

前 言

20 世纪 80 年代中期,摩托罗拉公司在质量觉醒的过程中创立了六西格玛管理。90 年代,通用电气、联合信号、福特、卡特彼勒、陶氏、杜邦、ABB、三星、LG、柯达、西门子、花旗银行等公司将六西格玛管理演绎成为提升企业核心竞争力的经营战略。时至今日,六西格玛管理的热潮依然不减,其应用范围从制造业到服务业,从教育、医疗机构到政府乃至军队,越来越多的组织在尝试它,越来越多的企业家、管理学者在关注它、研究它。成功实施六西格玛管理的企业在业界,尤其是供应链上所起的示范作用则带动了更多的企业对六西格玛管理跃跃欲试。

正确实施六西格玛管理需要正确的理念和方法的指导,为适应时代的需求,全国六西格玛管理推进工作委员会做了大量的工作,包括成功举办了四届全国六西格玛大会;2003 年确定了六西格玛黑带知识大纲,建立了六西格玛黑带注册考试制度,自 2004 年至今已举行了三次注册六西格玛黑带全国统一考试;举办了多次六西格玛教师培训班;2004 年出版了第一版《六西格玛管理》教材;2006 年组织制定了《六西格玛管理评价准则》。所有这些,对我国六西格玛管理的推进工作起了关键的支撑作用。

《六西格玛管理》(第一版)出版后,社会反映很好,出版社多次印刷。同时在使用过程中也发现内容上有许多不够完善的地方,因此,我们决定对第一版重新修订。在《六西格玛管理》第二版编写过程中,全国六西格玛管理推进工作委员会专家委员会多次开会研究,编委会成员也反复讨论,广泛征求了六西格玛实施企业和有关专家的意见,对第一版的内容做了比较大的调整、修改、补充和完善。

《六西格玛管理》(第二版)由马林、何桢担任主编。参加编写工作的有何桢、马逢时、龚晓明、马义中、杨跃进、孙静、牛占文、邵家骏。岳刚、王丽林、赵建坤参加了本书的修订、讨论和编写工作。张敏、施亮星、韩亚娟参加了部分书稿的校对工作。全书最终由何桢负责审校。

本书编写过程中得到了中国质量协会副会长、全国六西格玛



管理推进工作委员会主任弋辉的大力支持。全国六西格玛管理推进工作委员会委员单位，宝山钢铁股份有限公司、中国航空工业第一集团公司、中国航空工业第二集团公司、太原钢铁集团有限公司、戴尔中国有限公司、TCL 集团有限公司等为本书的修改提供了大量建设性的意见。

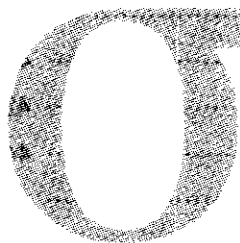
本书是在第一版的基础上修订完成的，马逢时、杨跃进、王金德、彭娟、邵家骏、孙静等参加了本书第一版编写工作，在此对他们付出的辛劳表示诚挚的感谢。

另外本书的部分成果得到了国家自然科学基金项目“实施六西格玛质量改进的关键技术研究”（编号：70372062）、“制造业六西格玛设计方法和应用研究”（编号：70572044）和“六西格玛管理的保障机制和技术研究”（编号：70672088）的资助，在此一并表示感谢。

最后，我们要特别感谢中国人民大学出版社对本书给予的高度重视和支持。

由于时间和水平所限，书中难免存在疏漏和错误，恳请读者批评指正，具体意见可发送至 wll@caq.org.cn。借本书出版之际，向广大读者表示最衷心的感谢。

《六西格玛管理》编委会
中国质量协会六西格玛管理推进工作委员会



目 录

第1篇 总 论	1
第1章 六西格玛管理概论	3
1.1 六西格玛管理的发展	3
1.1.1 质量概念的演进和质量管理的发展	3
1.1.2 六西格玛管理的起源和发展	9
1.2 六西格玛的概念和作用	13
1.2.1 六西格玛的概念	13
1.2.2 六西格玛管理的作用	15
1.3 六西格玛管理的领导、文化和战略	19
1.3.1 高层领导在六西格玛管理中的作用	19
1.3.2 六西格玛价值观与企业文化	22
1.3.3 六西格玛管理与企业战略	24
1.4 六西格玛管理的组织和推进	28
1.4.1 六西格玛管理的组织结构	28
1.4.2 六西格玛管理的推进步骤	32
1.5 六西格玛管理方法论	34
1.5.1 六西格玛改进的模式——DMAIC	34
1.5.2 六西格玛设计的模式	36
第2章 六西格玛与过程管理	38
2.1 过程管理基础	38
2.1.1 过程的定义	38



2.1.2	过程负责人和相关方	40
2.1.3	价值链与过程类别	41
2.1.4	过程与职能、项目的关系	42
2.2	顾客需求分析	44
2.2.1	顾客的识别	44
2.2.2	顾客满意与顾客忠诚	45
2.2.3	顾客数据的收集	47
2.2.4	顾客数据的分析	49
2.2.5	关键顾客要求的转化	50
2.3	经营结果	50
2.3.1	过程绩效度量指标	50
2.3.2	水平对比	54
2.3.3	财务收益	56
第3章	六西格玛项目管理	66
3.1	六西格玛项目选择	66
3.1.1	六西格玛项目选择原则	66
3.1.2	六西格玛项目选择流程	67
3.1.3	六西格玛项目选择需注意的问题	70
3.2	六西格玛项目立项表和计划	71
3.2.1	项目立项表和计划概述	71
3.2.2	项目规划工具	73
3.2.3	项目文档	75
3.3	六西格玛项目团队建设	75
3.3.1	团队的组建和授权	75
3.3.2	团队发展阶段	77
3.3.3	团队动力与绩效	78
3.3.4	团队工具	80
3.4	六西格玛项目监控与促进变革	83
3.4.1	项目跟踪和监控	83
3.4.2	促进变革	85
3.5	六西格玛项目管理与策划工具	88
3.5.1	亲和图	88
3.5.2	关联图	89
3.5.3	树图	90
3.5.4	矩阵图	91

3.5.5 优先矩阵图	92
3.5.6 过程决策程序图	93
3.5.7 网络图	95
3.6 六西格玛项目总结与成果评审	97
3.6.1 六西格玛项目总结	98
3.6.2 六西格玛项目成果评审与分享	99
第2篇 DMAIC 流程	103
第4章 界定	105
4.1 界定项目范围	105
4.1.1 确认顾客关键要求	105
4.1.2 过程的 SIPOC 分析	106
4.1.3 关键过程输出变量的确定——排列图法	109
4.1.4 项目的利益相关方分析	112
4.2 确定项目测量指标	113
4.2.1 关键质量特性 CTQ 的树图展开	113
4.2.2 建立测量指标	114
4.2.3 测算项目的西格玛水平	115
4.3 编制和完善项目立项表	115
4.3.1 项目背景	115
4.3.2 问题/机会和目标的陈述	116
4.3.3 项目范围、约束和团队任务	117
4.3.4 项目计划	118
4.3.5 项目立项表样例	118
第5章 测量	121
5.1 过程分析与文档	121
5.1.1 流程图	122
5.1.2 因果图与因果矩阵	124
5.1.3 其他过程分析工具与文档	126
5.2 概率与数理统计基础	128
5.2.1 概率论基础知识	128
5.2.2 随机变量及分布	132
5.2.3 数学期望、均值与方差	137
5.2.4 常用的离散分布	139
5.2.5 常用的连续分布	142
5.2.6 中心极限定理	147



5.2.7 统计量与抽样分布	148
5.3 数据的收集和整理	151
5.3.1 数据类型与测量尺度	151
5.3.2 收集数据的方法	152
5.3.3 抽样方法	153
5.3.4 描述性统计方法	155
5.3.5 数据的图示方法	158
5.4 测量系统分析	166
5.4.1 基本概念	166
5.4.2 测量系统的分辨力	168
5.4.3 测量系统的偏倚、线性和稳定性	170
5.4.4 测量系统的重复性与再现性	173
5.4.5 破坏性试验的测量系统分析	181
5.4.6 属性值数据的测量系统分析	184
5.4.7 测量仪器的校准和检定	187
5.5 过程能力分析	190
5.5.1 过程统计控制状态	190
5.5.2 过程能力和过程绩效	191
5.5.3 过程能力指数 C_p 与 C_{pk}	192
5.5.4 过程能力指数 C_{pm} 和 C_{pmk}	195
5.5.5 过程绩效指数 P_p 与 P_{pk}	197
5.5.6 过程能力指数与缺陷率的关系	199
5.5.7 长期能力与短期能力	201
5.5.8 非正态数据的变换	202
5.5.9 属性值数据的西格玛水平估算	205
第6章 分析	208
6.1 探索性数据分析和过程分析	208
6.1.1 探索性数据分析	209
6.1.2 过程分析	215
6.1.3 点估计和区间估计	219
6.2 假设检验	224
6.2.1 假设检验的基本概念	225
6.2.2 均值、方差和比例的假设检验	230
6.2.3 配对数据检验	240
6.2.4 拟合优度检验	242

6.3	方差分析	244
6.3.1	单因子方差分析	244
6.3.2	两因子方差分析	250
6.4	列联表	256
6.4.1	独立性检验概念	256
6.4.2	列联表的独立性检验方法	257
6.5	非参数检验	263
6.5.1	符号检验	264
6.5.2	曼-惠特尼-威尔科克森检验	269
6.5.3	威尔科克森符号秩检验	272
6.5.4	克鲁斯卡尔-沃利斯检验	275
6.5.5	数据独立性的键检验	277
6.6	相关分析与回归分析	279
6.6.1	相关分析	279
6.6.2	一元线性回归模型	282
6.6.3	多元线性回归模型	290
第7章	改进	296
7.1	试验设计基础	298
7.1.1	试验设计中的基本术语	298
7.1.2	试验设计的基本原则	303
7.1.3	试验设计的必要性	304
7.1.4	试验设计的类型	305
7.1.5	试验设计的策划与安排	306
7.1.6	试验设计的基本步骤	307
7.2	单因子试验设计与分析	308
7.2.1	单向分类设计	308
7.2.2	多项式回归	310
7.3	全因子设计与分析	314
7.3.1	二水平全因子试验概述	314
7.3.2	全因子设计的计划	316
7.3.3	全因子设计的分析	318
7.3.4	全因子设计实例	323
7.4	部分因子试验	340
7.4.1	部分因子试验概论	340
7.4.2	部分实施因子试验的计划	345



7.4.3	部分实施因子试验的分析实例	347
7.4.4	Plackett-Burman 设计——另一类筛选因子设计	356
7.4.5	三水平部分因子试验的分析	356
7.5	响应曲面设计与分析	359
7.5.1	响应曲面设计概论	359
7.5.2	响应曲面设计的计划	364
7.5.3	响应曲面设计的分析及实例	365
7.6	稳健参数设计	375
7.6.1	稳健参数设计的模型	375
7.6.2	稳健参数设计的计划	380
7.6.3	稳健参数设计实例分析	382
7.7	混料设计与分析	390
7.7.1	混料设计概论	390
7.7.2	混料试验的计划	392
7.7.3	混料试验的分析	396
7.8	调优运算	400
7.8.1	调优运算概论	400
7.8.2	调优运算方法	400
第 8 章	控制	402
8.1	控制阶段概述	402
8.1.1	过程改进成果的文件化	402
8.1.2	建立过程控制计划	403
8.1.3	实施持续的过程测量和控制	404
8.2	常规控制图	404
8.2.1	控制图的基本概念	405
8.2.2	计量控制图	413
8.2.3	计数控制图	419
8.3	特殊的控制图	423
8.3.1	累积和控制图	423
8.3.2	指数加权滑动平均控制图	427
8.3.3	短期生产统计过程控制技术	429
8.3.4	预控制	429
第 3 篇	精益生产与六西格玛设计	431
第 9 章	精益生产	433
9.1	精益生产概述	433



9.1.1	精益生产的产生及发展	433
9.1.2	精益生产的核心思想	434
9.1.3	精益生产与传统生产模式的比较	435
9.1.4	精益思想的五项基本原则	435
9.2	精益生产的技术体系及主要支撑技术	437
9.2.1	精益生产的技术体系	437
9.2.2	看板拉动式生产	438
9.2.3	均衡化生产与设备的快速换型调整	440
9.2.4	设备的合理布置与多技能作业员	442
9.2.5	标准化作业	443
9.2.6	5S与目视管理	445
9.2.7	自働化与差错预防	447
9.2.8	全面生产维护	448
9.2.9	价值流图分析	452
9.3	精益与六西格玛的结合	458
9.3.1	精益生产与六西格玛的特点及比较	458
9.3.2	精益与六西格玛的有机结合	461
9.3.3	精益六西格玛的实施路线/技术路线	462
第 10 章	六西格玛设计	465
10.1	六西格玛设计概述	465
10.1.1	六西格玛设计的产生和发展	465
10.1.2	六西格玛设计的重要意义	466
10.1.3	六西格玛设计流程	468
10.1.4	六西格玛设计部署	470
10.1.5	六西格玛设计的指导思想	471
10.1.6	六西格玛设计的主要工具	472
10.1.7	产品全生命周期的六西格玛解决方案	474
10.2	质量功能展开	475
10.2.1	概述	475
10.2.2	QFD 的基本方法	476
10.2.3	QFD 的工作程序	482
10.2.4	应用于制造业之外的 QFD	487
10.3	系统设计	489
10.3.1	系统设计的基本原理与方法	489
10.3.2	解决创造性问题的理论	490



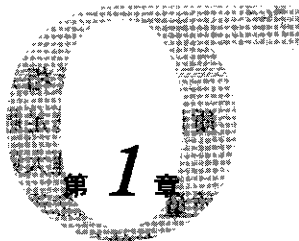
10.3.3	TRIZ 方法的结构体系	493
10.3.4	创造性问题的分析方法	495
10.3.5	创造性问题的解决方法	499
10.4	静态参数设计	503
10.5	容差设计	503
10.5.1	质量损失函数	503
10.5.2	容差的确定方法	504
10.5.3	望目、望小特性的容差设计	509
10.5.4	望大特性的容差设计	510
10.5.5	贡献率法	511
10.6	FMEA 分析	517
10.6.1	FMEA 方法的适用范围	517
10.6.2	FMEA 方法的分类	518
10.6.3	FMEA 的一般程序	518
10.6.4	DFMEA 分析	519
10.6.5	PFMEA 分析	524
10.7	面向 X 的设计	530
10.7.1	面向 X 设计的原理	530
10.7.2	面向制造和装配的设计	534
10.7.3	面向成本的设计	535
10.7.4	面向试验的设计	537
10.7.5	面向可靠性和维修性的设计	538
10.7.6	面向售后服务的设计	539
10.7.7	面向环境的设计	540
10.7.8	面向其他因素的设计	541
附 表	542
常用六西格玛中英文名词索引	565
参考文献	577
中国质量协会注册六西格玛黑带考试样题	580

第1篇

总论

σ

- 第1章 六西格玛管理概论
- 第2章 六西格玛与过程管理
- 第3章 六西格玛项目管理



六西格玛管理概论



六西格玛管理的发展

1.1.1 质量概念的演进和质量管理的发展

人类社会的质量活动可以追溯到远古时代，而现代意义上的质量管理活动则是从 20 世纪初开始的。从此，人类跨入了以加工机械化、经营规模化、资本垄断化为特征的工业化时代。

质量管理历经百年发展，各种方法风生水起、纷至沓来。六西格玛管理是质量管理在 20 世纪末最具魅力的新发展之一。六西格玛管理的起源、发展，正是在质量概念演进和质量管理发展的大背景下进行的。

1. 质量概念的演进

随着质量管理、技术以及市场不断发展，质量的概念也逐渐地拓展、深化和完善。在 ISO 9000:2000 标准中，质量被定义为“一组固有特性满足要求的程度”。

(1) 质量概念中主体的演进。质量概念中的主体指的是“什么的质量”，依照演进的顺序，可表达为：

- 产品的质量。包括性能、可信性等实物质量。
- 产品和服务的质量。扩展到包括准时交付、周期时间等服务质量。
- 产品、服务和过程的质量。其中过程质量（也称工作质量）涉及 5M1E（人、机、料、法、环、测），这时的质量已是体现 Q（实物质量）、C（成本）、D（交付）、E（环境）、S（安全）等的综合质量。
- 产品、服务、过程和体系的质量。这里的体系质量即管理系统的质量，或称管理质量、经营质量，可以是一个组织的经营管理质量，甚至可以是一个跨组织共同体的运作质量（如与顾客和供应商的伙伴关系质量，社区经济运行质量）。

显然，这是一个质量由“小”至“大”的演进过程。

(2) 质量概念中客体的演进。质量概念中的客体指的是“满足什么要求的质量”，依照演进的顺序，可表达为：

- 符合性质量。满足标准或规范要求，即合格，反之则不合格。
- 适用性质量。质量就是满足顾客要求的程度，质量好不好关键看对顾客是否适用。





● 顾客及相关方综合满意的质量。此时的质量已经是产品、服务、过程和体系的“大质量”综合满足顾客、股东、员工、供应商及合作伙伴、社会等利益相关方的程度。

质量管理大师朱兰在《朱兰质量手册》(第5版)中这样提到,20世纪80年代,“大质量”的概念逐渐广为人们所接受,且成为不可逆转的趋势。正是在这样一个质量概念剧烈演变的年代,六西格玛管理诞生了。因此,六西格玛管理深深地打上了“大质量”概念的烙印。21世纪是质量的世纪,亦即“大质量”的世纪。在这样一个“大质量”的时代,六西格玛管理恰逢盛世,在世界各地蓬勃发展。

2. 质量管理的发展

伴随着质量概念的演进,现代质量管理也在演进。从全球工业发达国家的情况看,质量管理历经了质量检验、统计质量控制和全面质量管理三大历史阶段。第二次世界大战以前可以看作第一阶段,通常称为质量检验阶段;20世纪40—50年代为第二阶段,通常称为统计质量控制阶段;第三阶段为20世纪60年代开始的全面质量管理阶段。

(1) 质量检验阶段。质量检验源远流长,在家庭作坊制生产条件下,生产职能和检验职能没有分开。20世纪初,随着机器化大生产的出现,检验职能从生产职能单独分离出来。这一阶段主要是通过检验的方式来控制和保证产出或转入下道工序的产品质量,主要特点是事后把关。其演进历程为:工人自检、工长监督检查、检验员专检。其倡导者是美国泰勒(F. W. Taylor),他也是“科学管理运动”的奠基人,提出了科学管理理论,要求按职能的不同进行合理的分工,首次将质量检验作为一种管理职能从生产过程中分离出来,建立了专职检查部门和检验制度。同时,基于大批量生产的产品技术标准的建立和公差界限的规定,也为质量检验奠定了基础。

当时,质量检验存在两个主要问题:一是事后检验,无法在生产过程中进行预防和控制,而且属于非增值活动,是低质量的代价,因此人们也称之为“死后验尸”;二是全数检验,成本太高,特别是在破坏性检验条件下更是无法进行。由此导致统计质量控制理论的诞生。

(2) 统计质量控制阶段。统计质量控制(statistical quality control, SQC)阶段酝酿于20世纪20—30年代,形成于40—50年代。这一阶段的特征是数理统计方法与质量管理的结合,从单纯依靠质量检验“事后把关”,发展到过程控制,形成了质量的预防性控制与事后检验相结合的管理方式。

在生产的推动下,统计学的应用有了很大进展。20世纪20年代英国数学家费希尔(R. A. Fisher)结合农业实验提出方差分析与试验设计等理论,为近代数理统计学奠定了基础。与此同时,美国贝尔实验室成立了两个课题研究组:一个是过程控制组,学术负责人是休哈特(W. A. Shewhart),创建了统计过程控制(statistical process control, SPC)理论,实现了应用统计技术对生产过程的监控;另一个是产品控制组,学术负责人是道奇(H. F. Dodge),他和同事罗米格(H. G. Romig)进行抽样检验的探索,在1929年发表了《抽样检验方法》,并设计了实用的抽样检验表,解决了全数检验和破坏性检验在应用中的困难。

在20世纪20—30年代提出过程控制理论与抽样检验理论之际,恰逢西方发达国家处于经济萧条时期,所以这些新理论乏人问津。直到第二次世界大战期间,由于国防工业迫切需要保证军火质量,这些理论才得到广泛应用。上述理论的实际应用效果显著,得到了



广泛的认可，战后风行全世界。

统计方法的应用减少了不合格品，降低了生产费用。但统计质量控制过分强调统计方法，使人们误认为质量管理就是统计方法，而且在缺乏计算机和数理统计软件支持的情况下，许多人感到难度大。随着现代化大规模生产的演进，影响产品质量的因素越来越多，单纯依靠统计方法不可能解决一切质量问题。质量管理由此走向系统工程的道路。

(3) 全面质量管理阶段。20 世纪 60 年代以来，随着科学技术和工业生产的发展，对质量的要求越来越高，这就需要人们运用“系统工程”的概念，将质量问题作为一个有机整体加以综合分析研究，实施全员、全过程、全企业的管理。美国通用电气公司的费根堡姆 (Feigenbaum) 首先提出了全面质量管理 (total quality control, TQC) 的概念，1961 年在其《全面质量管理》一书中指出：“全面质量管理是为了能在最经济的水平上并考虑到充分满足顾客需求的条件下进行市场研究、设计、生产和服务，把企业各部门的研制质量、维持质量和提高质量的活动构成一体的有效体系。”

日本在 20 世纪 50 年代引进美国的质量管理方法后，对其进行了发展创新，提出了全公司质量管理 (company wide quality control, CWQC)，首创了质量管理小组 (quality control circles, QCC) 方法、田口方法 (Taguchi method)、5S 管理 (日文 Seiri, Seiton, Seisou, Seiketsu, Sitsuke，即整理、整顿、清洁、清扫和素养)、全面生产维护 (total productive maintenance, TPM)、质量功能展开 (quality function deployment, QFD) 和丰田生产方式 (Toyota production system, TPS) 等，归纳了“老七种”、“新七种”工具并普遍用于质量改进和质量控制，使全面质量管理充实了大量新的内容。质量管理的手段不再局限于数理统计，而是全面地运用各种管理技术和方法。日本企业应用全面质量管理获得了极大的成功，引起了世界各国的关注，全面质量管理的观念在全球范围内得到广泛传播，各国结合各自的国情及实践皆有所创新与发展。

1979 年，美国质量管理专家克罗斯比 (Crosby) 出版了他的开山之作《质量免费——确定质量的艺术》，确立了“第一次就把事情做对”和“零缺陷”的理论。“零缺陷”的四项基本原则是：明确需求、做好预防、一次做对、科学衡量。

1987 年是质量管理发展史上最重要的一年：国际标准化组织发布了其第一套管理标准——ISO 9000 系列标准，适应全球化贸易的质量体系认证由此拉开帷幕；摩托罗拉公司在总结 70 年代竞争失利的基础上提出六西格玛方法并在公司正式实施；美国国会于 1987 年通过了《马尔科姆·波多里奇国家质量提高法——“公共法案 100—107”》，决定启动波多里奇国家质量奖评审，这是自 20 世纪 80 年代初美国重新审视和借鉴日本的发展、将 TQC 发展到 TQM (total quality management) 以来的一个里程碑，为 TQM 建立了一个从过程到结果的卓越绩效评价框架。

由 TQC 向 TQM 的演进，实质上是质量概念由“产品和服务质量满足顾客需要”向“大质量综合满足顾客及相关方需要”、质量管理由“全面的质量管理”向“全面质量的管理”的演进。国际标准化组织在 ISO 8402:1994 中将 TQM 定义为“一个组织以质量为中心，以全员参与为基础，目的在于通过让顾客满意和本组织所有成员及社会受益而达到长期成功的管理途径”，视其作为一种卓越经营的哲学和方法。但如何构建组织的 TQM？如何评估 TQM 做到什么程度，成效如何？如何识别经营管理的优势和改进机会，并推动持续改进和创新，不断提高管理的成熟水平？于是卓越绩效（质量奖）模式应运而生。几乎



所有经济发达和强劲发展的国家和地区均建立了各自的卓越绩效（质量奖）模式，以推动所在国家、地区的经营管理进步和核心竞争力提升。最经典的卓越绩效模式是三大质量奖：美国波多里奇国家质量奖、欧洲质量奖和日本戴明奖。其中波多里奇国家质量奖在使美国经济恢复活力以及在提高美国国家竞争力和生活质量等方面起到了重要作用，影响最为广泛。

为了适应我国市场经济发展的形势，应对入世后激烈竞争的挑战，在学习借鉴美国波多里奇国家质量奖和日本戴明奖成功经验的基础上，中国质量协会于 2001 年启动了全国质量管理奖评审工作，并从 2005 年开始，采用了新颁布的国家标准 GB/T19580《卓越绩效评价准则》和 GB/Z19579《卓越绩效评价准则实施指南》（其框架如图 1—1 所示）。

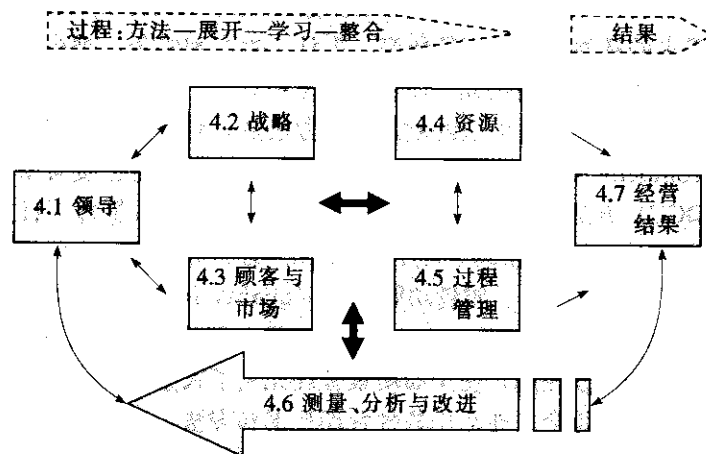


图 1—1 GB/T19580 和 GB/Z19579 框架图

今天，卓越绩效模式已成为现代 TQM 的实施框架和评价准则，成为企业经营管理事实上的国际标准。正如朱兰博士所言，卓越绩效模式的广泛传播，使其“成为关于 TQM 内容的最广为接受的定义”。

3. 质量大师的贡献

在质量管理发展过程中出现了许多质量管理大师，如美国的休哈特、戴明、朱兰、费根堡姆、克劳斯比，瑞典的桑德霍姆，日本的石川馨、田口玄一等，他们使人们对质量及质量管理有了更进一步的认识，对质量管理的发展和进步起了巨大的推动作用。这里简要介绍其中最重要的几位。

(1) W. A. 休哈特 (1891—1967 年)。沃特·阿曼德·休哈特 (W. A. Shewhart) 出生于美国伊利诺伊州的新坎顿，1917 年获得加州大学伯克利分校的物理学博士学位，1918—1924 年在西方电气公司 (Western Electric) 任工程师，1925—1956 年任贝尔实验室研究员，期间先后在伦敦大学、斯帝文理工学院、美国农业部研究生院和印度讲学。

休哈特博士基于对西方电气公司所制造产品的特性变异的关注和对抽样结果的研究，创立了统计过程控制 (SPC) 理论。1924 年 5 月 16 日他提出了世界上第一张控制图，1931 年发表经典著作《制成品质量的经济控制》，并将控制图应用在西方电气公司霍桑工厂的保险丝、加热控制和电站装置的生产上。

休哈特博士在研究中观测到，自然界以及工业产品中的所有事物都会发生变异，研究

这些变异并减少变异是质量改进的主要手段。虽然不存在完全相同的两件事物，但大量的观测会形成可预测的形态。他总结出两条重要的原理：（1）变异是不可避免的；（2）单一的观测几乎不能构成客观决策的依据。为了判断观测值的形态，可以将观测值以若干方式加以描绘：一种方式是将观测值绘制成直方图，这可以展示观测值的分布情况；另一种方式是按照观测值的顺序将之在图上打点，这样便形成线状图，这对于观察数据趋势和周期非常有用。休哈特主张用链状图来观测数据，他进一步指出存在着两类引起变异的原因：其一是偶然原因，是指只会引起数据很小波动的那些因素，其影响较小，但会大量聚集呈现某种形态，根据中心极限定理和观测经验推断它们通常近似服从正态分布；变异的另一种原因是系统性原因，是导致数据显著偏离偶然原因所形成分布形态的变异根源。休哈特控制图基于偶然原因所形成的正态分布的 3σ 原则建立了一组控制界限，任何落在这些界限之外的或者呈某种异常趋势的观测值都表明可能存在着系统性原因，由于观测数据是按所发生顺序描点，所以若存在趋势或异常形态很容易被观测出来。因此，控制图是对过程所产生的统计量的变异的一种图示描述，运用控制图进行生产过程监控，可以及早发现不合格的萌芽，并采取措施予以预防和控制。

休哈特的 PDCA（计划—执行—检查—处置）循环的观点被戴明等人广泛应用，至今仍然是质量管理最经典的持续改进管理模式和思维方式。

（2）W. E. 戴明（1900—1993 年）。威廉·爱德华兹·戴明（W. A. Deming）博士 1921 年获得怀俄明大学工程学士学位，1925 年在科罗拉多大学深造并获数学及物理学硕士学位，1928 年获耶鲁大学物理学博士学位。

1950 年开始，戴明应日本科学技术联盟邀请多次访问日本，讲授质量管理课程。戴明帮助听众理解统计质量控制（SQC）在制造业的重要性，对统计质量控制在日本的发展和深化起了巨大的作用。他强调，引起效率低下和不良质量的原因中有 85% 在于企业的管理系统而只有 15% 是由员工造成的。戴明质量战略的核心就是使用统计质量控制来识别变异的特殊原因和偶然原因。为了促进日本质量管理的发展，纪念戴明对日本的友好和贡献，日本于 1951 年设立了戴明奖。但直到 1980 年 6 月，美国国家广播公司（NBC）作了题为《如果日本可以，为什么我们不能？》的报道之后，戴明才在美国名声大震。

戴明的主要思想理念体现在他提出的质量管理 14 要点中：

- 1) 制定改进产品和服务的目标和实施的计划，致力于超过竞争对手。
- 2) 采用新的质量管理思想。
- 3) 停止依靠大量检验来提高质量。
- 4) 不要仅凭价格选择供应商，要以总成本最低为目标。
- 5) 发现问题并致力于改进工作体制。
- 6) 采用现代的在岗培训方法。
- 7) 提升领导能力，采用新的领导方式。
- 8) 消除员工的畏惧感。
- 9) 打破部门封锁，倡导产品设计、销售、生产等部门团队合作。
- 10) 消除那些要求员工做到零缺陷及高生产力水准的口号、劝诫及目标，低质量和低生产率是制度造成的而不是员工的问题。
- 11) 取消工作定额，代之以领导职能强化。

12) 消除各种影响员工为自己工作质量而自豪的障碍。

13) 设立生动活泼的教育和自我提高计划。

14) 建立使高层管理者能够推动每个员工按上述 13 条努力工作的机制。

(3) J. M. 朱兰 (1904—)。约瑟夫·M·朱兰 (Joseph M. Juran) 1924 年获得明尼苏达大学电子工程专业学士学位, 并就职于芝加哥西方电气公司霍桑工厂 (Hawthorne Works) 检验部。1926 年, 一个来自贝尔实验室的团队 (其中包括休哈特和道奇) 参观了该工厂, 并推行贝尔实验室开发的工具和方法。朱兰是被选中的 20 个受训者之一, 且作为最终入选的两个工程师之一, 到新创建的检验统计部工作。

1928 年, 朱兰撰写了他的第一部关于质量的著作, 一本名为《生产问题的统计方法应用》小册子。1937 年, 朱兰将帕累托原理概念化并应用于质量改进。1945 年, 朱兰离开华盛顿和西方电气公司, 作为独立的咨询师开展工作。1951 年, 《朱兰质量控制手册》出版, 为他赢得了国际威望。

他认为质量来源于顾客的要求, 在他的《朱兰质量手册》(第 5 版) 中对质量的定义是:

1) “质量”意味着能够满足顾客的需要从而使顾客感到满意的那些产品特性;

2) “质量”意味着免于不良, 即没有那些需要重复工作 (返工) 或会导致现场失效、顾客不满、顾客投诉等的差错。

朱兰认为要获得质量, 最好从建立组织的愿景以及方针和目标开始, 目标向成果的转化 (使质量得以实现) 是通过管理过程来进行的。他认为在质量管理活动中频繁应用的三个过程是: 质量策划、质量控制和质量改进, 即著名的朱兰质量管理三部曲。他将“人的因素”加入质量管理过程, 使质量管理超越了过去的统计学。

1954 年, 朱兰博士应日本科学技术联盟的邀请, 到日本为中高层管理人员讲授质量管理课程。1979 年, 朱兰建立了朱兰学院, 如今它已成为世界上领先的质量管理咨询公司。半个多世纪以来, 朱兰的质量管理理念和方法对整个世界质量管理学界产生了巨大的影响。

(4) 石川馨 (1915—1989 年)。石川馨 (Kaoru Ishikawa), 1939 年毕业于东京大学工程系, 1960 年获工程博士学位。他的《质量控制》一书获“戴明奖”、“日本 Keizai 新闻奖”和工业标准化奖。1971 年, 其质量控制教育项目获美国质量协会“格兰特奖章”。1968 年, 石川馨出版了《质量控制指南》。他是因果图的发明者, 质量管理小组 (QCC) 的奠基人之一, 是将国外先进质量管理理论和方法与日本实践相结合的专家。

石川馨认为, 全面质量管理 (TQC) 在日本就是全公司质量管理 (CWQC), 具体内容包

1) 所有部门都参加的质量管理, 即企业所有部门的人员都学习、参与质量管理;

2) 全员参加的质量管理, 即企业的经理、董事、部课长、职能人员、工班长、操作人员、推销人员等全体人员都参加, 并延伸到供应商、代理商;

3) 综合性质量管理, 即以质量管理为中心, 同时推进成本管理 (利润、价格管理)、数量管理 (产量、销量、存量)、交货期管理。

(5) 田口玄一 (1924—)。田口玄一 (Genichi Taguchi) 博士是数理统计应用技术方面的著名学者, “田口方法”的创始人, 戴明奖的获得者。

“田口方法”包含四大技术：线外质量控制、线内质量控制、计量管理技术和试验设计技术。“田口方法”提出了质量损失函数（QLF）的概念，把质量和经济两个范畴的概念统一起来，为质量波动的定量统计分析奠定了基础。他还提出了信噪比（SNR）的概念，为试验设计技术充实了新的内容，为解决产品或生产过程的动态性、稳健性以及测量误差等问题开辟了新途径。

“田口方法”提出了在新产品开发过程中进行三阶段设计的思想，即系统设计、参数设计、容差设计。也就是系统地考虑问题，选择参数间的最优水平组合，以及最恰当的容差。

“田口方法”既重视经济效益，又不等同于经济学；既运用数理统计方法，又不完全以统计观点来研究质量管理，他以工程学和技术观点研究质量管理理论和方法，具有通用性和边缘性的特点，因此又被称为“质量工程学”。“田口方法”在日本工业界得到了广泛应用，并带来了巨大的经济效益，目前世界上越来越多的国家在使用“田口方法”。尽管田口试验设计方法本身存在一些不足之处，但这丝毫不影响他成为20世纪著名的质量管理专家。

1.1.2 六西格玛管理的起源和发展

1. 六西格玛管理的起源

在戴明、朱兰的帮助下，以及费根堡姆 TQC 概念和理论的引导下，日本人结合其国情对质量管理进行了大量的实践和创新。20 世纪 70—80 年代，经过 20 多年的持续改进，日本产品的质量显著提高，占领了大部分美国市场，而美国企业的产品则逐步失去自己的市场，许多美国企业面临生死存亡的问题。

摩托罗拉（Motorola）公司同样面临着生死存亡的考验。20 世纪 70 年代初期，摩托罗拉已经成为全球无线通讯产品的领导者，并与得克萨斯仪器公司以及英特尔公司一起争夺半导体产品的最大销售商的位置。1974 年，8 个最大的半导体厂商有 5 个来自美国，3 个来自欧洲。但很快半导体市场的竞争变得异常激烈，仅仅在 5 年后的 1979 年，8 个最大的芯片生产商中就有 2 个来自日本。摩托罗拉在 1974 年将电视机业务卖给了日本松下，1980 年在日本竞争者面前又失去了音响市场。日本人也对摩托罗拉在美国的寻呼机市场的领导地位构成了威胁。于是摩托罗拉公司开始反思自己的问题，认识到提高产品质量实质上会降低成本，并不是当时绝大部分美国公司所认为的提高质量会多花美元。于是摩托罗拉公司决定认真采取质量战略，开始了其质量改进之路。在首席执行官鲍勃·高尔文（Bob Galvin）的领导下，一个特别工作组开始为摩托罗拉的创新和业务增长制定计划。这项工作因阿特·萨恩德赖（Art Sundry）而得到加速发展，因为他得到了摩托罗拉通讯产品顾客和用户的意见反馈，并大声疾呼“我们的质量糟糕透了”。高尔文的工作组和萨恩德赖的研究使一个四点计划产生并于 1980 年实施，其目的在于确保摩托罗拉在全球的领导地位。

（1）全球竞争力。通过与竞争对手进行水平对比，设计面向全球市场的产品，确保优胜地位。

（2）参与式管理。吸取全面质量管理之精华，将 QCC 的原则和方法引入摩托罗拉的企业文化，在全公司广泛推进顾客完全满意（total customer satisfaction, TCS）小组活动。



(3) 质量改进。将改进目标定为 5 年内改进 10 倍，将质量改进目标与所有管理人员的奖励计划挂钩——这个创意播下了六西格玛理念的火种。

(4) 摩托罗拉培训与教育中心。形成摩托罗拉大学的雏形，主要任务是通过培训使员工的能力满足质量流程与管理方式的巨大变化的需求。

经过几年的实践和酝酿，摩托罗拉在 1987 年全面推行六西格玛，点燃了六西格玛管理的火种。迈克尔·哈瑞 (Mikel Harry)、比尔·史密斯 (Bill Smith) 和理查德·施罗德 (Richard Schroeder) 等是六西格玛方法最重要的创始人。在当时的首席执行官鲍勃·高尔文的大力支持下，六西格玛在全公司范围内得到了广泛施行、推广。六西格玛产生的强大动力使得摩托罗拉制定了以前看上去几乎不可能实现的目标：20 世纪 80 年代早期的目标是每 5 年改进 10 倍，后来改为每 2 年改进 10 倍，到 1992 年产品和服务质量达到六西格玛质量水平。由于实施六西格玛，公司于 1988 年获得美国波多里奇国家质量奖，从开始实施的 1987—1997 年的 10 年间，销售额增长了 5 倍，利润每年增加 20%，通过实施六西格玛管理所带来的收益累计达 140 亿美元，股票价格平均每年上涨 21.3%，效果十分显著。

2. 六西格玛管理的推广

不论走到哪里，高尔文总要提到六西格玛的目标和成功的故事。在他的听众中就有拉里·博西迪 (Larry Bossidy)，原通用电气公司副总裁，通用电气首席执行官杰克·韦尔奇 (Jack Welch) 的助手与最要好的朋友。1991 年，博西迪结束了他在通用电气的任职，接手了刚刚由若干公司合并而成的联合信号公司 (Allied Signal)，出任 CEO。从通用电气这样一个“言出必践”的公司来到联合信号，博西迪吃惊地发现，表面上看起来，联合信号和通用电气有着相似的核心业务和管理流程，但在联合信号，这些流程大都只是一些空洞的形式，没有实际效果或效率很低。人们很少关注“提高生产力、扩大市场份额、提高产品质量”等具体问题。尽管有着很多聪明和勤奋的员工，但是在计划和行动之间存在很大的差距。更糟糕的是，企业的人员、战略和运行都采用了各自为政的方式，每个业务部门都强调自己的文化，不肯融入联合信号之中。博西迪面临的最重要的问题是如何将联合信号真正集成为一个整体，并且快速地改变公司的“执行”能力和“执行”文化。听了摩托罗拉的故事，他马上意识到，六西格玛正是他想要的可以将联合信号塑造成他心目中企业的方法。

1992 年，博西迪将六西格玛引入联合信号公司。“我们不但要给人们提出目标，还要提供合适的工具和方法”，联合信号的管理层这样说。人们从各地赶来，学习六西格玛方法。然后回到自己的岗位，将所学用于解决实际问题。事实上，对联合信号来说，不仅需要改进质量，更需要变革人们做事的方式。不论人们来自哪个公司，具有什么样的企业文化背景，在联合信号就要形成统一的语言和文化。为了使六西格玛更好地适合联合信号的目的，在博西迪的带领下，许多被称为“软工具”的内容补充到六西格玛方法中。其中包括大量的关于组织变革、领导力提升和变革企业文化方面的内容。

博西迪成功了。六西格玛改变了联合信号的经营与运作方式，公司的业绩出现了快速增长的势头，销售额和利润持续实现每年两位数的增长。公司的收益从 1991 年的 3.42 亿美元增长到 1997 年的 11.7 亿美元，在短短的 6 年内几乎翻了两番。而且连续 31 个季度保持每股利润 13% 以上的增长，公司的股价增长了 8 倍。正是因为博西迪和联合信号，华

尔街第一次听说了六西格玛。此后，得克萨斯仪器等一批公司相继引入了六西格玛，同样取得了成功。

美国通用电气公司在20世纪70—90年代制定的战略计划取得了成功，这使得通用电气适应了公司规模和经营多样化的发展，因而给公司带来了巨大利益。在这个时期，通用电气重点关注的是业务的高速增长、提高效率，而质量问题排在后面，产品质量大约为3~4个西格玛水平。随着业务的高速增长，质量问题越来越不容乐观，杰克·韦尔奇在其自传中回忆说：1995年4月在通用电气内部员工调查中发现，质量问题已经为许多员工所担忧。

1995年夏，杰克·韦尔奇见到了他的朋友博西迪，他们谈到了联合信号成功的原因。这次谈话给韦尔奇留下了深刻的印象。自上任以来，韦尔奇已经对通用电气进行了一系列大刀阔斧的变革，使通用电气在其所有业务领域几乎都是“数一数二”的。但韦尔奇希望通用电气这个百年老店能够永远成功。韦尔奇意识到，把通用电气公司打造成为六西格玛企业，将是他2001年退休时给通用电气留下的最宝贵的遗产。

在韦尔奇的带领下，通用电气公司开始了它的六西格玛之旅。但是，在韦尔奇眼中，此时的六西格玛已经不是彼时的六西格玛。它已不仅仅是关于统计学的，尽管它需要统计技术；它也不仅仅是关于改进质量的，尽管它最终使质量大幅度提升。韦尔奇要将六西格玛提升到打造通用电气公司核心竞争力的战略层次。在1996年1月通用电气高层博卡会议上，他将六西格玛称为通用电气前所未有的最雄心勃勃的工作，他说：“质量可以真正地使通用电气从最了不起的公司之一这个位置，上升到全球商界绝对最了不起的公司。这个房间的所有人都必须抓质量，这个问题是没有投机可言的，摩托罗拉用10年时间办到的，我们必须在5年内实现——不是通过走捷径，而是通过学习他人。”

韦尔奇认为六西格玛的含义远远超出了质量和统计学，它通过提供对付难题的方法，驱使领导层把工作做得更好，六西格玛的核心是使一个公司彻底转变。在他看来，六西格玛“是通用电气公司从来没有经历过的最重要的发展战略”。1995年底，随着200个六西格玛项目的实施，这种管理方法在通用电气大规模地开展起来。1997年，这样的项目超过了6000个。根据通用电气2000年度的报告：1999年通用电气公司的利润为107亿美元，比1998年增长了15%。其中，实施六西格玛获得的收益就达到了15亿美元。到21世纪初，这个数字达到了50亿美元。而2000年，在通用电气获得绿带认证的员工已达到90%以上，公司中层以上的经理中有过六西格玛黑带经历的人员已达15%。

在通用电气，六西格玛方法演变为一个管理系统。六西格玛被作为公司四大战略（全球化、服务、六西格玛和电子商务）之一来实施；公司建立了从“倡导者”（champion）、“资深黑带”（master black belt, MBB）、“黑带”（black belt, BB）到“绿带”（green belt, GB）的六西格玛组织结构；六西格玛实施结果与管理人员晋升和奖金紧密结合。

韦尔奇领导通用电气公司的20年时间里，通用电气的股票市值增长了30多倍，达到了4500亿美元，排名从世界第十位提升到了第二位，在他的带领下这个“百年老店”焕发了前所未有的青春活力。

在通用电气公司应用六西格玛取得巨大成功之后，六西格玛为全世界企业，特别是许多《财富》500强企业所关注、认识并接受。很多企业发现六西格玛同样可以对自己产生深远而重大的影响，它们也开始大力推行六西格玛，包括福特、卡特波勒、陶氏、杜邦、



ABB、3M、东芝、惠尔浦、三星、LG、西门子、爱立信、洛克西德-马丁、庞巴迪、花旗银行、J. P. 摩根、美国运通、英特尔、微软等跨国公司。今天，六西格玛的应用已经从摩托罗拉、通用电气走向了全世界：从西方走向了东方，从世界 500 强跨国公司走向了普通企业乃至中小企业，从电子、机械、化工、冶金等制造业走向了银行、保险、航空、电子商务等服务业。

3. 六西格玛管理的新发展

六西格玛诞生于全面质量管理蓬勃发展的 20 世纪 80 年代中期，是对全面质量管理特别是质量改进理论的继承和新发展。20 多年来，六西格玛已经走过了很长的发展路程。作为一种降低缺陷的方法，它在实践中得到了不断的充实和发展，六西格玛不再仅是一种质量改进的方法，而是已经发展成为可以使企业保持持续改进、增强综合领导能力、不断提高顾客满意度及经营绩效并带来巨大利润的一整套管理理念和系统方法。近年来，摩托罗拉提出了“新六西格玛”方法——一种由沟通、培训、领导艺术、团队合作、度量和以顾客为中心等价值观驱动的变革方法，旨在提高企业竞争力和变革企业文化。

传统的六西格玛被广泛地运用于改善产品的质量，尤其适用于生产制造业，是以降低缺陷和减少变异为核心的，而新六西格玛是一个领导力管理程序，是关于总体业务改进的方法，它解决了管理人员所面临的两难问题：一方面要通过快速的业务改进项目达到短期的财务目标；另一方面还要在关键人才和核心流程方面为未来的发展积蓄能力。

将平衡计分卡、业务流程再造、高效率团队、对核心业务流程进行持续不断的监控等工具和六西格玛改善流程整合起来，新六西格玛方法提供了一套实用的方法和切实可行的工具，以便管理人员在业务实施过程中达到平衡。表 1—1 列出的是新六西格玛方法的领导力原则。

表 1—1 新六西格玛方法的领导力原则

核心领导力原则	解 释
整合 (align)	<ul style="list-style-type: none">●用卓越绩效模式（以波多里奇质量奖标准为依据），将顾客的需求与业务战略以及核心业务流程相结合。●确定战略实施目标，制定具有挑战性的具体目标和恰当的衡量标准。目的是提供持续不断的、可衡量的、最基本的经营结果要求，以便推动经营总目标的实现。
调动 (mobilize)	<ul style="list-style-type: none">●向团队授权，通过高层经理人员选择的项目，利用项目管理方法和六西格玛方法达到改进的目标。●用项目立项表、成功准则和严格的评估等方法将团队组织起来。●向团队提供及时的培训服务并鼓励他们将其付诸行动。
加速 (accelerate)	<ul style="list-style-type: none">●采用学以致用 (action learning) 的方法，将系统的培训与实际的项目工作和指导结合起来，迅速消除由学习到实践的差距。让员工行动起来的激励性措施很容易失去效果，但是，为了及时达到项目的目标，激励性措施又是必不可少的。
控制 (govern)	<ul style="list-style-type: none">●对绩效计分卡模型进行控制，促进战略的实施。系统的流程审核包括对结果控制卡 (dashboard) 的审核，以及在必要时审核流程和项目的细节。当领导们分享最佳实践经验时，障碍就可以被消除。

1.2 六西格玛的概念和作用

1.2.1 六西格玛的概念

六西格玛是一套系统的、集成的业务改进方法体系，是旨在持续改进企业业务流程，实现客户满意的管理方法。它通过系统地、集成地采用业务改进流程，实现无缺陷的过程设计（design for six sigma, DFSS），并对现有过程进行过程界定（define）、测量（measure）、分析（analyze）、改进（improve）、控制（control）——简称 DMAIC 流程，消除过程缺陷和无价值作业，从而提高质量和服务、降低成本、缩短运转周期，达到客户完全满意，增强企业竞争力。

1. 六西格玛的统计含义

σ 是一个希腊字母，读作“西格玛”，在数理统计中表示“标准差”，是用来表征任意一组数据或过程输出结果的离散程度的指标，是一种评估产品和生产过程特性波动大小的参数。

西格玛质量水平则是将过程输出的平均值、标准差与质量要求的目标值、规格限联系起来进行比较，是对过程满足质量要求能力的一种度量。西格玛水平越高，过程满足质量要求的能力就越强；反之，西格玛水平越低，过程满足质量要求的能力就越低。六西格玛质量水平意味着百万出错机会（DPMO）中不超过 3.4 个缺陷。

如图 1—2 所示的正态分布中，当上下规格限之差为 12σ ($\pm 6\sigma$)，且过程无漂移，即实际分布中心与规格中心重合时，低于下规格限 LSL 和高于上规格限 USL 的面积（概率）均为 0.001ppm，总缺陷概率为十亿分之二。图中 μ 为正态分布的中心值， M 为规格中心， σ 为标准差。

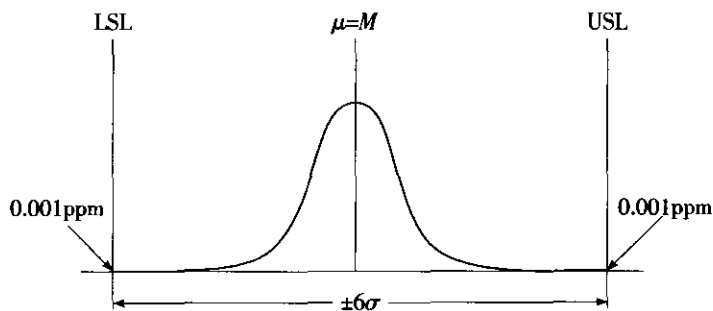


图 1—2 过程输出特性正态分布图（无漂移）

但实际上，过程输出质量特性的分布中心与规格中心重合的可能性很小，对于典型的制造过程，由于影响过程输出的基本质量因素（人、机、料、法、环、测）的动态变化，过程输出的均值出现漂移是正常的，如图 1—3 所示。在计算过程长期运行中出现缺陷的概率时，一般考虑将上述正态分布的中心向左或向右偏移 1.5σ ，此时一侧的缺陷为 3.4ppm，另一侧因数量级极小可忽略不计，总缺陷概率为百万分之 3.4，如图 1—4 所示。

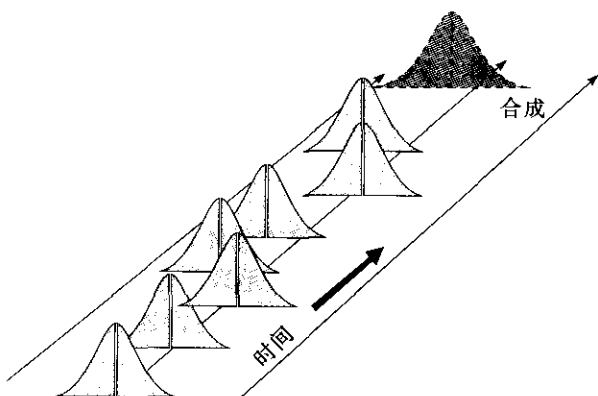
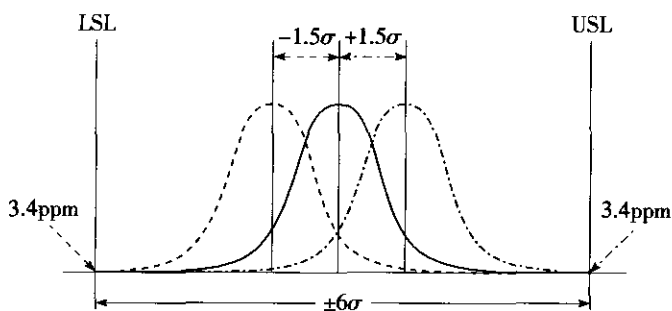


图 1—3 长期过程绩效——各种短期过程能力的合成图

图 1—4 过程输出特性正态分布图 ($\pm 1.5\sigma$ 漂移)

因此通常所说的六西格玛质量水平代表 3.4DPMO，是考虑了过程在长期运行中，分布中心相对规格中心漂移 $\pm 1.5\sigma$ 后出现缺陷的概率。表 1—2 显示了不同西格玛水平时的 DPMO。

表 1—2 西格玛水平与合格率及缺陷数之间的关系 (分布中心漂移 $\pm 1.5\sigma$)

西格玛水平	百万机会缺陷数 (DPMO)
1.0	690 000
2.0	308 700
3.0	66 810
4.0	6 210
5.0	233
6.0	3.4

2. 六西格玛的管理含义

从六西格玛的统计含义我们知道： σ 是一希腊字母，在数理统计中表示“标准差”；“西格玛水平”用于衡量产品、服务和过程的缺陷水平；六西格玛水平是指 100 万次出错机会中有 3.4 个缺陷的缺陷水平。

今天，六西格玛已远远超出其统计含义，成为一种客户驱动下的持续改进的管理模式。企业推行六西格玛，也不仅仅把六西格玛 (3.4DPMO) 作为一种目标或指标，六西

格玛还有更为深刻的多重管理含义,包括:

(1) 六西格玛目标。使过程趋于目标值并减少波动,追求零缺陷,追求完美。

实际上,实施六西格玛并不一定要达到六西格玛水平的质量,而在于对过程进行突破性的改进和创新。图1—5显示了“飞镖”过程:精而不准——波动小,但均值漂移大造成了较多的缺陷;准而不精——均值漂移小,但波动大也造成了较多的缺陷;又精又准——波动和均值漂移均极小,趋于零缺陷。六西格玛目标就是“又精又准”:使过程趋于目标值并减少波动,追求零缺陷,追求完美。

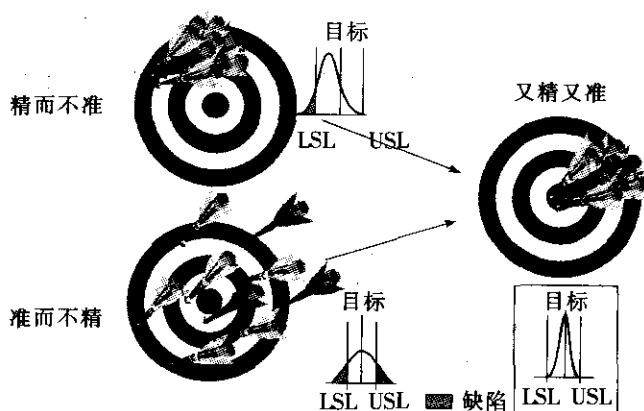


图1—5 六西格玛目标:又精又准

(2) 六西格玛方法。六西格玛在方法层面强调系统集成与创新,这里特别要指出,六西格玛绝非仅仅应用统计技术解决问题,它是一套系统的业务改进方法体系,其工具和方法包括现代质量管理技术、应用统计技术、工业工程和其他现代管理技术、信息技术等。

(3) 六西格玛文化和战略。将六西格玛价值观和改进方法融入企业文化,列为企业的战略,提升企业战略执行力,促进组织完成其使命,实现其愿景和战略目标。

1.2.2 六西格玛管理的作用

实施六西格玛是“一箭多雕”的、多赢的战略选择。实施六西格玛的好处多种多样,包括减少成本、提高生产力、增加市场份额、留住顾客、减少运作周期时间、减少错误、改变公司文化、改进产品和服务等。本节将实施六西格玛的作用归纳为以下四个方面:

- 问题解决和成本降低。解决困扰公司的重要而复杂的难题,降低不良质量成本。
- 文化变革。建立持续改进和创新的企业文化,消除沟通壁垒。
- 战略实施。全面提升公司的核心竞争力和经营管理成熟度。
- 员工发展。培养下一代领导者,促进员工职业发展。

1. 解决困扰公司的重要而复杂的难题,降低不良质量成本

一个有志于成功、有志于追求卓越的公司首先要明确自己的定位、未来的发展愿景以及行为准则,即确立公司的使命、愿景和核心价值观。其次要基于使命、愿景和核心价值观,确定自己的战略目标和战略方案。其三要设立能驱动战略目标实现、监测战略规划的关键绩效指标(KPI),并将关键绩效指标横向分解到相关职能部门和过程,纵向层层分



解到团队和员工。

这些关键绩效指标对公司应对动态的竞争环境、取得长期的成功至关重要。然而，要达成关键绩效指标常常是富有挑战性的，有些还涉及困扰公司多年、导致外部及内部顾客不满意的复杂问题，包括：

- 产品和服务质量问题；
- 运营成本问题；
- 生产率、流程周期问题；
- 市场和顾客流失问题；
- 环境和安全问题。

这些问题如果不彻底加以解决，势必会影响公司战略的实现。而要解决这些难题，必须由高层领导自上而下地推进，必须由一些具有很强的问题解决能力和技巧的人员，通过科学的流程，应用精确的数理统计技术和其他管理工具来解决。

这些问题的解决，要花费更多的投入吗？当谈到六西格玛管理时，常常会听到人们这样说：“要达到那么高的质量，企业是否要花费巨资？”“我们的质量水平与六西格玛差得太多了。要实现六西格玛，公司就得花费大量资金来更新设备或改进技术，这样做企业是否划算？”其实，这是对六西格玛管理的一种误解。成功的六西格玛管理为企业带来的是实实在在的经济效益。这一点已为摩托罗拉、联合信号、通用电气等“世界级”企业的实践所证明。

下面，我们从不良质量成本（cost of poor quality, COPQ）的角度谈谈为什么六西格玛管理能给企业带来如此显著的回报。大多数公司处于 3~4 西格玛水平，其不良质量成本可占到销售额的 20%~40%，而六西格玛企业的不良质量成本仅占销售额的 1%~5%。通用电气认为，公司从 3~4 西格玛水平提高到六西格玛水平，其减少的不良质量成本每年可达 80 亿美元~120 亿美元。

这些成本损失来自哪里呢？不良质量成本可以分为直观的和隐含的两大类，就像冰山一样，露在水面上的是我们通常统计的那些由于产品或服务不良而造成的损失，也就是传统质量成本统计中通常作为内部与外部故障成本所统计的部分。对于大多数企业来说，这些成本损失占销售额的 4%~5%。但冰山还有隐藏在海面下的部分，这是我们通常不去统计或不为人们重视但又实实在在存在于企业中的成本损失。这些“隐藏成本”损失要比露出的部分大得多，可高达销售额的 16%~35%。这些直观的和隐藏的不良质量成本已经远远超过了销售利润。根据美国著名管理咨询公司 Imberman & Deforest 在 20 世纪 90 年代初所作的调查，一般企业的平均利润水平仅占销售额的 1%~4%。

六西格玛管理就是要通过业务过程的改进和优化，降低这些直观的和隐含的成本损失，把“冰山”变为“金山”，使“更高的质量、更低的成本、更短的开发与生产周期、更好地满足顾客的要求”变为现实。

通用电气公司将六西格玛管理应用于企业经营管理活动的各个方面，并取得了巨大的收益。比如：人力资源部用六西格玛管理方法改进人员的配置；财务部用其提高付款的及时性与准确性；工程部改进产品设计的可靠性和降低图纸缺陷；制造部降低制造缺陷并且缩短生产周期；服务部缩短服务响应时间和维修时间；信息部提高信息系统的效率等。其中，一个六西格玛项目小组完成了改进产品交付周期的项目。他们了解到，顾客希望其产

品交付期不超过 10 天，而实际上他们的产品交付期平均为 33 天。通过运用六西格玛方法，他们将交付期缩短为平均 2.3 天，仅此一项每年为企业创造直接的经济效益 50 万美元。通用电气通讯部门通过六西格玛管理项目，将其拥有的 12 颗卫星的利用率从 63% 提高到 97%，每年增加收入 130 万美元。在通用电气公司中，这样的案例数不胜数。这也就是六西格玛为通用电气公司带来了如此之高的收益和增长速度的一个重要原因。

统计资料表明：如果一个 3 西格玛水平的企业组织其所有资源改进过程，大约每年可以提高一个西格玛水平，直到达到 4.7 西格玛水平，其投入主要是培训及咨询费用，而无须大的资本投入。这期间，利润率的提高十分显著。而当达到 4.8 西格玛水平时，要再提高西格玛水平，则需要对过程进行重新设计，增加资金投入，但此时产品/服务竞争力、市场占有率将得到明显的提高。对一个 3 西格玛水平的企业来说，提高一个西格玛水平可获得下述收益：

- 利润率增长 20%；
- 产出能力提高 12%~18%；
- 劳动力减少 12%；
- 资本投入减少 10%~30%。

美国质量协会指出：“六西格玛方法要求在一个合理的时间内得到显著的回报”是其成功的十大因素之一。因此，成功的六西格玛管理为企业带来的将是实实在在的经济效益。

2. 建立持续改进和创新的企业文化，消除沟通壁垒

对企业文化完整的定义是：企业在长期的生存和发展中形成的，为本企业所特有的，且为企业多数成员共同遵循的宗旨（使命）、最高目标（愿景）、价值标准、基本信念和行为规范（价值观）等的总和及其在企业活动中的反映。

企业文化具有精神层、制度层和物质层三个层次的结构，如图 1—6 所示。

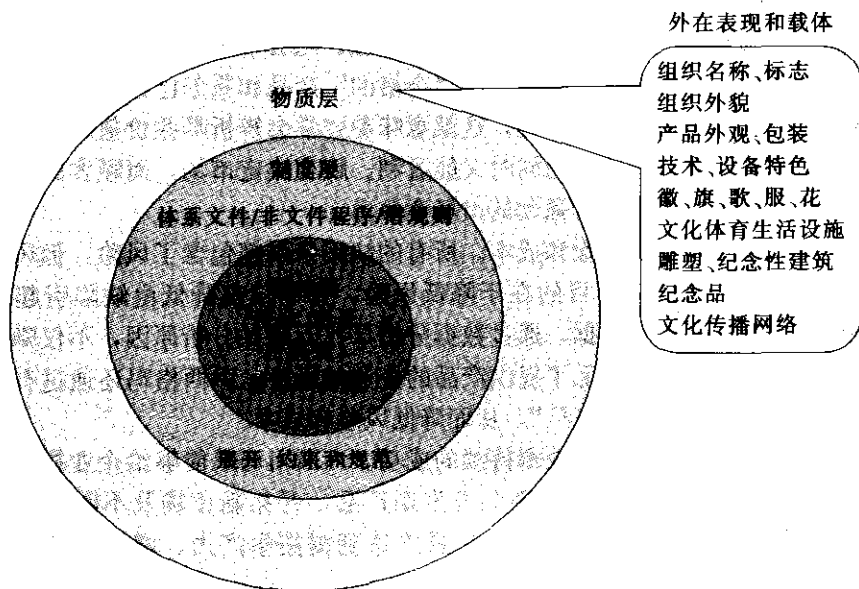


图 1—6 企业文化的三重结构



通俗地说,企业文化是企业中独特的做事的方式方法。影响企业文化的主要因素是企业高层领导。不论企业领导个人的言行举止有多好或多糟,都会被企业员工放大,形成企业整体的工作风气和行为准则。

六西格玛不仅仅是一个解决问题的技术方法,同时也是一种处世和处事的哲学,它的标准就是“完美”,在更广泛的意义上讲是一种文化,一种持续改进和创新的文化。它在不断地转变着人们的工作方式。随着六西格玛的推进,企业文化也会得到不断的完善,会形成一种同心协力克服障碍,人人积极关注、参与改进和创新的文化氛围。

企业的核心业务流程就像是一条横向流淌的河,而现有的职能式组织结构犹如纵向的大坝,部门之间由于业务目标的差异导致了沟通壁垒,这种沟通壁垒又大大地影响了流程运行的效率和效果。实施六西格玛,还有助于消除沟通壁垒,增进无边界合作的文化。六西格玛扩展了跨职能合作的机会,消除了部门间及上下级间的障碍,促进了组织内部横向和纵向的合作。六西格玛项目往往是跨职能的。在六西格玛跨职能团队中,团队成员有着共同的团队使命和目标,并且能意识到自己的职责与企业的使命、愿景、战略目标和关键绩效指标的关系,意识到业务流程各部分的相互依赖性,能够创造出一种能真正支持团队合作的管理结构和环境。而这种无边界合作的“纽带”正是有着强大使命感的黑带,黑带要想成功就必须打破部门之间的障碍。

六西格玛的“无边界”文化是通用电气成功的秘诀之一。韦尔奇致力于消除部门及上下级间的障碍,促进组织内部横向和纵向的合作,改善了过去仅仅是由于彼此间的隔阂和企业内部部门间的竞争而损失大量金钱的状况。这种做法改进了企业内部的合作,使企业获得了许多受益机会。

3. 全面提升公司的核心竞争力和经营管理成熟度

六西格玛还可以是一种战略,一种全面提升公司竞争力和经营管理成熟度的战略。六西格玛作为公司战略,主要通过以下作用来实现公司的战略目标。

(1) 提高顾客满意度,留住顾客,增加市场份额。随着各个行业竞争的日趋激烈以及顾客要求的不断提高,仅提供“好的”或“合格的”产品和服务已经不能保证企业的成功。六西格玛以顾客为真正的关注中心,这就意味着要学会辨析哪些价值对顾客是有意义的及计划如何给顾客提供这些价值的同时又能盈利。顾客决定市场,而顾客的选择标准是产品质量,因此六西格玛质量具有强劲的市场竞争力。

(2) 减少缺陷错误,降低风险和成本。所有的缺陷错误都包含了风险,但风险不一定体现在缺陷错误上。六西格玛的目的在于降低风险,并非仅仅降低由缺陷引起的不合格率。六西格玛强调基于事实和数据,通过数据分析寻找问题并分析原因,不仅降低了顾客购买产品时要承担的风险,还降低了提供产品的企业的风险。六西格玛还通过有效的方法查找不良质量成本,消除“隐蔽工厂”,从而降低风险和成本。

(3) 改进产品及服务,使企业获得持续的成功。六西格玛是能够给企业提供不断创新的技巧和文化的的方法,通过不断地革新和开发新产品、开拓新市场及不断地改进组织结构,使企业价值最大化,提高顾客满意度,最终达到提高生产力、增加市场竞争力的目的,从而使企业得到持续的发展。

(4) 加快改进的速度。摩托罗拉将1987—1991年“4年改进100倍”的成效归功于六西格玛的实施。顾客的需求是动态变化的,随着现代社会生活节奏的加快,人们对产品性

能改进的需求越来越强烈,改进速度最快的企业将在竞争中得到持续发展。六西格玛通过借鉴其他学科的工具和思维,可以帮助企业在改进绩效的同时加速改进流程。

4. 培养下一代领导者,促进员工职业发展

通用电气将六西格玛管理与人力资源管理密切结合。1997年3月22日,韦尔奇专门向通用电气全球的管理人员发送了一份传真,明确规定管理人员的提升将直接与六西格玛挂钩;5月22日,韦尔奇与他的两位副董事长联合签发了一条关于将六西格玛培训结果与晋升机会相联系的命令;1998年1月1日,通用电气新版本的管理人员晋升制度正式生效:从即日起,基层管理人员必须事先通过六西格玛绿带或黑带的培训,才能够获得晋升中级管理职位或高级管理职位的资格。从此,六西格玛管理成为通用电气未来领导人的摇篮,一大批训练有素、擅长基于数据与事实解决问题和进行决策的人员走上了通用电气各层次的管理岗位,为通用电气多年来的强劲发展奠定了良好的人力资源基础。

六西格玛认为员工是企业获取竞争优势的根本,六西格玛管理的推行能够有效地提高员工的素质,促进员工的职业发展。“学习型组织”这个很多人感兴趣但很难付诸行动的概念产生于20世纪90年代,联合信号的领导层对此评价说:“每个人都在谈论学习,但很少有人把它贯穿于大多数员工的日常生活中。”六西格玛是一种可以在组织内增强和加速新思维的发展和分享的方法。

六西格玛带给员工的是解决问题的方法。员工不仅仅通过培训学到知识,而且要将所学知识应用到实践中,通过实践完全掌握解决问题的科学方法,为组织带来丰厚的回报,同时提高员工的个人能力。通过员工个人行为的改变,促进学习和相互指导,提高人员素质,进而改变企业整体的文化氛围,使企业成为一个学习型组织。

六西格玛管理为企业实施持续、突破性的改进和创新提供了所必需的管理工具和操作技巧,也为企业培养了具备组织能力、激励能力、项目管理技术和数理统计诊断能力的领导者,这些人是企业推进变革和竞争的核心力量,使企业降低质量缺陷和服务偏差并保持持久性的效益,促进快速实现突破性绩效,帮助企业实现战略目标。

1.3 六西格玛管理的领导、文化和战略

1.3.1 高层领导在六西格玛管理中的作用

1. 高层领导在企业发展中的作用

一个企业的成功固然需要全体成员的共同努力,但不可否认,高层领导才是决定其兴衰成败的关键,古今中外概莫能外。

如果把企业比作一艘舰船,高层领导就是舵手,要领导着组织驶向一个光辉的彼岸。高层领导最基本的职能定位和作用就是建立一个企业未来的发展方向和绩效期望,即确定企业的使命、愿景和核心价值观,在企业中营造一种氛围,一种促使人们为了实现目标而全力以赴的企业文化氛围。

使命(mission)是指企业的角色、任务或总体功能,反映了一个企业之所以存在的理由或价值,以及企业的灵魂之所在。任何企业的存在都有其目的或使命,如工商企业的



共同使命是向社会提供有经济价值的产品或服务。任何一个企业只有明确自己独特的使命，才能着力“做正确的事”。具体企业的使命是对其存在的规定，应有其独特性，决定了它之所以是它而不是任何别的机构。惠普公司创始人之一戴维·帕卡特在1960年给惠普员工所做的演讲中指出，企业的使命触及的是一种除了赚钱之外的更深层次的公司存在理由，是做一些有价值的事。企业是为了包括股东、顾客、员工、供方和合作伙伴、社会在内的五大利益相关方而存在的，因此从宏观的角度讲，企业应当为所有的利益相关方创造平衡的价值。

愿景（vision）是指企业所渴望的未来图景和境界，是一个企业的整体发展方向和所要追求的目标。高层领导应当具备远见卓识，经过深思熟虑、周密制定一个现实、可信、诱人、远大的，需要花五年甚至十几年来实现的远景目标，并向人们清晰明确地指出，这种目标建立在当前条件基础上，只要经过努力就会实现。一旦人们有效地确定和实施这种愿景，就会产生巨大动力，并通过凝聚各方技能、才干和资源而推动人们奔向未来。愿景应当是企业全体成员的共同心愿，也应反映各利益相关方对企业的绩效期望。

价值观（values）是指期望企业及其员工如何运作的指导原则和/或行为准则。价值观反映和增强企业所渴望的文化。价值观以适当的方式支持和指引每一位员工做决定，帮助企业完成其使命，达成其愿景。价值观是一个企业所拥护和信奉的东西，是一个企业最重要的和最永恒的信条，是一小部分不随时间和环境的变化而改变的原则，应当经得住时间的考验，即便因拥有某一价值观而遭受失败也不会改变它。它有别于企业的经营战略和经营活动等应当随着经营环境的变化而改变的东西。价值观决定了企业对于好与坏、对与错、赞赏或是不屑等问题的判断。价值观不是一种功利性的选择，而是一个企业的信仰。价值观必须以高层领导的行动和行为作为支持，否则它的发布只会在企业中造成人们的冷嘲热讽。成功的价值观应当体现于企业全体员工的言行举止，融入企业的骨骼、血液和经络系统，成为企业成功的遗传基因密码，并通过价值链影响着企业的相关方。

当高层领导确立了企业的使命、愿景和核心价值观，为企业的未来建立了发展方向和宏伟愿景之后，最重要的就是要通过战略策划，勾画出通往愿景的路径，将愿景转化为可行动的规划。

2. 高层领导在六西格玛推进过程中的承诺和关键角色

六西格玛不仅仅是一种改进方法，还是一种文化和战略。企业实施六西格玛，可以促进企业文化的变革和战略目标的实现，最终促进企业使命的实现和愿景的达成。而企业领导在六西格玛管理中扮演着关键的角色。实施六西格玛的企业以及六西格玛专家经过多年的实践，在总结六西格玛成功实施的关键要素时，不约而同地把领导层的支持和参与作为成功的第一关键要素。

成功推行六西格玛管理并获得丰硕成果的企业都拥有来自高层的高度重视与卓越领导，拥有一批精力充沛、直言不讳、知识丰富的管理者，更重要的是他们能参与到六西格玛管理中。他们的管理风格可能各有千秋，诸如通用电气公司杰克·韦尔奇的活力，联合信号公司拉里·博西迪的严谨，摩托罗拉公司鲍勃·高尔文的无所不在，但重要的是他们都认识到了支持和积极参与是六西格玛管理成功的关键。六西格玛管理是自上而下推行的，因此六西格玛的实施应始终取决于企业的最上层，因为六西格玛管理最终要变革企业的文化，因此，必须获得高层执行领导的认同。高层领导要指导制定六西格玛推行的计

划,选择倡导者,分派使用资源,安排政策支持六西格玛行动,并把六西格玛管理与企业不断发展的战略目标相结合。

具体来讲,最高管理层的作用包括:

(1) 制定2~5年的六西格玛战略目标。这些目标必须针对公司具体情况而制定,并且由公司战略来驱动。这里的关键是,这些目标必须反映出只有从六西格玛中才能获得的收益,而不是从公司现有的体系中即可以安全获得的收获。下面是几个典型的目标:

- 将毛利润提高5%~10%;
- 将投资回报率增加5%~15%;
- 每年收入增加10%;
- 削减20%的生产管理费用和质量成本;
- 市场份额每年增长4%。

这些目标看似简单,但会给所有公司带来巨大的挑战。这样做的目的就是清楚地扩展企业的思想,并清晰地将目标和财务绩效的关键驱动力连接起来。

(2) 授权一个推进小组。六西格玛开始实施,会影响到企业的方方面面。领导层的介入就是承担这样的责任:定义并管理六西格玛在企业中的形象,以及如何将其要求和利益传达给他人。在大多数情况下,最高管理者将这些责任授权给推进小组,而这个小组主要负责制定详细的实施计划并实施。

推进的整体目标就是确立目标和预算以及制定政策和基础架构,以确保与最高管理层的战略相吻合。根据公司规模大小以及整个企业的复杂程度,推进小组需要不断对计划进行修改,在2个星期至3个月间,推进小组应该能够初步制定一个推进方案。

(3) 制定推进方案。推进小组由最高管理层负责管理(一般由倡导者担任组长),其职责包括:

- 确定企业现有绩效和预期绩效之间的差距;
- 判定如何实施六西格玛才能缩短这一差距;
- 为实施六西格玛制定一个方案。

最初的差距分析可以做得很复杂,也可以做得很简单,这取决于企业的需要。在分析之前应获得现有绩效的可靠数据,尽管得到可靠的数据要花费一些力气。预期绩效要从长期目标中得出。明确现有绩效与预期绩效之间的差别很重要,这使企业领导层有足够的勇气认识到进行变革的必要性。

一旦推进小组完成了差距分析,就要制定一个初步的方案,方案应该包括:

- 六西格玛组织结构和人员需求;
- 实施目标和计划(启动的时机、所需的人员数量和培训类型等);
- 财务衡量标准及目标(成本和利润)。

项目的初步计划将确保高层领导理解六西格玛,以及它对公司产生的经营和财务利益。此时需要有足够深度的分析,从而使高层领导和董事会理解必要的投资和涉及的风险。

(4) 领导层亲身参与。六西格玛管理推行成功最关键的因素是领导重视。领导重视不仅体现在投入的资源上,还表现在其亲身参与上,如以各种行动带头营造六西格玛氛围,包括通过邮件、讲话、网络等各种方式表达自己对六西格玛的态度。有些企业,高层领导



甚至亲自参与或指导六西格玛项目。

1.3.2 六西格玛价值观与企业文化

1. 六西格玛价值观

六西格玛价值观亦即六西格玛的理念、哲学,是企业推进六西格玛管理的指导原则和行为准则。经过20年的实践和提炼,形成了以下六个方面的六西格玛价值观。

(1) 以顾客为中心。“以顾客为中心”不但是六西格玛管理最基本的价值观,也是现代企业管理理论和实践的基本原则。

企业依存于顾客,获得高的顾客满意度和忠诚度是企业所追求的主要目标。然而顾客只有在其需求和期望得到充分理解并获得满足后,才会满意和忠诚。企业应当深入了解顾客当前和未来的要求和期望,并关注其动态变化,以此驱动质量改进,消除缺陷,为顾客解决问题,满足顾客要求并争取超越顾客期望,从而建立良好的顾客关系,不断增进顾客满意和忠诚。

在六西格玛管理中,以顾客为中心是最优先的事,强调“倾听顾客的声音”(voice of customer, VOC)。六西格玛管理从倾听顾客的声音开始,基于顾客要求驱动改进和评价改进的效果,即一切以顾客满意和为顾客创造价值为中心。

(2) 基于数据和事实的管理。“基于数据和事实的管理”是现代管理与传统经验管理的分水岭。

越来越多的企业开始重视信息系统建设和知识管理,但是它们做出的许多经营决策仍然是基于感觉和经验,或是一些自以为是的观点和假设。六西格玛管理把“基于数据和事实的管理”的理念和实践提到了一个更高的层次。六西格玛管理最大特点之一就是强调一切用数据和事实说话,一开始就界定和测量过程输出关键绩效指标及其基线值,然后应用统计方法进行数据的分析和探测,确定显著影响关键绩效指标的过程因素,并通过改进获得优化的结果。

(3) 聚焦于过程改进。理解过程,精研过程,“聚焦于过程改进”,为顾客创造价值,是建立竞争优势的有效方法。

在六西格玛管理中,过程是采取改进行动的主要对象。设计产品和服务,测量绩效并进行分析、改进和控制,甚至经营企业等,都是通过过程而进行的。

过程在六西格玛中是成功的关键。一切活动都是过程,所有的过程都有变异,六西格玛帮助我们有效减少过程的变异。

(4) 有预见的积极管理。“积极”是指主动地在事情发生之前进行管理,而不是被动地处理那些令人忙乱的危机,或称为“救火”。“有预见的积极管理”意味着我们应当关注那些常被忽略的业务运作,并养成习惯:确定远大的目标并且经常加以检视;确定清晰的工作优先次序;注重预防问题而不是疲于处理已发生的危机;经常质疑我们做事的目的,而非不加分析地维持现状。

六西格玛管理包括一系列工具和实践经验,它用动态的、即时反应的、有预见的、积极的管理方式取代那些被动的习惯,促使企业在当今追求几乎完美的质量水平而不容出错的竞争环境下,能够快速向前发展。

(5) 无边界合作。“无边界合作”是企业消除职能之间、层级之间乃至合作伙伴之间

的沟通壁垒,面向业务流程,提高过程运作效率和效果的关键。

基于跨职能团队改进的六西格玛管理扩展了“无边界合作”的机会,而联结这种无边界合作的“纽带”就是那些有着强烈使命感的黑带。由于黑带是项目改进团队的负责人,而黑带项目往往是跨部门的,要想获得成功就必须由黑带率领他的团队打破部门间的障碍,通过“无边界合作”完成六西格玛项目。

当人们确认了如何使自己的职责与企业的愿景相适应时,就会意识到并且能够衡量出工作流程各部分的相互依赖性。在六西格玛管理中无边界合作并不意味着无条件的个人牺牲,而是要在为顾客创造价值的基础上,使各利益相关方同时受益。由于六西格玛管理是建立在广泛沟通基础上的,因此六西格玛管理能够营造一种真正支持团队合作的管理结构 and 环境。

(6) 追求完美,容忍失败。为什么在“追求完美”的同时还要“容忍失败”?二者看上去似乎有些矛盾。从本质上讲,这两方面是互补的。作为一个以追求卓越为目标的管理方法,六西格玛管理为企业提供了一个近乎完美的努力方向,不存在不执行新方法、贯彻新理念就能实施六西格玛管理的企业,而这样做总会带来风险。在推行六西格玛管理的过程中,可能会遇到挫折和失败,企业应建立鼓励创新和变革、容忍失败的文化环境,以积极应对挑战的心态,敢于面对失败并从中汲取经验教训,为成功奠定基础。

2. 六西格玛价值观的融入与企业文化变革

企业文化最核心的部分是其价值观。将六西格玛价值观融入企业原有的价值观中,就是一场企业文化的变革。

高层领导具有确立价值观、营造企业文化的职责。当企业推行六西格玛时,高层领导应当将六西格玛价值观融入企业原有的价值观,变革企业文化和基因密码,使员工的信念、态度和期望与六西格玛管理同步;高层领导要对六西格玛的实施高度负责,并将这种精神渗透到整个企业中去;高层领导要有一种紧迫感,必须意识到要解决那些降低企业的获利和顾客满意度的问题,就要与关键人员一起参与到执行六西格玛的活动中来;高层领导要将新六西格玛方法的领导力原则付诸实施(见表1-1),不断提升和优化企业六西格玛管理的领导力。

六西格玛卓越的价值观可以强化企业好的文化,变革不利于企业的风气。如六西格玛强调通过运用严谨的科学方法和工具,对数据和事实进行分析,为决策提供依据,而不是“拍脑袋”。当六西格玛成为企业通用的一种语言时,将从根本上消除企业会议和决策争论中产生的“对人还是对事”的矛盾。它要求不断地质疑企业的信条和传统的工作方法,勇于尝试,在追求完美的同时容忍失败,能够接受和处理偶然的挫折。

这种文化的变革也是六西格玛管理取得成功所必需的。企业文化是一个组织长期以来形成并沉淀下来为大家所共识并遵循的“做事方式”。试想,没有做事方式的改变,哪来经营业绩的持久改进?当你试图改变人们的做事方式时,文化便显示出巨大的阻力。因为当战略与文化冲突时,文化恒胜;当企业文化与变革的精神不相容时,变革的努力将遭到失败。

六西格玛管理主要是关于企业经营业绩改进的,但最终要变革企业的文化。那些成功地实施了六西格玛管理的企业不仅在运行层面,包括过程、管理活动、系统和技术等推动变革,而且在更深的层次(涉及整个企业意识形态的变化)推动变革。六西格玛为企业带



来的不仅是工作过程或经营业绩的改善,更重要的是改变组织各个层面人们的工作方式——怎样确定他们的工作目标,怎样测量他们的工作业绩,他们的工作内容如何构成;如何处理组织内部的关系,怎样处理与顾客、供应商的关系等等。这些变革涉及组织中的每一个人、每一个部门。企业要实施六西格玛管理,就意味着必须改变人们的工作方式,这就是企业面临的挑战:企业文化的变革。

1.3.3 六西格玛管理与企业战略

1. 企业战略的制定和部署

战略是指组织为适应未来环境的变化,追求长期生存和发展而进行的整体谋划和决策。战略是达成愿景、实现使命的手段,反映了组织自身能力同外界环境中所孕育的机会与威胁的一种现实的结合,确定了诸如价格、价值、技术、市场占有率和组织文化等关键的成功因素。它使竞争优势达到最大,竞争劣势减到最小。

(1) 战略制定。战略制定一般采用 SWOT 分析,SWOT 是优势(strengths)、劣势(weaknesses)、机会(opportunities)、威胁(threats)的缩写,其中优势、劣势分析主要着眼于组织自身的实力及其与竞争对手的比较,而机会与威胁分析将注重外部环境的变化及其对企业可能产生的影响。多年的实践证明,SWOT 分析是战略策划中使用最广泛、最持久的分析工具。SWOT 分析不仅要找出公司的核心竞争力,即找到所拥有的资源和能力以及运用它们的独到的妙方,而且还要找出由于缺乏资源当前还不能利用的机会以及对劣势和威胁的改进、规避或遏制方法。总之,要找到内部的优势、劣势与外部的机会、威胁之间的战略匹配。战略制定包括以下步骤:

- 1) 评估当前绩效:
 - a. 用投资收益率、盈利性等评估当前绩效;
 - b. 评估组织当前战略态势,包括:使命、愿景、战略、方针。
- 2) 评价组织治理:评价组织高层领导(包括董事会和高层管理者)的绩效;
- 3) 分析外部环境:寻找显示出机会(O)和威胁(T)的战略因素;
- 4) 分析内部环境:寻找决定优势(S)与劣势(W)的战略因素;
- 5) 确定战略目标;
- 6) 进行 SWOT 战略因素综合分析,形成和确定战略方案。

“知己知彼,百战不殆”。组织在开始制定战略之时,要全方位、多方法地收集大量的内外部数据和信息,并通过对数据和信息的分析研究,分析外部环境以发现已存在的或可预见的机会与威胁,分析内部环境以找到促进组织成长的优势和制约发展的劣势。

环境分析就是监测、评价、分析来自外部和内部环境的信息,发现所有可能对组织成败有强烈影响的关键战略因素,以使组织战略能够充分、周密考虑内外部环境变化,确保在环境需求与组织供给、组织需求与环境供给之间具有良好的战略适配,形成竞争优势,避免战略意外,促进和谐发展,获取长期成功。图 1—7 显示了环境分析的主要方面和因素。

在环境分析中,情景策划是一种用来设想将来可能会遇到的一系列问题的方法。为了确立将来的趋势和不确切的走向,通过情景策划,可构造一系列情节和情景,包括好的未来、差的未来和不可预测的状况,使企业像“照镜子”一样,看到从未看到过的自己,有

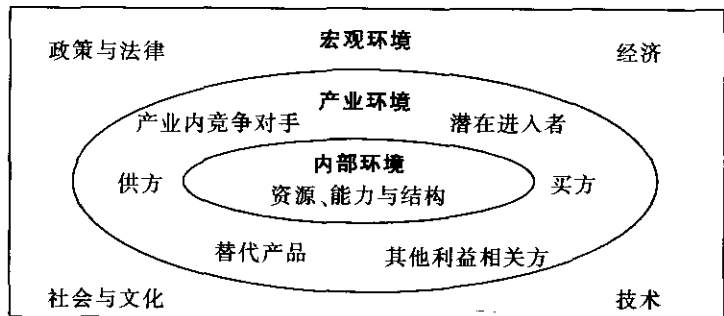


图 1—7 环境分析的主要方面和因素

助于弥补决策过程中过度自信和视野狭窄导致的常见错误，作出应对的行动计划。

在完成了外部和内部两方面的环境分析后，可以使用 SWOT 方法，对外部、内部战略因素进行综合归纳，以确定战略目标、形成战略方案。

(2) 战略部署。当组织的战略目标和战略方案制定后，最重要的就是战略部署，使战略方案具体化，制定在战略规划长短期内的各职能战略的详细规划，包括行动计划及关键绩效指标、预算和相关新程序等，沿管理层级、时间区间两个维度展开，并予以资源配置。

孙子兵法说：“上下同欲者胜。”组织要做到“上下同欲”，一是仰仗强有力的组织文化，二是有赖于以关键绩效指标系统为经络的战略部署过程。战略部署过程是一个复杂的系统工程，需要通过科学、系统而有序的方法，将组织的战略目标和战略方案落到每一个层级、每一个过程和每一个时间段，确保组织上下左右的协调一致、融合互补。

平衡计分卡（balanced scorecard, BSC）是战略部署的典型方法。

1992 年，美国哈佛商学院的罗伯特·卡普兰和复兴方案公司总裁戴维·诺顿在《哈佛商业评论》杂志上发表的一篇绩效管理系统的论文中提出了平衡计分卡的概念，认为任何单一的绩效指标都难以反映组织的绩效全貌，必须用一个平衡的指标体系来要求组织才能使之健康地发展。此后，平衡计分卡由“绩效考核阶段”发展到“战略考核和管理阶段”。今天，平衡计分卡作为一种战略执行和绩效管理工具，在全世界广泛应用。

如图 1—8 所示，这个平衡的指标体系包括四套指标：财务、顾客、内部过程以及学习与成长，源自组织的愿景和战略，力图平衡财务和非财务目标，平衡股东、顾客、员工等利益相关方的价值，平衡短期和长期目标，平衡领先性和滞后性指标，并层层展开，层层制定行动计划，形成各部门、团队和员工个人的平衡计分卡。

其中，“学习与成长”维度指组织和员工两方面的学习与成长，主要涉及培养员工能力、优化组织文化以及改善信息获取环境等，包括人力资源、组织文化、信息和知识、管理体系以及创新与改进等方面的关键绩效指标。

一般地，“学习与成长”维度的指标是战略基础和最领先的指标，驱动着“内部过程”指标；而“内部过程”指标又驱动着“顾客”维度指标；“顾客”维度指标驱动最滞后的“财务”指标。领先性（驱动性）指标体现了组织获得未来成功的竞争能力。

平衡计分卡与传统的目标管理相比，更加强调“平衡”和“战略”，强调组织各职能的目标和行动都要平衡各相关方的利益、平衡长期和短期目标、平衡领先性和滞后性指

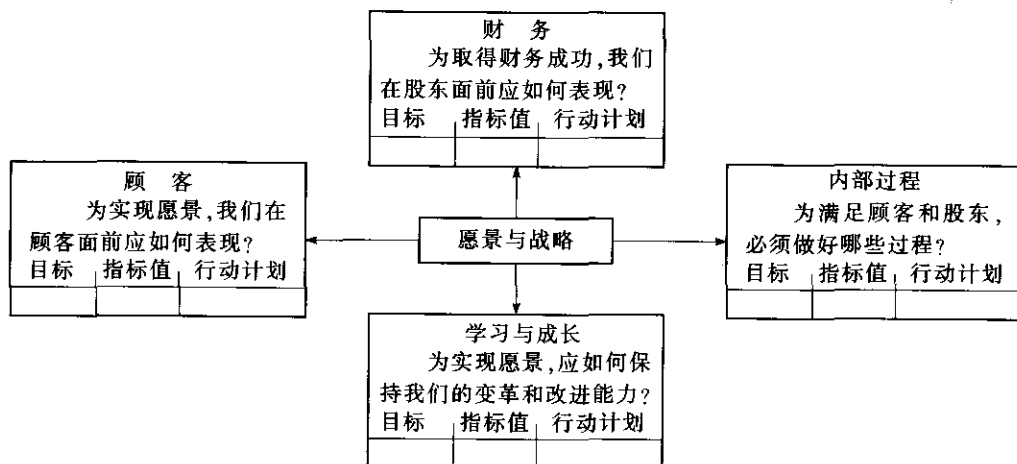


图 1—8 平衡计分卡示意图

标,都要置于组织总体愿景和战略的指导之下。

(3) 关键绩效指标测量系统。企业运用平衡计分卡方法进行层层分解的战略部署后,关键绩效指标就已经确定了。建立完善的关键绩效指标测量系统,有利于监测战略的部署,通过六西格玛等方法分析和改进战略绩效,对战略实施情况进行评估和控制。

完整的关键绩效测量系统应包括以下 5W1H 要素:

为什么测量 (why): 是关键的指标,能够反映组织的核心竞争力。

测量什么 (what): 除了测量指标名称、指标值之外,还包括指标定义和计算公式,以避免出现歧义、争议。

如何测量 (how): 即测量方法,包括数据来源(来自现场、信息系统或其他统计渠道)、获取方法(如检验、统计、抽样调查、审核评价等)以及测量规程等。

何时测量 (when): 即测量时间或频次,如按周、月、季、半年或年测量。

在何处测量 (where): 即测量点所在地点、部门或过程。

谁测量、谁分析和改进 (who): 包括负责数据测量及整合、分析和改进的责任部门或责任人。

2. 六西格玛: 作为一种企业战略

在过去的许多年里,六西格玛仅仅被视作一个统计学上的概念或是一种过程改进工具,用来帮助企业改进制造过程,降低产品缺陷。而在通用电气公司,六西格玛已经演变为一种应对动态的竞争环境、提升企业竞争力、取得长期成功的企业战略,成为公司发展的四大战略之一。韦尔奇说:“通过不断地升级和更新,我们已经大大地提高了通用电气产品和服务的质量。但是,我们做得仍然远远不够,我们也仍然远远不能够与那些拥有世界顶级产品质量的公司相提并论。这些公司已超越了与其他对手竞争的境界,而是自己向自己提出挑战,从而推动产品质量达到更新、更高的层次。”因此,他确信质量的改进将成为通用电气发展战略方面的新突破,动员整个公司全力以赴,并且使六西格玛战略与通用电气公司正在实施的其他战略相结合,同时借助六西格玛促进公司的“群策群力”和“无边界合作”。六西格玛管理终于像熊熊烈火一样燃烧着整个通用电气,使之成为富有竞

争力、敏捷灵活的卓越公司。

今天,企业面临着前所未有的经营压力,这种压力不仅来自降低成本的需要,也来自经营增长、加速创新和生产力持续提升的需要,来自市场竞争、全球化、信息技术以及经营变革步伐的加快。对企业而言,产品的高品质、对市场的快速反应、缩短周期、供应链运作的高效、建立持久的顾客关系等等,已经成为竞争的关键,这不仅是为了企业的成功,更是为了生存。在巨大的压力面前,企业需要不断提高经营绩效,寻找改进经营业绩的机会。六西格玛作为一种突破性企业战略的意义就在于,以其严格的科学方法和对经营绩效的突破性改进模式,将一个企业的所有组成要素紧密地连接起来,集中资源实现底线结果(波动、生产率、周期时间、成本等)和顶线结果(顾客满意、市场份额、营业收入等),帮助企业实现经营绩效的突破,并最终成为组织持久的绩效改进基因,提供一种战略方案、一个追求卓越和走向成功的阶梯。

3. 六西格玛:作为战略绩效改进的方法

如果将六西格玛作为一种战略是推行六西格玛管理的最有魅力的选择,那么,将六西格玛作为战略绩效改进方法,将六西格玛项目与组织目标密切联系,则是推行六西格玛管理所必需的。

六西格玛项目与组织目标的链接,意味着六西格玛项目来源于企业的关键绩效指标系统,以确保六西格玛能够支持企业战略的实现。通过关键绩效分析、评审,识别六西格玛的改进机会,以战略目标和关键绩效指标驱动六西格玛的实施。

关键绩效评审是在战略实施过程中,高层领导掌控公司战略实施和运作状况,寻找改进决策点的重要手段:基于绩效的分析和预测,评审组织的绩效和应变能力,包括组织的成就、与竞争对手和标杆绩效的比较以及长短期目标的进展,明确组织的竞争地位,确定改进关键业务的优先次序,并识别创新的机会。可以从以下四个方面进行评审:

- (1) 评审组织绩效的当前水平;
- (2) 评审组织绩效的趋势;
- (3) 竞争绩效和标杆绩效对比;
- (4) 评审绩效结果是否达到了关键的绩效要求。

一般地,六西格玛管理应着重解决那些与公司战略相关的重要而复杂的、需要突破性改进的难题,而那些员工身边存在的一般性问题,则可以通过 QCC 等方法进行解决。

4. 六西格玛战略风险分析

风险的定义是:对目前所采取的行动,在未来没有达到预期结果(失败)的可能性。其大小可用失败的概率和失败的后果两个变量来标识。

六西格玛战略风险分析指对企业推行六西格玛战略的风险分析。企业可以通过情景策划、应用 SWOT 分析,对推行六西格玛战略的内部能力、资源和文化等方面的优势和劣势以及外部环境的机会和威胁进行分析和归纳,并针对所识别的内部劣势、外部威胁等潜在风险,采取各种措施来缓解、规避或监控。

一般地,六西格玛管理战略的风险可包括:

(1) 六西格玛战略与企业其他战略的协调性。六西格玛管理的实施必须与企业其他战略保持良好的协调性,有效地促进企业其他战略的达成。例如,在通用电气,六西格玛战略与全球化、服务和电子商务战略的密切结合,使得四大战略齐头并进,相得益彰。



而如果六西格玛管理不能很好地与企业其他战略协调,所选择的六西格玛项目不能有效地支持战略目标的达成,那么就无法通过推行六西格玛管理提升企业的核心竞争力,也很难获得高层领导的整体认同和支持,使得六西格玛管理不能持续下去。

(2) 六西格玛管理对企业文化带来的挑战。文化是水,战略是舟;水既能载舟,亦能覆舟。六西格玛管理是对业务运作过程的变革,当企业文化与变革的精神不相容时,变革的努力将遭到失败。有些企业忽略了通过六西格玛创建持续改进的质量文化,没有高层领导的承诺和全体员工的共识,习惯于原来的做法而不愿意改变,相关部门之间不易协调,而仅仅是把六西格玛活动当作一场运动,轰轰烈烈推进了一阵子,便又恢复了从前的样子。

(3) 六西格玛在管理和技术方面的阻力。在管理方面的阻力包括:

- 缺乏科学合理的项目实施规划。有些企业推行六西格玛,认为只要选派几个人参加六西格玛学习班,拿到六西格玛黑带或绿带证书就可以了。离开企业中的六西格玛项目实施,任何公司的六西格玛黑带或绿带证书都是废纸一张。

- 机械地模仿。有些曾在某些大公司获得黑带(甚至资深黑带)的人往往以六西格玛专家自居,倾向于把大公司的做法强加于一些中小企业,不根据企业的具体实际而机械地模仿。任何一种管理模式,其科学性的理论和实践是可以学习借鉴的,但在具体应用上,要从企业自身实际出发,照抄照搬注定是要失败的。

- 错误地认为六西格玛管理仅仅是做项目,没有建立六西格玛组织结构,未能通过有效的培训形成包括六西格玛倡导者、资深黑带、黑带和绿带的关键群体,忽视团队建设,不能有效地传递六西格玛领导力和促进六西格玛在公司的广泛实施。

- 没有建立包括项目选择、立项、跟踪和总结全过程的六西格玛管理程序,以及项目成果发布、分享、认可和奖励制度,激励力量不足;人力、时间、资金等方面的资源配置不当或不足;与现行管理模式冲突,管理系统、方法之间不能协调、整合。这些都将给六西格玛的有效实施带来风险。

在技术方面的阻力主要是缺乏对六西格玛管理的专业培训和咨询。有些人错误地认为六西格玛就是统计方法在企业中的应用。统计方法固然非常重要,但是要完成企业中的实际项目往往需要多种分析方法和工具。如果缺乏正确、有效的培训和咨询指导,生搬硬套地应用统计方法,其结果往往是统计方法的误用。

(4) 六西格玛管理与企业实际情况的适应性。从原理上讲,六西格玛的管理模式适用于所有类型的企业,但是如果企业的基础管理薄弱,基础数据不完善,甚至是空白,这样的企业操之过急地推广六西格玛管理可能难以达到预期结果。

1.4

六西格玛管理的组织和推进

1.4.1 六西格玛管理的组织结构

企业实施六西格玛活动的首要关键任务是创建一个致力于流程改进的专家团队,并确定团队内的各种角色及其责任,形成六西格玛的组织体系。这是实施六西格玛管理的基本

条件和必备的资源。以黑带团队为基础的六西格玛组织是实施六西格玛突破性改进的成功保证。图 1—9 是六西格玛管理组织结构示意图。

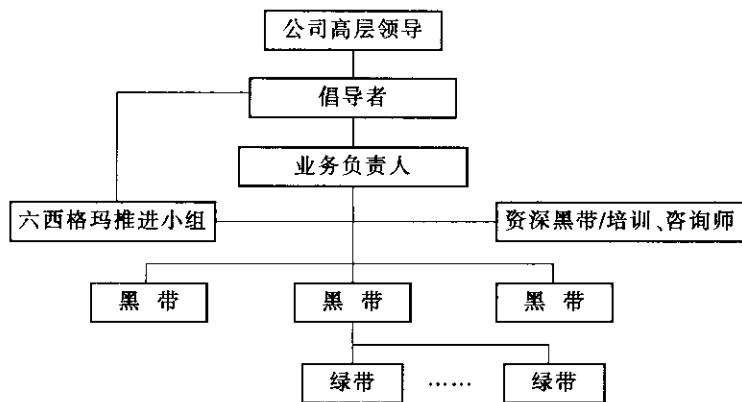


图 1—9 六西格玛管理组织结构示意图

六西格玛组织由高层领导、倡导者、资深黑带、黑带、绿带等构成。他们的职责与权限如下：

1. 高层领导

高层领导是推行六西格玛管理获得成功的关键因素。成功推行六西格玛管理并获得丰硕成果的企业都拥有来自高层的高度认同、支持参与和卓越领导。

2. 倡导者

倡导者发起和支持黑带项目，是六西格玛管理的关键角色。倡导者通常是企业推行六西格玛领导小组的一员，或者是中层以上的管理人员，其工作通常是以战略视角对六西格玛管理进行全面的战略部署、项目及目标确定、资源分配与过程监控，最终对六西格玛活动整体负责。其核心任务包括：

- 充分认识变革，为六西格玛确定前进方向；
- 确认和支持六西格玛管理全面推行，制定战略性的项目规划；
- 决定“该做什么”，确定任务的实施优先顺序；
- 合理分配资源、提供必要的支持；
- 消除障碍；
- 检查进度，确保按时、按质完成既定目标；
- 了解六西格玛管理工具和技术的应用；
- 管理及领导资深黑带和黑带。

倡导者在六西格玛组织中起着承上启下的作用，黑带应积极争取倡导者的支持。

3. 资深黑带

资深黑带的职责在不同的企业有不同的规定。例如：在通用电气公司，更多地强调其管理和监督作用；在霍尼韦尔（Honeywell）公司，资深黑带主要起协调作用，负责日程调整、项目领导和指导工具的运用。两者都很有效。

在一些企业中，资深黑带更多的是扮演企业变革的代言人角色，其工作更加具有管理性质，因为他们经常负责在全公司或特定领域、部门开展六西格玛管理工作；他们是六西



格玛管理的高参兼专家，是运用六西格玛管理工具的高手。六西格玛资深黑带的主要职责为：

- 担任公司高层领导和倡导者的六西格玛管理高级参谋，具体协调、推进六西格玛管理在全公司或特定领域、部门的开展，持续改进公司的运作绩效。
- 担任培训师，为黑带学员培训六西格玛管理及统计方面的知识。
- 帮助倡导者、管理者选择合适的人员，协助筛选最能获得潜在利润的项目。
- 为参加项目的黑带提供指导和咨询。
- 作为指导者，保证黑带及其团队保持在正确的轨道上，能够顺利地完成任务。
- 具体指导和协助黑带及其团队在六西格玛改进过程中完成每个步骤的关键任务。
- 为团队在收集数据、进行统计分析、设计试验及与关键管理人员沟通等方面提供意见和建议。

4. 黑带

黑带是六西格玛管理中最为重要的一个角色。他们专职（或兼职，但专职是最佳选择）从事六西格玛改进项目，是成功完成六西格玛项目的技术骨干，是六西格玛组织的核心力量。他们的努力程度决定着六西格玛管理的成败。黑带的主要任务是：

- 领导。在倡导者及资深黑带的指导下，界定六西格玛项目，带领团队运用六西格玛方法完成项目。
- 策划。决定项目每一个步骤需要完成的任务，包括组织跨职能的工作。
- 培训。具有培训技能，为项目团队成员提供新的战略和最有效的工具及技术应用的专门培训。
- 辅导。为组员提供一对一的支持，带领绿带队友快速有效地达到改进的目标。
- 传递。在各种形式的培训、案例研究、工作座谈会和交流活动中，将新的战略和新的工具方法传递给团队的其他成员。
- 发现。在内部或外部（如供应商和顾客等）找出新战略和新工具方法运用的机会，与资深黑带一起确定有价值的项目，解决一些有关资源的问题。
- 确认。通过与其他组织的合作，发现新的商业机会。
- 影响。拥有良好的人际关系和组织技巧，令团队始终保持高昂的士气与稳定的情绪。
- 沟通。项目完成后向最高管理层提供项目报告。

在六西格玛项目中，黑带组织、管理、激励、指导一支特定的六西格玛项目团队开展工作，负责团队运作的启动，管理团队的进展，并最终使项目获得成功。在推行六西格玛管理的企业里，如果没有一些具有实力且不怕辛苦的黑带，六西格玛项目通常不会取得最佳效果。

作为一名黑带必须拥有以下多项技能：

- 管理和领导能力。黑带必须能够运用权力和职责来指导项目的执行，要能够综合运用自己的管理能力和领导才能，并且能够熟练运用项目管理的方法和技巧。
- 决策制定。在六西格玛项目中，黑带可能要做出无数次的决策。为制定可靠、及时的决策，黑带必须随时掌握和了解项目的每一个方面，能够平衡成本、时间和效果。

● 沟通。将项目活动内容与结果及时与相关人员(团队成员、上层管理者、项目倡导者和组织的关键利益相关方等)沟通。

● 团队建设和谈判。黑带必须能够与不同的人建立持久的联系,如管理层、顾客、团队成员、项目倡导者及供应商等,这是由上级认同的特权。一个优秀的黑带必须能够经常与上级领导沟通和谈判,使六西格玛项目的进行获得优先权。

● 策划、调度和行动。与其他项目管理活动相同,六西格玛项目管理包括:目标建立、项目细化、绘制工作流程、任务调度、成本预算、协调团队、组员沟通等活动。黑带必须进行有效的策划和高效的行动,平衡项目规划和进度安排,这些是项目成功的关键。

● 关注全局。一个成功的黑带要能够回顾和预见项目任务的所有方面,而对项目细节的过分关注,可能会失去对项目整体上的判断。

● 人际交往的能力。作为项目领导,黑带必须具有一定的人格魅力:诚实、有能力、可信赖、有包容心;与项目倡导者和组织的关键利益相关方建立良好的关系;将具有不同背景的人员组成一个统一的团队。

5. 绿带

绿带是黑带项目团队的成员或较小项目的团队负责人,他们接受六西格玛技术培训的项目与黑带类似,但内容所达层次略低。一些实施六西格玛的企业,很大比例的员工都接受过绿带培训,他们的作用是把六西格玛的新概念和工具带到企业的日常活动中去。

在六西格玛管理中,绿带是人数最多,也是最基本的力量。他们的职责是:

- 提供相关过程的专业知识;
- 建立绿带项目团队,并与非团队的同事进行沟通;
- 促进团队观念转变;
- 把时间集中在项目上;
- 执行改进计划以降低成本;
- 与黑带讨论项目的执行情况及今后的项目;
- 保持高昂的士气。

当绿带作为项目团队负责人完成绿带项目时,也应具有黑带在项目团队负责人方面的职责、权限和技能。

6. 业务负责人

除了要选择培养好项目负责人——黑带之外,成功的六西格玛项目还需要相关业务部门负责人(过程管理者)的支持和配合,没有他们的协调和帮助,六西格玛很难取得丰硕的成果。业务负责人不需独立完成项目,他们在六西格玛管理中的职责是:

- 达成对六西格玛的共识;
- 协助选择黑带、绿带;
- 为黑带、绿带提供资源支持;
- 关注黑带、绿带的项目实施过程;
- 协调所管辖范围内的黑带、绿带项目,保持与业务方向的一致性;
- 确保过程改进能够落实,保持改进成果。

六西格玛团队建设也是六西格玛组织工作的关键环节,具体内容详见第3章。



1.4.2 六西格玛管理的推进步骤

如图 1—10 所示,企业推行六西格玛一般可分四个阶段,即导入期、加速期、成长期和成熟期,大约可用 4~5 年的时间完成从导入期到成熟期的全过程。

在六西格玛推进的过程中,每一阶段都会遇到三类阻力:技术阻力(对方法的恐惧、技术力量的不足等)、管理阻力(部门间的沟通壁垒、激励机制和资源缺乏等)和文化阻力(观念上不认同、靠经验和感觉做决策、变革动力缺失等)。当推进的动力难以抵御这三方面阻力的合力时,就会出现图 1—10 所示的“关键转折点”。如果不能有效地增进动力、降低阻力,六西格玛管理就会在某一阶段“夭折”。而如果成功地越过这些转折点,六西格玛就能在企业深入、持久地开展下去,为企业创造越来越多的效益,创造越来越强劲的成功能力。

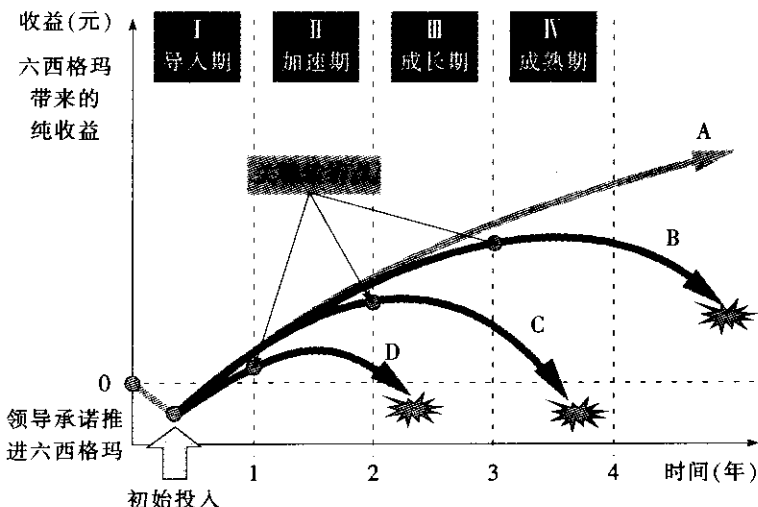


图 1—10 六西格玛管理的推进轨迹

资料来源: Strategic Six Sigma, Best Practices From The Executive Suite, Dick Smith and Jerry Blakeslee with Richard Koonce, Wiley.

1. 导入期

导入期又可分为起步、培训与改进实践、坚持不懈与获得成功等步骤。

(1) 起步。当企业决定要实施六西格玛时,会打破企业看似平静的现状。这时需要:

- 公司高层领导(首席执行官、总裁)支持六西格玛,是六西格玛的信仰者;
- 公司已经有了相当于倡导者、资深黑带的推动者(或选择了一个有经验的咨询服务方);
- 有了领导的承诺,为公司实施六西格玛建立了愿景或找到了一个切实的实施理由;
- 将六西格玛定为公司达成战略目标战略举措,或者制定了一个具体的六西格玛推进行动计划,拟订首批项目和黑带培训种子人选,有了初期投入的财务预算。

(2) 培训与改进实践。六西格玛管理的培训与项目实践是嵌套式的、融为一体的。一般是先培训第一批黑带“种子选手”,再由这些种子选手负责培训绿带、黄带。相关培训的内容和目的如表 1—3 所示。

表 1—3 六西格玛管理培训

类别	内容	目的
高层领导	六西格玛理念	了解六西格玛的理念和作用，统一高层的思想
倡导者	六西格玛领导与推进、团队合作、项目选择	熟悉如何选择和确定六西格玛项目，如何推进六西格玛
黑带	六西格玛全套课程，包括完整的六西格玛工具箱	掌握实施六西格玛项目的方法、工具和技术
绿带	六西格玛基础知识，包括各阶段常用工具和技术	熟悉实施六西格玛项目的常用方法、工具和技术
黄带	六西格玛普及知识	了解六西格玛基础知识

(3) 坚持不懈与获得成功。在首批黑带培训和绿带培训完成后，应有一部分项目已经完成，也会有一批项目正在进行中。在这一时段内，贵在坚持和恒心。只要坚持到底，就能够在年内收获一批成功的果实。只要初期投入不是太大，都能在年内收回所有投入并获得一定量的回报。

当高层领导一时还未对六西格玛管理做出承诺，企业难以采用上述的全面导入方式时，可以采用局部驱动方式，即在一些部门、区域或产品上小范围推行，为将来的全面展开积累经验、做出示范，用其成果说服企业中的其他人。这种方式的特点是：容易起步，仅需要有限的管理层关注，所需投入的资源比较少，因此风险也较小。这种方式只是作为将六西格玛引入企业的一种切入方式，只有及时转向在全公司范围的充分展开，六西格玛管理才能枝繁叶茂、硕果累累，并取得长期的成功。

局部导入也有两种方法：一是选派 3~5 名黑带或绿带外出培训，结合培训实施若干六西格玛项目，但需要克服缺乏直接的项目辅导的困难，同时要注意对项目相关人员的六西格玛宣导，以便获得相应的理解和支持；二是引入外部咨询顾问，进行小范围的绿带或黑带培训，结合培训辅导实施若干六西格玛项目，同时对管理层和项目相关人员进行六西格玛导入培训。

2. 加速期

第一年导入期的成功之日，也是新的转折之时。虽然经过第一轮项目工作，企业获得了初步的结果，也有了实施的热情和积极的参与者，一些冷眼观望乃至反对的人也开始转变原有的观念，但这时是一个关键的转折点。如果没有下一步的正确部署，六西格玛就会是一个短期的流行并走向失败，而且一旦热情冷却，就犹如一盘“冷饭”，今后要再次“蒸热”的难度就会更大。在这一转折点上应当引入“加速实施过程”，使六西格玛从试验性实施向企业的一项长期管理活动过渡。

要实现这一转折，企业应当：

- 制定六西格玛财务预算、核算和审计办法，使财务人员介入六西格玛活动。
- 建立项目成果发布、分享、认可和奖励制度，激励六西格玛黑带和团队。
- 加大培训力度，形成六西格玛倡导者、资深黑带和黑带这一关键群体，以传递六西格玛领导力，促进六西格玛在公司的广泛实施。
- 建立六西格玛管理程序，包括六西格玛组织结构以及项目选择、立项、跟踪和总结的全过程管理程序。



3. 成长期

对一个导入了六西格玛管理并成功实施了约两年工作后的企业,仍然会出现六西格玛管理“断流”的趋势。所有实施六西格玛的企业都会遇到这样的困难,其中最重要的一个原因是经营环境在不断变化,总有新技术、新方法和新政策等出现。为了获得持续发展,需要不断地将六西格玛工作拓展到公司的所有方面,包括用六西格玛促进新技术的应用,促进创新和新市场的开发。要成功地在这一点上实现转换,公司必须完善其支持基础:

- 完善六西格玛管理的组织结构,将其对六西格玛的管理职能充分展开,强化最高管理层对六西格玛管理的系统、定期评审,并使已关闭的项目持续产生效益。
- 拓展六西格玛的实施领域,如加大六西格玛管理方法在非制造领域的应用,用六西格玛设计促进创新和开发工作,将六西格玛管理沿供应链向供应商和顾客延伸等。
- 完善六西格玛培训体系,扩大培训范围,加大黑带、绿带占员工总数的比例。
- 使六西格玛管理与企业战略策划、部署和经营过程结合,强化六西格玛与顾客要求 and 市场发展趋势的结合。

4. 成熟期

最后一个转折也许是最困难的。将六西格玛融入企业,成为人们的一种工作和思维方式,确实是很难用时间表来预计其实现时间的。实际上,前面几个阶段的努力,都是在为这个阶段打基础。这个转折的关键是将六西格玛与公司其他管理活动有效地整合,进一步强化经营管理过程,建立完善的绩效改进体系,强化人们观念和行为方式的改变。要实现这一转折,公司应当:

- 使六西格玛价值观与公司的使命、愿景和核心价值观高度融合,强化人们观念和行为方式的改变。
- 将六西格玛与组织其他管理战略、管理体系和改进方法整合,建立高度整合的全面质量管理或卓越绩效管理体系,高度整合的持续改进、创新和知识分享体系。
- 使六西格玛成为日常工作的一部分。



1.5 六西格玛管理方法论

1.5.1 六西格玛改进的模式——DMAIC

前面讲到,六西格玛自20世纪80年代诞生于摩托罗拉以来,经过20年的发展,现在已经演变成为一套行之有效的解决问题和提高企业绩效的系统的的方法论,其具体实施模式为DMAIC。

DMAIC代表了六西格玛改进活动的五个阶段:

- 界定阶段 (define)
- 测量阶段 (measure)
- 分析阶段 (analyze)
- 改进阶段 (improve)
- 控制阶段 (control)

在六西格玛项目选定之后，团队成员一起合作，依照这个过程的五个步骤，可以有效地实现六西格玛突破性改进。团队的工作从一个问题的陈述到执行解决方案，中间包括了许多活动，通过 DMAIC 过程的活动方式，团队成员可最有效地发挥作用，完成项目使命。

DMAIC 是一个逻辑严密的过程循环，它是在总结了全面质量管理几十年的发展及实践经验的基础上产生的，是由项目管理技术、统计分析技术、现代管理方法等综合而成的系统方法。DMAIC 强调以顾客（外部和内部）为关注焦点，并将持续改进与顾客满意以及企业经营目标紧密地联系起来；它强调以数据的语言来描述产品或过程绩效，依据数据进行管理，并充分运用定量分析和统计思想；它追求的是打破旧有习惯、有真正变化的结果和带有创新的问题解决方案，以适应持续改进的需要；它强调面向过程，并通过减小过程的变异或缺陷实现降低风险、成本与缩短周期等目的。

1. DMAIC 过程活动

DMAIC 过程共分五个阶段实施，每个阶段的工作内容如下：

界定阶段：确认顾客的关键需求并识别需要改进的产品或流程，组成项目团队，制定项目计划，决定要进行测量、分析、改进和控制的关键质量特性（CTQ），将改进项目界定在合理的范围内。

测量阶段：通过对现有过程的测量和评估，制定期望达到的目标及绩效衡量标准，识别影响过程输出 Y 的输入 X_s ，并验证测量系统的有效性，确定过程基线。

分析阶段：通过数据分析确定影响输出 Y 的关键 X_s ，即确定过程的关键影响因素。

改进阶段：寻找最优改进方案，优化过程输出 Y 并消除或减小关键 X_s 的影响，使过程的缺陷或变异降至最低。

控制阶段：对改进成果进行固化，通过修订文件等方法，使成功经验制度化，通过有效的监测方法，维持过程改进的成果并寻求进一步提高改进效果的持续改进方法。

2. DMAIC 过程活动要点及其工具

各阶段使用的工具和技术如表 1—4 所示。

表 1—4 DMAIC 过程活动要点及其工具

阶段	活动要点	常用工具和技术	
D（界定阶段）	明确问题 确定 Y （CTQ/CTP）	头脑风暴法 亲和图 树图 流程图 SIPOC 图 平衡计分卡	力场图 因果图 顾客的声音 质量功能展开（QFD） 不良质量成本 项目管理
M（测量阶段）	确定基准 测量 Y ， X_s	排列图 因果图 散布图 流程图 测量系统分析 FMEA 过程能力指数	不良质量成本 水平对比法 直方图 趋势图 检查表 抽样计划



续前表

阶段	活动要点	常用工具和技术	
A (分析阶段)	确定要因 确定 $Y=f(X)$	头脑风暴法 因果图 FMEA 水平对比法 方差分析	试验设计 抽样计划 假设检验 多变异分析 回归分析
I (改进阶段)	消除要因 优化 $Y=f(X)$	试验设计 响应曲面法 调优运算 (EVOP)	FMEA 测量系统分析 过程改进
C (控制阶段)	保持成果更新 $Y=f(X)$	控制图 统计过程控制 防差错措施	过程能力指数 标准操作程序 (SOP) 过程文件控制

1.5.2 六西格玛设计的模式

DMAIC 流程对产品的质量优化仍具有局限性。众所周知,质量首先是设计出来的,实践表明,至少 80% 的产品质量是在早期设计阶段决定的,所以没有六西格玛设计,仅采用 DMAIC 流程来提高产品的质量,其成效是有限的。若想真正实现六西格玛的质量水准,就必须考虑六西格玛管理战略实施的另外一种途径——六西格玛设计 (DFSS)。六西格玛设计是一种实现无缺陷的产品和过程设计的方法。它基于并行工程和 DFx (design for X) 的思想,面向产品的全生命周期,采用系统的问题解决方法,把关键顾客需求融入产品设计过程中,从而确保产品的开发速度和质量,降低产品生命周期成本,为企业解决产品和过程设计问题提供有效的方法。

与六西格玛改进中的 DMAIC 流程一样,DFSS 也有自己的流程,但目前还没有统一的模式。迄今,研究者已提出的 DFSS 流程有多种,如:DMADV 流程,即界定、测量、分析、设计、验证;DMADOV 流程,在 DMADV 流程的基础上增加了优化环节;还有 DCCDI 流程,指的是界定、识别顾客需求、概念设计、产品和过程设计、实现;DMEDI 流程,指的是界定、测量、调查、开发、实现。此外还有其他的六西格玛设计流程,如 DMCDOV, DOCV, DCCDI, DMEDI, DMADIC, RCI 等。虽然这些流程的表述不同,内容却大同小异,这里就不逐一介绍了。下面我们重点介绍在业界广泛应用的由美国质量管理专家乔杜里 (Subir Chowdhury) 提出的 IDDOV 流程,即识别 (identify)、界定 (define)、开发 (develop)、优化 (optimize)、验证 (verify)。

(1) 识别。DFSS 在识别阶段的主要任务是收集和确定待开发产品的顾客需求,并论证即将开展的 DFSS 项目的可行性。DFSS 的特点之一在于产品设计之初就充分考虑顾客的需求,聆听客户的声音,利用一些评价工具对顾客需求进行识别和优先级排序,以保证设计出的产品满足客户的需要。在此基础上系统地考虑外部环境、市场和企业内部满足此项目开发所需的资源等,如市场上的竞争对手的产品状况、企业的工艺水平、人员情况、开发费用等,以确定项目的可行性。这一阶段对整个 DFSS 项目是非常重要的,所谓“差之毫厘,谬以千里”,若此阶段不能充分收集和分析顾客的需求、客观评价项目的可行性,将给项目带来致命的打击。在此阶段的成果一般包括:顾客的需求分析报告、项目的成本

分析及可行性报告、产品的功能要求等资料。

(2) 界定。界定阶段是 DFSS 实施的核心过程, 此阶段的任务是进一步细化展开顾客的需求, 并将海量的、具有模糊和不确定性甚至矛盾的 VOC 利用一些模糊信息处理技术转化为准确 VOC, 再通过质量功能展开将 VOC 逐层展开为设计要求、工艺要求、生产要求, 并提炼出顾客的关键需求, 准确地识别、量化顾客需求和期望并针对需求和期望进行产品设计, 最后生成产品的设计方案和工艺要求说明书。

(3) 开发。在 DFSS 的开发阶段, 所做的工作是对新产品进行详细的局部设计, 在前期工作给定的解决方案框架以及关键质量特性 (critical to quality, CTQ) 和关键过程特性 (critical to process, CTP) 尺度之内进行新产品的设计, 即可以把设计过程看作满足一定功能约束与设计约束的优化过程。本阶段结束后应完成样品的的设计, 并为采购、生产和售后服务提供一定的参考标准供后阶段讨论, 如原材料和产品的验收准则, 规定安全和正常使用所必需的产品特性以及初步的售后质保体系等。

(4) 优化。此阶段是对产品和过程设计参数的优化, 其目标是在质量、成本和交付时间允许的基础上达到企业利益的最大化。首先设计应该是稳健的, 在 DFSS 中强调预防和稳健性; 其次设计应尽量消除产品或服务失效的潜在可能, 通过失效模式和效应分析 (failure mode and effects analysis, FMEA) 分析潜在的失效模式和功能变异性, 从而在设计阶段就尽量减少产品和过程失效的可能。此阶段结束后应有详细的生产流程图, 并对生产的各环节有相应的生产要求标准, 完善的售后质保体系。

(5) 验证。DFSS 的验证阶段的任务是对产品设计是否满足顾客要求、是否达到期望的质量水平的确认过程。通过试生产等手段营造一个仿真的生产环境, 测试设计的能力、稳健性和可靠性。最后提交设计的验证报告和 DFSS 项目的总报告。

以上是 DFSS 实施阶段的简要介绍, 需要说明的是虽然 DFSS 有固定的流程可以遵循, 但在实施的过程中也不可拘泥于流程或工具的运用, 而要从 DFSS 的本质上去把握其精髓, 根据产品和流程的差异有选择地应用 DFSS 中的工具。此外, 在产品设计的过程中还必须考虑并行的产品和过程设计, 利用并行工程的思想并行设计产品和过程, 每个阶段的工作都要充分考虑其后续阶段, 在相邻的阶段之间需要有一定的交叉。有关 DFSS 的支撑工具将在第 10 章详细阐述。

本章小结

本章从质量概念的演进和质量管理的开始, 让读者了解六西格玛管理的起源和发展, 理解六西格玛的概念和作用, 掌握六西格玛管理的实施方法, 包括六西格玛管理的领导、文化和战略, 六西格玛管理的组织和推进, 以及六西格玛改进和设计模式。六西格玛管理以其严格的科学方法和对经营绩效的突破性改进模式, 将一个企业的所有组成要素紧密地连接起来, 为集中资源实现底线和顶线结果, 帮助企业实现经营绩效的突破, 并最终成为组织持久的绩效改进基因, 提供了一种战略方案, 一个追求卓越和走向成功的阶梯。

六西格玛与过程管理

2.1 过程管理基础

过程概念是现代管理最基本的概念之一，过程管理也是现代管理最基本的方式之一。从组织的使命、愿景、价值观，到战略目标和战略规划，都要通过过程管理来落实；所有的改进和创新也都是通过过程改进而实施的。现代管理重在结果和创造价值，但结果和价值都是通过过程而取得的，因此它必须面向过程。对于一个追求卓越的组织而言，过程是因，结果应是有因之果，应当通过卓越的过程取得卓越的结果。

六西格玛价值观之一就是聚焦于过程改进，减少或消除过程变异。六西格玛管理不应该是一系列相互割裂的项目组合，任何改进的努力都要与企业发展战略、顾客和其他利益相关方的需求相链接。而战略的实施和顾客及其他相关方需求的满足都要通过过程来落实，因此，六西格玛管理应当基于组织的业务过程和业务系统，聚焦于过程和系统的整体改进。

2.1.1 过程的定义

1. 过程的定义

在 ISO 9000: 2000《质量管理体系：基础和术语》中，将过程定义为：一组将输入转化为输出的相互关联或相互作用的活动。

在过程示意图（如图 2—1）中：

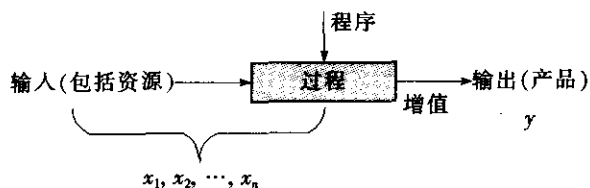


图 2—1 过程示意图

- 过程的任务在于将输入转化为输出，而输出就是过程的产品（包括服务），输入、输出以及过程特性应是可测量的。输入和预期的输出可以有形的（如设备、材料和元器件），也可以是无形的（如能量或信息）。输出也可能是非预期的，如废料或污染。

- 对形成的产品是否合格不易或不能经济地进行验证的过程，通常称为“特殊过程”。
- 增值是对过程的期望，即每一个过程均应该是增值的过程，否则应进行改进或删除。
- 为了使过程增值，组织应当对过程进行策划，即识别过程及其要求，进行过程设计并形成程序，建立过程绩效测量和过程控制方法。过程程序的实施能够使组织稳定受控地获得增值，但为了使过程具有更强的增值能力，组织还应当对过程进行持续改进和创新。
- 将输入转化为输出的动因是活动，而且是一组相互关联或相互作用的活动。
- 过程具有伸展性：一个过程可以分解成若干更小的子过程；若干小过程可集成为一个较大的过程。
- 根据所含活动的多少、大小，过程可以是一个大过程（如产品实现过程），也可以是一个小过程（如某一个冲裁生产过程）。

过程输入可以是人力、设备设施和材料，或者是决策和信息等。输入包括可控因素，也包括噪声因素，即那些被认为不可控制、不期望控制或控制费用昂贵的因素。实用而简单的过程概念模型是： y 是 x 的函数，即

$$y=f(x_1, x_2, \dots, x_n)$$

式中， y 代表结果变量； x 代表输入变量或过程变量。 x 是自变量，是独立的因素， y 是因变量，是非独立因素。上式表达的信息是，通过选取不同的 x 值，能够改进过程输出 y 的值。其中关键输入变量和关键输出变量分别称为 KPIV (key process input variables)，KPOV (key process output variables)。

2. 过程链和过程网络

过程链是指组织内或组织间一个个首尾串接的过程构成的链条，而过程网络则是由多个过程链串接、并接而构成的网络。如图 2—2 所示：

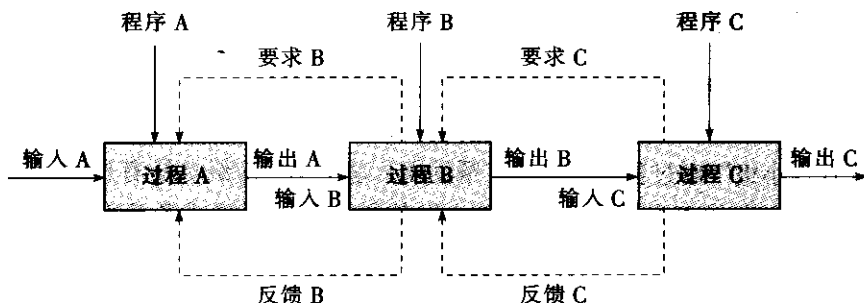


图 2—2 过程链示意图

- 前过程的输出就是后续过程的输入，由此构成过程的串联链接，构成了过程链，也称为“流程”。过程之间是相互关联、相互作用的。
- 前过程应当主动识别和确定后续过程的要求。
- 后续过程应当向前过程提出要求，并向前反馈过程输入质量等相关信息，前过程据此进行过程设计、控制和改进。
- 如果过程链反映的是组织之间的大过程链接，则该过程链就是链接供应商（过程 A）、组织（过程 B）和顾客（过程 C）的供应链。如图 2—3 所示：

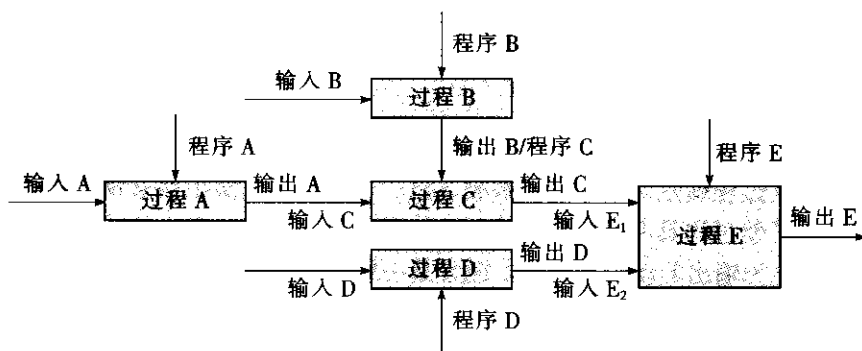


图 2—3 过程网络示意图

● 事实上，过程之间的链接既包括串联链接，也包括并联链接（若两个过程同时进行，其输出均为后一过程的输入）。

● 一个过程会有多个输入，同时也会产生多个输出。

● 过程的混联链接是很常见的。图 2—3 的过程 B 是一个过程策划过程（如工艺设计过程），其输出就成为过程 C 的程序（如工艺规程或作业指导书），而过程 C 和 D 是并联的，其输出同时都是过程 E 的输入。

● 过程网络就构成了所谓的“系统”或“体系”。通过建立和理解过程及其相互作用形成的网络，把过程作为一个系统进行管理，其协调运行通常称为“管理的系统方法”。系统也有大小之分，如一个班组、车间的系统，一个公司的系统，跨公司的系统。

● 系统是过程的集成，系统的目标通过过程而完成。

● 在过程网络中起主导作用的过程称为关键过程。

2.1.2 过程负责人和相关方

为确保实施、保持和改进过程及其相互关系，管理层应当规定每个人的作用和职责。负责某一过程的人或团队通常称为该过程的负责人。为了管理过程之间的相互关系，建立一个“过程管理小组”对过程改进是很有帮助的。这个小组应纵览所有过程，且每一个相互作用的过程都应有代表参加，那么，该小组就是这个过程链或者流程的负责人。在图 2—4 中，蒸汽生产过程负责人就包括负责加热、燃烧作业的操作工，以及负责操作和运营管理的班长、车间主任。

每一个过程都有顾客和受过程输出影响的其他相关方（他们可以是组织内部的，也可以是外部的），他们是过程输出（即产品）的接受者，即广义的顾客。利益相关方指组织的行动和成功所影响的所有群体。关键的利益相关方通常包括顾客、股东、供应商及合作伙伴、员工以及社区等。在图 2—4 蒸汽生产过程的利益相关方包括预期产品过热蒸汽、热水的顾客，非预期产品烟尘、废气和废渣的影响者社区（环境），过程信息输出（报表、记录）和过程绩效（产值、成本、效率）的影响者股东及受股东委托管理公司的管理层，过程安全健康的影响者员工，还包括过程输入的外部供应商水厂、煤矿和内部供应商空压机站、计划部、技术部、财务部、人力资源部等。

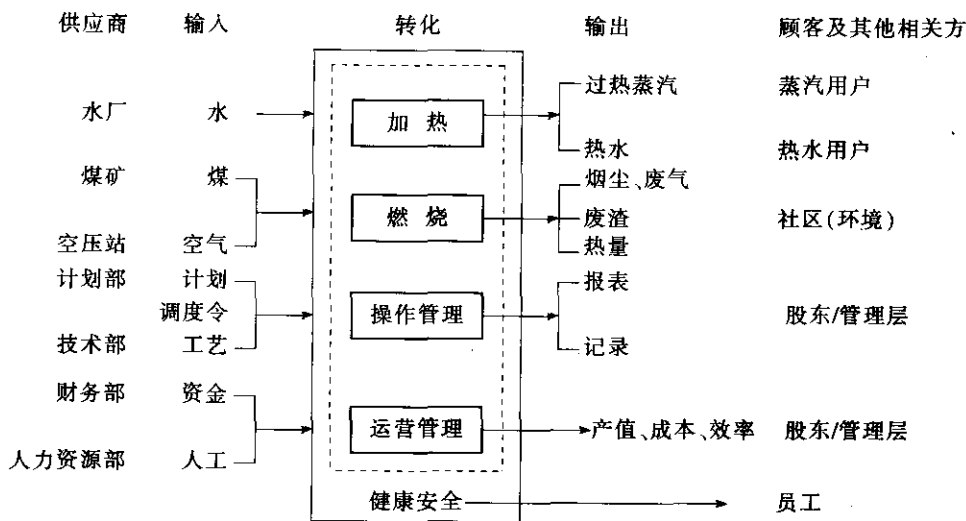


图 2-4 过程要素分析

2.1.3 价值链与过程类别

1. 价值链

组织的生存和发展依赖于顾客，而顾客希望从购买的产品或服务中获得某种价值，所以组织必须提供能够吸引和保持顾客的价值。

那么，价值是如何提供给顾客的呢？是通过将原材料和其他要素转变为顾客在特定时间、特定地点，以特定方式需要的产品或服务来实现的。这些将原材料和其他要素转变为顾客需要的价值以及顾客愿意为此付钱的过程看似简单，却包括了一大串由许多不同参与者（供方、员工、顾客）参与的过程，也就是前面所述的过程链、过程网络。从增值的角度看，其中具有增值作用的过程链（网络）就构成了价值链。广而言之，价值链就是在产品实现全过程（对于制造业来说，就是从原材料加工到产成品送达最终顾客）中，所有增值步骤组成的全部有组织、有内在联系的一系列活动。完整的价值链可能会包括从供方的供方到顾客的顾客的所有部分。

迈克尔·波特在 1985 年出版的《竞争优势：创造和保持卓越绩效》中提出了价值链的概念：把公司（组织）描述成“设计、营销、配送和支持产品等一系列活动的集合体”。他想要让经理们理解为顾客创造价值的组织活动的序列，并强调经理们必须理解他们的组织是怎样适应行业的整个价值创造活动的。价值链概念的三个特性最终把“核心过程”转变为企业管理人士关注的焦点。

- 价值链增强了业务活动内部的关键联系，也增大了组织成功的可能性。不同的职能部门在组织的基本目标中起着不同的作用：向顾客传递某种独特的价值，价值链的任何一个链条中断或弱化（不同职能部门之间的不协调，不增值的活动）都会影响价值传递过程。

- 价值链确定了职能部门的贡献率。当每个职能部门都贡献价值时，其中一些起主导作用，另外一些起次要作用。核心过程的贡献率比较显著。

- 价值链被细化到公司的操作单位层面。

价值链管理就是要达成链中成员的无缝链接,以团队的工作方式增值,表现为:快速组装、更准确的信息、更快的顾客反应速度和更好的服务等,以满足和超越顾客的需求和期望,为顾客创造价值并促进顾客的成功。这样,顾客就会持续地为此付出,为价值链上的成员创造财富。

组织应当通过价值链管理,来实施其战略规划,达成其战略目标,进而实现其使命和愿景。

2. 过程类别

组织在为顾客创造价值的同时,也为所有者、员工、供方和社会等其他利益相关方创造价值。

组织的增值分为两类:一类是直接为顾客增值;另一类是为组织自身增值,即不断增强自己应对竞争环境,取得长期成功的核心竞争力。

基于上述两类增值,组织的增值过程也分为两类,即为顾客增值的产品实现过程和为组织自身增值的经营过程。这些增值过程就构成了组织的价值链,其中增值最大、最显著的称为“核心过程”,在卓越绩效模式中也称为“主要价值创造过程”,典型如设计开发、生产运作和顾客服务过程,以及供应管理、市场营销、战略发展等过程。

在组织的过程链、过程网络中,除了上述直接创造价值的过程外,还有一类过程虽然不直接创造价值,但却对价值创造过程的实施和组织日常运作起支持、保证作用,我们称之为“支持过程”。其中最关键的又称为“关键支持过程”,典型的如人力资源、财务、基础设施以及信息技术管理过程。

六西格玛管理聚焦于过程改进,特别是核心过程的改进,也包括对关键支持过程的改进。

2.1.4 过程与职能、项目的关系

1. 过程与职能的关系

组织结构是组织中正式确定的使工作任务得以分解、组合和协调的框架体系。最典型、传统的组织结构是职能型结构,即将相似或相关职业的员工组合在一起,形成职能部门。其优点是分工明确,有利于专业化和知识共享,并减少人员和资源的重复配置,节约成本。其缺点是由于专业化的发展,使得各部门容易过分强调本部门的重要性和利益,追求职能目标而忽视整体目标,部门之间相互隔离而导致看不到整体利益,形成“隧道视野”,难以横向协调和沟通。

图2—5是职能制组织结构图与流程图的组合,阴影部分为一个业务流程图。业务流程、过程是横向的,就像一条横向流淌的河,而职能部门、职能犹如纵向的大坝,部门之间由于业务目标的差异导致了沟通壁垒,这种沟通壁垒又大大地影响了流程运行的效率和效果。

传统的职能制组织结构已不能适应动态和复杂化的环境,现代组织的领导者试图寻找创造性的办法来构建和安排组织的工作,建立过程导向的组织结构,打破部门封锁的团队。六西格玛项目是靠跨职能团队完成的,六西格玛团队要从顾客和企业的整体流程出发,以顾客需求和企业战略为导向,解决影响流程顺利执行的关键环节中的问题,优化整

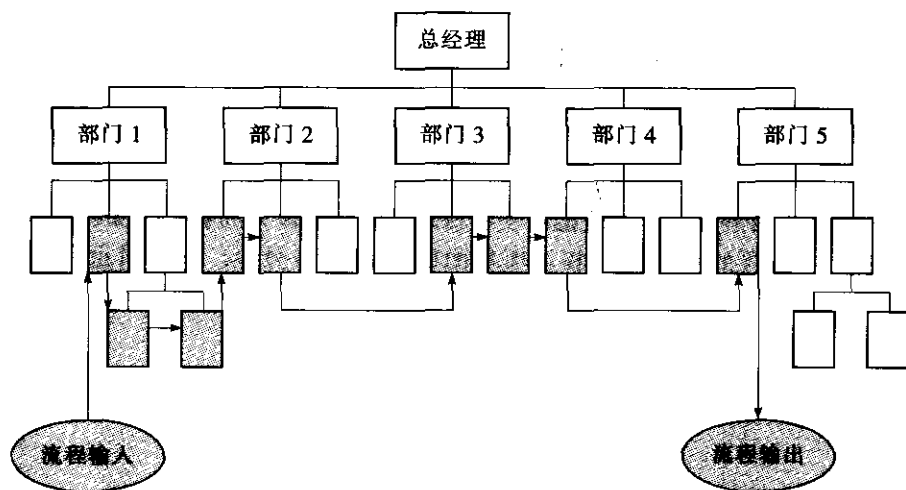


图 2—5 组织结构图与流程图的组合

体流程，实现持续改进。

2. 过程与项目的关系

职能制组织结构的缺点并不是新的发现，早在 20 世纪 20 年代，通用汽车公司就建立了“部门间关系委员会”来处理生产线与其他部门之间的以及公司分支机构内部不同职能部门之间的摩擦。通过“流程再造”、“结构重建”等措施，试图打破组织内部沟通的障碍的例子，在现代工商业史上不胜枚举，直到今天仍在进行之中。

当组织开始理解过程和职能之间的差异，并跨越不同职能部门的界限来描述过程时，实现部门间真正合作的关键要素就出现了。越来越多的组织开始导入基于团队的结构、矩阵制结构、无边界组织和学习型组织等现代组织结构模式，建立了越来越多的与业务流程方向一致的跨职能团队，以打破不同部门之间的沟通壁垒，提高过程的效率、效果和敏捷性。

今天，业务流程改进早已超越了自发的、非正式的组织形式，采取了有组织的逐个项目改进的方式。建立六西格玛、QCC、技术攻关等跨职能的项目团队，通过项目管理来实施业务流程改进是一条有效的路径。

项目是指“由一组有起止时间的、相互协调的受控活动组成的独特过程，该过程要达到符合时间、成本和资源的约束条件在内的规定要求的目标”（ISO 9000）。单个项目可作为一个较大项目结构中的组成部分，其结果可以是一种或几种产品。项目基于过程和 PD-CA 循环，是通过一组经过策划和相互配合的过程（包括启动、计划、实施、监控和收尾等五大过程）来实现的，但项目与一般的过程不同，是一种有时间限定的、临时性的独特过程，项目团队也是临时性组织，在项目寿命周期内存在。

美国项目管理协会（PMI）在《项目管理知识体系指南》中指出：项目管理就是把各种知识、技能、手段和技术应用于项目之中，以达到项目的要求。项目管理是通过应用和综合诸如启动、规划、实施、监控和收尾等项目管理过程来进行的。项目经理是负责实现项目目标的个人。

在项目管理过程中，应建立有关时间、成本和范围等的过程和产品质量指标，即项目



管理关键绩效测量指标,以监测、控制和评审项目管理过程与结果;应建立包括项目章程、项目计划等一系列的文档,并在项目收尾阶段总结归档,积累知识。

国内外许多组织的实践证明,有效地采用现代项目管理的技术和方法,至少可以在几个方面得到好处:促进顾客满意和忠诚、缩短项目周期、降低成本、减少风险、提高员工的专业技能和成就感以及企业的应变能力,为顾客、投资者、员工、供应商和社会创造价值。但是,管理项目并不意味着就能够使项目给经营带来最大利益,要实施有效的项目管理,有赖于企业高层的承诺和授权,项目团队的素质和努力,以及项目利益相关方的协同。

2.2 顾客需求分析

过程输出的接受者是顾客。六西格玛管理价值观之一是以顾客为中心。提高顾客满意度和降低经营成本是六西格玛管理所追求的。要提高顾客满意度首先应真正了解顾客的要求,并且要真正把握顾客的关键要求,除顾客明示的要求外,还包括顾客的潜在要求。并且,企业要运用质量功能展开等工具把顾客的关键要求进行多层次演绎分析,转化为产品的技术特性、设计要求等。这是目前世界先进企业的通行做法,也是六西格玛管理中最困难的一步,需要从以下几方面去实践。

2.2.1 顾客的识别

1. 顾客及其细分

顾客是指“接受产品的组织或个人”(ISO 9000 定义),这是一个广义的定义,如图2—4 过程要素分析所示,除了传统意义上的顾客外,过程输出(产品)的接受者还包括社区、股东、员工等其他利益相关方。

(1) 按照产品接受者的情况,顾客可以分为外部顾客和内部顾客两类。

- 外部顾客。是指组织外部接受产品或服务的组织或个人,如消费者、最终使用者、委托人、零售商等。

- 内部顾客。是指组织内部依次接受产品或服务的人员或部门。通常我们描述为过程输出的内部接受者。如产品生产流水线上的装配工,下道工序的操作者是上道工序的顾客,依此类推。

(2) 按照接受产品的提供方式,顾客可以分为中间顾客和最终顾客两类。

- 中间顾客。经销商、分销商或其他渠道伙伴。

- 最终顾客。组织产品或服务的使用者。

(3) 按照接受产品的时间顺序,可分为过去顾客、当前顾客和潜在顾客三类。

- 过去顾客。已接受过组织产品或服务的顾客。

- 当前顾客。正在接受组织产品或服务的顾客。

- 潜在顾客。尚未接受过组织产品或服务,但有购买期望的顾客群或者是竞争者的顾客。

(4) 按照顾客层次、业务量、市场区域、产品特征等,顾客还可进行细分,如大、

中、小型顾客,关键顾客和一般顾客等。

顾客要求是指顾客“明示的、通常隐含的或必须履行的需求和期望”(ISO 9000 定义),是一种特定的要求,由明示的或者是不言而喻的惯例及一般做法所考虑的顾客需求或期望组成,其潜在的、动态变化的特点使顾客要求越来越复杂而且多变。

细分顾客的根本目的在于:针对不同顾客群体采用不同的方法,倾听和了解不同顾客群的具体、差异化的要求,建立与之相适应的顾客关系,进而增进顾客满意和顾客忠诚。

2. 六西格玛项目的顾客

对于一个六西格玛项目而言,项目产品的接受者也就是项目的顾客。

首先,应当识别特定项目的顾客,包括内部顾客和外部顾客,并根据项目实际和顾客要求的差异程度进行进一步的细分,从中确定出关键的细分顾客群。例如,我们如果要实施一个改进蒸汽生产过程的六西格玛项目,基于图 2—4 过程要素分析可能就会识别该项目的顾客:蒸汽用户、热水用户、股东和管理层、社区、岗位员工。

其次,分析在该项目中各种顾客的需求是什么?有哪些影响因素?哪些是关键需求?如何将顾客的需求转化为项目目标?等等。

2.2.2 顾客满意与顾客忠诚

1. 顾客满意

顾客满意是指“顾客对其要求已被满足的程度的感受”(ISO 9000)。顾客抱怨是一种满足程度低的最常见的表达方式,但顾客不抱怨并不一定表明顾客很满意。即使达到顾客规定的要求,也不一定确保顾客很满意。

顾客满意的程度取决于顾客的价值观和期望值(认知质量)与所接受的产品或服务状况的比较。顾客的价值观决定了顾客的期望值,而组织提供的产品或服务(过程的输出)形成顾客的感知效果(感知质量),两者对比的符合程度决定了顾客是否满意。

由于顾客满意是顾客通过对一个产品(或一项服务)的可感知的效果(感知质量)与他的期望(认知质量)相比较后所形成的感觉状态,因此,满意水平是感知的效果和期望值之间的差异函数。顾客可以经历三种不同感觉状态中的一种。如果效果低于期望,期望得不到满足,则顾客不满意;如果可感知效果与期望相匹配,期望得到满足,顾客就满意;如果可感知效果超过期望,顾客就会惊喜,超出期望,直至产生忠诚(如图 2—6 所示)。

一般来说,顾客的要求是变化的,反映产品的质量特性也会随着变化,企业不仅要考虑顾客当前的要求,还应考虑顾客未来的要求,以适应顾客不断变化的要求。顾客的要求必须包括明示的、通常隐含的或必须履行的三个方面。由于顾客是根据感受来理解产品或服务的质量,所以很难提出准确的、量化的需求指标,这就需要组织主动了解顾客的期望,确定其真正的需求,从而定义自己的产品或服务的质量。

2. 卡诺(Kano)质量模型

日本质量专家狩野纪昭把质量依照顾客的感受及满足顾客要求的程度分成三种质量:基本质量、一元质量、魅力质量,如图 2—7 所示。

(1) 基本质量特性。当其特性不充足(不满足顾客要求)时,顾客很不满意;当其特性充足(满足顾客要求)时,顾客认为是应该的,充其量也就是不抱怨。

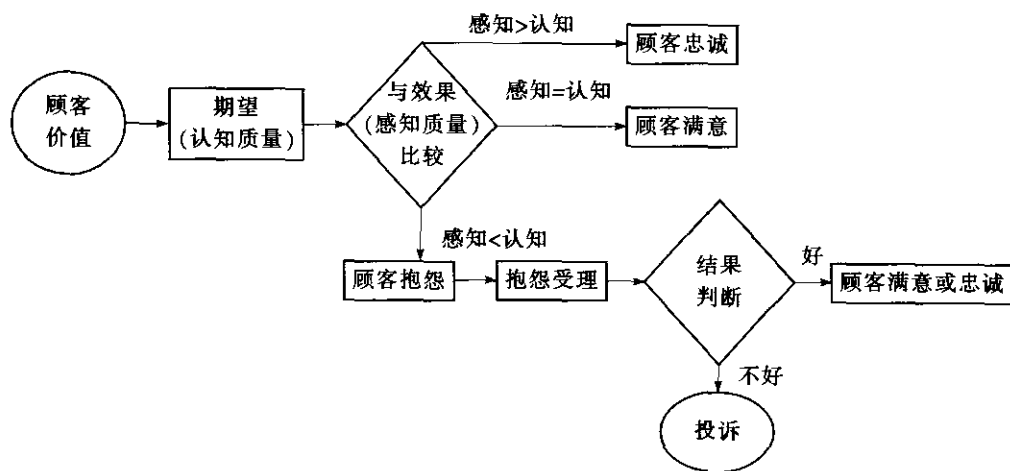


图 2—6 顾客满意示意图

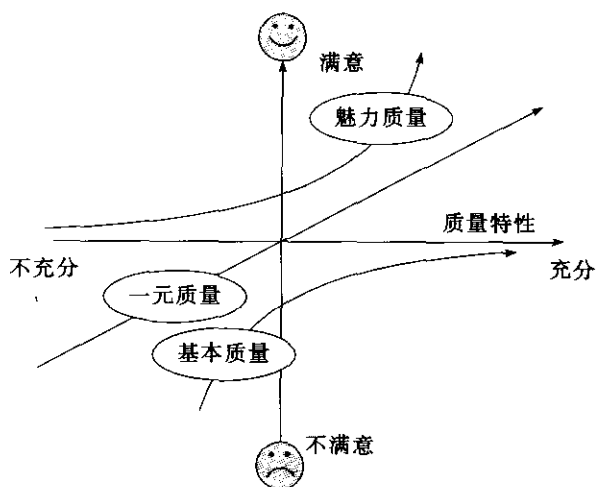


图 2—7 卡诺质量模型

基本质量特性是基线质量要求。以手机为例，手机通话这一特性是理所当然的质量，若手机不能通话，顾客肯定会很不满意，因为通话是手机最基本的质量，否则就不能称为手机了。同样，安全自然也是手机最基本的质量，这是隐含或必须履行的需求。若出现漏电、短路乃至爆炸、燃烧等，人们会非常不满意，因为安全是最起码的要求。

(2) 一元质量特性。当其特性不充足时，顾客不满意；充足时，顾客就满意。充足程度与满意度呈线性关系。

一元质量是质量的常见形式。同样以手机为例，“待机时间”、“接收信号”是手机的质量特性之一，如果待机时间长，可以方便顾客，减少顾客的充电时间，电池的寿命也长；接收信号强，手机的接通率就高，顾客就会很满意。相反，待机时间短、接收信号弱，顾客就会不满意。

(3) 魅力质量特性。当其特性不充足并且是无关紧要的特性（非重要特性）时，顾客

无所谓。当其特性充足时，顾客感到惊奇，并超出期望的满意。

有魅力的质量往往是质量的竞争性元素（或者是产品的魅力点、卖点），通常有以下特点：

- 具有全新的功能，以前从未见过；
- 性能极大提高；
- 引进一种以前没有见过的新机制、服务新政策等，顾客忠诚度得到极大提高；
- 一种非常新颖的风格。

如手机在有较强通话功能的同时，还具有电子词典功能、照相功能等。假如无这一功能，无关紧要，顾客无所谓，但有其功能，就能帮助顾客查询英语单词、拍照片等。一机多能，顾客就很喜欢，十分满意。因此，有魅力的质量是产品或服务具有竞争力（卖点）的质量保证。

但是，三种质量会相互转化，当有魅力的质量失去其特点和竞争性时，就开始变为一元的质量，甚至最后会变为理所当然的质量。

卡诺模型的三种质量划分，为六西格玛改进提供了方向。如果是基本质量，就要保证基本质量特性符合规格（标准），满足顾客的基本要求，项目团队应该集中在如何降低故障率上；如果是一元质量，项目团队关心的就不是质量特性是否符合规格（标准）问题，而是怎样提高规格（标准）本身，不断提高质量特性，促进顾客满意度的提升；如果是魅力质量，则需要满足顾客潜在需求，使产品或服务达到令顾客意想不到的新质量。项目团队应关注的是如何在维持前两个质量的基础上，挖掘顾客潜在需求，创造新产品和增加意想不到的新质量。

3. 顾客忠诚

顾客忠诚是指顾客在对某一产品或服务的满意度不断提高的基础上，重复购买该产品或服务，以及积极向他人推荐该产品或服务的表现。忠诚的顾客对组织及其产品或服务有明显的情感倾向性，能承受组织有限的涨价和抵制竞争者的降价或倾销价。

市场营销调查研究表明，保留现有顾客的成本比赢得新顾客的成本要低得多（1比6）。顾客忠诚程度越高，意味着销售成本的大幅度降低、营业收入和利润的稳定增长，也就意味着能够给企业带来更好的财务绩效。提高顾客忠诚度对企业的稳定和发展有着重要的意义，主要体现于：

- 忠诚顾客重复购买和推荐购买，能够带来更多的营业收入和利润；
- 忠诚顾客能够更积极地参与企业的产品、服务和管理的改进或提出建议；
- 忠诚顾客会主动宣传企业的品牌和形象，有利于市场的拓展；
- 借助于忠诚顾客的影响，有助于企业化解不满意顾客的投诉或抱怨。

2.2.3 顾客数据的收集

要真正把握顾客要求，首先要建立顾客反馈系统，倾听和了解“顾客的声音”。这是了解顾客要求的关键一步，必须给予充分重视。应注意“顾客的声音”中的“顾客”是一个广义的概念，除了产品使用者和潜在使用者，必要时还应包括分销商、产品维修人员等在产品寿命周期内关系密切的外部及内部的组织和人员。对于大型复杂产品的开发，“顾客的声音”将来自更多的方面。另外，环境法规、安全标准等国家和行业的法令、法规以



及标准、规范,由于构成了产品开发的约束条件,也应列入顾客要求的范畴。从企业长远的利益考虑,还必须深入地分析研究和考虑顾客的潜在需求和产品的更新换代规划、企业的发展方向和发展战略等。

可以从不同的顾客那里收集反馈的数据和信息,包括:当前感到满意的顾客,当前感到不满意的顾客(包括那些抱怨和没有抱怨的顾客),流失的顾客,竞争对手的顾客,潜在的顾客等。

用于收集“顾客的声音”的方法包括:

(1) 顾客调查。通过调查表、顾客代表座谈会等形式了解和归纳顾客对未来产品的要求。

(2) 顾客抱怨和投诉的处理。积极主动地收集顾客的抱怨和投诉,通过及时有效的处理,增进顾客满意和忠诚,改进产品、服务和管理,并了解顾客的真实要求。

(3) 同类产品质量跟踪和售后服务信息分析。了解现有产品中令顾客满意或抱怨的质量特性,将其转化为以顾客要求的形式表述。

(4) 将有关的政策法规等纳入顾客要求或作为产品开发的约束条件。

(5) 分析公司的战略和策略在产品开发中的贯彻方式,提炼出必要的顾客要求。

(6) 产品发展现状与趋势分析。通过媒体及专业杂志等手段收集信息,分析处理,从全局上把握产品发展方向,结合 QFD 小组的头脑风暴会议,对上述方式得出的顾客要求进行筛选和补充。

下面介绍几种常用的顾客调查方法。

1. 书面邮件调查

最常用的调查方法是邮寄书面问卷进行顾客满意度调查。许多公司定期地寄出调查问卷,把结果制成表格,为组织的所有领域提供反馈。有些组织按年度进行满意度调查,另一些组织则每月随机抽取顾客发送问卷。

实施邮件调查比电话调查和专人访问需要的时间长,但这是在所有的调查技术中强迫性最小的技术,因为回答者可以自己选择时间完成调查。邮件调查比专人访问费用低,通常比电话访问的费用也低。返回率低是邮件调查主要担心的问题。预先通知、撰写良好的介绍信、追踪调查在问卷中没有回答的问题,是提高返回率的有效技巧。在满意度调查中,一般可以达到 30%~50% 的返回率。

这种调查成本低,有助于统计分析。邮件调查中既可以包括定量提问,也可以包括定性提问。但开放式定性提问数据的质量不如电话访问或专人调查好,因为在邮件调查中,没有访问人员对不完整的回答或模糊的回答提出质疑或做出澄清。

为提高效率,书面邮件调查必须比较简短,提问不能超过几页,以减少接受调查者放弃调查的可能性,这限制了提问的数量。书面邮件调查的回应不是随机的,因为最满意的和最不满意的顾客比没有强烈感受的人做出回应的概率更大。为了克服这个问题,公司可以通过电话与随机抽取的顾客联系,测量和纠正书面调查结果中的偏差。

2. 电话调查

许多公司利用电话调查补充或替代书面调查。电话调查对于收集有关公司整体的信息,或关于如销售和服务这样的职能领域的信息特别有效。访问人员打电话给客户,识别需求并了解公司满足顾客需求方面做得如何。电话调查便于管理,并且调查结果可以很快

使用。与邮件调查相比,不回答的较少,由于返回率较高,电话调查会更准确。样本量可以较小,公司可以肯定地知道回答调查问题的人。电话调查比书面调查有更大的灵活性,因为有经验的调查员可以偏离准备好的访问指南,追寻顾客特别感兴趣的领域。如果有经过适当培训的访问者,还可以保证开放式评论数据有较好的质量。电话调查中需要注意的是确保所有的访问人员按照一致的方式实施调查,以减少由于访问人员的影响产生的偏差。有时,通过电话请受访者评价复杂概念比较困难。

电话访问有时间限制,调查时间一般比邮件调查长,但比专人访问时间短。电话调查比执行良好的书面调查费用高,但比专人访问费用低。电话访问注定会打扰被访问者,并且阻力日益增加,特别是来自消费者的阻力。因此,广泛地采用有声邮件和录音电话成为必需,通过多种努力以获得合理的回收率。最后,在某些消费品市场上,顾客可能没有列出电话号码或没有电话。然而,即使在这种不利条件下,电话调查仍越来越普遍,成为销售和顾客服务质量改进过程出色的输入来源。

3. 专人访问

专人访问或面对面访问是最强有力的信息收集技术之一。它可用于识别对顾客最重要的事项,以便在书面或电话调查中强调这些事项。专人访问还是一种收集定性信息的高效方法。一个训练有素的访问者可以迅速地提取顾客的关键信息。访问人员可以追寻最重要的领域,探究细节,讨论新见解和方法。无论是从内部顾客还是外部顾客获得信息,专人访问都是一种有效的工具。

专人访问的优点还有:如果进行了足够数量的调查,定量的结果就可以明确地反映被调查人群的意见。与焦点小组调查方法不同的是,被调查人的回答表明了他自己的意见,而不以任何方式受小组多数意见或占支配地位的参与者的影响。

实施大量的专人访问成本很大,特别是当公司拥有大量的顾客或顾客分布较分散时更是如此,因为需要出差,导致了大量的成本和时间付出。专人访问还要求使用训练有素的、没有偏见的访问人员,通常这意味着需要利用公司以外的人员。

4. 焦点小组

焦点小组是定性的小组讨论,通常有8~10个参与者,一般按照一定的细分准则,如产品使用的层面、年龄、收入、性别等。讨论之前准备好问题提纲作为讨论指南。各个问题由焦点小组主持人提出并进行推动。焦点小组可以识别出重要的反复发生的主题。这些主题与顾客对产品或服务的满意程度、关键顾客需求以及绩效改进的机会有关。

焦点小组的好处是回答每个问题的成本较低。焦点小组通常作为进行书面或电话调查的第一步,但也单独用于获得顾客对销售和营销计划的反馈。内部顾客焦点小组可以为业务改进、制造、开发、行政管理甚至销售过程提供输入。

2.2.4 顾客数据的分析

收集顾客数据只是第一步,还需要分析顾客数据。那么如何分析顾客反馈的数据呢?“顾客的声音”提供了原始的顾客需求,应加以规范化,进行分解、归并、筛选,并进行确认和分级,通过调查分析,确定各顾客要求的重要程度。

对顾客要求的表述有一定要求,主要是:

- 用语简洁,无歧义;



- 一项顾客需求只表达一个特定的意思；
- 不把对应的工程技术解决方案纳入顾客需求；
- 便于工程人员理解；
- 同一级别的顾客需求彼此独立，内容无覆盖及交叉。

为便于工程人员理解，应对用语进行规范化处理，使工程人员可据此提出相应的解决方案。为了建立顾客需求间的层次关系，亲和图法是一个形象有效的手段。具体步骤如下：

(1) 把每项顾客需求分别填在一张小纸条上，去掉内容重复的纸条，再把所有纸条排列起来；

(2) 把内容相近的纸条聚为一堆，起一个可以概括其内容的名字，作为高一级的顾客需求，另外写在纸条上；

(3) 将新写的纸条按内容相近程度聚堆、命名，作为更高一级的顾客要求，并另外写在纸条上；

(4) 如有必要，继续上一过程，直到顾客需求被系统而分层次地组织起来。

对于复杂一些的产品，为了深入细致地分析顾客对产品的要求，可能需要建立多级顾客需求，将它们填入需求质量展开表中（见表 2—1）。

表 2—1 某无线电远距操纵装置的需求质量展开表

1 次	2 次
1. 操作方便	1.1 携带方便
	1.2 操作中不感觉疲劳
	1.3 对操作易于理解
	1.4 能轻松愉快地操作
	1.5 能处理复杂问题
⋮	⋮
5. 安全	5.1 无误动作
⋮	⋮

对收集、整理的顾客数据，可采用图表方法、统计方法、定性的工具来分析和理解顾客的需求。

2.2.5 关键顾客要求的转化

应用质量功能展开（QFD）将顾客（用户、使用方）对产品的需求进行多层次的演绎分析，利用量化评估方法，识别关键顾客需求（critical customer requirements, CCR）并逐层转化为产品的设计要求、零部件特性、工艺要求、生产要求，以指导产品的设计和保证产品的质量（详见第 10 章）。

2.3 经营结果

2.3.1 过程绩效度量指标

过程绩效即过程输出的绩效，用来度量过程绩效的数据包括离散型数据（也称计数值

数据或属性值数据)和连续型数据(也称计量值数据)。

1. 离散型数据度量指标

六西格玛管理中常用的过程绩效离散型度量指标有:单位缺陷数、机会缺陷率和百万机会缺陷数等计点值指标,最终合格率、一次合格率、流通合格率等计件值指标。

这里,先介绍以下一些基本概念。

- 单位(unit):过程加工过的对象,或传递给顾客的一个产品及一次服务。
- 缺陷(defect):产品(或服务)没有满足顾客的要求或规格标准。
- 缺陷机会:单位产品上可能出现缺陷的位置或机会(缺陷机会与缺陷的度量应保持一致)。一般在确定缺陷机会数时可以分三步:列出可能出现缺陷类型的清单;判定哪一种缺陷是顾客特别关心的、真实发生的缺陷;用其他标准来复核产生缺陷机会的数目。

(1) 单位缺陷数(defects per unit, DPU)。单位产品上平均的缺陷个数。它由以下公式计算:

$$DPU = \frac{\text{检测发现的缺陷数}}{\text{抽取的单位产品数}}$$

例3-1

抽取100块电路板,检查出5个缺陷,则

$$DPU = \frac{5}{100} = 0.05$$

(2) 机会缺陷率(defects per opportunity, DPO)。每次机会中出现缺陷的比率,表示样本中缺陷数占全部机会数的比例。它由下式计算:

$$DPO = \frac{\text{缺陷数}}{\text{产品数} \times \text{单位产品的平均缺陷机会数}}$$

假定这100块电路板中,每一个电路板都含有100个缺陷机会,若在制造这100个电路板时共发现21个缺陷,则

$$DPO = \frac{21}{100 \times 100} = 0.0021 = 0.21\%$$

(3) 百万机会缺陷数(defects per million opportunity, DPMO)。DPO常以百万机会的缺陷数表示,即

$$DPMO = DPO \times 10^6$$

上例中,0.0021的DPO即为2100的DPMO。



例 1-4 某物料清单可能会发生四种错误，即 4 个机会，它们是：有多余项目、缺少项目、项目选错、参数写错，则

$$DPMO = \frac{\text{总的缺陷数} \times 10^6}{4 \times \text{抽取的单位产品数}}$$

假如在 1 376 张物料清单上发现 41 个缺陷，则

$$DPMO = \frac{41}{1\,376 \times 4} \times 10^6 = 7\,449$$

即每百万个机会中有 7 449 个缺陷。

(4) 最终合格率 (process final yield, PFY)。过程最终的合格率通常是指通过检验的最终合格单位数占过程全部投产单位数的比例。但是，此种质量评价方法不能计算该过程的输出在通过最终检验前发生的返工、返修的损失。这里我们把返工等叫做“隐蔽工厂” (hidden factory)。隐蔽工厂不仅出现在制造过程，同时也出现在服务过程。

(5) 一次合格率 (first time yield, FTY)。第一次就把事情做对，由没有经返工返修便通过的过程输出单位数而计算出的合格率。

(6) 流通合格率 (rolled throughput yield, RTY)。对于一个串行生产过程，假定各过程彼此独立，则其流通合格率为每个子过程的 FTY 的乘积，表明由这些子过程串联构成的大过程的一次合格率。

$$RTY = FTY_1 \times FTY_2 \times \cdots \times FTY_n$$

式中， FTY_i 是各子过程的一次合格率； n 是子过程的个数。

例 1-5 某产品的生产过程有三道工序：首先是下料，一次合格率为 95.5%；其次是车削，一次合格率 97%；第三是钻孔，一次合格率为 94.4%。请计算该过程的流通合格率。

$$\begin{aligned} RTY &= FTY_1 \times FTY_2 \times \cdots \times FTY_n = FTY_1 \times FTY_2 \times FTY_3 \\ &= 95.5\% \times 97\% \times 94.4\% = 87.4\% \end{aligned}$$

流通合格率就是一种能够找出隐蔽工厂的地点和数量的测量方法。

最终合格率与流通合格率的区别是，RTY 充分考虑了过程中子过程的存在，即隐蔽工厂的因素。若过程有 n 个子过程，而子过程的合格率分别为 y_1, y_2, \cdots, y_n ($i = 1, \cdots, n$)，则 $RTY = y_1 y_2 \cdots y_n = \prod_{i=1}^n y_i$ ($i = 1, \cdots, n$)，这样就充分考虑了过程中各子

过程的因素，比较客观地反映过程运作的实际。而最终合格率没有考虑过程的返修，因而掩盖了隐蔽工厂，从而掩盖了质量问题。但是在实际生产过程中，不少企业计算的合格率是最终合格率，并没有统计流通合格率，这是需要改进的地方。

图 2-8 流通合格率 (RTY) 计算实例

某生产过程投产 1 000 单位产品, 包含五个子过程, 每个子过程都有独立的合格率 (y_i) ($i=1, \dots, 5$), 分别计算 PFY 和 RTY (见图 2—8)。

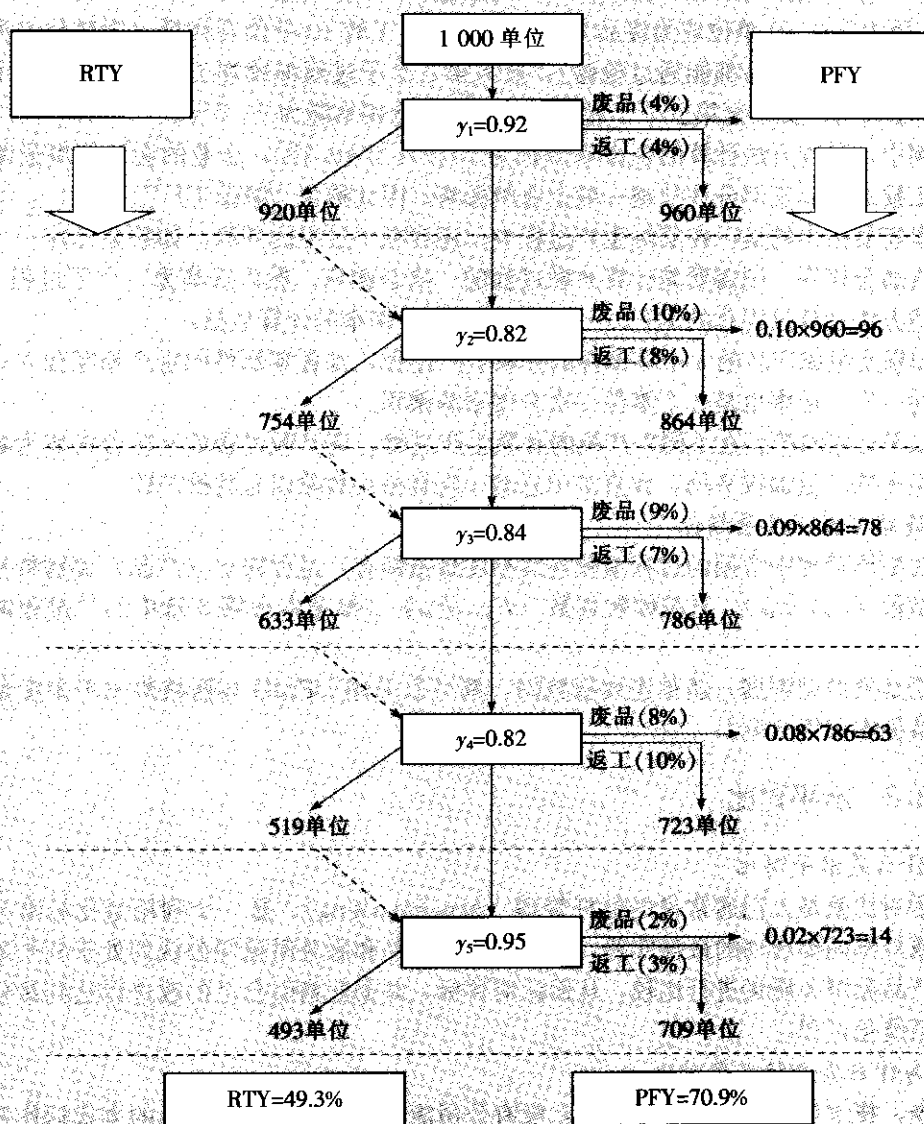


图 2—8 流通合格率 (RTY) 计算实例

本例中 RTY 也可以这样计算:

$$RTY = y_1 y_2 y_3 y_4 y_5 = 0.92 \times 0.82 \times 0.84 \times 0.82 \times 0.95 = 0.493 = 49.3\%$$

在图 2—8 左边, 我们发现通过每一个子过程的单位数自 1 000 逐渐递减至 920, 754, 633, 519, 直至 493 单位。当 1 000 单位投入第一个子过程时, 只有 920 单位一次合格。在图的右边, 我们可以看到另外 80 单位的去向: 40 单位成为废品, 40 单位需要返工并重



新回到生产循环。在第一个子过程中,共 920 单位一次合格而直接进入第二个子过程,其一次合格率为 82%,得到 754 个一次合格的在制品。如此继续运作,直到完成所有五个子过程,得到 493 个全过程均不需返工的合格品,其流通合格率(RTY)为 49.3%。

图 2—8 的右边则说明如何计算 PFY。经过第一个子过程,原 1 000 个投入单位减少为 960,因为其中 40 单位成为废品无法返工。可返工的 40 单位合格后(为简化起见,我们假定所有返工的产品都能通过检验),投入第二个子过程继续加工,这样第一个子过程的 PFY 为 96%。类推至第五个子过程,过程最终合格率降为 70.9%。

本例中,两种方法计算的合格率分别为 70.9%与 49.3%,企业的会计制度通常未能显示返工损失,返工部分往往被一般企业所忽略,因此称为“隐蔽工厂”。

这里需要指出的是,在实际生产过程中,还存在子过程的并联。如果要计算一个混联的过程流通合格率,则需要先计算并联过程的一次合格率,然后视其为一个子过程 FTY,采用上述方法计算过程总的 RTY。并联过程一次合格率的计算包括:

- 如果是组成产品的 n 个零部件的并联生产过程,且各零部件的生产相互独立,则并联过程的一次合格率也是各个零件一次合格率的乘积。
- 如果是相同产品在不同生产线的并联生产过程,则并联过程的一次合格率为各机台一次合格率按产量加权平均,或直接用总的一次合格品数除以总的投产数。

2. 连续型数据度量指标

六西格玛管理中常用的过程绩效连续型度量指标有:过程输出(产品)关键特性的过程能力指数(C_p , C_{pk})、过程绩效指数(P_p , P_{pk}) (参见本书第 5 章中的“过程能力分析”)。

无论是离散型数据,还是连续型数据,都可以采用 DPMO 与西格玛水平来度量过程绩效(参见本书第 5 章)。

2.3.2 水平对比

1. 什么是水平对比

水平对比就是人们通常说的标杆管理(benchmarking),是一个利用量化标准寻找行业内部或外部同类活动的最佳实践,即将过程、产品和服务质量同公认的处于领先地位的过程、产品和服务质量进行比较,从而认清目标,并据此指引自己的改进以达到最佳绩效水平的系统化过程。

2. 为什么要进行水平对比

当今,技术和商业竞争十分激烈,没有公司能够以孤立、浪费成本的方式运作还能够获得利润。客户需要的是反应及时、高效率、高质量的产品和服务。行业内的竞争使得企业只有做到最好,才能长久生存与发展。事实证明,水平对比法的作用是寻找最佳水平,并与企业实际对照,找出差距,促进质量改进,使企业在竞争中立于不败之地,甚至成为同行之冠。水平对比法是一种非常有效的提升企业竞争力的方法。施乐公司率先在美国运用标杆对比,分析、学习和借鉴日本复印机企业的生产效率,大幅度提升其绩效。在实施六西格玛的过程中,许多企业通过水平对比,发现自身存在的改进机会,为六西格玛项目确定方向。

水平对比超越了单纯竞争的含义,其目的不仅在于显示竞争地位本身,“比”完还得

“学”，更重要的是去发现、探索最佳绩效背后的最佳实践，用以借鉴和学习，为重大的、突破性的改进和创新以及战略环境分析提供支持，最终实现“赶”和“超”的目的。

3. 水平对比的作用和好处

水平对比的作用和好处如下：

- 通过与先进企业的对比，明确改进机会；
- 利用外部资源确立有效的行动目标；
- 引入最好的行为方式；
- 建立正确的绩效测量方法；
- 激励和激发团队；
- 减少变革的障碍；
- 使组织变得更具有竞争力；
- 使组织可以更好地满足最终顾客的需求。

4. 水平对比的主要应用范围

(1) 公司战略。水平对比会成为公司整体运作的战略指导。水平对比的结果有助于公司的战略投资、市场运营、质量管理及制定技术战略，通过水平对比可以指导公司进行战略策划、制定和调整战略目标及关键绩效指标。水平对比在大型、中型或是部门的项目中均可应用，范围很广。

(2) 顾客需求。在产品和服务中满足客户的关键需求至关重要。选择哪一方面去做水平对比，取决于一个企业关键成功因素（即一个企业成功的主要决定因素）。而对关键成功因素的理解是：顾客为什么要选择我们的产品？他们为什么信任我们？主要原因是我们是所有竞争者中最好的，或是我们要力争达到这一点。因此，顾客所关注的关键需求也是水平对比的重点。

5. 水平对比的一般模型和步骤

水平对比的一般模型和步骤如图 2—9 和图 2—10 所示。

我 们	
1. 对比什么？	2. 我们是如何做的？
3. 谁是最好的？	4. 他们是如何做的？
他 们	

图 2—9 水平对比的一般模型

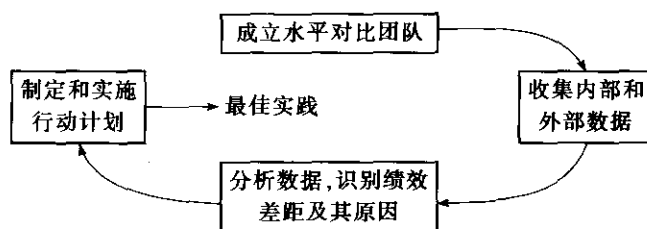


图 2—10 水平对比的一般步骤



6. 水平对比在六西格玛管理中的应用

首先,水平对比是黑带学员必须了解并基本掌握的一套理论和方法,它在六西格玛管理中有广泛的应用,主要体现在以下方面:

(1) 用于六西格玛管理的导入。组织首先应了解和学习业界以及世界上导入六西格玛管理最成功的企业(如通用电气公司等)的做法,然后再决定自己如何做,主要从以下几点思考:

- 我们为什么要导入六西格玛管理?
- 业界标杆企业为什么要导入六西格玛管理?
- 它们是在什么背景下导入六西格玛管理的?
- 它们是以什么方式导入六西格玛管理的?
- 它们是如何实施六西格玛管理的?
- 六西格玛管理为它们带来的绩效如何?
- 它们的成功经验是什么?
- 哪些值得我们借鉴?
- 根据我们的现状,我们应该如何做?

(2) 用于六西格玛项目的选择。六西格玛项目主要来源于企业的发展战略、顾客的需求、内部流程中存在的问题、竞争对手的比较等方面,因此,需要在行业内外以及组织内部进行水平对比,找出突出的问题,确立改进方向和具体的六西格玛项目(详见第3、4章)。

(3) 用于确定和达成六西格玛项目的目标。一个项目目标确立的正确、科学与否是影响项目成功的关键因素,因此,要使项目的目标科学合理,可从以下几个方面考虑:

- 这个项目问题目前可量化的绩效水平如何?
- 本组织内部不同业务区域之间此方面的现状如何?差异多大?
- 是什么原因导致内部不同业务区域之间的差异?
- 对于此项目问题,业界的最佳水平是多少?
- 我们处在业界的什么水平?
- 为什么我们与业界有差异?可能的原因在哪里?
- 标杆企业在这方面的做法如何?
- 我们目前的六西格玛项目目标应定在什么水平才合适?
- 近几年将如何去努力?
- 什么时间能达到业界的最佳水平?

2.3.3 财务收益

1. 经济性——六西格玛的核心特征

目前的经营环境要求我们增进对大质量的理解,使企业的质量活动同时为顾客、员工、所有者和整个公司创造价值和经济效益。六西格玛正是这样一种管理实践,它注重质量的经济性。当我们投资改进有缺陷的过程时,不但降低了不良质量成本,而且上升的顾客满意度又可以挽回部分原来失望的顾客。同样,它又会促进顾客对其产品的购买,从而带来年收入的增加。

因此，六西格玛管理的核心特征是：高顾客满意度和低资源成本。ISO/TR10014《质量经济性指南》给出了质量经济性管理的两个基本要素：提高顾客满意度和降低资源成本。而这恰恰是六西格玛管理的基本原则和核心特征（见图2—11）。

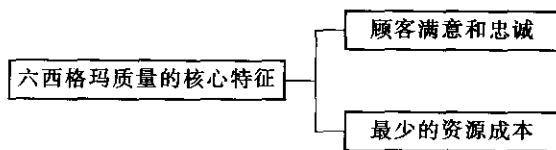


图 2—11 六西格玛质量的核心特征

降低资源成本是六西格玛管理的一个重要原则。六西格玛追求零缺陷，其最终目的在于降低风险，一方面可以降低顾客购买产品或服务风险，另一方面也降低了产品或服务提供者的风险。换言之，应用六西格玛来降低风险意味着绩效的提高，如质量、能力、周期时间、库存以及其他关键因素。这样，客观上降低了企业为纠正缺陷和防范风险所需的成本，从而实现了真正意义上的低资源成本。

2. 六西格玛管理与财务收益

通过六西格玛管理，顾客和企业可以同时获得满意。对顾客而言，以最低可接受的价格及时获得了最好的产品；对企业而言，则是以最小的成本和最短的周期实现了最大的利润。当这些全部实现时，“质量”这一词汇才对企业有了真正的意义。

（1）增加收益。从增加收益方面看，实施六西格玛管理对顾客而言，可以减少费用、改进适用性；对企业而言，可以提高利润和市场占有率。实现途径如下：

- 开发新的、特有的产品和服务，或改进现有的产品和服务，并减少进入市场的时间。
- 开拓现有产品或服务的市场，提高满意度和忠诚度，赢得信誉，增加市场份额。

下面列出在缩短周期时间、加快资金周转以及提高准时交付率方面的增加收益的计算方法：

- 缩短周期时间，加快资金周转（应收款、库存）。

$$\text{效益} = \frac{\text{年销售额} \times \text{加权资金成本} \times \text{缩短的天数}}{360}$$

式中：

$$\text{加权资金成本} = \frac{\text{借款利率} \times (1 - \text{所得税率}) \times \text{借款}}{\text{借款} + \text{权益}} + \frac{\text{权益报酬率} \times \text{权益}}{\text{借款} + \text{权益}}$$

- 提高准时交付率。

$$\text{效益} = \frac{\text{订单总金额} \times (1 - \text{平均交付率}) \times \text{资金回报率} \times 365 \text{ 天}}{\text{交付期}}$$

某六西格玛项目将从送货到货款回笼的平均时间由 30.6 天降低到 26.6 天。如果公司 2005 年的预计销售额为 400 000 000 元，预计净利润 50 751 000 元，期初权益 123 839 053 元。公司的借款利率为每年 5.94%，公司的权益报酬率为每年 40.98%，所得税率为每年 12%。该项目全年的财务收益是多少？

$$\text{权益报酬率} = \frac{\text{净利润}}{\text{期初所有者权益}}$$

$$\text{项目收益} = \frac{2005 \text{ 年销售额} \times \text{加权资金成本} \times \text{缩短的天数}}{360}$$

经过计算,该公司的加权资金成本为每年 26.89%,由于该项目缩短了回笼天数 4 天,则经核算全年共计节约 1 195 111.11 元。

(2) 降低成本。从降低成本方面看,实施六西格玛管理对顾客而言可以减少购置费、运行费、保养费、停机损失费和修理费以及可能的处置费用;对企业而言,可以降低产品寿命周期全过程的缺陷,包括不满意的产品返工、返修、更换、生产损失、担保和现场修理等发生的费用,以及承担产品责任和索赔风险。实现途径如下:

- 降低符合性成本:重新设计以获得更高的过程能力,提高现行过程能力和提高技能等。
- 降低非符合性成本:减少浪费,减少废品和返工返修,减少停工损失,降低超支,减少污染和顾客退货等。

下面列出在降低成本方面的计算方法:

- 降低成本。

$$\text{效益} = \text{产量} \times \text{单位成本降低额}$$

- 降低不合格率。

$$\text{效益} = \text{产量} \times (\text{改进后 RTY} - \text{改进前 RTY}) \times \text{单位成本}$$

图 2—12 显示了六西格玛的质量经济性,我们可以看出实施六西格玛取得六西格玛质量效益分两方面:一是通过提高产品的质量特性,优质优价和扩大市场份额,获得更高的收益;二是通过降低缺陷,缩短周期,降低出厂前的废品和返工返修品以及出厂后的担保成本,进而花费更低的总成本。

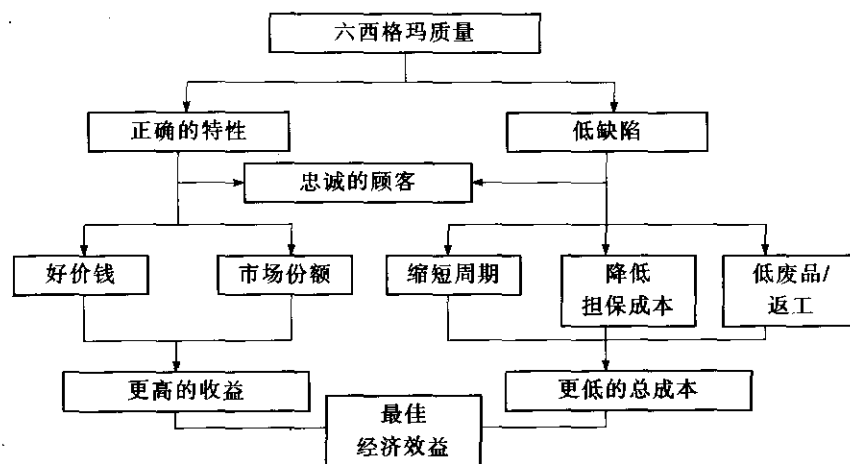


图 2—12 六西格玛质量与财务结果

在测算六西格玛效益时,存在着硬效益和软效益两类。硬效益能够用货币表达,如运营成本、生产成本、销售费用、管理费用、人事成本的减少,由于产量增加导致的收入增

加等；而软效益往往难以转化为金钱的语言，如员工满意度和顾客满意度的增加，环境保护、职业健康和安全的改进，法律法规的符合等。

3. 基本财务模型

由于投资可以获得升值，因此财务理论认为货币有时间价值。货币时间价值是财务决策理论的基础，对应于今天可用的一定金额的钱就是“现值”（PV），在未来可收回或可支付的一定金额就是“终值”（FV）。

（1）终值与现值。六西格玛管理可以提高企业经济效益，实际上六西格玛项目也是“投资”，当然也有“回报”，即“终值与现值的差额”。

$$FV = PV \times (1+i)^n$$

式中， i 为利率； n 为有效（计息）期。

（2）净现值（NPV）。评价六西格玛项目的成本和收益时，成本和收益通常是以现金流的方式产生的，而不是一下子全部出现。一般现金流不是均匀的，每次现金流的金额可能都不相等。净现值（NPV）是指将投资项目在有效期内或寿命期内的净现金流量按一定的折现率全部折算到零期的累计现值之和，其表达式为：

$$NPV = \sum_{t=0}^n (CI_t - CO_t)(1+i_0)^{-t}$$

式中， CI_t 为第 t 年的收益（现金流入）； CO_t 为第 t 年的费用支出或投资（现金流出）； $(CI_t - CO_t)$ 为第 t 年的净现金流量； i_0 为项目的收益率。

如果净现值大于零，六西格玛项目在经济上是可行的。净现值越大，六西格玛项目的回报就越大。净现值小于零，则该项目的经济性不好，不可行。

（3）内部收益率（IRR）。内部收益率是考虑货币的时间价值的财务分析评价方法，它是根据方案的现金流量计算出的方案本身的投资报酬率，是指能够使未来现金流入量现值等于未来现金流出量现值的贴现率，或者说是使投资方案净现值为零的贴现率。其计算方法通常需要使用“插值法”：首先估计一个贴现率，用它来计算方案的净现值。如果净现值为正数，说明方案本身的报酬率超过估计的贴现率，应提高贴现率后进一步测试；如果净现值为负数，说明方案本身的报酬率低于估计的贴现率，应降低贴现率后进一步测试，直至寻找出使净现值接近零的贴现率，即为方案本身的内部收益率。

如果某团队改进方案用 18% 的贴现率计算，财务净现值为 -499 元；用 16% 的贴现率计算，财务净现值为 9 元，则项目的财务内部收益率为：

$$16\% + \{ (18\% - 16\%) \times [9 / (9 + 499)] \} = 16.04\%$$

有关净现值（NPV）和内部收益率（IRR）的计算可用 Excel 软件来实现。

（4）投资收益率或回报率（ROI）。

$$ROI = \frac{\text{项目预计收益}}{\text{项目预计成本}} \times 100\%$$



某团队项目效果核算如表 2—2 所示：

表 2—2 某项目六西格玛改进效果核算表（减少不合格品的收益）

年产量	10 000 件
单件成本	1 200 元
不合格率降低	由原来的 10% 降到 2%
报废的单件产品价值（按回收废品处理）	100 元
年节约额	$8\% \times 10\,000 \times (1\,200 - 100) = 880\,000$ 元
项目改进成本（第一年的咨询、培训和夹具制作等）	20 000 元
第 1 年 ROI	$(880\,000 / 20\,000) \times 100\% = 4\,400\%$

4. 质量成本和不良质量成本

（1）质量成本概念的演进。质量成本的概念是 20 世纪 50 年代由美国质量管理专家朱兰、费根堡姆等人首先提出的。其定义是：为了确保和保证满意的质量而发生的费用以及没有达到满意的质量所造成的损失。它是企业生产总成本的一个组成部分。

传统的质量成本由以下四部分构成：

1) 预防成本。指为了预防故障所支付的费用，一般包括：质量策划费用、过程控制费用、顾客调查费用、产品设计评审和鉴定费用、体系研究和管理费用、供应商评价费用、培训费用以及其他预防费用。

2) 鉴定成本。指为评定质量要求是否被满足而进行试验、检验和检查所支付的费用，一般包括：检验和试验费用、计量服务费用、质量审核费用以及顾客满意度调查等其他鉴定费用。

3) 内部故障（损失）成本。指产品在交付前，不能满足质量要求所造成的损失，一般包括：报废损失、返工返修损失、降级损失、停工损失、质量事故处理费用、纠正措施费用以及其他内部故障费用。

4) 外部故障（损失）成本。指产品在交付后，不能满足质量要求所造成的损失，一般包括：产品售后服务及保修费用、顾客投诉处理费用、产品责任费用等外部故障费用。

传统的质量成本在故障成本方面的收集范围过窄，而且忽视了三个方面的“隐藏成本”：一是工作和过程质量（特别是非生产过程的工作和过程质量）故障损失；二是尽管生产着合格产品但过程低效率的损失，这种低效率来自过大的资源消耗和非增值的过程步骤等；三是由于不良质量而导致的销售损失，包括因质量问题以及产能不足而导致的顾客流失损失等。

这导致人们对故障成本的过低估计，以至于认为在低于 100% 合格水平的某处存在着一个“适宜的质量成本”。当我们将三个方面的“隐藏成本”分别加入内部、外部故障成本后，可以得出如图 2—13 所示的现代质量成本模型。

组织在推行质量成本管理的初期，注重通过增加预防和鉴定成本来降低内外部故障成本，随着质量水平和经营管理成熟度的提升，组织将着力于通过增加预防成本以降低鉴定成本和故障成本。完美无缺无疑是我们未来的向往，从长远和总体的角度看，质量水平越高，其质量成本越低，即所谓“质量是免费的”。

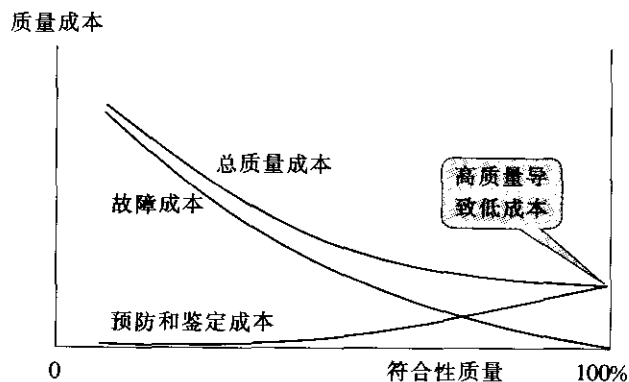


图 2—13 现代质量成本模型

资料来源：朱兰：《朱兰质量手册》（第5版），北京，中国人民大学出版社，2003。

(2) 符合性成本、非符合性成本与不良质量成本。质量成本根据其性质又可分为符合性成本（cost of conformity）和非符合性成本（cost of nonconformity）。符合性成本是指在现行过程无故障情况下完成所有顾客需求所支付的费用；非符合性成本是由于现行过程的故障而造成的损失。符合性成本包括预防成本和鉴定成本（预先检验预防部分）；非符合性成本包括鉴定成本（查明故障原因部分）和故障成本（内部+外部）。

六西格玛管理要求降低经营资源成本，就是要降低非符合性成本和符合性成本。非符合性成本是不增值的，而符合性成本由增值和不增值两个部分组成。要通过有效的手段，找出不增值的部分加以改进，即消灭所谓的隐蔽工厂，从而降低成本。于是人们又构筑了不良质量成本的概念（见表 2—3），即

不良质量成本=预防成本和鉴定成本中的不增值部分+现代质量成本中的故障成本

表 2—3 不良质量成本的构成

不良质量成本	符合性质量成本	预防成本（不增值的部分）
		鉴定成本（预先检验预防的不增值部分）
	非符合性质量成本	鉴定成本（分析故障原因）
		故障成本（内部和外部损失）

这一概念已广泛地用于六西格玛管理，是指由于质量不良而造成的成本损失，或者说是由于我们没有第一次就把事情做对、做好而额外付出的成本。由于质量不良而造成的成本损失十分惊人，遗憾的是这部分成本往往不为人们所知。

我们已经讨论过这样的问题：六西格玛值得推行吗？问题的反面是：假定不推行六西格玛，代价如何？维持现状，不作改进的成本又如何？对组织而言，维持现状或许是正确的选择。但组织应该先比较维持现状的成本与改进现状的成本，才能做出正确的决定。因此，即使是符合性成本，也应该认真加以研究，探索改进的空间。我们应该同时考虑对效率（efficiency）和有效性（effectiveness）的改进。

不良质量成本的概念将有助于我们挑选或者界定六西格玛项目，因为在成本节省方面是否有潜力是我们选择项目的重要考虑事项之一。在项目初期进行不良质量成本测算，可以作为项目选择的依据，也有利于明确项目改进的方向，同时为项目总结时评估项目财务



效益奠定基础。如果项目团队能够根据各项目的不良质量成本金额大小制作排列图，然后制定改进的优先顺序，那将是十分有效的，当然要花费很大的工夫。

(3) 不良质量成本科目与费用归集。

1) 预防成本（非增值部分）。为预防故障发生所支付的费用，其中一部分由于不增值，或者效果不理想，可计入不良质量成本。常见的预防成本科目有：

质量策划费用。有关部门或人员用于策划所需时间的费用支出。例如规划质量体系的具体细节所需的时间；根据产品设计和顾客对质量的要求，编制用于材料、工序和产品质量控制的方法、程序、指导书等所需的时间。质量策划费用还包括从事其他质量策划工作所需时间的费用支出，如可靠性研究，试生产质量分析，为编制试验、检验和工序控制的指导书等所需时间。

过程控制费用。包括为质量控制和改进现有过程能力的研究和分析制造过程（包括供应商的制造工序）所需全部时间的费用支出；为有效实施或执行质量规划而对车间工作人员提供技术指导所需的费用支出；在生产操作过程中自始至终进行控制所支出的费用。

顾客调查费用。为了掌握顾客的需求所开展的相关调查研究和分析所花费的费用。

质量培训以及提高工作能力的费用。用于改进和提高质量水平所花费的相关费用（不包括指导员工达到标准熟练程度的训练费）。

产品设计鉴定/生产前预评审费用。为了鉴定设计的质量、可靠性和安全性而评价试制产品、审批产品规范时所支出的费用。此外，还包括生产前预评审费。

质量管理体系的研究和管理费用。用于整个质量管理体系的设计和管理费用，以及辅助费用。

供应商评价费用。为实施供应链管理而对供方进行评价的活动费用。

其他预防费用。包括质量及可靠性组织机构的行政管理费用，以及零缺陷计划、厂房设备维护等预防性措施费用。

2) 鉴定成本（非增值部分）。为评审质量是否满足要求而进行的试验、检验和检查所支付的费用，其中相当大的部分不增值，应计入不良质量成本（强制性标准规定的以及顾客明确要求并愿意付钱的检验、试验成本，可视为增值的鉴定成本）。常见的鉴定成本科目有：

外购材料的试验和检验费用。包括检验、试验单位评价外购材料质量的费用，以及有关管理人员及办公室人员可能用到的任何费用，还包括检验人员到供货厂评价所购材料时支出的差旅费。

计量服务费用。计量仪器的校准和维修等费用。

检验费。检验人员评价厂内产品技术性能时支出的费用，以及管理人员和办公室人员可能支出的有关费用。

试验费。试验人员用于评价厂内产品技术性能时支出的费用，以及管理人员和办公室人员可能支出的有关费用。

检查工作费。包括以下工作所需时间的费用支出——操作人员按照质量计划的要求而检验自己的工作质量；在制造过程中按要求自检产品和工序是否合格；挑选废品、次品；进行加工过程中的产品质量评价等。

试验、检验装置的调整费。有关人员为了进行性能试验而调整有关设备所需时间的费用支出。

试验、检验消耗及小型质量设备费用。试验、检验设备的动力消耗，如蒸汽、油以及在进行破坏性试验（如寿命试验或拆卸检查）时消耗的材料和物品；列为非固定资产的小型质量设备费用。

质量审核费用。产品、过程和体系审核费，包括内审和第三方评审。

外部担保费用。包括外部实验室的酬金、保险检查费等。

顾客满意度调查费。为了解顾客（包括内部）满意程度而进行相关调查分析的费用。

产品出厂复核费用。在发货装运之前再次审查试验和检验数据所支出的费用。

现场试验费。按照顾客指定的场所试验产品时所造成的损失，包括有关差旅费。

其他鉴定费用。

3) 内部故障（损失）成本。由于产品或服务在交付前不能满足质量要求所造成的损失，均可计入不良质量成本。常见的内部故障成本科目有：

报废损失费。因产成品、半成品、在制品达不到质量要求且无法修复或在经济上不值得修复造成报废所损失的费用，以及外购元器件、零部件和原材料在采购、运输、仓储、筛选等过程中因质量问题所损失的费用（不包括由于其他原因而废弃的材料）。

返工或返修损失费。为修复不合格品使之达到质量要求或预期使用要求所支付的费用（包括重新投入运行前的再次检验费用）。

降级损失费。因产品质量达不到规定的质量等级而降级或让步所损失的费用。

停工损失费。因质量问题造成停工所损失的费用。

产品质量事故处理费。因处理内部产品质量事故所支付的费用，如重复检验或重新筛选等支付的费用。

内审、外审等的纠正措施费。解决内审和外审过程中发现的管理和产品质量问题所支出的费用，包括防止问题再发生的相关费用。

其他内部故障费用。包括工作和过程质量损失、过程低效率损失、资源闲置等方面的费用。

4) 外部故障（损失）成本。由于产品或服务在交付后未满足质量要求所造成的损失，均可计入不良质量成本。常见的外部故障成本科目有：

顾客投诉处理费。在保单约定范围内，对顾客投诉的调查研究、修理或更换等所支出的费用，或在保单期满之后用于顾客特殊投诉的调查处理所支出的费用。

产品售后服务及保修费。直接用于校正误差或特殊试验、保修产品或零件以及用于纠正非投诉范围的故障和缺陷等所支出的一切费用（不包括安装服务费及合同规定的维修费用）。

产品责任费。因产品质量故障而造成的有关赔偿损失费用（含退货、法律诉讼、仲裁等）。

其他外部损失费。包括由失误引起的服务及付款延迟、坏账、库存损失，由顾客不满意而引起的成交机会丧失、顾客流失损失，以及顾客评审导致的纠正措施等费用。

六西格玛项目不良质量成本归集、计算与分析。

表 2-4

六西格玛项目质量成本计算表

项目名称: 提高非标系统产品物料清单 (BOM) 的准确性		
过 程: 系统产品设计、自制/外协/外购件确定、BOM 编制/修改、下订单		
类	内容描述	计算方法/公式
预防成本	BOM 编制培训费	共 10 位工程技术人员, 每人全年参加 8 小时的培训, 成本总计: $30.63 \text{ 元/小时} \times 10 \text{ 人} \times 8 \text{ 小时/人} = 2\,450 \text{ 元}$
	合计	2 450 元
鉴定成本	BOM 互审	每个 BOM, 机械、电控各互审 1 小时, 共计 2 小时, 一年 105 套, BOM 互审时间为: $105 \times 2 = 210 \text{ (小时)}$, 总成本为: $30.63 \text{ 元/小时} \times 210 \text{ 小时} = 6\,432 \text{ 元}$
	合计	6 432 元
内部故障成本	常用物料 (如标准件、电路接头等) 的年备货 3 000 元	考虑到公司年资金运转次数为 6, 盈利能力为 20%, 故年平均备货 3 000 元所带来的故障成本为: $3\,000 \text{ 元} \times 6 \times 20\% = 3\,600 \text{ 元}$
	BOM 审核后的修改	每个非标产品 BOM 审核后的修改时间约为 0.5 小时, 去年共计 105 套非标产品, 总成本为: $30.63 \text{ 元/小时} \times 105 \text{ 套} \times 0.5 \text{ 小时/套} = 1\,608 \text{ 元}$
	合计	5 208 元
外部故障成本	等工待料 392 小时	$30 \text{ 元/小时} \times 392 \text{ 小时} = 11\,760 \text{ 元}$
	错误物料采购	5 000 元
	推迟交货期所带来的直接损失	运费增加约 3 000 元
	推迟交货期所带来的间接损失 (信誉)	约 30 000 元
	合计	49 760 元

注: 上述统计的时间段为 2001/1/1—2001/12/31。

填表人: _____

项目主管: _____

日期: _____

日期: _____

表 2—5 六西格玛项目质量成本与不良质量成本分析报告

项目名称: 提高非标系统产品物料清单 (BOM) 的准确性					
过 程: 系统产品设计、自制/外协/外购件确定、BOM 编制/修改、下订单					
预防成本	内容	金额 (元)	不增值的符合性成本	内容	金额 (元)
	BOM 编制培训	2 450			
	小计	2 450			
鉴定成本	BOM 互审	6 432		BOM 互审	6 432
	小计	6 432		小计	6 432
内部故障成本	常用物料 (含标准件) 的年备货	3 600		常用物料 (含标准件) 的年备货	3 600
	BOM 审核后的修改	1 608		BOM 审核后的修改	1 608
	小计	5 208		小计	5 208
外部故障成本	等料待工	11 760	非符合性成本	等料待工	11 760
	错误物料采购	5 000		错误物料采购	5 000
	推迟交货期所带来的直接损失	约 3 000		推迟交货期所带来的直接损失	约 3 000
	推迟交货期所带来的间接损失 (信誉)	约 30 000		推迟交货期所带来的间接损失 (信誉)	约 30 000
	小计	49 760		小计	49 760
质量成本合计		63 850	不良质量成本合计		61 400

本章小结

六西格玛管理聚焦于过程改进, 强调以顾客为中心, 追求完美的经营结果。本章介绍了过程管理基础、顾客需求的分析和经营结果等三个部分与六西格玛管理相关的基础知识; 阐明了过程的定义, 以及过程负责人和相关方、价值链与过程类别、过程和职能与项目的关系; 说明了如何倾听、分析和转化顾客的声音, 以增进顾客满意和忠诚; 详细讨论了过程绩效的各种度量指标、水平对比法及其应用, 以及六西格玛财务收益的测算方法。

六西格玛项目管理

项目管理包括启动、规划、执行、监控和收尾五大过程。在六西格玛项目管理中，启动过程包括六西格玛项目选择和六西格玛项目立项表制定，规划过程即六西格玛项目计划制定，执行和监控过程包括六西格玛项目团队建设、六西格玛改进 DMAIC 或六西格玛设计的具体实施和监控，收尾过程则包括六西格玛项目总结、成果评审与分享。

3.1 六西格玛项目选择

六西格玛管理通过有组织、有计划地实施六西格玛项目而实现其经济效益，同时也通过六西格玛项目的实施来促进人们观念和行为方式的转变。六西格玛项目工作是黑带成长过程中的必经之路。因此，六西格玛项目的选择和实施在变革企业文化、促进企业的持续改善、培养具有实践经验的六西格玛黑带及获得实际的经济效益等方面，都有着十分重要的意义。

3.1.1 六西格玛项目选择原则

1. 六西格玛项目的定义

六西格玛项目指由职责明确的团队通过运用六西格玛方法（DMAIC 或 DFSS），在规定时间内，寻找最佳方案并实现预定目标的特定过程。

2. 六西格玛项目的选择原则

一般来说，六西格玛项目的选择遵从以下原则：

（1）有意义、有价值（meaningful）。六西格玛项目的意义和价值体现在以下几个方面：

- 六西格玛项目要支持顾客满意的改善。以顾客为中心是六西格玛价值观之一。六西格玛有两个基本点：一是产品或服务让顾客满意直至忠诚；二是减少缺陷直至彻底消除缺陷。因此，所解决的问题必须来自对顾客需求、顾客反馈等信息的分析，找出顾客真正关心的问题，在此基础上确定关键质量特性。

- 六西格玛项目要支持企业战略目标的实现。六西格玛管理是企业实现战略目标的有效手段，每一个六西格玛项目都应支持企业战略的实现。也就是说，通过六西格玛项目为企业架设实现未来战略目标的桥梁，促进企业战略目标实现。

- 六西格玛项目目标要有挑战性。一般来说，六西格玛项目都要将缺陷显著降低，只

有这样，才能实现过程绩效的突破性改进。

- 六西格玛项目要强调过程的改进。聚焦于过程改进也是六西格玛价值观之一，六西格玛是通过应用过程方法来实现过程绩效突破性改进的。

- 六西格玛项目要为企业带来较大的经济效益。六西格玛管理要用财务的语言阐述现状水平和改进后的绩效，用财务指标将绩效转化成财务效益。

(2) 可管理 (manageable)。每个六西格玛项目都应是可管理的，也就是说，六西格玛项目欲解决的问题应清晰且可测量。六西格玛改进不仅关注产品质量，而且关注缩短制造或服务周期、改善服务或交付的响应时间、提高生产能力、提高产品开发过程能力、提高效率、降低不良质量成本等。无论对哪种问题，都必须先定义“缺陷”，换句话说，就是先定义评判“好”“坏”的标准及其测量方法，然后才能评估欲解决问题的现状水平。

六西格玛项目的范围应清晰可控。每个项目都要有一个明确要解决的问题，每个问题的范围应界定清楚，并且适当，一般来说项目应在 4~6 个月内完成。但在许多情况下，一个问题的解决会涉及多个方面，比如：缩短产品的交付时间可能会涉及制造过程的质量，也会涉及原材料采购周期、产品运输周期、运输质量等一系列问题。这时，可分解成几个小项目，在相对小的可控范围内解决，便于管理。

六西格玛项目应得到管理层的支持和批准。项目必须得到管理层的支持和批准，才能获得适当的资源（人、财、物等）支持。

3.1.2 六西格玛项目选择流程

选择六西格玛项目的流程，也就是通过分析确定六西格玛项目的过程。一般来说，项目选择需要经过四个步骤，具体包括：

步骤一：确定项目的大方向——项目的 Y；

步骤二：将 Y 分解为若干 y，如： y_1, y_2, \dots, y_n ，并确定本项目针对哪个方面进行改进；

步骤三：针对选定的需要改进的 y，明确顾客关注的关键质量特性 (critical to quality, CTQ)；

步骤四：根据 CTQ 确定项目课题——具体项目名称。

六西格玛项目的选择流程如图 3—1 所示。

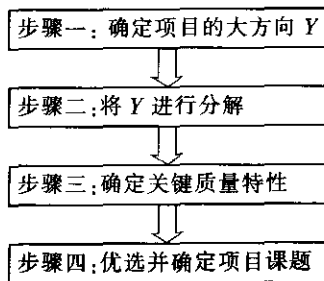


图 3—1 六西格玛项目选择流程



1. 确定项目的大方向 Y

确定项目的大方向，就如同在大海中找到指明航向的灯塔。只有有了正确的项目方向，我们才能在正确的方向上做正确的事——实施项目。那么，如何确定项目的大方向呢？一般来说，可以从以下几个方面进行考虑：

- (1) 根据卓越绩效评价中发现的改进机会，确定项目的方向；
- (2) 根据公司或部门的平衡计分卡中的指标弱项确定项目的改进方向；
- (3) 通过竞争对比和水平对比，找到存在的突出差距；
- (4) 根据对内外部顾客的声音（顾客高度关注的关键需求和期望）的分析，确定主要的改进方向。

从一些研究和企业的实践来看，从以下八个要点导入六西格玛项目，成功的把握将会增大：

- 战略实施的关键点；
- 目标展开的问题点；
- 顾客关注或投诉的热点；
- 统计数据的异常点；
- 部门间的矛盾点；
- 长期困扰企业的难点；
- 财务效益的增长点；
- 与竞争对手比较的薄弱点。

下面，着重说明一下平衡计分卡在六西格玛选项中的应用。

平衡计分卡中的四个维度中的每一个维度指标都可以分别作为项目选择的大方向，这就保证了六西格玛项目与组织的当期目标及未来的发展相一致。需要注意的是，在确定项目具体指标时需考虑四个维度的平衡，避免一种指标的改进可能导致另外一种指标产生不良。

图 3—2 是某企业的平衡计分卡指标。

财务	顾客
<ul style="list-style-type: none"> ● 库存水平 ● 单位成本 ● 生产成本 ● 隐藏工厂 ● 不良质量成本 (COPQ) 	<ul style="list-style-type: none"> ● 顾客满意和 CTQ ● 准时交付 ● 产品质量 (过程输出 KPOV) ● 沟通
内部业务过程	学习与成长
<ul style="list-style-type: none"> ● 缺陷 ● DPMO, SIGMA 水平 ● RTY (流通合格率) ● 过程周期 ● 关键过程输入 (KPIV) 	<ul style="list-style-type: none"> ● 六西格玛方法、工具应用 ● 专题培训 ● 六西格玛绿带、黑带培养 ● 组织活动技巧 ● 参加六西格玛项目员工人数

图 3—2 某企业的平衡计分卡指标

平衡计分卡中的四个维度中较弱的维度指标应作为项目选择的大方向，以保证企业的平衡发展。

项目的大 Y 一般表现为公司级的指标，如企业的顾客满意度低、生产效率低等，为了进一步明确问题，应对影响大 Y 的方面进行分解。

2. 将 Y 进行分解

在确定了项目的大方向 Y 后，基于 Y 是综合因素的反映以及涉及的方面较广等原因，一般难以针对 Y 直接进行改进。为此，需要分析影响 Y 的主要方面，通过逐层分解，确定要改进的主要方面 y。根据范围及影响度的不同，一般 y 需跨部门解决的，则可定义为黑带（BB）项目；在部门内或一定可控范围内可以解决的，往往可作为绿带（GB）项目；对于范围可控、难度较小的不需要高级统计工具来分析解决的项目（用老七种工具或新七种工具等即能分析解决的问题），可定义为一般项目，通过 QCC 或现场改进小组等方式实施。如造成顾客满意度低的一个主要原因是返修品的处理周期太长，因此返修品的处理时间就是一个 y，而对返修品的处理涉及市场部、质量部、技术部、制造部和车间，问题的解决需要一个跨职能的团队。

3. 明确对应于 y 的关键特性

在选定了项目 y 以后，有时还不能直接就确定 y 为改进项目。因为前面已经说过，六西格玛的价值观之一是以顾客为中心，因此需要针对 y 去听取顾客的心声，调查并了解顾客对 y 的需要，从而确定顾客最关心的关键质量特性 CTQ。对于返修品处理周期长的问题，顾客关注的 CTQ 包括顾客从把需要返修的产品交付给企业到接到问题处理结果为止的时间，以及问题解决率、对顾客损失的赔付等。

4. 优选和确定项目课题

通过上面三个步骤，已经明确了项目的改进点 y 和在改进点 y 上顾客最关心或最关注的是什么特性。在第四步，就特性 y 和顾客的 CTQ 优选并最终确定项目课题。在确定项目课题时通常要考虑以下问题：

（1）优选项目。如果有很多 CTQ 需要改进，而且改进资源有限，则应根据组织的战略发展需要和改进项目给组织带来的绩效，选择优先改进的机会。可采用帕累托优先级指数（Pareto priority index, PPI）进行优先级排序：对六西格玛项目带来的收益或节约的费用、成功的可能性、项目花费的成本和项目花费的时间等方面进行综合考虑，用计算项目 PPI 的方法来优选项目。项目 PPI 计算公式如下：

$$PPI = \frac{\text{项目节约的费用} \times \text{项目成功的概率}}{\text{项目需投入的成本} \times \text{项目完成时间（年）}}$$

例 3—1

根据项目带来的收益或节约的费用、成功的可能性、项目需投入的成本和项目花费的时间对 4 个项目的 PPI 分析如表 3—1。

表 3—1 六西格玛项目帕累托优先级指数（PPI）排序表

项目	收益（KRMB）	成功的概率	需投入的成本（KRMB）	完成时间（年）	项目 PPI
项目 1	175	0.70	50	0.33	7.4
项目 2	200	0.90	20	0.50	18.0
项目 3	185	0.75	50	0.50	5.6
项目 4	250	0.90	70	0.58	5.5

根据项目 PPI 的计算结果，项目 2 应作为优先选择的项目。



(2) 确定项目。确定项目的方式有两种：一是直接将顾客的 CTQ 确定为改善项目，例如：将缩短返修品处理时间作为改进项目；二是通过矩阵图，找出影响顾客 CTQ 的相关流程，确定一个或几个流程的改善为改进项目。

(3) 进行完整的问题陈述。言简意赅地定义问题并使之量化。

(4) 确定项目课题名称。

(5) 描述项目目标，做到 SMART：

Specific 具体的

Measurable 可测的

Attainable 可行的

Relevant 相关的（与企业战略一致的）

Time Bound 有时间限制的

SMART 原则是对六西格玛项目进行评估的重要方法。第一，项目目标必须具体，比如，降低库存量就不是一个具体的目标，没有指出是什么类型的库存，在制品、原材料还是产成品，不同类型的库存成因是不同的，不能一概而论。第二，项目目标必须可以测量，能够量化出来。有的项目目标可能比较容易量化，有的就比较困难，如顾客满意本身就是一个难以量化的指标，但是现在一般采用问卷调查和统计分析工具把它量化出来。项目目标不能量化，就难以考核目标是否达成。第三，项目提出的指标必须是可行的，是通过项目组的努力可以达到的，项目目标既要有挑战性，又要有可行性。一般在制定目标时，要考虑历史上的本流程的最佳水平以及与同行最佳水平比较的结果。第四，项目的目标必须与企业的战略相联系，与战略的大方向保持一致或者能够支持企业战略目标的实现。最后，项目目标的实现要给出时间节点，如果总目标下有分目标，也要给出分目标实现的具体时间节点。

3.1.3 六西格玛项目选择需注意的问题

在项目选择过程中经常会出现一些问题和错误，例如：

(1) 项目欲解决的问题与企业发展重点或关键顾客需求等没有联系，因此体现不出项目价值，无法得到管理层的支持和承诺，此类问题不适合选为六西格玛项目。

(2) 项目改进内容不是针对顾客的 CTQ，项目实施后看不到对 Y 的改善，不见成效，此类问题也不适合选为六西格玛项目。

(3) 没有针对 Y 进行分析、分解，直接将 Y 作为改进项目，造成因素太多，项目太大，抓不住重点，项目难以达成目标，应该将 Y 分解成若干个具体的六西格玛项目。

(4) 欲解决的问题原因已经明确，行动措施已经初步确定，此种问题应该立即采取行动，而不需要再选作六西格玛项目，避免比葫芦画瓢走过场，搞形式主义。

(5) 项目衡量指标不明确或项目目标没有挑战性，此类问题浪费资源、无意义，不适合选为六西格玛项目。

(6) 项目难度太大，超出项目团队的能力或授权，不适合选为六西格玛项目。

(7) 项目改进空间太小，预期收益太低，企业得不到适当的回报，此类问题也不适合选为六西格玛项目。

总的来说，好的开始是成功的一半，选择好的项目是企业六西格玛管理成功的关键。初次导入六西格玛管理的企业在考虑选择什么项目时，可能不是那么顺利，会占用一些时

间，甚至会有一些反复的过程，然而，这却是六西格玛管理活动中必不可少的一项重要工作，多花一些时间也很必要。

3.2 六西格玛项目立项表和计划

3.2.1 项目立项表和计划概述

在按照六西格玛项目选择的原则选出了项目后，就要制定项目立项表和计划，委派项目组长和组成项目团队，明确项目的目标和各种环境因素以及总体进度等，指明团队努力的方向，为项目的成功实施奠定基础。

1. 项目立项表概述

项目立项表是正式批准项目的文件，该文件授权六西格玛项目组长（黑带或绿带）在项目活动中动用组织的资源。项目立项表包括以下要素：

- 项目名称；
- 项目背景及选择理由；
- 问题/机会陈述；
- 目标陈述；
- 项目团队组成及职责分工；
- 项目涉及的过程和职能范围、约束和假定；
- 项目利益相关方及其影响；
- 总体里程碑进度表（阶段性任务及时间安排）；
- 倡导者的批准和授权。

项目立项表的制定是一个商议的过程，一般由项目组长组织项目成员一起讨论、分析，制定出初稿，报企业负责六西格玛协调、推进的资深黑带或承担类似职责的主管人员，或培训、咨询师审议，最后报送发起该项目的倡导者批准。

项目立项表的完善是一个动态的过程：在六西格玛项目的界定阶段，应对项目立项表要素进行进一步的明确和更新；当测量阶段结束后，如果项目过程绩效基线数据与界定阶段不同，也应对问题或机会以及目标陈述予以必要的更新和完善；当项目总结时，基于项目立项表评价项目目标的完成情况等。

2. 项目计划概述

实施项目还必须有计划，这个计划为项目提供了一个指导框架，以此协调团队的活动和资源的使用。六西格玛团队应基于总体里程碑进度表，制定细化的项目计划。六西格玛项目计划一般以进度计划为主，必要时，可制定成本费用、人力资源、沟通、风险管理等专项计划。

无论怎么做计划，有一点是可以确信的：项目计划与实际情况不会完全一致。但项目计划仍是必需的框架性文件，能为项目组织和协调提供有效的指导；同时，随着项目的进展必须对项目计划不断进行更新。也就是说，在制定项目计划时必须采取以目标为导向的策略，项目计划必须与目标紧密结合。

切实可行的计划是实现项目目标的重要保证。在制定项目计划的过程中最重要的是团队成员的参与并达成一致意见。目前，在企业里以团队的形式来开展工作、解决问题的实践越来越普遍。每个人都负责一个甚至多个项目，项目计划制定的工作往往落在团队负责人的肩上。这种项目计划制定的方式一般存在两方面的问题：一是计划考虑得不够全面；二是执行过程中团队成员的不认同与执行不到位。因此，最终的结果是计划不能得到有效的实施，即便实施了也无法达到预期的项目目标。因此，团队成员参与项目计划的制定就变得尤为重要。

在制定项目计划时，团队成员可以参照以下步骤完成：

(1) 任务分解 (work breakdown structure, WBS)。就是将为达成项目目标而需要完成的 DMAIC 或六西格玛设计的阶段性工作分解为可执行、可跟踪的工作单元 (任务、活动或关键阶段)，为进一步明确、区分各工作单元，还可以制定工作任务分解表 (见表 3—2)。

表 3—2 工作任务分解表

WBS 编号	任务名称	工作内容	输入	输出	标准	责任人

(2) 估算任务时间并确定任务之间的关系。明确了各个工作单元之后，就要估计每个单元所需的时间及彼此之间的先后关系，明确相关责任人。时间估算的方法很多，这里只介绍其中的一种：三点法。

$$E = (O + 4M + P) / 6$$

式中， E 为时间估算结果； O 为乐观估计时间； M 为正常估计时间； P 为悲观估计时间。

各工作单元之间关系的确定主要是考虑各任务之间的衔接，如哪项工作应该在哪项工作完成后才能开始等。

(3) 编制项目工作计划。为了使六西格玛项目有效实施，团队在界定项目时，就应该注意项目计划的编制和制定。依照六西格玛改进的五个阶段，可以相应地确定计划项目完成时间，制定“项目计划工作表” (见表 3—3)。然后按照各阶段任务分解 (WBS) 得出的工作任务，采用甘特图 (Gantt chart) 等方法把计划加以细化，便于团队活动。制定计划时，应让团队成员共同参与。由团队共同参与并听取成员的反馈意见将有助于增进团队的向心力。

表 3—3 项目计划工作表

项目名称	项目组长		
活动里程碑	责任团队成员	目标完成日期	实际完成日期
D 界定阶段			
M 测量阶段			
A 分析阶段			
I 改进阶段			
C 控制阶段			

3.2.2 项目规划工具

有很多工具和技术都可以协助项目团队制定项目时间表，按时间分配资源，在项目计划的实施过程中跟踪项目的进展。下面介绍最常用的两种：甘特图和网络计划技术（即通常所称的网络图、箭条图或矢线图）。

1. 甘特图

甘特图显示了随着时间的推移，项目任务之间的关系。甘特图的横轴是时间单元（天、周、月等），纵轴是需要完成的活动。条形显示了各个活动的估计开始时间和持续时间。表 3—4 是某企业六西格玛项目界定阶段的甘特图。

表 3—4 六西格玛项目计划甘特图

阶段	周（5月7—31日）			
任 务	第1周 5月7—14日	第2周 5月14—21日	第3周 5月21—28日	第4周 5月28—31日
评价项目立项表和确定项目				
确定顾客需求的关键质量特性（CTQ）				
确认关键过程输出变量（KPOV）				
界定不良质量成本（COPQ）				
估算 σ 质量水平，评价现有数据				
完成界定阶段，着手测量计划				

甘特图可以进行一些功能上的增强，如：增加里程碑事件的标志、项目进展图示、任务负责人等。

甘特图是在 20 世纪早期发展起来的经典的计划制定工具，具有形象、直观、简明、易懂、作图简单等优点，至今仍是行之有效的一种制定计划的方法。用甘特图制定项目计划时，必须清楚任务之间的相互关系，即哪些任务在其他任务开始之前必须完成，哪些任务可以同时进行。传统甘特图的主要缺点之一是不能在图上清晰和严密地显示各项任务之间相互关联、互为条件、互为因果的依存关系，以及在时间上的先行和后续的衔接关系。因此，如果一项任务被延误，其他哪些任务将会受到影响就不能明显地表示出来。不过，大多数项目管理软件包（如 Microsoft Project）能够提供通过应用连接箭头表示任务间相互依存关系的甘特图。

2. 网络计划技术

甘特图除了有上面谈及的缺点之外，还存在不能找出关键的活动和路径、不易对周期时间的缩短和资源的利用进行优化等缺点，包括：（1）难以给出极详细的计划；（2）在计划阶段不便于反复推敲与思考；（3）进入实施阶段后的情况变化与计划变更难以处理；（4）不能获得有关某项作业迟滞对整个计划影响的正确信息；（5）设计规模稍大即难以掌握计划全貌；（6）难以判断进度上的重点。

对于较大、较复杂的项目，可能需要使用更科学的工具，协助项目的计划、日程安排



和进行大量相关活动的协调工作，因此要求有一种新的、更好的编制计划的方法和计划的表达方式。

20 世纪 50 年代以来，很多人都在探索如何制定一种新的计划方法。1956 年，美国杜邦公司（Dupont Co.）在制定协调企业不同业务部门的系统规划时，运用网络方法制定出第一套网络计划。这种计划借助于网络表示各项工作及其所需要的时间，并表示出了各项工作之间的相互关系，从而找出编制与执行计划的关键路线，所以把这种方法称为“关键路线法”（critical path method），简称 CPM。1957 年，杜邦公司将 CPM 应用于一个大型化工厂的建设计划与控制，第一年就节约 100 万美元，相当于该公司用于研究开发 CPM 方法费用的 5 倍。1958 年，美国海军武器局在制定北极星导弹计划时，同样应用了网络方法和网络形式，但它注重对各项任务安排的评价和审查，所以把这种方法称为“计划评审技术”（program evaluation & review technique），简称 PERT。由于应用了 PERT，工程工期由原来计划的 10 年缩短为 8 年。实践再一次证明了网络计划方法是一种行之有效的科学管理方法，得到了管理界的广泛承认和采用。这两种方法通常被统称为网络计划技术，即通常所称的网络图、箭条图或矢线图，在我国也称为统筹法。

使用网络计划技术进行项目安排包括四个基本的阶段：规划、日程安排、改进和监控。

- 规划阶段需要应用 WBS 树图将项目分解为具体的可执行、可工作的任务或活动，然后估计这些活动的开始时间和完成这些活动所需的时间，把每个活动都用一个箭头来表示，构造项目的网络图。

- 项目安排阶段需要构造出标有每个活动的起止时间以及与其他活动之间关系的时间图。同时必须识别出关键活动，也就是那些必须按时完成的活动，否则将会延误整个项目的进度。关键活动构成了关键路径。

- 改进阶段需要着手对关键路径、关键活动进行改进。项目团队不应当仅仅只是接受项目的日程安排，同样也可以改进项目的日程安排。可以应用帕累托分析来识别影响关键活动周期时间的关键因素，通过改进以显著地缩短整个项目的完成时间。成本数据可用于对时间数据的补充，并可用帕累托分析对成本数据进行分析。

- 监控阶段需要使用网络图来进行周期性的评估和控制。包括：密切注意关键路径上的每个活动；根据项目的实际进展，对网络图进行动态的修订，特别关注关键路径的可能变迁；针对突发事件准备应急计划。

网络计划技术可用于：

- 辅助项目规划和监控；
- 确定在预定日期前完工的可行性；
- 识别项目实施中最可能出现瓶颈的环节；
- 估计项目需求或日程安排的变化对项目的影响；
- 估计偏离日程安排对项目的影响；
- 估计调走部分资源或增加资源对项目的影响。

网络图的构造和应用参见本章 3.5 节。

3.2.3 项目文档

作为六西格玛项目管理的一部分，规范一致的项目文档是必需的。这些规范的文档对于持续改进活动及知识的积累都会起到很好的作用。在项目实施过程中，应有专人负责对这些文档进行动态的归集、管理和维护，进行阶段性的审查，并在项目完成后交由六西格玛管理机构或档案管理部门存档。随着信息化时代的到来，使用电子文档来保存项目全生命周期的信息越来越普遍。

不同的企业会制定不同的六西格玛项目文档的模板及记录要求，但通常会包含以下内容：

- (1) 项目立项表及其更新；
- (2) WBS、项目计划（甘特图、网络图等）及其更新；
- (3) 不良质量成本测算报告；
- (4) 项目收益预测与资源需求、预算；
- (5) 团队规则；
- (6) 会议纪要、电子邮件及学习笔记等；
- (7) 项目状态分析资料、报告及演讲材料；
- (8) 项目效果测评（包括成本—收益分析）和确认；
- (9) 行之有效的措施的标准化（程序、图纸等的更改）；
- (10) 项目总结报告（包括经验教训）、成果发布和分享材料。

一般来说，推行六西格玛管理的企业都要建立六西格玛项目的审批流程及项目管理制度，其中应规定对项目文档管理的具体要求。

3.3 六西格玛项目团队建设

3.3.1 团队的组建和授权

1. 团队组成要素

拥有高度热忱的团队成員与拥有受过专业训练的领导人同样重要。团队应做好准备工作，建立对团队目标的共识，然后决定如何实现这一目标。表 3—5 列出了团队成员必须确定的内容。

表 3—5 六西格玛团队组成要素

要素	要求
使命	团队成立或存在的目的
基础	团队的使命如何与企业目标或计划配套
目标	对现状及绩效的挑战
角色	团队成员（黑带、绿带）
职责	根据项目分配给每位成员的职责和任务
主要里程碑	项目活动时间表、项目报告日期
授权	获得管理层的授权和支持

2. 选择团队成员

六西格玛团队通常由以下人员组成：

(1) 团队领导/组长。由黑带或绿带担任，对整个项目负责，具体职责参见第1章。

为了使团队领导/组长履行好其职责，保证项目的顺利开展，在项目启动前，需要做好相应准备工作，表3—6和表3—7列出了可供团队领导/组长参考的准备的注意事项和活动建议。

表3—6 团队领导/组长核查表：六西格玛项目团队首次会议前准备工作

<input type="checkbox"/> 与业务负责人及发起人一起起草项目立项表 <input type="checkbox"/> 如需要，挑选小组成员 <input type="checkbox"/> 联系并欢迎成员加入 <input type="checkbox"/> 起草首次会议日程 <input type="checkbox"/> 安排会议后勤事项 <input type="checkbox"/> 选择一个启动会议时调节气氛的小技巧 <input type="checkbox"/> 与过程负责人建立联系，并让其了解项目的整体方向 <input type="checkbox"/> 作相关方分析 <input type="checkbox"/> 开始罗列小组所需外部支持人员名单等
--

表3—7 项目小组首次会议

与会人员：[名单]		
日期：	时间：	地点：
会议目的：[六西格玛项目名称] 启动会		
日程安排： <ul style="list-style-type: none"> ● 致欢迎辞 ● 小组成员介绍 ● 讨论小组成员的目标、期望及角色，团队活动规则、章程等 ● 审核项目立项表 ● 列出主要相关方名单 ● 根据小组成员提出的疑问、项目立项表及下次会议所需准备，列出行动事项 ● 确定下次会议时间和议程 ● 评估会议效果 ● 结束会议 		

(2) 核心成员。核心成员是团队领导/组长选择来实施项目计划的人员，一般由绿带担任，他们对团队领导/组长负责，全程参与项目，具体职责参见第1章。

(3) 扩展成员。扩展成员即根据项目的需要，只需部分参与项目的成员，如团队中财务核算人员、供应商代表、个别的技术专家等。

(4) 业务负责人。业务负责人是项目所在部门的负责人和过程的管理者，具体职责参见第1章。

(5) 倡导者。他们有时也直接参与团队活动，具体职责参见第1章。

(6) 项目指导人或教练。他们是为团队成员提供六西格玛方法、工具指导的人员，一般由资深黑带或咨询顾问担任。其职责包括：

- 1) 根据项目各阶段的需要提供六西格玛方法、工具培训；
- 2) 指导项目按六西格玛方法、思路推进；
- 3) 为团队成员正确合理地应用六西格玛方法、工具提供指导。

3.3.2 团队发展阶段

随着团队成员共同工作，团队逐步进入成长和成熟阶段。理解这些发展阶段对有效的管理团队过程是很有价值的。这些阶段可能表现出不同的强度和持续期，这取决于团队的类型、团队工作环境以及构成团队的个体。达克曼描述了团队发展阶段的一般模型（见图3—3）。

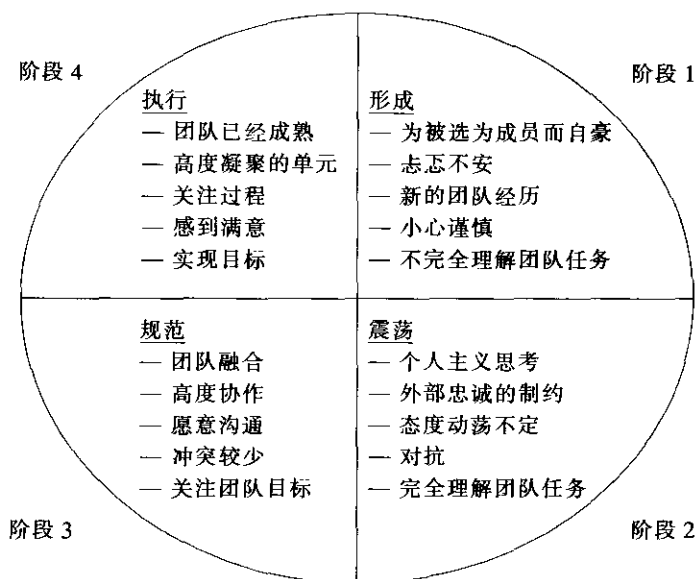


图 3—3 项目团队发展的四个阶段

阶段 1：形成期。团队首次聚在一起。每个团队都是一个新的经历，甚至对那些以前曾参加过团队的成员来说也是如此。每个人都小心翼翼地进入这个环境，他们在这个新的经历中对角色和绩效感到忐忑不安。因此，在形成阶段，团队通常要阐明其目的，确定每个成员的角色，进行团队培训，以及制定可接受的行为准则。

阶段 2：震荡期。在此期间，团队成员对团队任务的真实情况已完全理解，但团队成员仍首先作为个体来思考，并往往根据自己的经历做出决定，而不是与其他成员协商广泛听取建议。随着团队工作的进展，他们对于实现项目目标的最佳方法存在分歧，甚至对项目的总体目标也可能存在分歧。大家还会发现每个人在项目中的工作方法不同，这些分歧可能引起团队内的争论甚至矛盾。对于团队来讲，如果不能处理好这一阶段的问题，很可能就会导致项目最终失败。对此，团队领导/组长应该事先估计到这是团队组建过程中必然出现的正常阶段。他可以通过与团队成员的个别沟通，重申团队的目标、鼓励成员发表解决问题的意见等方式，排除障碍，带领团队跨过这一阶段。必要时可以请业务负责人或



倡导者出面支持。

阶段3：规范期。在此阶段，团队成员已从关心个人问题转变为关注实现团队目标，个人已融入团队中。团队成员间会达到一定程度的彼此适应。团队成员开始就各种问题达成一致，并愿意为了团队而协商和消除分歧，因而在这个阶段团队有了更多的协商和对话。团队领导/组长应予以支持和赏识。

阶段4：执行期。在此阶段，团队已经成为一个具有高度凝聚力的整体。团队成员对他人的优缺点以及他们如何完成使命了如指掌，并能够群体解决问题。在此期间，团队在实现目标上取得重大进展，团队领导/组长可予以更多的授权。

虽然发展的四个阶段表明了时间的逻辑结果，但每个团队的实际发展过程千差万别。例如，前进到阶段3或阶段4的团队可能退回到阶段1或阶段2。如果他们认识到之前的假设并不成立，他们已用于决策的信息并不确切，或是由于工作的调动而改变团队的成员关系，这种情况就可能发生。一些团队可能在原阶段停滞不前，但项目的时间有限，必须在动态中及时调整、解决团队的问题。

3.3.3 团队动力与绩效

1. 建立团队的技巧

团队领导/组长负责组建团队，并对项目目标负责。一个跨部门的团队通常由4~7名核心成员及其他必要的扩展成员组成。以下是选择核心成员时除了专业知识及背景外还要考虑的因素：

- 具有团队精神；
- 团队工作的经验；
- 良好的沟通能力；
- 愿意接受挑战；
- 勇于揭露潜在的问题。

一个成功的团队应该是一个高度团结、独立自主的团队。这些不仅仅是在选择团队成员时要注意的，也是一个团队在开展工作的过程中应该去努力营造的。当然，其中团队领导/组长要负主要的职责。通常团队建设的步骤如下：

- (1) 仔细分析任务，确定所需的技能组合和工作风格，以形成互相补充、互相促进的组合；
- (2) 向每一位团队成员讲述团队的愿景、目标以及对个人的任务要求；
- (3) 提供必要的技能培训以及必要的处理工具（软件、硬件等）；
- (4) 共同建立团队规则（如图3—4），包括团队的例会制度、问题的沟通解决机制、编制项目立项表等；
- (5) 监控进度、保证团队凝聚力，并不断向目标迈进；
- (6) 注意团队工作中的成绩，并给予祝贺。

六西格玛团队规则

我们一致同意从我们团队自己做起，最有效地利用时间。我们承诺以下行为准则：

1. 会议准时开始和结束。
2. 如果有人不能参加会议，要至少提前 1 天通知组长。
3. 每次会议的内容都要提前通知：做一个议程，或在会前作一讨论。
4. 准时完成所分派的工作。如果做不到，要至少提前 1 天通知组长。
5. 对任务坚持不懈。
6. 会议仅通知那些需要参加的人。
7. 会议以开放式论坛的方式进行。我们彼此挑战，坦诚相见。
8. 可能时，我们将利用午饭时间开会。

图 3—4 团队规则举例

2. 指导团队的技巧

面对六西格玛管理的挑战，六西格玛团队必须讲求团队技巧。人们希望努力工作去争取成就，但个人的成就总是有限的，团队的绩效要优于个人成就。前面我们已经提到黑带不仅必须具备使用统计方法的能力，同时必须拥有卓越的领导力与亲和力，使六西格玛团队成员相互依存、相互帮助，项目才能取得真正的成功。那么黑带如何领导团队取得成功呢？这需要一些方法和技巧。

(1) 团队激励。作为团队的负责人，欲使其成员共同为项目负责，激励技巧是十分重要的，这里有一种务实的团队激励办法可供参考，即 TARGET，具体如下：

- T (truth, 事实)。团队成员必须知道事实真相，黑带应与团队成员坦诚沟通，分享经验，这有助于建立“勇于面对事实真相”的工作环境。

- A (accountable, 负责)。团队成员应该为自己的绩效负责。黑带可以通过诸如审查项目计划及定期检查项目实施情况并予以指导等，增强团队成员的责任感。

- R (respect, 尊重)。团队成员应该本着正直与坦诚的原则，相互学习，交换心得。直接提出问题，然后制定一个了解问题、探索问题的解决方案。

- G (growth, 成长)。团队成员及团队本身都必须能够经由学习而成长，以便有能力挑战更困难的工作任务。黑带可以协助团队成员认识通过六西格玛活动而成长的重要性，让团队成员愿意经由学习而成长。

- E (empowered, 授权)。成功的群体仍需个别的行动，尤其是专业技术背景方面。黑带应协调业务负责人或倡导者，让团队成员取得授权，落实个别的具体行动，发挥个人的特长。

- T (trust, 信任)。作为一个有效的群体，六西格玛团队成员之间应该相互信任，并且同心协力完成指定的任务。黑带可以协助团队成员增进彼此间的相互信任。

有了上述保证，六西格玛团队就可以表现出下列高效六西格玛团队的共同特征：

- 跨职能团队。团队成员应来自不同的职能领域，同时拥有互助互补的技术能力。跨职能团队的多数成员仍需兼顾自己部门的工作。

- 挑战性绩效。更有挑战意义的绩效改进，将激励团队士气。在根据内外部顾客需求制定的挑战性目标指导下，团队成员之间充分沟通，团队成功机会将大幅度增加。

- 有意义的目标。团队的每一位成员都必须认同团队存在的价值及其对组织整体的贡



献，成功的团队必须共享目标，并得到所有成员的认可。

- 清楚的方法。团队必须使用共同语言，每位成员都必须了解并接受每一个问题解决办法。每位成员都必须彻底认识其所负责的过程作业。

- 开放的沟通。团队成员间的开放、诚实和经常沟通是一个基本要求。六西格玛的价值观之一是无边界合作，这是建立在广泛沟通的基础之上的。

- 公正的回报。为了发挥团队的激励作用，对工作卓越的团队或成员给予表彰和激励，使辛勤付出能够有相应公正的回报。

(2) 团队培训。六西格玛团队组建之后，必须结合项目的情况实施专题培训，教导“为何”(why)，指导“如何”(how)。培训是六西格玛项目推行的核心，重点对象是团队成员，特别是绿带。团队培训任务一般由资深黑带和黑带担任。

3. 团队绩效评估

在项目团队的发展过程中，应当不断对项目团队的绩效进行评估，从中及时发现团队发展中的问题。同时，团队绩效评估也是向管理层反馈团队在实现目标过程中的进展情况的一种有效手段，它对项目团队获得管理层的支持十分重要。

在项目实施的过程中，除了跟踪项目目标的实现情况，并由此评价项目团队的绩效外，也可以采用内部导向的测量指标，如团队会议的出勤率，团队成员的满意程度，以及其他表征团队动力的因素等来评价。项目团队应当进行自我评价，并向倡导者和业务负责人报告这些评价指标的测评结果。

3.3.4 团队工具

1. 头脑风暴法

头脑风暴法是团队工作中经常运用的方法。头脑风暴法一般可以分为三个阶段。第一阶段关注的是小组创造出的点子的数量。第二阶段需要审视这些点子，删除与实现目标无关的点子。第三阶段包括对筛选下来的点子做进一步的审视，并按照小组的意见，对它们进行优先排序。

头脑风暴会议过程如下：

(1) 头脑风暴的主题为小组接受，并用清晰的语句，以小组成员都能看到的方式写出来。

(2) 组织者向小组成员征求想法。

(3) 成员写下每个想法，不讨论、分析或批评。

(4) 这个过程延续到没有新的想法出现。

再次审视想法表，保证每个人都理解它，然后将类似的想法合并，形成可用于投票的想法表。

头脑风暴也可以做得更加结构化，如轮回法：要求小组的每个成员每次提出一个想法，如果没有想法则简单地跳过，接着询问下一个人的想法，然后是再下一个，依此类推，当最后一圈每个人都没有新的想法时，头脑风暴会议结束。也可以使用类似的过程，用几张纸片贴在房间的四周，在每张纸片的顶端写上特定的主题或问题，小组成员从一张纸走到另一张纸，把他们的想法写下来。在桌旁围坐的小组同样可以用传递纸片的方法来完成类似的过程。

头脑风暴还有一种特别的方法可用于以下场合：所处理的是特别敏感的主题，或小组成员间还没有建立高度的信任。这个方法叫做 Crawford 纸片法，请每个人制作他（或她）自己想法列表的纸片，接着，递给一个可信赖的人（例如，团队负责人），这个人将其编辑成为单一的列表。这种方法的匿名性可以使人们自由提出他们的想法，在小组中经常会发现有几个人有着相同的想法，从此开始建造共同的平台。

2. 名义组技术

处理头脑风暴所产生想法的一般方法是名义组技术（nominal group technique）。应用这个技术可以帮助团队从他们提出的想法中选择出少数重要的想法。名义组技术由下列步骤构成：

- （1）对想法列表进行整理，并简化/合并，使想法列表完整、清晰，且不重复。
- （2）请每个参加者为各个想法进行等级评价（例如，1 表示最好，8 表示最坏）。
- （3）在每项旁边记录全部参加者评出的等级。
- （4）对每项等级评分求和。那些总分最低的项目即团队认为最佳选择。

表 3—8 是名义组技术应用的一个简单案例。案例中某小组成员要决定他们的改进方案。通过头脑风暴他们提出了 4 个供选择的方案：方案一、方案二、方案三和方案四。小组组长请每个人为这 4 个选项打分，将他或她的选择按最佳选择至最末选择从 1~4 评分，将每个人的评分填入表 3—8 中。

表 3—8 名义组技术应用举例

方案	个人与排序					合计
	组员 A	组员 B	组员 C	组员 D	组员 E	
方案一	4	4	4	4	3	19
方案二	3	3	3	3	4	16
方案三	1	2	1	2	1	7
方案四	2	1	2	1	2	8

在四个选择中，方案三总分最低（优先级最高）成为首选。

3. 多重投票法

多重投票法是使团队成员将他们的想法统一起来的另一种方法。首先，将团队成员的所有想法列表。然后，要求他们投票选出他们认为最适宜的想法，通常允许每人投的票数是总数的一半左右。在全部参加者都做出投票选择之后，项目负责人将逐项进行统计。按照帕累托多数原则，保留得票较多的，将没有得票或得票很少的挑出并从列表中剔除。投票过程重复进行，直到达到预期的项目数量。

下面仍以某小组成员要决定他们的改进方案为例，第一次每人可以有 3 个选择；第二次每人可以有 2 个选择；第三次每人可以有 1 个选择（见表 3—9）。

表 3—9 多重投票法举例

方案	投 票		
	第一次投票	第二次投票	第三次投票
方案一	1		
方案二	3		



续前表

方案	投 票		
	第一次投票	第二次投票	第三次投票
方案三	3		
方案四	4	3	
方案五	5	4	
方案六	6	6	2
方案七	8	7	8

通过三轮投票，最终大家确定选择方案七。

4. 力场分析

力场分析是库尔特·莱文（Kurt Lewin）提出的一种方法，能够用于察看各种关系及影响其变化的因素。采用这个概念，一个组织中各种事情的当前状态被看成是一种平衡状态，维持这种平衡状态的是相反方向上作用的两组力：驱动力（推动、促进）推动变化，而阻止力（阻碍、妨碍）阻碍变化。一项变革不可能在阻力强过动力时发生。增加动力无疑会促进变革的发生，但减少阻力可能会取得更好的效果。

我们以企业整体推进六西格玛为例介绍力场分析的应用（见图 3—5）。

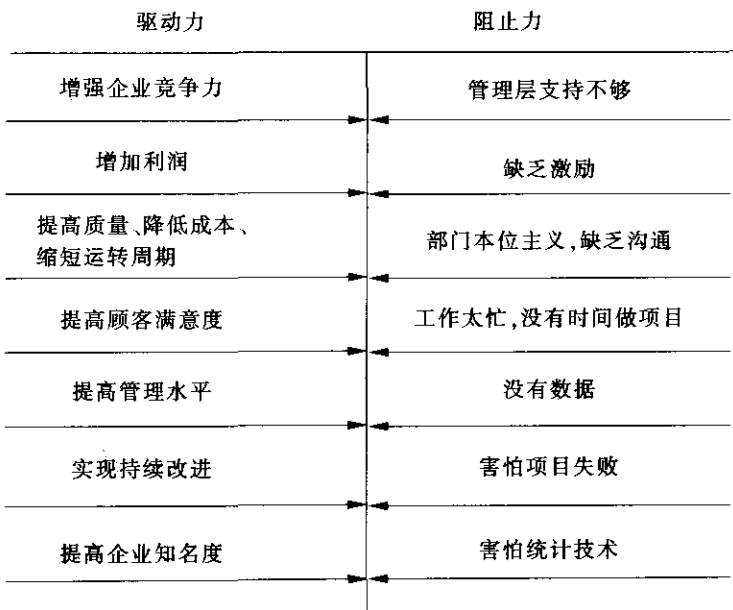


图 3—5 企业实施六西格玛力场分析图

第一步：把影响六西格玛实施的所有因素尽可能全面地罗列出来，如管理层的认识和支持程度、员工的认识程度、企业的战略、管理现状等。

第二步：对已列出的因素一一进行分析，判断其对推行六西格玛管理的过程是有利的还是不利的。比如，六西格玛对企业竞争力和管理水平的提升有帮助，这些有利于六西格玛的开展。而管理层支持力度不够、员工害怕项目失败等属于不利因素。

第三步：对有利因素和不利因素的“力量”大小进行估计，了解动力与阻力的力量对比。通常，力量的大小可在0~10之间取值。将有利因素的力量累加即为动力的大小，把不利因素的力量累加即为阻力的大小。当动力明显弱于阻力时，六西格玛不妨暂缓推行，待时机成熟再行实施；当动力大于阻力时，六西格玛可以比较顺利地推行。

力场分析可以帮助团队选择实施变革或新的措施的时机，它有助于管理者明确阻力的来源，做好有针对性的工作。如上例中，如果是由于管理层支持不够，企业就应当首先加强对管理层进行六西格玛管理的培训，组织高层管理者参观实施六西格玛管理的先进企业等。

3.4 六西格玛项目监控与促进变革

3.4.1 项目跟踪和监控

1. 项目跟踪和监控原则

在项目开始阶段，团队往往会投入较多的时间精力来界定项目，编制项目立项表和计划。一旦项目立项表和计划被批准了，就需要采取一系列措施，应用甘特图和网络计划技术等，跟踪、监督和控制项目的实施状况，包括监测项目进程，预测、发现并解决问题以及更新计划，以使项目恢复正常运作。

项目监控的目的是在预算内按时完成任务。在项目进行过程中，肯定会出现与计划不一致的事情，对此，团队需要快速地识别问题，解决问题，尽快使项目恢复正常。如何对项目进行有效的监控呢？图3—6是项目监控的常用流程。

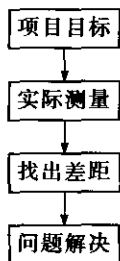


图 3—6 项目监控流程

项目目标是在制定项目计划时产生的。在制定项目计划时，为了便于以后的跟踪控制，项目计划中应该有里程碑（milestone）计划，对每一项工作应该有明确的输出要求。实际测量是指团队通过一些手段来明确项目目前的进展状况。这些手段包括向业务负责人或倡导者报告、团队会议、甘特图检查等。了解了目前项目进展状况，就可以与项目计划中的标准进行对比，找出差距，有效解决存在的问题。

2. 项目监控工具和方法

（1）项目柔性分析。项目监控的目的是使项目在预算内按时完成。但事实上，可以在规定的时间和预算内完成预定任务的项目并不多。这样，项目控制就涉及项目的调整问



题。是调整项目的时间、预算还是工作范围？为了能够使项目得到有效控制，建议在制定项目计划时，对项目柔性进行分析，根据分析结果，有针对性地调整项目（见表 3—10）。

表 3—10 某项目的柔性分析结果

不具有柔性	中等柔性	柔性最大
项目实施时间 项目目标	项目工作范围	项目所需资源 团队构成 培训

在以上分析结果中，柔性最大的是资源，最小的是项目实施时间。也就是说，当项目遇到问题需要调整时，团队首先考虑的是增加资源，确保项目在预定时间内完成。

（2）风险管理计划。项目不可能完全按照预期的那样顺利进行，一些意外情况的发生会给项目带来很大的影响，甚至会导致项目的失败。这就需要提前识别风险并有效管理风险。因为团队无法预测未来，所以最好的做法就是在制定项目计划时，以已知信息为基础制定风险管理计划。风险管理计划的制定可以参照风险识别、风险评估、风险管理三个步骤完成。

- 风险识别。一般地，项目风险来源于两个方面，一是项目本身的，二是项目外部的，如企业战略的调整、行业政策的变化等，团队可以运用头脑风暴法从以上两个方面去识别项目风险。

- 风险评估。团队对于已经识别的风险进行评估，评估每项风险对项目威胁的大小，针对可能对项目造成较大威胁的风险制定预防及应急计划，进行风险管理。一般对风险的评估从三个方面进行：风险发生后后果的严重性、风险发生的概率、及时发现风险的难易程度（见表 3—11）。

表 3—11 对某风险的评估

风险	后果的严重性	发生的概率	及时发现的难易程度	总体威胁
1. 订单量激增	M	H	L	L
2. 订单量严重不足	H	M	L	M
3. 熟练员工大量流失	H	L	H	M
4. 工时统计不及时	H	M	L	H
5. 新产品推出慢	H	H	L	M
6. 订单模式的变化大	M	H	M	M
7. 工艺的重大变化	M	M	H	M

说明：H 代表高；M 代表中等；L 代表低。

总体威胁亦即风险管理的优先顺序。对于总体威胁中等及总体威胁大的，要制定风险预防计划。

- 风险管理。主要从三个方面考虑来制定风险预防计划，包括预防措施、应急措施、风险发生的临界定义，最好指定一名风险负责人（见表 3—12）。

表 3—12 对某风险的预防

风险	预防措施	应急措施	临界定义	责任人
2. 订单量严重不足	培养一岗多能的弹性用工	优化生产安排	任务量低于 2 000 台	王小明
4. 工时统计不及时	专人负责统计	由生产处经理主管监督	连续两次发现有损失工时不报的情况	周大海
5. 新产品推出慢	试生产阶段做好工程准备	调整项目计划	新品试生产上线一周后,效率达不到正常水平	赵明
6. 订单模式的变化大	研究批次、每批任务量与效率之间的关系	调整项目计划	小批次特配订单任务量超过 50%	孙峰
7. 工艺的重大变化	加强对质量问题的监控	调整工艺及人员	批次性质量问题,加严测试	王小明

(3) 建立阶段或里程碑汇报制度。建立 DMAIC 五个阶段的项目进展汇报机制, 根据项目计划工作表, 通过各阶段目标达成情况的回顾和评估, 使问题在过程中被发现, 及时调整资源和计划, 保证项目目标的达成, 这是所有企业组织的惯常做法, 也是最有效的项目监控措施。

3.4.2 促进变革

1. 变革管理

六西格玛项目的实施过程是一个流程变革、系统变革和文化变革的过程。项目的负责人要带领团队迎接变革, 倡导变革。项目管理是一门复杂的科学, 其成败取决于很多因素, 项目管理中的沟通与变革管理是关系到项目成败的关键之一。成功的六西格玛项目不仅仅是一项技术工作, 也是一项变革管理。

变革管理的根本目的是取得项目实施成功所必需的利益相关者的支持与承诺, 同时促使企业全体员工接受并适应新的系统与业务流程。

具体而言, 变革管理必须实现以下目标: 项目团队与业务部门领导、公司决策层之间能进行开诚布公、及时有效的沟通, 从而获得他们的支持、参与和推动; 项目团队内部能进行高效的沟通, 以保证项目团队成员的工作协调一致, 按时、保质、保量实现预期的目标, 并得到认同。

为使整个组织能清楚地理解项目实施的目的、影响与进度, 应做到所有员工都理解项目实施的原因、意义及其对整个组织及组织内部每个部门或个人的影响; 广大员工能看到公司高层领导通过实际行动所表现出来的对于项目实施的支持与承诺; 保证组织合理安排员工的工作职责和角色转换, 以及可能发生的组织结构调整。

对项目所涉及的人员要进行必要的教育与培训, 使其以积极主动的心态迎接可能的变革, 并具有相应的技能来适应这种变革; 加强内外部的宣传与沟通, 为项目顺利推进营造一种适宜的组织氛围。

变革管理工作的核心是沟通。沟通是取得那些会受到变革影响的人的支持的基础。这种支持只有在关键决策者认识到变革的作用并推动变革时才会取得。所以项目负责人不仅是变革的倡导者, 还应该具有良好的沟通能力及影响力。



2. 项目成功的障碍

在推进六西格玛项目的过程中，会遇到各种影响项目成功的障碍，包括：

- 敌意/冲突——对新事物的敌意；
- 退却——知难而退，害怕承担责任；
- 各自为政——意见不一致，各行其是；
- 不适当、不完整的运用——不能做到学以致用；
- 宿命论——没信心，“这不可能成功”；
- 路径依赖——对现有规则的过分遵从或强调。

3. 谈判与解决冲突的技巧

项目的进展不会是一帆风顺的，团队总会面临一些问题及各种障碍，甚至是一些冲突。这时对项目负责人而言就不仅仅是要有管理与控制项目的的能力，还要有消除障碍、解决冲突及谈判的技巧。项目面临的冲突可能来自团队内部，也可能来自团队外部。无论来自哪里，项目负责人的主要任务是对各种冲突持积极主动的正视态度，认真构思冲突解决方案，在必要的时候甚至有意识引导冲突来达到自己的目的。也就是说，通过调控自己在冲突中的角色和地位来驾驭冲突，而不是消极地对待冲突。

项目负责人解决冲突时首先要求对冲突持积极主动的正视态度。也就是说，当冲突发生时，要敢于正视冲突，而不能消极地回避冲突。作为整个项目的负责人，如果不能正视冲突、勇敢面对冲突，结果就会导致团队成员对自己的不信任，甚至会导致冲突的升级直至无法处理。因此，勇敢面对冲突是项目负责人首先要做的。接下来就要冷静地分析冲突产生的根源。冲突是因为人际技巧的缺乏造成的？是因为团队成员间彼此关系不平等造成的？还是因为项目需求变化造成的？……只有找到了冲突产生的根源，才能采取有效的解决方式。

最后，要恰当地对冲突做出反应，有效解决冲突。解决冲突的方式有很多：最好的方法是沟通谈判，使冲突双方尽可能达成一致。而谈判的技巧又有很多，在六西格玛团队中很重要的一条就是团队成员要有妥协的精神。

4. 激励技巧

关于激励的传统观点是马斯洛（Maslow）的需求层次，即人们具有五个基本需求：

- 生存需要；
- 安全需要；
- 归属感；
- 尊重；
- 成就或自我实现。

人们最初的激励来自最低层次的需求。然而，当需求被满足后，这个需求的重要性就降低，人们的激励便来自高一层次的需求。现在许多人在工作中首先要满足的不再是归属需求，这种需求已经通过工作得到满足。他们追求的是自己的尊重和成就的需求。这在知识员工中尤为突出，由此得到五个新的激励因素：

- 目的。人们相信他们的工作是重要的，而且有助于组织的发展。因此，在现阶段矩阵型组织中，明确项目目的可以克服双重报告的不确定性。

- 积极主动。人们希望管理自己的事业发展，而不是呆板服从，通过目标管理向下放

权,让团队成员按计划自主地管理个人发展,进一步激发团队成员各自的内在动力。人们常说,内因是决定因素。当然,项目负责人也要注意对团队成员工作计划的过程监控。

- 分享收获。鼓励员工主动分析问题、解决问题,表扬和鼓励团队成员,并在团队中营造分享收获的氛围,让团队成员将他们在分析问题、解决问题方面的经验进行充分的交流,可以体现员工的成就感,同时让其他成员学习借鉴他人的经验,避免走弯路。

- 个人发展。在马斯洛需求层次顶端的人具有较高的自我实现需求,他们重视学习的机会,让他们参加新项目就是增加他们学习新经验的机会,以增加受尊重感和自我成就感,因此项目负责人要给成员多创造一些施展他们才能的机会。

- 专业上的认可。另一个成就感的度量是专业上的认可。因此,项目负责人应该带领团队成员在推进项目的同时,提高团队成员的六西格玛专业水平和能力,多给他们一些锻炼和展现自己六西格玛专业水平的机会,以获得广大员工和领导的认可。

5. 沟通

对于一个项目团队而言,良好的沟通很有必要,但同时又很困难。项目负责人要对项目结果负责,在项目推进过程中,带领团队营造一个良好的沟通环境是非常重要的。

- 什么时候沟通?是不是出了问题才沟通?当然不是。成功的团队应该随时保持良好的沟通。立项时团队要充分沟通,确保每个团队成员都能够认同项目的意义,愿意为团队的成功贡献自己的智慧。确定项目目标、计划时需要沟通,让每个团队成员都认可项目的目标与计划,才能保证大家努力的方向一致,才能保证项目按计划推进。出现问题后要沟通,要迅速消除大家心中的隔阂,找到问题的解决方案。项目成功后要沟通,使大家一起分享荣誉,增强大家的团队归属感。

- 什么地点沟通?良好的沟通不是只在会议室才能进行。一个轻松的环境有助于团队的沟通取得更好的效果。所以餐桌上、茶水间、运动场等都是团队沟通的场所,“沟通无处不在”。

- 以什么方式沟通?开会并不是最好的沟通方式。大多数团队习惯于以开会的形式沟通。其实沟通的方式有很多,并且有些效果比开会还要好。不同的团队可以根据自己团队的特点,采用不同的沟通方法。因为每个人对沟通的需要不同,每个人的沟通习惯不同,作为项目负责人需要了解每个人的沟通特点并采取有针对性的沟通方式。有的人属于外向型,喜欢与别人讨论,对于这样的人可以采取面对面的沟通;有的人属于内向型,喜欢思考,对于这样的人采取邮件的沟通方式可能会更有效。

沟通的目的是使团队统一思想,采取正确的行动。因此,无论何时、何地、以何种方式,双向互动式的沟通都非常重要。这种交互式的沟通需要双方的投入。项目负责人和团队成员以可以理解的方式,明白无误地传递信息。团队成员应积极处理信息。如果交流中某个环节出现了偏差,沟通双方应该意识到这一点,迅速要求对方给出反馈,或使用其他方法来确定哪些信息没被理解。为了确保高效和交互式的沟通,项目负责人必须:

- 以身作则,做一个良好的沟通模范;
- 肯定员工中优秀的沟通技巧;
- 建立团队的沟通制度;
- 不断进行培训和训练,提高沟通技巧;
- 提供面对面的、电子的以及其他形式的沟通途径,鼓励频繁的、坦诚的沟通;



- 提供充分的非正式交流机会，培养合作关系。

3.5

六西格玛项目管理与策划工具

20 世纪 70 年代以来，日本的一些质量学者提出要转向“思考性的 TQC”，而思考性的 TQC 要求在开展全面质量管理时注意进行多元分析、多元评价，要树立从一开始就不应该失败的概念。这就对质量管理的方法提出了新的要求：要有助于充实计划的内容，要有利于团队协作，要有助于整理语言文字等。在这样的背景下，产生了新七种工具：亲和图、关联图、树图、矩阵图、矩阵数据分析法、PDPC 法、网络图。与老七种工具强调数据分析不同，新七种工具更注重对文字资料的整理，更加面向经营管理和策划活动，是团队使用的重要工具。由于矩阵数据分析图以因素分析和主成分分析为基础，对于统计方法的依赖使得这种方法成为常规情况下非统计人员无法使用的工具，所以后来人们选用优先矩阵图代替它作为新七种管理和策划工具之一。

在使用新七种工具时，最需要注意的是尊重团队成员的意见，以团队的形式来分析。这也是与老七种工具最大的差异。

3.5.1 亲和图

亲和图（affinity diagrams）是由日本学者川喜田二郎（Kawakita Jiko）于 1970 年前后研究开发并加以推广的一种质量管理方法，又称为 KJ 法。所谓亲和图法，就是针对某一问题，充分收集各种经验、知识、想法和意见等语言、文字资料，通过图表进行汇总，并按其相互间的亲和性归纳整理这些资料，使问题得以明确，统一认识和协调工作，以利于问题的解决。

亲和图的主要用途：

（1）归纳思想、认识事物。对未知的事物或领域，广泛收集实际资料，并从杂乱无章的资料中整理出事物的相互关系和脉络，就某件事情达成共识。

（2）打破现状和可能的束缚。通过集思广益产生新的想法，并加以归纳整理。

（3）参谋筹划。不同观点的人集中在一起，很难统一意见。最好能由相互理解的人组成计划小组。为着共同的目标，小组成员提出自己的经验、意见和想法，然后将这些资料编成卡片并进行整理。

（4）贯彻方针。向下级贯彻管理人员的想法和方针。靠强迫命令不会取得好的效果，亲和图可以帮助人们进行讨论，集思广益，从而将方针自然地贯彻下去。

亲和图是用来把大组的项目重组成较小的、容易理解和处理的部分的方法。这种方法通常用在头脑风暴法之后，对罗列出的想法进行组织。可以通过首先创建分类，然后将想法按分类组合，或首先将所有的想法按相似性分组，在组合完之后给分组命名两种方式做亲和图。这个过程可以借助写卡片完成，每个想法写一张卡片，小组成员移动卡片进行组合，直到小组全体成员对组合满意为止。当确实需要时，一些条目可以放在多个组中。

亲和图的主要优点是建立了每个人对各种想法参与讨论的平台，所产生的图表是小组成员联合建立的对所分析问题的概念模型。图 3—7 是“改善团队行为方法”的亲和图例子。

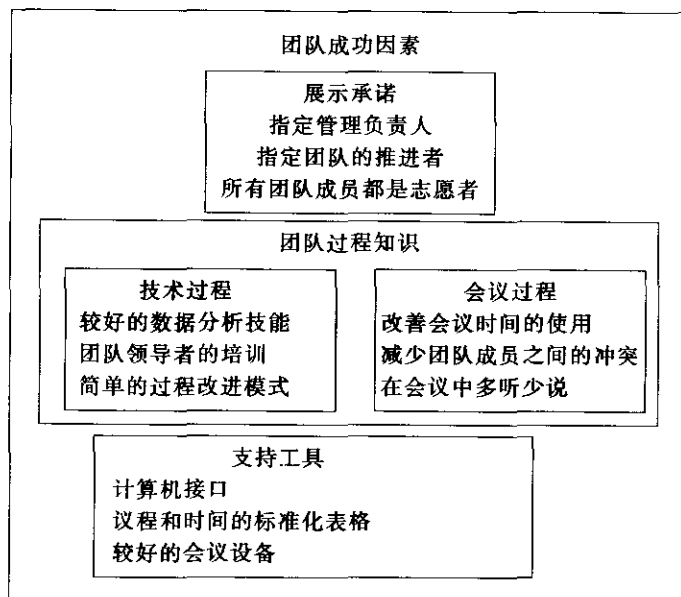


图 3—7 “改善团队行为方法”的亲和图

在亲和图法的使用过程中，资料的收集是至关重要的一环。语言文字资料收集的方法将随亲和图的用途与目的的不同而异。

3.5.2 关联图

关联图（interrelationship graphs）也称关系图（见图 3—8），用于将关系纷繁复杂的因素（因素与因素或因素与问题之间互相缠绕），按原因—结果或目的一手段有逻辑地连接起来，理清复杂问题、整理语言文字资料的一种图形方法。关联图主要用于：

- 界定六西格玛项目的范围和边界；
- 原因分析：因素之间相互缠绕，或者一个因素同时影响两个及两个以上问题；
- 确认关键影响因素；
- 拟订制造过程中预防不良品的措施；
- 提出解决市场问题的措施；
- 改进企业的日常管理活动。

关联图的应用步骤一般如下：

- 确认问题的存在，用图形圈起来；
- 小组成员自由地提出导致存在问题的有关因素（包括原因、理由、个人看法等）；
- 用明确的语言标示出问题和因素的名称，用图形圈起来；
- 确定问题和各种因素间的因果关系，并用箭头线连接起来；
- 根据图形进行分析讨论，检查有无不够确切或遗漏之处，复核和认可上述各种因素之间的逻辑关系；
- 将重要问题或重要因素用双线或粗线醒目地标示出来，并确定从何处入手解决问题。

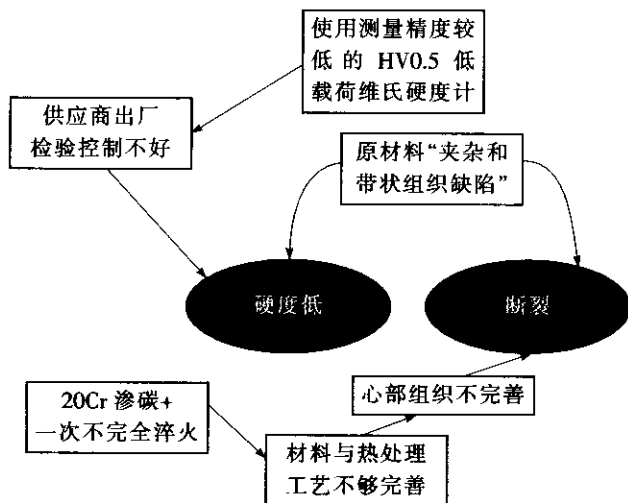


图 3—8 硬度低与断裂的原因分析关联图

3.5.3 树图

树图（tree diagram）又称系统图。它按照“主题—主要类别—组成要素—子要素”的顺序，将主题分解或者分层，以不断地增加细节内容，旨在使笼统的主题分解成较小的部分，使之变得更易于理解和解决。可以自左而右或自上而下地展开。

树图有以下几种类型及用途：

- (1) “目标—手段”类别的树图。将欲实现的目标与需要采取的措施或手段系统地展开，以寻求最佳手段或措施。采用自左而右展开的较多。
- (2) “问题—原因”类别的树图。用于分析质量问题与其影响因素之间的因果关系，以寻求根本原因和对策措施。采用自左而右展开的较多，实际上是因果图的另一种表现方式，如图 3—9 所示。
- (3) 组织结构图。用于描绘组织结构，一般采用自上而下的展开形式。
- (4) 项目管理中的任务分解（WBS）。用于分解项目管理的任务范围，一般采用自上而下的展开形式。

绘制系统图的一般步骤如下：

- 简明扼要地讲述清楚要研究的主题（目标、问题等）；
- 确定该主题的主要类别，即主要的层次。这时可以利用亲和图中的主卡片，也可以利用头脑风暴法中确定的主要层次来确定；
- 构造树图。把主题放在最左侧或上侧的开始框内，把主要类别放在右面或下面的框内；
- 针对这个主要类别确定其组成要素和子要素，并依次画出；
- 评审画出的树图，确保无论在顺序上或逻辑上均没有差错和空当。

在实施六西格玛的过程中，树图可以用来将复杂的改进机会进行分解，从而确定具体的六西格玛项目。如某企业针对提高顾客满意度进行分解得到图 3—10 所示的树图，根据

顾客满意度调查的数据发现对售后服务的抱怨最多，而且绝大部分抱怨的是缺陷产品的处理周期过长，最后确定将缩短缺陷产品处理周期作为六西格玛项目。

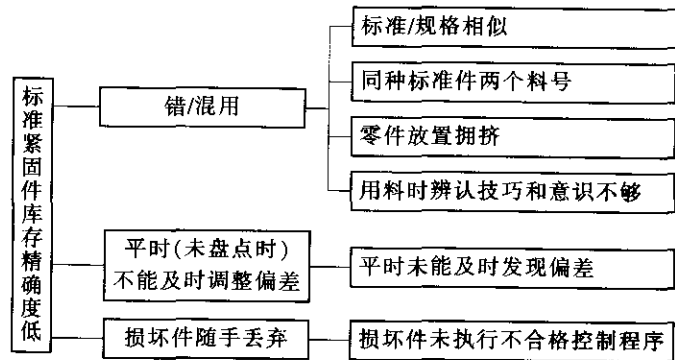


图 3—9 用于原因分析的树图示例

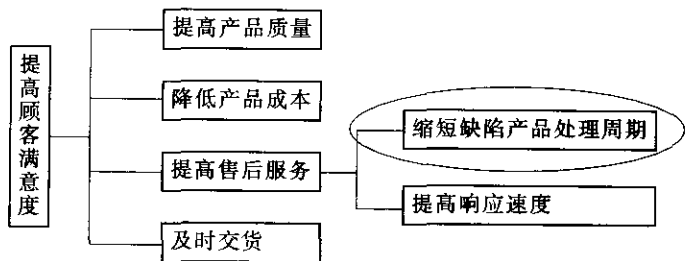


图 3—10 利用树图选择六西格玛项目

3.5.4 矩阵图

所谓矩阵图（matrix diagrams），就是从问题的各种关系中找出成对要素，并按数学上矩阵的形式，把问题及与其有对应关系的各个因素按行和列排成图，并在其交点处标出两者之间的关系，从中确定关键点的方法。

在寻求问题的解决手段时，若目的（或结果）能够展开为一元性手段（或原因），则可用树图法。然而，若有两种以上的目的（或结果），则其展开用矩阵图法较为合适。

在分析质量问题的原因、整理用户需求、分解质量目标时，将问题、用户需求、质量目标（设为 F）放在矩阵图的左边，将问题的原因、用户需求转化而来的质量目标或针对质量目标提出的质量措施（设为 R）列在矩阵图的上方，用不同的符号表示它们之间关系的强弱，一般称 L 型矩阵图。

矩阵图，尤其是因果矩阵，在六西格玛实施中得到大量应用。在六西格玛项目选择过程中，一般用矩阵图进行项目选择和排序，如表 3—13 所示。

表 3—13 矩阵图在六西格玛项目选择中的应用

	对质量的影响	对运转周期的影响	对成本的影响	对管理水平的影响	项目排序
项目 1	◎	○	◎	△	1
项目 2	○	◎	○	○	2



续前表

	对质量的影响	对运转周期的影响	对成本的影响	对管理水平的影响	项目排序
项目 3	○		△	◎	3
项目 4		○	○	△	4

◎—强相关 ○—较强相关 △—弱相关

除此以外，矩阵图法还可用于：

- (1) 设定系统产品开发或改进的着眼点；
- (2) 产品的质量展开以及其他展开，被广泛应用于质量功能展开（QFD）之中；
- (3) 系统核实产品的质量与各项操作和管理活动之间的关系，从而便于全面地对工作质量进行管理；
- (4) 发现制造过程中产生不良品的原因；
- (5) 了解市场与产品的关联性分析，制定市场产品发展战略；
- (6) 明确一系列项目与相关技术之间的关系；
- (7) 探讨现有材料、元器件、技术的应用新领域。

3.5.5 优先矩阵图

区分优先次序是指按照重要性的不同进行安排或者处理。优先矩阵图（prioritization matrices）是树图和矩阵图的结合，它可以帮助决策者确定所考虑的活动或者目标的重要程度。描绘优先矩阵图的目的是促使团队重点关注对组织最重要的关键事项，帮助我们在有着不同收益的多种选择间进行优选。每个行动方案的相对优点和相关影响构建了一张复杂的图，它很难在一个人的头脑中完成优选。例如，一个人在两辆汽车之间进行选择，其中一辆：提供了平均油耗里程、四轮驱动、车窗及座椅控制的全自动装置；而另一辆：有优越的音响系统、乘坐平稳并且有较好的燃料经济性，他将如何选择？对于一个驾驶者来说，四轮驱动可能是在其他所有考虑之上的首要选择；但对于另一个驾驶者来说，可能认为燃料经济性在其他选择之上。

优先矩阵图的建立和使用遵循以下步骤：

- (1) 建立一个矩阵图，将树图的末端项目列为评价项目，并作为行和列的表头标题，然后比较相互之间的重要性，计算每一项目的权重得分（行的相对百分比），如表 3—14 所示。

表 3—14 优先矩阵——评价项目的相关重要度

	便于操作	加速性	舒适性	价格	行总和	行%
便于操作		1	10	5	16.0	42.3
加速性	1		10	5	16.0	42.3
舒适性	1/10	1/10		1/5	0.4	1.1
价格	1/5	1/5	5		5.4	14.3

10—非常重要

5—重要

1—相同

1/5—不太重要

1/10—非常不重要

(2) 对评价项目给出权重后,对所有可能的选择(指参与选择的项目)对照每一项目进行评分。这时需要建立新的矩阵,用来基于每一项目,对各个选择进行评估(例如,表3—15 基于四个评价项目之一的“加速性”,对各个选择进行了评价)。

表 3—15 优先矩阵——基于同一评价项目 (加速性)

	第一辆车	第二辆车	第三辆车	第四辆车	行总和	行%
第一辆车		5	1/10	1/10	5.2	10.2
第二辆车	1/5		1/5	1/10	0.5	1.0
第三辆车	10	5		1/10	15.1	29.7
第四辆车	10	10	10		30	59.1

10—非常好
5—比较好
1—相同
1/5—有点不好
1/10—非常不好

(3) 最终的矩阵是将要作比较的各个选择放在左边的表头上,评价项目放在顶部表头。将来自上述矩阵的权重得分(行的相对百分比值)填入并相乘,对每个选择的得分进行加和,并求得其相对百分比,得分最高的被认为是最好的选择。如表3—16 所示。

表 3—16 优先矩阵——汇总表

	便于操作	加速性	舒适性	价格	行总和	行%
第一辆车	(项目的行百分比× 选择的行百分比)	(42.3×10.2) = 431				
第二辆车		(42.3×1.0) = 42				
第三辆车		(42.3×29.7) = 1 256				
第四辆车		(42.3×59.1) = 2 500				

优先矩阵能够增加决策的客观性(相对于情绪化的决策)。当存在着众多的可能性并具有多种复杂的选择和评价项目时,这个工具可供团队系统地讨论、识别,优选出对做出决策最具影响的评价项目,并评估所有可能,同时他们能够将全貌时时保持在头脑中。

3.5.6 过程决策程序图

在质量管理活动中,要达到目标或解决问题,总是希望按计划推进原定的各实施步骤。但是,随着各方面情况的变化,当初拟订的计划不一定行得通,往往需要临时改变计划。特别是解决困难的问题时,修改计划的情况更是屡屡发生。为应付这种意外事件,一种有助于使事态向理想方向发展的解决问题的方法——PDPC 法被提出。

PDPC 法(process decision program charts)也称为过程决策程序图法,其工具就是 PDPC。PDPC 法于 1976 年由日本学者提出,是运筹学中的一种方法。所谓 PDPC 法,是为了完成某个任务或达到某个目标,在制定行动计划或进行方案设计时,预测可能出现的障碍和结果,并相应地提出多种应变计划的一种方法。这样在计划执行过程中遇到不利情况时,仍能按第二、第三或其他计划方案进行,以便达到预定的计划目标。如图 3—11 所

示, 当 A 方案在 A2 步骤遇到困难时, 可改用 B 方案; 当工作刚开始时 (A0), A1 就受到严重阻碍, 只有使用 C 方案, 而在 C3 步骤也有可能受阻, 则可实施 D 方案, 也能够实现理想目的 Z。

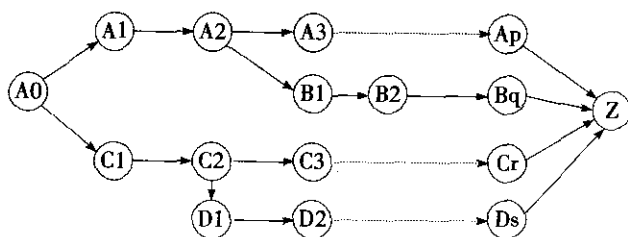


图 3—11 PDPC 法概念图

PDPC 法具有如下特征:

- (1) 从全局、整体掌握系统的状态, 因而可作全局性判断;
- (2) 可按时间先后顺序掌握系统的进展情况;
- (3) 密切注意系统进程的动向, 掌握系统输入与输出间的关系;
- (4) 情报及时, 计划措施可被不断补充、修订。

使用 PDPC 法的基本步骤如下:

- (1) 召集有关人员讨论所要解决的课题;
- (2) 从自由讨论中提出达到理想状态的手段、措施;
- (3) 对提出的措施, 列举出预测的结果及遇到困难时应采取的措施和方案;
- (4) 将各研究措施按紧迫程度、所需工时、实施的可能性及难易程度予以分类;
- (5) 决定各项措施实施的先后顺序, 并用箭条将理想状态方向连接起来;
- (6) 落实实施负责人及实施期限;
- (7) 不断修订 PDPC 图。

图 3—12 是某设备维修改进团队应用 PDPC 法进行设备维修计划安排。

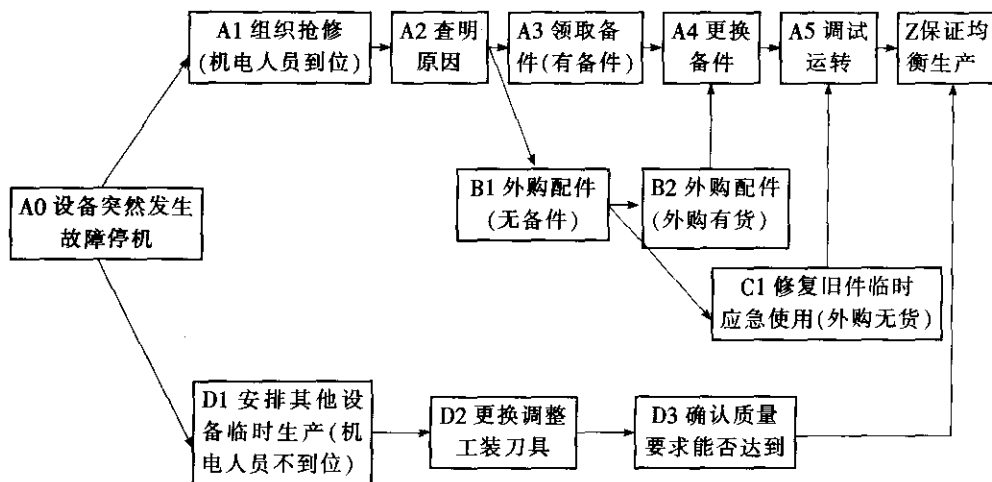


图 3—12 某设备维修改进团队 PDPC 法应用示例

3.5.7 网络图

网络图也称箭条图或矢线图，利用网络图进行计划安排和优化的方法称为网络计划技术，在我国也称为统筹法，是把推进计划所必需的各项工作的时间顺序和从属关系，用网络形式表示的一种“矢线图”（见图 3—13）。一项任务或工程，可以分解为许多作业，这些作业在生产工艺和生产组织上相互依赖、相互制约，网络图可以把各项作业之间的这种依赖和制约关系清晰地表示出来。通过网络图，能找出影响工程进度的关键和非关键因素，从而进行统筹协调，合理地利用资源，提高效率与效益。

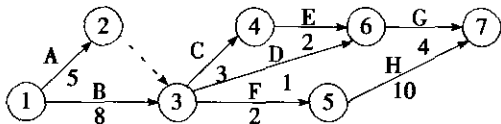


图 3—13 网络图概念图

网络图是一张有向无环图，显示了所有活动的先后关系，也就是任务被完成的先后顺序。其中节点表示事件（如 1, 2, 3, …），箭头表示活动（如 A, B, C, …），箭头边上的权值表示活动持续的时间（如 5, 6, 9, …）。在网络图中，路径最长（权重数值之和最大，图中用粗线表示）的路径称为关键路线，它的长度代表完成整个工程的最短时间，称为总工期。由于只有通过压缩关键路线上的活动时间，才能使整个工期缩短，因此关键路线上的活动是影响整个工程的主要因素，这就是“关键”一词的由来。

关键路线是网络图中一个极其重要的概念。关键路线又称主要矛盾线，其周期决定了整个作业进度的周期。关键路线上的延迟或提前，将直接导致整个项目总工期的拖延或提前。关键路线上的作业称关键作业。关键作业在时间上没有回旋的余地。因此，要缩短总工期，必须抓住关键路线上的薄弱环节，采取措施，挖掘潜力，以压缩工期。关键路线能使管理者对工程做到心中有数、明确重点。

如图 3—14 所示：如果事件 1 是事件 2 的紧前事件，则箭头由事件 1 指向事件 2。事件 2 可能存在若干个紧前事件。有时虽然两个事件之间没有活动存在，但事件 2 必须等待事件 1 完成后才可以开始，则用虚线箭头连接，虚箭头代表一个虚活动。完成虚活动不需要时间，它仅仅用来表示时间发生的先后顺序。

所有活动都指向的那个节点标志着项目的最后完成，称为网络图的交汇点。

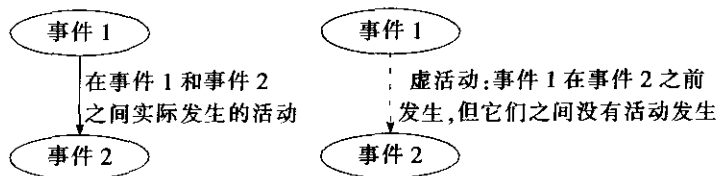


图 3—14 网络图相关术语和作图规范

构造网络图时应遵循下列原则：

(1) 规则 1：网络图中每个活动都由一个箭头表示，且只有一个箭头。一个活动在网



络图中不能出现两次,但是,一个活动可以分解为多个部分。

(2) 规则 2: 不能有两个活动从同一个事件指向同样的另一个事件。正确的做法是引入虚事件和虚活动, 见图 3—15。

(3) 规则 3: 为了保证网络图中事件和活动的先后关系正确, 每添加一个活动时都应先回答下列问题:

- 该活动开始前必须完成哪些活动?
- 哪些活动是该活动的直接后继活动?
- 哪些活动必须和该活动同时发生?

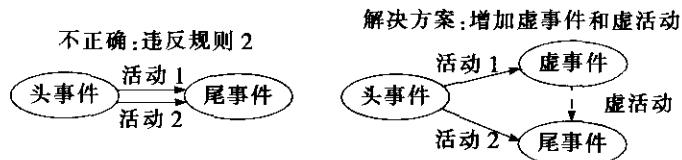


图 3—15 平行活动在网络图中的表示方法

某房屋建设工程中的活动如表 3—17 所示。试构造其网络图。

活动	完成所需的天数	活动	完成所需的天数
开挖	2	地板	4
打地基	4	内部粉刷	5
筑墙	10	内部装修	6
电线铺设	7	屋顶	6
外部管道	4	外墙	7
内部管道	5	外部粉刷	9
墙板	8	外部装修	2

(1) 首先, 要弄清这些活动间的先后次序, 例如只有先筑好墙后才能建屋顶。

(2) 描绘网络图, 计算每个事件的最早结束时间和最晚结束时间。

● 最早结束时间是指该时间的所有先驱活动都及早完成的情况下, 出现该时间的最早时间。从第一个事件开始, 从前往后, 依次计算每个事件发生的最早时间, 要求紧前事件在最早时间发生。

● 最晚结束时间是在不影响整个项目完工时间的情况下, 出现该事件的最晚时间。从最后一个事件开始, 从后往前, 依次计算每个事件的最晚时间, 要求紧后事件在最晚结束时间发生。

(3) 计算事件的延迟时间。事件的延迟时间是事件的最晚结束时间和最早结束时间的差值, 意味着在其他活动和事件都按日程安排正常进行的情况下, 在不影响整个项目完工事件的情况下, 该事件可容忍的拖延时间。

(4) 计算活动的延迟时间。活动的延迟时间是活动的最晚结束时间和最早结束时间的差值，意味着在其他活动和事件都按日程安排正常进行的情况下，在不影响整个项目完工事件的情况下，该活动可容忍的拖延时间。活动的延迟时间=紧后事件的最晚结束时间—(紧前事件的最早结束时间+该活动的时间)。见表 3—18。

表 3—18 房屋建造工程中活动延迟时间的计算

活动	延迟时间	活动	延迟时间
开挖 (1, 2)	$2 - (0 + 2) = 0$	地板 (9, 11)	$38 - (33 + 4) = 1$
打地基 (2, 3)	$6 - (2 + 4) = 0$	内部粉刷 (9, 12)	$38 - (33 + 5) = 0$
筑墙 (3, 4)	$16 - (6 + 10) = 0$	内部装修 (12, 13)	$44 - (38 + 6) = 0$
电线铺设 (4, 7)	$25 - (16 + 7) = 2$	屋顶 (4, 6)	$26 - (16 + 6) = 4$
外部管道 (4, 5)	$20 - (16 + 4) = 0$	外墙 (6, 8)	$33 - (22 + 7) = 4$
内部管道 (5, 7)	$25 - (20 + 5) = 0$	外部粉刷 (8, 10)	$42 - (29 + 9) = 4$
墙板 (7, 9)	$33 - (25 + 8) = 0$	外部装修 (10, 13)	$44 - (38 + 2) = 4$

(5) 寻找关键路径。项目的关键路径就是网络图上延迟时间为零的活动组成的路径（见图 3—16 粗箭头组成的路径），这些活动就是关键活动，其时间总和就是项目完成的总时间。

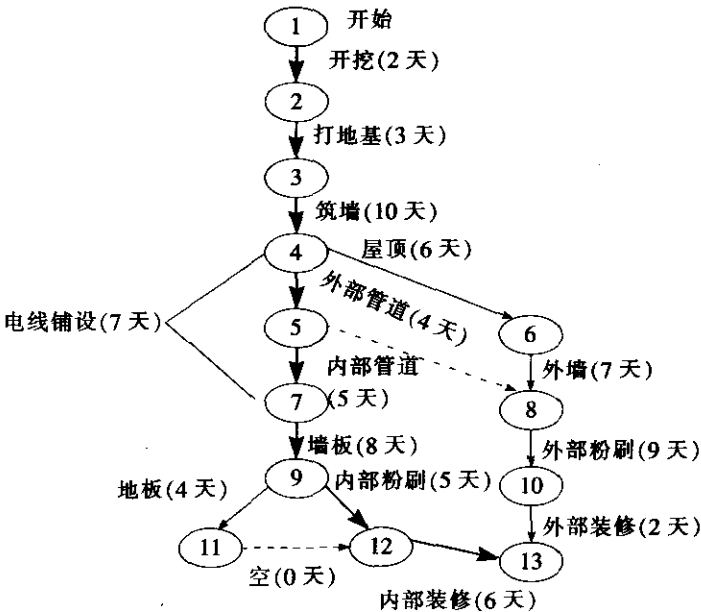


图 3—16 房屋建造工程的网络图

3.6 六西格玛项目总结与成果评审

项目管理的最后一个过程是“项目收尾”，包括行政收尾和合同收尾。对于六西格玛项目而言，主要包括六西格玛项目总结、项目成果评审与分享。

项目总结、评审和分享的过程是六西格玛项目管理 PDCA 循环的 A 阶段，也是实现



“闭环评价和知识管理”的过程，即：通过对达到目标的项目进行文件化并予以闭环评价，将获得的实践经验和教训进行管理，以便识别新的机会，包括与其他相关部门、类似过程以及顾客、供应商、社会等利益相关方分享项目的知识成果，进而使项目的成果倍增，从而实现六西格玛项目知识资产的确认、积累和分享。

3.6.1 六西格玛项目总结

1. 项目总结报告的编写

为交流和承认成果，推进六西格玛项目的持续开展，项目团队应基于项目立项表和项目文档，对活动记录进行整理归纳，撰写成果报告书，要求文字精练，条理清楚，尽量选用图表表达。一般可包括以下内容：

- (1) 前言。
- (2) 界定。项目背景、项目目标、项目计划、项目团队、流程分析等。
- (3) 测量。过程输出绩效的测量、过程因素分析、过程因素测量、测量系统分析等。
- (4) 分析。潜在失效模式与效应分析、关键过程因素与输出绩效的回归和相关分析、假设检验、方差分析等。
- (5) 改进。试验设计、解决方案的确定和实施等。
- (6) 控制。改进效果验证、收益评估（硬的和软的），控制图与过程能力分析、标准化等。
- (7) 经验教训、遗留问题和下一步打算等。

2. 项目总结报告的审核

项目总结报告应报请业务负责人、财务主管、资深黑带（MBB）和倡导者审核。业务负责人从过程的角度，核实项目过程及效果的真实性；财务主管核实项目财务收益的计算方法和结果；资深黑带对项目全过程进行全面的核实；而倡导者则更多地从公司文化和战略的宏观视角审核六西格玛项目的成果。表 3—19 是一家公司的六西格玛项目成果审核表，可供参考。

表 3—19 六西格玛项目成果审核表

项目名称：		组长：	团队成员：
项目目标：		项目启动时间：	项目结束时间：
改进前后状况及数据		改进内容及涉及文件	
改进前：			过程负责人审核 年 月 日
改进后：			
潜在年收益计算：			
实际收益计算：			
财务审核： 年 月 日			
资深黑带审核 年 月 日		倡导者审核 年 月 日	

3. 项目移交

在项目成果得到审核确认后，项目团队应将之移交给过程所有部门或区域，由后者对项目成果进行日常监测和控制。同时，做好项目文档的最终归档（移交给公司档案管理部门或六西格玛管理归口部门）。

3.6.2 六西格玛项目成果评审与分享

企业应当对六西格玛项目成果进行科学、全面的评价，建立符合企业自身特点的激励政策，使改进活动步入良性循环。

1. 成果发布、评审与分享

企业应根据六西格玛管理的推进情况，定期举行六西格玛项目成果发布、评审与分享会，参加者包括：评审人员、项目团队成员以及各相关过程、部门的人员。

企业可建立由倡导者、资深黑带以及与项目相关的业务负责人等组成的六西格玛项目评审委员会，对参与发布的成果进行无记名评分，对各评委的评分进行平均，必要时可结合进行现场核实。评价应鼓励和突出互相学习与交流，注重活动过程的有效性，倡导有特色的成果，兼顾经济、环境与社会效益。对项目成果的评价一是要科学，采用科学的评价方案，要实事求是，不能夸大项目的成果；二是要全方位地评价，不要仅局限于经济效益方面的成果，对不足之处也要指出。项目评审一般包括以下方面的内容：

(1) 项目选择的合理性。

- 选项是否有意义和有价值；
- 项目目标是否具有挑战性和符合 SMART 原则，是否达成目标。

(2) 应用六西格玛理念方法和统计技能。

- 是否正确运用六西格玛的思路、方法解决问题；
- 是否正确、合理运用统计工具和技能推进问题的解决；
- 发布资料是否完整和具有较高的质量。

(3) 项目收益。

- 项目收益计算是否合理、正确，是否为公司取得可观的项目收益或取得了显著的管理效果。

(4) 项目范围和推广应用前景。

- 在项目涉及的范围方面，团队成员是否与项目本身所涉及的职能部门一致，并承担适宜的项目工作，是否体现跨职能合作；

- 项目是否有推广应用价值。

(5) 发布人的思辨、表达、沟通、组织协调、回答问题等软性技能表现。

- 表述是否清晰；
- 时间掌握是否恰当；
- 回答问题是否正确；
- 团队负责人是否具备组织、领导项目组的能力。

除了在企业内部进行发布、分享外，在不违背公司商业机密的前提下，还可以与顾客、供应商和合作伙伴，以及在行业内外发布和分享（如参加全国或地区、行业六西格玛成果发布会和六西格玛优秀项目评选，编写资料或书籍等），促进社会进步和发展。分享



的目的在于进一步扩大改进和创新的效应,在更广泛的意义上降低改进的成本和提高改进的效率,扩大收益区域,为利益相关方创造更多的价值。

2. 成果激励

六西格玛成果是追求卓越绩效的体现,值得公开承认和奖赏,以增强参与员工的自尊和成就感。企业需要根据所在行业的特点以及自己的条件,制定多样化的激励方式和手段,包括物质的和精神的激励。

企业可依据项目评审的结果,设立优秀六西格玛项目、六西格玛成就奖、明星黑带和绿带等各种奖项,授予项目团队和成员。

(1) 物质激励。可包括:

- 奖金;
- 奖品;
- 晋升与调薪。

要基于个人贡献的大小来评价团队成员在六西格玛团队中的绩效,进行奖金分配。

(2) 精神激励。可包括:

- 让项目黑带或绿带有机会登台发布成果,可邀请其亲属、朋友参加;
- 高层领导签署表彰令,亲自参与成果发布会议、成果表彰仪式或宴会,授予奖状、奖章或奖牌(可刻上项目名称和黑带等团队成员姓名);
- 将项目报告及团队成员照片刊登在公司的新闻媒介上;
- 完成六西格玛管理培训也是一种成果,因此可给完成课程的人员授予证书、徽章等。

3. 黑带与绿带认证

黑带与绿带认证也是一种激励,一种对员工职业生涯发展的激励。企业可以自行认证,也可以鼓励员工参加权威的第三方机构(如中国质量协会)组织的资格考试、项目评审和注册。

一般地,企业内部的黑带和绿带认证应具备以下条件:

- 完成黑带知识课程培训,至少完成两个项目,取得显著收益,并为通过企业内部成果评审的项目组长,予以公司内部黑带认证,发给黑带证书;
- 完成绿带知识课程培训,组织完成或参与完成年收益较显著的若干个项目(如1~2个),并通过企业内部项目成果评审,予以公司内部绿带认证,发给绿带证书。

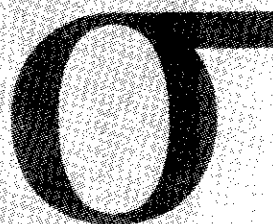
本章小结

本章从项目管理整体过程的角度,重点介绍了六西格玛项目选择,项目立项表和计划,项目团队建设,项目监控和促进变革,项目管理和策划工具,以及项目总结与成果评审。六西格玛管理是通过有组织、有计划地实施六西格玛项目而实现其经济效益的。对一个企业来讲,通过六西格玛项目的实施可以促进人们观念和行为方式的转变;同时,六西格玛项目又是黑带、绿带成长过程中的必经之路,黑带、绿带除了接受一系列知识、方法、工具的学习之外,还必须经受六西格玛项目从寻找改进机会到启动立项、规划、执行、监控、收尾总结全过程的锻炼,才能成长为一个真正的黑带、绿带。因此,六西格玛项目的实施在变革企业文化、培养统计技术方面的专家、促进企业的持续

改善及获得实际的经济效益等方面，都有着十分重要的意义。特别是在企业导入六西格玛管理的初期，六西格玛项目的成功与否，还关系到企业能否以较小的阻力引入这种新的管理方法并持久地进行下去。可以说六西格玛项目的选择与实施，是六西格玛管理的一个关键环节。

第 2 篇

DMAIC 流程



第 4 章 界定
第 5 章 测量
第 6 章 分析
第 7 章 改进
第 8 章 控制

界 定

界定 (define) 是六西格玛 DMAIC 方法的第一个步骤,也是非常重要的一步。六西格玛项目是从界定阶段开始的,一般说来,在初选出项目时,对欲解决的问题往往仅有比较宏观的考虑。项目团队需要通过界定阶段的工作,明确问题或者流程输出 Y 及其测量,确定 Y 的标准。项目团队还需要明确项目的关注领域和主要流程,将项目界定在一个比较合理的、团队可以把握的范围之内。通过界定阶段的工作,项目团队要明确项目的目标,测算项目预期收益,确定项目核心团队成员等。一些质量工具和方法,将帮助黑带及团队完成上述工作。常用的有:顾客需求 (VOC) 分析、卡诺模型 (Kano model)、SIPOC 分析、相关方分析、排列图等。所有这些工作的结果,将纳入项目立项表 (project charter) 或特许任务书的文件,并得到主管领导以及倡导者的批准。

4.1 界定项目范围

4.1.1 确认顾客关键要求

确定顾客的关键要求,是六西格玛项目工作中重要的一步。整个项目工作是在这个步骤的基础上展开的。因此,在六西格玛项目工作的界定阶段,需要特别关注这一点。顾客 (内部和外部) 只有在其需求得到充分理解并获得满足后,才会满意和忠诚。如何准确地定义顾客的需求及其要求,是黑带及其团队的职责。当顾客的需求被正确理解并被恰当地转换为过程输出的关键质量特性 (critical to quality, CTQ) 时,项目才具备了展开的基础。

1. 识别顾客

确认顾客关键要求的第一步是识别关键顾客,在任何过程中,产品或服务的接受方即顾客。在本书第 2 章,已经详细介绍了识别顾客及其关键要求的方法。这里需要强调的是,在项目界定阶段,团队需要明确项目所关注的顾客,并且确定这些顾客的需求和关键要求。

一个过程可能会有许多类型的顾客同时存在。正像前面所介绍的那样,在过程的顾客中,有外部顾客和内部顾客;中间顾客和最终顾客;当前顾客和未来或潜在顾客;以及关键顾客和一般顾客等。在项目界定阶段,需要识别项目需要关注的主要顾客是谁以及他们的要求是什么。如果项目的关键顾客不能识别清楚的话,可能会误导整个项目的方向,使



项目最终误入歧途。

可以借助 SIPOC 分析法, 识别项目关注的顾客, 包括内部顾客和外部顾客等。另外, 许多专家建议, 团队在识别项目关注的顾客时要注意避免片面性, 因为在任何细分市场上, 组织的顾客将由以下部分构成:

- 当前满意的顾客;
- 当前不满意的顾客;
- 失去的顾客;
- 竞争对手的顾客;
- 未来的潜在顾客。

2. 确定顾客关键要求

在识别项目的顾客的基础上, 需要确定顾客关键要求。这项工作是在顾客调查的基础上完成的。常用的顾客调查的方法有:

- 问卷调查;
- 焦点小组调查;
- 顾客访问;
- 从顾客投诉系统获得信息;
- 市场调查等。

除上述方法外, 近几年随着信息化管理技术的发展, 一些新的顾客信息分析方法也可以帮助项目团队获得顾客要求的信息, 例如:

- 顾客计分卡;
- 数据仓库;
- 数据挖掘等。

本书第 2 章第 2 节中已经详细介绍了常用的顾客调查以及顾客数据的分析方法。关于这部分内容可参考前面有关章节的内容。

许多情况下, 需要在顾客调查的基础上, 对顾客需求进一步展开, 以便将其转化为具体可测量的要求, 并细化到团队可以把握的层次, 便于项目工作的开展。本书第 2 章介绍了应用亲和图法归纳整理不同层次的顾客要求, 形成多级展开的顾客要求表的方法, 并给出了顾客要求展开表的例子。图 4—1 所展示的是某服务过程的顾客要求的树图展开, 它是顾客需求展开的另一个例子。

采用树图展开顾客需求的一个好处是为设计调查顾客满意度问卷提供了框架, 而且基于问卷结果的分析, 可以发现哪些方面是急需改进的项目。

4.1.2 过程的 SIPOC 分析

在项目的界定阶段, 需要将内部流程与顾客关键要求联系起来, 识别项目重点关注的内部流程。这个识别过程首先从宏观的流程分析 (也被称为高阶流程分析) 开始。而 SIPOC 分析方法则是通常使用的宏观流程分析方法。

1. SIPOC 图的构成

SIPOC 图, 也称高端程序图 (见图 4—2), 名字来自于供方、输入、过程、输出和顾客的第一个英文字母的缩写。其中:

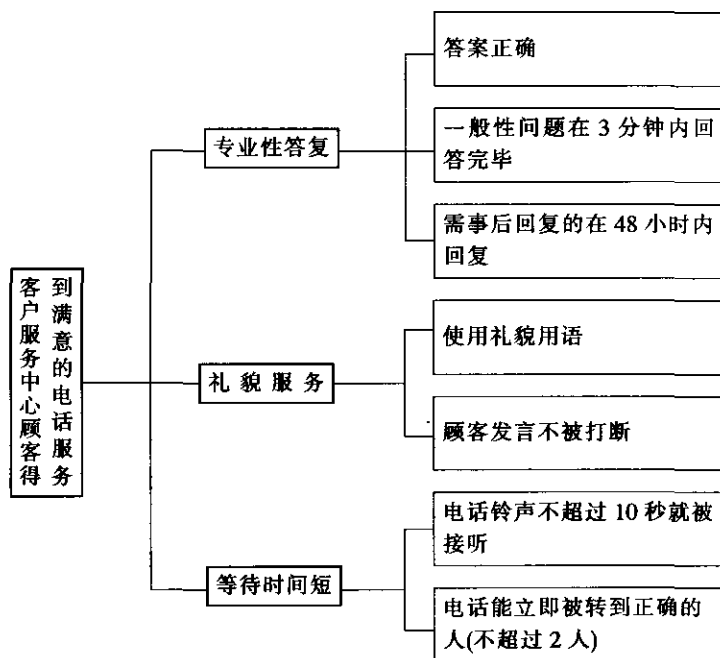


图 4—1 顾客要求的树图展开

供方 (supplier): 提供输入的组织和个人, 在六西格玛项目管理中, 专指向过程提供关键信息、材料或其他资源的个人和组织, 供方可以是内部的或外部的。

输入 (input): 供方提供的信息和资源, 包括人员、机器、材料、方法、环境等。

过程 (process): 将输入转化为输出的活动。过程是使输入发生改变的一组步骤, 理论上, 这个过程 (由这些步骤组成的过程) 将增加输入的价值。

输出 (output): 过程的结果。

顾客 (customer): 接受输出的人、组织或过程。

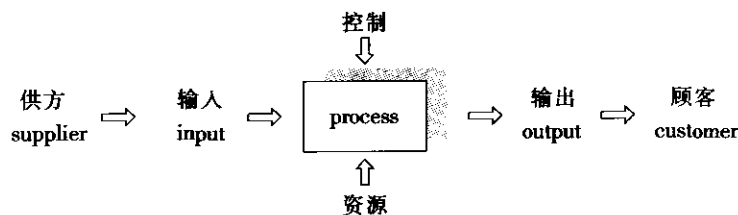


图 4—2 SIPOC 图

SIPOC 图上还可以加上过程输入和过程输出的基本要求, 用来表示一个业务流程或产品 (服务) 实现过程中的主要活动或子过程, 帮助项目团队界定过程的范围和过程的关键因素, 确定关键输入变量 (KPIV) 和关键输出变量 (KPOV)。SIPOC 图的优点是: 能在一张简单的图中展示出从输入到输出的跨职能活动。因此, 在项目界定阶段应使用 SIPOC 分析和 SIPOC 图, 来识别项目涉及的主要业务流程和相关职能, 确定项目范围, 识别项目相关方并组织项目团队。在界定阶段, 通过 SIPOC 分析, 使黑带们保持对流程的宏观



认识是非常重要的。

2. SIPOC 图绘制步骤

为了提供一份宏观的过程步骤和关键的供方、输入、输出和顾客关联关系的 SIPOC 图表（见表 4—1），需要确定项目所关注的过程起点和终点，在此基础上，确定相应的供方和输入，以及输出和顾客（包括组织内部的和外部的顾客）。

表 4—1 SIPOC 工作表

供方	输入	过程	输出	顾客

西蒙（Simon）建议，采用下述办法带领团队来绘制 SIPOC 图：

（1）用一块足够大的墙面或白板，供团队成员用小纸条来讨论和列写过程的供方、输入、过程、输出和顾客，直到最后确定它们。

（2）过程的步骤不宜列写得过细，用 4~5 个关键宏观步骤来表达从输入到输出的核心业务流程。

（3）团队先来讨论过程的输出是什么。过程有哪些结果（产品或服务）产生？应该包括哪些内容？过程的输出点或结束点应该在什么地方？

（4）列出过程输出的顾客。谁是过程结果的用户/使用者？

（5）列出过程的输入。过程的输入来自何处？

（6）列出提供输入的供方。谁是关键的供方？

（7）顾客的主要要求是什么？

（8）与团队的负责人、流程主管、倡导者以及其他相关方一起，确认项目的 SIPOC 图。

表 4—2 是一个设备租赁过程的 SIPOC 工作表的例子。

表 4—2 SIPOC 工作表（样例）

供方 S	输入 I	过程 P	输出 O	顾客 C
申请人	租赁申请	(见下表)	批准的申请表	申请人
	资质证明		出租的设备	
	信用证明		随机文件	
信用调查部门	信用报告		服务信息	

过程（process）

顾客信用调查	⇒	设备确认与准备	⇒	随机文件的准备	⇒	收抵押金	⇒	交付
--------	---	---------	---	---------	---	------	---	----

3. SIPOC 图表的检查

为了保持 SIPOC 图形的正确性，可采取下面一些检验活动：

(1) 如果在 SIPOC 图形的“P”中,步骤数量超过了7~8个,则项目团队对过程的考虑可能过于详细了。在项目界定阶段,需要在比较宏观的层面上把握项目范围,审视项目需要关注的流程,需要以宏观视角将包括在项目工作中的相关过程识别出来。过细的流程图可能会导致团队的视角变窄,忽略了应该在项目中考虑的跨职能合作的问题。

(2) 如果 SIPOC 图形显示出需要对项目涉及的步骤、输出或输入进行重新确定,则意味着需要调整项目的范围。

4.1.3 关键过程输出变量的确定——排列图法

在前述过程的基础上,项目团队应当进一步确定项目应关注的关键过程输出变量(KPOV)或称为关键质量特性(CTQ)。

1. 关键过程输出变量

对符合顾客需求极为重要的过程输出项目,称为关键过程输出变量(KPOV)。需要根据对顾客要求的分析,依照过程输出特性对顾客要求的影响程度的高低确定项目关注的输出变量和项目的优先顺序。

通过前面介绍的 SIPOC 图的分析,可以识别出过程的所有输出,以及对于输出的测量,即过程的输出变量,如顾客满意度、交货周期、产品的合格率以及缺陷数等。通过收集数据,可以观察这些输出变量的变化及其对顾客的影响,并进一步确定项目应关注的过程的关键输出变量。

在六西格玛管理中,经常把结果或输出变量记为Y,把输入变量记为X;X对Y的影响以及它们之间的相关关系,可以用下述函数关系来描述,即

$$Y=f(X_1, X_2, \dots, X_n)$$

有时,可直接把 KPOV 以 Y 表示。

2. 用排列图确定过程的关键输出变量

采用排列图方法是确定过程关键输出变量Y的常用方法。排列图是建立在帕累托(Pareto)原则之上的,即80%的结果源于20%的原因。排列图的目的是比较不同的缺陷类型所导致的结果或对顾客的影响,以便找出最重要的需要优先关注的问题;设置优先权并定义问题/机会;确定关键输出变量。

(1) 排列图作图步骤。通常采用的排列图的作图步骤如下:

步骤1:收集过程输出在某时间段内的数据,并确定过程输出中都有哪些不符合顾客要求的缺陷项。

步骤2:将各项缺陷发生的频数按从大到小的顺序排列,计算各自占总缺陷数的比例(%)和累计比例(%)。

步骤3:将横坐标按从大到小的顺序,依次列出各种缺陷项。

步骤4:以左侧纵坐标为缺陷发生频数,右侧纵坐标为比例(%)。

步骤5:在横坐标上的每个缺陷项处,画出与其发生频数对应的矩形。

步骤6:由左至右累加每个缺陷项的比例,画出累积频率曲线。

表4—3和图4—3是某服务过程输出缺陷排列图作图和分析的一个例子。

表 4—3 某顾客服务过程输出缺陷排列图计算表

缺陷类型	发生频数 (2005 年 1—10 月)	累积缺陷数	比例 (%)	累积比例 (%)
A: 顾客等待时间长	51	51	51	51
B: 迟于规定时间回复	25	76	25	76
C: 回复不准确	13	89	13	89
D: 遗漏顾客信息	7	96	7	96
E: 账单错误	2	98	2	98
F: 收费错误	2	100	2	100
合计	100		100	

图 4—3 是根据上述排列图计算表作出的排列图。

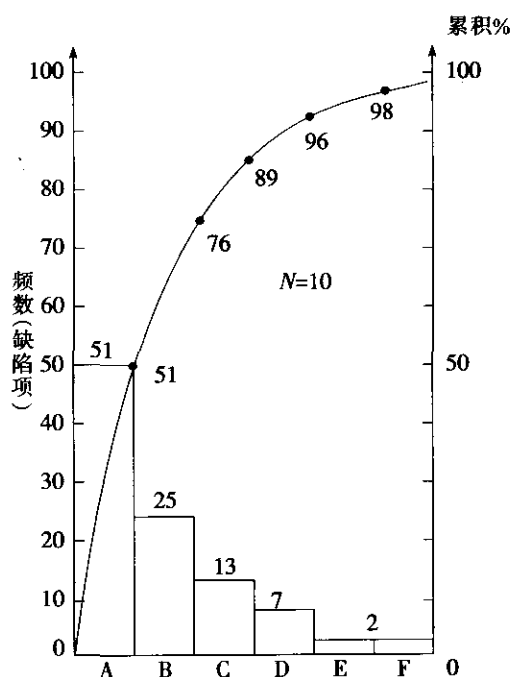


图 4—3 顾客服务流程缺陷的排列图

从排列图上可以看出, 占发生频数总数的 80% 左右的缺陷是 A 类缺陷和 B 类缺陷, 即顾客等待时间长和迟于规定时间回复。因此, 项目需要改善的过程关键输出变量应该是: 顾客等待时间; 回复时间。

采用计算机作排列图时, 作图步骤将大大简化。图 4—4 所示的是用 MINITAB 软件作出的某过程输出缺陷的排列图。从这张排列图上可以看出, 项目应当关注的过程的关键输出变量 KPOV 或 Y 应当是“气孔”和“夹渣”, 这两种缺陷占过程输出总缺陷数的 82% 左右。

(2) 关于加权排列图。排列图在帮助团队识别项目需要关注和解决的重点问题或过程关键输出变量方面, 是非常有效的工具。在前面介绍的排列图作图与分析过程中, 我们按照缺陷发生的频数来识别“关键的少数”和“非关键的大多数”。但是, 在一些情况下, 仅就缺陷发生的频数来识别“关键”与“非关键”并不合理。比如, 对那些“一旦发生将

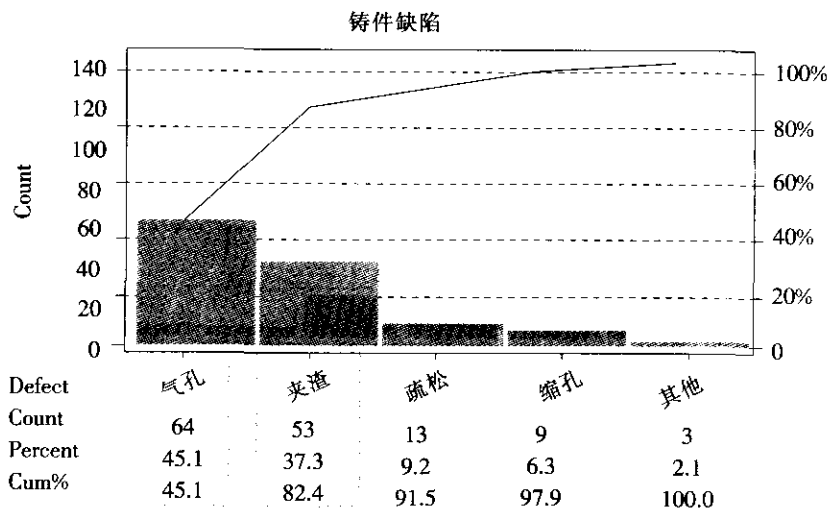


图 4—4 某铸造过程输出缺陷的排列图分析

造成顾客极端不满意”的缺陷与“即便发生顾客也不会太在意”的缺陷，按它们发生的频数进行排列来识别改进机会，则是非常不合理甚至会误导项目改进的大方向的。在这种情况下，可以采用加权排列图来识别改进机会。

仍然以前面提到的某服务过程输出缺陷排列图作图和分析作为例子。这里为各种输出缺陷类型对顾客不满意的影响设置权重，则在考虑权重之后的加权缺陷数、比例和累积比例如表 4—4 所示。

表 4—4 某顾客服务过程输出缺陷的加权排列图计算表

缺陷类型	权重	发生频数 (2005 年 1—10 月)	加权频数	累积加权 缺陷数	比例 (%)	累积比例 (%)
A: 顾客等待时间长	5	51	255	255	41.7	41.7
B: 迟于规定时间回复	5	25	125	380	20.4	62.1
C: 回复不准确	10	13	130	510	21.2	83.3
D: 遗漏顾客信息	10	7	70	580	11.4	94.7
E: 账单错误	8	2	16	596	2.6	97.3
F: 收费错误	8	2	16	612	2.6	100
合计		100	255			

图 4—5 所示的是用上述加权数据作出的排列图。

从图 4—5 中可以看出，对顾客满意度影响最大的关键缺陷类型有：A——顾客等待时间长，C——回复不准确，B——迟于规定时间回复。由此可以确定，项目应当关注的键过程输出变量 KPOV 或 Y 应当是：顾客等待时间；回复的准确性；回复时间。

当观察排列图时，需要确定是否“关键的少数”引起了问题的绝大部分。如果没发现帕雷托效应，则需要寻找或采用其他分析模式。

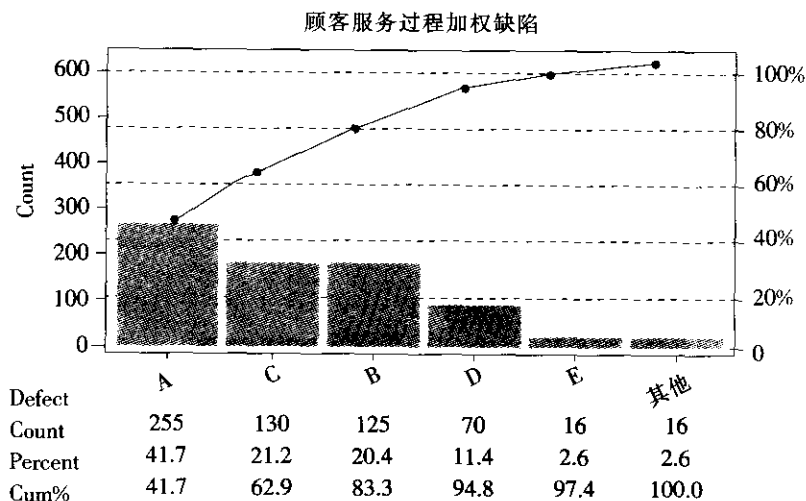


图 4—5 某顾客服务过程加权缺陷的排列图

4.1.4 项目的利益相关方分析

每一个过程都有顾客和受过程输出影响的其他相关方，他们可以是组织内部的，也可以是组织外部的。利益相关方指受组织的行动以及成功与否影响的所有群体，关键的利益相关方（stakeholder）通常包括顾客、股东、供应商、合作伙伴以及社区等，在组织内部，过程的相关方还包括管理者和员工等。

六西格玛团队常常需要跨职能开展工作，要确保团队活动的有效开展及项目的顺利进行，需要对项目的利益相关方进行确认和分析。项目的利益相关方就是与项目有着密切关系和相互关联的部门或人员，通常包括：

- 被项目结果所影响的部门或人；
- 影响项目结果的部门或人；
- 有决定权的人；
- 提供资源的部门或人；
- 流程的相关专家；
- 为项目提供数据/信息的部门或人。

为确保六西格玛项目的顺利进行，识别上述利益相关方，让他们对六西格玛项目产生正确的认识，并提供有效的支持是十分重要的工作。

瑞斯和斯特朗（Rath & Strong）以及贝克哈德和哈里斯（Beckhard & Harris）都提出了识别项目利益相关方并制定项目沟通和影响策略的方法。表 4—5 所示的就是项目利益相关方分析的一个示例。

表 4—5 项目利益相关方分析表

承诺水平	销售	管理	员工	顾客 A	供应商 A
支持		×○			
顺从	○		○		○

续前表

承诺水平	销售	管理	员工	顾客 A	供应商 A
无关紧要				×○	
不合作			×		
反对	×				
敌对					×
不需要					

注：×表示当前的承诺水平；○表示需要达到的承诺水平。

在此基础上，团队应当制定减小或消除阻力的项目相关方的沟通与影响策略，以使所有项目相关方的承诺水平达到期望的水平。比如，上表中表明，销售部门对本项目持反对态度，而本项目的成功实施需要销售部门的承诺从“反对”至少转变到“顺从”，因此项目团队需要制定相关的沟通与影响策略，消除来自销售部门的阻力，使其态度得到相应的转变并达到期望的承诺水平。同理，项目团队应当完成其他相关方的分析并制定相应的策略。

4.2 确定项目测量指标

4.2.1 关键质量特性 CTQ 的树图展开

界定阶段常常使用树图分析技术将关键质量特性 CTQ 进行适当地展开。CTQ 树图将帮助项目团队选择出与顾客满意密切相关的关键变量及其测量。一个 CTQ 展开树图可以将最初的顾客要求转化成可量化的产品或服务的要求。而这些可量化的变量可以视为过程的输出 Y 的关键测量项。

一个 CTQ 树图的展开过程将从一般的顾客要求开始逐层展开至具体的过程输出要求，或者说从难以测量的特性逐步展开至容易测量的特性。一般说来，将一般需求展开到过程的关键质量特性 CTQ，需要经过 2~3 个层次的展开。

构建 CTQ 树图的步骤如下：

- 识别顾客；
- 识别顾客需求；
- 识别第一层顾客要求；
- 逐层细化到适当的层次；
- 确认这些要求，以确保 CTQ 树图反映顾客要求。

某快餐店的 CTQ 树图的展开

某快餐店的顾客是那些希望方便快捷地解决吃饭问题的消费者。他们的主要需求是：方便快捷地订餐。各层 CTQ 树图展开见图 4—6。

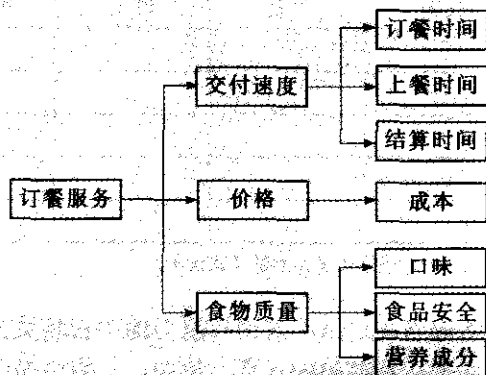


图 4—6 快餐店 CTQ 树图展开示例

4.2.2 建立测量指标

在 CTQ 树图展开的基础上，团队需要针对项目关注的 CTQ 选择适当的测量指标，以便建立测量、确定标准、定义缺陷。可以按下述步骤选择出适用的 Y 的测量指标。

1. 选出要测量什么

在界定阶段，团队确认了 DMAIC 项目要解决的主要问题，以及关键的顾客要求。这时要根据测量来获得对问题严重程度及出现频率的了解，以及对于顾客要求的达成程度的评判。在选择测量指标时应注意测量指标对问题分析的价值，以及收集数据的可行性和难易程度。表 4—6 是选择测量指标时的一些参考建议。

表 4—6 测量指标选择时的一些参考建议

可用性/测量指标的价值	可行性/难易程度
● 与重要的顾客要求有关	● 是否有历史数据供分析使用
● 能准确表达顾客要求的满足程度	● 测量周期是否很长
● 易于暴露问题与改进机会	● 测量是否复杂
● 可以作为与其他组织进行对比的标杆	● 是否与其他测量指标冲突
● 可以持续测量并提供有价值的信息	● 是否可以重复和再现

2. 用 CTQ 树图来确认测量

通过 CTQ 树图展开，Y 的测量会进一步明确。可以通过 CTQ 树图进一步明确项目的测量。例如图 4—6，对快餐店的顾客来说，顾客要求的一个重要方面是交付速度，这个问题也是团队重点要解决的问题。那么，团队选择的测量指标可以是：（1）订餐时间；（2）上餐时间；（3）结算时间。对这些测量指标，团队可以考察它们的可用性和可行性，以便确认它们是否是合适的测量指标。

3. 测量指标的细化及测量评估

更加详细的测量以及数据收集方法的确定，对整个项目的实施会非常有帮助。但是，在项目界定阶段，关于测量指标的选择可以在比较高阶的层面上展开。更加详细的测量以

及数据收集的方法，将在测量阶段展开，这里不再赘述。

4.2.3 测算项目的西格玛水平

选择了适当的测量指标之后，就可以在此基础上计算项目所关注的过程、产品和服务的基线西格玛水平。这个基线西格玛水平可以作为基准，据此制定项目目标和评价改进效果。计算基线西格玛水平的过程由三个基本步骤构成。

(1) 确定“单位”及“缺陷”和“缺陷机会”。关于“单位”、“缺陷”和“缺陷机会”的概念，在第2章中已经做了阐述，这里强调如下：

1) 单位。是指被处理的项目或提供给内部顾客与外部顾客的产品或服务。例如，一个产品、一个零备件、一项贷款、一个电脑平台、一次医疗诊断、宾馆的一夜住宿或一次信用卡服务等。

2) 缺陷。凡未达到顾客要求或过程输出绩效标准之处，均为缺陷。例如，拙劣的烤漆、贷款手续延迟、错误的处方、预订错误或订单出错等。

3) 缺陷机会。产品或服务未达顾客要求或绩效标准的机会。在这些用语中，缺陷机会最难分辨，但却是计算西格玛水平的重要变量。计算缺陷机会的优点在于可以让我们比较复杂度不同的过程。

(2) 收集历史数据，按照确定的单位、缺陷和缺陷机会，计算 DPU，FTY，DPO，DPMO 等。

关于计算 DPU，FTY，DPO 和 DPMO 的方法，请参考第2章。

(3) 计算过程的西格玛水平。可以根据统计计算出的过程的 DPMO 查表或计算出过程对应的西格玛水平。相关内容，请参见本书第5章“过程能力分析”一节。

4.3 编制和完善项目立项表

编制和完善项目立项表或项目特许任务书是界定阶段工作的主要输出。项目立项表表述项目的范围，阐明团队使命，陈述问题和机会，明确项目的目标和预期收益，并且帮助团队沟通他们的分工和计划。立项表是项目工作的第一个并且通常是最重要的一个里程碑，团队开始实施项目之前，立项表必须先由倡导者等领导批准。

作为界定阶段团队活动标志性的成果，立项表直接反映了项目团队活动的质量，直接关系到项目的进展和 DMAIC 过程的实施。

一个完备的项目立项表主要包括项目背景、改进机会、项目目标、项目范围、项目计划和团队构成等几个部分组成。

4.3.1 项目背景

在选择六西格玛项目时，黑带和倡导者们首先要回答这样一个问题：为什么要实施这个六西格玛项目？为了进一步取得管理部门和团队成员的共识，有必要在项目立项表中对



此做出明确的书面陈述,表明企业及其过程的经营情况,阐述项目对企业发展战略的意义,从而说明为什么要做这个六西格玛项目。

某企业的立项表是这样阐述其项目背景的:进入新一代数码产品国际市场是企业“十一五”计划期间的重要战略增长点。2005年,作为首批进入国际市场的 SM300 系列产品,遇到了较强的市场竞争。竞争压力主要来自于交付周期。主要竞争对手的交付周期达到了 10 天,而本企业 SM300 的交付周期平均为 18 天。由此而带来的直接市场损失为 2 600 万美元/年。

在阐述项目背景或业务经营情况时,需要回答以下几个问题:

● 什么是本企业或经营战略所关注的问题?为了达到企业的战略目标,当前遇到的首要问题是什么?

● 为什么这是值得关注的问题?

● 如果业务流程或经营情况不改变的话,将会产生什么后果?企业的经营损失是什么?

4.3.2 问题/机会和目标的陈述

在立项表中,问题/机会和目标的陈述主要是阐明:本项目要解决的问题是什么?项目的目的是什么?团队的使命是什么?预期达到怎样的目标?

1. 问题/机会的陈述

(1) 问题陈述准则。一个有效的问题/机会的陈述必须遵守以下几个准则:

● 具体的。明确地说明问题是什么。

● 客观的。描述了问题/机会可见的迹象,问题发生在哪里?

● 可测量的。这问题有多严重?已经发生了什么影响?用量化的语言来回答这些问题。

● 可控的。阐述的问题是否属于本企业职能范围内的?能否在规定的时间内(一般为 4~6 个月)解决?如果一个问题太大,它应该分解成若干个较小的可控的项目。

(续例 4—2) 某企业的六西格玛项目立项表中的问题/机会描述是这样的:统计数据表明 2005 年 6—12 月, SM300 数码仪的平均交付周期为 18 天,且波动较大。最快的交付批其交付周期可以达到 8 天,但最迟的交付批达到了 32 天。如果以交付周期不超过 12 天为标准,则大约有 82% 的交付批次超过了这个标准。

这个陈述是:

● 具体的。指明了 SM300 系列产品,且关注的问题是交付周期。

● 客观的。从内部的报告和交付记录中可获得有关数据和事实。

● 可测量的。以交付时间(即从订货到发货的天数)来测量。

● 可控的。交付周期问题在本企业的职责范围内。

(2) 问题陈述注意事项。下列问题在陈述时应引起团队注意：

- 陈述观点时，不应包含造成缺陷的原因。要由项目团队经过相应阶段的工作之后，再对问题产生的原因做出客观准确的判断。预先形成的有关原因的想法可能是不准确的、不完全的或者是错误的，并且可能误导团队。
- 不应提出一个改进方案。在不清楚根本原因的情况下，制定的解决方案是主观的，不是基于数据和事实的决策。
- 陈述问题时，不对相关责任人员加以指责。项目工作的目的是为了改进流程，通过流程的优化达到解决问题的目的，而不是追究相关人员的责任。六西格玛项目应当是“对事不对人”的。
- 不要将几个问题集中在一个问题的陈述中，不同的问题应当由不同的项目来解决。

2. 目标陈述

目标陈述将明确团队为之努力的方向，以及通过六西格玛项目将取得什么样的成效。项目的目标陈述与问题的陈述应当对应。项目目标必须符合 SMART 原则，有关 SMART 原则的内容详见第 3 章。

(续例 4—3) 该项目立项表中的目标陈述为：以交付周期不超过 12 天为标准，到 2005 年底，将 SM300 系列产品的交付周期超过 12 天的比例（缺陷率）从 82% 降到 25%。

在此例中，问题是以 SM300 的交付周期超过标准的比例（相当于流程的缺陷率）来表达的，目标也是以这个比例来表述的，与问题陈述相对应，目标具体、可测，而且基于流程分析和历史最好水平，该目标是可以实现的。

3. 目标的度量

项目目标的度量可分为：缺陷率、周期时间、费用成本三种类型，而通常采用的度量指标又有以下几种（见图 4—7）：

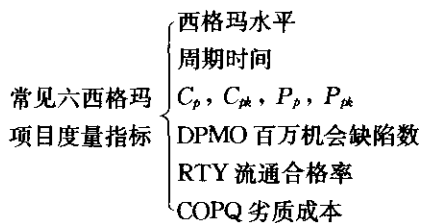


图 4—7 项目目标的度量

4.3.3 项目范围、约束和团队任务

1. 项目范围

项目团队要对项目范围有一个清楚的定义。项目范围的确定可借助于 SIPOC 图分析



方法,通过对项目所涉及的“S——供方;I——输入;P——过程;O——输出;C——顾客”的分析,识别项目所涉及的范围,并加以确定。

2. 项目约束

项目的约束和假定必须清楚。在确定约束和假定条件时,可通过回答下列问题加以考虑:

- 项目计划能够接受的时间多长?什么时候是最后期限?
- 团队有多少工作时间能够投入到项目计划中?
- 团队能使用的资源有多少?
- 团队能否吸引企业内部的其他人来参加项目?企业外部的呢?
- 设备、设施的使用是否满足项目实施的要求?

完成项目工作的时间取决于分配给它的资源,作为项目团队的领导,黑带应该解决资源分配的问题,且应该得到倡导者和过程管理者或部门负责人的支持和配合。

3. 团队的任务

任务与目标是不同的概念。目标是一种想要的结果状态,任务是为达到最终目的所进行的行动。决定必须进行什么任务或行动以达到一个目标,要把解决问题的过程与制定的目标分开。

黑带可与团队的成员一起通过回答下列问题来确定团队的任务。

- 问题陈述是否明确指出了项目欲解决的问题?
- 问题陈述描述了哪些缺陷和现象?
- 问题陈述暗示了原因,提出改进方法或给予责任人员责任了吗?
- 用什么措施来保证团队的职责与项目范围相适应?
- 团队成员来自于项目涉及的相关部门吗?(若没有,应建议团队组织作怎样的改变?)

- 项目是可控的吗?
- 项目目标是否符合 SMART 原则?
- 如何测量(用通用术语描述团队将如何测量它)?

4.3.4 项目计划

为了使六西格玛项目有效地实施,在团队界定项目时,就应该注意项目计划的制定。依照六西格玛改进的 DMAIC 五个阶段,可以相应地确定计划项目完成时间,采用甘特图等方法对计划加以细化,便于团队活动。制定计划时,应让团队成员共同参与。由团队共同参与并听取成员的反馈意见将有助于增进团队的向心力。

4.3.5 项目立项表样例

项目立项表也称为团队宪章,是界定阶段的主要输出项,主要包括六部分的内容:选题背景、改进机会、项目目标、项目范围、项目计划和团队构成。表 4—7 展示的是某项目立项表的样例。

表 4—7

某六西格玛项目立项表样例

DMAIC 项目立项表			
项目名称: 缩短 SM300 系列出口产品交付周期, 提高市场竞争力			
项目领导: 李向阳		团队成员	
选题背景 进入新一代数码仪产品国际市场是企业“十一五”计划期间的重要战略增长点。2005 年, 作为首批进入国际市场的 SM300 系列产品, 遇到了较强的市场竞争。竞争压力主要来自于交付周期。主要竞争对手的交付周期达到了 10 天, 而本企业 SM300 的交付周期平均为 18 天。由此带来的直接市场损失为 2 600 万美元/年。		李向阳: 项目负责人 (黑带) 于一周: 生产计划部 副部长 (绿带) 孙代红: 供配部 成品采购组组长 (绿带) 张成雨: 装配车间 主任 (绿带) 王德西: 成品总库 副主任 (绿带) 李 红: 工程部 副总工程师 (绿带) 章法强: 质量保证部 SM300 经理 (绿带) 强 为: 财务部 业务经理	
问题/机会陈述 统计数据表明 2005 年 6—12 月, SM300 数码仪的平均交付周期为 18 天且波动较大, 最快的交付批其交付周期可以达到 8 天, 但最迟的交付批达到了 32 天。如果以交付周期不超过 12 天为标准, 则大约有 82% 的交付批次超过了这个标准。		目标陈述 本项目的预期目标是: 以交付周期不超过 12 天为标准, 将 SM300 系列产品的交付周期超过 12 天的比例 (缺陷率) 从 82% 降到 25%。	
项目范围、约束和假定 本项目范围是从生产计划、采购/配套、装配调试到包装发运的业务流程。		相关部门 生产计划部、供配部、装配车间、成品总库、工程部、质量保证部	
预期计划	目标日期	实际日期	评审
开始日期			
界定 (define)	3 月 28 日		
测量 (measure)	4 月 15 日		
分析 (analyze)	5 月 15 日		
改进 (improve)	6 月 15 日		
控制 (control)	7 月 15 日		
完成日期	8 月 30 日		
项目主管: 李向阳		倡导者: 王广昌	
业务主管: 洪 莹		日期: 2003-4-1	
日 期: 2003/3/28		日 期: 2003-3-30	

本章小结

界定阶段的主要工作是识别顾客需求, 明确要解决的问题, 明确过程输出变量的测量和标准, 确定项目涉及的内部流程, 从而确立项目责任以及项目目标和项目实施关键步骤或里程碑。在此基础上, 组织六西格玛项目团队, 发挥六西格玛组织体系的作用, 使六西格玛项目实现其预期目标。

界定阶段的成果应该包括: 清晰完备的问题陈述, 符合 SMART 原则的改进目标, 适当的项目范围, 职责明确的项目团队和填写完备的项目立项表。

界定阶段的主要输出是项目立项表。通过项目立项表来全面回答以下问题:

- 为什么企业需要关注本项目提出的问题? 它与企业战略目标和顾客满意的关系是



什么？

- 如果这个问题不改变，将会产生什么后果？给企业带来的损失是什么？
- 项目欲解决问题的陈述是否明确？
- 问题陈述描述了哪些缺陷和现象？
- 项目目标陈述符合 SMART 原则吗？
- 项目团队的职责与项目范围是否相适应？
- 团队成员来自于项目涉及的相关部门吗？
- 项目是可控的吗？
- 如何测量（用通用术语描述团队将如何测量它）？

在界定阶段，常见的相关工具有：顾客需求分析、卡诺模型、SIPOC 分析、相关方分析、排列图、甘特图、项目计划、项目立项表和沟通计划等。

测量阶段是 DMAIC 过程的第二个阶段，它既是界定阶段的后续活动，也是连接分析阶段的桥梁。测量是项目工作的关键环节，是以事实和数据驱动管理的具体体现。

从测量阶段起就要开始收集数据，并着手对数据进行分析。通过测量阶段的数据收集和评估工作，可以获得对问题和改进机会的定量化认识，并在此基础上获得项目实施方面的信息。

正如前面所指出的，任何过程的输入与输出关系均可以表达为：

$$Y=f(X)$$

式中，Y 为过程的输出结果；X 为影响 Y 的过程输入。正是那些关键的输入变量 X 决定了输出变量 Y。六西格玛项目的实施过程，就是不断地揭示两者之间关系的过程。

测量阶段的工作重点是在界定阶段工作的基础上，进一步明确 Y 的测量，并通过收集 X 和 Y 的测量数据，定量化地描述 Y。特别是通过过程分析，认识 Y 的波动规律，揭示过程改进的机会，识别实现项目目标的可能途径和改进方向。

在这个阶段，需要大量的工具和方法帮助项目团队完成上述工作。比如，各种过程分析工具，文档和图形分析工具以及过程能力分析等方法。为了保证测量数据的准确可靠，项目团队还需要对测量系统的能力作出评估。

本章主要包括五个方面的内容：过程分析与文档、概率与数理统计基础、数据的收集和整理、测量系统分析和过程能力分析。

5.1 过程分析与文档

过程分析是六西格玛管理活动中的重要内容，也是测量阶段的重要工作。测量阶段正是从过程分析开始的。

在前面的章节中，我们已经讨论了什么是过程以及关注过程的重要性。正像前面介绍的那样，过程是“一组将输入转化为输出的相互关联或相互作用的活动”。如果展开来说的话，过程有这样一些特征：过程将输入转化为输出；过程是一组相互关联的活动，它将一个概念、一项要求或一个订单转化为一种可交付的产品或服务；过程是一个或一组活动，它接受输入并增加价值，然后向内部或外部顾客提供输出。这些特征构成了过程的基本特点。

不管是生产活动，还是服务或管理活动，都可以视为由一系列过程构成。正是基于这一点，DMAIC 方法可以广泛地应用于制造过程、服务过程以及各种经营和管理过程。从过程的观点来看，产品与服务的质量、生产效率、经营管理活动的绩效等都是过程的输



出，它们的好坏取决于过程方法。而改善过程输出的最有效途径，就是改善产生这些输出的过程方法。

在测量阶段，需要对准备改善的过程进行分析，需要考察这些过程的构成和步骤，目的是：

- 使项目团队对准备改善的过程达成统一的认识；
- 对产生问题或缺陷的区域进行定位；
- 识别不增值步骤，以便加以改进；
- 将过程步骤的现状记录并形成文档，以便与改善后的状况进行对比。

下面列出的是用于理解和分析过程的一些基本工具及常用文档。

5.1.1 流程图

流程图（flow chart 或 flow diagram）是展现过程步骤和决策点顺序的图形文档，是将一个过程的步骤用图的形式表示出来的一种图示技术。图 5—1 给出的就是一个流程图的例子。

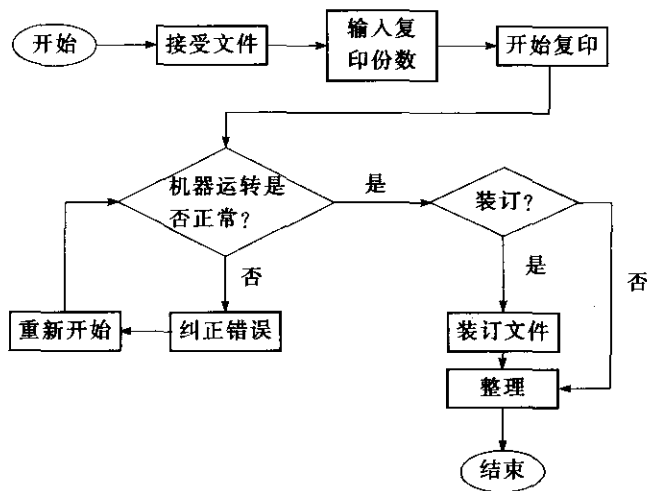


图 5—1 流程图示例

在流程图中，用矩形显示流程中的活动或事件；用菱形显示决策点，表明需做出的判断，以及根据判断结果采取的不同路径；用椭圆显示流程的起点和终点；用带有箭头的线段连接流程的各个步骤，表明流程中各个步骤发生的顺序。绘制流程图时应使用规范符号，常用符号见表 5—1。

表 5—1 绘制流程图的常用符号

符号名称	符号形状	符号的意义
操作，阶段或步骤	矩形 □	显示各步骤采取的活动或描述各阶段的活动或事件
决策点	菱形 ◇	显示可能的分支途径
端点	椭圆 ○	指示阅读者进入或退出流程
流程线	带箭头线条 →	显示流程方向
连接点	小圆 ○	显示流程线的转换

一般来说,绘制流程图应遵从以下步骤:

- (1) 判定过程的开始点和结束点;
- (2) 观察从开始到结束的整个过程;
- (3) 识别过程中的步骤(包括主要活动、判断和决策点等),以及各个步骤或活动的流向和相互关系;
- (4) 将上述步骤绘制出来,形成流程图的草图;
- (5) 团队成员就这张草图进行充分沟通并达成一致;
- (6) 形成正式文档。

除提供基本信息外,根据项目的不同目的,还可以在流程图中附加周期或责任者等信息,我们称之为跨职能流程图。图5—2是将过程中所涉及的责任者加以标注的跨职能流程图,它可以对涉及不同职能部门的比较复杂的过程进行说明。

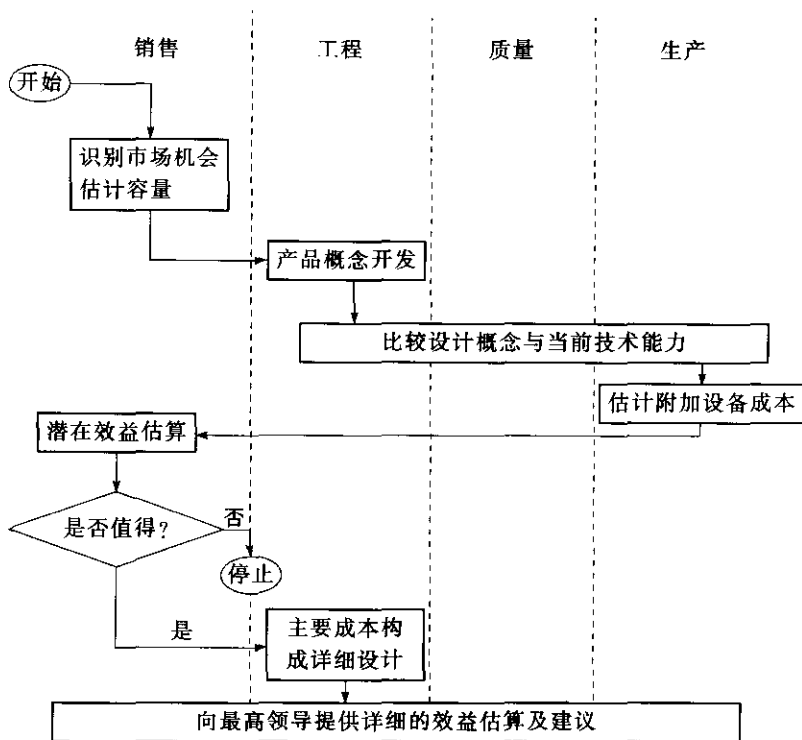


图 5—2 跨职能流程图示例

对于项目团队而言,可以用讨论并绘制过程的流程图作为开端,它可以帮助团队成员对过程涉及的方方面面有一个共同的理解。每一个团队成员基于自己的工作角色,对过程都有自己的认识,而对于整个过程往往了解不够全面。因此,应充分利用绘制流程图的过程,在团队成员中进行充分的沟通,以便对所改进的过程达成共同的理解,使所有团队成员能更有效地参与到项目工作中来。

项目团队在绘制完流程图之后,应对流程图进行充分的分析,重点关注以下几个



方面:

- (1) 产生过程输出缺陷或问题的重点关注区域在哪些环节或步骤上。
- (2) 流程中的非增值步骤或环节在何处, 如返工/返修环节或步骤等。
- (3) 流程中是否有“瓶颈”存在。这里“瓶颈”是指某点的工作负荷大大超过此处的处理能力, 从而延缓了整个工作进度。“瓶颈”现象也指准时把产品或服务按足够的数量传递给顾客的薄弱环节。
- (4) 流程中是否有缺失、冗余或者错误的步骤等。

在绘制流程图的过程中, 应避免闭门造车、凭想当然来绘制过程的流程图, 事实上, 对过程分析来说, 最重要的是要根据过程的实际情况绘制流程图。这样的流程图才有助于识别问题领域和关键影响因素产生的环节, 有助于对问题的测量、分析和改进。在许多情况下, 无法找到问题产生的原因, 是因为团队未能对过程的实际流程取得一致的认识, 使后续工作失去了重要的基础。

流程图在 DMAIC 项目工作的各个阶段都将用到。它是过程分析的重要工具, 也是项目工作中产生的重要文档。

流程图的方式很多, 如宏观流程图、详细流程图等。

5.1.2 因果图与因果矩阵

因果图 (cause and effect diagram) 又称为石川馨图 (Ishikawa chart) 或鱼刺图 (fishbone chart) 等。它是揭示过程输出缺陷或问题与其潜在原因之间关系的图表, 也是表达和分析其因果关系的重要工具和文档。制造类问题因果图见图 5—3, 服务与管理类因果图见图 5—4。

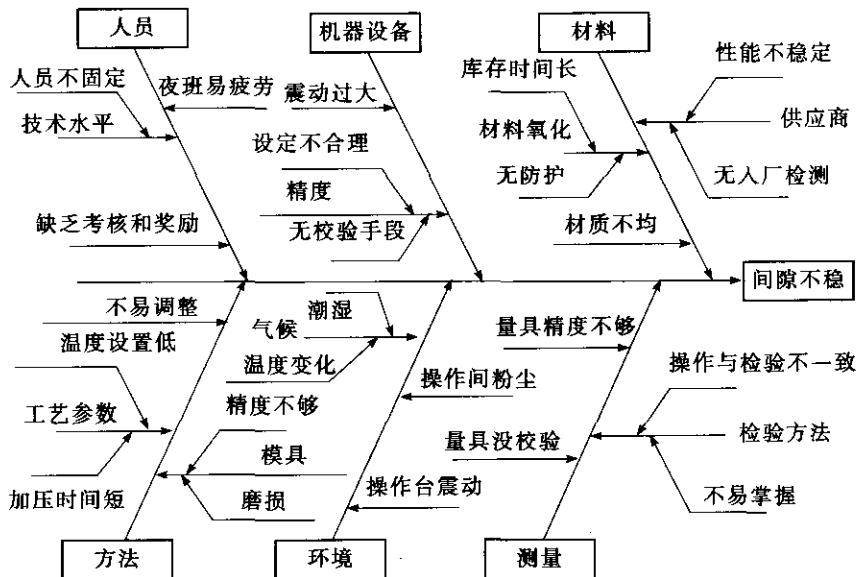


图 5—3 制造类因果图示例

在因果图中, 问题 (结果) 填在图右侧的方框中, 也就是通常所说的“鱼头”位置。造成结果的可能原因填在各个主分类线周围, 它们构成了“鱼刺”。因果图的主要作用是

分析和表达存在的问题与原因之间的因果关系，它是项目团队在揭示问题的根本原因的过程中，需要完成的重要工作。绘制因果图的过程需要项目团队成员的积极参与，因此在绘制因果图的过程中，通常需要采用头脑风暴法，将团队成员的认识和想法全部收集上来，并用因果图将它们展示出来。

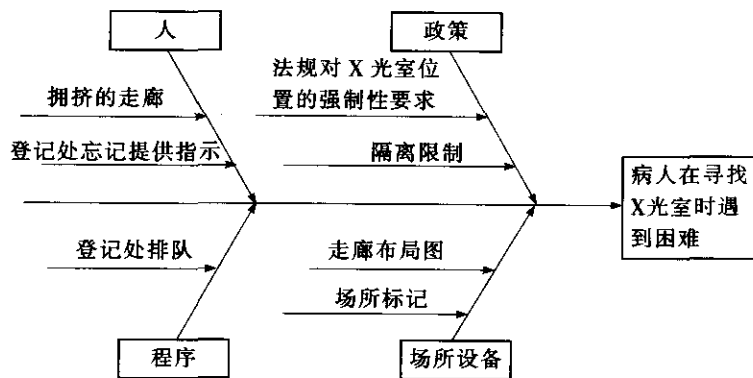


图 5—4 服务与管理类因果图示例

一般来说，绘制因果图应遵从以下步骤：

- (1) 简明扼要地阐明要解决的问题，并将其填入因果图右侧的方框——“鱼头”中。
- (2) 确定其可能原因的主要类别，画出因果图的各个“主枝”。一般来说，对制造类的问题，可以采用的典型分类方法是：人员、机器设备、材料、方法、测量和环境，即 5M1E 分类法。对服务类或管理类的问题，经常使用的分类方法是：人、政策、程序和场所设施，即 4P 分类法。
- (3) 采用头脑风暴法，将产生问题的所有可能原因，按其不同的分类填入各个主枝中。根据需要，可在各个主枝上继续分枝。位于各个分枝上的，是其下一层次的原因。
- (4) 检查和整理因果图，调整阐述含糊的内容，合并重复的内容等。
- (5) 团队成员对可能的原因进行充分讨论，确定少数可能性较高的原因，作为下一步调查和收集数据的重点。

当预期解决的问题比较复杂，有多种缺陷形式且它们的影响因素相互关联，无法将它们分开来考察和解决时，采用因果矩阵（cause and effect matrix）将是一种有效的分析工具。因果矩阵可以帮助项目团队选择重点关注的过程输入或影响因素，以便有针对性地收集数据进行分析。因果矩阵还可以与其他过程分析工具结合，用于对多个输出特性的过程进行分析和改进（见图 5—5）。

在绘制因果矩阵时，一般应遵从以下步骤：

- (1) 在矩阵图的上方填入过程输出缺陷的形式或关键过程输出变量。
- (2) 确定每一输出特性或缺陷形式的重要度，并给定其权重（1~10，10 代表的重要度最高）。
- (3) 在矩阵图的左侧，列出输入变量或所有可能的影响因素。
- (4) 评价每一输入变量或影响因素对各个输出变量或缺陷的相关关系，矩阵图中的单元格用于表明该行对应的输入变量对该列对应的输出变量的相关程度。一般将这种相关程



度分为四类，并分别赋予 0, 1, 3, 9 或 0, 1, 3, 5 的分值，表明其不同的相关程度。

(5) 评价过程输入变量或影响因素的重要程度。将每一单元的相关程度的分值乘以该列对应的输出变量的权重数，然后将每一行的乘积加起来，这个结果代表了该输入变量或影响因素的权重。

(6) 考察每个输入变量或影响因素的权重数，权重较高的将是项目重点关注的对象。

重要度	5	8	10	5	3	输入重要度排序
输出 输入	绝缘强度低	耐压击穿	功率大	转速低	启动性能差	
绝缘漆浓度低	◎	○				69
预供时间短		○	○			54
定子性能差			◎	◎	◎	162
转子缺陷	○		◎	◎		150
风叶不配套			○		○	39
风叶角度与电机不匹配			◎		△	93
轴承不合格		△	△	○	△	36
精加工精度差	◎	△		○		58

注：图中◎为 9 分，○为 3 分，△为 1 分，空格为 0 分。

图 5—5 因果矩阵示例

因果图或因果矩阵为后续的测量以及收集数据工作奠定了基础，它是过程分析的重要工具，也是项目工作中产生的重要文档之一。

5.1.3 其他过程分析工具与文档

除了上述用于过程分析的基本工具与文档外，还有一些常用于过程分析的文档。其中，在测量阶段使用较多的有过程失效模式与影响分析（process failure mode and effect analysis, PFMEA）。

PFMEA 是关于产品或过程的一种风险分析工具和文档。最初是用来对设计方案的风险进行评估的。在 DMAIC 的测量阶段，常使用 PFMEA 对过程的输入和影响因素进行评估，寻找那些对过程输出影响较大的输入或影响因素，作为测量和分析的重点（见表 5—2）。关于 PFMEA 分析方法，请参见本书第 10 章的内容。简单地说，项目团队在做 PFMEA 分析时需完成下列工作：

- (1) 识别过程的功能和要求；
- (2) 使用头脑风暴法分析潜在失效模式及其后果；
- (3) 评定每一后果的严重度等级（S）；
- (4) 使用头脑风暴法分析失效的潜在原因；
- (5) 评定每一原因的频数等级（O）；
- (6) 识别当前的过程控制方法；
- (7) 评定不可探测度等级（D）（它是指在现行控制下无法发现问题的可能性）。

表 5—2

过程失效模式与影响分析 PFMEA

PFMEA 编号:10554

页数:1 至 1

项目:降低备件发货成本

过程职能:发货过程

准备人员:王华

组员:李光伟 赵朋 沈冰仪 夏群

日期:4/3/2003

操作序号、过程 功能/要求	潜在失 效模式	潜在失 效后果	严重 度 S	等级	潜在失 效原因	频度 O	现行过 程控制	不可探 测度 D	风险等 级数 RPN	改进 措施	责任人 /完成日期	措施结果				
												采取的措施	S	O	D	RPN
向顾客发送备件	备件发送 错误	顾客不满 意,重新 发货增加 成本,因 错发赔偿 顾客损失	8		订单上填 写的备件 信息不详 细	2	由订货部 门核对信 息	4	64							
					顾客地址 不准确详 细	6	无	9	432							
					发货票据 填写错误	4	无	9	288							
					备件编码 信息不准	2	无	9	144							



(8) 针对每一行计算一个风险等级数 (RPN) (它等于严重度、频度和不可探测度三者的乘积);

(9) 具有高风险等级数的项, 将是收集数据进行分析和改进的重点。

在 DMAIC 项目中, 还有一些有助于过程分析的文档。从测量阶段开始, 应善于应用这些文档。这些文档有:

- 程序文件。通过了 ISO 9000 标准认证的组织都建立了这样的文件体系, 用以规范生产、服务和管理活动。

- 检查单。用来显示过程的每一步或检查的每一项内容等。例如, 飞行员使用的检查单、建筑作业需要的材料检查单、用来确保每一间客房准备就绪的酒店清洁检查单等。

- 照片。显示产品或服务在过程的关键阶段的外观性状。例如, 电信装置中带有色标的导线在塑封前的终端位置、装配操作的工作站布局、零件组装的照片等。

- 图表。用来显示多个部件之间的关系。例如, 一个“部件展开”或“安装”图表显示了各个部件是如何安装成一个更大组件的。

- 录像。用来显示过程的某一阶段的操作过程。

- 计算机数据库。操作者在工作站用来“提取”工作所需的图纸、程序/作业指导书和规范。

- 交互的计算机程序。在工作现场的计算机终端, 允许操作者直接输入数据或从过程中接收数据等。

5.2 概率与数理统计基础

5.2.1 概率论基础知识

许多决策都涉及对不确定性问题的处理, 例如, 过程能否满足要求? 用于过程控制的标准度量是否准确? 是否具有可重复性? 接收到一批产品的质量是否满足合同的质量要求? 统计方法可以回答这些问题。本节先讨论一些基本的概率概念, 它们是常用统计方法的基础。

1. 随机事件

随机事件是概率论中的一个基本概念。为了说明随机事件, 先看什么是试验, 什么是事件。在同一组条件下, 对某事物或现象所进行的观察或实验叫试验, 把观察或实验的结果叫事件。

抛掷一枚质地均匀的骰子就是一次试验, 骰子落地, 可能出现 1 点, 2 点, ..., 6 点, 或为奇数点或为偶数点, 点数大于 3 等等, 这些就是一个个事件。这些事件在一次试验中可能出现也可能不出现。与此不同, 还有两种事件, 如“点数大于 6”, 在每次抛掷骰子时一定不会出现, 而“点数小于 7”, 在每次抛掷骰子时一定会出现。于是, 明确下面三个概念:

- 随机事件。在同一组条件下, 每次试验可能出现也可能不出现的事件。
- 必然事件。在同一组条件下, 每次试验一定出现的事件。

- 不可能事件。在同一组条件下，每次试验一定不出现的事件。

必然事件和不可能事件可以作为两种极端的随机事件来看待。随机事件用大写字母 A, B, C 等表示，必然事件用 Ω 表示，不可能事件用 Φ 表示。如果一个事件不能分解成两个或更多的事件，则称为基本事件或简单事件。如“点数为1”，“点数为2”， \dots ，“点数为6”是抛掷骰子这个试验的6个基本事件，又可记为 $\{1\}, \{2\}, \dots, \{6\}$ ，其中：1, 2, \dots , 6 称为样本点。试验中一切可能的样本点构成样本空间，记为 Ω 。抛掷骰子试验中， $\Omega = \{1, 2, 3, 4, 5, 6\}$ 。

2. 事件间的关系与运算

设 A, B 表示两个随机事件，下面给出随机事件间的关系与运算。

(1) 事件的包含与相等。如果事件 A 出现，一定会导致事件 B 出现，则称事件 A 包含于事件 B 。用记号 $A \subset B$ 表示（见图 5—6）。例如，“电话交换台一分钟内接到的电话呼叫恰为指定次数”是一个随机事件，可能是 0 次，1 次，2 次， \dots 。“接到呼叫 2 次”和“接到呼叫不超过 5 次”这两个事件有明显的关系，前者发生，后者一定发生，前者包含于后者，或者说后者包含了前者。

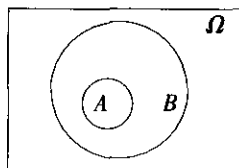


图 5—6 $A \subset B$

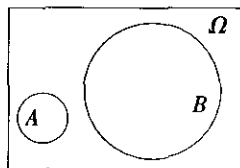


图 5—7 A 与 B 不相容

如果 $A \subset B$ 同时 $B \subset A$ ，则称事件 A 与事件 B 相等，用记号 $A = B$ 表示。

(2) 事件不相容。事件 A 与事件 B 没有共同的样本点，则称事件 A 与事件 B 不相容或称事件 A 与事件 B 是互斥事件（见图 5—7）。例如，抛掷硬币时，正面朝上和背面朝上的结果不可能同时出现，即“正面朝上”和“背面朝上”是互斥事件。检查一批产品，“恰有 1 件不合格品”和“恰有 2 件不合格品”是互斥事件。

(3) 事件的并。由事件 A 与事件 B 中所有样本点组成的新事件称为事件 A 与事件 B 的并，记为 $A \cup B$ （见图 5—8），是逻辑上“或”的关系。例如，电话交换台， $A =$ “接到呼叫 1 次”， $B =$ “接到呼叫 2 次”，则 $A \cup B =$ “接到呼叫 1 次或 2 次”。

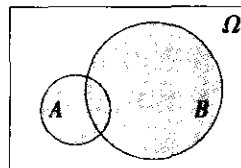


图 5—8 A 与 B 的并

(4) 事件的交。由事件 A 与事件 B 中公共的样本点组成的新事件称为事件 A 与事件 B 的交，记为 $A \cap B$ 或 AB （见图 5—9），是逻辑上“与”的关系。例如，电话交换台， $A =$ “接到呼叫多于 3 次”， $B =$ “接到呼叫少于 5 次”，则 $A \cap B =$ “接到呼叫 4 次”。

(5) 事件的差。由在事件 A 中而不在事件 B 中的样本点组成的新事件称为事件 A 对事件 B 的差，记为 $A - B$ （见图 5—10）。例如，电话交换台， $A =$ “接到呼叫多于 3 次”， $B =$ “接到呼叫少于 5 次”，则 $A - B =$ “接到呼叫多于 4 次”， $B - A =$ “接到呼叫少于 4 次”（见图 5—11）。

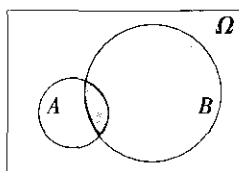
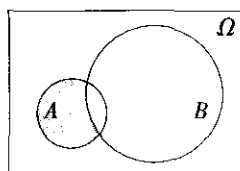
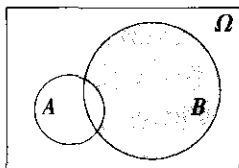
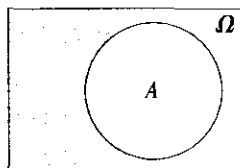


图 5—9 A 与 B 的交

图 5—10 $A-B$ 图 5—11 $B-A$ 图 5—12 A 的对立事件 \bar{A}

(6) 对立事件。由在 Ω 中而不在事件 A 中的样本点组成的事件称为事件 A 的对立事件，记为 \bar{A} (见图 5—12)。例如，检查布匹时，事件“至少有 1 个疵点”的对立事件是“没有疵点”。

3. 概率

事件 A 的概率是描述事件 A 在试验中出现的可能性大小的一种度量，表示事件 A 出现可能性的数值 $P(A)$ 称为事件 A 的概率。抛掷一枚质地均匀的硬币，“正面朝上”与“背面朝上”的可能性各为 $1/2$ 。抛掷骰子时，“恰为 6 点”的可能性为 $1/6$ 。概率是一个介于 $0 \sim 1$ 之间的数。概率愈大，事件发生的可能性愈大；概率愈小，事件发生的可能性愈小。不可能事件的概率为 0，必然事件的概率为 1。

4. 概率的性质和运算法则

概率具有以下基本性质：

性质 1：对任一随机事件 A ，有 $0 \leq P(A) \leq 1$ 。

不可能事件的概率为 0，必然事件的概率为 1， $P(\Phi) = 0$ ， $P(\Omega) = 1$ 。

性质 2：事件 A 的对立事件 \bar{A} ，有 $P(A) + P(\bar{A}) = 1$ 。

性质 3：若 $A \supset B$ ，则 $P(A-B) = P(A) - P(B)$ 。

性质 4：加法法则。

事件 A 与事件 B 的并的概率为 $P(A \cup B) = P(A) + P(B) - P(AB)$ 。

当事件 A 与事件 B 不相容时， $P(AB) = 0$ ，则 $P(A \cup B) = P(A) + P(B)$ 。

此性质可推广到多个两两互不相容的随机事件 A_1, A_2, \dots, A_n ，则

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

性质 5：乘法法则。

每一个随机试验都是在一定条件下进行的，这里要讨论的条件概率，是当试验结果的部分信息已知时，如当某一事件 B 已经发生时，事件 A 发生的概率，称这种概率为事件 B 发生条件下事件 A 发生的条件概率，记为 $P(A | B)$ 。计算公式为：

$$P(A|B) = \frac{P(AB)}{P(B)}, P(B) > 0$$

对任意两个事件 A 与 B , 有

$$P(AB) = P(B|A)P(A) = P(A|B)P(B)$$

这就是概率的乘法法则。其中, 第一个等式成立要求 $P(A) > 0$, 第二个等式成立要求 $P(B) > 0$ 。

当两个事件中不论哪一个事件发生与否都不影响另一个事件发生的概率时, 则称这两个事件相互独立。即 $P(A|B) = P(A)$, $P(B|A) = P(B)$ 。其乘法原则简化为:

$$P(AB) = P(A)P(B)$$

例 5-1

某地有甲、乙两家著名的报纸, 该地的成年人有 20% 读甲报, 有 16% 读乙报, 8% 两种报纸都读, 问成年人中至少读一种报纸的概率是多少?

解: 设事件 A 为“读甲报”, 事件 B 为“读乙报”, 事件“至少读一种报纸”可以用事件 $A \cup B$ 表示, 其概率为 $P(A \cup B) = P(A) + P(B) - P(AB)$ 。

由题意知, $P(A) = 0.2$, $P(B) = 0.16$, $P(AB) = 0.08$ 。则

$$P(A \cup B) = 0.2 + 0.16 - 0.08 = 0.28$$

故有 28% 的成年人至少读一种报纸。

例 5-2

100 件产品中, 有 80 件正品, 20 件次品; 而 80 件正品中一等品 50 件, 二等品 30 件。现从 100 件产品中任取一件, 问:

(1) 取到正品的概率是多少?

(2) 取到一等品的概率是多少?

(3) 如果取到的是正品, 那么, 恰好是一等品的概率是多少?

解: 设事件 A 为“取到正品”, 事件 B 为“取到一等品”。

(1) 取到正品的概率: $P(A) = \frac{80}{100} = 0.8$

(2) 取到一等品的概率: $P(B) = \frac{50}{100} = 0.5$

(3) 由于 $A \supset B$, 故 $P(AB) = P(BA) = P(B)$ 。利用条件概率的计算公式知, 如果取到的是正品, 恰为一等品的概率是:

$$P(B|A) = \frac{P(BA)}{P(A)} = \frac{P(B)}{P(A)} = \frac{5}{8}$$

同样, 也可以根据题意直接得到, 如果取到的是正品, 恰为一等品的概率是:

$$P(B|A) = \frac{50}{80} = \frac{5}{8}$$



10 件产品中, 有 8 件正品, 2 件次品。问:

(1) 某检验员从 10 件产品中依次抽取了 2 件, 2 件都是次品的概率是多少?

(2) 甲、乙两位实习检验员各自独立从这 10 件产品中抽取 1 件产品进行检测, 且抽取后皆放回。两位实习检验员都抽到次品的概率是多少?

解: 设事件 A_1 为“某检验员第 1 次抽到的是次品”, 事件 A_2 为“某检验员第 2 次抽到的是次品”, 事件 B_1 为“甲实习检验员抽到的是次品”, 事件 B_2 为“乙实习检验员抽到的是次品”。

(1) $P(A_1) = \frac{2}{10} = 0.2$, $P(A_2 | A_1) = \frac{1}{9}$, 故都是次品的概率为:

$$P(A_1 A_2) = P(A_1) P(A_2 | A_1) = 0.2 \times \frac{1}{9} \approx 0.022$$

(2) $P(B_1) = \frac{2}{10} = 0.2$, 因为, 甲、乙两位检验员抽取产品彼此独立, 故

$$P(B_2 | B_1) = P(B_2) = \frac{2}{10} = 0.2$$

都是次品的概率为:

$$P(B_1 B_2) = P(B_1) P(B_2) = 0.2 \times 0.2 = 0.04$$

5.2.2 随机变量及分布

1. 随机变量

先看一个熟悉的例子: 抛掷一枚质地均匀的硬币, 有两种可能的结果。若用 ω_0 表示正面朝上, 用 ω_1 表示背面朝上, 则所有可能出现的结果组成样本空间 $\Omega = \{\omega_0, \omega_1\}$ 。引入变量 X , 将可能出现的结果数量化, 令

$$X = X(\omega) = \begin{cases} 0, & \omega = \omega_0 \\ 1, & \omega = \omega_1 \end{cases}$$

X 是定义在样本空间 Ω 上的函数。对于每个可能的结果, 都有一个不同的实数 $X(\omega)$ 与之对应。而每次抛硬币之前, 结果是无法预料的。换言之, X 取 0 还是取 1 是随机的, 故称 X 为随机变量。并且, 随机变量 X 以确定的概率取不同的结果:

$$P(X = 0) = 0.5$$

$$P(X = 1) = 0.5$$

随机变量通常用大写字母 X, Y, Z 等表示, 随机变量的取值常用小写字母 x, y, z 等表示。通常, 我们只研究两类随机变量: 连续型随机变量 (continuous random variable) 和离散型随机变量 (discrete random variable)。例如:

(1) 设 X 为某铸件上的缺陷数, 则 X 是离散型随机变量, 可取 0, 1, 2, \dots 可列个可能值, 使用随机变量 X 的取值来表示事件。如 “ $X = 0$ ” 表示事件 “铸件无缺陷”,

“ $X=2$ ”表示事件“铸件上有两个缺陷”，“ $X>2$ ”表示事件“铸件缺陷多于两个”。这些事件可能发生，也可能不发生。 X 取 $0, 1, 2, \dots$ 值是随机的，但 X 以确定的概率取不同的值。

类似地，玻璃单位面积上的气泡数、布匹上的疵点数、车床一天内发生的故障数等都是取值为 $\{0, 1, 2, 3, \dots\}$ 的离散型随机变量。

(2) 设 X 为某电视机的寿命(单位：小时)，则 X 是在 $[0, \infty)$ 上取值的连续型随机变量。“ $X=0$ ”表示事件“电视机开箱即发生故障”，“ $X<10\,000$ ”表示事件“电视机寿命小于 $10\,000$ 小时”，“ $X>40\,000$ ”表示事件“电视机寿命超过 $40\,000$ 小时”。

类似地，测厚仪测量所得的涂料厚度、某超市一天内售出的鸡蛋重量等都是在 $[0, \infty)$ 上取值的连续随机型变量。

(3) 设 X 为产品检测的结果，可能是合格品，也可能是不合格品。若用“ $X=0$ ”表示合格品，“ $X=1$ ”表示不合格品，则 X 是只能取 0 或 1 的随机变量。

2. 离散型随机变量和分布

如果随机变量 X 只取有限个或可列个可能的值： x_1, x_2, x_3, \dots ，且 $P(X=x_i)=p_i$ ($i=1, 2, 3, \dots$)，则称 X 为离散型随机变量，其中： $P(X=x_i)=p_i$ 为离散型随机变量 X 的概率函数。可用分布列表示如下：

$X=x_i$	x_1	x_2	\dots	x_n	\dots
$P(X=x_i)=p_i$	p_1	p_2	\dots	p_n	\dots

显然，有 $p_i \geq 0$ ， $p_1 + p_2 + \dots + p_i + \dots = 1$ 。这就提供了离散型随机变量 X 的概率分布。

一批产品分为一、二、三级，其中一级品率为 50% ，二级品率为 35% ，三级品率为 15% 。从这批产品中随机抽取一件，观察其级别。

解：设 X 表示“随机抽取一件产品，所观察到的产品级别”，显然，有

$$P(X=1)=0.5$$

$$P(X=2)=0.35$$

$$P(X=3)=0.15$$

可用分布列表如下：

$X=x_i$	1	2	3
$P(X=x_i)=p_i$	0.5	0.35	0.15



抛掷一枚质地均匀的骰子，观察其结果。

解：设 X 表示“抛掷骰子的结果”，则 X 是取值为 1, 2, ..., 6 的离散型随机变量。它的分布为：

$X=x_i$	1	2	3	4	5	6
$P(X=x_i)=p_i$	1/6	1/6	1/6	1/6	1/6	1/6

3. 连续型随机变量和分布

我们将在本章后面部分介绍直方图，图 5—13 是一个根据一批样本数据绘制频率直方图的示例。得到的频率直方图如图 5—13 (a) 所示。当数据个数很多、组距很小、组数很多时，连接直方图中每个矩形上边中点的折线就接近一条光滑的曲线，图 5—13 (b) 给

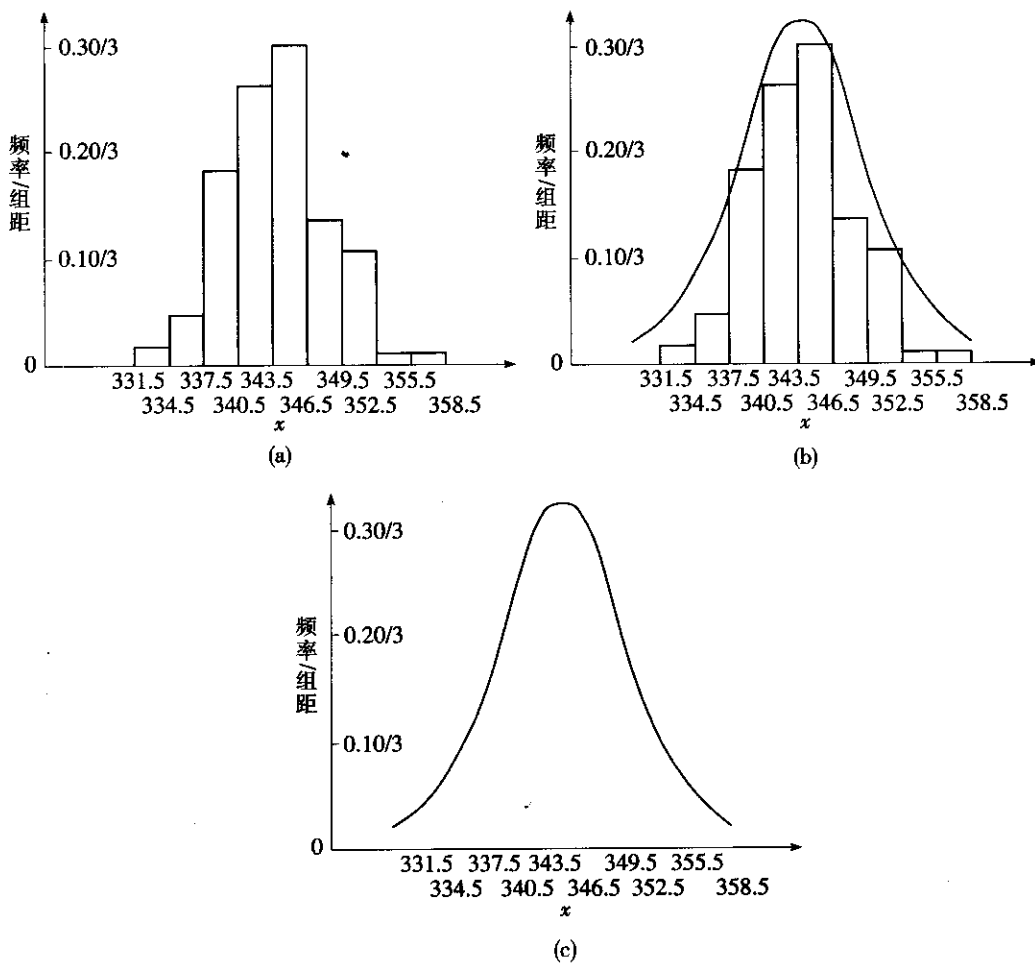


图 5—13 概率密度函数曲线的解释

出了粗略近似的描述。随着数据的不断增加, 频率趋于稳定, 频率的稳定值就是概率, 单位长度上的概率简称为概率密度, 这条曲线的函数即为概率密度函数 (probability density function, pdf), 如图 5—13 (c) 所示。

概率密度函数是描述连续型随机变量分布的最重要和最基本的工具。令 $p(x)$ 为连续型随机变量 X 的概率密度函数, 则 $p(x)$ 满足:

- $p(x) \geq 0$, 概率密度函数为非负函数。
- $\int_{-\infty}^{+\infty} p(x) dx = 1$, 概率密度函数曲线与实数轴所围成的面积为 1, 如图 5—14 所示。

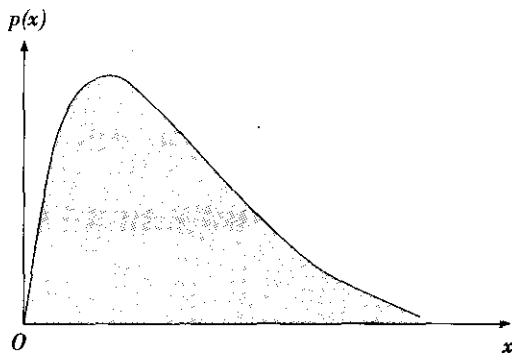


图 5—14 概率密度函数所围成的面积等于 1

• $P(a \leq X \leq b) = \int_a^b p(x) dx$, 区间 $[a, b]$ 上的概率可由概率密度函数在该区间上求积分得到, 如图 5—15 阴影部分所示。

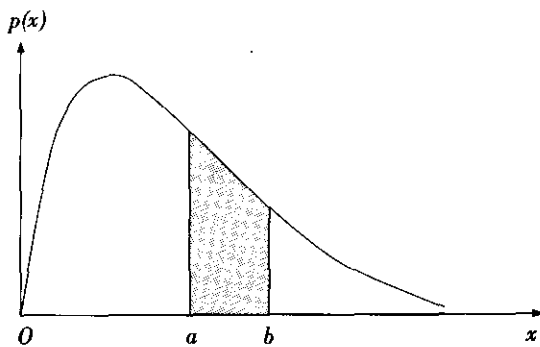


图 5—15 区间 $[a, b]$ 上的概率

其中, 随机变量 X 小于或等于实数 x 的概率 $F(x) = P(X \leq x)$, 有

$$P(X \leq x) = \int_{-\infty}^x p(t) dt$$

$F(x)$ 又称为累积分布函数 (cumulative distribution function, cdf) 或分布函数。如图 5—16 阴影部分所示。

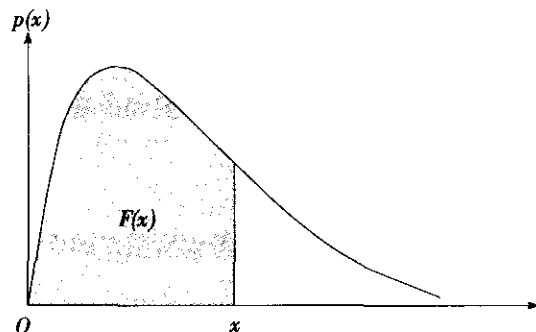


图 5—16 分布函数

由于在计算面积时, 直线的面积为 0, 即 $P(X=a) = 0$, 有

$$P(a \leq X \leq b) = F(b) - F(a)$$

指数分布 (exponential distribution) 是一种常见的连续分布, 其概率密度函数为:

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

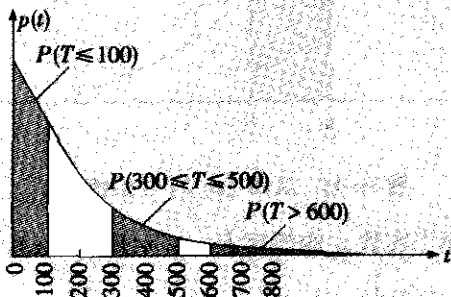
记为 $E(\lambda)$, 式中, $\lambda > 0$ 为参数。实际的工作生活中, 不少产品首次发生故障的时间, 或发生故障后需要维修的时间都服从指数分布。

图 5—17

某厂生产的推土机首次发生故障的时间 T (单位: 小时) 服从指数分布 $E(0.002)$ 。概率密度函数如图 5—17 所示。

$$p(t) = \begin{cases} 0.002e^{-0.002t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

- 问: (1) 该厂生产的推土机在 100 小时内需要维修的概率是多少?
 (2) 该厂生产的推土机在 300~500 小时内需要维修的概率是多少?
 (3) 该厂生产的推土机在 600 小时以后才需要维修的概率是多少?

图 5—17 指数分布 $E(0.002)$ 的概率密度曲线

解: (1) 该厂生产的推土机在 100 小时内需要维修的概率, 如图 5—17 左侧阴影所示, 可用积分计算:

$$P(T \leq 100) = \int_0^{100} 0.002e^{-0.002t} dt = -e^{0.002t} \Big|_0^{100} = 1 - e^{-0.2} = 0.183$$

(2) 该厂生产的推土机在 300~500 小时内需要维修的概率, 如图 5—17 中间阴影所示。

$$P(300 \leq T \leq 500) = \int_{300}^{500} 0.002e^{-0.002t} dt = 0.1809$$

(3) 该厂生产的推土机在 600 小时以后需要维修的概率, 如图 5—17 右侧阴影所示。

$$P(T > 600) = \int_{600}^{+\infty} 0.002e^{-0.002t} dt = 0.3012$$

5.2.3 数学期望、均值与方差

1. 数学期望

租赁公司规定租赁某种昂贵的建筑设备, 顾客必须至少提前三天预订。租赁费是 500 元。如果下雨, 顾客将无法使用该设备, 但是, 顾客仍要支付 500 元。如果没有下雨, 顾客使用建筑设备完成工作可获得收入 1 000 元。若降雨的概率是 20%, 租赁设备获得收益的期望 (expectation) 值是多少?

使用下面的方程式:

$$E = (1/5) \times (-500) + (4/5) \times 1\,000 = 700$$

顾客可以期望收入 700 元。注意, 实际上, 如果只租赁一天, 则获利 700 元是不可能发生的, 租赁人要么失去 500 元, 要么获得 1 000 元。期望值给出了顾客长期租赁可获利的平均水平, 概率作为“权重”被应用于不同的结果。因此, 数学期望可看作各种可能结果的加权平均。

均值 (mean)、方差 (variance) 与标准差 (standard deviation) 都是重要的期望值。均值用来表示分布的集中情况 (中心位置), 方差与标准差用来表示分布的散布情况。

2. 均值、方差与标准差

● 均值。表示分布的中心位置, 反映分布的集中情况, 用 $E(X)$ 表示。均值的计算公式为:

$$E(X) = \begin{cases} \sum_i x_i p_i, & X \text{ 为离散型随机变量} \\ \int_{-\infty}^{+\infty} x p(x) dx, & X \text{ 为连续型随机变量} \end{cases}$$

式中, x_i , p_i 和 $p(x)$ 的含义参见 5.2.2 中离散型随机变量和连续型随机变量的介绍。

● 方差。可用来表示分布的波动大小, 用 $\text{var}(X)$ 表示。方差大, 意味着分布的散布较宽、较分散; 方差小, 意味着分布的散布较窄、较集中。方差的计算公式为:



$$\text{var}(X) = \begin{cases} \sum_i [x_i - E(X)]^2 p_i, & X \text{ 为离散型随机变量} \\ \int_{-\infty}^{\infty} [x - E(X)]^2 p(x) dx, & X \text{ 为连续型随机变量} \end{cases}$$

● 标准差：为使表示分布波动情况的量与 X 具有相同的单位，常对方差求开方，记正平方根为 σ ，称为标准差：

$$\sigma = \sigma(X) = \sqrt{\text{var}(X)}$$

由于 σ 与 X 的单位相同，实际工作中，常使用标准差来表示分布的散布大小，但标准差的计算要通过计算方差来获得。

例 5-7

计算例 5-5 中分布的均值、方差和标准差。

解：设 X 表示“抛掷骰子的结果”，则 X 是以确定的概率 $1/6$ 取观测值为 1, 2, ..., 6 的离散型随机变量。它的均值、方差和标准差为：

$$\begin{aligned} E(X) &= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} \\ &= (1+2+3+4+5+6) / 6 = 3.5 \end{aligned}$$

$$\begin{aligned} \text{var}(X) &= (1-3.5)^2 \times \frac{1}{6} + (2-3.5)^2 \times \frac{1}{6} + \dots + (6-3.5)^2 \times \frac{1}{6} \\ &= (2.5^2 + 1.5^2 + 0.5^2 + 0.5^2 + 1.5^2 + 2.5^2) / 6 = 2.92 \end{aligned}$$

$$\sigma(X) = \sqrt{\text{var}(X)} = \sqrt{2.92} = 1.71$$

例 5-8

计算例 5-6 中分布的均值。

解：设 T 表示“推土机首次发生故障的时间”，服从指数分布 $E(0.002)$ 。其概率密度函数为：

$$p(t) = \begin{cases} 0.002e^{-0.002t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

它的均值为：

$$E(T) = \int_{-\infty}^{\infty} tp(t) dt = \int_0^{\infty} 0.002te^{-0.002t} dt = (0.002)^{-1} = 500$$

这表明推土机首次发生故障的平均时间是 500 小时。可以证明，对于指数分布，其均值与标准差是相等的，即 T 的标准差 $\sigma = 500$ 小时。

3. 均值与方差的性质

均值与方差具有下列性质：

● 设 X 为随机变量， a 与 b 为任意实数，则有

$$E(aX+b) = aE(X) + b$$

$$\text{var}(aX+b) = a^2 \text{var}(X)$$

- 对任意两个随机变量 X_1 和 X_2 , 有

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

这个性质可以推广到三个或更多个随机变量的场合。

- 设随机变量 X_1 和 X_2 独立 (即 X_1 的取值不影响 X_2 的取值, 相当于两个试验的独立性), 则有

$$\text{var}(X_1 \pm X_2) = \text{var}(X_1) + \text{var}(X_2)$$

这个性质也可以推广到三个或更多个相互独立随机变量的场合。

5.2.4 常用的离散分布

常见的离散分布有:

- 0—1 分布 (0-1 distribution)
- 二项分布 (binomial distribution)
- 泊松分布 (Poisson distribution)
- 超几何分布 (hypergeometric distribution)

1. 0—1 分布

抽取某产品, 检验其是否合格, 若检验合格记为 0, 不合格记为 1, 且知该产品的合格率为 95%。问检验结果如何?

解: 设 X 表示“抽检产品的结果”, 则 $P(X=0) = 0.95$, $P(X=1) = 0.05$ 。可用分布列表如下:

X	0	1
$P(X=x_i) = p_i$	0.95	0.05

当每次试验中, 只有两种可能的结果, 例如: 正面朝上与背面朝上, 产品合格与不合格, 检验通过与不通过, 目标命中与不命中, 具备某特性与不具备某特性等, 或是只关注试验的两种不同结果, 例如: 手机月话费是否高于 100 元, 布匹疵点数是否多于 5 个等等, 可以将这两种结果称为成功与失败。若成功的概率为 p , 则失败的概率为 $1-p$, 设随机变量 X 表示“试验的结果”, 则 X 服从两点分布。若把成功记为 1, 失败记为 0, 则称 X 服从 0—1 分布, 记为 $X \sim B(1, p)$ 。

X	0	1
$P(X=x_i) = p_i$	$1-p$	p

0—1 分布的均值、方差与标准差为:

$$E(X) = p, \text{var}(X) = p(1-p), \sigma(X) = \sqrt{p(1-p)}$$



2. 二项分布

例 2-12

从一批产品中随机抽取进行检测, 根据历史数据知, 产品的不合格率为 10%。假设要求产品检测人员每抽取一件产品, 检测完毕后, 要放回这批产品中 (又称为有放回抽样)。检验人员共检测了 6 件产品, 问检测到的不合格品数分别为 0, 1, 2, ..., 6 的概率?

解: 设随机变量 X 表示 “6 件检测产品中不合格品的个数”, 显然, X 可能的取值为: 0, 1, 2, ..., 6, 取这些值的概率为:

$$P(X=x) = C_6^x 0.1^x (1-0.1)^{6-x}, x=0, 1, 2, \dots, 6$$

称 X 服从二项分布, 记为 $X \sim B(6, 0.1)$ 。

没有不合格品的概率为:

$$P(X=0) = C_6^0 0.1^0 (1-0.1)^{6-0} = \frac{6!}{0! (6-0)!} 0.9^6 \approx 0.5314$$

式中, “!” 称为阶乘: $6! = 1 \times 2 \times 3 \times 4 \times 5 \times 6$, $0! = 1$, $C_6^0 = \frac{6!}{0! (6-0)!} = 1$ 。

恰有 1 件不合格品的概率为:

$$P(X=1) = C_6^1 0.1^1 (1-0.1)^{6-1} = \frac{6!}{1! (6-1)!} \times 0.1 \times 0.9^5 \approx 0.3543$$

类似地可计算有 2, 3, 4, 5, 6 件不合格品的概率, 有分布列表如下:

X	0	1	2	3	4	5	6
P	0.5314	0.3543	0.0984	0.0146	0.0012	0.0001	0.0000

这里, $P(X=6) \approx 0.000001$, 取前四位小数为有效数字, 故为 0.0000。

下面仔细分析二项分布应满足的条件: 考察由 n 次随机试验组成的随机现象。

(1) 试验次数固定为 n , 即重复进行 n 次随机试验。譬如, 连续抛掷一枚质地均匀的硬币 n 次, 检查 n 件产品的质量, 对一个目标连续射击 n 次等。

(2) n 次试验间相互独立, 即一次试验的结果不能对其他各次试验的结果产生影响。

(3) 每次试验仅有两种可能的结果: 成功与失败、好与坏等。

(4) 每次试验中成功的概率均为 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的次数”, 显然 X 是可以取 0, 1, ..., n 等 $n+1$ 个数值的离散随机变量, 而且取这些数值的概率为:

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

称 X 服从二项分布, 记为 $X \sim B(n, p)$ 。

二项分布的均值、方差与标准差分别为:

$$E(X) = np, \text{var}(X) = np(1-p), \sigma(X) = \sqrt{np(1-p)}$$

二项分布除了作为连续生产过程中不合格品数的精确分布外, 当样本容量小于批量的

10%时,二项分布还可以作为超几何分布的近似分布。

3. 泊松分布

质量控制中常遇到这样的情况:不仅要关注不合格品,而且要关注每件不合格品所包含的不合格项的情况。例如,全面检查生产的计算机总会发现一些不合格项,即使有些不合格项不重要(如仪表贴纸不整齐等)。此时所关注的不再是产品合格与否的问题,而是产品出现的不合格项。设 X 表示“产品不合格项的个数”,显然, X 的取值为可列个可能值: $0, 1, 2, \dots$, 取这些值的概率为:

$$P(X=x) = \frac{\lambda^x}{x!} e^{-\lambda}, x=0, 1, 2, \dots$$

则称 X 服从泊松分布,记为 $X \sim P(\lambda)$ 。式中, $e = 2.718\ 281\ 8$; λ 是泊松分布的重要参数,它给出了产品的平均不合格项数。

泊松分布的均值与方差相等,均为 λ 。有

$$E(X) = \lambda, \text{var}(X) = \lambda, \sigma(X) = \sqrt{\lambda}$$

泊松分布应用广泛,可用来描述不少随机变量的分布。如:

- (1) 一定时间内接错电话的次数;
- (2) 一定时间内某操作系统发生的故障数;
- (3) 一只铸件上的缺陷数;
- (4) 一平方米玻璃上的气泡数;
- (5) 一件产品擦伤留下的痕迹数;
- (6) 一页书上的错字数。

一条生产导弹的流水线,当每枚导弹制造完成时,要通过一个空气动力模型进行检查,对照设计要求把各处不合格记录下来。任何严重的不合格项都可能成为拒绝接受的理由。军方还要求对次要的不合格项进行控制,像模糊的文字、小毛刺等问题在审核时都要求明确记录下来。过去的历史数据显示,每枚导弹平均有三个次要的不合格项。问:生产的导弹没有次要不合格项的概率是多少?

解:设 X 表示“导弹次要不合格项的个数”,由题意知, $\lambda = 3$ 。则生产的导弹没有次要不合格项的概率是:

$$P(X=0) = \frac{3^0 e^{-3}}{0!} = \frac{1 \times e^{-3}}{1} \approx 0.05 = 5\%$$

换言之, $100\% - 5\% = 95\%$ 的导弹至少有一处次要不合格项。

4. 超几何分布

某批发商提供了一批产品,共 12 件,其中 3 件有轻微碰撞的痕迹。现用户取出 4 件使用。问:4 件全部完好无损的可能性有多大?

解:首先,注意这里涉及无放回抽样,即从这批产品中随机地取出一件,再从剩余产

品中选出另一件,这就是无放回抽样。而且,这批产品只有12件,抽取4件产品占了总量的33%左右。

设 X 表示“4件产品中不合格品的个数”,则 X 服从超几何分布。可以用下列公式来计算其概率:

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

式中, N 是批量; M 是该批产品中的不合格品数; n 是样本量; x 是样本中不合格品的个数; $P(X=x)$ 为在样本中恰好有 x 件不合格品的概率,记为 $X \sim H(n, N, M)$ 。

由题意知, $N=12, M=3, n=4$,则全部完好无损的概率为:

$$P(X=0) = \frac{C_{12-3}^{4-0} C_3^0}{C_{12}^4} = \frac{126 \times 1}{495} \approx 0.255$$

故用户使用的4件产品全部完好无损的可能性为25.5%。

超几何分布的均值、方差分别为:

$$E(X) = \frac{nM}{N}, \text{var}(X) = \frac{n(N-n)}{N-1} \cdot \frac{M}{N} \left(1 - \frac{M}{N}\right)$$

当样本量小于批量的10%时,二项分布可作为超几何分布的近似分布。

5.2.5 常用的连续分布

常用的连续分布有:

- 正态分布 (normal distribution)
- 均匀分布 (uniform distribution)
- 指数分布 (exponential distribution)
- 对数正态分布 (logarithmic normal distribution)
- 威布尔分布 (Weibull distribution)

1. 正态分布

质量管理中最常遇到的连续分布是正态分布,能描述很多质量特性 X 随机取值的统计规律性。正态分布的概率密度函数为:

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

如图5—18所示,概率密度函数曲线可形象地描述为“中间高、两边低、左右对称、

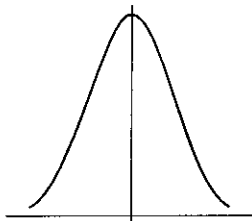


图5—18 正态分布的概率密度函数曲线

延伸到无穷”的钟形曲线。正态分布又称高斯分布，记为 $X \sim N(\mu, \sigma^2)$ 。式中， μ 和 σ^2 为正态分布的参数； μ 是正态分布的均值，给出了正态分布的中心； X 在 μ 附近取值的机会较大； σ^2 是正态分布的方差， σ ($\sigma > 0$) 是正态分布的标准差。

图 5—19 (a) 给出了相同的标准差 σ 、不同的均值 μ ($\mu < \mu'$) 时的情况，钟形曲线的中心由 μ 向右移动到 μ' ，而分布的形状不变。图 5—19 (b) 给出了相同的均值 μ 、不同的标准差 σ 时的情况，钟形曲线的中心不变。 σ 越大，分布越分散； σ 越小，分布越集中。正态分布的两个参数均值 μ 与标准差 σ 无关， μ 的变化不会影响到分布形状， σ 的变化也不会影响到分布中心，一般称 μ 为位置参数， σ 为形状参数。

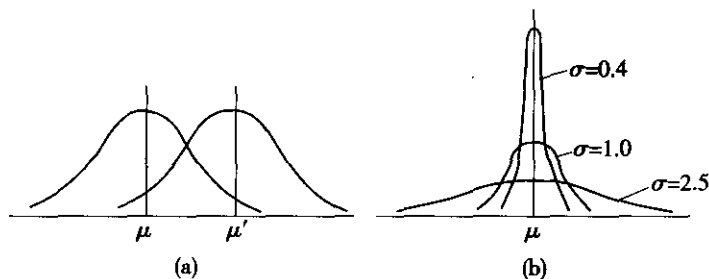


图 5—19 均值 μ 与标准差 σ 对分布的影响

$\mu=0$ 且 $\sigma=1$ 的特殊正态分布称为标准正态分布 (standard normal distribution)，记为 $N(0, 1)$ 。设 $X \sim N(\mu, \sigma^2)$ ，则 $\frac{X-\mu}{\sigma} \sim N(0, 1)$ 。式中， $\frac{X-\mu}{\sigma}$ 常被称为标准变换。经过标准变换，正态分布都可以转化为均值为 0、标准差为 1 的标准正态分布。利用标准正态分布函数表 (如附表 1 所示)，可以查表得到 $P(X \leq x)$ 的概率值，如图 5—20 所示的阴影面积。

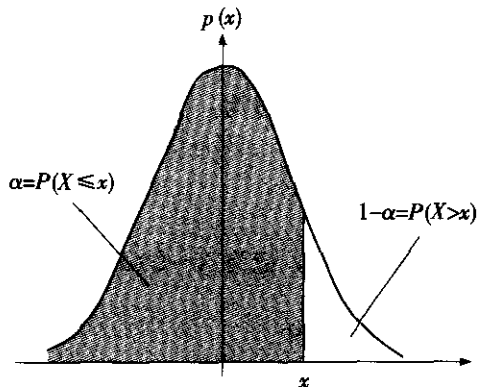
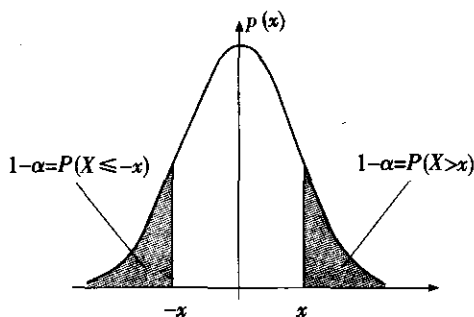


图 5—20 $P(X \leq x)$ 与 $P(X > x)$

设 $P(X \leq x) = \alpha$ ，则概率密度函数与实数轴之间的面积 1 被数值 x 分割为两部分，左侧阴影面积为 $\alpha = P(X \leq x)$ ，右侧面积为 $1 - \alpha = P(X > x)$ 。故 x 又记为 z_α ，称为 α 分位数。另外，标准正态概率密度曲线的对称特征决定了 $P(X \leq -x) = P(X \geq x)$ ，而 $P(X = x) = 0$ ，故 $P(X \leq -x) = P(X > x)$ ，进而 $P(X \leq -x) = 1 - \alpha = 1 - P(X \leq x)$ ，如图 5—21 所示。

图 5—21 $P(X \leq -x) = 1 - \alpha = 1 - P(X \leq x)$

检查微电路生产中关键线路黏合过程的阻断力。过程平均力度是 19，标准差为 4。过程力度服从正态分布。如果工程规格下限为 13，则过程力度低于规格下限的比例为多少？

解：设 X 表示过程力度， $X \sim N(19, 4^2)$ ，则 $\frac{X-19}{4} \sim N(0, 1)$ 。考虑过程力度低于规格下限的概率，首先，进行标准变换 $\frac{13-19}{4} = -1.5$ ，然后，查标准正态分布函数表，得到 6.68%，即

$$P(X \leq 13) = P\left(\frac{X-19}{4} \leq \frac{13-19}{4}\right) = P\left(\frac{X-19}{4} \leq -1.5\right) = 6.68\%$$

故 6.68% 的阻断力低于规格下限 13。质量控制通常要求均值距离规格界限至少 3 倍标准差。为了实现这个目标，需要改进过程，要么提高阻断力的均值，要么减小过程的标准差，或者同时改进这两个方面。

2. 均匀分布

设连续型随机变量 X 具有概率密度函数为：

$$p(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases}$$

则称 X 在区间 (a, b) 上服从均匀分布，记为 $X \sim U(a, b)$ ，如图 5—22 所示。

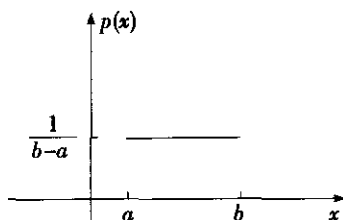


图 5—22 均匀分布的概率密度函数和分布函数

在区间 (a, b) 上服从均匀分布的随机变量 X ，意味着 X 落在区间 (a, b) 中任意等

长度的子区间内的可能性相同。换言之, X 落在子区间内的概率只依赖于子区间的长度而与子区间的位置无关。对任一长度 l 的子区间 $(c, c+l)$, $a \leq c < c+l \leq b$, 有

$$p(c < X \leq c+l) = \int_c^{c+l} p(x) dx = \frac{l}{b-a}$$

均匀分布 $U(a, b)$ 的均值、方差与标准差分别为:

$$E(X) = \frac{a+b}{2}, \text{var}(X) = \frac{(b-a)^2}{12}, \sigma(X) = \sqrt{\frac{(b-a)^2}{12}}$$

3. 指数分布

指数分布是一种常见的连续分布, 其概率密度函数为:

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

记为 $X \sim E(\lambda)$, 式中, $\lambda > 0$ 为参数。实际的工作生活中, 不少产品首次发生故障的时间或发生故障后需要维修的时间都服从指数分布。

指数分布的分布函数为:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

指数分布的均值、方差和标准差分别为:

$$E(X) = \frac{1}{\lambda}, \text{var}(X) = \frac{1}{\lambda^2}, \sigma(X) = \frac{1}{\lambda}$$

例 5-14

一家城市水资源管理公司, 平均每年发生 500 次泄漏。假定两次泄漏的间隔时间服从指数分布。问: 周末 (从周五晚上 6 点至周一早上 6 点) 工作人员没有接到报修电话的概率是多少?

解: 这里给出的是每年平均的泄漏次数, 需要转化为两次泄漏之间的平均间隔时间。每年 365 天, 每天 24 小时, 即 8 760 小时。两次泄漏之间的平均间隔时间 (MTBF) 为 $8\,760/500 = 17.52$ 小时。设 X 表示两次泄漏的间隔时间, 则 X 服从均值为 17.52 小时的指数分布, 故 $\lambda = \frac{1}{17.52} = 0.071$, 即 $X \sim E(0.071)$ 。周五晚 6 点至周一早 6 点之间共有 60 个小时, 这 60 小时内发生泄漏的概率为:

$$P(X \leq 60) = 1 - e^{-60/17.52} \approx 0.967 = 96.7\%$$

因此, 这 60 小时没有发生泄漏, 即工作人员周末空闲的概率仅为 $1 - 96.7\% = 3.3\%$ 。

4. 对数正态分布

在地质勘探中, 岩石的某种化学成分的对数服从正态分布, 故得名对数正态分布。对数正态分布出现在很多领域, 如针刺麻醉的阵痛效果、英语单词的长度、流行病的蔓延时间、某些电器的寿命、化学反应时间、绝缘材料的被击穿时间、产品维修时间等等。

设 X 是取值为正数的随机变量。若 $\ln X \sim N(\mu, \sigma^2)$, 则称 X 服从对数正态分布, 记为 $X \sim N(\mu, \sigma^2)$ 。X 的大量取值在左侧, 长尾在右侧, 并且很分散, 这样的分布称为



“右偏分布”（见图 5—23）。如机床维修中，大量机床在短时间内都可修理好，只有少量机床需要长时间维修，个别机床可能需要更长的修理时间。

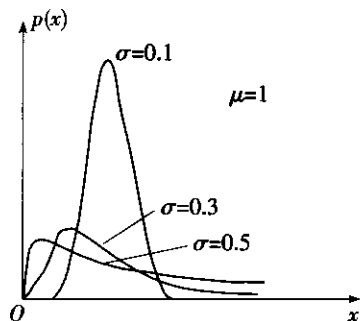


图 5—23 对数正态分布的概率密度函数曲线 ($\mu=1$, σ 不同)

对数正态分布的均值和方差为：

$$E(X) = e^{\mu + \frac{\sigma^2}{2}}, \quad \text{var}(X) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$$

5. 威布尔分布

威布尔分布是寿命试验和可靠性理论的基础，它是瑞典科学家威布尔于 1939 年首先提出的。

威布尔分布的概率密度函数为：

$$p(x) = \begin{cases} \frac{\alpha}{\beta} (x-\gamma)^{\alpha-1} e^{-\frac{(x-\gamma)^\alpha}{\beta}}, & x \geq \gamma \\ 0, & x < \gamma \end{cases}$$

式中， $\alpha > 0$ 称为形状参数； $\gamma \geq 0$ 称为位置参数； $\beta > 0$ 称为尺度参数，记为 $X \sim W(\alpha, \beta, \gamma)$ 。当 $\gamma=0$, $\alpha=1$ 时， $p(x)$ 简化为：

$$p(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

令 $\lambda = \frac{1}{\beta}$ ，则

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

这是大家熟悉的指数分布。

威布尔分布的形状参数 α 往往与失效机理相联系。例如，工作寿命分布，常将 $\alpha < 1$ 的情况称为“早期失效期”， $\alpha=1$ 为“偶然失效期”， $\alpha > 1$ 为“耗损失效期”。尺度参数 β 起到放大与缩小比例常数的作用。位置参数 γ 是一个平移参数，有时又称为最小保证寿命。产品在 γ 以前不会失效，即失效的概率为零。

5.2.6 中心极限定理

1. 随机变量独立同分布的概念

随机变量 X_1 和 X_2 独立, 是指 X_1 的取值不影响 X_2 的取值, X_2 的取值也不影响 X_1 的取值。随机变量 X_1 和 X_2 同分布, 意味着 X_1 和 X_2 具有相同的分布形状和相同的分布参数, 对离散型随机变量具有相同的概率函数, 对连续型随机变量具有相同的概率密度函数, 有着相同的分布函数, 相同的均值、方差与标准差。反之, 若随机变量 X_1 和 X_2 是同类型分布, 且分布参数全相同, 则 X_1 和 X_2 一定同分布。如 X_1 和 X_2 同为正态分布, 且有相同的均值 μ 和相同的方差 σ^2 , 则 X_1 和 X_2 具有相同的分布。如 X_1 和 X_2 同指数分布, 则 X_1 和 X_2 具有相同的指数分布参数 λ 。一般来说, 在相同条件下, 进行两次独立试验, 则这两次试验结果所对应的随机变量是独立同分布的。比如, 将一枚质地均匀的硬币抛掷两次, 设 X_1 为第 1 次抛掷硬币的结果, X_2 为第 2 次抛掷硬币的结果。显然, 第 1 次抛掷硬币的结果对第 2 次的结果没有影响, 反之亦然, 故 X_1 和 X_2 相互独立。同时, X_1 和 X_2 都只有两种试验结果: 正面朝上和背面朝上, 以 0 代表正面朝上, 1 代表背面朝上, 则

$$P(X_1=0) = P(X_2=0) = 0.5, P(X_1=1) = P(X_2=1) = 0.5$$

故 X_1 和 X_2 是独立同分布的随机变量。

随机变量独立同分布的特性可以推广到三个或更多个随机变量。

2. 独立同正态分布随机变量的重要性质

定理 1 设 X_1, X_2, \dots, X_n 是 n 个独立同正态分布的随机变量, $X_i \sim N(\mu, \sigma^2)$, $i=1, 2, \dots, n$, 则

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i$$

仍为正态分布, 其均值不变仍为 μ , 其方差缩小 n 倍, 若把 \bar{X} 的方差记为 $\sigma_{\bar{X}}^2$, 则 $\sigma_{\bar{X}}^2 = \sigma^2/n$ 。即 $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ 。

如设 X_1, X_2, \dots, X_9 是 9 个独立同分布的随机变量, 且 $X_i \sim N(10, 25)$, 则

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_9}{9} \sim N(10, \frac{25}{9})$$

这表明: \bar{X} 的均值为 10, \bar{X} 的方差为 $25/9=2.78$, \bar{X} 的标准差为 $\sigma_{\bar{X}} = \sqrt{25/9}=5/3$ 。

3. 中心极限定理

定理 2 (中心极限定理) 设 X_1, X_2, \dots, X_n 是 n 个独立同分布的随机变量, 且有有限的数学期望和方差, 分布均值为 μ , 方差为 σ^2 , 则在 n 较大时, 有

$$(1) X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i \text{ 近似服从均值为 } n\mu, \text{ 方差为 } n\sigma^2 \text{ 的正态分布 } N(n\mu, n\sigma^2);$$

$$(2) \bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i \text{ 近似服从均值为 } \mu, \text{ 方差为 } \sigma^2/n \text{ 的正态分布 } N(\mu, \frac{\sigma^2}{n}).$$

当 X_i 的分布对称时, 只要 $n \geq 5$, 那么, 近似效果就比较理想; 当 X_i 的分布非对称



时, 要求 n 值较大, 一般 $n \geq 30$ 近似效果较理想。

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 \bar{X} 的分布都可近似为正态分布。这一结论意义深远。

统计学中把均值 \bar{X} 的标准差称为均值的标准误, 记为 $\sigma_{\bar{X}}$ 或 SEM (standard error of the mean), 无论是正态还是非正态, 均值 \bar{X} 的标准误都有

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

SEM 随着 n 的增加而减少。

常常对一个零件的质量特性只观测一次, 就用该观测结果去估计过程输出的质量特性。这里建议一种简单有效的减少测量系统误差的方法。对同一个零件的质量特性作两次或更多次重复测量, 用其观测结果的平均值去估计过程输出的质量特性, 就可以减少标准差。当然, 这不是回避使用更精密量具的理由, 而是一种提高现有量具精度的简易方法, 多次测量值的平均值要比单次测量值更精确。

5.2.7 统计量与抽样分布

1. 总体与样本

研究对象的全体称为总体 (population), 构成总体的每个基本单位称为个体。例如, 在研究某厂生产的灯泡寿命时, 该厂生产的全部灯泡的寿命就是总体, 每只灯泡的寿命就是个体。研究该厂某天生产的灯泡寿命时, 总体是该天生产的全部灯泡的寿命。研究某高中男生身高时, 该校全体男生的身高是总体; 研究高三男生身高时, 总体是全部高三男生的身高。

总体依其包含的个体数分为有限总体和无限总体。例如, 研究某厂生产的灯泡寿命时, 总体是该厂生产的全部灯泡的寿命, 这包括了过去生产的灯泡的寿命、正在生产的灯泡的寿命和今后将要生产的灯泡的寿命, 这是一个无限总体。研究某天生产的灯泡寿命时, 总体是该天生产的全部灯泡的寿命, 个体数是有限的, 这是一个有限总体。

总体可以用随机变量 X 表示, 随机变量 X 的分布就是总体分布, 不同的总体可用不同的分布去描述。如果可以对总体中每一个个体进行观察, 当然可以了解该总体, 但是, 许多情况下这是不可能的或是没有必要的。因此, 常常从总体中抽取若干个体, 对这些个体进行观察, 利用观察结果对总体进行推断。

从总体中抽取一部分个体进行观察, 被抽到的个体组成了总体的一个样本。若总体中的每一个个体都有相同的机会被抽到, 那么, 这样得到的样本称为简单随机样本, 简称样本 (sample)。本书讨论的样本都是简单随机样本。

样本中所含的个体又称为样品, 样品的个数称为样本容量或样本量, 常用 n 表示。容量为 n 的样本记为 X_1, X_2, \dots, X_n 。有理由认为 X_1, X_2, \dots, X_n 是独立且与总体具有相同分布的随机变量。样本的观测值用 x_1, x_2, \dots, x_n 表示。

2. 常见的统计量

样本来自总体, 样本中包含了丰富的总体信息。但是这些信息是分散的, 为了把分散的信息集中起来以反映总体的特征, 就需要对样本进行加工。一种有效的办法就是构造样

本的函数,不同的函数反映总体的不同特征。不含未知参数的样本函数称为统计量 (statistic), 统计量的分布称为抽样分布。

(1) 描述中心位置的统计量。设 X_1, X_2, \dots, X_n 是来自总体的一个样本, 则

- 样本均值 (sample mean): $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
- 样本中位数 (sample median): $\tilde{X} = \begin{cases} X_{(\frac{n+1}{2})}, & n \text{ 为奇数} \\ \frac{1}{2} [X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}], & n \text{ 为偶数} \end{cases}$

式中, $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ 为 X_1, X_2, \dots, X_n 按照从小到大排列的有序样本 (又称为次序统计量, order statistic)。

(2) 描述波动情况的统计量。

- 样本极差 (sample range): $R = X_{(n)} - X_{(1)}$

式中, $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ 为 X_1, X_2, \dots, X_n 按照从小到大排列的有序样本。

- 样本方差 (sample variance): $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
- 样本标准差 (sample standard deviation): $s = \sqrt{s^2}$
- 变异系数 (coefficient of variation): $CV = \frac{s}{\bar{X}}$

上述描述波动情况的统计量中, 样本极差 R 计算最简便, 但对信息的利用较少; 样本方差 s^2 和样本标准差 s 能充分利用样本所提供的信息, 应用广泛; 样本标准差 s 可保持计量单位的一致性; 变异系数 CV 常用于不同数据的离散程度的比较。

3. 抽样分布

抽样分布是区间估计、假设检验的理论基础。

(1) 样本均值的分布。设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, 则

1) 方差 σ^2 已知时, 有 $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$;

2) 方差 σ^2 未知时, 有 $\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$, 即服从自由度为 $n-1$ 的 t 分布。图 5-24 为 t 分布。

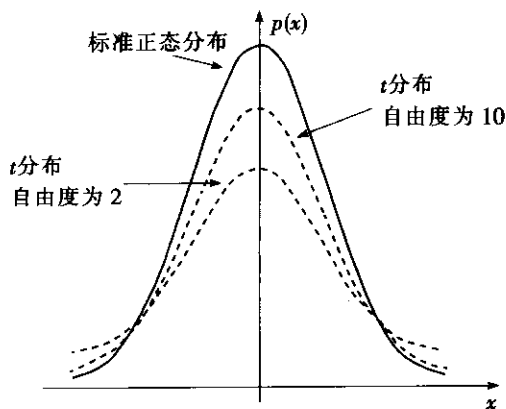


图 5-24 t 分布



t 分布的形状同正态分布很相像, 具有“中间高、两边低、左右对称、延伸到无穷”的特点。 t 分布受自由度的影响, 自由度越大, t 分布越接近于标准正态分布。当自由度大于 30, 二者差别已不大。实际上, 正态分布可以看成是自由度无穷大的 t 分布。

(2) 正态样本方差 s^2 的分布—— χ^2 分布 (χ^2 distribution)。设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, 则

$$\frac{(n-1) S^2}{\sigma^2} = \sum_{i=1}^n (X_i - \bar{X})^2 / \sigma^2 \sim \chi^2(n-1)$$

χ^2 分布的形状如图 5—25 所示。可见, χ^2 分布同样受自由度的影响, 自由度越大, 图形越向右移。

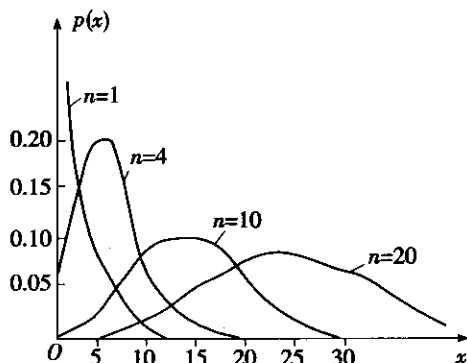


图 5—25 χ^2 分布

(3) 两个独立的正态样本方差之比的分布—— F 分布 (F distribution)。设有两个独立的正态总体 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$, 它们的方差相等。 X_1, X_2, \dots, X_n 是来自 $N(\mu_1, \sigma^2)$ 的一个样本, Y_1, Y_2, \dots, Y_m 是来自 $N(\mu_2, \sigma^2)$ 的一个样本, 而且这两个样本相互独立。则它们的样本方差之比为:

$$\frac{s_1^2}{s_2^2} \sim F(n-1, m-1)$$

式中, $n-1$ 称为分子自由度; $m-1$ 称为分母自由度。 F 分布的形状如图 5—26 所示。

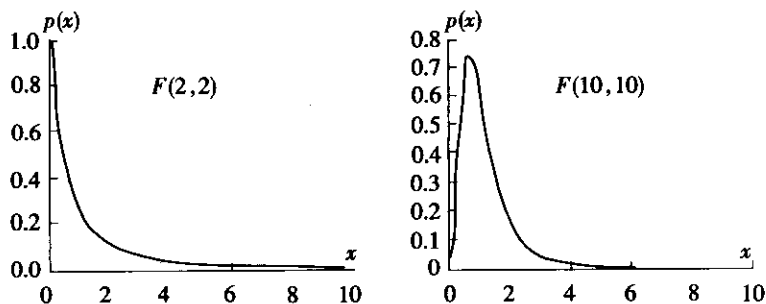


图 5—26 F 分布

5.3 数据的收集和整理

5.3.1 数据类型与测量尺度

1. 数据的类型

在六西格玛项目实施的过程中,需要不断与数据打交道,需要不断依据数据作出决策,需要处理形形色色的数据。但是从统计的角度来说,这些测量数据可以分为两种基本类型,即连续型数据和非连续或离散型数据。

用连续坐标进行测量并得出的数据是连续型数据,也称计量值数据。或者说,用测量仪器或量具测量出的可以连续取值的数据类型,是连续型数据,如长度、温度和湿度等。连续型数据的特点是,测量数据可以比较敏感地反映过程的变化,包含的信息丰富。在对过程进行统计分析时,可以用较少的样本量获得分析结论。但一般来说,连续型数据对测量手段的要求较高,或测量成本较高。

非连续型数据或离散数据,也称计数数据。如合格/不合格、通过/失败、是/否、接受/不接受、出席/缺席、好/坏以及缺陷个数等。非连续型数据在反映过程变化方面不如连续数据那样敏感。在对过程进行统计分析时,往往需要较大的样本量或较长的测量周期,才能得出分析结论。但一般说来,非连续数据对测量手段和测量精度的要求不高,测量成本也比较低。

在测量时,区分连续型数据还是离散型数据是十分重要的。项目团队应根据项目的具体情况,确定采用适当的数据类型。一般来说,在收集测量数据时,应尽可能采用连续型数据,以便提供尽可能多的可用于过程改进的信息。

2. 测量尺度

测量是按照某种规则,用数据来描述观察到的现象,即对事物作出量化描述。测量是对非量化学物的量化过程。

数据包含多少信息取决于测量所使用的尺度。测量的尺度决定了研究这些数据时应使用什么类型的统计分析方法。选定了相应的测量尺度,便确定了所产生的数据类型,也就确定了在项目实施过程中可使用的统计分析方法。

有四类测量尺度:定类、定序、定距和定比。

(1) **定类(名义)测量尺度** (nominal measurement scale)。定类尺度的数据是一些数字形式的名义值。它可以指示所属的类别,例如,0=白色,1=非白色;又如,白色、黑色、其余色。可以是两类,也可以是多类。定类尺度是测量形式中最简单的一种。定类尺度中类别划分是有要求的,那就是集合中的所有项都应属于一类且仅属于一类。以定类尺度收集的数据被称为属性数据。对于定类尺度,能识别的关系只有“=”(表示物体具有某属性)或者“≠”。能允许的变换(或换算)为任意的单值0—1函数。用属性数据所作的统计工作只有“计数”一项,即计算出属于某类的个数为多少。

(2) **定序测量尺度** (ordinal measurement scale)。定序变量是对可能的取值进行排序。例如,以“好”、“更好”和“极好”来划分顾客对某种服务的偏好;将降雨量按“旱”、

“偏少”、“中等”、“偏多”、“涝”分为5类；按某项业绩将各单位排名次等。定序变量能识别的关系为“=”（相等）、“≠”（不等）、“>”（多于）和“<”（少于）。能允许的变换（或换算）为任意的单调函数。用定序数据所作的统计工作有“计数”和“排序”两项，但没有距离的概念，因而不能对定序数据（包括名次）进行算术平均。

在质量管理中，有时为了简化而把定序数据转变成定类数据，然后用二项分布模型或泊松分布模型进行分析。例如，如果用差—好—很好的次序把零件分类，然后绘制出属于“差”这一类零件的不合格率 p 图。但这样做将损失某些信息，应尽量选用更细致的分析方法。

(3) **定距测量尺度** (interval measurement scale)。通常我们所进行的测量得到的数据都是连续性数据。但要细致分析，仍然要分为两类。第一种是定距尺度数据。例如温度（摄氏或华氏），记录事件发生的日期及时间等。这种数据中，“0”是没有意义的，因而没有倍数的概念。摄氏度的0度是以水结冰为标志，它只有相对意义，我们不能说“60℃是20℃的3倍”。定距变量能识别的关系除“=”、“≠”、“>”和“<”外，还能识别差距的大小，如“60℃与20℃间的差距比40℃与10℃的差距大”，“60℃与30℃间的差距与40℃与10℃的差距是相等的”等。用定距数据所作的统计分析可以使用算术平均，因而可以使用各种统计量，但没有“比值”、“比率”的概念。

对定距数据，能允许的变换（换算）只能为线性函数 $y=ax+b$ ($a>0$)。例如，摄氏度与华氏度间可以换算： $^{\circ}\text{F}=32+9/5(^{\circ}\text{C})$ ，其中 $a=9/5$ ， $b=32$ 。换算前后得到的统计结论一定是相同的，它不会受变换的影响。

(4) **定比测量尺度** (ratio measurement scale)。连续型数据的第二种是定比尺度数据，这比定距尺度数据更高了一个层次。例如，绝对温度、长度、质量、电流强度、压力、断裂强度、亮度、能量等。这种数据中，“0”是有意义的。定比尺度是指对同一个事物的两个不同测量结果之间的比值是有意义的，可以讨论它们的比值。例如，说“分子运动能量在绝对温度为10°K时是2°K时的5倍”，“电流强度为6安培是电流强度为2安培时的3倍”是有意义的。定比尺度数据不但能识别差距的大小而且可以识别和比较比值的大小。用定比测量尺度数据所作的统计工作不但可以进行算术平均、使用各种统计量，而且可以用除法求出倍数。对定比数据，能允许的变换（换算）只能为正比函数 $y=ax$ ($a>0$)。例如，长度可以用mm也可以用英寸 (inch) 为单位，它们之间可以换算：1英寸=25.4毫米，其中 $a=25.4$ 。换算前后得到的统计结论一定是相同的，它不会受变换的影响。只要使用适当的转换方法，不管采用何种定比尺度，基于数据的统计分析结果都是一样的。

5.3.2 收集数据的方法

有许多收集数据的方法，最典型的是使用检查表。检查表是收集过程数据时最常用的工具，常常用于DMAIC改进方法的测量阶段。

检查表有各种形式。针对不同的测量对象，需要使用不同的检查表。要有针对性地设计检查表，以便能够按照项目团队的意图，将重要的数据收集上来。在检查表上要注明被测变量，同时还应注明由谁来收集数据以及采用的测量间隔。

表5—3是一个最终装配过程的缺陷检查表。从该例中也可以看出油漆是最常发生的

缺陷，这张检查表也为减少缺陷提供了改进的基础。

表 5—3 某最终装配过程的缺陷检查表（示例）

	套管	高压开关	漏油	油量	油漆	低压开关	RP架	罐	活栓开关	活栓开关架	帽	机壳	高压仪表盘	压力试验	电镀RMU架	BOM	图纸	销售	零件材料
1				/															
2																/			
3																			
4					/							/							
5																			
5																		/	
7					/														
8	/																		
9									/										
∴																			
28																			
29																			
30																			
合计	4	0	0	1	9	0	0	1	1	0	0	2	0	0	0	1	3	2	0

表 5—4 所示的是检查表的另一个示例。该检查表既记录了抽样数量和检测到的不合格品的数量，又记录了各种缺陷发生的频次。

表 5—4 某成品抽样检验及外观不合格品项目检查表（示例）

批次	产品号	成品量 (箱)	抽样数 (支)	不合格 品数 (支)	批不合格 率 (%)	外观不合格项目								
						切口	贴口	空松	短烟	过紧	钢印	油点	软腰	表面
1	烤烟型	10	500	3	0.6	1					1			1
2	烤烟型	10	500	8	1.6			2	1					
3	烤烟型	10	500	4	0.8		1					1		
4	烤烟型	10	500	3	0.6		2			1				
5	烤烟型	10	500	5	1.0	1		1						1
⋮	⋮	⋮	⋮	⋮	⋮									
250	烤烟型	10	500	6	1.2			1		1				
合计		2 500	125 000	990	0.8	80	297	458	35	28	10	15	12	55

在应用检查表时，可将这些检查表存入计算机，这样数据就能直接保存到适当的数据库中，节省大量的时间和资源。

在评价过程改进潜力时，时间周期和过程波动的历史数据常常是有用的。历史数据能够很好地表明改进潜力，这也就是在 DMAIC 改进方法中重视测量阶段的一个主要原因。

5.3.3 抽样方法

为了确保数据的准确性和代表性，能从样本中正确推断总体，必须采用适当的抽样方法，常用的是简单随机抽样、分层抽样等方法。

1. 简单随机抽样

随机抽样也就是简单随机抽样，它是按随机原则直接从总体中抽取样本，构造统计量，并对总体相应的指标做出统计推断。



简单随机抽样必须满足两个基本条件:

- 等可能性。是指总体中的每个个体有同等机会被抽到。若总体有 N 个个体, 每个个体被抽到的概率为 $\frac{1}{N}$ 。

- 独立性。是指每次抽样是相互独立的。

常用的随机抽样方法有:

(1) 抽签法。将总体每个个体编上序号 (1, 2, ..., N), 并把号码写在结构均匀的签上, 掺和拌匀后从中抽取, 若有 N 个个体, 就要有 N 个签。

(2) 滚球法。这也就是人们熟悉的摇奖方法。

(3) 计算机模拟法。在中大型计算机中, 一般都带有随机数字发生器。

(4) 随机数表法。这种方法需要利用随机数表, 该表上数字的出现及其排列是随机形成的。0, 1, 2, ..., 9 这 10 个数字的概率各约为 $\frac{1}{10}$, 而且由表上组成的各种多位数 (如二位数, 三位数, ...), 也大体有相同的机会出现。

2. 分层抽样

分层抽样, 又称为类型抽样或分类抽样。这种抽样形式是先将总体按一定的标识加以分层, 例如按不同的设备、材料、供应商、班次、年龄组等分层。然后在各层中按简单随机抽样抽取若干个样本, 再由各层的样本组成一个总体样本。若设总体含有 N 个样本, 将总体划分为 k 层, 每层总体样本为 N_i ($i=1, 2, \dots, k$), 则有 $\sum_{i=1}^k N_i = N$, 从各层总体 N_i 中随机抽取 n_i 个样本, 则样本 $n = \sum_{i=1}^k n_i$ 。

根据不同要求, 对于样本容量 n , 可以有三种不同的分配方法。

(1) 比例分配法。样本数按各层总体数的多少比例分配, 即 $\frac{n_1}{N_1} = \frac{n_2}{N_2} = \dots = \frac{n_k}{N_k} = \frac{n}{N}$, 则各层的样本数为:

$$n_i = n \frac{N_i}{N}, i=1, 2, \dots, k$$

(2) 适度分配法。比例分配法显然没有考虑波动程度的差异, 适度分配法考虑到波动程度较大的层应该多取样, 而波动程度小的层应少取样, 这样可以减少抽样误差, 记 σ_i ($i=1, 2, \dots, k$) 为第 i 层的标准差。适度分配法要求各层样本数 n_i 要与各层总体与各层标准差乘积的比例相等, 即

$$\frac{n_1}{N_1 \sigma_1} = \frac{n_2}{N_2 \sigma_2} = \dots = \frac{n_k}{N_k \sigma_k} = \frac{n}{\sum_{i=1}^k N_i \sigma_i}$$

则各层的样本数为:

$$n_i = n \frac{N_i \sigma_i}{\sum_{i=1}^k N_i \sigma_i}, i=1, 2, \dots, k$$

(3) 经济分配法。经济分配法是指对于抽样费用较高的层, 相对取样少一些, 而费用较低的层则可以多取样, 记 C_1, C_2, \dots, C_k 为各层的抽样费用。由于样本数与费用的平

方根成反比例关系，应该使下列比例保持相等，达到在一定的抽取费用情况下抽样误差最小，或在一定的抽样误差情况下抽样费用最少，即

$$\frac{n_1}{N_1\sigma_1/\sqrt{c_1}} = \frac{n_2}{N_2\sigma_2/\sqrt{c_2}} = \cdots = \frac{n_k}{N_k\sigma_k/\sqrt{c_k}} = \frac{n}{\sum_{i=1}^k N_i\sigma_i/\sqrt{c_i}}$$

则各层的样本数应为：

$$n_i = n \cdot \frac{N_i\sigma_i/\sqrt{c_i}}{\sum_{i=1}^k N_i\sigma_i/\sqrt{c_i}}, i = 1, 2, \cdots, k$$

5.3.4 描述性统计方法

当项目团队从过程中抽样并获得样本数据后，需要对数据分布情况有基本了解。比如，分布的位置或者中心趋势、散布程度以及它的形状等。描述性统计方法就是用来描述分布的这些特性。表 5—5 给出了常见的描述性统计方法。

表 5—5 常见的描述性统计

样本统计量	说明	公式/记号
分布的位置或者中心趋势的度量		
样本均值	取自分布的样本的重心或质心	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ，其中： x_i 是观测值； n 是样本含量
中位数	有序样本中，排在中间位置的数据。数据集中的一半数据高于中位数，一半低于中位数	$\tilde{x} = \begin{cases} x_{(\frac{n+1}{2})}, & n \text{ 为奇数} \\ \frac{1}{2} [x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}], & n \text{ 为偶数} \end{cases}$ $x_{(1)}, x_{(2)}, \cdots, x_{(i)}, \cdots, x_{(n)}$ 为有序样本
众数	出现最频繁的数值。如果将数据分组，众数是频数最大的组	Mode
散布程度的度量		
极差（全距）	样本极值之间的距离	$R = \text{最大值} - \text{最小值}$
样本方差	对围绕均值波动的度量	$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}$
样本标准差	对围绕均值波动的度量	$s = \sqrt{s^2}$
形状的量度		
偏度	对称性分布的 β_3 为零，若 $\beta_3 > 0$ ，表示右偏态；若 $\beta_3 < 0$ ，表示左偏态	$\beta_3 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n (x_i - \bar{x})^3 / s^3$
峰度	峰度是分布平坦程度的度量。正态分布的峰度为 $\beta_4 = 0$	$\beta_4 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{s^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$

为了解某地区成年人心脏健康状况，项目小组对该地区的成年人进行了随机抽样，并且测量了评价心脏健康水平的若干项指标。表 5—6 是小组测量的脉搏（次数/分钟）的数据。从这组数据中可以获得以下信息：

表 5—6 小组测量的脉搏数据表

序号	脉搏	性别	序号	脉搏	性别	序号	脉搏	性别
1	64	男	16	82	女	31	82	男
2	96	女	17	84	男	32	61	女
3	78	女	18	68	男	33	64	女
4	58	男	19	62	男	34	74	男
5	62	男	20	76	男	35	68	男
6	66	男	21	100	女	36	94	女
7	96	女	22	68	女	37	60	女
8	62	女	23	88	女	38	64	男
9	78	女	24	70	男	39	90	男
10	82	女	25	62	女	40	92	男
11	64	男	26	74	男	41	60	男
12	74	男	27	66	男	42	84	女
13	78	女	28	70	男	43	72	女
14	68	女	29	80	女	44	66	女
15	72	女	30	54	男	45	80	女

使用 MINITAB 软件可以得到全体人员、男性和女性的三张计算结果示意图（见图 5—27、图 5—28 和图 5—29）。

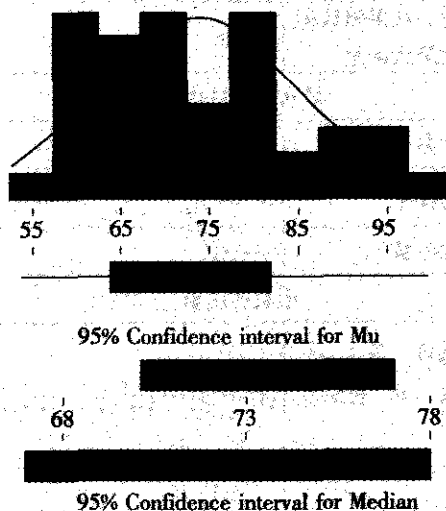


图 5—27 描述性统计

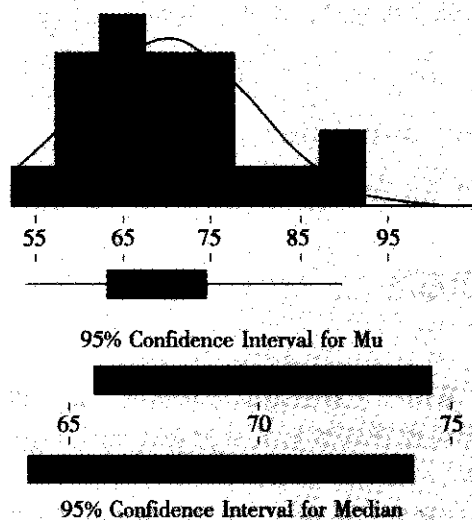


图 5—28 描述性统计 (男)

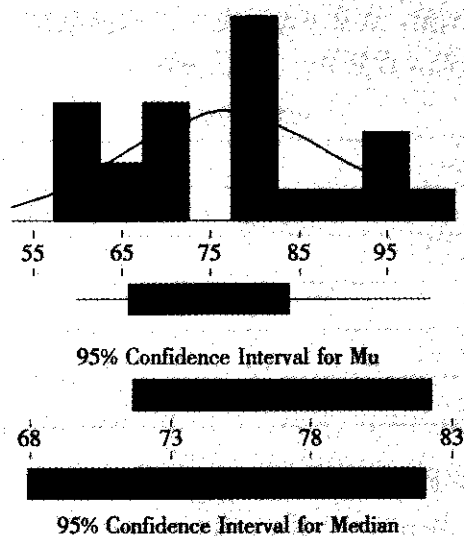


图 5—29 描述性统计 (女)

(1) 分布的位置或中心趋势。

样本均值 $\bar{x}=73.62$ (次/分), 其中: 男性 $\bar{x}=70.09$ (次/分), 女性 $\bar{x}=77.00$ (次/分)。

样本中位数 $\tilde{x}=72$ (次/分)。

(2) 分布的分散程度。

样本极差 $R=100-54=46$ (次/分), 其中: 男性 $R=92-54=38$ (次/分), 女性 $R=100-60=40$ (次/分)。

样本方差 $s^2=134.05$, 其中: 男性 $s^2=99.20$, 女性 $s^2=149.08$ 。



样本标准差 $s=11.58$ (次/分), 其中: 男性 $s=9.96$ (次/分), 女性 $s=12.21$ (次/分)。

(3) 分布的形状。

偏度 $\beta_3=0.57$, 其中: 男性 $\beta_3=0.76$, 女性 $\beta_3=0.32$ 。

峰度 $\beta_4=-0.53$, 其中: 男性 $\beta_4=-0.12$, 女性 $\beta_4=-0.90$ 。

5.3.5 数据的图示方法

1. 直方图

直方图 (histogram) 常用于了解数据的分布情况, 是一组数据的图形表示。这种展示数据的方法使我们较容易地看到数据的分散程度和中心趋势, 并与要求的分布进行比较。

直方图的构造一般由计算机自动进行, 详细步骤及要求如下:

(1) 从 n 个样本中找出最大值 $a_{(n)} = \max_{1 \leq i \leq n} x_i$ 和最小值 $a_{(1)} = \min_{1 \leq i \leq n} x_i$, 并计算极差。

(2) 对样本进行分组, 决定组数 k 和组距 d 。一般将样本分为 7~15 组为宜。具体组数可根据样本量 n 的大小而定, 通常可参看表 5—7 选择组数。组距用极差 R 与组数 k 来确定, $d=R/k$ 。通常还要把 d 近似为一个方便的数值。

表 5—7 组数 k 的参照表

样本量 (n)	组数 (k)
40~90	5~8
100~200	8~10
201~500	9~11
501~1 000	10~13
1 000 以上	12~15

(3) 调整确定各组的区间端点 $a_1, a_1+d=a_2, a_1+2d=a_3, \dots$, 形成如下区间:

$$[a_1, a_2], (a_2, a_3], \dots, (a_{n-1}, a_n]$$

(4) 统计样本落在每个区间中的频数 n_i 及频率 $f_i=n_i/n$ 。

(5) 画出频率直方图。以样本 x 为横坐标, 以频率/组距为纵坐标, 在每个小区间上竖起长条矩形。它们相连形成一张频率直方图。纵轴的单位也可以用频率来表示, 这里之所以用频率除以组距来表示, 是为了保证这些矩形面积之和为 1。

图 5—30

图 5—30 是根据表 5—8 构建的一个生产过程的频率直方图。特别注意的是该直方图中每个条形图的高度与该区间内落入的点数成正比。

表 5—8 分组的频数频率分布表

组号	区间	组中值	频数统计	频数	频率
1	(331.5, 334.5)	333	—	1	0.01
2	(334.5, 337.5)	336	正	4	0.04
3	(337.5, 340.5)	339	正正正	17	0.17
4	(340.5, 343.5)	342	正正正正正	27	0.27
5	(343.5, 346.5)	345	正正正正正正	30	0.30
5	(346.5, 349.5)	348	正正正	12	0.12
7	(349.5, 352.5)	351	正	7	0.07
8	(352.5, 355.5)	354	—	1	0.01
9	(355.5, 358.5)	357	—	1	0.01
合计				100	1.00

从这张频数频率表上可以看出,两头的区间频数少,而中间的区间频数高。从频率的这一列上可读出:大约有 74% 的数据集中在 (337.5, 346.5) 之间。

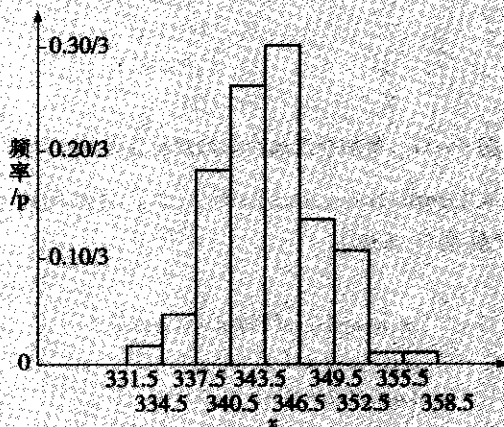


图 5—30 样本的频率直方图

使用直方图应注意的问题:

- (1) 如果直方图上画有规格限,那么直方图可用来比较过程与要求。要达到这一目标,应确认直方图具有合适的比例。
- (2) 直方图最好能与其他图,如趋势图或控制图联合使用。这是因为直方图的数据不是按时间顺序给出的,无法反映过程随时间变化的情况。
- (3) 根据数据来源的不同,应对数据分层,分别绘制直方图。

2. 茎叶图

茎叶图 (stem-and-leaf plots) 是直方图的变种,适用于较小的数据集 ($n < 200$)。和直方图相比,茎叶图的一个主要优点是全部或者部分地保留了原始数据的信息,从而可以使我们掌握有关总体的更多信息。

在画茎叶图时,要把一个数据分成两个部分:高位部分称为“茎”,低位部分称为“叶”。下面通过一个例子来说明茎叶图的作法。

图 5-31

对某型号的 20 辆汽车记录了每加仑汽油各自行驶的里程数如下：

29.8 27.5 28.3 28.7 27.9 29.9 30.1 28.0 28.7 27.9
28.5 29.5 27.2 25.9 28.4 27.9 28.0 30.0 29.5 29.1

由于数据基本在 25~31 之间，因此对数据作如下划分，以第一个数据为例：

数据	分开	茎与叶
29.8	29 8	29 与 8

在画茎叶图时，先将茎依从小到大的次序写在一条竖线的左边，然后将每个数据的叶写在竖线的右边，并将它们按从小到大的次序排列，再给出叶的单位，这便构成了一张茎叶图（见图 5-31）。

茎叶图 ($n=20$, 叶的单位=0.10)

频数	茎	叶
1	25	9
5	27	25999
7	28	0034577
5	29	15589
2	30	01

图 5-31 每加仑汽油汽车行驶里程数的茎叶图

用 MINTAB 软件从 Graph-Stem and Leaf 入口，制定间距为 1，则可得到图 5-32（注：带括号的组表示中位数所在组）。

```

Stem-and-Leaf Display, miles
Stem-and-leaf of miles N=20
Leaf Unit=0.10
1  25  9
1  26
6  27  25999
(7) 28  0034577
7  29  15589
2  30  01
  
```

图 5-32 茎叶图

我们若将上述茎叶图按逆时针方向旋转 90° ，则可将之看成组距为 1 的直方图。同样从图上可以看出中间一组数据较多，左右基本对称，具有正态分布的特征。除此以外，由于图上还标有具体数据，所以从中还可以了解数据细节。例如该组数据的最小值是 25.9，最大值是 30.1，中位数位于右边数第三组，在第 10 名与第 11 名之间，因而中位数为 $(28.4 + 28.5) / 2 = 28.45$ 。

茎与叶的分割应该视具体数据而定。若数据最小单位是 1，将数据分割时，可将百位、十位数看成是茎，个位数为叶；也可将百位数作为茎，十位数作为叶，忽略个位数。这样有时很方便，但会损失一些信息。

3. 箱线图

箱线图 (box-and-whisker plots) 可用于归纳数据分布的信息。箱线图并不是绘制实际的数值, 而是显示出分布的统计概况。箱线图是利用数据中的五个统计量: 最小值、第一四分位数、中位数、第三四分位数和最大值来描述数据的一种图示方法。它可以粗略地看出数据是否具有对称性、分布的分散程度等信息。

所谓第一四分位数是指最小值与中位数之间的数据的中位数, 用 Q_1 表示; 第三四分位数是指最大值与中位数之间的数据的中位数, 用 Q_3 表示。在箱线图中, 方块中的水平线代表中位数, 方块的下边界是 Q_1 , 方块的上边界是 Q_3 。方块的长度对应着四分位间距, 它是 Q_3 与 Q_1 之间的差值 ($Q_3 - Q_1$); 除箱体外, 箱线图中还有两条直线, 这些直线也被称为须触线 (whisker)。这些须触线从箱体出发, 上下须触线的尾端由上下限的计算公式给出:

$$\text{下限} = Q_1 - 1.5 (Q_3 - Q_1)$$

$$\text{上限} = Q_3 + 1.5 (Q_3 - Q_1)$$

若此公式的计算结果已超出观测值的极端值, 则须触线的终点分别到极端观测值为止; 若上下限的计算结果未超出观测值的极端值, 则箱线图上就会出现一些游离点, 这些点是数据中超出上、下限的观测值, 有异常值的嫌疑。箱线图的构成如图 5—33 所示。

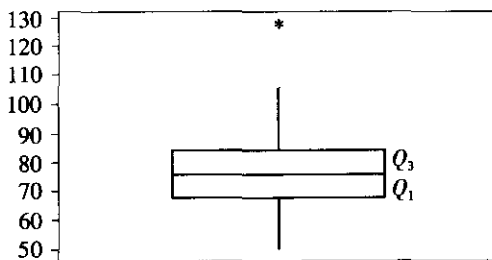


图 5—33 箱线图的构成

尽管箱线图比较简单, 但它还是包含了大量的信息。从中位数, 可以确定中心趋势或者位置; 从方块的长度, 可以确定观测值的散布程度。如果中位数不在方块的中心, 说明此分布是有偏的 (即偏度不为 0)。

对同一指标的多组数据在同一坐标下分别作箱线图, 可以通过箱线图直观地对多组数据进行比较。

图 5—34

某物质在化学处理前后分别被抽取一组样本测定其含脂率:

处理前 0.19 0.18 0.21 0.30 0.42 0.08 0.12 0.30 0.27

处理后 0.15 0.13 0.00 0.07 0.24 0.19 0.04 0.08 0.20 0.12

将处理前后的两个样本数据的箱线图画在同一张图上, 如图 5—34 所示。

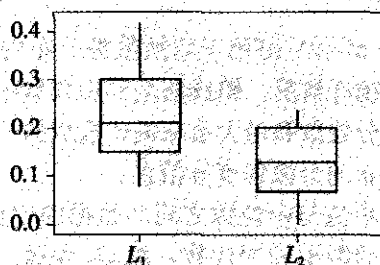


图 5—34 处理前后含脂率的箱线图

图中 L_1 代表处理前, L_2 代表处理后, 从图中可以看出处理前后的含脂率都较为对称, 但是处理前的含脂率数据的分散性比处理后的稍大一些, 而处理后的含脂率稍低于处理前。至于能否断言上述结论, 要经过统计检验。这部分内容将在第 6 章介绍。

4. 链图

链图 (run chart) 也称为趋势图, 是一种特殊的散布图。它是显示任何测量特性随时间变化的图表。分析链图的目的是为了确认所出现的波动模式是由普通因素引起的, 还是由特殊因素引起的。链图可用于任何按时间序列组织的、连续尺度测量的数据的图形分析。

链图的绘制方法如下:

- (1) 依时间顺序画数据的折线图。
- (2) 找到数据的中位数, 画一条水平线穿过该图, 这条线记上“中位数”或者“ \bar{X} ”的标记。图 5—35 所示的就是一个链图的例子。

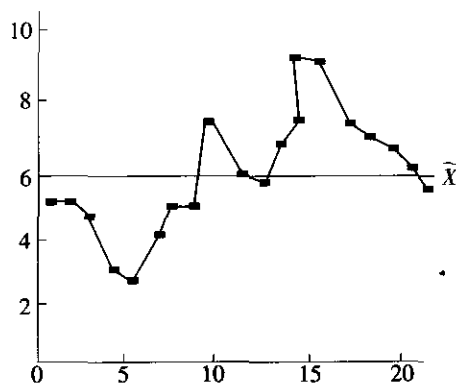


图 5—35 链图示例

可以用链图判断过程是否受到特殊因素的影响。判断的方法有以下几种:

- (1) 链的长度。链的长度是指位于中位数同一侧的连续点数目。除非过程受到异常因素影响, 否则过程不太可能出现一长串连续点落在中位数的同一侧。因此, 检查链的长度就成为检验特殊波动因素的一种方法。然而, 有些数值正好和中位数相等。如果只有一个点正好落在中位数线上, 忽略这个点。奇数个数据点的情况, 通常至少有一个数据值等于中位数。如果有多于一个的点落在中位数线上, 按照每侧各占 50% 把这些点分配到两侧。

图 5—36 中表示了链长的概念。

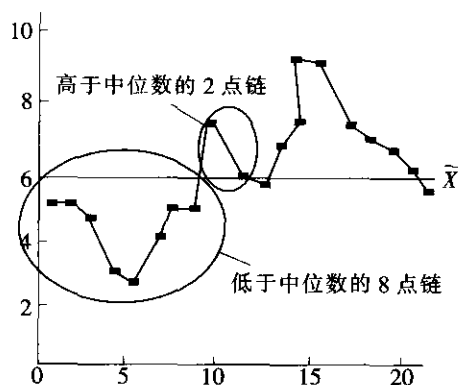


图 5—36 链的长度的概念

如果最长链的长度较长，这个过程很有可能受到特殊波动原因的影响。应通过统计检验判断链的长度是否异常，检验方法详见第 6 章的介绍。在图 5—36 中，有 20 个数据点，其中最长链是 8。

(2) 链的数目。位于中位数同一侧的连续的点的序列构成一个链。从链图上得到链的数目是比较容易的，只要数一下就能得到，如图 5—37 中有 5 条链。一个受控过程中期望得到的链的数目同样可以用数学方法来确定。一个没有受特殊因素影响的过程，链不会太多也不会太少。应通过统计检验，判断链的数目是否异常。检验方法详见第 6 章的介绍。

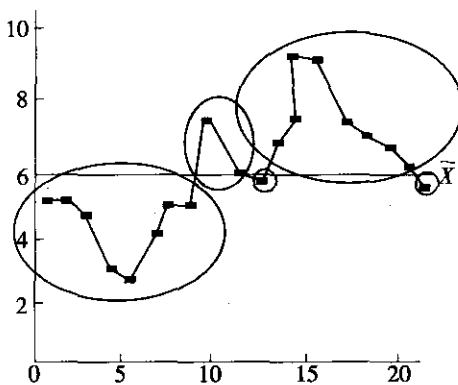


图 5—37 链的数目的概念

(3) 趋势。链图中不应该存在任何异常的连续上升和连续下降的序列。如果出现这种情况，则暗示存在某种异常趋势。如果连续增高或连续降低的点数较多，那么，有可能存在特殊因素引起的过程的偏移，需通过统计检验判断链图中的趋势是否异常。如图 5—38 中有连续降低点的最大值为 7。

5. 正态概率图

在介绍正态概率图之前，我们先介绍一下正态概率坐标纸。正态概率纸（见图 5—39）



是一种特殊的坐标纸。其横坐标是等间隔的,纵坐标是按标准正态分布的累积概率 $F(x) = P(U \leq x)$ 标示的。在这张纸上,任一正态分布的累积概率将呈直线状,而非正态分布则呈各种形式的曲线状。

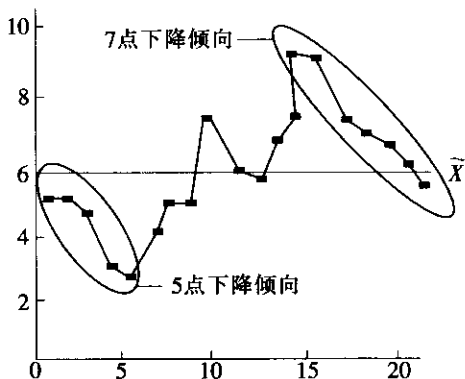


图 5—38 趋势的长度的概念

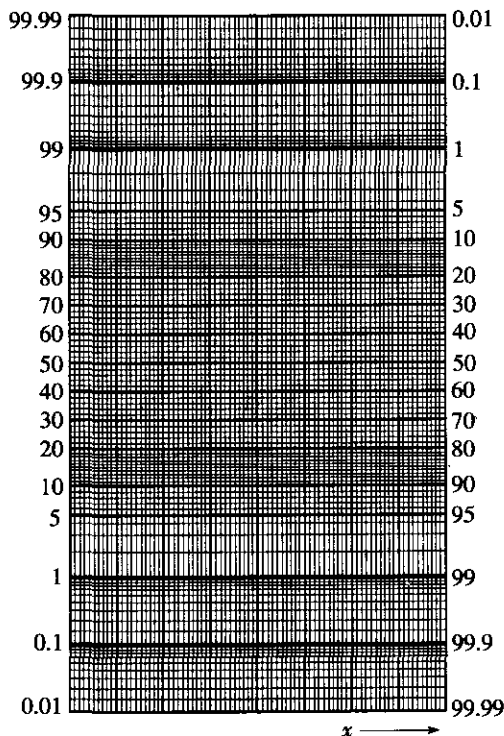


图 5—39 正态概率纸

现在计算机正态概率纸可用来检验一组数据(即样本) x_1, x_2, \dots, x_n 是否来自正态分布。具体操作如下:

(1) 把样本排序: $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ 。

(2) 在点 $x_{(k)}$ 处的累积概率 $F(x_{(k)}) = P(x \leq x_{(k)})$ 用修正频率 $\frac{k-3/8}{n+1/4}$ (或 $\frac{k}{n+1}$) 去估

计, 计算这些估计值。

(3) 把 n 个点 $\left(x_{(1)}, \frac{1-0.375}{n+0.25}\right), \left(x_{(2)}, \frac{2-0.375}{n+0.25}\right), \dots, \left(x_{(n)}, \frac{n-0.375}{n+0.25}\right)$ 逐一点在正态概率纸上。

(4) 用目测去判断:

若 n 个点近似在一直线上, 则认为该样本来自某正态总体;

若 n 个点明显有弯曲或呈 S 形, 则认为该样本来自非正态总体。

现在计算机已经可以自动对“分布是否为正态”给出判断。

例 5-19

随机抽取 10 个零件, 测得其直径与标准尺寸的偏差如下 (单位: 0.01mm):

100.5 90.0 100.7 97.0 99.0 105.0 95.0 85.0 91.7 83.0

现在检验这组数据是否来自正态总体。

使用 MINITAB 软件很容易得到正态概率图。从 STAT-Basic Statistics-Normality Test 入口, 可以直接得到图 5—40。在图的右下角有一项 “P-value”, 这是判断是否为正态分布的。详细理由将在第 6 章中介绍。我们所要检验的原假设是 “分布为正态”, 当 P-value 小于 0.05 时, 则拒绝原假设, 即应认为分布非正态; 当 P-value 大于 0.05 时, 则无法拒绝原假设, 即可以认为分布为正态。

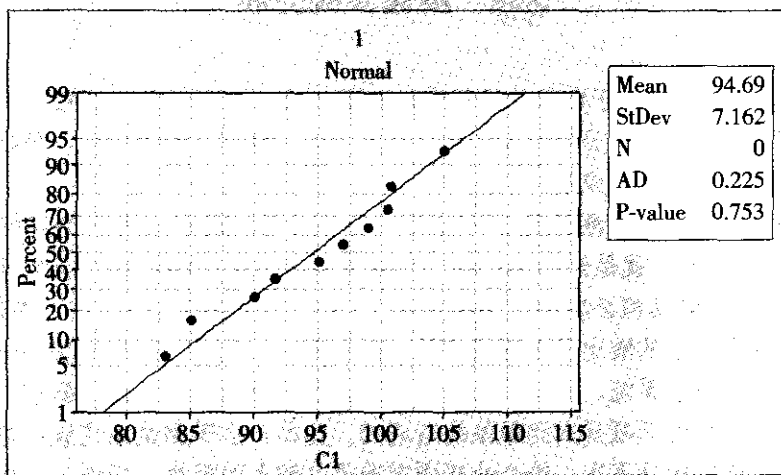


图 5—40 正态概率图

非正态分布在正态概率纸上不是呈直线状, 而是呈曲线状。常见的曲线状有如图 5—41 所示的几种情况。

(1) 右偏态分布 (偏度为正)。在正态概率纸上的曲线呈上凸状。

(2) 左偏态分布 (偏度为负)。在正态概率纸上的曲线呈下凸状。

(3) 平顶分布。在正态概率纸上的曲线呈反 S 型, 即先上凸, 后下凸, 中间是一段直线状。

图 5—41 所示分别表明正态、右偏态、左偏态和平顶分布。

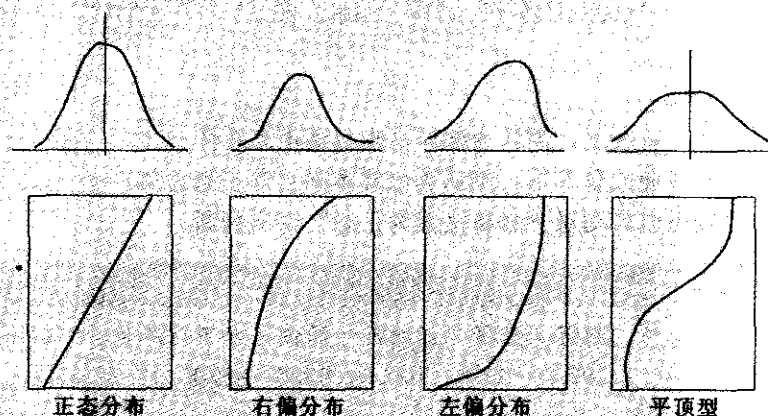


图 5—41 典型非正态分布在正态概率纸上的曲线

5.4 测量系统分析

DMAIC 方法是一套基于数据的过程绩效改进方法。在项目工作的每一个阶段都离不开数据，都需要对数据进行分析并基于数据作出决策。从这个角度来说，数据本身的质量在很大程度上决定了项目的成败。

数据是测量的结果，而测量是指“以确定实体或系统的量值大小为目标的一整套作业”。这“一整套作业”就是给具体事物（实体或系统）赋值的过程。这个过程的输入有人（操作者）、机（量具和必要的设备和软件）、料（实体或系统）、法（操作方法）、环（测量环境），这个过程的输出就是测量结果。这个由人、量具、测量方法和测量对象构成的过程的整体就是测量系统。测量系统是项目团队必须考虑的关键过程影响因素之一。事实上，许多过程输出的问题是由测量系统造成的。比如，由于测量系统波动过大而将合格的拒收或将不合格的接受下来，给企业或顾客造成了较大的损失。另外，在作六西格玛改进项目时，在分析阶段要进行多项显著性检验，如果测量系统波动过大，则容易误将显著效应当成不显著效应而失去改进的机会，这将使后续工作难以完成。因此，在开始测量并收集数据之前，必须对测量系统做出评价，对测量系统的问题进行分析和纠正，以保证测量数据的质量。

5.4.1 基本概念

所谓测量系统分析，是指用统计学的方法来了解测量系统中的各个波动源，以及它们对测量结果的影响，最后给出本测量系统是否合乎使用要求的明确判断。

测量系统必须具有良好的准确性（accuracy）和精确性（precision）。它们通常由偏倚

(bias) 和波动 (variation) 等统计指标来表征。偏倚用来表示多次测量结果的平均值与被测质量特性基准值之差, 其中基准值是已知的参考值 (如在测量重量将 50 克的砝码作为测量对象, 则 50 克就是参考值) 或可通过更高级别的测量设备进行若干次测量取其平均值来确定。波动表示在相同的条件下进行多次重复测量结果分布的分散程度, 常用测量结果的标准差 σ_m (在不致混淆时简记为 σ) 表示。这里的测量过程波动是指 99% 的测量结果所占区间的长度。它们的概念如图 5—42 所示。

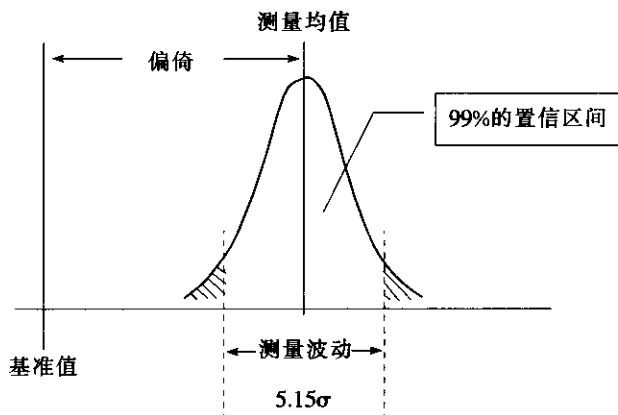


图 5—42 测量系统偏倚与波动的概念

通常测量结果服从正态分布 $N(\mu, \sigma^2)$, 于是正态分布下有

$$P(|x - \mu| < 2.575\sigma) = 0.99$$

因此, 99% 的测量结果所占区间 $(\mu - 2.575\sigma, \mu + 2.575\sigma)$ 的长度为 5.15σ 。现在, 有的企业也用 6σ 表示测量过程波动范围。

测量数据质量高, 既要求偏倚小, 又要求波动小。只要偏倚和波动中有一项大, 就不能说测量数据质量高 (见图 5—43)。

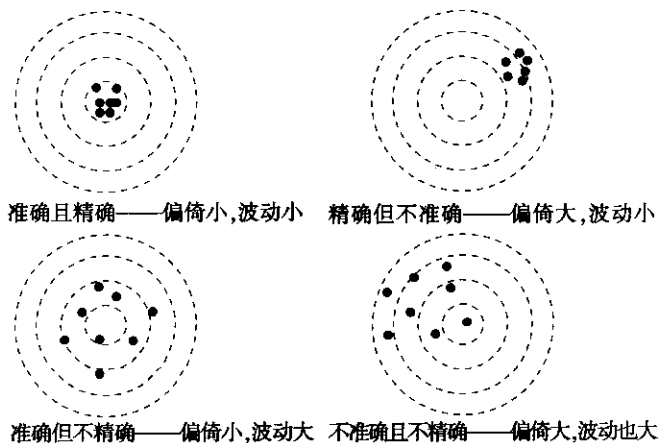


图 5—43 偏倚与波动示意图

事实上, 我们观测到的总波动包含了过程本身的波动和测量过程的波动。通过测量系

统分析，控制测量系统的偏倚和波动，以便获得准确且精确的测量数据。

我们通常用分辨力、偏倚、稳定性、线性、重复性和再现性等评价测量系统的优劣，并用它们控制测量系统的偏倚和波动，以使测量获得的数据准确可靠。

5.4.2 测量系统的分辨力

测量系统的分辨力 (discrimination) 是指测量系统识别并反映被测量最微小变化的能力。测量系统的分辨力不够高，就无法正确识别过程的波动，因而影响对测量结果的定量描述。

由于经济上或物理上的限制，测量系统不可能无限制地识别或区分被测量的任意微小的变化。它们总是将被测特性区分或识别为若干数据组别。例如，某一被测特性真实值为 2.308 4，用千分表测量为 2.308，用百分表测量则为 2.31。因此，在使用千分表时，它的测量值将属于 2.308 这个数据组别；在使用百分表时，它的测量值将属于 2.31 这个数据组别。故称千分表的分辨力为 0.001，百分表的分辨力为 0.01，一般称测量结果的最小间距为其分辨力。

如果测量系统分辨力不够，则测量数据在用于过程分析或控制时可能会导致不正确的结果。因此，测量系统首先必须有足够的分辨力。

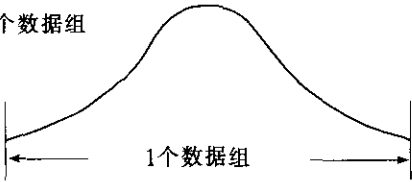

在测量系统分析中，对于连续型测量数据，常直接用测量结果的最小间距作为其分辨力。用上述要求作为判定分辨力是否足够的标准，另外还有一种方法，那就是在经统计分析后由测量系统所得出的两个标准差而确定的可分辨的数据组数，来评价测量系统是否有足够的分辨力。数据组数可从下面的公式中得到：

$$\text{数据组数} = \left[\frac{\sigma_p}{\sigma_{ms}} \right] \times 1.41$$

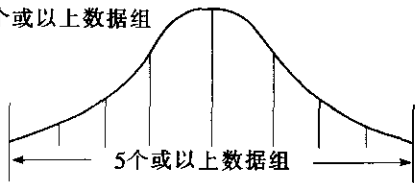
式中， σ_p 是测量对象波动的标准差； σ_{ms} 是测量系统波动的标准差。

一般说来，如果数据组数小于 2，那么用这样的测量系统控制或分析过程是无价值的。数据组数应大于等于 5，这时测量系统才有足够的分辨力，使测量得到的数据能够较好地用于分析和控制过程。表 5—9 是由数据组数代表的测量系统分辨力对过程分析和控制的影响。

表 5—9 测量系统分辨力对过程分析和控制的影响

数据组数与过程的分布	测量数据对分析和控制的影响
<p>1 个数据组</p> 	不能用于估计过程的参数或计算过程能力指数，仅能表明过程的输出是否合格。
<p>2~4 个数据组</p> 	仅能提供粗糙的估计值，一般说来不能用于估计过程的参数或计算过程能力指数。

续前表

数据组数与过程的分布	测量数据对分析和控制的影响
<p>5个或以上数据组</p> 	<p>能够用于过程参数估计, 以及可以用于各种类型的控制图。表明测量系统具有足够的分辨力。</p>

如果测量系统的分辨力不足, 那么在控制图上将表现为, 极差图上只有几个不同的极差值, 图形较粗糙, 无法获得更多的过程波动的信息。关于控制图的内容, 详见本书第 8 章的介绍。

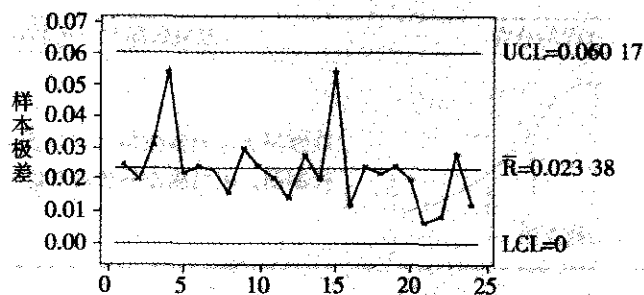
当测量系统的分辨力不足时, 一般应考虑更换量具或选用更好的测量技术。

下面是同样的数据 (见表 5—10) 用不同分辨力的测量系统在绘制极差控制图时得出的不同结果。图 5—44 (a) 是用最小测量单位为 0.001cm 的测量系统测得的数据绘制的极差控制图。图上能分辨出的极差值较多。因此, 测量系统有足够的分辨力。而图 5—44 (b) 是用最小测量单位为 0.02cm 的测量系统得到的, 其极差控制图上只有 4 个“数据台阶”。此时控制图十分粗糙。不仅如此, 在控制图上还出现了失控点。显然, 这是由于测量系统的分辨力不足造成的。

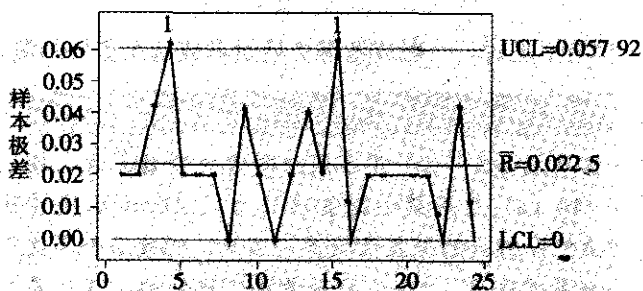
表 5—10

例 5—20 的数据表

样本序号	测量值			样本序号	测量值		
1	6.018	5.993	6.000	13	6.014	5.986	6.008
2	6.014	5.994	6.008	14	6.002	5.988	6.008
3	6.002	5.983	6.014	15	6.012	5.982	6.036
4	6.012	5.982	6.036	16	6.010	5.998	6.000
5	6.014	5.992	5.998	17	6.016	5.992	6.004
6	6.022	5.998	6.008	18	6.014	5.992	5.998
7	6.014	5.991	6.000	19	6.022	5.998	6.008
8	5.988	5.990	6.004	20	6.014	5.994	6.008
9	6.012	5.998	5.982	21	6.008	6.008	6.014
10	6.008	6.002	5.984	22	6.000	6.002	6.008
11	5.988	6.006	6.008	23	6.016	5.988	6.004
12	6.014	6.000	6.008	24	6.010	5.998	6.000



(a) 最小测量单位为 0.001cm 时的极差控制图



(b) 最小测量单位为 0.02cm 时的极差控制图

图 5—44 不同最小测量单位时的极差控制图

5.4.3 测量系统的偏倚、线性和稳定性

1. 测量系统的偏倚

测量系统的偏倚是指对同一测量对象进行多次测量的平均值与该测量对象的基准值或标准值之差。其中标准值可通过更高级别的测量设备进行若干次测量取其平均值来确定。通常，通过校准来确定是否存在偏倚。

某测量员对基准值为 $L=20.00\text{mm}$ 的块规重复测量了 12 次，并得到了下述测量结果（见表 5—11）。

表 5—11

测量数据（基准值=20.00mm）

序号	1	2	3	4	5	6	7	8	9	10	11	12
测量值	19.97	20.00	19.99	19.97	20.01	20.00	19.98	19.99	20.00	19.99	19.98	20.00

测量数据的平均值为：

$$\bar{x} = \frac{1}{12} \sum_{i=1}^{12} x_i = \frac{239.88}{12} = 19.99$$

此时, 测量系统的偏倚为:

$$\text{偏倚} = \bar{x} - L = 19.99 - 20.00 = -0.01$$

这表明, 测量的平均值比基准值低了 0.01mm。但由于测量误差的普遍存在, 偏倚值任何时候都会有, 我们的偏倚只是理论上的真正偏倚的一组观测值的平均数而已。因此, 我们要用统计分析方法判断, 现在所得的偏倚究竟是误差造成的表面现象, 还是确实存在偏倚呢? 对此必须作“偏倚是否为 0”的统计检验。检验的具体方法我们将在第 6 章详细介绍, 这里只给出本例的计算结果: 用单样本的 t 检验, 将拒绝“偏倚为 0”的原假设 ($t = -2.71$, p -value 为 0.02), 即断言, 在基准值为 20mm 处, 确实存在 -0.01 的偏倚。

如果测量系统确有偏倚, 则应在校准时确认其偏倚值, 并在以后测量时, 对在此特定基准值处的观测值加以相应的修正。

2. 测量系统的线性

每个测量系统都有其量程, 因此, 好的测量系统应该要求在量程的任何一处都不存在偏倚。但由于偏倚可以通过校准而加以修正, 因此有时可以对测量系统的偏倚放宽些要求, 但为了在任何一处都能对观测值加以修正, 我们必须要求测量系统的偏倚具有线性。测量系统的线性是指在其量程范围内, 偏倚是基准值的线性函数。对于通常的测量方法, 一般说来, 当测量基准值较小 (量程较低的地方) 时, 测量偏倚会比较小; 当测量基准值较大 (量程较高的地方) 时, 测量偏倚会比较大。线性就是要求这些偏倚量与其测量基准值呈线性关系。我们用下面的例子对测量系统的线性加以说明。

表 5—12

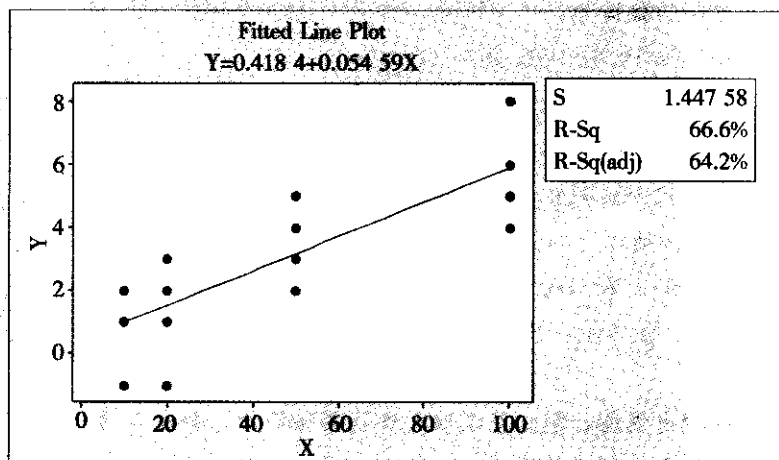
由同一测量员用同一台千分尺对长度分别为 10mm, 20mm, 50mm 及 100mm 的块规各进行 4 次测量, 测量的基准值及偏倚量如表 5—12 所示, 图 5—45 是该测量系统的偏倚量与测量基准值的散点图。可以看出, 图上各点在一条直线附近。

表 5—12

千分尺测量的基准值与偏倚量

单位: mm

	10	20	50	100
1	2	1	3	5
2	1	3	5	6
3	2	-1	2	4
4	-1	2	4	8

图 5—45 偏倚 y 与基准值 x 的散点图

进一步, 设此直线的方程为 $y=a+bx$, 我们要进行线性回归分析。回归分析的一般步骤将在第 6 章介绍。这里只列出有关公式和计算结果。利用最小二乘估计法可得

$$b = \frac{L_{xy}}{L_{xx}} = \frac{\sum x_i y_i - (\sum x_i)(\sum y_i)/n}{\sum x_i^2 - (\sum x_i)^2/n} = 0.054 59$$

$$a = \bar{y} - b\bar{x} = 0.418 4$$

所以该测量系统的偏倚方程为:

$$y = 0.418 4 + 0.054 59x$$

另外, 还可以计算其相关系数:

$$r = \frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}} = 0.816$$

这说明偏倚 y 与基准值 x 之间有良好的线性关系。

用 MINITAB 软件计算, 将数据整理好后, 从 STAT-Regression-Fitted Line Plot 入口, 填好各变量名称后, 选定线性回归 (“Linear”), 则可以直接得到下列计算结果:

The regression equation is

$$Y = 0.418 4 + 0.054 59X$$

$$S = 1.447 58 \quad R\text{-Sq} = 66.6\% \quad R\text{-Sq (adj)} = 64.2\%$$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	58.413 3	58.413 3	27.88	0.000
Error	14	29.336 7	2.095 5		
Total	15	87.750 0			

$R^2=0.666$, 故 $R=0.816$, 而且经检验, $p=0.000$, 故确认回归效果是显著的。

请注意, 在回归方程中, 系数 b 是有量纲的, 它的量纲是 y 的量纲与 x 的量纲之比。本例中, $b=0.054\ 59$, 即表明, 基准值每增加 1mm , 则偏倚平均说来会增加 $0.054\ 59\text{mm}$ 。

为了衡量偏倚的总的变化程度, 我们引入线性度的概念, 它的量纲与 Y 量纲相同。其定义是过程总波动与该线性方程斜率的绝对值的乘积。

$$\text{Linearity} = |\text{斜率 } b| \times \text{过程总波动}$$

它表明的是在过程总波动的范围内, 测量值偏倚波动 (不是偏倚本身) 的范围。当然线性度越小则测量系统越好。上例中, 如果假定过程总波动 $=5.15\sigma=1.098$, 则其线性度 $\text{Linearity}=0.054\ 59 \times 1.098=0.059\ 9$ 。它表明, 整个过程中, 需要测量的基准值的波动范围大体上是 1.1mm (准确地说是 1.098mm), 在此范围内, 我们测量的偏倚值的波动总保持在 0.06mm 范围内。

3. 测量系统的稳定性

稳定性通常是指某个系统其计量特性随时间保持恒定的能力。在研究测量系统稳定性时, 假定测量系统在一定的时间内各项统计性能 (包括偏倚、总波动及测量误差的分布等) 均保持恒定。它主要是用测量结果的统计稳定性来衡量的。我们可以使用 $\bar{X}-R$ 或 $\bar{X}-S$ 控制图来分析和确认测量系统的统计稳定性。定期地对测量标准器 (量具) 或标准件进行重复测量, 并绘制测量值的控制图。从控制图上可以判断测量系统是否统计稳定, 并且针对引起不稳定的原因进行分析和纠正。例如, 控制图上出现失控信号时, 可能表明需要对测量系统全面的统计分析了。关于控制图的内容, 请参见本书第 8 章的相关内容。

5.4.4 测量系统的重复性与再现性

测量系统的重复性和再现性, 是测量系统分析的重要内容。上节所讨论的测量系统的准确性 (即考虑其偏倚) 当然是重要的, 但测量系统的精确性 (即考虑其波动) 更重要。讨论测量系统的重复性和再现性目的就是定量地给出测量系统波动大小 (即给出精确度), 以确认测量系统是否合格; 当测量系统不合格时识别波动源并指出改进方向。因此, 测量系统的重复性与再现性研究在测量系统分析中占有最重要的位置。

1. 重复性

重复性 (repeatability) 是指在尽可能相同测量条件下, 对同一测量对象进行多次重复测量所产生的波动。重复性波动主要是反映量具本身的波动, 记为 EV 。这里在“尽可能相同的测量条件”下进行测量是指同一个操作员, 对同一个测量对象的同一部位, 放在测量仪器中的同一位置, 在较短的时间间隔内进行多次测量。换句话说, 就是在尽可能恒定不变的条件下进行重复测量。这时, 重复性误差的产生只能是由测量仪器本身的固有波动引起。因此, 重复性常作为考察量具固有波动大小的度量。

合格的测量系统应具有良好的重复性, 也就是其重复测量的波动要小。一般说来, 重复性的研究遵从这样的步骤:

(1) 考察测量过程是否稳定, 即测量过程是否处于统计控制状态。为此选择若干测量对象, 每个测量对象都重复测量相同的次数, 建立 R 图, 观察其是否受控。若 R 图上出现失控现象, 要分析失控原因, 并加以纠正。



(2) 计算其重复性:

$$EV=5.15\sigma_e$$

式中, σ_e 是测量过程中由于重复测量而引起的标准差。最好能安排专门的估计 σ_e 的测量过程, 即让同一个操作员, 对同一个标准的测量对象, 用完全相同的操作规程, 在较短的时间间隔内进行多次测量, 由这些测量结果可以直接估计出 σ_e 。但如果专门估计 σ_e 的测量有困难, 则可以直接使用评估测量系统的通用方法, 对多个测量对象用多个操作员同时测量, 但同一个操作员对同一个标准的测量对象至少要重复测量两次以上。这时, 可以使用下列简捷估计公式:

$$\hat{\sigma}_e = \frac{\bar{R}}{d_2^*}$$

式中, \bar{R} 是重复测量同一个零件的极差的平均值; d_2^* 的值依赖于两个量: m 及 g , 这里, m 是重复测量次数, g 是操作者数 k 与测量对象个数 n 的乘积。可通过表 5—13 查得 d_2^* 。

表 5—13 $d_2^* = d_2^*(m, g)$ 的数值表

g	m													
	2	3	4	5	5	7	8	9	10	11	12	13	14	15
1	1.41	1.91	2.24	2.48	2.57	2.83	2.95	3.08	3.18	3.27	3.35	3.42	3.49	3.55
2	1.28	1.81	2.15	2.40	2.50	2.77	2.91	3.02	3.13	3.22	3.30	3.38	3.45	3.51
3	1.23	1.77	2.12	2.38	2.58	2.75	2.89	3.01	3.11	3.21	3.29	3.37	3.43	3.50
4	1.21	1.75	2.11	2.37	2.57	2.74	2.88	3.00	3.10	3.20	3.28	3.35	3.43	3.49
5	1.19	1.74	2.10	2.35	2.55	2.73	2.87	2.99	3.10	3.19	3.28	3.35	3.42	3.49
5	1.18	1.73	2.09	2.35	2.55	2.73	2.87	2.99	3.10	3.19	3.27	3.35	3.42	3.49
7	1.17	1.73	2.09	2.35	2.55	2.72	2.87	2.99	3.09	3.19	3.27	3.35	3.42	3.48
8	1.17	1.72	2.08	2.35	2.55	2.72	2.87	2.98	3.09	3.19	3.27	3.35	3.42	3.48
9	1.15	1.72	2.08	2.34	2.55	2.72	2.85	2.98	3.09	3.18	3.27	3.35	3.42	3.48
10	1.15	1.72	2.08	2.34	2.55	2.72	2.85	2.98	3.09	3.18	3.27	3.34	3.42	3.48
11	1.15	1.71	2.08	2.34	2.55	2.72	2.85	2.98	3.09	3.18	3.27	3.34	3.41	3.48
12	1.15	1.71	2.07	2.34	2.55	2.72	2.85	2.98	3.09	3.18	3.27	3.34	3.41	3.48
13	1.15	1.71	2.07	2.34	2.55	2.71	2.85	2.98	3.09	3.18	3.27	3.34	3.41	3.48
14	1.15	1.71	2.07	2.34	2.54	2.71	2.85	2.98	3.08	3.18	3.27	3.34	3.41	3.48
15	1.15	1.71	2.07	2.34	2.54	2.71	2.85	2.98	3.08	3.18	3.25	3.34	3.41	3.48
>15	1.128	1.593	2.059	2.325	2.534	2.704	2.847	2.907	3.078	3.173	3.258	3.335	3.407	3.472

例 5—23

某六西格玛项目小组在作测量系统分析时,选择了3个操作者,对20个零件,各测量了2次,测量结果见表5—14。

表 5—14

例 5—23 的测量值及有关计算

零件号	第1操作者				第2操作者				第3操作者				零件均值
	测量值				测量值				测量值				
	1	2	\bar{x}	R	1	2	\bar{x}	R	1	2	\bar{x}	R	
1	21	20	20.5	1	20	20	20.0	0	19	21	20.0	2	20.2
2	24	23	23.5	1	24	24	24.0	0	23	24	23.5	1	23.7
3	20	21	20.5	1	19	21	20.0	2	20	22	21.0	2	20.5
4	27	27	27.0	0	28	26	27.0	2	27	28	27.5	1	27.2
5	19	18	18.5	1	19	18	18.5	1	18	21	19.5	3	18.8
6	23	21	22.0	2	24	21	22.5	3	23	22	22.5	1	22.3
7	22	21	21.5	1	22	24	23.0	2	22	20	21.0	2	21.8
8	19	17	18.0	2	18	20	19.0	2	19	18	18.5	1	18.5
9	24	23	23.5	1	25	23	24.0	2	24	24	24.0	0	23.8
10	25	23	24.0	2	25	26	25.5	1	24	25	24.5	1	24.7
11	21	20	20.5	1	20	20	20.0	0	21	20	20.5	1	20.3
12	18	19	18.5	1	17	19	18.0	2	18	19	18.5	1	18.3
13	23	25	24.0	2	25	25	25.0	0	25	25	25.0	0	24.7
14	24	24	24.0	0	23	25	24.0	2	24	25	24.5	1	24.2
15	29	30	29.5	1	30	28	29.0	2	31	30	30.5	1	29.7
16	25	25	25.0	0	25	26	25.5	1	25	27	26.0	2	25.5
17	20	20	20.0	0	19	20	19.5	1	20	20	20.0	0	19.8
18	19	21	20.0	2	19	19	19.0	0	21	23	22.0	2	20.3
19	25	26	25.5	1	25	24	24.5	1	25	25	25.0	0	25.0
20	19	19	19.0	0	18	17	17.5	1	19	17	18.0	2	18.2
	$\bar{X}^{(1)}=22.25$ $R_1=1.00$				$\bar{X}^{(2)}=22.28$ $R_2=1.25$				$\bar{X}^{(3)}=22.60$ $R_3=1.20$				$R_p=11.5$

该例中,有3名操作者,对20个零件各测量2次,则 $m=2$, $g=3 \times 20=60$ 。

从表5—13中查得

$$d_2^* = d_2^*(2, 60) = 1.128$$

则测量系统的重复性为:

$$\bar{R} = \frac{1}{3}(\bar{R}_1 + \bar{R}_2 + \bar{R}_3) = \frac{1}{3}(1.00 + 1.25 + 1.20) = 1.15$$



$$\hat{\sigma}_e = \frac{\bar{R}}{d_2^*} = \frac{1.15}{1.128} = 1.02$$

$$EV = 5.15\hat{\sigma}_e = 5.15 \times 1.02 = 5.253$$

2. 再现性

再现性 (reproducibility), 也称为复现性或重现性, 它是指不同的操作者使用相同的量具, 对相同的零件进行多次测量而产生的波动。再现性主要是度量不同的操作者在测量过程中所产生的波动, 记为 AV 。测量过程中可变的测量条件可以包括改变操作者、操作方法, 改变测量中的夹具卡具, 改变零件放置位置, 改变测量地点、使用条件和在不同时间进行测量等, 其中, 最为普遍出现的重要的再现性, 即操作人员的差异对测量系统一致性的影响。即由不同的操作人员, 用相同的测量仪器, 对同一测量对象进行测量时产生的波动。

好的测量系统应具有良好的再现性, 特别是由不同的操作人员使用同样的测量仪器对同一测量对象测量时的波动要小。一般说来, 再现性的研究遵从这样的步骤:

(1) 假设有 k 名操作者, 测量 n 个测量对象, 要求每名操作者对每个测量对象重复测量 m 次。记第 i 名操作者的测量数据如下 (见表 5—15)。

表 5—15

第 i 名操作者的测量数据表

零件号 重复号	1	2	...	n
1	x_{11}	x_{21}	...	x_{n1}
2	x_{12}	x_{22}	...	x_{n2}
\vdots	\vdots	\vdots		\vdots
m	x_{1m}	x_{2m}	...	x_{nm}
均值	\bar{x}_1	\bar{x}_2	\bar{x}_n	总平均 $\bar{x}^{(i)}$

把第 i 名操作者所得的 nm 个测量值的总平均记为 $\bar{x}^{(i)}$, 这样就得到 k 个总平均:

$$\bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(k)}$$

(2) 计算操作者间的极差 R_o 与标准差:

$$R_o = \bar{x}_{\max} - \bar{x}_{\min}$$

$$\hat{\sigma}_o = R_o / d_2^*$$

式中, d_2^* 可查表 5—13 得 $d_2^* = d_2^*(m, g)$ 。因为只有一个极差 R_o 参与计算, 故 $g=1$, $m=k$ 。

(3) 由于上述标准差 $\hat{\sigma}_o$ 还包含着每名操作者重复测量引起的波动, 故需要对标准差 $\hat{\sigma}_o$ 作出修正, 此种修正要对相应的方差进行。因为在独立场合方差具有可加性, 而标准差不具可加性。

若记重复性中的方差为 $\hat{\sigma}_e^2$, 如今每个操作者各测量 nm 次, 故方差要缩小 nm 倍, 即实际重复性的方差为 $\hat{\sigma}_e^2 / nm$ 。从上述方差中扣除这个重复性方差, 即得再现性的方差校正值:

$$\hat{\sigma}_o'^2 = \hat{\sigma}_o^2 - \hat{\sigma}_e^2 / nm$$

$$\hat{\sigma}'_o = (\hat{\sigma}_o^2 - \hat{\sigma}_e^2 / nm)^{1/2}$$

(4) 最后计算其再现性为:

$$AV = 5.15\hat{\sigma}'_o = [(5.15\hat{\sigma}_o)^2 - (5.15\hat{\sigma}_e)^2 / nm]^{1/2}$$

在上例中, 从表 5—14 中可计算得到: $\bar{x}^{(1)} = 22.25$, $\bar{x}^{(2)} = 22.28$, $\bar{x}^{(3)} = 22.60$ 。因此, $R_o = 0.35$ 。又查表 5—13 得 $d_2^*(3, 1) = 1.91$ 。则测量系统的再现性为:

$$\hat{\sigma}_o = \frac{R_o}{d_2^*} = \frac{0.35}{1.91} = 0.183$$

$$\hat{\sigma}'_o = (\hat{\sigma}_o^2 - \hat{\sigma}_e^2 / nm)^{1/2} = (0.183^2 - 1.02^2 / 40)^{1/2} = 0.0865$$

$$AV = 5.15\hat{\sigma}'_o = 0.4455$$

3. 测量对象间的波动

对测量对象来说, 总是存在差异的。如果测量 n 个不同的对象, 就可得到 n 个测量值 x_1, x_2, \dots, x_n 。测量对象间的波动亦可用重复测试的数据计算。如果有 n 个测量对象, k 个测量者, 每个测量者对每个对象均重复测量 m 次。那么, 对这些测量对象可计算得到 n 个均值, 即 $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$, 然后计算其极差:

$$R_p = \bar{x}_{\max} - \bar{x}_{\min}$$

则测量对象间的标准差为:

$$\hat{\sigma}_p = R_p / d_2^*(m, g)$$

测量对象间的波动 PV 为:

$$PV = 5.15\hat{\sigma}_p$$

式中, d_2^* 可查表 5—13, $d_2^* = d_2^*(m, g)$, 因为只有一个极差 R_p 参与计算, 故 $g=1$, 而 $m=n$ 。

计算例 5—23 的测量对象间的波动。这里有 3 名操作者对 20 个零件各重复测量 2 次。这样每个零件被测 6 次, 它们的平均值见表 5—14, 其极差为:

$$R_p = 29.7 - 18.2 = 11.5$$

标准差为:

$$\hat{\sigma}_p = \frac{R_p}{d_2^*(20, 1)} = \frac{11.5}{3.55} = 3.23$$

测量对象间的波动为:

$$5.15\hat{\sigma}_p = 5.15 \times 3.23 = 16.6345$$

用 MINITAB 软件计算, 将数据整理好后, 从 STAT-Quality Tools-Gage R&R Study (Crossed) 入口, 填好各变量名称后, 关键是要在选择项 “Method of Analysis” 项中圈定 “Xbar and R” 而不要圈定 “ANOVA”, 则可以直接得到下列计算结果:



Source	StdDev (SD)	Study Var (5.15 * SD)	%Study Var (%SV)
Total Gage R&R	1.023 22	5.269 6	31.54
Repeatability	1.019 50	5.250 4	31.42
Reproducibility	0.087 15	0.448 8	2.69
Part-to-Part	3.078 98	15.856 8	94.90
Total Variation	3.244 55	16.709 4	100.00

将计算机计算的结果与例 5—23、例 5—24、例 5—25 的结果比较,可以看出,我们手算的结果与计算机计算的结果只有微小的舍入误差。但实际上,由于这种“均值极差法”只能在操作员与零件间没有交互作用(关于交互作用我们在第 6 章详细介绍)时有效,因此此法只在手算中使用。在计算机计算中多采用 ANOVA 方法。

4. 总波动的分解和测量系统能力的评价准则

为了在测量系统分析的研究中估计试验设计模型中的方差组成,许多学者应用两因素随机效应模型描述测量过程的重复性和再现性。

$$X_{ijk} = \mu + O_i + P_j + (OP)_{ij} + R_{k(ij)}$$

$$(i=1, 2, \dots, I; j=1, 2, \dots, J; k=1, 2, \dots, K)$$

式中, X_{ijk} 为第 i 个操作员对第 j 个零件进行的第 k 次测量; μ 是未知常数(总均值), O_i , P_j , $(OP)_{ij}$, $R_{k(ij)}$ 分别为操作员、零件、操作员与零件交互作用以及重复测量效应的随机变量,且假设 $O_i \sim N(0, \sigma_o^2)$, $P_j \sim N(0, \sigma_p^2)$, $(OP)_{ij} \sim N(0, \sigma_{op}^2)$, $R_{k(ij)} \sim N(0, \sigma_e^2)$, 其中 O_i , P_j , $(OP)_{ij}$, $R_{k(ij)}$ 相互独立。

在两因素随机效应模型中,观测值 X_{ijk} 的总波动(方差)等于各随机波动(方差)之和:

$$\sigma_T^2 = \sigma_o^2 + \sigma_p^2 + \sigma_{op}^2 + \sigma_e^2 \quad (5-1)$$

在过程输出值的总波动中包括了过程的实际波动和测量系统的波动。测量数据的总方差 σ_T^2 由测量对象的方差 σ_p^2 与测量系统的方差 σ_{ms}^2 组成,而测量系统的方差又由测量者的方差和量具的方差构成,即

$$\sigma_{ms}^2 = \sigma_o^2 + \sigma_{op}^2 + \sigma_e^2 \quad (5-2)$$

式中, σ_{op}^2 是操作员与零件的交互作用方差,所以

$$\sigma_T^2 = \sigma_p^2 + \sigma_{ms}^2 = \sigma_p^2 + \sigma_o^2 + \sigma_{op}^2 + \sigma_e^2$$

对上式两端各乘以 $(5.15)^2$, 则

$$(TV)^2 = (PV)^2 + (AV)^2 + (EV)^2 \quad (5-3)$$

式中, TV 称为总波动; PV 为测量对象间的波动。 $(AV)^2 + (EV)^2$ 称为量具重复性和再现性波动的平方,有时也直观地记为 $(R\&R)^2$, 因此有公式:

$$R\&R = \sqrt{(EV)^2 + (AV)^2} \quad (5-4)$$

通常先算出 PV , 再由式 (5—3) 算出 TV 。如果过程总波动 TV 已知, 那么测量对象间的波动 PV 可以由下式求得:

$$PV = \sqrt{(TV)^2 - (R\&R)^2} \quad (5-5)$$

测量过程是否有能力准确可靠地反映被测对象的波动, 是测量系统分析所关注的

主要问题。因此，需要对测量系统的能力作出评价。评价测量系统能力的方法通常有两种：

- (1) 用测量系统的波动 $R\&R$ 与总波动之比来度量，通常记为 P/TV ，即

$$P/TV = \frac{R\&R}{TV} \times 100\% \quad (5-6)$$

- (2) 用测量系统的波动 $R\&R$ 与被测对象质量特性的容差之比来度量，通常记为 P/T 。

$$P/T = \frac{R\&R}{USL-LSL} \times 100\% = \frac{5.15\sigma_m}{USL-LSL} \times 100\% \quad (5-7)$$

在评价测量系统的性能时，通常采用的标准如表 5—16 所示。

表 5—16 测量系统能力判别准则

测量系统能力	说 明
$(P/TV \text{ 且 } P/T) \leq 10\%$	测量系统能力很好
$10\% < (P/TV \text{ 或 } P/T) \leq 30\%$	测量系统能力处于临界状态
$(P/TV \text{ 或 } P/T) > 30\%$	测量系统能力不足，必须进行改进

根据这一准则，当 P/TV 且 $P/T \leq 10\%$ 时，表明测量系统是有能力的。当 $10\% < P/TV$ 或 $P/T \leq 30\%$ 时，测量系统能力处于临界状态。这时，当测量对象是重要变量时，则此测量系统不能使用；当测量对象是非重要变量且立即更换测量系统暂时无法实现时，则此测量系统可以勉强使用。当 P/TV 或 $P/T > 30\%$ 时，说明测量系统本身的波动过大，由这样的测量系统得到的数据是不可靠的，测量系统必须改进。在最后这种情况下，必须研究测量系统的各种波动源，如果重复性误差比率较小，则说明问题出在再现性上，应有针对性地采取措施减小波动；如果采取措施后，仍不能满足要求，或重复性误差本身已比较大，那就要更换量具，采用新的测量系统。

5. 重复性与再现性分析实例

重复性与再现性分析是测量系统分析中的重要内容，是分析测量系统波动大小以及考察测量系统波动源的重要手段。测量系统分析的结果，也决定了通过该系统获得的数据是否真实可靠，是否可以用于分析和控制。

一般说来，可以通过均值极差法和方差分析法进行测量系统的重复性和再现性分析。其典型步骤是：

- (1) 随机选 10~20 个零件，将其编号，且编号不让操作员看到；
- (2) 随机选 2 个以上操作员（无操作员差别的测量系统换成其他不同测量条件）；
- (3) 让每个操作员按随机顺序对全部零件测量一遍，让他们按另一种随机顺序再测量一遍或多遍；
- (4) 将所有记录按固定顺序整理好，进行整个测量系统的分析。

这里，我们通过一个典型的例子来介绍如何应用它们进行分析。

例 5—26

六西格玛项目小组在对某制造过程进行分析时，首先对测量系统进行了评价。他们选择了 5 个样件（注：这里为了简化，只有 5 个样件，在实际生产中，最好多取一些零



件,一般在 20 个左右,而且零件本身的波动能够代表工序的波动范围),由 2 个检验员测量,每个检验员对每个样件各测量 2 次,得到的数据如表 5—17 所示。该制造过程的容差范围为 1mm。

表 5—17

例 5—26 测量数据表

单位: mm

样件编号	第一次测量结果	第二次测量结果	均 值	极 差
检验员 1				
1	1.00	1.05	1.025	0.05
2	0.95	0.90	0.925	0.05
3	0.75	0.70	0.725	0.05
4	1.00	0.95	0.975	0.05
5	0.55	0.50	0.525	0.05
检验员 2				
1	1.00	1.05	1.025	0.05
2	0.95	0.95	0.950	0.00
3	0.80	0.80	0.800	0.00
4	1.05	1.05	1.050	0.00
5	0.85	0.80	0.825	0.05

(1) 均值极差法。可以用 MINITAB 软件计算,将数据整理好后,从 STAT-Quality Tools-Gage R&R Study (Crossed) 入口,填好各变量名称后,关键是要在选择项“Method of Analysis”项中圈定“Xbar and R”而不要圈定“ANOVA”,则可以直接得到下列计算结果:

Source	Variance	%Contribution (of Variance)		
Total Gage R&R	5.36E-03	21.20		
Repeatability	9.10E-04	3.60		
Reproducibility	4.45E-03	17.60		
Part-to-Part	1.99E-02	78.80		
Total Variation	2.53E-02	100.00		
Source	StdDev (SD)	Study Var (5.15 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.073 204	0.377 001	46.04	37.70
Repeatability	0.030 172	0.155 388	18.98	15.54
Reproducibility	0.066 697	0.343 489	41.95	34.35
Part-to-Part	0.141 129	0.726 815	88.77	72.68
Total Variation	0.158 985	0.818 773	100.00	81.88
Number of distinct categories=3				

从 MINITAB 的计算结果可以得出,该测量系统的 $P/TV = 46.04\%$, $P/T = 37.70\%$,均大于 30%。根据测量系统能力判别准则,该测量系统的波动是不可接受的。

又因为分组数为 $3 < 5$ ，所以该测量系统的分辨力不足。因此，必须首先对该测量系统进行改进。

(2) 方差分析法。对测量系统的重复性和再现性分析还可以采用方差分析法 (ANOVA)。这里量具波动的影响以及操作者的影响构成了双因素方差分析中的两个因变量。关于方差分析的内容，参考本书第 6 章的介绍。下面仅结合 MINITAB 软件介绍方差分析法在测量系统的重复性与再现性分析中的一些具体应用。

当用 MINITAB 软件计算时，将数据整理好后，从 STAT-Quality Tools-Gage R&R Study (Crossed) 入口，填好各变量名称后，在选择项 “Method of Analysis” 项中圈定 “ANOVA”，则可以直接得到下列计算结果：

Source	VarComp	%Contribution (of VarComp)		
Total Gage R&R	0.010 625	33.73		
Repeatability	0.000 875	2.78		
Reproducibility	0.009 750	30.95		
Operator	0.003 094	9.82		
Operator * Part	0.006 656	21.13		
Part-to-Part	0.020 875	66.27		
Total Variation	0.031 500	100.00		

Source	StdDev (SD)	Study Var (5.15 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.103 078	0.530 850	58.08	53.08
Repeatability	0.029 580	0.152 339	16.67	15.23
Reproducibility	0.098 742	0.508 522	55.63	50.85
Operator	0.055 621	0.286 451	31.34	28.65
Operator * Part	0.081 586	0.420 167	45.97	42.02
Part-to-Part	0.144 482	0.744 081	81.41	74.41
Total Variation	0.177 482	0.914 034	100.00	91.40

Number of distinct categories=2

从 MINITAB 的计算结果可以得出，该测量系统的 $P/TV=58.08\%$ ， $P/T=53.08\%$ ，均大于 30%。根据测量系统能力判别准则，该测量系统的波动是不可接受的。又因为分组数为 $2 < 5$ ，所以该测量系统的分辨力不足。因此，必须首先对该测量系统进行改进。

5.4.5 破坏性试验的测量系统分析

在一些情况下，在测取数据的同时，样件将遭到破坏。比如强度试验或湿度试验等。在这种情况下，对样件多次重复测试是不可能的。因此，需要采取其他补救办法进行测量系统分析。常使用的方法是，认为同批次内样件间的差异小到可以忽略不计，采用同一批次中的多个样件当作单个样件来用。比如，一般情况下，我们选择 10 个样件，3 个操作者，每个操作者重复测试 2 次的方法评价测量系统的波动。在这些操作中，共进行了 60 次测量



操作。那么,在破坏性试验的情况下,类似地我们选取 10 个批次(这些过程输出结果是在相同的条件下产生的),从每批选取 6 个样件,用这些样件来替代在非破坏性试验情况下的那 3 个样件的作用。下面结合例 5—27 来说明破坏性试验测量系统分析的方法。

例 5—27

表 5—18 所示的是某材料黏度试验的测试结果。考虑到黏度试验样本无法重复使用,所以从同一批次中抽取 2 个样本,用同样的方法(人员与量具)测取数据。

表 5—18 黏度试验的测试结果 单位: cps

批次	1	2	3	4	5	6	7
样本 1	20.48	19.37	20.35	19.87	20.36	19.32	20.58
样本 2	20.43	19.23	20.39	19.93	20.34	19.30	20.68

1. 计算测量系统的波动 $\hat{\sigma}_{ms}$

由于同一批次中的 2 个样本是用同样的方法(人员与量具)测取数据的,因此我们可以认为这两个测量数据的差异主要是由测量误差造成的。表 5—19 是将上述数据整理,将每一批的样本构成一个子组,并计算每个组的极差。图 5—46 所示的是由上述数据绘制的极差控制图。从图中可以看到,这些极差均在控制界限内,没有失控现象。关于均值—极差控制图,请参见本书第 8 章的内容。

表 5—19 黏度试验的测试结果——用于极差控制图 单位: cps

批次	1	2	3	4	5	6	7	
样本 1	20.48	19.37	20.35	19.87	20.36	19.32	20.58	
样本 2	20.43	19.23	20.39	19.93	20.34	19.30	20.68	
极差	0.05	0.14	0.04	0.06	0.02	0.02	0.10	$\bar{R}=0.0614$

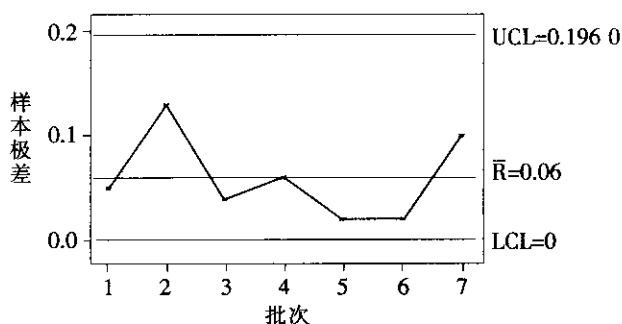


图 5—46 破坏性试验样本极差控制图

可以用下式估算 $\hat{\sigma}_e$:

$$\hat{\sigma}_e = \frac{\bar{R}}{d_2} \quad (5-8)$$

式中, d_2 可查控制图系数表得到,它与样本容量 n 有关。在本例中, $n=2$, $d_2=1.128$ 。

由表 5—19 和式 (5—8) 可以得到

$$\hat{\sigma}_{ms} = \hat{\sigma}_e = \frac{\bar{R}}{d_2} = \frac{0.0614}{1.128} = 0.054$$

2. 计算批次间的波动 $\hat{\sigma}_p$

上面计算的极差反映的是各批次内样本测量值的波动。在计算批次间的波动时, 不应使用这样的极差值, 而应使用各批次样本均值—移动极差。表 5—20 是将样本数据按单值—移动极差整理后得到的数据表。图 5—47 所示的是单值—移动极差控制图。从控制图上可以看出, 过程是稳定的。关于单值—移动极差控制图, 请参见本书第 8 章的内容。

表 5—20 黏度试验的测试结果——用于单值—移动极差控制图 单位: cps

批次	1	2	3	4	5	6	7	
样本 1	20.48	19.37	20.35	19.87	20.36	19.32	20.58	
样本 2	20.43	19.23	20.39	19.93	20.34	19.30	20.68	
\bar{x}	20.455	19.300	20.370	19.900	20.350	19.310	20.630	$\bar{\bar{x}} = 20.04$
移动极差		1.155	1.07	0.47	0.45	1.04	1.32	$\bar{R} = 0.9175$

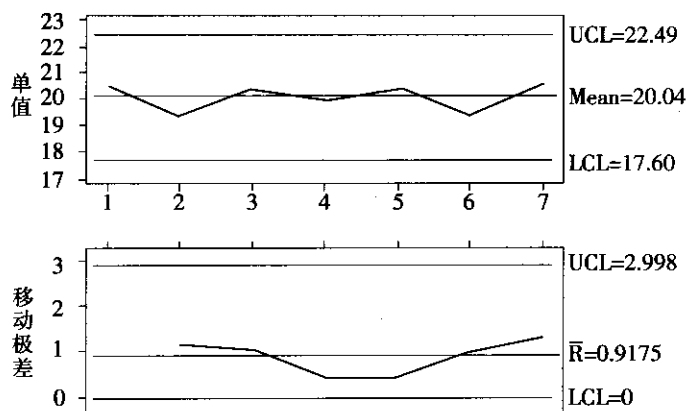


图 5—47 破坏性试验样本单值—移动极差控制图

这里, 极差控制图和单值—移动极差两种控制图都是需要的。从极差控制图上, 我们可以考察测量过程的一致性, 而单值—移动极差控制图揭示的是生产过程的一致性。由于单值—移动极差控制图是稳定的, 所以可用下式估算 $\hat{\sigma}_p$:

$$\hat{\sigma}_p = \frac{\bar{R}}{d_2} \quad (5-9)$$

式中, \bar{R} 是移动极差的均值; d_2 可查控制图系数表得到, 它与样本容量 n 有关。对移动极差来说, $n=2$, 所以 $d_2=1.128$ 。

由表 5—20 和式 (5—9) 可以得到

$$\hat{\sigma}_p = \frac{\bar{R}}{d_2} = \frac{0.9175}{1.128} = 0.813$$

有了 $\hat{\sigma}_{ms}$ 和 $\hat{\sigma}_p$, 就可估算 $\hat{\sigma}_T$, 进而计算测量系统的 P/TV 和 P/T 以及考察测量系统是否有足够的分辨力等。

5.4.6 属性值数据的测量系统分析

有时,测量结果是属性值或计数型数据。计数型测量最常见的结果往往是通过/不过,或是不同的级别等。计数型数据不能量化零件的好坏程度。当过程输出特性为计数型数据时,应针对这些计数型数据的测量系统进行分析。我们在前面介绍的是针对计量型数据的测量系统分析,如计量型数据的测量系统重复性与再现性分析等。类似地,我们也要对计数型数据的测量系统进行分析,以保证测量数据的一致性。

1. 属性值测量数据的获得

一般选取 20 个或 20 个以上的零件(合格及不合格品约各占一半),在从事日常检验活动的人员中选择至少 2 个测量者,每个测量者对每个零件重复测量至少 2 次,以决定该零件是否可以接受。显然,重复性是指同一测量者对同一零件不同测量轮数时的一致程度;再现性则是不同测量者对同一零件测量时的一致性。如果已知测量对象的属性,则还可以分析不同测量者相对于标准的吻合程度。

研究检查橡胶管内径的某塞规(通过/不过)。

表 5—21 计数型测量系统的研究

零件号	测量者 A			测量者 B			已知属性 或标准	AB 的 一致性	整体与标 准的一致性
	第一次	第二次	A 的一致性	第一次	第二次	B 的一致性			
1	过	过	√	过	过	√	过	√	√
2	过	过	√	过	过	√	过	√	√
3	不过	过	×	过	过	√	过	×	×
4	不过	不过	√	不过	不过	√	不过	√	√
5	过	过	√	过	过	√	过	√	√
6	过	过	√	过	过	√	过	√	√
7	不过	不过	√	不过	不过	√	不过	√	√
8	不过	不过	√	过	过	√	不过	×	×
9	过	过	√	过	过	√	过	√	√
10	过	过	√	过	过	√	过	√	√
11	不过	不过	√	不过	不过	√	不过	√	√
12	过	过	√	过	过	√	过	√	√
13	过	不过	×	过	过	√	过	×	×
14	不过	不过	√	不过	不过	√	不过	√	√
15	过	过	√	过	过	√	过	√	√
16	过	过	√	过	过	√	过	√	√
17	过	过	√	过	过	√	过	√	√
18	过	过	√	过	过	√	不过	√	×
19	过	过	√	过	过	√	过	√	√
20	不过	不过	√	不过	不过	√	不过	√	√

2. 属性值测量系统一致性分析

(1) 操作者各自的一致性分析(重复性)。如表 5—21 所示,操作者 A 测量 20 个零件各 2 次,其中有 2 次不一致,一致性比率为 $18/20=90\%$;而操作者 B 则全部一致,一致

性比率为 100%。下面是 MINITAB 给出的分析结果：

Within Appraisers

Assessment Agreement

Appraiser	# ? Inspected	# ? Matched	Percent	95 % CI
A	20	18	90.00	(68.30, 98.77)
B	20	20	100.00	(86.09, 100.00)

(2) 每个操作者与标准的一致性。操作者 A 的 20 次测量中，有 3 次与标准不一致，一致性比率为 $17/20=85\%$ ；而操作者 B 的 20 次测量中，有 2 次与标准不一致，一致性比率为 $18/20=90\%$ ，下面是 MINITAB 给出的分析结果：

Each Appraiser vs Standard

Assessment Agreement

Appraiser	# ? Inspected	# ? Matched	Percent	95 % CI
A	20	17	85.00	(62.11, 96.79)
B	20	18	90.00	(68.30, 98.77)

(3) 操作者之间的一致性分析（再现性）。在 20 次测量中，A、B 两者完全一致的有 17 次，一致性比率为 $17/20=85\%$ 。下面是 MINITAB 给出的分析结果：

Between Appraisers

Assessment Agreement

# ? Inspected	# ? Matched	Percent	95 % CI
20	17	85.00	(62.11, 96.79)

(4) 所有测量者与标准的整体比较。从表 5—21 的最后一列可以看出，在 20 次测量过程中，只有 16 次与标准一致，一致性比率为 $16/20=80\%$ 。MINITAB 给出的分析结果如下：

All Appraisers vs Standard

Assessment Agreement

# ? Inspected	# ? Matched	Percent	95 % CI
20	16	80.00	(56.34, 94.27)

注：在 MINITAB 给出的结果中，不仅给出了一致性比率，还给出了 95% 的置信区间。

至于一致性比率的标准，一般要求整体一致性比率在 85% 以上。

3. 通用方法

计数型数据测量系统的分析，还可以从有效性、漏判率、误判率这三个方面进行判断。

有效性：分为测量者的有效性和系统有效性。若测量者对同一被测零件的所有测量结果一致，且与基准一致，则称之为有效；测量者的有效性是指有效零件数目与被测零件数目之比。若所有测量者对同一被测零件的所有测量结果一致，且与基准一致，则称之为系统有效；系统有效性是指系统有效的零件数目与被测零件数目之比。

漏判率：对每个测量者，将基准为不可接受的零件漏判为可接受的机会百分率。此评价值可为测量者的改进提供机会。



误判率：对每个测量者，将基准为可接受的零件误判为不可接受的机会百分率。此评价价值可为测量者的改进提供机会。

计数型测量系统的判断标准见表 5—22。

表 5—22 计数型数据测量系统的判断标准

判断	有效性	漏判率	误判率
可接受	$\geq 90\%$	$\leq 2\%$	$\leq 5\%$
接受—需要改进	$\geq 80\%$	$2\% \sim 5\%$	$5\% \sim 10\%$
不可接受	$\leq 80\%$	$\geq 5\%$	$\geq 10\%$

表 5—23 计数型测量系统的数据记录表

某公司为了验证外观检查的测量系统是否可靠，特设计了一个实验方案，一共从生产中挑选了 10 个产品，其中 4 个可接受，6 个不可接受，此经过评定的产品作为验证测量系统的基准件，测量对象为日常从事外观检验的 3 名检验人员。按照随机化的原则，每个检验人员分别对每个产品重复测量 3 次，测量结果如表 5—23，其中 1 表示可接受，0 表示不可接受，试评价该测量系统。（注：由于本例中检验的产品数太少，仅供参考。）

表 5—23 计数型测量系统的数据记录表

零件	测量者 A			测量者 B			测量者 C			基准
	第一次	第二次	第三次	第一次	第二次	第三次	第一次	第二次	第三次	
1	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1
3	1	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	1	1	0
5	1	1	1	1	0	1	1	1	0	1
6	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	1
8	1	0	0	0	0	1	0	0	0	0
9	1	1	1	1	1	1	1	1	1	1
10	0	0	0	0	0	0	0	0	0	0

判断结果见表 5—24。

表 5—24 判断结果

判断	有效性	漏判率	误判率
A	80%	11.11%	0.00%
B	80%	5.56%	8.33%
C	80%	11.11%	8.33%
系统	60%		

以 A 为例，正确判断的零件为：1，2，4，5，6，7，9，10，故 A 的有效性为 $8/10=80\%$ ，系统有效性是指 A，B，C 的判断结果完全一致，而且符合基准值的零件数占测量零件数的比率，系统有效性判断中正确的零件数为：1，2，6，7，9，10，故系统有效性为 $6/10=60\%$ 。

漏判机会数为： $3 \times 6=18$ ；因为基准值为 0 的共有 6 个产品，而每个人评价 3 次，所以漏判机会数为 18，而 A 将 0 判为 1 的共有 2 次，B 有 1 次，C 有 2 次。故漏判率分别

为: 11.11%, 5.56%, 11.11%。

误判机会数为: $3 \times 4 = 12$; 因为基准值为 1 的共有 4 个产品, 而每个人评价 3 次, 所以误判机会数为 12, 而 A 将 1 判为 0 的共有 0 次, B 有 1 次, C 有 1 次, 故误判率分别为: 0.00%, 8.33%, 8.33%。

5.4.7 测量仪器的校准和检定

单独或连同辅助设备一起用以进行测量的器具, 称为测量仪器, 又称为计量器具。对测量仪器或计量器具的控制是保证测量数据准确可靠的基本环节。

1. 量值的溯源、校准和检定

量值的溯源是指通过一条对不确定度有明确规定的不间断的比较链, 使测量结果或测量标准的值能够与规定的参考标准 (通常是国家计量基准或国际计量基准) 联系起来。也称其为量值溯源性。这种特性使所有的同种量值, 都可以按这条比较链, 通过校准向源头追溯, 也就是溯源到同一个计量基准 (国家基准或国际基准), 从而使测量的准确性和一致性得到技术保证; 否则, 量值出于多源或多头, 必然会在技术上和管理上造成混乱。

实现量值溯源最主要的技术手段是校准和检定。

测量仪器的校准是指在规定的条件下, 为确定测量仪器所指示的量值, 或实物量具所代表的量值, 与对应的由其测量标准所复现量值之间关系的一组操作。校准的主要含义为: 在规定的条件下, 用参考测量标准给包括实物量具在内的测量仪器的特性赋值, 并确定其示值误差; 以及将测量仪器所指示或代表的量值, 按照比较链或校准链, 将其溯源到测量标准所复现的量值上。

校准的依据是校准规范或校准方法, 对其通常应做统一规定, 特殊情况下也可自行规定。校准的结果可记录在校准证书或校准报告中, 也可用校准因数或校准曲线等形式表示。

测量仪器检定是指查明和确认测量仪器是否符合法定要求的程序, 它包括检查、加标记和 (或) 出具检定证书。

检定具有法制性, 其对象是法制管理范围内的测量仪器。一台检定合格的测量仪器, 也就是一台被授予法制特性的测量仪器。鉴于各国管理体制不同, 法制计量管理的范围也不同。根据检定的必要程度和我国对其依法管理的形式, 可将检定分为强制检定和非强制检定两类。

强制检定是指由政府行政主管部门所属的法定计量检定机构或授权的计量检定机构, 对某些测量仪器实行的定点定期检定。强制检定的特点是, 由政府计量行政部门主管, 指定的法定或授权技术机构具体执行, 固定检定关系, 定点送检; 检定周期由执行强制检定的技术机构按照计量检定规程, 结合实际使用情况确定。

非强制检定是指由使用单位自己或委托具有社会公用计量标准或授权的计量检定机构, 对强制检定以外的其他测量仪器依法进行的一种定期检定。其特点是使用单位自主管理, 自由送检, 自求溯源, 自行确定检定周期。

强制检定和非强制检定均属于法制检定, 是我国对测量仪器依法管理的两种形式, 都



要受到法律的约束。不按规定进行周期检定的,要负法律责任。检定的依据是按照法定程序审批公布的计量检定规程。我国《计量法》规定:“计量检定必须按照国家计量检定系统表进行。国家计量检定系统表由国务院计量行政部门制定。计量检定必须执行计量检定规程。国家计量检定规程由国务院计量行政部门制定。没有国家计量检定规程的,由国务院有关主管部门和省、自治区、直辖市人民政府计量行政部门分别制定部门计量检定规程和地方计量检定规程,并向国务院计量行政部门备案。”因此,任何企业和其他实体是无权制定检定规程的。

在检定结果中,必须有合格与否的结论,并出具证书或加盖印记。从事检定工作的人员必须经考核合格,并持有有关计量行政部门颁发的检定员证。

随着改革开放和经济的发展,在强化检定法制性的同时,对大量非强制检定的测量仪器,为达到统一量值的目的,应以校准为主。过去,一直没有把校准作为实现单位统一和量值准确可靠的主要方式,而常用检定取而代之。这一观念目前正在改变中,校准在量值溯源中的位置已逐步确定。

2. 测量仪器的计量学特性

测量仪器的计量特性是指其影响测量结果的一些明显特征,其中包括测量范围、偏倚、重复性、稳定性、分辨力、鉴别力(阈)和示值误差等。为了达到测量的预定要求,测量仪器必须具有符合规定要求的计量学特性。

(1) 标称范围、量程和测量范围。测量仪器的操纵器件调到特定位置时可得到的示值范围,称为标称范围。此时的示值范围是与测量仪器的整体相联系的,是指标尺所指示的被测量值可得到的范围。标称范围通常以被测量的单位表示,而不管标尺上所标的单位是什么。例如,一台万用表,把操纵器件调到 $\times 10$ 一挡,其标尺上、下限的数码为 $0\sim 10$,则其标称范围为 $(0\sim 100)\text{V}$ 。标称范围一般用上限和下限说明,例如, $(100\sim 200)^\circ\text{C}$ 。当下限(即最小值)为零时,标称范围一般只用其上限(即最大值)来表示,例如, $(0\sim 100)\text{V}$ 的电压表,其标称范围可表示为 100V 。

标称范围的上限与下限之差的绝对值,称为量程。例如:某温度计的标称范围为 $(-30\sim 80)^\circ\text{C}$,则其量程为 $|80 - (-30)|^\circ\text{C} = 110^\circ\text{C}$;某电压表的标称范围为 100V ,则其量程为 $|100 - 0|\text{V} = 100\text{V}$ 。

测量范围,也称为工作范围,是指测量仪器的误差处于规定的极限范围内的被测量的示值范围。在这一规定的测量范围内使用,测量仪器的误差必定处在允许极限范围内;而若超出测量范围使用,示值误差就可能超出允许极限。换言之,测量范围就是在正常工作条件下,能确保测量仪器规定准确度的被测量值的范围。

有些测量仪器的测量范围与其标称范围相同。例如体温表、电流表、压力表、密度计等。而有的测量仪器处在下限时测量误差会急剧增大。例如地秤,这时应规定一个能确保其示值误差处在规定极限内的示值范围作为测量范围。

应正确区分和掌握示值范围、标称范围、测量范围和量程的概念。示值范围是指测量仪器标尺或显示装置所能指示的范围,可用标在标尺或显示器上的单位表示;标称范围是对测量仪器而言的,通常用被测量的单位表示;测量范围是指能保证规定的准确度、使误差处于规定极限内的量值范围;量程则是指标称范围上、下限之差的绝对值。

(2) 额定操作条件、极限条件和参考条件。额定操作条件是指测量仪器的正常工作条

件,也就是使测量仪器的规定计量特性处于给定极限内的使用条件。在这些条件中,一般包括被测量的范围或额定值,只有在规定的范围或额定值下使用,测量仪器才能达到规定的计量特性或规定的示值允许误差。例如:工作压力表测量范围的上限为10MPa,则压力的最大值只能加到10MPa;额定电流为10A的电流表,其输入电流不能超过10A。在使用测量仪器时,识别额定操作条件十分重要,它们是保证测量结果准确可靠的基本条件。

测量仪器的规定计量特性不受损也不降低,其后仍可以在额定条件下运行所能承受的极端条件,称为极限条件。

参考条件是指测量仪器在性能试验或进行检定、校准、对比时的使用条件,即标准工作条件,这些条件一般应对作用于测量仪器的影响量的参考值或参考范围作出明确规定,以真正反映测量仪器的计量性能和保证测量结果的可比性。

应注意正确区分和掌握额定操作条件、极限条件和参考条件。前者是测量仪器正常使用的条件,后者是确定测量仪器本身计量性能所规定的标准条件,中者则是测量仪器不受损坏和不降低准确度所允许的极端条件。在这三者中,参考条件的要求最严,额定操作条件则较宽,而极限条件的范围和额定值为最大。

(3) 示值误差和最大允许误差。示值就是由测量仪器所指示的被测量值。测量仪器的示值误差是测量仪器示值与对应的输入量的真值之差,它是测量仪器最主要的计量特性之一,反映了测量仪器准确度的大小。

示值误差是相对于真值而言的。由于真值不能确定,实际上使用的是约定真值或实际值。为确定测量仪器的示值误差,当接受高等级的测量标准对其进行检定或校准时,该测量仪器复现的量值即为约定真值,通常称为实际值、校准值或标准值。所以,指示式测量仪器的示值误差=示值-实际值,实物量具的示值误差=标称值-实际值。例如,被检测电流表的示值 I 为40A时,用标准电流表检定,其电流实际值为 $I_0=39$ A,则示值40A的误差 Δ 为:

$$\Delta = I - I_0 = (40 - 39) \text{ A} = 1 \text{ A}$$

又如,某工作玻璃量具的容量的标称值 V 为1 000ml,经标准玻璃量具检定,其容量实际值 V_0 为1 005ml,则量具的示值误差 Δ 为:

$$\Delta = V - V_0 = (1\,000 - 1\,005) \text{ ml} = -5 \text{ ml}$$

即该工作量具的标称值比其约定真值小5ml。

测量仪器的示值误差,通常简称为测量仪器的误差,可用绝对误差形式表示,也可以用相对误差的形式表示。确定测量仪器示值误差的大小,是为了判定测量仪器是否合格,并获得其示值的修正值。

对给定的测量仪器,由规范、规程等所允许的误差极限值,称为测量仪器的最大允许误差,有时也称为测量仪器的允许误差限,可简写为 mpe 。

应正确理解和区别测量仪器的示值误差和最大允许误差之间的关系。示值误差和最大允许误差均是对测量仪器本身而言的:最大允许误差是指技术规范(如标准、检定规程、校准规范)所规定的允许的误差极限值,它是一个判定测量仪器合格与否的规定的要求;而示值误差则是指测量仪器某一示值的误差的实际大小,它是通过检定、校准所得到的一个或一组值,用以评价测量仪器是否满足最大允许误差的要求,从而判断其是否合格,或者根据实际需要提供修正值,以提高测量结果的准确度。

5.5 过程能力分析

5.5.1 过程统计控制状态

任何一个过程都受到两类因素的影响，一类是人们无法控制或难以控制的随机因素 (random cause)，也称偶然因素。在随机因素作用下，导致过程输出的波动，称之为随机波动。由于这种波动的变化幅度较小，工程上是可以接受的。即使这种较小的随机波动，我们也不希望它存在，因为它毕竟会对最终产品的质量产生一定的影响。但是，我们又不能从根本上消除它，就不得不承认它存在的合理性。也就是说，随机因素存在于任何过程中是一种正常现象，从这种意义上讲，随机因素为固有因素或通常因素 (common cause)。由此，我们称仅有随机因素影响的过程为正常的过程。在随机因素影响过程的同时，还存在另外一类相对稳定的因素作用于过程，制约着过程的输出结果，这类相对稳定的因素称之为系统因素 (assignable cause)。正是由于这两类因素的共同作用，使得过程输出的结果呈现出内在的统计规律性。通过过程输出结果的规律性，可以探测过程是否处于控制状态，即系统因素是否发生变异。一旦系统因素发生变异，过程输出结果的规律将遭到破坏，从而过程失控，因此，这类因素又称为异常因素 (special cause)。

统计控制状态 (statistically in control) 是指所观察到的输出结果的波动可归因于只有随机因素影响的状态，而这种随机因素影响的结果通常表现为犹如来自同一总体的简单随机样本一样。受控过程 (process in control) 或稳定过程 (stable process) 是指其每一质量特性值 (如产品和服务的平均值和方差，不合格率或平均不合格数) 均处于统计控制状态的过程。

当我们说过程处于统计控制状态时，是指过程不存在异常因素。图 5—48 展示的是过程处于“不受控”和“受控”状态时的情况。

只有当过程处于统计控制状态时，对过程能力进行分析才是有意义的。如何判断过程处于统计过程控制状态，可参考第 8 章控制 (C) 阶段。

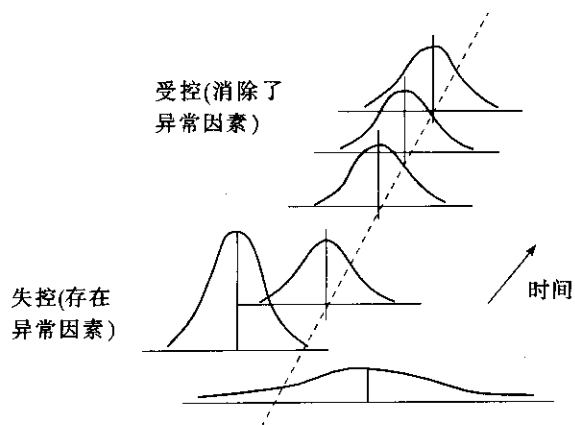


图 5—48 过程控制

当过程在统计意义上是受控的,它并不意味着过程产出的产品就不会超出规定的规格、符合质量要求。这种情况被认为是过程绩效问题,而不是过程控制问题。一个过程在统计意义上是受控的,但并不保证生产出的产品满足规格要求,如图5—49所示。出现这种情况是因为过程的均值过度偏离目标值,或者是因为过程的波动过大。我们可以通过研究过程绩效来解释这种现象。

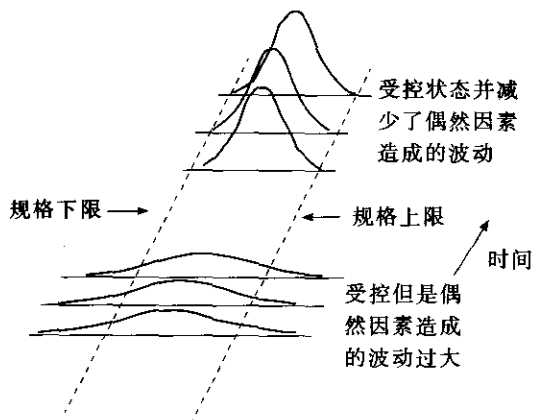


图5—49 过程能力

5.5.2 过程能力和过程绩效

过程能力和过程绩效分析是评价过程满足预期要求的能力及其表现的方法。过程能力分析在DMAIC项目中十分重要,它是评价过程基线及改进方向和目标的重要工具。因此,过程能力分析是测量阶段的一项重要工作。

在着手过程能力分析时,必须明确如下要素:

- (1) 过程输出特性。这是项目工作从界定阶段就已明确的。
- (2) 对过程输出特性的要求。在进行过程能力分析时,必须识别并明确顾客(内部的或外部的)对过程输出特性的要求,包括目标值和规格限或容限。通常将规格上、下限记为USL和LSL。对制造过程来说,识别目标值和规格限是比较容易的。因为工程上对此一般都有明确的规定。但对非制造类的过程来说,需要项目团队投入精力识别并明确这些要求。

(3) 抽样方案。不同的抽样方案,反映了过程的不同情况和状态。比如,在研究过程的短期能力时,抽取的样本应尽可能仅受到随机因素的影响。

(4) 过程是否稳定或具有可预测的分布。过程能力分析的假设前提是输出服从正态分布。因此,过程应是稳定或统计受控的。对那些非正态分布的情况,应进行适当的坐标变换,将其转换为正态分布的情况。

过程能力分析就是在确认上述要素的基础上,运用统计工具展开的。为此,我们定义下述概念:

- 过程短期波动 (inherent process variation), 也称样本内的波动, 是仅由短期内随机因素影响而产生的过程波动。这部分波动可以通过计算样本内部的极差 R_i 或标准差 s_i ,

进一步求出平均的极差 \bar{R} 或综合标准差 (pooled standard deviation) s , 利用 \bar{R}/d_2 或 s/c_4 估计过程短期波动的标准差 σ_{within} 。

● 过程的总波动, 是由随机因素和系统因素影响而产生的波动。它可以由所有样本标准差 s 估计长期的总波动的标准差 $\sigma_{overall}$ 。这里 $s = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)}$, x_i 是各单个测量值, \bar{x} 是所有测量值的均值, n 是所有测量值的总个数。

- 过程能力 PC , 是过程固有波动的 $6\sigma_{within}$ 范围。
- 过程绩效 PP , 是过程总波动的 $6\sigma_{overall}$ 范围。

5.5.3 过程能力指数 C_p 与 C_{pk}

1. 过程能力指数 C_p 的意义与计算

假定过程输出特性 $y \sim N(\mu, \sigma^2)$, 考察正态分布上的 3 个典型区间 $[\mu - \sigma, \mu + \sigma]$, $[\mu - 2\sigma, \mu + 2\sigma]$, $[\mu - 3\sigma, \mu + 3\sigma]$, 它们有这样一些特点 (见图 5—50):

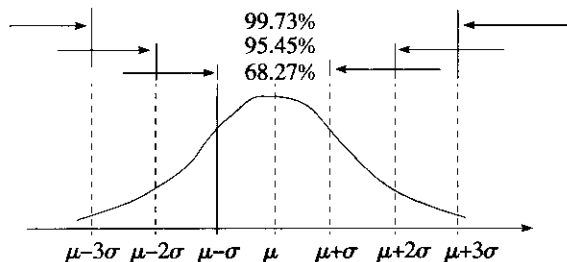


图 5—50 正态概率分布的特点

过程输出的特性 Y 落入上述 3 个区间的概率分别为:

$$P(\mu - \sigma \leq Y \leq \mu + \sigma) = 68.27\%$$

$$P(\mu - 2\sigma \leq Y \leq \mu + 2\sigma) = 95.45\%$$

$$P(\mu - 3\sigma \leq Y \leq \mu + 3\sigma) = 99.73\%$$

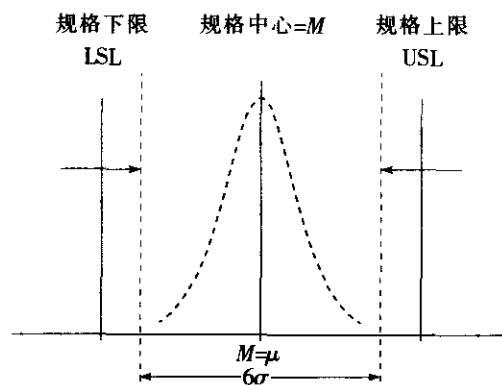
也就是说, 过程输出特性落到 $[\mu - 3\sigma, \mu + 3\sigma]$ 区间之外的概率仅为 0.27%, 是小概率事件。因此, 我们将 6σ 范围视为过程的自然波动范围。为了把过程的自然输出能力与要求的容差进行比较, 著名质量专家朱兰引入了能力比的概念, 即过程能力指数 C_p 。

如图 5—51 所示, 若过程输出 Y 服从正态分布, 即 $y \sim N(\mu, \sigma^2)$, 其中 μ, σ 分别为 y 的均值和标准差。当过程处于统计控制状态时, 则定义过程能力指数 C_p 为容差与过程波动范围之比。

$$C_p = \frac{\text{容差}}{\text{过程能力}} = \frac{USL - LSL}{6\sigma} = \frac{T}{6\sigma} \quad (5-10)$$

式中, USL, LSL 分别是质量特性的上、下规格限。 $M = \frac{1}{2} (LSL + USL)$, 称为规格中心。容差 $T = USL - LSL$, 反映了对过程的要求。在这个定义中, 容差 T 一般不能轻易改变, 因此, σ 越小, C_p 值越大。

C_p 是我们描述过程能力的最重要指标。过去通常我们称 $C_p < 1$ 时, 过程能力不足; $1 \leq$

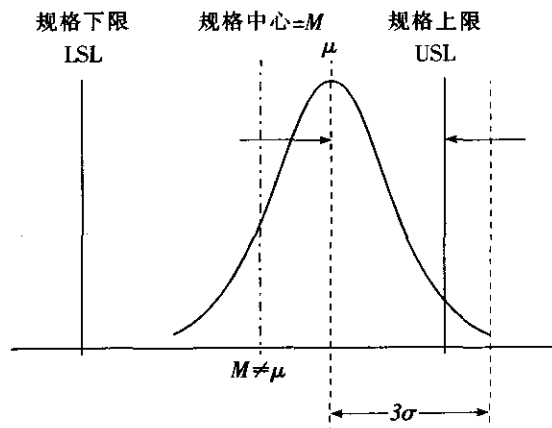
图 5—51 过程能力指数 C_p 示意图

$C_p < 1.33$ 时, 过程能力尚可; $1.33 \leq C_p < 1.67$ 时, 过程能力充足。但在经济繁荣、产品数量急剧增长、高科技蓬勃发展的今天, 原来的质量标准已不能适应当前的需求。若过程质量水平达到六西格玛, 就是要求过程能力指数 $C_p = 2$, 即要求 $USL - LSL = 12\sigma$ 。

需要注意的是, C_p 的计算与过程输出的均值无关, 它是假定过程输出的均值与规格中心重合时的过程能力之比。因此, C_p 指数只是反映了过程的潜在能力, 当我们设法把过程输出均值逐渐移向规格中心时, 这种潜力便会得到充分体现。所以在一般场合下, C_p 指数称为潜在过程能力指数。当 $\mu \neq M$ 时, 过程输出的不合格品率将增加。这就造成了尽管 C_p 值较大, 但不合格品率仍很高的情况, 因此要引入另一个过程能力指数 C_{pk} 。

2. 过程能力指数 C_{pk} 的意义与计算

对大多数情况来说, 过程输出的均值 μ 不会恰与规格中心或目标值重合。因此, 在进行过程能力分析时, 应当将均值 μ 的影响考虑进来。引入过程能力指数 C_{pk} 就是为了解决这个问题。由于过程中心 μ 通常在规格限 (LSL, USL) 之间, 因此, 用过程中心 μ 与两个规格限最近的距离 $\min \{USL - \mu, \mu - LSL\}$ 与 3σ 之比作为过程能力指数, 记为 C_{pk} , 如图 5—52 所示。

图 5—52 过程能力指数 C_{pk} 示意图

其计算公式为:



$$C_{pk} = \min \left\{ \frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma} \right\} \quad (5-11)$$

$$= \min (C_{pu}, C_{pl})$$

式中, $C_{pu} = \frac{USL - \mu}{3\sigma}$, 称为单侧上限过程能力指数, 仅有上规格限场合即可使用; $C_{pl} = \frac{\mu - LSL}{3\sigma}$ 称为单侧下限过程能力指数, 仅有下规格限的场合即可使用。

由 C_{pk} 的上述表达式可看出, 当 $\mu = M$ 时, $C_p = C_{pk}$; 当 $\mu \neq M$ 时, $C_{pk} < C_p$ 。只要双侧规格限都给定, C_p 有意义, 则应同时考虑 C_p 及 C_{pk} 两个指数, 以便对整个过程的状况有较全面的了解。例如, 当 C_p 及 C_{pk} 都较小而二者差别不大时 (比如 $C_p = 0.72$ 及 $C_{pk} = 0.69$), 说明过程的主要问题是 σ 太大, 改进过程应首先着眼于降低过程的波动。若 C_p 较大, 而 C_{pk} 很小 (比如 $C_p = 1.43$ 及 $C_{pk} = 0.72$), 二者差别较大, 说明过程的主要问题是 μ 偏离 M 太多, 改进过程应首先着眼于移动 μ 值, 使之更接近于 M 。如果 C_p 本身不够好, C_{pk} 更小 (比如 $C_p = 0.84$ 及 $C_{pk} = 0.35$), 而且二者差别较大时, 说明过程的 σ 及 μ 都有问题, 通常改进过程应首先移动 μ 值, 使之更接近于 M , 然后设法降低过程的波动。总之, 不要只单独使用这两个过程能力指数之中的一个。

当 C_p 值一定时, C_{pk} 将随着过程输出中心 μ 与规格中心 M 偏离的增大而减小。事实上, 利用 $T = USL - LSL$, $M = \frac{USL + LSL}{2}$, 则 C_{pk} 可表示为另一种形式:

$$C_{pk} = \frac{T}{6\sigma} - \frac{|M - \mu|}{3\sigma} \quad (5-12)$$

若对上式第二项的分子、分母分别乘以 $\frac{T}{2}$, 可得到 C_{pk} 的第三种形式:

$$C_{pk} = (1 - K) C_p, \quad K = \frac{|M - \mu|}{\frac{T}{2}} = \frac{2|M - \mu|}{T} \quad (5-13)$$

式中, $K > 0$ 称为偏离度。

需要特别强调的是, C_p 和 C_{pk} 是由处于统计受控状态下的过程波动的大小和均值偏离决定的。因而首先要判明过程是否处于统计受控状态, 这需要通过控制图等统计工具进行, 关于如何应用控制图分析判断过程的受控状态, 请参见本书第 8 章的内容。

例 5-30

某产品的关键加工尺寸要求为 $\Phi = 6.00 \pm 0.02 \text{mm}$ 。为了分析该加工过程的过程能力, 项目团队跟踪收集了一些数据 (见表 5-25)。试估算该加工过程的 C_p 和 C_{pk} 。

表 5-25

例 5-30 数据表

单位: mm

样本 序号	测量值			\bar{x}	R	样本 序号	测量值			\bar{x}	R
	x_1	x_2	x_3				x_1	x_2	x_3		
1	6.028	6.003	6.020	6.017 00	0.025	8	5.978	5.980	5.994	5.984 00	0.016
2	6.014	5.994	6.008	6.005 33	0.020	9	6.012	5.998	5.982	5.997 33	0.030
3	6.002	5.983	6.014	5.999 67	0.031	10	6.008	6.002	5.984	5.998 00	0.024
4	6.012	5.982	6.036	6.010 00	0.054	11	5.968	5.986	5.988	5.980 67	0.020
5	6.024	6.002	6.008	6.011 33	0.022	12	6.014	6.000	6.008	6.007 33	0.014
6	6.022	5.998	6.008	6.009 33	0.024	13	6.034	6.006	6.028	6.022 67	0.028
7	6.014	5.991	6.000	6.001 67	0.023	14	6.002	5.988	6.008	5.999 33	0.020

续前表

样本 序号	测量值			\bar{x}	R	样本 序号	测量值			\bar{x}	R
	x_1	x_2	x_3				x_1	x_2	x_3		
15	6.012	5.982	6.036	6.010 00	0.054	20	6.014	5.994	6.008	6.005 33	0.020
16	5.990	5.978	5.980	5.982 67	0.012	21	5.988	5.988	5.994	5.990 00	0.006
17	6.016	5.992	6.004	6.004 00	0.024	22	6.000	6.002	6.008	6.003 33	0.008
18	6.014	5.992	5.998	6.001 33	0.022	23	6.036	6.008	6.024	6.022 67	0.028
19	6.032	6.008	6.018	6.019 33	0.024	24	6.010	5.998	6.000	6.002 67	0.012
										$\bar{\bar{x}}$	6.003 5
										R	0.023 375

(1) 过程能力指数 C_p 的计算。根据表 5—25 的数据可作均值—极差控制图，判断过程是否统计受控。本例的均值—极差控制图显示该过程处于统计受控状态（分析过程略，关于均值—极差控制图请参见本书第 8 章的内容）。因此，过程固有波动的 σ 可由下式估计：

$$\hat{\sigma} = \bar{R}/d_2 = \frac{0.023\ 375}{1.693} = 0.013\ 8$$

式中， d_2 是控制图系数，它与样本组的容量 n 有关。在本例中， $n=3$ ，查附表 6 控制图系数表， $d_2=1.693$ 。将 $\hat{\sigma}$ 代入式 (5—10) 可得

$$C_p = \frac{USL - LSL}{6\sigma} = \frac{T}{6\sigma} = \frac{0.04}{6 \times 0.013\ 8} = 0.48$$

(2) 过程能力指数 C_{pk} 的计算。此时过程输出的均值为 $\bar{x}=6.003\ 5$ ，由式 (5—11) 可得

$$\begin{aligned} C_{pk} &= \frac{\min(USL - \mu, \mu - LSL)}{3\sigma} \\ &= \frac{\min(6.02 - 6.003\ 5, 6.003\ 5 - 5.98)}{3 \times 0.013\ 8} = \frac{0.016\ 5}{3 \times 0.013\ 8} = 0.40 \end{aligned}$$

用 MINITAB 软件，在用控制图验证了过程统计受控且验证了数据服从正态分布后，可以计算各项过程能力指数。从 STAT—Quality Tools—Capability Analysis (Normal) 入口，填入变量名称，指明 3 个数据为一组，填写 USL 及 LSL 后即可得到图 5—53 中的结果。

5.5.4 过程能力指数 C_{pm} 和 C_{pmk}

为了强调质量特性偏离目标值造成的质量损失，当把目标值 m 引入过程能力指数中时，就得到两个新的过程能力指数 C_{pm} 和 C_{pmk} 。

$$C_{pm} = \frac{USL - LSL}{6\sigma'} \quad (5-14)$$

式中， $\sigma'^2 = E[y - m]^2 = \sigma^2 + (\mu - m)^2$ 。

$$C_{pmk} = \frac{C_{pk}}{\sqrt{1 + \left(\frac{\mu - m}{\sigma}\right)^2}} \quad (5-15)$$

式中， $\sigma^2 = \text{var}(y)$ ； $\mu = E(y)$ ； m 为质量特性的目标值。

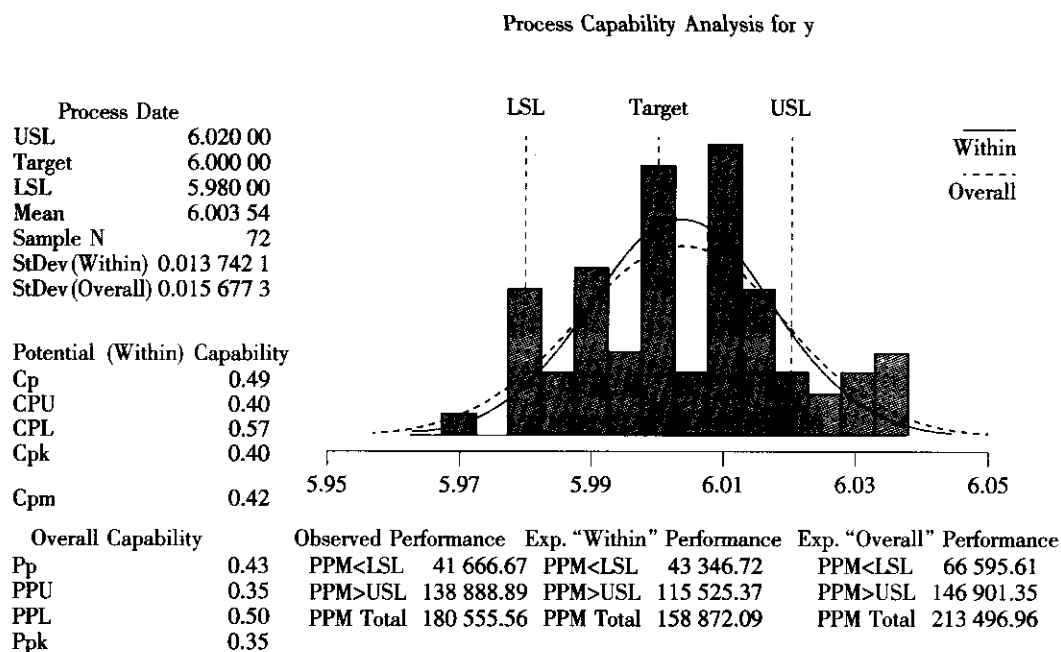


图 5—53 过程能力分析

质量特性 Y 偏离其目标值 m 而导致的损失, 通常被认为近似于对称的平方误差损失函数, 即质量损失函数。从式 (5—15) 可知, 当 $\mu = m$ 且 m 与规格中心重合时, $C_{pm} = C_p$ 。因此, C_{pm} 指数可以看作 C_p 指数的推广。同时使用 C_p 及 C_{pm} 可以更全面地分析过程的能力。有时也称 C_{pm} 为第二代过程能力指数。 C_{pm} 指数反映了过程输出 μ 和目标值 m 之间的偏差, 主要用于反映过程的期望损失, 与田口的质量观相一致, 体现了田口的质量损失函数的理念, 而且在目标值并不恰好为公差限中心时, 比 C_{pk} 能更好地体现过程输出 μ 和目标值 m (而不是公差限中心 M) 之间的偏差。

C_{pmk} 也称为混合能力指数, 它是在 C_{pk} , C_{pm} 能力指数的基础上, 为了更加灵敏地反映过程输出均值 μ 与目标值 m 之间偏差的基础上而提出的。它强调了向目标值靠近的重要性。强调了目标值, 淡化了规格限。

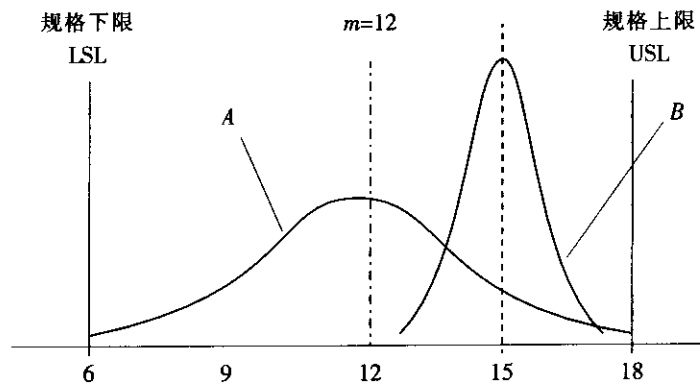


图 5—54 对两种不同输出的过程能力评价

假设过程的输出如图 5—54 所示, 图中分布 A 的均值 $\mu=12$, 标准差 $\sigma=2$; 分布 B 的均值 $\mu=15$, 标准差 $\sigma=0.57$ 。当用 C_{pk} 评价它们的优劣时, 会得到如下结果:

$$A: C_{pk} = \min \left\{ \frac{12-6}{3 \times 2}, \frac{18-12}{3 \times 2} \right\} = 1.00$$

$$B: C_{pk} = \min \left\{ \frac{18-15}{3 \times 0.57}, \frac{15-6}{3 \times 0.57} \right\} = 1.75$$

显然, B 比 A 要好, 因为 B 的 C_{pk} 要高于 A 的 C_{pk} , B 的合格品率要高于 A。但事实上, A 的优等品率比 B 的多, 因为大多数 A 的输出在目标值 m 附近。因此对 A 满意的顾客要比 B 多。如果我们用 C_{pm} 来评价的话, 可以得到这样的结论:

$$A: C_{pm} = \frac{18-6}{6 \times \sqrt{2^2 + (12-12)^2}} = 1.00$$

$$B: C_{pm} = \frac{18-6}{6 \times \sqrt{0.57^2 + (15-12)^2}} = 0.655$$

因此, A 达到目标值的能力比 B 要高。对于关心合格品率的企业, 无疑会认为过程 B 的能力较好, 而对于注重达到目标值的企业, 则会认为过程 A 的能力较好。值得注意的是, 在市场竞争日益激烈的今天, 企业越来越关注顾客要求的目标值, 关注提高优等品率, 因此, C_{pm} 和 C_{pmk} 的评价方法, 更能反映他们的要求。

5.5.5 过程绩效指数 P_p 与 P_{pk}

过程绩效指数是从过程总波动的角度考察过程输出满足顾客要求的能力。有时也将其称为长期过程能力指数。在过程输出的总波动中, 既包含了过程固有波动, 也包含了过程受到其他因素影响而产生的波动。在考察过程绩效时, 不要求过程稳定, 即不要求过程输出的质量特性 Y 一定服从某个正态分布。因为过程在较长时期内产生的数据很难保证具有正态性。很多波动源在短期观察中可能不会出现, 或很少出现。而经长期收集到的数据则会包含它们。例如: 机器性能的漂移或老化、不同操作者之间的技术差异、设备的调整、仪表的校准、更换材料批次或供应商的变动、环境因素的变化等。

计算过程绩效指数 P_p 和 P_{pk} 的方法与计算 C_p 和 C_{pk} 的方法类似。所不同的是, 它们是规格限与过程总波动的比值。这里, 过程总波动通常由标准差 s 来估计, 而标准差 $s =$

$\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)}$, x_i 是各单个测量值, \bar{x} 是所有测量值的均值, n 是所有测量值的总个数。对照过程能力指数 C_p 和 C_{pk} , 它们是通过样本组内波动来估算 σ 的, 通常取 $\hat{\sigma} = \bar{R}/d_2$ 。图 5—55 所示的是这两种不同波动的意义。

过程绩效指数也有多种。它们分别与过程能力指数对应, 差别仅在标准差 σ 上。它们的计算公式是:

- 潜在过程绩效指数:

$$P_p = \frac{USL - LSL}{6s} \quad (5-16)$$

- 单侧上限过程绩效指数:

$$P_{pu} = \frac{USL - \bar{x}}{3s} \quad (5-17)$$

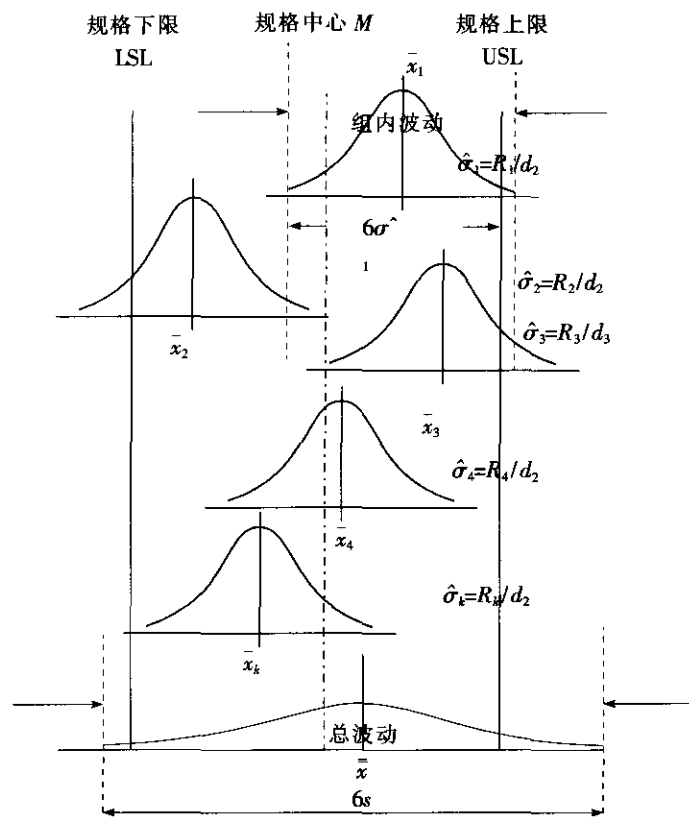


图 5—55 总波动与样本组内波动示意图

- 单侧下限过程绩效指数:

$$P_{pk} = \frac{\bar{x} - LSL}{3s} \quad (5-18)$$

- 实际过程绩效指数:

$$P_{pk} = \min \{P_{pu}, P_{pk}\} \quad (5-19)$$

对例 5—30 的过程, 数据如表 5—25 所示, 计算其绩效指数 P_p 和 P_{pk} 。

(1) 过程绩效指数 P_p 的计算。对表 5—25 的所有数据求均值和标准差, 可得到以下结果:

$$\bar{x} = 6.0035$$

$$s = \sqrt{\sum (x_i - \bar{x})^2 / (n - 1)} = 0.015622$$

将上述结果代入式 (5—16), 得到

$$P_p = \frac{USL - LSL}{6s} = \frac{0.04}{6 \times 0.015622} = 0.43$$

(2) 过程绩效指数 P_{pk} 的计算。由式 (5—17)、式 (5—18)、式 (5—19) 可得到

$$P_{pu} = \frac{USL - \bar{x}}{3s} = \frac{6.02 - 6.0035}{3 \times 0.015622} = 0.35$$

$$P_{pl} = \frac{\bar{x} - LSL}{3s} = \frac{6.0035 - 5.98}{3 \times 0.015622} = 0.50$$

$$P_{pk} = \min \{P_{pu}, P_{pl}\} = \min \{0.35, 0.50\} = 0.35$$

一般来说, 过程绩效指数值要比相应的过程能力指数值小。

5.5.6 过程能力指数与缺陷率的关系

当受控过程的质量特性 y 服从正态分布 $N(\mu, \sigma^2)$ 时, 其不合格品率为:

$$\begin{aligned} p(d) &= P_L + P_U = P(y < LSL) + P(y > USL) \\ &= \Phi\left(\frac{LSL - \mu}{\sigma}\right) + \left[1 - \Phi\left(\frac{USL - \mu}{\sigma}\right)\right] \end{aligned} \quad (5-20)$$

式中, Φ 为标准正态分布的概率分布函数。

假如过程中心 μ 位于规格中心 M 与上规格限 USL 之间, 即 $M \leq \mu \leq USL$, 则

$$C_{pk} = C_{pu} = \frac{USL - \mu}{3\sigma}$$

另外

$$\frac{LSL - \mu}{3\sigma} = \frac{(USL - \mu) - (USL - LSL)}{3\sigma} = C_{pk} - 2C_p$$

把以上两式代入不合格品率 p 的表达式, 可得

$$\begin{aligned} p(d) &= \Phi[3(C_{pk} - 2C_p)] + 1 - \Phi(3C_{pk}) \\ &= \Phi[-3(2C_p - C_{pk})] + \Phi(-3C_{pk}) \end{aligned} \quad (5-21)$$

当 $LSL \leq \mu \leq M$ 时, 亦可类似推得

$$p(d) = \Phi[-3(1+K)C_p] + \Phi[-3(1-K)C_p] \quad (5-22)$$

由上式可见, 仅有 C_{pk} 还不足以确定不合格品率 $p(d)$, 事实上, 假设一个过程输出均值 μ 位于规格中心, 即 $\mu = \frac{USL + LSL}{2}$, 且 $\sigma = \frac{USL - LSL}{8}$, 则 $C_{pk} = 1.333$, 在正态分布

下, 每百万个零件的不合格数 $DPMO = 53ppm$; 而另一个过程中, $\sigma = \frac{USL - LSL}{9}$, 而 $\mu = \frac{USL + LSL}{2} + 0.5 \frac{USL - LSL}{9}$, 在这种情况下, 过程输出中心并不在规格中心, 而是有

0.5σ 的偏差, 同样得到 $C_{pk} = 1.333$, 但在正态分布下, 每百万个零件的不合格品数却为 $DPMO = 32ppm$ 。从这个例子中, 我们可以看到: 两个过程具有相同的 C_{pk} 值, 但却有不同的不合格品数, 但由 C_{pk} 可以确定不合格品率 $p(d)$ 的一个范围。

$$\Phi(-3C_{pk}) \leq p \leq 2\Phi(-3C_{pk}) \quad (5-23)$$

若某过程的 $C_{pk} = 1$, 则知不合格品率 $p(d)$ 介于 $\Phi(-3)$ 和 $2\Phi(-3)$ 之间, 又 $\Phi(-3) = 0.00135$, 所以 $0.00135 \leq p(d) \leq 0.0027$, 若又知 $C_p = 1.1$, 则

$$\begin{aligned} p(d) &= \Phi(-3(2 \times 1.1 - 1)) + \Phi(-3) = \Phi(-3.6) + \Phi(-3) \\ &= 0.00016 + 0.00135 = 0.00151 \end{aligned}$$

在六西格玛管理活动中,除了使用上述过程能力的评价和表达方式外,为了和属性值数据进行横向比较,可以使用西格玛水平 Z_{bench} 来评价过程能力。西格玛水平 Z_{bench} 就是从过程均值到技术规格界限的距离内所含的标准差的数目。应用西格玛水平 Z_{bench} 来评价过程能力的优点是,它与过程的不合格品率 $p(d)$ 或 DPMO 是一一对应的。一般我们所说的西格玛水平考虑了过程的长期波动,西格玛水平等于 $1.5 + Z_{\text{bench}}$ 。西格玛水平 Z_{bench} 有以下一些表达方式:

- 仅有单侧上规格限时:

$$Z_{\text{bench}} = \frac{USL - \mu}{\sigma} \quad (5-24)$$

- 仅有单侧下规格限时:

$$Z_{\text{bench}} = \frac{\mu - LSL}{\sigma} \quad (5-25)$$

- 双侧规格限时:

$$Z_{\text{USL}} = \frac{USL - \mu}{\sigma}, \quad Z_{\text{LSL}} = \frac{\mu - LSL}{\sigma} \quad (5-26)$$

此时,双侧规格限下,综合的西格玛水平 Z_{bench} 还需通过总缺陷率进行折算。

Z 正是标准正态分布中对应分位点,如图 5-56 所示。通过查标准正态分布表,可以得到此 Z 变量下的合格率或缺陷率。

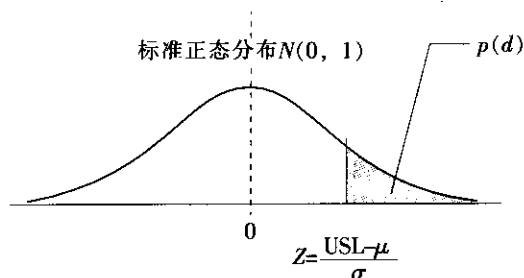


图 5-56 Z 变量与缺陷率一一对应

对例 5-30 的过程,数据如表 5-25 所示,计算其变量 Z_{USL} 和 Z_{LSL} ,以及对应的总缺陷率。

- (1) 计算 Z_{USL} 和 Z_{LSL} :

根据式 (5-26) 可得

$$\begin{aligned} Z_{\text{USL}} &= \frac{USL - \mu}{\sigma} = \frac{6.02 - 6.0035}{0.0138} = 1.196 \\ Z_{\text{LSL}} &= \frac{LSL - \mu}{\sigma} = \frac{5.98 - 6.0035}{0.0138} = -1.702 \end{aligned}$$

(2) 计算总缺陷率:

$$ppm_{USL} = [1 - \Phi(Z_{USL})] \times 10^6 = 0.115\,070 \times 10^6 = 115\,070$$

$$ppm_{LSL} = \Phi(Z_{LSL}) \times 10^6 = 0.044\,565 \times 10^6 = 44\,565$$

$$ppm_{Total} = 115\,070 + 44\,565 = 159\,635$$

5.5.7 长期能力与短期能力

所谓过程的短期能力是指过程仅受到随机因素的影响时,过程输出特性波动的大小。它是过程的固有能。而长期能力是指过程在较长的时期内所表现出的过程输出波动的大小,过程不仅受到随机因素的影响,而且受到其他因素的影响。由于短期能力仅含随机波动的影响,所以短期标准差 σ_{within} 较小;而长期能力不仅含随机因素的影响,还可能含有异常因素的影响,所以长期标准差 $\sigma_{overall}$ 较大。

在许多情况下,通过对过程的长期与短期能力进行具体分析和测算,可能获得更准确和有效的信息。

有许多计算和分析过程长期能力和短期能力的方法,这里仅结合下例做一简单介绍。

例 5-34

对例 5-30 的过程,数据如表 5-25 所示,计算其长期能力与短期能力。

解: (1) 计算该过程短期的总缺陷率 ppm_s 。一般来说,过程能力指数 C_p 和 C_{pk} 表示的是过程的短期能力指数,因为估计过程能力指数 C_p 和 C_{pk} 时,要求过程是稳定的,只受到随机因素的影响。

在前面关于 C_p 与 C_{pk} 的计算中,我们已经求得 $\hat{\sigma}_{within}$ 为:

$$\hat{\sigma}_{within} = \bar{R}/d_2 = \frac{0.023\,375}{1.693} = 0.013\,8$$

根据式 (5-26) 可得

$$Z_{USL} = \frac{USL - \mu}{\hat{\sigma}_{within}} = \frac{6.02 - 6.003\,5}{0.013\,8} = 1.196$$

$$Z_{LSL} = \frac{LSL - \mu}{\hat{\sigma}_{within}} = \frac{5.98 - 6.003\,5}{0.013\,8} = -1.702$$

分别计算缺陷率:

$$ppm_{USL} = [1 - \Phi(Z_{USL})] \times 10^6 = 0.115\,070 \times 10^6 = 115\,070$$

$$ppm_{LSL} = \Phi(Z_{LSL}) \times 10^6 = 0.044\,565 \times 10^6 = 44\,565$$

$$ppm_{Total} = 115\,070 + 44\,565 = 159\,635$$

所以该过程的总缺陷率为:

$$ppm_s = 115\,070 + 44\,565 = 159\,635$$

(2) 计算该过程长期的总缺陷率 ppm_L 。与过程的绩效指数 P_p 与 P_{pk} 对应的是过程的长期能力。在前面关于 P_p 与 P_{pk} 的计算中,我们得到 $\hat{\sigma}_{overall}$ 为:

$$\hat{\sigma}_{overall} = s = \sqrt{\sum (x_i - \bar{x})^2 / (n-1)} = 0.015\,622$$

所以,根据式 (5-26) 可得



$$Z_{USL} = \frac{USL - \mu}{\sigma_{overall}} = \frac{6.02 - 6.0035}{0.015622} = 1.056$$

$$Z_{LSL} = \frac{LSL - \mu}{\sigma_{overall}} = \frac{5.98 - 6.0035}{0.015622} = -1.504$$

分别计算缺陷率:

$$ppm_{USL} = [1 - \Phi(Z_{USL})] \times 10^6 = 0.146859 \times 10^6 = 146859$$

$$ppm_{LSL} = \Phi(Z_{LSL}) \times 10^6 = 0.066827 \times 10^6 = 66827$$

$$ppm_{Total} = 146859 + 66827 = 213686$$

所以该过程的总缺陷率为:

$$ppm_u = ppm_{Total} = 146859 + 66827 = 213686$$

5.5.8 非正态数据的变换

1. 偏态分布及其产生的原因

对许多过程来说, 总体分布呈现出不对称形态, 这种不对称性无论表现为左偏态, 还是表现为右偏态 (如图 5—57) 都是常见的。出现偏态分布的原因是多方面的, 当出现偏态时, 通常采用各种变换, 将其转换成近似正态分布。

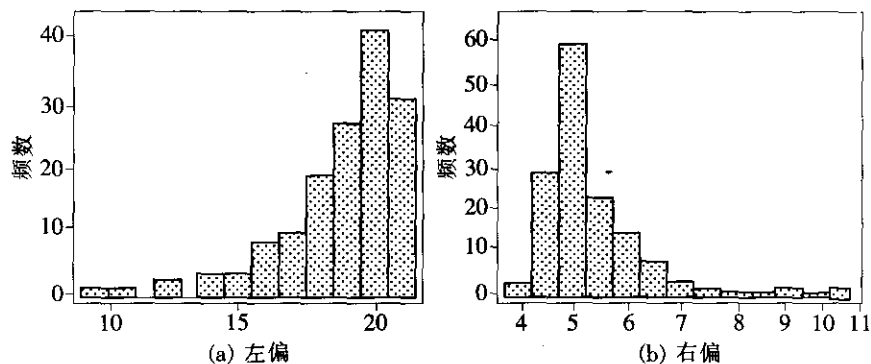


图 5—57 分布的尾部在左边 (右边) 称为左 (右) 偏

2. 数据变换

对数据作变换会改变直方图的形状, 这意味着, 通过适当的变换可使偏态分布转化为正态分布。

元素铍的照射会引起动物的细胞分裂, 从而引发身体损伤。细胞分裂时间 (IDT) 是重要指标, 现记录 40 个细胞的分裂时间 IDT (见表 5—26), 由这些数据构建的频率直方图如图 5—58 (a) 所示, 从图上看是偏态分布, 若对每个 IDT 取对数 (见表 5—26) 再作直方图, 则新的直方图就近似于正态分布 (见图 5—58 (b))。

表 5—26

细胞分裂 IDT 及其对数

序号	IDT	$\text{Log}_{10}(\text{IDT})$	序号	IDT	$\text{Log}_{10}(\text{IDT})$	序号	IDT	$\text{Log}_{10}(\text{IDT})$
1	28.1	1.45	15	50.1	1.78	29	21.0	1.32
2	31.2	1.49	16	23.7	1.37	30	22.3	1.35
3	13.7	1.14	17	18.5	1.27	31	15.5	1.19
4	45.0	1.55	18	21.4	1.33	32	35.3	1.55
5	25.8	1.41	19	25.5	1.42	33	19.1	1.28
6	15.8	1.23	20	25.2	1.42	34	38.4	1.58
7	34.8	1.54	21	32.0	1.51	35	72.8	1.85
8	52.3	1.79	22	43.5	1.54	36	48.9	1.59
9	28.0	1.45	23	17.4	1.24	37	21.4	1.33
10	17.9	1.25	24	38.8	1.59	38	20.7	1.32
11	19.5	1.29	25	30.5	1.49	39	57.3	1.75
12	21.1	1.32	26	55.5	1.75	40	40.9	1.51
13	31.9	1.50	27	25.5	1.41			
14	28.9	1.45	28	52.1	1.72			

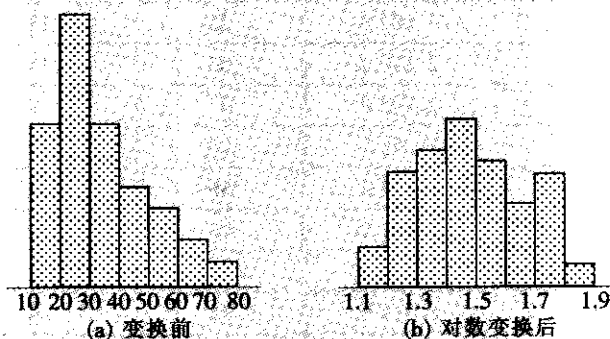


图 5—58 细胞分裂时间 (IDT) 直方图

这个例子显示, 变换可以改变频率直方图。这就告诉我们, 当遇到偏态分布时, 通过适当的变换, 就可以使变换后的频率直方图近似于中间高、两边低、左右对称。

在确认样本来自非正态分布后, 对数据作变换后若分布近似于钟形曲线, 则可以认为变换后的数据来自于某一正态总体, 常用的变换有:

$$y = \ln x, y = x^{-1}, y = x^2, y = \sqrt{x}$$

若变换后的数据服从正态分布, 可利用前边各节公式估计其过程能力指数和不合格品率。

例 5—36

从一个总体中抽取了 50 个样本见表 5—27, 给定上公差限为 100, 求 C_{pk} 。

绘制直方图 (见图 5—59 (a))。我们知道, 平方根变换能够把右偏的分布转化为正态, 然而, 该分布是高度右偏的, 经过平方根变换后仍表现为右偏。事实上, 我们应该采用 $y = x^{\frac{1}{r}}$ ($r > 2$) 变换, 但为了方便和准确, 我们采用了变换 $y = x^{\frac{1}{4}}$ 。图 5—59 (b) 给出数据变换后的直方图。变换后的直方图接近于单峰, 钟形曲线, 意味着是一个近似

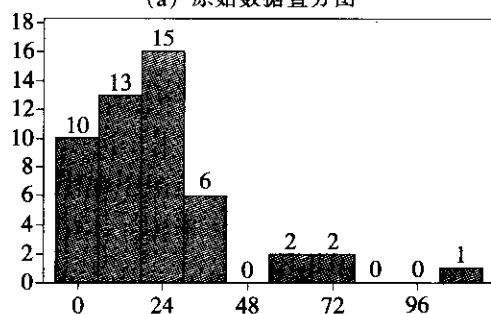
的正态分布。

表 5—27

抽取的 50 个样本及变换后的数值

序号	观测值	$X^{1/4}$	序号	观测值	$X^{1/4}$
1	18	2.059 8	26	13	1.898 8
2	1	1.000 0	27	4	1.414 2
3	19	2.087 8	28	54	2.828 4
4	32	2.378 4	29	11	1.821 2
5	28	2.300 3	30	12	1.851 2
6	8	1.581 8	31	2	1.189 2
7	4	1.414 2	32	21	2.140 7
8	27	2.279 5	33	34	2.414 7
9	72	2.913 0	34	15	1.958 0
10	8	1.581 8	35	19	2.087 8
11	41	2.530 4	36	10	1.778 3
12	75	2.942 8	37	5	1.555 1
13	29	2.320 5	38	7	1.525 5
14	4	1.414 2	39	107	3.215 2
15	11	1.821 2	40	29	2.320 5
16	30	2.340 3	41	4	1.414 2
17	32	2.378 4	42	35	2.449 5
18	15	1.958 0	43	1	1.000 0
19	29	2.320 5	44	20	2.114 7
20	10	1.778 3	45	23	2.189 9
21	25	2.235 1	46	28	2.300 3
22	4	1.414 2	47	24	2.213 4
23	10	1.778 3	48	55	2.723 3
24	20	2.114 7	49	7	1.525 5
25	19	2.087 8	50	3	1.315 1

(a) 原始数据直方图



(b) 数据变换后的直方图

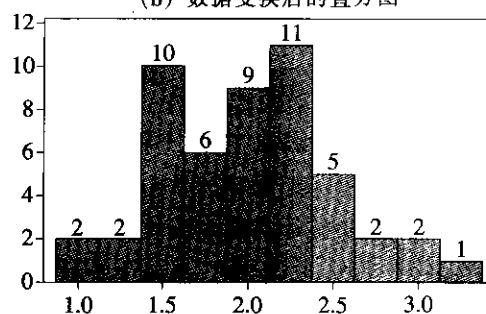


图 5—59 直方图

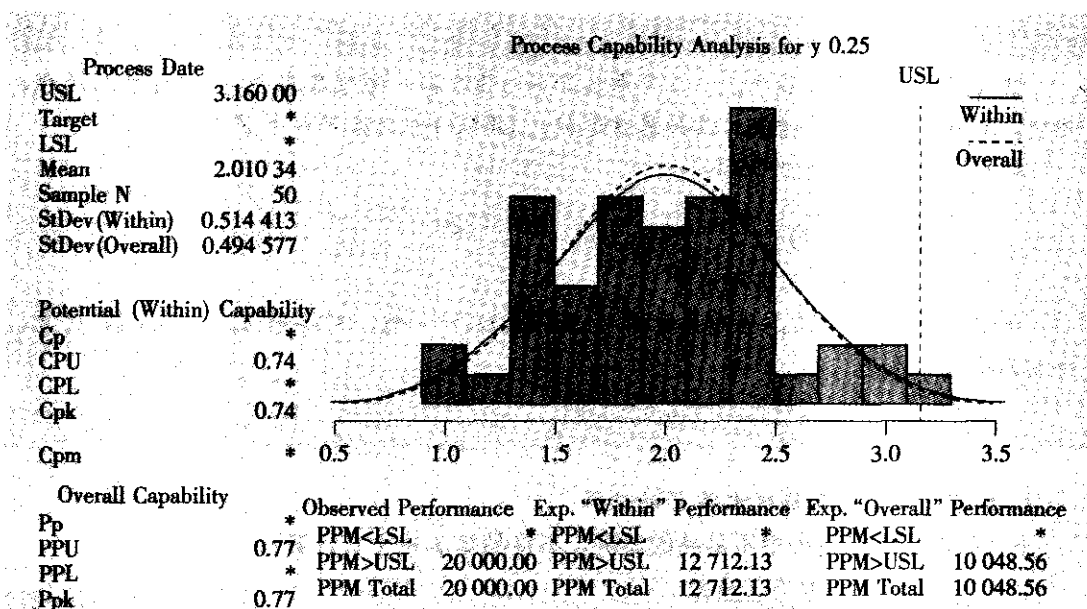


图 5—60 过程能力分析

对于变换后的数据计算过程能力指数时,应将原来的公差限也进行同样的变换。在本例中,原 $USL=100$,变换后,新公差上限为 $USL^*=3.162$,代入公式即可得到最后结果。用 MINITAB 软件可得如图 5—60 的结果。

其实,在 MINITAB 软件计算中,只要确定了变换的方幂为 0.25,不用手工计算新公差限,计算机就可以直接得到最后结果。

5.5.9 属性值数据的西格玛水平估算

前面我们介绍的均是过程输出特性为连续型数据的情况。对那些输出特性为非连续型数据的过程,也需要进行过程能力分析,测算它们的西格玛水平 Z ,评价它们的长期能力和短期能力。对属性值数据的情况,又分计点值数据和计件值数据,它们所对应的过程能力的估算方法有所不同。

1. 从 DPMO 到西格玛水平 Z 的计算

DPMO 是基于下面的参数计算得到的:

- (1) D =缺陷数
- (2) O =单位缺陷机会
- (3) U =单位数

$$(4) DPMO = \frac{D}{U \times O} \times 10^6 \quad (ppm)$$

在计算得到 DPMO 后,可通过查标准正态分布表,找到对应的 Z_{bench} ,然后令

$$Z = Z_{bench} + 1.5$$

即为其对应的西格玛水平 Z 。



某公司开单据的过程中，一个月中共开出 2 500 张，每张各有 10 处需要填写的栏，其中，共有 8 处出现了错误。请计算该过程的 DPMO 和西格玛水平 Z 。

(1) 计算该过程的 DPMO:

$$DPMO = \frac{D}{U \times O} \times 10^6 = \frac{8}{2\,500 \times 10} \times 10^6 = 320$$

(2) 计算该过程的西格玛水平 Z 。

由 DPMO 查标准正态分布表中的 $1 - DPMO \times 10^{-6}$ 值，可得到 $Z_{\text{bench}} = 3.41$ ，因此，得到西格玛水平:

$$Z = Z_{\text{bench}} + 1.5 = 3.41 + 1.5 = 4.91$$

2. 从不良品率到西格玛水平 Z 的计算

不良品率是生产过程中不良品的比例，不良品率 p 由下式计算:

$$p = \frac{\text{检验发现的不良品数}}{\text{检验的单位产品数}} \quad (5-27)$$

根据 p 值，可通过查标准正态分布表中的 $1 - p$ 值，找到对应的 Z_{bench} ，然后令

$$Z = Z_{\text{bench}} + 1.5$$

即为其对应的西格玛水平 Z 。

某公司开单据的过程中，一个月中共开出 10 000 张，其中，共有 105 张填写错误。请计算该过程的不良品率 p 和西格玛水平 Z 。

计算该过程的 p 值:

$$p = \frac{105}{10\,000} = 0.0105$$

由 $p = 1.05\%$ ，即 $1 - p = 98.95\%$ ，查标准正态分布表，可得到 $Z_{\text{bench}} = 2.308$ ，因此，得到西格玛水平:

$$Z = Z_{\text{bench}} + 1.5 = 2.308 + 1.5 = 3.808$$

本章小结

测量阶段是六西格玛项目工作中一个非常重要的阶段。没有测量，也就没有改进。正像我们在前面所指出的那样，从测量阶段起就要开始收集数据并着手对数据进行分析，通过测量阶段的数据收集和评估工作，可以获得对问题和改进机会的定量化认识，并在此基础上获得项目实施方向的信息。

帮助项目团队和黑带完成测量阶段工作的工具有很多。从这个阶段开始，需要正确地应用统计学的知识收集和处理数据，并从中获取有价值的信息。这个阶段所涉及的主要知识有：关于过程分析和文档的知识，包括流程图的绘制和分析方法、因果图和因果矩阵分析方法、失效模式和影响分析 (FMEA) 等；关于概率与统计学的基础知识，这些知识是

数据分析的基础；关于数据的收集和整理的方法，黑带应能正确地应用这些方法，将实际问题转化为可测量和定量描述的问题；关于测量系统分析的基本概念和方法，这些方法是保证获得可靠的测量数据所不可缺少的；以及过程能力分析的工具和方法，过程能力分析是测量阶段最重要的工作内容，也是测量阶段应产生的输出结果。

分析阶段是 DMAIC 各个阶段中最难以预见的阶段。项目团队所使用的方法将在很大程度上取决于所涉及的问题与数据的特点。在这个阶段中, DMAIC 团队应详细研究资料, 增强对过程和理解, 进而识别问题的原因, 使用各分析步骤来寻找“问题根源”。

有时, 问题的根本原因是显而易见的, 此时, 团队能很快通过分析找到; 有时, 影响问题的原因很多, 而且非常复杂, 多年来许多人都根据他们习惯的方法去做工作, 而没有考虑和证明其合理性。当这种情况发生时, 团队可能要花几周或更长的时间, 使用各种工具和试用各种不同的观点去分析, 才可能得出正确的结果。

DMAIC 团队用循环分析方法来实现他们对原因的探索。这个循环从数据的测量开始, 通过对过程的分析, 提出对原因的初始推测或者假设; 接着收集和关注更多数据和其他看得见的证据, 对这些推测或者假设做出进一步的判断; 分析循环继续进行, 各种假定不断地被确认或被拒绝, 直到真正的问题根源通过严谨的数据分析被明确识别出来。分析步骤中最大的挑战之一是正确使用工具, 能够用简单工具找到根本原因, 就不用复杂工具。有时原因是深藏而不易识别的, 有时问题与其他较多因素相关联而被纠缠混杂在一起, 这时就需要用到更高级的统计技术或其他管理技术来确定和验证根本原因。本章介绍的统计方法将有助于团队除了使用常用的图示工具以外, 还通过运用统计推断、假设检验、方差分析和回归分析等方法来确定项目问题的根本原因, 确定和验证 $Y=f(X_i)$ 的关系, 为改进找到有效的方向。这里需要说明两点, 一是尽管本章介绍了很多统计分析工具, 但六西格玛的分析工具并不仅指这些工具, 六西格玛的工具箱是开放性的, 任何适用于实际项目的分析工具都可以被整合和应用, 作为一名六西格玛黑带或绿带, 要善于持续学习。二是分析和测量两个阶段之间有很多交叉, 在对某一个原因进行深入分析时, 有可能发现已有的测量数据远远不够, 需要补充测量大量数据。

6.1 探索性数据分析和过程分析

一个六西格玛团队应该在现场收集足够多的各种形式的数据证据, 然后才能分析出究竟是什么引起缺陷的真正原因。要注意的是, 绝不应该接受表面上的假设, 必须证明确实是此原因造成了所研究的问题。有时寻找原因还得经历一个过程, 这个星期得出的结论可能会被下一星期的数据所推翻; 再一个星期收集的新数据可能又暗示了一个新的问题等等, 直到团队成员能够用数据真正证实某个假设, 捕获到了真正的原因。



有许多调查缺陷原因的有力工具可以使用，有两类不同的分析方法可用于研究问题的真正原因。

(1) 探索性数据分析。利用测量值和有关数据（已经收集的数据或在分析阶段收集的新数据）来发现、建议、支持或排除缺陷原因的模式、趋势和其他的异常，分辨问题模式、问题趋势或其他一些有关因素，这些因素可以是推测出来的，也可以是已证明或未证明的可能因素。

(2) 过程分析。深入研究并分析过程工作是如何开展的，从而识别出与过程目标不一致、可能引起问题发生或导致问题发生的环节。通过增值分析，可以判别过程中哪些环节是对顾客增值的，哪些环节虽对顾客非增值但对过程是增值的，哪些环节是不增值，可以考虑删除的，哪些是不能确定是否增值的，最终找到过程不能满足顾客需求的真正原因，找出改进的方向。

大多数团队在项目的分析阶段同时使用这两类工具。一般先做探索性数据分析，也可以先使用过程分析工具，还可以把团队分成两组，同时进行数据和过程分析。不论团队采用以上哪种模式，缺陷原因的分析是个循环过程，一般有三个阶段：

(1) 推测。以开放的态度调查数据和过程，目的是找出可以研究的问题或因素。

(2) 提出关于原因的假设。运用已有的知识，提出最有可能的缺陷原因。

(3) 证实或排除原因。利用数据、试验或进一步的过程分析对原因的假设做出判断、证实，确认哪个因素或环节是引起缺陷的最重要原因。

事实上，这三个阶段也不可能截然分开。团队可能在推测之前就已经有了关于原因的假设；如果你不能提出假设，就得重新回到推测阶段。事实上，大多数团队在正确定位问题关键原因之前，都要经历好几次推测——假设——证实的循环。团队会议的讨论从来都是十分必要的，即使你不按照正常顺序经历这些步骤，或者需要重新回到某一步骤也要经过讨论来决定，同时每次会议开始时都要明确本次会议的重点：“我们在这儿是为了提出关于原因的假设……”，“我们在这儿讨论如何证实这些原因……”或“今天的会议重点是我们将如何分析数据”。

6.1.1 探索性数据分析

数据分析的目的是用各种方法将数据转化为信息，使数据变得有意义。但是这种转化并不那么简单，尤其是对六西格玛团队而言，面对的问题是非常复杂的，牵涉的原因是多方面的。如前所述，数据分析应该分为三个阶段：推测阶段、提出关于原因的假设阶段、证实或排除原因阶段。下面分别加以叙述。

1. 数据分析的第一阶段：推测

为了能进行有效的推测，我们首先要学习数据分析的原则，然后学习些简单分析工具，最后学会一个重要分析工具——多变异图的分析。

(1) 数据分析原则。要时刻牢记三个要点，这会帮助你理清思路并有效地进行数据的分析。

1) 明确要深入了解的方向。众所周知，在工作中存在着大量数据，它们很容易让你觉得无从下手。正确的做法是经常重新翻看项目特许任务书和问题陈述，时刻牢记团队的目标。只有这样才能牢牢抓住分析的大方向。



2) 不断提出假设。有很多方法可以分析数据, 如果不当心的话, 六西格玛团队很有可能在摸索方法上浪费很多时间。较好的方法是, 先结合团队任务做出一个假设, 这会帮助你决定如何分析数据。例如:

假设: 餐馆的顾客抱怨增多是来自新的、不熟练的服务生。

分析方法: 把顾客抱怨数据分成两组——来自新服务生招待的顾客和来自有经验服务生招待的顾客, 寻找两组数据的系统差别。

注意: 要充分接受“假设是有缺陷的”这种可能性, 不能视而不见那些与假设不符的数据。有经验的数据收集者知道, 事实上, 在最开始阶段, 他们的绝大多数假设都会有缺陷的。但是, 正是认识到要注意不断改进缺陷, 才帮助他们最终确定出正确的假设。

3) 注意关于事件发生的频率、影响程度以及与问题缺陷症状相关的问题。如果把调查局限于一个假设或一个问题, 你甚至都不敢肯定是否抓住了正确的问题。较好的方法是提出更多问题, 通过分析找到其中最重要的问题。以下是上述餐馆六西格玛团队可能提出的一些问题:

新服务生招待的顾客是否比其他顾客抱怨得多些? (问题频率)

新服务生招待的顾客抱怨什么? 与有经验服务生招待的顾客抱怨相比, 有何不同? (观察到的问题)

如果新服务生招待的顾客抱怨得更多, 是否意味着这些顾客将不太可能重新光顾? (问题的影响)

团队需要对问题有深入的了解, 以便合理地决定究竟应该把时间花在哪里, 解决方案在哪里, 如何实施解决方案; 否则, 你很有可能浪费三个月时间解决一个不常出现或对顾客没什么影响的问题。

运用这些原则将帮助你使团队的数据收集工作紧紧围绕如何发现问题根本原因来进行。你可以依据这些原则来利用你在测量阶段识别和收集到的分层数据。例如, 对于刚才提到的餐馆, 六西格玛团队需要把关于抱怨的数据分为“新的”和“有经验的”服务生, 以开展分析。

上述原则其实都包括在一个称为逻辑原因分析的过程中, 只是大家可能平时对此没有特别留心罢了。任何曾经试图修理过某种家电故障(如面包烤箱)的人都会熟悉此过程的一些基本构成。

例如, 厨房里的面包烤箱一直正常运转, 但是今天早晨你却发现不能把面包塞进去了, 更别说烤面包了。假设“可能是断电”了, 但是“数据”马上把这个假设推翻了, 烤炉连着电源, 炉侧的红灯亮着, 表明有电流通过。然后, 再假设里面的“电线松了”, 于是拿起一把螺丝刀, 将线头螺丝拧紧(当然, 不要忘记拔电源), 再进行测试。逻辑原因分析的道理很简单: 当你的数据不支持你的假设时, 你的假设很可能是有缺陷的, 你就接着提出下一个假设。

尽管对于复杂的设备或者过程, 分析可能要复杂得多, 步骤可能几十个, 但基本概念还是同样简单的, 问一些问题, 这些问题能帮助你识别过程、材料、方法或人员在有故障或无故障之间的差异。我的烤箱与新的相比有什么不同? 它旧了, 里面掉进了面包屑? 还是线头松动? 或者电线被烧断了? 六西格玛问题解决方法就是要通过逻辑原因分析层层剖析, 直至找到根本原因。

下面是一些有用的逻辑问题，帮助分析问题产生的原因：

- 这些缺陷属于同一类吗？缺陷在人员、方法、过程步骤方面有哪些不同？
- 这些问题是否在某处更容易出现呢？这些出现故障的地方有哪些不同之处？如果你的产品在上海的门店平均销售额与深圳比要低，你需要确定上海这个城市有什么不同，而不是仅仅看销售额的不同。销售人员、店面设计、过程或顾客等因素有哪些不同？有什么想法吗？你要仔细寻找问题或过程之间潜在的不同之处。
- 是否在某个时间段内这些缺陷是非常普遍的？这个时间段与平常比有什么不同？比如烤炉经常在前一天晚上还正常运转，第二天早晨却不能使用，是什么原因呢？
- 当缺陷出现或数据发生波动时，是否有其他事物或者变量也发生波动？我们正在寻找一种相关性，一种绝非出于偶然的联系。例如，A市每月销售额是否与季节有关？同样的关系也存在于B市吗？

思考这些问题能帮助我们提出假设并检验假设，逐个排除，直到一个或几个假设被证明是站得住脚时为止。

(2) 初步数据分析。团队可以使用许多不同的方法和工具来初步分析数据。但通常，以下几种简单工具能帮助大家分析数据：

1) 排列图。一种由大到小排列好的柱状图，帮助团队关注有最大影响的问题组成部分，适用于离散型或属性数据。

2) 趋势图。一种以时间作为横轴表明某指标变化状况的工具，帮助团队观察随时间发展所关心的指标数值是否呈现一定的模式，适用于连续型数据。

3) 直方图（频率图）。表示所关心的指标其数值在各组间分布的状况。适用于连续型数据或属性数据。

(3) 多变异图。多变异图（multi-vari chart）是本节重点介绍的方法。它可以直观地提供与过程有关系的多个因素如何对过程输出响应变量发生影响的图形。六西格玛团队在研究多个变量时，可使用多变异图形象地描述因素间的关系。通常，这些图应该在方差分析等定量的数值分析之前做。团队可以通过对数据做简单的定性分析，获得初步的了解，提出变异产生原因的假设，然后进一步做定量分析。

项目团队研究三种材料在不同的通电时间下其超导强度的变化状况。数据收集是在三个通电时间长度分别为0.5分钟、1.0分钟和1.5分钟里测量三种类型材料（材料1、材料2和材料3）的各三个样品得到的。希望考察材料类型及通电时间长度两个因子对超导强度是否有显著的影响（见表6—1）。

表6—1

超导强度数据表

A 通电时间长度	B 材料类型	超导强度
0.5	1	23
0.5	1	20
0.5	1	21
0.5	2	22
0.5	2	19



续前表

A 通电时间长度	B 材料类型	超导强度
0.5	2	20
0.5	3	19
0.5	3	18
0.5	3	21
1.0	1	22
1.0	1	20
1.0	1	19
1.0	2	24
1.0	2	25
1.0	2	22
1.0	3	20
1.0	3	19
1.0	3	22
1.5	1	18
1.5	1	18
1.5	1	16
1.5	2	21
1.5	2	23
1.5	2	20
1.5	3	20
1.5	3	22
1.5	3	24

利用 MINITAB 软件, 从 STAT-Quality Tools-Multi-vari chart 入口, 可以绘出多变异图 (见图 6-1)。

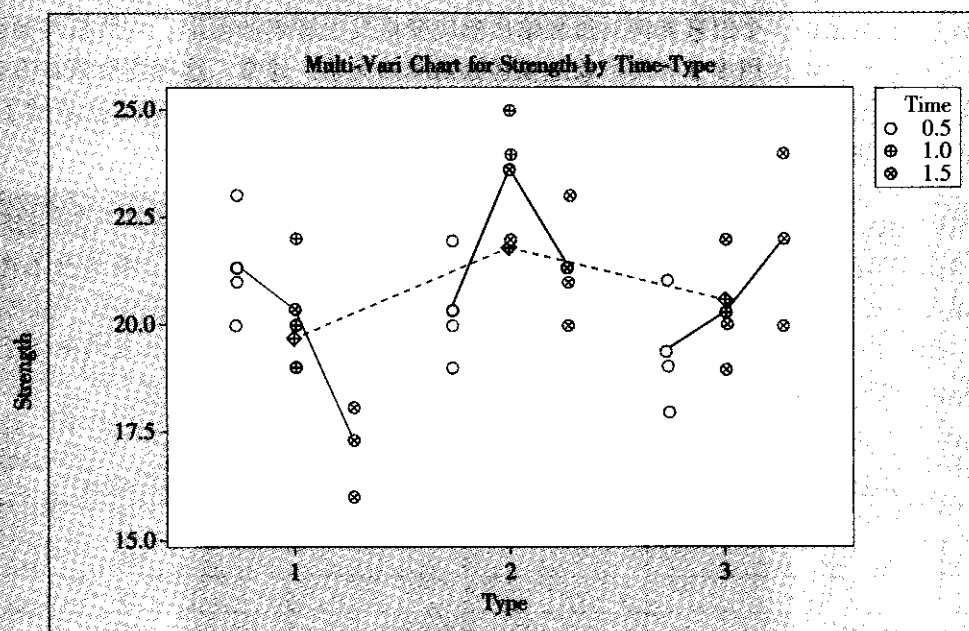


图 6-1 超导强度多变异图

在输出的多变异图中,每一种材料上都有连接三个点的连线,这三个点表示在对应的材料类型中,各种通电时间长度下超导强度的平均值,反映了各个材料类型组内误差的信息。图中虚线连线上的各个点分别代表三种材料的超导强度的平均值,反映了各个材料组间差别的信息。从此图来看,组间虽有差别,但并不严重。但是,同种材料内分别对应的通电时间长度(分别为:0.5分钟、1.0分钟、1.5分钟)的超导强度差异较大,而且不同材料对应各种通电时间的图像形状差别很大,说明材料的种类与通电时间可能还有交互作用。

2. 数据分析的第二阶段:提出关于原因的假设

(1) 寻找有效解决方案的秘诀。必须精确地知道正在试图解决的是什么问题。这个任务可不简单!想想所在组织中所有一而再、再而三地被“解决”的问题吧。显然,提出的许多解决方案并没有针对真正的问题,因而也就没有很大的作用。

(2) 发现问题潜在原因的秘诀。必须透过问题的表面现象进行深入研究,找出问题的真正潜在原因。在你完成分析过程的推测之前,你的团队已经对何时、何地、问题如何暴露出来等有了一定的了解,同时对潜在原因也有了许多想法,以便以后能真正找到解决办法。

在寻找原因的这一阶段,最常用的两大工具是因果图和关系图。它们为找到真正的根本性原因提供了关键链接。使用这些工具时有两点要注意:

一是它们仅仅帮助团队进行系统的思考并发现问题可能的潜在原因,仍需收集数据才能证实究竟什么才是问题的真正原因。

二是它们的有效性直接与你思考的创造性和深度有关。这就是为什么这些工具在团队一起用时最有效,即需要许多人同时进行头脑风暴,经过认真讨论列出一张广泛深入的潜在原因分析检查表。

3. 数据分析的第三阶段:证实或排除原因

验证引起缺陷的原因,可以通过三种途径:因果逻辑分析、统计假设检验和试验验证。下面分别予以介绍。

(1) 因果逻辑分析。假定团队认为许多弄错的订单是由于销售人员在订单的管理过程中出了差错而引起的,这也许能解释部分出错订单,但它不能解释不熟悉订单细节的顾客(小商人和计算机新个人用户)会出现更多差错这一现象。如果订单管理系统出了问题,你应该想到所带来的影响对于所有顾客都是差不多的。因此,假定订单管理系统是引起问题的原因,这或许能解释那些可见的问题(如出现有缺陷的订单),但是它不能解释那些不可见的问题:小的商业订单和计算机新用户订单差错更多。

要使假定的原因真的成立,它必须通过合乎逻辑的测试来解释从数据反映出的可见问题和潜在问题,当然,这些潜在问题现在并没有显现出来。

(2) 统计假设检验。用来分析并确定因果关系的最常用统计方法就是散点图(或称散布图)和相关分析。判断原因 X 和结果 Y 之间的相关程度,通常可从散点图获得一个大概的印象,然后再用精确的计算来进行回归分析和相关分析以获得确定的统计模型。当然也可能最后判明这两者间完全无关,没有任何模型。散点图提供了验证因果假设的一种途径,从若干成对数据去验证自变量 X 与因变量 Y 之间是否存在相关关系。成对数据顾名



意义是指 X , Y 是来自同一个观测个体,“成对”本身就反映了两个变量间的关系,它们可能只是表面上的相关,也可能是一种潜在的原因和结果。例如,在某项目中要构建一个体系,体系的复杂度(X 为变量或是潜在的原因)与构建该体系所花时间(Y 为输出或结果)之间潜藏着一种因果关系。一般来说,将一批成对数据(X , Y) 作为平面坐标系内的坐标,可以构成散点图,在图中显示其相关性,如图 6—2 所示。

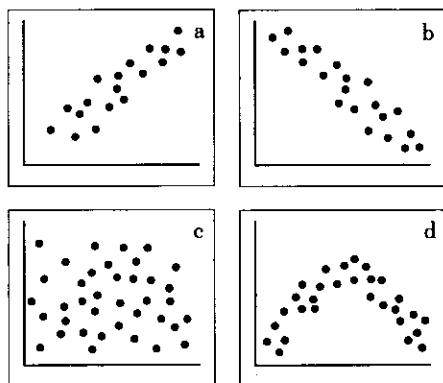


图 6—2 散点图

如果数据点分布的形式是从左下往右上以一定的角度展开(图 6—2a),表示自变量 X 和因变量 Y 是一种正相关关系。简单而言,这种走向表示“随 X 的增大, Y 有增大的趋势”。其他的类型则是负相关(图 6—2b)、无关(图 6—2c),或者是更复杂的相关性(该例是抛物线型,见图 6—2d)。

散点图显示出的很强的正相关性并不能百分之百地保证 X 的增加一定会引起 Y 的增加,但确实表明它们以某种方式互相关联并且不是完全随机地同时发生的。请特别注意:在任何时候都不能从“ X 与 Y 高度相关”得出“二者有因果关系”的结论。例如,商店中“雨伞日销售量”与“当日道路交通事故次数”高度相关,但二者间并无因果关系,其高度相关的真正原因是当日在下雨,这导致雨伞日销售量和当日道路交通事故次数同时增加。因此,团队在用散点图来说明相关性时要小心,但有时也确实会找到些因果关系的成果,因此这种方法还是很有用的。若考虑到的变量不止两个,则相关分析可能会很复杂,这时,可以对变量分组,进行两两组合,得到多个散点图来验证猜测的正确性,也可以将所有变量的两两数据散点图画在同一张矩阵散点图(matrix scatter plot)上,对它们同时进行分析。

画散点图还可以先对数据进行整理加工和分类,对同一个问题画出不同种类、不同层次的散点图。通过分层分析,团队常常会明白一个很有价值的结论:并非所有缺陷都是同一个背景原因,测量阶段找到的分类因素在确定因果模型过程中都可能是有用的。

当团队某个队员对所收集的数据提出假设时,其他成员也都要协助他再仔细分析一下,找出那些有助于验证或否定该假设的更细致的分层或分类信息,从而得出更准确的判断。

(3) 试验验证。在某些情况下,想要收集到现存数据以证实导致失误的准确原因是很困难的,甚至是不切实际的。此时,一个可行的方法是做些小的试验,先去稍稍改变过程

以消除可疑的原因，并且监测结果，最终判断这个原因是否为过程出现缺陷的真正原因。当然，做这样的试验不能光凭个人意愿去改变！只有把这种改变看成一次常规的局部检测试验才能起作用。要注意做到：

- 明确改变的内容和方法；
- 清楚改变的目的；
- 先在小范围内进行试点；
- 仔细考虑潜在的副作用并制定应对措施。

6.1.2 过程分析

上节介绍的数据分析是以过程中收集到的数据为依据的。本节将介绍如何对过程进行分析。在六西格玛团队组成后，一定要很清醒地认识到，不是团队所有成员对整个过程中都非常熟悉。即使他们自认为对过程非常了解，借助详细的流程图或过程影射图（process mapping），通过系统的过程分析，他们也会非常吃惊地发现有很多自己不知道的事。这里提供的过程分析工具和方法将会帮助团队全体成员对过程的认识达成一致，并且通过集体讨论指出过程中可能引起问题的疑点和波动。与数据分析方法一样，过程分析也分为三个阶段：推测阶段，提出关于原因的假设阶段，证实或排除原因阶段。

1. 过程分析的第一阶段：推测

在项目的界定阶段，团队可能提出一个用于界定项目过程边界的高端（或称宏观）过程图（SIPOC图），SIPOC图着重分析宏观流程，比较粗糙，许多过程中出现的细节问题在SIPOC图中都找不到。因此，在分析阶段，团队必须更细致地去分析过程中的各个细节部分，以便能够引导你发现某些隐藏着的X变量（原因）。为此，我们要绘制详细流程图并证实它。

（1）绘制详细流程图。本书5.1节已经讨论了流程图的绘制，这里不再赘述。只是从过程分析的角度要特别指出，详细的过程流程分析或过程影射对问题分析是至关重要的。当然流程图绘制的详细程度取决于项目的目标和问题本身，如对于一个生产线而言，如果进行工作的安排，只要将生产过程所需的作业及其前后关系描述清楚就可以了；而对于如何降低生产线的节拍，则可能需要将工人的动作都要描述出来，如绘制工人双手作业程序图。一个详尽的过程流程图的绘制本身，对团队成员而言就是一个有意义的经历，在这个过程中他们慢慢地认识到别人的工作，意识到本流程中所使用的方法其实可以有很多的变动，尤其是服务行业。

该阶段另一个有用工具是跨职能流程图（参见图5—2），即在通常的流程表中增加具体的职能部门，在图上标明哪个部门负责完成哪个环节。如果一个业务流程需要多个部门的参与（许多服务行业都是这样的），那么跨职能流程图将非常有用。

（2）证实流程图。一旦团队完成了一张详尽的流程图，就应该花时间去证实从图上反映的内容是否就是实际中真实发生的，因为有时候它们只是团队成员头脑中的产物。为了证实这些，还要花几天时间去跟踪某些典型的产品或服务小组的处理过程。应该使用一张过程检查单来做这个工作，务必以足够的细节来描述过程以获取每个重要步骤，补充流程图并证实流程图的真实性。



2. 过程分析的第二阶段：提出关于原因的假设

团队成员在绘制流程图中，潜在的问题就会被自然而然提出来。如果三个人处理同一过程的某一步骤，但他们对如何完成该步骤的方法不能达成一致，那么实际上就找出了潜在的原因。当然，还需要证实这种混乱的确是过程中发生缺陷的一个重要原因。如果有人处理程序的顺序是 A, B, C, 另一个人则是 B, A, C, 那就要查看这种顺序的不同是否会对产品或服务产生影响。

在本阶段中，可以细分为两层：先要识别过程中明显存在的问题，再进行量化增值分析。

(1) 第一层分析：识别过程中明显存在的问题。流程图一旦确定下来就可能成为团队会议的焦点，成员们可以用手中的记号笔来标记如下一些问题。讨论时要记住，我们是在寻找那些经常发生以至于已经成为过程中的一些典型性问题。

脱节：过程处理中存在于轮班之间、顾客和供应商之间或者经理和职员之间的使信息交流中断的步骤。

瓶颈：过程中的某个步骤的工作量超过了其本身原设定的容量，此步骤就会减慢整个进程的流动。可以设想，如果某项工作必须等待张三休假回来才能再进行，你就会看到瓶颈作用。

冗余：不同过程中的不同步骤提出了相同的活动和结果。例如，两个步骤中提出了相同的信息并且带着相同的信息进入相同的地方。

返工循环：过程中某步骤的输出产品有缺陷，必须送回其上游步骤，或另设“修复”步骤直到必需的工作完成为止。检查步骤往往是返工循环的起始点。

决定点/检查点：这是过程中的一个步骤，在那里进行一系列的检查和评估工作：对于合格产品进入正常程序的下一步；对不合格产品进入另设的“修复”阶段或进入返工循环。对于质量水平高的过程，这样的检查点是很少的；而在缺陷严重的过程中，有大量这样的检查点。由于每一种类型的问题都可能成为在过程中导致缺陷的一个潜在的原因，为了减少缺陷送到顾客手中的可能性，必然要设立很多检查点。但是从顾客的角度看，由于检查步骤没有任何有价值的输出，而且延长了产品和服务送达顾客的时间，对生产的进行造成了阻碍，而这些增加的成本时常会通过某种形式转嫁到顾客的头上，顾客当然希望这种检查点越少越好。

(2) 第二层分析：量化增值分析。当过程较复杂时，人们往往会关注过程本身的顺利进行而忽略了顾客为什么会花钱买你的产品。在高度发达的商业社会中，任何一个单独的集团为市场和社会提供的价值总是微不足道的，但它确实为顾客提供了某种价值。因此，“价值分析”作为一种途径，它通过外部顾客的视角重新审视工作流程并再三强调识别顾客需求。在量化增值分析的过程中，我们将每个步骤分成三种类型：

1) 顾客增值步骤 (customer value added)。这些任务或活动在外顾客看来是有价值的。这很重要，因为几乎任何步骤在某些操作员的眼里都可以被证明是有理的。“老板要我们做，我们就做”并不意味着这项工作对顾客而言是增加价值的。

对顾客而言，增加价值的三个标准是：

- 顾客关心的，顾客承认它的价值并愿意付钱给我们。
- 改进提供的服务和提高了产品质量。显然，货物运输只移动物体位置，并不增加

价值。

● 顾客希望一次做对。要注意：修补、返工、替换等只是改正以前造成的缺陷，它们并不增加价值。

顾客增值步骤必须同时满足上述三项标准。

2) 运行增值（或过程增值）（operational value added）。有时我们采取的一些活动，没有同时满足顾客增值步骤三个标准，但是具备潜在的增值可能。比如对于在流水线上为每件产品贴上编号条码，顾客不会因为我们贴了条码而愿意多付钱。不过贴条码可以使生产过程的效率大大提高，这就是使运行（或过程）达到增值。对于员工的培训也是这样，有效的培训可以提高员工的意识和技能，在适当的时机，可能开发出新的产品和服务。因此称这些活动是运行增值或潜在增值的步骤。

3) 不增值（non-value added）。在许多过程里都存在大量的不增值的步骤。属于这种类型的活动包括返工、延迟、检查、回顾、安装和准备、传输（从一个地点或过程中的一个步骤传到下一个地点或步骤）、内部报告和内部调整等。不增值类型的判定看起来相当残忍。但仔细研究会发现，许多组织所干的事情大部分在顾客眼里确实并没有增加价值。如组织内部流程中大量的审批、报告，生产过程中大量的设备调整、检验、返修、维护等，诸如此类的事很多很多，你的顾客会愿意为这些步骤多支付金钱吗？

上述分析是把过程的步骤分成三种与价值相关的类型，可以帮助你识别出哪个步骤从价值的角度对顾客而言是最具重要性的。与此类似的是，把过程的步骤分成三种与时间相关的类型，这就是时间分析法。

过程的时间分析可能会给你另一个震惊：商业过程中往往有许多时间是空闲着的，并不是人闲坐在那里，而是那些该做的事情衔接不上。员工们都忙在输送设备的箱柜里、在文件夹里、在篮筐里、在 E-mail 的附件里，并不真正做任何事情。在优先考虑生产过程循环时间的地方，时间分析的巨大作用使得整个过程耗时由天削减到小时、由月削减到天。激烈的商业竞争从“及时的递送”到快速产品周期再到多种基于时间的服务竞争，这些都是速度的需要。在过去的 15 年里，在全世界范围内，时间竞争已经给人们留下了深刻的印象，但是直到现在，仍然有很多过程尚没有被触及，因而还隐藏着巨大的潜力。

过程的时间由两个部分组成：

工作时间：实际花费的用于生产产品或向顾客提供服务的时间。

等待时间（或称“排队时间”或“堆栈时间”）：产品或服务为等待某些事情发生而花费的时间。想象一下，有一批零部件，有一堆的请求，或者一车皮的产品都等在那里，等着有人来处理它们或者移动它们，这是多大的浪费！本书在第 9 章的精益六西格玛中，会介绍消除浪费的方法。

3. 过程分析的第三阶段：证实或排除原因

一般说来，从过程分析中辨认并查实原因要比从数据分析中辨认并查实来得容易些。为什么呢？因为过程分析的判断准则很简单，即任何非增值过程步骤，任何增加耗时但并不增加价值的步骤都是在浪费时间，应尽量删减。

但是，你将不得不做更多的调查工作，因为有些缺陷对顾客而言既重要又很容易被他们注意到，而你并不明白在过程步骤中存在的疑惑或变数是否真的会导致那些缺陷，你就必须分析清楚并作出结论。如果在物流运输过程中，张三总是将航班标签贴在箱子的正



中，而李四总是将它们贴在右下角，那结果会怎样呢？它们对运输效率是否有影响？只要这种状况实际存在于运输过程之中，它就可能成为一种影响运输效率的原因。

(1) 证实过程原因的提示。

- 使用因果关系的逻辑，就像对待数据分析一样。
- 如果你还没有这样做，绘制一张详细的过程流程图来看是否在过程步骤中存在重要区别。

● 定点测试中的“试验”是另一种验证过程的缺陷是否确实带来影响的好方法。只要你在过程中，人为“试验”发生一个预定的波动（波动规模当然要小些），你去测量波动造成的影响，就可以断定此缺陷是否为真正的原因了。例如，你认为物流部门在操作员总是把标签贴在确定的不变位置的情况下能够以更快的速度完成任务的话，用这种方法来测试一个或两个星期，看过程时间是否减少了。

当然，最终目的是完全去除那些在顾客眼里不能增加价值的任何步骤，包括花费时间去返工和检查等等。去除许多上面列举的不增加价值的步骤，你将很可能缩短整个过程循环的时间以增加利润，这也会令你的顾客感到高兴。六西格玛团队常常可能去除掉 70%~80% 在顾客看来是被用来花费在返工上和其他不增加价值的步骤上的时间。但是，仅仅加速了问题处理的过程对团队并不是最重要的，团队必须设法消除那些已经分析揭示出的缺陷原因，然后才能进入提高阶段。

(2) 高级分析工具。大多数情况下，六西格玛团队利用本章前面所描述的基本工具完全能够把事情做好。利用排列图进一步挖掘关键事实，利用流程图查明特殊的原因，并且通过团队深入讨论关于如何能使过程真正走势良好（和如何使它走势更好），这仅仅是开始，减少缺陷的路还很长。

还有更有效的工具来测试变量 X 和输出 Y 之间的关系。团队应考虑下列问题：

- 测试几批数据之间是否存在明显的区别。
- 关于问题的原因要提出一个有效的假设。
- 证实或者否定关于原因的各种假设。

为了回答这些问题，六西格玛团队可以利用下面系列的统计推断来检验假设，包括：

- χ^2 检验
- t 检验
- 方差分析 (ANOVA)
- 多元分析

虽然这些检验对不同种类的数据使用不同的技术，它们的目的是为了回答这样一个问题：在数据中所看到的差异（或原因）是真的吗？有没有统计上的显著差异？

例如，有一个团队决定分析使用两种不同方法是否会导致整个过程运行时间不同。在一周时间内，分别用两种方法各试验若干次，测量出每次的过程运行时间。团队可以用 t 检验来分析测量到的两批平均运行时间之间是否有统计意义上的显著差异。现在这些检验都是用计算机程序来完成，例如 MINITAB 应用软件，只要输入了正确的数据，它就会完成所有与之有关的统计计算。

6.1.3 点估计和区间估计

统计学的应用非常广泛,但概括起来说,统计推断无非是“估计”与“检验”两大类。由于“检验”在六西格玛管理中有更多的应用,所以从6.2节起,我们将对其进行系统的讨论。本节将简要地介绍有关“估计”方面的知识。先介绍点估计,即对总体未知参数用一个数字来估计;再介绍区间估计,即对总体未知参数用一个区间来估计。这两种估计方法都很有用。

1. 点估计

在正态分布中,均值 μ 、方差 σ^2 、标准差 σ 等都被称为总体的参数(parameter)。在实际问题中,这些参数都是未知的,常常需要选用适当的统计量(statistic)作为未知参数的估计。用于估计未知参数的统计量称为点估计量,简称点估计,或估计(estimator)。

同一个未知参数的估计量可以有很多个,一般选用其中常用的一两个作为参数的估计量,选取的标准之一便是无偏性。

设 θ 为未知参数, $\hat{\theta}=\hat{\theta}(x_1, x_2, \dots, x_n)$ 是 θ 的估计量,若 $E(\hat{\theta})=\theta$,则称 $\hat{\theta}$ 是参数 θ 的无偏估计,即从长期大量的情况来看,用这种方法进行估计之平均值与真正的参数值是一致的,既不偏高也不偏低。这里无偏估计的条件 $E(\hat{\theta})=\theta$ 等价于 $E(\hat{\theta}-\theta)=0$,其中 $\hat{\theta}-\theta$ 是估计量 $\hat{\theta}$ 与真值 θ 的偏差。由于样本的随机性,导致这类偏差也是随机的,可大可小,可正可负,所以无偏估计的含义也可以解释为,每次使用 $\hat{\theta}$ 去估计 θ 都可能有偏差,但是多次使用它,偏差的平均是0。

(1) 正态均值的无偏估计。设 x_1, x_2, \dots, x_n 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本,其中参数 μ, σ^2 都未知。

正态总体均值 μ 常用的无偏估计有两个,一是样本均值 \bar{x} ;一是样本中位数 \tilde{x} ,即

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$
$$\tilde{x} = \begin{cases} x_{(\frac{n+1}{2})}, & n \text{ 为奇数} \\ \frac{1}{2} [x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}], & n \text{ 为偶数} \end{cases}$$

式中, $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ 是有序样本。

当样本量 n 为1或2时,这两个无偏估计相等,但是当 $n \geq 3$ 时, \bar{x} 优于 \tilde{x} ,这是因为样本中位数 \tilde{x} 仅使用了样本的部分信息,而样本均值 \bar{x} 使用了样本的全部信息,所以在实际中,应该优先选用样本均值 \bar{x} 作为 μ 的估计。有时在现场,为了简便快捷,当 $n < 10$ 时,也可选用样本中位数 \tilde{x} 作为 μ 的估计。另外,当样本中的某些观测值受到较大的特殊影响,远远偏离了原正态分布时(常把这种值称为异常值),由于样本均值依然会受到其影响因而对它们太“敏感”,而中位数则更具有稳健(robust)特性,所以两种估计都是很有用的。二者通常相差不多,一旦发现二者相差悬殊,则提示我们可能分布偏斜得非常厉害,或者样本中有异常值出现。

在第8章的控制图中, \bar{x} 图与 \tilde{x} 图便是它们的应用。

(2) 正态方差 σ^2 的无偏估计。正态方差 σ^2 常用的无偏估计就是样本方差 s^2 ,即

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

理论研究表明,它是 σ^2 的所有无偏估计中最好的。

(3) 正态标准差 σ 的无偏估计。人们常常认为,既然在正态分布中 s^2 是 σ^2 的所有无偏估计中最好的,那么 s 就应该是 σ 的最好的估计。其实不然,如果用 s 来估计 σ ,则通常会偏小些,它并不是无偏的,只有对它稍加修偏才行。对于用极差来估计标准差,则更需要考虑到样本量对极差的影响。因此,估计标准差 σ 常用的无偏估计有两个;一个是将样本极差除以一个与 n 有关的常数而得到;另一个是对样本标准差进行修偏得到。

$$\hat{\sigma}_R = R/d_2, R = x_{(n)} - x_{(1)}$$

$$\hat{\sigma}_s = s/c_4, s = \sqrt{s^2}$$

式中,系数 d_2 与修偏系数 c_4 是与样本本身无关而仅与样本容量 n 有关的常数,其数值列在附表6中。

当 $n=2$ 时,上述两个估计相同,当 $n \geq 3$ 时,它们是不同的,理论研究表明, $\hat{\sigma}_s$ 优于 $\hat{\sigma}_R$,因为 $\hat{\sigma}_s$ 比 $\hat{\sigma}_R$ 波动较小,所以在实际中应该优先选用 $\hat{\sigma}_s$ 。但由于 $\hat{\sigma}_R$ 的计算简便快捷,所以在现场,当 $n < 10$ 时,也可选用 $\hat{\sigma}_R$ 。

当 n 较大(如 $n > 30$)时,修偏系数 c_4 近似为1,因此在样本量较大时,无须修正,可以将样本标准差 s 直接作为 σ 的估计。

在第8章的控制图中, s 图与 R 图便是它们的应用。

零件DF的间隙是一个重要特性,经过正态检验,间隙服从正态分布 $N(\mu, \sigma^2)$ 。现从生产线上随机取5个零件DF测量其间隙,得数据如下:

0.77 0.72 0.67 0.72 0.62

试给出参数 μ, σ^2, σ 的无偏估计。

解:用样本均值 \bar{x} 估计 μ ,用样本方差 s^2 估计 σ^2 :

$$\hat{\mu} = \bar{x} = (0.77 + 0.72 + 0.67 + 0.72 + 0.62) / 5 = 0.70$$

$$\hat{\sigma}^2 = s^2 = \frac{1}{5-1} (0.07^2 + 0.02^2 + 0.03^2 + 0.02^2 + 0.08^2) = 0.0325$$

标准差的估计,可以选用 $\hat{\sigma}_s$,由于 $n=5$,从附表6查得 $c_4=0.940$,则

$$\hat{\sigma}_s = \frac{s}{c_4} = \frac{\sqrt{0.0325}}{0.940} = 0.192$$

由于 $n=5$,也可以选用 $\hat{\sigma}_R$,从附表6查得 $d_2=2.326$,则

$$\hat{\sigma}_R = R/d_2 = (0.77 - 0.62) / 2.326 = 0.064$$

这里两者相差不大。

注:本例仅供显示计算方法使用。在实际工作中,为了估计均值,样本量 n 应大于15;为了估计标准差或方差,样本量 n 应大于30,否则波动太大,相应置信区间范围太宽。

2. 区间估计

当获得一个具体的样本后,上面叙述的点估计仅仅给出参数一个具体的估计值,但是

没有给出估计的精度,为此又提出了参数估计的另一种形式,即区间估计的概念。

设 θ 是总体的一个待估参数,其一切可能取值组成的参数空间为 Θ ,从总体中获得容量为 n 的样本是 x_1, x_2, \dots, x_n ,对给定的 α ($0 < \alpha < 1$),确定两个统计量,即估计量下界 $\theta_L = \theta_L(x_1, x_2, \dots, x_n)$ 与估计量上界 $\theta_U = \theta_U(x_1, x_2, \dots, x_n)$ 。若对任意 $\theta \in \Theta$ 有 $P(\theta_L \leq \theta \leq \theta_U) \geq 1 - \alpha$,则称随机区间 $[\theta_L, \theta_U]$ 是 θ 的置信水平 (confidence level) 为 $1 - \alpha$ 的置信区间 (confidence interval),也简称 $[\theta_L, \theta_U]$ 是 θ 的 $1 - \alpha$ 置信区间, θ_L 与 θ_U 分别称为 $1 - \alpha$ 的置信下限与置信上限。

$1 - \alpha$ 置信区间的含义是所构造的多个随机区间 $[\theta_L, \theta_U]$ 中有 $(100(1 - \alpha))\%$ 的区间包含未知参数 θ 的真值。由于这个随机区间会随样本观察值的不同而不同,它有时盖住了参数 θ ,有时没有盖住 θ ,但是用这种方法作区间估计时,平均说来,100次中大约有 $100(1 - \alpha)$ 次置信区间能盖住未知参数 θ 。如果 $P(\theta < \theta_L) = P(\theta > \theta_U) = \alpha/2$,则称这种置信区间为等尾置信区间。

有时在总体参数的估计中,我们关心的只是其中的单侧问题。例如,钢筋的平均抗拉强度,我们只关心它的下限是多少,这就是单侧置信区间。

只有下侧置信限的是: $P(\theta_L \leq \theta) \geq 1 - \alpha$

只有上侧置信限的是: $P(\theta \leq \theta_U) \geq 1 - \alpha$

(1) 正态总体参数的 $1 - \alpha$ 置信区间。下面给出单个正态总体均值、方差、标准差的 $1 - \alpha$ 的置信区间。

设分布为 $N(\mu, \sigma^2)$,从总体中抽取的样本为 x_1, x_2, \dots, x_n ,样本均值为 \bar{x} ,样本方差为 s^2 ,样本标准差为 s 。

1) 总体均值 μ 的置信区间的求法。

μ 的估计一般用样本均值 \bar{x} ,从 \bar{x} 的分布来构造置信区间。

(i) 当总体标准差 σ 已知时,利用正态分布性质可知:

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

由此可得 μ 的 $1 - \alpha$ 置信区间为:

$$\bar{x} - Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

记为 $\bar{x} \pm Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$,其中 $Z_{1-\alpha/2}$ 是标准正态分布的 $1 - \alpha/2$ 分位数。

(ii) 当总体标准差 σ 未知时,将 σ 用其估计 s 代替,由 t 分布性质可得

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(n-1)$$

故可以得到 μ 的 $1 - \alpha$ 置信区间为 $\bar{x} \pm t_{1-\alpha/2}(n-1) \frac{s}{\sqrt{n}}$,式中 $t_{1-\alpha/2}(n-1)$ 表示自由度是 $n-1$ 的 t 分布的 $1 - \alpha/2$ 分位数。

当样本量大于 30 时, t 分布已与正态分布近似相等,因而可得 μ 的 $1 - \alpha$ 置信区间为:

$$\bar{x} - Z_{1-\alpha/2} s/\sqrt{n} \leq \mu \leq \bar{x} + Z_{1-\alpha/2} s/\sqrt{n}$$



记为 $\bar{x} \pm Z_{1-\alpha/2} \frac{s}{\sqrt{n}}$, 式中 $Z_{1-\alpha/2}$ 是标准正态分布的 $1-\alpha/2$ 分位数。

2) 总体方差 σ^2 与标准差 σ 的置信区间的求法。

(i) 当总体均值 μ 已知时, 利用 $\chi^2(n)$ 分布性质可知:

$$\frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \sim \chi^2(n)$$

故可以得到 σ^2 的 $1-\alpha$ 置信区间为:

$$\left[\frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi_{1-\alpha/2}^2(n)}, \frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi_{\alpha/2}^2(n)} \right]$$

式中, $\chi_{\alpha/2}^2(n)$ 与 $\chi_{1-\alpha/2}^2(n)$ 分别是 $\chi^2(n)$ 分布的 $\alpha/2$ 分位数与 $1-\alpha/2$ 分位数。两边开方, 可得 σ 的 $1-\alpha$ 的置信区间为:

$$\left[\frac{\sqrt{\sum_{i=1}^n (x_i - \mu)^2}}{\sqrt{\chi_{1-\alpha/2}^2(n)}}, \frac{\sqrt{\sum_{i=1}^n (x_i - \mu)^2}}{\sqrt{\chi_{\alpha/2}^2(n)}} \right]$$

(ii) 当总体均值 μ 未知时, σ^2 的估计常用样本方差 s^2 , 因此从 s^2 的分布来构造置信区间。利用 $\chi^2(n-1)$ 分布可以得到 σ^2 的 $1-\alpha$ 置信区间为:

$$\left[\frac{(n-1)s^2}{\chi_{1-\alpha/2}^2(n-1)}, \frac{(n-1)s^2}{\chi_{\alpha/2}^2(n-1)} \right]$$

式中, $\chi_{\alpha/2}^2(n-1)$ 与 $\chi_{1-\alpha/2}^2(n-1)$ 分别是 $\chi^2(n-1)$ 分布的 $\alpha/2$ 分位数与 $1-\alpha/2$ 分位数。两边开方, 可得 σ 的 $1-\alpha$ 的置信区间为:

$$\left[\frac{s \sqrt{n-1}}{\sqrt{\chi_{1-\alpha/2}^2(n-1)}}, \frac{s \sqrt{n-1}}{\sqrt{\chi_{\alpha/2}^2(n-1)}} \right]$$

以上讨论总结在表 6-2 中, 可供选用。

表 6-2 单个正态总体均值、方差、标准差 $1-\alpha$ 的置信区间

参数	条件	$1-\alpha$ 的置信区间
μ	σ^2 已知	$\bar{x} \pm Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$
μ	σ^2 未知	$\bar{x} \pm t_{1-\alpha/2}(n-1) \frac{s}{\sqrt{n}}$
σ^2	μ 已知	$\left[\frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi_{1-\alpha/2}^2(n)}, \frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi_{\alpha/2}^2(n)} \right]$
σ	μ 已知	$\left[\frac{\sqrt{\sum_{i=1}^n (x_i - \mu)^2}}{\sqrt{\chi_{1-\alpha/2}^2(n)}}, \frac{\sqrt{\sum_{i=1}^n (x_i - \mu)^2}}{\sqrt{\chi_{\alpha/2}^2(n)}} \right]$
σ^2	μ 未知	$\left[\frac{(n-1)s^2}{\chi_{1-\alpha/2}^2(n-1)}, \frac{(n-1)s^2}{\chi_{\alpha/2}^2(n-1)} \right]$
σ	μ 未知	$\left[\frac{s \sqrt{n-1}}{\sqrt{\chi_{1-\alpha/2}^2(n-1)}}, \frac{s \sqrt{n-1}}{\sqrt{\chi_{\alpha/2}^2(n-1)}} \right]$

图 6—3

有一种可用于培训从事设备维修员工的计算机辅助软件,团队对其进行研究,希望了解使用这种计算机辅助软件后,能缩短多少培训维修员工的时间。为此,团队首先要估计这种计算机辅助程序所需培训时间的均值。

假定对 15 位员工用这种计算机辅助程序方法进行培训,样本中每名员工所需培训时间列于表 6—3 中。假设培训时间总体是正态分布,要求培训时间均值的 95% 的置信水平的区间估计。

表 6—3 员工培训时间

员工	天数	员工	天数	员工	天数
1	52	6	59	11	54
2	44	7	50	12	58
3	55	8	54	13	60
4	44	9	62	14	62
5	45	10	46	15	63

下图是利用 MINITAB 软件运算的结果。从 Stat-Basic Statistic-Graphical Summary 入口,只要填写上变量名称,计算机将自动计算出下列结果(如图 6—3 所示)。

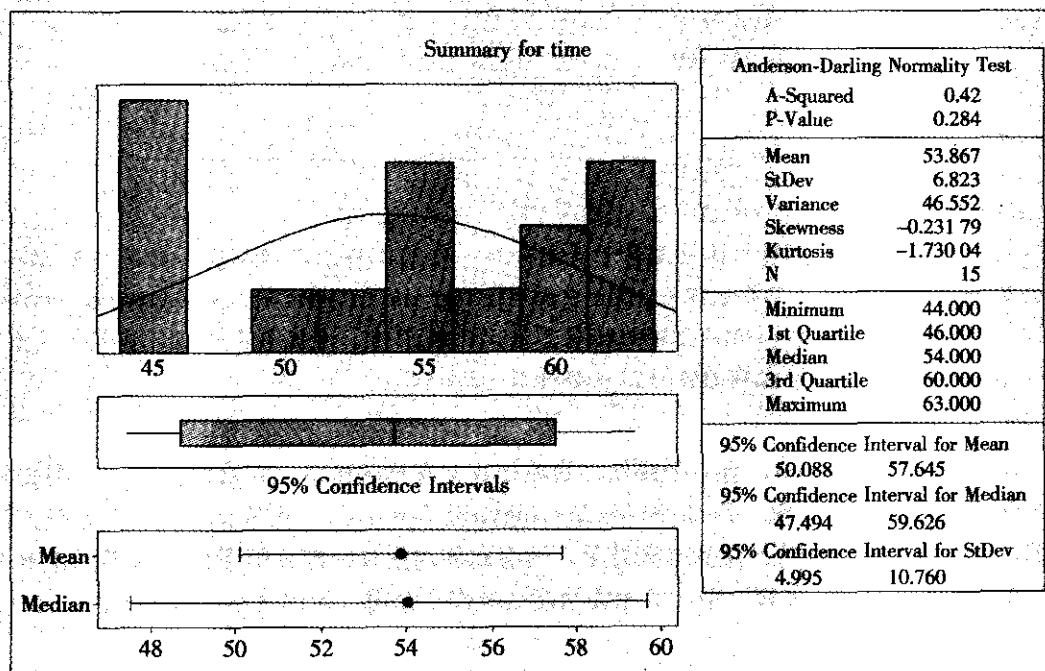


图 6—3 员工培训时间的分布分析

从此图的右侧区域内,可以看到描述性统计量的结果:样本均值为 53.87 天,样本标准差为 6.82 天,均值的 95% 的置信区间为 (50.09, 57.65),标准差的 95% 的置信区间为 (4.995, 10.760)。



(2) 比率 p 的置信区间。设总体 $X \sim b(1, p)$, 样本为 x_1, x_2, \dots, x_n , 样本之和为 k , 样本均值为 $\bar{x} = k/n$, 这便是 p 的点估计, 在样本容量 n 较大 (至少要 $np \geq 5$) 时, 由于 \bar{x} 的分布可以近似看成正态分布 $N(p, p(1-p)/n)$, 因此 p 的 $1-\alpha$ 的置信区间为:

$$\bar{x} \pm Z_{1-\alpha/2} \sqrt{\bar{x}(1-\bar{x})/n}$$

式中, $Z_{1-\alpha/2}$ 是标准正态分布的 $1-\alpha/2$ 分位数。

在某产品使用率情况的顾客调查中, 调查了 400 人, 其中 100 人使用了该产品, 试对该产品使用率 p 作置信水平为 0.95 的区间估计。

$n=400$, $\bar{x}=100/400=0.25$, 在 $\alpha=0.05$ 时, $Z_{0.975}=1.96$

由此得 p 的置信水平为 0.95 的置信区间为:

$$\bar{x} \pm Z_{0.975} \sqrt{\bar{x}(1-\bar{x})/n} = 0.25 \pm 1.96 \times \sqrt{0.25(1-0.25)/400} = 0.25 \pm 0.0424$$

即 $[0.2076, 0.2924]$ 。

用 MINITAB 软件, 从 STAT-Basic Statistics-1Proportion 入口, 在 Option 选项中选择 use test and interval based on normal distribution, 对于软件默认选定的双侧检验不予修改, 则可以直接得到下列结果:

Sample	X	N	Sample p	95% CI	Z-Value	P-Value
1	100	400	0.250 000	(0.207 566, 0.292 434)	-10.00	0.000

即 p 的 95% 置信区间为 $(0.2076, 0.2924)$ 。

注意: 上述置信区间的估计是基于正态分布计算的, 称为近似算法。当不满足 $np \geq 5$ 或需要精确算法时, 不要选择 Option 中的 use test and interval based on normal distribution 即可。关于基于二项式分布的精确算法, 比较烦琐, 这里不再介绍。在实际生产中, 为了得到 p 值较为准确的估计, 一般要求 $np \geq 5$ 。

在 MINITAB 软件中, 有一个操作很简单却非常有用的窗口, 它可以提供一个图文并茂的计算结果, 这就是 Stat-Basic Statistic-Graphical Summary。此窗口不但有直方图、箱线图, 还有重要的描述统计量的计算结果, 还有总体均值及标准差的置信区间等。在本书 5.3 节“数据的收集和整理”中, 对相关功能已经作了介绍, 这里就不再重复了。

6.2 假设检验

假设检验是六西格玛团队项目中应用最多的统计工具。诸如要判断下列结论是否正确: “新员工比老员工得到更多的投诉”, “改进工作后平均产量有提高”, “加工温度为 180 度时比 160 度时垫圈断裂强度要高”等等。由于我们观测到数据总会带有误差, 不能从简

单的样本统计量的结果下定论,必须使用严格的统计假设检验方法才能得出准确的判断结论。本节着重介绍有关假设检验的基本概念和简单的检验方法,这是我们今后几章讲解统计方法的理论基础,大家必须理解掌握。

6.2.1 假设检验的基本概念

1. 假设检验问题

参数估计和假设检验是统计推断的两个重要方面。参数估计是以“数”为其输出结果,而假设检验是以“判断”为其输出结果。为了说明其基本想法先看一个例子。

原来的冷拉钢筋生产线上的钢筋平均抗拉强度为 2 000kg,标准差为 300kg。希望经过调整参数后,钢筋平均抗拉强度能有所提高。项目团队实施改进后抽取了 25 根钢筋,测得钢筋平均抗拉强度为 2 150kg。问:能否断言,钢筋平均抗拉强度确有提高?

解:用 μ 代表总体的钢筋抗拉强度的平均值,这个值是我们不知道的。我们从抽样中得到的只是样本均值,而统计学的目的就是要用样本去推断总体。对于这个问题,我们知道,若总体均值 $\mu > 2\,000$,则认为平均抗拉强度有提高;若均值 $\mu \leq 2\,000$,则认为没有提高。为此可以建立两个命题,在假设检验中称它们为假设(hypothesis),一个为原假设(或零假设, null hypothesis),记为 H_0 ;另一个称为备择假设(或对立假设, alternative hypothesis),记为 H_1 ,即

$$H_0: \mu \leq 2\,000 \quad H_1: \mu > 2\,000$$

问题就是要根据样本的观测值去判断 H_0 是否为真。通常我们总是选要证明的命题作为备择假设,而把正常情况下成立的,一般不需证明而且是不证自明的作为原假设。在对某个参数进行检验时,我们把要证明的、不相等的、有差别的结论作为备择假设;相等的、无差别的、等号成立的结论作为原假设。因而在对于参数的检验问题中,备择假设中永远不会包含等号,只可能是“大于”、“小于”或“不等于”这三种情况之一。

2. 基本概念与步骤

可以把假设检验的步骤归纳为以下几点:

(1) 建立假设。假设检验的第一步便是建立假设,通常需要建立两个假设:原假设 H_0 和备择假设 H_1 。

在对总体均值进行检验时,有三类假设:

$$H_0: \mu = \mu_0 \text{ (或 } \mu \leq \mu_0), H_1: \mu > \mu_0$$

$$H_0: \mu = \mu_0 \text{ (或 } \mu \geq \mu_0), H_1: \mu < \mu_0$$

$$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$$

称前两个为单边假设检验,后一个为双边假设检验。

在例 6—5 中用的是单边检验 $H_0: \mu \leq 2\,000, H_1: \mu > 2\,000$ 。

假设检验的任务便是根据样本 x_1, x_2, \dots, x_n 来判断原假设是否为真。

(2) 选择检验统计量,确定拒绝域的形式。若对总体的均值进行检验,那么我们将用样本均值 \bar{x} 引出检验统计量;若对正态总体的方差进行检验,我们将从样本方差 s^2 引出

检验统计量。

根据统计量的值把整个样本空间分成两个部分：拒绝域 W 与非拒绝域 A 。当样本统计量的值落在拒绝域中就拒绝原假设，否则就无法拒绝原假设。所以在假设检验中我们必须找出拒绝域。

根据备择假设的不同；拒绝域可以是双边的也可以是单边的。在确定了拒绝域的类型后，还要确定临界值 c ，这应根据允许犯错误的概率来确定。

(3) 给出检验中的显著性水平 α 。在对原假设是否成立进行判断时，由于样本的随机性，判断可能产生两类错误，其示意列在表 6—4 中。第 I 类错误是当原假设为真时，由于样本的随机性，使样本观测值落在拒绝域 W 中，从而做出拒绝原假设的决定，这类错误称为第 I 类错误，其发生的概率称为犯第 I 类错误的概率，也称为弃真概率，记为 α ，即 $P_{H_0}(W) = \alpha$ 。第 II 类错误是当原假设为假时，由于样本的随机性，使样本观测值落在非拒绝域 A 中，从而做出无法拒绝原假设的决定，这类错误称为第 II 类错误，其发生的概率称为犯第 II 类错误的概率，也称为取伪概率，记为 β ，即 $P_{H_1}(A) = \beta$ 。

若要求犯第 I 类错误的概率不超过 α ，由此给出的检验称为水平为 α 的检验，称 α 为显著性水平，常取 α 为 0.05，有时也可能取 0.10 等。

表 6—4 假设检验中两类错误

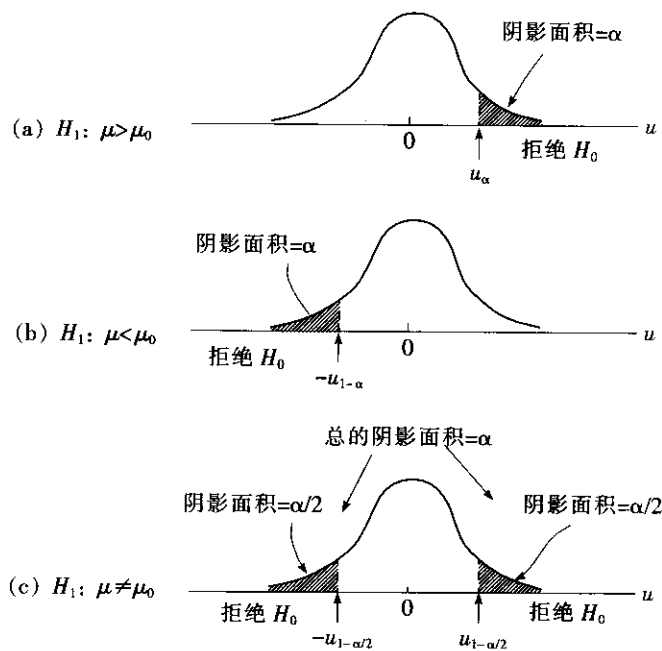
判断结论 \ 总体情况	H_0 是真的	H_1 是真的
	接受 H_0	第 II 类错误
	拒绝 H_0	第 I 类错误

以例 6—5 为例，因为要求检验钢筋平均抗拉强度是否超过 2 000kg，因此 H_0 的要求就是“钢筋平均抗拉强度未提高”。如果 $H_0: \mu \leq 2\,000$ 被拒绝，则表明钢筋平均抗拉强度比原来的 2 000kg 确实有提高。为了真正理解假设检验结论的含义，我们应该具体地理解犯两类错误的实际意义。

关于第一类错误的说明：如果钢筋平均抗拉强度为 $\mu \leq 2\,000$ ，平均抗拉强度并未提高，但我们拒绝了它，即把“未提高”当作了“有提高”。一般来说就是，当 H_0 成立时，我们却拒绝了 H_0 ，这就是第 I 类错误。犯第 I 类错误的概率我们用 α 来表示，通常我们取 $\alpha = 0.05$ 作为犯第 I 类错误的风险概率。

关于第 II 类错误的说明：如果钢筋平均抗拉强度比原来真有提高，这时钢筋平均抗拉强度已经不是原来的 2 000kg 了，但我们没有拒绝 H_0 ，误认为没提高，即把“已提高”误认为“未提高”。一般来说就是，当 H_0 不成立时，我们却没有拒绝 H_0 ，这就是第 II 类错误。犯第 II 类错误的概率我们用来 β 表示。通常我们取 $\beta = 0.10$ 。

(4) 给出临界值，确定拒绝域。有了显著性水平 α 后，可以根据给定的检验统计量的分布，查表得到临界值，从而确定具体的拒绝域。在不同的备择假设下，拒绝域、临界值与显著性水平 α 的关系是不同的，其示意图见图 6—4。



(图中曲线为 $N(0, 1)$ 的密度函数曲线)

图 6—4 备择假设、拒绝域和显著性水平

(5) 根据样本观测值, 计算检验统计量的值。收集样本数据, 计算检验统计量的值。

(6) 根据检验统计量的值是否落在拒绝域中做出判断。

1) 将检验统计量的值与拒绝域的临界值相比较, 当它落在拒绝域中就做出拒绝原假设的结论, 否则就做出不能拒绝原假设的结论。

2) 由检验统计量计算 p 值。所谓 p 值, 就是当原假设成立时, 出现目前状况的概率 (严格说是: 当原假设成立时, 出现目前状况或对原假设更不利状况, 即对备择假设更有利状况的概率)。当这个概率很小时 (例如小于 0.05), 这个结果在原假设成立的条件下就不该在一次试验中出现; 但现在它确实出现了, 因此我们有理由认为“原假设成立”的这个前提是错的, 因而应该拒绝原假设, 接受备择假设。因此可以有个最一般的规则: 如果 $p < \alpha$, 则拒绝原假设。目前大多数统计软件都提供了与假设检验对应的 p 值, 不必再查统计表确定拒绝域就可以根据 p 值做出判断结论。

3) 根据样本观测值可以得到总体参数的置信区间, 如果原假设的参数值未落入此置信区间, 就做出拒绝原假设的结论, 否则就做出不能拒绝原假设的结论。目前大多数统计软件都提供了相应的置信区间, 不必自己计算, 因此用这个方法判断也很方便。

例 6—5

设立原假设和备择假设。 $H_0: \mu \leq \mu_0 = 2000$ $H_1: \mu > 2000$

由于钢筋平均抗拉强度 X 为正态分布, 故 \bar{x} 也服从正态分布, 且标准差已知, 检验统计量为:



$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

在 H_0 成立时, 它应该为标准正态分布。当总体均值 μ 大于 μ_0 时, Z 有变大的趋势, 拒绝域应该取右侧, 即当样本值代入后计算出的 Z 值太大时应该拒绝原假设。取显著性水平 $\alpha=0.05$ 时, 由于 Z 的右侧 α 分位数, 也即是 Z 的左侧 $1-\alpha$ 分位数是 $Z_{1-\alpha}=1.645$, 因此, 检验统计量 Z 的拒绝域为: $Z \geq Z_{1-\alpha}$, 也就是说 $Z \geq 1.645$, 则拒绝 H_0 。

计算检验统计量:

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{2\,150 - 2\,000}{300 / \sqrt{25}} = 2.5$$

由于 $Z=2.5 > 1.645$, 故拒绝 H_0 。

用 MINITAB 软件可以得到下述结果:

One-Sample Z					
Test of mu = 2 000 vs > 2 000					
The assumed standard deviation = 300					
N	Mean	SE Mean	95% Lower Bound	Z	P
25	2 150.00	60.00	2 051.31	2.50	0.006

对于 MINITAB 输出的结果必须学会解释清楚。这里首先给出样本量为 25, 样本均值 \bar{x} 为 2 150, 总体标准差 σ 为 300, 均值的标准误 (即 $\frac{\sigma}{\sqrt{n}}$) 为 60。除此之外, 这里还提供了三处可以直接判断原假设是否成立的计算结果。

(1) $Z=2.50$, 应该拒绝原假设。前面对此已经作了分析。

(2) $p\text{-value}=0.006$ 。判断法则是: 当 $p\text{-value}$ 小于给定的显著性水平 α 时, 则拒绝原假设。本例, $p\text{-value}$ 比 0.05 小, 故应该拒绝原假设。

(3) 置信区间。在本例中, 样本均值为 2 150, 我们以 95% 的把握断言, 总体的均值应大于 2 051.31 (即 95% 置信下限), 因为原假设的均值 2 000 并未落入此置信区间内, 因此应该拒绝原假设。

这三处的结论当然是一致的, 但它们有不同的侧重点。 $p\text{-value}$ 着眼于考虑这个事件出现“多么不寻常”, 以此来拒绝原假设; 置信区间则着重于考虑原假设的参数值“离我们的 $1-\alpha$ 的置信区间有多远”, 以此来拒绝原假设。出发点略有不同, 而结论是殊途同归的。我们分析问题时, 应全面掌握这些信息, 以更好地理解所得到的结论。

3. 对总体均值进行假设检验时样本容量的确定

如果我们对总体均值进行假设检验, 使用者通常事先指定显著性水平以确定发生第 I 类错误的概率。以此为基础, 再通过控制样本容量, 也可以对发生第 II 类错误的概率进行控制。下面, 介绍在总体均值的单边检验中如何确定样本容量。

$$H_0: \mu \leq \mu_0 \quad H_1: \mu > \mu_0$$

式中, μ_0 为原假设总体均值, 则总体均值的单边假设检验中所推荐的样本容量

$$n = \left[\frac{(Z_{1-\alpha} + Z_{1-\beta})^2 \cdot \sigma^2}{(\mu_0 - \mu_a)^2} \right]$$

式中, $Z_{1-\alpha}$ 为标准正态分布的 $1-\alpha$ 分位数; $Z_{1-\beta}$ 为标准正态分布的 $1-\beta$ 分位数; σ 为总体的标准差; μ_0 为原假设总体的均值; μ_a 为备则假设总体的均值。

记号 $[A]$ 是取大于或等于 A 值的最小整数, 例如 $[3.9]=4$, $[4.1]=5$, $[4]=4$ 等。

注意: 在总体均值的双边假设检验中, 只需以 $Z_{1-\alpha/2}$ 代替 $Z_{1-\alpha}$ 即可。

由于两类错误造成的损失是不同类型的, 其严重性也是不同的, 因此不同的人可能会对两类错误的概率做出不同的限制。但是, 在确定样本容量之前, 必须对发生两类错误所允许的概率做出明确规定。

在假设检验中还常用到一个名词: 检出力 (power of test)。它的定义就是 $1-\beta$, 相当于“在备择假设成立时不犯第 II 类错误”的概率, 或说“在备择假设成立时拒绝原假设”的概率。

在例 6—5 中, 如果我们限定 $\alpha=0.05$, $\beta=0.10$ 。利用标准正态概率分布, 我们得到 $Z_{0.95}=1.645$, $Z_{0.90}=1.28$, 从关于错误的概率的说明中, 我们已知 $\mu_0=2\ 000$, $\mu_a=2\ 150$, 总体标准差的值已知, $\sigma=300$ 。因此, 建议使用的样本容量为:

$$n = \frac{(1.645 + 1.28)^2 \times 300^2}{(2\ 150 - 2\ 000)^2} = 34.2$$

取整后, 样本容量为 35。可见, 我们这个问题检验结果可以保证犯第 II 类错误的概率 $\beta < 0.10$ 。

用 MINITAB 软件可以从 STAT-Power and sample size-1sample z 入口, 直接得到下列结果 (注意在计算中不输入 β , 而是输入检出力 $1-\beta$)。

Power and Sample Size

1-Sample Z Test

Testing mean = null (versus > null)

Calculating power for mean = null + difference

Alpha = 0.05 Assumed standard deviation = 300

Difference	Sample Size	Target Power	Actual Power
150	35	0.9	0.905 440

这里得到的建议样本量也是 35。检出力的目标是 0.9, 但由于我们对样本量 34.2 向上取了整, 增大到了 35, 因此实际检出力为 0.905 44, 稍高于设定目标 (即犯第 II 类错误的概率稍小于 0.1)。

当 $n=35$ 时, 由于第 I 类错误和第 II 类错误的概率已经控制在可接受的水平内, 所以在假设检验中目前质量控制管理者只要在收集数据后, 在拒绝 H_0 和接受 H_0 之间作出判断就行了。

在使用 MINITAB 计算样本量时, 我们可以观察到 α , β 和样本容量 n 三者中有两者



已知时,即可计算得到第三者。如果我们仍然希望样本量为 25,我们可以计算出此时犯第Ⅱ类错误的概率。对于本例情况,我们可以假定对立假设的总体均值就是抽取的样本均值,调用 MINITAB 程序,可得下列结果。

1-Sample Z Test

Testing mean = null (versus $>$ null)

Calculating power for mean = null + difference

Alpha = 0.05, Assumed standard deviation = 300

Difference	Sample Size	Power
150	25	0.803 765

这时的检出力为 0.803 765,也就是说,此时犯第Ⅱ类错误的概率约为 0.2。

从软件计算结果中很容易看出:对于给定的显著性水平 α ,增大样本容量将会减小 β ;对于给定的样本容量,减小 α 将会使 β 增大,增大 α 将会使 β 减小。

当样本容量 n 给定时,我们不能毫无必要地选择太小的显著性水平 α ,因为选择较小的显著性水平 α 意味着将使发生第Ⅱ类错误的风险增大,通常我们都固定选取 α 为 0.05。当样本容量 n 给定, α 也取定时,两种检验方法如何比较优劣呢?这就要比较它们犯第Ⅱ类错误的概率。所谓一种方法比另一种方法“灵敏”,指的就是犯第Ⅱ类错误的概率 β 小,或说检出力 $1-\beta$ 较高,即当备择假设成立时能够很明确地拒绝原假设而接受备择假设。

6.2.2 均值、方差和比例的假设检验

1. 单个正态总体均值、方差的显著性水平为 α 的检验

设总体分布为 $N(\mu, \sigma^2)$,从总体中抽取的样本为 x_1, x_2, \dots, x_n ,样本均值为 \bar{x} ,样本的方差为 s^2 ,样本标准差为 s 。

(1) 关于 μ 的显著性水平为 α 的检验。

1) 当 σ 已知时关于正态均值 μ 常用的三对假设为:

$$\left. \begin{array}{l} H_0: \mu \leq \mu_0, H_1: \mu > \mu_0 \\ H_0: \mu \geq \mu_0, H_1: \mu < \mu_0 \end{array} \right\} \text{单边假设检验问题}$$

$$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0 \text{ 双边假设检验问题}$$

检验统计量都用统计量

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

在 H_0 成立,即 $\mu = \mu_0$ 时, $Z \sim N(0, 1)$ 。

给出显著性水平 α ,定出拒绝域 W :

$$W = \{Z \geq Z_{1-\alpha}\}, W = \{Z \leq Z_{\alpha}\}, W = \{|Z| \geq Z_{1-\alpha/2}\}$$

2) 当 σ 未知时,可以用 σ 的估计 s 代替,大样本 ($n \geq 30$) 时,采用检验统计量 $Z =$

$$\frac{\bar{x} - \mu_0}{s / \sqrt{n}}, \text{拒绝域分别为:}$$

$$W = \{Z \geq Z_{1-\alpha}\}, W = \{Z \leq Z_{\alpha}\}, W = \{|Z| \geq Z_{1-\alpha/2}\}$$

小样本 ($n < 30$) 时, 采用检验统计量 $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$, 在 H_0 成立条件下, $t \sim t(n-1)$ 。

相应于上述三个检验问题的拒绝域改用 t 分布获得, 分别为:

$$W = \{t \geq t_{1-\alpha}(n-1)\}, W = \{t \leq t_{\alpha}(n-1)\}, W = \{|t| \geq t_{1-\alpha/2}(n-1)\}$$

其中 $t_{1-\alpha/2}(n-1)$ 是自由度为 $n-1$ 的 t 分布的 $1-\alpha/2$ 分位数, 其他意义相同。

(2) 关于 σ^2 的显著性水平为 α 的检验。可以采用检验统计量 $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$, 在不同的检验问题中拒绝域如下确定:

若原假设与备择假设为: $H_0: \sigma^2 \leq \sigma_0^2$ (或 $\sigma^2 = \sigma_0^2$), $H_1: \sigma^2 > \sigma_0^2$, 那么拒绝域为 $\{\chi^2 \geq \chi_{1-\alpha}^2(n-1)\}$ 。

若原假设与备择假设为: $H_0: \sigma^2 \geq \sigma_0^2$ (或 $\sigma^2 = \sigma_0^2$), $H_1: \sigma^2 < \sigma_0^2$, 那么拒绝域为 $\{\chi^2 \leq \chi_{\alpha}^2(n-1)\}$ 。

若原假设与备择假设为: $H_0: \sigma^2 = \sigma_0^2$, $H_1: \sigma^2 \neq \sigma_0^2$, 那么拒绝域为 $\{\chi^2 \leq \chi_{\alpha/2}^2(n-1) \text{ 或 } \chi^2 \geq \chi_{1-\alpha/2}^2(n-1)\}$, 这里 $\chi_{\alpha/2}^2(n-1)$ 是自由度为 $n-1$ 的 χ^2 分布的 $\alpha/2$ 分位数, 其他意义相同。

将上述结果列在表 6-5 中以便查用, 通常将 Z , t , χ^2 作检验统计量的检验称为 Z , t , χ^2 检验。

表 6-5 单个正态总体均值、方差的显著性水平为 α 的检验

检验法	条件	H_0	H_1	检验统计量	拒绝域
Z 检验	σ 已知	$\mu \leq \mu_0$	$\mu > \mu_0$	$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	$\{Z \geq Z_{1-\alpha}\}$
		$\mu \geq \mu_0$	$\mu < \mu_0$		$\{Z \leq Z_{\alpha}\}$
		$\mu = \mu_0$	$\mu \neq \mu_0$		$\{ Z \geq Z_{1-\alpha/2}\}$
Z 检验	σ 未知 大样本 $n \geq 30$	$\mu \leq \mu_0$	$\mu > \mu_0$	$Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$\{Z \geq Z_{1-\alpha}\}$
		$\mu \geq \mu_0$	$\mu < \mu_0$		$\{Z \leq Z_{\alpha}\}$
		$\mu = \mu_0$	$\mu \neq \mu_0$		$\{ Z \geq Z_{1-\alpha/2}\}$
t 检验	σ 未知 小样本 $n < 30$	$\mu \leq \mu_0$	$\mu > \mu_0$	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$\{t \geq t_{1-\alpha}(n-1)\}$
		$\mu \geq \mu_0$	$\mu < \mu_0$		$\{t \leq t_{\alpha}(n-1)\}$
		$\mu = \mu_0$	$\mu \neq \mu_0$		$\{ t \geq t_{1-\alpha/2}(n-1)\}$
χ^2 检验	μ 未知	$\sigma^2 \leq \sigma_0^2$ $\sigma^2 \geq \sigma_0^2$ $\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\{\chi^2 \geq \chi_{1-\alpha}^2(n-1)\}$ $\{\chi^2 \leq \chi_{\alpha}^2(n-1)\}$ $\left\{ \begin{array}{l} \chi^2 \leq \chi_{\alpha/2}^2(n-1) \\ \text{或 } \chi^2 \geq \chi_{1-\alpha/2}^2(n-1) \end{array} \right\}$

例 6-4

某零件,其厚度在正常生产下服从 $N(0.13, 0.015^2)$ 。某日在生产的产品中抽查了 10 件,其观测值为: 0.112, 0.130, 0.129, 0.152, 0.138, 0.118, 0.151, 0.128, 0.158, 0.142。发现平均厚度稍有变化,如果标准差不变,试问生产是否正常? ($\alpha=0.05$)

解: (1) 建立假设。 $H_0: \mu=0.13$, $H_1: \mu \neq 0.13$

(2) 由于 σ 已知,故选用 Z 检验。

(3) 根据显著性水平 $\alpha=0.05$ 及备择假设知拒绝域为:

$$\{|Z| \geq Z_{1-\alpha/2}\} = \{|Z| \geq 1.96\}$$

(4) 由样本观测值,求得

$$Z = \frac{\bar{x} - 0.13}{0.015/\sqrt{10}} = \frac{0.1358 - 0.13}{0.015/\sqrt{10}} = 1.22$$

由于样本观测值未落在拒绝域中,所以不能拒绝原假设,可以认为该天生产正常。

由 MINITAB 计算,从 STAT-BasicStatistics-1Sample Z 入口,填写变量名后,要特别注意在“Options”窗口选择双侧检验“Not Equal”,得到结果如下:

Test of mu = 0.13 vs mu not = 0.13				
The assumed sigma = 0.015				
Variable	N	Mean	StDev	SE Mean
x	10	0.1358	0.01511	0.00474
Variable	95.0% CI		Z	P
x	(0.12650, 0.14510)		1.22	0.221

因为 $p=0.221>0.05$,则不能拒绝 H_0 ,即认为平均厚度变化到 0.1358 与 0.13 并无显著差别,即这批产品平均厚度合格。由于观测的样本均值是 0.1358,置信区间为 (0.127, 0.145), $\mu=0.13$ 落入此置信区间,因而以 0.05 的显著水平接受原假设。MINITAB 还给出非常直观的图形(见图 6-5)帮助判断:当 H_0 所对应的均值未落入

Histogram of X
(with Ho and 95% Z-confidence interval for the mean, and sigma=0.015 000)

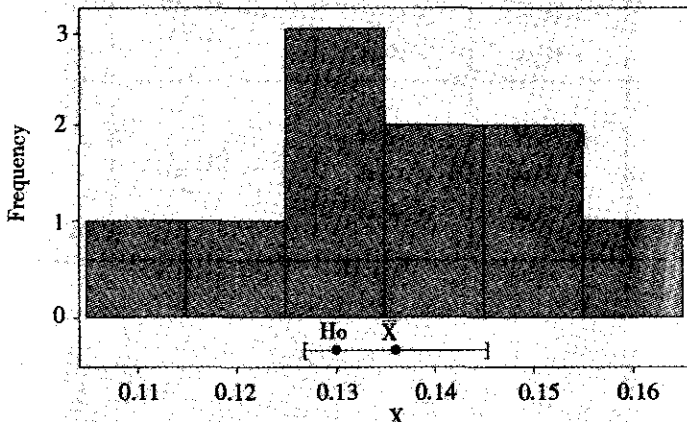


图 6-5 假设检验用的直方图和置信区间

均值的置信区间时,则拒绝原假设;否则不能拒绝原假设。本例的结论是:不能拒绝原假设 H_0 ,即平均厚度与 0.13 无显著差别。

例 6-7

某宾馆六西格玛团队评定某项服务的等级,最大可能的级别为 10。团队规定服务等级的总体均值“达到 7”的服务项目可以接受,总体均值“超过 7”的认为有提高。随机调查 12 位顾客,要求他们对此项服务评定等级,得到 12 个级别数据为:7, 8, 10, 8, 6, 9, 6, 7, 7, 8, 9 和 8。假使总体级别近似服从正态分布,在显著性水平 $\alpha=0.05$ 的情况下,问服务项目确实比规定的可接受标准有提高吗?

解:建立假设。 $H_0: \mu=7, H_1: \mu>7$

如果拒绝 H_0 ,则可以认为服务项目确实比规定标准有提高。由于是小样本,并且 σ 未知,故选用 1 sample t 检验。

下面是 MINITAB 软件计算的 1 sample t 的结果:

One-Sample T: 项目级别				
Test of mu = 7 vs mu > 7				
Variable	N	Mean	StDev	SE Mean
级别	12	7.750	1.215	0.351
Variable	95.0% Lower Bound	T	P	
级别	7.120	2.14	0.028	

计算结果显示 $\bar{x}=7.750, s=1.215, t=2.14$, 而 $t_{1-\alpha}(n-1)=t_{0.95}(11)=1.796$ (t 分布分位数表), 由于 $t=2.14>1.796$, 故拒绝 H_0 , 可以认定服务项目确实比规定标准有提高。

同样, 因为 $p=0.028<0.05$, 则拒绝 H_0 , 获得同样结论。

例 6-8

某种导线的电阻服从 $N(\mu, \sigma^2)$, μ 未知, 但要求千米长导线电阻的标准差不得超过 0.005Ω 。现随机抽取了 9 根千米长导线, 测得其电阻数据为:

10.010, 9.997 7, 9.989 8, 10.000 8, 10.002 4, 10.004 8, 9.999 6, 10.007 8, 9.993 2
测得样本的标准差为 $s=0.006 6$, 试问在 $\alpha=0.05$ 的显著水平上能否认为该批导线电阻标准差发生变化?

解: (1) 建立假设。

$$H_0: \sigma=0.005, H_1: \sigma \neq 0.005$$

(2) 由于 μ 未知, 故选用 χ^2 检验。

(3) 根据显著性水平 $\alpha=0.05$ 及备择假设知拒绝域为:

$$\{\chi^2 \leq \chi_{\frac{\alpha}{2}}^2(n-1)\} \text{ 或 } \{\chi^2 \geq \chi_{1-\frac{\alpha}{2}}^2(n-1)\} = \{\chi^2 \leq 2.18\} \text{ 或 } \{\chi^2 \geq 17.53\}$$

(4) 由样本观测值, 求得

$$\chi^2 = \frac{8 \times 0.0066^2}{0.005^2} = 13.94$$

由于样本统计量的值未落在拒绝域中, 所以不能拒绝原假设, 认为该批导线电阻标准差无显著变化。

在用 MINITAB 软件计算时, 我们可以借用 STAT-Basic Statistics-Graphical Summary 入口, 求出标准差的 95% 的双侧置信区间为 (0.004, 0.013), 而 0.005 落入此置信区间, 因此不能拒绝原假设, 认为该批导线电阻标准差合格。

一般来说, 估计总体均值大约 $n=15$ 就够了, 而估计标准差大约要 $n=30$ 才能估计得比较好。本例 $n=9$, 对于估计样本均值已经不够, 当然对于估计标准差就更不够了, 所以才导致最终结果中其双侧置信区间太宽了。

2. 两个总体均值、方差的显著性水平为 α 的检验

设有两个独立总体, $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ 。从总体 X 中抽取的样本 x_1, x_2, \dots, x_n , 样本均值为 \bar{x} , 样本方差为 s_x^2 , 样本标准差为 s_x ; 从总体 Y 中抽取的样本为 y_1, y_2, \dots, y_m , 样本均值为 \bar{y} , 样本方差为 s_y^2 , 样本标准差为 s_y 。

(1) 关于两个正态总体均值的显著性水平为 α 的检验。

1) 当 σ_1, σ_2 已知时, 可以采用检验统计量 $Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$ 。

对检验问题 $H_0: \mu_1 \leq \mu_2$ (或 $\mu_1 = \mu_2$), $H_1: \mu_1 > \mu_2$, 拒绝域为 $\{Z \geq Z_{1-\alpha}\}$ 。

对检验问题 $H_0: \mu_1 \geq \mu_2$ (或 $\mu_1 = \mu_2$), $H_1: \mu_1 < \mu_2$, 拒绝域为 $\{Z \leq Z_\alpha\}$ 。

对检验问题 $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$, 拒绝域为 $\{|Z| \geq Z_{1-\alpha/2}\}$ 。

2) 当 σ_1, σ_2 未知, 但 $\sigma_1 = \sigma_2$ 时, 可以采用检验统计量 $t = \frac{\bar{x} - \bar{y}}{s_w \sqrt{\frac{1}{n} + \frac{1}{m}}}$, 其中:

$$s_w^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$$

对上述三个检验问题的拒绝域可以用 t 分布的分位数得到, 它们分别是: $\{t \geq t_{1-\alpha}(n+m-2)\}$, $\{t \leq t_\alpha(n+m-2)\}$, $\{|t| \geq t_{1-\alpha/2}(n+m-2)\}$ 。

3) 当 σ_1, σ_2 未知, 但 n 与 m 都较大时, 可以采用检验统计量 $Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$, 对上述

三个检验问题的拒绝域仍可以用 $N(0, 1)$ 分布的分位数获得, 它们分别是: $\{Z \geq Z_{1-\alpha}\}$, $\{Z \leq Z_\alpha\}$, $\{|Z| \geq Z_{1-\alpha/2}\}$ 。

(2) 关于两个正态总体方差的显著性水平为 α 的检验。可以采用检验统计量

$$F = \frac{s_x^2}{s_y^2}$$

1) 检验问题 $H_0: \sigma_1^2 = \sigma_2^2$, $H_1: \sigma_1^2 \neq \sigma_2^2$

拒绝域为: $\{F \leq F_{\alpha/2}(n-1, m-1)\}$ 或 $\{F \geq F_{1-\alpha/2}(n-1, m-1)\}$

2) 检验问题 $H_0: \sigma_1^2 \leq \sigma_2^2$ (或 $\sigma_1^2 = \sigma_2^2$), $H_1: \sigma_1^2 > \sigma_2^2$

拒绝域为: $\{F \geq F_{1-\alpha}(n-1, m-1)\}$

3) 检验问题 $H_0: \sigma_1^2 \geq \sigma_2^2$ (或 $\sigma_1^2 = \sigma_2^2$), $H_1: \sigma_1^2 < \sigma_2^2$

拒绝域为: $\{F \leq F_{\alpha}(n-1, m-1)\}$

将上述关于两个总体均值、方差的显著性水平为 α 的检验列在下表中, 同样将 Z, t, F 作检验统计量的检验称为 Z, t, F 检验。

表 6—6 两正态总体均值、方差的显著性水平为 α 的检验

检验法	条件	H_0	H_1	检验统计量	拒绝域
Z 检验	σ_1, σ_2 已知	$\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$ $\mu_1 = \mu_2$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$	$\{Z \geq Z_{1-\alpha}\}$ $\{Z \leq Z_{\alpha}\}$ $\{ Z \geq Z_{1-\alpha/2}\}$
t 检验	$\sigma_1 = \sigma_2$ 未知	$\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$ $\mu_1 = \mu_2$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	$t = \frac{\bar{x} - \bar{y}}{s_w \sqrt{\frac{1}{n} + \frac{1}{m}}}$	$\{t \geq t_{1-\alpha}(n+m-2)\}$ $\{t \leq t_{\alpha}(n+m-2)\}$ $\{ t \geq t_{1-\alpha/2}(n+m-2)\}$
近似 Z 检验	σ_1, σ_2 未知 m, n 大样本	$\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$ $\mu_1 = \mu_2$	$\mu_1 > \mu_2$ $\mu_1 < \mu_2$ $\mu_1 \neq \mu_2$	$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$	$\{Z \geq Z_{1-\alpha}\}$ $\{Z \leq Z_{\alpha}\}$ $\{ Z \geq Z_{1-\alpha/2}\}$
F 检验	μ_1, μ_2 未知	$\sigma_1^2 \leq \sigma_2^2$ $\sigma_1^2 \geq \sigma_2^2$ $\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$F = \frac{s_x^2}{s_y^2}$	$\{F \geq F_{1-\alpha}(n-1, m-1)\}$ $\{F \leq F_{\alpha}(n-1, m-1)\}$ $\left\{ \begin{array}{l} F \leq F_{\alpha/2}(n-1, m-1) \\ \text{或 } F \geq F_{1-\alpha/2}(n-1, m-1) \end{array} \right\}$

例 6—9

在改进工艺前后, 各测量了若干钢条的抗剪强度, 数据如下:

改进后: 525, 531, 518, 533, 546, 524, 521, 533, 545, 540

改进前: 521, 525, 533, 525, 517, 514, 526, 519

问: 可以认为改进工艺后钢条的平均抗剪强度有提高吗?

解: 设在改进后钢条的抗剪强度 $X \sim N(\mu_1, \sigma_1^2)$, 改进前钢条的抗剪强度 $Y \sim N(\mu_2, \sigma_2^2)$ 。从两组产品中随机抽取 $n=10, m=8$ 的样品, 测得其抗剪强度的均值和方差分别为:

$$\bar{x} = 531.60, s_x^2 = (9.78)^2$$

$$\bar{y} = 522.50, s_y^2 = (6)^2$$

首先在 $\alpha=0.05$ 的显著性水平上检验改进前后的方差是否相等, 若可以认为相等的话, 再在 $\alpha=0.05$ 的显著性水平上进一步检验改进后均值是否增大。

(1) 首先检验方差是否相等。

1) 设 $H_0: \sigma_1^2 = \sigma_2^2, H_1: \sigma_1^2 \neq \sigma_2^2$

2) 用 F 检验。

3) 据显著性水平 $\alpha=0.05$ 及备择假设可知拒绝域为:

$$\{F \leq F_{\alpha/2}(n-1, m-1) \text{ 或 } F \geq F_{1-\alpha/2}(n-1, m-1)\} \\ = \{F \leq 1/4.20 \text{ 或 } F \geq 4.82\}$$

4) 由样本观测值, 求得

$$F = \frac{s_x^2}{s_y^2} = \frac{(9.78)^2}{(6)^2} = 2.66$$

由于样本观测值未落在拒绝域中, 所以不能拒绝原假设, 可以认为改进前后的方差相等。

(2) 再检验均值是否有提高。

1) 设 $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 > \mu_2$

2) 由于两总体方差相等, 但未知, 故选用 t 检验。

3) 据显著性水平 $\alpha=0.05$ 及备择假设知拒绝域为:

$$\{t \geq t_{1-\alpha}(n+m-2)\} = \{t \geq 1.746\}$$

4) 由样本观测值知:

$$\bar{x} - \bar{y} = 531.45 - 522.44 = 9.01$$

$$s_w = \sqrt{\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}} = \sqrt{\frac{9 \times (9.84)^2 + 7 \times (5.88)^2}{10+8-2}} = 8.34$$

$$\text{则 } t = \frac{\bar{x} - \bar{y}}{s_w \sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{9.01}{8.34 \times \sqrt{1/10 + 1/8}} = 2.28, \text{ 由于统计量的值落在拒绝域中, 应拒}$$

绝 H_0 , 即可认为改进后的钢条的平均抗剪强度确有提高。

用 MINITAB 软件, 先检验方差相等性, 从 STAT-Basic statistics-2 Variances 入口, 即可得下列图形 (见图 6-6), F 检验的 p -value 为 0.188, 没有充足的证据判定两总体方差不等, 可认为两总体方差相等。

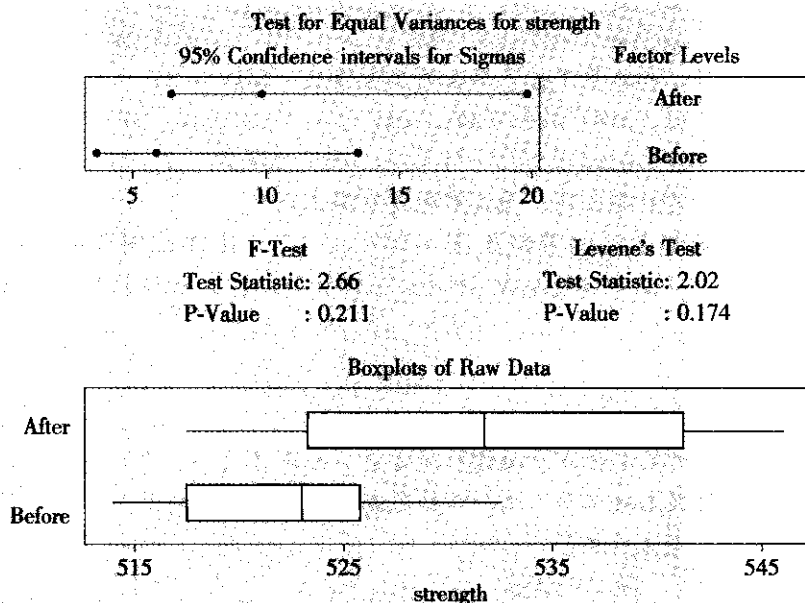


图 6-6 两总体方差相等性检验

再检验均值，从 STAT-Basic Statistics-2 Sample t 入口，选定单侧检验作为备择假设，圈定“两总体方差相等”(Assume Equal Variances)，即可得到下列结果。

Two-Sample T-Test and CI: strength, time				
Two-sample T for strength				
time	N	Mean	StDev	SE Mean
After	10	531.60	9.78	3.1
Before	8	522.50	6.00	2.1
Difference = mu (After) - mu (Before)				
Estimate for difference: 9.10				
95% lower bound for difference: 2.19				
T-Test of difference = 0 (vs >); T-Value = 2.30 P-Value = 0.018 DF = 16				
Both use Pooled StDev = 8.34				

其检验结果的示意图见图 6—7。

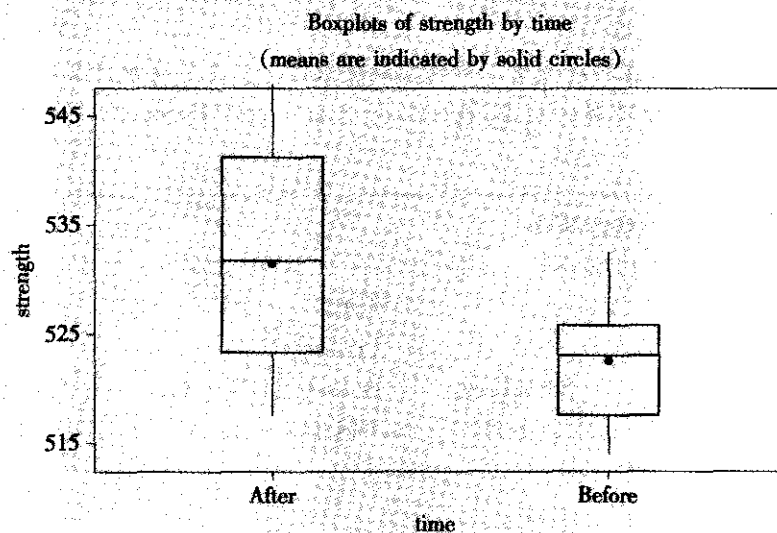


图 6—7 改进工艺前后钢条的抗剪强度比较图

3. 比率 p 的假设检验

(1) 单个比率 p 的显著性水平为 α 的检验。设样本 x_1, x_2, \dots, x_n 来自二点分布总体 X , $X \sim b(1, p)$, 在小样本场合只能使用精确的二项分布计算 (计算机软件一般都可以给出结果); 在大样本场合可以用近似的 z 检验。关于参数 p 检验的近似正态统计量为

$$Z = \frac{\bar{x} - p_0}{\sqrt{p_0(1-p_0)/n}}, \text{ 显著性水平为 } \alpha \text{ 的拒绝域如表 6—7 所示。}$$

表 6—7 单个比率 p 的显著性水平为 α 的检验

检验法	H_0	H_1	检验统计量	拒绝域
Z 检验	$p \leq p_0$ $p \geq p_0$ $p = p_0$	$p > p_0$ $p < p_0$ $p \neq p_0$	$Z = \frac{\bar{x} - p_0}{\sqrt{p_0(1-p_0)/n}}$	$\{Z \geq Z_{1-\alpha}\}$ $\{Z \leq Z_\alpha\}$ $\{ Z \geq Z_{1-\alpha/2}\}$

某厂规定产品必须经过检验合格后才能出厂, 其不合格品率 p_0 不得超过 5%。现从一批产品中随机抽取 200 个进行检验, 发现有 16 个不合格品, 问该批产品能否出厂? (取 $\alpha=0.05$)

解: (1) 假设 $H_0: p \leq 0.05$, $H_1: p > 0.05$

(2) 采用近似算法, 选用近似 Z 检验。

(3) 根据显著性水平 $\alpha=0.05$ 及备择假设知拒绝域为: $\{Z \geq Z_{1-\alpha}\} = \{Z \geq 1.645\}$

(4) 由样本观测值, 求得

$$Z = \frac{16/200 - 0.05}{\sqrt{0.05 \times 0.95/200}} = 1.95$$

由于统计量的值落在拒绝域中, 所以拒绝原假设, 不允许这批产品出厂。

用 MINITAB 软件计算, 从 STAT-Basic Statistics-1 Proportion 入口, 可得下列结果:

Test and CI for One Proportion						
Test of $p = 0.05$ vs $p > 0.05$						
				95% Lower		
Sample	X	N	Sample p	Bound	Z-Value	P-Value
1	16	200	0.080 000	0.048 446	1.95	0.026

(2) 两个比率的显著性水平为 α 的检验。设样本 x_1, x_2, \dots, x_n 来自二点分布总体 X , $X \sim b(1, p_1)$, 样本 y_1, y_2, \dots, y_m 来自二点分布总体 Y , $Y \sim b(1, p_2)$, 两样本独立, 要对参数 p_1 与 p_2 进行比较, 记 $\hat{p}_1 = \bar{x}$, $\hat{p}_2 = \bar{y}$, $\hat{p} = \frac{n\hat{p}_1 + m\hat{p}_2}{n+m}$ 。在小样本情况下, 只能借助于计算机使用精确检验方法; 在大样本情况下, 可以使用检验的近似正态方法, 其统计量为:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\hat{p}(1-\hat{p})}}$$

在不同的假设下, 显著性水平为 α 的拒绝域如表 6—8 所示。

表 6-8 两个比率的显著性水平为 α 的检验

检验法	H_0	H_1	检验统计量	拒绝域
Z 检验	$p_1 \leq p_2$	$p_1 > p_2$	$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{(\frac{1}{n} + \frac{1}{m})\hat{p}(1-\hat{p})}}$	$\{Z \geq Z_{1-\alpha}\}$
	$p_1 \geq p_2$	$p_1 < p_2$		$\{Z \leq Z_\alpha\}$
	$p_1 = p_2$	$p_1 \neq p_2$		$\{ Z \geq Z_{1-\alpha/2}\}$

用 A 与 B 两种不同的方法制造某种零件, 从各自制造的零件中分别随机抽取 100 个, 其中 A 中有 10 个不合格品, B 中有 6 个不合格品。在 $\alpha=0.05$ 水平上, 能否认为不合格率与方法有关?

解: 设 A 的不合格率为 p_1 , B 的不合格率为 p_2 , 此时要检验的假设为:

$$H_0: p_1 = p_2, H_1: p_1 \neq p_2$$

由所给出的备择假设, 利用大样本的正态近似性, 得出在 $\alpha=0.05$ 的显著水平上的拒绝域为 $\{|Z| \geq 1.96\}$ 。

由样本数据知 $n=100, m=100$, 有

$$\hat{p}_1 = \frac{10}{100} = 0.10, \hat{p}_2 = \frac{6}{100} = 0.06, \hat{p} = \frac{10+6}{100+100} = 0.08$$

则得

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{(\frac{1}{n} + \frac{1}{m})\hat{p}(1-\hat{p})}} = \frac{0.10 - 0.06}{\sqrt{(1/100 + 1/100) \times 0.08 \times (1-0.08)}} = 1.04 < 1.96$$

未落在拒绝域中, 所以在 $\alpha=0.05$ 的显著性水平上认为不合格率与制造方法无关。

用 MINITAB 软件从 STAT-Basic Statistics-2 Proportion 入口, 可得到下列结果:

Test and CI for Two Proportions			
Sample	X	N	Sample p
1	10	100	0.100 000
2	6	100	0.060 000
Difference = p (1) - p (2)			
Estimate for difference: 0.04			
95% CI for difference: (-0.034 992 6, 0.114 993)			
Test for difference = 0 (VS not = 0): Z = 1.04 P-Value = 0.297			

这里 p 值为 0.297, 大于给定的 $\alpha=0.05$, 因此无法拒绝原假设, 得到同样的结论: 两种方法间无显著差异。这就给我们很大的启示, 表面上看, 10% 当然比 6% 要大, 但是在样本量较小时竟然得不到有显著差别的结论。统计工作者不相信简单的百分数, 100 分之 6 与 1 000 分之 60 是不同的。如果上例的数字都扩大十倍, 则二者的差别就是显著的了。

**Test and CI for Two Proportions**

Sample	X	N	Sample p
1	100	1 000	0.100 000
2	60	1 000	0.060 000

Difference = p (1) - p (2)

Estimate for difference: 0.04

95% CI for difference: (0.016 285 2, 0.063 714 8)

Test for difference = 0 (vs not = 0): Z = 3.30 P-Value = 0.001

6.2.3 配对数据检验

x_i 和 y_i 是分别来自 $N(\mu_1, \sigma_1^2)$ 与 $N(\mu_2, \sigma_2^2)$ 的样本观测值, 要检验这两个分布的均值是否一致。比如, 某团队对测量矿石中二氧化锰含量的两种分析方法进行对比, 随机从过程产品中抽样, 将同一个样品用两种实验分析方法分别测量二氧化锰含量各一次。如果我们将此问题当作两总体的均值检验问题, 很容易得到“差别不显著”的结论, 这是因为矿石间有很大差异, 两组的方差都很大, 而这种差别又不是由于测量误差造成的。鉴定测量方法差别的最好方法是将同一块矿石的两个结果之差求出来, 看此差值的均值是否为零。

我们采用如下方法处理:

令 $d_i = x_i - y_i$ ($i = 1, 2, \dots, n$), 那么可以把它们看成是来自如下正态总体的样本的观察值:

$$D = X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2) \stackrel{ic}{=} N(\mu, \sigma^2)$$

则检验 μ_1 与 μ_2 是否相等的问题就转化为检验如下假设:

$$H_0: \mu = 0, H_1: \mu \neq 0$$

这是单个正态总体均值是否为 0 的检验问题。

由于 σ 未知, 因此对此问题用 t 检验, 检验统计量变成:

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

因此, 在 $\alpha = 0.05$ 的显著水平上拒绝域为 $\{|t| \geq t_{1-\alpha/2}(n-1)\}$ 。

例 6-9

某团队对测量矿石中二氧化锰含量的两种分析方法进行对比, 随机地从过程产品中抽样, 将同一个样品用两种实验分析方法分别测量二氧化锰含量各一次。表 6-9 是测量的记录, 问两种分析方法的结果在 $\alpha = 0.05$ 的显著性水平上是否有显著差异?

表 6-9

二氧化锰含量测量的记录

序号 i	x_i (实验方法 A)	y_i (实验方法 B)	$d_i = x_i - y_i$
1	7.2	6.6	0.6
2	2.6	2.4	0.2
3	9.4	9.6	-0.2
4	15.4	15.0	0.4
5	6.7	7.0	-0.3
6	7.3	7.3	0.0
7	5.2	4.9	0.3
8	7.9	7.4	0.5
9	1.3	1.0	0.3
10	9.7	9.6	0.1
11	11.1	10.7	0.4
均值	$\bar{x}=7.62$	$\bar{y}=7.4095$	$\bar{d}=0.2091$
标准差	$s_x=3.90$	$s_y=3.8977$	$s_d=0.2844$

这里的数据属于配对数据, 每次测定的结果不仅与实验方法有关, 还与该产品中的二氧化锰含量有关。 d_i 便是两种分析方法的差异, 其中 $d_i = x_i - y_i$ ($i=1, 2, \dots, n$)。 $n=11$, 在 $\alpha=0.05$ 时的拒绝域是:

$$\{|t| \geq t_{1-\alpha/2}(n-1)\} = \{|t| \geq 2.228\}$$

由于 $n=11$, $\bar{d}=0.2091$, $s_d=0.2844$, 可求得 $t = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{0.2091}{0.2844/\sqrt{11}} = 2.44$, 样本落在拒绝域中, 所以可以认为两种分析方法测定的结果有显著差异。

用 MINITAB 软件, 对于差值进行单样本均值是否为 0 的检验, 得到下列结果:

Test of mu = 0 vs mu not = 0				
Variable	N	Mean	StDev	SE Mean
d	11	0.2091	0.2844	0.085763
Variable	95.0% CI		T	P
d	(0.017998, 0.400184)		2.44	0.035

如果将此问题误认为是普通的两样本均值检验, 我们可以使用 MINITAB 软件直接用“2 Sample t”方法得到下列结果:

Two-Sample T-Test and CI: A, B				
Two-sample T for A vs B				
	N	Mean	StDev	SE Mean
A	11	7.62	3.90	1.2
B	11	7.41	3.89	1.2

Difference = $\mu_A - \mu_B$

Estimate for difference: 0.209 091

95% CI for difference: (-3.256 860, 3.675 041)

T-Test of difference = 0 (vs not =): T-Value = 0.13 P-Value = 0.901 DF = 20

Both use Pooled StDev = 3.896 7

p -value 高达 0.907, 不能拒绝原假设, 即不能认为测量结果有显著差异。可见对于配对观测数据如果按配对数据方法来分析, 可以得到较为精确的结论; 误认为普通两样本数据来分析则容易犯第 II 类错误而得不到正确结论。

6.2.4 拟合优度检验

1. 拟合优度检验问题

我们考虑这种情形: 总体中的每个个体被分配到几个类别中的一个, 且仅被分配到一个类别中。这样的总体被称作多项总体。多项概率分布可看作二项分布在三个及三个以上类别情形的推广。在每次多项试验中, 有且仅有一个结果发生。每次试验被假定为独立的, 而且每次试验中落入各类别的概率都不变。

例 6—13

市场调查公司研究各型号产品的市场份额, 在过去的一年中, 型号 A 的市场份额稳定于 30%, 型号 B 为 50%, 型号 C 为 20%。最近型号 C 进行了改进, 开发了一种改进产品, 即将取代其当前占有市场的产品而且会扩大市场份额。调查公司希望判断此新产品是否使市场份额发生了改变。

本例中, 待检验总体是一个多项总体; 每个顾客按照他是购买型号 A、型号 B 还是型号 C 来分类。于是, 我们有了一个具有三个类别的多项总体。采用下列记号来表示比率:

p_A = 型号 A 的市场份额 p_B = 型号 B 的市场份额 p_C = 型号 C 的市场份额

进行抽样调查, 计算各个产品的偏好比率, 然后进行假设检验以确定新产品是否导致市场份额的改变。假定 C 的新产品尚未改变市场份额, 建立如下零假设和对立假设。

$H_0: p_A = 0.30, p_B = 0.50, p_C = 0.20$

$H_1: p_A = 0.30, p_B = 0.50, p_C = 0.20$ 中至少一项不成立

如果样本结果导致拒绝 H_0 , 则市场调查公司有证据表明新产品的引进对市场份额有影响。

调查公司用一组 200 名顾客的群体进行研究, 向每个人询问他们对于型号 A, B 及新型号 C 的购买偏好。200 份答卷汇总结果见表 6—10 的第 3 列“观测频数”中。

现在进行拟合优度检验, 确定 200 名顾客购买偏好与零假设是否相符。先计算出零假设为真的条件下应期望的顾客数, 再将观察结果与期望结果做比较, 再计算出描述差异程度的统计量。因此, 先要在假设 $p_A = 0.30, p_B = 0.50, p_C = 0.20$ 下, 计算 200 名顾客购买偏好, 只要将每个种类的假设比率乘以样本容量 200 即可求得每个种类的期望频数。将得到期望的结果列在表 6—10 的第 4 列“期望频数”中。

2. 拟合优度检验

现在, 拟合优度检验重点考察观测频数与期望频数之差。观测频数与期望频数之差较大则会引起对假设的比率或市场份额的正确性产生怀疑。观测频数与期望频数之差是否过大, 可以用以下检验方法来衡量。现将此类问题概括为一般形式:

$$H_0: p_i = p_{i0} \leftrightarrow H_a: p_i \neq p_{i0} \text{ 至少一项成立, } i=1, \dots, k$$

拟合优度的检验统计量:

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$$

式中, o_i 为类别 i 的观测频数; e_i 为假设 H_0 为真时类别 i 的期望频数; k 为类别总数。 e_i 的计算公式为 $e_i = np_{i0}$, 式中 $n = \sum o_i$ 。

注意: 当观测样本足够多, 且所有种类的期望频数均大于或等于 5 时, 检验统计量近似服从自由度为 $k-1$ 的 χ^2 分布。实际观测数据与期望数据偏差较大时, 检验统计量 χ^2 会变大, 因此当检验统计量超过 χ^2 分布上侧 α 分位数时, 拒绝原假设。在拒绝原假设后, 还可以分析 χ^2 统计量的构成, 我们将第 i 项计算结果 $\frac{(o_i - e_i)^2}{e_i}$ 称之为第 i 项的卡方贡献量。找出卡方贡献量较大的若干项, 则可以肯定它们是使拟合不佳的“罪魁祸首”。

我们仍考虑三个型号的市场份额研究的表 6—10 数据。由于期望频数都大于或等于 5, 我们可以进行 χ^2 检验统计量的计算。拟合优度检验中, χ^2 检验统计量的计算过程见表 6—10, 我们得到检验统计量的值为 7.34。

表 6—10 市场份额研究中 χ^2 检验统计量的计算过程

类别	假设比例 (p_{i0})	观测频数 (o_i)	期望频数 (e_i)	差 ($o_i - e_i$)	差的平方 ($o_i - e_i$) ²	卡方贡献量 ($o_i - e_i$) ² / e_i
A	0.30	48	60	-12	144	2.40
B	0.50	98	100	-2	4	0.04
C	0.20	54	40	14	196	4.90
总计		200				7.34

假设显著性水平为 $\alpha=0.05$, 自由度为 $k-1=3-1=2$, 查 χ^2 分布表, 得到相应的 $\chi_{0.05}^2=5.99$ 。因为 χ^2 检验统计量 $7.34 > 5.99$, 所以拒绝 H_0 , 即得出结论认为市场份额确有变化。

计算机软件包, 如 MINITAB 和 Excel, 能大大简化多项总体拟合优度检验所需要的计算。另外, 计算机也可以使我们直接得到 p -value, 进而立即得到检验结果。本检验的 p -value 为 0.025 5。于是, 对于 $\alpha=0.05$, 拒绝 H_0 。

前面已经指出, 拟合优度检验仅能够得出结论认为型号 C 新产品的引进将改变当前市场份额结构。作为检验结果的进一步推论, 我们能从卡方贡献量中找出贡献最大者, 并通过比较观测频数和期望频数来得到市场份额结构发生变化的原因。考虑型号 C, 它的卡方贡献量最大, 又发现观测频数 54 比期望频数 40 大, 观测频数较大则表明改进的型号 C 将对型号 C 的市场份额有正的影响; 同理可知, 改进的型号 C 将对型号 A 的市场份额有负的影响; B 基本上未变; 而且可知 C 增加的市场份额对型号 A 的影响要大于对型号 B 的影响。

6.3 方差分析

6.3.1 单因子方差分析

有时我们会遇到需要比较多个总体的均值问题，下面便是一个例子。

现有四条生产线生产同一种型号垫片，为了解不同生产线的垫片的断裂强度有无明显差异，现分别从每条生产线随机抽取 5 个垫片测定其断裂强度，数据如表 6—11 所示，试问四条生产线生产的垫片的平均断裂强度是否相同？

表 6—11 四条生产线的垫片断裂强度

生产线	1	2	3	4
断裂强度	86.5	93.4	88.6	94.3
	92.0	87.9	93.2	93.3
	85.2	90.6	88.8	92.0
	87.9	85.5	92.7	89.2
	86.0	88.4	90.9	92.5

在这个问题中，我们遇到需要比较 4 个总体均值的问题。这类问题可以用方差分析方法来解决。

1. 基本概念

为了以后叙述方便，这里先给出在本节和以后在试验设计中经常遇到的几个概念。

(1) 因子。对指标有影响的因素称为因子，因子常用大写字母 A, B, C, D 等表示。在例 6—14 中，生产线对指标垫片的断裂强度有影响，因此生产线便是因子，记为 A 。

(2) 水平。在试验中因子所处的状态称为因子的水平，用因子的字母加下标来表示，譬如因子 A 的水平用 A_1, A_2 等表示。例 6—14 中有四条生产线，这便是过程这一因子的 4 个水平，分别记为 A_1, A_2, A_3, A_4 。

(3) 试验条件（也称处理）。在一次试验中，每个因子总取一个特定的水平，若干因子各取一个特定的水平构成的组合，称其为一个试验条件。在例 6—14 中仅有一个因子，所以每一个水平便是一个试验条件。

(4) 指标。衡量试验条件好坏的量称为指标，用 y 表示，它是一个随机变量。在例 6—14 中垫片的断裂强度就是试验指标。

2. 单因子方差分析的模型

如果一个试验中所考察的因子只有 1 个，那么这是单因子试验问题。

设在一个试验中只考察一个因子 A ，它有 r 个水平，在每一水平下进行 m 次重复试验，其结果用 $y_{11}, y_{12}, \dots, y_{1m}$ ($i=1, 2, \dots, r$) 表示。记第 i 水平下的数据和为 T_i ，数据均值为 \bar{y}_i ，总的均值为 \bar{y} 。此时共有 $n=r \times m$ 个数据。常常把数据列成如下表格形

式, 见表 6—12。

表 6—12 单因子试验数据表

水平	试验数据	和	均值
A_1	$y_{11}, y_{12}, \dots, y_{1m}$	T_1	\bar{y}_1
A_2	$y_{21}, y_{22}, \dots, y_{2m}$	T_2	\bar{y}_2
\vdots	\vdots	\vdots	\vdots
A_r	$y_{r1}, y_{r2}, \dots, y_{rm}$	T_r	\bar{y}_r

为了对这些数据进行统计分析, 需要对数据作三项假定:

(1) 在水平 A_i 下, $y_{i1}, y_{i2}, \dots, y_{im}$ 是来自正态分布 $N(\mu_i, \sigma_i^2)$ 的一个样本, 其中诸 μ_i 就是要比较的对象。

(2) 在不同水平下的方差相等, 即 $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_r^2 = \sigma^2$, 只要诸试验是在相同条件下进行, 方差相等一般可以满足;

(3) 各数据 y_{ij} 相互独立, 这通常只要把试验次序随机化即可得到满足。

在上述三项假定下, 诸总体均值是否相等的问题归结为一个假设检验问题, 其原假设与备择假设分别为:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_r, \quad H_1: \mu_1, \mu_2, \dots, \mu_r \text{ 不全相等}$$

检验这一对假设的统计方法便是方差分析 (analysis of variance, ANOVA)。

当 H_0 不真时, 表示不同水平下指标的均值有显著差异, 此时称因子 A 是显著的, 否则称因子 A 不显著。

如果在一个试验中所要考察的影响指标的因子有 2 个, 那么是两因子试验的问题, 数据分析可以采用两因子方差分析方法 (two-way ANOVA)。

如果在一个试验中所要考察的影响指标的因子更多, 那么试验往往要事先进行设计, 以便用尽可能少的试验去获得数据, 再对数据进行分析。

3. 方差分析的基本思想和步骤

(1) 平方和分解。上述 n 个数据不全相等, 它们的波动可以用总的偏差平方和 SS_T 表示:

$$SS_T = \sum_{i=1}^r \sum_{j=1}^m (y_{ij} - \bar{y})^2$$

引起数据波动的原因不外乎如下两个:

一是由于因子 A 的水平不同, 当假设 H_0 不真时, 各个水平下指标的均值不同, 这必然会使试验结果不同, 我们可以用组间偏差平方和来表示, 也称因子 A 的偏差平方和:

$$SS_A = \sum_{i=1}^r m(\bar{y}_i - \bar{y})^2$$

这里乘以 m 是因为每一水平下进行了 m 次试验。

二是由于存在随机误差, 即使在同一水平下获得的数据间也有差异, 这是除了因子 A 的水平外的一切原因引起的, 我们将它们归结为随机误差, 可以用组内偏差平方和表示:

$$SS_e = \sum_{i=1}^r \sum_{j=1}^m (y_{ij} - \bar{y}_i)^2$$

可以证明有如下平方和分解式:

$$SS_T = SS_A + SS_e$$



(2) 自由度与均方和(平均偏差平方和)。上述诸平方和的大小与数据个数有关,一般讲,数据个数越多,偏差平方和会越大一些;水平数越多,因子A的偏差平方和也会越大一些。所以为了进行比较,还需要引入自由度的概念。这里只给出自由度的计算公式。 SS_T , SS_A , SS_e 的自由度分别用 df_T , df_A , df_e 表示,那么:

在 SS_T 的 $n=rm$ 个偏差中有一个关系式 $\sum_{i=1}^r \sum_{j=1}^m (y_{ij} - \bar{y}) = 0$,故仅有 $n-1$ 个偏差是独立的,因而总的偏差平方和的自由度 $df_T = n-1$;

在 SS_A 的 r 个偏差中有一个关系式 $\sum_{i=1}^r (\bar{y}_i - \bar{y}) = 0$,故仅有 $r-1$ 个偏差是独立的,因而因子A的偏差平方和的自由度为 $df_A = r-1$;

在 SS_e 的 $n=rm$ 个偏差中有 r 个关系式 $\sum_{j=1}^m (y_{ij} - \bar{y}_i) = 0 (i=1, \dots, r)$,故仅有 $r(m-1)$ 个偏差是独立的,因而 $df_e = r(m-1) = n-r$ 。

它们之间也有分解式 $df_T = df_A + df_e$ 。

我们将因子或误差的偏差平方和与相应的自由度之比称为因子或误差的均方和,并分别记为:

$$MS_A = SS_A / df_A, MS_e = SS_e / df_e$$

(3) F 比与拒绝域。当原假设 H_0 成立时,诸 μ_i 间无差异,从而因子A的均方和较小,即 MS_A 与 MS_e 相差不大;而当 MS_A 相对于 MS_e 大得多时,则可以认为因子A是显著的。这一比较可以用两者的比值来表示,记为: $F = MS_A / MS_e$ 。

在 H_0 成立时,统计量 F 的分布为自由度是 df_A , df_e 的 F 分布。当 $F > F_{1-\alpha}(df_A, df_e)$ 时认为因子A是显著的,即诸均值不全相等;当 $F \leq F_{1-\alpha}(df_A, df_e)$ 时,认为因子A不显著。其中 $F_{1-\alpha}(df_A, df_e)$ 是自由度为 df_A , df_e 的 F 分布的 $1-\alpha$ 分位数。

(4) 方差分析表。以上计算过程可列成一张方差分析表,见表6—13。

表 6—13 单因子方差分析表

来源	偏差平方和	自由度	均方和	F 比
因子A	SS_A	$df_A = r-1$	$MS_A = SS_A / df_A$	$F = MS_A / MS_e$
误差e	SS_e	$df_e = n-r$	$MS_e = SS_e / df_e$	
总计T	SS_T	$df_T = n-1$		

在以上计算中,关键是计算各个偏差平方和,通过代数运算有:

$$SS_T = \sum_{i=1}^r \sum_{j=1}^m (y_{ij} - \bar{y})^2 = \sum_{i=1}^r \sum_{j=1}^m y_{ij}^2 - \frac{T^2}{n}, df_T = rm - 1$$

$$SS_A = \sum_{i=1}^r m(\bar{y}_i - \bar{y})^2 = \sum_{i=1}^r \frac{T_i^2}{m} - \frac{T^2}{n}, df_A = r - 1$$

$$SS_e = SS_T - SS_A, df_e = rm - r$$

式中, T_i 是第 i 个水平数据的和; T 为 $n=rm$ 个数据的总和。

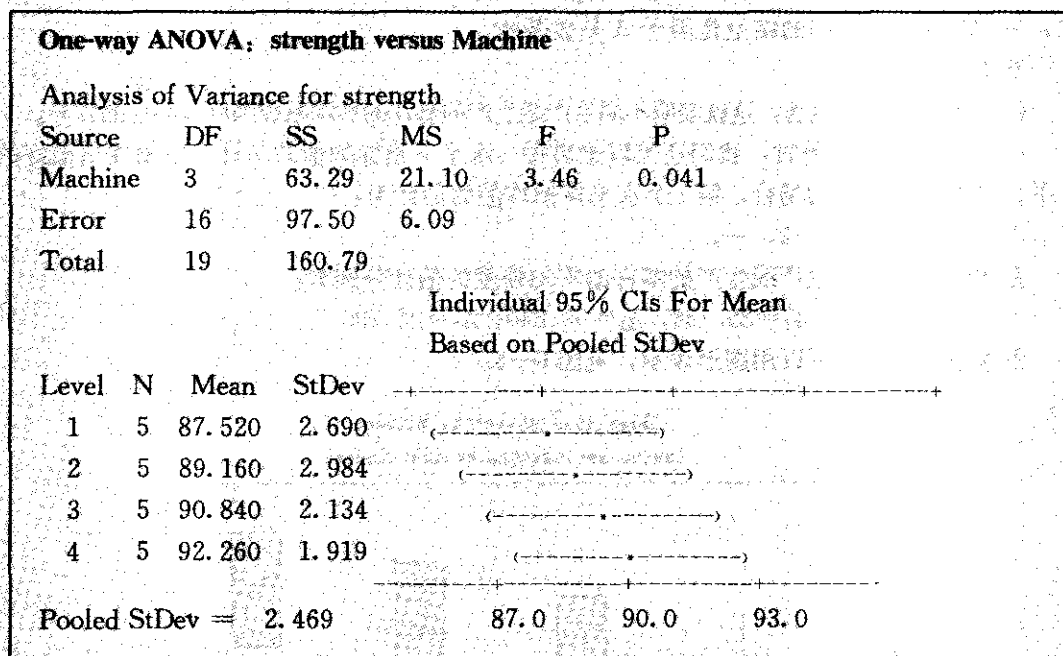
(5) MINITAB软件的运算。如果采用MINITAB软件运算,从ANOVA-One Way ANOVA的窗口进入,可计算出各项偏差平方和,并给出ANOVA表。在ANOVA表中若 $p\text{-value} < \alpha$,则拒绝 H_0 。

4. 方差分析的具体步骤

综上所述, 进行方差分析的步骤如下:

- (1) 计算因子 A 的每一水平下数据的和 T_1, T_2, \dots, T_r 及总和 T ;
- (2) 计算各类数据的平方和 $\sum_{i=1}^r \sum_{j=1}^m y_{ij}^2$, $\sum_{i=1}^r T_i^2$ 和总和的平方 T^2 ;
- (3) 依次计算 SS_T, SS_A, SS_e ;
- (4) 填写方差分析表;
- (5) 对于给定的显著性水平 α , 将求得的 F 值与 F 分布表中的临界值 $F_{1-\alpha}(df_A, df_e)$ 比较, 当 $F > F_{1-\alpha}(df_A, df_e)$ 时认为因子 A 是显著的, 否则认为因子 A 是不显著的。

对表 6—11 的数据进行方差分析, 下框是利用 MINITAB 软件运算的结果:



p 值为 $0.041 < 0.05$, 所以在 $\alpha = 0.05$ 水平上我们的结论是因子 A 是显著的。

我们也可以用前述公式手工计算。

- (1) 计算各类和。每一水平下的数据和为: $T_1 = 437.6, T_2 = 445.8, T_3 = 454.2, T_4 = 461.3$, 总和为:

$$T = 1798.9$$

- (2) 计算各类平方和。

$$\text{原始数据的平方和为: } \sum_{i=1}^4 \sum_{j=1}^5 y_{ij}^2 = 161962.85$$

$$\text{每一水平下数据和的平方和为: } \sum_{i=1}^4 T_i^2 = 809326.73$$

(3) 计算各偏差平方和。

$$SS_T = 160.79, \quad df_T = 4 \times 5 - 1 = 19$$

$$SS_A = 63.29, \quad df_A = 4 - 1 = 3$$

$$SS_e = 160.79 - 63.29 = 97.50, \quad df_e = 19 - 3 = 16$$

(4) 填写方差分析表 (见表 6—14)。

表 6—14

例 6—14 方差分析表

来源	偏差平方和	自由度	均方和	F 比
因子 A	$SS_A = 63.29$	$df_A = 3$	$MS_A = 21.10$	$F = 3.46$
误差 e	$SS_e = 97.50$	$df_e = 16$	$MS_e = 6.09$	
总计 T	$SS_T = 160.79$	$df_T = 19$		

(5) 如果给定 $\alpha = 0.05$, 从 F 分布表查得 $F_{0.95}(3, 16) = 3.24$, 由于 $F > 3.24$, 所以在 $\alpha = 0.05$ 水平上我们的结论是因子 A 是显著的。

结论:

(1) 因为因子 A 显著, 这就表明不同生产线生产的垫片的平均断裂强度有明显的差异。

(2) 当因子 A 显著时, 我们还可以给出每一水平下指标均值的估计, 以便找出最好的水平。在单因子试验的场合, 第 i 个水平指标均值的估计为:

$$\hat{\mu}_i = \bar{y}_i, \quad i = 1, 2, \dots, r$$

在本例中, 四条生产线生产的零件的平均强度的估计分别为:

$$\hat{\mu}_1 = 87.520, \quad \hat{\mu}_2 = 89.160, \quad \hat{\mu}_3 = 90.840, \quad \hat{\mu}_4 = 92.260$$

各水平下的均值可以用图形表示, 见图 6—8。

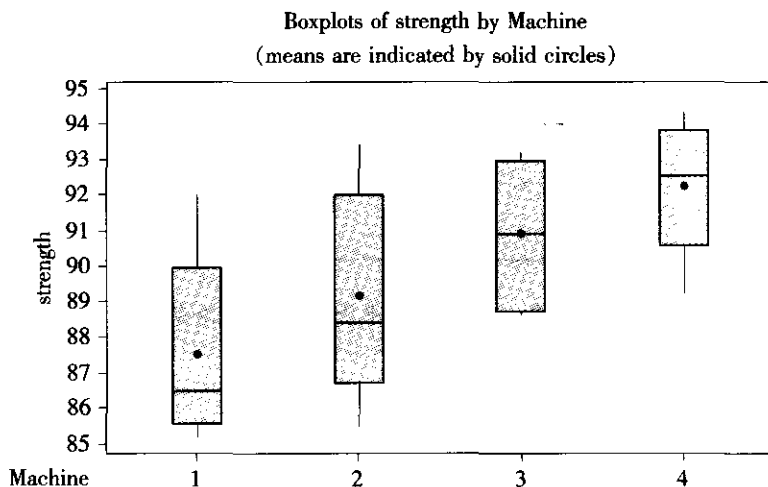


图 6—8 例 6—14 的各水平的箱线图

由此可见, 4 号生产线的垫片的断裂强度均值最大, 如果需要平均断裂强度大的垫片, 那么 4 号生产线的为好。而从过程来讲, 1 号与 2 号线应该设法改进。

(3) 前面提到, 即使是同一条生产线 (过程), 生产的产品特性也是有波动的, 若波动用方差度量, 则通过数据分析还可以给出误差方差的估计, 这里方差 σ^2 的估计是 MS_e 。

在本例中 σ^2 的估计是 6.09, σ 的估计是 $\sqrt{6.09} = 2.469$ 。

5. 重复数不同的单因子方差分析

若在一水平下重复试验次数不同, 假定在 A_i 水平下进行了 m_i ($i=1, 2, \dots, r$) 次试验, 那么进行方差分析的步骤仍然同上, 只是在计算中有两个改动:

一是此时 $n = \sum m_i$; 二是 SS_A 的计算公式改为:

$$SS_A = \sum_{i=1}^r \frac{T_i^2}{m_i} - \frac{T^2}{n}$$

如果采用 MINITAB 软件运算, 仍可从 ANOVA-One Way ANOVA 入口, 判断法则同样是: 若 $p\text{-value} < \alpha$, 则拒绝 H_0 。

例 6-15

某团队项目是在仪器制造中降低材料消耗, 对用各种结构的仪器分别测定其材料消耗量, 数据如表 6-15, 想要了解仪器的结构 (记为因子 A) 对材料消耗量的影响是否显著 (这里假定每一种结构下的材料消耗量服从等方差的正态分布)。

表 6-15

不同结构仪器耗材量

水平	材料消耗量						
结构 A_1	61.0	62.8	57.6	58.3	54.7	55.5	59.3
结构 A_2	52.8	54.5	48.5	50.2			
结构 A_3	54.3	56.1	51.4	53.6			

现在对这批数据作方差分析。

One-way ANOVA: consume versus structure					
Analysis of Variance for consume					
Source	DF	SS	MS	F	P
Structur	2	155.65	77.82	11.86	0.001
Error	13	85.34	6.56		
Total	15	240.98			

				Individual 95% CIs For Mean	
				Based on Pooled StDev	
Level	N	Mean	StDev		
A_1	8	58.688	2.742		
A_2	4	51.500	2.670		
A_3	4	53.850	1.943		

Pooled StDev =	2.562	49.0	52.5	56.0	59.5
----------------	-------	------	------	------	------

由于 $p\text{-value} = 0.001 < 0.05$, 故拒绝原假设, 或者从 F 分布表查得, 由于求得的 F 为 $11.86 > 3.81$, 也同样拒绝原假设。所以在 $\alpha = 0.05$ 水平上, 因子 A 是显著的。



结论:

(1) 由于因子 A 显著, 这就表明不同结构生产仪器的材料消耗量有明显的差异。

(2) 我们还可以给出不同结构生产仪器的材料消耗量的估计:

$$\bar{\mu}_1 = 58.69, \bar{\mu}_2 = 51.50, \bar{\mu}_3 = 53.85$$

由此可见, 从特性指标的角度看, 结构 A₂ 最理想, 结构 A₃ 次之, 结构 A₁ 最不好。

(3) 在本例中误差方差的估计为 $\hat{\sigma}^2 = 6.56$, 标准差的估计为 $\hat{\sigma} = 2.562$ 。

6.3.2 两因子方差分析

如果在一个试验中需要同时考察两个因子 A 与 B, 并设因子 A 有 r 个水平, 因子 B 有 s 个水平, 这时共有 $n = rs$ 个不同的试验条件, 也就是说有 n 个总体。

现有 4 条生产线生产同一种垫片, 为了解不同生产线的垫片的断裂强度有无明显差异, 现分别用 5 种不同的温度进行试验, 测得数据如表 6—16。在假定不同条件下垫片的断裂强度分别服从等方差的正态分布时, 分别分析不同生产线及不同温度对垫片的断裂强度均值有无显著影响, 如果有影响, 那么在什么条件下垫片的断裂强度达到最大?

表 6—16 垫片断裂强度两因子试验结果

生产线		1	2	3	4
断 裂 强 度	温度 700	88.6	89.5	95.7	90.6
	温度 750	86.0	88.0	86.6	92.6
	温度 800	87.0	89.1	88.8	92.4
	温度 850	89.3	91.9	93.1	91.8
	温度 900	81.2	84.0	85.7	95.1

为作一般数据分析, 现在作如下假定: 每一个总体的分布是正态分布, 其均值为 μ_{ij} , 它与因子 A 及 B 的水平有关, 方差相同, 都是 σ^2 ; 数据相互间是独立的。数据整理成表 6—17 格式。

表 6—17 两因子试验数据表 (每个条件下进行一次试验)

A \ B	B				行和 $T_{i.}$	行平均 $\bar{y}_{i.}$
	B_1	B_2	...	B_s		
A_1	y_{11}	y_{12}	...	y_{1s}	$T_{1.}$	$\bar{y}_{1.}$
A_2	y_{21}	y_{22}	...	y_{2s}	$T_{2.}$	$\bar{y}_{2.}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
A_r	y_{r1}	y_{r2}	...	y_{rs}	$T_{r.}$	$\bar{y}_{r.}$
列和 $T_{.j}$	$T_{.1}$	$T_{.2}$...	$T_{.s}$	总和 T	
列平均 $\bar{y}_{.j}$	$\bar{y}_{.1}$	$\bar{y}_{.2}$...	$\bar{y}_{.s}$		总平均 \bar{y}

现在需要我们去分析因子 A 的水平不同对指标的均值有无显著影响, 因子 B 的水平不同对指标的均值有无显著影响, 有时还需要回答两个因子不同水平的搭配对指标的均值有无显著影响, 这种影响如果存在就称为因子 A 与 B 有交互作用。

1. 交互作用概念

因子 A 与 B 的交互作用可以用图形直观地表示。如图 6—9a 表示因子 A 与 B 不存在交互作用, 这时不管因子 B 取什么水平, 因子 A 的效应 (即 A 在高水平时指标均值与 A 在低水平的指标均值的差) 总为相同值 h 。在图 6—9b 中, 当因子 B 取不同水平时, 虽然因子 A 的高水平的指标均值总比低水平时的高, 但高的程度有所不同, 这表示因子 A 与因子 B 间有正向的交互作用。在图 6—9c 中, 当因子 B 取低水平时, 因子 A 的高水平的指标均值比 A 的低水平的高, 但是当因子 B 取高水平时, 因子 A 的高水平的指标均值却比 A 的低水平的低, 这表示因子 A 与因子 B 之间有反向的交互作用。

如图 6—9b 与 c 所显示的, 一个因子的效应与另一因子所处水平有关时, 称为因子 A 与 B 有交互作用, 记为 AB 或 BA。

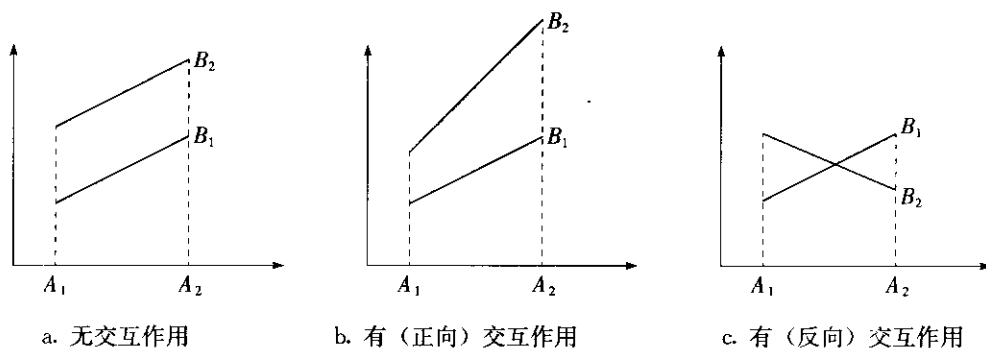


图 6—9 因子 A 与 B 的交互作用示意图

2. 无交互作用时的两因子方差分析

当凭经验和专业知识知道因子 A 与 B 不存在交互作用时, 我们可以在因子 A 与 B 的每一种搭配下只进行一次试验, 由此即可分析因子 A 的水平不同对指标的均值有无显著影响, 因子 B 的水平不同对指标 A 的均值有无显著影响, 分析方法基本与单因子方差分析相同。

设在 A_i 与 B_j 条件下的试验结果用 y_{ij} 表示, 它们相互独立, A_i 水平下的均值用 $\bar{y}_{i\cdot}$ 表示 (即表 6—17 的行平均), B_j 水平下的均值用 $\bar{y}_{\cdot j}$ 表示 (即表 6—17 的列平均), 总的数据均值用 \bar{y} 表示。类似于单因子方差分析, 可以进行平方和分解, 只是现在数据的总偏差平方和 SS_T 可以分解成三项:

$$SS_T = SS_A + SS_B + SS_e, \quad df_T = df_A + df_B + df_e$$

式中 SS_A , SS_B 及 SS_e 分别称为因子 A、因子 B 及误差的偏差平方和, df_A , df_B 及 df_e 分别表示其相应的自由度, 它们的表达式和计算公式如下:

$$SS_T = \sum_{i=1}^r \sum_{j=1}^s (y_{ij} - \bar{y})^2 = \sum_{i=1}^r \sum_{j=1}^s y_{ij}^2 - \frac{T^2}{n}, \quad df_T = rs - 1$$

$$SS_A = \sum_{i=1}^r s(\bar{y}_{i\cdot} - \bar{y})^2 = \sum_{i=1}^r \frac{T_{i\cdot}^2}{s} - \frac{T^2}{n}, \quad df_A = r - 1$$

$$SS_B = \sum_{j=1}^s r(\bar{y}_{\cdot j} - \bar{y})^2 = \sum_{j=1}^s \frac{T_{\cdot j}^2}{r} - \frac{T^2}{n}, \quad df_B = s - 1$$



$$SS_e = \sum_{i=1}^r \sum_{j=1}^s (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y})^2 = SS_T - SS_A - SS_B$$

$$df_e = df_T - df_A - df_B$$

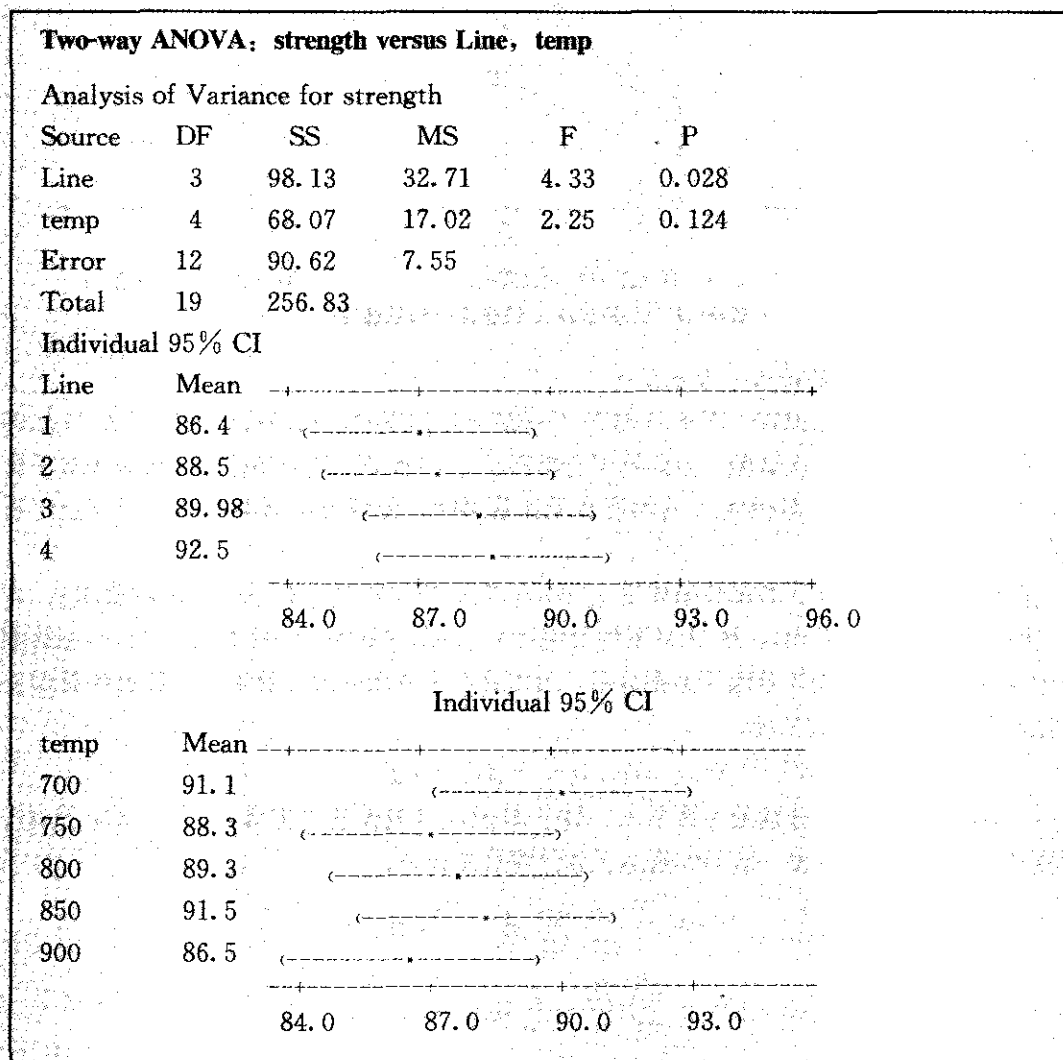
$$= (r-1)(s-1)$$

式中的 $T_{i.}$, $T_{.j}$ 与 T 分别表示 A_i 水平下的数据和 (即行和)、 B_j 水平下的数据和 (即列和) 与数据总和。

例 6-16

对数据进行方差分析, 其步骤如下:

在 MINITAB 中, 从 ANOVA-Two Way ANOVA 入口, 由于本例中两个因子每次搭配时都只进行了一次试验, 因此这里不可能考虑两因子的交互作用, 在下方小方格中 (Fit Additive Model) 不论选择与否, 计算时都将只拟合“可加模型”。其计算结果如下:



结论:

(1) 因子B(生产线)的 $p=0.028<0.05$,在 $\alpha=0.05$ 水平上因子生产线对垫片的断裂强度有显著影响;因子A(温度)的 $p=0.124>0.05$,在 $\alpha=0.05$ 水平上因子温度对垫片的断裂强度没有显著影响。

(2) 为了寻找最好的条件,需要给出不同条件下均值的估计。在没有交互作用的情况下,可以认为:

$$\mu_{ij} = \mu + a_i + b_j$$

式中, μ 称为一般平均,其估计为 $\hat{\mu} = \bar{y}$ 。

a_i 称为因子A的第*i*水平的效应,在A不显著时,它们全为0,取任意水平均可;在A显著时,其估计是 $\hat{a}_i = \bar{y}_{i.} - \bar{y}$,此时我们可以从 A_i 各水平的平均值去寻找最好的水平。

b_j 称为因子B的第*j*水平的效应,在B不显著时,它们全为0,取任意水平均可;在B显著时,其估计是 $\hat{b}_j = \bar{y}_{.j} - \bar{y}$,此时我们可以从 B_j 各水平的平均值去寻找最好的水平。最后 μ_{ij} 的估计为:

$$\hat{\mu}_{ij} = \hat{\mu} + \hat{a}_i + \hat{b}_j$$

将上述找到各自的最好水平组合起来,便是最好的条件。

在本例中,因为B显著,A不显著,故

$$\hat{\mu}_{ij} = \hat{\mu} + \hat{a}_i + \hat{b}_j = \bar{y} + 0 + (\bar{y}_{.j} - \bar{y}) = \bar{y}_{.j}, i=1, 2, \dots, r; j=1, 2, \dots, s$$

所以只要找出B的最好水平,便是最好的条件。从所给出的数据可见,在4号生产线上,垫片的平均断裂强度达到最大,此时均值的估计是92.5。

(3) 还可以给出误差方差 σ^2 的估计 $\sigma^2 = MS_e = 7.55$,标准差 σ 的估计是 $\sigma = \sqrt{MS_e} = 2.746$ 。

(4) 由于因子A不显著,因此应该将不显著的因子A删去,然后再拟合一个仅含单因子B的ANOVA模型。不要以为在多因子问题中,因子个数的选择总是“多多益善”。实际上,把不显著的因子选入模型只会增大误差,因而还是不选入为好。很显然,单个因子的方差分析这部分内容前面已作过介绍,在这里就不再重复了。

3. 有交互作用时两因子的方差分析

当需要考察两因子A与B的交互作用时,重复试验是不可少的,在A与B的每一种搭配下都需进行*m*次试验(注意这里重复次数是相等的,且至少两次),那么不仅可以分析因子A的水平不同对指标的均值有无显著影响和因子B的水平不同对指标的均值有无显著影响,而且还可以分析A与B的交互作用对指标均值有无显著影响。数据的一般表达形式如表6—18。

设在 A_i 与 B_j 条件下的*m*个试验结果用 $y_{ij1}, y_{ij2}, \dots, y_{ijm}$ 表示,它们相互独立, $A_i B_j$ 条件下的数据均值用 \bar{y}_{ij} 表示, A_i 水平下的均值用 $\bar{y}_{i.}$ 表示, B_j 水平下的均值用 $\bar{y}_{.j}$ 表示,总的的数据均值用 \bar{y} 表示。



表 6—18

两因子试验数据表 (每个条件下进行 m 次试验)

$A \backslash B$	B_1	...	B_s	行和 $T_{i.}$	行平均 $\bar{y}_{i..}$
A_1	y_{111}, \dots, y_{11m} (T_{11})	...	y_{1s1}, \dots, y_{1sm} (T_{1s})	$T_{1.}$	$\bar{y}_{1..}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
A_r	y_{r11}, \dots, y_{r1m} (T_{r1})	...	y_{rs1}, \dots, y_{rsm} (T_{rs})	$T_{r.}$	$\bar{y}_{r..}$
列和 $T_{.j}$	$T_{.1}$...	$T_{.s}$	总和 T	
列平均 $\bar{y}_{.j.}$	$\bar{y}_{.1.}$...	$\bar{y}_{.s.}$		总平均 \bar{y}

类似前面进行平方和分解, 现在数据的总偏差平方和 SS_T 可以分解成四项:

$$SS_T = SS_A + SS_B + SS_{A \times B} + SS_e, \quad df_T = df_A + df_B + df_{A \times B} + df_e$$

式中 SS_A , SS_B , $SS_{A \times B}$ 及 SS_e 分别称为因子 A、因子 B、交互作用 $A \times B$ 及误差的偏差平方和, df_A , df_B , $df_{A \times B}$ 及 df_e 分别为其自由度, 它们的表达式和计算公式如下:

$$SS_T = \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^m (y_{ijk} - \bar{y})^2 = \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^m y_{ijk}^2 - \frac{T^2}{n}, \quad df_T = n - 1 = rsm - 1$$

$$SS_A = \sum_{i=1}^r s(\bar{y}_{i..} - \bar{y})^2 = \sum_{i=1}^r \frac{T_{i.}^2}{sm} - \frac{T^2}{n}, \quad df_A = r - 1$$

$$SS_B = \sum_{j=1}^s r(\bar{y}_{.j.} - \bar{y})^2 = \sum_{j=1}^s \frac{T_{.j}^2}{rm} - \frac{T^2}{n}, \quad df_B = s - 1$$

$$SS_{A \times B} = \sum_{i=1}^r \sum_{j=1}^s m(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y})^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{T_{ij.}^2}{m} - \frac{T^2}{n} - SS_A - SS_B, \\ df_{A \times B} = df_A \times df_B = (r-1)(s-1)$$

$$SS_e = \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^m (y_{ijk} - \bar{y}_{ij.})^2 = SS_T - SS_A - SS_B - SS_{A \times B},$$

$$df_e = df_T - (df_A + df_B + df_{A \times B}) = rs(m-1)$$

式中, $T_{ij.}$, $T_{i.}$, $T_{.j}$ 及 T 分别表示 $A_i B_j$ 条件下的数据和, A_i 水平下的数据和, B_j 水平下的数据和及数据总和, $n = rsm$ 。方差分析表只要再增加一行交互作用项即可。

例 6—17

为了寻找最好的温度与生产线的配合, 使得垫片的断裂强度达到最大, 选取两因子的水平如下:

因子 A (温度) 有 2 个水平: 700, 800

因子 B (生产线) 有 4 个水平: 1, 2, 3, 4

在每个 $A_i B_j$ 条件下各进行 5 次试验, 所得数据见表 6—19, 试对数据进行分析, 并找出使垫片的断裂强度达到最高的条件。

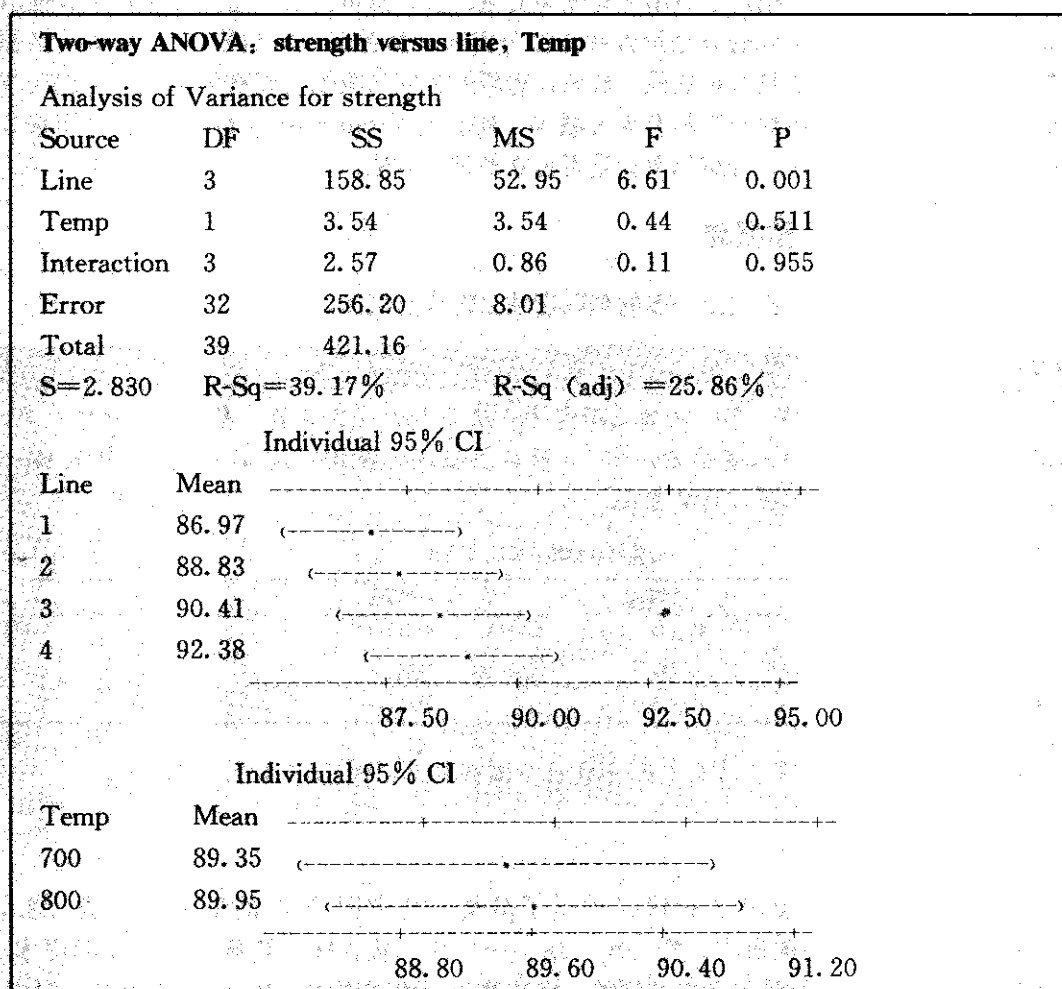
表 6-19

有交互作用的垫片断裂强度两因子试验数据

生产线		1	2	3	4
断裂强度	温度 700	88.6	89.5	95.7	90.6
		86.0	88.0	86.6	92.6
		87.0	89.1	88.8	92.4
		89.3	91.9	93.1	91.8
		81.2	84.0	85.7	95.1
断裂强度	温度 800	86.5	93.4	88.6	94.3
		92.0	87.9	93.2	93.3
		85.2	90.6	88.8	92.0
		87.9	85.5	92.7	89.2
		86.0	88.4	90.9	92.5

在本例中 $r=2$, $s=4$, $m=5$, $n=rs m=2 \times 4 \times 5=40$ 。

采用 MINITAB 软件运算, 仍然是从 ANOVA-Two Way ANOVA 入口, 注意不要选中“可加模型”(Fit Additive Model), 则可得到下列结果。



所以在 $\alpha=0.05$ 水平上:

因子 B (生产线) $p=0.001<0.05$, 显著。

因子 A (温度) $p=0.511>0.05$, 不显著。

交互作用 $p=0.955>0.05$, 不显著。

由此结果可以得到的结论是: 生产线因子显著、温度及交互作用 $A \times B$ 都不显著。因此, 我们应先删去 $A \times B$, 拟合只含 A, B 两因子可加模型。由于 A 因子仍不显著, 最后拟合只含 B 因子的单因子 ANOVA 模型, 得到的结论是: B 因子 (生产线) 的效应是显著的, 使垫片的断裂强度达到最高的条件是生产线 4, 其垫片的平均断裂强度可以达到 92.38。

6.4 列联表

本章前几节我们已介绍过了适用于连续型变量的统计分析方法, 这些方法对于属性数据就不适用了。我们收集到的只是按各种方法分类的计数数据 (即频数), 要分析的问题是判断两种分类方式是否具有独立性。例如, 想研究肺癌与吸烟之间有无关联, 工艺改革与合格品率有无关联, 事故与工龄有无关联等。列联表 (contingency tables) 就是用来分析两种方式 (即两个因子) 分组的数据是否独立的常用方法。

6.4.1 独立性检验概念

我们先看一个例子, 再引出一般的列联表独立性检验概念。

例 6—18

调查了 2 000 人, 在 1 200 名不吸烟者中, 有 1 100 人不咳嗽, 有 100 人咳嗽; 在 800 名吸烟者中, 有 700 人不咳嗽, 100 人咳嗽。将这些数据列成如下表格 (见表 6—20), 我们想分析吸烟是否与咳嗽有关。

表 6—20

吸烟与咳嗽 2×2 列联表

	B_1 : 不咳嗽	B_2 : 咳嗽	行和
A_1 : 不吸烟	1 100	100	1 200
A_2 : 吸烟	700	100	800
列和	1 800	200	2 000

这张表就被称为列联表, 表中因子 A 表示吸烟状况, 因子 B 表示咳嗽状况。

1. 列联表

一般地, 假设对两个因子的状况进行了分类, 其中因子 A 的取值有 r 类: A_1, A_2, \dots, A_r ; 因子 B 的取值有 c 类: B_1, B_2, \dots, B_c 。从总体中获得容量为 n 的样本, 按两个变量分类的状况统计了各自的频数, 将这些频数值排列成一张 r 行 c 列的表, 这就

是一张二维 $r \times c$ 列联表 (见表 6—21)。

表中 n_{ij} 是因子 A 取 A_i 且因子 B 取 B_j 的样品数。行和与列和分别为:

$$n_{i\cdot} = \sum_{j=1}^c n_{ij}, \quad i = 1, 2, \dots, r$$

$$n_{\cdot j} = \sum_{i=1}^r n_{ij}, \quad j = 1, 2, \dots, c$$

式中, 行和 $n_{i\cdot}$ 表示样本中 A 取 A_i 的样品数, 列和 $n_{\cdot j}$ 表示样本中 B 取 B_j 的样品数。

表 6—21 $r \times c$ 列联表

	B_1	B_2	...	B_c	行和
A_1	n_{11}	n_{12}	...	n_{1c}	$n_{1\cdot}$
A_2	n_{21}	n_{22}	...	n_{2c}	$n_{2\cdot}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
A_r	n_{r1}	n_{r2}	...	n_{rc}	$n_{r\cdot}$
列和	$n_{\cdot 1}$	$n_{\cdot 2}$...	$n_{\cdot c}$	总和 n

2. 独立性假设

列联表要研究的问题是这两个因子之间是否有关联, 这种关联性的研究在统计上可以描述为如下的假设检验问题:

H_0 : 因子 A 与因子 B 独立; H_1 : 因子 A 与因子 B 不独立

6.4.2 列联表的独立性检验方法

将上面提到的假设问题, 称为列联表的独立性检验。

令 $p_{ij} = P(A_i, B_j)$, 又记 $p_{i\cdot} = \sum_{j=1}^c p_{ij}$, $p_{\cdot j} = \sum_{i=1}^r p_{ij}$, 则当两个因子独立时, 对一切 i, j , 应该有 $p_{ij} = p_{i\cdot} p_{\cdot j}$ 。

所以列联表的独立性检验可以等价地转化为检验如下的假设:

H_0 : 对一切 i, j , 有 $p_{ij} = p_{i\cdot} p_{\cdot j}$

H_1 : 至少有一对 i, j , 有 $p_{ij} \neq p_{i\cdot} p_{\cdot j}$

由于在实际问题中, p_{ij} , $p_{i\cdot}$, $p_{\cdot j}$ 都是未知的, 因此常采用如下估计:

$$\hat{p}_{i\cdot} = \frac{n_{i\cdot}}{n}, \quad i = 1, 2, \dots, r$$

$$\hat{p}_{\cdot j} = \frac{n_{\cdot j}}{n}, \quad j = 1, 2, \dots, c$$

当两个因子独立时, 对一切 i, j , 有

$$\hat{p}_{ij} = \hat{p}_{i\cdot} \hat{p}_{\cdot j} = \frac{n_{i\cdot}}{n} \times \frac{n_{\cdot j}}{n}$$

在例 6—18 中, 不吸烟的比例 (不管是否咳嗽) 为 $1\,200/2\,000 = 0.6$, 不咳嗽的比例 (不管是否吸烟) 为 $1\,800/2\,000 = 0.9$ 。因此, 如果吸烟与咳嗽无关的话, 则不吸烟且不咳嗽的人占总人数的比率应该 $0.6 \times 0.9 = 0.54$, 相当于应该有 $2\,000 \times 0.54 = 1\,080$ 人不吸烟且不咳嗽。但实际上不吸烟且不咳嗽的人有 1 100 人, 这中间有 20 人的差距。当这个

差距不大时,我们可以认为这是观测误差,仍可以认为吸烟与咳嗽无关;但当这个差距太大时,我们就不能认为吸烟与咳嗽无关了。对此要有统计的检验方法。在 6.2 节末尾讨论分布拟合优度检验时曾提及,对于此类拟合检验,统计学家 K. Pearson 建议采用如下的检验统计量:

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - n\hat{p}_{i\cdot}\hat{p}_{\cdot j})^2}{n\hat{p}_{i\cdot}\hat{p}_{\cdot j}} = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - \hat{n}_{i\cdot}\hat{n}_{\cdot j}/n)^2}{\hat{n}_{i\cdot}\hat{n}_{\cdot j}/n}$$

并且证明了,在 H_0 成立的条件下,当 n 较大时, χ^2 检验统计量近似服从自由度为 $(r-1)(c-1)$ 的 χ^2 分布。

由于在 H_0 为真时,对一切 i, j , 观测频数 n_{ij} 与期望频数 $n\hat{p}_{ij} = n\hat{p}_{i\cdot}\hat{p}_{\cdot j} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{n}$ 不应相差很大,故当 χ^2 统计量的值偏大时,倾向于 H_0 不真,即应该拒绝 H_0 。

所以在显著性水平为 α 时,拒绝域取为:

$$\{\chi^2 \geq \chi_{1-\alpha}^2((r-1)(c-1))\}$$

如果采用 MINITAB 软件运算,对于已整理好的数据表,可以从 STAT-Tables-Chi Square Test 入口,将列联表中的各列作为响应变量输入,则可直接得到计算结果。若有 $p < \alpha$, 则拒绝 H_0 。例 6—18 的计算结果如下:

Chi-Square Test: NC, C

Expected counts are printed below observed counts

Chi-Square contributions are printed below expected counts

	NC	C	Total
1	1 100	100	1 200
	1 080.00	120.00	
	0.370	3.333	
2	700	100	800
	720.00	80.00	
	0.556	5.000	
Total	1 800	200	2 000

Chi-Sq = 9.259, DF = 1, P-Value = 0.002

由于 p -value 只有 0.002, 因此拒绝原假设, 即应认为吸烟与咳嗽有关, 或说吸烟与不吸烟的咳嗽率是有显著差别的。上述 MINITAB 计算表格中, 在每个格中都列出了 3 个数字, 从上而下, 分别是实际观测值 o_i 、期望值 e_i 及对卡方统计量的贡献。本例中, 总卡方统计量是 9.259, 其中右下角的贡献 5.000 为最大, 它的含义是, 若吸烟与咳嗽无关的话, 在吸烟又咳嗽的这类人数应该为 80 人, 而实际上, 这类人达到 100 人, 它远远高于应该有的人数, 这说明: 吸烟的人患咳嗽的人数太多了, 或说吸烟的人患咳嗽的比率确实较高。

下面举例说明列联表在多个总体比例比率中的应用。

例6—19

国家有关部门规定,对于工商企业申请营业执照要在10个工作日内办理完毕。统计了A, B, C三个城市实际办理营业执照的及时与不及时状况,得到下表(见表6—22)。

表6—22 营业执照办理情况数据表

	及时户数	不及时户数
A	205	103
B	312	113
C	150	90

问:三个城市办理营业执照的及时率有显著差异吗?

解:首先计算期望频数 $n\hat{p}_{i \cdot} \cdot \hat{p}_{\cdot j} = \frac{n_{i \cdot} \cdot n_{\cdot j}}{n}$, 结果见表6—23。

表6—23 营业执照办理期望频数表

	及时	不及时	总计
A	211.14	96.86	308
B	291.34	133.66	425
C	164.52	75.48	240
合计	667	306	973

由以上数据可计算得

$$\chi^2 = \frac{(205-211.14)^2}{211.14} + \frac{(103-96.86)^2}{96.86} + \frac{(312-291.34)^2}{291.34} + \frac{(113-133.66)^2}{133.66} + \frac{(150-164.52)^2}{164.52} + \frac{(90-75.48)^2}{75.48} = 9.301$$

现在 $r=3$, $c=2$, $(r-1)(c-1)=2$, 在 $\alpha=0.05$ 时, $\chi_{0.95}^2(2)=5.991$, 故拒绝域为 $\{\chi^2 \geq 5.991\}$, 现在样本统计量的值落在拒绝域中, 所以认为三个城市办理营业执照的及时率方面有显著差异。

用 MINITAB 计算, 可得下列结果:

Chi-Square Test: ON, NO			
Expected counts are printed below observed counts			
Chi-Square contributions are printed below expected counts			
	ON	NO	Total
1	205	103	308
	211.14	96.86	
	0.178	0.389	
2	312	113	425
	291.34	133.66	
	1.465	3.193	
3	150	90	240
	164.52	75.48	
	1.282	2.794	
Total	667	306	973
Chi-Sq = 9.301, DF = 2, P-Value = 0.010			



从上述结果中也可以看出应拒绝原假设的结论,即三个城市办理营业执照的及时率方面有显著差异。另外,还可以再进行更具体的分析:在总计 χ^2 统计量 9.301 中,贡献最大的是城市 B 的“不及时”格中的 3.193,如果三个城市办理营业执照的及时率方面无显著差异,则 B 城不及时的应有 133.66 户,但实际上只有 113 户,说明 B 城办事效率高,不及时的少于预期;贡献次大的是城市 C 的“不及时”格中的 2.794,如果三个城市办理营业执照的及时率方面无显著差异,则 C 城不及时的应有 75.48 户,但实际上却高达 90 户,说明 C 城办事效率太低,不及时的多于预期。A 城的贡献几乎为 0,说明 A 城办事效率居于中等,与三城市的平均水平没什么差距。

例 6-20

对某计算机软件的质量情况进行用户调查,请他们对产品的质量情况选择回答:差、一般、好。随机抽取 70 人,对他们进行询问,并发现其中 40 人接受过有关的广告宣传,另 30 人则不太关心这类广告,回答的结果列在表 6—24 中。

表 6—24 软件质量调查结果数据表

观测频数	差	一般	好	合计
听过广告宣传	11	18	11	40
未听过广告宣传	10	13	7	30
合计	21	31	18	70

试问广告与人们对产品的评价间有无联系?(取 $\alpha=0.05$)

解:首先计算期望频数 $n\hat{p}_{i\cdot}\hat{p}_{\cdot j}=\frac{n_{i\cdot}\cdot n_{\cdot j}}{n}$,结果见表 6—25。

表 6—25 软件质量调查期望频数表

期望频数	差	一般	好	合计
听过广告宣传	12	17.7	10.3	40
未听过广告宣传	9	13.3	7.7	30
合计	21	31	18	70

由以上数据可计算得

$$\chi^2 = \frac{(11-12)^2}{12} + \frac{(10-9)^2}{9} + \frac{(18-17.7)^2}{17.7} + \frac{(13-13.3)^2}{13.3} + \frac{(11-10.3)^2}{10.3} + \frac{(7-7.7)^2}{7.7} = 0.321$$

现在 $r=2$, $c=3$, $(r-1)(c-1)=2$, 在 $\alpha=0.05$ 时, $\chi_{0.05}^2(2)=5.991$, 故拒绝域为 $\{\chi^2 \geq 5.991\}$, 现在样本统计量的值未落在拒绝域中, 所以认为广告与人们对产品的评价间无联系。

用 MINITAB 计算, 可得下列结果:

Chi-Square Test: bad, midd, good				
Expected counts are printed below observed counts				
Chi-Square contributions are printed below expected counts				
	bad	midd	good	Total
1	11	18	11	40
	12.00	17.71	10.29	
	0.083	0.005	0.050	
2	10	13	7	30
	9.00	13.29	7.71	
	0.111	0.006	0.066	
Total	21	31	18	70
Chi-Sq = 0.321, DF = 2, P-Value = 0.852				

要注意的是当期望频数 $n\hat{p}_{ij}$ 过小 (一般当 $n\hat{p}_{ij} < 5$) 时, 通常需要将相邻的组进行合并后再进行独立性检验。

例 6—21

某厂有三条流水线, 统计每条流水线上每天生产的不合格品数, 经过整理, 在下面的列联表 6—26 中给出了每条流水线在 300 个工作日中, 每天生产的不合格品数为 0 个、1 个、2 个等的天数。

表 6—26

流水线不合格品数列联表

流水线	不合格品数					合计
	0	1	2~3	4~5	≥ 6	
甲	230	52	12	4	2	300
乙	236	40	14	9	1	300
丙	197	67	22	11	3	300
合计	663	159	48	24	6	900

试问在显著性水平 0.05 上不合格品数与流水线是否有关?

解: 首先计算期望频数 $n\hat{p}_{i \cdot} \hat{p}_{\cdot j} = \frac{n_{i \cdot} \cdot n_{\cdot j}}{n}$, 结果列在表 6—27 中。

表 6—27

流水线不合格品数期望频数表

流水线	不合格品数					合计
	0	1	2~3	4~5	≥6	
甲	221	53	16	8	2	300
乙	221	53	16	8	2	300
丙	221	53	16	8	2	300
合计	663	159	48	24	6	900

由于这里不合格品数≥6 这一列的期望频数仅为 2, 小于 5, 所以将该列与前一列合并, 列联表修改为表 6—28。

表 6—28

流水线不合格品数列联表 (合并后)

流水线	不合格品数				合计
	0	1	2~3	≥4	
甲	230	52	12	6	300
乙	236	40	14	10	300
丙	197	67	22	14	300
合计	663	159	48	30	900

重新计算期望频数 $np_{i \cdot} p_{\cdot j} = \frac{n_{i \cdot} n_{\cdot j}}{n}$, 结果见表 6—29。

表 6—29

流水线不合格品数期望频数表 (合并后)

流水线	不合格品数				合计
	0	1	2~3	≥4	
甲	221	53	16	10	300
乙	221	53	16	10	300
丙	221	53	16	10	300
合计	663	159	48	30	900

这里所有的期望频数都大于 5, 由以上数据可计算得

$$\chi^2 = \frac{(230-221)^2}{221} + \dots + \frac{(14-10)^2}{10} = 17.597$$

现在 $r=3$, $c=4$, $(r-1)(c-1)=6$, 在 $\alpha=0.05$ 时, $\chi_{0.95}^2(6)=12.592$, 故拒绝域为 $\{\chi^2 \geq 12.592\}$, 由于样本统计量的值落在拒绝域中, 所以认为不合格品数与流水线是有关的。

同样, 通过 MINITAB 计算也可以得到同样结果, p -value 为 0.007, 故应拒绝原假设, 即认为不合格品数与流水线有关。

Chi-Square Test: 0, 1, 2~3, >4

Expected counts are printed below observed counts

Chi-Square contributions are printed below expected counts

	0	1	2~3	>4	Total
1	230	52	12	6	300
	221.00	53.00	16.00	10.00	
	0.367	0.019	1.000	1.600	
2	236	40	14	10	300
	221.00	53.00	16.00	10.00	
	1.018	3.189	0.250	0.000	
3	197	67	22	14	300
	221.00	53.00	16.00	10.00	
	2.606	3.698	2.250	1.600	
Total	663	159	48	30	900

Chi-Sq = 17.597, DF = 6, P-Value = 0.007

列联表的用途是很广泛的。我们常用它进行对离散记录的分析，从中可以得到很多有意义的信息。

6.5 非参数检验

我们先介绍有关非参数检验的概念。

从检验的问题本身，可以分为两大类：参数检验问题和非参数检验问题。例如，对于均值的检验，对于比率的检验等，这都是关于参数的假设检验。例如，分布的正态性检验、数据的独立性检验、两总体分布是否相同的检验等则是非参数假设检验。

我们下面要讨论的所谓非参数检验方法，讨论的仍是参数问题，特别是有关均值的检验，但使用的方法是非参数的。普通的参数检验所使用的方法常常是求平均值、标准差等统计量，但这些统计量的性质特别是分布强烈地依赖于原始数据“服从正态分布”这个前提条件。我们下面所用的方法只限于使用“符号”(sign)、“秩”(rank)及“游程”(run)等几种与原始数据分布完全无关的工具。

很显然，本章介绍过的各种分析方法其使用条件除了要求各数据间相互独立之外，一定还要求数据为正态分布。如果原始数据非正态，则原来使用的 Z 、 t 、 F 及 χ^2 检验就都不能使用了。为获得检验结果，只能使用不依赖于分布的非参数方法。

非参数方法的使用范围广、简便易行，但我们必须明确回答，如果一组数据符合正态



分布,那么对这些数据使用非参数统计方法效果如何呢?实际上,同组数据、同样问题用非参数检验效果是不如参数方法的,更准确地说就是,非参数检验的检出力会比参数方法稍小,换言之,非参数检验所犯的第Ⅱ类错误概率会比参数方法稍大。因此,应该在条件允许的情况下(确认分布为正态)尽量使用前面介绍过的参数方法。

6.5.1 符号检验

在统计分析中常遇到只取两类的计数结果(例如“合格”与“不合格”、“满意”与“不满意”等),有时也会将连续数据根据某种准则分为两类(例如“比中位数大”与“比中位数小”、“上升”与“下降”等)。如何对这类数据进行统计分析呢?一种简单有效的方法就是符号检验(sign test)。例如,在通常的市场研究应用中,找 n 名顾客来调查并最后确认对诸如方便面、咖啡、软饮料和清洁剂等商品的两个品牌中喜欢哪一个,或者两个方案中赞成哪一个。这 n 个顾客对喜欢的表达是名称数据,因为顾客只需说出所喜欢商品的名称或品牌。根据这些数据,我们的目标是确定人们对于被比较的两个事物是否存在喜欢程度上的差异。符号检验就是回答这一问题的一个最简单方便的非参数统计方法。

1. 小样本情形

当 $n \leq 30$ 时,称之为小样本情形。不同样本量的情况下使用符号检验的步骤是不同的。下面让我们对例6—22进行研究,演示一下如何使用小样本符号检验法。

例 6—22

某企业生产一种在市场上名为A的食品。该企业的竞争对手也已开始生产一种新的名为B的食品。在一项关于消费者对这两个品牌的偏爱的研究中,给12个人未贴标签的每种食品样本。每个人首先品尝哪个品牌是随机挑选出来的,在品尝过两种食品之后,这些人被要求说出在两个品牌中所偏爱的那个品牌。结果是2人喜欢A品牌,10人喜欢B品牌。问能否确定消费者确实偏爱两种食品中的一种。

解:设 p 表示消费者总体中偏爱A的比率,我们想检验以下假设:

$$H_0: p=0.50, H_1: p \neq 0.50$$

若无法拒绝 H_0 ,我们将没有证据表明对两种产品品牌的偏爱间存在差异。然而,若是拒绝了 H_0 ,我们可以认为消费者对两种品牌的偏爱间存在差异。在这种情况下,大多数消费者选出的品牌被认为是受欢迎的品牌。

在下面的讨论中我们将展示如何使用符号检验的小样本形式来检验这些假设,并得出有关消费者偏好的结论。为了记录参与这项研究的12个人的偏好数据,我们用加号来表示对A的偏好,用减号来表示对B的偏好。由于数据是用加号或减号来记录的,这一非参数检验被称为符号检验。

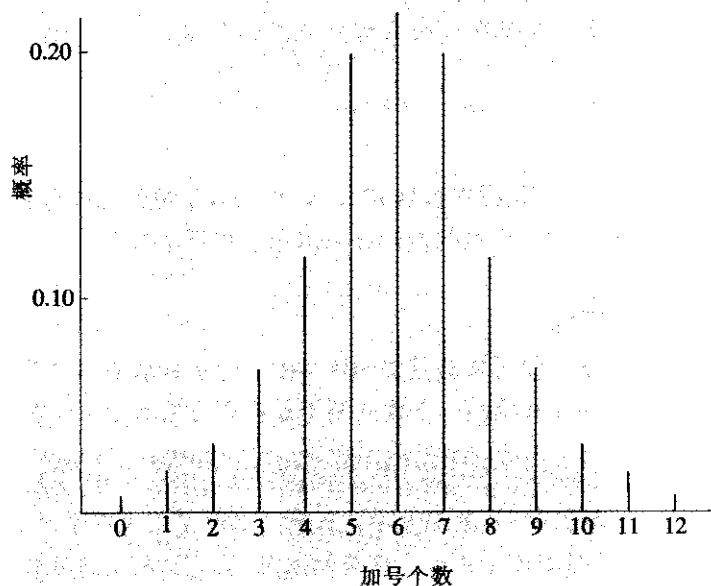
可以看出,如果考虑“加号”出现的比率,则上述问题其实也可以表达为单样本的比率是否恰为0.50的检验问题(见本章6.2节),但当时我们使用的是大样本近似正态的分析方法。如果实际问题中样本量较小,又该怎样进行呢?下面所介绍的就是解决此问题的思路 and 具体步骤。

在 H_0 为真 ($p=0.50$) 的假设下, 加号的个数服从 $p=0.50$ 的二项概率分布。当样本容量 $n=12$ 时, 用二项分布公式可以计算出概率分布值 (见表 6—30), 由此可以画出 $p=0.50$ 的二项分布的概率分布图 (见图 6—10)。它显示了在 H_0 成立的假设下, 加号个数的概率分布律。使用这一分布可以给出拒绝 H_0 的法则。例如, 若 $\alpha=0.05$, 我们将把图 6—10 中分布两端各约 0.025 的部分定义为拒绝域。在此分布的低端 (左端), 我们发现加号个数为 0, 1 或 2 的概率为 $0.000\ 2+0.002\ 9+0.016\ 1=0.019\ 2$ 。注意到我们停留在加号个数为 2 的位置上, 因为加上了加号个数为 3 的概率, 将使得此低端区域等于 $0.019\ 2+0.053\ 7=0.072\ 9$, 超过了所要求的 0.025 的区域。在此分布的高端 (右端), 我们发现相对于 10, 11 或 12 个加号的概率同样为 0.019 2。于是, 我们可以距离 $\alpha=0.05$ 最近, 同时又不超过 $0.019\ 2+0.019\ 2=0.038\ 4$, 所以我们采用以下的拒绝法则: 若加号个数小于 3 或大于 9, 则拒绝 H_0 。

表 6—30

 $n=12, p=0.5$ 的二项分布

加号个数	概率	加号个数	概率
0	0.000 2	7	0.193 4
1	0.002 9	8	0.120 8
2	0.016 1	9	0.053 7
3	0.053 7	10	0.016 1
4	0.120 8	11	0.002 9
5	0.193 4	12	0.000 1
6	0.225 6		

图 6—10 $n=12, p=0.5$ 时加号个数的二项分布概率值

为了简化分析步骤, 不必每次计算出概率分布来。我们明确检验问题及解决步骤是这样的:

$H_0: p=0.50$, $H_1: p \neq 0.50$, 其中 p 是出现“+”号的概率。

在样本观测值中, 记: n_+ 为取正号个数, n_- 为取负号个数, $n=n_++n_-$, 令 $s=\min\{n_+, n_-\}$, 查本书后的附表 8 可得其下限临界值 s_α , 当 $s \leq s_\alpha$ 时, 拒绝原假设。

例 6-22

对食品调查结果进行分析, 已知: 喜欢 A 者 2 人, 喜欢 B 者 10 人, 问消费者是否偏爱两种食品中的一种。

解: 在本例中, $n_+=2$, $n_-=10$, $n=n_++n_-=12$, $s=\min\{n_+, n_-\}=\min\{2, 10\}=2$, 当 $\alpha=0.05$ 时, 查附表 8 可得下限临界值 $s_\alpha=2$, 由于本组结果 $s \leq s_\alpha$, 故拒绝原假设。此研究提供证据表明消费者对两种品牌产品的偏爱确实存在差异。我们会告知该企业, 消费者更喜欢对手的 B 牌产品。这个结果与我们用二项分布计算的结果完全一致。

在上述例子中, 此项研究中全部的 12 个参与者都能够表述自己的偏好。在许多情况下, 样本中的一个或多个个体无法表述出明确的偏爱。更一般的情况是在样本中, 有些取“+”, 有些取“-”, 还有些取“0”。在这种情况下, 取“0”者不应计入总数, 也就是说应将这些样本除去不计。

附表 8 给出的临界值可用于构造样本容量小于 90 的任何符号检验的决策判断。对于原假设 $p=0.50$ 以及样本容量 n , 可建立对应于 0.01, 0.05, 0.10, 0.25 共 4 种显著性水平的决策准则。当然, 这里仅给出双侧检验情况。若仅考虑二项概率分布的低端或高端, 我们将上述显著水平减半即可。附表 8 中未提供样本容量大于 90 的临界值。在这种情况下, 我们可以使用二项分布的大样本正态近似法, 下段介绍大样本的符号检验法。

2. 大样本情形

使用原假设 $H_0: p=0.50$ 且样本容量 $n>30$, 则加号个数的抽样分布可用正态概率分布来近似。原假设成立时, 加号个数的抽样分布可用下列正态近似:

均值: $\mu=0.50n$

标准差: $\sigma=\sqrt{0.25n}$

由于在大样本的情况下, 符号检验其实与单样本的比率检验是一回事, 所以在用计算机计算时实际使用 1-Proportion 检验。下面的例子是将符号检验用于顾客调查。

例 6-23

在近期市长选举活动的一项民意调查中, 询问了 200 名人大代表, 目的是求得评价在甲候选人与乙候选人中谁的政绩更佳。调查结果显示, 72 人认为甲候选人更好, 103 人认为乙候选人更好, 25 人弃权。这一民意调查是否显示出公众对两个候选人政绩评价存在着显著不同?

解: $H_0: p=0.50$, $H_1: p \neq 0.50$, 其中 p 是出现“甲候选人的政绩得到较高评价”的概率。

利用符号检验, 我们看到 $n=72+103=175$ 人能够指出他们所认为的具有最佳政绩的候选人。利用均值和标准差的计算结果, 我们发现加号个数的抽样分布图形是如图 6—11 这样的状况。

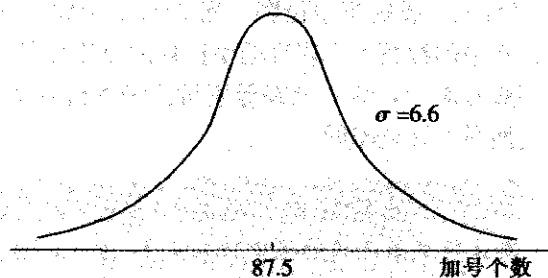


图 6—11 $n=175$ 的符号检验中加号个数的概率分布

$$\mu = 0.50n = 0.50 \times 175 = 87.5$$

$$\sigma = \sqrt{0.25n} = \sqrt{0.25 \times 175} = 6.6$$

当 $n=175$ 这样大时我们当然可以认为抽样分布已近似正态分布, 可以使用标准正态概率分布面积表来建立此检验的拒绝法则。若 $\alpha=0.05$, 这一双边检验的拒绝法则是:

如果 $Z \leq -1.96$ 或 $Z \geq +1.96$, 则拒绝 H_0 。

用“甲候选人的政绩得到较高评价”的次数作为加号的个数 ($x=72$)。我们将得到如下检验统计量的值:

$$Z = \frac{x - \mu}{\sigma} = \frac{72 - 87.5}{6.6} = -2.35$$

由于 $Z = -2.35 < -1.96$, 则在 0.05 的显著性水平下, 我们将拒绝两位候选人的政绩无差异的假设。此项研究表明乙候选人被认为拥有更好的政绩。

利用 MINITAB 软件可以用下列步骤进行。

从 Stat-Basic Statistic-1 Proportion 入口, 将总次数 175 及甲 (或乙) 候选人的得票数 72 作为“成功”次数, 在“Option”窗口选择对 $p=0.5$ 做双侧检验, 可得到下列结果:

Test and CI for One Proportion							
Test of $p = 0.5$ vs $p \text{ not } = 0.5$							
Sample	X	N	Sample p	95.0% CI	Z-Value	P-Value	
1	72	175	0.411 429	(0.338 520, 0.484 337)	-2.34	0.019	

p -value 为 0.019, 拒绝原假设, 这与我们手工计算的结果完全相同。

下面我们介绍一些符号检验法的应用实例。最简单而又有用的例子是中位数的符号检验法。

我们曾阐述了怎样将假设检验用于对总体均值的推断。当总体分布可以被认为是对称

时,均值与中位数是重合的;当分布不能被认为是对称时,均值与中位数的偏离通常也很小。因此,对于均值的检验也可以近似化为对中位数的检验。实际上,我们讨论的是怎样用符号检验来进行有关总体中位数的假设检验。回想一下,中位数将一个总体分割成为两部分,且总体中 50% 的值将大于或等于中位数,而 50% 的值将小于或等于中位数。我们可以采用符号检验,把样本中的数据大于假设的中位数的值用加号表示,把样本中的数据小于假设的中位数的值用减号表示,任一恰好等于假设的中位数的值的数据应被剔除掉。符号检验的计算过程与前面所述完全相同。

例 6-24

公众普遍认为 A 地新建独居单元的价格中位数是 13 万元,经调查,在一个由 62 所新住宅组成的样本中,34 所住宅的价格高于 13 万元,26 所住宅的价格低于 13 万元,有两所住宅的价格恰好等于 13 万元。可以认为新建独居单元的价格中位数是 13 万元吗?

解: H_0 : 中位数=13 万元, H_1 : 中位数 \neq 13 万元

对于 $n=60$ 所其价格不等于 13 万元的住宅,我们得到

$$\mu = 0.50n = 0.50 \times 60 = 30, \sigma = \sqrt{0.25n} = \sqrt{0.25 \times 60} = 3.87$$

因为加号个数 $x=34$, 所以检验统计量为:

$$Z = \frac{x - \mu}{\sigma} = \frac{34 - 30}{3.87} = 1.03$$

采用双边检验以及 $\alpha=0.05$ 的显著性水平,当 $Z \leq -1.96$ 或 $Z \geq +1.96$ 时,拒绝 H_0 。由于此检验统计量 $Z=1.03$,不能拒绝 H_0 。在这些数据的基础上,不能拒绝 A 地新建独居单元的价格中位数为 13 万元的原假设。

用 MINITAB 软件,从 Stat-Basic Statistic-1Proportion 入口,输入汇总数据,将总试验次数 60,“成功”次数 34 输入,则可以得到下列结果:

Test and CI for One Proportion					
Test of p = 0.5 vs p not = 0.5					
Sample	X	N	Sample p	95.0% CI	Exact P-Value
1	34	60	0.566667	(0.432410, 0.694119)	0.366

由于此检验 p -value 为 0.366,不能拒绝 H_0 。

我们再举一个实例说明符号检验的步骤及对符号检验方法的改进途径。

例 6-25

装配车间每日完成记录仪的台数如表 6-31 所示。车间主任宣称他们平均产量已超过日产 35 台水平,能确认他的说法吗?

表 6—31

日装配记录仪台数记录表

日期	1	2	3	4	5	6	7	8	9	10	11	12
产量	32	38	39	33	40	32	41	34	38	40	39	42
比较	—	+	+	—	+	—	+	—	+	+	+	+

解：我们所要检验的假设是：

$$H_0: \text{中位数} \leq 35, H_1: \text{中位数} > 35$$

从数据表中容易看出，超过 35 的为 8 个，低于 35 的为 4 个。查附表 8，看 $n=12$ ， $\alpha=0.10$ 的那列（双侧 0.1 相当于单侧 0.05），得临界值为 2。由于 $4 > 2$ ，故无法拒绝原假设，即不能认为日产量比 35 大。

用 MINITAB 软件，仍从 Stat-Nonparametrics-1 Sample Sign 入口，输入原始数据，可得下列结果：

Sign Test for Median: x						
Sign test of median = 35.00 versus > 35.00						
	N	Below	Equal	Above	P	Median
x	12	4	0	8	0.1938	38.50

由于 $p\text{-value}=0.1938$ ，故无法拒绝原假设。

符号检验简单易算，但效率不高，即不容易检验出显著差别来。对于例 6—25 我们在本节 6.5.3 威尔科克森符号秩检验中还要再算一次，由于采用更精密的方法，将分析出更准确的结论。

6.5.2 曼-惠特尼-威尔科克森检验

在前面我们对两个总体均值间差异进行了参数检验。检验了下面的假设：

$$H_0: \mu_1 - \mu_2 = 0, H_1: \mu_1 - \mu_2 \neq 0$$

使用两个样本的 t 检验法可以解决问题。但要注意，此方法的使用基于一个基本条件，即两个总体均为正态分布。如果这个条件不满足，又如何检验呢？下面我们给出另一种可用于研究两总体间是否存在差异的非参数方法。此检验使用的条件是两个独立样本即可，其中每个样本可以来自任何类型的分布，我们通常把它称为曼-惠特尼-威尔科克森 (Mann-Whitney-Wilcoxon, MWW) 检验。

非参数 MWW 检验不需要上述两个假设。MWW 检验唯一的要求是两个独立随机样本中产生的数据的测度尺度至少是顺序的。MWW 检验确定两总体是否相同，而不是检验两总体均值间的差异。下面是 MWW 检验的假设：

$$H_0: \text{两总体相同}, H_1: \text{两总体不相同}$$

只要两个总体的样本容量皆小于或等于 10，就可以使用 MWW 检验的小样本检验法。我们将通过一个例子来演示 MWW 检验在小样本情形下的应用。

例 6-25

纱厂挡车工的操作者大部分都曾经接受过新工艺 A 和 B 的专项培训。团队成员提出一个问题, 即 A、B 两种工艺的培训在提高技术能力上是相同的吗? 考虑下面的假设:

H_0 : 两个总体在技术能力上是相同的, H_1 : 两个总体在技术能力上是不同的

利用培训的记录, 团队挑选了 4 名曾接受过 A 种工艺培训的员工组成随机样本和另一个由 5 名接受过 B 种工艺培训的员工组成的随机样本。本项研究中所使用的 9 名员工现今的过程操作技术能力表现被记录下来, 9 名员工在全厂挡车工中的名次排队情况列在表 6-32 中。

MWW 过程的第一步是把来自两个样本的混合数据由低到高进行排序, 样本中名次的最佳者(全厂排为第 8 名)的秩为 1, 而样本中名次的最劣者(全厂排为第 202 名)的秩为 9。全部 9 个员工的排队结果列在表 6-33 中。

表 6-32 挡车工操作名次数据

A 工艺培训		B 工艺培训	
员工	班级名次	员工	班级名次
1	8	5	70
2	52	6	202
3	112	7	144
4	21	8	175
		9	146

下一步是分别求出每个样本的秩和。在 MWW 过程中原本可使用任一个样本的秩和, 但附录表 9 要求只对样本量较小者计算秩和。在下面讨论中, 我们使用由 4 名接受 A 工艺培训的员工组成的样本的秩和, 用 T 表示。对于本例, $T=1+2+3+5=11$ 。

表 6-33 挡车工操作排队情况

员工	工艺	班级名次	混合样本秩
1	A	8	①
4	A	21	②
2	A	52	③
5	B	70	4
3	A	112	⑤
7	B	144	6
9	B	146	7
8	B	175	8
6	B	202	9

让我们来考虑 A 工艺的秩和的特性。由于此样本中有 4 名员工, 则 A 工艺培训可能会拥有了此项研究中最好的 4 名员工。若情况确实如此, $T=1+2+3+4=10$ 将是秩和 T 的最小可能值。相反地, 也可能拥有最差的 4 名员工, 在此种情况下, $T=6+7+8+9=30$ 将是 T 的最大可能值。于是, 样本的 T 值必为介于 10~30 之间的一个值。

两个样本容量均小于或等于 10 时的 MWW 检验的临界值列在附表 9 中。在此表中, n_1 代表了在此检验中样本量较小者的样本的容量, 下界 T_L 的值和上界 T_U 的值可直接从此表中读出。当 T 小于等于 T_L 或大于等于 T_U 时将拒绝“两总体相同”的原假设。

利用附表 9 中双侧 0.05 的显著性水平, 我们看到 $n_1=4$ (A 工艺)、 $n_2=5$ (B 工艺) 的 MWW 统计量的低端临界值 $T_L=12$, 高端临界值 $T_U=28$ 。这就是说, 若第一样本 (A 工艺) 的秩和小于等于 12 或大于等于 28, 则两个总体相同的原假设被拒绝。拒绝法则可写为:

如果 $T \leq 12$ 或 $T \geq 28$, 则拒绝 H_0 。

我们看到本例的实际结果是 $T=1+2+3+5=11$ 。于是, 拒绝原假设 H_0 , 我们可得以下结论说在技术能力方面, A 工艺员工总体与 B 工艺员工总体有显著不同。A 所得到的较高的名次表明, 与 B 工艺相比, A 工艺员工做得更好。

使用 MINITAB 软件, MWW 检验被简称为 Mann-Whitney 检验。从 STAT-Nonparametric-Mann Whitney 入口, 可以直接得到下列结果:

Mann-Whitney Test and CI: score_A, score_B

score_A N = 4 Median = 36.5

score_B N = 5 Median = 146.0

Point estimate for ETA1-ETA2 is -108.5

96.3 Percent CI for ETA1-ETA2 is (-181.0, -18.0)

W = 11.0

Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.037 3

这里 p -value 为 0.037 3, 因此结论同样是拒绝“两总体相同”的原假设。

当两总体至少有一个样本量超过 10 时, 附表 9 已无法使用, 这时, 计算机将自动使用正态近似法进行计算, 最后结尾 (is significant at) 给出的数字就是检验的 p -value, 如果此值小于 0.05, 则应拒绝“两总体相同”的原假设。

对于两总体中位数相等性的检验, 除了上述的 Mann-Whitney 检验外, 还可以使用 Mood 的中位数检验法。

仍以例 6—26 的数据来说明此方法。

将两组样本合并后, 求出合并样本的中位数, 计算两个样本各有多少个观测值比此中位数大及比此中位数小。最后, 将这些数据构成列联表, 用 χ^2 检验法检验即可。此例的中位数是 112, 得出比较结果列在表 6—34 中。

表 6—34

挡车工操作的中位数检验表

	不超过中位数	大于中位数
A	4	0
B	1	4

用 MINITAB 从 STAT-Nonparameter-Mood's Median Test 入口, 使用 χ^2 检验法, 结



果如下:

Mood median test for score				
Chi-Square = 5.76 DF = 1 P = 0.016				
Individual 95.0% CIs				
group	N<=	N>	Median	Q3-Q1
A	4	0	37	86
B	1	4	146	82
Overall median = 112				
* NOTE * Levels with < 6 observations have confidence < 95.0%				

结论是拒绝“两总体在技术能力上是相同的”的原假设,认为两总体有较大差别。当然,在本例中由于数据太少,用 χ^2 检验误差较大,应收集更多数据来说明问题。

6.5.3 威尔科克森符号秩检验

对于单个总体均值的检验问题,我们已在6.2.1作了介绍,当标准差已知时,我们可以用 Z 检验;当标准差未知时,可以用 t 检验。当然,这两种检验法都必须要求总体分布为正态。如果分布非正态,我们可以使用符号检验法(见6.5.1节),但符号检验法太粗糙,非常容易犯第Ⅱ类错误,也就是说,即使明明已有显著差异仍然不能拒绝。下面介绍效果相当好的非参数威尔科克森符号秩检验法。这种方法适用于单个总体的中位数检验(在分布对称的假定下,也就是均值等于给定值的检验),当然也适用于配对检验,因为成对或相匹配的观测值间的差异提供了两个总体间差异的信息,最终转化为单总体均值是否为0的检验。

例 6—35

装配车间每日完成记录仪的台数如表6—35所示。车间主任宣称他们平均产量已超过日产35台水平。能确认他的说法吗?

解:我们所要检验的假设是:

$$H_0: \text{中位数} \leq 35, \quad H_1: \text{中位数} > 35$$

先求出各观测值与35之差(见表6—35第4行),然后按其绝对值排序,并将其秩次列在表6—35最后一行(注意观测值相同时要进秩的修正)。容易看出负号共4项(见表6—35第5行中带圆圈者),对应负号的秩和 T 为 $1+2+4.5+4.5=12$ 。查附表9,对应 $n_1=4$, $n_2=8$ 的 $\alpha=0.05$ 的 T 的单侧下限为16,我们观测值为12,故应拒绝原假设,即认为目前产量确实比日产35台有所提高。

表 6-35 挡车工操作的符号秩检验表

日期	1	2	3	4	5	6	7	8	9	10	11	12
产量	32	38	39	33	40	32	41	34	38	40	39	42
比较	-	+	+	-	+	-	+	-	+	+	+	+
差值	-3	3	4	-2	5	-3	6	-1	3	5	4	7
绝对值序	(4.5)	4.5	7.5	(2)	9.5	(4.5)	11	(1)	4.5	9.5	7.5	12

用 MINITAB 计算, 从 STAT-Nonparametric-1Sample Wilcoxon 入口, 可以得到下列结果:

Wilcoxon Signed Rank Test: x					
Test of median = 35.00 versus median > 35.00					
	N	N for Test	Wilcoxon Statistic	P	Estimated Median
x	12	12	66.0	0.019	37.25

这里计算机给出的 Wilcoxon 统计量 (Wilcoxon Statistic) 为 66, 它计算的是对应正号观测值的秩和 (因为数据总共 12 个, 故正负号总秩和为 78)。由于算出的 p -value=0.019, 故应拒绝原假设, 即认为日产量比 35 有显著提高。

注意我们在前面已经用符号检验法处理过此问题, 当时的结论是“不能拒绝原假设”。这说明符号检验法较为粗糙, 不易发现显著差异, 而我们使用的符号秩检验法的检出效率大大提高了。

当正负号中有任一组个数超过 10 时, 附表 9 就不能用了, 这时, 计算机会自动使用正态近似方法计算之, 因此不论样本量如何, 对单个总体的中位数检验都可以使用效果比较灵敏的威尔科克森符号秩检验法。

配对样本的参数分析方法 (配对差异的 t 检验) 需要是定距类数据并假定其成对观测的差服从正态分布。在这一假定下, 可以采用 t 分布来检验总体均值间无差异的原假设。若服从正态分布的假定不成立, 则可使用威尔科克森符号秩检验。我们将通过比较两种不同方法的有效性来演示这一非参数检验。

例 6-27

一家制造企业试图确定两种生产方法在完工时间上是否存在差异。选出了一个由 11 人组成的样本, 且每个工人都分别使用两种生产方法完成了同样一项生产任务。每个工人首先使用哪种生产方法是随机挑选的。于是, 每个工人提供了一对观测值, 如表 6-36 所示。若差异为正, 表明方法 1 需要更多的时间; 若差异为负, 则表明方法 2 需要更多的时间。请判断, 两种方法在完工时间上是否存在差异?

解: 事实上, 我们有两个完工时间的总体, 一个总体对应于一种方法。设立下面的假设:

H_0 : 两总体相同, H_1 : 两总体不同

若 H_0 不能被拒绝, 我们将没有证据说明两种方法的完工时间存在差异; 若拒绝, 则可以说两种方法在完工时间上存在差异。

表 6—36 两种方法完工时间比较表

工人	方法 1	方法 2	差异
1	10.2	9.5	0.7
2	9.6	9.8	-0.2
3	9.2	8.8	0.4
4	10.6	10.1	0.5
5	9.9	10.2	-0.3
6	10.2	9.3	0.9
7	10.9	10.5	0.4
8	10.0	10.0	0.0
9	11.2	10.6	0.6
10	10.7	10.2	0.5
11	10.6	9.8	0.8

威尔科克森符号秩检验的第一步要求把两种方法间差异的绝对值进行排序。我们将剔除一切 0 差异, 接着把剩下的差异绝对值由低到高进行重排。相同的差异将被赋予它们在混合数据列中位置的平均秩。差异绝对值的排序情况列在表 6—37 中的最右列。应注意到工人 8 的 0 差异被从排序中剔除出去了, 于是最小的差异绝对值 0.2 的秩被定为 1, ..., 0.9 的秩被定为 10。工人 3 和 7 的相同的绝对差的秩被定为平均秩 3.5, 同时工人 4 和工人 10 的相同的绝对差的秩被定为平均秩 5.5。

从表 6—37 可以看出, 差异为负的只有 2 项, 它们分别具有秩 1 和 2, 因此负号的秩和为 $1+2=3$ 。正号的项数是 8 项, 秩和为 52。查附表 9, 对于 $n_1=2$, $n_2=8$, 双侧临界值是 (3, 19), 我们例中对应较少项的符号为负号, 其秩为 3, 未落入上述区间 (此区间不含两侧端点), 因此我们拒绝原假设, 即认为二者有显著差异。

表 6—37 两种方法完工时间绝对差的秩

工人	差异	差异的绝对值	秩
1	0.7	0.7	8
2	-0.2	0.2	1
3	0.4	0.4	3.5
4	0.5	0.5	5.5

续前表

工人	差异	差异的绝对值	秩
5	-0.3	0.3	2
6	0.9	0.9	10
7	0.4	0.4	3.5
8	0.0	0.0	—
9	0.6	0.6	7
10	0.5	0.5	5.5
11	0.8	0.8	9
		负符号秩之和	3
		正符号秩之和	52

用 T 代表威尔科克森符号秩检验中的符号秩的和。当两个总体相同且两组数据中至少一组个数超过 10 时, T 的抽样分布可以认为近似于正态概率分布, 计算机将直接给出近似结果的 p -value, 利用此值就能得出检验的结论。

6.5.4 克鲁斯卡尔-沃利斯检验

MWW 检验可以用于检验两个总体是否相同, 而克鲁斯卡尔和沃利斯将检验扩展到三个或更多总体的情形。 $k \geq 3$ 的总体克鲁斯卡尔-沃利斯 (Kruskal-Wallis) 检验的假设可写成如下的形式:

H_0 : 所有总体均相同, H_1 : 并非所有总体均相同

克鲁斯卡尔-沃利斯检验基于对来自于 k 个总体中每一个总体的独立随机样本的分析。

在前面, 我们表明方差分析 (ANOVA) 可被用于检验 3 个或 3 个以上总体的均值的相等性。ANOVA 过程要求的条件是: 所有数据为定距或定比数据, 所有总体为正态分布, 且各总体的方差均相等。

非参数的克鲁斯卡尔-沃利斯检验可被用于顺序数据, 或是定距数据以及定比数据。而且, 克鲁斯卡尔-沃利斯检验不要求参数的方差分析过程所要求的正态分布和等方差的假设。于是, 只要当来自于 $k \geq 3$ 的独立随机样本的数据是顺序的, 或者数据为定距的但正态性假设及等方差假设存在问题时, 克鲁斯卡尔-沃利斯检验就提供了一种可用于检验总体是否相同的替代统计方法。我们将它应用于一项关于雇员挑选的问题, 以此来演示克鲁斯卡尔-沃利斯检验。

例 6—28

公司从 3 所当地的大学聘用管理工作人员。最近, 该公司的人事部收集并评出了雇员的年度表现等级, 试图确定从 3 所大学招聘来的管理人员的表现是否存在差异。表现等级数据可以从由 7 名来自 A 大学的雇员、6 名来自 B 大学的雇员以及 7 名来自 C 大学的雇员组成的三个独立样本中得到。表 6—38 汇总了这些数据, 每名管理人员的总体表现等级在 0~100 之间, 其中 100 是最高的可能表现等级。

表 6—38

20 名雇员的表现等级

A 大学	25	70	60	85	95	90	80
B 大学	60	20	30	15	40	35	
C 大学	50	70	60	80	90	70	75

假设我们想检验在表现的评价方面这三个总体间是否存在差异。克鲁斯卡尔-沃利斯检验统计量以每个样本的秩和为基础,用下列公式算出:

$$W = \left[\frac{12}{n_T(n_T+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} \right] - 3(n_T+1)$$

式中: k ——总体的个数;

n_i ——样本 i 中的个体数;

$n_T = \sum n_i$ ——所有样本的个体总数;

R_i ——样本 i 的秩和。

克鲁斯卡尔和沃利斯证明了,在各总体相同的原假设下, W 抽样分布近似自由度为 $k-1$ 的卡方分布。若每个样本的容量均大于等于 5,则这一近似误差很小。

为了计算例子中的 W 统计量,必须首先对全部 20 个数据项进行排秩,来自于大学 B 样本的 15 是最小数据值,其秩为 1;而来自于大学 A 样本的 95 是最高数据值,其秩为 20,下表中给出了数据的值、其相应的秩以及三个样本各自的秩和。应注意到对于同分值,我们将其秩赋为秩的平均数。例如,数据值 60, 70, 80 和 90 都有同分值。排秩结果列在表 6—39 中。

表 6—39

20 名雇员的混合秩

A 大学	秩	B 大学	秩	C 大学	秩
25	3	60	9	50	7
70	12	20	2	70	12
60	9	30	4	60	9
85	17	15	1	80	15.5
95	20	40	6	90	18.5
90	18.5	35	5	70	12
80	15.5			75	14
秩和	95	秩和	27	秩和	88

样本容量为:

$$n_1 = 7, n_2 = 6, n_3 = 7$$

$$n_T = \sum n_i = 7 + 6 + 7 = 20$$

计算 W 统计量:

$$W = \frac{12}{20 \times 21} \left(\frac{95^2}{7} + \frac{27^2}{6} + \frac{88^2}{7} \right) - 3 \times (20+1) = 8.92$$

卡方分布表(本书附表 4)给出在自由度为 $k-1=2$ 时,右侧的 $\alpha=0.05$ 的卡方临界值为 $\chi^2=5.991$,由于本检验的统计量 $W=8.92$,大于 5.991,我们拒绝三个总体相

同的原假设。作为结果,我们得出结论:来自不同大学的管理人员业绩显著不同。而且,由于表现等级最低的是大学B,所以此公司将减少从大学B招聘的员工数或者至少是更深入地评价此大学的毕业生。

用 MINITAB 软件,从 STAT-Nonparametric-Kruskal Wallis 入口,可以直接得到下列结果:

Kruskal-Wallis Test: Grade versus univ				
Kruskal-Wallis Test on Grade				
univ	N	Median	Ave Rank	Z
A	7	80.00	13.6	1.70
B	6	32.50	4.5	-2.97
C	7	70.00	12.6	1.15
Overall	20		10.5	
H = 8.92 DF = 2 P = 0.012				
H = 8.98 DF = 2 P = 0.011 (adjusted for ties)				

这里最后两行都是计算结果,带有 adjusted for ties 的最后一行,是在计算秩时,对于观测值过多相等情况的秩的计算作了修正。而我们通常在实际使用此计算机程序时只看最后一行就行了。

从最后一行结果中可以看出, p -value 为 0.011, 因此拒绝原假设, 即应认为各总体是不同的。

6.5.5 数据独立性的链检验

在实际工作中常常要检验观测数据是否独立。所谓数据间相互独立指的是: 连续观测的一些数据之间, 后面的数据不受前面数据的影响。比较典型的不独立数据是股票市场上的数据, 后一日的股票价格将明显受到前一日股票价格的影响。对于不能保证相互独立的数据进行统计分析是相当困难的, 通常要使用“时间序列分析”工具。如果数据间不相互独立, 则我们前面所介绍的各种方法与结论都将失效, 例如中心极限定理中介绍的样本均值的标准差将减小为原标准差的 $\frac{1}{\sqrt{n}}$ 也将不能成立。那么如何检验数据是否相互独立呢? 这

方面有很多方法, 本节将只介绍最简单的链 (run) 检验法, 也称为游程检验法。

在一串字符组成的序列中, 接连相同的符号组成的一串相同字符被称为一个“链”。例如, 一串由正负号组成的序列是这样的:

++++--++---+-----

可以看出本序列中, 正号一共有 7 个, 负号一共有 9 个, 而连在一起的正号序列有 3 个, 称为有 3 个正号链; 连在一起的负号序列也有 3 个, 称为有 3 个负号链, 总共有 6 个链。如果一个序列由 n_1 个正号和 n_2 个负号组成, 记 r 为其链的总个数, 那么正常情况下



该有多少个链呢?在本书的附表 10 中的左上角和右上角的两个表中,分别给出了链的个数的上下界。在本书的附表 11 中的下半部给出了链的长度的上界。

可以设想,如果一个序列数据相互独立时,它们的链的总个数不会太多也不会太少,最长的链的长度也不应太长。如果一个序列数据不相互独立,则它们的链的总个数会不正常,其最长的链的长度也可能太长。比如,考察某月的上证指数,其月中位数为 1 300,把 26 个交易日的每日上证指数记录下来,把高于 1 300 者记为+号,把低于 1 300 者记为一号,则得到的序列为:

+++++---+-----

由于股市总有这样的特点:某天指数较高者则下一天指数也常常较高,某天指数较低者则下一天指数也常常较低,因而链的个数通常偏少,而其最长的链的长度也可能太长。数一下此序列,可以得知总的链的个数只有 4 个,最长的一个链长度为 8,这就显示数据不独立。另一种情况则正好相反,比如,为了把一个球投入筐中,如果第一次偏远,为了争取下次投掷入筐,常常会“矫枉过正”,下次会偏近,再下次又会偏远……这样一来,链的总个数会偏多。因此,对于链的检验可以从链的总个数和最长的链长两方面进行。

我们所要检验的是:

H_0 : 数据相互独立, H_1 : 数据不独立

像上面的那个例子,正负号各 13 个,查附表 10 左上角,可以确定下界值为 8,查附表 10 右上角,可以确定上界值为 20。也就是说,如果此序列的链的个数小于或等于 8 则拒绝原假设,判数据不独立;如果此序列的链的个数大于或等于 20 则同样拒绝原假设,判数据不独立。本例因样本链的个数只有 4,故应认为数据不独立。在一般数据例题中,计算链的个数本身就很麻烦,MINITAB 软件内的 Stat-Quality Tools-Run Chart 可以自动计算出链的个数。

对于链长可以同时进行检验,按正负号个数,查附表 11 中的下半部,可以得知链的长度的上界,如果数据中的最长链长大于或等于表中的数,则应该拒绝原假设,即应认为数据是不独立的。在本例中,最长的链长为 8,按正负号各 13 个查附表 11 中的下半部,可以得知链的长度的上界为 8,我们的链长已经达到上界,因而也能判定数据不独立。

对于链的总数和链长的检验,只要任何一个检验结论是拒绝原假设,则我们的结论就是拒绝原假设。

第 5 章已经讲过链图,我们采用第 5 章图 5—36 的链图为例,进行下列检验:

(1) 链最大长度检验:序列总长度 $N=20$,链最大长度 $L=8$,查附表 11 下半部的表($\alpha=0.05$),对应 (10, 10) 的最大长度为 7,本例最大长度 $L=8$,故对 H_0 应予以拒绝,即认为数据并不独立。

(2) 链的数目检验:序列总长度 $N=20$,链总数只有 5 个。查附表 10 左上角表($\alpha=0.05$),对应 (10, 10) 的下界为 6,本例链总数只有 5 个,小于下界,故对 H_0 应予以拒绝,即认为数据并不独立。

(3) 趋势检验:连续下降的点达到 7 个,查附录表 8 符号检验表,对于 $N=7$, $\alpha=0.05$ 时,表中列出下界为 0,故对 H_0 应予以拒绝,即认为数据并不独立。

样本量超过 40 时, 查表给不出临界值, 这时可以使用近似正态的方法。由于软件功能的限制, 我们通常要先求出中位数, 再从 Stat-Nonparametric-Run Test (游程检验) 入口, 以中位数为界来作链检验。计算机在求出游程总数的同时, 还用正态近似法计算出 p -value, 当 p -value 小于 0.05 时, 拒绝原假设。

6.6 相关分析与回归分析

在六西格玛管理的分析阶段 (analyze) 中, 分析问题出现的原因是至关重要的。这时, 统计学中的相关分析及回归分析就是很有力的工具。

如果同时获得了两个或多个连续型变量的观测值, 就可以使用相关分析和回归分析了。例如, 在一定的范围内, 反应罐中的温度与最终的产量间, 就呈现某种关系。分析并确认二者是否有关系, 这就是相关分析的任务; 把二者的关系用方程的形式表达出来, 这就是回归分析的任务。回归分析的任务可能更复杂, 影响最终产量的, 除了温度之外, 还可能有反应时间、反应罐中的压力等多个因素, 因此回归分析又分单自变量的回归 (称一元回归) 和多自变量的回归 (称多元回归)。除了线性回归外, 还可能有二次回归、三次回归等多项式回归以及更多自变量的逐步回归等。限于篇幅, 我们下面将只介绍简单线性回归分析方法。

6.6.1 相关分析

如果拥有 X 与 Y 的成对数据, 则可以在散点图中直观地看见这种关系。 X 与 Y 的关系可能是密切的, 也可能是不密切的。

判断 X 和 Y 之间的相关程度, 通常先从散点图可以获得一个大概的印象。我们在 6.1 节中曾经描述过, 如果数据像在图 6—2 中展示的那样, 就可以看出两个变量间是正相关还是负相关或是基本上无关。当然, 如果只是获得些初步印象, 那是远远不够的, 我们要定量地对相关关系加以分析。

某团队在分析产品加工温度与产量之间的关系时, 收集了以下数据 (见表 6—40)。

表 6—40

温度 X 与产量 Y 数据

序号	温度 X	产量 Y
1	2	58
2	6	105
3	8	88
4	8	118
5	12	117
6	16	137
7	20	157
8	20	169
9	22	149
10	26	202



我们可以使用散点图初步分析出过程中的产量与反应温度有关。产量与反应温度的散点图如图 6—12 所示。除了直观上所能看到的存在某种关系之外，我们当然更希望能够量化出这种相关性的程度——究竟密切的程度如何？相关分析将给出明确的结果。更进一步，如果在一定的范围内，对某个给定的温度点能够预测出产量，那将是有实际用途的。解决这个问题的有力工具是回归分析，它能够帮助我们确定所研究的两个或多个变量间的函数关系。我们先来讨论相关分析方法。

首先画出产量与温度的散点图（见图 6—12）。

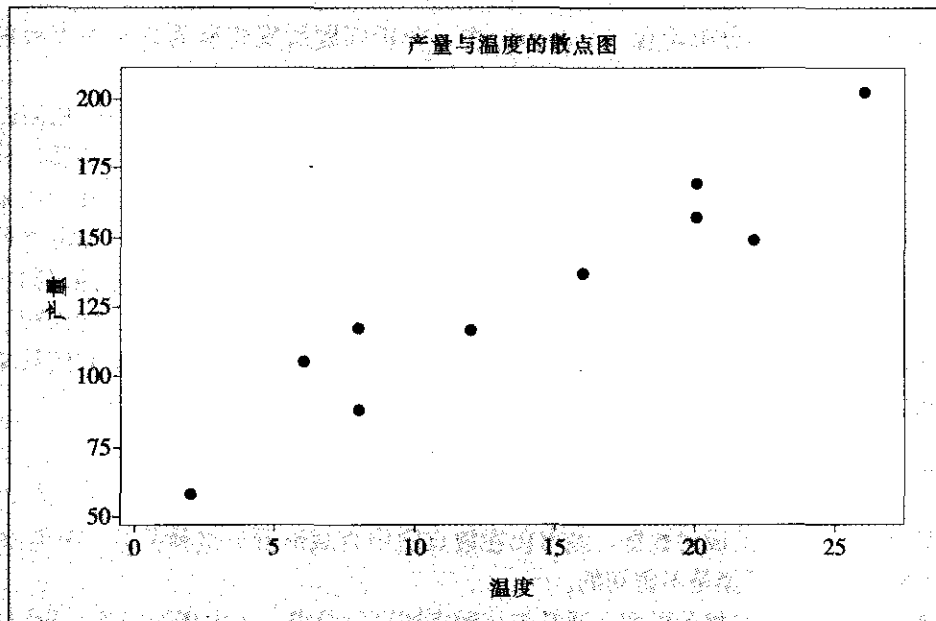


图 6—12 温度对产量的影响

从图中可以明显地看出，随着温度的增长，产量也显著地增长。那么如何度量相关的程度？这就要引入相关系数的概念。

1. 相关系数的概念

相关系数 r （通常是指 Pearson 相关系数）是用来描述两个变量线性相关程度的一种度量。由两个变量的 n 组观测值可以画出多种不同形状的散点图，我们希望用相关系数 r 来表示出它们是正相关还是负相关以及它们以何种程度相关。

当散点图呈现图 6—12 或图 6—2a 的形状，即随 x 的值增加时， y 有明显增加的趋势，应该得到一个正的相关系数，而且两者的关系越接近于一条直线，则此系数越接近于 1；反之，当散点图呈现图 6—2b 的形状，相关系数应该是个负数。因此，样本相关系数 r 被定义为：

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}$$

式中: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ 分别为变量 x 与 y 的平均值;

$L_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}$, 是诸 x_i 的偏差与诸 y_i 的偏差的乘积和;

$L_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2$, 是诸 x_i 的偏差平方和;

$L_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2$, 是诸 y_i 的偏差平方和。

我们不必仔细研究如何按上述公式具体计算相关系数, 但要清楚了解相关系数 r 的具体意义。数学上可以证明, 一定有关系式: $-1 \leq r \leq 1$ 或 $|r| \leq 1$ 。因此, 当 $r = \pm 1$ 时, 表明 n 个点全在一条直线上, 这时两个变量间完全线性相关; $|r|$ 越接近于 1, 则表明各个点靠近直线越近; 随着 $|r|$ 的逐渐减小, 各个点就越来越远离直线, 呈现非常分散的状况。总之, 当 $r > 0$ 时, 称两个变量间为正相关, 这时当 x 的值增加时, y 的值也有增大的趋势 (如图 6—2a), r 的值越大说明二者关系越密切; 当 $r < 0$ 时, 称两个变量间具有负相关, 这时当 x 的值增加时, y 的值有减小的趋势 (如图 6—2b), r 的绝对值越大说明二者关系越密切。

2. 相关系数的检验

若 x 与 y 无关时, 总体的相关系数 ρ 应该为 0, 样本的相关系数 $|r|$ 会很小, 这时, 散点图上 n 个点会毫无规律地形成“乱作一团”的形状 (如图 6—2c)。但是反之, 如果总体的相关系数 ρ 为 0, 样本的相关系数 $|r|$ 很小时, 我们却不能肯定 x 与 y 无关, 因为很可能二者形成某种曲线关系 (例如图 6—2d 所显示的那样是抛物线型的二次函数关系)。我们只能肯定地说: 如果总体的相关系数 ρ 为 0, 样本的相关系数 $|r|$ 很小时, 则 x 与 y 无线性相关关系。

问题是样本相关系数大到什么程度可以认为两个变量间确实存在相关关系呢? 必须指出, 判断样本相关系数是否足够大的标准与样本量有非常密切的关系, 不可能有一个统一的标准。一般说来, 样本量超过 9 时, 只要相关系数绝对值达到 0.7, 那么就应当认定两变量间是确实相关的; 当样本量超过 25 时, 只要相关系数绝对值达到 0.4, 那么就应当认定两变量间是确实相关的。手工计算相关系数的检验方法可以查看附表 12, 计算机的检验方法将在下面介绍。

下面对例 6—29 进行相关系数 r 的具体计算。

计算均值: $\bar{x} = 14, \bar{y} = 130$

计算平方和与乘积和: $\sum x_i^2 = 2\,528, \sum x_i y_i = 21\,040, \sum y_i^2 = 184\,730$

计算: $L_{xx} = 2\,528 - 10 \times 14^2 = 568$

$$L_{yy} = 184\,730 - 10 \times 130^2 = 15\,730$$

$$L_{xy} = 21\,040 - 10 \times 14 \times 130 = 2\,840$$

计算:

$$r = \frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}} = \frac{2\,840}{\sqrt{568 \times 15\,730}} = 0.950\,1$$

用 MINITAB 软件从 STAT-Basic Statistics-Correlation 入口, 可以直接得到下面的



结果:

Correlations: 温度, 产量

Pearson correlation of 温度 and 产量=0.950

这里确实见到 X 与 Y 间高度正相关的情况。

如果要对 X 与 Y 之间相关系数是否为 0 进行检验, 可以记 ρ 为总体相关系数, 进行假设检验。

$$H_0: \rho=0, H_1: \rho \neq 0$$

检验方法与一般参数检验方法是相同的。用 MINITAB 计算时, 只要在上述计算过程中增加输出 p -value 即可判断出结论。

Correlations: 温度, 产量

Pearson correlation of 温度 and 产量=0.950

P-Value=0.000

此检验得到的 p 值小于 0.05, 故拒绝原假设, 即应该认为 X 与 Y 是线性相关的。

6.6.2 一元线性回归模型

假定有两个变量, 其中:

x 是自变量, 其值是可以控制或精确测量的, 它是非随机变量。

y 是因变量, 对给定的 x 值, y 的取值事先不确定, y 是随机变量。

假设 (x, y) 的散点图显示有直线关系, 则我们可以认为观测值 y 由两部分迭加而成: 一是 y 的平均值随 x 的变化而呈线性变化的趋势, 用 $b_0 + b_1 x$ 表示; 二是其他随机因素影响到 y 值本身会偏离平均值, 其误差用 ϵ 表示, 常设 $\epsilon \sim N(0, \sigma^2)$, 故有如下的数据结构式:

$$y_i = b_0 + b_1 x_i + \epsilon_i, \quad i=1, 2, \dots, n$$

1. 回归系数的最小二乘估计

按最小二乘法来估计, 先定义:

$$Q(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

若 \hat{b} 与 \hat{b}_1 满足如下等式:

$$Q(\hat{b}_0, \hat{b}_1) = \min_{b_0, b_1} Q(b_0, b_1)$$

即估计量 \hat{b}_0, \hat{b}_1 可使 $Q(b_0, b_1)$ 达到极小, 则称 \hat{b}_0, \hat{b}_1 为 b_0, b_1 的最小二乘估计。

用求导函数的方法得

$$\begin{cases} \frac{\partial Q}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0 \\ \frac{\partial Q}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) x_i = 0 \end{cases}$$

解之,得

$$\begin{cases} \hat{b}_1 = \frac{l_{xy}}{l_{xx}} \\ \hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x} \end{cases}$$

可以验证: \hat{b}_0, \hat{b}_1 确实可使 $Q(b_0, b_1)$ 达到最小,故 \hat{b}_0, \hat{b}_1 为回归系数的最小二乘估计法计算结果。

2. 手工计算的步骤

(1) 求出 $\sum x_i, \sum y_i$ 和 \bar{x}, \bar{y} ;

(2) 求 $\sum x_i^2, \sum x_i y_i, \sum y_i^2$;

(3) 计算 $L_{xx} = \sum x_i^2 - n\bar{x}^2, L_{xy} = \sum x_i y_i - n\bar{x}\bar{y}, L_{yy} = \sum y_i^2 - n\bar{y}^2$;

(4) 代入公式即可求出 b_0, b_1 的最小二乘估计。

例 6-29

进行 b_0, b_1 的计算,利用 MINITAB 应用软件,由 STAT-Regression-FittedLine-Plot 可以直接得到回归系数的计算结果: $b_0=60, b_1=5$ 。

回归直线方程 $Y=60+5X$, 其中 $R-Sq=90.3\%=r^2$

Regression Analysis: 产量 (Y) versus 温度 (X)

The regression equation is

产量 = 60.0 + 5.00 温度

Predictor	Coef	SE Coef	T	P
Constant	60.000	9.226	6.50	0.000
温度	5.000 0	0.580 3	8.62	0.000
S=13.83	R-Sq=90.3%	R-Sq (adj) =89.1%		

STAT-Regression-Fitted Line Plot 窗口只对求一个自变量的回归方程有效,但它的优点是可以显示出回归直线的图形(见图 6-13)。

这里特别要说明,回归系数 b_1 是有明确的物理意义的。那就是,当自变量 X 增加一个单位时, Y 平均说来会增加 b_1 个单位。很显然, b_1 是有量纲的,它的量纲是 $[Y] / [X]$ 。比如本例,回归系数 $b_1=5$,说明温度每升高 1 度,平均说来产量会增加 5 千克。另外,回归直线还有一个特点,回归直线总是会经过 $(0, \hat{b}_0)$ 和 (\bar{x}, \bar{y}) 这两点的。这可以作为计算结果的校验使用。

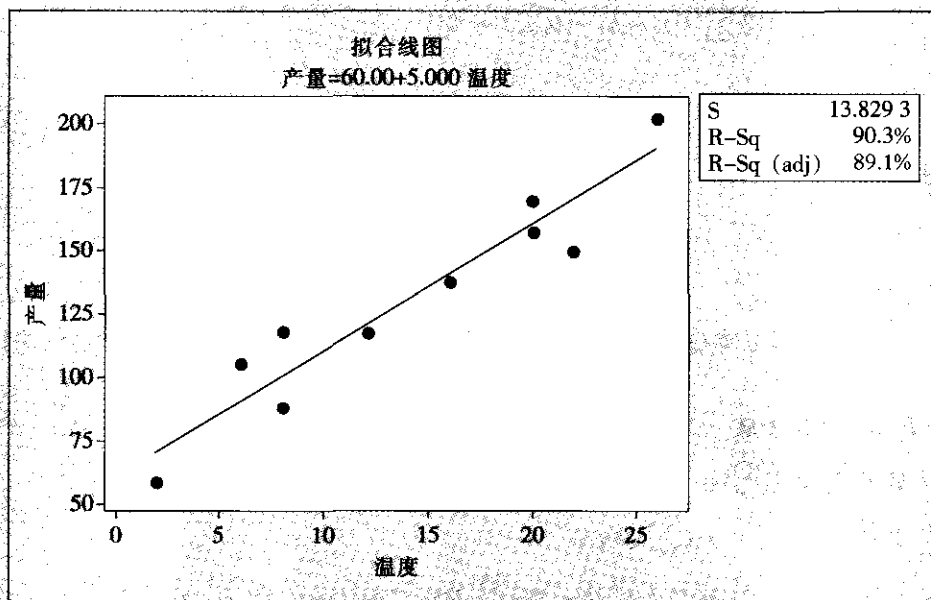


图 6—13 最小二乘拟合的回归直线图

3. 回归方程的显著性检验

我们建立回归方程的目的是去表达两个具有线性相关的变量间的定量关系, 因此, 只有当两个变量确实具有线性相关关系时所建立的回归方程才是有意义的。两个变量间是否存在线性相关关系, 有两种检验方法。

方法之一是上一小段所叙述的求两个变量间的相关系数, 这时我们是将自变量 X 也看作随机变量。对于给定的显著性水平 α , 当相关系数 r 的绝对值大于临界值 $r_{1-\alpha}(n-2)$ 时 (其数值列在附录表 12 中), 便认为两个变量间存在线性相关关系, 所求得的回归方程是有意义的。

方法之二是用方差分析法。对给出的 n 个 y 的观测值求出其总的波动, 如同方差分析中一样, 用 SS_T 表示总偏差平方和:

$$SS_T = \sum (y_i - \bar{y})^2$$

造成这种波动的原因有两个方面: 一是由于自变量 x 取值不同, 当变量 y 与 x 线性相关时, x 的变化会引起 y 的变化, 反映在 y 的预测值的波动上, 即回归分析的效应方面; 另一个是各实测 y 值与预测 y 值间的差别, 即随机误差方面。我们可以用回归平方和 SS_R 与残差平方和 SS_E 分别表示由这两个原因引起的数据波动:

$$SS_R = \sum (\hat{y}_i - \bar{y})^2$$

$$SS_E = \sum (\hat{y}_i - y_i)^2$$

同样可以证明有平方和分解式:

$$SS_T = SS_R + SS_E$$

它们的自由度也有分解式: $df_T = df_R + df_E$, 式中:

$$df_T = n - 1, df_R = 1 \text{ (即自变量的个数)}, df_E = df_T - df_R$$

如同方差分析中一样, 计算 F 比:

$$F = (SS_R/df_R) / (SS_E/df_E)$$

对给定的显著性水平 α , 当 $F > F_{1-\alpha}(df_R, df_E)$ 时, 认为回归方程是有意义的。

例 6-29

下面我们利用 MINITAB 应用软件对回归方程的显著性进行检验。

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	14 200	14 200	74.25	0.000
Residual Error	8	1 530	191		
Total	9	15 730			

这里, p -value 比 0.001 还小, 故认为回归方程是有意义的。

用手工计算, 其中方差来源:

$$\text{回归平方和 } SS_R = \sum (y_i - \bar{y})^2 = 14\,200$$

$$\text{均方差 } MS_R = SS_R/df_R = 14\,200$$

$$\text{误差平方和 } SS_E = \sum (y_i - \hat{y}_i)^2 = 1\,530$$

$$\text{均方差 } MS_E = SS_E/df_E = 191.25$$

$$F = (SS_R/df_R)/(SS_E/df_E) = 74.25$$

在 $\alpha = 0.01$ 时, $F_{1-\alpha}(1, 8) = 11.26$, 现在 $F > 11.26$, 这表明在 $\alpha = 0.01$ 水平上方程是有意义的。

4. 利用回归方程作预测

当求得了回归方程 $\hat{y} = \hat{b}_0 + \hat{b}_1 x$, 并经检验确认回归方程是显著的, 则可以用回归方程做预测。

所谓预测是指当 $x = x_0$ 时对相应的 y 的取值 y_0 所作的推断。由于 y 是随机变量, 其实际取值是无法准确预测的, 我们只能对其平均取值做出估计, 这便称为 y 的预测值。显然, 如果 $x = x_0$, 那么 y 的预测值的估计值为:

$$\hat{y}_0 = \hat{b}_0 + \hat{b}_1 x_0$$

但是, 由于方程中的系数也是由样本估计出来的, 因此系数本身就有波动性, 这导致此估计值 \hat{y}_0 也就有波动性, 即 y_0 的真正的理论上的预测值与 \hat{y}_0 就可能有偏离, 再加上模型中本来就假设 y 具有围绕其均值波动的随机性, 为了得到 y_0 的预测值, 我们要求这种绝对偏差 $|y_0 - \hat{y}_0|$ 不超过某个 δ 的概率为 $1 - \alpha$, 其中 α 是事先给定的一个比较小的数 ($0 < \alpha < 1$), 即

$$P(|y_0 - \hat{y}_0| \leq \delta) = 1 - \alpha$$



或

$$p(\hat{y}_0 - \delta \leq y_0 \leq \hat{y}_0 + \delta) = 1 - \alpha$$

$[\hat{y}_0 - \delta, \hat{y}_0 + \delta]$ 就称为 y_0 的概率为 $1 - \alpha$ 的置信区间。其中 δ 的表达式为:

$$\delta = \delta(x_0) = t_{1-\alpha/2}(n-2) \hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{L_{xx}}}$$

式中, $\hat{\sigma} = \sqrt{SS_E/(n-2)}$, $SS_E = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = L_{yy} - \hat{\beta}_1^2 L_{xx}$ 。

$t_{1-\alpha/2}(n-2)$ 是自由度为 $n-2$ 的 t 分布的 $(1-\alpha/2)$ 分位数, 可查表给出。由上式可以看出预测区间的长度 2δ 与样本量 n , x 的偏差平方和 L_{xx} , x_0 到 \bar{x} 的距离 $|x_0 - \bar{x}|$ 有关, x_0 越远离 \bar{x} , 预测精度就越差。当 x_0 位于所有自变量观测区间之外时, 预测精度可能变得很差, 在这种情况下作外推, 需要特别小心。

图 6—14 给出在不同 x 值处预测区间的示意图: 在 $x = \bar{x}$ 处预测区间最短, 随着离 \bar{x} 的距离越来越远, 预测区间越来越宽, 其图形呈喇叭状。

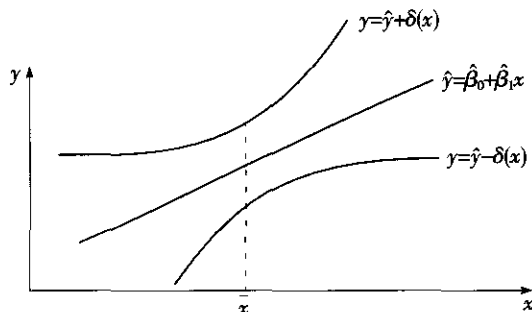


图 6—14 预测区间的示意图

当 n 较大时 (如 $n > 30$), t 分布可以用正态分布来近似, 进一步, 若 x_0 与 \bar{x} 相差不大时, δ 可以近似取为:

$$\delta \approx \hat{\sigma} Z_{1-\alpha/2}$$

式中, $Z_{1-\alpha/2}$ 是标准正态分布的 $1-\alpha/2$ 分位数。

例 6—19

用手工方法计算预测区间

如果取 $x_0 = 10$, 则得预测值为:

$$\hat{y}_0 = 60 + 5 \times 10 = 110$$

求概率为 $1-\alpha$ 的预测区间:

(1) σ 的估计 $s = 13.83$

(2) 由给定的 α , 查 t 分布表的分位数 $t_{1-\alpha/2}$, 比如取 $\alpha = 0.05$, 则 $t_{0.975}(8) = 2.306$;

(3) 按公式计算 δ 的值。本例中 $\bar{x} = 14$, $L_{xx} = 568$, 故

$$\delta = 13.83 \times 2.306 \times \sqrt{1 + \frac{1}{10} + \frac{(10-14)^2}{568}} = 33.87$$

(4) 写出预测区间 $(y_0 - \delta, y_0 + \delta)$, 本例中为 $110 \pm 33.87 = (76.13, 143.87)$ 。

如果希望计算预测区间只能使用功能更强大的另一窗口: STAT-Regression-Regression。此窗口对于任意多个自变量都可以使用, 输出的功能也更全面, 缺点是没有图形的输出。为了对于新自变量值进行预测, 只要在“Option”窗口内填入自变量值 10, 即可得到下列结果:

Predicted Values for New Observations				
New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	110.00	4.95	(98.58, 121.42)	(76.13, 143.87)
Values of Predictors for New Observations				
New Obs	x			
1	10.0			

在计算机显示的结果中此单个观测值的置信区间用“Prediction Interval”表示。另一个置信区间考虑的是: 由于回归方程中的系数也是用样本观测值估计出来的, 因而系数是有波动的, 在回归线上的点的值也会随之波动, 因此, 预测的在该点处的理论均值也会有波动。“Confidence Interval, CI”表示的是, 预测的理论均值的 95% 置信区间, 其置信区间半长公式为下面的公式:

$$\delta_0 = \delta_0(x_0) = t_{1-\alpha/2}(n-2) \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{L_{xx}}}$$

5. 残差分析: 证实模型假定

使用回归方程的上述统计检验可以获得有关 X 与 Y 的丰富信息, 但仍然不能确定设定的回归模型是否与数据拟合得很好。为了确认拟合效果, 必须采用残差分析方法来进行诊断。定义在 x_i 处的残差 r_i 是因变量的观测值 (y_i) 与因变量的估计值 (\hat{y}_i) 之差, 即

$$r_i = y_i - \hat{y}_i$$

换言之, x_i 处的残差是利用估计的回归方程去预测 y_i 而引起的误差。对例 6—29, 残差的计算在表 6—41 中, 因变量的观测值在第 2 列, 利用估计的回归方程 $\hat{y} = 60 + 5x$ 得到因变量的估计值在第 3 列, 对应的残差在第 4 列。对这些残差进行分析将帮助我们确定, 对回归模型做出的这些假定是否成立。

表 6—41

例 6—29 的残差

温度 x_i	产量 y_i	估计的产量 $\hat{y} = 60 + 5x$	残差 $y_i - \hat{y}_i$
2	58	70	-12
6	105	90	15
8	88	100	-12
8	118	100	18
12	117	120	-3



续前表

温度 x_i	产量 y_i	估计的产量 $\hat{y}=60+5x$	残差 $y_i-\hat{y}_i$
16	137	140	-3
20	157	160	-3
20	169	160	9
22	149	170	-21
26	202	190	12

现在让我们重温一下上面例子的回归假定。假设简单线性回归模型为：

$$y=b_0+b_1x+\epsilon$$

这个模型表示，我们假定产量（ y ）是温度（ x ）的一个线性函数加上一个误差项 ϵ 。

我们对误差项 ϵ 做了以下假定：

- (1) $E(\epsilon)=0$ ，即 ϵ 的均值为 0；
- (2) 对所有的 x 值， ϵ 的方差 σ^2 都是相同的；
- (3) ϵ 的值是相互独立的；
- (4) 误差项 ϵ 服从正态概率分布。

我们之所以能利用 t 检验和 F 检验去确定 x 和 y 之间的关系是否显著，并得出置信区间估计和预测区间估计，其理论上的依据皆来源于上述关于误差项 ϵ 的几项基本假定。如果关于误差项 ϵ 的假定不那么可靠，有关回归关系的显著性假设检验和区间估计的结果也就站不住脚。

对于残差的诊断都可以通过对残差图的分析来进行。所谓残差图，其纵坐标总是残差，而横轴可能取不同的值，可以是观测时间，可以是响应变量预测值，也可以是自变量的取值。在对残差图的结果进行解释之前，让我们首先考虑在任意一个残差图中可能观测到的正常形状。如果假定对所有的 x 值， ϵ 的方差都是相同的，并且假定描述变量 x 和 y 之间关系的回归模型是合理的，那么残差图给出的一个总印象就是，所有的散点都应随机地落在一条以 0 为中心的水平带中间。如果不是这样，就意味着这里有问题了。下面详细介绍具体分析。我们将讨论以下几种残差图，这些图通常都是由计算机自动生成的，学习的重点不是如何画这些图，而是如何分析这些图。

(1) 关于按观测顺序的残差图。关于按观测顺序的残差图是用水平轴表示观测顺序，用纵轴表示对应的残差值。这些残差点应在横轴（即残差为 0）上下随机波动，不应有任何上升、下降、摆动、跳跃等趋势。如果有某种趋势存在，则说明数据观测过程中受到某个未知的因素的强大的影响，应该找出此因素并加以控制，此图列在图 6—15 的右下方。

(2) 关于因变量的预测值 \hat{y} 的残差图。第二种残差图是在水平轴上表示因变量的预测值 \hat{y} 和在纵轴上表示残差值。每个残差用图上的一个点来表示。每个点的横坐标由 \hat{y}_i 给出，纵坐标由对应的第 i 个残差值 $y_i-\hat{y}_i$ 给出。对于例 6—29，利用表 6—41 的数据，对应 $\hat{y}_1=70$ 和 $y_1-\hat{y}_1=-12$ ，第一个点的坐标是 (70, -12)；对应 $\hat{y}_2=90$ 和 $y_2-\hat{y}_2=15$ ，第二个点的坐标是 (90, 15) 等等，同样计算可以得到整个残差图（见图 6—15 右上方）。由于我们假定残差的标准差是个常数，它不随预测值而变化，因此，这个残差图上的点应该分布在一条水平的带子中。如果在图中有明显的“喇叭口”形状，即表明残差的标准差

不是常数，而是在随预测值变化，这提示我们原来的模型假定可能有问题。如果有“喇叭口”的形状，可以通过对 Y 作变换加以解决。在本图中没有发现任何不正常之处。

(3) 残差的正态性检验。残差应该服从正态分布，我们对此可以加以检验。一方面我们可以画出残差的直方图（见图 6—15 左下方），也可以将残差画在正态概率纸上（见图 6—15 左上方）。在本图中没有发现任何不正常之处。如果对正态性有怀疑，可以直接对残差数据进行正态性检验。

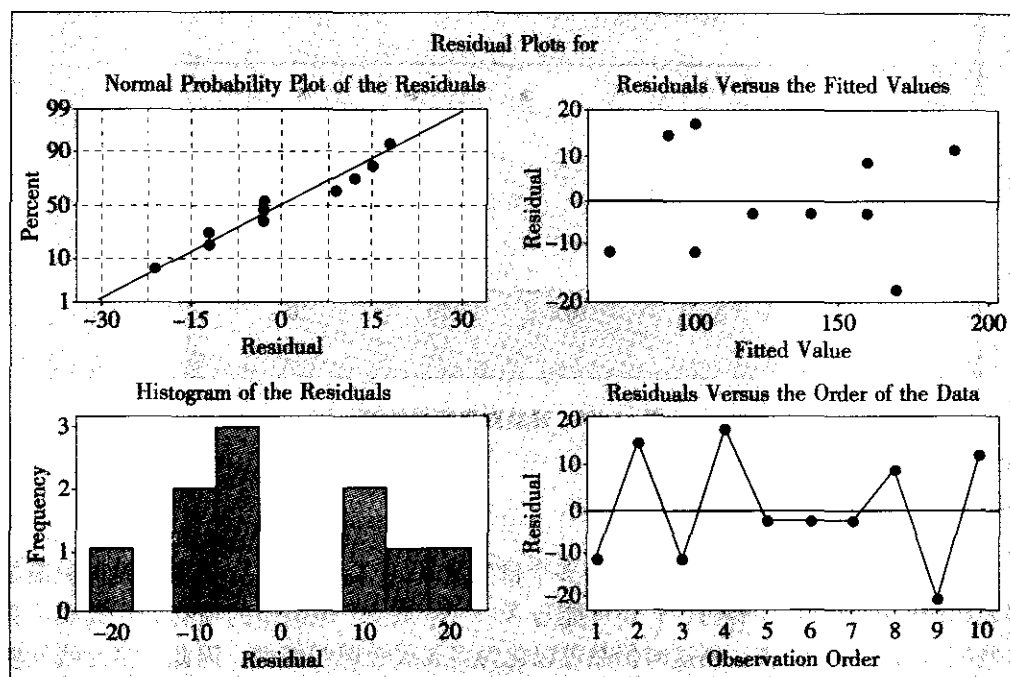


图 6—15 残差图

(4) 关于自变量 x 的值的残差图。关于自变量 x 的值的残差图是用水平轴表示自变量的值，用纵轴表示对应的残差值。每个残差用图上的一个点来表示。每个点的横坐标由 x_i 的值给出，纵坐标由对应的残差值给出。对于例 6—29，根据表 6—41 的数据可以做出关于 x 的残差图。对应 $x_1=2$ 和 $y_1-\hat{y}_1=-12$ ，第一个点的坐标是 $(2, -12)$ ；对应 $x_2=6$ 和 $y_2-\hat{y}_2=15$ ，第二个点的坐标是 $(6, 15)$ 等等。图 6—16 是得到的对自变量的残差图。对于此图分析的重点是看残差的标准差是否能保持常数（即各残差点是否有“喇叭口”）或是否有弯曲的形状。如果对 y 的预测值的残差图也不正常，首先考虑对 y 要做变换；而如果对 y 的预测值的残差图正常，这时说明在回归方程中仅有线性项不够，要增加 X 的高阶项或其他项。图 6—16 完全正常。

现在让我们总结图 6—15 及图 6—16 所示的两张残差图。这些残差看来都是正常的。因此我们的结论是，残差图并没有提供足够的证据，使我们对例 6—29 的回归模型所做的假定表示怀疑。这将导致我们在最后能够确信，例 6—29 的简单线性回归模型是合理的。

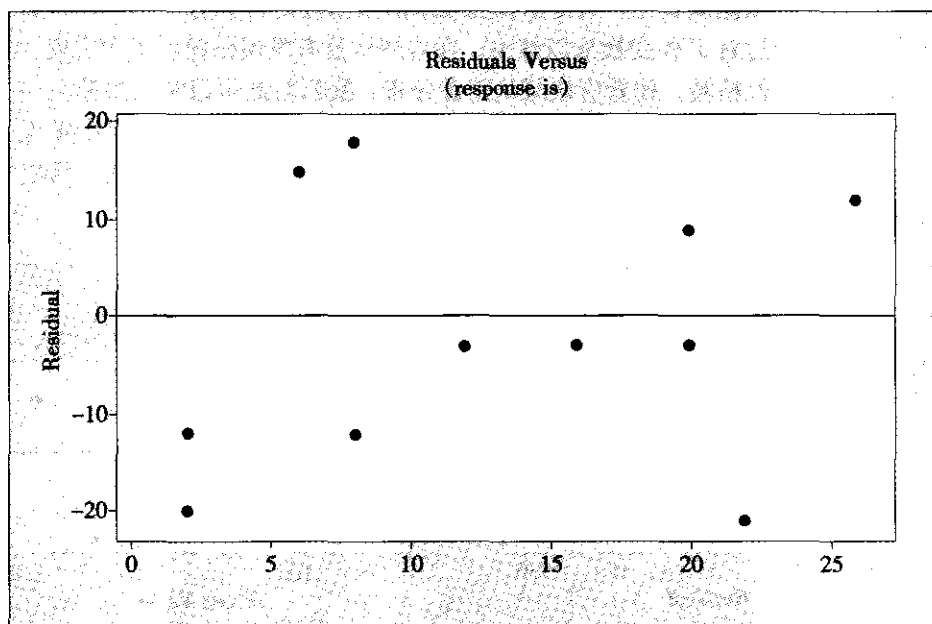


图 6—16 残差对自变量的图

6.6.3 多元线性回归模型

在实际工作中，因变量 Y 通常会与多个自变量 X_1, X_2, \dots, X_k 有关，这时就需要建立多元的线性回归方程。有时， Y 还可能与 $X_1 \times X_2$ 或 X_1^2 等类型的项有关，其实，只要令这些项为新的自变量项，则这时仍然可以化为多元线性回归模型。因此，多元线性模型是一种用途非常广泛的模型。

考察亚硝酸氨的产量提高问题。我们知道产量（以千克为单位）与温度和反应时间有关系。查阅到产量与温度及反应时间的数据如表 6—42 所示。

表 6—42 亚硝酸氨的产量数据表 单位：千克

	温度	时间	产量 Y
1	60	20	23.7
2	60	23	29.6
3	60	26	28.3
4	65	20	33.5
5	65	23	39.4
6	65	26	38.4
7	70	20	43.7
8	70	24	50.0
9	70	28	43.5

续前表

	温度	时间	产量Y
10	75	20	53.6
11	75	25	59.6
12	75	30	45.7
13	80	20	63.6
14	80	25	69.7
15	80	30	55.6

从 STAT-Regression-Regression 入口, 可以得到Y 对于温度和时间两个自变量的回归方程。其结果汇总如下:

回归方程是:

$$y = -72.2 + 1.79 \text{ temp} - 0.320 \text{ time}$$

方差分析表可以用来检验回归的总效果:

H_0 : 回归效果不显著, H_1 : 回归效果显著

从本方差分析表中的 p 值来看, 它几乎为 0, 应拒绝原假设, 即可以认为回归的总效果是显著的。从回归系数的检验中, 可以看出, 温度的效应显著 (p 值为 0.000), 而时间的效应不显著 (p 值为 0.381)。

The regression equation is

$$Y = -72.2 + 1.79 \text{ temp} - 0.320 \text{ time}$$

Predictor	Coef	SE Coef	T	P
Constant	-72.17	13.00	-5.55	0.000
temp	1.786 5	0.172 7	10.34	0.000
time	-0.320 4	0.352 5	-0.91	0.381

S=4.585 36 R-Sq=90.1% R-Sq (adj) =88.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	2 309.0	1 154.5	54.91	0.000
Residual Error	12	252.3	21.0		
Total	14	2 561.3			

从残差的四合一图 (见图 6—17) 中, 粗略看一切似乎是正常的, 但从对各自变量的残差分析中可以看出, 残差对温度是正常的, 而残差对时间有明显的弯曲趋势, 这提示我们应该增加时间的二次项。

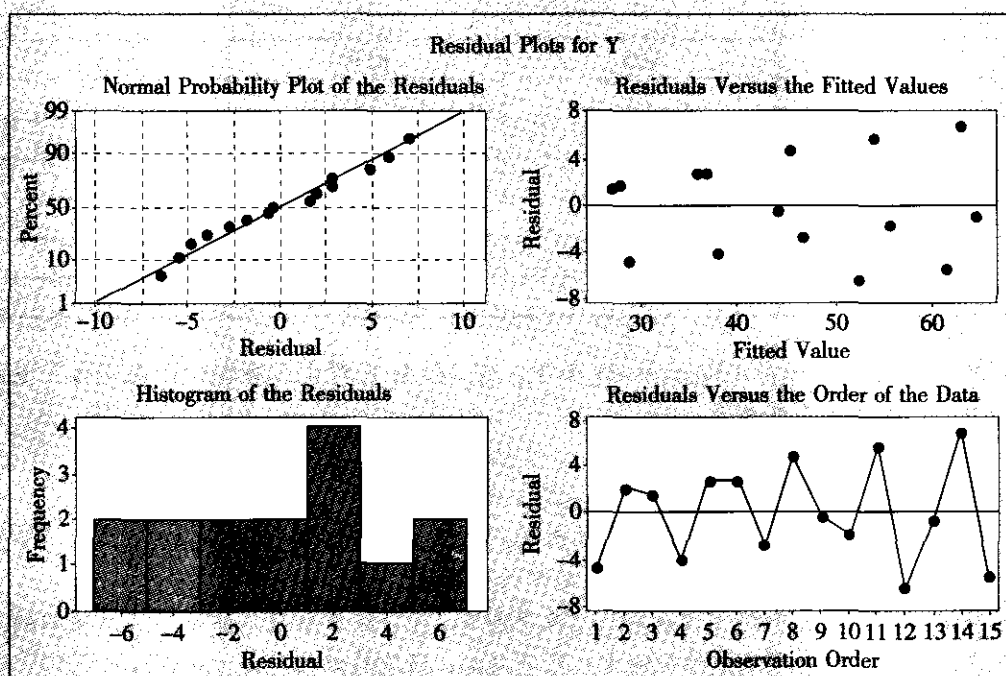


图 6—17 例 6—30 残差四合一图

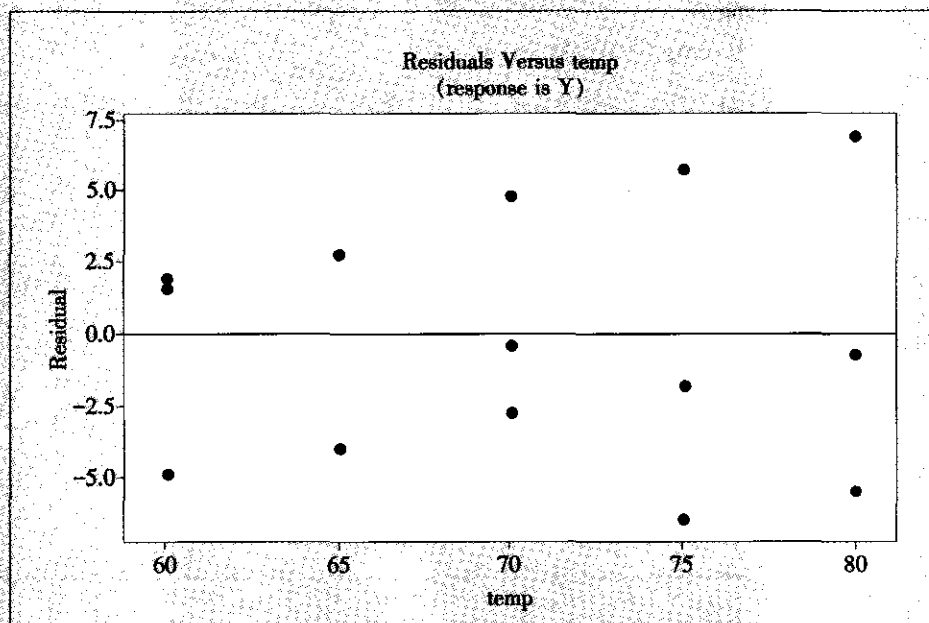


图 6—18 例 6—30 残差对于温度的散点图

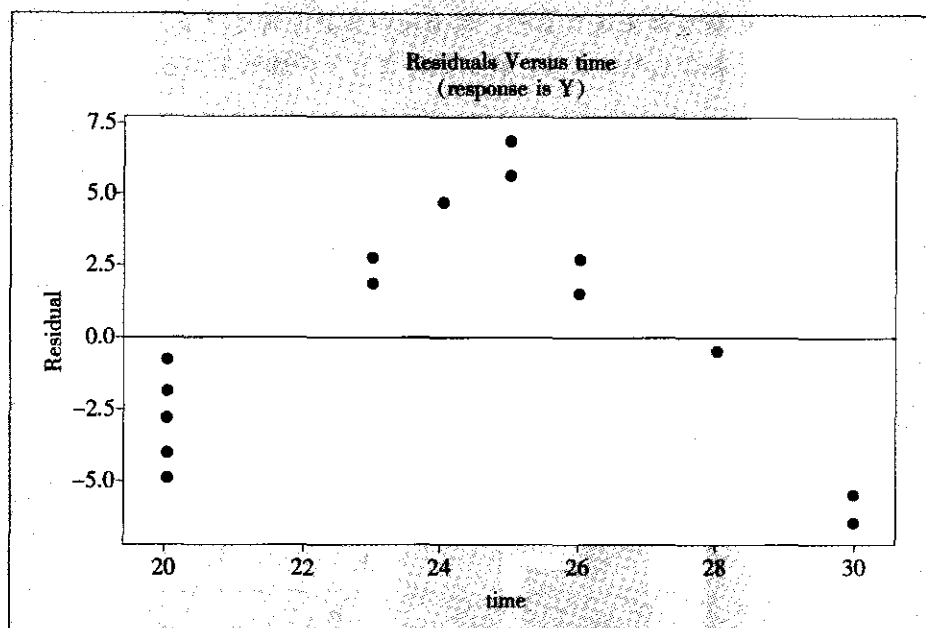


图 6-19 例 6-30 残差对于时间的散点图

我们增加时间的平方项作为 time2，再次进行回归，则可以得到下列结果：

Regression Analysis: Y versus temp, time, time2

The regression equation is

$$Y = -320 + 2.00 \text{ temp} + 19.1 \text{ time} - 0.398 \text{ time}^2$$

Predictor	Coef	SE Coef	T	P
Constant	-319.572	1.443	-221.48	0.000
temp	2.003 39	0.003 65	548.76	0.000
time	19.113 0	0.111 7	171.07	0.000
time2	-0.398 273	0.002 285	-174.28	0.000

S = 0.091 125 5 R-Sq = 100.0% R-Sq (adj) = 100.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	2 561.22	853.74	102 812.29	0.000
Residual Error	11	0.09	0.01		
Total	14	2 561.31			



此结果证明我们的模型效果是显著的, 各项回归系数也都是显著的, 从残差图 (见图 6—20 至图 6—22) 来看也都很正常, 我们最终找到了合适的回归模型。

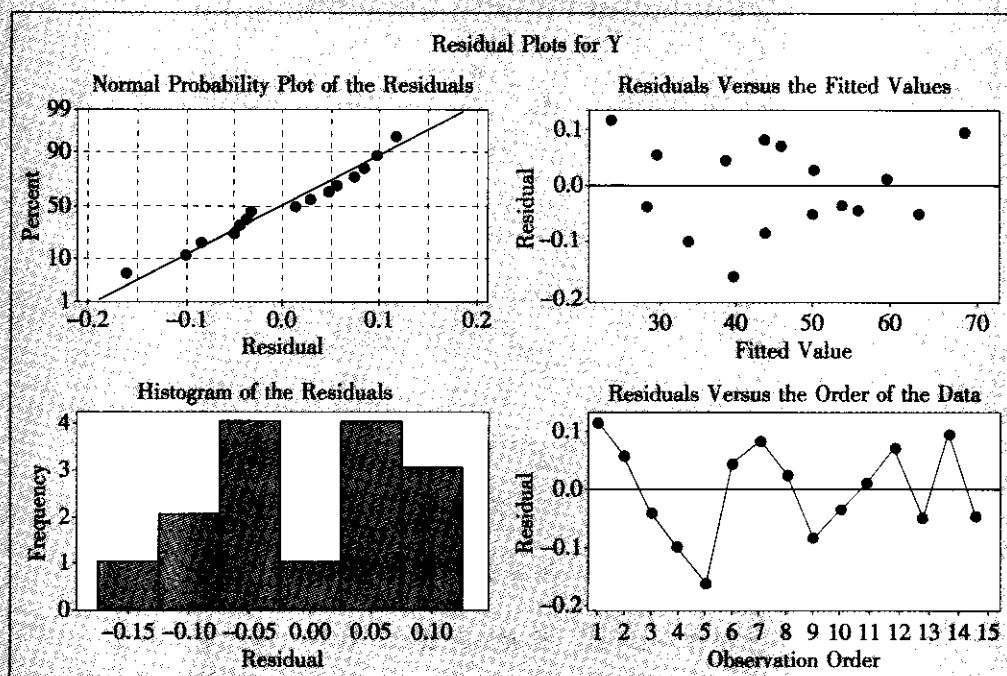


图 6—20 例 6—30 拟合二次方程的残差四合一图

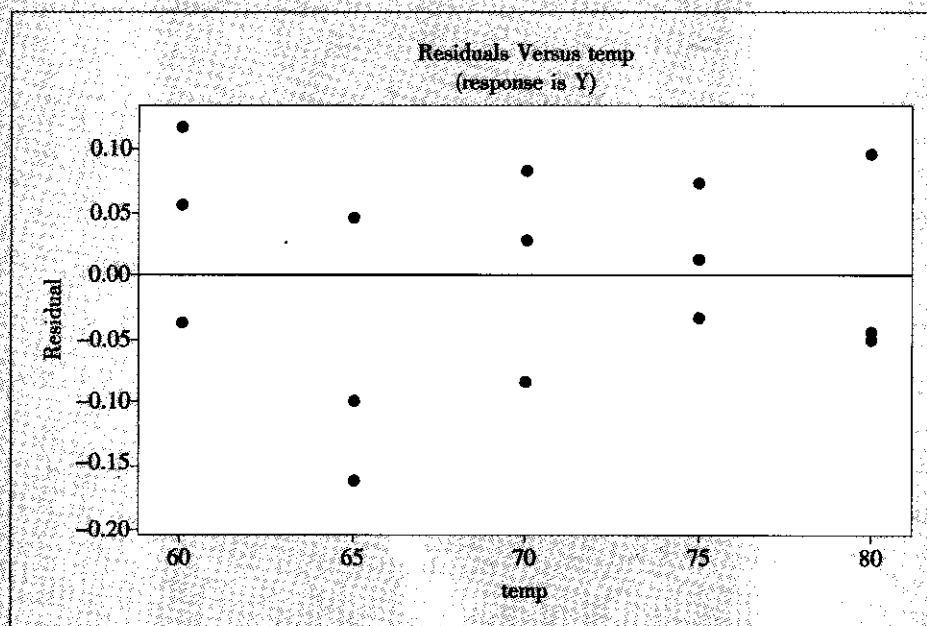


图 6—21 拟合二次方程的残差对于温度的散点图

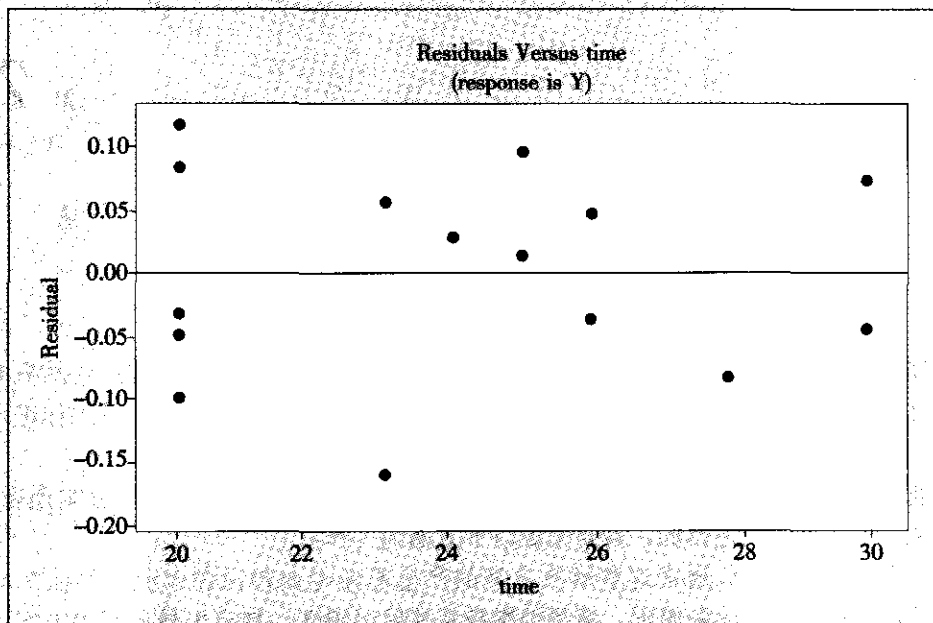


图 6—22 拟合二次方程的残差对于时间的散点图

本章小结

分析阶段的目标或目的是收集有关现状的数据,使项目焦点更加集中在了解 X 和 Y 变量之间的关系,找出变异产生的潜在根源,减少改进阶段的变量个数,确认那些造成高风险的输入变量。

分析阶段的成果体现在:通过假设检验、回归分析、方差分析等方法,对 X 变量与 Y 变量的关系进行统计分析,展示 X 与 Y 的因果关系图,找出问题产生的根本原因,对需改进的方面进行简要描述。

分析阶段需要解决的问题是收集到的数据是否可靠?如何确认潜在原因?用什么工具收集特定的数据?哪些因素是根本原因,或者是对问题影响最大的原因?有几个重要的变量影响过程?是哪些变量?已确认出哪些变量是变异的来源?如何验证这些确实是根本原因或少数关键重要变量?如何分析如此大量的数据?如何判断已经找出的确实是根本原因而不是表面现象?哪些因素需要进一步深入分析?根据什么进行这样的选择?哪些表或图能够支持你的选择?对根本原因的确认是否要调整团队的成员?是否要修正设定的最终目标?如是,该如何修改?

分析阶段在结束之前还应该明确的问题是能否找到任何可以实现快速改进的方案?过程能力可以进行什么改进?在最终解决方案建立和实施之前,针对某些缺陷,可以采取哪些过渡性问题隔离措施?改进计划是怎样的?下一步要怎样做(包括时间安排、责任落实和目标期望)?改进效果量化的根据是什么?

分析阶段使用的相关工具有:过程图等图表工具;数据和过程分析;抽样;变异源分析;假设检验;相关和回归分析;方差分析;非参数检验等统计和管理工具。

通过前面三个阶段的项目工作，团队对要解决的问题以及引起该问题产生的根本原因等有了比较准确的把握，奠定了从根本上解决这些问题的基础。至此，项目工作进入了关键性的“改进”阶段。

改进阶段的目标是形成针对根本原因的最佳解决方案，并且验证这些方案是有效的。一般来说，为了达到上述目的，需要完成以下方面的工作：

(1) 产生解决方案。通常这些解决方案的产生需要专业知识和对流程的认知与经验等。但是，也有一些方法和原则，将帮助黑带及其项目团队，制定出最佳解决方案。试验设计 DOE 是帮助团队产生最佳解决方案的强有力工具。本章的后续章节中，将着重介绍 DOE 及其应用。此外，还有一些设计工具或设计原则，可以帮助团队产生或完善改进方案，比如：SSD 设计。

SSD 是英文 smart simple design 的缩写，从字面上理解，它是一种简化设计原则与理念，也就是说要尽可能用简化的方法来实现改进方案，复杂产生浪费，复杂产生缺陷，好的改进方案应该是 SSD 的。即：

- 流程的步骤比较简化，以最少的步骤产生要求的输出；
- 尽可能地使用标准化的过程方法（表格、文件、软件）；
- 流程中的传递或转换比较少，特别是不同部门之间的传递比较少；
- 没有不增值的环节或步骤，没有周期或成本的浪费产生。

因此，在产生解决方案的时候，应当：减少各个不必要的流程步骤；优化各个步骤间的顺序；尽可能合并流程中的一些功能；对流程标准化。当然，这个过程需要不断挑战自己，挑战流程的各种限制，以达到突破的目的。

(2) 评价解决方案。改进方案产生后，需要对其优劣进行评价。一般将从改进方案的可行性、成本投入、周期等方面，评价改进方案的优势与劣势。在许多六西格玛项目实践中，团队将产生多个改进方案，并且选用因果矩阵的评价方法，从项目综合优势上，选择一个最好的方案加以验证实施。

当然，对改进方案的评价，不仅从技术方法上评价，还应该考虑到这些方法的接受程度，即它们将对使用这些方法的人员（包括操作人员、管理人员等）造成什么影响，他们是否愿意接受这样的改进等。有时，需要退而求其次。因为没有好的接受程度，是不会有好的改进效果的。有时，团队选用的方案不一定是技术上最佳的，而是最容易被大家接受的。

(3) 完成改进方案的风险评估。一个改进方案的实施是否会给顾客以及流程带来其他

问题？这是每一个黑带和项目团队需要把握的问题。一个好的改进方案，是不能够以给顾客或企业带来较高的风险为代价的。评价改进方案的风险，并且对高风险项加以有效控制，是改进阶段项目团队必须完成的一项工作。对改进方案做风险评估的方法比较多，但最常用的是使用 FMEA 分析方法。本书其他章节中已经介绍了 FMEA 方法，这里不再赘述，但需要强调的是，对改进方案进行风险分析是非常必要的。而改进方案的 FMEA 分析结果，也将是控制阶段产生控制计划的基础或者输入。

(4) 改进方案有效性验证。任何改进方案均需要验证其有效性，才能最终被确认为团队将采取的方案。而其有效性必须通过量化结果加以证明。这些量化结果包括：通过对样本的测量和统计分析得到的改进方案实施效果的数据以及作出推断的置信程度；改进后缺陷率降低的情况或者改进后过程能力的评估数据等。

这里需要特别指出的是试验设计 DOE 方法对产生或形成最佳改进方案的作用。许多项目在形成改进方案时面临的一个重要问题是：哪些自变量 X 显著地影响着 Y ，怎样去改变这些自变量 X 或者这些自变量 X 取什么值时将会使 Y 达到最佳值。这时，我们可以使用的最主要的手段和工具就是试验设计 (design of experiment, DOE)：计划安排一批试验，并严格按照计划在设定的条件下进行这些试验，获得新数据，然后对之进行分析，获得我们所需要的信息，从而获得改进的途径。试验设计就是这样一门科学，它是研究如何以最有效的方式安排试验以获得含有最大信息量的数据。

如上所述，DOE 所要达到的目的是分析出“哪些自变量 X 显著地影响着 Y ，这些自变量 X 取什么值时将会使 Y 达到最佳值”，我们前面所介绍的统计方法中也有与之类似的目的，那么 DOE 与这些方法有什么关系呢？

要注意的是，在 A 阶段，我们的分析完全依赖于未改变当前生产条件下所获得的数据；而在改进阶段，我们可能要考虑全新的工艺，也可能要打破目前生产中条件的限制，大胆考虑那些从未使用过的值，因而有可能使 Y 获得更佳值。总之，两个阶段的条件限定是不同的。我们对于过去已经获得的结论要重新考虑其适用性，对于新情况更要开创性地探索以求得其规律性，以前的结论不能再直接套用。因此，这个阶段是一个最富挑战性的阶段，也是五个阶段中最困难的阶段，要做好足够的精神准备。同时，既然下决心要做试验，期望获得比过去出色得多的好结果，那也一定要承受可能得到比过去坏得多的结果，加上任何试验的设计、实施及分析都要付出大量的财力、物力和人力，就必然要万分慎重地对待，按照团队工作条例，经过讨论并获得批准后才能付诸实施，切记不要贸然行事。

如果我们能够利用历史数据获得一些信息，我们决不应轻易放弃。通常我们可以使用多元回归分析 (multiple regression) 方法对历史数据加以分析，获得相应的回归方程，得到 Y 与各个自变量 X 间的关系式。但这种关系的获得是“被动”的，因为我们使用的是现成的数据，几乎无法控制适用范围，无法控制方程的精确度，只能处于“有什么算什么”的状况。而在改进阶段处于一种全新的状况，几乎没有可用的数据，我们面对的首要问题是如何获得更有价值的信息，这就必然要用到新工具 DOE。当然，我们在试验中获得了数据之后要进行一系列分析工作，而其中最重要的一项是判断响应变量 Y 是否受各 X 的显著影响，并进一步找出 Y 与 X 间的关系，这时我们就要使用方差分析和多元回归分析方法。因此，方差分析和多元回归分析都是 DOE 分析中所使用的工具。这些工具在前



面各章中已有介绍，在此不再重复。

7.1 试验设计基础

7.1.1 试验设计中的基本术语

一般实际问题都是复杂的，任何数学模型都只是它的某种抽象概括，试验设计也不例外。我们先给出一些有关试验设计的简单基本概念的描述，稍后将给出一些更深入的概念的定义。

1. 因子：可控因子与非可控因子

我们将过程模型简化为图 7—1，其中 Y_1, Y_2, \dots, Y_r 是我们关心的输出变量，称为“**响应变量**”（response）或“**指标**”。一般情况下，常需要考虑多方面的指标，但在本书的试验设计中，只考虑单个响应变量的情况。如果问题中确实要同时考虑多项指标，这就使问题变得非常复杂，我们常常也是设法将之转化为单指标的情形来处理。我们将影响响应变量的那些变量称为试验问题中的**因子**（factor）。假定， X_1, X_2, \dots, X_k 是人们在试验中可以加以控制的因子（称**可控因子**（controlled factor）），它们是输入变量，是影响过程最终结果的。这些变量可以是连续型的（通常是这样），也可以是离散型的。影响过程及结果的变量除了这些可控因子外还可能包含一些可以记录但不可控制的**非可控因子**（uncontrolled factor）： U_1, U_2, \dots, U_m ，它们通常包括环境状况、操作员、材料批次等等。这些变量可能取连续值，也可能只取离散值。对于这些变量，通常很难将它们控制在某个精确值上，实际问题中它们确实也可能取不同的值。我们也把这些非可控因子称为**噪声因子**（noise factor），因为常把它们当作误差来处理。

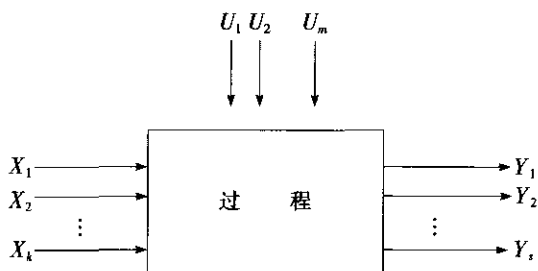


图 7—1 过程模型示意图

2. 水平及处理

为了研究因子对响应的影响，需要用到因子的两个或更多个不同的取值，这些取值称为因子的**水平**（level）或**设置**（setting）。各因子皆选定了各自的水平后，其组合被称为**处理**（treatment）。一个处理的含义是：按照设定的因子的水平组合，我们就能进行一次试验，可以获得一次响应变量的观测值，因此处理也可以称为一次“**试验**”（trial 或 experimental run），也简称一次“**运行**”（run）。

3. 试验单元与试验环境

指对象、材料或制品等载体，处理（即试验）应用其上的最小单位称之为**试验单元**（experiment unit）。例如，按因子组合规定的工艺条件所生产的一件（或一批）产品，接受治疗的一个病人等。

以已知或未知的方式影响试验结果的周围条件，称之为**试验环境**（experiment environment）。这里通常包括温度、湿度、电压等。

4. 模型与误差

考虑到影响响应变量 Y 的可控因子是 X_1, X_2, \dots, X_k ，因此，我们在试验设计中建立的数学模型是：

$$Y = f(X_1, X_2, \dots, X_k) + \epsilon$$

式中， Y 是响应变量； X_1, X_2, \dots, X_k 都是可控因子； f 是某个确定的函数关系。本式中的误差 ϵ 除了包含有非可控因子（或噪声）所造成的“**试验误差**”（experimental error）外，还可能包含“**失拟误差**”（lack of fit）。这里，失拟误差是指我们所采用的模型函数 f 与真实函数间的差异。试验误差与失拟误差这两种误差性质是不同的，分析时也要分别处理。有时为了简化，常假定函数关系 f 是准确的，从而可以忽略失拟误差。从上述概念中还可以看到，试验误差本身包含了测量误差。为了不使测量误差影响分析结果，通常要在试验进行前，先进行测量系统分析，只有测量误差满足了对测量系统的最低要求后，试验才能开始进行。有关测量系统的分析我们在第5章中已作了细致的讨论。

5. 主效应和交互效应

我们举一个最简单的因子设计来说明主效应和交互效应的概念。

在农田试验中，考虑两个因子，每个因子皆2水平。 A ：浇水。低水平：水少；高水平：水多。 B ：施肥。低水平：肥少；高水平：肥多。以产量 Y 为响应变量（单位：kg），列表如下（见表7—1）：

表 7—1 可加模型数据表

<div><div><div>A</div><div>B</div></div></div>	A	
	水少	水多
肥少	100	120
肥多	130	150

如何分析因子 A 浇水的效应呢？由于 A 处于低水平时（不考虑因子 B ），得到的产量的平均值是 $(100+130)/2=115\text{kg}$ ， A 处于高水平时（不考虑因子 B ），得到的产量的平均值是 $(120+150)/2=135\text{kg}$ ，产量由 115kg 提高到 135kg 完全是因子 A 的作用，这时，我们称因子 A 的**主效应**（main effect）为 $135-115=20\text{kg}$ 。

A 的主效应 = A 处于高水平时 Y 的平均值 - A 处于低水平时 Y 的平均值
同样可以算出：

$$\text{因子 } B \text{ 的主效应} = [(130+150)/2 - (100+120)/2] = 140 - 110 = 30\text{kg}$$

不但如此，我们还发现，当 B （肥料）处于高水平时，因子 A 的效应为 $150-130=20$ ，当 B （肥料）处于低水平时，因子 A 的效应为 $120-100=20$ ，二者完全相同。如果



以因子 A 为横轴，以响应变量（产量）为纵轴作图（见图 7—2），可以看出，两条线是平行的。

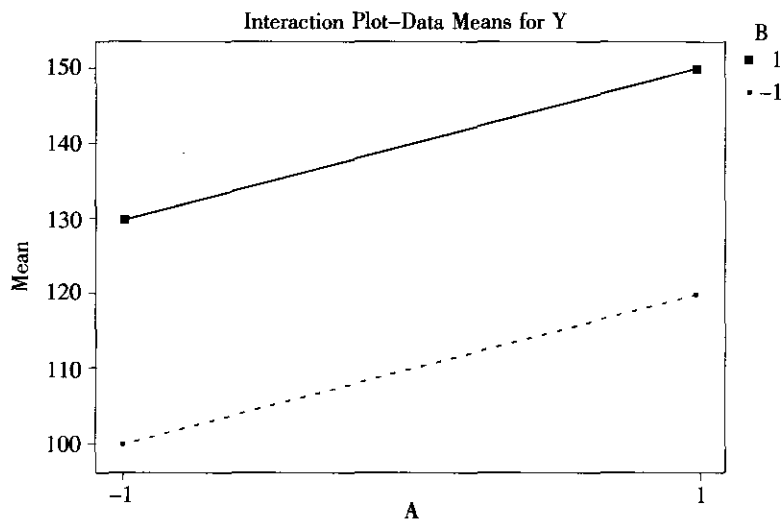


图 7—2 无交互作用时的效应图

如果数据换成另一组（见表 7—2）：

表 7—2

有交互作用的数据表

B \ A	A	
	水少	水多
肥少	100	120
肥多	130	170

由于 A 处于低水平时产量的平均值是 $(100+130)/2=115\text{kg}$ ，A 处于高水平时产量的平均值是 $(120+170)/2=145\text{kg}$ ，我们称因子 A 的主效应为 $145-115=30\text{kg}$ 。同样可以算出：

$$\text{因子 B 的主效应} = [(130+170)/2 - (100+120)/2] = 150 - 110 = 40\text{kg}$$

我们发现：当 B（肥料）处于高水平时，因子 A 的效应为 $170-130=40$ ，当 B（肥料）处于低水平时，因子 A 的效应为 $120-100=20$ ，二者大不相同。仍以因子 A 为横轴，以响应变量（产量）为纵轴作图（见图 7—3），可以看出，两条线是不平行的。

我们定义两因子间有交互作用为：如果因子 A 的效应依赖于因子 B 所处的水平时，则称 A 与 B 之间有交互作用（interaction）。

显然，在前组数据中，因子 A 与因子 B 没有交互作用。在后组数据中，因子 A 与因子 B 是有交互作用的。那么，如何来度量交互作用的大小呢？如果没有交互作用，当因子 B 处于不同水平时，因子 A 的效应是固定不变的，因此我们定义交互效应的出发点就是，当因子 B 处于不同水平时，因子 A 的效应到底差了多少？因此，我们定义交互效应（interaction effect）为下式：

$$AB \text{ 交互效应} = (B \text{ 处于高水平时 A 的效应} - B \text{ 处于低水平时 A 的效应})/2$$

在后一组数据中， $AB \text{ 交互效应} = [(170-130) - (120-100)]/2 = 10$ 。同样，交

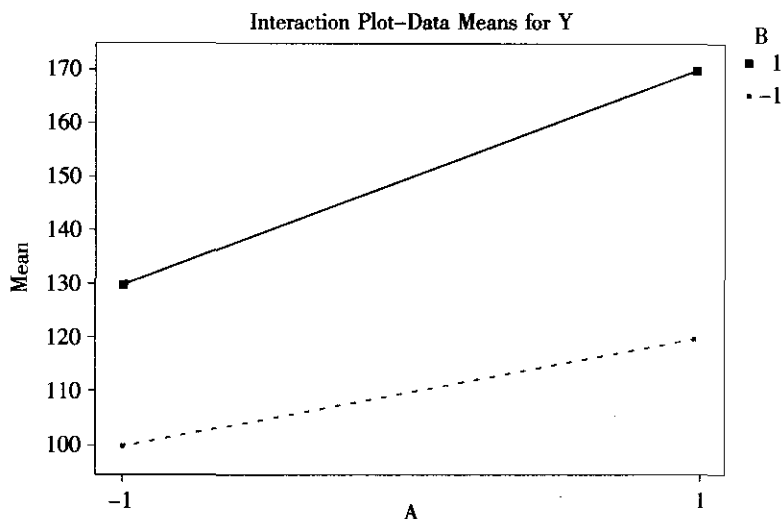


图 7—3 有交互作用时的效应图

换 A 与 B 的顺序，可以得到公式：

$$BA \text{ 交互效应} = (A \text{ 处于高水平时 } B \text{ 的效应} - A \text{ 处于低水平时 } B \text{ 的效应})/2$$

$$BA \text{ 交互效应} = [(170 - 120) - (130 - 100)]/2 = 10$$

显然二者是相同的。以后我们不再区分 $A \times B$ 或 $B \times A$ 。另外，我们还可以得到变形的公式：

$$AB \text{ 交互效应} = [(A \text{ 高 } B \text{ 高} + A \text{ 低 } B \text{ 低}) - (A \text{ 高 } B \text{ 低} + A \text{ 低 } B \text{ 高})]/2$$

用本例数据代入，即为：

$$AB \text{ 交互效应} = [(170 + 100) - (130 + 120)]/2 = 10$$

此公式可以帮助我们理解在后面讲到的从正交表中直接分析交互效应的计算公式。

我们注意的是，如果两个因子间存在显著的交互作用，就不能只用主效应大小作为该因子是否重要的判断依据。有时一个因子主效应很小，只要某个含它的交互作用效应显著，这个因子就是重要的，就应予以保留。

6. 试验设计示例简介

为了说明试验设计有关概念，我们先研究一个非常简单的例子。

合成氨纯度试验

在提高合成氨纯度(%)的工艺研究中，发现有三个因子很重要，它们是因子 A——温度；因子 B——压力；因子 C——反应时间。对每个因子都设定了高低两个水平。我们希望考察这三个因子中，哪些因子效应及交互效应是显著的。其具体取值如下：

A 因子——温度，低水平：460 (度)；高水平：500 (度)

B 因子——压力，低水平：250 (大气压)；高水平：270 (大气压)

C 因子——时间，低水平：20 (分钟)；高水平：30 (分钟)

如何安排试验呢？三个因子按 2 水平全面搭配应该有 8 种组合。先看下面这张正交表 (见表 7—3)：

表 7—3 三因子正交试验代码表

	A	B	C
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1

我们将按照此表来安排试验,具体是这样进行的:凡表中列为-1者,试验时取低水平;凡表中列为1者,试验时取高水平,这样就可以进行8次试验。

我们之所以称这张表为正交表(orthogonal array),是因为它有两个特点:

(1) 任一列中正负号出现次数各占一半,即在试验中,每个因子取低水平、高水平次数相同。

(2) 任两列中,++,-+,+,-,-四种搭配出现的次数都是相等的。“搭配出现次数都相等”这种性质,用数学上的方法表示,就是两列乘积和为0,即代数上说的两列“正交”。这种“正交性”将导致对于试验结果的分析有“均衡分散,整齐可比”的特点,因而具有很多优良性质,而且很容易计算出相应的回归方程。这种试验设计方法常称为“正交试验设计法”(orthogonal experimental design)。

在本例中,三个因子的水平都被取定后,我们就可以用它们的组合条件来进行试验了。例如,第1号试验条件是三个因子皆取低水平,温度取460度,压力取250大气压,时间取20分钟,生产一反应罐合成氨,从中取一个样品,化验出它的纯度 Y_1 (%)。依此类推,总计可以得到8个纯度: Y_1, Y_2, \dots, Y_8 。这里的试验单元就是化学反应中使用的“罐”,我们要分析的是三个因子的主效应和交互效应。

把正交试验的代码表列在表中的前3列,分别记为Ac, Bc和Cc,把各因子所取的真实值列在紧接着的三列中,将8次试验结果获得的纯度 Y (%)记录在最后一列,数据如下表(见表7—4)。

表 7—4 三因子正交试验结果表

	Ac	Bc	Cc	A	B	C	Y
1	-1	-1	-1	460	250	20	94.53
2	1	-1	-1	500	250	20	97.17
3	-1	1	-1	460	270	20	93.84
4	1	1	-1	500	270	20	95.85
5	-1	-1	1	460	250	30	94.59
6	1	-1	1	500	250	30	97.50
7	-1	1	1	460	270	30	94.17
8	1	1	1	500	270	30	96.54
M-	377.13	383.79	381.39				764.19
M+	387.06	380.40	382.80				95.52
Effect	2.482 5	-0.847 5	0.352 5				

从此表中,很容易得到各因子的主效应。首先讨论A因子,将Ac这列中对应“-1”的各次试验(即第1,3,5,7各次)响应变量Y的总和计算出来,填入M-行(本例, $M- = 94.53 + 93.84 + 94.59 + 94.17 = 377.13$)。同样将Ac这列中对应“+1”的各次试验(即第2,4,6,8各次)响应变量Y的总和计算出来,填入M+行(本例, $M+ = 97.17 + 95.85 + 97.50 + 96.54 = 387.06$)。因此,可以得到,在A处于高水平时的平均值为 $387.06/4 = 96.765$,在A处于低水平时的平均值为 $377.13/4 = 94.2825$,因而得知,A因子的主效应是 $\text{Effect}(A) = 96.765 - 94.2825 = 2.4825$ 。同样方法可得, $\text{Effect}(B) = -0.8475$, $\text{Effect}(C) = 0.3525$ 。

具体的分析方法,我们将在第2节更详细地给出,这里不再讨论了。

7.1.2 试验设计的基本原则

有三个基本原则在试验设计中必须要考虑:重复试验(replication)、随机化(randomization)和区组化(blocking)。

所谓重复试验是指一个处理施于多个试验单元。这些单元是我们在统计推断中一个处理所形成的总体的代表,它可以使我们得以估计试验误差的大小。通常的显著性检验都是拿不同处理间形成的差别与随机误差相比较,只有当处理间这种差别比随机误差显著地大时,我们才说“处理间的差别是显著的”。没有随机误差就无法进行任何统计推断,因此在试验设计中安排重复试验是必不可少的。需要注意的是:一定要进行不同单元的重复(replicate),而不能仅进行同单元的重复(repetition)。换言之,一定要重新做试验即重复试验,而不能仅是重复观测或重复取样。比如在上例中,一定要用同样的工艺条件生产两罐或多罐合成氨,而不能只是从同罐合成氨中分次取不同的样品来测试纯度。显然,同罐合成氨中分次取不同的样品所测试出的纯度间差异要小,而不同罐合成氨中取不同的样品来测试的纯度间差异要大。以同单元重复得到的差异来估计随机误差将会低估试验误差,所得的结论就都是不可信的。在试验中一定要包含有真正的重复。

随机化是第二个原则。随机化的含义是以完全随机的方式安排各次试验的顺序和/或所用试验单元。这样做的目的是防止那些试验者未知的但可能会对响应变量产生的某种系统的影响。在例7-1中,假使我们在同一天内进行的8次试验之顺序就按照表上所列顺序进行的话,会有什么问题呢?如果当天的电压有一种由高向低变化的趋势,而恰好电压的降低将导致纯度的降低,那么很明显,前4次试验是在电压较高的情况下进行的,后4次试验是在电压较低的情况进行的。如果后4次纯度明显低于前4次,而前4次试验恰好是因子C(反应时间)处于低水平(20分钟)时进行的,后4次试验恰好是因子C(反应时间)处于高水平(30分钟)时进行的,那么后4次平均值变低,究竟是因子C的作用还是电压的作用,是无论如何也分辨不清的。如果将这8次试验顺序完全打乱,则不会再出现上述问题了。随机化并没有减少试验误差本身,但随机化可以防止出现未知的但可能会对响应变量产生的某种系统影响。

区组化是第三个原则。各试验单元间难免会有某些差异,如果我们能按某种方式把它



们分成组，而每组内可以保证差异较小，即它们具有**同质齐性**（homogeneous），则我们可以在很大程度上消除由于较大试验误差所带来的分析上的不利影响。一组同质齐性的试验单元称为一个**区组**（block），将全部试验单元划分为若干区组的方法称之为**分区组或区组化**。通过在同一个区组内比较处理间的差异，就可以使区组效应在各处理效应的比较中得以消除，从而使对整个试验的分析更为有效。例如，假定在上（下）午时段内差异不大，而上午、下午差异可能较大，那我们就把上午、下午当作两个区组。这时在分析中就可以去除掉上午、下午间的差异的影响，或尽可能把试验全都安排在上午（或下午）进行。如果分区组有效，则这种方法在分析时，可以将区组与区组间的差异分离出来，这样就能大大减少可能存在的未知变量的系统影响，这就是分区组的好处。当然，在区组内还应该用随机化的方法进行试验顺序及试验单元分配的安排。什么时候用分区组，什么时候用随机化呢？在试验的设计中应遵照下列原则：“**能分区组者则分区组，不能分区组者则随机化**”（Block what you can and randomize what you cannot）。

7.1.3 试验设计的必要性

人类在认识自然界的过程中，进行着多方面探索。试验是构成学习过程（learning process）的一个要素。通过试验的学习是一种综合了人们的期望、需要、知识和资源的复杂过程。在试验的统计设计方法开始形成的 20 世纪 20 年代以前，科技工作者在试验中走了不少弯路。

发明家爱迪生（Thomas Edison）一生艰苦奋斗，经历了无数次的失败之后，为人类发明了许多重要科技产品。他的座右铭是“天才靠的是百分之一的灵感和百分之九十九的汗水”。他的助手特勒萨（Nicola Telsa）在 1931 年 10 月 19 日（即爱迪生去世后第二天）的《纽约时报》发表纪念评论写道：“如果在爱迪生工作的黑屋中能有一支蜡烛照亮他前进的方向的话，以他蜜蜂般的勤奋，他将会获得远比 he 已发明的东西多得多的成果”，“我非常同情地观察到他的工作状况，但凡有一点点理论和计算能帮助他的话，将节省他百分之九十的精力”。爱迪生是靠苦干拼出来的，他是在边试验边分析后确定下次试验该怎么做的。这种方法速度太慢，而且只能从已经得到的偶然出现的好结果出发，摸索着继续前进，无法形成理论上的知识，无法预测何处将有更好的成果。这种凭直观猜测逐步探索的方法已经无法适应当代快速发展的需求了。

另一种在多因子分析试验中常用的方法是“**单因子变化法**”或称“**一次一因子法**”（one-factor-at-one-time）。其做法是：多个因子中，每个因子依次改变，而其他因子则保持在固定的或选定的水平上。这样做所需要的试验次数肯定能比全面搭配时少，但它有许多缺点：与统计试验设计方法相比，要达到同样的效应估计的精度，需要更多的试验次数；不能估计某些交互效应；不能在试验区域内进行系统而全面的搜索，因而它的分析结论缺乏普遍性；它最后的结论对于固定因子的初始值的选定太敏感，因而可能错过最优的因子设置。这一方法现已弃之不用了。

我们将**全部因子全部水平的全部搭配都进行至少一次试验**的安排方法称为**全因子试验法**（full factorial design）。这是人们容易想到的一个方法，而且可以获得相当多的信息。但是，是否我们永远可以使用全因子试验法呢？答案是否定的，因为这样做的试验次数太多，人们无法接受。如果有 8 个因子，每个因子只取 2 水平，那么全因子试验要

$2^8=256$ 次; 每个因子取 3 水平, 那么全因子试验要 $3^8=6\,561$ 次。这在实际工作中是办不到的。我们只能从中选择一部分来进行。那么怎样来选择条件呢? 我们要寻求好的设计, 就是要用最少的试验次数, 获得尽可能多的信息, 这就需要运用统计分析工具。如果不使用统计分析工具, 即使是用全因子试验法, 在拿到试验结果后, 也只是从中挑选那个最优值就结束, 那么我们就浪费了大量的有用信息。试验设计的统计分析方法不但能从试验结果中找到最优值, 而且可以判明哪些因子影响显著, 哪些因子影响不显著, 还可以得到有关的变化规律, 预测将要达到的最佳值是多少和这个最佳值将在什么范围内波动, 而这个最佳值可能根本未在我们选定的条件中试验过。这些就是统计分析方法的威力。

7.1.4 试验设计的类型

根据不同的研究内容, 可以对试验设计进行多种方法的分类。

根据试验的因子个数, 可以分为单因子和多因子。根据试验的目的, 可以分为因子设计和回归设计。在不考虑区组的设计中, 常用的有完全随机化设计 (completely randomized design); 在考虑区组的设计中, 常用的有: 配对比较设计 (paired comparison design)、随机区组设计 (randomized block design)、平衡不完全区组设计 (balanced incomplete block design)、部分平衡不完全区组设计 (partial balanced incomplete block design) 等。因子效应可以分为固定效应和随机效应两大类。在固定效应中, 又可以分为单向分类 (one-way layout)、双向分类 (two-way layout)、多向分类 (multi-way layout)。在随机效应中, 主要是应用嵌套设计 (nested design) 或称方差分量模型 (variance component modeling)。这些试验设计用到的理论和方法都比较复杂, 工程师们用得较少。除了单因子的单向分类试验设计将在第 2 节介绍外, 别的内容本书就不再介绍了。

根据试验目的不同可分为两大类: 因子设计和回归设计。本章开始的叙述中已经说明, 我们进行试验有两个基本目的: 一是明确哪些自变量 X 显著地影响 Y ; 二是找出 Y 与 X 间的关系式, 从而进一步找出自变量 X 取什么值时会使 Y 达到最佳值。第一种试验的目的是为了确定在相当多的自变量中, 哪些自变量 X 并不显著地影响 Y , 应予以删除; 哪些自变量 X 显著地影响 Y , 应予以保留; 我们称其为“因子筛选设计” (screening design)。由于这种试验的目的是针对因子的, 因此这种试验设计属于“因子设计” (factorial design) 或称“析因设计”或“因析设计”。第二种试验目的是为了确定 Y 与 X 间的关系式, 找出 Y 对于 X 的回归方程。由于这种试验的目的是针对回归关系的, 因此这种试验设计被称为“回归设计” (regression design)。

当然, 这两类设计也有相通之处: 一方面, 筛选因子的方法其实也是先建立一个 Y 与 X 间的简单的线性回归方程, 然后根据各项系数的显著性来筛选 (详细叙述参见 6.6 节)。要注意的是, 在试验设计中所说的“线性”, 已经与通常数学概念中的“线性”有所不同。在“试验设计”中说的“线性”指的是在回归方程中除了可以包含各自变量的一次项外, 还允许包含两个或多个自变量的乘积项, 例如可以含有 x_1x_2 , $x_1x_2x_3$ 等, 而通常的数学概念中的“线性”是不允许包含这些项的。在建立了线性回归方程后, 除可以判断变量是否显著外, 对于求最大值或最小值的问题也可以求出最佳值, 以及达到此最佳值的自变量的最佳设置, 这在实际工作中也常常是有用的。总之, 筛选变量也是



通过建立回归方程来实现的。另一方面,我们建立了回归方程,特别是建立了含平方项的响应曲面方程后,也可以在方程中判断是否有效应不显著的因子,可以删除它们,达到筛选因子的目的。因此,因子设计和回归设计间确有相通之处:它们都要建立回归方程。但因子设计只要线性的,而这里的回归设计指的是二阶的。总的说来,筛选的要求是较粗糙的,试验次数较少;建立回归曲面方程要求就细致多了,试验次数要大增。在因子设计中,又可以按因子水平的个数分为二水平因子设计、三水平因子设计和混合水平因子设计几类。实践证明:在因子设计中,使用二水平正交试验法,再加若干中心点的设计方法最简单有效,因此本书只介绍二水平的试验设计。再细分,又有**全因子试验设计**(full factorial design)和**部分因子设计**(fractional factorial design)两大类,我们将在本章的7.3和7.4两节中分别加以介绍。对于回归设计,也以建立二次回归方程为主要工具,介绍**响应曲面方法**(response surface methodology, RSM),这部分内容将在7.5节详细介绍。

另一类很重要的试验目的是寻求系统的**稳健性**(robust)。所谓稳健性,是指系统的抗干扰能力要强,即当系统受到难以控制的因子(或称为“噪声”)的严重影响时,系统输出的**变异性**(variation)要足够的小。为做到这一点,我们尽量选择那些使得系统对噪声变化不敏感的控制因子的某种水平的组合来达到目的,这就是**稳健参数设计**(robust parameter design)。在国内,这类设计通常被称为“田口参数设计方法”。这类问题在六西格玛改进工作中也是有重要意义的,我们将在7.6节中介绍。

如果讨论的是配方问题,例如在橡胶、造纸、药品生产等行业中,我们研究的是在整个产品中各个分量所占的比例问题,显然,这些比例的总和应该为100%。研究这类问题的试验设计被称为“**混料设计**”(mixture design),我们将在7.7中简要介绍。

有时,现有的生产条件已基本上满足要求,但是我们希望获得更好的结果,这时可以在原有生产条件的基础上稍加调整来寻求解决,这就是“**调优运算**”(evolutionary operation, EVOP),这部分内容将在7.8节简单介绍。

7.1.5 试验设计的策划与安排

人类认识自然界的过程是个循序渐进的过程。一般先根据已有的知识提出某种设想,设计一个试验去验证或否定它,从试验中获得的数据帮助人们验证或修正初始的设想,然后又提出了一个更新更深入的设想,再设计新的试验,这个反复的过程会一直持续下去,直到形成了较为完整的理论结果为止。我们进行试验也是一个学习过程,不可能一蹴而就。我们不要企图“毕其功于一役”,进行一次试验就结束战斗。一般来说,试验要进行好几批,一般采用下面几个步骤。

1. 用部分因子设计进行因子的筛选

最开始,情况不很清楚,考虑到影响响应变量的因子个数可能较多(大于或等于5),这时应在较大的试验范围内,先进行因子的筛选(screening),通常应使用部分实施的因子试验设计法,这样获得的结果可能较为粗糙,但试验次数可以大大减少,也能够达到筛选的目的。如果认为部分实施的因子试验费用仍然太昂贵,则可以使用试验次数更少的“Plackett-Burman设计”方法(详细内容见7.4.4节)来筛选因子。

2. 用全因子试验设计法对因子效应和交互作用进行全面的分析

当因子的个数被筛选到小于等于5个之后,我们可以进一步在稍小范围内进行全因子试验设计以获得全部因子效应和交互作用的准确信息,并进一步筛选因子直到因子个数不超过3个。

3. 用响应曲面方法(RSM)确定回归关系并求出最优设置

当因子个数不超过3个时,我们就有条件采用更细致的响应曲面设计分析方法,在包含最优点的一个较小区域内,对响应变量拟合一个二次方程,从而可以得到试验区域内的最优点。

以上所说的当然是典型的步骤,在实际工作中,可能跳过某个环节,也可能在某个步骤上反复进行几次。总之,要不断地筛选因子,不断调整试验的范围和因子水平的选择,经过几轮试验后才能最终达到我们试验的总目标。

7.1.6 试验设计的基本步骤

粗略地说,试验的计划与实施应该包含计划、实施及分析三阶段。

1. 计划阶段

这里又可以分为下面几个步骤:

(1) 阐述目标。所有团队成员都要投入讨论,明确目标及要求。究竟是为了筛选因子还是为了找寻关系式?最终要达到什么要求?

(2) 选择响应变量。在一个试验中若有多种响应,则要选择起关键作用的。能用连续型指标做度量的响应变量远比只有二元响应(成、败)的响应变量好得多。

(3) 选择因子及水平。用流程图及因果图或鱼骨图先列出所有可能对响应变量有影响的因子清单,然后根据数据和各方面的知识进行细致分析作初步筛选。不能确定该删除者就应该保留。对于水平的选择也要仔细处理,一般来说,各水平的设置应足够分散,这样效应才能检测出来,但也不要太分散以至将各种其他的物理机械因素都包括进来,这会使统计建模和预测变得困难。

(4) 选择试验计划。根据试验目的,选择正确的试验类型,确定区组状况、试验次数,并按随机化原则安排好试验顺序及试验单元的分配,排好计划矩阵(planning matrix)。

2. 实施阶段

严格按计划矩阵的安排进行试验,除了记录响应变量的数据外,还要详细记录试验过程的所有状况,包括环境(气温、室温、湿度、电压等)、材料、操作员等。试验中的任何非正常数据也应予以记录,以便后来分析使用。

3. 分析阶段

对数据的分析方法应与所应用的设计类型相适应。分析中应包括拟合选定模型、残差诊断、评估模型的适用性并设法改进模型等。当模型最终选定后,要对此模型所给出的结果作必要的分析、解释及推断,从而提出重要因子的最佳设置及响应变量的预测。当认定结果已经基本达到目标后,给出验证试验(confirmation run)的预测值,并做验证试验以验证最佳设置是否真的有效。关于分析阶段的具体方法和所使用的统计工具请参见以后各节的详细叙述。

7.2 单因子试验设计与分析

单因子试验在实际工作中经常会遇到，它比较简单，所用到的方法和理论在多因子试验中也会遇到。

单因子试验通常有两个目的：一是想比较一下因子的几个不同设置间是否有显著差异，如果有显著差异，哪个或哪些设置较好；二是建立响应变量与自变量间的回归关系（通常是线性、二次或三次多项式），判断我们建立的回归关系是否有意义。下面分别介绍这两种要求。

7.2.1 单向分类设计

我们先举一个实例。

比较四个车工 A, B, C, D 所加工的丝杠的直径是否相等。将 20 根坯料随机编好顺序号，再采用随机抽取方法，让每个车工选取 5 根，按顺序号在同台车床上加工。获得单向分类设计直径数据表（见表 7—5）。

表 7—5 单向分类设计直径数据表 单位: mm

A	B	C	D
50.0	49.8	50.7	50.6
50.4	50.2	50.5	51.0
49.8	49.9	50.9	50.8
50.3	50.0	50.8	51.2
50.7	50.4	50.3	50.5

本试验是单因子试验，只有“车工”这个因子，它取 4 个水平（即 4 个车工）。这就是单向分类设计（one-way layout），每个车工取 5 个试验值，这就是“重复”。在安排试验时，采用了随机化方法，将试验单元（每根坯料加工成丝杠是一个单元）及顺序号两方面皆采用了随机化处理。所选取的 4 个车工是固定的，要比较的是他们各自加工的丝杠直径，这是固定效应的问题，可以直接采用单因子方差分析法（one way ANOVA）。使用 MINITAB 软件可以先将数据经 Stack 合并成一列，然后从 STAT-ANOVA-One Way 入口，得到下列结果：

One-way ANOVA: A, B, C, D					
Analysis of Variance					
Source	DF	SS	MS	F	P
Factor	3	1.844 0	0.614 7	7.66	0.002
Error	16	1.284 0	0.080 3		
Total	19	3.128 0			

这里要检验的假设是：

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D \leftrightarrow H_1: \text{至少一对 } \mu_i \neq \mu_j$$

由 AVOVA 结果可以看出，组内差（随机误差）离差平方和为 1.284，自由度为 16；组间差（操作工间差异）离差平方和为 1.844，自由度为 3。算得 F 比为 7.66，其 p -value 为 0.002，显然应拒绝原假设，即应认为操作工间有显著差异。

计算中还输出示意图（见图 7—4）。

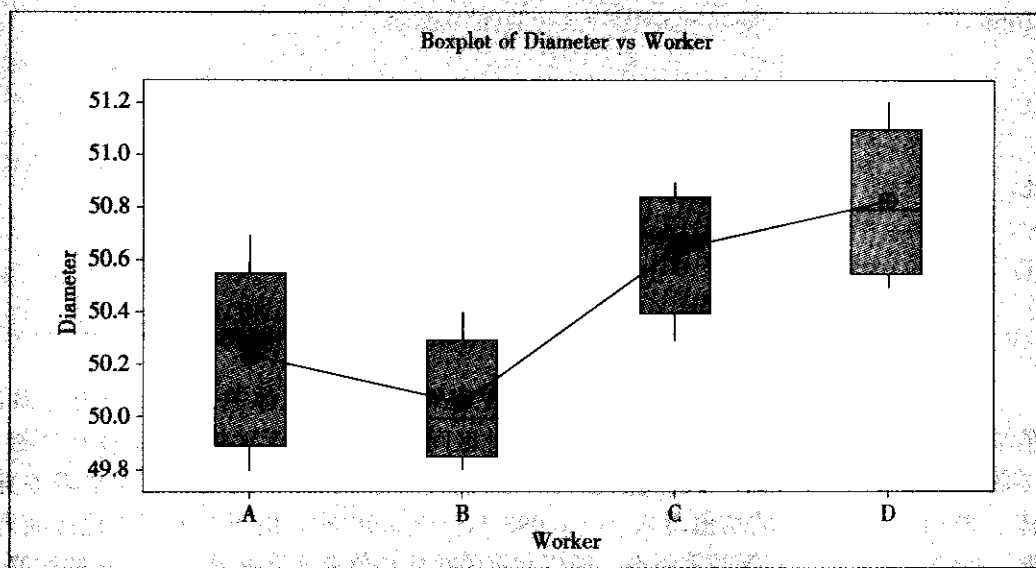


图 7—4 车工加工丝杠直径箱线图

从图中可以看出，4 个车工加工出的丝杠直径间有显著差异。

本例的分析只是方差分析法的一次运用，但在安排试验时注意到了试验设计的随机化及重复两条原则。实际上，如果我们不能假定所有 20 根坯料性能基本相同，那就还必须安排划分区组，这里就不再讨论了。

在断言各组间有显著差异后，一个很自然的问题就是要回答：究竟哪些组间有显著差异，而哪些组间没有显著差异？这就要用到统计方法：多重比较（multiple comparison）。

由于要比较的总体共有 k 个（超过 2 个），使用比较两总体均值的 t 检验在这里是不适用的。这是因为，在零假设成立的条件下，如果设定误判任意两总体均值不等的犯第 I 类错误的概率为 α_0 ，则最后总的犯第 I 类错误的概率会变大。因为，这时要两两比较的总体个数为 $s = k(k-1)/2$ 个，最后总的犯第 I 类错误的概率为 $\alpha_Z = 1 - (1 - \alpha_0)^s$ 。当 k 较大时， s 也较大， α_Z 会变得很大。例如， $k=4$ ， $s=6$ ， $\alpha_0=0.05$ ， $\alpha_Z = 1 - (1 - \alpha_0)^s = 1 - (1 - 0.05)^6 = 0.265$ ，这当然是太大了。为此，有一种较保守的方法（Bonferroni），它可以保证最后 α_Z 不超过原来设定的 α ，只要取 $\alpha_0 = \alpha/s$ 就行了。更精确的方法是 Tukey 法，它根据最后的总的犯第 I 类错误的概率，再根据参加比较的总体的个数，反过来算出比较两总体时容许犯第 I 类错误的概率 α_0 ，再由此算出任意两总体均值差的置信区间。我们可以根据此区间是否包含 0，判断出两总体间是否有显著差异。

在 MINITAB 软件中, 在选定 ANOVA 后, 点击打开 “Comparisons” 窗口, 选中 “Tukeys, family error rate” 缺省使用 5%, 则可得到下列结果:

Tukey's pairwise comparisons			
Family error rate=0.050 0			
Individual error rate=0.011 3			
Critical value=4.05			
Intervals for (column level mean)-(row level mean)			
	A	B	C
B	-0.333 1		
	0.693 1		
C	-0.913 1	-1.093 1	
	0.113 1	-0.066 9	
D	-1.093 1	-1.273 1	-0.693 1
	-0.066 9	-0.246 9	0.333 1

上述计算结果所列之数值是两总体之差的置信限。例如, 总体 A 的均值—总体 B 的均值落入 $(-0.333\ 1, 0.693\ 1)$, 由于 0 落入此区间内, 因此, 无法拒绝两总体均值相等的零假设, 即应判定两总体均值无显著差异。同样, AC 间、CD 间均值也无显著差异。总体 A 的均值—总体 D 的均值落入 $(-1.093\ 1, -0.066\ 9)$, 由于 0 未落入此区间内, 因此, 拒绝两总体均值相等的零假设, 即应判定两总体均值有显著差异。总体 A 的均值—总体 D 的均值总是负数, 即总体 A 的均值小于总体 D 的均值, 简记为 $A < D$ 。同样, $B < C$, $B < D$ 。总之, CD 间无显著差异, B 比它们显著地小, A 比 D 显著地小, 而与 B, C 间均值无显著差异。

7.2.2 多项式回归

通常单因子试验的第二个目的是希望建立响应变量与因子(自变量)间的回归关系。我们先看一个例子(见例 7—3)。

例 7—3

锻压温度显著影响锻件最终断裂强度。采用 800, 850 和 900 (摄氏度), 各锻造 3 根钢轴, 将钢轴的断裂强度(单位: 吨)记录下来, 得到下列数据(见表 7—6)。

表 7—6 锻压温度与断裂强度数据表

温度/摄氏度	800	850	900
1	35.8	39.4	40.1
2	36.2	39.6	39.6
3	37.1	38.5	38.9

先进行 ANOVA, 可以看出锻压温度确实显著地影响着锻件最终断裂强度。

Analysis of Variance for strength

Source	DF	SS	MS	F	P
temp	2	18.002	9.001	23.48	0.001
Error	6	2.300	0.383		
Total	8	20.302			

由于 p -value 为 0.001, 拒绝均值相等的原假设, 即判不同温度的各总体断裂强度显著不同 (见图 7—5)。

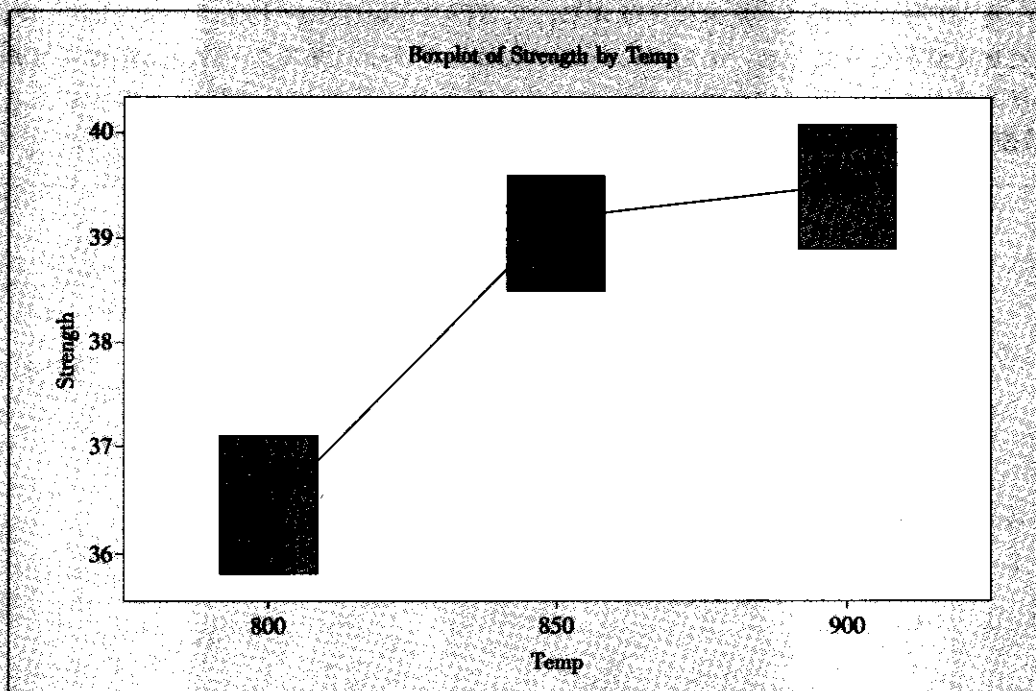


图 7—5 锻压温度与断裂强度间的关系图

那么, 断裂强度与锻压温度间具有什么关系呢? 我们希望建立断裂强度与锻压温度间的回归方程。将数据重新整理为下列形式 (见表 7—7):

表 7—7 锻压温度与断裂强度数据表 (计算机数据格式)

温 度	强 度
800	35.8
800	36.2
800	37.1
850	39.4

续前表

温 度	强 度
850	39.6
850	38.5
900	40.1
900	39.6
900	38.9

用 MINITAB 软件, 先拟合线性方程: 从 STAT-Regression-Fitted Line Plot 入口, 设定自变量和响应变量后, 选定 “Linear” (线性) 得下图 (见图 7—6)。粗略看即可知数据有弯曲趋势, 因此拟合线性方程回归效果不够好。判断拟合好坏的准确方法要用 “残差诊断” (residual diagnostic)。

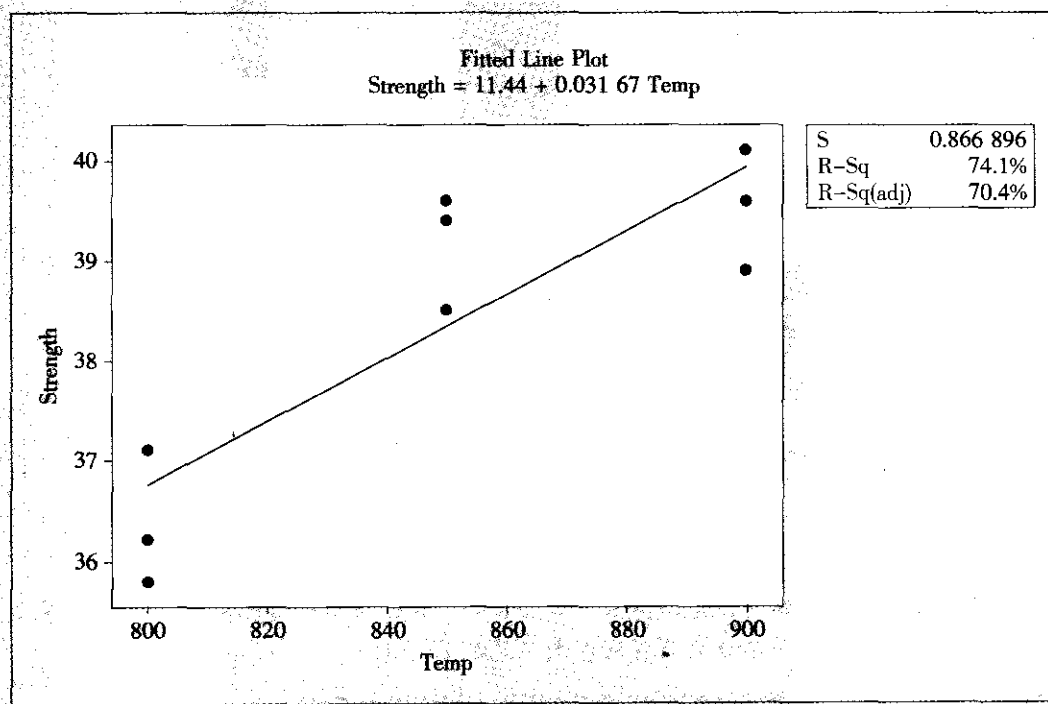


图 7—6 锻压温度与断裂强度间的线性回归示意图

由于自变量取值已达到 3 个不同的值, 因此我们可以拟合二次函数, 同样从 STAT-Regression-Fitted Line-Plot 入口, 选定 “Quadratic” (二次函数), 可得下列结果 (见图 7—7):

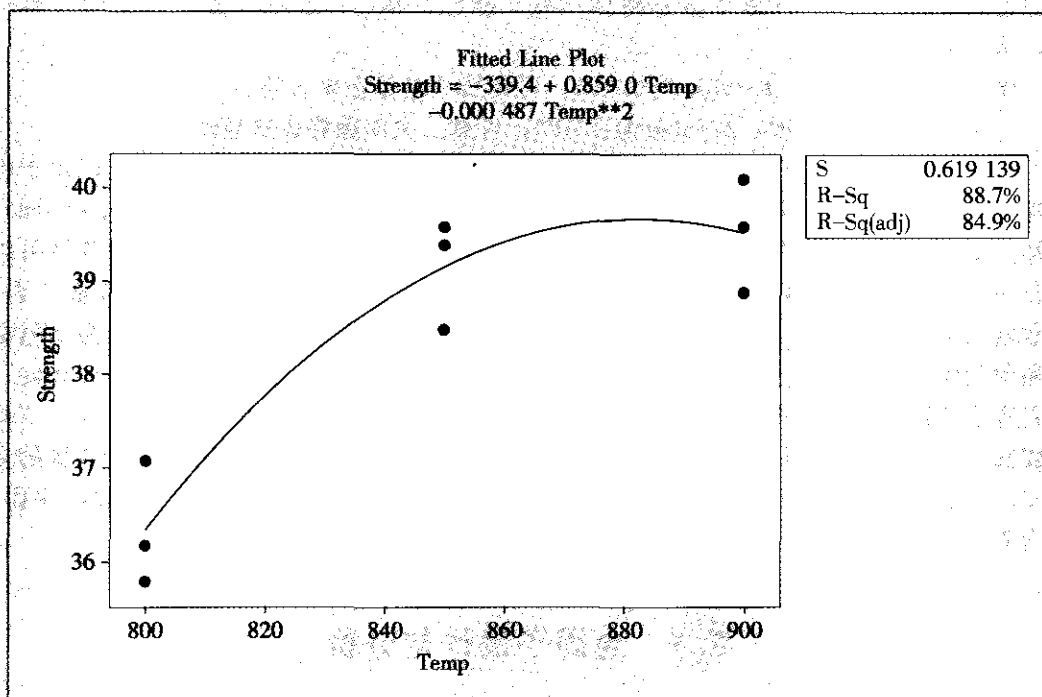


图 7—7 锻压温度与断裂强度间的二次回归示意图

Polynomial Regression Analysis: strength versus temp

The regression equation is

$$\text{strength} = -339.367 + 0.859 \text{ temp} - 0.0004867 \text{ temp}^2$$

S=0.619 139 R-Sq=88.7% R-Sq (adj) =84.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	18.002 2	9.001 11	23.481 2	0.001
Error	6	2.300 0	0.383 33		
Total	8	20.302 2			

Source	DF	Seq SS	F	P
Linear	1	15.041 7	20.015 3	0.003
Quadratic	1	2.960 6	7.723 2	0.032

先看 ANOVA 表中的总效果。建立假设检验：

 H_0 : 模型无效 $\leftrightarrow H_1$: 模型有效

如果对应回归项 (regression) 的 p -value < 0.05 , 则表明应拒绝原假设, 即可以判定本模型总的说来是有效的; 如果对应回归项的 p -value > 0.05 , 则表明无法拒绝原假设, 即可以判定本模型总的说来是无效的。本例中, regression 项的 p -value 为 0.001, 可见



模型是显著的。ANOVA 表下方是将回归按“线性”和“二次项”分开列出的计算结果，从中可以看到：

线性（Linear）项， p -value 为 0.003，可见线性趋势是显著的。

二次（Quadratic）项， p -value 为 0.032，可见二次项趋势也是显著的。

上述方法还可以推广到更高阶的情形。为了让回归方程中的各项保持独立以便于检查各项效应的显著性，最好选用“正交多项式回归”（orthogonal polynomial regression）。另外，从拟合的多项式的阶数上来说，一个因子取了 k 个水平，对于所获得的数据可以拟合一个 $k-1$ 阶多项式。但实际上，4 次以上的多项式一般是不使用的。表面上看，次数增高可以使“拟合效果更好”，但这样的拟合模型缺乏好的预测能力，因为这时估计和预测值的方差都变大了。在回归分析中，称这类现象为“超拟合”（overfitting）。我们应尽量避免这种情形的出现。如果低次多项式与数据确实拟合不好，可以采用“样条回归”（spline regression）或分段样条回归的方法，先在小区间上拟合最高为 3 次的多项式，然后将这些多项式拼接修补在一起，形成一个在整个区域上光滑的函数或曲面。这需要专门的工具，本书不作专门介绍。

7.3 全因子设计与分析

全因子试验设计是指所有因子的所有水平的所有组合都至少要进行一次试验。由于包含了所有的组合，全因子试验所需试验的次数会较多，但它的优点是可以估计出所有的主效应和所有的各阶交互效应。所以在因子个数不太多，而且确实需要考察较多的交互作用时，常常选用全因子设计。

当因子水平超过 2 时，由于试验次数随因子个数的增长呈指数速度增长，因而通常只作二水平的全因子试验。如果确实需要作三水平或更多水平全因子试验时，计算机软件中也有与二水平完全相同的试验设计及分析方法，本节对此不再单独叙述。但通常认为，加上了中心点之后的二水平试验设计在工程实践中已经足够了，在相当大程度上它可以代替三水平的试验，而且分析简明易行，现已被工程师们普遍使用。

7.3.1 二水平全因子试验概述

将 k 个因子的二水平的全因子试验记为： 2^k 试验。这是整个全因子试验的记号，而不仅仅是试验次数。当然， 2^k 也恰好是 k 个因子的二水平的全因子试验所需要的最少试验次数。这对下节理解部分实施因子设计的记号也是有用的。

1. 试验目的

什么时候需要进行 2^k 试验呢？一般来说，进行任何试验都要进行好几批，大约是：先用部分实施的因子设计进行因子的筛选，让因子个数最后不超过 5 个，然后用全因子试验设计进行因子效应和交互作用的全面分析，进一步筛选因子直到因子个数不超过 3 个时，最后用响应曲面方法（RSM）确定回归关系并求出最优设置。因此，部分实施因子试验通常只是为了筛选因子，可以稍粗糙些，因而试验次数也较少些；而响应曲面试验设计是为了得到包含平方项的非常精细的回归方程，试验次数最多，因而只对较少因

子(个数 ≤ 3)有实际意义。全因子试验既不像部分实施因子试验,也不像响应曲面试验设计,它可以兼有筛选因子和建立回归方程两方面目的,它可以分析出所有因子的主效应和各因子间的各阶交互作用的效应,回归方程中将包含全部一次项以及各因子的乘积项,试验次数也较适中,它可以适用于因子个数不超过5的情况。由于全因子试验设计方法和理论是理解部分实施因子试验方法和响应曲面试验设计方法的基础,因此本节详细介绍全因子试验设计。

2. 试验的安排及中心点的选取

在例7-1中已经介绍了一个三因子的全因子试验的安排,当时只是作为例题展示了基本思路,还有很多细节需要加以补充,主要是研究如何在试验设计中考虑到三个基本原则:重复试验、随机化和划分区组。

如何实现“重复试验”呢?一种办法就是将每一个试验条件都重复2次或更多次,这样做的好处是对于试验误差估计得更准确,但代价却是大大增加了试验次数因而增加了试验成本。另一种更巧妙的办法就是只在“中心点”处安排重复试验,通常是在中心点重复做3次或4次试验。

中心点在所有因子都是连续变量时容易找到,那就是各因子皆取其高水平与低水平的平均值。当因子全部是离散变量时,可以选取它们各种搭配中的某一个组合作为“伪中心点”;当因子中既有连续变量又有离散变量时,则可以对连续变量选取其平均值,离散变量选取某一个组合作为“伪中心点”(这时不要求试验的平衡,强调的是确实要有重复)。

选取“中心点”并在此处安排重复试验的好处,主要是为了进行完全相同条件下的重复,因而可以估计出试验误差即随机误差。另外,由于每个连续因子的取值原来只有2个值(高水平及低水平),现在增加到了3个值,因而增加了对于响应变量可能存在的弯曲趋势估计的能力,这也是简单全面重复所不能达到的效果。而且,在把因子点试验的顺序随机化之后,如果再把中心点处所进行的3次或4次试验安排在全部试验的开头、中间和结尾,那么这几个点的试验结果间只应存在随机误差。如果这几个试验结果呈现非常明显的上升、下降或其他不正常的趋势,则有可能帮助我们发现在试验过程中出现的不正常状况。

总之,安排每个因子取二水平,再加上中心点,则可以构成较好的全因子试验安排。7.4节的部分实施的因子试验中,可以采用相同方法来安排中心点。

3. 代码化及其计算

无论试验的目的是筛选因子还是建立关系式,都要建立回归方程。那么,是对原始的自变量数值进行回归分析好呢,还是先将这些数值进行代码化后再进行回归分析好呢?所谓代码化(coding),就是将该因子所取的低水平的代码(code)值设定为-1,高水平代码值设定为1,中心水平设定为0。经过理论上的分析后发现,将自变量代码化后有很多好处,因此应该对代码化的数据进行回归分析。

(1) 代码化后的回归方程中,自变量及交互作用项的各系数可以直接比较,系数绝对值大的效应比系数绝对值小的效应更重要、更显著。对自变量原始数据所进行的回归方程中的回归系数是有单位的。在 $Y=a+bx$ 中, b 的单位是 $[Y]/[X]$,显然, X 更换单位后,其系数也会变化。如果自变量不止一个,多个自变量间含义不同,量纲也不同,其回归系数之间显然是不可比的。如果代码化后,每个自变量都化为无量纲的 $[-1, 1]$ 间的



数据,这时,各自变量间具有相同的“尺寸”,各系数之间就可以比较了。

(2) 代码化后的回归方程中各项系数的估计量间是不相关的。很明显, x_1 与 x_1x_2 之间是相关的,它们的回归系数的估计量之间也是相关的。比如,在回归方程中,保留 x_1x_2 项及删除此项时, x_1 的回归系数肯定要发生变化,这造成了使用中的诸多不便。一旦将自变量全部代码化,则没有这个问题了,删除或增加某项,对于其他项的回归系数将不会发生任何影响。

(3) 在自变量代码化后,回归方程中的常数项(或称“截距”)就有了具体的物理意义。代码“-1”与“1”的中点恰好为“0”,而将全部自变量以“0”代入方程得到的响应变量预测值则恰好是截距值。因此,截距值是全部试验结果的平均值,也是全部试验范围中心点上的预测值。

用代码数据得到的回归方程是重要的,但用原始数据得到的回归方程有时也是有意义的,因此我们应该熟悉真实值与代码值间的换算。

举例说明之。假定,真实值的低水平是 100,高水平是 200,相应的代码值是-1 和 1。列出表格(见表 7—8)。

表 7—8 真实值与代码值换算表

	低水平 (low)	中心	高水平 (high)
真实值	100	150	200
代码值	-1	0	1

记 中心值 $M = (\text{low} + \text{high})/2$

半间距 $D = (\text{high} - \text{low})/2$

则有 代码值 $= \frac{\text{真实值} - M}{D}$

或 真实值 $= M + D \times \text{代码值}$

在本例中

$$\text{代码值} = \frac{\text{真实值} - 150}{50}, \text{真实值} = 150 + 50 \times \text{代码值}$$

7.3.2 全因子设计的计划

全因子设计的计划中,最关键的是选定因子及确定它们的水平。我们先举例说明。

例 7—4

压力成型塑胶板工艺条件问题

在压力成型塑胶板生产中,经过因子的初步筛选后,最后得知,影响成型塑胶板强度的因子有三个:压模间距(distance)、成型压力(pressure)及压力角(angle)。我们要判断哪些因子的主效应及哪些交互效应是显著的,哪种生产条件下可以获得最大的成型塑胶板强度(strength)。记

A: 压模间距,低水平: 60mm; 高水平: 70mm

B: 成型压力,低水平: 300 (Pa); 高水平: 400 (Pa)

C: 压力角, 低水平: 20 度; 高水平: 24 度

准备做全因子试验并安排 4 个中心点 (即 $2^3 + 4$) 的试验, 如何安排试验计划呢?

有了计算机的帮忙, 制定试验计划就是件简单的事情了。以 MINITAB 软件为例, 制定计划的入口是 STAT-DOE-Factorial-Create Factorial Design。虽然要生成的是全因子试验设计, 但它实际上是部分实施的因子设计中的特例。因此, 在下面填写补充信息时, 要选择试验次数是全因子的试验次数, 别的就没有问题了。

在 “Designs” 对话框内, 选中 “Full Factorial”, 填入中心点个数 (Number of Center Points)。在 “Options” 对话框内, 删除选项 “Randomization”, 这样就可以获得一个标准顺序的试验设计表格了。其中, 最后一列是空白的 (见表 7—9)。

表 7—9 全因子试验计划表 (标准顺序)

StdOrder	RunOrder	CenterPt	Blocks	Press	Dis	Angle	Strength
1	1	1	1	300	60	20	
2	2	1	1	400	60	20	
3	3	1	1	300	70	20	
4	4	1	1	400	70	20	
5	5	1	1	300	60	24	
6	6	1	1	400	60	24	
7	7	1	1	300	70	24	
8	8	1	1	400	70	24	
9	9	0	1	350	65	22	
10	10	0	1	350	65	22	
11	11	0	1	350	65	22	
12	12	0	1	350	65	22	

下面要进行的是随机化试验顺序的工作。在 “Run Order” 一列中产生随机数, 并可以将最后几个中心点的试验顺序调整为大体在试验的开始、中间和结尾。然后再以 “Run Order” 为序, 将上述表格重排为计划矩阵 (planning matrix), 列表如下 (见表 7—10)。

表 7—10 全因子试验计划矩阵

StdOrder	RunOrder	CenterPt	Blocks	Press	Dis	Angle	strength
10	1	0	1	350	65	22	
8	2	1	1	400	70	24	
1	3	1	1	300	60	20	
2	4	1	1	400	60	20	
12	5	0	1	350	65	22	
4	6	1	1	400	70	20	
7	7	1	1	300	70	24	
6	8	1	1	400	60	24	
9	9	0	1	350	65	22	
5	10	1	1	300	60	24	
3	11	1	1	300	70	20	
11	12	0	1	350	65	22	

按此表格安排试验就可以进入实施阶段了。



7.3.3 全因子设计的分析

全因子设计的分析方法是一般的试验设计分析的典型代表。共有五大步骤，其流程图见图 7—8。

这里所说的第一步“拟合选定模型”，其主要任务是根据整个试验的目的，选定一个数学模型。通常首先可以选定“全模型”。由于三阶及三阶以上的交互作用通常都可以忽略不计，在因子设计的问题中，全模型就是在模型中包含全部因子的主效应及全部因子的二阶交互作用。在经过细致分析后，如果可以断言某些主效应及某些因子的二阶交互作用是不显著的，则在下次选定模型时，将只保留那些效应显著的项。

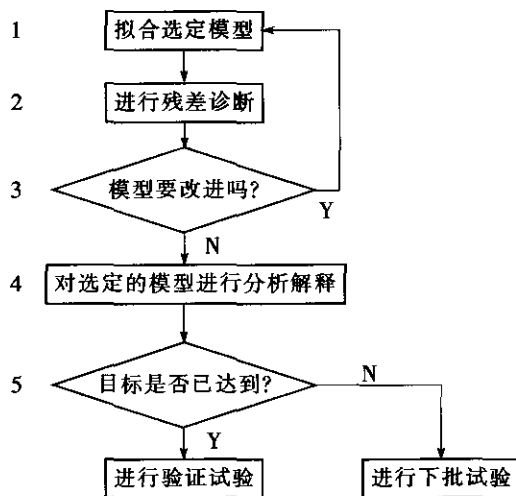


图 7—8 试验设计分析五步法流程图

由于计算机的计算全部是自动进行的，通常见到的软件都会有下列一些结果输出，我们将指出分析的要点。

第一步要注意观察下面几处。这些在 MINITAB 软件中从 STAT-DOE-Factorial-Analyze Factorial Design 窗口中选定全模型后，在运行窗口中可以看到。

1. 看 ANOVA 表中的总效果

H_0 : 模型无效 $\leftrightarrow H_1$: 模型有效

如果对应回归项 (regression) 的 p -value < 0.05 ，则表明应拒绝原假设，即可以判定本模型总的说来是有效的。如果对应回归项的 p -value > 0.05 ，则表明无法拒绝原假设，即可以判定模型总的说来是无效的。遇见这种情况就比较麻烦，这说明整个试验没有有意义的结果。造成这种情况的常见原因可能如下：

(1) 试验误差太大。由于 ANOVA 检验的基础是将有关各项的离差平方和与随机误差的平方和相比较，取其比值形成 F 统计量。如分母的随机误差平方和太大，则将使 F 变小，从而得不到“效应显著”的结论。这时候，应仔细分析误差产生的各项原因，是否能设法降低，这是能找到显著因子的关键措施。当然，也有可能“试验误差太大”是由测量系统不好造成的，这时就要设法改进测量系统。

(2) 试验设计中漏掉了重要因子。这当然与上项有某种关系。漏掉了重要因子必然会使试验误差增大, 这时更应仔细分析因子的选定。通常我们在试验前考虑因子时, 很难有把握说某因子重要或不重要, 这时候, 应该“宁多勿漏”, 而不是相反。因子多选了, 将来删去它很容易, 一旦漏掉再想找回来就困难了。

(3) 有可能是模型本身有毛病, 例如模型有失拟 (lack of fit), 或数据本身有较强的弯曲性 (curvature), 这时也可能导致判断为“模型无效”。下面将讨论这两方面的问题。

2. 看 ANOVA 表中的失拟现象

本项所检验的假设是:

$$H_0: \text{无失拟} \leftrightarrow H_1: \text{有失拟}$$

在 ANOVA 表中, 如果失拟项 (lack of fit) 的 $p\text{-value} > 0.05$, 则表明无法拒绝原假设, 即可以判定, 本模型并没有失拟现象; 反之, 就说明模型漏掉了重要的项, 应该补上。本项计算的依据是: 最初是以重复试验间的差异作为试验误差的估计 (pure error), 将缺失的项 (例如高次项、高阶交互项等) 所造成的误差平方和与之相比, 经过 F 检验即可判明。以后, 将判明为不显著的各项都归并为随机误差项, 重新计算失拟项是否显著。

3. 看 ANOVA 表中的弯曲项

本项所检验的假设是:

$$H_0: \text{无弯曲} \leftrightarrow H_1: \text{有弯曲}$$

在 ANOVA 表中, 如果弯曲项的 $p\text{-value} > 0.05$, 则表明无法拒绝原假设, 即可以判定模型并没有弯曲现象; 反之, 就说明数据呈现弯曲, 而模型中并没有平方项, 应该补上。本项计算的依据是: 最初是以重复试验间的差异作为试验误差的估计; 将高低水平的 2 个数据连同中心点的试验数据, 构成自变量的 3 个不同的观测值, 扣除线性项后可得二次项的平方和; 将二次项平方和与试验误差相比较, 经过 F 检验即可判明是否呈现弯曲。

4. 拟合的总效果多元全相关系数 (multiple correlation coefficient) R^2 (即 R-Sq) 及修正的多元全相关系数 (adjusted multiple correlation coefficient) R^2_{adj} (即 R-Sq(adj))

由回归分析中的平方和分解公式可知:

$$SS_{\text{Total}} = SS_{\text{Model}} + SS_{\text{Error}}$$

考虑 SS_{Model} 在 SS_{Total} 中的比例, 定义 R 平方 (R-Square, 简记 R-Sq):

$$R^2 = \frac{SS_{\text{Model}}}{SS_{\text{Total}}}$$

显然, 此数值越接近于 1 就越好。容易看出, 它有另一种写法:

$$R^2 = 1 - \frac{SS_{\text{Error}}}{SS_{\text{Total}}}$$

如果将自变量的这种可控的普通变量数据也看成随机变量, 则可以求出二者间的相关系数 (correlation coefficient)。而 $R\text{-sq}$ 恰好就是相关系数的平方。因此, 它的含义是很好理解的。对于多个自变量的情况, 定义不变, 它被推广为“多元决定系数”, 仍然表示 SS_{Model} 在 SS_{Total} 中的比例。但它也有一个缺点: 当自变量个数增加时, 例如只增加一个新自变量, 不管增加的这个自变量效应是否显著, R^2 (R-Sq) 都会增加一些, 因而在评价是否该增加此自变量进入回归方程时, 使用 R^2 就没有价值了。为此, 我们引入修正的 R^2 , 即 R^2_{adj} , 它的定义是:



$$R_{\text{adj}}^2 = 1 - \frac{SS_{\text{Error}}/(n-p)}{SS_{\text{Total}}/(n-1)}$$

式中, n 为观测值总个数; p 为回归方程中的总项数 (包含常数项在内)。也就是说, R_{adj}^2 (即 $R\text{-Sq}(\text{adj})$) 是扣除了回归方程中所受到的包含项数的影响的相关系数, 因而可以更准确地反映模型的好坏。它同样是越接近于 1 就越好, 而且在实际应用中, 由于回归方程所含项数 p 总会大于等于 1, 因而很容易看出, R_{adj}^2 总比 R^2 要稍小一些。因此, 要判断两个模型的优劣可以从 $R\text{-Sq}(\text{adj})$ 与 $R\text{-Sq}$ 的接近程度来判断: 二者之差越小则说明模型越好。我们常常比较包含所有自变量有关项的“全模型”与删去所有影响不显著的项后的“缩减模型”, 看看究竟哪个更好, 如果将影响不显著的项删去之后, 二者更接近, 则说明删去这些项确实使模型得到改进。

5. 对于 s 值或 s^2 的分析

考虑到所有的观测值与理论模型之间可以有误差, 但总是假定这个误差应该是以 0 为均值, 以 σ^2 为方差的正态分布。在 ANOVA 表中, 对应于残差误差 (residual error) 那行中的平均离差平方和 (adj MS) 的数值则恰好是 σ^2 的无偏估计量, 我们将其记为均方误差 MSE (mean square of error), 而有些软件会将其平方根 s 一并输出, 我们可以认为 s 值是 σ 的估计。粗略地说, 在预测值的基础上, 加减 2 倍 s , 则可以得到预测值的 95% 的置信区间。因此这里的 s 在分析模型的好坏中起着关键性的作用, 显然 s 值越小说明模型越好。因此, 比较两个模型的优劣最关键的指标就可以选择 s 或 s^2 , 哪个模型能使之达到最小, 哪个模型就最好。

6. 各项效应的显著性

在计算结果的最开始部分, Estimated Regression Coefficients for y 中, 列出了各项的效应及检验结果。这是对各项的分别的检验, 很可能有些项是显著的, 而另一些项不显著, 将来修改模型时, 应该将这些不显著项删除。这里要注意的是: 如果一个高阶项是显著的, 则此高阶项所包含的低阶项也必须包含在模型中。例如, 二阶交互作用 BC 项显著, 则 B 及 C 这两个主效应项也一定要包含在模型中, 即使表面上看这两个主效应项本身并不显著。

对于各项效应的显著性, 计算机还输出一些辅助图形帮助我们判断有关结论。这里最重要的就是 **Pareto 效应图** (Pareto effect plot) 及 **正态效应图** (normal effect plot)。

Pareto 效应图是将各效应的 t 检验所获得的 t 值作为纵坐标, 按照绝对值的大小排列起来, 根据选定的显著性水平 α , 给出 t 值的临界值, 绝对值超过临界值的效应将被选中。

用 Pareto 效应图来判断因子效应的显著性是非常直观的, 但它有个重要缺点, 那就是进行各效应的 t 检验时, 首先要用 s^2 估计出 σ^2 来, 而通常 s^2 并不一定可靠。将各因子的效应按由小到大 (正负号考虑在内) 排成序列, 将这些效应点标在正态概率图上, 这就是 **正态效应图** (normal effect plot)。可以假定, 大多数因子中只会有极少数因子效应是显著的, 即 **效应稀疏原则** (effect sparsity principle)。因此, 当挑选位于中间的一些效应的点群拟合一条直线后, 就可以判定任何一个远离该直线的点所对应的效应是显著的。因为如果所有的效应无效, 则这些点应该服从正态分布, 点全在直线上; 某些效应确实非零时, 相应的估计效应绝对值应会偏大, 且一定会远离直线。对于正效应, 估计的效应将落在直线的右 (上) 方; 对于负效应, 估计的效应将落在直线的左 (下) 方。

经过上述各项计算,就完成了对数据的初步分析,即第一步“拟合设定模型”的任务。

DOE 分析“五步法”里所说的第二步是进行残差诊断。我们的主要目的是基于残差的情况来诊断我们的模型是否合适。单纯从 ANOVA 表及回归系数的估计与检验两方面结果来分析整个结果是远远不够的。数学上可以举出这样的一些例子,其 ANOVA 表及回归系数结果完全相同,但它们的真实状况相差甚远。为了弥补第一步所得结果的不足,我们要进行残差的诊断。如果数据与模型的拟合是正常的,则残差应该是正常的。

如何判断残差是正常的呢?我们通常假定,数据是相互独立的,且其分布为 $N(f(x), \sigma^2)$,即这些数据是以模型为均值,以固定常数 σ^2 为方差的正态分布。残差(residual)是观测到的响应变量数据与代入回归模型后的预测值之差,因此,应该有: $\text{residual} \sim N(0, \sigma^2)$ 。我们应该观察一下,残差是否真的这样分布。如果确实是这样分布,则可以相信我们所选定的模型是正确的,否则就说明我们选定的模型不正确,要对模型进行某些修改。由于事先对选定的模型真的没有什么绝对的把握,因此,残差诊断还是很有必要的。

具体地说,残差诊断应包括下列 4 个步骤,观察计算机自动输出的四个图形。在 MINITAB 软件中可以从回归分析 STAT-Regression-Regression 内或从试验设计 STAT-DOE-Factorial-Analyze Factorial Design 内的对话框 Graphs 中选定并得到它们。

(1) 观察残差对于以观测值顺序为横轴的散点图,重点考察此散点图中,各点是否随机地在水平轴上下无规则地波动着。

(2) 观察残差对于以响应变量拟合预测值为横轴的散点图,重点考察此散点图中,残差是否保持着等方差性,即是否有“漏斗型”或“喇叭型”。

(3) 观察残差的正态性检验图,看残差是否服从正态分布。

(4) 观察残差对于以各自变量为横轴的散点图,重点考察此散点图中是否有弯曲趋势。

残差诊断的四个图都应是正常的。什么是最常见的情况而被认为“不正常”呢?

第一种最常见残差图不正常情况出现在第二个图中,即残差对响应变量拟合值的图中,残差未保持等方差,散点明显呈“漏斗型”或“喇叭型”,这说明对响应变量 y 作某种变换后才会与模型拟合更好。例如,取 $y^* = \sqrt{y}, y^2, \frac{1}{y}, \frac{1}{\sqrt{y}}, \ln y$ 等作为新响应变量,可能会好得多。一般规则是这样的:如果从图中可以看出,残差的标准差 σ 大体上与拟合值 \hat{y} 的 u 次方成正比时,则可以进行 Box-Cox 变换:

$$y^* = \begin{cases} y^\lambda, & \lambda \neq 0 \\ \ln y, & \lambda = 0 \end{cases}$$

式中, $\lambda = 1 - u$ 。由于人们用肉眼来判断 u 的取值很不容易,不少计算机软件会自动给出对 λ 的建议值来。要注意的是:虽然这里变换的形式及名称都是 Box-Cox 变换,但这与将非正态数据转换为正态数据的 Box-Cox 变换完全是两回事,MINITAB 14 版也并未提供相应的窗口。

第二种最常见残差图不正常情况在第四个图中,即残差对自变量的散点图中。残差虽保持等方差,但散点明显呈 U 型或反 U 型弯曲,这说明对响应变量 y 而言,对该自变量



x 仅取线性项已经不够了, 应增加 x 的平方项或立方项, 将会使模型拟合得更好。

总之, DOE 分析的第二步只是进行残差诊断, 看“模型是否有问题”。如果确实发现了问题, 又怎么办呢? 这就转入了第三步。

DOE 分析“五步法”里所说的第三步是判断模型是否要改进。我们考虑采取措施的主要依据是基于数值计算及残差诊断的两方面结果。如果残差诊断的四个图都是正常的, 则说明模型是正常的。因此, 如果残差诊断提示我们, 要对响应变量 y 作某种变换后才行, 或要增加 x 的平方项或立方项才会使模型拟合更好, 那么我们就一定要修改模型。另外, 基于各项效应及回归系数计算的显著性分析中可以发现, 有些主效应项或交互效应项并不显著, 这些项应该从模型中删除, 模型的拟合要重新进行。总之, 凡是发现模型需要修改时, 就要返回最初的第一步, 重新建立模型, 再重复前面所有各步骤。

经过前三步的多次反复, 可以获得一个我们认为最满意的方程, 将它定为选定的模型。DOE 分析“五步法”里所说的第四步是“对选定模型进行分析解释”, 主要是在拟合选定模型后输出更多的图形和信息, 并作出有意义的解释。主要有列三方面, 而这些要求在一般的计算机软件中都会自动给出。

(1) 输出各因子的主效应图和交互效应图。在 MINITAB 软件中可以从 STAT-DOE-Factorial-Factorial Plot 窗口中得到它们。我们应该从主效应图、交互效应图的图形上进一步确认所选中的那些主因子和交互作用项是否真的很显著以及未选中的那些主因子和交互作用项是否真的不显著, 从而更具体更直观地确认我们选定的模型。

(2) 输出等高线图、响应曲面图等。在 MINITAB 软件中可以从 STAT-DOE-Factorial-Counter Surface Plot 窗口中得到它们。我们应该从等高线图、响应曲面图上进一步确认, 响应变量是如何受所选中的那些主因子和交互作用项的影响的, 它的变化规律如何。如果目标是望大 (或望小或望目), 那么自变量在什么设置时, 将会达到响应变量最大 (最小或与目标最接近)? 我们要从直观上看到整个试验范围内的最佳值的位置。由于等高线图、响应曲面图都只能对两个自变量的情形给出, 因此当自变量个数超过两个时, 要分别对选定的某两个自变量作图。其实, MINITAB 或其他很多软件都可以自动给出所有两个自变量的组合的等高线图和响应曲面图。

(3) 实现最优化。按照具体问题的望大、望小或望目在数值上求出在整个试验范围内的最佳值。在 MINITAB 软件中可以从 STAT-DOE-Factorial-Response Optimizer 窗口中得到它们。虽然在因子设计阶段, 试验设计的主要目标是筛选变量, 但实际上, 根据 DOE 分析第一步, 就已经可以判断出哪些自变量效应是显著的, 哪些自变量效应是不显著的, 我们可以在使用这些信息的基础上, 获得最佳值。计算机提供的“响应变量优化器” (response optimizer) 是个功能强大的工具, 它可以自动给出最优设置。通常只要在选定响应变量后, 再对最优的目标 (goal) 予以设定 (setup) 即可。

DOE 分析“五步法”里所说的第五步是判断“目标是否已经达到”。这主要是将预计的最佳值与原试验目标相比较。如果离目标尚远, 则应考虑安排新一轮试验, 通常是在本次获得的或预计的最佳点附近, 重新选定试验的各因子及其水平, 继续做因子设计 (DOE) 或回归设计 (RSM), 以获得更好的结果。如果已基本达到目标, 则要做验证试验以确保将来按最佳条件生产能获得预期效果。试验设计大师博克斯 (G. Box) 有句名言: “所有统计模型都是错的, 只不过有些模型有用罢了。” 我们根据试验数据所建立的任

何统计模型是否符合实际规律,在未验证前都是没有把握的。通常的做法是先算出在最佳点的观测值的预测值及其变动范围,然后在最佳点做若干次验证试验(通常要3次以上),如果验证试验结果的平均值落入事先计算好的范围内,则说明一切正常,模型是正确的,预测结果可信;否则就要进一步分析发生错误的原因,改进模型,再重新验证,以求得符合实际数据的统计模型。

MINITAB软件在DOE栏中给出预测区间的功能较弱,只好借助于回归分析(regression)栏。为此,要先将所有交互作用的乘积(或将来在响应曲面方法中会出现的平方项)用计算器(calculator)计算后形成新列,然后进行回归(STAT-Regression-Regression),在其对话框Options选项内,严格按选定模型中自变量的顺序,填写上最优设置的数值,并挑选输出所有补充计算的结果:预测值(fits),预测值处的标准差(SE of fits),预测值处的回归结果的置信区间(confidence limit,输出时缩写为95% CI),预测值处的单个观测值的置信区间(prediction limit,输出时缩写为95% PI),则计算机将自动给出上述数值。

前一个置信区间(95.0% CI)表明的是回归方程上的点的置信区间,这是由于回归系数是以样本观测值为基础估计出来的,它必然会有误差,而回归系数的误差必然导致回归方程上的预测值的误差。对此区间可以理解为,当按照此自变量的设置无限多次地重复运行下去将会获得的理论均值的95%置信范围,此值可以作为改进的结果的预报写在总结报告中。后一个置信区间(95.0% PI)表明的是以上述回归方程上的点的置信区间为基础,加上观测值具有的方差为 σ^2 的波动而给出的置信区间。这里我们假定波动的方差为不依赖于自变量位置的常数 σ^2 ,并已经求出其无偏估计(mean square of error, MSE)。这里,(95.0% PI)是将来作一次验证试验时将要落入的范围,可供作验证试验时使用。

以上经回归分析所算出的也只是无限多次平均值及单个观测值的95%的置信区间,如果需要求出 m 个(例如 $m=3$ 或 $m=5$)观测值的95%置信区间,则只能在MINITAB条件下编写宏指令,或直接用手算。 m 个观测值的95%置信区间计算公式是:

$$\hat{y} \pm t_{1-\alpha/2}(n-p) \sqrt{(\text{SE of fits})^2 + \frac{\text{MSE}}{m}}$$

式中, n 是试验总次数; p 是最终模型中所包含的项数(常数项要计算在内); m 是验证试验的次数;SE of fits是在回归方程预测值时输出的拟合值的标准误;MSE是在ANOVA表中输出的误差的MS项,也就是 σ^2 的无偏估计量。显然,当 $m=1$ 时,此置信区间就是95.0% PI;当 m 趋于无穷大时,此置信区间就是95.0% CI;当需要求出 m 个(例如 $m=3$ 或 $m=5$)观测值的95%置信区间时,只要代入相应的 m 的数值就行了。

7.3.4 全因子设计实例

案例7-4

压力成型塑胶板工艺条件问题

我们实施全因子试验,按计划矩阵做试验,就得到了相应数据记录,结果见表7-11。对于这批数据进行分析,其步骤如下:

表 7-11

压力成型塑胶板工艺条件试验结果数据表

StdOrder	RunOrder	CenterPt	Blocks	Press	Dis	Angle	strength
10	1	0	1	350	65	22	60.2
8	2	1	1	400	70	24	61.3
1	3	1	1	300	60	20	61.8
2	4	1	1	400	60	20	55.3
12	5	0	1	350	65	22	59.1
4	6	1	1	400	70	20	63.5
7	7	1	1	300	70	24	94.8
6	8	1	1	400	60	24	49.6
9	9	0	1	350	65	22	73.3
5	10	1	1	300	60	24	55.2
3	11	1	1	300	70	20	89.1
11	12	0	1	350	65	22	69.5

第一步：拟合选定模型。我们首先将全部备选项列入模型。这里包含 Press（压力），Distance（距离），Angle（角度）及它们之间的交互作用项：Press * Dis，Press * Angle 和 Dis * Angle。注意，在这里只选了全部主效应和二阶交互作用项，并不含三阶交互作用项。其计算结果如下：

Factorial Fit: strength versus Press, Dis, Angle

Estimated Effects and Coefficients for strength (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		66.058	1.619	40.80	0.000
Press	-17.800	-8.900	1.983	-4.49	0.006
Dis	21.700	10.850	1.983	5.47	0.003
Angle	-2.200	-1.100	1.983	-0.55	0.603
Press * Dis	-11.750	-5.875	1.983	-2.96	0.031
Press * Angle	-1.750	-0.875	1.983	-0.44	0.677
Dis * Angle	3.950	1.975	1.983	1.00	0.365

S=5.60846 R-Sq=92.35% R-Sq (adj) =83.17%

Analysis of Variance for strength (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	1585.14	1585.14	528.380	16.80	0.005
2-Way Interactions	3	313.46	313.46	104.485	3.32	0.115
Residual Error	5	157.27	157.27	31.455		
Curvature	1	1.71	1.71	1.707	0.04	0.844
Lack of Fit	1	9.68	9.68	9.680	0.20	0.686
Pure Error	3	145.89	145.89	48.629		
Total	11	2055.87				

从 ANOVA 表中可以清楚地看出, 主效应项 (Main Effects) 中, p -value 为 0.005, 显示我们所选定的模型总的效果是显著的、有效的。在弯曲 (curvature) 一栏中, p -value 为 0.844, 显示这里响应变量 strength 没有明显的弯曲趋势。在失拟 (lack of fit) 一栏中, p -value 为 0.686, 显示这里对响应变量 strength 的拟合没有明显的失拟。

计算机软件通常直接给出下列结果:

$$R\text{-Sq}=0.923\ 5, R\text{-Sq (Adj)}=0.831\ 7, S=5.608\ 46$$

在 ANOVA 表之前得到的 Estimated Effects and Coefficients for strength (coded units) 是各项主效应及各交互效应的结果。从显著性来看, 因子 A (Press), B (Distance), AB (Press*Dis) 对应的 p -value 都小于显著性水平 0.05, 因此可以判定, 这三项效应是显著的, 而其余各项皆不显著。

另外, 从因子效应的 Pareto 图 (见图 7—9) 及因子正态效应图 (见图 7—10) 中, 也可以验证这一点。

DOE 分析的第二步是残差诊断。按规定的四个步骤可以进行残差的分析:

(1) 观察残差对于以观测值顺序为横轴的散点图 (见图 7—11), 重点考察对观测顺序的散点图中看各点是否随机地在水平轴上下无规则地波动, 是否有不正常的升降趋势。

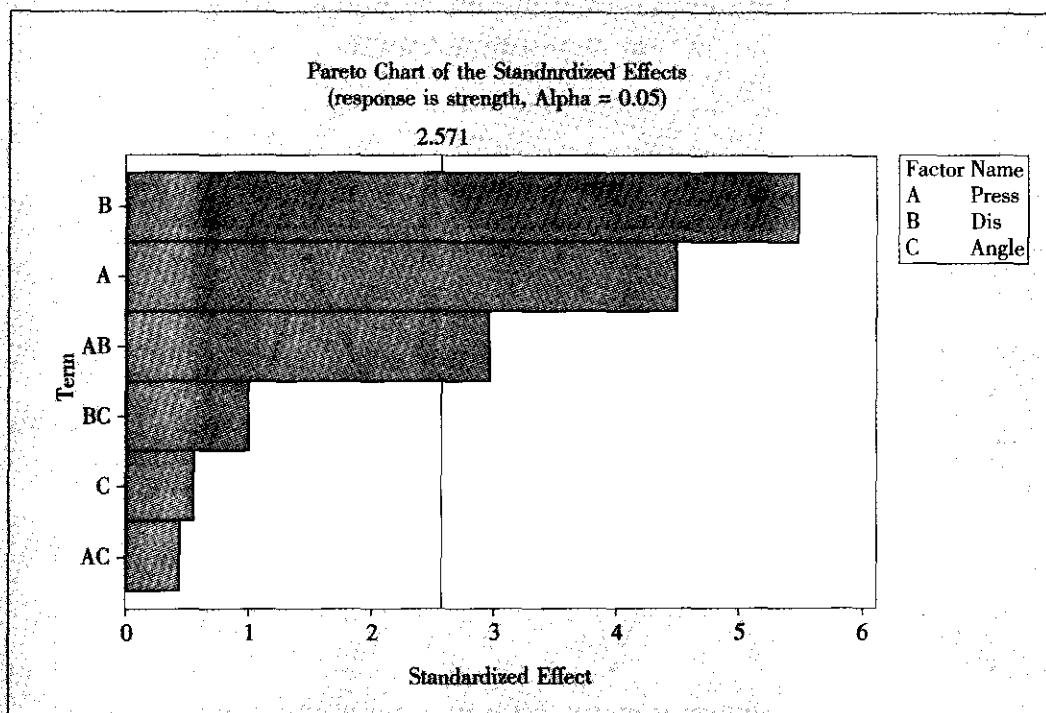


图 7—9 例 7—4 因子 Pareto 效应图

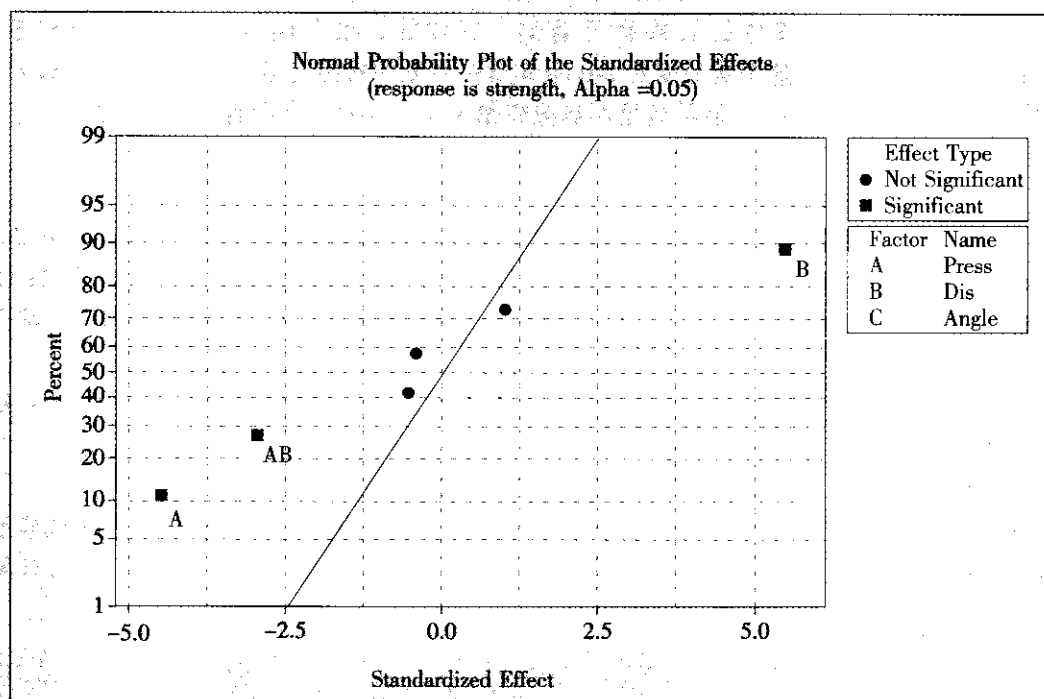


图 7—10 例 7—4 因子正态效应图

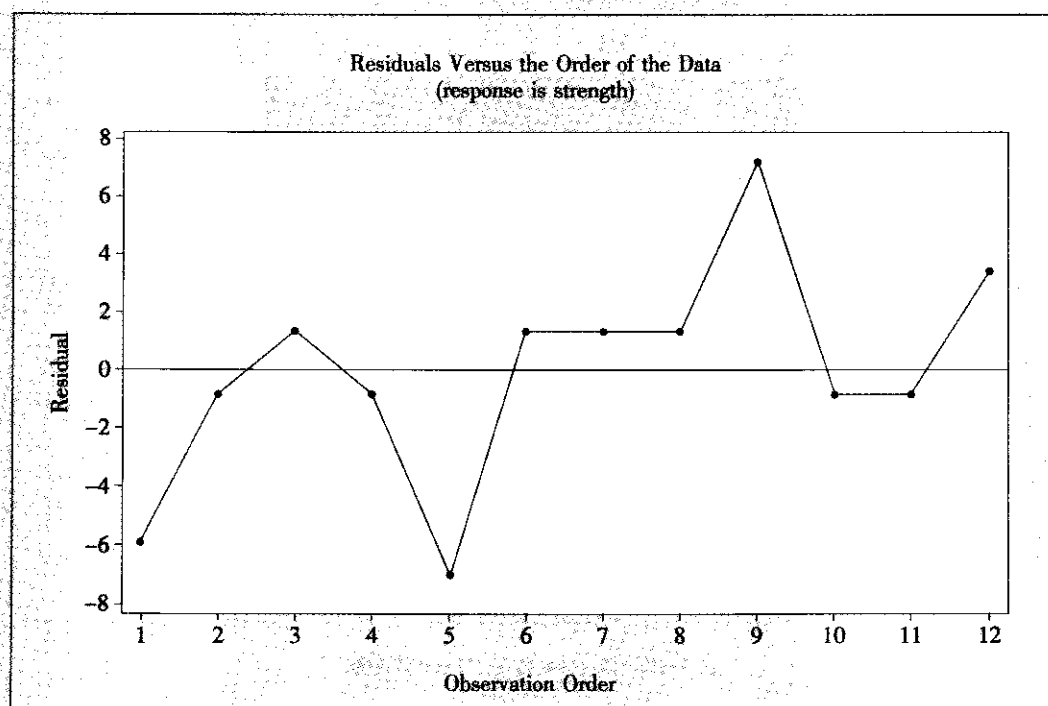


图 7—11 残差对于观测值顺序的散点图

本例中，此图是正常的。

(2) 观察残差对于响应变量拟合值的散点图（见图 7—12），重点考察残差是否保持着等方差，即是否有“漏斗型”或“喇叭型”。

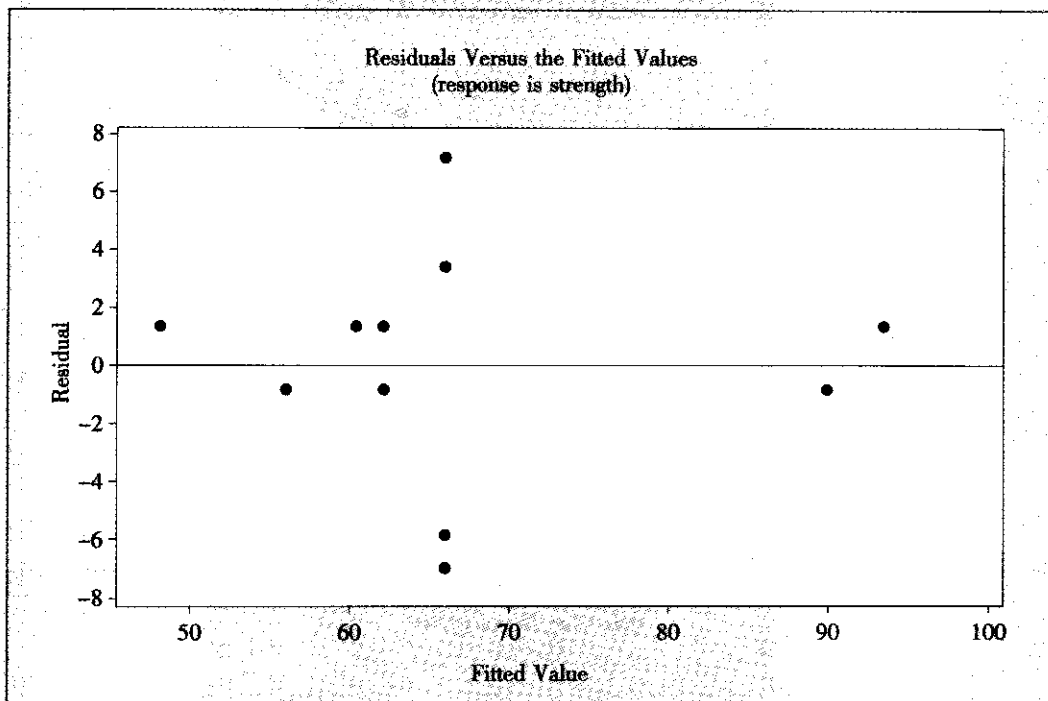


图 7—12 残差对于响应变量拟合值的散点图

本例中，此图是正常的。

(3) 观察残差的正态性检验图（见图 7—13），看残差是否服从正态分布。

这里只有粗糙的散点图，要想得到准确的检验结果，可以直接对残差进行正态性检验（见图 7—14）。

本例中，因为这里正态性检验的 p -value 为 0.097，因此残差可以被认为服从正态分布。如果这里不能服从正态分布也不用着急，因为现在还只是初步模型，还要经过模型的修改，很可能将来残差对于最终模型是正态的。

(4) 观察残差对于各自变量的散点图（见图 7—15、图 7—16 及图 7—17），重点考察是否有弯曲趋势。

本例中，这三个图都是正常的，看不出有什么弯曲的趋势。

第三步：判断模型要改进吗？

从残差诊断中看出，模型基本上是好的，只是在检验各项效应中，发现三个自变量主效应中，只有 A: Press（压力），B: Distance（距离）是显著的，C: Angle（角度）作用不显著。它们之间的三个交互作用项中，仅 AB（Press * Dis）显著，其余两个交互作用都不显著。因此修改拟合模型中的“选项”（Term），重新进行计算，即又回到第一步。



新第一步：拟合选定模型。

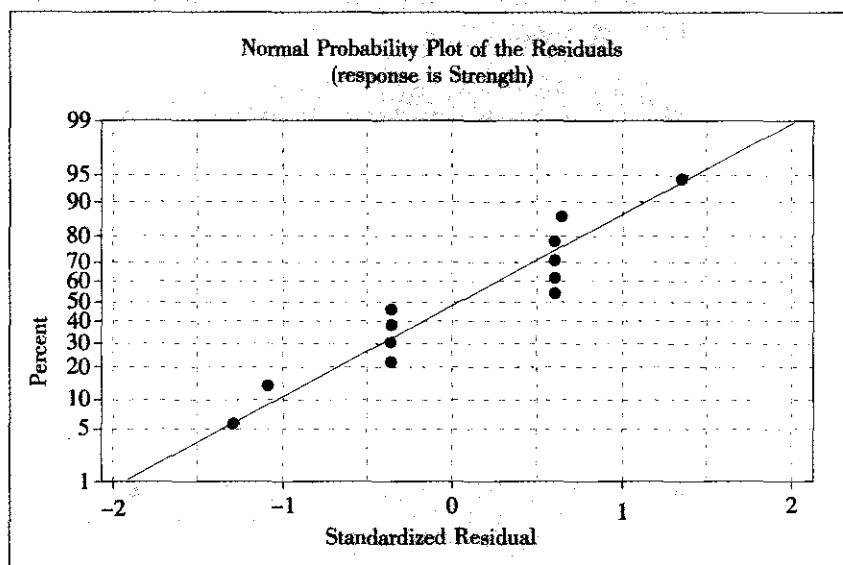


图 7—13 残差的正态分布散点图

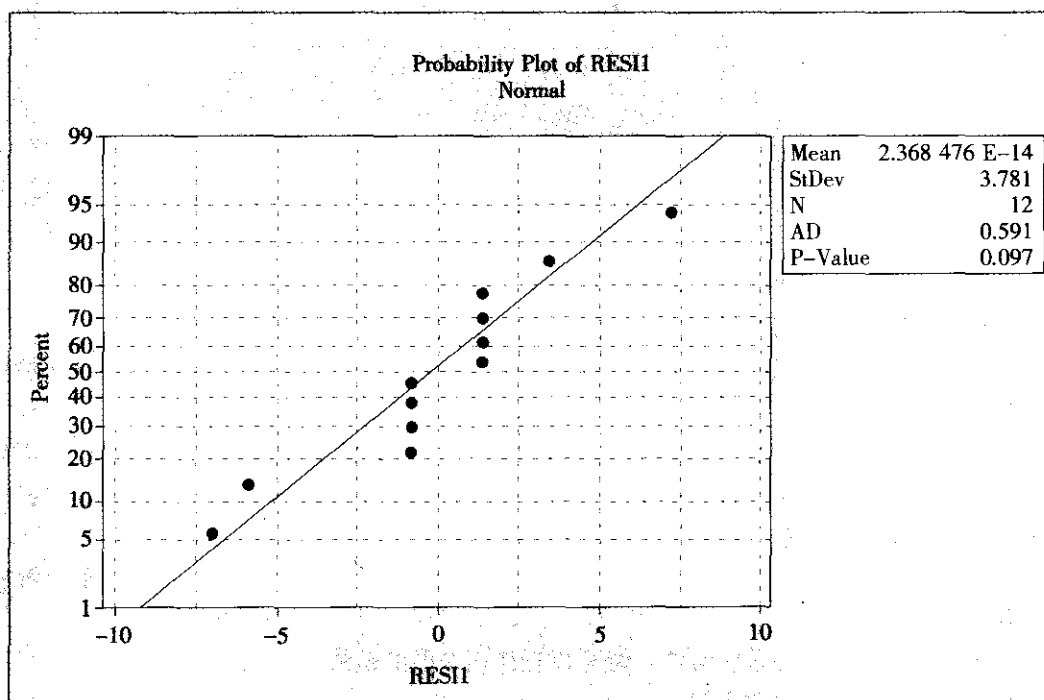


图 7—14 残差的正态分布检验图

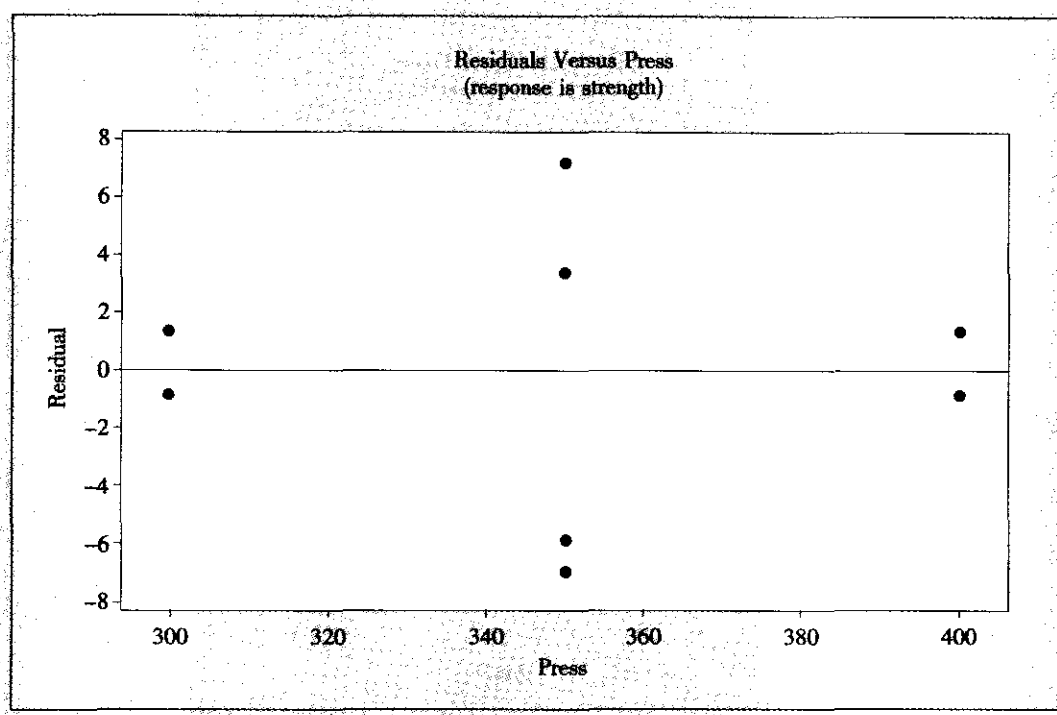


图 7—15 残差对自变量“压力”的散点图

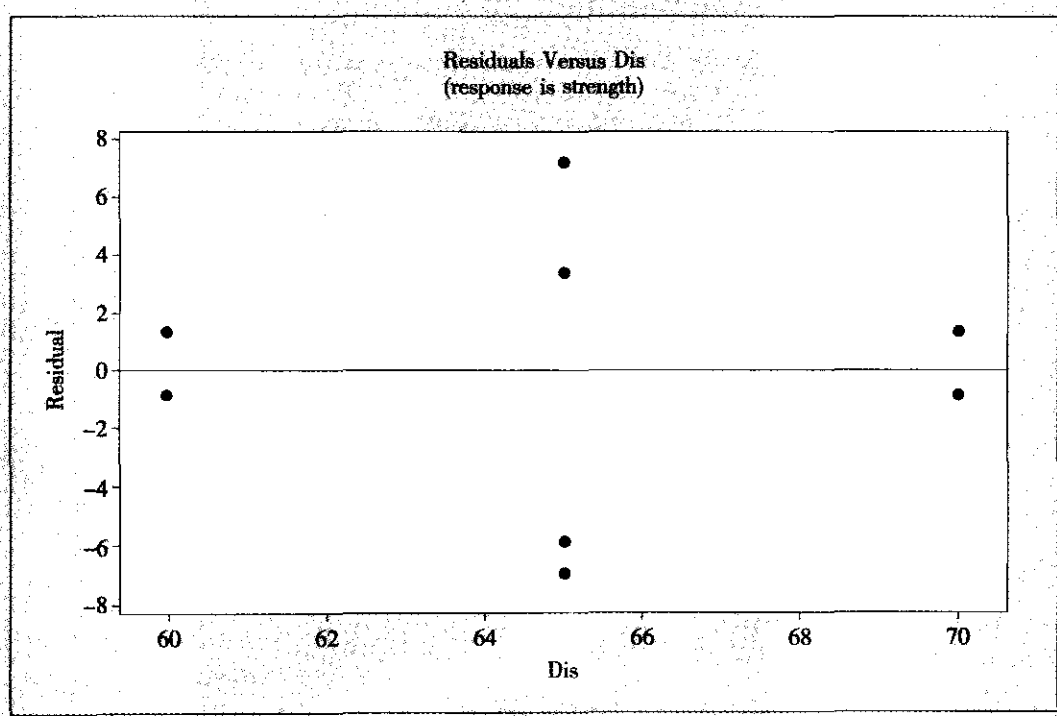


图 7—16 残差对自变量“距离”的散点图

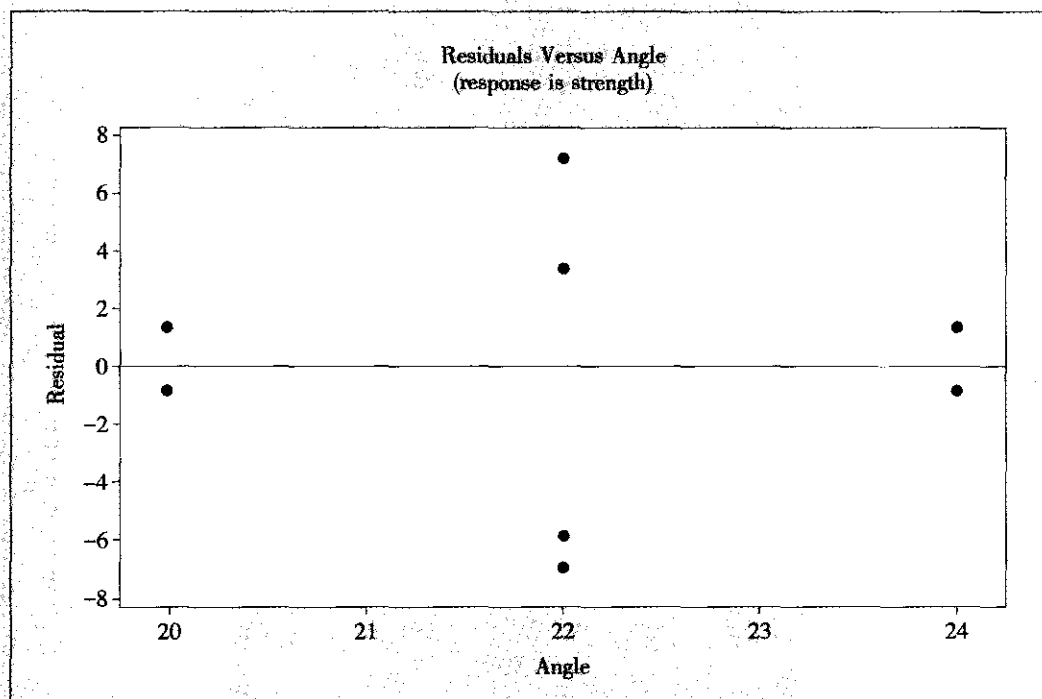


图 7—17 残差对自变量“角度”的散点图

Factorial Fit: strength versus Press, Dis

Estimated Effects and Coefficients for strength (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		66.058	1.459	45.28	0.000
Press	-17.800	-8.900	1.787	-4.98	0.001
Dis	21.700	10.850	1.787	6.07	0.000
Press * Dis	-11.750	-5.875	1.787	-3.29	0.011

S=5.053 27 R-Sq=90.06% R-Sq (adj) =86.34%

Analysis of Variance for strength (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	2	1 575.46	1 575.46	787.730	30.85	0.000
2-Way Interactions	1	276.12	276.12	276.125	10.81	0.011
Residual Error	8	204.28	204.28	25.536		
Curvature	1	1.71	1.71	1.707	0.06	0.815
Pure Error	7	202.58	202.58	28.940		
Total	11	2 055.87				

Estimated Coefficients for strength using data in uncoded units

Term	Coef
Constant	-547.317
Press	1.349 50
Dis	10.395 0
Press * Dis	-0.023 500 0

对于这次的计算结果，仍用前面介绍过的方法与步骤来进行分析。

此时观察到下面几处要点：

先看 ANOVA 表中的总效果。在本例中，对应主效应项及二阶交互项的 p -value 分别为 0.000（即 <0.001 ）和 0.011，表明应拒绝原假设，即可以判定模型总的说来是有效的。在本例中，对应失拟项（Lack of Fit）的 p -value 为 0.815，其值远比临界值 0.05 大，表明无法拒绝原假设，即可以判定模型删去了好多项，但并没有造成失拟现象。看删减后的模型是否比原来有所改进，要手算多元全相关系数 R -Sq 和修正的多元全相关系数 R -Sq (adj)，然后把两个模型计算的 R -Sq 和 R -Sq (adj) 以及标准差的估计量 s 汇总成表（见表 7—12）。

表 7—12 全模型与删减模型效果比较表

	全模型	删减模型
R -Sq	0.923 5	0.900 6
R -Sq (adj)	0.831 7	0.863 4
s	5.608 46	5.053 27

可以看出，由于模型项数减少了三项， R -Sq 通常会有微小的降低（本例由 0.923 5 降低到 0.900 6），但关键是看修正的 R -Sq (adj) 是否有所提高。本例中， R -Sq (adj) 由 0.831 7 提高到 0.863 4，可见删除不显著的主因子及交互作用项后，回归的效果更好了，而 s 的值由 5.608 5 降为 5.053 3，再次证明删除不显著的主因子及交互作用项后，回归的效果更好了。

新第二步：残差诊断。

残差诊断仍按规定的 4 个步骤进行。

(1) 观察残差对于观测值顺序的散点图（见图 7—18）。

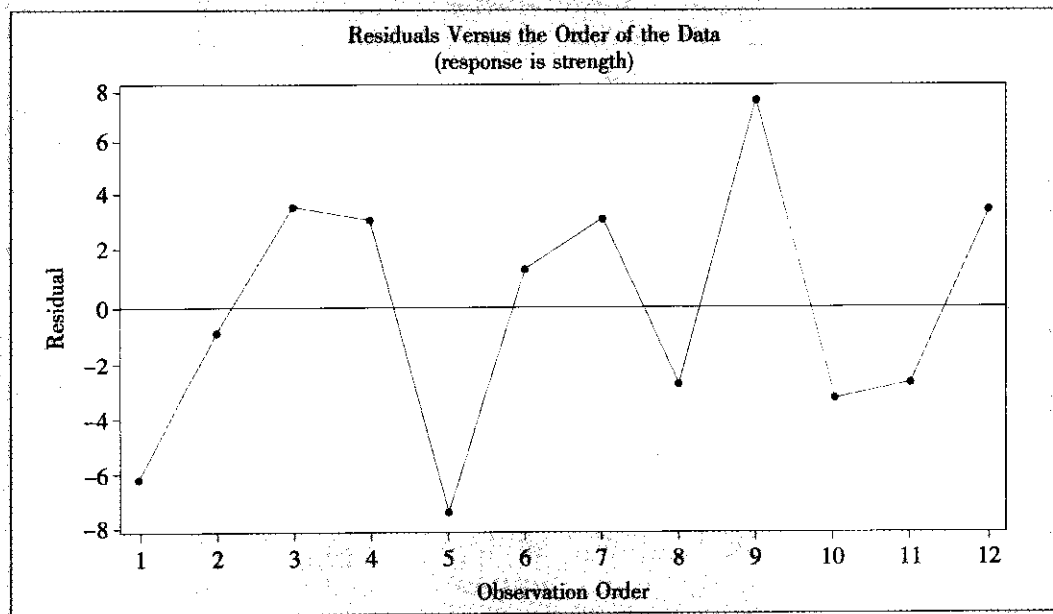


图 7—18 残差对于观测值顺序的散点图

本例中，此图是正常的。

(2) 观察残差对于响应变量拟合值的散点图（见图 7—19）。

此图重点考察残差是否保持着等方差，即是否有“漏斗型”或“喇叭型”。本例中，此图属于正常。

(3) 观察残差的正态性检验图（见图 7—20），看残差是否符合正态分布。

这里只有粗糙的散点图，要想得到准确的检验结果，可以直接对残差进行正态性检验（见图 7—21）。

本例中，因为这里正态性检验的 p -value 为 0.451，因此残差可以被认为是正态的。

(4) 观察残差对于以各自变量为横轴的散点图（见图 7—22、图 7—23 及图 7—24），重点考察是否有弯曲趋势。

对这三张图重点考察残差是否有弯曲趋势。本例中，此三图皆属于正常，没有弯曲趋势。

新第三步：判断模型是否要改进。

由上述分析，可以认为已经选定了最终的模型。根据计算结果提供的数值，可以写出最后确定的代码化数据的回归方程：

Term	Effect	Coef	SE Coef	T	P
Constant		66.058	1.459	45.28	0.000
Press	-17.800	-8.900	1.787	-4.98	0.001
Dis	21.700	10.850	1.787	6.07	0.000
Press * Dis	-11.750	-5.875	1.787	-3.29	0.011

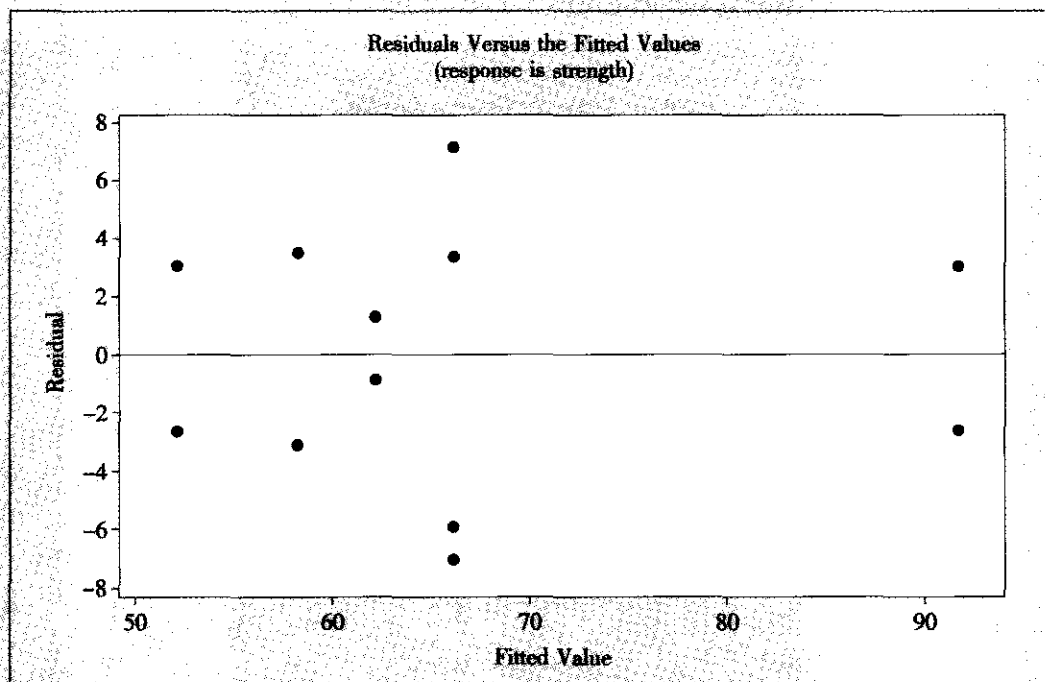


图 7—19 残差对于响应变量拟合值的散点图

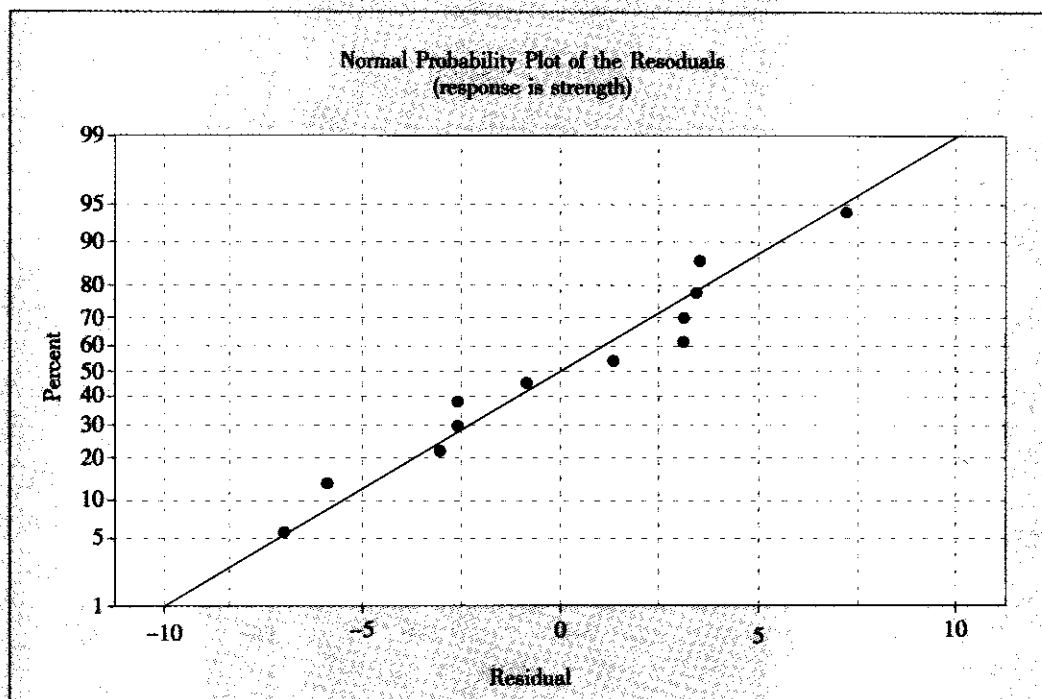


图 7—20 残差的正态分布散点图

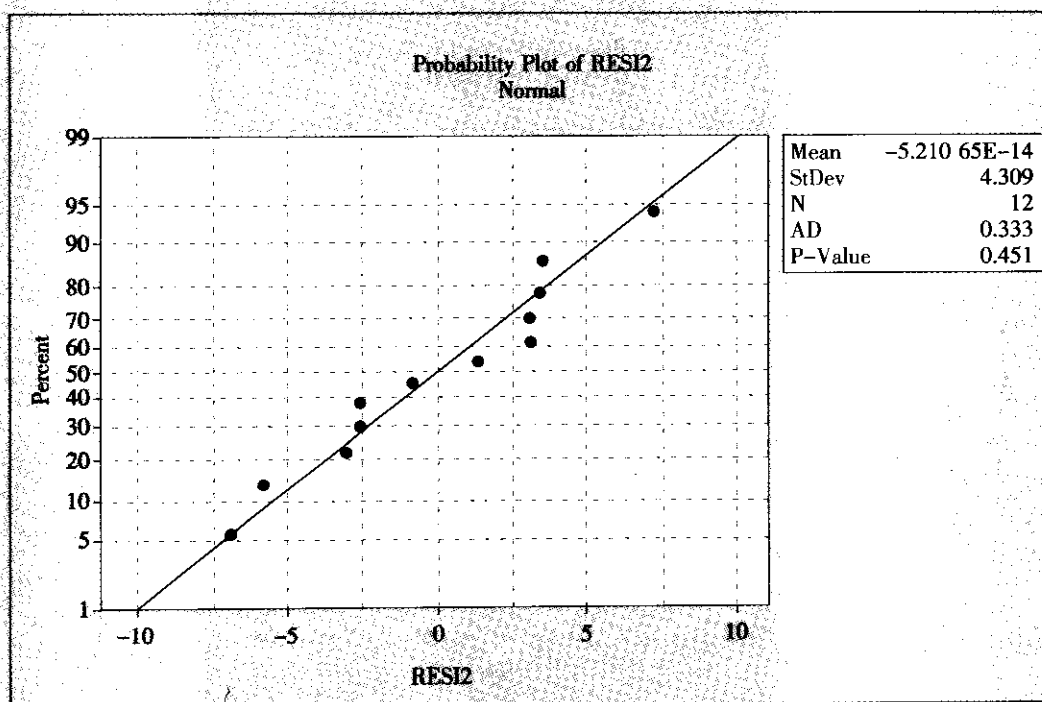


图 7—21 残差的正态分布检验图

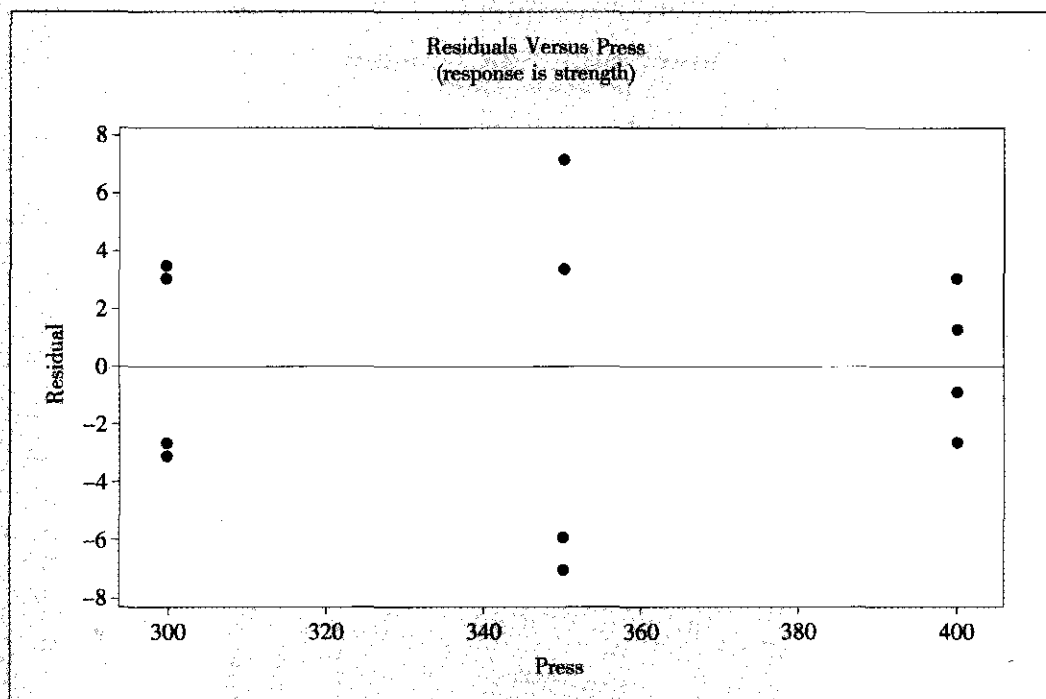


图 7—22 残差对自变量“压力”的散点图

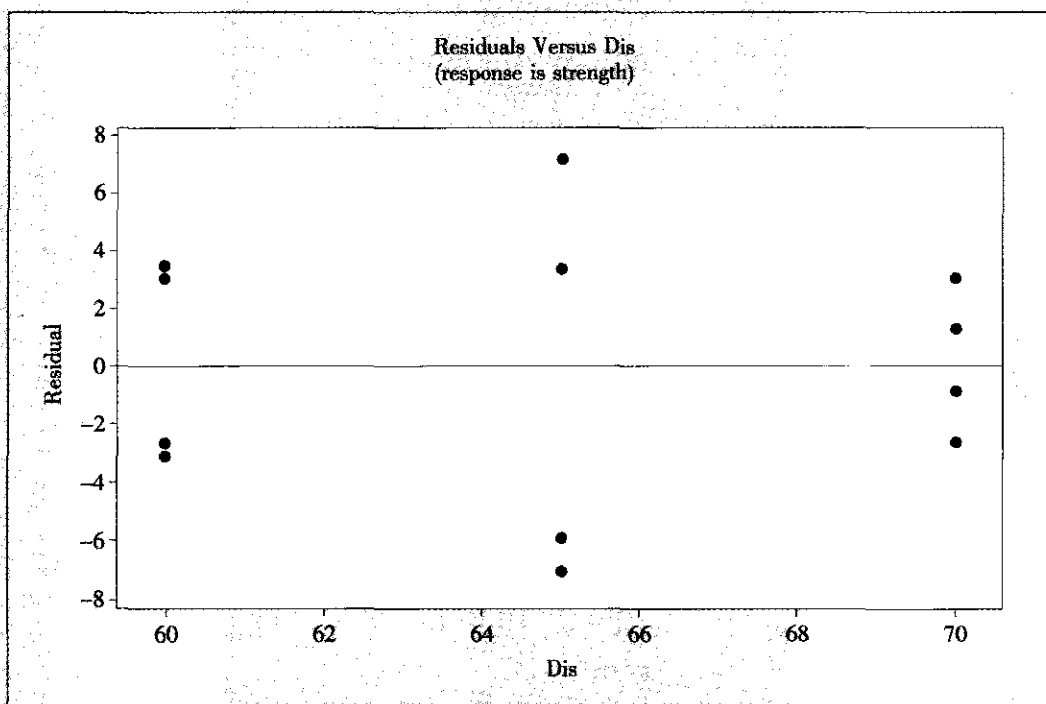


图 7—23 残差对自变量“距离”的散点图

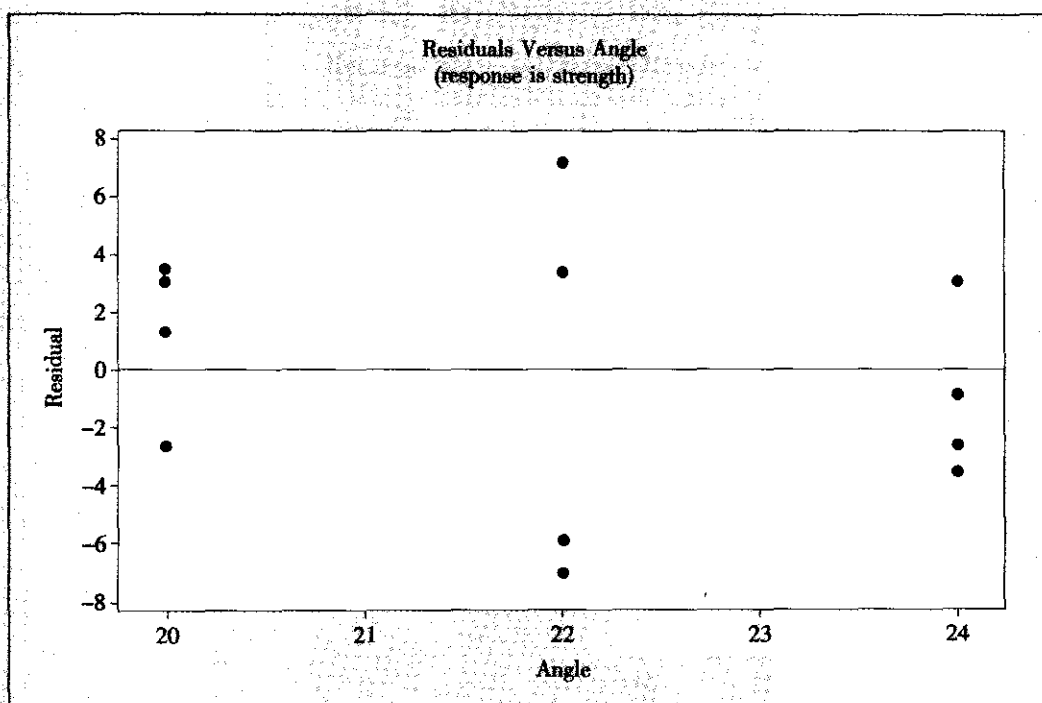


图 7—24 残差对自变量“角度”的散点图

$$y = 66.06 - 8.9\left(\frac{\text{Press} - 350}{50}\right) + 10.85\left(\frac{\text{Dis} - 65}{5}\right) - 5.875\left(\frac{\text{Press} - 350}{50}\right) * \left(\frac{\text{Dis} - 65}{5}\right)$$

计算机也提供最后确定的对原始数据的回归方程：

$$y = -547.32 + 1.3495\text{Press} + 10.395\text{Dis} - 0.0235\text{Press} * \text{Dis}$$

第四步：对选定模型进行分析解释。

我们要在拟合选定模型后输出更多的信息，并给出有意义的解释。

(1) 输出各因子的主效应图、交互效应图。主效应图见图 7—25。

从图中可以看出，因子 Press 和因子 Dis 对于响应变量 Strength 的影响确实是很显著的，而因子 Angle 对于响应变量 Strength 的影响确实是不显著的。而且还可以看出，为使 Strength 取值更大，应该让 Press 尽可能小，让 Dis 尽可能大。

交互效应图见图 7—26。

从图中可以看出，因子 Press 与因子 Dis 的交互作用对于响应变量 Strength 的影响确实是很显著的（两条线非常不平行），而其他交互效应对于响应变量 Strength 的影响确实是不显著的（两条线几乎平行）。

(2) 输出等高线图、响应曲面图等。在 MINITAB 软件中可以从 STAT-DOE-Factorial-Counter Surface Plot 窗口中得到它们（见图 7—27 及图 7—28）。

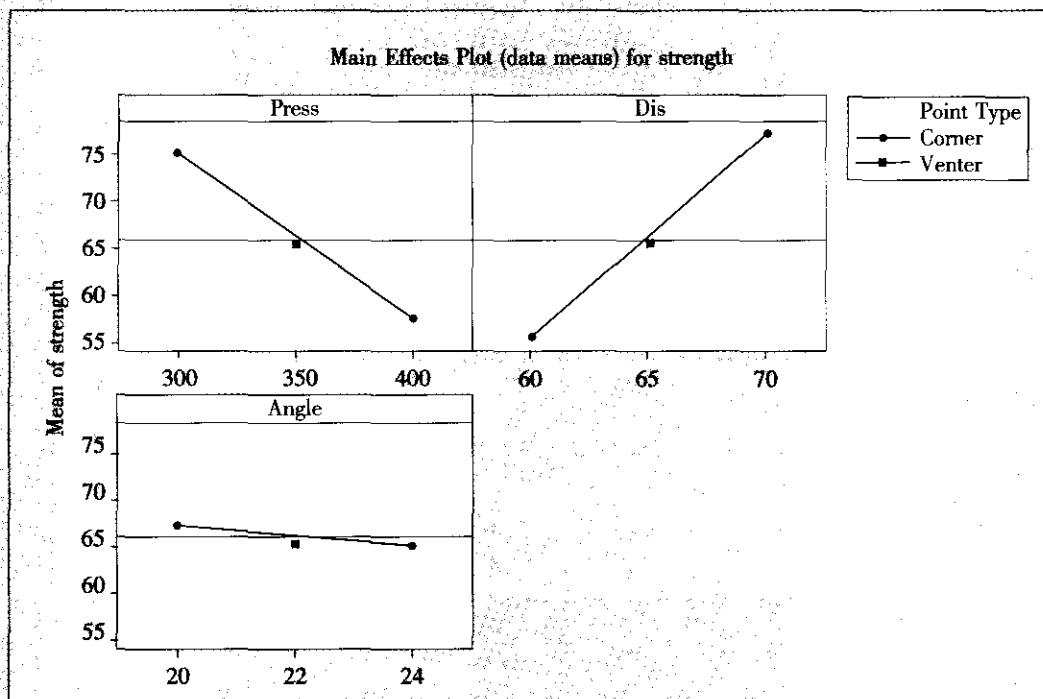


图 7—25 各因子的主效应图

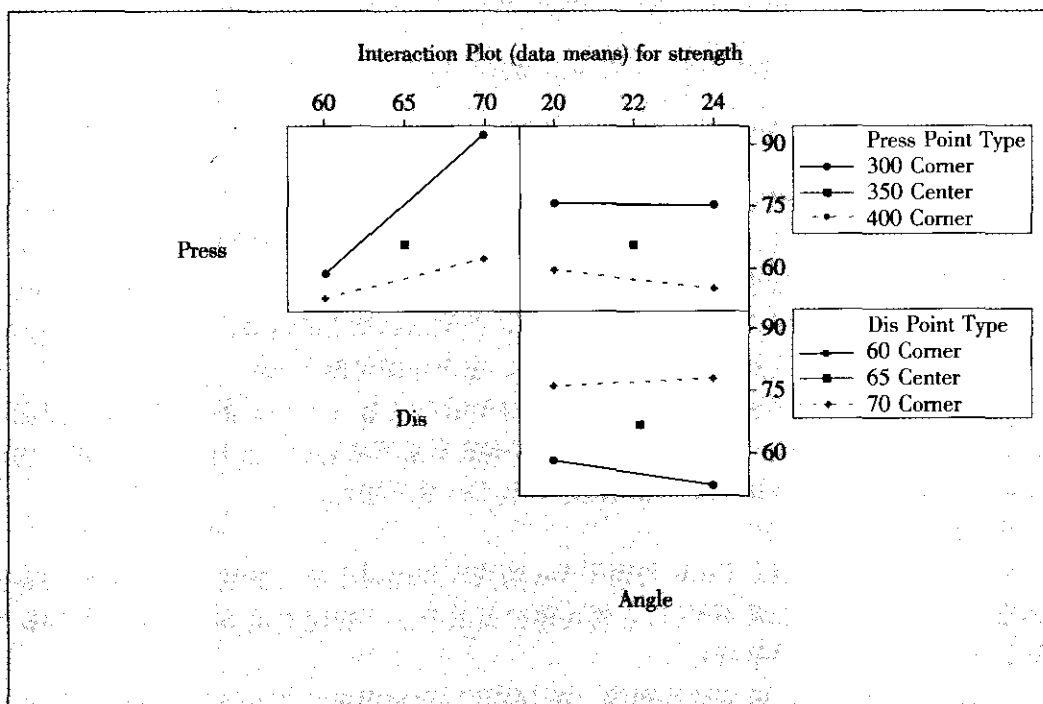


图 7—26 各因子间的交互效应图

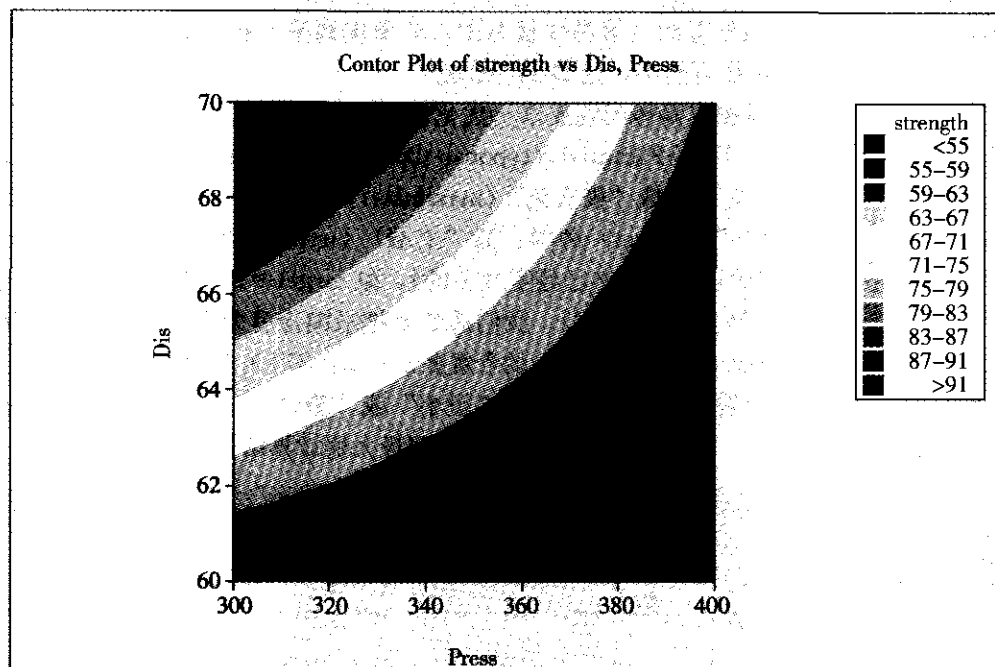


图 7—27 响应变量等高线图

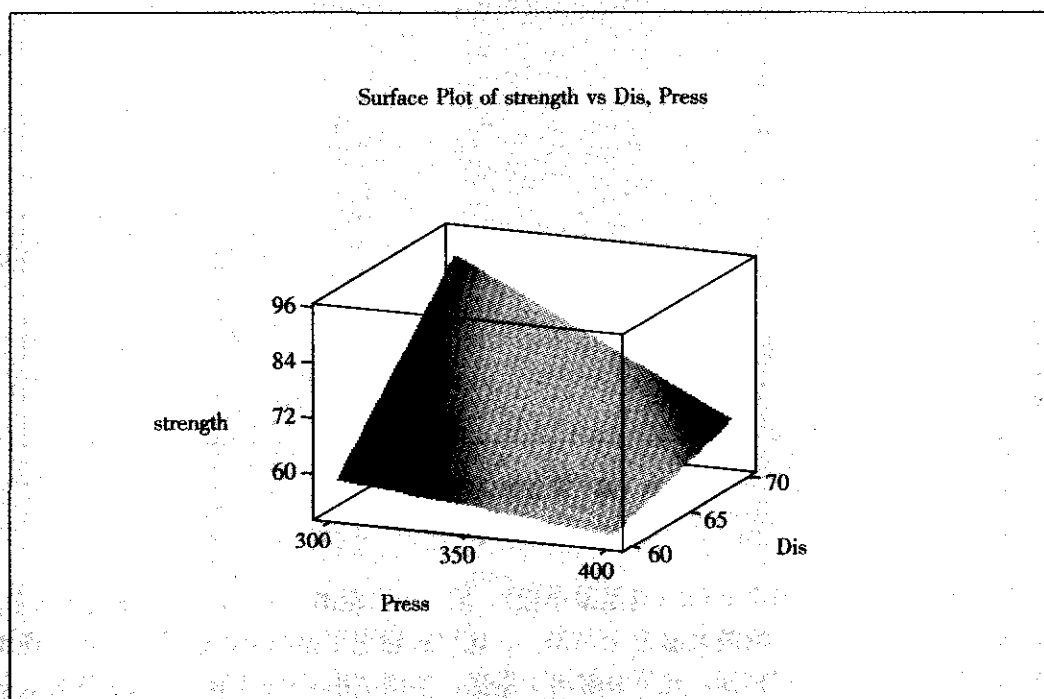


图 7—28 响应变量曲面图

从这两张图中可以看出，因子 Press 及因子 Dis 主效应及其交互作用对于响应变量 Strength 的影响确实是很显著的（等高线很弯曲，曲面偏离平面很严重），为使 Strength 取值更大，应该让 Press 尽可能小，让 Dis 尽可能大。

（3）实现最优化。根据本具体问题要求，优化目标是属于“望大”型的。在 MINITAB 软件中可以从 STAT-DOE-Factorial-Response Optimizer 窗口中得到它们。这时，在对最优的目标（goal）设定上，取“极大化”（maximize），在设定（setup）中，只需填写“下端”（lower）及“目标”（target）两项，而“上端”（upper）留为空白即可。我们取 lower=80（这个值是在做过的试验中已经实现了的），取 target=100（这个值是在做过的试验中未能达到的，是我们的较高理想）。我们通常要把求响应变量 Y 达到最优的问题转化为求一个渴求函数（desirability function）的合意度 $d(0 \leq d \leq 1)$ 达到最大的问题。这个渴求函数的具体设置是计算机根据“望大”、“望小”或“望目”三种不同情况自动设置的，我们不必过问其细节。计算机自动搜索后，得到最大值计算结果（见图 7—29）。

New D 0.584 17	Hi Cur Lo	Press 400.0 [300.0] 300.0	Dis 70.0 [70.0] 60.0
strength Maximum y = 91.683 3 d = 0.584 17			

图 7—29 响应变量优化器输出结果图

当因子 A（压力）取 300 Pa（这是最小值），因子 B（距离）取 70 mm（这是最大值）时，Strength 将会达到平均最大值为 91.683 3。这里还输出了指标 d ，这是渴求函数所取得的最大值。越接近设定目标，此值就越向 1 靠拢，本例结果 $d=0.584 1$ ，这是按最大值理想是 100 而言的。最大值的设定不相同时，完全相同的最佳结果也会使 d 值不同，对此使用者不必关心它。

大家可以看见，在试验过程中，第 7 号试验，取 Press=300，Dis=70，Angle=24

时,曾经达到过 $\text{Strength}=94.8$,为什么最后计算出的最优值还不如已经得到的结果呢?这是因为,试验总是会有误差的,第7号试验的结果有偶然性,它是偏大的。仔细分析,还可以看见另一个结果:第3号试验,取 $\text{Press}=300$, $\text{Dis}=70$, $\text{Angle}=20$ 时,却只使得 $\text{Strength}=89.1$ 。而经过分析,因子 Angle 效应并不显著,因此,这两个值都应考虑到。事实上,最优预测值,是综合了全部试验结果,仔细估计了误差状况,又考虑了显著因子的效应才计算出来的。大体上的估算可以取 94.8 与 89.1 的平均值 (91.95),此值已经与准确的最后综合预测值 $91.683\ 3$ 很接近了。

第五步:判断目标是否已经达到。

本例得到预计的最佳值为 $91.683\ 3$,应将它与试验目标相比较。如果认为离目标尚远,则需考虑安排新一轮试验。例如,可以考虑以本次获得的最佳点 $\text{Press}=300$, $\text{Dis}=70$ 为中心,在其附近重新选定试验的各因子水平,继续作试验。由于最后只有2个因子,而且根据实际经验,最优点就在附近,因此最好选择使用响应曲面方法(RSM),其试验次数虽稍多(2因子要13次,3因子要20次),但结果要精细得多,这样可以获得更好的成果。关于响应曲面方法,本章第5节有详细介绍。

如果认为强度达到 91.68 已基本达到目标,则可以结束试验,但此时不要忘记做验证试验,以确保将来按最佳条件生产能获得预期效果。在最佳点 $\text{Press}=300$, $\text{Dis}=70$ 处做若干次验证试验(次数记为 m ,通常要3次以上)。现在的任务就是计算出,将来的每一次试验结果应该落在什么范围内,进一步计算出 m 次验证试验结果的平均值应落入什么范围内,如果 m 次验证试验结果的平均值落入事先计算好的范围内,则说明一切正常,模型正确,预测结果可信。

MINITAB 14 版以上的软件在 DOE 分析中可以很方便地给出预测值和预测区间,只要选 STAT-DOE-Factorial-Analyze Factorial Design 后,点击 Prediction 选项,在 Factor 下的空栏依次输入优化的主因子的取值,就可以在 Session 报告中得到预测值和预测区间。

由于我们讲述的试验设计本身就是回归设计,另外一种方法是采用回归分析(regression)。由于最终选定的模型中包含 Press , Dis 和 $\text{Press} * \text{Dis}$ 共三项,而原始数据表中没有交叉乘积项,为此,先要将此项作为数据表中新的一列,然后进行回归分析计算(STAT-Regression-Regression),在其内 Options 选项内,严格按选定模型中自变量顺序 Press , Dis 和 $\text{Press} * \text{Dis}$,填写上最优设置的数值 300 , 70 , $21\ 000$,并选中输出所有补充计算的结果:预测值(fits即 \hat{y}),预测值处的标准差(SEs of fits),预测值处的回归结果的置信区间(confidence limit,输出时缩写为 $95\% \text{ CI}$),预测值处的单个观测值的置信区间(prediction limit,输出时缩写为 $95\% \text{ PI}$),则计算机将自动给出上述数值,结果如下:

Predicted Values for New Observations				
New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	91.68	3.42	(83.79, 99.57)	(77.61, 105.76)
Values of Predictors for New Observations				
New Obs	Press	Dis	PD	
1	300	70.0	21 000	



前一个置信区间 (95.0% CI) 结果为 (83.79, 99.57), 表明的是回归方程上的点的置信区间。它的含义是: 按照此自变量的设置, 无限多次地重复运行下去将会获得的理论均值的 95% 置信区间。后一个置信区间 (95.0% PI) 结果为 (77.61, 105.76), 表明的是单个观测值的置信区间。

要想计算 m 个观测值的平均值的 95% 的置信区间, 其计算公式是:

$$\hat{y} \pm t_{1-\frac{\alpha}{2}}(n-p) \sqrt{(\text{SE of Fits})^2 + \frac{\text{MSE}}{m}}$$

式中, n 是试验总次数; p 是最终模型中所包含的项数 (常数项要计算在内); m 是验证试验的次数。本例中, $n=12$, $p=4$, $m=3$, $\text{SE of Fit}=3.42$ (见刚才的回归计算结果), $\text{MSE}=25.536$ (见 ANOVA 表)。查 8 个自由度的 t 分位数表可得

$$t_{1-\alpha/2}(n-p) = t_{0.975}(8) = 2.306$$

变动半径的公式如下所示:

$$\begin{aligned} \delta &= t_{1-\alpha/2}(n-p) \sqrt{(\text{SE of Fits})^2 + \frac{\text{MSE}}{m}} = 2.306 \times \sqrt{3.42^2 + \frac{25.536}{3}} \\ &= 10.36 \end{aligned}$$

因此可知, 三个观测值的平均值的 95% 的置信区间为:

$$\hat{y} \pm \delta = 91.68 \pm 10.36 = (81.32, 102.04)$$

7.4 部分因子试验

7.4.1 部分因子试验概论

首先介绍进行部分实施的因子试验的必要性。

进行二水平全因子试验设计时, 大家都知道, 全因子试验的总试验次数将随因子个数的增加而急剧增加。例如, 5 个因子需要 32 次试验, 8 个因子就需要 256 次试验。但仔细分析所获得的结果可以看出, 所建立的回归方程包括哪些项呢? 除常数项外, 估计的主效应项有 8 项, 二阶交互项 28 项, 三阶交互项 56 项, ..., 八阶交互项 1 项, 详细结果列成下表 (见表 7-13)。

表 7-13

全因子试验系数分布表 (8 因子)

项别	常数	1	2	3	4	5	6	7	8
项数	1	8	28	56	70	56	28	8	1

容易看出, 回归方程中除了常数、一阶及二阶项外, 共有 219 项是三阶及三阶以上的交互作用项, 而这些项实际上已无具体的物理意义了。这自然会提出一个问题: 能不能少做些试验, 但又照样能估计方程中的常数、一阶及二阶项系数呢? 如果能够这样, 那就有很重要的应用价值了。部分实施的因子试验就是使用这种方法, 它可以用在因子个数较多 (例如 5 个以上), 只需要分析各因子主效应和二阶交互效应是否显著而并不需要考虑高阶交互效应时使用。

下面,用简单例子说明部分实施的因子试验的方法原理。

有 A, B, C, D 共 4 个可控的试验因子,每个因子都为二水平。如何能在 8 次试验中分析出每个因子的主效应?

方案 1: 删节试验方法。设想,4 个试验因子,每个因子都为二水平,作全因子试验设计要 16 次。从这 16 次试验中,选出 8 次来做,希望照样能分析主效应,是否可行?如何做?

先列出全因子设计的 16 次试验的计划表(见表 7—14)。

表 7—14 4 因子全因子试验计划表

	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD
1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1
2	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1
3	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	-1	1	-1
4	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1
5	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1
6	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1
7	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1
8	1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	-1	-1
9	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	1	-1
10	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1
11	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
12	1	1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	-1	-1
13	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1
14	1	-1	1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	-1
15	-1	1	1	1	-1	-1	-1	1	1	1	-1	-1	-1	1	-1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

先想:随机挑 8 次,行吗?显然不行。因为很可能挑成这样:某些因子在 8 次试验中,取高水平及低水平次数并不恰好相等。这时,原来正交试验的“均衡分散,整齐可比”的优点不复存在,分析方法要另搞一套才行,这当然是我们不希望看到的。我们知道,在上述正交表中,任何一列都与另外一列“正交”,因此,固定将某列(比如最后一列“ABCD”)取“1”的 8 行予以保留,而删去取“-1”的 8 行,这样可以保证,保留的 8 行表中, A, B, C, D 这 4 列中皆有 4 行取“1”,4 行取“-1”,且各列间仍然保持“均衡分散,整齐可比”,即保持了正交性。设定取 $ABCD=1$,结果见表 7—15。

表 7—15 减半实施的 4 因子全因子试验计划表 ($ABCD=1$)

	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD
1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1
4	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1
6	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1
7	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1



续前表

	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD
10	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1
11	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
13	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

仔细分析可以发现,原来的 16 行的正交表 7—14 中,15 列是完全不同的。但删去 8 行后(见表 7—15),除去一列全为 1 外,另 14 列中,每列都有与之成对的另一列是完全相同的。例如, $ABCD=1$ 的表中, D 与 ABC 就完全相同,记为 $D=ABC$ 。完全相同的两列,在作分析时,计算出效应或计算出的回归系数结果就完全相同,这两列的效应就被称作“混杂”(confounded)。也可以换个说法:这时, D 与 ABC 互为别名 (D is the alias of ABC)。

是否可以选别的条件作为删去 8 行的标准呢?例如,选 $BCD=1$,也是可以的。这时,由于 $D=BC$,显然不如 $D=ABC$ 好。经比较后,可知 $ABCD=1$ 的这种安排方法是所有安排中效果最好的。

混杂总是坏事,能否不产生混杂呢?答案是:任何部分实施的因子试验,混杂是不可避免的。我们只是希望混杂安排得更好些,尽量让感兴趣的因子或交互作用只与更高阶的交互作用相混杂,而在通常情况下,高阶的交互作用项可以忽略不计,这时,我们感兴趣的因子或交互作用就都可以估计了。

方案 2: 增补因子方法。设想,总计 8 次试验,由于每个因子都为二水平,做全因子试验可以安排 3 个因子,假设 A, B, C 这 3 个因子已安排在前 3 列了,现在有 4 个因子要安排,如何安排这新的第 4 个因子呢?如何做才好?此问题的描述见表 7—16。

表 7—16

3 因子全因子试验计划表

	A	B	C	AB	AC	BC	ABC	D
1	-1	-1	-1	1	1	1	-1	?
2	1	-1	-1	-1	-1	1	1	?
3	-1	1	-1	-1	1	-1	1	?
4	1	1	-1	1	-1	-1	-1	?
5	-1	-1	1	1	-1	-1	1	?
6	1	-1	1	-1	1	-1	-1	?
7	-1	1	1	-1	-1	1	-1	?
8	1	1	1	1	1	1	1	?

原来的这个表有 8 行 7 列,任意两列间是相互正交的。希望增加一列来安排因子 D ,而且希望此列仍然能与前面各列保持正交性。能否找出一个与前 7 列不同的列而与前 3 列保持正交呢?数学上可以证明,这是不可能的。换言之, D 这列必然要与前面第 4, 5, 6, 7 列中某列完全相同。权衡之下,我们认为,取 $D=ABC$ 是最好的安排。将 D 取值设定与 ABC 列相同,并将其前移至第 4 列,可以得到下表(见表 7—17)。

表 7—17 减半实施的 4 因子试验计划表

	A	B	C	D	AB	AC	BC	D=ABC
1	-1	-1	-1	-1	1	1	1	-1
2	1	-1	-1	1	-1	-1	1	1
3	-1	1	-1	1	-1	1	-1	1
4	1	1	-1	-1	1	-1	-1	-1
5	-1	-1	1	1	1	-1	-1	1
6	1	-1	1	-1	-1	1	-1	-1
7	-1	1	1	-1	-1	-1	1	-1
8	1	1	1	1	1	1	1	1

将 $ABCD=1$ 的表 7—15 重新排序后, 与表 7—17 完全相同。

很明显, $ABCD=1$ 这个约定非常重要。这将导致 $A=BCD$, $B=ACD$, $C=ABD$, $AB=CD$, $AC=BD$ 及 $AD=BC$, 即某些主效应将与三阶交互效应混杂, 某些二阶交互效应将与另一些二阶交互效应混杂。从上述结果很容易发现混杂的规律。可以用下列法则来表述: 在 $ABCD=1$ 这个等式两边都乘以 A , 由于 $AA=1$, 因此就能得到: $BCD=A$ 。同样, 也可以得到 $D=ABC$ 及 $AB=CD$ 等。这种运算规则简单易用, 相当于任何字母在等式两侧可以随意移动, 只要不“无中生有”就行。这个法则对二水平的部分因子试验设计总是成立的。

如果约定改为 $ABC=1$, 这将导致 $A=BC$, $B=AC$, $C=AB$, 某些主效应将与某些二阶交互效应混杂, 这个设计 $ABC=1$ 将比上个设计 $ABCD=1$ 的混杂情况要差很多。在部分因子试验中, 想使混杂情况尽可能地好且照顾到自己问题的实际需求, 是要用到较多的试验设计知识和技巧的。现在, 计算机已能自动提供最佳的设计, 因此, 一般使用试验设计的工程师不必掌握这些烦琐的内容。但有关的概念很重要, 即使直接使用计算机也要理解好概念才行。

我们先认识一个记号: 2^{k-p} 。这里, k 代表全部因子的个数, p 代表新安排的因子个数。例如前面的例子, 4 个因子安排在 8 次试验中, 此试验应记作 2^{4-1} , 这里 4 代表全部因子的个数, 如果作全因子试验, 8 次试验只能安排 3 个因子, 原有因子个数为 3, 因此新安排的因子个数为 1。当然也可以这样解释: $2^{4-1} = \frac{2^4}{2}$, 分子为因子个数是 4 时的全因子试验次数 16, 分母为“2”的 1 次方, 表示只实施了 1/2。两种理解都有意义。因此, 2^{k-p} 是部分因子试验设计的很恰当的记号。

我们称 $D=ABC$ 为“生成元”(generator), 称 $ABCD=1$ (或写为 $I=ABCD$) 为“定义关系”(defining relation), 简称“字”(word)。一旦给定了全部生成元 (2^{k-p} 中共有 p 个生成元), 则此部分因子试验设计就完全确定了。接着, 由它可以形成全部的字 (2^{k-p} 中全部的字共有 2^p-1 个)。每个字都有其“字长”(wordlength), 例如, $I=ABCD$ 的字长为 4。我们称所有的字中字长最短的那个字的长度为整个设计的分辨率(resolution)。例如, 2^{4-1} 设计中, 若定义关系为 $I=ABCD$, 则这是一个分辨度为 IV 的设计 (通常分辨率用罗马数字给出), 我们记这个设计为 2^{4-1}_{IV} , 右下角是分辨率。一般地, 分辨度为



R 的部分实施因子设计记为 $2k^p$ 。

我们有必要再详细解释一下分辨度的含义。

分辨度为Ⅲ的设计：各主效应间没有混杂，但某些主效应可能与某些二阶交互效应相混杂。

分辨度为Ⅳ的设计：各主效应间没有混杂，主效应与二阶交互效应间也没有混杂，但主效应可能与某些三阶交互效应相混杂，某些二阶交互效应可能与其他二阶交互效应相混杂。

分辨度为Ⅴ的设计：某些主效应可能与某些四阶交互效应相混杂，但不会与三阶或更低阶交互效应混杂；某些二阶交互效应可能与三阶交互效应相混杂，但各二阶交互效应之间没有混杂。

以下依此类推，但常用的到分辨度为Ⅴ的设计就可以了。

如果能够认为三阶以上的交互作用全都可以忽略不计，与三阶及三阶以上的交互作用相混杂的因子效应或交互效应则称为是可以估计（estimable）的，也称是纯净的（clear）。那么，分辨度为Ⅳ的设计中各主效应都是可以估计的，那些未相互混杂的二阶交互效应也是可以估计的；分辨度为Ⅴ的设计中，全部主效应及二阶交互效应都是可以估计的。

怎样才能根据 k 和 p 的数值确定分辨度的数值呢？这是个很难给出一般结论的问题，也没有简单的公式可用，只能查下列这张表（见表 7—18），这是进行部分实施的因子试验必须熟悉的一张表。

表 7—18 中，第一行代表试验因子的个数，第一列代表试验的总次数（不含中心点），表中所列出的数值就是最佳设计所具有的分辨度值。例如，第一行选因子个数为 4，第一列选试验总次数 8（安排全因子试验可以安排 3 个因子），这就是部分因子设计 2^{4-1} 。表中所列的就是最佳设计所具有的分辨度值Ⅳ，也就是说我们可以使用设计 $2k^{p-1}$ 。至于这里如何选定生成元，全部的定义关系（字）是什么，哪些效应间会产生混杂，这些都将由计算机自动给出，详细使用方法见 7.4.2 节。除非设计者有特别要求，由计算机所给定的设计通常是合用的、最好的。

表 7—18 部分实施的因子试验分辨度表

	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4	Full	Ⅲ												
8		Full	Ⅳ	Ⅲ	Ⅲ	Ⅲ								
16			Full	Ⅴ	Ⅳ	Ⅳ	Ⅳ	Ⅲ	Ⅲ	Ⅲ	Ⅲ	Ⅲ	Ⅲ	Ⅲ
32				Full	Ⅵ	Ⅳ	Ⅳ	Ⅳ	Ⅳ	Ⅳ	Ⅳ	Ⅳ	Ⅳ	Ⅳ
64					Full	Ⅶ	Ⅴ	Ⅳ	Ⅳ	Ⅳ	Ⅳ	Ⅳ	Ⅳ	Ⅳ
128						Full	Ⅷ	Ⅵ	Ⅴ	Ⅴ	Ⅳ	Ⅳ	Ⅳ	Ⅳ

表 7—18 还可以有更多的用途。例如，要考察 8 个因子，做多少次试验可以保证分辨度不低于Ⅳ？从表中对应因子数为 8 的那列，可以看到，做 16 次或 32 次试验分辨度都是Ⅳ，做 64 次试验分辨度才能达到Ⅴ。这样一来，就可以决定，做 16 次试验就够了。再例如，如果条件限定最多做 16 次试验，如果保证分辨度不低于Ⅳ，最多可以安排多少因子？从表中对应试验次数为 16 的那行，可以看到，安排 6，7，8 个因子都可以使分辨度为Ⅳ，

故最多可以安排 8 个因子。

7.4.2 部分实施因子试验的计划

我们通过一个例题来说明如何安排部分实施因子试验的计划。

例 7—5

陶粒混凝土工艺条件试验

在陶粒混凝土工艺条件试验中，考察影响其抗压强度的因子，并求出使抗压强度达到最大的试验条件。

共考察 6 个因子：

A 因子：水泥用量 (Amt. Cement)。低水平：180 (kg)；高水平：200 (kg)

B 因子：水泥标号 (St. Cement)。低水平：400 (号)；高水平：500 (号)

C 因子：陶粒用量 (Ceramsite)。低水平：150 (kg)；高水平：170 (kg)

D 因子：含砂率 (Sand)。低水平：38%；高水平：40%

E 因子：搅拌时间 (Mix)。低水平：1.5 分钟；高水平：2 分钟

F 因子：养护时间 (Curing)。低水平：2 天；高水平：3 天

要求是：连中心点在内，不得超过 20 次试验，考察各主效应和 AB, AC, CF, DE 各二阶交互效应是否显著。

解：由于试验次数的限制，我们在因子点上只能做试验 16 次，另 4 次取中心点。这就是 $2^{6-2} + 4$ 试验。由表 7—18 可以查得，这时 $R=IV$ ，可达分辨度为 IV 的设计。各主效应间没有混杂，主效应与二阶交互效应间也没有混杂，但某些二阶交互效应可能与其他二阶交互效应混杂，因此，只要保证所要考察的 AB, AC, CF, DE 各二阶交互效应间没有混杂就行了。

下面具体生成试验设计表。以使用 MINITAB 为例，从 Stat-DOE-Factorial-Create Factorial Design 入口即可。先选定因子个数为 6，从对话框 Design 中选定 16 次试验，4 个中心点，从对话框 Factor 中，设定各因子名称及水平，在对话框 Options 中先选定非随机化，则可生成下表（见表 7—19）。

表 7—19

例 7—5 部分实施因子设计计划表

StdOrder	RunOrder	CenterPt	Blocks	Amt Cem	St Cem	Ceramsite	Sand	Mix	Curing
1	1	1	1	180	400	150	38	1.50	2.0
2	2	1	1	200	400	150	38	2.00	2.0
3	3	1	1	180	500	150	38	2.00	3.0
4	4	1	1	200	500	150	38	1.50	3.0
5	5	1	1	180	400	170	38	2.00	3.0
6	6	1	1	200	400	170	38	1.50	3.0
7	7	1	1	180	500	170	38	1.50	2.0
8	8	1	1	200	500	170	38	2.00	2.0
9	9	1	1	180	400	150	40	1.50	3.0

续前表

StdOrder	RunOrder	CenterPt	Blocks	Amt Cem	St Cem	Ceramsite	Sand	Mix	Curing
10	10	1	1	200	400	150	40	2.00	3.0
11	11	1	1	180	500	150	40	2.00	2.0
12	12	1	1	200	500	150	40	1.50	2.0
13	13	1	1	180	400	170	40	2.00	2.0
14	14	1	1	200	400	170	40	1.50	2.0
15	15	1	1	180	500	170	40	1.50	3.0
16	16	1	1	200	500	170	40	2.00	3.0
17	17	0	1	190	450	160	39	1.75	2.5
18	18	0	1	190	450	160	39	1.75	2.5
19	19	0	1	190	450	160	39	1.75	2.5
20	20	0	1	190	450	160	39	1.75	2.5

在运行窗中，可以看到下列混杂情况表：

Factorial Design

Fractional Factorial Design

Factors: 6 Base Design: 6, 16 Resolution: IV

Runs: 20 Replicates: 1 Fraction: 1/4

Blocks: none Center pts (total): 4

Design Generators: E=ABC F=BCD

Alias Structure

I+ABCE+ADEF+BCDF

A+BCE+DEF+ABCDF

B+ACE+CDF+ABDEF

C+ABE+BDF+ACDEF

D+AEF+BCF+ABCDE

E+ABC+ADF+BCDEF

F+ADE+BCD+ABCEF

AB+CE+ACDF+BDEF

AC+BE+ABDF+CDEF

AD+EF+ABCF+BCDE

AE+BC+DF+ABCDEF

AF+DE+ABCD+BCEF

BD+CF+ABEF+ACDE

BF+CD+ABDE+ACEF

ABD+ACF+BEF+CDE

ABF+ACD+BDE+CEF

此表告诉我们,生成元(Design Generators)是: $E=ABC$, $F=BCD$ 。后面的 Alias Structure 中列出的是“别名结构”,描述的是每列中互为别名的有哪些。例如, $A+BCE+DEF+ABCD F$ 就说明: A 所在的列,其实不仅 A 在起作用,同时起作用的还有 BCE , DEF 和 $ABCD F$ 。有时也可将此式写成 $A=BCE=DEF=ABCD F$,表明这些项是互为别名的,或说它们是相互混杂的。要检查一下我们所感兴趣的二阶交互作用的混杂情况。将三阶及三阶以上交互作用忽略不计,这里混杂的有: $AB=CE$, $AC=BE$, $AD=EF$, $AF=DE$, $AE=BC=DF$, $BD=CF$, $BF=CD$ 。本问题所要求估计的4个二阶交互作用 AB , AC , CF , DE 之间,没有发生混杂,因此本设计方案是可行的。万一出现“所要求的可以估计的二阶交互作用间在给定设计中出现混杂”,那么有两种情况可能出现。一种是可以自行解决的,其办法有两条:(1)将因子名称相互交换;(2)自行选定设计生成元(这时就要改变计算机给出的缺省的生成元,自己另行指定)。这些要用到较多的技巧,可以向统计学家咨询。也有另一种可能,那就是该问题所要求的那么多二阶交互作用都不能与别的项混杂是根本办不到的。如何区分这两种情况,也只能向统计学家咨询。当然,不管遇到哪种情况,只要增加试验次数总是可以解决的。

上述试验计划表给定后,要将 RunOrder 列随机化,然后再按随机化后的 RunOrder 的数值顺序将表重新排好形成计划矩阵(planning matrix),就可以按此表进行试验了。

7.4.3 部分实施因子试验的分析实例

部分实施因子试验设计也是试验设计的一种,其分析方法当然应该与一般的试验设计的分析步骤相同,其流程图已列在7.3.3节。

由于部分实施因子试验分析方法与全因子试验分析方法基本相同,这里就不系统解释了,以下给出一个例子。

图 7-16 降低硫代硫酸钠产品的杂质率问题

降低硫代硫酸钠产品的杂质率问题

在硫代硫酸钠生产中,经过头脑风暴发现,影响其杂质率的原因有很多,至少有4个因子:成分 A 含量、成分 B 含量、反应罐内温度及反应时间,而且成分 A 与成分 B 在反应中根本不相遇,因而可以认为 AB 间无交互作用。由于试验成本很高,研究经费只够安排11次试验。

A: 成分 A 含量。低水平取12%,高水平取16%。

B: 成分 B 含量。低水平取2.4%,高水平取2.8%。

C: 反应罐内温度。低水平取200度,高水平取220度。

D: 反应时间。低水平取40分钟,高水平取50分钟。

由于试验次数的限制,本例只能采用 $2^{4-1}+3$ 设计。由部分实施因子试验分辨率表(见表7-18)中可以看出,8次试验(不包含中心点)可以实现 $Resolution=IV$ 的计划。这时候,计算机自动取生成元 $D=ABC$,这时, $AB=CD$, $AC=BD$, $AD=BC$,总计将有3对交互作用相混杂。

安排试验计划后,严格按随机化后的试验顺序进行了试验,其结果列成下表(见表7—20),其中杂质率(impurity)的单位是万分之一。

表 7—20 降低产品杂质率试验结果表

StdOrder	RunOrder	Center	Block	A	B	temp	time	impurity
1	9	1	1	12	2.4	200	40	26.9
2	10	1	1	16	2.4	200	50	27.4
3	3	1	1	12	2.8	200	50	16.9
4	7	1	1	16	2.8	200	40	38.2
5	11	1	1	12	2.4	220	50	28.3
6	6	1	1	16	2.4	220	40	39.0
7	4	1	1	12	2.8	220	40	26.8
8	2	1	1	16	2.8	220	50	36.2
9	1	0	1	14	2.6	210	45	32.1
10	5	0	1	14	2.6	210	45	28.7
11	8	0	1	14	2.6	210	45	29.8

对于部分实施因子试验的数据进行分析,其方法与全因子试验设计的五步分析法完全相同。其具体步骤及结果如下:

第一步:拟合选定模型。

在 MINITAB 软件中同样从 STAT-DOE-Factorial-Analyze Factorial Design 窗口中选定全模型后,在运行窗中可得到计算结果。

首先将全部备选项列入模型。这时,全模型也只有因子 A:成分 A 含量(A);因子 B:成分 B 含量(B);因子 C:反应罐内温度(temp);因子 D:反应时间(time),及 AB, AC, AD 共 7 项。注意,另三项交互作用项 BC, BD, CD 并未列入,这是因为现在是部分实施,效应是有混杂的,在列选项时,只能将不混杂的交互作用项放入模型中。此处最好把它们理解为是“代表”而已。例如,这里列出的 AB,实际上应该是 AB+CD,原因是“由于 AB 与 CD 是混杂的,本列的效应实际是两项效应之和”。表面上写的是 AB,但它也可能是 CD 的作用。因此,如果这一项效应是显著的,我们就无法分清 AB, CD 二者中究竟是哪个真正起了作用,这就是混杂。当然,如果这一项效应是不显著的,我们就可以断言这二者都没有显著作用。

计算结果的分析与全因子试验设计相同,现列出如下:

Factorial Fit: impurity versus A, B, C, D

Estimated Effects and Coefficients for impurity (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		30.027	0.431 5	69.60	0.000
A	10.475	5.237	0.505 9	10.35	0.002
B	-0.875	-0.437	0.505 9	-0.86	0.451
C	5.225	2.613	0.505 9	5.16	0.014
D	-5.525	-2.762	0.505 9	-5.46	0.012
A * B	4.875	2.438	0.505 9	4.82	0.017
A * C	-0.425	-0.212	0.505 9	-0.42	0.703
A * D	-1.275	-0.638	0.505 9	-1.26	0.297

S=1.430 97 R-Sq=98.44% R-Sq (adj) =94.80%

Analysis of Variance for impurity (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	336.635	336.635	84.158 7	41.10	0.006
2-Way Interactions	3	51.144	51.144	17.047 9	8.33	0.058
Residual Error	3	6.143	6.143	2.047 7		
Curvature	1	0.123	0.123	0.123 1	0.04	0.858
Pure Error	2	6.020	6.020	3.010 0		
Total	10	393.922				

Effects Pareto for impurity

Alias Structure

I+A*B*C*D

A+B*C*D

B+A*C*D

C+A*B*D

D+A*B*C

A*B+C*D

A*C+B*D

A*D+B*C

MINITAB计算中输出了因子 Pareto 效应图 (见图 7—30)。

(1) 先看 ANOVA 表中的总效果。在本例中, 对应主效应项 (main effects) 的 p -value 为 0.006, 表明本模型总的说来是有效的。

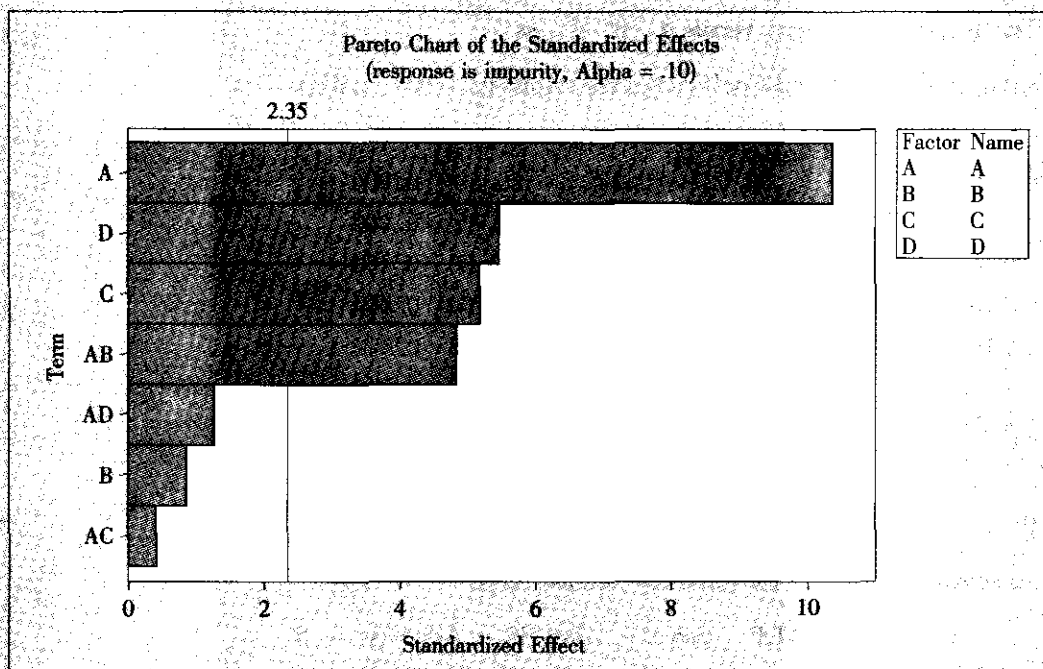


图 7—30 降低产品杂质率试验的因子 Pareto 效应图

(2) 看 ANOVA 表中的弯曲项 (curvature)。在本例中, 对应弯曲项 (curvature) 的 p -value 为 0.858, 大于 0.05, 则表明本批数据并没有弯曲现象。

(3) 各项效应的显著性。从计算结果的最开始参数估计项中可以看出, 因子 A, C, D 及 AB 是高度显著的 (参见图 7—30), 而其余项皆不显著。仔细查看本设计的“别名结构表” (alias structure) 可以看出混杂的状况。AB 所在列的效应其实是 $AB + CD$, 即 AB 与 CD 混杂, 而由问题的背景材料分析中已经明确 AB 交互作用并不存在, 因此, 这个显著项只能是 CD 的作用。这样一来, 我们最终将选入 A, C, D 及 CD。

部分实施因子试验的数据分析与全因子试验设计的数据分析, 其差别只在这里。即当数据分析结果中有某些二阶交互作用效应显著时, 不能仅从表面上的结果来决定取舍, 要仔细分析混杂结构, 查看在结构表中, 此显著项是与哪个 (或哪些) 二阶交互作用效应相混杂的, 再根据背景材料予以判断, 最终决定谁入选。有时没有相关背景材料提供, 这时判断确实可能有困难, 只好再做进一步的试验来区分这些混杂的效应。有时实在没办法, 还可以“猜”一下。例如, A, C 和 D 这三个因子主效应显著, 而 $AB = CD$, 这时候很容易想到: B 因子并不显著, 而 C, D 两因子皆显著, 当然, “CD 显著”的可能性要比“AB 显著”的可能性大得多, 因此可以判 CD 显著。这当然没有绝对的把握, 严格地说, 必须再做进一步的试验才能予以判别。

第二步: 进行残差诊断。

残差诊断分析法与全因子试验设计完全相同, 本例残差诊断中未发现任何问题, 此处从略。

第三步: 判断模型是否要改进。

从残差诊断中看出,模型基本上是好的,改进模型主要是删除不显著项。因此,实际上,又要返回第一步。

本例中,在重新拟合模型的计算时保留 A , C 和 D 及二阶交互作用项 CD ,再次计算所得的结果如下:

Factorial Fit: impurity versus A, C, D

Estimated Effects and Coefficients for impurity (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		30.027	0.413 5	72.61	0.000
A	10.475	5.238	0.484 9	10.80	0.000
C	5.225	2.613	0.484 9	5.39	0.002
D	-5.525	-2.762	0.484 9	-5.70	0.001
C * D	4.875	2.438	0.484 9	5.03	0.002

S=1.371 55 R-Sq=97.13% R-Sq (adj) =95.22%

Analysis of Variance for impurity (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	335.104	335.104	111.701	59.38	0.000
2-Way Interactions	1	47.531	47.531	47.531	25.27	0.002
Residual Error	6	11.287	11.287	1.881		
Curvature	1	0.123	0.123	0.123	0.06	0.824
Lack of Fit	3	5.144	5.144	1.715	0.57	0.687
Pure Error	2	6.020	6.020	3.010		
Total	10	393.922				

Estimated Coefficients for impurity using data in uncoded units

Term	Coef
Constant	424.052
A	2.618 75
C	-1.932 50
D	-10.790 0
C * D	0.048 750 0

对于这次的计算结果,仍用前面介绍过的方法与步骤来分析。

此时观察到下面几处要点:

(1) 先看 ANOVA 表中的总效果。在本例中,对应主效应项 (main effects) 的 p -value 为 0.000,表明本模型总的说来是有效的。

(2) 看 ANOVA 表中的弯曲项 (curvature)。在本例中,对应弯曲项 (curvature) 的 p -value 为 0.824,大于 0.05,表明本批数据并没有弯曲现象。

(3) 看 ANOVA 表中的失拟项 (lack of fit)。在本例中, 对应失拟项 (lack of fit) 的 p -value 为 0.687, 大于 0.05, 表明删减模型并没有失拟现象。

(4) 各项效应的显著性。从参数估计中可以看出, 因子 A, C, D 及 CD 确实是高度显著的。而且, 全模型的 $S=1.430\ 97$, 而删减模型中 $S=1.371\ 55$, 因此可以断言, 删减模型效果是好的。

第四步: 对选定模型进行分析解释。

我们要在拟合选定模型后输出更多的信息, 并给出有意义的解释。

(1) 输出各因子的主效应图和交互效应图。主效应图见图 7—31。从图中可以看出, 因子 A, C, D 对于响应变量 impurity 的影响确实是很显著的, 而因子 B 的影响确实是不显著的, 而且还可以看出, 为使 impurity 取值更小, 应该让 A, C 尽可能小, 让 D 尽可能大。

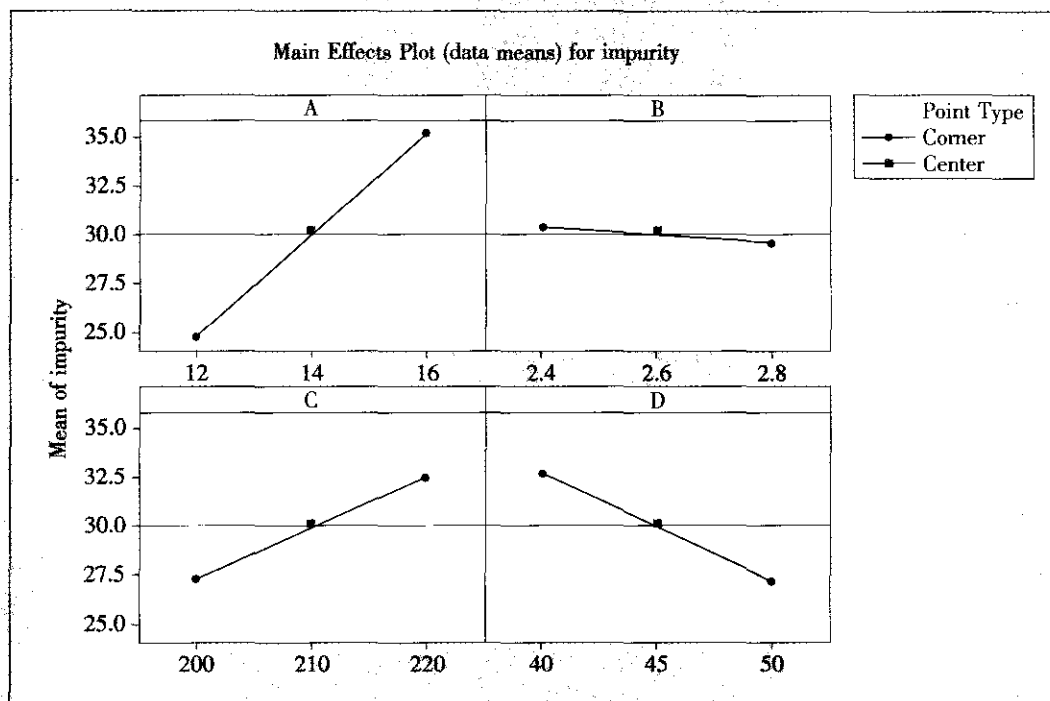


图 7—31 因子主效应图

交互效应图见图 7—32。

从图 7—32 中可以看出, 因子 C 与因子 D 的交互作用对于响应变量 impurity 的影响确实是很显著的 (两条线非常不平行), 而其他交互效应对于响应变量 impurity 的影响确实是不显著的 (两条线几乎平行)。

(2) 输出等高线图、响应曲面图等。在 MINITAB 软件中可以从 STAT-DOE-Factorial-Count Surface Plot 窗口中得到它们。我们注意到, 在选定有交互作用的 C 和 D 作为两个坐标轴时, C 为横轴的曲面图不易观察, 我们改将 C 作为纵轴, D 为横轴, 图形见图 7—33 及图 7—34。

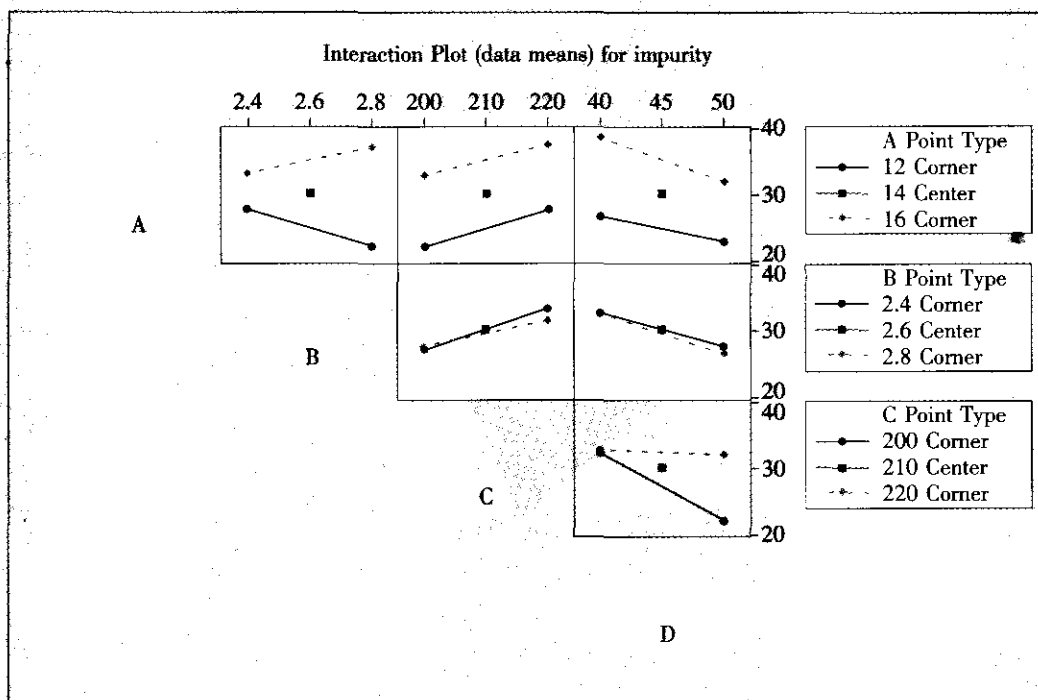


图 7—32 各因子交互效应图

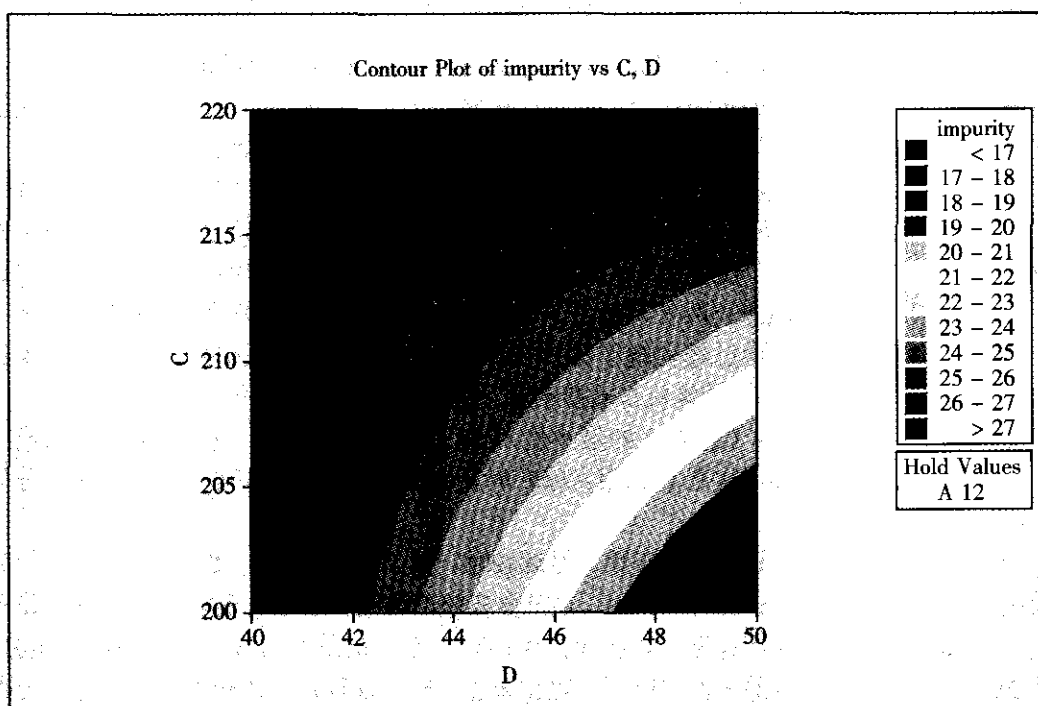


图 7—33 杂质率对于 C, D 二因子的等高线图

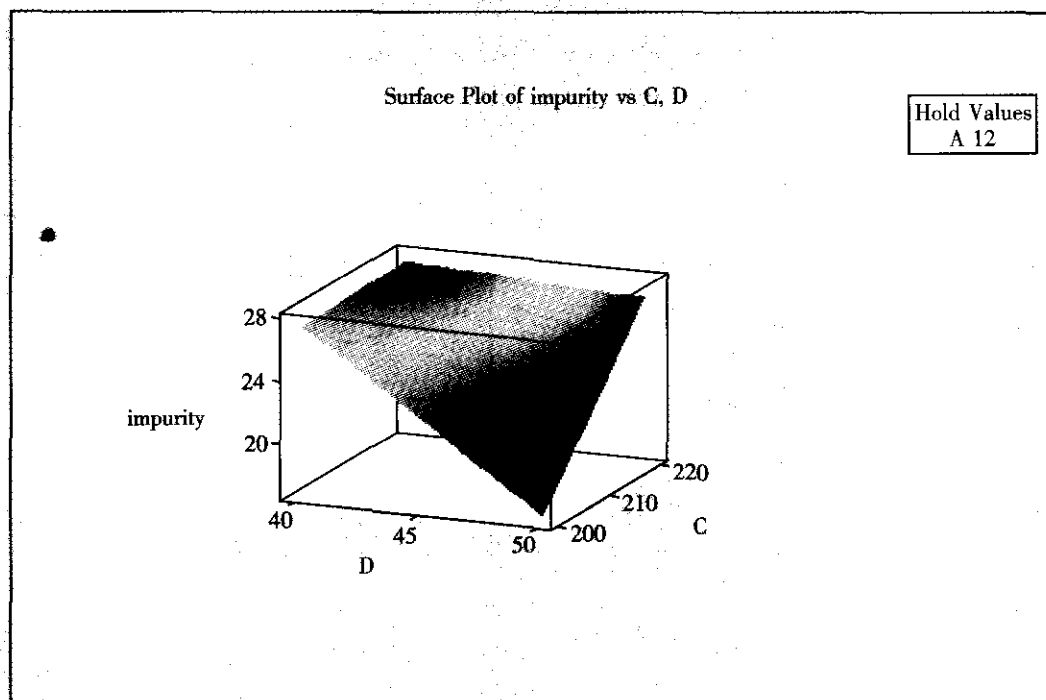


图 7—34 杂质率对于 C, D 二因子的曲面图

从图 7—33 及图 7—34 中可以看出, 交互作用 CD 对于响应变量杂质率 (impurity) 的影响确实是很显著的 (等高线很弯曲, 曲面偏离平面), 为使 impurity 取值更小, 应该让 C 尽可能小, 让 D 尽可能大。A 则由主效应图上确定, 应取尽可能小。

(3) 实现最优化。根据本具体问题, 是属于“望小”型的。在 MINITAB 软件中可以从 STAT-DOE-Factorial-Response Optimizer 窗口中得到最优化结果。这时, 在对最优的目标 (goal) 选择上, 取“极小化” (minimize), 在设定 (setup) 中, 只需填写“上端” (upper) 及“目标” (target) 两项, 而“下端” (lower) 留为空白即可。我们取 upper = 25 (这个值在做过的试验中已经实现了的), 取 target = 15 (这个值在做过的试验中未能达到, 是我们的较高理想)。计算机自动搜索后, 得到最小值计算结果, 见图 7—35。

当成分 A 含量取 12% (这是 A 的最小值), 反应罐内温度取 200 度 (这是 C 的最小值), 反应时间取 50 分钟 (这是 D 的最大值) 时, impurity 将会达到平均最小值为万分之 16.977 3。

可以看到, 在试验过程中, 进行第 3 号试验时, 在 A=12, B=2.8, C=200, D=50 的条件下, 曾经达到过 impurity=16.9。在部分实施的因子试验中, 能够正巧在最优设置做过一次试验的情况 (B 不显著, 取值无关大局) 并不一定总能出现。由于是部分实施, 有些试验条件的搭配组合并未做过试验。现在通过分析可以看到, 即使未在最优设置处做过试验, 我们的分析也确实可以保证一定能找到这个最佳设置。这正是试验设计的统计分析方法的价值所在。如果只安排试验而不进行统计分析, 是无论如何也做不到这一点的。

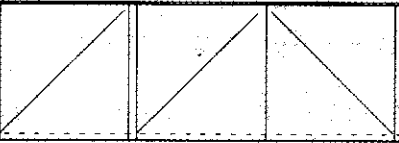
Optimal D 0.802 27	Hi Cur Lo	A 16.0 [12.0] 12.0	C 220.0 [200.0] 200.0	D 50.0 [50.0] 40.0
impurity Minimum y = 16.977 3 d = 0.802 27				

图 7—35 杂质率的最优化结果图

第五步：判断目标是否已经达到。

本来，部分实施的因子试验的目的是筛选因子，能够由各因子的主效应图和交互效应图分析出哪些因子（哪些因子间）的效应是显著大的，应予以保留，这就达到目的了。通常是以这些结果为依据，选定因子并确定因子水平，进行下一轮试验。但实际上，我们还可以更充分地挖掘信息，进一步求出最佳值，甚至可以求出最佳值的预测值的置信区间。这些计算方法与全因子的相应部分完全相同，具体细节，就不在这里详细叙述了。

借助回归分析（Regression）栏来进行计算。由于最终选定的模型中包含 A，C，D 及 CD 共 4 项，而原始数据表中没有 CD 交叉乘积项，为此，先要将 CD 项作为数据表中新的一列，然后进行回归（Regression-Regression），在其中 Options 选项内，严格按选定模型中自变量顺序 A，C，D，CD，填写上最优设置的数值 12，200，50，10 000，并选中输出所有补充计算的结果，则计算机将自动给出结果如下：

Predicted Values for New Observations				
New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	16.977	1.054	(14.397, 19.558)	(12.743, 21.211)
Values of Predictors for New Observations				
New Obs	A	C	D	CD
1	12.0	200	50.0	10 000



与全因子试验设计的相应部分完全相同,也可以计算出 m 个观测值的平均值的 95% 置信区间,此处从略。

7.4.4 Plackett-Burman 设计——另一类筛选因子设计

前面介绍的各种部分实施因子设计在一般的应用中已经足够了。它的一般理论也比较完整,易理解、易应用。但有时试验经费非常昂贵,节省每一次试验都是有价值的,这时就可以使用试验次数更少的方法。Plackett-Burman 设计就是最常见的一种试验次数最少的方法。

如前所述,部分实施因子设计的试验次数都是 2 的整方幂: 4, 8, 16, 32, …, 这些试验计划保证了试验安排的正交性,除中心点外,任一列中正负号出现次数各占一半,任两列中,++, +-, -+, -- 四种搭配出现的次数都是相等的。这种“正交性”导致试验结果“均衡分散,整齐可比”。但是,保持正交性的设计其试验次数 n 是否必须是 2 的整方幂呢? 数学上容易看出, n 确实不一定非要是 2 的整方幂,但它一定得是 4 的整倍数。Plackett-Burman 设计就是针对所有恰为 4 的整倍数的 n 的这些情况给出的设计。若此时 n 又恰为 2 的整方幂时,其结果与前面介绍的部分实施因子设计完全一样,因此没有讨论的必要。总之,Plackett-Burman 设计最有用的是 $n=12, 20, 24, \dots$ 。下面就是一个 L12 表(见表 7—21)。

表 7—21 Plackett-Burman 设计表 ($n=12$)

	A	B	C	D	E	F	G	H	J	K	L
1	1	-1	1	-1	-1	-1	1	1	1	-1	1
2	1	1	-1	1	-1	-1	-1	1	1	1	-1
3	-1	1	1	-1	1	-1	-1	-1	1	1	1
4	1	-1	1	1	-1	1	-1	-1	-1	1	1
5	1	1	-1	1	1	-1	1	-1	-1	-1	1
6	1	1	1	-1	1	1	-1	1	-1	-1	-1
7	-1	1	1	1	-1	1	1	-1	1	-1	-1
8	-1	-1	1	1	1	-1	1	1	-1	1	-1
9	-1	-1	-1	1	1	1	-1	1	1	-1	1
10	1	-1	-1	-1	1	1	1	-1	1	1	-1
11	-1	1	-1	-1	-1	1	1	1	-1	1	1
12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

用 Plackett-Burman 设计来安排试验的最大好处是节省试验次数。例如,我们用 L12 时,用 12 次试验,可以安排最多 11 个因子。当然,这种方法的分辨率只有 III,换言之,主效应将与二阶交互效应混杂。若任何一个二阶交互效应是显著的,则将导致不能准确分析主效应。这种设计不像一般的部分实施因子设计还可能提供别的更多信息,它只能用于筛选因子,因此也专称 Plackett-Burman 设计是“筛选试验设计”。它的使用范围和条件是:因子个数较多,试验昂贵,不必考虑任何交互作用。因此,不到万不得已,此方法通常是不采用的。由于其设计及分析方法与前面介绍的部分实施因子设计完全相同,这里从略。

7.4.5 三水平部分因子试验的分析

对于全因子试验设计,本节前面只考虑了二水平试验的设计与分析。如果实际工作中

确实需要考虑三水平或多水平时，可以自行定义一个全因子试验，计算机软件也有相应的设计与分析的功能。由于它与二水平的状况相仿，故不再加以叙述。但如果需要进行三水平的部分因子试验时，就要自己动手安排和计算。

高、中、低三水平的代码可以有多种选择方法，选 1, 2, 3 可以，选离散的 A, B, C 也行，但应以连续型变量 -1, 0, 1 作为代码最好。这时，我们可以把对因子效应的方差分析和对自变量的回归结合起来进行。

下面举例说明。

例 7-7

烟灰砖试验

为了环保，利用粉煤灰及煤矸石经碾压制成烟灰砖。为了考察如何能使其强度达到最大，要进行试验：将烟灰砖制成边长为 15cm 的立方体后，将其压碎，以其承载力 (kg) 为指标 y ，希望选定最优工艺条件使 y 达到最大。共有 3 个因子，选择其水平如下：

A. 成型水分 (%)：A1=9, A2=10, A3=11

B. 碾压时间 (分钟)：B1=8, B2=10, B3=12

C. 料重 (kg/盘)：C1=330, C2=360, C3=390

找出重要影响因子及最佳搭配，使承载力 y 达到最大。

下面我们以 MINITAB 软件为例说明整个操作过程。

在计划阶段，注意使用一般因子设计 (STAT-DOE-Factorial-Create Factorial Design) 入口，选中 “General Full Factorial Design”，并选定因子个数为 2 (因为只进行 9 次试验)，在对话框 “Design” 中，填写 A, B 二因子，且水平数皆为三。在对话框 “Factor” 中，填写 A, B 二因子的各水平为 -1, 0, 1。在对话框 “Options” 中，选中非随机化。这样就可以得到含 A, B 二因子各三水平的正交表。自己在表中再补充一列 C，按部分因子三水平正交表的第 3 列填写，三水平仍取 -1, 0, 1。顺序随机化后 (本例题省略此步)，就得到完整的试验计划表了。

在实际试验后，将数据 y 填写在 A, B, C 之后，结果见表 7-22。

表 7-22 烟灰砖强度的三水平试验结果表

StdOrder	RunOrder	Blocks	A	B	C	y	A2	B2	C2
1	1	1	-1	-1	-1	172	1	1	1
2	2	1	-1	0	0	182	1	0	0
3	3	1	-1	1	1	180	1	1	1
4	4	1	0	-1	0	192	0	1	0
5	5	1	0	0	1	183	0	0	1
6	6	1	0	1	-1	198	0	1	1
7	7	1	1	-1	1	159	1	1	1
8	8	1	1	0	-1	160	1	0	1
9	9	1	1	1	0	181	1	1	0

在分析时，不能使用 “DOE” 窗，而要用 STAT-ANOVA-General Linear Model 作为入口。这时，各因子都应当作 “固定效应” (因而不要填写入 Random Effect 窗内)，得到 ANOVA 结果如下：

General Linear Model: y versus A, B, C

Factor Type Levels Values

A fixed 3-1 0 1

B fixed 3-1 0 1

C fixed 3-1 0 1

Analysis of Variance for y, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
A	2	889.56	889.56	444.78	108.19	0.009
B	2	272.89	272.89	136.44	33.19	0.029
C	2	197.56	197.56	98.78	24.03	0.040
Error	2	8.22	8.22	4.11		
Total	8	1368.22				

可以得知, 3个因子的效应都是显著的(显著性水平取为0.05)。为了得到回归方程, 要先在数据表中加上3个因子的平方项, 然后再进行回归分析。这是因为, 因子取三水平后, 不单有线性效应(main effect), 而且包含二阶效应, 不加入二阶项所得到的因子效应是不完全的、不准确的。从STAT-Regression-Regression入口, 选入A, B, C及它们的平方项共6项, 结果如下:

Regression Analysis: y versus A, B, C, A2, B2, C2

The regression equation is

$$y = 194 - 5.67A + 6.00B - 1.33C - 18.7A^2 + 5.33B^2 - 9.67C^2$$

Predictor	Coef	SE Coef	T	P
Constant	193.889	1.788	108.43	0.000
A	-5.6667	0.8278	-6.85	0.021
B	6.0000	0.8278	7.25	0.019
C	-1.3333	0.8278	-1.61	0.249
A2	-18.667	1.434	-13.02	0.006
B2	5.333	1.434	3.72	0.065
C2	-9.667	1.434	-6.74	0.021

S=2.028 R-Sq=99.4% R-Sq (adj)=97.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	6	1360.00	226.67	55.14	0.018
Residual Error	2	8.22	4.11		
Total	8	1368.22			

这说明, 应取 A, B, A^2, B^2 (p -value 为 0.065, 处于边界状态, 这里也视为显著项) 及 C^2 共 5 项, 即取 2 个主因子及全部 3 个因子平方项。

由 STAT-ANOVA-Main Effect Plot 入口, 可以得到下图 (见图 7—36)。

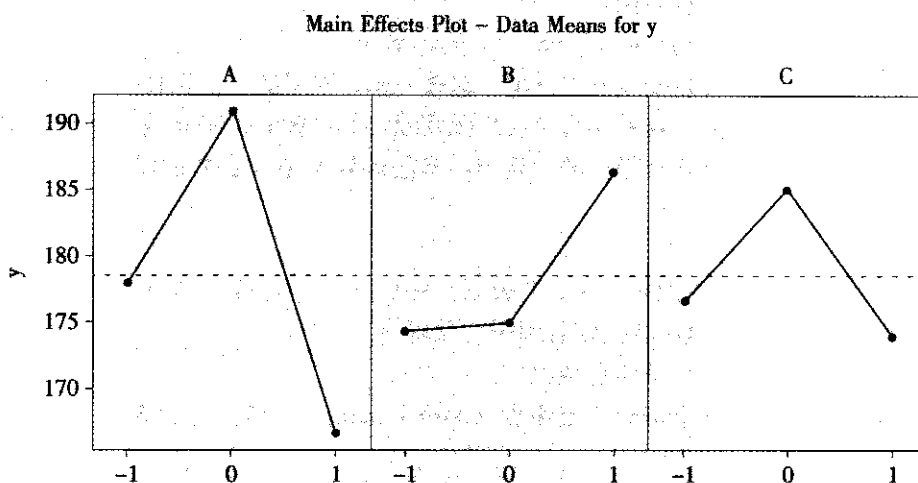


图 7—36 三水平试验主效应图

由图 7—36 可以断言, A 取中水平 0, B 取高水平 1, C 取中水平 0, 将使响应变量 y 取值最大。最后再从 STAT-Regression-Regression 入口, 选定 A, B, A^2, B^2 及 C^2 共 5 项, 以 0, 1, 0, 1, 0 代入方程, 得到最优值的预测结果。

Predicted Values for New Observations					
New Obs	Fit	SE Fit	95.0% CI	95.0% PI	
1	205.222	2.213	(198.180, 212.265)	(194.575, 215.870)	
Values of Predictors for New Observations					
New Obs	A	B	A2	B2	C2
1	0.000 000	1.00	0.000 000	1.00	0.000 000

7.5 响应曲面设计与分析

7.5.1 响应曲面设计概论

在实际工作中, 常常需要研究响应变量 y 究竟如何依赖于自变量, 进而找到自变量的设置使得响应变量得到最佳值 (望大、望小或望目)。如果自变量的个数较少 (通常不超过 3 个), 则响应曲面方法 (response surface methodology, RSM) 是最好的方法之一。本方法特别适合于响应变量望大或望小的情形, 而参数稳健设计 (robust parameter de-



sign) 方法则更适合于响应变量望目的情形。

通常的做法是这样的：先用二水平的因子试验的数据，拟合一个线性的回归方程（可以含交叉乘积项），如果发现有弯曲的趋势，则拟合一个含二次项的回归方程。其一般模型是这样的（以两个自变量为例）：

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{11} x_1^2 + b_{22} x_2^2 + b_{12} x_1 x_2 + \varepsilon$$

选项比过去增加了各自变量的平方项。这样一来，原来的一些设计点的数据就不够用了，要再增补一些试验点。这种先后分几阶段完成全部试验的策略就是“序贯试验”的策略。适用于这种策略的方法有很多种，其中最常用的就是中心复合设计（central composite design, CCD）。

1. 中心复合设计

下面的图形就是三维空间（即 3 个因子）中的一个中心复合设计示意图（见图 7—37）。以下叙述中给出的坐标都假定各因子已代码化。

整个试验由下面三部分试验点构成：

（1）立方体点（cube point）[或角点（corner point）]。各点坐标皆为 1 或 -1，这就是组成因子试验的部分。

（2）中心点（center point）。各点坐标皆为 0。

（3）星号点（star point）[或轴点（axial point）]。除一个坐标为 $\pm \alpha$ 外，其余坐标皆为 0。在 k 个因子情况下，共有 $2k$ 个星号点。

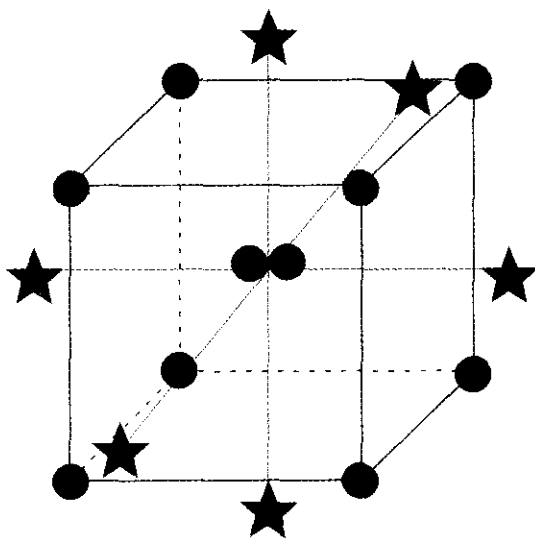


图 7—37 三因子中心复合设计布点示意图

上述立方体点和部分中心点构成本章第 3 节中介绍过的一个普通的全因子设计，星号点和另一些中心点将其扩展为二阶设计。当然，如果可以肯定试验区域已接近于最优区域，则可以不经过序贯试验，而直接进行中心复合设计，三部分试验点同时进行。这里有三个问题需要详细讨论：

- （1）如何选择全因子设计部分；
- （2）如何确定星号点的位置（即确定 α 的数值问题）；

(3) 如何确定中心点的个数问题。

在因子设计部分中,我们通常都取全因子试验的安排方法,当因子个数大于5时才考虑部分实施的因子设计,这时通常都要求设计的分辨率在V以上。

在 α 的选取上可以有多种出发点,旋转性(rotatability)是个很有意义的考虑。所谓旋转性指的是:将来在某点处预测值的方差仅与该点到试验点中心的距离有关,而与其所在方位无关,也即响应变量的预测精度在以设计的中心为球心的球面上是相同的。可以证明,这时应取:

$$\alpha = F^{\frac{1}{4}}$$

(F 为因子点试验次数,对于两因子问题, $F=4$, $\alpha=1.414$;对于三因子问题, $F=8$, $\alpha=1.182$)

对于 α 值的选取不必很精确,设计只要有近似旋转性(near rotatability)就够了。按上述公式选定的 α 值来安排CCD设计是最典型的情形,它可以实现试验的序贯性,此种试验特称为中心复合序贯设计(central composite circumscribed, CCC),它是CCD中最常用的一种。

如果要求进行CCD设计,但又希望试验水平安排不能超过立方体边界时,可以将星号点设置为1及-1,则计算机自动将原CCD缩小到整个立方体内,这种设计也称为中心复合有界设计(central composite inscribed design, CCI)。但这种设计失去了序贯性,因为前一次在立方体点上已经做过的试验结果,在后续的CCI设计中已经没有用了。

带区组的CCD设计中,对于 α 值的选取就要从另一个角度出发了。由于我们把区组也作为一个因子(它是可控的,但又难以完全随机化)来安排,我们在最开始估计各项效应显著性时,要把“区组”作为一个因子放入回归方程以减小误差,从而增加参数估计及判定的精度;但后来在作预报时,又应将“区组”这个因子从回归方程中剔除。此时我们首先考虑的是,希望在“保留”和“删除”区组因子这两种不同选择下,对其他因子效应的估计保持不变,也就是说,要保持“区组”因子与其他因子间效应估计的独立性或这两部分因子间的正交性。为此,要调整 α 值的选取标准,这时顾及不到旋转性而只能顾及正交性的要求了。具体的计算公式这里不详细叙述,计算机自动按正交性的要求对不同的因子个数计算出相应的 α 值来。

对于 α 值的选取的另一个出发点也是有意义的,就是取 $\alpha=1$ 。这就意味着将星号点的位置向中心收缩而设在立方体的表面上,这样的设计称为中心复合表面设计(central composite face-centered design, CCF)。这样做的好处是每个因子取值的水平数只有3个(-1, 0, 1),而一般的CCD,因子取值的水平数是5个(- α , -1, 0, 1, α),这在更换因子水平较困难的情况下是有意义的。另一种情况下也将有用:如果希望试验设计有序贯性,即允许先进行因子设计,其结果在二阶回归中仍然要被采用,做CCF就是有意义的,但代价是CCF失去了旋转性。其示意图见图7-38。

在中心点的个数(N_c)问题上,更是有许多细节要分析。在满足旋转性的前提下,如果适当选择 N_c 则可以使在整个试验区域内的预测值都有一致均匀精度(uniform precision)。这当然是好事,特别是在试验前,我们无法预知最优点究竟位于何处,在整个试验区域内的预测值都有一致均匀精度很重要。为此,中心点的试验要重复很多次,具体次数见表7-23。但有时认为,这样做的代价太大, N_c 其实取2以上就够了,若中心点的选取主要是为了估计试验误差,则可能需要 N_c 取4以上。

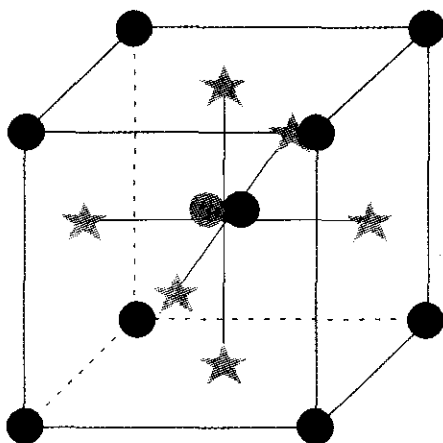


图 7—38 三因子中心复合表面设计布点示意图

下面是 Box 建议的一致均匀精度的 CCD 设计在不同因子数的情况下总计需试验点的个数表（见表 7—23）。

这里，因子数为 5 时，有两种选择，前一种在因子点上进行的是全因子试验 2^5 ，而后一种是进行的部分因子试验 2^{5-1} ，相应的中心点个数要求也不同。

总之，当时间和资源条件都容许时，进行 CCD 试验设计时尽量按照表 7—23 中所给出的试验计划去安排试验，可以达到一致均匀的精度，设计结果和推测出的最佳点都比较可信。实在需要减少试验次数时，中心点至少也要 2~5 次。当因子水平更换有困难且试验水平安排不能超过立方体边界时可采用 CCF 设计。必须要保证一致均匀精度时，只能牺牲序贯性而保持旋转性，这时可以采用 CCI 设计。这些对于具体问题的处理原则将会在下一段响应曲面设计的计划阶段体现出来。

表 7—23 CCD 设计试验点数表

因子数	立方体点	星号点	中心点	总计
2	4	4	5	13
3	8	6	6	20
4	16	8	6	30
5	32	10	10	52
5	16	10	7	33

2. Box-Behnken 设计

响应曲面设计的另一种方法就是 Box-Behnken 设计。这种设计安排是将因子各试验点取在立方体的棱的中点上。三因子的 Box-Behnken 设计取点示意图见图 7—39。

这种设计所需点数比 CCD 要少，它也具有近似旋转性，它最适应的试验区域是球形域（而不是立方体域），但其最大缺点是设计没有序贯性，上批进行过的试验结果数据几乎对下批试验没有用，每批试验都要全部重新做。因子数为 3 时，试验次数为 $12+3=15$ ；因子数为 4 时，试验次数为 $24+3=27$ 。除非极端重视试验次数，通常不采用这种设计。

下面以 3 个因子为例, 给出 4 种设计的数值表 (见表 7—24)。第 1 种是普通的 CCD, 第 2 种是 CCI, 第 3 种是 CCF, 第 4 种是 Box-Behnken 设计。前 3 种试验都要 20 次, Box-Behnken 设计需要 15 次。这些设计都假定一开始就希望完成全部响应曲面设计, 未作序贯设计的安排, 而且表中所列数值假定已代码化。

[illegible]

续前表

	CCD			CCI			CCF			BB		
	A	B	C	A	B	C	A	B	C	A	B	C
18	0	0	0	0	0	0	0	0	0			
19	0	0	0	0	0	0	0	0	0			
20	0	0	0	0	0	0	0	0	0			

7.5.2 响应曲面设计的计划

整个响应曲面设计的计划应该包含两个阶段。

一旦识别出来的因子个数较少,就可以准备进行响应曲面分析方法。第一阶段要先考虑识别目前试验区域的状况,它是否已经接近或达到了能使响应变量达到最佳值的最优区域,还是目前仍然远离最优区域?我们以望大型(望小型完全相同)问题为例来讨论,望目型问题可以留在7.6节参数稳健性设计中解决。由于通常响应变量的极大值是在有曲面的弯曲的“山顶”部分达到的,因此,我们首先要特别认真分析数据是否显示了弯曲性,或在数据拟合线性回归方程时的ANOVA表中的“失拟”现象是否显著。如果可以断言说弯曲或失拟并不严重,则说明目前试验区域的位置仍然远离最优区域。就像爬山时,如果脚下虽然倾斜但仍然是较为平坦的区域,则说明此时仍然仅在半山坡上,离山顶依然遥远。这时,首要任务是先沿最陡峭的方向爬上去,当可以肯定达到山顶区域时,再建立细致的曲面方程来描述其数学规律。因此,我们可以归结响应曲面试验的两个阶段如下:

第一阶段:用最速上升法(steepest ascent search)寻找试验的最优区域。由于这时的回归方程是线性的,其等高线是些近似的平行直线,选取与等高线垂直的方向作为“最速前进方向”(path of steepest ascent),边沿此方向前进边做试验(对每个选定的位置只做一次即可)。若响应变量数值持续增加,则继续前进,直到响应变量数值出现了下降的情况才停止。选定刚才已得到的最佳值之处作为新一批响应曲面设计的中心点,转入下面的第二阶段。

第二阶段:在已确认为最优区域的范围内进行响应曲面试验。要根据实际条件来安排CCD设计或Box-Behnken设计,在CCD设计中又有CCC(中心复合序贯设计)、CCI(中心复合有界设计)或CCF(中心复合表面设计)三种选择可能。

各计算机软件都能自动生成上述各种设计的表格,下面介绍使用MINITAB软件的主要步骤。

从STAT-DOE-Response Surface>Create Response Surface Design入口后,首先在对话框中可以选定CCD或Box-Behnken设计,并选定因子个数。在Display Available Design框中可以看到不同要求下所需要的试验次数。

在Design对话框中选定中心点试验次数及 α 值,计算机提供缺省的试验次数选择,这是在可能情况下的最佳选择,除非资源限制使试验次数实在达不到这些要求可以自行选择外,通常不要改变试验次数。在 α 值的选取上有三种选择:(1)缺省,这时计算机会按因子个数、旋转性还是正交性要求自动求出最佳的 α 值。(2)选 $\alpha=1$,即安排CCF(中心

复合表面设计)。(3) 自选 α 值。

在 Factor (因子) 对话框中, 除了要设定各因子的实际水平外, 要注意选取星号点的设置。当 Levels Define 栏中选定 “Cube Points” 时, 表示这时设定的水平作为立方体点, 即这时选择的是 CCC (中心复合序贯设计) 设计, 星号点将超出立方体。选定 “Axial Points” 时, 表示这时设定的水平作为轴上的星号点, 即这时选择的是 CCI (中心复合有界设计) 设计, 星号点将达到立方体边界而不超出, 立方体点将向内收缩。

在 Options 对话框中选定 “非随机化顺序”, 这样我们就可以看到标准化的安排方法。其中, 开始部分是因子设计点 (立方体点), 中间部分是星号点部分, 最后是中心点。真正进行试验时, 再将试验顺序随机化。这些与其他试验设计相同, 不再赘述。

7.5.3 响应曲面设计的分析及实例

响应曲面设计也是试验设计的一种, 其分析方法当然应该与一般的试验设计的分析步骤相同。其流程图如图 7—8 所述, 也分为五大步骤。

响应曲面试验的分析方法与一般的试验设计分析的差别, 就在于这里的选项。响应曲面试验设计要拟合的是二次曲面, 其一般模型是 (以两个自变量为例):

$$y = b_0 + b_1x_1 + b_2x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{12}x_1x_2 + \epsilon$$

选项比过去增加了各自变量的平方项。另外, 在计算机的输出中, 输出的表达上也有些细微的差别, 在此不多加叙述, 详细比较可以从下面例题中看出。

例 7—8

黏合剂生产条件优化问题

在黏合剂生产中, 经过因子的筛选, 最后得知, 反应罐内温度及反应时间是两个关键因子。在本阶段的最初全因子试验时, 因子 A 温度的低水平及高水平取为 250 度及 300 度, 因子 B 时间的低水平及高水平取为 40 秒及 70 秒。在中心点处也做了 3 次试验, 试验结果见表 7—25。

表 7—25

黏合剂生产条件优化问题第一批试验数据表

StdOrder	RunOrder	CenterPt	Blocks	temp	time	stick
1	5	1	1	200	40	23.8
2	7	1	1	300	40	34.1
3	1	1	1	200	70	31.8
4	2	1	1	300	70	42.8
5	6	0	1	250	55	48.2
6	4	0	1	250	55	46.5
7	3	0	1	250	55	49.8

对于这批数据进行分析, 其结果如下:

首先将全部备选项列入模型, 这里包含 temp (温度), time (时间) 及它们的交互作用项 temp * time。

Fractional Factorial Fit: stick versus temp, time

Estimated Effects and Coefficients for stick (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		39.571 4	4.328	9.14	0.003
temp	10.650 0	5.325 0	5.725	0.93	0.421
time	8.350 0	4.175 0	5.725	0.73	0.519
temp * time	0.350 0	0.175 0	5.725	0.03	0.978

S=11.450 0 R-Sq=31.79% R-Sq (adj) =0.00%

Analysis of Variance for stick (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	2	183.145	183.145	91.572	0.70	0.564
2-Way Interactions	1	0.122	0.122	0.122	0.00	0.978
Residual Error	3	393.307	393.307	131.102		
Curvature	1	387.860	387.860	387.860	142.42	0.007
Pure Error	2	5.447	5.447	2.723		
Total	6	576.574				

从 ANOVA 表中可以清楚地看出, 在弯曲 (curvature) 一栏中, p -value 只有 0.007, 显示响应变量 stick 有明显的弯曲趋势。另外, 在残差分析中, 由残差对各自变量的图 (见图 7—40 及图 7—41) 中也可以验证这一点。

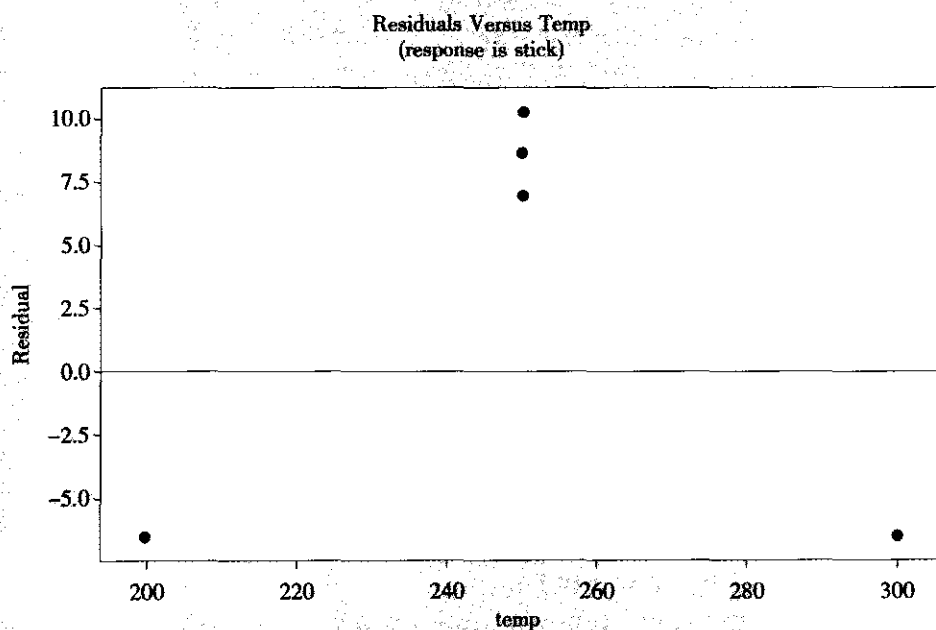


图 7—40 黏度残差对自变量“温度”的散点图

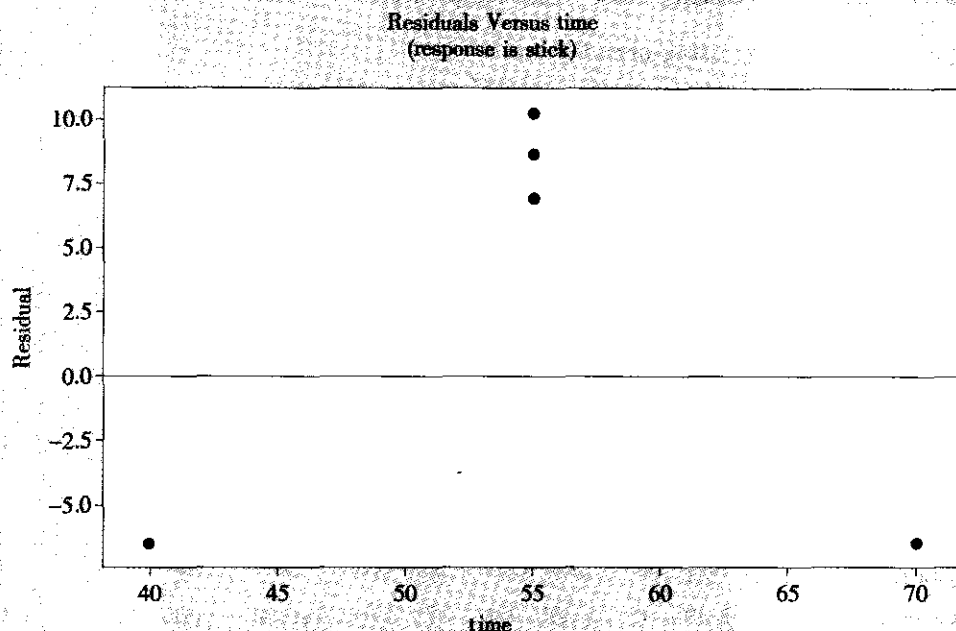


图 7—41 黏度残差对自变量“时间”的散点图

这些结果说明：在本问题中，试验数据有明显的弯曲，这也提示我们，现在进行的试验区域已达到响应变量的最佳区域了。这时对响应变量 stick 单纯拟合一阶线性方程不够了，要再补充些星号点，构成一个完整的响应曲面设计，拟合一个含二阶项的方程就可能解决问题。补做 4 次星号点上的试验，而且由于我们确信这批新做的试验，其各方面条件都与上批相同，因此直接将它们并在一起进行分析。如果没有这种假定，两批试验应该当作两个“区组”（block）来对待才能使分析更准确。补充的 4 个星号点试验结果见表 7—26。

表 7—26 黏合剂生产条件优化问题第二批试验数据表

No	StaOrder	RunOrder	Blocks	temp	time	stick
1	5	2	1	179	55.0	13.0
2	6	4	1	321	55.0	29.3
3	7	3	1	250	33.8	35.7
4	8	1	1	250	76.2	49.1

对全部 11 个点构成的 CCC（中心复合序贯设计）设计进行分析，拟合一个完整的响应曲面模型。分析结果如下：

第一步：拟合选定模型。

我们首先将全部备选项列入模型。这里包含 temp（温度），time（时间），及它们的平方项，temp * temp，time * time 以及它们的交互作用项 temp * time。获得的计算结果如下：

Response Surface Regression: stick versus temp, time

The analysis was done using uncoded units.

Estimated Regression Coefficients for stick

Term	Coef	SE Coef	T	P
Constant	-354.6	21.958 0	-16.148	0.000
temp	2.7	0.126 5	21.475	0.000
time	1.5	0.367 4	4.077	0.010
temp * temp	-0.0	0.000 2	-22.671	0.000
time * time	-0.0	0.002 6	-4.425	0.007
temp * time	0.0	0.000 9	0.253	0.810

S=1.381 R-Sq=99.3% R-Sq (adj) =98.6%

Analysis of Variance for stick

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	5	1 395.95	1 395.950	279.190	146.40	0.000
Linear	2	404.79	960.830	480.415	251.92	0.000
Square	2	991.04	991.042	495.521	259.84	0.000
Interaction	1	0.12	0.122	0.122	0.06	0.810
Residual Error	5	9.54	9.535	1.907		
Lack-of-Fit	3	4.09	4.089	1.363	0.50	0.719
Pure Error	2	5.45	5.447	2.723		
Total	10	1 405.49				

在第一步要注意观察下面几处要点:

(1) 看 ANOVA 表中的总效果。在本例中, 对应回归项 (regression) 的 p -value 为 0.000 (即 <0.001), 表明应拒绝原假设, 即可以判定本模型总的说来是有效的。

(2) 看 ANOVA 表中的失拟现象。在本例中, 失拟项的 p -value 为 0.719, 其数值远比临界值 0.05 大得多, 表明无法拒绝原假设, 即可以判定本模型并没有失拟现象。

(3) 看拟合的总效果多元全相关系数 R^2 (即 R-Sq) 及修正的多元全相关系数 R^2_{adj} (即 R-Sq (adj))。本例中, R-Sq 为 99.3%, R-Sq (adj) 为 98.6%, 二者已经很接近了, 如果将影响不显著的效应删去, 二者会更接近。

(4) s 值的分析。本例中, $s=1.381$ 。

(5) 各项效应的显著性。在计算结果的最开始参数估计中, 列出了各项效应及检验结果。可以看出, temp (温度), time (时间), 及它们的平方项, temp * temp, time * time 都是高度显著的, 而它们的交互作用项 temp * time 则不显著, 将来修改模型时, 应该将此交互作用项删除。

第二步: 进行残差诊断。

利用计算机自动输出的残差图可以很容易进行残差诊断。残差诊断包括4项要求（见图7—42至图7—45，第3项检验正态性省略）。

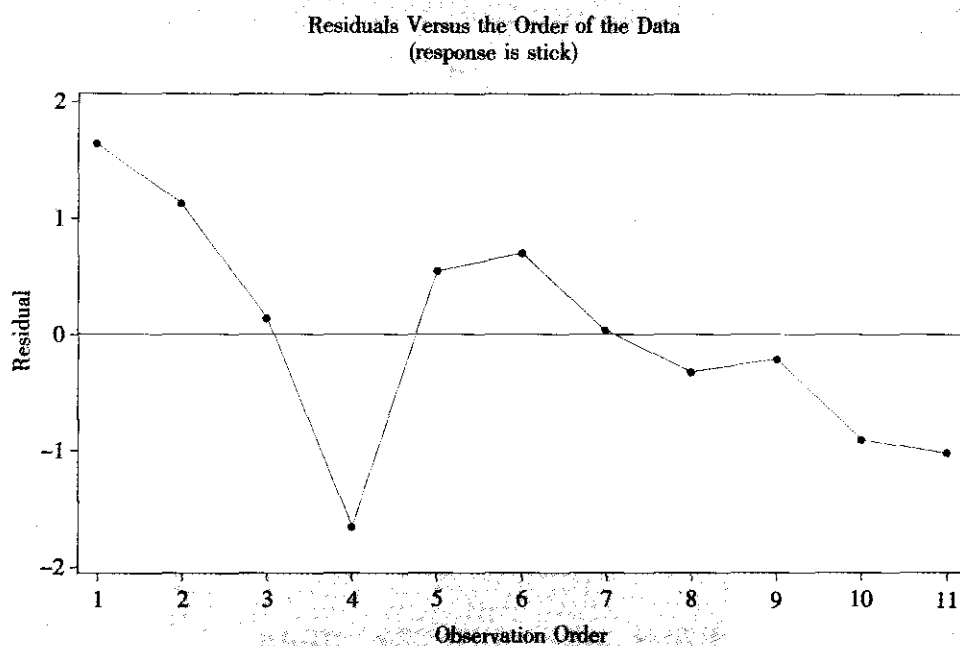


图7—42 残差对于观测值顺序的散点图

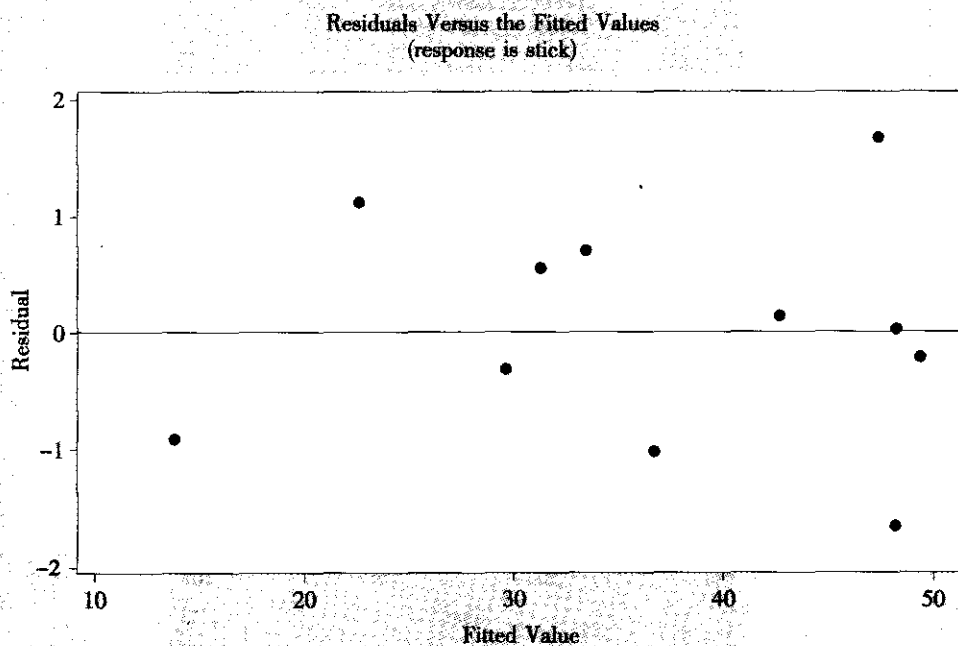


图7—43 残差对于响应变量拟合值的散点图

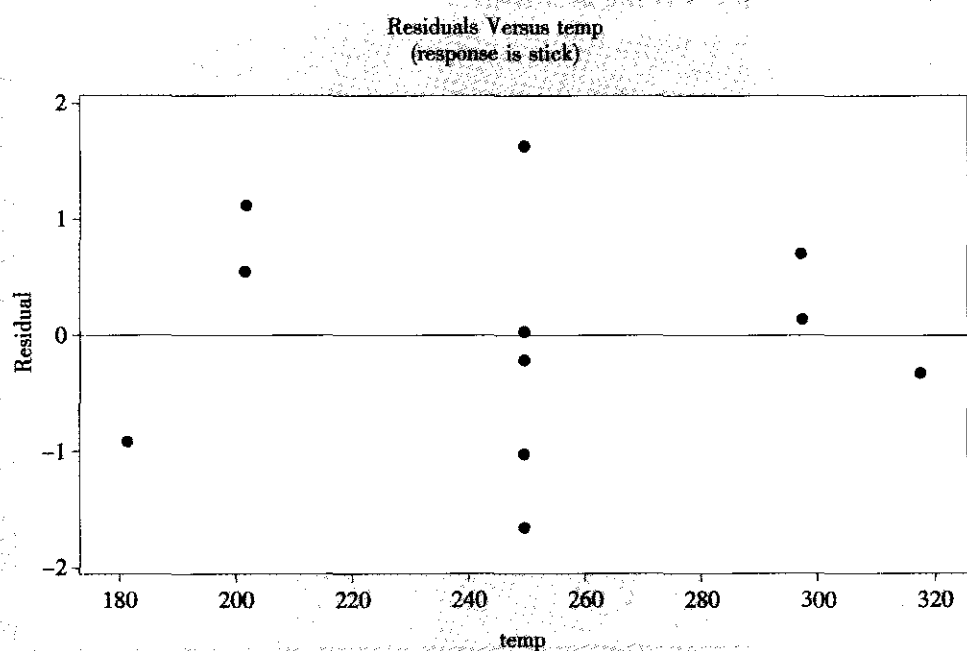


图 7—44 残差对于自变量“温度”的散点图

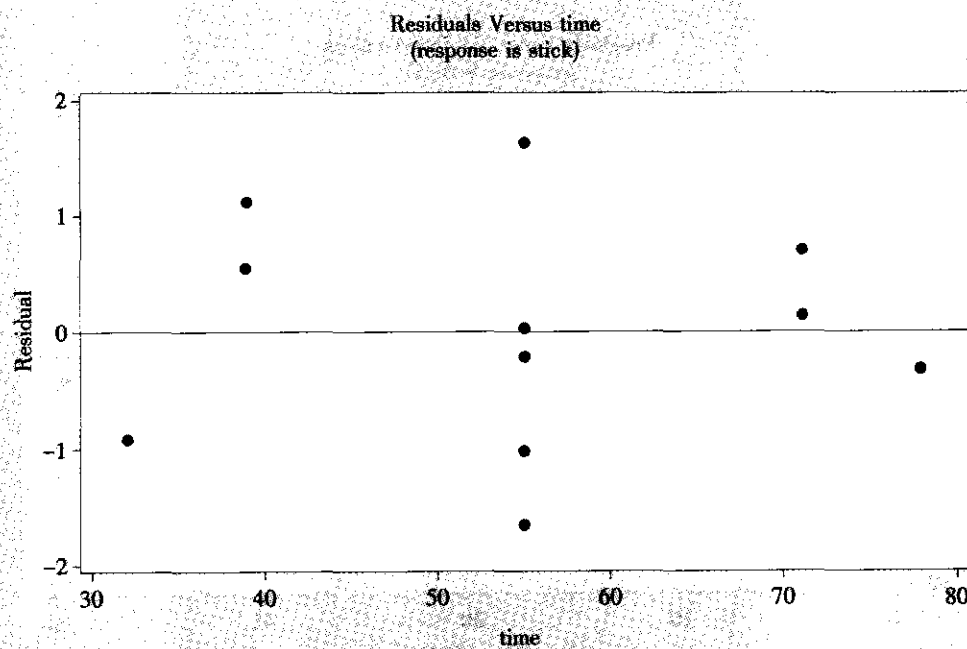


图 7—45 残差对于自变量“时间”的散点图

从上述4张图中可以看出,残差的状况是正常的。

第三步:判断模型是否要改进。

从残差诊断中看出,模型基本上是好的,只是在检验各项效应中,发现两自变量间的交互作用项不显著,因而改进模型主要是删除此不显著项。因此,实际上,又要返回第一步。

新第一步:拟合选定模型。

本例中,在重新拟合模型的计算时要删去交互作用项 $\text{temp} * \text{time}$,而保留 temp (温度), time (时间),及它们的平方项, $\text{temp} * \text{temp}$, $\text{time} * \text{time}$ 。再次计算所得的结果如下:

Response Surface Regression: stick versus temp, time

The analysis was done using uncoded units.

Estimated Regression Coefficients for stick

Term	Coef	SE Coef	T	P
Constant	-357.8	16.483 4	-21.706	0.000
temp	2.7	0.106 5	25.629	0.000
time	1.6	0.263 1	5.915	0.001
temp * temp	-0.0	0.000 2	-24.677	0.000
time * time	-0.0	0.002 4	-4.817	0.003

S=1.269 R-Sq=99.3% R-Sq (adj) =98.9%

Analysis of Variance for stick

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	4	1 395.83	1 395.83	348.957	216.79	0.000
Linear	2	404.79	1 061.70	530.850	329.80	0.000
Square	2	991.04	991.04	495.521	307.85	0.000
Residual Error	6	9.66	9.66	1.610		
Lack-of-Fit	4	4.21	4.21	1.053	0.39	0.810
Pure Error	2	5.45	5.45	2.723		
Total	10	1 405.49				

对于这次计算结果,我们仍用前面介绍过的方法与步骤来进行分析。

此时观察到下面几处要点:

(1) 先看 ANOVA 表中的总效果。在本例中,对应回归项 (regression) 的 p -value 为 0.000 (即 <0.001),表明应拒绝原假设,即可以判定,本模型总的说来是有效的。

(2) 看 ANOVA 表中的失拟现象,在本例中,对应失拟项的 p -value 为 0.810,其值远比临界值 0.05 大得多,表明无法拒绝原假设,即可以判定,本模型删去了一项,但并不

没有造成失拟现象。

(3) 看删减后的模型是否比原来有所改进。我们把两个模型计算的多元全相关系数 $R-Sq$ 和修正的多元全相关系数 $R-Sq(adj)$ 以及标准差的估计量 s 汇总成表 7—27。

表 7—27

全模型与删减模型效果比较表

	全模型	删减模型
R-S	0.993	0.993
R-Sq (adj)	0.986	0.989
s	1.381	1.269

可以看出, 由于模型项数减少了一项, $R-Sq$ 通常会有微小的降低 (本例由于计算的舍入误差, 未看出降低), 但关键是看修正的 $R-Sq(adj)$ 是否有所提高 (本例 $R-Sq(adj)$ 由 0.986 提高到 0.989), s 的值是否有所降低 (本例 s 的值由 1.381 降为 1.269)。本例结果证明删除了不显著的交互作用项后, 回归的效果更好了。

除了上述常用结果外, 为了获得更精密的数值结果, 我们还可以输出更精确的对于原始数据的回归系数值。

Estimated Regression Coefficients for stick using data in uncoded units

Term	Coef
Constant	-377.429
temp	0.055 632 4
time	0.109 353
temp * temp	-2.094 87E-06
time * time	-5.087 57E-05

由上述数值, 我们可以写出最后确定了回归方程:

$$y = -377.43 + 0.056temp + 0.109time - 0.000\ 002temp^2 - 0.000\ 051time^2$$

新第二步: 进行残差诊断。

没有发现任何不正常情况 (详细结果从略), 可以确认上述模型为我们最终选定的模型, 不用再进行第三步模型修改了。

第四步: 对选定模型进行分析解释。

本问题是希望黏度越大越好, 这里两个平方项系数皆为负, 又没有交叉项, 可以肯定将可以从回归方程中 (令两个偏导数为 0) 求得最大值 (但不能保证此最大值一定落在原试验范围内)。在求解前, 先看一下等高线图及曲面图 (见图 7—46 及图 7—47)。从图中可以清楚地看到, 在原试验范围内确实有个最大值。

利用手工解方程的方法, 也可以得到最优解: 当温度 = 260.56 度, 时间 = 67.98 秒时, 所获得的黏度最高, 最佳值可以达到 50.68。

计算机提供了自动求最优解的功能, 利用“响应变量优化器” (response optimizer) 可以直接获得最佳点的设置及最佳值, 同时还提供可以进行人机对话的界面, 显示出有关

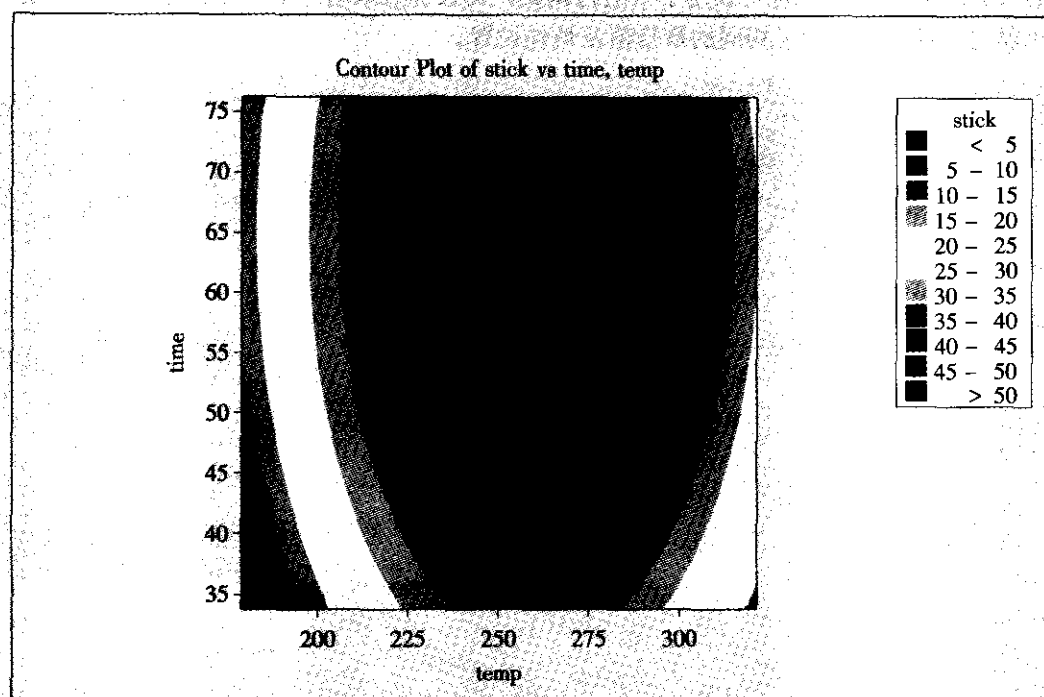


图 7—46 响应变量的等高线图

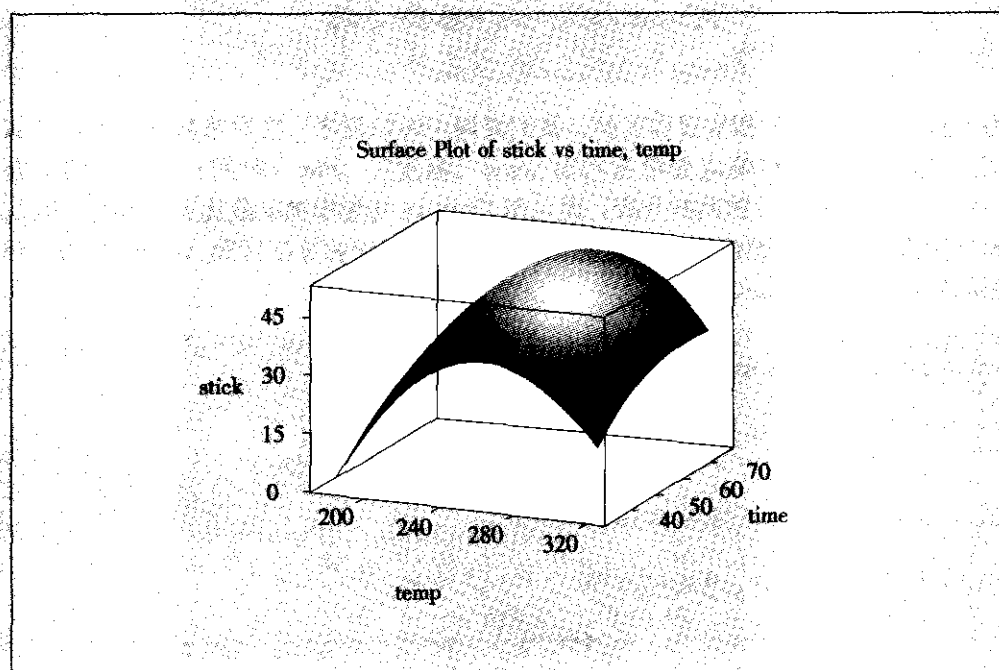


图 7—47 响应变量的曲面图

图形(见图7—48),并且可以用手工进行调整,对最优点取整等等,非常方便。响应变量优化器的计算结果显示了相同的最佳值及最佳设置。

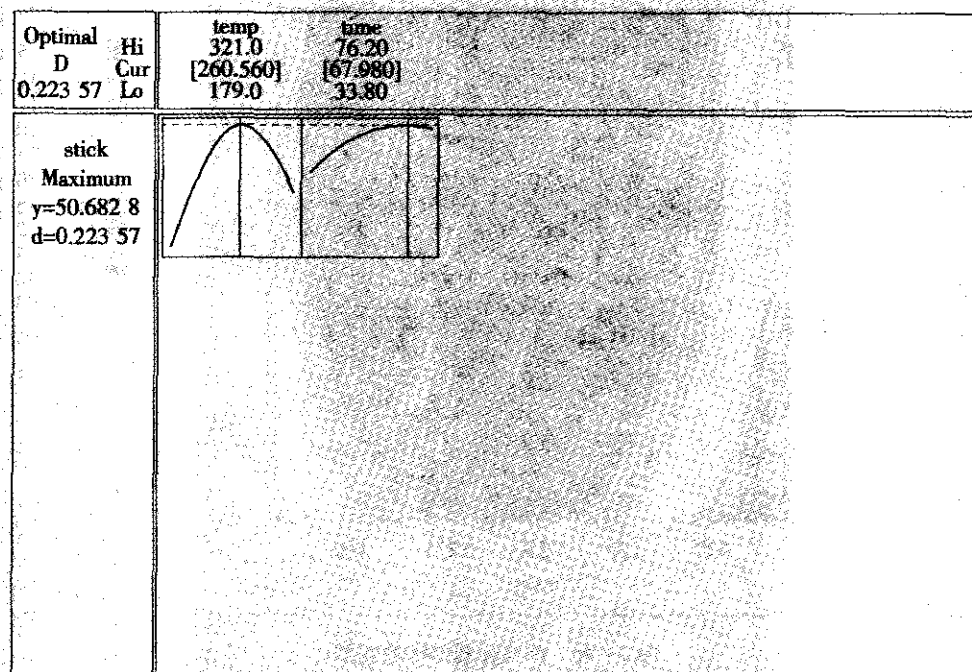


图7—48 响应变量优化器输出结果图

为了获得预测结果的相应置信区间,我们要调用回归程序。在原来数据表中,增加两列自变量平方的数值,将响应变量进行对 temp(温度)、time(时间)及它们的平方项 temp * temp, time * time 的回归分析,而且用 Options 功能增加在最优点的预测计算,即可得到预测值(fit)、预测值的标准误差(SE fit)、预测值的置信区间(95.0% CI)和单个观测值的置信区间(95.0% PI)。

Predicted Values for New Observations					
New	Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	50.082	0.660	(49.067, 52.297)	(47.183, 54.181)	
Values of Predictors for New Observations					
New Obs	temp	time	temp * temp	time * time	
1	261	68.0	68 121	4 624	

前一个置信区间(95.0% CI)表明的是回归方程上的点的置信区间,此值可以作为改进的结果的预测写在总结报告中。后一个置信区间(95.0% PI)表明的是以上述回归方程上的预测值的置信区间为基础,加上观测值固有的波动所给出的置信区间,这就是将

来做一次验证试验时将要落入的范围,可供做验证试验时使用。

7.6 稳健参数设计

稳健参数设计 (robust parameter design) (也称健壮设计、鲁棒设计,简称参数设计)是工程实际问题中很有价值的统计方法。它通过选择可控因子的水平组合来减少一个系统(或产品、过程)对噪声变化的敏感性,从而达到减少此系统性能波动的目的。大家知道,过程的输入变量(因子)有两类:可控因子(control factor)和噪声因子(noise factor)。可控因子是指其值一旦选定就保持不变的变量,它包括产品或生产过程设计中的设计参数,而噪声因子是在正常条件下难以控制的变量。在作参数设计时,就是把可控因子的设置当作研究的主要对象,与此同时让噪声因子按照设定的系统改变的方法来表示正常条件下的变化,最终按照我们预定的望大、望小或望目的目标选出最佳设置。从方法角度而言,可以使用乘积表(cross array)(即内外表)法和单一表(single array)(两类因子排在同一张表内)法两种方法。从建模策略而言,可以使用相应的位置及散度建模和响应建模两种建模策略。日本田口玄一博士(Genichi Taguchi)在参数设计方法方面贡献非常突出,他在设计中引入SN比(信噪比)的概念,并以此作为评价参数组合优劣的一种测度,这是很有价值的,甚至很多文献和软件都把稳健参数设计方法称为田口设计(Taguchi design)。

我们上面讨论到的响应变量都只是单个变量。在所研究的系统中,这个响应变量有特定的最佳值(望大、望小或望目),而且这个最佳值是不变的,这类响应被称为简单响应(simple response),这类系统被称为简单响应系统(simple response system)。在更复杂的系统中,响应变量显著地依赖于一个有效的信号因子,这个系统被称为信号—响应系统(signal-response system)。例如,研究自动车床生产质量问题,其重要指标就是加工出的零件直径的波动性。如果限定此零件直径是50 mm,则可以将其生产工艺条件设计看成是望目特性的问题,波动越小越好,这里我们把生产的零件直径看成简单响应。但实际上,零件直径是由切削深度决定的,我们可以对切削深度加以控制,不同的切削深度会生产出不同直径的产品来。我们希望的是在各种直径零件情况的波动性都能很小。这里,切削深度就是个信号因子(signal factor),整个生产过程就是用切削深度与零件直径间的信号—响应关系来刻画的,细致研究这种关系将比简单的望目特性研究要深入得多、复杂得多。在田口方法的术语中,简单响应关系被称为静态(static)特征的,对应的参数设计问题被称为静态参数设计(static parameter design)。信号—响应关系被称为是有动态(dynamic)特性的,对应的参数设计问题被称为动态参数设计(dynamic parameter design)。本书将只讨论不含信号因子的简单响应问题,即静态情况下的稳健参数设计问题。

7.6.1 稳健参数设计的模型

产品性能指标除了受可控因子的影响外,还受到噪声因子的影响。一般的试验设计对误差的分析比较笼统,全都归为随机误差或试验误差。但在稳健参数设计中,为了能达到产品或过程的稳健性,一定要细致地分析这些变差是如何形成的。首先要明确噪声因子的



具体状况,对噪声因子的状况进行细致的分析,然后才能予以准确描述,进而设法在试验中反映这些变差,最后讨论如何通过稳健参数设计来实现控制这些变差的目标。

1. 噪声因子的进一步分析

噪声因子 (noise factor) 就是在正常生产过程或使用条件下难以控制的因子,在本章中将不区分“噪声”与“难以控制”。噪声因子可能有很多类型,下面是一些常见的类型。

(1) 参数的变化。生产过程中的某些要素,由于在试验设计过程中并不发生设置的改变,我们就并未将它们设定为可控因子。在实际工作中,任何参数的控制都不可能完全准确,这就造成了误差。例如,反应罐中的温度很重要,如果要考虑它的变化影响,则我们可将它处理为可控因子;否则可以设定为常量。但实际上它是有波动的,因此一般情况要将温度这个参数的波动看成噪声因子。

(2) 原材料参数的变化。加工成产品的零部件有其名义值 (nominal value),但是其实际值与名义值总会有偏离,这就造成了变异。由于原材料参数通常被我们选定作为可控因子,这类误差所形成的噪声因子也常称为“内噪声” (internal noise)。

(3) 环境的变化。环境条件对响应变量也会有直接的影响,而我们通常又未将它们放入可控因子范围内,例如手表运行快慢随温度的变化而波动,电视机的清晰度与输入电压的大小有密切关系,很多生产过程与室内条件有关,例如温度 (或湿度)、静电粒子数、输入电压等。我们要在参数稳健设计中考虑如何使产品对这些环境的变化不敏感。我们称这种使用条件和环境条件的波动为外干扰,也称为“外噪声” (external noise)。

(4) 载荷因子。是指产品所承受的外部载荷。例如,洗衣机的设计中要考虑洗衣量的变化,汽车要考虑载重量的变化,冰箱要考虑开门次数的变化等。

(5) 单元间差异与空间差异。原料的不同批次、操作工的不同安排、反应炉中不同位置会有不同温度等,这些差异也造成响应变量的波动。

(6) 时间差异。时间差异是指一段较短的时间区间内,上下午的差异、同一周内不同的几天等产生的差异。

(7) 耗损降级。耗损降级是指在一个长时期内,各种条件的缓慢衰变、元器件老化磨损、生物活性降级、材料耗损等导致性能的长期变异。要让系统对此类变异的灵敏度减小。

本章前几节的试验设计中,一般是靠在单元上或在时间上进行重复试验来估计试验误差的大小,试验的最终目的还是针对响应变量本身 (例如望大或望小)。而稳健参数设计则把响应变量的变差作为研究对象,寻找并识别出直接影响性能变差的变量,这比原来的要求大大深入了一步,功能也大为扩充了。

2. 试验与建模的策略

响应变量的变差如何才能减小呢?很自然的想法是,通过减小噪声的变差来实现减小响应变量的变差,但通常这要付出较高的经济代价。稳健参数设计则是更好的一种策略选择。这种策略是通过探索可控因子与噪声因子间的交互作用,从而用改变可控因子的水平组合的办法来减小响应变量的变差。因为可控因子通常易于改变,所以稳健参数设计比直接减小噪声变差更经济、更方便。

如果可控因子本身受到噪声的影响而有波动,响应变量与这个可控因子的关系是非线性的,则我们可以选择斜率较小的平坦区域从而使响应变量的变差减小。这样减小变差的方法比直接减小可控因子的噪声波动要便宜得多。其示意图可参见图 7—49。一般人们是

对乘积表试验的数据该怎样建模和分析呢？一个很自然的想法就是位置和散度分别建模。

位置与散度建模法 (location and dispersion modeling) 就是分别建立位置和散度的度量值关于可控因子主效应和交互效应的模型。对每个控制水平的组合，用噪声重复试验的样本均值作为位置的度量，用样本方差的对数或样本方差本身作为散度的度量。对这两种度量，分别找出对它们有显著影响的因子来。

凡对位置度量有显著影响者，称为**位置因子** (location factor)；

凡对散度度量有显著影响者，称为**散度因子** (dispersion factor)；

是位置因子但又非散度因子者，称为**调节因子** (adjustment factor)。

对于望目型问题，我们解决问题的两步程序是这样的：

- (1) 选择散度因子的水平使散度最小化；
- (2) 选择调节因子的水平使位置达到目标值。

在第二步中，若只使用一个调节因子仍不足以使位置达到目标值，则可使用两个或多个调节因子来达到目的。若没有调节因子存在，只能使用同时是散度因子的位置因子，这时候，有可能当位置调整到目标时散度却增大了，这就需要重新调整散度因子，并且要在上述两步间进行迭代。如果第二步总也不能使位置达到目标值而且偏差明显时，第一步所得到的结果可能要推翻重来。甚至也有可能仅用这两步找不到可控因子的最佳水平组合。

对于望大或望小型问题，我们解决问题的两步程序是：

- (1) 选择位置因子的水平使位置达到最大（小）；
- (2) 选择非位置因子的散度因子的水平使散度最小化。

之所以可以使用位置与散度建模法来解决稳健参数设计问题，是因为在乘积表中，对于位置和散度的度量都是在相同的噪声表中进行的，位置（或散度）度量间的任何差异都可以归因于可控因子水平的组合。因此，乘积表的使用自然会导致用位置与散度建模法来寻求最优解。

3. 位置及散度的度量

取什么指标来作为位置及散度的度量是最好的呢？前面所说的样本均值及样本方差是常见的选择，但我们的响应变量优化的目标可能有望大、望小和望目三种形式，统一使用“信噪比”及“灵敏度”是田口提出的建议之一。粗略说来，对于望大、望小和望目三种形式，用不同的公式来定义信噪比后，我们的位置与散度建模法的优化步骤的第一步，都可以归结为信噪比极大化，第二步再根据不同的目标选不同的因子予以调节，不一定都选“灵敏度”作指标。下面给出不同特性情况下的**信噪比** (signal to noise ratio, SN ratio) 定义。

- (1) 望目特性信噪比。望目特性的信噪比与变异系数有着密切的关系。

设望目特性 y 为随机变量，它的期望值为 μ ，方差为 σ^2 ，它的目标值为 m 。对于望目特性 y 来说，我们希望：

- 1) $\mu = m$
- 2) σ^2 越小越好

在概率论中，我们常用变异系数 $\gamma = \frac{\sigma}{|\mu|}$ 作为随机变量的离散性指标，即变异系数 γ

越小,说明随机变量(质量特性)可能值的密集程度越高。变异系数的优点是既考虑了标准差 σ 的影响,又考虑了期望值 μ 的影响。如果把标准差理解为绝对误差,则变异系数代表相对误差。

望目特性信噪比定义如下,令

$$\eta = \frac{\mu^2}{\sigma^2}$$

可见,望目特性信噪比 η 等于信号 μ^2 与噪声 σ^2 之比,也就是变异系数平方的倒数。因此, η 是随机变量的一个关于变异性的优良指标,其值越大越好。

η 的分子 μ^2 的估计 $\hat{\mu}^2$ 由 $\hat{\mu}^2 = \frac{1}{n}(S_m - V_e)$ 确定,分母 σ^2 由

$$V_e = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

确定,因此有

$$\hat{\eta} = \frac{\hat{\mu}^2}{\hat{\sigma}^2} = \frac{\frac{1}{n}(S_m - V_e)}{V_e}$$

在实际计算时,通常将估计 $\hat{\eta}$ 取常用对数再乘以10,化为分贝值。在不致引起混淆的情况下,我们仍记为 η 。有

$$\eta = 10 \lg \frac{\frac{1}{n}(S_m - V_e)}{V_e} \quad (\text{dB})$$

式中:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$V_e = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

$$S_m = n\bar{y}^2$$

有时更简化为:

$$\eta = 10 \lg \frac{\bar{y}^2}{V_e} \quad (\text{dB})$$

另一种信噪比 η 值的定义是只考虑样本的方差,定义为:

$$\eta = -10 \lg V_e \quad (\text{dB})$$

(2) 望小特性信噪比。当产品的质量特性 y 为望小特性时,一方面希望其数值越小越好,即等价于希望期望值 μ 的绝对值或 μ^2 越小越好;另一方面,希望 y 的波动越小越好,即相当于希望方差 σ^2 越小越好,也就是 $\mu^2 + \sigma^2$ 越小越好,其倒数越大越好。因此,望小特性 y 的信噪比定义如下:

$$\eta = \frac{1}{\mu^2 + \sigma^2}$$

注意,随机变量 y 的二阶原点矩 $E(y^2)$ 为:

$$E(y^2) = \mu^2 + \sigma^2$$

因此 $\eta = \frac{1}{E(y^2)}$



这说明望小特性 y 的信噪比 η 应取二阶原点矩 $E(y^2)$ 的倒数。

二阶原点矩 $E(y^2)$ 的无偏估计称为均方值 V_T :

$$V_T = \frac{1}{n} \sum_{i=1}^n y_i^2$$

因此, η 的估计公式为:

$$\hat{\eta} = \frac{1}{V_T} = \frac{n}{\sum_{i=1}^n y_i^2}$$

取常用对数再乘以 10, 化为分贝值, 则得到望小特性信噪比的估计公式为:

$$\eta = -10 \lg \frac{1}{n} \sum_{i=1}^n y_i^2 \quad (\text{dB})$$

(3) 望大特性信噪比。设 y 为望大特性, 则 $\frac{1}{y}$ 为望小特性。因此, 将望小特性信噪比的估计公式中 y_i 变换成 $\frac{1}{y_i}$, 可分别得到望大特性信噪比的估计公式:

$$\hat{\eta} = \frac{n}{\sum_{i=1}^n \frac{1}{y_i^2}}$$

$$\eta = -10 \lg \frac{1}{n} \sum_{i=1}^n \frac{1}{y_i^2} \quad (\text{dB})$$

以上介绍了用乘积表进行位置与散度建模的方法。这是目前最通用的做法, 计算机软件也多按此思路编写程序。但是位置与散度建模方法也有其缺点, 这就是用可控因子建立的位置与散度模型可能会掩盖可控因子与噪声因子之间的某些重要关系, 因为建立的各项效应分析都只考虑了可控因子本身, 噪声因子并没有连同可控因子一起予以考虑, 特别是未能考虑二者的交互作用。另外, 散度还可能并不是可控因子的线性函数, 更有可能是其非线性函数, 这些使建立度量散度的模型更加困难。再一点就是, 乘积表所用的试验次数实在是太多了, 因此提出一种全新的设想, 即用单一表直接进行响应建模。

在这种响应建模 (response modeling) 方法中, 不是计算噪声因子在重复试验上的方差, 而是将响应变量 y 拟合为可控因子及噪声因子的函数, 将这种拟合模型称为响应模型 (response model)。安排试验时, 也是将可控因子与噪声因子统一安排在一张表中, 这将大大减少试验次数。其设计及分析用到较多统计知识, 本书从略。详细内容可以参看本书后所附参考文献中吴建福教授 (C. F. Jeff. Wu) 的著作。

7.6.2 稳健参数设计的计划

对于实际应用问题, 稳健参数设计的主要步骤包括:

1. 制定可控因子水平表

同前几节介绍的一般的试验设计一样, 首先要确定本问题中的可控因子名单, 并确定它们各取多少水平。如果确实需要考虑某些因子间的交互作用, 则最好选定所有因子取相同水平的对称正交表 (symmetrical orthogonal array), 这时交互效应容易得到估计,



但水平数受到限制。在一般情况下,我们不应特别强调交互作用,可以选用混合水平的非对称正交表(asymmetrical orthogonal array),这使各因子水平数的选择有更多的考虑余地。

2. 制定控制表(或称内设计或内表)

一旦可控因子名单和它们的水平确定下来,就可以选定控制表。全部二水平或全部三水平的全因子试验或部分实施因子试验都可以使用。例如二水平的L4, L8, L16, L32;三水平的L9, L27等。如果有多种水平的混合,则可以选用L12, L18, L36等设计表。在稳健参数设计时,很少使用全因子试验,而多是采用部分实施而且不考虑其交互作用,因为这可以在较少次数的试验中考察相当多的因子,是很经济和有效的。

使用MINITAB软件可以自动形成内表。主要步骤是:从STAT-DOE-Taguchi>Create Taguchi Design入口,根据自己选定的是对称型(各因子水平数相同)设计还是混合型设计,并考虑到可控因子个数,先可以从“Display Available Design”中观察和确认自己选定的试验次数是否可行,然后,从对话框“Design”中,选定试验次数。注意这里的圈选框“Add a signal factor for dynamic characteristics”是针对动态型稳健参数设计问题而给出的,在静态问题中,先不要管它。在对话框“Factor”中,输入因子名称,水平设置。如果有交互作用需要特别设定的话,注意在这里选择按键“Interaction”后,对要考虑的交互作用项名称予以设定,否则一律不考虑交互作用。全部输入完成后,即可输出设计表(即控制表,或称内表)。

3. 制定误差因子水平表

由于误差因子有很多种类型,要仔细分清每种误差的来源,并设定正确的计算公式。例如,可控因子的误差(用相对误差给出)也是应该作为误差因子来对待的,其取值的计算通常是依赖于可控因子本身的数值的,而环境误差等则常取固定的设定值作为“外噪声”。

4. 制定噪声表(或称外设计或外表)

当误差因子取定后,如何安排噪声表(或称外表)是件非常重要的事。一般情况下,噪声表也采用正交表,因为它可以保证布点“均衡分散”,具有较好的均匀性,这时实施一次噪声表试验相当于在噪声变差范围内实施一次全面模拟试验。而实际上,相当于对内表的每一个试验都要安排一个全面的噪声表。

为了实现对内表的每一个试验都安排一个全面的噪声表,就要把内表(控制表)和外表(噪声表)“相乘”而得到,相乘的方法如表7-28所示。内表是用L9安排4个可控因子,每个因子皆取三水平。外表是用L8安排4个误差因子,其中每个因子为二水平,按二分之一实施安排的。这样一来,搭配在一起共要做 $9 \times 8 = 72$ 次试验。

在用MINITAB软件进行稳健参数设计时,并不真正列出乘积表,误差表(外表)要自己另行安排,其试验结果按顺序填入响应变量的多个列中即可。

能够安排完成全面的乘积表试验当然会得到很好的结果,但噪声因子个数较多或噪声因子水平较多时,试验次数会很大,这时可以用较少次数的试验安排来代替。通常有两种方法可以解决此问题。第一种方法就是综合误差法,考虑在噪声表中选择少数几点(通常3点或4点)能使误差达到最大的最具代表性的试验结果作为全部试验误差的代表,这就

可以大大节省试验次数。第二种方法是第一种方法的特例,就是最不利综合误差法,在噪声表中只选2点最不利情况(一个正偏,一个负偏)作为全部试验误差的代表,这就使试验总次数更少了。具体如何选择噪声表要根据实际情况慎重确定。

安排好噪声表后,经过试验,将所有试验结果作为响应变量的多个取值并列排在试验表最右侧,这就完成了全部的试验实施及收集数据步骤。

7.6.3 稳健参数设计实例分析

下面通过两个例子,分别说明望目和望大(小)类型问题的稳健参数设计的分析方法。

例 7—9 塑料袋口封装机封装强度分析

响应变量是封装强度,若封装强度过低,则被包装物容易脱落受损;若封装强度过高,则顾客拆解包装困难。封装强度目标定为18(kg)。我们已经知道影响封装强度的因子有3个:温度、压力及厚度,希望能各取三水平。分析3个因子效应的显著性,寻求最佳搭配,以使封装强度与目标值18(kg)最为接近。为此,我们选用L9正交表,如表7—29所示。

表 7—29 三水平正交表 L9

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

将温度(取60,75及90摄氏度)、厚度(取32,36及40微米)和压力(取1.00,1.25及1.50千克)分别安排在上述表格的前3列,形成内表如表7—30所示。

表 7—30 安排各因子的控制表(内表)

	温度	厚度	压力
1	60	32	1.00
2	60	36	1.25
3	60	40	1.50
4	75	32	1.25

续前表

	温度	厚度	压力
5	75	36	1.50
6	75	40	1.00
7	90	32	1.50
8	90	36	1.00
9	90	40	1.25

对于误差因子, 选定温度 ($\pm 5\%$)、压力 ($\pm 5\%$) 及厚度 ($\pm 10\%$)。

我们要用 MINITAB 软件生成田口试验计划表, 其步骤如下:

从 STAT-DOE-Taguchi-Create Taguchi Design 入口, 控制表的 3 个因子, 每个因子皆三水平, 故选水平数为 3, 因子数为 3。在对话框 “Design” 中选定 9 次试验, 在对话框 “Factor” 中选定因子名称, 并确定具体水平, 为简化叙述, 我们仍以代码 1, 2, 3 作为各因子的水平。这就可以形成田口设计的内表了 (见表 7—31 的前 4 列)。

用综合误差法, 对内表中的每行选取最不利状况, 进行 2 次试验。结果列在数据文件表内 (见表 7—31 的第 5, 6 两列) 中, 表中最后 3 列是计算机自动算出的结果。

表 7—31

封装试验结果数据表

	Temperature	Thickness	Pressure	y1	y2	SNRA	STDE	MEAN
1	60	32	1.00	19.0	16.0	18.328 6	2.121 32	17.50
2	60	36	1.25	18.4	18.0	36.170 5	0.282 84	18.20
3	60	40	1.50	17.5	17.0	33.766 7	0.353 55	17.25
4	75	32	1.25	18.6	17.5	27.312 0	0.777 82	18.05
5	75	36	1.50	19.3	17.0	20.953 3	1.626 35	18.15
6	75	40	1.00	19.1	18.5	32.930 4	0.424 26	18.80
7	90	32	1.50	18.4	16.0	20.116 6	1.697 06	17.20
8	90	36	1.00	17.0	16.5	33.511 2	0.353 55	16.75
9	90	40	1.25	15.9	14.5	23.724 6	0.989 95	15.20

对于以上数据的分析, 我们可以直接使用 MINITAB 软件来进行。

计算机软件提供了全部要用手工计算的结果, 包括信噪比、均值等。用 MINITAB 的操作如下:

从 STAT-DOE-Taguchi-Analyze Taguchi Design 入口, 将两列结果 (如果有多列计算方法也相同) 都放入响应变量 “Response” 内。在对话框 “Graphs” 中选定要输出信噪比、均值图形。在对话框 “Storage” 中选定要输出信噪比、均值、标准差数值。在对话框 “Options” 中选定信噪比计算公式: 对于望目者, 可以选用通常信噪比 ($\eta = 10 \lg \frac{\bar{y}^2}{V_e}$),

也可以单纯考虑样本方差选用 ($\eta = 10 \lg V_e$), 本例选定 ($\eta = 10 \lg \frac{\bar{y}^2}{V_e}$)。计算结果如下:

Taguchi Analysis: y1, y2 versus Temperature, Pressure, Thickness			
Response Table for Signal to Noise Ratios			
Nominal is best ($10 * \text{Log} (Ybar ** 2 / s ** 2)$)			
Level	Temperature	Pressure	Thickness
1	29.42	21.92	28.26
2	27.07	30.21	29.07
3	25.78	30.14	24.95
Delta	3.64	8.29	4.12
Rank	3	1	2
Response Table for Means			
Level	Temperature	Pressure	Thickness
1	17.65	17.58	17.68
2	18.33	17.70	17.15
3	16.38	17.08	17.53
Delta	1.95	0.62	0.53
Rank	1	2	3

在第一个信噪比表 Nominal is best ($10 * \text{Log} (Ybar * * 2 / s * * 2)$) 分析中, 可以看到每个因子的各水平的信噪比均值和极差 (delta), 从极差的大小排序 (rank) 中, 可以看出在各因子对于信噪比的影响中, 以因子 B 为最重要, 因子 C 次之, 因子 A 最小。

在第二个均值表 response table for means 分析中, 可以看到各因子的各水平均值的平均值和均值的极差 (delta)。从极差的大小排序 (rank) 中, 可以看出在各因子对于平均值的影响中, 以因子 A 为最重要, 因子 B, C 都很小。

可以初步判定, 因子 B, C 对望目特性信噪比影响最大, 因而为散度因子; 因子 A 对平均值影响最大, 因而为位置因子, 而且由于 A 对于望目特性信噪比影响很小, 可以认定因子 A 为调节因子。

再来看有关图形 (见图 7—50 及图 7—51), 也证实了上述分析结果。

根据对于望目特性的问题解决程序, 首先选择散度因子的水平使散度最小化。在本例中, SN 最佳水平应确认为: B (Thickness) 取二水平, C (Pressure) 取二水平。此时试验条件为 2 号 (A=1); 实际结果是 SN=36.17, Mean=18.20。

下一步是确定均值的最佳水平及预测。对于田口试验设计的预测方法, 通常是使用主效应迭加方法来计算的 (因为通常在田口设计中不考虑交互作用), 这样预测的结果一般还是可信的, 但也有很多情况甚至与实际试验结果有较大偏离, 为了更深入地了解这方面的问题以及解决方法, 可以参看茆诗松、马逢时和吴建福 (1990) 的论文。此处只给出计算结果。

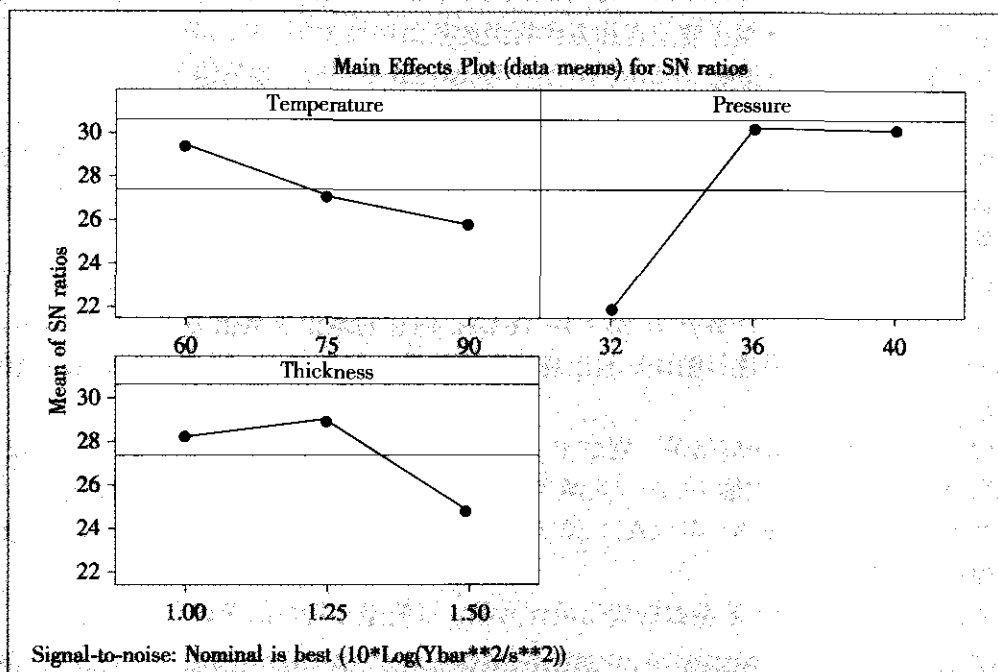


图 7—50 信噪比的主效应图

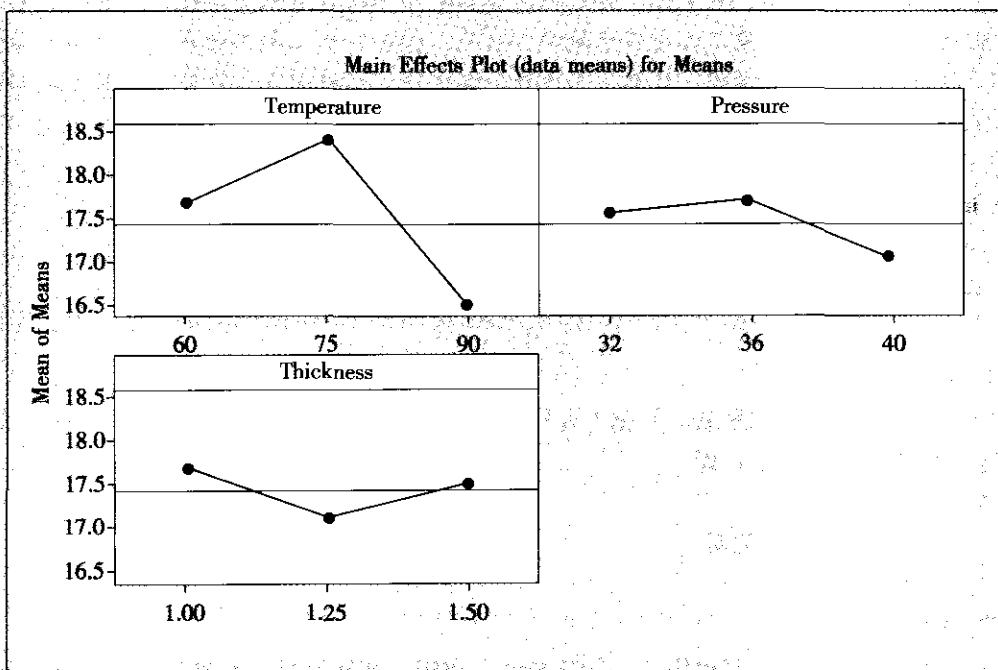


图 7—51 响应变量均值的主效应图



(1) 现有的在搭配 (A1B2C2) 处 (Temperature=60) 的 2 次实际试验结果之平均为 18.20, 而 MINITAB 根据主效应迭加方法得到的预测结果为 $Y=17.5889$;

(2) MINITAB 试算搭配 (A2B2C2) 处 (Temperature=75), 预测结果 $Y=18.2722$;

(3) MINITAB 试算搭配 (A3B2C2) 处 (Temperature=90), 预测结果 $Y=16.3222$ 。

注意, 我们对于最后的两个试验条件并未进行试验, 这里只是根据主效应的迭加而得到的预测值, 并不一定准确, 要经过验证试验才能肯定。从上述结果来看, 我们可以有下列两种最终选择:

(1) 如果试验结果的数值 18.20 (实测平均结果) 与目标 18 的误差可以容忍, 则可选 $A=60$ (A1), 即选定生产条件为 A1, B2 及 C2; 但它的理论预测值为 $Y=17.5889$, 与目标 18 相差较大, 要再进行验证试验以确认最终结果, 但数值 18.20 是实际试验的结果, 应该比较可信。

(2) 希望得到更精确的结果, 则由于 A2B2C2 的预测平均值为 18.2722, 离 18 较近, 而 A1B2C2 试验结果的数值 18.20 (实测平均结果) 也与目标 18 接近, 故可以固定 B, C 两个因子取 B2, C2, 在 $A=60$ (A1) 及 $A=75$ (A2) 之间再做进一步试验, 以寻求更接近目标值的更好结果。

下面以钛合金磨削工艺参数的优化设计为例, 说明望小特性的参数设计方法。

例 7-10

钛合金磨削工艺参数的优化设计

钛合金以其强度高、重量轻、耐热性好和具有良好的抗腐蚀性等优点, 被人们誉为“未来的钢铁”, 目前已被广泛应用于航空、航天、造船和化工等工业部门。但是, 钛合金的导热系数小、黏附性强、抗氧化能力低, 致使磨削性能极差。即使采用特制的砂轮磨削钛合金, 其表面粗糙度也只能达到 $Ra > 0.6(\mu m)$, 即相当于 $\Delta 6 \sim \Delta 7$ 。为了进一步降低表面粗糙度, 今用稳健参数设计方法优化钛合金磨削工艺参数。

试验目的: 优化钛合金磨削工艺参数, 将表面粗糙度降至 $0.2\mu m$ 以下。

质量特性: 表面粗糙度 y (即 Ra), 望小特性。

试验指标: 以望小特性选定信噪比 η , 则 η 越大越好。

1. 试验计划的制定

(1) 制定可控因子水平表。根据专业知识, 选用对表面粗糙度影响较大的因子作为磨削工艺参数中的可控因子, 即

A: 工件转速

B: 修整砂轮时的走刀量

C: 工件纵向走刀量

D: 磨削深度

为了减少试验次数, 其他因子, 如冷却液、磨床、磨削用量及修整用量中的其他参数均固定不变。

选取可控因子水平, 见表 7-32。

表 7—32 钛合金磨削工艺试验的可控因子水平表

因子 水平	A (转/分)	B (mm/转)	C (mm/转)	D (mm)
1	112	0.03	0.82	0.005
2	160	0.06	3.30	0.002 5
3	80	0.09	1.65	0.001 25

表中因子的水平为随机排列, 因子间交互作用可以忽略。选用三水平正交表, 进行 9 次试验。

(2) 确定综合误差因子及其水平。本例质量特性表面粗糙度 y 是不可计算的, 只能通过试验测出其值。为了减少试验次数, 外设计采用最不利综合误差因子法。

对于望小特性, 综合误差因子 N' 取如下两种水平:

N'_1 : 标准条件

N'_2 : 正侧最坏条件

本例对下表中的每号方案, 分别在综合误差因子 N' 的两个水平 N'_1 , N'_2 下各测得一个数据 y_{i1} , y_{i2} ($i=1, 2, \dots, 9$), 全部试验结果列在表 7—33 中。

表 7—33 钛合金磨削工艺试验结果数据表

A	B	C	D	y_1	y_2	SNRA1	MEAN1	STDE1
1	1	1	1	0.162	0.184	15.221 6	-0.173 0	0.015 556 3
1	2	2	2	0.259	0.313	10.834 1	0.286 0	0.038 183 8
1	3	3	3	0.178	0.206	14.310 9	0.192 0	0.019 799 0
2	1	2	3	0.204	0.211	13.658 4	0.207 5	0.004 949 7
2	2	3	1	0.226	0.244	12.572 3	0.235 0	0.012 727 9
2	3	1	2	0.167	0.178	15.259 8	0.172 5	0.007 778 2
3	1	3	2	0.213	0.228	13.126 8	0.220 5	0.010 606 6
3	2	1	3	0.157	0.188	15.229 3	0.172 5	0.021 920 3
3	3	2	1	0.238	0.271	11.868 0	0.254 5	0.023 334 5

我们以 MINITAB 为例说明分析过程。从 STAT-DOE-Taguchi-Analyze Taguchi Design, 将 y_1 , y_2 放入响应变量, 注意在对话框 “Options” 选中 “望小” 特性, 即以式 $\eta =$

$10 \lg \frac{1}{n} \sum_{i=1}^n y_i^2$ 为信噪比, 最后可得下列结果:

Response Table for Signal to Noise Ratios

Smaller is better

Level	A	B	C	D
1	13.455 5	14.002 3	15.236 9	13.220 6
2	13.830 2	12.878 6	12.120 2	13.073 6
3	13.408 0	13.812 9	13.336 7	14.399 5
Delta	0.422 1	1.123 7	3.116 7	1.326 0
Rank	4	3	1	2

Response Table for Means

Level	A	B	C	D
1	0.217 000	0.200 333	0.172 667	0.220 833
2	0.205 000	0.231 167	0.249 333	0.226 333
3	0.215 833	0.206 333	0.215 833	0.190 667
Delta	0.012 000	0.030 833	0.076 667	0.035 667
Rank	4	3	1	2

Response Table for Standard Deviations

Level	A	B	C	D
1	0.024 513 0	0.010 370 9	0.015 084 9	0.017 206 3
2	0.008 485 3	0.024 277 3	0.022 156 0	0.018 856 2
3	0.018 620 5	0.016 970 6	0.014 377 8	0.015 556 3
Delta	0.016 027 8	0.013 906 4	0.007 778 2	0.003 299 8
Rank	1	2	3	4

2. 注意到对于望大或望小型问题，解决问题的两步程序

- (1) 选择位置因子的水平使位置达到最大(小)。
- (2) 选择非位置因子的散度因子的水平使散度最小化。

从上述 S/N 及 MEAN 结果看出，这两方面的评价是一致的，即影响位置的因子顺序首先是 C，然后是 D 和 B，而 A 就次要多了。选 S/N 最大同时使平均值最小的最佳搭配是 C1D3B1，散度因子的顺序是 A，B，C 及 D。因此，A 是非位置因子的散度因子，它取 A2 将使散度最小。因此最佳搭配应选 A2B1C1D3。

从下面的三个效应图（见图 7—52 至图 7—54）上可以得到同样结论。

利用 MINITAB 可以给出在此最佳条件下的预测值。从 STAT-DOE-Taguchi-Predict Taguchi Result 入口，选定输出 Mean, S/N Ratio 及 Std Deviation，然后在因子 (Terms) 中选定 A, B, C, D。在水平 (Levels) 中选定 A2B1C1D3，即可得到下列结果：

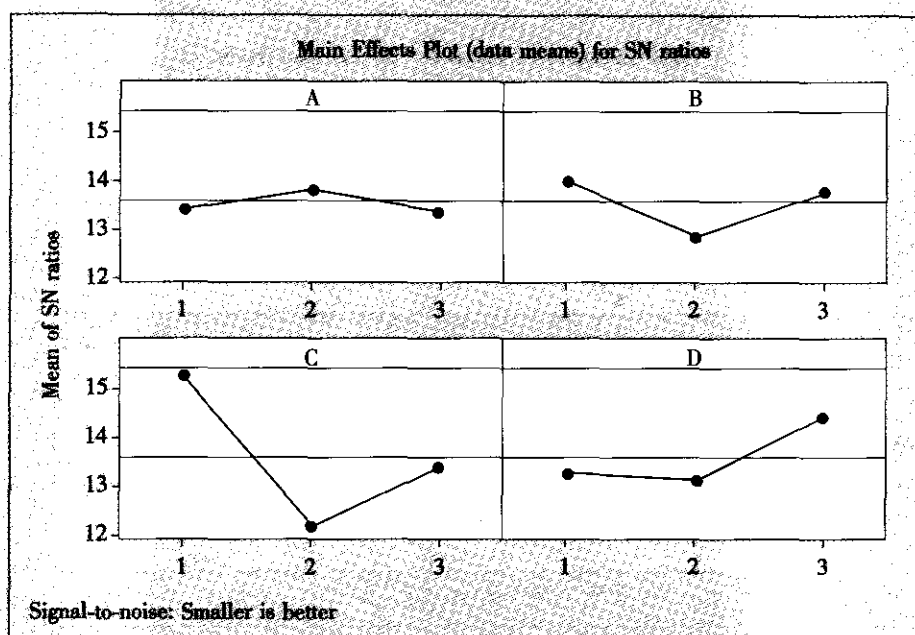


图 7—52 钛合金磨削工艺试验信噪比的主效应图

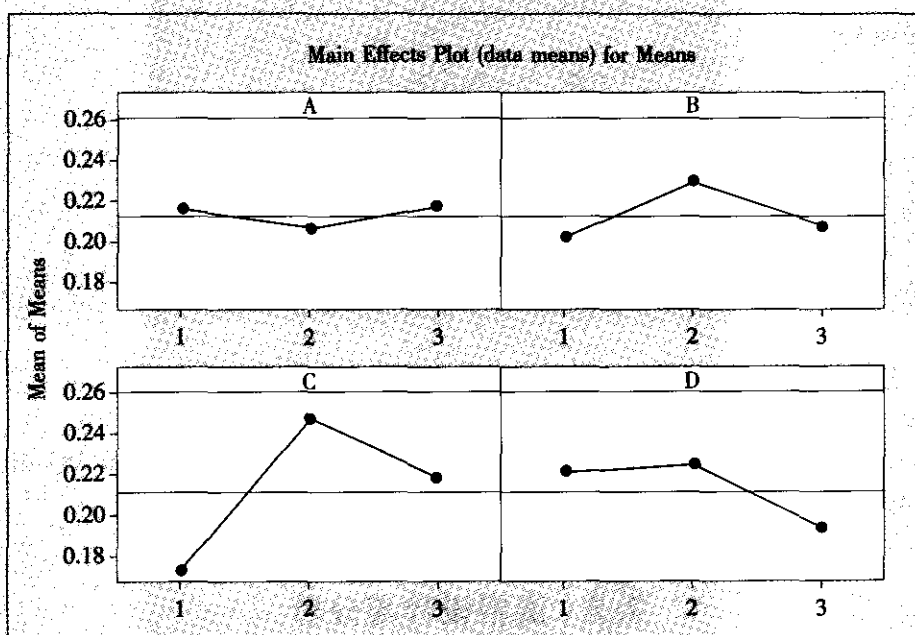


图 7—53 钛合金磨削工艺试验响应变量均值的主效应图

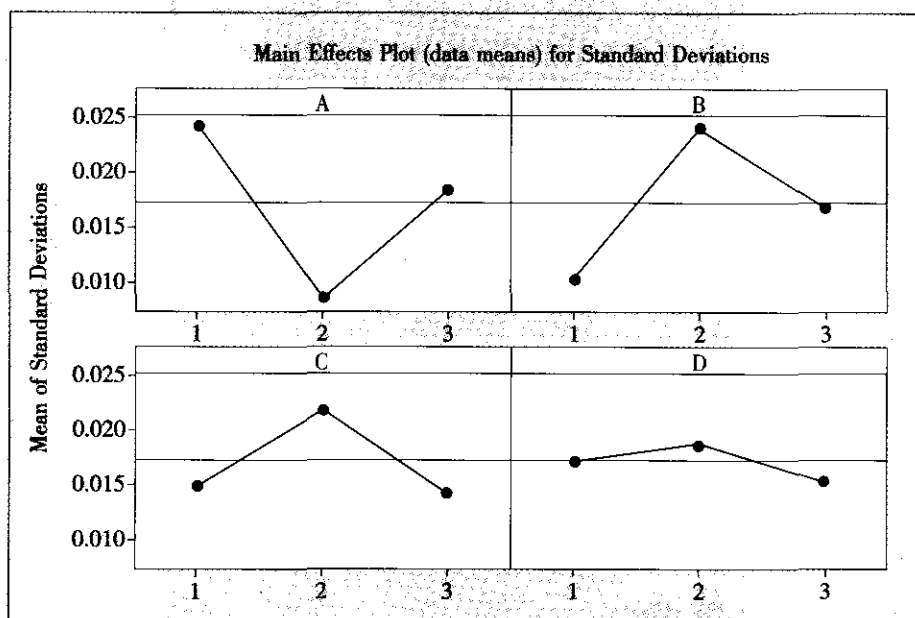


图 7—54 钛合金磨削工艺试验响应变量标准差的主效应图

Predicted values

S/N Ratio	Mean	StDev	Log (StDev)
16.775 1	0.130 833	0.002 121 3	-5.477 91

Factor levels for predictions

A	B	C	D
2	1	1	3

按工艺参数 A2B1C1D3 做五次验证试验，测得其表面粗糙度如下：

0.138, 0.139, 0.159, 0.145, 0.166 (μm)

这 5 个结果都达到预期目的，粗糙度都在 0.2 (μm) 以下，其均值为 0.149 (μm)， η 值为 16.49 (dB)。

7.7 混料设计与分析

7.7.1 混料设计概论

在实际工作中，常常需要研究一些配方配比试验问题。这种问题常出现在橡胶、化工、制药、冶金等课题中。例如不锈钢是由铁、镍、铜和铬 4 种元素组成；闪光剂由镁、

硝酸钠、硝酸锶及固定剂组成；复合燃料、复合塑料、混纺纤维、混凝土、黏接剂、药品、饲料等都是由多种成分按相应比例制作而成的，统称为混料。组成混料的各种成分称为混料成分或分量，也就是混料试验中的因子。这里我们关心的是各种分量的比例而不是其绝对数值；而且显然所有分量的比例之和总是为1。这里对分量的约束条件将使本节内容与以前所讨论的试验设计类型有所差别。对于这种分量的比例之和总是为1的试验设计，我们称为混料设计（mixture design）。在本节中，混料的成分至少应有三种，其约束如图7—55所示。

为了更直观显示各分量的组成状况，引入“三线坐标系”，其原理是：平面几何的知识告诉我们，等边三角形内的任何一点，到三边的（垂线）距离的和等于该三角形的高。如果把三角形的高设为1，则任何一点就可以由其到三个边的三个距离给出三个坐标。当然，这三个坐标并不独立，三者之和恒为1。三线坐标系的示意图见图7—56。

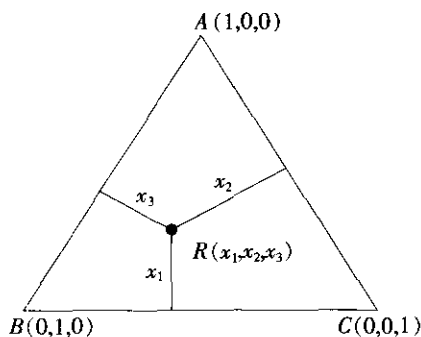


图7—55 混料设计的约束图

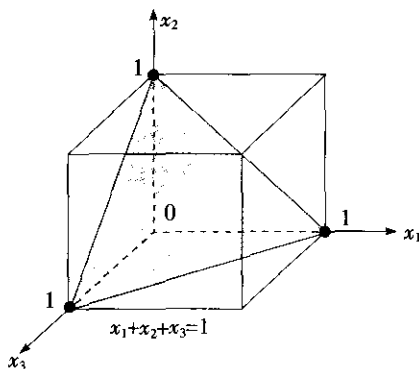


图7—56 三线坐标图

从图中可以看出：三角形的三个顶点坐标分别为 $A(1, 0, 0)$ ， $B(0, 1, 0)$ 及 $C(0, 0, 1)$ 。三角形内任意一点都有三个坐标，可以理解为每个点到三个顶点对应边的“距离”，越靠近某个顶点，这项坐标越接近于1，越远离某个顶点，这项坐标越接近于0。如果点移到了三角形的边上时（图中未显示），则表示三种成分中缺少一种（此缺少成分的名称，在此边所对的顶点处）。三角形的重心到三顶点距离相同，所以坐标为 $(1/3, 1/3, 1/3)$ 。图7—56只是因子个数为3的示意图，4个因子将画为正四面体，5个以上则无直观图形了。以下我们将因子个数记为 p （或称 p 维），对于设计及分析将对一般的 p 值来讨论，但仍然以三因子的图作为示意。

当然，混料设计还可以有更细致的分类。第一种混料（mixture）型最简单，最终结果只与各分量的比例有关；第二种混料—总量（mixture-amounts）型，最终结果与各分量的比例及混料总量有关；第三种混料—过程变量（mixture-process variable）型，最终结果与各分量的比例及某些过程变量有关，而这些过程变量并不是混料的一部分，但可能影响混料的搅拌性质。本节只讨论第一种类型问题。



7.7.2 混料试验的计划

为了全面考察各分量对于响应变量的影响，我们要在整个试验区域内布置一些点来做试验。点的选取方法有三种。

第一种：单纯形重心法 (simplex centroid)

在 p 个因子的混料设计问题中，单纯形重心法的基本思想是：试验点由下列 p 批点组成。第一批是各顶点 (共 p 个)，第二批是上述顶点中，每两个顶点的重心 (共 $p(p-1)/2$ 个)，第三批是上述顶点中，每三个顶点的重心 (共 $p(p-1)(p-2)/3!$ 个)， \dots ，总之， p 批点总计有 $2^p - 1$ 个。

三因子单纯形重心法设计示意图见图 7—57 (a)。这时，列出试验计划表中各试验点 (总共 7 个) 的坐标是这样的：

(1, 0, 0), (0, 1, 0), (0, 0, 1) 顶点；

(0.5, 0.5, 0), (0.5, 0, 0.5), (0, 0.5, 0.5) 两个顶点的重心；

(1/3, 1/3, 1/3) 三个顶点的重心。

在这种类型试验结果下，将要拟合的是下列方程 (以三因子为例)：

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3$$

图 7—57 是单纯形重心法设计示意图。

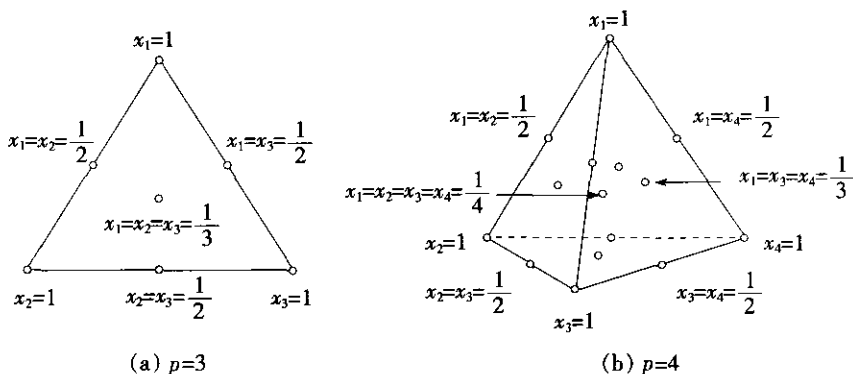


图 7—57 单纯形重心法设计示意图 (三维及四维)

下面给出四因子单纯形重心法设计计划表 (见表 7—34)。

表 7—34

四因子单纯形重心法设计计划表

	x_1	x_2	x_3	x_4	
1	1	0	0	0	顶点
2	0	1	0	0	顶点
3	0	0	1	0	顶点
4	0	0	0	1	顶点
5	0.5	0.5	0	0	2 顶点重心
6	0.5	0	0.5	0	2 顶点重心

续前表

	x_1	x_2	x_3	x_4	
7	0.5	0	0	0.5	2 顶点重心
8	0	0.5	0.5	0	2 顶点重心
9	0	0.5	0	0.5	2 顶点重心
10	0	0	0.5	0.5	2 顶点重心
11	0.333	0.333	0.333	0	3 顶点重心
12	0.333	0.333	0	0.333	3 顶点重心
13	0.333	0	0.333	0.333	3 顶点重心
14	0	0.333	0.333	0.333	3 顶点重心
15	0.25	0.25	0.25	0.25	4 顶点重心

第二种：单纯形格点法 (simplex lattice)

在 p 个因子 (或称 p 维) 的混料设计问题中, 单纯形格点法的基本思想是将全部格子点集内每个点依次选中。格子点是由维数和阶数 (degree) 这两个参数给定的, 此格子点集记为 $\{p, d\}$ 。仍以 $p=3$ 为例。三维一阶格子点集合 (记作 $\{3, 1\}$) 就是 3 个顶点全体 (见图 7—58B)。三维二阶格子点集合 (记作 $\{3, 2\}$) 就是将 3 条边各二等分, 由 3 个顶点及 3 边中点全体组成的集合 (见图 7—58C)。三维三阶格子点集合 (记作 $\{3, 3\}$) 就是将 3 条边各三等分, 过各分点画与另两边平行的直线, 由平行线所交而形成的格点 (共 10 个) 组成的集合 (见图 7—58D)。依此类推, 但阶数通常到 3 为止就够了。

$\{3, 3\}$ 格子点坐标如表 7—35 所示。

表 7—35 三因子三阶单纯形格子点设计计划表

	x_1	x_2	x_3	
1	1	0	0	Vertex (顶点)
2	0	1	0	Vertex
3	0	0	1	Vertex
4	0.667	0.333	0	Double blend (双混)
5	0.333	0.667	0	Double blend
6	0.667	0	0.333	Double blend
7	0.333	0	0.667	Double blend
8	0	0.667	0.333	Double blend
9	0	0.333	0.667	Double blend
10	0.333	0.333	0.333	Center (中心)



一般来说, p 维 d 阶格子点集 $\{p, d\}$ 中, 共包含 $\binom{p+d-1}{d}$ 个点 (注: 这是现在国际上更通行的组合数记号, 是从 $p+d-1$ 个元素中取 d 个元素的组合数)。例如, $p=3, d=2, N=6$; $p=3, d=3, N=10$ 。

在这种类型试验结果下, 将要拟合的方程与阶数还有关, 这里就不详细讨论了。

这些设计方法是普通的未增广 (unaugmented) 的设点方法, 其示意图见图 7—58。

如果有条件, 可以加密些取点, 从而拟合更高阶方程, 我们可以使用增广设计 (augmented) 方法, 其设点的示意图参见图 7—59。

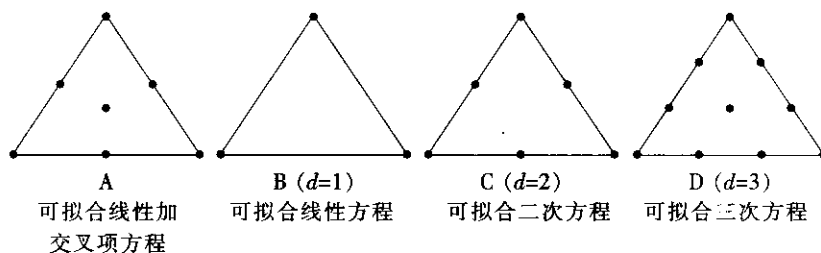


图 7—58 单纯形重心及单纯形格点设计示意图 (未增广情形)

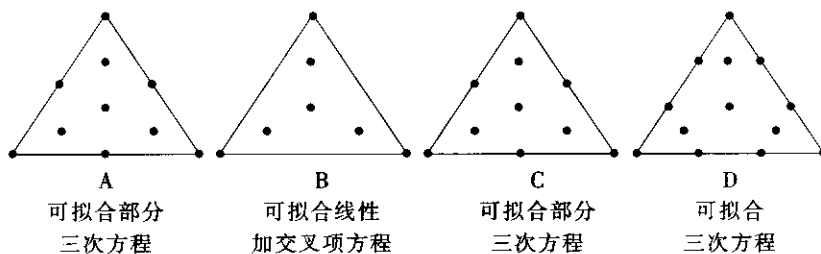


图 7—59 单纯形重心及单纯形格点设计示意图 (增广情形)

为了在单纯形格点设计法中比较不同的因子数 (维数) 和阶数的含义, 特将 4 组单纯形格点设计法的图一起列出, 见图 7—60。

第三种: 极端顶点设计法 (extreme vertices design)

在很多实际的混料设计问题中, 各分量常要受到上下界的限制。对这类问题有一种最简便的极端顶点设计法。各限制平面的交点处称为极端顶点。利用极端顶点集所构成的混料试验方案称为极端顶点设计。见图 7—61。

我们准备讨论计算顶点的细节, 因为在实际工作中, 只要输入约束条件, 计算机可以自动形成全部顶点坐标。

下面介绍用 MINITAB 软件对混料试验方案计划表的形成步骤。

从 DOE-Mixture>Create Mixture Design 入口后, 可以选择 Simplex Centroid (单纯形重心法), Simplex Lattice (单纯形格点法) 或 Extreme Vertices (极端顶点法) 三者之一, 再填写因子个数后, 即可得到相应设计。

选定 Simplex Centroid (单纯形重心法) 后, 在 “Design” 对话框中, 要进一步加选 “要增广” (Augmented) 或 “不要增广” (Unaugmented), 并要注明在哪些类型的点上要

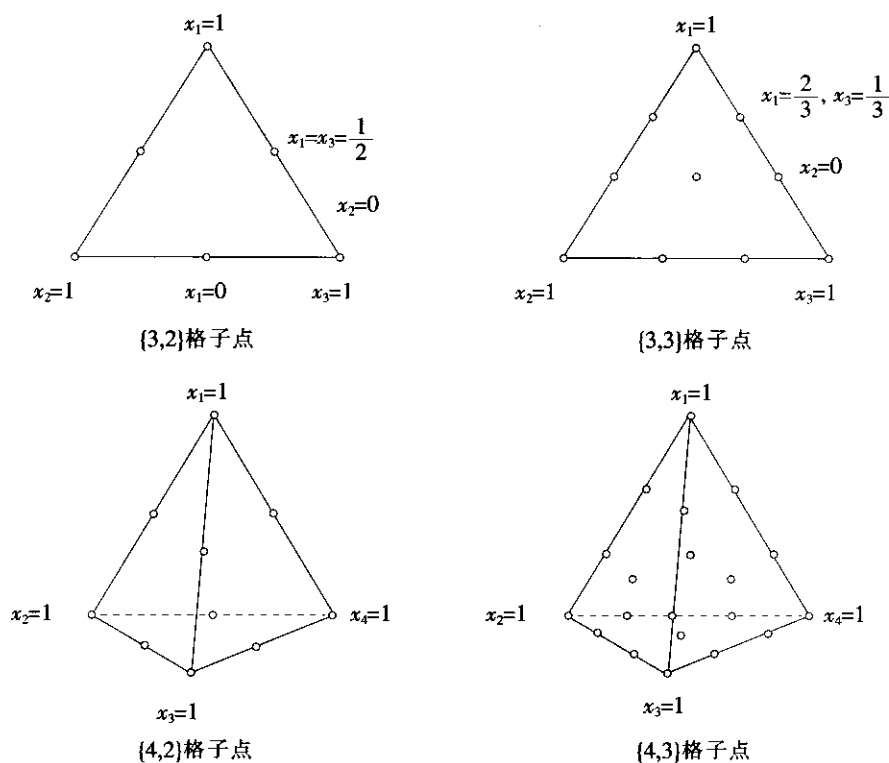


图 7—60 单纯形格点设计法示意图

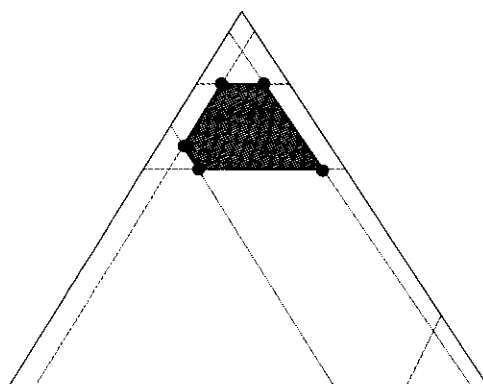


图 7—61 极端顶点设计法示意图

做重复，这些类型分“顶点”（Vertex）、“二阶点”（Double Blend）、“三阶点”（Triple Blend）、“中心点”（Center Point）等等。

选定 Simplex Lattice（单纯形格点法）后，在“Design”对话框中，要进一步指定阶数（Degree Lattice），加选“要增广”（Augmented）或“不要增广”（Unaugmented），并要注明在哪些类型的点上的要做重复，这些类型同样分“顶点”（Vertex）、“二阶点”（Double Blend）、“三阶点”（Triple Blend）、“中心点”（Center Point）等等。

选定 Extreme Vertices (极端顶点法) 后, 在 “Design” 对话框中, 也要进一步指定阶数 (Degree Lattice), 加选 “要增广” (Augmented) 或 “不要增广” (Unaugmented), 并要注明在哪些类型的点上要做重复。这里最重要的是要在 “Component” 对话框中, 进一步给出各分量的直接约束条件和 (或) 线性约束条件不等式。

7.7.3 混料试验的分析

下面举例说明混料设计的计划及分析步骤。

图 7-11 新型高强度合金钢配方设计问题

新型高强度合金钢配方设计问题

在新型高强度合金冶炼技术中, 关键是添加剂的配比问题。假设添加剂总量占合金总量的 2%, 而添加剂由 A, B, C 及 D 共 4 种成分构成。试安排一个混料试验设计, 求出 4 种分量的配比使断裂强度 y 达到最大。

本问题采用四因子二阶单纯形格子点设计, 共进行 10 次试验, 其试验结果见表 7-36。

表 7-36 新型高强度合金钢配方设计试验结果表

StdOrder	RunOrder	PtType	Blocks	A	B	C	D	y
1	9	1	1	1.0	0.0	0.0	0.0	142
2	6	2	1	0.5	0.5	0.0	0.0	142
3	4	2	1	0.5	0.0	0.5	0.0	138
4	2	2	1	0.5	0.0	0.0	0.5	145
5	5	1	1	0.0	1.0	0.0	0.0	136
6	10	2	1	0.0	0.5	0.5	0.0	133
7	3	2	1	0.0	0.5	0.0	0.5	135
8	1	1	1	0.0	0.0	1.0	0.0	133
9	7	2	1	0.0	0.0	0.5	0.5	133
10	8	1	1	0.0	0.0	0.0	1.0	132

利用 MINITAB 软件中混料试验分析功能, 从 DOE-Mixture-Analyze Mixture Design 入口, 在最初的模型选项中, 先选定全模型 (A, B, C, D 及 6 项二阶交互项)。但实际上, 由于只有 10 个观测值, 只能刚好估计出全部效应, 因而没有误差估计项。删除效应最小的三项 AC, BD, CD 后, 重新拟合回归方程, 得以下结果:

Regression for Mixtures: y versus A, B, C, D**Estimated Regression Coefficients for y (component proportions)**

Term	Coef	SE Coef	T	P	VIF
A	142.152	0.492 2	*	*	1.450
B	136.319	0.492 2	*	*	1.450
C	133.240	0.456 0	*	*	1.244
D	132.407	0.456 0	*	*	1.244
A * B	11.059	2.570 3	4.30	0.023	1.412
A * D	30.882	2.556 9	12.08	0.001	1.397
B * C	-7.118	2.556 9	-2.78	0.069	1.397

S=0.540 82 PRESS= *

R-Sq=99.55% R-Sq (pred)=*% R-Sq (adj)=98.64%

Analysis of Variance for y (component proportions)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	6	192.023	192.022 5	32.003 8	109.42	0.001
Linear	3	145.500	72.836 8	24.278 9	83.01	0.002
Quadratic	3	46.523	46.522 5	15.507 5	53.02	0.004
Residual Error	3	0.877	0.877 5	0.292 5		
Total	9	192.900				

从上述结果可以看出, 4 个主效应和 3 项二阶交互效应都是显著的 (显著水平取 0.1)。下面的图 7—62 为效应轮廓图 (response trace plot), 可以显示各单个因子的效应。本例题中, 看出欲使 y 增大, A 应取稍大些值 (例如 0.6)。

从下面的等高线图 (见图 7—63) 及曲面图 (见图 7—64) 上可以看出整个变化趋势。令 $B=0$, $C=0$ 将会使 y 取最大值。

为了准确地求出能使强度最大的设置, 使用响应变量优化器, 设定目标 (Goal) 为 Maximize, 取 Lower 值为已实现的 140, 取 Target 为 160 (Upper 不填写), 则可得到下列结果 (见图 7—65)。

最后可知, 当 A 取 0.663 6, D 取 0.336 4 时, 可得最大强度 145.77。这里自变量的设置是相对的比例, 注意到所给条件, 4 种分量合计占 2%, 因此, 分量 A 应占全部原料的 $2\% \times 0.663 6 = 1.327 2\%$, 分量 D 应占全部原料的 $2\% \times 0.336 4 = 0.672 8\%$ 。由于本例题中只做了 10 次试验, 回归方程的阶数较低, 未含各变量的二次项, 区域内部无极大值。如果可以认定在内部区域内 (而不是边界上) 将有极大值时, 应该选用更多试验点, 拟合更高阶方程。

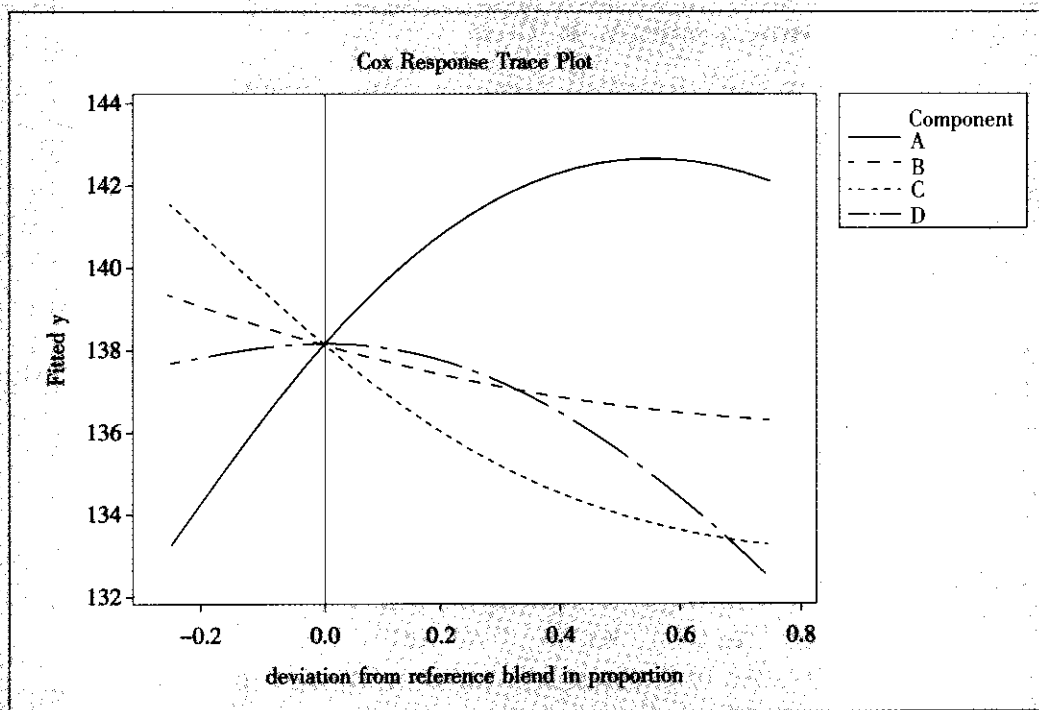


图 7—62 混料设计的效应轮廓图

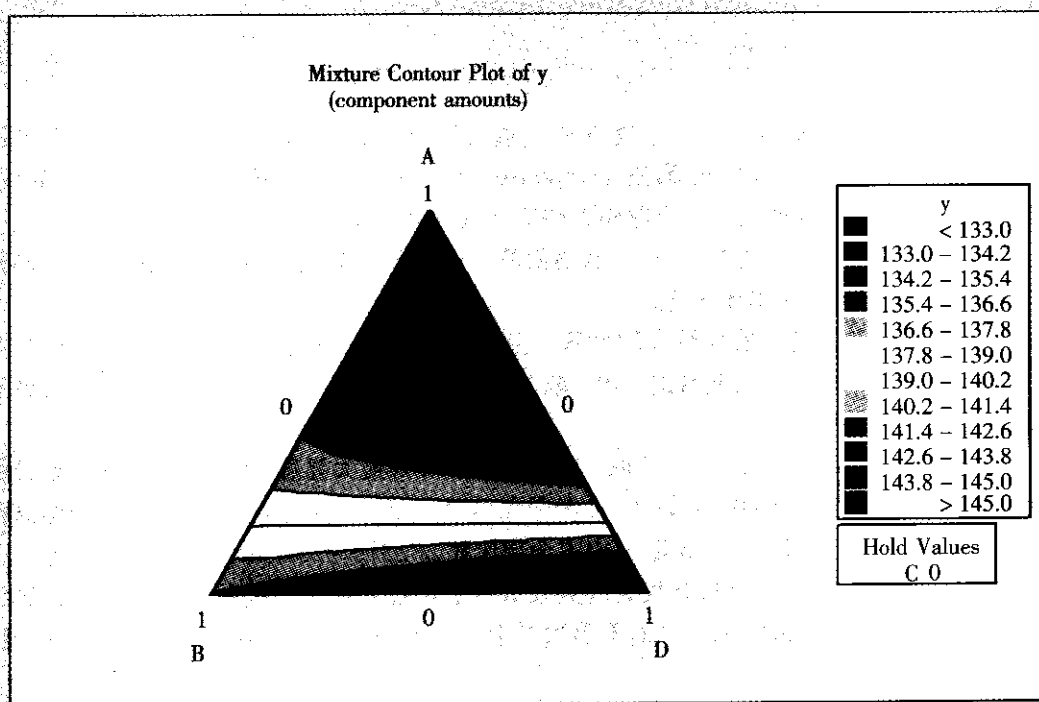


图 7—63 混料设计响应变量的等高线图

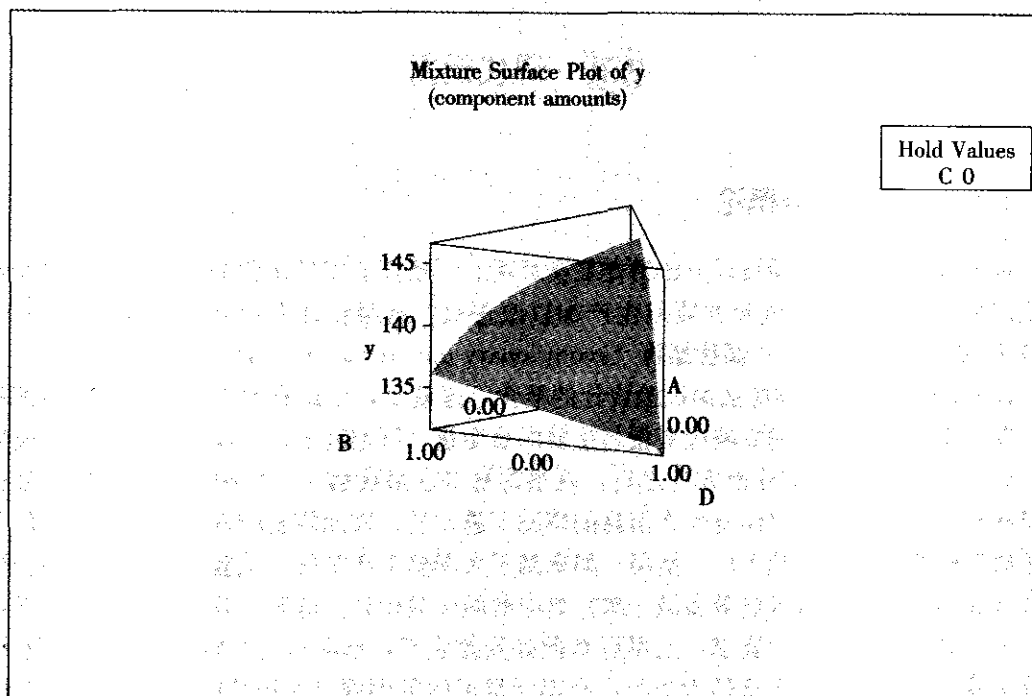


图 7—64 混料设计响应变量的曲面图

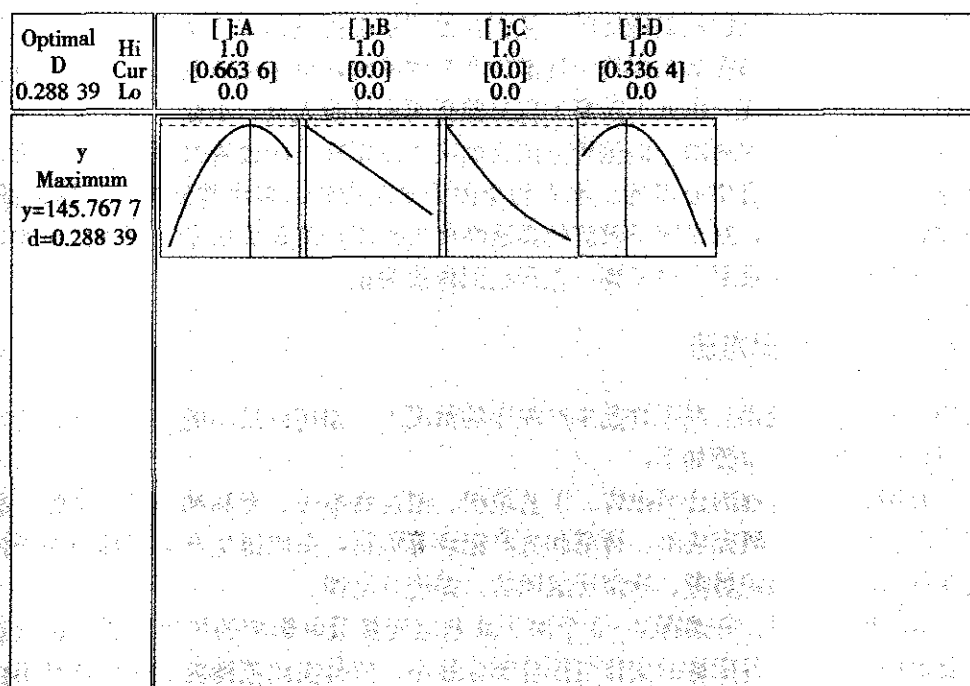


图 7—65 混料设计的响应变量优化图



7.8 调优运算

7.8.1 调优运算概论

前面几节介绍的试验设计方法都是在新环境、新条件下安排试验的。如何在不改变当前生产环境条件下进行试验并能对生产加以改进呢？试验设计大师博克斯（G. Box）于1957年提出一种方法——**调优运算**（evolutionary operation, EVOP）。

7.5节已经介绍了响应曲面试验设计的概念，这就是，以最有效的方式安排一些试验点，获得有关信息后，建立响应变量与重要的自变量（可控因子）之间的回归方程（通常指二阶以上）。得到响应曲面方程之后，就很容易寻求最优解了。但响应曲面试验设计通常是在对试验区域内几乎没有什么信息的情况下进行的，因而需要较多的试验点，一般是处于研发阶段或小型试验工厂中使用。如果到了大规模生产阶段，试验程序相对复杂，试验成本又高，通常不做（或最多做一次）响应曲面试验设计。这时，只要可以确认在某点上已达到我们对生产的要求，就可以在此处安排生产，不必再进行响应曲面试验设计。调优运算就是以当前生产条件为中心，在周围很近的范围内（过程性能可能有微小变化，但要保证不会出现任何问题）选定并安排一批简单的重复试验，对这些试验结果进行分析。分析的结论可能有两种：或者断言对生产条件的微调会使生产有改进；或者断言在一定的精度内目前生产条件已经是最优的，从而在当前不必再进行改进生产的尝试。调优运算的目的在于使生产过程在加工出符合标准的产品的时候，也生成有关其自身的信息。过程因子只允许有小的改变，因此不需要专门试验费用和试验人员，不会对生产有任何大的干扰，因而也不会有任何风险，而这种方法又确实可以发现可能改善过程性能的因素。其方法通常是安排一个具有中心点的二水平的全因子试验设计，多次重复进行，形成循环，在分析比较结果数据中，对于因子效应的显著性作出判断，将显著因子水平加以调整，使得生产获得改进。不断重复上述步骤，直到达到最优为止。

7.8.2 调优运算方法

我们所面对的问题是，在不扰乱生产秩序的情况下，如何以最小的工作量和风险来改进一个过程的性能。其步骤如下：

（1）查阅关于该过程的公司报告、工艺说明、制造商报告、专利和公开的文献、教科书和技术百科全书等，研究成本、得率和生产记录等资料，在取得生产管理部门领导的同意和支持的条件下，组织黑带、绿带形成团队，并进行培训。

（2）确定响应变量，并选择2~3个对于此响应变量最具影响的可控因子。选定这些变量变化范围及水平，各变量的变化范围应充分地小，使得在此范围内的任一点都不会让响应变量值显著变坏。在此范围内选取的两个水平通常是在目前生产条件两侧对称取值。

（3）进行第1周相的第1次循环。将现有生产条件作为中心点（编号为0）进行试验，将其他因子按设定水平进行一次 2^2 或 2^3 全因子试验，将结果记录在调优运算计算表中。

(4) 进行第1周相的第2次循环。重复第1次循环步骤,将结果记录在调优运算计算表中,并比较两次循环的平均值、标准差,计算出主效应及交互效应,计算出“均值变化(change-in-mean, CIM)效应”,再计算出对平均值、效应及 CIM 的误差限。

(5) 判断各效应是否显著。若有某项效应显著,则对各水平均值进行比较,选定一个更好水平作为新的中心点,确定新的范围及水平,开始第2周相(即下一个周相)第1次循环;若无任何效应显著,则进行本周相下次循环。第3次循环、第4次循环……若某次循环后出现显著效应,则选定新中心,开始新周相。若循环至6~10次仍无显著效应,则可认为对于所选因子而言,本中心点为最优条件,停止试验;或换选其他因子,重新进行整个步骤。总之,无显著效应就在本周相内进行下一次循环;若有显著效应就选定新中心点,开始新周相内的第1次循环。

由于 EVOP 本身并未使用任何新概念及新方法,在此我们就不给出更多细节了。有兴趣者可以参看有关参考书。虽然每次微小改进可能并不引人注目,但在黑带的指导下,累积改进效果可能是巨大的,而且由于每次改进都是渐进的,这保证了 EVOP 中过程不可能“从悬崖上跌落”。再者,黑带应帮助设立过程的 EVOP 流程,并且训练需要经常使用它的操作人员和监管人员。如果出现问题,也可以用周期性访问的方式提供现场支持。黑带及其团队决定在把改进的流程传授给操作人员之后,还要注意保持他们成果的方法,同时要通过研究基本的过程和变动产品设计来获得更大的改进。

本章小结

在改进阶段中,试验设计是最重要的统计工具。从设计的类型来看,有因子试验设计(包括全因子试验设计及部分实施因子实验设计)、响应曲面设计、稳健参数设计和混料设计。从六西格玛改进策略来看,首先应将有可能影响响应变量的大量因子用部分实施的因子设计进行筛选(这时不必过多考虑交互作用问题,取二水平加上若干中心点足够);然后,当因子个数不超过5时,就可以进行全因子试验,更细致地考虑因子间的交互作用。如果不能判定目前的试验区域是否已经达到最优区域,可以试着沿最速上升方向前进,直到达到最优区域。在此区域,用因子设计法所得到的线性回归方程可能出现失拟(即响应变量呈现明显的弯曲),应增加试验点,用响应曲面方法拟合一个二次回归方程,然后求出最优解。这对于望大或望小目标的问题特别有效。如果问题是望目特性的,则可选用稳健参数设计方法,效果可能更好。对于混料类型问题,则应选用混料设计。已有较好基础而又不宜任意改变因子水平时,可以选用对各因子只进行微调的调优运算。

控制阶段是项目团队维持改进成果的重要步骤。作为 DMAIC 过程的最后一个阶段，控制十分关键。要保持改进的成果，需要将改进阶段对流程的修改或新的流程作业指导书纳入作业标准和受控的文件体系，并建立过程控制系统。

控制阶段的主要目的是避免回到旧的习惯和程序，对人们的工作方式形成长期影响并得以保持。我们不仅需要测量和监视结果，还要不断说服和推销观念，两者都是必要的。六西格玛团队控制细节应该包括：

- 建立监视过程，明确已经取得的改进；
- 制定应变方案；
- 确定关键控制点、控制参数和控制方法；
- 形成新的程序文件或作业标准。

8.1 控制阶段概述

控制的目的在于保持项目取得的成效并实现持续改进。没有控制的努力，改进的过程很容易恢复到以前的状态，破坏团队认为已经实现的收益，影响团队的士气，对组织的六西格玛管理实施造成严重的伤害。在控制阶段有三个要素：

- (1) 过程改进成果的文件化；
- (2) 建立过程控制计划；
- (3) 持续的过程测量（控制）。

8.1.1 过程改进成果的文件化

维持一个稳定的、可预测的过程需要对个人和组织两方面的约束。在很多企业，作业程序文件没有得到有效的执行，一个重要的原因在于企业对员工严格执行程序缺乏明确的纪律，当程序文件和作业执行成为两张皮的时候，过程控制就成了一句空话。

1. 明确纪律

同样的管理模式在不同的企业往往会产生不同的结果，其原因主要在于企业在执行力上的不同。员工能否严格按照文件、制度、标准去工作是衡量执行力的一个重要方面。因此，企业在员工招聘、培训、跟踪、考核评估和奖励的时候，要考虑员工是否能够严格执行纪律。要求员工执行不能仅靠空洞的说教，偶尔的“全体集合”的会议通常是浪费时

间,在这些会议上管理人员会力劝员工“更加负责任”、“更加注意细节”。为了避免这种情况,你的组织应该明确确定一个“过程负责人”,他对过程负责——不仅仅指他要完成自己的工作——而且要保证整个过程是被持续监督、研究和改进的。

要求所有员工做到严格执行制度绝非易事,除非员工都能理解监督、控制和改进的原因和益处。不理解这一点,收集数据和绘制图表将被视为不必要的烦琐工作。因此,企业在实施六西格玛时,对全体员工进行六西格玛的基础培训,让员工清楚自己的职责是必要的。

2. 改进过程文件化

一旦改进方案实施后看起来是有效的、可维持的,要想使改进能真正坚持下去,就必须使已经改进的过程详细记录在册,并形成操作规程及固定程序,将所有这些新措施文件化。以下是一些在改进过程文件化中应该注意的问题:

(1) 要用最简洁明确的语言,运用过程图、程序文件、图表、照片、录音带等形象化的辅助工具将过程要求逐一解释清楚准确,尽可能使用通俗易懂的表达方式,避免运用过于专业的术语或专有名词,要保持文件的简单化。

(2) 要特别注意在改进初期显露出的问题、潜在的困难,制定应急预案,并在文件中发出警告信号。

(3) 如果过程太复杂,操作规程文件不易马上全部确定,可以吸收过程负责人、技术人员、操作人员一起讨论,反复论证,直到文件细化到可以准确顺利执行,并在实施过程中适时地进行修订。

(4) 让那些将来会应用这些文件的人帮助起草文件,这样能使文件更实用。向所有有关人员展示概略草稿,并请他们提出建议。

8.1.2 建立过程控制计划

在实际中,要对过程实施有效的监督和控制,需要建立一套系统的过程控制计划。一个好的过程控制计划一般包括以下内容:

(1) 正确的过程图。过程负责人需要掌握过程中各种行动和决策的每个细节,正确、详细的过程图是必不可少的。

(2) 行动警报。首先要制定明确的过程应变计划,过程应变计划可以清楚标明在过程中的哪些地方可以准确测量过程输入、过程操作和过程输出的关键数据。一旦开始显示质量下降,过程负责人会根据过程应变计划做出反应。例如,连续三个顾客的订单没有按时完成,就需要在生产线管理方面启动意外事故处理计划,如外包。

(3) 紧急事件抢险。一旦发出了行动警报,已经启动的应急预案或补救计划就显得十分重要,员工们事先已经了解这些紧急操作步骤,可以毫不犹豫、有条不紊地圆满完成相关任务。

(4) 持续改进的计划。除了跟踪过程中问题的发生,过程应变计划也为六西格玛团队彻底检查过程的薄弱环节提供了基础。一旦建立起来足够的过程应变计划,在实践中又积累了足够的经验,就可以进一步发现改进机会,这就有了广泛的项目资源,未来的六西格玛团队可以针对这些项目进行改进。

另一个有助于过程控制的工具是过程仪表板。大家可能熟悉那些应用于管理执行阶段



的仪表板，它们汇集了大量代表过程顺利运转的指示灯。在过程控制中仪表板也是同样的原理，需要设置一些过程反应、过程重要操作状况的指示灯。通常包括输出项的测量（例如顾客满意度）、关于生产速度或效率的测量（如周期时间）以及一些过程质量的测量（如缺陷水平）。

8.1.3 实施持续的过程测量和控制

在 DMAIC 的开始阶段，团队通过测量来定义问题、统计缺陷、分析变化的原因，并测量改进是否成功。在控制阶段，团队仍然要努力关注测量，必须重新明确地找出需要进行测量的关键点，不仅要包括对响应变量的测量，而且往往需要扩展到对生产条件中关键因子的测量，团队寻求测量和控制时有三个关键。

1. 完善过程的 SIPOC 图

对已经改进的过程，完善其过程 SIPOC 图，其中要特别注意“顾客的需要”这个关键问题。六西格玛团队必须实时地测量过程输出是否与顾客需要相一致，要做到及时发现产品缺陷或过程变化。

2. 对过程关键变量进行控制

确定出测量哪个上游的过程变量会与改进紧密相连，持续测量这个过程将可以预测下游的结果。控制了过程的所有关键变量，才可以确保过程的输出能够满足要求。例如：如果“准时”是一个关键顾客需求，并且过程测量显示其中的两个关键过程步骤正变得越来越长，团队就需要在它导致结果延迟（即出现缺陷）之前寻找和评估出现这种趋势的原因。

3. 关注关键输入测量

关键输入测量能帮助预测过程关键步骤的运行和关键输出结果的质量。在控制阶段中可能会花一段时间来建立对一系列变量进行测量的测量链，这是很值得的工作，测量链最能体现过程的特征。一旦测量链建立，它就变成了计分卡的根据，任何负责改进过程的人员都可以在固定的基础上（例如，每天或每月）用计分卡来监督。

确定最佳测量项以后的一个重要问题是如何根据收集的数据建立过程控制方法，在六西格玛工具集里，有许多过程控制方法，其中控制图是最为常用的工具。

8.2 常规控制图

20 世纪 20 年代，贝尔电话实验室成立了以休哈特为首的过程控制研究组和以道奇为首的产品控制研究组。经过研究，休哈特提出了过程控制理论和监控过程的工具——控制图。世界上第一张控制图是休哈特在 1924 年 5 月 16 日提出的 p 控制图。道奇与罗米格则提出了抽样检验理论和抽样检验表。这两个研究组研究工作的影响是深远的，在他们之后，虽然有成百上千篇论文出现，但至今仍未脱其窠臼。

休哈特创建的过程控制理论，以 1931 年出版的名著《工业产品质量的经济控制》为标志，该部名著在其出版半个世纪以后的 80 年代在美国再版。从质量控制理论的发展史来看，休哈特可称为统计过程控制理论（statistical process control, SPC）的奠基人。



统计过程控制就是应用统计技术对过程中的各个阶段进行监控,从而达到改进与保证产品质量的目的。这里统计技术泛指任何可以应用的数理统计方法,以控制图理论为主。统计过程控制理论的研究虽是从加工过程开始的,但其研究成果适应于各种过程,如设计过程、管理过程、生产过程、服务过程。

当年休哈特提出的几种控制图,统称为休哈特控制图,我国亦称为常规控制图。在休哈特之后,又提出了诸多的控制图,其中累积和控制图、指数加权滑动平均控制图等都是值得注意的,本章随后将予以介绍。

8.2.1 控制图的基本概念

1. 波动理论

人们对波动有以下认识:

- (1) 过程中有许多导致波动的因素存在;
- (2) 每种因素的发生是随机的且无法预测,然而这些因素都影响着过程的输出,即质量特性;
- (3) 质量特性有波动是正常现象,无波动是虚假现象或是测量仪器分辨力太低;
- (4) 彻底消灭波动是不可能的,但减少波动是可能的;
- (5) 控制过程就是要把波动限制在允许的范围内,超出范围就要设法减少波动并及时报告,迟到的报告有可能引起损失,是失职行为。

导致质量特性波动的因素根据来源的不同,可分为人(man)、机(machine)、料(material)、法(method)、环(environment)、测(measurement)六个方面,简称为5M1E。从对质量影响的大小来分,质量因素又可分为偶然原因(common cause)和可查明原因(assignable cause)两大类。偶然原因,简称为偶因,又称为一般原因。可查明原因,又称为特殊原因(special cause)或系统性原因,俗称异常原因,简称为异因。国家标准 GB/T4091—2001《常规控制图》对此做了如下解释:

控制图理论认为存在两种变异。第一种变异为随机变异,由“偶然原因”(又称为“一般原因”)造成。这种变异是由种种始终存在的且不易识别的原因所造成,其中每一种原因的影响只构成总变异的一个很小的分量,而且无一构成显著的分量。然而,所有这些不可识别的偶然原因的影响总和是可度量的,并假定为过程所固有。消除或纠正这些偶然原因,需要管理决策来配置资源,以改进过程和系统。

第二种变异表征过程中实际的改变。这种改变可归因于某些可识别的、非过程所固有的,并且至少在理论上可加以消除的原因。这些可识别的原因称为“可查明原因”或“特殊原因”。它们可以归结为原材料不均匀、工具破损、工艺或操作的问题、制造或检测设备的性能不稳定等。

偶然原因引起质量的偶然波动(简称偶波),异常原因引起质量的异常波动(简称异波)。偶波是不可避免的,但对质量的影响微小,故可把它看作背景噪声而听之任之。异波则不然,它对质量的影响大,且采取措施不难消除,故在过程中异波及造成异波的原因是关注的对象,一旦发生,就应该尽快找出,采取措施加以消除,并纳入标准,保证它不

再出现。将质量因素区分为偶然原因与异常原因,质量波动区分为偶然波动与异常波动,并分别采取不同的处理策略,这是休哈特的贡献。

2. 控制图的基本原理

偶然波动与异常波动都是质量的波动,如何能发现异常波动的到来呢?经验与理论分析表明,当生产过程中只存在偶然波动时,质量特性值将形成某种典型分布。例如,在车制螺丝时,其直径形成正态分布。如果除去偶然波动以外还有异常波动,则此异常波动将迭加在偶然波动形成的典型分布上,故质量特性值的分布必将偏离原来的典型分布。因此,根据典型分布是否偏离就能判断异常波动是否发生,也即是否出现了异常因素,而典型分布的偏离可由控制图检出。

正态分布有一个事实在质量控制中经常要用到,即不论 μ 与 σ 如何取值,落在 $[\mu-3\sigma, \mu+3\sigma]$ 范围内的概率为99.73%。这是经过严格概率计算得到的精确值,参见图8—1。因而,落在 $[\mu-3\sigma, \mu+3\sigma]$ 范围外的概率为 $1-99.73\%=0.27\%$,而落在大于 $(\mu+3\sigma)$ 一侧或者小于 $(\mu-3\sigma)$ 一侧的概率为 $0.27\%/2=0.135\%\approx 1\%$ 。休哈特就是根据正态分布的这一性质构造了休哈特控制图,亦称为常规控制图。

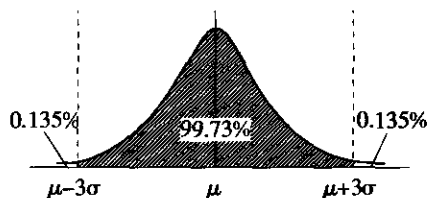


图 8—1 正态分布的性质

将图 8—1 按顺时针方向转 90°, 如图 8—2 (a) 所示。由于图中数值上小下大不符合常规, 故再将图 8—2 (a) 上下翻转 180°, 成为图 8—2 (b), 这样就得到了一张控制图, 具体说是一张单值控制图, 参见图 8—3。图中的 $UCL=\mu+3\sigma$ 为上控制界, $CL=\mu$ 为中心线, $LCL=\mu-3\sigma$ 为下控制界。

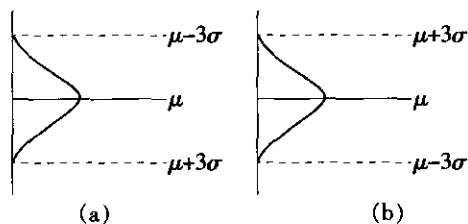


图 8—2 控制图的形成

在图 8—3 所示的车制螺丝事例中, 由于发生了车刀严重磨损的异常因素, 导致螺丝直径的分布偏离了原来的正态分布而向上移动, 于是点子超出上控制界的概率大为增加, 从而点子频频出界, 表明过程存在异常波动。因此, 控制图的控制界限就是区分偶然波动与异常波动的科学界限。只有偶然因素没有异常因素的状态, 称为统计控制状态 (state in statistical control), 简称稳态, 是控制阶段实施过程控制所追求的目标。

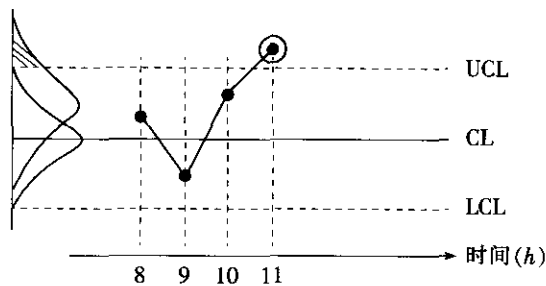


图 8—3 单值控制图

谨慎！很多人认为如果过程数据保持在控制界限以内，就是正在为顾客生产好的产品。这么想就错了！控制界限是通过过程数据计算出来的，是区分偶然波动与异常波动的科学界限，与顾客需求和产品服务的规格规范无关。例如，某汽车修理点应用控制图来监控换油时间，发现整个过程都没有发出异常信号，显示换油过程是受控的。然而，承诺顾客 30 分钟取车，而实际平均换油时间是 49 分钟。如何将换油时间稳定地控制在 30 分钟以内，达到对顾客的承诺正是六西格玛团队进行质量改进的一项目标。

3. 控制图的设计思想

利用控制图进行监控，不可避免地要面对两类错误。

第Ⅰ类错误，亦称为虚发警报的错误。过程正常而点子偶然超出控制界外，根据点出界就判异的原则，判断过程处于异常，于是就犯了第Ⅰ类错误，即虚发警报的错误。通常将犯第Ⅰ类错误的概率记为 α ，参见图 8—4。

第Ⅱ类错误，亦称为漏发警报的错误。过程已经出现了异常，但仍会有部分产品，其质量特性的数值会位于控制界内。如果抽取到这样的产品，点子就会落在控制界内，而不能判断过程出现了异常，从而犯了第Ⅱ类错误，即漏发警报的错误。通常犯第Ⅱ类错误的概率记为 β ，参见图 8—4。

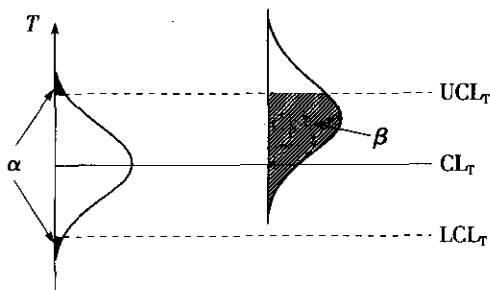


图 8—4 控制图的两类错误

一般，常规控制图的控制界限为：

$$UCL_T = \mu_T + 3\sigma_T$$

$$CL_T = \mu_T$$

$$LCL_T = \mu_T - 3\sigma_T$$

式中， μ ， σ 为统计量的总体参数，下标 T 泛指某统计量。可见，常规控制图的设计思想

是先定 α ，再看 β 。休哈特为增强使用者的信心，把常规控制图的虚发警报概率 α 取的特别小，仅为 $\alpha=0.27\%$ 。优点是：过程处于统计控制状态时，大约 1 000 个点子中仅有不到 3 个点子会超出控制界限。然而，缺点是：漏发警报的概率会增加。因此，下面讨论判断过程是否出现异常的八种模式。

4. 异常波动的判断

国家标准 GB/T4091—2001《常规控制图》明确给出了判断异常波动的八种模式。下面对这八种模式一一进行分析。

上、下控制限分别位于中心线的上、下 3σ 距离处。为了便于应用这八种模式，将控制图分为 6 个区，每个区的宽度为 1σ ，如图 8—5 所示。6 个区的标号依次为 A, B, C, C, B, A，分别位于中心线两侧，关于中心线对称。

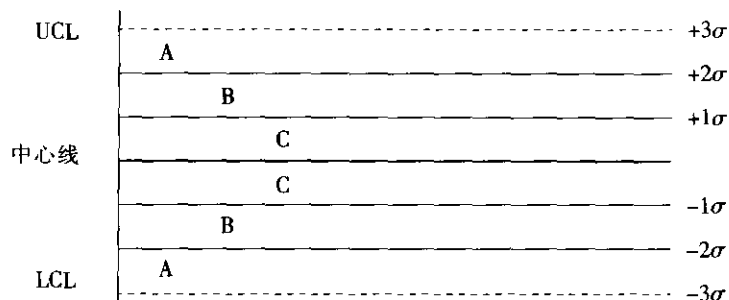


图 8—5 可查明原因的检验

(1) 模式 1。检验 1 给出了对控制图的基本解释：点出界就判异。当过程处于统计控制状态时，点子落在控制限内的概率为 99.73%，落在一侧控制限外的概率为 $(1 - 99.73\%) / 2 = 0.135\%$ 。

(2) 模式 2。当过程处于统计控制状态时，连续 9 点落在中心线同一侧的概率为：

$$P(\text{连续 9 点落在中心线同一侧}) = 2 \times (0.5)^9 = 0.003\ 906 = 0.390\ 6\%$$

即虚发警报的概率为 0.390 6%，第 I 类错误的概率为 $\alpha=0.390\ 6\%$ 。

在控制图中心线同一侧连续出现的点称为链，链可能位于中心线的上侧也可能位于下侧。注意：链必须由“连续”出现的点子构成。链中包含的点子数目称为链长。若链长 ≥ 9 ，判断出现了异常。

(3) 模式 3。当过程处于统计控制状态时，连续 6 点递增或递减的概率为：

$$P(\text{连续 6 点递增或递减}) = \frac{2}{6!} \times (0.997\ 3)^6 = 0.002\ 733 = 0.273\ 3\%$$

即虚发警报的概率为 0.273 3%，第 I 类错误的概率为 $\alpha=0.273\ 3\%$ 。

点子逐点上升或下降的状态称为倾向或趋势。注意，递减的下降倾向，后面的点子一定要低于或等于前面的点子，否则倾向中断，需要重新起算。递增的上升倾向也是同样。

(4) 模式 4。对于“连续 14 点中相邻点交替上下”这种模式，国家标准 GB/T4091—2001《常规控制图》的参考文献 (Nelson, L. S., Interpreting Shewhart \bar{X} Control Chart, *Journal of Quality Technology*, 17, No. 2, April 1985, pp. 114 - 116) 有这样一段描述：“例如：两台设备、两只纺锤、两位操作工、两台自动售货机，它们交替使用

所带来的系统影响就会引发这样的现象。在记录交替倒置的鸡蛋计时器流出的沙子的时间时,也会观察到这样的现象。因此,这种模式又名‘鸡蛋计时器效应’。点子数的选择依据蒙特卡洛(Monte Carlo)的试验结果,因为14点时虚发警报的概率(大约为0.004)与模式1基本一致。”

(5) 模式5。当过程处于统计控制状态时,连续3点中有2点落在中心线同一侧的B区以外的概率为:

$$\begin{aligned} & P(\text{点子落在中心线一侧B区以外的概率}) \\ &= \begin{cases} P(X > \mu + 2\sigma) = 1 - P(X \leq \mu + 2\sigma) = 1 - 0.9772 = 0.0228 \\ P(X < \mu - 2\sigma) = 0.0228 \end{cases} \\ & P(\text{连续3点中有2点落在中心线同一侧的B区以外}) \\ &= 2 \times C_3^2 \times (1 - 0.9772)^2 \times 0.9772 \\ &= 6 \times 0.0228^2 \times 0.9772 \\ &= 0.003048 = 0.3048\% \end{aligned}$$

即虚发警报的概率为0.3048%,第I类错误的概率为: $\alpha = 0.3048\%$ 。

(6) 模式6。当过程处于统计控制状态时,连续5点中有4点落在中心线同一侧的C区以外的概率为:

$$\begin{aligned} & P(\text{点子落在中心线一侧C区以外的概率}) \\ &= \begin{cases} P(X > \mu + 3\sigma) = 1 - P(X \leq \mu + 3\sigma) = 1 - 0.8413 = 0.1587 \\ P(X < \mu - 3\sigma) = 0.1587 \end{cases} \\ & P(\text{连续5点中有4点落在中心线同一侧的C区以外}) \\ &= 2 \times C_5^4 \times (1 - 0.8413)^4 \times 0.8413 \\ &= 10 \times 0.1587^4 \times 0.8413 \\ &= 0.005331 = 0.5331\% \end{aligned}$$

即虚发警报的概率为0.5331%,第I类错误的概率为 $\alpha = 0.5331\%$ 。

(7) 模式7。当过程处于统计控制状态时,连续15点落在中心线两侧的C区内的概率为:

$$\begin{aligned} & P(\text{连续15点落在中心线两侧的C区内}) \\ &= (1 - 0.1587 \times 2)^{15} \\ &= 0.00326 = 0.326\% \end{aligned}$$

即虚发警报的概率为0.326%,第I类错误的概率为 $\alpha = 0.326\%$ 。

初学者往往认为这样的点子形式看上去很不错。然而,这种“看上去不错”的点子形式依赖于所使用的缩放比例。所以,我们并不关心“看上去”的情况,而是关注于这种点子形式所隐含的非随机性。蒙特卡洛试验结果也指出以连续15点为模式7的临界值,因为15点的虚发警报概率与模式1基本相等。

模式7可用于检测分层问题。为了简化,假设子组大小为2,每个子组中的一个个体来自分布A,子组中的另一个个体来自分布B。如果分布A和分布B有不同的均值,那么,就使得基于极差的控制限间隔非常大。因此,并不是点子的变化不大,而是控制限太宽了。因此,应该注意:模式7的错误是经常出现的,模式7所发出的警报其原因很可能是分层问题。

(8) 模式 8。当过程处于统计控制状态时, 连续 8 点落在中心线两侧且无一在 C 区内的概率为:

$$\begin{aligned} P(\text{连续 8 点落在中心线两侧且无一在 C 区内}) \\ &= (0.1587 \times 2)^8 \\ &= 0.000103 = 0.0103\% \end{aligned}$$

即虚发警报的概率为 0.0103%, 第 I 类错误的概率为 $\alpha = 0.0103\%$ 。

如果考虑模式 7 描述的两个分布, 假设子组中的数据要么来自分布 A, 要么来自分布 B, 而不是同时来自这两个分布, 那么, 就会出现模式 8 所显示的情况。

5. 统计控制状态的判断

国家标准 GB/T4091—2001《常规控制图》给出了过程改进的策略, 如图 8—6 所示, 图中 PCI 为过程能力指数。

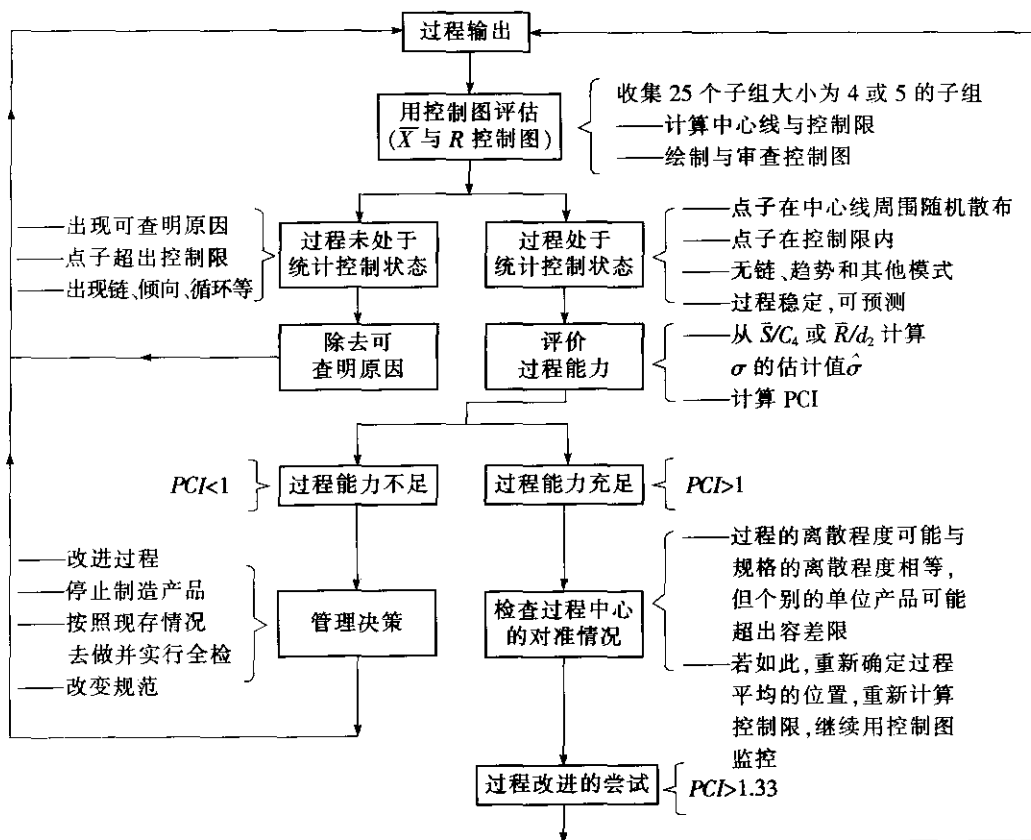


图 8—6 过程改进的策略

从图 8—6 可以看出, 对数据的要求是: 收集 25 个子组大小为 4 或 5 的子组。

过程处于统计控制状态的条件是:

- 点在中心线周围随机散布;
- 点在控制限内;
- 无链、趋势和其他模式;

- 过程稳定，可预测。

在下列情况下，可认为过程未处于统计控制状态：

- 出现可查明原因；
- 点子超出控制限；
- 出现链、趋势、循环等。

6. 分析用控制图与控制用控制图

控制图可分为分析用控制图与控制用控制图两个阶段。一个过程开始实施控制图时，几乎都不会恰巧处于统计控制状态，总存在着异常波动。如果就以这种状态下的参数来建立控制图，上、下控制限的间隔一定较宽，这会导致判断失误。因此，过程开始控制时，总需要将失控状态调整到统计控制状态。这就是分析用控制图的阶段。

分析用控制图阶段主要做到：

- (1) 调整过程，以便使过程受控；
- (2) 使过程能力指数 C_p 或 C_{pk} 能达到顾客要求。一般要求 $C_{pk} \geq 1.33$ 。

一旦项目进入控制阶段，即表明过程实现了上述两点，就可延长控制限作为控制用的控制图，进入控制用控制图阶段，以保持过程的统计控制状态。

7. 控制图的设计制作基本程序

(1) 确定少数重要的关键过程测量值，参考表 8—1 和图 8—7，决定所使用的控制图。根据过程数据的性质，常规控制图主要有两类控制图：一类是针对连续数据的计量控制图；另一类是针对离散数据的计数控制图。

(2) 采集数据，制作简单实用的数据采集表。即使控制图是借助计算机软件制作完成的，也必须进行这个步骤。对六西格玛团队来说，如何采集数据至关重要。对于初始建立控制图，一般至少需要采集 20 组以上的样本。

(3) 根据需要控制的统计量计算控制限。

(4) 绘制控制图。如果使用计算机软件，务必注意数据的正确性。如果手工绘制，请预先准备高约 6~8 公分，宽度能容纳所有数据的图表。可以使用一般的绘图纸或特别的统计过程控制作业表。

在许多时候，可借助统计软件（如 MINITAB）来做控制图。但如果是首次使用控制图，建议先手工绘制简单的控制图，以帮助团队成员了解控制图及其应用。

(5) 应用控制图，利用判断异常以及统计控制状态的准则，确定过程的状态。如果使用计算机软件，控制图上会自动显示过程的状态；如果手工绘制，需要我们自己完成对过程状态的判断，确定过程是否出现了异常；如果出于控制的目的建立控制限，描点后若有点超出控制限或点的排列有缺陷，应分析原因，找到原因后应剔除数据，重新计算控制限。

表 8—1 常规控制图

图表类型	适用数据	使用说明
$I-MR$ 图	单件的连续数据。标出单件的观测值以及移动极差，即相邻各件观测值之差。	I 图最容易绘制，可用于多种状况
$\bar{X}-R$ 图与 $\bar{X}-S$ 图	子组的连续数据。标出子组平均值和子组极差或子组标准差。	\bar{X} 图使得异常因素带来的波动更凸显

续前表

图表类型	适用数据	使用说明
p 图与 np 图	计件离散数据, 例如: 不合格品数、销售中的流失数等。	p 图: 允许样本大小变化 np 图: 样本大小不变
u 图与 c 图	计点离散数据, 例如: 缺陷点数、瑕疵数、数目较小的误差数或错误数等。	u 图: 允许样本大小变化 c 图: 样本大小不变

控制图的选用流程如图 8—7 所示。

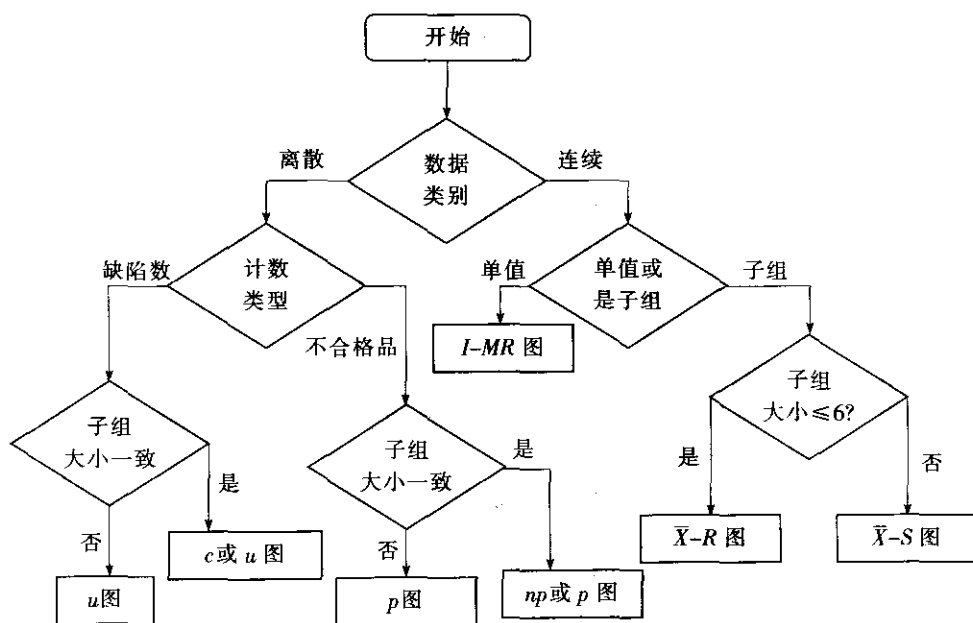


图 8—7 常规控制图的选择

8. 控制图判断准则的使用

国家标准 GB/T4091—2001《常规控制图》明确指出, 判断异常波动的八种模式适用于 \bar{X} 图和单值 (X) 图。这里假定质量特性 X 的观测值服从正态分布。

应该注意, 极差和标准差的分布并不是正态的, 尽管在计算控制限估计常数时, 对极差和标准差的分布作了近似正态性的假设, 这种假设对于检验决策程序而言还是令人满意的。因此, 对于极差 R 控制图, 国家标准 GB/T4091—2001《常规控制图》指出, 与控制限进行对比, 检查数据点是否有失控点以及有无异常的模式或趋势。对于极差数据中关于可查明原因的每一个征兆, 分析过程的运行, 以便找出原因, 进行纠正, 并防止再次出现。

对于不合格品率 p 控制图, 国家标准 GB/T4091—2001《常规控制图》指出, p 图用来确定在一段时间内所提交的平均不合格品百分数。该平均值的任何变化都会引起过程操作人员和管理者的注意。 p 图判断过程是否处于统计控制状态的判断方法与 \bar{X} 和 R 控制图相同。若所有子组点都落在试用控制限之内, 并且未呈现出可查明原因的任何迹象, 则称此过程处于统计控制状态。这种情况下, 取平均不合格品率 \bar{p} 为不合格品率 p 的标准值, 记为 p_0 。

综上所述, 国家标准 GB/T4091—2001《常规控制图》给出的八种模式, 对于 \bar{X} 图和单值 (X) 控制图, 完全适用; 对于极差控制图 and 标准差控制图, 因为存在极差和标准差的近似正态性假设, 所以可以近似使用; 对于不合格品率 p 控制图、不合格品数 np 控制图、单位缺陷数 u 控制图以及缺陷数 c 控制图, 同样因为存在着近似正态性假设。所以可以近似使用八种模式, 以判断异常原因是否出现以及何时出现。

8.2.2 计量控制图

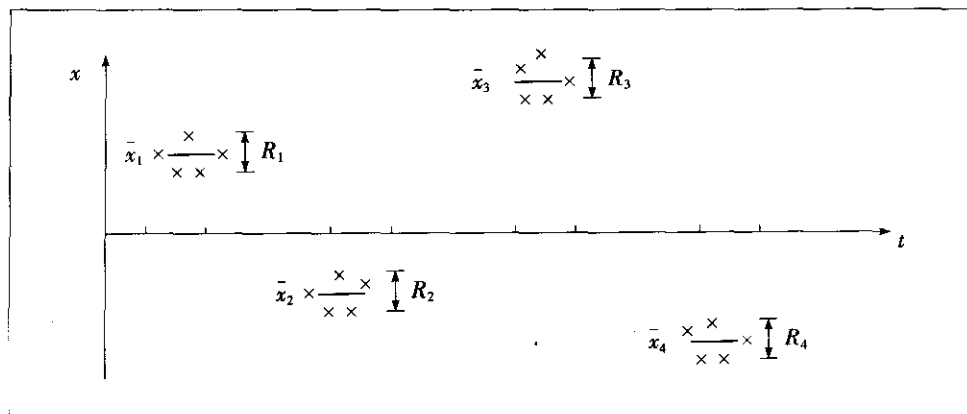
计量控制图的基本假定是质量特性 X 的观测值服从正态分布 $N(\mu, \sigma^2)$, 其中: μ 为正态分布的均值, σ 为正态分布的标准差, μ 与 σ 彼此独立。因此要控制计量值需要两张控制图: 一张用于控制 μ ; 另一张用于控制 σ 。根据样本大小 n 和用于估计 μ 与 σ 的统计量的不同, 计量控制图共有四对, 它们是:

- (1) 均值—极差图 ($\bar{X}-R$ 图): 精度尚可, 使用方便, $n=2\sim6$ 。
- (2) 均值—标准差图 ($\bar{X}-S$ 图): 精度最高, 计算量大, $n\geq 2$ 。
- (3) 中位数—极差图 ($\tilde{X}-R$ 图): 精度较差, 计算量小, $n=2\sim6$ 。
- (4) 单值—移动极差图 ($I-MR$ 图): $n=1$ 。

1. 合理子组原则

合理子组原则是休哈特提出的控制图基础理论之一。在抽取样本时要使: 组内波动仅由偶然原因引起; 组间波动主要由异常原因引起。休哈特称这样得到的样本为子组。由于子组内无异常波动, 故用其估计标准差可得到较为精确的估计, 且估值一般较小。进而上下控制限的间隔也不大, 从而检出异常波动较为灵敏。

为了实现合理子组原则, 一个最简单的方法就是在短时间内抽取一个子组所需的全部个体, 或者对连续生产产品进行“块抽样”。由于抽样间隔时间短, 就尽可能避免了异常因素进入子组。具体可见图 8—8。



注: \bar{x}_i 与 R_i 为第 i 个子组的均值与极差。

图 8—8 合理子组原则示意图

2. 计量控制图的控制限

计量控制图有四对, 各自中心线 CL、上控制限 UCL、下控制限 LCL 的计算公式列于表 8—2。其中 $\bar{X}-R$ 图最为常用。



表 8—2 计量控制图的中心线和上、下控制限

控制图的名称与符号		CL	UCL	LCL
均值—极差控制图 ($\bar{X}-R$ 图)	\bar{X} 图	$\bar{\bar{x}}$	$\bar{\bar{x}}+A_2\bar{R}$	$\bar{\bar{x}}-A_2\bar{R}$
	R 图	\bar{R}	$D_4\bar{R}$	$D_3\bar{R}$
均值—标准差控制图 ($\bar{X}-S$ 图)	\bar{X} 图	$\bar{\bar{x}}$	$\bar{\bar{x}}+A_3\bar{s}$	$\bar{\bar{x}}-A_3\bar{s}$
	s 图	\bar{s}	$B_4\bar{s}$	$B_3\bar{s}$
中位数—极差控制图 ($\tilde{X}-R$ 图)	\tilde{X} 图	\bar{M}_e	$\bar{M}_e+A_4\bar{R}$	$\bar{M}_e-A_4\bar{R}$
	R 图	\bar{R}	$D_4\bar{R}$	$D_3\bar{R}$
单值—移动极差控制图 ($I-MR$ 图)	I 图	$\bar{\bar{x}}$	$\bar{\bar{x}}+E_2\bar{R}$	$\bar{\bar{x}}-E_2\bar{R}$
	MR 图	\bar{R}	$D_4\bar{R}$	$D_3\bar{R}$

注：(1) 表中 $A_2, A_3, D_3, D_4, B_3, B_4$ 与样本量 n 有关，可在附表中查得。

(2) 表中 $E_2=3/d_2$ ，对于单值—移动极差控制图，由于采用移动极差，样本含量为 2， $d_2=1.128$ ， $E_2=2.66$ 。

(3) 表中 A_4 的值见表 8—3。

表 8—3 A_4 的值

n	2	3	4	5	6	7	8	9	10
A_4	1.88	1.19	0.80	0.69	0.55	0.51	0.43	0.41	0.36

(1) $\bar{X}-R$ 图。 $\bar{X}-R$ 图是应用最为广泛的一对图，其他控制图（如 $\bar{X}-S$ 图、 $\tilde{X}-R$ 图）可参照进行。

1) 以子组为单元收集数据，确定子组大小、子组个数与子组间隔。

子组大小 n ：一般以 4~5 个为宜。

子组个数 k ：一般以 20~25 组为宜。

子组间隔：没有统一规定，要视产量而定。譬如每小时生产 10 个以下产品，子组间隔可定为 8 小时；若每小时产量在 10~19 个之内，子组间隔可定为 4 小时；若每小时产量在 20~49 个之内，子组间隔可定为 2 小时；若每小时产量在 50 个以上，子组间隔可定为 1 小时。

按合理子组原则在生产线上抽样及测量，抽样要尽快完成，以减少对生产线的干扰。取控制图用纸（见图 8—9），填写表头，记录测量数据。

计算每个子组的均值 \bar{x} 和极差 R ：

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$R = x_{\max} - x_{\min}$$

在 \bar{X} 图和 R 图上选择合理的尺度，使点子尽量在图的中间部位。

在控制图上点出均值 \bar{x} 和极差 R ，然后连成折线。

2) 计算控制限。计算 k 个子组均值的平均值 $\bar{\bar{x}}$ ：

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_k}{k}$$

作为 \bar{X} 图的中心线，又称为过程平均。

计算 k 个子组极差的平均值 \bar{R} ：

计量值控制图

零件名										质量特征值										工序名称										零件号					图号				
操作者										设备号					测量仪器					测量单位					公差					零等于									
日期																																							
次数																																							
样本	1																																						
	2																																						
	3																																						
	4																																						
	5																																						
合计																																							
均值 \bar{X}																																							
极差 R																																							
操作者																																							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25													
\bar{X}																																							

图 8—9 控制图表的一般格式



$$\bar{R} = \frac{R_1 + R_2 + \cdots + R_k}{k}$$

作为 R 图的中心线, 又称平均极差。

计算 \bar{X} 图的上、下控制限:

$$UCL_{\bar{X}} = \bar{\bar{x}} + A_2 \bar{R}$$

$$LCL_{\bar{X}} = \bar{\bar{x}} - A_2 \bar{R}$$

计算 R 图的上、下控制限:

$$UCL_R = D_4 \bar{R}$$

$$LCL_R = D_3 \bar{R}$$

完成 \bar{X} 图和 R 图中的中心线 (实线) 与上下控制限的绘制。

这样绘制的 \bar{X} 图与 R 图供分析使用, 考察过程是否受控, 若过程失控需对控制图进行修改或补充。

控制图上点的散布状态是对生产过程运行状况的缩影, 各种波动 (正常波动或异常波动) 都可通过点的散布状态表现出来。应从图上判断过程是否存在异常波动, 并对每个异常波动逐个分析, 寻找原因, 及时纠正, 纳入标准, 以免再现。 \bar{X} 图显示子组间的波动, 并表明过程的稳定性。 R 图显示子组内的波动, 也反映了所考察过程的波动程度。 R 图的失控将会影响到 \bar{X} 图, 因为 \bar{X} 图的上下控制限依赖于平均极差 \bar{R} 。所以应先分析 R 图, 后分析 \bar{X} 图。

表 8-4

某团队项目进入控制阶段, 对生产过程的工艺参数进行控制, 收集数据如表 8-4 所示。

表 8-4

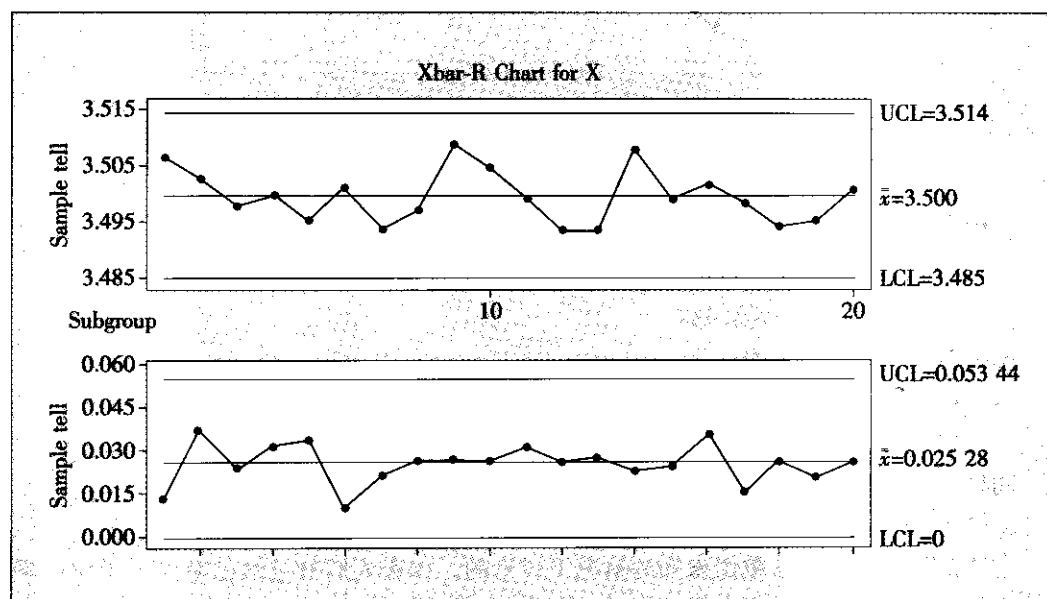
工艺参数数据

样本序号	观察值					样本均值 \bar{x}_i	样本极差 R_i
	x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}		
1	3.506 5	3.508 6	3.500 9	3.514 4	3.503 0	3.506 5	0.013 5
2	3.488 2	3.508 5	3.488 4	3.525 0	3.503 1	3.502 6	0.036 8
3	3.489 7	3.489 8	3.499 5	3.513 0	3.496 9	3.497 8	0.023 3
4	3.515 3	3.512 0	3.498 9	3.490 0	3.483 7	3.500 0	0.031 6
5	3.505 9	3.511 3	3.501 1	3.477 3	3.480 1	3.495 1	0.034 0
6	3.497 7	3.496 1	3.505 0	3.501 4	3.506 0	3.501 2	0.009 9
7	3.491 0	3.491 3	3.497 6	3.483 1	3.504 4	3.493 5	0.021 3
8	3.499 1	3.485 3	3.483 0	3.508 3	3.509 4	3.497 0	0.026 4
9	3.509 9	3.516 2	3.522 8	3.495 8	3.500 4	3.509 0	0.027 0
10	3.488 0	3.501 5	3.509 4	3.510 2	3.514 6	3.504 7	0.026 6
11	3.488 1	3.488 7	3.514 1	3.517 5	3.486 3	3.498 9	0.031 2
12	3.504 3	3.486 7	3.496 9	3.501 8	3.478 4	3.493 2	0.025 9
13	3.504 3	3.476 9	3.494 4	3.501 4	3.490 4	3.493 5	0.027 4
14	3.500 4	3.503 0	3.508 2	3.504 5	3.523 4	3.507 9	0.023 0
15	3.484 6	3.493 8	3.506 5	3.508 9	3.501 1	3.499 0	0.024 3

续前表

样本序号	观察值					样本均值 \bar{x}_i	样本极差 R_i
	x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}		
16	3.514 5	3.483 2	3.518 8	3.493 5	3.498 9	3.501 8	0.035 6
17	3.500 4	3.504 2	3.495 4	3.502 0	3.488 9	3.498 2	0.015 3
18	3.495 9	3.482 3	3.496 4	3.508 2	3.487 1	3.494 0	0.025 9
19	3.487 8	3.486 4	3.496 0	3.507 0	3.498 4	3.495 1	0.020 6
20	3.496 9	3.514 4	3.505 3	3.498 5	3.488 5	3.500 7	0.025 9

利用 MINITAB 软件来做控制图, 从 STAT-Control Chart-Xbar-R 入口, 填写过程变量名即可得到下列控制图 (见图 8—10):

图 8—10 \bar{X} -R 图

其中 \bar{X} 图的上、下控制限为:

$$UCL_{\bar{X}} = \bar{\bar{x}} + A_2 \bar{R} = 3.514$$

$$LCL_{\bar{X}} = \bar{\bar{x}} - A_2 \bar{R} = 3.485$$

R 图的上、下控制限为:

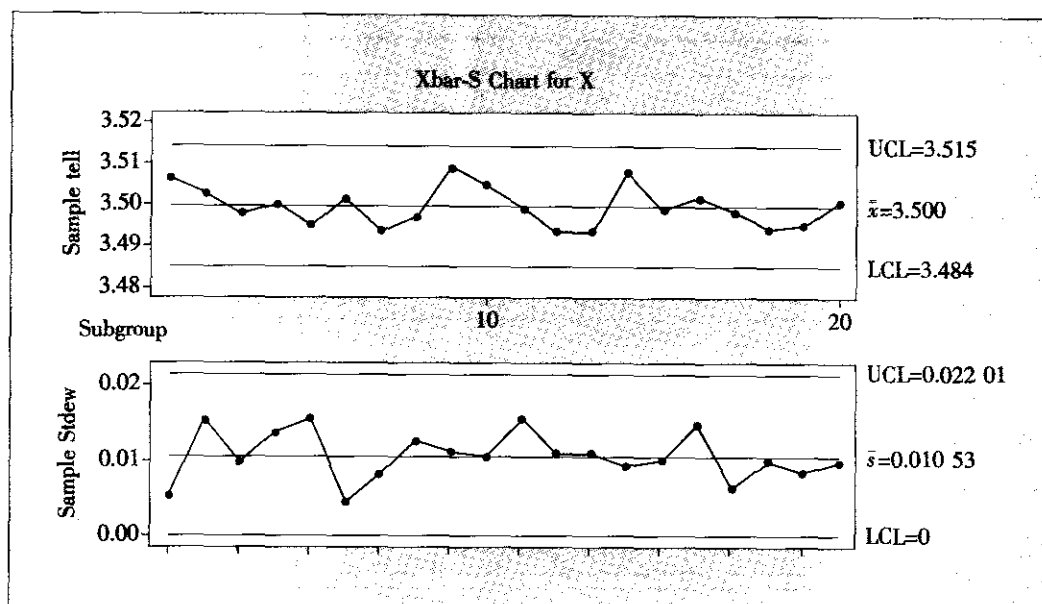
$$UCL_R = D_4 \bar{R} = 0.053\ 44$$

$$LCL_R = D_3 \bar{R} = 0$$

\bar{X} -R 图上显示: 过程受控。

(2) \bar{X} -S 图。同样程序, 可以制作 \bar{X} -S 图。

在例 8—1 中, 利用 MINITAB 软件来做均值标准差 (\bar{X} -S) 控制图, 从 STAT-Control Chart-Xbar-S 入口, 填写过程变量名即可得到下列控制图 (见图 8—11):

图 8—11 \bar{X} -S 图

其中 \bar{X} 图的上、下控制限为:

$$UCL_{\bar{X}} = \bar{\bar{x}} + A_3 \bar{s} = 3.515$$

$$LCL_{\bar{X}} = \bar{\bar{x}} - A_3 \bar{s} = 3.484$$

R 图的上、下控制限为:

$$UCL_S = B_4 \bar{s} = 0.02201$$

$$LCL_S = B_3 \bar{s} = 0$$

同样, \bar{X} -S 图上显示: 过程受控, 而且精度更高些。

3. 参数值给定时计量控制图的控制限

上述控制图都是在参数值未知的情况下讨论的。如果人们对过程已有深入了解并掌握了较多信息, 可以认为这些参数值给定。然而, 对于不了解的过程或新过程, 应从应用“参数值未知时的控制图”开始。参数值给定时, 控制限的计算见表 8—5。

表 8—5 参数值给定的计量控制图中心线和上下控制限

控制图的名称与符号		CL	UCL	LCL
均值—极差控制图 (\bar{X} -R 图)	\bar{X} 图	X_0 或 μ	$X_0 + A\sigma_0$	$X_0 - A\sigma_0$
	R 图	R_0 或 $d_2\sigma_0$	$D_2\sigma_0$	$D_1\sigma_0$
均值—标准差控制图 (\bar{X} -S 图)	\bar{X} 图	X_0 或 μ	$X_0 + A\sigma_0$	$X_0 - A\sigma_0$
	S 图	s_0 或 $c_4\sigma_0$	$B_6\sigma_0$	$B_5\sigma_0$
单值—移动极差控制图 (I-MR 图)	I 图	X_0 或 μ	$X_0 + 3\sigma_0$	$X_0 - 3\sigma_0$
	MR 图	R_0 或 $d_2\sigma_0$	$D_2\sigma_0$	$D_1\sigma_0$

其中: X_0 , R_0 , σ_0 , s_0 和 μ 为给定的参数值。系数 A , D_2 , D_1 , B_6 , B_5 见附表 6。

在例 8—1 中, 如果团队经过研究, 认定参数值为: $\mu=3.505$, $\sigma_0=0.020$, 则 $\bar{X}-S$ 控制图调整为图 8—12 所示。

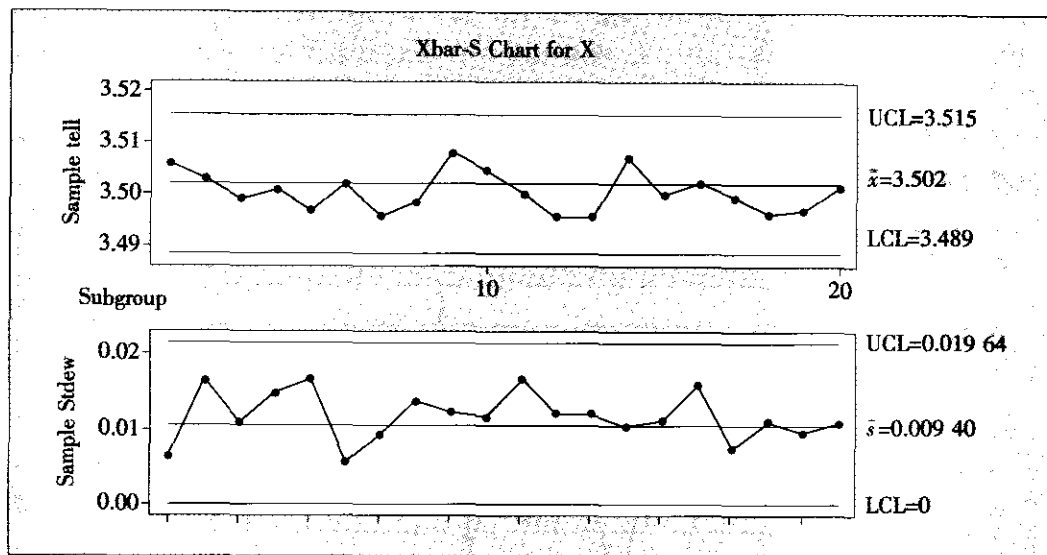


图 8—12 参数值给定的 $\bar{X}-S$ 控制图

8.2.3 计数控制图

当数据为离散数据时, 控制图可采用计数控制图。计数控制图又可分为计件控制图与计点控制图。

1. 计件控制图

计件控制图的背景是二项分布 $B(n, p)$ 。当 n 给定时, 只含一个参数 p 。 p 常用于表示不合格品率。控制不合格品率只需一张控制图。根据样本大小是否相同, 分别是:

- (1) 不合格品率控制图 (p 图): 样本大小一般较大, 可不相等。
- (2) 不合格品数控制图 (np 图): 样本大小相等。

计件控制图的中心线和上下控制限的计算如表 8—6 所示。其中: 参数值未给定时, $\bar{p}=\frac{X}{n}$; 参数值给定时, 即 p_0 已知。如果 LCL 计算结果为负, 则 $LCL=0$ 。

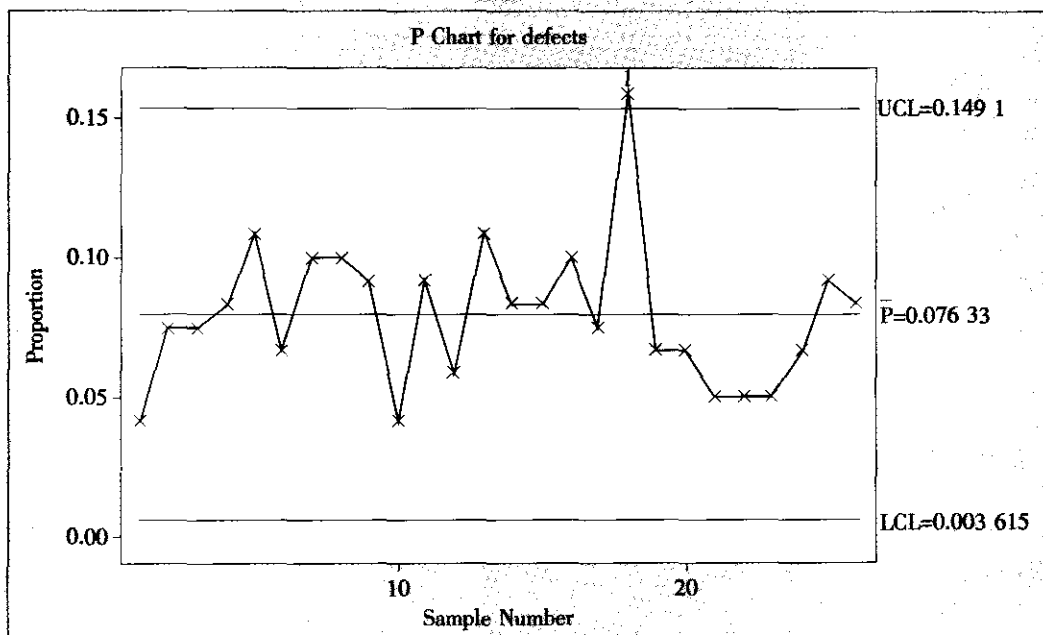
表 8—6 计件控制图的中心线和上下控制限

控制图的名称与符号		CL	UCL / LCL
不合格品率 控制图 (p 图)	参数值 未给定	\bar{p}	$\bar{p} \pm 3\sqrt{\bar{p}(1-\bar{p})/n}$
	参数值 给定	p_0	$p_0 \pm 3\sqrt{p_0(1-p_0)/n}$
不合格品数 控制图 (np 图)	参数值 未给定	$n\bar{p}$	$n\bar{p} \pm 3\sqrt{n\bar{p}(1-\bar{p})}$
	参数值 给定	np_0	$np_0 \pm 3\sqrt{np_0(1-p_0)}$

图 8-12

在硅片生产线上进行抽检, 每日临下班时抽取 120 片, 3 月份 26 天的记录中发现不合格硅片数量如下: 5, 9, 9, 10, 13, 8, 12, 12, 11, 5, 11, 7, 13, 10, 10, 12, 9, 19, 8, 8, 6, 6, 6, 8, 11, 10。准备设计一个控制图来监控硅片生产中的不合格品率。

利用 MINITAB 软件制作不合格品率控制图, 从 STAT-Control Chart-p 入口, 填写过程变量名即可得到下列控制图 (见图 8-13)。

图 8-13 不合格品率 p 控制图

用同组数据做 np 控制图, 从 STAT-Control Chart- np 入口, 填写过程变量名即可得到下列控制图 (见图 8-14)。

图 8-14 的图形与 p 图完全相同, 只是纵坐标的度量单位不同而已。

图中明显可以看出第 18 天不正常。经检查发现, 当天曾因两次停电而中断生产, 后来虽然恢复工作, 但破坏了正常生产过程, 处于不正常状态, 这天数据应删除。删除后只剩 25 个数据。得 p 控制图 (见图 8-15)。

此处只给出 p 图, np 图省略。注意, 这里删去了 18 号观测值, 因此以后的观测值的序号提前了一号。从图中显示结果看, 过程是受控的。

假如六西格玛团队希望硅片合格率控制在 94.5% ($p_0 = 1 - 94.5\% = 0.055$, 这是参数值给定情形), 再设计一个控制图来监控硅片生产。

利用 MINITAB 软件制作不合格品率控制图, 从 STAT-Control Chart-p 入口, 在 “historical p ” 处填写 p_0 值, 即可得到参数值给定的 p 控制图 (见图 8-16)。

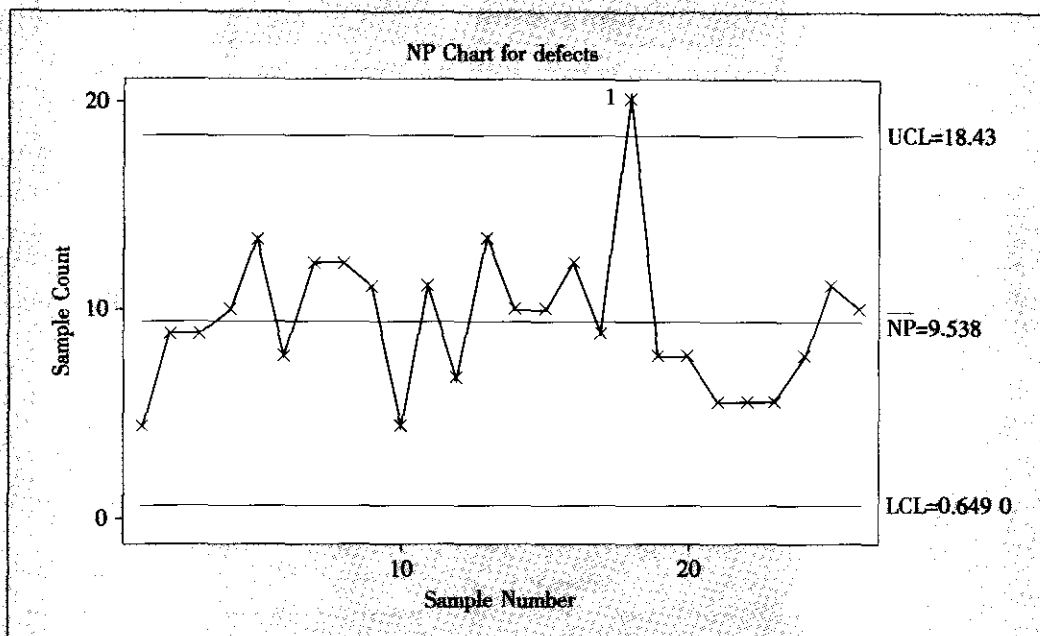


图 8—14 不合格品数 np 控制图

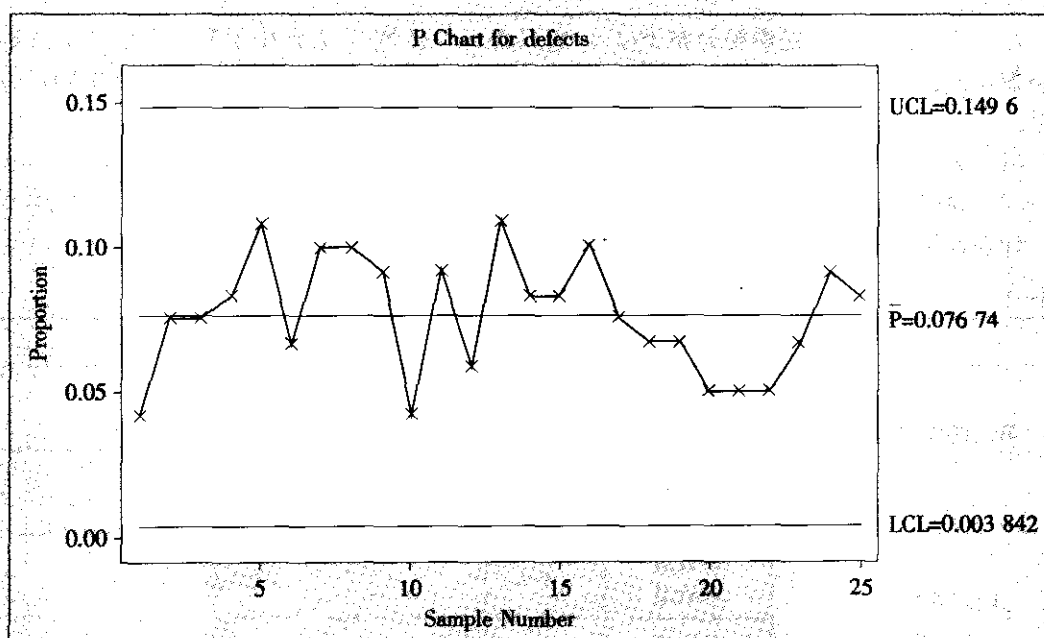
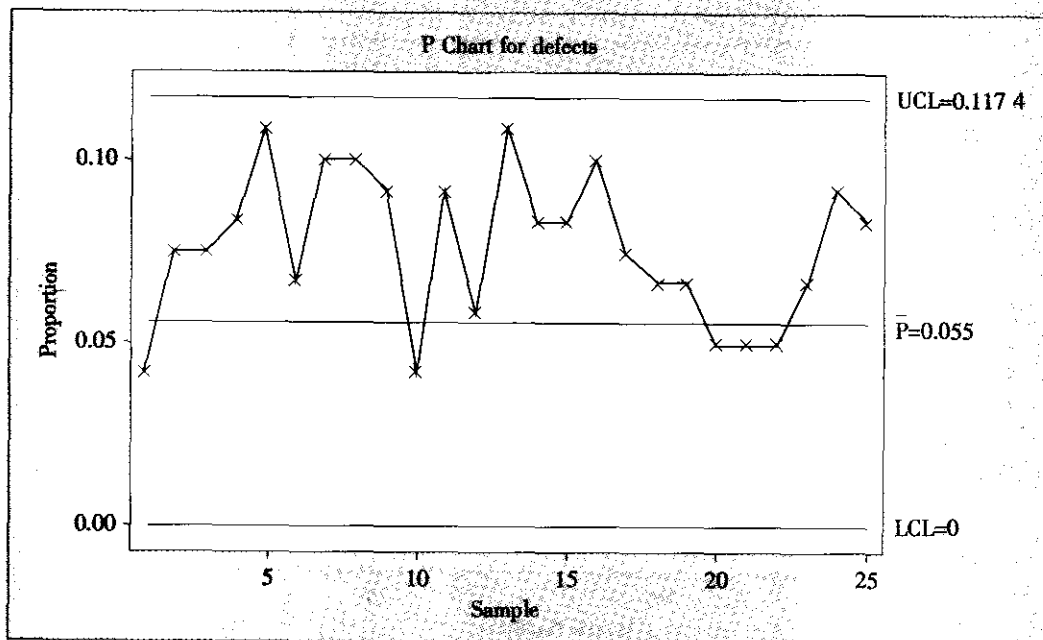


图 8—15 去除异常观测值后的 p 控制图

图 8—16 参数值给定的 p 控制图

2. 计点控制图

计点控制图的背景是泊松分布 $P(\lambda)$ ，它只含一个参数 λ ， λ 常用于表示单位产品上的缺陷数，控制缺陷或单位产品上的缺陷只需一张控制图。根据样本大小是否相同，分别是：

(1) 单位缺陷控制图 (u 图)：样本大小可不相等。

(2) 缺陷数控制图 (c 图)：样本大小要相等。

计点控制图的中心线和上下控制限的计算如表 8—7 所示。

表 8—7 计点控制图的中心线和上下控制限

控制图的名称与符号		CL	UCL / LCL
单位缺陷数控制图 (u 图)	参数值 未给定	\bar{u}	$\bar{u} \pm 3\sqrt{\bar{u}/n}$
	参数值 给定	u_0	$u_0 \pm 3\sqrt{u_0/n}$
缺陷数控制图 (c 图)	参数值 未给定	\bar{c}	$\bar{c} \pm 3\sqrt{\bar{c}}$
	参数值 给定	c_0	$c_0 \pm 3\sqrt{c_0}$

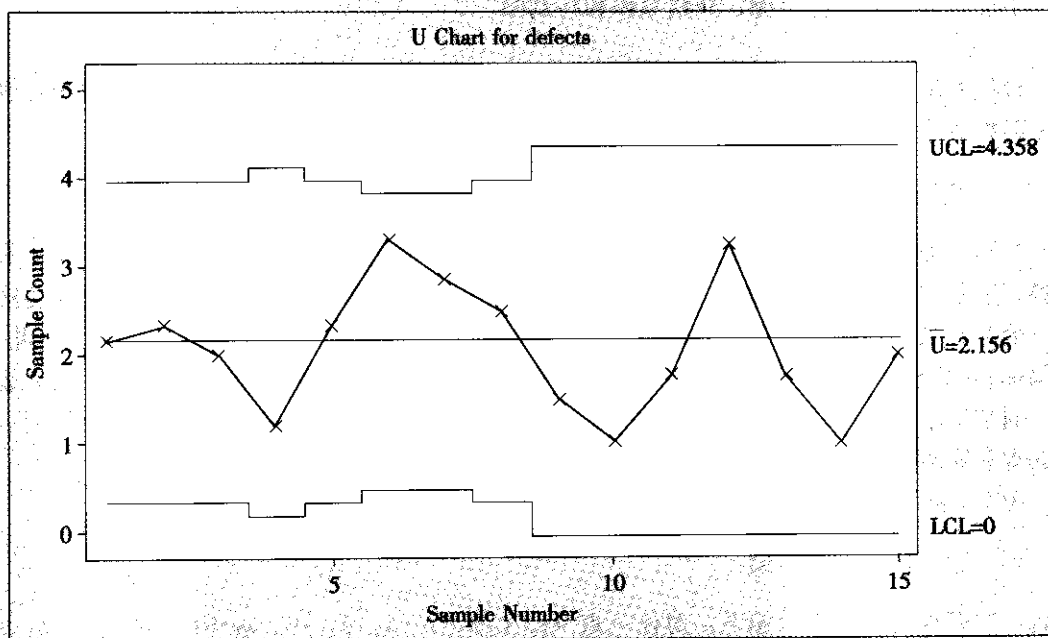
在给硅片镀膜的生产过程中,如果硅片上有了斑点,则每个斑点都被认为是缺陷。每天检测的硅片数不全相同。设计一个控制图来监控每片硅片的平均斑点数(见表8—8)。

表 8—8

数据表

日期	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
缺陷数	13	14	12	6	14	23	20	15	6	4	7	13	7	4	8
检查片数	6	6	6	5	6	7	7	6	4	4	4	4	4	4	4

根据问题的叙述,这里应该使用 u 图。用 MINITAB 软件,从 STAT-Control Chart- u 入口,输入“变量名”及“检查单元数”,直接得到下列 u 控制图(见图8—17)。由于每天检查硅片数不全相同,上下控制限变成了“城墙”型,受检查的片数越多,则上下控制限间的范围越窄。

图 8—17 硅片平均斑点数的 u 控制图

8.3 特殊的控制图

8.3.1 累积和控制图

利用常规控制图进行监控,不可避免地要面对两类错误:第 I 类错误,亦称为虚发警报的错误。过程正常而点子偶然超出控制限外,根据点出界就判异的原则,判断过程处于



异常,于是就犯了第Ⅰ类错误,即虚发警报的错误,通常记为 α 。第Ⅱ类错误,亦称为漏发警报的错误。过程已经出现了异常,但仍会有部分产品,其质量特性的数值会位于控制限内。如果抽取到这样的产品,点子就会落在控制限内,而不能判断过程出现了异常,从而犯了第Ⅱ类错误,即漏发警报的错误,通常记为 β 。常规控制图的设计思想是先定 α ,再看 β 。休哈特为增强使用者的信心,把常规控制图的虚发警报概率 α 取得特别小,仅取 $\alpha=0.27\%$ 。优点是:过程处于统计控制状态时,大约1 000个点子中仅有不到3个点子会超出控制界限。然而,缺点是:漏发警报的概率会增加。为了克服这个缺陷,1954年人们提出了累积和(cumulative sum)控制图,简称CUSUM控制图。

累积和控制图用“平均链长”(average run length, ARL)去控制两类错误,对过程的判断是以历次观测结果为依据,对观测数据加以累积。这样可以充分利用数据所提供的信息,对发现过程的小偏移特别有效。人们正是借鉴了累积和控制图的设计思想,才逐步完善形成了常规控制图的八种检验模式。

下面以计数型数据为例,介绍累积和控制图。

1. 累积和

设被检测的质量指标为样本中的不合格品数或缺陷数,累积和是指:逐次抽取的容量为 n 的样本中检出的不合格品数或缺陷数对其目标值 T (或参考值)的偏差之和。

$$C_{mT} = \sum_{i=1}^m (x_i - T)$$

当指标为不合格品数或缺陷数时,如果累积和图中的样本特性值 C_{mT} 的轨迹有向上的趋势,表示质量有变坏的倾向,若呈水平趋势或平缓向上趋势,则表示质量有变好的倾向。因此,当样本特性值有向上的趋势时,要及时报告质量是否已经发生了异常。为此,在国标GB4887—85(计数型累积和图)中规定了一个警戒值,称它为“参考值”,用 k 表示,通常要求 $k>T$,再定义一个参数 h ,称为“决策区间”。如果图中任意一点与最后一点的垂直距离大于等于 h ,则判断过程已经发生异常。

用 C_m 表示对参考值 k 的累积和,即

$$C_m = \sum_{j=1}^m (x_j - k)$$

那么,决策规则可以表示如下:对任一点 (i, C_i) ($i=1, 2, \dots, m-1$),有 $C_m - C_i \geq h$,即当

$$\sum_{j=i+1}^m (x_j - k) \geq h$$

则判断过程发生异常。因此,一组参数 (h, k) 决定了一个过程的检验方案。决定 (h, k) 与平均链长有关。平均链长是指对给定的质量水平(譬如,不合格品率 p ,平均缺陷数 λ ,过程的均值 μ 等)累积和控制图从开始到发出警报为止所抽取的平均样本数。

若一个过程的质量特性用过程的不合格品率来表示,这便是计件的特性。下面就用计件特性来说明对两种平均链长的要求。

(1) 若要求一个好的过程的不合格品率不大于 p_0 (即为合格质量水平AQL),那么当过程的不合格品率为 p_0 时表示过程处于受控状态。此时如果报警则属于误报。用 L_0 表示受控过程的平均链长,即在受控状态下,平均经过 L_0 次抽样才误发一次警报,在制作累

积和控制图时，总希望 L_0 大。

(2) 若规范要求过程的不合格品率不得大于 p_1 （即为拒收质量水平 RQL），那么当过程的不合格品率为 p_1 （ $p_1 > p_0$ ）时表示过程已到达失控状态，应该发出警报。用 L_1 表示失控过程的平均链长，即平均抽取 L_1 个样本就发一次警报，而不发警报相当于犯第Ⅱ类错误。为及时发现失控，在制作累积和控制图时总希望 L_1 小。

在国标 4887—85 中，给出了两种常用的 L_0 与 L_1 的值，一是 $L_0 \approx 1\,000$ ， $L_1 \approx 10$ ；二是 $L_0 \approx 200$ ， $L_1 \approx 5$ 。我们可由 L_0 与 L_1 制定 (h, k) ，在国标中分别称它们为 C_1 方案与 C_2 方案，如表 8—9 表示。

表 8—9 累积和控制图的参数

	C ₁ 方案				C ₂ 方案			
	参数值	特性值			参数值	特性值		
目标值 T_0	h	k	$\lambda_{10}/\lambda_{1\,000}$		h	k	λ_5/λ_{200}	
⋮								
0.5	3	1.5	3.0		2	1.5	3.3	
0.8	5	1.5	2.3		3.5	1.5	2.7	
1.0	5	2	2.2		5	1.5	2.5	
1.25	4	3	2.2		5	2	2.3	
1.6	5	3	2.0		4	3	2.15	
2.0	7 或 8	3	1.80 或 1.79		5	3	2.10	
2.5	7	4	1.73		5	4	1.92	

表中， $T_0 = np_0 = \lambda_0$ ； λ_{10} （或 λ_5 ）表示对应失控过程的平均缺陷数； $\lambda_{1\,000}$ （或 λ_{200} ）表示对应受控过程的平均缺陷数。

2. 累积和控制图的步骤

制定计数累积和控制图的步骤如下：

(1) 给出 L_0 或 L_1 ，以确定用 C_1 或 C_2 方案。通常用 $L_1 \approx 10$ 或 5，前者用 C_1 方案，后者用 C_2 方案。

(2) 根据产品的规格给出两个不合格品率。确认过程受控的不合格品率 p_0 与过程已经失控的不合格品率 p_1 。在给出这两个参数时要注意 p_1/p_0 的值不能太大。

(3) 确定 n, h, k 。当用 C_1 方案时，根据 p_1/p_0 的值，从 $\lambda_{10}/\lambda_{1\,000}$ 中找出最接近的比值，从该行对应的 T_0 值可确定样本容量 n ，即 $n = T_0/p_0$ ，而该行对应的 h 与 k 即为所求的参数。

当用 C_2 方案时，则从 λ_5/λ_{200} 中找出最接近的比值，然后确定 n, h, k 。

图 8—4

项目团队在夹具磨统工序的过程控制中，每半小时抽检一个样本，按控制测量的要求，产品的不合格品率不应超过 6%，并希望对异常的生产过程能在半天中检测出来。按此要求制定一个累积和控制方案。

按上面所述的步骤来制定累积和控制方案。

由于该工序半小时检测一次, 又根据要求希望在半天内能检测出异常, 则要求 $L_1 < 8$, 所以采用方案 C_2 ; 由于要求不合格品率不能超过 6%, 故令 $p_1 = 5\%$, 确定 p_0 时, 应使 p_1/p_0 的值不至于过小, 根据过去的生产所能达到的较好水平, 取 $p_0 = 2.1\%$; 从 C_2 方案中找出与 $p_1/p_0 = 2.33$ 最接近的值为 2.3 或 2.5, 取其中稍大的 2.5, 则 $T_0 = 1.0$, $h = 5$, $k = 1.5$ 。则 $n = T_0/p_0 = 40$ 。

即每半小时抽检 40 个产品, 记录其中的不合格品数 (见表 8—10)。求其对参考值 k 的偏差的累积和, 当累积和小于 0 时, 下一次重新计算累积和; 当累积和大于 h 时判断过程异常, 需要查找原因, 加以纠正。

表 8—10 检查结果及累积和方案的计算 ($k=1.5$, $h=5$)

序号 m	不合格品数 x_m	对参考值的偏差 $x_m - k$	偏差的累积和 C_m
1	1	-0.5	-0.5*
2	2	0.5	0.5
3	2	0.5	1.0
4	0	-1.5	-0.5*
5	0	-1.5	-1.5*
6	0	-1.5	-1.5*
7	0	-1.5	-1.5*
8	2	0.5	0.5
9	0	-1.5	-1.0*
10	1	-0.5	-0.5*
11	2	0.5	0.5
12	4	2.5	3.0
13	2	0.5	3.5
14	2	0.5	4.0
15	3	1.5	5.5**

当 $C_m < 0$ 时, 表示过程正常, 用 * 表示, 从下一个样本开始重新计算累积和; 当 $C_m > h$ 时, 判断过程异常, 发出报警信号, 用 ** 表示。

从第 11 号到第 15 号样本的局部累积和超过了判定距 h , 故判生产过程异常, 需要查找原因。从第 11 号到第 15 号样本为止, 这段时间生产过程的平均不合格品率为:

$$\frac{2+4+2+2+3}{40 \times 5} = 0.065 = 6.5\%$$

利用 MINITAB 软件制作不合格品率控制图, 从 STAT-Control Chart-CUSUM 入口, 输入变量名及子组个数 (本例为 1), 在 “Option” 窗口输入 k 和 h 数值, 则可得到累积和控制图 (见图 8—18)。

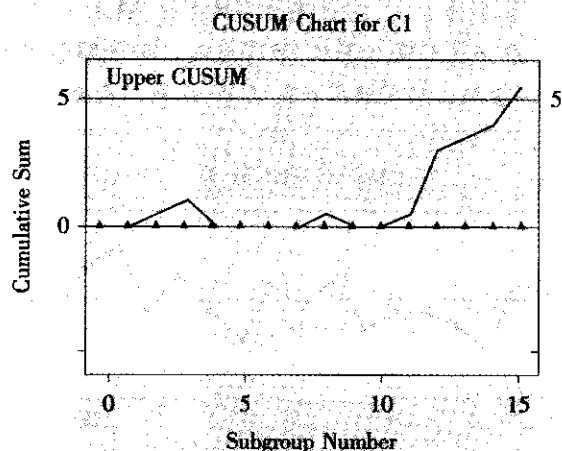


图 8—18 累积和控制图

8.3.2 指数加权滑动平均控制图

继 1954 年人们提出累积和控制图之后, 1959 年又提出了另一种能够有效控制过程小偏移的方法: 指数加权滑动平均 (exponentially weighted moving average, EWMA) 控制图。EWMA 控制图同样充分利用了所有的历史数据, 而且对数据的处理更有特色。下面以单个观测值为例, 介绍 EWMA 控制图。

EWMA 控制图关注的是当前值与历史数据的加权平均, 即

$$Z_t = \lambda x_t + (1 - \lambda) Z_{t-1}$$

式中, Z_t 为当前时刻的 EWMA 统计量; Z_{t-1} 为上一时刻的 EWMA 统计量; $Z_0 = \mu$, 为 EWMA 统计量的初值; x_t 为当前时刻的观测值; λ 为权重因子。

EWMA 控制图的控制限为:

$$UCL = \bar{\mu} + 3\sigma \sqrt{\lambda / (2 - \lambda)}$$

$$LCL = \bar{\mu} - 3\sigma \sqrt{\lambda / (2 - \lambda)}$$

式中, $\bar{\mu}$ 和 σ 分别为观测值序列 x_t 的均值和标准差的估值。

绘制 EWMA 控制图时, 一般要求给定权重 λ 。Montgomery (1991) 推荐使用 $0.05 \leq \lambda \leq 0.25$ 。一般 λ 常取 0.08, 0.10, 0.15, 0.2。

例 8—5

团队收集了自动车床生产的 25 根轴棍的直径数据 (单位: 毫米):

18.5, 20.9, 20.8, 19.5, 20.7, 21.1, 19.1, 19.8, 20.2, 20.2, 19.7, 21.2, 20.4, 21.0, 19.4, 20.2, 21.3, 20.3, 22.1, 21.5, 20.0, 20.7, 21.7, 22.2, 22.9
试分析生产状况是否异常。

解: 先作单值—移动极差控制图如图 8—19 所示。

最后几个点虽有上扬的趋势, 但用八条判异准则, 都没发现任何问题, 只能判生产正常。

用 EWMA 控制图, 取 λ 为 0.2, 对上述数据做 EWMA 控制图 (见图 8—20)。从该图看到在 22 号点以前, 生产一直正常, 但从 22 号点起, 均值明显上扬。从 EWMA 控制图上可看出, 样本点对目标值向上明显漂移, 到了 25 号点已跃出控制限, 应判生产异常。

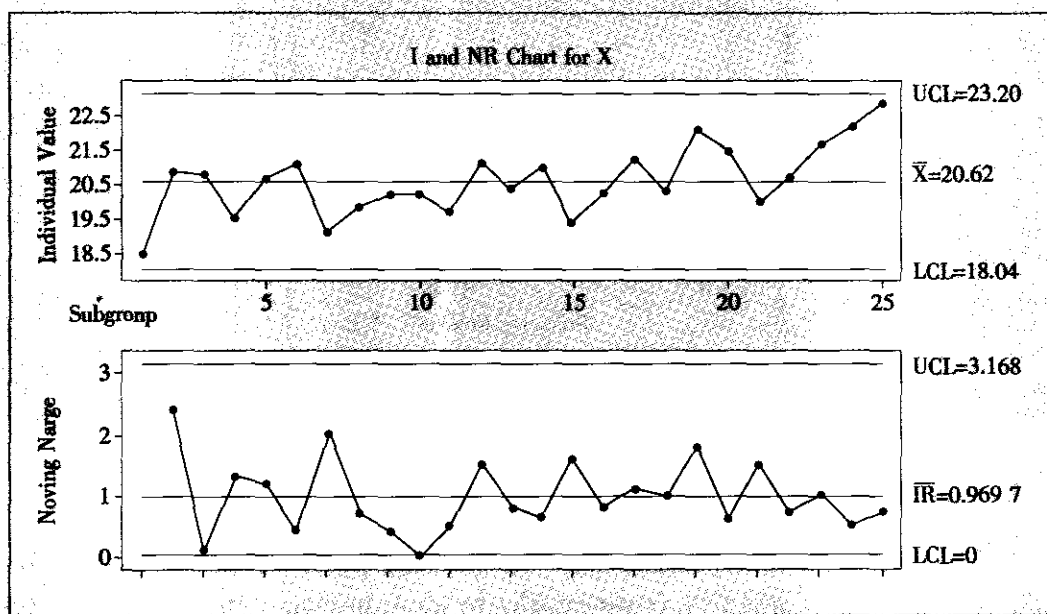


图 8—19 轴棍直径单值—移动极差控制图

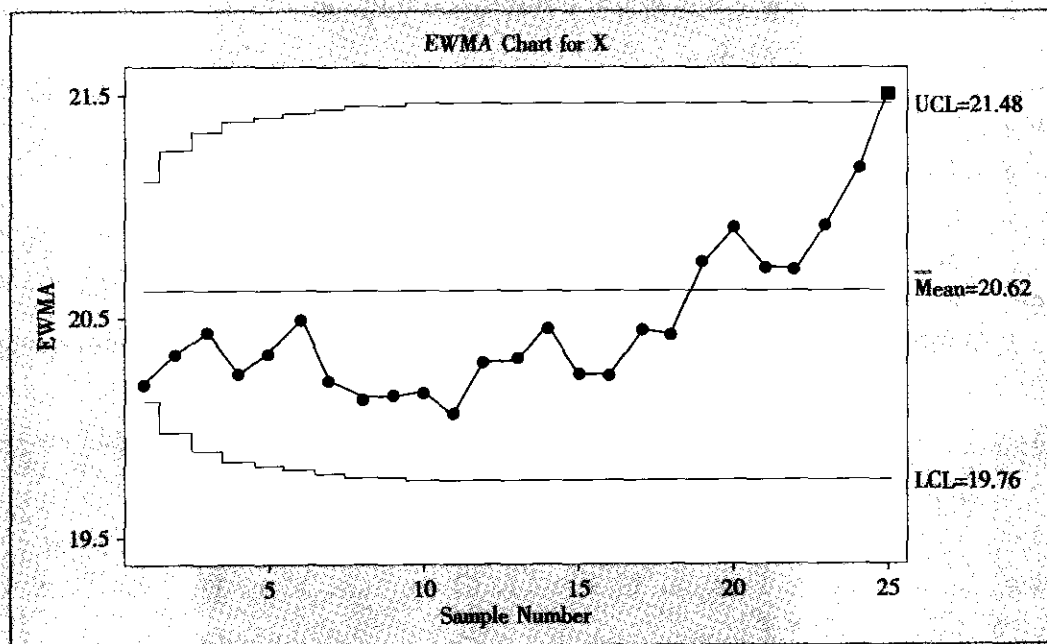


图 8—20 轴棍直径 EWMA 控制图

8.3.3 短期生产统计过程控制技术

短期生产是许多制造公司采用的生产方式。这种制造趋势已经发展为适应顾客个性化要求的小批量生产。将来，这种生产方式会更加普遍。亨利·福特的“顾客可以喜欢任何颜色，但我们只提供黑色汽车”的时代已经过去。

经典 SPC 方法，例如 \bar{X} 和 R 图，诞生于大批量生产的时代。产品生产经常要持续几周、几个月，甚至几年。很多当前正在使用的“SPC 规则”都是在大批量生产的情况下设计的。例如，当容量为 5 的子组至少得到了 25 组时，才能够计算控制界限。这项规则在 1930 年可能根本不成问题，然而，今天确实会出问题。事实上，存在许多的生产线，其整条生产线生产的产品比开始使用一张标准控制图所要求的还要少。

多数情况，可以对常规 SPC 方法进行简单调整以用于短期、小批量的生产。例如：可以使用移动均值和移动极差来构造 \bar{X} 和 R 图。然而，有些 SPC 方法非常适合在短期和小批量生产中应用，如精确方法、代码值图、变量的稳定控制图等。

8.3.4 预控制

预控制最初是由杜兰·沙宁 (Dorian Shainin) 在 20 世纪 50 年代提出的。按照沙宁的提法，预控制是基于容差来控制过程的简单运算法则。假定过程正在生产一种产品，该产品具有可度量的、可调整的、服从某种分布的质量特性。对于分布的峰度和稳定性没有做任何假设。在容差界限内设定警戒区域。采集单个观测值的连续样本，直到一行中在两个落在警戒区域之前，五个都落在中心区域，那么过程是合格的。为了简化预控制图的应用，通常将预控制图用不同的彩色作为编码。在这样一张图上，中心区域是绿色的，警戒区域是黄色的，容差范围以外的区域是红色的，如图 8—21 所示。预控制并不等于 SPC。

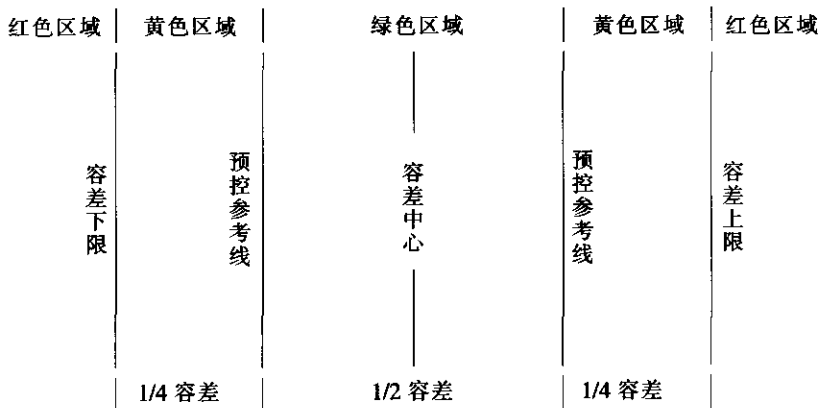


图 8—21 双侧容差的预控制区域

应用预控制的第一步是设定条件。测量每件产品，直到连续五件产品落在绿色区域。如果遇到一件产品落在黄色区域，那么重新开始计数。如果连续两件产品都落在黄色区域或是出现落在红色区域的情况，那么，应当对过程进行调整，然后重新开始计数。

设定条件之后，就进入了下一步：使用。周期性地测量两件连续的产品 (A, B)。如

果两件都落在同一侧的黄色区域,那么应该进行调整。如果落在两侧的黄色区域,需要采取措施减小过程的变异。如果出现了落在红色区域的情况,显然,应该进行过程调整。注意:当两件产品皆落在黄色区域时,必须立刻进行调整以避免不合格品的产生。当落在红色区域时,停止过程,因为一直在生产不合格品。依照设定的程序,找出所有不合格品。

沙宁(Shainin, 1988)建议对检验频率进行调整,使得在两次过程调整之间可以平均检验六对A与B,简单的计算公式如下:

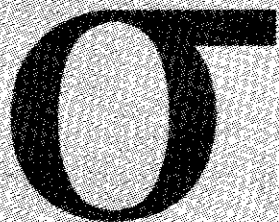
$$\text{两次测量之间的分钟数} = \text{两次过程调整之间的小时数} \times 10$$

本章小结

本章主要介绍了在项目控制阶段的主要工作和常用的工具,重点介绍了几种用于过程控制的控制图。这里要特别指出,控制是项目工作的一个重要阶段,一定要把改进阶段的成果通过建立过程控制体系保持下来,要注意那些修改了的文件、程序、作业指导书,及时纳入现有的质量管理体系。另外,项目控制阶段结束后,还有必要对整个项目工作进行总结和团队评估,为以后的项目开展积累经验。

第 3 篇

精益生产与 六西格玛设计



第 9 章 精益生产

第 10 章 六西格玛设计

9.1 精益生产概述

精益生产 (lean production) 源于日本丰田汽车公司所创造的丰田生产方式, 是继单件小批量和福特大批量流水线生产方式之后诞生的生产方式。它的基本理念是在产品设计、制造、销售以及零部件库存等各个环节消除一切不必要的浪费。它提出的准时化生产 (just in time) 和自働化 (jidoka) 等思想改变了日本企业的经营方式, 极大地促进了日本制造业的飞速发展, 被称为“改变世界的机器”。

9.1.1 精益生产的产生及发展

精益生产起源于日本丰田汽车公司的“丰田生产方式”, 第二次世界大战后日本重新开始发展汽车工业, 战后的日本经济萧条, 缺少资金和外汇。怎样建立日本的汽车工业? 是照搬美国的大量生产方式, 还是按照日本的国情, 另谋出路? 丰田选择了后者, 以大野耐一为代表的丰田人进行了一系列探索和实验, 根据日本的国情, 经过 30 多年的努力, 终于形成了完整的丰田生产方式, 并使日本的汽车工业赶上并超过了美国。

为揭开日本汽车工业成功之谜, 1985 年美国麻省理工学院制定了一个名为“国际汽车计划”的研究项目, 组织了 53 名专家、学者, 历时 5 年对 14 个国家的近 90 个汽车制造厂进行了实际考察, 并对西方的大量生产方式与日本的丰田生产方式进行分析对比, 最后作为研究结果, 于 1990 年出版了《改变世界的机器》一书, 第一次把丰田生产方式称为 lean production, 即精益生产方式。“lean”原意是“瘦”, 因为与大量生产方式相比, 该生产方式的一切投入都大为减少。“精益生产”概念的出现在世界范围内掀起了一股研究先进制造模式的浪潮。

20 世纪 90 年代美国进行了一系列针对精益生产的研究和实践。这其中包括美国军方 1993 年出台的美国“国防制造企业战略”、“精益航空计划 (lean aerospace initiative)”等政府指令性的活动。除了汽车行业又有更多的美国企业, 如波音、洛克希德马丁、惠普等投入到实施精益生产的实践中来。在这个过程中, 日本人提供了基本的思考方法, 用出色的实践证明了精益生产的强大生命力。实际应用和研究表明, 精益思想反映了在现代竞争条件下企业对多品种、小批量、高质量、快交货、低成本要求下制造管理模式的先进理念, 尽管它源于日本, 但它实际上反映的是几十年来人们对现代制造所面临的新环境所作

的管理变革,精益制造在世界范围的广泛应用,则证明了精益思想在世界上的普遍意义,并升华为新一代的生产哲理。1996年詹姆斯·沃迈克(James Womack)和丹尼尔·琼斯(Daniel Jones)所著的《精益思想》(Lean Thinking)一书问世,成为精益生产方式的里程碑。

在《精益思想》一书发表前后,精益思想跨出了它的诞生地——制造业,作为一种普遍的管理哲理在各个行业传播和应用,先后出现了精益建筑(lean construction)、精益服务(lean services)、精益物流(lean logistics)、精益医疗(lean healthcare)、精益软件开发(lean programming)和精益政府(lean government)的概念,精益思想的应用取得了飞跃性的发展。

近年来,我国也有越来越多的企业正在实施或准备实施精益生产,尤其是我国企业在面临消耗严重、资源短缺的情形下,精益思想日益得到社会的认可和接受。严格来讲,精益与六西格玛管理都是旨在实现持续改进的活动,都强调顾客满意和系统集成。六西格玛的出发点是减少变异,精益生产的核心是消除一切不必要的浪费,严格来讲,消除变异的过程本身也消除浪费。两者目标的一致性、工具的互补性、管理模式上的相互借鉴和补充使得精益和六西格玛的结合成为必然。本章将主要介绍精益的思想和方法体系,以及如何实现精益和六西格玛的结合。

9.1.2 精益生产的核心思想

1. 以消除浪费为核心思想

精益生产和浪费(muda)直接对立。浪费包括很多类型,如:闲置的库存、不必要的工序、不必要的运输、超过需求的生产、人员的不必要动作、各种等待等,所有这些日常生产中很少为人们所注意但却大量存在的活动(在很多制造流程中甚至超过了90%),不仅不能为企业创造价值,反而增加了成本,都是精益生产反对和致力消除的。从企业获取订单到生产过程,再到最后的销售,消除浪费的思想贯穿始终。

沃迈克和琼斯把企业内部的浪费分为两种:(1)生产中不创造价值,但在现有技术与生产条件下不可避免的作业内容,称为Ⅰ型muda,如必要的在线检验、物料运输等。(2)不创造价值且可以立即去掉的作业内容,称为Ⅱ型muda,如不必要的等待、多余的动作、不必要的审批等。精益生产审查生产特定产品的所有活动,首先努力消除Ⅱ型muda,降低成本,然后使产品在整个流程中流动起来。通过连续流动,进一步消灭Ⅰ型muda浪费,优化流程,并且流动越快,所暴露出的浪费会越多,越有利于精简流程和降低成本。

2. 暴露问题,解决问题

精益生产也是一个企业实现业务流程持续改进的过程,与六西格玛实现持续改进类似,这个过程也需要发现企业在流程中存在的问题。在发现问题方面,精益思想特别强调暴露问题,找到引起该问题的根本原因,然后才能彻底解决问题。如在很多企业里,过多的库存量非常普遍,究其原因,可能是由于供应商供货不及时或时间不确定、机器需要较长时间的换模或换线、某台设备机器经常出故障等。为了保证整个生产表面上看起来顺畅,企业不得不增加大量的库存以备不测,即使采用统计分析或经济批量模型求得最佳的批量从而控制库存,也是治标而非治本。从精益思想的角度看,库存是万恶之源,库存不仅占用了资金,消耗了库存管理成本,更重要的是,库存掩盖了这些问题,使这些问题成

为企业的顽疾。因此,精益生产要求尽量减少库存,才能暴露出这些问题,然后调动团队,确定改进项目,找到根本原因,解决问题,保持持续改进。

9.1.3 精益生产与传统生产模式的比较

精益生产与传统的生产模式相比,从考虑问题的层次、范围、角度和具体的操作方法等方面都有较大的区别(见表9—1)。

表 9—1 精益生产与传统生产模式的比较

比较内容	传统生产模式	精益生产模式
企业目标	足够好的原则:容忍一定的废品率,最低限度的库存,产品单一	尽善尽美的原则:不断降低成本,追求零缺陷、实现零库存、产品多样化
工作强度	体力和精神的紧张	有节奏感、创造性的紧张
人事管理 组织结构	以个人为单位进行工作实绩考核	以集体为单位进行工作实绩考核
	组织机构以集中式或多级递阶控制为主	组织结构以分布式或适度递阶控制为主
	面向职能,按职能划分部门	面向过程,按过程划分项目组
	被动管理	主动管理,以人为本
产品开发	顺序开发	并行开发
销售	以市场预测和订货为基础的大规模生产	把销售纳入生产体系,客户和供应商也参加到综合小组中,按订单组织生产
生产计划 管理	面向设备负荷编制生产计划,使设备尽可能满负荷运转	面向实际需要编制生产计划
	大批量优先	最小批量
	按事先制定的生产计划控制生产过程——“推”的管理方式	按合同组织生产过程——“拉”的管理方式
	自顶向下,分层生产控制	分布式自适应控制
采购管理	零件成本最低的采购原则	产品总成本最低的采购原则
	按零件或材料类型采购	以子系统为单位进行采购
财务管理	被动管理,注重事后成本核算	主动管理,注重面向事前整个生产过程的成本控制
	以降低成本为目标	以提高生产率、改善产品增值过程为目标
质量管理	专职的质量管理人员	质量自检
	提高产品质量必然会增加产品成本	提高产品质量并非一定会增加产品成本
信息处理	集中处理	分布处理,将信息作为生产要素得到更好的应用

由表9—1可见,精益生产是对原有的生产模式、质量控制、产品开发、管理技术、组织结构等进行不断改进,逐步达到尽善尽美的理想境界。

9.1.4 精益思想的五项基本原则

沃迈克和琼斯在《精益思想》中,将由丰田开创的精益生产方式总结出5个基本原则,成为所有踏上精益道路的组织不厌其烦地理解和实行的基本原则。这些原则是:

1. 正确地确定价值

●精益生产中经常使用“价值”的概念，这里的价值是从顾客的角度定义的，所谓有价值的活动是能为顾客增加效用、顾客承认它的必要性并愿意为其付款的活动，正确地确定价值就是以顾客的观点来确定企业从设计到生产到交付的全部过程，实现顾客需求的最大满足。

●以顾客的观点确定价值还必须将生产全过程的多余消耗减至最少，不将额外的花销转嫁给顾客。

●精益价值观将生产方、顾客和供应商的利益统一起来，而不是过去那种对立的观点。

●以客户为中心的价值观审视企业的产品设计、制造过程、服务项目就会发现太多的浪费，从不满足客户需求到过分的功能和多余的非增值消耗。当然，消灭这些浪费的直接受益者既是顾客也是生产方。

2. 识别价值流

价值流是指从原材料转变为成品并给它赋予价值的全部活动。这些活动包括：从概念设计到产品设计，工艺设计，再到投产的技术过程，从订单处理、计划到送货的信息过程和从原材料到产品的物质转换过程，以及产品全生命周期的支持和服务过程。精益思想识别价值流的含义是在价值流中找到哪些是真正增值的活动，哪些是可以立即去掉的不增值活动。精益思想将所有业务过程中消耗了资源而不增值的活动叫做浪费。识别价值流就是发现浪费和消灭浪费。识别价值流的方法是“价值流分析”（value stream map analysis）——首先以产品族为单位画出当前的价值流图，再从顾客的观点分析每一个活动的必要性。价值流分析成为实施精益思想最重要的工具。

价值流并不是从自己企业的内部开始的，多数价值流都向前延伸到供应商，向后延长到向客户交付的活动。按照最终用户的观点全面地考察价值流，寻求全过程的整体最佳。

3. 流动

如果正确地确定价值是精益思想的基本观点，识别价值流是精益思想的准备和入门的话，“流动”（flow）和“拉动”（pull）则是精益思想实现价值的中坚。精益思想要求创造价值的各个活动（步骤）流动起来，强调的是不间断地“流动”。“价值流”本身的含义就是“动”的，但是由于根深蒂固的传统观念和做法，如部门的分工（部门间交接和转移时的等待）、大批量生产（机床旁边等待的在制品）等等阻断了本应动起来的价值流。精益将所有的停滞作为企业的浪费，号召“所有的人都必须和部门化的、批量生产的思想作斗争”，用持续改进、JIT、单件流（one-piece flow）等方法在任何批量生产条件下创造价值的连续流动。

当然，使价值流流动起来，必须具备必要的环境条件。这些条件是：

●过失、废品和返工都造成过程的中断、回流。实现连续的流动要求每个过程和每个产品都是正确的。

●环境、设备的完好性是流动的保证。3P（production, preparation, process），5S，全面生产性维护（total productive maintenance, TPM）都是价值流动的前提条件。

●系统地进行流程设计和能力规划，避免瓶颈造成的阻塞。

4. 拉动

“拉动”就是按客户的需求投入和产出，使用户精确地在他们需要的时间得到需要的东西。实行拉动以后用户或制造的下游就像在超市的货架上取到他们所需要的东西一样，而不是把用户不太想要的产品强行推给用户。拉动原则由于将生产和需求直接对应，消除了过早、过量的投入，而减少了大量的库存和现场在制品，大量压缩了提前期。拉动原则更深远的意义在于企业具备了当用户一旦需要，就能立即进行设计、计划和制造出用户真正需要的产品的能力，最后实现抛开预测，直接按用户的实际需要进行生产。

实现拉动的方法是实行 JIT 生产和顺畅生产。当然，JIT 和顺畅生产的实现最好采用单元布置，对原有的制造流程进行改造。应用实践表明，流动和拉动一般可以使产品开发时间减少 50%、订货周期减少 75%、生产周期降低 90%。

5. 尽善尽美

奇迹的出现是由于上述 4 个原则相互作用的结果。改进的结果必然是价值流动速度显著地加快。这样就必须不断地用价值流分析方法找出更隐藏的浪费，做进一步的改进。这样的良性循环成为趋于尽善尽美（perfection）的过程。近来沃迈克又反复阐述了精益制造的目标是：“通过尽善尽美的价值创造过程（包括设计、制造和对产品或服务整个生命周期的支持）为用户提供尽善尽美的价值。”

尽善尽美是难以达到的，但对尽善尽美的持续追求，将造就一个永远充满活力、不断进步的企业。

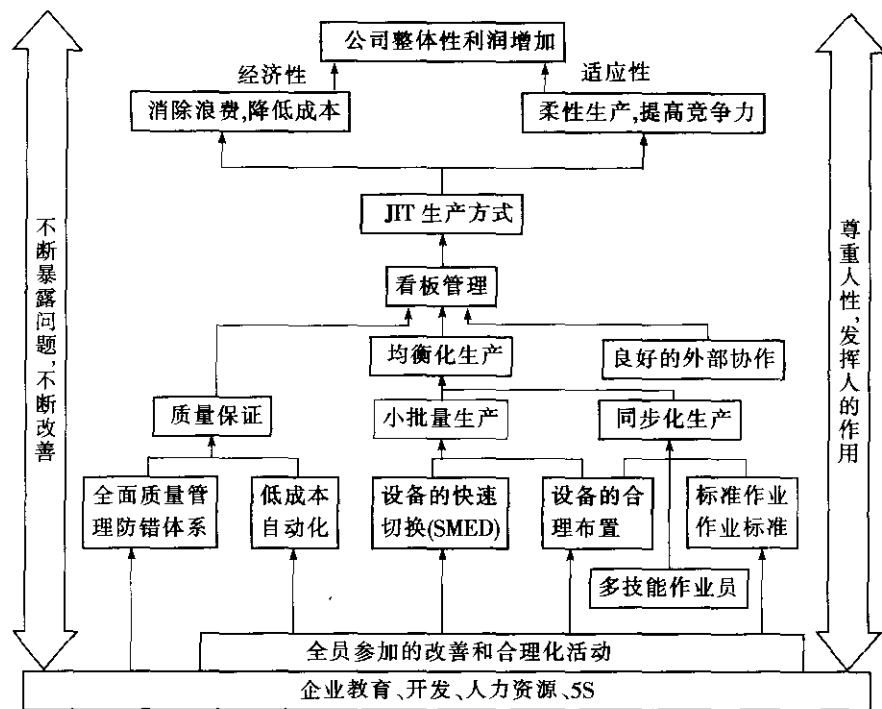
9.2 精益生产的技术体系及主要支撑技术

9.2.1 精益生产的技术体系

企业追求精益的过程实质上就是一个持续改进的过程，而且是一个长期的过程，一场没有终点的比赛。改善过程需要综合应用现场管理、设施布置与优化、人因工程、作业研究、质量控制与保证、流程管理与控制、现代制造技术等一系列方法与工具，因此精益生产是一个包含了多种制造技术和管理技术的综合技术体系。实际上，精益生产就是工业工程技术在企业中的具体应用。图 9—1 表示了精益生产的体系构造，同时表明了该体系的目标以及实现目标的各种技术、手段和方法及其相互间的关系。

要实现“消除浪费，增加利润”这一基本目标，就必须彻底杜绝过量生产以及由此产生的在制品过量和人员过剩等各种直接浪费和间接浪费。如果生产系统能够具有足够的柔性，能够适应市场需求的不断变化，即实行准时化生产，在必要的时刻生产必要数量的必要产品或零部件，创造出能够灵活地适应市场需求变化的生产系统，这种生产系统就能够从经济性和适应性两个方面来保证公司整体性利润的不断提高。此外，这种生产系统具有一种内在的动态自我完善机制，即在准时化生产的激发下，通过不断缩小加工批量和减少在制品储备，使生产系统中的问题不断暴露出来，使生产系统本身得到不断完善，从而保证精益生产的顺利进行。

下面对精益生产技术体系的主要支撑技术进行简要说明。



9.2.2 看板拉动式生产

1. 看板拉动生产的基本概念

看板是对生产过程中各工序生产活动进行控制的信息系统。通常，看板是一张在透明塑料袋内的卡片。经常被使用的看板主要有两种：取料看板和生产看板。取料看板标明了后道工序应领取的物料的数量等信息，生产看板则显示前道工序应生产的物品的数量等信息。精益生产以逆向“拉动”方式控制着整个生产过程即从生产终点的总装配线开始，依次由后道工序“在必要的时刻领取必要数量的必要零部件”，而前道工序则“在必要的时刻生产必要数量的必要零部件”，以补充被后道工序领取走的零部件。这样，看板就在生产过程中的各工序之间周转，从而将与取料和生产相关的时间、数量、品种等信息从生产过程的下游传递到了上游，并将相对独立的工序个体联结为一个有机的整体，如图 9-2 所示。

使用看板所遵循的基本规则是：

- (1) 不向后一道工序传送次品；
- (2) 在必要的时间由后一道工序来领取必要数量的必要零部件；
- (3) 前道工序只按后一道工序的取货量进行生产；
- (4) 进行均衡生产，各工序均匀地领取零部件；
- (5) 根据看板进行微小的调整；
- (6) 要使生产工序稳定化、合理化，不断减少看板数量。

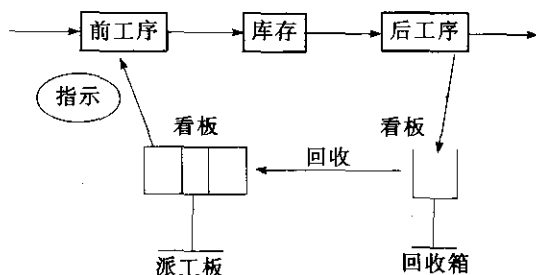
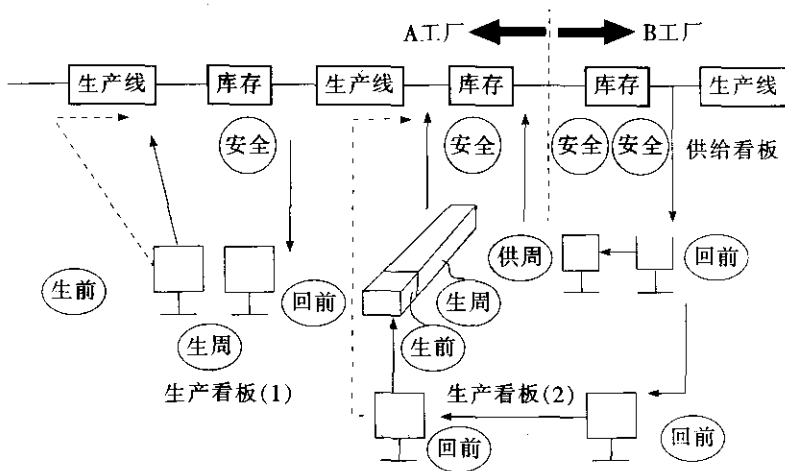


图 9—2 看板拉动生产模型

2. 看板的设置及计算

看板的设置以生产过程中的产品和零部件的搬运单位作为看板设置的基准，搬运的工具通常使用托板、拖车、零件箱以及专用物流器具等。



注：安全——安全库存；生前——生产提前期；生周——生产周期；回前——回收提前期；供周——供应周期。

图 9—3 看板张数计算方法

拉动生产过程中每种产品所需的看板张数通常使用式（9—1）进行计算，式中各参数的含义和关系如图 9—3 所示。

$$\text{看板张数} = \frac{\text{每班最大产量} \times (\text{生产间隔期} + \text{生产提前期} + \text{回收提前期} + \text{安全库存})}{\text{单位容器容量}}$$

（9—1）

式中：

生产间隔期：从对生产线做生产指示到下次生产指示的间隔。

生产提前期：从对生产线做生产指示到生产结束的间隔。

回收提前期：回收放在回收箱内的看板，用来对前工序做生产指示的间隔。

安全库存：为应付突发事件所储备的库存，这里用安全库存量能够维持的生产时间表示。

单位容器容量：单位物流容器盛放的零件数。

图 9—4 所示为看板张数计算的例子, 生产间隔期=0.5 日, 生产提前期=0.2 日, 回收提前期=0.6 日, 安全库存=0.2 日产量, 单位容器容量=5 个/托板。1 天两班次 (早班+晚班) 生产, 每班作业时间=8 小时, 产品 A, B, C 每班最大的产量分别为: A 产品=100 个/班, B 产品=150 个/班, C 产品=150 个/班。则每班产品 A, B, C 所需的看板张数分别为:

$$A=100 \times (0.5+0.2+0.6+0.2) / 5=30 \text{ 张}$$

$$B=150 \times (0.5+0.2+0.6+0.2) / 5=45 \text{ 张}$$

$$C=150 \times (0.5+0.2+0.6+0.2) / 5=45 \text{ 张}$$

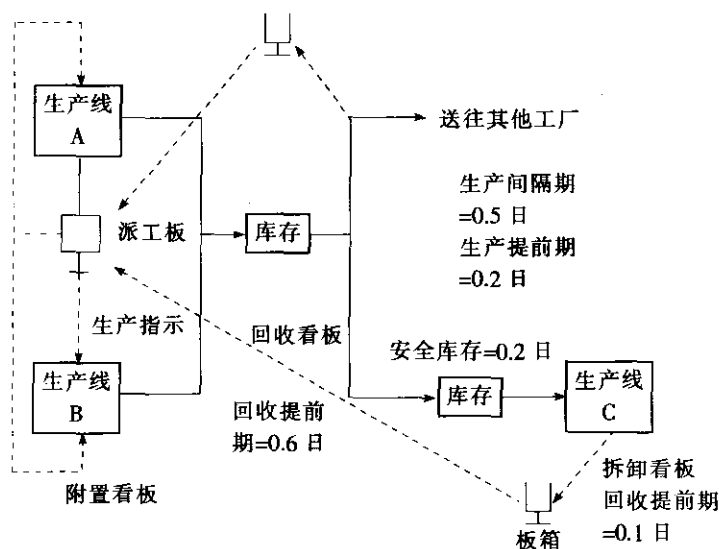


图 9—4 看板张数计算算例

实施看板拉动是有条件的, 如生产的均衡化、作业的标准化、设备布置合理化等。如果这些先决条件不具备, 看板拉动就不能发挥应有的作用, 从而难以实现精益生产。

9.2.3 均衡化生产与设备的快速换型调整

1. 均衡化生产

用看板拉动控制生产过程, 生产的均衡化是最重要的前提条件。均衡化生产要求的是生产数量的均衡和产品种类的均衡, 即总装配线向各前工序领取零部件时, 要均匀地领取各种零部件, 实行混流生产。要防止在某一段时间内集中领取同一种零部件, 以免造成前方工序的闲忙不均, 以及由此引发的生产混乱。生产的均衡化使得零部件被领取时的数量变化达到最小程度, 即各工序每天如一地以相近似的时间间隔领取数量相近的零部件, 如图 9—5 所示。这样, 各工序得以一定速度和一定数量进行生产, 这是实施看板拉动的首要条件。事实上, 在最终装配线没有实现均衡化生产的情况下, 看板拉动也很难实施。

除此之外, 精益生产把均衡化生产作为使生产适应市场需求变化的重要手段。通过均衡化生产, 任何生产线都不大批量地制造单一种类的产品。相反, 各生产线必须每天同时生产多种类型的产品, 以满足市场的需要。这种多品种、小批量的混流生产方式具有很强

的柔性，能迅速适应市场需求的变化。

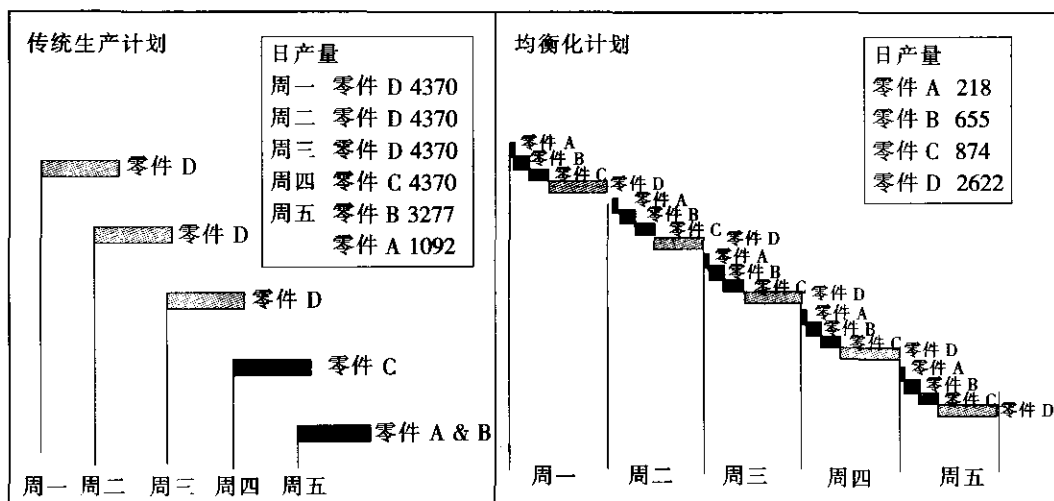


图 9—5 传统生产与均衡化生产的比较

为了实现以“多品种、小批量”为特征的均衡化生产，就必须缩短生产提前期，以利于迅速且适时地生产各类产品。于是，为了缩短生产提前期，则必须缩短设备的换型调整时间，以便将生产批量降到最小。

2. 设备的快速换型调整

实现以“多品种、小批量”为特征的均衡化生产最关键和最困难的一点就是设备的快速换型调整问题。

丰田公司发明并采用的设备快速换型调整的方法是 SMED 法 (single minute exchange of die)，即“单分钟内换模法”或快速换模。这种方法的要领就是把设备换型调整的所有作业划分为两大部分，即外部换型调整作业和内部换型调整作业，如图 9—6 所示。所谓“外部换型调整作业”是指那些能够在设备运转之中进行的换型调整作业，而“内部换型调整作业”是指那些必须或只能够在设备停止运转时才能进行的换型调整作业。为了缩短换型调整时间，操作人员必须在设备运行中完成所有的外部换型调整作业，一旦设备停下来则应集中全力于内部换型调整作业。最重要的一点就是要尽可能地把内部换型调整作业转变为外部换型调整作业，并尽量缩短这两种作业的时间，以保证迅速完成换型调整作业。

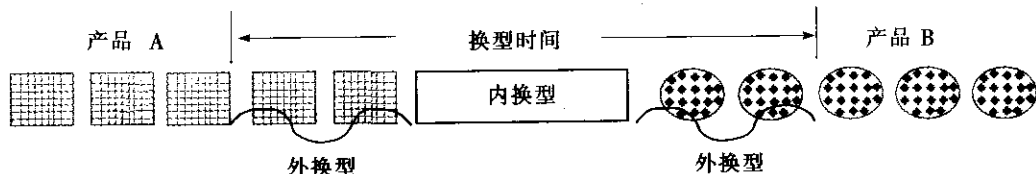


图 9—6 内部、外部换型调整作业

为了缩短作业转换时间，重要的是事先在设备运行当中把必要的夹具、工具、下一次使用的模具和材料准备好，然后在设备停下来后安装上这副模具，设备一开始运转，就把



拆下的模具和夹具移走。作业人员在设备停止运转时，必须集中全力进行换型作业。最重要的一点是，把内部换型和外部换型明确地区别开来，尽量把内部换型变成外部换型。

为了减少切换时间，通常依据图 9—7、表 9—2 所示的主要步骤方法，循序渐进地进行。

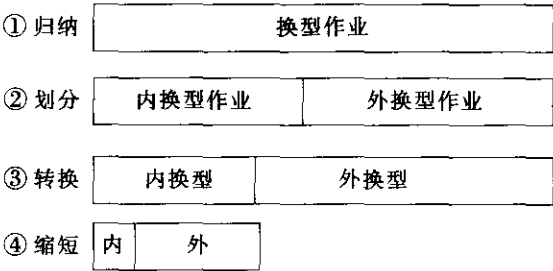


图 9—7 快速换型的主要步骤

表 9—2 快速换型方法和手段

方法与手段	目 的
模具、刀具、工具及材料的规格标准化和摆放	减少辅助时间
模具高度标准化	减少调整冲床时间
使用快速螺栓或快速夹具	减少装夹时间
使用辅助工具	减少内变换时间
旋转式台车	减少模具的搬运时间
使用特制的导轨和导销	减少吊装和定位时间
加强模具的通用性，免除整换调整时间	节省整换时间

设备的快速换型调整所带来的效果如下：

- (1) 减少每次生产的批量，降低库存；
- (2) 缩短生产周期；
- (3) 减少次品；
- (4) 增强作业人员责任心；
- (5) 增强生产系统柔性。

丰田公司把“设备的快速换型调整”视为提高企业竞争力的关键因素之一。

9.2.4 设备的合理布置与多技能作业员

1. 设备的合理布置

设备的快速换型调整为满足后工序频繁领取零部件制品的生产要求和“多品种、小批量”的均衡化生产提供了重要的基础。但是，这种频繁领取制品的方式必然增加运输作业量和运输成本，特别是如果运输不便，将会影响精益生产的顺利进行。可见，生产工序的合理设计和生产设备的合理布置是实现小批量频繁运输和单件生产单件传送的另一个重要基础。

丰田公司改变了传统的设备布置方式，采用了图 9—8 所示的 U 型生产单元，即按零

件的加工工艺要求,把功能不同的机器设备集中布置在一起组成一个个小的加工单元。这种设备布置方式可以简化物流路线,加快物流速度,减少工序之间不必要的在制品储量,减少运输成本。

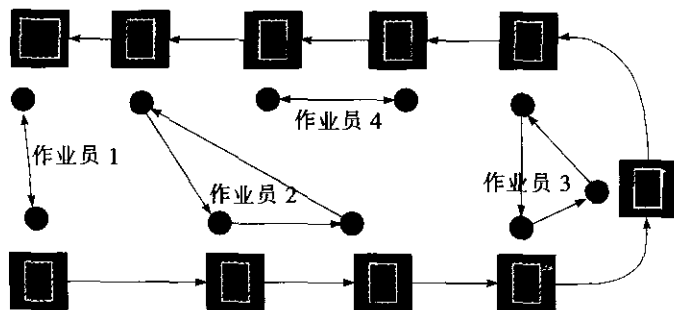


图 9—8 U型生产单元

U型单元中,由于作业员是属于可以操作很多不同工作站的多技能作业员,所以不但能缩减人员成本也可以使生产线变得更有弹性;同时也因为作业员之间的距离较近,若系统中有问题发生,作业员可以很快地反应并相互配合解决问题,进而减少损失。图 9—8 所示的 U 型生产单元物料的投入和成品的产出由同一个作业员 1 控制,可以有效控制 U 型单元的生产节拍和中间在制品数量,实现标准化作业。

2. 多技能作业员

多技能作业员(或称“多面手”)是指那些能够操作多种机床的生产作业工人。多技能作业员是与设备的单元式布置紧密联系的。在 U 型生产单元内,由于多种机床紧凑地组合在一起,这就要求并且便于生产作业工人能够进行多种机床的操作,同时负责多道工序的作业,排除了工序间不必要的在制品,加快了物流速度,有利于生产单元内作业人员之间的相互协作等。特别是,多技能作业员和组合 U 型生产线可以将各工序节省的零星工时集中起来,以便整数削减多余的生产人员,有利于提高劳动生产率。

9.2.5 标准化作业

所谓标准作业是指把生产过程中的各种要素(人、机、料、法、环、测)进行最佳组合,按“准时化”生产的要求制定出相适应的作业标准。各类作业人员按照作业标准进行操作。作业标准一方面是操作者的行为规范,另一方面是管理者检查指导工作的依据。

工人操作时所遵循的作业标准顺序,称作“标准操作规程”(standard operating procedure, SOP),这是标准化作业的第一个目标。标准化作业的第二个目标是在生产时间上,各制造过程达到平衡化。第三个目标,是将半成品的标准库存量,控制在最小数量范围内。

作业标准一旦决定下来,并不是一成不变的,随着技术的发展、工艺的改进以及持续不断的改善,需要反复修订作业标准。

标准作业的目的,是通过必要的、最小数量的作业人员进行生产。它由下面的程序确定:

- (1) 确定生产节拍(循环时间);

上加工的半成品。如果没有这些数量的在制品,那么生产单元内的一连串机器将无法同步作业。但是,应设法尽量减少在制品存量,使之维持在最低水平。

根据标准化作业的要求(通常用标准作业组合表表示),所有作业人员都必须在标准周期时间内完成单位制品所需要的全部加工作业,并以此为基础,对作业人员进行训练和对工序进行改善。

9.2.6 5S 与目视管理

5S 和目视管理是精益生产的基础。5S 是整理(seiri)、整顿(seiton)、清扫(seiso)、清洁(seiketsu)、素养(shitsuke)五项活动的统称。由于这五项活动每一个词的第一个字母都是“S”,所以简称“5S”。一个没有 5S 造就的良好现场管理基础的企业实施精益生产是不可想象的。

5S 是现场管理的基础。在一个没有实施 5S 管理的企业,浪费现象随处可见,无论是精益生产还是 MRP,ERP 等信息管理系统的实施,都要求全体人员具有按规定做事的习惯,什么物品应该放在哪里、数量应该是多少、何时出入库等等,否则,再好的生产系统也无法取得好的效果,甚至实施后其结果更糟,企业内部陷入混乱之中。国内近几年推行 MRP,ERP 失败的企业数远远超过成功的企业数,失败的原因有很多,其中员工没有养成“有规定按规定去做”的职业素养是失败的重要因素之一。

5S 管理是通过现场现物的规范,明确“场所、方法、标识”,确定“定点、定容、定量”及大量使用目视管理等方法、手段,构筑一个整洁、明朗、一目了然的工作现场,其最终目的是提升人的素养,让全体员工养成消除马虎之心、认真对待每一件小事、有规定按规定去做事的良好习惯。如此,追求完美的精益生产方式才可能成功导入。

近来,有些企业将 5S 进一步推广,倡导 7S,增加了安全(safety)和节约(save)两项内容,将安全生产和消除浪费、节水节电、保护环境等内容同时纳入进来。

目视管理是利用形象直观、色彩适宜的各种视觉感知信息来组织工作场所的活动,是一种以公开化和视觉显示为特征的管理方式。图 9—10 是生产现场和办公室实施了 5S 和目视管理后的示意图。

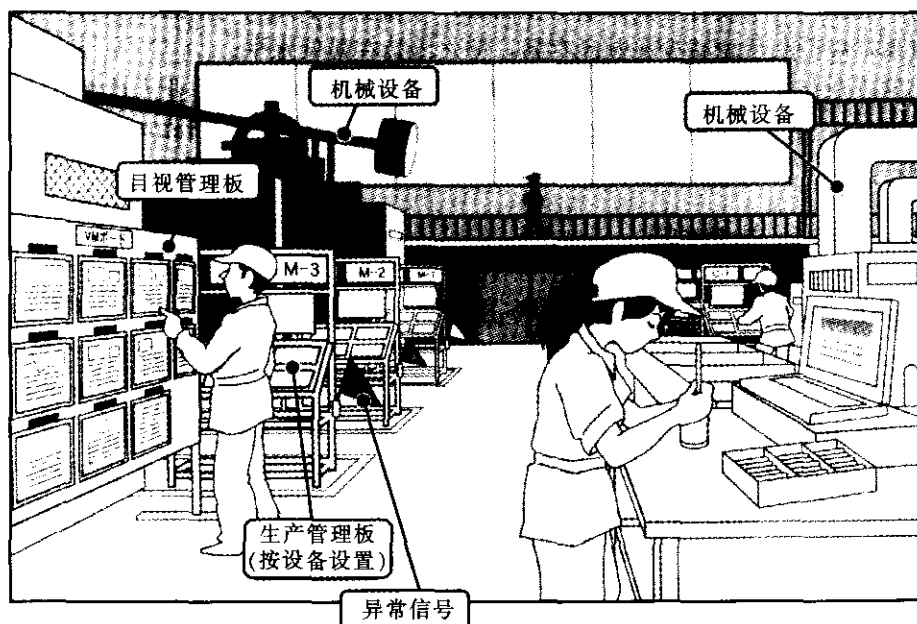
目视管理的内容与形式是:

(1) 生产/工作任务和完成情况要公开化和图表化。公开化就是要将生产/工作任务、计划指标、完成情况均应按车间、班组和个人落实,并列表公布张贴在墙上,并用作图法,使大家都能形象地看出各项指标完成中的问题和发展趋势,以便都能按期、按质、按量地完成各自的任务。

(2) 与现场/办公室密切相关的规章制度和工作标准、时间标准要公之于众。与岗位工人直接有关的部分,应分别展示在岗位上,如岗位责任制指导书等,并要始终保持完整、齐全、正确。

(3) 以清晰、标准化的视觉显示信息落实定置管理。采用清晰、标准化的信息显示符号,将各种区域、通道、物品的摆放位置鲜明地標示出;机器设备和各种辅助器具均应运用标准颜色,不得任意涂抹。

(4) 要采用与现场工作状况相适应的信息传导信号。为能及时地控制生产作业,生产环节和工种之间也要设立方便实用的信息传导信号,以尽量减少工时损失,提高生产的连



(a) 生产现场目视管理示意图



(b) 办公室目视管理示意图

图 9—10 生产现场与办公室目视管理示意图

续性。在各质量管理点，要有质量控制图，清楚地显示质量波动情况；车间要公布不良品的统计情况，出现的废品要陈列在展示台上，以利分析、诊断和改良。

(5) 现场/办公室各种物品的码放和运送要标准化。在现场中，各种物品要按标准化运送及码放，以达到过目知数。

(6) 统一规定现场人员的着装,实行挂牌制度。统一着装可以进一步体现正规化、标准化。单位挂牌和个人佩戴标志可以起到激励和推动的作用。

(7) 现场/办公室的各种色彩运用要实行标准化。为利于生产和工人的身心健康,现场中的色彩也要标准化。有关色彩的内容应符合工效学中的原则。

目视管理的基本要求:

- (1) 统一:即目视管理要实行标准化。
- (2) 简明:即各种视觉显示信号应易看懂,一目了然,简单明确。
- (3) 醒目:即各种视觉显示信号要清晰、位置适宜。
- (4) 实用:少花钱、多办事、讲求实效。
- (5) 严格:严格遵守和执行有关规定。

9.2.7 自动化与差错预防

自动化(jidoka)是丰田生产方式的两大支柱(JIT和自动化)之一,也被称为“智能自动化”或者“具有人的判断能力的自动化”,其产生可以追溯至丰田佐吉和他制作自动织布机的长期发明工作。他发明的自动织布机有一个装置可以发现断掉的纱线,一旦发现纱线断掉,织布机就会立即停止运转,等接好纱线后再重新启动织布机,这样便可以避免重复产生瑕疵,造成大量质量浪费。由这样的发明而产生的概念和方法被丰田在汽车生产过程中大量使用,并成为丰田生产方式的两大支柱之一。

自动化的主要理念是:应该在作业流程中及时进行源头质量管理,必须用有效的方法在发生异常或质量缺陷时立即探测到,并立即使生产暂停,使员工可以立即解决问题,避免异常扩大或质量缺陷进入到下游流程。这种品质管理比事后检验出问题再设法纠正问题更具成效且成本更低。

自动化在很多时候有“自动化”的含义,但绝不仅仅限于用机械化设备代替手工作业,二者之间的比较如表9—3所示。

表9—3 自动化与自动化的比较

自动化	自动化
机器自身可以发现异常并及时停机	出现异常需有人停机,否则将连续运行
不产生不良品,可以防止设备、模具、夹具出现故障损坏	不能及时发现不良品产生,易造成设备、模具、夹具出现故障损坏
容易发现异常原因,防止再次发生	难以及时发现异常原因,容易再次发生
省人	省力

自动化的理念主要通过各种差错预防装置(英文 error proofing 或者日文 poka-yoke)实现,Poka-Yoke由日本人新乡重夫(Shigeo Shingo)最早提出。在日语中 Yoke 是防止的意思,而 Poka 是疏忽错误之意。其主要目的在于预防生产过程中出现的可能失误,并避免由于失误而产生的质量问题。

Poka-Yoke 方法的内容包括两个基本方面:一是检测出现的差错;二是 Poka-Yoke 防错装置。检测差错包括三种形式,一是后继检测,即下一道工序的生产者检查上一道工序的质量问题并立即反馈给上一道工序的生产者,让其停止生产纠正错误;二是自我检测,



即生产者检查自身的工作；三是预防性检测，即在错误未发生之前发现它并采取相应措施。所有的检查都必须是100%的逐个检测，而不是抽查或其他非完全检查。

Poka-Yoke 防错装置主要包括两种类型：信号型 Poka-Yoke 和控制型 Poka-Yoke。前者是对生产失误所发出的警示，后者是对出现的错误所采取的保险性措施。如生产作业中，一个待加工零件被放在机器上一个不正确的位置，那么信号型 Poka-Yoke 可能是一个指示灯，提醒工人要纠正位置。而控制型 Poka-Yoke 则可能是连接机器电源的一个装置，它会自动关闭切断电源，使生产停止，直至零件被拿走或放于正确位置。

Poka-Yoke 的设置可在发生生产失误时对生产者发出及时警告，提醒其改正错误。同时它也可以通过一定的装置或手段对失误自动采取应急措施，防止生产失误转化成产品质量问题，因而成为预防质量问题的“保险”。

Poka-Yoke 防错装置实现的主要功能如下：

1. 出现操作错误时工件无法安装；
2. 工件出现问题时，机器无法开始加工；
3. 出现操作错误时，机器无法开始加工；
4. 自动修正错误操作使加工继续进行；
5. 后道工序检查前道工序出现的问题，防止出现不良品；
6. 丢序漏序时下一工序无法开始。

图 9—11 所示切削加工时经常出现零件错装夹的情况，通过在机床上加装防错挡板（图 9—11 下图所示）防止零件错装夹；图 9—12 所示是焊接加工时如果漏装了工件，随焊机主轴上下移动的防错装置会及时按下停止按钮，防止错误发生。

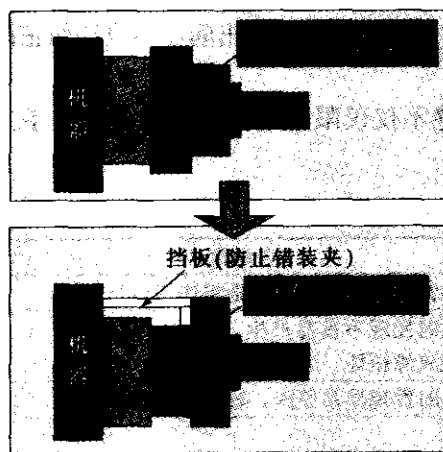


图 9—11 切削加工防错装置

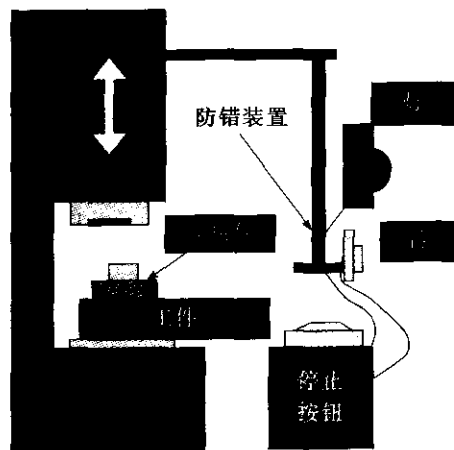


图 9—12 焊接加工防错装置

9.2.8 全面生产维护

全面生产维护（TPM）最重要的特性在于，由团队活动引导的公司上至最高管理层下至现场工人在内的所有员工的参与，操作工的自动维护，推行机器寿命期内的生产设备维护的系统，努力使设备效率最大化。

1. 全面生产维护 (TPM)

全面生产维护中的“全面”包括以下含义:

- 全面致力于追求经济效益及利润;
- 全面维护系统包括维护预防、可维护性及预防性维护;
- 员工全面地参与, 包括操作工自动维护及团队活动。

设备效率最大化的目标要求完全消除因机器操作产生的故障、缺陷、浪费及损失, TPM 的目标是零停机及零缺陷。通过 TPM 的实施, 一些公司报告停机只有原来的 2%, 过程缺陷减少了 90%。

有六大损失会对设备效率带来负面影响:

- 设备故障。停机造成时间损失使得生产力降低, 要加强预防维修工作, 改变事后修理的思维方式, 要求做到零故障停机。
- 设备调整。频繁的设备调整会带来损失, 应尽量减少设备调整时间甚至无须调整。
- 闲置与小停顿。如传感器缺陷导致流水线产品卡住, 由于生产线不平衡、换班、材料(零件)供应不及时等原因造成短时间的生产中断或速度放慢, 从而造成产能损失, 对于这些问题, 要通过加强现场管理水平和基础管理工作来保证。
- 速度减慢。是因设计运行速度与实际运行速度之间差异造成的损失, 造成速度减慢的原因一般是由于机器的老化或维修、维护不到位造成的。
- 过程缺陷。设备出错造成废品及质量缺陷, 过程缺陷造成设备工时损失, 同时造成材料和人工浪费。
- 产出减少。因机器停机及启动造成的产品损失。

表 9—4 显示相关六大损失的目标。

表 9—4 六大损失	
六大损失	目 标
停机损失	0
设置及调试损失	小于 10 分钟
速度损失	0
闲置与小停顿损失	0
质量缺陷损失	0
产出损失	最小化

如果能做到完全消除六大损失, 则会给工厂带来显著的改善。如果没有 TPM, 精益生产或精益生产系统根本不可能实现。

2. TPM 指标

设备综合效率。设备综合效率 (overall equipment efficiency, OEE) 是用于衡量 TPM 的主要指标, 它考虑了三个方面: 设备的工作时间、性能和质量水平。用公式表示为:

$$\text{设备综合效率} = \text{时间开动率} \times \text{性能开动率} \times \text{合格品率}$$

式中, 时间开动率反映了设备的时间利用情况; 性能开动率反映了设备性能的发挥情况;

合格品率反映了设备的有效工作情况。其中：

时间开动率 = 实际工作时间 / 计划工作时间

计划工作时间 = 每班可用时间 - 计划停机时间

式中，计划停机时间主要包括计划性维护、晨会等所需时间。

实际工作时间 = 计划工作时间 - 非计划停机时间。

式中，非计划停机时间包括设备故障、产品换型、设备调整等消耗的时间。

例如：若每天单班生产，每班工作 8 小时，每班可用时间为 480 分钟，其中有 15 分钟的换型时间，10 分钟的班前会时间，30 分钟的非计划设备故障，计划工作时间及时间开动率为多少？

计划工作时间 = 每班可用时间 - 计划停机时间 = $480 - 10 = 470$ （分钟）

实际工作时间 = 计划工作时间 - 非计划停机时间 = $470 - 30 - 15 = 425$ （分钟）

时间开动率 = $425 / 470 = 90.4\%$

性能开动率是指理论产出与实际产出的比值，它考虑两个方面：净开动率和速度运转率，用公式表示为：

性能开动率 = 净开动率 × 速度运转率

净开动率 = (产量 × 实际节拍) / 实际工作时间

净开动率用以测量设备稳定性、小停顿造成的损失、小问题及调试造成的损失。

速度运转率 = 理论节拍 / 实际节拍

速度运转率反映了设备由于老化或维护不良，达不到设计的或理论的节拍所造成的损失。

仍采用前面的例子，若加工量是 300 件，实际节拍时间是每件 1 分钟，理论节拍时间是每件 40 秒，实际工作时间是 425 分钟，净开动率、速度开动率和性能开动率是多少？

净开动率 = $(300 \times 1) / 425 = 70.59\%$

速度开动率 = $40 / 60 = 66.7\%$

性能开动率 = $0.667 \times 0.7059 = 47.08\%$

若质量合格率为 95%，设备综合效率是：

总体设备效率 = $0.904 \times 0.4708 \times 0.95 = 40.4\%$

在这个例子中，设备综合效率（OEE）很差，为 40.4%。TPM 奖获者的 OEE 一般大于 85%。理想的情况是：

- 工时利用率大于 90%；
- 性能开动率大于 95%；
- 质量合格率大于 99%。

3. 实施 TPM 的步骤

日本工厂维护研究所（JIPM）每年颁发预防性维护（PM）奖，评奖标准是基于 TPM 的改进效果，包括：

- 生产效率提升；
- 质量提高；
- 成本降低；
- 库存降低；

- 事故降低或消除；
- 污染控制；
- 工作环境。

将 TPM 原理与公司实际有效结合将带来工人与设备利用的提高，改进员工态度，提高员工技能，改变工作环境和提高员工士气。

(1) 实施步骤。要在几年时间内达到获奖 TPM 的效果，建议实施 12 个步骤。

步骤 1：最高管理层公布进行 TPM 的承诺（报纸、公告）；

步骤 2：通过公司交流程序介绍 TPM（集中的教育课程）；

步骤 3：组织每个职能层面推行 TPM（以团队形式）；

步骤 4：制定 TPM 方针及目标；

步骤 5：制定详细的 TPM 主计划；

步骤 6：启动 TPM（开始实际工作）；

步骤 7：建立提高设备效率的项目团队（优化努力以达到最大成果）；

步骤 8：培养技能以制定自动维护程序（培训操作工日常维护要求）；

步骤 9：制定计划性维护程序（维护部门开始计划性维护活动）；

步骤 10：操作工及维护技能培训；

步骤 11：制定早期设备管理程序（确保通过对新设备的评估、验证及改进使其可运用于生产）；

步骤 12：实施 TPM，旨在完美。

(2) 自动 TPM 团队活动。团队活动，包括与职能部门的跨职能活动，是 TPM 的核心，部门主管通常是团队组长，设备清洁、润滑、紧固、检查等活动就是团队活动。进行 TPM 的活动与时间都有监控。团队的发展过程分四个阶段：

- 自我发展：成员学习；
- 改进活动：完成团队改进活动；
- 问题解决：团队开始积极地选择问题进行解决；
- 自动管理：团队设定高层目标并独立管理工作。

(3) 面向可设计可维护性与可用性的设备设计。在大部分设备故障维修或预防性维护时，设备的可维修性都会影响到时间、材料与费用。易于维护性就是影响这些因素的设计特征。当然对过去初始设计状态的修正也需要计划与成本。

设计可维护性的指导原则，将提高可维护性和可用性：

- 标准化。寻找配件的兼容性，减少系统中不同零件类型的数量，这会减少备件库存。

- 模块化。建立模块尺寸、形状的标准，使得装配及拆卸程序标准化。

- 功能包装。将同一内容所需的零件集中在一个包装中。

- 互换性。控制尺寸及功能性公差，这是指当插入式装置有故障零件时，备件可以随时进行互换，一个产品同时可以用在其他装置上。

- 可达性。有一定空间使操作者能够进行适当工作，所有零件都伸手可得。零件能易于够到且易于更换，不需要先将好的零件拆下后才能更换故障零件。

- 故障报警。当发生故障时用一定方法通知操作工。这包括量规、仪表、灯光或



声音。

- 问题隔离。错误能被追踪，这是所有维护工作中最耗时的工作，通过预防性维护程序、自带检测设备（BITE）、设计简化及人员培训，这个问题可被降至最低。
- 标识。所有零件都有特定标识，并有记录修理及预防性维护的方法。

9.2.9 价值流图分析

在生产过程中，从原材料开始，经过多个工序的作业，加工成为零件、装配成组件和完整的产品。在整个工艺流程中，安排是否最合理，能否把在制品库存降至最低，非增值活动是否降到最低？这些需要进行科学的分析，价值流图就是解决上述问题的有效工具。

1. 价值流

价值流是指产品通过其生产过程的全部活动，包括从概念设计、产品设计、过程设计直到投产的设计流和从原材料、制造过程直到产品交到顾客手中的生产流。在企业实践中，一个典型的现象是，谁也说不准增值活动占企业生产活动的比例。当一位老总或厂长被告知，他的企业的增值活动还不到所有活动的5%时（实际上大部分工厂离5%还差得远！）他会大吃一惊，甚至怀疑你可能搞错了。导致这种现象的基本原因有二：其一，在大量生产模式下，部门按职能设置，彼此画地为牢，管理人员缺乏对企业运作的宏观了解。其二，管理人员缺乏对增值和非增值活动的认识和判别，难以运用系统的观点对企业的浪费现象进行分析和改善。因此，在精益生产中提出了“价值流管理”的概念。这种管理模式彻底打破了大量生产模式下的职能割据，由“价值流团队”对产品从供应商到顾客整个过程进行管理，实现对运作过程的系统而有效的控制，从而提高满足顾客需求的能力及反应速度。价值流经理由直接向最高管理者汇报的管理人员担任，享有改革的充分授权和自主权。价值流经理组建由跨部门成员组成的价值流团队，以顾客为导向选定产品族，并跟随其生产路径，从头到尾分析和描绘每一个工序的状态、工序间的物流、信息流和价值流，此为“当前状态图”。找出需要改善的地方，再描绘一个“未来状态图”，显示价值流改善的方向和结果。在此基础上，团队决定改善的计划和行动。整个价值流管理实行全成本核算，价值流经理对整个产品族的流程、资源配置及成本运作担负着主要领导、协调和控制的责任。

对一个产品来说，以下两条主要流动路径是至关重要的：一是从原材料到达顾客手中的生产流程；二是从概念到正式发布的产品设计流程。价值流就是使一个产品通过这些主要流程所需要的全部活动，包括增值活动、必要但非增值活动和非增值活动（即浪费）三类。研究表明，企业用于增值活动的时间仅占整个流程的极小部分，大部分时间都花在非增值的活动中。

2. 价值流图

价值流图是一种帮助人们了解物流和信息流的可视化工具。它体现了从原材料到成品，再到客户的所有活动，有助于观察和理解产品通过价值流过程时的物料流动和信息流动，以及其中的增值和非增值活动，从而发现浪费和确定需要改善的地方，为改善活动确定蓝图和方向。同时也便于员工了解企业的状态，提供参与改善的机会。价值流图析研究的是生产流程，它是指将现场观察到的生产流程中物料和信息的流动情况用简图描绘下



来,对其进行分析找出产生浪费的原因,并应用精益生产原理提出改进方案,然后对生产过程实施改进的系统技术。

应用价值流图分析企业生产流程,意味着要从全盘看待问题,而不是集中于某个单独的过程;意味着将改变整体,而不仅仅是优化某个部分。价值流图分析可以是针对企业内部(又称“四堵墙以内”)的活动进行分析和改善,也可以针对“四堵墙以外”,即从供应商出货起到顾客收货为止的整个价值流的分析和改善。对于价值流图,需要从以下几点进行深入理解:

(1) 物料流和信息流是生产过程价值流的两个不可分割的重要方面。

(2) 观察研究价值流的范围可以从一个过程、一个车间、一个工厂直至整个公司、整个供应链。一般从一个车间(从进口到出口)着手,然后向下对每一个过程进行研究并采取具体的改进措施,向上扩展到整个工厂、公司直至整个供应链,改进整个系统的价值流。

(3) 价值流图析技术强调研究人员亲自到现场观察、了解、记录和画图,而不只是凭书面资料。

(4) 价值流图析包括的主要步骤是:

- 了解并记录顾客的要求。如图9—13中右上角顾客要求框所示,主要应了解顾客的需求量、种类、交付频次和要求等。

- 了解并画出工厂内材料流图。如图9—13中下部的线框所示,也是工厂内的基本生产过程,将能够连续进行的过程列入一个框内,无法连续的则在两框之间用库存三角分开。

- 收集并记录每个生产过程的数据(数据框)。在每个过程线框下记录的主要数据包括生产节拍、换型时间、操作人数、有效工作时间、设备使用率、废品率等与过程改进有关的数据。注意,这些数据应是现场收集的,而不是某些资料记载的。

- 了解库存情况。小组应对所有库存(包括线上在制品库存)进行盘点,然后记载库存于三角下面。注意,是点数而非查账!

- 原材料采购和交付的情况。如图9—13中左上角所示,了解主要供应商的供货情况。

- 了解、记录顾客订货、生产计划、原材料订货过程的信息传递途经及信息,画出信息流,如图9—13上部所示,标示出顾客订单、材料订单和生产信息传递途径和方法。

- 画出生产时间线和计算相关数据。将库存数量按照顾客需求节拍转化为时间,与生产过程时间数据一同画在时间线上,求出生产过程时间占整个时间的百分比。

以下将简要介绍价值流现状图和价值流未来图的绘制。

3. 价值流现状图

图9—13是一个价值流现状图的例子。这是一个生产汽车转向柱支架的工厂,转向柱分左置和右置两种。顾客是表示在右上角的汽车装配工厂,供应商是表示在左上角的钢铁厂。图的下部描述物料流。每一个方框表示一个单独的过程,如冲压、点焊、装配和发运。方框的下面是数据框,如加工周期、换模时间、设备的完好率、几班制和有效工作时间等,方框左右的三角形表示库存,例如在冲压过程的前面有5天的原材料库存,在它和第一个点焊过程之间有4600件左置和2400件右置的半成品,等等。

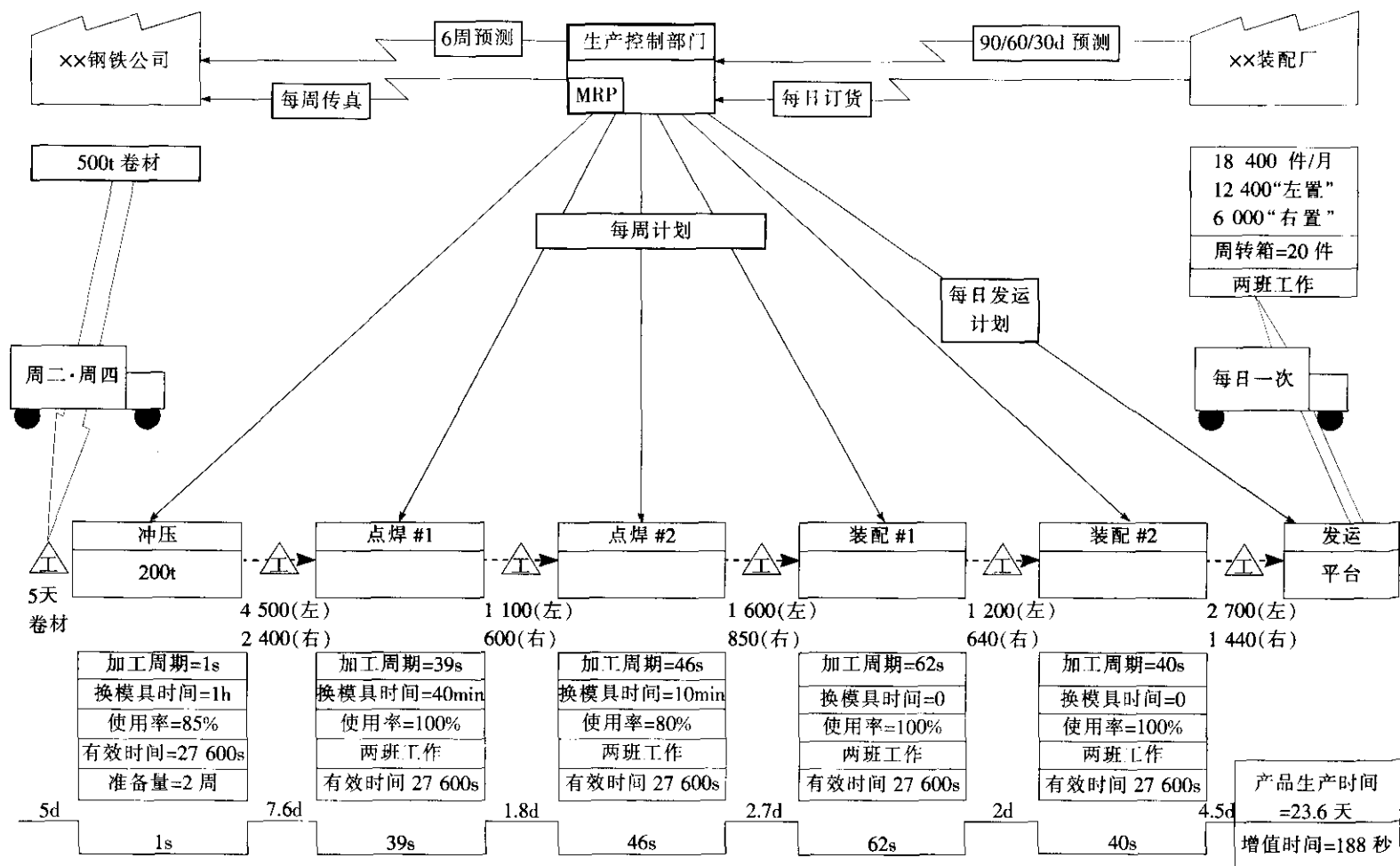


图 9—13 的上部描述信息流。生产计划部门根据顾客 90 天/60 天/30 天的预测订单和每日送货要求,应用 MRP(物料需求计划)系统制定每周生产计划,下达给每一个过程,MRP 系统按 6 周的预测数量通知钢铁工厂进行原材料生产,按每周的传真组织发货。

从价值流现状图可以看出,这个工厂当前实行的生产方式是一种按预测进行的“推动”式的批量生产。这种方式存在显而易见的浪费:为保证不中断对顾客的供应,过程之间不得不保持较大数量的库存。其他的浪费还有:人员负荷不满、设备完好率低、换模时间长、过多的搬运、缺陷率高等。如图最下面的数字表示,物料由进入工厂到变成产品的生产周期长达 23.6 天,而实际用在增值加工的时间仅为 188 秒。

4. 价值流未来图

为了减少(完全消除是很困难的)当前价值流中的浪费,需要应用精益生产的原理和方法,并考虑实际实施的可能性,改进小组首先应该勾画出一个价值流未来图。图 9—14 是转向柱支架工厂的价值流未来图。在勾画价值流未来图时,主要应用以下几个原理:

(1) 努力实现按节拍时间生产。在确定节拍时间时,应考虑生产节拍与销售节拍相一致,即节拍时间=每班有效工作时间/每班顾客需求量。在本例子中,节拍时间等于: $27\ 600\text{秒}/460=60\text{秒}$ 。

(2) 在可能的地方实现连续流动。连续流动是指产品不经中间停顿从一个过程传递到下一个过程,它是消除过程间库存和提高生产效率的有效方法。例如,把原来分割为相互独立的冲压、焊接和装配过程组合成一个连续流动的过程,并实现按节拍生产。

(3) 连续流动无法向上游扩展时,使用“超市”方式控制生产。实际中有些过程由于生产周期过快或过慢以及与其他的过程距离远等原因难以实现与其他过程节拍一致的连续流动生产。为了防止通过生产计划部门单独下达生产计划的管理方式,一个很好的经验方法是,采用类似“超市”的拉动系统。例如在本例中,在冲压过程前、冲压过程和其他过程中间以及发货过程前设置超市系统,冲压过程是按照从超市送来的看板指定的数量(一般为一个材料搬运箱,例如本例中的 20 件)和品种进行生产,不再有另外的库存,从而把库存减到最少。这要求冲压过程实现快速换模,以实现多品种、小批量生产。

未来状态图是以精益思想为指导,按照企业的实际情况,为未来的运作模式指明方向,设计新的精益流程。所谓“未来状态”,也仅仅是基于当前的技术和改善水平,在一定时间内可以达到的较为理想的目标。随着生产技术和改善水平的提高,原来的目标又变得不理想了,人们又进入了一个更高层次的改善循环。如此往复,正是精益思想中“与完美竞争,永无止境”精髓之所在。价值流未来图为制定实现精益生产方式的计划提供了基础,但真正实施还需要周密的计划和大量的工作。管理者的决心和支持当然是最重要的,员工的理解和参与也不可缺少,需要价值流经理和精益计划推动小组具体指导和推动计划的实施。按照合理的步骤实施计划并在适当的阶段组织评审,根据评审结果,对评审进行必要的修改和调整。

5. 价值流制图图标

价值流图中所使用的图标如表 9—5 所示。

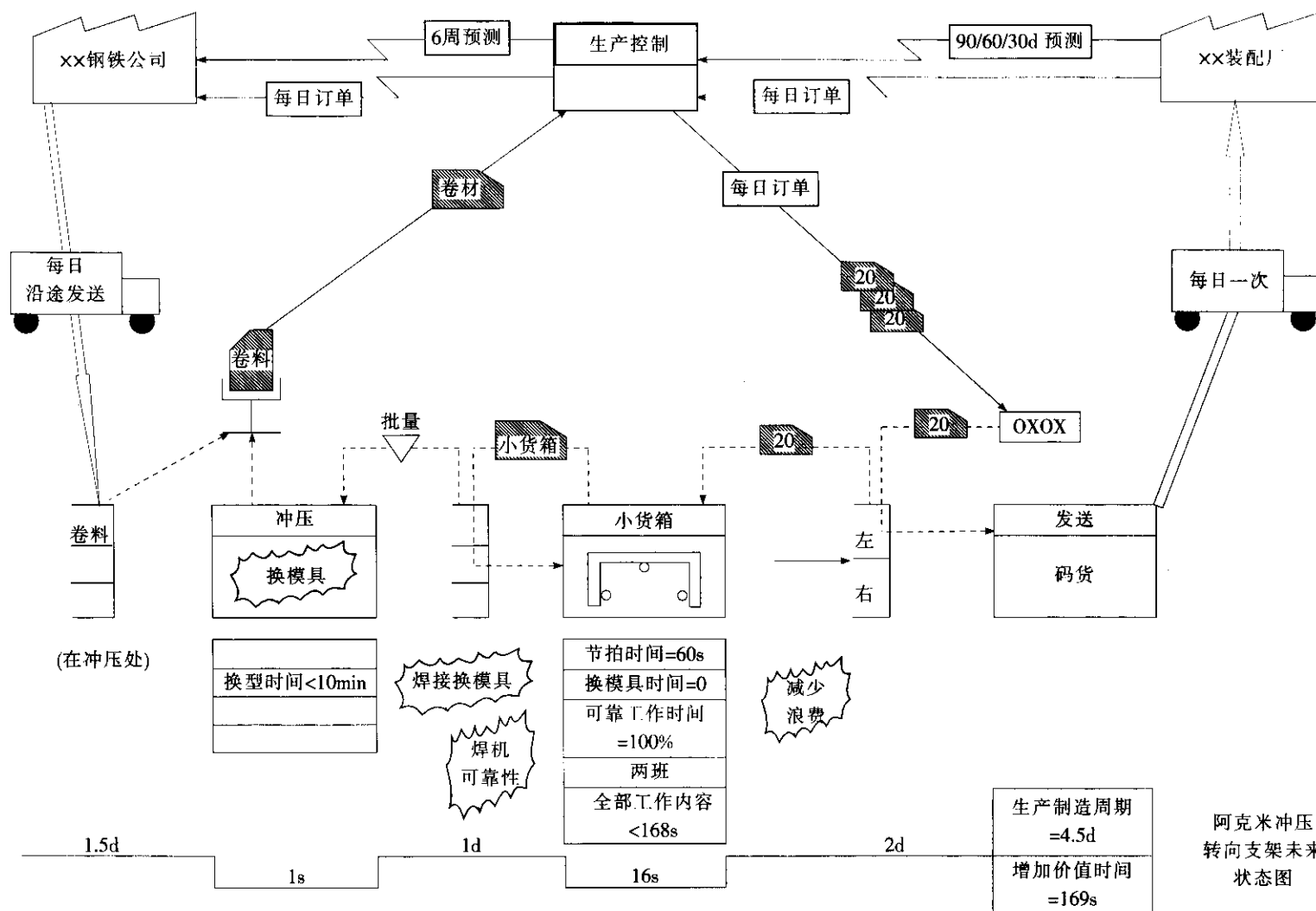


表 9—5

价值流制图图标

电子信息流	先进先出	成品	观察
库存	改善点	看板批量	看板箱
生产看板	看板信号	取料看板	均衡
手工信息流	操作工	过程框	拉动箭头
拉动圈	推动箭头	计划框	供应商
超市	卡车运输	缓冲库存	数据框

正如本章开始所述,精益本身是一个致力于发现问题、持续改进的过程,精益的思想是非常朴素的,许多精益的工具也比较简单,但是企业要真正实现精益生产却需要长期的管理积累,尤其是现场基础管理工作,这也是我国不少企业在过去推进过精益生产(有的称为丰田生产方式、JIT或看板管理等)后效果并不十分理想的原因。读者不难看出,图9—1中的精益生产的技术体系所涵盖的内容也是十分广泛的,严格来讲,精益的核心是工业工程在企业的应用,由于历史原因,我国企业对工业工程的学科内容 and 应用价值还不甚了解,工业工程的应用广度和深度都不够,因此,现场和基础管理相对落后,这一点,不仅是对精益生产产生制约,同时也是六西格玛推广的一大障碍,很难想象一个基础管理水平很差、现场管理混乱、基础数据不全的企业能在短期内推进六西格玛并取得明显收益。工业工程的学科体系也很庞杂,内容和技术很多,如动作研究和时间研究、流程分析、人因工程、设施规划和布置等,同时还有运筹学等支撑学科和技术,鉴于篇幅所限,这里无法全部讲述。但是作者想特别指出,这些综合的知识对于在企业推进精益或精益六西格玛是非常必要的,作为六西格玛的黑带,要进一步提高自己的问题解决能力,应该特别注意学习这方面的知识。

9.3 精益与六西格玛的结合

9.3.1 精益生产与六西格玛的特点及比较

1. 六西格玛管理的特点

六西格玛管理是一种基于顾客驱动的追求卓越绩效和持续改进的管理哲学。它以产品、流程持续改进和设计为基本策略,通过消除过程变异和提高质量,大量减少不良品,降低成本,强调提高顾客价值和顾客满意度,综合提高企业的竞争能力和盈利水平。它已经成为一种理念、文化和方法体系的集成。六西格玛方法管理通过 DMAIC 的持续不断的循环,不断地消除过程变异,提高过程质量,从而降低成本,增加顾客价值,获得顾客完全满意。

(1) 强调以数据和事实为依据。六西格玛将所有工作都作为一种流程,重点以数据和事实为依据,对实际问题进行量化分析,确定影响企业绩效和顾客需求的关键质量特性,并找出主要原因,寻求满意的解决方案,进而迅速逼近实际解决方案。

(2) 减少变异,优化流程,提高质量。从统计质量管理角度来看,变异是质量问题的主要原因,因为变异会导致生产失去均衡与协调,引起生产周期不稳定,交货期变得无法预测;另一方面,过程变异也会引起产出质量不稳定,无法满足顾客要求,造成巨大损失。

六西格玛管理采用统计过程控制、实验设计、质量功能展开等多种质量工具,通过产品和流程的系统化设计和严格监控,大大提高了设计、生产过程的能力,消除了包括从设计到投产、从生产到销售所有流程的缺陷与波动,从而显著提高产品和服务质量。

(3) 注重思想观念转变,改进行为方式。六西格玛并非单纯的质量标准,它是一种工作哲学,六西格玛推行的首要目的就是要通过教育培训,改变员工以及组织的思维模式、行为方式直至企业文化。

六西格玛的执行,首先是公司从上至下改变“我一直都这样做,而且做得很好”的思想,员工以主动参与取代被动的工作习惯;实施六西格玛,要努力培养员工“预防为主,第一次就把事情做对”、“决不能容忍缺陷存在”的意识,使员工形成一种观念:产品返修率哪怕只有千分之一,对顾客来说就是百分之百。并把这种意识和观念融入企业文化中,通过规范化的工作方法,力图使每一道工序、每一件产品和每一次服务都有接近完美的质量,也就是追求零缺陷。

(4) 追求持续改进。六西格玛方法是一个连续性的上升螺旋圈,通过对整个流程周而复始、反复不断地实施界定、测量、分析、改进、控制,使流程各个环节得到修正和补充,实现流程的螺旋式提升,产品质量得到持续改进。

六西格玛不是一个静止的目标,而是一个持续改进的方向。它是顾客驱动下的持续改进模式,顾客需求的动态变化使企业必须追求持续改进。今天企业在某个质量关键点上达到六西格玛水平,明天顾客可能会提出更严格和更多更新的要求,企业要持续努力才能达到并超越六西格玛绩效。

(5) 顾客驱动。顾客驱动就是以顾客需求为导向,它是六西格玛核心价值观之一。它贯穿六西格玛改进的始末,在六西格玛分析阶段绘制 SIPOC 图 (supplier-input-process-output-customer; 供应商—输入—处理—输出—顾客),目的就是流程进行整体考虑,从顾客源头出发,反向分析所做的每一件事情能否提高顾客价值。就流程而言,六西格玛要求企业深度理解顾客当前和未来要求及期望,找出“顾客关键需求”(critical customer requirements, CCR),完全以外部顾客驱动流程设计与改进,评估流程有效性。因此,唯有顾客驱动的六西格玛管理,才能实现顾客完全满意和忠诚,企业才会得到相应回报。

(6) 跨职能合作。六西格玛采用项目小组的工作方式,通过企业内部的研发、采购、财务、生产和营销等部门的跨职能合作和及时沟通与协调,能够整体上更快速、更彻底地消除变异,实现改进。

通过大量企业的实践,人们逐渐发现六西格玛管理在实施中也存在一些问题,如六西格玛强调精英的贡献。六西格玛的成功实施离不开资深黑带、黑带等一批精英人员的贡献,他们的强力支持和推动一方面保证了六西格玛项目的成功;另一方面由于缺乏对普通员工全员参与的强调和重视,也使企业普通员工的积极性不高,这些问题都会降低六西格玛实施的应有效果,甚至导致实施失败。另外,有些企业的黑带在项目中对现场的关注不够,大部分时间是在电脑前进行统计分析,这也很容易使企业的其他人员对六西格玛黑带的作用产生怀疑。

2. 精益生产的特点

精益生产的指导思想是,从顾客需求出发,精确地确定顾客价值,识别和优化产品价值流,并通过顾客订单拉动生产,追求尽善尽美。其最终目的是通过流程整体优化,平顺物流,高效利用资源,消灭一切库存和浪费,达到用最少的投入(人员、设备、时间和场地等)向顾客提供最完美价值的目的,即持续不断地追求尽善尽美和精益求精。精益方法强调“适时、适量、适物”。

精益方法主要具有以下优势:

(1) 关注顾客,创造完美价值。价值只能由最终顾客来确定,价值只有由具有特定价格、在特定时间内满足顾客需求的产品来表达时才有意义。关注顾客,企业真正要做的是站在顾客立场,重新审视价值,消灭一切浪费,向顾客提供完美价值。精益方法从顾客角度审视从设计到生产到交付的全部流程,将全过程消耗减至最少,消除一切不增加顾客价值的流程和产品功能,不将额外花销转嫁给顾客,实现顾客需求的最大满足。

(2) 消除浪费,优化流程,降低成本。精益方法审查生产特定产品的所有活动,首先努力消除Ⅱ型 muda,减低成本,然后使在制品在整个流程中流动起来。通过连续流动,进一步消灭浪费,优化流程,并且流动越快,所暴露出的浪费会越多,越有利于精简流程、降低成本。

(3) 缩短流程周期,提高响应能力。时间是决定库存水平的重要变量,时间变量的升高往往和效率成反比,导致库存增多。减少流程周期会带来两方面的好处:1) 企业库存和成本下降;2) 缩短交货期,提高企业柔性。精益方法以最终顾客需求为起点,采用看板管理,以后道工序准时拉动前道工序,使价值连续流动,通过生产单元之间的均衡与协调,快速而可靠地减少流程周期和提前期,提高效率,同时减少大量浪费,加速资金流转,最终大大提高企业对市场变化的响应能力。精益方法通过消除流程瓶颈和等待以及并

行工程和价值连续流动来加快流程速度，这种方法适应了多品种、小批量和定制化生产模式。

(4) 强调全员参与。精益方法非常强调全员参与。员工是组织的根本资源，全员参与有效地集中了员工集体智慧和才干，为组织带来了巨大收益。全员参与使员工的积极性和热情度很高，更加关注为顾客创造价值；能主动地发现流程中存在的问题，进行持续改进并强化自己的技能、知识和经验，承担起对组织目标和任务的责任；使得员工能够有效地对改进组织方针和战略目标做出贡献；也更容易获得对工作的满足感和对组织的自豪感。

精益方法的主要不足如下：

(1) 缺乏严谨的定量分析。精益方法解决问题的特点是更多地依赖专家经验与直觉，不能使用量化方法与专业工具管理流程，难以解决复杂的、综合的以及不明确的问题，决策也不易做到科学、准确和高效，难以真正实现“精益”。

(2) 缺乏系统性改进方法的整合。虽然精益能够从企业整体考虑消除浪费，但是它缺乏系统性改进方法的整合，过多追求短平快效果和短期利益，所以容易出现“非精益化反弹”现象。

(3) 易受波动影响。企业实施精益方法失败绝大多数是因为波动造成的，这些波动因素概括为人员、机器、材料、方法、测量、环境（5M1E），实施精益生产前之所以要先导入5S和TPM等方法，就是要使5M1E因素的波动最小化。如果不能实现流程稳定，也就很难真正实现“精益”。

表 9—6 是六西格玛管理与精益方法的综合比较。

表 9—6 六西格玛管理与精益方法的比较

名 称	精益方法	六西格玛管理
假 定	(1) 消除浪费可以改善绩效 (2) 大量的小改进更有利于组织成长	(1) 问题总是存在 (2) 测量是重要的 (3) 随着变异减少，系统产出得到改进
文化基础	东方以人为本的文化和管理思想	西方制度、工具化的文化和管理思想
直接目标	(1) 消除一切浪费，降低成本 (2) 缩短流程周期，增强响应能力 (3) 多品种小批量生产，增加柔性	(1) 消除变异，优化流程 (2) 提高质量，增加价值
关注焦点	价值流	问题
工具方法	5S现场管理、准时生产、快速换模、Poka-Yoke、看板、并行工程、视觉控制、自动化、平顺化、TPM、约束理论、持续改进、DFM/A（面向可制造性/可装配性设计）、价值工程和标准化作业等	分层法、散布图、排列图、因果图、关联图、系统图、亲和图、矩阵图、矩阵数据分析法、过程决策程序图、SIPOC图、QFD、FMEA、DOE、SPC、Benchmarking、测量系统分析、方差分析、响应曲面方法和平衡计分卡等
实施步骤	(1) 精确地确定价值 (2) 识别价值流 (3) 流动 (4) 拉动 (5) 尽善尽美	(1) 界定 (2) 测量 (3) 分析 (4) 改进 (5) 控制

续前表

名 称	精益方法	六西格玛管理
共 同 点	(1) 关注顾客满意, 顾客驱动 (2) 关注财务成果 (3) 注重持续的系统整体改进 (4) 都重视改变思想观念和行为方式 (注: 精益要求把“部门”和“成批排队”思想转换成“团队”和“连续流动”概念) (5) 全员参与, 团队相互合作与协调 (6) 管理层的大力支持与参与 (7) 注重人、系统和技术集成	
特 点	(1) 工具软性, 但很有效 (2) 注重柔性, 灵活性, 迅速机动 (3) 强调节流	(1) 工具精良, 功能强大 (2) 注重系统性、规范化 (3) 强调开源与节流
实施方式	自下而上推动	自上而下推动
主要效果	(1) 减少一切浪费 (库存、缺陷等) (2) 优化流程, 缩短交货期 (3) 提高生产率 (4) 降低成本, 改善资本投入	(1) 减少变异, 统一产出 (2) 消除缺陷, 改进质量 (3) 增加顾客价值, 提高利润 (4) 顾客满意与忠诚
长 处	(1) 持续的全面创新和变革 (2) 强调连续流动和拉动 (3) 与相关利益主体全面合作关系 (4) 整体优化, 追求尽善尽美 (5) 见效快	(1) 应用大量统计工具, 精确界定问题 (2) 流程彻底改进和设计 (3) 追求完美和精益求精 (持续改进)
不 足	(1) 过多依赖经验管理, 缺乏定量分析 (2) 对波动处理不力, 难以“精益” (3) 疏于人才培训和系统方法整合 (4) 急功近利, 出现“非精益化”反弹	(1) 无法提高流程周转速度 (2) 不鼓励创新和变革
精益六西格玛优势	通过持续快速改进, 消除浪费与缺陷, 低成本地快速满足顾客需求, 获得竞争优势	

9.3.2 精益与六西格玛的有机结合

通过对六西格玛管理与精益生产的优势与不足的分析, 可以看出, 二者具有很强的互补性。六西格玛管理是解决问题的方法论, 有一个非常好的解决问题的框架。它以数据分析为基础, 旨在通过消除过程变异、持续改进获得近乎完美的质量, 进而获得顾客完全满意和顾客忠诚。六西格玛关注于质量和价值, 其核心目的是增加价值, 满足顾客。精益方法是一种消除浪费、优化流程、准时制造的方法。它把何处都可作为改善起点, 其最终目的是用尽善尽美的流程为顾客创造尽善尽美的价值。精益关注于成本和速度, 其核心是降低成本、提高效率。如果能够把精益生产和六西格玛方法进行整合, 就可以同时获得二者的优势, 摒弃它们的不足。

精益六西格玛是精益方法与六西格玛管理的有机融合, 如图 9—15 所示。作为一种先进管理模式, 精益六西格玛不仅能通过六西格玛管理大幅度提升产品质量, 增加顾客价值, 同时能利用精益方法减少资本投入、提高效率和市场响应能力。

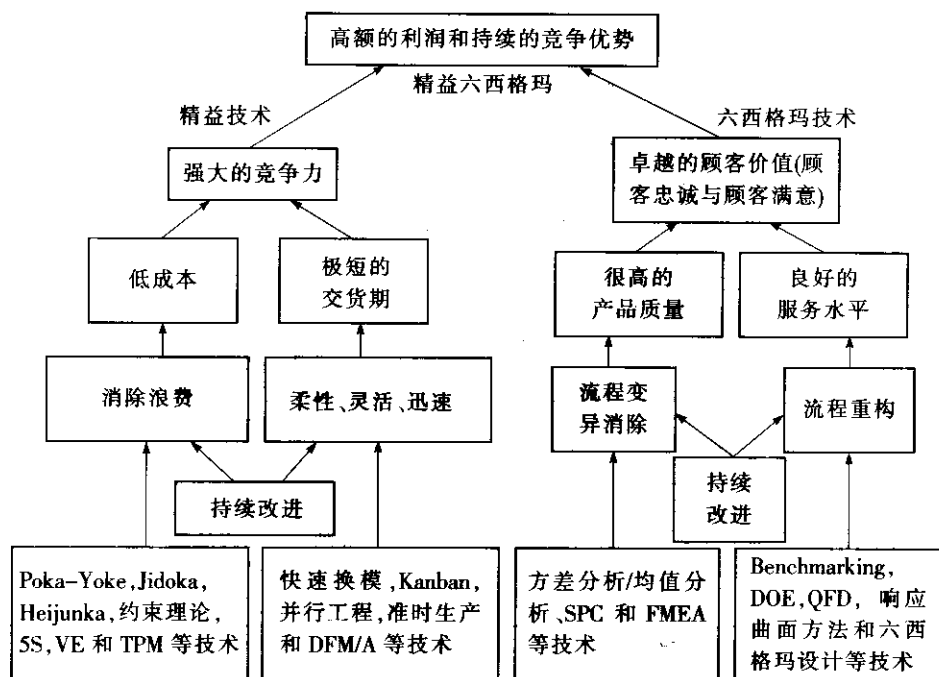


图 9-15 精益六西格玛技术体系

9.3.3 精益六西格玛的实施路线/技术路线

基于上面的分析，可以归纳总结精益六西格玛的实施路线如图 9-16 所示，第一阶段首先通过市场调研以及竞争对手分析，确定影响顾客价值和顾客需求的关键质量特性以及需要改进的领域。

第二阶段，通过头脑风暴、因果分析以及失效模式与效应分析找到问题的原因，并利用 QFD 将关键顾客需求转化为产品特性措施以及生产工艺措施，找出影响 TQC 的关键因素，并确定因素水平。

第三阶段，通过 VE/VA 或价值流分析不断地消除产品和生产过程中的非增值环节以及瓶颈，平衡物流。

第四阶段，首先进行测量系统分析和多变异分析，在此基础上确定抽样方案，进行抽样与数据测量。

第五阶段，通过试验设计，找出各个因素的最优水平进行优化，同时对产品和生产过程进行顾客参与模块化设计。通过不断的优化人、机、料、法、环 5M1E 因素，减少它们的波动。

在此基础上，通过看板管理、快速换模、单件流、防错法以及生产与物料控制，实行拉动生产。最后一个阶段是控制与标准化，通过设定标准作业程序，实行统计过程控制等方法优化流程、巩固改进成果，并通过全企业的经验推广与学习交流，把改进成果进一步扩大。

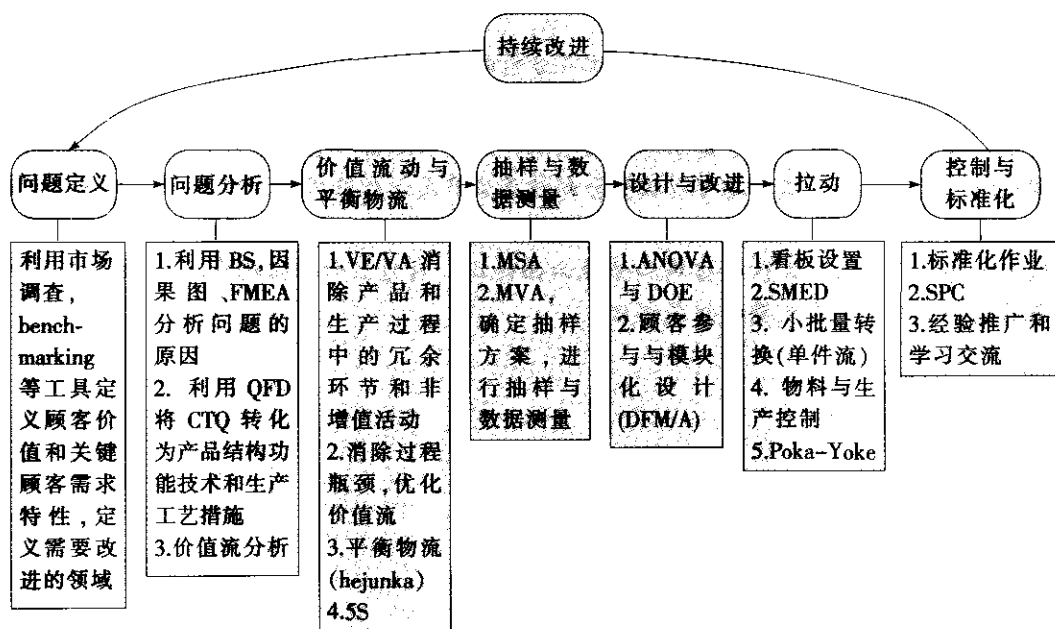


图 9-16 精益六西格玛实施技术路线

通过以上步骤周而复始不断推行, 精益六西格玛最终可以帮助企业获得以下四方面的优势:

- (1) 消除库存和浪费, 降低成本;
- (2) 消除过程波动, 提高质量;
- (3) 实现多品种小批量生产, 增加企业柔性;
- (4) 推行看板拉动生产, 实现快速准时交货。

实际上, 目前许多企业在推进六西格玛的过程中已经将精益的一些思想和工具应用于六西格玛项目, 不管名称是六西格玛或精益六西格玛, 精益和六西格玛的结合是企业六西格玛管理应用于企业系统的业务流程改进的必然。企业中需要精益和六西格玛的结合来解决的项目很多, 如:

- 减少原材料和零部件采购成本;
- 减少备品备件采购成本;
- 缩短新产品开发周期;
- 缩短产品线换型时间;
- 减少在制品占用;
- 提高设备利用率;
- 改善设备布置;
- 改善生产系统的物流。

另外需要指出, 许多服务业在推进六西格玛的过程中, 也特别注意引入精益的思想和方法, 读者不难看出, 与六西格玛一样, 精益的思想在服务业也是大有作为的, 关于精益服务等其他话题, 本书不再赘述。

本章小结

本章在对精益生产和六西格玛管理比较分析的基础上主要介绍了精益生产和六西格玛融合的理论基础、必要性和可能性，同时简要介绍了常用的精益工具和方法。需要特别指出的是，精益生产的根本在于建立精益的思想，正如六西格玛管理一样，任何可以应用于过程改进、消除浪费的工具都可以被整合到精益生产或六西格玛管理中，所以本书没有也不可能完全覆盖或详尽阐述每一种工具。精益和六西格玛管理的融合是一个大趋势，笔者在这里不再从字面上区别六西格玛和精益六西格玛，因为在六西格玛管理的知识体系中，精益思想和工具是一个重要的组成部分。

六西格玛设计

10.1 六西格玛设计概述

10.1.1 六西格玛设计的产生和发展

六西格玛改进（DMAIC 方法）是对现有流程的改进，即针对产品/流程的缺陷产生原因采取纠正措施，通过不断改进，使流程趋于“完美”。然而通过六西格玛 DMAIC 方法对流程的改进是有极限的，即便发挥 DMAIC 方法的最大潜力，产品的质量也不会超过设计的固有质量。试想如果产品设计时并没有赋予其高的质量（六西格玛的质量），怎么可能通过 DMAIC 改进来实现六西格玛的质量呢？当然通过增加投入在一定程度上也可提高绩效，但可能得不偿失。据国外调查统计，当改进使流程水平达到约 4.8 西格玛水平时，就再难以突破，这就是我们常说的所谓“五西格玛墙”。面对“五西格玛墙”，六西格玛设计便应运而生了，因为此时最明智的选择就是：对现有产品进行以六西格玛质量为目标的设计改进；对新开发的产品和业务流程，一开始就采用六西格玛设计。实践证明经过六西格玛设计的产品/流程在运行中质量波动很小，六西格玛改进工作量将会大大减少。唯有进行六西格玛设计赋予产品高的质量，才能实现真正意义上的六西格玛质量。

通用电气公司（GE）是实施六西格玛的先驱，1995—1998 年期间开展六西格玛改进取得了巨大成功，但是，推动这场革命的 GE 公司总裁杰克·韦尔奇则说，如果当时同步开展六西格玛设计，通用电气将会获得更加辉煌的成就。韦尔奇对六西格玛改进和六西格玛设计的理解也非常形象，他说六西格玛改进是引进了修理工，六西格玛设计则是引进了设计工程师，六西格玛设计的启动应先于六西格玛改进。1998 年 GE 宣布未来所有新产品和新业务都要采用六西格玛设计，1999 年 GE 医疗系统的利润增长达 25%，有 7 种产品应用了六西格玛设计。2000 年 GE 又有 22 种产品采用了六西格玛设计，并宣布 2001 年 50% 以上销售额将来自六西格玛设计的产品。除了 GE，摩托罗拉、陶氏、霍尼韦尔等世界著名公司都加入了六西格玛设计行列，2003 年初微软也宣布，今后所有的新产品设计必须采用六西格玛设计。今天，可以高兴地看到人们用六西格玛改进来调整当前的状况，用六西格玛设计来开发完美的未来产品，六西格玛改进和六西格玛设计已成为企业追求卓越的左膀右臂，在企业不同领域发挥着重要的作用。

10.1.2 六西格玛设计的重要意义

1. 从技术角度来说明

设计质量决定了产品的固有质量。从产品研制的时间序列来看,产品设计、工艺设计、生产控制等不同阶段对产品质量的影响是不同的,影响最大的是产品设计,其次是工艺设计,再次是生产控制。因为一旦产品技术规范和图纸形成后,固有质量已基本形成,其后的工艺设计是为了保证设计要求的实现,生产控制则是为了达到技术规范的要求(符合性),几乎没有进一步提高产品固有质量的可能。目前进行的六西格玛改进活动大都集中在生产控制阶段,也就是对现有产品/流程进行测量、分析、改进并将改进成果固化(控制),以减少缺陷,达到产品的固有质量。为了真正实现六西格玛质量,必须开展六西格玛设计,只有在设计阶段就赋予产品很高的固有质量,才有可能实现六西格玛的质量目标。

2. 从管理角度来分析

朱兰博士提出的质量管理三部曲是质量策划/质量设计、质量控制、质量改进。质量策划/质量设计是三部曲中最重要,对质量的贡献起着四两拨千斤的作用,质量的理论成本也是在质量设计阶段形成的。但是企业界对质量策划/质量设计的重视远不及对质量控制、质量改进的重视。20世纪40年代以来在质量控制、质量改进上花了很多力气,从休哈特的控制图,到各种质量改进工具、方法,如QC七工具等,林林总总不断推出并得到广泛应用,但主要是在质量控制和质量改进的圈子里忙碌(见图10—1)。

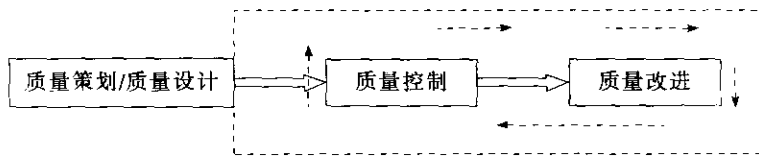


图 10—1 传统基于质量改进工作模式

随着社会生产力的发展,质量学界开发了许多适用于设计阶段的非常有效的方法,如质量功能展开、系统设计、参数设计、容差设计、FMEA、DFX等技术方法,但这些方法却没有像SPC、QC七工具那样得到广泛应用。实际上只有把质量管理向产品的源头延伸,在设计上采用更好的方法,才能在管理上赢得主动。目前大部分企业把在现场忙于“救火”的质量人员看作企业的英雄,对产品开发的质量管理却不够重视,造成设计阶段遗留的质量问题被带到生产、使用阶段,由此埋下的祸根难以消除。我们必须改变长期以来形成的错误的工作导向,变“救火”为“防火”。六西格玛设计正是这样的防火工程,可以减少或杜绝“失火”,根除隐患。

3. 从经济角度来透视

(1) 质量改进行动启动得越早,质量成本就降得越低。同样一个问题,发现得越早,付出的代价越小。据国外调查统计,如果某个质量问题在草图设计中被发现,采取措施进行改进的代价是1美元的话,那么该问题在产品生产阶段被发现,采取纠正措施进行质量改进的代价将是100美元;在出厂检验时被发现,采取措施的代价将要花去10 000美元;在顾客使用中被发现,甚至在顾客使用时发生了严重质量事故,解决问题的代价可能会达

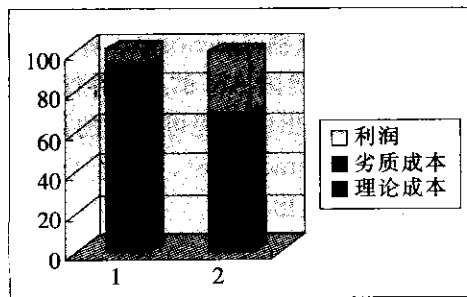
到10万甚至100万美元。这种损失是呈几何级数增加的。六西格玛设计就是在产品/流程设计一开始就寻找潜在缺陷,将一切可能的问题消灭在萌芽状态,努力创造一个新的更好的产品/流程,这种预防性改进的效益当然是最高的。

(2) 六西格玛设计为企业带来了多方面的财务收益,包括:

- 如果某产品能很好地满足顾客需求,该产品在市场上的销售量必然增加,从而为企业带来利润的增加。

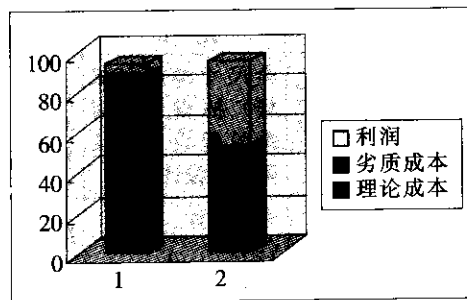
- 如果产品具有魅力质量,超出了顾客的预期,企业就有提高价格的适当理由,将为企业带来额外的利润。

- 采用六西格玛设计中的稳健设计方法将使产品实现低成本下的高质量,产品的理论成本必然下降,即原材料、制造、装配等成本下降为企业带来了直接的经济效益(见图10—2)。此外高质量意味着产品质量稳定,质量故障等引起的不良质量成本下降,为企业带来的利润也是显而易见的(见图10—3)。六西格玛改进可以降低产品的不良质量成本;而六西格玛设计则既可以降低产品的理论成本,又可以降低产品的不良质量成本。



1—DFSS 前 2—DFSS 后

图 10—2 六西格玛设计前后理论成本变化示意图



1—DFSS 前 2—DFSS 后

图 10—3 六西格玛设计前后不良质量成本变化示意图

- 六西格玛的稳健设计还使产品具有很高的抗干扰能力,这是稳健设计的又一显著特点,比如说同一产品不论在西藏还是在海南的环境下都能很好地实现其功能,并能长期保持功能的稳定。产品的抗干扰能力使产品的可靠性提高,使制造企业的维修费用和使用者的运行费用同时下降,对企业有好处,对社会也有好处。

- 六西格玛设计使产品的工程修改次数减少,研发周期大大缩短,不仅使产品能及时投放市场,为企业带来新的效益增长点,还降低了产品的开发成本,一举多得。

● 开展六西格玛设计的产品在制造、装配等过程中降低了故障和缺陷的发生概率,减少了排除故障和返工的工作量,使企业质量管理工作进入良性循环。

10.1.3 六西格玛设计流程

1. 现有的六西格玛设计流程

六西格玛设计就是按照合理的流程、运用科学的方法准确理解和把握顾客需求,对产品/流程进行稳健设计,使产品/流程在低成本下实现六西格玛质量水平。同时使产品/流程本身具有抵抗各种干扰的能力,即便环境恶劣或操作不当,产品仍能满足顾客的需求。

六西格玛改进的 DMAIC 流程已经成为美国许多开展六西格玛管理企业的标准流程。与 DMAIC 相似,六西格玛设计也有自己的流程,但目前还没有统一模式。迄今为止,专家们已提出的六西格玛设计流程主要有以下几种:(1) DMADV 流程,即界定(define)、测量(measure)、分析(analyze)、设计(design)、验证(verify),这种模式可以更好地利用 DMAIC 流程的基础,但一般只适用于现有产品或流程的局部重新设计;(2) DMEDI 流程,指的是界定(define)、测量(measure)、探索(explore)、研发(develop)、实现(implement);(3) IDDOV 流程,即识别(identify)、界定(define)、研发(develop)、优化(optimize)、验证(verify),是由美国 ASI 的总裁乔杜里先生提出的,它更适合新产品和新流程的开发要求,但乔杜里未给出该流程可操作的详细工作内容;(4) ICOV 流程,即识别(identify)、特性实现(characterize)、优化(optimize)、验证(verify),是由美国维恩州立大学的 Kai Yang 教授提出的。

2. 本书提出的六西格玛设计流程

本书借鉴 ASI 的 IDDOV 流程,但将研发(development)阶段改为设计(design)阶段,并提出了 IDDOV 流程各个阶段的主要工作内容、技术方法和设计输出(见表 10—1)。

表 10—1 六西格玛设计 IDDOV 流程的主要工作内容、技术方法和设计输出

阶段	I (identify) 识 别	D (define) 界 定	D (design) 设 计	O (optimize) 优 化	V (verify) 验 证
主要工作内容	寻找市场机会,识别顾客需求,进行项目论证	顾客需求的确定和展开,产品方案论证和设计	全尺寸样机(试样)的设计、制造过程设计、保障资源设计	产品和过程设计参数的优化,正样的设计	样机的试制,产品设计方案的正确性和达到的质量水平的验证
主要技术方法	QFD 卡诺分析 新 QC 七工具 风险分析	QFD 系统设计 DFX 功能 FMEA 新 QC 七工具 风险分析 LCC 分析	系统设计 QFD FMEA DFX DOE 参数设计 容差设计 CAD/CAM 研发试验 FRACAS LCC 分析 DFSS 计分卡	DOE 参数设计 容差设计 DFX FMEA LCC 分析 CAD/CAM 仿真 优化试验 FRACAS DFSS 计分卡	仿真试验 V&V 试验 可靠性试验 寿命试验 鉴定试验 小子样 SPCATP FRACAS MTBF S/N 比 DFSS 计分卡

续前表

阶段	I (identify) 识别	D (define) 界定	D (design) 设计	O (optimize) 优化	V (verify) 验证
主要设计输出	1. 项目可行性研究报告 2. DFSS 项目特许任务书	1. 产品设计方案 2. 技术规范	1. 样机生产用图纸和工艺文件 2. 产品规范 (试运行稿) 3. 售后保障体系 (含保障资源) 设计方案初稿	1. 产品生产图纸和工艺文件 2. 产品规范 3. 售后保障体系 (含保障资源) 设计方案	1. 设计验证试验报告 2. 设计鉴定报告 3. 过程能力分析 4. DFSS 项目绩效报告

3. IDDOV 流程各阶段的工作步骤

(1) 识别 (I) 阶段又可划分为“寻找市场机会”、“识别顾客需求”、“制定项目特许任务书”三个步骤,要应用 QFD、卡诺分析、新 QC 七工具、风险分析等方法寻找市场机会、识别顾客需求、论证和确定要开展六西格玛设计的项目、组织项目团队、落实人员和职责,并编制和批准六西格玛设计项目特许任务书。

(2) 界定 (D) 阶段又可划分为“顾客需求的确定和展开”、“产品总体设计方案的论证和确定”两个步骤,我们要通过质量功能展开深入分析,将顾客需求逐层地展开为设计要求、工艺要求、生产要求,并采用系统设计 (包括 TRIZ)、DFX、功能 FMEA、新 QC 七工具、风险分析、全寿命周期费用 (life cycle cost, LCC) 分析等方法,通过创造性思维和自顶向下的设计,形成一个可以实现顾客需求的总体设计方案。

(3) 设计 (D) 阶段又可划分为“初步设计”、“全尺寸样机的设计”、“过程设计和样机的试制”三个步骤,要采用系统设计、QFD、FMEA、面向 X 的设计 (DFX)、参数设计、容差设计、CAD/CAM 等方法进行产品的初步设计 (技术设计)、产品各子系统、部件、设备和供应商的确定、全尺寸样机及其制造过程的设计、保障资源的设计等。

(4) 优化 (O) 阶段又可划分为“产品设计的优化”、“过程设计的优化”两个步骤,要通过稳健设计 (DOE、参数设计、容差设计) 和 FMEA、DFX 等方法,使产品质量特性稳定在目标附近 (变异小),在使用中抗干扰,并进行过程设计的优化。

(5) 验证 (V) 阶段又可划分为“设计质量的验证”、“制造质量的验证”、“产品的验证与确认”三个步骤,要通过小子样 SPC、验收检验规程 (acceptance testing procedure, ATP) 等方法进行过程能力的分析、制造质量的验证,通过仿真试验、V&V 试验、可靠性试验、寿命试验、鉴定试验等方法进行六西格玛设计产品的验证与确认,以及通过平均故障间隔时间 (MTBF) 和信噪比 (S/N 比) 的统计及六西格玛设计计分卡等来考察产品的质量可靠性水平,并请顾客试用来验证六西格玛设计是否达到了希望的目标。当然,最优化的方案还应当通过技术状态控制的方法固化下来,以保证设计的产品在后续加工过程中完全符合顾客的需求。

总之,六西格玛设计 IDDOV 是一整套不同于常规设计的方法,它以顾客需求为导向,以质量功能展开 (QFD) 为纽带,深入分析和展开顾客需求,综合应用系统设计 (包括创造性问题解决理论 TRIZ)、试验设计、参数设计、容差设计以及普氏矩阵、FMEA 等设计分析技术,大幅度地提高产品的固有质量,从而更好地满足顾客的需求。六西格玛设计

还体现了激发创新能力、加快创新速度、预防为主等思想理念。六西格玛设计工作一般结束于新产品设计鉴定并投入批量生产时,在投入批生产和投放市场后出现少量不满足顾客需求的问题,还可以通过六西格玛改进 DMAIC 来解决。

10.1.4 六西格玛设计部署

1. 管理承诺

实施六西格玛设计就像实施六西格玛改进一样,关键在于从高层到中层管理者的认同。无论一个计划有多好,若没有高中层管理者的全力支持,实施起来都会极其困难。因此必须确保管理层充分认识到实施六西格玛设计的意义、价值,并且相信六西格玛设计项目绝不会对他们的职务和在公司的地位构成威胁。管理层的强有力支持会使企业的每个人都感受到高层领导实施六西格玛设计的决心。如果高层领导不仅支持而且还有政策鼓励,中层的领导也会效仿,六西格玛设计才能得到自上而下的推进。

2. 公司领导和项目团队

一旦各级组织都认同了六西格玛设计的实施办法,挑选六西格玛项目领导者(倡导者、资深黑带和黑带)就成为当务之急。倡导者在公司层面上发起和支持黑带项目,资深黑带负责培训黑带,提供指导、咨询、进行管理、监督和协调。黑带在倡导者和资深黑带的指导下,界定六西格玛设计工作项目,带领团队完成项目。项目领导者必须是高效的 manager,能自我激励并激励他人,他能让别人感到有依靠,对项目的目标承担责任,敢于授权、注重细节、严格监督。简而言之,必须找一个领军人物来领导项目。组织应该确信为项目选择了合适的人物,此时应举行一个由高层、中层和项目领导参加的公开会议。项目领导应熟悉六西格玛设计项目所涉及的各种因素,对各部门工作充分了解,能够行使权力、掌握项目管理技巧等,并受人尊敬。另一个影响项目成功的重要因素就是要建立一个充分授权的、多专业的、跨部门的项目团队。六西格玛设计项目应该由市场营销、研发、工艺、生产、质量管理、售后服务等方面的人员组成跨部门的团队,团队的成员应接受绿带培训、协同工作;为确保实施的成功,组织应制定并贯彻激励机制。

3. 有效沟通

确立准确有效的沟通方式是跨部门项目团队的职责。整个组织的每一个成员对于相同的信息应具有相同的理解,公司应该提供关于六西格玛设计的词汇表,最好用人人都能明白的语言来表达。团队的成员都要理解这些概念、定义,如六西格玛设计的含义是什么?如何通过 QFD 来诠释顾客的需求? TRIZ 方法是什么?稳健设计的含义是什么?如此等等。让每个成员通过高效的交流渠道进行交流,一旦收到信息就能做出快速反应,准确理解信息的内容和含义。六西格玛设计要求清晰、简洁、及时的信息交流,这对团队至关重要,误会或曲解信息将导致工作绩效大大降低。

4. 教育和培训

与六西格玛改进的培训相同,六西格玛设计也应结合项目进行。分段进行培训和实践,在经过较长一段时间例如半年或一年后,培训结束,六西格玛设计项目也完成了。训练是六西格玛教育和培训的核心,如果只在导师指导下进行课堂练习,那你掌握不了多少东西,即使在六西格玛设计每次集中培训学习时,我们仍要花大量时间让学员交流、参与研讨并训练他们的展示(呈现)能力。六西格玛设计对培训导师也提出了更高的要求,导

师的角色将变成教练,他们不仅要懂项目管理、统计技术、六西格玛设计方法,更要有丰富的工程实践经验,能胜任指导团队实现目标的要求。实践证明,这种理论联系实际的学习方式非常有效。

5. 整合战略

国外企业的实践证明,六西格玛设计同其他工作及 ISO 9000 体系的实施整合在一起,将会显著提高企业的绩效,否则可能带来不必要的矛盾和冲突,企业业务必予以高度重视。六西格玛设计是一种全新的思路,要固化到企业的常规业务中去,固化到质量体系文件中。一切有效的做法都要文件化,这样既可以巩固六西格玛设计成果,又会促进企业变革。

6. 提高盈利能力

项目与盈利能力的结合会促成企业文化变革,试想当企业领导人发现应用新方法的财务数据明显比以前要好,他有什么理由不进行变革呢?全面质量管理也为企业赢得了显著的收益,但企业领导无法区分哪些是全面质量管理的贡献,哪些是其他因素促成的;全面质量管理的投资回报率难以计算,企业推行全面质量管理主要是基于对这种方法的信任而不是所看到的财务收益。而六西格玛管理强调投资回报,并将其作为项目确认和项目选择流程的一部分,然后一直跟踪到实现项目设定的盈利能力。这种方式将会改变领导的决策心理,他们会毫不犹豫地投资支持和推动六西格玛设计项目。

六西格玛设计的组织结构、人员角色等与六西格玛改进基本相同,不赘述。

10.1.5 六西格玛设计的指导思想

1. 以顾客为关注焦点

六西格玛设计的本意是要减少产品/过程的质量波动,要求在产品质量特性均值偏离设计目标值 1.5σ 时,不合格(差错)率小于 3.4ppm ,其隐含的前提是:设计目标值必须与顾客的要求完全一致,并且质量特性的规格限必须是顾客可以接受的。因此,开展六西格玛设计必须以顾客为关注焦点,首先采用 QFD 方法分析和确定顾客的需求(设计目标值),并初步确定质量特性的规格限。如果忽视了顾客的需求,所确定的设计目标值和规格限并不能使顾客满意,六西格玛设计也就失去了意义。

以顾客为关注焦点还意味着应满足顾客对产品全寿命期间的各种要求,如使用方便、维修服务周到、备件和耗材供应充足、寿命足够长、价格适当、不污染环境等。

2. 提高产品抗干扰能力,减少质量波动

产品的质量波动是由于使用环境的变化(外干扰)、零部件磨损(内干扰)和加工误差(产品间干扰)引起的,为了减少质量波动,最好的办法是通过六西格玛设计提高产品的抗干扰能力(健壮性)。

图 10-4 形象地表示了产品质量特性的概率密度曲线因三种干扰因素的影响而恶化的情况。 $\phi(y)$ 为产品性能 Y (随机变量)的概率分布密度函数(pdf); μ 为 Y 的数学期望; σ 为 Y 的标准差; m 为设计目标值; S_U, S_L 分别为公差带上、下限值; T_U, T_L 分别为产品失效的上、下临界值。

按传统的“门柱法”检验准则,只要产品性能在公差带上、下限范围内,即为毫无区别的合格,超出公差带即为毫无区别的不合格。如产品 A 刚好在公差界限内,而产品 B

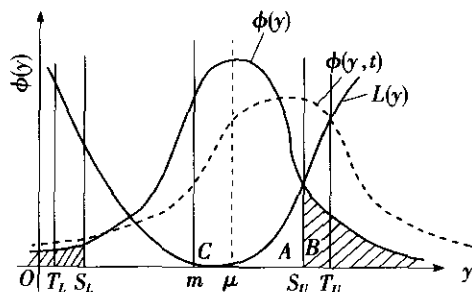


图 10—4 产品质量特性（性能）概率密度函数和质量损失函数

刚好超出公差界限值（如图 10—4 中所示），两者性能几乎一样，但是 A 为合格品，B 为废品。又如图 10—4 中，产品 C 离目标值 m 很近，但 C 与 A 在“门柱法”检验中并无区别。六西格玛设计要求提高产品的抗干扰能力、减少质量波动，形象地说就是要减肥（减小标准差 σ ）、纠偏（减小偏差 $(\mu - m)$ ），使质量特性的概率密度曲线很陡，且其均值尽可能靠近设计目标值。

另外，六西格玛设计还要使产品的质量损失曲线尽可能平坦，即相同的质量波动引起的产品的质量损失尽可能小。

3. 缩短周期、降低成本

脱离研制周期和产品成本来讨论产品质量是没有意义的。缩短周期和降低成本都是增强顾客满意的重要方面，实施六西格玛设计应当在提高产品质量的同时，尽可能缩短周期、降低成本，以提高企业的绩效。

10.1.6 六西格玛设计的主要工具

我们把应用于 DFSS 的主要技术工具整合为市场需求分析、系统设计、稳定性优化设计、面向 X 的设计、适用的可靠性工程和设计验证六个模块，见图 10—5 中的实线方框及流程。每个技术工具的模块中又包括若干个技术工具。图中的虚线框表示 DFSS 的相关技术。DFSS 技术与 IDDOV 五个阶段的大致对应关系如图 10—5 所示。

需要强调的是，IDDOV 五个阶段有先后的次序，但不是串行关系；在实施中必须贯彻并行工程，在产品研发的初期就要面向市场和顾客，考虑和着手解决产品全寿命周期中可能遇到的所有问题；每个阶段都要面向后续阶段开展研发；在不同的阶段之间需要有一定的重叠，验证阶段应当是对研发全过程的分阶段的验证。贯彻并行工程有利于缩短周期、提高质量、降低成本，实现质量、成本、进度的三位一体。

六西格玛设计的主要技术工具简介如下：

1. 质量功能展开

质量功能展开是开展六西格玛设计必须应用的最重要的方法之一。为了保证设计目标值与顾客的要求完全一致，质量特性的规格限满足顾客的需求，在六西格玛设计的第一步识别（I）阶段就要采用 QFD 方法分析和确定顾客的需求（设计目标值），并初步确定质量特性的规格限。在界定（D）阶段，需要应用 QFD 技术将顾客的需求科学地转化为设计要求，并确定关键质量特性 CTQ 和瓶颈技术。在产品的设计（D）和优化设计（O）阶

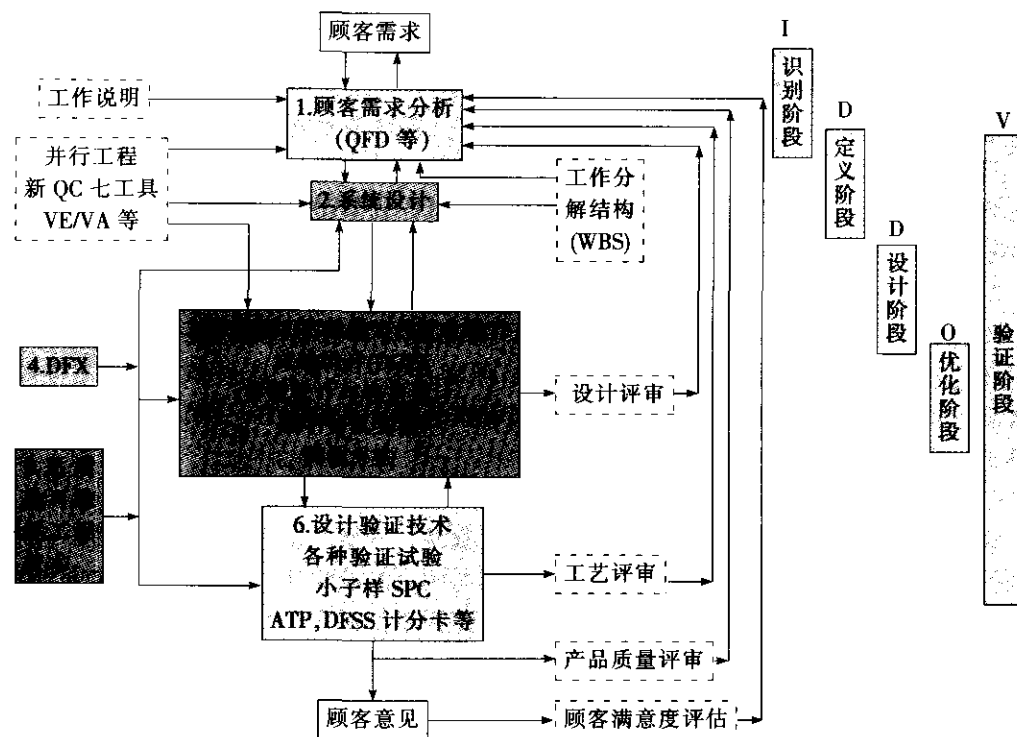


图 10—5 六西格玛设计技术与流程示意框图

段，QFD 也可以发挥辅助的作用。

2. 系统设计

系统设计 (system design) 在六西格玛设计中有着十分重要的作用。在顾客需求明确以后，如何有针对性地开发出技术含量高、生命力强、适销对路的产品，从根本上决定了产品的质量，也将直接影响企业的成败。

近年来，在质量学界的不懈努力下，人们对系统设计的过程及其一般规律有了深入的理解，提出了一些新的方法，主要有西欧流派的理论、公理性设计原则、解决创造性问题的理论 (TRIZ) 以及自顶向下的设计方法等。系统设计适用于界定 (D) 和设计 (D) 阶段。

3. 参数设计

参数设计 (parameter design) 在系统设计之后进行。参数设计的基本思想是通过选择系统中所有参数 (包括原材料、零件、元件等) 的最佳水平组合，尽量减少外部、内部和产品间三种干扰的影响，使所设计的产品质量特性波动小、稳定性好。另外，在参数设计阶段，一般选用能满足使用环境条件的最低质量等级的元件和性价比高的加工精度来进行设计，使产品的质量和成本两方面均得到改善。参数设计主要适用于优化设计 (O) 阶段。

4. 容差设计

容差设计 (tolerance design) 在完成系统设计和由参数设计确定了可控因素的最佳水平组合后进行，此时各元件的质量等级较低，参数波动范围较宽。

容差设计的基本思想如下：根据各参数的波动对产品质量特性贡献（影响）的大小，从经济性角度考虑有无必要对影响大的参数给予较小的容差（例如用较高质量等级的元件替代较低质量等级的元件）。这样做，一方面可以进一步减少质量特性的波动，提高产品的稳定性，减少质量损失；另一方面，由于提高了元件的质量等级，使产品的成本有所提高。因此，容差设计阶段既要考虑进一步减少在参数设计后产品仍存在的质量损失，又要考虑缩小一些元件的容差将会增加成本，要权衡两者的利弊得失，采取最佳决策。容差设计主要适用于优化设计（O）阶段。

5. FMEA 分析

通过 FMEA 分析，找出影响产品质量和可靠性的各种潜在的质量问题和故障模式及其危害度和原因（包括设计缺陷、工艺问题、环境因素、老化、磨损和加工误差等），经采取设计和工艺的纠正措施，提高产品的质量和抗各种干扰的能力。FMEA 分析主要适用于界定（D）和设计（D）阶段。

6. 面向 X 的设计（DFX）

顾客对于产品全寿命周期内的特性，例如可靠性、寿命、使用维护、保修期、备件耗材的保障、不污染环境、全寿命周期的费用等均有明示的或隐含的要求。产品质量特性的实现和成本的形成也受到结构设计方案以外的许多因素如工艺、制造、装配、检验、使用维护、保障服务、研制周期、成本控制等的影响和制约。因此，为了在产品全寿命周期内增强顾客满意，必须针对有关的各种要素 X，进行面向 X 族的设计（DFX）。所谓 DFX，本质上就是面向产品全寿命周期的设计。DFX 技术主要适用于界定（D）、设计（D）和优化（O）阶段。

7. 设计验证技术

主要包括设计评审、小子样 SPC、仿真试验、双 V 试验、可靠性试验、寿命试验、鉴定试验、DFSS 计分卡等。应用设计验证技术在 IDDOV 的界定（D）、设计（D）和优化（O）三个阶段对设计输出是否符合设计输入的要求进行验证；在设计验证（V）阶段，对样机制造的过程能力和样机的功能、性能和可靠性等进行全面的验证，以确保产品的研制质量达到预期的目标、满足顾客的要求。

10.1.7 产品全寿命周期的六西格玛解决方案

企业在实施六西格玛改进（DMAIC 流程）和六西格玛设计（IDDOV 流程）项目的工程实践中，经常遇到孤立地实施六西格玛改进和六西格玛设计均不适用的情况，例如在 DMAIC 流程进行到分析阶段时发现不改动设计已经无法解决问题，但整个项目又不是新产品的的设计，有时只好放弃该六西格玛项目。其实，为了实现产品的六西格玛质量水平，应当根据产品寿命周期的不同阶段和遇到的实际问题，正确地选择六西格玛的方法。

我们建立了 DMAIC 流程和 IDDOV 流程之间的联系，提出了产品全寿命周期的六西格玛解决方案。图 10—6 为包括 IDDOV 和 DMAIC 的完整的六西格玛流程图，对于产品的全寿命周期而言，六西格玛管理始于市场需求的调研和分析，首先开展 DFSS，应用 IDDOV 流程，进行市场机会的识别（I）、产品的界定（D）、设计（D）、优化（O）和样机的设计验证（V）。研发成功的产品投入批量生产，产品进入市场，然后收集和分析顾客的反馈，如果产品基本上能满足顾客需求，应本着不断改进的原则，针对存在的质量问

题和不足之处,应用 DMAIC 流程,进行改进项目的选择和界定(D)、测量(M)、分析(A),在分析过程中增加了一个决策点,判断是否必须进行产品的再设计(即局部更改设计方案),如果回答是否定的,即不必改进设计方案,只要改进过程就能解决问题,则继续 DMAIC 的流程,进行过程改进(I)和控制(C)。经过验证的控制措施要贯彻落实到批生产线上。如果必须进行产品的再设计,则转而执行 IDDOV 流程的后三个阶段即设计(D)、优化(O)和验证(V)阶段。

这样通过在 DMAIC 的分析(A)阶段增加一个判断,将 IDDOV 和 DMAIC 流程有机地结合了起来,并形成了一个派生的流程 DMADOV,即界定(define)、测量(measure)、分析(analyze)、设计(design)、优化(optimize)、验证(verify)。

如果在聆听顾客声音时发现由于市场需求的发展或技术的进步,或产品存在质量问题,需进行新产品的设计,则流程返回到“市场需求”分析。

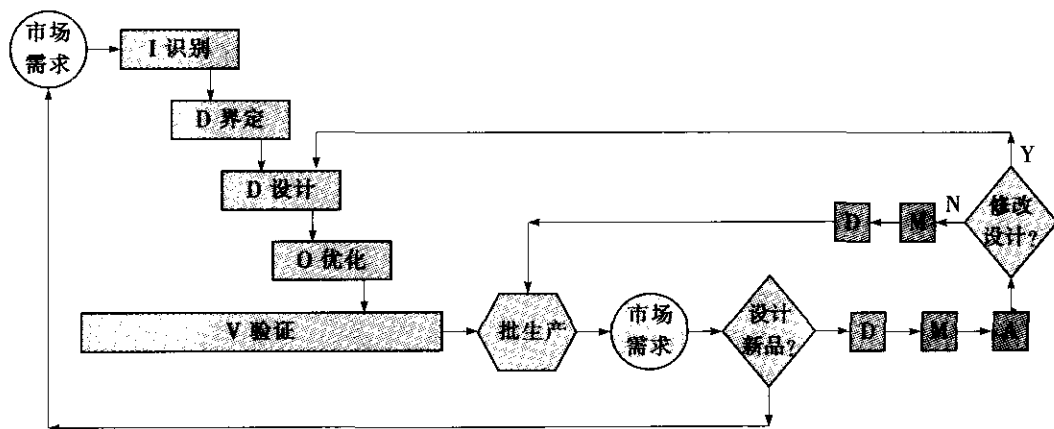


图 10—6 包括 IDDOV 和 DMAIC 的完整的六西格玛流程图

10.2 质量功能展开

10.2.1 概述

质量功能展开(QFD)是把顾客(用户、使用方)对产品的需求进行多层次的演绎分析,转化为产品的设计要求、零部件特性、工艺要求、生产要求的质量工程工具,用来指导产品的稳健设计和质量保证。QFD 技术产生于日本,在美国得到进一步发展,并在世界范围内得到广泛应用。QFD 产生初期,主要用于产品设计和生产的质量保证,但几十年来不断向管理、服务业等各个领域渗透,表现出广泛的适应性。

QFD 体现了以市场为导向,以顾客要求为产品开发唯一依据的指导思想。在六西格玛设计方法体系中,QFD 技术占有举足轻重的地位,它是开展六西格玛设计的先导步骤,通过对顾客需求的逐层展开来确定产品研制的关键质量特性(CTQ)和关键过程特性(CTP),从而为六西格玛设计的具体实施确定了重点,明确了方向。

采用 QFD 技术使产品的全部研制活动与满足顾客的要求紧密联系,从而增强顾客满

意,提高产品的市场竞争能力,保证产品开发一次成功。根据文献报道,运用 QFD 方法,产品开发周期可缩短 1/3,成本可减少 1/2,质量大幅度提高,产能成倍增加。

本章介绍的 QFD 方法,是消化吸收日、美两国 QFD 方法的精华(见有关参考文献),结合我国国情并总结了大量工程实践经验而形成的。下面将侧重介绍适用于中小规模产品开发的 QFD 方法和阐述应用于非制造业的 QFD 技术。

10.2.2 QFD 的基本方法

为了适应市场竞争,必须以顾客需求为导向进行产品开发。QFD 的基本原理就是用“质量屋”(quality house)的形式,量化分析顾客需求与工程措施之间的关系,经数据分析处理后找出对满足顾客需求贡献最大的工程措施,即关键措施,或称关键质量特性,从而指导设计人员抓住主要矛盾,开展稳定性优化设计,开发出满足顾客需求的产品。

1. 质量屋的建立

质量屋也称质量表(quality chart or quality table),是一种形象直观的二维矩阵展开图表,如图 10—7 所示。其基本结构要素如下:

- (1) 左墙——顾客需求及其重要度;
- (2) 天花板——工程措施(设计要求或质量特性);
- (3) 房间——关系矩阵;
- (4) 地板——工程措施的指标及其重要度;
- (5) 屋顶——相关矩阵;
- (6) 右墙——市场竞争能力评估矩阵;
- (7) 地下室——技术竞争能力评估矩阵。

质量屋的结构借用了建筑上的称谓,好懂易记,形象地喻示 QFD 方法的结果是使顾客可以在质量大厦的庇护下,满意地享用他们所需要的产品或服务。采用质量屋的形式进行矩阵展开,不但直观易懂,具有吸引力,而且处理和分析的信息量比 QC 老七种工具中的鱼骨图(因果图)等要大得多,在处理的深入程度和量化程度上也要好得多。

下面以圆珠笔的开发为例,使读者对 QFD 有一个初步的了解。

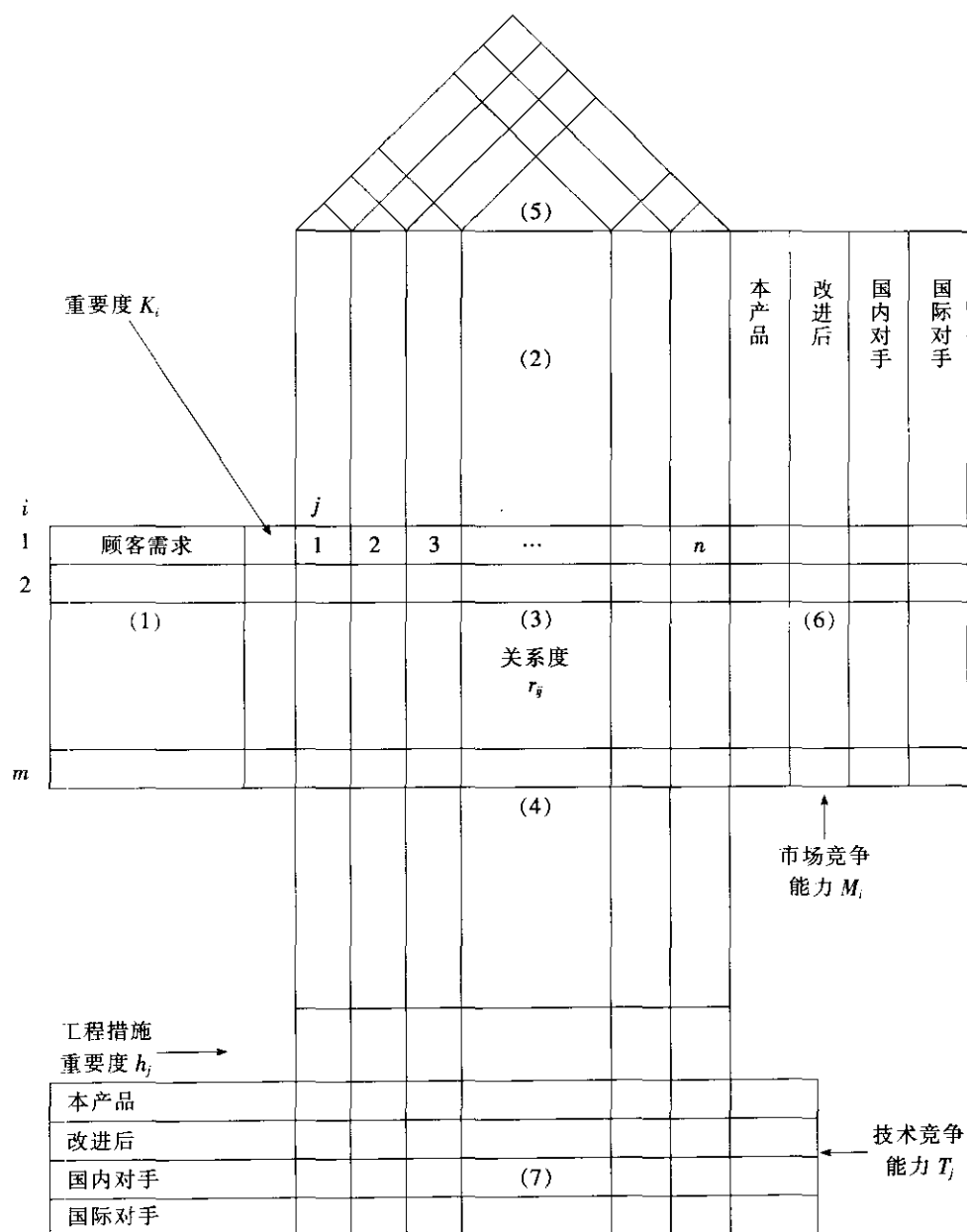
2. 顾客需求分析与工程措施的确定

顾客或市场的需求往往比较笼统、定性化,有些意见可能比较片面。另外,随着时间的推移、经济和技术的发展、消费环境的变化,市场需求也是不断变化的。应当尽可能完整、及时地收集第一手市场信息。在此基础上,对这些原始信息进行整理、加工和提炼,形成系统的、有层次的、有条理的、有前瞻性的顾客需求。这项工作是极其重要的,它是一个企业正确制定产品开发战略、设定产品质量目标的基础。

经过广泛调研,顾客对圆珠笔的要求主要有:书写要流利、字迹永不褪色、外形美观、使用方便、价格适中、有适当的耐用性。将这六条整理后作为顾客需求填入质量屋左墙。

顾客需求重要度 K_i ($i=1, 2, \dots, m$) 分为下列 5 个等级:

- 1: 表示不影响功能实现的需求;
- 2: 表示不影响主要功能实现的需求;
- 3: 表示比较重要的影响功能实现的需求;



说明：在实践中，质量屋的结构可以灵活地进行剪裁或扩充。

图 10—7 质量屋的结构

4：表示重要的影响功能实现的需求；

5：表示基本的、涉及安全的、特别重要的需求。

从技术的角度出发，应针对顾客的需求，进行产品质量特性（设计要求）的展开（需要时可以把质量特性划分层次），按隶属关系整理成表格，形成质量屋中的天花板部分。

圆珠笔的设计要求包括：笔尖组件设计、油墨浓度选择、油墨成分的确定、收放机构

设计、外形设计、成本控制和材料。这七项要求没有层次上的隶属关系，作为同级工程措施并列填入质量屋的天花板。

3. 关键措施与瓶颈技术的确定

为了从上述七项工程措施中挑选出具有关键意义的几项，首先要确定顾客需求与工程措施两两之间的关系度 r_{ij} （关系矩阵），关系度的等级规定如下：

- 1：表示该交点所对应的工程措施和顾客需求间存在微弱的关系；
- 3：表示该交点所对应的工程措施和顾客需求间存在较弱的关系；
- 5：表示该交点所对应的工程措施和顾客需求间存在一般的关系；
- 7：表示该交点所对应的工程措施和顾客需求间存在密切的关系；
- 9：表示该交点所对应的工程措施和顾客需求间存在非常密切的关系。

根据实际情况，必要时也可采用中间等级：

- 2：表示介于 1~3 之间；
- 4：表示介于 3~5 之间；
- 6：表示介于 5~7 之间；
- 8：表示介于 7~9 之间；
- 空白即为 0，表示不存在关系。

有时，也可只采用 1, 3, 9 三个关系度等级，此时，可用符号◎表示（9），○表示（3），△表示（1）。

加权后工程措施的重要度：

$$h_j = \sum_{i=1}^m k_i r_{ij}, \quad j = 1, 2, \dots, n \quad (10-1)$$

以顾客需求的重要度作为加权系数，分别计算每项工程措施与全部顾客需求的加权关系度之和并进行比较，加权关系度之和大（亦即对满足顾客需求贡献大）的那些工程措施就是所谓的关键措施或 CTQ。我们将每项工程措施对顾客需求的加权关系度之和称作工程措施的重要度，根据该重要度明确重点，集中力量实现关键的工程措施，把好钢用在刀刃上，最大限度地发挥人力、物力的作用。

关键措施的重要度应明显高于一般工程措施的重要度。如可将重要度高于所有工程措施的平均重要度 1.25 倍以上的工程措施列为关键措施。

图 10—8 给出了开发优质圆珠笔的质量屋。通过建立质量屋进行量化评估，确定了两项关键措施：油墨成分和笔尖组件设计。

4. 市场竞争能力和技术竞争能力的分析

在图 10—8 的质量屋中，对新产品预期的竞争能力（市场竞争能力和技术竞争能力）也作了分析，帮助决策者了解产品的竞争态势。

市场竞争能力： M_i （ $i=1, 2, \dots, m$ ）可取下列 5 个数值：

- 1：表示无竞争能力可言，产品积压，无销路；
- 2：表示竞争能力低下，市场占有率递减；
- 3：表示可以进入市场，但并不拥有优势；
- 4：表示在国内市场竞争中拥有优势；
- 5：表示在国内市场竞争中拥有明显优势，并具有一定的国际市场竞争能力。

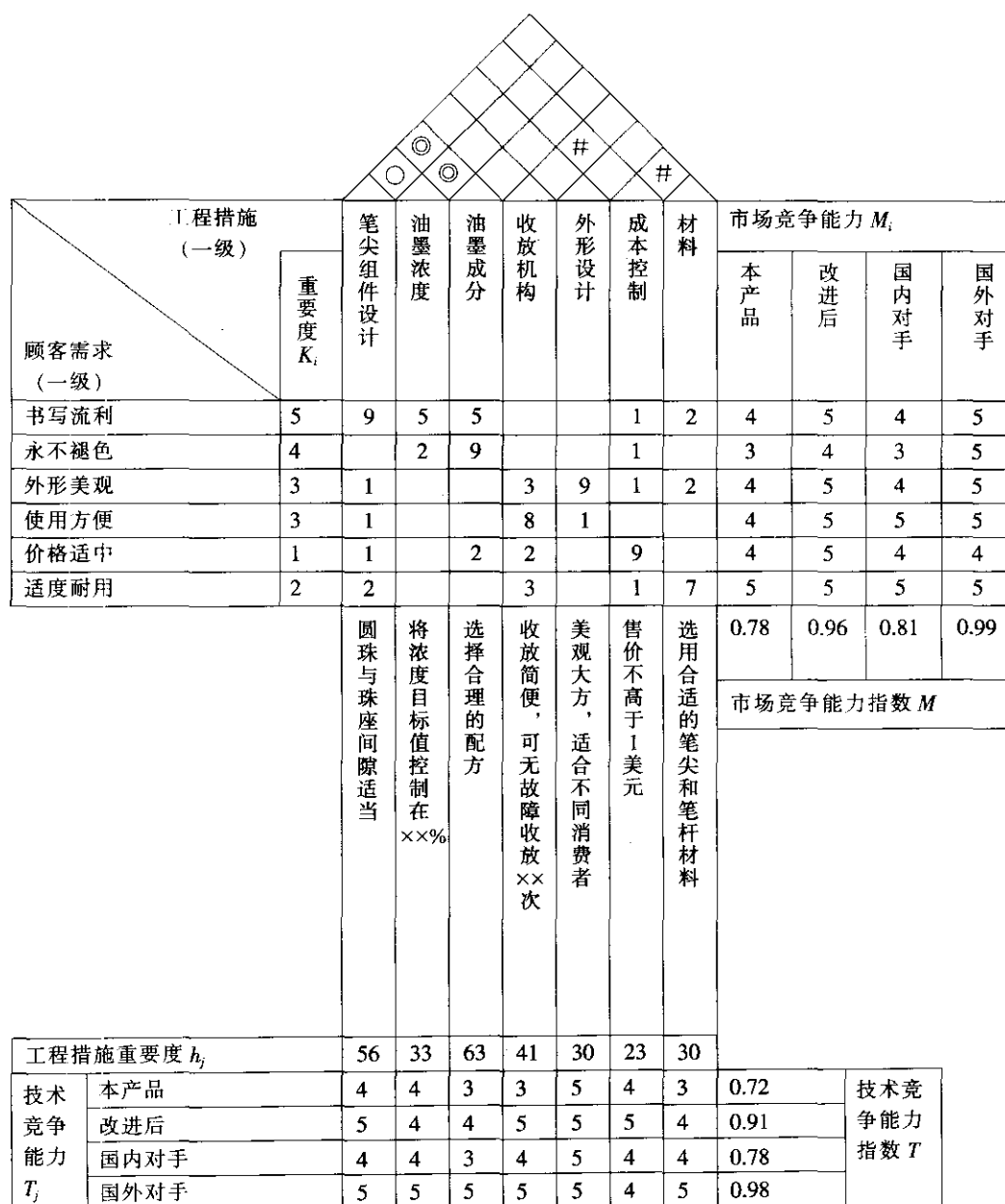


图 10—8 开发优质圆珠笔一级质量屋

对市场竞争能力 $M_i (i=1, 2, \dots, m)$ 进行综合后，获得产品的市场竞争能力指数 M ：

$$M = \sum_{i=1}^m k_i M_i / (5 \sum_{i=1}^m k_i) \quad (10-2)$$

M 值越大越好。

技术竞争能力 $T_j (j=1, 2, \dots, n)$ 表示第 j 项工程措施的技术水平。所谓技术水平

包括指标本身的水平,本企业的设计水平、工艺水平、制造水平、测试检验水平等,可取下列5个数值:

- 1: 表示技术水平低下;
- 2: 表示技术水平一般;
- 3: 表示技术水平达到行业先进水平;
- 4: 表示技术水平达到国内先进水平;
- 5: 表示技术水平达到国际先进水平。

对技术竞争能力 $T_j (j=1, 2, \dots, n)$ 进行综合后,获得产品的技术竞争能力指数 T :

$$T = \sum_{j=1}^n h_j T_j / (5 \sum_{j=1}^n h_j) \quad (10-3)$$

T 值越大越好。

综合竞争能力指数是市场竞争能力指数与技术竞争能力指数的乘积:

$$C = MT \quad (10-4)$$

C 值越大越好。

我们将现有技术很难解决的技术关键称为“瓶颈技术”,在质量功能展开的过程中必须找出瓶颈,并攻克瓶颈技术。

5. 工程措施之间的交互作用评估

通常用下列符号表示工程措施两两之间的交互作用或相关度:

- : 正相关,表示该交点所对应的两项工程措施间存在互相加强、互相叠加的作用;
- ◎: 强正相关,表示该交点所对应的两项工程措施间存在很强的互相叠加的作用;
- ×: 负相关,表示该交点所对应的两项工程措施间存在互相减弱、互相抵消的作用;
- #: 强负相关,表示该交点所对应的两项工程措施间的作用强烈排斥,有很大矛盾;
- 空白表示该交点所对应的两项工程措施间不存在交互作用。

6. 四个阶段的质量功能展开

产品开发一般要经过产品规划、零部件展开、工艺计划、生产计划四个阶段,因此有必要进行四个阶段的质量功能展开。根据下一道工序就是上一道工序的“顾客”的原理,各个开发阶段均可建立质量屋,且各阶段的质量屋内容有内在的联系。上一阶段质量屋天花板和地板的主要项目(分别为关键工程措施及其指标)将转换为下一阶段质量屋的左墙。质量屋的结构要素在各个阶段大体通用,但可根据具体情况适当剪裁和扩充。第一阶段(产品规划阶段)质量屋一般是最完整的,其他阶段的质量屋有可能将右墙、地下室等要素剪裁。对圆珠笔的开发而言,可以将“油墨成分”、“笔尖组件设计”作为下一阶段即零部件展开阶段的质量屋的左墙,进一步展开对零部件设计的分析,以便将顾客的要求深入地贯彻到产品的详细设计中去。在圆珠笔的工艺计划和生产计划阶段,也应类似地进行质量功能展开。

图10—9显示了四个阶段的质量功能展开。其中零部件展开阶段质量屋左墙的顾客需求应是产品规划阶段质量屋中关键的工程措施(设计要求)及其指标,天花板是为实现设计要求而提出的零(部)件特性。与此相仿,工艺计划阶段质量屋的左墙应为零件特性及其参数,天花板是工艺要求;生产阶段质量屋的左墙应为工艺要求及其参数,天花板是生产要求。

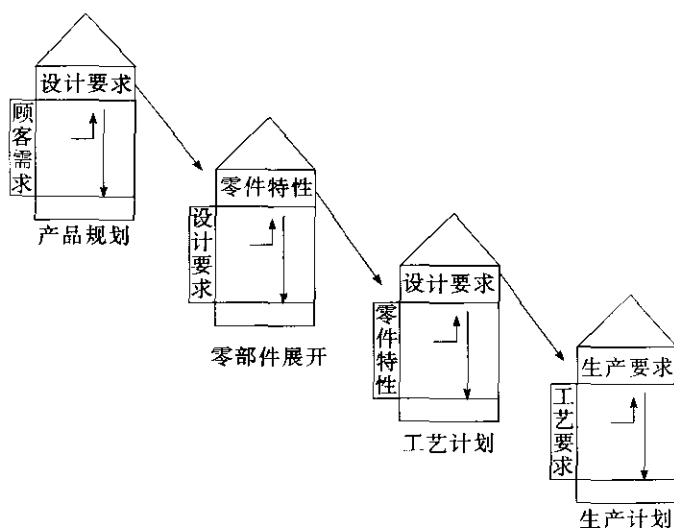


图 10—9 四个阶段的质量功能展开

并不是所有的质量功能展开都需要完整地包括上述四个阶段。根据 QFD 工作对象的复杂程度，可以按如下原则对四个阶段的质量功能展开进行剪裁或扩充：

(1) 每一阶段质量屋的工程措施应足够具体和详细，适于作为下一阶段质量屋的顾客要求（左墙）。例如，若产品规划质量屋中关键的工程措施不够具体和详细，可能需要在进行零部件展开前增加一层质量屋；反之，若产品规划阶段工程措施对于工艺计划阶段已足够详细，则可省略零部件展开阶段。

(2) 质量屋的规模不宜过大，即顾客需求和工程措施的数量不宜过多，以便于操作。一般顾客需求不应多于 20 项，工程措施不应多于 40 项。要特别指出，四个阶段的质量屋必须按照并行工程（concurrent engineering）的原理在产品方案论证阶段同步完成，以便同步地规划产品在整个开发过程中应该进行的所有工作，确保产品开发一次成功。质量功能展开是贯彻实施并行工程思想的有力工具。

7. 并行工程与 QFD 的结合运用

并行工程的发展与市场竞争的推动和信息技术的发展密切相关。一方面，由于竞争的激化，出现了经济全球化的趋势，有实力的企业纷纷提出了全球营销战略，要求在大范围和短时间内将产品投放市场并尽可能降低成本；另一方面，随着生产和装配向自动化方向发展，计算机技术的广泛应用，CAD/CAM 技术的深入发展，要求产品设计和工艺人员加强合作以改进产品的可生产性，保证产品的质量。为适应这一环境，需要对产品设计、工艺设计、制造等活动进行并行的分析和实施，研制全过程的信息数据应在整个企业内发布并由各个部门共享，从而推动了并行工程的发展。这意味着产品设计、工艺、生产和其他研制工作并行地开展，包括使有关的、有用的和所有潜在的信息在全公司各部门间流动，在方案论证阶段即并行地考虑安排各项有关工作，在产品的设计阶段充分考虑工艺、制造、运输、维修和售后服务的需要，以便最大限度地缩短产品开发周期，并保证一次成功。

由于 QFD 方法有效地支持了产品开发的策划工作，在组织结构上采用跨专业、多功

能综合小组的形式,它的实施为并行工程的开展提供了一种载体,成为直观、形象、功能强大的工具。四个阶段的质量屋是按照并行工程原理,在产品开发早期就同步完成的,规划了产品全寿命周期的全部工作,尽可能暴露各种矛盾并予以解决,这样就避免了返工和报废,缩短了产品的研制周期,降低了成本,提高了产品的质量,保证产品研制一次成功。

并行 QFD 的另一个表现是,在较复杂产品的开发中,在零部件展开阶段可能不止建立一个质量屋,而是相互平行的各专业或各子系统分别根据产品规划阶段的输出建立自己的零部件展开质量屋,并行地进行质量功能展开,从各自的角度对产品的设计要求进行全面系统的演绎分析。对于工艺计划及生产计划阶段也是如此。

10.2.3 QFD 的工作程序

本节以四阶段质量功能展开为背景,给出了 QFD 的工作程序,按本节的步骤实施质量功能展开,能很好地辅助中小规模的产品开发。

1. 顾客需求分析

顾客需求分析是六西格玛设计中识别(I)阶段的主要工作。

(1) 调查顾客需求。顾客需求分析是质量功能展开的关键环节,必须给予充分的重视。具体内容前已述及,此处不再重复。

(2) 顾客需求的整理。建立质量屋时,采用亲和图法等进行顾客需求的整理,提取前两级或前三级顾客需求即可。质量屋中顾客需求和工程措施的项数都不宜过多,否则影响工作效率。必要时,可增加质量屋的层次或划分为几个并列的质量屋加以展开;或删除一些影响轻微的顾客需求和工程措施;或将它们作一定的归并,以减小质量屋的规模。

按量化评估准则,结合顾客调查(含顾客抱怨调查)的结果,量化给出各项顾客需求的重要度。借鉴企业收集的顾客在使用同类产品时产生的抱怨,检查是否已把这些抱怨转化为相应的顾客需求,分析是否适当提高这类顾客需求的重要度,解决这些抱怨所对应的技术要求应纳入工程措施。

对于工程措施,也可以采用亲和图法进行整理。

(3) 市场竞争能力分析。在需求质量展开表的基础上,对新产品在市场上的定位进行策划。方法是通过进一步征询意见,调查研究,与竞争对手的产品进行水平比较(benchmarking),策划新产品对每一项顾客需求的满足程度,并求出原产品、新产品及竞争对手产品的市场竞争能力,进行竞争能力分析。这一过程也称为计划质量的设定过程,形成了质量屋的右墙。

表 10—2 是某无线电远距操纵装置的计划质量设定表。

首先,进行市场竞争能力比较分析,评定现有同类产品和竞争对手产品的竞争力。在可能的情况下,把这些产品摆在一起,客观地评估它们对各项顾客需求的满足程度,量化打分。

然后,对新产品的市场竞争能力进行定位。方法是采用同样的评分准则,设定新产品对各项顾客需求应达到的满足程度,并给出量化分值。

表 10—2 计划质量设定表

需求质量		重要度 K_i	市场竞争能力分析				水平 提高率 L_i	修正 系数 α_i	权 值	
第一级	第二级		本公司		X 公司	Y 公司			绝对 权值	相对 权值
			现有 产品	新 产品						
1. 操作方便	1.1 携带方便	3	3	5	4	4	1.67	1.5	7.5	8.4
	1.2 操作中不感觉疲劳	3	4	5	5	4	1.25		3.8	4.2
	1.3 对操作易于理解	3	4	5	5	3	1.25	1.2	4.5	5.1
	1.4 能轻松愉快地操作	3	3	4	3	3	1.33		4.0	4.5
	1.5 能处理复杂的问题	3	4	5	4	3	1.25	1.2	4.5	5.1
∴										
5. 安全	5.1	4	5	5	4	4	1.0	1.5	6.0	6.8
∴										
								合计	88.8	100

最后，利用公式，求出各产品的市场竞争能力分值并加以比较和确认，若得出的新产品市场竞争能力分值不满足公司的要求，则对新产品重新进行量化评分。

在工程实践中，可对本部分的内容加以扩充如下：为明确新产品相对于现有产品的水平提高程度，可计算新产品对应于每项顾客需求的水平提高率 L_i （见表 10—2）：

$$L_i = \text{新产品的市场竞争能力} / \text{现有产品的市场竞争能力}$$

为了取得竞争优势，可能需要在产品的设计中突出对某项顾客需求的满足程度，使产品具有鲜明的特色，成为产品的卖点（sales points），为此引入修正系数 α_i 。如果要重点突出对某项顾客需求的满足程度，可将该系数值定为 1.5；如果只是一般的突出，可将系数值定为 1.2。相应地，如果某项顾客需求被设置了修正系数，则其重要度应有所提高，为此，应对各项顾客需求的重要度进行修正，修正后重要度的绝对权值为顾客需求原重要度值 K_i 、水平提高率 L_i 和修正系数 α_i 三者的乘积。相对权值是每项绝对权值占绝对权值总和的百分比。在质量屋中可以用修正后的顾客需求的权值进行圆整后代替原有的重要度值，进行质量屋的分析计算。

2. 工程措施的确定与瓶颈技术的攻关

（1）工程措施的确定。对工程措施按以下要求进行分析和确认：

1) 工程措施应从产品整体着眼提出，而不是从现有产品的零件及工艺的技术要求中总结得出，以免限制产品的设计方案，影响创造力发挥。

2) 同一级的工程措施应相互独立。

3) 工程措施的组合应是全面系统的，可据此产生完整的设计方案。

4) 对于所选择的工程措施，应有助于提出量化的指标，以便对该项工程措施的实现方法和可实现程度进行科学评估。用 QFD 辅助大型、复杂产品的开发时，对其顶层质量屋的工程措施难以量化，此时工程措施及其指标的组合应能为后续的方案开发等工作指明

方向,使设计人员可据此判断设计工作是否偏离轨道。

5)在提出工程措施时,应调查同类产品在售后服务中发现的设计缺陷,检查是否有对应的工程措施进行改进,避免同类故障的发生。

确认后的工程措施可用亲和图法进行分级。

(2)质量屋要素的量化评估。按量化评估方法对各项顾客需求与对应的工程措施的相互关系进行打分,完成质量屋的房间部分——关系矩阵,计算得出各项工程措施的重要度。在关系矩阵完成后,应作如下检查:

1)若某项顾客需求与所有工程措施关系值都是0(空白),则应重新评估,或增加可满足该顾客需求的工程措施。

2)若某项工程措施与所有顾客需求的关系值都是0,应检查该工程措施是由哪一项顾客需求推导出来的,是否应取消。

3)若一项顾客需求与大多数工程措施有较强的关系,应分析量化分值的科学性。也可能该需求是高一级的顾客需求,此时应将其分解为几项子需求。

4)若一项工程措施与大多数顾客需求有较强的关系,应与第3)条类似的处理。

按10.2.2节的准则进行技术竞争能力评估,分析现产品及竞争对手的产品对各项工程措施的满足程度,结合售后服务调查结果,初步确定工程措施指标。由于工程措施是从技术的角度提出的,通常表示为各种具体的设计要求,因此针对某一项工程措施评估各产品达到的技术水平时,应考虑是否能找到技术上的评价标准,以提高量化评分的可信度。结合本企业的现有技术能力及技术发展策略设定新产品对每项工程措施的实现水平,用量化的分值表示,如将分值设为4分,表示新产品开发成功后,在该工程措施对应的技术领域应达到国内先进水平;设为5分,则应达到国际先进水平。因此,在设定新产品的技术竞争能力分值时,应考虑技术上的可行性,并对工程措施的指标进行相应修正,对需要进行的技术改造和技术攻关进行初步规划。完成量化评分后,计算这些产品的技术竞争能力和总体竞争能力,如得出的新产品的竞争力值不符合公司的产品发展战略,则重新设定新产品的竞争能力分值,确定相应的技术保证措施。

按10.2.2节的要求进行工程措施的相互关系评估,完成相关矩阵(尖屋顶)。

(3)工程措施指标的确定。由于工程措施指标的确定是质量功能展开的重要一环,直接指导后续技术工作的开展,在相关矩阵完成后,应参照以下原则对工程措施指标进行必要的评估和完善:

1)如不存在成本、资源、进度等限制,可参照业界领先水平或世界领先水平取值,以开发世界领先和国内领先的产品。

2)对于重要度高或对企业竞争优势作用重大的工程措施取值要高标准,必要时对此导致的成本和工作量的增加寻求管理层支持。

3)为彼此负相关或强负相关的工程措施确定技术指标时应进行权衡,因为它们对应的技术要求互相矛盾,不可能都按高标准取值。

4)如受到本企业技术条件限制,则工程措施的指标设定要实事求是,着眼于总体方案的优化。

5)对重要度不高的工程措施,应结合成本控制确定其取值。

应该指出,工程措施及指标的选择与产品技术方案的确是相互影响的。通过工程措

施的组合形成了产品的初步技术方案,应对此方案进行全面的评估与优化,并根据优化的结果对工程措施进行必要的调整。先进的系统设计方法就是辅助产品设计方案总体性、全局性优化和进行技术创新的有力工具;运用系统设计方法,还有助于将顾客需求科学地映射为功能要求、设计参数和工艺变量。

(4) 关键措施与瓶颈技术的确定。工程措施重要度确定以后,根据其重要度的大小可找到关键工程措施,作为控制重点。根据工程经验,另有一部分工程措施虽然重要度不高,但技术实现上难度较大,对此类工程措施也必须重点攻关。瓶颈技术在多个阶段的质量屋中都存在,可按下列准则加以分析确认:

- 1) 现有技术不能实现的关键工程措施;
- 2) 成本过高的零部件;
- 3) 质量和可靠性过不了关的零部件;
- 4) 新开发的工艺和原材料等。

找出关键工程措施和瓶颈技术后,应组织力量攻关,应用有关的工程技术,包括本章介绍的其他六西格玛设计方法,使瓶颈技术获得突破,使关键工程措施的设计方案实现稳定性优化,使产品开发中的技术障碍得以解决,以免延误研制周期并给产品的内在质量留下隐患。

3. 各级质量屋的建立

通常应与产品规划阶段质量屋一起,按并行工程原理协调、同步建立零部件展开、工艺计划、生产计划阶段质量屋。

产品规划阶段的 QFD 指导了产品的总体方案设计,输出了关键的设计要求或关键质量特性 CTQ 及指标,适用于六西格玛设计的界定 (D) 阶段。

在零部件展开阶段(对应于六西格玛设计的设计 (D) 和优化 (O) 阶段),以上一阶段输出的设计要求作为顾客需求,必要时从技术角度将这些设计要求未能覆盖的产品的功能与性能要求补充为顾客需求,运用头脑风暴法分析为满足这些顾客需求应具备的零部件特性,如技术参数、关键尺寸、材料等,形成产品的设计方案,并筛选重要的、对产品最终质量特性影响大的零部件特性,建立质量屋。在本阶段综合采用 FMEA, FTA 分析及参数设计等质量与可靠性工程方法对优化设计方案有很好的作用。本阶段的 QFD 用于指导产品详细设计和有关技术要求的制定,输出是关键零部件特性及指标;同时,应制定出质量保证要求,其主要内容是应保证的关键零部件特性指标(相当于关键过程特性 CTP),作为制定 QC 工序表 (quality control process chart) 的基础。

工艺计划质量屋是从设计到生产阶段的转移,其顾客需求是上一级质量屋输出的关键零部件特性,并可根据实际需要加以补充。从顾客需求入手,推导相应的工艺变量,形成工艺方案,小组成员还要根据这些工艺变量的组合建立工艺流程图。选择重要的工艺变量作为工程措施建立质量屋,进行量化评估。为保证生产过程的稳定,应进行工艺的稳健设计,使关键的零部件特性值保持在允许的波动范围内,还需要进行过程能力指数 C_p 及 C_{pk} 分析,运用 SPC 等质量工具控制过程波动。工艺计划阶段的 QFD 指导了工艺方案的编制,本阶段的输出是重要度高的工艺变量及指标。

在最后一个阶段,生产计划阶段,顾客需求是关键的工艺参数,工程措施则是一线技术人员的制造操作要求,主要是操作程序、操作工人的培训,检验、试验,计量保证,生

产设备的维护等，为防止已知故障的重复发生，还应设定预防性维护要求。本阶段 QFD 的目的是策划如何减少生产操作的成本，将生产的波动最小化，同时提高产量。一般而言，操作要求的制定应与 SPC 及质量管理体系密切协调，以便最大限度地发挥作用。根据这一级质量屋，制定 QC 工序表、操作说明书、预防性维护计划等文件，把顾客的需求落实到最底层。

工艺计划阶段和生产计划阶段的 QFD 技术应用于六西格玛设计的验证（V）阶段。应通过在线统计过程控制（SPC）动态地记录和监控产品是否稳定地达到设定的目标水平，若发现超出产品规格限的趋势，必须采取调整工艺参数的措施，以恢复生产线的质量水平。

表 10—3 给出了 QC 工序表的主要结构。表中的“控制项”可根据工艺计划质量屋得出，操作说明书可从生产计划质量屋得出，“控制表形式”包括控制图、检查单、图表、测试数据等。

表 10—3 QC 工序表框架

部门名称：				部门经理：						制表人：				
日期：														
零件名	工艺流程			工序名	操作说明书号	控制项	控制方法					检验项目	检验方法	备注
	原料处理工序	预处理工序	主工序				控制说明书编号	控制表形式	责任人	检验人	抽样频率及测试方法			

4. 落实关键环节的稳定性优化设计和强化控制

通过四个阶段的质量功能展开，确定了关键环节，为深入的产品开发指明了方向，对于瓶颈技术攻关及产品和工艺的稳健设计则有赖于其他质量工程工具和可靠性工具，因此，有必要将 QFD 与可靠性工程及其他质量工程工具结合使用，开发出优化的产品，根据情况可采用系统设计（如 TRIZ 等）、试验设计、参数设计与容差设计、故障模式与影响分析（FMEA）、故障树分析（FTA）、价值工程与价值分析（VE/VA）、数字仿真、以顾客为中心的可靠性设计分析等各种方法，以便实现稳定性的设计优化、工艺优化，充实和完善质量屋。对关键零部件和关键工艺，在生产制造阶段还要采用与稳健设计相适应的监控方法，如实施统计过程控制（SPC）、建立过程失控行动方案、完善质量管理体系等。

5. 质量屋的不断迭代与完善

第一轮四个阶段的质量屋大致在产品初步设计结束、技术状态冻结时完成。随着产品研制工作的深入，需要对各阶段的质量屋及时、不断地进行迭代、完善，直到产品进入市场时形成最终的四个阶段的质量屋，成为产品技术归档资料的一部分。质量屋的迭代与完善可结合设计评审、工艺评审和产品评审进行。

在新产品上市后，应继续应用 QFD 方法开展和优化售后服务，收集、研究顾客的意

见,应用 QFD 方法不断改进产品,提高产品质量,推出新的款式、型号,满足市场新的需求。

6. 计算机辅助质量功能展开

质量屋的绘制、填写、量化计算和修改等工作如果用手工完成,会给 QFD 小组带来一定的负担。另外,对企业而言,不同的 QFD 小组用不同的方式编制质量屋,也会给这些资料在企业内的保存、传递及再利用带来困难。采用计算机辅助质量功能展开软件,可以帮助工程人员在计算机上建立工程的质量屋模型,并对其进行一系列的量化评估、迭代分析及运算比较,最终产生一套完整的质量屋,同时生成详尽的可检索调用、可保存修改的工程信息记录,有效地支持了 QFD 的工程应用,并促进了 QFD 技术的规范化。从长远看,计算机辅助 QFD 软件产生的信息将融入企业内部的产品数据和管理信息流,在网络环境支持下,实现在各部门间的传递、共享和重用。我国自行开发的商品化计算机辅助质量功能展开软件已经具备上述所有功能。

10.2.4 应用于制造业之外的 QFD

QFD 除了在硬件产品开发和改进中广泛应用外,还渗透到了制造业之外的广泛领域。这是因为 QFD 本质上是一种用矩阵展开的方法来处理目的和手段二维关系的分析工具,在产品开发、批量生产、经营、管理、科学研究等活动中,都可以发挥它的作用。因此,质量功能展开方法将适用于各种领域和形式的工程和管理问题。

下面重点介绍 QFD 在服务业、软件开发和管理上的应用。

1. QFD 在服务业的应用

服务与产品的生产不同,是一个无形的不可逆过程。企业通过提供这一过程满足顾客的需要,获取顾客的认可,为企业带来效益。QFD 在服务业领域得以应用的背景是行业内竞争的激化。服务行业要在这样的条件下生存和发展,必须科学地规划自己的行为,不断提高服务质量。

由于服务行业自身的特殊性,QFD 的运用与在制造业中也有所不同。首先,进行广泛的顾客调查,掌握顾客对服务的要求,并利用质量屋将顾客需求转换为服务的要素。然后,以服务的要素(可称为质量要素)为质量屋的左墙,以业务功能为天花板加以展开,明确为了达到顾客的要求应该开展的业务及各项业务所提供的服务的具体内容,确定重要业务,为服务的改进和提高指明方向。为了把对业务的要求贯彻到服务活动中,还应考虑进行服务过程分析,找出提供服务所需的必要技术,确定重要技术,评估竞争范围内的技术水平现状,设定企业的技术水平目标,反映到业务手册中去。要根据顾客需求的变化,及时调整业务,并开发有特色的服务。

里兹-卡通酒店是一家跨国公司,1992 年曾获美国波多里奇国家质量奖。为了向顾客提供更优质的服务,其在密歇根州的下属公司针对“房间整理”这一课题运用 QFD 进行了研究,优化服务程序,以最大限度地满足房客的要求,取得成功经验后再加以标准化,向全公司推广。图 10-10 是它针对顾客需求作的质量屋。根据这一质量屋,明确了努力的目标,经过对目标的讨论,得出问题的关键在于安排好服务的流程。最后,确定每套客房由 3 个服务员同时整理,分工合作,其中一人兼管处理优惠卡服务事项。这一方案在降低失误率、减少对房客的打扰时间和打扰次数等方面获得了成功。

里兹的措施		顾客需求						
		重要性	失误率(随机抽样)	房间整理时间	可能的打扰次数	固定的打扰次数	优惠事项与折扣率处理	房客不在时一次打开的
顾客的要求	房间供应充足	4	◎				○	
	及时服务	5	◎		○	○	○	
	打扰尽可能少	3			◎	◎		
	打扰时间尽可能短	4		◎				△
	服务质量始终如一	5	◎				○	
	房客到达时房间已准备好	5		○				○
	在房客方便时清扫房屋	4		○		◎		○
	优惠事项准确及时兑现	5	△				◎	
	随身物品安全保险	5		△	◎			◎
	减少房间整理的成本与劳力	5		◎	△			◎
里兹的要求								
目标			() 次失误	() 分钟	每日 () 次	每日 () 次	每千房客比率	开 () 扇门

图 10—10 旅店业的质量功能展开

可通过 QFD 将顾客的要求进行排序, 与竞争对手比较, 以决定如何行事。优化服务流程, 以得到最大的竞争优势, 并避免将有限的财力、时间和人力用于顾客并不需要的地方, 或自己已经遥遥领先的地方。

在国防工业系统, 洛克威尔公司下属的柯林斯航电与通讯部, 主要负责向政府和军方提供产品。为了更好地满足顾客要求, 管理层曾确定了一项关键服务指标——顾客设备平均定期修理时间, 为更好地贯彻这一目标, 其服务中心组建了 QFD 小组, 进行综合调研, 以改进自己的服务。通过 QFD 分析, 他们发现, 该指标与顾客要求无关, 公司为此调整了业务, 节约了大量经费。

公司内部的辅助部门, 如财务部门、营销策划部门等, 也可视为服务业的一部分。要充分发挥它们的作用, 为公司创造效益, QFD 方法也大有用武之地。惠而浦公司将 QFD 方法用于大型会议的组织; ABB 集团将 QFD 用于为是否从事某项新事业提供帮助; 还有一家公司将 QFD 用于广告和公关策划活动, 节约成本 20%, 缩短准备时间 40%。

2. QFD在管理等领域的应用

QFD在管理领域的应用方式很灵活，其中一个通用模式是：从工作环境入手确定管理的目标和要求，作为顾客需求，逐级推导各管理层应承担的职责，把责任落实到人，并确定相应的检查和验证措施，确保任务完成。图10—11给出了一个用于策略规划的QFD的框架。

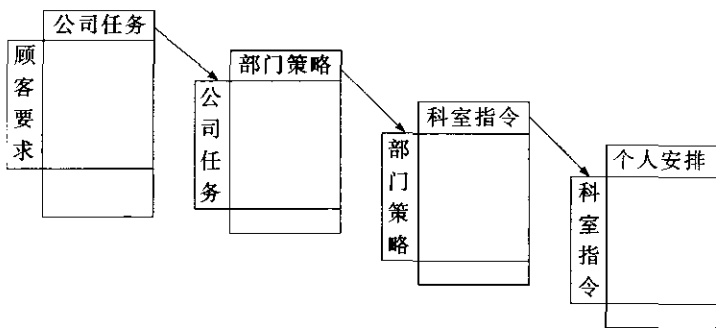


图 10—11 用于策略规划的 QFD 框架

10.3 系统设计

系统设计主要应用于 IDDOV 的界定需求 (D) 和研发 (D) 阶段。为了达到六西格玛质量水平，首先必须论证和确定先进可行的总体设计方案，从某种意义上说，产品的成败、产品能不能达到六西格玛水平，首先取决于总体设计方案。界定和研发阶段的主要任务急需采用的主要方法是：深入分析确定顾客需求（质量功能展开）、论证确定总体设计方案、攻克瓶颈技术、完成产品的结构和零部件设计并制定加工工艺等（系统设计包括 TRIZ, FMEA 等）。研发阶段是开展六西格玛设计效益最高和潜力最大的阶段。系统设计的内涵包括自顶向下的设计、公理性设计和解决创造性问题的理论（TRIZ）等。

限于篇幅，本节重点介绍 TRIZ 理论，对系统设计的其他内容只作简略介绍。

10.3.1 系统设计的基本原理与方法

1. 设计是一个映射过程

在产品开发初期，由于产品工作原理与结构尚待确定，其技术方案犹如一个黑箱，设计输入通过黑箱映射（mapping）为设计输出，设计过程是使黑箱透明化的过程。能够完成这一映射的技术方案是无穷的，系统设计的目的是找到一个优化的技术方案。公司通过对顾客或市场需求（包括潜在需求）的调查，整理出对产品的要求，形成需求规范，构成产品开发中最初的和最基本的设计输入。以此为基础，分析系统的工作环境，明确系统的基本输入和输出，确定系统的功能结构，把对产品的设计要求落实到功能要求上，实现从需求规范到功能要求的映射。然后，根据产品的功能要求确定产品的层次体系及零部件结构，完成产品的设计方案。最后，把产品的设计方案映射为对应的工艺方案（见图 10—12）。

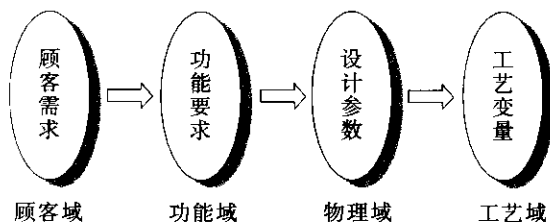


图 10—12 设计的映射过程

2. 自顶向下的设计

自顶向下设计的基本思想是深入分析确定顾客需求，从系统顶层开始，进行功能的分配、系统原理框架的展开、信息的传输与处理及综合、接口的控制、冗余度和系统重构设计、工程专业综合（engineering specialty integration）等，开展自上而下的设计和自下而上的综合。通过有机地管理各子系统、零部件的并行开发，协调技术、进度、成本的矛盾，保证产品的开发按规范的程序、规定的时间节点和成本完成，并保证产品的高质量和高可靠性。

3. 公理性设计

为优化映射过程，提出了相应的设计公理，即：

公理 1：独立公理（模块化公理）——保持功能要求的独立性。

好的设计应具有模块化的特点，特定的子系统（部件）实现特定的功能，互不干扰，最大限度地减少各功能要求之间的“耦合”。对于完全贯彻了独立公理的设计方案（无耦合设计），在为更好地实现某一功能而调整设计参数时，将不会影响其他功能的实现。这将使设计思路更加明晰，避免设计参数间复杂的交互作用带来的不确定性，有利于故障的诊断和产品的维修，也有利于产品功能的扩充和升级换代。

公理 2：信息公理（简化公理）——信息含量最小化。

4. 设计公理的推论

以上述两项设计公理为基础，推导出了七个推论，构成了系统化的设计原则，对设计实践提供指导。设计人员可以据此开展各阶段映射，并检查设计的合理性，在产品投入试制和生产前消除设计缺陷，完善设计。

10.3.2 解决创造性问题的理论

1. TRIZ 方法的基本思想

TRIZ 方法的提出源于以下认识：大量发明面临的基本问题和矛盾（在 TRIZ 中称之为系统冲突和物理矛盾）是相同的，只是技术领域不同而已。同样的技术发明原则和解决方案一次次地在多年后被重新使用。将这些有关的知识进行提炼和重新组织，就可以指导后来者的创新和开发。TRIZ 体系正是基于这一思路开发的，打破了我们思考问题的心理惰性和知识面的制约，避免了创新过程中的盲目性和局限性，指出了解决问题的方向和途径，并得到了计算机辅助软件的支持。

2. 理想的技术系统

TRIZ 理论认为，对技术系统而言，重要的不在于系统本身，而在于如何更科学地实

现功能。较好的技术系统应是在构造和使用维护中都消耗资源较少,而能完成同样功能的系统;理想技术系统则是不需要用材料来构造、不耗费能量和空间、不需维护也不会损坏的系统,即它在实体上不存在(隐含在已有的技术系统中),却能完成所需要的功能。这一思想与信息公理是切合的,是 TRIZ 所追求的理想目标。在现实中,理想技术系统的状态通常是不能实现的,但这一概念为创造性问题的解决指出了努力方向,在 TRIZ 理论中有着重要的地位。为了趋近于理想技术状态的目标,可在解决问题的初始阶段,设定要努力实现“理想最终结果”,这是后文所述“创造性问题解决步骤”的基础。在特殊的情况下,理想状态有时也能实现:假如能把应由某个系统完成的功能完全交由已经存在的子系统、高一级系统或同一层次的相邻系统去实现,就达到了理想系统状态。

例 10-1

为研究酸液与合金材料的相互作用,将合金加工成试样,放入密封的充满酸液的容器中,在各种温度与压力条件的组合下进行试验。在这种情况下,酸液不但与试样反应,还会腐蚀容器的内壁。为解决这一问题,在容器内壁敷设了玻璃层以防腐蚀,但却带来了新的问题:在某些试验中由玻璃组成的内壁很容易碎裂,需要频繁更换。

为解决这一难题,引入了“理想最终结果”的概念。首先,分析问题区域的所有部件及其功能。对该课题而言,酸液与试样的反应是主要功能,为考察它们的相互作用,两者都不能缺省,而容器的功能是把酸液维持在试样周围,提供必要的试验环境,属辅助功能。为解决容器内壁防腐与抗碎裂的矛盾,根据“理想最终结果”的思想,考虑能否取消现有容器的物理实体,而同样具有容器的功能。据此开发了创造性的解决方案:将待试验的合金试样加工成密封的容器,将酸液封闭在容器中进行试验(见图 10-13),既保证了试验条件,又免除了对容器内壁腐蚀的担心,因为本来就要考察酸液与合金的反应情况。

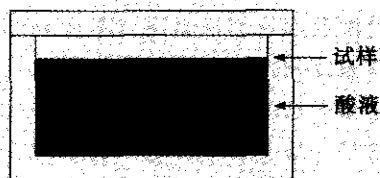


图 10-13 酸液对合金腐蚀试验的优化

3. 缩小的问题与扩大的问题

在解决问题的初期,面对需要克服的缺陷,可以有很多不同的思路,例如,改变系统、改变子系统或其中某一部件、改变高一层次的系统等,都可能使问题得到解决。思路不同,所思考的问题及对应的解决方案也会有所不同。

TRIZ 方法把所有的问题分为两类:缩小的问题(mini-problem)和扩大的问题(maxi-problem)。缩小的问题致力于使系统保持不变甚至简化,而消除系统的缺点,完成改进;扩大的问题则不对可选择的改变加以约束,因而可能为实现所需功能而增加零部件,甚至开发一个新的系统,使解决方案复杂化,甚至使解决问题所需的耗费与解决的效

果相比得不偿失。TRIZ 建议采用缩小的问题。这一思想是实现“理想技术系统”的途径。

驱动小轿车的四缸发动机通常有很强的二阶振动,在发动机低速运转的情况下(空挡状态),这种振动的频率较低,无法通过发动机底座加以隔离,而且某些小轿车还会产生结构的共振,影响了驾驶的舒适度,使部件的故障率及相应的保修费用提高。驾驶杆的共振就是其中的一个典型问题。

汽车公司对此进行了调查,发现由于驾驶杆的固有频率接近于发动机空挡时的二阶谐振频率,导致驾驶杆在空挡状态下剧烈振动,即使安装了减振器,其振动情况仍使驾驶员操作时不舒适。另外,驾驶杆的抖振也和发动机的负载有关,发动机还兼有驱动液压系统、为车上的用电附件(电机、空调等)供能等任务,这些负载越高,抖振现象越厉害。公司成立了攻关组,成员包括制造这些附件及传动设备的高级工程师,及车身与底盘的工程师。攻关组将这些普通小轿车与高档车作了比较,发现普通车电机与空调的效率比高档车低得多,驱动液压系统所需牵引扭矩在空挡状态比高档车大得多,车身硬度及车身与驾驶杆的固有频率则比高档车低得多,这些都是造成抖振的重要原因。攻关组回到各自相应的部门,有针对性地开发更好的系统,结构设计人员为此还造出一个加强的车身样品,并进行了测试,结果表明抖振得到了明显的改善。但这些办法工作量太大,花费太多,一时间无法投入应用。这是一个常规的解决问题的方式,也是典型的问题扩大化的情况。

TRIZ 专家参与攻关后,经过对问题背景知识的了解,提出两点建议:

(1) 如果试图不大幅改动系统而使问题得以解决,建议按缩小的问题处理,以便争取简化系统;

(2) 尽量使用已有的系统资源。

大家把注意力集中到不更改引起抖振的系统(车身和有关的附件等)而降低驾驶杆的振动程度,针对抖振本身来解决问题。在驾驶杆减振上,发现增加减振器惯性块的重量可改善减振能力,为了尽量使用已有的系统资源,大家对车内可用作惯性块的大块物体作了统计,前提是不影响它们的主要功能。这些物体有:散热器、电瓶、气囊、备用轮胎等。通过分析,把车前部防撞用的气囊兼作惯性块集成到驾驶杆的减振器上,顺利地解决了抖振问题。结果表明,采用该方法后驾驶杆的抖振比高档车还要小。

4. 系统冲突及其处理

系统冲突是 TRIZ 方法的一个核心概念,用于描述隐藏在问题背后的固有矛盾。如果要改进系统的某一部分属性,其他的某些属性就会恶化,就像天平一样,一端翘起,另一端必然下沉,这种问题就称作系统冲突。在飞机的结构设计中,结构的重量和强度是一对矛盾,通常,减轻结构的重量就必然削弱结构的强度;相反,要增强结构的强度则必须增加结构的重量。二者都是飞机结构的关键属性,构成了一对系统冲突。典型的系统冲突除了重量—强度冲突,还有形状—速度冲突、可靠性—复杂性冲突等。TRIZ 认为,发明可视为系统冲突的解决过程。

对于属性间的彼此冲突,传统的解决办法是折中,即寻找一个可接受的平衡。如果要求一个物体既要雪白又要漆黑,传统方法就把它设定为灰色。这样的解决方案并不完美,有时还无法接受,这就需要以创造性的思维,把冲突彻底消除。

TRIZ理论的创始人G. 奥特舒勒(G. Altshuller)对大量发明作了研究,发现虽然它们所属技术领域及所处理的问题千差万别,但隐含其中的系统冲突数量是有限的,典型的系统冲突只有1 250种。解决这些冲突所需的典型技术则更少,计40种。在此基础上产生了创造性问题的解决原则。

5. 物理矛盾及其处理

如果互相对立的属性集中于系统的同一元素上,就称为存在物理矛盾(有的文献称之为内部系统冲突)。物理矛盾的定义是:同一物体必须处于互相排斥的物理状态。也可以表述为:为实现功能 F_1 ,元素应有属性 P ;为实现功能 F_2 ,元素应有对立的属性 $-P$ 。

根据TRIZ理论,物理矛盾可以用三种方法解决:把对立属性在时间上加以分隔,把对立属性在空间上加以分隔和把对立属性所在的系统与部件分开。

在超音速飞机的发展过程中曾经必须解决气动力上的一个难题:为了优化飞机的亚音速性能,要求采用小后掠角、大展弦比的机翼;为了满足超音速的要求,则要求采用大后掠角、小展弦比的机翼。两种要求使机翼必须处于互相排斥的物理状态。这一矛盾可以用第一种方法(即时间分隔法)解决,即采用变后掠翼设计。通过在飞行的不同阶段(不同时间)相应调整后掠角,满足了飞机在亚音速与超音速状态下的不同气动要求。

采用第一种方法的思路的还有起落架的设计。在起降过程中要求飞机有起落架,支持飞机在地面的滑行过程;在飞行中则要求不要有起落架,以免增加飞行阻力。为此设计了可收放的起落架,在起降时伸出机体外,飞行时则收回起落架舱中。

为了使煎锅很好地加热食品,要求煎锅是热的良导体;而为了避免从火上取下煎锅时烫手,又要求煎锅是热的不良导体。为了解决这一矛盾,设计了带手柄的煎锅,把对导热的不同要求分隔在锅的不同空间。这是第二种方法(即空间分隔法)的体现。

某塑料管加工工艺中,使用旋转刀具把塑料管从要求的长度尺寸处切开。为了加快工艺流程,塑料管不断向前运动,刀具在切割过程中保持随动。由于塑料管和刀具之间难以精确同步,造成塑料管被切割部位出现毛刺。管子切割系统面临的物理矛盾是:必须运动,以加快工艺流程;必须静止,以保证切割精度。为了解决这一问题,采用了第二种方法的思路,将塑料管被切割部位暂时固定,保证切割精度,管的其余部分则继续前进,因被切割部位不能移动而弯成弧形,切割完成后解除固定。在实际应用中,为了引导管的移动,还安装了导稳滚轮,保证切割过程的稳定性。

第三种方法是指把矛盾的一方置于系统层面上,另一方则置于系统所属的部件或子系统上。如双体船,单一的船身不稳定,但通过两个船身的连接,系统有较好的稳定性。

10.3.3 TRIZ 方法的结构体系

TRIZ的结构体系可用图10—14表示。

作为TRIZ的理论基础,技术系统的演化规律概括了不同领域产品和工艺更新换代的一般规律,如增加理想度的规律、子系统演化不一致的规律、系统向更高层次发展的规律等。原则上,可以根据这些规律指导问题的解决,但对具体的问题而言,这些规律可能显

得太笼统,因此在 TRIZ 理论中,开发了更专业化的创造性问题分析工具和问题解决工具。20 世纪 70 年代后,以技术系统的演化规律为基础,开发了 TRIZ 预测技术,指导具体问题的解决,弥补了这一缺陷。

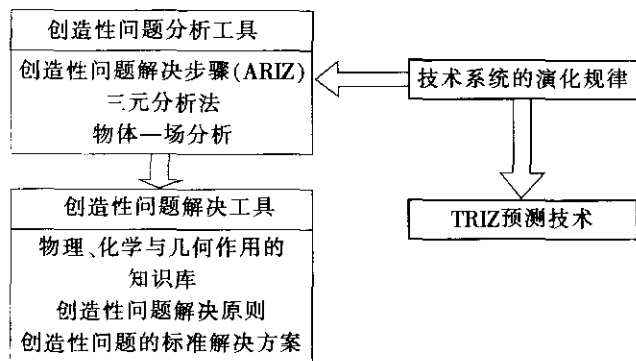


图 10—14 TRIZ 方法的结构体系

问题分析工具是一种建模技术,包括三元分析 (triads) 和物体—场分析 (substance-field analysis), 通过分析问题中的关键因素, 寻找系统内部矛盾, 使工程人员把握问题的根源。

创造性问题解决步骤 (algorithm for inventive problem solving, AIPS 或 ARIZ) 提供了解决问题的逻辑流程。由于矛盾和冲突往往隐藏在现实问题当中, 并不直接表现出来, 使人无法确认在何处及如何运用 TRIZ 工具, 因此需要遵循该流程来挖掘技术问题中的固有矛盾, 进行表述及解决。需要指出的是, 对较为复杂的问题而言, 正确地把握矛盾是极为重要的。虽然完整地运行 ARIZ 也可能实现问题的最终解决, 但考虑到该解决步骤主要进行的是问题的分析过程, 本章将其归入创造性问题的分析方法。

创造性问题的解决原则共 40 条, 与各种系统冲突模式分别对应, 直接指导创造者对新设计方案的开发。

创造性问题的标准解决方案是对特定问题经常使用的解决方案, 比解决原则更为具体, 它建立在物体—场分析及对不同领域技术问题通用解决方法的研究的基础上。

为了便于工程人员学习和应用 TRIZ 方法, TRIZ 研究组开发了计算机辅助软件, 其内核包括一个很大的知识库。该知识库建立在全球上百万件专利的收集、归纳和总结之上, 分门别类地记载了各类物理、化学、几何作用在工程上的应用。

以上方法从不同角度为创造性问题的解决提供了指南, 工程技术人员可根据需要从中选择适用的方法, 开发相应的解决方案。总的来说, 运用 TRIZ 方法解决问题的流程如图 10—15 所示。

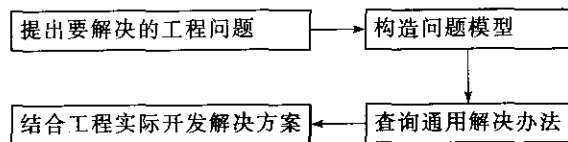


图 10—15 TRIZ 实施流程

10.3.4 创造性问题的分析方法

对于创造性问题的解决而言,技术上的辛勤攻关固然重要,对事物加以科学的分析,找出问题的突破口则更为关键。对每个试图解决科学和技术难题的人来说,首要的问题是必须决定向哪个方向努力。什么途径使问题的解决更快、更有效?如果没有找准突破口,可能大部分时间都会花在错误的方向和错误的方法上;如果找到了解决问题的正确方向,任何任务都可完成得更快,技术方案也更出色。TRIZ 通过对发明创造规律的总结,提出了三种独特的方法,帮助创造者分析问题的本质,找出问题的着眼点,从而利用创造性问题的解决办法,事半功倍地实现产品和工艺的改进和开发。

1. 创造性问题的解决步骤

为了争取达到技术系统的理想状态,用创造性的方法解决问题,就必须用科学的方式分析问题后面隐藏的根本矛盾,即 TRIZ 所定义的系统冲突和物理矛盾,从而根据该理论总结的创造性问题的解决原则启迪思维,开创突破性的解决办法。为此可采用创造性问题解决步骤 (ARIZ),该方法提出的流程如图 10—16 所示。

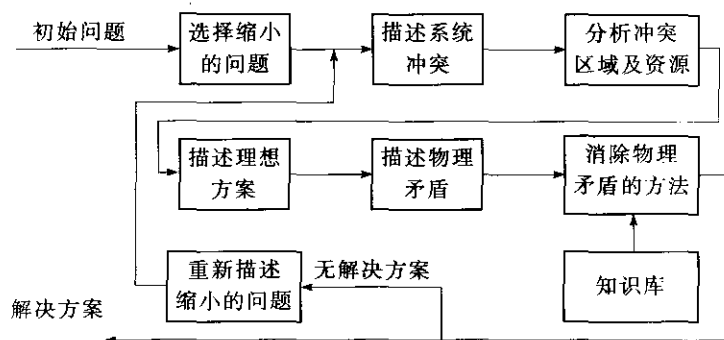


图 10—16 ARIZ 流程

按照 TRIZ 的基本观点,对同一问题的解决可能有多个不同的方案,方案的难易及可行性与对问题的描述方法息息相关。把实践中的矛盾描述为缩小的问题,将指引问题的解决者朝着理想最终结果的方向前进,从而找出既简便又有效的方法,以最小的代价使问题得到解决。创造性问题解决步骤就是为实现这一目的而开发的,它综合了 TRIZ 关于问题缩小化、理想产品及冲突与矛盾的有关观点,是一个连续的逻辑流程。考虑到现实问题通常有复杂的表象,创造者不一定在第一次分析时就能对问题作出正确的描述,该流程是一个循环结构。根据流程图,首先,对问题的初始描述一般比较模糊,创造者通过对问题的深入理解,将矛盾集中到较小的层面,描述一个缩小的问题。然后,以此为着眼点分析隐藏在系统中的冲突,找出冲突发生的区域,明确区域中有哪些固有资源,建立一个对应的理想方案。一般而言,可以在冲突区域发现互相矛盾的物理属性,即物理矛盾。为此,应分析系统面临的物理矛盾,找出矛盾所在的部件,作为问题解决的关键。最后,利用创造性问题的解决方法,在知识库(专家系统)的支持下开发具体的设计方案。在不同的时间、空间或不同的层次上分隔物理矛盾,是问题的解决途径之一。为了使解决方案尽可能接近理想方案,在具体方案的策划上,要尽量利用系统已有的资源,少增加额外的资源,

在对系统改动最小的情况下达目标。

如果对一个缩小的问题作了全面分析仍找不到解决方案,通常是因为对问题的初始描述或对缩小的问题的描述有误或不准确。因此,如果完整地进行了该流程而问题没有解决,建议回到分析的起点,进行更深入的调查研究,重新定义一个缩小的问题,再次按图 10—16 的流程寻找解决办法。

例 10—3

某纸厂用圆木作造纸原料。圆木卸在海边的传送带上,并运往切削机械进行加工。为使切削流程顺利完成,对圆木送到机器时的轴线方向作了规定。由于圆木卸下时杂乱地堆放在传送带上,所以需要在传送过程中增加一个圆木定向的工序,要求使圆木轴线方向与传送带轴线方向一致。这一操作如果由机器人完成,则结构复杂,占据大片场地,可靠性也不高。有没有简单、可靠、成本低廉的解决方案呢?对这一问题用 TRIZ 方法作了分析。

- (1) 缩小的问题:不对系统做主要改动,实现圆木定向。
- (2) 系统冲突:定向需要将圆木按要求方向加以排列的机构,但这使系统复杂化。
- (3) 问题模式:应利用系统中已有的要素实现定向功能。
- (4) 冲突区域及资源分析:冲突区域是传送带表面。该区域系统的唯一资源是传送带。
- (5) 理想最终结果:利用传送带自身实现圆木定向。
- (6) 物理矛盾:为实现圆木定向,传送带表面的不同点应有不同的速度;但为传送圆木,传送带表面应以同一速度运动。
- (7) 消除物理矛盾:将互相矛盾的要求分隔在不同的层面上,整个传送带以生产所需的速度向前运动,它的部件则以不同的速度运动。

工程方案:将传送带设计成三个部分。中间的主体部分以生产速度运动,把圆木送往切削机,两边的传送带则向相反的方向错动,通过摩擦力作用在圆木上,调整圆木的姿态,使其轴线方向与传送带轴向一致,达到定向的目的,见图 10—17。

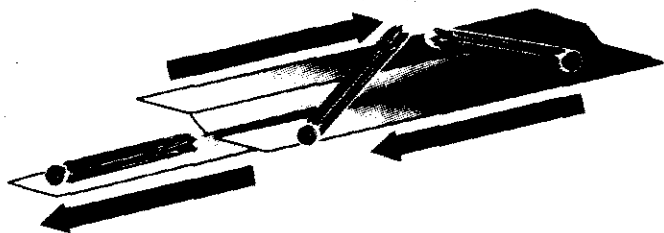


图 10—17 对圆木定向的传送带

2. 三元分析法

这是一种较新的分析方法,认为所有系统功能的实现有赖于三个物体的相互作用,即主动物体、被动物体和使动物体之间的相互作用。三元分析法就是对每一具体问题找出这

三个关键物体，分析它们彼此之间的关系，使工程人员明白现有系统这三者间好的与坏的作用，全面把握问题的实质，以便采取针对性的解决措施。

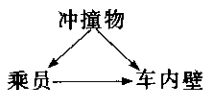
以汽车的气囊为例。气囊的功能是保护乘员免受突然冲撞的伤害，这一功能实施时，主动物体是空气袋，被动物体是乘员，那么使动物体是什么？考虑到使动物体在过程中的角色是使主动物体和被动物体互相作用，如没有这一物体，则作用不会发生，可知该过程中使动物体是与汽车发生撞击的冲撞物，如路旁的栏杆或另一辆车等。三者是功能得以实现的关键物体，构成了一个三角：



下面分析这一三角中的彼此关系。冲撞物使乘员产生与车座相脱离的运动，危及乘员的安全；同时冲撞物使气囊膨胀，气囊则对撞向车内壁的乘员起保护作用。现有的气囊对于乘员在大部分情况下有良好的保护作用，但在某些条件下也会对乘员带来严重伤害，例如，如果车祸发生时，乘员正好在座位前，离气囊的位置特别近，就会被气囊击伤；又如，对小孩或矮个乘员，气囊正好撞击在头部，伤害可能是致命的。现有气囊系统面临的问题是：乘员没有得到彻底有效的保护。

通过三元分析，对问题的现状有了较为深刻的了解，如何利用这一三角来开启问题的解决思路呢？由于三个物体间有着内在的密切联系，理论上只要改进了任两者之间的相互作用，都可以使问题得到缓解或解决，这就为解决问题提供了广阔的思路。例如，冲撞物使乘员产生与车座脱离的运动是有害的，因为乘员如果因此撞上车内的内壁，就会造成伤害，那么有什么办法增加乘员与车内壁的距离，使乘员脱离车座到达车内壁的运动时间长一些，多给乘员一些作出反应的机会？又如，鉴于气囊在某些情况下对乘员有伤害作用，那么如何改进气囊系统，尽最大可能减小伤害？对每一个问题的解答都是一种或多种设计方案的选择与确定过程。

当然，针对乘员在冲撞中的保护问题，也可以从另一个角度进行三元分析。建立如下的三角：



这一三角中，乘员是主动物体，车内壁是被动物体，冲撞物是使动物体。相互作用为：冲撞物使乘员产生脱离车座的运动，使车内壁减速，乘员撞击车内壁。面临的问题变为：如何减小或消除乘员与车内壁的强烈碰撞？能否使乘员作脱离运动的时间长一些？如何使车内壁的减速过程足够平缓？为此进行的分析将开启全新的思路。

可见，只要正确地选定了主动物体、被动物体和使动物体，运用三元分析法可以为问题的解决开辟广阔的空间。

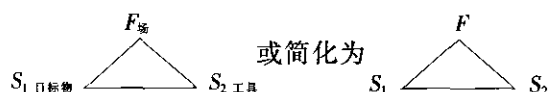
3. 物体一场分析

根据 TRIZ 理论，尽管技术系统实现的功能千变万化，但都可表示为物质和能量的转换过程。有鉴于此，TRIZ 专家开发了物体一场分析这一建模技术，用于描述技术系统中物体和能量的相互作用，对技术系统的运作原理进行深入分析。

技术系统功能的实现,表现为特定的目标物体发生了特定的作用,上述圆木定向系统功能的实现,即表现为圆木改变了传送状态。功能的实现方式有两种:一种是将能量直接施加在目标物上,使作用发生,称为直接法。以汽车为例,为了使车辆行驶,必须驱使车轮转动,直接法就是把发动机产生的能量直接传递到车轮上,使之转动。这一方法有内在的缺陷,即目标物的作用过程难以控制。为了解决这一问题,可以采用另一种方式,即引入新的物体,通过它与目标物的相互作用来实现控制,称之为间接法。新引入的物体可称为工具。对汽车而言,工具就是控制车轮转向与加减速的机构。兼有能量、目标物和工具的技术系统才是完整和稳定的,三者之间构成了一个三角形。



TRIZ 把这三要素分别称作物体和场。“目标物”和“工具”属于“物体”,构成所分析的技术系统的实物部分。“物体”可以是一个单独的实体,如船锚,也可以是一个由很多零部件组合而成的系统,如船。“场”则是指物体之间实现相互作用或控制所需的能量。在 TRIZ 中,上述的三角模式被描述为:



这一结构称为系统的物体一场结构。这一系统构架可以描述为最小的技术系统的模式。对任一技术系统,都可以建立其物体一场结构,区别在于,部分系统的物体一场结构是不完整的,而 TRIZ 理论根据对大量技术系统的分析,认为包含了这三要素的系统,才是一个完整而功能稳定的系统。三要素中任一部分的缺失,都会给系统带来内在的缺陷。

该结构中的“物体”主要有 13 类,即固体、液体、气体、结构系统、松散类、表面属性类、机械属性类、热属性类、化学属性类、电属性类、磁属性类、光学属性类、微分子类物体。

“场”有 8 大类,即机械能、声场、热力场、化学能、电能、磁场、电磁辐射、核能。机械能包含 11 个子类:机械力、重力(与地球相关)、惯性力、摩擦力、液压、气压、固体的动能、松散物体流、液体流、气体流、机械振动。需要指出,不同的场有不同的作用模式和效果,在明确了技术系统中发挥作用的场以后,改变场的作用模式也是改进技术系统的思考方向之一。

考虑到问题的发生可以归结为特定的物体不能产生特定的作用,或产生作用的同时,附带了有害的作用,而问题可望通过对物体更科学合理的控制来解决,我们便可以在问题解决的开始阶段,建立系统的物体一场结构,分析这一三角形的完整性及各部分之间的相互作用,探求这一结构的不足之处,并有针对性地加以改变,从而使问题得到解决。这就是物体一场分析的基本原理。应用物体一场结构可对技术系统进行本质性的描述,因而可望通过物体一场分析揭示问题产生的深层的根本原因,理清事物内部复杂的相互关系,使创新者把握问题的本质,为创造性地解决问题开拓思路。这是一种通用的分析方法,在 TRIZ 体系中占有重要的地位。物体一场分析完成后,可用标准的解决

方法来解决问题。

物体一场分析的步骤是：

- (1) 确认目标物 S_1 ；
- (2) 确认工具 S_2 （如存在）；
- (3) 确认场（如存在）；
- (4) 描述系统当前的物体一场模型；
- (5) 建立系统期望的物体一场模型；
- (6) 开发解决方案。

对物体一场结构的改善而言，基本的规律是：结构中至少一个部分的缺乏表明结构需要完善。

10.3.5 创造性问题的解决方法

为了创造性地解决问题，TRIZ 开发了颇具特色的解决方法，提出了系统冲突的解决原则和问题的标准解决方案，还建立了对解决问题有着重要支持作用的知识库。

1. 处理创造性问题的 40 个原则

前面已经介绍了系统冲突的概念。创造性问题的解决原则就是特定的系统冲突所对应的问题解决技术。典型的问题解决原则共 40 种，这些原则还可细分出子原则，约 100 项。下面列出了其中的一部分：

原则 1：分割的原则。

- (1) 将系统划分为多个彼此独立的部件；
- (2) 使系统可分解；
- (3) 增加系统被分割的程度。

原则 17：转移到新的维度。

(1) 将物体在一维空间的问题转移到二维空间解决，将物体在二维空间的问题转移到三维空间解决；

- (2) 将层次单一的系统转变为多层次的系统；
- (3) 将物体转动、倾斜；
- (4) 反映到物体的相邻区域或反面上。

原则 28：向非机械系统转变。

- (1) 将机械系统转变为光学、声学或嗅觉系统；
- (2) 在机械系统中运用电场、磁场或电磁场；
- (3) 加入铁磁粒子，运用磁场或电磁场。

原则 32：改变颜色。

- (1) 改变物体或其环境的颜色；
- (2) 改变物体或其环境的透明度；
- (3) 使用带色添加剂，以观察难以看清的物体或过程；
- (4) 若此类添加剂已经使用，则采用发光轨迹或示踪元素。

格雷戈里·叶泽尔斯基（Gregory Yezersky）是系统化研究组织的一员，他介绍了自己运用其中的一条原则（原则 5）“将相同或相似的物体结合或合并”的案例。

表 10—4

1993 年, 一家著名航空公司为了改进火箭气动外形, 引入了“蜂腰”结构, 但发现由此带来了负面影响: 这种改变将引起强烈的振动, 使安装在“蜂腰”结构内壁的传感测量设备无法工作。受已有设计的限制, 对问题的解决不允许在系统中进行大量更改。公司无法解决这一问题, 请叶泽尔斯基应用 TRIZ 方法加以分析和解决。

在美国专利中有一项“给薄玻璃板打圆角”的专利, 内容是, 给一块薄玻璃板打圆角容易把玻璃震碎, 可以把几块薄玻璃板临时黏合在一起, 共同打圆角, 就避免了问题的发生。从这一类专利中可以抽象出一条原则: 如果有有害作用加于系统, 可将相似系统彼此结合, 以加强系统的薄弱环节。

叶泽尔斯基巧妙地运用了这一原则, 提出用带子把该处的设备捆在一起, 结果避免了振动带来的危害, 使问题得以解决。

为了直观地反映对应于每一对典型系统冲突的问题解决原则, 构造了矛盾矩阵表, 表 10—4 是其局部示意图。

这张表为具体问题的解决提供了重要的向导。表的上部是在问题中需改善的系统属性, 左侧则是改善了这些属性后会恶化的系统属性。对具体问题进行分析, 明确要改善的是表中第 j 项属性, 且由此将导致表中第 i 项属性恶化, 则第 i 项属性所在的行与第 j 项属性所在的列相交的格子中的数字就是对应于这一对系统冲突的问题解决原则的编号。

表 10—4

矛盾矩阵表

改善的属性 (j)		1	2	3	4	5	...	22	...	30	...	39
		运动物 体重量	静止物 体重量	运动物 体长度	静止物 体长度	运动物 体面积		能量 损失		影响物 体的有 害因素		生产 能力
恶化的属性(i)	1			15,8 29,34		29,17 38,34		6,12 34,19		22,21 18,27		35,3 24,37
	2				10,1 29,35			18,19 28,15		2,19 22,37		1,28 15,35
	3	8,15 29,34				15,17 4		7,2 35,39		1,15 17,24		14,4 28,29
	4		35,28 40,29					6,28				30,14 7,26
	5	2,17 29,4		14,15 18,4				15,17 30,26				10,26 34,2
	⋮											
	33	25,2 15,13	6,13 1,25	1,17 13,12		1,17 13,16		2 19,13				15,1 28
	⋮											
	39	35,26 24,37	28,27 15,3	18,4 28,38	30,7 14,26	10,26 34,31		28,10 29,35		22,35 13,24		

例 10-5

当前玻璃杯的制造工艺是将熔融状态的玻璃注入相应的模具中成形。外观带棱角的玻璃杯如果用整体模成形,模具的制造会很困难;如果模具精度不好,会影响杯子的外形。在矛盾矩阵表的指引下,怎样解决这一难题呢?

首先,分析该课题要改善的属性,本案例要改善的是玻璃杯的“形状”,是矩阵表中的第12项属性。其次,分析会在改进后恶化的属性,本案例中,恶化的属性为玻璃杯(模具)的“可制造性”,是表中的第32项属性。然后,在表中查找TRIZ建议的解决原则,在该表这两项属性所在的列和行的相交处,格子里的数字为:1,32,17,28。即可选用原则1、原则32、原则17、原则28来解决问题。

根据原则1,可将模具分成几块,增加可制造性,再组合起来供玻璃杯成形,以保证杯子形状的精度。大多数厂家采用的就是这种办法。

原则32似与本问题的解决无关,暂不考虑。

原则17的第四条为:反映到物体的相邻区域或反面上。这使技术人员换个角度思考,能否把杯子的棱角由在杯子外壁改为设计在杯子内壁,而外壁仍保持圆滑?如果这么做同样可满足顾客要求,就会带来制造上的很大方便:模具遇到的上述精度问题不复存在,只要加工一个外壁的棱角与杯子的内壁棱角一致的芯模,用来在杯子成形时保证杯子的内形即可。

原则28的第二条建议使用电场、磁场或电磁场,据此可考虑把模具从水平方向分割开,分为杯底和杯壁两部分,用电蚀法取代机加法加工棱角。

2. 创造性问题的标准解决方案

物体一场分析揭示了系统冲突中本质性的问题。在得出系统的物体一场结构以后,可以利用问题的标准解决方案进行物体一场结构的综合与转换,以便解决或至少避开系统冲突。

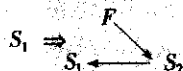
奥特舒勒归纳了76项标准,下面结合其中的两项介绍用标准解决方案解决问题的过程。

标准之一:如果对系统的要求难以通过对现有物体的更改来满足,对系统引入新的物体和场也没有限制,则问题可望通过引入物体一场结构中缺少的部分,合成一个完整的结构而得以解决。

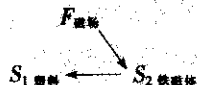
例 10-6

制造带刷毛的塑料块,传统工艺是使用定制的模式。模式是一个附有一套针形凸起的金属块,针的尺寸和样式随刷毛的尺寸和样式而定。生产时,模式浸入熔化的塑料中再拉起,带动附在针上的塑料从塑料液中拉出,形成刷毛的形状,刷毛长度达到要求后,用气冷法冷却塑料,再从针的末端把塑料切下。这一工艺的缺点是会有部分塑料黏在针上,在刷毛组合不同时需要频繁清洗。

对此进行物体一场分析。将熔化的塑料定为 S_1 (目标物),发现传统工艺的物体一场结构中缺少工具 S_2 与场 F ,为此引入 S_2 和 F 来控制 S_1 。这一变化反映在物体一场结构上:

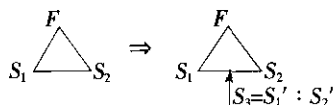


刷毛成形的过程是部分熔化的塑料受机械力而拉伸的过程。考虑到该工艺中如将 F 定为机械力则不好控制,故选用了磁场作用, S_2 则是将磁力转化为机械力的铁磁体。对应的物体—场结构为:



具体方案为:根据对刷毛的要求在熔化的塑料中加入磁粉,将原金属模替换为已磁化的铁磁体,置于塑料块带刷毛区的上方,磁粉在磁场作用下连同附带的塑料一起向上伸,形成刷毛。这一方法可以适应刷毛尺寸和样式的迅速改变。

标准之二:如果物体一场结构中两物体间既有好的作用又有坏的作用,而没有必要保持该两物体的直接联系,且不希望或不允许引入新的物体,则将两物体加以修正,组合成第三个物体,使问题得以解决。反映在物体一场结构上的转换为:



例 10—7

钢丸发送机的弯管部分是强烈磨损区,在弯管部分添加保护层的效果很有限。

对发送机进行物体—场分析,钢丸为目标物 S_1 ,管子为工具 S_2 , F 为机械力。经分析发现,管子和钢丸间既有好的作用(管子为钢丸导向),又有坏的作用(钢丸冲击和磨损弯管部分)。采用本解决标准处理这一问题,修正物 S_3 可以是钢丸、管子或这两者。

解决方案:经过分析,选取 $S_3=S_1$,即用钢丸本身兼作保护层。实施办法为,在弯管外壁放置磁体,将飞行中的钢丸吸附在弯管内壁,形成保护层(见图 10—18)。

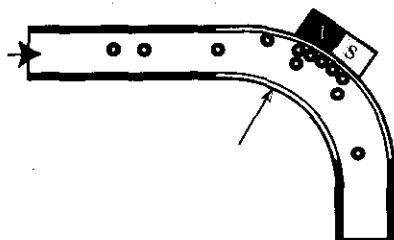


图 10—18 用钢丸自身保护弯管部分

10.4 静态参数设计

静态参数设计是六西格玛设计最主要的技术工具之一,在 QFD 和系统设计的基础上,针对 CTQ 和 CTP 深入开展静态参数设计,寻找工作参数(或过程参数)的优化匹配,以减少质量特性(或过程特性)的波动,提高产品(或过程)抵御各种干扰的能力(即健壮性)。静态参数设计也可应用于六西格玛 DMAIC 的改进(I)阶段,其主要内容在 7.6 节已作了简要介绍,本节不再重复。关于静态参数设计更深入的内容,可见章后相关文献。

10.5 容差设计

10.5.1 质量损失函数

按照田口提出的理论,要评价产品设计质量的好坏,只要计算损失函数就够了。本章主要介绍运用质量损失函数确定产品容差的方法。

1. 望目特性的损失函数

设产品的质量特性 y 为望目特性, m 为目标值。称下述函数为望目特性的损失函数:

$$L(y) = k(y - m)^2 \quad (10-5)$$

式中, k 为与 y 无关的常数,它可以按下述方法确定:

$$k = A_0 / \Delta_0^2 = A / \Delta^2 \quad (10-6)$$

式中, Δ_0 为产品的功能界限, A_0 为产品丧失功能时的损失; Δ 为产品的容差, A 为产品不合格的损失。

若已知质量特性 y 的 n 个观测值: y_1, y_2, \dots, y_n 时,则产品的平均质量损失为:

$$L(y) = k \frac{1}{n} \sum_{i=1}^n (y_i - m)^2 = kV_T \quad (10-7)$$

$$\text{式中} \quad V_T = \frac{1}{n} \sum_{i=1}^n (y_i - m)^2 \quad (10-8)$$

2. 望小特性的损失函数

设产品的质量特性 y 为望小特性,则 y 不取负值,且理想值为零。此时,只要在式(10-5)至式(10-8)中,令 $m=0$ 便可得到望小特性的损失函数和平均质量损失。

$$L(y) = ky^2 \quad (10-9)$$

$$k = A_0 / \Delta_0^2 = A / \Delta^2 \quad (10-10)$$

$$\bar{L}(y) = k \frac{1}{n} \sum_{i=1}^n y_i^2 = kV_T \quad (10-11)$$

$$\text{式中} \quad V_T = \frac{1}{n} \sum_{i=1}^n y_i^2 \quad (10-12)$$

3. 望大特性的损失函数

设产品的质量特性 y 为望大特性,则 y 不取负值,且理想值为无限大。此时,只要将

$1/y$ 视为望小特性, 便可得到望大特性的损失函数和平均质量损失。

$$L(y) = k/y^2 \quad (10-13)$$

$$k = A_0 \Delta_0^2 = A \Delta^2 \quad (10-14)$$

$$\bar{L}(y) = k \frac{1}{n} \sum_{i=1}^n \frac{1}{y_i^2} = k V_T \quad (10-15)$$

$$\text{式中 } V_T = \frac{1}{n} \sum_{i=1}^n \frac{1}{y_i^2} \quad (10-16)$$

10.5.2 容差的确定方法

1. 由安全系数确定容差

本节引入的安全系数的定义是田口博士提出的, 与通常工程技术中的安全系数不是完全相同的概念, 应注意两者的区别。

(1) 安全系数定义。设 A_0 为达到功能界限时的平均损失, 主要是用户的损失; A 为不合格品时的工厂损失。则安全系数 Φ 为:

$$\Phi = \sqrt{\frac{A_0}{A}} \quad (10-17)$$

由于 $A_0 > A$, 因此安全系数 Φ 大于 1。安全系数 Φ 越大, 说明丧失功能时的损失也越大。因此, 对于要求有很高安全性的产品, 相应的安全系数 Φ 比较大。一般采用安全系数 $\Phi = 4 \sim 5$ 。

(2) 望目、望小特性情形。望目和望小特性的容差 Δ 和丧失功能的界限 Δ_0 之间具有如下关系式:

$$\Delta = \sqrt{\frac{A}{A_0}} \Delta_0 \quad (10-18)$$

由式 (10-17) 和式 (10-18) 可知, 当已知功能界限 Δ_0 和安全系数 Φ 时, 则容差 Δ 为:

$$\Delta = \frac{\Delta_0}{\Phi} \quad (10-19)$$

例 10-8

设某电视机电源电路的直流输出电压的目标值 $m = 120\text{V}$, 功能界限 $\Delta_0 = 25\% \times m$, 丧失功能后用户的平均损失 (包括修理费及修理后电视机仍不能使用的损失等) $A_0 = 500$ 元, 工厂内超出规格的产品, 可采用改变电路中电阻值进行调整 (包括电阻费用和人工费用等), $A = 2$ 元, 试求安全系数 Φ 和容差 Δ 。

解:

$$\Phi = \sqrt{\frac{A_0}{A}} = \sqrt{\frac{500}{2}} = 15.8$$

$$\text{容差 } \Delta = \frac{\Delta_0}{\Phi} = \frac{0.25 \times 120}{15.8} = 1.9(\text{V})$$

$$\text{合格品范围 } m \pm \Delta = 120 \pm 1.9 = (118.1\text{V}, 121.9\text{V})$$

(3) 望大特性情形。望大特性的容差 Δ 和丧失功能的界限 Δ_0 之间具有如下关系式:

$$\Delta = \sqrt{\frac{A_0}{A}} \Delta_0 \quad (10-20)$$

由式 (10-17) 和式 (10-20), 当已知功能界限 Δ_0 和安全系数 Φ 时, 则容差 Δ 为:

$$\Delta = \Phi \Delta_0 \quad (10-21)$$

例 10-2

用硬聚氯乙烯型材加工塑料门窗。当材料的拉伸强度低于 31MPa 时, 门窗就会断裂, 此时造成的损失 $A_0 = 500$ 元; 而因材料不合格, 工厂作报废处理的损失 $A = 120$ 元, 试求硬聚氯乙烯型材的安全系数 Φ 与容差 Δ 。

解: 质量特性拉伸强度为望大特性, 已知 $\Delta_0 = 31\text{MPa}$, $A_0 = 500$ 元, $A = 120$ 元, 因此

$$\Phi = \sqrt{\frac{A_0}{A}} = \sqrt{\frac{500}{120}} = 2.04$$

$$\Delta = \Phi \Delta_0 = 2.04 \times 31 = 63.24 \text{ (MPa)}$$

因此所用型材的强度下限为 63MPa。

2. 下位特性容差的确定

现在让我们来看一看如何根据上位特性 (结果特性) 的容差或功能界限, 确定下位特性的容差。

设产品的上位特性为 y , 下位特性为 x , 考虑最简单的情况, 设当下位特性 x 变化单位量时, 相应上位特性 y 的变化量为 b , 即 y 与 x 之间存在如下线性关系:

$$y = a + bx \quad (10-22)$$

记: Δ_y 为上位特性的容差;

A_y 为上位特性的不合格损失;

Δ_{0y} 为上位特性的功能界限;

A_{0y} 为上位特性丧失功能的损失;

Δ_x 为下位特性的容差;

A_x 为下位特性的不合格损失;

m_y 为上位特性的目标值;

m_x 为下位特性的目标值;

b 为 x 每变化一个单位时, 相应 y 的变化量;

a 为 $x=0$ 时, 相应 y 的值。

下位特性 x 的容差 Δ_x 有以下两个计算公式:

$$\Delta_x = \sqrt{\frac{A_x}{A_y}} \times \frac{\Delta_y}{|b|} \quad (10-23)$$

$$\Delta_x = \sqrt{\frac{A_x}{A_{0y}}} \times \frac{\Delta_{0y}}{|b|} \quad (10-24)$$

收录机稳压电源电路的质量特性为输出电压 y ，公差为 $m_0 \pm 1.5V$ 。电路中某个电阻的中心值为 m ，阻值变化对输出电压的影响为 $b=0.5V/k\Omega$ 。当输出电压超出公差，导致产品不合格的损失 $A_y=20$ 元。在组装前发现电阻不合格时，工厂损失 $A_x=0.5$ 元，求该电阻的公差 Δ_x 。

解：已知 $\Delta_y=1.5V$ ， $A_y=20$ 元， $A_x=0.5$ 元， $b=0.5V/k\Omega$ ，由公式 (10—23)，电阻的公差 Δ_x 为：

$$\Delta_x = \sqrt{\frac{A_x}{A_y}} \times \frac{\Delta_y}{|b|} = \sqrt{\frac{0.5}{20}} \times \frac{1.5}{0.5} = 0.474(k\Omega)$$

因此，电阻的公差为 $0.474k\Omega$ 。

3. 老化系数公差确定

(1) 老化特性和老化系数。随时间推延而向同一倾向发生变化的特性称为老化特性（或劣化特性）。例如，电阻的阻值随着时间的延长而逐渐变大；机械零件的磨损量随时间延长越来越大；化学电池的电压随储存时间的增加渐渐降低等均为老化特性。

另外，像钟表的快慢，如在一个月內要快多少或慢多少；炮弹引信中火药延时机构的延时特性随温度高低而变化等质量特性，虽然不是老化特性，但也可以用下面提供的方法，类似老化特性来处理。

由于特性随时间推延而老化，老化量在不同截止时间是不同的，所以首先必须确定设计寿命。

为了简便起见，只讨论老化特性随时间呈线性变化的情况。设产品老化特性 y 的设计寿命为 T ，老化系数为 β ， α 为初始值， y 随时间 t 呈线性变化，即

$$y = \alpha - \beta t, \quad 0 < t < T \quad (10-25)$$

老化系数 β 的公差 Δ^* 的含义是：在整个寿命期间，若 $\beta > \Delta^*$ ，表示为不合格品，此时相应的损失为 A^* 元。

本节主要讨论当老化特性 y 的功能界限 Δ_0 、丧失功能的损失 A_0 和设计寿命 T 已知时，如何确定老化系数 β 的公差。下面分初始值等于目标值和不等目标值两种情形进行讨论。

(2) 初始值等于目标值的情形。设老化特性 y 的目标值为 m ，当 $t=0$ 时， $y=m$ ，则公式 (10—25) 简化为：

$$y = m - \beta t, \quad 0 < t < T \quad (10-26)$$

其图形如图 10—19 所示。

老化系数 β 的公差 Δ^* 的计算公式如下：

$$\Delta^* = \sqrt{\frac{3A^*}{A_0}} \times \frac{\Delta_0}{T} \quad (10-27)$$

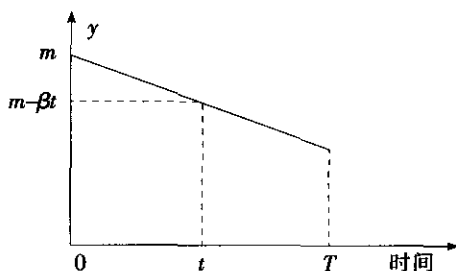


图 10—19 当 $y_0=m$ 时的老化特性

例 10—11

某机械零件尺寸 y 的设计初始值为目标值 m ，设计寿命 $T=10$ 年，当磨损量达到 $300\mu\text{m}$ 时就不能正常使用，即功能界限 $\Delta_0=300\mu\text{m}$ ，此时的损失 $A_0=80$ 元。若每年平均磨损量 β 不合格，产品降级使用的损失 A^* 为 $=5$ 元，求 β 的容差 Δ^* 。

解：

$$\Delta^* = \sqrt{\frac{3A^*}{A_0}} \times \frac{\Delta_0}{T} = \sqrt{\frac{3 \times 5}{80}} \times \frac{300}{10} = 13.0 (\mu\text{m}/\text{年})$$

因此，每年平均磨损 β 的容差为 $13\mu\text{m}/\text{年}$ 。

(3) 初始值不等于目标值的情形。设老化特性 y 的初始值为 $y_0 \neq m$ 。不失一般性，设 $y_0 > m$ ，且当 $t=T/2$ 时， $y=m$ 。则式 (10—25) 简化为：

$$y = m + \beta T/2 - \beta t, \quad 0 < t < T \quad (10-28)$$

其图形如图 10—20 所示。

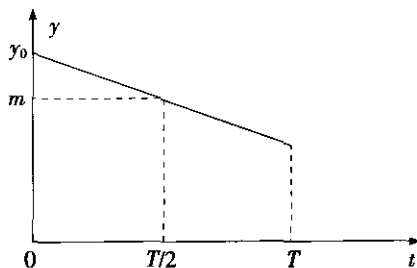


图 10—20 当 $y_0 > m$ 时的老化特性

初始值不等于目标值的老化系数 β 的容差 Δ^* 的计算公式如下：

$$\Delta^* = \sqrt{\frac{12A^*}{A_0}} \times \frac{\Delta_0}{T} \quad (10-29)$$

比较式 (10—27) 和式 (10—29) 可知，初始值大于目标值情形下老化系数 β 的容差 Δ^* 较初始值等于目标值时要大。故初始值选取得比目标值大为好。

例 10—12

在例 10—11 中, 设初始值 y_0 不等于目标值, 当 $t=5$ 年时, 零件尺寸 y 等于目标值 m , 其他条件不变, 求每年平均磨损 β 的容差 Δ 。

解:

$$\Delta = \sqrt{\frac{12A^*}{A_0}} \times \frac{\Delta_0}{T} = 2\sqrt{\frac{3A^*}{A_0}} \times \frac{\Delta_0}{T} = 2\Delta^* = 2 \times 13 = 26(\mu\text{m}/\text{年})$$

因此, 每年平均磨损 β 的容差为 $26\mu\text{m}/\text{年}$ 。

4. 下位特性的老化系数容差的确定

设产品的上位特性为 y , 下位特性为 x , y 与 x 之间具有线性关系。此处介绍下位特性 x 的老化系数 β 的容差的确定方法。具体符号如下:

Δ_y 为上位特性的容差;

A_y 为上位特性的不合格损失;

m_y 为上位特性的目标值;

Δ_x 为下位特性的容差;

A_x 为下位特性的不合格损失;

m_x 为下位特性的目标值;

b 为 x 每变化一个单位时, 相应 y 的变化量;

a 为 $x=0$ 时, 相应 y 的值;

T 为下位特性的设计寿命;

β 为下位特性的老化系数;

Δ^* 为下位特性的老化系数的容差;

A^* 为 β 不合格时的损失。

上位特性 y 与下位特性 x 之间具有下列关系式:

$$y = a + bx \quad (10-30)$$

由式 (10—23) 可知, 下位特性 x 的初期容差 Δ_x 可由下式给出:

$$\Delta_x = \sqrt{\frac{A_x}{A_y}} \times \frac{\Delta_y}{|b|} \quad (10-31)$$

下位特性 x 的老化系数 β 的容差的计算公式为:

$$\Delta^* = \sqrt{\frac{3A^*}{A_y}} \times \frac{\Delta_y}{|b|T} \quad (10-32)$$

例 10—13

在车间照明中, 当灯泡的照度变化 $\Delta_y = 50\text{lx}$ 时, 因质量方面发生故障的修理费用 $A_y = 150$ 元。而当灯泡的发光强度变化 1cd (坎德拉) 相应照度要变化 0.8lx , 即 $b = 0.8\text{lx}/\text{cd}$ 。灯泡的初期发光强度超出规格时可进行调整, 调整费用 $A_x = 3$ 元。对于老化情形, 当超出规格时则报废处理, 损失 $A^* = 32$ 元。设计寿命为 $T = 20\,000\text{h}$, 试求灯泡初期发光强度的容差 Δ 及发光强度老化系数的容差 Δ^* 。

解：初期发光强度的容差为：

$$\Delta = \sqrt{\frac{A_x}{A_y}} \times \frac{\Delta_y}{|b|} = \sqrt{\frac{3}{150}} \times \frac{50}{0.8} = 8.8(\text{cd})$$

发光强度老化系数的容差为：

$$\Delta^* = \sqrt{\frac{3A^*}{A_y}} \times \frac{\Delta_y}{|b|T} = \sqrt{\frac{3 \times 32}{150}} \times \frac{50}{0.8 \times 20\,000} = 0.002\,5(\text{cd/h})$$

10.5.3 望目、望小特性的容差设计

容差设计的主要工具是质量损失函数。运用质量损失函数计算产品的质量损失，并按照“使社会总损失（即质量损失与成本之和）最小”的原则来确定合适的容差。

例 10—14

设计某机械产品，其材料可从 A_1 、 A_2 、 A_3 三种材料中任选。三种材料的线性膨胀系数 b_i （温度每变化 1°C 时 A_i 的延伸量）、每年磨损量 β 和价格 P_i 如表 10—5 所示。设产品功能界限 $\Delta_0 = 6\mu\text{m}$ ，丧失功能时的损失 $A_0 = 702$ 元，设计寿命 $T = 20$ 年，温度标准差 $\sigma_x = 15^\circ\text{C}$ ，试问选择哪种材料最合适？

表 10—5

例 10—14 数据表

i	材料	b_i ($\mu\text{m}/^\circ\text{C}$)	β ($\mu\text{m}/\text{年}$)	价格 P_i (元)
1	A_1	0.08	0.15	7.02
2	A_2	0.03	0.06	13.65
3	A_3	0.01	0.05	24.57

解：设产品在标准温度下出厂时的尺寸等于目标值，因此三种材料由于温度变化和老化引起的质量特性波动的均方偏差 V_{Ti} 为：

$$V_{Ti} = b_i^2 \sigma_x^2 + \frac{1}{3} T^2 \beta_i^2, \quad i = 1, 2, 3 \quad (10-33)$$

而三种材料的质量水平 L_i 为：

$$L_i = \frac{A_0}{\Delta_0^2} V_{Ti} = 19.5 V_{Ti} \quad (10-34)$$

材料 A_1 的 V_{T1} 和 L_1 计算如下：

$$\begin{aligned} V_{T1} &= b_1^2 \sigma_x^2 + \frac{1}{3} T^2 \beta_1^2 \\ &= 0.08^2 \times 15^2 + \frac{20^2}{3} \times 0.15^2 \\ &= 4.44(\mu\text{m}) \end{aligned}$$

$$L_1 = 19.5 V_{T1} = 19.5 \times 4.44 = 86.58 (\text{元})$$

可以类似计算出 V_{T2} 、 V_{T3} 、 L_2 、 L_3 等，具体结果见表 10—6。

表 10—6

例 10—14 最终计算结果

i	材料	$b_i (\mu\text{m}/^\circ\text{C})$	$\beta (\mu\text{m}/\text{年})$	价格 P_i (元)	$V_{Ti} (\mu\text{m})$	质量水平 L_i (元)	总损失 (元)	L_i/P_i
1	A_1	0.08	0.15	7.02	4.44	86.58	93.60	12.3
2	A_2	0.03	0.06	13.65	0.68	13.31	26.96	1.0
3	A_3	0.01	0.05	24.57	0.36	6.94	31.51	0.3

表 10—6 中, 总损失为:

$$L_T = P_i + L_i, \quad i=1, 2, 3 \quad (10-35)$$

从表 10—6 中可看出, 材料 A_2 的总损失 L_{T2} 最小, 是最佳材料。而且采用 A_2 时, 质量损失和成本之间达到平衡, 两者之比接近 1。

通过容差设计, 得到产品的成本, 即不合格品损失 A 为:

$$A = A_2 = 13.65 \text{ (元)}$$

因此, 最佳条件下的安全系数为:

$$\Phi = \sqrt{\frac{A_0}{A}} = \sqrt{\frac{702}{13.65}} = 7.2$$

10.5.4 望大特性的容差设计

某系统的设计中需选用一种树脂管, 希望其强度越大越好。设树脂管的强度和价格均与管子的截面积成正比, 单位截面积强度 $b=80\text{N}/\text{mm}^2$, 单位截面积价格 $a=40 \text{ 元}/\text{mm}^2$ 。又当应力 $\Delta_0=5000\text{N}$ 时, 树脂管会断裂, 此时的损失 $A_0=30 \text{ 万元}$ 。试设计这种管子的最佳截面积。

解: 设截面积为 x , 其价格 $P=ax$, 其强度 $y=bx$, 相应损失函数为:

$$L = L(y) = \frac{A_0 \Delta_0^2}{y^2} = \frac{A_0 \Delta_0^2}{(bx)^2} \quad (10-36)$$

因此, 总损失 L_T 为:

$$\begin{aligned} L_T &= P + L \\ &= ax + \frac{A_0 \Delta_0^2}{(bx)^2} \end{aligned} \quad (10-37)$$

因 x 的最佳解是使 L_T 达最小值, 故有 $\frac{dL_T}{dx} = 0$, 即

$$a = \frac{2A_0 \Delta_0^2}{b^2 x^3} \quad (10-38)$$

可得管子的最佳截面积为:

$$x = \left(\frac{2A_0 \Delta_0^2}{ab^2} \right)^{\frac{1}{3}} \quad (10-39)$$

将 $a=40$ 元/ mm^2 , $b=80\text{N}/\text{mm}^2$, $A_0=30$ 万元, $\Delta_0=5\,000\text{N}$ 代入上式, 得到

$$x = \left(\frac{2 \times 300\,000 \times 5\,000^2}{40 \times 80^2} \right)^{\frac{1}{3}} = 388 (\text{mm}^2)$$

因此, 成本、质量损失和总损失分别为:

成本 $P=ax=40 \times 388=15\,520$ (元)

质量损失 $L=L(y)=\frac{300\,000 \times 5\,000^2}{(80 \times 388)^2}=7\,784$ (元)

总损失 $L_T=P+L=15\,520+7\,784=23\,304$ (元)

这样就得到最佳截面积 $x=388\text{mm}^2$, 它是使生产成本与质量损失之和达到最小的截面积。

最佳条件下, 成本 P 等于不合格品的损失 A , 因此安全系数为:

$$\Phi = \sqrt{\frac{A_0}{A}} = \sqrt{\frac{300\,000}{15\,520}} = 4.4$$

从而可得到强度 y 的容差 Δ , 有

$$\Delta = \Phi \Delta_0 = 4.4 \times 5\,000 = 22\,000 \text{ (N)}$$

10.5.5 贡献率法

1. 单因素情况

例 10—16

设乳剂黏度 x (单位: P) 与涂布厚度 y (单位: μm) 之间存在线性关系。涂布厚度的目标值 $m=24\mu\text{m}$, 容差 $\Delta=10\mu\text{m}$, 当涂布厚度不合格时, 该涂布为不合格品, 此时损失为每单位面积 $A=16$ 元。现引进某种自控装置, 对黏度进行自动控制, 将使黏度的容差 Δ_r 减小一半。该自控装置的价格为 50 万元, 每年经费 (利息、偿还、运转费等) 占其 40%, 涂布每年产量为 6 万个单位面积, 问引进黏度自控装置是否可行? 为了进行贡献率分析, 现每隔 2 小时, 把黏度为 x 的乳剂涂布一次, 共涂 15 次, 测得涂布厚度 y 的数据如表 10—7 所示。

表 10—7

例 10—16 数据表

I	$x(p)$	$y(\mu\text{m})$	i	$x(p)$	$y(\mu\text{m})$	i	$x(p)$	$y(\mu\text{m})$
1	6.5	20	6	8.8	25	11	9.2	34
2	10.6	43	7	9.2	39	12	8.9	26
3	9.3	39	8	7.0	17	13	9.0	30
4	8.9	23	9	7.9	22	14	8.3	23
5	7.6	19	10	7.5	21	15	6.8	15

解:

(1) 回归直线的确定。对表 10—7 中的数据进行计算, 得

$$\begin{cases}
 \bar{x} = \frac{1}{15} \sum_{i=1}^{15} x_i = 8.37 \\
 \bar{y} = \frac{1}{15} \sum_{i=1}^{15} y_i = 26.4 \\
 L_{xx} = \sum_{i=1}^{15} x_i^2 - \frac{1}{15} \left(\sum_{i=1}^{15} x_i \right)^2 = 17.77 \\
 L_{xy} = \sum_{i=1}^{15} x_i y_i - \frac{1}{15} \left(\sum_{i=1}^{15} x_i \right) \left(\sum_{i=1}^{15} y_i \right) = 120.6 \\
 b = \frac{L_{xy}}{L_{xx}} = \frac{120.6}{17.77} = 6.79 \\
 a = \bar{y} - b\bar{x} = 26.4 - 6.79 \times 8.37 = -30.43
 \end{cases} \quad (10-40)$$

因此, 所求回归直线为:

$$y = a + bx = -30.43 + 6.79x \quad (10-41)$$

(2) 试验数据的统计分析 (见表 10-8)。

表 10-8

方差分析表

来源	偏差平方和 S	自由度 f	方差 V	纯波动 S'	贡献率 ρ(%)
m	S _m =86.4	1	V _m =86.4	S' _m =68.53	ρ _m =6.0
x	S _x =819.3	1	V _x =819.3	S' _x =801.43	ρ _x =70.4
e	S _e =232.3	13	V _e =17.87	S' _e =268.04	ρ _e =23.6
T	S' _T =1 138	15	V _T =75.87	S' _T =1 138	100

(3) 系统偏差的校正。由于 $m = 24\mu\text{m}$, $\bar{y} = 26.4\mu\text{m}$, 故存在系统偏差 $\bar{y} - m = 26.4 - 24 = 2.4(\mu\text{m})$ 。

由式 (10-41) 可得

$$\bar{x} = \frac{m - a}{b} = \frac{24 + 30.43}{6.79} = 8.02(\text{P})$$

上式说明, 只要将黏度 x 的均值从原来的 $\bar{x} = 8.37\text{P}$ 调整到 $\bar{x}' = 8.02\text{P}$, 则可望使涂布平均厚度 $\bar{y}' = m = 24\mu\text{m}$, 从而使贡献率 $\rho'_m = 0$ 。

(4) 损失函数与平均质量损失函数的确定。由条件 $m = 24\mu\text{m}$, $\Delta = 10\mu\text{m}$, $A = 16$ 元, 则损失函数 $L(y)$ 和平均质量损失函数 \bar{L} 分别为:

$$L(y) = \frac{A}{\Delta^2} (y - m)^2 = 0.16 (y - 24)^2$$

$$\bar{L} = \frac{A}{\Delta^2} \left[\frac{1}{n} \sum_{i=1}^n (y_i - m)^2 \right] = \frac{A}{\Delta^2} V_T = 0.16 \times 75.87 = 12.14(\text{元})$$

(5) 容差设计。引进自控装置后 $\Delta'_x = \Delta_x / 2$, 又 $\rho_m = 0.06$, $\rho_x = 0.704$, 于是新的质量水平 L_N 为:

$$\begin{aligned}
 L_N &= \bar{L} \left\{ 1 - \rho_m - \rho_x \left[1 - \left(\frac{\Delta'_x}{\Delta_x} \right)^2 \right] \right\} \\
 &= 12.14 \times \{ 1 - 0.06 - 0.704 \times 0.75 \} = 5.00(\text{元})
 \end{aligned}$$

可得到容差设计表如下(见表10-9)。

表 10-9

例 10-16 计算结果

方案	成本 P (元)	平均质量损失 L (元)	总损失 $L_T = P + L$ (元)
原方案	0	$L = 12.14$	$L_T = 12.14$
自控装置	$P' = 3.33$	$L_N = 5.00$	$L'_T = 8.33$

说明引进自控装置后,每件产品可减少损失 3.81 元。以年产 6 万件计,年增益 22.86 万元,因此是可行方案。

2. 多因素情况

要设计一个电感电路,此电路由电阻 R (单位: Ω) 和电感 L (单位: H) 组成,线路图如图 10-21 所示。

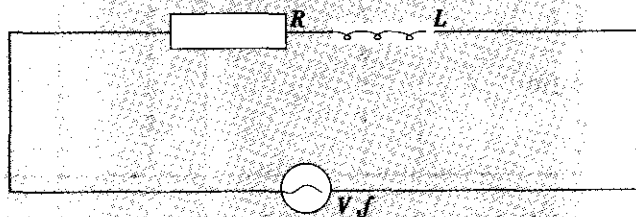


图 10-21 电感电路图

当输入交流电压为 V (单位: V) 和电源频率为 f (单位: Hz), 输出电流的强度为:

$$y = \frac{V}{\sqrt{R^2 + (2\pi fL)^2}} \quad (10-42)$$

质量特性 y 的单位为安培 A。

设质量特性 y 的目标值 $m = 10\text{A}$, 用户对输出电流强度的容许范围为 $m \pm 4\text{A}$, 即 $\Delta_0 = 4\text{A}$; 超出此范围的售后服务等损失为 $A_0 = 160$ 元。

设以最佳参数 $R = 9.5\Omega$, $L = 0.01\text{H}$ 进行容差设计, 原方案采用三级品的电阻和电感, 问是否存在更好的容差设计方案?

(1) 试验方案的设计。本例有 4 个误差因素, 即电源电压 V , 电源频率 f , 电阻 R 和电感 L 。可利用正交表来设计试验方案。

误差因素都取三水平, 误差因素 V 和 f 根据外界客观条件确定:

$$V_1 = 90\text{V}, V_2 = 100\text{V}, V_3 = 110\text{V}$$

$$f_1 = 50\text{Hz}, f_2 = 55\text{Hz}, f_3 = 60\text{Hz}$$

电阻 R 和电感 L 以最佳参数为第二水平, 即 $R_2 = 9.5\Omega$, $L_2 = 0.01\text{H}$, 都采用三级品, 波动范围为 $\pm 10\%$, 则可得到误差因素水平, 见表 10-10:

表 10—10

例 10—17 数据表

因素 水平	V(V)	f(Hz)	R(Ω)	L(H)
1	90	50	8.55	0.009
2	100	55	9.50	0.010
3	110	60	10.45	0.011

选用正交表 $L_9(3^4)$ ，将误差因素 V, f, R, L 顺序上列，得到试验方案，见表 10—11。

表 10—11

例 10—17 正交表

因素 水平	V(V)	f(Hz)	R(Ω)	L(H)
1	1(90)	1(50)	1(8.55)	1(0.009)
2	1	2(55)	2(9.50)	2(0.010)
3	1	3(60)	3(10.45)	3(0.011)
4	2(100)	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3(110)	1	3	2
8	3	2	2	3
9	3	3	1	1

再由式 (10—42) 计算表 10—11 中 1~9 号条件下的质量特性值 y ，其结果见表 10—12。

表 10—12

例 10—17 试验结果及分析

因素 水平	V(V)	f(Hz)	R(Ω)	L(H)	y_i	$y'_i = y_i - 10$
1	1(90)	1(50)	1(8.55)	1(0.009)	9.99	-0.01
2	1	2(55)	2(9.50)	2(0.010)	8.90	-1.10
3	1	3(60)	3(10.45)	3(0.011)	8.01	-1.99
4	2(100)	1	2	3	9.89	-0.11
5	2	2	3	1	9.17	-0.83
6	2	3	1	2	10.70	0.70
7	3(110)	1	3	2	10.08	0.08
8	3	2	2	3	11.76	1.76
9	3	3	1	1	10.90	0.9
T_{1j}	26.90	29.96	32.45	30.06	$T=89.4$	
T_{2j}	29.76	29.83	29.69	29.68	$T'=-0.6$	
T_{3j}	32.74	29.61	27.26	29.66		

注意, 第 j 列的部分和 T_{1j}, T_{2j}, T_{3j} 是对 y_j 求和, 而不是对 y'_j 求和。

(2) 试验数据的统计分析。由于因素的水平是等间隔的, 所以可利用正交多项式回归理论把因素引起的波动平方和分解为一次项、二次项引起的波动平方和, 并求出相应的贡献率。具体步骤如下:

1) 偏差平方和。

$$S'_T = \sum_{i=1}^n (y_i - m)^2 = \sum_{i=1}^9 (y'_i)^2 = (-0.01)^2 + (-1.10)^2 + \cdots + (0.90)^2 \\ = 10.28 (A^2)$$

2) 均值偏差平方和。

$$S_m = n(\bar{y} - 10)^2 = \frac{1}{n} \left(\sum_{i=1}^n y'_i \right)^2 = \frac{(-0.6)^2}{9} = 0.04 (A^2)$$

3) 各因素一次项、二次项引起的波动平方和。本例采用正交表 $L_9(3^4)$, 总试验次数 $n=9$, 水平数 $k=3$, 水平重复数 $\gamma=3$ 。从附表正交多项式表中可查出:

$$W_{11} = -1, W_{21} = 0, W_{31} = 1, \lambda_1^2 S_1 = 2$$

$$W_{12} = 1, W_{22} = -2, W_{32} = 1, \lambda_2^2 S_2 = 6$$

因此, 可算出:

$$S_{v1} = \frac{(W_{11}T_{11} + W_{21}T_{21} + W_{31}T_{31})^2}{\lambda_1^2 S_1} = \frac{(-1 \times 26.90 + 32.74)^2}{3 \times 2} = 5.68 (A^2)$$

$$S_{vq} = \frac{(W_{12}T_{11} + W_{22}T_{21} + W_{32}T_{31})^2}{\lambda_2^2 S_2} = \frac{(26.9 - 2 \times 29.76 + 32.74)^2}{3 \times 6} \\ = 8 \times 10^{-4} (A^2)$$

可类似计算 $S_{f1}, S_{R1}, S_{Rq}, S_{L1}, S_{Lq}$, 将这些结果整理成下面的统计分析表 (见表 10—13)。

本例有 4 个误差因素, 选用 $L_9(3^4)$ 时没有空列作为误差列。为了估计试验误差的方差, 把方差较小的项, 即 $V < 0.01$ 有 “ Δ ” 记号的项合并成误差项, 并得到误差平方和 S_e 。

表 10—13

例 10—17 计算结果

来源	S	f	V	S'	$\rho(\%)$
m	0.04	1	0.04	0.036 4	0.35
V_e	5.68	1	5.68	5.676 4	55.22
ΔV_q	$\Delta 0.000\ 08$	$\Delta 1$	—	—	—
f_1	0.02	1	0.02	0.016 4	0.16
Δf_q	$\Delta 0.000\ 045$	$\Delta 1$	—	—	—
R_1	4.49	1	4.49	4.486 4	43.64
ΔR_q	$\Delta 0.006\ 05$	$\Delta 1$	—	—	—
L_1	0.03	1	0.03	0.026 4	0.26
ΔL_q	$\Delta 0.007\ 2$	$\Delta 1$	—	—	—
e	0	0	—	—	—
(e)	-0.014 5	4	-0.003 6	-0.031 1	-0.32
T	10.28	9	1.14	10.28	100

(3) 正交多项式回归方程的确定。从表 10—13 中可以看出, V_e 和 R_1 的贡献率较大, 均为显著项。根据正交多项式回归理论, 质量特性 y 与显著项之间的回归方程为:

$$\hat{y} = \bar{y} + b_{1v}(V - \bar{V}) + b_{1R}(R - \bar{R}) \quad (10-43)$$

式中

$$\bar{y} = \frac{1}{9} \sum_{i=1}^9 y_i = \frac{89.4}{9} = 9.93(\text{A})$$

$$\bar{V} = 100, \bar{R} = 9.5(\Omega)$$

回归系数

$$b_{1v} = \frac{W_{11}T_{11} + W_{21}T_{21} + W_{31}T_{31}}{\lambda_1 S_1 h_v}$$

$$b_{1R} = \frac{W_{11}T_{13} + W_{21}T_{23} + W_{31}T_{33}}{\lambda_1 S_1 h_R}$$

从附表 13 “正交多项式系数表”中可查出 $\lambda_1 S_1 = 2$, 而水平间隔 $h_v = 10\text{V}$, $h_R = 0.95\Omega$, 因此有

$$b_{1v} = \frac{-1 \times 26.90 + 1 \times 32.74}{3 \times 2 \times 10} = 0.097$$

$$b_{1R} = \frac{-1 \times 32.45 + 1 \times 27.26}{3 \times 2 \times 0.95} = -0.910$$

所以回归方程为:

$$\hat{y} = 9.93 + 0.097(V - 100) - 0.910(R - 9.5)$$

当 $V=100\text{V}$, $R=9.5\Omega$ 时, 质量特性 y 的估计 $\hat{y} = 9.93\text{A}$, 它与目标值 $m=10\text{A}$ 相差不多, 可以不进行调整, 说明在参数设计阶段已对系统偏差进行校正。

(4) 损失函数与质量水平的确定。由已知条件 $\Delta_0=4\text{A}$, $A_0=160$ 元, $m=10\text{A}$, 可得到电流强度 y 的损失函数为:

$$L(y) = \frac{A}{\Delta_0^2} (y - m)^2 = \frac{160}{4^2} (y - 10)^2 = 10 (y - 10)^2$$

n 件产品 y_1, y_2, \dots, y_n 的质量水平为:

$$\bar{L} = \frac{A_0}{\Delta_0^2} V_T = 10V_T = 10 \times 1.14 = 11.4(\text{元})$$

说明在原方案中, 当 R 和 f 都采用三级品时, 每件产品的平均质量损失为 11.4 元。

(5) 容差设计。本例中只有 $\rho_{v1}=0.5522$, $\rho_{R1}=0.4363$ 较大, 且电压为环境因素, 不易压缩容差, 所以只考虑对电阻 R 的容差进行压缩。

现有两个容差设计方案: 电阻 R 采用二级品, 容差 $\Delta'_R = \Delta_R/2$, 增加成本 $P'=3$ 元; 电阻 R 采用一级品, 容差 $\Delta''_R = \Delta_R/10$, 增加成本 $P''=5$ 元。

对第一方案, 其质量水平为:

$$\begin{aligned} L'_N &= \bar{L} \left\{ 1 - \rho_m - \rho_R \left[1 - \left(\frac{\Delta'_R}{\Delta_R} \right)^2 \right] \right\} = 11.4 \times \{ 1 - 0.0035 - 0.4364 \times 0.75 \} \\ &= 7.6(\text{元}) \end{aligned}$$

对第二方案, 其质量水平为:

$$L_N'' = L \left\{ 1 - \rho_m - \rho_{Re} \left[1 - \left(\frac{\Delta_R''}{\Delta_R} \right)^2 \right] \right\} = 11.4 \times \{ 1 - 0.0035 - 0.4364 \times 0.99 \} \\ = 6.4 (\text{元})$$

于是得到容差设计表 (见表 10—14),

表 10—14

两个方案比较表

方案	成本 P(元)	平均质量损失 L(元)	总损失 $L_T = P + L$ (元)
原方案 Δ_R	0	11.4	11.4
第一方案 $\Delta_R' = \Delta_R / 2$	3	7.6	10.9
第二方案 $\Delta_R'' = \Delta_R / 10$	5	6.4	11.4

可见, 第一方案为最佳容差设计方案。综上所述, 我们选取电阻 R 的中心值为 9.5Ω 的二级品, 电感 L 的中心值为 0.01H 的三级品, 以此组装电感电路, 输出电流强度的期望值为 10A , 达到目标值, 且总损失最少。

10.6 FMEA 分析

10.6.1 FMEA 方法的适用范围

六西格玛设计的目的是减少产品的质量波动, 从而减少各种质量问题, 包括故障。在产品研制和设计优化、过程优化、质量改进、售后服务等工作中, FMEA 是一种复杂度介于因果分析图法 (鱼刺图) 和试验设计/田口方法之间的质量问题分析方法。FMEA (failure mode effects analysis) 在我国被译为“故障模式与影响分析”, 迄今基本上也只应用于故障的分析。应当注意, 名词“故障”来自于英文单词 failure, 就英文本意而言, failure 的含义是“失败、失效、故障、事故、中断、损坏、折断、衰退、缺少、不足等”, 可见 failure 的含义本不限于“故障”, 而是包含了各种质量问题在内。FMEA 方法的适用范围包括产品和工艺设计中的各种质量问题的分析。在国际标准 ISO 9004: 2000《质量管理体系 业绩改进指南》中, 已把 FMEA 作为对“产品和过程的确认和更改”以及对“设计和开发”进行风险评估的工具。

为了避免出现更多的名词, 也为了充分利用 20 年来 FMEA 方法在可靠性工程中的应用的经验和基础, 本节仍沿用“故障模式与影响分析”的名称, 在分析表格中, 被分析的问题仍称为“故障模式”, 但强调本章中所称“故障”均泛指各种质量问题, “故障模式”是指质量问题模式。另外, 本节介绍的 FMEA 方法实际上已包括以估算风险度 (风险优先级数) RPN 为主的简化的危害性分析 (CA) 方法。

10.6.2 FMEA 方法的分类

1. 硬件法和功能法

当产品可按设计图纸及其他工程设计资料明确确定时,一般采用硬件法。这种分析方法适用于从零件级开始分析质量问题,再扩展到部件级、产品级、系统级,即自下而上地进行。然而也可从任一层次开始进行分析。硬件法 FMEA 是较为严格的。

当产品处于研发初期,设计图纸尚未形成时,一般采用功能法进行 FMEA 分析。这种方法认为产品应具备若干功能,而功能可按输出的质量特性分类,质量特性可按一定的技术指标来度量。使用功能法时,将输出的质量特性一一列出,并对其故障进行分析。

采用功能法时,一般自上而下进行分析,即从约定层次向下分析。然而也可在产品的任一层次开始向任一方向进行。

2. DFMEA 和 PFMEA

FMEA 还可以划分为设计 FMEA (DFMEA) 和过程 FMEA (PFMEA)。DFMEA 适用于产品设计阶段的故障模式与影响分析,而 PFMEA 则适用于工艺设计和生产制造过程的故障模式与影响的分析。

进行 FMEA 的目的是分析产品或过程的故障模式及其产生的影响或后果,并对可能出现的各种故障模式采取设计、工艺或操作方面的改进或补偿措施。

10.6.3 FMEA 的一般程序

1. 建立产品的框图或过程的流程图

(1) 收集分析产品(或系统,下同)的技术规范,明确产品的任务阶段及工作方式,产品的结构及功能描述,产品的工作环境及任务时间等。

(2) 根据需建立产品的功能框图。在产品研制初期采用功能法开展 FMEA 分析或被分析的产品较复杂时,应建立产品的功能框图以表示产品及其各功能单元的工作情况和相互关系,以及产品和每个约定层次的功能逻辑顺序。

(3) DFMEA 框图。进行 DFMEA 前,应把产品分割成具有独立功能的部件,分析其余度结构,按照完成任务和满足顾客需求的要求,建立用于 DFMEA 的框图。根据需要,DFMEA 框图可以包括原材料、机器设备、环境条件、人的因素、计算机软件等与产品故障模式与影响有关的要素,这是 DFMEA 框图与可靠性框图的不同之处。

(4) 过程流程图。由于大量的过程一般都是成熟的过程,为了突出重点,在进行 PFMEA 前,应建立过程流程/风险评估图,列出所有的工序,进行风险评估,选择高风险的工序进行 PFMEA。

2. 填写 FMEA 表格

DFMEA 是以设计工程师为主的团队分析产品每一个潜在的故障模式,并确定其发生原因和机理的一种分析技术,通过填写 DFMEA 表格来完成;而 PFMEA 则是以制造主管工程师为主的团队分析过程的每一个潜在的故障模式,并确定其发生原因和机理的一种分析技术,通过填写 PFMEA 表格来完成。通用的故障模式与影响分析表和过程故障模式与影响分析表,分别见 10.6.4 节和 10.6.5 节。

在分析产品的某一故障模式时,一般可只考虑对于产品工作的影响,但如果某几个故障模式同时发生的概率不能忽略不计时,还应考虑这几个故障模式同时发生的情况,作为一个新的故障模式来处理。如果发现严重度 $S \geq 8$ 的故障模式,必须采取设计和工艺的更改措施,以消除这些故障模式或明显降低其发生的概率;对于 RPN ,应规定某一临界值 C ,当某一故障模式的故障原因的 $RPN \geq C$ 时,应采取纠正或补偿措施; C 的取值应根据组织的具体要求来定,例如可定为 120~150 以上。

DFMEA 和 PFMEA 分析应形成按技术责任制审签的技术报告,其重点是:本产品(或过程)的薄弱环节是什么?与竞争对手比较,本产品(或过程)有哪些不足?针对顾客当前的和潜在的需求,有何改进措施?必须暂时保留的 $S \geq 8$ 的故障模式应列出清单。

随着产品研制的进展,DFMEA 和 PFMEA 分析应不断地进行修订和迭代,FMEA 报告应与产品文实相符,并成为积累排故经验、不断改进产品质量的有力工具。

10.6.4 DFMEA 分析

广义质量包括产品的功能、性能、可靠性、维修性、安全性、保障性、环境适应性、价格、外观、经济性、按时交货、售后服务等所有的质量特性。从工程具体操作而言,质量是指符合技术规范(包括产品规范、工艺规范、材料规范等)、工作说明和技术经济合同的要求,符合国家和行业的有关法规与标准,符合用户需求(含潜在需求)和参与市场竞争的要求。凡不符合这些要求的都是质量问题,在本章中,都称之为“故障”和“故障模式”。

1. DFMEA 分析用表格及填写说明

DFMEA 分析用表见表 10—15。填表说明如下:

第(1)栏(产品或部件/功能):填写被分析产品或部件的名称,该名称应与产品图纸或技术规范一致;并填写该产品或部件的功能,包括与接口设备的相互关系。

第(2)栏(潜在的故障模式):分析人员应根据产品技术规范和国家、行业的有关规定和标准所确定的各质量特性的指标要求,结合工程实践经验和类似产品所发生的故障,分析确定所有可能出现的故障模式。

第(3)栏(故障影响):填写每个潜在的故障对产品使用、功能或状态所导致的后果。如果被分析的产品不是最终产品,必要时还需考虑故障对更高层次的产品直至最终产品的影响。

第(4)栏(严重度 S):按照表 10—16 所推荐的产品故障严重度(S)评分标准确定故障模式的严重度 S 。

第(5)栏(故障原因/故障机理):确定并说明与潜在的故障有关的各种原因,包括设计缺陷,工艺缺陷,人机料法环因素,特定的物理、化学过程等。必要时还应考虑相邻层次的故障原因。

第(6)栏(频数 O):估计每个故障模式和故障起因发生的可能性,按照表 10—17 所推荐的产品故障频数(O)评分标准来确定故障模式或故障起因的频数 O 。

表 10—15

故障模式与影响分析表

初始约定层次

完成部门:

DFMEA 编号:

第 页共 页

约定层次

分析：

审核:

批准:

填表日期： 年 月 日

[illegible]

表 10—16 推荐的产品故障严重度 (S) 评分标准

影响	评分标准	严重度
无警告的严重危害	在没有任何预兆的情况下发生的、影响使用安全或违反有关法律法规的极其严重的故障模式。	10
有警告的严重危害	在有故障预兆的情况下发生的、影响使用安全或违反有关法律法规的极其严重的故障模式。	9
很大	产品（或系统）不能工作、丧失基本功能的严重故障模式，顾客非常不满。	8
大	产品（或系统）能工作、但性能下降的故障模式，顾客不满。	7
中等	产品（或系统）能工作、但司职舒适和方便的一些部件损坏，顾客感觉不便。	6
小	产品（或系统）能工作、但司职舒适和方便的一些部件降级工作，顾客感觉有些不便。	5
很小	加工精细、外观、噪声等方面不符合要求，大部分顾客能发现有缺陷。	4
轻微	加工精细、外观、噪声等方面不符合要求，有一半顾客发现有缺陷。	3
很轻微	加工精细、外观、噪声等方面不符合要求，但很少有顾客发现缺陷。	2
无	无影响。	1

表 10—17 推荐的产品故障频数 (O) 评分标准

故障概率	设计寿命期内可能的故障率	频数
很高：故障几乎不可避免	$\geq 1/2$	10
	1/3	9
高：反复发生的故障	1/8	8
	1/20	7
中等：偶尔发生的故障	1/80	6
	1/400	5
	1/2 000	4
低：相对很少发生的故障	1/15 000	3
	1/150 000	2
极低：几乎不可能发生的故障	1/1 500 000	1

第 (7) 栏 (设计控制)：说明在现行设计方案中已经采取的设计控制措施。控制措施大致有五类：1) 能阻止故障起因或故障模式的发生，或减少其发生概率。应尽可能选择此类措施。2) 虽不能阻止故障模式的发生，但能有效地减轻故障影响的严重度。汽车采用的安全带和安全气囊即属此类。3) 虽不能阻止故障模式的发生，但能自动检测、隔离故障，切除故障部件，使系统在降级的情况下继续工作。4) 虽不能阻止故障模式的发生，但能自动检测、隔离故障并报警，以便操作人员采取防范措施。5) 虽不能阻止故障模式的发生，但能自动或人工检测、隔离故障并记录有关信息，以方便维修。

第 (8) 栏 (检测难度 D)：取决于现行设计控制措施的有效程度，可按照表 10—18 所推荐的产品故障检测难度 (D) 评分标准来确定故障模式或故障起因的检测难度 D，需注意设计控制措施的有效程度越低，检测难度就越高。

表 10—18 推荐的产品故障检测难度 (D) 评分标准

检测可能性	利用现行设计控制方法检测故障模式的可能性	检测难度
绝对不可能	没有设计控制方法能检测故障模式	10
很微小	利用现行设计控制方法检测故障模式的可能性很微小	9
微小	利用现行设计控制方法检测故障模式的可能性微小	8
很小	利用现行设计控制方法检测故障模式的可能性很小	7
小	利用现行设计控制方法检测故障模式的可能性小	6
中等	利用现行设计控制方法检测故障模式的可能性中等	5
中等偏上	利用现行设计控制方法检测故障模式的可能性中等偏上	4
高	利用现行设计控制方法检测故障模式的可能性高	3
很高	利用现行设计控制方法检测故障模式的可能性很高	2
几乎肯定能	利用现行设计控制方法几乎肯定能检测故障模式	1

第 (9) 栏 (风险度 RPN): 计算并填写表示故障模式的风险度等级的 RPN 数值:

$$RPN = S \times O \times D$$

S, O, D 的数值应由 DFMEA 小组充分讨论后确定。RPN 的引进提供了一种量化评估故障模式的风险及改进措施的有效性的手段。

第 (10) 栏 (建议的改进或补偿措施): 首先针对 $RPN \geq C$ 的故障模式的起因提出纠正和预防措施, 以消除该故障或降低故障发生的风险。若故障模式的根本起因不详, 应通过试验设计等手段来确定主要起因。若对某一特定起因尚无纠正和预防措施, 应提出设计、工艺、质控和操作上的补偿措施。若没有任何可采取的措施, 在该栏填写“无”字予以明确。对于危及人员安全的故障模式, 若不能消除, 必须提出防护措施。

第 (11) 栏 (责任部门/责任人及完成日期): 对于所提出的改进或补偿措施, 应同时建议责任部门/责任人及完成日期。

第 (12) 栏 (实施的改进/补偿措施和日期): 填写实际执行的改进/补偿措施和日期。

第 (13)、(14)、(15)、(16) 栏: 参照第 (4)、(6)、(8)、(9) 栏分别确定并填写改进后的严重度 S、频数 O、检测难度 D 和 RPN 值, RPN 值的降低就是六西格玛设计或六西格玛改进的绩效的一种体现。

DFMEA 表头的“初始约定层次”是要进行 FMEA 的总的、完整的产品所在的层次, “约定层次”是初始约定层次下的相继层次。这些约定层次表明了从总体到局部、直至较简单的组成部分的有顺序的系列。

2. DFMEA 分析的例子

表 10—19

汽车车门的 DFMEA 分析

DFMEA 分析是针对产品的约定层次来进行的, 在本例中, 初始约定层次是某型汽车的车体, 约定层次是汽车前门组件。前门组件的潜在故障模式有很多, 表 10—19 只给出了汽车车门的 DFMEA 分析表的第 1 页, 该页有一个潜在故障模式“内侧底部面板被腐蚀”, 从表 10—19 可以了解 DFMEA 分析的具体方法。

表 10—19

故障模式与影响分析表

初始约定层次:某型汽车车体

完成部门:车体部/设计室

DFMEA 编号:D 010

第 1 页共 1 页

约定层次:汽车前门 CT/QM 000 12

分析:×××

审核:×××

批准:×××

填表日期:××年××月××日

(1) 产品或 部件/功能	(2) 潜在 的故障模式	(3) 故障影响	(4) 严重度 S	(5)故障原因/ 故障机理	(6) 频数 O	(7)现有设计控制	(8) 检测 难度 D	(9) 风险 度 RPN	(10)建议的改进 或补偿措施	(11)责任人 门/责任人 及完成日期 改进后	改进后					(16) 风险 度 RPN
											(12)实施的 改进/补偿 措施和日期	(13) 严重度 S	(14) 频数 O	(15) 检测 难度 D		
汽车前门 CT/QM- 000-12 功能: ● 进出汽车 ● 使车内人员不受天气、噪声的影响和外部冲撞的伤害 ● 承载包括镜子、门铰链、门插销和车窗调整器等金属构件 ● 使汽车美观	1. 车门内侧底部面板被腐蚀	车门寿命品质降低导致: ● 随着时间的推移,水气透过涂层使车门生锈,造成不满意的外观 ● 车门内金属构件的功能被削弱	7	内门板上的防护蜡的上缘太低 蜡层厚度不够 使用了不合适的蜡 气流的作用使蜡未能进入角落/边缘地带 喷涂的蜡阻塞了门的排水孔 门板之间未给喷头留下足够的通路	6 4 2 5 3 4	通用的车辆耐久试验 T-118,T-109,T-301 通用的车辆耐久试验(同上) 进行物理和化学试验,完成试验报告 无 最糟的喷涂状态对孔的大小进行试验室测试 对于新车门设计,用计算机辅助三维空间协调的方法进行喷头通道评估	7 7 2 8 1 4	294 196 28 280 21 112	增加试验室加速腐蚀测试 增加试验室加速腐蚀测试,对蜡层厚度进行 DOE 无 对生产用喷射设备和蜡进行气流作用的实验和评估 无 对使用辅助装配架和喷头进行团队评估	车体部 ××年××月××日 车体部 ××年××月××日 无 车体部 ××年××月××日 ……	防护蜡层上缘提高 125mm,××年××月××日 试验证明蜡层厚度够了 实验证明基本不存在此现象,××年××月××日 已评估,通路足够,××年××月××日	7 7 7 7 …	2 2 1 1 …	2 2 3 1 …	28 28 21 7 …	
2. ……	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	



10.6.5 PFMEA 分析

PFMEA 分析中使用的“故障”一词与可靠性工程中的故障定义有所差别，过程故障包括了更广的含义，例如各种情况的超差、组装错误、热处理和表面处理不合格以及环境变化所引起的过程波动等。PFMEA 的目标是：

- 确定潜在的与过程故障模式有关的产品；
- 评估过程故障对顾客的潜在影响；
- 确定引起故障的潜在的制造或装配过程原因，并确定哪一类过程故障是应该重点避免的以及相关的关键过程；
- 找到和发现一系列过程故障模式，为今后的分析工作打下良好的基础，建立一个优选的系统。

PFMEA 最早应在可行性论证前或可行性论证时进行，先于对产品的加工，按照从单独的元部件到装配件的过程，考虑所有的加工操作。在制造计划阶段，新过程（或修改过程）的早期检查和分析可以帮助预测、解决和监控潜在的过程关键问题。在产品整个研制过程中，应适时对过程 FMEA 进行迭代完善。

1. PFMEA 分析用表格及填写说明

在进行 PFMEA 前，首先应了解“顾客”的定义，PFMEA 的“顾客”可以是“最终用户”，也可以是后续的制造或装配操作以及维修活动。

另外，进行 PFMEA 的过程中，应提倡团队合作精神，负责 PFMEA 的工程师希望 FMEA 尽可能地包含所有受到影响的工作范围，如设计、装配、制造、材料、质量控制、维修和供货商以及与下一步装配有关的工作。PFMEA 起到一种催化剂的作用，在各个受影响的功能之间进行权衡协调，集思广益提出解决问题的措施，从而促进团队合作。

进行 PFMEA 的前提是假定产品的设计方案均能实现，设计的目标均能达到，由于设计缺陷而产生的潜在故障，在难以判别时，可以包括在 PFMEA 中，但不是必须包括的；这些潜在故障的影响和改进措施原则上应涵盖在 DFMEA 中。

在进行 PFMEA 时，应先建立过程流程/风险评估图表，该流程图应明确每一步操作的工艺特征。

过程故障模式与影响分析表见表 10—20。

填表说明如下：

第（1）栏（过程名称、功能要求）：填写被分析过程的名称，该名称应与工艺规程一致；并填写该过程的功能、过程操作的目的、操作要求和与其他过程的关系等。

第（2）栏（潜在的过程故障模式）：分析人员应根据工艺规程和国家、行业的有关规定和标准所确定的各质量特性的指标要求，结合工程实践经验、专业技术知识和类似过程所发生的故障，分析确定所有可能无法满足过程要求和/或设计意图的故障模式。应注意，在 FMEA 的编写中，应假设正在加工的零部件/材料是没有问题的。

第（3）栏（过程故障影响）：尽可能使用表示产品功能的词汇，将过程缺陷对部件及系统的影响提出来。故障影响是故障模式对顾客的影响，这里的顾客可以是下一过程、部件、系统、最终用户等，在进行分析时，应考虑到对所有顾客的影响，填写每个潜在的故障对过程输出以及生产线所导致的后果。

表 10—20

过程故障模式与影响分析表

产品名称/型号

完成部门:

PFMEA 编号:

第 页 共 页

过程名称

分析:

审核:

批准:

填表日期: 年 月 日

(1) 过程名称、功能要求	(2) 潜在的故障模式	(3) 过程故障影响	(4) 严重度 S	(5) 过程的故障原因/故障机理	(6) 频数 O	(7) 过程控制	(8) 检测难度 D	(9) 风险度 RPN	(10) 建议的改进或补偿措施	(11) 责任部门/责任人及完成日期	改进后				
											(12) 实施的改进/补偿措施和日期	(13) 严重度 S	(14) 频数 O	(15) 检测难度 D	(16) 风险度 RPN

第（4）栏（严重度 S）：按照表 10—21 所推荐的过程故障严重度（S）评分标准确定故障模式的严重度 S。

表 10—21 推荐的过程故障严重度（S）评分标准

影响	评分标准	严重度
无警告的严重危害	在没有任何预兆的情况下发生的、影响操作人员和机器安全或违反有关法律法规的极其严重的故障模式。	10
有警告的严重危害	在有故障预兆的情况下发生的、影响操作人员和机器安全或违反有关法律法规的极其严重的故障模式。	9
很大	对生产线造成较大的破坏，产品可能必须 100% 报废，使产品或系统丧失主要功能的故障模式，顾客非常不满。	8
大	对生产线造成较小的破坏，产品需分选，部分产品报废，导致产品或系统降级工作，顾客不满。	7
中等	对生产线造成较小的破坏、部分产品报废（但产品不需分选），司职舒适和方便的一些部件不工作，顾客感觉不方便。	6
小	对生产线造成较小的破坏，产品可能必须 100% 返工，司职舒适和方便的一些部件降级工作、顾客感觉有些不满。	5
很小	对生产线造成较小的破坏，产品需分选，部分产品返工，顾客不满，加工精细、外观、噪声等方面不符合要求，大部分顾客能发现有缺陷。	4
轻微	对生产线造成很小的破坏，部分产品站外在线返工，加工精细、外观、噪声等方面不符合要求，有一半顾客发现有缺陷。	3
很轻微	对生产线造成很小的破坏，部分产品站内在线返工，加工精细、外观、噪声等方面不符合要求，但很少有顾客发现缺陷。	2
无	无影响。	1

第（5）栏（过程的故障原因/故障机理）：确定并说明与潜在的故障有关的各种原因，包括过程设计的缺陷，工艺参数的量值，人机料法环因素，检验试验方法，特定的物理、化学过程等。必要时还应考虑相关过程的故障原因。尽可能列出每个故障模式所对应的任何可以想到的故障原因。故障原因的定义应详细而明确（如操作人员未安装密封垫），不能使用模糊不清的定义（如操作人员出错、机器故障等）。

第（6）栏（频数 O）：估计每个故障模式和故障起因发生的可能性，按照表 10—22 所推荐的过程故障频数（O）评分标准来确定故障模式或故障起因的频数 O。

表 10—22 推荐的过程故障频数（O）评分标准

故障概率	设计寿命期内可能的故障率	C_{pk}	频数
很高：故障几乎不可避免	$\geq 1/2$	< 0.33	10
	$1/3$	≥ 0.33	9
高：反复发生的故障	$1/8$	≥ 0.51	8
	$1/20$	≥ 0.67	7
中等：偶尔发生的故障	$1/80$	≥ 0.83	6
	$1/400$	≥ 1.00	5
	$1/2\ 000$	≥ 1.17	4
低：相对很少发生的故障	$1/15\ 000$	≥ 1.33	3
	$1/150\ 000$	≥ 1.50	2
极低：几乎不可能发生的故障	$1/1\ 500\ 000$	≥ 1.67	1

第(7)栏(过程控制):说明在现行设计方案中已经采取的过程控制措施,如工作环境的控制、工序的检验、特种工序的参数控制、统计工序控制、5S管理、检测不同故障的测试装置及其检测方法等。过程控制措施的分类可参照10.6.4节中第(5)栏的有关设计控制措施分类的说明。

第(8)栏(检测难度 D):取决于现行设计控制措施的有效程度,可按照表10—23所推荐的过程故障检测难度(D)评分标准来确定故障模式或故障起因的检测难度 D ,需注意过程控制措施的有效程度越低,检测难度就越高。

表 10—23 推荐的过程故障检测难度(D)评分标准

检测可能性	利用现行过程控制方法检测故障模式的可能性	检测难度
绝对不可能	没有过程控制方法能检测故障模式	10
很微小	利用现行过程控制方法检测故障模式的可能性很微小	9
微小	利用现行过程控制方法检测故障模式的可能性微小	8
很小	利用现行过程控制方法检测故障模式的可能性很小	7
小	利用现行过程控制方法检测故障模式的可能性小	6
中等	利用现行过程控制方法检测故障模式的可能性中等	5
中等偏上	利用现行过程控制方法检测故障模式的可能性中等偏上	4
高	利用现行过程控制方法检测故障模式的可能性高	3
很高	利用现行过程控制方法检测故障模式的可能性很高	2
几乎肯定能	利用现行过程控制方法几乎肯定能检测故障模式	1

第(9)栏(风险度 RPN):计算并填写表示过程故障模式的风险度等级的 RPN 数值, RPN 的计算与 $DFMEA$ 的相同。

第(10)、(11)、(12)栏的确定和填写与 $DFMEA$ 的相同。

第(13)、(14)、(15)、(16)栏可参照表10—20第(4)、(6)、(8)、(9)栏分别确定并填写经过程改进后过程故障的严重度 S 、频数 O 、检测难度 D 和 RPN 值, RPN 值的降低就是过程六西格玛设计或六西格玛改进的绩效的一种体现。

2. 过程 FMEA 分析的例子

图 10—19

某兼容磁头装置的极尖加工过程故障模式和影响分析

磁头是磁带记录仪的关键部件之一,它是一种电磁、磁电换能器,其质量直接关系到磁带机记录及重放信号的好坏。某兼容磁头装置是 IRIG 标准宽频带计测记录—重放仪的磁头,信道数为 14 道/英寸,要求在 3 米/秒带速下记录重放 400Hz~2MHz 的宽频信号。

此类磁头录放过程的特点是:工作频带宽、偏磁频率高、信号波长短、带速快、头—带压强、对失真度和信噪比要求苛刻。其工艺复杂,制造精度要求很高。

单个磁头的结构如图 10—22 所示,根据类似磁头的加工经验,其中加工难度最大的是磁头极尖的生产加工,且从功能上进行分析,磁头极尖质量直接关系到磁头的质量以及磁带记录仪录放信号的质量。由于极尖加工难度大,极易造成报废,因此,需要进行该型兼容磁头装置加工过程的 PFMEA 分析。

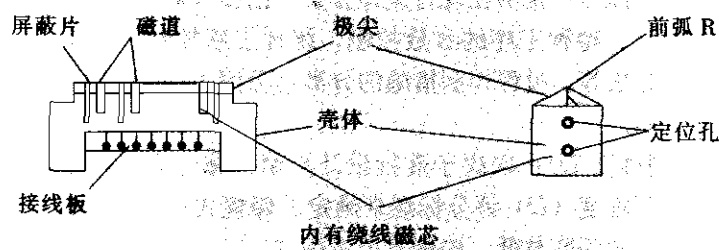


图 10—22 磁头结构示意图

首先进行过程流程和风险评估分析，结果如表 10—24 所示。

表 10—24 磁头加工过程流程/风险评估表

	过程步骤	风险评估		过程步骤	风险评估
1	磁头壳体的加工	低风险	8	磁头极尖复合面研抛	中风险
2	屏蔽片的加工	中风险	9	磁头后磁芯组复合面研抛	中风险
3	接线板的加工	低风险	10	极尖与后磁芯的复合组装	低风险
4	磁头后磁芯的加工	中风险	11	灌环氧树脂固封磁头	中风险
5	磁头后磁芯绕线	低风险	12	磁头前弧面加工	中风险
6	磁头后磁芯入壳	低风险	13	前弧 R 研抛	中风险
* 7	磁头极尖的加工	高风险	14	打定位孔	中风险

对标 * 号的高风险过程“磁头极尖的加工”进行 PFMEA 分析，结果见下页表 10—25。

对表中风险度 RPN 值进行分析发现，故障模式 2（缝隙宽度不均匀且超差）的故障原因 1（缝隙平面度不满足要求）的 RPN 值最高，为 280；故障模式 4（缝隙面烧伤或裂纹）的故障原因 1（缝隙平面研抛加工应力偏大）的 RPN 值其次，为 168，均大于 150。而这两项故障原因的纠正措施均为改进缝隙平面的研抛过程，因此，再对缝隙平面研抛过程作下一层次的 PFMEA 分析。

例 10—20

某兼容磁头装置磁头极尖缝隙平面研抛过程的 PFMEA 分析。通过例 10—19 磁头极尖加工过程的 PFMEA，确定关键过程为磁头极尖缝隙平面研抛过程，现对磁头极尖缝隙平面研抛过程进行 FMEA 分析，根据分析结果，采取改进措施。

首先进行过程流程和风险评估分析，结果如表 10—26 所示。

表 10—26 磁头极尖缝隙平面研抛过程流程/风险评估表

	过程步骤	风险评估		过程步骤	风险评估
1	用丙酮、酒精清洗待磨极尖料块	低风险	* 6	用磨床细磨极尖缝隙平面	高风险
2	用 502 胶将极尖料块粘在工艺垫块上	中风险	7	从夹具上取下工艺垫块并清洗	低风险
3	在磨床上调校夹具的平行度和垂直度	低风险	8	手工精抛极尖缝隙平面	中风险
4	将工艺垫块装在夹具上	低风险	9	清洗	低风险
5	用磨床粗磨极尖缝隙平面	中风险	10	检测平面度	低风险

表 10—25

某兼容磁头装置过程故障模式与影响分析表

产品名称/型号:某兼容磁头装置

完成部门:工艺室

PFMEA 编号:P-003

第 5 页共 17 页

过程名称:磁头极尖加工

分析:

审核:

批准:

填表日期:2005 年 5 月 8 日

(1) 过程名称、功能要求	(2) 潜在的过程故障模式	(3) 过程故障影响	(4) 严重度 S	(5) 过程的故障原因/故障机理	(6) 频数 O	(7) 过程控制	(8) 检测难度 D	(9) 风险度 RPN	(10) 建议的改进或补偿措施	(11) 责任部门/责任人及完成日期	改进后				
											(12) 实施的改进/补偿措施和日期	(13) 严重度 S	(14) 频数 O	(15) 检测难度 D	(16) 风险度 RPN
磁头极尖加工: 要求形成满足技术条件的高磁阻间隙。	1. 缝隙开裂。	极尖报废。	8	1. 缝隙玻璃选用不当, 粘接强度低。	3	1. 目视。	4	96	1. 选用满足要求的粘接玻璃。	2003 年 5 月 31 日前工艺室提出改进措施及培训要求, 人教处组织上岗培训。	2003 年 5 月 28 日前有关单位已实施第 10 栏所有措施, 其中对磨床细磨极尖缝隙平面的措施又做了 PFMEA 分析, 见修改工艺后的 PFMEA 表 008 号。	8	2	4	64
	2. 缝隙宽度不均匀且超差。	极尖报废。	8	2. 烧结工艺不合理。	2	2. 用高倍显微镜检测。	2	32	2. 改进烧结工艺。						
				1. 缝隙平面度不满足要求。	7		5	280	1. 改进缝隙平面研磨工艺。						
				2. 缝隙中夹有杂质。	2		4	64	2. 保持加工间的洁净度。						
				3. 合缝工艺问题。	2		3	48	3. 改进合缝工艺。						
				4. 合缝夹具不满足要求。	3		2	48	4. 修整合缝夹具。						
	3. 缝隙过大或过小。	极尖报废。	8	1. 膜层厚度控制不好。	5	3. 用高倍显微镜检测。	2	80	1. 调整膜层厚度。						
				2. 缝隙粗糙度不满足要求。	3		5	120	2. 加强缝隙面的抛光并进行检测。						
				1. 缝隙平面研磨加工应力偏大。	4	4. 用高倍显微镜检测。	7	168	1. 改进研磨加工工艺。						
				2. 研磨工艺操作人员技术不过关。	2		5	60	2. 对操作工人进行培训。						
	4. 缝隙面烧伤或裂纹。	料块磁性下降。	6												

标 * 号的过程为高风险过程, 因此, 对用磨床细磨极尖缝隙平面过程进行 PFMEA 分析, 结果见下页表 10—27。

本案例通过对极尖加工过程的 PFMEA 分析, 找出影响极尖加工成品率的 4 个潜在故障模式及所有的故障原因, 并采取改进设计、改进工艺、改进管理等措施, 使所有 RPN 值均有所降低, 并全都低于 120; 使磁头极尖的成品率从类似磁头极尖的 20% 上升到 50% 左右, 并且与类似磁头相比, 技术指标也明显提高, 使整机顺利达到预定的技术指标要求。

表 10—21

汽车车门内侧的上蜡过程故障模式与影响分析过程流程/风险评估图表详见表 10—28。

表 10—28 汽车车门内侧上蜡过程流程/风险评估表

过程	过程步骤	风险评估	过程	过程步骤	风险评估
1	从支架上拿起上蜡器	低风险	5	移开上蜡器	中风险
2	打开车门	低风险	6	关上车门	低风险
* 3	将上蜡器喷头插入车门并拉动启动装置 12 秒	高风险	7	将上蜡器放回支架上	低风险
4	松开启动装置等待 3 秒	中风险	8		

标 “*” 号的过程为高风险过程, 需进行 FMEA 分析, 结果见表 10—29。

10.7 面向 X 的设计

10.7.1 面向 X 设计的原理

设计方案的实施受到工艺、制造、装配、试验检验、环境保护、使用维护、成本、交付周期等诸多因素的制约。面向 X 的设计 (design for X, DFX) 是为了满足顾客对产品全寿命周期的所有要求, 针对产品全寿命周期的各种要素进行设计的方法的统称, 包括面向制造的设计 (design for manufacturing, DFM)、面向装配的设计 (design for assembly, DFA)、面向成本的设计 (design for cost, DFC)、面向试验的设计 (design for test, DFT)、面向可靠性和维修性的设计 (design for reliability & maintainability, DFRM)、面向售后服务的设计 (design for service, DFS)、面向环境的设计 (design for environment, DFE), 以及面向其他因素的设计等一系列面向某一需重点考虑或保证的领域的设计技术。

由于专业分工和知识面的局限, 设计人员通常着重关心和着力去实现的主要是产品性

表 10—27

某兼容磁头装置过程故障模式与影响分析表

产品名称/型号:某兼容磁头装置

完成部门:工艺室

PFMEA 编号:P-008

第 15 页共 17 页

过程名称:磁头极尖缝隙平面研抛

分析:

审核:

批准:

填表日期:2003 年 6 月 2 日

(1) 过程名称、功能要求	(2) 潜在的过程故障模式	(3) 过程故障影响	(4) 严重度 S	(5) 过程的故障原因/故障机理	(6) 频数 O	(7) 过程控制	(8) 检测难度 D	(9) 风险度 RPN	(10) 建议的改进或补偿措施	(11) 责任部门/责任人及完成日期	改进后					(16) 风险度 RPN
											(12) 实施的改进/补偿措施和日期	(13) 严重度 S	(14) 频数 O	(15) 检测难度 D		
磨床细磨极尖缝隙平面,要求极尖缝隙平面具有良好的平面度和粗糙度。	1. 缝隙表面划伤。	1. 影响磁头某些磁道的性能。	4	1. 研磨时未按要求进行冷却。	3	1. 目视。	1	12	1. 加冷却并提高加工人员水平。	2003 年 6 月 30 日前工艺室修改工艺,人教处组织了补充培训。	2003 年 6 月 25 日工艺室修改了工艺,人教处组织了补充培训。	4	2	1	8	
	2. 缝隙表面产生微裂纹。	2. 影响磁头的质量和使用寿命。	7	2. 砂轮进刀量偏大或未使用研磨冷却或操作人员技术问题。	2	2. 用高倍显微镜检测。	7	98	2. 提高加工人员加工技术水平和责任心。			7	1	6	42	
	3. 未达到要求的平面度(如塌边等)。	3. 影响磁头某些磁道的性能。	4	3. 砂轮平面度不满足要求。	5	3. 平晶检测。	1	20	3. 加强对砂轮平面度的检测并定期修磨砂轮。			2	5	1	10	
	4. 未达到要求的粗糙度。	4. 造成磁头实际缝隙宽度增大,影响磁头性能。	7	4. 加工时间不够或未使用规定粒度的砂轮。	6	4. 目视或粗糙度仪检测。	2	84	4. 提高加工人员加工技术水平和责任心。			7	3	2	42	

表 10—29

过程故障模式与影响分析表

产品名称/型号: 汽车前门 CT/QM-000-1

完成部门: 车体部/装配车间

PFMEA 编号: P-1450

第 1 页共 1 页

过程名称: 车门内侧人工上蜡过程

分析:

审核:

批准:

填表日期: 2003 年 5 月 17 日

(1) 过程名称、功能要求	(2) 潜在的过程故障模式	(3) 过程故障影响	(4) 严重度 S	(5) 过程的故障原因/故障机理	(6) 频数 O	(7) 过程控制	(8) 检测难度 D	(9) 风险度 RPN	(10) 建议的改进或补偿措施	(11) 责任部门/责任人及完成日期	改进后				
											(12) 实施的改进/补偿措施和日期	(13) 严重度 S	(14) 频数 O	(15) 检测难度 D	(16) 风险度 RPN
将上蜡器喷头插入车门并拉动启动装置 12 秒, 要求用尽可能薄的蜡层覆盖车门内部较低位置的表面以防止腐蚀。	未能在规定的表面上覆盖上足够的蜡。	1. 车门寿命品质降低导致。	7	1. 人工插入上蜡器喷头时插的不够深。	8	1. 每班每小时目视检查薄膜厚度和覆盖层。	5	280	1. 为喷射工人增加深度挡块。	制造工程 03.10.15	增加了挡块和在线喷射器检查。	7	2	5	70
		2. 随着时间的推移, 通过涂层产生车门生锈, 造成不满意的的外观。	7	2. 喷头阻塞。 ● 黏度太高 ● 温度太低 ● 压力太低	5	2. 在启动时和闲置一段时间后检测喷射式样, 采用预防维修方法维护喷头。	3	105	2. 自动喷射。	制造工程 03.12.15	03.10.02	7	1	3	21
		3. 车门内金属构件的功能被损害。	7	3. 由于碰撞造成喷头变形。	2	3. 采用预防维修方法维护喷头。	2	28	3. 对黏度、温度、压力进行试验设计。	制造工程 03.10.01	确定温度和压力界限并进行控制, 过程受控, $C_{pk}=1.85$	7	2	2	28
		4. 喷射时间不够。	7	4. 喷射时间不够。	8	4. 严格按照用法说明并且大子样 (10 个门/班) 检查临界区域的覆盖情况。	7	392	4. 安装喷射计时器。	维修部 03.09.15	安装自动喷射计时器, 控制图显示过程受控 $C_{pk}=2.05$	7	1	7	49

能,而对于制造与装配技术、成本、开发周期、可靠性与维修性、环境保护、全寿命周期的产品保障和维护使用、售后服务以及产品的升级、产品的报废处置等缺乏深入的了解和关心,因此导致产品设计与制造、装配、成本控制、开发周期、可靠性与维修性、使用保障、售后服务、报废处置等方面不同程度地脱节。

系统工程、并行工程和价值工程的发展提供了解决上述问题的基础。系统工程是应用科学和工程技术的成就,研究、设计、实现、运转复杂系统的一门综合性工程技术。它的主要内容是:

(1) 通过界定、综合、分析、设计、试验和评价等反复迭代的过程,将顾客的要求转换为对系统性能参数技术状态的描述。

(2) 综合有关的技术参数,确保所有物理、功能和计划的接口间具有相容性,以便优化和确认整个系统的总体设计方案。

(3) 将可靠性、维修性、安全性、人因工程、可生产性、使用保障、质量保证、环境工程、全寿命周期费用(LCC)等专门工程专业因素综合到整个工程中去,使费用、进度和技术性能达到总目标。系统工程所要综合的专门工程专业因素大体就是 DFX 所要考虑的因素。系统工程的目标是实现系统总体效果最优,同时实现这一目标的具体方法或途径也要求达到最优。

并行工程是集成地、并行地设计产品及其相关过程(包括制造过程和支持过程)的系统方法,这种方法要求产品研制人员将原来顺序进行的工作在时间和空间上交叉和部分重叠,在一开始就考虑产品整个寿命周期中从概念形成到产品报废的所有因素,包括质量、成本、进度计划和用户要求,以便提高产品设计质量,保证产品一次成功,缩短产品研制周期。并行工程的理论和方法为提高产品质量、降低成本、缩短产品开发周期和上市时间、更好地满足顾客的需求创造了条件。据文献报道,实施并行工程将使早期生产中工程更改次数减少 50% 以上,产品研制周期缩短 40%~60%,生产成本降低 30%~40%。

价值工程是以功能分析和功能评价为手段,以最低的总成本可靠地实现产品(或劳务)的必要功能为目的而进行的一种有组织的技术经济分析综合活动。产品总成本或寿命周期成本为产品研制成本与顾客使用成本之和。价值分析时应站在顾客的立场上,努力满足顾客需要的必要功能,努力去掉不必要的、过剩的功能;对完成产品必要功能的方案设计、结构设计、材料选择、强度和安全性设计、工艺设计、生产制造、供应链管理、项目管理等各方面进行分析,努力降低总成本。

面向 X 的设计本身是一项系统工程,是追求产品研制方案最优的一个过程。在产品总体方案设计阶段,要大力贯彻质量功能展开、系统设计等先进技术,紧扣顾客的需要,运用各种系统设计技术,追求最简约、可靠、高效的设计方案。简化设计不仅减少了制造、装配的工作量,提高了产品的可靠性,缩短了研制周期,还大量地节省了产品材料费、工时费,以及包装、库存、运输、维护、使用、售后服务等支持费用。在产品开发的整个过程中,应对面向 X 的设计进行补充、完善和迭代。

面向 X 的设计中需考虑的各种要素 X,可能是相互矛盾、相互制约的,在实施 DFX 时,必须进行分析权衡,以期总体上更好地满足顾客的需求和获取最高的效益。

DFX 是一种优化产品设计的方法,也是一种先进的设计理念。其作用方式有两种: DFX 设计准则和 DFX 分析评价工具。前者指将 X 要素方面专家的有关知识和经验整理成

具体的设计准则,然后在它们的指导下进行产品设计,相当于在这些专家的直接帮助下选择设计方案,确定产品结构。后者指产品设计进行到一定程度后,通过系统分析影响产品的X要素的各种因素,对产品的X要素进行评价,在此基础上给出改进设计的建议。DFX在产品开发过程中的作用和地位主要表现在根据顾客对产品全寿命周期的要求进行优化设计、减少零件数(从而精简产品结构)、改进可制造性和可装配性、改善使用维护和售后服务、降低产品成本等方面。

10.7.2 面向制造和装配的设计

面向制造和装配的设计(DFMA)也称作设计制造一体化,是DFX中最重要的部分。DFMA的指导思想是:产品制造、装配和试验检验中的很多问题和困难源于设计不当。在产品设计阶段,工程师除了要满足性能要求外,还必须尽可能地考虑和满足可制造性、可装配性的要求。

1. DFMA 总则

(1) 简化设计。在总体方案设计中,首先应当贯彻简化设计的原则,即采用尽可能简洁的结构来实现产品的功能。方案设计的简化不仅减少了零部件的数量,节省了设计费、材料费,更节省了大量的加工费、装配费、库存费、运输费、维护保养费等,在源头采取措施将事半功倍、一本万利。必须纠正“简单的设计方案体现不出水平,先进的方案一定比较复杂”的错误思想,切忌把简单的事情复杂化。最优秀、可靠性最高、效益最高的方案是能够满足顾客需求的最简约的方案。

(2) 模块化、系列化、通用化(三化)设计。

模块化设计:在产品研制中,尽可能地将产品中功能相对独立的组成部分设计和制造成通用性很强的、可以在今后新产品研制中重用的模块或单元,以便在新产品的研制中根据需要直接应用通用模块(或做适应性修改后)进行结构的搭建。

系列化设计:根据顾客当前的和潜在的需求以及不同顾客群的需求,针对某类产品的基本结构和参数,选择代表性强、规格适中、结构先进、用量最大的一种牌号产品进行技术攻关和突破,形成一个可扩展的基本产品或技术平台,在此基础上进行一系列改进、改型,开发出基于某个技术平台的多姿多彩的系列产品,以满足不同顾客群、不同应用环境的需要。

通用化设计:对各类产品中大量存在的功能相同、结构类似的零部件进行归并统一,形成可以彼此互换的标准件或标准化程度较高的通用零部件。

三化设计的推广应用,将极大地提高可制造性、可装配性,节省人力、物力和时间。

(3) 互换性设计。产品互换性设计是采用适当的尺寸公差,以便控制加工精度,实现同一种零部件不同个体之间的无条件互换。

(4) 继承性设计。尽可能采用先进的、经过使用考验的、成熟的设计,包括工作原理、计算公式、设计方案、现成的零部件、成熟的工艺、现成的工装设备和生产线、行之有效的试验检验方法等,以便最大限度地减少设计、制造、安装、维护、保障的工作量和费用,最大限度地降低产品研制的风险。

(5) 贯彻并行工程,实施团队工作法。DFMA的实施需组织包括设计、工艺、试验检验、计划管理、质量保证、财务等部门的跨专业、跨部门的团队,自设计早期开始就全

面深入地考虑与制造和装配有关的所有问题,提出解决措施,并在整个研制过程中不断完善,以保证产品的可制造性、可装配性、可靠性、维修性和保障性。

(6) 与 DFMA 有关的数字化技术。DFMA 的实施需要广博的知识和大量的数据和信息,如能应用相应的计算机辅助软件可以提高工作的效率。适用的计算机辅助软件有美国 BDI 公司的 DFMA 软件等。DFMA 软件将设计、装配、材料和加工工艺的知识集成在一起,设计者可以充分利用这些知识立刻作出合理的决策,而不必再费力去获取第一手资料或咨询专家;开发团队则可以利用 DMFA 让大家都在同一时间思考同一问题。与当前产品开发方法相比,应用 DFMA 的结果是以更低的成本和更短的时间得到了更高质量的产品,促进产品并行设计和创新设计的实现。

DFMA 的设计理念对美国工业界产生了巨大的影响,DFMA 的应用对美国制造业竞争优势的形成做出了重大的贡献。据报道,一般来说,DFMA 软件在一个项目上就可收回投资。其主要依据如下:DFMA 可以缩短产品从设计到投产的周期高达 50%,减少零件数量平均达 30%~70%,减少装配时间达 50%~80%。1991 年,鉴于 BDI 公司在 DFMA 领域的创意、研发和市场化的杰出工作,美国总统布什曾给 BDI 公司创始人颁发了美国国家技术奖。

另外,随着信息技术和 CAD 技术的发展,产品开发已越来越多地采用数字化设计和制造技术。在产品阶段,应开发数字样机,进行工作仿真和维修演示,在计算机上进行空间关系和运动件的演示、协调、检查,通过虚拟设计和虚拟制造实施 DFMA。

2. 面向制造的设计

DFM 方法的主要内容是:在产品设计的早期阶段即考虑与制造有关的约束,指导设计师对同一零件进行不同材料和工艺的选择,提高产品的可制造性;对不同制造方案进行制造时间和成本的快速定量估计,全面比较与评价各种设计与工艺方案等。设计团队根据这些定量的反馈信息,在零件的早期设计阶段就能够及时改进设计,确定一种最满意的设计和工艺方案。

为了提高可制造性,应在保证满足设计要求的前提下尽可能选择易于加工的材料、最低的尺寸公差和形位公差要求、合适的尺寸标注基准、适当的热处理和表面处理要求、成熟的工艺方案等。

3. 面向装配的设计

DFA 指在产品早期阶段考虑并解决装配过程中可能存在的问题,以确保零件快速、高效、低成本地进行装配。DFA 是一种针对装配环节的统筹兼顾的设计思想和方法,就是在产品设计过程中利用各种技术手段如分析、评价、规划、仿真等,充分考虑产品的装配环节以及与其相关的各种因素的影响,在满足产品性能与功能的条件下改进产品的装配结构,使设计出的产品是可以装配的,并尽可能降低装配成本和产品总成本。DFA 在产品开发过程中的作用和地位主要表现在减少零件数(从而精简产品结构)、改进装配性能、降低产品成本等方面。

10.7.3 面向成本的设计

1. 面向成本设计(DFC)的原理

质量、进度和成本三者之间存在着永恒的矛盾,DFX 技术可以帮助我们兼顾三者,

寻找既能提高质量与可靠性,又能缩短周期、降低成本的先进的设计方案。如 DFMA 技术以及后面章节要介绍的 DFRM, DFT, DFS, DFE 等技术中,都要考虑成本的因素。设计、制造技术的重大进步更是大幅度降低产品成本的重要途径。例如,微电子设计技术和加工工艺的突飞猛进带来计算机类产品的迅速更新换代和成本的快速下降。本节所要介绍的面向成本的设计是针对已经形成的设计方案具体地分析研究降低产品成本的方法。

DFC 是指在满足用户需求的前提下,通过分析和研究产品在设计、制造、装配、试验检验、使用、维修、回收或报废的全寿命周期中各个部分的成本组成情况并进行评估后,对原设计中成本过高的部分进行修改,以降低产品成本的设计方法。

根据文献报道,产品全寿命周期成本包括设计成本(约占 10%~15%)、制造成本(约占 30%~35%)、使用和维修成本(约占 50%~60%),以及其他成本(所占比例一般小于 5%)。需要指出的是,产品设计方案将影响几乎所有的成本因素,在很大程度上决定了产品全寿命周期的成本,例如,对罗尔斯·罗伊斯公司(Rolls Royce)的调查表明,2 000 个部件在产品的设计阶段决定了最终产品成本的 80%;通用汽车公司(General Motors)经理称卡车变速器 70%的制造成本是由设计决定的;福特汽车公司(Ford Motor)估计在设计、材料、劳动力和制造费用四个因素中,70%的生产节约源于设计的改进。因此,DFC 有着广阔的用武之地。

2. DFC 的主要工作内容

(1) 建立基于并行工程的全寿命周期成本模型。此模型应能在产品的不同设计阶段对成本进行相应的估算。

(2) 成本数据的收集与成本估算。应建立成本数据收集与评估系统。不同的估算方法有可能产生不同的结果,怎样合理地确定估算是值得研究的问题。一般来说在并行工程中,由于不同的设计阶段所产生的信息的完整程度不同,所采用的成本估算方法也不同,例如,在设计初期,可采用参数化估算方法、神经网络等方法;在详细设计阶段,可采用手册估算、详细估算等方法。而且,生产批量、产品类型、产品生命周期长短的不同在选择估算方法时也是需要考虑的因素,同时所选方法应该是灵活易用、可靠准确的。

(3) 与其他 DFX 工具的集成。由于 DFC 是面向产品全寿命周期的,这就要求它必须能够与其他 DFX(如 DFM, DFA, DFT, DFS 等)工具协调一致进行工作,即需要建立不同评价标准的协调原则和方法。

(4) 设计结果评价和改进设计建议的生成。这一工作需要其他 DFX 技术的支持下,即不能只考虑成本一个要素,需综合考虑其他 DFX 的评价模型提供的评价结论,才能生成合理的设计建议。

(5) 目标成本的确定。对于顾客有明确的价格限制要求或企业将产品的售价作为参与市场竞争的重要筹码时,应设置产品目标成本。在目前还缺乏精确方法的情况下,一般考虑市场、利润、工厂条件和生产批量等因素,利用历史资料确定目标成本。

3. 估算制造成本的主要方法

(1) 成组技术的应用。将品种繁多的机械零件,按其形状、尺寸、工艺等相似性分类归组,进行成组编码和识别,利用相似性原理进行成本估算,以便充分利用产品成本的历史信息,并增加成本信息处理时的样本量。

(2) 作业成本法(activity based cost, ABC)。以作业为核算对象,通过成本动因来

确认和计算作业量,进而以作业量为基础分配间接费用的成本计算方法。作业成本包括产品单位成本(如原材料、能耗、直接人工等)、生产批次成本(如工装夹具、模具、检测装置研制、产品例行试验等)、产品维持成本(如获取生产许可、生产线维护等)、工厂级成本(如生产线折旧、厂房维护等)。

4. 估算非制造成本的主要方法

非制造成本主要包括:

R&D成本:产品预研、技术储备和样机研制等方面的费用。

设计成本:用于市场调研和顾客需求分析、可行性研究、总体方案设计、初步设计、详细设计(完成产品和保障资源的成套生产图样和技术条件)、研制试验、设计更改、鉴定试验等的费用。

销售成本:用于广告宣传、促销、包装、储存、运输以及有关税金等费用。

保修成本:在保修期内提供产品维修,包括更换零件的费用。

使用成本:用于培训、能耗、耗材、维护保养、更换备件、保修期外的维修、环境控制、操作人员等的费用。

回收报废成本:产品报废处理和再生的费用。

管理费用:用于非直接制造人员的办公和试验费用(包括设施折旧)及工资,全体人员的通讯、福利、文化等方面的费用等。

非制造成本主要通过历史成本信息的收集、分析和统计获得。

5. 减少成本设计

在不改变总体设计方案的前提下,除了应用各种 DFX 技术外,还可以采用最低消耗设计,即在满足顾客需求,保证安全、强度和适当的设计裕度的原则下,采用合适的材料、最小的尺寸、最低的功耗、最低的加工精度等,以实现最低消耗设计。避免过功能、过强度、过寿命设计。

10.7.4 面向试验的设计

开展六西格玛设计的产品,其设计方案必须经过试验验证并确认达到了预期的性能和质量水平,才能投产;批生产中也必须按照设计确定的试验和检验方法才能进行验收和交付,因此,产品设计时,必须考虑设计验证、例行试验和产品检验的可行性和经济性。无法进行验证和检验的产品是不能保证质量,因而不能投产的;过于昂贵的验证和检验也必然会影响企业的绩效。

除了计算分析外,设计验证主要通过试验来进行,包括研制试验和鉴定试验。研制试验可以是各种仿真试验(包括半实物半仿真试验),也可以是真实条件下的实物试验;鉴定试验则是对新设计产品的性能、质量、可靠性和寿命进行真实环境下的全面考核,一般应由顾客代表和行业权威部门参加和认可。试验检验是根据设计确定的验收规程对批生产的产品进行全数或抽样的试验或检验,以获得产品的质量特性的量值,并判断合格与否。

DFT 要求在设计早期就论证和确定设计验证的方案及其设施,产品制造和装配过程中的测试、试验、检验和计量的方案及其设施。应当尽可能地应用现有的经过实际使用考验的设计验证和试验检验设施,但是对于大型复杂产品而言,现有设施往往不符合要求,需要在确定总体设计方案的同时确定设计验证和试验检验的方案并及早地启动相应设施的

组建。产品设计时,还必须考虑和确定原材料、元器件、外包工序和外包部件等的验收试验或检验规范。DFT 是保持产品一致性,减少质量波动的重要方法。

DFT 必须考虑的另一个方面是故障的检测和排除。应在深入到零件一级的 FMEA 分析的基础上列出产品的所有故障模式及其症状,针对每个故障模式确定需监控的参数和故障判断逻辑,并给出检测、隔离和排除的方法,以便当产品在使用中出现故障时能迅速准确地排除。

10.7.5 面向可靠性和维修性的设计

1. 面向可靠性的设计

可靠性作为质量的时间延续特性,已越来越多地受到人们的重视。如本章概论所说,六西格玛设计的核心是稳健设计(包括 QFD、系统设计、实验设计、参数设计、容差设计等方法),其宗旨是提高产品抵御环境变化、制造误差和磨损老化等各种干扰的能力,减少产品质量波动。而实质上,稳健设计在减少产品质量波动的同时,也肯定提高了产品的可靠性。前面已介绍了稳健设计的主要技术方法,包括通常也归于稳健设计范畴的 DFMEA 和 PFMEA 分析,本节主要介绍面向可靠性的经典设计方法。

可靠性设计的目标是在顾客所要求的寿命期内不出或尽可能少出故障,即满足顾客关于寿命和平均故障间隔时间的要求并降低全寿命周期费用(LCC)。这个目标只有从产品研制开始就紧密结合产品研制深入开展可靠性设计和分析工作才有可能达到。可采用的可靠性设计方法包括可靠性指标论证与确定,可靠性分配与预计,制定和贯彻可靠性设计准则,开展简化设计、热设计、降额设计、余度设计、耐环境设计等;可采用的可靠性分析方法包括 FMEA 分析以及故障树分析(FTA)、事件树分析(ETA)、热分析、容差分析等。

(1) 可靠性指标论证与确定。对于电子产品等偶然失效占统治地位的产品,应论证与确定平均故障间隔时间 MTBF 的指标,对于耗损失效占统治地位的产品,应论证与确定寿命指标;对于兼有偶然失效和耗损失效的产品应论证与确定平均故障间隔时间 MTBF 和寿命两种指标。产品的寿命不是越长越好,应当根据顾客的需求来确定。在产品的寿命期间,平均故障间隔时间 MTBF 应尽可能长。

(2) 可靠性分配与预计。为了保证产品能满足顾客对可靠性的指标要求,应当在设计早期自顶向下地将基本可靠性和任务可靠性指标分配到各部件和零件,并自下而上地对零部件和整机的基本可靠性和任务可靠性指标进行预计,以便评估在实现可靠性指标方面设计方案的可行性。通过可靠性预计,可以发现可靠性的薄弱环节,对这些薄弱环节应采取设计和工艺的改进措施,以提高产品的可靠性水平。

(3) 可靠性设计准则。可靠性设计准则是有助于提高产品可靠性的定性设计要求的归纳总结,应制定并要求设计员贯彻可靠性设计准则,在设计评审时进行可靠性设计准则的符合性检查。

(4) 经典可靠性设计方法。

简化设计:产品的结构越简单,零部件越少,可靠性肯定越高。

降额设计:限制元器件、零部件的载荷(包括电流、电压、应力、温度、功耗等)水平,使其低于额定值的设计方法。

冗余设计：用多于一种途径来完成一个规定功能的设计方法。

耐环境设计：采用如热设计、减振、空调、防尘、防沙、防霉菌、避光、防湿、隔噪声等局部改善环境的设计方法。

(5) 可靠性分析方法。

FMEA 分析：详见 10.6 节。

故障树分析 (FTA)：通过对可能造成产品故障的硬件、软件、环境、人为因素进行分析，画出故障树，从而确定产品故障原因的各种可能组合方式和（或）其发生概率的一种分析技术。

热分析：对电子产品及其各组成部分的温度及分布进行计算、测量和分析。

容差分析：针对元器件的容差积累进行最坏情况分析或蒙特卡罗分析的方法。

(6) 故障数据管理系统的设计。制定和实施故障收集、机理分析、故障纠正、故障预防（包括举一反三）、故障统计分析、故障数据库的建立等。

2. 面向维修性的设计

维修性 (maintainability) 也是产品的一种质量特性，即由产品设计赋予的使其维修简便、迅速和经济的固有特性。提高维修性也是降低全寿命周期费用 (LCC)，以满足预期的费用和维修活动的重要途径。维修性的指导思想是：一旦产品出现故障能迅速、经济地发现和排除，因此，维修性设计是可靠性设计的重要补充。很多情况下，维修性已经成为当今商品的重要卖点。

面向维修性的设计缩写为 DFMT，以区别于面向制造的设计 DFM，其工作主要包括：

- (1) 平均修复时间 MTTR 的确定和分配；
- (2) 维修策略；
- (3) 可达性设计；
- (4) 防错设计（参见 DFMA）；
- (5) 人因工程设计等；
- (6) 测试性设计等。

10.7.6 面向售后服务的设计

六西格玛设计的目标是在产品寿命周期内最大限度地满足顾客的需求，这就必须在设计阶段就考虑和解决售后保障和维修服务的所有问题。主要包括产品维修服务策略的确定，维修性的设计，售后保障资源的设计，质量可靠性信息系统的建立等。面向售后服务的设计 (DFS) 的效益在于：降低保修成本、提高顾客满意度和经济地维护对环境敏感的长寿命产品。

1. 维修服务策略的确定

产品售后维护、修理、服务的活动和费用无论对于顾客还是制造商都是很大的负担，但是归根结底产品全寿命周期的费用都是由顾客支付的，因为制造商的售后服务费用也已经包括在产品售价中。精心实施 DFS 对于提高顾客满意度和为整个社会节省资源都是十分重要的。开展六西格玛设计时，需进行产品可靠性与维修性的权衡，制定维修服务策略，并进行全寿命周期费用 (LCC) 的估算。LCC 由两部分组成：产品研制费用和维护使用费用。提高产品可靠性会增加研制费用，但可减少维护使用费用，应通过 LCC 的分析

以确定合适的可靠性和维修性水平,进而确定维修服务策略。

根据产品的可靠性水平(故障发生频率)和寿命,以及维修性(含测试性)的设计方案,进行产品的维修分析,确定产品的所有维修活动及时机,提出保障资源需求和维修服务策略:

- (1) 售后保障资源的方案及设计需求;
- (2) 免费保修的期限和保修范围;
- (3) 维修网络的规模;
- (4) 备件的需求量及供应;
- (5) 产品的报废处理。

2. 维修性的设计

根据可靠性与维修性权衡的结果进行面向维修性的设计。

3. 售后保障资源的设计

售后保障资源包括:

- (1) 随机装箱的保障资源,如使用说明书,培训资料,随机配送的附件、备件、耗材、专用工具等;
- (2) 维修站配备的测试和维修设备、备件、零部件、元器件、技术说明书、维护使用说明书、故障诊断和排故资料、维修人员培训资料等;
- (3) 远程咨询服务系统。

售后保障资源的设计是产品开发工作的重要组成部分,应当与产品的设计同步进行。

4. 质量可靠性信息系统的建立

产品的质量可靠性信息系统是收集传递顾客的意见、愿望、抱怨的主要渠道,也是产品故障信息的主要来源。应确定信息系统的组织机构,规范信息收集的内容,设计适用的表格,明确信息收集人员的职责,建立先进的信息收集、传递、统计、分析、检索、报表生成等手段,以便充分地了解顾客的意见和需求,作为改进设计、工艺和服务的依据。

10.7.7 面向环境的设计

面向环境的设计包括两方面的含义:环境保护设计和环境适应性设计。

1. 环境保护设计

随着全球性环境立法的出现,需要新的工具帮助制造商设计绿色的产品,以便在相关法律颁布以前能够顺利过渡到生产合法的产品,避免对环境的污染。

产品的设计应尽可能采用不会造成环境污染的材料和制造技术。例如,冰箱、空调制造业已经逐渐摒弃了把氟里昂作为冷却剂的传统技术。在产品制造过程中,必须考虑有害废气、废液、废渣的净化排放和操作工人的防护。在产品使用中,必须使可能产生的污染环境的噪声、闪光、电磁辐射、高温、振动、有害气体、异味等低于法定的限度。在许多国家,产品报废时的处理责任也已经交给了制造商,因此产品设计时,还必须考虑产品报废时对环境的影响。许多公司开始回收报废的产品,从计算机等电子产品到各种消费类产品。通过使用 DFE 软件对这些产品进行面向环境和成本效益的设计,制造商能够获得该产品的同业竞争优势。DFE 模拟产品报废时的拆卸过程,揭示产品设计对于成本和环境的影响。这些定量的信息可以辅助设计早期的各种平衡性决策。

DFE 可以与 DFA 软件无缝集成在一起发挥作用, 初始的拆卸顺序可以由 DFA 软件自动生成。用户可以编辑拆卸列表以定义产品报废时不必拆卸就可以回收或处理的组件。基于拆卸的顺序, 可以完成两个主要的分析: (1) 拆卸、处理、再利用和回收的利益回报分析; (2) 原产品制造、处理、再利用或回收所引起的环境影响分析。

利益回报分析显示了在不同拆卸阶段的利益影响。它等于每个零部件的拆卸成本与可再利用零部件价值之差, 再减去产品剩余零部件的处理成本。

产品的环境影响分析可采用荷兰人研究和开发的一个称为 MET 点系 (materials (材料), energy (能源), toxicity (毒性)) 的价值分析模型。该模型来源于欧洲的生命周期分析项目, 后来研究结果被总结成一个快速、有效的分析工具, 并成为 DFE 软件的一部分。它使得没有专门生态知识背景的用户能够在设计中考虑环境的因素。

2. 环境适应性设计

使产品本身具有很强的抵御环境变化的能力 (参见稳健设计), 或改善局部工作环境, 减少环境变化对产品性能和可靠性的影响 (参见耐环境设计)。

10.7.8 面向其他因素的设计

根据 DFX 的原理, 参照以上几种 DFX 的方法, 结合工程实际需要, 还可以考虑面向其他因素的设计, 例如面向开发周期的设计、面向某一顾客群的设计、面向安全的设计、面向便携的设计、面向人因工程的设计等。

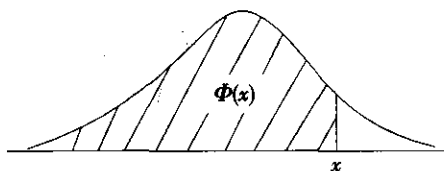
本章小结

本章介绍了六西格玛设计的产生和发展, 从技术、管理、经济等方面阐述了六西格玛设计的重要意义, 提出了产品全寿命周期的六西格玛解决方案, 并以主要的篇幅介绍了六西格玛设计的 IDDOV 流程和主要技术方法。

10.1 节介绍了 IDDOV 流程的五阶段: 识别 (identify)、界定 (define)、研发 (design)、优化 (optimize) 和验证 (verify), 以及每个阶段的工作内容和适用的技术方法, 提出了产品全寿命周期的六西格玛解决方案, 并将 IDDOV 流程、DMAIC 流程和可靠性工程有机结合, 首次提出了能覆盖产品全寿命周期的完整的六西格玛解决方案。10.2~10.7 节分别介绍了六西格玛设计的主要技术方法: 质量功能展开、系统设计、参数设计、容差设计、FMEA 分析和面向 X 的设计。

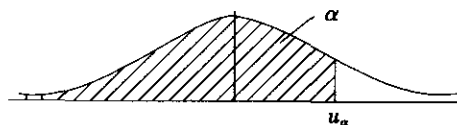
六西格玛设计的内容十分丰富, 开展六西格玛设计也是当前六西格玛管理深入发展的方向。但由于本教材的重点是六西格玛改进 (DMAIC), 因此第 10 章只是对于六西格玛设计的简要介绍。需要深入学习了解六西格玛设计的读者, 可以阅读本书后所列的参考文献。

附表 1

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$


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	0.5190	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8970	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9270	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	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
x	0.00	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
3.0	0.9 ² 8650	0.9 ³ 0324	0.9 ³ 3129	0.9 ³ 5166	0.9 ³ 6631	0.9 ³ 7674	0.9 ³ 8409	0.9 ³ 8922	0.9 ⁴ 2765	0.9 ⁴ 5190

附表 2

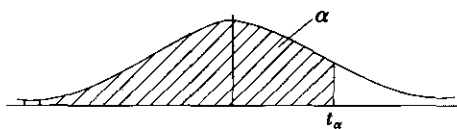
标准正态分布的 α 分位数表

α	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	—	-2.33	-2.05	-1.88	-1.75	-1.64	-1.55	-1.48	-1.41	-1.34
0.10	-1.28	-1.23	-1.18	-1.13	-1.08	-1.04	-0.99	-0.95	-0.92	-0.88
0.20	-0.84	-0.81	-0.77	-0.74	-0.71	-0.67	-0.64	-0.61	-0.58	-0.55
0.30	-0.52	-0.50	-0.47	-0.44	-0.41	-0.39	-0.36	-0.33	-0.31	-0.28
0.40	-0.25	-0.23	-0.20	-0.18	-0.15	-0.13	-0.10	-0.08	-0.05	-0.03
0.50	0.00	0.03	0.05	0.08	0.10	0.13	0.15	0.18	0.20	0.23
0.60	0.25	0.28	0.31	0.33	0.36	0.39	0.41	0.44	0.47	0.50
0.70	0.52	0.55	0.58	0.61	0.64	0.67	0.71	0.74	0.77	0.81
0.80	0.84	0.88	0.92	0.95	0.99	1.04	1.08	1.13	1.18	1.23
0.90	1.28	1.34	1.41	1.48	1.55	1.64	1.75	1.88	2.05	2.33

α	0.001	0.005	0.010	0.025	0.050	0.100
u_α	-3.090	-2.576	-2.326	-1.960	-1.645	-1.282

α	0.999	0.995	0.990	0.975	0.950	0.900
u_α	3.090	2.576	2.326	1.960	1.645	1.282

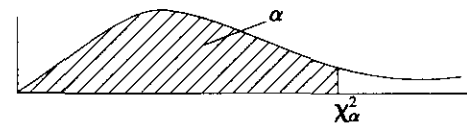
附表 3

 t 分布的 α 分位数表

df	$t_{0.60}$	$t_{0.70}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.861
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576

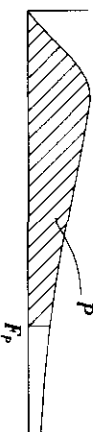
附表 4

χ^2 分布的 α 分位数表



df	$\chi^2_{0.05}$	$\chi^2_{0.01}$	$\chi^2_{0.025}$	$\chi^2_{0.05}$	$\chi^2_{0.10}$	$\chi^2_{0.90}$	$\chi^2_{0.95}$	$\chi^2_{0.975}$	$\chi^2_{0.99}$	$\chi^2_{0.995}$
1	0.000039	0.00016	0.00098	0.0039	0.0158	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.1026	0.2107	4.61	5.99	7.38	9.21	10.60
3	0.0717	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.5	29.05	51.81	55.76	59.34	63.69	66.77
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
120	83.85	86.92	91.57	95.70	100.62	140.23	146.57	152.21	158.95	163.64

附表 5

F 分布的 0.90 分位数 $F_{0.90}(n_1, n_2)$ 表 n_1 = 分子的自由度, n_2 = 分母的自由度

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.70	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.50	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76

续前表

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.80	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

n_1 = 分子的自由度, n_2 = 分母的自由度

F 分布的 0.95 分位数 $F_{0.95}(n_1, n_2)$ 表

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76

续前表

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

n_1 = 分子的自由度 n_2 = 分母的自由度

F 分布的 0.975 分位数 $F_{0.975}(n_1, n_2)$ 表

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.12	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97

续前表

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00



F 分布的 0.99 分位数 $F_{0.99}(n_1, n_2)$ 表

n_1 = 分子的自由度, n_2 = 分母的自由度

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6016	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26

续前表

n_1 m_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

附表 6

计量控制图计算控制限的系数表

子组 中观 测值 个数 n	控制限系数											中心线系数			
	A	A_2	A_3	B_3	B_4	B_5	B_6	D_1	D_2	D_3	D_4	c_4	$1/c_4$	d_2	$1/d_2$
2	2.121	1.880	2.659	0.000	3.267	0.000	2.606	0.000	3.686	0.000	3.267	0.7979	1.2533	1.128	0.8865
3	1.732	1.023	1.954	0.000	2.568	0.000	2.276	0.000	4.358	0.000	2.574	0.8862	1.1284	1.693	0.5907
4	1.500	0.729	1.628	0.000	2.266	0.000	2.088	0.000	4.698	0.000	2.282	0.9213	1.0854	2.059	0.4857
5	1.342	0.577	1.427	0.000	2.089	0.000	1.964	0.000	4.918	0.000	2.114	0.9400	1.0638	2.326	0.4299
6	1.225	0.483	1.287	0.030	1.970	0.029	1.874	0.000	5.078	0.000	2.004	0.9515	1.0510	2.534	0.3946
7	1.134	0.419	1.182	0.118	1.882	0.113	1.806	0.204	5.204	0.076	1.924	0.9594	1.0423	2.704	0.3698
8	1.061	0.373	1.099	0.185	1.815	0.179	1.751	0.388	5.306	0.136	1.864	0.9650	1.0363	2.847	0.3512
9	1.000	0.337	1.032	0.239	1.761	0.232	1.707	0.547	5.393	0.184	1.816	0.9693	1.0317	2.970	0.3367
10	0.949	0.308	0.975	0.284	1.716	0.276	1.669	0.687	5.469	0.223	1.777	0.9727	1.0281	3.078	0.3249
11	0.905	0.285	0.927	0.321	1.679	0.313	1.637	0.811	5.535	0.256	1.744	0.9754	1.0252	3.173	0.3152
12	0.866	0.266	0.886	0.354	1.646	0.346	1.610	0.922	5.594	0.283	1.717	0.9776	1.0229	3.258	0.3069
13	0.832	0.249	0.850	0.382	1.618	0.374	1.585	1.025	5.647	0.307	1.693	0.9794	1.0210	3.336	0.2998
14	0.802	0.235	0.817	0.406	1.594	0.399	1.563	1.118	5.696	0.328	1.672	0.9810	1.0194	3.407	0.2935
15	0.775	0.223	0.789	0.428	1.572	0.421	1.544	1.203	5.741	0.347	1.653	0.9823	1.0180	3.472	0.2880
16	0.750	0.212	0.763	0.448	1.552	0.440	1.526	1.282	5.782	0.363	1.637	0.9835	1.0168	3.532	0.2831
17	0.728	0.203	0.739	0.466	1.534	0.458	1.511	1.356	5.820	0.378	1.622	0.9845	1.0157	3.588	0.2787
18	0.707	0.194	0.718	0.482	1.518	0.475	1.496	1.424	5.856	0.391	1.608	0.9854	1.0148	3.640	0.2747
19	0.688	0.187	0.698	0.497	1.503	0.490	1.483	1.487	5.891	0.403	1.597	0.9862	1.0140	3.689	0.2711
20	0.671	0.180	0.680	0.510	1.490	0.504	1.470	1.549	5.921	0.415	1.585	0.9869	1.0133	3.735	0.2677
21	0.655	0.173	0.663	0.523	1.477	0.516	1.459	1.605	5.951	0.425	1.575	0.9876	1.0216	3.778	0.2647
22	0.640	0.167	0.647	0.534	1.466	0.528	1.448	1.659	5.979	0.434	1.566	0.9882	1.0119	3.819	0.2618
23	0.626	0.162	0.633	0.545	1.455	0.539	1.438	1.710	6.006	0.443	1.557	0.9887	1.0114	3.858	0.2592
24	0.612	0.157	0.619	0.555	1.445	0.549	1.429	1.759	6.031	0.451	1.548	0.9892	1.0109	3.895	0.2567
25	0.600	0.153	0.606	0.565	1.435	0.559	1.420	1.806	6.056	0.459	1.541	0.9896	1.0105	3.931	0.2544

资料来源:ASTM, Philadelphia, PA, USA.

附表 7

西格玛水平换算表

每百万机会缺陷数 (DPMO)	西格玛水平 (σ)
933 200	0
915 450	0.125
894 400	0.25
869 700	0.375
841 300	0.5
809 200	0.625
773 400	0.75
734 050	0.875
691 500	1
660 082	1.1
645 650	1.125
621 378	1.2
598 700	1.25
581 814	1.3
549 750	1.375
541 693	1.4
501 349	1.5
461 139	1.6
450 250	1.625
421 427	1.7
401 300	1.75
382 572	1.8
354 350	1.875
344 915	1.9
308 500	2
274 112	2.1
242 071	2.2
265 950	2.215
226 600	2.25
211 927	2.3
190 800	2.375
184 108	2.4
158 700	2.5
135 686	2.6
130 300	2.625
115 083	2.7
105 600	2.75
968 09	2.8
845 50	2.875
807 62	2.9

续前表

每百万机会缺陷数 (DPMO)	西格玛水平 (σ)
668 00	3
548 01	3.1
521 00	3.125
445 66	3.2
401 00	3.25
359 31	3.3
304 00	3.375
287 16	3.4
227 00	3.5
178 64	3.6
168 00	3.625
139 03	3.7
122 00	3.75
107 24	3.8
880 0	3.875
819 7	3.9
620 0	4
466 1	4.1
435 0	4.125
346 7	4.2
300 0	4.25
255 5	4.3
205 0	4.375
186 5	4.4
130 0	4.5
967	4.6
900	4.625
687	4.7
600	4.75
483	4.8
400	4.875
336	4.9
230	5
159	5.1
180	5.125
107	5.2
130	5.25
72	5.3
80	5.375
48	5.4
30	5.5

续前表

每百万机会缺陷数 (DPMO)	西格玛水平 (σ)
20	5.6
23.35	5.625
13	5.7
16.7	5.75
8	5.8
10.05	5.875
5	5.9
3.4	6

附表 8

符号检验表 $P(S \leq S_\alpha) = \alpha$

$\alpha \backslash n$	0.01	0.05	0.10	0.25	$\alpha \backslash n$	0.01	0.05	0.10	0.25	$\alpha \backslash n$	0.01	0.05	0.10	0.25	$\alpha \backslash n$	0.01	0.05	0.10	0.25
1					24	5	6	7	8	47	14	16	17	19	69	23	25	27	29
2					25	5	7	7	9	48	14	16	17	19	70	23	26	27	29
3				0	26	6	7	8	9	49	15	17	18	19	71	24	26	28	30
4				0	27	6	7	8	10	50	15	17	18	20	72	24	27	28	30
5			0	0	28	6	8	9	10	51	15	18	19	20	73	25	27	28	31
6		0	0	1	29	7	8	9	10	52	16	18	19	21	74	25	28	29	31
7		0	0	1	30	7	9	10	11	53	16	18	20	21	75	25	28	29	32
8	0	0	1	1	31	7	9	10	11	54	17	19	20	22	76	26	28	30	32
9	0	1	1	2	32	8	9	10	12	55	17	19	20	22	77	26	29	30	32
10	0	1	1	2	33	8	10	11	12	56	17	20	21	23	78	27	29	31	33
11	0	1	2	3	34	9	10	11	13	57	18	20	21	23	79	27	30	31	33
12	1	2	2	3	35	9	11	12	13	58	18	21	22	24	80	28	30	32	34
13	1	2	3	3	36	9	11	12	14	59	19	21	22	24	81	28	31	32	34
14	1	2	3	4	37	10	12	13	14	60	19	21	23	25	82	28	31	33	35
15	2	3	3	4	38	10	12	13	14	61	20	22	23	25	83	29	32	33	35
16	2	3	4	5	39	11	12	13	15	62	20	22	24	25	84	29	32	33	36
17	2	4	4	5	40	11	13	14	15	63	20	23	24	26	85	30	32	34	36
18	3	4	5	6	41	11	13	14	16	64	21	23	24	26	86	30	33	34	37
19	3	4	5	6	42	12	14	15	16	65	21	24	25	27	87	31	33	35	37
20	3	5	5	6	43	12	14	15	17	66	22	24	25	27	88	31	34	35	38
21	4	5	6	7	44	13	15	16	17	67	22	25	26	28	89	31	34	36	38
22	4	5	6	7	45	13	15	16	18	68	22	25	26	28	90	32	35	36	39
23	4	6	7	8	46	13	15	16	18										

秩和检验表 $P(T_1 < T < T_2) = 1 - \alpha$ [illegible]

游程总数检验表

$$R_{2,\alpha}$$
 表示满足 $P(R \geq R_2) \leq \alpha$ 的 R_2 中之最小整数
$$R_{2,\alpha}$$
 表示满足 $P(R \geq R_2) \leq \alpha$ 的 R_2 中之最小整数 $R_{1,0.025}$ $R_{2,0,025}$ [illegible] $R_{1,0.05}$ $R_{2,0.05}$

n_1	n_2	n_1	n_2
2	3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20	2	3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
3		3	
4	2	4	7 8
5	2 2 3	5	9 9
6	2 3 3 3	6	9 10 11
7	2 3 3 4 4	7	9 10 11 12
8	2 2 3 3 4 4 5	8	11 12 13 13
9	2 2 3 4 4 5 5 6	9	11 12 13 14 14
10	2 3 3 4 5 5 6 6 6	10	11 12 13 14 15 16
11	2 3 3 4 5 5 6 6 7 7	11	13 14 15 15 16 17
12	2 3 4 4 5 6 6 7 7 8 8	12	13 14 15 16 17 17 18
13	2 3 4 4 5 6 6 7 8 8 9 9	13	13 14 15 16 17 18 18 19
14	2 3 4 5 5 6 7 7 8 8 9 9 10	14	13 14 16 17 17 18 19 20 20
15	2 3 4 5 6 6 7 8 8 9 9 10 10 11	15	15 16 17 18 19 19 20 21 21
16	2 3 4 5 6 6 7 8 8 9 10 10 11 11 11	16	15 16 17 18 19 20 21 21 22 23
17	2 3 4 5 6 7 7 8 9 9 10 10 11 11 12 12	17	15 16 17 18 19 20 21 22 22 23 24
18	2 3 4 5 6 7 8 8 9 10 10 11 11 12 12 13 13	18	15 16 18 19 20 21 21 22 23 24 25
19	2 3 4 5 6 7 8 8 9 10 10 11 12 12 13 13 14 14	19	15 16 18 19 20 21 22 23 23 24 25 25 26
20	2 3 4 5 6 7 8 9 9 10 11 11 12 12 13 13 14 14 15	20	15 17 18 19 20 21 22 23 24 25 25 26 27 27

附表 11

游程最大长度检验表 $(P\{L \geq La\} \leq \alpha)$

$\alpha=0.01$

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
7				7	7	7																		
8				8	8	8	8																	
9			9	9	8	8	8	8																
10			10	10	9	9	8	8	8															
11			11	10	10	9	9	9	9	9														
12		12	12	11	11	10	10	9	9	9	9													
13		13	13	12	11	11	10	10	9	9	9	9												
14		14	13	13	12	11	11	10	10	10	9	9	9											
15		15	14	13	13	12	11	11	11	10	10	10	10	10										
16		16	15	14	13	13	12	11	11	11	10	10	10	10	10									
17		17	16	15	14	13	13	12	12	11	11	10	10	10	10	10								
18		18	17	15	15	14	13	13	12	12	11	11	11	11	10	10	10	10						
19		19	17	16	15	14	14	13	13	12	12	11	11	11	11	10	10	10	10					
20		19	18	17	16	15	14	14	13	13	12	12	11	11	11	11	10	10	10	10				
21		20	19	16	17	16	15	14	14	13	13	12	12	11	11	11	11	11	11	10	10			
22		21	20	18	17	16	15	15	14	14	13	13	12	12	12	11	11	11	11	11	11	11		
23	23	22	20	19	18	17	16	15	15	14	14	13	13	12	12	12	11	11	11	11	11	11	11	
24	24	23	21	20	19	18	17	16	15	15	14	13	13	13	12	12	12	11	11	11	11	11	11	11
25	25	24	22	21	19	18	17	16	16	15	14	14	13	13	13	12	12	12	11	11	11	11	11	11

$\alpha=0.05$

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
5			5	5																				
6		6	6	6	6																			
7		7	7	6	6	6																		
8		8	8	7	7	7	7																	
9		9	8	8	7	7	7	7																
10	10	10	9	8	8	7	7	7	7															
11	11	10	10	9	8	8	8	7	7	7														
12	12	11	10	10	9	9	8	8	8	8	8													
13	13	12	11	10	10	9	9	8	8	8	8	8												
14	14	13	12	11	10	10	9	9	8	8	8	8	8											
15	15	13	12	11	11	10	10	9	9	9	8	8	8	8										
16	16	14	13	12	11	11	10	10	9	9	9	8	8	8	8									
17	17	15	14	13	12	11	11	10	10	9	9	9	9	8	8	8								
18	17	16	14	13	12	12	11	11	10	10	9	9	9	9	9	9	9							
19	18	17	15	14	13	12	12	11	11	10	10	9	9	9	9	9	9	9						
20	19	17	16	15	14	13	12	11	11	10	10	10	9	9	9	9	9	9	9					
21	20	18	16	15	14	13	13	12	11	11	10	10	10	10	9	9	9	9	9	9				
22	21	19	17	16	15	14	13	12	12	11	11	10	10	10	10	9	9	9	9	9	9			
23	22	20	18	16	15	14	13	13	12	12	11	10	10	10	10	10	9	9	9	9	9	9		
24	23	20	19	17	16	15	14	14	13	12	12	11	11	10	10	10	10	9	9	9	9	9	9	9
25	24	21	19	17	17	15	14	14	13	12	12	12	11	11	10	10	10	10	9	9	9	9	9	9

附表 12

相关系数检验表

$$P(|r| > r_\alpha) = \alpha$$

α $n-2$	0.10	0.05	0.02	0.01	0.001	α $n-2$
1	0.98769	0.99692	0.999507	0.999877	0.9999988	1
2	0.90000	0.95000	0.98000	0.99000	0.99900	2
3	0.8054	0.8783	0.93433	0.95873	0.99116	3
4	0.7293	0.8114	0.8822	0.91720	0.97406	4
5	0.6694	0.7545	0.8329	0.8745	0.95074	5
6	0.6215	0.7067	0.7887	0.8343	0.92493	6
7	0.5822	0.6664	0.7498	0.7977	0.8982	7
8	0.5494	0.6319	0.7155	0.7646	0.8721	8
9	0.5214	0.6021	0.6851	0.7348	0.8471	9
10	0.4973	0.5760	0.6581	0.7079	0.8233	10
11	0.4762	0.5529	0.6339	0.6835	0.8010	11
12	0.4575	0.5324	0.6120	0.6614	0.7800	12
13	0.4409	0.5139	0.5923	0.6411	0.7603	13
14	0.4259	0.4973	0.5742	0.6226	0.7420	14
15	0.4124	0.4821	0.5577	0.6055	0.7246	15
16	0.4000	0.4683	0.5425	0.5897	0.7084	16
17	0.3887	0.4555	0.5285	0.5751	0.6932	17
18	0.3783	0.4438	0.5155	0.5614	0.6787	18
19	0.3687	0.4329	0.5034	0.5487	0.6652	19
20	0.3598	0.4227	0.4921	0.5368	0.6524	20
25	0.3233	0.3809	0.4451	0.4869	0.5974	25
30	0.2960	0.3494	0.4993	0.4487	0.5541	30
35	0.2746	0.3246	0.3810	0.4182	0.5189	35
40	0.2573	0.3044	0.3573	0.4032	0.4896	40
45	0.2428	0.2875	0.3384	0.3721	0.4648	45
50	0.2306	0.2732	0.3218	0.3541	0.4433	50
60	0.2108	0.2500	0.2948	0.3248	0.4078	60
70	0.1954	0.2319	0.2737	0.3017	0.3799	70
80	0.1829	0.2172	0.2565	0.2830	0.3568	80
90	0.1726	0.2050	0.2422	0.2673	0.3375	90
100	0.1638	0.1946	0.2331	0.2540	0.3211	100

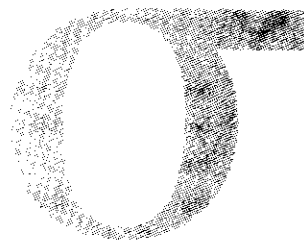
附表 13

正交多项式系数表
常用正交多项式表

$n=1$	$n=2$		$n=3$			$n=4$			
p_1	p_1	p_2	p_1	p_2	p_3	p_1	p_2	p_3	p_4
-1	-1	1	-3	1	-1	-2	2	-1	1
1	0	-2	-1	-1	3	-1	-1	2	-4
	1	1	1	-1	-3	0	-2	0	6
			3	1	1	1	-1	-2	-4
						2	2	1	1
$\frac{1}{2}$	2	$\frac{2}{3}$	5	4	$\frac{9}{5}$	10	14	$\frac{72}{5}$	$\frac{288}{35}$
2	1	3	2	1	$\frac{10}{3}$	1	1	$\frac{5}{6}$	$\frac{35}{12}$
$n=6$					$n=7$				
p_1	p_2	p_3	p_4	p_5	p_1	p_2	p_3	p_4	p_5
-5	5	-5	1	-1	-3	5	-1	3	-1
-3	-1	7	-3	5	-2	0	1	-7	4
-1	-4	4	2	-10	-1	-3	1	1	-5
1	-4	-4	2	10	0	-4	0	6	0
3	-1	-7	-3	-5	1	-3	-1	1	5
5	5	5	1	1	2	0	-1	-7	-4
					3	5	1	3	1
$\frac{35}{2}$	$\frac{112}{3}$	$\frac{324}{5}$	$\frac{567}{7}$	$\frac{400}{7}$	28	84	216	$\frac{3\ 168}{7}$	$\frac{4\ 800}{7}$
2	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{7}{12}$	$\frac{21}{10}$	1	1	$\frac{1}{6}$	$\frac{12}{20}$	$\frac{7}{20}$
$n=8$					$n=9$				
p_1	p_2	p_3	p_4	p_5	p_1	p_2	p_3	p_4	p_5
-7	7	-7	7	-7	-4	28	-14	14	-4
-5	1	5	-13	23	-3	7	7	-21	11
-3	-3	7	-3	-17	-2	-8	13	-11	-4
-1	-5	3	9	-15	-1	-17	9	9	-9
1	-5	-3	9	15	0	-20	0	18	0
3	-3	-7	-3	17	1	-17	-9	9	9
5	1	-5	-13	-23	2	-8	-13	-11	4
7	7	7	7	7	3	7	-7	-21	-11
					4	28	14	14	4
42	168	594	$\frac{12\ 672}{7}$	$\frac{31\ 200}{7}$	60	308	$\frac{7\ 128}{5}$	$\frac{41\ 184}{7}$	20\ 800
2	1	$\frac{2}{3}$	$\frac{12}{7}$	$\frac{10}{7}$	1	3	$\frac{5}{6}$	$\frac{7}{12}$	$\frac{3}{20}$

续前表

$n=10$					$n=11$				
p_1	p_2	p_3	p_4	p_5	p_1	p_2	p_3	p_4	p_5
-9	6	-42	18	-6	-5	15	-30	6	-3
-7	2	14	-22	14	-4	6	6	-6	6
-5	-1	35	-17	-1	-3	-1	22	-6	1
-3	-3	31	3	-11	-2	-6	23	-1	-4
-1	-4	12	18	-6	-1	-9	14	4	-4
1	-4	-12	18	6	0	-10	0	6	0
3	-3	-31	3	11	1	-9	-14	4	4
5	-1	-35	-17	1	2	-6	-23	-1	4
7	2	-14	-22	-14	3	-1	-22	-6	-1
9	6	42	18	6	4	6	-6	-6	-6
					5	15	30	6	3
$\frac{165}{2}$	528	$\frac{15\ 444}{5}$	$\frac{82\ 368}{5}$	78 000	110	858	$\frac{30\ 888}{5}$	41 184	249 600
2	$\frac{1}{2}$	$\frac{5}{3}$	$\frac{5}{12}$	$\frac{1}{10}$	1	1	$\frac{5}{6}$	$\frac{1}{12}$	$\frac{1}{40}$



常用六西格玛中英文名词索引

- 1-sample proportion test (单一样本比率检验)
1-sample t-test (单一样本 t 检验)
2-sample proportion test (双样本比率检验)
2-sample t-test (双样本 t 检验)
5S (SEIRI, SEITON, SEISO, SEIKETSU, SHITSUKE; 即整理、整顿、清扫、清洁、素养)
 6σ (六西格玛)
accelerated change (加速变革)
acceptable quality level, AQL (合格质量水平)
acceptance testing procedure, ATP (验收检验规程)
accuracy (准确性; 准确度)
action learning (学以致用)
activity based cost, ABC (作业成本法)
adjusted multiple correlation coefficient (调整的多元全相关系数)
adjustment factor (调节因子)
advanced quality system, AQS (先进质量体系)
affinity diagram (亲和图)
algorithm for inventive problem solving, AIPS or ARIZ (创造性问题解决步骤)
alias (别名)
alias structure (别名结构)
align (整合)
alpha error (α 错误, 也称第 I 类错误)
alternative hypothesis (备择假设)
analysis of control chart (控制图分析)
analysis of variance, ANOVA (方差分析)
Anderson-Darling test (Anderson-Darling 检验, 一种正态性检验方法)
artificial limits (人为界限)
asymmetrical orthogonal array (非对称正交表)
attribute measurement system analysis (属性值测量系统分析)
average quality loss (平均质量损失)
axiomatic design (公理性设计)

balanced scorecard (平衡计分卡)
Bartlett test (Bartlett 检验, 一种检验两个或多个总体方差是否相同的方法)

- benchmarking (水平对比)
- beta error (β 错误, 也称第 II 类错误)
- bias (偏倚)
- big Y (大 Y, 指影响组织发展和生存且必须满足的关键性指标)
- binomial distribution (二项式分布)
- black belt, BB (黑带)
- blitz team (突击队, 指为快速完成改进项目而成立的改进团队)
- block (区组)
- blocking (划分区组)
- bottleneck technique (瓶颈技术)
- box-and-whisker plots (箱线图)
- Box-Behnken design (Box-Behnken 设计)
- Box-Cox transformation (Box-Cox 转化, 一种将非正态数据转化为正态的方法)
- box plot (箱线图)
- brainstorming (头脑风暴法)
- business case (经营情况)

- capability maturity model, CMM (能力成熟度模型)
- cause and effect diagram (因果图)
- cause and effect matrix (因果矩阵)
- C chart (C 图, 缺陷数控制图)
- cash flow (现金流)
- cause and effect diagram (因果分析图)
- cause and effect matrix (因果矩阵)
- CDAM, combine, delete, add, or modify (组合、删除、增加、修改; 一种用于筛选、整理一系列观点的方法)
- center line, CL (中心线)
- center point (中心点)
- central composite circumscribed design, CCC (中心复合序贯设计)
- central composite design (中心复合设计)
- central composite face-centered design, CCF (中心复合表面设计)
- central composite inscribed design, CCI (中心复合有界设计)
- central limit theorem (中心极限定理)
- central tendency (数据分布中心)
- champion (倡导者)
- change acceleration framework (变革加速框架)
- change plan (变革计划)
- charter (团队或项目宪章, 或称特许任务书)
- check sheet (检查表)
- Chi square tests (卡方检验)
- coded unit (编码单位, 用 -1 和 +1 表示的因子位级)
- code of conduct (行为规范)
- coding (代码化)

- common cause (偶然原因)
- comparative methods (比较方法)
- complete randomized design (完全随机化设计)
- computer simulation (计算机模拟)
- concentration diagram (示意图, 一种用实物的图形作为收集数据的检查表, 在图形上标明出现了什么问题)
- concurrent engineering (并行工程)
- confidence coefficient $1-\alpha$ (置信水平 $1-\alpha$)
- confidence interval (置信区间)
- confirmation run (验证试验)
- confounded (混杂)
- contingency tables (列联表)
- continuous flow manufacturing, CFM (连续流生产)
- control array (控制表)
- control chart (控制图)
- control limits (控制界限)
- controllable factor (可控因子)
- critical to quality, CTQ (关键质量特性)
- corporate responsibility (公司责任)
- correlation coefficient (相关系数)
- cost benefit analysis (成本收益分析)
- cost of conformity (符合性质量成本)
- cost of nonconformity (非符合性质量成本)
- cost of poor quality, COPQ (不良质量成本或劣质成本)
- cost of quality, COQ (质量成本)
- critical business requirement, CBR (业务关键要求)
- critical to process, CTP (关键过程特性)
- critical to quality, CTQ (关键质量特性)
- critical customer requirement, CCR (顾客关键需求)
- critical value (临界值)
- cross array (乘积表; 交叉表)
- customer (顾客)
- customer requirements (顾客要求)
- customer value added activity (顾客增值活动)
- CUSUM control chart (累积和控制图)
- cycle time, CT (周期时间)
- cycle-time reduction (缩短周期)
- data analysis (数据分析)
- defect (缺陷)
- defects per million opportunity, DPMO (百万机会缺陷数)
- defects per opportunity, DPO (机会缺陷率)
- defects per unit (单位缺陷数)

defining relation (定义关系)
design for X, DFX (面向 X 的设计)
design failure mode and effect analysis, DFMEA (设计 FMEA)
design for assembly, DFA (面向装配的设计)
design for cost, DFC (面向成本的设计)
design for environment, DFE (面向环境的设计)
design for manufacturing, DFM (面向制造的设计)
design for reliability & maintainability, DFRM (面向可靠性和维修性的设计)
design for service, DFS (面向售后服务的设计)
design for six sigma, DFSS (六西格玛设计)
design for test, DFT (面向试验的设计)
design mapping process (设计的映射过程)
design of experiment, DOE (试验设计)
detection (检测难度)
digital simulation (数字仿真)
direct costs (直接成本)
discrete data (离散数据)
discrimination ratio (分辨力)
dispersion factor (散度因子)
distribution (分布)
DMADV (define, measure, analyze, design, and verify) (一种六西格玛设计流程, 代表界定、测量、分析、设计和验证)
DMAIC (define, measure, analyze, improve, and control) (六西格玛改进流程, 代表界定、测量、分析、改进和控制)
defects per million opportunities, DPMO (百万出错机会缺陷数)
defects per unit, DPU (单位缺陷数)
dynamic parameter design (动态参数设计)

effect sparsity principle (效应稀疏原则)
effect (效应)
engineering specialty integration (工程专业知识综合)
estimable (可估的)
event tree analysis (事件树分析)
evolutionary operation, EVOP (调优运算)
experiment environment (试验环境)
experiment unit (试验单元)
experimental error (试验误差)
experiment run (试验)
exponentially weighted moving average, EWMA (指数加权滑动平均)
EWMA control chart (指数加权滑动平均 (EWMA) 控制图)
external customers (外部顾客)
external noise (外噪声)
extreme vertices design (极端顶点设计法)



factor (因子)
factorial design (因子设计)
failure (故障; 失效)
failure mode (故障模式)
failure mode and effects analysis, FMEA (失效模式与影响分析)
failure tree analysis, FTA (故障树分析)
families of variation (变异类别)
financial metrics (财务指标)
five why's (五个为什么)
flow chart (流程图)
flow diagram (流程图)
force field analysis (力场分析)
fractional factorial design (部分因子设计)
full factorial design (全因子试验法)
F-test (F 检验)
functional deployment process map (跨职能流程图)

gage (测量仪器)
Gantt chart (甘特图)
generator (生成元)
goal statement (目标陈述)
goodness-of-fit test (拟合优度检验)
green belts, GB (绿带)

hidden factory (隐性工厂)
histogram (直方图)
hypothesis testing (假设检验)

ideal technical system (理想技术系统)
in control (受控)
independence assumption (独立性假设)
independent axiom (独立性公理)
independent sample (独立样本)
indirect costs (间接成本)
industrial engineering (工业工程)
information axiom (信息公理)
inherent process variation (过程固有波动)
input variable (输入变量)
institutionalizing (制度化)
integrated review process (全面的评审过程)
interaction (交互作用)
interaction effect (交互效应)

interim containment action, ICA (临时应急措施)

internal noise (内噪声)

internal estimate (区间估计)

internal customers (内部顾客)

interquartile range, IQR (四分位间距)

interval measurement scale (定距测量尺度)

inter-outer array (内外表)

is/is not (是非分析)

ISO 14001 (环境管理的国际标准)

ISO 9000 (质量管理体系国际标准)

KANO analysis (KANO 分析)

key process input variable, KPIV (关键过程输入变量)

key process output variable, KPOV (关键过程输出变量)

killer phrases (扼杀性言语)

Kruskal-Wallis test (Kruskal-Wallis 检验, 一种用于多组样本数据的中位数之间是否存在差异的非参数方法)

lack of fit (失拟)

lateral thinking (水平思考)

lean concepts (精益思想)

lean six sigma (精益六西格玛)

lean tool (精益工具)

level (水平)

level of confidence (置信水平)

Levene test (Levene 检验, 一种用于分析多组数据的方差之间是否存在差异的非参数检验)

lower control limit, LCL (下控制限)

life cycle cost, LCC (全寿命周期费用)

linearity (线性)

location and dispersion modeling (位置和散度建模法)

location factor (位置因子)

logistic regression (logistic 回归)

main effect (主效应)

Mann-Whitney test (Mann-Whitney 检验, 一种分析两组独立的样本数据之间是否存在差异的非参数检验)

master black belt, MBB (资深黑带)

maxi-problem (扩大的问题)

mean (均值)

mean square of error, MSE (均方误)

mean time between failures, MTBF (平均故障间隔时间)

measure performance (测量绩效)

measurement system analysis, MSA (测量系统分析)



- median (中位数)
method time measurement, MTM (方法时间测量)
milestones (里程碑)
mini-problem (缩小的问题)
mission (使命)
mistake proofing (错误预防)
mixture design (混料设计)
mixtures of distributions (分布的混合)
mixture-amounts (混料—总量型)
mixture-process variable (混料—过程变量型)
mode (众数)
multiple comparison (多重比较)
multiple correlation coefficient (多元全相关系数)
multiple regression (多元回归分析)
multi-vari chart (多变异图)
multivoting (多轮投票)
multi-way layout (多向分类)
- natural limits (自然界限)
near rotatability (近似旋转性)
nested design (嵌套设计)
net present value, NPV (净现值)
noise array (噪声表)
noise factor (噪声因子)
nominal measurement scale (名义测量尺度)
non value-added activities (非增值活动)
nonparametric test (非参数检验)
normal distribution (正态分布)
normal effect plot (正态效应图)
null hypothesis (原假设或零假设)
number of distinct categories (可区分的类别数)
- objective tools (客观工具)
occurrence (频度)
one-factor (单因子)
one-factor-at-a-time, OFAT (单因子试验, 即每次只有一个因子改变其他因子固定不变的试验方法)
one-factor-at-one-time (单因子变化法; 一次一因子法)
one-way-layout (单向分类)
operational definition (作业定义)
ordinal measurement scale (定序测量尺度)
orthogonal array (正交表)
orthogonal experimental design (正交试验设计法)
orthogonal polynomial regression (正交多项式回归)

outcome indicators (输出指标)

outlier (异常值)

out of control (失控)

output requirements (输出要求)

overfitting (超拟合)

p , np , c , u control chart (p , np , c , u 控制图)

paired comparison design (配对比较设计)

paired-comparison test (配对数据检验)

Paired t -Test (配对 t 检验)

parameter design (参数设计)

Pareto effect plot (帕雷托效应图)

Pareto chart (排列图; 帕雷托图)

partial balanced incomplete block design (部分平衡不完全区组设计)

path of steepest ascent (最速前进方向)

PDPC (过程决策程序图法)

Pearson product moment correlation coefficient (皮尔逊相关系数)

percent precision to tolerance ratio, %P/T (精度/公差比率, 也称作 %P/T)

percent repeatability and reproducibility, %R&R (重复性与再现性百分比, %R&R)

physical mapping (物理映射)

pilot plan (试运行计划)

Plackett-Burman design (Plackett-Burman 设计)

planning matrix (计划矩阵, 实施矩阵)

point estimate (点估计)

Poisson distribution (泊松分布)

Poka-Yoke (防差错)

polynomial regression (多项式回归)

power (功效或检出力)

precision (精度; 精密度)

pre-control (预控制)

predetermined time system, PTS (预定时间制)

process control system (过程控制体系)

probability density function, pdf (概率密度函数)

problem/opportunity statement (问题/机会描述)

process (过程)

process analysis (过程分析)

process capability indices, PCI (过程能力指数)

process control plan (过程管理计划)

process failure mode and effect analysis, PFMEA (过程 FMEA)

process mapping (过程映射)

profit and loss (P and L) (损益表, 展示一段时期内净收益与亏损的会计报表)

program evaluation review technique, PERT (计划评审法)

project scope (项目范围)



Pugh matrix (普氏矩阵)
pull system (拉式系统)
push system (推式系统)
p-value (p-值)

qualitative analysis (定性分析)
quality control circle (质量管理小组)
quality function deployment, QFD (质量功能展开)
quality house (质量屋)
quality loss function (质量损失函数)
quality system standards (质量体系标准)
quantitative data (定量数据)

random words (随机词汇)
randomization (随机化)
randomized block design (随机区组设计)
range (极差)
ratio measurement scale (定比测量尺度)
rational subgrouping (合理子组化)
real-life simulation (模拟现实)
region of rejection (拒绝域)
regression and correlation analysis (回归和相关分析)
regression design (回归设计)
rejectable quality level, PQL (拒收质量水平)
repeatability (重复性)
repetition (相同试验条件的连续重复)
replicate (仿行, 相同试验条件下的非连续重复)
reproducibility (再现性)
requirement mapping (需求映射)
residual diagnostic (残差诊断)
resolution (分辨率)
response model (响应模型)
response modeling (响应建模)
response surface methodology, RSM (响应曲面方法)
response (响应变量)
return on investment, ROI (投资回报率)
return on net assets, RONA (净资产回报率)
risk priority number, RPN (风险度; 风险系数)
robust design (稳健设计; 健壮设计)
robust parameter design (稳健参数设计; 健壮参数设计)
rotatability (旋转性)
rolled throughput yield, RTY (流通合格率)
root cause (根本原因)



六西格玛管理

run chart (链图)
runs test (游程检验)

sample size (样本含量)
sampling (抽样)
scatter plots (散布图)
screening design (因子筛选设计)
service requirements (服务要求)
setting (设置)
severity (严重度)
short-run SPC (小批量生产统计过程控制技术)
sigma calculation (西格玛水平计算)
sigma level (西格玛水平)
signal factor (信号因子)
signal to noise ratio (信噪比; S/N 比)
signal-response system (信号—响应系统)
sign test (符号检验)
simple linear regression (一元简单线性回归)
simple response (简单响应)
simple response system (简单响应系统)
simplex centroid (单纯形重心法)
simplex lattice (单纯形格点法)
single array (单一表)
SIPOC analysis and map (SIPOC 分析和图)
six sigma design scorecard (六西格玛设计计分卡)
single minute exchange of die, SMED (快速换模)
SMART (specific, measurable, attainable, relevant, and time bound) criteria (SMART 准则, 代表具体、可衡量、可实现、相关性和时间限制)
Software Engineering Institute, SEI (卡内基梅隆大学的软件工程研究所, 他们提出了评价软件开发商的软件成熟度概念框架)
sources of variation (SOV) studies (变异源研究)
special cause (特殊原因)
specification limits (规格限)
spline regression (样条回归)
sponsor (资助者)
stability (稳定性)
stakeholders (利益相关方)
standard deviation (标准差)
standard work (标准作业)
static parameter design (静态参数设计)
statistical inference (统计推断)
statistical process control, SPC (统计过程控制)
statistical quality control, SQC (统计质量控制)



- statistics (统计学)
- steepest ascent search (最速上升法)
- stem-and-leaf (茎叶图)
- stratified random sampling (分层随机抽样)
- substance-field analysis (物体—场分析)
- sum of squares (平方和, 统计中指离差平方和)
- supplier-input-process-output-customer map, SIPOC (SIPOC图; 高端流程图; 宏观流程图)
- symmetrical orthogonal array (对称正交表)
- system design (系统设计)
- systematic process sampling (过程系统抽样)
- Taguchi method (田口方法)
- task-oriented roles (面向任务的角色)
- team charter (团队宪章, 或团队特许任务书)
- test statistic (检验统计量)
- TL 9000 (电信行业的质量管理标准)
- the effects of interactions between input's (输入变量间交互作用的影响)
- the effects of time relationships (时间关系效应)
- time study (时间研究)
- tips for innovation problem solving, TRIZ (创造性问题解决理论)
- theory of constraints, TOC (约束理论)
- tolerance (容差)
- tolerance design (容差设计)
- top-down chart (自上而下流程图)
- top-down design (自上向下的设计)
- total quality control, TQC (全面质量控制)
- total quality management, TQM (全面质量管理)
- total productive maintenance, TPM (全面生产维护)
- Toyota production system, TPS (丰田生产方式)
- treatment (处理)
- triads analysis (三元分析)
- TS 16949 (汽车制造行业执行的一套过程管理系统, 以保证产品的质量。它取代了汽车制造行业的众多国家标准)
- two-sided comparison (双侧比较)
- two-way ANOVA (两因素方差分析)
- two-way layout (双向分类)
- type I error (第I类错误)
- type II error (第II类错误)
- type of design of experiment (试验设计类型)
- unbiased (无偏的)
- uncontrolled factor (非可控因子)
- uniform precision (一致均匀精度)



upper control limit, UCL (上控制限)

value engineering/value analysis, VE/VA (价值工程与价值分析)

variance (方差)

variance component modeling (方差分量模型)

verification & validation testing (双 V 试验; V&V 试验)

vision (愿景)

voice of the business, VOB (公司要求)

voice of the customer, VOC (顾客声音)

Weibull distribution (威布尔分布)

Wilcoxon rank test (Wilcoxon 秩检验)

Wilcoxon signed rank test (Wilcoxon 符号秩检验)

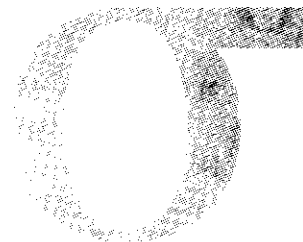
word (字)

wordlength (字长)

work study (工作研究)

X-bar and R control chart (均值—极差控制图)

X-bar and S control chart (均值—标准差控制图)



参考文献

1. Pyzdek, T. *The Complete Guide to Six Sigma QA Publishing*, L. L. C. 1999
2. Pyzdek, T. *Defining Six Sigma Projects*. Retrieved August 19, 2001
3. Rath, & Strong. (2000). *Rath & Strong's Six Sigma Pocket Guide*. Lexington, MA: Rath & Strong Management Consultants
4. Simon, K. *Sample SIPOC Diagram*. Retrieved August 19, 2001
5. 全国六西格玛管理推进工作委员会. 六西格玛管理评价准则. 北京: 中国标准出版社, 2007
6. 马林. 全国质量奖对提升我国企业国际竞争力的意义. 中国质量, 2003 (5)
7. [美] S·托马斯·福斯特著, 何桢译. 质量管理: 集成的方法. 北京: 中国人民大学出版社, 2005
8. [美] 弗兰克·M·格里纳著, 何桢主译. 质量策划与分析. 北京: 中国人民大学出版社, 2006
9. 上海质量管理科学研究院编著. 六西格玛核心教程: 黑带读本. 北京: 中国标准出版社, 2002
10. Berenson, *Basic Business Statistics*, Prentice-Hall, 1999
11. 上海质量管理科学研究院编著. 六西格玛自由论坛: 论文汇编. 北京: 中国标准出版社, 2002
12. 全国质量专业技术人员职业资格考试办公室组织编写. 质量专业理论与实务、质量专业综合知识 (中级). 北京: 中国人事出版社, 2003
13. [美] 潘德著, 王金德译. 六西格玛是什么. 北京: 中国财经易文出版社, 2002
14. [美] 朱兰学院著, 王金德译. 六西格玛基础教程. 北京: 中国财经易文出版社, 2002
15. [美] 潘德著, 王金德译. 六西格玛团队实战手册. 北京: 中国财经易文出版社, 2003
16. Box, G. E. P. and Draper, N. R. *Evolutionary Operation*. New York: John Wiley, 1969
17. Box, G. E. P., Hunter W. G and Hunter, J. S, *Statistics for Experiments*, New York: John Wiley, 1978
18. Box, G. E. P. and Draper, N. R. *Empirical Model Building and Response Surfaces*. John Wiley, New York, 1987
19. Montgomery, D. (1991). *Design and Analysis of Experiments*, John Wiley & Son, 3rd Edition (中译本: 实验设计与分析. 汪仁官等译. 北京: 中国统计出版社, 1998)
20. Wu, C. F. J. and Hamada, M (2000), *Experiments: Planning, Analysis and Parameter Design Optimization*. Wiley, New York (中译本: 试验设计与分析及参数优化. 张润楚等译. 北京: 中国统计出版社, 2003)
21. MINITAB User's Guide 2: Data Analysis and Quality Tools, Release 13 for Windows, 2000
22. 任露泉编著. 试验优化设计与分析. 长春: 吉林科学技术出版社, 2001
23. 焦叔斌等译. 朱兰质量手册 (第五版). 北京: 中国人民大学出版社, 2003
24. 上海质量管理科学研究院编著. 六西格玛核心教程黑带读本. 北京: 中国标准出版社, 2002
25. 茆诗松, 马逢时, 吴建福. 序贯淘汰水平法——一种利用正交表的搜索方法. 应用概率统计,



第6卷(2)期, 185-204

26. 何桢, 刘豹, 齐二石. 基于多变异分析的工序控制方法研究. 管理工程学报, 2002, Vol. 16, No. 2, 1-4
27. [日] 门田安弘著, 王瑞殊译. 新丰田生产方式. 石家庄: 河北大学出版社, 2001
28. [美] 杰弗里·莱克著, 李芳龄译. 丰田汽车案例 14 项原则. 北京: 中国财政经济出版社, 2005
29. [美] 詹姆斯·沃麦克, 丹尼尔·琼斯著, 沈希瑾等译. 改变世界的机器. 北京: 商务印书馆, 1999
30. [美] 詹姆斯·沃麦克, 丹尼尔·琼斯著, 沈希瑾等译. 精益思想. 北京: 商务印书馆, 1999
31. 何桢, 魏明, 周善忠. 我国企业推广六西格玛应注意的关键问题. 工程机械, 2004, 35 (12): 34-37
32. 何桢, 车建国. 精益六西格玛: 新竞争优势的来源. 天津大学学报(社科版), 2005, Vol. 7, No. 5, 321-325
33. 胡亚琴, 何桢. 论六西格玛的本质属性. 科学学与科学技术管理, 2004, 25 (10): 137-139
34. 何桢, 张志红. 精益与六西格玛的比较研究. 工业工程, Vol. 9 No. 1, 2006, 1-4
35. 何桢, 张志红. 基于并行质量工程的六西格玛公差设计方法应用研究. 中国机械工程, Vol. 16, No. 9, 2005, 827-830
36. 方海萍, 魏青江译. 精益六西格玛. 北京: 机械工业出版社, 2003
37. 肖智军等编著. 精益生产方式 JIT. 海口: 海天出版社, 2003
38. 魏大鹏编著. 丰田生产方式研究——准时化生产方式的技术支撑体系. 天津: 天津科学技术出版社, 2003
39. [美] 迈克·鲁斯, 约翰·舒克著, 胡师金等译. 学习观察——增加价值、消除浪费的价值流程图. 北京: 中国劳动社会保障出版社, 2005
40. 邵家骏. 质量功能展开(卓越质量文库, 六西格玛设计系列丛书之2). 北京: 机械工业出版社, 2004
41. [美] 乔杜里·苏著, 方海萍, 魏青江译. 六西格玛设计. 北京: 机械工业出版社, 2003
42. [美] 坦南特·杰著, 吴源俊等译. 六西格玛设计——新产品/新服务完美投放市场. 北京: 电子工业出版社, 2002
43. 邵家骏主编. 健壮设计手册. 北京: 国防工业出版社, 2002
44. Chapman, Ross L. and Hunt, Robert edited, *Proceedings of the World Innovation & Strategy Conference*, University of Western Sydney, Macarthur, 1998
45. Fey, Victor R. and Rivin, Eugene I., *the Science of Innovation*, TRIZ Group, 1997
46. King, Bob, *Better Design in Half the Time*, GOAL/QPC, 1989
47. Mizuno, Shigeru and Akao, Yoji, *QFD: The Customer-Driven Approach to Quality Planning and Deployment*, Asian Productivity Organization, 1994
48. Phadke, Madhav S., *Quality Engineering Using Robust Design*, Prentice Hall, Englewood Cliffs, New Jersey, 1989
49. Reference Manual, *Potential Failure Mode and Effects Analysis (FMEA)*, Chrysler Corporation, Ford Motor Company, General Motors Corporation, 1995
50. Suh, Nam Pyo, *Axiomatic Design: Advances and Applications*, Oxford University Press, 1999
51. Shao, Jiajun, *Research and Applications of Quality Function Deployment*, Proceedings of 2nd Sino-America Quality Conference, Beijing, 2001
52. QS 9000, "Quality System Requirements", Third Edition, 1998

53. [美] 波士, 艾伦和哈密尔顿公司编著, 王若松, 章国栋等译. 美国系统工程管理. 北京: 航空工业出版社, 1991
54. 韩之俊, 章渭基. 质量工程学. 北京: 科学出版社, 1991
55. 韩之俊. 三次设计. 北京: 机械工业出版社, 1992
56. [美] 尤列加, 威廉等著, 邵家骏译. 质量功能展开概论. 北京: 国防工业出版社, 1991
57. [日] 田口玄一著, 中国兵器工业质量管理协会译. 开发设计阶段的质量工程学. 北京: 兵器工业出版社, 1990
58. [日] 田口玄一著, 魏场禄译. 质量工程学概论. 北京: 中国对外翻译出版公司, 1985
59. 章渭基、韩之俊等. 质量工程学应用手册. 北京: 兵器出版社, 1993
60. 张性原主编. 设计质量工程. 北京: 航空工业出版社, 1999
61. 张宗斌, 邵家骏. QFD 技术在计算机软件改造中的应用. 世界标准化与质量管理, 1998 (10)
62. 邵家骏, 赵胜. 系统设计的基本原理和方法. 中国质量, 2001 (12)
63. 航天工业总公司. “三 F” 技术培训教材, 1995
64. GJB 768.1~768.3-89. 故障树分析, 1989
65. GJB 1391-92. 故障模式、影响及危害性分析程序, 1992
66. 陈晓川, 张暴暴, 冯辛安. 面向成本的设计的关键技术及其概念模型. 大连理工大学学报, 1999
67. 邵家骏. 开展六西格玛设计, 创造世界一流产品. 中国质量, 2005 (3)
68. 刘书庆, 杨水利编著. 质量管理学. 北京: 机械工业出版社, 2003
69. [美] 格雷格·布鲁著, 柴金平译. 六西格玛经理人. 北京: 机械工业出版社, 2003
70. [美] W·爱德华·戴明著, 钟汉清, 戴永久译. 戴明论质量管理. 海口: 海天出版社, 2003
71. [美] 马特·巴尼, 汤姆·麦卡锡著, 摩托罗拉大学译. 总裁读: 新六西格玛. 北京: 中国人民大学出版社, 2003
72. 马林, 罗国英. 全面质量管理基础知识. 北京: 中国经济出版社, 2001
73. 彭娟. 浅谈 GE 质量经营战略、六西格玛与质量经营. 联想质量探索, 2003 (1)
74. Thomas Pyzdek 著, 孙静译. 六西格玛手册. 北京: 清华大学出版社, 2000
75. 孙静主编. 最新国家标准 GB/T4091-2001《常规控制图》理解与实施. 北京: 中国标准出版社, 2002
76. [美] 罗纳德·D·斯尼, 罗格·W·何瑞尔著, 祝可嘉译. 领导者读六西格玛. 北京: 中国人民大学出版社, 2003

中国质量协会注册六西格玛黑带考试样题

(此样题基本按《六西格玛管理》的章节排列，题号后的中括号内所列数字即为章号。试题全部为选择题，有单选和多选两类。对于“多选题”，本样题中用“多选”加注，未加注者皆为单选题。)

1. [1] 确定项目选择及项目优先级主要是下列哪个角色的责任：
A. 黑带 B. 资深黑带 C. 绿带 D. 倡导者
2. [1] 在分析 \bar{X} -R 控制图时应：
A. 先分析 \bar{X} 图然后再分析 R 图 B. 先分析 R 图然后再分析 \bar{X} 图
C. \bar{X} 图和 R 图无关，应单独分析 D. 以上答案都不对
3. [1] 质量管理大师戴明先生在其著名的质量管理十四条中指出“停止依靠检验达成质量的做法”，这句话的含义是：
A. 企业雇用了太多的检验人员，对经营来说是不经济的
B. 质量是设计和生产出来的，不是检验出来的
C. 在大多数情况下，应该由操作人员自己来保证质量，而不是靠检验员保证
D. 人工检验的效率和准确率较低，依靠检验是不能保证质量的
4. [1] (多选) 六西格玛管理方法：
A. 起源于摩托罗拉，发展于通用电气等跨国公司
B. 其 DMAIC 改进模式与 PDCA 循环完全不同
C. 是对全面质量管理特别是质量改进理论的继承性新发展
D. 可以和质量管理小组 (QCC) 等改进方法，与 ISO 9001、卓越绩效模式等管理系统整合推进
5. [1] (多选) 推行六西格玛管理的目的是要
A. 将每百万出错机会缺陷数降低到 3.4
B. 提升企业核心竞争力
C. 追求零缺陷，降低劣质成本
D. 变革企业文化
6. [2] 在质量功能展开 (quality function deployment, QFD) 中，首要的工作是：
A. 客户竞争评估 B. 技术竞争评估 C. 确定客户需求 D. 评估设计特色
7. [2] 在某检验点，对 1 000 个某零件进行检验，每个零件上有 10 个缺陷机会，结果共发现 16 个零件不合格，合计 32 个缺陷，则 DPMO 为：

- A. 0.003 2 B. 3 200 C. 32 000 D. 1 600

8. [2] (多选) 顾客需求包括:

- A. 顾客及潜在顾客的需求 (VOC) B. 法规及安全标准需求
C. 竞争对手的顾客需求 D. 供货商的需求

9. [3] 哪种工具可以用于解决下述问题:

一项任务可以分解为许多作业, 这些作业相互依赖和相互制约, 团队希望把各项作业之间的这种依赖和制约关系清晰地表示出来, 并通过适当的分析找出影响进度的关键路径, 从而能进行统筹协调。

- A. PDPC (过程决策程序图) B. 箭条图 (网络图)
C. 甘特图 D. 关联图

10. [3] 下述团队行为标示着团队进入了哪个发展阶段?

团队的任务已为其成员所了解, 但他们对实现目标的最佳方法存在着分歧, 团队成员仍首先作为个体来思考, 并往往根据自己的经历做出决定。这些分歧可能引起团队内的争论甚至矛盾。

- A. 形成期 B. 震荡期 C. 规范期 D. 执行期

11. [3] 在界定阶段结束时, 下述哪些内容应当得以确定?

1. 项目目标
2. 项目预期的财务收益
3. 项目所涉及的主要过程
4. 项目团队成员

- A. 1 B. 1 和 4 C. 2 和 3 D. 1, 2, 3 和 4

12. [3] 在项目特许任务书 (team charter) 中, 需要陈述“经营情况” (business case, 也被称为项目背景)。该项内容是为了说明:

- A. 为什么要做该项目 B. 项目的目标
C. 公司概况 D. 问题产生的原因

13. [3] (多选) 亲和图 (affinity diagram) 可应用于以下场合:

- A. 选择最优方案 B. 用于归纳思想, 提出新的构思
C. 整理顾客需求 D. 评价最优方案

14. [4] 有关质量成本的描述, 正确的是:

- A. 质量成本包括内部损失和外部损失成本
B. 质量成本包括预防成本+鉴定成本+内部损失+外部损失成本
C. 质量越高, 质量成本也越高
D. 以上答案皆不对

15. [4] (多选) 以下哪些是一个好的项目问题陈述所共有的组成部分? 选择所有可能的回答:

- A. 问题对象描述具体 B. 有清楚的时间描述
C. 结果可测量 D. 含有解决方案

16. [4] (多选) 高端过程图 (SIPOC) 能令员工了解企业的宏观业务流程是由于:

- A. 它描述了每个详细流程 B. 它确认过程之顾客



C. 它确认过程之供方

D. 它阐明过程的结果

17. [5] 下表是一个分组样本:

分组区间	(35, 45]	(45, 55]	(55, 65]	(65, 75]
频数	3	8	7	2

则其样本均值 \bar{X} 近似为:

A. 50

B. 54

C. 62

D. 64

18. [5] 在某快餐店中午营业期间内, 每分钟顾客到来人数为平均值是 8 的泊松分布。若考虑每半分钟到来的顾客分布, 则此分布近似为:

A. 平均值是 8 的泊松分布

B. 平均值是 4 的泊松分布

C. 平均值是 2 的泊松分布

D. 分布类型将改变

19. [5] 一批产品分一、二、三级, 其中一级品是二级品的二倍, 三级品是二级品的一半, 若从该批产品中随机抽取一个, 此产品为二级品的概率是:

A. 1/3

B. 1/6

C. 1/7

D. 2/7

20. [5] (多选) 在测量系统分析计算重复性和再现性 (R&R) 时, 相对于极差法而言, 采用方差分析和方差估计法的优点是:

A. 计算简便

B. 可以估计交互作用的影响

C. 可以进行深层次的统计分析

D. 是精确算法, 计算结果没有误差

21. [5] 在钳工车间自动钻孔的过程中, 取 30 个钻孔结果分析, 其中心位置与规定中心点在水平方向的偏差值的平均值为 1 微米, 标准差为 8 微米。测量系统进行分析后发现重复性 (repeatability) 标准差为 3 微米, 再现性 (reproducibility) 标准差为 4 微米。从精确度/过程波动比 (R&R%) 的角度来分析, 可以得到结论:

A. 本测量系统从精确度/过程波动比 (R&R%) 来说是完全合格的

B. 本测量系统从精确度/过程波动比 (R&R%) 来说是勉强合格的

C. 本测量系统从精确度/过程波动比 (R&R%) 来说是不合格的

D. 上述数据不能得到精确度/过程波动比 (R&R%), 从而无法判断

22. [5] 对于正态分布的过程, 有关 C_p , C_{pk} 和缺陷率的说法, 正确的是:A. 根据 C_p 不能估计缺陷率, 根据 C_{pk} 才能估计缺陷率B. 根据 C_p 和 C_{pk} 才能估计缺陷率C. 缺陷率与 C_p 和 C_{pk} 无关

D. 以上说法都不对

23. [5] 对于一个稳定的分布为正态的生产过程, 计算出它的工序能力指数 $C_p = 1.65$, $C_{pk} = 0.92$ 。这时, 应该对生产过程作出下列判断:

A. 生产过程的均值偏离目标太远, 且过程的标准差太大

B. 生产过程的均值偏离目标太远, 过程的标准差尚可

C. 生产过程的均值偏离目标尚可, 但过程的标准差太大

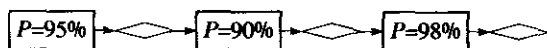
D. 对于生产过程的均值偏离目标情况及过程的标准差都不能做出判断

24. [5] 假定轴棒生产线上要对轴棒长度进行检测。假定轴棒长度的分布是对称的(不一定是正态分布), 分布中心与轴棒长度目标重合。对于 100 根轴棒, 将超过目标长度者记为“+”号, 将小于目标长度者记为“-”号。记 $N+$ 为出现正号个数总和, 则 $N+$ 的分布近似为:

- A. (40, 60) 间的均匀分布
- B. (45, 55) 间的均匀分布
- C. 均值为 50, 标准差为 10 的正态分布
- D. 均值为 50, 标准差为 5 的正态分布

25. [5] 某生产线有三道彼此独立的工序, 三道工序的合格率分别为: 95%, 90%, 98%。

如下图所示:



每道工序后有一检测点, 可检出前道工序的缺陷, 缺陷不可返修, 问此时整条线的初检合格率是多少?

- A. 90%
- B. 98%
- C. 83.79%
- D. 83%

26. [5] 对一批数据绘制直方图的结果显示, 其分布为右偏态分布, 则可以推断均值和中位数的关系是:

- A. 均值大于中位数
- B. 均值小于中位数
- C. 均值等于中位数
- D. 无法确定它们的关系

27. [5] 某瓷砖生产企业出厂检验结果显示, 在 1—3 月份生产的某类瓷砖中抽取了 1 000 块, 检验发现 200 个瑕疵, 瑕疵的出现是完全随机的, 则估计该类瓷砖的初检合格率是:

- A. 80%
- B. 85%
- C. 81.87%
- D. 90%

28. [6] 某药厂最近研制出一种新的降压药, 为了验证新的降压药是否有效, 实验可按如下方式进行: 选择若干名高血压病人进行实验, 并记录服药前后的血压值, 然后通过统计分析来验证该药是否有效。对于该问题, 应采用:

- A. 双样本均值相等性检验
- B. 配对均值检验
- C. F 检验
- D. 方差分析

29. [6] 为了判断 A 车间生产的垫片的变异性是否比 B 车间生产的垫片的变异性更小, 各抽取 25 个垫片后, 测量并记录了其厚度的数值, 发现两组数据都是正态分布。下面应该进行的是:

- A. 两样本 F 检验
- B. 两样本 T 检验
- C. 两样本配对差值的 T 检验
- D. 两样本 Mann-Whitney 秩和检验

30. [6] 为了判断两个变量间是否有相关关系, 抽取了 30 对观测数据。计算出了它们的样本相关系数为 0.65, 对于两变量间是否相关的判断应该是这样的:

- A. 由于样本相关系数小于 0.8, 所以二者不相关
- B. 由于样本相关系数大于 0.6, 所以二者相关
- C. 由于检验两个变量间是否有相关关系的样本相关系数的临界值与样本量大小有关,

所以要查样本相关系数表才能决定

- D. 由于相关系数并不能完全代表两个变量间是否有相关关系, 本例信息量不够, 不可能得出判定结果

31. [6] 响应变量 Y 与两个自变量 (原始数据) X_1 及 X_2 建立的回归方程为:

$$Y = 2.2 + 30.000X_1 + 0.0003X_2$$

由此方程可以得到结论是:

- A. X_1 对 Y 的影响比 X_2 对 Y 的影响要显著得多
- B. X_1 对 Y 的影响与 X_2 对 Y 的影响相同
- C. X_2 对 Y 的影响比 X_1 对 Y 的影响要显著得多
- D. 仅由此方程不能对 X_1 及 X_2 对 Y 影响大小做出判定

32. [6] 为了判断改革后的日产量是否比原来的 200 (千克) 有所提高, 抽取了 20 次日产量, 发现日产量平均值为 201 (千克)。对此可以得到判断:

- A. 只提高 1 千克, 产量的提高肯定是不显著的
- B. 日产量平均值为 201 (千克), 确实比原来 200 (千克) 有所提高
- C. 因为没有提供总体标准差的信息, 因而不可能做出判断
- D. 不必提供总体标准差的信息, 只要提供样本标准差的信息就可以做出判断

33. [6] 六西格玛团队分析了历史上本车间产量 (Y) 与温度 (X_1) 及反应时间 (X_2) 的记录。建立了 Y 对于 X_1 及 X_2 的线性回归方程, 并进行了 ANOVA、回归系数显著性检验、相关系数计算等, 证明我们选择的模型是有意义的, 各项回归系数也都是显著的。下面应该进行:

- A. 结束回归分析, 将选定的回归方程用于预报等
- B. 进行残差分析, 以确认数据与模型拟合得是否很好, 看能否进一步改进模型
- C. 进行响应曲面设计, 选择使产量达到最大的温度及反应时间
- D. 进行因子试验设计, 看是否还有其他变量也对产量有影响, 扩大因子选择的范围

34. [6] 回归方程 $\hat{Y} = 30 - X$ 中, Y 的误差的方差估计值为 9, 拟合回归方程的数据点超过 30 个, 且 $\bar{X} = 1$ 。当 $X = 1$ 时, Y 的 95% 的近似预测区间是:

- A. (23, 35) B. (24, 36) C. (20, 38) D. (21, 39)

35. [6] (多选) 在下列哪些情况中可以使用方差分析方法:

- A. 比较多个正态总体的均值是否相等
- B. 比较多个正态总体的方差是否相等
- C. 比较多个总体的分布类型是否相同
- D. 分解数据的总变异为若干有意义的分量

36. [7] 在 4 个因子 A, B, C, D 的全因子设计中, 增加了 3 个中心点的试验。分析试验结果, 用 MINITAB 软件计算, 其结果如下:

Factorial Fit: y versus A, B, C, D

Analysis of Variance for y (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	8.161 08	8.161 08	2.040 27	22.87	0.000

2-Way Interactions	6	0.676 59	0.676 59	0.112 76	1.26	0.369
Residual Error	8	0.713 61	0.713 61	0.089 20		
Curvature	1	0.025 58	0.025 58	0.025 58	0.26	0.626
Lack of Fit	5	0.404 63	0.404 63	0.080 93	0.57	0.735
Pure Error	2	0.283 40	0.283 40	0.141 70		
Total	18	9.551 27				

在正交试验中, 假定数据在拟合线性模型后, 试验数据的残差有共同的方差, 对于方差的估计量应该是 MSE (mean square error, 即平均误差均方和), 在本题中是:

- A. 0.089 20 B. 0.141 70 C. 0.713 61 D. 0.283 40

37. [7] 下列哪种响应曲面设计肯定不具有旋转性 (rotatability)?

- A. CCD (中心复合设计, central composite design)
B. CCI (中心复合有界设计, central composite inscribed design)
C. CCF (中心复合表面设计, central composite face-centered design)
D. BB (BB 设计, box-behnken design)

38. [7] 经过团队的头脑风暴确认, 影响过程的因子有 A, B, C, D, E 及 F 共 6 个, 且皆为连续性变量。其中除因子的主效应外, 还要考虑 3 个二阶交互效应 AB, AC 及 DF, 所有三阶以上交互作用可以忽略不计。由于试验成本较高, 限定不可能进行全面的重复试验, 但仍希望估计出随机误差以准确检验各因子显著性。在这种情况下, 应该选择进行:

- A. 全因子试验
B. 部分实施的二水平正交试验, 且增加若干中心点
C. 部分实施的二水平正交试验, 不增加中心点
D. Plackett-Burman 设计

39. [7] 在部分实施的因子试验设计中, 考虑了 A, B, C, D, E 及 F 共 6 个因子, 准备进行 16 次试验。在计算机提供的混杂别名结构表中, 看到有二阶交互作用效应 AB 与 CE 相混杂, 除此之外还有另一些二阶交互作用效应相混杂, 但未看到任何主效应与某二阶交互作用效应相混杂。此时可以断定本试验设计的分辨率是:

- A. 3 B. 4 C. 5 D. 6

40. [7] (多选) 对于响应曲面方法的正确叙述是:

- A. 响应曲面方法是试验设计方法中的一种
B. 响应曲面方法是在最优区域内建立响应变量与各自变量的二次回归方程
C. 响应曲面方法一般用于因子筛选
D. 响应曲面方法一般用于参数优化

41. [7] (多选) 在两水平因子试验时, 增加若干个中心点的优点是:

- A. 可以得到纯误差项
B. 检验模型的弯曲性
C. 使模型系数的估计更准确
D. 不破坏正交性和平衡性

42. [7] (多选) 在二水平全因子试验中, 通过统计分析发现因子 C 及交互作用 $A \times B$

是显著的, 而 A, B, D 均不显著, 则在选取最佳方案时, 应考虑:

- A. 找出因子 A 的最好水平
- B. 找出因子 C 的最好水平
- C. 找出因子 A 和 B 的最好水平搭配
- D. 找出因子 D 的最好水平

43. [8] 在计算出控制图的上下控制限后, 可以比较上下控制限与上下公差限的数值。这两个限制范围的关系是:

- A. 上下控制限的范围一定与上下公差限的范围相同
- B. 上下控制限的范围一定比上下公差限的范围宽
- C. 上下控制限的范围一定比上下公差限的范围窄
- D. 上下控制限的范围与上下公差限的范围一般不能比较

44. [8] 一位工程师每天收集 100~200 件产品, 每天抽样数不能保证相同, 准备监控每天不合格品数, 他应当使用以下哪种控制图?

- A. \bar{u}
- B. np
- C. c
- D. p

45. [8] (多选) $\bar{X}-R$ 控制图比 $X-MR$ (单值移动极差) 控制图应用更为普遍的原因在于:

- A. $\bar{X}-R$ 图可适用于非正态的过程
- B. $\bar{X}-R$ 有更高的检出力
- C. $\bar{X}-R$ 图作图更为简便
- D. $\bar{X}-R$ 图需要更少的样本含量

46. [8] (多选) 在 \bar{X} 图中, 下列情况可判为异常:

	UCL
A	
B	
C	CL
C	
B	
A	
	LCL

- A. 连续 3 点中有 2 点落在中心线同一侧的 B 区以外
- B. 连续 15 点落在中心线两侧的 C 区内
- C. 连续 9 点落在中心线同一侧
- D. 连续 4 点递增或递减

47. [9] 假设每次轮班可用时间为 7.5 小时, 30 分钟调整时间, 15 分钟计划停工时间, 15 分钟用于设备意外。请问设备的时间开动率为:

- A. 87%
- B. 93%
- C. 90%
- D. 85%

48. [9] 有关全面生产性维护 (TPM) 的描述, 不正确的是:

- A. TPM 应是团队工作来完成

- B. TPM 强调一线员工积极参与
C. TPM 的目的是消除因机器操作产生的故障、缺陷、浪费和损失
D. TPM 就是缩短故障维修时间
49. [9] 精益生产的核心理念是：
A. 加速流程
B. 精益求精
C. 准时交付
D. 消除一切不必要的浪费
50. [9] 在质量功能展开（QFD）中，质量屋的“屋顶”三角形表示：
A. 工程特征之间的相关性
B. 顾客需求之间的相关性
C. 工程特性的设计目标
D. 工程特征与顾客需求的相关性
51. [9]（多选）在下列项目中，属于防错设计（Poka-Yoke）的是：
A. 带有防盗器的汽车停车后，车主未锁车前，防盗器发出警报声
B. Word 文件编辑后退出 Word 时询问是否保存文件
C. 打印机卡纸后不工作
D. 微波炉在门打开时不工作
52. [10] 应用面向六西格玛的设计（DFSS）是因为：
A. 过程质量和产品质量受设计的影响，而六西格玛改进（DMAIC）的作用是有限的
B. 质量首先是设计出来的
C. DFSS 的方法可以替代 DMAIC
D. DFSS 是从源头抓起，及早消除质量隐患，从根本上解决问题
53. [10]（多选）进行 FMEA 分析时对于风险度大的故障模式必须：
A. 提供备件以便在出现该故障模式时更换
B. 规定在出现该故障模式时安排抢修
C. 采取设计和工艺的改进措施消除该故障模式或降低其风险度
D. 采取措施降低该故障模式的严重度、发生频率和检测难度
54. [10]（多选）关于四个阶段质量屋的正确表述是：
A. 四个阶段的质量屋应当于产品研发进行到各自阶段的时候分别建立
B. 质量功能展开的四个阶段可根据产品的规模和复杂程度等实际情况增加或减少
C. 四个阶段的质量屋是各自独立的，互相之间没有关系
D. 四个阶段的质量屋在产品规划阶段就应同步建立，以后不断进行迭代和完善
55. [10]（多选）关于面向制造和装配的设计（DFMA）的表述，正确的是：
A. 产品设计不必考虑企业现行的工艺及其设施
B. 简化设计、三化设计、互换性设计、防错设计、虚拟设计和虚拟制造等方法都是 DFMA 的方法
C. 产品设计早期就应考虑与制造、装配有关的约束和可能存在的问题，提高产品的可制造性和可装配性
D. DFMA 应贯彻并行工程的原则和采用团队工作的方法

答案:

1. D	2. B	3. B	4. ACD	5. BCD	6. C	7. B	8. ABC	9. B	10. B
11. D	12. A	13. BC	14. B	15. ABC	16. BCD	17. B	18. B	19. D	20. BC
21. C	22. B	23. B	24. D	25. C	26. A	27. C	28. B	29. A	30. C
31. D	32. D	33. B	34. A	35. AD	36. A	37. C	38. B	39. B	40. ABD
41. ABD	42. BC	43. D	44. D	45. AB	46. ABC	47. C	48. D	49. D	50. A
51. ABD	52. ABD	53. CD	54. BD	55. BCD					