) 掌握 QC 七工具的制作方法

### [教学难点]

- (1) 统计学 QC 工具的基本概念的掌握
- (2) 统计学 QC 工具的具体用途的掌握
- (3) 统计学 QC 工具的柏拉图的掌握

## 第一节 QC 七大手法的背景

### 一、QC 七大手法的起源

QC 七大手法又称为 QC 七工具来源与英文 “The seven basic tools of quality control”，最初是由日本质量管理专家石川馨（见下图 3-1 所示）首先提出来的，石川馨将统计方法进行了难度的区分为三级。其中基础的就是 QC 七大手法，他的多年经验认为公司内部 95% 的问题可以通过 QC 七大手法来解决。



图 3-1 石川馨

### 二、QC 七大手法的意义与数据

- (1) 使用真实依据和事实说话。
- (2) 全面预防。
- (3) 全因素、全过程的控制。
- (4) 依据 PDCA 循环突破予以改善。
- (5) 层层分解、重点管理。

### 三、QC 七大手法的构成

QC 七大手法主要有：层别法、帕累托图（排列图）、鱼骨图、检查表、直方图、散点图（散布图）、控制图，下节开始对七大手法进行一一讲解。QC 七大手法构成如下图 3-2 所示。

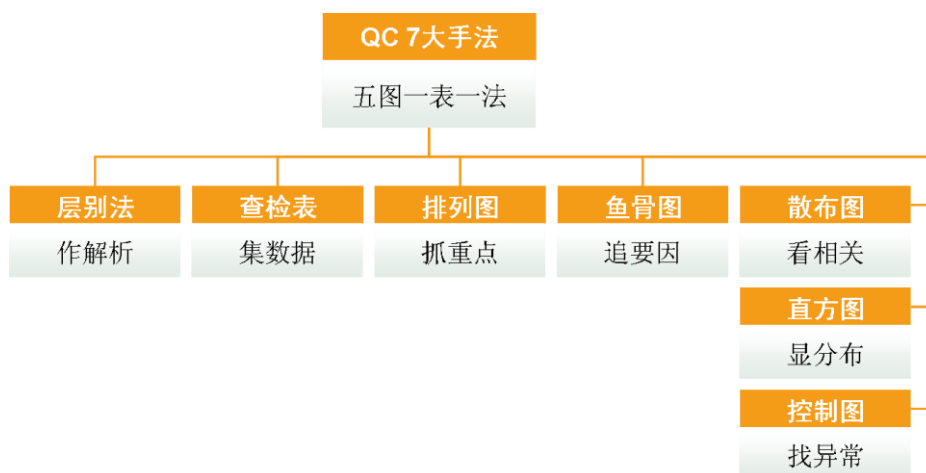


图 3-2 QC 七大手法

## 第二节 QC 七大手法

### 一、层别法

#### 1. 层别法的背景

层别法又称为分层法，在生产中质量问题往往是多种因素的综合结果。要把具体原因分析清楚，问题判断准确，行之有效的方法就是将数据分层，即将数据按各种影响质量的因素分别整理，层层分析，从而使分析准确无误。这种分别整理数据的分析方法就叫做分层，分层分得越仔细，判断就越准确。为使分层能达到应有的作用，分层时，应使同一层次内数据波动尽可能小，而层与层之间的差别尽可能大。

层别法的应用示例图如下图 3-3 所示。

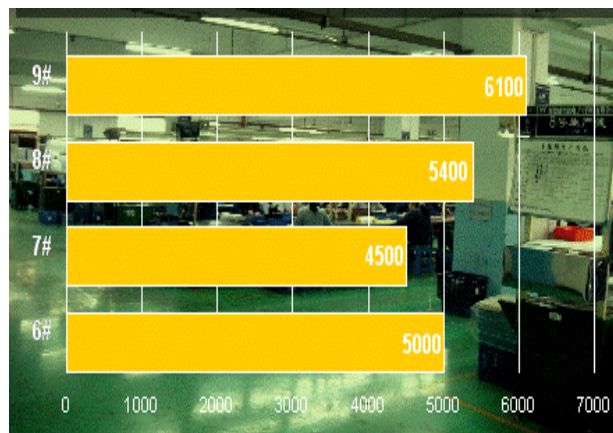


图 3-3 层别法示例图

#### 2. 分层的类别

可根据具体情况采用不同的分层类别，常见的有：

- (1) 操作人员：按个人、现场、班次、经验等分；
- (2) 机器设备：按机器、工夹具等分；
- (3) 材料：按来源、品种、生产批次分；
- (4) 加工方法：按加工、装配、测量、检验等各种工作条件分；
- (5) 时间：按上下午、年月季度分；
- (6) 环境：按气象情况、室内温度、磁场等环境分；

(7)其它：按发生情况、发生位置分。

分层是一种很有用的分析方法，我们在作直方图、因果图、排列图、相关图等其他一些图表时，都常结合分层的方法进行，否则，就难以得出有益的结论。

分层的思想图如下所示。

工程管理（见下图 3-4）：

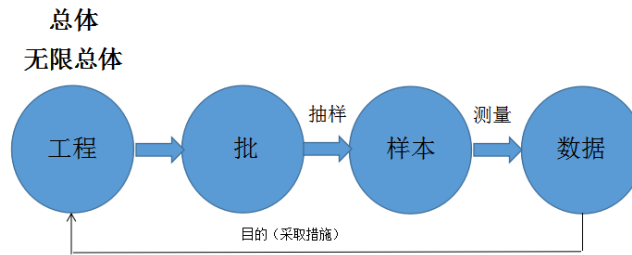


图 3-4

抽样检验（见下图 3-5）：

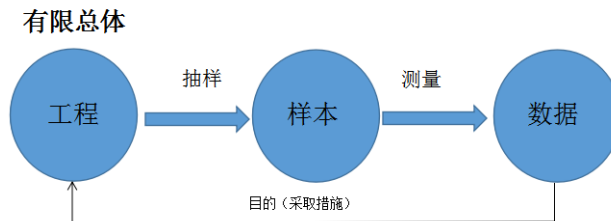


图 3-5

### 3. 抽样方法：

①简单随机抽样：对于总体中的每一个个体，其被抽到的概率为同等的，优点是抽样误差小，缺点是操作繁琐，花费时间较长。

**【例题一】**某生产线体，一天生产电芯总数为 500 个（电芯有编号），IPQC 随机抽取 50 个编号，按照编号对电芯进行质量检查。

②分层抽样：首先将总体划分几层，然后在划分的几层中在进行随机抽样。优点：减少层内差异，抽样误差小，增加样本的代表性。缺点：操作最繁琐。

**【例题二】**某生产线体，一天生产电芯总数为 500 个，分为 5 箱出货，在这 5 箱中每箱按照简单随机抽样的方法抽取 5 个电芯进行质量检测。





③系统抽样：称为等距抽样或机械抽样，它是按照一定的时间间隔抽取样本的一种方法。其优点：操作简单，缺点：不易使用在周期性变化的场合。

【例题三】某生产线体。一天生产电芯总数为 500 个（电芯已编号），对电芯随机抽签得到编号。如抽到 3 那么个位数是 3 的电芯全部作为样本检验。

④整群抽样：将整体分成若干群，然后随机抽取若干群的样本，优点：操作简单，缺点：抽样误差大、不具有样本代表性。

【例题四】某生产线体，一天生产电芯总数为 500 个，分为 5 箱出货，在这 5 箱中抽取一箱电芯全部作为样本检验。

## 二、帕累托图（Pareto 图）

### 1. 帕累托图的背景

1879 年，意大利经济学家维弗雷多·帕累托（Vilfredo Pareto）提出：社会财富的 80% 是掌握在 20% 的人手中，而剩余下的 80% 的人只占有 20% 的财富。渐渐的，这种“关键的少数和次要的多数”的理论，就是常说的二八定律，被称之为 Pareto 原则。也常常被称为 80/20 原则，或称帕累托法则，又或者被翻译为“帕拉图”法则。

#### 帕累托图（Pareto 图）的运用

柏拉图是统计工作中的一个重要的工具，其理念是：柏拉图把数据按从大到小次序排序，给我们展示了各个部分的分布情况，直条的排列是由多至少，柏拉图表折线是用一条累积相加，代表直条的百分比总数，从左到右一直累积上去，直到百分百。

柏拉图又称为排列图是一种按照事件发生的频率排序而成的，显示由于各种原因引起的缺陷数量不一致的排列顺序，是找出影响项目产品或服务质量的的主要因素的方法。在生产过程中，该定律指绝大多数问题或缺陷产生于相对有限的起因。

排列图是工厂中质量管理活动中会经常使用到的工具，柏拉图是本章的重点学习内容之一，需要达到熟练运用的程度方可。

柏拉图的制作 8 步骤：

- （1）确定分类的名称；
- （2）收集数据，统计某时间段或者某区域内的各分类的数量；



- (3) 汇总数据，计算各分类的数量和总数量；
- (4) 按照数量从多到少排序；
- (5) 计算各类的百分比；
- (6) 计算累计百分比；
- (7) 确定纵坐标高度和横坐标长度；
- (8) 绘制柏拉图。

2. 下面举例说明如何制作柏拉图

【例题一】三月份某线体设备故障时间如下，做出相应的柏拉图并进行分析。下表为三月份某线体设备故障时间表如下表 3-1。

表 3-1 表 3-2

三月份某线体		三月份某线体	
设备名称	故障时间/min	设备名称	故障时间/min
上料机	13	包外观胶机	130
极耳裁切机	32	贴VHB和热压机	90
焊接机	50	扣合机	70
贴PI胶机	23	装CAF机	60
装CAF机	60	焊接机	50
扣合机	70	极耳裁切机	32
包外观胶机	130	卸载具机	31
功能测试机	24	贴ID胶机	26
卸载具机	31	功能测试机	24
捡漏机	13	贴PI胶机	23
贴ID胶机	26	APMT机	19
APMT机	19	下料机	15
贴VHB和热压机	90	上料机	13
下料机	15	捡漏机	13

柏拉图制作步骤如下：

- (1) 数据整理，按照故障时间从多到少往下排列，如表 3-2 所示。
- (2) 算各组数据的百分比后在得到累计百分比，下表 3-3 所示。

表 3-3

三月份某线体		
设备名称	故障时间/min	累计百分比
包外观胶机	130	21.81%
贴VHB和热压机	90	36.91%
扣合机	70	48.68%
装CAF机	60	58.72%
焊接机	50	67.11%
极耳裁切机	32	72.48%
卸载具机	31	77.68%
贴ID胶机	26	82.05%
功能测试机	24	86.07%
贴PI胶机	23	89.93%
APMT机	19	93.12%
下料机	15	95.64%
上料机	13	97.82%
捡漏机	13	100.00%

- (3) 制出故障时间柱状图，如图 3-6 所示。

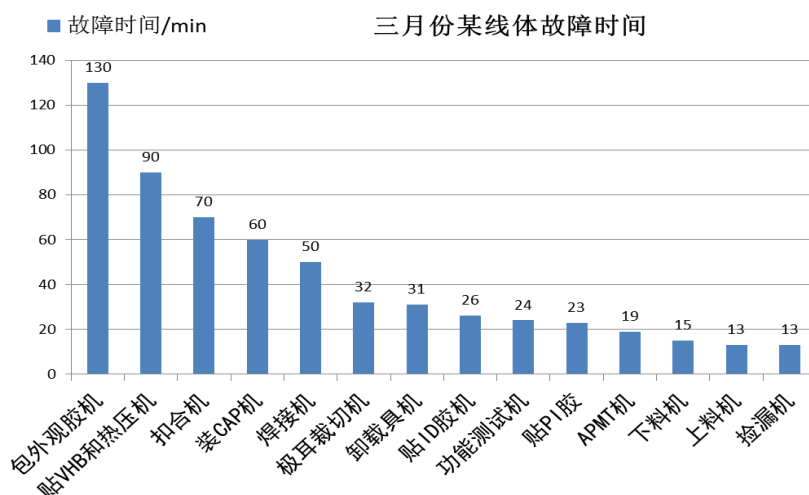


图 3-6 三月份某线体故障时间

(4) 柱状图上添加上累计百分比折线图，如下图 3-7 所示。

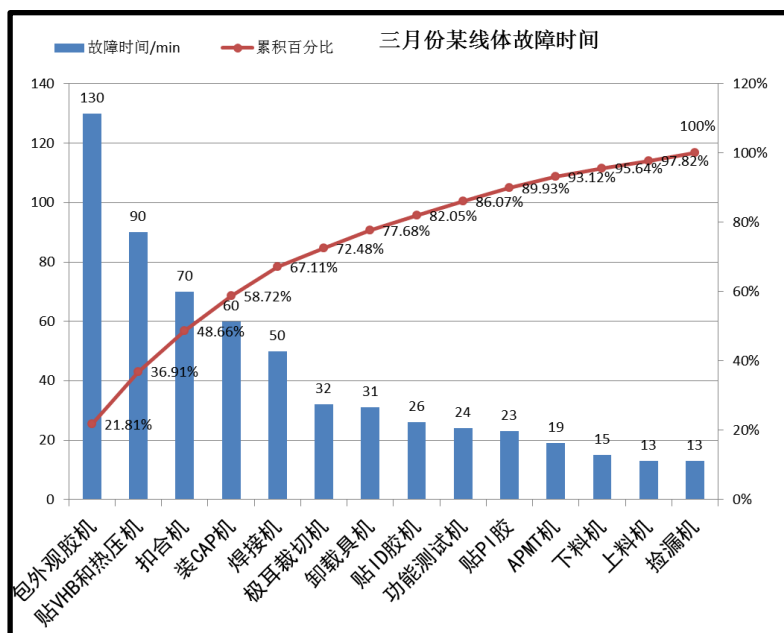


图 3-7 三月份某线体故障时间柏拉图

(5) 对柏拉图的坐标格式以及字体大小进行美化。

(6) 对完成的柏拉图进行分析，如上图所知，包外观胶机、贴 VHB 和热压机、扣合机、装 CAP 机和焊接机的故障时间相对而言较多，应重点关注。

**【例题二】**某工厂生产电芯产量一天可达到 2 万以上，每日的不良品也较多，其中生产过程中产生的不良品种类有：气泡、外观胶褶皱、焊点焊穿、少焊点等坏品（见下

表 3-4),工厂为了提高产品品质,减少成本,收集数据制成柏拉图来进行分析解决问题。

表 3-4 PACK 制成不良品数

坏品	数量	电池图片
气泡	278	
外观胶褶皱	199	
焊偏	89	
少焊点	75	
焊穿	24	
炸焊	17	
掉墨	15	
漏铝	14	
漏装Cap	14	
漏防水胶	11	
漏ID胶	9	
离型纸未撕离	7	
脏污	5	
功能测试不良	2	
OCV不良	2	

柏拉图的制作具体步骤如下:

(1) 建立表格,并且按照降序的方式排列,并且计算累积百分比。如下表 3-5 所示。

表 3-5

坏品	数量	累计百分比	电池图片
气泡	278	37.3%	
外观胶褶皱	199	64.0%	
焊偏	89	76.0%	
少焊点	75	86.0%	
焊穿	24	89.3%	
炸焊	17	91.5%	
掉墨	15	93.6%	
漏铝	14	95.4%	
漏装Cap	14	97.3%	
漏防水胶	11	98.8%	
漏ID胶	9	100.0%	
离型纸未撕离	7	100.9%	
脏污	5	101.6%	
功能测试不良	2	101.9%	
OCV不良	2	102.1%	

(2) 插入表格,利用表格中两个系列的数据分别做成柱状图(坏品数量)和折线图(累加百分比)选择第一个数据系列,生成一个柱状图。如下图 3-8 所示

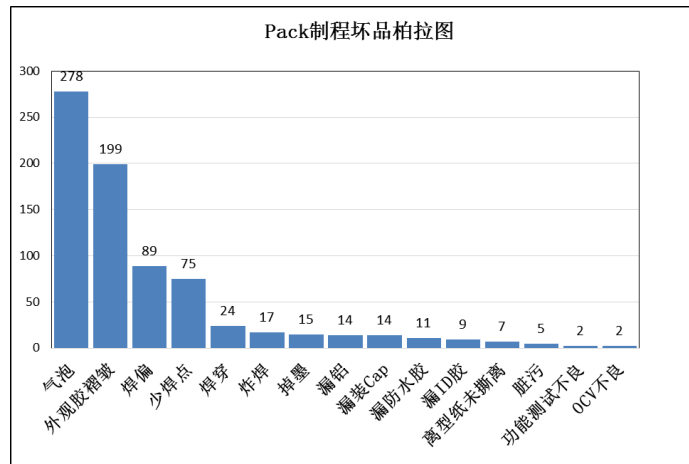


图 3-8 Pack 制程坏品柱状图

(3) 制作第二个系列数组，单击图表选中，鼠标右键打开菜单，在菜单中单击“选择数据”如下图 3-9。

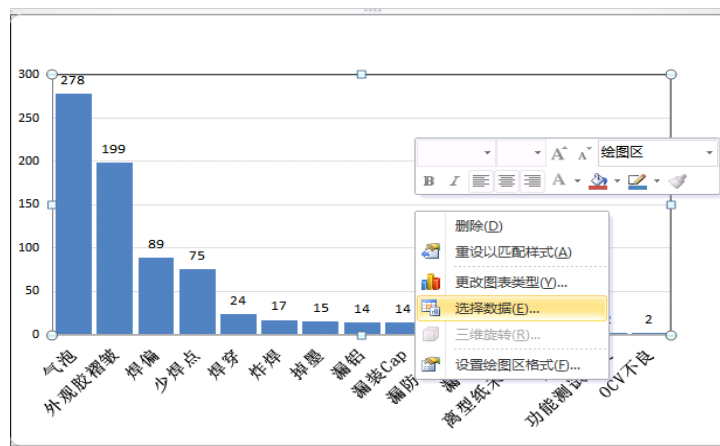


图 3-9 柱状图

(4) 在“选择数据源”窗口单击添加数据，出现编辑数据系列，系列名称选择表中“累积百分比”，系列值选择图示红色圈内单元格数据，如图 3-10 所示。

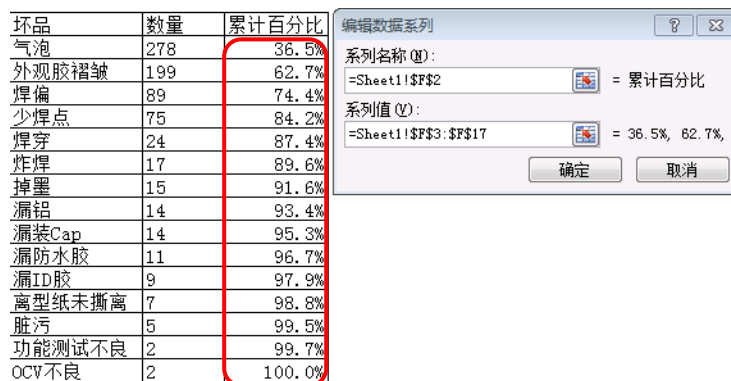


图 3-10

(5) 更改图表类型，选中“累积百分比”数据，单击右键菜单选择“更改图表类型”，选择折线，单击确定退出。如图 3-11 所示

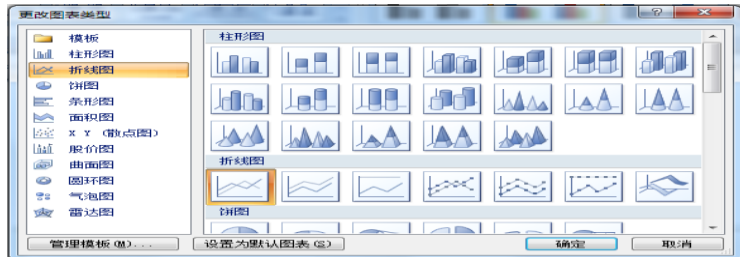


图 3-11

(6) 将系列二图表格格式设置为“次坐标”，选中“系列二”，鼠标单击右键打开“设置数据系列格式”。在系列选项中选择“次坐标轴”。如下图 3-12。

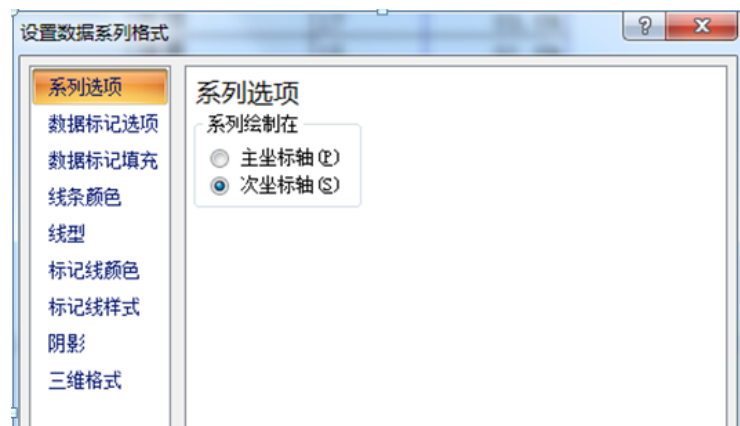


图 3-12

(7) 设置完成后按“关闭按钮”，单击“累积百分比”右键选择添加数据标签。设置字体格式并调整位置，如图 3-13 所示

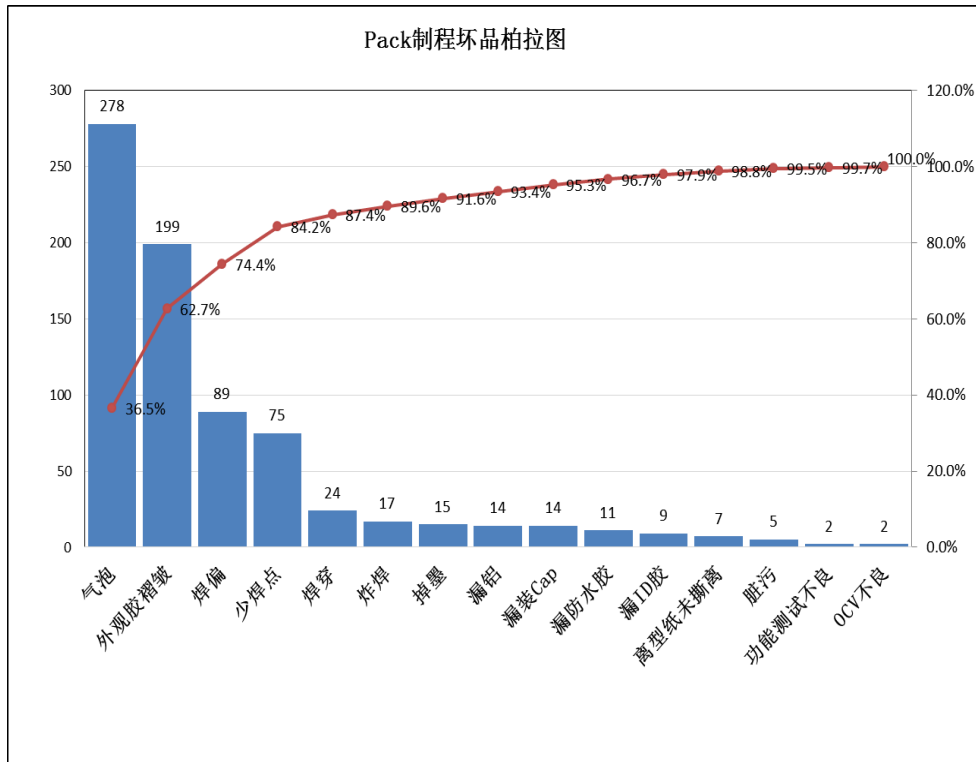


图 3-13 柏拉图

### 3. 柏拉图的用途

柏拉图能够在生产中出现的品质问题时能快速找出影响产品质量的主要因素并识别质量改进的机会。在制作柏拉图分析问题时，应用鱼骨图和对策表同时对问题进行分析，会起到非常大的帮助，本书下一节将对鱼骨图进行介绍。

## 三、鱼骨图

### 1. 鱼骨图背景介绍

鱼骨图是由日本管理大师石川馨 (Ishikawa) 先生所发明，鱼骨图又名因果图，是一种发现问题根本原因的分析方法，因其形状很像鱼骨，被称为鱼骨图。

鱼骨图的作用：针对柏拉图中显现出来的问题进行深入解剖，从六大层面分析造成故障、缺陷的原因，进行针对性的改善和改良，减少百分占比，达到零故障的目标。



石川馨

鱼骨图的使用步骤：首先确立问题点，收集相关数据，并对数据按六方面（人、机、

料、法、环、测) 进行分类, 接着对层次中的数据进行分类, 分辨从属关系, 画出层次鱼骨图。最后对各个层次进行汇总, 形成鱼骨图, 查看内容是否遗漏或重复, 完成一幅鱼骨图。

要因图如下图 3-14 所示

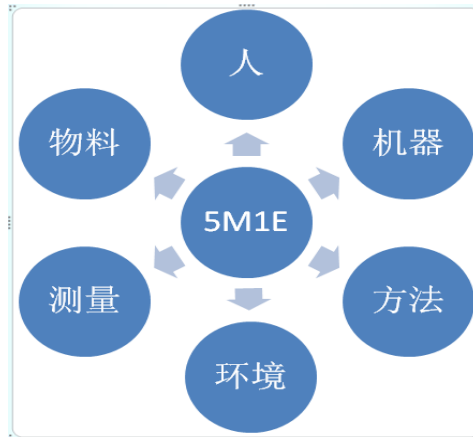


图 3-14 要因图

## 2. 鱼骨图的应用范围

分析因果关系, 表达因果关系;

通过识别症状分析原因, 寻找措施, 促进问题解决, 如图 3-15 所示。



图 3-15

## 3. 如何绘制鱼骨图

工业生产中, 鱼骨图经常会用来分析问题的主要原因, 所有学会绘制鱼骨图是非常重要的。绘制鱼骨图的具体步骤如下:

1. 填写鱼头, 也就是问题点, 画出鱼的主要骨头;
2. 画出大骨, 填写大要因, 通常选择人, 机, 料, 法, 环, 测作为鱼骨图的大骨,

见下图 3-16;

3. 画出中骨, 小骨, 填写中小要因 (见下图 3-16);



4. 关键的要点需要重点标识。

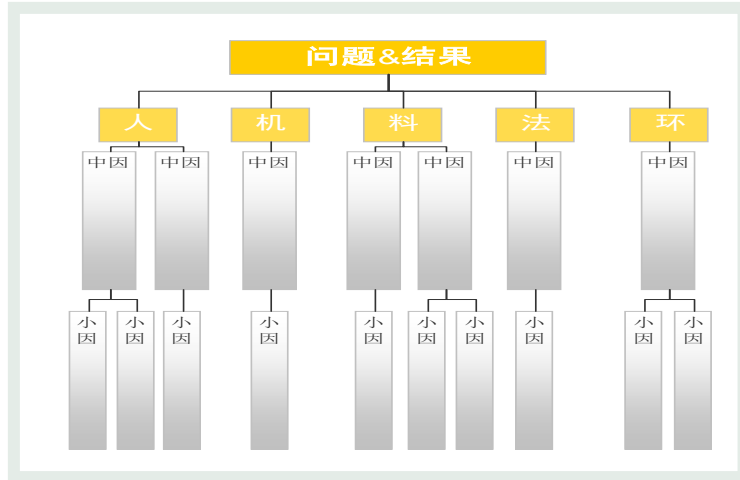
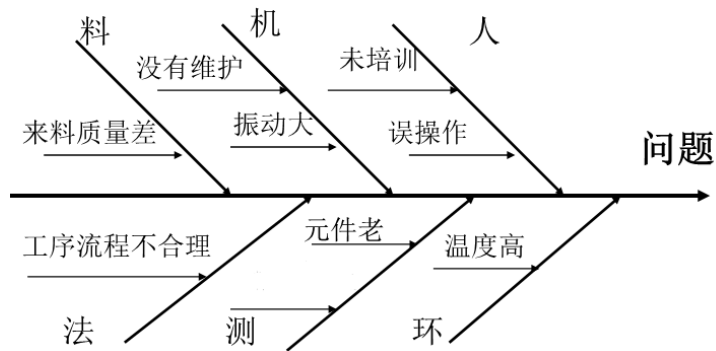


图 3-16

常见的鱼骨图如下图 3-16 所示。



3-17 常见的鱼骨图

【例题 1】某车间来料电芯表面油漆破损，针对电芯表面喷漆不良，维护人员从人机料法环测六方面展开分析，做出鱼骨图如下图 3-18 所示

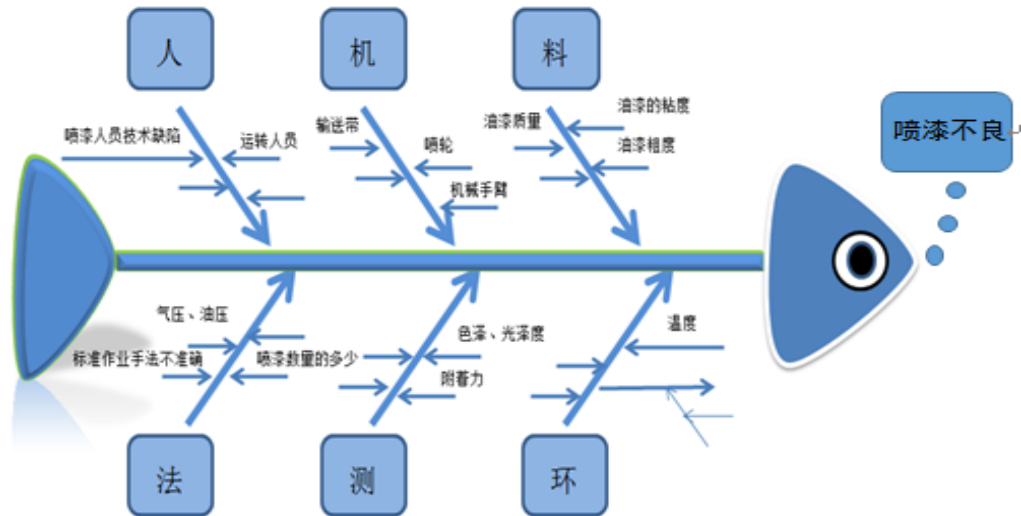


图 3-18 喷漆不良鱼骨图

#### 4. 画鱼骨图需要注意的事项

(1) 首先要明确结果，比如“电芯不合格”，未明确哪个地方不合格，应该指出具体的如尺寸不合格、喷漆不合格等。

(2) 区分目的，要做的是改善还是保持现有状态。如果要改善那么就要改变平均值，如果是保持就要缩小波动。由于寻找因素的着眼点不同，对改善和保持现状要明确区分。

(3) 确保每一个因素都分析到位，积极听取众人的意见特别是现场人员的意见，力求分析结果无一遗漏。

(4) 一个问题一个鱼骨图，比如电芯头部胶不良和表面脱漆，这里就要做两个鱼骨图来分析原因。

(5) 将鱼骨图中的要因都区分出来，对原因审查，确定重要的因素，并尽快实施解决措施。

## 四、检查表

### 1. 背景介绍

检查表也称为调查表、核对表是指特别为分层并采集数据而使用的表。

检查表是可以单独针对 Y 或是针对 X 的，也可以同时针对 Y 和 X 收集数据以下是几种常用的检查情况示例。如下表 3-6 所示。



表 3-6

采集数据对象	针对Y		同时针对Y和X	
	计量型	计数型	单个X, 单个Y	多个X, 多个Y
数据类型	计量型	计数型	单个X, 单个Y	多个X, 多个Y
用途	频数直方图	柏拉图	按位置不同记录的不良散点图	散点图矩阵

2. 检查表的使用目的?

- ①用于记录;
- ②用于调查;
- ③用于日常管理。

3. 常见的统计检查表

(1) 不合格品项目调查表

应用区域主要是调查生产车间中存在的合格品的频数和不合格率, 整理好的不合格项目调查表方便于研究、分析柏拉图等。

例如, 表 3-7 是某加工厂, 检验成品电池, 对电池尺寸不合格做了检查记录。检查了 100 批电芯, 平均不合格率为 0.5%。从 APMT 上可以看出其中 FAI3、FAI5、FAI6 尺寸的问题较多。

表 3-7 成品尺寸抽样项目调查表

批次	产品编号	不合格品数	批不合格率	尺寸不合格项					
				FAI3	FAI4	FAI5	FAI6	FAI9	FAI49
1	XWD001	5	0.5	1		2	1		1
2	XWD002	6	0.6			1	3		
3	XWD003	3	0.5		1		1		1
4	XWD004	8	0.8	1		1	1	1	
5	XWD005	6	0.6		1	1	2		1
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
100	XWD00100	6	0.6	1		1	3	1	2
合计		50	0.5	1	1	18	15	5	10

(2) 缺陷位置调查表

缺陷位置调查表在工厂中运用于成品外观的检测脏污、破损程度、产品表面的斑点、



气孔等，可以从缺陷位置调查表中记录数据，分析成品产品的质量问题的部位和密集程度，从中找出其规律。例如，某工厂，成品电池的头外观胶纸褶皱、破损、露耳朵、气孔、破损等缺陷的质量调查表如表 3-8。

表 3-8 成品电池外观缺陷表

序号	种类	频数	累计	累计频率
1	褶皱	3367	3367	69.14%
2	破损	521	3888	79.84%
3	划伤	382	4270	87.69%
4	露耳朵	201	4471	91.82%
5	气孔	156	4627	95.02%
6	脏污	120	4747	97.48%
7	其他	123	4870	100%

### (3) 点检表

点检表在工厂中会经常运用到，在确认点检表时制表人员需要对必要点检的项目进行周密的计划，详细列出所需要点检的项目，有效的防止遗漏。点检表适用于设备维修、5S 管理、TPM 活动小组情况自查等，点检表对工厂的日常生产有着极为重要的意义，下表 3-9 为某工厂 APMT 设备点检表。

表 3-9 APMT 点检表

APMT点检表																						
产线编号: _____												文件编号: F-W1-FE-E-BU2-0047-01/X0			年/月: _____							
责任部门: _____												设备名称: APMT			班次: _____							
工艺流程: 电池上料→电池扫码/定位→OMM测量电池→OIM测量电池→PPG测量电池→电芯下料																						
序号	点检项目	点检内容	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1	清洁	机台内无电芯																				
		治具、吸嘴和激光镜头用无尘布擦拭干净，无杂物残留																				
		作业员签名																				
2	设备运行	设备全轴找原点成功，设备正常启动																				
		X1441Pack与自动放料检测软件能够正常打开																				
3	真空气压值	安全门正常开启																				
		电池方案 气压值(0.3-0.45)																				
4	标准块	FAI2-OMM (44.5966-44.6366)																				
		FAI93-OMM (100.0979-100.1379)																				
		FAI83-OMM (36.1041-36.1441)																				
		FAI87-OMM (44.5966-44.6366)																				
		FAI88-L-OMM (0.6804-0.7204)																				
		FAI88-R-OMM (0.6807-0.7207)																				
		点检结果 (OK/NG)																				
		点检人																				
注1: FAI尺寸由点检人将实际数据填入点检栏内并判定点检结果,其余: "√" 正常 "×" 异常 "△" 未使用 注2: 确认人由当班负责人确认																						

### 4. 检查表的使用要点:

- (1) 观察事实，观物要细心，客观；
- (2) 收集数据应能获得层次的情况；
- (3) 注意样本取得之随机性与代表性；
- (4) 观察数据是否集中在某些项目，各项目有何差异；
- (5) 明确检查的责任者，并告诉其收集数据的目的及方法。

## 五、直方图

### 1. 直方图的概念

直方图又称为质量分布图，是一种统计报告图，是柱状图的特殊图形，在图中各个数据间的距离为零或接近零，作法简单，且又形象、直观，在企业或工厂中应用广泛。一般用横轴表示数据值，纵轴表示频数情况。直方图如下图 3-19 所示。

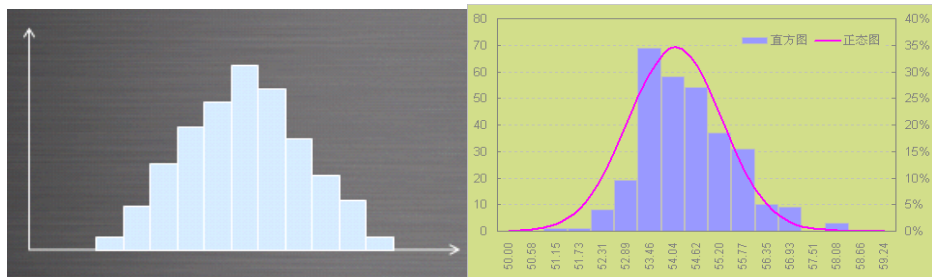


图 3-19 直方图

### 2. 直方图的特点

①突出显示了数据的集中趋势；②了解数据的离散程度；③确定了样本分布是对称的还是偏斜的。

### 3. 直方图在应用上必须注意事项：

- (1) 生产中的制程是否出现异常，可以通过直方图来观察。
- (2) 产品规格分布图案可与标准，标准规格作比较，有多大的差异。
- (3) 根据直方图分析是否需要进一步的深入了解。

### 4. 直方图的制作步骤

频数直方图的制作是建立在数据频数分布表的基础上的，它的制作方法如下：

- ①确定要分析的数据；
- ②计算极差值： $R = \text{Max} - \text{Min}$ ；



③根据样本数，确定组数；

④确定组距： $H=R/K$ ；

⑤确定各组的临界点和中心点；

⑥绘制频数分布表；

⑦建立坐标系，以各组边界值做横轴，纵轴为频数，在坐标系上画出以组距  $h$  为宽，频数为高的直方。下面以某工厂实际案例讲解对直方图的制作步骤进行介绍。

**【例题一】**某工厂的机台零件规格为  $310 \pm 8\text{mm}$ ，今抽验 50 件零件长度数据如下表 3-10，分析其加工长度的情况。



表 3-10 零件长度规格表

零件长度（单位mm）				
308	317	306	314	308
315	306	302	311	307
305	310	309	305	304
310	316	307	303	318
309	312	307	305	317
312	315	305	316	309
313	307	317	315	320
311	308	310	311	314
304	311	309	309	310
319	312	312	312	318

解题思路：

①根据表中数据，计算数据的极差：

Min=320，Max=为 302，

根据极差公式： $R=320-302=18$

②选择组数；

根据样本数  $n=50$ ，查下表 3-11 选取组数  $k$ ，这里组数  $k$  范围为  $6\sim 10$ ，我们这里取  $k=7$ 。

表 3-11 组数选取表

n	k
50	5~7
50~100	6~10
100~250	7~12
250	10~12

③计算组距；

组距公式为： $H=R/K$ ；

根据公式得  $R/K=18/7=2.5$ ，这里取整数  $H=3$ 。

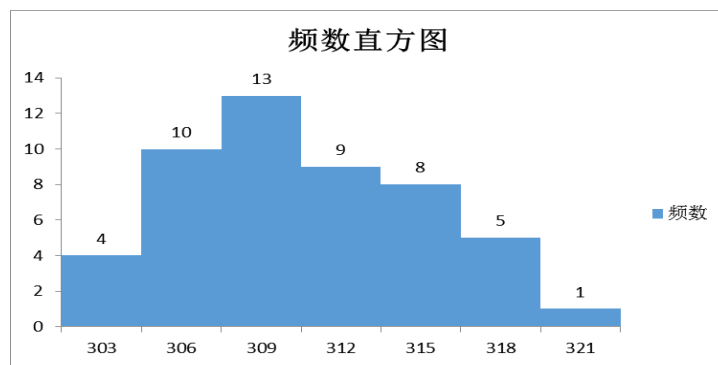
④确定各组临界点和中心点；

临界点：为避免数据落在边界上，要使得最小值落在第一组内，故取第一组的下限等于最小值减去最小测量单位的一半，固第一组下侧边界值为 301.5，依次可得出如下表 3-12 所示。

3-12 频数分布表

组数	组界	中心值	频数
1	301.5~304.5	303	4
2	304.5~307.5	306	1
3	307.5~310.5	309	1
4	310.5~313.5	312	9
5	313.5~316.5	315	8
6	316.5~319.5	318	5
7	319.5~322.5	321	1

根据频数分布表，制作例题的频数直方图如下图 3-20 所示。



3-20 频数直方图

### 5. 数据的频率直方图

频率直方图如下图 3-21 所示。

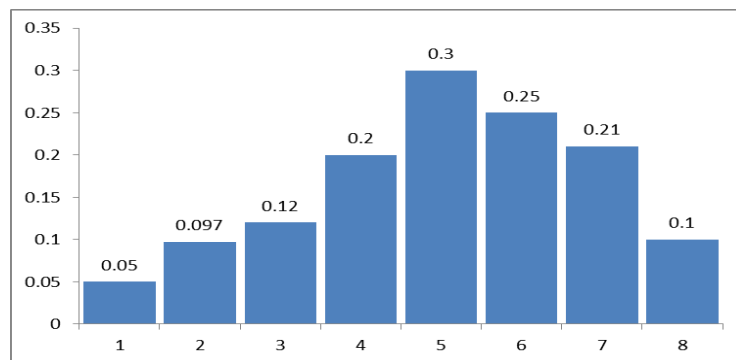


图 3-21 频率直方图

特点：频率直方图可以帮助我们分析数据在每一组中所占的比例的变化情况，其中与频数直方图的区别在于纵轴不相同。

【例题二】下表是某工厂一周生产 A 产品的数据表，算出每天生产频率和画出相关频率直方图，A 产品数据表如下 3-13 所示。

表 3-13 A 产品生产统计表



日期	周一	周二	周三	周四	周五
数量	4500	4800	5600	4900	4700

解题思路：首先计算总数 Z； $Z=4500+4800+5600+4900=24500$

其次分别用每天的数量除以总数得到相应频率计算 A 产品频率图如下表 3-14 所示。

表 3-14 A 产品统计表

日期	周一	周二	周三	周四	周五	总数
数量	4500	4800	5600	4900	4700	24500
频率	18.37%	19.59%	22.86%	20.00%	19.18%	100%

绘制频率直方图如下图 3-22 所示。

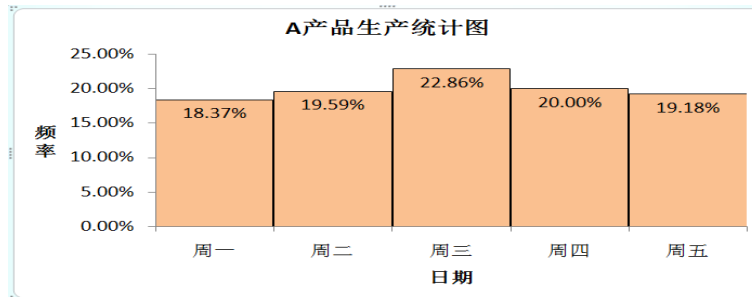


图 3-22 生产统计图

## 6. 常见直方图

下面是过程控制中常见的几种直方图，因产品类别、设备、加工等方法不同而有异，以下分析仅供参考。

(1) 正常型：特点是中间高、两边低、左右基本对称。这是数据服从正态分布的特征，也是大多数产品质量特性所具有的图形特点是中间高、两边低、左右基本对称。如下图 3-23 所示。

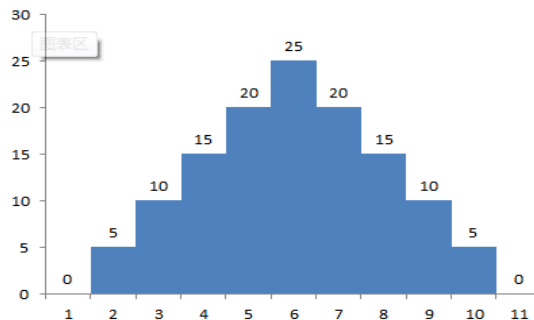


图 3-23 正常型

(2) 偏向型：特点也是中间高、两边低、但高峰偏向一侧，形成不对称形状。（人为控制）如下图 3-24 型直方图所示。

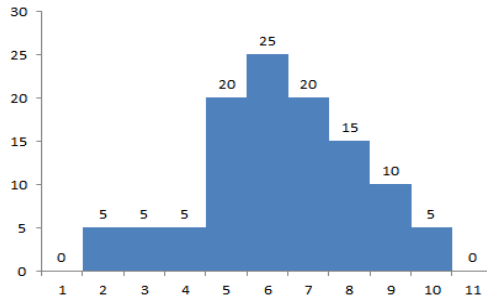


图 3-24 偏向型

(3) 双峰型：特点是两个高峰。这往往由于来自两个总体的数据混在一起所致，如可能两个线体加工的产品混在一起，不利于作数据分析，建议区分后作数据分析。如下图 3-25 双峰型直方图所示。

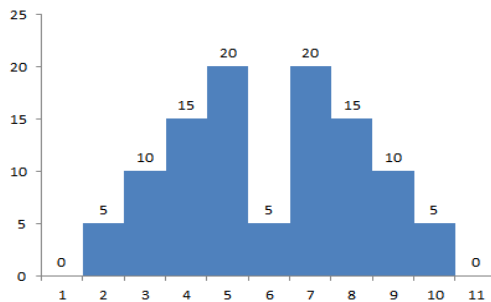


图 3-25 双峰型

(4) 低峰型：可能由于过程中某种倾向性因素缓慢作用所致，如下图 3-26 所示。

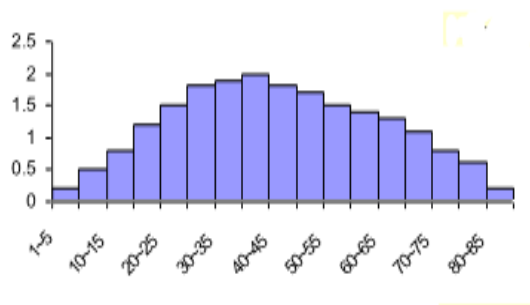


图 3-26 低峰型

(5) 锯齿型：特点是直方图内各直方高低参差不齐，其原因可能是直方图分组不当(过多)或测量误差过大而所致，如下图 3-27 所示。

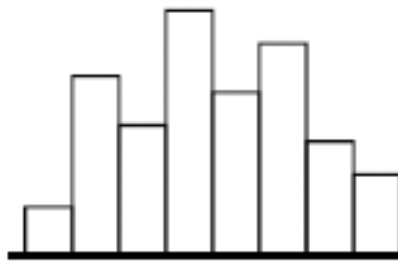


图 3-27 锯齿型

## 六. 散点图

### 1. 散点图的背景

在做大数据时，采用数据表格说明的方式不够直观，很难一眼看出问题的重点，而采用散点图在做大数据分析前可以直观的观察数据集中分布。

散点图又称相关图也叫散布图，它是指用来研究两个变量之间是否存在相关关系的一种图形。散点图如下图 3-28 所示。

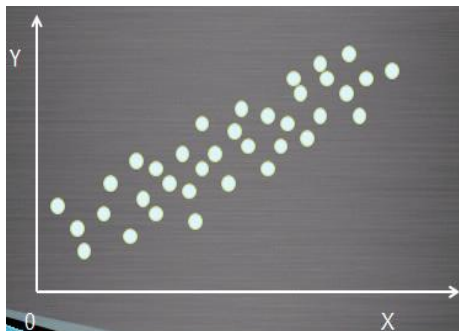


图 3-28 散点图

### 2. 适用于有较大数据或难以发现规律的数据中。

①在一组数据中，确定上下极限。筛选出不在范围内的数据，用散点图与之上上下极限作对比，进行分析得出结论。

②在两组数据中，两组数据放在同一散点图中，观察散点图的线性情况，进行对比分析，确定数据稳定情况，得出结论。

### 3. 散点图的制作

**【例题一】**某工厂生产的电芯，在电芯边距的改善前后做了大量的数据发现运用图表（见下表 3-15）的方式不够直观很难发现其中的突出点，可通过运用散点图表达。

表 3-15 电芯边距前后改善数据表

描述数据	改善前		改善后	
	左边距	下边距	左边距	下边距
均值	50.98	9.98	51.02	9.99
标准差	0.94	0.89	0.25	0.31
总投入数	900	900	900	900
最大值	53.60	12.61	52.17	10.91
最小值	48.09	6.77	50.15	9.04
超下限数量	136	10	0	0
超上限数量	132	10	1	0
良率	70.2%	97.8%	99.9%	100.0%

这里介绍 Excel 工具表如何制作散点图步骤如下：

第一步：将收集到的数据整理，并且选中改善前所有数据，如下表 3-16 所示。

表 3-16 电芯边距前后改善数据表

改善前		改善后	
左边距	下边距	左边距 1	下边距 1
49.75	9.00	51.21	10.00
50.24	10.03	50.96	9.80
50.43	10.28	50.64	9.94
48.60	10.44	51.17	10.28
51.24	9.99	51.53	9.93
53.83	9.83	51.20	10.23
51.65	9.79	50.17	9.86
52.26	9.18	50.55	9.70
51.22	9.41	50.76	9.96
51.64	8.60	50.55	9.94
50.54	9.43	51.50	10.20
50.96	10.37	50.77	9.95
51.22	10.85	50.78	10.39
51.13	10.21	50.88	10.38

第二步：创建改善前的散点图表，在选中改善前的所有数据后，单击“插入”选项，选择“推荐图表”，在图表下拉列表中选择“散点图”。如图 3-29 所示。

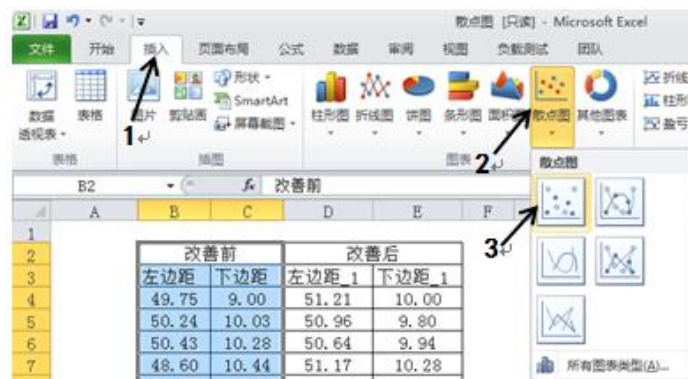


图 3-29

第三步：添加第二个系列数组，选中“改善前的散点图”，单击右键在菜单中选择“选择数据”，弹出“选择数据源”窗口，单击添加按钮。如下图 3-30 所示。

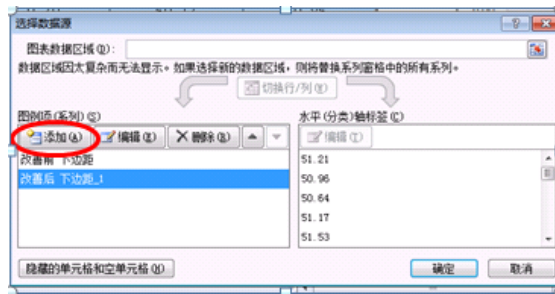


图 3-30

第四步：在弹出编辑数据窗口中，选择“系列名称”选择“改善后”单元格，X 轴系列值 (x) 选择区域为“左边距 1”所有数据，Y 轴系列值 (Y) 选择区域为“下边距 1”所有数据。如下图 3-31 所示

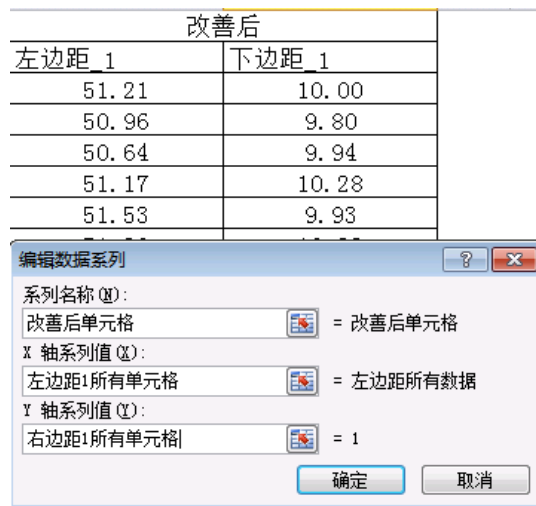


图 3-31

第五步：更改标题，双击选中标题文本框，输入“喷码位置改善前后对比”。

第六步：设置坐标轴格式，散点图分布主要集中在 X 轴方向的 48 到 54 之间。设置坐标轴格式的具体步骤为：①单击 X 轴坐标，鼠标右键打开菜单，选择“设置坐标轴格式”后弹出对话框，在最小值输入 48，最大值输入 54 即可。见图 3-32。

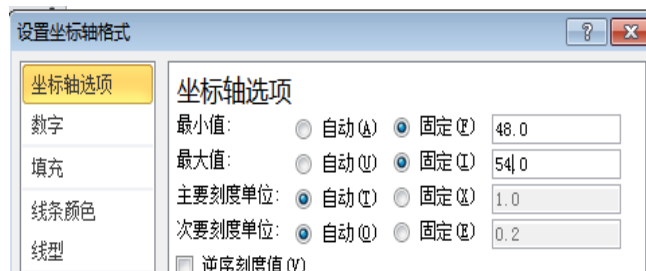


图 3-32

第七步：完成后的散点图如下图 3-33 所示。

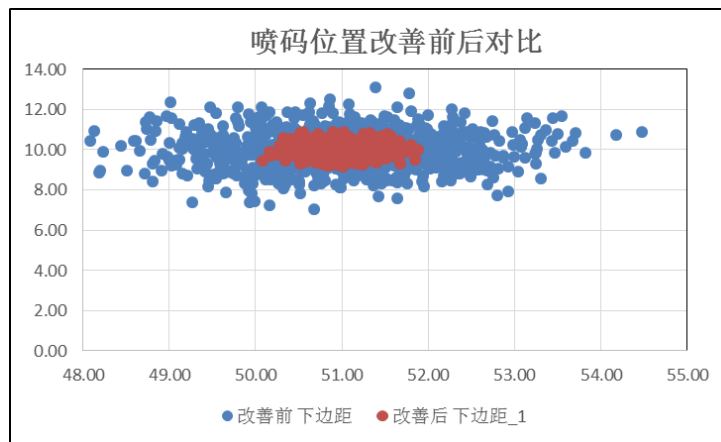


图 3-33 散点图

### 3. 散点图与表格统计数据的对比

表格统计数据，较难马上发现改善后左下边距的位置变化。散点图，一眼就可以看出红色区域为改善后边距的位置，较为集中。

【例题一】以下是收集部分省市离婚比例与人均寿命的数据，如下表 3-17 所示，请在纸上绘制一张散点图。

表 3-17 部分省市离婚比与人均寿命关系

序号	省市	离婚结婚比%	人均寿命
1	安徽	28.09%	75.08
2	北京	50.60%	80.18
3	广东	24.59%	76.49
4	广西	26.10%	75.11
5	江苏	35.11%	76.63
6	青海	17.06%	69.96
7	山东	33.98%	76.46
8	上海	43.61%	80.26
9	西藏	10.53%	68.17
10	浙江	36.32%	77.73

解：首先建立坐标系，纵坐标轴表示为人均寿命数据，横坐标轴表示离婚比例。

第二步：在坐标轴上依次将 (75.08, 28.09), (80.18, 50.60), (76.49, 24.59),

(75.11, 26.10), (76.63, 35.11) (69.96, 17.06) (76.46, 33.98) (80.26, 43.61)  
(68.17, 10.53) (77.73, 36.32) 的点标出, 得出如下图 3-34 所示, 可以得到, 人均  
寿命与离婚比率的相关关系。

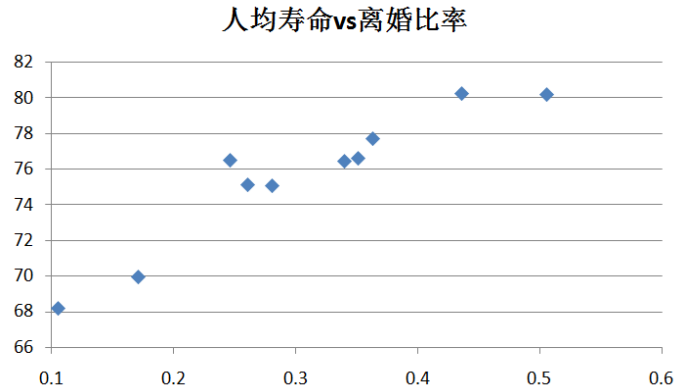


图 3-34 离婚比率与寿命相关性

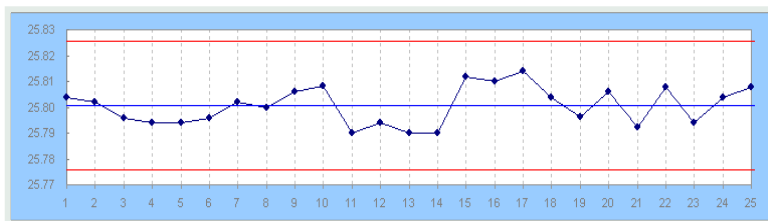
## 七、控制图

### 1. 控制图的起源与发展

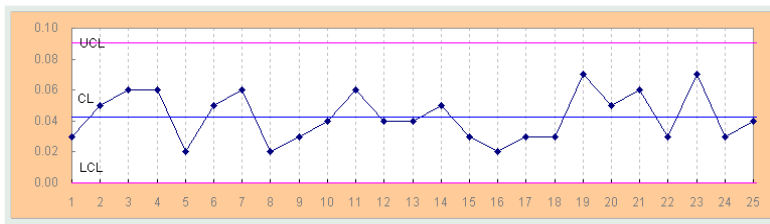
1924 年美国贝尔电话研究所的休哈特制作了世界上第一张控制图, 并且在 1931 年的时候, 休哈特教授正式发表了一份《工业产品质量的经济控制》之后, SPC 应用于各种制造过程改善便就此展开。

1941-1942 年形成美国行业标准, 主要有质量控制指南、数据分析控制方法, 并在生产过程中应用过程控制。

全世界推广, 逐步推广到美国民用工业区域全世界广泛用于各行各业, 下图 3-35 为两种常用类型的 SPC 图。



X-bar图



R 图

图 3-35 SPC 图

## 2. 控制图的定义与分类

控制图也称 SPC 图，是将一个过程定期收集的样本数据按顺序点绘而成的一种图示技术。它可展示过程变异并发现异常变异，从而促使人们采取预防措施。控制图已成为过程控制的重要手段。

控制图按照类型划分可以分为两大类型：计量值控制图和计数值控制图。如下图 3-36 所示。

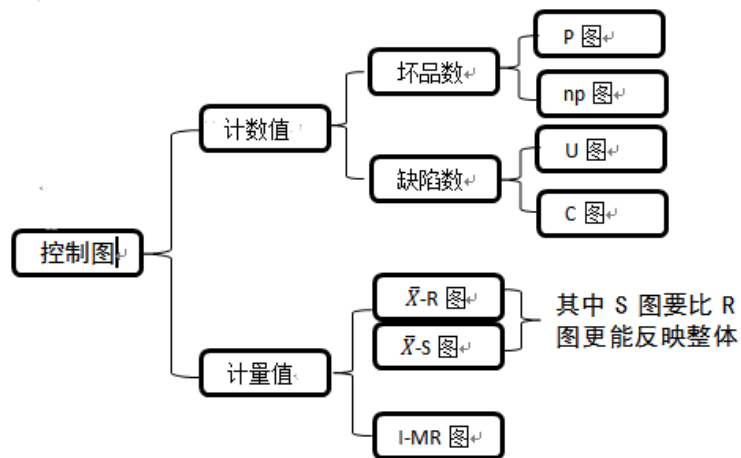


图 3-36 控制图的分类

### (1) 计数值控制图

计数值控制图是离散的数值，主要在数轴上呈现为整数形式。比如，某工厂一天生产的不合格品的件数。计数值控制图的目的是通过数据分析，对产品的制作工序加以控制并达到稳定状态。为了让产品的不合格率降低，保证产品的质量。它适用于产品的质量特性为计数的情形。例如，某工厂生产电芯，其中的不合格品、不合格率、缺陷数、单位紧缺数等的离散变量。常用的计数值控制图及应用有：

1. 不合格品率控制图 (P 图)
2. 不合格品数控制图 (np 图)





### 3. 单位缺陷数控制图 (U 图)

### 4. 缺陷数控制图 (C 图)

#### (2) 计量值控制图

计量值控制图是连续型的数值，在数轴上表现为所有点。在测量精度达到要求时，它的特征值的表现形式为任意精度表示，比如，某工厂生产一批零件的长度、重量、时间等连续的变量。常用的计量值控制图及应用有：

1. 均值-极差控制图，也称为 $\bar{X}$ -R 图，在统计计量数据的范围中， $\bar{X}$ -R 图是最常见的、最基本的控制图。主要应用场合是控制对象为：长度、时间、厚度、生产量的计量值场合。

2. 中位数-极差控制图，也称为 $\tilde{X}$ -R 图。

3. 单值-移动极差控制图，也称为 I-MR 图。

4. 均值-标准差控制图，也称为 $\bar{X}$ -S 图。与 $\bar{X}$ -R 图相似，因为极差的计算公式极为简单，所以 $\bar{X}$ -R 图能够被广泛的应用，但如果样本量 N 大于 10 的时候，用 $\bar{X}$ -R 图估计标准差 $\sigma$ 显得效率不够，这时候需要应用到 $\bar{X}$ -S 图。同时，现在计算机的高速发展，使得 $\bar{X}$ -S 图的计算越来越方便，故 $\bar{X}$ -S 图的应用也会越来越广泛。

### 3. 控制图的用途

控制图的主要用途：

①在生产过程中，会因为许多因素影响生产的制程，影响生产过程的稳定，控制图可以通过分析和判断影响生产的过程的因素，保证生产过程处于统计控制状态。

②当生产过程中出现缓慢变异或者异常现象时，在控制图中可以及时的发现这些异常，及时解决问题。

③控制图可以为评定产品质量提供依据。

④查明准确的设备和工业设备实际精度，以便做出正确的计数决定。

### 4. 控制图的基本构造

#### (1) 标题要素

在绘制控制图时，标题部分，比如产品、编号、质量特性、操作员、设备、数量、控制图的名称等，我们通常会将其内容编成表格的形式。例如，某工厂的紫铜管控制图数据表 3-18 所示。

表 3-18 数据表

样组	测定值					$\bar{X}$	R	样组	测定值					$\bar{X}$	R
	X1	X2	X3	X4	X5				X1	X2	X3	X4	X5		
1	50	50	49	52	51	50.4	3	14	53	48	47	52	51	50.2	6
2	47	53	53	45	50	49.6	8	15	53	48	49	51	52	50.6	5
3	46	45	49	48	49	47.4	4	16	46	50	53	51	53	50.6	7
4	50	48	49	49	52	49.6	4	17	50	52	49	49	49	49.8	3
5	46	48	50	54	50	49.6	8	18	50	49	50	49	51	49.8	2
6	50	49	52	51	54	51.2	5	19	52	49	52	53	50	51.2	4
7	47	49	50	48	52	49.2	5	20	50	47	50	53	52	50.4	6
8	48	50	46	49	51	48.8	5	21	52	49	51	53	50	51.0	4
9	50	50	49	51	53	49.0	4	22	55	54	51	51	50	52.2	5
10	49	51	51	46	48	49.2	5	23	50	54	52	50	49	51.0	5
11	51	50	49	46	50	49.2	5	24	47	51	51	52	52	50.6	5
12	50	50	49	52	51	50.4	3	25	53	51	51	50	51	51.2	3
13	49	49	49	50	55	50.4	6							1250	120

(2) 控制图要素

下图为控制图的基本格式（见图 3-37）

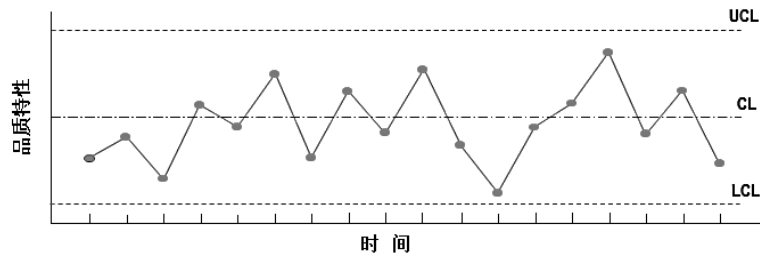


图 3-37 控制图

上图为控制图的基本格式，由图中可以看出，构成控制图的要素有：

横坐标作为时间的顺序，纵坐标作为品质的质量特性。

实线 CL 为中心线；虚线 UCL 为上控制限的临界线；虚线 LCL 为下控制限的临界线。

(3) 控制图的制作步骤如下：

- 1、收集数据；
- 2、建立控制限；
- 3、绘制控制图；
- 4、控制图的观察与分析。

下面以例题来对  $\bar{X}$ -R 图的制作步骤进行详细分析。

【例题一】某工厂加一标准件，其规格为 0.5~0.9mm，现在为了评估过程质量，工厂在 7 天的时间内总共测量了 25 组数据。数据图如下表 3-19 所示。



表 3-19

样组	测量值				
	X1	X2	X3	X4	X5
1	30	32	33	31	32
2	27	29	28	33	32
3	31	33	32	30	32
4	30	32	32	31	34
5	29	27	28	29	28
6	31	32	29	29	28
7	31	29	30	32	31
8	29	28	31	30	29
9	28	30	31	30	29
10	29	28	31	30	31
11	33	30	32	31	29
12	31	30	31	28	29
13	27	31	31	29	31
14	31	30	30	28	31
15	32	29	31	31	33
16	29	29	30	30	29
17	31	32	29	29	28
18	31	29	30	32	31
19	29	28	31	30	29
20	28	30	31	30	29
21	29	28	31	30	31
22	31	27	30	31	31
23	32	31	32	29	30
24	30	32	29	28	29
25	28	28	31	29	28

第一步：收集数据，将表格中的数据整理并且计算每个子组的样本的控制统计量（例如均值、极差等）。

计算公式如下：

平均值（Mean）

$$\bar{X} = X_1 + X_2 + X_3 + \dots + X_N / N$$

极差（Range）

$$R = \text{MAX} - \text{Min}$$

将计算结果整理制成如下表 3-20 所示。

表 3-20



样组	测量值					$\bar{X}$	R
	X1	X2	X3	X4	X5		
1	30	32	33	31	32	31.6	3
2	27	29	28	33	32	29.8	6
3	31	33	32	30	32	31.6	3
4	30	32	32	31	34	31.8	4
5	29	27	28	29	28	28.2	2
6	31	32	29	29	28	29.8	4
7	31	29	30	32	31	30.6	3
8	29	28	31	30	29	29.4	3
9	28	30	31	30	29	29.6	3
10	29	28	31	30	31	29.8	3
11	33	30	32	31	29	31	4
12	31	30	31	28	29	29.8	3
13	27	31	31	29	31	29.8	4
14	31	30	30	28	31	30	3
15	32	29	31	31	33	31.2	4
16	29	29	30	30	29	29.4	1
17	31	32	29	29	28	29.8	4
18	31	29	30	32	31	30.6	3
19	29	28	31	30	29	29.4	3
20	28	30	31	30	29	29.6	3
21	29	28	31	30	31	29.8	3
22	31	27	30	31	31	30	4
23	32	31	32	29	30	30.8	3
24	30	32	29	28	29	29.6	4
25	28	28	31	29	28	28.8	3

第二步：建立控制限，根据公式计算出控制图的中心线和控制限，其计算公式如下：

$$\begin{aligned}
 CL_{\bar{X}} &= \bar{\bar{X}} & CL_R &= \bar{R} \\
 UCL_{\bar{X}} &= \bar{\bar{X}} + A_2 \bar{R} & UCL_R &= D_4 \bar{R} \\
 LCL_{\bar{X}} &= \bar{\bar{X}} - A_2 \bar{R} & LCL_R &= D_3 \bar{R}
 \end{aligned}$$

其中， $A_2$ ， $D_3$ ， $D_4$ 查阅附表二计量型控制图系数表。

将 $\bar{X}$ 图的计算结果整理如下表 3-21 所示。

表 3-21

$\bar{X}$	$\bar{\bar{X}}$	UCL	LCL
	30.07	31.985	28.154

同理，根据公式计算出中心线及计量型控制图系数表，将 R 图的计算结果整理如下表 3-22 所示。

表 3-22

R	$\bar{R}$	UCL	LCL
	3.32	7.018	0

第三步，绘制控制图如下图 3-38 所示。

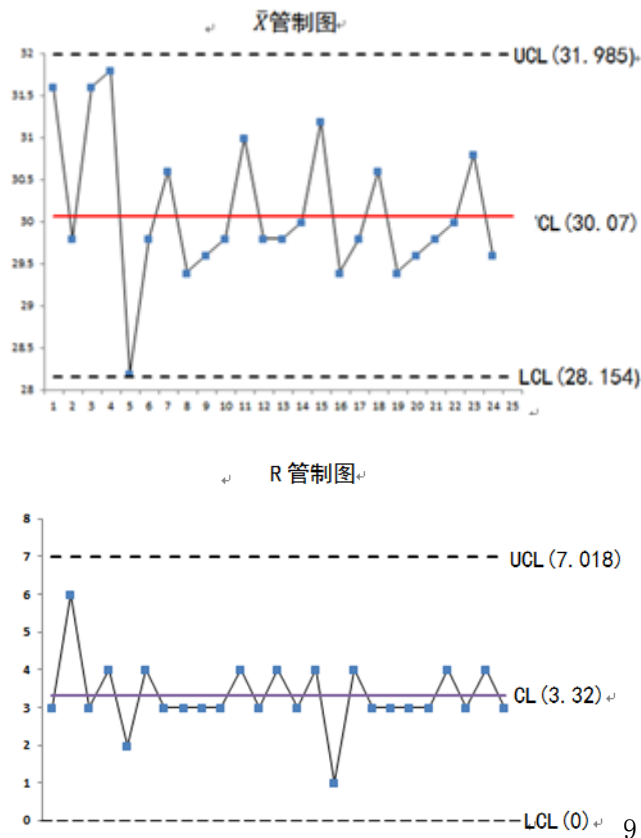


图 3-38  $\bar{X}$ -R 图

#### 第四步，控制图的观察与分析

通过观察数据点的排列与位置判断是否处于生产稳定状态，如，查看数据点是否在控制限之外，是否符合附录中的八大异常现象。在观察与分析时，应执行“查明原因，采取措施，加以消除，不再出现，纳入标准”这 20 字方针。

由上图可以看出数据点都分布在控制限之内且不符合八大异常现象，故该生产处于统计控制状态。

**【例题二】**根据某工厂生产的一批电芯，试试制作其不合格品数的控制图。

第一步：收集数据，将表格中的数据整理并且计算每个子组的样本的控制统计量，统计资料见表 3-23 所示。

表 3-23



样本组号	样本大小	不合格品数
1	100	2
2	100	3
3	100	4
4	100	0
5	100	4
6	100	3
7	100	4
8	100	5
9	100	0
10	100	2
11	100	3
12	100	2
13	100	0
14	100	0
15	100	9
16	100	2
17	100	1
18	100	1
19	100	1
20	100	2
21	100	3
22	100	4
23	100	0
24	100	0
25	100	2
合计	2500	48
p	0.019	
np	1.92	

第二步：建立控制限，根据公式计算出控制图的中心线和控制限，其计算公式如下：

$$CL = \bar{P}n = 2.72 \quad UCL = \bar{P}n + 3\sqrt{\bar{P}n(1 - \bar{P})} = 7.62 \quad LCL = \bar{P}n - 3\sqrt{\bar{P}n(1 - \bar{P})} < 0$$

第三步：绘制控制图如下图 3-39 所示。

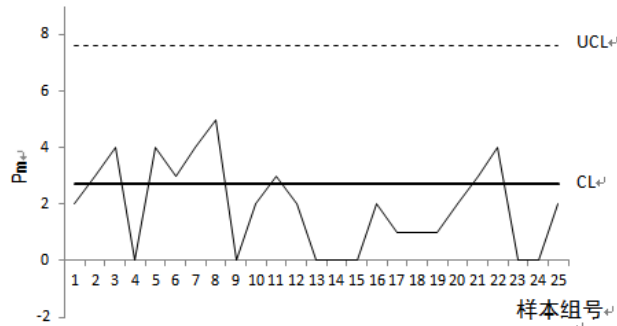


图 3-39 np 控制图

第四步：控制图的观察与分析

观察上图，数据点都在上下控制限之内，且不符合八大异常现象，此生产过程处于统计控制状态。



### 第三节 复习题

#### 一、单项选择题

- 直方图的特点有 ( )
  - 集中性
  - 简洁性
  - 连续性
- 下列选项中错误的是 ( )
  - 直方图又称为质量分布图。
  - 调查表是检查表
  - 控制图也被称为 SSP
- 常说的二八定律, 被称之为
  - Pareto 原则
  - 质量分布原则
  - 统计报告原则

#### 二、多项选择题

- 控制图分为哪几类 ( )
  - 计数值控制图
  - 计值控制图
  - 计点值控制图

#### 三、简述题

- 常用的层别法分类有哪些?
- 某注塑厂统计今天早, 中, 晚三个班次的注塑外观缺陷, 请汇总并绘制柏拉图进行来发现主要缺陷

缺陷分类	班次	数量	缺陷分类	班次	数量
气纹	早班	29	气纹	中班	28
黑边	早班	3	黑边	中班	8
毛边	早班	12	毛边	中班	9
拉伤	早班	16	拉伤	中班	11
色差	早班	3	色差	中班	2
流痕	早班	8	流痕	中班	3

缺陷分类	班次	数量
气纹	晚班	33
黑边	晚班	2
毛边	晚班	5
拉伤	晚班	15
色差	晚班	1
流痕	晚班	12

- 以下是收集某产品质量特性的 30 个数据如下表所示, 请绘制一张频数分布直方



图

75	46	64	50	47	21
1	81	41	34	23	39
50	53	45	68	57	54
11	29	88	39	53	47
44	52	67	53	77	99

4. 用你所选的项目或者公司的一个问题点，练习创建一个鱼骨图。

5. 某厂生产一种零件，其长度要求为  $49.50 \pm 0.10\text{mm}$ ，生产质量过程要求为过程能力指数不小于 1，为对该过程实施连续控制，试设计  $\bar{X}$ -R 图。

6. 某厂测得钢的淬火温度与硬度之间的成对数据如下表所示，现用散点图对着 30 对相关数据的相关程度进行分析研究。

序号	淬火温度	硬度	序号	淬火温度	硬度
1	810	47	16	850	43
2	890	45	17	840	48
3	850	42	18	830	52
4	890	59	19	880	51
5	830	50	20	9870	47
6	840	48	21	850	46
7	830	45	22	840	49
8	830	60	23	880	53
9	820	62	24	830	52
10	870	54	25	820	50
11	820	53	26	850	46
12	810	53	27	840	48
13	840	50	28	820	51
14	830	59	29	840	52
15	870	45	30	830	55





## 第四章 质量变异的规律性分析

### [教学目标]

- (1) 掌握数据的随机概率的概念
- (2) 掌握排列与组合的计算
- (3) 掌握离散性和连续性分布的计算与运用

### [教学难点]

- (1) 掌握离散性和连续性分布的计算与运用
- (2) 排列组合的常用技巧



## 第一节 随机概率及其分布

在自然界和人类社会普遍存在着两类现象，一类是在一定的条件下某些事情一定会发生，或者一定不会发生，例如，在常压下，水加热到  $100^{\circ}\text{C}$  就会沸腾，氢气和氧气混合物遇到火就会爆炸，我们向上抛一颗石子，它一定会落下来而不会飞走，我们称这类现象为确定性现象，另一类现象则不然，在一定条件下某些事情可能发生也可能不发生，例如，某人骑车上班需要通过 10 个路口，那么他会遇到几次红灯？如果一个人买了三张体育彩票，那么他一定会中奖吗？向上抛一枚硬币，下落后是正面朝上还是反面朝上，我们称这一类现象为随机现象，概率统计就是研究随机现象的数量规律的一门学科。

### 一、随机现象

#### 1. 研究对象

- ① 概率论与数理统计研究的对象为随机现象。
- ② 概率论研究的对象是概率分布。
- ③ 数理统计研究的对象是数据收集与处理

#### 2. 随机现象与确定性现象的定义

① 随机现象：按照结果来分，在一定条件下，有相同结果也有不同结果的现象称为随机现象。例如，掷骰子出现 6 点与买彩票是否会中奖。

② 确定性现象：按照结果来分，确定性分布只有一个结果。例如，水在标准大气压强下加热到  $100^{\circ}\text{C}$  会沸腾；掷一粒骰子，点数必定小于 7。

#### 3. 举例说明随机现象

随机现象的例子：

- ① 向空中抛一枚硬币，它掉落后可能会正面朝上，也可能会反面朝上。
- ② 掷一颗骰子，出现的点数。
- ③ 一副牌，抽到大小王。

确定性现象的例子：

- ① 口袋中放有 3 个相同的白色糖果，从里面抽取必然是白色糖果。
- ② 氢气和氧气混合物遇到火就会爆炸。

随机试验：对在相同条件下可以重复的随机现象的观察、记录、实验称为随机试验。



#### 4. 随机现象的特点

- ①从结果分析，不止一个结果。
- ②结果具有未知性。

## 二、样本空间与样本点

1. 样本空间：随机现象产生的所有基本结果所构成的集合，我们称之为样本空间，记为 $\Omega$ 。

2. 样本点：随机现象产生的基本结果称之为样本点，它是抽样中的基本单元，记为 $\omega$ 。

关系式： $\Omega = \{ \omega \}$ 。

3. 举例说明：下面给出随机现象的样本空间。

①一枚硬币的样本空间为 $\Omega = \{ \text{硬币正面朝上}, \text{硬币反面朝上} \}$ 。

②掷一粒骰子的样本空间 $\Omega = \{ \omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6 \}$ 。其中， $\omega_1$ =掷骰子数为1点， $\omega_2$ =掷骰子数为2点， $\omega_3$ =掷骰子数为3点... $\omega_6$ =投掷骰子数为6点。

③设备的寿命样本空间为 $\Omega = \{ t: t \geq 0 \}$ 。

4. 注意事项：

①样本空间中的元素不一定是数字。

②最基本的样本空间是含有两个样本点，如若只含有一个不能称之为样本空间。

③根据样本空间样本点数分类，可以分为有限和无限两类，在例题①、②中样本点数是可数的有限个，在③中的样本点个数为无限个。

## 三、随机事件

随机事件的定义：随机事件是从 $\Omega$ 中的某些样本点组成的集合称之为随机事件。例如，掷一枚骰子，得到偶数的事件记为A。即 $A = \{ 2, 4, 6 \}$ 是从 $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$ 某些样本点组成。

维恩图：概率论中常用于表示任何一个事件是相应的样本空间的一个子集，如下图4-1所示。

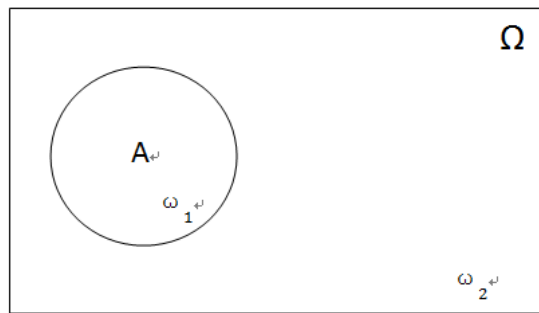


图 4-1 维恩图

基本事件：样本中单个基本元素所构成的子集称为基本事件。

必然事件：样本空间  $\Omega$  的最大子集即 ( $\Omega$  本身)。

不可能事件：样本空间中  $\Omega$  的最小子集 (即空集)。

举例说明：掷一枚骰子，它的样本空间为  $\Omega = \{1, 2, 3, 4, 5, 6\}$ 。

事件  $A =$  “出现大于 6 点” 在  $\Omega$  中，不存在大于 7 点的样本点，那么也就是说，事件  $A$  为空集，是不可能事件。

事件  $B =$  “出现的点数小于 7 点”，在样本空间  $\Omega$  中，任意一个样本点都小于 7 点，所以事件  $B$  为必然事件。

事件  $C =$  “出现 6 点” 在样本空间由一个样本点所组成，即基本事件。

## 四、概率的定义

### 1. 概率的定义

概率就是随机事件发生的可能性的的大小。

① 概率是有大小之分的，在一次抽奖活动中，一等奖 1 名，二等奖 10 名，三等奖 100 名，人们的共识是，抽到三等奖的概率比抽到一等奖的概率要大。

② 随机事件的的发生的可能性是可以无法度量的，例如一枚硬币正反面的可能性是相同的，都为  $1/2$ 。

③ 概率的原型：新生婴儿是男的还是女的，可以通过男婴出生率衡量；买彩票中奖与未中奖可以用中奖率度量。

2. 概率的公理化的定义：设  $\Omega$  为一个样本空间， $\mathcal{y}$  为  $\Omega$  中的事件域，且事件  $A, P(A)$  满足：

① 正则性公理， $P(\Omega) = 1, P(\Phi) = 0$ ;



②非负性公理：对任何事件  $A$ ,  $P(A) \geq 0$ ;

③可列可加性公理，若事件  $(i=1, 2, \dots)$  两两互斥，则有：

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

概率的公理化定义是判定是否为概率事件的标准，当满足上述三个条件时，就称之为概率，如果不满足，那么是不被称为概率的。

## 五、随机变量

(1) 随机变量的定义：我们常用大写的字母  $X$ ,  $Y$ ,  $Z$  来表示随机现象的结果的变量，称之为随机变量。

(2) 离散型随机变量：假如一个随机变量仅可能取有限个或可列个数值，我们则称为离散型随机变量。

(3) 连续型随机变量：假如一个随机变量在数轴上的一个区间  $(a, b)$ ，且它的取值范围无限，则称为连续型随机变量，

特点：

离散型随机变量的取值可以一一列举；

连续型随机变量的取值不能一一列举；

这个定义表明：随机变量  $X$  是样本点，这个函数可以是不同样本点对应不同的实数。也允许多个样本点对应同一实数，这个函数的自变量（样本点）可以是数也可以不是数，但是因变量一定要是实数。

与微积分中的内容不同，概率论中的随机变量  $X$  是一种随机取值的变量且伴随一个分布概率，以离散随机变量为例子，我们不仅仅要知道随机变量可能取哪些值，而且还需要知道它的概率是多少，这就需要分布的概念，有没有分布是区分一般变量与随机变量的主要标志。

## 六、频数、频率

频数：随机事件在一组数据或多次试验中出现的次数，或不同数据落在某区间的个数称为频数。例如：有 3 台设备的总气压的值落在  $0.4 \sim 0.6$  区间，即总气压在  $0.4 \sim 0.6$  这一随机事件的频数为 3。

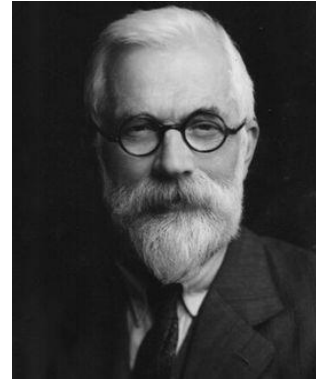
频率：随机事件发生的个数(次数)在总和观察数中所占的比率称为频率。通常记作 $f_n$ 。如上例事件若是从总数为100台设备中抽取的，则总气压落在0.4~0.6之间的事件A的频率为 $f_n(A) = 3/100 = 3\%$ 。

频率的稳定性的意义就在于它表明概率的存在，从而使得概率概念在现实世界有其参照物。

## 七、小概率事件

当某一事件发生的概率很小时(通常小于0.05)称小概率事件

著名的英国统计学家费希尔(Ronald Fisher)把二十分之一作为标准，也就是0.05，从此0.05或比0.05小的概率都被认为是小概率。



Ronald Fisher 是统计学家、生物进化学家、数学家、遗传学家和优生学家。人物如上图。



## 第二节 排列与组合

### 一、排列

1. 排列的定义：排列是设定有  $n$  个不同的元素，在  $n$  个元素中取出  $m$  ( $m$  个元素，将取出的  $m$  个元素按照一定的顺序排列成一列，称为从  $n$  个不同元素中取出  $m$  个元素的一个排列。

注意，当  $m=n$  时，在  $n$  个不同的元素中，将它全部取出，按照顺序排成一列，叫做  $n$  个不同元素的全排列。

2. 定义的理解，注意以下几点

①取出元素必须为  $m$  个不同元素，( $m \leq n$ )；

② $m$  个元素按照一定的顺序排列好；

③相同的排列是指元素相同且排列顺序也相同，要与相同的集合区别来，相同的集合只要求元素相同并不要求考虑顺序。

#### 【例题一】

1. 写出 A、B、C、D 中选取 2 个元素取出的所有排列。

解：A 开头的排列：AB；AC；AD；

B 开头的排列：BA；BC；BD；

C 开头的排列：CA；CB；CD；

D 开头的排列：DA；DB；DC。

2. 排列数的定义：假设有  $n$  个不同的元素，在  $n$  个不同的元素中取出  $m$  ( $m$ ) 个元素进行排列，对排列的个数称为从  $n$  个不同元素中取出  $m$  个元素的排列数，用符号  $A_n^m$  表示。

简单的说，排列数就是在满足条件下，有多少种排列方法。

3. 排列数的公式

$$A_n^m = n(n-1)(n-2) \cdots (n-m+1), \quad n, m \in \mathbb{N}^*, \text{ 且 } m \leq n.$$

4. 全排列数公式

当  $m=n$  时，叫做全排列数。

$A_n^n = n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1$ ，叫做自然整数  $n$  的阶乘，用  $n!$  表示。注意  $0! = 1$ 。



5. 重复排列从  $n$  个不同元素中每次抽取一个，取出后在放回去后再取下一个，按照这样的方法连续取  $m$  次所得到的排列称为重复排列，排列数为  $n^m$  个。

#### 6. 有限制条件的排列问题

(1) 注意事项：

- ①某些元素必须排列在或不能排列在某些位置；
- ②某些元素要求连排（必须相邻）
- ③某些元素要求分离（即不能相邻）

(2) 解决排列问题的常用方法：

①优限法：有特殊元素或特殊位置的排列问题，通常是先排列元素或特殊位置，称为优先处理特殊元素法也称为元素位置分析法。优限法当有两个以上约束条件时，往往实在考虑一个约束条件的同时还要兼顾其他条件。

②间接法：也叫排除法，对于含“至多”或“至少”等的排列问题，若直接解答需要进行复杂的讨论，可以考虑整体去杂，即将总体中不符合条件的排列或组合删除掉，从而计算出符合条件的排列个数的方法。

③捆绑法：某些元素要求必须相邻时，可以先将这些元素看作一个整体，与其他元素排列后，在考虑相邻元素内部排列的方法。

④插空法：某些元素不相邻排列时，可以先排列其他的元素，再将这些不相邻的元素插入空档的方法。

⑤缩倍法：在排列问题中限制几个元素必须保持一定顺序，可用缩小倍数的方法处理。

## 二、组合

1. 组合的定义：假设有  $n$  个不同元素，从  $n$  中取出  $m$  个元素合成一组，叫做从  $n$  个不同元素中取出  $m$  个元素的一个组合。

2. 排列与组合的区别：①排列是有顺序，组合没有顺序。②排列与组合都是从  $n$  个不同元素中取出  $m$  个元素。

3. 组合数的定义：假设有  $n$  个不同元素，从  $n$  中取出  $m$  个元素所有的不同组合的个数，叫做从  $n$  个不同元素中取出  $m$  个元素的一个组合数，符号表示为  $C_n^m$ 。





#### 4. 组合数公式

(1)  $C_n^m = A_n^m / A_m^m = n(n-1)(n-2)(n-3) \cdots (n-m+1) / m!$  其中,  $n, m \in$  正整数,  $n \geq m$ 。

(2)  $C_n^m = n! / m!(n-m)!$  其中,  $C_n^0 = 1$ 。

#### 【例题一】

1. 判断下面问题是组合问题还是排列问题并计算:

(1) 高中篮球比赛, 有 12 支队伍进行单循环比赛, 需要进行多少场比赛?

解: 12 支队伍进行单循环比赛, 没有顺序, 为组合问题。

需要进行  $C_{12}^2 = 66$  (场)。

(2) 10 个人聚会, 见面后每两个人之间相互握手, 总共握手多少次。

解: 每两个人之间握手, 没有顺序, 为组合问题。

所以总共握手:  $C_{10}^2 = 45$  (次)。

### 三、数学期望

#### 1. 数学期望的来源

在 17 世纪中叶, 一位赌徒向法国数学家帕斯卡提出一个使他苦恼很久的分赌本的问题: 甲、乙两赌徒赌技不相上下, 各出 50 法郎, 他们约定谁先赢三局则获得全部赌资。当甲赢了两局, 乙赢了一局时, 因故要中止赌博, 现在该如何分这 100 法郎。

①甲得 2/3, 乙得 1/3。这是基于已赌的局数甲赢了两局乙赢了一局;

②帕斯卡的分法提出, 设想甲乙继续在赌博下去, 则甲所获  $X$  是一个随机变量。  $X$  的取值为 0 或者 100。而在进行两局必可结束, 结果有以下四种结果:

(甲甲)、(乙乙)、(甲乙)、(乙甲)

只有当乙胜两局时甲才分 0 法郎, 那么甲获得 100 法郎的概率为 0.75, 获得 0 法郎的概率为 0.25。即  $X$  的分布列为图 4-2 所示。

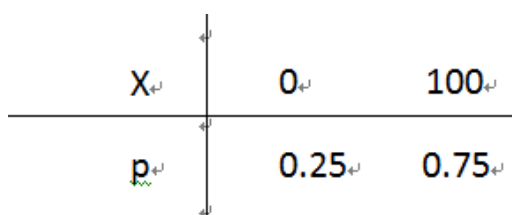


图 4-2  $X$  的分布列



那么甲获得的期望应得  $0 \times 0.25 + 100 \times 0.75 = 75$ 。

## 2. 数学期望的定义

数学期望又称为均值，用来描述一个随机变量的分布的中心位置。

(1) 离散型随机变量  $X$  的数学期望：

$$E(X) = \mu = \sum_i x_i p_i$$

(2) 连续型随机变量  $X$  的数学期望：

$$E(x) = \mu = \int_{-\infty}^{\infty} g(x) p(x) dx$$

(3) 数学期望主要的数学性质：

①若  $k$  是一常数，则：

$$E(kX) = kE(X) ;$$

②对于任意两个随机变量  $X$ 、 $Y$ ，有：

$$E(X+Y) = E(X) + E(Y) ;$$

③若两个随机变量  $X$ 、 $Y$  相互独立，则：

$$E(XY) = E(X) E(Y) ;$$

## 四、方差与标准差

### 1. 随机变量的方差

方差是它的各个可能取值偏离其均值的差平方的均值，记为  $D(x)$  或  $\sigma^2$ ；

公式： $D(X) = \sigma^2 = E(X - \mu)^2$ ；

离散型随机变量的方差： $D(X) = \sigma^2 = \sum (x_i - \mu)^2 p_i$ ；

连续型随机变量的方差： $D(x) = \sigma^2 = \int_{-\infty}^{\infty} [x - \mu]^2 f(x) dx$

### 2. 方差的主要性质

①若  $k$  是一常数，则： $D(K) = 0$ ； $D(KX) = K^2 (DX)$ ；

②若两个随机变量  $X$ 、 $Y$  相互独立，则

$$D(X+Y) = D(X) + D(Y) ;$$

### 3 随机变量的标准差

标准差是它的各个可能取值偏离其均值的差平方的均值的算术平方根即

标准差 = 方差的算术平方根



#### 4 标准差的性质

方差和标准差都反映随机变量取值的分散程度。它们的值越大，说明离散程度越大，其概率分布曲线越扁平。



### 第三节 常用离散型分布

每个随机变量都有一个分布，不同的随机变量可以有不同的分布，也可以有相同的分布，常用的分布分为两类：一类是离散型分布，另一类是连续型分布。本节讲解离散型分布，主要包含了二项分布、泊松分布、超几何分布，主要讲解二项分布，其余分布只作介绍。

#### 一、常用离散分布之二项分布

有时候一个事物，只有两种可能的状态或结果，例如：一件产品的检验，要么合格要么不合格，一个待发射的卫星，要么发射成功，要么发射不成功，虫子吃了农药要么死去，要么活着。等等，要么成功要么失败。此时我们都可以用二项分布来分析这类问题。

二项分布公式为： $P(X=k) = C_n^k p^k (1-p)^{n-k}$ ， $k=0, 1, 2, 3 \dots n$ ，记为  $X \sim b(n, p)$ 。

**【例题一】**以产品检验为例，虽然结果只有合格与不合格两种情况，但抽到不合格品(或合格)的概率显然决定于该批产品的固有不合格率。

如果我们用  $p$  和  $q$  分别代表不合格率和合格 率，则  $p+q=1$ ， $(p+q)^n$  的展开式也应为 1；从而得出：

第  $x=r$  项即  $n$  个产品中出现  $r$  个不合格的概率为

$$C_n^r p^r q^{n-r} \text{ 或 } C_n^r p^r (1-p)^{n-r}$$

写成一般形式为：

$$C_n^x p^x (1-p)^{n-x}, (x=1, 2, \dots, n)$$

二项分布的均值和标准差公式为：

$$\mu = np$$

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

二项分布图形的特点：二项分布属离散型分布，其图形由横坐标上孤立点的垂直线表示如下图 4-3 所示。

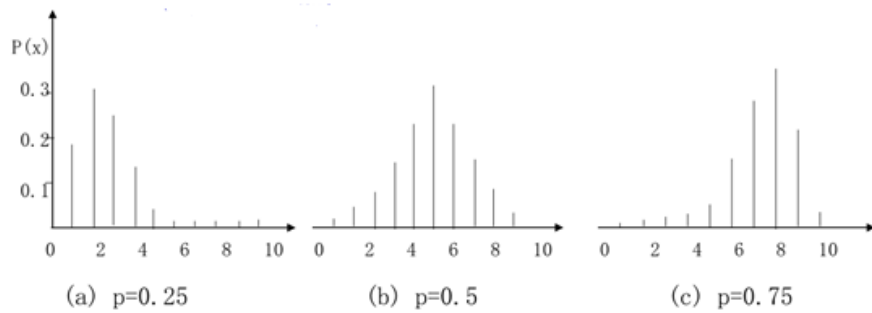


图 4-3 二项分布变化图

由图形分析可知：

- (a) 当  $p=0.25$  时图形是偏向于左边；
- (b) 当  $p=0.5$  时图形是两边对称；
- (c) 当  $p=0.75$  时图形是偏向于右边。

### 【例题二】

有一批不合格率为  $p=0.03$  的产品，按照规定白班与夜班，抽取 5 件并且被抽的 5 件产品不允许有不合格品。试计算产品被接收的概率。

解：

由公式，得：

$$P(x \leq c) = \sum_{x=0}^c C_n^x p^x (1-p)^{n-x}, \text{ 根据所给条件 } x=0, \text{ 代入得:}$$

$$P(x=0) = 0.859;$$

所以产品被接收的概率为 85.9%。

## 二、常用离散分布之泊松分布

泊松分布是由法国数学家泊松在 1837 年首次提出的，公式为：

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots, \text{ 记为: } X \sim P(\lambda);$$

泊松分布也是一种离散型分布，在单位时间上的计数应用较多，例如：

- ① 在一天内，来到地铁站的乘客；
- ② 1 平方米，玻璃上的气泡数；
- ③ 一个铸件上的砂眼数；



### 三、常用离散分布之超几何分布

超几何分布的定义：超几何分布在从一个有限总体中进行不放回抽样会运用到，超几何分布是一种离散型的分布。其公式为：

$$P(X=k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, K=0, 1, 2 \dots r$$

超几何是一种常用的离散分布，主要应用在一些不需要放回样本的抽样中运用到，它在抽样理论中占有重要地位。

## 第四节 常用连续分布

在连续分布场合密度函数与分布函数是等价的，含有相同的信息，各有各的用处，但在图形上，密度函数对各种函数的特征能得到直观显示，如正态与偏态、单峰与平顶都是依据密度函数命名的，因而人们对密度函数更为注意。本节主讲正态分布。

### 一、正态分布背景

正态分布是概率论与数理统计中最重要的一个分布，是由著名数学家，高斯（Gauss，1777-1855）在那个时候通过研究误差理论时，运用整体分布来进行刻画误差的分布。所以正态分布又称之为高斯分布。

正态分布是最常见的以及应用最广泛的一种分布，主要广泛应用于轴径的加工尺寸、化工产品的化学成分、一些电子产品的电参数、测量误差，以及一些自然现象等。

### 正态分布

高斯(德国)



### 二、正态分布的密度函数和分布函数

#### ①密度函数

若随机变量  $X$  的密度函数为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < +\infty,$$

其中， $\mu$  -- 总体均值；

$\sigma$  -- 总体标准差；

则称  $X$  服从正态分布，简单记法： $X \sim N(\mu, \sigma^2)$  其中，如下图  $P(X)$  是一条钟形的曲线，中间高两边低，左右关于  $\mu$  对称， $\mu$  是正态分布的中心，且在  $x = \mu$  附近取值的可能性大，在两侧取值的可能性小，如下图 4-4 所示。

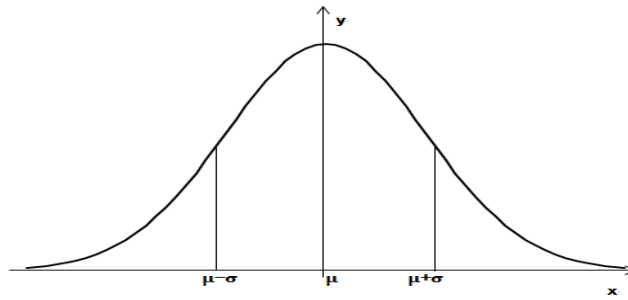


图 4-4 密度函数

图 4-4 为草帽状的正态分布密度函数图形，其特点是在图形的最高点所对应的横坐标值  $\mu$  是正态分布的均值，由图形分析可知当  $X$  的取值越接近  $\mu$ ，它的概率就越大，反之越小。

如图 4-5 给出的  $\sigma$  的变化对正态密度曲线的变化情况。

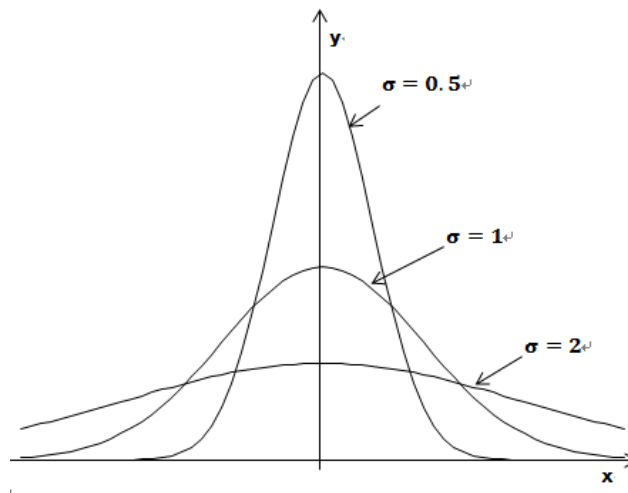


图 4-5 正态密度曲线

由图形可知，当  $\mu$  不变密度曲线由高瘦转变为矮胖，分布情况由集中转变为分散。也就是说正态密度函数的尺度由  $\sigma$  参数来确定，因此也称  $\sigma$  为尺度参数。



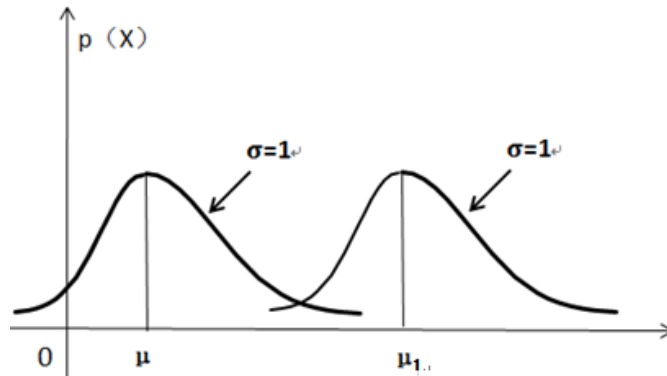


图 4-6 密度函数图形

由图形 4-6 可知，当 $\sigma$ 的位置不变时，改变 $\mu$ 取值，可以看出，图形沿着 X 轴进行平移，而曲线的形状不发生变化。那么，正态密度函数的位置是与 $\mu$ 密切相关的，即也称 $\mu$ 为位置参数。

## ② 正态分布的分布函数

正态分布的分布函数

$$P(a < X < b) = \int_a^b \varphi_{\mu, \sigma}(x) dx$$

如函数图形如下图 4-7 所示

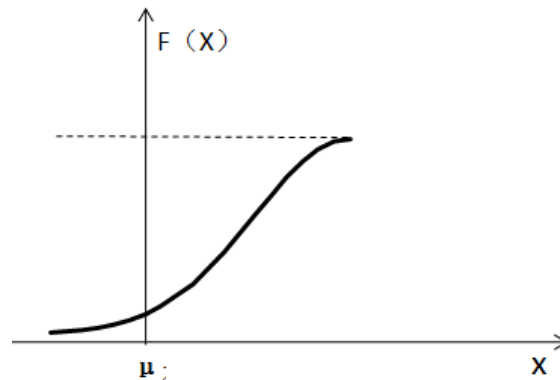


图 4-7 正态函数分布函数图

由图形可以看出它是一条光滑上升的 S 型曲线。

## 三、标准正态分布

### 1、正态分布的累积概率

累积正态分布的累积概率是类似于累积频率一样的，是从最小的频率一直累加到一

个特定的值或位置的直方频率。那么在累积正态分布中累积函数的概率也是指从概率最小的位置（理论为 $-\infty$ ）一直累加到一个特定的值或位置的概率。

累积概率的计算实质上是计算 $-\infty$ 到某一特定点（正态分位点）的正态曲线的所包含的面积。那么累积概率的图形如下图 4-8 中阴影部分。

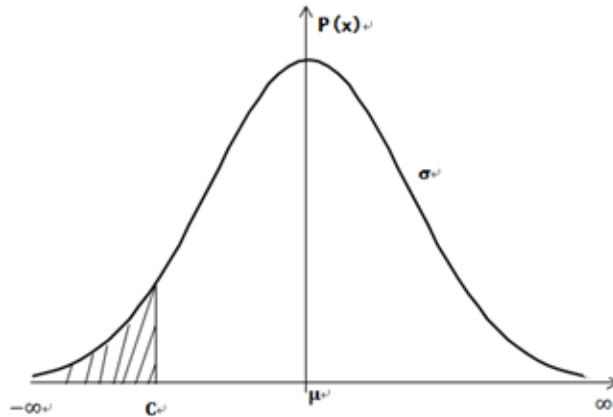


图 4-8 累积正态分布

累积正态分布函数表达式为：

$$\Phi(C) = \int_{-\infty}^C \phi(X) dx$$

## 2. 标准正态分布

我们将当 $\mu=0, \sigma=1$ 时的正态分布称之为标准正态分布。标准正态分布记为 $N(0, 1)$ 。

标准正态分布的密度函数记为：

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, -\infty < x < +\infty,$$

当 $\mu \neq 0, \sigma \neq 1$ 时的正态函数转化为标准正态函数。则令 $Z = \frac{X - \mu}{\sigma}$ ，通过公式 $Z = \frac{X - \mu}{\sigma}$ 将一般正态分布转换为标准正态分布如下图 4-9 所示。

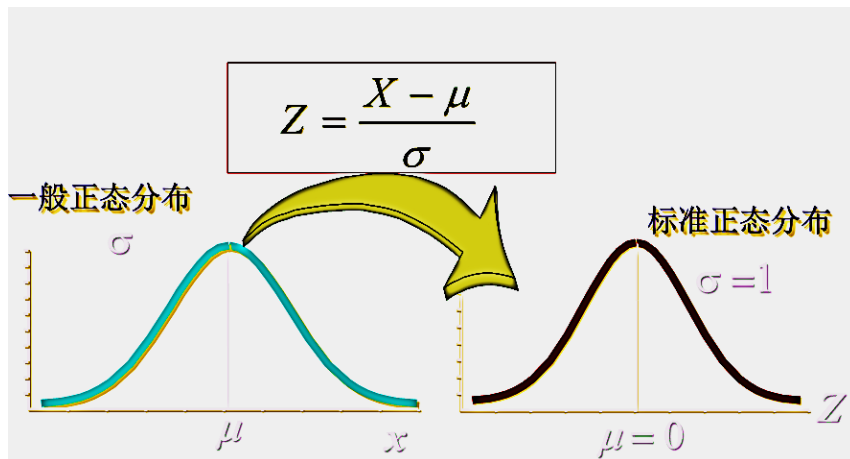


图 4-9 正态分布转换为标准正态分布图步骤

由图 4-9 可知, 根据公式  $Z = \frac{X - \mu}{\sigma}$ , 即将正态分布的每一个值减去  $\mu$  并缩小  $\sigma$  倍。

### 3. 标准正态分布函数表

下表中的正态分布函数表是以标准正态分布设计的。为便于排版, 0.9<sup>2</sup>8650、0.9<sup>3</sup>0324 等写法, 分别表示小数点后有 2 个 9 或 3 个 9, 即 0.998650 和 0.9990324。

### 4. 正态分布的数学期望与方差

设随机变量  $X \sim N(\mu, \sigma^2)$ , 由于  $U = \frac{X - \mu}{\sigma} \sim N(0, 1)$ , 所以  $U$  的数学期望:

$$E(U) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u e^{-\frac{u^2}{2}} du;$$

判断该函数为奇函数, 所以其积分值等于 0, 即:

$$E(U) = 0;$$

又因为  $X = \mu + \sigma U$ , 所以正态分布数学期望:

$$E(X) = \mu + \sigma \times 0 = \mu;$$

也就是说正态分布  $X \sim N(\mu, \sigma^2)$  中的  $\mu$  是数学期望。

$$\text{Var}(U) = E(U^2) - (E(U))^2 = E(U^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \times \sqrt{2\pi} = 1;$$

且  $X = \mu + \sigma U$ , 由方差的性质可以得:  $\text{Var}(X) = \text{Var}(\mu + \sigma U) = \sigma^2$  同理可以得到参数  $\sigma^2$  是  $X$  的方差。

$X$  的标准化随机变量  $X^*$ :  $X = \mu + \sigma U$  公式进行换算可得  $X^* = \frac{X - \mu}{\sigma}$ , 同时可得:

$$E(X^*) = 0, \text{Var}(X^*) = 1;$$

### 5. 样本均值的分布

统计推断是根据样本统计量( $\bar{X}$ 、 $s$  等), 去对总体作出判定, 若已知总体的概率分布, 则通常可以确定由所抽取的样本统计量的概率分布。统计量的概率分布称抽样分布。

可以证明, 不论总体分布如何, 样本均值 可以证明, 即,  $N(\mu, \frac{\sigma}{\sqrt{n}})$  均值不变, 而标准差为  $\sigma_x = \frac{\sigma}{\sqrt{n}}$ 。

### 6. 正态性检验的方法

有的直方图虽然很像正态分布, 但不一定是稳定的, 利用正态分布解决问题(常直接使用正态分布表), 也不一定要画直方图。通常需要直接验证数据是否服从正态分布, 即所谓的正态性检验问题。数据的正态性检验有多种方法, 如  $\chi^2$  检验法、偏态、峰态检验法等。

### 7. 3 $\sigma$ 原则

#### (1) 3 $\sigma$ 的背景介绍

3 $\sigma$ 原则是在正态变量的取值范围在  $(-\infty, +\infty)$ , 它的 99.73% 的值都落在  $(\mu - 3\sigma, \mu + 3\sigma)$  内, 被称为正态分布的 3 $\sigma$  原则。如下图 4-10 所示。

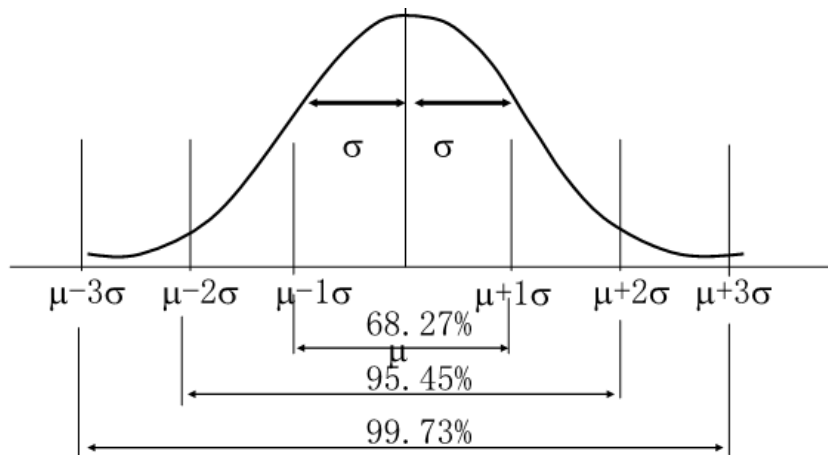


图 4-10 3 $\sigma$  原则

其中,

当  $\mu \pm 1\sigma$  界限范围内的概率为 68.27%;

当  $\mu \pm 2\sigma$  界限范围内的概率为 95.45%;

当  $\mu \pm 3\sigma$  界限范围内的概率为 99.73%。

其计算公式为：设随机变量  $X \sim N(\mu, \sigma^2)$ ，则

$$P(|X - \mu| < k\sigma) = 2\Phi(k) - 1$$

当  $k=1$  时， $2\Phi(k) - 1 = 0.6826$ ；

当  $k=2$  时， $2\Phi(k) - 1 = 0.9545$ ；

当  $k=3$  时， $2\Phi(k) - 1 = 0.9973$ 。

正态分布的  $3\sigma$  原则作用非常大，比如，在一些工业生产上用到的控制图，和一些产品质量指数（过程能力指数  $C_p$ ,  $C_{pk}$ ）。

**【例题一】**

设  $x$  服从  $N(80, 4)$  的正态分布，试求：

(1)  $x \leq 75$  的概率；

解：①因为  $x$  服从正态分布，首先画出正态分布图，根据  $Z = \frac{X - \mu}{\sigma}$  转化为标准正态分布，由图 4-11 可以理解左边画斜线的区域就是  $x \leq 75$  的概率。

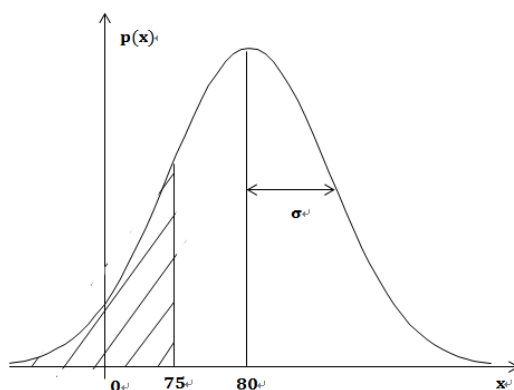


图 4-11

②根据公式  $Z = \frac{X - \mu}{\sigma} = \frac{75 - 80}{4} = -1.25$ ，根据标准正态分布函数表可查得当  $Z = -1.25$  时， $\Phi(-1.25) = 0.1056$ 。所以  $x \leq 75$  的概率为 10.56%。

除了查标准正态分布函数表，还有一种快捷的方式，在 Excel 中计算  $\Phi(-1.25)$  的具体操作步骤如下图 4-12 所示。

第一步，在表格中输入函数公式：NORMSDIST(-1.25)

第二步，输入公式后回车键确认，此时表格中计算出的标准正态分布数值为 NORMSDIST(-1.25) = 0.1056。

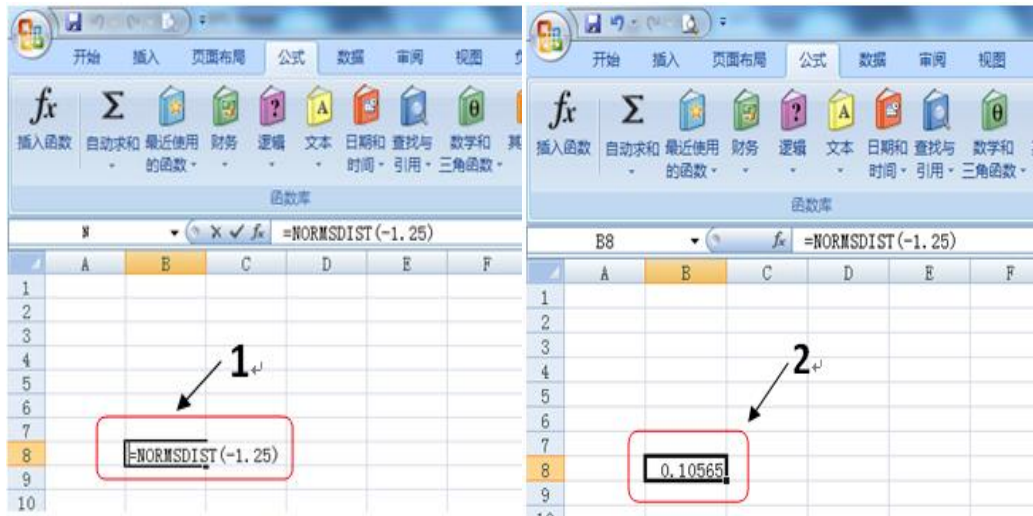


图 4-12

(2)  $x \leq 89$  的概率

解: ①因为  $x$  服从正态分布, 首先画出正态分布图如下图 4-13 所示, 根据  $Z = \frac{X - \mu}{\sigma}$  转化为标准正态分布, 由图 4-13 可以理解画斜线的区域就是  $x \leq 89$  的概率。

②根据公式  $Z = \frac{X - \mu}{\sigma} = \frac{89 - 80}{4} = 2.25$ , 根据标准正态分布函数表可查得当  $Z = 2.25$  时,  $\Phi(2.25) = 0.9878$ 。所以  $x \leq 89$  的概率为 98.78%。

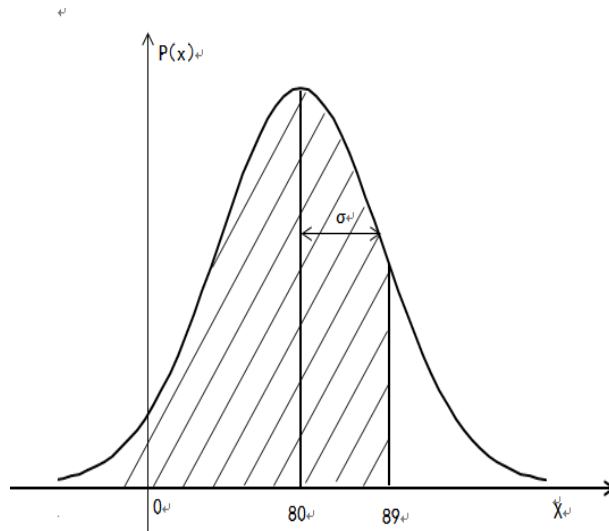


图 4-13

由 (1) 式方法在 Excel 中计算 (2.25), 结果如下图 4-14 所示。

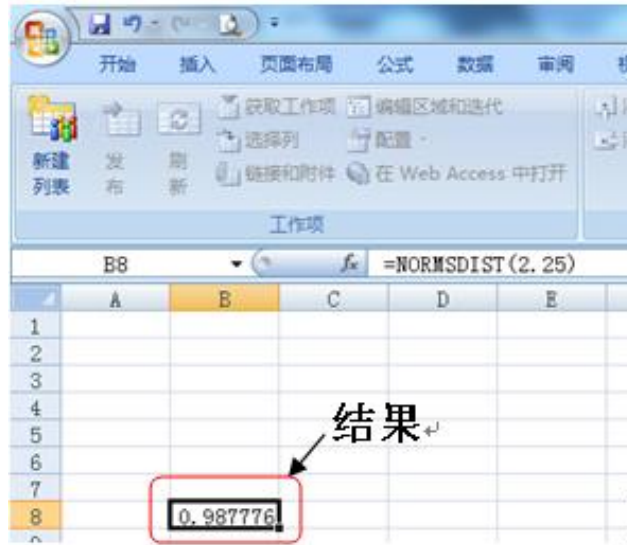


图 4-14

(3)  $75 \leq x \leq 89$  的概率

解：①因为  $x$  服从正态分布，首先画出正态分布图如下 4-15 所示，根据  $Z = \frac{X - \mu}{\sigma}$  转化为标准正态分布，由图可以理解画斜线的区域就是  $75 \leq x \leq 89$  的概率。

②由图形可以理解  $P(75 \leq x \leq 89) = (2.25) - (-1.25) = 0.8822$ 。所以  $P(75 \leq x \leq 89)$  的概率为 88.22%。

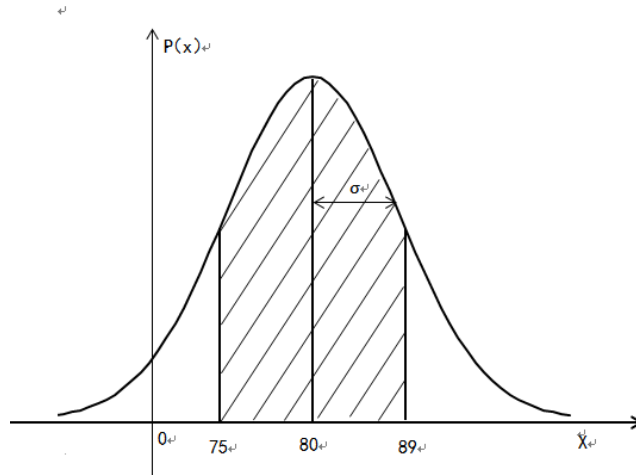


图 4-15

(4)  $x \geq 89$  的概率

解：①同理，画出正态分布图 4-16 所示，根据  $Z = \frac{X - \mu}{\sigma}$  转化为标准正态分布，由图可知画斜线部分就是  $x \geq 89$  的概率，如下图。

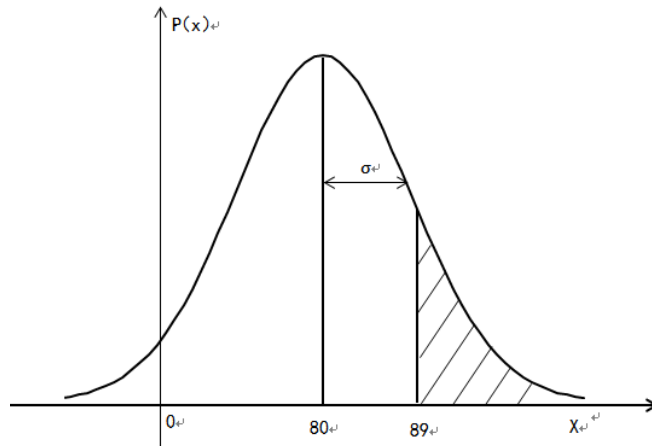


图 4-16

②计算得  $p(x) = 0.122$ ，即大于 89 的概率为 12.2%

#### 四、中心极限定理

这个定理表明：无论随机变量服从何种分布，可能是离散分布，也可能是连续分布，连续分布可能是正态分布，也可能是非正态分布，只要独立同分布随机变量的个数  $n$  较大，那么，随机变量之和的分布，随机变量均值的分布都可以近似为正态分布。

统计学中把子组样本平均值的标准差称为均值的标准差，记为  $\sigma_{\bar{x}}$  或者 SEM。无论是正态还是非正态，子组均值的标准差都有：

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

SEM 伴随着  $n$  的增加而减少，如图 4-17 表示，当  $n$  小于 10 的值时， $\sigma_{\bar{x}}$  会迅速减小，当  $n$  大于 10 时  $\sigma_{\bar{x}}$  减小缓慢。

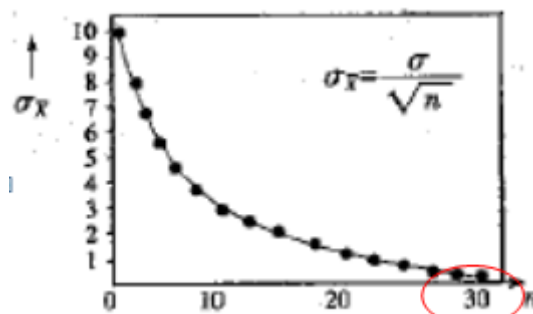


图 4-17  $\sigma_{\bar{x}}$  与  $n$  的关系







## 第五节 复习题

### 一、单选题

1. 若  $n \in N^*$ , 且  $n < 10$ , 则  $(10 - n)(11 - n) \cdots (100 - n)$  等于 ( )  
A.  $A_{100-n}^{10-n}$       B.  $A_{100-n}^{90}$       C.  $A_{100-n}^{91}$       D.  $A_{100-n}^{91}$
2. 由数字 1, 2, 3, 4, 5 组成没有重复数字的五位数, 其中小于 50000 的偶数共有 ( )  
A. 24 个      B. 36 个      C. 48 个      D. 60 个
3. 3 名医生和 6 名护士分配到 3 所学校为学生体检, 每校分配 1 名医生和 2 名护士, 则不同的分配方法共有 ( )  
A. 90      B. 180      C. 270      D. 540
4. 某仪表显示屏上一排有 7 个小孔, 每孔可显示出 0 或 1, 若每次显示其中的 3 个孔, 但相邻的两孔不能同时显示, 则这个显示屏可以显示的不同信号的数目 ( )  
A. 10      B. 48      C. 60      D. 80

### 二、多选题

1. 在密度函数中, 下面符号所代表的含义正确的是 ( )  
A.  $\mu$  是总体均值    B.  $\mu$  是标准差    C.  $\sigma$  是总体标准差    D.  $\sigma$  是总体均值
2. 在正态分布的  $3\sigma$  原则中下面选项正确的是 ( )  
A. 当  $\mu \pm 1\sigma$  界限范围内的概率为 68.27%  
B. 当  $\mu \pm 2\sigma$  界限范围内的概率为 95.45%  
C. 当  $\mu \pm 3\sigma$  界限范围内的概率为 99.73%

### 三、简述题

1. 用数字 0, 1, 2, 3, 4, 5 可以组成没有重复数字, 并且比 20000 大的五位数共有多少个
2. 12 名同学合影, 站成前排 4 人后排 8 人, 现摄影师要从后面 8 人抽两人调到前排, 若其他人的相对顺序不变, 则不同的调整方法有多少种?
3. 某地区 18 岁女青年的血压  $X$  (收缩压) 服从  $N(110, \cdot)$ , 试求该地区 18 岁女青年的血压在 100 至 120 的可能性有多大?
4. 8 个人坐在前后两排, 每排 4 人, 问:  
(1) 有多少种不同的做法?



(2) 如果某 2 人一定要坐在前排，有多少种做法？

(3) 如果某一人必须做在前排，某 2 人必须坐在后排，有多少种做法？

5. 7 个人到 7 个地方旅游，甲不去 A 地，乙不去 B 地，丙不去 C 地，丁不去 D 地，  
则有不同旅游方案多少种？



## 第五章 过程质量控制

### [教学目标]

- (1) 掌握过程控制的概念
- (2) 掌握过程能力指数的计算
- (3) 掌握了解流程控制的影响因素

### [教学难点]

- (1) 掌握过程能力指数的计算
- (2) 了解流程控制的因素



## 第一节 过程质量

### 一、影响过程质量的因素

生产过程中，产品的质量偶尔会出现不良，对于这些不良到底是由于什么原因造成的，是一个工程人员应该了解的问题，是工程工作者的基础知识和必备功课。通常来说对过程质量起主要作用的因素是5M1E，即操作者、材料、设备、工艺方法和环境、测量等6大主因素，这些因素是造成产品不良的可能原因。

操作者：操作者的情绪、技术水平、质量意识

材料：主要指构成产品的原材料和生产中使用的辅助材料

设备：机床、辅助装置、工装夹具或模具的精度和状况

工艺方法：加工工艺流程和操作方法是否合理

环境：合适的温度和光线、整齐、清洁的场所

测量：测量的设备、试验手段、测量方法

在现实生产中，面对突发的设备故障，首先应该通过生产过程进行分析，运行专业的工程能力技术和相关的或相类似的经验，找出影响的因素，找出其中的主要因素，将起主要作用的因素作为质量控制的重点。其他因素也要时常关注，减少或消除其他因素的影响。

### 二、过程质量的波动

在现实生产的产品中，单单从统计学的角度进行讨论和分析，可以把设备生产的产品质量的波动分为正常波动和异常波动两种。

正常波动：在生产过程中大量存在由随机原因引起的较小的产品质量波动，这些随机因素对产品质量发生细微影响，但它造成的质量特性波动往往较小，对产品的合格并不造成影响，可以忽略不计，这种波动称为产品的正常波动。一般情况下，这些质量波动在生产过程中是允许存在的，正如世界上没有相同的两片树叶，任何产品都是不尽相同的，有波动是正常的，统计学参数公差就是承认这一种波动的产物。

异常波动：由生产过程中未及时发现的系统性原因引起的产品质量波动叫异常波动，这些影响因素在正常生产过程中不常见，只是在某时刻特殊条件下影响因素发生改变而

较大质量波动显现或爆发出来，由原先的正常波动变为异常波动，出现较多产品不良，对产品质量的影响是比较显著的。

### 三、过程质量的状态

在生产中，过程质量的状态的了解是很重要的，当处于统计控制状态时，生产的条件未发生变异，生产无异常。当过程质量的状态处于非统计过程状态时，那么生产有异常，我们就要从生产的各个因素去分析哪里出现了变异导致控制状态不受控制了，我们需要根据控制图与生产因素分析给出改善措施，让生产顺利进行。

#### 1. 过程控制中的普通变异和特殊变异

(1) 普通变异：在生产过程中，在我们所能够控制的范围内，流程处于稳定、可控和预测的状态。

(2) 特殊原因：当流程处于特殊原因时，会造成制程的变异，从而导致流控制状态。普通变异与特殊变异与控制图的关系如下图 5-1 所示。

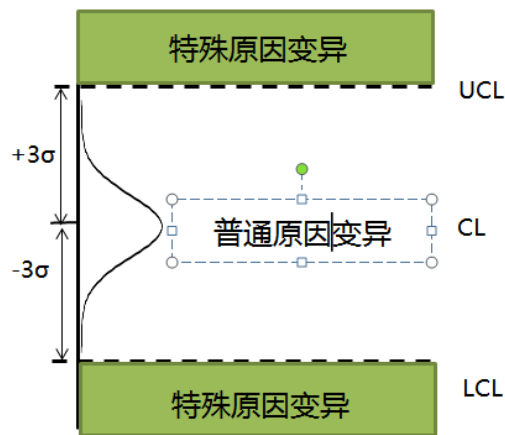


图 5-1

#### 2. 过程控制图原理

通过控制图的界限来判定生产中的过程质量是否异常，这种质量异常误判错误有两种，第一种是错误将正常判为异常，第二种错误是将异常判为正常。

在应用控制图的过程中，犯错误是避免不了的。事实上，在控制图上所能变动的不外乎是上下控制界限的间距。但是要改动上下限间距会发生两种情况，一种是拉大间距，造成犯第二种错误的可能性增大，第一种错误的可能性减小，第二种是缩小间距，造成

犯第一种错误的可能性增大，第二种错误的可能性减小。这两者是矛盾的。因此只能根据第一种错误和第二种错误所造成的总损失为最小者为准则来确定上下控制界限。

根据经验证明，当  $UCL = \mu + 3\sigma$ ， $LCL = \mu - 3\sigma$  时的所谓  $3\sigma$  方式就是两种错误所造成的总损失最小的界限。见下图 5-2 所示。

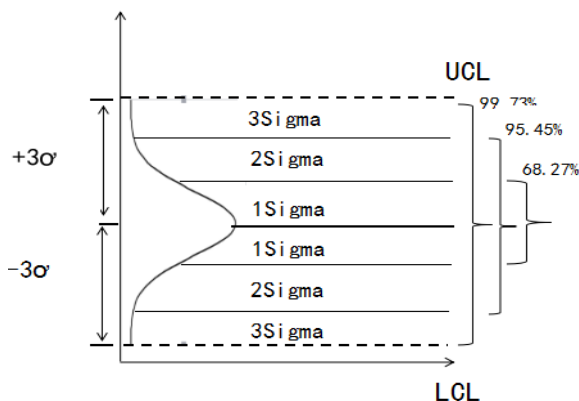


图 5-2 控制图原理

### 3. 控制图的判异准则

统计学中，按照国标 GB/T 4091-2001，存在 8 种判异现象的概率大体都等于或接近 0.27%。

在 8 种判异准则中，前四项其实与分区无关，后面四项法则都是在正态分布的条件下，将  $\pm 3\sigma$  区域分成 6 个子分区，按照正态分布在各区中应该出现的概率来制定的法则。故而，后四项判断法则只对单值  $X$  及小组均值的控制图使用，其他控制图皆只使用前四项法则。所以这里我们只介绍前面四种判异准则，后四项法则不做介绍，剩余判异规则可在附表 2 中查看。

(1) 检验一：1 个点落在控制限之外，如图 5-3。

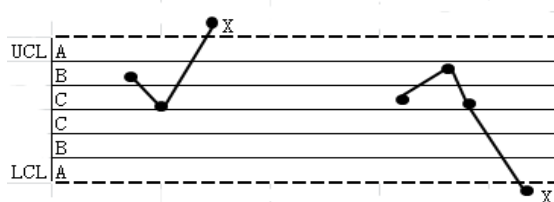


图 5-3 检验一

根据控制图原理可知数据点落在控制限之内的概率为 99.73%，落在控制限之外的概率为 0.27%。所以检验一：1 个点落在控制限之外的概率等于 0.27%。

(2) 检验二：连续 9 点落在中心线同一侧，如图 5-4。

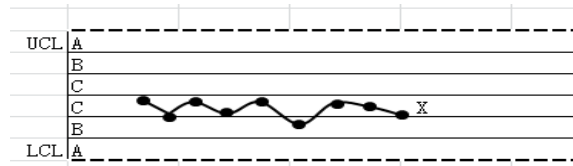


图 5-4 检验二

概率计算为： $P=0.5^9=0.195\%$ 。

(3) 检验三：连续 6 点递增或递减。如图 5-5。

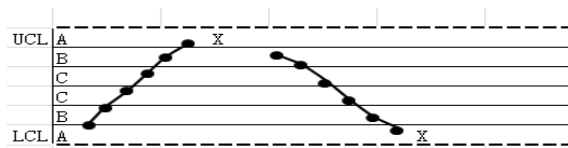


图 5-5 检验三

概率计算为： $P=2/A_6^6=0.278\%$ 。

(4) 检验四：连续 14 点相邻点上下交错。

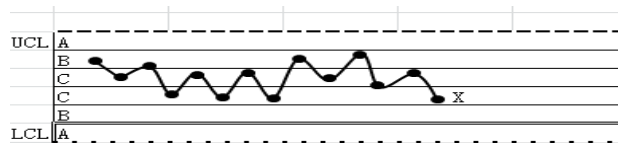


图 5-6 检验四

概率计算为： $P=(1/A_7^7)*14=0.278\%$ 。

#### 4. 判断控制状态

根据前面介绍的控制图原理以及八大判定准则，可以进行分析过程控制质量处于什么样的状况了，过程控制质量状态分为 2 种分别是控制状态、失控状态。而失控状态又分为：稳定失控状态和不稳定失控状态。

##### (1) 控制状态

由图形 5-7 所示的生产过程控制状态图，其中  $\mu_0$  和  $\sigma_0$  是在调整过后控制的状态图，调整过后符合质量的要求。在图中可以发现，图形随着时间的增多，逐渐平稳在上下控制界限内并且均匀分布，这就是生产过程控制的目的，也就是生产控制稳定的状态。此时，数据线在  $\pm 3\sigma$  控制限之内，并且与上述中提到的 8 大判异准则无相关。所以此时的生产状态是处于控制状态，无异常。



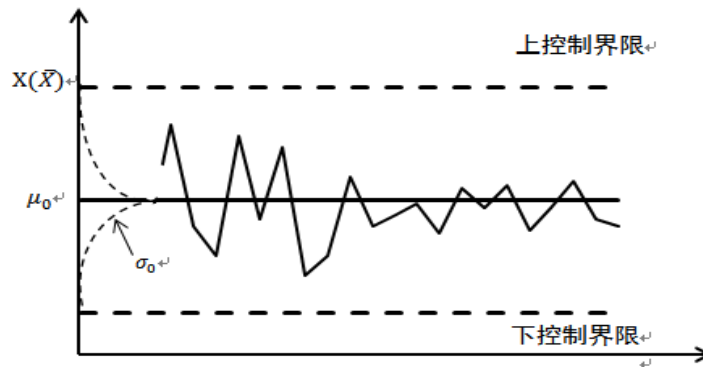


图 5-7 生产过程控制状态图

失控状态：生产过程中  $\mu$  和  $\sigma$  两者之一或都不符合质量规格要求，或随时间变化而出现不稳定状态，就认为过程处于失控状态，失控状态分为稳定失控状态和不稳定失控状态。

稳定失控：总体均值  $\mu$  和总体标准差  $\sigma$  的值不随时间变化，但部分质量特征值的分布接近或超出了上下控制界限。下图表示的是样本均值  $\mu$  大于规定总体均值，加工的产品质量特征数据在总体均值上方，并不在总体均值附近呈现正态分布。这种情况可能是设备参数调整不当造成的，应及时采取措施将产品的均值  $\mu$  调整回原本的  $\mu_0$  分布中心位置上来，减少出现不良的几率，做到稳定地生产。

稳定失控状态图如下图 5-8 所示：

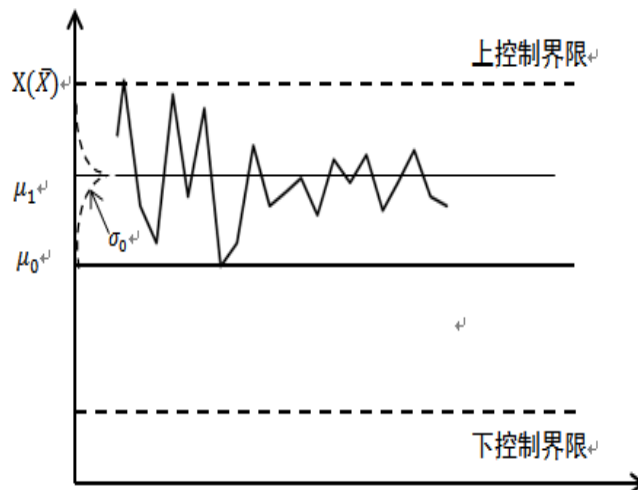


图 5-8 稳定失控状态图

此时，由图形可以看出，数据线已经快超出  $\pm 3\sigma$  范围了，那么也就是说生产的产品

有可能会出现在控制图中 0.27% 的不良品，这种小概率事件要发生，但是还是在可以控制的状态下，故应该要注意生产过程中是否有变异。

不稳定失控：产品的质量特征值超出了上下控制界限，样本均值  $\mu$  和样本标准差  $\sigma$  的值随时间变化而变化，偏离了规定的标准值，产生较多的不良造成较大影响，

下图表示的是样本均值  $\mu$  变大的情况，说明生产过程中有某一因素或某几个因素产生了变化，而这因素变化对设备生产情况产生极大影响，不能正常的运行。

不稳定失控状态图见下图 5-9：

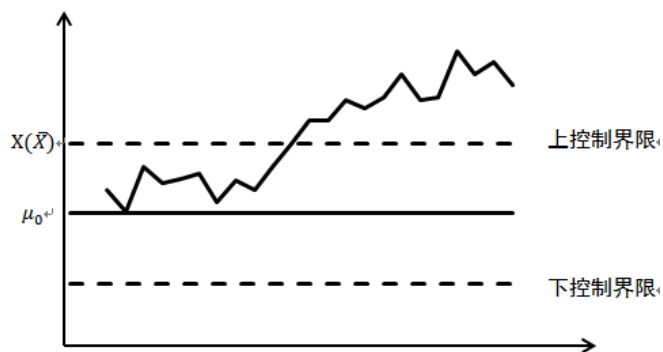


图 5-9 不稳定失控状态示意图

由图形可知，数据点已经超出了  $\pm 3\sigma$  控制界限了，属于 8 大判异现象，生产过程中的某些因素已经发生变异，此时的生产状态是不稳定的失控状态，所以发生产品质量特征值的集体偏离原先设定的标准值，出现失控状态，此时应该马上停止生产，对设备运行情况进行重点排查，查明造成失控的具体原因，对原因进行分析，并及时确认就是此原因造成的失控，提出针对此失控的解决方案并进行验证，消除影响恢复正常运行状态，继续生产产品。

#### 四、过程质量分析

生产活动中通过观察产品制造过程中其特有的质量特征值的上下波动情况，记录原始的数据，分析数据找出影响产品质量的各类因素，从中找出主导性或重要性因素，调查并验证这些因素与生产过程及结果之间的关系，然后在过程分析的基础上发现因素的变异范围，建立过程因素管理标准，这一系列活动就叫做过程分析。



## 五、过程质量控制

在生产过程中，一旦发现过程，尤其是关键过程质量失控，就必须按鱼骨图思路从人、机、料、法、环、测量等方面进行过程分析，进而对影响因素进行控制，使过程回到控制过程。所以在生产中要求操作者记录过程质量的原始数据，随时掌握和分析质量的变化趋势，并进行过程能力指数的计算和评价，确保过程始终处于良好的受控状态，尽可能不出或少出不合格品。

## 六、过程质量分析和控制关系

过程质量需要严密的监控，发现异常情况时，要进行质量分析，然后根据分析的结果调节和控制生产要素，因此，过程质量分析和过程质量控制是相辅相成的，而且分析和控制过程是不间断的、反复进行的。

## 第二节 过程能力及其指数

### 一、过程能力

过程能力：过程能力的称法有好几种，从早期工业开始的工序能力，到现今称为过程能力、工程能力或工艺能力。它主要是指在控制状态下的过程，即在控制状态下衡量过程加工的一致性，也可以理解为，当过程处于统计控制养状态，过程符合容差范围的输出能力。一般用特性值散布的 6 倍标准差 ( $6\sigma$ ) 衡量。

过程能力与质量之间的关系：一般情况下，过程能力和产品的质量实际波动称成反比，即过程能力越高，质量波动就越小。下图为几种不同的过程能力，由此可发现， $\sigma$  是过程能力的一个重要的参数， $\sigma$  越大过程能力就越低， $\sigma$  越小过程能力就越高。如下图 5-10 不同  $\sigma$  状态下的过程能力图。

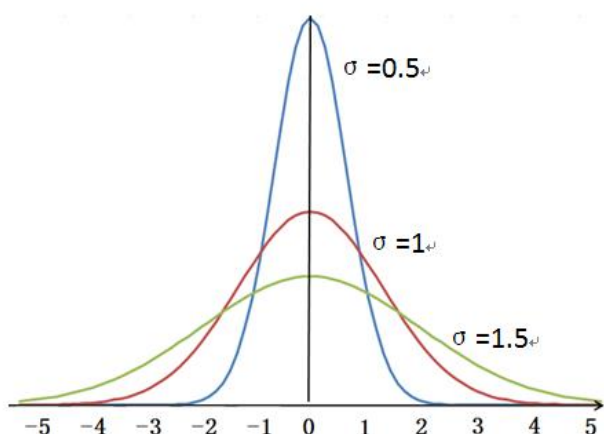


图 5-10 不同的过程能力图

根据过程统计规律，一般采用“ $3\sigma$ ”来描述过程能力的大小。

从公式： $B=2\times 3\sigma$  中，其中  $B$  为过程能力， $\sigma$  为处于正态分布状态下的过程质量特征值的标准差。

这里的  $3\sigma$  是指的是  $3\sigma$  范围内，它的 99.73% 的质量特征值就在  $6\sigma$  这个范围内。也就是说，生产中，符合  $3\sigma$  的产品，那么它即有 99.73% 的合格品。当时  $3\sigma$  的创建者在取这个  $3\sigma$  区间时，考虑到的就是产品的合格率，即，当我取值范围为  $1\sim 2$  时，则包含的概率太低，也就是合格品率太低；如果，取值范围为  $(4\sim 5)$  时，则达到这样高合格率的产品所需要耗费的成本将不划算。所以就有了过程能力定义为  $6\sigma$ 。



注意：过程能力  $B=6\sigma$ ，为什么？它是有条件的：首先，质量特征必须服从正态分布，其次，控制的结果是产品的合格率能够达到 99.73%，所以， $B=6\sigma$  适用于一般的情况，要灵活的去运用，机械的套用  $B=6\sigma$ ，会产生较大的偏差。

## 二、过程能力指数

1. 过程能力指数：过程能力只是对一种过程固有的实际加工能力，而已产品的技术要求无关，为了反映过程能力是否满足顾客的客观计数要求，需要将两者进行比较，从而引入了过程能力指数的概念。通常将允许的容差范围除以  $6\sigma$  的比值，称为过程能力指数。也就是，用来表示过程能力对产品质量计数要求的保证程度，记为  $C_p$ 。

过程能力指数公式： $C_p=T/(2\times 3)\sigma$

式中， $T$ ——公差范围； $\sigma$ ——总体标准差；

### 2. 过程能力指数的计算

过程能力的计算根据情况有不同的计算方式，主要有以下几种情况：

当双向差的中心值  $M$  与数据分布中心值  $\mu$  相一致时，即  $\mu=M$ ，称过程能力“无偏”，用  $C_p$  表示；不一致时，称“有偏”，用  $C_{pk}$  表示。

(1) 双向公差，且过程分布中心  $\mu$  与标准公差中心  $M$  重合的情况

分析当  $\mu=M$  时，称过程能力为“无偏”，用  $C_p$  表示。

公式为： $C_p=T/(2\times 3)\sigma$ ；

其中， $T$  为容差范围： $T=T_u-T_L$ ；

式中， $T_u$ ——公差上限； $T_L$ ——公差下限；

由公式，可推论出当  $T$  过大时值  $C_p$  较大，当数据的散布 ( $\sigma$ ) 越大，数据  $C_p$  的值就越低。

**【例题一】**工艺规范要求，某零件热处理的温度为  $450^\circ\text{C}\pm 5$ ，经长期测试结果，炉温服从  $N(450, 2)$  正态分布，请计算这种条件下的过程能力指数。

解：①因为炉温服从正态分布，首先画出炉温的正态分布图如下图 5-11 所示，由图形分析问题。

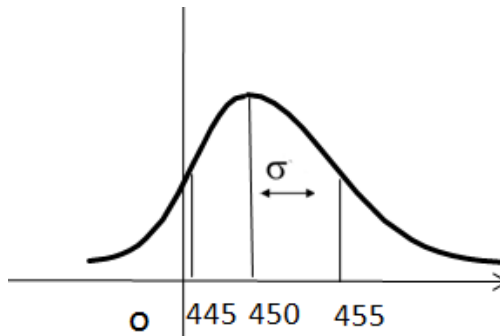


图 5-11 正态分布图

②因为 $\mu=M=450$ ，容差的上下限范围为 $T=T_u-T_L=455-445=10^\circ\text{C}$ 。代入公式 $C_p=T/(2\times 3)\sigma$ ，求得 $C_p=(455-445)/12=0.83$ 。

**【例题二】**在例题一中如果将温度范围控制为 $450^\circ\text{C}$ ，那么 $C_p$ 值有啥变化？

解：此时温度范围控制在 $450^\circ\text{C}$ 那么 $T=T_u-T_L=459-441=18^\circ\text{C}$ ，根据公式 $C_p=T/(2\times 3)\sigma$ 得 $C_p=18/12=1.5$ 。

由上述例题可知如果过程能力( $\sigma$ )不变， $C_p$ 值将决定于公差范围的大小。

(2) 当 $\mu \neq M$ 时，

如果容差中心 $M$ 与数据分布中心不一致，即当 $\mu \neq M$ 时，上式需乘上一个修正系数 $(1-K)$ ，这时 $K=\epsilon/(T/2)=2\epsilon/T$ 。其中 $\epsilon=|M-\mu|$ 。计算公式为：

$$C_{pk} = (1-k) C_p = (T-2\epsilon)/6\sigma$$

**【例题三】**工艺规范要求，某零件热处理的温度为 $450^\circ\text{C} \pm 9$ ，经长期测试结果，炉温服从 $N(452, 2)$ 正态分布，请计算这种条件下的过程能力指数 $C_{pk}$ 的值。

解：①炉温服从 $N(452, 2)$ 正态分布，画出正态分布图 5-12 所示进行分析。

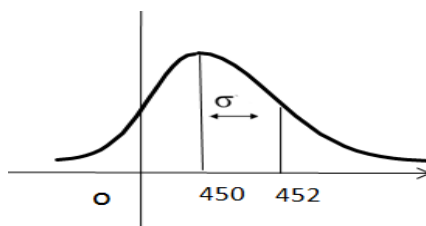


图 5-12 正态分布图

②由于 $M \neq \mu=450$ ，根据公式 $C_{pk} = (1-k) C_p = (T-2\epsilon)/2 \times 3\sigma$ ， $\epsilon=|M-\mu|=452-450=2$ ； $T=T_u-T_L=18^\circ\text{C}$ ，即 $C_{pk}=(18-4)/12=1.667$ ；

显然，由于 $M \neq \mu$ ， $C_{pk}$ 值总是要比 $C_p$ 值小，而数值减小的幅度取决于 $M$ 与 $\mu$ 的差值。

(3) 双边容差情况

①当  $\mu$  与  $M$  重合 ( $\mu=M$ ) 时,  $C_p=T/6\sigma=(T_u-T_L)/6\sigma$ .

如下图 5-13 所示

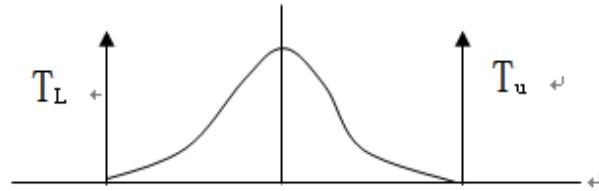


图 5-13  $\mu$  与  $M$  重合 ( $\mu=M$ )

②当  $\mu$  与  $M$  不重合 ( $\mu \neq M$ ) 时

$$C_{pk} = (1-K) C_p = (T-2\varepsilon) / 6\sigma$$

如下图 5-14 所示

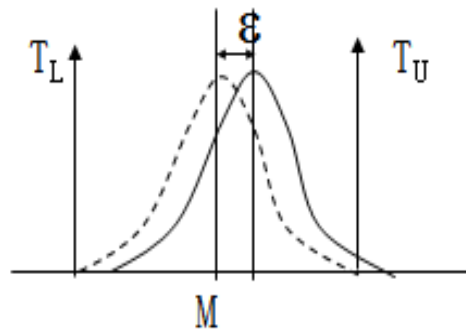


图 5-14  $\mu$  与  $M$  不重合 ( $\mu \neq M$ )

(4) 单边容差的情况

①单侧上限, 如下图 5-15 所示

$$C_p=T/3\sigma=(T_u-T_L)/3\sigma$$

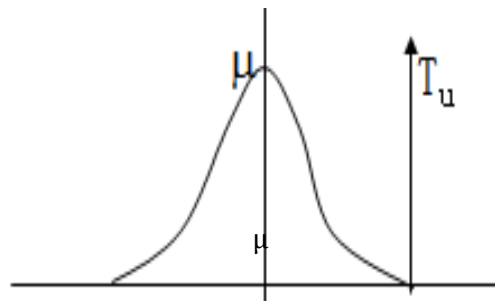


图 5-15 单侧上限图

②单侧下限, 如下图 5-16 所示。

$$C_p = T / 3\sigma = (\mu - T_L) / 3\sigma$$

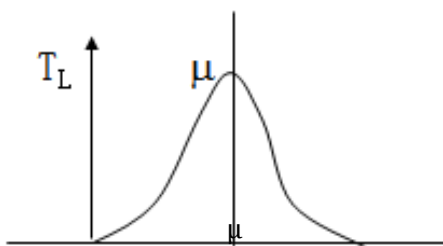


图 5-16 单侧下限图

### 3. 过程能力的分级及 $C_{pk}$ 值的调整

#### (1) 过程能力分级

为便于比较，通常将过程能力指数分成 5 级，如下表 5-1 所示：

表 5-1 过程能力分级表

过程能力指数	分级
$C_p > 1.67$	特级（过高）
$1.33 < C_p \leq 1.67$	一级（充足）
$1 < C_p \leq 1.33$	二级（满足）
$0.67 < C_p \leq 1$	三级（不足）
$C_p < 0.67$	四级（严重不足）

#### (2) 对有偏过程能力的调整

当有偏情况 ( $M \neq \mu$ ) 时，是否需将数据的分布中心调整到与容差中心一致，则取决于下列多种因素：

- ①  $C_p$  值的富裕程度；
- ② 调整的难易程度；
- ③ 对最终产品的影响；
- ④ 调整的经济性；

下表给出的情况可供参考，见下表 5-2。

表 5-2

$C_{pk}$ 值	偏离系数 K	建议的措施
$1.33 < C_{pk}$	$0 < K$	不必调整
$1.33 < C_{pk}$	$0.25 < K < 0.50$	密切关注
$1.00 < C_{pk} \leq 1.33$	$0 < k < 0.25$	密切关注
$1.00 < C_{pk} \leq 1.33$	$0.25 < k < 0.50$	应要调整

#### (3) 过程能力与容差限

标准差反映了在正常生产条件下产品特性的散布，是衡量质量是否稳定的标志，是



特性值散布规律的反映。而容差限是根据实际情况需要人为规定的一个区间。容差限区间定的范围宽，不合格品率就低，定的范围窄，不合格品率就高。

**【例题四】**加工  $\Phi 80$  毫米的轴径。根据统计分析，已知轴径的加工尺寸服从  $N(80, 0.1)$  正态分布。如果将容差范围分别定在  $\Phi 80^{\pm 0.35}$  和  $\Phi 80^{\pm 0.2}$  毫米，试分析两种情况的不合格品率的变化。

解：若容差范围定在  $\Phi 80^{\pm 0.35}$  毫米，则  $T = \pm 0.35 = 0.7$ ，是标准差  $\sigma = 0.1$  的 7 倍，亦即容差的上、下限 ( $T_u, T_l$ ) 定在  $\mu \pm 3\sigma$  之外，根据正态分布的特点可知，轴径加工尺寸在  $\mu \pm 3\sigma$  之外的概率应小于 0.3%，在生产正常时，几乎没有不合格品。

若容差范围定在  $\Phi 80^{\pm 0.2}$  毫米，则容差上、下限恰好定在  $\mu \pm 2\sigma$  处，此时大体上就会出现 5% 的不合格品。这 5% 的不合格品，在正常生产条件下也是不能避免的 (如下图 5-17 所示)。

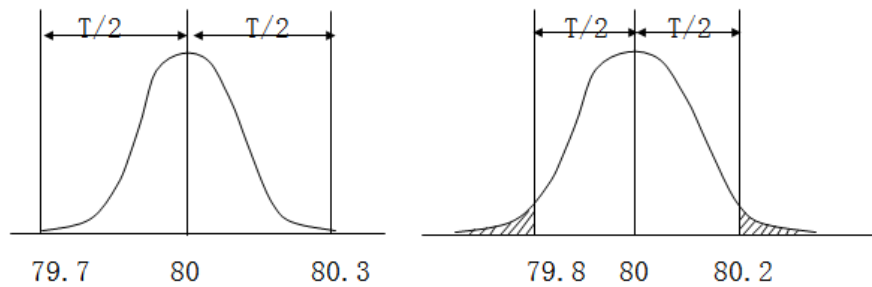


图 5-17  $\sigma$  不变时容差限变化对不合格品率的影响

如图 5-18 在过程稳定状态，即数据分布  $N(\mu, \sigma)$  不变的条件下，不合格品率只与容差限有关，其关系如下表所示：(其中  $L$  是容差的一半， $\sigma$  是总体标准差。

L / $\sigma$	不合格品率 (双边)%	不合格品率 (单边)%
3.29	0.1	0.05
3.09	0.2	0.10
2.81	0.5	0.25
2.58	1.0	0.50
2.33	2.0	1.0
1.96	5.0	2.5
1.64	10.0	5.0
1.44	15.0	7.5
1.28	20.0	10.0
1.15	25.0	12.5
1.04	30	15
0.84	40	20
0.67	50	25
0.52	60	30
0.39	70	35
0.25	80	40

图 5-20 容差限与不合格率的关系图

所谓“双边”不合格品率，就是把大于容差上限和小于容差下限的产品都作为不合格品时，它们占全部产品的百分率。“单边”不合格品率，则是仅把大于容差上限或小于容差下限的产品看作不合格品时，它们占全部产品的百分率。可以看出，上表的双边不合格品率实际上是相当于化简了的《正态分布双侧分位数表》。

【例题五】如果将设计上下限定在距 $\mu \pm 1.96\sigma$ 处，其在正常生产条件下的不合格品率是多少？

解：查  $L/\sigma$  为 1.96 行，得双侧不合格品率为 5%，单侧不合格品率 2.5%。

【例题六】若喷嘴外径的均值为 8 毫米，标准差 $\sigma=0.1$  毫米，如果希望把废品率控制在 2%左右，那么在质量稳定的情况下应把设计上下限定在哪里？

解：从表查出当不合格品率为 2%(双侧)时， $L/\sigma=2.33$ ，

因此， $L=2.33 \times \sigma=2.33 \times 0.1=0.233$  毫米；

于是，容差上限 $\mu+L=8+0.233=8.233$  毫米；

容差下限 $\mu-L=8-0.233=7.767$  毫米；

#### 4. 过程能力指数与不合格品率

正态分布的双侧分位数表，主要用于已知一个小概率 $\alpha$ ，再去查正态分位点(如显著性检验)，但在考虑过程能力(指数)与不合格品率  $P$  的关系时，直接查正态分布表(附表)就更方便。当容差中心  $M$  与数据散布中心 $\mu$ 重合( $M=\mu$ )时，可根据下式计算并查附表得出



不合格品率 P:

$$P=2-[2\Phi(3C_P)]$$

当容差中心 M 与数据散布中心  $\mu$  不重合 ( $M \neq \mu$ ) 时, 不合格品率 P 由过程能力指数  $C_P$  以及 K 值决定, 算法如下:

$$P=2-\{\Phi[3C_P(1+K)]+\Phi[3C_P(1-K)]\}$$

【例题七】用公式求  $C_P=0.8$  时的 P 值 ( $M=\mu$ )

解: 由于是无偏 ( $M=\mu$ ), 可用公式  $P=2-[2\Phi(3C_P)]$ ;

由  $C_P=0.8$ , 得  $3C_P=2.4$ ,

查附表, 得到,

$$\Phi(2.4)=0.991802;$$

故不合格品率为:

$$P=2-2 \times 0.991802=0.016396=1.6396\%。$$

【例题八】在上例中, 如果  $M \neq \mu$ , 且  $K=0.2$  时, 求不合格品率 P。

解: 将有关数值代入公式并查表, 得到,

$$\begin{aligned} P &= 2 - \{\Phi[3C_P(1+K)] + \Phi[3C_P(1-K)]\} \\ &= 2 - \{\Phi[3 \times 0.8(1+0.2)] + \Phi[3 \times 0.8(1-0.2)]\} \\ &= 2 - \{\Phi[2.4 \times 1.2] + \Phi[2.4 \times 0.8]\} \\ &= 2 - \{\Phi[2.88] + \Phi[1.92]\} \\ &= 2 - \{0.998012 + 0.97257\} \\ &\approx 2.94\% \end{aligned}$$

下表是根据公式求出的不同 K 值的  $C_P$  ( $C_{Pk}$ ) 值对应的 P 值供参考:

K- $C_P$ -P 的数表 (%)

表 5-3 K- $C_P$ -P 的数表



K	0.00	0.04	0.08	0.12	0.16	0.20	0.24	0.28	0.32	0.36	0.40	0.44	0.48	0.52
0.50	13.36	13.43	13.64	13.99	14.48	15.10	15.86	16.75	17.70	18.92	20.19	21.38	23.09	24.71
0.60	7.10	7.26	7.48	7.85	8.37	9.03	9.85	10.81	11.92	13.18	14.59	16.01	17.85	19.69
0.70	3.57	3.64	3.83	4.16	4.63	5.24	5.99	6.89	7.94	9.16	10.55	12.10	13.84	15.74
0.80	1.64	1.69	1.89	2.09	2.46	3.00	3.55	4.31	5.21	6.28	7.53	8.98	10.62	12.48
0.90	0.69	0.73	0.83	1.00	1.25	1.60	2.05	2.62	3.34	4.21	5.27	6.53	8.02	9.75
1.00	0.27	0.29	0.35	0.45	0.61	0.84	1.14	1.55	2.07	2.75	3.59	4.65	5.94	7.49
1.10	0.10	0.11	0.14	0.20	0.29	0.42	0.61	0.88	1.24	1.74	2.39	3.23	4.31	5.66
1.20	0.03	0.04	0.05	0.08	0.13	0.20	0.31	0.48	0.72	1.06	1.54	2.19	3.06	4.20
1.30	0.01	0.01	0.02	0.03	0.05	0.09	0.15	0.25	0.40	0.63	0.96	1.45	2.13	3.06
1.40	0.00	0.00	0.01	0.01	0.02	0.04	0.07	0.13	0.22	0.36	0.59	0.93	1.45	2.19
1.50			0.00	0.00	0.01	0.02	0.03	0.06	0.11	0.20	0.35	0.59	0.96	1.54
1.60					0.00	0.01	0.01	0.03	0.06	0.11	0.20	0.36	0.63	10.70
1.70						0.00	0.01	0.01	0.03	0.06	0.11	0.22	0.40	0.72
1.80							0.00	0.01	0.01	0.03	0.06	0.13	0.25	0.48
1.90								0.00	0.01	0.01	0.06	0.07	0.15	0.31
2.00									0.00	0.01	0.02	0.04	0.09	0.20
2.10										0.00	0.01	0.02	0.05	0.13
2.20											0.00	0.01	0.03	0.08
2.30												0.01	0.02	0.05
2.40												0.00	0.01	0.03
2.50													0.01	0.02
2.60													0.00	0.01
2.70														0.01
2.80														0.00

### 三、流程质量及其控制

流程质量及其控制,主要取决于6个因素,即人(Man)、机(Machine)、料(Material)、法(Method)、测(Measurement)和环(Environment/Mother nature),简称5M1E(或者6M)。在这6个方面,GB/T19001:2000标准都有明确的要求。

#### 1. 人一员工

“员工是一个组织之本”,是过程控制的主体,对管理者而言,至少应做到:

(1) 不断提高人员素质,培训员工的知识技能,因为员工是组织最根本的力量,只有员工强大了公司才会更加的强大。

(2) 重视人-机系统:在现代化生产中,人-机配合状态对过程质量至关重要,设备及其使用者同属一个系统,对人-机系统运行状态及其结果的分析是现代管理的一项内容,也是过程改进的重要信息。现代人体工程学、工业工程等都为改进人-机系统提供了适用的方法和理论基础。

(3) 注重人际关系:人际关系的作用能够为公司营造一种积极奋进、和谐互助、良好的工作氛围,同时它也是能够提升公司员工之间的凝聚力的重要因素。总而言之,人际关系的影响对于组织或者公司的影响是非常大的。对于管理者而言要积极改善人际关系,增强公司的凝聚力刻不容缓。

#### 2. 机即标准中的基础设施

GB/T19001:2000要求:组织的管理者应能提供并维护为实现产品符合性所需的各种基础设施,包括过程设备、工作场所和支持性服务等。这里有两层意思:



(1) 选择和提能高满足产品要求的各种设备，设备的过程能力应能满足过程综合能力 $C_p$ 的总要求。

(2) 通过诸如日点检，定期检查和维护保养，使设备始终处于良好状态。

### 3. 料，包括组织所采购的各种物料

组织应通过采购信息，明确采购产品(包括原材料、元器件、半成品等)要求，应根据对产品质量的影响采用适宜的方法控制并验证采购产品质量，其措施包括：

(1) 新产品开发过程，应提供需采购物料的清单。清单应包括明确的技术要求和检验要求。

(2) 通过进货验证或供方货源处抽检等手段，保证不合格物料不进入行产现场。

(3) 对提供物料的供方实现动态评价，并作必要的调整。

(4) 对物料流动的全过程进行状态标识，防止不合格物料的误用。

### 4. 法，包括管理方法和运行方法

从大的方面，应包括 2000 版质量标准中的流程方法、管理的系统方法，基于事实的决策方法，以及质量管理体系方法、统计方法等。另一方面，从流程角度，还包括质量体系或产品实现的各个阶段所需方法，如生策划、设计、工艺、检验、改进等具体方法。这些具体方法一般反映在组织的各类文件中，与文件控制要求密切相关。

### 5. 测，包括测量装置和测量保证

过程和产品的监视和测量，是质量管理体系的基本要求，也是质量改进的重要依据。对测量装置和测量过程应重点控制：

(1) 测量装置的选择和使用应能满足被测量产品的要求。

(2) 量值的溯源性和合理的校准周期。

(3) SPC 方法对测量过程进行分析和控制。

(4) 标识和记录。

### 6. 环，主要指工作环境和贮存条件

2000 版质量标准规定：组织应确定和管理为实现产品要示

所需的工作环境，并“应针对产品的符合性提供防护”，应控制的环境因素包括：

(1) 物理环境因素，如温度和湿度、噪音和振动、色彩和照明等。

(2) 化学环境因素，如粉尘和有害气体等。



(3) 心理环境因素，如创造一个和谐、协调的工作气氛等。

(4) 作业环境因素，如文明生产、安全生产、均衡生产等。

流程分析方法，包括统计分析方法和技术分析方法，大多数统计技术都可用于流程分析，如频数(频率)分布表、直方图、显著性分析、控制图、散点图、回归分析、方差分析、试验设计、多变量分析，等等。

所以强调用图表对过程分析，这是因为这些图表应用简单，但作用却很大，它能帮助激发思维，充分发挥大脑的潜能。实践证明，在过程分析中，用图(或表)分析要比文字性分析更加直观有效，这就是为什么总强调：能用图不用表，能用表不用文字的原因。



### 第三节 复习题

#### 一、单项选择题

1. 工程能力指数的计算公式中 T 指的是 ( )  
A. 产品的公差范围                      B. 产品的标准差  
C. 质量的数据分布中心                D. 中心偏移量
2. 在诸多因素中, 其处于受控状态, 那么通常用标准差的 ( ) 倍来表示工程能力的大小  
A. 1                      B. 3                      C. 6                      D. 9

#### 二、多项选择题

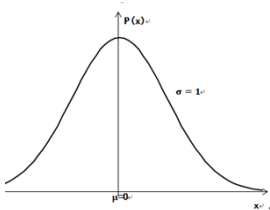
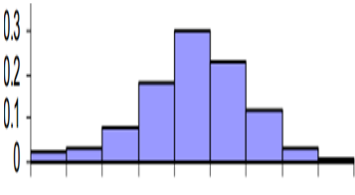
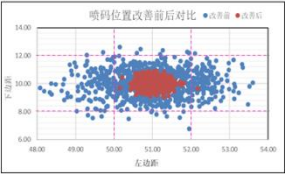
1. 下列选项中错误的是 ( )  
A. 过程能力: 过程能力的称法有好几种, 从早期工业开始的工序能力, 到现今称为过程能力、工程能力或工艺能力。  
B. 过程能力指数: 过程能力只是对一种过程固有的实际加工能力, 而已产品的技术要求无关  
C. 过程能力与质量之间的关系: 一般情况下, 过程能力和产品的质量实际波动成正比, 即过程能力越高, 质量波动就越高。

#### 三、简述题

1. 某线体头部胶工站需加工包胶一滚轮, 尺寸要求为  $\phi 20_{-0.010}^{+0.025}$ , 加工数量为 100 件以后, 计算得  $\bar{X} = 20.0075, S = 0.005$ , 求该过程的过程能力指数及不合格率。
2. 某机器加工的零件尺寸要求为:  $\phi 20 \pm 0.023$  现经随机抽样, 测得样本平均值  $\bar{X} = 19.997$  样本标准偏差为  $S = 0.007\text{mm}$ , 求  $CPk$  和  $P$ 。
3. 某工厂生产了一批整形下压块, 现要求径向跳动不超过  $0.05\text{mm}$ , 已知其样本平均值为  $\bar{X} = 0.02\text{mm}$ , 样本标准偏差  $S = 0.015\text{mm}$ , 求该过程的过程能力指数, 并且估计其不合格品率。

## 期末测试

### 一、单项选择题（每题 3 分）。

- 古典统计学时期是哪位科学家以叙述国家显著事项和国家政策关系为特征。（）
  - 威廉·佩第博士
  - 康令博士
  - 拉普拉斯
- 某人去某商场购买饮料，那么总体是（）
  - 某商场的所有饮料
  - 某商场的部分饮料
  - 某商场每一瓶饮料
- 在工厂、车间、生产现场经常会碰到统计数据，下面那个选项属于计数型的数据（）
  - 不合格品数
  - 化学成分
  - 电阻
- 下列选项中属于离散特征值得是（）
  - 极差
  - 平均值
  - 最大值
- 下列选项不属于 QC 七大工具的是（）
  - 控制图
  - 鱼骨图
  - 饼图
- 中位数-极差控制图指的是（）
  - $\bar{X}-R$
  - I-MR
  - $\bar{X}-R$
- 著名英国统计学家 Ronald Fisher 确立了（）小概率作为标准。
  - 0.05
  - 0.025
  - 0.03
- 以下图形属于正态分布图形的是（）
  - 
  - 
  - 
- 流程质量控制中的 5M1E，也称为（）
  - 6M
  - 6E
  - 6ME
- 生产过程中检验控制状态图，需要用到 8 种异常判定，下面不属于 8 种判异的是（）
  - 一个点出现在控制限之外





B. 连续 6 点递增或递减

C 连续 8 个点在同一侧

## 二、多项选择题（每题 5 分）

1. 统计学中指标体系的特点是（）

A. 同质事物的可量性    B. 量的综合性    C. 具有时间、地点和条件性

2. 计数数据又分为（）

A. 计点数据    B. 计件数据    C. 计量数据

3. QC 七大工具有（）

A. 鱼骨图    B. 直方图    C. 柏拉图

4. 下列属于随机现象的是（）

A. 向空中抛一枚硬币，它掉落后可能会正面向上，也可能会反面向上。

B. 买彩票中奖

C. 掷骰子不出现 7 点

5. 过程控制状态分为（）

A. 失控状态    B. 控制状态    C. 非稳定控制状态

## 三、简述题（共 45 分）

1. 某工厂，电池生产线生产的某型号的电池宽度标准值是  $10 \pm 1\text{cm}$ ，现从生产线随机抽取了 20 个电池测量其宽度，数据如下，试分析宽度的分布情况（20 分）。

7	11	11	14	10	13	12	10	10	9
10	9	11	13	15	13	10	10	11	9

2. 在下表中填写平均值与标准差的值，发现什么规律？（25）

	1	2	3	4	5	平均值	标准差
数据 1	2	3	5	7	8		
数据 2	4	5	7	9	10		
数据 3	4	6	10	14	16		



# 习题解答

## 第一章 复习题答案

### 一、单选题

1. A

2. C

### 三、简述题

1. 统计学的定义具有两种，一种是《不列颠百科全书》中的定义：统计学是“用以收集数据、分析数据和由数据得出结论的一组概念、原则和方法”。另外一种是“普通高等学校本科专业目录和专业设置”中给出的定义为：“统计学是研究如何收集数据，分析数据，以便得出正确认识结论的方法论科学。”

2. 凡是客观存在的，在同一性质基础上结合起来的许多事物（个体）的整体，就是统计总体。组成统计总体的个体称为总体单位。

3. 总体：该商场卖的所有吸尘器。总体单位：该商场每一台吸尘器。



## 第二章 复习题答案

### 一、单选题

1. C
2. C

### 二、多选题

1. ABC
2. AC

### 三、判断题

- 1 (×) 2 (×) 3 (√) 4 (√) 5 (√)

### 四、简述题

1. 统计方法分为两大类：推断型统计技术和描述型统计技术

推断型统计技术：主要解决从样本如何推断总体。概率论和数理统计研究的对象大多属于此类

描述型统计技术：主要是利用数据的特征或有关图表描述事物

2. 平均数、最大值、最小值、中位数
3. 计量数据是连续取值、计数数据不是连续取值；计量数据服从泊松分布，计数数据服从反二项分布。
4. 数据具有六大性质：统一性、针对性、完整性、准确性、及时性、连续性、统一性。
- 5 有利于提高管理的效率，促进质量管理体系的持续改进和产品质量的不断提高。



## 第三章 复习题答案

### 一、单项选择题

1. A      2. C      3. A

### 二、多项选择题

1. AB

### 三、简述题

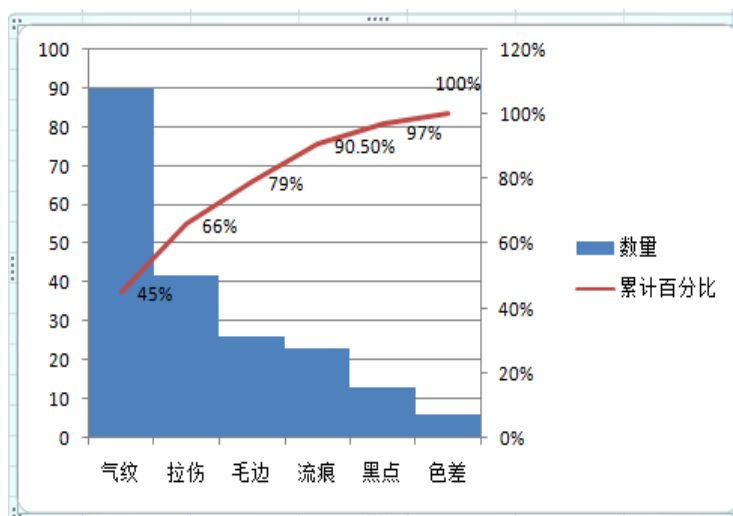
1. 可根据具体情况采用不同的分层类别，常见的有：

- (1) 操作人员：按个人、现场、班次、经验等分；
- (2) 机器设备：按机器、工夹具等分；
- (3) 材料：按来源、品种、生产批次分；
- (4) 加工方法：按加工、装配、测量、检验等各种工作条件分；
- (5) 时间：按上下午、年月季节分；
- (6) 环境：按气象情况、室内温度、磁场等环境分；
- (7) 其它：按发生情况、发生位置分。

2. 解：①整理数据制成汇总数据表

排序	分类名称	计算	百分比	累计百分比
1	气纹	90	45%	45%
2	拉伤	42	21%	66%
3	毛边	26	13%	79%
4	流痕	23	12%	91%
5	黑点	13	7%	97%
6	色差	6	3%	100%
	合计	200		

②制作柏拉图如下

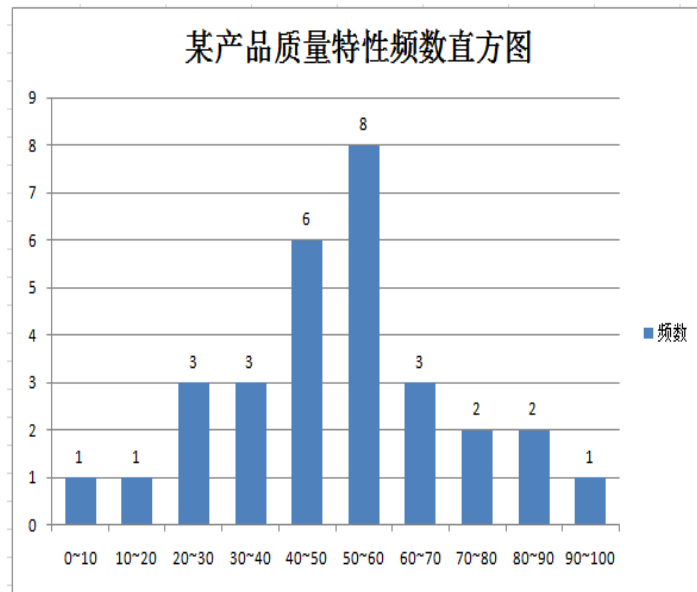


3. 频数分布直方图制作步骤如下

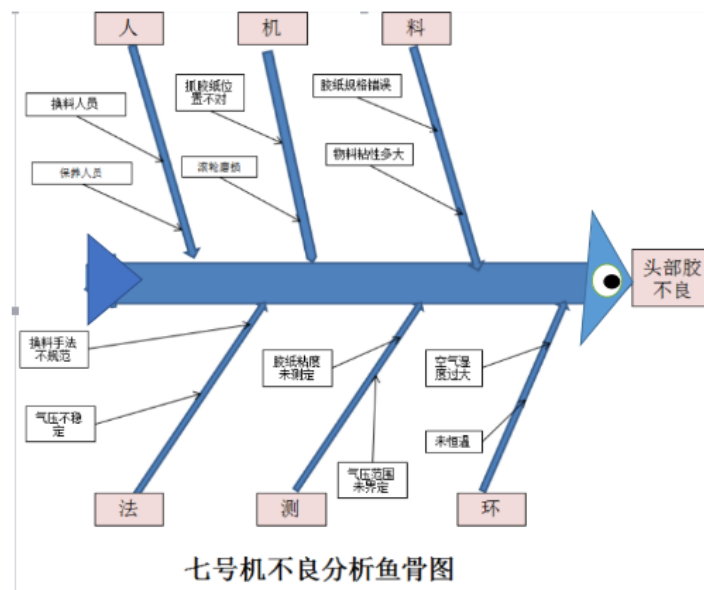
解①制作频数分布表

组号	组的中心值	组的左右临界点	频数
1	5	0~10	1
2	15	10~20	1
3	25	20~30	3
4	35	30~40	3
5	45	40~50	6
6	55	50~60	8
7	65	60~70	3
8	75	70~80	2
9	85	80~90	2
10	95	90~100	1

②制作频数直方图



4. 某产线头部胶不良原因分析鱼骨图如下



5. (1) 收集数据并加以分组，在 5M1E 充分固定，并标准化的情景下从生产过程中收集数据。

(2) 每个 2 小时，从生产过程中抽取 5 个零件，测量其长度，组成一个大小为 5 的成本，

(3) 计算总平均和极差平均;  $\bar{\bar{X}}=49.506$ ;  $\bar{R}=0.0800$

(4) 计算控制限

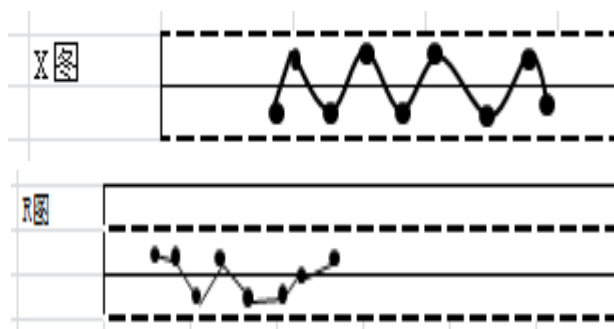
$\bar{X}$ 图: UCL=49.555; CL=49.506; LCL=49.460

R图: UCL=0.49553

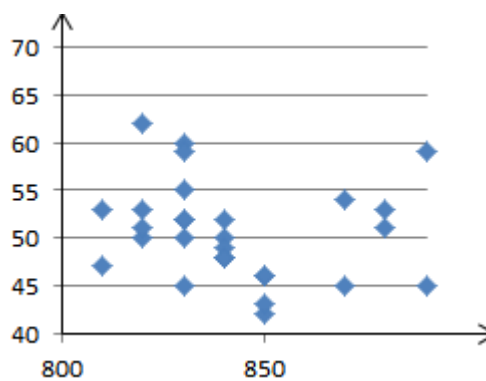
CL=0.0800

LCL=  $D_3\bar{R}$

根据数据制作控制图如下。



6. 散点图制作如下所示





## 第四章 复习题答案

### 一、单选题

1. C
2. B
3. D
4. D

### 二、多选题

1. AC
2. ABC

### 三、简述题

1. 个位是 0 的有  $C_4^1 A_4^3 = 96$  个，个位是 2 的有  $C_3^1 A_4^3 = 72$  个，个位是 4 的有  $C_3^1 A_4^3 = 72$  个，所以共有  $96+72+72=240$  个
2. 解：第一步，从后排 8 人选择 2 人，共有  $C_8^2$  种选法，第二步将选出的 2 人排在前排 6 个位置的 2 个，有  $A_6^2$  种排法，第三步，将前排 4 人相对位置不变，坐在前排剩余 4 个位置，只有一种坐法，故有  $C_8^2 A_6^2$  种方法。
3. 0.5934
4. (1) 40320      (2) 8640      (3) 5760
5. 2790 种





## 第五章 复习题答案

### 一、单项选择题

1. A

2. B

### 二、多项选择题

1. C

### 三、简述题

1. 1.17; 0.00046

2.  $C_{PK}=0.95; P=0.229\%$

3. 0.67; 7.75%



## 期末测试答案

### 一、单项选择题

1. B 2. A 3. A 4. A 5. C 6. A 7. A 8. A 9. A 10. C

### 二、多项选择题

1. ABC 2. AB 3. ABC 4. AB

### 三、简述题

1. 解题步骤如下：

①计算极差=15-7=8

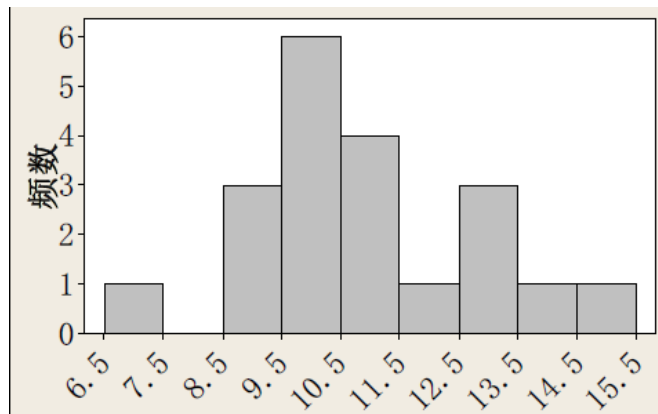
②样本组数选择为 k=7

③组距计算为  $h=1.14$  取值为 1

④确定边界绘制数据频数分布表

组号	组界	组中值	数据归类	频数	频率	累计频率
1	6.5~7.5	7	7,	1	0.05	0.05
2	7.5~8.5	8		0	0	0.05
3	8.5~9.5	9	9, 9, 9	3	0.15	0.2
4	9.5~10.5	10	10, 10, 10, 10, 10, 10	6	0.3	0.5
5	10.5~11.5	11	11, 11, 11, 11,	4	0.2	0.7
6	11.5~12.5	12	12,	1	0.05	0.75
7	12.5~13.5	13	13, 13, 13	3	0.15	0.9
8	13.5~14.5	14	14,	1	0.05	0.95
9	14.5~15.5	15	15,	1	0.05	1

⑤画出频数直方图如下



2.

	1	2	3	4	5	平均值	标准差
数据1	2	3	5	7	8	5	2.08
数据2	4	5	7	9	10	7	2.08
数据3	4	6	10	14	16	10	4.16

规律：数据每一项加固定值，平均值也加固定值，标准差不变。数据每一项乘以 2，



---

平均值也乘以 2，标准差也是乘以 2。



# 附表 1 标准正态分布表

标准正态分布表

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt (x \geq 0)$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5398	0.5438	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5793	0.5832	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6179	0.6217	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6554	0.6591	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6915	0.6950	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7257	0.7291	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7580	0.7611	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7881	0.7910	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8159	0.8186	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8413	0.8438	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8643	0.8665	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8849	0.8869	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9032	0.9049	0.9082	0.9099	0.9115	0.9133	0.9147	0.9162	0.9177
1.4	0.9192	0.9192	0.9207	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9332	0.9345	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9452	0.9463	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9554	0.9540	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9713	0.9649	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9772	0.9719	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9821	0.9776	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9861	0.9826	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9893	0.9864	0.9871	0.9875	0.9878	0.9880	0.9884	0.9887	0.9890
2.3	0.9893	0.9918	0.9896	0.9901	0.9904	0.9906	0.9907	0.9909	0.9911	0.9912
2.4	0.9918	0.9938	0.9920	0.9925	0.9927	0.9929	0.9930	0.9932	0.9934	0.9936
2.5	0.9938	0.9953	0.9940	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9965	0.9955	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9974	0.9966	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9981	0.9975	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000									

## 附表 2 控制图判异规则

判异规则:

按国标 GB/T 4091-2001, 8 种判异现象如下:





附表 3 计量值控制图系数表

子组大小n	均值控制图			标准差控制图						极差控制图						
	控制线系数			中心系数		控制线系数				中心线系数			控制线系数			
	A	A <sub>2</sub>	A <sub>3</sub>	C <sub>4</sub>	1/C <sub>4</sub>	B <sub>3</sub>	B <sub>4</sub>	B <sub>5</sub>	B <sub>6</sub>	d <sub>1</sub>	1/d <sub>2</sub>	d <sub>3</sub>	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>
2	2.121	1.880	2.659	0.7979	1.2533	0	3.267	0	2.606	1.128	0.8865	0.853	0	3.686	0	3.267
3	1.732	1.023	1.954	0.8862	1.1284	0	2.568	0	2.276	1.693	0.5907	0.888	0	4.358	0	2.574
4	1.500	0.729	1.628	0.9213	1.0854	0	2.266	0	2.088	2.059	0.4857	0.880	0	4.698	0	2.282
5	1.342	0.577	1.427	0.9400	1.0638	0	2.089	0	1.964	2.326	0.4299	0.864	0	4.918	0	2.114
6	1.225	0.483	1.287	0.9515	1.0510	0.030	1.970	0.029	1.874	2.534	0.3946	0.848	0	5.078	0	2.004
7	1.134	0.419	1.182	0.9594	1.0423	0.118	1.882	0.113	1.86	2.704	0.3698	0.833	0.204	5.204	0.076	1.924
8	1.061	0.373	1.099	0.9650	1.0363	0.185	1.815	0.179	1.751	2.847	0.3512	0.820	0.388	5.306	0.136	1.864
9	1.000	0.337	1.032	0.9693	1.0317	0.239	1.761	0.232	1.707	2.970	0.3367	0.808	0.547	5.393	0.184	1.816
10	0.949	0.308	0.975	0.9727	1.0281	0.284	1.716	0.276	1.669	3.078	0.3249	0.797	0.687	5.469	0.223	1.777
11	0.905	0.285	0.927	0.9754	1.0252	0.321	1.679	0.313	1.637	3.173	0.3152	0.787	0.811	5.506	0.136	1.864
12	0.866	0.266	0.886	0.9776	1.0229	0.354	1.646	0.346	1.610	3.258	0.3069	0.778	0.922	5.594	0.283	1.717
13	0.832	0.249	0.850	0.9794	1.0210	0.382	1.618	0.374	1.585	3.336	0.2998	0.770	1.025	5.647	0.307	1.693
14	0.802	0.235	0.817	0.9810	1.0194	0.406	1.594	0.399	1.563	3.407	0.2935	0.763	1.118	5.696	0.328	1.672
15	0.775	0.223	0.789	0.9823	1.0180	0.428	1.572	0.421	1.544	3.472	0.2880	0.756	1.203	5.741	0.347	1.653
16	0.750	0.212	0.763	0.9835	1.0168	0.448	1.552	0.440	1.526	3.532	0.2831	0.750	1.282	5.782	0.363	1.637
17	0.728	0.203	0.739	0.9845	1.0157	0.466	1.534	0.458	1.511	3.588	0.2787	0.744	1.356	5.820	0.378	1.622
18	0.707	0.194	0.718	0.9854	1.0148	0.482	1.518	0.475	1.496	3.640	0.2747	0.739	1.424	5.856	0.391	1.608
19	0.688	0.187	0.698	0.9862	1.0140	0.497	1.503	0.490	1.483	3.689	0.2711	0.734	1.487	5.891	0.403	1.597
20	0.671	0.180	0.680	0.9869	1.0133	0.510	1.490	0.504	1.470	3.735	0.2677	0.729	1.549	5.921	0.415	1.585
21	0.655	0.173	0.663	0.9876	1.0126	0.523	1.477	0.516	1.459	3.778	0.2647	0.724	1.605	5.951	0.425	1.575
22	0.640	0.167	0.647	0.9882	1.0119	0.534	1.466	0.528	1.448	3.819	0.2618	0.720	1.659	5.979	0.434	1.566
23	0.626	0.162	0.633	0.9887	1.0114	0.545	1.455	0.539	1.438	3.858	0.2592	0.716	1.710	6.006	0.443	1.557
24	0.612	0.157	0.619	0.9892	1.0109	0.555	1.445	0.549	1.429	3.895	0.2567	0.712	1.759	6.031	0.451	1.548
25	0.600	0.153	0.606	0.9896	1.0105	0.565	1.435	0.559	1.420	3.931	0.2544	0.708	1.806	6.056	0.459	1.541

$$\text{当 } n > 25 \text{ 时, } A = \frac{3}{\sqrt{n}}, A_3 = \frac{3}{C_4 \sqrt{n}}, C_4 = \frac{4(n-1)}{4n-3}, B_3 = 1 - \frac{3}{C_4 \sqrt{2(n-1)}}, B_4 = 1 + \frac{3}{C_4 \sqrt{2(n-1)}}$$

$$B_5 = C_4 - 3\sqrt{1 - C_4^2}, B_6 = C_4 + 3\sqrt{1 - C_4^2}。$$



## 参考文献

1. 于义良. 概率论与数理统计. 北京: 中国人民大学出版社, 2004
2. 宇传华. Excel 与数据分析. 北京: 电子工业出版社, 2003
3. 陈吉象. 文科数学基础. 北京: 高等教育出版社, 2003
4. 茆诗松, 程依明, 濮晓龙. 概率论与数理统计教程第二版. 北京: 高等教育出版社, 2011.
5. 杨振明. 概率论. 北京: 科学出版社, 1999
6. 周润兰. 应用概率统计. 北京: 科学出版社, 1999
7. 李贤平. 概率论基础. 3 版. 北京: 高等教育出版社, 2010
8. 邢文英. 2004. QC 小组基础教材 (修订版). 北京: 中国社会出版社
9. 李为柱. 2003. 统计技术应用教程. 北京: 企业管理出版社
10. 徐国祥. 管理统计学. 上海财经大学出版社, 1995
11. 伍爱. 2001. 质量管理学 (第二版). 广州: 暨南大学出版社
12. 吴建伟. 2004. ISO9000:2000 认证通用教程. 北京: 机械工业出版社
13. 张公绪, 孙静. 2003. 质量工程师手册 (第二版). 北京: 企业管理出版社
14. 秦现生. 2002. 质量管理学. 北京: 科学出版社
15. 李明荣. 质量管理. 北京: 科学出版社, 2007



欣旺达电子股份有限公司  
Sunwoda Electronic Co., Ltd.

地址: 深圳市宝安区石岩街道石龙社区颐和路2号欣旺达新能源产业园

电话: 86-755-2951 6888 传真: 86-755-2951 6999

网址: [www.sunwoda.com](http://www.sunwoda.com)