

工商管理经典译丛·产品管理系列

Business Administration Classics · Product Management Series

World Class Quality:
Using Design of Experiments to Make it Happen

世界级质量管理工具

(第2版)

[美] 凯克·博特

(Keki R. Bhote)

著

阿迪·博特

(Adi K. Bhote)

遇今石柱 译

遇今 审校

中国人民大学出版社

我带着嫉妒、兴奋和希望的心情阅读了凯克·博特的惊人之作，嫉妒是因为他能将如此复杂的问题变成简单和易于理解的技术，我真是望尘莫及；兴奋是因为他强调了实际应用；希望是因为想改进质量的任何人都可以读这本书——这点非常重要。

——美国生产力和质量中心主席C. Jackson Grayson Jr.

在Maytag，我喜欢凯克·博特对质量工作的执着，他确立了大胆的质量改进目标，提供了简单、高效的分析工具，以识别和消除问题的产生根源。

——Maytag设备公司经理William L. Beer

不像许多质量培训辅助教材，凯克·博特是用简单、程式化、有趣的方式教我们掌握复杂和有效的工具。我非常感谢他在摩托罗拉，Honeywell，AlliedSignal和现在在Cessna做出的贡献。

——Cessna航空公司材料部副部长Mike Katzorke

凯克·博特是我们的顾问，试验设计对于综合工程师的理论知识与工人的实践经验是非常重要的。

——N. V. 菲律宾电器工业合作政策副董事长Willy Hendricks

本书第一版的意大利译著在意大利非常受欢迎。本书的再版更加详细，将进一步促进意大利公司正确应用试验设计。我必须感谢凯克·博特及他的儿子阿迪·博特做出的新的和重要的工作。

——Galgani公司总裁Alberto Galgani

《世界级质量管理工具》是一项非常实际的工作！它既不是关于基本原理，也不是关于培训，而是关于应用的结果！凯克·博特高效的21世纪工具箱将会使公司的绩效得到进一步的提高。这些工具非常简单和易于应用……

——Iomega公司高级副总裁Wayne Stewart

ISBN 7-300-05661-X



9 787300 056616 >

ISBN 7-300-05661-X/F·1780

定价：35.00元

工商管理经典译丛·产品管理系列
Business Administration Classics · Product Management Series

World Class Quality:
Using Design of Experiments to Make it Happen

世界级质量管理工具

[美] 凯克·博特

(第2版)

(Keki R. Bhole)

著

阿迪·博特

(Adi K. Bhole)

遇·今 石柱 译

遇·今 审校

 中国人民大学出版社



前 言

在计划经济时期，许多生活必需品都曾实行过配给制，而如今，人们可以随心所欲地选择自己所需要的商品和服务。经过 20 多年的改革开放，中国人亲身经历了告别短缺经济，进入过剩经济，以至国际商品涌入国内市场的巨变时代。中国企业在由计划经济向市场经济转轨的过程中，越来越体会到产品的开发与管理关系着企业今天的生存和明天的发展。随着经济全球化的浪潮涌来，我们的企业正面临着跨国公司和世界级产品参与竞争的严峻挑战。

据国内最大的本土管理咨询公司——和君创业研究咨询有限公司对数百家不同行业企业的调查研究表明，我国大多数企业在理



解和满足客户需求、制定产品开发战略、组织和配置战略资源，以及用不同产品适应市场竞争的变化等方面，普遍存在以下主要问题：

1. 大多数企业还未将产品开发与管理作为企业生存和发展的战略看待，企业缺乏基于战略的可持续发展产品规划。许多产品在昙花一现之后，很难具有持久的生命力。

2. 企业高层领导对产品管理关注和重视不够，长期依赖技术部门和少数技术人员响应和争夺市场，不能形成科学系统的研发体系与完善的营销体系间的协同配合，无法从制度上保证产品开发的成功率和持续性。

3. 多数企业对产品研发的投资力度不够，产品研发后劲不足。不能保证产品同时构思一代，开发一代，销售一代。

4. 由于产品研发的质量不高，生产流程设计欠合理，以及受供应商管理水平所限等原因，大多数企业的生产成本普遍较高，库存量较大，产品缺乏竞争力，不能有效引导市场并超越竞争对手。

企业间的竞争表现为产品的竞争，而产品竞争的实质是人才的竞争。企业产品可持续发展的动力和源泉，来自于企业家对市场走势的准确判断和把握，来自于企业中科学、领先的研发组织体系，来自于研发人员的先进理念和技能。视野决定思路，思路决定出路。在国际竞争日趋激烈，全世界都推崇“产品领先战略”，跨国公司纷纷投巨资建立产品研究院的今天，我们策划和出版本丛书，目的就是为那些每天思考竞争态势的中国企业家和管理者，以及从事产品研发的技术人员，提供基于战略的产品开发和管理的国际化前沿理念、实用技能和有效方法，使我国企业的产品开发和管理工作真正成为提高核心竞争力的有力武器。

本丛书的策划和出版得到了北京新华信管理顾问有限公司董事长赵民先生、中国人民大学商学院黄卫伟教授和陈俊副教授，以及许多企业界朋友的热情鼓励和支持；麦格劳-希尔教育集团、约翰·威利父子出版公司、美国AMA出版公司和哈佛商学院出版社等世界著名出版机构都为本书提供了版权，我们在此深表感谢。



需要说明的是，本丛书中涉及了大量未引进的英文产品品牌和名称，尽管我们在翻译和编辑的过程中尽了很大的努力，但由于译者和编辑的学识所限，不可能将这些品牌和产品的名称全部翻译到位，敬请读者原谅。但是，我们希望关心本丛书出版的读者们来信指点，帮助我们解决书中的问题，进一步提高丛书的出版质量。

中国人民大学出版社

2004年8月



译者前言

凯克·博特 (Keki R. Bhote) 是美国著名的质量管理专家。他曾经是世界上 380 家公司的质量管理高级顾问，并在国际质量管理界享有盛誉。特别是他在摩托罗拉公司做质量管理顾问时推广的先进的质量管理方法，使摩托罗拉公司的产品质量改进了 10 倍，甚至 100 倍，而且令他因此荣获了美国国家质量奖。

凯克·博特总结长期的工作实践经验，于 1991 年出版发行了《世界级质量管理工具》一书。该书在美国已经销售了十多万册，而且被翻译成四种欧洲语言出版。如今，应各国读者的要求，原出版作者又再版了该书。



长期以来,我国在开展质量管理工作时,由于一直缺少相应的质量管理工具的支持,影响了质量管理的实际效果。因此,闻洁工作室策划了《工商管理经典译丛·产品管理系列》,并委托我翻译其中的这本《世界级质量管理工具》,在翻译过程中,译者深感本书可为国内从事质量管理的技术人员提供有效的帮助。全书具有如下显著特点:

1. 全面、系统地介绍了 21 世纪的质量管理工具,理论联系实际,便于应用;
2. 应用案例极其丰富,所有重要的概念、方法和实施步骤都通过案例反复说明,避免了高深的数学理论,强调方法的实用性;
3. 尽量用简单的图表来说明概念、方法和应用的结果;
4. 每章都以提问的方式启发读者正确地使用质量管理的方法,可操作性强。

本书适用于企业质量管理人员、产品开发人员、质量监督部门人员、大专院校质量管理专业教师、高年级本科生和研究生,以及从事质量管理研究的其他人员。本书中文译本的面世,有助于先进的质量管理方法在我国的进一步推广应用,有助于我国产品质量的进一步提高,我们将感到莫大的欣慰。

本书的第 1 章至第 13 章由遇今翻译;第 14 章至第 24 章由石柱翻译。全书由遇今审校。由于译者水平有限,译文若有不妥之处,恳请读者批评指正。

遇 今

2002 年 5 月





英文版序言

凯克·博特 (Keki R. Bhote) 对质量工作的长期奉献精神鼓舞我推荐您拜读这本极具启发性和实用性的书。

《世界级质量管理工具》一书极其通俗易懂。它不负众望，适用于任何客户、买方和组织机构。

世上万物都在不停地变化，当然不仅仅是地理的变化，而凯克·博特教会了有关影响我们企业赢得客户忠诚度的首要的工作、目标、过程、输入、产出、服务和产品的各个方面的知识。

这些成就是我们公正、忠诚和自信的回报，同时又促进了世界经济的健康发展，带来了世界级的质量标准。



凯克·博特总结了为 21 世纪设计的 10 种质量管理工具，阐释了世界级质量的观点。甚至质量专业人员也不了解其中的大部分工具。这些工具以试验设计为中心，工具非常简单、费效比最佳和统计有效。在摩托罗拉公司，由他开发并应用了这些技术。通过实践的检验，质量改进了十倍，甚至百倍。

他用了大量案例和实践经验解释了这些技术，不仅是从设计到制造，而且还囊括了保障服务、农业、医院和大学的案例。

在试验过程中，他指出，试验设计不仅仅面向专业人员，还应推广到整个工厂，使工厂所有人员都能轻松地解决问题，让人们体会到 W·爱德华·戴明（W·Edward Deming）博士所称的“工作在轻松的气氛里”。

鲍勃·高尔文
(Bob Galvin)





第二版前言

《世界级质量管理工具》第1版于1991年出版。这些年来，该书已售出十万多册，并被翻译成四种欧洲语言，名列一些城市的畅销书排行榜。

重要的是我们收到了几百家公司的订单，客户们认为，在这本书中，他们找到了关于质量的信息，特别是试验设计。读过第1章后他们不再会对试验设计感到困难和困惑。他们赞扬这些解决问题的方法简单易行，不需要专业术语，可操作性强，费效比最佳。书中告诉他们如何解决长期的质量问题，其效果远比在日本广泛采用，又被许多西方公司盲目效仿的7种质量控制工具好得多。

本书介绍的方法比应用广泛的田口方法和西方工厂仍在采用的传统设计方法更容易、更省钱、更有效。我们的顾问将公司从应用田口方法和传统方法的困惑（使用这些方法困难和结果不理想）中解脱出来，令它们转向有效的解决问题的办法。

令人高兴的是，一些大学用《世界级质量管理工具》作为本科生和研究生的质量管理课程的教材。我们在南美、欧洲、亚洲和澳大利亚等地 350 多家公司举办了一至两天的学术会议，教授这些解决问题的高效技术。使用这些技术的人员不仅包括技术娴熟的工程师，而且也包括不想知道，也不需要知道那些毫无效果的复杂技术的技师、蓝领工人、供应商们。

那么为什么要再版本书呢？其理由如下：

- 在全世界客户的强烈要求下，我们增添了许多解决问题的案例，包括成功的经验和失败的教训（事实上，人们从失败中总结的经验要比从成功中获得的经验多）。

- 再版的另一个原因是增加了案例研究。这使得我们从课堂到实践更加容易。在本书的案例中（包括成功的和失败的案例），应用者会从中找到与解决自己具体问题类似的办法，即为他们解决问题提供了指南。

- 尽管通过阅读本书并模仿许多公司的实际案例可以学到应用试验设计（DOE）的方法，但是，专家顾问的指导仍是一个重要因子。若干年前的研究表明：要培养优秀的外科手术医生，实习医生必须在作为教练、教师的主刀外科医生的指导下，进行大量的反复实践。作者正是这种教练，为上百家采纳这种技术的公司提供了帮助。

- 我们为质量管理提供了容易理解和内容更丰富的试验设计技术工具箱。

- 书中将介绍称为多环境强化应力试验（MEOST）的技术。该技术类似于质量中的试验设计技术，它在可靠性方面取得了很大成就。将试验设计与 MEOST 相结合是一种有效的方式。在许多情况下仅靠这种结合便可解决长期的现场质量问题，并在设计的早期



避免问题发生。

- 我们增加了一章专门总结有关的技术，例如质量功能展开（QFD）、全面生产维修（TPM）、水平比较、防差错（Poka-Yoke）、客户管理（NOAC）、供应管理、价值工程（VE）和缩短生产周期。这些方法围绕着试验设计和 MEOST，构成了 21 世纪强有力的质量管理工具箱。

- 我们修改并完善了第 1 版中的试验设计技术，目的是使这些技术更好用。文中采取了提问的方式，以启发读者的思路，便于更好地解决实际问题。这里再一次强调教练的重要。

- 我们还将试验设计技术推广到服务业和管理领域。

- 本书能帮助公司取得 ISO 9000 和 QS 9000 认证，不仅仅是简单的质量实践，而是达到真正的世界级质量标准。

- 书中每章都附有小结，其中包括本章重点和参考文献。

- 对所有高层领导重要部门，新版本通过直接影响组织、团队、基线结果和质量文化创建而保证其领导核心作用。

凯克·博特
(Keki R. Bhote)



致谢

首先我最应该感谢的是我尊敬的母亲——著名的印度作家。她谆谆地告诫我：“当你长期艰苦地思考某一个问题，并从各个角度找到了解决的办法时，就应该把它写下来，这样你会感觉更好。”她的鼓励使我在九岁时便在杂志上发表了第一篇文章。我的每一次成功都归功于我的母亲。

在生活中第二个给我巨大鼓励的是摩托罗拉公司管理局主席鲍勃·高尔文，他是工业界的杰出代表之一。他那“走别人没有走过的路”的洞察力体现了其真正的领导才能。他将雇员看做助手，充分发挥他们的潜能，这些使我一生受益匪浅。

我还应感谢的是 Dorian Shainin，他是一位解决质量问题的专家。他的告诫“与部件对话，它们比工程师更聪明”是本书中许多试验设计的基础。他的观点一直持续到今日。我还要感谢我的两个出色的儿子，Peter 和 Dick，他们继承了父业。

我的儿子 Adi K. Bhote 是我的骄傲，他不仅继承了我的专业，而且还与我一起编著了本书。他的经历和实践使得试验设计 (DOE) 在学术价值上更加完善。

从某些方面来说，本书是我的学术文集，它是对我已故和受人尊敬的 Adolph Hitzelourger 和 Carlou Braun 两位同事的纪念。我们三个当中的这两位土耳其人，早在 40 年前年轻时起就从事试验设计研究。我们一起与许多公司的典型官僚主义和怀疑主义作斗争。我要感谢的还有慈父般的摩托罗拉大学校长，他为我们的试验设计培训奠定了基础。我还要感谢我的合作者 Dick Wendt, Rick Kiska, Bill Schmidt, John Lupinski 和 Thad Woronowicz，他们解决了许多难以解决的问题。

我要特别感谢一些国际企业的总经理们，是他们点亮了试验设



计的火炬，是他们成为试验设计的冠军。其中包括受人尊敬的 Oded Erlich（以色列），Shane Yemm（英国），Bernard Forgeot D'Arc（法国），Y. F. Tan（中国）和 T. T. Yew（马来西亚）。

自 1992 年从摩托罗拉退休，我就成立了自己的咨询公司。我们将 DOE 推广到 33 个国家，其中最杰出人物是飞利浦公司的 Willy Hendrickx 和 Frans Wouters，是他们将试验设计推广到 44 个飞利浦子公司。我们得到了许多人的支持，其中包括 Brian Boling（美国），Sid Dasgupta（印度），T. Van Der Hoek（荷兰），Mark Dale（英国），Chia Beng Chye（新加坡），Jeff Van Houdt（比利时）和 Alberto Chinen（巴西）。对他们为试验设计所做出的杰出贡献，我深表敬意。

在美国，要感谢的人不胜枚举，其中最著名的有 Mike Katzorke（希纳公司），Ted Taber 和 Carl Saunders（直升机公司），Wayne Stewart（奥米伽公司），J. C. Anderson 和 Jeff Postl（涡轮发动机公司），Bernie Sergesketter 和 Mary Rodino（美国电报电话公司），Chuck James（美国产品），Hemant Buch（第维可公司），Dr. C. Jackson Grayson（美国产品和质量中心），Harvey Kaylie（微电路公司），Bill Beer，Jeff Winter 和 Tony Hair（米塔公司），Doug Kilian（西门子公司）。

我非常感谢我的 350 家客户的大部分工作，是他们对试验设计技术的应用反复进行试验，不断完善。没有美国管理学会的高级编辑 Ray O'Connor 的鼓励和指导，本书今天恐怕还不能与读者见面。我还要感谢我的助手 Jean Seeley 是她打印了本书的全部手稿。

我衷心感谢我的家人——我的女儿 Safeena 和 Shenaya，我的儿子 Adi 和 Xerxes，他们对父亲充满了信心。最后我要感谢我的妻子 Mehroo，她是我写作的动力和源泉，她为本书从构思直到成稿的整个写作过程作出了许多牺牲。





目 录

第 1 篇 引言

- 第 1 章 试验设计的需求、目标与益处 (3)
- 第 2 章 50 年质量运动的风风雨雨 (11)
- 第 3 章 21 世纪的强大质量工具包 (18)

第 2 篇 试验设计绪论

- 第 4 章 工序能力的衡量 (47)
- 第 5 章 偏差:行业的通病 (63)



第 6 章	试验设计的三种方法:经典方法、田口方法与谢恩方法	(71)
第 7 章	10 种强有力的试验设计工具	(86)

第 3 篇 “与部件对话”——组线索生成工具

第 8 章	多变量分析:自动寻找红 X	(108)
第 9 章	集中图:重复出现问题的精确定位	(150)
第 10 章	部件搜索:简便而平滑的交换	(160)
第 11 章	成对比较:一种精巧而通用的工具	(200)
第 12 章	产品/过程搜索:精确定位过程变量	(226)

第 4 篇 表示产品/过程特性试验技术有效设计

第 13 章	变量搜索法:罗尔斯·罗伊斯对红 X 的研究	(243)
第 14 章	全析因:识别与量化每个交互影响的最佳技术	(295)
第 15 章	B vs. C:卓越的确证技术	(323)

第 5 篇 试验设计优化

第 16 章	获得现实的技术规格和公差散布图	(354)
第 17 章	响应曲面方法(RSM):交互作用的优化	(376)



第6篇 从试验设计到统计过程控制的转换:重要的学科

第18章 正向控制:冻结过程增益	(392)

第19章 过程确认:消除不良质量的外部原因	(404)

第7篇 设计过程控制:用于监控质量

第20章 控制图:一项过时的技术	(422)

第21章 预先控制:一项全新的技术	(429)

第8篇 试验设计与可靠性的联系

第22章 多环境强化应力试验:对可靠性的有效性如同试验设计对质量的有效性一样	(455)

第9篇 设计试验工具使用的逻辑顺序

第23章 解决长期性问题的序贯试验设计工具的案例研究	(481)

第10篇 从课堂教育到实践中学习

第24章 实践出真知	(507)

参考文献	(520)



第 1 篇

引 言



第1章

试验设计的需求、目标与益处

► 1. 需求

中国农民以前杀蝗虫的方式

有一个发生在中国“文化大革命”时期的故事。某村庄的农田遭受了蝗灾，眼见庄稼就要颗粒无收，村民们想从毛主席语录中寻找攻克这场灾难的答案，但是只有一句看上去最为合适。毛主席说，人民在没有方针指引的情况下，应该自己解决问题！于是，全村男女老少齐动员，一起到田间徒手将蝗虫一个一个地除掉。数百村民艰辛地劳作了七天，才把蝗虫全部消灭。若同样的事发生



在今天的美国，则只消一小时，撒上几袋农药便能解决问题。

亨利法则

上面的例子说明了工具的重要性——正确的工具。许多出发点很好的解决问题的团队如质量循环、小组改进活动总是出现问题，就是由于没有使用正确的工具。此原理称为亨利法则。^[1]按照该法则，团队的主要特性是以指数的形式增长的；如果忽视或误解了主要特性，那么这种忽视或误解也会呈指数形式增长。如果主要特性是知识和较好的信息，那么它们也会以指数形式增长。即使公司有好的领导和乐于奉献的热心员工，但没有好的工具，也只会使人感到困惑、混乱和失去信心。

对试验设计（特别是谢恩法）的需求是迫切的，范围也是广泛的

1. 美国工厂中 90% 不知道如何解决长期性的质量问题。这种说法不是针对所有问题而言的。有些最简单的问题能用质量控制（QC）的 7 种工具、工程判断、头脑风暴法、统计过程控制或凯普纳—特拉格（Kepner Trago）方法解决。但是，也有这些传统工具无法解决的长期的质量问题，即已存在了数天、数星期、数月甚至数年的质量问题。本书详细介绍的谢恩（Shainin）试验设计（design of experiments, DOE）正好能解决这些问题。

2. 公司利用的是生产线上工人的体力而不是其智慧。利用试验设计的第二个迫切原因是其应用的广泛性：所有工人，包括计时工人、维修工人、技师、工程师和专家都可以使用。在某些公司中，只安排工程师去解决问题，对工人则仅利用其体力，而将其智慧拒之门外，这是巨大的浪费。在咨询了几百家公司后，我们得出这样的结论：领班工人的智商与工程师一样高，他们只不过没有工程师那样的崭露头角的机会而已。如果将 DOE 技术的简单工具传

授给工人，予以培训、鼓励和指导，他们会做得特别好。我们曾看到过许多工厂在解决质量问题后人人轻松的局面。

3. 90%的技术规范和公差是错的。工程师对技术规范和公差的制定有脱离实际的倾向。其中一些规范和公差是基于过去的图纸、供应商的推荐、传统的规则和类似的信息而制定，很少使用最坏情况下的公差或几何公差。然而，几乎所有的工程师都忽视了利用 DOE 来确定实际的技术规范和实际的公差，而这样做可以取得产品质量方面的突破，避免了产品延期交付，并节省大量费用。

4. 不知道对产品/工艺参数进行优化。在产品或工艺的设计阶段，工程师通过公式、计算机模拟、电路分析、供应商信息或者仅凭经验来猜测重要的参数。如果猜测是正确的，那么就会生产出百分之百的生产率。但是这种情况极少发生。从不重要的参数中挑选出重要参数的最好工具就是试验设计。利用这种方法，重要参数的公差被用于保持高产，而非重要参数公差则用于节省费用——甚至节省一半。

5. 认为缺陷和偏差是不可避免的。工厂与缺陷为伴的生活如此之长，以至于它们认为缺陷是生活中的一部分，就像呼吸一样。分类和筛选、检测和纠错已成为惯例。这就导致了产品费用的增加，交给客户的只能是高价产品。类似地，偏差也被视为是不可避免的。“既然自然界有偏差，产品为什么就没有理由有偏差呢？”这样的看法成了合理的推断。

过去，我们没有消除产品和工艺的缺陷并减少其偏差的工具。现在，利用 DOE 就有能力把缺陷从百分之几降低到百万分之几，甚至十亿分之几；并且通过减少客户返修、摩擦、分析、重复工作、检验和测试，我们可以向目标价值推进，继而节省大量的费用，而且偏差可减少到工序能力指数 C_{pk} 为 2.0，甚至更好（见第 4 章）。

6. 强化可靠性试验只是对不可靠的互换。可靠性也许比质量更重要，因为它有两个附加的条件——时间和应力。传统的可靠性设计利用的工具有可靠性预计研究、故障模式影响分析（FMEA）、故障树分析（FTA）。这些工具对产品的可靠性设计就像质量管理的 7 种工具对质量管理一样有效。换句话说，最多不过是勉强有



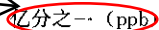
效。美国国防部长期利用大量的寿命试验来改进可靠性，正像工厂利用大量的检验来改进质量一样，但结果是没有用！甚至到今天，耗费在外场维修、保障军事装备方面的成本竟是其采办费的 11 倍！现在已有一种较好的办法，即在设计阶段，通过综合利用 DOE 和 MEOST，能够模拟并防止外场失效，从而达到极大地改进外场可靠性的级别的目的。

B. 目标

本书的目标是：

(1) 用简单、容易和切合实际的非数学语言描述解决长期质量问题的有效的统计技术，这些问题用传统的方法又是无法解决的。这些技术应用领域极为广泛，范围涵盖从莒莒的加工到核能的利用，从微观的半导体晶片到宏观的结构设备。

(2) 适用于公司内各级人员——从工程师、技术员到主管和生产线操作者，从经理到客户和供应商——都可使用这些工具。

 (3) 从当前不可接受的高缺陷水平，例如百分之一的可接受质量水平 (AQLs)，降低到百万分之一 (ppm)、 亿分之一 (ppb) 甚至零缺陷，最终实现零偏差目标。

(4) 在产品设计阶段和工艺设计阶段使用 DOE 可以避免把质量问题带到产品的生产或使用现场中。若产品的生产过程存在问题，这些问题就是工程设计失败的标志！这样的设计缺陷是所有质量问题的最大的隐患。

(5) 已经证明，统计过程控制 (SPC) 并不是解决问题的工具，而仅是监控和维护工具。而且，在 SPC 领域，仍广泛用于西方的控制图已经过时，其工作量不但繁重且耗费大量资金，在统计方面也有缺陷。而预控制图可以说是具有生命力的技术。它比控制图简单，费效比更佳，且在统计方面远比控制图更充分。

(6) 总结 21 世纪的 10 种有效工具，并说明它们是如何共同影



响世界级质量和产品生产力的。

(7) 将基本的可靠性技术（例如 FMEA 和强化应力试验）提升至完美的可靠性技术（DOE 和 MEOST 的综合应用），能够使可靠性的改进提高一到两个数量级。

(8) 通过加入案例研究，包括成功的经验和失败的教训，同时为读者提供一些能用于实践的练习，加强读者使用解决问题的正确方法，从而避免一开始就出错。

(9) 提供一份与每一项 DOE 技术相关的清单，其目的是：

- 避免出错；
- 为解决问题的团队提供详细的指南；
- 给出管理的通用指南。

(10) 强调为整个公司内提升质量文化、管理支持、涉及和承诺的所有重要任务。

C. 益处

北美、欧洲、亚洲和澳大利亚的 200 多万人的实践已经证明了这些技术的益处，特别是我们在 33 个国家中的 350 多个客户尤为受益。这些益处主要表现在：

(1) 质量改进从 2 : 1 达到 10 000 : 1，且 1 天时间相当于过去的 6 个月时间。

(2) C_{pk} 从 0.5 和 1.0 分别增加到 2.0 和 5.0。

(3) 在检测和试验过程中，几乎消除了产品的擦伤、返工和降级，向零缺陷和 100% 的生产率稳步前进。

(4) 质量差的产品造成的浪费迅速下降了一个数量级，而此前其占产品销售额的 10%~20%，每天每人 100 美元~200 美元。

(5) 外场可靠性增长 9 倍，甚至 99 倍。

(6) 提高了客户满意度、客户忠诚度，并加强了与客户的长期联系。



(7) 从解决生产中的问题转向在产品和工艺设计阶段预防问题。

(8) 延长了设备的正常工作时间，增加了产量，明显改善了全面生产维修（TPM）结构的有效性，使工厂总有效率（FOE）从小于 50% 提高到大于 90%。

(9) 生产周期时间缩短到原来的 1/5。

(10) 调动了员工的积极性，激发了员工的热情。避免了因挫折造成的士气低落，放弃了较差的工具。

(11) 对供应商来说，DOE 进一步的益处在于改进了质量，降低了成本和缩短了周期。从客户方面来说，增加了客户，加强了客户对供应商的理解。

(12) 在企业方面则改善了企业的利润情况、投资回报率、市场占有率和总的生产率。

图 1—1 给出了 DOE^[2] 在 12 个重要方面对企业的积极影响。

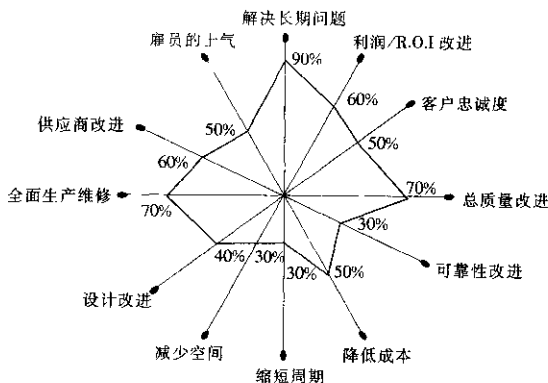


图 1—1 DOE 对企业的积极影响：蜘蛛图

注：蜘蛛图显示出特定技术对若干参数的相对影响。图中，每项分支的总长代表 100%；百分数表示 DOE 技术对相应参数的积极影响。



1. 解决长期问题 (90%)。DOE 是惟一能解决长期质量问题的工具，这些问题很长时间以来一直无法解决。

2. 利润/R.O.I. 改进 (60%)。在一般的公司中，不良质量成本 (COPQ) 使产品销售额缩减 10%~20%。如果能将该成本削减一半，那么公司的利润将会翻一番。DOE 技术对这种改进的影响率为 60%。

3. 客户忠诚度 (50%)。客户的忠诚度 (维系) 超出了客户的满意度。对一般公司来说，客户与产品的维系增加 5%，就能使公司利润增加 35%~120%。通过注重产品质量、满足客户要求，而不仅仅是为了质量而追求质量，DOE 能够增进与客户的联系。

4. 总质量改进 (70%)。在所有针对质量改进的技术中，只有 DOE 能够实现突破性的质量改进。

5. 可靠性改进 (30%)。MEOST 与 DOE 的组合能够大大改进可靠性。

6. 降低成本 (50%)。用 DOE 技术实现 100% 的产出是降低成本的最佳方式之一。

7. 缩短周期 (30%)。DOE 是缩短周期的重要技术，特别是在制造业。

8. 缩减空间 (30%)。利用 DOE 解决所有的质量瓶颈，可以缩减维修所需空间，减少检验所需场地，并降低库存。

9. 设计改进 (40%)。在设计阶段使用 DOE 是惟一能确保防止质量问题进入产品生产过程和使用现场的方法。

10. 全面生产维修 (70%)。作为 TPM 的一种度量，工厂总效率 (FOE) 是产出的百分比、正常工作时间百分比与设备效率百分比的乘积。DOE 对产出百分比和正常工作时间百分比影响很大，而对设备效率的影响很小。

11. 供应商的改进 (60%)。伙伴供应商的开发取决于供应商实现对客户公司的质量、成本和周期的改进的能力。DOE 是改进这三方面的关键。

12. 雇员的士气 (50%)。提供给工人较差的解决问题的工具



是对其士气的最大挫伤。应该为工人提供有关使用正确的工具去解决其问题的知识。DOE 的简单性和有效性往往使车间重新焕发生机。

在深入研究质量改进工具之前，我们先介绍一下质量改进那充满挫折的曲折历史。

【注释】

[1] 以工业哲学家韦斯·亨利（Wes Henley）的名字命名。

[2] 本书所谓的 DOE 是指谢恩的 DOE 技术，而不是指效果稍逊的传统的 DOE 技术或出口试验（见第 6 章）。





第2章

50 年质量运动的风风雨雨

质量工具的潮起潮落

20 世纪下半叶，质量运动以十足的摩登模式——每隔十年变换一次潮流风向，不断寻求着质量工具，新技术层出不穷。

● **抽样。**20 世纪 50 年代，风行的是各式各样的抽样——接收质量水平 (AQLs)、平均超出质量水平 (AOQLs) 和批容差百分缺陷 (LTPDs) 的抽样技术。后来，人们开发了许多这样的抽样技术，每种抽样技术都围绕着质量控制中的 α 和 Ω 作文章。（作者后来在某重点大学教授一项 16 学时的质量控制课程时惊异地发现，竟有 14 学时贡献给了各式各样的抽样。）幸好，抽样在其



流行过程中渐渐失色。

● **零缺陷运动**。20 世纪 60 年代则风行零缺陷运动。苏联的史泼尼克号 (Sputnik) 横空出世并成功的发射到太空, 而美国的发射却连连在太平洋受挫, 这一切迫使美国国防部不得不支持零缺陷运动。它强调的是只要工人们保证零缺陷, 那么质量就有保证。某公司曾发生这样一件事, 公司要求所有的工人签订零缺陷保证书, 但有一名工人拒绝签字。他的工头恳求他签, 结果不行, 然后他的经理又一次劝说也未见效。最后他被拖到了总经理办公室, 总经理对他说: “约翰, 你要么签字, 要么被解雇。” 约翰只好签了字。总经理又问他为什么不愿意签, 约翰回答说: “总经理先生, 到目前为止, 没人能说清楚签了保证书的好处。” 最后, 零缺陷运动演变成为一场集体“秀”, 而没有任何实质成果。

● **ISO 9000**。^[1]1987 年许多国家兴起了 ISO 9000, 它充其量不过是非常基本的质量体系——由 45 个国家起草的最小的通用标准。它那隐含着约束的规程带来的是繁文缛节的负担, 它只是合同的质量水平而不是不断改进的方式。它很少关注消费者, 其目标只是“冻结”现有水平的缺陷, 而不是消除它们。正如有人所说: “如果你要造废纸, 那么 ISO 9000 会自始至终地帮助你完成!” 如今它已成为逼迫企业就范的关卡, 成为与其他具有同样的遵从压力的供应商做生意的护照。然而, 至今几乎没有任何公司从 ISO 9000 中看到效益底线。事实上, 它几乎使质量运动倒退了 20 年。

● **马尔科姆·鲍德里奇奖**。1988 年, 美国国会为了迎接全球的质量挑战, 特别是日本戴明奖的挑战, 设立了马尔科姆·鲍德里奇 (Malcolm Baldrige) 质量奖。其指南远胜于实施 ISO 9000 企业的规程, 但它与世界级质量体系仍然相距甚远。其用辞也是含糊不清、令人迷惑——即使在质量专业人士看来也是这样, 结果让试图使用它的公司兴趣索然, 而让某些外行的顾问成了富翁。它的确没有充分注意到那些能够将高目标变成结果的有效工具。它在供应商、客户忠诚度 (以区别客户满意度)、领导力 (只注重上层领导) 和基层企业方面做得不够充分。进一步地讲, 因为鲍德里奇奖对公

Quality Department
2006.05.18
SongYongSheng



司来讲是随意的,而 ISO 9000 正演变成带有强制色彩的,特别是对那些想与欧洲共同体国家做生意的公司,所以申请该奖的公司越来越少,与几年前相比已下降了 50%。

● **欧洲质量奖。**1990 年,欧洲共同体设立了欧洲质量奖。它以马尔科姆·鲍德里奇奖作为模板,并进行了一些结构性和装饰性的改进,其中最重要的改进是强调结果。

● **QS 9000。**QS 9000 是美国三大汽车公司为了统一其对供应商所要求的质量体系而于 1995 年设立的。其目的是改进 ISO 9000,强调减少偏差,最小 C_p 为 1.33,质量体系审核与产品质量紧密联系。类似于鲍德里奇奖,其缺点是用辞含混、语气专断;工具细节描述不充分;领导力管理及雇员激励方面是弱项;在与供应商合作中不能双赢。供应商们不得不投入巨大成本,而带来的费效比却令人怀疑。许多供应商抱怨,三大公司只是 QS 9000 的传教士而非它的实践者,如果在他们的公司应用 QS 9000 注定会失败。

● **全面质量管理。**20 世纪 80 年代和 90 年代,全面质量管理(TQM)成为花样迭出的质量运动的新宠。在 15 年前其诞生之际,它就被预言为一个大救星——最后的解决办法。然而,它也迫使人们不得不放弃了幻想。表 2—1 所示为一些权威组织最近对全面质量管理效果的调查结果。

表 2—1 全面质量管理效果的调查结果

调查者	
美国电子学会 ¹	73%的公司开展了全面质量管理,其中 63%的公司减少了小于 10%的质量缺陷。
Arthur Little ²	仅有 1/3 的公司认为全面质量管理提升了它们的竞争力。
麦肯锡公司	67%的公司没有获得真正的改进或者深陷泥潭。
美国质量基金会 ³	77%的公司没有达到预期效果和企业目标。
质量杂志 ⁴	“公司组织了数千名雇员开展各种各样的质量活动,然而,结果却非常令人失望。除了良好的愿望外,这些公司弄得一团糟。它们认识到,以革命式的速度进行一年一度的质量改进的需求,并不能达到它们预期的目标。”

* 在本书中的上角标的标号均指本书参考文献的序号。



有效的质量体系

6 σ 。1987年，摩托罗拉引进了著名的6 σ 过程^[2]作为 C_p 和 C_{pk} 指数的备选（见第4章）。它的主要目的是在公司的“全部客户都满意”的要求下，达到迄今为止达不到的质量水平高度。它变成了所有雇员一起与缺陷作斗争的统一标志。

摩托罗拉的6 σ 的成功使得通用电气、ABB、Allied Signal和宝丽来等公司产生了许多“6 σ 族”计划，甚至美国质量协会也想“克隆”省钱的“6 σ 族”计划，因为这些计划需要大量的投资——上百万美元的外部咨询费和上千万美元的内部成本。而这些公司声称，其质量改进的投资回报为2:1到4:1。然而，业内人士怀疑，其中大部分投资是否做到了有的放矢。不管怎样，这些“6 σ 族”计划在大胆想像、实施范围和结果方面尚存在不足。

终极6 σ ——“大Q”。即使是摩托罗拉的6 σ 过程也需要在许多重要方面进行改进。当我们赢得了马尔科姆·鲍德里奇国家质量奖后，我们认为，若只停留在荣誉面前，就会停滞不前。我开始研究和开发真正世界级质量体系，远远超过全面质量管理和摩托罗拉的6 σ ，⁵而后者受到了狭义的产品质量（例如设计和制造）的限制。广义地说，在企业的责任和质量之间应该是有联系的。为了质量而追求质量是徒劳的。为了客户忠诚度和利润而追求质量是没有错误的。我的最终6 σ 体系注重企业全体员工的聪明才智，即以“大Q”来取代所有其他质量体系表示的产品质量的“小Q”。在一些明智的公司中，“大Q”体系已经变成了为企业走向辉煌而指明道路的“圣经”。

表2-2比较了各种质量体系相对的强势和弱点，并分别给出了有效性评级（1~100），其中，1是最无效的，而100是最有效的。

表2-3则比较了某公司在12个关键领域中传统的全面质量管理、“6 σ 族”计划和我的“大Q”。



表 2—2

各种质量体系的相对有效性

质量体系	优点	缺点	有效性
• ISO 9000	• 现在成为一项国际标准 • 对大多数公司来说是强制性的 • 是通向“欧洲堡垒”的护照	• 繁文缛节, 其规程是在可以消除缺陷前冻结它们 • 体现了合同质量水平而非连续的改进水平 • 未强调领导力、组织、工具、现场、支持服务和人员这些方面, 不适用于基层	5
• QS 9000	• 美国“三大”汽车公司规范的统一质量要求 • 减少偏差	• 前进一步的 C_{pk} 为 1.33, 但是不充分 • 专断, 冗长 • 三大公司不能满足它们自己的 QS 9000 标准 • 不强调领导力、雇员、组织工具和支持服务这些方面	10
• 马尔科姆·鲍德里奇奖指南	• 强调领导力和客户这些方面	• 带有强制色彩的 ISO 9000 的一致性遮掩这种自愿但出色的系统 • 内容含糊不清而混乱 • 以牺牲企业各个方面的优秀性为代价追求狭义的质量	25
• 欧洲质量奖	• 基于马尔科姆·鲍德里奇奖设立 • 强调结果	• 与马尔科姆·鲍德里奇奖的缺点相同, 但程度稍弱	30
• 全面质量管理 • 6 σ	• 强调管理 • 力达目标 • 在摩托罗拉收到了很好的效果	• 许多效仿它的公司还没有学到体系的真谛	50
• 终极 6 σ ——大 Q	• 强调企业所有方面的优秀性, 重于狭义的产品质量 • 与世界级公司的规章相似	• 新引进 • 内容在准备中, 但还未出版	90

注: 有效性范围, 1=最低; 100=最高。



表 2—3

传统的全面质量管理、6 σ 族和大 Q 的比较

方面	传统的全面质量管理	6 σ 族——“小 Q”	终极 6 σ ——“大 Q”
1. 客户	客户满意	客户满意	客户、雇员和投资商的忠诚度
2. 管理	不强调领导力	含糊地提到领导力	鼓励领导帮助雇员充分发挥其潜力
3. 雇员	泰勒主义，因担心驱动，辛苦工作	有黑带专家其他人被动	切实放权，自由，工作场所气氛愉快，充分调动积极性
4. 组织	垂直架构的大金字塔	部门自定规范	平金字塔，共同组建
5. 体系	ISO 9000 QS 9000	提升鲍德里奇奖意义上的质量意识	世界级体系，企业取得超常的价值
6. 工具	7QC 工具 PDCA, 8-D	经典的试验设计	谢恩 DOE, MEOST, QFD, TPM, NOAC, VE, 循环时间
7. 设计	隔离设计	计算机模拟	设计时间、缺陷量、费用和人力减半
8. 供应商	独断，远程控制关系	无政策指导	基于相互信任、热诚而积极帮助的双赢的伙伴关系
9. 制造	SPC 和控制图	长达四个月的昂贵的黑带培训	解决问题容易，不仅仅是黑带问题
10. 现场	克服产品缺陷的服务合同	FMEA 族	MEOST, 内在的诊断技术
11. 支持服务	静止的商业过程	没有量化质量	客户管理，流程图；关键度量周期时间，避免空想
12. 测量/结果	仅为质量而改进质量	鼓励质量投资回报从 2:1 到 4:1	质量投资回报超过 10:1；利润改进最少为 2:1



【注释】

[1] ISO 9000、马尔科姆·鲍德里奇奖、QS 9000 和欧洲质量奖都定期加以修订。有些明显的不足之处在修订时得到了纠正。

[2] 统计意义上的 6σ 与摩托罗拉的 6σ 有很大的差别。统计上的 6σ 是指每十亿 (ppb) 个部件有两个部件缺陷水平，而摩托罗拉的 6σ 是假设非中心的 1.5δ 每百万个部件 (ppm) 有 3~4 个缺陷水平。摩托罗拉通过加上整条生产线从开始到结束的总数提高了缺陷数，但是又通过除以这些单元的总缺陷数乘以该单元部件数，再乘以每个部件的缺陷数降低了具有挑战的 ppm 水平。这样做的原因是公平地与其他的产品、生产线或工厂作比较，而不考虑产品的复杂度，这一过程被称为“规范化”过程。





第 3 章

21 世纪的强大质量工具包

随着全面质量运动的不断推进，质量工具也经历了交替变换的历史过程，例如 QC 的 7 种工具、7 种质量管理工具、福特的解决质量问题的 8D 方案和工程方法。本章我们描述 21 世纪应用的 10 种强大工具及其突破性的成果。⁷

基本工具

QC 的 7 种工具——幼儿园工具

将近 30 年来，西方国家一直盲目地照

搬日本人总结和包装的一系列质量技术——QC 的 7 种工具，以便解决产品的质量。表 3—1 分别列出了这些工具的目标、方法及其应用的时机和场合。公平地说，日本在培训其应用这些工具方面取得了成功。通过其质量循环、持续改善小组（Kaizen）和雇员的建议，利用这些工具处理每天的问题，从而动员了工厂的所有工人而不仅仅是专家。但是 QC 的 7 种工具仅能解决最基本的质量问题。对于解决长期的质量问题，它们是完全没有用的，因此，这些方法在西方不具有吸引力。

表 3—1

基本 SPC 工具

工具	目标	方法	使用 时机	典型 使用者
1. PDCA (计划， 做，检查， 采取措施)	通过试验和 修正解决问 题	计划工作；执行；检查结果；如果在期望的目标与实际结果之间有偏差，则采取措施。不断循环上述步骤直到偏差降到零为止	当不知 道更多 的有效 工具时	大多 数生 产线 上的 工人
2. 数据收 集和分析	• 评估质量 • 控制产品 • 规范过程 • 接收/拒收 产品 • 整理观察 结果	定义收集数据的具体原因；确定性能度量准则（属性、变量、等级）；确保测量设备的精度（至少是产品所需精度的 5 倍）；随机化；对数据收集进行分类（时间、材料、设备、操作者和缺陷的类型及部位）；分析使用若干 SPC 工具得到的数据（大多数的数据是庞杂的、任意地收集的，没有经过组织，使用受限）	任何 时间	通用
3. 图	• 显示趋势 • 压缩数据 • 向其他人 解释	选择两个或多个欲显示的参数；确定显示方法（大多数为柱型图，线型图或圆圈图）；选择显示参数的最合适范围	任何 时间	通用



续前表

工具	目标	方法	使用 时机	典型 使用者
4. 检查单	对原始数据进行分类	确定数据的归类（例如：缺陷的类型、缺陷的部位和天数）。量化每类数据	为直方图或频率分布图作准备	通用
记单据	以组群、单元的半图形方式标记	可将记录数据分成 10 档或更多单元。绘制出单元的边界或中值。在每个单元中标记出观测数（打叉）		
直方图/ 频率分布图	将数据转化成某种质量特性的平均图和散布范围图	将标记单上的数据转换成柱型图（直方图）或线型图（频率分布）以表示质量特性的各种价值和观察量之间的关系。（或占总量的百分比）。至少要求有 30~50 个观测值	在生产前或生产中用于过程能力研究	工程师，技师，生产线工人
5. 帕累托法则	将重要因子分离出来，以便集中管理	尽量识别众多问题的原因及对已知影响（美元、百分比等）的贡献。用 X 轴表示原因，用 Y 轴以递增或递减的方式表示影响。对主要影响的几个原因（通常 20% 或更少的原因对影响的贡献达 80% 或更多），制定采取措施的优先顺序	任何时间	通用——对制造业和白领工作是有用的工具
6. 头脑风暴法	利用群众的智慧尽量创造解决问题或改进过程的办法	针对最关心的问题，组织群众；准确定义问题；要求每个人写出原因或问题，及改进的办法；不容许批评；记录这些办法；畅所欲言；对最可能的原因进行投票，并付诸实施（该方法称为民主解决问题）	- 初期解决问题 - 过程改进	- 质量循环 - 改进小组



续前表

工具	目标	方法	使用时机	典型使用者
原因和后果(石川; 鱼骨图)	• 为了便于纵览所有的原因, 将原因进行分类, 然后确定在何处采取纠正措施	• 确定问题; 构造一个鱼骨图, 用主要“分枝”表示主要原因(例如材料、机器、方法、人员等); 用“细枝”细化每个主要原因; 量化技术规格; 确定每个原因的公差、实际值和对问题的影响; 如果原因与后果之间的关系能够量化, 围绕原因画方框图; 如果这种关系难以量化, 则在原因下面画一条线; 如果不能证明原因与后果有关系, 则不要标记原因; 用饼图表示主要原因的比例关系; 在发现根原因之前, 用 PDCA 的方式进行试验	• 解决问题的最初尝试	在日本的质量循环中广泛使用, 仅能解决简单问题
CEDAC (带有卡片的因果图)	• 与原因和后果类似, 早期原因的识别, 较好的工人参与	• 工人在现场识别原因。工人可以随时更改识别这样原因的卡片	• 与原因和后果相同	• 与原因和后果相同
7. 控制图	在找到主要原因并采取措施后, 保持最小变化的参数	详见控制图一章的内容	• 不是为了解决问题 • 生产	工程师, 技师, 生产线工人

7 种质量管理工具——无用的应用

日本人习惯于将简单的事情复杂化。例如，20 世纪 80 年代他们借用结构化开发原理的抽样理论开发了质量管理工具，称之为 7 种质量管理工具，并大肆吹嘘该技术如何的好。表 3-2 给出了这些工具的目标、主要方法、使用的时机及地点。尽管这些工具在策划和相关性研究中有用，但与解决问题没有关系。

表 3-2

7 种质量管理工具

工具	目标	方法	使用时机	使用地点
1. 亲和图 (K-J)	将种类不同的数据综合成有意义的类别	在一张卡片上记录一些数据，然后将其分成若干类别或族	将客户要求、质量问题原因、可靠性原因等数据进行分类	• QFD 方法 • 因果图使用者
2. 相关图	确定相互影响	在影响的方向上画箭头。方案的输出与输入用箭头表示，并进行排序。按照箭头方向可以找到根本的方案	• 总策划 • 推测根原因	• 头脑风暴会议 • 总策划 • 因果图使用者 • 作为亲和图的并行方法
3. 树图	将亲和图（最初的）分解成第二级和第三级的类别	细化方案，分解出方案的主要部分	确定和分解客户需求	• QFD 方法 • 作为亲和图的后续方法
4. 矩阵图	考虑两种方案或数据的相关性	比较两种数据或方案，确定相关程度	在 QFD 中比较“什么”与“如何做”	• 散布图 • 测量设备的精度 • QFD 方法

续前表

工具	目标	方法	使用时机	使用地点
5. 矩阵数据分 析图	市场和客户数 据	将 X 轴和 Y 轴 上的数据四等 分以确定市场/ 客户的策略	在商业/客户策 略开始时	• 业务量分析 • 客户牵引和 公司竞争
6. 过程决 策计划图	原因—后果— 纠错分析	类似故障树分 析, 但要考虑 原因和故障发 生的时间	早期设计	现场服务策划
7. 有向 图	找出项目的较 短的试验时间	类似项目评价 评审技术的部 分	新产品开发	任何复杂活动 的策划

资料来源: Keki R. Bhole, "A Powerful New Tool Kit for the 21st Century".

8 - D 方法——费力不讨好的方法

20 世纪 90 年代早期, 福特汽车公司引进了名为 8 - D (eight disciplines) 的解决工厂和供应商问题的技术。该技术包括 8 个步骤:

- 利用集体智慧;
- 描述问题;
- 包含的征兆;
- 发现和验证根原因;
- 选择纠正措施并验证;
- 采取永久的纠正措施;
- 防止再次发生;
- 奖励团队成员。

许多公司迫于福特的名望及其僵化而烦琐的质量管理方法, 不得不采用 8 - D 方法作为解决问题的魔法, 既费时又费钱, 结果却是无效和失望。



~~福特~~的8-D原理可以被认为是双倍的戴明 PDCA（计划、做、检查、采取措施）循环，最多提供了解决问题的程序和通用结构。但是它没有告诉人们如何解决既有的问题。这种基于推测和猜测的解决问题的技术已经使用了一个世纪，然而效果却非常令人遗憾。例如，在步骤4中（发现和验证根原因），提出了如下问题：

是否利用了因果图？是否表述了真正的根原因，而不是仅仅表述了征兆或效果？是否列出了过程根原因？是否将系统根原因列表表示？是否验证了真正的根原因，并列表表示？是否打破砂锅问到底？

8-D 仅仅用到了两项技术，即因果图和“五个为什么”，而且都是边缘值。因此，我们不再讨论8-D技术。

工程方法（观察、思考、尝试、解释）

如果工程师“与部件进行对话”（以下各章将详细介绍），那么“观察”这一步骤就完成了。但是，他们不知道如何与部件对话！“思考”是指针对长期问题的无效的猜测、建议、主观臆断和理论。“尝试”是指在某一时刻保持其他因子是常量，而仅仅变化一个因子（或变量）。这除了需要花费大量时间外，还会错失重要的相互作用结果，导致试验产生缺陷或造成错误的试验结果。“解释”通常是指试图解释结果及其与理论相符的合理性，并不想通过解决问题与否来验证改进程度。

工人的参与

一线工人比远离现场的工程师和经理更清楚实际发生的事，但是，即使是工人也有可能导致错误的意见和判断。

作者的原则是：（1）与部件对话；（2）与一线工人对话；（3）最后与工程师对话。

行为科学家指出，团队的概念是一个优秀的组织性构件。特别

是在日本，他们常采用团队命名的方法，例如质量循环、Kaizen 小组改进活动（SGIA）。但是，为了解决问题，团队中的工人仍然需要有效的工具。如果没有有效的工具，就如同把他们绑在轮子上而无法前进，毫无动力。

21 世纪 10 种有效的工具

如果公司雄心勃勃想取得世界级的质量，则其领导层必须理解、消化、吸收和指导利用这些新的、简单且非常有效的工具，这些质量工具远比 QC 的 7 种工具、7 种质量管理工具、工程方法、工人参与和 8-D 有效。这 10 种工具是：

- 试验设计（DOE）；
- 多环境强化应力试验（MEOST）；
- 质量功能展开（QFD）；
- 全面生产维修（TPM）；
- 水平对比；
- 消除操作人员控制的错误；
- 客户管理（NOAC）；
- 供应管理；
- 全面价值工程；
- 缩短生产周期。

因为第一个工具——试验设计是本书的重点，后面各章几乎都涉及到试验设计，因此，本章将不讨论试验设计。多环境强化应力试验（MEOST）也不在本章进行讨论，但鉴于它非常重要，所以我们单独把它作为一章来专门讨论。本章将概述其他 8 种工具。尽管对每种工具的详细描述超出了本书的范围，但是简要的介绍也可为读者提供指南。对于每一种工具，我们将重点强调工具的应用范围、需求、目标、益处、方法和基于我们在实际应用中总结出的注意事项。



工具 3：质量功能展开——倾听客户的意见⁴

应用范围

用于新产品/服务概念阶段的客户—市场—设计接口。

需求

产品设计长期受工程师和管理层所左右，而不是倾听客户的意见。结果，十有八九的新产品在市场上堆积如山。有人说：“你也许生产了世界上最好的狗食，但是如果连狗都不吃它，它又好在哪儿呢？”

Quality Department

2006.05.18

SongYongSheng

目标

1. 在概念设计转到模样设计前，确定客户的需求、要求和期望。
2. 客户根据重要性和你与最佳竞争对手的类似产品的比较对每项需求进行评级。
3. 基于“质量屋”矩阵，确定你的设计中重要的、困难的和比较新的方面。
4. 利用类似的“质量屋”矩阵，将产品技术规范分解成零部件技术规范、工艺技术规范和生产技术规范。

益处

1. 与以前的设计比较，设计时间减半、缺陷量减半、成本减半、人力减半。
2. 客户由失望或仅仅是满意转向非常满意，并对你的产品情有独钟。
3. 具有优越的竞争力。
4. 在客户、设计、制造和供应商之间建立了有效的关联。
5. 使新工程师易于接受。

方法

在 QFD 方法中使用的工具是质量屋，如图 3—1 所示。在图的左侧，按重要性程度列出了与客户最相关的要求（什么）；在图的右侧，是客户对每项要求的评价及你与对手竞争的表现的评价；中间是关系矩阵，用来比较每项客户要求与设计技术规范（如何做）之间的关联（强、中、弱）。用一个简单的计算就可以指出，那些必须既满足客户要求又迎合市场竞争需要的生产规范。图的底部是根据目标值和市场竞争制定的技术规范的比较，它是基于反求工程（竞争性分析）得到的。顶部是相关矩阵，表示每对工程技术规范是强相关还是弱相关。

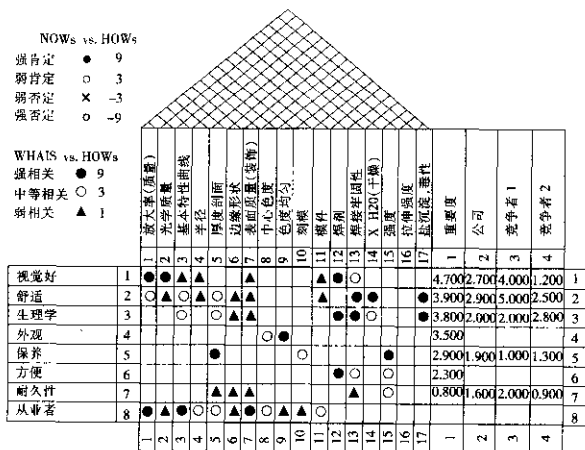


图 3—1 质量屋（隐形眼镜）

资料来源：Bhote, "A Powerful New Tool Kit...", 77

类似质量屋的矩阵可将技术规范的什么转换成零部件技术规范
的如何做，相似的转换还有：将零部件转换成工艺、将工艺转换成
产品、将产品转换成试验、将试验转换成质量要求。



注意事项

1. 大多数 QFD 研究不注重核心客户的需求, 因为这些客户过于个性化。仅靠一小部分攻关人员和专家会诊的研究也不是解决之道。**大量定制新的规范才是最好的办法。**这样能确定出每一个核心客户的独特要求, 然后按照客户的要求定制产品, 再利用信息技术和计算机综合制造技术, 就可以生产出价廉物美的大量产品。

2. 大多数客户趋于重视性能要求。客户也许更关心可靠性、交付日期、售后服务、**价格或其他问题。**

3. **不要列出多达 40~70 个的冗长和无法管理的“项目清单”, 建议不要超过 20 个。**

4. 通常客户不止一家, 例如, 客户、OEM 制造商和分销商。也许他们之间的要求互相矛盾, 需要分别进行研究。

5. 在第一次应用 QFD 时, 手工填写“质量屋”可能会“失败”, 因为填写者会因该工具的复杂性而丧失信心。

6. 90% 的 QFD 实践者在第一次将客户要求的什么转化成工程师如何做的技术规范, 或者将技术规范转化成零部件、工艺和试验的技术规范时, 都会遇到困难。

7. 大多数 QFD 的实践者在产品从方案到样机, 再到交付使用的设计阶梯中, 不愿意重新确定客户的要求和期望。这将形成产品与客户要求间的“鸿沟”。因此, 在产品的方案阶段、样机阶段和预产阶段都应按照客户的需求进行若干次 QFD 研究。

工具 4: 全面生产维修——改进设备生产率⁹

应用范围

制造过程/设备。

需求

1. 维修费用占销售额的 9%~15%。
2. 在工厂普遍存在“如果没坏, 就没有必要修”的观念。

3. 美国大多数公司的工厂总有效率 (FOE) ($\text{产量}\% \times \text{正常运行时间}\% \times \text{机械效率}\%$) 是 40%~50%。机械效率等于运行时间除以 (运行时间+准备时间)。

4. 美国 50% 的工厂中的维修人员仍然将其一半以上的时间用在维修上, 而不是用在防止问题的发生上。

目标

1. 极大地改进工艺/设备质量和生产率。
2. 提高工厂的生产率; 降低生产周期和存货。
3. 建立以预防为主, 而不只是纠正和解决问题的维修队伍。
4. 减少设备生命周期成本。
5. 掌握工艺特性并优化工艺, 用正向控制冻结关键参数。

益处

1. 劳动生产率可提高 40%~50%。
2. 设备故障可减少 90%。
3. 生产线能力可提高 25%~40%。
4. FOE 增加超过 85%。
5. 每单元维修费用减少 6%。
6. 共节省销售额的 5% 到 14%。

方法

1. 利用试验设计提高产量。
2. 利用四种重要的 DOE 工具改善正常运行时间 (即减少停机时间)。这四种工具是: (a) 说明工艺特性的变量搜索法; (b) 优化工艺的散布图法; (c) 冻结工艺的正向控制法; (d) 消除“噪音”因子的工艺验证法。
3. 在不牺牲产品质量和设备寿命的条件下, DOE 试验可提高工艺速度。

4. 提高机械效率 [即: $\text{运行时间} / (\text{运行时间} + \text{准备时间})$]。



这主要是利用非常确定的组装时间减少方法、流程图、实时监控、强化实践管理等方法将准备时间的减少达到 50 : 1 以上（正如赛车后勤人员作准备工作一样）。

5. 建立操作者——维修人员队伍，以便实施 DOE 和其他连续改进 FOE 的试验。

6. 鼓励操作者实施更多的预防规范，例如清洁、润滑、保持记录、正向控制日志和预控制。

7. 培训操作人员和维修人员解决问题的技巧，特别是试验设计和价值工程方面。

8. 以“**实践，实践，再实践**”^[1]的指导思想来量化和优化关键工艺参数，减少准备时间并进行应力试验。

9. 建立较好的量化尺度来追踪 FOE、平均故障间隔时间（MTBF）和零备件的使用。

10. 利用可以监控关键工艺变量（例如温度、振动、噪音和润滑）的诊断和报警信号改进预计的维修。

注意事项

1. 应赋予操作者和维修工人强大的解决问题的工具，例如 DOE，而非效果稍逊的工具，如因果图、CEDAC 和 PDCA 等。

2. 在组建团队时，要预先明确每个成员的责、权、利。

3. 应与新设备/工艺的供应商一起在工厂的基层单位开展 DOE 工作。在开始全面生产前，应直到取得高产量，再付给供应商资金。供应商会从这样的试验设计工作中受益，同时能加强其他客户对它们产品的兴趣。

4. 应在产品开发前半年到一年就开始开发工艺。

工具 5：水平对比——学习并采纳最好的经验¹⁰

应用范围

通用，包括制造业、商业、服务业、公共事业和政府等。

需求

1. 设定内部目标，在时间和水平上内部目标不够充分的过程。
2. 公司通常不希望学习他人——“非我族类”（NIH）症结。
3. 尚未将水平对比看做一项重要的合作要求。

目标

1. 水平对比与关键的合作策略紧密结合。
2. 消除公司与业内最佳公司间在关键功能、管理、技术方面的鸿沟。
3. 在某项关键功能、管理或技术方面成为最佳。
4. 在日常工作中，将水平对比制度化。

益处

1. 蛙跳式竞争。
2. 缩短学习时间。
3. 在产品和服务中，对质量、成本和生产周期的改进来说，都是一个高效的工具。
4. 扩展视野产生全球化意识。

方法：水平对比的 12 个步骤

1. 确定什么是水平比较以及为什么进行水平比较——与关键的合作策略密切结合。
2. 建立你自己公司的性能，作为公司的基线。
3. 在你自己的公司（其他部门）和相邻的公司中开展活动。
4. 确定比较的目标——国家级的还是世界级的。
5. 参观比较的公司。
6. 确定比较的公司与你公司之间在业绩方面的差距。
7. 确保合作承付款项的安全，防范虚假收购。
8. 设立目标和制定行动计划。
9. 实施计划并监控结果。

10. 重复/循环上述过程。
11. 利用“横向思维”重新设计过程。
12. 确立比较基准。

注意事项

1. 如果没有一个公司指导委员会和一个进行水平比较的权威机构，那么工作将缺乏重点，在组织方面也会存在脱节现象。
2. 在水平比较和关键商业结果之间必须建立关联。
3. 应既涵盖进行水平比较研究的“内部客户”，又能够支持服务的合作和帮助。
4. 对团队成员来说，水平比较方面的在职培训比在课堂上的专职培训更重要。
5. 在与其他公司进行水平比较前，必须知道自己公司的基线。
6. 需要认真研究什么公司才是比较的基准。非竞争对手的公司当然更好，但也不应放弃比较竞争对手。
7. 必须落实先前提出的问题清单，而且现场拜访前应先进行电话访问。
8. 确保在水平比较公司访问的人是合适的人。对其公司的成功和失败应予以同样的重视。
9. 至少每年或每两年重复一遍比较水平过程。
10. 应把比较水平输出作为进行进一步研究和行动的弹性模板。总有方法来改善一家公司的最佳表现。

工具 6：防差错 (Poka-Yoke) ——清除操作者可控差错¹¹

应用范围

制造业流水线工人。

需求

1. 无论如何严格管理和给多么高的酬金，人都难免会出错。



2. 在劳动密集型操作过程中，流水线工人会造成质量问题。

目标

提供电子式、机械式或可视的传感器，以警告操作者已发生了一个错误；更佳的形式是能预警出错，并能避免出错。

益处

1. 直接、迅速，对工人没有威吓和责斥。
2. 可提高质量、生产能力和客户满意度。
3. 实际上省去了统计过程控制。
4. 较好的简化生产性的设计。

方法

最好举例说明（见图 3—2）。

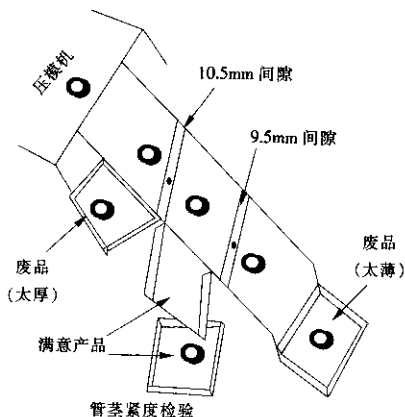


图 3—2 自检系统例证

资料来源: Bhote, "A Powerful New Tool Kit..."



注意事项

1. 传感器必须简单，而非复杂的设计。
2. 利用工人的聪明才智来帮助设计传感器。
3. 最佳的 Poka-Yoke 解决方案是通过试验设计或设计来减少变量。

在未用 Poka-Yoke 解决问题前

绝缘材料粉末被倒入入压模机中。被压缩粉末的模片的内径和外径是一个常量，但厚度在变化。某工人有三台机器需要观察，因此当这些模片通过槽时，该工人无法控制模片的厚度。

采用 Poka-Yoke 解决问题后

规定的绝缘材料粉末模片的厚度是 $10\text{mm} \pm 0.5\text{mm}$ 。将一个计量传感器/指标器 A 附在槽的上端，计量传感器与槽的空隙设为 10.5mm ，再将一个计量传感器 B 附在槽的底端，计量传感器与槽的空隙设为 9.5mm 。当成形的模片通过槽时，厚度大于 10.5mm 的模片就不能通过计量传感器 A，并由 A 引导到槽的左侧的废品箱；而厚度小于 10.5mm 的产品则经计量传感器 B 检测。同样，小于 9.5mm 的模片被引入另一个废品箱。而被计量传感器 B 截留的处于 $9.5\text{mm} \sim 10.5\text{mm}$ 之间的合格产品则进入可接受部件箱。

最佳的解决办法

进行 DOE 试验，看一看压模机中为什么会产生模片厚度的变化，如何防止发生这种变化。

工具 7：客户管理——白领质量、成本和周期时间的改进^{12,13}

应用范围

制造业公司中所有的服务组织和保障服务。



需求

1. 服务业的生产率低于 40%（而制造业则高达 80% 以上）。
2. 质量对于大多数白领人士来说是“超出屏幕”。
3. 在任何服务性操作中，周期时间好像是一个“外来语”。
4. 美国的全国生产率被服务业拖了后腿。

目标

1. 改善所有服务性操作中的质量、成本和周期时间。
2. 将垂直管理变成水平管理。
3. 打破部门界限，组建跨职能小组。
4. 革新商业过程。
5. 在性能评价方面，用内部客户的评价取代老板的评价。

益处

1. 高利润、高投资回报率、高市场占有率和高生产率。
2. 通过工人愉快而高效的生产劳动，获得更高的客户满意度和忠诚度。
3. 员工满意度和工作热情得到提高。
4. 极大地改进商业过程。

方法：客户管理的 10 个步骤

1. 设立指导委员会、过程管理者和若干个改进团队。
2. 确定商业过程问题；量化对质量、成本、周期时间和士气的影响。
3. 确定主要的内部客户及其要求的优先次序。根据测量方法，就满足这些要求的内部供应商的能力达成协议。
4. 确定内部客户反馈频率，以此作为满足或未满足客户要求的记录和结果。
5. 绘制整个“宏”过程流程图。
6. 为过程中和全部周期时间内的每个步骤确定所需的平均周



期时间。

7. 将有附加价的步骤与无附加价的步骤相分离。若可以将无附加价的步骤去掉，估算出可减少的步骤和可缩短的时间。

8. 利用过程改进工具，例如强化外场分析、价值工程、试验设计、过程重新设计和工作重新设计，来去掉或减少无价值步骤。

9. 利用“思想箱”、价值工程、创新工具，来检验一个完全不同的解决商业过程或服务的方法，包括该方法的淘汰。

10. 实施对内部客户分数管理的评审，并根据健全的商业记录追踪过程。

注意事项

1. 客户管理是迈向商业过程重组工程（business process re-engineering, BPR）进程的一步，但是 BPR 是一个太过先锋的概念，以至于难于实现。该方法在员工的雇佣方式、评价、报酬、激励和提升方面真正体现了革命。因此，BPR 完全被公司采用至少尚需 10~15 年。NOAC 就是达到这种前景的实践桥梁，而且一些有远见卓识的公司已采用了该技术。

2. 在步骤 6、7、8 上不需花费过多的时间，因为解决问题的完全不同的革命化方法（步骤）甚至能完全剔除过时的过程。

工具 8：供应管理——供应商质量、成本和改进周期的转折点^{[2], 14}

应用范围

想达到双赢关系的关键供应商。

需求

1. 50%以上的销售额来自外部的供应商提供的材料，而仅有 3%来自厂家的直接劳动力。

2. 公司越来越认识到他们应进行“核心竞争”，而不是在所有的商业活动中竞争。

3. 在市场上，平均价格每年下降2%~7%（对于高科技产品，这个数字更高）。然而，供应商对客户公司的价格每年上涨5%，最终不得不退出商业圈。

4. 公司限制了每个零部件的采购价格，更需要找到能够缩短设计周期和降低生产成本的子系统供应商。

目标

1. 与关键的供应商进行真正的合作（90%以上的合作者仅是名义上的），以获得互惠的利益。

2. 为供应商提供积极的、具体的帮助，以改进供应商的质量、成本和生产周期，同时降低每年的产品价格，让供应商早期介入设计。

益处

1. 改进合作的供应商质量，幅度从10:1到100:1。

2. 每年平均降低合作的供应商价格5%，同时帮助它们提高效益。

3. 缩短合作供应商的生产周期和原材料的库存量，改进程度达10:1。

4. 加强所有重要的客户及供应商之间的忠诚度和联系。

方法

1. 建立高层管理指导委员会，指导供应管理的主要企业工作。

2. 与合作的供应商共同建立严格的、有意义的、相互接受的技术规范。

3. 对每个主要商品建立分类商品组，以便帮助有关的合作供应商改进质量、成本和生产周期，同样也不断地降低产品价格。

4. 使得从“制造”到大量“购买”，最后到“成批购买”更加容易。

5. 简化供应商的复杂的情况，与供应商一起并行设计，实现



成本目标。

注意事项

1. 公司不能用取消供应商的供应资格来驱动供应商降低成本。必须采取双赢的解决方案，即公司每年都能得到更低的价格，而供应商年年都获得更高的利润。

2. 公司和供应商应基于高度的信任和合作的基础之上。

3. 公司成员若想帮助供应商开发产品，他们首先必须是质量、成本和生产周期改进工具方面的专家。

4. 合作的供应商不要第一次遇到困难就退缩，而应从长远利益来考虑。

工具 9：全面价值工程——远远超越了传统价值工程¹⁵

应用范围

设计、制造、服务和供应商。实际上，价值工程可广泛用于任何经济活动领域。

需求

1. 客户要的是价值，而不仅仅是价格。
2. 螺旋式上升的劳力、材料和管理成本。
3. 不理想的利润。
4. 传统的成本降低大多是无效的，因为它要求产品/服务配置保持一致。

目标

1. 远远超越了成本的降低和传统的价值工程。
2. 远远超过增强客户满意度所要求的所有方面的质量改进。
3. 加强了公司、学校和政府间的作用。
4. 为持续不断地改进提供了结构框架。

益处

1. 成本平均降低 25%，至少为 10%。
2. 投资回报达 10 : 1。
3. 改进了客户要求的质量、可靠性和其他方面。
4. 保护资源和环境。

方法

技术的变化，包括价值方法、功能分析、价值工程工作计划、“五个为什么”、工作重设计、过程重设计和思想库。

注意事项

1. 在低成本的驱动下，避免牺牲质量。
2. 必须认真评估真正的客户要求，技术规范不能被模糊地接受。
3. 调动客户各方面的积极性，不能仅限于成本要求（见图 3—3）。
4. 供应商必须全身心投入到全价值工程的工作中。

工具 10：缩短生产周期——质量、成本、交付期和效益的综合体现¹⁴

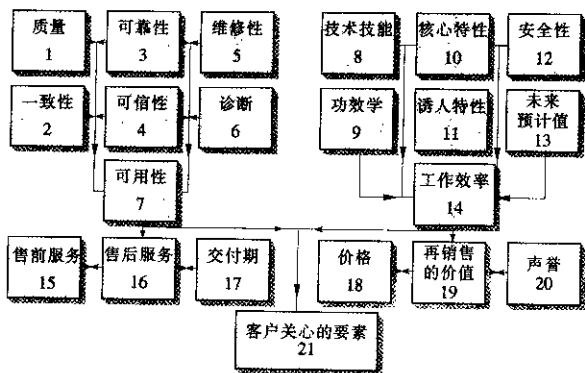
应用范围

制造、商业服务和设计。

要求

1. 公司在所有方面都存在由于浪费时间而导致的浪费成本现象。
2. 在过去，不需要减少库存和缩短生产周期，这是因为在产品清单中没有这一项要求。
3. 公司用大量的安全贮存以避免延期交付产品。





图中:

1. 质量: 面向零缺陷	12. 安全性: 针对用户的产品的安全性, 预防不利的产品条件
2. 一致性: 面向零偏差	13. 未来预计值: 预先考虑的需求
3. 可靠性: 面向现场零失效	14. 工作效率: 8 到 13 个部件的集成
4. 可信性: 面向寿命时间保证	15. 售前服务: 销售、合作、沟通
5. 维修性: 面向准确、迅速、低成本维修	16. 售后服务: 持续的和售后的注意事项
6. 诊断: 面向客户的自我诊断	17. 交付期: 短时间
7. 可用性: 面向 100% 的正常运行时间	18. 价格: 低价竞争
8. 技术性能: 先进的技术	19. 再销售价值: 购买价格的高百分比
9. 功教学: 式样新、颜色好、操作“用户友好”	20. 声誉: 印象、感觉的质量
10. 核心特性: 由用户预定	21. 客户关心的要素: 价值、合意、诚信
11. 诱人特性: 使客户兴奋的特性	

图 3—3 客户关心的要素图

资料来源: Keki R. Bote, *Going Beyond Customer Satisfaction to Customer Loyalty*, New York: American Management Association, 1997.



目标

1. 检查浪费时间的所有方面——等待、启动、运输、存储、批准、检验、试验等，实际上是一场针对所有浪费的战争。
2. 利用过程图（流程图），但更重要的是，利用“思想箱”使制造、业务和设计工艺改头换面。
3. 需要“推”与“拉”系统。

益处

1. 公司的新产品快速占领市场，在竞争中获胜。
2. 缩短生产周期，减少幅度达 10 : 1，甚至 100 : 1。
3. “库存周转”增加，增幅为 3 : 1，6 : 1，甚至 100 : 1。
4. 当生产周期缩短时，质量、成本、产品交货和效益同时得到改善。
5. 当生产周期缩短时，预报变得不太必要，而且完全放弃了老的进度月报计划系统。

方法

1. 利用诸如 DOE 技术来减少产品缺陷，缩短检验时间、试验时间、返工时间和“争议”时间。
2. 利用全面生产维修（详见本章工具 4）改进工厂的总效率。
3. 利用蜘蛛图绘制工厂产品的物理流程，从不经济的过程转向生产水平高的过程。
4. 利用工厂的中心方案，关注重点客户产品、专用设备和专职人员的范围。
5. 减少记录时间，减少启动/更改时间达 10 : 1 倍，甚至 100 : 1 倍。
6. 利用“看板”拉系统来取代主要进度计划和进度月报计划（MRP-2）。
7. 利用制造单元和 U 形状的布局。
8. 相互协作。



9. 培养多面手。
10. 与供应商和客户一起开发类似的缩短生产周期技术。

注意事项

1. 在工厂的小范围内开始——W. 威利 (Wee Willy) 方法几乎适合于任何时间, B. 鲁斯 (Babe Ruth) 偶尔也适用。
2. 澄清质量, 缩短生产周期的必要条件。
3. 为缩短启动时间/交接时间而实践、实践、再实践。
4. 摆脱进度月报计划、主要进度计划和过分依赖计算机的约束。
5. 测量底线结果——缩短生产周期, 提高库存周转率, 改善投资回报。

对 21 世纪工具箱的认识和应用

功能强大的工具箱鼓舞着创立世界品牌的公司。但是一般公司和优秀公司对这些工具的认识和应用情况怎样? 如表 3—3 所示, 笔者对其应用情况进行了估计, 虽然不是很科学, 但是仍可以说明问题。认知这些工具的公司非常少, 甚至所谓的优秀公司也是如此, 更不用说应用这些工具了。正如人们所说: “任重而道远!”

表 3—3 对 21 世纪工具箱的认识和应用

工具	一般公司		优秀公司	
	认知 (%)	应用 (%)	认知 (%)	应用 (%)
1. 试验设计*	0.01		10.00	1.00
2. 多环境强化应力试验	—		1.00	0.05
3. 质量功能展开	1.00	0.01	10.00	1.00
4. 全面生产维修	5.00	0.02	15.00	3.00
5. 水平对比	10.00	0.50	50.00	20.00
6. Poka Yoke	1.00	—	10.00	1.00

续前表

工具	一般公司		优秀公司	
	认知 (%)	应用 (%)	认知 (%)	应用 (%)
7. 客户管理	0.10	—	3.00	0.10
8. 供应管理**	10.00	1.00	40.00	10.00
9. 全面价值工程***	2.00	0.01	20.00	4.00
10. 缩短生产周期	5.00	0.20	30.00	5.00

* 试验设计指谢恩方法，正如本书所介绍的，而不是效果不佳的传统或出口方法。

** 供应管理是指与关键供应商真正的双赢，以及真正对供应商有帮助的具体活动。

*** 全面价值工程远远超越了传统的价值工程（传统的价值工程注重成本的降低），考虑到了客户各方面的要求。

资料来源：Bhote, “A Powerful New Tool Kit...”

【注释】

[1] 在我们介绍 6σ 的会上，有一组外科医生参观了摩托罗拉公司，目的是想提高他们的生产率。会上讨论的技术之一就是水平对比法（工具 5）。为了看一看谁的速度最快，他们参观了印第 500 汽车大奖赛比赛，比赛中维修站的工作人员仅用 12 秒就更换了赛车的四个轮胎。问其成功秘诀，用他们自己的话来讲就是“实践、实践、再实践”。外科医生回去后研究了外科手术规程的每一个步骤。通过实践、实践、再实践，结果是每天的手术量提高了一倍，为病人的健康做出了贡献。

[2] 在公司合作的尺度内，供应管理是管理的第一步，除了公司的质量、成本和周期外，该方法对于供应商的发展和成长也是一个非常有效的工具。

Quality Department

2006.05.18

SongYongSheng

第2篇

试验设计绪论



第4章

工序能力的衡量

为什么说变量有害

变量是大自然的产物，因此可以顺理成章地认为变量在工业界同样是与生俱来、不可避免的。然而，今天我们有了一个新的范例——变量是有害的，至少在工业界如此。使用试验设计的新工具能够大幅度减少变量，从而削减成本，而不是增加成本，并由此提高客户的满意度。

源自足球的类推

图 4—1A 是一个旧的范例。只要一个企业的产品在技术规格的上限和下限之间，

就认定客户对此是满意的。有关足球的推理是恰当的。当射门时，无论球落在何处，只要它在两个球门柱之间就行。比如，球可以落在靠近任何一侧门柱处，亦或落在球门中间，踢球的一方都算射门得分。我们在工业界也有这种足球心理。规格有如门柱。如果一个企业产品处于规格的界限之内并接近其中的一端，我们就认定客户对此是满意的。如果另一个企业产品刚好落在了规格界限之外，我们就假定客户的满意度降至零。实际上，对于客户的满意度没有一个数字化的衡量。两个产品都在最外缘，处于临界状态，不可接受。

图4—1B是一个新的范例。客户希望他们接收的产品具有一致性和连贯性。目标值（通常位于规格宽度的正中间）是最佳的。相对于这种目标值或设计中心值的任何偏离都代表经济损失的不断增大，一个组产品向规格界限的一端或另一端移动，会给公司、客户甚至社会带来经济损失。

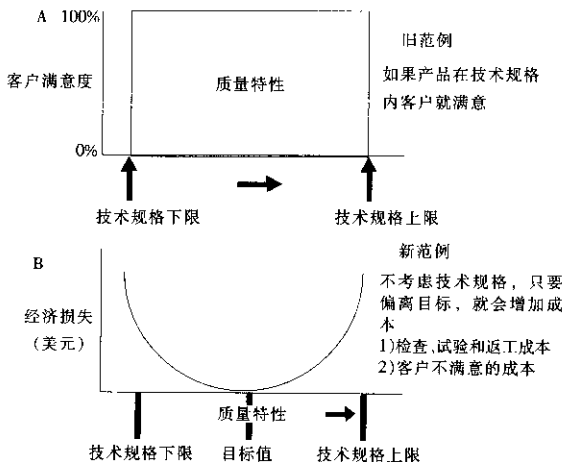


图 4—1 技术规格限与目标值

案例研究

福特公司与马自达公司的比较案例研究

此范例通过著名的福特公司的案例研究加以说明，其内容已经编入录像带用于培训。几年以前，福特公司发现其自动变速装置存在严重问题，噪音大、换挡性能差、客户怨声载道，公司为此花费了巨额保修成本。福特决定将其自动变速装置与马自达公司的进行比较。因为马自达公司的该装置有着完全相反的效果——噪音低、换挡平顺、客户满意而且保修成本非常低。福特决定把这两家的变速装置拆卸下来进行比较。它们有相同的设计和规格，而且，福特的每一个零部件都满足规格，那么究竟错在何处呢？福特发现，其零部件的尺寸分布遍及各处，有些在规格的这一端，有些在另一端，还有一些在中间。与之相反，马自达公司的零部件规格则紧紧围绕着一个目标值、一个设计中心值。结果，唐纳德·彼得森先生（Donald Peterson），也就是后来的福特公司董事长，向福特公司的工程师发布了一道命令，让他们按照目标值进行设计，而不是让尺寸在宽泛的规格界限周围左右偏离。

日本电报电话公司的案例研究

日本电报电话公司（N. T. T.）是日本政府所有的大型公司之一，有6家日本供应商和1家美国供应商为其提供传呼机，其灵敏度的最大规格是12微伏。美国供应商很轻松地就达到了这一技术要求，它的技术条件范围是2微伏~11微伏（数值越低越好）。然而，N. T. T. 公司希望灵敏度水平一致，而不考虑规格，这样，每一个传呼机都将有一个相同的接听范围。这表明，其两家最好的日本供应商的产品有一个较窄的范围——4微

伏~8 微伏。N. T. T. 公司竭尽全力声明，如果美国供应商不能满足这样窄的灵敏度范围，那么合同将被迫取消。美国供应商痛苦地采取了权宜之计，它使用 100% 的筛选作业操作，只把灵敏度在 4 微伏~8 微伏之间的产品选出来——成本高达 200 万美元！后来，在笔者的帮助下，美国公司开始系统的努力，使用 DOE 技术将灵敏度的分布减小到 5 微伏~6.5 微伏。结果如何呢？美国供应商击败了 6 家日本竞争对手，成为 N. T. T. 公司的首选和最受欢迎的供应商。

F-16 喷气战斗机的案例研究

近年来，由于政治因素，日本的公司为美国国防部制造了一些 F-16 喷气战斗机，其技术规格对于两国的供应商来说完全相同。但是，实际的历史表明，日本 F-16 战斗机的可靠性（平均故障间隔时间）是美国 F-16 战斗机的两倍！原因在于，美国的生产企业使用规范公差的范围，而日本企业则力图使尺寸波动接近于零。事实上，可靠性方面得到的教训之一就是，减少生产中的缺陷和偏差是保证可靠性的极好方法，尽管生产中的缺陷与实际中的可靠性之间并没有绝对的数学关系。

公差的建立

如果只考虑一个零部件，刚刚满足规格界限还可以勉强接受。但是如果两个零部件每个都占据整个的规格宽度，那么组件的总公差将超出允许的规格宽度。不合格水准将随着一个组件中所包含的零部件数量 5、10、20 或 50 地不断递增。不幸的是，“刚好满足规格”这种理念已经在制造业的灵魂中根深蒂固，并形成了一种永恒的文明，即进行不必要的检查和试验以及高成本的返工和报废。

检查、试验、返工以及报废的巨大成本

除了客户满意度方面的巨大损失以外，尺寸变化之所以有害，还在于不良质量造成的可怕损失。80%的公司不知道这个度量，15%的公司不进行测量，4%的公司进行了分析但并没有去降低它，剩下1%的公司将其转化为基础利润。不良质量成本（COPQ）中最基本的成分包括保修、报废、分析、返工、检查和试验，其中没有任何一项是增值的。它们占据了公司销售额的10%~20%。笔者进行的分析表明，没有开始质量革命的公司，每个雇员每天造成的不良质量成本在100美元~200美元之间。这是一笔多么大的浪费啊。减少这些成本的最佳方法是减少试验设计过程中的变化。

工序（加工）能力简介： C_p

在研究偏差的根源和成因以及减少偏差的方法时，我们必须先测量偏差的衡定标准。 C_p （一个工序能力的意思）和 C_{pk} （一个工序能力，但要考虑分布中心与公差中心偏离程度）以其过去20年来最基本的或最具参数的水准，已经成为质量的标准语言。然而即使到了今天，仍有30%的公司和20%的质量专业人士不知道 C_p 和 C_{pk} ，或者没有系统地用其测量变量并减少它。

C_p 的定义是规格宽度（S）除以工序宽度（P）或范围。它是一种对散布范围的衡量。图4-2描述了6种比较规格宽度（通常是40—20=20）的频率的分布，以改变过程宽度。

图4-2中的过程A的工序宽度是30，由此得出 C_p 值为0.67。该过程未在控制之内，两端均有2.5%的瑕疵部分。这在20世纪70年代，在统计过程控制（SPC）出现之前形成了公司（仅有少数公司进行测量）的一种标准，即便在今天，30%的公司中重要参数的 C_p 值也仅等于或低于0.67。它们对这种失去控制的情况的补救就是强行拣选、报废和返工。



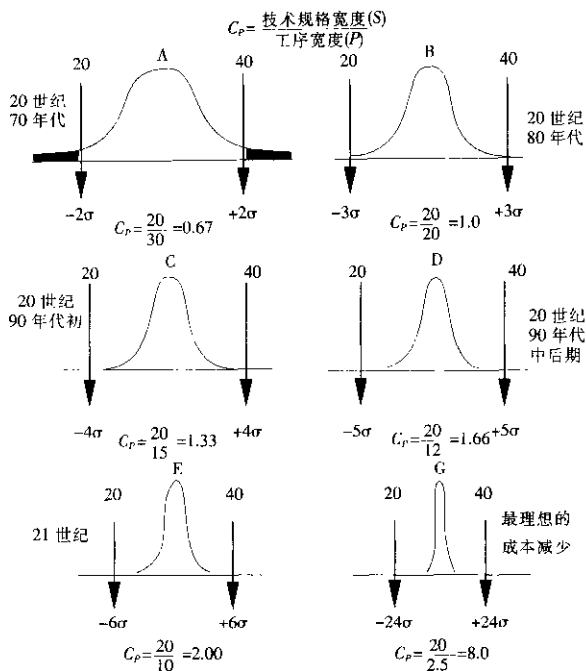


图 4-2 变量的测量—— C_p

注：技术规格上限=40；技术规格下限=20

过程 B 的工序宽度等于已知的规格宽度 C_p 1.0。它是一种临界情况，至多有两种原因：(1) 它的两端各有 0.13% 的瑕疵部分或每百万个零部件中有 1 300 个是次品。在当今充满竞争的世界里，许多公司正将不合格品的上限建立在每百万 100 个、10 个甚至每 10 亿个有几个零部件的基础上，每百万出现 1 300 个次品的时代已成为过去。(2) 任何偏离工序中心的略微偏移都将造成不合格水准的明显上升。

过程 B 代表了 20 世纪 80 年代许多公司的典型情况。SPC 曾风行一时。美国汽车工业曾经大规模地使用 SPC 以求质量。然而，到 90 年代中期， C_p 1.0 陷入了困境。我们自吹自擂的 SPC 同样也难逃厄运。

过程 C 的情况比较好一点， C_p 值为 1.33。其中在严密的工序界限和规格界限之间有一个保护段。它代表了 90 年代早期稳步发展的一些公司的典型情况。QS 9000——由美国三大汽车公司制定的质量标准，要求将 C_p 1.33 作为其汽车供应商的最低标准，尽管它们自己的制造厂还达不到该要求。不错，这确实很虚伪。

过程 D 的 C_p 值是 1.66，在工序界限和规格界限之间有了一个更宽的安全裕度。90 年代中后期只有几家公司达到该水平。

过程 E 的 C_p 值是 2.0，其工序宽度仅为规格宽度的一半。它正迅速成为 21 世纪头 10 年统一的质量标准，目前仅有少数世界级的公司达到了该标准（笔者在摩托罗拉汽车及工业电子集团公司工作时，早在 1984 年便在一些重要的质量参数上达到了 C_p 值 2.0）。

另一个与 C_p 有关的公制单位是 σ （或称标准方差）。表 4-1 表明了 C_p 、 σ 与相关不合格品率之间的关系。这是有实际统计意义的 6σ ，不是对未来情况的每百万 3.4 个不合格品的统计量，而是每 10 亿（ppb）中有 2 个次品的目标。

表 4-1 C_p 、 σ 与不合格品率

C_p	σ	不合格品率
0.67	$\pm 2\sigma$	5%
1.0	$\pm 3\sigma$	0.13%
1.33	$\pm 4\sigma$	60ppm
1.66	$\pm 5\sigma$	1ppm
2.0	$\pm 6\sigma$	2ppb（每 10 亿个中的部件）

资料来源：Motorola Six Sigma Seminars.

过程 F 的 C_p 值为 8.0（ $\pm 24\sigma$ ），不仅是最理想的，而且也是可以达到的，总成本最低。实际上，对于越来越高的 C_p 值没有限制，可以达到 10、15 甚至更高，只要产品或制造中没有重复增加成本，而只是 DOE 成本的增加——正如同一种投资。我们的经验



表明, C_p 值越高, 不仅不会增加成本, 实际上成本还会降低。此外, C_p 值从 1.0 或更低增大到 2.0 或更高, 不是以年、月来衡量, 而是以周、天来计算的, 通常利用不超过 1、2 或 3 个设计完好的 DOE 便可实现。

C_{pk} , 一种更好的测量偏差和工序能力的方法

C_p 仅仅用于简单地介绍生产加工能力这一概念, 它没有考虑与一个参数的规格界限有关的任何生产过程尺寸分布偏离中心的情况。这种偏离中心的结果是降低了安全裕度, 因此带来损失, 称为 K 或修正因数。公式如下:

$$C_p = S/P$$

$$K = \frac{D - \bar{X}}{S/2} \text{ 或 } \frac{\bar{X} - D}{S/2} \quad (K > 0)$$

$$C_{pk} = (1 - K) C_p$$

式中: S 为规格宽度; P 为加工 (尺寸分布) 宽度 ($\pm 3\sigma$ 限制); D 为设计中心 (D 不一定在规格宽度的中点); $\bar{X}$ 为制造平均值。

当加工 (尺寸) 平均值 $\bar{X}$ 和设计中心 D 或目标值一致时, K 为零, 这时 C_p 等于 C_{pk} , 但是如果加工 (尺寸) 平均值斜向规格界限的一端或另一端, 偏离设计中心, K 值增大, 这时引起 C_{pk} 相对于 C_p 减小。

图 4—3 对此有所说明。图形 A 分布宽, C_p 为 0.71。因为它的设计中心 D 和平均值 $\bar{X}$ 一致, C_p 和 C_{pk} 也同等于 0.71。图形 B 分布窄, C_p 值高达 2.5。但是因为它靠近规格界限的下限端, 因此 K 修正后, 得到较差的 C_{pk} 值 1.0。图形 C 比图形 B 的展形宽, C_p 值较低, 为 1.67。但是与图形 B 相比, 由于它靠近设计中心 D , 因此 K 的损失较小, C_{pk} 为 1.33, 比图形 B 理想。图形 D 是最理想的, 分布窄, 加工结果又定中, 此时 C_p 和 C_{pk} 同等于 5.0。

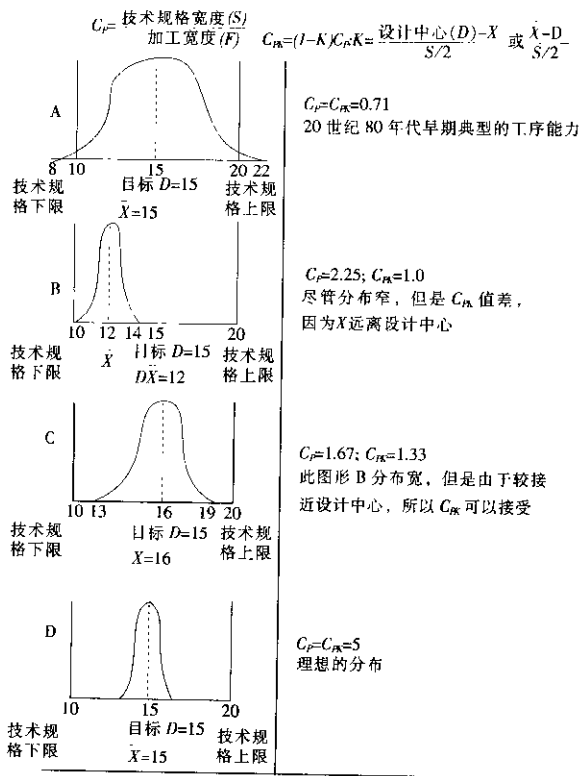


图 4—3 工序能力

注: 关键参数: 最低 $C_{pk}=1.33$
 期望的 $C_{pk}=2.0$
 理想的 $C_{pk}=5.00$



C_{pk} 还有一个更简单的公式，该公式对于单边规格界限尤其有用：

$$C_{pk} = (\bar{X} - \text{最近的规格界限}) / 0.5 \text{ 加工（尺寸）宽度}$$

或 $C_{pk} = (\text{最近的规格限制} - \bar{X}) / 0.5 \text{ 加工（尺寸）宽度}$

C_{pk} 为正值的那一个。

C_{pk} 是一种极好的测量可变性和加工能力的方法，因为它同时考虑了加工（尺寸）散布和偏心情况。[在加工控制中，定中加工比减小加工（尺寸）散布图容易得多。定中只需要简单的调整，而加工（尺寸）散布图的减小通常要求耐心地应用试验技术的设计。]在 C_p 中，目标是获得越来越高的 C_{pk} 。 C_{pk} 为2.0只不过被看做从零缺陷到接近零偏差的万里长征中的一个里程碑。

测量 C_p ， C_{pk} 时需避免的错误

错误 1：衡量所有的参数

许多急于开始推行 C_p 文化的公司通常所犯的错误是衡量所有的产品参数，而不考虑其重要性。这种方法只能增加成本而不能增加价值。只有最重要的参数才需要2.0甚至更高的 C_p 值，其余的则无所谓。其余参数的 C_p 值可以是1.0甚至0.5，而且对于大多数零部件来说甚至不需要测量。要想区分重要和不重要的参数，工程上的猜测并不是依据，而应该在产品或制造的设计阶段使用变量搜索作为主要技术的试验设计（见第13章）。

错误 2：经常地测量 C_p ， C_{pk}

另外一个过分积极的缺点是不断地测量一些给定参数的 C_p 和 C_{pk} 值。测量一个 C_p 值的结果是非好即坏——高于2.0或低于2.0。

如果工序能力以前已经被验证过了，那就无须不断地进行再测量；可以通过预先控制加以监控（见第 12 章）。如果 C_p 值大大低于 2.0，参数就在控制之外，反复进行测量也不会奇迹般地令其处于控制之内。最好的方法是使用 DOE。

一家倾心于 C_p 测量的、主要的、享有名望的制造商讲到，它的每一家工厂每周都要对 200 多个参数进行 C_p 和 C_{pk} 的测量并将结果报告公司总部。其 C_p 值的范围在 1.0~1.3 之间——全部低于最低值 1.66。如果某一个特殊 C_p 值从 1.3 降到 1.2，工厂就得拿出为何有这种错误的说明！笔者在担任该公司顾问期间，曾经主要使用 DOE 来试图替代这一程序，结果为该公司节约了几十万美元并减少了变量（只有针对参数才是真正重要的）。

错误 3：如果规格界限不正确，那么 C_p 和 C_{pk} 就没有意义

C_p 和 C_{pk} 以规格界限为基础。如果不修正后者， C_p 和 C_{pk} 的值就是错误的且没有意义的。正如我们将在“散布图”一章中说明的那样，工程师们制定的 90% 的技术规格不是太紧就是太松，而且总是错误的。当考核一个产品或工序的质量时，必须使用变量搜索将重要的参数和不重要的参数区分开来。接下来，必须使用“散布曲线图”确定重要参数所需的公差和规格界限。只有如此，才能切合实际地确定 C_p 和 C_{pk} 。

错误 4：将 C_p 测量扩大至供应商而不在本公司执行

有这样一种说法：“能为者，为之；不能为者，教之。”有数家公司要求其供应商的产品具备较高的 C_p 值，而它们自己却不进行测量或者达不到它们对供应商提出的要求。三家大汽车公司在这一点上声名狼藉。在 QS 9000 标准下，它们要求第一层汽车供应商产品的 C_p 值最低为 1.33（正准备修订 QS 9000，要求最低 C_p 值为



1.66 甚至是 2.0)。然而，大多数它们自己的汽车制造厂不测量 C_p 值或者达不到 C_p 值 1.33。在供应商行业内，福特公司以如下说法而闻名——“照我说的做，不要照我做的做！”

► 小组练习 1：折弯机

装配一个折弯机以生产一个尺寸为 3 英寸 ± 0.005 英寸的成型零部件。工序能力分析表明，加工界限为 3.002 英寸 ± 0.006 英寸，即最小 2.996 英寸，最大 3.008 英寸。修正之后，加工界限处于 3.001 英寸 ± 0.002 英寸的控制之内。

问题

- 问题 1：计算旧工序能力的 C_p 和 C_{pk} 。
- 问题 2：计算修正后的 C_p 和 C_{pk} 。

答案

- 问题 1：规格宽度 (S) = 0.010 英寸；加工（尺寸）宽度 (P) = 0.012 英寸

因此 $C_p = S/P = 0.010/0.012 = 0.833$

$\bar{X} = 3.002$ 英寸；设计中心 (D) = 3.000 英寸

因此 $K = (\bar{X} - D) / (S/2) = (3.002 - 3.000) / 0.005 = 0.002/0.005 = 0.4$

- 问题 2：规格宽度 (S) = 0.010 英寸；加工（尺寸）宽度 (P) = 0.004 英寸

因此 $C_{pk} = (1 - 0.4) 0.833 = 0.5$

$C_p = S/P = 0.010/0.004 = 2.5$

$\bar{X} = 3.001$ 英寸；设计中心 (D) = 3.000 英寸

$$K = (\bar{X} - D) / (S/2) = (3.001 - 3.000) / 0.005 = 0.001 / 0.005 = 0.2$$

$$\text{因此 } C_{pk} = (1 - K) C_p = (1 - 0.2) 2.5 = 2.0$$

使用更简单和备用的公式计算 C_{pk} ：

$$\text{问题 1: } C_{pk} = (3.005 - 3.002) / 0.006 = 0.003 / 0.006 = 0.5$$

$$\text{问题 2: } C_{pk} = (3.005 - 3.001) / 0.002 = 0.004 / 0.002 = 2.0$$

小组练习 2：供应商的公差与成本比较^[1]

一家制造商向其供应商订购一种零部件，公差为 $\pm 5\%$ 。制造商的管理层指示采购部门要寻找到最低价格。经与供应商协商，后者同意如果公差增大到 $\pm 10\%$ ，每个零部件的价格可以降低 10% ；如果公差增大到 $\pm 20\%$ ，价格可降低 25% 。供应商的基本原理和加工（尺寸）分布如图 4—4 所示。

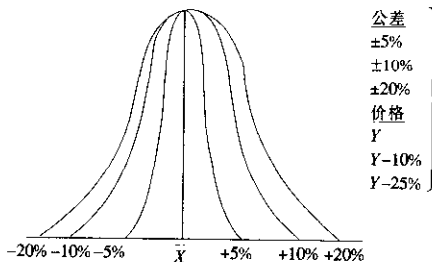


图 4—4 供应商公差与成本比较

但是工程部门坚持 $\pm 5\%$ 的公差以满足客户的要求，即与客户要求一致。应该采取什么样的措施呢？

1. 以较高的价格接受 $\pm 5\%$ 的公差？



2. 向工程部门挑战以证明 $\pm 5\%$ 是基本的?
3. 不理睬工程部门, 以 10% 的折扣接受较高的公差值 $\pm 10\%$?
4. 不理睬工程部门, 以 25% 的折扣接受较高的公差值 $\pm 20\%$?
5. 任何其他措施?

答案

解决采购部门的两难问题有两种方法。

首先, 挑战任何工程上的规格和公差总是可以的。应该执行散布曲线图(见“散布图”章节)以确定能够容纳客户建立的规格公差的最宽的参数公差。

其次, 假定工程部门给出的 $\pm 5\%$ 的公差是正确的, 图 4-4 清楚地表明供应商产品的变量太大。为客户提供更严格公差的惟一方法就是通过测试和筛选、检测和修正、报废和返工——全部都是不增值运作。

制造商的解决办法就是深入供应商, 培训它们 DOE 技术并与它们一道就该零部件制造方面共同工作, 这样就会轻易获得 $\pm 5\%$ 的公差。之后, 公司在成本上的降低不止 25% , 而是 $30\% \sim 35\%$ 。基本原理就是供应商能够消除报废和返工部分, 也就大大减少了检查和测试过程, 减少了循环时间和制造总量, 减少了要求的空间, 提高了生产能力, 利润被公正地予以分配。

小组练习 3: 按客户要求选配发光二极管 (LED)

一家供应商的 LED 产品偏差变化非常宽。为了匹配 LED 的该偏差, 客户就把该供应商的 LED 进行分类:

1. 5 个箱柜用于不同的灯的强度。
2. 3 个箱柜用于不同的正向电压—— V_f 。

3. 2 个箱柜用于不同的温度变化。
4. 3 个箱柜用于不同的波长（彩色）。

这就意味着客户不得不拥有 90 个箱子，按照以上的要求从箱柜中挑选合适的 LED！客户和供应商应该采取什么样的措施来消除这种多余的客户规格化要求？

答案

供应商给客户长期带来的这种多余的选配行为不应再继续下去了。在所有四个产品中（或称绿 Y，下一章中将有解释），这些 LED 的偏差太大了。客户必须选择另外一家更好的供应商。如果这种 LED 的先进技术没有更好的供应商了，那么客户必须培训供应商 DOE 技术。然后，它们必须在供应商的设施处对每一种绿 Y 执行 DOE 试验以减小偏差，使生产的 LED 具有相同的强度、正向电压、温度和波长——这些质量特性的每一个的 C_p 值都不得低于 2.0。最终的结果应该是给客户的价格降低了，而且与现有的 90 个箱柜相比，也就需要一到两个。

顶层管理者需要考虑的问题

1. 规格是否真的是从客户要求中衍生出来的？
2. 目标值（设计中心）是否是减小变量的长期目标？
3. 最低值 $C_{pk} 2.0$ 是重要参数的目标吗？
4. 仅对 C_p 、 C_{pk} 做一至两次测量是否就可以确定一个生产能力？
5. 如果有生产能力，下一步是 SPC（预先控制）吗？
6. 如果没有生产能力，要采取 DEO 吗？
7. 在供应商的图纸和规格上， $C_{pk} 2.0$ 是指定的最低值吗？
8. 在减小变量方面是否为供应商提供了积极、实在的帮助？



9. 作为上述帮助的回报，在供应商获得更高利润的同时，客户能要求并获得成本的降低吗？

【注释】

[1] 小组练习的内容曾经是该公司向笔者提出的问题。DOE 小组与供应商一道成功地执行了笔者建议的解决方法。





第5章

偏差：行业的通病

◆多种偏差源

第4章通过介绍 C_p 和 C_{pk} 量化了偏差，但是测量偏差并不能减少偏差。在系统地解决偏差之前，应该先了解一下工业内部造成偏差的多种来源。偏差在工业界非常普遍，以至于成为一种通病，因此新的质量任务就是在产品和制造中解决偏差问题。实际上，对于所有技术人员——研发工程师、制造工程师、质量专业人士、技术人员以及维修人员，其主要任务是系统地减小偏差。

本章将概述偏差的诸多来源、主要原因以及减小偏差的一般途径。表5-1概述了偏差的来源、原因及减少方法。偏差的主要



来源可分为七类：(1) 管理差；(2) 产品/生产规格低；(3) 部件规格低；(4) 质量体系不健全；(5) 实际生产能力弱；(6) 供应商材料差；(7) “操作员”的错误。

表 5—1 偏差来源、原因及减小方法

来源	造成偏差的原因	减小偏差
管理差	• 不了解减小偏差方面的知识或策略 • 没有将资源和时间分配到 DOE 上 • 没有展开 DOE 竞赛或投入到 DOE 中 • 未培训或未履行 DOE • SPC 和控制图，特别是在解决问题方面	• 加强管理层对 DOE 的培训 • 成立 DOE 管理层指导委员会 • 对技术人员进行 DOE 培训和分组讨论 • 监控 DOE 过程而不仅是设定目标和检查结果
产品/制造规格低	• 没有倾听客户的意见 • 销售而不仅是营销 • 促进现代化的设计 • 宽公差与目标值的比较 • 可靠性不是一个技术条件 • 系统试验中没有 DOE	• 部署质量职能并广泛征求客户意见 • 比较渐近的与变革的设计 • 目标值以及获得目标值的 DOE • 优化老设备而不是使其变为废品 • 针对可靠性进行多环境强化应力试验 • 将 DOE 推广到客户应用上 • 全面生产维修
部件规格低	• 迷恋技术 • 不加选择和过分苛刻的公差 • 样板规格；供应商公布的规格 • “蒙特卡罗”和最坏情况分析 • 没有与偏差相关的公式、错误或不能确定相互作用和结果	• 在试运行阶段采用 DOE 技术将重要偏差与不重要的区分开 • 实际公差的平行四边形散布曲线图

续前表

来源	造成偏差的原因	减小偏差
缺乏质量工具和体系	• ISO 9000、QS 9000、马尔科姆·鲍德里奇奖 • 欧洲质量奖以及 TQM 都缺乏 • 忽视质量辅助设备	• 建立一个“大 Q”质量体系的基础设施 • 正向控制过程论证、预控制
实际生产能力弱	• 标准操作程序：不足/太困难 • 预防维修不够 • 没有环境控制 • 缺少计量制	• 工人参与标准操作程序的制定 • 全面生产维修 • 见第 19 章
供应商材料差	• 供应商太多 • 以协商和图表的办法进行控制 • AQL (合格质量水平), 引入的验收方法	• 与“同类中最佳”的供应商合伙 • 直接接触, 持续帮助 • DOE 培训 • C_{pk} 最低值为 2.0
“操作员”的错误	• 指导、培训差 • 制造、材料、设备差 • 对无制造能力进行规划 • 外观检查 • “双方出现的故障”	• 加强 DOE 培训 • 鼓励、支持与管理 • 自行检查和 poka-yoke (防差错) • 收获共享 • 授权

管理差

戴明和朱兰 (Joseph M. Juran) 声称, 85% 的质量问题是由管理造成的, 只有 15% 是由生产线工人造成的。他们对于管理造成的质量问题所占比例估计得过低, 实际的比例应该是 95% 对 5%。尽管由管理造成的质量问题的原因有许多, 但是以下原因是在偏差问题上遗漏了一些该做的事。

- 缺乏有关偏差对质量、成本、周期时间和利润的影响的知识。
- 在减少偏差方面的政策不一致。



● 没有将资源或时间分配到 DOE 上，而是把大量的资源和时间花费在“救火”上。

● 在减少偏差的目标设定、资助、竞赛以及参与方面没有统一的领导。

● 没有 DOE 培训或实施时没有对培训进行跟踪。

● 使用 SPC 和控制图推进质量均一。

正如福特公司的工程和制造副总裁威廉·斯科拉德（William Scollard）的内部对话所诠释的那样，一个文明的管理哲学体系是截然不同的：“我们新的质量思路应该是，把减少围绕标称的加工变量作为不断改进质量的运行哲学。”

在管理层着手解决偏差问题之前，首先需要了解减少偏差的经济性。这就需要高层的管理委员会发动技术人员参加 DOE 的培训、分组讨论以及“从实践中汲取经验”。委员会还应该参与到 DOE 过程中以便于改进，而不仅是高喊响亮的口号和官僚地跟踪结果。

产品/生产规格低

多数产品的规格不是不清楚、武断，就是错误。有时候还会遗失某个重要的质量特性的要求，其适当程度也从未被怀疑过。加工技术规格甚至更差！

造成这种偏差的一个主要原因在于销售和市场营销的不同。销售时，管理层或工程师孤立地决定产品的要求，然后通过华而不实的广告和其他高压策略迫使客户接受。而在市场营销方面，公司首先经过艰苦的努力去发掘客户的期望，然后设计产品以满足那些需要。在朱利叶斯·罗森沃尔德（Julius Rosenwald）和罗伯特·伍德（Robert Wood）奠定了市场营销的基础，并且在此基础上使西尔斯·罗巴克（Sears Roebuck）成为一个大商店之后 80 年的今天，多数美国人仍然只重产品的“销售”而不重“市场营销”实在是可悲。这种偏差源最致命的一点就是有效地设计并制造出了产品，甚至是零缺陷，但却不是客户所需要的！

必须用“客户的心声”代替“工程师的意见”。只有如此，才能通过诸如质量功能展开（QFD）和广泛征求客户意见这类方法，将前者转化为有意义的产品规格。

产品/生产规格低的其他原因还有：

- 工程师在创造永远铭刻其名字的先进的设计时表现出的自负。

- 使用宽的规格界限和公差，而不是集中在目标值或产品/工艺参数的设计中心。

- 不常使用可靠性——按照平均故障间隔时间（MTBF）或平均辅助间隔时间（MTBA）——作为一种技术规范。

- 缺少用 DOE 方法对客户的操作进行测试的系统，以识别重要的和相互影响的变量。

减少偏差的纠正方法有：

- 使用渐进的而不是变革的方法进行产品/生产设计，在给定的时间内对设计的改动量不要超过 1/4；

- 建立目标值；

- 通过 DOE 节省并优化旧的生产加工设备，而不是将其废弃，有利于将资本集中到新设备上；

- 使用多环境强化应力试验作为在可靠性方面减少产品/加工偏差的最有利工具；

- 将 DOE 扩大到客户基地进行现场试验。

部件规格低

即使假设产品规格已经如上文所述被优化了，还有另外一个错误，那就是没有充分地把产品规格转换为部件规格。原因如下：

- 工程上对技术过于迷恋。

- 工程上倾向于苛刻的公差。

- 工程上依赖以前的部件图纸、样本条件或供应商公布的规格。

- 依靠计算机确定部件公差。如果控制产品（或依赖变量）和



独立部件变量之间的关系的公式已知，那么使用计算机是可以的。但是在许多复杂的设计当中，包括独立变量的计算以及推导一个公式几乎是不可能的。这是许多蒙特卡罗模拟练习中的一个主要弱点。

- 对在实际情况中发生概率极低的最坏情况进行分析和设计，这会明显地增加成本而不增值。

- 尽管有数学公式可以了解偏差之间的关系，但是没有方法去获悉两个部件偏差或两个以上部件之间的未知的、叠加的相互影响。

通过完好设计的试验，在样机、工程试运行阶段或产品试运行阶段能够克服所有这些过多的变量。第3篇中有对全部细节的说明，并指出了重要变量及其理想值（或水准）和实际配合公差。

缺乏质量工具与体系

传统质量工具的缺乏已经在第3章中讨论过。表5-2列出了传统质量工具和三种DOE工具的有效性（1=最差，100=最好）。从该表中可以看出，谢恩DOE的有效性至少是传统工具的20~40倍，是经典工具和田口DOE工具的3~5倍。同样，各种质量体系的缺乏在第2章中也有阐述。

表5-2 各种质量工具的相对有效性

质量工具	有效性度量尺度
QC7种工具	3
7种管理QC工具	2
SPC	5
DOE	
经典	30
田口 (Taguchi) DOE	20
谢恩 (Shainin) DOE	100

注：度量尺度：1=最坏；100=最好。

实际制造能力差

这一类是非常重要的，以至于我们用了整章篇幅来讲述生产认证。实际上，我们强烈建议生产认证的“清查”——一种带有补救措施的审查——应在 DOE 项目开始和 DOE 研究结束时进行。

供应商材料差

紧随设计之后，供应商材料的差异也是造成大多数不良质量的原因。同一个零部件拥有多家供应商以保证质量、交货、成本的传统方法，已经过时并且有碍生产。协商、按图表办法以及通过命令和遥控以求改进质量也遭遇了同样的命运。面临同样处境的还有接受质量水平（AQL）、抽样计划、质量审查和外部检查。改进供应商质量并减少偏差的惟一方法就是遵守以下步骤：

- 确保供应商既承诺改进又能成为你长期的合作伙伴；
- 首先证明你的公司总体来说在质量领域特别是 DOE 方面非常专业；
- 选择一家离你较近，足够小又渴望接受你的 DOE 技术培训的供应商；
- 指定重要参数的 C_{pk} 值在 2.0 以上，绕过不必要的 AQLs、ppms、抽样计划和事后外部检查；
- 在质量、成本和周期时间的改进方面提供积极而具体的帮助，以求持续降低价格。

“操作员”的错误

用户的变动和矛盾通常被传统的管理层作为质量问题的原因。但是这种引证总是反映了对于质量的无知——管理层的质量意识最差，工人的错误只是其结果。主要原因可能是以下几点：

- 缺乏指导、目标、培训和监督；



- 生产、材料和测试设备差；
- 对于生产能力的设计差；
- 使用外部检查作为支撑；
- 仅把工人看作不用头脑的“双手”。

当清除这些有碍质量的障碍之后，99%的工人都开始被激发出热情并一如既往地生产出高性能产品。一般地说，在这些范围内，鼓励、支持、排除害怕情绪以及管理层与工人的积极互融，有助于减少产品的偏差。也要认真考虑改进性能对经济的刺激，例如利益分享。

在特定条件下，从自检、外部检验、到互检直到最后检验，可以借助半自动的 poka-yoke（防差错）方法，例如使用自动化设备和传感器装置的目视检验。重要的是工人们没有采用被日本生产线工人广泛应用的质量控制 7 种工具，而是采用在第三篇叙述的较容易的 DOE 工具。我们的经验是：某些更冒险和更值得做的 DOE 工作是由经过培训的生产线工人和技师完成的，而不是那些更谨慎、保守的工程人员。



第6章

试验设计的三种方法：经典方法、 田口方法与谢恩方法

无效的解决问题和减少 变量的方法

第4章测量了变量，第5章概述了变量的起源，但是仅测量变量丝毫不能减少变量。以节食的人们打个比方，为了与肥胖作斗争，成千上万的人每天都要称体重；但是如果仅靠测量体重就能达到目的，美国人将会是地球上最瘦的人！减肥都这样困难，减少变量就更是如此。

曾有一种传统的减少变量的方法流行于世，但它们均告无效。

1. 工程判断已经应用了将近一个世纪。但是，如果它是有效的，为什么我们



仍有很多长期的质量问题呢？这些问题有的持续数周，有的持续数月，有的持续数年甚至数十年！有些问题存在的时间长得可以参加选举，我们的一个客户的质量问题甚至可以申请社会福利！

2. 计算机仿真是解决问题方面的最新时尚。较早有蒙特卡罗仿真、E—芯片以及许多其他的软件程序，它们都被声称有解决根本原因的办法。但是，要使计算机有效，必须已知控制自变量与因变量之间关系的数学方程式。如果该方程式或公式还未知，就无法为计算机编制程序。遗憾的是，对于许多复杂的产品和过程，甚至连爱因斯坦式的人物也不能推导出这样一个方程式。例如，有些科学家用他们的数学模型预言：今后 100 年内温度的上升将带来全球变暖的毁灭性灾难；另外一些科学家用他们的“水晶球”模型预言：在同一时期，地球将进入另一个冰河期。所有这些计算机仿真共同预测的可能结果是，我们地球人在未来 100 年内仍将过着舒适的日子！

3. 统计过程控制 (SPC) 已被看做改进质量的重要工具。但是，SPC 只是一个监测工具而不是解决问题的工具。就像在某个过程的上方悬挂一个计量表，计量表只是一个可视的指示器，它并不解决问题或改进过程。SPC 只应在问题已经得以解决之后使用，而不应在问题得以解决以前使用。

4. 福特 8-D 已在第 3 章中加以详述，它至多只是为解决问题提供一个管理的框架。令人遗憾的是，它缺乏“如何”解决问题的具体办法，没有提供解决问题所需的专门工具。至于著名的 PDCA 循环，它是一个更为原始的框架。

5. 凯普纳-特拉格。有些公司已经掌握由凯普纳-特拉格的促进者们教授的所谓的侦测方法。凯普纳-特拉格也许是写侦探小说时可以接受的一种方法，但它不能代替解决工业问题的方法。



试验设计三种方法的背景

所以，只有试验设计才是解决长期质量问题和减少变量的最佳手段。它有下列三种方法。

经典 DOE

这一方法基于罗纳德·费雪（Ronald Fisher）爵士的开拓性工作，早在 20 世纪 20 年代他就把 DOE 应用到农业领域。很难想出一个具有与农业同样多的变量的应用。农业中的变量有土壤、雨、水、太阳、气候、种子、肥料、地势等，但是，费雪仅仅用了完全析因法就提高了英国农田的生产率，并由于他的巨大贡献而被授予爵士爵位。事实上，费雪利用 DOE 减少了变量，是现代质量控制之父。而现在被誉为质量控制之父的沃尔特·休哈特（Walter Shewhart）则是费雪的学生。

费雪的继承人把费雪浓缩的完全析因法冲淡成微弱得多的析因法——称为经典方法，但是它被限定在农业和化学工业领域。直到最近，在这些应用之外，它仍然一直被保留在学术界范围内。大学教授们胡乱创造出一点儿枯燥的析因法，并不期望他们的学生将其用于“现实世界”。

田口 DOE

日本的田口采用了经典的方法，用他自己的正交表将其简化。然而，他在自己的国家并没有取得多大成就。事实上，日本科学家和工程师联合会（日本主要的培训机构）的执行理事对这位创始者的评论是：“田口是出口商品的”。田口在美国则受到热情的欢迎。他免费向美国电话电报公司的贝尔实验室提供服务，以作为对美国曾经给予日本帮助的回报。于是在 20 世纪 80 年代出现了对田口的



崇拜，美国电话电报公司、福特、施乐以及 ITT 等公司都成为田口的宣传者。成千上万人怀着“如果它是日本的，它必定是好的”的盲从心态，涌进田口研讨会。但是田口方法在最好的情况下是取得了适度的成功，而在最坏的情况下则是彻底失败。本章将在后面进行解释。

谢恩 DOE

第 3 种方法是由美国多利安·谢恩（Dorian Shainin）发明或者说完善的一些简单但强大的方法。谢恩是 800 多个最主要公司的顾问。谢恩无疑是世界上第一流的质量问题专家，他是美国人的偶像，曾经获得过美国质量协会颁发的所有主要奖金和奖章。事实上，在摩托罗拉公司有这样一句名言：“没有戴明，美国就不会有质量哲学；没有朱兰，美国就不会有质量的方向；没有谢恩，美国就无从解决质量问题。”这句名言总结了美国最伟大的三位质量大师的贡献。

遗憾的是，谢恩的这些方法没有得到应有的充分推广和广泛应用，因为应用了这些方法并体验到卓越成果的那些公司不愿意与别人分享这些方法。事实上，这位创始者曾获准在其第一部也是惟一的一部关于这些方法的教科书中发布了这些方法，因为他所在的公司——摩托罗拉公司获得了马尔科姆·鲍德里奇国家质量奖^[1]，该奖规定它所采用的方法要与其他美国公司共享。

◆ 摩托罗拉案例研究：10 : 1、100 : 1、1 000 : 1 的质量改进

20 世纪 70 年代，摩托罗拉面临着来自日本的激烈竞争。它们在我们所处的商业领域中无处不在，并正虎视眈眈地想吞食我们的市场份额。“迎接日本人的挑战”成为我们重整旗鼓的口号。摩托



罗拉的董事会主席鲍勃·卡尔文确定了质量就是这次挑战的焦点。

1981年，他制定了五年内10:1的质量改进目标。摩托罗拉以往的质量记录是良好的，并一直以每年10%的速度增进。现在，鲍勃·卡尔文的挑战是五年内不是50%或100%，而是1000%地提高质量。许多怀疑者认为这是一个不可能达到的目标。但是，到1986年，摩托罗拉的大多数部门已经达到了这一目标。这位时任摩托罗拉自动化部门的集团理事的创始人，三年内取得了10:1的质量改进。事后，我们拿自己与日本人进行比较，发现他们仍在质量方面占据优势。所以，鲍勃·卡尔文在1987年制定了另一个10:1质量改进目标，但这次的实现日期是二年后，即1989年。1989年，他又提升了质量目标的高度，即在二年之后——1991年，再取得一个10:1的改进。所以，从1981年开始，摩托罗拉必须取得1000:1的质量改进。就摩托罗拉整体分布很广的业务而言，这一目标并未完全达到，但是平均质量改进从1981年的良好基础开始已经达到不可思议的800:1。

聚宝盆

许多传媒界人士百般挑剔摩托罗拉对质量的“执著”。他们得意地笑道：“对于这么巨大的质量改进，难道不会产生繁重的经费负担吗？”而事实是，质量并不耗费成本，它有所回报，而且是慷慨的回报。自1979年以来，摩托罗拉跟踪了它为质量低劣付出的成本（见第4章）。10年后，通过减少为低劣质量付出的成本，摩托罗拉节约了90多亿美元！这是一个聚宝盆，它使摩托罗拉以较低价格的形式把节约的一部分资金转移到忠实用户手中，一部分资金转移到雇员身上（业内最高薪酬），另一部分则以13年后股票增值24倍的形式转移给了股东。几年后，一家主要的新闻期刊问道：“卡尔文先生，你已经带领公司勇攀许多质量高峰并获得很多荣誉，包括马尔科姆·鲍德里奇奖。那么在你追求质量的历程中，你最感遗憾的是什么呢？”鲍勃·卡尔文的惊人回答是：“我定的目标还不



够高!”

鲍勃·卡尔文的激励式领导风格

摩托罗拉之所以取得惊人的质量成就有许多原因。成功有多位父亲，失败则是一个孤儿。首先，最主要的是卡尔文的领导及其对提高产品质量的不懈努力。没有他的远见、激励以及魅力，我们可能已陷入平庸改进的泥沼。他领导了质量改革，但并没有规定用以达到这些质量高度的方法。他坦诚地承认他并不知道怎样达到他制定的目标，但是他对属下怀有始终不渝的信心，通过带领他们进入以前从未有人涉足的领域，通过为他们指引方向，他就能激励他们开拓出通向成功的路径图。

工具

获得成功的第二个原因是我们拥有工具。这些工具已在第3章中概述过，其中最重要的是试验设计。早在1982年，培训和教育部门（现为摩托罗拉大学）的副总经理就找到这位创始人，商量有什么办法可以实现艰巨的10:1质量改进目标，他知道惯常的改进方法胜任不了这一任务。他说，我们需要突破旧有模式。我提议应该采用谢恩的DOE，不是像过去那样零碎地、间断地采用，而是全面地采用。于是，DOE时代开始了。

DOE在阿凯德试点

两个月之后，首次谢恩DOE试点在摩托罗拉设在纽约阿凯德的工厂进行。从技术人员和一线工人中抽调的60名人员接受了为期3天的DOE培训。研讨会结束后，培训人员被分成了12个组，分别解决12个主要质量问题。两个月以后，我们检查了他们的结果，而更重要的是检查他们使用的DOE方法。各组均完成得很好，



但任务完成得最好的组，其人员全部是从一线岗位抽调出来的，他们仅在一个试验中就将厚膜基片的缺陷率从 12.5% 降至 0.5%。^[2]

这是该创始人第一次可以加强他的这样一个信心：DOE（谢恩）不必是研发工程师和过程工程师们的专属领域。一线工人们对新概念和新技术的理解力和接受程度总是为管理层所忽视。我至少在 12 家一线工人广泛使用 DOE 的公司中得到证实：若给一线工人以培训、鼓励和支持，他们会做得和工程师同样好，也许还要更好一些，因为他们急于解决问题，并且不会由于对技术类型持保守的怀疑态度而停顿。

在美国 5 家工厂的试运行

阿凯德的成功促使摩托罗拉在其美国的其他 5 家工厂进行了 DOE 试运行。它们都获得了成功，但是会计部门对这些项目的高昂成本表示反对。为了消除这种否定态度，倡议这些项目的全员培训和教育部门引进了一个独立的审计公司来监查项目结果。审计部门确实揭示了培训的高成本——328 000 美元，但是总体节约在 700 万美元以上，投资回报比大于 1:20。从那以后，我们再也没有遭到来自 19 世纪的会计部门的批评和攻击！

DOE 扩展到全球

取得成功经验后，我们先后在美国、欧洲和亚洲的所有工厂——总计 50 多家都创建了 DOE。几年以后，摩托罗拉开展了它著名的顾客全部满意（TCS）竞赛。在这一竞赛中，多达 5 000 个团队，即 140 000 名雇员中的 65 000 名职工为不断改进质量、成本、周期时间以及其他关键测量参数而努力。DOE 一直是这些被授权的团队所采用的主要工具之一。公司估计通过全部顾客满意竞赛，每年可节约 24 亿美元。这种节约主要表明：每年高科技产品的市场价格可降低 15%~30%。



经典方法和田口方法的基本弱点

DOE 的这三种方法——经典方法、田口方法以及谢恩方法——都优于工程判断、计算机仿真、SPC、8-D 以及凯普纳-特拉克等解决问题的传统方法。这三种方法也远优于过去常在大学中讲授并仍被传统的工程师们广泛实践的老式试验，在这些老式试验中，一次仅有一个变量变动，其他变量均保持恒定。这种试验除了需要过量的时间之外，它的主要弱点是经常不能把主效应从它们的相互作用效应中分离出来，结果是不断受挫、恶性循环和增加成本。

然而，DOE 的经典方法和田口方法也有其主要结构上的弱点和和其他弱点。表 6-1 概述了十个范畴内各种方法的弱点以及谢恩方法的优点，还给出了用 10 分制评价指标评价经典方法、田口方法和谢恩方法的评分结果。

表 6-1 试验设计的三种方法的比较

特征	经典	评分	田口	评分	谢恩	评分
1. 方法	一或两种方法——分析因法和响应曲面法	3	一种方法——正交表	2	最少 10 种方法（见图 7-1）	10
2. 线索生成能力	差：猜测	1	差：猜测	1	强：“与部件对话”	10
3. 效果	• 中度改进，范围 2:1 至 10:1 • 有可能倒退	3	• 低度改进，范围 2:1 至 5:1 • 有倒退的可能	2	• 高度改进，范围 5:1 至 100:1 • 不退化	10
4. 成本	• 高：30~60 次试验	4	• 高：50~100 次试验	2	• 低：2~30 次试验	8



续前表

特征	经典	评分	因口	评分	谢恩	评分
5. 复杂度	- 困难的方案 - 全面的方差分析要求	2	- 困难的方案 - 内表和外表相乘: 方差分析, 信噪比	1	- 容易的、简单的、合乎逻辑的方案 - 由一线工人和工程师进行实验	8
6. 时间 (a) 了解(b) 实践(做)	(a) 长: 3 天或 2 个星期 (b) 长: 思路不充分, 意味着全面再启动	2	(a) 长: 1~2 个星期 (b) 长: 思路不充分, 结果相差大, 需数次试验	1	(a) 短: 1 天 (b) 短: 实验在 1 天至 3 星期内完成	9
7. 统计 有效度	- 弱 - 混淆主要作用和相互作用	2	- 十分弱 - 未随机化 - 混淆主要作用和相互作用	1	- 强 - 清晰地分离和量化主要作用和相互作用	8
8. 适用性	需要硬件主要用于生产	2	可与计算机仿真一起用于论文研究, 但有可能得到错误结果	4	- 需要硬件 - 可以通用在产品/过程设计、生产、现场、供应商和行政管理过程	8
9. 执行的 难易度	- 难 - 如果相互作用强烈, 会得到错误结果	2	- 难 - 由于分级过多, 很可能产生错误结果	1	- 易 - 有了线索生成方法, 不需要重复实验	9
10. 干扰生 产程度	实验期间停止生产	1	实验期间停止生产	1	在思路创新实验中不停产, 线索生成实验能解决 70% 的长期质量问题	8

* 1~10 的度量尺度: 1—最差, 10—最好。



1. 方法：经典/田口方法有局限性且无效；谢恩方法用途多且有效

- 经典 DOE 绝大部分仅利用分析因法（虽然在经典的和谢恩的 DOE 中均采用全析因法）来解决问题并用响应曲面法来优化。

- 田口 DOE 只采用正交表（级数）——参数的内表和噪声的外表均是如此。在经典方法和田口方法中，如果这些单一的路径失败了，你就处于困境，别无出路。

- 谢恩 DOE 采用了十种截然不同的方法，每种方法适合一类特殊的问题或用途。如果一种路径不适合，还可以另辟许多其他路径。

2. 线索生成：“部件比工程师更聪明”

这是谢恩方法和经典/田口方法最重要的差别之一。

- 在经典/田口 DOE 中，工程师和团队猜测问题可能的原因，他们采用头脑风暴法和举手表决的方法确定最可能的原因。这叫做民主解决问题！如果猜测错了，试验就失败。

- 在谢恩 DOE 中，“我们和部件、工艺对话。它们处理问题比工程师更聪明”。这不是贬低工程师，我们需要他们的知识，但是，我们要求他们搁置他们的猜测、直觉、见解、偏见和理论，直到他们已与部件对话——意指他们必须首先从四种线索方法中获得有效且合理的线索，这些线索比猜测更为实际，并具有远远高于猜测的侦测能力。

3. 效益：经典/田口方法成功率低，谢恩方法的成功率高

- 经典 DOE：如果不存在强烈的交互影响效应，就有可能得到适度的质量改进，其范围为 2:1 至 10:1，平均为 4:1。但是因为它没有随后的有效性确认试验就可能会返回到原来的老毛病。



进一步的防护措施如正向控制和过程确认等未知。

- 田口 DOE：如果不存在强烈的交互影响效应，可得到适度改进，其范围为 1.5 : 1 至 5 : 1，平均为 2 : 1（田口 DOE 比经典 DOE 更差，原因是设计试验存在较大的离散）。但是，由于缺乏有效性确认试验，就可能会返回到原来的老毛病。同样，进一步的防护措施如正向控制和过程确认等未知。

- 谢恩 DOE：取得一个量级（10 : 1）的改进是很平常的事。平均改进是 20 : 1、50 : 1、100 : 1，甚至 1 000 : 1 的改进也是能够达到而且已经达到的。由于有了“B 与 C 比较”鉴定试验以及正向控制和过程确认等防护措施，就避免了倒退。

4. 成本：经典/田口方法中失败的高成本与谢恩方法中成功的低成本比较

- 经典 DOE 利用分析因作为初步筛选工具。如果试验不成功，必须重做。采用主因子和相互作用因子的组合可能要试 30~60 次，从而增加了费用。此外，由于生产总是被打乱，就会造成与生产线停工有关的巨大损失。

- 田口 DOE 的成本更高，原因有二：第一，需要反复试验而带来的分级过多；第二，内表和外表必须一起相乘，大大地增加了试验的总次数。和经典 DOE 一样同样要打乱生产。

- 谢恩 DOE 就样本规模和所需试验次数来说是最经济的。因为“与部件对话”优于猜测，所以它对原因的诊断要准确得多，很少需要进行反复试验。最重要的是，它不打乱生产——这是思路创新的一个明显的优势。

5. 复杂程度：经典/田口方法中很难理解的统计术语，谢恩方法中简单的逻辑概念

经典 DOE 需要充分的方差分析（ANOVA），甚至经过严格的



培训之后，工程师仍很难理解它，因而很少施行。试验中的因子是猜测和直觉的。田口 DOE 不仅对内表因子需要进行猜测，对外表因子和它们的相乘也需要猜测。它还需要 ANOVA 以及进行信噪比 (S/N) 计算和图形分析。因为太过复杂，因而大多数工程师对它大感头痛。

谢恩 DOE 则相反，它很简单，因而一线工人都能理解并使用它。(有时我们开玩笑说，即使工程师也能理解和使用，甚至没多久经理也能理解和使用)。

6. 教授时间：经典/田口方法研讨会需 3 天至 2 周，谢恩方法仅需 1 天；实施过程：经典/田口方法开始易错，谢恩方法思路清晰

●就理解而言，经典/田口方法的研讨会需要 3 天至 2 周。即使这样，工程师们还会目光呆滞，毫无理解能力。就使用情况而言，即使假定他们已开始启用（很多人尚未开始），进行试验也还需要很长时间，尤其是当他们猜测的因子往往是条死胡同，因而试验必须重新进行时。

●我们举办的谢恩方法研讨会只需 1 天就能使参加者理解各种不同的方法。第 2 天，参加者就能实际设计 DOE 试验，开始解决他们自己的长期的质量问题。所需时间也短得多，因为从与部件对话中获得的思路非常可靠。

7. 统计有效性：经典/田口方法中主效应和交互影响效应混淆不清并且处于饱和状态，谢恩方法中二者明确分离

从统计学观点看，经典/田口 DOE 方法的主要弱点是主效应和二阶、三阶以及更高阶的交互影响效应严重混淆不清（对上述情况的统计和技术处理感兴趣的读者，可参见第 13 章提供的详细说

明)。(严重混淆不清或者用专业术语“污染”。)

●在经典 DOE 中,当自变量的全部可能组合只有 3%~25% 成为试验的组成部分时,就会发生分级过多的情况,我们就会得到临界的或者完全错误的结果,因为主因子效应不能与交互影响效应相隔离。经典 DOE 学派试图这样解释这一弱点:在分析因试验中,试验分离只是一种筛选机制,应当接着进行第二个实验;并且这种方法如果不能使更高阶纯交互影响,至少可使二阶纯交互影响。这两种说法都是错误的。如果试验的基本设计是错的,试验的结论多半也是错的。其次,它没有办法把二阶交互影响效应与主效应分离。这种设计方法的基础本身是脆弱的,其结果犹如建立在沙滩上的大厦。

●在田口 DOE 中,所发生的试验离散甚至比经典 DOE 中的更加显著,更加有害。例如,如果有 10 个因子需要考查,每个因子有高、低两个水平,在全部析因设计中,每一因子必须与其他的每一个因子的每一个水平一起进行试验。这种组合将需 2^{10} 即 1 024 个组合或 1 024 个试验。田口的一个正交表只能抽查 8 个或者 16 个这样的组合,那么,你选哪些组合呢?田口派说:“选最重要的”。但是如果你不知道哪些是最重要的因子,你怎么能够确定那些未知因子的最重要组合呢?这好比玩俄罗斯轮盘赌,甚至比玩俄罗斯轮盘赌更糟糕。因为,在俄罗斯轮盘赌中,你只有 1/6 的机会一次输掉,而正交表,几乎每一回都有一次输掉的机会。

在谢恩 DOE 中,经过“一线工人告诉你哪些是重要的因子”的线索生成试验以后,全部析因或变量搜索可在一次析因分析中干净利落分离主因子、二阶及更高阶的交互影响,并可以进行量化(见第 13、14 章)。

8. 适用性:在产品或过程的设计阶段,在生产、供应商、使用现场,甚至在行政管理应用中,谢恩方法最佳

●经典 DOE 必须需要硬件。它不用来作初步的论文研究,而



主要用于解决生产问题。

- 田口 DOE 已经与计算机仿真一起用作论文研究，但是这种方法是危险的，因为计算机仿真必须要知道决定自变量和因变量之间的相互关系的数学方程式，大多数设计者缺乏这种知识。

- 谢恩 DOE 和经典 DOE 有相同的局限性，即它需要硬件——在设计的样机阶段。但是它在评价新设计方面比其他两种方法具有显著的优点。它干净利落地把重要变量和非重要变量分开。此外，除在产品或过程的设计阶段之外，它在生产中、在供应商中、在用户中、在现场以及在行政管理过程中都普遍适用。

9. 应用难易程度：经典/田口方法——难而慢，枯燥； 谢恩方法——易而快，使人振奋

- 经典/田口方法：这种方法需要很多准备时间。必须停止生产，必须收集每个因子的两个水平样本量。此外，如果结果不好——因为错误的猜测或交互影响效应与主效应混淆不清——试验必须再重复一次或更多次，又要再进行 8、16、32 或 64 次试验。

- 谢恩方法：因为线索生成方法远优于工程猜测和直觉，这些方法能够很快得到预期的结果（至少 70% 的时间，不需更多的正式试验）。此外，应用它的变量搜索方法，你在一至两次试验中就能确定是否走上了正轨。

10. 干扰生产程度：经典/田口方法——干扰生产； 谢恩方法——不干扰生产

- 经典/田口方法：生产与用户要求的进度及运输有关，工厂不希望因为试验而使生产线停工。在经典/田口方法中，在进行某个试验期间，生产线被迫停工，而第一次试验甚至第二次很可能会不成功。这就意味着用户的愤怒、周期时间的损失以及高昂的库存费用。

● 谢恩方法：至少 70% 的时间不停产，因为只有线索生成方法是对正在生产的产品进行抽样的。

公开的挑战

我们深信谢恩 DOE 的能力，因此我们提出挑战。任何用经典或田口方法能够解决的问题，我们用谢恩方法能解决得更好、更快、更省。

【注释】

[1] 摩托罗拉公司是 1988 年第一家获得马尔科姆·鲍德里奇奖的公司，也是该奖的 11 年历史中，企业界获得该奖的惟一一家大公司。该公司在为来自其他公司的数千人举办著名的六西格玛研讨会上，非常慷慨地与人共享了它们的成果。

[2] 该团队中的某个工人开始在她的工作中经常地使用 DOE，结果她的事业迅速走向成功——她升任了领班、组长、督察员、经理。她现在是不负众望的、非常受人尊敬的规模为 1 500 多人的工厂的经理。





第 7 章

10 种强有力的试验设计工具

概述

我们已在第 6 章中说明，谢恩 DOE 明显有别于经典/田口方法，其截然不同的特点是它在使用 10 种不同的 DOE 工具时的多面性，每种 DOE 工具仅限于解决特定的问题，而在经典/田口方法中仅有一到两种直接套用的方法。

图 7-1 是一个介绍性的框图。它在一个完美地解决问题的连续过程中将 10 种工具连接了起来。开始可能是一个有 1 000 个变量的问题（在本书中，变量、原因、因子是同义词），可以系统地消除掉这些变量，直到提取出 1 个、2 个、3 个主要的变量，

用以采取纠正措施和预防控制。

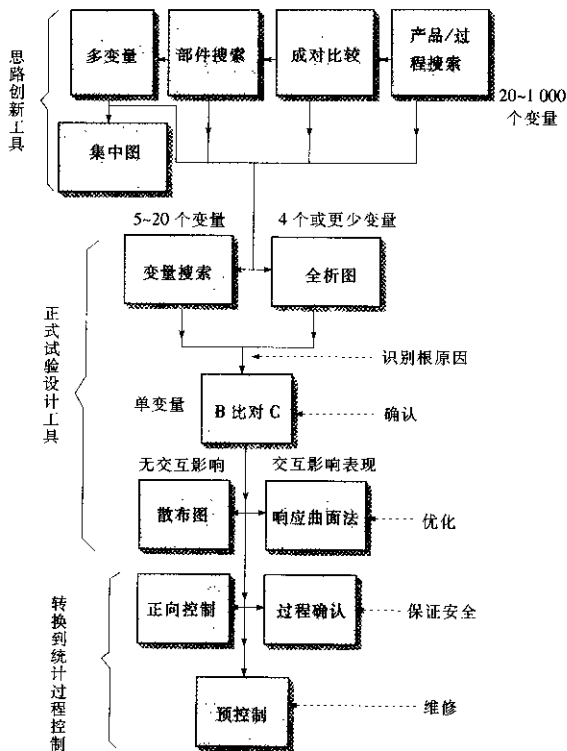


图 7—1 减少变量框图简介



线索生成工具——滤掉不重要的变量

顶部是 4 个主要的线索生成工具——多变量、部件搜索、成对比较和产品/过程搜索以及第 5 个派生的工具——集中图。它们的优点是：（1）它们通过与各部件对话提供思路，这种方法比经典/田口方法的基础——工程猜测、理论、见解、偏见有效得多；（2）它们不打乱生产，让生产继续进行，同时从生产过程中仅抽取少量样本；（3）它们把大量未知的、难以处理的变量——20~1 000 个——减少到少量——1~20 个可处理变量族。

Quality Department

2006.05.19 正式的 DOE 工具——针对红 X

SongYongSheng

从上述有意义的思路出发，我们既可以返回去求助于工程判断，以找出问题的根本原因，也可以采用正式的 DOE。如果仍认为 5~20 个变量值得注意，则最佳的 DOE 方法是变量搜索。如果认为 2~4 个变量重要，则完全析因是优先选用的方法。如果只剩下一个变量，根本原因就知道了，于是只须采用 B 与 C 的 DOE 比较方法，该方法的主要目的是确认纠正措施的有效性和持久性。下一步是优化重要的参数，即它们的理想水平或理想值以及实际的容许偏差。如果没有交互影响效应，则应采用散布图 DOE 方法；如果有交互影响效应，则应采用响应曲面法。

逆向防护措施

在 DOE 完成和生产统计过程控制的开始之间还有两个非常重要的步骤。99% 的公司完全未注意到这些规范，而那些确实知道这些规范的公司中，只有不到 10% 的公司使用了这些方法。这些方法的目的是“冻结”以前 DOE 试验所取得的成果，并防止滑入原来的缺陷中。

（1）正向控制保证重要的过程参数始终保持在其各自的由



DOE 最优化试验所确定的容差范围以内。

(2) 过程确认可保证外部质量问题——例如操作员可控制的问题、计量以及环境因子等——得以提交并确认，以防止 DOE 试验前和试验后产生缺陷。

(3) 统计过程控制：预控制。最后一步是预控制，它比老的控制图更好、更快、更省。老的控制图是目前生产中逐渐被淘汰（或应该被淘汰）的维护和监控方法。

解决问题的总框架

表 7-1 所示为一个 10 步骤解决问题的框架。该框架综合了图 7-1 中所示的全部工具，形成了一个解决某个问题的循序渐进的综合性方法，尤其是长期的质量问题。它远比 PDCA 这个空洞的口号或者 8-D 这个空洞的技术（而它们的拥护者错误地声称“找到了根本原因”）的意义要深远得多。

表 7-1 解决问题的框架：10 步骤法

1. 定义问题（绿 Y）
2. 量化并测量绿 Y
• 测量散布图（而不是圆弧图）
• 李克特度量尺度将属性转换成变量
3. 问题的历史（问题存在的时间、故障率、费用）
4. 创新思路
• 多变量（包括集中图）
• 部件搜索
• 成对比较
• 产品/过程搜索
5. 试验的正式设计
• 变量搜索
• 完全析因
• B 与 C 比较
6. 问题的引出及解决——保证改进的持久性



续前表

- B 与 C 比较
- 7. 建立实际的规范和容许偏差（最优化）
 - 散布图
 - 响应曲面法（RSM）
- 8. “冻结”过程的改进
 - 正向控制
- 9. 确认过程：明确减少所有外部的质量问题
 - 过程认证
- 10. 用 SPC 保持成果
 - 预控制

10 种 DOE 工具的简要介绍

表 7-2 提供了 10 种 DOE 工具的简要介绍——目标是什么、在何处使用、于何时应用，同时列出了在试验中使用的令人难以置信的小样本量。下面的几章中我们将针对表 7-1 中的每一种工具进一步进行解释，以使读者熟悉这些工具。

表 7-2 10 种 DOE 工具

工具	目的	用于何处	何时应用	样本量
多变量图	• 把大量没有联系的原因减少到一族数量较少且相关的原因，例如时间到时间、部件到部件、部件内部、机器到机器、测试位置到测试位置 • 检测非随机趋势	• 确定产品、过程正在怎样运行：没有大量用途有限的历史数据时的迅速处理 • 在一些白领阶层的应用中，取代过程能力研究	在工程试运行、生产试运行、在生产中，甚至在现场	抓取至少 9~15 个或直至 80% 的历史变量
集中图	多变量图的续篇。以单元或部件正确地指出重复的缺陷	同多变量图	当红 X 在单元内	同多变量图

续前表

工具	目的	用于何处	何时应用	样本量
部件搜索	从几十万个部件/组件中, 针对红 X, 捕捉所有重要的主效应和交互影响效应的量	在有两个不同地方用可互换部件组装的场合 (标注“好”和“差”)	在样机、工程试运行、生产试运行或在现场	2
成对比较	以高置信度确定“好”和“差”单元之间的重复差异, 以向红 X 提供线索	- 不能拆卸的产品以不同的方式组装 (标注“好”和“差”) 的配套装置处 - 为许多管理者或白领职员应用	样机、工程试运行、生产试运行或在现场运行	6~8 对“好”和“差”的产品
产品/过程搜索	- 用成对比较发现重要的产品变量 - 发现与 8 个好和 8 个差的产品有关联的重要过程变量	仅用多变量图难以分离重要过程	样机、工程试运行、生产试运行或正式生产	单元数量足以通过一个过程产生 8 个“好”产品和 8 个“差”产品以及它们相关的过程参数
变量搜索	指出红 X、粉红 X 等。捕捉所有重要主效应和交互影响效应的量。放开所有不重要变量的允许偏差以减少费用	- 有 5~20 个需研究的变量处 - 卓越的问题预防工具 - 应用在白领职员工作中	在研究与开发以及开发工程中卓有成效, 在生产中产品/过程表示特征。也用于多变量图表或成对比较之后指出红 X	1~20
完全析因	同变量搜索	仅在 2~4 个变量处有实效	同变量搜索	1~20



续前表

工具	目的	用于何处	何时应用	样本量
B vs. C	• 以期望的统计置信度(通常为 95%) 确认新的或更好的 (B) 产品/过程 优于 现行的 (C) 产品/过程 • 评估工程革新 • 降低费用	• 在以上 7 种工具中的一或多种的引导下进行 • 在某些白领职员的应用中, 当问题容易解决时, B 与 C 比较可以越过以上这些工具 (即可以不用以上工具直接进行 B 与 C 比较) • 可用于两个以上的选择, 即 “B 与 C” 与 “D 与 E” 比较等	样机、工程试运行或生产时	通常 3 个 B 和 3 个 C
散布图(实际容许偏差平行四边形)	• 确定红 X、粉红 X、变量的最优值(水平)以及它们的最大允许偏差	当重要变量之间有微弱交互影响或者没有交互影响时	工程试运行或产品/过程	30
响应曲面法 (RSM)	同散布图	当重要变量间有强交互影响时	同散布图	5~25

绿 Y、红 X、粉红 X、浅粉红 X

谢恩 DOE 的基础建立在一个古老而普遍的理论——帕累托定律之上。帕累托是一位研究意大利经济的意大利经济学家，他发现人们的收入不是平均分配的，少数人的收入总量占社会全部收入总量的绝大部分份额。

朱兰博士将帕累托定律转化到工业，并得到普遍应用。原因和

效果并不是成线性相关的，少数原因造成了绝大部分效果，这就是20/80定律，即20%或更少的原因造成了某个特定效果的80%或更多。帕累托定律在如此多的工业和现实生活的场合中都很正确，真是不可思议。20%或更少的部分占了总成本的80%或更多——此原理适用于价值工程和库存量控制，在管理中，被用于“特例管理”。同样，在质量问题中，20%或更少的原因产生特定质量问题总量的80%或更多。

图7—2所示为帕累托定律。有几个术语将在全书中多次使用。

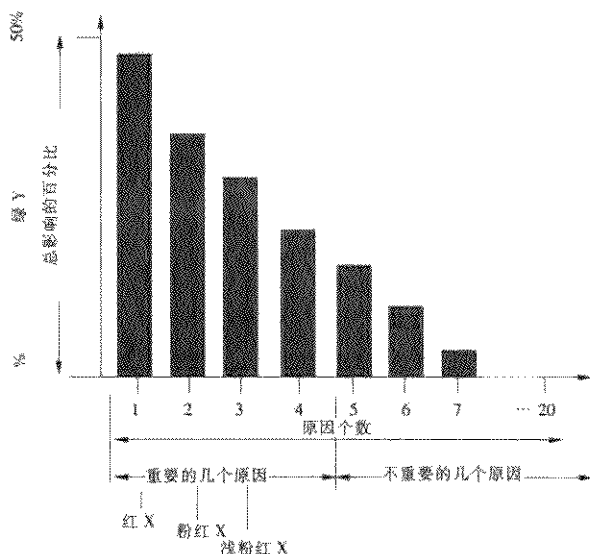


图7—2 帕累托定律：绿 Y 和红 X

注：解决红 X、粉红 X 和浅粉红 X 能够

- 减少变量；
- 用 1、2 或 3 个试验使 C_p 达到 2~10。



●效果被称为绿 Y，它代表必须解决的问题（或输出或响应）的量。

●第一位的或最主要的原因称为红 X，它一般至少占全部绿 Y 的 50%。

●第二位的原因称为粉红 X，它约占全部绿 Y 的 20%~30%。

●第三位的原因称为浅粉红 X，它约占全部绿 Y 的 10%~15%。

通常，用 DOE 找到某特定绿 Y 的红 X、粉红 X、浅粉红 X 的解决办法后，变量可以减少 75%~95%。我们假定开始时 C_p 为 1.0，如果在此过程中变量减少 75%， C_p 即可以达到 4——许多公司努力为之奋斗的世界级标准的 2 倍；如果变量减少 95%， C_p 即可达 20——世界级标准的 10 倍。与经典/田口方法 C_p 仅能达到 2.0 的微弱结果相比较，这正是另一个量化谢恩方法的优点的方面—— C_p 可以达到 20。

图 7-1 的框图显示了各种 DOE 工具的关系。但是，要系统地解决问题，还需要一个更详细的路径图来指导应用人员搜索红 X。图 7-3 就是这样一张路径图，其起点为绿 Y，以解题方式连续运用各种工具，着手提取红 X，然后用预控制进行监视和控制。（在此路径图中，变量搜索和完全析因并列在一起，因为它们的目的相同。变量搜索是一个多用途的工具，因为它能处理 5 个或更多的变量，而完全析因最适合处理 4 个或更少的变量。）

当我们沿着搜索红 X 的路径行进时，必须注意图 7-3 中的每一个步骤。起点是定义绿 Y（阴影区域）。

定义并量化问题：绿 Y 检查单

在 DOE 开始之前，重要的是说明、定义并量化问题，即绿 Y。DOE 工作组应使用下面的检查单作为起点，以查明涉及的所有问

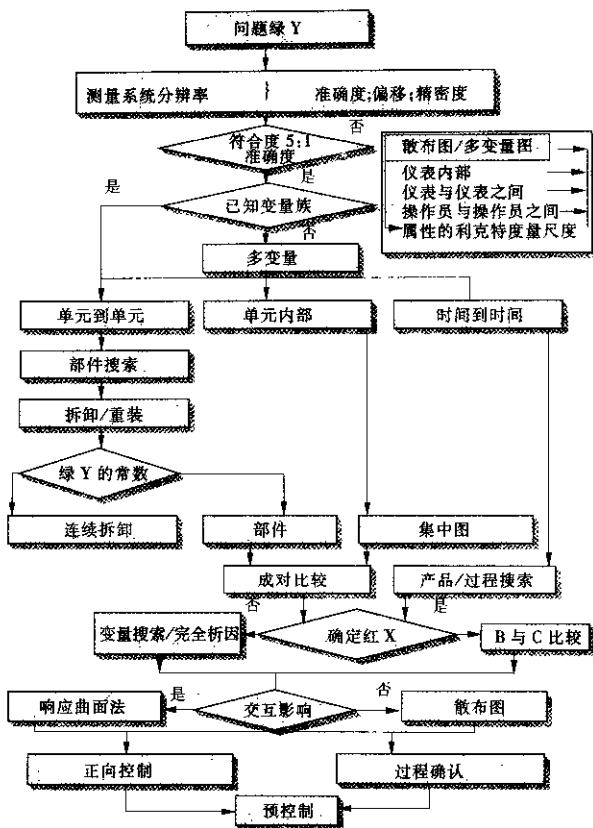


图 7—3 搜索红 X: 解决问题路径图



题。在 DOE 项目评审中，管理层也可采用该检查单。

(1) 问题是否已清楚地说明——最多用一句话（或一段话）？如果问题不能以简短的话语进行说明，那么工作组就很可能在此起点步履蹒跚。

(2) 绿 Y 是否已根据下列内容定义并量化？

- 缺陷水平（百分率、百万分率或 C_p 、 C_{pk} ）或现场故障水平？

- 成本、安全性或环境影响？

- 寿命（周、月、年）？

(3) 如果绿 Y 是一属性（通过/通不过），它是否能转换成一种以某种量化尺度表示的人为的变量，比方说从 1 到 10——1 为最劣，10 为最佳？这就是利克特量化尺度。

(4) 绿 Y 是否多于一个（在极少数情况下可多达 7 个）？

(5) 是否有一个与最终绿 Y 密切相关的较早的或较容易的绿 Y？例如，零售商店抱怨来自某加工厂的食品过早变质，但变质试验需时 14 天——对于计划的试验来说，时间太长了。比变质更早的绿 Y 是包装和密封食品的袋子的渗漏，于是渗漏成为绿 Y。但渗漏难以测量，所以封闭在袋中密封区的小块食品就成为较容易和较早的绿 Y。DOE 试验在三天内而不是持续几周的试验中就给出了红 X 原因。

(6) 是否已尽可能在流程中尽早查明绿 Y，以便从源头就抓住问题，不让问题在过程中继续堆积？流程是否已制成图表？

(7) 如果问题出现在现场而不是工厂，绿 Y：

- 是否设计了多环境强化应力试验以便在短时间内加速绿 Y 的出现？从长远观点来看，它是否能复制？

- 是否已找到一个或多个外部用户并与之接触？

- 是否已考察了用户的应用？

- 绿 Y 位于用户链中的何处——在运输、安装、批发商、销售商/零售商、服务商或用户之中？

- 用户的规范和容许偏差的有效性和现实水平如何？



- 是否已根据用户的需求得出内部规范和容许偏差？

(8) 如果问题与可靠性有关：(时间加应力，以区分在零时刻和零应力下的问题与质量有关)

- 是否绘制了浴盆曲线以确定故障是“早期失效的”(一个月之内在现场)，还是恒定故障率(1~3年之内在现场)，或者是一个耗损问题(远远超出保修期)？

- 问题是否是间歇性的？如果是，是否设计了强化间歇性失效变成永久失效的应力试验？

- 在发运之前或在工厂应力试验之中，在工厂中是否发生过类似的可靠性失效？

- 是否能用应力试验使问题加速，使绿 Y 变成导致失效的应力水平或导致失效的应力时间？

表面看来，这些问题的答案有许多是显而易见的，不需列入检查单。但是，许多 DOE 项目正是由于未精确地定义、量化或从头到尾地思考过绿 Y 而失败。

测量准确度

在进入第 3 篇——强大的线索生成 DOE 工具之前，应当强调一下与产品有关的测量系统的问题，因为经常会发生这种问题。与绿 Y 相联系的红 X，即首要考虑的问题是测量系统的弱点和变量，而不是产品内部的变量。

在定义绿 Y 之后，第一个应问的问题是：测量仪器好到什么程度？它相对于产品变量的变量是多少？可建立一些准则作为指南。例如：某指南规定了产品变量：仪表变量最小应为 10:1，有的指南则规定 4:1 到 6:1 为适宜范围。

总的变量(或总容许偏差 T_T)由两部分组成：产品容许偏差 T_P 和仪表容许偏差 T_I 。两个容许偏差并不是算术相加。它们遵循均方根(RMS)定律，由下式决定：

$$T_T = \sqrt{T_P^2 + T_I^2}$$



例如，如果产品容许偏差 (T_p) 为 5 个单位，仪表容许偏差 (T_I) 为 1 个单位——比例为 5:1，总容差为：

$$T_T = \sqrt{5^2 + 1^2} = \sqrt{26} = 5.1$$

如果现在把仪表容许偏差减至零，总容许偏差将仍然高达 5.0——5 中仅改变了 0.1 或仅是 2% 的变化。换言之，产品容许偏差：仪表容许偏差为 5:1 时，意味着总容许偏差的 98% 来自产品，仅有 2% 来自仪表，这样的不准确度可以忽略不计。

然而，如果上例中的仪表容许偏差为 3（即产品容许偏差：仪表容许偏差为 1.66:1），总容许偏差将是 $\sqrt{5^2 + 3^2} = \sqrt{34}$ 或 5.83，5 中有 0.83 的改变或不准确度为 16.6%，这是不能忽略不计的不准确度。

总之，10:1 的产品与仪表的容许偏差比，仪表准确度难以达到，尤其当工业产品容许偏差愈密集时，更是如此；而 6:1 的比率又将导致不可接受的准确度，所以，5:1 是一个合理的工业标准。容许偏差按均方根定律相加的另一特点是：必须不相称地结合在一起的两个、三个或四个容许偏差中的主要容许偏差（本例中为产品容许偏差）影响总容许偏差。

计量学中的精密度、偏移、准确度、分辨率

精密度的定义是围绕着频率平均值的频率分布的概率分布。在分布中，可以称它为范围 R ；如果是对一个规格宽度测得的，则称为 C_p 。**偏移的定义**是，在频率分布中平均值偏离目标值或设计中心的偏移量。在 C_{pk} 计算中，它是偏离中心的 ($\bar{X} = D$)，此处为 $\bar{X}$ 平均值， D 为设计中心。**准确度**结合了精密度和偏移。如与一规格宽度相联系，它可以称为 C_{pk} 。**分辨率**是产品概率分布与测量概率分布之比，最小比值为 5:1。图 7—4 为这些名词的图示。

A 栏所示的四个频率分布，它们的精密度（范围）恒定、 C_p 恒定，但偏移从上部的高偏移变至下部为零； C_{pk} 也相应改善。

B 栏所示的四个频率分布，其所有偏移均为零，精密度、 C_p 和 C_{pk} （二者始终相等）在上部为差，而到下部改善为极佳。

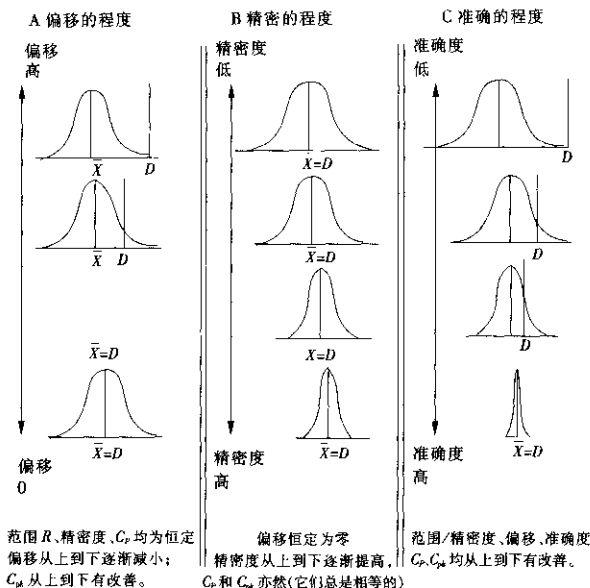


图 7—4 计量学：偏移、精密度、准确度

注：精密度=重复性；概率分布；范围 C_p （对于规格而言）

偏移=偏离设计中心 D 的量； $X-D$ 此处

D =设计中心或目标

$\bar{X}$ =平均值； R =范围

准确度：结合精密度和偏移； C_{pk} （对于规格而言）

分辨率：产品概率分布与测量概率分布之比

C 栏所示为四个不同准确度的频率分布——上部的为差而下部的为优——因为精密度、范围、偏移、 C_p 、 C_{pk} 均不断改善。



减少仪表变量

前面已经指出，如果总分辨率比值小于 5:1，减少测量仪表中的固有变量应先于减少产品中的变量。总仪表容差 T_i 由三部分组成——仪表内部的容差 (T_{in})、仪表之间的容差 (T_{i-i}) 以及操作员之间的容差 (T_{o-o})。

总容差由均方根定律决定：

$$T_i = \sqrt{T_{in}^2 + T_{i-i}^2 + T_{o-o}^2}$$

因为这三个容差不太可能相等（帕累托定律是普遍适用的），三者中的主要容差将对 T_i 有重大影响，因而是需要解决并减少的容差。

减少仪表内部的变量

造成同一仪表内部的变量的原因很多，当然这里假定操作员技术变量除外。这些原因包括：

- 在逐次测量过程中，环境温度、湿度等的差别。
- 电磁场、瞬态电压、电压、外部噪声、位置、材料、振动的影响。
- 滞后（例如齿轮的游隙）随着输入的减小而增加。
- 非线性。
- 装夹：连接、销钉、机械连结、负载、定位等。
- 多变量研究（集中图表用于找出单元内的变量）或 B 与 C 比较可以用作 DOE 工具，以找出并纠正特定的原因。

减少仪表之间的变量

两个或两个以上相同型号的仪表之间产生变量的原因有：不同的制造商；缺乏校准；材料、部件、线路、机械装配的差异；连接方法等。

可用的 DOE 方法是最佳和最差单元之间的部件搜索（如果仪表可以拆卸并重装，并且它们初始的绿 Y 仍相同）以及成对比较。



减少操作员之间的变量

两个或两个以上操作员之间产生变量的原因包括：缺少规程；程序和步骤顺序不统一；技术不熟练；判读误差；身体的差别（视力、高度、熟练等）。

可用的 DOE 方法是流程图，记录下每一个步骤并用成对比较法寻找差别。

测量准确度检查单

在 DOE 开始之前，解决问题小组应当查阅下列关于测量设备准确度的检查单。

- 绿 Y 是否可测量？
- 某个属性（即开动/停止，通过/失败）的测量是否可用李克特度量尺度转换成一变量（见下面的案例研究）？
- 这样一个李克特度量尺度在不同操作员之间可否统一评定？
- 是否进行了散布图研究以确定 5 : 1 的最小分辨率比？
- 仪表内部的、仪表之间的以及操作员之间的变量（容差）是否已相对于产品容差进行了量化？
- 是否已对三个变量根源中最主要的原因进行了试验，是否已确定了一劳永逸的解决办法？
- 如果技术水平不可能达到 5 : 1 的比值，是否有较早的、较易的而且可以更加一致地测量的替代绿 Y？

略过这一检查表，且不评定测量系统的充分程度，就会使一个解决问题的项目从一开始就出错。

将属性转换成变量

在量化和测量绿 Y 时的另一个重要步骤是，需要客观地将一属性转换成一变量。有三种类型的数据——变量、属性和等级。

● 变量数据可包括非常多的，几乎是无穷多的某参数的读数，这取决于所采用的度量尺度。例如，变量数据有尺寸、电压、重量、抗拉强度等。狭义地说，甚至缺陷数、成品率以及百万分率都



能近似于变量。

●**属性数据**是数字数据 要么好要么差、要么通过要么失败、要么接收要么拒收——在两个极端之间没有读数。属性数据包括各种装饰缺陷，如油漆、镀覆、气泡、腐蚀、颜色、焊接缺陷、砂眼以及工具伤痕等。

●**等级数据**变量不是用确切的值来记载的，而是用等级来分类——从最小到最大，从最低到最高，从最好到最差（或反之）。等级数据的使用我们将在第 11 章关于图基检验一节中进行讨论。

李克特度量尺度

因为属性数据在好与差之间没有等级，所以我们面临的挑战就是人为地把一个属性转换成一个变量，并给予它一个等级。这一任务由李克特度量尺度来完成，其中，1 是最差，10 是最佳。由用户、销售商、工程界、制造业、质量部门组成的委员会把 10 个实物样本分定等级，从不可接受的 1 到非常完美的 10 分成 10 个等级。这些实物样本（有时照片就可以）并不构成一个对生产中的相似缺陷确定恰当的分数的“变量度量尺度”。这需由不同检查员的独立主观判断产生。它防止了生产部门对譬如装饰缺陷要求接收而检验部门要求拒收之类的无休止的争吵。有时一个 0~5 或 1~6 的度量尺度就够了。李克特度量尺度也可倒过来，以 0 为最佳，5（或 10）为最差。图 7—5 是李克特度量尺度用于凸轮端轴承的工具伤痕的实例，0 为最佳，5 为最差。

李克特度量尺度达到了较少的样本量

李克特度量尺度的另一优点是减少了 DOE 工作中的大样本量。如果不用李克特度量尺度，就需要计算缺陷率、成品率等。例如，如果我们要区别两个成品率——譬方说 70% 和 80%——在 DOE 试验中就需要大量部件（通常每批至少抽取 30~50 个样本）以确定这一属性的差别。相反，采用李克特度量尺度，每个成品仅需 3~



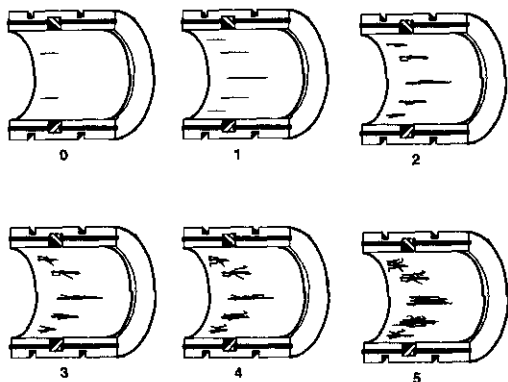


图 7—5 利克特度量尺度：凸轮端轴承分级

10 个样本量就能发现差别，甚至量化它（见第 11 章和第 15 章中的图基检验和 B vs. C）。

测量案例研究：电动剃刀

某环球公司的一家工厂的电动剃刀产品的噪声过大。但在确定红 X 原因之前，必须建立一个准确的测量系统。典型的噪声测量工具是频谱分析仪和分贝表。然而，这些仪表的读数前后不相一致，而且它们的分辨率也很低。

有人提议进行听觉试验。经过数次反复试验后，一个从 1 到 10 的利克特度量尺度——1 最差、10 最佳建立了起来。10 名操作员接受了听觉试验培训。结果，有噪声的剃刀的平均值是 3，变化范围仅为 0.2；无噪声的剃刀的平均值为 8.6，变化范围为 0.3。分辨率比值在两种情况下均远优于 5 : 1。管理人员非常惊奇于由经过培训的操作员所进行的主观的听觉试验的结果竟然比更加高级的仪表更一致、更准确。

第3篇

“与部件对话”——一组线索生成工具

走向红 X 的连续过程

在第 7 章我们看到，伴随着问题——绿 Y 的定义和量化，走向红 X 之旅开始了。接下来要强调的是，为了成功地解决某个既定的绿 Y 问题，必须有一种可行的测量系统，其准确度与产品的容差或技术规格要求说明书规定的宽度（不管其是否很大）相比至少为 5:1。

第 3 篇由一组 5 种线索生成工具所构成——多变量分析、集中图、部件搜索、成对比较和产品/过程搜索。这些工具相互结合，形成了“与部件对话”的能力，且远比猜测、直觉、偏见、见解、经典试验和田口试验的判断要优越。它们所提供的线索是可靠的，并且不会干扰生产，因为它们不中断生产，抽取样品仅仅是为了检验。

在接下来的 5 章里，我们将深入地研究这 5 种工具，每章包括：

- (1) 详述给定的工具和程序的每个步骤。
- (2) 成功案例的研究。
- (3) 不成功案例的研究及从中吸取的教训。
- (4) 工厂实践，回答读者实践问题。
- (5) 在白领管理问题中的应用。
- (6) DOE 小组的指导问题。
- (7) DOE 计划评审时的管理问题。





第 8 章

多变量分析：自动寻找红 X

目的与原理

多变量研究的目的是为了把大量未知的（或仅仅是怀疑的）和不可处理的变量的原因（原因、因子和变量是同义词）减少为少得多的一族相关变量，其中包含红 X，即主要原因。这是一种通过去掉一些无关紧要的变量原因来定位最可能的原因的图表技术。请在图 7-1 解决问题路径中，找出其位置。

用扑克戏法说明用多变量分析法寻找红 X

作者曾在 DOE 研讨会上，通过扑克戏



(11) 你会很神气地从第 2 列的第 5 行中抽出那张牌（总是中间的那行），并声明梅花 A 就是红 X。

扑克戏法背后的多变量原理

在给定的问题中，我们说 27 张牌就是 27 个变量，其中任何一个都有可能是红 X。如果猜红桃 Q 是红 X，猜对的可能性 4 仅仅是 $1/27$ ，这种可能性比工程师所遇到的问题情况要好，因为他们猜对的可能性只会是 $1/50$ 或 $1/100$ ！

●在布牌 1 情况下（图 8-1A 栏），当那位选出的观众指示第 2 列含有所选的牌时，你马上就得到一个信号，即这张牌不在第 1 列和第 3 列，这就取消了 18 个变量，把红 X 的范围缩小到仅有 9 张牌的第 2 列。

●在布牌 2 情况下（图 8-1B 栏），原先第 2 列中的 9 张牌，处在 3 个中间行上，即图 8-1B 栏中的第 4、5 和 6 行。当观众又指示第 1 列时，你又会得到一个信号：红 X 仅仅可能处于第 1 列的第 4、5、6 行中，这又取消了与第 2、3 列相关的第 4、5、6 行中的 6 张牌。这样一来，红 X 就缩小到仅是 27 张牌中的 3 张牌。

●在布牌 3 中（图 8-1C 栏），刚才第 1 列的第 4、5、6 行中的这 3 张牌，现在处于中间或者说第 5 行上。在你发问后，观众指示第 2 列时，红 X 就肯定是梅花 A。

这就是多变量研究原理。它是一种过滤性技术，通过它，大量的无关变量被排除，直到红 X 被系统地过滤出来。大多数应用情况下，多变量作为第一过滤器，通常按下列规则进行：

●如果仍有太多的残留相关变量，那么应寻找其他思路创新技术。

●如果残留的相关变量在 5~20 之间，那么应进行变量搜索；如果在 2~4 之间，那么应进行全析因分析。

●如果相关变量仅有 1 个，并且已确定要进行永久改善，那么



应进行 B 与 C 比较。

●既然令人头晕目眩的猜想已大大减少，那么就要进行工程判断了。

Quality Department

2006.05.20

SongYongSheng

因果（石川）图的有效性

许多试验是由因果图开始的，即罗列出每个能够想到的、由头脑风暴会成员所发掘出的给定效果的原因。但是数目庞大的因子，会令人深感困惑。图 8—2 就是一个令人踌躇的因果图，该图试图解决某微电子电路生产过程中的烧结电阻的变量问题。竟有 80 多个原因被精心罗列出来！你从哪里开始？你将选取哪个原因着手试验？专家组典型的一种处理方法是头脑风暴会成员举手表决选定一个因子。民主可以是一种可靠的政治体制，在解决加工过程中的长期的质量问题方面也是有用的。基于图 8—2 中的 80 多串猜测，田口试验在该公司就显得无能为力，没能找到问题所在。

相比之下，使用一种渐进式多变量研究，再结合其他线索生成工具将利索地找出红 X、粉红 X、浅粉红 X，并把烧结电阻的变量由 55%~90% 减少到 2% 以下。

三个变量族：位置的、周期的、暂时的

扑克戏法中 27 张牌被分成 3 族：第 1 列、第 2 列和第 3 列。但是，在工业管理中，也能将扑克戏法作为解决问题的方法。分解问题时，最常用的方法是将其分为 3 族：位置上的或称部件内的变量；周期的或称部件对部件的变量；暂时的或称时间对时间的变量。



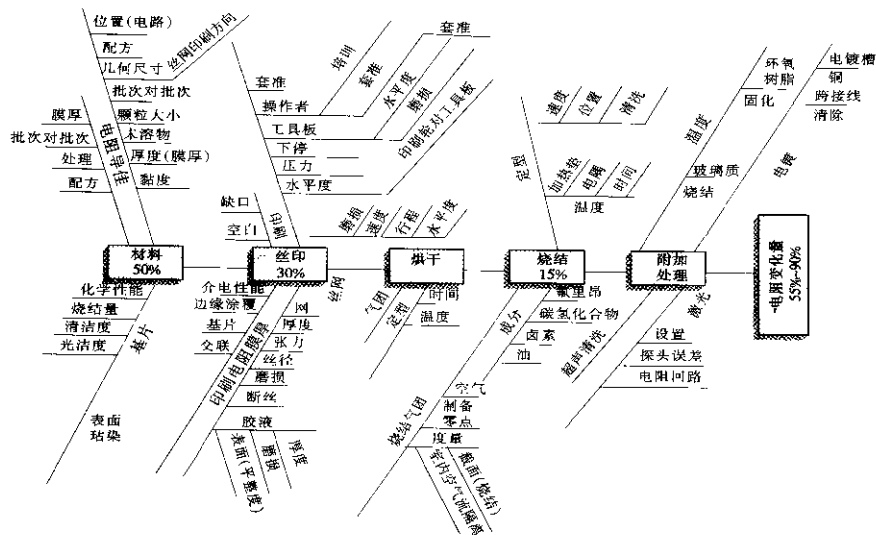


图 8—2 烧结电阻值变化量的因果图

资料来源:An early Motorola microcircuit experiment.

如果最大的变量是暂时性的，则可排除与位置和周期性变量有关的原因，或者在深入调研时不对其进行优先考虑。每一族变量的实例如下。

位置变量（部件内的族系）

- 在一个部件内的变量（例如左面对右面；顶面对底面；中心对边缘；锥度；不圆度；偏转；铸件壁厚）；

- 包含许多部件的单一部件中的变量（例如包含许多芯片的半导体硅晶片；某种元件类型或数值的缺陷反复出现的印刷电路板电路）；

- 在成批加料时出现的位置或方位的变量（如模压时，模腔对模腔的变量）；

- 机器对机器的变量；

- 试验位置对试验位置的变量，夹具对夹具的变量；

- 操作者对操作者的变量；

- 生产线对生产线和工厂对工厂的变量。

周期性变量（部件对部件族系）

- 在同样的时间框架内，从一个生产过程中抽取的连续的部件间的变量；

- 部件组中的变量；

- 批次对批次的变量；

- 批量对批量的变量。

暂时性变量（时间对时间族系）

- 小时对小时；

- 班次对班次；



- 每日对每日；
- 每周对每周。

DOE 研究概述

作者见过太多的公司绘制的帕累托图表实例，这些图表描绘了各种缺陷：大量的模型和大量的缺陷类型。可是公司却被所要解决的问题数量和范围累垮。在这种情况下，重要的是应遵从下列 4 条简单的规则：

- (1) 仅选择具有最大数量缺陷的模型作为起点；
- (2) 在此模型中，应选取具有最大数量缺陷的模式；
- (3) 用一个或几个 DOE 工具来解决此问题；
- (4) 对于具有第二多的缺陷的模型，重复上述的第 1、2、3 步，以此类推。

设计、实施和分析一个多变量的研究：一种路径图

在设计、运行、解释和分析一个多变量的研究过程时，必须按一定步骤进行，现归纳如下。

1. 设计多变量研究

- (1) 鉴别绿 Y（要解决的问题）。如果绿 Y 是一项属性，则试着利用一种李克特度量尺度将它们转化为一种变量。
- (2) 要保证测量仪器的精度至少为产品精度（即它的允差）的 5 倍。
- (3) 确定可出现的变量的族系数目。



- (4) 画出族谱。
- (5) 估计所要求的时间对时间采样的数目。
- (6) 确定在加工过程中连续抽取的部件对部件的数目（一般为3~5）。
- (7) 确定在部件内族系的各子族系的采样数目，如方位、机器和模腔的数目。
- (8) 将第3、4、5步中的数量相乘，以确定需要研究的部件的总数量。
- (9) 设计一个图表，以简化多变量数据的收集。

2. 进行多变量试验

- (1) 不要混淆一个给定的产品内的模型，仅对此产品中最差的模型进行试验。
- (2) 进行多变量研究，扩展时间对时间采样的数量，直到找到80%以上的重要的变量或技术规格要求公差（即使很小）为止。
- (3) 在多变量研究过程中，应使过程中的各种调整的次数最小化。
- (4) 要特别注意任何不连续性，例如：休息喝咖啡、吃午饭、换班、更换操作人员、改变设置、换工具、预防性的维修等，这在进行多变量研究过程中是不可避免的。在进行时间对时间的采样时，要尽可能挑选在这些不连续因子之前或之后的时间进行。

3. 解释和分析多变量图

- (1) 确定重要变量的族系（注意：红 X 仅可能存在于几个族系之一中，虽然粉红 X 或浅粉红 X 可能存在于另一个族系中）。
- (2) 如果红 X 的族系是时间对时间的，要检验温度、湿度、工具磨损、休息和午饭期间的改变以及在多变量运行过程中的调整 and 任何加工参数的改变。随后的产品/加工过程搜索的后续工作就会



顺利（见第12章）。

（3）如果红X族是处于部件对部件族系，就要检验周期图形、灰尘及管理，这些因子可以影响一个部件，但不会连续影响其他部件。随后的部件搜索和（或）成对比较就会顺利。

（4）如果红X是处于部件内的族系，就要建立一个集中图（见第9章）以确定绿Y的重现方位或部件。

（5）寻找非随机的趋势或其他线索。

（6）寻找一个或几个存在不寻常图形的样本。不相等的灵敏度说明可能存在交互的影响。

（7）在族谱中，列出每一个变量族系所有可能的原因，以开始随后一系列的调研。

案例研究1：设计多变量——印刷电路板钻孔

印刷电路板在钻孔时产生的过量毛刺就是绿Y。有一种利克特度量尺度（见第7章）用于测量印刷电路板的毛刺，毛刺的级别被分为1（无毛刺）到10（最大毛刺）。然后就是对各种变量族系进行甄别。

● 决定仅用1天时间进行试验，期望在1天之内就能充分找到至少80%的重要变量；

● 每天3班倒；

● 有13台同样的钻床；

● 每台机器有4个钻削夹头；

● 每班8个工人操作机器；

● 每台机器有3个装在印刷电路板上的控制面板；

● 每台机器有10种钻孔尺寸。

这3个主要变量族系和7个子族系，以族谱的形式显示在图8—3中。随着多变量调查的进展，对毛刺变量无关紧要的子族系可予以排除。



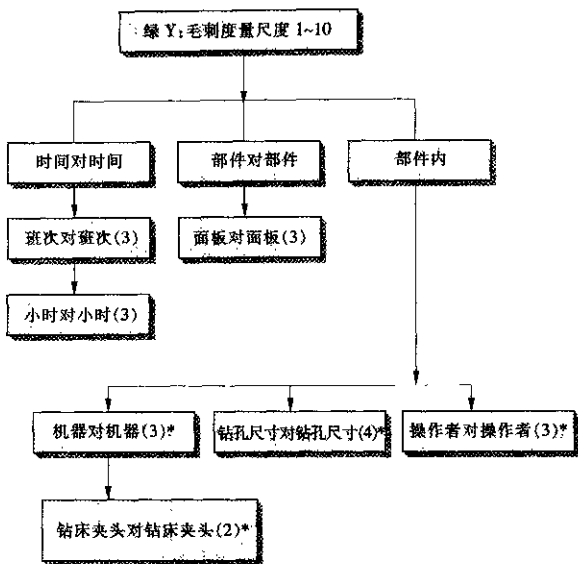


图 8—3 某项多变量计划的族谱

* 样本量。

注：在给定的子族系里，不必取得全部部件数（除非没有得到 80% 的重要的变量）。这里，13 台机器中仅有 3 台被取样；4 个钻床夹头中仅有 2 个被采样；10 种钻孔尺寸中仅有 4 个被采样；8 个操作者中仅有 3 个被采样。

资料来源：Lika Corporation, Stockton, Calif.

确定采样频率和所需部件的数量

一项多变量研究应持续多长时间？正如前面说过的那样，有经验法则表明：应该持续在一个过程中进行周期采样，直到采集到

80%的重要的变量或技术规格要求公差（不管其是否很小）。如果采用较低的样本量，例如 50%或 60%，则快速找到红 X 的机会就会减少。总体而言，多变量研究可以持续不到 1 小时至 4 周或更长，以便采集到 80%以上的重要的变量。但是，比较典型的是多变量研究不会持续超过 1~3 天。需要多少样本呢？在部件对部件族系里，样本量可为 3~5 个连续部件。

在部件内族系里，每个子族系的采样既可以是 100%，也可以从每个子族系中采集合理数量的样本；目标是将所要求的部件总数量保持在实用的上限，以便有足够的机会采集到红 X。

例如，在印刷电路板的例子里，如果我们以 3 为部件对部件变量采样的样本量，从 10 台机器、4 个钻床夹头、8 个操作者、10 个钻孔尺寸中采集这样的样本量，每小时所要求的部件总数为： $3 \times 10 \times 4 \times 8 \times 10$ ，即 9 600 件/小时；再乘以 3 小时和 3 班，则总数为 86 400 件。所以，可以作一个决策：仅对 10 台机器中的 3 台取样，即以往最差的、最好的和中间的；同样，也可只选择 3 位操作者、3 个钻孔尺寸和 2 个钻床夹头。这将使每小时所需部件总数减少到 $3 \times 3 \times 3 \times 3 \times 2$ ，即 162 个；再乘以 3 班、每班 3 个样本，所要求的部件的总数将仅为 1 458 件。图 8-3 所示为每个子部件中的这些较小的数量。

换一种方案，即可以只由 1 班开始。如果采集到了 80%以上的重要的变量，则可删除第 2 班和第 3 班，这样就将部件的总数减少到 162 件。

我们特意选择了一个具有多变量族系、每个族系里有许多部件的案例研究。对于一个多变量而言，绝对的最低值是 3 个部件和 3 个时间周期，或者说整个试验要进行 9 次，很少有总部件数量超过 100 件的。表 8-1A 和表 8-1B 所示为根据变量族收集数据的数据表，表 8-1B 是表 8-1A 的子表。这样的表格可以很容易地交给操作者或检验员去作记录。



表 8—1A

多变量制备表：印刷电路板钻孔

变量族	1 天								
班 #	1 班			2 班			3 班		
小时 #	1 小时	2 小时	3 小时	1 班	2 班	3 班	1 小时	2 小时	3 小时
小时 #	1	2	3	1	2	3	1	2	3
部件 #									

表 8—1B

变量族	部件 1											
机器 #	1			2			3					
钻床夹头 #	1		2	1		2	1		2			
操作者 #	1	2	3	1	2	3	1	2	3	1	2	3
钻头 #	1	2	3	1	2	3	1	2	3	1	2	3

资料来源：Lika Corporation, Stockton, Ca.

小组练习 1：半导体硅晶片多变量计划

在一个多变量研究中，要在一个半导体硅晶片上（包括几百个芯片）进行厚度测量，以确定哪些变量族对厚度的差别影响最大。厚度的变量就是绿 Y。假定测量仪器的精度是硅晶片规范厚度公差 6 倍多。设计一个多变量计划，以表述下列计划方针所罗列的问题。

- (1) 为以下 8 个计划方针中的每一个鉴别重要变量族（即时间



对时间、部件对部件、部件内)。

●在每个硅晶片上,要对5个部分进行测量(北、南、东、西和中间)。

●从镀膜后的每个批次里采集3个样本。

●在该批次镀膜过程中,从真空室的左、中、右3个方位采样。

●使用2个真空镀膜室。

●每批次需2小时,每班4批次。

●从1班和2班中2个连续的批次里采样。

●采样是在周一、周二和周三完成的。

●采样是在连续的3周内进行的。

(答案将在下文给出。)

(2) 画出族谱。

(3) 在整个多变量运行中,应包括多少硅晶片?

(4) 如果过去的经验表明,周对周、天对天、班对班中很少产生有影响的变量时,所要求的总硅晶片采样数目会减少多少?

答案

(1) 问题1

●在每个硅晶片上,东、西、南、北、中5部分	部件内
●每个批次3个硅晶片	部件对部件
●每个批次里的3个方位	部件内
●2个镀膜室	部件内
●每班4批次	部件对部件
●取自1班和2班的批次	时间对时间
●于周一、周二、周三取样	时间对时间
●在连续3个星期内取样	时间对时间

(2) 问题2

见图8—4。



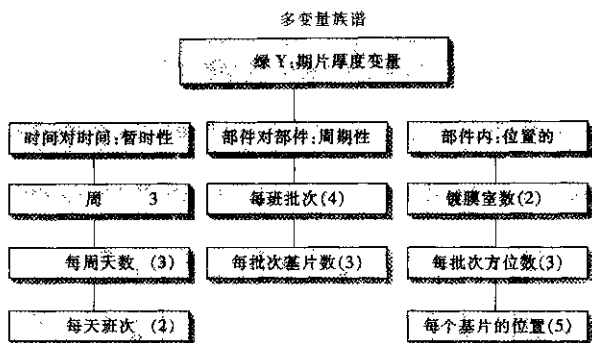


图 8—4 多变量族谱

(3) 问题 3

周 天数/每周 班数/每天 批次数 硅晶片数/批 镀膜室方位数/批

$$3 \times 3 \times 2 \times 4 \times 3 \times 2 \times 3 = 1\,296; \text{读数} = 1\,296 \times 5 \text{ (部分数/硅晶片)} = 6\,480.$$

(4) 问题 4

如果时间变量 (周、天、班) 是最小值, 硅晶片总数量将是:
 $1\,296 / (3 \times 3 \times 2) = 72$, 而读数将是: 72×5 (即 5 个部分) = 360。

► 编制多变量图表

表格数据总是可以用于解释和分析一个多变量研究, 但是人的眼睛是不能很容易地消化这些数据的。因此要采用一种图形式的方法来简化解释, 并量化每种变量族的影响。

水平线表示时间——可以是周对周、天对天、班对班或小时对小时, 但是由于从每个时间段中仅有样本是周期性地采集样——每次



3~5 个样本——所以可把水平轴划分为具体的时间周期。

垂直轴表示正在研究的绿 Y，且部件内的变量（比如说从顶部到底部，或从侧面至侧面）以垂直的小短划线表示。多变量不是人为试验，它并不中断生产——这是一个优点。在一定意义上，它是在给定时间内或一个短的周期内抢拍到的变量快照。这有些像心脏病患者的心脏 EKG 图，各种传感器监视着患者胸部的不同部位，以确定心脏功能的大致情况。图 8—5 表示了可能应用于度量某项质

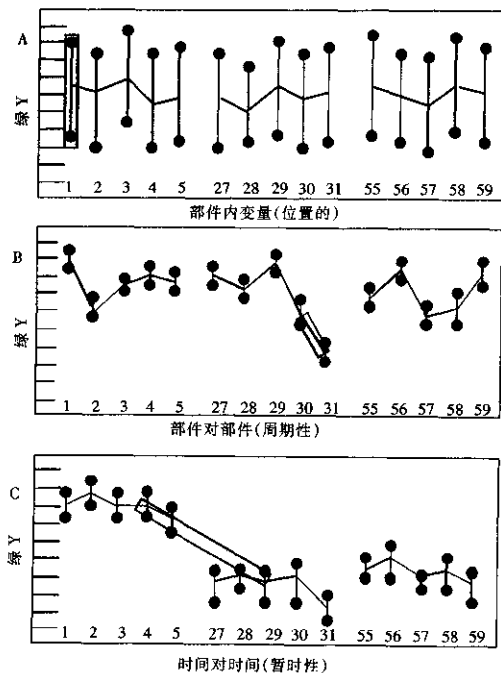


图 8—5 多变量图表：根据重要族系分离



量特性的 3 种变量类型。需在给定的时间内对少数几个部件，通常是 3~5 个，连续进行观测（图 8—5 表示了 5 个连续的部件），在其后的 1 或 2 个小时中再对另 5 个连续部件进行监测。采样是周期性重复进行的，直到至少采集到过程中的 80% 的重要变量或整个规范公差为止，不管其是否很小（参见“进行多变量试验”一节）。

通过绘制多变量运行所得到的结果，你可以确定最大变量是位置性的（部件内），如图 8—5A 所示，还是周期性的（部件对部件），如图 8—5B 所示，或是暂时性的（时间对时间），如图 8—5C 所示。多变量是 DOE 技术中的一种，它不应当被随机化，它是连续性的。许多人第一次接触多变量图时，会把它与控制图相混淆，但控制图仅可区分出时间对时间变量，它不能探测出部件对部件或部件内的变量及其子族系变量。

案例研究 2：转子轴

某加工者在制造直径要求为 0.0250 英寸 $\pm$ 0.001 英寸的圆柱体转子轴时，出了许多废品。对加工过程能力的研究表明：整体误差分布为 0.0025 英寸，而要求的误差分布为 0.002 英寸，即 C_{pk} 仅为 0.8。工段长准备废弃老式车床，去买一些新机床（70 000 美元），这些新机床可维持 $\pm$ 0.008 英寸的误差，即 C_{pk} 值为 1.25。但工厂的管理者在经过咨询之后，要求在购买新机床之前进行多变量研究。

图 8—6 以图表的形式给出了多变量研究的结果。有 3 个轴是在上午 8:00 加工的，另 3 个轴是在上午 9:00、10:00、11:00 和中午 12:00 加工的；没有必要超过中午，因为规则规定至少 80% 的重要的变量——0.002 5 英寸要在 11:00 前采集到。对每一个轴要进行 4 次测量——两次在左面，两次在右面。对每一个轴应从左至右测量锥度；从上至下测量不同心度，让轴旋转，以测量最大直径点



和最小直径点。这样一来，在部件内变量中，就会出现两个子族系，换言之就是锥度和不同心度。

连接每个轴的 4 个读数平均值的细线表示的是周期性变量，即部件对部件变量；双点划线则表示的是暂时性的即时间对时间的变量。

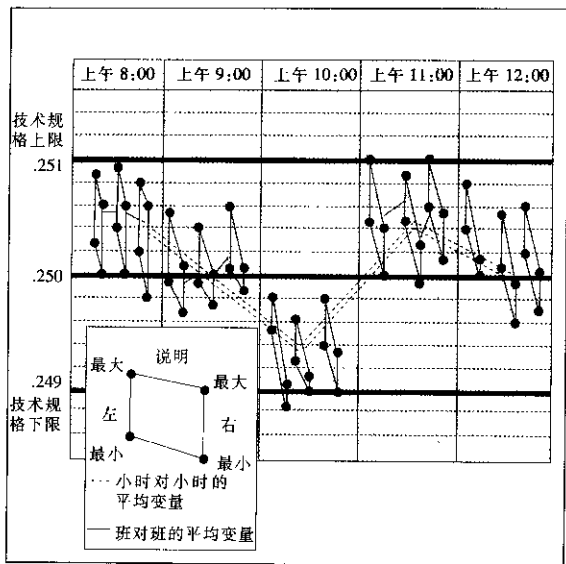


图 8—6 转子轴：多变量图表

即使是对多变量技术不了解的人，对图表一瞥，也会发现最大的变量似乎是时间对时间的变量。从 8:00 到 9:00，直径减小，到 10:00 进一步减少；但在 10:00 和 11:00 之间，情况向反方向发展。是何种原因造成了时间对时间的变量？工段长怀疑是工具磨损，但

如果工具磨损是原因的话，读数在 8:00 到 10:00 这段时间内应上升，可它们反而下降，所以就排除了工具磨损这一原因。从 10:00 到 11:00 发生的反向现象，为工段长提供了一条有效的线索。在一般情况下 10:00 左右发生了什么问题？是喝咖啡休息！当下一批 3 个轴在 11:00 被采样时，读数情况如同 8:00 开始生产时一样。这样一来，工段长就可将时间变量与温度变量等同起来。随着车床自 8:00 起开始升温，到 9:00，再到 10:00，读数变得低了；但在工人们出去喝咖啡，关掉车床后，较低的温度条件导致了 11:00 时样本较高的读数。然后工段长猜测温升可能是由于箱中冷却液的量不足造成的。他确认冷却液液位降低了。当他将冷却液添加到指定的液位之后，这种为数占允许的误差数量达 50% 的时间对时间的变量，几乎降至 0。这表明了我们是怎样由时间线索引导到温度线索，再由温度线索引导到冷却液液位——即红 X 的！

部件对部件变量只占总允许误差的 5%，是不值得进行调研的。但是部件内的变量（即正向控制）则有两个显著的子变量族系，一个是不同心度，它占允许误差量的 30%；另一个是锥度，它占允许误差量的 10%。每个轴的不同心度的存在，是由于卡盘轴的轴承磨损造成的。换上新轴承，连工时都算上才花费了 200 美元，这样就又消除了上述变量的 30%。

每个轴的锥度显示出是一个非随机性的变量，所有 15 个轴的左面都高于右面（从每个多变量图中搜寻这样的非随机性的趋势，以便提供有力的线索是很重要的）。由此可得到结论：刀具在从左至右切削轴时，它不平行于轴的轴线。稍稍调整一下导轨，就可将锥度减少至几乎为 0，使上述变量又减少了 10%。

每个变量族系和子族系的量化

对每个变量的族系和子族系进行量化是很重要的。

● 时间对时间族变量是最高时间平均值和最低时间平均值的差（或范围）；



● 部件对部件族变量, 是最高部件内平均值和最低部件内平均值的差 (或范围);

● 部件内族变量, 是给定部件的最长垂直点划线的高度。

有时要对每种族系的平均 $\bar{X}$ 值和标准偏移量进行严密的分析, 但在大多数情况下不是必须的。

表 8—2 是转子轴案例研究中每个变量族和子变量族系的小结——变量所占的百分比、原因、所做的修正和变量降低的百分比。

表 8—2 转子轴案例研究: 变量的量化

变量族	变量子族	占总变量的比例 (%)	变量原因	变量修正	变量减少比例 (%)
时间对时间	小时对小时	50	冷却液液位低	冷却液加量	接近 50
部件内	不同心度	25	轴承磨损	新轴承	接近 25
部件内	锥度	15	设置不当	调整导轨	接近 15
部件对部件	轴对轴	5	?	—	—

图 8—7 为转子轴案例研究的族谱图。

概括而言, 转子轴多变量研究获得了下述结果。在随后的生产过程中, 总变量由 0.002 5 英寸下降到 0.000 3 英寸, 新的 C_{pk} 值是 0.002/0.000 3, 即等于 6.7, 超过了世界标准 C_{pk} 2.0。所得到的效益是: 废品为 0, 且节省了购买新设备的 70 000 美元投资。

从此案例中我们得到一个教训, 即企业常常没有对暗藏的变量原因进行调查, 就把旧机器或生产过程当作“垃圾”而丢弃。我们的经验表明, 90% 以上的类似案例是可以鉴别出并修正故障原因的, 从而无需承受沉重的资本投资负担。而日本人, 一边在旧机器中仔细地寻找变量的原因, 一边向西方人推销新式而昂贵的机器! 但他们采用的降低变量的方法是粗糙的, 例如因果图——它无法与多变量法或我们后而要讲的其他线索生成技术相比拟。



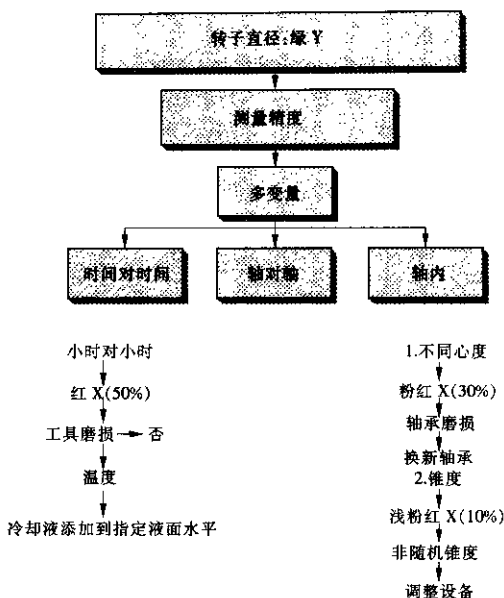


图 8—7 转子轴案例研究的族谱图

两种多变量图

在多变量图表里，各种变量族系可以综合进一个图表中，每个族系和/或子族系也可分别进行制图。

当族系和子族系的总数不超过 3 个或 4 个时，第 1 种方法是最好的。否则，图表会变得太复杂，在解剖单个变量族时会太忙乱。



即使是只有4个变量的图8—6，对于一个初次制作多变量图的人来说，似乎也太复杂。

当变量族和子变量族系的数量超过4个时，分开显示每一个变量族是一种可取的方法。图8—8以转子轴案例研究为例，说明了这一点，其中红X族（时间对时间变量）、粉红X族（部件内的不同心度和锥度变量）被分别作图。图8—8表明：最小的（可忽略的）变量是轴对轴变量，它小于0.0001英寸。

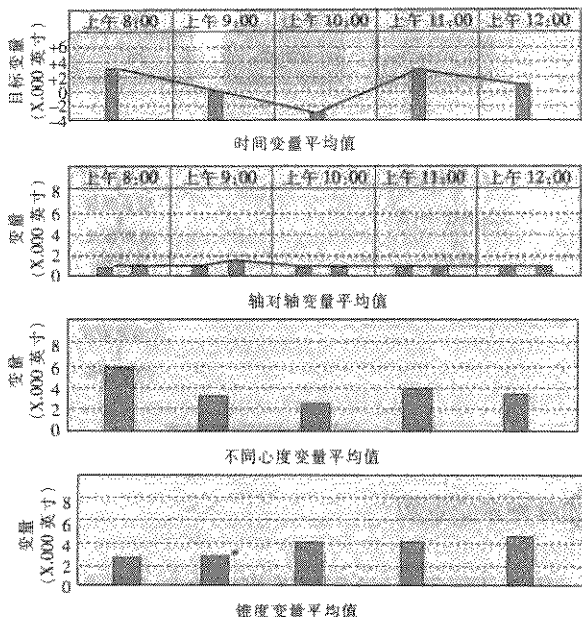


图8—8 另一种多变量图：转子轴



案例研究 3：收音机拒收——“找不到故障”的类型

几年以前，在某大型无线电工厂，收音机是由计算机检测站测试几个电器参数来进行测试的。经过这些测试后所产生的废品被送去分析，以确定报废的原因。在这些废品中有 22% 以上分析员找不到任何毛病，所以这些收音机又被送回计算机检测站重新测试。大多数被送回的收音机又被计算机检测站定为废品。于是又被送回来分析，接着又送到检测站检验，就这样来来往往，转来转去，直到它们不得不走出厂门。

为了不再这样兜圈子，工厂成立了一个多变量小组。小组每天抽查 100 台收音机，并按下述的变量族检验“找不到故障”收音机 (no-trouble-found, NTF)：天(3)、班(2)、分析者(2)、检测站(7)、无线频段(3)、电器参数(12)。

图 8—9 显示了除电器参数之外的其他上述变量族的多变量图。它清楚地表明，从第 1 天到第 2 天，再到第 3 天；每天从第 1 班到第 2 班；或从每班的 2 位分析人员中都看不到 NTF 在数量上有什么变化。但是在计算机检测站中，检测站 A 和 C 所观察到的 NTF 数量是检测站 B 和 E 的 10 倍，这肯定是红 X。此外，频段 A02 和 A03 的 NIF 是 A01 的 5 倍，这是一个粉红 X。

问题：在计算机检测站中有 22% 的废品收音机经分析被判定为 NIF

分类：(1) 天对天
(2) 班对班
(3) 频段对频段
(4) 站对站
(5) 测试参数对测试参数

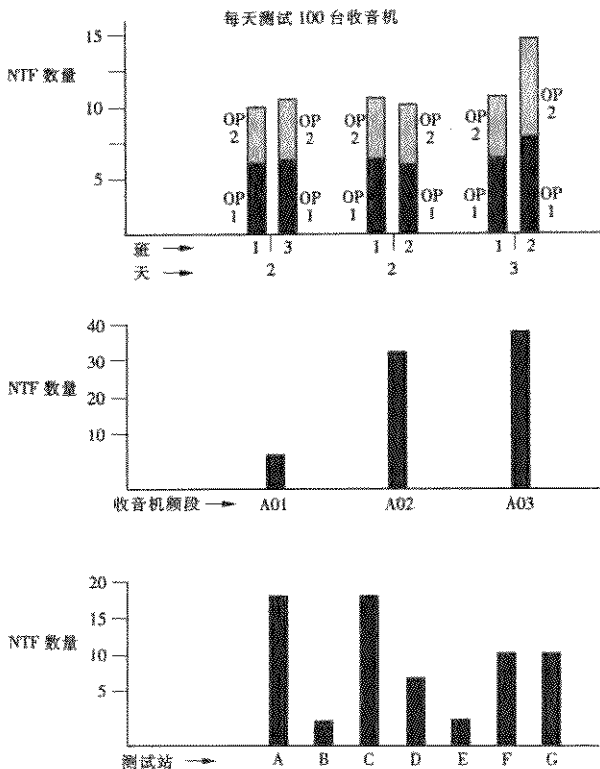


图 8—9 多变量图：NTF 案例研究

资料来源：Motorola, Bayton Beach, Florida.

经过对检测站的调查得知，检测站 A、C 的电源不标准；检测站 A、C、D 测试插销有缺陷；检测站 F、G 的同轴电缆有缺陷。当这些故障被纠正之后，正向控制、生产过程的确认、预先控制等（参见第 18、19、21 章）都得到了切实的控制。对频段 A02 和 A03 调研之后，确认 12 个电器参数中有 3 个（呼叫灵敏度、I—V B、天线调谐）要进行少量的更改设计。

结果，NTF 由 22% 下降至 2%，每年可节约 150 000 美元。调查又持续了两周，结果 NTF 又降低至 0.5%，每年可节约 50 万美元！

从这个案例研究中，我们可以得到下述教训：

（1）不要满足于微小的、细微的、递增的改善。（如果你满足于微小的改善，就利用经典或田口方法吧！）如果你不能得到 5:1、10:1、50:1 的改善，那么你仅仅是触及皮毛而已。

（2）经典的 DOE 和田口 DOE 是怎样解决问题的？去猜原因？真是浪费时间！

（3）实际消耗在多变量研究上的时间仅仅 3 天而已，外加 1 周的准备时间。生产过程完全没有搞乱。

小组练习 2：洗衣机桶的喷漆缺陷

在某洗衣机生产线上，喷漆缺陷是最大的缺陷类别。公司决定利用多变量图区分出变量族。重要的变量族在 7% 左右。公司认为，1 天的数据足以采集到至少 80% 的重要的变量。生产线的生产率为 250 台/每小时。公司采用 100% 的每小时累积采集的检验数据，把缺陷按桶的 3 个部位进行区分，即顶部、中间和底部——用时钟的方式指明圆筒的圆周上的位置，如：12:00、3:00、6:00 和 9:00 等位置。

图 8—10 所示为按族分类的缺陷图。从中你能得到什么结论？在多变量分析中犯了什么错误？



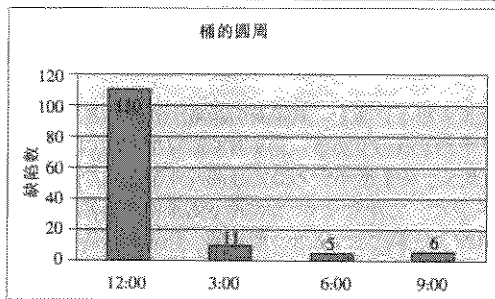
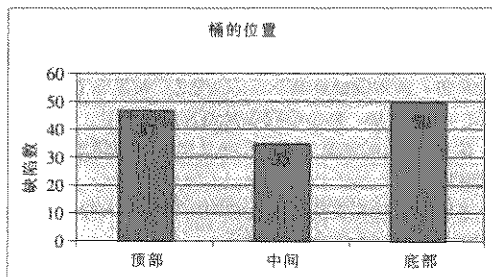
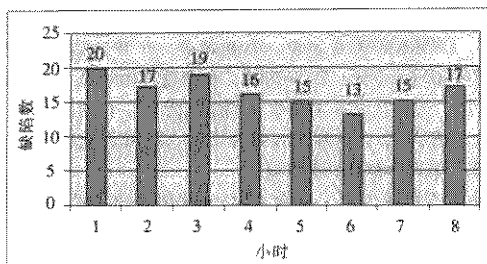


图 8—10 多变量图：洗衣机桶的喷漆缺陷

资料来源：Vitromatic Co., Celaya, Mexico.

洗衣机桶喷漆缺陷练习中的结论和所犯错误

(正确的方法划√；错误的方法划×)

√ (1) 一天的多变量运行，足以使研究小组采集到7%重要的变量中的80%以上。

× (2) 但是，绿Y仅被当成是一种属性——所有喷漆缺陷的总数被漏下了。根据尺寸、颜色或缺陷的形状，应该可以通过利克特度量尺度将其转变为一个变量，以便使其对缺陷有较大的灵敏度。

× (3) 每小时都对时间对时间变量族进行跟踪，但是应采用100%的检验。采样量是2 000件，真是一个庞大数字。总采样量应该可减少至400件，应能找到28个缺陷。另外，如果利用1到100分级的利克特度量尺度，加权的缺陷数量应可扩展到最大200。这将在变量族之间提供充分的差异。每小时应可采样50个部件，每次5件，这样一来，每小时要对5个部件采样10次。

× (4) 部件对部件变量族——即3~5个连续部件——没有进行鉴别，所以1个关键的变量族被遗漏了。

√ (5) 在部件内变量族里，对2个子变量族系进行了适当的鉴别，即顶部、中间、底部以及桶内径处等部位。

√ (6) 随时间产生的变量以及顶部对底部的变量不是重要的变量。

√ (7) 红X是在12点钟，这一点是找对了。

小组练习3：瓦片的黏结

在对安放在一条垫带上的瓦片的黏结进行多变量研究中，表8—3显示了每条垫带内的变量、垫带对垫带的变量、时间对时间的变量（读数越高，黏结的强度越高）。

表 8—3

多变量研究：瓦片的黏结

	上午 8:30			下午 1:00			下午 3:00		
垫带号	11	12	13	267	268	269	314	315	316
瓦片号									
1	66	59	54	60	57	47	38	14	56
2	56	58	32	53	37	45	09	43	39
3	58	56	59	44	46	48	54	08	60
4	65	48	48	50	44	49	57	38	58
5	67	63	72	58	52	56	60	60	60
垫带平均值	62.4	58.8	53.0	53.0	47.2	51.4	43.6	32.6	54.6
平均时间	58.1			50.5			43.6		

图 8—11 为根据上述数据绘制的数据图 (H 和 L 分别表示每个垫带中的最高和最低拉力; X 是每个垫带的平均值, O 是在每个时间间隔里所有 3 个垫带的平均值)。

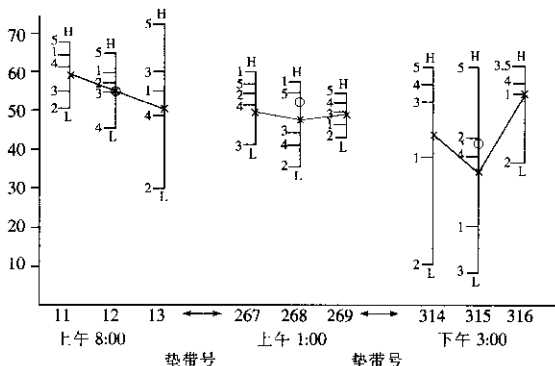


图 8—11 多变量研究：瓦片的黏结

注:

O—一次测试 15 片瓦的平均强度;

X—每条垫带上瓦的平均强度;

1, 2, 3, 4, 5—每个瓦片的强度。



问题

- (1) 红 X 变量族是什么？校验你的选择。
- (2) 粉红 X 变量族是什么？校验你的选择。
- (3) 你所探测到的非随机趋势是什么？利用图 8—11，请你至少找出 3~4 条线索。

答案

(1) 垫带内的变量是红 X 变量族，它具有最大 61 个部件的变量（见表 8—3 和图 8—11）。

(2) 垫带对垫带变量是粉红 X 变量族，具有最大 22 个部件的变量（在下午 3:00 垫带 2~3 之间）。

(3) 趋势和线索：

- 在红 X 变量族里（垫带内），最有可能的非随机趋势是瓦片 5 的黏结（通常是最高值的）和瓦片 2（通常是最低值的）。

- 每个垫带里具有最高黏结力的瓦片之间有非常好的吻合性（60~70 个部件），而最低黏结力的瓦片则呈现出大的变量（8~56 个部件）。

这指明，试验设计具有极大的应用性。

试想一下，为什么同样的设计、同样的生产过程、同样的材料、同样的操作人员、同样的测量用具，我们会同时得到好部件和差部件呢？好部件和差部件的差别是什么？工程师和解决问题的人员的传统做法是仅仅调查差部件——第 1 和第 5 个瓦片，即与缺陷的 10% 相关联的瓦片。他们的条件反射般的反应是改变设计，改变生产过程，改变供应，并且责骂可怜的操作人员，而不想去调查为什么应用同样的“配方”会有 99%、95% 或 90% 的好产品。他们从不想费心思去调查好产品和差产品之间的差别。

然而，这种好、差部件间的差别，形成了后面 3 个思路创新技术的基础——即部件搜索、成对比较及生产/过程搜索。



●时间对时间变量是浅粉红 X 变量族，其最大变量发生在下午 3:00，最小变量发生在下午 1:00（这也指明了垫带内和下午 3:00 的读数之间的相互作用）。

●垫带 1 和 2 有同样数量的变量，而垫带 3 的变量不同。图 8—12 描述的是此练习的族谱。

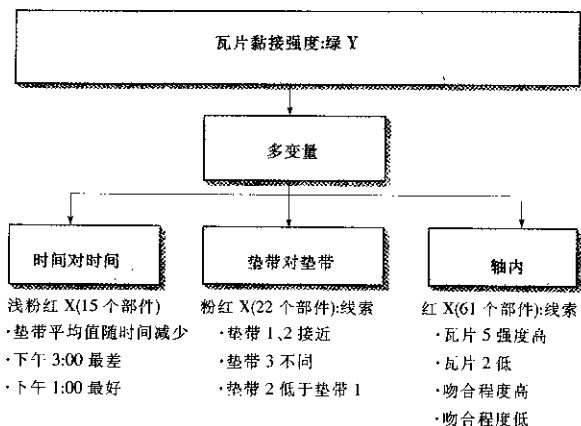


图 8—12 瓦片的黏结：族谱

小组练习 4：不成功的结果——皮带钝化剂多变量的案例研究

在一家半导体制造厂的硅晶片生产部门，CVD 二氧化硅玻璃被用做两层金属连接线间的绝缘层。对玻璃膜的厚度的控制是很严格的，且目前的生产设备生产的是无误差硅晶片。



有 3 个托盘，每个托盘装 3 片硅晶片。托盘安放在传送带上，由传送带将它们送进反应室。在反应室内，气体起反应，将掺杂磷的玻璃沉积在硅晶片上。此化学反应的速率是个常数，假设玻璃厚度取决于皮带经过反应室的速度。

生产过程的技术规格要求是 5 000 埃 $\pm$ 1 000 埃。^[1]在托盘中间位置上的硅晶片的中心部位进行测量，如果这一读数超过 4 000 埃~6 000 埃的规格限度，就要对传送带的速度调整一下。

以往的制造过程显示，钝化膜厚是个重要的变量，这种变量被认为是皮带速度随时间的改变而造成的。

后来展开了多变量的研究。在一整天内，研究人员每过 4 个小时对 3 个托盘里的硅晶片进行测量，并对每个托盘里的 3 个硅晶片的两个部位进行测量，如图 8—13 所示。表格化的数据列于表 8—4 中，图 8—14 是多变量图。

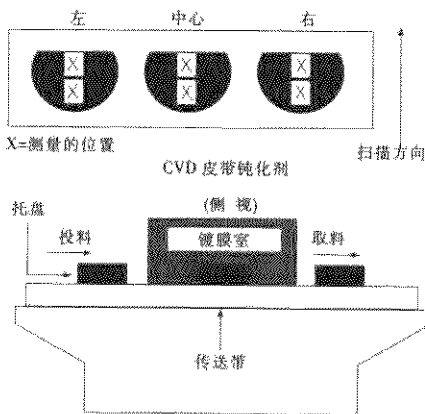


图 8—13 小组练习 4：皮带钝化剂

资料来源：Motorola, Phoenix, Ariz.



表 8-4 小组练习之四：皮带钝化剂多变量数据

硅晶片位置												
时间 托盘 次序		左			中			右			托盘平 时间平 均值 均值	
		上	下	平均	上	下	平均	上	下	平均	均值	均值
上午 8:00	1	5200	5350	5275	5000	5200	5100	5300	5350	5325	5233	
	2	4750	4950	4850	4900	5150	5025	5200	5300	5250	5042	5339
	3	4950	5350	5150	5750	6050 *	5900	5950	6400	6175	5742	
中午	1	4150	4300	4225	4050	4150	4100	4350	4550	4450	4258	
	2	4250	4300	4275	4250	4400	4325	4350	4550	4450	4350	1367
	3	4400	4650	4525	4200	4350	4275	4650	4700	4675	4492	
下午 4:00	1	3750	4050	3900	3950 *	4250	4100	4700	5000	4850	4283	
	2	5350	5550	5450	5700	5750	5725	5650	5800	5725	5633	5211
	3	5450	5600	5525	5650	5350	5500	5800	5850	5825	5617	
下午 8:00	1	5800	5950	5875	5700	5850	5825	5850	6050	5950	5883	
	2	5400	5500	5450	5550	5600	5575	5500	5700	5600	5542	5608
	3	5250	5450	5350	5300	5400	5350	5450	5650	5500	5400	

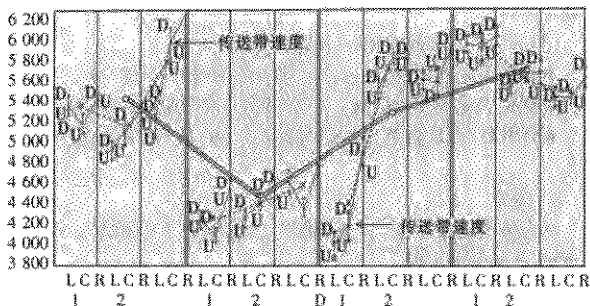
* 无论何时，在测量中间的硅晶片出现超差时，都要对传送皮带传送速度进行调整。

资料来源：Motorola, Phoenix, Ariz.

问题

- (1) 绿 Y 是什么？它可以量化和测量吗？
- (2) 相对仪器精度的规格宽度而言，测量系统的精度可以满足 5 : 1 这一要求吗？
- (3) 多变量的运行是否长得足够采集到至少 80% 的重要的变量或规范公差？
- (4) 在这次多变量研究中，变量族系是什么？
- (5) 红 X 变量族系是什么？
- (6) 粉红 X、浅粉红 X 族系是什么？
- (7) 在每个族系里，变量的量化值是什么？
- (8) 多变量图可以探知到什么非随机趋势？
- (9) 试验进行的方法有什么不妥之处？
- (10) 较好的多变量试验方法是什么样的？





托盘时间

图例:

D 和 U=单个硅晶片下,上的读数

O=单个硅晶片平均值

X=单个托盘里左(L),中(C),右(R)位置基片的平均值

O=平均时间

变量

l=硅晶片内

——=硅晶片对硅晶片

....=托盘对托盘

——=时间对时间

图 8—14 小组练习 4: 皮带钝化剂多变量图

资料来源: Motorola, Phoenix, Ariz.

答案

(1) 玻璃厚度变量。它是以埃为单位来量化及测量的。

(2) 是的。厚度的测量可以接近至 50 埃 (见表 8—4 中以增量 50 区分各个读数), 规范的宽度是 2 000 埃 ($\pm 1\,000$ 埃), 得到了 40:1 的精度。

(3) 是的。规范公差是 $\pm 1\,000$ 埃, 即一个 2 000 埃的分布范围。多变量的范围是 3 750 埃~6 400 埃, 或 2 650 埃。该值大于规范宽度的 100%。

(4) 时间对时间 (4); 托盘对托盘 (3); 硅晶片对硅晶片 (3);



硅晶片内(2)。

(5) ~ (7) 根据表 8—4 和图 8—14:

- 红 X 是托盘对托盘, 4 253 埃~5 633 埃=1 350 埃。
- 粉红 X 是时间对时间, 4 367 埃~5 608 埃=1 241 埃。
- 浅粉红 X 是硅晶片对硅晶片, 5 150 埃~6 175 埃=1 025 埃。

埃。

8. 趋势:

● 在每个硅晶片里, “下读数”总是高于“上读数”, 72 次观测中的 71 次都是这样。

● 在 12 次观测里, 每个托盘里右面的硅晶片总是高于左面和中间的硅晶片。

● 每当调整传送带速度时, 玻璃膜厚度都会出现很大的波动, 可以认为主要是时间对时间变量平均值方面的差别。

● 除了调整传送带速度期间外, 托盘对托盘变量的平均值相对较小。

● 2 个差托盘 (#3 号盘上午 8:00 和 #1 号盘下午 4:00) 都曾要求对传送带速度进行调整; 这些差托盘中的每一个都显示出从左面的低厚度值到右面的高厚度值这样一种趋势。

9. 试验中的失误:

● 传送带速度的调整太频繁。一个好的多变量研究无须进行调整或少进行调整就能看到过程中变量的本质。传送带速度的调整引起了膜厚的很大波动。

● 调整仅仅基于中间硅晶片的读数。因为从硅晶片至硅晶片变量(浅粉红 X)改变很大, 所以即使是必要的调整, 也要基于所有 3 个硅晶片的平均值或中值。

● 在每个硅晶片内, 只有 2 个读数, 即上读数和下读数; 每个硅晶片有 100 多个位置, 要至少采样 5 个位置才能确定硅晶片内的变量。

10. 较好的方法:

- 多变量不需要运行整整 12 小时。在 4 小时之内, 100% 的规

范变量或允许的误差都可采集到。

- 不要在刚刚启动时（早 8:00）调整传送带速度。这是一种无用的过分调整的典型情况，这样做只会搞乱所要观测的变量的自然属性。

- 采样频率（每 4 小时一次）太低。对 3 个托盘的采样要每小时做一次，以便较好地确定时间对时间变量族。

- 每个硅晶片采样 2 次是不够的。应至少在每个硅晶片上的 5 个点，比如说北、南、东、西和中间等 5 个位置进行采样。

- 红 X、粉红 X 和浅粉红 X 族系变量的所有结论，都是经过人工调整传送带速度之后观察到的；如果不进行调整，就会得出不同的结论。

- 有两种明显的变量，它们与带速调整无关，这是一个预警信号，应该马上采取行动。这两个变量是：（1）在每个托盘右面的硅晶片总是高于托盘中左面和中间的硅晶片；（2）每个硅晶片的“下读数”总是高于“上读数”。

案例研究 4：“墓碑现象”

经常会发生这样的现象：在很长的生产线上，红 X 变量被埋没在该生产线上的某个过程里。

在此情况下，进行了一个给定测试点的初始多变量研究之后，就有必要对生产过程中的每一步进行渐进式多变量研究。这就像在医生的办公室或医院里，采取 EKG 医疗手段一样有效：把数个探头放置于心脏、手臂和腿部区域，以便精确地定位病变的部位。同样，在长生产链中的多变量研究中，多变量总是定位于最先出现问题和经常产生问题的地方。图 8—15 是一个生产流程的典型例子，它的主要问题（绿 Y）是：在长生产链中的某个环节里，某部件引线的另一头与母板脱离。在这种现象里，部件不是平放在那里，而是像一个墓碑立在那里，所以叫做“墓碑现象”。此时废品率在 7 000ppm 左右。



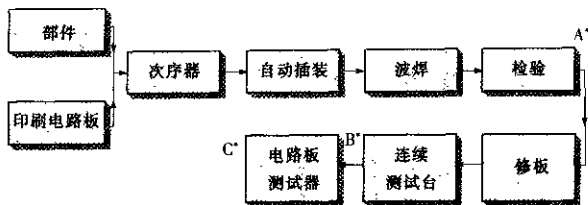


图 8—15 一个印刷电路板的工作顺序

* 检测点。

资料来源：Philips, inc., Singapore.

上面的例子中，生产过程的每一道工序在检验后的以及在两个检测点（A、B 和 C）测量后的初始多变量研究之后都进行了渐进式多变量研究。这些初始的多变量研究表明：有两种电容在印刷电路板的 3 个部位重复出现缺陷。怀疑其原因可能是：电容引线的直径、电容引线变形、印刷电路板上一个指定位置上孔的尺寸弯曲变形的电路板顺序器上电容的朝向；自动插装机失调、自动插装机编程错误、波焊机的参数变化或在连续检测台（针式检测台）上插针接触点的压力不均衡等问题所致。

渐进式多变量表明：在生产链上的一个环节上有一个小的变量，找到它之后，自动插装机的缺陷水平就再也没有显著增加。必须调整自动插装机上与电容及其相关的孔位有关联的滴管，这样做之后，缺陷率才下降至 0。于是，可利用正向控制（见第 18 章）来定期地校准滴管。

多变量在管理工作中的应用

在过去的 10 年里，作者已把 DOE 技术扩展到非生产性的应用领域以及管理层的白领领域中，每种 DOE 技术在这些领域里都可得到应用。



本节将围绕多变量技术列举 3 个例子。而生产过程中的 3 个主要变量族——时间对时间、部件对部件、部件内——不再适用于管理工作，在这个领域里似乎有更多的变量。

案例研究 5：医院记账单的错误

医院因提供给病人的记账单有误，常被弄得声名狼藉。病人对医院、医疗保险公司和医生的投诉不断，有些案件甚至拖几个月都解决不了。

在大城市里的大医院中，出错的记账单占全部病人记账单总数的 30% 以上。作为初期的调研，医院决定在可能包含红 X 变量的 300 个纠纷投诉中，建立几个变量族，这些变量被划分为以下几个种类或“族”：

- (1) 病人住院的时间；
- (2) 对病人的治疗（复杂—简单）；
- (3) 记账员；
- (4) 主治医生；
- (5) 护士；
- (6) 保险公司；
- (7) 账单日期；
- (8) 处方药品。

在这些族系里，存在着一些变量，但在关于主治医生可以采取哪些医疗检查的问题上，保险公司是有异议的。这些异议是用红 X 表示的。于是就把病人抛在了一场打乒乓球似的拉锯战中。而后，医院建立了一套对医生和病人都公平的基本规章，同时又兼顾到保险公司想把与日俱增的成本降低下来的要求。

案例研究 6：饭店里的顾客服务问题

调查顾客的满意度是各大饭店的通常做法。虽然饭店的每



间客房都有征询顾客反馈意见的表格，但反馈情况并不理想，反馈率不到 10%，而且即便顾客填写了反馈意见，也很少有认真填写完各项内容的。饭店的管理部门决定由训练有素的经理亲自与顾客个人进行深切的会晤。采样量是 200 名，受访者将回答以下 10 种问题：

- (1) 靠近/进入饭店/泊车；
- (2) 前台；
- (3) 看门人；
- (4) 房间的环境/房间整理；
- (5) 客房服务（食品和饮料）；
- (6) 餐厅；
- (7) 饭店设施（包括娱乐）；
- (8) 商务服务；
- (9) 关怀的气氛；
- (10) 未预料到及未曾指望的经历。

然后，利用上述分类法对这些自由自在的会晤进行概括综合，由顾客依重要性按 1~3 个等级打分，而后顾客再按每种类别分为 1~5 个级别为饭店评分。

这样征询意见的结果如同前述的表格方法一样能得到顾客满意度的大致情况，但关怀的气氛和未预料到的经历这两条是个例外，顾客给这两条在重要性和等级上打的分数，要高于其他评分。这个事实说明：饭店的管理人员抽出时间，不怕麻烦进行个人会晤，本身就表达了饭店对顾客的关怀，这也是一种未曾预料到的惊喜。另一个未曾预料到的惊喜是：顾客入住 10 分钟之后，前台来探询房间里是否有不合意之处。第三个惊喜是：饭店的员工随时准备灵活掌握饭店的规章，以满足顾客的需要。

这些关怀和未预料到的惊喜所产生的长期影响是，客房的入住率和回头客大幅度增加。



案例研究 7: 雇员的频繁调离

某公司是劳动密集型企业, 管理部门认为员工的调动过于频繁。由于公司自身的劳动力资源相当稳定, 员工的流动性很低, 再加上公司具有良好声誉, 所以这一点尤为令人不安。

人事部门决定针对离职的员工开展多变量研究, 并把收集来的数据按下列类别划分多变量族或种类:

- (1) 年龄;
- (2) 婚姻状况;
- (3) 上班的距离;
- (4) 在公司的服务年限;
- (5) 离职时所陈述的理由;
- (6) 工资级别;
- (7) 上班的班次;
- (8) 所属部门的类型。

从年龄、婚姻状况、上班的距离、离职的理由^[2]或上班的班次等多变量研究的结果来看, 没有较大的区别。最主要的变量在于服务的年限, 6 个月至 1 年是一个主要的周期; 另外还有工资级别和工作部门等两个主要变量。

在上述的 3 个原因当中, 工资待遇问题是可以理解的, 大量的辞职雇员属于低工资阶层。公司不想改变其工资政策, 因为在其行业领域里, 该公司已经属于最高工资企业。在工作部门类型的类别里, 油漆商店的问题最突出, 其原因是恶劣的工作条件和保健问题。公司采取强有力的措施——应注意的是, 还包括其他的 DOE 技术——以改善内务管理、工作环境和员工精神状态。

于是就剩下 6 个月至 1 年的工作年限这样一个子变量种类问题作为重点因素来加以探讨。公司采用了另一个 DOE 技术——成对比较, 以确定其潜在原因 (见第 11 章)。

关于多变量技术, 我们现在以提一些问题的方式来为本章作结论。这些问题包括一个 DOE 小组在开展多变量研究时应问自己的



问题，也包括主管经理在检查 DOE 小组的工作时要问的问题。这些问题为解决实际问题提供了帮助，以免误入歧途。（某些问题与绿 Y 有关，并已在第 7 章加以阐述，为了强调其重要性和保持连贯性，在此我们又重复一次。）

DOE 小组的问题

(1) 对于绿 Y，已经清楚地定义、量化和度量了吗？

(2) 测量仪器的精度是否至少为产品误差分布范围或规范允许误差的 5 倍，无论哪一个更大些？

(3) 如果绿 Y 是一个属性，那么利用利克特度量尺度，可以把它转化为一个变量吗？

(4) 是否有一个更快、更便宜、更简单的绿 Y 来近似地代替最初的、更恰当的绿 Y？

(5) 如果为了确定绿 Y 所花费的时间太长（典型的是，当一个问题在出厂后几个月才在现场发现），为简化一个重要的 DOE 试验，是否建议了利用多环境强化应力试验来加速绿 Y 的出现（例如一个现场的失效）？这样可以在 1 小时或 1 天内最大限度地模拟出绿 Y，而不是等上几周或几个月。在应用 DOE 之前，可以用 MEOST 把一种间断的失效转化为永久的失效。

(6) 绿 Y 的数目不止一个吗？如果是这样，多变量（或任一 DOE 技术）可以从中选出最重要的绿 Y 来。

(7) 如果有一个以上的绿 Y，是否已根据最高成本、最大缺陷和/或缺陷率以及对顾客影响等因子来选择一个最重要的绿 Y？

(8) 绿 Y（或需要解决的问题）对于顾客真是那么重要吗？没有什么比把钱花在对顾客无效的事情上更浪费了。例如在轻工业，化妆品的缺陷是一个老大难问题。其中一项研究表明：它的典型的做法是每一种化妆品要淘汰掉顾客不喜欢的 25%~40% 的缺陷。建立一个有顾客参与的化妆品标准，可以节省几百万美元和大量时

间。

(9) 如果对于高缺陷率或低产量有一个帕累托原因分布图, 优先将最普遍的缺陷模式作为选择绿 Y 的手段了吗?

(10) 在一条特定的生产线上, 如果有许多模式正在研究, 拥有最高废品率的模式被单独划出作为 DOE 试验的起点了吗?

(11) 在时间对时间变量族里, 总试验时间是按 80% 规则 (即运行时间采样, 直至采集到 80% 的重要的变量或规范公差) 来确定的吗? 就像许多想解决问题的新手们所试图做的那样, 在处理范围宽广的各种模型和大范围的缺陷模式时, 很难抓住要点, 其结果只能是无效、挫折和失败。

(12) 如果已知在同样的时间段里, 变量来自部件对部件, 可以不用多变量技术, 而进行部件探索或成对比较。

(13) 在部件对部件族里, 连续采样的数日 (通常是 3, 4 或 5 件) 确定下来了吗?

(14) 在部件内变量族里, 为了组成变量族谱而指定合适的采样数量, 那么确定了生产线、机器、电机、测试仪器、操作者以及每个部件内的位置的数目了吗?

(15) 在每次时间采样里, 用连续部件的数量、部件内采样方位数来乘时间采样数所确定的数量, 是整个多变量研究所要求的总数量吗?

(16) 这个总数量 (见上述第 15 个问题) 在低成本与必须满足 80% 的规则间采用了合理的折衷方案吗? 一种建议方案是在时间对时间族里仅采样 3 个, 然后在必要的情况下扩展时间采样, 以实现 80% 规则。

(17) 已经把显示所选择的族系及子族系的族谱画出来了吗?

(18) 已经绘制出用于记录由重要变量族和子族系里收集到的数据的表格了吗?

(19) 在分析趋势时, 利用旧有的、重要的数据是否不被提倡? 大多数重要的数据值得怀疑, 除了净缺陷百分数之外, 价值不大。没有人能准确把握是谁收集的数据, 是怎么收集到数据的, 是在什

32
40

么条件下收集到数据的，又是在多长的时间内收集到数据的。多变量的目的是重新开始并在严密控制的条件下收集合理的数据。

(20) 对生产维修人员已发过指示，让他们在多变量研究期间不要调整任何过程控制吗？如果这是绝对不可避免的，就要精确地记录这样一种调整的性质和时间。

(21) 在确定时间样本的精确时间时，要重视与典型的非连续性相关的时间段，如休息喝咖啡的前后、午餐、换班等，因为在这些间断过程中会发生对产品的重大影响。

(22) 在分析多变量图，寻求线索时，检验过非随机趋势吗？

DOE 项目评审时高层管理的问题^[3]

(1) 什么是绿 Y？量化和测量过绿 Y 吗？仪器精度是产品精度的 5 倍吗？

(2) 绿 Y 对顾客有多么重要？

(3) 绿 Y 问题的持续周期是多少？缺陷的级别如何？对公司的成本影响如何？

(4) 谁是工作小组的成员？是否有小组的协调部门？

(5) 为什么选用这种 DOE 技术，而不用其他 DOE 技术？

(6) 如何选择变量的子族系？

(7) 绿 Y 是在工厂或现场使用时发现的吗？如果仅是在现场使用时发现的，是否建议过采用多环境强化应力试验来加速并模拟同样的在工厂的失效模式，以便较早地检测到失效部件并采取预防措施？

(8) DOE 试验时间多长？准备时间多长？运行时间多长？

(9) 红 X 和粉红 X 被定位了吗？或为了追踪由多变量所产生的线索，必须在工厂或者对供应商进行随后的 DOE 试验吗？如果是那样，哪一种 DOE 技术最适当？

(10) 工作小组所经历的最大障碍是什么？是来自管理部门还

是来自保障部门？

(11) 管理层提供的资源充分吗？

(12) 从这个计划得到什么教训？是针对工作小组还是针对管理人员？

【注释】

[1]：1 埃等于百万分之一厘米的百分之一。

[2]：在与辞职的雇员进行最后会谈时，有一个自然的趋势：辞职的雇员总是很小心、很警惕，他们不想自断退路，所以你往往发现不了辞职的真正理由。

[3]：这些高层管理人员要提问他们的 DOE 小组的一般性问题，对每种 DOE 技术都是一样的。后面章节里将不再重复。在后面每一章的结尾，高层管理人员要提的问题肯定出不了这些与指定技术有关的问题的范围。



第9章

集中图：重复出现问题的精确定位

目的

如果多变量研究表明重要变量族是在部件内，那么下一步就应当是绘制部件内问题的精确位置。用集中图就可以很好地做到这一点——其通常的称谓是“斑点图表”。在解决问题路径图（见图7—3*）中我们可以找到它的位置所在。

集中图表明：（1）重现的故障问题没有一个具体的位置（即它是一个随机分布）；（2）它的确显示出在某个特定位置有缺陷集中的现象。经验告诉我们：根据通用的帕累

* 原文误为图7—1，全书还多处引用该图，不一一指出。 译者注

托规则，造成许多现象的原因从来都不相同。下面的结论也几乎源于同样的真理：在部件内的某个具体位置上有一个问题集中的限度，如果该部件是一个组件的话，又可把它说成是一个部件内指定的部件或元器件。

在上一章的多变量案例研究中，我们已经看到了部件内的红 X 族变量的例子。在第 8 章小组练习 2 中，红 X 族是在垫条内——特别是瓦片 5（重复出现率高）和瓦片 2（重复出现率低）。在第 8 章的案例研究 3 中，红 X 族是计算机检测站，检测站 A、C 所检测出的“找不到故障”的收音机数量是检测站 B、E 的 10 倍。

制作集中图

- (1) 制作一张包括重现缺陷的图或模板。
- (2) 如果需要，画一些网格，这样可以找出问题的精确位置。
- (3) 在检查每个部件时，请检验人员标出如下内容：
 - 每种缺陷类型的位置以及每种缺陷类型的适当的编码；
 - 在每个位置上每种缺陷类型的数量（不必记录每种缺陷出现的时间，除非时间对时间变量被认为是重要的）。

下面几个实例可以用来很好地说明集中图技术。

案例研究 1：冰箱门上泡沫保温材料的漏气

某家电制造厂的冰箱生产线在用泡沫保温材料填充冰箱的门和衬里时，出现了漏气现象。该问题已存在了几个月，当时的缺陷率为 8%，返修成本是每年 120 000 美元。操作员曾因此备受指责。第 2 班和第 3 班曾经是怀疑的重点。

该厂请作者去解决这个问题。于是，作者对 3 个班次的工作进行了多变量研究，并将其分为时间对时间族、门对门族和门内族 3

Step 3

个变量族系。在时间对时间（包括班对班、小时对小时的采样）以及门对门族系变量里没有发现显著的变化。在操作者对操作者族系里也没有显著的变化。泡沫塑料漏气最经常出现在门内变量里。在门的右上角、右下角和左下角这些填充泡沫的地方几乎没有漏气现象，集中图定位于左上角的位置。经确认：操作者总是由门的左上角开始填充泡沫塑料，然后沿顺时针填充其他各部位。这使研究的重点转向填充泡沫的起始过程，结果发现在左上角发泡的初期阶段，编程上有 2 秒钟的延时。当消除该延时之后，缺陷降为 0。整个试验，包括采取纠正措施，仅用了不到 2 周时间。

案例研究 2：波焊的缺陷

作者曾在其公司位于澳大利亚的某工厂主持了一个 DOE 研讨会后，受邀解决该厂在某种高容量印刷电路板中出现的波焊缺陷问题。该印刷电路板上的缺陷率为 1 500ppm 左右，而当时，即十多年前，该公司就为其生产工艺制订了 500ppm 的最大容许缺陷率。

于是，作者着手进行了下述多变量研究：（1）分别在上午 9:00、9:30 和 10:00 3 次采样；（2）采样 10 个控制面板，每次每块控制面板放 5 块印刷电路板；（3）绘制一个集中图，用以说明每块印刷电路板内焊接缺陷的分布情况。挑选资深的检验人员来标注出有缺陷的焊接处的位置和数量。图 9-1 显示的是所考虑到的变量的族谱。

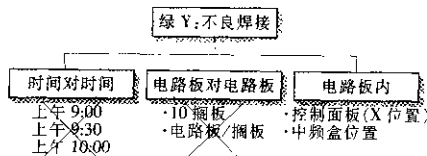


图 9-1 印刷电路板焊接的多变量族谱

资料来源：Motorola Inc., Melbourne, Australia.

检验人员的报告表明：上午 9:00 至 10:00 的焊接总缺陷率是 1 450ppm，这样一来，采集到 80% 或以上的 1 500ppm 的重要缺陷这一规则的要求，仅仅在 1 个小时内就实现了。经多变量研究发现各印刷电路板之间在 9:00、9:30 和 10:00 并没有显著的变化。同时，控制面板对控制面板、印刷电路板对印刷电路板之间也没有任何显著的变化。

最大的变量是在印刷电路板内，主要是在 3 个位置。首先是焊接处的针孔成簇状分布在印刷电路板的一个边棱的中间部位，如图 9—2 所示；另一些则位于与中频盒有关的两处焊接薄弱部位（见图 9—3）。其余的 990 多处焊接处都是完好的。

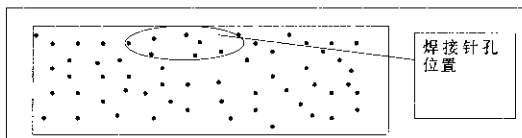


图 9—2 焊接处的针孔（印刷电路板底面）

资料来源：Motorola Inc. , Melbourne, Australia.

对于图 9—2 所示的集中图，明显会产生这样的疑问：为什么在板的其他 3 个边棱没有薄弱焊接缺陷？多变量研究小组认定，问题在于波焊时用来放置控制面板的夹具。夹具的水平性不好，向一面倾斜，导致这个边上的焊剂没有很好地浸润。将夹具调正之后，针孔缺陷随即消除。第 2 次修正后的变量表明焊接缺陷由 1 450ppm 降至 500ppm。仅仅用了 3 小时，缺陷降低率就达到 3 : 1。

然后，小组的注意力又集中于与中频盒缺陷有关的部件的焊接缺陷上。如图 9—3 所示，电路板上的孔允许焊剂抽吸到孔里去可后来焊剂又塌下去了。小组试用了两种解决方法，一种是中频盒引线预先涂焊剂，另一种是往引线上涂助焊剂。

没有一种方法能显著减少 500ppm 这样的缺陷率。孔的大小与引线直径大小的比例在印刷电路板设计中是很重要的参数。在生产

中希望两者的比例要大一些，以便接插部件；而在质量控制中则要求两者的比例要小，以便得到良好的焊接质量。

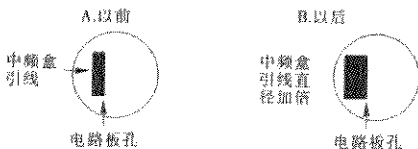


图 9-3 印刷电路板中中频盒的可焊性

资料来源：Motorola Inc., Melbourne, Australia.

中频盒的案例令人感到两者的比例太大了（见图 9-3A）。于是小组试着将中频盒的引线直径加大一倍，以减小比例（见图 9-3B）。结果缺陷降低为 0。运用多变量技术后表明是 0 缺陷。又生产了半天时间，肯定了 0 缺陷的事实！有人感叹：什么样的经典的或田口 DOE 试验能解决这样的问题，而且它们既浪费时间又没有成果！而多变量集中图仅用了 1 天半的时间就解决了问题！

小组练习 1：喷漆缺陷

在控制面板上进行喷漆的过程中，成品率平均仅为 82%。每 1% 的产量的提高，每年会给公司节约 45 000 美元。于是公司开展了多变量研究。经确定，重要变量不是时间对时间或控制面板对控制面板，而是控制面板内的变量。研究人员制作了一个集中图（见图 9-4）图中显示了 4 种缺陷（通过显微镜和截面检测）以及每种缺陷的数目和位置。

(1) 通过集中图，你从每种缺陷类型的分布情况中能得出什么结论？

(2) 根据各种缺陷的类型，你能猜测出其形成原因吗？



答案

(1) 铁铬镍合金缺陷不但是最常发生的缺陷类型（占缺陷总数的 75%），而且集中于控制面板顶边的中部，集中图清楚地说明了这一点。DOE 小组发现，控制面板是用铁铬镍合金制的钩子沿面板上棱的中部悬挂起来的；由于这些钩子没有进行定期清洗，所以钩子上的碎屑就掉到了控制面板上。应采取的预防性措施是，由过去的每季度清洗一次改为每两周清洗一次，同时重新设计钩子，这样就使总缺陷率由 19% 减少至 4.8%。

(2) 其他缺陷种类：玻璃、有机物和铁等，占缺陷总数的 25%，且都没有集中的位置。DOE 小组从控制面板到悬挂控制面板进行追踪，将“玻璃”缺陷归结于珐琅质。即使这些缺陷，没有重现的性质，但重新设计了小一些的钩子可使这类缺陷减少为 3:1。

			I: 9	I: 16	I: 18		
							G: 2
		G: 2					
		O: 2					
	F _E : 1						
					G: 1		
			O: 2				
							F _E : 1

缺陷代码	总缺陷数
I—铁铬镍合金缺陷	33
G—“玻璃”缺陷	5
O—有机物缺陷	4
F _E —铁缺陷	2

图 9—4 喷漆缺陷集中图

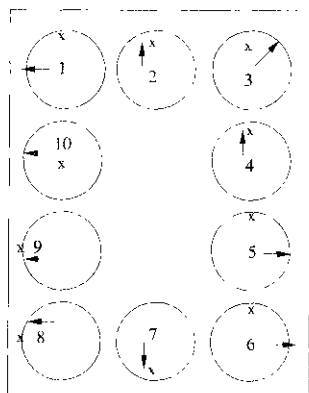


(3) 有机物缺陷可能是由于在进行打底漆或光亮喷漆时产生漏气而造成的；至于铁缺陷，被认为是控制面板的返工造成的，后来没有对其进行跟踪。

(4) 集中图促使产量提高了近 20%，每年大致可节约 90 万美元。

小组练习 2：热处理后的轴变形

某工厂的轴在经过热处理和淬火之后，发生了变形。工厂每次将 10 个轴放在一个容器里（见图 9-5），在华氏 1 500 高温条件下热处理 1 个小时，然后再进行大约 3 分钟的淬火。绿 Y 为消退，在热处理之前它是合格的，但在热处理之后就超差了。同时，在热处理和淬火之后，消退的方向也有所改变。



图例：“X”——热处理前径向跳动的方向
“↑”——热处理后径向跳动的方向

图 9-5 热处理中轴的摆放位置及热处理前后径向跳动的方向

资料来源：Caterpillar, Decatur, III.

某 DOE 小组研究了图 9—5 所示的集中图，结论如下：

(1) 热处理引起的消退方向的改变，轴的一半发生了改变，而另一半没有改变。

(2) 怀疑引起改变的原因为：

- 淬火：淬火溶液、淬火时间、淬火温度。
- 热处理：热处理时间和热处理温度。

最初认为淬火溶液是引起改变的原因。通过进行 B vs. C 的试验（见第 15 章）后，并未发现热处理前后跳动的变化；同样，在改变了热处理时间和温度之后，也没有什么变化。

问题

研究小组从集中图中，找到了什么重要的线索？应该采取了什么行动？

答案

实际上，有 5 个轴发生了消退方向的改变，但是其余的 5 个（见图 9—5）并没有变化。然而，10 个轴中有 9 个轴的消退的方向指向容器的外缘，这个重要的线索提示了轴外缘的温度与内侧的温度不同。实际的温度测量肯定了这一温度上的差别。于是对热处理室进行了改造，加上了一台风机，以使室内温度均匀，这使得消退在热处理之后，能达到规范的要求。

在管理工作中的应用

正如在多变量研究中的作用一样，在管理工作中，集中图也可很好地定位重现问题。

案例研究 3：短路的行输出晶体管

某电视机生产公司面临着一连串的行扫描问题，问题是出在短路的行输出晶体管上。失效分析表明，大量的短路源于一些无法解释的现象。在试验室里试图模拟这种短路的努力没有成功。为了找到线索，公司决定精确地绘制每种失效的地理区域图。

公司准备的一张集中图显示了在美国各地出现的行扫描失效的地点，该图清楚地显示：沿海地区的失效率远比内陆高。而沿海一带的雷雨天气远比内陆地区多，所以行输出晶体管必须重新设计以耐受电冲击。这个恼人的行扫描问题最终得到了解决。

案例研究 4：大学招生工作

某小型大学所获的赞助费不多，因而想增加申请人学者的名额。它没有降低入学标准，而是把生源目标转向了过去能吸引到大量申请者的地区。它把过去 5 年中所招来的学生分为 20 个地理区域，然后制作了一张集中图，以确定 20 个区域里哪一个是生源最多的区域。

与校方的预料相反，集中图显示，20 个区域中最吸引学生的 3 个地区都是离学校较远的地区。经过研究，校方发现了两个原因：第一个原因是，这些地区的中学生顾问劝导学生们到这所学校读书；第二个原因是学生们有选择与朋友们同校的趋向。

DOE 小组的问题

- (1) 在多变量研究中，红 X 族是在“部件内”吗？
- (2) 有 1 个以上的绿 Y（缺陷模式）要考虑吗？
- (3) 在确定绿 Y（缺陷模式）位置之前，准备了图纸、模板或计划表吗？

(4) 如果绿 Y 是一个属性，可以利用利克特度量尺度将它转化为变量吗？

(5) 故障重现位置和重现部件的出现频率有记录吗？

(6) 在集中图处理之后还要进行工程判断或其他 DOE 试验以找出红 X 吗？

DOE 项目评审时高层管理问题

(1) 在集中图技术中，是否用到成熟的多变量技术？红 X 是在“部件内”族系吗？

(2) 要用多长时间才能把“部件内”族系的 80% 或 80% 以上的绿 Y 找出来？

(3) 有与绿 Y 相关的位置或部件方面的集中吗？

(4) 要进行的后续工作是什么？



第 10 章

部件搜索：简便而平滑的交换

引言

部件搜索是我们用于“与部件对话”的第三个线索生成技术。正如在多变量技术及其辅助技术——集中图的案例中一样，部件搜索不打乱生产秩序，也不要求对生产技术有高深的了解，所以一线工人可以在不到一小时的时间内学会并应用于实践。而且，样本量也令人难以置信的小——只要两个部件，一个是“好”的，一个是“差”的。

一个幽默的类推：发生在英国火车上的故事

可以用四名乘英国的火车由伦敦的尤斯

顿车站出发到纽卡斯尔去的乘客的故事来说明部件搜索的原理。四名乘客坐在火车的包厢里，如图 10—1 所示：一位老妇人 (L)；旁边是一位年轻妇女 (YW)；年轻妇女的对面是一位军队的将军 (G)；将军的旁边是一个士兵 (P)。

老妇人 (L)	年轻妇女 (YW)
英俊的士兵 (P)	将军 (G)

图 10—1 在火车包厢里的 4 个乘客

英国人很古板，没有适当的引见，不会互相搭话。大家就这样在车厢里坐着。当列车经过一条长长的隧道时，不知是什么原因，包厢里的灯突然灭了一段时间。就在灯灭不久，四个人都听到了一声亲吻声，随后就是一记响亮的耳光声音！于是四个人都开始猜测谁应对这起事件负责。

●老妇人想：“一定是这个将军。这个老山羊，他亲了年轻妇女，而年轻妇女打了他一个嘴巴。”这是老妇人的理论。

●年轻妇女想：“真该死，我是个漂亮人儿，那个英俊的士兵想亲吻我，却错亲了那个老太婆，于是就挨了她一巴掌。”这是年轻女人的理论。

●那位将军就是挨了一记嘴巴的人，他一面揉着脸颊一面在想：“一定是那个年轻的士兵干的。这个军队的渣滓，我奇怪这种人怎么会混到军队里来！大概就像以前那个副总统奎尔，在国家卫队里呆过，或者更像克林顿总统从国民卫队里混出来一样。总而言之，他就是冒犯者，是他亲吻了年轻女子，而我却因坐在年轻女子对面，首当其冲受到怀疑挨了打！”这是将军的理论。

●包厢里只有这个士兵知道事情的真相。实际上，他趁车厢里漆黑一团和特别安静之机，亲了自己的手背一下，同时又打了将军一个嘴巴。任何人都未发现他所做的这一切。

这就是部件搜索的原理，只是你不知道这是交换理论。实际

上，你已将部件或子部件，从好的至差的或由差的至好的进行了交换，看一下问题是由被交换的部件带走，还是留在剩下的部件中。部件搜索原理就这么简单！

通过部件搜索分流多变量

多变量研究通常可确定 3 个主要的变量族系：时间对时间、部件对部件、部件内。但是如果你已经知道重要的变量是存在于部件对部件时，该怎么做？如果在同样的运行下和同样的时间里都存在着“好”和“差”的部件时^[1]，又该怎么做？此时，可以绕开多变量，开始部件搜索过程。如图 10-2 所示，该图是图 7-3 的复制品，这里只是强调部件搜索的步骤。

消除对部件搜索的批评

(1) 部件搜索利用令人难以置信的少而又简单的两个部件——在其重复出现的绿 Y 值中，这两个部件应是尽可能相离较远的。许多经典的 DOE 实践者反对这样少的采样规模。他们问道，“仅仅用两个部件，你怎么能确定一个给定的参数（绿 Y）的分布？”他们那费力、费时又费钱的方法是在 2~3 天的时间里要对 50~300 个部件运动控制图以进行测定。

而在部件搜索技术里，仅需考虑极端的分布情况，如图 10-3 所示。采集到两个极端，就可以采集到整个绿 Y 变量的范围，然后再通过交换部件或子部件的方法把这些原因系统地过滤掉。

(2) 传统主义者的第二个顾虑是在部件搜索过程中发现的故障部件的重现性问题。他们会问，“如果选择的是不同的一对好、差部件，你怎么能保证同样的故障部件会出现？”答案在于所选择的是特别好的部件和特别差的部件。只要我们调查的是同样的绿 Y，红 X 肯定会从极端的变量中找到。几百次部件搜索试验已证实了这

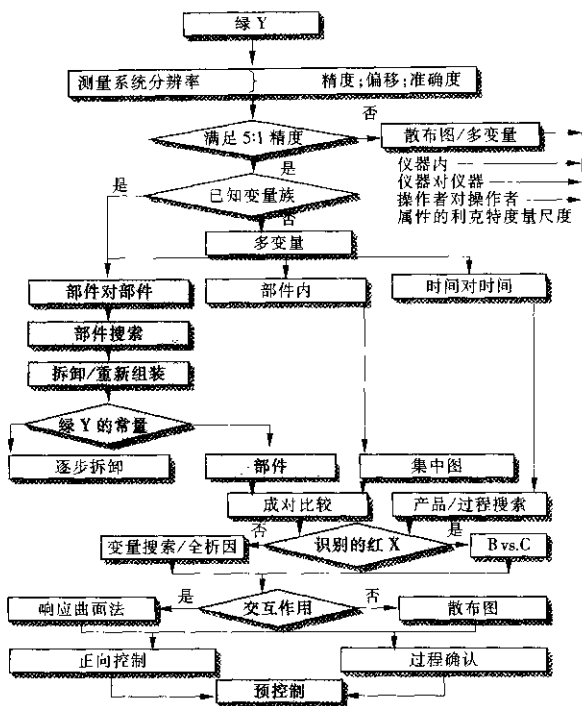


图 10—2 搜索红 X: 解决重点部件搜索问题的路径图

一点。在我们的经历中，只有两次不同的红 X 是从第二次部件搜索试验中找到的，这是由于时间因子造成的。

(3) 还有一些对部件搜索的批评认为：它是另一种类型的分析因。这显示出对部件搜索的优点及其诊断能力缺乏了解。分析因不能将主要效应和相互效应区别开来，这些“乱麻”可能会导致分析

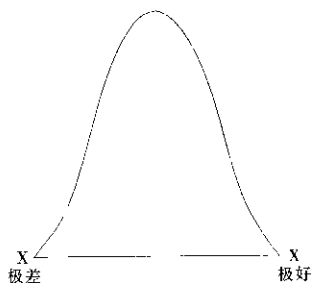


图 10—3 一种变量分布的两个极端：
重点部件研究实践

因产生错误的结果。

比起一味地肯定或否定一组预选的原因——正如经典 DOE 或田口 DOE 所尝试的那样，部件搜索技术实际是与部件对话，其系统的、对称的设计发掘出一些意外的控制因子，它们既可以是主要的原因，也可以是一些高阶的相互作用。

► 部件搜索的 4 个阶段

理解部件搜索试验中的 4 个阶段，将有助于消除批评者的误解。表 10—1 所示为每个阶段及其要达到的目标。

阶段 1：球场阶段。这个阶段里要确定在被研究的部件中是否选择了正确的变量，亦即它们是否在球场中。如果这个阶段得以通过，统计上在 95% 的置信水平下，试验者就可以转向第 2 个阶段。

阶段 2：排除阶段。第 2 阶段的日的不仅是要排除所有不重要的主要原因或因子，而且要排除与这些非重点原因相关联的相互作

用的效应。该目标可以利用成对交换试验来达到。随着每次交换，大量潜在的相互作用及其主效应（单个因子被改变）会同时被“激活”。在每个成对试验中，除了红 X 之外，就有 95% 的把握排除掉大量的主效应及其相互作用。

阶段 3：求交运算阶段。这个阶段可以证实第 2 个阶段中所选择的重要原因综合之后仍然重要；而在第 2 个阶段鉴别为非重要的原因，可被肯定为不重要。

阶段 4：析因分析。这个阶段不是进行另一次试验，而仅仅是对第 1、2、3 阶段进行分析，通过一个全析因矩阵去量化并分离出重要效应及其相互影响效应的幅度和方向。

在上述 4 个阶段中，每个阶段的方法将用后面的实例研究加以阐述。

表 10—1

部件搜索的 4 个阶段

阶段	目标
1. 球场	确定红 X 和粉红 X 是否包括在所考虑到的变量*中；这也保证了拆卸/重新组装的故障复现
2. 排除	排除非重要原因及相关交互影响的效应
3. 求交运算	确认重要变量确实重要，非重要变量确实不重要
4. 析因分析	量化重要原因及其交互影响效应的幅度和方向

* 整个文章中所提到的变量、原因和因子都是同义同，可以互换使用。

部件搜索：先决条件

● 该技术主要用于装配操作，从而可区别于面向过程的操作（尽管它可以用于具有变化的绿 Y 的过程和机械，也可以用于在这些机械中互换的部件）。

● 性能（输出或绿 Y）必须是可测量的，测量仪器的精度至少为技术规格要求公差或产品误差分布的 5 倍。

● “好”和“差”两种部件必须能够拆卸和重新组装而又不会

33

造成初始输出的显著改变。

12 个步骤程序

(1) 从 1 天的产品里仅仅采样 2 个部件，这 2 个部件（装配）距离有待研究的绿 Y 要尽可能地远，即很好的部件和很差的部件（简言之，1 个是最好中的最好——BOB，1 个是最差中的最差——WOW，BOB 与 WOW 相离越远，越容易找到红 X），还要测量它们的绿 Y 值。

(2) 阶段 1。拆卸并重新组装 BOB 和 WOW 两次，再测量它们重现的绿 Y 两次以上。

(3) 显著性检验。这个检验中要测 2 个部件，以便确定从统计上看是否 BOB 与 WOW 间的差别是显著的。

● BOB 的 3 个绿 Y 值必须全都好于 WOW 的绿 Y，并且它们之间没有交叉覆盖；

● $D/\bar{d}$ 的比率必须大于 1.25，或者最小要等于 1.25。[D 是 BOB 中值与 WOW 之间距离， $\bar{d}$ 是 3 次 BOB 测量和 3 次 WOW 测量的测量值范围间的平均差（或不重现性），下面的案例研究将进一步阐明显著性检验。]

(4) 如果 $D/\bar{d}$ 的比率小于 1.25，部件搜索的第 1 阶段就是失败的。这意味着绿 Y 不能保持稳定不变，说明问题出在组装过程中（即部件组装的顺序中），而不是出在部件本身。这就要求开展逐步的步进式的拆卸和再组装，以便确定组装过程中的哪一步是红 X（见图 10-2）。如果 $D/\bar{d}$ 的比率大于或等于 1.25，红 X 就会是 1 个以上的部件或子部件。此时，阶段 1 是成功的，可以开始进行部件或子部件的交换步骤。

(5) 最好由子部件开始，以递减的次序排列子部件；如果没有显著的子部件，那么就按递减的次序排列部件。

(6) 阶段 2。将最高一级的子部件或部件（称为部件 A）从



BOB 转接到 WOW，并将其对应者由 WOW 转接至 BOB；测量并记录两个新绿 Y 值。

(7) 在第 6 步，可能会有 3 个结果：

(a) BOB 依然是 BOB，即好的；WOW 依然是 WOW，即差的。这个结果意味着所转换的部件 A 是不重要的。

(b) BOB 变成了 WOW，而 WOW 变成了 BOB。这个结果意味着交换的部件 A 是重要的，并且是一个固定的红 X。于是部件搜索结束。

(c) BOB 部分地降级到一种 WOW，但是还没有完全降低到那种程度，WOW 部分地朝 BOB 改善，但还没有完全达到。这种结果意味着交换的部件 A 或许是重要的，但还不是整个问题的症结所在，还有其他一些部件会引起部分改变，这需要在随后的交换中予以发现。这种部分的改变也表明 A 与其他部件间的相互作用。

为了确定这种部分改变的程度是重要的还是不重要的，必须计算出判断极限（下一章将予以叙述）。任何交换过程，如果引起 BOB 或 WOW 有一点处于判断极限之外，就意味着部件 A 肯定是重要的，与其在一起的还有一个以上其他有待发现的部件。

(8) 对于第 7 步所产生的 3 个可能的结果中的每一个，将原先来自 BOB 的部件 A 归还 BOB，把原先来自 WOW 的部件 A 归还 WOW，以保证原始的 BOB 和 WOW 的绿 Y 值被复原。如果没有得到这样的结果，那么拆卸/重装过程就值得怀疑，要进行调查。

(9) 对于随后最有可能的子部件或部件 B，重复第 6 步，7 (a) 或 7 (c) 以及第 8 步；然后是 C 再后是 D，依次类推。

(10) 一旦在 7 (c) 和第 9 步有两个或多个部件被识别为是重要的，就要同时把它们在 BOB 和 WOW 之间交换，直到反向效果出现，即 BOB 变成 WOW，或者相反。那么，现在红 X 就是这些重要部件及其相互作用的组合。

(11) 阶段 3。通过对所有未试验过的和不重要的子部件/部件（如 WOW），相对所有重要的子部件/部件（如 BOB）来检测，以便完成试验，其结果应当接近第 1 阶段的 BOB；然后反过来将



BOB 当作不重要的部件, 而把 WOW 作为重要的部件进行检测, 其结果应当接近第 1 阶段的 WOW。

(12) 阶段 4。最后, 利用在第 1 和 2 阶段得到的所有数据进行完全析因分析, 以便量化和确定主效应和相互影响效应的方向和幅度。阶段 4 只是个计算, 不是新的试验。

案例研究 1: 计时器

某电子工厂生产一种计时器, 用户要求这种计时器能在 -40°C 的条件下可靠地工作。但是废品率是 2 位数, 问题拖了几个月也没解决。最糟糕的部件甚至在 0°C 时就不再工作了。²

计时器的组成部分有一个线圈和一个屏蔽 (用于集中电荷, 定时发出脉冲)。脉冲触发一个线圈的销, 销又引起一个边臂或电铃曲柄去触发计数器, 计数器步进一位。计数器连到一个含有数字轮的数字轮轴上, 数字轮由惰轮分隔开来, 惰轮在惰轮轴上转动。惰轮和数字轮都连在主框架上, 主框架由白色的硬塑料制成。脉冲由电路板提供。

第 1 阶段: 拆卸与重组装

选出两个部件, 一个是很好的 (在 -40°C), 一个是很差的 (在 0°C), 用于部件搜索。表 10-2 显示了第 1 阶段拆卸与重组装的结果。

表 10-2

第 1 阶段: 建立显著的重现差别

结果	高 (好) 的组件	低 (差) 的组件
初始	-40°C	0°C
第一次拆卸/重组装	-35°C	-5°C
第二次拆卸/重组装	-37°C	-7°C

资料来源: Motorola Automotive, Schaumburg, Ill.



显著的重现差别的检验

(a) BOB 的 3 个绿 Y 好于 WOW 的 3 个绿 Y, 检验通过。

(b) $D/\bar{d} \geq 1.25^{[3]}$, 其中:

$D=3$ 个高值和 3 个低值的中值⁻⁴ 的差

$d = (\text{高值组件的不重现性范围} + \text{低值组件的不重现性范围}) / 2$

这里 $D = -37^\circ - (-5^\circ) = -32^\circ$

$\bar{d} = [(-40^\circ + ^\circ 35^\circ) + (-7^\circ - 0^\circ)] / 2$

$= [(-5^\circ) + (-7^\circ)] / 2 = -6^\circ$

所以 $D/\bar{d} = -32^\circ : -6^\circ = 5.33 : 1$, 该值超过 1.25。

结论

(a)、(b) 的显著性是吻合的, 因此阶段 1 是成功的。这意味着红 X 和粉红 X 是在部件中, 而不在组装技术里。

表 10—3 显示了对计时器中的可疑部件和子部件的排序情况, 排列顺序以同类降序排出。当然, 这属于工程判断, 可以是正确的, 也可以是不正确的。由于通过了阶段 1, 我们知道一个或多个这样的部件/子部件肯定是红 X 或粉红 X 等。

表 10—3 部件排序 (以同类降序排列)

排列	部件	标签
1	线圈、销、轴	A
2	惰齿轮轴	B
3	数字轮轴	C
4	主框架	D
5	曲柄	E
6	惰齿轮	F
7	数字轮	G
8	电子电路板	H
9	其余部件	R

资料来源: Motorola Automotive, Schaumburg, III.

* 原文误为一 (减号)。-- 译者注。

曲线图

阶段 2：下一步是由好的计时器到差的计时器（或者相反地）交换每个部件/子部件。开始时是由部件 A——即线圈、销和轴先进行交换，因为好的继续保持是好的，差的依然是差的，所以 A 不是重要的。交换之后，要把部件分别归回原先好的和差的部件中去。（在图 10—4 中没有显示，为避免图表过于拥挤，对于其他交换的部件如 B，C，D 等都这样处理。）

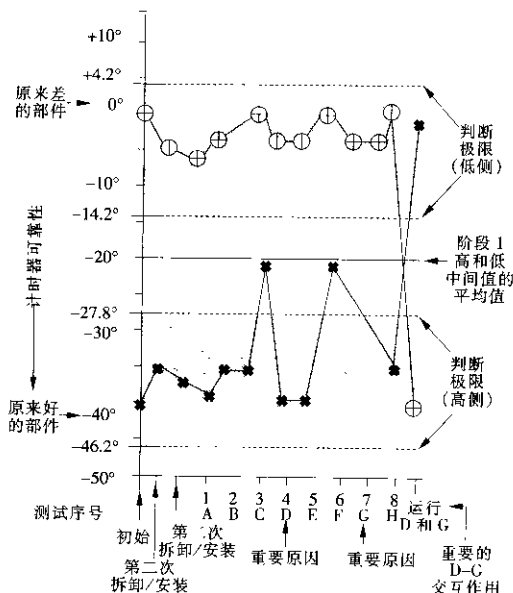


图 10—4 部件搜索案例：计时器试验——一个曲线图

资料来源：Motorola Automotive, Schaumburg, III.



由图 10-4 可见, 在交换过程中, 部件/子部件 B, C, E, F 和 H 没有产生由好变差或者由差变好的反向改变, 所以它们是不重要的; 但是 D, G 产生了部分的反向转变, 可以认为是重要的部件/子部件。无论何时, 在有两个以上的部件/子部件有部分的反向转变时, 要把它们同时由好至差或由差至好地交换。

阶段 3: 试验。此时, 重要的部件 D, G 审慎地保持在好 (高) 的水平, 其余的部件 R 保持在其差 (低) 的水平, 符号是 $D_H G_H R_L$ 。原先出现在阶段 1 的好的水平可以再现。接下来, 重要的部件 D, G 被审慎地保持在其差 (低) 的水平, 此时其余的部件 R 被保持于好 (高) 的水平上, 符号为 $D_L G_L R_H$ 。在阶段 1 出现的差水平可以再现。

判断极限

第 2 阶段判断极限的目的是, 确定一个部件 (被交换时) 及与之有关的相互作用的影响是重要的 (处于判断极限之外) 还是非重要的 (处于判断极限之内), 从而可以排除非重要的部件。

阶段 2 的判断极限来自阶段 1, 由下式给定:

$$\text{判断极限} = \text{中值} (H \text{ 或 } L) \pm t_{0.95} \bar{d} / \bar{d}_2$$

式中, 中值高 (h), 为阶段 1 出现的高值组件的中值 (见表 10-1); 中值低 (l), 为阶段 1 出现的低值组件的中值 (见表 10-1); t 为相应于一个 0.95 或 95% 置信度的值。

$\bar{d} / \bar{d}_2$ 是西格玛 (σ) 估计值。 ($\bar{d}$ —— 在检测显著的可重现差数时已予以定义, 它是高值和低值组件的不可重现性相加在一起时的平均范围。)

对于好、差部件的 3 次拆卸/重组装, 对每个部件有 2 个自由度或共有 4 个自由度; 在 t 分布表中^[16], 自由度为 4, 截尾数为 2, 置信度为 95% 时, t 等于 2.776; $\bar{d}_2$ 是一个统计常数, 等于 1.81。

所以, 判断极限 (高) = 中值 (高) $\pm 2.776 \bar{d} / 1.81$,

判断极限 (低) = 中值 (低) $\pm 2.776 \bar{d} / 1.81$ 。



在计时器的案例中：中值(高)=-37°；中值(低)=-5°； $\bar{d}=-6^\circ$ 。

所以，判断极限（高）=-37°±2.776×(-6/1.81)
=-46.2°和-27.8°；

判断极限（低）=-5°±2.776×(-6/1.81)
=+4.2°和-14.2°。

图 10-4 所示的不仅是交换部件 A~H 的效果，也包括这些部件高水平的判断极限，也就是-46.2°和-27.8°，以及这些部件的低水平判断极限，即+4.2°和-14.2°。

●它显示部件 A, B, C, E, F 和 H 都还处于高或低判断极限内，因而它们是不重要的；

●但是部件 D, G 超出了高水平的判断极限，但不是低判断极限，所以 D, G 是重要的。（即使一个部件超出了一个判断极限，例如高极限，而没有超出另一个，就可以认为它是重要的，同时，其交互影响的效应也被认为是重要的。）

表 10-4 综合了计时器案例研究的阶段 1、阶段 2 和阶段 3，它包括判断极限（来自阶段 1）、分析和结论。

表 10-4 部件搜索：计时器案例研究

试验号	接通的部件	高组件	结果	高判断极限	低组件	结果	低判断极限	分析
初始		所有零件，高	-40°		所有零件，低	0°		
第一次拆卸/重组		所有零件，高	-35°		所有零件，低	-5°		
第二次拆卸/重组		所有零件，高	-37°		所有零件，低	-7°		
1	A	A _L R _H	-40°	-27.8°和-46.2°	A _H R _L	-5°	-14.2°和+4.2°	A 不重要
2	B	B _L R _H	-35°	-27.8°和-46.2°	B _H R _L	0°	-14.2°和+4.2°	B 不重要

试验号	接通的 部件	高组件	结果	高判断 极限	低组件	结果	低判断 极限	分析
3	C	C _L R _H	-35°	-27.8°和 -46.2°	C _H R _L	-5°	-14.2°和 +4.2°	C不 重要
4	D	D _L R _H	-20°	-27.8°和 -46.2°	D _H R _L	-5°	-14.2°和 +4.3°	D重要
5	E	E _L R _H	-40°	-27.8°和 -46.2°	E _H R _L	0°	-14.2°和 +4.3°	E不 重要
6	F	F _L R _H	-40°	-27.8°和 -46.2°	F _H R _L	-5°	-14.2°和 +4.2°	F不 重要
7	G	G _L R _H	-20°	-27.8°和 -46.2°	G _H R _L	-5°	-14.2°和 +4.2°	G重要
8	H	H _L R _H	-35°	-27.8°和 -46.2°	H _H R _L	0°	-14.2°和 +4.2°	H不 重要
试验	R	D _H G _H R _L	-40°	-27.8°和 -46.2°	D _L G _L R _H	9°	-14.2°和 +4.2°	R不 重要

* 判断极限 = 中值 $\pm d/1.81$

高组件中值 = -37°，高组件中值 - 5°

$d = (\text{高组件范围} + \text{低组件范围}) / 2 = (7 + 5) / 2 = 6$

判断极限 (高) = -37° $\pm (2.776 \times 6 / 1.81) = -27.8^\circ$ 和 -46.2°

判断极限 (低) = -5° $\pm (2.776 \times 6 / 1.81) = +4.2^\circ$ 和 $+14.2^\circ$

结论:

(1) 部件 A, B, C, E, F, H 在高和低控制极限内, 它们是不重要的。

(2) 部件 D, G 超出高控制极限, 所以它们是重要的。

(3) 试验确认: 部件 D 与 G 的组合超出高和低控制极限, 所以 D, G 以及它们之间的相互作用也是重要的。

资料来源: Motorola Automotive, Schaumburg, Ill.

评估阶段 2 的部件交换重要性的被选经验

许多实际工作者会发现, 与确定这些判断极限相关的统计计算是冗长繁杂的。笔者开发了一种实验法, 用来评估阶段 2 的部件交换是否重要。这种方法不如判断极限精确, 但也可以作为经验加以应用。

(1) 选取阶段 1 的高中值和低中值, 在计时器案例研究中它们分别是 -37° 和 5°。

(2) 计算两者的平均值等于 -21° 。

3. 在第2阶段的部件交换中，如果一个部件交叉平均值或很接近交叉平均值($\pm 10\%$)，那么该部件就是重要的。如果不是这样，它就是不重要的。在本案例的阶段2中，A, B, C, E, F, H从不接近 -21° 的平均值。但当由高至低交换时，D和G在 20° 时接近平均值 -21° ($\pm 5\%$)，因此是重要的，其交互影响效应也如此。在试验时，从高至低或由低至高，D, G进行交换都超出平均值，那么D, G肯定都是重要的；同时，由于有完全的反向转变发生，所以它们的交互影响效应也是重要的。

主效应与交互影响效应

在讨论第4阶段(即析因分析)之前，我们要依次阐明一下主效应和交互影响效应。可以举一个简单的例子，一个人只喝酒可能不感到有任何副作用，他只吸毒可能也没有任何反应，但当他同时喝酒并吸毒时就会致死。这就是交互影响的效应，即只有两种项目综合起来才发生的结果。

另一个例子是，一个“发疯”的科学家用一瓶氢气和一瓶氧气进行试验，他点起火柴，分别放在两个瓶子处，没有什么事情发生。后来他把两种气体按一定的比例混合，再点火柴时，他马上被掀翻在地，到另一个世界去了。这就是一个交互影响的效应。

在人类相处的世界里，有许多交互影响的效应。通常，每个人各怀有有限的一技之长，若大家合在一起工作，常常会得到不寻常的结果。在体育世界，一个激励起来的球队往往会战胜一个未能很好协调的超级球队，这是一种交互影响——协作效应。在1998年世界杯足球赛的最后决赛中，法国队被认为处于下风，其对手巴西队，由许多耀眼的明星组成，包括超级巨星罗纳尔多，但是法国队以令人难以置信的3:0战胜了巴西。这就是团队的交互影响，团队的协作，团队的共同奋斗。

至于一个产品，我们可以量化其交互影响。试取2个组件A,



B，它们分别具有 3 个和 7 个单位的变量。当它们结合到一起时，正常时要有 10 个变量，此时 A，B 被认为是自主变量，分别单独地影响结果 10（被称作主效应）。但是如果 A，B 结合到一起产生的变量数为 16 时，这就是交互影响，通常是一种加法影响，此时总数大于其和数。化学反应里的催化剂也是加法交互影响。另一方面，如果 A，B 结合后，变量数变为 2，这就是减法交互影响，此时总数小于其和数。在政治和实业界，我们都可以看到许多加法和减法式的交互影响的实例。在一些罕见的例子里，民主党的总统和共和派的国会紧密协作，就是加法交互影响，国家繁荣；但经常是两者互相掣肘，交互影响是减法式的，国家就变糟了。

交互影响可以用图解法表示，在图 10—5A 中，我们设定输出（绿 Y）从 5 上升至 15 个单位，此时，输入变量 A 是从低变高。为了方便，在本书以后的代数计算中，低用负号（-）表示，高用正号（+）表示。上述例子是在一个输入这样的条件下：即另一个变量 B 保持在恒定的低（负）值。那么如果 B 现在由低值（-）变为高值（+），于是绿 Y 就会又增加 5 个单位。不管 B 是（-）或（+），A 对绿 Y 的影响恒定为 10 个单位；不管 A 是（-）或（+），B 对绿 Y 的影响恒定为 5 个单位。换言之，A 对 B 是独立的，B 对 A 是独立的，也就是 AB 的交互影响等于零。图中可见：两条线总是平行的，交互影响为零。

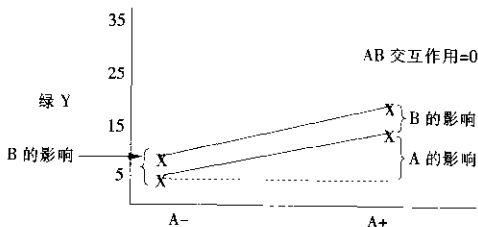


图 10—5A 主效应：A 和 B 无交互影响



在图 10—5B 中，当变量 A 由（-）变为（+）时，绿 Y 由 5 变为 15 个单位，和以前 B 为（-）时一样。但是，当 B 由（-）变为（+）时，如果 A 是（-），绿 Y 仅仅增加至 8；如果 A 是（+），它增大到 35 个单位。即在 A 是（-）时，B 对绿 Y 的影响仅是 3 个单位；而如果 A 是（+），则 B 对绿 Y 的影响是 20 个单位。此时，AB 的交互影响不为零。图中可看到：两条线不平行；交互影响越大，不平行度就越大。

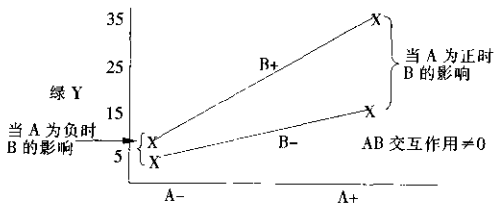


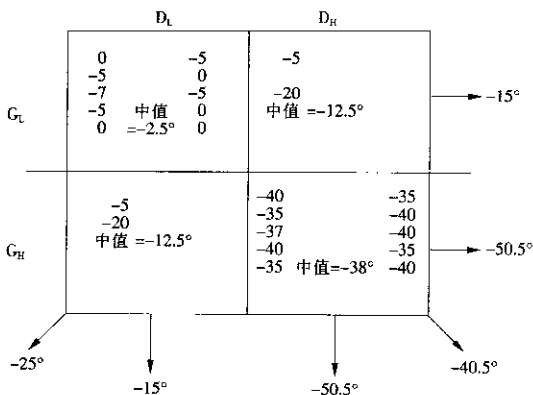
图 10—5B 主效应：A 和 B 有交互影响

它是这样的一种交互影响：当与主影响没有隔离时，使得分析因方法（可以是经典 DOE 或田口 DOE）产生完全错误的结果。

阶段 4：析因分析

析因分析^[5]不是一种新的试验方法，它从前 3 个阶段（第 1，2，3）采集数据，以便量化在部件搜索的第 2、3 阶段已经鉴别出的重要的主效应和交互影响的方向和幅度。

最简单的析因分析是一个 2×2 因子，也就是只有 2 个水平的 2 个因子。图 10—6 显示出计时器案例研究中的 2 个因子（D 和 G）在第 2 和第 3 阶段中已经判明是重要的。对于因子 D，它有 2 个水平 D_L 和 D_H ；对于因子 G 也有 2 个水平 G_L 和 G_H 。一个 2×2 因子有时指定为 2^2 ，其中底数 2 表示 2 个水平，而幂 2 表示变量的数目。因此有 4 种组合或 4 个单元（见图 10—6），需要予以填充。



主影响 $D = -50.5^\circ - (-15^\circ)/2 = -35^\circ$
 主影响 $G = -50.5^\circ - (-15^\circ)/2 = -35^\circ$
 DG 交互影响 $= -40.5^\circ/2 - (-25^\circ) = 9.5^\circ/2$

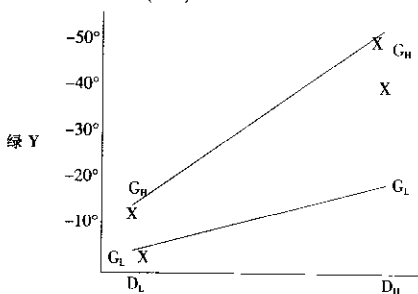


图 10—6 析因分析：计时器

资料来源：Motorola Automotive, Schaumburg, IL.

* 原书误为 15.5°。——译者注

在 $D_L G_L$ 单元 (双低), 有着总数为 10 个的绿 Y, 3 个来自阶段 1, 6 个来自低值组件阶段 2, 1 个来自阶段 3 (此时 D 和 G 是双低)。参照表 10—4 可以看出: 读数为 0° 的都是来自初始“低”的组件, -5° 和 -7° 是来自可重现的低值组件, 上面的 3 个读数都来自阶段 1。在阶段 2, 当 D 和 G 是双低时, 也有 6 种组合出现。这些读数是 -5° 、 0° 、 -5° 、 0° 、 -5° 和 0° 。在阶段 3 (即试验), 当 D 和 G 是双低时, 在 0° 处有多于一个的组合。

同样, 在阶段 2 时 (G 为低值, D 为高值), $D_H G_L$ 有 2 个读数, 即 -5° 和 -20° 。对于 $D_L G_H$ 单元在阶段 2 (D 为低值, G 为高值) 也有 2 个读数, 即 -5° 和 -20° 。最后, 在 $D_H G_H$ 单元, 有 10 个读数, 3 个来自第 1 阶段, 6 个来自高组件阶段 2, 1 个来自试验阶段 3, 此时 D, H 皆为高值, 它们的数值分别是: -40° , -35° , -37° , -40° , -35° , -35° , -40° , -40° , -35° 和 -40° 。

接下来, 确定每个单元中的中值, 如图 10—6 所示。在图 10—6 的因子矩阵中, 如果加上 D_L 列的中值, 结果是 D_L 单独影响绿 Y, 因为 G_L 和 G_H 平衡, 从效果上说是互相抵消。类似的是, 当加上 D_H 列的中值时, D_H 将单独影响绿 Y 值。其次, D_H 和 D_L 之间的差数被称为因子 D 的主影响。同样, 将 G_L 和 G_H 行的中值相加, 然后再从 G_H 减去 G_L , 可以得出因子 G 的主影响。为了确定 D 和 G 的交互影响, 把对角中值 $D_L G_H$ 和 $D_H G_L$ 相加, 再从 $D_L G_L$ 和 $D_H G_H$ 对角中值里减去。交互影响效应的曲线图显示在图 10—6 中。其中 $D_L G_L$ 单元读数 (中值) 是 -2.5° ; $D_H G_L$ 单元读数 (中值) 是 -12.5° ; $D_L G_H$ 单元读数 (中值) 是 12.5° ; $D_H G_H$ 单元读数 (中值) 是 -38° ; 当绘出 4 个点之后, 通过 G_L , G_H 连线, 会看到一个不平行的影响, 这表明在因子 D 和 G 之间有一种中等程度的交互影响。

“与部件对话”

运用部件搜索研究计时器的可靠性的工程师在其报告的最后写道：“解决此问题竟用去了我们 18 个月的时间。

- 我们与供料商对话；
- 我们与工程师和设计师对话；
- 我们与工程管理者对话；
- 但是我们从没有与部件对话！

利用部件搜索技术，仅仅 3 天我们就解决了问题……”

计时器案例研究的最终解决方案

线索生成技术——部件搜索

主框架和数字轮导致性能下降。

工程分析

- 制作 60×实物大小的模型。
- 为了解决问题，将第 1 数字轮与主框架隔离开来。
- 在温度临界点进行测量。

结果

●主框架有高达 0.002 英寸的收缩，这使得数字轮和惰轮轴过于接近。

- 第 1 数字轮偏离中心 0.005 英寸。
- 当收缩率与偏心率相吻合时，计数器卡住了。

可选用的解决方案

- 重新设计主框架（要耗费 50 000 美元）——此方案未被采纳。
- 改变数字轮的技术规格要求和公差（低成本）——采用了这种方案。

► 让部件而不是让工程师来决定正确的技术规格要求

在传统的解决问题的方法中，有缺陷的部件被送回到前方的检验者或供应商那里，以便反复检验它们是否符合技术规格。结果经常是千篇一律的回答：部件符合技术规格要求。调查结束！烦恼开始了！

部件搜索（及将在下一章详述的与之相关的成对比较技术）是一种比通过工程判断来建立正确的技术规格的好得多的办法。作者的看法是（同时得到与我合作多年的工程师们的支持）：大多数技术规格都是随意制定出来，且脱离实际的，结果使得技术规格太过严格或太过松散，总之，往往是错误的。

在部件搜索技术中，部件告诉我们哪个是好的部件，哪个是差的部件。应该根据好的部件或好的部件组，来确定正确的技术规格。作者看到一个又一个的实例中差的部件符合技术规格的要求，而好的部件却不符合技术规格的要求。那么，我们该怎么办？还有什么要问的，当然要修改任意定制的技术规格要求，以便肯定这些好部件的读数。

► 案例研究 2：点不着的燃烧器

一家中等规模的跨国公司生产烧油的锅炉，产品中有将近 3% 的燃烧器不起火。检查各种部件后，发现其功能良好，但燃烧器就是不起火。

于是公司开展了部件搜索，此时仅存在着一个“通过/不通过”的绿 Y，换言之就是点得着与点不着的燃烧器。

阶段 1

将 BOB 和 WOW 部件拆卸/重组装，其各自表示的绿 Y 没有变化，如图 10—7 所示。

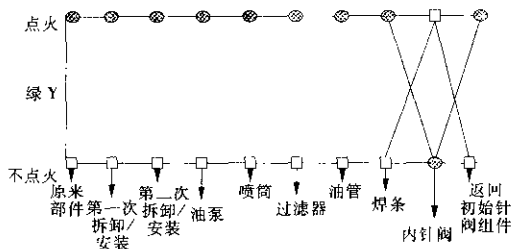


图 10—7 部件搜索案例研究：燃烧器

资料来源：Thermax, Pune, India. III.

阶段 2

可疑部件以递减的顺序排列如下：

- 油泵；
- 喷嘴；
- 过滤器；
- 油管；
- 电极；
- 内针阀组件。

当对上述部件进行交换时，经常会发生这样的情况：最不被怀疑的内针阀组件结果却成了红 X。而另一些部件对问题的影响不大，因此也就没有交互影响，所以就不需要进行第 3、4 阶段了。

随后，对内针阀组件进行部件搜索（这里未给出），其部件按递减顺序排列如下：

- 弹簧；
- 针阀；
- 喷嘴支架；
- 底座。

幸运的是，第一次交换就发现弹簧有 BOB 和 WOW 的反向改变，所以无需对其他部件再进行交换。弹簧被送回供应商那里进行调查。（应当在本公司对比好的弹簧和坏的弹簧进行多次检验，如下章所述，因为供应商们通常不知道如何应用 DOE 技术解决长期存在的问题。仅有专业知识是不够的。）

小组练习 1：某汽车制造公司的失败教训

几年以前，某汽车制造公司所生产的新车遇到了一个特殊问题，这个问题是在某些特定条件组合下出现的。典型的例子是一位年轻的太太的经历。这位太太开着新车到冰激凌店去买蛋卷香草冰激凌。当她拿着冰激凌回到车上，想发动汽车时，点不着火；等了好一会儿，才好歹打着火。

几天之后，她又驾车去那家店买蛋卷香草冰激凌，然后拿着冰激凌上了车，可汽车又发动不起来了，她只好又等了一会才能把车开走。于是，她坚信蛋卷香草冰激凌就是症结所在！所以当她第三次光顾冰激凌店时，她更换了口味，要了一筒果仁冰激凌。当她拿着果仁冰激凌，回到汽车上后，第一次发动就把汽车开走了！她很自信地把这个发现告诉了丈夫：“我把问题解决了，一定是什么东西怕香草的气味！”丈夫对此感到可笑，可妻子不服气地说：“如果你觉得你聪明，你就去试试看！”

果真如此，每次丈夫拿着香草冰激凌蛋卷上车时，汽车就是无法启动，可如果换上果仁冰激凌，汽车马上就能开走。这真怪了。夫妻俩来到卖汽车的商人那里投诉。商人听后心里嘀咕：这城里真是住满了怪人！可是商人自己一试，也愣了！此时，汽车公司已接到了许多

类似的投诉，而美国交通部则正在考虑大规模的召回问题。

但汽车制造公司已开展 DOE 工作以期解决这一问题。简单地说，为什么香草蛋卷冰激凌不能启动汽车，而果仁冰激凌却能，其原因不在于两者的味道，而在于商贩制作这两种冰激凌所花时间的长短。工业界都知道所谓“热不稳定”问题。当新车跑路尚未热起来时，假如汽车点火器熄灭的时间很短，那么就没有充分的力矩重新启动；但当汽车完全冷下来后，就可以重新启动了。香草味的冰激凌很容易制备，商贩仅需 2 分钟就可以给你，这对“热不稳定”问题就显得时间太短了。而果仁冰激凌是一种不寻常口味的冰激凌，商贩必须到店的后面，把桶拿过来，等它稍微溶化一些再挖取。这需要 5~10 分钟，而这段时间足以使发动机冷却下来，并重新启动。

图 10—8 是汽车制造公司的实际部件搜索图示。第 1 步是把有属性特征的绿 Y——启动/不启动——转化为可量化、可测量的绿 Y。比较“好”车和“差”车（有“热不稳定”问题的车）上散热器管道处的水温，但是相关性不很明朗。可是油温却变成了用以监测发动机温度的较为可重现的绿 Y。油温为 240°F~250°F 的汽车没有“热不稳定”问题，而油温在 190°F~200°F 的汽车却总是出现“热不稳定”问题。

问题

查看一下图 10—8 并回答如下问题：

- (1) 绿 Y 被正确量化和测量了吗？
- (2) 正确实施了部件搜索的第 1 阶段吗？
- (3) 阶段 2 有 12 次部件交换，这些交换进行得正确吗？

(4) 在阶段 2，什么是重要部件，什么是不重要部件？根据“评估阶段 2 的部件交换重要性的被选经验”，利用好汽车的油温和有“热不稳定”问题的汽车油温的平均值来进行评价。

- (5) 进行阶段 3 的检验了吗？

- (6) 应用析因分析确定交互影响的幅度和方向了吗？



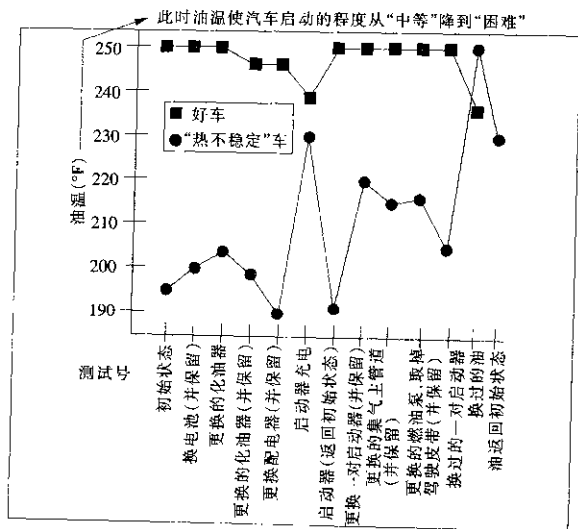


图 10—8 部件搜索小组练习 1: “热不稳定”的汽车

资料来源: General Motors Corporation.

答案

概述。在起始之际我们必须声明的是: 此部件搜索试验没有根据“教课书”的路径图去做。在谢恩 DOE 技术发展的早期曾做过这种试验。这种方法是暂时性的、试验性的。

答案 1: 基本上可以肯定地回答“是的”。汽车公司试图通过水温来区分好汽车和有“热不稳定”问题的汽车, 但两者相关性不是很强。而油温则被证明是好汽车之间和“热不稳定”汽车之间始终如一差异——温差近 60°F。

答案 2: 不。好汽车与“热不稳定汽车”都没有被拆卸和重组装,以肯定问题是出现在部件上,而不是出现在组装过程中。正如所发生的情况那样,试验是幸运的,问题是出在一个或几个部件上,而不是出在组装过程里。

答案 3: 不。当电池、化油器、主集气管道、燃料泵、驾驶皮带进行交换之后,它们没有返回交换开始时的原汽车中。为了确定是否存在交互影响(原始的绿 Y 是否会再次出现),这一步是必须的。试验又一次是幸运的,这些部件在交换时,似乎没有产生显著的改变,而且也没有产生可疑的重要交互影响。另一方面,当交换启动器时,“热不稳定”汽车有一个显著的改善,启动器被放回到来原来的汽车上,原始情况又再现了。其他部件也应当被返回到原来的汽车上,不管其交换结果如何。

另一个错误是把两辆汽车都换上了新的启动器。这引入了新的变量,会导致错误的结论。在这种情况下,由于装上了新启动器,“热不稳定”汽车呈现了某些改善。最终,交换发动机油才使得“热不稳定”汽车有了戏剧性的改善,而好汽车只有轻微的变差。但是当把油返回到原来的汽车上,“热不稳定”汽车出现了轻微的变差,不过在好汽车上进行的这一最后的交换结果没有记录下来。

答案 4: 发电机油是最重要的部件——一个准红 X,后面跟着一个粉红 X——启动器。在油和启动器之间还显示出有强烈的交互影响。原始的油温 250°F 和 195°F 之间的中线线是在 222.5°F ,通过“常规的经验法则”判断可知,再也没有其他任何部件是重要的了。

答案 5: 没有。这是阶段 3,其目的是验证重要变量确实重要,非重要变量确实不重要。此时发动机油和启动器应同时被交换,其他的部件应保持反的方向。即试验 1 为 $O_A S_A R_B$, 试验 2 为 $O_B S_B R_C$; 其中 O_A 和 O_B 分别是好车和差车的油, S_A 和 S_B 分别是好车和差车的启动器, R_C 和 R_B 分别是好车和差车的其余部件。

答案 6: 没有。仅仅当有迹象表明阶段 2 有某种交互影响时,析因分析(阶段 4)才是必须的。

进行进一步的 DOE 试验

即使汽车制造公司没有按照教科书方法执行部件搜索，也要利用全析因法进行进一步的试验设计（见第 14 章）。它们显示出如下的结果：

- 油方面的差别：“热不稳定”车是 10W 油，而好车是 20W 油，说明在可启动性方面有 55°F 的差异。
- 低限启动器和高限启动器之间的差别改善了后者的启动性约 26°F。
- 把压缩比由 1 : 9 提高至 10 : 5，仅使可启动性改善约 4°F。
- 把点火角由 6° 提高至 13°，仅改善可启动性 1°F。

分析

较稀薄的 10W 油失去了润滑性，并从啮合表面的壁上排出。如果发动机在热状态下关闭并仅冷却几分钟，“热不稳定”性就最为显著。发动机在运行状态时，由于啮合表面是处于油槽中，稀薄的油就变得不那么重要了。而且，新车在经过磨合之后，比如说跑了 1 000 英里，由于启动器在长期运用之后所要求的力矩降低，因此“热不稳定”问题就消失了。

结论

“热不稳定”现象仅在某些变量组合的条件下才会发生：一辆新车、稀薄的油、发动机配合较紧、启动器的力矩较低，但是试验设计探明了根本原因。

小组练习 2：振荡器的时间延迟

某电子仪器在电源交换信号和部件实际开始运行之间有长时间的延迟。一个“好”仪器的延迟时间约为 16 毫秒 (ms)，而“差”

仪器的延迟时间是 35 毫秒。振荡电路被确认为时间延迟的根源。阶段 1 和阶段 2 (针对 10 个可怀疑的部件) 如表 10—5 所示。R 表示组件中不需要被交换的部件。

表 10—5 部件搜索小组练习 2: 振荡器的时间延迟

阶段 1	高值组件 (H)		低值组件 (L)	
	配置	结果 (ms)	配置	结果 (ms)
初始	所有部件, 高	13	所有部件, 低	34
第一次拆卸与重组	所有部件, 高	16	所有部件, 低	38
第二次拆卸与重组	所有部件, 高	15	所有部件, 低	35

阶段 2 试验	部件	配置	结果 (ms)	控制极限 高		控制极限 低	
				配置	结果 (ms)	配置	结果 (ms)
1	A: 晶体	A _L R _H	16	A _H R _L	19		
2	B: 微处理器	B _L R _H	16	B _H R _L	35		
3	C: 晶体管	C _L R _H	14	C _H R _L	33		
4	D: 电容 C ₂	D _L R _H	15	D _H R _L	37		
5	E: 电容 C ₁	E _L R _H	16	E _H R _L	16		
阶段 3 试验	A 和 E	A _L E _L R _H	32	A _H E _H R _L	17		

资料来源: Motorola Inc., Arcade, N. Y.

问题

(1) 阶段 1。确定在高的 (H) 或好的组件与低的 (L) 或差的组件之间, 是否有显著的差别。

(2) 计算 D/d 比率。阶段 1 证明了什么?

(3) 将表 10—5 的结果制成图, 计算高、低判断极限。什么是不重要的部件? 你会得出什么结论?



(4) 阶段 3 的求交运算成功吗?

(5) 构建析因分析。确定重要部件的主效应和交互影响效应。绘制表现这些交互影响范围的曲线图。

答案

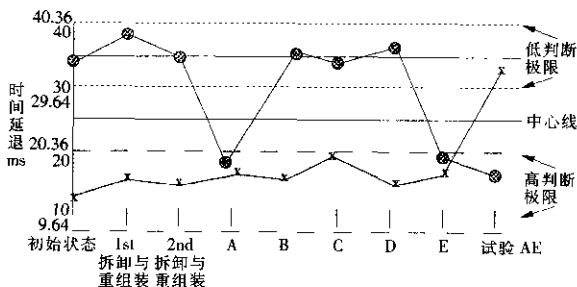
(1) 在高读数 13、16 和 15 与低读数 34、38 和 35 范围, 呈现出相当连贯的一致性。这一点将被问题 2 有关 D/d 比率的答案所证实。

(2) 3 个高值组件的中值是 15, 3 个低值组件的中值是 35, 于是 $D-35-15=20$ 。

高值组件的范围是 3, 低值组件的范围是 4, 因而平均范围 $\bar{d}$ (或缺少重现性) 等于 $(3+4)/2=3.5$ 。于是 $D:\bar{d}$ 比率是 $20:3.5=5.7:1$, 该值超过了比率为 $1.25:1$ 的最低要求。

阶段 1 证明, 延迟问题的原因不在于组装过程, 而在于部件本身。

(3) 图 10-9 是表 10-5 的曲线图。



图例: X=除指定部件, 全都是高值
O=除指定部件, 全都是低值

图 10-9 高低判断极限曲线图

判断极限 (高值) 是:

$$\begin{aligned}\text{中值(高)} \pm 2.776\bar{d}/1.81 &= 15 \pm (2.776 \times 3.5/1.81) \\ &= 15 \pm 5.36 = 20.36 \text{ 和 } 9.64\end{aligned}$$

判断极限(低值)是:

$$\begin{aligned}\text{中值(低)} \pm 2.776\bar{d}/1.81 &= 35 \pm (2.776 \times 3.5/1.81) \\ &= 35 \pm 5.36 = 40.36 \text{ 和 } 29.64\end{aligned}$$

部件 B, C 和 D 是不重要的, 它们既不造成任何初始时为好的(或差的)读数的反向改变, 也不超出高的或低的控制极限。

可以得出下述结论: 第一, 当从高至低值进行交换时, 在阶段 2 被交换的部件中没有产生任何显著的改变, 因为 A, B, C, D 和 E 没有超出它们的高判断极限。第二, 另一方面, 在第 2 阶段进行交换时, 部件 A, E 超出了它们的低判断极限。第三, 这样的单向改变是在一组控制极限里, 但是超出了另一组控制极限, 这肯定是一种交互影响效应的表现。

(4) 是的。当部件 A、E 保持高值, 而其余(R)部件是低值时, 结果是高值(超出了高判断极限); 反过来则相反。

(5) 表格 10—6 显示的是部件 A, E 的两因子的析因分析。

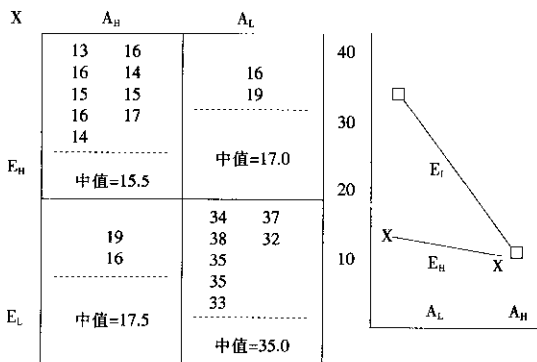


表 10—6 两因子析因分析

结论

对时间延迟问题的进一步工程调查发现：晶体的串联阻抗处于低端，而电容漏电处于高端。当上述情况出现时，振荡回路将给微处理器加载，从而引起较长时间的延迟。通过与电容供应商协调，并利用DOE技术降低电容漏电问题，可解决时间延迟问题，而且未给公司增加额外的成本。

我们可以粗略地称E为一个红X，称A为粉红X，而称AE的相互作用为一个浅粉红X。在红X、粉红X和浅粉红X之间缺乏很大的间隔时，这个结果是很不寻常的。

从表10—6可见：

主效应 $A = (17.0 + 35.0) / 2 - (15.5 + 17.5) / 2 = 26 - 16.5 = 9.5$ ，

主效应 $E = (17.5 + 35.0) / 2 - (15.5 + 17.0) / 2 = 26.25 - 16.25 = 10$ ，

AE交互影响效应 $= (15.5 + 35.0) / 2 - (17.0 + 17.5) / 2 = 25.25 - 17.25 = 8$ 。

当拆卸/重组装是不可重现性时，部件搜索的实践

迄今为止，在我们所涉及的案例中，完成拆卸与重组装之后，好部件照旧是好部件，差部件照旧是差部件。但如果情况不是这样，又该怎么办呢？那么，就应一步接一步地进行渐进式的拆卸与重组装，以便确定具体在哪一步需要分析和修正。下面用案例来说明这种技术。

案例研究3：电冰箱门隙

某家电制造商遇到了如下问题：关闭电冰箱箱门时，箱体和箱门之间出现缝隙。这不是一个功能性的问题，而是遭到顾客指责的

外观问题。此问题已存在数月，需要大量返工。

部件搜索显然是 DOE 首要采用的技术。小组选择了一个具有最小门缝的 BOB 和一个具有最受指责的门缝的 WOW。但是，在对每个部件进行了拆卸与重组装之后，BOB 具有变成了 WOW，而 WOW 变成了 BOB。这在漫长的部件搜索技术历史中，是一个极不寻常的现象。于是 DOE 小组依次列出了拆卸/重组装的各个步骤：

- (1) 旋开并重新旋定托架；
- (2) 取掉并重新插入铰链和铰链销；
- (3) 取掉并重新插入箱门和箱体上的凸缘角铁；
- (4) 取下箱门并重新安装到铰链上。

小组发现，是在第 4 步——即把箱门重新安装到铰链上——缺少与前面步骤的连贯性。装配工人在生产时有其自己的独特装配方法。对这些方法的成对比较表明了每名装配工必须严格遵循的正确安装程序。后来，所有的 WOW 都变成了 BOB，没有再出现过次品。

小组练习 3：“行走”的机器

在另一条生产线上，某公司在进行加速寿命试验的过程中，由于振动，遇到了“机器行走”的问题。由于有“好”机器和“差”机器，所以部件搜索就是逻辑上要采用的 DOE 技术了。小组从某一大的产品里选定一个 BOB 和一个 WOW。但是在进行了拆卸/重组装之后，“行走”的数量是不可重现的。又对其他几个机器在拆卸/重组装之前，试着去重现行走的数量，但是读数完全是前后不一致的。在第 2 次或第 3 次试验时，“行走”的机器会停下来或者情况会反过来。（实际生产部门把“行走”的机器送到质量检测试验室，如果它们通过了——不走了——就把它们邮寄给用户。但是来自应用现场的对“行走机器”的投诉照例困扰着公司。）



问题

(1) 在进行部件搜索之前, 应该应用什么 DOE 技术?

(2) 是否要应用一种渐进式的拆卸/重组装技术来观察哪一个拆卸步骤对“行走”问题的影响最大?

答案

(1) 对于机器“行走”的方向和幅度, 绿 Y 已经得到了正确的量化。但是 DOE 调查的第 2 步没有跟上, 换言之, 就是测量系统要求精度要达到“行走”公差的 5 倍。测量系统是最不可靠的, 在指定的“地面”上, 无论是在生产中还是质量测量试验室里, “行走”的读数都前后不一致。使问题更复杂的是, 工厂“地面”上找不到“行走”的部件, 可是来自应用现场的对“行走”的投诉还是源源不断。

当这种不一致性出现在各试验站时, 第一件要做的工作就是建立一个具有一致性的试验平台, 同时制定一种新的试验方法: 每台机器的支脚上的力沿 X、Y 和 Z 轴都经过精确测定。应用新的测量方法之后, BOB 和 WOW 间的差别消失了。所有机器的“行走”都呈现了一致性, 可是都比竞争对手公司的同类产品差。这意味着有严重的设计问题, 必须通过变量搜索试验方法(见第 13 章)加以解决。

(2) 不用。当正确的试验方法已经揭示了没有部件不存在行走问题时, 由于不存在 BOB 部件, 所以没有必要进行拆卸/重组装这一步。

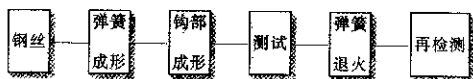
► 部件搜索技术在过程/机器/生产线中的应用

正如本章通篇所说的那样, 对于组件来说, 部件搜索是一种自然而强大的 DOE 工具。它也可以用来比较看似相同的两个过程、机器或生产线——一个是好的, 一个是差的。此方法也许不像应用

于组件时那么精确。例如，或许不会有几个 BOB 或 WOW，每种也许只有一个。还有，拆卸/重组也不是总能行得通的，但是部件搜索的通用原理确是适用的。

案例研究 4：两条生产线上弹簧扭矩的差别

两条同样的生产线在生产弹簧。工艺流程图如下：



仅有一个生产班次，在最后的再检测站上，报告生产线 1 在扭矩方面有着高废品率，而生产线 2 的废品率几乎为零。经理起先以为是两条线上的主管不同。生产线 1 的主管责备工人，工人怪罪于成形机器。如果机器能说话，它们或许会指责钢丝材料呢！

两条生产线之间的部件搜索发现了红 X。交换主管、操作者、成形机器，甚至交换了应力消除设备后发现，扭矩问题仍然存在于生产线 1 上。最后将两条线上的再检测用的夹具进行交换，此时生产线 2 变成了有问题的线，而生产线 1 变成了完美的生产线。人们发现生产线 1 的测试夹具未经维修和校准，造成力矩读数不准。

案例研究 5：操作者之间的差异

在激光调阻过程中，对于两台同样的机器，两名操作者却给出了完全不同的结果。第 1 台机器的废品率为零，而第 2 台的废品率却很高。对激光调阻机的参数进行的检测表明两台机器相同。没有调阻的电阻阻值也均匀。接下来交换两名操作者，结果两台调阻机的废品率为零。

进一步的调查集中于机器—操作者界面。令人惊讶的结果！第



2 名操作者在其他工作中有着优秀的质量记录，可是在第 2 台机器上却出了许多废品。这是因为她身材矮小，在进行调阻时与观察窗有一个视角，视角误差使她错过靶面。而第 1 台机器上的观察窗低 6 英寸，2 号操作者没有任何困难就可达到零废品率。解决的办法是给 2 号操作者一个高凳子，仍在 2 号机上工作，或让她转到 1 号机上工作。她选择了去 1 号机上工作。

部件搜索技术在管理工作中的应用

案例研究 5 展示了部件搜索技术是怎样用于区分两个操作者的。同样的原理可以用于在管理工作中区分 2 名售货员、2 名安装工、2 名服务员、2 名秘书、2 名银行出纳员和 2 名部门经理，只要他们中的好与差之间存在一个显著的绿 Y 差异。在很多此类情况下，我们发现，WOW 人员很少像 WOW 系统那样备受责难。

小组练习 4 (游戏): 找出红 X 球 (1)

有 80 个钢球。除了 1 个较重之外，其余的重量相同。给你 1 个双盘天平（无量度），请通过少于 4 次的称量找出红 X 球。

小组练习 5 (游戏): 找出红 X 球 (2)

有 12 个钢球，除有 1 个球比其它球或轻或重之外，其余的重量都相同。给你 1 个双盘天平（无量度），请你通过少于 3 次的称量找出红 X 球。

小组练习 4 的答案: 找出红 X 球 (1)

通常人们趋向于把 80 个球分成两组，每组 40 个；这将造成结

果至少是 5 次而不是 4 次称量。正确的方法是将其分成 27、27 和 26 的 3 组（思路是 $3^4=81$ ，这意味着分成 3 组，称量 4 次），如图 10--10 所示。

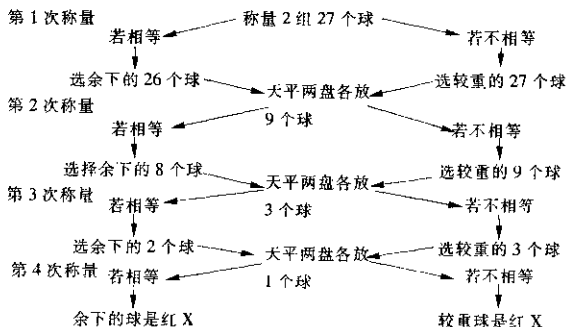


图 10—10 小组练习 4 的答案

小组练习 5 的答案：找出红 X 球 (2)

提示：将球 1~12 编号（见图 10-11）。

DOE 小组的部件搜索问题

正如对多变量技术那样，下列问题是提供给 DOE 小组在进行部件搜索研究时，作为记忆的引导员和准则的。在讨论多变量的章节中的前 10 个问题，涉及了确定绿 Y 和保证测量仪器的精度问题。同样的 10 个问题也可由 DOE 小组针对所有 DOE 技术进行提问（为节省时间和空间，此处不再重复）。



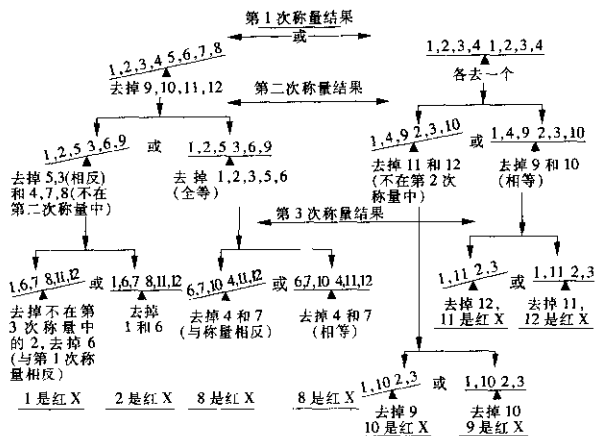


图 10—11 小组练习 5 的答案

注：天平向左侧倾斜，反之亦然。

概述

(1) 如果已知重要变量族是部件对部件（不是时间对时间或部件内的），那么我们回避的多变量与选择的部件搜索会不会是重要变量族？这种部件搜索的目的是研究在同一时间下运行的部件间的差别。

(2) 产品/组件是可以进行拆卸/重组装，而不会从根本上改变或损坏的吗？

(3) 在一天或一周内，你尝试过竭力选取最好或最差的部件了吗？（这些部件之间的差别越大，就越容易找出红 X。）



阶段 1

(4) 在初始的绿 Y 值之后，你是否将 BOB 和 WOW 拆卸/重组 2 次，以满足 2 个规则：第一，3 个 BOB 读数一定要比 3 个 WOW 读数重要；第二， $D:\bar{d}$ 比率必须大于 1.25。

(5) 如果没有满足上述两个规则，那么，是否在一步一步的拆卸和重组的过程中有一步呈现为一个可能的红 X？

阶段 2

(6) 你列出了一个以重要性递减的顺序排列需要交换的部件的清单了吗？

(7) 从 BOB 到 WOW，或者以相反的方向进行交换的部件，在每一次交换之后，它们都返回到原始的 BOB 或 WOW 中了吗？

(8) 如果从 BOB 到 WOW，或者以相反的方向交换，只产生了部分的改变，那么还对没有进行交换过的部件继续进行部件搜索了吗？

(9) 如果 1 个 BOB 变成了 WOW，但是 WOW 依然保持是 1 个 WOW（或相反的情况），造成一个交互影响，部件搜索应当继续进行吗？而且这表明在被交换的部件和未知的部件之间有一个交互影响吗？

(10) 部件搜索技术需要继续对一个或几个部件进行交换，直到出现 BOB 或 WOW 的完全反向的改变吗？

(11) 为了找出哪一个部件交换超出了其判断极限，计算过 BOB 和 WOW 的判断极限吗？

(12) 作为判断极限可选用的方案，阶段 1 的 BOB 和 WOW 之间的中值连线被用于确定部件的重要性了吗？



阶段 3

(13) 当所有的重要部件都保持在其高值水平，而不重要的部件保持在其低值水平（或者相反的情况下），进行过阶段 3 的试验吗？

阶段 4

(14) 用析因分析来量化主效应和交互影响效应了吗？

高层管理人员的问题

(1) 为什么把部件搜索选作 DOE 技术？

(2) 部件搜索技术遵从早期的 DOE 技术（例如多变量）吗？

(3) 要用多长的时间才能找到真正的最好的部件（BOB）和真正最差的部件（WOW），以开始部件搜索？

(4) BOB 和 WOW 之间的差距足够大吗？（差距不是足够大，是 DOE 小组中的共同缺点。）

(5) 仅发现了一个单一原因吗？一个单一的红 X？

(6) 或者是有几个重要原因？

(7) 进行过阶段 3 试验，以便确定并区分重要因子和不重要因子了吗？

(8) 如果有数个重要因子，那么，进行过析因分析以便量化交互影响了吗？

【注释】

[1] 在本章和后面的章节里，“好”和“差”的部件，不必是



“合格”和“废品”部件。它们两者都是可接受的，但在绿 Y 测量的概念里，它们必须是远远分隔的。“好”和“差”仅仅是简化的标签。

[2] 这个案例研究比一般的部件搜索要复杂一些，选择它是为了展示整个技术所要用到的所有部分。

[3] 根据 0.05 传统的 F 表，每个组件有 3 种构成。

[4] 可以用平均值，但中值（在中值的两侧，有相等的读数）更精确。

[5] 析因分析与部件搜索、变量搜索和完全析因法都是通用的技术。





第 11 章

成对比较：一种精巧而通用的工具

引言

在第 10 章我们说过：无论何时（或多或少是在同一时间内），生产出好的和差的部件时，就要回避多变量，而应用部件搜索技术。但是如果不符合部件搜索的先决条件之一——好的和差的部件都应能拆卸/重组装时，该怎么办？如果拆卸会损坏或会引起好的或差的部件的根本性改变时，又该怎么办？那么，部件搜索就不能解决问题。

幸运的是，“与部件对话”的第 4 个线索生成技术可以在这些场合大显身手。这就是成对比较技术。和前 3 项技术一样，这种



技术不中断生产过程，其方法基于12~16个部件数据的极为简单的布局——通常是6~8个好部件和6~8个差部件——按高低次序排列。而且当与部件相关的一个参数或质量特性被宣布是重要的时候，其置信度较高（90%以上）（参照图7—3，看一下这一步骤适合路径图中的哪一个位置）。

成对比较的通用性很强，以至于可以用于新产品和过程的设计、生产、现场使用、后勤服务、管理工作、农场、医院和学校——简言之，可以用于任何经济活动。当红X从系统、子系统和子部件的部件搜索滤出，不能再进行拆卸时，成对比较还是部件搜索的后续工作。

先决条件

- 性能（输出或绿Y）必须是可测量的，测量仪器的精度至少是规范公差或产品分布的5倍。

- 在一个大致恒定的时间框架内，每一次尝试都应当选取最好的部件（BOB）和最差的部件（WOW）。

- 如果质量特性或参数是一个属性，就应当利用在第7章讨论过的李克特度量尺度将其转化为一个变量（如果可能的话）。

方法：方案A

- （1）选择采样量。选取相对于被调查的绿Y尽可能远的6个或8个好的部件（BOB）以及同样数量的差的部件（WOW）。BOB与WOW相隔越远，越容易找出红X。

- （2）尽可能多地列出可以表达BOB和WOW的绿Y值差异的多个参数或质量特性。肯定有些原因造成好部件之所以为好，而差



部件之所以为差。所以，某些参数里的差别必定就在其中。（有些 DOE 试验人员在利用成对比较时，找出 4~5 个参数还没有找到差别时，就会失去耐心继而放弃。）差别应该是可视的、有尺寸的、外观方面的、机械的、电气的、化学的和冶金等的。测量时可以利用肉眼、尺子、X 射线、扫描电子显微镜、CAT 扫描、有限单元分析或破坏性试验（利用 MEOST）。

（3）选择 1 对部件——1 个好的和 1 个差的；如第 2 步所述，记下差别。

（4）再选择第 2 对部件——1 个好的和 1 个差的，记下第 2 对的差别。

（5）继续选择第 3 对、第 4 对、第 5 对、第 6 对（如果有必要，可以有第 7 对和第 8 对），重复上述搜索过程，直到出现 1 个或多个参数能够在同方向上显示出 1 个可重现的差别为止。

（6）通常在第 5 对或第 6 对的几个重要参数中，就可能出现前后一致的、可重现的差别，这将为找出变量的主要原因提供有力的线索。

方法：方案 B

虽然这种方案利用的是成组的比较（好部件和差部件的比较），而不是成对的比较，我们还是保留成对比较这个通用的名称。

- 方案步骤在第 1 步和第 2 步是同样的。但是选择的是 6 或 8 个单独的对，记下 6 或 8 个好部件及 6 或 8 个差部件的每个质量参数的读数，将读数按由小至大（或相反）的次序排列，而不管它们是好还是差的。

- 应用图基检验（下面将详述图基检验的步骤和实例）。

- 如果整个终结计数是 6 或大于 6 时，则该特定的质量参数在解释好与差的部件差别方面的重要性上，就具有 90% 以上的置信

度。

- 如果整个终结计数是 5 或小于 5 时，这样一个质量参数在解释好部件与差部件的差别方面的重要性上，就不具有充分的置信度。

图基检验

图基检验 (Tukey Test) 是以前著名的创始人——普林斯顿大学的约翰·图基 (John Tukey) 的名字来命名的，检验的目的是确定一个特定的质量参数是否重要，其重要性具有很高的置信度。它被应用于成对比较、完全析因法和 B vs. C 等 DOE 工作中，并作为一种非参数性质的比较技术而得以应用。此时，数据是按高低次序来组织的，而不像变量数据那样按绝对值来组织，也不按照完全不充分的非鉴别属性数据来组织。

图基检验的程序

- 不管好与差，把与指定的质量特性或参数相关的一组读数从高到低（或相反方向）排列起来。

- 为 12 个或 16 个排列好的原始数据读数中的每一个，标注“好的部件” (G) 或“差的部件” (B)。

- 当“全都是差的”第一次改变为“全都是好的”（或相反）时，划一条由这些读数的顶端开始的直线。这些是终结计数。

- 同样，当“全都是好的”第一次改变为“全都是差的”（或相反）时，划一条由这些读数的底端开始的直线。

- 将顶端和底端的终结计数相加以确定合计终结计数。[如果顶端终结计数只包括“全都是好的”，那么底端的终结计数就必须仅包括“全都是差的”（或相反）；如果顶端终结计数和底端终结计数有同样的标注，即两者全都是好的或两者全都是差的，那么合计终结计数下降为零]。

合计终结计数与置信度

根据组合方程，图基检验表明，合计终结计数与置信度之间有下列关系：

合计终结计数的数目	置信度
6	90%
7	95%
10	99%
13	99.9%

一个图基检验实例

通过实例可以最好地说明图基检验的简易性及强大能力。在对一种用于漏油的推力式柱塞密封进行调查时，测量了 8 个不漏油的（好的=G）和 8 个漏油的（差的=B）部件。测量的参数包括涡轮摇杆机构、油漏、涡轮表面光洁度、密封高度和垫圈的厚度。下面的读数是涡轮摇杆的：

0.015B	0.019B	0.017B	0.024G
0.018B	0.018G	0.019B	0.023G
0.014B	0.016G	0.011B	0.021G
0.022B	0.023G	0.007B	0.017G

（1）将以上涡轮摇杆机构的 16 个读数建成一个图基排列。

（2）确定顶端、底端以及合计终结计数。

（3）区分漏油与不漏油的部件时，涡轮摇杆机构是重要的，那么置信度是什么？

图基检验中的统计原理

（1）组合方程。图基检验的统计基础是组合方程：

$$\text{组合数目（无任何重复）} = n! / [(n-r)! r!]$$



式中, n 为数据中部件总数; r 为好部件的数目; $(n-r)$ 为差部件的数目。

因此, 没有任何重复的组合总数等于:

$$\begin{aligned} & \frac{16!}{(n-r)! r!} \\ &= \frac{16 \times 15 \times 14 \times 13 \times 12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1} \\ &= 12.870. \end{aligned}$$

所给出的这些组合, 只有两个完全不存在重叠机会的途径; 或者是所有 8 个好的部件的排列高于所有 8 个差的部件, 或者是反过来。这样一来, 得出错误结论的风险只有 $2/12\,870$ 或 0.016% 。因此, 16 的合计终结计数的置信度是 99.984% , 简言之, 是非常非常高的。

(2) 置信度在很大程度上与采样规模无关。令人惊奇的是, 与图基检验相关的置信度, 大多与好、差部件的采样数量无关, 只要每种部件的数量分别不小于 6 和 6 即可。例如, 如果采样数量是 12 个好的和 12 个差的, 那么有较大的重叠区也不碍事。因为总是由顶端和底端终结计数的和来确定置信度水平的。如果合计终结计数比较大, 比如说 20, 那么置信度甚至高于 99.9% 。增加到这样高的置信度, 并不证明采样越多就越好。

(3) 与组合方程相关的统计学原理要求好部件和差部件的数量为彼此的 20% 即可。

(4) 计算由“好”部件向“差”部件的初次和最后一次过渡时的终结计数。在一组读数排列中, 如果在由好至差初次过渡点或由好至差的最后过渡点处, 出现好和差的读数都一样的情况时, 该怎么办? 无论何时, 如果有一个结点——即在每两个过渡点上好与差间的每个相等的读数——则终结计数必须由 1 降低为 $1/2$ 。让我们考虑一下涡轮摇杆机构的读数 (表 11—1), 如果代换表 11—1 的读数, 就可得到:

表 11-1

图基检验排序

差	好	
0.007 0.011 0.014 0.015		顶端终结计数（全差）=4
0.017 0.018 0.019 0.022	0.016 0.017 0.018 0.019 0.021	
	0.023 0.023 0.024	底端终结计数（全好）=3

重叠区域

合计终结计数=4+3=7。因此，在解释推力式柱塞的漏油与不漏油的差别时，有95%的置信度说明涡轮摇杆机构是重要的。

方案 A		方案 A	
0.007	差	0.023	差
0.011	差	0.023	好
0.014	差	0.023	好
0.015	差	0.024	好
0.015	好		

- 在方案 A 中，顶端终结计数不是 4，而是 $3+1/2=3.5$ 。
- 在方案 B 中，底端终结计数不是 3，而是 $2+1/2=2.5$ 。

案例研究 1：微型马达的噪音



在一种静音寻呼机（该种寻呼机不是通过声音而是通过振动信

号来呼人)的生产中,存在扰人的噪音问题。部件搜索技术确定问题是出在日本厂商供应的微型马达上。但由于马达本身太小,所以不可能对其进行部件搜索。鉴于法律方面的原因,日本供货商对造成产品不合格的质量问题不承担责任,但被取消供货资格。于是,50 000个这样的马达(每个5美元)就变成了一堆废铁。

DOE小组在8个最好的(最安静的)和8个最差的(噪音最大的)马达之间进行了成对比较试验,他们选择了8个质量参数来研究,其结果见表11—2。

表 11—2 成对比较:微型马达

比较的参数	重复/非重复的差别
多层垫	差异轻微且不重复
含油轴承	差异轻微且不重复
填料轴承座	差异轻微且不重复
轴间隙——轴向和径向	差异轻微且不重复
前后轴承尺寸	未试验
配重	未试验
电阻和电流	未试验
转速(转/分)	重复的差异:好部件是4 800~5 000,差部件是6 900~7 200

资料来源: Motorola Inc., Boynton Beach, Florida.

结论:红X出现在好马达与差马达间的马达转速方面,转速有显著的40%的差异。

采取的措施

为了调查好马达与差马达间在转数方面的差别,需要拆卸微型马达。但是这不可能,于是DOE小组决定采取一种印刷电路板外的措施来解决差马达的噪音问题。他们在差马达的电池上串联一个18欧姆的电阻,使电池电压由1.3伏~1.5伏降低为1伏左右,这就使转速降低到4 500转/分~5 000转/分,结果噪音减弱了。你

看，加上一个 1 美分的电阻就拯救了 5 美元一个的小马达，整体节约（包括劳动力）达 248 000 美元以上。

关于微型马达案例的评述

(1) 没有对单对的好马达和差马达进行检测，也没必要。

(2) 没有对所列举的 8 个参数的差异进行量化。

(3) 没有这样的量化，就不能进行图基检验。

(4) 但是在好马达和差马达之间有一个能形成终结计数等于 16 的明显的区别（没有重叠），所以可以认定马达转数是解释好、差马达差别的重要因子，其置信度为 99.9%。

(5) 差马达的根本成因还没有解决，因为一个新的马达供应商取代了不合作的日本供应商。现在的目标是挽救差马达并大幅度降低废品的成本。

小组练习 1：隐形眼镜的处方缺陷

一个隐形眼镜制造商发现，其某些型号产品有 3% 的镜片不符合光学规范的要求。隐形镜片是由两片塑料曲面组成，中间用一种称作 hema 的化合物黏合。镜片一旦成形，不能拆开，所以部件搜索技术在此就不能应用。

DOE 小组选取了 6 个最接近处方技术规格要求的镜片和 6 个离处方要求甚远的镜片。要测试的质量参数是：

- (1) 圆柱度：镜片的前曲面；
- (2) BP 侧圆柱度：镜片后曲面；
- (3) UV 吸收性：吸收的紫外线的数量；
- (4) 偏光性：经过镜片光线衍射的模式；
- (5) 模具：镜片成形的模具。

结果如表 11—3 所示。



表 11—3

成对比较：隐形眼镜处方缺陷

	测量的参数				
	圆柱度	BP 側圓锥度	UV 吸收性	偏光性	模具
好	0.030mm	0.074mm	9.8%	彩虹/左/无模式	凸起
好	0.020mm	0.043mm	8.8%	彩虹/左/无模式	不凸起
好	0.018mm	0.069mm	10.9%	彩虹/左/无模式	凸起
好	0.016mm	0.058mm	11.2%	彩虹/左/无模式	凸起
好	0.030mm	0.063mm	9.9%	彩虹/左/无模式	不凸起
好	0.026mm	0.072mm	8.9%	彩虹/左/无模式	不凸起
差	0.055mm	0.053mm	9.3%	彩虹/左/无模式	不凸起
差	0.053mm	0.048mm	7.8%	光/左/模式	不凸起
差	0.051mm	0.049mm	7.4%	光/无/模式	不凸起
差	0.048mm	0.049mm	8.7%	光/右/模式	凸起
差	0.051mm	0.077mm	8.8%	光/左/模式	不凸起
差	0.056mm	0.070mm	11.2%	彩虹/全部/无模式	凸起

资料来源：CIBA Corp., Atlanta, Ga.

问题

- (1) 确定红 X 和粉红 X 参数，证明你的选择。
- (2) 对这些参数进行图基检验，哪一个参数在解释好、差部件差异时有着 90% 以上的置信度？
- (3) 你怎样量化两个并非变量而是属性的参数？

答案

- (1) 红 X 是圆柱度，粉红 X 是偏光性。

● 关于圆柱度，所有好的读数都低于差的。好镜片的最高的读



数低于差镜片的最低读数，其间无重叠。

●关于偏光性——一种属性——好或差镜片都只有一种共同的光学模式。其余的5个好镜片的光学模式与差的镜片完全不同。

(2) 参见表 11—4。

表 11—4 隐形眼镜小组练习的图基检验

排序	参数			
	圆柱度 (mm)	BP 侧圆柱度(mm)	UV 吸收性 (%)	偏光性 (模式)
1	.016G	.043G	7.4 B	1G
2	.018G	.048B	7.8 B	1G
3	.020G	.049B	8.7 B	1G
4	.026G	.049B		1G
5	.030G	.053B	$\begin{bmatrix} 8.8 \\ 8.8 \end{bmatrix} \begin{matrix} B \\ G \end{matrix}$	1G
6	.030G	.058G	$EC = \frac{1}{2}$	
7	.048B	.063G	8.9 G	$\begin{bmatrix} 1B \\ 1G \end{bmatrix} EC = \frac{1}{2}$
8	.051B	.069G	9.3 B	
9	.051B	.070B	9.8 G	
10	.053B	.072G	9.9 G	2B
11	.054B	.074G	10.9 G	2B
12	.056B	.077B	11.2 B	3B
			11.2G	4B
				5B
合计终结计数	6+6=12	1+1=2	$3\frac{1}{2} + \frac{1}{2} = 4$	$5\frac{1}{2} + 5 = 10\frac{1}{2}$
置信度	99.7%	无	无	99.2%

●圆柱度的合计终结计数是 12，在解释好、差镜片的差别时，圆柱度是重要的参数，其置信度为 99.7%。

●BP 侧圆柱度合计终结计数是 2，在解释好、差镜片的差别时，BP 侧圆柱度作为重要的因子的置信度是比较低的。

●关于 UV 吸收性：对于合计终结计数为 4 的情况，顶端终结计数是 $3\frac{1}{2}$ (B 和 $G=1/2$ 之间 8.8% 处的一个结点)，底端终结计数也是 $1/2$ (B 和 $G=1/2$ 之间 11.2% 处的一个结

点)，所以，在解释好与差的差别时，UV 吸收性的重要性是不大的。

(3) 至于偏光性，可以建立任一种利克特度量尺度用于各种不同的模式，彩虹/左/无模式为第 1 种，光线/左/模式为第 2 种，光线/无模式为第 3 种，光线/右/模式为第 4 种，彩虹/全部/无模式为第 5 种。图基检验发现了一个顶端终结计数 5.5（在好与差之间的一个结点）及一个底端终结计数 5 或一个合计终结计数 10.5，因此可以认为偏光性是解释好与差差异的重要参数，其置信度为 99.2%。

(4) 关于模具，如果是依照一种利克特度量尺度，“凸起”为 1，而“不凸起”为 2，图基检验（表 11-4 未予显示）表明：几乎全部都存在重叠，模具这一参数完全不可能是重要的。

案例研究 2：4 兆的动态 RAM

一家大型半导体制造公司的设计部开发了一种 4 兆的动态随机存储器（RAM），但是由于质量试验总未过关，产品未能投产。其中一种质量试验是老化（高温处理），试验后有近 7% 的产品不合格。

失效模式是一种突然式的失效，比如有一半的存储（2 兆）被抹去。使问题更加复杂的是，在工程师们将许多失效的器件由老化架上取下后，它们会复原。这意味着问题是间歇式的。详细的失效分析没有任何结果。设计工程师的压力越来越大，因为竞争对手也生产了相似的器件。调查工作由几天延伸到几周，又由几周延伸至几个月，仍毫无结果。绝望中，公司请来了两位厂外顾问，还是未能解决问题。最后，工程师们联系到笔者，看看 DOE 是否能有所帮助。

我把多变量排除在外，因为大量的 2 兆失效器件没有准确的位置，而且好、差部件同处于同一时间。本想用部件搜索技术，



但由于器件被封在壳内，所以实行起来不太容易。我推荐采用成对比较方法，找出 8 个好的和 8 个差的，但工程师认为可能很难找出 8 个差的部件，因为老化之后它们会复原。我向他们肯定，这种间歇性无伤大碍。差部件中的某一些东西与好部件不同，我们必须坚持找到它们。我们列出下列参数清单用以检验好、差部件的差别：

- (1) 成形化合物；
- (2) 引线框架；
- (3) 金属丝（细如发丝）；
- (4) 芯片端的压焊点；
- (5) 硅晶片端的压焊点；
- (6) 老化架上的插座。

工程师们围绕着这 6 个参数展开了工作。如果在好、差部件之间没有发现重现性的差别，就接着对其他参数进行调查。

在一周的时间内，红 X 找到了。在前 5 个参数中，没有找到有价值或重复性的差别，问题在于老化架上的插座。差的插座上有一层绝缘膜——其某个部位存在漏点——它是用作电气隔离层，但是它很薄，在与部件接触的时候这个隔离层就击穿了，结果就变成了电气接触。原来产品本身没有任何问题，问题出在试验装置上。这个案例给了我们这样的启示：与部件对话，与部件对话，与部件再对话！这就是本书所要反复强调的。

小组练习 2：变形的栅格

某家用产品制造厂面临着一个恼人的长期问题：在焊接之后，产品上的栅格方形产生形变。问题已存在了 7 年。根据我们的 DOE 讨论会的建议，一个 DOE 小组利用成对比较方法找出了好栅格与差栅格间变量的重要原因。

图 11-1 是产品制造和装配流程图。表 11-5 显示了测量 6 个



好部件和 6 个差部件得到的 17 个参数 (列) 是绿 Y, 此处方栅格变形是以 C, D 间尺寸的差别来度量的)。

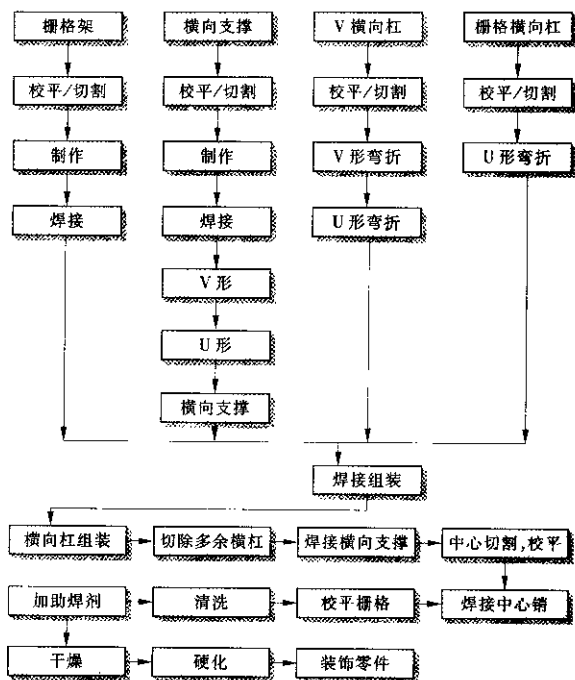


图 11—1 栅格制造和装配流程图

资料来源: Vitromatic Corp., Celaya, Mexico.

- (1) 与钩子的距离——左和右;
- (2) 总横向长度——前、中、后;
- (3) 总纵向长度——左、中、右;

表 11—5 栅格参数的成对比较

抽棉参数的成对比较																		
形状	C/D 对	重量按框架平直度										4 个角的角度				弓形		
		总纵向长度					总横向长度					4 个角的角度						
		左	右	前	中	后	左	中	右	前	后	左	右	前	后			
好	0	25	45	217	216	216.5	484	484	184	X	V	S	S	89.5	89.5	90.5	90.0	N
好	-1	50	60	216	215	216	484	485	484	X	V	S	S	89.5	89.5	91.0	90.0	N
好	1	28	28	216.5	215	216	484	484	181	X	V	S	S	89.5	90.0	90.5	90.0	N
好	0	26	27	216.5	215	216	184	485	184	X	V	S	S	89.5	90.0	90.5	90.5	N
好	0	45	60	217.5	216	216	484	485	185	X	V	S	S	89.5	89.0	90.5	90.5	N
好	0	28	30	217	216	216	484	484	184	X	V	S	S	90	89.0	90.0	90.0	N
坏	9	41	15	218	215	216	184	485	486	V	X	V	S	91	88.5	88	92	R
坏	8	50	55	217	214	215	481	486	486	V	X	X	V	89	89.5	93	89.5	L
坏	8	83	23	216	215	215	484	486	484	V	X	S	V	88.5	91	91.5	89.5	L
坏	10	90	20	217	215	216	183	484	484	V	X	X	V	88	91.5	92.5	87.5	L
坏	9	85	21	216	215	216	481	485	484	V	X	X	V	88.5	91	92.5	89	L
坏	10	56	45	216	213	215	184	485	485	V	X	S	V	88.5	90.5	92.5	88.5	L

注: 1. 金属框架平直度

2. 弓形方向

L=左

R=右

N=无弓形

S=平直

资料来源: Vitromatic Corp., Celazo, Mexico.

- (4) 金属丝框架的平直度：短边（前和后）、长边（左和右）；
- (5) 四个角的角度；
- (6) 弓形形变的方向。

问题

在所测量的参数里，哪些是重要的和不重要的？进行图基检验，以便确定每个参数的置信度。

答案

表 11-6 是 17 个栅格参数进行图基检验后的汇总。它显示在表 11-7 中有着大于 6 的合计终结计数，因此，它们在确定好与差部件的差别时是重要的参数，其置信度为 90% 以上。

评论

该 DOE 试验主要是由一家跨国公司的一线劳动小组进行的，他们对英语一知半解，只参加了一天我们的学术研讨会。这次试验堪称一个精彩的“与部件对话”实例——得到了有力的从何处入手的线索，从而把下阶段的工作集中于排除这一拖延了 7 年的问题。更重要的是，它说明了我们曾多么低估了一线劳动者的能力和效率，他们过去只是迫于漫不经心的管理层的要求付出体力，而不是脑力。

当中心好而两头差的时候，图基检验的排列方法

图基检验的数据通常根据产品质量特性或参数由最小到最大（或反过来）排列，其中最小的是好的，最大的是差的（或反过来）。

表 11—7

具有高置信水平的参数

重要参数	合计终结计数	置信度
1. 金属丝框架的平直度		
a. 前: 凸出现象好于凹进现象	12	99.7%
b. 后: 凹进现象好于凸出现象	12	99.7%
c. 左: 平直度好于凹进现象	12	99.7%
d. 右: 平直度好于凹进现象	12	99.7%
2. 弓形方向: 无弓形好于左面弓出形变	11	99.2%
3. 距左钩子的距离	$8\frac{1}{2}$	97%
其余 11 个参数	<6	无置信度

但是, 如果在排列顺序里, 中间的是好的, 而两头是差的, 这又该怎么办呢? 这是一个典型的参数数据分布问题, 其中排在好的队列里的反映实际的变量数据处在中间部位, 而排在反映这些变量的差的队列里的数据分布在尾端, 如图 11-2 所示。

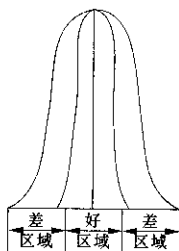


图 11—2 参数的分布与排列的关系

根据前面讨论过的图基检验的通用准则, 如果排列的两个尾端都是差的 (或都是好的), 则终结计数等于零, 且此时如果参数是重要的, 则是没有置信度的。但是有一个方法能把中间是好的、两

端是差的这种排列方式转化为图基格式。这时，不是由最小到最大这样排列，而是根据某个参数相对其目标值的偏差（正或负）来排列。于是，好的部件趋向于有一个较小的偏差，而差的部件趋向于有一个较大的偏差。下面我们通过实例来说明这种技术。

表 11—4 所示为隐形镜片的图基检验，让我们假定圆柱度参数的排列与从前一样，但是好（G）与差（B）镜片的分布如表 11—8 的列 1 所示。我们还假定圆柱度的目标值是 0.040 毫米。列 2 表示的是与每个所观察的镜片与其目标值（0.040mm）的偏差（正或负都没有关系）。然后将这些偏差量排序于列 3，其中顶部的 5 个部件的偏差值较小，是好的；底部的 5 个部件具有大的偏差值，是差的；得到顶部的终结计数是 5，底部终结计数是 5，其合计终结计数是 10，置信度为 99%。

表 11—8 图基检验的实例：中间好，两头差

初始 排序	列 1	偏离目标值 0.040 的列子	偏离目标值 的新排序
1	.016B	.024B	.008G
2	.018B	.022B	.010G
3	.020B	.020B	.010G
4	.026G	.014G	.011G
5	.030G	.010G	.011G
6	.030G	.010G	.013B
7	.048G	.008G	.014G
8	.051G	.011G	.015B
9	.051G	.011G	.016B
10	.053B	.013B	.020B
11	.055B	.015B	.022B
12	.058B	.016B	.024B
终结计数			5+5=10

成对比较在管理工作中的应用

如同多变量和部件搜索技术一样，成对比较技术也可用于非生产性的场合以及管理工作。在这些应用领域中，它是如此通用，以至于成为整个线索生成系列的“老黄牛”。本书作者将成对比较方法用于：

- 销售与市场营销；
- 人力资源；
- 商务、就业和服务；
- 医院；
- 农场。

机会是无穷的！两个实例足以说明成对比较技术在这些应用中的简易性和强大能力。

案例研究 3：牛奶场的细菌指标

一家著名的跨国公司为美国和国外的牛奶场提供制奶设备，它的一个子公司则提供化学药剂，用于牛奶场的生产线在实际制奶前的消毒清洁工作。

公司接到一些牛奶场的投诉，称化学药剂没能充分地降低细菌数量水平，而这些细菌的指标远低于美国政府规定的最大值要求。但是，奶农争辩说，牛奶场的合作者（他们的顾客）将根据牛奶中所含细菌数的降低规模来付款，所以投诉者责备公司所提供的化学药剂使他们的收入减少了。

本书作者作为公司制奶设备的常年顾问，应邀调查了公司的化学子公司的细菌问题。

我建议展开 DOE 成对比较试验，并选择了其中 6 个经常投诉的牛奶场和 6 个最好的牛奶场（无投诉）——它们全都在威斯康星，然后对它们的牛奶场的状况展开调查。一位训练有素、经营家

庭牛奶场多年的公司代表，被选中利用如下 25 家牛奶场的参数，采集这 12 家奶场的综合数据。

A. 绿 Y（所有 6 家“差”牛奶场的细菌计数比 6 家“好”牛奶场要高）。

（1）细菌计数

（2）巴氏灭菌法灭菌后培养计数（可以引起酸臭味，货架寿命较短）

（3）红/白血细胞计数

B. 输入参数：这里实际的变量是不可能存在的，建立一个 1~10 的利克特度量尺度，其中 1=最差，10=最好。

（1）奶牛/牛乳头清洗

- 奶牛的清洗
- 牛乳房清洗
- 牛乳房的卫生处理（预先和随后）
- 牛乳头的清洗
- 挤奶前挤奶装置的清洗
- 挤奶后挤奶装置的清洗
- 过滤器的清洗

（2）水分析

- 硬度
- 缓冲剂（重碳酸盐）
- pH
- 铁质

（3）设备清洗

- 系统的配置
- 化学品的类型
- 消毒
- 闲置周期
- 预先冲洗温度
- 洗涤剂清洗周期



- 洗涤剂清洗温度（起始和终结）
- 酸洗周期
- 酸冲洗（起始温度和时间）
- 线的斜度
- 卫生消毒

对 25 个参数的每一个都进行了图基检验。表 11—9 即为试验所观察到的 7 个重要参数的合计终结计数。

正如期望的那样，化学药剂的类型（最初投诉的根源）的合计终结计数仅为 1。其余的 17 个参数的终结计数较低，为不重要参数。

表 11—9 案例研究 3 的图基检验结果

参数	合计终结计数
消毒	12
预先冲洗	12
温度	10
牛奶过滤器	10
洗涤剂清洗周期	9
牛乳头清洗	9
过滤器清洗	6.5

整个研究工作耗时 2~3 个月，揭示出牛奶场必须严密控制其生产操作。公司代表回到有问题的牛奶场，指明了应改进的地方；有 4 家改进了，后来没有再被投诉，另两家则拒绝改变。他们坚持说：“我就是这样办牛奶场的，我爸爸就是这样办牛奶场的，我爷爷也是这样做的。你凭什么告诉我应该怎样办牛奶场！”

我劝总公司的老板在牛奶场的“质量审核”方面投点本钱，把它作为一种业务，在服务时收取一定的费用。他同意了，但没有完全采纳我的建议，而是免费为奶场提供质量审核服务，条件是买他们的设备和化学药剂。结果是：公司赢得了竞争对手的 60% 的客户！

案例研究 4：临时雇员的高流动率

作者在一家有名的计算机外围设备制造厂作 DOE 咨询时，被人力资源部经理请去，看看是否 DOE 技术能用到他主管的部门中以解决一个恼人的问题，也就是在一线劳动的临时雇员的高流动率问题，估计每年公司因此要损失 80 万美元。按政策，公司有一个恒定的一线雇员的核心，其余的人即所谓的“临时工”。但是这个公司发展得太快，根本就没有临时工被解雇，而且他们之间的摩擦也是要考虑的事情。

我提议采用成对比较方法，对 8 个受雇之后很快离去的临时雇员 (WOW) 和 8 个能在公司呆上两年以上的临时雇员 (BOB) 进行采样。人力资源部门选择了一些人，并为成对试验建立了如下的参数：

- (1) 通勤的距离——由家至公司；
- (2) 驾驶执照；
- (3) 受教育水平；
- (4) 在其他公司时的跳槽率（流动率）；
- (5) 对管理层的看法；
- (6) 工作日程的方便性（公司对一线雇员安排倒班工作）；
- (7) 对于社会上同等工作的薪金的看法；
- (8) 相对于固定员工，有关待遇的看法。

这里，要想测定实际的变量是不可能的，所以要定制一个 1~10 的利克特度量尺度（1=最差，10=最好）。

对 8 个参数进行了图基检验后，只有两个参数的终结计数达到了 6 以上。

(1) 相对于固定员工的待遇问题：终结计数为 7，置信度为 95%；

(2) 通勤距离：终结计数等于 6，置信度为 90%（没有一个通勤距离超过 15 英里的雇员能在公司呆上 6 周以上）。

(3) 第 3 个参数（受教育水平）的终结计数是 5，但是没有评

定价值。(具有与职位相应的学位或更高学位的人,似乎流动率更高。)

在雇用职员时,公司在通勤距离和学历方面不可能做更多的事情,但是临时雇员相对于固定员工的待遇问题——这更多的是前者的自身感受而不一定是实际情况,对人力资源部门则是很大的挑战,但也是一个有所作为的项目。

DOE 小组的问题

(1) 绿 Y 被量化和测量过了吗? 测量仪器的精度至少是产品允差的 5 倍以上吗?

(2) 如果绿 Y 是个属性,它被转化为李克特度量尺度的变量了吗?

(3) 部件搜索——拆卸/重组过程中绿 Y 没有显著的改变——这种技术被排除掉,需要应用成对比较吗?

(4) 是否选取了 6 或 8 个最好的部件 (BOB) 和 6 或 8 个最差的部件 (WOW),以便进行成对比较?(在不可避免的情况下,量化非常受到局限时,可以通过仅仅 3 个 BOB 与 3 个 WOW 进行成对比较,其终结计数等于 6,但是其后不允许有重叠现象产生。)

(5) 如果 BOB 与 WOW 不可能有相等的数目,它们重现的计数是分别处于彼此的 20%之内吗?

(6) BOB 与 WOW 之间的分隔是尽可能的大吗? 比如说,是一天或一周之内的产品吗?

(7) 如果用成对比较调查处于顾客手里的部件,需要进行现场调查,你是在现场暴露的可比较的时间里进行 BOB 与 WOW 的选择吗?

(8) 列出一张含所有可能的质量参数(这些参数可以解释 BOB 和 WOW 之间的差别)的清单了吗?

(9) 清单以递减顺序排列可能产生差别的原因吗?

(10) 清单是被分离成数个非破坏性试验（在为探求差别而有可能进行的破坏性试验之前，首先在 BOB 和 WOW 部件上进行）吗？

(11) 是否有适当的方法用以验证所确定的 BOB 和 WOW 差别——例如，SEMs、X 射线法、硬度计、频谱分析法、有限单元法、形态图形法和光谱法？

(12) 把图基检验用于分离重要参数（终结计数为 6 以上，置信度为 90% 以上）和不重要参数（终结计数为 5 以下，置信度小于 90%）了吗？

(13) 把与好部件相关的参数用于确定或修改与部件有关的工程技术规格要求了吗？

(14) 把与差部件相关的参数用于置疑或修改与部件有关的技术规格要求了吗？

评审时高层管理评审的问题

除了与 DOE 技术相关的一般的问题之外，应提出下列有关成对比较的一些特定问题：

(1) 好、差部件是在大致同样的时间内生产出来的吗？

(2) 在不损坏部件甚至不改变其绿 Y 值的条件下，这些部件就不能进行拆卸/重组装吗？

(3) 比起其他技术，为什么用成对比较技术较好？

(4) BOB 与 WOW 之间有足够的隔离吗？

(5) 所制定的清单，能够尽可能地包括所有涉及 BOB 和 WOW 的参数吗？

(6) 非破坏性试验是在破坏性试验之前进行的吗？

(7) 如果鉴别出几个重要变量，那么在成对试验之后，还跟随有其他技术，如变量搜索或全析因技术吗？

(8) 随后是否进行了最终的“B vs. C”技术，以便肯定成对比

较试验的结果？

成对比较试验的后续工作：注意事项

成对比较是一种线索生成技术，有时线索（特别是已经鉴别出一个单独的红 X 后）能精确地定位根本性原因，并修正它。但是经常要通过一个随后的试验来量化重要的变量及其交互影响，例如用变量搜索法或全析因法以及“B vs. C”确认法等。图基检验并不能保证 100% 的置信度。90% 的置信度意味着还有 10% 的风险：标定的重要参数并不一定是真正的重要的参数。试验者必须准备好应对这种风险，即使它们只是偶尔发生。





第 12 章

产品/过程搜索：精确定位过程变量

产品/过程搜索是作者新近发展的一种技术，它是在不打断生产过程的前提下解决问题的另一项突破。因此，它成为我们的第 5 种线索生成技术。

◆ 产品/过程搜索的目标

正像成对比较是将重要的产品参数与不重要的产品参数相分离一样，产品/过程搜索技术是将重要的过程参数与不重要的过程参数相分离。

产品/过程搜索的原理

(1) 生产中一个部件中的任一变量有着两个通用的因子，或者是产品本身有变异（由材料引起），或者是一个以上能影响产品的过程参数有变异。

(2) 如果在已完成的产品中有变量，部件搜索或成对比较技术就能够探测到这些产品变量的原因，这就是产品/过程搜索的产品部分。

(3) 但是如果原因是过程参数的变异：移动、漂移波动或随时间而改变，就可以利用产品/过程搜索的过程部分来解决。它利用与成对比较技术相同的原理，即利用 6 个或 8 个好产品和 6 个或 8 个差产品，但它是把与 6 个或 8 个好产品相关的过程参数与 6 个或 8 个差产品的过程参数相比较。

(4) 产品/过程搜索是多变量研究的自然的后续工作，其中时间对时间族系是红 X，产品参数是随假设时间而波动或改变的（此项技术在路径图上的位置，请参见图 7—1）。

产品/过程搜索的方法

(1) 如果怀疑过程参数变量是造成产品好或差的可能原因，就把这些过程参数以同类递降顺序列出一个清单。

(2) 确定怎样测量每个过程参数，谁来测量，在何处进行精密测量。

(3) 要保证测量仪器的精度至少为过程参数（允差）的 5 倍以上。例如，如果炉温是 $1\,200^{\circ}\text{C} \pm 5\%$ ，那么测量仪器的精度最大只能是 $\pm 1\%$ 。

(4) 要保证测量的是实际的过程参数，而不仅仅是设定值。

(5) 如果一个特定的过程参数在监测过程中并不变化，就可在

进一步的考虑中将其排除掉。

(6) 运行 100% 的部件采样 (特别是在高废品周期中, 如果废品不断) 或部件的多变量采样, 一直到:

- 在过程的终结阶段, 最少能采集到 8 个好部件和 8 个差部件;

- 最好部件和最差部件之间的分布范围, 应该占在产品生产过程中观察到的有意义的变化的 80%。

(7) 测量所有与每个部件相关的、已经指定的过程参数 (或者是 100% 部件, 或者是在多变量中选择的样本)。

(8) 在产品制造过程中, 直到过程结束前, 是不可能对部件是好或差作出结论的。这意味着许多潜在的好部件必须与潜在的差部件同时被测量, 以最终达到最少 8 个差部件的标准, 特别是如果有意义的缺陷率比较低时——比如说是 1%。平均而言, 在可能出现 1 个差部件之前, 针对每一个过程参数都必须测量 99 个好部件。所以, 为了得到最少 8 个差部件, 必须测量将近 800~900 个好部件, 即在它们经过生产过程时逐一进行测量。

这是一个烦琐而耗时的的工作。不过, 现代的许多过程都配备有自动数据识别和控制系统, 可以快速、自动地记录每个部件的过程参数。

(9) 然后, 运行与 8 个好部件和 8 个差部件相关的过程参数的成对比较。对每个参数都要进行图基检验, 然后计算其合计终结计数。如果终结计数是 6 以上, 则该过程参数是重要的, 其置信度为 90% 以上。

(10) 如果经鉴别后, 有数个过程参数具有 90% 以上的置信度, 那么就进行 “B vs. C” 试验 (见第 15 章), 以便验证哪个是真正的重要参数。说到底, 毕竟 90% 的置信度不等于 100% 的置信度, 还是可能有某些参数被遗漏或者某些参数未重复出现。

(11) 接下来, 要进行变量搜索 (见第 13 章) 或完全析因法 (见第 14 章), 以量化最重要参数和它们的交互影响效应。

(12) 再通过散布图或调优运算, 随后再进行正向控制、过程



确认和预先控制等步骤进一步优化处理（见第 16、17、18 和 19 章）。

(13) 可以扩展第 8 步的不重要参数的公差，以便减少成本，虽然这样或许需要进行某些试验才能确定要扩展到什么程度。

产品/过程搜索：当不能测量到个别部件的过程参数时

在某些过程中，有个别产品的部件在生产时不能测量到与之相关的参数。例如铸造和玻璃加工中的批量过程或加热就属于这种情况。此时，与作为总体的每批相关的过程参数必须予以记录，对与 8 批好的和 8 批差的产品相关的过程参数进行成对比较。这种方法不像每个单独的部件的过程读数那么灵敏，但是在这种环境下，它是惟一可行的选择。

案例研究 1：发蔫的莴苣/包装袋漏气

一家为世界快餐连锁店供应莴苣的公司，不断接到几家连锁店的投诉，说是收到的莴苣都变蔫了。莴苣暴露于空气中，1 天就会变色、失水和变味，所以必须在运输前将莴苣放进真空包装袋里，这样可以使莴苣的保质期从 1 天延长到 2 周。

公司惟一的顾客就是快餐连锁店，所以它很怕失去全部业务。笔者接到公司董事长打来的紧急电话，让我去帮助解决问题。

过程

莴苣是装在冷藏卡车中由农场运来的。在加工厂，温度保持在



较冷的 $34^{\circ}\text{F} \sim 37^{\circ}\text{F}$ ，莴苣经过去皮、清洗、干燥，在大型机器上按顾客的尺寸规格要求进行切割。下一步就是包装：切好的莴苣放进一个滑道，然后落进另一个滑道的聚乙烯袋中，充满袋子的 $2/3$ 之后抽成真空，最后封口，装入纸箱，运到快餐店。

一个早期的绿 Y

在与公司的管理人员初次会谈时，我的第一个任务就是定义和量化绿 Y。最重要的绿 Y 就是变蔫的莴苣，这很难量化，要花费相当长的时间去测量。公司知道问题的原因是真空漏气——这是一个早期的绿 Y。但是，在我们量化该参数时，一位经理确定了漏气袋子与一小块莴苣之间的关系，这一小块莴苣卡在莴苣床之间的区域里，特别是处于袋子的顶部、需要密封的区域时情况更严重。于是，我们确定莴苣的块数是更早期的绿 Y。

快速多变量运行表明：变量随时间而变，而且是在袋内。这表明要进行产品/过程搜索试验，因为：（1）怀疑有几个变量在包装过程中是随时间而变化的；（2）集中图显示，在热密封过程中，在密封区内的特定区域里有大量莴苣块堆积现象。

DOE 小组画出了主要的过程变量清单：

- （1）抽气管的位置（插入莴苣床的深度）。
- （2）抽真空时间。
- （3）抽气管的移动速度。
- （4）密封时间。
- （5）密封时的加热。
- （6）冷却时间。
- （7）机器的 PSI（压力）。
- （8）断电的延迟。
- （9）抽拔抽气管的延迟。

另一些过程变量则作为备份，以备主要的过程变量的合计终结计数值小于 6 时再予应用。它们是：



- 真空 PSI。
- 真空压力。
- 传送带的高度。
- 密封宽度。
- 抽气管的角度。
- 袋子的振动。
- 袋子的重量。
- 袋子的摆放。

结果

绿 Y：在每个包装袋中发现的莴苣切片的数目的变化是从无（好袋子）到 2~4 块（差袋子）。对 9 个过程参数的产品/过程搜索一直进行到产生出 8 个好袋子和 8 个差袋子。这些过程参数的成对比较如表 12-1 所示。

与抽气管的位置、抽真空时间、抽气管的移动速度及机器的 PSI 等参数有关的终结计数分别是 16、6、6.5 和 9，这 4 个过程参数是重要的参数（置信度为 90% 以上）。抽气管的位置（插入莴苣床的深度）是红 X。这是逻辑上的结论，因为抽气管深入到莴苣切片床处，会把莴苣片吸到绿 Y 处，但如果抽气管仅仅是贴近莴苣床，也可能不造成缺陷。由于每一批莴苣的密度和湿度都不一样，所以很难控制莴苣在袋中的高度。但是包装机供应厂商进行了改进，使得抽气管不再深入莴苣床中。

随后的试验与结论

下一步是对上述的四个变量以及包装机供应商推荐的另两个变量进行变量搜索试验。结果是缺陷率由 12% 降为零，每年可节约 350 000 美元。更重要的是，公司挽救了业务，留住了客户。公司甚至制作了一套演示幻灯片送给客户的上层管理人员，以显示我们





表 12-1

产品/过程搜索:案例研究——包装机

过程 参数	油气管 位置	抽真空 时间	抽气管 速度	密封 时间	密封 温度	冷却 时间	机器 压力	断电 延迟	抽拔抽气管 延迟时间
参数范围	6~12英寸 插入深度	0.85分	100秒	0.5分	290°F	0.3分	95 PSI	0.1分	0.1分
1	1/2B	0.85B	100B	0.9B	307B	0.5G	95G	0.1B	0.1B
2	1/2B	0.84B	100B	0.9G	307G	0.5G	95G	0.1B	0.1B
3	1/2B	0.84B	98B	0.9G	305G	0.5B	95G	0.1G	0.1G
4	1/2B	0.83B	99G	0.9B	305B	0.5B	97G	0.1G	0.1G
5	1/2B	0.81G	98B	0.9G	305G	0.5B	96B	0.1B	0.1G
6	1/4B	0.80G	97G	0.9G	304B	0.5G	98B	0.1B	0.1G
7	1/4B	0.77B	97G	0.9B	304G	0.5B	100G	0.1G	0.1B
8	1/4B	0.77B	95G	0.9B	302G	0.5G	100G	0.1G	0.1G
9	0G	0.75G	95B	0.9G	300B	0.5B	101B	0.1B	0.1G
10	0G	0.75G	95B	0.9B	300B	0.5B	103G	0.1G	0.1B
11	0G	0.75G	94B	0.9B	300G	0.5B	103G	0.1B	0.1B
12	0G	0.72B	93B	0.9G	300G	0.5G	104B	0.1G	0.1G
13	0G	0.70G	92G	0.9G	300B	0.5G	104B	0.2B	0.1G
14	0G	0.69B	92G	0.9B	299B	0.5G	105B	0.2G	0.1B
15	0G	0.68B	90G	0.9G	298B	0.5B	105B	0.2G	0.1B
16	0G	0.68G	90G	0.9B	297G	0.5G	105B	0.2B	0.1B
终结计数	8+8	4+2	$2\frac{1}{2}+4$	0	1+1	0	4+5	0	0
合计终结 计数	16	6	$6\frac{1}{2}$	0	2	0	9	0	0

注:每种参数的变数是以递增或递减顺序来排列的。

B=差部件,G=好部件;阴影区为终结计数。

的 DOE 技术的简易性和强大功能。

案例研究 2：玻璃屏幕上的气泡

一家跨国公司的拉美工厂制造玻璃屏幕，其最终产品存在气泡，客户对此很不满，缺陷率约为 13%。

公司的工程师进行过田口试验，缺陷率由 13% 改善到 6.2%。但工厂的管理层仍不满足于这种程度的改善，于是请我去解决这个问题。

流程图

加工过程是从加热炉开始的，在这里用很高的温度把玻璃熔化，然后经过两个供料器传送到两个压模里并成形，其后屏幕被传送到退火工段，最后进行检测。

对有意义的数据的检验表明：变量是随时间而改变的，有的时段缺陷率为 0，而有的时段缺陷率为 10%，甚至有着更高的跳跃。这种现象表明必须回避多变量技术，而要尝试产品/过程搜索试验。工厂对关键过程参数有着优良的监测系统，这一事实也支持了我们采取的这个决定。

因为玻璃生产是一种批量生产过程，对于单个的屏幕，是不可能监测到它们的重复过程参数的，所以部件是一批而不是单个的屏幕。我们选取了下列 8 个过程参数：

熔炉/供料器	压模
炉冠部温度	压力 #1
炉底温度	压力 #2
供料器温度（进口）	箱体高度 #1
供料器温度（出口）	箱体高度 #2



绿 Y 是每批所有屏幕的气泡数量。（建立了一个李克特尺度，以便区分气泡大小、气泡的深度和气泡的数量。将 8 批好的和 8 批差的各自记录下来，其重复出现的过程参数见表 12-2。）每个参数的 16 个读数不是真正的读数，而是偏离最佳预期值的偏差（+ 或 -）。至于两个炉温，最接近预期最佳值的屏幕是好的，处于温度尾部的屏幕是差的，这表明必须用辅助方法计算终结计数，即把上尾部（差的）、下尾部（差的）和中间的（只有好的）相加。此方法还可用于箱体高度 #2 的计算。其他 5 个参数从好至差有着单向的趋势，每个参数在数值上是增加的，所以可用标准的图基检验法来计算终结计数。

结果

表 12-2 显示：

- 红 X 是炉底温度，终结计数是 11，应该将其维持在（好的区域）1 353°F~1 359°F 之间；
- 粉红 X 是箱体高度 #2，终结计数是 10，应该维持在（好的区域）39cm~40cm 之间；
- 浅粉红 X 是炉冠部的温度，终结计数为 9，应该将其维持在（好的区域内）1 567°F~1 569°F；
- 浅浅粉红 X 是压力 #2，终结计数是 6，维持在 520°F；
- 一项“B vs. C”试验表明：由于改进了参数，使 B 批的气泡缺陷降低到 0.4%，相比之下 C 批的缺陷率是 6.1%。

还需要进一步的 DOE 试验——通过一个 2^4 全析因试验来量化相互作用的效应、用散布图或调优运算进行参数优化——最终将缺陷水平降为零。

小组练习 1：塑料注塑成形

一家大型玩具制造厂在用注塑成形机制造塑料部件时，有着很

表 12-2

产品/过程搜索: 案例研究 2——玻璃屏幕上的气泡

过程 参数	炉顶部 温度	底部 温度	给料部温 度(进)	给料部温 度(出)	压力 #1	压力 2#	箱体高 度#1	箱体高 度2#
参数	1533°F	1350°F	1098°F	1030°F	445PSI	515PSI	21cm	36cm
范围	1570°F	1370°F	1126°F	1050°F	515PSI	520PSI	29cm	44cm
期望的 最佳值	1565°F	1360°F	1123°F	1040°F	475	515	29	39
袋号								
1	-10B	-15B	-5G	-2G	-12B	-4B	-7B	-3B
2	-10B	-13B	-5G	-2B	-10B	-4B	-6B	-3B
3	-9B	-13B	-5B	-2B	-5G	-3B	-5G	-2G
4	-8G	-12B	-5G	-1G	-2B	-2G	-4B	-2G
5	-8G	-10B	-5B	-1G	0G	-2G	-4G	-2B
6	-5G	-7G	-5B	-1B	+5G	-1B	-4G	0G
7	-3B	-6G	-5B	-1B	+8G	-1G	-3B	0G
8	-1G	-4G	-5G	-1G	+12G	0G	-3G	0G
9	0B	-4G	-5G	0G	+12B	0B	-3B	+1G
10	+1B	-2G	-5B	0B	+14G	0B	-2B	+1G
11	+2G	-1G	-5G	0B	+20G	0G	-2G	+3B
12	+3G	0B	-5G	0G	+22B	+2B	-2G	+3B
13	+3G	+1G	-5B	0G	+25B	+2B	-2B	+3G
14	+4G	+2B	-5B	0B	+26G	+3G	-2B	+4B
15	+5B	+3B	-5B	+1G	+27G	+4G	-1G	+5B
16	+5B	+5B	-5G	+1B	+28G	+5G	-1G	+5B
总计数	3+2+4	2+3+6	0	0	2+3	3+3	2+2	2+3+5
合计终 结计数	=9	=11	=0	=0	=5	=6	=4	=10

注: B=差部件, G=好部件; 阴影区表示终结计数。

高的废品率（超过 20%）。绿 Y 是注料不足，对 8 个过程参数进行产品/过程 DOE 试验：

- 模具温度；
- 材料温度；
- 初始压力；
- 维持压力；
- 注塑速度；
- 螺杆速度；
- 模具排气；
- 注塑时间。

注塑成形机装备有先进的自动数据获取系统，通过它可以把每次注料的关键数据记录下来，其结果如表 12—3 所示。

表 12—3 产品/过程搜索：案例研究——塑料注塑成形机

过程参数	模具温度	熔化温度	压力	维持压力	注塑速度	螺杆速度	模具排气	注塑时间
推荐的	80°F ~	430°F ~	1 000PSI ~	75PSI ~	1 英寸/	2 英寸/	0.000 5 ~	
参数范围	100°F	450°F	1 100PSI	90PSI	秒	秒	0.001 5	2 秒
实际情况								
1B	85	405	1 055	82	1.0	2.0	0.000 5	2.0
2B	88	420	1 040	85	1.0	1.9	0.000 5	2.1
3B	101	409	1 060	83	1.0	1.9	1.000 5	2.1
4B	103	403	1 020	77	1.1	2.0	0.000 5	2.0
5B	95	430	1 010	85	1.0	2.0	0.000 5	2.1
6B	90	416	1 050	82	1.2	2.0	0.000 5	2.0
7B	100	401	1 095	84	1.2	1.9	0.000 5	2.0
8B	101	406	1 100	80	1.1	2.0	0.000 5	2.0
9G	89	432	1 190	77	1.0	2.0	0.000 5	2.1
10G	104	416	1 055	81	0.9	2.0	0.000 5	2.0
11G	101	438	1 075	76	1.0	2.0	0.000 5	2.0
12G	93	408	1 015	80	0.9	1.9	0.000 5	2.1
13G	101	422	1 075	78	1.0	2.0	0.000 5	2.0
14G	104	418	1 050	75	1.0	1.9	0.000 5	2.0
15G	92	440	1 055	79	1.0	2.0	0.000 5	2.0
16G	99	445	1 070	83	1.0	2.0	0.000 5	2.0
终结计数	4	8	0		6	0	0	0

注：B=差部件，G=好部件。



问题

- (1) 8 个过程参数中的每一个参数的终结计数是多少?
- (2) 哪些过程参数是重要的? 哪些是不重要的?
- (3) 机器供应厂商所推荐的每个参数的范围合理吗?
- (4) 如果不合理, 那么在下一轮试验里, 应该在什么样的范围里尝试?

答案

(1) 题和 (2) 题。见表 12-4。

(3) 题和 (4) 题。厂家推荐的模具温度范围是 $80^{\circ}\text{F} \sim 110^{\circ}\text{F}$, 显然是太宽、太低了。好部件所显示的温度是 $89^{\circ}\text{F} \sim 104^{\circ}\text{F}$ 。再严格一点, $95^{\circ}\text{F} \sim 105^{\circ}\text{F}$ 比较合理 (见表 12-5)。

表 12-4 过程参数和它们的重要性

参数	终结计数	置信度	重要性
模具温度	4	--	不重要
材料温度	8	97%	重要
初始压力	0	0	不重要
维持压力*	$3\frac{1}{2} + 2\frac{1}{2} = 6$	90%	重要
注塑速度	6	90%	重要
螺杆速度	0	0	不重要
模具排气	未改变	—	不重要
注塑时间	0	0	不重要

* 对于维持压力, 从差至好以及从好至差的第一次和最后一次过渡时有结点, 产生 $2\frac{1}{2}$ 终结计数。



表 12—5

参数范围

过程参数	厂商推荐范围	评论	好读数的范围	下一轮目标值
模具温度	80°F~110°F	终结计数低	89°F~104°F	不需改变
材料温度	410°F~ 450°F	太宽	408°F~445°F	44°F
初始压力	1 000PSI~ 1 100PSI	终结计数为 0	1 055PSI~ 1 090PSI	不需改变
维持压力	75PSI~ 90PSI	太宽	75PSI~ 83PSI	77PSI
注塑速度	1 英寸/秒	太窄	0.85 英寸/秒~ 1 英寸/秒	0.95 英寸/秒
螺杆速度	2 英寸/秒	终结计数为 0	1.9 英寸/秒~ 2 英寸/秒	不需改变
模具排气	0.001~ 0.001 5	终结计数为 0	0.000 5	不需改变
注塑时间	2 秒	终结计数为 0	2 秒	不需改变

作为产品/过程搜索技术指南的 DOE 小组问题

(1) 对绿 Y 进行定义、量化和测量了吗 (仪器的精度是否至少超出产品允差 5 倍)?

(2) 如果绿 Y 是一个属性, 是否可以通过李克特度量尺度把它转化为变量, 从而把缺陷用李克特度量尺度相乘, 使其达到一个较大又可权衡缺陷的依据?

(3) 可用产品/过程搜索作为逻辑 DOE 技术来产生线索是时间对时间的变量族吗?

(4) 是否绘制了一个以能影响最终产品性质的能力为基础的降序排列的过程参数清单? (过程参数如果没有显示变化, 就无须继续监测。)

(5) 是否建立了一个能精确地测量每个过程参数的仪器系统?

(6) 所监测的参数是经过实际测量的而非仅仅是设置的吗?

(7) 产品/过程搜索的应用是在缺陷的高发阶段而非缺陷率为最小或等于 0 时进行的吗?

(8) 所完成的最好产品和最差产品的分布, 能够包括 80% 最终产品的最有意义的变量吗?

(9) 如果不能对每个部件(产品)的相关过程参数进行监测, 接下来能有最小批量或时间采样用于获得 8 个最好的批次和 8 个最差的批次(或时间采样)吗?

(10) 如果过程参数显示出好部件分布于中间, 差部件在尾部, 那么能用辅助图基检验方法来确定终结计数吗?

(11) 一旦鉴别出重要的过程参数, 还要进行“B vs. C”试验, 以验证其改善程度吗?

(12) 是否应用后续试验, 如散布图、响应曲面方法等来优化重要的过程参数?

(13) 进行过深入的试验, 以观察究竟不重要的参数的允差可以扩展到什么程度来达到降低成本的目的吗?

(14) 使用过正向控制、过程确认和预先控制等技术, 以保证在每天的生产活动中有持续的改善吗?

在后续 DOE 小组工作中的高层管理的问题

(1) 为什么选定这个特定的绿 Y? 可以再选一个更早、更容易的绿 Y 吗? 绿 Y 的测量精度能保证 5:1 吗? 对于属性绿 Y 是否建立了李克特度量尺度?

(2) 为什么选用产品/过程搜索作为线索生成技术?

(3) 在过程进行中, 可以针对每个产品的部件测量其过程参数(是实际值而非设定值)吗? 如果不能, 什么是可以区分 8 个最好和 8 个最差批次或时间采样的最小批次或时间采样?

(4) 在产品/过程搜索之后, 进行过“B vs. C”试验以验证过



程的改善吗？

(5) 利用过其他一些 DOE 技术（如散布图、正向控制法）去优化过程参数，然后再将其“冻结”于进行中的生产过程里吗？

(6) 是否应用过程确认技术以保证所有质量外围条件，例如在生产开始前（甚至在 DOE 开始前）的条件包括制造经验、标准化、环境因子等？

(7) 预控制是监控紧密控制的重要过程参数的最好的 SPC（统计过程控制）工具吗？



第 4 篇

表示产品 / 过程特性试验技术有效设计



第 13 章

变量搜索法：罗尔斯·罗伊斯 (Rolls Royce) 对红 X 的研究

从思路创新到红 X 的提取

在上一章我们阐述了 4 种有可能与部件对话的线索生成技术——多变量法、部件搜索法、成对比较法和产品/过程搜索法。作者以及我们的很多委托人，都有这样的经验：70%~80% 长期存在的质量问题可以按顺序采用或同时采用一种或多种这些线索生成技术方法去解决。这些方法还有不影响生产的独特优势。

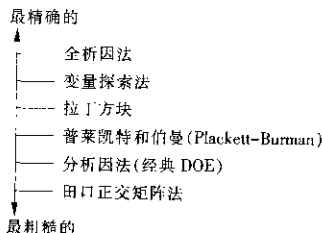
然而，在某些时候，当通过思索滤去很多非重要的变量后，还遗留下 2~10 个甚至 15 个重要变量，需要进一步提取出红 X、粉红 X 和浅粉红 X，同时量化其交互影响效



应，所以需要继续采用更多的有效的 DOE 技术，即使的确要中断生产。

有效的 DOE 技术

目前，正流行几种有效的 DOE 技术。最常用的一些是（按照精确度而不是实用性排序）：



在这些方法中，全析因法是最完美的，因为它精确地将主要效应从其他交互影响的效应中分离出来——将所有的二阶、三阶、四阶甚至更高阶的交互影响效应分隔开。那么，试验人员为什么不只是使用全析因法，而忘记不那么纯正的 DOE 技术呢？答案就在于实用性。

全析因法

如我们将在第 14 章中看到的，设计全析因法的方法是为每个因子选取两个水平——高水平和低水平，然后每个因子的每个水平都与其他所有因子的每个水平搭配进行试验，于是，创建了 2^n 矩

阵或 2^n 个组合（即试验）。如果有 4 个因子，而每个因子各有 2 个水平，则有 2^4 或 16 个组合，这意味着要做 16 次试验。这种情况是切实可行的。但如果有 10 个因子，各有 2 个水平，将会怎样呢？将会有 2^{10} 或 1 024 个组合（即 1 024 次试验）。等年轻的试验人员完成这些工作时，就已老态龙钟了。

所有类型的分析因法的局限性

作为对全析因法的简化，其他的有效的 DOE 技术只对大量组合中的一部分进行试验。最终是拉丁方块——一种简捷的 DOE 技术，在 20 世纪 60—70 年代与其作者一样有名，但随后就被放弃不用了，因为它通常要对三种因子进行试验，每个因子有两个水平，且只做 4 次试验来替代所要求的 8 次试验。这种半因子法——或半重复试验法是分析因法系类中受污染最少的。

田口方法

另外受污染最重的是田口正交矩阵法，它有多达 10 个因子需要进行试验，比如说一个 L_8 矩阵或一个 L_{16} 矩阵，即只需进行 8 个或 16 个试验。正如我们将在下一节中阐明的，这意味着主效应和交互影响效应可能受到严重污染——统计学上的术语为混淆，将带来临界的或直接的错误结果。田口方法的信仰者反驳说，他们已将那些最重要的交互影响效应包括在适当的“ L ”矩阵内了。然而，他们如何在未知的变量中筛选出最重要的组合呢？这很像玩俄罗斯轮盘，实际上，甚至比玩俄罗斯轮盘的情况更糟，因为俄罗斯轮盘的死亡机率只有 $1/6$ ，而田口方法却时时都有死亡机会。

经典分析因法——仅稍好于田口法

非常信赖分析因法的经典 DOE 学派采用比田口正交矩阵的组合（或所需试验）总数中更高的比例，所以其污染——或混淆程度要稍好些，但是它在试验设计本身方面仍有根本的缺陷。在这里，

经典学派的试验者也反驳说，他们采用分析因法作为一种“筛选”实践，至少可以确定出主要效应和某些二阶交互影响效应，此后可以用全析因法区分出真正的变量。然而，正如我们下一节要说明的，如果试验的基本设计存在缺陷，那么其结果同样会存在缺陷。建立在沙堆上的房子终究会倒塌。

►主效应与交互影响效应混淆的案例

假定有 4 个因子：A，B，C 和 D，要求对它们的主效应和真正的交互影响效应进行量化。如用全析因法——这是最完善的技术，将要求 16 次试验（见表 13—1）；而使用平衡设计则要求 15 个对照物——或 15 列（规则为对照物的数量等于试验数减 1）。

- 有 4 列主效应：A，B，C 和 D。
- 有 6 列两因子交互影响：AB，AC，AD，BC，BD 和 CD。
- 有 4 列三因子交互影响：ABC，ABD，ACD 和 BCD。
- 有 1 列四因子交互影响：ABCD。

现在我们假定：决定采用半重复的分析因法（该法在分析因法族系中最少受到反对）以代替全析因法。表 13—2 所示为一个典型例子。

半重复分析因法的缺点

(1) 只有一半的数量，即做 8 次试验代替在用全析因法时的 16 次试验。

(2) 设计中包括 4 个主要因子（A，B，C 和 D），每个因子有 2 个水平——高（+）和低（-）——所以每个因子都要做 4 次高水平 和 4 次低水平的试验。

(3) 但对于 8 次试验而言，在设计中只有 7 个对比的机会。A，B，C 和 D 各占用 4 列，只给 3 个交互影响留下机会，如 AC，BC

表 13-1

2⁴ 全析因设计

单元	因子的主效应				两因子的交互影响				三因子的交互影响				四因子的交互影响				结果/输出
组	A	B	C	D	AB	AC	BC	AD	BD	CD	ABC	ABD	ACD	BCD	ABCD	输出	
1(1)	-	-	-	-	+	+	+	+	+	+	-	-	-	-	-	+	
2a	+	-	-	-	-	-	+	-	+	+	+	+	+	+	-	-	
3b	-	+	-	-	-	+	-	-	-	+	+	+	-	+	+	-	
4ab	+	+	-	-	+	-	-	+	-	+	-	-	+	+	+	+	
5c	-	+	+	-	-	+	-	+	+	-	+	-	+	+	-	-	
6ac	+	-	+	-	-	+	-	-	+	+	-	+	-	+	+	+	
7bc	-	+	+	-	+	-	+	+	-	-	+	+	-	-	+	+	
8abc	+	+	+	-	+	+	+	-	-	-	+	-	+	-	-	+	
9d	-	-	-	+	-	+	+	-	-	-	-	+	-	+	-	-	
10ad	+	-	-	+	-	-	+	+	-	-	+	-	+	-	+	+	
11bd	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+	+	
12abd	+	+	+	+	-	-	-	+	+	-	-	+	-	-	-	-	
13cd	-	-	+	+	+	-	-	-	-	+	+	-	-	-	-	-	
14acd	+	-	+	+	-	+	-	+	-	+	-	+	+	+	-	-	
15bcd	-	+	+	+	-	+	+	-	+	+	-	-	-	-	-	-	
16abcd	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	

注: (1) 单元

单元(1)是指 A, B, C 和 D 均处于()水平。

较低状态 A(a)表示 A 处于正(+)水平; B, C 和 D 同理。

较低状态 A(kab)表示 A 和 B 均处于正(+)水平。

(2) 交互影响。

A, B 的正负号等于代数相乘的结果符号。

● 如果 A 为(-), B 为(-), 则 AB 为(+), 即(-) × (-) = (+)。

● 如果 A 为(-), B 为(-), C 为(-), 则 ABC 为(-), 即(-) × (-) × (-) = (-)。

(见第 14 章的全析因的说明。)

表 13—2

半重复的分析因法设计

试验号	A	B	C	D	AC	BC	ABC	结果
1	+	+	+	+	+	+	+	
2	+		+	-	+	-	-	
3	-	+	-	-	+	-	+	
4	-	-	+	+	-	-	+	
5	+	+	-	-	-	-	-	
6	-	-	-	+	+	+	-	
7	+	+	+	-	+	+	+	
8	-	-	-	-	+	+	-	

及 ABC。对于下列情况该怎么办呢？

● 遗留下两因子的交互影响 (AB, AD, BD 和 CD) 该怎么办？

● 遗留下三因子的交互影响 (ACD, ABD 和 BCD) 该怎么办？

● 四因子的交互影响 (ABCD) 该怎么办？

(4) 十分现实的是：15 个交互影响中有 8 个没有被包含在试验中，而它们仍可以找到影响 4 个主要因子之一（或更多）或 3 个被选中的交互影响的路径，从而完全搞乱了结果。

(5) 因为高 (+) 和低 (-) 的数量不相等，它们之间的交互影响可能不平衡（见表 13—2）。

(6) 最后的结果是混淆的，会导致计划落空，幻想破灭。

统计学家谈到假信号，这意味着假设或错误的名称。分析因法充满了假信号。例如：有 7 个因子，各有 2 个水平。全析因法需要进行 27 或 128 次试验。如果以分析因法取代，只进行 8 次试验，其结果是非常混淆，请参阅表 13—3。这表明：每个主效应可能被 15 个交互影响效应搞乱，总共产生 112 个假信号。

表 13—3 对 7 个因子仅做 8 个试验所产生的假信号

A	-	BD=CE=BCG=CDF=BFF=FG=DEG
	=	BCDEFG=ACEFG=ABDFG=ABCF
	=	ADEF=ACDG=ABEG=ABCF
B	-	AD=EG=ACG=CDE=AEF=DFG
	=	ACDEFG=BCFG=ABDEG=ABCDF
	=	BDEF=BCDG=ABFG=ABCE
C	-	AE=BF=ABG=BDE=ADF=DG=DFG
	=	ABDEFG=BCDFG=ACDEG=ABCEG
	=	CDEF=BCFG=ACFG=ABCD
D	=	AB=CG=BCE=ACF=EF=AEG=BFG
	=	ABCEFG=CDEFG=ABDEF=ABCDG
	=	BDEG=BCDF=ADFG=ACDE
E	-	AC=BG=BCD=DF=ABF=ADG=CFG
	=	ABCDGF=BDEFG=ACDEF=ABCEG
	=	CDEG=BCFG=AEFG=ABDE
F	=	BC=AG=DF=ACD=ABE=BDG=CEG
	=	ABCDEG=BCDEF=ADEFG=ABCFG
	=	CDFG=BEFG=ACEF=ABDF
G	=	ABC=CD=BE=AF=ADE=BDF=CEF
	=	ABCDEF=BCDEG=ACDFG=ABEFG
	=	DEFG=BCFG=ACEG=ABDG
	112	假信号
	15	永远丢失
	127	全部对比

这可以和一位芝加哥女人的故事相提并论——她为了从山姆大叔那里搜集到福利支票，用了 100 个假名字，一直到被捕入狱。她被称为假福利皇后艾拉斯。在统计学家的大量言词中，分析因法也可被称为 DOE 的艾拉斯皇帝。

变量探索法：优于所有分析因法

如果因子数量超过 4，全析因法是太花时间、太复杂了，且如果整个分析因法族在统计学方面又太弱了，那么用什么方法可以代替呢？变量搜索法正好满足了这方面的要求。它有试验数量较少的

优势 甚至对于 10 个因子最多 26 次，同时它可对所有的主效应和全部二阶、三阶和大多数四阶的交互影响效应进行简练的分离。这些将在本章后面的部分说明。此外，通过因子从 3:1 到 10:1，它比经典法和田口法更节省试验费用。其在解决问题的路径图中的位置见图 7—3。

变量搜索法的目标

(1) 从较早的线索生成 DOE 技术中，减少大量的遗留原因——高达实际上最大的 20，并提炼出红 X、粉红 X、浅粉红 X（前面述及的原因、变量及因子，都是同义术语）。

(2) 将重要的因子从不重要的因子中分离出来。

(3) 放宽非重要因子的误差范围（即这些因子的 C_p 和 C_{pk} 值都能很容易地减少到 1.0 以下），并节省许多费用。

(4) 对重要因子及它们的交互影响效应的大小和要求的方向进行量化，并将那些重要变量的误差范围压缩到最小，使 C_p 、 C_{pk} 值为 2.0。

(5) 当变量搜索法能够且已经用作解决生产问题的工具，并作为 4 种线索生成工具的后续工具时，它的必不可少的功能是在产品或过程设计阶段中，把预防影响生产的问题放在首位。当在生产中发现质量问题时，就已经太晚了，而到了用户手中才发现产品质量问题就更晚了。实际上，生产中存在问题，必然是由于对设计功能缺乏考虑造成的。

(6) 关于在概念设计阶段确定重要因子的文献有很多，并且还有许多计算机软件工具。它们包括蒙特卡罗仿真法、E—芯片技术和电路分析法以及其他方法。但这些方法普遍都存在一个弱点：需要知道决定自变因子与因变量（绿 Y）之间关系的数学方程式或结构式；如果不知道这些公式，那么，计算机和论文研究至多不过是猜谜游戏，甚至毫无用处。在这种情况下，变量搜索是惟一的答案。

(7) 甚至在已知自变量和因变量之间关系的方程式时，进一步在模样阶段用变量搜索法进行计算机仿真，可对方程式的精确度进行必要的改进（计算机建模是在方程式的基础上进行的）。根据作者的经验，计算机鉴别造成错误确认重要变量的可能性至少为20%。

(8) 对于标准化产品，变量搜索法也可用于放宽误差范围和降低费用。

二元搜索法的原理

变量搜索法是建立在二元搜索法的基础上的。用“20问”游戏来解释最好不过了。让某人挑选一个字，字的含义可以在一本800页的字典中查到。然后，他写下这个字，并记下字典上出现这个字的页数。在游戏中，提问者通过提出不超过20个的问题来找到那个字，并且问题的答案只能为“是”或“否”。让我们假定：那个人在第272页上挑选了一个字。提问的问题如表13—4所示。

表 13—4 二元搜索法原则：20 问

序号	问题	答案
1	该页数比 400 大吗？	否
2	比 200 页大吗？	是
3	比 300 页大吗？	否
4	比 250 页大吗？	是
5	比 290 页大吗？	否
6	比 275 页大吗？	否
7	比 263 页大吗？	是
8	比 269 页大吗？	是
9	比 272 页大吗？	否
10	比 271 页大吗？	是
11	是否是 272 页？	是



通常，提问不超过 11 个问题就可以确定那个页数。然后，提问者问，所选的字是否在那一页的左边？如果答案为否，它即是在右边。下一个问题：那个字是否在右边的上半部分？如果是，再问它是否在右上 1/4 处，依此类推，直至在少于 20 问的情况下找到那个字。

这个游戏表明了让一位工程师从超过 10 万字的字典中猜测出那个字的差异。那样的话他要为此永远难眠！而二元搜索法是一个系统的消除过程。当回答第一个问题后，400 个页码就被排除了；在第二个问题后，600 个页码被排除了。依此类推，直到第 11 个问题，当 799 个页码被排时，红 X——即第 272 页就被找到了。

这个二元搜索法原理的改进形式被应用于变量搜索法中。变量搜索方法与部件搜索法相似，部件搜索法用于处理离散的部件、部件或组件的质量好坏，而变量搜索法常用于处理连续变化的过程或材料参数。

变量搜索的方法

与部件搜索法相似，变量搜索法分 4 个阶段（见表 13—5），每个阶段都有一个名称和目标陈述。

表 13—5 变量搜索的阶段

阶段	目标
1. 球场	决定试验用的每一个变量是否是正确的变量、处于正确的水平
2. 分离重要的和非重要的因子	● 将重要变量从非重要变量中分离出来 ● 消除非重要变量及其相关的交互影响
3. 求交运算(确认)	验证重要变量是重要的，非重要变量是非重要的
4. 析因分析	对重要变量及其有关的交互影响的大小和期望水平进行量化

第 1 阶段 球场

(1) 绿 Y。在初始阶段（球场），试验者必须要首先决定绿 Y

(即问题、输出或响应)。在某些调查中,可能有不止一个绿 Y。作者在变量搜索试验中最多曾用到 7 个绿 Y。绿 Y 必须是定量的和可测量的,同时,测量仪表的精度至少要高出绿 Y 技术规定允差的 5 倍。如果绿 Y 是一个属性,试验人员应采用李克特度量尺度将其转换为一个变量。这种转换减少了将狭窄缺陷范围扩大为加权缺陷范围所需的样本量(见第 7 章)。应尽各种努力使较早的绿 Y 在过程中定位,这个绿 Y 应易于测量,并且很可能与最终的绿 Y 有合理的关联。

(2) 试验中变量的选择。这是第 1 阶段的重要部分。列出一张对于因子 A, B, C, D, E, F, G, H 等最重要的输入变量清单,并根据重要性递减排序。(这样排序的理由是:如果这是合理而正确的,就可大大缩短试验。)

进行这种选择的最佳时间是在实施一种或多种的 4 种线索生成技术以后,拥有从大量未知的、难以处理的因子清单缩小而成的一张只有较少的、相对好处理的因子族表格。例如,终结计数在成对比较法或产品/过程搜索中是一个确定重要性的实用方法——终结计数愈大,重要性可能也越大。

然而,若产品或过程是新的,尚处于模型阶段时,没有可能采用多变量法、部件搜索法、成对比较法或产品/过程搜索等方法,那么:

- 利用计算机仿真或电路分析以选择可能重要的因子。(然而,必须承认:对于很多复杂的产品或过程,决定自变量和因变量之间关系的方程式是未知的。在这种情况下,大型计算机也只能起到猜测的作用!即使已知这个方程式,人们也总是愿意用硬件变量搜索试验来确认计算机的仿真结果。)

- 组织一次头脑风暴研讨会议,征求工程界、操作人员、维修人员和供应商们的不同意见,绘制一张列有可能的因子的清单。然而,这是一个很少有成果的方法,因为它是建立在直觉、猜测、主观和理论的基础上的。

(3) 每个变量的最佳水平和临界水平。在选定试验用因子后,



对每个因子赋予两个水平：最佳水平（B），它有可能成为最佳的绿 Y（有希望比现有的绿 Y 好）；临界水平（M），它是在正常维护条件下，在每天的生产中，表现出与最佳水平有一定的偏移 [有时把符号（+）赋予每个最佳水平，而把符号（-）赋予每个临界水平]。如果是在模型阶段实施变量搜索法，最佳水平通常是该因子的设计中心或日标值，临界水平是对该因子能偏离（双边）最佳水平多远的工程判断，以记录下大的、在实际极限内能重复的差异。如果试验人员不能确定哪个水平（最佳的或临界的）较好时，他们可以随意指定最佳水平和临界水平，变量搜索试验将在随后阶段中确定哪个水平较好。

第 1、2、3 步为变量搜索试验的设计。下一步就进入了试验的实施。

(4) 样本量。与在部件搜索法中一样，样本量难以置信地小——指定两个：一个是“BOB”，而另一个是“WOW”。变量搜索法的样本量也同样难以置信地小。表 13—6 给出了一些准则。

表 13—6 变量探索法的样本量

绿 Y	因子	条件	样本量
变量	部件或过程参数	部件能插入和重新插入而不干扰绿 Y	1
用利克特度量尺度将属性转换为变量	部件或过程参数	同上	5~10
属性 (即：无缺陷%，缺陷%，产量等)	部件或过程参数	同上	16~502

注：如果某些部件不能在第 1 阶段或第 2 阶段中重复使用，应重新选用一组新的具有相同数量的部件。

- 如果缺陷百分比比较高，较小的样本量也能检测出差异。
- 如果缺陷百分比比较低，则要求较大的样本量。
- 如果缺陷为 100ppm 或更低，而且不能转为利克特度量尺度，变量搜索法也不适用，应采用成对比较法或产品/过程搜索。

(5) 早期和快速评估发展趋向。实施两组试验，第一组采用全部处于最佳水平的因子，第二组采用全部处于临界水平的因子。



(a) 如果因子的全部最佳组合和全部临界组合的绿 Y 之间似乎存在巨大的差异, 这是一个早期征兆, 表明在因子表中已采集到正确的因子, 可以继续进行下一步, 见下面第 6 步。

(b) 如果全部最佳水平都比全部临界水平好, 但它们之间似乎只存在微小的差距, 这说明没有收集到正确的因子, 或这些因子的正确水平, 或红 X 被强大的粉红 X 抵消, 或红 X 是在偶数因子组中的交互影响。

(c) 如果全部最佳水平显示比全部临界水平有所改进, 但其差异没有比绿 Y 的过去水平高很多, 这说明在收集正确因子或所选定因子的正确水平方面, 还做得不够。

- 如果第 5 步的结果是 (b) 或 (c), 对于变量搜索法, 仅在两组试验中失利, 而对于经典法或田口试验, 你不知道你是否成功了, 直到完成 16 组、32 组或 64 组试验才算结束。

- 如果第 5 步的结果是 (b) 或 (c), 第 2 步的因子表必须重新评估。首先, 试用某些因子的不同水平; 或者取消某些因子, 并插入一些新的因子。

- 如果第 5 步的结果是 (a), 则继续进行第 6 步。

(6) 重复。用所有因子中两个以上的全部最佳水平 and 所有因子中两个以上的全部临界水平重复第 5 步试验。这意味着: 现在总共用三个全部最佳和三个全部临界的水平。然而, 必须注意: 实施三个最佳和三个临界水平的程序 (或顺序) 是随机的。否则, 有可能导致读数的偏离 (见第 15 章)。

(7) 重要性试验。为了确定 3 个全部最佳的绿 Y 是否确实比 3 个全部临界的绿 Y 好, 要进行两组重要性试验。

(a) 所有 3 个全部最佳的绿 Y 应该比所有 3 个全部临界的绿 Y 好, 而且没有重叠 (如果 3 个全部临界的绿 Y 反而比 3 个全部最佳要好——这是经常发生的! 那么, 只要将全部最佳水平的上部换成全部临界就可以, 反过来也是一样)。

(b) $D:\bar{d}$ 的比率应大于或等于 1.25 : 1 (与部件搜索中的规则相同), 这里

- D 是最佳和临界绿 Y 中值之差。
- $\bar{d}$ 是全部最佳的绿 Y 和全部临界的绿 Y 间差值（或范围）

的平均值（即每个都缺少可重复性）。

(8) 重要性试验得到满足。如果这些重要性试验获得通过，第 1 阶段就告结束；正确的因子已收集到，虽然红 X ，粉红 X 等尚未精确定位（这要在第二阶段中明确）。

(9) 重要性试验没有得到满足。如果两组重要性试验都未通过，将一对最有可能是因子的最佳水平转换成临界水平（反之亦然）以观察是否消除了影响。如果仍无显著差别，将第 2 对因子进行转换。除非最佳工程判断和临界水平相反，否则，几乎没必要这样做。

如果可重复性 $\bar{d}$ 很差，这就表明在第 2 步中，一个重要因子——可能是红 X ——被遗漏在表外，这时要检查思路创新过程。粗心大意的思路创新常常是导致第一阶段失败的原因。寻找更好的思路——在表中增加一个或两个以上的因子，并重复运行第 1 阶段。

第 2 阶段 分离重要的和非重要的因子

(1) 实施一对试验。利用最重要因子 A 的临界水平即 A_M ，以及所有余下因子的最佳水平——称为 R_B 。然后进行镜像试验：利用因子 A 的最佳水平即 A_B 以及所有余下的因子的临界水平——称为 R_M 。使用与部件搜索中同样的公式计算高限和低限的判断极限。

判断极限为：中值 $\pm 2.776 (\bar{d}/1.81)$ 。（这与部件搜索法的判断极限相似。）

(2) 可能的结果

(a) 如果这对试验中的两组—— $A_M R_B$ 和 $A_B R_M$ ——显示的结果都在判断极限的低限和高限内，因子 A 以及所有与它有关的交互影响效应都是非重要的，可以在进一步的研究中排除。

(b) 如果完全颠倒过来，即如果 $A_M R_B$ 成为初始全部最佳水平，而 $A_B R_M$ 成为初始全部临界水平，那么， A 为惟一的红 X ，其余的因

子 B, C, D 等都是非重要的, 并可以排除。变量搜索到此结束。

(c) 如果这对试验—— $A_M R_B$ 和 $A_B R_M$ ——的显示结果都超出各自判断极限的低限和高限, 但并未完全颠倒, 因子 A 以及与之有关的交互影响效应不能被排除, 必须考虑增加一些其他因子及其交互影响。

(3) 如果第 2 步的试验结果为 (a) 或 (c), 用因子 B 重复第 1 步。如果结果是 (a), 那么 B 是非重要的并可排除。如果结果是 (b), 那么 B 是红 X, 结束变量搜索。如果结果是 (c), 那么 B 不能被排除, 必须考虑加上 B 和其他一些因子及其交互影响。

(4) 如果第 3 步的结果是 (a), 用因子 C 重复第 1 步。如果结果是 (a), 那么 C 是非重要的并可排除。如果结果是 (b), 那么 C 是红 X, 变量搜索结束。如果结果是 (c), 那么 C 不能被排除, 必须考虑加上 C 和其他一些因子及其交互影响。

第 3 阶段 求交运算

(1) 如果因子 A 和 B 显示出部分颠倒 [结果 c], 且读数超过判断极限, 用这些因子—— $A_B B_B R_M$ 和 $A_M B_M R_B$ ——进行求交运算 (即确认试验或验证试验), 以观察余下的因子 R 是否可排除。如果一个或更多的结果超出了判断极限, 这将显示搜索还不完整。在此情况下, 用下一个的单个因子——C, 然后是 D, E 等——按字母顺序继续实施第 2 阶段, 直至另外的因子显示出超越控制的结果。

(2) 实施 3 因子的求交运算 (一般很少需要 4 因子求交运算)。

第 4 阶段 析因分析

这不是物理试验, 而只是利用在第 1、2、3 阶段产生的、在第 2 阶段又不能排除的重要因子的数据来拟定析因分析的数据整理。这与在部件搜索法中的案例研究所述的析因分析技术相似, 是对重要因子的主效应和交互影响效应进行定量分析。

评估第 2 阶段因子重要性的备选经验法则

这个经验法则与在部件搜索法中所述的相仿。很多专业人



员发现与确定判断极限有关的统计计算工作是冗长而混乱的。本书的作者们开发了同样的经验方法以评估第2阶段的因子是否重要。它虽不如判断极限一样有严格的统计学理论依据，但也能用作近似方法。

(1) 首先，在第1阶段中，取最佳中值和临界中值。

(2) 计算这两个中值之间的平均值，以作为中心线。

(3) 在第2阶段中，如果因子穿过这条中心线或非常接近穿过它（在10%以内），则因子就是重要的。反之，它就不是重要的。这个交叉方法将在变量搜索的案例研究中加以叙述。

变量搜索法的交互影响效应

怀疑变量搜索法的人抱怨此方法只是一个老方法的陈旧版本：每次只变化一个变量，而保持其他一切不变，是不是有点太离谱？但变量搜索确定是一个巧妙的方法，其中：

(1) 可以证明所有非重要的主要因子及其相关的交互影响效应是非重要的（详见第一个案例）。

(2) 所有重要的主要因子以及它们的：

(a) 二阶交互影响效应被简洁地分离了；

(b) 三阶交互影响效应被简洁地分离了；

(c) 62.5%的全部四阶交互影响效应被简洁地分离了。如果一个全因子能够仿效，其余37.5%的四阶交互影响效应也能被分离。

(3) 五阶及更高阶的交互影响效应可能会被混淆，但是五阶及更高阶的交互影响效应只可能在10 000次实际试验中出现一次。

上述说法的证据如下：

(a) 2个因子A和B，每个都有2个水平，能产生4个组合，如图13—1的4个角所示。变量搜索法能否覆盖全部4个角（交互影响）？答案是：能。



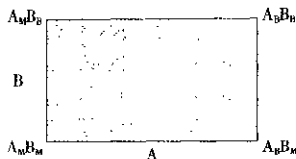


图 13—1 两个因子的交互影响效应：示意图

- $A_M B_M$ 和 $A_B B_B$ 是在第 2 阶段的首批两组试验中测量的。
- 如果 A 和 B 是非重要的，在图 13—1 上的图示曲线将为两条平行线，表示交互影响效应为零（或近乎零）。
- 如果 A 和 B 是不重要的，图 13—1 的图示曲线将是两条不平行线，表示有交互影响存在。
- 对于其他因子，每次两个因子，都能得到相同结论。

(b) 3 个因子 A、B 和 C，各有 2 个水平，能够产生 8 个组合，如图 13—2 中的 8 个角所示。变量搜索能否覆盖全部 8 个角（交互影响）？答案是：能。

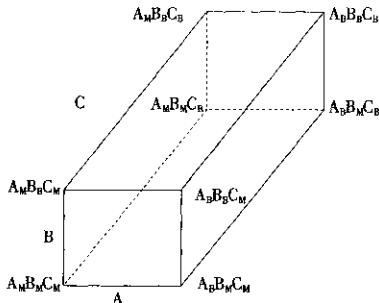


图 13—2 三因子交互影响效应：示意图

- $A_B B_B C_B$ 和 $A_M B_M C_M$ 在第 1 阶段测量。



● $A_M B_B C_B$, $A_M B_M C_M$, $A_M B_B C_M$, $A_B B_M C_B$, $A_B B_M C_M$ 和 $A_B B_B C_M$ 全部都在第 2 阶段首批 6 组试验中测量。

● 如果 A 和 B 是不重要的, 其 3 相图形曲面 (难以描绘) 将是平行的平面, 表明交互影响为零或趋于零。

● 如果 A, B 和 C 是重要的, 其 3 相图形曲面将是不平行 (扭曲) 的平面, 这时存在交互影响。

● 对于其他因子 (每次 3 个), 都能得到相似结论。

(c) 4 个因子, A, B, C 和 D, 各有 2 个水平, 能产生 16 个组合, 如表 13-7 所示。

表 13-7 四因子交互影响效应

单元	组合	
(1)	$A_M B_M C_M D_M$	● 变量搜索在第 1 阶段测量单元 (1) 和 abcd。
a	$A_M B_B C_B D_B$	● 在第 2 阶段它 (变量搜索) 测量单元 a、b、c、abc、d、abd、acd 和 bcd。
b	$A_B B_M C_B D_B$	
ab	$A_M B_M C_B D_B$	● 在第 1 阶段或第 2 阶段不测量单元 ab、ac、bc、ad、bd 和 cd。
c	$A_B B_B C_M D_B$	
ac	$A_M B_B C_M D_B$	● 因此, 如果在未经测量的 6 个单元或组合中交互影响效应是可疑的, 第 3 阶段——求交运算可能不会成功。
bc	$A_B B_M C_M D_B$	
abc	$A_M B_M C_M D_B$	
d	$A_B B_B C_B D_M$	
ad	$A_M B_B C_B D_M$	
bd	$A_B B_M C_B D_M$	● 在此情况下, 在 4 个因子中间安排全因子试验。
abd	$A_M B_M C_B D_M$	
cd	$A_B B_B C_M D_M$	● 然而, 如果第 3 阶段是成功的, 并且第 1 阶段的结果很接近地在第三阶段重复出现, 那么在未经测量的 6 个单元中的交互影响效应可以被认为是不重要的。
acd	$A_M B_B C_M D_M$	
bcd	$A_B B_M C_M D_M$	
abcd	$A_M B_M C_M D_M$	

注: 下标 B 表示最佳; 下标 M 表示临界。



产品/过程的特征

正如在本章前面所述,虽然变量搜索法是跟在一个或更多的4个线索生成工具后继续使用的强有力的DOE技术,但其最大的用途是作为开发模型阶段中的产品或过程的设计技术。

近年来,术语“产品/过程特征”已用于描述设计意图。这意味着:

- 将重要变量(或因子)从非重要的变量中分离出来;
- 收紧重要变量的公差范围,使 C_{pk} 等于或大于2;
- 扩大非重要变量的公差范围(有可能还要求做一些试验,以观察公差范围可放宽到多大),以明显地降低费用。

达到这些目的的最佳方法就是变量搜索。

案例研究与小组练习

案例研究 1: 冲压成型

在进行金属冲压/成型工序时,冲压成型的零件不能保证在0.005英寸的公差范围内或0.010英寸的加工宽度。加工尺寸变量高达0.010英寸(或 C_{pk} 值仅为0.5)。工长坚持认为:这个冲压成型是一个“黑色魔法”过程。他抱怨供应商提供的金属厚度和硬度变量太大,还抱怨操作人员不稳定、粗心大意,同时他曾试图采用较新的冲压成型机,但收效甚微。

有人向公司管理层推荐变量搜索法(时间的浪费已使工长烦透了)。目标设定为通过控制,使公差范围缩小到0.005英寸或更小。选择了6个被认为有重要影响的因子按降序排列,确定了每个因子的最佳和临界水平(出于保密的原因,确切的水平未给出)。表13-8列出了第1阶段和第2阶段的结果,其数字表示偏离正常值的量(用0.001英寸的倍数),数字低的为较好,数字高的为较差。



第 1 阶段 球场

3 个全部最佳水平和 3 个全部临界水平（按随机顺序实施）表明：第一，3 个全部最佳的确比 3 个全部临界的好（偏离较小）；第二，最佳中值和临界中值相应为 4 和 61， $D/\bar{d}$ 比率为 57 : 11 或 5.21，比最低要求 1.25 大。因而第 1 阶段的结果表明：A 到 F 都被选中为合适的因子。

表 13—8 变量搜索：冲压成型案例研究

因子	最佳	临界
A. 冲头与凹模调整	调整	没有调整
B. 金属板材厚度	厚	薄
C. 金属板材硬度	硬	软
D. 金属板材弯度	平整	弯曲
E. 冲床冲程	精整成形	气压成形
F. 材料夹持	水平	有角度

注：以下结果表明偏离正常值的量（为 0.001 英寸的倍数）。

第 1 阶段	全部最佳水平	全部临界水平
初始	4	47
第一次重复	4	61
第二次重复	3	68

第 2 阶段

试验	组合	结果	中间值	判断极限	结论
1	A _M R _B	3	4	-12.87~20.87	A 不是重要的
2	A _B R _M	72	61	44.13~77.87	
3	B _M R _B	5	4	-12.87~20.87	B 不是重要的
4	B _B R _M	47	61	44.13~77.87	
5	C _M R _B	7	4	-12.87~20.87	C 不是重要的
6	C _B R _M	72	61	44.13~77.87	
7	D _M R _B	23	4	-12.87~20.87	D 与其他因子是重要的
8	D _B R _M	30	61	44.13~77.87	
9	E _M R _B	7	4	-12.87~20.87	E 不是重要的
10	E _B R _M	50	61	44.13~77.87	



11	$F_M R_B$	73	4	-12.87~20.87	F 与其他因子 是重要的
12	$F_B R_M$	18	61	44.13~77.87	
求交运算	$F_M F_M R_B$	4	4	-12.87~20.87	● DR 交互影响 是重要的
求交运算	$F_M F_M R_B$	70	61	44.13~77.87	● 全部颠倒 ● 试验结束

注：最佳和临界集的中值为 1 和 61，相应地， $D=61-4=57$ 。 d =每个集系的平均不可重复性 $=(1+21)/2=11$ 。所以 $D/d=57/11=5.21$ ，大于 1.25 ± 1 ，因而第 1 阶段是成功的。

控制极限—中值 $\pm 2.776d/1.81$ = 中值 ± 19.87

所以，控制下限 = 44.13~77.87，

控制上限 = -12.87~20.87。

资料来源：Qualitron Corp., Chicago, Ill.

第 2 阶段 消元法

其判断极限为：

(1) 临界中值— $-2.776d/1.81 = -12.87 \sim 20.87$ ；

(2) 最佳中值— $-2.776d/1.81 = 44.13 \sim 77.87$ 。

● 位于 3 的因子 A_M 没有超出它的临界控制极限，而位于 72 的因子 A_B 也没有超出它的最佳控制极限，因此，因子 A 是不重要的。

● 同样，B、C 和 E 也是不重要的。

● 位于 23 的 D_M 超出了它的 -12.87~20.87 的临界判断极限。

● 位于 30 的 D_B 也超出了它的 44.13~77.87 的最佳判断极限。

● 因此，D 是重要的，但在 23 及 30 处，它仅表示了部分而不是全部逆转，所以其他的因子必须包括在内，并进行研究。

● 同样，F 是重要的，但在 73 和 18 处，它虽表示出有较大部分但并不是全部逆转，所以 D 和 F 都是重要的，同时它们之间存在交互影响。

用取舍选择法判断极限

● 在第 1 阶段中，全部最佳水平和全部临界水平之间的中心线



是 $(4+61)/2=32.5$ 。经验法则表明：如果一个因子在第 2 阶段穿过中心线，则该因子是重要的；如果没有，则是不重要的。因子 A, B, C 和 E 没有通过其临界水平和最佳水平之间的中心线，因此其是不重要的。

● 因子 D 没有通过 $D_M(23)$ 的 32.5 中心线，但却经过 $D_B(30)$ 的其他途径通过中心线，所以 D 有一定的重要性（表明有一个单向的交互影响）。

● 因子 F 通过 $F_M(73)$ 的 32.5 中心线， $F_B(18)$ 也同样通过，所以 F 是很重要的。

● 因此，因子 F 和 D 在第 3 阶段的求交运算中应予以考虑。

第 3 阶段 求交运算

图 13—3 的求交运算表明：因子 D 和 F 处于相应的最佳水平和临界水平（当其他 4 个因子也处于相应的最佳水平和临界水平时），第 1 阶段的结果是很近似的。因此，现在可以确认 D 和 F 是重要的。

	D 最佳	D 临界	
F 最佳	4 3	23	→ 24.5
	4 5	18	
	3 7		
	4 7		
	中值=4	中值=20.5	
F 临界	73	47 72	→ 116.0
	30	61 47	
		68 72	
		70 50	
	中值=51.5	中值=64.5	
	72.0	55.5	85.0
			68.5

$$\text{主效应 } D = \frac{(20.5 + 64.5) - (4.0 + 51.5)}{2} = \frac{85 - 55.5}{2} = 14.75$$



$$\text{主效应 } F = \frac{(51.5+64.5)-(4.0+20.5)}{2} = \frac{(116-24.5)}{2} = 45.75$$

$$\text{DF 交互影响} = \frac{(20.5+51.5)-(4.0+64.5)}{2} = \frac{(72-68.5)}{2} = 1.75$$

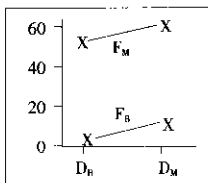


图 13—3 冲压成型的析因分析法

注：上述结果清楚地表明：由于 45.75 的主效应影响过程进展，所以红 X 是因子 F；由于 14.75 的主效应影响过程进展，所以因子 D 是粉红 X；同时，DF 交互影响的 1.75 对过程进展起作用。图形表明存在相对微弱的交互影响。

第 4 阶段 析因分析

（请参看第 10 章有关部件搜索的案例中所补充的析因分析的研究。）

变量搜索 DOE 试验可产生以下结论：

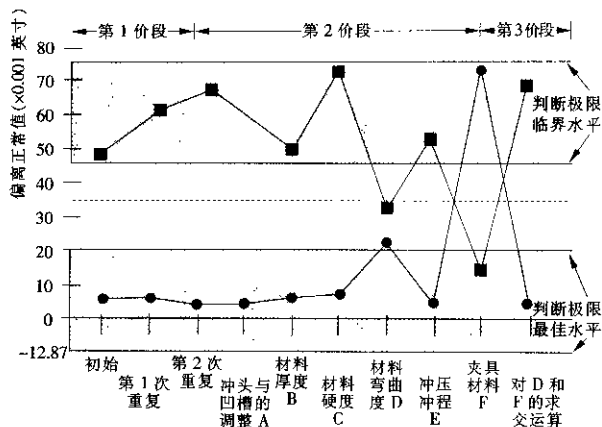
- （1）金属板材的弯曲度应控制平直（粉红 X）。
- （2）金属的厚度和硬度并不重要，因而对它们公差的范围可以放宽。冲头与凹模的调整和冲压过程也不重要。
- （3）板材的夹持（红 X）是非常重要的。在夹具设计时，要考虑保持板材平行，并且应排除与操作者有关的变量。
- （4）在采取这些措施后，冲压成形就能满足 ± 0.002 的公差要求，而不是原先的 $\pm 0.005 \sim \pm 0.010$ ，其相应的 C_{pk} 为 2.5。

案例研究 1 的图解曲线

如同部件搜索法一样，在变量搜索法中也可用单一、简明的图来解释上述文字。图 13—4 是表 13—8 所叙述的冲压成型案例研究中



第1、2和3阶段结果的图解。它清楚地表明：



图例：● 所有因子(除 A, B 等外)最佳, 在第 2, 3 阶段中处于其临界水平。

■ 所有因子(除 A, B 等外)临界, 在第 2, 3 阶段中处于其最佳水平。

图 13—4 变量搜索：冲压成型——图解曲线

● 因子 A, B, C 及 E 对最佳和临界水平的绿 Y 的逆转没有影响。这是由于这些因子的进一步变化都没有超出最佳判断极限和临界判断极限的范围。

● 因子 D 和 F 两者都显示出部分逆转 (F 比 D 更大), 两者都超出了最佳的和临界的判断极限。

● 取舍选择法判断极限显示出中心线没有通过因子 A, B, C 或 E, 这表明这些因子都不是重要的。但因子 F 通过中心线, 而因子 D 在其最佳水平时通过中心线。

● 所有这些都在图中十分清楚, 图解法明显地可以增强对如何去做的理解。

变量搜索试验中常犯的错误

变量搜索法若使用正确，就是强大的 DOE 工具。然而作者发现：变量搜索法常比应用其他 DOE 工具时所犯的错误要多。下面列出了尤其应防范的常见错误：

第 1 阶段

(1) 第 1 阶段完全被跳过去，而将第 2 阶段作为起点。

(2) 只进行一组全部最佳的试验和一组全部临界的试验，而不是三组都做。

(3) 如果全部最佳的绿 Y 并不如历史水平好，不要试图一切从头开始，无论是对同一因子的不同水平，还是不同的因子。

(4) 如果绿 Y 不只一个，不要试图分辨哪一个绿 Y 是最佳的，而其他都不是的情况。

(5) 三组全部最佳的试验和三组全部临界的试验的顺序不是随机的。

(6) 在决定一个因子的临界水平时存在不确定性。

(7) 即使 $D: \bar{d}$ 的比值小于 1.25，也执行第 2 阶段。

第 2 阶段

(8) 在决定重要因子和非重要因子时有些混乱。

(9) 判断极限由于太冗长而被跳过（此时，通过中心线的因子的取舍选择法可用作经验法则）。

第 3 阶段

(10) 未进行求交运算以鉴别并将重要因子从非重要因子中分离出来。



第4阶段

(11) 未进行析因分析计算以确定重要因子的主效应的和交互影响效应。

后续变量搜索

(12) 未将优化、正向控制和工艺定型等作为变量搜索的后处理方法（见第16章到第19章）。

小组练习1：发动机控制模块

某家大公司的工程部为一个重要的汽车客户开发一块基于微处理器的发动机控制模块（其电路上有600个电子部件）。此项目的开发十分复杂，模块要对客车发动机的20~25个参数实施监控并对发动机进行系统优化，以求油耗行驶里程最大化和污染最小化。

怠速行驶电流是客户要求的重要技术条件之一，它必须保持在650毫安（ma）~800毫安（ma）之间。据客户反映，如果怠速行驶电流低于650ma时，汽车可能失速或熄火；如果超过800ma，某些部件可能会烧毁。

由于该模块尚处于模型阶段，所以仅制造了少数几块模块，而且这些模块在怠速行驶电流方面有不可接受的变量，所以不能运用线索生成技术，工程小组选用了变量搜索法。他们在电路分析的基础上，根据对绿Y（被称为怠速电流）的影响，按重要性降序列出7个因子。每个因子的最佳水平都要求处在各部件或因子的设计值中心，临界水平被确定在每个部件公差范围的一端或另一端（无论哪个都被认为是较差的）。其结果如表13-9所示。

问题：

(1) 第1阶段是否成功？是否选中了合适的因子？ $D:\bar{d}$ 的比率是多少？试验能否进入第2阶段？



表 13—9

变量搜索：发动机控制模块

			因子水平		
因子说明	正常因子值	因子公差	最佳	临界	
A. 电阻 R85	0.68 Ohms	±5%	0.68 Ohms	0.65 Ohms	
B. 电源电压 V_{cc}	5.0 Volts	±5%	5.0 Volts	4.75 Volts	
C. 电阻 R77	100 Ohms	±1%	100 Ohms	99 Ohms	
D. 电阻 R75	787 Ohms	±1%	787 Ohms	729 Ohms	
E. X轴 QB 饱和电压	75M. V.	150 M. V. Max	75 M. V.	150 M. V.	
F. 电阻 R79	43 Ohms	±5%	43 Ohms	40.185Ohms	
G. 等成电路 IC4 补偿电压	0 M. V.	±8 M. V.	0 M. V.	-8 M. V.	
第 1 阶段		全部因子处于最佳水平		全部因子处于临界水平	
		742 ma		1 053 ma	
		738 ma		1 050 ma	
		725 ma		1 024 ma	
第 2 阶段	试验号	组合	结果	控制极限	结论
	1	$A_M R_B$	768		
	2	$A_B R_M$	1 020		
	3	$B_M R_B$	704		
	4	$B_B R_M$	1 051		
	5	$C_M R_B$	733		
	6	$C_B R_M$	1 028		
	7	$D_M R_B$	745		
	8	$D_B R_M$	1 018		
	9	$E_M R_B$	726		
	10	$E_B R_M$	1 022		
	11	$F_M R_B$	733		
	12	$F_B R_M$	1 020		
	13	$G_M R_B$	1 031		
	14	$G_B R_M$	718		

资料来源：Motorola Inc., Schaumburg, Ill.

(2) 在第 2 阶段，确定重要的和非重要的因子。最佳水平和临界水平的判断极限是多少？它们是否具有显著的交互影响效应？



- (3) 将第 1 阶段和第 2 阶段的结果绘制成图解曲线。
- (4) 是否需要进行求交运算和析因分析?
- (5) 非重要因子的公差范围将怎样扩展?
- (6) 如果在筛选因子时, 重要性顺序发生逆转, 将会发生什么情况?

答案

(1) 第 1 阶段是成功的, 因为有关重要性的两组试验已经通过。

● 3 个最佳水平都优于 3 个临界水平, 前者都符合技术条件, 而后者都不符合。

● D 值为 $1\ 050 - 738 = 312ma$; $\bar{d}$ 值为: $\frac{17+29}{2} = 23$; 所以 $D:\bar{d}$ 的比率为 $312:23$, 远比最小值 1.25 要大得多。第 1 阶段是成功的。能得出的惟一结论是工程小组选择了正确的因子 (但在第 1 阶段尚不能确定哪个是红 X、哪个是粉红 X 等)。

(2)

● 判断极限 (最佳) 为: 6 个最佳中值 $\pm 2.776 (d/1.81) = 738 \pm 2.776 \times 7.5/1.81 = 773 \sim 703$ 。

● 判断极限 (临界) 为: 临界中值 $\pm 2.776 (d/1.81) = 1\ 085 \sim 101$ 。

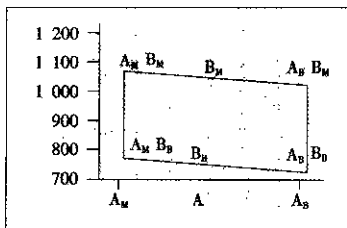
● 取舍选择法, 最佳水平中值和临界水平中值之间的中心线为 $\frac{1\ 050+738}{2} = 894$ 。

● 在判断极限法的基础上, 从因子 A 到 F 都保持在它们相应的判断极限范围内。因此, 这些因子都不是重要的。同样, 根据取舍选择法, 从因子 A 到 F 都没有通过 894 的中心线, 因此, 这些因子都不是重要的。

● 因子 G 是惟一超过最佳和临界两个判断极限范围的因子, 它也通过中心线, 因而, 它是惟一的、确定无疑的红 X。



●根据“如果主要因子是非重要的，它的有关的交互影响效应也能被证明是非重要”的准则，那么就不存在明显的交互影响效应。（见图 13—5，由图可见：在本案例中研究了两个因子 A 和 B，同时还有 4 个可能组合—— $A_M B_M$ ， $A_M B_B$ ， $A_B B_M$ 和 $A_B B_B$ 。图中数字是从表 13—9 得来。图中有两条平行线，这就意味着：A、B 的交互影响实际上为零。）



从第 1 阶段： $A_B B_B=738$

$A_M B_M=1\ 050$

从第 2 阶段： $A_M B_B=768$

$A_B B_M=1\ 020$

AB 交互影响=0

图 13—5 二因子交互影响效应试验

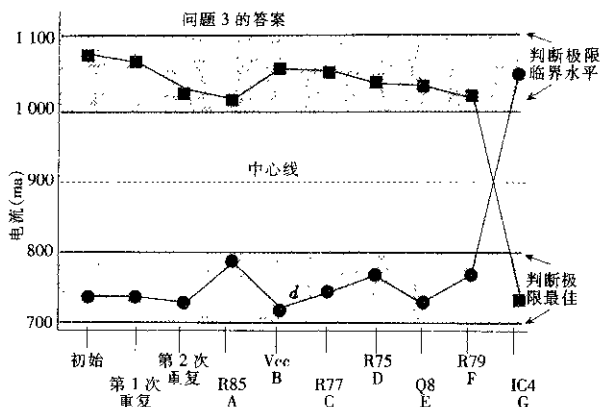
(3) 见图解曲线（如图 13—6 所示）。

(4) 因为只有一个因子（G）就可实现完全逆转，所以不必进行求交运算。同样，也不必进行析因分析，因为只有一个因子，不存在交互影响效应。

(5) 由于因子 G 是惟一的红 X，其余 6 个因子的公差范围可以放宽以减少费用。（这要求进一步试验。初期的第 1 阶段和第 2 阶段的试验是由一名技术人员花 2 天时间完成的。）集成电路的补偿电压公差下降一半，为 $\pm 4\text{mv.}$ ，同时也向供应商展示如何以低成本利用 DOE 方法来减少过程变量。公差范围从对 $\pm 1\%$ 的部件放宽到 $\pm 5\%$ 。其余部件的公差范围不变，主要是因为更大的公差范围已无价格优势。由于节省了用于改进产品质量的费用和实际上的成本的降低，全部节约费用在第一年就达到 450 000 美元。

(6) 如果部件（因子）的选择顺序发生逆转，同时集成电路的

27



图例: ● 全部因子最佳(除 A、B 等), 在第 2 阶段处于它们的临界水平。
 ■ 全部因子临界(除 A、B 等), 在第 2 阶段处于它们的最佳水平。

图 13-6 变量搜索: 发动机控制模块——图解曲线

补偿电压(因子 G)成为首位, 第 2 阶段的 14 组试验将减少到 2 组, 因为全部逆转的结果只是由于对换了 G。

由于变量搜索法是十分重要和实用的技术, 本章将再研究几个案例和小组练习。这些案例用表列举了变量搜索法的多种用途, 同时也列举了在实施变量搜索法时一些有经验的专业人员也可能犯的典型错误。

案例研究 2: 环氧树脂破裂

本案例涉及到我们一家在远东的客户, 它曾碰到过这样一个问题: 黏结底座和盖罩的环氧树脂在固化过程中发生破裂。废品率高达 10% 左右, 而且三年来问题一直没有得到解决。一个 DOE 小组



采用了变量搜索法来解决这个问题（见表 13—10）。

表 13—10

案例研究 2：环氧树脂破裂

第 1 阶段			
因子代码	因子名称	因子水平	
		最佳 (B)	临界 (M)
A	固化温度 (°C)	150	120
B	固化时间 (分钟)	30	45
C	垂直度	0°	1.5°
D	环氧树脂用量 (导柱覆盖率)	100%	75%
E	调谐孔的对准	无制约	制约
F	环氧树脂黏度 (CPS)	199 000	180 000
样本量: 20 件; 加权故障 = (故障产品数 × 李克特度量尺度)			
		最佳 (B)	临界 (M)
运行 1		0	11
运行 2		0	29
运行 3		0	33
$D = 33$; $\bar{d} = 12/2 = 6$; $D/\bar{d} = 33 \div 6$. (因此大于 1.25)			
第 2 阶段			
试验编号	组 合	结 果	需要性
1	$A_M R_B$	37	A: 红 X
2	$A_B R_M$	0	
3	$B_M R_B$	0	B
4	$B_B R_M$	25	
5	$C_M R_B$	0	C
6	$C_B R_M$	19	
7	$D_M R_B$	0	D
8	$D_B R_M$	22	
9	$E_M R_B$	12	E: 粉红 X
10	$E_B R_M$	0	
11	$F_M R_B$	0	F
12	$F_B R_M$	27	
第 3 阶段: 求交运算之 1			
	组 合	结 果	
	$A_M E_M R_B$	30	
	$A_B E_B R_M$	0	

然而，因子 E 必须保持临界水平，以防止基座孔与扭曲线圈的调谐孔之间的轴线不重合。因而，要进行第二次求交运算。

求交运算之 2

组合	结果
$A_M E_M R_B$	24
$A_B E_M R_M$	0

注：~ 不重要。

资料来源：Motorola, Penang, Malaysia.

绿 Y

采用 0~10 的利克特度量尺度来确定环氧树脂的破裂程度，其中 0 表示没有破裂，10 表示 100% 破裂。

结论

- (1) 红 X 为固化温度，固化温度保持在 150℃。
- (2) 其他因子是不重要的，但其他因子应继续受到限制。
- (3) 减少环氧树脂的用量和黏度可降低成本。
- (4) 废品率从 10% 降低为 0。
- (5) 每年可节约 40 296 美元。

小组练习 2：丝网印刷（未成功）

在丝网印刷过程中，印刷油膏的重量应满足 $1.0 \text{ 克} \pm 0.1 \text{ 克}$ 的技术要求。为了减少生产质量的不稳定，DOE 小组设定了有 6 个因子的变量搜索试验，每组试验测试 50 个网点，并计算平均用膏重量和标准差。第 1 阶段和第 2 阶段试验的结果见表 13-11。



表 13—11

小组练习 2: 丝网印刷

因子代码	因子特征	因子水平	
		最佳 (B)	临界 (M)
A	油墨黏度 (Mcps)	1.32	0.95
B	刮片速度 (英寸/秒)	1.6	3.8
C	刮片压力 (磅)	8	20
D	超程 (mils)	0	0.5
E	丝网擦净	是	否
F	丝网提升速度	慢	快

阶段	试验序号	组合	平均重量 (克)	标准差
1	1	全部最佳	1.071	0.012 3
	2	全部临界	0.965	0.061 2
2	1	$A_M R_B$	0.936 9	0.018 0
	2	$A_B R_M$	1.162 3	0.017 7
	3	$B_M R_B$	1.131 2	0.035 9
	4	$B_B R_M$	0.980 4	0.023 7
	5	$C_M R_B$	0.936 6	0.024 1
	6	$C_B R_M$	1.208 8	0.038 7
	7	$D_M R_B$	1.046 4	0.032 6
	8	$D_B R_M$	0.984 3	0.036 8
	9	$E_M R_B$	1.031 0	0.015 7
	10	$E_B R_M$	1.087 5	0.019 7
	11	$F_M R_B$	0.959 1	0.026 1
	12	$F_B R_M$	1.123 7	0.105 9

问题:

- (1) 变量搜索法是否是正确的起始点?
- (2) 是否选定了正确、最佳、临界水平?
- (3) 是否只应有一个绿 Y——即印刷油墨的重量?
- (4) 50 个网点的样本量是否合适?
- (5) 采用平均用量和标准差是否是正确的方法?
- (6) 第 1 阶段试验是否正确实施了?
- (7) 第 2 阶段试验是否正确实施了?
- (8) 是否应进行第 2、3 阶段的试验?
- (9) 小组对红 X 和粉红 X 的结论是否正确?
- (10) 你将如何进行合适的 DOE 试验?

答案

(1) 由于在每个单元上有很多网点要研究, 应推迟使用变量搜索法, 直到多变量法明确指出族系。

(2) 假定最佳水平是合适的, 而某些因子的临界水平却不同寻常, 达到 $2.5:1$, 这个范围实在是太大了。此外, 有一个因子——丝网提升速度并未定量。“慢”和“快”应是客观测定, 而不是主观臆断的。

(3) 可以设想: 油墨网点的大小和形状可能与油墨的重量同样重要。

(4) 样本量太大。如果首先实施多变量法, 含有每个单元分成 5 个区域 (东、南、西、北、中) 的一个样本量就足够了。

(5) 采用平均值是不合适的, 它冲淡了变量。同样, 标准差、成品率和缺陷百分率包括了大量属性类型的数据, 它们缺乏变量数据的敏感性。

(6) 第 1 阶段没有正确完成。初始的全部最佳和全部临界水平根本没有重复性, 它们应重复两次。虽然有很大的标准差, 但全部最佳水平并未高出全部临界水平。这样, 就应该终止试验并应试用其他因子。

(7) 暂时撤下标准差，而集中注意中值。在第 2 阶段中，已显示了因子 B（刮片的速度）和因子 D（丝网的超程）有一定的重要性；另一方面，将标准差作为绿 Y，因子 A（油墨黏度）和因子 E（丝网擦净）也有某些重要性。但在整个第 2 阶段中，存在太多的“噪音”，没有足够的“信号”可用来得出明确的结论。

(8) 由于第 1 阶段的失败，不应再继续进行第 2、3 和 4 阶段。

(9) 小组有关因子 B 是红 X 和因子 A 是粉红 X 的结论就绿 Y 的平均值和标准差而言，仅具有部分的推理方面的正确性，参见答案 7。

(10) 正确的 DOE 技术应从多变量法入手，随后是产品/过程搜索。只有通过这两个 DOE 规则产生思路后，再进行变量搜索。

案例研究 3：“墓碑”缺陷（错误技术）

“墓碑”是一种发生在印制电路板上的不正常现象：当部件芯片的一根引出线被回流焊焊上，而另一根却翘起，从而使水平部件垂直地矗立在印制板上，就像墓碑一样。公司在改进了它的工艺过程后，缺陷水平下降了 10ppm~20ppm（考虑了印制板上的全部部件）。然而，公司的目标是 1ppm 或更少。公司曾试用表 13-12 中的因子进行变量搜索试验。DOE 小组认识得太晚了：要想收集到 10ppm~20ppm 的缺陷水平，只有 75 块印制板的样本量实在是太小了，甚至 10 000 块印制板的样本量也难以收集到如此低的缺陷水平。因此，用变量搜索法是错误的。

小组折回到成对比较法，收集了 8 个好的印制板和 8 个带“墓碑”的印制板（而不管产量多高每天平均出现这些数）。他们测试了芯片部件和焊点的几个特性，发现只有在焊膏与焊盘的精度上有些重复差异。质量好的组件，其芯片部件的焊膏透过焊盘中心，而质量不好的组件却有 0.3mm 的偏离（芯片上焊膏与焊盘不对中）。最终的解决办法是将原来的三洋牌即时印刷机换成具有较高焊膏一

焊盘精度的 MPM 印刷机。结果，在以后的生产中连续几个月没有出现“墓碑”缺陷。

表 13—12

案例研究 3：“墓碑”缺陷

因子编码	因子特征	因子水平	
		最佳	临界
A	焊膏重量（克）	0.312	0.334
B	位置精度	对中心	偏心 0.25mm
C	回流焊型面		
	(1) 斜率	1.2℃/秒	3.5℃/秒
	(2) 吸收数	2	1
	(3) 吸收温度	100℃和 175℃	只为 100℃
D	焊膏型号	Alpha	SCM
		1005205 A15	1005205 A20
	结果	0ppm	0ppm

注：样本量=75 块印制板。

资料来源：Motorola, Penang, Malaysia.

案例研究 4：隐型眼镜圆柱度没有满足技术条件

本案例研究是一个说明在第 2 阶段只产生部分逆转的因子组合，经继续工作，在最终求交运算时发生全部逆转的很好的例子。

在第 11 章关于用成对比较法的小组练习 1 中，谈到了不能满足光学技术条件的隐形眼镜。在该练习中，红 X 被确定为圆柱度的尺寸，它超过了最大技术要求 0.025mm。进一步的变量搜索研究被引入，以确定有重要影响的因子。表 13—13 列出了第 1、2 和 3 阶段的结果（出于保护公司利益的原因，未列出具体的最佳水平和临界水平）。图 13—7 用最佳水平和临界水平的判断极限和区别重要的和非重要的因子的中线方法，以图解形式生动地表示了结果。

表 13-13

变量搜索:眼镜成型

试验目的:确定红 X 变量,该变量引起圆柱度在凹模上超差 0.025mm。

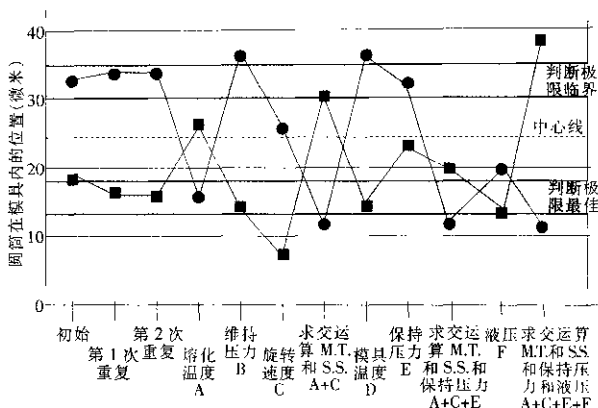
第 1 阶段

试验结果			
试验	全部最佳	全部临界	备注
1-B 和 1-M	0.018mm	0.032mm	
2-B 和 2-M	0.016mm	0.033mm	
3-B 和 3-M	0.016mm	0.032mm	
* 3 个全部最佳和 3 个全部临界以随机顺序试验			
$D: d$	$D - (B - M) / 2$	$\bar{d} - 0.002 + 0.001 = 0.0015$ so	
$D = 0.032 - 0.016 = 0.016$		$D / \bar{d}$ is $> 1.25 : 1$	

第 2,3 阶段

试验变量识别	设定变量		试验结果			
	最佳值	临界值	试验 #	全部最佳, 除了	试验 #	全部临界, 除了
熔化温度	最佳值	临界值	A-1	0.026mm	A-2	0.016mm
维持压力(P17—21)			B-1	0.015mm	B-2	0.035mm
螺杆速度(D2)			C-1	0.007mm	C-2	0.026mm
模具温度			D-1	0.015mm	D-2	0.035mm
保持压力(P7 16)		保持	E-1	0.024mm	E-2	0.031mm
液压(P4)			F-1	0.014mm	F-2	0.020mm
求交运算 熔化温度与螺杆速度			G-1	0.030mm	G-2	0.012mm
求交运算 熔化温度、螺杆速度与保持压力			H-1	0.020mm	H-2	0.013mm
求交运算 熔化温度、螺杆速度与液压			I-1	0.038mm	I-2	0.011mm
求交运算 保持压力与螺杆速度			J-1	0.039mm	J-2	0.015mm
求交运算 保持压力与液压			K-1	0.022mm	K-2	0.012mm

资料来源: Ciba Corp., Atlanta, Ga.



图例：●全部因子最佳，除了 A, B 等，在第 2, 3 阶段处于临界水平。

■全部因子临界，除了 A, B 等，在第 2, 3 阶段处于最佳水平。

资料来源: Ciba Corp. Atlanta, Ga.

图 13—7 变量搜索：眼镜成型——图解曲线

结论与评论

(1) 第 1 阶段是成功的，因为 $D:d$ 的比率为 $11:1$ ，远比 1.25 大。

(2) 如表 13—13 和图 13—7 所示，因子 B（反压）和因子 D（模具温度）是不重要的（虽然从技术上讲，这两个因子恰好都超出临界水平判断极限的上限）。

(3) 因子 A（熔化温度）和因子 C（螺杆速度）在第一次求交运算时，发生部分逆转，要求交换另一个因子 E（保持压力）。因子 E 只超出一个判断极限（最佳），而没有超出另一个（临界）。

(4) 因此，对因子 A, C 和 E 进行第二次求交运算。但结果仍然是部分逆转，需要交换另一个因子 F（液压）。因子 F 超出一个

判断极限（临界），而没有超出另一个（最佳）。

（5）最终，由因子 A，C，E 及 F 进行求交运算，产生了巨大逆转。这表明在这些因子之间存在交互影响效应，需要用析因分析法进行定量。

（6）实际上，当存在 3 个以上因子时，采用 2^4 全因子试验法作为变量搜索法的后续，要比只用析因分析法进行数字处理要好。

小组练习 3：点焊——两个参数（即两个绿 Y），机械的和目测的

该小组练习阐述了采用变量搜索法以优化两个或更多的输出（或响应，或绿 Y），同时解释了如果输入因子对产生一个绿 Y 是最佳的，而对另一个绿 Y 并不是最佳因子时，该如何处理。

采用变量搜索试验改进一个具有难以接受的缺陷率的点焊工艺。有两个绿 Y 被认为是重要的——首先是机械的——一个好的焊点被定义为：当焊点被断开后，其焊接部位在直径上至少有 1/4 英寸以上的焊柱。第二个绿 Y 是目测的——最小的焊接变形。对于第二个绿 Y，可建立李克特度量尺度：其中 1 代表最佳，10 代表最差。5 分或更小是可接受的，6 分或更多则是不能接受的。表 13—14 所示为 5 个焊接参数。

问题

- （1）是否量化和测量了两个绿 Y？
- （2）第 1 阶段的 6 组试验是按随机顺序进行的吗？
- （3）哪些是非重要因子，为什么？
- （4）哪些是重要因子，为什么？
- （5）是否对一个绿 Y 是重要因子而对另一个绿 Y 却不是？
- （6）是否要进行求交运算？
- （7）如果问题 5 的答案为“是”，下一步如何工作？

表 13—14

小组练习 3: 点焊

		因子水平		
因子编码	因子说明	最佳	临界	
A	焊接数量	15	22	
B	保持时间 (秒)	25	37	
C	加压时间 (秒)	25	37	
D	焊接压力 (PSI)	25	35	
E	通用焊接 (%)	80	90	
阶段	试验	组合	机械的	目测的
1	全部最佳		通过	1
	全部临界		失效	9
	全部最佳		通过	1
	全部临界		失效	10
	全部最佳		通过	2
	全部临界		失效	8
2	A _M R _B		通过	4
	A _B R _M		失效	6
	B _M R _B		通过	3
	B _B R _M		失效	10
	C _M R _B		通过	2
	C _B R _M		失效	9
	D _M R _B		失效	1
	D _B R _M		通过	8
	E _M R _D		通过	7
	E _B R _M		失效	5

资料来源: Craftsman Co., Melrose Park, Ill.

答案

(1) 两个绿 Y 都被量化。机械的绿 Y 用一个变量 (在直径上至少 1/4 英寸焊柱), 而目测的绿 Y 用利克特度量尺度。然而, 如果能用实际的计量数据表示, 则比只用“通过”或“失效”表示好。

(2) 第 1 阶段应始终按随机顺序进行 (见第 15 章), 否则, 不



可控制的或“噪音”因子会以循环方式变化或在它们内部偏移，并影响其结果。在本案例中，序列是随机的，但未示出。

(3) 第2阶段表明：因子 A、B 和 C 是不重要的——两个绿 Y 在第1阶段的读数都没有发生逆转。

(4) 和 (5)

- 因子 D 是重要的， D_B 对机械的绿 Y 产生好的结果，但对目测的绿 Y 却产生差的结果。

- 因子 E 在与因子 D 的相反方向是重要的， E_B 对机械的绿 Y 产生差的结果，但对目测的绿 Y 却产生好的结果。

(6) 应该进行求交运算，但未做。

(7) 有 4 种处理方法：

- 用全部最佳因子和全部临界因子进行的求交运算可以重复，与第1阶段一样。这不是常规做法，但如果全部最佳因子产生好的结果，而全部临界因子产生差的结果（如在第1阶段所进行的），这表明：因子 E_B 可能被认为是具有次要影响（浅粉红 X），而因子 D_B 则具有显著影响（红 X）。

- 应定期对 $D_B E_M R_M$ 和 $D_M E_B R_B$ 进行求交运算。如果其结果与第1阶段重合，这表明：为了使两个绿 Y 都得到最佳结果，因子 D 应保持在最佳水平，而因子 E 应处于临界水平。

- 如果求交运算表明一个或两个绿 Y 都变坏，将需要 2×2 因子作为有各种水平的不同组合的后续试验。比如说，取因子 D 为 28 和 32 磅/平方英寸，同时 E 为 83 和 86，以观察通过这 4 种组合之一能否使机械的和目测的绿 Y 都能优化。

- 如果 2×2 因子试验没有成功，将需要采用调优运算（见第 16 章），以确定对机械的和目测的绿 Y 来说都最佳的 D 和 E 的值。

变量搜索法：软件应用

如以下案例所描述，变量搜索法已用于软件应用方面。

案例研究 5: 软件应用

对一台新的具有紫外线可擦性微控制器功能和储存试验容量的微控制器试验装置, 设计了变量搜索试验, 以检查微控制器的存储容量和功能。软件程序的结构组成如下:

- (1) 可编程只读存储器 (PROM) 的容量;
- (2) 计时器;
- (3) 随机存取存储器 (RAM) 和输入/输出试验。

每次运行时, 主导装置 UVC8 向从属的 UVC8 RAM 发送一段试验数据 (字节), 在接到这些数据后, 从属的 UVC8 进行自我测试, 并在运行下一段数据前, 将结果转发回主导装置 UVC8。其结果通过 4 个发光二极管 (LED) 显示, 标明: 空白、不空白、好和坏。

选择了 4 种类型的 UVC8 状态, 每种状态由两种水平组成, 如表 13-15 所示。

表 13-15 案例研究 5: 软件应用

因素代码	UVC8 状态	水平	
		真 (T)	伪 (F)
A	空白	空白	不空白
B	不空白	不空白	空白
C	可编程	可编程	不可编程
D	不可编程	不可编程	可编程
测试	组合	结果	
1	A ₁ R ₁	真	
	A ₁ R ₂	伪	
2	B ₁ R ₁	伪	
	B ₁ R ₂	真	
3	C ₁ R ₁	伪	
	C ₁ R ₂	真	
4	D ₁ R ₁	伪	
	D ₁ R ₂	真	

DOE 技术的派生物：变量搜索与成对比较的组合

近几年来，本书作者根据变量搜索和成对比较的原理，开发了 DOE 技术，实践证明它在生产和管理工作的某些与众不同的方面是非常有用的。

结论：红 X 是空白试验部分，因而排除了来自其他试验部分的可能。

在调整程序后，将程序在仿真程序器上模拟，然后用现在的试验装置和用于生产的新装置对 UVC8 进行重复试验，其结果为：

试验	装置	结果
空白	1 和 2	真实
不空白	1 和 2	真实
可编程	1 和 2	真实
不可编程	1 和 2	真实

结论：该程序是全功能的，同时新装置可用于生产。

产品应用

如果组装的产品十分复杂，将会发生什么情况呢？

(1) 不能对好的和差的部件进行非破坏性的成对比较试验，将会发生什么情况呢？

(2) 破坏性成对比较不能作为检测部件或子组件差异的后续试验（因为它们可能在分解过程中受到损坏），将会发生什么情况呢？

在此种场合下，在既成事实前即在组装和制成之前，选择最佳部件和临界部件，比如说带有 10~20 个最佳部件和 10~20 个临界部件的 20~40 个组件。在完成装配后，选出 8 个最佳的绿 Y 和 8 个最差的绿 Y，然后，对与 8 个最佳和 8 个最差组件有关的部件的原始数据进行图基检验。在最佳部件和临界部件之间，如果某些部

件的终结数为 6 或更多, 则这些部件是重要的; 如果终结计数小于 6, 则这些部件是不重要的。

用取舍选择法制造 30 个组件, 装在其上的部件性能指标范围为理想值 $\pm 3\sigma$ 。然后, 选择 8 个最佳和 8 个最差的绿 Y。接着, 对每个不同部件进行如上所述的成对比较图基检验, 以确定合计终结计数。这种取舍选择法较少优先选用, 因为它并不集中在最佳和最差的部件量值——变量搜索法——因而潜在的结果有较少的敏感性。

变量搜索/成对比较法案例研究 1: 旅行车

某家生产结构设备的大型工厂曾面临旅行车套筒轴在保修期内在外场存在不可接受的故障率的问题。为了在问题暴露于现场和客户之前找到有相似故障模式的套筒轴, 工厂决定进行加速强化应力试验, 这样, 可以在运行的最初 20~80 小时内查出故障。绿 Y 就是出现重大故障的小时数。

一辆发生故障的汽车损坏得十分严重, 以至分解汽车只能揭示出毁坏的部件或组件, 因而既不能用无损检测法, 又不能用破坏性的检测方法进行好、差汽车的成对比较试验。

工厂决定装配 30 辆装有 8 个“最佳”部件和 8 个“临界”部件的汽车。怀疑会引起问题的部件特征为:

- 套筒的球形表面;
- 平板的球形表面;
- 套筒与平板之间球形表面的重叠量;
- 定位销的长度;
- 外筒花键尺寸;
- 内筒花键尺寸;
- 花键轴的重叠量;
- 弹簧夹的深度。

汽车装配好后, 进行加速强化应力试验。试验的结果为: 6 辆

汽车是好的，10 辆是差的。上述 8 个部件特征中有 5 个显示图基最终计数为 6 或更多，同时，套筒与平板之间球形表面的重叠量为红 X。下一步的工作是研究部件平板的材料。

整个研究工作进行了 5 个月，结果是富有成效的。从此以后，汽车不仅在保修期内，而且即使超过保修期，也不再发生这类故障。

变量搜索/成对比较法案例研究 2：仪表的间隙现象

某家生产高度复杂仪表（在机械方面比瑞士表还复杂得多）的工厂曾经历过其仪表由于间隙现象和在仪表范围内缺乏精确的调整，在现场出现很高的、不可接受的故障率。通过大量工作，工厂决定在工厂内进行一次严格的最终试验，通过 1 小时的强化应力试验来模拟现场故障。

失效的仪表不能取出，同时因为在产品试验过程中取消了一些调整，所以各种部件和子组件只能通过成对比较法进行试验。

工厂装配了 6 个装有 13 种最佳部件/参数/调整的仪表和 6 个装有 13 种临界部件/参数/调整的仪表，并对其使用变量搜索/成对比较 DOE 技术（当然，最好试验数量更多些，但是工厂要满足生产周期），这 13 种部件/参数/调整如下：

- 洁净度（10 000 级工作室与通常的工作环境相比）；
- 齿轮和轴的表面光洁度；
- 齿形；
- 齿轮啮合调整；
- 防游隙齿轮；
- 轴的端头间隙；
- 宝石轴承的质量；
- 齿轮偏心；

- 润滑；
- 膜片的交叉载荷；
- 平衡；
- 齿轮尺寸；
- 调整顺序。

装配好的仪表经受了最终试验和强化应力试验，并产生了 5 个好的仪表和 7 个差的仪表。每种部件/参数/调整的数值都用成对比较/图基检验法进行检验。

详细的结果是保密的，但在 13 种部件/参数/调整中有 4 种具有 6 或 6 以上的合计终结计数，包括进入超净工作室的必要性（10 000 级或最终 1 000 级）。结果现场故障为 0，同时由于取消了 3~4 次的试验-拒收-返工再重新试验等枯燥无味的反复循环，仪表的产量增加了 4 倍。

变量搜索/成对比较法在管理方面的应用

第 11 章举出了一些案例以说明成对比较技术能同样出色地用于一些非生产性应用领域，诸如销售和市场营销、人力资源、服务行业、商业、农场、学校、医院和政府部门等。

这些是建立在“后验”的基础上——在对事实进行考虑以后，也就是说，在问题描述出来以后，能区别出 8 个好的和 8 个差的绿 Y（如入）的参数或因子已作为原因被确定。例如第 11 章案例研究 4（临时雇员的高流动率）中，因子涉及到通勤距离、对管理层的看法、受教育的水平等。这些因子都是通过回想可能的原因而确定下来的。

如果似乎有必要在进行 DOE 试验之前引入两种截然不同水平的通勤距离、两种截然不同的教育水平、两种截然不同的以前的工作的流动率，以观察它们对临时雇员的相对效果，成对比较试验技术也可由变量搜索法替代，以设计各种因子的最佳与临界水平，然

后再观察其效果——人们（绿 Y）喜欢的或不喜欢的。案例研究 4 可以利用这项技术重新进行设计并稍加修改，如表 13—16 所示。

其次，假定有 20~40 个临时雇员具有从 A 到 E 的最佳水平的特征，并处于最佳的工作地位——F、G 和 H。同样，有 20~40 个临时雇员具有从 A 到 E 的临界水平的特征，并处于临界的工作地位——F、G 和 H。

人力资源部门可对这两组人员进行一定时期（如一年或更长）的观察，当然对于那些很早就离职的人应深入了解。然后，人力资源部门选出 8 个留任时间最长的人和 8 个离职最早的人，并对上述每个特征进行变量搜索/图基检验，以观察哪一种特征的终结计数为 6 或 6 以上（因而是重要的），哪一种特征的终结计数不到 6。如果没有一个特征是重要的，就要开发新的特征。假如有些特征是不重要的，则不必对这些特征再浪费金钱和时间，只对一些重要特征给予重视。

表 13—16 变量搜索/成对比较法：在管理方面的应用

因子代码	因子描述	最佳水平	临界水平
A	通勤距离（英里）	<10	>15
B	驾驶执照	有	无
C	教育水平（上学年数）	<12	>12
D	过去在其他公司的流动率（次/年）	<1	>2
E	对报酬的感受（与其他公司相比）	更高	更低
F	工作变动	稳定	变动
G	领导层的素质	授权	独断专行
H	与固定职员相比的待遇	相等	“二等公民”

医学/药物研究的相似性

医学研究常常需要对大量人群进行长期的试验，以确定独特的药物如阿斯匹林、抗胆固醇等药物的疗效。人群常被分成两组，一组给予独特的药物，而另一组仅给予安慰剂。然后进行几个月甚至

几年的观察以确定特殊的疗效（绿 Y），如对心脏病的疗效。之后作出报告，通过有规律地服用药物的小组与只服安慰剂的小组的对比，指出是否在防治心脏病发作方面有明显的区别（这两组人都不知道他们服用的是什么药）。这样的研究存在以下缺点：

- （1）只有一个绿 Y（输出），如心脏病发作；
- （2）只有一个感兴趣的输入变量——药物与安慰剂的对比；
- （3）研究涉及了大量人群——通常为 1 000~5 000 人；
- （4）其他输入变量，诸如年龄、健康状况、饮食习惯、环境因子、气候条件以及大量其他变量的差异，都可能单独或集中地对结果产生足够重要的影响，但可能没有被分割开；
- （5）试验需要进行很长时间，而上述变量可能随时间和人而改变；
- （6）结果是用相对小的改善百分率计量的。

作者并非医学专家，然而，在下列医学研究中，仍有可能采用经过改进的变量搜索/成对比较试验。

（1）可能存在一个以上的绿 Y——如心脏病、中风、血压、贫血及其他相关边缘效应；

（2）可以保留主要的、感兴趣的输入变量——药物与安慰剂的对比；

（3）其他原因的因子可分成最佳的和临界的，如表 13—17 所示。

表 13—17 变量搜索/成对比较法：在医学上的应用

因子代码	因子描述	最佳水平	临界水平
A	药物	是	安慰剂
B	年龄（岁）	<65	>65
C	性别	女	男
D	体重（磅）	低于 10 到 20	超过 20 到 40
E	饮食	按规定的	无节制的
F	抽烟	否	是
G	胆固醇水平	低	高
H	家族病史	无	有
I	环境	低污染	高污染



样本量将由 40~80 个具有最佳水平特征的人和 40~80 个具有临界水平特征的人组成，每一组将在那些服用药物的和服用安慰剂的人中等分出来。

在指定的试验和观察期结束后，选出 8 个具有最佳结果（试验前后对比）的人和 8 个具有最差结果（恶化）的人。然后，对上述每个因子进行成对比较 t 检验，只有合计终结计数为 6 或 6 以上（90% 以上的置信度）的因子才被认为是重要的。

因此，通过非常小的样本量即可获得有较大意义和价值的结果，并有一定的置信度，这远远超过由常规实践所获得的结果。这项技术建议虽然尚未试用过，但正处于与一家制药厂协商的过程中，以便推出一种能降低费用、节省时间、提高质量的方法。

管理案例研究：设备的（售后）服务

修理服务业是一个声名狼藉的行业，这缘于其以次充好的行为、漫天要价、毫无道理的返工，以及素质很差的商人、粗心大意的装配工、消极被动的服务和永不满意的用户！再加上 20%~40% 的上门服务的维修结果归于被称为“无明显缺陷”（NAD）或“未找到故障”（NTF）的范畴，更使得用户备受挫折。

一家设备制造厂面对保修期内高昂的维修费用，决定建立起选择其销售商和服务站的规范。工厂决定采用变量搜索/成对比较方法。该厂选定了一个大城市地区，那里有它的好几家销售店。

输出或绿 Y 是设备制造厂最终用户——即顾客的满意指数。确定指数的依据是：（1）一些与机器特性、安装、服务、价格等相关的参数的重要性（I）；（2）用户对每个参数的评价（R）；（3）每个参数的得分（I×R）；（4）总分——表示满意程度的百分比，0~100。输入的因子估量了销售商、安装队、服务站以及产品特性，如表 13—18 所示。为每个因子选择两级水平，即最好水平和临界水平。根据制造商销售量和其产品的声誉，选择 15 个最好的和 15 个临界的销售商。

表 13—18

定变量搜索/成对比较法：售后服务

因子代码	因子描述	最佳水平	临界水平
	销售商		
A	规模	小	大
B	代理	独家生产	几家制造商
C	销售人员流动率	低	高
	安装人员/服务人员		
D	流动率	低	高
E	培训	高	低
F	教育水平	高	低
G	工资	高	低
H	达到高的顾客满意指数的奖金	是	否
	产品		
I	复杂性	低	高
J	可靠性	高	低
K	安装/服务难度	低	高
L	现有的诊断	是	否

经过一年的观察，取得了用户满意指数的得分情况，并为成对比较/图基检验评估选定了 8 家最佳销售商和 8 家最差销售商。初步的结果表明：缺点更多地在于制造厂的产品可靠性和维修性方面，而不是原先所怀疑的销售商/安装队/服务站环节上的弱点。

作为 DOE 小组行动准则的问题

(关于绿 Y 的确定和精度问题与前面所述的 DOE 技术相同。)

(1) 是否采用了线索生成技术以确认重要变量，并将它们集中，在第 1 阶段绘制成表？

(2) 是否采用计算机模拟或电路分析技术以确定所选的变量（这类技术必须始终紧跟变量搜索法，并使用现行硬件作为安全保障的方针）？

(3) 列表时，是否将选定的变量实际按照从最可信的到最不可

的顺序排序？（这能明显地减少第 2 阶段进行试验的数量。）

（4）如果以前的历史表明存在低水平的故障水平，是否已将每个变量（或因子）的最佳水平作为它的目标值来确定？

（5）如果以前历史表明有不可接受的故障率，是否已根据其现有水平的改进情况来决定每个变量（或因子）的最佳水平？

（6）每个变量（或因子）的临界水平是否已作为其现行水平被确定，以逐步提高到最佳水平？

（7）如果有些特殊因子的最佳水平或临界水平没有确定，是否采用了试验性的“B vs. C”的比较试验（见第 15 章），作为预先步骤来确定哪一个水平较好（B）、哪一个水平较差（C）？在进行这组试验时，其他变量应保持不变。

（8）样本量如何确定？是否进行了努力以选定绿 Y 的变量，或者至少用李克特度量尺度将一个有属性特征的绿 Y 转换成变量？其目的是保持小的样本量。

（9）确保第 1 阶段没有被跳越——这是一个常犯的错误。

（10）确保在第 1 阶段使用了 3 个全部最佳和 3 个全部临界的变量进行运算，而不是只用了 1 个全部最佳和 1 个全部临界的变量。

（11）确保 3 个全部最佳和 3 个全部临界的试验是按随机顺序进行的（见第 14 章随机化）。

（12）第 1 阶段的重要性试验是否成功？（ $D : \bar{d} > 1.25$ ）。

（13）如果不是，是否采用某些因子水平的修正值或因因子本身的修正值，将第 1 阶段的试验重新做一遍？

（14）是否在第 1 阶段计算了最佳和临界水平的判断极限？

（15）或者，至少是否在最佳水平中值与临界水平中值之间划出中心线，以在重要与非重要因子之间作出区别？

（16）在第 2 阶段，对每个因子的重要性是否作出了正确的决定——重要的（红 X）、不重要的或有些重要的？

（17）如果在第 2 阶段结果只有部分逆转，是否还有其他因子被确定为有一定重要性，而且有 2 个或 2 个以上的因子参加试验以



求全部逆转？

(18) 是否进行了求交运算，以判别重要或不重要的因子？

(19) 是否设计了析因分析，以使重要的主效应和交互影响效应定量化？

(20) 为了降低成本，在变量搜索试验后，是否继续工作以确定那些不重要因子的公差范围可以放宽到多大？

DOE 项目评审时对高层管理提出的问题

(1) 为什么选择变量搜索法比用其他技术（如完全析因法、析因法、田口正交表法、计算机模拟等）要好？

(2) 试验将持续或已经做了多久？

(3) 试验结果是否比绿 Y 的现有水平有所改进？

(4) 是否有一个以上的绿 Y？如果是，是否在某些因子对一个绿 Y 的重要性与这些因子对另一个绿 Y 的重要性方面有矛盾？

(5) 如何解决这些矛盾？

(6) 是否进行了求交运算？

(7) 有多少因子是重要的？是否设计了析因分析？主效应和交互影响效应的相对量是什么？

(8) 在变量搜索试验后，是否接着进行散布图、“B vs. C”试验、正向控制和过程确认等技术？

(9) 变量搜索试验是否放宽了非重要因子的公差范围，从而降低了成本？

(10) 变量搜索试验的结果是什么——在质量提高、费用节约和周期缩短等方面？



第 14 章

全析因：识别与量化每个交互影响的最佳技术

第 13 章将全析因介绍为所有正规的试验设计中的最佳技术，因为全析因可以简捷而又正确地识别和量化每一个一阶影响、二阶影响、三阶影响及四阶交互影响效应。（实际上，试验设计的研究先驱和鼻祖费雪爵士仅使用全析因而没有借助于计算机。他决不利用分析因之类的捷径来玷污他的成果！）本章将探讨全析因的用途、局限、目标、原则和方法，并通过案例研究和小组练习的方式给出一些实例。（欲了解全析因在详细的解决问题路径图中的位置，请参见图 7—3。）



局限

正如第 13 章阐述的那样, 对一个随机的、重复的且平衡的设计而言, 若 n 为要研究的因子或变量总数, 则全析因需要 2^n 次试验。若 n 的取值为 5、6、7、8、9 和 10, 则全析因分别需要 32、64、128、256、512 和 1 024 次试验。要完成如此多的试验需要大量的时间, 并且出错的可能性和不可控制的外部 (或噪声) 因子会成指数地增加。因此, 从实用的角度出发, 全析因应限于 4 个或 4 个以下的因子。于是变量搜索便成了选择 5~20 个因子的试验。

目标

(1) 在由 1 种或多种 4 线索生成技术挑选出来的 2、3 或 4 个变量中, 确定哪些变量重要, 哪些变量不重要。

(2) 扩展不重要变量的公差, 以减低费用。

(3) 量化不重要变量及其交互影响效应的大小和期望的方向, 并收紧这些变量的公差, 以达到 2.0 及更高的 C_p , C_{pk} 。

(4) 尽管全析因可以用作解决问题的工具, 但应防止在研究问题的早期就试图使用该技术。相反, 在开始时应优先使用一种或多种 4 线索生成技术。

(5) 然而, 在设计开始时, 如果样本数量对于线索生成工具还不足够, 且设计小组可以研究不多于 4 个的变量, 那么全析因肯定会成为主要的试验设计技术。

原则

大多数试验设计研究中的 3 个基本原则是: 平衡、重复以及随机化。

平衡

全析因的能力是对 4 个（或少于 4 个）被选变量中的每个变量都用其他变量的所有水平（通常为两个水平）进行检验。由于要对所有的主效应和交互影响效应进行系统识别和量化，因此，应按平衡设计的方式对变量和水平的所有可能组合进行检验。平衡设计的规则是：试验数必须比对照数或列数要多。这正是分析因支持者们所违反的规则。

重复

在一次试验中，往往不可能包括所有不可控制的原因、因子或者变量。与所选因子的纯“信号”相反，这些原因可能会将“噪声”或不一致性带入结果之中。在每一组合或部件中，重复（即复现）的目的是确定每个部件中的偏差或不一致性。如果不一致性太大（比方说，在每个部件中的两个读数的偏差大于 10%），那么，试验的噪声将会“模糊”掉信号，因此应当消除这些噪声，以使其其他试验得以在更为受控的条件下进行。

因此， 2^2 、 2^3 和 2^4 试验应当重复或复现两次，以确定试验的信噪比和有效性。一个关于重复的例外情况是，若在试验中有 16 个部件，那么，可以对每一列执行一个终结计数大于或等于 6 的图基检验，以确定重复的置信水平。根据该规则，一个具有 16 个部件的 2^4 析因无须重复，一个仅具有 8 个部件的 2^3 析因应重复一次（即每个部件两个读数），一个仅具有 4 个部件的 2^2 析因应重复 3 次以上才能得到每个部件的 4 个读数。

随机化

随机化原则最好用下列游戏来解释：

- 至少邀请 20 位参试者，每人在 4 个数字（1、2、3 和 4）中选

出其中一个，并记下每位参试者的选择。绝大多数人会选择了数字“3”。本书作者在处于不同地理位置、具有不同文化背景和宗教信仰的33个国家中重复了该试验，结果是98%的参试者选择了数字“3”。

- 至少邀请20位参试者选择他心目中的一件家具，大多数参试者会选择“椅子”。

- 按类似的方法让参试者选择一种鲜花，大多数参试者会选择“玫瑰”。

我们人类在选择某种特定的响应时，既存在倾向于某种选择的偏见、情绪、预感和意见，也存在群体本能的成分。我们无论如何都不具有随机选择数字的能力。如果我们能够做到随机选数的话，就会有25%的参试者选择“1”，25%的参试者选择“2”，25%的参试者选择“3”，25%的参试者选择“4”。

那么，人们或许会问，这与试验设计有什么关系呢？答案是试验的次序或顺序不必遵循某一固定模式。例如，正如第15章将要说明的那样，如果有3个B（较好产品）和3个C（当前产品），一个合理的顺序可能是：先检验3个C，然后作出一些改变并检验3个B，如下所示。

C1, C2, C3 : B1, B2, B3
开始 >----- 结束

但是，这会给试验带来偏见。这些偏见可能是：

- 诸如温度、湿度、静电之类的外部因子或噪声因子会随着时间变化，比如，在上午检验C，在下午检验B。

┌──────────┐
└──────────┘
●上述外部因子随着时间改变的某种趋势。

或



- 上述外部因子随着时间循环。



外部因子的上述任何改变都会引入可能对 C 与 B 的变化结果造成过分影响的“噪声”。

因此，试验必须以随机的次序进行。随机的目的是使这些不可控制的“噪声”因子有相等的机会进入或不进入试验。对田口试验的一个严厉批评是它不相信随机化（另一个主要过失）。可以绕过随机化的惟一条件是，当难于转换新旧过程或方法或可以引入的偏差要大于两个受检过程之间的差别时。

选择随机数的方法有如下 3 种：

(1) 第 1 种方法是使用随机数表，例如由剑桥大学编制的随机数表（见表 14—1）。表中的任何一个数都与其他数没有任何关系。步骤是以随机次序选出 6 个或 6 个以上的数。在表 14—1 上抛出一支铅笔或飞镖。假定它落在数 60 上，然后从左到右、从右到左、从上到下、从下到上或者按对角线的方式读取数字。假定我们从左到右读取 6 个数字来指定 3 个 B 和 3 个 C。然后以某种顺序，例如先检验最小数（或最大数），排列这 6 个数，如下所示：

C_1	C_2	C_3	B_1	B_2	B_3
60	45	92	97	33	44
④	③	⑤	⑥	①	②

正确的随机次序是 B_2, B_3, C_2, C_1, C_3 和 B_1 。

(2) 第 2 种方法是使用计算器。选择 5 个个位数字，例如，4、1、8、3 和 9。按下除法符号，并另行选择 5 个个位数字，例如，2、8、6、7 和 3。按下等号，得到商 1.459 177 623 54。选择一个随机次序序列 1、4、5、9、7 和 6（忽略一个重复的个位数字）。从数学的角度讲，除法运算将产生一个随机数列。

(3) 第 3 种方法是使用带随机数功能的计算器，利用该功能自动产生随机数集。



表 14—1

随机数表

80 14 67 29 70	44 69 53 51 58	40 45 4 31 85	25 6 31 71 14	55 13 31 95 34
48 58 6 90 36	35 19 94 38 13	25 42 21 79 14	94 13 4 36 70	27 67 42 34 39
69 63 85 3 17	82 5 22 26 54	84 78 47 0 91	29 87 90 47 74	32 27 54 83 66
39 65 78 11 40	48 40 23 30 25	45 32 15 9 3	12 14 4 28 68	89 49 73 50 87
61 18 41 7 27	3 83 48 10 88	22 66 22 32 45	30 6 85 5 80	33 72 10 21 7
15 66 33 12 4	90 82 6 33 70	83 57 49 96 12	47 9 73 18 89	80 80 95 24 73
79 12 39 88 47	37 8 18 99 69	31 89 46 64 6	50 48 47 81 51	66 16 10 83 50
27 95 81 3 65	75 84 46 62 60	92 95 15 44 89	41 61 31 28 11	56 61 47 62 37
34 62 68 17 22	27 56 90 53 45	21 84 83 43 71	57 86 34 64 31	55 72 44 19 75
57 16 83 35 96	13 39 71 72 93	42 3 71 92 50	63 24 59 37 34	49 80 31 87 49
3 74 9 96 37	29 11 25 26 30	44 85 78 39 31	50 75 7 35 22	78 66 71 82 30
21 49 58 38 12	72 74 55 91 52	59 25 79 39 10	73 73 13 38 19	56 79 10 23 6
11 8 72 1 8	11 19 88 12 53	3 46 91 4 72	58 26 90 69 37	96 69 43 77 71
7 92 88 46 16	1 14 31 9 43	85 28 54 31 99	1 21 42 89 87	90 5 10 66 1
70 35 91 61 58	51 71 83 74 61	91 8 15 42 95	96 23 86 42 82	44 16 97 91 51
69 65 46 7 6	41 49 47 49 35	47 5 54 15 36	8 80 8 71 18	28 87 3 32 67
91 11 32 74 42	38 72 55 49 63	27 68 23 4 70	8 52 87 6 76	45 25 35 4 66
90 12 32 72 44	80 14 83 88 71	74 88 72 99 80	46 29 2 19 95	90 4 84 79 97
39 91 70 7 15	72 84 78 86 96	33 50 5 30 39	55 86 65 96 26	55 90 14 49 77
42 16 79 69 40	1 93 70 59 12	30 30 45 26 5	67 29 77 7 2	7 14 59 57 49
16 49 20 58 56	75 44 82 68 78	34 55 25 55 37	96 71 4 43 34	21 37 49 68 10
8 73 64 39 27	99 97 54 58 63	98 71 95 15 19	90 55 54 11 34	10 72 30 18 38
85 2 70 67 40	94 74 38 49 33	29 82 94 51 6	8 89 74 42 81	95 25 29 27 0
18 45 98 50 14	3 57 15 14 90	52 60 45 92 97	33 44 90 94 76	95 81 33 17 49
77 27 24 53 8	73 76 28 93 74	49 62 57 47 67	55 47 33 23 3	43 47 19 9 73
43 40 76 93 60	45 2 81 51 24	56 89 90 75 88	1 13 31 66 69	45 60 7 7 76
5 67 50 60 7	69 77 74 54 37	32 28 7 96 40	37 58 57 53 63	73 0 96 7 19
30 35 40 31 60	53 58 76 92 77	86 97 4 13 34	29 59 96 9 75	54 54 85 24 91
38 40 85 73 33	27 79 42 41 54	39 73 48 45 4	32 62 9 1 70	37 75 20 71 31
26 53 35 39 64	82 61 1 55 35	71 77 76 41 17	23 60 78 37 37	61 9 73 92 72
56 83 50 74 40	22 50 35 34 40	35 7 41 34 35	14 66 78 87 83	43 77 88 59 57
37 47 15 8 1	65 9 41 94 52	40 19 62 84 64	43 89 21 77 54	56 94 57 17 72
3 93 15 95 92	40 20 5 92 91	97 99 45 4 43	87 80 30 32 52	96 97 84 7 66
32 66 85 76 53	14 4 51 43 11	69 70 35 32 11	39 91 95 55 55	85 36 5 79 0
82 82 59 19 21	24 71 64 65 81	11 45 14 31 73	97 11 66 62 5	67 87 68 89 20
42 57 30 94 10	98 25 52 45 93	69 16 76 34 62	9 32 93 6 11	69 36 79 37 13
41 56 71 3 9	35 21 28 22 8	74 78 81 76 21	83 3 93 34 37	76 35 43 53 50
20 24 77 27 5	9 21 7 20 52	14 11 1 89 54	22 96 29 26 82	73 94 85 32 0
19 62 31 92 88	76 14 49 65 8	71 69 91 66 86	56 66 50 13 74	55 54 25 78 23
48 40 52 61 27	67 1 4 20 62	52 33 44 51 79	40 45 74 83 59	83 32 80 13 13

在多变量、成对比较或产品/过程搜索研究中并不要求随机，因为其目标是一个精心选择的时间顺序。然而，对部件搜索和变量搜索而言，在第 1 阶段的顺序应当是随机的（即不应首先检验所有好的/所有最好的，然后再检验所有坏的/所有临界的，或者与之相反）。为了确定试验设备的精度，在全析因试验、“B vs. C”研究（详见第 15 章）、散布图和响应曲面方法（详见第 16 章）、散布图/多变量研究中也要求随机化。

方法

下面给出一个关于 2^4 全析因的例子，与 2^2 或 2^3 全析因相比，其过程要简单些。

绿 Y

在选择绿 Y 时，应使用与变量搜索（详见第 13 章）相同的规则，即：寻找一个早期的绿 Y，测量仪器的精度至少 5 倍于该绿 Y 的规范公差，并将绿 Y 的一个属性用李克特度量尺度转换为一个变量。

矩阵

(1) 根据在变量搜索中阐述的同样的方法，选择待研究的 4 个输入变量或因子，将这些因子指定为 A、B、C 和 D。

(2) 为每个因子选择两个水平。第 1 水平标记为 (—)，通常（但不一定）就是该因子的当前水平；第 2 水平标记为 (+)，假定（但不一定）会产生较好的结果。另一个可供选择的实用方法是每天都测量在制造厂中使用的零部件或因子的实际范围，并将最好的零部件标记为 (+)，将一般的零部件标记为 (—)。



(3) 绘制一个表示 16 种组合的矩阵 (见表 14-2), 以使每一

表 14-2

波焊 2⁴ 因子全析因试验

绿 Y: 焊点缺陷数

		(A) 焊剂				
		(A-)A19		(A+)880		
		(B-)速度 4 英尺/分	(B+)速度 6 英尺/分	(B-)速度 4 英尺/分	(B+)速度 6 英尺/分	
C- 角 5°	(D-) 预热 60°F	1 --- 21 (19) 17	3 --- 14 (15) 16	2 +--- 104 (108) 112	4 +++ 8 (8) 8	→150
	(D+) 预热 220°F	9 ---+ 17 (16) 15	11 -++ 64 (61) 58	10 ++- 1 (1) 1	12 +++ 0 (0) 0	→78
	(D-) 预热 160°F	5 --+- 4 (4) 4	7 ---+ 43 (45) 47	6 +-+- 44 (41) 38	8 +++- 3 (3) 3	→93
	(D+) 预热 220°F	13 --++ 32 (33) 34	15 -+++ 14 (13) 12	14 +++ 10 (10) 10	16 ++++ 0 (0) 0	→56
		↓72	↓134	↓160	↓11	

A-=72+134=206 | A-比 A+差 35 个缺陷

A+=160+11=171

B-=72+160=232 | B-比 B+差 87 个缺陷

B+=134+11=145

C-=150+78=228 | C-比 C+差 79 个缺陷

C+=93+56=149

D-=150+93=243

D+=78+56=134 | D-比 D+差 109 个缺陷

(-)A19 焊剂(+)880 焊剂
(-)4 英尺/分(+)6 英尺/分
(-)5° 角(+)6° 角
(-)160° 预热(+)220°

资料来源: Motorola Inc., Seguin, Tx.

因子都用其他因子的每一水平进行检验。

(4) 随机选择检验每一组合或部件的顺序。

(5) 按随机次序表指出的顺序用每一个组合来进行试验，并记录每一部件的绿 Y (输出)。

(6) 使用第二检验顺序的另一个随机次序重复步骤 4 和 5。

(7) 计算每一部件的两个读数的平均值。

(8) 若每一部件的两个读数相差太大 (例如，超出 $\pm 10\%$ 的范围)，则表明该试验的“噪声”太大。可能需要选择其他因子或水平或者需要通过过程确认 (详见第 18 章)，使这些不可控制的 (噪声) 因子更加受控。

(9) 不要仅根据某个部件的最优结果就轻率地作出结论，这样做会忽略其余 15 个部件的有用数据。通过比较所有最佳 (+) 部件和所有一般 (-) 部件，就会得到一个放大的效应，它容许在每个因子的适当水平及主效应和交互影响效应间的关系有更高置信度。

(10) 对于 32 组读数，将每个部件中 A 为 (-) 的平均部件读数相加，并将每个部件中 A 为 (+) 的平均部件读数相加。因为所有其他因素 (B、C 和 D) 都是相互平衡的，也就是说，它们对两个水平都加上或减去了一个常数，因而没有改变它们的初始差别，所以说，A (-) 和 A (+) 之间的差别是由因子 A 引起的。与此类似，将每个单元中 B 为 (-) 的平均部件读数相加，并将每个部件中 B 为 (+) 的平均部件读数相加。它们的差别是由因子 B 引起的。用类似的方法可得到 C (-) 和 C (+)、D (-) 和 D (+) 平均数之间的差别。

(11) 构造一个方差分析 (ANOVA) 表，在必要时，还应构造一个图形分析来表示交互影响效应 (将在下面的案例研究中阐述)。

案例研究及小组练习

案例研究 1: 摩托罗拉波焊

15 年前, 一个指导工作小组承担了波焊工艺的改进工作, 该小组的成员仅参加过一次作者关于试验设计的讲习班。当时的可焊性缺陷水平高达 10 000ppm。该小组雄心勃勃, 其目标是将缺陷水平降低到 200ppm 以下, 即改进高达 50 : 1。实际上, 在 20 世纪 80 年代早期, 世界上还没有一个波焊工艺的缺陷水平能低于 10 000ppm~15 000ppm。

该小组决定进行具有 4 个因子、每个因子 2 个水平的全析因试验, 具体如下:

因子代码	因子描述	因子水平	
		(-)	(+)
A	焊剂类型	A19	A880
B	传送带速度 (英尺/分)	4	6
C	倾角	5°	7°
D	预热温度 (°F)	160	220

* (-) 水平表示当前水平。

全析因矩阵 (见表 14—2) 给出了结果。在“方法”一节的“矩阵”中简要地阐述了该过程。结果表明, 表示当前工艺水平的所有 4 个 (-) 因子都要比这些 (+) 因子差。然而, 为了确定交互影响效应, 必须使用在第 13 章变量搜索介绍中给出的方法构造一个充分的方差分析 (ANOVA)。方差分析表的结构 (见表 14—3) 如下:

表 14-3

方差分析表：波焊 24 因子全析因试验

单元组	因子 (主效应)				两因子 交互作用				三因子 交互作用				四因子 交互作用				输出
	A	B	C	D	AB	AC	BC	AD	BD	CD	ABC	ABD	BCD	ABCD			
1(1)	-	-	-	-	+	+	+	+	+	-	-	-	-	+	19		
2a	+	-	-	-	-	-	+	-	+	+	+	+	-	-	108		
3b	-	+	-	-	-	+	-	+	-	+	+	-	+	-	15		
4ab	+	+	-	-	+	-	-	-	-	+	-	-	+	+	8		
5c	-	-	+	-	+	-	-	+	+	-	+	-	+	+	4		
6ac	+	-	+	-	-	+	-	-	+	-	-	+	-	+	41		
7bc	-	+	+	-	-	-	+	+	-	-	-	+	+	-	45		
8abc	+	+	+	-	+	+	+	-	-	-	+	-	-	-	3		
9d	-	-	-	+	+	+	+	-	-	-	-	+	+	+	16		
10ad	+	-	-	+	-	-	+	+	-	-	+	-	-	+	1		
11db	-	+	-	+	-	+	-	-	+	-	+	-	+	-	61		
12adb	+	+	-	+	+	-	-	+	+	-	-	+	-	-	0		
13cd	-	-	+	+	+	-	-	-	-	+	+	-	-	+	33		
14acd	+	-	+	+	-	+	-	+	-	+	-	-	+	-	10		
15bcd	-	+	+	+	-	-	+	-	+	+	-	-	+	-	13		
16abcd	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0		
总和	35	87	79	79	211	47	33	189	115	35	73	139	127	181	39		

红 X
焊剂+速度

粉红 X
焊剂+预热

(1) 在“部件组”栏中，输入部件组合。约定如下：

- 部件 1 (1) 表示所有因子 A, B, C 和 D 都是 (-)。
- 部件 (a) 表示除 A 外所有因子均为 (-)。
- 部件 (b) 表示除 B 外所有因子均为 (-)。
- 部件 (ab) 表示除 A 和 B 外所有因子均为 (-)，等等。

(2) 在“主要因子”栏中，分别为因子 A, B, C 和 D 输入适当的符号 (-) 和 (+)。

(3) 在“输出”栏中，输入记录在表 14-2 中的每个部件的输出平均值。

(4) 在“两因子交互作用”栏中，将部件 1 (1) 中 A 和 B 的符号相乘，并将其积的符号记录在 AB 栏中。这里，因为 A 和 B 的

符号均为(-)，因此其积 AB 的符号为 (+)。与此类似，确定 A 与 C，B 与 C 等的代数积，并将它们记录在每一行的相应的“两因素交互作用”栏中。

(5) 在“三因子交互作用”栏中，将因素 ABC，ABD 等各个三因子的符号相乘，并在每一行相应的“三因子交互影响”栏中记录其积的符号。

(6) 同样，在“四因子交互作用”栏中，将四因子 ABCD 的符号相乘，并在每一行相应的“四因子交互作用”栏中记录其积的符号。

(7) 在“A”栏，将所有 A 为 (-) 的输出相加，所有 A 为 (+) 的输出相加，再将这两个和的差记入最后一行，标记为“主效应和交互影响效应的总和”。因为 A (-) 的和大于 A (+) 的和 35 个点，因此在“A”栏的“总和”行上方放置一个 (-) 号，这表示 A (-) 比 A (+) 要差 35 个点。这与表 14—2 的结果完全相同。

(8) 最后一行以准确的量化形式给出了每一个主要因子以及两因子、三因子和四因子的交互作用在数量 and 方向 (+ 或 -) 方面的影响。

结果如表 14—3 所示，表明在这 4 个主要因子中，几乎没有一个因子像两因子 AB 交互作用 (红 X)、两因子 AD 交互作用 (粉红 X) 和 - 因子 BCD 交互作用 (浅粉红 X) 那样重要。

最后一步是绘制主要交互影响效应图 (见图 14—1)，其步骤如下：

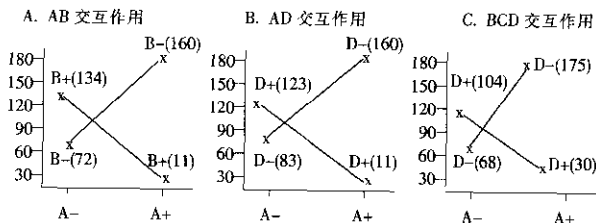


图 14—1 波焊试验：交互影响效应的图形描述

(1) 对于 AB 交互作用, 联合观察 A 栏和 B 栏。将 A 和 B 均为 (—) 的所有部件的输出相加; 将 A 为 (—) 和 B 为 (+) 的所有部件的输出相加; 将 A 为 (+) 和 B 为 (—) 的所有部件的输出相加; 将 A 和 B 均为 (+) 的所有部件的输出相加。

(2) 对于 AD 交互作用, 联合观察 A 栏和 D 栏, 并重复步骤 1。

(3) 对于 BCD 交互作用, 联合观察 B 栏、C 栏和 D 栏, 并重复步骤 1。

(4) 绘制 AB 交互作用的 4 个拐角——A—、B—、A+、B+; A+、B—; A+、B+ 的结果 (如图 14—1A 所示)。

(5) 与此类似, 绘制 AD 和 BCD 交互作用的结果 (分别如图 14—1B 和图 14—1C 所示)。

表 14—3 和图 14—1 的说明

我们现在解释上述结果的物理意义。

- 在因子 A (焊剂类型) 中, 880 焊剂 (新焊剂) 优于 A19 焊剂 (旧焊剂)。

- 在因子 B (传送带速度) 中, 快速 (6 英尺/分钟) 优于慢速 (4 英尺/分钟)。

- 在因子 C (倾角) 中, 大倾角 (7°) 优于小倾角 (5°)。

- 在因子 D (预热温度) 中, 220°F 优于 160°F 。

然而, 在这 4 个主要因子中, 没有任何一个因子像下列几个交互影响效应那样重要:

- 焊剂和传送带速度间的交互作用是红 X, 表示 A 和 B 应控制在其 (+) 水平。这在表 14—3 中得到了证实——若 A 和 B 均为 (—) 将会多出现 211 个缺陷, 且图 14—1 表明它们具有非常强的交互影响效应。(两条平行线表示没有交互影响; 两条非平行线表示有明确的交互影响; 两条交叉线, 如图 14—1A, B 和 C 所示, 表示非常强的交互影响。) 为得到最佳的结果 (最低的缺陷), A 和 B 应保持在 (+) 水平。

●与此类似，AD 交互作用是粉红 X，若 A 和 B 均为（-），则会多出现 189 个缺陷。这可以由图 14—1B 中的强交叉线得到证实。为得到最佳的结果（最低的缺陷），A 和 D 应保持在其（+）水平。

●类似地，BCD 交互作用是浅粉红 X，若 B、C 和 D 转为（-），则会多出现 181 个缺陷。这可以由图 14—1C 中的强交叉线得到证实。为得到最佳的结果（最低的缺陷），B、C 和 D 应保持在（+）水平。

确认试验：求交运算

使 A、B、C 和 D 保持其（+）水平，在 17 块新型电路板 L 上进行求交运算。每块电路板上 有 800 处连接，其缺陷率降低至 220ppm，尽管没有完全达到该小组的目标（即 50：1 的改进），却也取得了 45：1 的改进。在 20 世纪 80 年代早期，无论在美国还是在欧洲和日本，都没有人在波焊工艺上达到如此低的缺陷水平。（然而，目前摩托罗拉公司通过进行更多的试验设计工作，可以使缺陷率稳定在 10ppm 以下，并且还有一些其他公司也已取得了类似的成功。）

该小组继续对该工厂的其他 12 台波焊机进行类似的参数修改。他们减少了 20 名修理员及相同数量的质检员，每年节省开支 75 万美元。该小组的组长在工作中逐步迷上了试验设计，每周都主动进行 1 至 2 次试验设计。他原来是一名维修技师，后来晋升为拥有 1 800 名员工的得克萨斯工厂的首席试验员和督导。

小组练习 1：瓷漆产量

在一家大型设备生产厂，一种瓷漆工艺的瓷漆产量，在 68%~82% 之间波动，令人不太满意。该工厂的工程师们认为，在一年之中瓷漆的产量因时间的不同而发生很大的变化——在冬季较高，在夏季较低。工厂经理甚至对冬季的较高产量也不满意。他估计在瓷

漆产量上每改进一个百分点，每年就能节省 10 万美元。于是邀请作者来帮助解决这一长期困扰工厂的顽固问题。在将解决问题方向集中在过程确认（详见第 18 章）上之后，工厂成立了一个工作组来进行试验设计。由于该工作组深信该工厂的较高温度和较高湿度是造成瓷漆产量下降的原因，于是，小组决定进行一个具有两种温度水平和两种湿度水平的简单 2^2 全析因（详见图 14-2）。（出于保密需要，恕不说明精确的水平。）

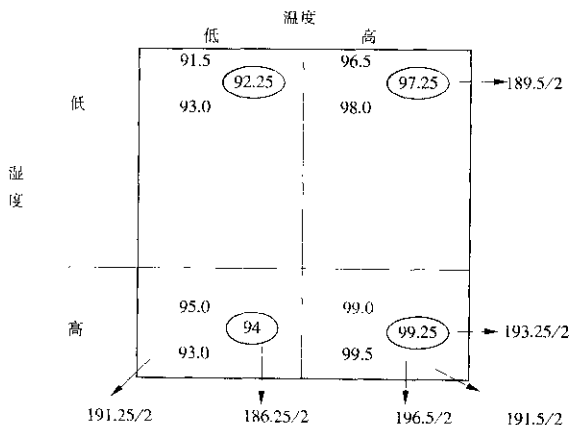


图 14—2 瓷漆产量： 2^2 全析因

问题

- (1) 温度和湿度及其交互影响效应的贡献是什么？
- (2) 在试验中信号与“噪声”的贡献是什么？
- (3) 该工作组关于温度和湿度的初始意见得到确认了吗？
- (4) 真正的红 X 是什么？
- (5) 可以从中得出什么结论？吸取什么教训？



答案

$$(1) \text{ 温度 (T) 的贡献} = \frac{196.5 - 186.25}{2} = 5.125\%$$

$$\text{湿度 (H) 的贡献} = \frac{193.25 - 189.5}{2} = 1.87\%$$

$$\text{交互作用 (T} \times \text{H) 的贡献} = \frac{191.5 - 191.25}{2} = 0.125\%$$

(2) 在每个部件中缺少可重复性，其变化在 0.5%~2% 之间，所有产量均超过了 90%。因此信噪比非常高且十分可以接受。为了在技术上做到正确无误，尽管在 2^2 全析因中的每个部件均应有 4 个（而非 2 个）读数以便充分重复，但是，一次重复就已经得到了高的信噪比，使我们能跳过最后一步。

(3) 尽管就温度和湿度这两个因子而言，该工作组对其重要性的初始意见是正确的，但就研究方向而言，其选择是完全错误的。实际上，较高的温度能得到较高而不是较低的瓷漆产量，并且较高的湿度在某种程度上也能得到较高而不是较低的瓷漆产量。此外，温度对提高瓷漆产量的贡献几乎是湿度的 3 倍。

(4) 令该小组惊讶的是，在这 4 种温度和湿度的组合中，没有一个组合的瓷漆产量低于 90%。他们此前从未得到过如此之高的瓷漆产量。因此，真正的红 X 是一个过程确认技术的建立（详见第 18 章）。过程确认技术明确所有的外部质量因子，包括设备预防性维修、校准、监控、清洁、保养等，该方法优于 2^2 全析因试验。

(5) 从这个案例中得到的经验是：首先，让零部件和工艺（而不是工程师）说话；其次，在开始一项试验之前，确保良好的生产习惯（例如，维修、校准和清洁等）已经就位，以确保这些不可控制的噪声因子保持最小并且不会影响信号的纯洁。

后记

该工厂的工程师们不愿相信工作小组的结果。他们说：“等到夏季完全到来时，瓷漆产量将会再次降低。”然而，经过一个夏季

的实践，瓷漆产量从未低于过 90%，该工厂节省开支高达 150 万美元。该工作小组的新目标是通过进一步的试验设计研究达到 99% 的瓷漆产量。

案例研究 2：最小化波纹效应

在某新型计算机显示器的开发过程中，设计小组必须最小化屏幕的波纹效应，即所谓的扫描波纹，它是一种由阴影图案和水平线图案造成的波浪型图案。

期望的技术规格（绿 Y）是：波纹效应的最大尺寸为 1mm。在显示器样机上，波纹效应的尺寸读数超过了 2mm。扫描波纹取决于水平扫描频率，且可以因选择了适当的水平扫描频率而减小。然而，显示器必须在某个范围的扫描频率上运行。

根据电路分析的结果，选择了如下所示的 4 个因子进行 2^4 全析因试验，以确定最小化扫描波纹的最佳组合。

因子编码	因子描述	因子水平	
		(-)	(+)
A	电阻 (R_1)	22r	5 r 6
B	电阻 (R_2)	470r	220r
C	B+电压 (V)	7	6
D	Q1 晶体管增益 (H_{fe})	12.6	16.9

表 14—4 给出了一个在高水平扫描频率下运行的、具有两个绿 Y（波纹尺寸和集电器电流 I_c ）的全析因矩阵。表 14—5 和表 14—6 分别给出了在中水平扫描频率和低水平扫描频率下运行的、具有两个绿 Y（波纹尺寸和集电器电流 I_c ）的类似的全析因矩阵。

表 14—4

全析因矩阵：波纹尺寸——高频
2 个绿 Y：波纹尺寸和集电器电流 (I_c)

112.5k/90 Hz 试验号	因子主效应				测量	
	A	B	C	D	波纹尺寸 (mm)	I_c (mA)
1	-1	-1	-1	-1	2	195.94
2	1	-1	-1	-1	0	182.99
3	-1	1	-1	-1	1.5	192.95
4	1	1	-1	-1	0	170.74
5	-1	-1	1	-1	2.5	195.98
6	1	1	1	-1	0.5	182.71
7	-1	1	1	-1	2	193.20
8	1	1	1	1	0	170.27
9	-1	-1	-1	1	2.5	198.98
10	1	-1	-1	1	0	185.11
11	-1	1	-1	1	4	203.05
12	1	1	-1	1	0	186.66
13	-1	-1	1	1	30	201.58
14	1	-1	1	1	0	185.16
15	-1	1	1	1	40	200.87
16	1	1	1	1	0	185.94

在高频下得出的测量数据

资料来源：Philips Corp., Chungli, Taiwan.

表 14—5

全析因矩阵：波纹尺寸——中频
2 个绿 Y：波纹尺寸和集电器电流 (I_c)

106.2k/85 Hz 试验号	因子主效应				测量	
	A	B	C	D	波纹尺寸 (mm)	I_c (mA)
1	-1	-1	-1	-1	1.5	186.26
2	1	-1	-1	1	0	173.68
3	-1	1	-1	-1	1	183.75
4	1	1	1	-1	0.5	167.14
5	-1	1	1	-1	1.5	185.71
6	1	-1	1	1	0	173.58
7	-1	1	1	-1	1.5	184.90
8	1	1	1	-1	0.5	166.85
9	-1	-1	-1	1	2	188.65
10	1	-1	-1	1	0	177.18
11	-1	1	-1	1	2	191.89
12	1	1	-1	1	0	179.59

续前表

106.2k/85 Hz 试验号	因子主效应				测量	
	A	B	C	D	波纹尺寸 (mm)	I_r (mA)
13	-1	-1	1	1	2.5	192.67
14	1	-1	1	1	0	184.91
15	-1	1	1	1	2.5	192.87
16	1	1	1	1	0	178.17

在中频下得出的测量数据

资料来源: Philips Corp., Chungli, Taiwan.

表 14—6 全析因矩阵: 波纹尺寸——低频
2 个绿 Y: 波纹尺寸和集电器电流 (I_r)

106.2k/70 Hz 试验号	因子主效应				测量	
	A	B	C	D	波纹尺寸 (mm)	I_r (mA)
1	-1	-1	-1	-1	1.5	120.06
2	1	-1	-1	-1	1	114.76
3	-1	1	-1	-1	1	119.80
4	1	1	-1	-1	0.5	114.01
5	-1	-1	1	1	1.5	120.75
6	1	-1	1	1	1	114.89
7	-1	1	1	1	1	120.95
8	1	1	1	-1	0.5	113.91
9	-1	-1	-1	1	3	117.11
10	1	-1	-1	1	0.5	117.13
11	-1	1	1	1	3	118.81
12	1	1	-1	1	0	118.13
13	-1	-1	1	1	3	118.67
14	1	-1	1	1	0.5	117.43
15	-1	1	1	1	2.5	118.45
16	1	1	1	1	0	110.90

在低频下得出的测量数据。

资料来源: Philips Corp., Chungli, Taiwan.

表 14—7、表 14—8 和表 14—9 分别给出了在上述三种水平扫描频率下只有一个输出或绿 Y (波纹尺寸) 的方差分析 (ANOVA) 表。为了简明扼要地说明问题, 我们没有给出集电器电流的方差分析表。当集电器电流为绿 Y 时, 红 X、粉红 X 等的结果类似于波纹尺寸为绿 Y 时的结果。



表 14—7

方差分析表：波纹尺寸——高频

分组	因子 主效应				两因子 交互作用						三因子 交互作用				四因子 交互作用		波纹 尺寸
	A	B	C	D	AB	AC	BC	AD	BD	CD	ABC	ABD	ACD	BCD	ABCD		
1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1	2.0	
2	1	-1	-1	-1	-1	-1	1	-1	1	1	1	1	1	-1	-1	0.0	
3	-1	1	-1	-1	-1	1	-1	1	-1	1	1	1	-1	1	-1	1.5	
4	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1	0.0	
5	-1	-1	1	-1	1	-1	-1	1	1	-1	1	-1	1	1	-1	2.5	
6	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1	0.5	
7	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1	2.0	
8	1	1	1	-1	1	1	1	-1	-1	-1	1	-1	-1	-1	-1	0.0	
9	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1	1	-1	2.5	
10	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	0.0	
11	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	4.0	
12	1	1	-1	1	1	-1	-1	1	1	-1	-1	1	-1	-1	-1	0.0	
13	-1	-1	1	1	1	-1	-1	-1	-1	1	1	-1	-1	-1	1	30.0	
14	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	1	-1	-1	0.0	
15	-1	1	1	1	-1	-1	1	-1	1	1	-1	-1	-1	1	-1	40.0	
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.0	
影响	-84	10	65	68	-11	-64	8	-69	13	62	-9	-12	-63	9	-8		

红 X

粉红 X

资料来源：Philips Corp., Chungli, Taiwan.

表 14—8

方差分析表：波纹尺寸——中频 (106.2k/85Hz)

分组	因子 主效应				两因子 交互作用						三因子 交互作用				四因子 交互作用		波纹 尺寸
	A	B	C	D	AB	AC	BC	AD	BD	CD	ABC	ABD	ACD	BCD	ABCD		
1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1	1.5	
2	1	-1	-1	-1	-1	-1	1	-1	1	1	1	1	1	-1	-1	0	
3	-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	0.5	
5	-1	-1	1	-1	1	-1	-1	1	1	-1	1	-1	1	1	-1	1.5	
6	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1	0	
7	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1	1.5	
8	1	1	1	-1	1	1	1	-1	-1	-1	1	-1	-1	-1	-1	0.5	
9	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1	1	-1	2	
10	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	0	
11	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	2	
12	1	1	-1	1	1	-1	-1	1	1	-1	-1	1	-1	-1	-1	0	

续前表

因子主效应																两因子交互作用								三因子交互作用				四因子交互作用		波纹尺寸
分组	A	B	C	D	AB	AC	BC	AD	BD	CD	ABC	ABD	ACD	BCD	ABCD															
13	-1	-1	1	1	1	-1	-1	-1	-1	1	1	-1	-1	-1	1	2.5														
14	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	1	-1	-1	0														
15	-1	1	1	1	-1	-1	1	-1	1	1	-1	-1	-1	1	-1	2.5														
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0														
影响	-14	0.5	1.5	2.5	1.5	-1.5	0.5	-4.5	-0.05	0.5	-0.5	-1.5	-0.5	0.5																

红 X

粉红 X

资料来源: Philips Corp., Chungli, Taiwan.

表 14-9

方差分析表: 波纹尺寸—低频

	因子 主效应				两因子 交互作用						三因子 交互作用				四因子 交互作用		波纹 尺寸
分组	A	B	C	D	AB	AC	BC	AD	BD	CD	ABC	ABD	ACD	BCD	ABCD		
1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1	1.5	
2	1	-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	1	
4	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1	0.5	
5	-1	-1	1	-1	1	-1	-1	1	1	-1	1	-1	1	1	-1	1.5	
6	1	-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	1	
8	1	1	1	-1	1	1	1	-1	-1	-1	1	-1	-1	-1	-1	0.5	
9	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1	1	-1	3	
10	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	0.5	
11	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	3	
12	1	1	-1	1	1	-1	-1	1	1	-1	-1	1	-1	-1	-1	0	
13	-1	-1	1	1	1	-1	-1	-1	-1	1	1	-1	-1	-1	1	3	
14	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	1	-1	-1	0.5	
15	-1	1	1	1	-1	-1	1	-1	1	1	-1	-1	-1	1	-1	2.5	
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	
影响	-13	-3.5	-0.54	5	0.5	0.5	-0.5	-8.5	0.5	0.5	-0.5	0.5	-0.5	0.5			

红 X

粉红 X

资料来源: Philips Corp., Chungli, Taiwan.

分析

(1) 方差分析表 14-7、表 14-8 和表 14-9 都表明: 对所有



三种水平扫描频率而言,因子 A (电阻 R_1) 均为红 X。因此,为最小化波纹尺寸,因子 A 应保持在其 (+) 水平。即使未给出集电器电流 (I_c) 的方差分析表,但对第 2 个绿 Y (集电器电流 I_c) 而言也是如此。

(2) 对所有三种水平扫描频率而言,AD 交互作用都是粉红 X。AD 交互作用在 (-) 水平而不是 (+) 水平时要高些,通常这表明 A 和 D 都应保持在其 (+) 水平。然而,在上述所有三种频率下 AD 交互作用的图形结构 (为简明起见,未在本书中给出) 表明: D (-) 的坡度没有 D (+) 的陡峭。因此, D 最好保持在其 (-) 水平。

(3) 除了在最高频率下的方差分析表中之外,因子 C 是不重要的。在任何情况下,在 C (-) 下的波纹尺寸要比在 C (+) 下的波纹要小些。这已由一些图形结构 (为简明起见,未在本书中给出) 得到了证实。

(4) 因子 B 对上述三种频率而言都不重要,可以保持在比较经济的水平上。

因此,最优的水平是 A (+)、B (+)、C (-) 和 D (-)。

结果

在最优水平上的求交运算表明: 波纹尺寸都令人满意地小于 0.5mm (改进大于 40:1), 且 C_p 高于 3.0。

对案例研究 2 最小化波纹效应的评述

(1) 即使该项目涉及一个新型设计, 仍有足够的数量来证明使用部件搜索或成对比较能更容易且更快捷地得到结果。

(2) 在上述矩阵的 16 个部件中, 一个都没有重复。然而, 对方差分析表中的重要因子进行的图基检验表明, 对 A 和 AD 因子 (红 X 和粉红 X) 而言, 其合计终结计数分别是 16 和 7, 刚好都大于最小值 6, 其置信水平大于 90%。

(3) 绘制一个散布图 (见第 16 章) 来确定 2 个绿 Y (波纹尺

寸和集电器电流)之间的相关性。结果表明,这2个绿Y在高水平扫描频率和中水平扫描频率时有非常好的相关性,而在低水平扫描频率时则没有相关性。

小组练习 2: 钻孔作业

某金属加工厂需在零件的两面十字交叉钻孔,每孔的孔径均为9/64英寸,通常会造成50%以上的零件出现严重的出口毛刺和较大的附着层。甚至在每一个零件都用高压空气进行处理之后,所有发送货物仍都被客户退回。因此,工厂对可能造成问题的4个关键因子设计了一个 2^4 全析因。

因子代码	因子描述	因子水平	
		(-)	(+)
A	转速(转/分)	2 850	5 000
B	钻头类型	GT100	Gehring
C	钻头进刀(英寸/转)	0.001	Gold
D	冷却剂类型	水溶雾	切削油

绿 Y

将有无毛刺及有无附着层这些属性转换成利克特度量尺度0~4。每个部件的缺陷得分分为各种缺陷数与其利克特度量尺度之积的和。

	尺度
最小毛刺——非缺陷	S0
轻微凸起的毛刺	S1
凸起的毛刺及一个容易脱落的附着毛刺	S2
凸起的毛刺及两个以上的附着毛刺	S3
过多的悬挂毛刺及多个附着层	S4

在每个部件中试验100件零件,并以随机的次序进行试验。为消除钻具的磨损,在每次实验中都使用一个新的钻具。表14-10给出了其全析因矩阵,表14-11给出了其方差分析表,图14-3

是其交互作用的图形。

表 14—10

全析因：钻孔试验

		(A) 主轴速度							
		(A-)2850 R.P.M.				(A+)5000 R.P.M.			
		(B-) GT 100		(B+) Guhring Gold		(B-) GT 100		(B+) Guhring Gold	
(C) 转速	(C-) 0.001	1 9 S1:37 S2:42 S3:3 S4:2 Tot: (168)		2 3 S1:9 S2:26 S3:54 S4:7 Tot: (251)		3 14 S1:42 S2:35 S3:3 Tot: (121)		4 5 S1:18 S2:30 S3:50 Tot: (228)	
	(D+) 油	5 11 S1:49 S2:41 S3:1 S4:2 Tot: (142)		6 10 S1:1 S3:1 S4:96 Tot: (396)		7 7 S1:34 S2:42 S3:6 S4:1 Tot: (140)		8 6 S1:3 S2:7 S3:46 S4:44 Tot: (128)	
	(D-) 油雾	9 8 S1:50 S2:43 S3:7 Tot: (157)		10 13 S1:15 S2:35 S3:45 S4:4 Tot: (236)		11 12 S1:53 S2:17 S3:3 S4:0 Tot: (96)		12 15 S1:62 S2:30 S3:2 Tot: (128)	
	(D+) 油	13 2 S1:22 S2:7 Tot: (36)		14 16 S1:35 S2:49 S3:10 S4:1 Tot: (167)		15 4 S1:5 S2:1 Tot: (7)		16 1 S1:11 Tot: (11)	

注：每个部件角上的第 1 个方块表示部件号；每个部件角上的第 2 个方块表示执行试验的次序。

问题

- (1) 每个部件都需要 100 个零件吗？
- (2) 将缺陷分成 0~4 级有什么优点呢？
- (3) 根据单一部件的结果，要使用的最佳水平是什么呢？
- (4) 根据方差分析表，什么是红 X 和粉红 X 呢？
- (5) 对每个因子而言，要使用的最优水平是什么呢？
- (6) 交互作用的图形结构肯定了这些选择吗？



表 11-11

全析因：钻式试验——方差分析表

单元组	因子主效应				两因子交互作用				三因子交互作用				四因子交互作用		输出
	A	B	C	D	AB	AC	BC	AD	BD	CD	BCD	ABD	BCD	ABCD	
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	168
2	1	-1	-1	-1	-1	-1	-1	1	-1	1	1	1	-1	-1	251
3	-1	1	-1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	121
4	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	228
5	-1	-1	1	-1	-1	-1	-1	1	1	-1	-1	1	-1	-1	142
6	1	-1	1	-1	-1	-1	-1	-1	-1	1	-1	1	1	1	396
7	-1	1	1	-1	-1	-1	-1	1	-1	-1	1	-1	-1	1	140
8	1	1	1	-1	-1	-1	-1	-1	-1	-1	1	-1	-1	-1	331
9	-1	-1	-1	1	1	1	1	-1	-1	-1	1	1	1	-1	157
10	1	-1	-1	1	-1	-1	-1	1	-1	1	-1	1	1	1	236
11	-1	1	-1	1	-1	-1	-1	-1	-1	1	-1	-1	-1	1	96
12	1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	-1	-1	128
13	-1	-1	1	1	1	-1	-1	-1	1	-1	-1	-1	-1	1	36
14	1	-1	1	1	-1	-1	-1	1	-1	1	1	1	-1	-1	167
15	-1	1	1	1	-1	-1	-1	-1	1	-1	-1	-1	1	-1	7
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	11
主效应和交互影响	-491	881	939-155-213-217-389	-13-279	-637	-135	-167	-231	+7						
效应的贡献															

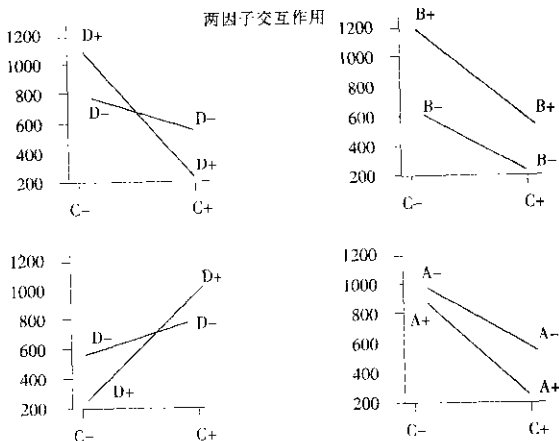


图 14—3 全析因：钻孔试验

答案

(1) 如果缺陷率较低，例如 5% 或者更低，那么，每个部件 100 件是合理的。若缺陷率为 50% 且经过利克特度量尺度的放大，那么，每个部件 10 件就够了。

(2) 缺陷分级有下述两点优点：第一，缺陷的得分范围从 7 放大到 396。正如在答案 1 中描述的那样，这可以减少每个部件的样本规模。第二，如果要分开某一特定类型的缺陷（绿 Y），例如，单独考虑毛刺或单独考虑附着层，那么，该矩阵可以很容易地分开这两个绿 Y。

(3) 根据一个单个部件，其最佳水平应为 A (+)、B (-)、C (+) 和 D (+)。由于它最终产生于方差分析表，因此，该组合即为最佳组合。然而，充分的方差分析练习还是必要的，以免丢失

其他 15 个部件中的有用数据。

(4) 在方差分析表的最后一行，具有最高读数的因子是 C，其最佳水平应为（-）的对立面，即（+）；具有次高读数的因子是 B，且其最佳水平应为（-）。

(5) 在方差分析表的最后一行，根据读数和符号，最优水平应为 A（+）、B（-）、C（+）和 D（+）。

(6) 图形结构可解释如下：在 CD 交互作用中，C（+）、D（+）可产生最少的缺陷。尽管这可能表明应使用 D（-），因为 D（-）的坡度没有 D（+）的坡度陡峭，但在实际生产中，D（冷却剂）偏移到另一种冷却剂并没有任何危险。类似地，在 BC 交互作用中，因为坡度最低，C（-）、B（-）可产生最少的缺陷。在 BD 交互作用中，B（-）、D（+）可产生最少的缺陷，并且选择 D（+）的理由与在 CD 交互作用中的一样。在 AC 交互作用中，A（+）、C（-）可产生最少的缺陷，且两者的坡度大致相等，表示交互作用很小。

结论

在进行全析因之前，人们认为较低的钻头进刀速度会好一些，但是，试验的结果却与人们的期望完全相反。在较高的钻头进刀速度、GT100 型钻头、较高的转速和切割油冷却剂的组合下，缺陷率降至 0。

◆ DOE 小组的问题

(1) 为了减少噪声因子和增加信噪比，在进行全析因试验之前，是否执行了所有好的生产实践（例如，设备的预防性维修、仪器校准、控制变化的环境、定期清洁等）？

(2) 坚持了平衡、重复和随机性原则吗？

(3) 每个因素的（-）水平是其当前水平吗？（-）水平可以认为是对当前水平进行了一定的改进吗？

(4) 在析因矩阵中的每个部件是否都被重复或重现? 其检验次序是随机的吗?

(5) 每个部件的两个读数具有合理的一致性(即具有较高的信噪比)吗? 如果在试验中存在 16 个部件, 对重要的参数进行了图基检验吗?

(6) 对 B vs. C 变量搜索、散布图、正向控制和过程确认还存在后续措施吗?

管理问题

(1) 在试验中, 为了确保良好的生产实践和高的信噪比, 在进行全析因试验之前需要使用过程确认技术(详见第 18 章)吗?

(2) 为什么选择全析因方法而不选择变量搜索技术?

(3) 为试验选择因子的依据是什么?

(4) 是否使用了方差分析表技术和交互影响效应的图形结构?

(5) 根据质量、费用和寿命时间改进, 试验的结果是什么?

(6) 有哪些教训可供以后的设计小组参考?





第 15 章

B vs. C：卓越的确证技术

► 确认的重要性

在所有解决问题的实践中，一个主要缺点是，一旦得到了改进，便认为该改进是本质的、真实的和持久的。然而，在解决问题的实践的历史中，最初取得成功而后来很快失败的例子俯拾皆是，且从不试图回到改进前的条件，并观察问题是否重新出现。

确认改进的程度及持久性的最佳方法是：回到改进前的产品、工艺或方法，并确保问题再现。然后，转用新改进的产品、工艺或方法，并确保问题“归零”。至少重复两次上述过程。这就像打开电灯的开关一样，复现问题，问题“归零”；再复现问题，



问题再“归零”。只有这样，试验人员才能得到关于改进持久性的足够的置信水平。

刹车片案例研究

下面的例子将说明进行这种验证的重要性和原因。一家生产飞机刹车片的厂商曾经历一次产品质量的危机。该厂商的客户几乎取消了其供货商的资格。但该厂商决定重整旗鼓。它通过了由外部合同方举办的一个“黑带”培训项目。黑带小组设计了一个经典的分析因试验，不但解决了该问题，并且计划每年节省 500 万美元的开支。然而，在引入改进的一个月之后，该问题又非常明显地出现了！该小组设计了一个后续的验证试验，结果发现该解决方法既不稳定又不持久。

于是该厂商匆忙邀请作者解决这个问题。我们放弃了分析因方法，完全从头开始，首先使用成对比较法，然后使用产品/过程搜索法，最后用“B vs. C”来确认改进的持久性。

什么是 B vs. C?

符号 B 和 C 分别代表两种不同的产品、过程、方法或商业政策/实践，其中，C (current) 表示当前产品，B (better) 相应地代表更好的产品。但 B 和 C 可以是两个互换的过程、方法或政策。B vs. C 的任务是确定哪个更优。

B vs. C 是一种非参数型比较试验方法，无需假设 B 或 C 产品或过程满足正态分布。术语非参数指在第 12 章中论述的第二种数据——秩评定数据。非参数秩幂是指仅需要极小的样本量（通常是 3 个 B 和 3 个 C），就可以比较两种产品、过程或方法，并以非常高的置信度确定哪种产品、过程或方法更好。欲了解 B vs. C 在详细的解决问题路径图中的位置，可参见图 7—3。

目标

B vs. C 应被用作一种验证工具，而不是一种初始的解决问题工具。该技术的基本目标是：使用非常小的样本量，根据质量和可靠性，以 90% 或更高的置信水平在两个产品或过程之间确定哪个更好。此外，还有如下其他目标：

(1) 以 90% 或更高的置信度，预测一个特定产品或过程比另一个产品或过程要好多少。

(2) 保证对原有产品或过程的改进的持久性。

(3) 在两种产品或过程中选择一种较好的产品或过程，即使其质量没有什么改进，也会有某些其他的实际利益（例如，费用或寿命周期时间等）。

(4) 同时评价多个产品、过程、材料（B, C, D, E 等）。

(5) 将 B vs. C 扩展到人类尝试的任一个领域。

(6) 在初步调查中用作一般的民意测验工具。

原则

1. 过程能力研究。过程能力研究通常用于比较两个过程（一个当前过程 C 及一个可能更好的过程 B），方法是对每个过程运行 50~100 个部件，其结果可能是图 15-1 中所示的 4 种频率分布之一。

- 在图 15-1A 中，B 和 C 没有差别，这称为零假设。

- 在图 15-1B 中，尽管 B 优于 C，但其改进存在一些不确定性，在其重叠区域，有些 C 部件实际上要优于某些 B 部件。这可能是粉红 X 或浅淡粉红 X 改进。

- 在图 15-1C 中，最差的 B 部件都等于最好的 C 部件。这是一个红 X 改进。



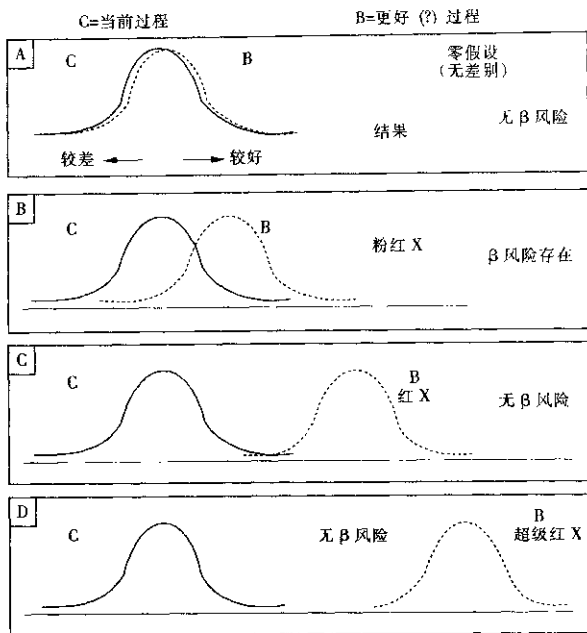


图 15—1 过程 B 和 C 的 4 种分布

●在图 15—1D 中，最坏的 B 部件都大大优于最好的 C 部件。这是一个超级红 X 改进。

然而，研究这样一个具有 100~200 个部件的过程能力需要很长的时间、巨大的经费和人力开支。

2. 六组合检验。本书阐述的所有试验设计，都以简单性、图形化和有效性为目标，还以小样本量为目标。在 B vs. C 中，我们



论及“六组合检验”(six-pack test)——3个横坐标和3个纵坐标!当然,这只是一个便于记忆的叫法,即一种简单方法——以记住在B vs. C检验中最常见的样本量。3个B和3个C即六组合检验。

3. 组合律。传统的统计学指出,从总体中抽取样本的绝对规模决定了样本描述总体的接近程度。样本的绝对规模(而非总体的百分比)越大,就越接近于总体。那么,3个B和3个C这样小的样本量能够以何种接近程度来描述两个总体中的每一个总体呢?答案在于组合律的能力,公式如下:

$$\text{组合数}^* = \frac{n!}{(n-r)! r!}$$

式中, n 为部件总数; r 为一种类型的部件总数; $(n-r)$ 为老类型的部件总数。

使用3个B和3个C时,组合数 $= \frac{6!}{3! 3!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1 \times 3 \times 2 \times 1} = 20$ 。

因此,在不重复相同次序的组合时,有20种排列3个B和3个C的方法(如下所示)。

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
↑ 最佳	B	B	B	B	B	B	B	B	B	B	C	C	C	C	C	C	C	C	C	C
	B	B	B	C	C	C	C	C	C	C	B	B	B	B	B	B	C	C	C	C
	B	C	C	B	C	B	B	C	C	C	B	B	B	C	C	C	B	B	B	C
↓ 最差	C	B	C	B	B	C	C	B	B	C	B	C	C	B	B	C	B	B	C	B
	C	C	C	C	B	B	C	B	C	B	C	B	C	B	C	B	B	C	B	B
	C	C	B	C	C	C	B	C	B	B	C	C	B	C	B	B	C	B	B	B

所以,如果在B vs. C检验中以降序(从最佳到最差)排列3个B和3个C,那么,在这20种组合中只有惟——种组合是3个B完全排列在3个C之上,这将主要取决于偶然性,完全取决于命运的随意摆布。也就是说,只有1/20的可能性观察到所有3个B均优于3个C的改进是假的,即实际上没有任何改进。即使3个B均

* 原文为:组合数 $= \frac{n!}{(n-r)! r!}$,有误,应为:组合数 $= \frac{n!}{(n-r)! r!}$ 。——译者注



超过 3 个 C，没有任何改进的可能性为 $1/20$ （即风险为 5%）。反过来说，如果 3 个 B 超过 3 个 C，那么，B 对 C 进行了一定改进的可能性为 $19/20$ （即置信度为 95%）。

4. α 和 β 风险。试验人员可能根据试验结果与实际（但未知的）情况作出下述两种决策：

根据实际试验作出的决策	实际（但未知的）情况	
	B 优于 C	B 与 C 相同
B 似乎优于 C	是	否
	对	I 型错误 α 风险
B 似乎与 C 相同	II 类错误 β 风险	对

如果试验表明 B 优于 C（3 个 B 均排列在 3 个 C 之前），那么，试验人员将拒绝关于 B 与 C 相同的零假设（ H_0 ）。如果 B 实际上真的优于 C（实际但未知的情况），那么，该试验人员作出了正确的决策。然而，如果实际上没有任何改进，那么，试验人员就犯了 I 类错误，其可能性称为 α 风险。

如果试验表明了一个混合的排序，其中并非所有 3 个 B 均优于 3 个 C，即 B 与 C 相同，那么，试验人员将不拒绝零假设。如果 B 实际上并不优于 C，那么，该试验人员作出了正确的决策。然而，如果 B 实际上真的优于 C，那么，试验人员就犯了 II 类错误，其可能性称为 β 风险。

因此，定义 α 风险为拒绝零假设（即在没有任何改进时假定存在改进）的风险；定义 β 风险为接受零假设（即当存在一个期望的改进时假定没有改进）的风险。

更简单地讲， α 风险用于检查 B 是否优于 C； β 风险用于检查 B 对 C 的改进程度是否充分。在图 15—1 中，只在图 15—1B 中给出了一个 β 风险，其中存在一个重叠区域且改进存在不确定性。图



15-1C 和图 15-1D 没有重叠区域, 也就没有 β 风险。图 15-1A 具有确定性, 但既没有改进, 也没有 β 风险。

方法

根据组合律, 表 15-1 给出了一个确定 B vs. C 检验的样本尺寸及适当的 α 和 β 风险的权威性指南。

表 15-1 B vs. C 样本量及 α , β 风险 (基于无重叠规则)

α 风险: 拒绝零假设 (无差别) 的风险。

β 风险: 在存在改进时接受零假设的风险。

错误决策的后果	随 机 检 验 (集) 总数		K 值 (均差) β 风险		
			0.50	0.10	0.05
α 风险 置信水平 0.001 0.999	B_s 2	C_s (43)	3.0*	4.0	4.3
			3.9*	5.1	5.5
超关键	3	16	2.5	3.6	3.9
			3.2	4.5	5.0
● 许多人员死亡	4	10	2.3	3.4	3.8
			2.9	4.3	4.8
● 食物中毒	5	8	2.2	3.4	3.7
			2.9	4.3	4.7
● 核反应堆	6	6	2.2	3.3	3.7
			2.8	4.2	4.7
α 风险 置信水平 0.01 0.99	2	13	2.3	3.4	3.8
			3.0	4.4	4.6
关键	4	7	2.0	3.2	3.6
			2.6	4.1	4.6
● 数名人员死亡	4	5	2.0	3.1	3.5
			2.5	4.0	4.5
● 高达 1 亿美元	5	4	2.0	3.1	3.5
			2.5	4.0	4.5

续前表

错误决策的后果	随 机 检 验 (集) 总数		K 值 (均差) β 风险		
			0.50	0.10	0.05
α 风险 置信水平 0.05 0.95					
重要	1	19	2.5 / 3.2	3.6 / 4.6	3.9 / 5.0
● 高达 10 万美元	2	5	1.7 / 2.2	3.0 / 3.8	3.4 / 4.3
	3	3	1.6 / 2.0	2.9 / 3.7	3.3 / 4.2
	4	3	1.7 / 2.2	3.0 / 3.8	3.4 / 4.3
α 风险 置信水平 0.10 0.90					
中等	1	9	2.1 / 2.6	3.2 / 4.1	3.6 / 4.6
● 高达 1 万美元	2	3	1.4 / 1.8	2.7 / 3.5	3.2 / 4.0
	3	2	1.4 / 1.8	2.7 / 3.5	3.2 / 4.0

* 在上面的值中, $\sigma_B = \sigma_C$ 。

** 在下面的值中, $\sigma_B \neq \sigma_C$ (通常, $\sigma_B = 1.5\sigma_C$ 或 $\sigma_C = 1.5\sigma_B$)。

资料来源: Copyright 1988 Red X Technologies, Inc.

风险与置信水平

选择适当的风险或置信水平, 风险的补就是置信水平。

● 中等重要性—高达 1 万美元的费用—0.10 或 10% 的风险 (90% 的置信水平)。

● 较高重要性—高达 1 万美元的费用—0.05 或 5% 的风险 (95% 的置信水平)。

● 关键重要性—高达 1 亿美元的费用及数名人员的死亡—0.01 或 1% 的风险 (99% 的置信水平)。

● 超关键重要性—许多人员的死亡—0.001 或 0.1% 的风

险 (99.9%的置信水平)。

B vs. C 的样本量

在每一个置信水平的基础上, 选择适当的 B 和 C 产品/过程样本数目。因为 C 是当前过程, 因此很可能用于检验的 C 部件要多于较新的 B 部件。

然而, 在工业界, 大多数 B vs. C 试验的一个简单的准则是使用六组合检验。该检验使用 3 个 B 和 3 个 C, 且具有 5% 的 (风险 (或 95% 的置信水平))。其优点有: (1) 易于记忆 (3 和 3); (2) 检验的部件总数最少。

随机化的重要性

正如在第 14 章中阐述的那样, 一个非常重要的预防措施就是随机安排检验的顺序。由于不可控制的因子或噪声因子没有等同的机会安排或者不安排试验, 因此, 先安排 3 个 C 然后安排 3 个 B 的逻辑次序可能会得出存在偏见的结果。

排列次序

使用可变数据, 或者至少使用李克特度量尺度将收集到的 B 和 C 的属性数据转换成可变数据。避免缺陷百分比或收益百分比。然后按最佳到最差的顺序排列部件 B 和 C。

决策规则

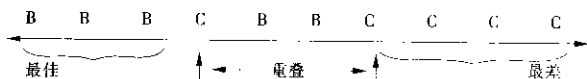
B vs. C 的决策方法有下述两种:

- **无重叠规则。**该规则不允许有重叠。更适当的叫法应为无重叠终结计数, 其 3 个 B 必须全部超过 3 个 C (或者 3 个 C 必须



全部超过 3 个 B)。在 95% 的置信限下, 判断 B 比 C 好 (或者 C 比 B 好)。但该无重叠规则必须事先 (即在试验之前) 使用。如果只有 1 个或 2 个 B 超过 C (即如果在排列中存在重叠), 那么, 既不能认为 B 优于 C, 也不能认为 C 优于 B。若使用小的样本量且存在重叠, 则没有足够的置信水平表明 B 是对 C 的改进。

● **重叠/终结计数规则。**这就是著名的图基检验, 该检验已在与成对比较和产品/过程搜索的关系中进行了详细的论述。在某些工业场合, 允许某些重叠并仍然保持非常低的错误决策风险也许更可取。例如, 如果 5 个 B 和 5 个 C 的排列如下:



B 终结计数=3, C 终结计数=4, 且总终结计数=7。其中存在一个 (为 3 的) 重叠。但是, 再一次值得指出的是, 该重叠规则必须事先 (即在试验之前) 使用, 而且 B 和 C 的样本数要大些 (每种 6 个或者更多)。B 的样本数 (即 n_B) 既可以小于也可以大于 C 的样本数 (即 n_C), 但 n_B 与 n_C 的比率不应大于 3:4 (或 4:3)。正如在成对比较和产品/过程搜索中论述的那样, 其风险和置信水平是相同的。

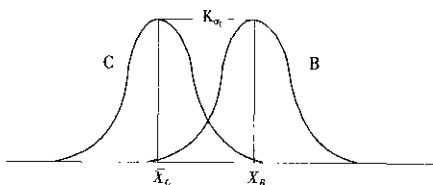
风险 (α)	置信水平	终结计数 ($\geq$)
0.10	90%	6
0.05	95%	7
0.01	99%	10
0.001	99.5%	13

β 风险: 确定 B 对 C 的改进程度

仅知道 B 是否优于 C 是不够的。无论是从经济的角度还是从技

术的角度讲，在试验之前就能预测出实际的改进程度都是很有用的。与改进程度相关的风险通过 B 部件均值与 C 部件均值之间的间隔来估计。该间隔或德耳塔 (Δ) 距离以 $K\sigma_c$ 为单位进行度量，其中 K 是事先由试验人员选择的一个约定的数值， σ_c 是 C 部件的标准差 (B 部件的数量可能太少，而无法精确地确定它的标准差 σ_B)。表 15-1 给出了 β 风险为 0.5、0.1 和 0.05 (或置信水平为 50%、90% 和 95%) 时的 K 值。上面的 K 值假定 $\sigma_B = \sigma_c$ ，下面的 K 值假定 $\sigma_B \neq \sigma_c$ (也可以是其中一个大于另一个)。

例如，如果检验了 3 个 B 和 3 个 C，且所有 3 个 B 都优于所有 3 个 C，那么，就有 95% 的置信水平表明 B 优于 C。但是，究竟好多少呢？为回答这个问题，可在表 15-1 的右边查出风险为 0.1 (即置信水平为 90%) 时在 3B、3C 行中的 K 值。假定 $\sigma_B = \sigma_c$ 。从表 15-1 查出的 K 值为 2.9。若 C 部件的均值为 $\bar{X}_C$ ，B 部件的均值为 $\bar{X}_B$ ，则间隔： $\bar{X}_C - \bar{X}_B$ 应等于或大于 $2.9\sigma_c$ ，以得到该改进将得到检验的置信水平为 90% (或者说该改进没有得到检验的风险只有 10%)。图 15-2 用图示的方式说明了这种情况。



图中： $\bar{X}_C$ 为 C 产品/过程的均值；

$\bar{X}_B$ 为 B 产品/过程的均值；

σ_c 和 σ_B 分别是 C 和 B 部件的标准差；

如果 $\sigma_B = \sigma_c$ ，置信水平为 90% 时， $K = 2.9$ ；

如果 $\sigma_B \neq \sigma_c$ ，置信水平为 90% 时， $K = 3.7$ ；

$K\sigma_c$ 是 $\bar{X}_C - \bar{X}_B$ 要求的最小距离。

图 15-2 β 风险：确定 B 对 C 的改进程度

案例研究与小组练习

案例研究 1：破裂的环氧化物

在第 13 章关于变量搜索的案例研究 2 中，我们描述了一种情形的成功结果：在当时，黏结底座和盖罩的环氧化物在固化处理期间破裂。红 X 为固化处理温度，在 150℃ 时产生零缺陷，而在 120℃ 时其缺陷率为 10%。此外，尽管环氧化物的含量及其黏性都不重要，却可以减少其容量和黏性以降低原材料的成本。

为了保证高固化处理温度的改进是持久的，并观察在环氧化物的含量和黏性降低时出现的环氧化物破裂，对 3 个 B 部件和 3 个 C 部件进行了 B vs. C 确认试验，如下所示：

参数	B 过程	C 过程
固化处理温度 (°C)	150°	120°
环氧化物的含量 (导柱覆盖率)	60%	75%
环氧化物的黏性 (CPS)	150 000	180 000

如前所述，绿 Y 是环氧化物破裂的李克特度量尺度，0 表示没有破裂，10 表示 100% 破裂，其得分为：加权缺陷 = 缺陷部件数 × 李克特度量尺度。

3 个 B 和 3 个 C (以随机次序执行) 的排列为：

部件类型	缺陷数	缺陷类型李克特度量尺度	加权缺陷得分
B	0	0	0
B	0	0	0
B	0	0	0
C	1	2	2
C	1	4	4
C	1	5	5

这证实 B 过程优于 C 过程，且置信水平为 90%；它还证实降低环氧化物的含量和黏性不会降低性能。



小组练习 1: 电线结合强度

为确定一个工艺改变 (B) 同当前工艺 (C) 相比是否能改进 IC 芯片中电线结合强度 (通过一个校准的拉力试验), 进行了一个 B vs. C 试验。在开始试验之前已经确定了下述内容:

- (1) α 风险为 5%;
- (2) β 风险为 5%;
- (3) 不允许重叠;
- (4) B 的标准差为 3.0gm;
- (5) 假定 σ_B 等于 σ_C ;

3 个 B 检验后跟 3 个 C 检验的结果 (gm.) 为:

B: 225, 223, 219

C: 217, 212, 210

问题

- (1) B 工艺优于 C 工艺吗? 置信水平是多少?
- (2) 假定重新进行检验, 且检验顺序为: B1、C2、C1、B3、C3 和 B2, 那么 B 工艺优于 C 工艺吗? 置信水平是多少?
- (3) 计算 $\bar{X}_C$ 和 $\bar{X}_B$ 之间的距离。
- (4) 工艺 B 对工艺 C 的改进程度足够吗? 置信水平是多少?

答案

(1) 否, 否, 否! B vs. C 检验顺序并不随机, 因此, 该检验无效且完全无用。

(2) 既然检验是以随机次序进行的, 它就是有效的, 且 3 个 B 优于 3 个 C, 其置信水平为 95% (5% 的风险)。

(3) $\bar{X}_C = 222.3$; $\bar{X}_B = 213$; 因此, $\bar{X}_C - \bar{X}_B = 9.3$ 。

(4) 最小距离 $\bar{X}_C - \bar{X}_B = K_{\alpha}$ 。因为 $\sigma_B = \sigma_C = 3$ 且 β 风险为 5%, $K = 3.3$, 因此, $K_{\alpha} = 3.3 \times 3.0 = 9.9$ 。因为实际距离 9.3 小于要



求的最小距离 9.9, 因此, B 工艺对 C 工艺的改进程度对置信水平 90% 来说是不充分的。

案例研究 2: 冲压成型

在第 13 章变量搜索案例研究 1: 冲压成型中, 为确保改进 (包括金属板材弯度和夹持) 是否确实持久, 进行了一个 B vs. C 确认试验。绿 Y 是尺寸公差, 以前它无法保持在英寸以内, 通过变量搜索后一直保持在英寸以内。B vs. C 检验以随机次序进行, 并以标称值偏差与 0.001 英寸的积来表示。选择的 α 风险为 5% (置信水平为 95%)。因为某些 C 的结果要优于 B, 并可能出现重叠, 因而使用重叠方法 (图基检验), 其样本量分别为 12 个 B 和 13 个 C。以随机次序得到的结果如图 15-3 所示。

C		
2.	5.	5. 4, 6, 9, 7. 1, 6. 4, 8, 2. 1
B		
2.	0.	1. 2, 2, 2, 1. 1, 0, 1, 2, 2
以降序排列, 结果如下:		
0	B	B 终结计数=3
0	B	
0	B	
1, 1, 1	B	1, 1C
2, 2, 2, 2, 2	B	2, 2C
重叠		

	4, 4C	C 终结计数=9
	5, 5C	
	6, 6C	
	7 C	
	8 C	
	9 C	

图 15-3 案例研究 2: 冲压成型

注: 合计终结计数=3+9。因此, B 过程是对 C 过程的改进, 其置信水平为 99.5% (大于要求的 95%)。

增加 B 和 C 的样本量, 以降低预测 B 对 C 改 进程度的不确定性

在采取 3 个 B 部件和 3 个 C 部件的无重叠决策中, 如果 3 个 B 都超过了 3 个 C, 则 B 产品优于 C 的置信水平为 95%。反过来, 如果 3 个 C 都超过了 3 个 B, 那么 C 优于 B 的置信水平为 95%。然而, 若终结计数小于 6 (即并非所有 3 个 B 都优于所有 3 个 C), 那么, 改进只能用 $\bar{X}_B$ 和 $\bar{X}_C$ 的距离来估计。对于 95% 的置信水平 (5% 的 β 风险) 且假定 $\sigma_B = \sigma_C$, 在表 15-1 中其 K 值为 3.3, 因此, 要求 $\bar{X}_B$ 和 $\bar{X}_C$ 的最小距离为 $3.3 \bar{X}_C$ 。但在实际中这种最小距离可能太宽, 因此, 要求的 K 值 3.3 太大。

当 $\bar{X}_B$ 和 $\bar{X}_C$ 的实际距离较小时, 为了恢复相同的置信水平, 必须减小要求的 K 值。如果每一组 3 个 B 和每一组 3 个 C 都增加到 2~10 个部件, 并且可以计算出每一个 B 和每一个 C 的均值, 那么便可做到这一点。(B 和 C 之间的随机只在顺序, 而非在每一个 B 或 C 之中。) 这表明一共可能需要 6~30 个 B 和 6~30 个 C, 选择样本数需要根据要求的 $\bar{X}_B$ 和 $\bar{X}_C$ 的距离来权衡。需要的样本量由中心极限定理确定:

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}$$

式中, n 为要求的样本量; $\sigma_{\bar{X}}$ 和 σ_X 分别为样本和总数的标准差。

重写上面的公式为:

$$\sqrt{n} = \frac{\sigma_X}{\sigma_{\bar{X}}} \text{ 或 } n = \frac{\sigma_X^2}{\sigma_{\bar{X}}^2} \quad (\text{方程 1})$$

用 K 替换 σ 得该集合的规模

$$n = \frac{\text{表中的 } K^2}{\text{要求的 } K^2} \quad (\text{来自表 15-1}).$$

$\bar{X}_B$ 和 $\bar{X}_C$ 的最小距离为 $K\sigma_{\bar{X}}$ (见表 15-1)。

对于 95% 的置信水平, K (来自表 15-1) 为 3.3, 因此最小距离为 $3.3\sigma_{\bar{X}}$ 。如果需要一个较小的距离, 且 K 减小到 2:

$$\text{由方程 1: } n = \frac{3.3^2}{2^2} = \frac{10.89}{4} = 2.72 \approx 3$$

这表明 3 个 B 中的每一个都应增加到 3 个部件, 即总共 9 个 B 部件, 且 3 个 C 中的每一个都应增加到 3 个部件, 即总共 9 个 C 部件, 以得到一个较低的 K 并允许一个较低的要求的 $\bar{X}_B - \bar{X}_C$ 。

表 15—2 是根据方程 1 构造出来的。它表明要减小 K 值并因此得到较小的距离, 应增加每一个 B 和 C 的样本量。该表基于 β 风险为 5% (或置信水平为 95%) 建立。当 β 风险为 10% 或者 $\sigma_B = \sigma_C$ 时, 其样本量略有不同。

表 15—2 增加样本量以减小 $\bar{X}_B - \bar{X}_C$

要求的最小距离 $\bar{X}_B - \bar{X}_C$	3 个 B 和 3 个 C 中每一个需要的样本量 (n)
$K=3.3$	1B, 1C
$K=2.33$	2B's, 2C's
$K=2.0$	3B's, 3C's
$K=1.65$	4B's, 4C's
$K=1.47$	5B's, 5C's
$K=1.34$	6B's, 6C's
$K=1.26$	7B's, 7C's
$K=1.16$	8B's, 8C's
$K=1.10$	9B's, 9C's
$K=1.03$	10B's, 10C's

例如, 在关于电线结合强度的小组练习 1 中, $\bar{X}_C$ 和 $\bar{X}_B$ 间的最小距离 9.9 得不到满足。如果每一个 B 和每一个 C 的样本量都增加到 2, K 将从 3.3 减少到 2.33, 要求的最小距离应为 $2.33 \times 3 = 6.99$, 那么实际距离 9.3 对 95% 的置信水平将是足够的。

修改确定 C 劣于 B 的检验

正如在本章及前一章中讨论的那样, 只要好与坏、(+) 与 (—) 或 B vs. C 的比率基本相等且不超过 3 比 4 (彼此的规模不超过 25%), 那么, 图基检验基本上与样本量无关。

该图基检验是一个排列两者的检验，即要么 B 可能劣于 C，要么 C 可能劣于 B。然而，假若所关心的仅仅是确保 B 不劣于 C，那么，可以修改图基终结计数，以便只反映 C 劣于 B 的情况。因此，可能还需使用组合定律和乘法定理。

例如，如果有 4 个 B 和 4 个 C，那么最劣部件是 C 的概率为 $\frac{4}{8}$ ，次劣部件还是 C 的概率是 $\frac{4}{8} \times \frac{3}{7}$ ，第 3 劣部件仍是 C 的概率是 $\frac{4}{8} \times \frac{3}{7} \times \frac{2}{6}$ ，所有 4 个较劣部件全是 C 的概率为：

$$\frac{4}{8} \times \frac{3}{7} \times \frac{2}{6} \times \frac{1}{5} = 0.014^*$$

因此，其风险为 0.014 或 1.4%，置信水平为 98.6%，可进行类似的计算得到表 15—3。

表 15—3

C 劣于 B 的最小终结计数
(B vs. C 的样本比率为 1:1)

置信水平	C 的数目	C 的最小底部终结计数
关键的	5~6	5
99%	7~19	6
	20~∞	7
重要的	3	3
95%	4~15	4
	16~∞	5

扩展到多于 2 个备选 (多于 B 和 C)

至此本章仅涉及两种选择或者两种备选——一个 B 和一个 C，但是比这两种选择要多的产品/过程/方法可能会同时进行评价。例如，B，C，D，E，F，G 等，可以使用图基检验方法的扩充，并增

* 此处原文误为 “ $\frac{4}{8} \times \frac{3}{7} \times \frac{1}{5} \times = 0.014$ ”，应为 “ $\frac{4}{8} \times \frac{3}{7} \times \frac{2}{6} \times \frac{1}{5} = 0.014$ ”。——译者注

加每一个 B, C, D, E 检验的样本数。

根据该图基检验扩充, 表 15—4 给出了为保证适当的置信水平所需要的最小终结(顶部和底部)计数总数:

- 两级置信水平: 99% 和 95%;
- 2、3、4、5 和 6 种选择, 即从 B 和 C 直至 B 到 G;
- 每种选择的不同样本量 n 。

表 15—4 多于 2 种选择 (即 B、C、D、E 等) 的排列次序

置信水平	选择/备选数 (B、C、D、E 等)									
	2		3		4		5		6	
	n	最小 EC	n	最小 EC	n	最小 EC	n	最小 EC	n	最小 EC
99%	6~14	9	4~9	12	4	14	4	17	3	18
	15	10	10~12	13	5	15	5	18	4	20
			13	14	6	16	6	19	5	23
					7	16	7	21	6	24
					8	17	8	24	7	26
					9	18	9	24	8	28
95%					10	19	10	25		
	4	6	4	9	3	12	3	13	3	16
	5~13	7	5~7	10	4, 5	13	4	15	4	18
			8~10	11	6, 7	14	5	16	5	20
			11~13	12	8	15	6	18	6	23
					9	16	7	19	7	25
					10	17	8	21	8	27
							9	22		
							10	24		

小组练习 2: 在 4 个供货商之间选择

某生产厂为选择一家供货商对其装配线中使用的一个关键部件进行表面精密加工而评价 4 家供货商 B, C, D 和 E。为此进行一个 B vs. C 与 D 与 E 试验, 并要求其置信水平达到 95%。于是从每一供货商那里选择 4 个部件, 并建立如表 15—5 所示的排列次序。在 1, 2, 3 和 4 列列出了 4 种可能的结果。

表 15—5

4 家供货商的排序：小组练习

		第 1 栏	第 2 栏	第 3 栏	第 4 栏最佳
最佳	1	B	B	B	B
	2	B	C	E	B
	3	E	D	B	B
	4	B	D	E	C
	5	B	D	C	B
	6	C	B	D	E
	7	D	C	B	E
	8	C	D	E	C
	9	E	B	D	C
	10	E	C	C	E
	11	C	B	D	D
	12	E	E	C	E
	13	D	E	D	C
	14	C	C	B	D
	15	D	E	C	D
最差	16	D	E	E	D
合计终结计数					

问题

- (1) 计算在每一栏中的最佳供货商的终结计数。
- (2) 如果结果如第 1 栏所示，将会得出什么结论？
- (3) 如果结果如第 2 栏所示，将会得出什么结论？
- (4) 如果结果如第 3 栏所示，将会得出什么结论？
- (5) 如果结果如第 4 栏所示，将会得出什么结论？

答案

如表 15—4 所示，对 95% 的置信水平和 4 种选择而言，其最小终结计数为 13。

(1) 如表 15—6 第 1 栏所示，2 个 B 在顶部，11 个 C、D 和 E 在底部，因此 B 从顶部排列的终结计数为 13 (2+11)，所以供货商 B 显然要比供货商 C、D 和 E 更好。又如表 15—6 第 1 栏所示，



2个D在底部, 6个非D在顶部, 因此D的终结计数为8, 不足以说明D为最差供货商。

(2) 在第2栏中, 1个B在顶部, 5个非B在底部, 因此B从顶部排列的终结计数为6, 不足以说明B为最佳供货商。在第2栏中, 2个E在底部, 11个非E在顶部, 因此E从底部排列的终结计数为13, 所以E显然是最差供货商。

(3) 在第3栏中, 各个供货商的终结计数分别为: $B=1+2$, $C=5+1$, $D=5+3$, $E=1+1$, 因此没有一个供货商的终结计数接近最小终结计数13。所以, 这4个供货商没有什么显著的差别。

(4) 在第4栏中, 3个B在顶部, 11个非B在底部; 10个非D在顶部, 3个D在底部。因此B从顶部排列的终结计数为14, D从底部排列的终结计数为13。所以, B显然是最佳供货商, D显然是最差供货商。

表 15-6

4家供货商的排序: 答案

第1栏		第2栏		第3栏		第4栏	
最佳	1	对B的终结计数	B	对B的终结计数	B	对B的终结计数	B
	2		B		F		B
	3	对D的终结计数	E	对E的终结计数	D		B
	4		B		E		C
	5		B		C		B
	6	对B的终结计数	C		D		E
	7		D		C		E
	8		C		D	对B的终结计数	C
	9		E		B		C
	10		E		C		E
	11		C		B		D
	12		E		E		E
	13		D	对D的终结计数	E		C
	14	对D的终结计数	C	对E的终结计数	C		D
	15		D		E		D
最差	16		D		E		D
合计终结计数:		对B: $2+11=13$ 对D: $6+2=8$		对B: $1+5=6$ 对E: $11+2=13$		对B,C,D,E: 0	对B: $3+11=14$ 对D: $10+3=13$

一种特殊的 B vs. C 失效力试验

尽管 B vs. C 是确认工业问题解决方法的改进程度的一种通用技术，然而，在下述 3 种情况下，前面阐述的传统 B vs. C 技术可能会不适当：

(1) 当难以将一个属性绿 Y 转换成一个变量，甚至使用李克特度量尺度也难以转换时。例如，若一个产品或过程具有许多种缺陷模式（例如，7~20 种），乃至无论是单独还是共同地将一个李克特得分与这些缺陷连接起来都不现实时，这些缺陷就必须组合成一个总的缺陷百分比或收益百分比。在这种情况下，3 个 B 和 3 个 C，甚至为每 1 个 B 和每 1 个 C 选择更大的样本都可能不足以检验出两种这类缺陷百分比或收益百分比之间的差异。

(2) 当 B 和 C 产品或过程的缺陷水平都较小（例如，50ppm 及以下），3 个 B 和 3 个 C，甚至为每 1 个 B 和每 1 个 C 选择更大的样本都可能不够敏感，以至无法检验出 B 和 C 之间的差异时。

(3) 当该问题在工厂里并未表现为一个质量问题，而在交付客户后经过客户数月的使用逐步发展成一个可靠性问题时。对于上述三种情况，一种特殊的技术就是开发一个称为失效力试验的绿 Y。该技术是一个称为多重环境强化应力试验 (MEOST) 的强大可靠性工具的组成部分。MEOST 可靠性工具已在第 3 章中介绍，并将在第 22 章中详细阐述。

简要地讲，它表示增加应力，例如，3 个 B 和 3 个 C，使用 1 个或多个环境或应力并继续增加超出设计极限的应力，直到所有 B 和所有 C 失效为止。因此，绿 Y 便成为失效力或失效前时间。与 3 个 C 相比，如果所有 3 个 B 都要求更高的应力才失效或者更长的时间才失效，那么 B 产品或过程就优于 C 产品或过程，其置信水平为 95%。类似地，改进的程度也可以估计出来。当然必须要认真考虑保证失效模式切合实际，同时又是现场长期的失效模式，而不应该是在现场不可能发生的模拟的失效模式。

这里也许可以充分证明因便宜而引进的新材料 B 并不比旧材料 C 差。

案例研究 3：传动列车发动机

一家厂商生产在铁路支线车辆上使用的设备，该设备在驱动传动列车的发动机具有非常低的失效率：每年 120ppm。该厂商希望更换另一家供货商生产的发动机，该马达符合当前供货商生产的发动机的所有规格。此计划每年能节省开支 225 000 美元，因而从经济的角度出发很有吸引力。但是，工程师们关心的是在铁路支线车辆通常运行的、崎岖不平的环境中经过 6~12 个月的使用后的可靠性。对该厂商的客户而言，当前的失效率 120ppm 是完全可以接受的。

由于现有发动机的失效率较低，并需要几个月的时间来检验失效，常规的 B vs. C 评价不能满足实际要求。因此，安排了一个置信水平高达 99% 的 B vs. C 检验，并选择了 5 台现有 (C) 发动机和 4 台新型 (B) 发动机。当所有 9 台发动机同时经受多种环境（应力），其中包括振动、撞击、极端温度、湿度和灰尘时，将绿 Y 确定为失效前时间。这些应力逐步增强，慢慢超过设计应力，直到在所有 9 个部件上都观察到失效为止。由于所有 9 台发动机都在相同的时间和相同的多种环境场所进行试验，因此不需要考虑随机性问题。

根据失效小时数，结果如下：

C 发动机：148, 163, 182, 225, 237。

B 发动机：171, 196, 203, 221。

在排列了 C 和 B 的次序后表明，在可靠性方面 B 发动机不逊于 C 发动机，因而批准使用 B 发动机。B 发动机在现场使用了 2 年后，证实的失效为每年 132ppm，比原先使用 C 发动机时 120ppm 的失效水平有较显著的增加。

B vs. C 技术的工业应用

B vs. C 技术在工业领域中的应用如同其在其他方面的表现一样引人关注。有下述 2 种结果：(1) B 优于 C，并具有一个期望的置信水平（通常为 95% 或更高）；(2) B 不优于 C，但从质量的角度看，B 不逊于 C。表 15—7 列出了每一结果的几种应用。

表 15—7

B vs. C 试验的应用

B 优于 C 时	B 不逊于 C 时
● 更改设计 ● 更改过程 ● 更改制造方法 ● 可靠性、寿命试验 ● 新设备 ● 新供货商/材料 ● 产量改进	● 费用降低 ● 周期时间的改进 ● 可变性降低 ● 安全性 ● 更便利的可制造性 ● 人类工程学（用户友好） ● 空间缩小 ● 环境改善 ● 低成本加工 ● 低成本设备 ● 增加正常运行时间 ● 机器效率 ● 扩大公差 ● 消除一次运行/检验

例如，由本书作者介绍的 B vs. C 的一个非常重要的应用就是他坚持在现行生产中作出的每一次改变都应使用 B vs. C 检验，以确认质量改进或者证明不存在质量的下降。

当然，不应盲目地固守“6 组合检验”（3 个 B 和 3 个 C）。在某些情况下，即使重叠方法需要稍大一些的样本量，它也比非重叠方法更适合些。在某些情况下，增加每 1 个 B 和每 1 个 C 的样本量（至 2~10），可以降低 B 改进程度的不确定性。在某些情况下，只需要证明 C 劣于 B，然后，仅需考虑 C 的底部



终结计数。在某些情况下，候选数可能会从 2（B 和 C）增加到 3~6 个或者更多。最后，失效应力试验可能是 B vs. C 方法在可靠性评价或者强行区分在 B vs. C 中非常小的缺陷水平方面的一个强大的扩充方法。

B vs. C 的管理应用

在产品和制造领域之外，B vs. C 技术的应用几乎是无限的。无论何时评价两个（或更多）备选方案，B vs. C 技术都可以得到较好的应用。表 15-8 中列出的几个应用仅仅是许多应用中的少数几个代表。一些公司将数百万美元花费在广告、市场营销和促销上。在政治领域，则进行了许多关于选民对候选人的选择意向的调查，其样本量从 500~2 000 人不等。在这种情况下，更快捷、更简单和更廉价的 B vs. C 技术可以给出关于趋势和选择意向的早期迹象，在必要时，还可以再继续使用更为传统的方法。

表 15-8 B vs. C 技术在管理服务中的应用

- | |
|---|
| <ul style="list-style-type: none">● 中心组、专家会诊、专家组● 调查：市场、政治、经济、社会、健康● 广告● 促销、销售预测● 工作致富：纵向和横向● 公司政策：工作条件、附加利益、保险、弹性上班制等● 医院：入院、急诊室、护理、结账等● 学校：招生、奖学金、教学方法、学习方法等 |
|---|

案例研究 4：汽车造型

在检验候选产品时，公司的传统做法通常是在检验市场中介绍产品，在报纸、电台和电视上发布广告或者对客户的偏好进行广泛

的调查。尽管上述做法都有其价值，但在花费数百美元甚至数百万美元之前结果是未知的。

某汽车公司决定对其两种汽车造型方法进行初步的 B vs. C 检验。在其中心组评价中，将 2 种候选造型提交给 20 个专家组成员，并要求每 1 位专家组成员按利克特度量尺度从 1 到 10 给这 2 种造型打分，10 为最佳，1 为最差。根据专家组成员的打分，造型 1 优于造型 2，其合计终结计数为 11，因而其置信水平为 99%。

该汽车公司不相信该技术的能力以及得出的高置信水平。在 B vs. C 检验之后又用传统市场研究方法进行了研究。在市场研究上花费了 45 000 美元之后，公司得出了与之相同的结论。

在舆论调查中的 B vs. C 技术：穷人的民意测验

本书作者曾多次在其公开的和非公开的讨论会上检验过在一次势均力敌的公职竞聘中，B vs. C 技术在两位候选人中或者在两个有争议的论点之间进行选择的确程度，其中游戏的成分多于严肃的方法。

通常地，从参与者中随机选出 2~16 人，要求他们对两个候选人或两个论点中的一个或另一个进行秘密投票，并用利克特度量尺度 0~10（10 表示最佳，1 表示最差）表示对所选的候选人的偏爱程度。在数十次这类竞选中，作者成功地预测出获胜的候选人或论点的比例高达 80%。这些练习之所以标称为“穷人的民意测验”，是因为其玩笑的成分多于严肃的民意测验，但它确实显示了小样本数和创新技术的巨大能力。（本人无意将该技术传授给民意调查组织或者赌博组织！）

► 作为 B vs. C DOE 小组指南的问题

- (1) B vs. C 检验仅被用作一个验证技术，以保证用先前的

DOE 技术作出的改进的持久性吗？

(2) 作为试验一个建议的或期望的改进的工具，在不事先使用其他 DOE 技术的情况下，B vs. C 检验的使用令人感到沮丧吗？因为这些期望的改进是以技术人员的意见、直觉、猜测和理论为基础的。给出工程师们猜中根本原因的能力的记录，直接进行 B vs. C 检验可能导致失败，且可能导致多次失败。

(3) 如何确定 α 风险（或相应的置信水平）？

(4) 选择重叠方法，还是非重叠方法？为什么？

(5) 如何确定 B 和 C 部件的适当数目？

(6) 检验 B 和 C 部件的次序是随机的吗？

(7) 在确定 B 是不是对 C 的改进时，还关心其改进程度吗？

(8) 选择何种 β 风险？为什么？

(9) 如果目标是减少 B 对 C 改进程度的不确定性，是否选择了一个适当低的 K 值且相应增加了每一个 B 部件和 C 部件的样本量？

(10) 如果目标是证明 C 劣于 B，是否从适当的表中确定了 C 的数目和 C 的最小终结计数？

(11) 如果存在 2 个以上的选择或备选（例如，B，C，D，E 等），是否使用了适当的表来确定每一个 B，C，D，E 部件的样本量和需要的最小终结计数？

(12) 如果绿 Y 仅可以用产量或缺陷百分比来度量，或者如果 B 和 C 的缺陷率都非常低（500ppm 以下），或者如果问题是一个现场可靠性问题而不是一个工厂质量问题，那么，是否应该使用失效应力来评估 B 对 C 的改进？

(13) 即使对 C 没有任何质量改进，其他经济的/人类工程学的/环境的因子会证明 B 是正当的吗？

(14) 在公司的行政管理和服务领域考虑过进行 B vs. C 研究吗？



评审 DOE 小组项目时，对高层管理提出的问题

- (1) 使用 B vs. C 技术作为初始试验设计的诱感受到过抵制吗？
- (2) B vs. C 工具主要用于验证产品/过程改进的持久性吗？
- (3) 为什么要选择适当的置信水平 ($1-\alpha$ 风险)？
- (4) 如何确定适当的样本量？
- (5) B vs. C 检验的顺序是随机的吗？
- (6) 是否存在两个以上的备选（例如，B，C，D，E 等）需要评价，且使用了适当的技术吗？
- (7) 在诸如缺陷率非常低、估计可靠性差别和一成不变地应用 B vs. C 技术之类的情況下，是否使用了失效应力方法？
- (8) 即使在质量方面 B 不是对 C 的改进，还会因为经济和环境方面的原因而使用 B 吗？
- (9) B vs. C 技术已经应用于白领/服务工作中吗？

第5篇

试验设计优化

第3篇论述了解决问题的5种思路创新技术：多变量、集中图、部件搜索、成对比较和产品/过程搜索。这些技术说明了如何“同部件对话”。这些技术在不干扰生产的同时提供了强大而有意义的思路。第4篇接着阐述了更多的正规的DOE解决问题技术，其目的是描述产品/过程的特征。这意味着区分重要变量与非重要变量。如果要研究的输入变量为5个或5个以上，则可用变量搜索来实现；如果要研究的变量为4个或4个以下，则可用全析因来实现。但这并不是解决问题之路的终点。接下来的一个重要步骤是，确保通过这些方法而取得的改进是持久的——B vs. C测试。

但这仍不是解决问题之路的终点。在解决了问题和产品/过程特性描述之后必须进行产品/过程优化。这意味着为前面进行特征描述的（即重要的）输入变量建立最优目标值、技术规格极限及公差。第5篇将阐述下列几种不同的优化技术：

- 散布图
- 响应曲面方法，包括：
 - ✓ 调优运算（EVOP）
 - ✓ 单纯形法
 - ✓ 随机调优运算（REVOP）
 - ✓ 最速上升法





第 16 章

获得现实的技术规格和公差的 散布图

► 目标

散布图的目标是：

- (1) 建立现实的技术规格和实际的公差；
- (2) 缩小重要变量的公差，以得到较高的 C_{pk} ；
- (3) 放宽不重要变量的公差，以减少开支。该技术在解决问题路径图中的位置如图 7-3 所示。

► 实际的技术规格和公差的重要性

为什么技术规格和公差如此重要？为什

么它们应当是现实的？理由非常简单，那就是成本高昂。表 16—1 列出了与放宽公差有关的不可接受的费用。表 16—2 列出了与缩小公差有关的不可接受的费用。

表 16—1 与放宽公差有关的不可接受的费用

● 失去市场份额	● 竞争劣势
● 客户流失	● 政府罚款
● 法律诉讼	● 环境破坏
● 产品召回	● 服务呼叫率高
● 现场可靠性低	● 保修费用高

表 16—2 与缩小公差有关的不可接受的费用

● 与客户实际需要无关	● 更多的材料评审委员会要求
● 废料多和返工多	● 更高的供货商费用
● 高达 100% 的审查和检验	● 更精密的设备以满足 5:1 的产品
● 更多的分析费用	设备比率
● 更严格的加工设备公差	● 更频繁的设备校准
● 更短的加工设备寿命	● 生产和质量间更多的对抗

至今，仍有些工程师以如下方法建立技术规格和公差：

- (1) 他们凭空给出一些数字，这称为凭空分析！
 - (2) 他们从以前的设计和制图中照搬数字。
 - (3) 他们使用严格的公差来保护自己。同产品不能使用相比，他们不大可能会因严格的公差和较高的费用而受到管理部门的批评。
 - (4) 他们使用已用滥了的、预先“消化过”的公差。
 - (5) 他们盲目地遵循供货商的建议。
 - (6) 他们用最坏的情景来确定公差。
 - (7) 他们并未向客户咨询什么对其重要或者什么对其不重要。
- 这样做的结果是，90% 的技术规格和公差都是错误的。这可以用下



面介绍的一个案例研究来很好地说明。

案例研究：隐形眼镜的边缘缺陷

某隐形眼镜生产厂家认为，应当消除隐形眼镜中的装饰性边缘缺陷。这些缺陷包括在眼镜外围的划痕、凹口、内含物等。用裸眼很难发现这些边缘缺陷，操作人员和检验人员必须使用高倍显微镜来找这些缺陷。为了确保最终产品在交付给客户之前没有这些缺陷，每一副隐形眼镜都必须经过 4~5 个检验站，这是一种蛮干的质量方法！

厂家已为解决这种缺陷花费了数百万美元，但仍正在将其市场份额丢失给其竞争对手。在一次“逆向工程”研究中厂家意外地发现，竞争对手生产的隐形眼镜中存在的边缘缺陷远远多于本厂隐形眼镜的边缘缺陷。厂家随后进行了一次市场调研，以确定隐形眼镜用户为什么喜欢其竞争对手生产的隐形眼镜。研究表明，对隐形眼镜用户来讲，因为这些边缘缺陷只在隐形眼镜的外围，且裸眼发现不了，又不影响视力（这才是隐形眼镜应关注的焦点），因而用户很少关心这些缺陷。竞争对手的产品优势在于：（1）更符合隐形眼镜的处方，因而可以提供更好的视力；（2）佩戴更舒适。该生产厂家很快将其关注焦点从边缘缺陷转移到处方的准确性（得到 2.5 或更高的 C_{pk} ）和佩戴的舒适性上，且在 18 个月后恢复了其原有的市场份额。

原则

● 在最终的分析中，技术规格和公差必须追溯到客户的实际需求。如果客户仅用模糊而又主观的方式来表达这些需求，那么应使用诸如质量功能展开（详见第 3 章）、中心组、专家组和专家会诊之类的更为正式的技术，以量化这些需求。



● 尽管可以使用诸如回归分析或多重回归分析之类的更为复杂的技术，但是，散布图具有所有 DOE 的优点：简单性、图形化和有效性。

● 在散布图中，纵向散布量是所有输入变量而不是所选输入变量对绿 Y 的总贡献的一个度量（参见下面的“方法”一节）。

● 只有在先前的 DOE 试验中（例如，变量搜索或全析因）表明两个或多个重要的输入变量没有强的交互作用时，才应使用散布图。如果在先前的 DOE 试验中（例如，变量搜索或全析因）表明两个或多个重要的输入变量有强的交互作用，那么应使用响应曲面方法（参见第 17 章）。

方法

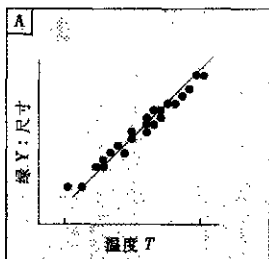
在诸如变量搜索之类的 DOE 研究中，假定一个特定红 X 变量的最佳水平产生一个最佳结果，其临界水平产生一个较差的结果。然而，我们如何知道较好的结果真的就是最佳水平呢？在这个水平的两侧是否还有一个水平能给出更好的结果呢？散布图的目的是微调这种情况并确定红 X 的最佳水平（即设计中心）及其现实的公差，以保证零缺陷。

散布图是一种图形化技术，其中 30 个读数^[1]以一个随机的时间顺序，对照一个相关的 Y 的相应值域，来表示要绘制的一个独立变量值域的 10 个或 10 个以上的水平。如果充分相关——一个窄小的平行四边形，则该独立变量是一个红 X 或强粉红 X，且其最适当的目标值和公差可以在图形上确定。如果没有什么关系——一个宽大的平行四边形，则该独立变量是不重要的，且它的值和公差可以设置为最经济的水平。

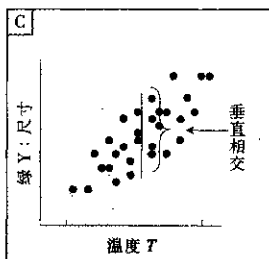
散布图可以用图 16-1 来很好地说明。假定要观察一个独立变量 T（温度）对一个相关变量（绿 Y）尺寸的影响。既然温度是可变的^[2]，比如说以 2℃ 的间隔在 80℃～140℃ 之间变化，并记录相



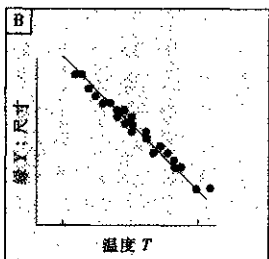
应的尺寸绿 Y。一般说来，绘制散布图的结果将如图 16-1 之 A、B、C 或 D 所示。在散布图 A 中，有非常强的证据表明温度和绿 Y（尺寸）之间存在正（细的铅笔线）相关。除了温度之外没有什么对绿 Y 有影响，温度可以称为红 X 变量。在散布图 B 中也是如此，其中温度也称为红 X 变量，但相关直线的斜率是负的而不是正的。在散布图 C 中，尽管存在相关，但这种相关较弱。如果画出两条平行线，就会得到一个平行四边形的影响。即使温度保持在某个常



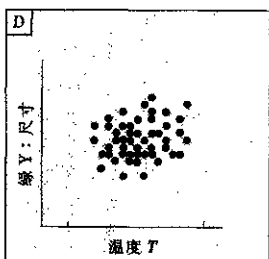
正相关
一个红 X 条件



不清楚的正相关
一个粉红或浅粉红条件



负相关
一个红 X 条件



不相关

图 16—1 典型的散布图



数，例如 110°C ，在绿 Y 中也有变化。因为温度保持在某个常数，因此，这些变化不可能来自温度。因此，在散布图 C 中的垂直相交表示除了温度以外的所有输入变量对绿 Y 的贡献。垂直相交越大，输入变量（即温度 T ）的贡献就越小。散布图 D 表明根本没有任何相关。

图 16-2 给出了散布图的其他几个例子。散布图 A 表示充分相关，但不是线性相关。相关不一定非得是线性的。如果只考虑图形中心部分，则散布图 B 表示弱相关，但如果将输入变量的范围扩展到两边，则散布图 B 表示适度相关。如果考虑所有的数据点，则散布图 C 表示不相关。但似乎是三个分开的和分层的相关，以表示三种影响。

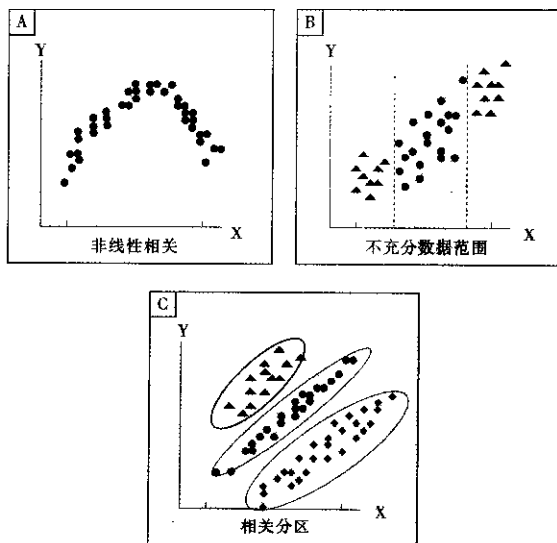


图 16-2 罕见的散布图



测验问题 1

图 16—3A 和图 16—3B 所示为一个输入变量 X 及其相关绿 Y 的两种不同的散布图。在何种情况下存在相关?

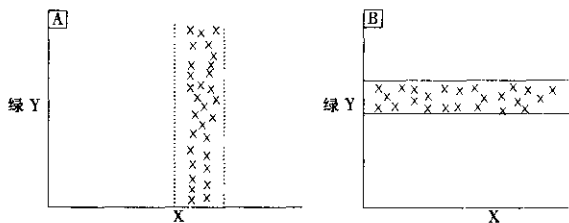


图 16—3 垂直的与水平的散布图

测验问题 2

与图 16—4B 相比, 图 16—4A 中 X 对绿 Y 的影响更大吗?

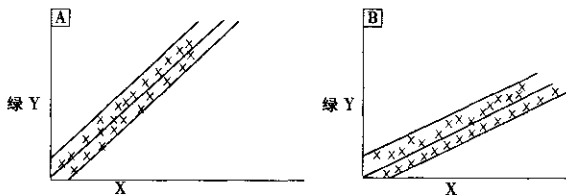


图 16—4 不同斜角的散布图



测验问题答案

(1) 无论是图 16—3A 还是图 16—3B 都不存在相关性。在图 16—3A 中, 即使输入变量 X 保持常量, 在绿 Y 中仍存在巨大的变化。因此, 绿 Y 中的变化必然来自其他输入变量而非 X。

在图 16—3B 中, 尽管输入变量 X 存在巨大的变化, 但在绿 Y 中却没有任何变化, 因此 X 对绿 Y 没有影响。

本练习的要点是, 若存在相关, 散布图中必有倾斜。垂直的和水平的绘图都不相关。

(2) 与图 16—4B 相比, 在图 16—4A 中, X 对绿 Y 的影响不一定更大。尽管图 16—4A 的倾斜度为 45° , 而图 16—4B 的倾斜度更小, 但是, 倾斜度取决于 X 和绿 Y 所使用的尺度。与倾斜度相比, 斜率更能说明相关性, 因为倾斜度可以改变, 它主要取决于所选择的尺度。

确定现实的技术规格和公差

(1) 经常参照客户要求的相应极限来选择绿 Y 及最大和最小的技术规格极限。

(2) 从以前的 DOE 技术中确定需要优化的红 X 和粉红 X 输入变量。使用 B vs. C 检验, 确保这些红 X 和粉红 X 变量已被验证为永久的改进。

(3) 确定这些重要变量之间的交互作用不存在或者可以忽略不计。如果不是这样的话, 使用响应曲面方法, 而不使用散布图。

(4) 为红 X 变量选择可以微调绿 Y 的一个值域 (或水平)。选择 30 个这样的红 X 值并记录对应的绿 Y 值。

(5) 绘制结果 (见图 16—5)。如果在图中是**倾斜的**且仅有较小的垂直散布, 则可进一步确认该红 X。如果只有很小的或根本没有倾斜且垂直散布较大, 则该变量既不是一个红 X 也不是一个粉红



X，并且该变量不重要。

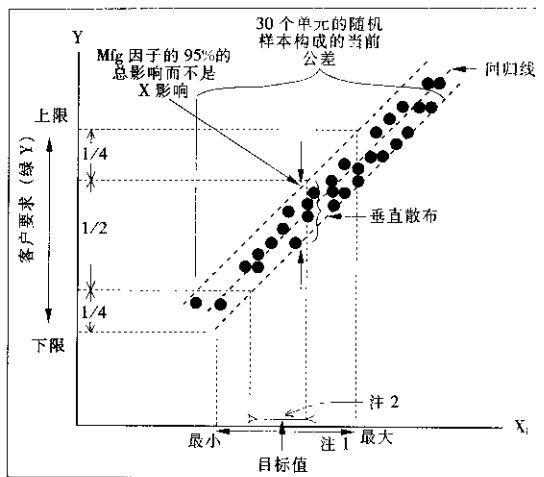


图 16—5 用散布图确定现实的技术规格和公差

注：(1) 最小和最大，对于 $C_{pk} = 1$ 的现实的输入变量公差，(2) 最小和最大，对于 $C_{pk} = 2$ 的现实的输入变量公差。

(6) 画一条通过这 30 个绘图点的中线，该中线称为回归线。在中线的两边各画一条平行线。这两条平行线离中线应当是等距离的，且几乎所有 30 个（最少 29 个）绘图点都应落在这两条平行线之间。这样生成的平行四边形的垂直截距是由除了红 X 之外的所有输入变量引起的绿 Y 变化。该垂直截距应当不大于绿 Y 技术规格的某个百分比（例如 20%），以证明该输入变量应标为红 X。如果垂直截距为绿 Y 技术规格的 20%~50%，则该输入变量可能是粉红 X 或浅粉红 X。

(7) 接下来从绿 Y 的技术规格上限处画一条水平线与该平行四

边形的顶线相交，然后在该交叉点处画一条垂线至 X 轴，这表示红 X 输入变量的最大的现实的公差。在该最大值右边的红 X 值将导致拒绝绿 Y（在最高技术规格极限的上方）。

(8) 类似地，从绿 Y 的技术规格下限处画一条水平线与该平行四边形的底线相交，然后在该交叉点处画一条垂线至 X 轴，这表示红 X 输入变量的最小的现实的公差。在该最小值左边的红 X 值将导致拒绝绿 Y（在最低技术规格极限的下方）。

(9) 该红 X 最大水平和最小水平的中间是该红 X 的目标值。同客户的技术规格宽度相比，这些最大值和最小值保证 C_{pk} 为 1.0。

另一方面，如果与客户的技术规格宽度相比，其期望的 C_{pk} 为 2.0，那么，客户的技术规格宽度应分成 4 等份。应从该技术规格宽度的中间一半处画两条水平线分别与该平行四边形的顶线和底线相交，并从这两个交点处画两条垂线至 X 轴。这样就可以确定为保证 C_{pk} 为 2.0 而非 1.0 时的最大和最小红 X 值。

(10) 这些正确的红 X 的最大/最小值和目标值应对照现有的值和公差进行比较，并作出必要的更改，以确保零缺陷和 100% 的成品率。

案例研究 1：点火放大器

在一家发动机点火放大器生产厂中，其产品在称为“关闭时间”的关键参数上的缺陷率大于 10%。关闭时间决定了在一个快速的关—开切换周期中关闭点火的时间（单位为毫秒）。以前的研究表明，电阻 R_3 和 R_4 分别是红 X 和粉红 X，且它们之间没有交互作用。

以关闭时间为绿 Y 并改变两个电阻的值而绘出的散布图如图 16--6 所示。



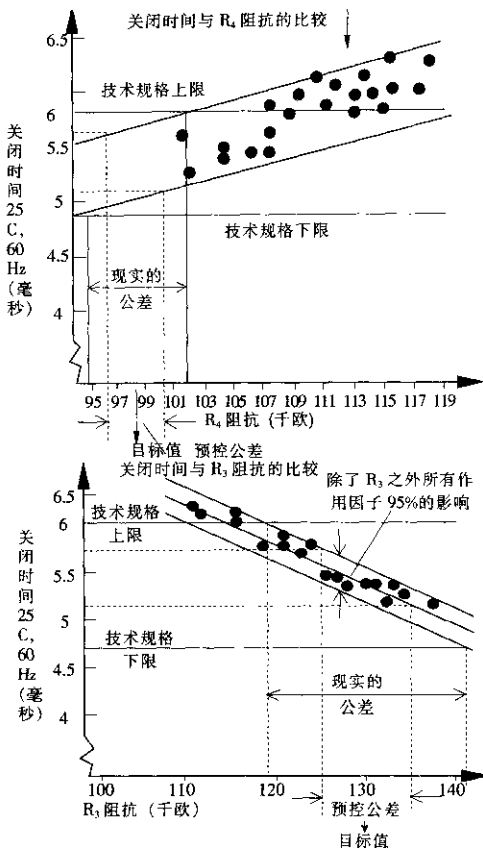


图 16—6 散布图案例研究：点火放大器

资料来源：Motorola, Schaumburg, 用。



电阻 R_4 的散布图确认 R_4 为浅粉红 X。尽管其平行四边形是倾斜的,但它较宽,其垂直截距大约是允许的绿 Y 技术规格宽度 1.2 毫秒 (4.8 毫秒~6.0 毫秒) 的 60%,这表明,绿 Y 变化的 60% 是由除了 R_4 之外的所有其他要素造成的。而且,该图以图形化的方式表明, R_3 的初始电阻值 110 千欧 $\pm$ 10%是错误的,并且这很可能就是造成高缺陷率的原因。该图还表明,现实的公差应为 94 千欧~102 千欧,目标值为 98 千欧可取得最小 C_P 值 1.0。然而,对于 C_P 值 2.0,该绿 Y 的技术规格宽度只能在 5.1 毫秒~5.7 毫秒之间变化,要求 R_4 不小于 96 千欧且不大于 100 千欧,即 98 千欧 $\pm$ 2%。

R_3 的散布图确认 R_3 为红 X。该平行四边形较窄,且其垂直截距只是该绿 Y 技术规格宽度 1.2 毫秒的 15%,这表明只有 15% 的该绿 Y 变化是由 R_3 之外的所有因子造成的。然而,该散布图再次清楚地表明, R_3 的初始电阻值 120 千欧 $\pm$ 10%是错误的,并且导致了点火放大器的高缺陷率。该图表明,现实的公差应为 118.5 千欧~141.5 千欧,目标值为 130 千欧可取得最小 C_P 值 1.0。如果希望得到 2.0 的 C_P 值, R_3 的电阻值范围应为 124 千欧~135 千欧,即 130 千欧 $\pm$ 4%。 R_3 和 R_4 的最终值既要考虑其理想目标值的折衷,又要考虑供货商能够提供的最接近的电阻值。在调整了电阻值后,产品的缺陷率降至 0。

小组练习 1: 红外焊接

从以往的经验来看,在某个复杂的电子模块中,中央处理单元 (CPU) 的红外焊接的成品率都很低。三个过程变量 (预热温度、焊接剂的黏性及 CPU 引线的平坦) 被认为是重要的,且它们之间的交互作用可以忽略不计。

改变上述变量值,并测量百分比结果绘出的散布图如图 16—7 所示。



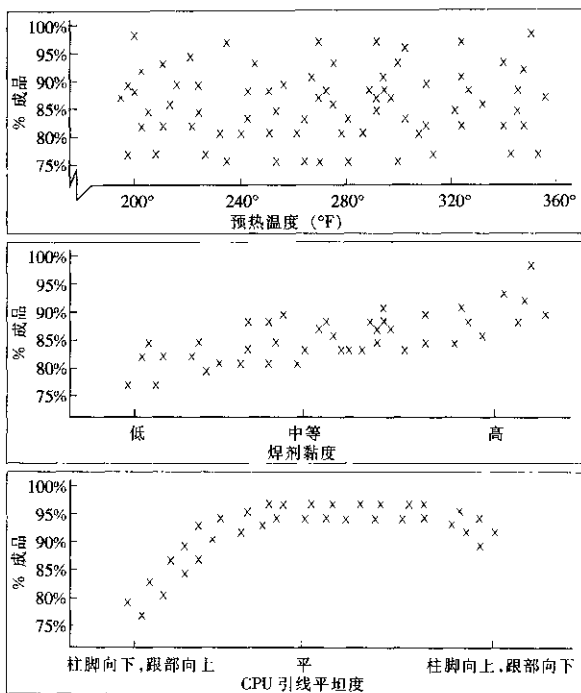


图 16—7 散布图小组练习 1: 红外焊接

问题

- (1) 预热温度重要吗? 为什么?
- (2) 焊接剂的黏性重要吗? 为什么?



(3) CPU 引线的平坦重要吗, 为什么?

(4) 在描述红 X 的图形中, 所有其他变量对成品率百分比的影响是什么?

(5) 如果期望的最小成品率是 90%, 那么, 要评价的 3 个变量的水平应是什么?

答案

(1) 和 (5) 参见图 16—8。预热温度不重要, 散布图没有倾斜, 且其垂直截距很大, 这表明所有其他输入变量的影响很大。对于最小的成品率 90%, 预热温度可以保持在 $200^{\circ}\text{F} \sim 360^{\circ}\text{F}$ 之间的任何温度, 但从经济的角度讲应保持在 200°F 。

(2) 和 (5) 焊接剂的黏性有点重要, 它是一个粉红 X 或浅粉红 X。其散布图有点倾斜 (尽管较小), 其平行四边形有些水平, 且其垂直截距适中, 表明大约 10%~20% 的成品率变化是由其它输入变量引起的。对于最小的成品率 90% 而言, 焊接剂的黏性应尽可能地高。

(3) 和 (4) CPU 引线的平坦非常重要, 它是一个红 X。尽管其散布图是非线性的 (散布图不必是线性的), 且有一个非常小的垂直截距, 表明只有 5% 的成品率变化是由其他输入变量引起的。对于最小的成品率 90% 而言, CPU 引线应当是平坦的。然而, 由于所有 64 个引线都完全平坦是不现实的, 柱脚向上、根部向下的方向也较好, 因为其倾斜度比柱脚向下、根部向上的方向的陡峭的倾斜度要平缓一些。

案例研究 2: 柱面配方与眼镜轴

一家隐型眼镜生产厂启动了一项研究, 以便从 “+” 组眼镜和 “-” 组眼镜中确定柱面配方 (光学精度) 和眼镜轴之间的关系。

367

从每一组中以 6 个不同的轴 (30、60、90、120、150 和 180) 选中了 10 副眼镜。

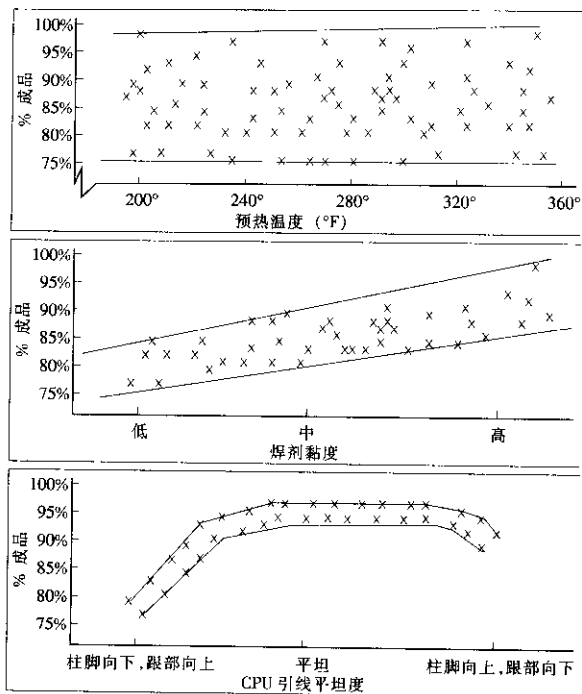


图 16—8 散布图小组练习答案 1: 红外焊接

资料来源: Astec Co., Manila, Philippines.

图 16—9 给出了“+”组眼镜的散布图结果, 图 16—10 给出了“-”组眼镜的散布图结果。



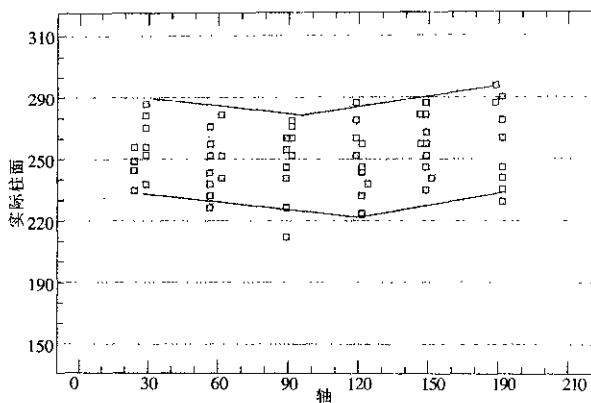


图 16—9 轴对柱面 Rx 的影响 (“+”组眼镜)

资料来源: Ciba Corp., Atlanta, Ga.

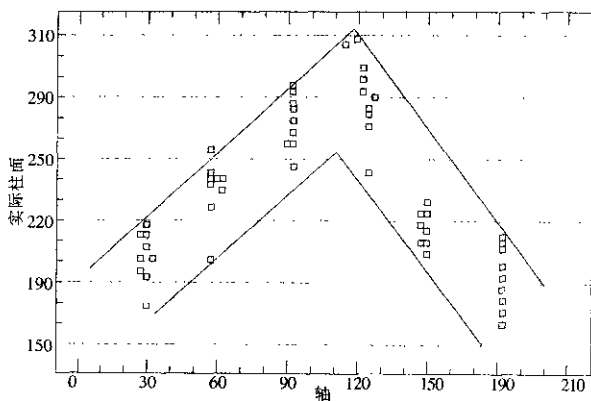


图 16—10 轴对柱面 Rx 的影响 (“-”组眼镜)

资料来源: Ciba Corp., Atlanta, Ga.

结论

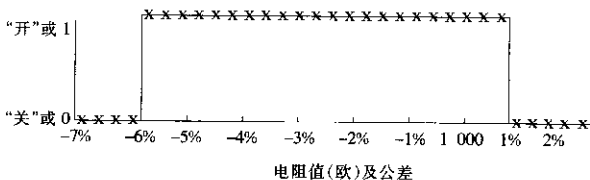
对于“+”组眼镜,其散布图表明有点倾斜,仅在 90~120 轴处有点弯曲,而且,每个轴的读数表明在实际的柱面配方中有一个较大的变化(变化的范围为 40)。这表明轴不是柱面配方变化的一个因子。

对于“-”组眼镜,其散布图显示了一个明确的趋势,与 30、60、150 和 180 轴相比,90 和 120 轴在实际柱面配方中有一个明显的增加。然而,这里在每个轴的 10 个读数中也有较大的范围——在柱面配方中有 30~40 的变化。为解释这种变化需要调查其他的因素。

小组练习 2: 数字电路

在某数字电路中,输出电压必须保持在“打开”或 1 状态方被认为是可接受的,“关闭”或 0 状态则被认为是不可接受的。以前的研究已经证实在该数字电路中的一个 1 000 欧的电阻是一个红 X,其公差规定为 $\pm 1\%$ 。

为检查电阻值及其公差,构造了一个散布图:



问题

(1) 1 000 欧是一个正确的值吗? 如果不是,那么该电阻的目标值应当是什么?

(2) $\pm 1\%$ 的公差正确吗? 如果不是,那么该电阻的正确的公



差应当是什么?

答案

(1) 在本例中, 由于只有惟一的属性 (1 为好, 0 为劣), 因此在散布图中没有倾斜。在 1 000 欧处, 电阻会给出好的结果。但它并不是中心, 其范围为 $-6\% \sim +1\%$, 因此, 其中心或目标应为 $7/2$ 或 -3.5% , 即在 965 欧处。

(2) $\pm 1\%$ 的公差太严格。对 $C_{pk} 1.0$, 它可以被展开为 $\pm 3.5\%$; 对 $C_{pk} 2.0$, 它可以被展开为 $\pm 1.75\%$ 。

散布图在行政管理中的应用

同其他 DOE 技术一样, 散布图已经广泛地应用于工业界、政府部门、学校和医院的行政管理和服务领域。实际上, 散布图广泛地应用于需要建立两个参数 [一个输出或从属变量 (绿 Y) 和一个输入或独立变量] 间关系的任何领域。表 16—3 给出了一些典型的例子。

表 16—3 散布图在行政管理中的应用

输出 (绿 Y)	输入变量	相关程度
国内生产总值	电力消费	紧密相关, 直到最近
国内生产总值	水资源消费	紧密相关, 直到最近
人均收入	教育程度	强相关
收入水平	学校声誉	强相关
癌症死亡	吸烟程度	强相关
公路交通死亡	速度限制	强相关
交通死亡	驾驶员酗酒	强相关
教学成绩	学校资金	中等相关
教学成绩	班级规模	中等相关
城市内部的贫困	领先启动计划	强相关
公司的寿命	公司的道德规范	强相关
客户忠诚度	雇员忠诚度	强相关



案例研究 3：客户满意度、客户忠诚度与利润

在工业界，人们一直认为较高的客户满意度将会带来更高的利润。然而，近期的研究表明，客户的满意度与公司的收益率没有什么关系。实际上，在公司的调查中发现，15%~40%的客户在他们表示满意后不久就转向了其他公司。这已从在仪表业进行的一个散点图研究中得到了证实。如图 16-11 所示，客户满意度与客户忠诚度之间的相互关系较差。作为对照，同样的研究表明，客户忠诚度与公司利润之间存在着较强的相互关系（如图 16-12 所示）。其差别在于：在客户满意度研究中度量的是客户的言辞，而在客户忠诚度研究中度量的则是客户的行动。

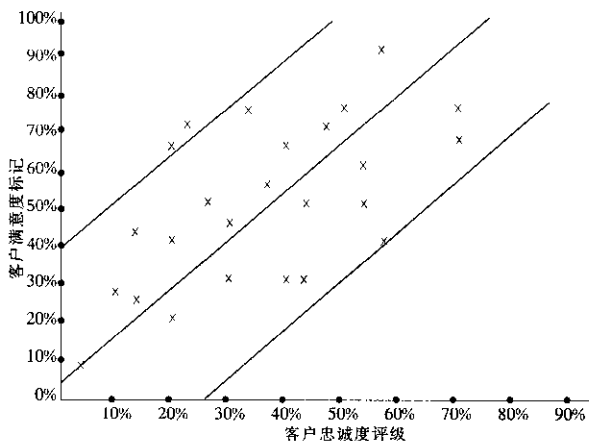


图 16-11 客户满意度并不是客户忠诚度的先知

资料来源：Keki R. Bhote, *Going Beyond Customer Satisfaction to Customer Loyalty*, New York: American Management Association, 1997.

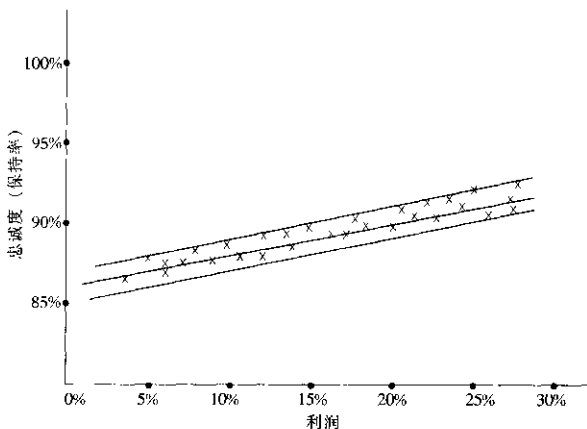


图 16—12 客户忠诚度与利润之间存在紧密的相互关系

资料来源: Keki R. Bhote, *Going Beyond Customer Satisfaction to Customer Loyalty*, New York: American Management Association, 1997.

多元回归分析

如果只有两个变量(一个从属的输出和一个独立的输入),那么,散布图及其更加数学化的版本——回归分析则是非常有用的。然而,通常有几个(多元)原因影响一个给定的输出。因此,需要一个更为严格的数学模型——多元回归分析,以指出这些变量之间的关系。

实例

一家航空公司需要确定:在几个独立的变量中,哪一个会导致



人们在以后乘坐飞机时选择该航空公司。在进行了一次多元回归分析之后，得到了下述结果：

变量	累积调整的 R^2
频繁的航班计划	0.43
航班时间的方便性	0.62
准点到达	0.75
机上服务	0.82
行李处理时间	0.85

这表明乘客在选择航空公司时的方差的 43% 是根据其频繁的航班计划作出的，而另外的 19% 是根据航班时间是否方便作出的。其余因子对乘客选择航空公司的决策影响较小。

► 作为散布图指南的 DOE 小组的问题

(1) 在试图使用散布图优化产品之前，已经使用了诸如变量搜索之类、前面阐述过的 DOE 来描述一个产品的特征吗？

(2) 如何确定当前的技术规格和公差？如何解决这一问题？

(3) 与客户需求相关的技术规格是否能满足要求？

(4) 使用了诸如质量功能展开、专业组、专家组及专家诊断之类的技术来评估客户的实际需求吗？

(5) 在最小 8~12 级水平中使用了输入变量的 30 个读数吗？（绝对最小值应为 20 个读数和变量 6 级水平。）

(6) 输入变量的读数和水平的检验顺序是随机的吗？

(7) 在获取一个红 X 输入变量的过程中，其垂直截距只占整个客户技术规格宽度的一个较小的百分比吗？

(8) 实际公差是根据 $C_{pk} 2.0$ 而不是 $C_{pk} 1.0$ 来确定的吗？

(9) 如果有 3~4 个具有强交互影响的输入变量，那么，是否



考虑使用响应曲面方法而不是散布图？

(10) 如果有几个输入变量，那么，是否考虑使用多元回归分析作为散布图的备选技术？

高层管理的问题

(1) 当前的技术规格和公差与客户的需求直接相关吗？在确定这些需求时同客户进行了交流吗？如何进行交流？

(2) 散布图中的 30 个读数在时间顺序上是随机的吗？

(3) 如果两个或多个输入变量之间存在交互作用，那么，尝试了响应曲面方法吗？

(4) 散布图方法已经扩展到了行政管理应用中吗？

【注释】

[1] 30 个读数最小可以分组成 8~12 个水平。

[2] 散布图用于输入变量之间没有强相关时，因此，当每一个输入变量变化时，其他变量保持在常量水平。





第 17 章

响应曲面方法 (RSM): 交互作用的优化

► 目标

● 上一章详细论述了散布图的目的，即确定没有交互作用的输入变量的最佳水平，以优化一个绿 Y。

● 响应曲面方法的目的是，确定两个或两个以上交互输入变量（在前面阐述的 DOE 试验中标识）的最佳水平组合，以得到一个最大、最小或最优的绿 Y（响应、输出和绿 Y 是同义词）。该技术在解决问题路径图中的位置如图 7-3 所示。

► 方法

有许多经验性的、统计的和数学的优化

方法以及响应曲面方法：

- 调优运算；
- 单纯形法；
- 随机调优运算；
- 最速上升法。

幸运的是，正如本书强调的那样，使用这些技术的人不需要渊博的统计学或数学知识。本章将重点介绍非数学的图形方法，该方法可以用少于 3~20 次迭代试验很容易地实现。

最速上升法将不在本书中论述，因为其数学的严密性在 90% 的响应曲面方法应用中是没有必要的。

概念：用爬山来类比

响应曲面方法的概念可以用爬山作为类比来进行解释。图 17—1 是一座山的二维等高线图，图中山周围的相同高度用等高线表示。山高逐渐上升，直到山峰为止。与寻找山峰类似，用 DOE 的术语来说，响应曲面方法是试图在一系列爬山试验中找到最优的响应（或最优的绿 Y）。调优运算、单纯形法、随机调优运算及最速上升法是到达顶峰的不同路径。**响应曲面**是重要的交互输入变量与

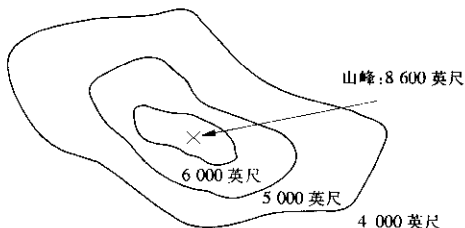


图 17—1 爬山

诸如产量、杂质之类的一个相关输出或响应、绿 Y 之间关系的数学（或图形）表示。等值的绿 Y 被描绘成等高线，就像山峰的等高线一样。

调优运算

调优运算（EVOP）是一种最简单的响应曲面方法优化技术。最基本的调优运算试验使用 2 个交互输入变量和 1 个响应变量或绿 Y。

过程

阶段 1。以 1 个输入变量的两级水平（高于和低于预先为该变量确定的某个适当的水平）及另 1 个输入变量的两级水平作为开始，这将导致 1 个 2^2 全析因试验。将这 4 个部件的响应或绿 Y 画成 1 个方框。还要在这 2 个输入变量的水平的中点进行 1 次附加试验，以确定在该方框内是否可能存在山峰，该读数称为中心点。

阶段 2。沿着阶段 1 创建的方框内的最高绿 Y 方向移动，用这两个输入变量的邻近水平进行第 2 次 2^2 全析因。再次沿着该中心点绘制一个表示这 4 个部件的方框。这可能产生 1 个更高的绿 Y。在阶段 3、4 和 5 等重新创建该过程，直到达到山峰或最优的绿 Y，且在任何方向都没有记录到进一步的改进为止。

案例研究：最高过程产量——肖特基二极管

在某个肖特基二极管工艺中已经确定了两个因素——时间和温度是重要的交互变量，这是事先在变量搜索试验中得出的特性。产量（绿 Y）以前为 82%。下一步的工作是进行一次调优运算，以确定产生最高产量的时间和温度水平。图 17—2 给出了调优运算的 4 个阶段。



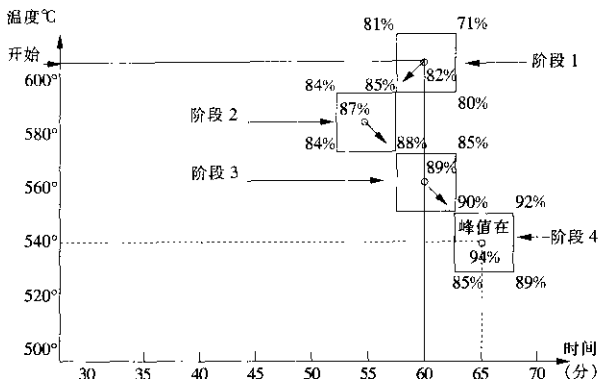


图 17-2 优化肖特基二极管产量的调优运算试验

资料来源: Motorola, Phoenix, Ariz.

阶段 1

调优运算的开始是用先前进行的变量搜索试验来确定这两个因子的适当水平。对于时间因子而言,其水平为 60 分钟;对于温度因子而言,其水平为 600°C。下一步将用两个临近的较高温度 610°C 和较低温度 590°C 以两个临近的较长时间 64 分钟和较短时间 56 分钟进行一次 2^2 全析因。(建议重复该 2^2 全析因 3 次以上,以检验在该试验中的残留错误或噪声。该试验的 16 个读数可以用图基检验来检验其重要性。大多数只关心读数的均值而不关心其方差的实验人员往往跳过这个较严格的方法。)图 17-2 表明阶段 1 在该方框的 4 角产生的产量分别为 80%、71%、81% 和 85%。

阶段 2

左下读数 85% 为最高产量。那么,调优运算规则要求阶段 2 从方框的中心沿对角线方向移动到 85% 及以上的高端。第 2 次 2^2 全



析因的结果表明,在较高点沿 85% 的产量为 85% (原点)、84%、84% 和 88%。阶段 2 东南角的产量比中心的产量要高,因此应在东南方向继续进行调优运算。

阶段 3 和阶段 4

下一阶段(即阶段 3 和阶段 4)重复该过程,直到阶段 4 结束为止。此时到达了顶峰,产量为 94%,其 4 个角的读数均低于中心的读数,表明该过程的最优时间为 66 分钟且最优温度为 540℃。

在该案例研究中,试验总数(不包括重复试验)为 17 次。一般说来,一次调优运算试验最少需要 2 个阶段,最多需要 7 个阶段。

应避免的缺陷

(1) 不要将调优运算用作筛选试验。在进行调优运算之前,应通过思路创新技术和变量搜索或全析因来确定产品或过程的特征,并正确确定交互作用输入变量。

(2) 在存在 4 个以上的交互作用输入变量时不要使用调优运算。在这种情况下,随机调优运算将是一种更好的响应曲面方法技术。

(3) 输入因子不应是某些属性,应为连续变量。

(4) 在调优运算的任何阶段,每一个输入变量的变化都应尽可能小。大的变化容易丢失峰值响应并可能导致回溯。

(5) 在每一阶段都应使 2^2 全析因的顺序随机化以避免偏差,并给噪声因子一个平等进入或不进入试验的机会。

(6) 若可行时,在每一个阶段重复 2^2 全析因以确定残留错误(即噪声与信号)。

(7) 不要无止境地搜索峰值绿 Y。例如与得到这么小的改进所需要的费用、时间和工作量相比,在该案例研究中产量为 94% 和产量为 100% 或 97% 的差别可能是微不足道的。

(8) 如果只需直接劳动者用较少的费用和较少的时间就可以进行一个简单的调优运算,那么,就不要诉诸于在响应曲面方法的最速上

升法中使用的、更加数学化的模型。不必将大量的工作花在区别二次方程式的响应和一次方程式的响应上，以试图从 99.5% 提高到 100%。

(9) 确保产品/过程不会随着时间的变化而变化。如果产品/过程随着时间的变化而变化，那么，要么必须重复调优运算，要么必须使用 DOE 和 MEOST 的组合方法进行进一步研究，以确定引起变化的原因。

(10) 确保在优化某个特殊的绿 Y 时，其他与性能、费用或可靠性有关的绿 Y 不会受到不利的影响。

将调优运算扩展到 3 个交互作用输入变量

如果优化某个绿 Y 的交互作用输入变量是 3 个或 4 个，而不是只有 2 个，那么可以应用与调优运算相同的原则。

图 17-1 和图 17-2 是三维输出的二维图示（2 个因子）。要估计第 3 个交互作用输入变量，需要一个三维图示来解释输出或绿 Y，而该输出或绿 Y 则成为第 4 维。

图 17-3 为一个具有 3 个输入变量 A、B 和 C 及一个响应曲面或绿 Y 的 3 维*图示，其响应曲面则为第 4 维。

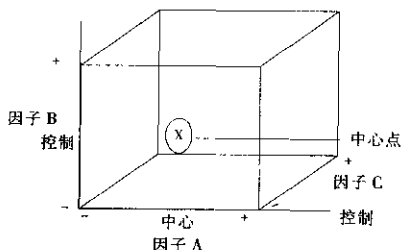


图 17-3 3 个输入变量的调优运算图示

* 本书原文为 2 维，应为 3 维。——译者注

其中，阶段1的调优运算实验以3个因子或变量A、B和C，且每个变量具有2个水平（-）和（+）的方式执行。在中心（位于A、B和C的中间水平）周围的8个角（而不是4个角）都是样本。只要图17-3的8个角中的1个或多个能得出比中心和其它所有角更好的结果，那么所有3个输入变量将以所希望的方向改变，以便向该4维表面的一个更好的等高线方向移动“检验区”。在进行了一系列这类 2^3 全析因之后，当中心的读数要好于最终的检验区的所有8个角的读数时，该绿Y便是最优值或者非常接近于最优值。表17-1是图17-3的8个角的表格式描述，其中心点如全析因矩阵所示。它基本是一个具有中心点的 2^3 全析因矩阵，外加因子A、B和C的中间点。

表 17-1 3个输入变量调优运算的析因矩阵

	A-		A	A+	
	B-	B+		B-	B+
C-			中心点		
C+					

将调优运算扩展到4个交互作用输入变量

要优化一个具有4个交互作用输入变量的绿Y，需要一个4维的图示来解释处于第5维的绿Y。尽管不可能用图形的方式来描绘，但可以使用相同的全析因矩阵概念。这需要一个具有中心点的 2^4 全析因矩阵（见表17-2）。

现在有16个角的读数并在因子A、B、C和D的中间水平有一



表 17—2

4 个输入变量调优运算的析因矩阵

		A			
		A-		A+	
		B-	B+	B-	B+
C-	D-				
	D+				
C+	D-				
	D+				

个中心点。此外，执行一系列的 2^3 全析因，直到中心点优于最终矩阵的其他 16 个角，以得到最优绿 Y。

单纯形法

虽然调优运算是到达某个绿 Y “顶峰”的一个简单的、图形化的和循序渐进的爬山过程，但它会产生大量的实验。例如，如果有 7 个阶段，每个阶段具有 2^2 全析因实验和 1 个中心点，那么，即使没有重复也需要进行 28 次实验。

单纯形法是一种较完善的技术，只需要较少的实验。因此，单纯形法比调优运算更为有效，且到达顶峰更快。单纯形法设计在每一阶段需要 $(n+1)$ 个点，其中 n 为交互作用输入变量的个数。与调优运算类似，单纯形法最好用一个涉及两个输入变量的案例研究来解释。

案例研究：油漆工艺的优化

在某家大型油漆生产厂，以前所做的变量搜索实验已经标识了两个参数为交互作用变量。这两个参数分别是：涂料粉剂的粒子大小

178

(研磨到以编码单位表示的纯度率)和烘箱温度(以编码单位表示)。变量搜索已经达到了84%的产量。为了进一步提高产量(达到目标值97%),启动了一个利用单纯形法的优化试验设计。图17-4给出了单纯形法的3个阶段,起点是由变量搜索达到的84%的产量。

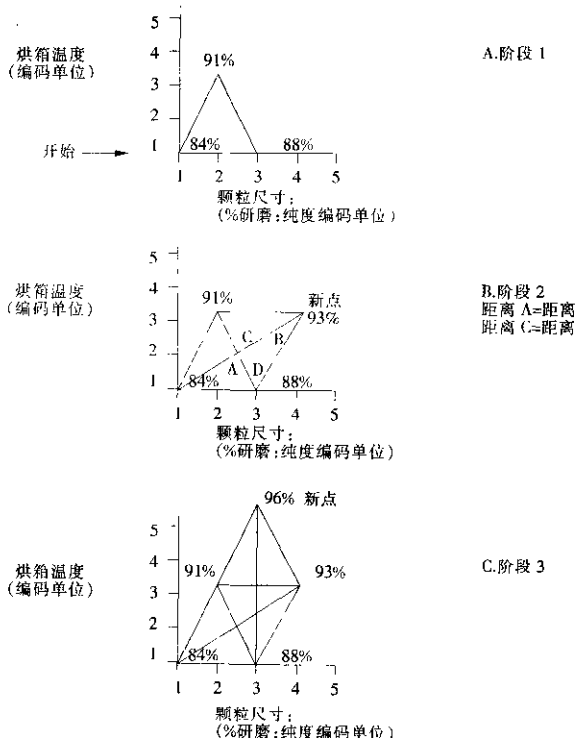


图 17—4 优化油漆工艺的单纯形法试验

阶段 1

试验了温度变量的 2 级水平和粒子大小变量的 3 级水平。阶段 1 的三角形结果表明 84% 的起点产量是最低的，而其他两个产量分别是 88%* 和 91%。

阶段 2

沿期望的响应方向继续进行优化。删除具有最低产量的单纯形法点 (84%)，并用阶段 1 三角形的其他两个点组成另一个单纯形法。(已经研究出了许多算法来确定新的阶段 2 三角形的第 3 个点，最简单的技术是构成连接 2 个剩余点的一条直线。) 新的尝试是从一个删除点上画一条二等分 2 个剩余点连线的直线，并延长该直线直到其距离等于删除点到等分线的距离为止^[1]。图 17-4B 给出了一条直线，该直线从删除的单纯形法点 84% 出发，等分 91% 和 88%* 连线，并到达一个更高的产量 93%*

阶段 3

删除阶段 2 的最低产量 88%，并以 91% 的产量建立阶段 3 的三角形 (见图 17-4C)，得到顶点 96%。尽管可以进行更多的阶段，但考虑到时间和费用，96% 的产量已经足够。需要指出的是，一个相应的调优运算方法 3 个阶段需要 12 次试验，而单纯形法仅需 $3+1+1=5$ 次试验。

随机调优运算

尽管调优运算有效且彻底，但在交互作用输入变量数目超过 4 时，它便变得十分费时且开支昂贵。这个问题导致开发了随机调优

* 此处原文误为 87%，应为 88%；原文误为 85%，应为 88%；原文误为 95%，应为 93%。——译者注。

运算技术。随机调优运算理论与调优运算理论基本相同，但它引入了定位改进方向的随机方向。随机调优运算的实践者声称，最优的绿 Y 可以在 16~20 次试验中得到。

步骤

(1) 如果可能的话，在变量搜索或全析因试验中使用全析因分析技术，尽量将实际的交互作用输入变量或因因子减少到 3 个或 4 个以下。这将不需要随机调优运算。

(2) 如果步骤 1 不可行，那么根据安全性、费用、以前的经验、客户/政府的要求等，列出这些输入因子，并确定每个因子的实际范围。

(3) 确定每个因子要改变的随机方向和随机变化量。

(4) 确定在每次试验中每个因子所使用的范围的最大部分。

(5) 根据以往试验的最好水平，选择一个起点。

(6) 绘制首次随机变化的结果。如果结果有利，那么以相同的方向继续。

(7) 如果结果不利，那么以相反的方向继续。

(8) 如果结果既不是有利的也不是不利的，或者两个方向的结果都是不利的，那么，选择另一个随机方向。

(9) 继续上述步骤，直到大约 4 个相继的、随机的、不利的方向和总量都已选择为止（与其对应的 4 个相反方向也是有利的），此时将得到一个最优的绿 Y。

图 17—5 是一次涉及 5 个输入变量的随机调优运算试验的图形表示，这 5 个输入变量分别是温度、压力、时间、浓度和体积。该试验的目的是最大化某个输出或响应。

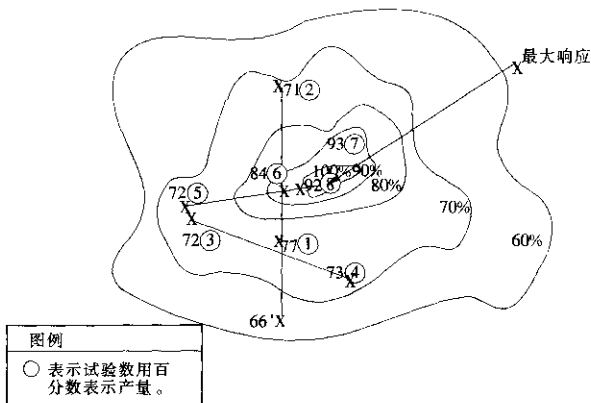


图 17—5 随机调优运算试验的图形表示

作为响应曲面方法 (RSM) 项目的 DOE 小组的问题

(1) 响应曲面方法是否仅用作红 X 搜索中的最后一步，且仅在已经描述了一个产品/过程的特征之后进行？

(2) 为证明使用响应曲面方法是正确的，在早期的试验设计中是否已经确认存在 2 个或 2 个以上的重要交互作用输入变量？

(3) 选择响应曲面方法的不同方法（即调优运算、单纯形法、随机调优运算和最速上升法）的理由是什么？

(4) 是否在每个输入变量中，有意地使水平的变化保持较小？

(5) 在调优运算中的 2^2 全析因试验是随机的和重复的吗？

(6) 在达到一个绝对峰值的绿 Y 和时间、人力、费用支出之间



存在一个合理的折衷吗？

(7) 对响应曲面方法而言，简单的图形方法优于复杂的数学模型吗？

(8) 存在可靠性问题（即绿 Y 随时间和应力变化）吗？如果存在，考虑使用试验设计/MEOST 方法的组合来解决这一问题吗？

(9) 除了所选的绿 Y 之外，还有其他需要优化的绿 Y 吗？

(10) 在优化某个绿 Y 时，另一个绿 Y 被局部优化，在这种情况下，输入变量之间必须进行某些折衷吗？

在 DOE 小组项目评审期间的 DOE 小组项目评估的高层管理的问题

(1) 真的需要响应曲面方法作为 DOE 项目的最后一步吗？

(2) 与从绿 Y 改进得到的利益相比，响应曲面方法的费用是多少？

(3) 生产线操作员和非技术人员能学会响应曲面方法吗？

(4) 与费用高昂的数学方法相比，应强调简单的图形方法吗？

(5) 如何解决优化两个或两个以上的绿 Y 之间的冲突（每一个绿 Y 都需要不同的输入变量水平）？

【注释】

[1] 如果新点低于单纯形法图中其他两个点，则规则为回到原先的单纯形法的方向。已经为单纯形法技术开发出了高级的计算机程序。



第 6 篇

从试验设计到统计过程控制的转换：
重要的学科

第3篇、第4篇和第5篇论述了DOE技术的万用工具箱，这些DOE可以共同解决任何长期的质量问题，更重要的是，可以在第一地点预防这类问题。

然而，这并非解决问题之路的终点。在我们介绍统计过程控制（statistical process control, SPC；统计过程控制最多是一个监察和维护工具，而不是一个解决问题工具）之前，必须运用这两门非常重要的学科，即正向控制和过程确认。目前关于质量的教科书甚至都不提及这两门学科，而且绝大多数质量工作者都乐于忽略它们的存在和功能。

（1）第1门学科是正向控制（positrol），它冻结过程，特别是在试验设计之后的过程增益。

（2）第2门学科是过程确认（process certification），它系统地审核和改正一些引起不良质量的外部原因。这些外部原因可能潜伏在工厂中，并且试验设计未将之考虑为因子。实际上，这些因子集合在一起可以由于累积噪声因子足以使其结果被认为不确定的或错误，以致否定任何试验。

作者正逐步得出这样一个结论：过程确认应当是任何试验设计的前提条件，而且是试验设计的继续。





第 18 章

正向控制：冻结过程增益

概念

工业的弱点之一是，工程师和技术人员试图通过检查其所生产的产品来控制一个过程。这种做法未免太晚了。以航海的术语来类比，这就像通过观察船艇的航迹来驾驶一艘船艇一样。与产品一样，一个过程也具有必须通过 DOE 工具来确定和分析的特性和内在变量。

- 产品过程搜索的思路创新技术是一项 DOE 技术，它可以区分出重要的过程变量和不重要的过程变量。

- 变量搜索和全析因可以使 DOE 进一步地提取重要变量以及它们之间的交互影响



效应。

● 散布图和响应曲面方法为每一个重要变量确立了现实的技术规格和现实的公差。

但是，如果在日常生产中没有控制这些重要变量及其公差，没有锁住和冻结它们，所有的努力都将落空。我们曾看到这样的先例，一些公司的 DOE 工作做得很好，其产品的缺陷水平大大降低，然而在以后的生产中产品的缺陷率却再次回升。

正向控制计划

必须制定一个称为正向控制^[1]计划的体系，来确保那些重要的变量受到严格的控制。这些变量即为控制对象，通常通过以往的试验设计来刻画和优化，并由包括“人员”、“方法”、“时间”和“地点”的计划进行严格控制。正向控制在解决问题路径图中的位置如图 7-3 所示。

该计划确定了哪些“人员”应该监控、测量和记录每一重要的过程参数。“方法”确定了测量这些重要参数的正确仪器 [观察 5:1 法则，该法则指明仪器的精度至少应该是参数精度（即公差）的 5 倍]。“地点”应该是测量过程参数的最佳位置，只有这样才能真实地反映其正确值。“时间”是测量的频率，最初由工程判断确定，而后来则由预先控制（详见第 21 章）确定。

表 18—1 是一个有关正向控制计划的例子，该计划是在成功地进行了波焊研究的 2⁴ 全析因（详见第 14 章）之后制定的。该计划必须由过程技术人员、维护人员以及生产线操作员正确无误地认真贯彻。如果重要参数及其公差已经通过 DOE 辛辛苦苦地建立起来，那么，就应当遵守，除非附加的 DOE 研究可以更进一步提高质量或产量。

表 18-1

波焊过程的正向控制计划

参数	规格及公差	测量			
		人员	方法	地点	时间
预热温度	220°F ± 5°	自动	热耦合	容器入口	连续
倾斜角度	7° ± 20%	过程技术员	角度盘	倾斜处	每次模型变更
传送带速度	6 英尺/分 钟 ± 10%	过程技术员	计数器	控制台 馈入	每次模型变更
880 焊剂 密度	0.864 gm./ cc ± 0.008	实验 技术员	专用 重力仪	实验室	每天一次

资料来源: Adi K. Bhote, Motorola, Scottsdale, Ariz.

正向控制图

一旦制定了正向控制计划,就必须有一个正向控制图紧随其后。该图是一个在过程中维护的具有较高可视性的运行日志,它必须由指定人员填写,并由监督人员、过程工程师和质量控制部门定期监控。

表 18-2 给出了一个有关喷涂机的正向控制图的例子。该喷涂机的用途是在玻璃的一面喷涂铬、镍和金涂层,而在另一面喷涂铬和金涂层。喷涂机以往的历史表明,玻璃上的金属附着层存在着连续的“弯曲”这一缺陷。尽管过程工程师一个接一个地旋转移动控制按钮,以试图控制金属的附着,但是,最后总是不得不结束这种徒劳无功的努力,并且完全迷惑不解甚至不知所措。DOE 研究确定了四个因素(即功率、气压、速度和真空)是重要因子。散布图确定了每一个参数的最大实际公差。在三班制操作中,由操作员对图进行一日四次的维护。通过 DOE,该缺陷降到零。通过正向控制,缺陷保持在零!



表 18-2

有关金属喷涂机的正向控制图

机器: 903

结束周: 6-20

机器参数	周 二						周 三						周 四						周 五						周 六						
	6h	10h	12h	14h	16h	18h	6h	10h	12h	14h	16h	18h	6h	10h	12h	14h	16h	18h	6h	10h	12h	14h	16h	18h	6h	10h	12h	14h	16h	18h	
800~900	Cr	820	800	820	820	832	865	820	820	820	820	820	820	861	882	861	861	861	861	861	861	861	861	861	861	861	861	861	861	861	861
1706~4800	Ni	1818	4830	4770	4779	4772	4797	4779	4779	4779	4779	4779	4779	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817	4817
2300~2400	Au	2400	2400	2400	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370
1020~1130	Cr	1066	1066	1066	1066	1066	1066	1066	1066	1066	1066	1066	1066	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111
2300~2400	Au	2400	2400	2400	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370	2370
气 压 (c)	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	
	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	
	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	
	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	
速 度 (IPM)	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	
	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	
	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	
	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	
3×10° 真空度	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	
	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰	3×10 ⁻¹⁰		
检查人																															
每架结束时的运行次数																															

备注 / 问题原因

资料来源: Motorola, Schaumburg, Ill.



在许多情况下，使用预先控制（详见第 21 章）比使用正向控制图更经济。预先控制的好处是：在两次纠正措施间隔时间足够长的情况下，可以延长检查每一过程参数的间隔时间。对过程而不是产品使用预先控制还有一个优点。当产品的缺陷水平下降到大大低于 100ppm 时，控制图就不起作用了，但即使使用预先控制来检查产品的特性，其代价也可能变得非常高。因此，解决的办法就是通过监控过程（而不是用相反的形式）来控制产品，通过正向控制和预先控制来监控关键的过程参数。

在更为现代化的过程中，通常以微处理器为基础，过程参数的公差可以设计进过程中，因此正向控制变成了自动的，所以既不需要正向控制图，也不需要预先控制。

倒退的原因

为什么 DOE 的良好初始结果会倒退、滑坡、烟消云散？其原因有如下几点：

（1）试验者没有用 B vs. C 检验（详见第 16 章）来确认 DOE 的改进。

（2）试验者通常是技术型的，他们把过程维护移交给维护人员、装配人员、技术人员以及生产线操作员，而没有对正向控制在维持所得增益方面的重要性提供指导。

（3）这些操作员是很好的，他们中的绝大多数是真正渴望做好质量工作的。但是，他们不是 DOE 解决方案中的一员，而在他们的心目中相信，他们可以适当地停止过程，并且会比工程师们做得更好。

（4）这些过程技术人员天生就是“欺骗的艺术家”，而且不能远离控制按钮。

（5）车间缺乏纪律要求。

（6）后果的责任不明确。



(7) 使生产线操作员觉得正向控制是一个监视他们（而不是过程状态）的“间谍系统”。

在某种程度上，正向控制的要求等同于 ISO 9000 的要求，而在 ISO 9000 中，步骤是严格按照文字来遵守的。但最大的差别是，ISO 9000 是在过程通过 DOE 得到改进之前采用的。这样做的好处是什么呢？ISO 9000 仅仅冻结了缺陷，而正向控制却冻结了改进！

通过一个案例研究可以更好地说明正向控制的重要性。

案例研究 1：机械手

一家著名跨国公司的新加坡工厂聘笔者为顾问，以便大幅度地提高其输出产品的质量。主要的推动力就是 DOE。在两年的时间里，通过定期访问并与该公司协作，我们将其输出产品的质量由 15 000ppm 提高到 120ppm。过程之一是一个机械手，该机械手以每分钟夹取 100 件的速度从一个卷轴传送带上挑选出元器件，并将它们转移到一块印刷电路板上，该过程的缺陷率为 3 000ppm。我们的 DOE 小组已经将该过程的缺陷率降低到 60ppm。

然而，在以后的访问中，工厂经理提到缺陷率虽然没有以前那样糟糕，但现在已经反弹到了 500ppm。我指导他说：“让我们与部件谈谈吧。”于是，每天绘制一张缺陷历史图。在某些天中，缺陷率为零；在另一些天中，缺陷率为 1 000ppm、200ppm 或 500ppm。图表像一张锯齿图，说明机械手没有受到控制。我问道：“你的正向控制图在哪里？”工厂经理拿我的问题去问技师，结果他的脸变白了。工厂经理这才知道正向控制的重要性，因此，他参加了我的为期两天的 DOE 研讨会。

该事件促使该厂在全厂范围内确立了一项纪律要求，即每台机械手都配置一张正向控制图，并显示在明显位置，使数个 DOE 方案所产生的增益能够被锁定。一个更进一步的证据发生在几年以



后。当新加坡的劳动力成本变高时，工厂在跨过马来西亚海峡的印度尼西亚建立了一个进料器生产厂，几个过程也随之转移到那里。工厂担心的问题是新的国家、新的工厂和新的人员会造成管理上的阻碍。但由于每一过程都已经过特征化、最优化并被冻结，因而过程的转移非常顺利，仅在一周之内就重现了原来在新加坡工厂的成品率。

小组练习 1：焊接 PC 板

一家大型电视机生产厂使用 IXOE 将焊接的缺陷率从 12 000ppm 降低到 85ppm。接着又制定了一个正向控制计划（见表 18—3），以监控关键的过程参数。

表 18—3 小组练习：焊接 PC 板的正向控制

技术规格					
设备			材料		
型号：	EPM CDD 300E	焊料合金：	SN 60 PB 40		
产品：	UV 900 双面板	焊剂：	ZEVA 590		
方法：	没有清洁	散热油：	COBAR 808		
空气：	过滤干燥空气				
正向控制计划					
参数 (对象)	技术规格	测量			
		人员	方法	位置	时间
焊剂	0.814 ~ 0.83 (PC 衬垫 0.818~0.826)	机器操作员	密度计	焊剂箱	每班 3 次
喷焊 压力	1.6Bar~1.8 Bar	机器操作员	调节器	机器进口	每班 3 次



续前表

技术规格					
设备			材料		
喷焊鼓 旋转	6.5rpm ~ 7.5rpm/ 8.0Sec/Rev ~ 9.23 Sec/Rev	机器 操作员	秒表	焊鼓	每班 3次
焊剂喷 射角度	85°~90°向右	技师/ 助工	角度计		每天 下午
预热温 度	90°C~110°C	技师/ 助工	莫尔 分布图	进口到 出口	每周
焊接厚 度	27mm~28mm 开泵 时	技师/ 助工	刻度尺/ 厚度传感 器	焊池	每班 4次
喷嘴缝 隙	0.8mm~0.9mm	技师/ 助工	进料 量计	谐波 喷嘴	每天 下午
谐波高 度	1至1.5气压计刻度	机器 操作员	气压计	谐波	每班 3次
主波高 度	2至3.0气压计刻度	机器 操作员	气压计	主波	每班 3次
焊接温 度	245C~250C	机器 操作员	温度计	主波 焊接	每班 3次
调谐器 底到谐 波喷嘴 顶距离	10.5mm~11.5mm	技师/ 助工	游标 卡尺	谐波 喷嘴	每周 下午
调谐器 底到主 波喷嘴 顶距离	5.5mm~6.5mm	技师/ 助工	游标 卡尺	主波 喷嘴	每周 下午
传送带 速度	50 秒/米~55 秒/米	机器 操作员	秒表	传送带/ 链上1米	每班 3次
不同的 泵油	25~35Gamma/ 4-W	技师/ 助工	油泵	主波	每班 3次
运载车 清洗	15 台在线	生产 操作员	超声波 装置	焊剂室	每班 4次

资料来源: Phillips, Singapore.



问题

- (1) 对于该产品及其有关波焊过程，你可能会使用什么方法？
- (2) 对正向控制计划你会使用什么样的改进措施？

答案

问题 (1) 和 (2)：正向控制计划表明有 15 个过程参数是重要的，并且需要通过正向控制来维护。这个过程参数清单似乎太大了。首先，使用产品过程搜索（一种最佳的 DOE 技术）来缩小这个由 15 个过程参数组成的清单。通过变量搜索，该清单可以进一步减少到 4~5 个真正重要的参数。

下面我们将给出 3 个附加的案例研究（3 个完全不同的过程），这些案例研究将说明：在使用 DOE 来描述过程的特征之后，正向控制是如何成为一个重要步骤的。

案例研究 2：计算机数控车床

正向控制计划（未给出参数公差）如表 18-4 所示。

表 18-4 案例研究 2：计算机数控车床

对象	人员	方法	地点	时间
冷却剂	自动	反射计	喷嘴	每隔 4 小时
进刀	自动	转速计	内部	连续
速度	自动	转速计	内部	连续
切削刀具	操作员	目视：10 倍率	工具架	每隔半小时

资料来源：S. K. F.，Gottenburg，Sweden.

案例研究 3: 塑料注模机

如表 18-5 所示。

表 18—5

案例研究 3: 塑料注模机

参数对象	技术规格及公差	人员	方法	地点	时间
模具温度	140°F~150°F	副工长	高温计	模腔表面	每班两次
启动压力	540PSI~560PSI	副工长	压力计	机器	每班两次
保持压力	570PSI~590PSI	副工长	压力计	机器	每班两次
熔化温度	520°F~530°F	副工长	高温计	熔化物	每班一次
材料干燥	180 秒~200 秒	副工长	温度计	干燥器	每班两次

资料来源: Pine River Plastics, Warren, Mich.

案例研究 4: 模具铸造

如表 18—6 所示。

表 18—6

案例研究 4: 模具铸造

参数对象	技术规格及参数	人员	方法	地点	时间
金属温度	800°±20°	操作员	热偶合	压机	每天两次
冲击压力	800PSI~1 000PSI	操作员	量计	压机	每次准备过程
水压机压力	800PSI~1 000PSI	操作员	量计	压机	每次准备过程
冷却剂	全部	操作员	目视	压机	每天两次
喷射脱模剂	全部	操作员	目视	压机	每天一次

资料来源: Racine Die Casting, Racine, Wisconsin.

作为正向控制项目指南的 DOE 小组的问题

(1) 是否使用产品过程搜索对过程进行研究,以挖掘出有关重要过程参数的线索?

(2) 是否在产品/过程搜索之后使用变量搜索或全析因来描述过程的特征?

(3) 是否在变量搜索或全析因之后使用散布图/响应曲面方法对过程进行最优化?

(4) 作为正向控制计划的开始,是否为每一个重要的过程参数(对象)确立了实际的技术规格和公差?

(5) 每一重要参数的**测量人员、方法、地点和时间**是否已经明确?

(6) 在制定了正向控制计划之后,是否建立了正向控制图?

(7) 是否使用预先控制来监控和控制重要的过程参数,尤其是当与过程相关的产品缺陷水平低于 100ppm 时?

(8) 是否警告过生产线操作员、维护小组以及技术人员千万不要背离正向控制图,并停止平时那种“欺骗和调整”的倾向,除非认为有必要进行更进一步的试验设计?

(9) 正向控制图是否显示在显著位置,以使所有人都能够看到并在过程参数失控时及时采取措施?

(10) 是否有警报灯光或其他视听信号,以便向操作员或管理人员发出参数失控的警告?

(11) 是否采用微处理器技术把警告界限置于过程内部,以便在过程参数失控时可以自动关闭过程?

(12) 是否向生产线操作员保证过,正向控制不是监视其表现的管理系统,只是一个监视过程表现好坏的系统?



在 DOE 项目评审期间的高层管理的问题

(1) 对 ISO 9000 (用于冻结缺陷) 的需要与对正向控制 (用于冻结改进) 的需求之间是否存在明显的区别?

(2) 管理部门是否已向所有操作员、维护小组、技术人员说明正向控制并非检查他们的表现, 而是检查过程的表现及其一致性?

(3) 是否由不同级别的管理部门进行定期的巡回检查, 以确定正向控制图表的要求是否得到满足?

(4) 是否采用预先控制与正向控制相结合的方式来监控所选的过程参数?

(5) 是否设计了警报灯光和其他视听信号, 以便向生产线操作员发出过程参数失控的警告?

【注释】

[1] 术语 positrol (正向控制) 是由 positive control (正向控制) 两词合成而来的。





第 19 章

过程确认：消除不良质量的外部原因

墨菲定律

如果圣·帕特里克是爱尔兰的守护神，那么，圣·墨菲毫无疑问就是工业界的守护神！实际上，令人感到幽默但又十分真实的是，过程确认的基础就是墨菲定律，即一句普遍的格言：“如果事情可能会变坏，那么，它就将会变坏。”墨菲定律在工业界是普遍存在的，过程确认所面临的挑战就是要消除墨菲定律。

偏差的主要原因可以通过 DOE 来大大减少，但仍有一些导致不良质量的外部原因和四处运动着的一些微小的墨菲现象，可能使过程在 DOE 研究之前或之后半途而废。



必须捕捉到这些原因并把它们禁闭起来。过程确认正是完成这一捕捉任务的警察。图 7—3 解决问题路径图中标明了过程确认的位置（注：过程确认必须在 DOE 研究之前使用）。

不良质量的外部原因分类表

表 19—1 是一个可能导致不良质量的各种外部原因的分类表。该表可以分为下述 5 大主要类别：

- 管理/监控不足；
- 违背良好的生产习惯；
- 对车间/设备的忽视；
- 忽视环境；
- 人为缺陷。

表 19—1 不良质量的外部原因分类表

(1) 管理/监控不足	(2) 违背良好的生产习惯 (GMP)
• 生产线工人中弥漫着恐惧心理 • 工人的思想观念受到抑制 • 排除错误原因未得到鼓励 • 独断的生产线监视 • 没有部门内部的或职能交叉的队伍 • 人员流动过于频繁 • 过高的缺勤率 • 没有利益共享 • 很少或没有培训 • 很少或没有 Poka-Yoke • 没有操作员确认 • 多技能操作员未得到鼓励 • 没有可达到的目标 • 测量： ◇ 未测量不良质量的代价	• 未编写标准操作过程或者太困难 • 工人和产品的安全性较差 • 不良的人类工程学 • 保管比较随意 • 过程流与生产流 • 推与拉系统 • 建立、转换时间太长 • 地面上堆积过多存货，走道拥挤 • 难以得到工具 • 频繁变换模型 • 频繁中断生产线 • 不完全的构造 • 不清楚的、混淆的、矛盾的指令 • 网络过多



◇未跟踪生产率/周期时间

◇未测量 C_p , C_{pk}

- 结果的反馈很少
- 没有视觉/听觉上的质量警告信号
- 数据污染 未对数据采取措施
- 工人没有关闭不良质量生产线的权力
- 没有正向控制
- 检验员追踪零件, 文件过多
- 把产量凌驾于质量之上
- 对出色完成的工作缺乏赏识
- 恶劣的工作条件

(3) 对车间/设备的疏忽

(4) 忽视环境

(5) 人为缺陷

- 未使用全面生产维修
- 预防性维护与损坏时修理的比例较低
- 忽视以下方面:
 - ◇润滑
 - ◇机器噪声
 - ◇机器振动
 - ◇过热
 - ◇电压波动
 - ◇传送带速度
 - ◇腐蚀
 - ◇空气软管压力
- 仪器:
 - ◇未达到 5:1 的精度
 - ◇没有跟踪国家标准的能力
 - ◇未经常校准或未校准
- 通风不足
- 不良的失效/安全控制
- 没有外部空气密闭阀

缺乏:

- 温度控制
- 湿度控制
- 水净化
- 空气净化
- 尘埃控制
- 化学物控制
- 充足照明
- 通风控制
- 静电负荷保护
- 电磁兼容保护
- 禁止吸烟

- 没有遵守标准操作过程
- 缺乏纪律
- “蒙骗专家”
- 严重的个人主义凌驾于团队协作之上
- 将个人问题带到工作当中
- 酗酒/吸毒
- 不合理的结盟要求



该列表并没有涵盖所有的情况, 其目的只是作为过程确认小组

在执行 DOE 之前和之后的一个检查表和指南。在每一分类中，一些项是显而易见的，另一些项则需要进行详细的说明。

管理/监控不足

戴明博士曾经说过：工业界最具腐蚀性的因素之一就是在雇员中蔓延着害怕直话直说的心理。仅次于零部件本身，工人是解决问题的思路的最丰富的源泉。但是，假如他们思想的水库没有被开源引流的话，那么势必会损失一笔极有价值的资源。而且，工人需要“参与”，以获得被承认的感觉，否则，他们将会抱着不交流、不参与的态度。

错误原因的消除 (ECR) 是让工人记录质量问题的方法之一。他们不必提供解决方法，但是，鼓励他们指出这些问题是向着促使他们参与和解决问题的方向所迈出的第一步。

操作员资格确认是一项必要的要求，尤其是对那些困难的工艺和装配。它包括对特殊工艺的操作员进行培训以及考试和确认。同时还应对操作员定期进行资格的重新确认，以确保正确的方法不会遗忘。通常要保留一个表格，该表列出了每一个操作员的姓名以及他们得到资格确认的各种过程。

在质量测量中，不良质量的代价是最重要的宏观参数。由于包含了担保、废料、分析、修理、检查、检验以及过多的库存——所有非增值项，不良质量的代价所造成的总浪费要占到销售量的10%~20%。**更明确地说，不良质量造成了一个公司每人每天 100 美元~200 美元的浪费！**更详细的测量包括成比率（目标是 100%）和寿命周期时间（目标是不超过两倍直接劳动时间）之间的关系以及其他测量，例如每个部件的缺陷总数（TDPU），即将从生产线的起点到末端之间的所有缺陷累加起来，并除以通过的部件总数；以及重要产品/过程能力指数 C_p 和 C_{pk} 。

违背良好生产习惯

良好生产习惯（CMP）在工业界是人尽皆知的。它们通过像质量环、小组改进活动（SGIA）、持续改善（Kaizen）、玻璃墙管理以及 ISO 9000 等计划得到加强。这些大纲应当用来确定以下内容：

- 是否需要标准操作过程。
- 如果需要的话，对操作员来说，该过程是否过于困难或过于墨守成规。
- 这些过程以简单的术语阐述。
- 是否鼓励操作员输入标准操作过程。
- 操作员是否遵守指令，而不是只在必要时才遵守指令。
- 是否强制实施正向控制。
- 标准操作过程是否在简单化、准确、一致等方面进行定期审查。

在过去的 10 年中，以丰田生产系统为基础的精益制造（lean manufacturing）日益流行起来，其中包括焦点工厂、产品与过程流、拉系统、小批量规模、短时设置时间等。精益制造还强调产品质量、过程质量和短时周期，这些都是其设计的本质。

对车间/设备的疏忽

对车间/设备的疏忽类属于全面生产维护（TPM），这已在第 3 章中讨论过了。大多数西方工业界都忽视了它的用途和巨大的利益。因此，人们对待过程和机器的态度就是：“如果它没坏，就不要修理它！”预防性维修与纠正性维修的比率还不足 1：10，更不用说反过来了。应当鼓励全体生产线操作员通过 DOE 技术承担更多的预防性维修工作。

此外，正如零部件会给出提示一样，过程和机器同样会给出提示。机器噪声、发热、摩擦以及其他警告信号都是未来故障的早期警告。



最后还应介绍一下计量学的内容。目前已经建立了这样的规则,即仪器的精度必须不少于被测量产品规格公差的5倍。此外,检验设备的校准必须符合国家标准,校准的频率及遵循校准时间表的过程必须受到监督。

忽视环境

在许多工厂中,环境控制不在生产管理思想的大屏幕上。工厂对于房间环境温度、湿度、静电放电等,任其随着自然的变化而变化,而其中任意一个环境条件都可能对许多产品的质量产生严重的影响。

以灰尘为例,在半导体工业中,灰尘是质量的大敌。灰尘的测量是按每立方英尺空气中杂质大于0.5微米的标准进行的。在医院的外科监护病房,允许的数字是10 000粒子。在半导体工业中,杂质所允许的最大颗粒数几年前就是1 000,现今房间的建造要求是少于1——医院外科病房洁净度的10 000倍。

人为缺陷

大多数的雇员并不是问题。像戴明和朱兰这样的质量大师指出,85%的质量问题是管理的责任,只有15%的质量问题是工人的责任。但是,在工人当中可能会出现这样的倾向,即回避纪律、寻求捷径、加速生产、“蒙骗”管理。尽管所有这类倾向的动机都无恶意,但它是牺牲良好的质量为代价的,这种倾向应当通过管理来监督和纠正。

过程确认方法

(1) 过程确认最好由一个交叉学科小组来全力实施,其成员来



自开发或工程部门、质量部门、制造工程部门、生产部门以及熟悉确认过程的其他部门。

(2) 该小组查阅前面描述过的外部质量问题分类表或者编制一个更有限的、具体到某个过程的列表。

(3) 然后开始审核该过程，确保该过程在得到确认之前通过彻底的“过程清洗”来解决和牢牢抓住所有潜在的外部质量问题。

(4) 根据时间安排，较早的建议是在 DOE 和正向控制之后、但在采用统计过程控制作为维护工具之前进行过程确认。作者当前的建议是，在 DOE 开始之前进行一次初步过程审核，并且在 DOE 之后还要进行一次试验设计后的过程审核。原因是，在开始时，一个过程可能包含许多这类外部质量问题，如果这些问题得不到解决，它们可能会使试验噪声累加到超过 DOE 分析信号的程度。在 DOE 之前进行一次过程“清洗”，将会保证 DOE 研究的成功。

(5) 只有在第 4 步之后一切就位，该过程才有资格开始进行第一个生产部件的确认。

(6) 应定期地（最好是每年一至两次）由该小组对过程进行重新确认，以确保既不会再次出现早期的外部质量问题，又不致产生新的外部质量问题。

案例研究 1：硅电容气压传感器金属喷镀

在硅电容气压（SCAP）传感器产品中，一台喷涂机把金属层涂覆在一只玻璃盘的两侧。下面的过程确认检查单是由过程确认小组起草的：

- (1) 喷涂机的正向控制计划和正向控制图准备就位了吗？
- (2) 温度及湿度控制装置准备就位了吗？
- (3) 存放和使用化学品的安全控制装置准备就位了吗？
- (4) 操作员对操作目标明确了吗？
- (5) 操作员有明确的指令吗？

- (6) 有质量目标以及目标未达到时的关闭准则吗?
- (7) 玻璃的清洁过程是足够的和明确的吗?
- (8) 所有的检验设备都有校准措施和校准频率吗?
- (9) 在变更引入之前, 工程变更的有效性经过了 B vs. C 试验的确认吗?
- (10) 黄金控制及废料处理的政策和措施是否就位, 是否得到了遵守?

案例研究 2: 冲压机

某车间在 DOE 分析之后编制了下述过程确认检查单:

- (1) 对在变量搜索设备中识别出的重要参数, 是否编制了正向控制计划和正向控制图?
- (2) 是否对原材料的厚度、硬度、宽度和弧度进行了定期检查, 并对每个批次都进行了检查?
- (3) 轧机和原料矫直机有预防性维护的时间表吗?
- (4) 检查了原料进料器的校准吗?
- (5) 板材缺陷检测就位了吗?
- (6) 压机控制:
 - 冲头进行校准了吗?
 - RPM 是正确的吗?
 - 零件弹出机制就位了吗?
 - 抗衡压力正确吗?
 - 安全电路正在起作用吗?
- (7) 模具组:
 - 快速切割的细节检查过了吗?
 - 成形细节检查过了吗?
 - 夹持压力正确吗?
- (8) 润滑就位了吗?



(9) 毛口清理在技术规格范围内吗?

(10) 表面除油受控吗?

评论

正如在“变量搜索”一章中描述的那样,有关冲压成形的DOE试验识别出了必须高度受控的两个重要变量。但是,正如前面指出的那样,在过程中仍有一些需要防止的“小墨菲现象”,以使产品在一段连续时间域内达到零缺陷且 C_{pk} 超过2.0。

小组练习 1: 瓷器喷漆过程

有关全析因一章中的小组练习1讨论了一个瓷器喷漆过程,该过程的成品率很低,通常在63%~82%之间。DOE小组推测温度和湿度是两个可能的原因,因为成品率在冬季高一些(大约为82%),而夏季低一些(大约为63%)。

该小组练习描述了一个由作者引入的 2^2 全析因。结果是令人吃惊的,两个温度和湿度的四种组合中没有一种产生的成品率低于91.5%! DOE小组无法相信这样好的结果——与温度或湿度无关。

问题

(1) 什么是真正的红X?

(2) 从该小组练习中得到的教训是什么?

答案 (1 和 2)

● 2^2 全析因表明,较高的温度实际上比较低温度所产生的成品率更高,与DOE小组的理论正好相反。



●较高的湿度实际上比较低的湿度所产生的成品率更高，同样与 DOE 小组的理论正好相反。

●温度对成品率的影响是湿度的三倍。

●无论温度还是湿度都不是真正的红 X。

●真正的红 X 是过程确认审核以及在 DOE 分析开始之前对审核所发现的外部质量问题所进行的“清洗”。

●工厂经理曾带作者参观生产线并讲解喷漆过程。他为这些复杂的设备而感到自豪。但是，在参观生产线时，我发现有很多方面（例如，保管、设备校准、校准仪的控制等）都偏离了良好的生产习惯。工厂经理非常尴尬。我建议，作为企业的首要任务是，在开始 DOE 研究之前，由 DOE 小组来实施一次过程确认审核，并领会、捕捉和排除“小墨菲现象”。

表 19-2 给出了过程确认审核的结果。实际情况比人们预料的还要糟糕。人们整整花了三个星期时间进行过程确认清洗以排除这些缺陷。但是，清洗消除了许多“噪声”因子，否则这些因子可能会进入 2^2 全析因实验，并污染信号的纯洁性。

表 19-2

过程确认审核结果：喷漆过程

1. 吊架未对准或遗漏	13. 清洗设备传送带需要封闭
2. 吊架在粉末间出口处摩擦墙壁	14. 生产线停机过多（一小时五次）
3. 基本涂层手动喷涂（无自动喷枪）	15. 传送带颤动
4. 气压计显示低压力	16. 传送带轨道需要清洁
5. 小室内无水平控制	17. 冷却管道底部脏了
6. 喷枪背面渗漏	18. 空调（AC）回风位置不正确
7. 泵渗漏	19. 输送区域人员太少
8. 输送装置渗漏	20. 吊架阻塞气球充满灰尘
9. 粉末从万能小室中末级过滤器中漏出	21. 万能小室软管过长
10. 输送站墙壁需要绝缘	22. 备件供给不足
11. 清洗设备的零件生锈	23. 没有偶氮屏蔽或偶氮废气收集器
12. 清洗设备温度 20°，太低	24. 外部空气渗入

从这次很有启发的工厂实习中得到的教训是：在 DOE 研究之



前，应进行一次过程确认审核，以建立一个更纯洁的和低噪声的基线，为 DOE 研究提供取得成功的机会。

小组练习 2：存储器芯片的低成品率

美国一家制造存储器芯片的半导体工厂，为了解决成品率低的问题进行了多种尝试，但其成品率总是不能突破 75%。而在该公司的一家生产类似存储器芯片的日本合资工厂中，却有着至少不低于 95% 的成品率。

于是，该公司将一个实地调查小组派往日本工厂，以调查造成成品率差异的可能原因。调查小组发现，在设计、过程、材料和检验设备方面没有什么差别，洁净室也具备相同的最大杂质颗粒要求。

问题

红 X 差别是什么？它是如何被发现的？

答案

调查由于未能找到造成两家工厂成品率差别的根本原因而受阻，公司决定进行一次对日本工厂的“巡视”，从工作日的开始就记录下操作习惯方面的差异。此记录将制作、装配、检验的不同步骤绘制成流程图。除了两个非常相关的观察结果之外，仍然没有发现在美国和日本的两家工厂之间的重大差别。这两个相关的观察结果是：

(1) 早晨，日本的工人会正好在同一时间走进超洁净室；上午休息时则刚好在同一时间走出超洁净室，并在休息后恰好同时走回超洁净室；午休时也是如此。美国的工人会在不同的时间里零零散散地走进超洁净室，在第一次休息之前则多次独自走出去喝咖啡或抽烟，并且在整个当班期间反复如此。因此，与日本的工厂相比，



美国工厂的超洁净室外充满灰尘的空气较为频繁地进入超洁净室。

(2) 在每次休息之后，日本工人手臂交叉坐在其工位上；重新开始工作之前，在整整五分钟的时间里不做任何工作。这是要让灰尘自然地落下。损失的五分钟生产时间将通过更高的成品率得到更多的补偿。这一令人困惑的习惯的理由是，人体以每分钟超过 100 000 的速度释放出灰尘颗粒，当人在走动或移动时这一数字跳跃到超过 1 000 000。在敏感的半导体设备上继续工作之前，让灰尘降落是目前的标准习惯。调查小组返回美国，着手对这些过程确认审核进行改进，结果取得了 95% 甚至更高的成品率。

作为 DOE 小组过程确认指南和在 DOE 项目 评审期间管理的问题

过程确认审核与清洗

(1) 在开始审核之前，是否成立了一个交叉学科小组来编制一个针对过程中待观察项目的检查单？

(2) 在开始 DOE 研究之前，该小组是否实施了一次过程确认审核，以排除或者降低 DOE 工作中的噪声因子？

(3) 该小组是否进行了一次过程确认清洗，以排除各种外部质量问题原因？

(4) 在 DOE 研究之后，该小组进行了过程确认重新审核？

管理/监控

(1) 在管理者和工人之间存在大胆交流的氛围吗？

(2) 工人在工作和使用简单的 DOE 工具解决问题方面得到了足够的培训吗？



(3) 错误原因的消除措施是否正式就位，以便工人指出重新出现的质量问题？

(4) 是否对工人进行了一种或多种关键过程的技术操作资格确认，是否定期地对工人重新进行资格确认？

(5) 是否建立了 Poka-Yoke（因果图），使工人可从传感器接收到一个错误将要发生的信号并预先纠正该错误？

(6) 有关过程的正向控制计划和正向控制图是否就位，并正得到仔细的维护？

(7) 像不良质量代价、成品率/周期时间图、每部件总缺陷以及 C_p 和 C_{pk} 这类的质量度量是否就位，以便跟踪改进措施？

(8) 团队协作是否通过像质量环和持续改善这样的纪律得到了鼓励？

良好生产习惯

(1) 标准操作过程是否在必要的地方就位？这些过程很容易遵守吗？工人的输入是否要求采用明确阐述的标准操作过程？

(2) 像拉系统、产品流、小批量规模以及短时设置时间之类的精益生产概念是否已被引入并加以实施？

计量学

(1) 仪器精度是否能够满足标准局或其他国家机构规定的要求？

(2) 是否有一个关于所有重要仪器的校准时间表并得到了遵守和审核？

全面生产维护

(1) 全面生产维护是否就位，并且是否度量了每一过程的工厂



总体效率？

- (2) 每一过程的预防性维修与纠正性维修的比例是多少？
- (3) 有无视听警告信号作为过程问题的指示？
- (4) 有无带有自纠正机制的过程内置诊断措施？

环境

(1) 在过程确认审核和 DOE 研究中，是否考虑了诸如温度、湿度、静电之类的周边环境？

(2) 对灰尘、空气、水、气体以及化学品的控制是否就位？

人员

一旦在过程中建立起了综合 DOE 以及正向控制，那么，是否禁止工人“蒙骗”控制装置和调整装置？



第 7 篇

设计过程控制：用于监控质量

统计过程控制（SPC）作为质量运动的标志——“银弹”而被广泛推销，它可以奇迹般地解决长期性的质量问题。然而，尽管数以千计的公司在过去 20 年中曾使用过它，但是它却产生了令人失望的结果—— C_p 和 C_{pk} 很少达到 1.0。

本书第 7 篇阐述的是统计过程控制的一项重要用途，而不是把它视为解决问题的一个工具。它只能在用 DOE 技术解决质量变化和质量问题之后使用，其主要作用是监控一个大大改进了的产品或过程，以使问题不再重新出现，主要用作维护工具。

统计过程控制有两个分支：

- 控制图——过时的、不太有效的技术；
- 预先控制——全新的、统计上更为有力的技术。

我们将在第 20 章和第 21 章分别介绍这两种技术。





第20章

控制图：一项过时的技术

► 控制图的曲折历史

控制图是 70 多年前由沃尔特·休哈特 (Walter Shewhart) 开发出来的，它很快就成为实验室的学术界和生产的实践界之间的桥梁。在第二次世界大战期间，控制图在生产中得到了广泛的应用。但是，在第二次世界大战之后，随着美国成为惟一的经济超级大国，工业界认为它并不需要这些统计技术。

但是，到了 20 世纪 70 年代末期，美国在工业上的领先地位逐渐丧失。美国国家广播公司制作了一部电视记录片：“如果日本能够做到，那么我们为什么不能做到？”该记录片的主题是：日本的质量是杰出的，因



为日本使用了统计过程控制和控制图；而美国的质量是落后的，因为美国没有使用统计过程控制和控制图。在制作该记录片时，人们还不知道早在几年前日本就已经因为控制图效率不高而放弃使用了！尽管如此，美国国家广播公司制作的这部记录片仍然成为了划分前 SPC 时代和 SPC 时代的历史性分水岭。控制图重新得到了广泛的应用。

那曾是一个专制主义盛行的时代。几个原始设备制造商的客户，尤其是福特公司，将使用控制图作为与设备制造商做买卖的准入证。然而不幸的是，控制图并没有在解决问题和减少质量变化方面起到什么作用。在 15 年里，工业界在控制图上花费了数十亿美元，而投资的回报只有数百万美元，投资回报率仅 0.1%。人们哪怕将钱存入已经倒闭的 S&L 银行都可以做得更好！其原因是：他们要做的只是维护一个通过 DOE 改进的产品或过程。

控制图的缺点

由于目前已经发表有大量有关控制图资料¹⁷，因此本书将不详细介绍其理论基础、原则和机制。相反，我们将重点讨论控制图的固有缺点，并通过两个案例研究加以充分说明。

案例研究 1：套管

在某机加工车间的作业中，要把一个套管加工到 0.500 英寸 ± 0.002 英寸的长度。表 20-1 给出了一些数据，利用这些数据可以构造出如图 20-1 所示的 $\bar{X}$ 图和 R 图。

这些图表明，所有子组的平均数 ($\bar{X}$) 都处在控制上限和控制下限之间的范围内，所有子组的值域 (R) 同样都处在控制上限和控制下限之间的范围内。附带地在一个半作业班中获取 60 个读数和 32 个计算值而完成的试验控制图表明该过程是稳定的（换句话说，它是一个定因

系统)。于是,可以全面开展生产,以加工成千上万的部件。

表 20—1

套管长度的控制图数据

套管长度技术规格 0.500 英寸 ± 0.002 英寸													
样本号	上午 8:00	上午 9:00	上午 10:00	上午 11:00	中午 12:00	下午 1:00	下午 2:00	下午 3:00	下午 4:00	下午 5:00	下午 6:00	下午 7:00	
1	0.501	0.501	0.502	0.501	0.501	0.500	0.500	0.500	0.501	0.502	0.501	0.500	
2	0.501	0.501	0.501	0.502	0.501	0.500	0.501	0.501	0.501	0.502	0.502	0.500	
3	0.500	0.501	0.502	0.501	0.501	0.502	0.501	0.501	0.501	0.501	0.501	0.501	
4	0.501	0.501	0.501	0.500	0.501	0.502	0.501	0.501	0.501	0.502	0.501	0.502	
5	0.502	0.502	0.501	0.500	0.501	0.502	0.500	0.500	0.501	0.501	0.501	0.501	
X 样本	2.505	2.506	2.507	2.504	2.505	2.506	2.503	2.503	2.505	2.508	2.506	2.504	
X ₁	0.501	0.501	0.501	0.500	0.501	0.501	0.500	0.500	0.501	0.501	0.501	0.500	
R ₁	0.002	0.001	0.001	0.002	0.000	0.002	0.002	0.001	0.000	0.001	0.001	0.002	

$$X_1 \text{ 之和} = 6.0128$$

$$R_1 \text{ 之和} = 0.0115$$

$$\bar{X} = \frac{\sum X}{N} = \frac{6.0128}{12} = 0.50107 \quad \bar{R} = \frac{\sum R}{N} = \frac{0.0115}{12} = 0.00096$$

控制限:

$$\bar{X} \pm A_2 \bar{R} = 0.50107 \pm (0.58)(0.00096)$$

$$UCL = 0.50180$$

样本均值: $LCL = 0.50034$

$$UCL_R = D_4 \bar{R} = (2.11)(0.00096) = 0.00204$$

值域: $LCL_R = D_3 \bar{R} = (0)(0.00096) = 0$

然而,如果已经画出了规格的控制上限(见图 20-1),那么,可以明显地看到,平均数 $\bar{X}$ 的控制上限危险地接近了单个读数的技术规格上限。甚至一个对统计了解不多的外行人都可以推断出,如果平均值如此接近于技术规格限,那么,组成平均值的单个值就可能超过技术规格上限。(许多控制图的纯化论者不让技术规格限显示在控制图上,因为他们担心这样做会抑制连续过程改进的必要性,他们相信只要通过制作控制图就可以神奇地实现过程改进!)



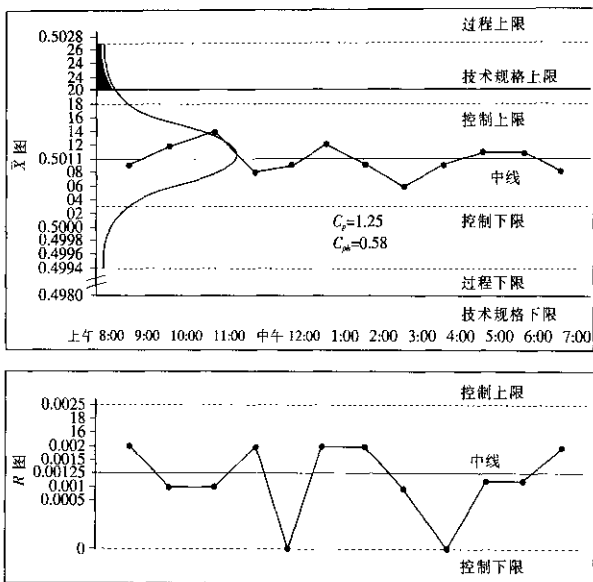


图 20—1 套管长度: $\bar{X}$ 和 R 图

资料来源: Wilson Sporting Goods Co., Chicago, Ill.

更准确地说, 单个值的预计散布可以通过过程限公式计算出来:

$$\text{过程上限} = \bar{X} + 3\sigma^l = \bar{X} + 3\bar{R}/d_2$$

$$\text{过程下限} = \bar{X} - 3\sigma^l = \bar{X} - 3\bar{R}/d_2$$

式中, σ^l 为全体标准偏差; $\bar{X}$ 为子组平均数的总平均值; $\bar{R}$ 为子组值域的平均值; d_2 为给定子组规模的一个常数。给定子组规模为 4 时, d_2 为 2.059, 给定子组规模为 5 时, d_2 为 2.326。

图 20-1 画出了这些过程限。图中清楚地表明过程上限为

0.007 英寸，高于技术规格上限，说明 7% 的套管可能存在缺陷（图 20-1 的阴影部分）。因此，这是一个表明一切正常且生产可以继续全速进行的控制图。而实际上，此时过程可能会产生一个完全不可接受的部件缺陷率。而且，在大多数制作控制图的工作中，人们甚至不知道过程限，更没有计算它或者用它来测量过程的能力。实际上，如果计算一下该过程的 C_p 和 C_{pk} 的话，将会得到很糟糕的 C_p (1.25) 和 C_{pk} (0.58)。对控制图来说，这是一个莫大的讽刺！

在质量控制文献中， β 风险被定义为接受一个应当废弃的产品的风险，而 α 风险被定义为废弃一个合格产品的风险（在该案例研究中， β 风险为 7%）。可以证明，在某些情况下 β 风险可能会高达 30%。

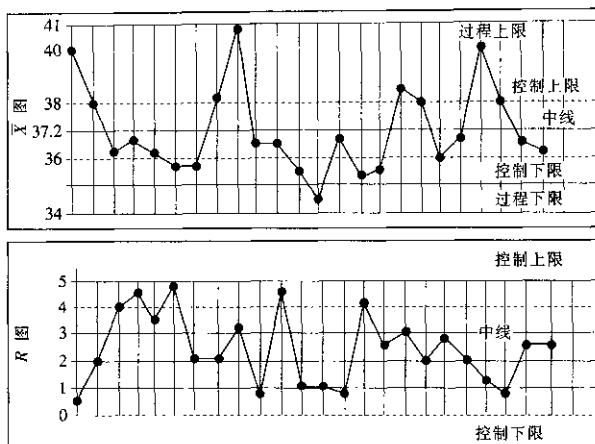
案例研究 2：传感器电容

一个在汽车上检测大气压力的电子元件对电容的要求为 31 皮法~45 皮法 (pf)。图 20-2 给出了该过程的 $\bar{X}$ 和 R 控制图。

与第一个案例研究相比， $\bar{X}$ 和 R 图都给出了在控制上限和控制下限以外的点，因而宣布该过程失去控制并停止了该过程。然而，生产表明，在过去一年半的时间内生产了 10 000 多个元件，且在输出质量审核或在客户的工厂中没有发现一个废品。浏览一下子组数据就会发现，甚至既没有单个读数接近于技术规格上限 45 皮法，也没有单个读数接近于技术规格下限 31 皮法。此外，预计的单个读数的过程限为 41 皮法和 35 皮法，刚好处在规格限范围之内。（预计的过程分布比实际的单个读数分布稍微窄一些，因为选择的子组规模为不同寻常的大——10。）

在本案例研究中，存在一个与第一个案例研究相反的情况：控制图表明该过程应当关闭，而生产却有种种合理的理由得以继续！其实，该过程的 C_p 和 C_{pk} 值分别是 2.0 和 1.92，这是 15 年前就已达到的世界级质量水平！





技术规格: 31 微微法拉~45 微微法拉

$\bar{X} = 37.2 \text{ pf}$; $R = 2.2$; $UCL_{\bar{x}} = 37.88$; $LCL_{\bar{x}} = 36.52$; $UCL_R = 3.95$;

$ICL_R = 0$

过程上限=41.0; 过程下限=35.0

$C_p = 2.0$; $k = 0.58$; $C_{pk} = 1.92$

图 20—2 传感器 $\bar{X}$ 和 R 图

资料来源: Motorola, Schaumburg, III.

随意的控制图实践

在实践中, 违背控制图规则和指南的情况比符合它们的情况还要多。为应付官僚制度, 在经济枪口下生存的公司 (尤其是供货商) 采取了下述方式使用控制图:



(1) 对失控的图不采取措施, 绘制图只是为了以某种方式暗示其一致性。

(2) 当一个点超出控制限时, 生产线操作员、检查员和技术员没有如何处理它们的头绪, 于是他们把成堆的控制图存档, 以证明它们的存在。

(3) 处理一切可能的参数, 不论其是否重要, 其结果只会使控制图激增, 好像仅用数量就可以抵补有意义的内容。

(4) 工厂的墙上乱七八糟地贴满了控制图, 为的是看谁能在相互竞争的工厂之间赢得拥有最多图表的奖赏。(某工厂的记录是 2 000 张图。)

(5) 在 75% 以上的控制图中, 确定控制限的初步试验从未得到再次校准, 控制限一旦画出, 它们就像雕在石头上的雕塑一样固化了。

(6) 在数个案例中, 工厂干脆在客户质量代表来厂参观的前一天凭空捏造控制图。

(7) 有一种错误的看法, 认为如果控制图显示出所有 $\bar{X}$ 和 R 点都在控制限范围之内, 那么一个定因系统就渗透到一个产品或过程之中。实际上, 这样的情况不能当作一个稳定的过程。设计会变更、过程会变更、材料会变更、环境会变更、人员会变更、设备会变更。在工业中惟一不变的事情就是变更!

(8) 制作的控制图未画出技术规格限, 因而没有废品水平的参照; 未计算出过程极限, 因而没有 C_p 和 C_{pk} 或者 α 和 β 风险的参考; 没有检验趋势的 $1\sigma_x$ 、 $2\sigma_x$ 极限。

DOE 小组和管理的问题

由于控制图已经渡过了其实用期, 因而 DOE 小组不应为了唤起对控制图问题的记忆而浪费自己宝贵的时间, 高层管理也不必评审控制图工作的进程。相反, 高层管理应把时间、金钱以及人力分配到更有用的项目中去。





第 21 章

预先控制：一项全新的技术

预先控制是一项比控制图更新、更简单、更好用、更廉价，在统计方面更强大的技术。在本章后面的部分中，我们将展示预先控制具有控制图的所有优点而没有其缺点。由于预先控制不大为人所知，因而我们对其的介绍要比控制图更加深入。预先控制在解决问题路径图中的位置如图 7-3 所示。

◆ 预先控制的发现

预先控制是由拉丝和强力 (rath and strong) 咨询公司在 20 世纪 50 年代为一家对烦琐而无效的控制图不再抱有幻想的《财

富》500 强公司开发的。

预先控制的奠基人弗兰克·斯凯瑟韦特 (Frank Scatherwaite) 是一位卓越的统计学家。40 多年前, 他在一篇综合性的论文中创立了预先控制的理论基础。不幸的是, 就在预先控制刚刚获得承认时 (20 世纪 50 年代末期笔者曾在摩托罗拉公司介绍预先控制), 美国工业在第二次世界大战后的那些年里因经济的成功而充满活力, 它抛弃了所有的统计学方法, 包括控制图和预先控制。接着, 随着统计过程控制在 20 世纪 80 年代的再次流行, 控制图又再次显露头角, 预先控制开始重新出现在统计学的视野之中。

在一些大公司中, 仍然存在着对预先控制的后防阻力, 它们宁愿生活在挥鞭拉车的时代也不愿步入喷气机的时代。但是, 它们的队伍正在明显地缩小。15 年前, 控制图与预先控制的用户之比为 99:1, 目前的比值为 80:20。而在 21 世纪, 随着预先控制的简单性和有效性得到较好的宣传, 控制图就如在日本一样将逐步成为历史, 而预先控制将会成为统计过程控制界的主要维护工具。

由 4 个简易步骤构成的预先控制机制

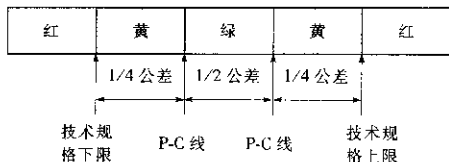
预先控制的机制是如此的简单, 以至于任何人 (例如生产线工人、供应商甚至工程师) 都可以在 10 分钟内学会预先控制。只需遵守下述 4 个简单的规则:

规则 1: 建立预先控制线 (控制限)

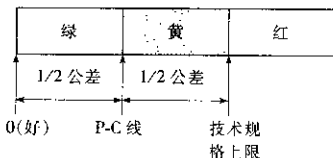
(a) 对于一个双边公差, 把绿 Y 参数的技术规格宽度¹⁾除以 4, 并在中间 1/2 处标出两个点, 即从下技术规格限起 1/4 宽度处标出一点, 并从上技术规格限起 1/4 宽度处标出另一点。于是, 便立即有了两个预先控制的控制限, 称为预先控制线, 或者简称为 P-C 线 (见图 21—1A)。(预先控制不需要一个正态分布或者任何有关其分

布的形状或稳定性的假设。)

(A) 对于一个双边公差



(B) 对于单边公差 (只有技术规格上限)



(C) 对于单边公差 (只有技术规格下限)

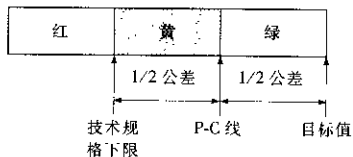


图 21—1 预先控制规则

资料来源: Rath S. Strong, Boston, Mass.

预先控制的规则 1 非常简单而又非常精巧! 只要除以 4, 就可以马上得到预先控制 (P-C) 线。这就是该项技术称为预先控制的原因。甚至在开始生产之前, 就可以得到控制限。这与控制图是完全不同的, 因为对于控制图来说, 用户必须做到:

- 在一天、二天或者三天的时间内运行 100 个、200 个或者 300 个部件。



● 然后仅得到**试验性**的控制限。

● 于是经常需要以大约每两周一次的周期重新校准这些试验性控制限，因为在工业中没有如此稳定的系统。

● 在一天、二天或者三天的时间里为每一次的重新校准完成另外的 100 个、200 个或者 300 个部件。

在两条 P-C（预先控制）线之间的区域称为绿区；在每条预先控制线和技术规格限之间的区域称为黄区；在技术规格限以外的两个区域称为红区。

(b) 对于单边公差（只有技术规格上限），此处只有一个技术规格上限而没有技术规格下限。把技术规格上限和零之间的距离平分，在中间绘制一条单一的预先控制（P-C）线，如图 21-1B 所示。

(c) 对于单边公差（只有技术规格下限），此处只有一个技术规格下限而没有技术规格上限。把技术规格下限和目标值或设计中心之间的距离平分，在中间绘制一条单一的预先控制（P-C）线，如图 21-1C 所示。（另一种备选方法是，有些预先控制的老手使用绿 Y 参数的平均值来代替目标值，但是，这需要预先知道该平均值，而且使用该平均值不如使用目标值那样准确。）

简单的预先控制规则：

(1a) 在技术规格宽度一半的中间绘制 2 条预先控制线。

(1b) 在技术规格上限和 0 的中间绘制一条单一的预先控制线。

(1c) 在技术规格下限和目标值的中间绘制一条单一的预先控制线。

(2) 为确定过程能力，在一行中的 5 个部件必须落在 P-C 线（即绿区）之中，否则，需要使用试验设计技术。

(3) 在生产中，连续而定期地抽样 2 个部件。

(4) 抽样频率。以两次停机间的间隔时间除以 6。

条件	措施
1. 2 个部件在绿区	继续
2. 1 个部件在绿区，1 个部件在黄区	继续
3. 2 个部件在黄区	停止
4. 1 个部件在红区	停止

* 要恢复生产，在一行中的 5 个部件必须全在绿区中。

规则 2：确定过程能力的捷径

过程能力通常是通过从给定的过程里提取 30~100 个部件的样本来确定的，然后就可以计算出 C_p 和 C_{pk} 。预先控制的规则 2 提供了一个从过程中提取一个仅由 5 个连续部件构成的样本的捷径。只有当全部 5 个部件都落在绿区时，才能得出结论（参见有关预先控制理论的章节）：过程能力是一个其值为 1.33 的最小 C_{pk} ，而且有可能还要高些。这给出了生产可以开始的信号。但是，如果 5 个连续部件中的某一个落在绿区以外，那么，过程就是失控的，且生产不允许开始。恰当的措施应当是通过 DOE 研究来确定和减少造成偏差的原因。

规则 3：生产的预先控制：样本规模；继续/停止

一旦生产开始了，样本规模就是从过程中定期提取（参见规则 4）的两个连续部件。下面的规则提供了继续或停止生产的预先控制抽样：

- (1) 如果所有部件都落在绿区内，则继续生产。
- (2) 如果一个部件落在绿区而其余部件落在黄区，则过程仍在控制之中，继续生产。
- (3) 如果所有部件落在同一黄区内，则表明有偏移，应暂停生产，对过程进行调整（根据以往 DOE 研究中确定的过程参数）。如果两个部件落在相反的黄区，则表明有较大偏移，应停止生产，并对产生偏差的原因进行调查。
- (4) 如果有一个部件落在红区，则表明已经产生了一个废品，必须停止生产，以研究产生废品的原因。
- (5) 无论何时停止生产（如按规则 3 和规则 4）并确定和纠正了偏差的原因，都应返回到规则 2，即在开始重新生产之前，5 个连续部件必须落在绿区之内。

规则 4：抽样频率

两个连续部件的抽样频率是通过把两次停机（即两对黄点）之间的平均时间间隔除以 6 来确定的。换言之，如果在上午 9:00 有一次停机（两个黄点），在纠正了过程之后，很快又重新开始该过程，紧接着在中午 12:00 又有一次停机（又有两个黄点），两次停机之间 3 个小时的时间间隔除以 6 得 0.5，因而表明抽样应当每半小时进行一次。另一方面，如果两次停机之间的时间间隔是 3 天，那么，抽样频率应降低到每半天（或每 4 小时）一次。因而，就质量控制的最好传统而论，预先控制提供了胡萝卜加大棒式的政策。胡萝卜就是对非常好的质量实施很少的抽样，大棒就是对较差的质量实施频繁抽样（因而代价高昂）。

在两次停机之间选择 6 个抽样，代表在频繁抽样的时间和代价与生产超公差产品的风险之间的一个折衷（参见有关预先控制理论的章节）。

需要修改预先控制步骤的特殊情况

1. 恒定抽样频率。有些预先控制老手忽略了有关抽样频率的规则（即抽样频率是根据质量水平而变化的），而使用一个恒定的抽样频率（例如，每小时两个连续部件）。理由是改变抽样频率可能使生产线操作员产生混淆，而且对稳定的生产量没有任何益处。但是，如果总是不因较差的质量而增加抽样频率的话，那么， β 风险（即接受一个应废弃的不良产品的风险）就增加了。此外，如果抽样频率永远不因非常好的质量而降低的话，那么，预先控制检查的代价就不必要地增加了。

2. 提高抽样频率的上限。使用规则 4 时，抽样频率不能无限制地加快。例如，如果抽样频率在开始时是每小时 2 个部件，而在 3 个小时内相继出现 2 对黄点，那么，抽样频率必须加快到每半小

时 2 个部件 (3/6)。因此, 如果接下来 2 对以上的黄点在 1 个小时内相继出现, 那么, 抽样频率必须进一步加快到每 10 分钟 2 个部件 (1/6)。该逻辑的下一个推论可能是每分钟 2 个部件。即使对一个新手来说, 这种情况也将表明整个过程严重地失控, 预先控制必须停止, 并马上使用工作量很大的试验设计。因此可以制定一项准则: 只要抽样频率过高 (例如, 高于 15 分钟一次), 就表明出现失控情况, 必须进行 DXE 研究。

3. 抽样频率降低的下限。也可能发生相反的情况, 在这种情况下, 质量水平允许不太频繁的抽样频率。如果在两对黄点之间的时间间隔是一个月, 那么情况会怎样呢? 在理论上, 抽样频率可能是每 22/6 天 1 次或者大致每 4 天 1 次。在这种情况下, 没有预先控制抽样的时间太长了。作为一条实用规则, 不论质量水平有多好, 抽样都不应降低到小于每天 1 次。

4. 从频繁抽样到较少抽样。假定日益恶化的质量水平迫使抽样更为频繁, 例如, 从每小时 2 个部件到每半小时 2 个部件。假定现在已经恢复了好的质量, 那么, 何时可以实施较低的抽样频率呢? 一种方法是以每半小时 2 个部件的速度提取后面的 6 个抽样。如果不再出现黄对, 那么, 就可以重新实施每小时 2 个部件的原始抽样频率。如果以该频率提取的另外 6 个抽样不再产生黄对, 那么, 就可以实施每 2 个小时 2 个部件的新抽样频率。依此类推, 直到达到每天 1 次的抽样频率的下限。

5. 生产间断期间的预先控制准则。生产间断可能是导致质量恶化的一个主要原因, 而预先控制必须对这种变化非常敏感。大规模间断的例子包括模型的变更、装配的变更、工具的变更、长时间停机后的开启以及某一产品的非频繁生产。一旦出现较大的间断, 那么, 预先控制应回到规则 2, 即连续提取 5 个部件, 以查看是否所有读数都落在绿区内, 以便确保较好的过程能力。

6. 缺陷水平低于 100ppm 的预先控制。尽管预先控制良好, 并且超过了控制图, 但是, 对于缺陷水平低于 100ppm 的产品来说, 其用途就很有有限了。为了能够检测到目前几个公司所达到的 50、10

或者 5ppm 这么低的缺陷水平, 2 个部件的定期抽样规模可能还不够大, 即使经常进行这样的抽样也是如此。在这种情况下, **控制产品的最好方法就是控制生产产品的过程**。因此, 就可以使用预先控制来控制 and 监控关键过程参数而不是产品参数。另一种方法就是把抽样规模由 2 个部件增加到 2 组, 每组 5~20 个部件。

预先控制理论

根据预先控制的机理, 什么是它的阿尔法 (α) 和贝塔 (β) 风险呢?

α 风险

α 风险就是废弃一个好过程的风险。图 21-2 是一个过程的最坏情况, 其技术规格宽度等于过程宽度, 即 C_{pk} 为 1.0。假定一个正态分布, 从正态分布表得到绿区中曲线包围的面积是 86%, 而在每一黄区和红区中的曲线所包围的面积是 7% (取自正态分布表)。

因此, 一个部件落在一个黄—红区的概率为 7%。另一部件落在同一黄—红区的概率也是 7%。这样, 使用乘法法则, 两个部件同时落在同一黄—红区的概率为 $7\% \times 7\% = 0.49\%$ 。这就意味着两个部件落在同一黄—红区的概率大约为 1/200。但是, 两个部件可能位于两个黄—红区的方式有四种, 即两个部件位于同一个黄—红区; 两个部件位于相反的黄—红区; 一个部件位于左黄—红区且一个部件位于右黄—红区; 或者一个位于右黄—红区且另一个位于左黄—红区。因此, 两个部件落在预先控制线以外的总概率大约为 1/50 或者 2%。换言之, 过度修正的风险为 2%; 过程没有问题且两个部件全都落在预先控制线以外完全是偶然的。但有 98% 的概率并非偶然 (或者调整需要的置信水平为 98%)。

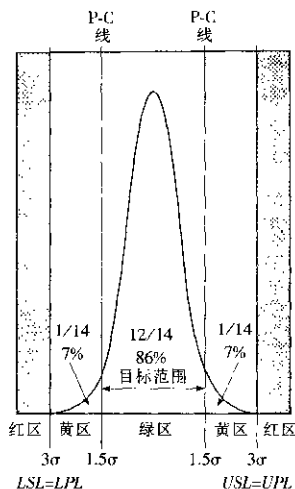


图 21—2 预先控制: α 风险

如果 C_{pk} 是 1.33, 红区缩小为零, 且黄区缩小到 4.6%, 那么, 两个部件落在同一黄区的概率为 $4.6\% \times 4.6\% = 0.21\%$ 。另外, 两个部件可能位于两个黄区的方式有四种, 因此, 两个部件完全由于偶然因子落在预先控制线以外的概率是 $0.21\% \times 4 = 0.84\%$, 得到的是一个比 C_{pk} 为 1.0 的情况更低的过度修正的风险。

β 风险

β 风险是接受某拒受过程的风险。表 21—1 是以百分比缺陷表示的、各种 β 风险的列表, 其中 C_p 分别为 1.0、1.33 和 0.8, 且对每一 C_p 值, C_{pk} 分别为 1.0、0.83、0.67 和 0.5。 C_{pk} 值是分别以目标值 0σ 、 0.5σ 、 1.0σ 和 1.5σ 的 $\bar{X}$ 偏移为基准的。

表 21—1 在不同 C_p 和 C_{pk} 值的情况下, 过程的 β 风险列表

参数	C_p	$D(\sigma) = 0$ $C_{pk} = 1.0$	$D(\sigma) = 0.5$ $C_{pk} = 0.83$	$D(\sigma) = 1.0$ $C_{pk} = 0.65$	$D(\sigma) = 1.5$ $C_{pk} = 0.433$
P_{RT}		0.27%	0.64%	2.28%	6.68%
$P_{黄}$		14.00%	17.51%	29.19%	43.45%
$P_{绿}$	1.0	85.73%	81.85%	68.53%	49.87%
$P_{(A) = q}$		97.77%	95.60%	86.97%	68.21%
$P_{\lambda = q^6}$		87.29%	76.58%	43.27%	10.07%
APQ = β 风险		0.23%	0.19%	0.99%	0.67%
参数	C_p	$C_{pk} = 1.33$	$C_{pk} = 1.1$	$C_{pk} = 0.83$	$C_{pk} = 0.67$
P_{RT}		0.00%	0.02%	0.13%	0.62%
$P_{黄}$		1.56%	7.28%	15.87%	30.25%
$P_{绿}$	1.33	95.41%	92.70%	84.00%	69.13%
$P_{(A) = q}$		99.77%	99.42%	99.22%	89.37%
$P_{\lambda = q^6}$		98.63%	96.57%	95.41%	52.69%
APQ = β 风险		0.00%	0.02%	0.12%	0.33%
参数	C_p	$C_{pk} = 0.8$	$C_{pk} = 0.65$	$C_{pk} = 0.54$	$C_{pk} = 0.4$
P_{RT}		1.64%	3.06%	8.11%	18.41%
$P_{黄}$		21.38%	25.60%	35.35%	43.73%
$P_{绿}$	1.33	76.98%	71.34%	56.54%	37.86%
$P_{(A) = q}$		92.18%	87.41%	71.95%	47.45%
$P_{\lambda = q^6}$		61.35%	44.60%	13.87%	1.14%
APQ = β 风险		1.01%	1.36%	1.13%	0.21%

注:

 $D(\sigma)$ 是把 $\bar{X}$ 与目标值分开标准偏差数。 P_{RT} , $P_{黄}$, $P_{绿}$ —红、黄、绿区分别占总面积(100%)的百分比。 $P_{(A) = q}$ 是(1—落在黄区和红区的概率)的概率。 $P_{\lambda = q^6}$ 是合格(即在两次过程停止之间的时间间隔里没有得到两对黄点)的概率。APQ—平均百分比缺陷 = $P_{\lambda} \times P_{RT}$; APQL—最大百分比缺陷。

资料来源: Rath & Strong, Boston, Mass.

● $P_{\text{红}}$ 、 $P_{\text{黄}}$ 和 $P_{\text{绿}}$ 是每一 C_p 、 C_{pk} 列中红、黄、绿区在正态分布中所占总面积 (100%) 的百分比。这些数字是从正态分布表中得到的。

● $P_{\text{允}}=q$ 是没有得到一对黄点或一个红点的概率 $= 1 - (\text{得到两个黄点的概率} + \text{一个红点的概率})$ 。得到两个黄点的概率是 $7\% \times 7\%$ (参见有关 α 风险的章节) 或者 0.195。由于得到两个黄点的方式有四种, 故 $P_{\text{黄}} = 0.49\% \times 4 = 1.96\%$ 。在一边或另一边得到一个红点的概率 $P_{\text{红}} = 0.135\% + 0.135\% = 0.27\%$ 。于是, $P_{\text{允}} = 1 - (1.96\% + 0.27\%) = 1 - 2.23\% = 97.77\%$ 。

● $P_{\text{A}} = q^2$ 是在两次停机之间的 6 对两部件点中没有得到两个连续黄点的概率。

● ΔPQ (类似于抽样方案中的 AQL) 是期望的平均质量缺陷水平, 且等于 $q \times P_{\text{允}}$, 即 $87.29\% \times 0.27\% = 0.23\%$ 。

预先控制的强大统计能力

表 21-1 清楚地说明了预先控制的强大统计能力和超过控制图的巨大优势。

● 如果 C_{pk} 的值达到 1.33, 那么, β 风险 (即接受不良产品的风险) 将降到零。

● 在 C_p 为 1.33 时, 最高 β 风险出现在 $\bar{X}$ 等于距目标值 (未表示在表 21-1 中) 两个标准偏差 (2σ) 的点处, 且仅为 0.34%。

● 在 C_p 为 1.0 且 C_{pk} 为 1.0 时, β 风险仅为 0.23%。

● 在 C_p 为 0.8 时, 最高 β 风险出现在距目标值 0.5 标准偏差 (0.5σ) 的点处, 且仅为 1.36%。

● 总之, 如果确保 C_{pk} 的值为 1.33, 即预先控制机理中规则 2 的基础, 那么实际上没有任何接受不良产品的风险。

● 在 C_{pk} 为 2.0 时 (表 21-1 中没有给出), 预先控制的安全界限是如此巨大, 以至制造成百上千个部件也不会产生一个废品。不产生废品的记录曾是 3 000 万个部件。

- 总之， α 风险（即拒绝一个好产品的风险）仅为 2%。

绘制预先控制图：减轻操作员的负担

与控制图的情况不同，在预先控制中图形记录并不是强制性的。过程操作员有最简单的规则：

- 两个绿点或者一个绿点和一个黄点：继续生产。
- 两个黄点位于同一黄区：调整。
- 两个黄点位于相反的黄区或者一个点处在红区：停止生产。

没有必要用既冗长又令人厌烦的数据项来分散操作员的注意力。但是，如果由于历史的目的或者作为向客户出示的控制证明以及是由供应商提出而需要预先控制图的话，操作员仅在预先打印的、划分了绿、黄和红区的表格上标记大幅削减了的数据项就可以了。图 21-3 是一个预先控制图的例子，用于控制喷涂机在玻璃上喷涂的铬、镍和金镀层的厚度。有 15 对读数可以很容易地计算出 C_p 和 C_{pk} 值。既不需要人工计算，也不需要手持计算器或昂贵的计算机程序。对于那些不想把钱投在既昂贵而又无效的控制图的过程中的小供应商来说，这是一个很重要的优点。

处理多质量特性

在控制多质量特性（多个绿 Y 值）方面，预先控制也是一个远比控制图更为经济的工具。例如，如果要在铸模机的运行中监控一个 36 模槽铸模的偏差，那么，建立控制图中的控制极限甚至需要 5400 个读数；相反，预先控制只需要每个模槽 5 个读数（或者总共 180 个读数）就可以确定过程的能力。

此外，在确定每一过程的初始过程能力之后，就不必监视所有的绿 Y 值。只有最重要或者变化最大的绿 Y 值需要定期抽样。例

如，在 36 模槽铸模中，只有 2 个或 3 个模槽有较大的偏差，因而需要经常监控，其余的模槽只需偶尔监控或根本就不用监控。

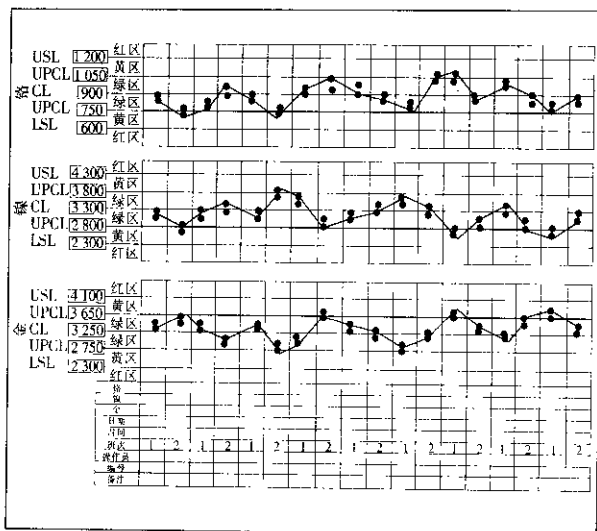


图 21—3 预先控制图：金属喷涂

资料来源：Motorola, Schaumburg, III.

一些公司反对预先控制的原因

正如前面指出的那样，仍有一些公司反对使用预先控制。对于那些非常明智地放弃控制图并采用预先控制的供应商，一些公司以取消它们的供货商资格相威胁。其原因如下：

- 大多数这类公司还不了解预先控制的简单性和精巧性。

- 它们的质量部门不具备较深入的统计学知识，无法比较控制图的缺点和预先控制的强大统计能力。

- 在这些公司中，其高层管理机构受命于控制图，而低层员工则担心说出“皇帝没穿衣服”式的话来。

- 大多数这类公司还没有进行并行研究，以了解预先控制是如何比控制图更快地放弃一个坏的产品且更长时间地保持一个好的产品的。

- 一些公司宣称使用技术规格限抑制了持续的改进。表面上看，这可能是对的。在控制图中，由于进行了改进，控制限确实变得更加接近和更加狭窄。但是，预先控制中并没有规则表明技术规格限的“门柱”不能向内移动。精明的预先控制实践者经常使用更为严格的过程限作为内部技术规格限，以便画出一半公差的预先控制线。因此，随着过程 C_{pk} 值的增加，越来越窄的预先控制线也能反映这种改进。

对那些坚持控制图胜过预先控制的顽固不化的客户来说，下述 3 种方法是行之有效的。

(1) 请他们证明控制图在统计方面的优越性。使用第 20 章和第 21 章的内容来整理你的答复。

(2) 同时使用控制图和预先控制对某个典型过程进行并行研究，向自己证明预先控制在废弃一个坏产品并保持一个好产品方面具有更高的敏感性。

(3) 针对内部控制，以两个部件的抽样规模使用预先控制。然后在抽样规模中加入另外的两个部件以组成一个规模为 4 的子组，并绘制一个控制图来满足外部的客户需求。但是，最后一步应当仅在前两个建议对保守客户未能奏效时才使用。

预先控制与控制图的简略比较

现在，我们能够真正地比较一下控制图的缺点和预先控制的优点了。表 21—2 所示为一个详尽的比较，只有少量的特征需要加以

详细说明。

表 21—2

预先控制胜过控制图的优势

特性	控制图	预先控制
1. 简单性	• 复杂——有关控制的计算	• 简单——预先控制限是技术规格宽度的中间二分之一
2. 易用性	• 困难——强制绘图、解释不清	• 容易——绿区、黄区和红区，对所有工人都是—种实用的方法
3. 数学方面	• 涉及——必须计算 $\bar{X}$ 、 R 、控制极限和过程极限	• 基础——仅需知道如何除以 1
4. 过程条件	• 25 个子组，每个子组需要 4~5 个部件	• 一行中的 5 个绿点确保最小 C_{pk} 值为 1.33
5. 小规模生产运行	• 不适用于 500 个部件以下的生产运行；在可以建立均匀的试验控制限之前，需要 80~150 个部件的抽样	• 可用于 20 个部件以上的生产运行，预先控制线通过技术规格（可以变窄）预先确定
6. 决策	• 滞后——需要绘图并画出许多点	• 实时——取决于绿区、黄区或者红区
7. 控制极限的重新校准	• 频繁——在工业中没有定因系统	• 不需要，除非技术规格的“门柱移内”
8. 机器调整	• 耗时——任何调整都需要另外 80~150 个部件的试验运行	• 即时——基于 2 个部件
9. 抽样频率	• 含糊，任意	• 简单规则——在 2 次停止/调整之间进行 6 次抽样

续前表

特性	控制图	预先控制
10. 辨别力	较弱——无废品时废弃合格品的 α 风险很高；有废品时接受废品的 β 风险很高；与技术规格没有多少关系	极好——无废品时废弃合格品的 α 风险很低，在最坏情况下小于 2%，C_{pk} 为 1.66 时为零；最坏情况下 β 风险小于 1.36%，C_{pk} 为 1.66 时为零
11. 属性曲线	P&C 图未把缺陷模式类型或重要性区别开	属性图可以通过加权缺陷模式和任意的评分尺度转换为预先控制图。
12. 经济性	昂贵——计算、文件、大量的样本、频繁地抽样和长时间的试验运行	廉价——计算简单、最少量的文件、少量样本、质量好时抽样很少、仅通过 5 个部件就能确定过程能力

平均值的使用

控制图使用 $\bar{X}$ 和 $\bar{\bar{X}}$ 的平均值。但是，平均值是靠不住的，就像一个关于猎人的故事中所描述的那样。猎人对准池塘里的一只鸭子，瞄准、射击，偏了，子弹轻轻飞向鸭子的左侧，于是他再次瞄准并射击，又没有射中。这一次子弹飞向鸭子的右侧。但是，平均来说，鸭子是死了！平均值往往冲淡了单个数值中所发生的偏差。预先控制只使用单个值，在测量单个偏差的程度方面更具优势。

结果的反馈

控制图需要在可能观测到趋势之前记录许多单个读数、它们的平均值、它们的范围并需要大量绘图。这是一种滞后的反应。太晚了！对于控制图来说，就好像为了控制一辆汽车的方向而现场调换方向盘一样。预先控制则完全相反，它在决定保持运行或者进行调

整时仅需要两个读数，就像驾驶员本能地对汽车方向盘进行细小的微调一样。

案例研究 1：波焊

在第 14 章有关全析因的案例研究中，一个 DOE 小组已经把印刷电路板上的焊接缺陷从 10 000ppm 降低到 220ppm。在进行了优化、正向控制以及过程确认之后，开始进行预先控制以确保 45：1 的改进能够在将来的生产中得以保持。

DOE 小组决定采取两项措施。为了追踪像 100ppm 这样低的缺陷水平，小组把抽样规模由两部件增加到两组，每组 10 块印刷板。其次，用利克特度量尺度来估算各种类型焊接缺陷的严重程度。缺陷类型被进一步细分为焊接过度缺陷和焊接不足缺陷（见图 21—4），以便可以模拟双边公差。缺陷类型的严重性乘以该类型的缺陷数目，得到一个缺陷分数。焊接过度 和焊接不足的技术规格限均设置为 100ppm，而焊接过度 和焊接不足的 P-C（预先控制）线均设置为 50ppm。

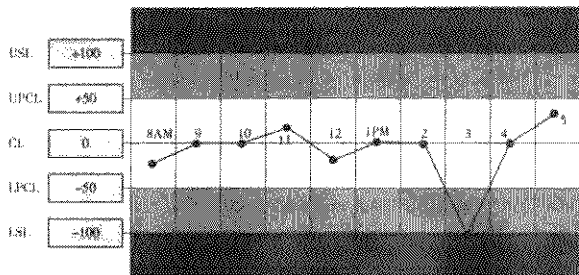
业已证明预先控制图是非常有用的。与它所替代的、烦琐的控制图相比，操作员很喜欢它的简单性。一旦出现了失控的情况，操作员几乎马上就可以发现，并可以查出原因，记录在预先控制图上，同时采取调整或者纠正措施。预先控制在这家工厂中已经成为一种生活方式。

小组练习 1：套管

在有关控制图的第 20 章中，案例研究 1 涉及一个长度要求为 0.500 英寸 $\pm$ 0.002 英寸的套管。该案例研究表明，尽管潜在的缺陷率为 7%，但控制图却显现出一种稳定的、受控的情况。

过焊编码 缺陷

焊接不足	+20				×					×	×
焊接稍微不足	+20										
过度焊接	+5										
覆盖	+5										
总计					+20						+40



欠焊编码 缺陷

拆焊连接	-100								×		
欠焊	-20	×	×								
除湿	-20										
吹孔	-5					×	×				
总计			-40				-10			-100	

图 21—4 预先控制图：波焊操作

资料来源：Motorola Inc., Seguin, Texas.

问题

(1) 如果在该过程中使用了预先控制，那么它会不会检测到某种失控的情况，而与此相对的是，控制图却检测不出该失控情况呢？

(2) 预先控制会在多长时间内停止该不良过程?

答案

(1) 套管的规格为 0.500 英寸 $\pm$ 0.002 英寸, 因此, 预先控制线将会在 0.500 英寸 $\pm$ 0.001 英寸处。在第 20 章中, 表 20—1 给出了几对落在预先控制线以外的单个读数, 表明该过程应被废弃。

(2) 预先控制的规则 2 指出, 在开始时, 一行中的 5 个部件应落在预先控制线之内, 以保证过程的能力。表 20—1 表明, 在最先的 5 个部件抽样中有 2 个部件位于 P-C (预先控制) 线的边界线上, 而一个部件位于 P-C (预先控制) 线以外。该不良过程一开始时就会被预先控制停止, 而不允许像控制图那样继续 60 个读数并作出继续生产的错误决定。

小组练习 2: 传感器电容

在第 20 章中, 案例研究 2 描述了一个最小技术规格为 31 皮法、最大为 45 皮法的传感器电容。图 20—2 给出了控制图数据以及 $\bar{X}$ 图和 R 图, 该图表明传感器电容严重失控并应立即停止该过程。

问题

如果使用预先控制来监视该产品, 是否会存在失控的情况并发出停止生产的命令?

答案

在该例子中, 预先控制限是 34.5 皮法和 41.5 皮法。除了

两个读数之外，没有单个读数超过 P-C（预先控制）线。该过程的 C_p 为 2.0， C_{pk} 为 1.92，表明有一个非常好的过程能力。预先控制永远也不会停止这样好的过程，而控制图却错误地废弃了该过程。

小组练习 3：导线连接器

自动导线连接器把一根头发粗细的导线连接到晶体管的管芯和接线柱上，已经证实该连接器要开始实施预先控制。导线连接的牢固性已通过对导线进行破坏性拉力试验而得到检验。不论在提起哪一侧接头之前，连接强度的技术规格为最小 6gm、最大 14gm。确定过程能力的最初 5 个连续抽样部件有以下几个读数：8.7gm，9.0gm，9.4gm，8.9gm 及 10gm。

问题

(1) 预先控制线的值是多少？

(2) 以 5 个部件的样本为基础，是否具有良好的过程能力？

(3) 假定在全面生产开始时，两次停机（两对黄点）之间的平均时间间隔为 12 小时，那么在开始时用于连接强度试验的两个连续部件的抽样频率应当是多少？

(4) 在接下来的生产中，从过程中定期提取的两个连续抽样部件的结果如表 21—3 所示。（为了便于练习起见，撇开了要求一行中的 5 个部件落在绿区范围内以重新建立过程能力的规则。）对于每一个样本，你将采取怎样的措施？是继续生产还是停止生产？

(5) 在第 4 问中你发现了什么样的非随机趋势？解释你的结论。

(6) 如果连接强度没有 14gm 的技术规格上限，且假定目标连

接强度为 11gm. 那么, 单一控制线将画在何处?

表 21—3

小组练习 3: 导线连接器

抽样号	部件 1	部件 2	措施
1	9.4	9.0	_____
2	9.0	8.8	_____
3	8.9	8.6	_____
4	8.5	8.1	_____
5	8.4	8.0	_____
6	8.0	8.0	_____
7	8.0	7.6	_____
8	7.5	7.3	_____
9	13.0	13.0	_____
10	12.0	12.0	_____
11	11.6	11.4	_____
12	11.0	10.8	_____

答案

(1) 预先控制线位于 8gm 和 12gm 处。

(2) 在最初抽样中的所有 5 个部件都位于绿区, 因此, 稳固地建立了过程的能力。

(3) 抽样频率应为 12/6 或者每两个小时一次。

(4) 在样本 1 到样本 7 以及样本 10 到样本 12 上继续生产。在样本 8 和样本 9 上停止生产。(在实践中, 在 8 号样本之后, 过程必须进行调整或者纠正, 在重新实施 2 部件抽样之前, 必须有 5 个以上的部件落在绿区范围内。9 号样本的情况也是如此。)

(5) 有下述 3 种非随机趋势: 第一, 连接强度正变得越来越

449

小，直至 8 号试验样本后采取了纠正措施为止（原因可能是连接头污染或失去连接力量）；第二，在每一个样本中，第二个读数几乎总是低于第一个读数；第三，9 号样本似乎有些矫枉过正了（连接强度太高）。

（6）在目标连接强度为 11gm 且只有一个单一技术规格下限的情况下，单一预先控制线应位于目标值 11gm 和技术规格下限 6gm 中间，即位于 8.5gm 处。

有关预先控制的 DOE 小组的预先控制的问题

（1）在开始统计过程控制之前，是否使用了所有必要的 DOE 工具来解决问题？

（2）在开始预先控制之前，是否应用了正向控制以及过程确认等技术？

（3）在着手实施预先控制之前，是否同用户一道利用像质量功能展开、散布图和响应曲面法这样的技术来确定正确的技术规格和公差？

（4）是否使用了预先控制的规则 2 来确定过程的能力（最小 C_{pk} 值为 1.33）？

（5）是否训练生产线操作员在其各自负责的过程中实施预先控制，并在简单的、预先打印的预先控制表格上记录他们所发现的问题？

（6）在生产期间的每次重大停顿之后或者在一次预先控制停机之后，是否重新建立了过程能力（规则 2）？

（7）抽样频率是否是以两次停机之间的时间间隔为变化依据的？

（8）是否对这样的抽样频率建立了上限和下限？

（9）是否只使用预先控制处理重要的参数（绿 Y）？

（10）根据预先控制图，是否有提供给客户的（或者根据供应



商的要求提供的) 客观证据, 以表明保持了很好的质量?

在 DOE 项目评审期间的高层管理的问题

(1) 为什么客户宁愿用复杂而无效的控制图却不用简单而更有效的预先控制? 采取了什么措施, 以使客户确信它们需要预先控制?

(2) 是否不再继续加重供应商们在控制图方面的负担, 并指导它们使用预先控制?

(3) 是否定期向客户提供预先控制图, 以作为我们保持良好质量的证明?

(4) 是否定期要求供应商提供预先控制图, 作为它们维持良好质量的证明?

【注释】

[1] 当然, 假定已经使用在第 7 篇中详细讨论过的优化技术, 通过试验设计, 建立了正确的技术规格和公差。

第 8 篇

试验设计与可靠性的联系



第22章

多环境强化应力试验： 对可靠性的有效性如同 试验设计对质量的有效性一样

◆质量与可靠性

质量与可靠性之间存在着根本的差别。质量是产品在其离开工厂时（即零时刻）的优良品质。可靠性有两个附加量纲：时间和应力。时间指产品的寿命；应力指各种应力或者环境因子，例如，温度、振动、湿度等，这些因子之间可能会相互影响、相互促进并加速现场失效。因此，尽管质量和可靠性是同义使用的两个术语，但是对客户、保修费用、产品召回以及赔付责任的法律诉讼来说，可靠性更为重要。

◆传统可靠性实践的基本缺点

令人遗憾的是，当前的可靠性实践是有缺陷的。而这些缺陷在每年的可靠性论文以及会议中被讨论得沸沸扬扬。下面就是一些传统的可靠性实践：

- 试图去预计不可靠性，而不是把注意力放在实际消除现场失效上。
- 使用数学模型、公式以及计算机软件来预报零部件和产品的可靠性特性。这些数学模型最多不过是对实际情况的一种逼近，在最坏的情况下则完全脱离实际。它们还会使可靠性的实践者感到迷惑。
- 以错误的信念进行毫无理性的试验：一个星期试验 200 个部件与一个部件试验 200 个星期会得到相同的可靠性估计！
- 盲目地遵循已经落后了 15 年的军用可靠性规范。
- 很少使用零部件降额（降低应力）。
- 没有加大产品的应力且没有在可靠性说明中结合环境。

◆适度的可靠性进展

在过去的 20 年间，可靠性研究尽管已经取得了一些进展，但是，这些进展与基准可靠性相差很远，并且很少被可靠性实践者使用。

1. 可靠性预计研究：以美国军用手册 217E、英国电信手册（BT）、法国国家电信中心（CNET）以及像“预计程序”（predictor）之类的软件为基础。这些预计可能会与实际的现场可靠性相差 1/25（最有利的情况下）到 4 倍（最不利的情况下）——摆幅为 100:1。

2. 故障模式影响分析（FMEA）：日本和欧洲比美国使用得更



广泛。故障模式影响分析只不过是一个基础性的理论研究，它只是把一个工程师的猜测、意见、直觉和理论进行了量化。

3. 环境科学研究所 (Institute of Environmental Science, IES) 指南:

- 他们在生产中要求 100% 的筛选。
- 他们限制热循环的温度极限以及振动谱系。
- 他们对“有效性”的调查评价基本是“一般”。

4. 单一应力试验:

- 热循环: 低速率的热变化, 大量的循环。
- 振动: 正弦波, 单一频率。

5. 单一环境序贯试验:

- 多个环境, 但是按顺序每次使用一个。
- 低应力加速。
- 完全失去的应力之间的相互作用。
- 可靠性推测随每次试验而变化很大。

革命性的可靠性试验原则

幸运的是, 人们正在抛弃在过去 20 多年间逐渐演变而来的可靠性实践而采用革新的技术。下面的原则就是巩固这种革新的基础。

● 以高的置信水平预计现场可靠性即使是可行的, 也是非常困难的。尽管多环境强化应力试验 (MEOST)¹⁸ 可以作出这种预计, 但是, 在 5 年、10 年或 20 年时间里作出可靠性结论之前, 该项技术还要在现场经历半年和一年的考验。

● 目标是消除失效、周期, 而不是过多地考虑量化不可靠性和失效率。因此, 可靠性数据将由其自行处理。

● 目标不是进行“成功的”试验。其目标不是通过某个试验, 以便使产品在未改变意图之前更快地投入生产。与此相反,



其目标是故意导致失效，因为只有通过失效才能查出设计中的薄弱环节。

- 每一个零部件都有其薄弱环节，一些零部件在其应力失效极限和试验应力之间存在一个较大的防卫带，其他零部件则只有一个较小的防卫带。后者便是薄弱环节，可以通过使其应力大大超过设计应力而消除该薄弱环节。

- 为了缩短失效前时间以及减少失效率的变化，可以用更高水平的强化应力来强制或者加速失效。图 22-1 给出了逐步升高的应力，直到多环境强化应力试验，并描述如何缩短通常较长的、

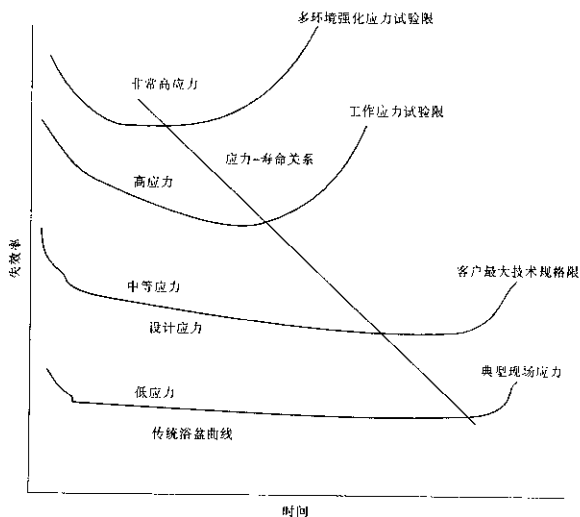


图 22-1 说明应力与期望寿命之间关系的浴盆曲线

资料来源：KeKi R. Bhote and James McLamm at Amway Corp., Ada, Mich.

恒定不变的浴盆可靠性曲线，该曲线描述了失效率与时间之间的关系¹⁸。

● 应力极限是：

◇把各个单一环境（应力）结合起来以促进由交互作用而引起的失效。一个简单的例子：温度极限可能不会单独导致失效，湿度极限也可能不会单独导致失效，但是，温度极限和湿度极限一起会导致失效。

◇大大超过设计应力水平以加速失效。后一个原则是多环境强化应力试验（MEOST）的基础。

►失效与应力的关系

产生失效的应力水平由 Miner 方程（或幂等律）来决定：

$$D \approx ns^{\beta},$$

式中， D 为损坏（失效）累积； $\approx$ 为近似于； n 为应力循环数； s 为应力； β 为乘方因子，通常为 1~12。

就可靠性而言， β 因子通常为 10，这意味着应力程度具有 10 倍于应力循环（通常为热循环）数的检测失效的能力。在作为应力代表的热循环和振动中都使用 Miner 方程。

热循环的应力加速

图 22-2 显示了在热循环的速率提高时，达到失效的热循环数是如何压缩的。¹⁹ 例如，

- 5℃/分钟的速率需要 400 个热循环就可以达到失效。
- 25℃/分钟的速率仅需要 4 个热循环就可以达到失效。

这是一个 100 : 1 的时间压缩！



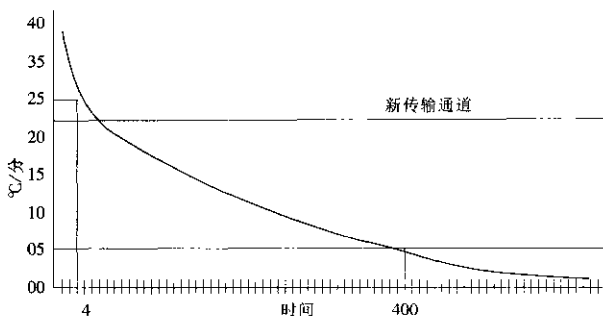


图 22—2 热循环速率 (应力) 与热循环数

资料来源: Smithson S. "Effectiveness and Economics Yardstick for ESS Decision," *Proceedings of the 1990 Annual Technical Meeting*, © Institute of Environmental Sciences and Technology, (847) 255-1561.

振动的应力加速

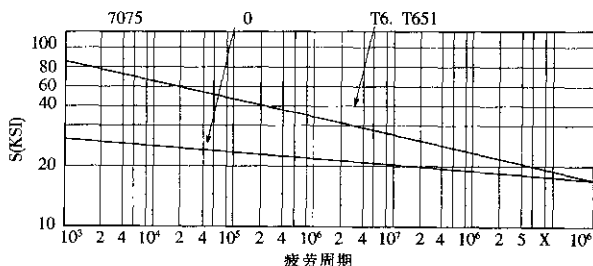
Miner 方程不仅应用于热循环,也应用于所有类型的应力,其中包括机械的、电气的、振动、湿度或任何其他激励源。图 22—3 所示为一个 S/N 图 (其中 S 为应力, N 为到达失效的循环数),摘自样品的抗拉伸疲劳试验。²⁰该图表明当拉伸力以线性方式提高时,对 T6 (热处理)、未退火的 T651 铝而言,到达失效的循环数按指数 (β 由曲线斜率获得) 方式降低。

- 40K PSI 的应力需要 2 000 000 个循环到达失效。
- 80K PSI 的应力仅需要 2 000 个循环到达失效。
- 拉伸应力的提高仅为 2 : 1, 却导致到达失效的循环 (时间) 数降低 1 000 倍。

由机械引起疲劳的加速因子比由电气或热引起疲劳的加速因子高得多。在现场失效的零部件具有某些可能导致应力增加的缺陷。



这种应力集中，即使是由一个很小的缺陷引起的，也可能是好零部件中应力的两倍或者三倍。这种应力的增加可能会导致边缘零部件的疲劳失效比没有多少固有应力的坚固零部件的疲劳失效快几个数量级。



成形铝—逆向混合

图例:0=退;T6=热处理

图 22—3 到达失效的拉力强度与要求的疲劳周期数

资料来源: Steinberg, *Vibration Analysis of Electronic Equipment*. Copyright © 1973, John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

正是这种以 2 : 1 和 4 : 1 的因子加速应力的能力，使得到达失效的时间能够在热循环这样的试验中以 10 : 1 加速，在振动这样的试验中以 1 000 : 1 加速。这正是高效可靠性试验的关键。热循环和振动是两个最为重要的应力代表，并得到了最优秀的可靠性实践者的广泛使用。

表 22—1 总结了基准热循环的显著特征，表 22—2 总结了基准振动的显著特征^[1]。

如果图 22—1 中的推断是以 40℃/分钟而不是 25℃/分钟作出的，那么，失效时间可能是一个循环或以 400 : 1 压缩的时间。对试验时间和试验设备而言，这是一个非常大的节约！

表 22—1

基准热循环的显著特征

- 较高的温度梯度：25℃/分直至 60℃/分。尽可能少的热循环。继续热循环，直到“最薄弱环节”的零部件失效为止。
- 最高的可能温度极限——仅不足“基本技术极限（FLT）”（即破坏极限）的 20%~30%。
- 热冲击：仅用于元器件（由热到冷的即时变化）。
- 使用液氮来加速热循环（与压缩机正相反）的速率。
- 空气扰动。
- 在每一温度极限处的 10 分钟最小停顿时间。
- 连续监视绿 Y（输出）参数以检测间歇。

表 22—2

基准振动的显著特征

- 随机振动。
- 具有 6 个自由度的三轴振动系统。
- 六轴激励（ x ， y ， z 和三个旋转）同时激活所有产品的响应模式，提供需要任何方向的矢量加速来强化潜在的缺陷。
- 从 2Hz 到 2.5KHz 及更高频率的响应。
- 最小产品振动为 20G（RMS）。

高加速寿命试验

在过去的 15 年中，一个称为高加速寿命试验（HALT）的系统曾被几家在工业界领先的大公司（尤其是航空航天和电子工业的大公司）所使用。在应力试验中，该系统仍在不断革新。

原则

- 不打算预计现场可靠性。
- 不打算通过某个应力试验，而是要使应力试验失败。
- 迫使一个产品刚好达到基本技术极限（FLT）。
- 目的是要激励失效而不管失效是否在现场发生。
- 通过加压的方法超越设计极限，使较弱的元器件失效，并

使坚固的元器件保持无损。

- 极高的强化应力也能缩短失效前时间。

方法

(1) 确定每一基本技术极限应力的较高破坏极限、较低破坏极限。

(2) 挑选尽可能多的样本。

(3) 挑选应力 (通常是热循环和振动)。

(4) 每次步进提高应力。

- 振动以 25%GRMS 为级距。

- 热循环以 5℃~10℃为级距, 由冷开始, 然后到热。

- 电压的级距等于操作范围的 20%。

- 在适当的情况下, 其他应力的级距等于操作范围的 20%。

(5) 继续步进提高应力, 直到一半或一半以上的样本失效为止。

(6) 分析失效的每一个原因, 直到根原因为止; 采取纠正措施。

(7) 继续步进提高应力, 直到达到基本技术极限为止。

(8) 处理下一个应力, 并重复上述步骤。

(9) 如果需要 (通常是热循环和振动) 的话, 运行复合应力。

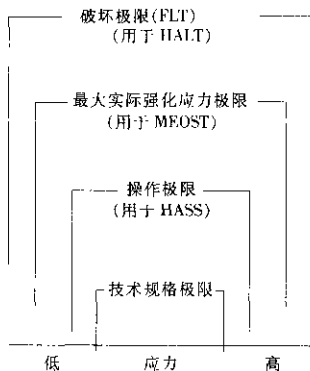
(10) 以破坏水平的应力重新进行整个循环, 一次、两次甚至多次评估纠正措施和设计变更。

HALT 的优缺点

HALT 的最大优点是, 它至少在一贯使用它的那些行业 (尤其是航空航天业) 中取得了几年的成功。根据其加快投入生产和投放市场的时间以及在投资回报方面的表现, HALT 的费效比是较高的。

另一方面, 在查找由机械引起的薄弱环节方面, 尤其对于大型产品, HALT 的用途却不大。此外, 在消除现场失效方面的作用也不及多环境强化应力试验 (详见后面章节的比较)。

其最大的缺点，也是最具争议之处是，它坚持把一个单一应力（或者一个复合应力）推至其终极水平，即直到基本技术极限（FLT）。图 22—4 描绘了在一个应力连续体上的产品规格极限、操作极限、多环境强化应力试验实际强化应力极限以及 HALT 破坏极限（即基本技术极限）之间的分布。首先，应明确有关这些极限的定义。



- 图例：1. FLT（基本技术极限）= 破坏极限 - HALT 应力
 2. 最大实际强化应力 = MEOST（多环境强化应力试验）应力
 3. 操作极限 = HASS 应力

图 22—4 技术规格极限与操作极限 (HASS)、最大实际过应力 (MEOST)、破坏极限 (HALT)

极限的定义

(1) 用于 HALT 中的破坏极限 (DL) 是：

- 当产品崩溃时达到基本技术极限的应力水平。
- 或者，当少量的应力增加将导致失效数较大增加时的应力水平。

(2) 操作极限 (OL) 是损坏薄弱零部件但不损坏好零部件的

应力水平。操作极限用于 HASS (下一个技术)。

(3) **最大实际过应力极限 (MPOSL)**, 该应力水平具有一个低于破坏极限的充足防卫带, 以防止人为失效的累积, 防卫带通常低于破坏极限 20%~33%。

最大实际强化应力极限用于多环境强化应力试验。加压到达基本技术极限可能会引入完全人为的 (或是愚蠢的) 失效, 这些失效即使在极限条件下也永远不会在现场发生。追踪每一个单独的、随机的、人为的失效就如同追踪鬼影一样, 纯粹是浪费时间、人力和财力。任何熟悉失效分析物理学的人都能体会到: 调查极原因, 并且一旦发现它就通过打开和关闭它来确认该原因是多么困难。

HALT 的成功是其支持者打出的王牌, 而 HALT 的失败却被隐藏起来。甚至在其成功当中, 10:1 的可靠性改进在给人深刻印象的同时, 也仅仅降低了失效, 而没有像 HALT 目标所宣称的那样消除失效。该技术的其他缺点是: 多半仅使用两个应力 (热循环和振动); 试验的样本规模较大, 以及不能预计现场可靠性。

表 22-3 总结了 HALT 的优点和缺点。

表 22-3 高加速应力试验的优缺点

优点	缺点
• 在航空航天及电子工业中有很好的跟踪记录 • 与大多数其他方法相比, 是一种更好的可靠性改进技术 • 缩短了整个设计周期时间 • 大大缩短了应力试验时间 • 便于更早投产和现场开工 • 降低成本: 非常好的费效比 • 可用于所有级别的产品 (从零部件、组件直到系统) • 可得到商用试验设备	• 对于机械产品, 尤其是规模较大的机械产品很少取得成功 • 可靠性的改进并不像宣称的那样惊人 • 分析每一个由应力导致的失效 (直到根原因) 并进行纠正的代价和时间过高 • 把应力推到技术极限可能产生在现场永远也不会重现的人为失效 • 主要依靠两个应力: 热循环和振动 • 推荐大量的试验部件 • 对每一设计更改, 都要重复 HALT 循环 • 部件无法从现场返回, 以作进一步的调查 • 无可靠性预计能力



高加速应力筛选 (HASS)

HALT 主要是用于产品的设计阶段。在大多数情况下, 高加速应力筛选是生产中的一个 100% 的试验。它的突出特征是: 先决条件是产品设计阶段中完全成熟的 HALT。定义: 高加速应力筛选是一种使用应力的筛选, 它所使用的应力远远高于在现场 (包括海运) 中的应力。应力的水平应高到足以加速暴露那些可能出现在现场的相关缺陷, 并把 80%~90% 的有用寿命保留在剩余产品中。

表 22-4 简略地总结了生产中高加速应力筛选应力试验的优缺点。

表 22-4 高加速应力筛选应力试验的优缺点

优点	缺点
1. 筛选试验, 高加速应力筛选原则的充分确认	• 生产中 100% 的高加速应力筛选试验是解决可靠性问题的一种非理性的方法
2. “注入样本”——故意引入缺陷来试验高加速应力筛选的筛选能力	• 尽管进行“筛选试验”, 也不能保证所有缺陷都能检测出来
3. 不论对振动还是对热循环, 筛选时间都很短	• 没有足够的证据表明, 有用的产品寿命不会被高加速应力筛选降低
	• 只使用一个或两个应力——热循环和振动
	• 筛选试验、注入样本和优化是冗长、烦琐的过程, 该过程可能增加开支并延长生产循环时间
	• 没有尝试也不可能预计可靠性

多环境强化应力试验 (MEOST)

历史

多环境强化应力试验的最早应用是在 20 世纪 60 年代美国国家宇航局 (NASA) 进行阿波罗系列太空探索的那一段紧张和不可靠



的日子里。苏联的太空发射仿佛翱翔着飞进了天空，而美国的太空发射却“扑通”一声落进了大西洋！除了登月舱，阿波罗太空船的每一个部分都失效了。它本应载着两名宇航员到达月球并返回地面。如果登月舱失效了的话，月球就会由于这两个人的到来而造成人口过剩。但是，登月舱并未失效，因为它曾在陆地上经受了多环境强化应力试验的考验。

从那时起，在直升飞机、飞机发动机、汽车、火车、水处理和空气处理系统、航空航天以及商务运作中，一直都在使用多环境强化应力试验来大大改善其可靠性。这就是为什么摩托罗拉公司用在地球与火星、木星和土星的远距离探测器之间以及在像哈勃这样的天文望远镜中实现主要无线通信的设备，可以在5年、10年乃至更长时间内航行数百万英里而不失效的原因之一。多环境强化应力试验目前正在应用于摩托罗拉雄心勃勃的铱星计划。在铱星计划中，66颗卫星可提供由地球上任何一点到任何其他点之间的无线电话通信而无需电话交换。笔者20世纪70年代在摩托罗拉推广多环境强化应力试验并把它应用于几个项目中，从其自动化和工业电子集团着手，取得了200:1的可靠性改进，并且在一个不便公开的重要客户的应用中，实现了7年间50万英里行车使用无失效的成果。

多环境强化应力试验的目标

(1) 开发一个应力试验，从中得到的可靠性远远超过所有其他技术；

(2) 在应力试验中重复相似产品在现场中遇到的相同失效；

(3) 急剧提高应力水平，迫使失效更早发生；

(4) 在生产中开发一个费效比高的试验筛选，防止生产和供应缺陷导致设计完整性的下降；

(5) 开发一个合理的有效方法来预计寿命时间的可靠性；

(6) 开发一个方法来评估与DOE紧密联系的工程更改的有效性；



(7) 开发一个方法与 DOE 一道解决长期性的质量或可靠性问题；

(8) 开发一个方法来对付持续的“未发现故障”问题；

(9) 开发一个方法来降低与价值工程学紧密联系的产品成本；

(10) 在公司范围外采用多环境强化应力试验，以帮助供应商及客户。

多环境强化应力试验的优点

(1) 它从实际上消除了现场失效，而不是试图预计产品的不可靠性。

(2) 通过半年及一年的现场暴露，它可以以 90% 的置信度推断出可靠性的一个量化数字，以满足管理的需要。

(3) 它减少了整个设计周期的时间、人力、成本、场地、试验设备以及能源损耗。

(4) 在促进新产品快速投放市场方面，它取得了决定性的竞争优势。

(5) 对于 DOE 来说，它是一个强有力的解决问题的工具。

(6) 对于 DOE 来说，它是一个强有力的、评估工程更改有效性的工具。

(7) 对于价值工程学来说，它是一个强有力的、削减未来产品成本的工具。

多环境强化应力试验方法：路径图

序言（适用于零部件、模块、组件和系统的路径图）

(1) 确定零部件降额（应力降低）的适当程度。降额意味着设计工程师应当把零部件应力从供应商推荐的最大定额至少降低到 25% 的定额。为了节约几个便士而把零部件应力推到最大定额，可能给公司带来的开支将是因现场失效和顾客不满意所带来的花费的

100 到 1 000 倍。决不应回避一个综合的降额研究。

(2) 列出并按优先级排列相似产品的现场失效。

应力与水平

(1) 尽可能多地列出现场可能出现的应力或环境。

(2) 把可能同时出现的应力与可能单独出现的应力分开。

(3) 按优先级排列同时发生的应力，并且在可能的情况下把这些应力减少到五个或六个。

(4) 确定在现场可能出现的最高应力水平（最好通过仪器）。

(5) 确定设计规格极限。把这些极限与客户要求的应力以及最高现场应力相比较，使用三者中最高的应力作为设计极限。

(6) 确定每一应力的基本技术极限，基本技术极限又称为破坏极限。

(7) 确定每一应力的操作极限，操作极限定义为可能导致薄弱零部件而非产品剩余部分失效的应力水平（通常比设计极限高出设计极限与破坏极限之间距离的三分之一）。

(8) 确定每一应力的最大实际过应力极限（MPOSL）（通常比破坏极限低出设计极限与破坏极限之间距离的三分之一）。

样本量

(1) 在原型阶段：3 个可修理部件；5~10 个不可修理部件。

(2) 在小批试产及后续阶段：10 个可修理部件；30 个不可修理部件。

多环境强化应力试验的 7 个步骤

步骤 1：单一应力到设计应力极限

● 逐级提高应力：由热循环或振动入手。

● 对于热循环：以 40℃/分为速率，10℃~15℃为级距；停顿时间：在每一极限处 10 分钟，从冷开始然后到热循环。

● 对于振动：25%GRMS 为级距。



- 如果没有失效发生，则进入步骤 2。
- 如果仍有一个以任何失效方式发生的单个失效，则分析并纠正每一个失效。利用 B vs. C 试验通过“打开”或“关闭”失效来加以确认，然后进入步骤 2。

步骤 2：单一应力直到最大实际强化应力极限

● 继续实施步骤 1 超过设计应力极限，如果部件可修理的话，就对步骤 1 的相同部件实施步骤 1；如果部件不可修理的话，就对新部件实施步骤 1，直到到达最大实际强化应力极限。不要到达破坏极限。

- 如果在强化应力区没有失效，则：
 - ◇ 应力类型或应力速率不适当。
 - ◇ 没有适当地进行试验。
 - ◇ 失效需要多种环境。
 - ◇ 对该应力来说，设计是健壮的。
- 如果每个失效模式只有一个失效，则忽略该失效。
- 如果每个失效模式有两个或更多失效，则分析并纠正失效，通过 B vs. C 试验加以确认。

步骤 3：多环境强化应力试验到最大实际强化应力极限

- 选择多种环境（参见应力和水平）。
- 准备一个多环境强化应力试验试验方案：应力水平和顺序（连续的或间断的）。
 - 如果可能的话，使用与步骤 1 和步骤 2 相同的部件，并组合应力。
 - 从设计应力开始，然后以 5~10 为间隔逐级提高应力，直到达到最大实际强化应力极限。
 - 该过程与步骤 2 的过程相似，只是具有多种环境。步骤 1、步骤 2 和步骤 3 成功的关键标志是能够再现相似产品在现场中所出现的相同失效模式。
 - 通过“注入缺陷”（故意引入的缺陷）完成另一轮步骤 3 的



有效性。

- 成功的结果表明，设计可以用于批量试产。

步骤 4：在批量试产中进行多环境强化应力试验

目的：为确保设计更改、工具供应商、工艺、装配等没有反过来影响设计的可靠性。

- 利用步骤 3 的指南，对新部件进行步骤 4 多环境强化应力试验。
- 一个成功的结果意味着设计可以用于完全投产。

步骤 5：在后续生产中进行删失多环境强化应力试验

目的：确保设计中可靠性的完整性不因生产过程、制作工艺以及原材料而降低。

应力：仅挑选步骤 3 中所发现的两个或三个最重要的应力来进行删失多环境强化应力试验，但要把应力加到最大实际强化应力极限。

抽样频率：永远不使用 100% 的抽样。每周抽样一次（或者至少每月抽样一次）。使用连续向上或向下的抽样方案来降低好结果的抽样频率和提高不良结果的抽样频率。

样本量：与步骤 4 相同——10 个可修理的部件和 30 个不可修理的部件。

证明删失多环境强化应力试验的有效性：在初始删失多环境强化应力试验之后，用“注入缺陷”进行一轮步骤 5。

步骤 6 和 7：良好的现场返回部件的多环境强化应力试验

● 从现场召回好的部件：服务 6 个月后召回（步骤 6）；服务一年之后召回（步骤 7）。

● 尽可能多地获得这些部件的现场历史数据，例如使用时间、操作情况、历史日志（如果可得到的话）。

- 对这些现场返回部件进行步骤 3 的多环境强化应力试验。

- 使用尽可能多的样本——步骤 6 为至少 10 个，步骤 7 为 10 个。



● 把应力水平与现场时间（即 0 时间、6 个月和 12 个月）投影到威布尔（Weibull）图上。

通过推断查看 90% 的置信水平是否满足目标可靠性（见图 22-5）。在图 22-5 中，垂直标度是复合应力，因而，没有给出用于初级单一应力的步骤 1 和 2。水平标度是现场暴露时间。工作矩形区是实际上不应发生失效的时间应力区域。这意味着一个投放市场并以完全设计应力维护的产品在目标产品寿命（例如 10 年）内应当没有失效。简言之，多环境强化应力试验应当消除黑点——在工作矩形区中的极点。

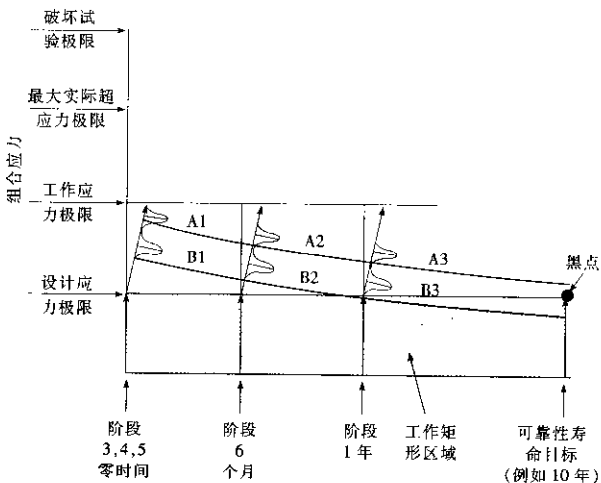


图 22-5 时间上的失效应力

步骤 3、4 和 5 中的多环境强化应力试验从设计应力和零（现场）时间开始，并连续进行直到达到最大实际强化应力极限。假定最高失效分布（在 3 个步骤中）为 A_1 。类似地，在步骤 6，当多环



境强化应力试验应用于在 6 个月和 1 年后收回的良好现场部件之后, 失效分布分别为 A_2 和 A_3 。在应力与时间的威布尔图中推断出这 3 个失效分布的较低尾部将会得到 1 个高于黑点且目标可靠性寿命为 10 年的点。另一方面, 如果失效分布在零时间、6 个月和 1 年分别为 B_1 、 B_2 和 B_3 , 威布尔图上的推断将不能消除黑点, 说明不充分的可靠性 (大致为 2 到 3 年)。

表 22—5 对多环境强化应力试验的 7 个步骤进行了简要的总结, 描述了每一步骤在产品投产的位置、应力类型、应力水平、样本量、允许的失效数以及试验频率。

表 22—5 多环境强化应力试验的七个步骤

步骤序号	在产品投产的位置	应力类型	应力水平 ^a	样本量 ^b	允许失效数	试验频率
1	原型	单一	直到设计极限	$R: 3 \sim 10$ $N.R.: 5 \sim 15$	0	一次
2	原型	单一	直到 MPOSL	$R: 3$ $N.R.: 5 \sim 15$	1	一次
3	原型	多种	直到 MPOSL	$R: 3$ $N.R.: 5 \sim 15$	1	两次: 第二次用注入缺陷
4	工程/生产 ^d 批量试产	多种	直到 MPOSL	$R: 10$ $N.R.: 30$	1	一次
5	生产 ^d	删失多种	直到 MPOSL	10 或尽可能多	1	1) 每周一次或每月一次 2) 使用注入缺陷缺陷再做一轮
6 和 7	现场	多种	直到 MPOSL	10 或尽可能多	1	1) 6 个月 2) 1 年后

a 应力等级: MPOSL=最大实际过应力极限。

b 样本量: R =可修理部件; $N.R.$: 不可修理部件。

c 允许失效数: 0 或 1 指仅一个失效模式。

d 步骤 3、4 和 5, 于设计应力开始, 然后逐级提高应力直到最大实际过应力极限。



总有效性比较: HALT/HASS 与 MEOST

在全面地描述了 HALT/HASS 以及 MEOST 的原则、目的、优点和方法之后,就可以比较这两种改进可靠性的主要方法的有效性了。当然,假定不考虑除了上述两种技术之外的其他方法(军队的、商业的和国际的等)。

表 22—6 是 HALT/HASS 和 MEOST 的比较,总共比较了 20 个特征。该表本身就可以说明问题。在几乎所有的特征方面,MEOST 都优于 HALT/HASS。

表 22—6 有效性比较: HALT/HASS 与 MEOST

特征	HALT/HASS	MEOST
● 可靠性预计能力	无	有: 每 6 个月和 1 年有两个现场返回部件
● 产品多样性	主要是航空航天以及电子	所有产品
● 适用于所有产品水平	是	是
● 最高应力水平	破坏试验	最大实际过应力
● 环境/应力数	一般两个	一般 3~6 个
● 允许失效数	无	直到最大应力无失效,每一失效模为过应力
● 暴露的交互作用	部分	几乎全部
● 样本数	“尽可能多”	3~10 个可修理产品
● 设计中的试验次数	每一次设计更改就试验一次	2 次: 仅对主要更改
● 失效分析/纠正	传统方法	试验设计
● 失效纠正费用	昂贵,费时	对于每个单一失效无费用需求
● 纠正措施的确认	传统方法	无参数的 B vs. C 试验



续前表

特征	HALT/HASS	MEOST
● 缩短设计周期时间	是	大大缩短
● 减少设计人力/成本	是	大大减少
● 批量试产评价	无具体措施	有
● 生产评价	HASS	删失 MEOST
● 生产中的样本量	100%	一般 10 个；每月一次
● 注入缺陷	有	有
● 已用尽产品寿命的百分数	最小 10%	无——仅使用样本
● 整体跟踪记录	较好	非常好

多环境强化应力试验与试验设计间的紧密联系

多环境强化应力试验正越来越多地被笔者用于纯可靠性改进以外的任务中。这些试验与 DOE 一起进行——它们之间有着一种紧密而又自然的联系。

1. 解决难以捉摸的问题

如果一个问题难以捉摸，比如间歇现象或可能在工厂中从未被发现、只有在现场才会被发现的问题，那么，DOE 与多环境强化应力试验的组合是一个一定会取得成功的方法。

这里的目的是要提取生产中的部件并使它们经受一次或两次组合应力试验，直到发生的失效与客户所见到的失效相同为止。于是失效前时间或者失效前应力成了响应、输出或者绿 Y。可用一种或



者多种上述技术来确定哪些输入变量导致了问题的发生。笔者已经解决了一些似乎颇令客户迷惑的这类问题。

2. 处理“未发现故障”或“无明显缺陷”

在工业中，最容易受挫的和最难以解决的问题是称为“未发现故障”(NTF)、“无明显缺陷”(NAD)和具有其它简称的缺陷/失效问题。通常这些问题占到现场问题的25%~50%。客户报告了一个失效，维修人员或维修站却没有发现故障。这些问题有时是由于客户缺乏理解，有时是维修人员分析得不够深入，有时则是一个“间歇故障”。任何情况下，对这类申诉的传统跟踪服务方法都是费时的。如果涉及到产品责任的法律诉讼，那么其费用可能就更加高昂。这类情况的历史表明，50%~60%的“未发现故障”都是间歇故障。因而处理“未发现故障”和间歇故障的最好方法就是让部件经受多环境强化应力试验，提高应力水平，直到一个间歇失效变成明显的、永久性的失效为止。

3. 评价工程更改的有效性

工程师们在不断更改方面是声名远扬的。更改以修订版A，B，C等表示。修订版的字母常常贯穿整个字母表，然后又从AB，AC等从头开始！有时需要进行三四次更改才能解决同一个问题。

这里，可以用B vs. C试验来评价每一项工程更改。B代表一个好的设计；C代表当前设计。要求试验3个B和3个C（以随机顺序），然后按从好到坏排列输出。如果3个B在优良品质上超过了3个C且无重叠，那么，可以以95%的置信水平断言B设计比C设计好。

当输出（或者响应、绿Y）是一个变量时，很小的样本量（3个B和3个C）就够了。如果输出是一个属性，以高成比率



百分数表示，或输出具有较低的缺陷水平（例如，100ppm 或更少），则 3 个 B 和 3 个 C 的样本就太小了，它们没有足够的区别能力来检测 B 和 C 之间的差别。此时，DOE 技术（B vs. C）与多环境强化应力试验的组合技术就派上用场了，其目的是通过使用多环境强化应力试验让这些 B 和 C 都经受逐级提高的应力水平以把一个属性（如成品率或缺陷水平）转换为一个变量。于是，失效前时间或失效前应力水平就成为一个变量，这样 3 个 B 和 3 个 C 的样本量就够了。

4. 成本的降低

多环境强化应力试验的主要目的是改进可靠性。然而，许多怀疑者认为，无论是对客户还是对公司来说，把可靠性贯彻到底都是得不偿失的。不要理会这些批评。质量和可靠性是没有终结的。它们并不耗费资金，它们节约资金。但是，通过多环境强化应力试验所达到的可靠性水平能够凭借降低产品成本而得到保持吗？答案是肯定的，且答案是通过多环境强化应力试验。

遵循多环境强化应力试验的步骤 6 和 7，当达到一个额定的可靠性目标时，可以增加步骤 8，具体如下：

- 制定一个产品的降低成本目标；
- 列出所有高成本的零部件、模块和配件的清单；
- 按优先顺序排列这些清单，针对每一项目，在成本降低的可能程度与可靠性降低的危险、加工、供应商风险以及试验成本之间进行权衡。
- 把价值工程技术¹⁵运用于每一项目，找出一种候选方法，以提供不受当前设计限制的、项目需要的功能。
- 通过该候选方法，使用多环境强化应力试验和 B vs. C 的一个组合，以检验设计更改的有效性。
- 以失效前时间或失效前应力作为输出（或者绿 Y），排列 3 个 B（降低成本的新设计）与 3 个 C（可靠的老设计）。如果 3 个 B



不全部劣于 3 个 C，那么，成本降低的新设计就没有降低产品的可靠性。

结论

本章利用已经过证实的多环境强化应力试验技术，为可靠性实践者提供了一种简单的、非数学的、非常及时且费效比高的方法。尽管如此，为了掌握该技术，实践者应该从零部件、插件板或者模块级开始，然后再将该方法扩展到系统级和供应商级。

【注释】

[1] 振动应力时间比热循环少一或两个数量级，与热循环的几小时相比只有 5~10 分钟。



第 9 篇

设计试验工具使用的逻辑顺序



第23章

解决长期性问题的序贯试验设计工具的 案例研究

►引言：初级实践者在实施 DOE 时遇到的困难

我们的几家客户公司表示，它们关心的问题之一是：尽管介绍性的 DOE 专题讨论看起来似乎相当简单、无须数学基础并且切实可行，但是，DOE 工具在解决现实问题时却是难以实施的。这种看起来似乎是二歧性的问题是有一些原因的。下面几节将分别阐述这些原因。

1. 根原因红 X 可能位于 5~6 层 深处

第一个原因是，每一个问题（每一个绿



Y) 都有一个红 X 原因。但是，一次 DOE 研究不足以发现红 X。被揭示的红 X 本身可能是一个新的绿 Y 或者一个派生问题，它可能需要进行第二次 DOE 研究以揭示其最底层的红 X。在某些解决问题的练习中，总共可能需要 2~3 个序贯试验。但在另外一些练习中，**根原因红 X** 可能埋藏在 5~6 层之深，需要 5~6 个序贯实验。

例如，假定一个部件搜索试验揭示某个配件是一个红 X。该配件需要一个后续的部件搜索试验来指出某个特定的元器件是一个红 X。如果该元器件不能拆卸并重新安装，那么，可能需要进行成对比较实验来标识该红 X 参数。这可能需要把 DOE 研究转移到元器件供应商那儿，通过多变量或者产品/过程搜索试验等方法来分析该问题。

2. 未“打开”和“关闭”试验

第二个非常普遍的原因是，没有通过验证性的 B vs. C 试验来确认一个 DOE 改进的持久性。试验员对感觉到的改进充满了侥幸心理，以为该改进是真实而又持久的。在历史上充满了这种乐观主义，随之而来的却是连续的失望，因为改进已经消失了。B vs. C 试验的目的就是故意地在改进的产品（或过程）与老产品（或过程）之间来来回回地打开和关闭问题。

3. 未建立现实的公差

第三个原因是，在变量搜索实验后，对每一个重要变量“好”水平的假定，而没有通过像散布图或响应曲面法这样的优化 DOE 研究来建立一个围绕着该中心值的实际公差带。因此，生产中不实际的公差可能造成没有改进。

4. 没有正向控制

第四个原因是，允许操作员、技术人员或维护人员通过灵感来控制一个过程。随意的调节违背了正向控制的规则，正向控制要求每一个在先前的 DOE 研究中指出的重要过程参数，均应通过人员、地点、方法和时间这样一套体制得到保护。笔者已经发现，尽管正确使用了前面的 DOE 技术，但是，缺乏正向控制纪律的现象随处可见。

5. 没有过程确认（实验开始时“噪声”太多）

不能过分地夸大过程确认在控制大量外部质量差错方面的重要性。实际上，我们认为，甚至在开始 DOE 研究之前就应当实施过程确认，这样，不良生产习惯、环境波动、测量仪器偏差等“噪声”因子都可以保持在严格的范围内，以确保 DOE 研究的成功。此外，在 DOE 研究和全面生产重新开始之后，应进行过程的再确认。

连接 DOE 工具的连续性

本章描述了两个综合的案例研究，它们把向零缺陷迈进的几个 DOE 工具以一种逻辑的顺序连接起来。每个案例的结果都不是一两天就能够得到的，需要进行数个序贯试验。要有耐心，要守纪律，要坚韧不拔！而结果是可观的——1 000：1 的质量改进、每年节约十多万美元、4：1 的周期时间缩短以及竞争上的明显优势。（为了保护公司的机密，一些包含重要变量水平的具体细节将不予提供。）

案例研究 1: 波焊

产品: 波焊
缺陷水平: 2 400ppm~3 500ppm; 平均值: 2 970ppm
成本: 106 800 美元/年
缺陷模式: 欠焊 80%; 虚焊 14%; 错焊和气泡等 6%
问题持续时间: 14 个月
迄今为止的解决 抓过程、头脑风暴、因果图和控制图
问题技术:

DOE 试验 1: 多变量

试验设计从多变量开始, 有四个偏差族:

- (1) 时间到时间: 上午 9:00; 11:00; 下午 2:00。
- (2) 部件到部件: 3 块控制面板; 每个控制面板 5 块电路板。
- (3) 电路板内: 左; 左中; 右中; 右。
- (4) 缺陷类型: 欠焊; 虚焊; 其他。

图 23--1 以图形方式给出了按偏差族划分的缺陷数。在总共 140 个缺陷的情况下, 缺陷率为 3 111ppm, 大大超过了捕捉最小值规则, 即 2 970ppm 缺陷率的 80%。

● 图 23--1A 说明从时间到时间或从控制面板到控制面板没有重大偏差。

● 相反, 图 23--1B 说明处在控制面板两端的电路板 A 和电路板 E 存在非常大的偏差 (5 : 1), 存在许多缺陷; 而处在控制面板中部的电路板 B、电路板 C 和电路板 D 只有很少的缺陷。

● 图 23--1C 也说明每一电路板的左右两个四分之一区域存在较大的偏差 (3 : 1 到 10 : 1), 存在许多缺陷; 处在中间的两个四分之一区域只有很少的缺陷。

● 图 23--1D 说明欠焊是主要的缺陷模式。



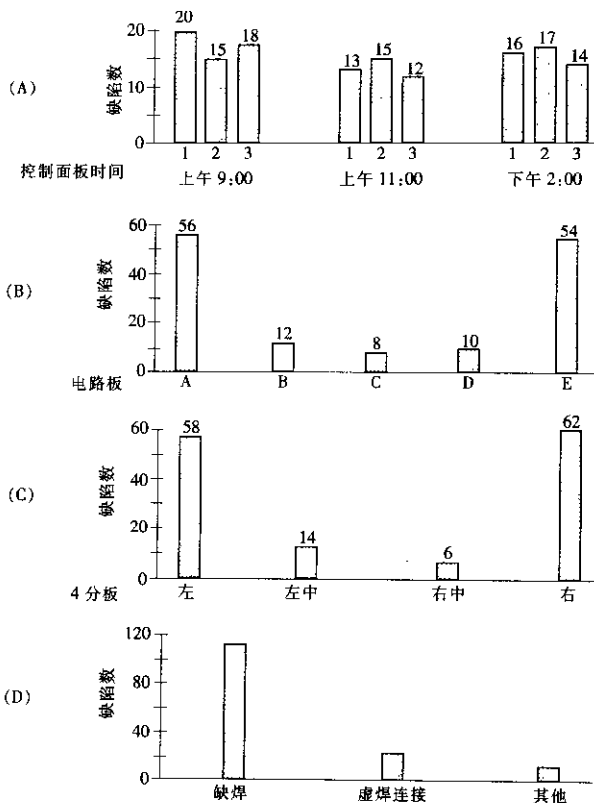


图 23—1 波焊过程的多变量图

(总缺陷数: 140)

资料来源: BRK Electronics, Aurora, III.



DOE 试验 2：集中图

多变量清楚地表明在控制面板中（每一端）和电路板中存在最多的缺陷（红 X 族）。构造一个集中图，用控制面板和电路板中的孔位置和元器件位置来描绘缺陷集中情况。每段时间间隔和每一缺陷类型都用不同的颜色进行编码，以便于分析。

集中图（更流行的叫法是麻疹图）揭示：

- 欠焊的最高数目位于每一电路板右端的边缘连接器区域。
- 第 2 个欠焊的集中区域位于每一电路板左端的 IC 引线处。
- 这两种欠焊类型暗示了波焊过程参数或许是可能的原因。
- 排在第 2 位的缺陷模式是虚焊的连接，集中在相同的两个孔的位置上。

DOE 试验 3：电路板的成对比较

从试验 1 中的 9 块控制面板中挑选 4 块控制面板，比较每块控制面板里面的“好”电路板以及外面的“不好”的电路板。在波焊完成时，控制面板的弯曲通过以弯曲度作为测度的绿 Y 来检验。

表 23—1 表明，外面的电路板具有重复的较高的弯曲（0.150 英寸~0.300 英寸），而里面的电路板具有较小的弯曲（0.008 英寸~0.015 英寸），它们之间几乎是 20:1 的差别。

纠正措施：弯曲表明，控制面板的固定不够，而且在预热区的波焊机温度可能过高。使用更结实的固定器对控制面板进行固定，并且把预热区温度降低 10°F。

DOE 试验 4：虚焊连接的成对比较

另一个成对比较研究是针对 4 块电路板进行的。在这 4 块电路板中，每块电路板在相同位置上均存在着重复性的虚焊连接，把这些位置与具有良好焊接连接的相邻位置进行比较。



表 23—1

里外电路板的成对比较

控制面板号	成对	弯曲 (从控制面板中心到每块电路板中心)	
		左边的电路板	右边的电路板
1	里面的电路板	0.008 英寸	0.010 英寸
	外面的电路板	0.150 英寸	0.017 0 英寸
2	里面的电路板	0.010 英寸	0.015 英寸
	外面的电路板	0.210 英寸	0.250 英寸
3	里面的电路板	0.015 英寸	0.012 英寸
	外面的电路板	0.300 英寸	0.260 英寸
4	里面的电路板	0.013 英寸	0.010 英寸
	外面的电路板	0.190 英寸	0.160 英寸

资料来源: BRK Electronics, Aurora, III.

根据 (1) 孔中的镀层或者 (2) 引线上的焊层, “良好的”和“不良的”焊接连接之间没有什么差别。但是, “不良的”焊接连接比“良好的”焊接连接具有更高的孔与引线的直径比, 其比例因子的范围为 1.4 : 1.0 到 1.7 : 1.0。

纠正措施: 已明确其原因是由于在印刷电路板上采用单一钻头尺寸钻孔造成的, 因此, 引线直径较小的元器件比大多数引线直径较大的元器件更容易产生虚焊。在后续电路板生产中规定使用较小的钻头尺寸。

DOE 试验 5: B vs. C 确认

接下来进行 B vs. C 试验, 以确定在试验 3 和试验 4 中采用的纠正措施的有效性。

C 过程: 当前的固定、预热区以及钻头尺寸。

B 过程: 坚实的固定、低 10°F 的预热区以及在有影响的孔位置使用较小的钻头尺寸。

在波焊机上以随机的顺序处理 3 个 C 控制面板和 3 个 B 控制面



板，绿 Y 为焊接缺陷数。结果是：

控制面板	B ₂	B ₃	B ₁	C ₃	C ₁	C ₂
缺陷数	5	3	2	12	16	13

结论：B 过程优于 C 过程，置信水平为 95%。缺陷水平从 3 111ppm 降到 660ppm，几乎是 5 : 1 的改进。

DOE 试验 6：变量搜索

此时，试验中的一个逻辑序列理应是使用产品/过程搜索来调查波焊过程参数对欠焊的作用。但是，在那些日子里，笔者还没有开发出产品/过程搜索，于是，研究小组继续进行变量搜索试验。表 23-2 所示为具有 8 个因子的试验设计。在最好水平和临界水平各挑选 10 个控制面板。原始的全部最好和全部最差以及两个反复实验（以随机顺序运行的）的结果如表 23-2A（设计）、表 23-2B（阶段 1）、表 23-2C（阶段 2）所示；表 23-2D（求交运算）、表 23-2E（析因分析）以及表 23-2F（ANOVA）描绘了变量搜索试验的剩余部分。

结论

（1）阶段 1 是成功的， $D : \bar{d}$ 的比值为 7 : 1，大于最小值 1.25。该结果表明试验中选择了正确的因子。

（2）阶段 2 表明因子 A, B, F, G 和 H 是不重要的，因而，它们的交互作用也是不重要的。

（3）因子 C, D 和 E 是重要的，且它们的交互作用是不能忽略的。应当使用全析因来量化它们，分析表明一个三因子的交互作用 CDE 是红 X，因子 C 是粉红 X，因子 E 是浅粉红 X。为得到最佳结果，应保持在最佳水平。

（4）成对比较研究使缺陷率已从 3 111ppm 降到 660ppm。变量搜索把缺陷水平降到大约 80ppm (4×10^6) / ($10 \times 100 \times 5$)，一个 39 : 1 的改进。



表 23—2

变量搜索：波焊

A: 试验设计			
编码	过程参数	最好水平	临界水平
A	热空气切割器压力 (psi)	14	10
B	预热区温度	断面 1	断面 2
C	焊剂密度 gm/cc	0.9	0.8
D	传送带速度 (英尺/分)	4	6
E	传送带角度 (°)	7	5
F	焊接温度 (°F)	480	450
G	焊接停顿时间 (秒)	3.5	3.0
H	焊泡高度	1.2	1.0

R. 阶段 1 (以随机次序进行): 缺陷

全部最好水平	全部临界水平
4	42
5	46
2	51

显著性检验

a) 3 个“全部最好”好于 3 个“全部临界”

b) $D-46-4=12$; $d-9+3)/2=6$ D: $\bar{d}$ 比值 = 42; $6-7$; 1 ; $>$ 最小值 1.25

C. 阶段 2

因子组合	绿 Y	因子重要性
$A_M R_B$	13	不重要
$A_B R_M$	38	
$B_M R_B$	12	
$B_B R_M$	39	
$C_M R_B$	32	重要, 与另一个变量
$C_B R_M$	15	
$D_M R_B$	20	
$D_B R_M$	21	
$E_M R_B$	25	重要, 与另一个变量
$E_B R_M$	22	
$F_M R_B$	10	
$F_B R_M$	40	



续前表

因子组合	绿 Y	因子重要性
$G_M R_H$	9	不重要
$G_H R_M$	42	
$H_M R_H$	8	不重要
$H_H R_M$	38	

判断极限—中值+2.776×d 1.81

中值最好的—4；中值临界的=46

于是，控制极限（最好的）—4+2.776×6 1.81=13.2和—5

控制极限（临界的）—46+2.776×6 1.81=55.2和35.8

两中值间的中线—25

D. 求交运算		
因子组合	绿 Y	是否证实?
$C_M D_M E_M R_B$	43	是——R 不重要
$C_B D_B E_B R_M$	7	是——R 被证实为不重要

E. 析因分析

		C			
		最好的	临界的		
D 最好的	E 最好的	4 13 9 5 12 8 中值 2 10 7 =8	32	40	
	E 临界的	25	21	D 最好的=86	
	E 最好的	20	22	4.6	
D 临界的			42 38 42 46 39 38 中值 51 40 43 =42	D 临界的=99	
		15		57	
		C 最好的=68	C 临界的=117	C 偏差=49	
		D 最好的=86	D 临界的=99	D 偏差=13	
		E 最好的=86+42=82	E 临界的=46+57=103	E 偏差=21	



F. 方差分析 (ANOVA) 表

C	D	E	CD	CD	DE	CDE	输出
+	+	+	+	+	+	+	8
-	+	+	-	-	+	-	32
+	+	-	+	-	-	-	25
-	+	-	-	+	-	+	21
+	-	+	-	+	-	-	20
-	-	+	+	-	-	-	22
+	-	-	-	-	+	+	15
-	-	-	+	+	+	-	42
-49	-13	-21	+9	-3	+9	-53	

资料来源: BRK Electronics, Aurora, III.

DOE 试验 7: B vs. C 确认

尽管变量搜索试验的求交运算显示出了成功,但是,还应当再在 10 天之后以随机顺序运行 10 个 B 和 10 个 C 控制面板,以使对改进作 B vs. C 验证。

C 过程: 坚实的固定、低热区、较小的钻头、标准波焊过程参数。

B 过程: 与上述相同,此外,焊剂密度为 0.9gm/cc,传送带为 4 英尺/分,传送带角度为 7°。

结果: (缺陷总数)

控制面板	B ₁	B ₂	B ₂	C ₂	C ₃	C ₁
	3	3	4	25	31	29

结论: B 波焊参数以 95% 的置信度好于 C。

DOE 试验 8: 通过全析因进行优化

试验 6 显示出焊剂密度 (B)、传送带角度 (E)、传送带速度 (D) 之间有一个三因子的交互作用。在这种交互作用之下,正确的

后续最优化试验应是响应曲面法或者调优运算技术 (EVOP)。但是, 研究小组不熟悉调优运算, 因而选择了使用 2^2 全析因来做一个小型的最优化。从试验 6 中选出 2 个最重要的因子, 每个因子具有两个水平:

(1) 焊剂密度: 为 0.9gm/cc (由变量搜索试验得到) 以及一个可能更好的水平 0.85gm/cc 。

(2) 传送带角度: 为 7° (由变量搜索试验得到) 和一个可能更好的水平 6° 。以反复和随机的方式使用每一种组合因子来运行 10 块控制面板。图 23-2 给出了 2^2 全析因的结果 (缺陷数)。

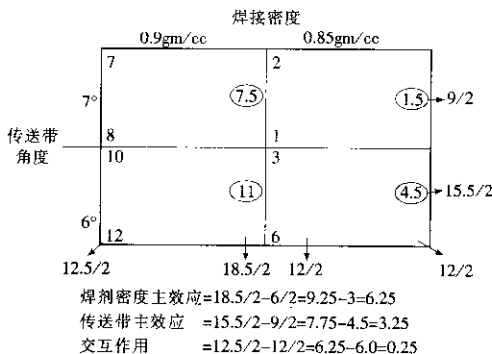


图 23-2 波焊过程的 2^2 全析因

结论:

- 焊剂密度应从原始变量搜索的最好水平 0.9gm/cc 改为 0.85gm/cc 。
- 但是, 传送带角度应保持在原始变量搜索的最好水平 7° 。
- 两个因子之间的交互作用是可以忽略的。



● 组合因子产生一个 30ppm 的焊接缺陷水平 一个 2.5 : 1 的进一步改进, 并且比原始的 3 111ppm 减少了 103 : 1。

DOE 试验 9: 用散布图进行最后的优化

试验 8 的全析因表明焊剂密度是红 X, 而且焊接缺陷的 ppm 数可以通过把焊剂密度从 0.9gm/cc 降低到 0.85gm/cc 而得到进一步降低。

最优化过程的下一个逻辑步骤是散布图。焊剂密度在 0.74gm/cc~0.9gm/cc (原始变量搜索水平) 范围内变化。图 23-3 所示为散布图结果, 其中绿 Y 直接用 ppm 表示。如果最大允许缺陷水平为 20ppm, 那么, 该散布图表明焊剂密度应限制在 0.78gm/cc~0.84gm/cc 之间的水平, 目标值为 0.81gm/cc。为了使用预先控制进行监控, 焊剂密度的 P-C (预先控制) 线应分别位于 0.795gm/cc 和 0.825gm/cc 之间。

正向控制

用一个正向控制计划 (见表 23-3) 来控制变量搜索试验中识别出的 3 个重要因子: 焊剂密度、传送带角度以及传送带速度。

表 23—3 正向控制计划: 波焊

对象 (关键因子)	控制人员	控制方法	控制地点	控制时间
焊剂密度 0.179 5gm/cc~ 0.925gm/cc	焊接技术人员	专用重力计	焊剂容器	每小时一次
传送带角度 7°±10%	焊接技术人员	机器设置	传送带	每天一次
传送带速度 4 英尺/分 ±10%	焊接技术人员	计数器	传送带	每次模型改变

资料来源: BRK Electronics, Aurora, III.

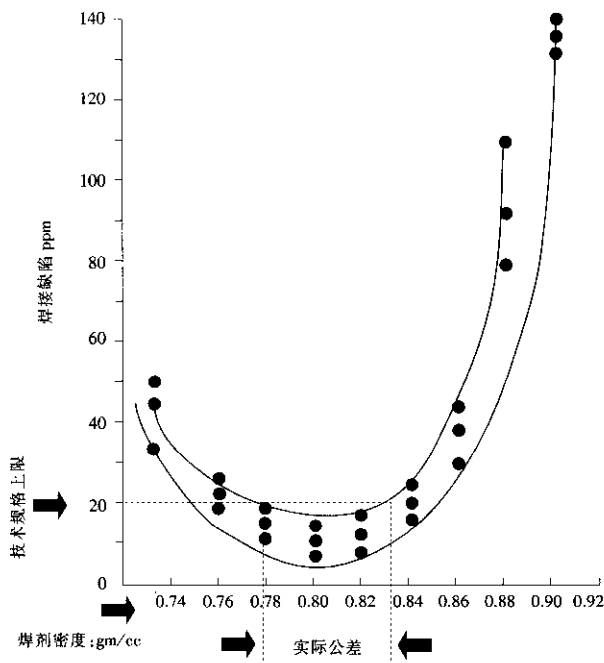


图 23—3 焊剂密度散布图

资料来源：BRK Electronics, Aurora, III.

过程确认

过程确认是最后一项要求，其作用是确保外部质量问题不脱离 DOE 增益的轨道。在许多情况下，过程确认应由跨学科小组在

DOE 研究之前实施,以使这些质量问题的不可控“噪声”不超过 DOE “信号”的纯度。在怀疑违背了良好的生产习惯时,这一点尤其重要。表 23-4 给出了与波焊过程相关的质量外部问题清单,这些质量外部问题必须在过程确认之前由跨学科小组进行审核。

表 23-4 质量问题:过程确认清单

操作员	设备
1. 是否禁止工人对过程进行未授权的调整?	13. 是否运用了全面的预防性维修?
2. 正向控制日志是否得到认真维护?	14. 预防性维修与“损坏时再修理”之比是多少?
3. 焊接技术人员是否经过资格认证并定期重新认证?	15. 仪器与产品精度是否具有 5:1 的最小比值?
管理/监控	16. 仪器校准是否就位?
4. 是否消除了害怕的氛围?	17. 是否注意到了摩擦、噪声、振动、过热、电压波动等?
5. 是否对操作员进行了 DOE 培训?	环境
6. 是否有视听方面的质量警报信号?	18. 环境/湿度控制是否满足要求?
7. 操作员是否有权关闭不良质量的生产线?	19. 化学品和灰尘控制是否满足要求?
良好生产习惯	20. 水和空气纯度是否满足要求?
8. 是否有书面的标准操作习惯?	21. 静电负荷保护是否满足要求?
9. 标准操作习惯是否太难遵守?	22. 是否禁止在工厂内吸烟?
10. 是否正在使用标准操作习惯?	
11. “保管”是否完善?	
12. 设置更改是否快速和准确?	

波焊案例研究的总结

试验	平均缺陷水平 (ppm)
1. 开始时	2 970
2. 多变量确认时	3 111
3. 成对比较纠正后 (B vs. C 确认)	660
4. 变量搜索后	80
5. 全析因后	30
6. 散布图后	<20

►案例研究 2：透声罩式高频扬声器的灵敏度

产品：汽车工业的透声罩式高频扬声器
技术规格：灵敏度：88.1db±1.5db 从 4.5kHz 到 20kHz
缺陷水平：40%需要返工
成本：>100 000 美元/年
问题持续时间：从几年前开始生产至今

DOE 试验 1：整套组件的部件搜索

详见表 23—5。

表 23—5 透声罩式高频扬声器灵敏度的部件搜索

BOB: 87.25db; WOW: 85.2db

部件	结果
第 1 次和第 2 次拆卸/重装	无变化
前端交换	无变化
磁交换	部分
透声罩交换	部分
磁和透声罩交换	完全相反
	红 X

DOE 试验 2：透声罩的成对比较

详见表 23—6。

表 23—6 透声罩式高频扬声器的成对比较

透声罩	活动机壳	机壳透声罩	机壳线圈	缠绕宽度	在空气间隙中的位置
#1 好	0.156	0.072	0.081	2.2	0.35
#1 坏	0.176	0.063	0.108	2.4	0.30
#2 好	0.159	0.081	0.081	2.3	0.35

续前表

透声罩	活动机壳	机壳透声罩	机壳线圈	缠绕宽度	在空气间隙中的位置
#3 好	0.162	0.075	0.090	2.4	0.35
#3 坏	0.179	0.081	0.100	2.2	0.30
#4 好	0.162	0.076	0.086	2.2	0.35
#4 坏	0.166	0.077	0.090	2.2	0.30
#5 好	0.160	0.070	0.089	2.2	0.30
#5 坏	0.173	0.078	0.092	2.35	0.25
#6 好	0.160	0.078	0.088	2.35	0.35
#6 坏	0.167	0.067	0.098	2.2	0.30
终结计数	12	0	6	0	9

资料来源: Philips, Dendermonde, Belgium.

成对比较试验清楚地表明:

- 活动机壳为红 X, 其终结计数为 12。
- 在空气间隙中的位置为粉红 X, 其终结计数为 9。
- 机壳线圈为浅粉红 X, 其终结计数为 6。

DOE 试验 3: 磁的成对比较

完成该试验的样本超过了要求的样本数。但是, 图 23—4 显示了良好的磁感应和不良的磁感应之间的明显区别, 前者很少降至 9.0 高斯以下, 而后者很少升至 9.0 高斯以上。显然, 磁感应就是

50

磁方面的红 X。

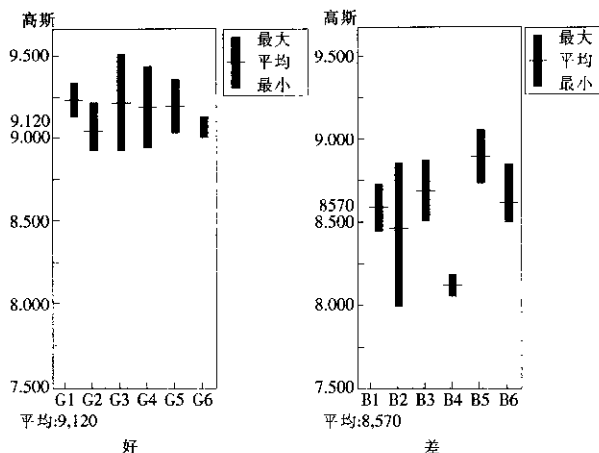


图 23—4 成对比较: 磁感应

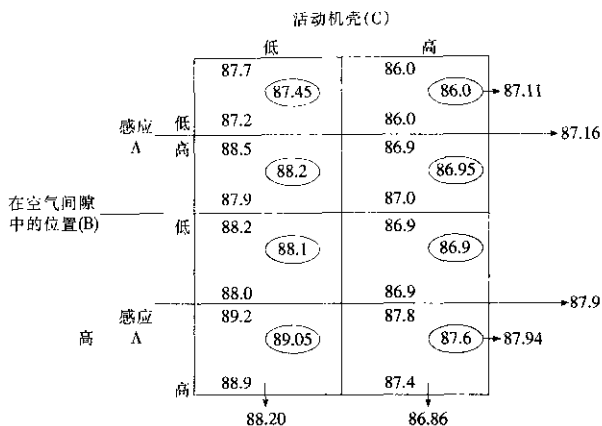
资料来源: Philips, Dendermonde, Belgium.

DOE 试验 4: 全析因

DOE 小组继续进行 2^3 全析因, 以便对透声罩的活动机壳、在空气间隙中的位置、磁感应以及它们之间的交互作用等的相对重要性进行量化 (见图 23-5)。

结论

必须对所有 3 个因子尤其是活动机壳 (红 X), 完成附加的 DOE 工作, 它表明需要把活动机壳降低到一个更低的机壳。



主效应：活动机壳 (C) -1.34

主效应：空气间隙位置 (B) -0.74

主效应：感应 (A) -0.83

交互作用 (析因分析未显示)

AB=0.06

AC=0.01

BC=0.18

ABC=0.09

图 23—5 全析因：透声罩式高频扬声器

DOE 试验 5：透声罩胶重量的多变量研究

既然确定了透声罩上的活动机壳为红 X，就需要对透声罩的胶重量偏差进行多变量研究（如图 23-6 所示）。研究表明，透声罩到透声罩而非时间到时间的偏差是红 X 族，偏差高达 25mg 或者大致分布于胶重量范围的 20%。

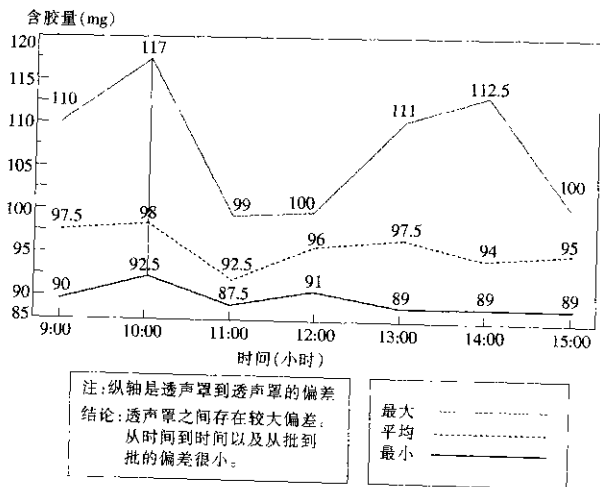


图 23—6 多变量:透声罩的胶重量

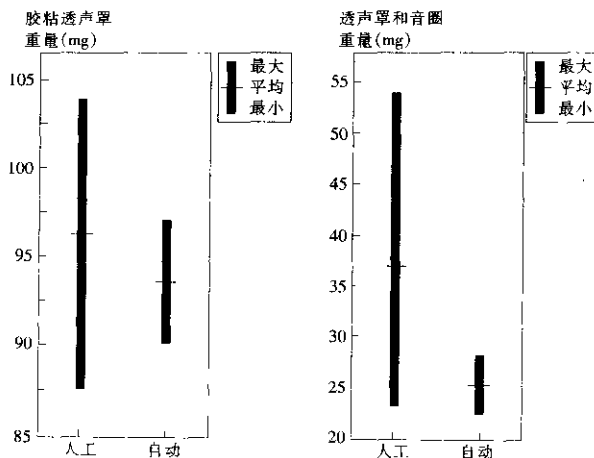
DOE 试验 6: 线圈和透声罩胶重量的多变量研究 (未给出)

一个相似的多变量研究是在音圈和透声罩的胶重量和总重量 (绿 Y) 的人工应用基础上完成的。结果与试验 5 相同,透声罩到透声罩的偏差再次成为红 X 族,分布于胶重量范围的 40%。

DOE 试验 7: B vs. C——胶的自动应用与人工应用

进行 B vs. C 试验 (重叠方法), 以确定对于透声罩以及透声罩/音圈来说胶的自动化应用是否会比人工应用导致更均匀的胶重量。绿 Y 是胶重量的分布范围或者数量。图 23—7 说明透声罩胶重量的

分布降低了 3 : 1 (或者 C_p 值的改进为 3 : 1), 透声罩和音圈胶重量分布降低了 6 : 1 (或者 C_p 值的改进为 6 : 1)。



在自动化喷胶的情况下, 平均值可以通过调整胶管喷嘴的空气压力来进行修改, 偏差保持大致相同。

自动化/半自动化更好, 因为 σ 偏差更小。

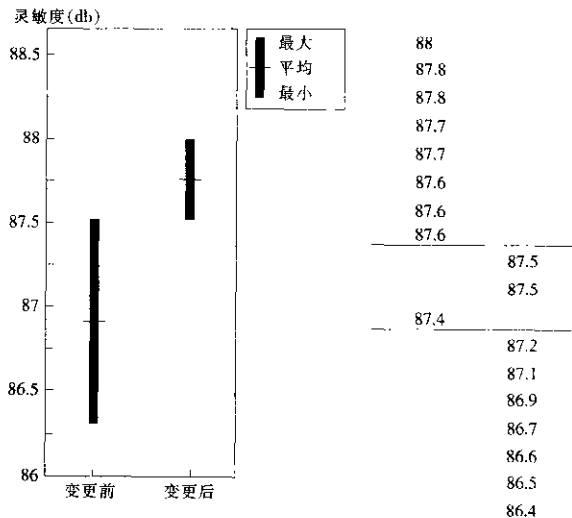
图 23—7 B vs. C 试验: 胶的自动应用与人工应用

DOE 试验 8: 音圈在空气间隙中的位置

我们发现音圈在空气间隙中的位置 (粉红 X) 并不在空气间隙的中间。把该参数调整到中心位置进一步提高了扬声器灵敏度, 并且通过 B vs. C 试验得到了进一步的证实。

DOE 试验 9：磁化部件的改变 (B vs. C 试验)

第三个 (浅粉红 X) 因子是磁感应。全析因表明一个更高的数字是必要的。修改磁化部件, 并且对 9 个老的 (C) 和 9 个新的 (B) 部件进行 B vs. C 试验。绿 Y 是灵敏度 (db)。图 23-8 说明:



变更后: 更高的感应以及具有更小偏差的灵敏度。

图 23-8 B vs. C 试验: 磁化部件变更前与变更后

α 风险

- B 比 C 更好。
- 图基检验终结计数为 14。
- 这意味着改进的置信水平为 99% 以上。

β 风险

- $\bar{X}_B$ 是 87.75, $\bar{X}_C$ 是 86.85。
- $\bar{X}_B - \bar{X}_C$ 是 0.90。
- 对于 90% 的置信水平, 当 $\sigma_B = \sigma_C$ 时, K 为 3.7。
- 因此, 最小差幅 $\bar{X}_B - \bar{X}_C = K\sigma_C = 3.7K_C$ 。
- σ_C 估算值为 1/6, 因此, 最小差幅 = 3.7/6 或 0.61, 该值小于实际差幅 0.9。
- 因此, 同样以 90% 的置信水平, B 足够好于 C。

DOE 试验 10: 最终结果

通过三种变更 (胶的自动化应用、音圈居中、更高的磁感应部件), 最终的灵敏度超过了客户最大技术规格 89.6db。

结论

客户很乐于接受更高的灵敏度, 这确实推动了工艺水平的发展。同样, 客户对于较低的分布 (C_p 值从 0.6 提高到 3) 也非常满意。在 7 家生产该产品的供应商中, 客户将它的供应商缩减为仅两家, 并以该公司作为主要的供应商。该公司不仅每年节省 10 万美元, 而且, 该产品的业务量也增加了 4.5 倍!

第 10 篇

从课堂教育到实践中学习



第24章

实践出真知

► 实践、实践、再实践

我们经常讲述这样一段故事，一群外科医生到摩托罗拉公司拜访我们，以便学习我们著名的“六西格玛过程”。在演示时所讨论的技术之一就是我们在基准方面取得的成功：在世界范围内寻找培训、职责和技术最好的公司，确定它们成功的因子，并在你自己的公司中引入或者采纳这些因子。

外科医生认为重要的是缩短手术周期时间。他们参观了一套不同寻常的基准套件——印第安纳波利斯 500 汽车拉力赛的维修班组。在比赛中，赛车的 4 个轮胎要在 12 秒内更换完毕。除了人力和设备之外，维修

班组的成功秘诀就是：实践、实践、再实践。医生们返回到医院，并观察他们的冗长而又复杂的手术过程中所涉及到的每一个时间环节。经过反复试验，他们可以把手术周期时间缩短一半，将每天的手术量翻倍，却丝毫不会降低质量或者有损健康。

在奥林匹克运动中，在体育冠军中，在“精益生产”中，成功的关键因子就是勤奋、培训和持之以恒的实践。在 DOE 中，这同样也是成功的关键。

- 揭示 DOE 是成功的必要条件，但不是充分条件。
- 阅读本书可以先行一步，但只是一个开端。
- 尽管同训练有素的专业人员以及外部顾问一起工作是学习 DOE 的捷径，但是，只有通过实践、实践、再实践才能真正学会并领悟 DOE 的真谛。

车间地板上堆满了成百上千的 DOE 机会。但是，在激发起 DOE 的全部工作动力之前，需要一步一个脚印地建立起一种有关 DOE 的解决问题的文化。

为解决问题的文化注入活力

IBM 的前任执行官和得克萨斯大学的荣誉退休教授达里尔·皮索尔 (Darryl Piersol)²² 创立了一种管理复杂更改的方法 (见图 24—1)。该方法有下述 5 个要素：眼光、技巧、激励、资源和行动计划。要实现持久的工业文化转变，5 个要素必须全部存在。应用于 DOE，图 24—1 表明：

- **没有眼光**，结果可能会因为不同的 DOE 开端引向不同的方向而引起诸多混乱。只有管理部门才能想像那种眼光；正如世界工业的前任官员之一工业巨头之一鲍勃·盖尔文经常指出的那样，“把眼光移向以前从未研究过的方向上”；并且鼓励人们具有这种眼光。

- **没有技巧**，人们会躁动不安，尽管人们反复地尝试一些方

法，但是，最终均以焦虑和失望而告终。在解决问题方面尤其如此，除了 DOE 之外，所有解决问题的方法都因太复杂、太困难而无法实施。

● **没有行动计划**，就不会有路径图，人们将无法捕捉到能够把目标转化为结果的过程。净输出将是一系列错误的开始。对于 DOE，表 7—1 和图 7—3 提供了其基本路径图。

● **没有资源**（例如，人力、资金，尤其是时间），人们将气馁和缺乏动力。这正是为了“救火”而放弃 DOE 的最经常原因。由于某种原因，尽管管理部门总是数次寻找时间以试图解决同一个长期性的问题，却不肯进行一整套全面的 DOE 研究来永久性地解决这个问题。

● **没有激励**，将仅有有限的变化。人们需要参与，需要事业和成就。团队的最佳协作只有与个人利益挂钩才能实现。



图 24—1 管理复杂的更改

资料来源：Dr. Darryl Piersol, lecture at Motorola Management Institute.

然而，激励并不一定总是金钱方面的。承认一项令人满意的工作是十分重要的。摩托罗拉公司的 6000 个全体客户满意（TCS）竞争团队得到了高层管理人员这样的认可：他们在追求团队目标方面的旺盛斗志几乎很难削弱！

在具备管理眼光、管理资源和管理激励的情况下，本文所提出的 DOE 方法为解决问题小组提供了完美的技巧和所需的行动计划。



表 7—1（一个通用的解决问题框架）和图 7—3（一个解决问题路径图）清晰地描述了这些内容。更具体地说，上层管理人员、设计工程师、质量专业人员、生产管理、供应商和生产线工人均应逐步到位。这些步骤的详细描述如下。

步骤 1：高层管理者的责任和困难

但是，高层管理者可能厌倦“一切好的开端都必须从高层管理者开始”这样的老生常谈。当存在许多有待解决的项目时，首席执行官（CEO）必须巧妙地处理它们。因此，首席执行官（CEO）必须非常仔细地选择需要优先考虑的事情。建议将 DOE 作为需要优先考虑的事情之一，这是一个关键的策略。本书第 1 章的图 1—1 以图解的形式详细描述了 DOE 对业务取得成功的贡献。在产品流程中没有 DOE 影响不到的区域。

首席执行官不必成为使用这些技术的专家，但是，他们必须知道这些技术的存在，必须理解这些技术，必须支持这些技术。更具体地说，首席执行官们、部门经理们和工厂经理们应当：

（1）从高层管理者阶层中寻求一位 DOE 目标的支持者。由于公司的巨额报酬，他可以全身心地埋头钻研 DOE 技术，并努力实现 DOE 技术。这样的管理者将成为“过程所有者”、企业文化变革的催化剂。

（2）强制公司的每一位技术人员参加一个为期一天的 DOE 专题讨论。如果高层管理者能够从头到尾地参加该专题讨论，而不是仅仅在专题讨论开始时宣布神圣的使命的话，那将是一种更好的身体语言信号。

（3）建立一个指导委员会来监督进展。它的任务是：

- 制定质量/可靠性改进目标以及目标节省额。跟踪培训和试验成本，并与节省额进行对比。如果节省额与成本之比未达到 10:1，则所有的努力都滞后了。一些公司（例如，通用电气公司）

近期公开的报告声称利益投资比为 3 : 1, 本书对比的评级是 C-。

- 识别在工厂及现场中的高层质量和成本问题, 并设立需要 DOE 小组处理的项目。

- 建立一个跨学科的 DOE 小组来处理每一个中选项目。世界前质量大师之一——朱兰, 断然指出解决长期性问题和降低不良质量代价的方法就是选择这些项目并指派工作小组——没有其他的方法!

- 必要时, 以重要设备和资金之类的资源来支持 DOE 小组。大多数 DOE 项目不需要耗费大量的资金。尽可能地利用原有设备。(多环境强化应力试验是一个例外, 其中, 高速应力加速需要复杂的设备, 其造价为 50 000 美元~250 000 美元不等。)

- 最重要的是, 以时间来支持 DOE 小组。除了抽调 DOE 项目的关键人员去参加应急处理和例行工作之外, 其他任何事情都不能阻碍 DOE 小组的工作。

- 定期地(至少每月一次, 最好每周一次)评审 DOE 小组的进展。管理部门通常只确定目标并监督结果, 很少评审把目标转化为结果的的所有的重要过程路径。指导委员会应出席 DOE 专题讨论, 并使用本书中每一 DOE 章节后附的 DOE 小组和高层管理提问题指南, 这将有利于管理者和指导委员会提出正确的问题。对预定过程的检查将有助于管理者对 DOE 的掌握, 促进周围的学习气氛, 赢得 DOE 小组成员的尊敬。

(4) 创建 DOE 工作会议, 以此作为 DOE 专题讨论的后续措施。笔者曾把这些工作会议作为促进学习的可靠方法。

- 通常在一天时间的 DOE 工作会议中, 每一位 DOE 小组领导提出的项目应包括:

- ◇问题的描述(简明扼要地描述绿 Y)。

- ◇问题的历史: 缺陷/故障等级; 预算费用; 持续时间。

- ◇有关问题位置(以及可能出现的物理抽样)的流程图。

- ◇问题的可能原因。

- ◇在 DOE 之前所采取的解决问题的方法。



◇选择特定的思路创新 DOE 技术（多变量、部件搜索、成对比较或产品/过程搜索）的原因。

◇到目前为止的 DOE 结果。

其他小组成员也参加每一小组的介绍，以补充观点和见解。

● 接下来，讨论向出席工作会议的其他 DOE 小组开放，尽管它们没有直接参与该项目。这类似于新产品启动时的设计评审，其中，“其他的”开发工程师们扮演了“魔鬼”代言人的角色，质疑所选择的方法。有时，这种质疑退化到通过工程判断来重新处理解决问题的传统方法。但是，最好让技术讨论尘埃落定，以便评价所有的方法。

● 最后，作为外部顾问，笔者评审了每一个小组的介绍，指导、教授并鼓励每一小组继续 DOE 项目的下一阶段。

● 鼓励填写“教训”日志，以便将来的项目和将来的小组可以从错误中吸取教训，而不是将这些疏忽掩盖起来。

● 经过两三次这样的工作会议之后，公司通常会变得非常自信，而 DOE 会成为其生存的一种方式。

(5) 建立一个能够在持续的基础上提供 DOE 连续性的专家队伍。有些公司用“黑带”和“黑带主管”这样的术语来表示这些专家。它们把“黑带”定义为已经通过正式的 DOE 培训，并通过至少三四个成功的项目证明了其能力的人。“黑带主管”指除了完成 DOE 项目之外，还可以教授正式 DOE 课程的人。

(6) 把整个工厂从解决问题中释放出来。尽管“黑带”的概念是合理的，但是，无须停止开明的管理。一个公司只有有限数量的技术人员，他们已经由于新的设计、新的过程、全面的生产维护、应急处理、供应商的不连续性以及文件工作等而不堪重负。

为什么不把全部劳动力从解决问题当中释放出来呢。如果 DOE 技术难以掌握，那么，让直接劳动者和蓝领工人使用这些技术即使是可行的，也的确是一项非常艰巨的任务。然而，本书详细阐述的 DOE 技术是非常容易的。在某种程度上讲，指导生产线工



人使用 DOE 要比教他们使用控制图更容易些。成功地在几家公司的直接劳动者阶层实施了 DOE 后，笔者确信生产线工人在经过管理部门的适当培训、支持、关心和鼓励后，是乐于接受 DOE 的。结尾的案例研究说明了把 DOE 延伸至直接劳动者会给公司带来的巨大利益。

(7) 通过下述方法度量 DOE 进展：

- 每月培训的技术人员数量。
- 每月培训的直接劳动者数量。
- 每月培训的伙伴供应商人员数量。
- 在以上三组人员中，每组承担的项目数量。
- 在以下方面给公司带来的资金节省：

◇质量改进：不良质量代价降低；成品率提高。

◇可靠性改进：降低保修费用、产品召回费用以及产品责任费用。

◇成本降低：通过工厂总有效率 (FOE) 的提高来降低产品成本、供应商成本和全面生产维护 (TMP) 成本。

◇周期时间缩短：进料、生产运行以及成品。

◇库存量减少（提高库存量的周转）。

◇节省额与投资的比值（最小：5 : 1；最大：超过 100 : 1；平均：20 : 1）。

步骤 2：设计工程师的转变

如果转变必须从高层管理者开始，那么，系统地降低“偏差”这一工业弊病就必须从设计工程师开始。实现这种转变的工具就是 DOE。但是，在许多方面，宣传 DOE 时最艰苦的工作就是转变那些保守的工程师们，这远比应付那些不得不忍受工程师们的产品问题且易于接受 DOE 的生产技师以及工厂生产线工人更加困难。在设计中具体的 DOE 培训包括：

- 质量功能展开和大量听取“客户意见”。



- 把客户的“意见”转变为实际技术规格和实际公差，而不是进行“空洞的分析”。

- 设计目标值，以得到最小为 2.0 的 C_{pk} 。

- 使用强大的变量搜索工具把重要的变量和不重要的变量分开。

- 严格控制重要变量，以得到 2.0 或者更高的 C_{pk} 。

- 放宽不重要变量的公差，以降低成本。

- 在设计中通过多环境强化应力试验强制失效，以强化可靠性改进。

- 认识到现场失效是生产车间中生产线失效的一个相关函数，生产车间已经变成了解和纠正现场失效的最大和最早的维修中心。

- 避免试图猜测问题的原因，并学会“与零部件对话”。

步骤 3：质量管理——从警察到教练

质量工作者的角色正在从生产车间的吹哨员和永不疲劳的人转变为教练、顾问和教师的角色。但是，这种转变是非常缓慢的。除了成为公司内部的客户拥护者之外，如果质量工作者要想对其设计和生产同事有用的话，他必须做到如下几点：

- 为了保住饭碗，成为 DOE 方面的真正专家。不掌握 DOE 解决问题工具，就不会对设计人员、生产人员或者生产线工人有多大的帮助。不能提供多大的帮助，就会很快失去这些生产线小组的尊重，并被视为官僚主义的无能。

- 即使不能成为专家，也应了解 21 世纪的 10 个强有力的工具。大多数质量工作者甚至只了解一两种工具，甚至很少使用这些工具。

- 帮助收集、分析和降低不良质量的成本（不良质量的成本是宏观性能的一个单一度量），并有能力提供这种帮助。

- 理解质量与周期时间之间的紧密联系，并了解这两门学科之间的关系。



步骤 4：生产管理

生产管理必须从下述一组原则开始：

- 生产线工人将作为一个高度积极的个体完成工作，除非受到不良管理的抑制以专横的态度或者被迫在充满顾虑的氛围中工作。

- 只要给予培训、支持和鼓励，生产线工人的头脑与管理人员的头脑一样富有想像力。

- 检查和检验不会增加产品的价值，因而应大大减少。

要成功地实施 DOE，生产管理必须：

- 坚持进入生产的新产品的每一个重要的质量参数都应达到 2.0 的 C_{pk} ，这样可以大大地减少生产中的“救火”质量问题。这可以通过预先控制的规则 2 快速而又容易地确定（5 个连续的部件必须落在预先控制线内，即落在绿区中）。

- 以铁的纪律建立正向控制，以监督过程的稳定性。

- 以铁的纪律建立过程确认，以控制所有的外部质量因子。

- 使用预先控制而不是控制图来维持降低的偏差和 C_{pk} 控制。

- 使用 B vs. C 这种快捷有效的 DOE 研究来验证每一工程更改的有效性。

- 用不良质量代价、成品率、每部件的缺陷数、 C_{dp} 、 C_{pk} 和周期时间（不大于直接劳动时间 2 倍的目标值，亦可称为理论周期时间）来监督进展。

步骤 5：供应商

在质量问题中，供应商的问题仅次于设计问题，占据了较大的百分比。一些公司采用的传统方法是强迫供应商符合质量要求或者寻求一家新的供应商，这是一种独裁的、遥控的方法，作用有限。

需要的是在客户公司和其关键供应商之间建立一种伙伴关系。在这种伙伴关系中，一个主要的原则是，客户公司必须帮助供应商



改进其质量、费用和周期时间。为什么必须这么做呢？因为这是帮助公司自己的最佳途径。作为对这种帮助的回报，供应商必须每年都持续地降低费用。尽管不应强迫供应商按照公司的要求进行，但是，通过帮助供应商实际增加了它们的利润。只有到那时，这种关系才是双赢的伙伴关系。

最有效地帮助一家供应商的方法之一是，通过 DOE。客户的商品小组必须实际生活在供应商的有关机构，培训和辅导它们，并解决它们的长期性质量问题。成功的标志是大大减少供应商的不良质量成本，使双方都从中受益。笔者在与其客户公司和它们的伙伴供应商的共同工作中，已有成功的记录表明，在费用方面已取得每年 5%~10% 的降低，大大超过了物价上涨的幅度。

步骤 6：利用“讨厌的”工作——谦逊地引导

但是，如果管理、设计、质量和生产部门在支持方面比较冷淡或者根本没有的话，情况又会如何呢？我们已经主张我们的客户成立一个（专门致力于 DOE 的）小而谨慎的“讨厌的工作”小组，以毫不夸张的方式尝试 2~3 个 DOE 项目、度量结果，然后向高层管理者介绍其利益。没有一个有理智的管理者会拒绝经过验证的节省！

步骤 7：一线工人的贡献——未开发的金矿

最后的，但不是微不足道的一步是，在解决问题时囊括全体直接劳动者。这种发挥劳动者主观能动性的基层结构已经存在，一些公司日益致力于像持续改善、小组改进活动（SGIA）和玻璃墙管理之类的运动。但是，我们需要超出这些基层结构。我们需要开发工人技术（见图 24-1）。我们需要为工人们提供简单而强大的工具，这些工具不是因果图、头脑风暴和控制图之类的基本工具，而是行之有效的工具，其中 DOE 是最容易且最卓有成效的工具。



下面给出了一个表明一线工人是一座金矿的案例研究。

案例研究：比利时的登德蒙德

几年前，笔者应邀到位于比利时的登德蒙德的飞利浦电子公司介绍 DOE。此前作者已在新加坡飞利浦公司取得了 DOE 的成功。登德蒙德工厂生产针对汽车工业和高保真客户的扬声器。

工厂当时正在亏损，每年出现 15% 的赤字。工人们感到很沮丧，工厂已经从 800 名工人中临时解雇了 300 人。新的工厂管理层必须尽快停止这种失血。于是工厂开展了一项由生产线工人共同参加的、称为“消除浪费”的运动。同时，工厂还认识到需要将 DOE 作为启动这项运动的火花。

开始时，我为工厂的 40 名技术人员举行了为期一天的专题讨论，并在第二天举行的工作会议上扼要地描绘了解决问题的 3 个 DOE 项目。在我第二次访问该工厂时，这 3 个 DOE 项目已经成功地完成，并且已经开始了 12 个新的 DOE 项目。在我第三次访问该厂时，DOE 项目已经增加到 30 个。在技术领域，该工厂正向着成功之路迈进。

根据我同其他公司一道工作的经验，我建议该工厂的经理为生产线工人举办一期 DOE 的专题讨论，将其作为解决整个工厂问题的起点。他有点为难，并提出了异议：“他们不能说英语，而您又不能说佛兰德语（Flemish）。”但他还是从生产线工人中选择了 12 名略懂英语的女工参加了我主持的为期一天的 DOE 专题讨论。这些女工都做得很好，解决了一些简单的小组练习，这些练习与我为技术小组提供的练习类似。这些女工为她们取得的成绩感到十分兴奋，她们希望全体职工都得到更为详细的培训。因而，工厂选择了一名能干的经理作为 DOE 目标的支持者来培训生产线工人。

后来，DOE 同“消除浪费”的活动结合了起来。组建了一些直接劳动者小组。他们选择自己的小组领导，检查员的角色变成了促进者。每天这些小组有半个小时的时间，既可以在班前也可以在班后解决前一天的问题，如果可能的话这些问题由他们自己解决，



必要时可在工程师们的帮助下解决。从生产率的角度来看，并没有出现半个小时的损失问题，因为大家都一致认为节省的开支大大超过了在时间上的投入。该工厂的经理在上午 6:00 的班前或晚上的班后，尽可能多地亲自参加这些小组的会议。

这为推动 DOE 打下了良好的发展势头。在 400 名生产线工人中，已有 300 名接受了 DOE 培训，且只强调下述 3 种技术：多变量、部件搜索和成对比较。一线工人小组能够依靠自己的能力解决他们遇到的问题中的 70%。实际上，他们对工程师们开玩笑地说：“去干其他的事吧，我们不需要你们！”

工厂经理和他的直接报告每隔两周要评审一次这些 DOE 项目（既包括技术人员的 DOE 项目，也包括一线工人的 DOE 项目）。起初，绘制了有关培训人数、DOE 项目数和以比利时法郎为单位的节省值的统计图表。已记录的质量改进从 2:1 直达 100:1，中值为 15:1，平均值为 28:1。然而，随着 DOE 逐渐成为一种生活方式，再也没有必要证明采用 DOE 是否正当了。

其底线如何？在不到两年之后，按照销售的百分比计算，该工厂从 15% 的赤字转为 4% 的盈利。

快乐的工作环境

杰出的质量大师戴明博士经常强调，建立“快乐的工作环境”是管理的职责。当工人将其一半以上的清醒时间花费在工作之上时，工作应当是有趣的。

取而代之的是，我们通常看到的是把工人当作机器人，日复一日、周复一周、月复一月地从事乏味的、重复的、令人厌烦和沉闷的工作。然后我们不安地说，普通工人已经失去了工作自豪感！我们这些坐在空调办公室中的技术人员中有谁会同生产线上的工人交换位置吗？

因此，当今工业的挑战是：



● 消除那些固守在自己的小圈子里的工人的思想恐惧，他们害怕做出会使自己受到严厉批评的事情。

● 根据他们的需要，培训并辅导他们使用简单而强大的工具。

● 关注他们——霍索恩效应*。倾听他们的呼声，支持他们的工作，鼓励他们的斗志。让他们了解更多与他们的工作相关的事情，而不仅限于了解他们所做的工作。

● 通过下述途径重新设计或丰富他们的工作：

◇ 合并任务；

◇ 形成自然工作部件和小组；

◇ 建立客户关系（下一道工作是作为客户）；

◇ 通过逐步增加他们的操作管理职责，建立丰富的纵向作业。

简言之，让我们改变我们的工作环境，以使我们能够真正地由讨厌工作转变为热爱工作。

* 指工人、学生等因受到研究人员的关注而增加产量或提高成绩。——译者注

参考文献

1. Schaffer and Thomson, "Successful Change Programs Begin with Results", *Harvard Business Review* 5/6, 1992.
2. Jay Matthews and Peter Katel, "The Cost of Quality", *Newsweek*, September 7, 1992.
3. Gilbert Fuchsberg, "Quality Programs Show Shoddy Results", *Wall Street Journal*, May 14, 1992.
4. A. Blanton Goddfrey, "Strategic Quality Management. Part 1", *Quality*, March 1990.
5. Keki R. Bhote, *Plan for Maximum Profit : The 12 Critical Success Factors That Guarantee Increased Profits from Total Quality*, Zurich: Strategic Directions, 1995.
6. Keki R. Bhote, *The Ultimate Six Sigma ; The Pursuit of Business Excellence*, New York: American Management Association, forthcoming.
7. Keki R. Bhote, "A Powerful New Tool Kit for the 21st Century", *National Productivity Review*, Autumn 1997.
8. Bob Kine, *Better Designs in Half the Time*, Naiman Press, 1987.
9. Seiichi Nakajima, *T. P. M. Development Program ; Implementing Total Productive Maintenance*, Cambridge, Mass. : Productivity Press, 1989.
10. Michael J. Spendolini, *The Benchmarking Book*, New York: American Management Association, 1992.
11. Nikkam Kogyo Shinbun, *Poka-Yoke ; Improving Product Quality by Preventing Defects*, Cambridge, Mass. : Productivity Press, 1988.
12. Keki R. Bhote, *Next Operation as Customer ; How to*



Improve Quality, Cost and Cycle Time in Support Services, New York; American Management Association, 1991.

13. James Harrington, *Business Process Improvement*, McGraw-Hill.

14. Keki R. Bhote, *Strategic Supply Management : A Blueprint for Revitalizing the Manufacturer-Supplier Partnership*, New York; American Management Association, 1989.

15. Lawrence D. Miles, *Techniques of Value Analysis and Engineering*, McGraw-Hill, 1972.

16. J. Duncan Acheson, *Quality Control and Industrial Statistics*, 5th ed. Homewood, Ill. : Richard D. Irwin Inc. , 1986. (See p.1007, Table D3)

17. Dorian Shainin and Peter Shainin, "Statistical Process Control", in *Quality Control Handbook*, ed. J. M. Juran and F. M. Gryna, section 24, McGraw-Hill, 1988.

18. Keki R. Bhote, "World Class Reliability through Multiple Environment Over Stress Tests (MEOST)", Asia Reliability Symposium, Hong Kong, February 1999.

19. Keki R. Bhote and James McLinn, "MEOST: An Approach to Quickly Evaluate Any New Product", Ada, Mich. : Amway Corp. , March 8, 1996.

20. Stephen A. Smithson, "Effectiveness and Economics; Yardsticks for ESS Decisions", Proceedings of the Institute of Environmental Sciences, 1990.

21. Dave Steinberg, *Vibration Analysis of Electronic Equipment*, Wiley, 1973.

22. Darryl Piersol, *Managing Complex Change*, Endicott, N. Y. ; IBM.

WORLD-CLASS QUALITY; *Using Design of Experiments to Make It Happen*, 2/E. Copyright © 2000 Keki R. Bhote & Adi K. Bhote. Published by AMACOM, a division of the American Management Association, International, New York, All rights reserved.

图书在版编目 (CIP) 数据

世界级质量管理工具: 第2版/ (美) 博特, (美) 博特著; 遇今等译.
北京: 中国人民大学出版社, 2004
(工商管理经典译丛·产品管理系列)

ISBN 7-300-05661-X/F·1780

I. 世…

II. ①博…②博…③遇…

III. 质量管理

IV. F273.2

中国版本图书馆 CIP 数据核字 (2004) 第 056950 号

工商管理经典译丛·产品管理系列

世界级质量管理工具 (第2版)

[美] 凯克·博特 阿迪·博特 著

遇今 石柱 译

遇今 审校

出版发行	中国人民大学出版社	
社 址	北京中关村大街31号	邮政编码 100080
电 话	010-62511242 (总编室)	010-62511239 (出版部)
	010-82501766 (邮购部)	010-62514148 (门市部)
	010-62515195 (发行公司)	010-62515275 (盗版举报)
网 址	http://www.crup.com.cn	
	http://www.itrnet.com (人大教研网)	
经 销	新华书店	
印 刷	北京东方圣雅印刷有限公司	
开 本	890×1240 毫米 1/32	版 次 2004年9月第1版
印 张	16.875 插页2	印 次 2004年9月第1次印刷
字 数	452 000	定 价 35.00 元

版权所有 侵权必究 印装差错 负责调换