测量系统分析

Measurement Systems Analysis

MSA

Third Edition



测量系统分析

参考手册

第三版

第一版，1990 年 10 月。第二版，1995 年 2 月；二次印刷，1998 年 6 月。第三版，2002 年 3 月

版权 © 1990，© 1995，© 2000 戴姆勒克莱斯勒公司、福特汽车公司、通用汽车公司

中文版版权 © 2003，品上公司

QS-9000 及相关手册中文版修订工作小组

翻译：　蒋红卫

编译：

方 俭	品士公司-台北
任德昌	品士公司-台北
钟樱珍	品士公司-台北
马至聪	品士公司-北京
孙玲玲	品士公司-北京
刘 菁	品士公司-北京
俞 琼	品士公司-北京
唐 詠	品士公司-上海
桂诗海	品士公司-上海
李锦幽	品士公司-重庆
余丽红	品士公司-重庆
胡永红	品士公司-重庆
蒋玉峰	品士公司-重庆

中文繁体版台湾地区总经销

品士股份有限公司
地址：台北市 111 天母北路 87 巷 12 号 1 楼
电话：+886 2 2873 2327，传真：+886 2 2876 3122
Email: plexus@plexus.org.tw
Website: http://www.plexus.org.tw

中文简体版大陆地区总经销

北京品士质量管理顾问有限公司
地址：北京市海淀区学院路蓟门饭店 3508 室
电话：+86 10 6203 5170，传真：+86 10 6200 5456
Email: info@plexus.com.cn
Website: http://www.plexus.com.cn
AIAG 服务专线：+1 248 358 3003

前言

本参考手册是在美国质量协会 (American Society for Quality, ASQ) 及汽车工业行动集团(Automotive Industry Action Group, AIAG)赞助下，由戴姆勒克莱斯勒、福特和通用汽车公司供方质量要求特别工作组认可的测量系统分析 (MSA) 工作小组所编写。负责第三版的工作小组成员是 David Banham（戴姆勒克莱斯勒）、Micheal Down（通用汽车公司）、Peter Cvetkoeski（福特汽车公司）、Gregoty Gruska (Third Generation, Inc.)、Tripp Martin (Federal Morgul)与 Steve Stahley (SRS Technical Services)。

以往，戴姆勒克莱斯勒、福特和通用汽车公司各有其确保供方一致性的指导文件和格式。但这些指导文件的差异导致了对供方资源的额外需求。为了改善这种状况，特别工作组被特许将戴姆勒克莱斯勒、福特和通用汽车公司所使用的参考手册、程序、报告格式以及技术术语进行标准化处理。

因此，戴姆勒克莱斯勒、福特和通用汽车公司同意在 1990 年编写并通过 AIAG 发行了 MSA 手册。第一版发行后，供方反映良好，并根据实际应用经验，提出了一些修改建议，这些建议都已纳入第二版和这次的第三版。本手册由戴姆勒克莱斯勒、福特和通用汽车公司批准并承认，是 QS-9000 的一份补充参考文件。

本手册可看作是对测量系统分析的一种介绍，*它并不试图去限制某特定的过程或商品所适用的分析方法的发展*。当这些指导文件意在涵盖测量系统通常发生的情况时，其中可能还有一些问题没有考虑到；这些问题应该由你的顾客的供方质量保证部门(SQA)提出指示。如果不知道如何与有关的 SQA 部门联系，在顾客采购部门的采购人员可以提供帮助。

MSA 特别工作组衷心感谢：戴姆勒克莱斯勒汽车公司副总裁 Tom Sidlik、福特汽车公司副总裁 Carlos Mazzorin，以及通用汽车公司副总裁 Bo Andersson 的领导和委任；感谢 AIAG 在编写、出版和发行本手册中提供的帮助；感谢特别工作组负责人 Hank Gryn（戴姆勒克莱斯勒）、Russ Hopkins（福特汽车公司）、Joe Bransky（通用汽车公司）的指导；以及 ASQ 代表 Jackie Parkhurst（通用汽车公司）和美国测试与材料协会(American Society for Testing and Materials, ASTM International) 的协助。因此，这本手册才得以编写出来，以满足汽车工业界的特殊要求。

本手册的版权由戴姆勒克莱斯勒公司、福特汽车公司、通用汽车公司保留所有权，2002 年。需要更多的手册可向 AIAG 订购；和 / 或通过 AIAG 联系电话：248-358-3570，取得在供方组织内使用本手册的部分复制内容的许可。

2002 年 3 月

前言

本参考手册是在美国质量控制协会 (ASQC) 汽车部及汽车工业行动集团 (AIAG)主持下，由克莱斯勒 / 福特 / 通用汽车公司供方质量要求特别工作组认可的测量系统分析(MSA)工作小组编写，负责第二版的工作小组成员是 Rey Daughterty（克莱斯勒），Victor Lowe, Jr.（福特汽车公司）、主席 Michae H. Down（通用汽车公司）、Gregoty Gruska (The Third Generation, Inc.)。

以往，克莱斯勒、福特和通用汽车公司各有其确保供方一致性的指导文件和格式。这些指南的差异导致了对供方资源的额外需求。为了改善这种状况，特别工作组被特许将克莱斯勒、福特和通用汽车公司所使用的参考手册、程序、报告格式以及技术术语进行标准化处理。

因此，克莱斯勒、福特和通用汽车公司同意在 1990 年编写并通过 AIAG 发行 MSA 手册。第一版发行后，供方反映良好，并根据实际应用经验，提出了一些修改建议，这些建议都已纳入第二版。由克莱斯勒、福特和通用汽车公司批准并承认的本手册可由供方在制造过程和满足 QS-9000 要求中用来实现 MSA 技术。

本手册可看作是对测量系统分析的一种介绍，它并不意图去限制某特定的过程或商品所适用的分析方法的发展。当这些指导文件意在涵盖测量系统通常发生的情况时，其中可能还有一些问题没有考虑到；这些问题应该由你的顾客的供方质量保证部门 (SQA) 提出指示。如果不知道如何与有关的 SQA 部门联系，在顾客采购部的采购人员可以提供帮助。

特别工作组衷心感谢：克莱斯勒汽车公司副总裁 Thomas T. Stallkamp、福特汽车公司副总裁 Norman F. Ehlers，以及通用汽车公司副总裁 Harold R. Kurner 的领导和委任；感谢 AIAG 在编写、出版、分发手册中提供的帮助；感谢特别工作组负责人 Russell Jacobs（克莱斯勒）、Stephen Welsh（福特）、Dan Reid（通用）的指导，以及 ASQC 给与的关心和帮助。因此，这本手册才得以编写出来，以满足汽车工业界的特殊要求。

本手册的版权由 AIAG 保留所有权，1994 年。需要更多的手册可向 AIAG 订购；和 / 或通过 AIAG 联系电话：248-358-3570，取得在供方组织内使用本手册的部分复制内容的许可。

1995 年 2 月

测量系统类型	MSA 方法	章节
基本的计量型	极差法、均值与极差法、方差分析 (ANOVA) 法、偏倚、线性、控制图法	III
基本的计数型	信号探测法 假设试验分析法	III
不可重复的 （如：破坏试验）	控制图法	IV
复杂的计量型	极差法、均值与极差法、方差分析 (ANOVA) 法、偏倚、线性、控制图法	III、IV
复合的系统、量具或试验标准	控制图法、方差分析 (ANOVA) 法、回归分析法	III、IV
持续的过程	控制图法	III
其它	替代的方法	V
其它	白皮书 — 可上网查询，网址： http://www.aiag.org/publications/quality/msa3.html	

注：关于 GRR 标准差的应用

以往为方便起见，一直用 99%的分布宽度 (spread) 来代表测量误差的“全”宽度，也就是乘以一个系数 5.15（用 σ_{GRR} 乘以 5.15 来代表 99%的全部分布宽度）。

99.73%分布宽度是由乘数 6 代表，即±3σ，其代表一“常态”曲线的全部分布宽度。

如果读者想把总测量变异的范围或分布宽度提高到 99.73%，在计算时请用乘数 6 代替 5.15。

应注意，使用哪一个相乘系数对方程序和结果的正确计算是极关紧要的。如果是用于测量系统变异和公差之间的比较，它尤为重要。

目录

第一章 — 测量系统总指南 1
第一章 — 第 A 节 3
引言、目的及术语 3
测量数据的质量 3
目的 4
术语 4
术语汇总 5
真值 10
第一章 — 第 B 节 11
测量过程 11
测量系统的统计特性 12
变差的来源 13
测量系统变差的影响 16
对决策的影响 16
对产品决策的影响 17
对过程决策的影响 18
新过程的接受 20
过程作业准备 / 控制（漏斗试验） 21
第一章 — 第 C 节 23
测量的策略和计划 23
复杂程度 23
识别测量过程的目的 24
测量生命周期 24
对测量过程设计的选用准则 24
研究不同的测量过程方法 26
概念和建议方案的开发和设计 26
第一章 — 第 D 节 27
测量资源的开发 27
数据的协调 28
先决条件及假设 29
量具资源的选择过程 29
详细的工程概念 29
预防性维护的考虑 30
规范 30
评估报价单 31
可交付的文件 32
在供方处的鉴定 33
运输 34
在顾客处的鉴定 34
文件交付 34
测量系统开发检查清单的建议要素 36
第一章 — 第 E 节 39
测量问题 39
测量系统变差的类型 40
变差的定义和潜在的来源 40
测量过程变差 48
位置变差 48
宽度变差 52
测量系统变差 56
评论 59
第一章 — 第 F 节 61
测量不确定度 61
总则 61

测量不确定度和 MSA 62
测量的可追溯性 62
ISO 对表达测量不确定度的指南 63
第一章 — 第 G 节 65
测量问题分析 65
第二章 — 用于评估测量系统的基本概念 67
第二章 — 第 A 节 69
背景 69
第二章 — 第 B 节 71
选择 / 开发试验程序 71
第二章 — 第 C 节 73
测量系统研究的准备 73
第二章 — 第 D 节 77
结果分析 77
第三章 — 对简单测量系统的推荐实施方法 79
第三章 — 第 A 节 81
试验程序范例 81
第三章 — 第 B 节 83
计量型测量系统研究 — 指南 83
用于确定稳定性的指南 83
确定偏倚的指南 — 独立样件法 85
确定偏倚的指南 — 控制图法 88
确定线性的指南 92
确定重复性和再现性的指南 97
极差法 97
平均值与极差法 99
平均值图 102
极差图 104
链图 105
散点图 106
振荡图 107
误差图 108
正常化直方图 109
平均值 — X-Y 图 110
X-Y 比较图 110
数值计算 111
结果分析 — 数值法 115
方差分析法 (ANOVA) 117
随机化和统计的独立性 117
第三章 — 第 C 节 125
计数型测量系统的研究 125
风险分析法 125
分析法 135
第四章 — 实施复杂的测量系统 141
第四章 — 第 A 节 143
实施复杂的或不可重复的测量系统 143
第四章 — 第 B 节 145
稳定性研究 145
S1：单一零件，每次实施单一测量 145
S2：n ≥ 3 个零件，每次每个零件实施单一测量 146
S3：从稳定过程中的大量样本 147
S4：分割样本（一般），每次采用单一样本 148
S5：试验台 148
第四章 — 第 C 节 151
变异研究 151
V1：标准 GRR 研究 151

V2：用 p≥ 2 台仪器的多个读值 151
V3：分割的样本 (m=2) 152
V4：分割样本（一般） 153
V5：与 V1 相同，使用稳定的零件 153
V6：时间序列分析 154
V7：线性分析 154
V8：时间与特性（属性）退化的关系 155
V9：V2 用于同时存在的多个读值，且 p≥ 3 台仪器 155
第五章 — 其它测量概念 **157**
第五章 — 第 A 节 **159**
识别过大的零件内部变差的影响 159
第五章 — 第 B 节 **161**
平均值和极差法 — 额外的处理 161
第五章 — 第 C 节 **169**
量具性能曲线 169
第五章 — 第 D 节 **175**
通过多次读值减少变差 175
第五章 — 第 E 节 **177**
聚集标准差法计算 GRR 177
附录 **185**
附录 A **187**
方差分析的概念 187
附录 B **191**
GRR 对能力指数 *Cp* 的影响 191
公式 191
分析 191
图形分析 193
附录 C **195**
d2 表 195
附录 D **197**
量具 R 研究 197
附录 E **199**
用误差修正术语来代替 *PV* 计算 199
附录 F **201**
P.I.S.M.O.E.A 误差模型 201
术语 **205**
参考文献 **211**
范例表格 **215**
索引 **219**
MSA 手册使用者反馈过程 **225**

表格目录

序号	标题	页码
1	控制理论和驱使的关注点	17
2	偏倚研究用数据	87
3	偏倚研究 — 偏倚研究的分析	88
4	偏倚研究 — 对偏倚的稳定性研究的分析	90
5	线性研究数据	94
6	线性研究 — 中途的结果	95
7	量具研究（极差法）	98
8	*ANOVA* 表格	120
9	*ANOVA* 分析法 —变差%及贡献率	121
10	*ANOVA* 分析法和平均值和极差法的比较	122
11	*ANOVA* 方法的 GRR 报告	122
12	计数型研究数据表	127
13	测量系统举例	143
14	基于不同类型的测量系统使用的方法	144
15	聚集的标准差分析数据组	181
16	变差元素的估计值	187
17	5.15σ 分布宽度	188
18	方差分析	189
19	*ANOVA* 结果的列表(零件 a 和 b)	190
20	观测的与实际的 Cp 比较	193

插图目录

序号	标题	页码
1	长度测量的可追溯性链举例	10
2	测量系统变差的因果图	15
3	不同标准之间的关系	42
4	分辨力	44
5	过程分布的区别分类数对控制与分析活动的影响	45
6	过程控制图	47
7	测量过程变差的特性	48
8	偏倚和重复性之间的关系	60
9	用于稳定性分析的控制图	84
10	偏倚研究 — 偏倚研究的直方图	87
11	线性研究 — 图示分析法	95
12	量具重复性和再现性数据收集表	101
13	平均值图 — “重迭画出”	103
14	平均值图 — “非重迭画出”	103
15	极差图 — “重迭画出”	104
16	极差图 — “非重迭画出”	105
17	零件的链图	105
18	散点图	106
19	振荡图	107
20	误差图	108
21	正常化直方图	109
22	平均值 X-Y 图	110
23	X-Y 比较图	111
24	完成的 GR&R 数据收集表	113
25	量具重复性和再现性报告	114
26	相互作用图	119
27	余数图	119
28	过程范例	126
29	与测量系统有关的灰色区域	126
30	*Pp*=*Ppk*=1.33 的过程范例	133
31	在常态概率图纸上描绘的计数型量具性能曲线	139
32	计数型量具性能曲线	140
33	(33a & b) 测量评价控制图	164 & 165
34	(34a & b) 评价一测量过程的控制图方法的计算	166 & 167
35	没有误差的量具性能曲线	172
36	量具性能曲线 — 范例	173
37	在常态概率图纸上绘制的量具性能曲线	174
38	(38a. b & c) 聚集的标准差研究的图示分析	180、183、184
39	观测的与实际的 Cp 关系（以过程为基础）	193
40	观测的与实际的 Cp 比较（以公差为基础）	194

致 谢

本手册是许多人多年来负责编写的成果，下面列举的只是那些为本手册的开发贡献了大量的时间和精力的少数人。

ASQ 及 AIAG 奉献了时间和设施为本手册的发行开发提供了帮助。ASQ 汽车部门的代表 Grey Gruska、以及前修订工作小组的主席 John Katona 一直是编写及以往出版本手册的主要贡献者。

本手册第三章的技术部分是在 Berney Flym 的指导和促进下，由雪佛兰 (Chevrolet) 产品保证部的 Katona Mirkhani 首先进行调查并编写的。计量型量具研究是依据 General Electic（1962 年 ASQC 会刊）中 R.W. Traver 所发表的论文，由 Jim McCaslin 校对。

这些概念，被雪佛兰汽车的 Jim McCaslin、Gregory Grucka 和 Tom Brunzell，扩展到计数型数据研究和量具性能曲线中（1976 年 ASQC 会刊）。1978 年 6 月，Bill Wiechec 将这些技术进行了总结和编辑，出版了雪佛兰汽车 *"测量系统分析"* 手册。

在过去的几年里，本手册追加了补充材料。特别是奥斯摩比汽车 (Oldmobile) 的 Sheryl Hansen 和 Ray Benner 编写了方差分析 (ANOVA) 方法和置信区间的文件；在八十年代，雪佛兰汽车的 Larry Marruffo 和 John Lazur 更新了雪佛兰汽车手册；John Lazur 和 Kazem Mirkhani 对该手册提出了新的章节并强化了一些概念，例如：稳定性、线性和 ANOVA；EDS 的 Jothi Shanker 为供方开发人员的准备工作进行更新。其它的新增修改包括：零件内变差的识别与鉴定的观念，以及对统计稳定性作了更全面的描述，通用汽车公司统计评审委员会完成了这两处修改。

最新的改进包括：为符合现行 QS-9000 文件要求而使用了新的格式；更多的说明和范例，使本手册让读者方便阅读；讨论了测量不确定度的概念；增加了原手册编写中没有包括的部分或不存在的内容。这次更新还包括测量系统寿命周期的概念，以及促使测量分析成为常规的过程分析。在本次修订中还包括通用汽车公司动力系统 (GM Powertrain)内部在 1993 年 4 月 28 日所出版的 *"测量过程：计划、使用或改进"* 手册中的一部分内容。

本手册重新编写小组委员会由通用汽车公司的 Mike Down 主持，成员包括：戴姆勒克莱斯勒公司的 David Benham、福特汽车公司的 Peter Cvetkovski、ASQ 汽车部门的代表 Greg Gruska、Federal Mogul 公司的 Tripp Martin、SRS Technical Services 公司的 Steve Stahley。其它来自 Minitab 公司的 Yanling Zuo、ASTM *International* 的 Neil Ullman 和 Rock Valley 大学技术部的 Gordon Skattum 也做出了重要贡献。AIAG 为本手册发行的开发贡献了时间和设施。

最后，代表了通用、戴姆勒克莱斯勒与福特汽车公司的 MSA 工作小组成员，一致同意本文件内容的效力，他们的批准签名如下：

Michael H. Down	David R. Benham	Peter Cvetkovski
通用汽车公司	戴姆勒克莱斯勒公司	福特汽车公司

第一章

测量系统总指南

第一章 — 第A节

引言、目的及术语

引言

测量数据的应用比以前更多更广泛了。例如：现在，是否对制造过程进行调整的决定通常以测量数据为基础，将测量数据或一些从它们所计算出的统计值与过程的统计控制限 (statistical control limits) 进行比较，如果该比较显示过程已超出统计控制，则进行某种调整；否则，该过程将被允许在没有调整的状态下运行。测量数据的另一个用途是确定在两个或多个变数之间是否存在重大的相互关系。例如，如果怀疑一个模塑零件的某一关键尺寸与材料的注塑温度有关，这种可能的相互关系可以通过利用一种称为回归分析 (regression analysis) 的统计程序研究，以比较关键尺寸的测量值与材料注塑温度的测量值。

进行这种举例的互相关系探测研究，被戴明博士 (Dr. W. E. Deming) 称为*分析研究法 (analytic studies)*。通常，分析研究法是不断增加对影响过程的系统原因知识的一种分析研究。分析研究是测量数据使用的重要应用之一，因为应用它们能使得对过程有更好的理解。

使用以数据为基础的程序的最大益处取决于所使用的测量数据的质量。如果数据质量低，程序的益处可能会较低；同样的，如果数据的质量高，利益也可能会较高。

为确保从使用测量数据得到的利益大到足以承担获得这些数据的成本，数据的质量需要特别的注意。

测量数据的质量

数据的质量取决于从处于稳定条件下进行操作的测量系统中，多次测量的统计特性。例如：假设使用某一在稳定条件下操作的测量系统对某一特定特性值进行了几次测量，如果这些测量值均与该特性的参考值 (master value)“接近”，那么，数据的质量被称为“高”；同样，如果部分或所有的测量值与参考值相差“很远”，则称数据的质量很“低”。

通常用来描述测量数据质量的统计特性是某测量系统的*偏倚 (bias)* 及*变差 (variance)*。被称为偏倚的统计特征指的是数据值相对于参考（基准）值的位置。而被称为变差的特性指的是数据的分布宽度 (spread)。

低质量数据最普遍的原因之一是变差太大。一组数据中的变差多是由于测量系统及其环境的*相互作用 (interaction)*造成的。例如：一个用来测量一罐液体容积的测量系统，可能对该测量系统所处的环境中的大气温度较敏感。在这种情况下，数据的变差可能是因为容积的变化造成的，也可能是因为环境温度变化造成的。因此，对测量的数据很难解释，因此，该测量系统不尽理想。

如果相互作用产生的变差过大，那么数据的质量会太低，从而造成测量数据无法利用。例如：具有较大变差的测量系统可能不适合用于分析制造过程，因为测量系统的变差可能掩盖制造过程的变差。管理一个测量系统的许多工作集中在监视和控制其变差，其它的还需要把重点集中在了解测量系统与其环境有什么样的相互作用，以便获得可接受质量的数据。

目的



本手册的目的是为评估测量系统的质量提供指南。尽管这指南一般而言可用于任何测量系统，但主要意图是用于工业界的测量系统。本文件并不意图成为对所有测量系统的一种分析总览；而是主要用于那些对于每个零件的数据可以重复读取的测量系统。许多分析对于其它形式的测量系统也是很有用的，且本手册的确包含了参考意见和建议，但对于更复杂的或不常用的方法本手册没有讨论，建议使用者参考适宜的统计资源。本手册也不涵盖顾客对测量系统分析方法所要求的批准。

术语

如果不建立一套术语来引述共同的统计特征和相关的测量系统要项，那么讨论测量系统的分析可能会造成混淆和误解。本节将用于本手册的术语汇总如下。

在本手册中使用了以下术语：

- 测量 (Measurement) 被定义为“对某具体事物赋予数字（或数值），以表示它们对于特定特性之间的关系”。这定义由 C. Eisenhart (1963) 首次提出。赋予数字的过程被定义为测量过程，而数值的指定被定义为测量值。

- 量具 (Gage) 是指任何用来获得测量的装置。经常是特别用在工厂现场的装置，包括通 / 止规 (go/no go device)。
- 测量系统 (Measurement system) 是对测量单元进行量化或对被测的特性进行评估，其所使用的仪器或量具、标准、操作、方法、夹具、软件、人员、环境及假设的集合；也就是说，用来获得测量结果的整个过程。

由以上这些定义可以将测量过程看成一个制造过程，其产生的输出就是数值（数据）。这样看待一个测量系统是很有用的，因为这样让我们明白已经说明的所有的概念、原理和工具，这在统计过程控制中早已被证实他们的作用。

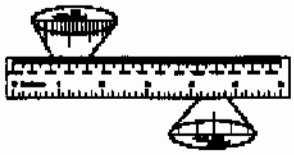
术语汇总[1]

标准 (Standard)

- 用于比较的可接受偏倚
- 接受的准则
- 一已知的值，在不确定度 (uncertainty)的指定范围内，被接受而为一真值 (true value)
- 参考值 (referencr value)

> 标准应该有一个可以操作的定义：该定义在由供方或顾客应用时，将会产生同样的结果，并且在过去、今天、将来都有同样的含义。

基本的设备 (Basic equipment)



- 分辨力 (Discrimination)、可读性 (Readability)、解析度 (Resolution)
  - ✓ 别名：最小可读单位、测量解析度、最小刻度极限、或探测的最小极限
  - ✓ 由设计所确定的固有特征
  - ✓ 一个仪器测量或输出的最小刻度单位
  - ✓ 通常被显示为测量单位
  - ✓ 10 比 1 的比例法则
- 有效解析度 (Effective resolution)
  - ✓ 特定应用条件下， 一个测量系统对过程变差的敏感度。

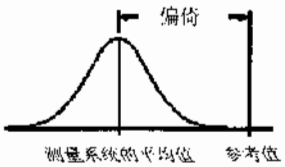
[1] 参见第一章第 E 节术语定义和讨论

- ✓ 可以导致测量有用的输出信号的最小输入
- ✓ 通常被描述为一种测量单元

- 参考值 (Reference value)
  - ✓ 某一个物品的可接受数的值。
  - ✓ 需要一个可操作的定义
  - ✓ 常被用来替代真值使用。

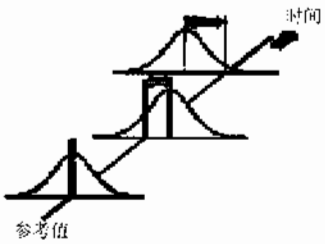
- 真值 ((True value)
  - ✓ 某一物品的真实数值
  - ✓ 不可知且无法知道的

位置变差 (Location variation)



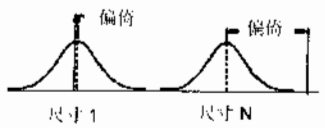
- 准确度 (Accuracy)
  - ✓ 与真值或可接受的参考值"接近"的程度
  - ✓ 在 ASTM 包括了位置及宽度误差的影响

- 偏倚 (Bias)
  - ✓ 观测到测量值的平均值与参考值之间的差值
  - ✓ 是测量系统的系统误差所构成



- 稳定性 (Stability)
  - ✓ 随时间变化的偏倚值
  - ✓ 一个稳定的测量过程在位置方面是处于统计上受控状态
  - ✓ 别名：漂移 (drift)

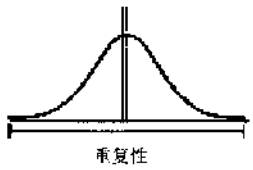
- 线性 (Linearity)
  - ✓ 在量具正常工作量程内的偏倚变化量
  - ✓ 多个独立的偏倚误差在量具工作量程内的关系
  - ✓ 是测量系统的系统误差所构成



宽度变差 (Width variation)

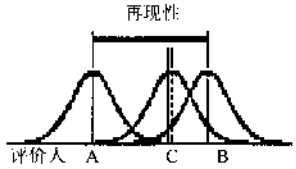
- 精确度[2] (Precision)
  - ✓ 每个重复读数之间的"接近"程度
  - ✓ 是测量系统的随机误差所构成

- 重复性 (Repeatability)

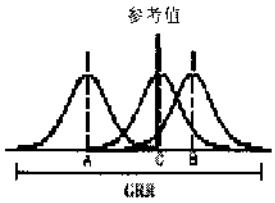


  - ✓ 一个评价人使用一件测量仪器，对同一零件的某一特性进行多次测量下的变差。
  - ✓ 是在固定的和已定义的测量条件下，连续（短期内）多次测量中的变差
  - ✓ 通常被称为 E.V. — 设备变差 (Equipment Variation)
  - ✓ 设备（量具）能力或潜能
  - ✓ 系统内部变差

- 再现性 (Reproducibility)



  - ✓ 不同评价人使用相同的量具，测量一个零件的一个特性的测量平均值的变差
  - ✓ 在对产品和过程进行鉴定时，误差可能是评价人、环境（时间）、或方法
  - ✓ 通常被称为 A.V. — 评价人变差 (Appraiser Variation)
  - ✓ 系统之间（条件）的误差
  - ✓ 在 ASTM E456-96ASTM E456-96 包括：包括：重复性、实验室、环境以及评价人影响



- GRR 或量具的重复性和再现性 (Gage R&R)
  - ✓ 量具的重复性和再现性：测量系统重复性和再现性联合估计值
  - ✓ 测量系统能力：取决于所用的方法，可能包括或不包括时间的影响

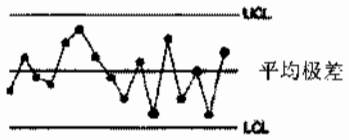
- 测量系统能力 (Measurement System Capability)
  - ✓ 测量系统变差的短期估计值（例："GRR"，包括图表法）

---

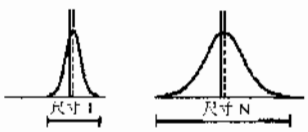
[2] 在 ASTM 文件中，没有测量系统的精密度；即，精密度不能由一个单值来解释。

- 测量系统性能 (Measurement System Performance)
  - ✓ 测量系统变差的长期估计值（例：长期控制图法）

- 敏感度 (Sensitivity)
  - ✓ 能导致可探测到的输出信号的最小输入
  - ✓ 测量系统对被测特性变化的感应度
  - ✓ 取决于量具设计（分辨力）、固有质量 (OEM)、使用期间的维修，以及测量仪器与标准的操作情况
  - ✓ 通常被描述为测量单元



- 一致性 (Consistency)
  - ✓ 随时间重复性变化的程度
  - ✓ 一致的测量过程是在宽度（变差）方面处于统计上受控状态



- 均一性 (Uniformity)
  - ✓ 在正常工作范围内重复性的变化
  - ✓ 重复性的同义词

系统变差 (System Variation)

测量系统的变差可分类为：

- 能力 (Capability)
  - ✓ 短期内读数的变化量

- 性能 (Performance)
  - ✓ 长期读数的变化量
  - ✓ 以总变差(total variation)为基础

- 不确定度 (Uncertainty)
  - ✓ 有关被测值的数值估计范围，相信真值都被包括在该范围内

测量系统必须稳定并且一致

测量系统的总变差的所有特征是假设该系统稳定并且一致。例如：变差的组成部分可能包括第 15 页图 2 所示项目中的任何组合。

标准及可追溯性

国家标准与技术协会 (National Institute of Standards and Technology, NIST) 是美国主要的国家测量协会 (National Measurements Institute, NMI)，它属于美国商业部 (U.S. Department of Commerce)。NIST 的前身是国家标准局 (National Bureau of Standards, NBS)，是美国度量衡 (metrology) 的最高权力机构。NIST 的主要职责是提供测量服务以及保存测量标准，从而帮助美国工业界建立可追溯的测量，最终能为产品和服务的贸易提供帮助。NIST 为许多行业直接提供这种服务，但主要是给那些需要高准确度产品以及在过程中结合目前测量科技进步水准的行业提供服务。

国家测量协会（NMI）

世界上许多工业化国家维持了他们与 NIST 相似的 NMI，也为他们的相关行业提供高水准的度量衡标准或测量服务。NIST 与其它国家的 NMI 合作研究，从而确保在一个国家进行的测量不会与另一个国家不同。该项工作通过各个 NMI 之间的互相认可协定 (Mutual Recognition Arrangements, MRAs)，并进行*互相实验比较 (interlaboratory comparisons)*来完成。有一点要注意，这些各国的 NMI 的能力有差别，并且不是所有类型的测量均在一个共同标准的基础上进行比较，所以会存在差别；这就是为什么要了解追溯到是谁的测量，以及是如何进行追溯的重要性。

可追溯性 (Traceability)

对商品和服务的贸易追溯是一个非常重要的概念。测量可以追溯到相同或类似的标准，比不能追溯的测量更将容易达成互相承认。可追溯的测量还可帮助减少重新试验的要求，以及好产品的拒收与坏产品的接收。

可追溯性在 *ISO 国际基本和通用的度量衡术语词汇 (ISO International Vocabulary of Basic and General Terms in Metrology, VIM)* 中的定义为：

“通过一个完整的比较链追溯到规定的参考标准（通常为国家或国际标准）的测量特性或标准值，都具有一定的不确定度。”

建立一个测量的可追溯性一般可以通过一条比较链追溯到 NMI。但是在工业界的许多情况下，测量的可追溯性可能追溯到顾客和供方双方同意的参考值或“一致的标准”。这些达成一致的标准和 NMI 之间的可追溯性也许不是总能清晰地被理解的，所以，最关键的是测量能追溯到满足顾客需求的程度。随着测量系统在工业界中的测量技术提升和目前科技进步水准，测量追溯到哪里和如何追溯的定义将成为一个不断进化的概念。

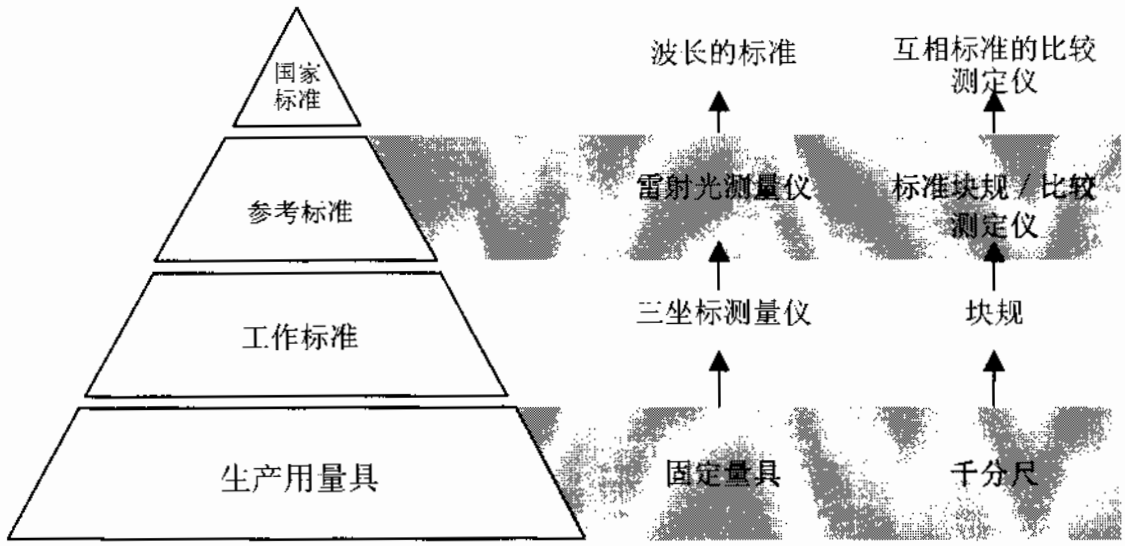


图 1：长度测量的可追溯性链举例

NMI 与不同国家实验室、量具供方、科技开发制造商等密切合作，从而确保他们的参考标准经过适当的校准，并且能够直接追溯到 NMI 所保存的标准。这些政府和私人工业机构然后可以使用他们的标准，来为他们的客户的度量衡或量具实验室、校准*工作*或其它的*初级*标准提供校准和测量服务。这些情况的连结或链将应用在工厂的现场，并提供基本的测量可追溯性。这个可以通过这种完整的测量链回溯到 NIST 的测量被称为*可追溯到 NIST*。

不是所有的组织都有自己的内部计量或量具实验室，因此，他们必须依靠外部的商业 / 独立实验室来提供可追溯的标准和测量服务，如果商业 / 独立实验室的能力已透过实验室认可来保证，以上方法是获得可追溯到 NIST 的可接受和适应的方法。

真值

测量过程目标是零件的“真”值。希望任何个别的读值能尽可能的（经济的）与真值接近。不幸的是，*真值*永远不能确切地得到。然而，基于作业上完善地定义一个特征的*参考值*，以及使用具有产生更高解析度并可追溯到 NIST 的测量系统，能够使这种不确定度减到最低。因为参考值常被用于对真值的替代，这些术语通常被互换使用；但不建议这种互换使用。

第一章 — 第 B 节

测量过程[3]

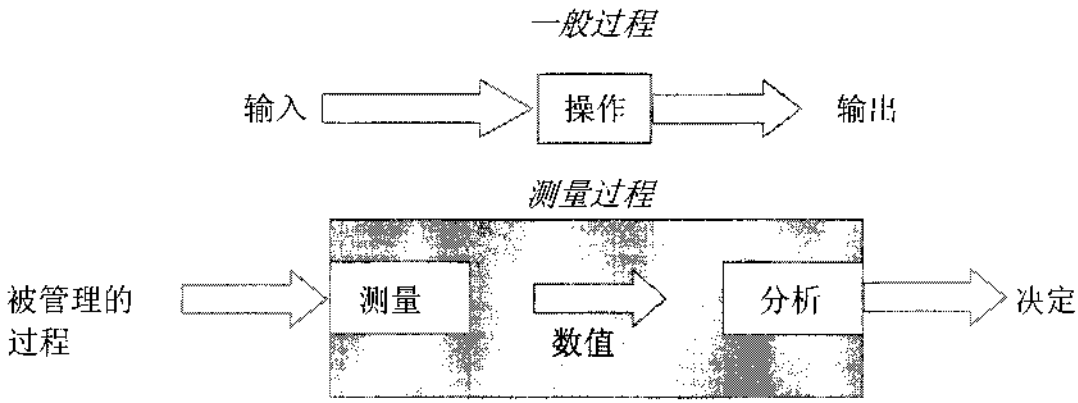
为有效地管理任何过程的变差，需要以下知识：

- 过程应该做什么
- 会出什么错
- 过程正在做什么

规范及工程要求决定过程应该做什么：

过程失效模式及后果分析[4] (PFMEA) 的目的是确定与潜在过程失效有关的风险，并且在这些失效发生之前提出要采取的纠正措施。PFMEA 的结果将转移到控制计划。

通过评估过程参数或结果可以获得过程正在做什么的知识。该活动通常称之为检验，就是透过适宜的标准和测量仪器来检查过程参数、过程中零件、装配完的子系统或完工的最终产品的活动，从而使观测者能够确认或否定该过程是在一个稳定条件下运行，其变差相对于顾客指定的目标是可接受的。且这检验本身就是一个过程。

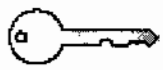


不幸的是，工业界传统上认为测量和分析活动是一种“黑箱作业”。设备是主要的重点 — 越“重要”的特性，就采用越贵重的量具。很少顾虑仪器的适用性，它与过程和环境的匹配性，以及它本身的可用性。因此这些量具通常没有得到适当的使用或简单地不去用它。

---

[3] 本章部分内容经允许改编自《*测量系统分析 — A Tutorial by G. F. Gruska and M. S. Heaphy, The Third Generation, 1987, 1998*》。

[4] 参见《*潜在失效模式和后果分析（FMEA）参考手册 — 第 3 版*》。



测量及分析活动是一个过程 — *测量*过程。在该过程中可以应用任何与所有过程控制的管理、统计及逻辑技术。

这意味着首先要识别顾客与顾客们的需求。顾客也就是过程的拥有者，希望用最小的努力获得正确的决定。管理者必须提供资源以采购必要的和足够的设备来完成工作，但购买了最好的或最新的测量技术不一定能确保能做出正确的生产过程控制决定。

设备只是测量过程的一部份。过程的拥有者必须了解如何正确地使用这些设备，以及如何分析和解释结果。因此管理者还必须提供清晰明了的操作规定和标准，以及提供培训和支持。接下来，过程的拥有者有义务监控和控制这测量过程，从而确保获得稳定和正确的结果，包括整个测量系统分析方法 — 对量具、程序、使用者和环境的研究，如：正常操作条件的研究。

测量系统的统计特性

一个理想的测量系统是每次使用时均能产生“正确的”测量结果；每个测量都会遵循某个标准[5]。能够产生这样的测量结果的测量系统被称为具有如下的统计特性：零变差、零偏倚，及对其所测量的产品被错误分析的可能性为零。不幸的，具有这样理想统计特性的测量系统很少存在；所以，过程管理者通常被迫使用统计特性不尽理想的测量系统。测量系统的质量通常仅仅取决于经过一段时间后产生数据的统计特性。当他们要促成一整体上令人满意的测量系统时，其它因素如成本、易于操作等也同样重要。但这是所产生数据的统计特性，是测量系统质量的决定因素。

对某种使用情况下非常重要的统计特性，其不一定是另一种使用情况下非常重要的统计特性。例如：在某些情况下使用三坐标测量仪 (coordinate measuring machine, CMM)，最重要的统计特性是“小”的偏倚和变差，具有这样的统计特性的 CMM 将产生“接近”可追溯标准的公认值的测量结果。用这样的仪器所获得的数据对于分析一个制造过程是非常有用的；但是，不管 CMM 仪器的偏倚和变差多么“小”，使用 CMM 这样的测量系统也许不能用来分辨好产品和坏产品的接受工作，因为测量系统的其它因素会产生额外的变差来源。

[5] 关于标准的完整讨论，参见 *out of the crisis*, W. Edwards Deming，1982, 1986, P279-281。



管理者有责任为最佳的数据应用，识别最为重要的统计特征；管理者也有责任确保使用这些特征作为选择测量系统的依据。为了达到以上的目的，需要将这些统计特性具体定义，并且规定测量它们的可接受方法。尽管每个测量系统可能被要求具有不同的统计特性，有一些基本的特性来定义一个“好”的测量系统。包括：

1) 具有足够的分辨力和敏感度。测量的增量(increment)应该小于测量目的相应的过程变差或规范限值(specification limit)。通常被称为 10 比 1 规则，也就是说仪器的分辨力应该能将公差（或过程变差）划分成 10 等份或者更多。这比例规则的意图是为选择量具时的一个实际最先遵守的原则。

2) 测量系统应处于统计受控状态[6]。这意味着在重复测量条件下，测量系统中的变差只能由普通原因(common cause)造成，而不能由特殊原因(special cause)造成。这种情况可称之为具有统计的稳定性，并且可以透过控制图法最佳地进行评价。

3) 为了*产品控制*，测量系统的变差必须小于规范限值。以特性的公差来评估测量系统。

4) 为了*过程控制*，测量系统的变差应该能证明具有有效的解析度，并且小于制造过程的变差。6σ 过程变差和／或 MSA 研究中的总变差可用来评估测量系统。

测量系统的统计特性随着被测量的项目不同可能会发生变化。如果这样，测量系统的最大（最坏）的变差必须小于过程变差或规范限值。

变差的来源

与所有过程类似，测量系统受到变差的随机原因和系统上原因的影响。这些变差的来源是由于普通原因和特殊原因造成的。为了控制测量系统的变差：

1) 识别潜在的变差来源。

2) 消除（如有可能的话）或监控这些变差的来源。

[6] 测量分析必须同时考虑实用性及统计上的特性。

尽管特殊原因取决于不同情况，但一些典型的变差来源可以被识别。有许多不同的方法来呈现与分类这些变差的来源，例如：因果图、失效树等。但本指南介绍的方法将专注在测量系统的主要因素。

S	标准
W	工作件（即零件）
I	仪器
P	人 / 程序
E	环境

缩写 S.W.I.P.E.[7] 被用来表示一个普遍化的测量系统为了达成被要求的目标，有六个必要因素。S.W.I.P.E.分别代表标准、工作件、仪器、人和程序，以及环境。可以考虑作为整个测量系统的一个误差模型[8]。

需要理解那些 6 个区域的影响因素，以便能够控制或消除它们。



图 2 显示了变差的某些潜在来源的原因与后果图。由于影响一个特定测量系统的实际变差来源对该系统将是唯一的。所以，这图可作为开发一个测量系统变差来源的思考启动器。

[7] 这些缩写由度量衡专家 Ms. Mary Hoskins 与 Honeywell, Eli Whitney Metrology Laband the Bendixcorporationg 首创提出，。

[8] 另一个供选择的误差模型见附件 F，P.I.S.M.O.E.A。

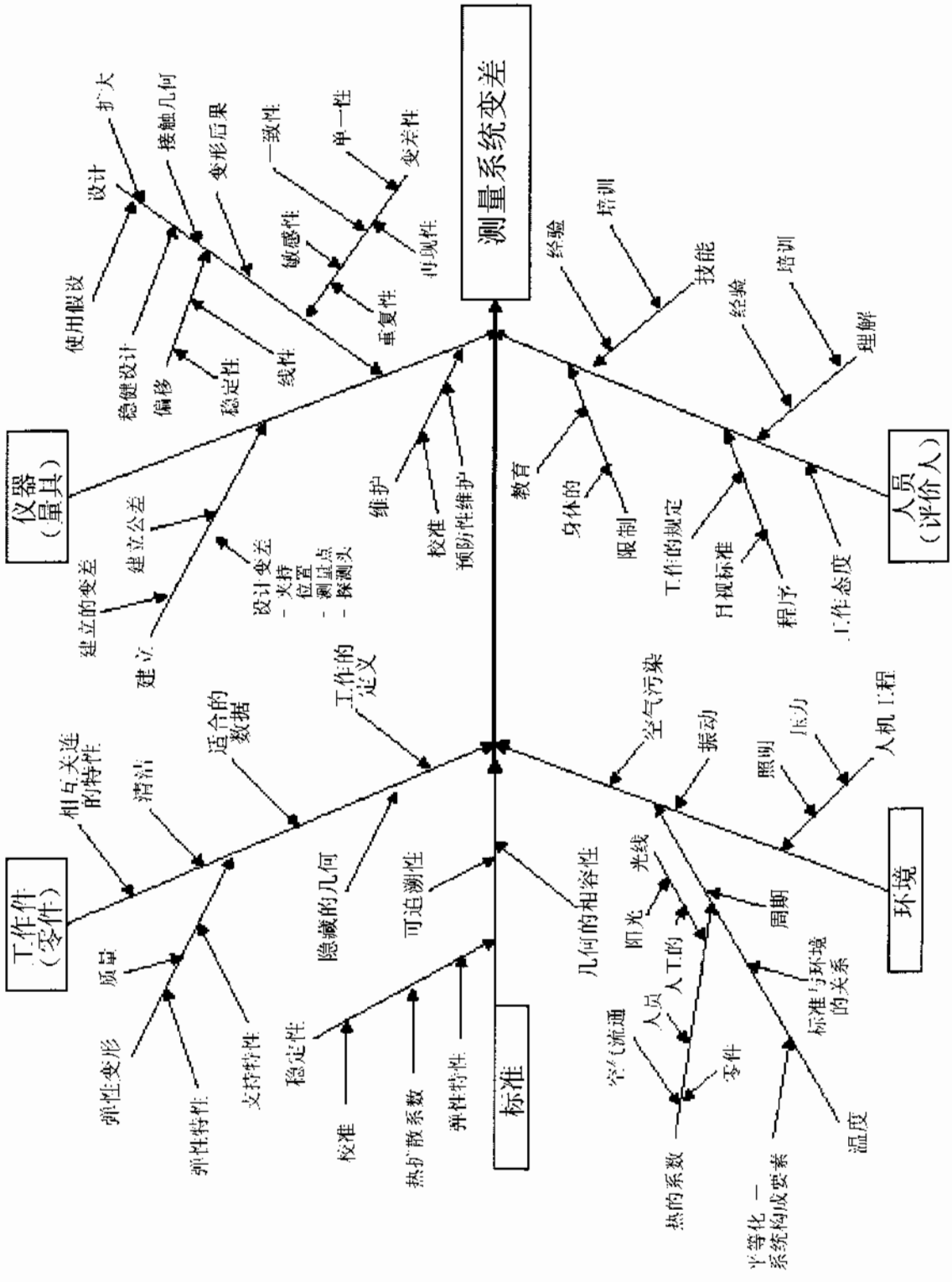


图 2：测量系统变差的因果图

测量系统变差的影响

由于测量系统会被不同的变差来源影响，相同零件的重复读值不会有相同的产出与结果，每个读值之间不同是因为普通和特殊原因变差造成的。

对测量系统变差的不同变差来源影响应该进行短期和长期的评价。*测量系统能力*是测量系统在短期间的（随机）误差，它是由线性、均一性、重复性与再现性结合的误差量。*测量系统性能*，像过程性能一样，是所有变差来源在长期的影响；这是经由确定了我们的过程是否处于统计受控状态（如：稳定并一致；仅有普通原因的变差）、对中性（无偏倚），以及具有期望结果在一定范围内的可接受变差（量具的重复性和再现性 (GRR)）的实现，这是在测量系统能力上增加了稳定性和一致性。

由于我们是用测量系统的输出来对某产品和过程做出决策，这所有变差来源的累积后果通常被称为*测量系统误差*，或有时简单地称为“*误差*”。

对决策的影响

测量一个零件后，可以对零件状态的确定采取一个行动措施。一般来说，可以用来确定该零件是否可接受（在规范内）或不可接受（超出规范）。另一个常用的情况是将零件分类为特定的类别（例如：活塞的尺寸）。

以一个范例作为接下来的讨论，有两种分类将被使用：超出规范（“坏”）和在规范内（“好”）；这并不为其它类别活动的讨论做出限制。

进一步的分类可能是可返工的、可补救的或报废。按照*产品控制理论 (product control philosophy)*，这种分类活动是测量一个零件的主要原因。但是，以*过程控制理论 (process control philosophy)*，集中的关注是在零件的变差是否由过程中的普通原因还是特殊原因造成的。

原理	关注点
产品控制	零件是否在指定的范围内？
过程控制	过程变差是否稳定并可接受？

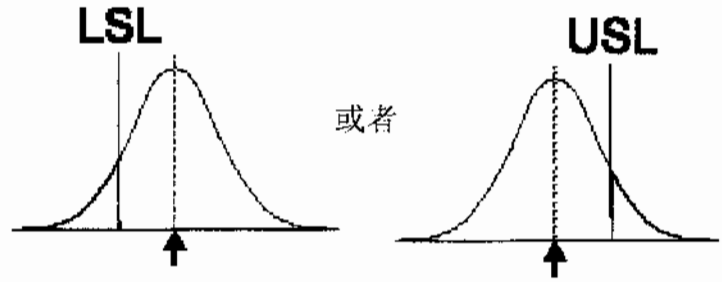
表 1：控制理论和驱使的关注点

下一部份将针对测量误差对产品决策的影响，接下来再讨论它对过程决策的影响。

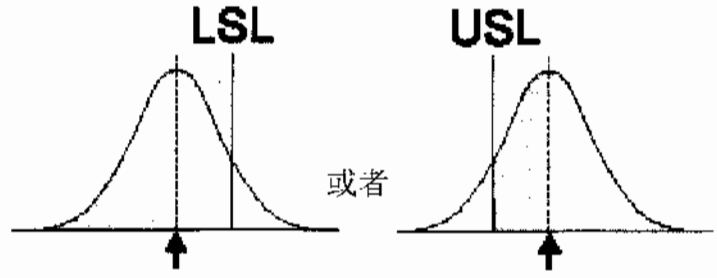
对产品决策的影响

为了更好的理解测量系统误差对产品决定的影响，考虑的案例是假设对某一个单一零件的多次读值的所有变差是由于量具的重复性和再现性引起的。也就是说，测量过程处于统计受控状态，并且零偏倚量。

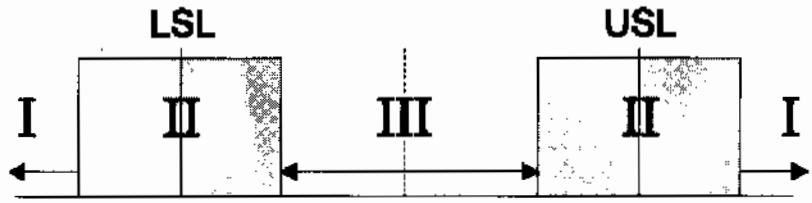
无论何时任何零件的测量分布区域与规范限值重迭时，有时可能会做出错误的决定。例如，如果出现以下情况，一个好的零件有时会被称为“坏”的（第 I 类误差，制造者的风险或错误警告）：



然后，如果出现以下情况，一个坏的零件有时会被称为“好”的（第 II 类误差，顾客的风险或过失比率）



也就是说，考虑到规范限值，只要当测量系统误差与规范限值相交时，就可能发生对零件做出浅在的错误决定。以下提出三个明显的区域。



目标

图中，

第 I 类区　坏零件永远被称为坏零件
第 II 类区　可能做出潜在的错误决定
第 III 类区　好零件永远被称为好零件

由于我们的目标是尽可能地对产品状态做出正确的决定，有以下两个选择：

1) 改进生产过程：减少过程变差，不会生产出落在区域 II 的产品。

2) 改进测量系统：减少测量系统误差以减低区域 II 的大小，这样，生产出的所有的产品将全部落在区域 III 内，且就能减少做出错误决定的风险。

以上讨论假设测量过程处于统计受控状态并且对准目标。如果以上两个假设中的任何一个不成立，那么很难相信任何观测值能带来正确的决定。

对过程决策的影响

对于过程控制，以下需求要被建立：

- 统计的控制
- 对准目标
- 具有可接受的变差

正如上一部份的解释，测量误差可能造成对产品的错误决定。测量误差对过程决定的影响如下：

- 将普通原因识别为特殊原因
- 将特殊原因识别为普通原因

测量系统变差能影响对于一个过程的稳定性、目标与变差的决定。实际的与观测到的过程变差之间的基本相互关系为：

$$\sigma^2_{观} = \sigma^2_{实} + \sigma^2_{测}$$

式中：

$\sigma^2_{观}$ — 观测到的过程变差值

$\sigma^2_{实}$ — 实际的过程变差

$\sigma^2_{测}$ — 测量系统的变差

能力指数[9] *Cp* 由下式计算：

$$Cp = \frac{\text{公差范围}}{6\sigma}$$

将上述公式进行替换可以得出观测过程指数与实际过程指数之间的关系，如下

$$(Cp)^{-2}_{观} = (Cp)^{-2}_{实} + (Cp)^{-2}_{测}$$

假设该测量系统处于统计受控状态并且对正目标，实际过程 *Cp* 与观测过程 *Cp* 的图形上比较大致相等。[10]

因此，观测到的过程能力结合了实际过程能力加上测量过程造成的误差。要想达到某一特定的过程能力目标，需要考虑测量变差中的因素。

举例说明，如果测量系统的 *Cp* 值为 2，实际过程 *Cp* 必须大或等于 1.79，才能得到 1.33 的计算值（观测值）。如果测量系统 *Cp* 值本身只有 1.33，则该过程需要没有任何变差才能使观测到的过程能力达到 1.33 — 显而易见，这是一个不可能存在的情况。

[9] 虽然讨论中使用的是 *Cp*，也可能用性能指数 *Pp* 来计算。

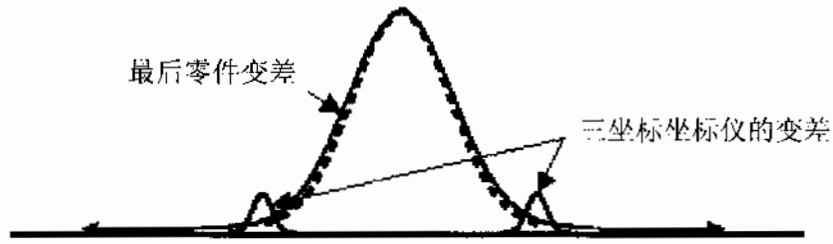
[10] 公式和图表见附录 B

新过程的接受

当一个新的过程中购买了如机械、制造、冲压、材料处理、热处理或装配，通常会有作为采购验收活动的一系列步骤需要完成；一般会在供方的场所进行的设备研究，然后在顾客的场所进行研究。

在正常条件下使用时，如果测量系统在某个地方使用与在另一地方使用不一致时，可能会产生混淆。使用不同的测量仪器的最常见情况是，在供方使用的仪器分辨力高过在生产场地使用的仪器（量具）。例如，在采购验收期间使用坐标仪来测量零件，然后在生产期间用高度规测量；再如，在采购验收用一电子天平或实验室里的机械式天平作测量（称重），但在生产期间用简单的机械式天平作测量。

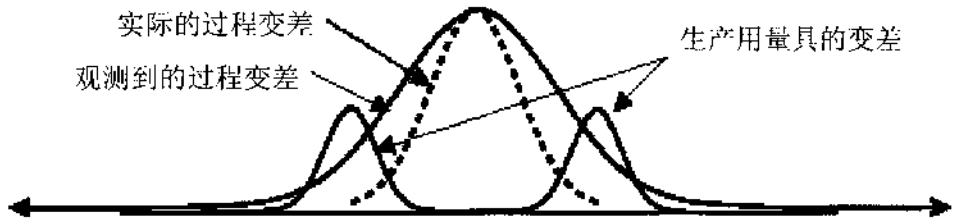
在这案例，当（较高分辨力）测量系统被使用在验收期间，其 GRR 为 10%，且实际过程 *Cp* 为 2.0，那么在验收期间观测到的过程 *Cp* 值将为 1.96。[11]



当这过程是在生产场地使用生产用的量具来研究时，会观测到更大的变差（也就是说一个较小的 *Cp* 值）。例如，如果生产量具的 GRR 为 30%，而实际过程能力 *Cp* 值还是 2.0，则观测到的过程能力 *Cp* 值将为 1.71。

最坏的情况是使用了不合格的生产用量具。如果测量系统的 GRR 是 60%（但是并不知道这一事实），那么观测到的 *Cp* 值将是 1.28。观测到的 *Cp* 值由 1.96 到 1.28 之间的区别就是由测量系统的不同所造成的。如果不能了解这情况，可能要努力地寻找新的过程在什么地方出错了，当然这些努力是白费的。

[11] 讨论这些是假设没有抽样的变差。在事实上，1.96 是期望值，但实际结果会在它的附近变化。



过程作业准备/控制（漏斗试验）

通常在制造场所会用每天开始工作时的第一个零件来验证过程是否对准目标值。如果被测零件没有对准目标值，则对过程进行调整；接下来，在某情形下另一个零件被测量，并且过程再度被调整。戴明博士将这种测量和做决定的方式称为“*擅改 (tampering)*”。

让我们思考一种情况：在某零件表面镀上一种贵重金属的重量被控制在目标值 5.00 克。假设不知道用来测量重量的天平偏离了±0.20 克，因为从来没有对测量系统进行过分析，所以人们不知道这一事实。作业指导书要求操作者在作业准备时和每间隔 1 小时用一件产品进行重量验证。如果结果超过了 4.90 至 5.10 的范围，则操作者重新进行过程的作业准备。

在进行作业准备时，该过程在 4.95 克下运行，但是由于测量误差使操作者观测到的值为 4.85 克；按照作业指导书的规定操作者设法将过程上调 0.15 克，目前这过程在目标值为 5.10 克的情况下生产。当操作者在下个时间进行作业准备验证时，他观测到 5.08 克，于是过程被允许生产，可是这过程已经是被加上变差的过度调整 (over-adjustment)，并将会继续下去。

这是漏斗试验 (funnel experiment) 的一个例子，戴明博士称它为擅改[12]影响的结果。测量误差加重了这种问题。

漏斗试验有以下四个规则：

规则 1：除非过程不稳定，否则不要进行调整或采取措施。

[12] Deming, W. Edwards, 《*Out of the Crisis*》, 麻省理工学院, 1982, 1986。

规则 2：从过程上一次被测量偏向的相反方向，对过程调整一个相等的量。

规则 3：将过程重新设置到目标值，然后从目标值往相反方向调整一个相等的量。

规则 4：将过程调整到上一次测量的位置。

为贵重金属镀层过程的作业准备指导书是规则 3 的一个范例。规则 2、3 和 4 将累进的增加更多变差，规则 1 是产生最小变差的最好选择。

漏斗试验的其它范例如下：

- 武断的以一个界限为基础来校准量具 — 也就是说界限值不能反映测量系统的变差。（规则 3）
- 武断的使用没有任何指标或来历的一项变更（特殊原因）的数值之后，设置（重新设置）过程控制的测量系统的基准。（规则 3）
- 以上一次生产的产品为基础，透过自动补偿的方法来调整过程（规则 2）。
- 在进行工作培训时，由工人 A 训练工人 B，再由工人 B 培训工人 C...，没有一个标准的培训材料。类似"邮局"的游戏。（规则 4）
- 测量了零件，并已发现偏离目标，但在控制图上的图形呈现过程是受控的 — 因此不采取措施。（规则 1）

第一章 — 第C节

测量的策略和计划

引言

在设计和采购测量设备或系统前，进行计划是很关键的。在计划阶段所做的许多决定会影响测量设备的倾向和选择。测量的目的是什么？如何来使用测量的结果？计划阶段将解决这些问题并且对测量过程的操作产生重要的影响，同时可以减少可能出现的问题以及以后的测量误差。

在某些情况下，由于被测零部件所涉及的风险或者由于测量装置的成本和复杂性，OEM 顾客可能要求使用 APQP 过程和承诺以决定对供方的测量策略。



不是所有的产品和过程特性都要求用如此复杂的方法来开发测量系统。诸如分厘卡或游标卡尺之类的简单标准测量工具可能不需要如此深奥的策略和计划。一个基本的原则是这零部件或次总成系统的被测特性是否已在控制计划中注明，或者是否是对产品或过程的接受决定的一个重要特性。另外，还可根据对某特定尺寸的指定公差等级来决定。在任何情况下，常识就是指导方针。

复杂程度

测量系统的类型、复杂程度和目的可能驱使不同程度的项目管理、策略计划、测量系统分析，或其它为测量系统选择、评估和控制的其它特别考虑。简单的测量工具或装置（例如：天平、刻度尺、固定的限规或计数性量具）也许不需要更复杂的或重要的测量系统所要求的管理、计划或分析等级（例如：基准或参考标准、CMM、试验标准、自动化的线上量具等）。任何测量系统需要进行更多或少的策略计划和详细程度，取决于产品和过程的情况，适当程度的决定必须由指派对测量过程和顾客负责的 APQP 小组来确定。在后续许多活动中涉及或实施的实际程度，应该由特定的测量系统、支持量具控制和校准系统的考量、对过程深切地了解和常识来决定。

识别测量过程的目的

第一步是要建立测量的目的以及如何应用测量。为了完成这工作，关键的是在测量过程开发之前组织一个跨功能小组，在审核、过程控制、产品和过程开发与“测量生命周期”分析的关连下作成特别的考量。

测量生命周期

“测量生命周期 (Measurement Life Cycle) ”这一概念是指随着时间的不同，对过程的了解以及过程的改进，测量方法的可能改变。例如：为了建立稳定和有能力的过程，可能开始对一个产品的特性进行测量；透过测量了解了对直接影响产品特性的关键过程控制特性。这了解意味了对产品特性信息的依靠减少了，并且可能减少抽样计划（从5件／小时，减少到1件／每班）；同时，测量的方法也可以从 CMM 测量改为用某种计数性量具测量；最后，可能发现可能只需要监测极少数的零件，只要过程被维护着或测量和监控这维护和工装，也许就是所有必要的工作了。测量的程度是依赖着对过程理解的程度。

对供方进料，大部份的测量和监控最终可能会取消。经过在相当长的时间后，在过程的同一区域对相同的特性进行同一测量，这是对测量过程缺乏学习或停滞的证据。

测量过程设计的选用准则

在采购一个测量系统之前，一个详细的测量过程的工程概念要被建立。利用上述的所开发的目的，为了符合设计要求，各个跨功能小组将开发测量系统的计划与概念。下面是一些指南：

小组需要评估子系统或零组件的设计，并识别其重要特性。重要特性的识别应以顾客的要求和子系统或零件在整个系统中的功能为基础。如果重要的尺寸早已被识别，则评估测量该特性的能力；例如，如果一个塑胶射出零件的重要特性是它的分模线，对这尺寸测量将很困难，测量的变差也会很高。

可用 FMEA 过程来解决上述类似的问题，即从测量零件的能力以及量具的功能两方面来分析量具设计的风险区域（设计和过程 FMEA）。这种方法同样可以用来帮助开发维护保养和校准计划。

画出过程流程图，以标示出零件或子系统的制造或装配过程中的关键过程步骤，识别该过程每个步骤中关键的输入和输出；这样有助于开发那些受过程场所影响的测量设备准则和要求。透过这样的调查，[13] 可以得到测量计划及测量类型的清单。

对于复杂的测量系统，要作成测量过程的流程图；包括被测零件或子系统的交付、测量本身，以及退回零件或子系统到过程。

接下来使用一些头脑风暴法来开发用于每个测量所要求的共通准则。使用因果图是一种简单的方法。[14] 参见图 2 的范例以作为思考的开始。

与测量计划有关的几个需考虑的其它问题：

- 谁应加入“需求”分析？流程图和最初的讨论将有利于关键项目的识别。
- 为什么要采取测量？及其如何被使用？数据是否用来控制、分类、资格验证等？测量的应用方式会改变测量系统的敏感度等级。
- 需要什么级别的敏感度？产品规范是什么？预期的过程变差是什么？需要用量具来检测的零件之间差别有多少？
- 需要对量具提供那些信息（例如：使用手册 — 操作、维护等）？需要操作者具备哪些基本技能？谁来实施这培训？
- 如何进行测量？是否可以进行人工测量？在移动的输送带上？离线测量？自动测量等？零件的位置和固定是否是可能的变差来源？接触或非接触测量？
- 这测量如何被校准？是否将以其它的测量过程来比较？谁负责这校准的基准？
- 测量在何时何地进行？零件是否被清洁、涂油、加热等？

---

13 可将这个看成一种初始的控制计划

14 参见 *“Guide to Quality Control”*, Kzoru Ishikana, Asian Productwity Organization 出版，1986。

记住，数据的使用是证实对于测量过程的一般假设；这比较保险因为这是在操作环境下收集的数据，胜过基于错误的信息下做出决定与开发一个基于非稳健环境问题下的系统。

研究不同的测量过程方法

在投资到新设备之前，应对目前使用的测量方法进行研究。对测量方法进行证实可以提供更可靠的操作，如有可能，使用有追踪记录证实的测量设备。

概念和建议方案的开发和设计

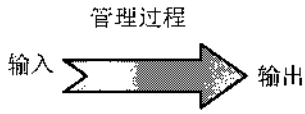
当进行概念和建议方案的开发和设计时，参考第一章第D节后面的“测量系统开发检查表的建议要素”。

在测量设备的制造和测量过程的开发（方法、培训、文件化等）之中和之后，将进行试验研究和数据收集活动，利用这些研究和数据来理解这测量过程，从而对这过程和将来的过程进行改进。

第一章 — 第 D 节

测量资源的开发

引言



本节介绍一测量过程的报价 / 采购过程的寿命时间范围。这已经建构了对于下列的一种自行控制的讨论：开发一测量过程报价方案的过程、获得对报价方案的回复、裁定项目、完成最终设计、开发测量过程，并在最后将已建立的测量过程与生产过程互相结合。强烈要求在没有完全阅读和理解对测量过程的所有讨论的情况下，不要使用本章节。为了从测量过程中获取最大的利益，用具有输入和输出的过程来研究并表述它。[15]

本章内容是以团队合作的哲理所编写。这不是为购买者或采购机构所使用的工作说明。成功地完成这里所说明活动将需要小组的参与，并且应该将它纳入先期产品质量计划 (APQP) 小组的整体工作时间架构下管理。这样可以使不同功能的小组之间良好的合作关系 — 在计划过程中提出的某些概念可能会在量具供方作出最后设计之前被修改，进而满足测量系统要求。



通常，“采购过程”始于顾客和供方之间为某一特定项目所进行的正式联系。直截了当的联系是项目成功的关键，因为建立一个有效的顾客 / 供方关系所必需的基础工作将在这个阶段完成。顾客正式地把项目的内容列在“报价需求通知 (Request For Quote, RFQ)”即是采购过程的开始，随之而来的是供方为了符合这些意图（报价）的正式计划案说明。顾客和供方需要彻底地理解项目的要求，例如：将要交付些什么及交付方法都需由双方达成一致的意见。这种理解要通过双方准确及时地沟通才能达成。

一旦这概念达成了某种协定，并且已经为该项目建立了顾客 / 供方关系，就可以开始着手详细地设计、测量过程的制造、以及开发活动。顾客与供方的沟通在这一阶段尤其是特别重要的，因为可能有几个概念阶段要进行批准、可能的环境变化或小组成员改变，测量过程项目可能会动摇甚至失败。如果在顾客和供方之间保持并记录了经常的、详细的沟通，并由双方指派正式负责人员（分别的）维护这些沟通的话，可以降低这种风险。APQP 过程是上述讨论和活动形式的理想方式。

测量过程已经被概略的设计完成之后，便可以开始测量过程 / 系统的相关采购活动。

---

[15] 参见第一章，第 B 节

数据的协调

理想上，最好使用目前普及的几何尺寸及公差 (Geometric Dimensioning & Tolerancing, GD&T)，在整个制造过程和测量系统中的数据需要被协调(例如：一致化)，并且应在 APQP 活动的非常早期建立。这项工作的初始负责人可能是产品设计工程师、尺寸控制等，依据特定的组织。如果数据方案 (datum schemes) 与整个制造过程，特别是测量系统不相匹配，那么可能测量了错误的事项，可能会发生适用上的问题等，从而导致不能对制造过程实施有效的控制。

有时用于最终装配的数据方案可能与用于子部件制造过程的数据方案不相匹配。如果出现了这种情况，要尽早地在 APQP 活动中提出来，以便小组所有的成员了解可能遇到的困难和冲突，从而尽早的利用一切机会来克服它们。在这个过程中，可能需要探索不同的数据方案以便理解这些差异的影响。

某些零件的显着特性会比其它特性产生更多的问题，如凸轮轴的中心，或其它圆形体，圆柱体或管状体的特性。例如：一凸轮轴必须加工成以中心线为中心，但它最重要的产品特征是在其凸轮上。因此，制造测量时可能要求用一种方法或数据方案，而在最终产品测量时则要求使用另一种方案。

先决条件及假设

在讨论开发量具供方之前，首先要假设“正确的”产品工程设计 (GD&T) 和“正确的”过程设计（在过程中适当的时间和地点进行测量）中的问题已经被解决。这些假设议题不应该被转移注意，在 APQP 活动早期由适当的小组成员来考虑。



假设量具供方将参与 APQP 过程的*小组*活动，这量具供方将对整个生产过程和产品使用有清晰的了解，所以，他的角色不仅能由他自己去了解，而且能由小组中的其它人员（制造、质量、工程等）参与。

某些活动中或活动的顺序可能会有稍微的重迭，这些活动取决于特定的计划 / 项目或其它约束。例如，APQP 小组可能没有从量具来源取得充足的输入来发展可靠的量具概念，而其它概念可能需要量具来源的专业知识。这将因为测量系统的复杂性和小组的决定来驱使它更合理。

量具资源的选择过程

开发报价方案

详细的工程概念

在正式向潜在供方提出一个正式的测量过程需求报价方案之前，需要开发一份测量过程的详细工程概念。小组中负责测量过程维护和持续改进的成员是开发详细工程概念的直接责任人员。他们可以是 APQP 小组中的一部分。为了更好的开发这些概念，需要回答几个问题。



小组可能研究不同的议题，有助于决定在设计这测量过程将遵循的方向和途径；有些概念可能在产品设计中就已经有规定或有隐含的规定。当开发这详细概念，需要由小组讨论的许多可能性问题可参见本节末的“测量系统开发检查表的建议要素”。

通常，顾客非常依赖供方来解决问题。在顾客要求供方为解决过程问题提供建议之前，负责该过程的小组需要好好地理解整个过程的基础和意图，如此，只有这样，过程才能被依此理解适当的被使用、支持和提高。

预防性维护的考虑

什么活动需要进行有计划的预防性维护（例如：润滑、振动分析、整体探查、更换新零件等）？大部份的这些活动将取决于测量系统、装置或仪器的复杂程度。简单的量具可能只需要进行周期性的检查，而复杂的系统可能需要持续进行详细的统计分析，并且有一组工程人员透过预知性 (predictive) 的方式来进行维护保养。

预防性 (preventive) 维护活动的规划应该与测量过程计划的初衷相一致。许多活动，例如每天清理空气滤清器、设备运行设定的若干小时后就润滑轴承等，可以在测量系统完成制造、开发和使用之前就予以规定。事实上，这样做是很完善并且能够进一步改进测量的计划与成本。与这些活动有关的数据收集方法和推荐的保养方法可以从原设备制造商处获取，或由工厂的工程、制造和质量人员来开发。在测量过程完工并投入使用后，需要经常的收集与测量过程功能相关的数据并进行画图，可以使用简单的分析方法（推移图 (run chart)、趋势分析）来确定这系统是否稳定。最后，依据系统稳定性规定的判断，能够计划性的进行规律的预防性维护。基于时间序列的信息对一稳定的系统进行预防性维护，将比用传统的技术对系统进行预防性维护更经济。

规范 (Specification)

在设计和制造过程中，规范可作为顾客和供方两者的指南，这些指南作为对可接受标准的沟通。可接受标准可能考虑了以下两个方面：

- 设计标准
- 制造标准

设计标准的形式可能取决于为项目付费的人而不同；成本问题将影响这形式。一般来说，具有足够文件化的设计细节是一个好主意，因为这样可能由任何合格的制造商来制造或修补以符合这设计的原始意图 — 但是，这种决定会受到成本和关键性的限制。必要的最终设计形式可能是一些从 CAD 或工程图面上取得，这可能包括从 OEM、SAE、ASTM 或其它组织所摘选的工程标准；量具的供方必须取得并理解这些标准的最新版本。OEM 可能要求在设计或制造阶段应用特定的标准，甚至可能要求在启用测量系统之前进行正式的批准。

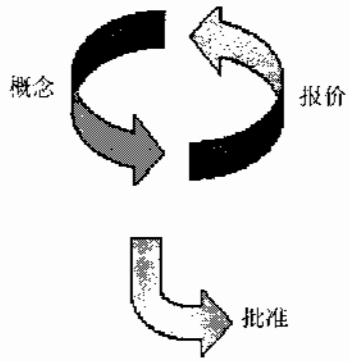
设计标准将会详细地规定将设计传达给制造者的方法（CAD — 例如：CATIA、Unigraphics、IGES、手绘文件等）。对于很复杂的测量系统，还可能包括性能标准。

制造标准将包括制造测量系统所允许的公差；制造公差应该以用于制造量具或量具零件的过程的综合能力，以及预定的被测量特性的关键性为基础。制造公差不应该仅仅是产品公差的某一设定的百分数。

> 如果需要生产多个夹具或系统，适应的计划和标准化可以使系统具有可互换性和便利性。

使用标准化的零件或次总成件也可以使系统具有可互换性、便利性、降低成本，并且通常能减低长期测量的误差。

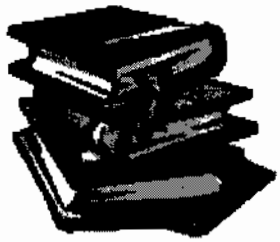
评估报价单



得到报价单后，小组成员要聚集在一起对它们进行评审和评价。要注意以下项目：

- ✓ 是否符合基本的要求？
- ✓ 是否有任何明显的顾虑？
- ✓ 是否有任何的供方提出了异常的情况，为什么？
  （异常情况可能是对价格和交付方面的重大偏离 — 没有必要把这个看成一个负面因素 — 一个供方可能已经发现了其它供方所忽视的地方。）
- ✓ 概念是否提升了简单易懂且可维修性？

可交付的文件



通常采购一测量过程有时会忽视文件的作用，不能理解文件在任何成功的项目中所担任的重要意义。文件的一般作用是为了测量过程硬件在交付时，提供一套对机械的或电子的设计的原始数据（CAD 或图面）。这些可以满足初始的应用要求，但这文件没有说明潜在的磨损部位、建议可能的故障区域，或没有说明如何使用该过程。因此，用于任何过程所必要的文件要能包括测量设备的总成图和详细的零件图，及更多的文件。

用于任何系统的有效文件的作用与一张好的旅行地图是一样的。例如，它可以告诉使用者如何从一个地方到另一个地方（使用者指导书或量具操作手册）；当主要的路径不通或关闭时，这些文件可以告诉使用者采用其它可能的方法来完成想达到的目的地（问题诊断指南或故障诊断分析）。

一套完整的文件可能包括：

- 一套可复制的总成和详细的机械图面（CAD 或文件）（包括任何要求的基准）。
- 一套可复制的电子电路图、逻辑图及软件。
- 大量使用或易磨损项目 / 部位的建议备用零件清单。该清单应该包括可能需要较长的时间采购的项目。
- 附有机械解剖图的维修手册，以能够依步骤正确的装配或拆卸机械零组件。
- 为作业准备、操作方法及机械搬运定义了有用的要求的手册（例如：安装轴承部份）。
- 故障诊断分析与问题诊断的指南。
- 认证报告（适用时，可追溯到 NIST）。
- 校准指导书。
- 可供技术支持人员、系统操作人员和维修人员使用的使用者手册。

上述可作为建立一份报价方案的检查清单，然而并不必包括所有内容。

这里讨论的中心议题是沟通。由于文件是沟通的一种形式，小组成员及其它人员要能参与测量过程文件开发的各个阶段。

在供方处的鉴定

适用情况下，在测量系统供方交付之前，应对量具或测量系统进行全尺寸的检验和功能试验。因此明显的，被选择的供方必须在生产现场具有合格的测量设备和人员，以便完成以上的工作；如果无法进行，必须事先安排一个合格的外部独立实验室来完成这些工作。这种全尺寸检验和 / 或试验应该是基于顾客的设计和制造标准要求下所执行的结果，并且应该完全记录下来以供顾客评审。

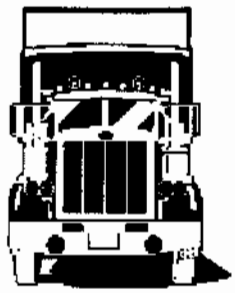
经过成功的全尺寸检验后，供方应该进行初期的正式测量系统分析。同样，事先要求供方具有有合格的人员、知识及经验来完成这适当的分析。顾客应该事先与供方（可能还有 OEM 参加）一起确定需要进行哪些种类的分析，并且应该了解供方可能需要的任何指导。需要进行讨论、协商或一致同意的事项包括：

- 初始 MSA 研究的目的：
  - ✓ 量具重复性 (GR[16]) 和再现性(GRR) 的比较
  - ✓ 偏倚和/或线性的评估
  - ✓ 顾客对测量目的评估
- 研究中所需零件、测量次数及操作者的数量
  - ✓ 接受准则
- 利用供方人员和顾客人员的比较
- 人员所需要的培训
  - ✓ 他们能够胜任吗？
  - ✓ 他们理解研究的意图吗？
  - ✓ 可能需要使用什么软件？

此时此刻不论得到什么样的结果，都应该意识到这些仅仅是初始分析，需要进行判断从而确认这结果是否可以接受。

[16] 见附录 D。

运输



检查表

- 设备应该在什么时候交运？
- 应该如何运输？
- 谁把设备从卡车或火车上卸下来？
- 是否需要保险？
- 文件是否应该与硬件一起交运？
- 顾客是否有合适的设备卸下硬件？
- 交运前，系统将存放在什么地方？
- 应用前，系统将存放在什么地方？
- 交运文件是否完整并且易于被装载、搬运、卸货及安装的人员理解？

在顾客处的鉴定

在上述供方交运之前对测量系统所做的鉴定工作，通常在交付完成后，在顾客场所应该用同样的方式重复做一次。因为这是测量系统在其打算工作的环境下进行的首次实际研究机会，所以要慎重地考虑使用接受标准和分析方法。各团体积极参与这部份对这测量系统的最后成功以及它所产生的数据的应用极为重要。

收到测量系统后，在进行任何测量分析之前，应该对其进行全尺寸测量，从而确定它是否满足制造要求 / 标准。这全尺寸检验工作的范围可能与测量系统在交运前在供方处所做的全尺寸结果进行比较，以获得在供方处所做的全尺寸结果的质量自信，并对在运输过程中没有损坏的信赖。当在对运输前后进行比较时，要意识到由于测量系统条件不同，两个测量结果会存在一些差异。

文件交付

为了帮助任何系统的实施或启用，至少需要以下信息（这些信息要在设备交付之前给提供给顾客）：

- 如果小组要求，CAD 或图面文件
- 适用时，该系统的流程图

- 使用者手册
  - ✓ 维修 / 服务手册
  - ✓ 备用零件清单
  - ✓ 问题诊断指南
- 校准说明书
- 其它特别需要考虑的事项

在一开始时，已交付的文件需要被注明为“初始”，这个时候可以不交付原始的或可复制的文件，因为在实施后可能需要进行修订。事实上，直到整个系统被执行之后，不提交原始文件的做法是明智的 — 在供方处更改文件比在顾客处更有效率。

测量系统开发检查清单的建议要素

本清单应该根据测量系统的情况和类型进行修改。最终检查清单的建立应该是顾客和供方合作的结果。

与测量系统设计和开发有关的问题：

- 被测特性是什么？特性的类别是什么？是机械上的特性吗？是动态的还是静态的？是一项电的特性吗？其零件内部变差大吗？
- 测量过程的结果（输出）将被应用的目的是什么？生产改进、生产监控、实验室研究、过程审核、出货检验、进货检验、对 D.O.E 的回应？
- 谁将使用该过程？操作者、工程师、技术员、检验员、审核员？
- 培训要求：操作者、维修人员、工程师；教室、应用实习、在职训练、学徒期间。
- 变差的原因是否已被识别？透过小组、头脑风暴法、渊博的过程知识、因果图或矩阵图等方法建立一个误差模型（S.W.I.P.E 或 P.I.S.M.O.E.A.）。
- 是否展开了测量系统的 FMEA？
- 弹性的或专用的测量系统：测量系统可以是固定的、专用的还是弹性的(flexible)，是否有测量不同类型零件的能力；例如爪型量具、夹紧量具、三坐标座标测量仪等。弹性的量具价格较贵，但从长远来看能节约成本。
- 接触式或非接触式：可靠性、特性的类型、抽样计划、成本、维护保养、校准、人员技能要求、兼容性、环境、速度、探头的类型、零件的变形、影像处理，以上内容可能由控制计划和测量的频率来确定（全接触式量具在连续抽样时可能会过度磨损）。整个表面接触的探头、探头的类型、空气回流喷嘴、影像处理、CMM 与光学比较仪等。
- 环境：灰尘、水分、湿度、温度、振动、噪音、电磁干扰 (electro-magnetic interference, EMI)、大气流动、空气杂质等。实验室、工场、办公室等。在小且严格的公差下以微米为单位的测量系统中，以及在 CMM、光学系统、超音波仪器等环境中，环境成为一个关键的问题。对线上自动反馈类型的测量也是一个影响因素。切削油、切削碎屑及极端温度也会成为问题。是否清洁环境的要求？
- 测量及固定点：使用几何尺寸与公差 (GD&T) 清晰地定义固定位置和夹紧点，以及在零件的什么部位进行测量。
- 固定方法：不固定或夹紧零件。
- 零件方位：主体的位置或其它的位置。
- 零件准备：在测量之前，零件是否应该清洁、储油、温度稳定等？

- 感测器的位置：从主定位器或定位系统的取向角度与距离？
- 相关问题#1 — 备份的量具：在工场内或不同工场之间是否需要备份的（或多个）量具支持？制造的考虑、测量误差的考虑、维修的考虑、标准是哪个的考虑？如何才能使每个考虑问题均符合要求？
- 相关问题#2 — 方法差异：在可接受的实施和操作极限内，由不同的测量系统设计对同一产品 / 过程进行测量的测量误差结果（例如：CMM 对手工量具或开放式设定量具的测量结果）。
- 自动或手动：线上、线外、操作者依赖性。
- 破坏性和非破坏性 (Nondestructive, NDT)测量：例如：拉力试验、盐雾试验、电镀 / 涂装的厚度、硬度、尺寸测量、影像处理、化学分析、应力、耐久性、冲击、扭力、扭矩、焊接强度、电特性等。
- 潜在的测量量程：可能的测量尺寸大小和期望的量程。
- 有效的分辨率：对应用在一特殊的应用场所的可接受性，如测量对物理变化是否敏感（探测过程或产品误差的能力）？
- 敏感度：最小输入量的信号能产生一个可探测的输出（可辨别的）信号，测量装置对应用这种情况的可接受性？敏感度是由量具的固有设计和质量 (OEM)、使用期间的维护，和操作条件所决定。

<u>与测量系统制造有关的问题（设备、标准、仪器）</u>

- 是否已在系统设计中针对变差来源的识别？设计评审；验证和确认。
- 校准和控制系统：推荐的校准计划和设备审核及其文件。频率、内部或外部、参数、生产过程中的验证检查。
- 输入要求：机械的、电子的、液压的、真空的、波动抑制器、干燥器、滤清器、作业准备和操作问题、隔离、解析度和灵敏度。
- 输出要求：类比或数位、文件和记录、档案、保存、存取、备份。
- 成本：开发、采购、安装、操作和培训的预算要素。
- 预防性维护：形式、计划、成本、人员、培训、文件。
- 可维修性：内部和外部、场所、支持程度、回应时间、服务配件的可取得性、标准零件清单。
- 人机工程学 (Ergonomics)：在长时间的装载和操作设备过程中，人员不被伤害的能力。测量装置的讨论需要着重在测量系统与操作者之间的相互关系。
- 安全的考虑：人员、操作、环境、切断。
- 贮存及场所：建立对测量设备的贮存及场所的要求。隔离、环境、安全、取得性（接近）有关的问题。
- 测量周期时间：测量一个零件或特性需要多长时间？测量周期要与过程和产品控制合并。

- 是否有任何对过程流程、批次完整性、纪录、测量和零件回复的干扰？
- 材料搬运：是否需要特殊的支架、支撑夹具、搬运设备或其它物料搬运设备来放置被测零件或对测量系统本身？
- 环境问题：是否有特殊的环境要求、条件、限制等影响本测量过程或临近的过程？是否要求特殊的排气？是否有必要控制温度或湿度？湿度、振动、噪音、电磁干扰、清洁？
- 是否有任何特别的可靠性要求或考虑？设备是否能够在任何时间下维持其状况？在生产使用之前是否需要进行验证？
- 备用配件：共享清单、适当的供应和订购系统、可取得性、导入期的理解与说明。是否有足够的安全库存？（轴承、软管、皮带、开关、插座、阀等）
- 使用者说明书：夹紧顺序、清洁程序、数据解释、图表、目视辅具、易于理解的。可取得性、适当的陈列。
- 文件：工程图面、诊断分析、使用者手册、语言等。
- 校准：与可接受的标准进行比较；可接受的标准的可取得性和费用。建议的频率、培训要求、停机时间的要求。
- 贮存：是否有与测量装置贮存有关的特别要求或考虑？隔离、环境、防止损坏／窃盗等。
- 防错：已知的测量程序错误是否容易由操作者更正（非常容易？）数据输入、设备误用、防错。

<u>与测量系统实施有关的问题（过程）</u>

- 支持：由谁来支持测量过程？实验室技术人员、工程师、生产、维护保养、与外部签订保养合同？
- 培训：为了使用和维护本测量过程，需要为操作者／检验人员／技术人员／工程师提供哪些培训？时间、资源和费用。谁来培训？在哪进行培训？导入期要求？与测量过程的实际使用相协调。
- 数据管理：如何管理从本测量过程输出的数据？人工、电脑化、汇总法、汇总频率、评审方法、评审频率、顾客的要求、内部的要求。可取得性、储存、存取、备份、安全。数据的解释。
- 人员：是否需要聘请人员以支持这测量系统？成本、时间、可取得性有关的问题。目前的或新的。
- 改进方法：由谁来对测量过程进行经常性的改进？工程师、生产人员、维护保养人员？使用什么评价方法？是否有一系统以识别需要的改进？
- 长期稳定性：长期性研究的评估方法、形式、频率、需求。漂移、磨损、污染、操作的完整性。长期误差是否能够被测量、控制、理解、预测？
- 特别的考虑：检验人员的特质、体能限制或健康问题：色盲、视力、身体强度、疲劳、耐力、人机工程。

第一章 — 第E节

测量问题

引言

当评估一个测量系统时必须讨论三个基本问题：

1) 测量系统必须证明有足够的敏感度。

   ✓ 首先，仪器（和标准）是否有足够的分辨力？选择测量系统的基本开始点是由设计所决定的分辨力（或等级）。通常应用 **10:1** 的原则，也就是说，仪器的分辨力应该把公差（或过程变差）细分为 10 分之一或更多。
   ✓ 第二，测量系统是否证明具有有效的解析度？与分辨力类似关系，有效解析度要确定测量系统在应用状况下具有敏感度以检测出产品和过程变差。

2) 这测量系统必须稳定

   ✓ 在重复性状况下，测量系统变差仅由普遍原因而不是由特殊原因（无秩序的）产生。
   ✓ 测量分析必须时时刻刻考虑实施和统计的重要性。

3) 统计的特性（误差）要一直保持在期望的范围内，并且足以满足测量的目的（产品控制或过程控制）。

长久以来，传统上只用公差的百分比作为测量误差的报告方法，这已不能适用于强调过程持续改进的市场竞争的策略。随着过程的变化和改进，一测量系统必须为它的使用意图重新评价。组织（管理者、测量计划者、生产操作者和质量分析人员）理解测量的目的并且进行适当的评价是不可或缺的。

测量系统变差的类型

测量通常被假设是精确的，而且许多后续的分析和结论都是以该假设为基础。个人可能无法意识到在测量系统中存在影响各个测量的变差，它进而又影响基于这些数据的结论。测量系统误差可分成五类：偏倚、重复性、再现性、稳定性和线性。

测量系统研究的目的之一是获得测量系统与其所处环境有相互作用时，其产生的测量变差的类型和结果的讯息。这些信息是很有价值的，因为对于一般的生产过程，确定重复性和校准偏倚，以及为这些制定合理的极限，比提供具有非常高重复性的准确量具更有用。应用这种研究可提供：

- 接受一新测量设备的准则。
- 一个测量装置与另一个测量装置的比较。
- 评价一个疑似不充分量具的依据。
- 测量设备维修前后的比较。
- 为计算过程变差以及生产过程可接受程度的必要构成元素。
- 描绘量具性能曲线 (Gage Performance Curve, *GPC*)[17] 必要的信息。量具性能曲线是指接受某一零件真值的概率。

下列定义帮助我们描述与一测量系统有关的误差或变差类型，以便在其后的讨论中清楚地理解各个术语。每个定义都提出了说明，以便生动的展现每项术语的含义。

变差的定义和潜在的来源

可操作的定义

可操作的定义系指人员可以进行其业务的定义。对安全、圆的、可靠的或其它任何质量（特性）的可操作性的定义必须是可互相交流的：对买卖双方都具有相同的含义，对生产工人来说在昨天和今天都具有相同的含义。例如：

1. 对一种材料或一个总成件的特定试验。
2. 判断用的准则
3. 决定：是或者不是；物体或材料符合或不符合准则。[18]

[17] 参见第V章，第C节

[18] W. E. Deming, *"Out of the Crisis"* (1982, 1986), 第 277 页。

标准 (Standard)

标准通常指全体同意作为一个比较的基础，一种可接受的形式。它可以是由权威机构设定或建立的一个*物品*或是一个*整体*（仪器、程序等），如用来度量数量、重量、长度、值或质量的规则。

这整体 (ensemble) 的概念是在 ANSI/ASQC 标准 MI-1996 中正式形成[19]。这术语强调需要考虑所有影响测量不确定度的实情；例如：环境、程序、人员等。“举一个简单的整体例子：校准块规的整体包括了标准块规、比较测定仪、操作者、环境及校准程序。”

参考标准 (Reference Standard)

指在某一指定地点存放的，通常具有最高度量衡的质量的标准。使用这样的标准所做的测量是可以追溯的。

测量和试验设备 (Measurement & Test Equipment, M&TE)

完成一项测量所需要的所有的测量仪器、测量标准、基准材料及辅助设备。

校准标准 (Calibration Standard)

进行日常校准时用来作为一项参考的标准。意图是作为校准工作量和实验室参考标准之间的缓冲。

传递标准 (Transfer Standard)

用来将某一已知数值的标准与即将进行校准的单元进行个别比较的标准。

基准 (Master)

作为在校准过程中的一项标准。也可被描述为参考或校准的标准。

工作标准 (Working Standard)

在实验室内进行日常测量的标准。其不作为校准标准使用，但可能会作为传递标准使用。

---

[19] 该定义后来根据军用标准更新成了测量和试验设备或 M&TE。

为物件选择一个标准应该很慎重。所采用的标准应能够反映该测量系统的用途和范围，以及以时间为基础的变差，如磨损及环境因素（温度，湿度等）。

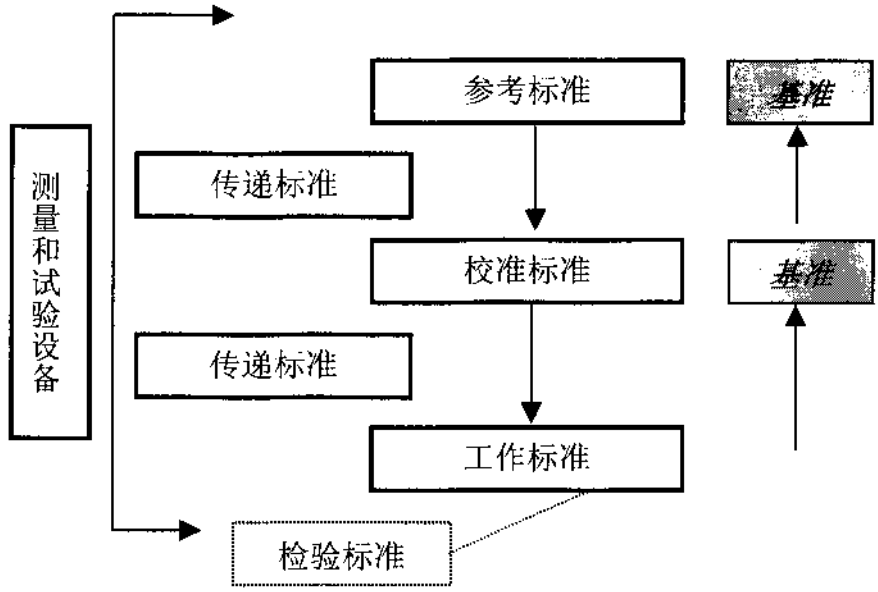


图 3. 不同标准之间的关系

检验标准 (Check Standard)

一个与被指定的测量过程非常相似的测量物品，但它比被评价的测量过程更稳定。

参考值 (Reference Value)

参考值也被称为可接受的参考值或基准值，它是一个被同意用作参考比较的一个物品或整体的值。可接受的参考值取决于以下几个方面：

- 用高一级的测量设备（例如：度量衡实验室或全尺寸检验设备）多次测量的平均值来确定。
- 法定值：由法律定义并强制执行的。
- 理论值：以科学原理为基础
- 指定值：以一些国家或国际组织的实验工作为基础（由健全的理论支持）。
- 一致同意的值：以一些科学或工程组织赞助的合作实验为基础，透过专业人士和商业组织等使用者一致同意其定义的值。

- 协议值：由受影响的各团体所协调一致得到的值。

在所有的情况下，参考值需要以一个可操作的定义为基础，并且由一个可接受的测量系统得到其结果。为了达到这目标，用来确定这参考值的测量系统应该包括：

- 测量仪器与正常评价的测量系统相比，具有较高分辨力和低的测量系统误差。
- 用可追溯到 NIST 或其它 NMI 的标准进行校准。

真值 (True Value)

真值是被测零件的“实际值”。尽管该值不被知道且无法知道，但它是测量系统的目标，所有个别的值都要尽可能地（经济地）与该值接近。不幸的是，永远不可能知道确切的真值。在所有分析中，参考值被作为真值的最佳近似值。由于参考值被用来代替真值，这些术语通常被可交替的使用。不推荐这样使用。[21]

分辨力 (Discrimination)

分辨力是指一测量仪器能够检测并忠实地显示相对于参考值的变化量。通常也被称为可读程度或解析度。

这种测量的能力通常是某测量仪器上最小刻度的值。如果仪器的刻度较“粗略”，则可以使用刻度值的一半作为解析度。

一个通用的比例规则是，测量设备的分辨力要至少等于被测量距离的十分之一，传统上把该距离看成是产品的规范；最近这 10：1 规则的意义被解释为：测量设备要能分辨出过程变差的至少十分之一以上。这与持续改进的构想是一致的。（例如：过程的关注焦点是一顾客指定的目标。）

[21] 参见 ASTM: E177-90a

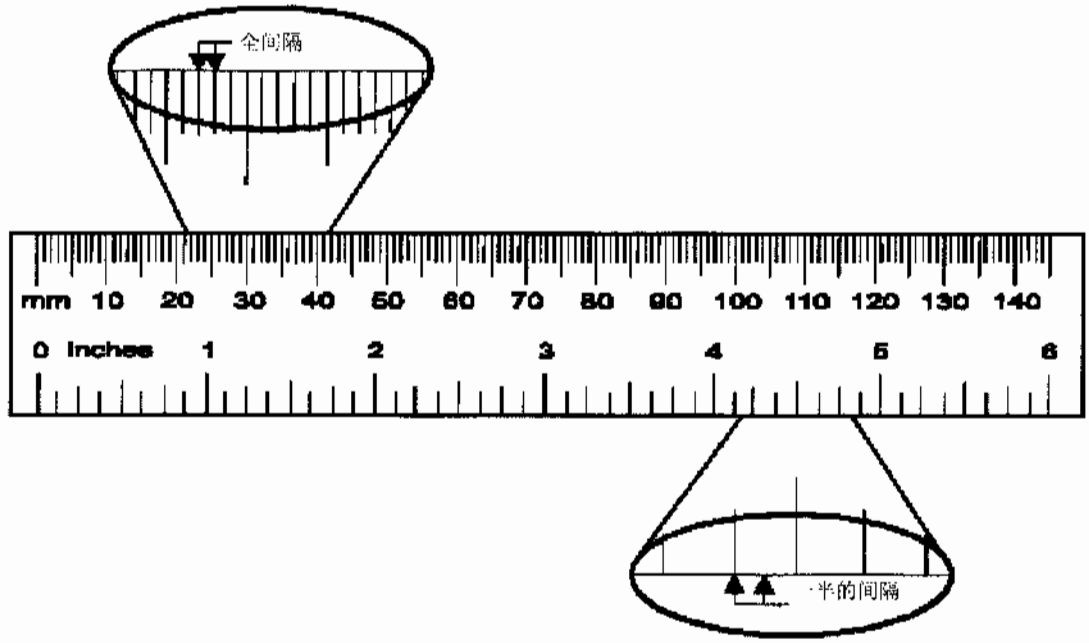


图 4. 分辨力

可以将上述的比例规则视为确定分辨力的一个起始点，因为这还没有包括测量系统变差的任何其它因素。

由于经济和物理上的限制，测量系统将不能识别过程分布成各个独立或不同测量特性的所有零件；取而代之的是以这些测量特性分组成不同的数据种类。在同一数据分类中的所有零件，对其被测特性来说具有相似的数值。

如果测量系统的分辨力（敏感度或有效解析度）不足，它将不是一个合适的系统来识别过程的变差，或对各个零件特性值定量。如果是这样的情形，应该使用更好的测量技术。

如果不能检测到过程的变差，则这分辨力用于分析是不可接受的；如果不能检测出特殊原因变差，它用于控制是不可接受的（见图 5）。

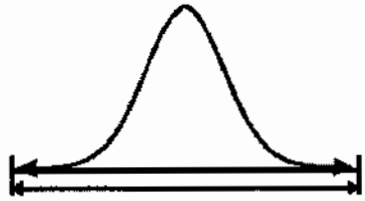
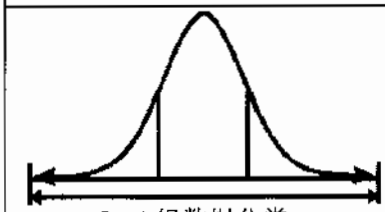
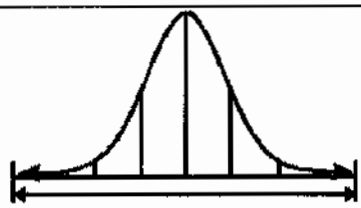
分类数	控制	分析
 1 组数据分类	只有在下列条件下才可用于控制： • 与规范相比，过程变差较小 • 在预期的过程变差上，损失函数很平缓 • 过程变差的主要来源导致了平均值偏移	• 用于过程参数及指数的估计是*不可接受的* • 只能指出过程是否正在产生合格或不合格的零件
 2 - 4 组数据分类	• 依据过程分布，可被用为准计量型控制技术 • 可产生不敏感的计量型控制图	• 用于过程参数及指数的估计一般是不可接受的，因为它仅提供粗劣的估计
 5 或更多组数据分类	• 可用于计量型控制图	• *建议使用*

图 5：过程分布的区别分类数 (Number of Distinct Categories, *ndc*) 对控制与分析活动的影响

不适用的分辨力可能会在极差图 (range chart) 中出现。图 6 是包含了从相同数据中取得的两组控制图。控制图 (a) 显示精确至千分之一英寸的原始测量值；控制图 (b) 显示这些百分之一英寸的原始测量值是四舍五入的。由于人为的紧缩限制，控制图 (b) 显示是失控的；这四舍五入的零极差超越了描述一产品的子组变差的象征。

不合适的分辨力的最佳象征，可由过程变差的 SPC 极差图看出来。尤其是，当极差图中只有一、二或三种可能的极差值在控制限值内时，这种测量正是在分辨力不足情况下进行的结果。另外，如果极差图显示有 4 种可能的极差值在控制限值内且超过四分之一以上的极差值为零，则该测量也是在分辨力不足情况下进行的。

回到图 6，控制图 (b)，我们只看到只有两种可能的极差值在控制限值内（数值 0.00 和 0.01），因此，上述规则正确地识别出失控的原因是分辨力不足（敏感度或有效分辨率）。

当然，可以透过提高这测量的分辨力，从而改变检测子组内变差的能力来补救这一问题。如果一个测量系统的可视解析度比过程变差小，该测量系统将具有足够的分辨力。因此，为得到足够的分辨力，建议可视解析度最多可以到总过程的 6σ 标准偏差的十分之一，以取代传统上设定可视解析度最多为公差范围的十分之一的规则。

最终情况是，能获得一个稳定的、高能力的过程，且该过程使用了一稳定的、在实施技术限制内最好等级的测量系统。有效解析度可能不够，那么对测量系统进行进一步的改进便成为不切实际的。在这些特别情况下，测量计划可能需要使用其它替代的过程监控技术；应该只有具有资格的、熟悉测量系统和过程的技术人员才能作成这种决定并予以文件化。这种情况要求必须有顾客的批准，并且必须在控制计划中予以规定。

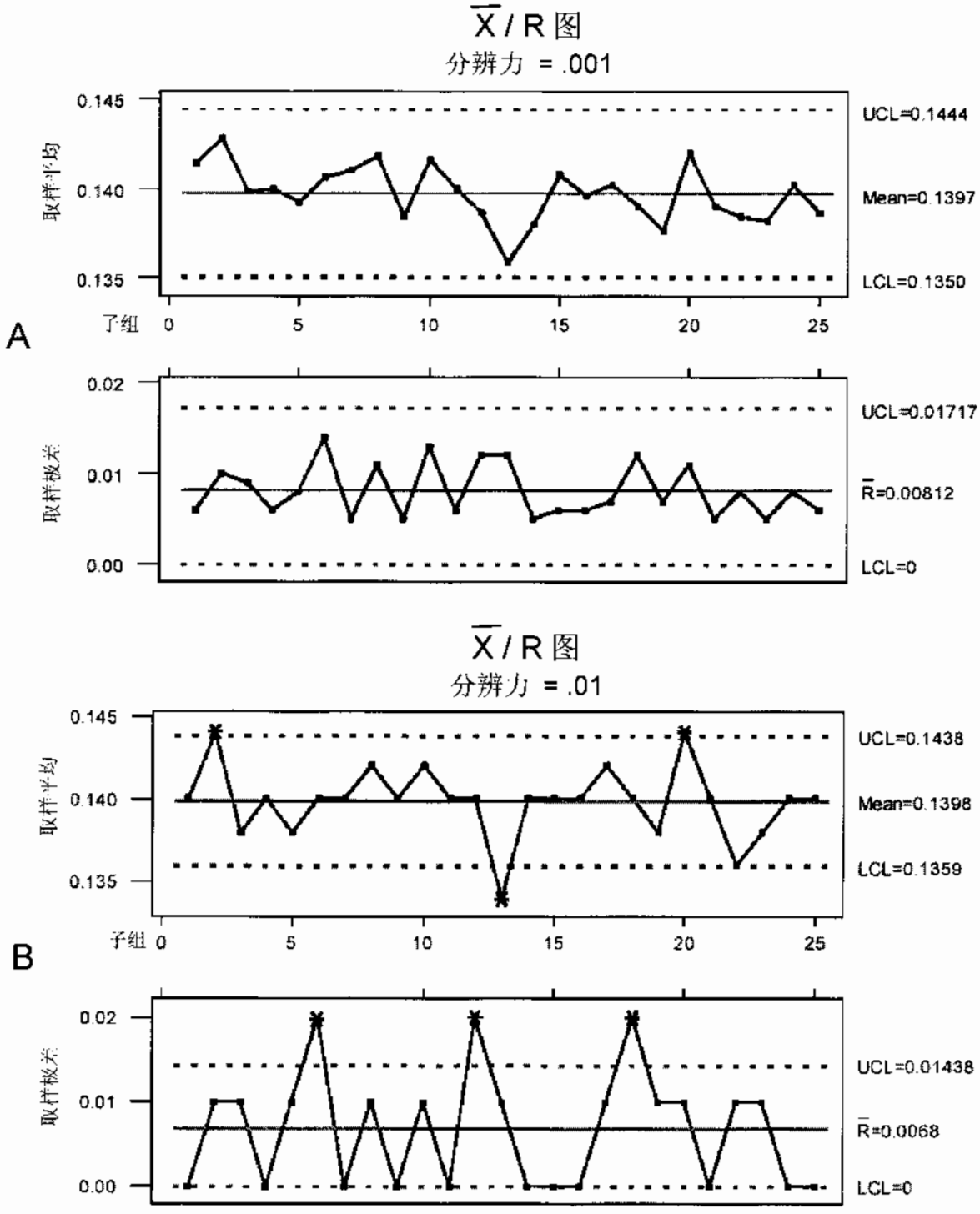


图 6：过程控制图 [22]

[22] 图 6 摘自《*Evaluating The Measurement Process*》，由 Wheeler 和 Lyday 编著， 1989, SPC Press , Inc., Knoxville, Tennessee.

测量过程变差

对于大多数测量过程，其所有的测量变差通常被描述为常态的分布；常态机率是对一个测量系统分析的标准方法的假设。事实上，有些测量系统不是常态分布，当这种情况发生时，而却以常态分布为假设，则 MSA 方法可能会过大地估计测量系统误差。因此，这测量分析必须认清该非常态分布的测量系统，并修改评价方法。

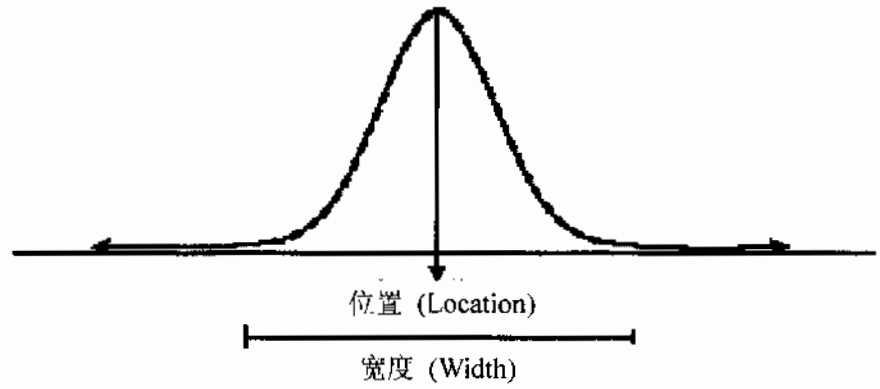


图 7：测量过程变差的特性

位置变差

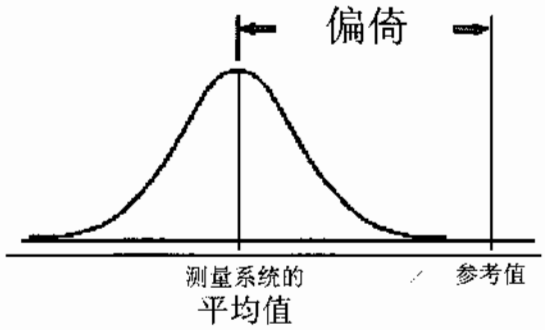
准确度 (Accuracy)

通用的正确说法，它是指一个或多个测量结果的平均值与一个参考值之间一致的接近程度。测量过程必须处于统计受控状态，否则谈过程的准确度没有意义。

在某些组织内，准确度通常与偏倚互用。ISO（国际标准化组织）和 ASTM（美国试验和材料协会）用这术语准确度来涵盖偏倚和重复性。为了避免*准确度*这一术语的混淆使用，ASTM 建议只用*偏倚*来描述位置误差。接下来的内容也将遵循这一原则。

偏倚 (Bias)

偏倚通常被称为“准确度”。由于“准确度”有多种意思，建议不要用准确度来代替偏倚。



偏倚是指对相同零件上同一特性的观测平均值与真值（参考值）的差异。

偏倚是测量系统的系统误差的测量。它是由所有已知或未知的变差来源所共同影响的总偏差所造成，这些促使在某一测量时期内重复的应用相同测量过程时，以有结果恒定的且可预测的呈现偏置(offset)的总偏差趋向。

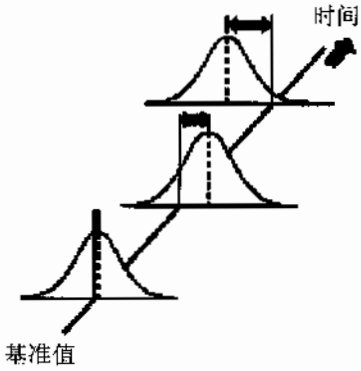
造成过大的偏倚的可能原因有：

- 仪器需要校准
- 仪器、设备或夹具磨损
- 基准的磨损或损坏，基准偏差
- 不适当的校准或使用基准设定
- 仪器质量不良 — 设计或符合性
- 线性误差
- 使用了错误的量具
- 不同的测量方法 — 作业准备、载入、夹紧、技巧
- 测量的特性不对
- 变形（量具或零件）
- 环境 — 温度、湿度、振动、清洁
- 错误的假设，应用的常数不对
- 应用 — 零件数量、位置、操作者技能、疲劳、观测误差（易读性、视差）

在校准过程所使用的测量程序（如：使用“基准”），应该尽可能地与正常操作的测量程序一致。

稳定性 (Stability)

稳定性（或漂移）是指经过一段长期时间下，用相同的测量系统对同一基准或零件的同一特性进行测量所获得的总变差。也就是说，稳定性是整个时间的偏倚变化。

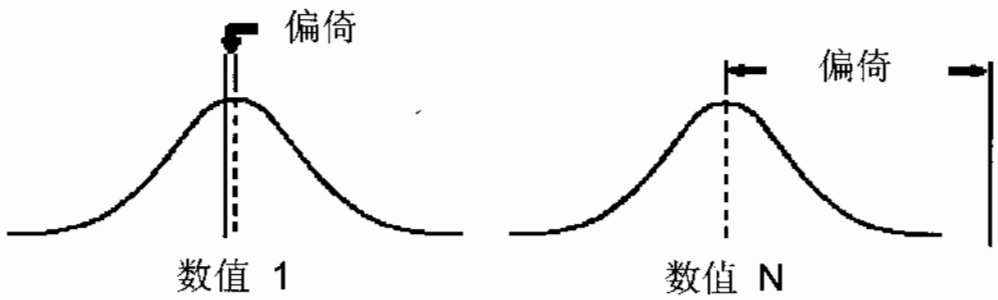


造成不稳定的可能因素有：

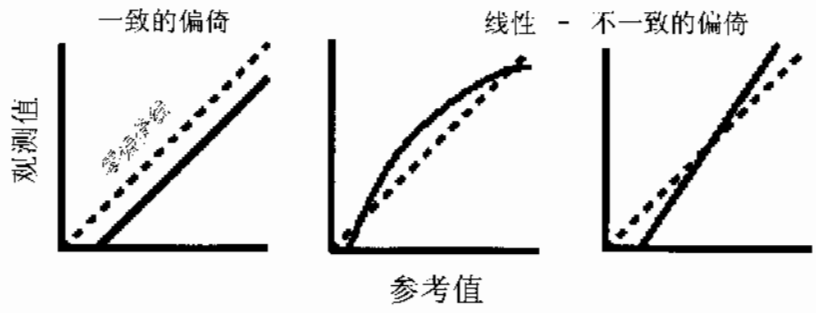
- 仪器需要校准，缩短校准周期
- 仪器、设备或夹具的磨损
- 正常的老化或损坏
- 维护保养不好 — 空气、动力、液体、过滤器、腐蚀、尘土、清洁
- 基准的磨损或损坏，基准的误差
- 不适当的校准或使用基准设定
- 仪器质量不好 — 设计或符合性
- 仪器缺少稳健的设计或方法
- 不同的测量方法 — 作业准备、载入、夹紧、技巧
- 变形（量具或零件）
- 环境变化 — 温度、湿度、振动、清洁
- 错误的假设，应用的常数不对
- 应用 — 零件数量、位置、操作者技能、疲劳、观测误差（易读性、视差）

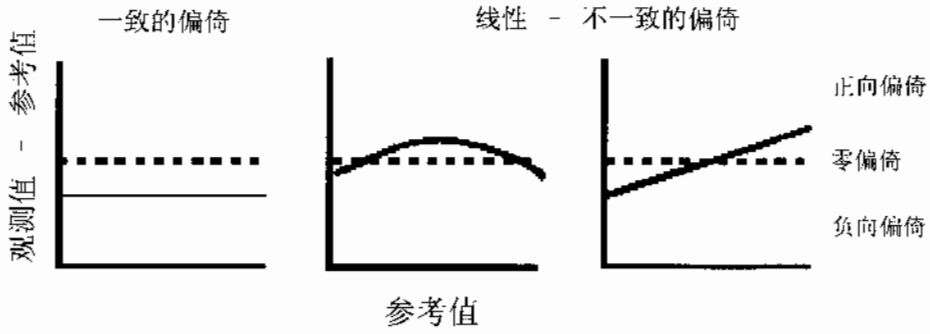
线性 (Linearity)

线性是在测量设备预期的工作（测量）量程内，偏倚值的差异。线性可被视为偏倚对于量程大小不同所发生的变化。



注意：不可接受的线性会增添不同结果。不要假设偏倚是不变的。





造成线性误差的可能原因如下：

- 仪器需要校准，缩短校准周期
- 仪器、设备或夹具的磨损
- 维护保养不好 — 空气、动力、液体、过滤器、腐蚀、尘土、清洁
- 基准的磨损或损坏，基准的误差 — 最小 / 最大

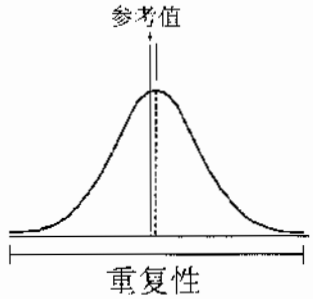
- 不适当的校准（没有涵盖操作范围）或使用基准设定
- 仪器质量不好 — 设计或符合性
- 缺乏稳健的仪器设计或方法
- 应用了错误的量具
- 不同的测量方法 — 作业准备、载入、夹紧、技巧
- 随着测量尺寸不同，（量具或零件）变形量不同
- 环境 — 温度、湿度、振动、清洁
- 错误的假设，应用的常数不对
- 应用 — 零件数量、位置、操作者技能、疲劳、观测误差（易读性、视差）

宽度变差

精确度 (Precision)

传统上，将测量系统整个作业量程范围内（尺寸、范围和时间）的解析度、敏感度和重复性的最终影响定义为精确度。有些组织将精确度与重复性互用。事实上，精确度通常用来描述在测量范围内重复测量的预期变差；这范围可能是尺寸或是时间（例如："某测量设备在高量程和低量程一样精确"或者"今天和昨天一样精确"）。精密度也是对重复性的线性偏倚量（虽然重复性是随机的，而偏倚是系统的误差）。ASTM 定义的精确度的范围更广，包括由于不同读数、量具、人、试验或条件造成的变差。

重复性 (Repeatability)



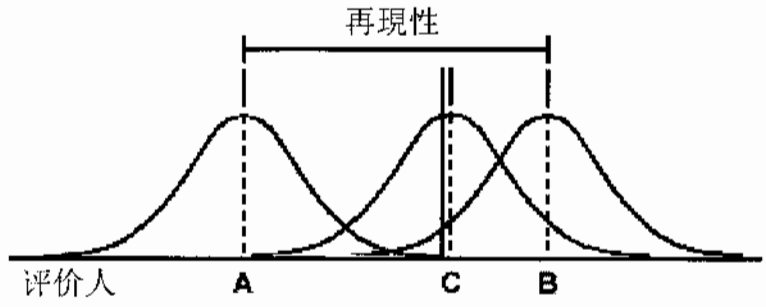
传统上将重复性称为"评价人内部"的变异。重复性是用一个评价人使用相同的测量仪器对同一零件上的同一特性，进行多次测量所得到的测量变差；它是设备本身的固有变差或能力。重复性通常被称为设备变差 (equipment variation, EV)，但这是一种误解；事实上，重复性是在指定的测量条件下连续测量的普通原因（随机误差）的变差。重复性的最佳措辞为：当测量条件已被确定和定义 — 以确定的零件、仪器、标准、方法、操作者、环境和假设之下，系统*内部*的变差。除了设备*内部*的变差之外，重复性还包括在误差模型中的任何条件下的*内部*变差（参见以下内容所述）。

造成重复性的可能原因包括：

- 零件*内部*（抽样样本）：形状、位置、表面光度、锥度、样本的一致性
- 仪器*内部*：维修、磨损、设备或夹具的失效、质量或保养不好
- 标准*内部*：质量、等级、磨损
- 方法*内部*：作业准备、技巧、归零、固定、夹持、点密度的变差。
- 评价人*内部*：技巧、位置、缺乏经验、操作技能或培训、*意识*、*疲劳*
- 环境*内部*：对温度、湿度、振动、清洁的小幅度波动
- 错误的假设 — 稳定，适当的操作
- 缺乏稳健的仪器设计或方法，一致性不好
- 量具误用
- 失真（量具或零件）、缺乏坚固性
- 应用 — 零件数量、位置、观测误差（易读性、视差）

再现性 (Reproducibility)

传统上将再现性称为“评价人之间”的变差。通常将再现性定义为用不同评价人使用相同的测量仪器对同一产品上的同一特性，进行测量所得的平均值的变差。手动仪器通常的确受操作者技能的影响；但是对于操作者不是变差主要原因的测量过程（例如自动的系统），上述说法是不正确的。因为这原因，再现性是指测量的系统*之间*或条件*之间*的平均值变差。



ASTM 的定义比上述定义更广，不仅包括评价人的不同，同时可能还包括：量具、试验室及环境（温度、湿度）的不同，除此之外，在再现性的计算中还包括重复性。

造成再现性误差的潜在原因包括：

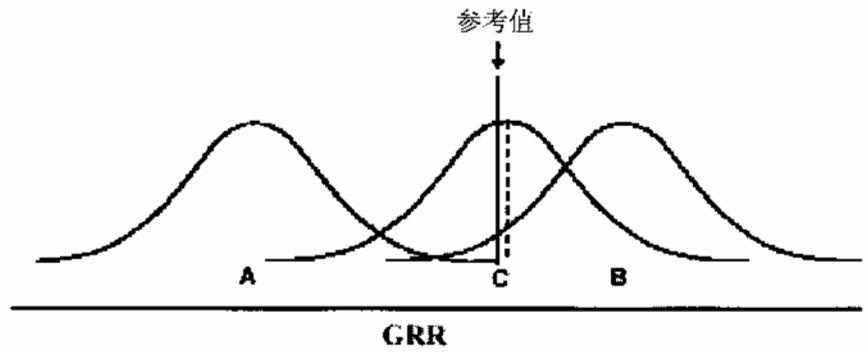
- 零件*之间*（抽样样本）：使用相同的仪器、操作者和方法测量A、B、C零件类型时的平均差异。
- 仪器*之间*：在相同零件、操作者和环境下使用A、B、C仪器测量的平均值差异。注意：在这种情况下，再现性误差通常还混有方法和/或操作者的误差。
- 标准*之间*：在测量过程中，不同的设定标准的平均影响。
- 方法*之间*：由于改变测量点密度、手动或自动系统、归零、固定或夹紧方法等所造成的平均值差异。
- 评价人（操作者）*之间*：评价人A、B、C之间由于培训、技巧、技能和经验所造成的平均值差异。推荐在为产品和过程鉴定和使用手动测量仪器时使用这种研究方法。
- 环境*之间*：在经过1、2、3等时段所进行的测量，由于环境周期所造成的平均值差异。这种研究常用在使用高度自动化测量系统对产品和过程的鉴定。
- 研究中的假设有误。
- 缺乏稳健的仪器设计或方法。
- 操作者培训的有效性。
- 应用 — 零件数量、位置、观测误差（易读性、视差）。

正如以上两个定义中介绍的那样，ASTM 所使用的定义与本手册所使用的定义有区别。ASTM 文献专注的是对室内实验室评价，其集中在实验室与实验室的差别，包括可能不同的操作者、量具和环境，以及实验室内部重复性的影响；因此，他们的定义需要包含这些差异。按照ASTM标准，重复性最好是是在设备当前壮况下（一个操作者、一个量具、短期时间内）；而再现性是当差异原因来自于多个来源时，能呈现更多典型的操作状况。

量具 R&R 或 *GRR*

量具的 R&R 是结合了重复性和再现性变差的估计值。换句话说，GRR 值等于系统内部变差和系统之间变差的和。

$$\sigma^2_{GRR} = \sigma^2_{再现性} + \sigma^2_{重复性}$$

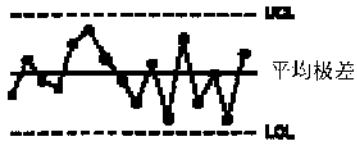


敏感度 (Sensitivity)

敏感度是指能产生一个可检测到（有用的）输出信号的最小输入。它是测量系统对被测特性变化的回应。敏感度由量具设计（分辨力）、固有质量 (OEM)、使用中保养，以及仪器操作条件和标准来确定。它通常被表示为一测量单位。

影响敏感度的因素包括：

- 一个仪器的衰减能力
- 操作者的技能
- 测量装置的重复性
- 对于电子或气动量具，提供无漂移操作的能力
- 仪器使用所处的条件，例如：大气条件、尘土、湿度



一致性 (Consistency)

一致性是系统随着时间变化，测量变差的差值。它可以看成是重复性随时间变化的差值。

影响一致性的因素都是特殊原因变差，包括：

- 零件的温度
- 电子设备必要的预热
- 设备磨损

均一性 (Uniformity)

均一性是量具整个工作量程内变差的差值。它可以看成在不同尺寸值下重复性的同质性（相同性）。

影响均一性的因素包括：

- 由于位置不同，夹具能允许更小 / 更大的尺寸
- 刻度的可读性不够
- 读数的视差

测量系统变差

能力 (Capability)

一测量系统的能力是基于在短期评估中测量系统误差（随机的和系统的）的组合变差的估计值。简单的能力包括下列组成要素：

- 不准确的偏倚或线性
- 重复性或再现性 (*GRR*)，包括短期的一致性

参考第三章对以上典型要素计算的方法和范例。

因此，一测量能力的估计值，表示某测量系统在指定条件、范围和工作量程下的预期误差（不像测量*不确定度*，其表示与测量结果有关的误差或数值的预期范围）。能力是一个组合变差，当测量误差互相不相关时（随机并且是独立的）的能力可由下公式计算得到：

$$\sigma^2_{能力} = \sigma^2_{偏倚（线性）} + \sigma^2_{GRR}$$

理解和正确地应用测量能力要记住以下两点：

首先，能力的估计总是与设定测量 — 条件、量程和时间的范围有关。例如：没有限定测量条件的范围或工作量程，而描述一个 25mm 毫米尺的能力是 0.1mm 是不完整的。再者，这也是为什么一个误差模型对定义一个测量过程来说是如此重要的原因。一估计测量能力的范围，要在有限的部分测量量程范围内或整个的测量量程内，对操作非常确切的或是一般的陈述。短期的意思是：对一系列测量周期的能力、完成 GRR 评价的时间、一特定的生产期间，或者以校准频率的时间表示。一测量能力的表达要尽可能完整，从而能够合理的呈现出测量的条件和范围。可以用文件化的控制计划来说明这意图。

第二，在一个能力估计中要包括测量工作量程范围内的短期一致性和均一性（重复性误差）。对于一个简单的量具，如一把 25mm 的分厘卡，用代表性的、熟练的操作人进行的量具整个量程内的测量重复性，可以看成是一致并且均一的。在本例中，一能力估计可能包括了在一般条件下，对量具整个测量量程内的多种特性。更大量程或更复杂的测量系统（例如：CMM）可能要证明（未修正的）线性、均一性、和在整个量程或尺寸内的短期一致性等测量误差；因为这些误差是互相关联的，他们不能用上面简单的线性公式来组合。当（未修正的）线性、均一性或一致性在整个量程内的变化很大时，测量计划者和分析人员仅有以下两个实用的选择：

1) 报告这测量系统整个指定的条件、范围和量程下最大（最坏情况）的能力，或
2) 确定并报告对测量量程中特定部分的多个能力评估值（例如：低、中、高量程）。

性能 (Performance)

像过程性能一样，测量系统性能是所有重大的和决定性的变差来源的长期总影响。性能是组合了测量误差（随机的和系统的）的长期评估计算。因此，性能包含了以下长期误差要素：

- 能力（短期误差）
- 稳定性和一致性

参考第三章对以上典型要素计算的方法和范例。

一测量性能的估计，表示某测量系统在指定条件、范围和量程下的预期误差（不像测量*不确定性*，其表示与测量结果有关的误差或数值的预期范围）。测量性能为组合变差，当这些测量误差互不相关时（随机并且独立），可由以下公式计算：

$$\sigma^2_{性能} = \sigma^2_{能力} + \sigma^2_{稳定性} + \sigma^2_{一致性}$$

再者，正如短期*能力*一样，长期*性能*总是与设定测量一条件、量程和时间的范围有关。一估计测量性能的范围，要在有限的部分测量量程范围内或整个的测量量程内，对操作非常确切的或是一般的陈述。长期意味着：一定时期内几个能力评估的平均值；由测量控制图计算出的长期误差的平均值、校准纪录的评估或多个线性研究，或者是对测量系统整个生命周期和量程内多次 GRR 研究的平均误差。表达一测量性能时需要尽可能完整，从而可以合理地呈现出测量条件和量程。

在一性能估计中要包括整个测量量程内的长期一致性和均一性（重复性误差）。测量分析人员必须意识到这些误差相互之间的潜在相关性，以免过分地估计这性能；这取决于各个误差要素是如何被确定。当长期（未修正的）线性、一致性和均一性在整个量程内的变化很大差距时，测量计划者和分析者只有以下两个实用的选择：

1) 报告该测量系统整个指定的条件、范围和量程中的最大性能（最坏的情况），或

2) 确定并报告这测量量程中特定部分的几个性能评估值（例如：低、中、高量程）。

不确定度 (Uncertainty)

由 VIM 定义的测量不确定度为："是一个与测量结果有关的参数，其特性是由于被测物 (measurand) 特征所可能合理造成的数值离散。"[24] 详细的解释参见第 一章 F 节。

[24] VIM 将 Measurand 定义为："用来测量的一定量的物体"。

评论

对于测量系统的参数、准确度和精密度，大部份的操作人员很熟悉，因为他们每天都要在技术或销售讨论中用到。但是，这些术语也很容易混淆，因为他们通常以为这些术语是可以互换的。例如，如果一个量具经过独立的代理机构认证它是很准确的，或这设备由其供方保证它是很精确的，而这会被不正确地认为所有的读数会与真值很接近。这说法不只是理解性的错误，更可能导致对产品和过程的错误决定。



对于偏倚和重复性也有同样的混淆（如准确的和精确的测量一样）。了解以下几点是很重要的：

- 偏倚和重复性是相互独立的（见图 8）
- 控制了一些误差来源中的一个原因，不能保证控制了其它的误差原因。因此，测量系统控制计划（传统上被称为量具控制计划）要对所有相关的变差来源进行量化和跟踪。[25]

[25] 参见第一章，第 B 节。

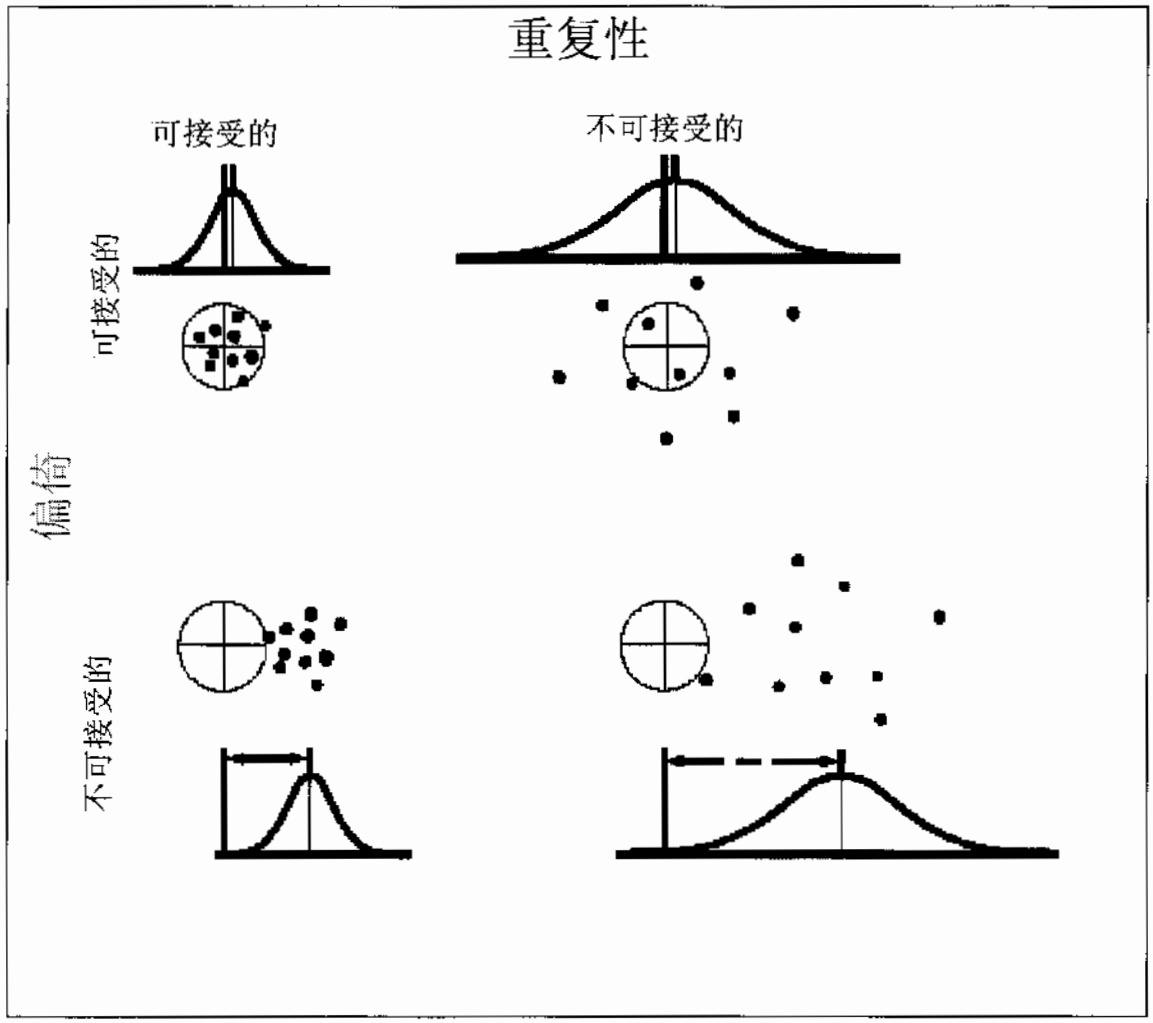


图 8：偏倚和重复性之间的关系

第一章 — 第F节

测量不确定度

总则

*测量不确定度*是国际上用来描述一测量值质量的术语。这术语传统上为许多在度量衡或量具实验室所进行的高准确度测量使用，如 QS-9000 或 ISO/TS 16949 之类的质量系统标准要求："要了解测量不确定度，并与其它任何检验、测量或试验设备所需的测量能力相一致。"[26]

本质上，不确定度是对一指定的测量结果范围的描述，在特定的置信度水准内，预期的包含了*实际*的测量结果的范围。不确定度是测量可靠性的一种量化的表达。这种概念可简单的表达为：

测量*实际值* = 测量的*观测值*（结果）± *U*

U 这术语即是这被测物和测量结果的"扩展的不确定度 (expanded uncertainty)"。扩展的不确定度是组合的标准误差 (u_c) 或这组合误差（随机的和系统的）的标准差，在测量过程中乘以一个涵盖常数 (k)，这是对一期望置信度水准的常态分布曲线面积的表示；记住，应用了对测量过程呈常态分布的原则性假设。ISO/IEC《*测量不确定度指南*》定义这涵盖常数是在常态分布下的 95%面积来充分的呈现不确定度。这通常被诠释为 $k = 2$

$$U = ku_c$$

这组合标准差 (u_c) 包括测量过程中所有重大变差的要素。在许多情况下，依据本手册所实施的测量系统分析方法，可以被视为一种用来对测量不确定度的多种来源的定量工具。通常，可以将这最大误差要素定量为 $\sigma^2_{\text{再现}}$。其他较

[26] QS-9000，第三版，章节 4.11.1。

大的误差来源可以根据测量的用途来应用。表达一不确定度时必须包括一个适当的范围，该范围能够识别所有的重要误差并且允许测量被重复进行。有些不确定度建立于长期测量系统误差，另一些是基于短期测量系统误差。然而，可以通过以下简单公式来计算：

$$u_c^2 = \sigma^2_{性能} + \sigma^2_{其它}$$

测量不确定度仅仅是用来估计在某测量时间下，测量可能的变化有多大，记住这一点很重要。应该考虑测量过程中所有重大的测量变差来源，加上校准、基准标准、方法、环境的重大误差，以及其它在测量过程中没有预先考虑到的误差。在许多情况下，将使用 MSA 和 GRR 方法来计算那些重大的标准误差来估计不确定度。为了确保这不确定度估计值的持续准确性，对一测量过程相关的不确定度要适当周期性地进行再评价。

测量不确定度和 MSA

不确定度和 MSA 的主要不同点在于，MSA 专注于理解某测量过程、确定这过程中误差的大小，并评估这测量系统是否适用于产品和过程的控制；MSA 提升理解和改进（减小变差）。不确定度是测量值的范围、通过一个置信度区间 (confidence interval) 的定义、与测量结果相关，并预期包括测量的真值。

测量的可追溯性

*可追溯性*是通过一个不间断且描述了不确定度的比较链，对一测量特性或一标准数值与规定参考值的比较，这所指的通常是国家或国际标准；所以理解该链中每个环节的测量不确定度是很有必要的。正如在测量过程和追溯性链所介绍的，这包括短期的和长期的测量变差来源；因此测量系统的测量不确定度可被评价，以确保考虑了可追溯性的所有影响；接下来，这也能减少与测量相关的其它问题。

ISO 对表达测量不确定度的指南

测量不确定度表达指南 (Guide to the Expression of Uncertainty of Measurement, GUM) 是如何评价并表达一测量的不确定度的指南。由于它对如何对测量不确定度的来源进行分类和组合，为使用者提供了理论上的解释和整套的指南；这应该被视为一本高水准的参考文件，而不只是一个“如何使用”的手册。它的确给使用者提供了一些更进一步议题的指南，例如：变差来源的*统计独立性 (statistical independence)*、敏感度分析、自由度等，当评价更复杂、多参数的测量系统时，这些议题是很关键的。

第一章 — 第G节

测量问题分析

引言

理解测量变差以及它对整个变差的贡献是需要一个问题解决的基本步骤。当在测量系统中的变差超过了其它任何变差，在对该系统进行其它分析研究之前，分析和解决这些问题便成有必要的工作。在有些情况下，造成测量系统的变差被忽略或忽视；这将会造成时间和资源的浪费，因为注意力被集中在过程本身，而事实上这呈现的变差是由测量装置所造成的。

本节将对问题解决的基本步骤进行探讨，并且说明它们的关系，以理解在一测量系统中存在的问题。每个公司可能都会使用这种顾客已经核准的问题解决过程。

如果测量系统是使用本手册介绍的方法所开发的，那么许多起始步骤将已经存在。例如，早已经存在的因果图可能会对某测量过程提出有价值的经验学习。在任何正式的问题解决开始前，要能收集并评价这些资料。

步骤 1

识别问题

与研究其它任何过程一样，当研究测量系统时，清晰地定义问题或事件是很重要的。对于测量问题，可以从准确度、变差、稳定性等状态来考虑。最重要的是试图从过程变差中，将测量变数及其影响个别分离出来（也就是说要研究过程，而不是研究测量装置）。问题的陈述需要是一个适当可操作的定义，这样任何人都能理解并能解决这一问题。

步骤 2

识别小组

在这种情况下，问题解决小组将取决于测量系统及问题的复杂程度。一个简单的测量系统也许只需要几个人。但是，如果系统和问题很复杂，小组成员数也会增加（最多的小组人数要被限制在 10 人以下）。小组成员以及它们代表的功能需要在这问题解决单上予以识别。

步骤 3

测量系统和过程的流程图

小组将要评审所有过去的测量系统和过程的流程图。这样，可以促进对测量和与其有关的过程的已知和未知信息的讨论。过程流程图的分析可能会识别出需要加入其它人员到这分析小组。

步骤 4

因果图

小组将要评审在测量系统中所有过去的因果图。在某些情况下，这种评审可能会产生解决或部分解决问题的方法，这种评审还会促进对已知和未知信息的讨论。小组将使用这内容的知识初步的识别出对问题影响最大的那些变差。可能会执行额外的研究以证实这些决定。

步骤 5



计划 — 实施 — 研究 — 行动 (PDSA)[27]

随后将采用 PDSA 循环，这一科学的研究方法的形式。计划实验、收集数据、建立稳定性、提出假设并证明，如此循环直到得到一个适宜的解决方法。

步骤 6

可能的解决方法及纠正的证明

步骤和解决方法被文件化，以纪录这决定。进行初始研究来确认解决方法。可以使用一些实验设计的方法来进行解决方法的确认。另外，随着时间发展，还可能要进行额外的研究，包括环境和材料变差等。

步骤 7

将变更制度化

将最终解决方法记录在报告中，然后由适当的部门或功能对过程进行适当改变，从而防止问题再发生。这样可能需要对程序、标准及培训材料作变更。这是整个问题解决过程中最重要的步骤。许多事件和问题在其它的时间会重复发生。

[27] W. E. Deming, "*The New Economics for Industry, Government, Education*". The MIT Press, 1994, 2000

第二章

用于评估测量系统的基本概念

第二章 — 第A节

背景

引言

需要对两个重要领域进行评估：

1) 验证在适合的特性位置上测量了正确的变数。如适用，验证夹具和固定装置，同时识别与测量相互依赖的任何关键环境因素。如果一直在测量错误的变数，那么无论该测量系统多么准确或多么精确，都是在浪费资源，无任何益处。

2) 确定该测量系统需要具备哪些可被接受的统计特征。为了作成这确定，知道将使用什么样的数据是很重要的，因为如果不了解这一点，就不能确定适当的统计特性。在确定了统计特性后，必须对测量系统进行评估，以确定实际上该测量系统是否具有这些特性。

第1&2阶段

理解测量过程，确定它是否满足要求？

第1阶段 试验是验证测量系统是否按照其设计规范要求，在适当的特性位置上测量正确的变数；（若适用，要验证夹具和固定装置）此外验证是否存在任何与测量相互依赖的重要的环境问题。第1阶段可使用一统计的设计实验来评估操作环境对测量系统参数（例如：偏移、线性、重复性和再现性）的影响。第1阶段试验的结果可能说明操作环境不会对整个测量系统变差产生重大的影响。另外，与重复性和再现性要素相比较，测量装置的偏倚和线性的影响应该较小。

从第1阶段试验中所得到的知识应该被用为开发测量系统维护计划的输入，同时应该被用在第2阶段试验类型的输入。环境问题可能驱使位置的改进，或对测量装置的环境控制。

例如，如果在整个测量系统变差中，重复性和再现性的影响很大，那么在第2阶段试验中可能要周期性简单的进行这两因素的统计实验。

随着时间推移，测量系统是否能持续满足要求？

第2阶段 试验是对变差的主要原因提供持续的监控，从而说明测量系统是持续可信的（以及所产生的数据）和/或随着时间推移，测量系统是否出现变坏的信号。

第二章 第B节

选择 / 开发试验程序

"如果理解和遵守技术的极限，则任何技术都可能是有用的。"[28]

有许多适当的程序可用于评估测量系统。选择使用哪些程序取决于许多因素，其中许多因素需要基于被评估的每个测量系统的个案来逐一确定。在某些情况下，为确定一个程序对一特定的测量系统是否合适，可能需要预先的试验。这种预先试验要是上一章节讨论中的第一阶段试验的整个部分。

当选择或开发一个评估程序时，一般要考虑的问题包括：

- 试验中是否应该使用诸如那些可追溯至 NIST 的标准？如果是，什么等级的标准是合适的？标准经常是评估一个测量系统的准确度所必要的。如果不使用标准，该测量系统的变差性仍有可能被评估，但不大可能按合理的可信度去评估该系统的准确度；缺乏这样的可信度可能是一个问题，例如，试图解决一个生产者的测量系统和一个顾客的测量系统之间明显的差别。
- 对于在第二阶段持续的试验，应考虑使用盲测 (blind measurement)。盲测是指在实际测量环境下，由一事先不知正在对该测量系统进行评估的操作者所获得的测量结果；通过适当的管理，根据盲测得到的试验结果通常不受众所周知的霍桑效应 (Hawthorne effect) [29] 所干扰。
- 试验的成本。
- 试验所需要的时间。
- 任何没有被普遍接受的术语应该被可操作的定义。这些术语包括如准确度、精确度、重复性和再现性等。

---

28 W. E. Deming, *"The logic of Evaluation"*, The Handbook of Evluation Research, Vol. 1, 由 Elmer L. Struening 与 Marcia Guttentag 编著

29 "霍桑效应"引用了 1924 年 11 月到 1932 年 8 月之间在 Hawthorne Works of Western Electric 所进行一系列工业实验的结果。在这实验中，研究人员有系统的修改了 5 位组装工人的工作条件并监控其结果。根据条件的改善，生产也提升了；然而，当工作条件被降低，生产仍持续的提升。这被认为：是这些工人独自开发了一更积极于工作的态度的结果，仅仅这部份他们对研究的结果，更胜于改变工作条件结果。参见*"A History of the Hawthorne Experiments"*, Richard Gillespie 著, Cambridge University Press, New York, 1991.

- 由这个测量系统取得的测量结果是否要与另外一个测量系统得到的测量结果比对？如果要比较，使用的程序应该考虑依赖使用诸如前面第 1 阶段中所讨论的标准。如果不使用标准，仍有可能确定这两个测量系统是否可以同时一起正常工作；然而，如果两个系统不能一起正常工作，那么不用标准，就不可能确定哪个系统需要改进。
- 第 2 阶段试验应该每隔多久进行一次？这个问题可能基于个别测量系统的统计特性和对该设施后续结果来决定；以及以该设施进行制造过程，但事实上由于一个测量系统没有正常工作而未受监控的制造过程顾客来决定。

除了这些一般性的问题外，正被试验的特殊测量系统的其它特定问题也可能是重要的。发现一特殊测量系统的具体问题是第 1 阶段试验重要的两个目的之一。

第二章 — 第C节

测量系统研究的准备

如任何的研究或分析，进行测量系统研究之前要进行充分足够的计划和准备。进行测量系统研究的准备一般包括：

1) 应该计划所使用的方法。例如：如果在校准或使用仪器时有评价人的影响，使用工程判断、目视观察、或量具研究来确定。有些测量系统可以忽略再现性的影响；例如，只需按一下按钮，测量结果就能打印出来的测量系统。

2) 应该事先确定评价人的人数、抽样零件的数量，及重复读数的数量等。选择以上内容时要考虑以下因素：

   (a) 尺寸的关键性 — 关键尺寸需要更多的零件和/或测量次数；原因是量具研究估计用的自信程度需求。

   (b) 零件的形态 — 大型或重型零件可能意味着样本较少但测量次数多。

3) 由于目的是评价全部的测量系统，评价人应该从那些正常操作该仪器的人员中选择。

4) 样件的选择对适当的分析是很关键的；其完全的取决于 MSA 研究的设计、测量系统的目的，及代表该生产过程的零件样本的可获得性。

对产品控制的情况下，当测量结果与决定准则是确定"符合或不符合某特性的规范"（如：100%检验或抽样），样品（或标准）必须被选择，但不需要包括整个过程范围。测量系统的评估是以特性公差为基础（如对公差的%*GRR*）。

在过程研究情况下，当测量结果与决定准则是确定"过程稳定性、方向以及是否符合自然的过程变差"（如：SPC、过程监控、能力及过程改进），在整个作业过程范围的样本可获得性变得非常重要。当评估一测量系统对过程控制的适用性时（如对过程变差的%*GRR*），推荐采用过程变差的独立估计法（过程能力研究）。

当不能使用为对过程变差的独立估算法时，或为了确定过程方向和测量系统对过程控制的持续适宜性时，必须从过程中选择样件，并且该样件能代表整个生产作业范围。使用 MSA 研究中所选择的抽样零件的变差 (*PV*) 来计算这研究的总变差 (*TV*)。这 *TV* 指数（如对 *TV* 的*%GRR*）是对过程的方向以及对过程控制的测量系统持续适宜性的指标值。如果抽样零件不能代表生产过程，在评估时必须忽略 *TV*。忽略 *TV* 不会影响使用公差的评估（产品控制），或一个过程变差的独立估算法（过程控制）。

零件抽样的选择可以是在许多天中每天抽一个样品。再者，这是必要的，因为如果零件代表过程中生产变差的范围，这些零件将用来作为分析。由于每个零件要被测量数次，因此，必须对每个零件进行编号作为标识。

5) 仪器应该有足够的分辨力，特性直接被读出的值至少是预期过程变差的十分之一。例如，如果特性的变差为 0.001，则设备应该能被"读出"0.0001 的变化量。

6) 确保测量方法（如评价人和仪器）是测量特性的尺寸，并遵循已定义的测量程序。

进行研究的方式是非常重要的。本手册介绍的所有分析方法均假设单个读值之间具有统计上的独立性 [30]。为了减少误导结果的可能性，需要采取以下步骤：



1) 测量应该以一随机次序 [31] 进行，从而确保可能发生的漂移或变化将随机地分布在整个研究中。评价人应该不能察觉正在检查的零件编号，从而避免任何可能因了解所造成的偏倚。但是，主导研究的人应该知道正在检查哪个零件，并依次记下测量数据。即：评价人 A，零件 1，第一次测量；评价人 B，零件 4，第二次测量等。

2) 对设备的读值，应该按测量仪器解析度的实际极限纪录这测量值。机械式测量装置必需读到并纪录解析度刻度的最小单位；对于电子读值，测量计划中必须规定共同原则以纪录最准确的有意义显示数值。

---

[30] 在读值之间没有相互关系。

[31] 参见第三章，第 B 节："随机化和统计的独立性"。

对于类比式装置，应该纪录最小刻度或灵敏度与分辨率极限值的一半。对于类比式装置，如果最小刻度值为 0.0001 吋，则测量结果应该被记录到 0.00005 吋。

3) 这研究工作应该由了解实施一个可靠研究的重要性的人员负责管理和监督。

在发展第 1 阶段和第 2 阶段试验计划时，需考虑以下数个因素：

- 评价者对这测量过程有哪些影响？如有可能，研究中应该包括正常使用该测量装置的评价者。

  每个评价者都应该使用程序 — 包括所有步骤 — 他们通常用来获取读值。评价者所使用的方法之间的不同将影响测量系统的再现性。

- 评价者对测量系统的校准是否可能是一重大的变差因素？如果是这样，评价者在每个读值分组之前应重新校准设备。

- 需要多少个取样零件和重复几次读值？所要求的零件的数量将取决于被测特性的重要性，以及测量系统变差的估计中所需要的自信度水准。

  当使用本手册推荐的方法进行研究时，尽管评价者人数、测量次数和零件数量可能不同，在第一阶段试验和第二阶段试验计划之间、或为同一个测量系统的连续第二阶段试验之间，通常保留使用相同的评价者人数、试验次数和零件数量。在试验计划之间及连续的试验之间保持一致，将提高在不同的试验结果之间的可比较性。

第二章 — 第D节

结果分析

结果应该被评价，从而确定这测量装置对其预期的应用来说是否为可接受。在进行任何有效的分析之前，测量系统应该稳定。

位置误差

接受性准则 — 位置误差

位置误差通常被定义为偏倚和线性的分析。

通常，如果一测量系统的偏倚或线性误差明显的异于零、或超过该量具校准程序中所建立的最大容许误差，则该测量系统为不可接受。在这种情况下，应该对这测量系统进行重新校准或进行一个补偿修正来减少这误差。

宽度误差

接受性准则 — 宽度误差

一测量系统的变差是否为满意的准则，取决于测量系统变差对制造过程误差或零件公差所占的百分比。对特定的测量系统的最终接受准则取决于测量系统的环境和目的，并且应该经顾客同意。

对于目的是用来分析一过程的测量系统，对测量系统可接受性的通用比例原则如下：

- 低于 10%的误差 — 通常被认为是一个可接受的测量系统。
- 10%到 30%的误差 — 根据应用的重要性、测量装置的成本、维修费用等，可能是可接受的。
- 大于 30%误差 — 考虑为不可接受 — 应该尽各种力量以改进这测量系统。

另外，由测量系统对过程进行划分的区别分类数 (*ndc*)[32] 要能大于或等于 5。

一测量系统的最终可接受性应该不仅仅取决于一些简单的指数，应该同时使用随时间变化的图表来分析测量系统的长期性能。

[32] 参见第三章第 B 节，结果分析 — 数字法。

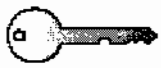
第三章

对简单测量系统的推荐实施方法

第三章 — 第A节

试验程序范例

引言



本章将介绍一些特定的试验程序范例。这些程序简单实用并可容易的使用于一生产环境。如前面所讨论的那样，试验程序应该被用来理解一测量系统并量化其变差，其取决于那些可能影响测量系统的变差来源。在许多情况下，主要的变差来源是由于仪器（量具／设备）、人（评价者）和方法（测量程序）造成的。本章介绍的试验程序是充分的对这种类型的测量系统分析。

当满足以下要求时，能适当地实用这些程序：

- ✓ 只研究两个因素，或测量的条件（即评价者和零件）加上测量系统重复性。
- ✓ 忽略每个零件内部变差的影响。
- ✓ 在评价人和零件之间没有统计上的相互影响。
- ✓ 在研究期间，零件的尺寸不变。

可以实施统计的实验设计和／或使用专业知识 (subject matter knowledge) 以确定这些程序是否适用于任何特定的测量系统。

第三章 — 第B节

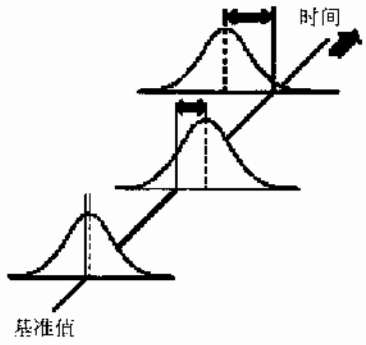
计量型测量系统研究 — 指南

引言

本节包含了在第一章第 E 节所述的测量系统技术的实施指南。建议在适当的应用这些指南之前，彻底复习一遍第 E 节的内容。

用于确定稳定性的指南

进行研究



1) 取得一样件并建立其可追溯到相关标准的参考值。如果无法取得这样的样件，则选择一件落在生产测量范围中间的生产零件[33]，指定它为基准样件以进行稳定性分析。跟踪测量系统稳定性时，不要求该已知的参考值。

   这也许会希望拥有位于预期测量结果的下限、上限和中间位置的基准件。推荐对每种基准件单独的进行测量和画控制图。

2) 以一定的周期基础（每天、每周）测量基准件三到五次。抽样的数量和频率应该取决于对测量系统的认识；可能考虑的因素包括要求重新校准或维修的频率如何、使用测量系统的频率，以及操作条件的重要性等。应该在不同时间下取得多次读值，以代表测量系统的实际使用情况。这将考虑了在一天之中因为热机、周遭或其它因素可能发生的变化。

3) 将数据按时间顺序画在 $\overline{X}$ &R 或 $\overline{X}$ & s 控制图上。

结果分析 — 图示法

4) 建立控制限，使用控制图分析法来评价是否有不受控或不稳定的情况。

[33] 应该注意，由于使用、材料及搬运的原因可能造成生产基准零件遭受过度的磨损。这将要求对该生产零件进行某些改变，例如电镀，以延长这基准零件的寿命。

结果分析 — 数值法

除了通常的控制图分析法外，没有用于稳定性分析的特定数值分析或指数。[34]

如果该测量系统为稳定，这些数值可用来确定测量系统的偏倚。

另外，测量的标准差可作为该测量系统重复性的一个近似值；重复性可以与过程的标准差相比较，以确定这测量系统重复性对这应用是适合的。

使用实验设计或其它问题解决的分析技术可能是必要的，以确定测量系统缺乏稳定性的主要原因。

范例 — 稳定性

为了确定某一新测量仪器的稳定性是否为可接受，过程小组选取了生产过程输出范围中接近中间值的一个零件。该零件被送到了测量实验室，经测量其参考值确定为 6.01。小组每班测量该零件 5 次，共测了四周（20 个子组）；收集到所有数据后，画出了 $\overline{X}$ &R 图（见图 9）

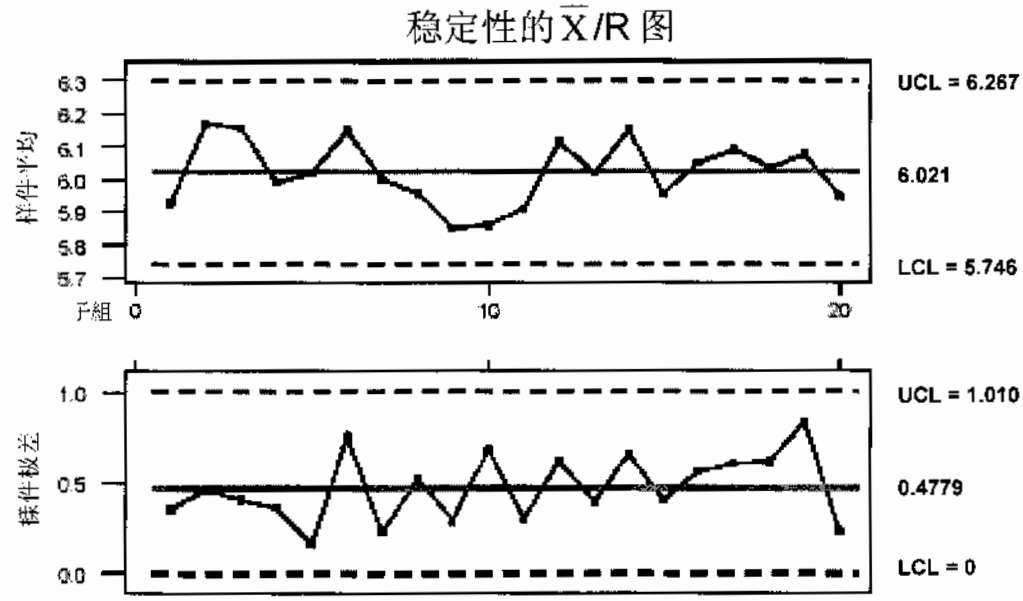


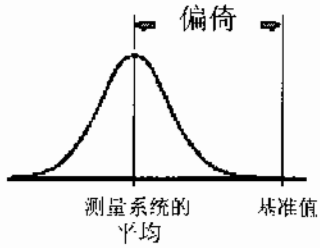
图 9：用于稳定性分析的控制图

控制图分析表明该测量过程处于稳定状态，因为没有明显可见的特殊原因结果发生。

[34] *参见 SPC 参考手册。*

确定偏倚[35]的指南 — 独立样件法

进行研究



1) 取得一个样件，并且建立其与可追溯到相关标准的参考值。如果不能得到这参考值，选择一件落在生产测量范围中间的生产件，并将它指定为偏倚分析的基准件。在工具室里测量该零件 $n \geq 10$ 次，并计算这 n 个读值的平均值。将该平均值视为"参考值"。

   也许会希望拥有位于期望测量结果的下限、中间及上限位置的基准样件，如果可做到这样，可以使用线性研究来分析这些数据。

2) 让一个评价者以正常方式测量样件 $n \geq 10$ 次。

结果分析 — 图示法

3) 画出这些数据相对于参考值的直方图。使用专业知识评审这直方图，从而确定是否存在任何特殊原因或异常点。如果不存在，继续分析。当 $n < 30$ 时，对任何的解释或分析时，要能特别注意。

结果分析 — 数值法

4) 计算 n 个读值的平均值

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

5) 计算重复性标准差（参见下面的量具研究，极差法）

$$\sigma_{\text{重复性}} = \frac{\text{最大值}(X_i) - \text{最小值}(X_i)}{d_2^*}$$

公式中 d_2^* 可从附录 C 查到
取，$g = 1$，且 $m = n$。

[35] 参见第一章第 E 节，潜在原因的可操作性定义及讨论。

如果已获得 GRR 研究（且为有效），重复性标准差计算应该取决于该研究结果。

6) 确定偏倚的 *t* 统计值 (*t*-statistic)[36]

偏倚 = 观测到的平均测量值 − 参考值[37]

$$\sigma_b = \sigma_r / \sqrt{n}$$

$$t = \frac{\text{偏倚}}{\sigma_b}$$

7) 如果 0 落在偏倚值附近的 1-α 自信度界线内，则偏倚在 α 水准上是可接受的。

$$\text{偏倚} - \left[\frac{d_2\sigma_b}{d_2^*}\left(t_{v,1-\alpha/2}\right)\right] \le 0 \le \text{偏倚} + \left[\frac{d_2\sigma_b}{d_2^*}\left(t_{v,1-\alpha/2}\right)\right]$$

公式中，d_2、d_2^*及 v 可以在附录 C 中查到
取 $g = 1$ 且 $m = n$ 並且
$t_{v,1-\alpha/2}$可以利用标准 *t* 分布表中查到。

所使用的 α 水准取决于敏感度的水准，敏感度水准对评价 / 控制过程是必要的，并且与产品 / 过程的损失函数（敏感度曲线）有关。如果 α 置信度水准不是使用预设值 0.05（95%置信度），则应该得到顾客的同意。

偏倚 — 范例

一个制造工程师评价了一个用于过程监控的新测量系统。测量设备的一项分析证明该测量系统应该没有线性误差的考量，所以该工程师只需对测量系统的偏倚进行评价。他基于一份已文件化的过程变差描述，在这测量系统操作范围内选取了一个零件；通过对该零件进行了全尺寸测量来确定它的参考值，然后由主要操作者测量该零件 15 次。

---

[36] 本方法用极差的平均值来估这计测量过程的标准差。由于这极差的使用效率随取样数量的增大而加速，如果取样总数大于 20，则建议将所有样本分散成多个子组并使用控制图法，或以传统的一个样本 *t* 试验，用来源平均平方 *(root mean squared, RMS)* 来计算标准差 。

[37] 偏倚的不确定度由 σ_b 产生。

		参考值 = 6.00	偏倚
测量次数	1	5.8	-0.2
	2	5.7	-0.3
	3	5.9	-0.1
	4	5.9	-0.1
	5	6.0	0.0
	6	6.1	0.1
	7	6.0	0.0
	8	6.1	0.1
	9	6.4	0.4
	10	6.3	0.3
	11	6.0	0.0
	12	6.1	0.1
	13	6.2	0.2
	14	5.6	-0.4
	15	6.0	0.0

表 2：偏倚研究用数据

通过使用散布图和统计软件，检验员得到了直方图和数值分析结果（参见图 10 和表 3）

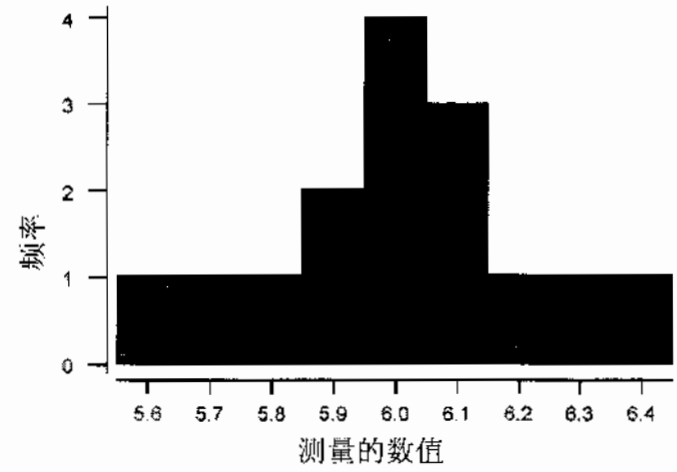


图 10：偏倚研究 — 偏倚研究的直方图

由于 0 落在偏倚自信度区间内 (-0.1185，0.1319)，该工程师可假设这测量的偏倚是可接受的，即在实际使用时，将不会带来额外的变差来源。

	n (m)	平均值，$\overline{\overline{X}}$	标准差，σ_r	平均值的标准误差，σ_b
测量值	15	6.0067	.22514	.05813

	参考值=6.00， α = .05， g = 1， d_2^* = 3.55					
	统计的 t 值	df	显著的 t 值（2-有尾数的）	偏倚	偏倚的95%置信度区间	
					下限	上限
测量值	.1153	10.8	2.206	.0067	-.1185	.1319

表 3：偏倚研究 — 偏倚研究的分析

确定偏倚的指南 — 控制图法

进行研究

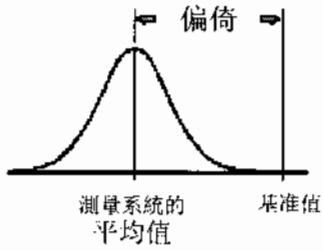
如果用一$\overline{X}$&R 图或用$\overline{X}$& s 图来衡量稳定性，其数据也可以用来进行偏倚的评价。在偏倚被评价之前，控制图分析应该表明这测量系统处于稳定状态。

1) 取得一个样件，并且建立可追溯到一相关标准的参考值。如果不能取得参考值，选择一件落在生产测量值范围中间的生产零件，将它指定为偏倚分析的基准样件。在工具室里测量该零件 $n \geq 10$ 次，并计算 n 个读值的平均值。使用该平均值作为"参考值"。

结果分析 — 图示法

2) 画出这些数据相对于参考值的直方图。对直方图进行评审，以专业知识确定是否存在特殊原因变差或出现异常。如果没有，继续这项分析。

结果分析 — 数值法



3) 从控制图上获取平均值 $\overline{\overline{X}}$

4) 用平均值 $\overline{\overline{X}}$ 减去参考值，计算得到偏倚

$$偏倚 = \overline{\overline{X}} - 参考值$$

5) 用极差的平均值来计算重复性标准差。

$$\sigma_{重复性} = \frac{\overline{R}}{d_2^*}$$

公式中，d_2^* 取决于子组的多少 (m) 和在控制图中的子组中的数量 (g)。（见附录 C）

6) 确定对偏倚的统计 t 值[38]：

$$\sigma_b = \sigma_r / \sqrt{g}$$

$$t = \frac{偏倚}{\sigma_b}$$

7) 若 0 落在偏倚值附近的 1-α 置信区间内，则偏倚在这 α 水准上是可接受的。

$$偏倚 - \left[\frac{d_2\sigma_b}{d_2^*}\left(t_{v,1-\alpha/2}\right)\right] \le 0 \le 偏倚 + \left[\frac{d_2\sigma_b}{d_2^*}\left(t_{v,1-\alpha}\right.\right.$$

公式中，d_2、d_2^* 及 v 可从附录 C 中查到 $t_{v,1-\alpha/2}$ 可以用标准 t 分布表中查到。

所使用的 α 水准取决于敏感度的水准，敏感度水准对评价 / 控制一过程是必要的，并且与产品 / 过程的损失函数（敏感度曲线）有关。如果 α 置信度水准不是使用预设值 0.05（95%置信度），则应该得到顾客的同意。

[38] 偏倚的不确定度由 σ_b 决定。

范例 — 偏倚

参见图 9，对一个参考值为 6.01 的零件进行了稳定性研究，所有样本（20 个子组）的整体平均值为 6.021；因此偏倚的计算值为 0.011。

使用散布图和统计软件，检查员得到了数值分析结果（见表 4）。

由于 0 落在这偏倚的置信度区间内 (-0.0800，0.1020)，该过程小组可以假设这测量系统的偏倚是可接受的，即在实际使用中将不会带来额外的变差来源。

	n	平均值，$\bar{\bar{X}}$	标准差，σ_r	平均值的标准误差，σ_b
测量值	100	6.021	.2048	.0458

	参考值=6.01，α = .05，m = 5，g = 20，d_2^* = 2.334，d_2 = 2.326					
	统计的 t 值	df	显著的 t 值（2-含尾数的）	偏倚	偏倚的 95%置信度区间	
					下限	上限
测量值	.2402	72.7	1.993	.011	-.0800	.1020

表 4：偏倚研究 — 对偏倚的稳定性研究的分析

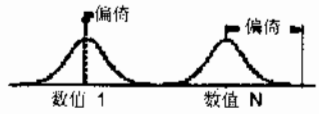
偏倚研究的分析

如果偏倚在统计上不等于 0，检查是否存在以下原因：

- 基准件或参考值有误，检查确定标准件的程序。
- 仪器磨损。这问题会在稳定性分析中呈现出来，建议进行维修或重新整修计划。
- 仪器产生错误的尺寸。
- 仪器所测量的特性有误。
- 仪器没有经过适当的校准。对校准程序进行评审。
- 评价者使用仪器的方法不正确。对测量指导书进行评审。
- 仪器纠正的指令错误

如果测量系统的偏倚不等于零。若有可能，应该采用硬体修正法、软件修正法或同时使用这两种方法来对量具进行重新校准以达到零偏倚。如果偏倚不能调整到零，通过变更程序（如：对每个读值根据偏倚进行修正）还可以继续使用该测量系统。由于存在评价人误差这一高度风险，因此这种方法只能在取得顾客同意后方能使用。

确定线性的指南[39]



进行研究

可使用以下指南进行线性评价：

1) 由于存在过程变差，选择 $g \geq 5$ 个零件，使这测量涵盖这量具的整个工作量程。

2) 对每个零件进行全尺寸测量，从而确定其参考值，并确定涵盖了这量具的工作量程。

3) 让经常使用该量具的操作者测量每个零件 $m \geq 10$ 次。

   ✓ 要*随机*的选择零件，从而减少评价人对测量中偏倚的“记忆”。

结果分析 — 图示法

4) 计算零件每次测量的偏倚，以及每个零件的偏倚平均值。

$$偏倚_{i,j} = X_{i,j} - (参考值)_i$$

$$\overline{偏倚_i} = \frac{\sum_{j=1}^{m} 偏倚_{i,j}}{m}$$

5) 在线性图上画出相对于参考值的每个偏倚及偏倚平均值。（参见图 11）

6) 应用以下公式，计算并画出最适合的线及该线的自信度区间。

   对最适合的线，用公式：$\overline{y_i} = ax_i + b$

   式中

   $x_i =$ *参考值*，

   $\overline{y_i} =$ *偏倚平均值*

   以及

[39] 参见第一章第 E 节，对潜在原因的可操作性定义及讨论。

$$a = \frac{\sum xy - \left(\frac{1}{gm}\sum x \sum y\right)}{\sum x^2 - \frac{1}{gm}\left(\sum x\right)^2} = \text{斜率 (slope)}$$

$$b = \bar{\bar{y}} - a\bar{x} = \text{中心 (intercept)}$$

对于一个已知的 x_0，α 置信度区间[40]为：

$$\text{式中 } s = \sqrt{\frac{\sum y_i^2 - b\sum y_i - a\sum x_i y_i}{gm-2}}$$

$$\text{上限：} b + ax_0 - \left[t_{gm-2,1-\alpha/2} \left(\frac{1}{gm} + \frac{\left(x_0 - \bar{x}\right)^2}{\sum \left(x_i - \bar{x}\right)^2} \right)^{1/2} s \right]$$

$$\text{下限：} b + ax_0 + \left[t_{gm-2,1-\alpha/2} \left(\frac{1}{gm} + \frac{\left(x_0 - \bar{x}\right)^2}{\sum \left(x_i - \bar{x}\right)^2} \right)^{1/2} s \right]$$

7) 画出“*偏倚 =0*”的线，并对图进行评审，以观察是否存在特殊原因，以及线性是否可接受。（见图 11）

如果“偏倚 = 0”的整个直线都位于置信度区间以内，则称该测量系统的线性是可接受的。

结果分析 — 数值法

8) 如果图示法分析表示该测量系统的线性是可接受的，则以下假设应该为真：

H_0：$a = 0$　　*斜率* = 0

如果下式成立，则不能被否定

$$|t| = \frac{|a|}{\left[\frac{s}{\sqrt{\sum \left(x_i - \bar{x}\right)^2}}\right]} \le t_{gm-2,1-\alpha/2}$$

[40] 参见第三章第 B 节，“确定偏倚的指南”中 α 程度选择中的注释。

如果以上假设为真，则测量系统对所有的参考值具有相同的偏倚。这个偏倚必须为 0，该线性才可被接受。

H_0：$b = 0$　　*中心（偏倚）* = 0

如果下式成立，则不能被否定：

$$|t| = \frac{|b|}{\left[\sqrt{\frac{1}{gm} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}}\right] s} \leq t_{gm-2,1-\alpha/2}$$

范例 — 线性

某工厂检验员对某过程引进了一套新测量系统。作为 PPAP[41] 的一部分，需要对测量系统的线性进行评价。根据已文件化的过程变差描述，在测量系统的全部工作量程范围内选取了五个零件。通过对每个零件进行全尺寸检验从而确定它们的参考值，然后由主要操作者对每个零件测量 12 次。在分析中，这些零件是随机抽取的。

	零件参考值	1 2.00	2 4.00	3 6.00	4 8.00	5 10.00
	1	2.70	5.10	5.80	7.60	9.10
	2	2.50	3.90	5.70	7.70	9.30
	3	2.40	4.20	5.90	7.80	9.50
	4	2.50	5.00	5.90	7.70	9.30
测	5	2.70	3.80	6.00	7.80	9.40
量	6	2.30	3.90	6.10	7.80	9.50
次	7	2.50	3.90	6.00	7.80	9.50
数	8	2.50	3.90	6.10	7.70	9.50
	9	2.40	3.90	6.40	7.80	9.60
	10	2.40	4.00	6.30	7.50	9.20
	11	2.60	4.10	6.00	7.60	9.30
	12	2.40	3.80	6.10	7.70	9.40

表 5：线性研究数据

使用散布图和统计软件，检查员画出了线性图（见图 11）。

[41] *生产件批准程序* 手册，第 3 版，2000。

	零件参考值	1 2.00	2 4.00	3 6.00	4 8.00	5 10.00
偏倚	1	0.7	1.1	-0.2	-0.4	-0.9
	2	0.5	-0.1	-0.3	-0.3	-0.7
	3	0.4	0.2	-0.1	-0.2	-0.5
	4	0.5	1	-0.1	-0.3	-0.7
	5	0.7	-0.2	0.0	-0.2	-0.6
	6	0.3	-0.1	0.1	-0.2	-0.5
	7	0.5	-0.1	0.0	-0.2	-0.5
	8	0.5	-0.1	0.1	-0.3	-0.5
	9	0.4	-0.1	0.4	-0.2	-0.4
	10	0.4	0.0	0.3	-0.5	-0.8
	11	0.6	0.1	0.0	-0.4	-0.7
	12	0.4	-0.2	0.1	-0.3	-0.6
	偏倚平均值	0.491667	0.125	0.025	-0.29167	-0.61667

表 6：线性研究 — 中途的结果

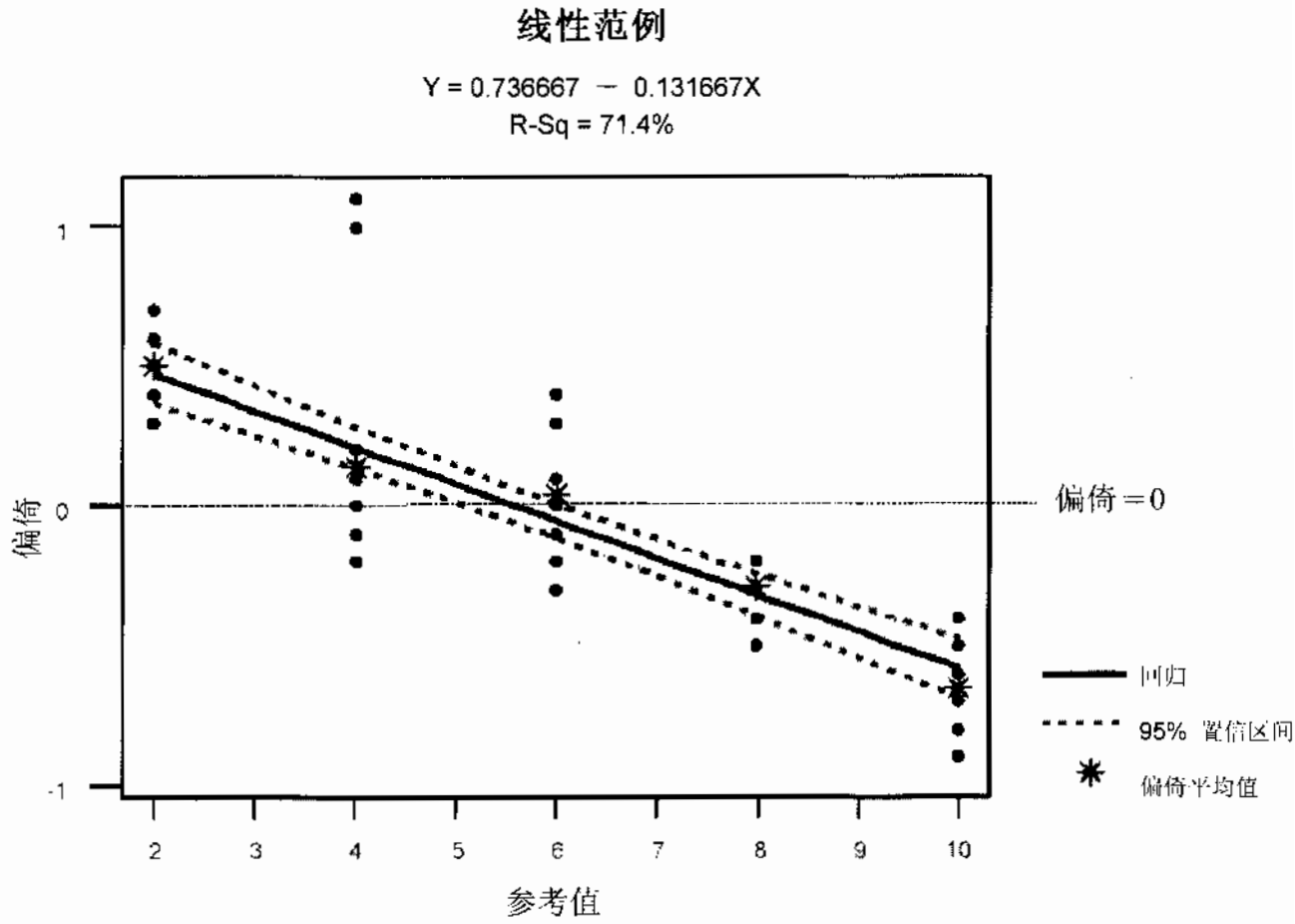


图 11：线性研究 — 图示分析法

图示法分析指出特殊原因可能会影响这测量系统。参考值为 4 的数据呈现两种统计模式。

即使不考虑参考值为 4 的数据，图示分析也明显地表示该测量系统存有一线性问题。R^2 值指出有一线性模型可能不是这些数据的适当模型[42]。即使该线性模式是可接受的，这"*偏倚 = 0*"的直线与置信度区间相交，而不是包含在置信度区间内。

此时，这检验员要开始分析并解决测量系统存在的问题，因为数值分析不能提供任何进一步的方向。然而，为确保不能半途而废，检验员对斜率和中心计算统计的 t 值 (*t*-statistic)：

$$t_a = -12.043$$

$$t_b = 10.158$$

取预设 α = .05；由 *t*-分布表，取 $(gm - 2) = 58$ 个自由度，并取概率为 .975。该督导者得出关键值：

$$t_{58,\ .975} = 2.00172$$

由于 $|t_a| > t_{58,\ .975}$，由图示法分析得到的结论经过了数值法分析的证实 — 本测量系统存在一线性问题。

在这种情况下，由于测量系统存在一线性问题，不用再去分析 t_b 与 $t_{58,\ .975}$ 的关系如何。线性问题的可能的原因可参考第一章，第 E 节"位置变差"。

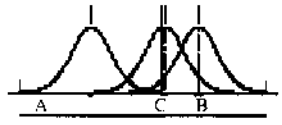
如果测量系统存在一线性问题，需要通过硬体修正、软件修正或对两者修正的方法，对该测量系统进行重新校准使其达到零偏倚。

如果不能将整个测量系统范围内的偏倚调整到零偏倚，它还是可能被用为对产品 / 过程的控制，只要这测量系统依旧稳定，可以不加以分析。

> 由于这有一个评价者误差的较高风险，该测量应该仅在得到顾客同意下被使用。

[42] 参见标准统计内容对使用一线性模式以描述两变数之间关系的适当分析。

确定重复性和再现性的指南 [43]



GRR

可以使用不同的方法进行计量型量具的研究。本节将详细讨论三种可接受的方法。它们是：

- 极差法 (Range method)
- 均值－极差法 (Average and Range method)
- 方差分析法 (ANOVA method)

除极差法之外，其它方法所用的研究数据的设计都很相似。如所呈现的，所有的方法在它们的分析时均忽视了零件内部变差（如：在第四章，第 A 节所讨论的圆度、锥度直径、平面度等。）

但是，整个测量系统不仅包括量具本身及其相关的偏倚、重复性等，还包括被测零件之间的变差。如何处理零件内部的变差，需要取决于对零件使用意图以及测量目的的合理理解。

最后，本章节中的所有技术均以过程处于统计的稳定状态这一前提条件。

> 尽管再现性通常被解释为评价者变差，但有些情况下该变差会由其它原因造成。例如对重复性研究是必要的，对于一些过程中没有人为评价人的测量系统，如果所有的零件由相同的设备来搬运、夹具及测量，则再现性为零。

极差法



极差法是一种经修正的计量型量具研究方法，它能对测量变差提供一个快速地的近似值。这方法只能对测量系统提供变差的整体情况，不能将变差分解成重复性和再现性。它通常用来快速地检查以验证 GRR 是否有变化。

> 使用这方法能够潜在的检测出测量系统为不可接受 [44] 的概率是：对于抽样次数是 5 的情况下，机率为 80%；对于抽样次数为 10 的情况下，机率为 90%。

[43] 参见第一章，第 E 节，对潜在原因的可操作定义和讨论。
[44] 如：%*GRR* > 30%

用极差法进行研究时通常选用两个评价人与五个零件。在这种研究中，两个评价人测量每个零件一次。由评价人 A 测量的每个零件的极差与由评价人 B 测量的每个零件的极差是决然不同的。计算极差之和以及极差的平均值 ($\overline{R}$)；总测量变差即为极差的平均值乘以 $\frac{1}{d_2^*}$，d_2^* 可在附录 C 中查到，取 m = 2，且 g = 零件的数量。

零件	评价人 A	评价人 B	极差 (A，B)
1	0.85	0.80	0.05
2	0.75	0.70	0.05
3	1.00	0.95	0.05
4	0.45	0.55	0.10
5	0.50	0.60	0.10

$$极差平均值\overline{R} = \frac{\sum R_i}{5} = \frac{0.35}{5} = 0.07$$

$$GRR = \left(\frac{\overline{R}}{d_2^*}\right) = \left(\frac{\overline{R}}{1.19}\right) = \left(\frac{0.07}{1.19}\right) = 0.0588$$

（过程标准差 = 0.0777 从之前的研究中取得）

$$\%GRR = 100 * \left(\frac{GRR}{过程标准差}\right) = 75.7\%$$

表 7：量具研究（极差法）

为了确定测量变差占过程标准差的多少百分比，可通过把 *GRR* 乘以 100，再除以过程标准差，即可将 *GRR* 转化成百分数。在以上范例中（参见表 7），该特性的过程标准差为 0.0777，因此：

$$\%GRR = 100 * \left(\frac{GRR}{过程标准差}\right) = 75.7\%$$

现在已确定了这测量系统的 %*GRR*，就应该对这结果进行解释。在表 7 中，%*GRR* *被确定*为 75.7%，于是结论是需对测量系统进行改进。

平均值和极差法

平均值和极差法 ($\overline{X}$ & R) 是一种可同时对测量系统提供重复性和再现性的估计值的研究方法。与极差法不同，这方法允许将测量系统的变差分解成两个独立的部分：重复性和再现性，但不能确定它们两者的相互作用[45]。

进行研究

尽管评价人的人数、测量次数及零件数量均可会不同，但下面的讨论呈现进行研究的最佳情况。参见图12中的GRR数据表，详细的程序如下：

1) 取得一个能代表过程变差实际或预期范围的样本，为 $n > 5$ 个零件[46]的样本。

2) 给评价人编号为A、B、C等，并将零件从1到 n 进行编号，但零件编号不要让评价人看到。

   参见第2章，第C节。

3) 对量具进行校准，如果这是正常测量系统程序中的一部分的话。让评价人A以*随机*顺序[47]测量 n 个零件，并将结果记录在第1行。

4) 让评价人B和C依次测量这些一样的 n 个零件，不要让他们知道别人的读值；然后将结果分别的记录在第6行和第11行。

5) 用不同的随机测量顺序重复以上循环，并将数据记录在第2、7和12行；注意将数据记录在适当的栏位中，例如：如果首先被测量的是零件7，然后将数据记录在标有零件7的栏位中。如果需要进行三次测量，则重复以上循环，并将数据记录在第3、8和13行中。

6) 当测量大型零件或不可能同时获得数个零件时，第3步到第5步将变更成以下顺序：

   ✓ 让评价人A测量第一个零件并将读值记录在第1行；让评价人B测量第一个零件并将读值记录在第6行；让评价人C测量第一个零件并将读值记录在第11行。

---

[45] 使用ANOVA分析法可以用来确定量具和评价人之间的相互作用，如果存在的话。

[46] 为了获得结果中最低限度的置信度水准，产生“极差”的总数要>15个。尽管数据表格被设计成只能最多容纳10个零件，这方法并不局限于这个数值。对于任何统计技术来说，样本数量越多，将会呈现小的抽样变差，产生的风险也越小。

✓ 让评价人 A 重新测量第一个零件并将读值记录在第 2 行；评价人 B 重新测量第一个零件并将读值记录在第 7 行；评价人 C 重复测量第一个零件并将读值记录在第 12 行。如果需要进行三次测量，则重复以上循环，并将数值记录在第 3、8 和 13 行中。

7) 如果评价人处于不同的班次，可以使用一个替代的方法。让评价人 A 测量所有 10 个零件，并将读值记录在第 1 行；然后让评价人 A 按照不同的顺序重新测量，并把读值记录在第 2 行和第 3 行。评价人 B 和评价人 C 也同样做。

[47] 参见第三章第 B 节“随机性及统计独立性”

量具重复性和再现性数据收集表

	评价人 / 测量次数	零件										平均值
		1	2	3	4	5	6	7	8	9	10	
1	A 1	0.29	-0.56	1.34	0.47	-0.80	0.02	0.59	-0.31	2.26	-1.36	
2	2	0.41	-0.68	1.17	0.50	-0.92	-0.11	0.75	-0.20	1.99	-1.25	
3	3	0.64	-0.58	1.27	0.64	-0.84	-0.21	0.66	-0.17	2.01	-1.31	
4	平均值											$\overline{X}_a =$
5	极差											$\overline{R}_a =$
6	B 1	0.08	-0.47	1.19	0.01	-0.56	-0.20	0.47	-0.63	1.80	-1.68	
7	2	0.25	-1.22	0.94	1.03	-1.20	0.22	0.55	-0.08	2.12	-1.62	
8	3	0.07	-0.68	1.34	0.20	-1.28	0.06	0.83	-0.34	2.19	-1.50	
9	平均值											$\overline{X}_b =$
10	极差											$\overline{R}_b =$
11	C 1	0.04	-1.38	0.88	0.14	-1.46	-0.29	0.02	-0.46	1.77	-1.49	
12	2	-0.11	-1.13	1.09	0.20	-1.07	-0.67	0.01	-0.56	1.45	-1.77	
13	3	-0.15	-0.96	0.67	0.11	-1.45	-0.49	0.21	-0.49	1.87	-2.16	
14	平均值											$\overline{X}_c =$
15	极差											$\overline{R}_c =$
16	零件平均值											$\overline{\overline{X}} =$ $R_p =$
17	$\overline{\overline{R}} = (\overline{R}_a =\ ] + [\overline{R}_b =\ ] + [\overline{R}_c =\ ]) / [$评价人数 $=\ ] =$											$\overline{\overline{R}} =$
18	$\overline{X}_{DIFF} =$ [最大值$\overline{X} =\ ]$ − [最小值$\overline{X} =\ ] =$											
19	$*UCL_R = [\overline{\overline{R}} =] \times [D_4 =] =$											

2 次测量时 $D_4 = 3.27$，3 次测量时 $D_4 = 2.58$。UCL_R 代表个别 R 值的限值。圈出那些超出限值的点，查明原因并采取纠正措施；让相同的评价人使用相同的量具用原来的方法重新读值，或剔除这些数值并由其余的数值重新平均和计算 $\overline{\overline{R}}$，以及控制限值。

注：

图 12：量具重复性和再现性数据收集表

结果分析 — 图示法[48]

使用图表工具是很重要的，使用哪种特定的图示取决于用于收集数据的实验设计。在进行其它的统计分析之前，应该使用图表工具对数据进行系统地筛选，从而找出变差的明显的特殊原因。

下面是一些被证明为有用的分析技术（另参见变差数分析法）。

从测量系统分析中得到的数据可通过控制图画显示出来。通过使用控制图来回答与测量系统有关的问题，这一见解已被Western Electric所采用（见参考文献表中*"AT&T Statistical Quality Control Handbook"*）。

平均值图 (Average Chart)

以零件编号顺序画出由每个评价人对每个零件多次读值的平均值。该图可以用来确认评价人之间的一致性。

如果以极差的平均值计算所确定的总平均值和控制限也画出来了，则这产生的平均值图可用来显示测量系统的"实用性"。

控制限以内的区域表示测量的敏感性（干扰）。由于研究中所使用的零件组代表了过程变差，大约一半或一半以上的平均值应该落在控制限之外。如果数据呈现这样的图形，则测量系统应该是适合进行检验出零件之间的变差，以及能为过程的分析和控制提供有用的信息；如果少于一半的数据点落在控制限之外，则测量系统的有效分辨率不足，或这样本不能代表预期的过程变差。

[48] 对这些分析的详细描述已经超出了本手册的范围。要想得到更多的资讯，可查阅参考书目和从权威的统计资料中查询。

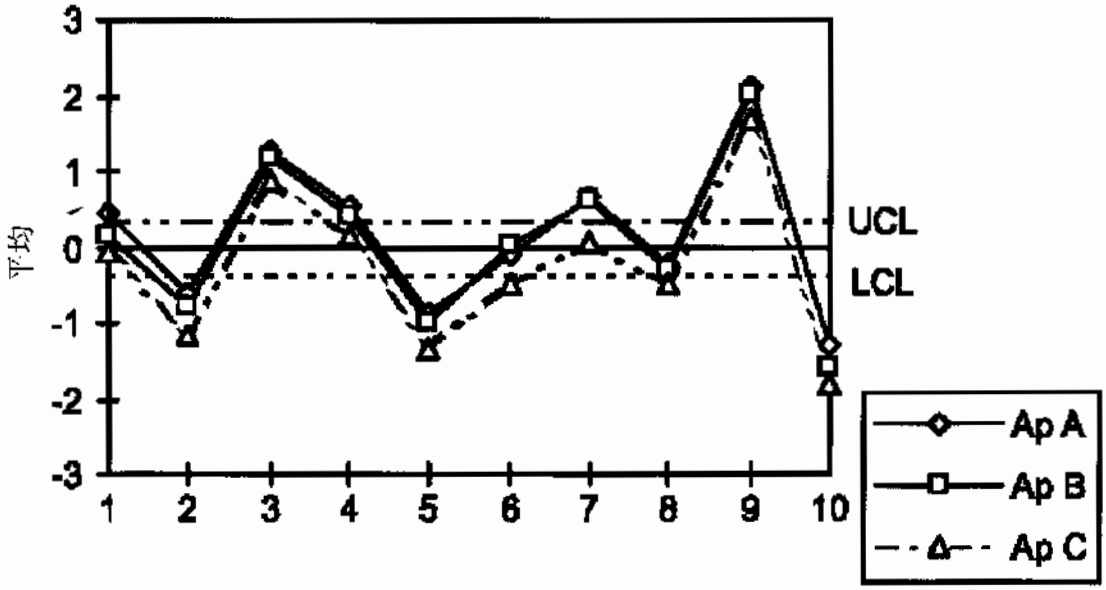


图 13：平均值图 —“重迭画出”[49]

对图进行评价可知：测量系统有足够的解析度来测量样本零件所代表的过程变差。没有发现明显的评价人与评价人之间的差别。

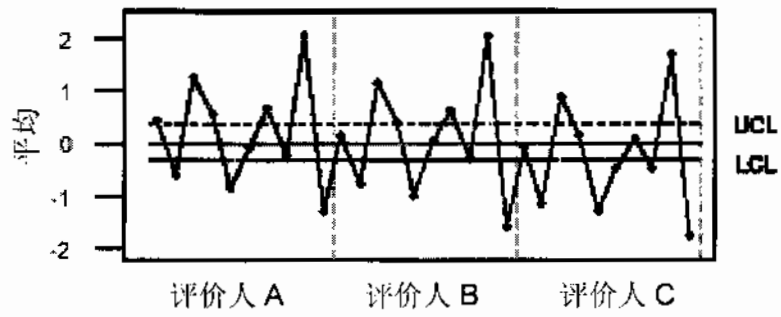


图 14：平均值图 —“非重迭画出”

[49] 在 *ANOVA* 变差数分析方法中，这也被称为是*评价人与零件相互作用*图。

极差图 (Range Chart)

极差图被用来确定过程是否受控。原因是不论测量误差可能有多大，控制限将包含该误差。这就是为什么需要在进行适切的测量系统研究之前，需要识别并消除特殊原因变差的原因。

将由每个评价人对每个零件多次测量读值的极差，画在一个包括了极差平均值和控制限的标准极差图上。从被画在图上数据的分析，可以得到一些有用的解释。如果所有的极差均受控，则说明所有评价人都进行了相同的工作。

如果某个评价人是在控制限之外，则说明他使用的方法与其它人不一致。

如果所有的评价人均有一些超出控制范围的点，则说明该测量系统对评价人的技巧较敏感，需要进行改进以获得有效的数据。

图表应该不是显示数据对于评价人或零件关系的图形。

极差不是对数据的排序。不能像一般控制图趋势分析来使用，即使画出来的数据点是用线条连接的。

稳定性是以一个点或多个点超出了控制限来确定；评价人之间或零件之间。稳定性分析要考虑到实用性和统计的含义

极差图可帮助确定：

- 与重复性有关的统计控制
- 评价人之间对每个零件的测量过程一致性。

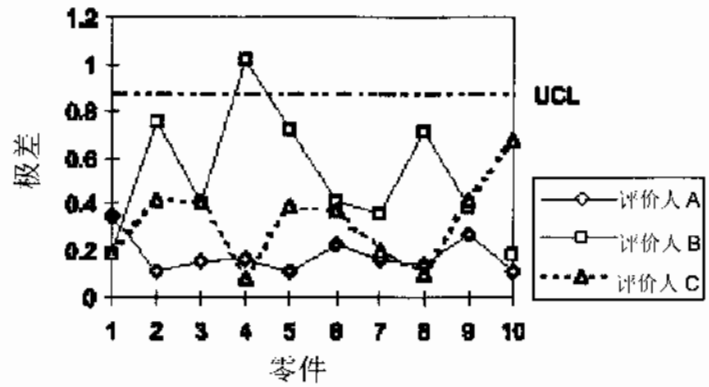


图 15：极差图 — “重迭画出”

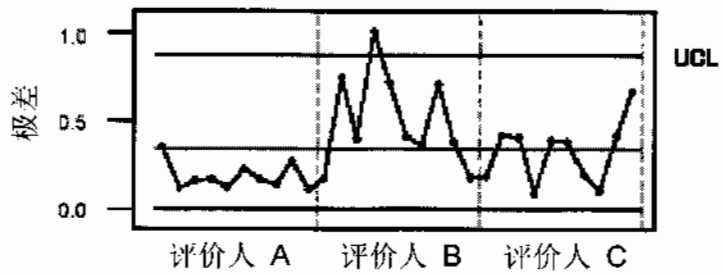


图 16：极差图 — “非重叠画出”

评审以上图表显示评价人的变差之间存在差异。

链图 (Run Chart)

画出所有评价人对零件的所有读值以获得对下列的理解：

- 个别零件在变差一致性上的影响
- 奇异读值的呈现（即不正常的读值）

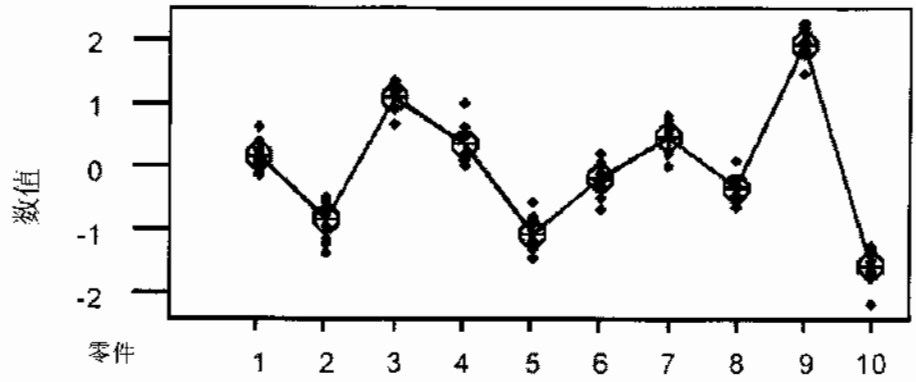


图 17：零件的链图

对上图进行分析可知：没有奇异数据或不一致的零件。

散点图 (Scatter Plot)

将个别的读值依评价人所测量的零件绘制图表，以获得下列的理解：

- 评价人之间的一致性
- 呈现可能的分离
- 零件 — 评价人之间的相互作用

对图 18 进行分析可知：没有指出任何明显的分离，但指出评价人 C 的读值可能比其它人的小。

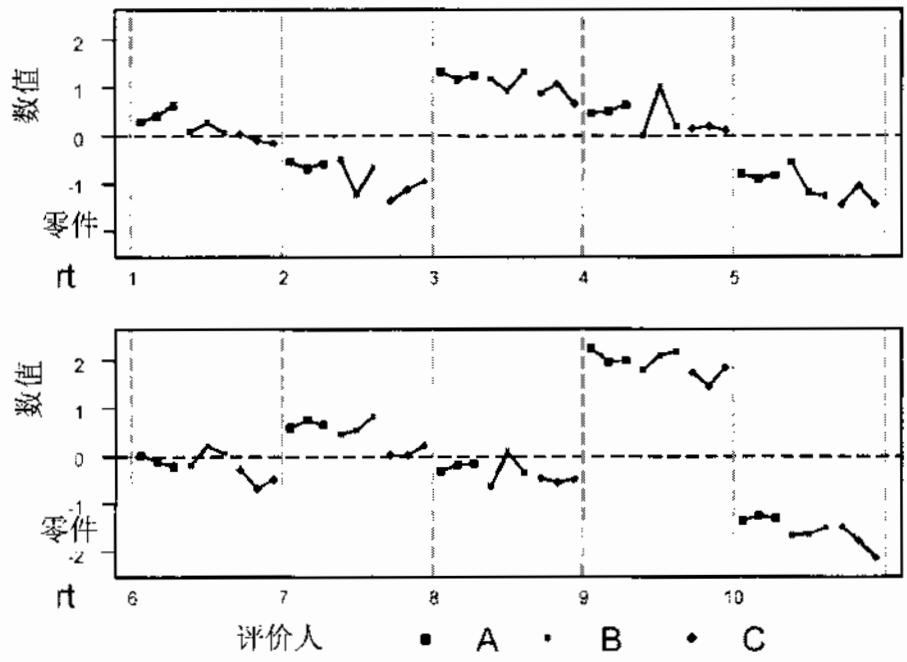


图 18：散点图

振荡图 (Whiskers Chart)

在振荡图中，依评价人所测量的零件画出读值中的最高值、最低值以及平均值（见图 19），通过这图可理解：

- 评价人之间的一致性
- 分离的呈现
- 零件 — 评价人之间的相互作用

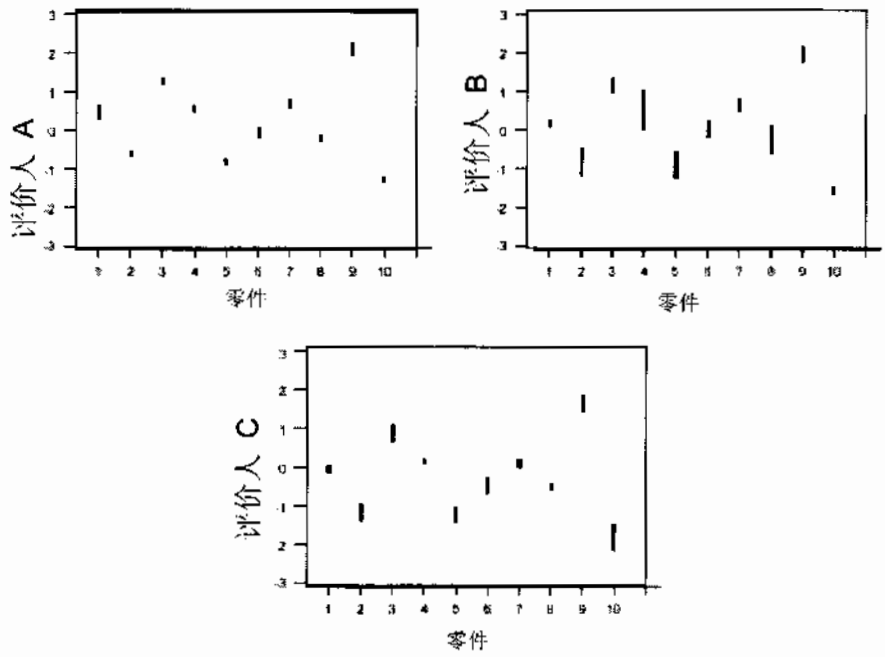


图 19：振荡图

对图 19 进行分析可知：没有呈现任何明显的分离，但指出评价人 B 可能有较大的变差。

误差图 (Error Chart)

测量系统分析的数据可以通过“误差图”（见图 20）来被分析，即画出个别读值相对于可接受的参考值的误差。每个零件的个别读值偏差或误差可用以下公式计算：

误差 = 观测值 − 参考值

或

误差 = 观测值 − 零件的测量平均值

选用哪个公式取决于是否能够得到被测零件的参考值。

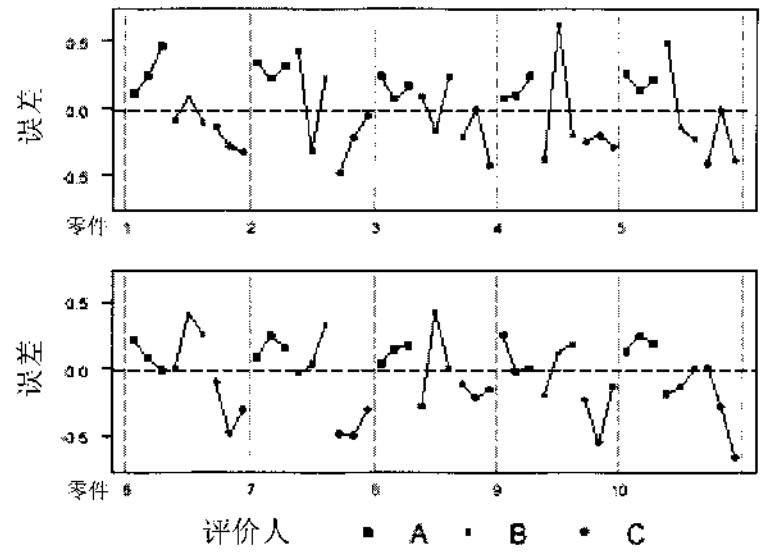


图 20：误差图

对上图进行评审可知：

- 评价人 A 有一整体性的正向偏倚
- 评价人 B 的变差最大，但没有明显的偏倚
- 评价人 C 有一整体性的负向偏倚

正常化直方图 (Normalized Histogram)

直方图（图 21）是呈现参加这研究的评价人误差的频率分布图示法。这图也可显示所有评价人联合的频率分布。

如果可以得到参考值，则

误差 = 观测值 − 参考值

否则

正常化的数值 = 观测值 − 零件平均值

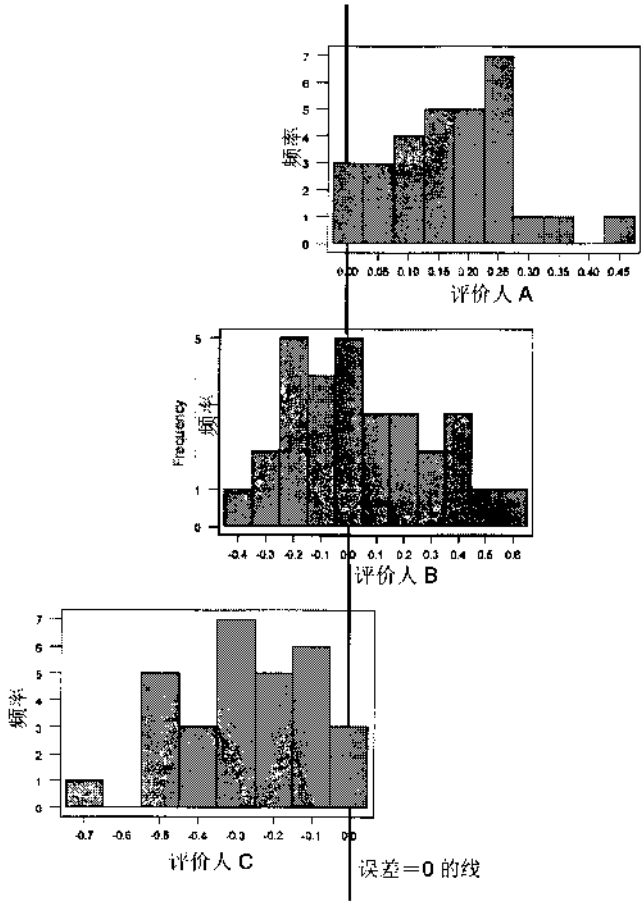


图 21：正常化的直方图[50]

直方图能提供一快速目视观察这误差是如何分布的。甚至在对数据分析之前，便能识别出评价人所进行的测量是否存在偏倚或不一致的问题。

对直方图进行分析（图 21）可以加强误差图分析的结论。它们同样指出只有评价人 B 呈现对称图形。评价人 A 及评价人 C 均由于产生偏倚而带来一系统上变差的来源。

[50] 注意：每个直方图的“0.0”数值是与其它的直方图对齐。

平均值 X—Y 图 (X—Y Plot of Averages by Size)

由各评价人对每个零件多次读值的平均值，其与参考值或零件总平均值作为相对指数的图示法（见图 22），这图可以帮助确定：

- 线性（如果使用了参考值）
- 评价人之间线性的一致性。

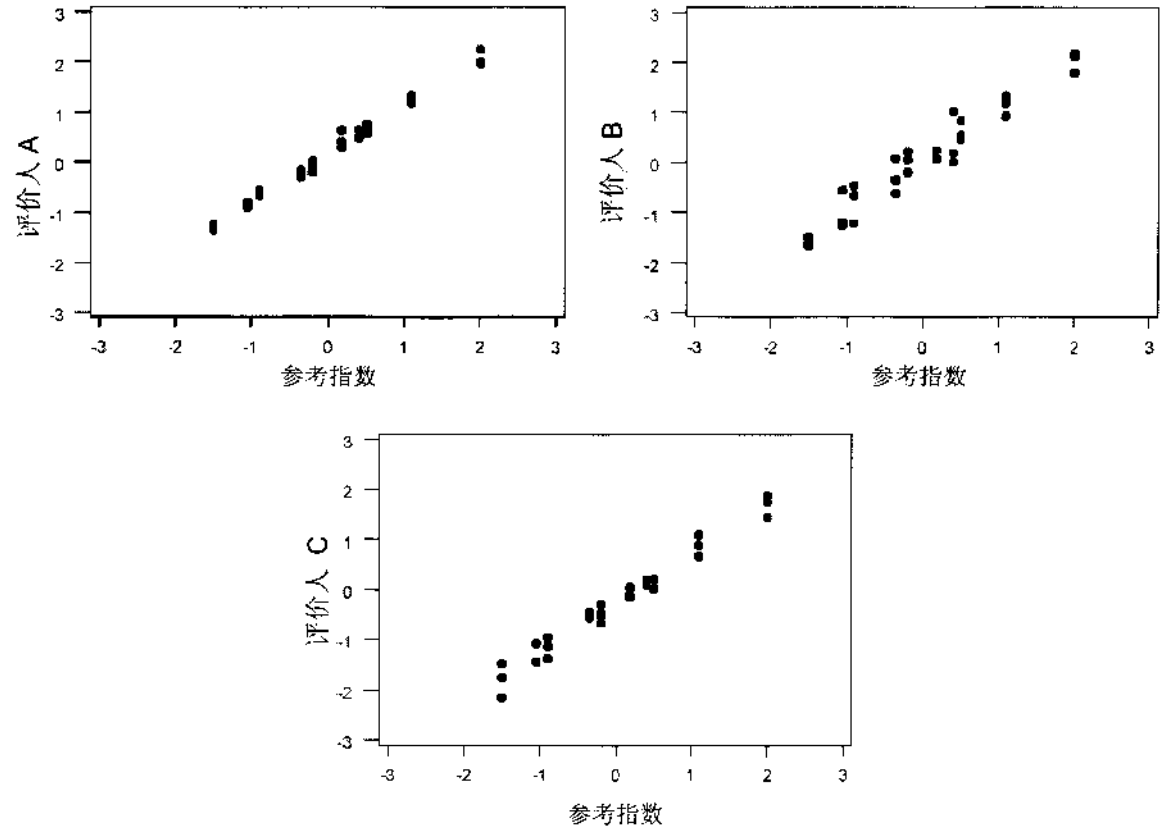


图 22：平均值 X—Y 图

X—Y 比较图 (Comparison X —Y Plots)

由每个评价人对每个零件多次读值的平均值，其与另一评价人的结果成为相对指数的图示法。本图是把一个评价人获得的数据与其它的评价人数据进行比较（见图 23）。如果评价人之间完全一致，则画出的点将形成一条与起源成 45 度角的直线。

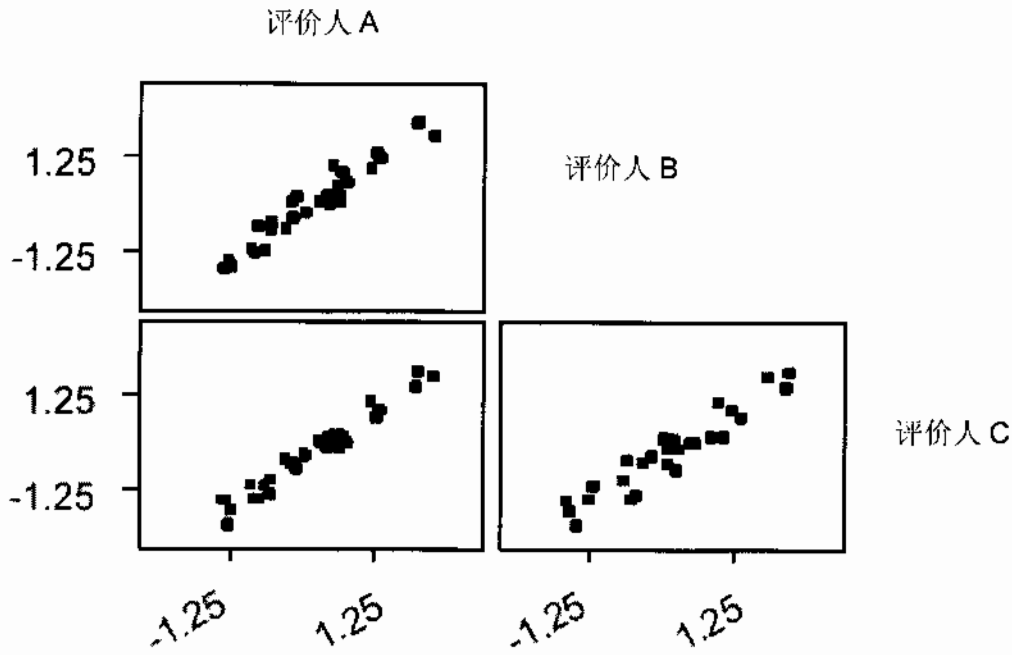


图 23：X—Y 比较图

数值计算

量具的重复性和再现性的计算如图 24 和图 25 所示。图 24 是数据收集表格，记录所有的研究结果；图 25 是报告表格，记录了所有识别信息和依据指定公式进行的所有计算。

在本手册的表格范例部份提供了可复制的空白表格。在收集数据之后所进行的计算程序如下：

（以下计算参见图 24）

1) 用 1、2、3 行中的最大值减去它们中的最小值，把结果记入第 5 行。在第 6、7、8 行和第 11、12、13 行重复以上步骤，并将结果分别记录在第 10 行和第 15 行。

2) 填入第 5、10 及 15 行的数据是极差，所以它们都总是正值。

3) 将第 5 行的数据相加，然后除以零件抽样数量即得到了第一个评价人测量的极差平均值 $\overline{R}_a$；对第 10 行和第 15 行进行同样的计算以得到 $\overline{R}_b$ 和 $\overline{R}_c$。

4) 将第 5、10 和第 15 行的平均值（$\overline{R}_a$，$\overline{R}_b$ 和 $\overline{R}_c$）转到第 17 行。用它们的和除以评价人数，结果记入 $\overline{\overline{R}}$ 栏（所有极差的平均值）。

5) 将$\overline{\overline{R}}$输入第19和20行，并乘以D_4[51]以得到控制限的上限和下限。注意如果进行两次测量，则D_4为3.27。将该单独极差的上控制限(UCL_R)填入19行，对于测量次数少于7次的情况下，极差的下控制限(LCL_R)为零。

6) 对于极差结果大于UCL_R计算结果的任何读值，让原来的评价人对原来的零件进行重新测量，或剔除那些读值，然后根据修改后的抽样数量重新平均并计算极差平均值$\overline{\overline{R}}$和控制限UCL_R。如果已使用了前面所讨论的控制图来绘制并分析数据，这种状况将已经得到了纠正，因此不会在这发生。

7) 加总第1、2、3、6、7、8、11、12和13行的读值，然后将每行的总和除以抽样的数量，并将结果填入该行最右边标有"平均值"的栏位中。

8) 将1、2、3行的平均值相加，用该总和除以测量次数，将所得结果填入第4行中的$\overline{X}_a$栏为中。对第6、7、8行和第11、12、13行重复以上步骤，并将结果分别记录在第9行和第14行中的$\overline{X}_b$和$\overline{X}_c$栏位中。

9) 找出第4、9及14行平均值（$\overline{X}_a$、$\overline{X}_b$和$\overline{X}_c$）中的最大值和最小值，并将它们填入第18行的适当位置并求出它们的差；将这差值填入第18行中标示$\overline{X}_{DIFF}$的空白处。

10) 计算每次测量每个零件的读值之和，然后除以总测量次数（测量次数乘以评价人人数）；将结果填在第16行为每个零件平均值提供的空白处。

11) 用零件平均值的最大值减去零件平均值的最小值，并将结果填入第16行中标示R_p的栏位。R_p是零件平均值的极差。

（以下步骤参见图25）

12) 将计算所得的$\overline{\overline{R}}$、$\overline{X}_{DIFF}$、R_p填入报告表格所提供的空白处。

13) 进行报告表格左侧标题为"测量单元分析"栏之下的计算。

14) 进行报告表格右侧标题为"总变差%"栏之下的计算。

15) 检查计算结果以确定没有错误。

---

[51] 常数表可参见*"统计过程控制(SPC)参考手册"*，1995，或其他统计学参考资料的常数表。

量具重复性和再现性数据收集表

	评价人 / 测量次数	零件 1	2	3	4	5	6	7	8	9	10	平均值	
1	A 1	0.29	-0.56	1.34	0.47	-0.80	0.02	0.59	-0.31	2.26	-1.36		0.1
2	2	0.41	-0.68	1.17	0.50	-0.92	-0.11	0.75	-0.20	1.99	-1.25		0.1
3	3	0.64	-0.58	1.27	0.64	-0.84	-0.21	0.66	-0.17	2.01	-1.31		0.2
4	平均值	0.447	-0.607	1.260	0.537	-0.853	-0.100	0.667	-0.227	2.087	-1.307	$\overline{X}_a =$	0.19
5	极差	0.35	0.12	0.17	0.17	0.12	0.23	0.16	0.14	0.27	0.11	$\overline{R}_a =$	0.1
6	B 1	0.08	-0.47	1.19	0.01	-0.56	-0.20	0.47	-0.63	1.80	-1.68		0.0
7	2	0.25	-1.22	0.94	1.03	-1.20	0.22	0.55	-0.08	2.12	-1.82		0.1
8	3	0.07	-0.68	1.34	0.20	-1.28	0.06	0.83	-0.34	2.19	-1.50		0.0
9	平均值	0.133	-0.790	1.157	0.413	-1.013	0.027	0.617	-0.297	2.037	-1.600	$\overline{X}_b =$	0.06
10	极差	0.18	0.75	0.40	1.02	0.72	0.42	0.36	0.71	0.39	0.18	$\overline{R}_b =$	0.5
11	C 1	0.04	-1.38	0.88	0.14	-1.46	-0.29	0.02	-0.46	1.77	-1.49		-0.2
12	2	-0.11	-1.13	1.09	0.20	-1.07	-0.67	0.01	-0.55	1.45	-1.77		-0.2
13	3	-0.15	-0.96	0.67	0.11	-1.45	-0.49	0.21	-0.49	1.87	-2.16		-0.2
14	平均值	0.073	-1.157	0.880	0.150	-1.327	-0.483	0.080	-0.503	1.697	-1.807	$\overline{X}_c =$	-0.25
15	极差	0.19	0.42	0.42	0.09	0.39	0.38	0.20	0.10	0.42	0.67	$\overline{R}_c =$	0.3
16	零件平均值	0.169	-0.851	1.099	0.367	-1.064	-0.186	0.454	-0.342	1.940	-1.571	$\overline{\overline{X}} =$ $R_p =$	.00 3.5
17	$\left([\overline{R}_a = 0.184] + [\overline{R}_b = 0.513] + [\overline{R}_c = 0.328]\right) / [\text{评价人数} = 3] =$											$\overline{\overline{R}} =$	0.34
18	$[\text{最大值}\overline{X} = 0.1903] - [\text{最小值}\overline{X} = -0.2543] = \overline{X}_{DIFF} = 0.4446$												
19	$*[\overline{\overline{R}} = 0.3417] \times [D_4 = 2.58] = UCL_R = 0.8816$												

2 次测量时 D_4 =3.27；3 次测量时 D_4 = 2.58。UCL_R 代表个别 R 值的限值。圈出那些超出这限值的点，查明原因并采取纠正措施；让相同的评价人使用相同的量具用原来的方法重新读值，或剔除这些数值并由其余的数值重新平均和计算 $\overline{\overline{R}}$，以及控制限值。

注：

图 24：完成的 GR&R 数据收集表

量具重复性和再现性报告

零件编号和名称：　　量具名称：　　日期：

特性：　　量具编号：　　操作人：

规范：　　量具型式：

操作表：$\overline{\overline{R}} = 0.3417$　　$\overline{X}_{DIFF} = 0.4446$　　$R_p = 3.511$

测量单元分析	总变差% (*TV*)
重复性 — 设备变差(*EV*) $EV = \overline{\overline{R}} \times K_1$ $= 0.3417 \times 0.5908$ $= 0.20188$ 试验次数 / K_1：2 → 0.8862；3 → 0.5908	$\%EV = 100[EV/TV]$ $= 100[0.20188/1.14610]$ $= 17.62\%$
再现性 — 评价人变差(*AV*) $AV = \sqrt{(\overline{X}_{DIFF} \times K_2)^2 - (EV^2/(nr))}$ $= \sqrt{(0.4446 \times 0.5231)^2 - (0.20188^2/(10 \times 3))}$ $= 0.22963$ *n*=零件数　*r*=测量次数 评价人数量 / K_2：2 → 0.7071；3 → 0.5231	$\%AV = 100[AV/TV]$ $= 100[0.22963/1.14610]$ $= 20.04\%$
重复性和再现性(*GRR*) $GRR = \sqrt{EV^2 + AV^2}$ $= \sqrt{(0.20188^2 + 0.22963^2)}$ $= 0.30575$ 零件数量 / K_3：2 → 0.7071；3 → 0.5231	$\%GRR = 100[GRR/TV]$ $= 100[0.30575/1.14610]$ $= 26.68\%$
零件变差(*PV*) $PV = R_p \times K_3$ $= 1.10456$ 4 → 0.4467；5 → 0.4030；6 → 0.3742	$\%PV = 100[PV/TV]$ $= 100[1.10456/1.14610]$ $= 96.38\%$
总变差(*TV*) $TV = \sqrt{GRR^2 + PV^2}$ $= \sqrt{(0.30575^2 + 1.10456^2)}$ $= 1.14610$ 7 → 0.3534；8 → 0.3375；9 → 0.3249；10 → 0.3146	$ndc = 1.41\left(PV/GRR\right)$ $= 100(1.10456/0.30575)$ $= 5.094 \sim 5$

对本表中所使用理论和常数的信息，请参考 *MSA 参考手册*，第三版

图 25：量具重复性和再现性报告

结果分析 — 数值法

图 24 和 25 的*量具重复性和再现性数据收集表及报告表*，提供了研究数据的数值分析[52]。这种分析将可以估计变差、整个测量系统变差占过程总变差的百分比，以及测量系统的重复性、再现性，和零件的变差。这些信息需要与图示分析法所得的全数结果进行比较。

在报告表（图 25）左侧的测量单元分析的下面，是对每个变差组成部分的标准差计算。

重复性或设备变差（*EV* 或 σ_E）是由极差平均值（$\bar{\bar{R}}$）乘以一个常数（K_1）来决定。K_1 取决于量具研究中的测量次数，其值为从附录 C 查到的 d_2^* 的倒数。d_2^* 取决于测量次数（*m*）和零件的数量乘以评价人的人数（*g*）*（为了计算 K_1 值，这值假设大于 15）*。

再现性或评价人变差（*AV* 或 σ_A）是由评价人平均值的最大差值（$\bar{X}_{DIFF}$）乘以一个常数（K_2）来决定。K_2 取决于量具研究中的评价人的人数，其值为从附录 C 查到的 d_2^* 的倒数。d_2^* 取决于评价人的人数（*m*），且 $g=1$，因为只有一个极差计算。由于评价人变差被包含在设备变差中，因此必须通过减去设备变差的一个分数来对其进行调整。因此，评价人变差（*AV*）可由下式计算得到：

$$AV = \sqrt{\left(\bar{X}_{DIFF} \times K_2\right)^2 - \frac{(EV)^2}{nr}}$$

式中：n = 零件的数量，以及 r = 测量次数

如果根号下所得的数值为负数，则评价人变差（*AV*）为零。

测量系统变差的重复性和再现性（*GRR* 或 σ_M）的计算为设备变差的平方加上评价人变差的平方，然后再开根号，如下式：

$$GRR = \sqrt{(EV)^2 + (AV)^2}$$

[52] 范例中的数值结果如是由人工计算，结果被四舍五入到小数点后一位。由电脑程式计算则这种中间值应该保留电脑 / 程式语言的最大精度。从一个有效的电脑程式产生的结果可能与范例中产生的第二个小数位或更后面的小数位有差别，但最终分析将保持一致。

零件与零件之间的变差（*PV* 或 σ_p）是由零件平均值的极差（R_p）乘以一个常数（K_3）所决定。K_3取决于量具研究中的零件的数量，其值为从附录 C 查到的d_2^*的倒数。d_2^*取决于零件的数量（*m*）与（g），在这情况 $g=1$，因为只有一个极差计算。

这研究的总变差（*TV* 或 σ_T）是加总了重复性和再现性变差的平方与零件之间的变差（*PV*）的平方，再开根号计算而得。即：

$$TV = \sqrt{(GRR)^2 + (PV)^2}$$

如果已知该过程变差，且它的值是以 6σ 为基础，则可用它来代替从量具研究数据中计算得到的总研究变差（*TV*）。也就是说可通过以下两个公式进行计算来完成：

$$1)\quad TV = \frac{\text{过程变差}}{6.00}$$

$$2)\quad PV = \sqrt{(TV)^2 - (GRR)^2}$$

以上两个值都可用来代替前面计算的值。

一旦确定了在量具研究中各个因素的变差后，可将它们与总变差（*TV*）进行比较。也就是完成量具报告表（图 25）中右侧"总变差%"下方的计算。



设备变差（%*EV*）占总变差（*TV*）的百分比，被计算为 100 [*EV*/*TV*]，其它因素占总变差的百分比计算方法相似，如下式：

$$\%AV = 100\,[AV / TV]$$

$$\%GRR = 100\,[GRR / TV]$$

$$\%PV = 100\,[PV / TV]$$

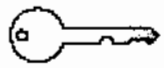


每个因素所占的百分比之和将不等于 100%。

需要对以上总变差的百分比结果进行评价，以确定该测量系统对其预期的使用是否为可接受。

如果分析是以公差为基础来代替以过程变差为基础的话，则可对量具重复性和再现性报告表（图 25）进行修改，使表格右边的总变差的百分比由公差的百分比来代替。在这种情况下，%*EV*、%*AV*、%*GRR* 以及 %*PV* 的

计算公式中的分母是由公差*除以 6* 来代替总变差 (*TV*)。无论使用哪一种或两人方法都用，是取决于这测量系统的期望用途以及顾客的期望。



在这数值分析的最后一步是确定区别分类数 (the number of distinct categories)，这能由该测量系统可靠地分辨，这是可以覆盖预期的产品变差的非重迭 97%自信度区间。[53]

$$ndc = 1.41\left(PV/GRR\right)$$

如果图示分析法没有发现特殊原因的变差，则可用第二章第 D 节的比例法则来确定量具和重复性和再现性 (*%GRR*)。

另外，*ndc* 应该四舍五入到整数，且要能大于或等于 5。

方差分析法 *(ANOVA)*

方差分析法 (anlysis of variance, *ANOVA*) 是一种标准的统计技术，可用它来分析测量误差和一测量系统研究中的其它变差来源。在变差的分析中，变差可分解成四类：零件、评价人、零件与评价人之间的相互作用，以及由于量具造成的重复误差。

与平均值和极差法相比，*ANOVA* 方法有以下的优点：

- 有能力解决任何实验的作业准备；
- 能更准确地估计变差；
- 可从实验数据中得到更多的信息（如：零件与评价人之间相互作用的影响）。

其缺点是：其数值计算更加复杂，而且要求一定程度的统计学知识以解释它的结果。以下章节所介绍的 *ANOVA* 方法是一建议的方法，尤其是可以使用电脑时。

随机化和统计的独立性

数据收集的方法在 *ANOVA* 法中是很重要的。如果数据不是以随机方法收集的，将会带来一偏倚值的误差来源。为了确保 (*n*) 个零件、(*k*) 个评价人、(*r*) 次测量的平衡设计，一个简单的方法就是通过随机化。一个随机化的通用方法

[53] *ndc* 与 *Wheeler* 的分级比率" 相同，在第一版 *"Evaluating the Measurement Process"* 中被定义，Wheeler 和 Lyday 著，1984。确定有效分辨力的一种可替代的方法 ("Discrimination Ratio")，见 1989 年所出版的的第二版。

是：在一张小纸条上写下 *A1*，代表第一个评价人对第一个零件的测量，写下 *A* (*n*) 代表第一个评价人对第 *n* 个零件的测量；用同样的程序对下一个评价人进行编号，直到第 *k* 个评价人。类似的标记将被用在 B_1、C_1 代表第二、第三个人对第一个零件的测量。等所有的 *nk* 个组合都被写好后，把这些纸条放入一个帽子或一个碗中；每次抽取一张纸条，这纸条的组合 (A_1、B_2、...) 就是这量具研究所进行的测量顺序。等所有 *nk* 个组合的纸条均抽出来后，把它们再全都放进帽子中，再重复以上程序；重复该程序 *r* 次以确定每次重复的实验顺序。

也可使用其它方法来进行随机抽样。但应该注意实施随机抽样、随意抽样与自由抽样[54]等的区别。

进行研究

使用与图 12 相似的表格，以随机的方式收集数据。在这个范例中，选取了 10 个零件、3 位评价人，且这实验在实施了随机的顺序下对每个零件和每位评价人的组合测量 3 次。

图示分析法

前面所讨论过的任何图示法，可被使用在 *ANOVA* 研究中对收集数据的图形分析的部份。用这些方法可对数据进行确认并进一步的提供理解（如趋势、周期等）。

建议采用一种被称为*相互作用图 (interaction plot)* 的图示方法。该图可以确认 *F* 试验对相互作用是否很大的结果。在图 26 描绘的相互作用图，其横向轴是零件编号 (1, 2,... 等)，纵轴是每个评价人对每个零件的测量平均值。将每个评价人对每个零件测量平均值的点连起来可以得到 *k*（评价人的人数）条线。解释该图的方法是：如果这 *k* 条线是平行，且没有交互关系；若这些线不平行，则这相互作用是有意义的。相交的角度愈大，相互作用愈大。应该采取适当的措施来消除造成相互作用的原因。在图 26 所示的范例中，这些线接近平行，表明没有明显的相互作用。

[54] 参见 Wheeler 与 Lyday 所著，*"Evaluating the Measurement Process"*，第二版，1989，第 27 页。

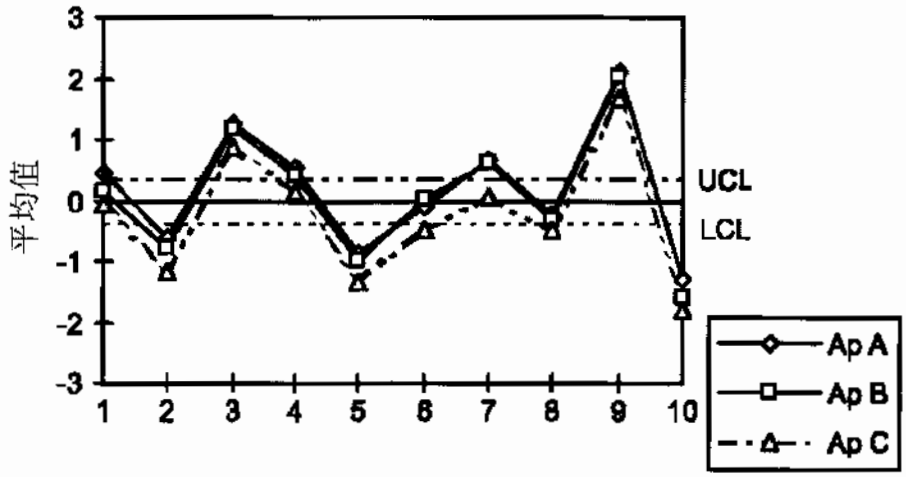


图 26：相互作用图

另外一种图有时是有意思的 - *余数图 (residuals plot)*。该图能更好地检查假设的有效性。这假设是：量具（误差）是从一个常态分布的一项随机变数。画在图上的余数是观测值与预测值之间的差值，预测值是每个评价人对每个零件重复读值的平均值。如果这些余数不是随机地散布在零（水平参考线）的上侧和下侧，则这可能是因为假设是错误的，建议对数据进行进一步的调查。

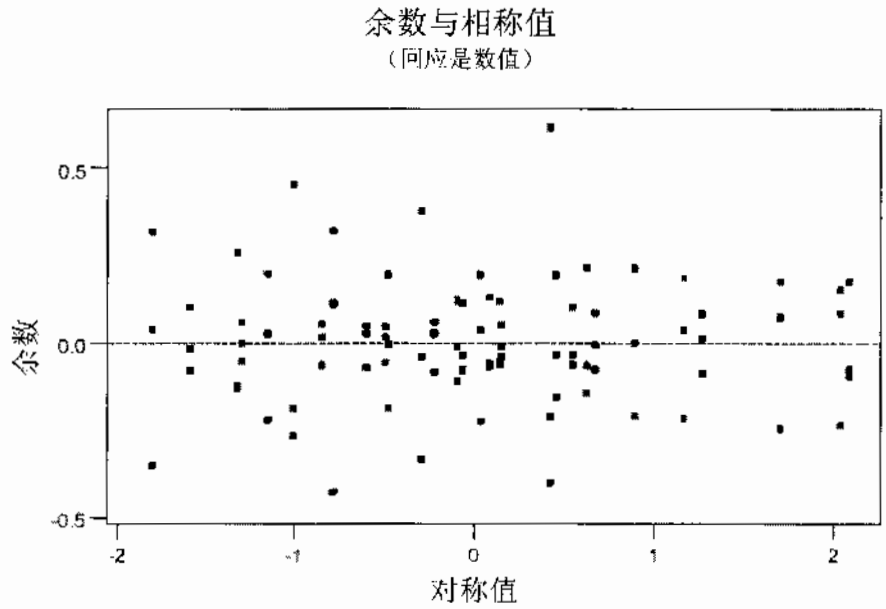


图 27：余数图

数值计算

虽然这些数值可以采用人工计算，但许多人利用电脑程式来产生所谓的方差分析法 (*ANOVA*) 表格。(参见附录A。)

这里的 *ANOVA* 表格是由五列所构成（参见表8)：

- *来源* 列是造成变差的原因。
- *DF* 列是关于这来源的 *自由度 (degree of freedom)*。
- *SS* 或 *平方的总和 (sum of squares)* 列为这来源平均值附近的偏差。
- *MS* 或 *平均平方 (mean square)* 列为平方的总和除以自由度。
- *F－比率 (F-ratio)* 列是一个计算值，用来确定来源数值的统计含义。

ANOVA 表用来将总变差分解成四个部分：零件、评价人、零件和评价人之间的相互作用，以及由于仪器造成的重复性。

为了便于分析，将负的变差要素设定为零。

如在平均值和极差分析法一样，这些讯息是用来确定测量系统的特性。

表8为从图24中的数据范例，用 *ANOVA* 计算的结果，假设是一个固定的影响模式。表10为 *ANOVA* 分析法与平均值和极差法的比较。表11是对 *ANOVA* 分析法的GRR报告。

变差来源	DF	SS	MS	F
评价人	2	3.1673	1.58363	34.44*
零件	9	88.3619	9.81799	213.52*
评价人和零件	18	0.3590	0.01994	0.434
设备	60	2.7589	0.04598	
总计	89	94.6471		

*意义是在 $\alpha = 0.05$ 水准

表8：*ANOVA* 表格

变数的估计值	标准差 (σ)	总变差%	贡献率%
τ^2 = 0.039973 （重复性）	*EV* = 0.199933	18.4	3.4
ω^2 = 0.051455 （评价人）	*AV* = 0.226838	20.9	4.4
γ^2 = 0 （相互作用）	*INT* = 0	0	0
系统 = 0.09143 $\tau^2 + \gamma^2 + \omega^2$	*GRR* = 0.302373	27.9	7.8
σ^2 = 1.086446 （零件）	*PV* = 1.042373	96.0	92.2
总变差	*TV* = 1.085	100.0	

表 9：*ANOVA* 分析法，变差%及贡献率
（变数的估计值是以没有相互作用的模型为基础）

$$ndc = 1.41\left(1.04233 \big/ .30237\right) = 4.861 \cong 4$$

$$\text{总变差（}TV\text{）} = 100\sqrt{GRR^2 + PV^2}$$

$$\%\text{总变差} = 100\left(\frac{\sigma_{(\text{要素})}}{\sigma_{(\text{全部})}}\right)$$

$$\left(\text{对总变差的贡献率}\right)\% = 100\left(\frac{\sigma^2_{(\text{要素})}}{\sigma^2_{(\text{全部})}}\right)$$

方法	下 90%*CL*[55]	标准差	上 90%*CL*	总变差%
GRR*				
EV	0.175	.202	0.240	17.6
AV	0.133	.230	1.016	20.1
相互作用	--	na	--	na
GRR	0.266	.306	0.363	26.7
PV		1.104		96.4
ANOVA				
EV	0.177	0.200	0.231	18.4
AV	0.129	0.227	1.001	20.9
相互作用	--	0	--	0
GRR	0.237	0.302	1.033	27.9
PV		1.042		96.0

*在平均值和极差法中，无法估计相互作用要素。

表 10：*ANOVA* 分析法与平均值和极差法的比较

零件编号及名称：　　量具名称：　　日期：
特性：　　量具号：　　执行者：
规范：　　量具类型：

	标准差	总变差%	贡献率
重复性 (*EV*)	0.200	18.4	3.4
再现性 (*AV*)	0.227	20.9	4.4
评价人与零件 (*INT*)	0	0	0
GRR	0.302	27.9	7.9
零件之间 (*PV*)	1.042	96.0	92.2

对于过程的控制和分析来说，测量系统是可接受的

注：
公差 = 不适用　　总变差 (*TV*) = 1.085
区别分类数 (*ndc*) = 4

表 11：*ANOVA* 方法的 GRR 报告

[55] *CL* = 置信度限值 (Confidence Limit)

GRR 研究的分析

平均值和极差法和 *ANOVA* 分析法两者均能提供关于测量系统或量具的变差信息。

例如，如果重复性大于再现性，原因可能是：

- 仪器需要维修。
- 可能需要对量具进行重新设计，以获得更好的严格度。
- 需要对量具的夹紧或固定装置进行改进。
- 零件内变差太大。

如果再现性大于重复性，原因可能是：

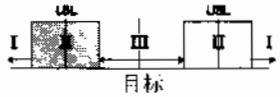
- 需要更好的对评价人进行如何使用和判读该量具仪器的培训。
- 量具校准，刻度的不清晰。

也许需要用某种形式的夹具帮助评价人更一致地使用量具。

第三章 — 第C节

计数型测量系统研究

引言



计数型测量系统是一种测量数值为一有限的分类数量的测量系统。它与能获得一连串数值结果的计量型测量系统截然不同。通 / 止规 (go/no go gage) 是最常用的量具，它只有两种可能的结果；其它的计数型测量系统，例如目视标准，可能产生五到七个分类，如非常好、好、一般、差、非常差。在前面章节中所介绍的分析方法不能被用来评价这样的系统。

如第一章第 G 节所讨论，使用任何测量系统来作结论都会存在有一定程度的风险；由于最大的风险位于分类的边界处，因此，最合适的分析方法是量化测量系统的变差成为量具性能曲线。

风险分析方法

在一些计数情况下，不可能充分获得附有计量型参考数值的零件。在这种情况下，可使用下述方法来评价作出错误或不一致[56]决定的风险[57]：

- 假设性试验分析
- 信号探测理论

由于这些方法没有量化测量系统变差，应该只有在顾客同意的条件下才能使用。选择和使用这样的技术应该基于一个良好的统计上的实践、了解影响产品和测量系统的潜在变差来源，以及了解一个错误决定对接续过程以及最终顾客的影响。

计数型测量系统的变差来源应该通过利用了人为因素和人机工程学的研究结果使之最小化。

可能的方法

情景

某生产过程处于统计受控状态，其性能指数为 $Pp = Ppk = 0.5$，这是不可接受的。由于过程正在生产不合格的产品，于是被要求采取遏制措施以便从生产过程中挑出不可接受的产品。

---

[56] 这包括了多位评价者的比较。

[57] 参见参考书目表。

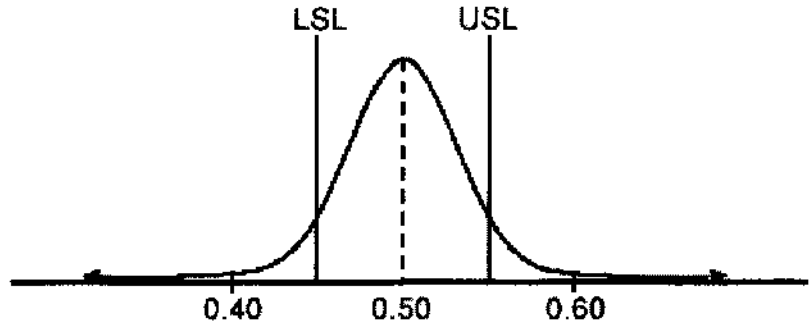


图 28：过程范例

具体的遏制措施是：过程小组采用一个计数型量具来对每个零件与一指定的限值进行比较，如果满足限值则可接受该零件，不满足的零件则拒收（如通 / 止量具）。许多这样的计数型量具基于一套基准零件来设定接受还是拒收；不像计量型量具，计数型量具不能显示一个零件多么好或多么坏，它只能指示该零件是被接受还是拒收（即 2 个分类）。

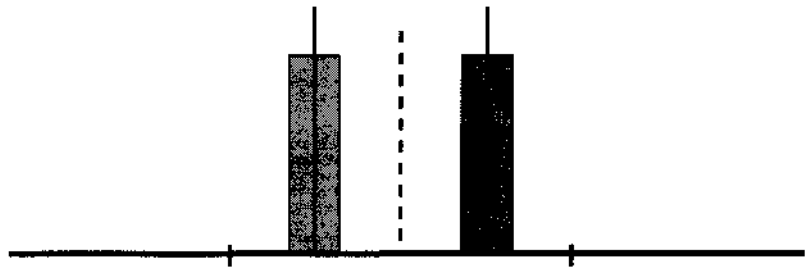


图 29：与测量系统有关的“灰色”区域

小组使用了一个%*GRR* 为公差[58]的 25%的特定量具。由于这还没有被小组文件化，于是需要对这测量系统进行研究。小组已决定从过程中随机选取 50 个零件，以获得涵盖了整个过程范围的零件。

使用 3 名评价人，每位评价人对每个零件测量 3 次。

[58] 由于过程变差大于公差，适当的方法是将测量系统与公差进行比较，而不是与过程变差进行比较。

零件	A-1	A-2	A-3	B-1	B-2	B-3	C-1	C-2	C-3	参考	参考值	代码
1	1	1	1	1	1	1	1	1	1	1	0.476901	+
2	1	1	1	1	1	1	1	1	1	1	0.509015	+
3	0	0	0	0	0	0	0	0	0	0	0.576459	–
4	0	0	0	0	0	0	0	0	0	0	0.566152	–
5	0	0	0	0	0	0	0	0	0	0	0.570360	–
6	1	1	0	1	1	0	1	0	0	1	0.544951	x
7	1	1	1	1	1	1	1	0	1	1	0.465454	x
8	1	1	1	1	1	1	1	1	1	1	0.502295	+
9	0	0	0	0	0	0	0	0	0	0	0.437817	–
10	1	1	1	1	1	1	1	1	1	1	0.515573	+
11	1	1	1	1	1	1	1	1	1	1	0.488905	+
12	0	0	0	0	0	0	0	1	0	0	0.559918	x
13	1	1	1	1	1	1	1	1	1	1	0.542704	+
14	1	1	0	1	1	1	1	0	0	1	0.454518	x
15	1	1	1	1	1	1	1	1	1	1	0.517377	+
16	1	1	1	1	1	1	1	1	1	1	0.531939	+
17	1	1	1	1	1	1	1	1	1	1	0.519694	+
18	1	1	1	1	1	1	1	1	1	1	0.484167	+
19	1	1	1	1	1	1	1	1	1	1	0.520496	+
20	1	1	1	1	1	1	1	1	1	1	0.477236	+
21	1	1	0	1	0	1	0	1	0	1	0.452310	x
22	0	0	1	0	1	0	1	1	0	0	0.545604	x
23	1	1	1	1	1	1	1	1	1	1	0.529065	+
24	1	1	1	1	1	1	1	1	1	1	0.514192	+
25	0	0	0	0	0	0	0	0	0	0	0.599581	–
26	0	1	0	0	0	0	0	0	1	0	0.547204	x
27	1	1	1	1	1	1	1	1	1	1	0.502436	+
28	1	1	1	1	1	1	1	1	1	1	0.521642	+
29	1	1	1	1	1	1	1	1	1	1	0.523754	+
30	0	0	0	0	0	1	0	0	0	0	0.561457	x
31	1	1	1	1	1	1	1	1	1	1	0.503091	+
32	1	1	1	1	1	1	1	1	1	1	0.506850	+
33	1	1	1	1	1	1	1	1	1	1	0.487613	+
34	0	0	1	0	0	1	0	1	1	0	0.449696	x
35	1	1	1	1	1	1	1	1	1	1	0.498698	+
36	1	1	0	1	1	1	1	0	1	1	0.543077	x
37	0	0	0	0	0	0	0	0	0	0	0.409238	–
38	1	1	1	1	1	1	1	1	1	1	0.488184	+
39	0	0	0	0	0	0	0	0	0	0	0.427687	–
40	1	1	1	1	1	1	1	1	1	1	0.501132	+
41	1	1	1	1	1	1	1	1	1	1	0.513779	+
42	0	0	0	0	0	0	0	0	0	0	0.566575	–
43	1	0	1	1	1	1	1	1	0	1	0.462410	x
44	1	1	1	1	1	1	1	1	1	1	0.470832	+
45	0	0	0	0	0	0	0	0	0	0	0.412453	–
46	1	1	1	1	1	1	1	1	1	1	0.493441	+
47	1	1	1	1	1	1	1	1	1	1	0.486379	+
48	0	0	0	0	0	0	0	0	0	0	0.587893	–
49	1	1	1	1	1	1	1	1	1	1	0.483803	+
50	0	0	0	0	0	0	0	0	0	0	0.446697	–

表 12：计数型研究数据表

设定用 1 表示可接受的决定，0 为不可接受的决定。表 12 中所示的参考决定和计量参考值在一开始还没有被确定。表 12 还显示了“代码”列，分别用“−”、“x”、“+”代表零件是否位于 I 区，II 区及 III 区。

假设试验分析 - 交叉表法

由于小组不知道零件的参考决定，他们展开了交叉表格 (cross-tabulations) 来比较每个评价人与其它人的结果。

A*B 交叉表

			B		总计
			.00	1.00	
A	.00	数量	44	6	50
		期望的数量	15.7	34.3	50.0
	1.00	数量	3	97	100
		期望的数量	31.3	68.7	100.0
总计		数量	47	103	150
		期望的数量	47.0	103.0	150.0

B*C 交叉表

			C		总计
			.00	1.00	
B	.00	数量	42	5	47
		期望的数量	16.0	31.0	47.0
	1.00	数量	9	94	103
		期望的数量	35.0	68.0	103.0
总计		数量	51	99	150
		期望的数量	51.0	99.0	150.0

A*C 交叉表

			C		总计
			.00	1.00	
A	.00	数量	43	7	50
		期望的数量	17.0	33.0	50.0
	1.00	数量	8	92	100
		期望的数量	34.0	66.0	100.0
总计		数量	51	99	150
		期望的数量	51.0	99.0	150.0

这些表格的目的在于确定评价人之间一致性的程度。为确定这一致性的程度，小组使用了（cohen 的）*Kappa*，这是用来衡量两个评价人对同一物体进行评价时，其评定结论的一致性。*Kappa* 为 1 时，表示有完全的一致性，为 0 时，表示一致性不比可能性来得好。*Kappa* 仅用于表格，表中两个变数有相同的分类值，且两个变数有相同的分类数量。

Kappa 一种对*评价人内部一致性*的衡量，它测试在诊断区（获得相同评定的零件）中的数量与那些基于可能性期望的数量是否有差别。

让 P_0 = 在对角栏框中，观测比例的总和

P_e = 在对角栏框中，期望部分的总和

则：

$$kappa = \frac{P_0 - P_e}{1 - P_e}$$

Kappa 是一种程度而不是一种试验[59]。通过使用一种渐近的标准误差以形成一个 *t* 统计值来判断其大小。通用的比例法则是 *kappa* 值大于 0.75 则表示有很好的一致性（最大的 *kappa*= 1）；*kappa* 值小于 0.40 则表示一致性不好。

Kappa 不考虑评价人之间的不一致量有多大，只考虑他们之间是不是一致[60]。

通过以上对评价人计算了 *kappa* 程度，小组得到以下结论：

Kappa	A	B	C
A	—	.86	.78
B	.86	—	.79
C	.78	.79	—

这分析表明所有评价人与其它评价人之间有良好的一致性。

[59] 在所有的这种分类评价中，需要以大量零件来涵盖所有的可能范围。

[60] 如果这观测是在一个固定的范畴级别上进行的测量，可用*加权的 kappa* 以更好地衡量一致性。按 *kappa* 来看待两个评价人的一致性，但不一致是通过测量评价人差异来衡量分类的数量。

这种分析用来对确定评价人之间是否有差异的需求，但不能告诉我们这测量系统从坏零件中挑选出好零件的能力。在本分析范例中，小组使用一计量型测量系统来评价零件，并应用其结果来确定其参考决定。

使用新的信息建立了另一组交叉表，以便将每个评价人与参考决定进行比较。

A*参考交叉表

			参考		总计
			.00	1.00	
A	.00	数量	45	5	50
		期望的数量	16.0	34.0	50.0
	1.00	数量	3	97	100
		期望的数量	32.0	68.0	100.0
总计		数量	48	102	150
		期望的数量	48.0	102.0	150.0

B*参考交叉表

			参考		总计
			.00	1.00	
B	.00	数量	45	2	47
		期望的数量	15.0	32.0	47.0
	1.00	数量	3	100	103
		期望的数量	33.0	70.0	103.0
总计		数量	48	102	150
		期望的数量	48.0	101.0	150.0

C*参考交叉表

			参考		总计
			.00	1.00	
C	.00	数量	42	9	51
		期望的数量	16.3	34.7	51.0
	1.00	数量	6	93	99
		期望的数量	31.7	67.3	99.0
总计		数量	48	102	150
		期望的数量	48.0	102.0	150.0

小组也计算出了 kappa 值以确定每个评价人与参考决定之间的一致性。

	A	B	C
kappa	.88	.92	.77

以上数据可被解释为每个评价人与标准之间有很好的一致性。

然后这过程小组计算这测量系统的有效性。

$$\text{有效性} = \frac{\text{作出正确决定的次数}}{\text{总决定次数}}$$

	评价人%[1]			结果%与归因的比较[2]		
变差来源	评价人 A	评价人 B	评价人 C	评价人 A	评价人 B	评价人 C
总检查数	50	50	50	50	50	50
相配数	42	45	40	42	45	40
错误的拒收（由于评价人偏倚造成的拒收）				0	0	0
错误的接受（由于评价人偏倚造成的接受）				0	0	0
不相配				8	5	10
95% UCI	93%	97%	90%	93%	97%	90%
计算所得的结果	**84%**	**90%**	**80%**	**84%**	**90%**	**80%**
95% LCI	71%	78%	66%	71%	78%	66%

	系统有效结果%[3]	系统有效结果%与参考的比较[4]
检查总数	50	50
一致的数量	39	39
95% UCI	89%	89%
计算所得的结果	**78%**	**78%**
95% LCI	64%	64%

注

(1) 在所有测量中，评价人本身是一致的。
(2) 评价人对所有测量与已知的标准一致。
(3) 所有评价人本身与其它人之间是一致的。
(4) 所有评价人本身与其它人之间一致，并与参考值一致。
(5) UCI 和 LCI 分别为置信区间边界的上限和下限。

每对评价人之间多次试验的假设可用零假设来表示：

H_0：两位评价人一致的有效性。

由于每位评价人结果的计算值均落在其它人的自信度区间内，小组决定不能拒绝零假设。这结论进一步证实了 *kappa* 测量得到的结论。

为进一步分析，小组中的一名成员得出下列表格，为每个评价人的结果提供指南：

决定 测量系统	有效性	错误率	错误警报率
评价人可接受的条件	≥ 90%	≤ 2%	≤ 5%
评价人可接受的条件 - 可能需要改进	≥ 80%	≤ 5%	≤ 10%
评价人不可接受的条件 - 需要改进	< 80%	> 5%	> 10%

对他们所已经得到的所有信息进行汇总，小组得到以下结论：

	有效性	错误率	错误警报率
A	84%	6.3%	4.9%
B	90%	6.3%	2.0%
C	80%	12.5%	8.8%

这些结果显示：各个评价者对于该测量系统，在有效性、错误率与错误警报率上都有不同程度的结果；在所有三个项目中，没有一位评价者是可被接受的。是否需要为这过程更改其接受标准？这些风险可以被接受吗？评价者是否需要更好的培训？测量的环境可不可以被改善？重要的是：顾客对这测量系统与其研究结果会有何看法？顾客原本预期的情况是什么？

顾客是否接受这些风险。

关注点：

1) 关于可接受的风险，并没有以理论为基础的决策准则。以上指南是探索性的，并且是基于怎样才是"可接受"的个别"信念"下所发展的。最终的决定准则应该取决于对后续过程及最终顾客的影响（如风险）。这是一个客观事物的决定 — 而不是统计上的决定。

2) 上述分析是以数据为依据的。例如：如果过程能力指数为 *Pp* = *Ppk* = 1.00，则所有的结论将可能正确，因为不会有零件落在测量系统的区域C（“灰色”区域）中。

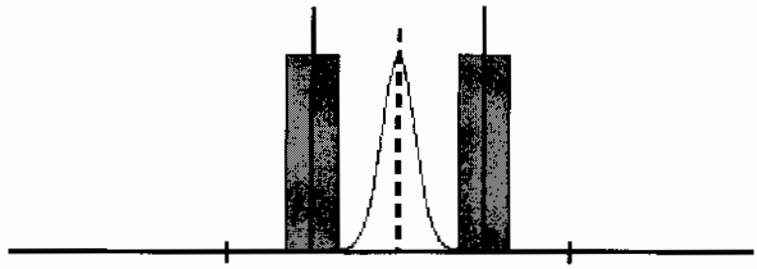


图 30：*Pp* = *Ppk* = 1.00 的过程范例

在这新的情况下，可以得出这样的结论：所有评价人都是可被接受的，因为将不会有决定误差。

3) 通常对于交叉的结果的实际意义有一误解。以评价人B的结果为例：

B * 参考交叉表

			参考		总计
			.00	1.00	
B	.00	数量	45	2	47
		在参考值内的%	93.8%	20.%	31.3%
	1.00	数量	3	100	103
		在参考值内的%	6.3%	98.0%	68.7%
总计		数量	48	102	150
		在参考值内的%	100.0%	100.0%	100.0%

由于检验的目的在于找出所有的不合格零件，许多人视左上处为一个测量找到坏零件的有效性。这百分比表示将已经是坏的零件判定为坏零件的可能性：

$$\Pr(\text{称为坏零件} \mid \text{一个坏零件})$$

假设过程已经被改进到 *Pp* = *Ppk* = 1.00，生产者关心的概率是：

$$\Pr(\text{零件是坏的} \mid \text{被称为坏的})$$

从上面的数据中确定以上结果，必须应用 Baye 的理论。

$$\Pr(\text{坏的}|\text{称为坏的}) = \frac{\Pr(\text{称为坏的}|\text{坏的}) * \Pr(\text{坏的})}{\Pr(\text{称为坏的}|\text{坏的}) * \Pr(\text{坏的}) + \Pr(\text{称为坏的}|\text{好的}) * \Pr(\text{好的})}$$

$$\Pr(\text{坏的}|\text{称为坏的}) = \frac{.938*(.0027)}{.938*(.0027)+.020*(.9973)}$$

$$\Pr(\text{坏的}|\text{称为坏的}) = .11$$

也就是说，以上这些结果指出如果某零件被判定成是坏的，实际上它只有十分之一的可能是坏的。

4) 这种分析不必使用计量型数据，即使参考决定值已被确定且为可获得时，也不需要安排这相关的资料。

信号检查 (Signal Detection) 方法

参考值	代码	参考值	代码
0.599581	–	0.503091	+
0.587893	–	0.502436	+
0.576459	–	0.502295	+
0.570360	–	0.501132	+
0.566575	–	0.498698	+
0.566152	–	0.493441	+
0.561457	x	0.488905	+
0.559918	x	0.488184	+
0.547204	x	0.487613	+
0.545604	x	0.486379	+
0.544951	x	0.484167	+
0.543077	x	0.483803	+
0.542704	+	0.477236	+
0.531939	+	0.476901	+
0.529065	+	0.470832	+
0.523754	+	0.465454	x
0.521642	+	0.462410	x
0.520496	+	0.454518	x
0.519694	+	0.452310	x
0.517377	+	0.449696	x
0.515573	+	0.446697	–
0.514192	+	0.437817	–
0.513779	+	0.427687	–
0.509015	+	0.412453	–
0.505850	+	0.409238	–

另一种替代的方法 — 信号检查理论来确定区域II的近似宽度，从而确定测量系统的 *GRR*。

让 d_i = 被所有评价人接受的最后一个零件与被所有评价人拒收的第一个零件之间的距离（对于每个规范）。

然后，

$d = d_i$ 的平均值

这是区域II宽度的估计值[61]，且 *GRR* 的估计值为 $5.15*\sigma_{GRR}$.

在本例中：

d_{LSL}=0.470832-0.446697=0.024135

d_{USL}=0.566152-0.542704=0.023448

d = 0.0237915

或者，%*GRR* 的估计值为：

%*GRR* = 29%

实际%GRR = 25%，因此这估计值将导引至测量系统的相同评价。

如果只能得到已整理过的数据信息，仍可用本方法，但需要相关知识来量化 *d* 值。

[61] 本估计值的“好”取决于样本容量及样本代表这过程有多接近。样本容量越大，对估计越好。

[61] 本估计值的“好”取决于样本容量及样本代表这过程有多接近。样本容量越大，对估计越好。

分析方法[62]

就像对于任何测量系统一样，应该验证过程的稳定性，如有需要，还应该进行监控。对于计数型测量系统，通常采用恒定样本的长期计数型控制图来验证其稳定性[63]。

对于一计数型测量系统，量具性能曲线（Gage PerformanceCurve，参见第四章，第C节）的概念是被用来展开一个测量系统研究，以评估这测量系统的重复性和偏倚的数值。这分析可用于单边限值和双边限值的测量系统；对于双边限值测量系统，只有一个限值需要用来检查差异的线性和一致性的假设；为方便起见，用其下限值进行讨论。

通常，计数型测量系统研究由获得了多个被选定零件的参考值所构成。应该记录每个零件被评价的次数 (m)，以及接受的总数 (a)。从这些结果中，可以评估重复性和偏倚。

计数型研究的第一步是选取零件。有必要已知研究中每个零件的参考值。应该按实际情况尽量等间隔地选取 8 个零件，其中的最大值和最小值应该能代表过程范围。尽管这种抽样方法不会影响结果的自信度，但它会影响为完成这量具研究所需的全部零件数量。8 个零件必须用量具测量，记录 $m = 20$ 次，接受的数量 (a)。

为完整的研究，最小的零件必须 $a = 0$，最大的零件必须 $a = 20$；其它 6 个零件则在 $1 \leq a \leq 19$。如果不满足这些准则，必须用该量具测量更多已知参考值 (X) 的零件，直到满足上述条件。如果最小的值 $a \neq 0$，则选取更小的零件并评价直到 $a = 0$；如果最大的值 $a \neq 20$，则选取更大的零件直到 $a = 20$。如果其它六个零件不满足 $1 \leq a \leq 19$，则在全范围内选取额外的零件，可以选择早已在研究中被测零件在中间的那些点。在 $a = 0$ 端，第一个间隔起始于 $a = 0$ 的最大测量处；在 $a = 20$ 端，第一个间隔始起于 $a = 20$ 的最小测量处。为了获得最好的结果，抽样应该在 $a = 0$ 和 $a = 20$ 两端进行，然后朝零件范围的中间选取；必要时可以重复上述程序直到满足准则要求。

---

[62] 经允许摘自 *"Analysis of Attribute Gage System"*, J. Mcaslin & G. Gruska, ASQC, 1976.

[63] 警告：$np > 4$

一但满足了数据收集准则后，必须按以下公式计算各个零件的接收概率：

$$P_a' = \begin{cases} \dfrac{a+0.5}{m} & if\ \dfrac{a}{m} < 0.5, \quad a \neq 0 \\ \dfrac{a-0.5}{m} & if\ \dfrac{a}{m} > 0.5 \quad a \neq 20 \\ 0.5 & if\ \dfrac{a}{m} = 0.5 \end{cases}$$

以上调整涵盖了 $1 \leq a \leq 19$ 的情况。对这情况，当 $a = 0$ 时，$P_a' = 0$ 除去 $a = 0$ 时的最大参考值，此时其 $P_a' = 0.025$；当 $a = 20$ 时，$P_a' = 1$，除去 $a = 20$ 时的最小参考值，此时其 $P_a' = 0.975$。

一旦每个 X_T 的 P_a' 被计算出来后，就可以建立量具性能曲线(*GPC*)。尽管 *GPC* 可以用图形表示（见图 32），但用常态概率图（见图 31）可以产生更准确的重复性和偏倚估计值。

将计算得到的概率绘制在常态分布概率图上，画出经过这些点的一条最佳适合直线。偏倚等于下限值减去 $P_a' = 0.5$ 对应的测量参考值。

或

$$偏倚 = 规范下限值 - X_T（在 P_a' = 0.5）$$

发现 $P_a' = 0.995$ 和 $P_a' = 0.005$ 之间对应的测量参考值的差异除以一个 1.08[64] 的调整系数即为重复性。

$$重复性 = \frac{X_T(P_a' = 0.995) - X_T(P_a' = 0.005)}{1.08}$$

为了确定偏倚是否明显的偏离零，采用下列统计量：

$$t = \frac{31.3 \times |偏倚|}{重复性}$$

如果这计算值明显大于 $2.093(t_{0.025,19})$，则这偏倚明显的偏离零。

下面将用一个范例来阐述计数型研究的数据收集和重复性及偏倚的计算。

---

[64] 1.08 调整系数是样本容量为 20 时的特定值，它是通过一个模拟的方法所确定。

范例：

用一个计数型的量具测量公差为 ±0.010 的尺寸。这量具是一个生产线末端100%自动检查的量具，它已受到重复性和偏倚的影响。为了完成这计数型研究，用该量具测量 8 个零件，每个零件各测量 20 次，这 8 个零件的参考值为从 -0.016 至 -0.002，间隔为 0.002。各零件的接受次数为：

X_T	a
-0.016	0
-0.014	3
-0.012	8
-0.010	20
-0.008	20
-0.006	20
-0.004	20
-0.002	20

由于只有 2 个参考值位于 $1 \le a \le 19$，必须至少再找 4 个零件；因此，需要测量那些参考值在现有间隔的中间处的零件，这些参考值及接受次数为：

-0.015	1
-0.013	5
-0.011	16

现在有 5 个参考值位于 $1 \le a \le 19$。程序要求再找到一个零件是可以位于 $1 \le a \le 19$；为此，评价了以下零件：

-0.0105	18

现在已经满足了数据收集准则，采用二项式调整可以计算它们的接受概率，如下所示：

X_T	a	P_a'
-0.016	0	0.025
-0.015	1	0.075
-0.014	3	0.175
-0.013	5	0.275
-0.012	8	0.425
-0.011	16	0.775
-0.0105	18	0.875
-0.010	20	0.975
-0.008	20	1.000

将这些概率画在常态概率图上，如图 31 所示，画出一条经过这些点的最佳适合直线，可确定偏倚和重复性。偏倚等于下限值减去 P_a' = 0.5 对应的测量参考值。

由图 31 可以得到：

偏倚 = -0.010 — (-0.0123) = 0.0023

重复性等于发现 P_a' = 0.995 和 P_a' = 0.005 对应的测量参考值之差值，再除以 1.08，从图 31 可以得到：

$$R = \frac{|-0.0163| - |-0.0084|}{1.08}$$

$$= \frac{0.0079}{1.08} = 0.0073$$

确定偏倚是否明显的偏离零，计算如下：

$$t = \frac{31.3 \times |\text{偏倚}|}{R}$$

$$= \frac{31.3 \times |0.0023|}{0.0073} = 9.86$$

由于这值大于 $t_{0.025,19}$ = 2.093，可见这偏倚明显的偏离零。

像第五章 C 节所示的计量型量具性能曲线一样，计数型 *GPC* 也可以绘制在空白的图纸上（见图 32），这可用两种方法中的任何一种方法来完成；一种方法将可以在另一个规范限值处进行抽样研究。本例中，计数型研究的大样法 (long method) 也必须在规范的上限处进行研究，并画出相应的计算值。

但是，基于前面提到的假设，将不需要进行再次研究，因为在上限值的曲线形状应该与在下限值的曲线成“对称影像”；唯一需要考虑的是对应于 X_T 值的曲线位置，该位置可由偏倚来确定。

曲线正确位置将被定义在 P_a' = 0.5 的点，且 X_T 值等于规范限值减去偏倚。在本例中，该点将是 X_T = 0.010 — 0.0023 = 0.0077。按这种方法绘制的 *GPC* 如图 31 所示。

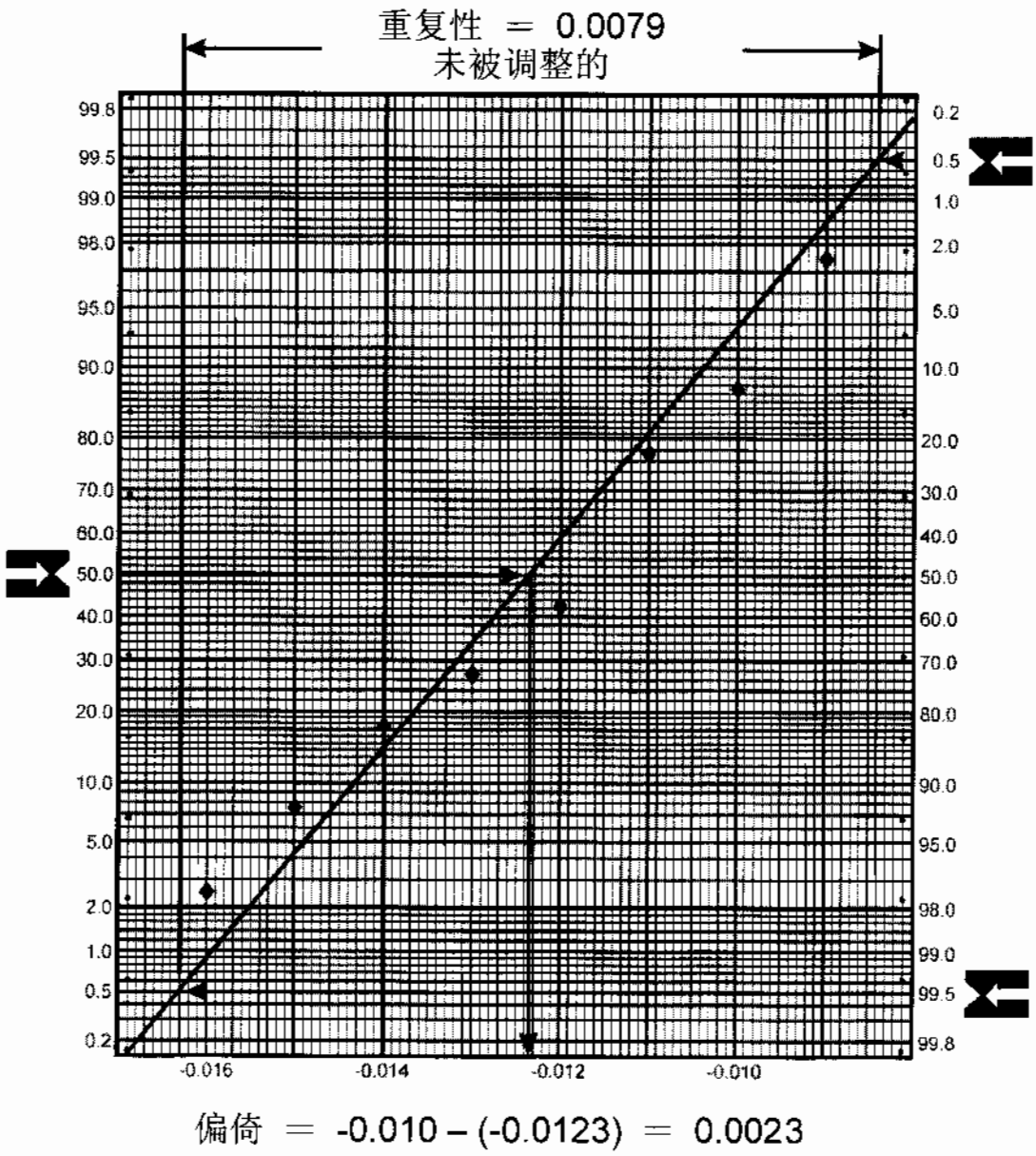


图 31：在常态概率图纸上描绘的计数型量具性能曲线

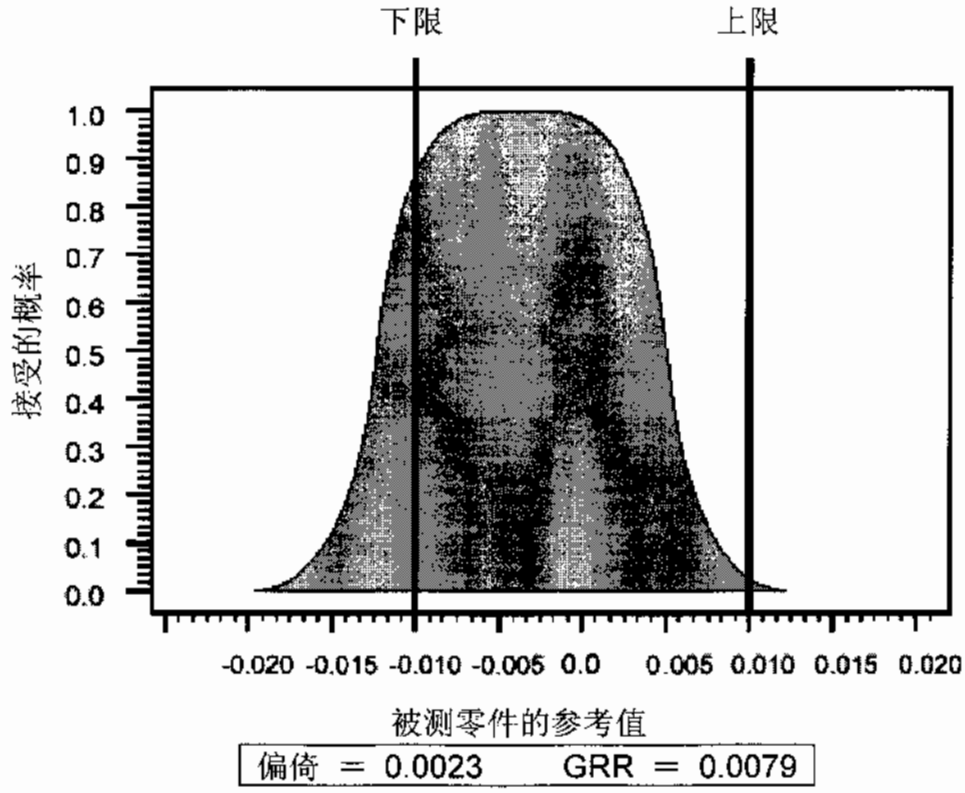


图 32：计数型量具性能曲线

第四章

实施复杂的测量系统

第四章 — 第A节

实施复杂的或不可重复的测量系统

引言

本参考手册的重点是对每个零件的读值能被重复的测量系统；但并不是所有的测量系统都有这种特性。例如：

- 破坏性测量系统
- 零件每次随使用／试验会有变更的系统，例如：发动机或变速箱动力计试验。

下面是一些测量系统分析方法的例子，包括本手册中前面未被讨论的那些系统。这里不打算全面列出涵盖各种型式的测量系统，而只是提出不同方法的一些例子。如果本手册不适用于你所具有的测量系统，建议你求助于合适的统计资源。

不可重复的测量系统	
情景 — 非破坏性测量系统	范例
在测量过程中零件不会被改变；即测量系统是非破坏性的，并可使用具有下列属性的零件（样本）： • 静态的属性，或 • 已经稳定的动态（变化）属性	• 用于不是新的车辆／动力系统进行测功试验 • 有计量型数据的泄漏试验
已知特性（属性）的定性期限，并能延续至超过期望的研究期间 — 即经过一段预期的使用后，被测特性不会改变。	• 从单一材料批次中抽取样本进行质量的光谱测量
情景 — 破坏性测量系统	范例
试验台	• 生产线末端 ✓ 发动机试验台 ✓ 变速箱试验台 ✓ 整车动力试验机 • 有定质数据的泄漏试验 • 盐雾试验／湿度试验 • 重力计
其它不可重复的测量系统	• 采用不允许重复的自动线上测量系统 • 破坏性的焊接测试 • 破坏性的电镀测试

表 13：测量系统举例

本章节描述的研究方法及不同的情况如下表所示：

稳定性研究					
情景	S1	S2	S3	S4	S5
在测量过程中零件不会被改变；即测量系统是非破坏性的（可重复），并且使用具有下列特性的零件（样本）： • 静态的属性，或 • 已经稳定的动态（变化）属性	√	√			
已知特性（属性）的定性期限，并能延续至超过期望的研究期间 — 即经过一段预期的使用后，测量特性不改变。	√	√			
破坏性测量系统			√	√	
不可重复的测量系统			√	√	
试验台					√

变差研究									
情景	V1	V2	V3	V4	V5	V6	V7	V8	V9
在测量过程中零件不会被改变；即测量系统是非破坏性的（可重复），并且使用具有下列特性的零件（样本）： • 静态的属性，或 • 已经稳定的动态（变化）属性	√								
上述情况的 $p \geq 2$ 仪器		√							
破坏性测量系统			√	√					
不可重复的测量系统			√	√					
具有动态特性的测量系统，例：试验台			√	√	√	√	√	√	
上述情况的 $p \geq 3$ 仪器									√

表 14：基于不同类型的测量系统使用的方法

第四章 — 第 B 节

稳定性研究

S1：单一零件[65]，每次实施单一测量

应用：

a) 零件不会由测量过程而被改变的测量系统；即测量系统是非破坏性的，并可使用具有下列属性的零件：

- ✓ 静态的属性，或
- ✓ 已经稳定的动态（变化）属性。

b) 已知该特性（属性）的定性期限，并延续超过预期的研究期间；就是说，在预期的使用期间中，被测特性不发生改变。

假设：

- 已知（已有文件说明）该测量系统在预期的特性（属性）范围内，具有一线性的回应。
- 零件（样本）涵盖该特性的过程变差的预期范围。

采用 *X & mR* 图表分析：

- 确定测量系统的稳定性
  - ✓ 将描绘的点和控制限值进行比较，
  - ✓ 寻找趋势（仅 *x* 图）。
- 将 $\sigma_e = \overline{R}/d_2^*$（总测量误差）和由变差研究（见下一章节）得到的重复性估计值 σ_E 来进行比较
- 如果已知参考值，则确定偏倚：

$$偏倚 = \bar{x} - 参考值。$$

[65] 如果对这过程是适用的话，可以使用一参考标准。

S2：$n \geq 3$ 个零件[66]，每次每个零件实施单一测量

应用：

a) 零件不会由测量过程而被改变的测量系统；即测量系统是非破坏性的，并可使用具有下列属性的零件：

✓ 静态的属性，或

✓ 已经稳定的动态（变化）属性。

b) 已知该特性（属性）的定性期限，并延续超过预期的研究期间；即在预期的使用期间中，被测特性不发生改变。

假设：

- 已知（已有文件说明）这测量系统在预期的特性（属性）范围内，具有一线性回应。
- 零件（样本）涵盖该特性的过程变差的预期范围。

采用 [z, R] 图表分析：式中 $z_i = x_i - \mu_i$
且 μ_i 是（参考）标准值，或者由对零件（样本）大量连续的读值平均值所确定的值。

- 确定测量系统的稳定性：
  - ✓ 将描绘的点和控制限值进行比较
  - ✓ 寻找趋势（仅 z 图）。
- 将 $\sigma_e = \overline{R}/d_2$ 与由变差研究所得出的重复性估计 σ_E 进行比较[67]。
- 若已知参考值，则确定偏倚：

  $$偏倚 = \bar{\bar{x}} - 参考值$$

- 若采用了 $n \geq 3$ 个零件，则确定线性：
  - ✓ 零件（样本）必须涵盖特性的预期范围。
  - ✓ 各零件（样本）应该被单独分析其偏倚和重复性。

[66] 如果该过程适用的话，可以使用一参考标准。

✓ 使用第 3 章，第 B 节所讨论的线性分析法来量化线性。

如果在本研究中，使用了一个以上的仪器，要确定这些装置的一致性（变差的同质性）。例如，使用 *F* 试验、Bartlett 试验、Levene 试验[68]等。

S3：从稳定过程中的大量样本

应用：

测量系统必须用*同质的独立可识别分布 ("independent identically distributed, iid")* 样本（收集并隔离）来评价。各个零件（样本）的测量不被重复进行，因此可以用破坏性的和不可重复的测量系统进行研究。

假设：

- 已知特性（属性）的定性期限并延续超过预期的研究期间，即在预期的使用和 / 或储存期间，被测特性不改变。
- 这些零件（样本）涵盖了该特性（属性）的过程变差的预期范围。
- 已知（已由文件说明）测量系统的线性涵盖了特性（属性）的预期范围。（如果这回应是非线性的，则必须相应地校正读值。）

进行分析：

- 通过对 $n \geq 30$ 个零件的能力研究，以确定总变差。（这种初步研究也应该被用来验证样本的一致性，即所有零件（样本）来自单边形式的分布。）
- $\sigma^2_{总} = \sigma^2_{过程} + \sigma^2_{测量系统}$
- 测量从每个时间周期的隔离样本中的一个或多个零件，使用带有控制限的 $\bar{x}$ & *R* 或 *x* & *mR* 图，这控制限值由能力研究所确定。
- 将描绘点和控制限值进行比较。
- 寻找趋势。
- 因为这些零件（样本）不会变化（一个隔离的样本），任何不稳定性迹象将归因于该测量系统的变化。

[68] *Dataplot*, National Institute of Standards and Technology, Statistical Engineering Divison(www.itl.nist.gov).

S4：分割样本（一般），每次采用单一样本

应用：

各个零件（样本）的测量部分不被重复，故可以采用破坏性或不可重复的测量系统进行研究。

假设：

- 已知特性（属性）的定性期限，并延续超过预期的研究期间；即在整个预期的使用和 / 或储存期间，被测特性不会改变。
- 零件（样本）涵盖该特性（属性）的过程变差的预期范围。
- 这些样本被分割成 *m* 部份。*m* = 2 部份时常被称为 *试验 — 再试验研究 (test-retest study)*。

用以下方法分析：

- 极差图，以便追踪测量的一致性（推翻“批量内部”一致性）。
- 将 $\sigma_e = \overline{R}/d_2^*$ 与由变差分析所得到的重复性估计 σ_E 进行比较。
- 这是上边界 (upper bound) 的研究：
  $\sigma_e^2 = \sigma^2_{\text{重复性估计}} + \sigma^2_{\text{样本之间}}$
- 用图表以便追踪生产过程的一致性。

S4 用于从不同批次中采用连续的（同质的）配对零件 — S4a

本研究与 S4 一样，但是是从不同批次所抽取的同质性的零件。

这是一种上边界研究：$\sigma_e^2 = \sigma^2_{\text{重复性估计}} + \sigma^2_{\text{样本之间}} + \sigma^2_{\text{批次}}$

S5：试验台

在这种情况下，由多台测量仪器（试验台）评价连续生产产品的同一特性；这些产品是随机的被分配到各试验台。

S5a：计数型数据回应

采用 *p* 图进行分析：

- 确定试验台之间（判断的）一致性：使用包含各试验台结果的一张图表
- 确定各个试验台内的稳定性：对每个试验台使用单独的图表。

采用 $\bar{p}$ & *mR* 图分析整个系统的稳定性，这里 $\bar{p}$ 是在某个设定工作日内所有试验台的平均值。

S5b：计量型数据回应

采用变异数分析 (*ANOVA*) 和图示技术分析[69]

- 以时间周期计算每个试验台（依特性）的 $\bar{x}$ &*s*。
- 确定试验台之间的一致性：一张包含所有试验台结果的 $\bar{x}$ & *s* 图。
- 确定各个试验台内的稳定性：每个试验台单独的 $\bar{x}$ &*s* 图。
- 对试验台之间的一致性（变差的同质性）定量；例如，采用 *F* 试验、Bartlett 试验、Levene 试验。
- 通过与试验台平均值的比较，确定所有试验台是否在同一目标上；例如，采用一单向 *ANOVA* 分析。如果存在任何差别，通过采用如 Tukey 的 *T* 试验来隔离不同的试验台。

[69] 也可参考 James, P. D., *"Graphical Displays of Gage R&R Data"*, AQC Transaction , ASQC, 1991

第四章 — 第 C 节

变差研究

所有描述的研究，是在研究期间对测量系统（包括环境的后果）本质的计算。由于测量系统是被用于作出对产品、过程或服务的未来判定，所以就需要对测量系统作出*分析的*结论。从计算到分析结果的转化需要专业知识和专家，以能：

- 确保某研究的设计和实施已考虑到所有预期的测量变差来源；
- 根据预期的用途、环境、控制、保养等，分析这些结果（数据）。

V1：标准 GRR 研究

这些研究已被包含在本参考手册中。
这些研究包括图形分析和数值分析。

V1a — 极差法 (*R&R*)

V1b — 极差法（*R&R* 和零件内部）

V1c — *ANOVA* 法

V1d — 修改的 *ANOVA* / 极差法

V2：用 $p \geq 2$ 台仪器的多个读值

这允许对多台仪器进行比较。

应用：

a) 零件不会由测量过程而被改变的测量系统；即测量系统是非破坏性的，并可使用具有下列属性的零件（样本）：

   1) 静态的属性，或
   2) 已经稳定的动态（变化）属性。

假设：

- 已知特性（属性）的定性期限，并延续超过预期的研究期间；即在整个预期的使用和 / 或储存期间，被测特性不会改变。
- 零件（样本）涵盖了该特性（属性）的过程变差的预期范围。

采用 Grubb[70]（或 Thompson[71]）估计的分析

- 过程变差性；
- 仪器变差性 = 重复性
- 自信度范围的计算是可获得的。

V3：分割的样本 ($m = 2$)

应用：

个别的零件（样本）部分的测量不可被重复，因此可用破坏性和不可重复的测量系统进行研究，并且可被使用在动态特性的测量系统分析。

假设：

- 已知特性（属性）的定性期限，并延续超过预期的研究期间；即在整个预期的使用和 / 或储存期间内，被测特性不会改变。
- 零件（样本）涵盖了该特性过程变差的预期范围。
- 这些样本被分割成 m 部份，当 $m = 2$ 时通常被称为*试验 — 再试验研究*。

采用回归技术分析：

- 用误差术语估计重复性：$\sigma_{\text{重复性估计}} = \sigma_e$
- 线性（通过估计线和 45°线进行比较）。

V3a — V3 用于连续的配对零件

本研究与 V3 相同，但使用连续零件的配对，而不是分割样本。本研究应用在当不破坏被测量的属性，该零件就不能被划分的情况。

这是一个上边界分析：$\sigma_{\text{重复性估计}} \leq \sigma_e + \sigma_{\text{样本之间}}$

---

70 参见参考文献表中第 15 项

71 参见参考文献表中第 38 项

V4：分割样本（一般）

应用：

个别的零件（样本）部分的测量不可被重复，因此可用破坏性和不可重复的测量系统进行研究，并且可被使用在动态特性的测量系统分析。

假设：

- 已知特性（属性）的定性期限，并延续超过预期的研究期间；即在预期的使用和／或储存期间，被测特性不会改变。
- 零件（样本）涵盖了该特性（属性）的过程变差的预期范围。
- 把样本分割成 m 等份，这里 m 为 2 或 3 的倍数；$m \geq 2$（如 $m = 3, 4, 6, 9.....$）。

采用以下方法进行分析：

- 包括图形法的标准 GRR 技术
- *ANOVA* — 随机化的分组设计（双向 *ANOVA*）

V4a — V4 用于从不同批次中的连续的（同质的）配成对零件

本研究与 V4 研究相同，但使用连续零件而不是分割样本。本研究应用在当不破坏被测量的属性，该零件就不能被划分的情况。

这是一个上边界研究：$\sigma_{\text{重复性估计}} \leq \sigma_e + \sigma_{\text{零件}} + \sigma_{\text{批次}}$

下列研究假设零件（样本）特性（属性）是动态的。

V5：与 V1 相同，使用稳定的零件

本研究中使用基于工程知识和专门技术为基础的稳定过程的零件；例如，磨合过的发动机对新发动机的比较。

V6：时间序列分析

假设：

- 在规定的时间间隔重复的读值。
- 已知特性（属性）的定性期限，并延续超过预期的研究期间；即在预期的使用和／或储存期间，被测特性不会改变。
- 零件（样本）涵盖了该特性（属性）的过程变差的预期范围。

通过确定了每个样本的退化模型来分析：

- $\sigma_{\text{重复性估计}} = \sigma_e$
- 退化的一致性（如果 $n \geq 2$）

V7：线性分析

假设：

- 在规定的时间间隔内重复的读值。
- 已知（已有文件说明的）在规定的时间间隔内，这测量系统中的退化具有线性回应。
- 已知特性（属性）的定性期限，并延续超过预期的研究期间；即在预期的使用和／或储存期间，被测特性不会改变。
- 零件（样本）涵盖了该特性（属性）的过程变差的预期范围。

用线性回归来分析：

- $\sigma_{\text{重复性估计}} = \sigma_e$
- 退化的一致性（如果 $n \geq 2$）

V7a — V7 用于同质性的样本

通过线性回归来分析：

- 这是一个上边界分析：$\sigma_{\text{重复性估计}} \leq \sigma_e + \sigma_{\text{零件之间}}$

V8：时间与特性（属性）退化的关系

可以修改 V6 和 V7，以确定这退化是否基于于时间（即定性期限）或活动。

V9：V2 用于同时存在的多个读值，且 $p \geq 3$ 台仪器

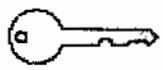
这分析与 V2 相同（另参见 Lipson& Sheth, sec 13.2）。

第五章

其他测量概念

第五章 — 第 A 节

识别过大的零件内部变差的影响



了解一测量系统的变差来源对于所有测量的应用是很重要的，特别是当有一个显著的零件内部变差时。零件内部变差，例如锥度或失去真圆度，会造成测量系统的评价提供误导的结果；这是因为未考虑零件内部变差影响了对重复性、再现性或两者的估计。也就是说，零件内部变差可能成为测量系统变差的重要要素。对目前的工作来说，了解产品中呈现的零件内部变差，将比了解测量系统的适用性有更大的意义。

可能遭遇到的零件内部变差的例子有：真圆度（失去真圆度）、中心度、锥度、平面度、表面粗糙度、圆柱度等[72]。在同一零件中同时存在一个以上的这些特性（组合的误差）是可能的，每个特性的强度和他们的相互依赖关系可能包含在数据中，并且影响对测量系统的理解。在测量系统分析过程中最主要的，是识别和分析这些额外的变差来源，以理解测量系统的真实变差。在这种情况下，用一个D.O.E.、ANOVA 或其它更复杂的统计技术进行分析可能是必要的。不论采用哪种技术，均应该征得顾客供方质量代表的核准。

另外，产品设计工程师所做的决定，会对如何进行零件测量、如何设计夹具产生很大的影响，而且其结果会影响最终的测量误差。例如：塑胶产品的分模线是一很关键的特性（两个模穴的模具通常有多余的塑胶材料注入，因此分模线是一个不好控制的表面）。这些因素要在设计 FMEA 中加以考虑。

一但理解了这零件内部变差要素，就有可能在测量系统内控制这些因素（例如，重新设计量具或使用不同夹紧方法／设备），以使得未来的数据不致被混淆。

---

[72] 许多这些特性在设计时通过*几何尺寸及公差 (GD&T)* 的方法来进行控制。GD&T 通过了一种*功能性*的方法提供一可操作的规定方法来检测零件。通常，功能性检验是一种计数型检验。当计量型数据是必要的，就会产生用一个为功能性检验所设计的量具来测量产生计量型数据的问题；在这种情况下，有时可能通过将这功能性量具作为一种夹紧装置来进行一项 CMM 研究。然而，当使用这种方法时，一定要确保这夹紧装置能紧紧地夹住零件，而且能在某些场所重复实施（如果不能，这 MSA 研究结果就会产生这种零件内部的误差）。

第五章 — 第 B 节

平均值和极差法 — 额外的处理

引言

关于测量系统评估的平均值和极差法，有一附加考虑值得一提。一控制图的范例经允许摘自于 Wheeler 和 Lyday 所着的"Evaluating the Measurement Process"（见参考文献）。

如其它的优秀测量系统分析设计一样，这种图形方法的主要目的是：确定某测量过程是否足以测量生产过程的变差和 / 或评价产品性能。

- 所有量具在做相同工作吗？
- 所有评价人在做相同工作吗？

顾及过程变差，测量变差是否可以接受？

- 从这测量过程所获得的数据有多好？或这数据能划分成多少非重迭的组或类别？

程序的步骤

1) 应该小心地按第二章第 C 节"测量系统研究的准备"进行；

2) 让每位评价人测量将被研究其特性的每个样本。在控制图最上面的数据行中记录第一次检查结果（见图 33）。

3) 重复测量并在控制图上第二行数据栏中记录数据。（*注：当进行这第二次测量时，不允许评价人看到他们的原始读值。*）现在由每位评价人对同一零件应该有两次测量数据。

4) 通过计算每个子组的平均值 ($\overline{X}$) 和极差值 (R)，分析这数据。

5) 在极差图上描绘极差值并计算平均极差 ($\overline{R}$)（包括所有评价人的所有子组极差 (R)）；在图上画出该极差的平均值。使用 $n = 2$ 的 D_4 常数以计算这极差图的控制限值。画出这限值并确定所有数值是否都受控。

   ✓ 如果所有极差均受控，则所有评价人在做同样的工作。

   ✓ 如果一位评价人失控，则他的方法不同于其它人。

   ✓ 如果所有评价人均有一些失控的极差值，则该测量系统对评价人技术是敏感的，且需要改进以获得有用的数据。

6) 下一步，在平均值图上描绘所有评价人的各个子组平均值 ($\overline{X}$)（见图 33）。该平均值代表了零件变差和测量变差。

计算总平均值 ($\overline{\overline{X}}$)（包括所有评价人的所有子组平均值 ($\overline{X}$)）；在图上画出该总平均值 ($\overline{\overline{X}}$)。

现在，用 $n = 2$ 的 A_2 常数和极差图上的平均极差 ($\overline{R}$)，计算该图的控制限值，并在平均值图上画出这些限值。注意本项研究中的极差图只包含测量变差；因此，平均值的控制限之间的面积代表在这系统的测量变差量。

如果所有平均值均落在控制限值内，则这测量变差蒙蔽了该过程变差。换句话说，这测量过程比生产过程有更大的变差，监控和控制该生产过程没有价值。

如果少于一半的平均值落在控制限值外，则这测量系统不适用于过程控制。

相反的，如果大多数的平均值落在控制限值外，这表明来自生产过程的信号大于测量变差，这测量系统能提供用于控制该过程的有用数据。

计算表举例

问题 — “这测量系统所收集的数据有多好？”完成了图 34a 的计算表就能得到答案。这计算表所需的所有数据能从上述平均值和极差图上找到。

以下是用于完成计算表范例（图 34a）的步骤：

1) 注明将被评价的测量和特性，谁是实施这研究的执行者和研究的日期。

2) 直接从控制图上获得子组平均极差 ($\overline{R}$)。

3) 重复次数 (r) 是每位评价人检查同一零件的次数。

4) 如图 34 所示，采用 r 所对应的 d_2 值来计算估计的重复性标准差 ($\hat{\sigma}_e$)。

5) 在提供的空白处填入评价人人数 (n_A)。

6) 在提供的空白处填入样本数量 (n)。

7) 平均每位评价人所获得的所有样本，计算每位评价人的平均值，并把这些平均值填入提供给每位评价人 (*A*、*B*、*C*) 的空白处。

8) 检查每位评价人 (*A*、*B*、*C*) 的平均值，由最高值减去最低值来确定评价人平均值的极差，并填入 (R_A)处。

9) 采用 n_A 值所对应的 d_2^* 值，计算估计的评价人标准偏差 $\hat{\sigma}_A$。

10) 通过平均每个样本的所有评价人所获得的数值，计算样本平均值。例如，(*评价人 A 的样本 1 平均值 + 评价人 B 样本 1 平均值 + 最后一位评价人样本 1 的平均值，其总和除以评价人人数*)。这是样本真值的最佳估计。在图 34b 中所提供的空白处 (1、2、3、……、9、10) 填入每个样本的平均值。

11) 观察样本平均值 (1、2、3、……、9、10)，并通过最高值减去最低值，计算样本平均值的极差 (R_p)。把该值填入提供的空白处。

12) 采用 *n* 值（*n* = 样本数量）对应的 d_2^* 值，计算估计的样本间标准偏差 ($\hat{\sigma}_p$)。

13) 通过样本标准差除以测量标准差，以计算 *"扰噪比 (Signal to Noise Ratio)" (SN)*，并填入提供的空白处。

$$SN = \frac{\sigma_p}{\sigma_{GRR}}$$

14) 确定区隔产品的区别分类数 (the number of distinct catehories)，这能由这些测量来区隔。将 *SN* 乘以 1.41，并填入提供的空白处（图 34）。

> 只需考虑该数值的整数部分，因为这定义了区别的分类（省略所有小数部分）（见图 34b）。

如果分类数量小于 2，则这测量系统对于控制该过程毫无价值；它全是干扰，一个零件不能被指为与另一个零件不同。

如果分类数量等于 2，它意味着数据只能分成高、低两组；然而这只等同于计数型数据。

如果分级数量等于 3，则数据能分成高、中、低三组；这是一个稍微好的测量系统。

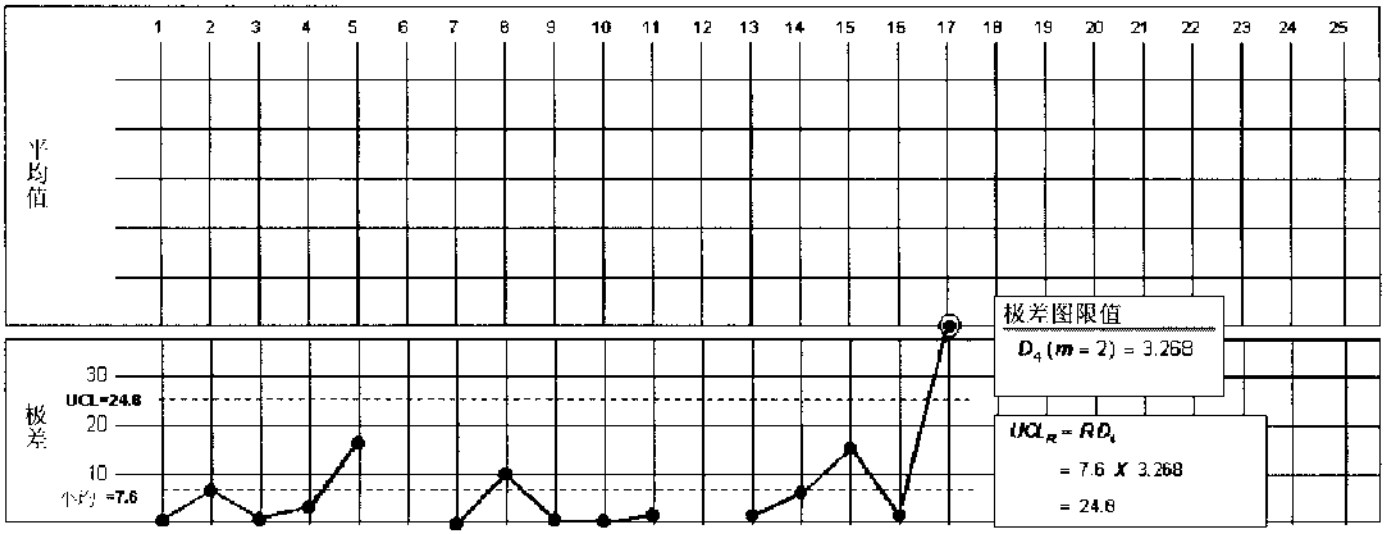
包含 4 个或更多分类的系统会比前面三个例子好得多。

计量控制图 ($\bar{X}$ & R)
测量评价

零件编号		图表编号
零件名称（产品） 接头	作业（过程）	规范限制
操作者	机器	量具 #130	测量单位 .001 mm	0=

日期		评价人(A)						(B)						(C)												
时间	零件	1	2	3	4	5		1	2	3	4	5		1	2	3	4	5								
样本测量	1	113	113	071	101	113		112	117	082	098	110		107	115	103	110	131								
	2	114	106	073	097	130		112	107	083	099	108		109	122	086	108	090								
	3																									
	4																									
	5																									
总和																										
平均值		113.5	109.5	072	099	121.5		112	112	82.5	98.5	109		108	118.5	94.5	109	110.5								
极差 R		1	7	2	4	17		0	10	1	1	2		2	7	17	2	41								
备注																										

总 $R = 114$; $\bar{\bar{R}} = \frac{114}{15} = 7.6$



极差图限值

$D_4 (m = 2) = 3.268$

$UCL_R = \bar{R} D_4$

$= 7.6 \times 3.268$

$= 24.8$

图 33a：测量评价的控制图（$\bar{X}$ & R）– 1

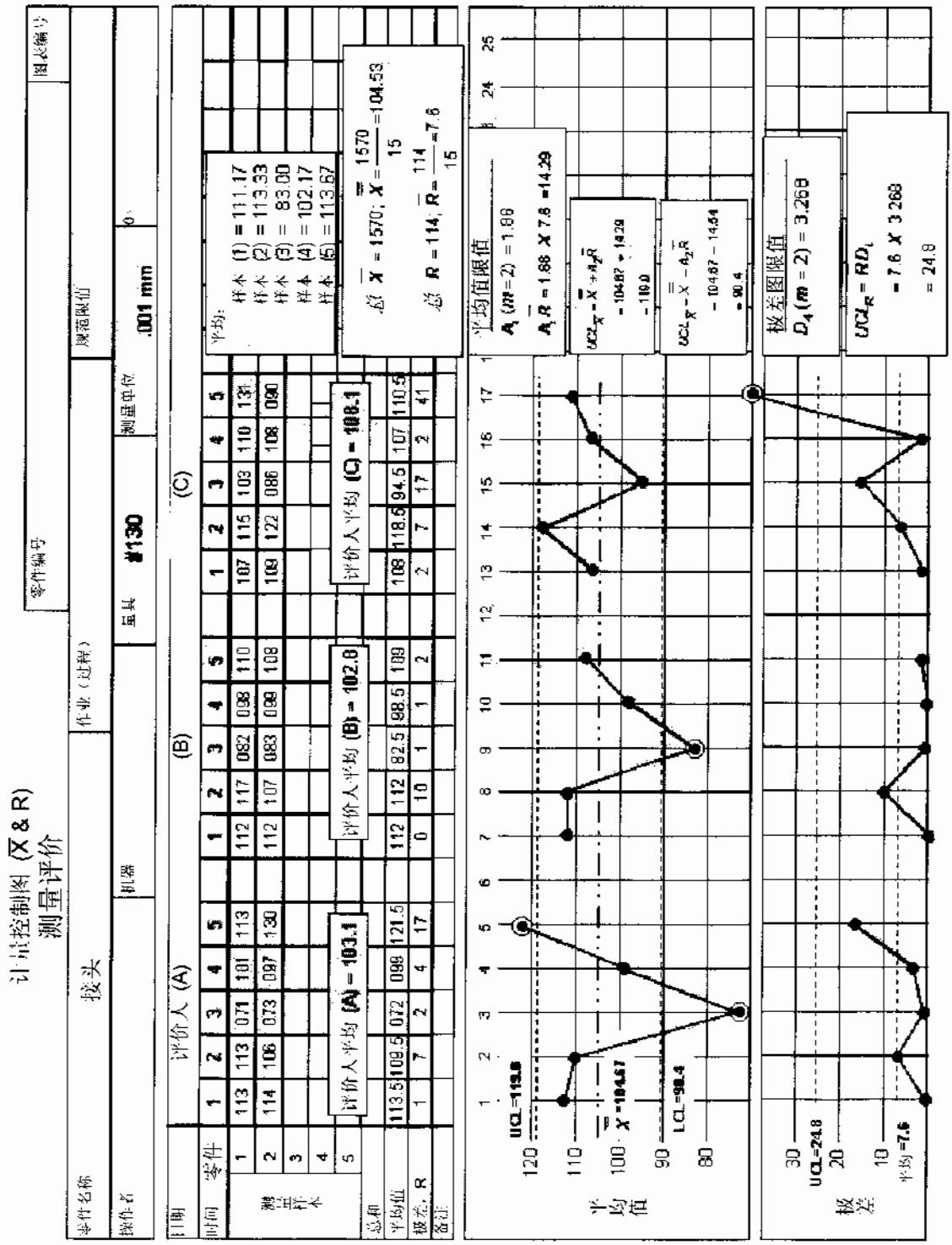


图 33b：测量评价的控制图 ($\bar{X}$ & R) - 2

被评价的测量 接头 完成者 R. W. L. 日期 MM-DD-YY

重复误差：平均子组极差 = $\bar{R}$ = 7.6

重复次数 = 子组大小 = r = 2

估计重复标准差 $\frac{\bar{R}}{d_2} = \hat{\sigma}_e = 6.738$

r	d_2
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704
8	2.847
9	2.970
10	3.078

评价人影响的计算：

n_A	d_2^*
2	1.410
3	1.906
4	2.237
5	2.477
6	2.669
7	2.827
8	2.961
9	3.076
10	3.178

评价人人数 = $n_A = 3$

样本数量 = n = 5

$108.1 - 102.8 = R_A$

评价人平均极差 = $R_A = 5.3$

$\frac{5.3}{1.906} = \hat{\sigma}_A$ $\frac{\bar{R}}{d_2^*} = \hat{\sigma}_A = 2.781$

下面写下评价人平均值：

评价人	平均值
A	103.1
B – 低	102.8
C – 高	108.1
D	-
E	-
F	-

图 34a：评价—测量过程的控制图方法的计算（第 1 部分，共 2 部分）

测量误差标准差的计算：
估计这测量误差标准差为：

$$\hat{\sigma}_m = \sqrt{\hat{\sigma}_e^2 + \hat{\sigma}_A^2}$$
$$= \sqrt{(6.7376)^2 + (2.7806)^2}$$
$$= 7.289$$

此处写下 $\hat{\sigma}_m$ 的适当数值：$\hat{\sigma}_m = 7.289$

对干扰信噪比 (Signal to Noise Ratio) 的计算：

n	d_2^*
2	1.410
3	1.906
4	2.237
5	2.477
6	2.669
7	2.827
8	2.961
9	3.076
10	3.178

在下面写下对每个单件或群组样件的平均值：

113.67 – 83.0 = R_P

这些样本平均值的极差 = R_P = 30.67

估计样本之间的标准差：

$\frac{30.67}{2.477} = 12.38$ $\frac{R_P}{d_2^*} = \hat{\sigma}_p = 12.38$

信噪比：

$\frac{12.38}{7.289} = 1.698$ $\frac{\hat{\sigma}_p}{\hat{\sigma}_m} = 1.698$

样本	平均值
1	111.17
2	113.33
3 – 低	83.00
4	102.17
5 – 高	113.67
6	-
7	-
8	-
9	-
10	-

因此，由这些测量 1.41×1.698，
产品区别分类的数量可以被可靠的区隔

$$1.41\left(\frac{\hat{\sigma}_p}{\hat{\sigma}_m}\right) = 2.395 \text{ 或 } 2$$

这是涵盖产品变差范围的 97%置信度区间数量。（∵97%置信度区间位于单次测量值的中心，其将包含实际的产品值，对于这个值有 97%的机会由该测量来呈现。）

图 34b：评价一测量过程的控制图方法的计算（第 2 部分，共 2 部分）

第五章 第 C 节

量具性能曲线[73]

建立一量具特性曲线 (*Gage Performance Corve, GPC*) 的目的是确定接受或拒收某零件一些参考值的概率。理论上，一个没有误差的测量，其 *GPC* 如图 35 所示；但这是对测量系统的理想而不是正常会发生的情形。

一旦确定了误差量，当采用某系统，就有可能要计算一些其接收某零件参考值的概率。

为完成这工作，首先假设这测量系统误差主要由差的重复性、再现性和偏倚组成，再进一步假设重复性和再现性已呈现常态的分布，另外有一些变异为 σ^2。因此，量具的误差是平均值 X_T（即参考值），加上偏倚和一些变差 σ^2 的常态分布。换句话说：

$$\textit{由量具得到的实际值} = N(X_T + b, \sigma^2)$$

接受某零件参考值的概率表示成如下关系式：

$$P_a = \int_{LL}^{UL} N(X_T + b, \sigma^2)\, dx$$

采用标准常态表：

$$p_a = \phi\left(\frac{UL - (X_T + b)}{\sigma}\right) - \phi\left(\frac{LL - (X_T + b)}{\sigma}\right)$$

式中：

$$\phi\left(\frac{UL - (X_T + b)}{\sigma}\right) = \int_{-\infty}^{UL} N(X_T + b, \sigma^2)dx$$

$$\phi\left(\frac{LL - (X_T + b)}{\sigma}\right) = \int_{LL}^{\infty} N(X_T + b, \sigma^2)dx$$

---

[73] 经允许，节选自 J. McCaslin 和 G. Gruska 所著的*"Analysis of Attribute Gage System"*, ASQC, 1976。

范例

当参考扭矩值为 0.5Nm、0.7 Nm、0.9 Nm 时，要确定接受一零件的概率。

采用前述所进行研究中的数据：

规范上限	=	*UL*	=	1.0 Nm
规范下限	=	*LL*	=	0.6 Nm
偏倚	=	*b*	=	0.05 Nm
σ_{GRR}	=	0.05 Nm		

将上面的数据代入前一页的公式：

$$P_a = \phi\left(\frac{UL-(X_T+b)}{\sigma}\right) - \phi\left(\frac{LL-(X_T+b)}{\sigma}\right)$$

$$P_a = \phi\left(\frac{1.0-(0.5+0.05)}{0.05}\right) - \phi\left(\frac{0.6-(0.5+0.05)}{0.05}\right)$$

$$P_a = \phi(9.0) - \phi(1.0)$$

$$= 1.0 - 0.84$$

$$= 0.16$$

因此，当该零件有一参考值为 0.5 Nm 时，大约有 84%的机会将被拒收。

量具性能曲线范例

对于 X_T= 0.7 Nm

$$P_a = \phi\left(\frac{1.0-(0.7+0.05)}{0.05}\right) - \phi\left(\frac{0.6-(0.7+0.05)}{0.05}\right)$$

$$P_a = \phi(5.0) - \phi(-3.0)$$

$$= 0.999$$

如果该零件的参考值为 0.7Nm，那么它被拒收的机会将少于 0.1%。

对于 X_T= 0.9 Nm

$$P_a = \phi\left(\frac{1.0-(0.9+0.05)}{0.05}\right) - \phi\left(\frac{0.6-(0.9+0.05)}{0.05}\right)$$

$$P_a = \phi(1.0) - \phi(-7.0)$$

$$= 0.84$$

如果该零件的参考值为 0.9 Nm，那它被拒收的机会大约为 16%。

如果对所有 X_T 的数值计算了接受概率并画成图形，然后就得到了量具性能曲线，如图 36 所示。

在常态概率图纸上能更容易绘制这相同的曲线，如图 37。一旦绘制出图形，这 *GPC* 曲线能得到对任何零件尺寸的一个零件接收的概率，

而且，一旦 *GPC* 被建立，它可被用于计算重复性和再现性误差，以及偏倚误差[74]。

这 5.15 *GRR* 范围可以经由找出概率 $P_a = 0.995$ 对应的 X_T 值，和 $P_a = 0.005$ 对应的 X_T 值，而确定两个限值；*GRR* 就是在这两个 X_T 值的差值，如图 37 图形所示。

偏倚的估计值是经由找出对应于概率 $P_a = 0.5$ 的上限值或下限值 X_T，并计算：

$$B = X_T - LL \quad 或 \quad BB = X_T - UL$$

这取决于选择了哪个限值的 X_T[75]。

[74] 参见第 3 章，第 C 节，“计数型量具研究”

[75] 这假设测量系统在其工作量程范围内是线性的。

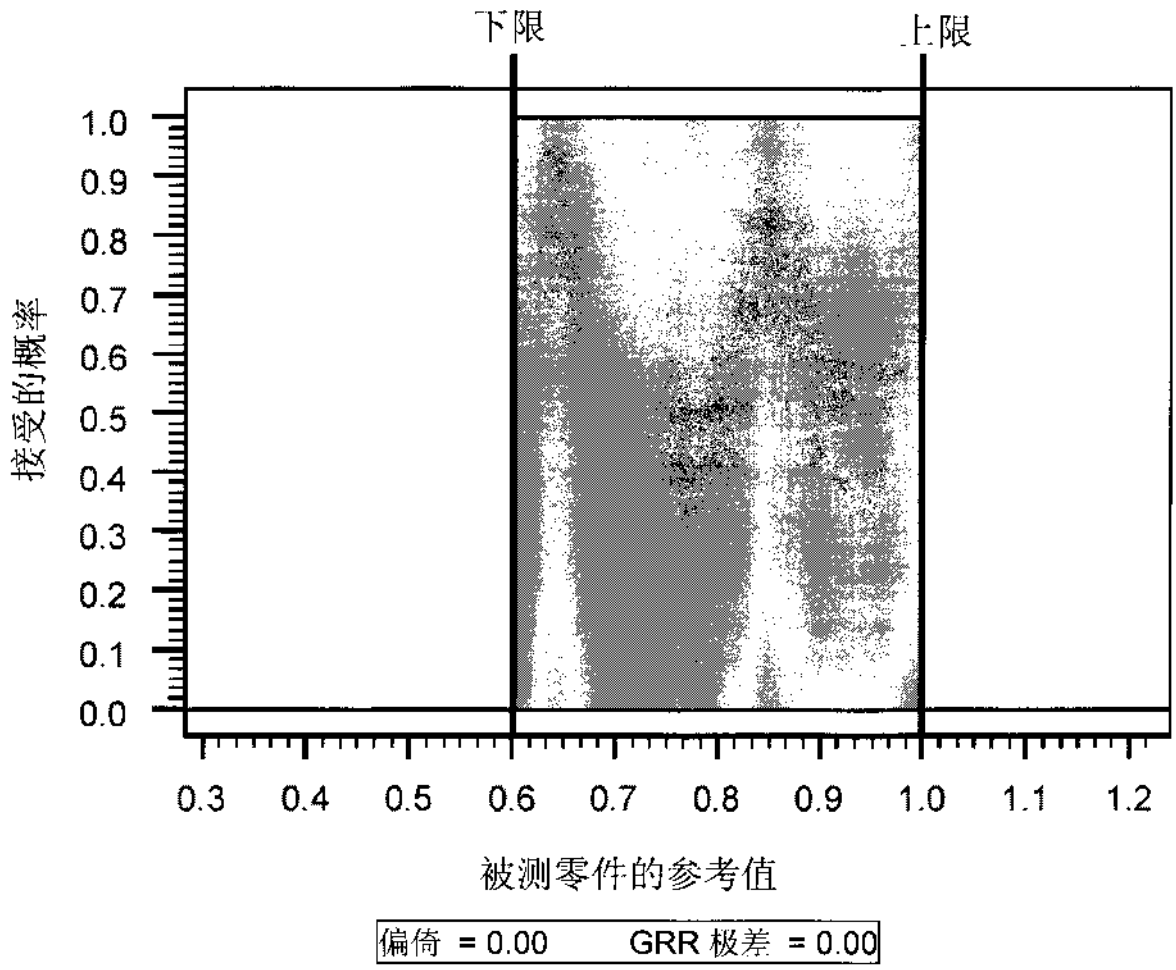


图 35：无误差的量具性能曲线

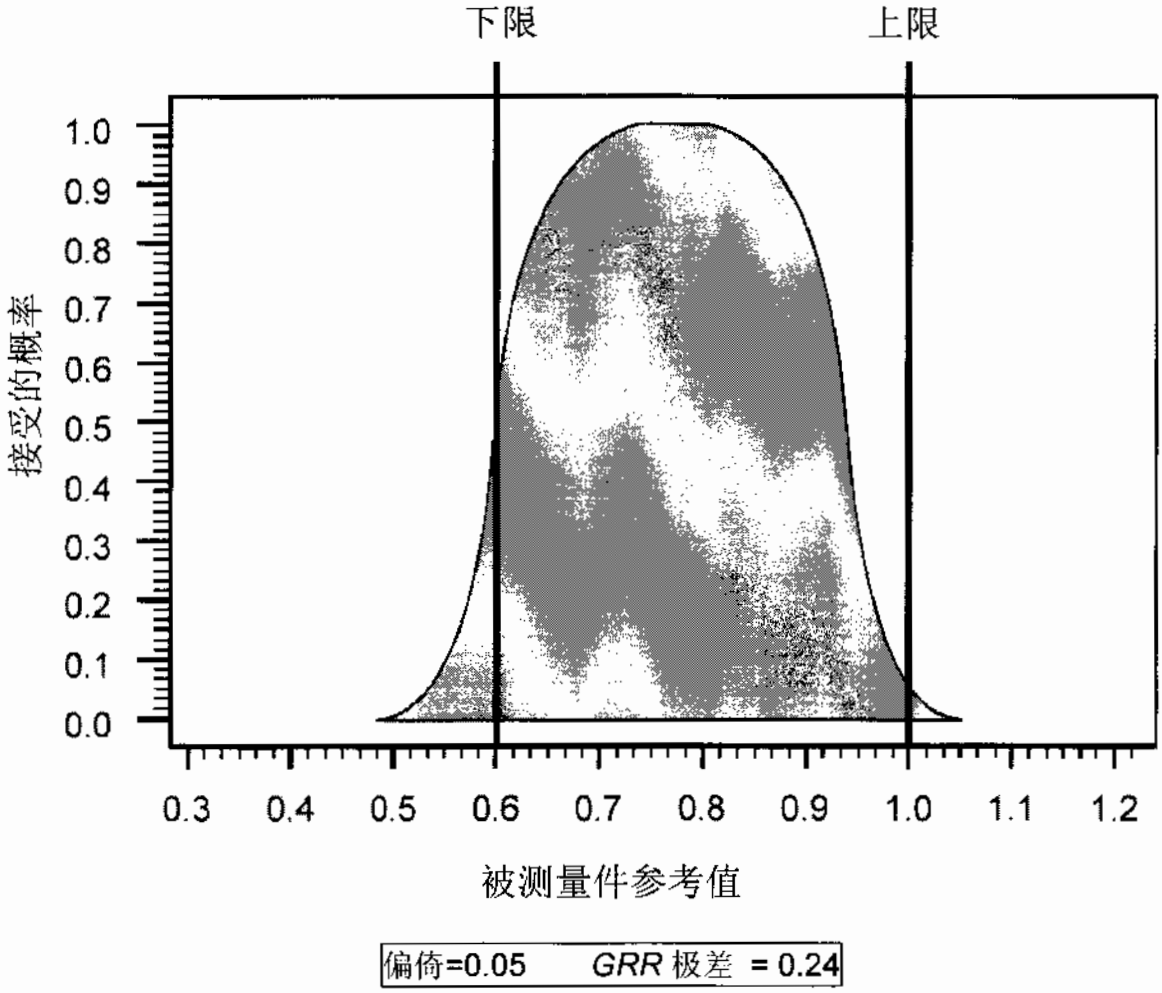


图 36：量具性能曲线 — 范例

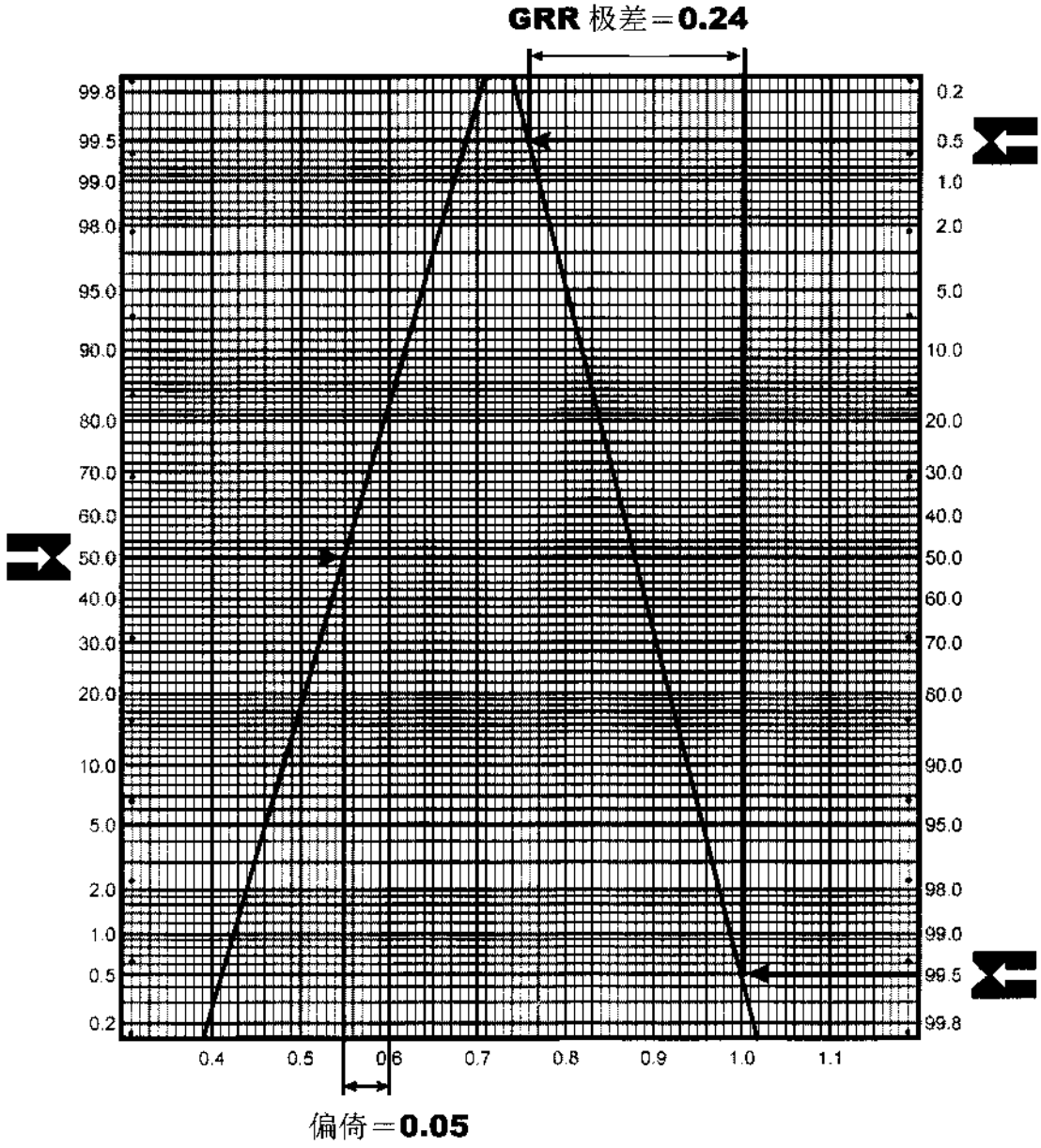


图 37：在常态概率图纸上绘制的量具性能曲线

第五章 — 第 D 节

通过多次读值减少变差

如果目前的测量系统的变差是不可接受的（大于 30%），有一种方法可被用来将变差减少至一个可接受的水准，直到已对测量系统进行了适当的改进。通过对被评价零件特性采取多次统计上独立（非相关）的测量读值，确定这些测量的平均值，并以这结果的数值代替各别的测量，这样可能减少了不可接受的变差。这种方法当然会费时，但是是一种替代的方法，直至已对量具系统进行改进（如：重新设计或购买一个新量具）。这种替代方法的程序如下：

1) 确定为了满足变差的可接受水准，必要的多次读值的数量；

2) 遵循在这本手册前面所讨论的量具研究程序

范例

在 XYZ 公司的范例中，公差 R&R 变差的百分比是 25.5%，有一个 6σ 宽度为 0.24。顾客要缩减该数值至最少 15%，其将有一个等同的 6σ 宽度为 0.14[76]。

为了满足希望的 15%准则，确定所需要的多次读值数量，首先必须明白，个别和平均测量值的分布具有相同平均值。其次，平均值分布的变异等于个别分布的变数除以样本大小。理解了以下所示的这种关系，就能确定所需的多次读值数量。

$$(6\sigma_{\bar{Y}})^2 = \frac{(6\sigma)^2}{n}$$

同时，该等式也可表示为：

$$(6s_{\bar{Y}}) = \frac{(6s)}{\sqrt{n}}$$

以及

[76] 参见第 vi 页的备注。

$$0.14 = \frac{0.24}{\sqrt{n}}$$

所以

$$\sqrt{n} = 1.714$$

以及

$$n = 3 \quad \text{（四舍五入至最接近的整数）}$$

因此，对零件特性的 3 次读值将减少整个测量系统变数至大约 0.14，%*GRR* 减至 15%。

这种方法应该被视为一个临时步骤，直至已对测量系统进行其它改进。该方法应该只有在顾客同意的情况下被使用。

第五章 — 第 E 节

聚集标准差法计算 GRR[77]

一测量系统的分析通常假设所有的评价人以随机的方式获得了所有零件 / 样本的可重复的数据。但这种假设不一定总是可能的。如果某测量系统分析了包括多个场所，逻辑上这很难做到随机的抽样。同样对一些试验，尤其是化学及金相分析（实验室内部和实验室之间的研究），可能需要采用不同样本的混和取样，这些样本不是一同质过程的零件，也可能不是在同一时间获得的。

这些情况可以使用一群体的 DOE 来处理这一问题。一种替代的方法是聚集的标准差法 GRR 研究，这是遵循在 ASTM E691 所讨论过的方法。

该方法将每个零件看成一个单独的材料，然后如 E691 计算重复性和再现性的标准差。这样，将产生多个单独的重复性和再现性值。由于被研究的零件的标识是不可或缺的，假设这些单独的估计值也被有效地标识。当然，它们从来不会相同，但它们的平均值将对重复性及类似再现性的真实水准，提出一个很好的估计值。

如果用这方法来评价一组实验室，有一个顾虑是："再现性"的本质到底是什么？如果大多数时间均大于 0（也就是大多数的材料均大于 0），这意味着应该解释为，有一个存在于操作者之间差异；例：对于内部实验室计划来说，实验室之间存在真正的差异。

连续的应用

尽管这 E691 方法通常是用于完整样本的情况，但也可把它用在连续样本的方法；对当所有样本不能在同一时间获得的情况下是很有用的。另外，它还可用于校准过程的一部分，以保持测量系统变差的信息。

以下的研究假设这种研究可以用于连续的情况。

进行研究

注意应该按照第二章，第 C 节所述的"测量系统研究的准备"来进行。

[77] 本章节的一部分，包括"一致性统计"全文，由 "American Society for Testing and Materials, ASTM *International*)" 的 Neil Ullman 所投稿。

接续自第二章第 C 节，74 页的第 6 步骤：

7) 选 $m \geq 2$ 个评价人，每人评价每个零件 $r \geq 3$ 次读值；将数据记录在数据收集表（参见范例表格）的适当行中。*（注意：在进行多次读值时，不要让评价人看到他们以前测量的读值）*。

8) 计算每个评价人每个零件的平均值（$\bar{X}$）及标准差（s）。

9) 将标准差数值描绘在标准差图表上，并计算标准差的平均值（$\bar{s}$）（包括对所有评价人的所有子组的标准差）；将这些标准差的平均值画在图上。用对应测量次数 r 的 D_4 常数计算标准差图表的上控制限；画出该控制限并判断是否所有的值均受控。（见图 38）。

10) 将所有评价人每个子组的平均值（$\bar{X}$）画在平均值图表上（见图 38），平均值代表过程变差和测量变差两者。

11) 计算总平均值（$\bar{\bar{X}}$）（包括所有评价人的所有子组平均值），在平均值图表上画出该总平均值。

12) 用对应 r 的 A_2 常数，以及从标准差图上的标准差平均值（$\bar{s}$）来计算平均值图的控制限；将这些控制限画在平均值图上。

13) 应用控制图和再平均值和极差图（见第三章）中所讨论的其它图形方法来分析数据。

14) 通过聚集了评价人的结果，评价每个零件的测量系统参数。

$$s_{\bar{X}_g} = \sqrt{\frac{\sum_{i=1}^{m} s_{\bar{X}_i}^2}{m}}$$

$$\text{重复性}_g = s_{E_g} = \sqrt{\frac{\sum_{i=1}^{m} s_{x_i}^2}{m}}$$

$$\text{再现性}_g = s_{A_g} = \sqrt{s_{\bar{X}_c}^2 - \frac{s_{E_g}^2}{m}}$$

$$GRR_g = s_{GRR_g} = \sqrt{s_{E_g}^2 + s_{A_g}^2}$$

E691 按照在 MSA 中的惯例，将再现性定义为评价人的变差，以及 MSA 的 GRR 被称为再现性，在这种情况下：

$$S_{\text{评价人}} = \sqrt{s_{\bar{x}}^2 - \frac{s_r^2}{3}}$$

$$S_R = \sqrt{s_r^2 + s_{\text{评价人}}^2}$$

式中，$s_r = s_E =$ 重复性，$s_R = GRR =$ ASTM 再现性

15) 通过聚集了零件结果，评价整体测量系统的参数。

$$\text{重复性} = S_E = \sqrt{\frac{\sum_{i=1}^{g} s_{E_i}^2}{g}}$$

$$\text{再现性} = s_A = \sqrt{\frac{\sum_{i=1}^{g} s_{A_i}^2}{g}}$$



$$GRR = s_{GRR} = \sqrt{\frac{\sum_{i=1}^{g} s_{GRR_i}^2}{g}}$$

当计算总变差的百分比时，过程的历史标准差应该被使用。

如果这零件涵盖一广泛多变个过程，例如：不同的金相或化学的样本，评价总变差的百分比时，应该以指定样本的过程变差为基础，而不是以所有样本的总变差为基础。

当评价人是隶属于不同的场所（例如实验室）时，在解释测量系统的参数时，应该特别注意。

重复性将包括设备间的变差和设备内部的变差。可通过计算和比较每个场所的重复性来进行评价。

再现性将包括不同场所之间和评价人之间的变差。在本研究中，不能将这两个要素分开。

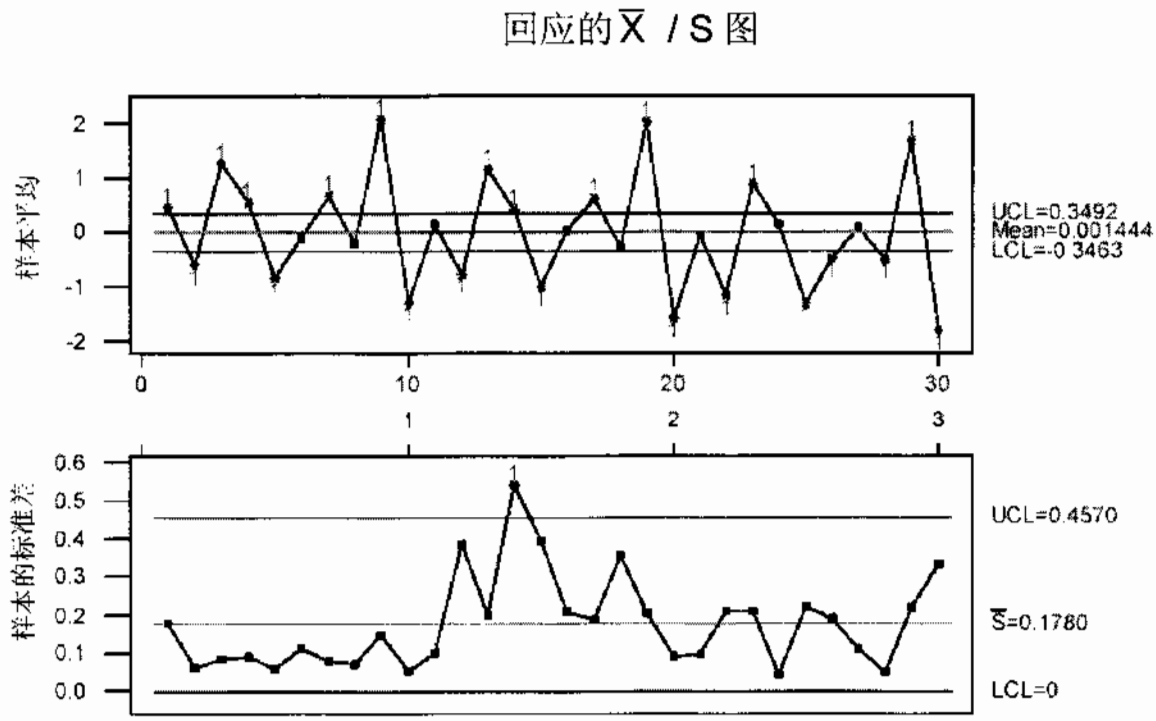


图 38a：聚集的标准差研究的图示分析

表 15：聚集的标准差分析数据组

			零件											横跨的零件
	测量		1	2	3	4	5	6	7	8	9	10		
A	1		0.29	-0.56	1.34	0.47	-0.8	0.02	0.59	-0.31	2.26	-1.36		
	2		0.41	-0.68	1.17	0.5	-0.92	-0.11	0.75	-0.2	1.99	-1.25		
	3		0.64	-0.58	1.27	0.64	-0.84	-0.21	0.66	-0.17	2.01	-1.31		
		$\bar{x}$	0.4467	-0.6067	1.2600	0.5367	-0.8533	-0.1000	0.6667	-0.2267	2.0867	-1.3067	$\bar{\bar{x}}$	0.19033333
		标准差	0.1779	0.0643	0.0854	0.0907	0.0611	0.1153	0.0802	0.0737	0.1504	0.0551	评价人标准差	0.10289153
B	1		0.08	-0.47	1.19	0.01	-0.56	-0.2	0.47	-0.63	1.8	-1.68		
	2		0.25	-1.22	0.94	1.03	-1.2	0.22	0.55	0.08	2.12	-1.62		
	3		0.07	-0.68	1.34	0.2	-1.28	0.06	0.83	-0.34	2.19	-1.5		
		$\bar{x}$	0.1333	-0.7900	1.1567	0.4133	-1.0133	0.0267	0.6167	-0.2967	2.0367	-1.6000	$\bar{\bar{x}}$	0.06833333
		标准差	0.1012	0.3869	0.2021	0.5424	0.3946	0.2120	0.1890	0.3570	0.2079	0.0917	评价人标准差	0.3017394
C	1		0.04	-1.38	0.88	0.14	-1.46	-0.29	0.02	-0.46	1.77	-1.49		
	2		-0.11	-1.13	1.09	0.2	-1.07	-0.67	0.01	-0.56	1.45	-1.77		
	3		-0.15	-0.96	0.67	0.11	-1.45	-0.49	0.21	-0.49	1.87	-2.16		
			-0.0733	-1.1567	0.8800	0.1500	-1.3267	-0.4833	0.0800	-0.5033	1.6967	-1.8067	$\bar{\bar{x}}$	-0.25433333
		标准差	0.1002	0.2113	0.2100	0.0458	0.2223	0.1901	0.1127	0.0513	0.2194	0.3365	评价人标准差	0.19056058
各个件的标准			0.26182	0.28005	0.19648	0.19751	0.24077	0.26555	0.32524	0.14385	0.21221	0.25125		聚集的标准差
	$\bar{x}$标准差													
	重复性		0.13153	0.25721	0.17534	0.31863	0.26388	0.17736	0.13524	0.21252	0.19494	0.20385	测定系统	
		聚集		0.204274	0.195107	0.23223	0.238896	0.229787	0.218795	0.218021	0.215578			0.21443466
	再现性		0.25056	0.23743	0.16839	0.07191	0.18644	0.24501	0.31573	0.07508	0.17991	0.22198		
		聚集		0.244086	0.221744	0.195372	0.193619	0.203089	0.222696	0.209998	0.206871			0.20843864
	GRR		0.28299	0.35004	0.24310	0.32664	0.32310	0.30247	0.34347	0.22540	0.26527	0.30138		
		聚集		0.318285	0.295359	0.303481	0.307505	0.306671	0.312194	0.302708	0.29878			0.29904106

一致性统计

ASTM 以及 ISO 方法[78]建议计算两个"一致性"统计量，计算 *h* 和 *k*。*h* 值按下式计算：

$$h = \frac{\overline{X}_{\text{评价人}} - \overline{\overline{X}}_{\text{零件}}}{s_{\overline{X}}}$$

对于评价人 A 和零件 1 来说，平均值（上式中的 $\overline{X}_{\text{评价人}}$）为 0.447，零件平均值（上式中的 $\overline{\overline{X}}_{\text{零件}}$）为 0.169。评价人之间的标准差（上式中的 $s_{\overline{X}}$）为 0.262，则：

$$h = \frac{0.447 - 0.169}{0.262} = \frac{0.278}{0.262} = 1.06$$

k 值是每个评价人测量每个零件的标准差与重复性标准差的比值。在本例中（评价人 A 和零件 1），为：

$$k = \frac{(\text{评价人}A，\text{零件}1)\ \text{的标准差}}{\text{重复性}} = \frac{0.178}{0.132} = 1.35$$

这些计算的原因之一是允许对非常不同的材料进行比较。

尽管在本例中没有收集差别较大的材料，这些材料有不同的水准，而且它们的标准差可能相差非常大。但仍可应用 E691 中的 *h* 和 *k* 的计算，以比较重复性标准差和评价人的回应数值。下表列出了各评价人的 *h* 和 *k* 值。

h

	零件											
评价人	1	2	3	4	5	6	7	8	9	10	平均 *h*	"z"
A	1.06	0.87	0.82	0.86	0.88	0.32	0.65	0.80	0.69	1.05	0.80	2.53
B	-0.14	0.22	0.29	0.24	0.21	0.80	0.50	0.32	0.46	-0.11	0.28	0.88
C	-0.93	-1.09	-1.11	-1.10	-1.09	-1.12	-1.15	-1.12	-1.15	-0.94	-1.08	-3.41

k

											中间的 *k*	"z"
A	1.35	0.25	0.49	0.28	0.23	0.65	0.59	0.35	0.77	0.27	0.42	-3.20
B	0.77	1.50	1.15	1.70	1.50	1.20	1.40	1.68	1.07	0.45	1.30	3.14
C	0.76	0.82	1.20	0.14	0.84	1.07	0.83	0.24	1.13	1.65	0.84	-0.17

[78] 参见 ISO 5725

上表的最后两列是平均值和"z-值"，用来观察评价人是否显着不同的。在他们对零件大小的读值中，*h* 值表明评价人A明显的大，评价人C明显的小。正是这种明显差异产生GRR标准差。

重复性标准差也可以通过检查 *k* 值的方法来评价。为此，计算 *k* 的中间值，然后得出一个估计的"z 分"。本研究中，期望的中间值约为 0.861，其标准差约为 0.439。评价人 A 的 *k* 值的中间值约为期望的水准之下的-3.2 标准差，而评价人 B 明显的大。因此我们在仅有的三个操作者的表现中就能看到很大的差别。

h 图（图 38b）和 *k* 图（图 38c）也有助于说明这些差别。评价人 C 比其它人有更小的结果；同样的，*k* 值表明评价人 A 对于有关重复性的变差有多小。这些是由评价人进行关于测量方法的表现时，必然要检查的事项。

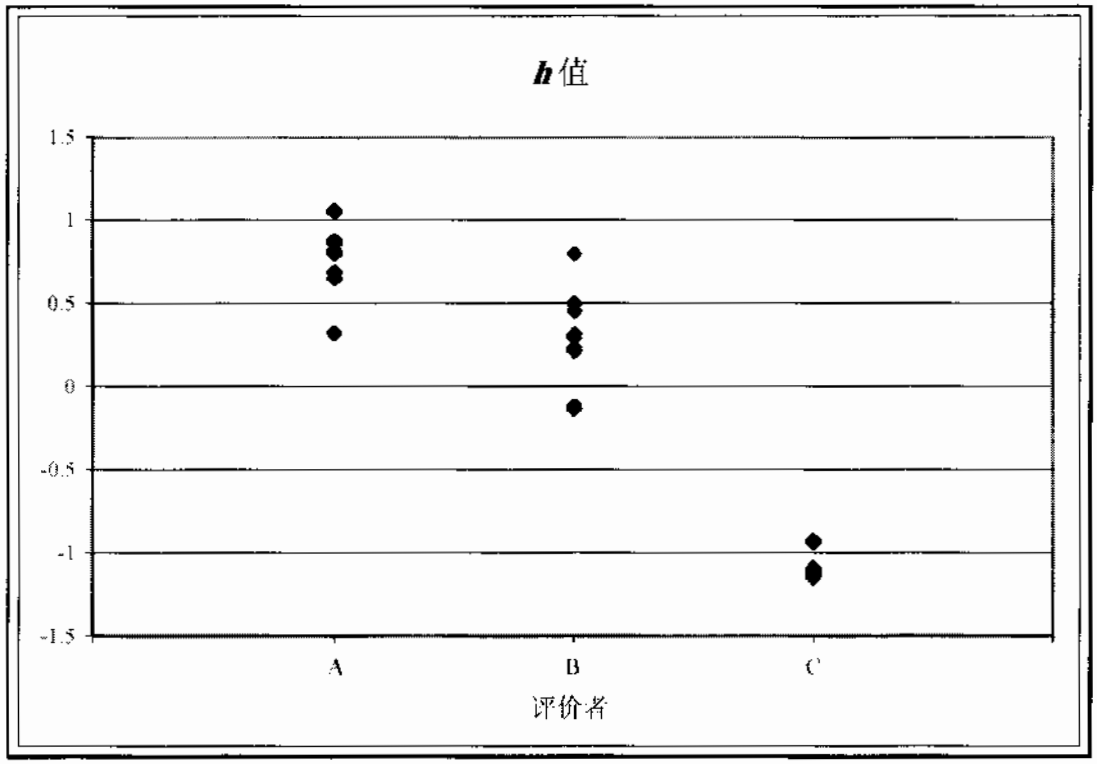


图 38b：*h* 值的点图

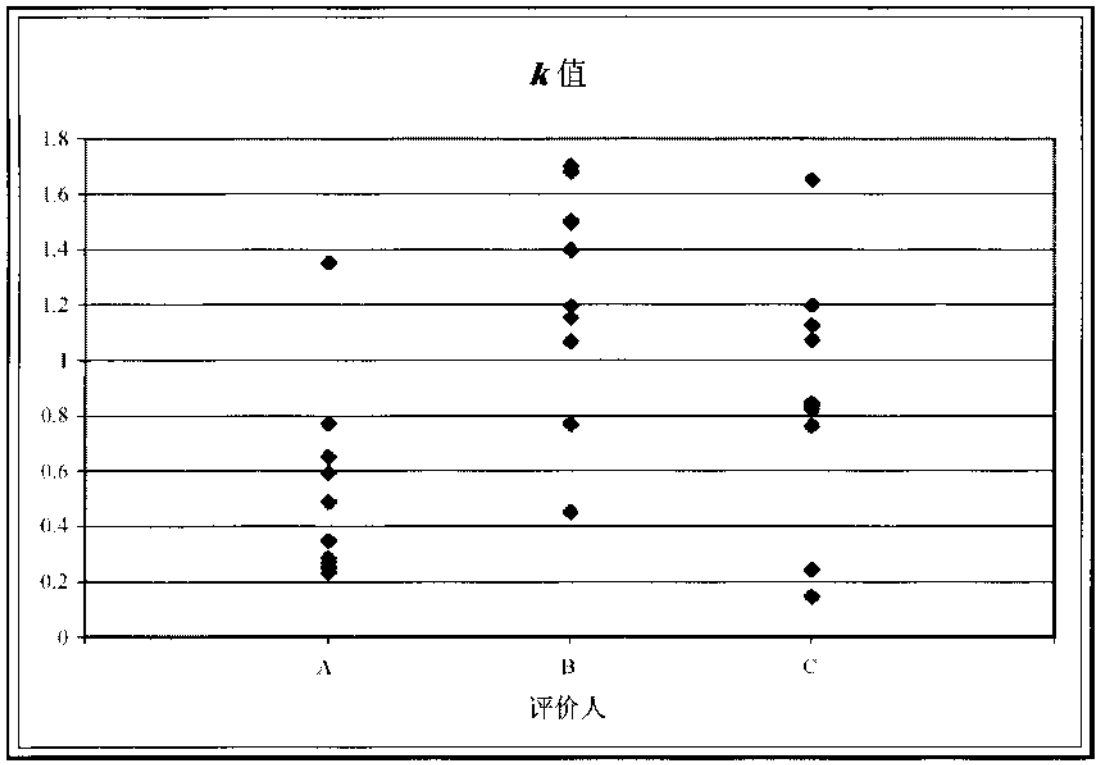


图 38c：*k* 值的点图

附录

附录 A

方差分析的概念

可依据表 18 中的公式进行 GRR 数值的分析：这就是被称为方差分析 (*ANOVA*) 的表。*ANOVA* 表包括 6 列：

- *来源* 列为变差的原因。
- *DF* 列为与变差来源有关的 *自由度* (*degree of freedom*)。
- *SS* 或 *平方的总和* (*sum of square*) 列为来源平均值附近的偏差。
- *MS* 或 *平均平方* (*mean square*) 列是平方的总和除以自由度。
- *EMS* 或 *预期的平均平方* (*expected mean square*) 列确定了每个 *MS* 的变数元素的线性组合。*ANOVA* 表将总变差来源分解成 4 个要素：零件、评价人、零件与评价人的相互作用，以及量 / 设备具重复性造成的重复误差。
- *F-比率* (*F-ratio*) 列只用于对 MSA ANOVA 的相互作用计算；它由相互作用的平均平方除以平均平方误差来确定。

对每个误差来源的变数元素估计值在表 16 中提出[79]。

	变数估计值
量具 (*EV*)	$\tau^2 = MS_e^{80}$
相互作用 (*INT*)	$\gamma^2 = \frac{MS_{AP} - MS_e}{\gamma}$
评价人 (*AV*)	$\varpi^2 = \frac{MS_A - MS_{AP}}{nr}$
零件 (*PV*)	$\sigma^2 = \frac{MS_p - MS_{AP}}{kr}$

表 16：变数元素的估计值

由于每个平均平方值是受到抽样变差影响的一个样本数量的值，并且这计算涉及到平均平方的差异，所以可能出现负的变差要素估计值。因为"基准"变数要素等于或接近于 0，或者有一个小的样本大小，所以这是一点小问题。为了分析，负的变数要素被设定为 0。

[79] 在这表中，假设变差的所有要素均为随机的影响。

[80] 在 ANOVA 应用于测量系统分析时，ANOVA 的误差术语与 MSA 变差 MS_E 的术语相同。

标准偏差比变差易于解释，因为它与原始观测使用相同的测量单位。实际上，基本分布宽度的测量被设定为 5.15 倍的标准偏差[81]。表 17 列出了被称为设备变差 (*EV*) 的重复性度量和被称为评价人变差 (*AV*) 的再现性度量，两者例用了 5.15σ 分布宽度的公式。如果零件和评价人的相互作用是显着的，则存有一非附加的模型，于是可以提出它的变数要素的估计值。表 17 中的 *GRR* 是总测量系统变差。

设备变差 = 重复性	
$AV = 5.15\sqrt{\frac{MS_A - MS_{AP}}{nr}}$	评价人变差 = 再现性
$I_{AP} = 5.15\sqrt{\frac{MS_{AP} - MS_e}{r}}$	零件对评价人的相互作用
$GRR = \sqrt{(EV)^2 + (AV)^2 + (I_{AP})^2}$	量具 R&R
$PV = 5.15\sqrt{\frac{MS_P - MS_{AP}}{kr}}$	零件变差

表 17：5.15σ 分布宽度

在附加的模型中，相互作用是不显着的，且各来源的变数要素被确定如下：第一，是将量具误差平方和（表 18 中的 SS_e）加到零件对评价人的相互作用的平方总和上（表 18 中的 $SS_{评价人}$），这等于具有（*nkr – n – k + 1*）[82] 自由度的聚集平方总和 （$SS_{聚集}$）。然后 $SS_{聚集}$ 将除以 (*nkr – n – k + 1*) 以计算出 $MS_{聚集}$。 然后 5.15σ 分布宽度限值将是：

$$EV = 5.15\sqrt{MS_{聚集}}$$

$$AV = 5.15\sqrt{\frac{MS_A - MS_{聚集}}{nr}}$$

$$GRR = \sqrt{(EV)^2 + (AV)^2}$$

$$PV = 5.15\sqrt{\frac{MS_P - MS_{聚集}}{kr}}$$

[81] 这是 99%的范围。参见第 vi 页的备注。
[82] 公式中 *n* 为零件数量，*k* 为评价人人数，*r* 为测量次数。

为了确定这相互作用是否显著，计算零件对评价人的相互作用的 *F* 统计（参见表 18）。将这 *F* 统计与得自 *ANOVA*（表 18）中以自由度为分母、*F* 分布为分子的一个上端百分比点相比较。

为了减小错误地做出没有相互作用的结论的风险，选择了一个较高的重要水准。一旦确定了 *GRR*，然后就可以计算与过程性能有关的%*GRR* 了。

$$SS_{零件} = \sum_{i=1}^{n}\left(\frac{x_{i..}^2}{kr} - \frac{x^2}{nkr}\right)$$

$$SS_{评价人} = \sum_{j=1}^{k}\left(\frac{x_{.j.}^2}{nr} - \frac{x^2}{nkr}\right)$$

$$TSS = \sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{m=1}^{r}\left(x_{ijm}^2 - \frac{x^2}{nkr}\right)$$

$$SS_{零件对评价人} = \sum_{i=1}^{n}\sum_{j=1}^{k}\frac{x_{ij}^2}{r} - \sum_{i=1}^{n}\frac{x_{i..}^2}{kr} - \sum_{j=1}^{k}\frac{x_{.j.}^2}{nr} + \frac{x^2}{nkr}$$

$$SS_{设备} = TSS - [SS_A + SS_p + SS_{AP}]$$

变差源	DF	SS	MS	F	EMS
评价人	*k-1*	SS_A	$MS_A = \frac{SS_A}{(k-1)}$		$\tau^2 + r\gamma^2 + nr\omega^2$
零件	*n-1*	SS_P	$MS_p = \frac{SS_p}{(n-1)}$		$\tau^2 + r\gamma^2 + kr\sigma^2$
零件的评价人	*(n-1) (k-1)*	SS_{AP}	$MS_{AP} = \frac{SS_{AP}}{(n-1)(k-1)}$	$\frac{MS_{AP}}{MS_e}$	$\tau^2 + r\gamma^2$
设备	*nk (r-1)*	SS_E	$MS_e = \frac{SS_e}{nk(r-1)}$		τ^2
总变差	*nkr-1*	*TSS*			评价人 ~ $N(0, \omega^2)$ 零件 ~ $N(0, \sigma^2)$ x 零件的评价人 ~ $N(0, r^2)$ 设备 ~ $N(0, \tau^2)$

表 18：变差数分析 (*ANOVA*)

表 19a 和 19b 是用 *ANOVA* 法计算在图 24 中的举例数据。

变差来源	DF	SS	MS	F	EMS
评价人	2	3.1673	1.58363	79.41*	$\tau^2 + 3r^2 + 30\omega^2$
零件	9	88.3619	9.81799	492.29*	$\tau^2 + 3\gamma^2 + 9\sigma^2$
零件的评价人	18	0.3590	0.01994	0.434	$\tau^2 + 3r^2$
设备	60	2.7589	0.04598		τ^2
总变差	89	94.6471			

*显着的在 α=0.05 水准

表 19a：*ANOVA* 结果的列表

变差的估计	标准差 (σ)	5.15 (σ)	总变差%	贡献度%
τ^2 = 0.039973 （设备）	0.199933	*EV* = 1.029656	18.4	3.4
ω^2 = 0.051455 （评价人）	0.226838	*AV* = 1.168213	20.9	4.4
r^2 = 0 (相互作用)		*INT* = 0	0	0
GRR = 0.09143 ($\tau^2 + r^2 + \omega^2$)	0.302373	*GRR* = 1.557213	27.9	7.8
σ^2 = 1.086447 （零件）	1.042327	*PV* = 5.367987	96.0	92.2
总变差	1.085	*TV* = 5.589293	100.0	

表 19b：*ANOVA* 结果的列表
（变差的估计是以没有相互作用的模式为基础）

$$ndc = 1.41(PV / GRR) = 1.41(5.37/1.56) = 4.85 \cong 4$$

$$\text{总变差 } (TV) = \sqrt{GRR^2 + PV^2}$$

$$\text{总变差\%} = 100\left(\frac{5.15\sigma_{(要素)}}{5.15\sigma_{(全部)}}\right)$$

$$\text{（对总变差）贡献度\%} = 100\left(\frac{5.15\sigma^2_{(要素)}}{5.15\sigma^2_{(全部)}}\right)$$

附录 B

GRR 对能力指数 *Cp* 的影响

公式：

$$\sigma_0^2 = \sigma_A^2 + \sigma_M^2 \tag{1}$$

式中：*O* = 观测到的过程变差
A = 实际的过程变差
M = 测量的系统变差

$$Cp_x = \frac{|U - L|}{6\sigma_x} \tag{2}$$

式中：*U*、*L* 是规范值的上、下限
x 为 (1) 式中所定义的 *O* 或 *A*

$$GRR\% = GRR_P \times 100\% \tag{3}$$

若以过程变差为基础：

$$GRR_p = \frac{k\sigma_M}{6\sigma_O} \tag{4}$$

注：$GRR_P \leq 1$，因为根据公式 (1)：$\sigma_0^2 \geq \sigma_M^2$。

若以公差范围为基础：

$$GRR_p = \frac{k\sigma_M}{|U - L|} \tag{5}$$

在公式 (4) 及公式 (5) 中，*k* 通常取 5.15。但是，如果没有损失，则通常这分析将 *k* 取为 6，以便简化计算。

分析：

$$Cp_O = Cp_A \times \frac{\sigma_A}{\sigma_O}$$

$$= Cp_A \times \frac{\sqrt{\sigma_O^2 - \sigma_M^2}}{\sigma_O} \qquad \text{使用 (1)}$$

GRR 基于过程变差

$$Cp_O = Cp_A \times \frac{\sigma_O\sqrt{1 - GRR^2}}{\sigma_O} \qquad \text{使用 (4)}$$

$$= Cp_A\sqrt{1 - GRR^2} \tag{6}$$

或者

$$Cp_A = \frac{Cp_O}{\sqrt{1-GRR^2}} \tag{6'}$$

GRR 基于公差范围

$$GRR = \frac{1}{Cp_O} \times \frac{\sigma_M}{\sigma_O}$$ 使用 (2) 和 (5)

因此

$$Cp_O = Cp_A \times \sqrt{1-(Cp_O \times GRR)^2} \tag{7}$$

以及

$$Cp_A = \frac{Cp_O}{\sqrt{1-(Cp_O \times GRR)^2}} \tag{7'}$$

图形分析

基于公式 (6)，Cp_O 对于 Cp_A 关系的系列直线如下：

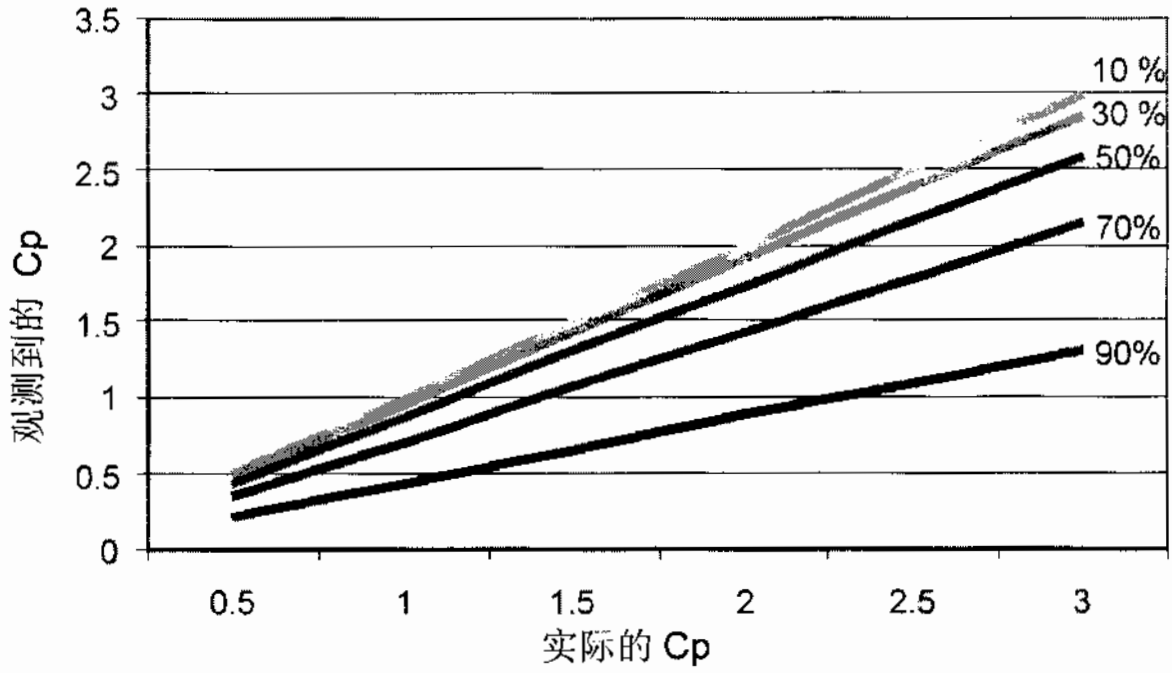


图 39：观测的与实际的 *Cp* 关系（以过程为基础）

实际的	*GRR*							
	10%	20%	30%	40%	50%	60%	70%	90%
	观测的 *Cp* 与基于过程范围的 *Cp*							
1.3	1.29	1.27	1.24	1.19	1.13	1.04	0.93	0.57
	观测的 *Cp* 与基于公差范围的 *Cp*							
1.3	1.29	1.26	1.20	1.11	0.99	0.81	0.54	不会发生这情况

表 20：观测的与实际的 *Cp* 比较

基于公式 (7')，Cp_A 对于 Cp_O 关系的系列直线如下：

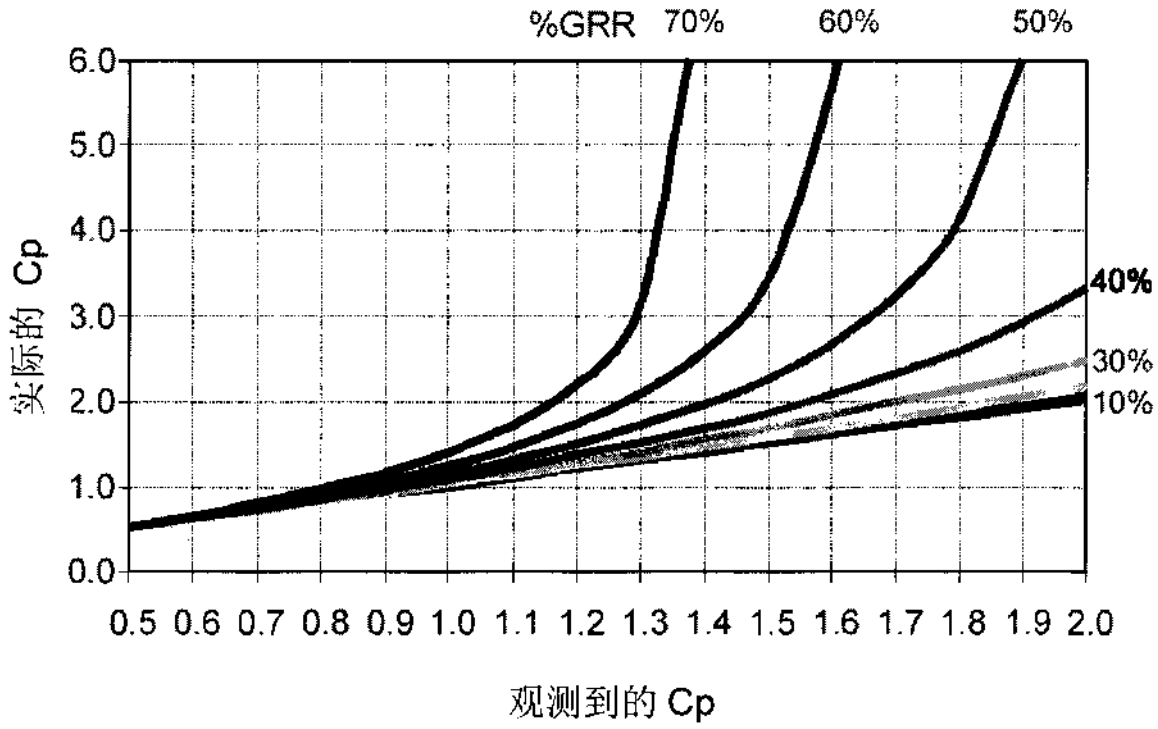


图 39：观测的与实际的 *Cp* 关系（以公差为基础）

与平均极差的分布有关的数值

子组的数量 (g)	子组的大小 (m)																		
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.9 1.41421	2.0 1.91155	2.9 2.23887	3.8 2.48124	4.7 2.67253	5.5 2.82981	6.3 2.96288	7.0 3.07794	7.7 3.17905	8.3 3.26909	9.0 3.35016	9.6 3.42378	10.2 3.49116	10.8 3.55333	11.3 3.61071	11.9 3.66422	12.4 3.71424	12.9 3.76118	13.4 3.80537
2	1.9 1.27931	3.8 1.80538	5.7 2.15069	7.5 2.40484	9.2 2.60438	10.8 2.76779	12.3 2.90562	13.8 3.02446	15.1 3.12869	16.5 3.22134	17.8 3.30463	19.0 3.38017	20.2 3.44922	21.3 3.51287	22.4 3.57156	23.5 3.62625	24.5 3.67734	25.5 3.72524	26.5 3.77032
3	2.8 1.23105	5.7 1.76858	8.4 2.12049	11.1 2.37883	13.6 2.58127	16.0 2.74681	18.3 2.88628	20.5 3.00643	22.6 3.11173	24.6 3.20526	26.5 3.28931	28.4 3.36550	30.1 3.43512	31.9 3.49927	33.5 3.55842	35.1 3.61351	36.7 3.66495	38.2 3.71319	39.7 3.75857
4	3.7 1.20621	7.5 1.74989	11.2 2.10522	14.7 2.36571	18.1 2.56964	21.3 2.73626	24.4 2.87656	27.3 2.99737	30.1 3.10321	32.7 3.19720	35.3 3.28163	37.7 3.35815	40.1 3.42805	42.4 3.49246	44.6 3.55183	46.7 3.60712	48.8 3.65875	50.8 3.70715	52.8 3.75268
5	4.6 1.19105	9.3 1.73857	13.9 2.09601	18.4 2.35783	22.6 2.56263	26.6 2.72991	30.4 2.87071	34.0 2.99192	37.5 3.09808	40.8 3.19235	44.0 3.27701	47.1 3.35372	50.1 3.42381	52.9 3.48836	55.7 3.54787	58.4 3.60328	61.0 3.65502	63.5 3.70352	65.9 3.74914
6	5.5 1.18083	11.1 1.73099	16.7 2.08985	22.0 2.35253	27.0 2.55795	31.8 2.72567	36.4 2.86680	40.8 2.98829	45.0 3.09467	49.0 3.18911	52.8 3.27392	56.5 3.35077	60.1 3.42097	63.5 3.48563	66.8 3.54522	70.0 3.60072	73.1 3.65253	76.1 3.70109	79.1 3.74678
7	6.4 1.17348	12.9 1.72555	19.4 2.08543	25.6 2.34875	31.5 2.55460	37.1 2.72263	42.5 2.86401	47.6 2.98368	52.4 3.09222	57.1 3.18679	61.6 3.27172	65.9 3.34866	70.0 3.41894	74.0 3.48368	77.9 3.54333	81.6 3.59888	85.3 3.65075	88.8 3.69936	92.2 3.74509
8	7.2 1.16794	14.8 1.72147	22.1 2.08212	29.2 2.34591	36.0 2.55208	42.4 2.72036	48.5 2.86192	54.3 2.98373	59.9 3.09039	65.2 3.18506	70.3 3.27006	75.2 3.34708	80.0 3.41742	84.6 3.48221	89.0 3.54192	93.3 3.59751	97.4 3.64941	101.4 3.69806	105.3 3.74382
9	8.1 1.16361	16.6 1.71828	24.9 2.07955	32.9 2.34370	40.4 2.55013	47.7 2.71858	54.5 2.86028	61.1 2.98221	67.3 3.08896	73.3 3.18370	79.1 3.26878	84.6 3.34585	90.0 3.41624	95.1 3.48107	100.1 3.54081	104.9 3.59644	109.5 3.64838	114.1 3.69705	118.5 3.74284
10	9.0 1.16014	18.4 1.71573	27.6 2.07746	36.5 2.34192	44.9 2.54856	52.9 2.71717	60.6 2.85898	67.8 2.98100	74.8 3.08781	81.5 3.18262	87.9 3.26775	94.0 3.34486	99.9 3.41529	105.6 3.48016	111.2 3.53993	116.5 3.59559	121.7 3.64755	126.7 3.69625	131.6 3.74205
11	9.9 1.15729	20.2 1.71363	30.4 2.07537	40.1 2.34048	49.4 2.54728	58.2 2.71600	66.6 2.85791	74.6 2.98000	82.2 3.08688	89.6 3.18174	96.6 3.26699	103.4 3.34406	109.9 3.41452	116.2 3.47941	122.3 3.53921	128.1 3.59489	133.8 3.64687	139.4 3.69558	144.7 3.74141
12	10.7 1.15490	22.0 1.71189	33.1 2.07436	43.7 2.33927	53.8 2.54621	63.5 2.71504	72.6 2.85702	81.3 2.97947	89.7 3.08610	97.7 3.18103	105.4 3.26620	112.7 3.34339	119.9 3.41387	126.7 3.47879	133.3 3.53861	139.8 3.59430	146.0 3.64630	152.0 3.69503	157.9 3.74087
13	11.6 1.15289	23.8 1.71041	35.8 2.07316	47.3 2.33824	58.3 2.54530	68.7 2.71422	78.6 2.85627	88.1 2.97847	97.4 3.08544	105.8 3.18037	114.1 3.26561	122.1 3.34282	129.8 3.41333	137.3 3.47826	144.4 3.53810	151.4 3.59381	158.1 3.64582	164.7 3.69457	171.0 3.74041
14	12.5 1.15115	25.7 1.70914	38.6 2.07213	51.0 2.33737	62.8 2.54452	74.0 2.71351	84.7 2.85562	94.9 2.97787	104.6 3.08487	113.9 3.17984	122.9 3.26510	131.5 3.34233	139.8 3.41286	147.8 3.47781	155.5 3.53766	163.0 3.59339	170.3 3.64541	177.3 3.69417	184.2 3.74002
15	13.4 1.14965	27.5 1.70804	41.3 2.07125	54.6 2.33661	67.2 2.54385	79.3 2.71290	90.7 2.85506	101.6 2.97735	112.1 3.08438	122.1 3.17938	131.7 3.26465	140.9 3.34191	149.8 3.41245	158.3 3.47742	166.6 3.53728	174.6 3.59302	182.4 3.64505	190.0 3.69382	197.3 3.73969
16	14.3 1.14833	29.3 1.70708	44.1 2.07047	58.2 2.33594	71.7 2.54326	84.5 2.71237	96.7 2.85457	108.4 2.97689	119.5 3.08395	130.2 3.17897	140.4 3.26427	150.2 3.34154	159.7 3.41210	168.9 3.47707	177.7 3.53695	186.3 3.59270	194.6 3.64474	202.6 3.69351	210.4 3.73939
17	15.1 1.14717	31.1 1.70623	46.8 2.06978	61.8 2.33535	76.2 2.54274	89.8 2.71190	102.8 2.85413	115.1 2.97649	127.0 3.08338	138.3 3.17861	149.2 3.26393	159.6 3.34123	169.7 3.41178	179.4 3.47677	188.8 3.53666	197.9 3.59242	206.7 3.64447	215.2 3.69325	223.6 3.73913
18	16.0 1.14613	32.9 1.70547	49.5 2.06917	65.5 2.33483	80.6 2.54228	95.1 2.71148	108.8 2.85375	121.9 2.97613	134.4 3.08324	146.4 3.17829	157.9 3.26362	169.0 3.34092	179.7 3.41150	190.0 3.47650	199.9 3.53640	209.5 3.59216	218.8 3.64422	227.9 3.69301	236.7 3.73890
19	16.9 1.14520	34.7 1.70480	52.3 2.06862	69.1 2.33436	85.1 2.54187	100.3 2.71111	114.8 2.85341	128.7 2.97581	141.9 3.08294	154.5 3.17801	166.7 3.26335	178.4 3.34066	189.6 3.41125	200.5 3.47626	211.0 3.53617	221.1 3.59194	231.0 3.64400	240.5 3.69280	249.8 3.73869
20	17.8 1.14437	36.5 1.70419	55.0 2.06813	72.7 2.33391	89.6 2.54149	105.6 2.71077	120.9 2.85310	135.4 2.97552	149.3 3.08267	162.7 3.17775	175.5 3.26311	187.8 3.34042	199.6 3.41103	211.0 3.47605	222.1 3.53596	232.8 3.59174	243.1 3.64380	253.2 3.69260	263.0 3.73850
d_2 cd	1.12838 0.876	1.69257 1.815	2.05875 2.7378	2.32593 3.623	2.53441 4.4658	2.70436 5.2673	2.8472 6.0305	2.97003 6.7582	3.07751 7.4539	3.17287 8.1207	3.25846 8.7602	3.33598 9.3751	3.40676 9.9679	3.47193 10.5396	3.53198 11.0913	3.58788 11.6259	3.64006 12.144	3.68896 12.6468	3.735 13.1362

数据表使用：每一栏的第一行是自由度(*v*)，每一栏的第二行是 d_2^*；d_2 是 d_2^* 的无限值；额外的 *v* 值可以从不同的 cd 常数来建立。

备注：在这表中所使用的备注是采用 Acheson Duncan 所着的*"Quality Control and Industrial Statistics"*, 5th edition McGraw-Hill, 1986.

$(\bar{R}/d_2^*)/\sigma$ 是当一个 X^2 分布于自由度 v 的近似的分布，$\bar{R}$ 是子组大小 *m*、子组数量 *g* 的平均极差

附录 D

量具 R 研究

应用：

- 只能为短期量具重复性研究提供一初步的结论。
- 可能用于从量具供方交运前的检查。
- 可能用于当零件可获得量很少时的生产前准备。
- 可能用于量具的开发期之中 — 例如：对不同的夹持位置，不同的方法的快速比较。
- 没有其它的更完整和更详细的 MSA 方法一起使用时，这方法不能被用于最终的量具接收。

假设：

- 在这时间点上，过程稳定且变差为未知。
- 线性和偏倚没有问题。
- 在此不考虑再现性。目的仅仅专注于量具的重复性。
- 用 *I/MR*（单值 / 移动极差）图来提供稳定性的最小评估。

用以下方法进行分析：

抽取一个零件、一个操作者；将零件固定在夹具上、测量；将零件从夹具上卸下来；让同一操作者重复测量同一零件 9 次以上。

在 *I/MR* 图上绘出数据；评价稳定性。如果数据显示不稳定，采取相应的纠正措施[83]。

如果稳定，用所有读值的 *s* 或用 $\overline{MR}/d_2^*$ 来计算个别的 σ；乘以 6；将所得的数值除以特性的公差；再乘以 100%。评审所得到的重复性%与以前所建立的量具可接受的标准，或将它用为量具开发过程中的比较目的。

[83] 这时可能需要进行一些调整，由单一数据组成的 10 个子组不足以建立稳定性；但可能评价出明显的不稳定性，并且为分析提供价值。

附录 E
用误差修正术语来替代 *PV* 计算

本手册将 *PV* 定义为 $\sqrt{TV^2 - GRR^2}$ 。由于这零件变差的定义中包括了 *EV*，所以当设备变差很重要时，可能常常需要摘取这 *EV* 对 *PV* 的影响。这可通过以下公式来完成(注意：同样可用于 *AV* 的公式中，当 *EV* 的影响被摘取出来)。

$$PV = \sqrt{(R_p \times K_3)^2 - \left[\frac{EV^2}{k \times r}\right]}$$

式中，R_p = 零件平均值的极差，k = 评价人人数，r = 测量次数

用本方法计算 *PV* 的认知是发表于 1997 年[84]。这里介绍的这方法比本手册前面所介绍通用的 *PV* 可接受定义更具有统计上的正确性。通常，当 *EV* 包含 *PV*，它仅为 1%或 2%的程度。

[84] *"Reliable Data is an Important Commodity"*, Donald. S. Ermor 和 Robin Yang E-Hok 著, University of Wiconsin, Madison. 在 *"The Standard"* 发表, ASQ Newsletter of Measurement Quality Division, Vol 97-1, Winter, 1997

附录 F

P. I. S. M. O. E. A. 误差模型

与其它所有过程一样，测量系统受随机的和系统的变差来源的影响。这些变差是由于普通原因和特殊（无次序的）原因造成的。为了理解、控制和改进一测量系统，首先要识别潜在的变差来源。尽管特定的原因会取决于所处的场景，可以用一个通用的误差模型来将影响任何测量系统的变差来源进行分类。用不同的方法来说明并对这些变差来源分类，如使用简单的因果图、矩阵图或诊断树。

首字字母缩写 P. I. S. M. O. E. A.[85] 是另一种用来定义测量系统变差的基本来源的模型。它不仅仅是一种模型，还可支持广泛的应用。

误差来源		同义词或元素	因素或参数
P	零件	生产零件、样本、被测物、被测单元 (UUT)、加工品、检查标准	不知道
I	仪器	量具、M&TE 单元、基准量具、测量机器、试验台	比较的意图
S	标准	比率、参考、加工品、检查标准、固有的标准、一致、标准参考材料 (SRM)、级别、可接受准则	已知的值被*真理* *、参考值，或接受准则所接受
M	方法	在职培训、口头的、作业指导书、控制计划、检验计划、试验计划、零件计划	如何做
O	操作者	评价人、校准或试验技术人员、评估者、检验员	谁来做
E	环境	温度、湿度、污染、清洁、光线、位置、振动、动力、电磁干扰 (EMI)、噪音、时间、空气	测量条件、杂讯
A	假设	统计的、可操作的、校准、恒定的、手册值、热稳定、弹性系数、科学定理	准则、常数，或支持可信赖测量

*真实的或物理的*真值*是不可知的。

同义的误差来源随测量应用及工业而不同。这些是一般的例子。然而，这些参数和特性影响总是相同的。

[85] P. I. S. M. O. E. A. 源自于 Mr. Gordo Skattun 所建立，他是 ASQ CQE 的资深专家。Rock Vally College Technology Center 整合制造的计量专家及总监。

所有测量系统分析的方法均为*质量工具*。所有的质量工具均以假设为基础。如果偏离了假设，该工具最好的情况下是变成不可预测，最坏的情况则将导致错误的结论。

量具 R&R 研究中最典型的统计假设包括：常态的过程、随机和独立的测量、稳定，以及试验 — 再试验的准则。当一个或多个假设不成立时（例如，非常态测量过程，操作者偏倚），这工具和分析最终将变得不稳定、混淆，并且导致误解；对产品和过程控制的*%GRR* 评价可能会被过度估计。另外还有与测量系统相关的非统计上的假设（例如：校准、可操作性、共同因素和扩张的比率、物理定律和常数）。测量的计划者应该能够识别、控制、纠正或更改 MSA 方法，以弥补在测量过程中假设的重大违背。

试验 — 再试验准则的违背通常是在破坏性或线上测量系统的一项考量，这种测量系统的特性是变化的。测量的计划者要对标准的测量系统研究考虑适当的混合、修改，或其它的技术。

假设有助于对总测量变差做出更大贡献，尤其是当：1) 使用更有效的方法或工具（例如 ANOVA、回归分析、设计实验、概率估计和控制图）；以及 2) 测量精密度提高。应用高精密度测量通常必须为测量系统中的共同因素和热膨胀、变形、蠕变或其它假设，实施计划和偶而的应用，以及纠正。

一质量分析者或测量计划者对假设的最大危害，是假设它们是不重要的或是恒定的，因此而常常疏忽了它们。

下表是三种不同测量情况下 PISMOEA 模型的展示范例。

误差来源	典型的生产，自动 MSA	自动化的线上或试验台	校准
P	随机的生产零件、整个过程范围	生产组件、试验样件、检查标准、加工品	量具、UUT、试验样件、被测物
I	单独的生产量具型式	DCC CMM、试验台	基准量具和设备
S	比率、基准标准或分类；符合“10：1 比例原则”	比率和几何学参考的试验标准	基准、参考、固有的或一致的、加工品
M	标准操作程序 (S.O.P)、经常口头的、可能被文件化；控制计划	文件化的 S.O.P、DCC 计划或自动化试验周期	文件化、正式的校准程序
O	（2-3 位）代表性的、经培训的、经常操作的人	受限的试验操作者、特殊的培训与技能	合格的技术员、ISO 17025 专业证据
E	稳定的生产和操作条件	通常受控	控制界限、最佳化、主要的误差来源
A	统计的，通常被忽略	统计的，特定的应用	不能被假设的、主要的误差来源
目的	过程控制 (SPC)	产品控制，100%检验	产品控制，校准公差

特定的误差来源对测量变差的影响程度及贡献率取决于应用的情景。矩阵表、因果分析、或失效树诊断将是识别和理解测量变差主要来源的有用工具，进而控制和改进测量系统。

测量系统分析开始于理解测量的目的与过程，要消除所有浑沌的和不合理的误差来源。测量研究是一个已计划的实验，遵循一些简单的原理：定义重大的误差来源、解决一部份、允许一个或多个控制因素的改变、测量多次、分析结果，以及采取措施。

术语

其它术语的定义见*统计过程控制 (SPC)手册*

5.15 与 $6\sigma_{GRR}$ 相乘系数	参见第 vi 页的备注。
准确度(Accuracy)	一观测值与可接受的参考值之间的一致接近程度。
方差分析 (Analysis of Variance)	通常被用于实验设计 (*DOE*) 的一种统计方法 (*ANVOA*)。用来分析多个群体中的计量型数据，以便比较变差的意义和分析其来源。
可视解析度 (Apparent Resolution)	测量仪器最小的增量大小即为可视解析度。该数值通常广泛的用在公告资料中，以划分测量仪器的等级。数据的分类数可通过该增量大小除以预期的过程分布宽度 (6σ) 来确定。注：显示的或呈现的数值大小并不一定表示该仪器的解析度。例如：零件的被测值分别为 29.075、29.080、29.095、……等，被记录为 5 个数字位量，但该仪器的解析度也许是 0.001，而不是 0.005。
评价人变差 (Appraise Variation)	不同评价人（操作者）使用相同的测量仪器和方法，在一稳定的环境下，对相同零件（被测物）进行测量所得的平均值的变差。评价人变差 (*AV*) 是测量系统变差（误差）的普通原因变差之一，它是由于评价人使用同一测量系统的技能和方法之间的差异所导致的变差。通常将评价人变差假设为与测量系统有关的"再现性误差"。但这种假设不总是正确的(见再现性术语)。
偏倚 (Bias)	测量观测平均值（在重复条件下的测量）与一参考值之间的差值；历史上被称为*准确度*。通过一个单值点是否落在测量系统工作范围内来评价和表述偏倚。
校准 (Calibration)	在规定条件下，建立测量装置与一可追溯且已知参考值和不确定度的标准之间关系的一整套操作活动。校准可能包括以下步骤：检验、纠正、报告，或通过调整来消除被比较的测量装置在准确度方面的任何偏差。
校准周期 (Calibration Interval)	在两次校准之间的特定时间或条件设定。在这段时间内，一测量装置的校准参数是被认为有效的。
能力 (Capability)	基于测量系统的一短期评估，对测量误差(随机的和系统的)的组合变差的一个估计值。
置信度区间 (Confidence Interval)	预期的包括了某一参数的真值的数值范围(在某些要求应用情况下被称为置信水准)。

控制图 (Control Chart)	在以时间为顺序所进行样本测量的基础上的一过程特性图。它用来显示一过程的表现、识别过程变差的模式、评估稳定性，并显示过程的走向。
数据 (Data)	在一定条件下观测值的集合。可以是*计量型*数据（量化的值，并有测量单位）或*抽象的*数值（计数型数据或总数值，例如通过 / 失败、好 / 坏、通 / 止等）。
实验设计 (Designed Experiment)	一种有计划的研究，包括一系列有意图性的对过程要素进行改变与其效果观测，对这些结果进行统计分析以便确定过程变差之间的关系，从而改进这过程。
分辨力 (Discrimination)	与*最小可读单位*同义。分辨力是测量装置和标准的测量解析度、刻度限制、或最小可检出的单位。它是量具设计的固有属性，并通常以测量或分类的单位来呈现。数据的分类数常称为*分辨比率*，因为它描述了对观测到的过程变差，能够可靠的被区隔为多少类别。
区别数据分类 (Distinct Data Categories)	数据的类别数或分类数，它可以被测量系统的有效解析度，以及在实际应用的观测过程中的零件变差来加以可靠的区隔。见 *ndc*。
有效解析度 (Effective Resolution)	当进行整个测量系统变差的数据分类数大小时，要考虑这有效解析度。通过基于测量系统变差的置信度区间范围来确定该分类的大小。区别分类的数量 *ndc*，可以通过把该分类的大小划分在预期的过程分布宽度中来确定。关于有效解析度，这 *ndc*（在 97% 置信水准时）的标准估计值为 1.41[*PV*/*GRR*]。（参见 wheeler 在 1989 年的另一种替代的解释。）
F 比率 (*F* ratio)	是一个统计的数学比率表示，它代表一套的数据组中，数据组之间的平均平方误差与数据组内部的平均平方误差的比值；这些数据是用来评价在一定的自信度水准下随机发生的概率。
量具 R&R (GRR)	一测量系统重复性和再现性的综合变差的估计值。*GRR* 变差等于系统内部和系统之间变差之和。
直方图 (Histogram)	分组数据的频率所显示的一种图示法（长条图）。提供数据分布的目视评价方法。
受控 (In Control)	表示一过程当它只呈现随机、普通原因变差（与混乱的、可查明的，或特殊原因变差相反）的状态。生产中只存在随机变差的过程是稳定的。
独立 (Independent)	一个事件或变数的发生不会影响另一个即将发生的事件或变数的概率。

独立及相等的分布 (Independent and Identically Distributed)	通常简称为 "*iid*"，指一个同质的数据组能独立地并且随机地分布在一普通的分布区域中。
交互作用 (Interaction)	由两个或多个变数产生明显的组合影响或结果。评价人和零件之间的不累加性。评价人的差别取决于被测量零件。
线性 (Linearity)	量具在预期工作范围内偏倚的差值。换句话说，线性是多个且独立的偏倚误差在整个量具的工作行程内的相关性。
长期能力 (Long-Term Capability)	对某过程在一长时间下所显示出其子组内部变差的统计度量。这不同于性能，因为它不包括子组之间的变差。
被测物 (Measurand)	在一定条件下被测量的特定数量或物体；为测量应用的一组已定义的规范。
测量系统 (Measurement system)	测量系统是用来对被测特性进行测量单位的定量或提供评价的仪器或量具、标准、操作、方法、夹具、软件、人员、环境及假设的集合；这整个过程是用来获得测量结果。
测量系统误差 (Measurement System Error)	由于量具偏倚、重复性、再现性、稳定性和线性的综合变差。
度量衡(Metrology)	关于测量的科学。
ndc	区别分类的数量 (number of distinct categories)。1.41 (*PV*/*GRR*)。
不可重复性 (Non-replicable)	由于被测物的动态特性，没有能力对相同样本或零件进行重复测量。
分级数 (Number of Distinct Categories)	见 *ndc*
失控 (Out-of-Control)	指过程存在混乱的、可查明的、或特殊原因变差时的状态。一失控的过程是统计上的不稳定。
零件变差 (Part Variation)	与测量系统分析有关，零件变差 (*PV*) 是指对一个稳定过程中不同零件和不同时间之间出现的预期变差。
零件之间变差 (Part-toPart Variation)	由于测量不同零件造成的变差
性能 (Performance)	基于对测量系统的长期评估，各种测量误差（随机的和系统的）的综合变差估计值；包括整个时期所有显着的和可确定的变差来源。

精密度
(Precision)

在测量系统整个操作范围内（尺寸、范围和时间）分辨力、敏感度和重复性的净影响。有些组织将精密度与重复性互换使用。事实上，精密度更经常被用来描述在测量的整个范围内，重复测量的预期变差；上述范围可能是尺寸或是时间。术语“精密度”通常用在叙述性的场合。

概率 (Probability)

基于某一特定收集数据的分布的一种估计（百分比或分数值），它表示一个特定事件发生的可能性。概率估计值的范围是从 0（不可能的事件）到 1（必然的事件）。共同作用产生某种结果的一组条件或原因。

过程控制
(Process Control)

应用于真实生产中测量目的及判定准则的可操作叙述，其用来评估过程稳定性，并测量或描述过程变差的特征。测量结果指出过程是处于稳定和“受控”或“不受控”的状态

产品控制
(Product Control)

当测量目的及判定准则是用来评估被测物或特性是否符合某一规范要求的可操作叙述。测量的结果为“在公差内”或“不在公差内”。

参考值
(Reference Value)

被认可并同意基于参考或基准值作为一被测物的数值比较，它可能是：

- 一个理论值或基于科学原理而建立的数值；
- 基于一些国家或国际组织的一个指定值；
- 基于在一科学或工程组织主持的合作研究实验工作下，一致确定的数值；或者
- 用于一特定用途，利用一可接受的参考方法所获得一致同意的可接受数值。

与某一特定量化定义并被接受的一致的数值，按照惯例有时被接受用于某已知的目的。

注：与参考值同义的其它术语有：
可接受的参考值
可接受值
惯用值
惯用真值
指定值
该数据的最佳估计值
基准值
基准测量值

回归分析
(Regression
Analysis)

两个或多个变数之间关系的一种统计研究。定义两个或多个变数之间数学关系的一种计算。

重复性 (Repeatability)	从指定测量条件下进行连续测量所产生普遍原因变差、随机变差的测量。重复性通常被称为设备变差 (*EV*)，虽然这是一种误解。重复性的最佳术语为：当测量条件已被固定且已定义 — 零件、仪器、标准、方法、操作者、环境和假设均被固定时，这系统的*内部*变差。除了设备*内部*的变差外，重复性还包括在误差模型中任何条件的*内部*变差。
可重复的 (Replicable)	对相同样本或零件进行重复测量的能力；对被测物或测量环境没有显着的物理影响。
重复 (Replication)	在重复性（相同的）条件下进行的多次测量。
再现性 (Reproducibility)	在测量过程中由正常的条件变化所造成的测量平均值的变差。通常它被定义为：在稳定环境下，用不同评价人（操作者）使用相同的测量仪器与方法，对同一零件（被测物）进行测量所得的平均值的变差。人工仪器的确经常受操作者技能的影响，但是对于操作者不是变差主要原因的测量过程（例如自动化系统），上述说法是不正确的。为这个理由，再现性是指测量的系统*之间*或条件*之间*的平均值变差。
解析度 (Resolution)	可应用于*测量*解析度或*有效*解析度。 测量系统检测并如实指示被测特性微小变化的能力（参见分辨力）。 如果任何零件的指示值与参考零件的差值小于 δ，这零件的指示值将等于参考零件的指示值，当存在一个这种相等的概率，则测量系统的解析度为 δ。测量系统解析度受测量仪器的影响，也受整个测量系统其它变差来源的影响。
散点图 (Scatter Diagram)	用来评价两个变数之间相互关系的数据 X—Y 描点图。
敏感度 (Sensitivity)	一个测量装置的最小输入信号产生可探测的（可识别的）输出信号。一仪器应该至少对其解析度单位刻度值敏感。敏感度取决于固有的量具设计和质量、使用中保养，以及操作条件来确定。它通常被指示为测量的单位。
显著水准 (Significance level)	对试验的随机结果概率所选择的统计水准；也与风险有关，用 α 风险来表示，这表示某一决策误差的概率。

稳定性 (Stability)

指测量过程的*统计*稳定性和长期的*测量*稳定性两种情况。这两种情况对测量系统是否能满足其预期目的来说都是很重要的。统计稳定性是指具有可预测的，测量过程在只有普通原因变差（受控）的情况下操作。测量稳定性（同义词漂移）是指测量系统在整个操作生命周期内，能够满足测量标准或参考的需要。

公差 (Tolerance)

维持配合、形状及功能在一标准或名词数值所允许的偏差。

不确定度 (Uncertainty)

与测量结果相关的一个参数，其描述被测物数值的分散可能被合理的归类 (VIM)；不确定度是对一指定的测量结果范围的描述，在特定的置信度水准内，预期的包含了*实际*的测量结果的范围。不确定度是测量可靠性的一种量化的表达。

单一模型 (Unimodal)

一组具有相同模式的连续数据。

参考文献

1. ASTM E691-99, *Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method*, (www.astm.org).

2. ASTM E177-90a, *Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods*.

3. ASTM E456-96, *Standard Terminology for Relating to Quality and Statistics*

4. *AT&T Statistical Quality Control Handbook*, Delmar Printing Company, Charlotte, NC, 1984.

5. Burdick, Richard K., and Larsen, Greg A., "Confidence Intervals on Measures of Variability in R&R Studies," *Journal of Quality Technology*, Vol. 29, No. 3, July 1997.

6. *Dataplot*, National Institute of Standards and Technology, Statistical Engineering Division, (www.itl.nist.gov)

7. Deming, W. E., *Out of the Crisis*, Massachusetts Institute of Technology, 1982, 1986.

8. Deming, W. E., *The New Economics for Industry, Government, Education*, The MIT Press, 1994, 2000.

9. Duncan, A. J., *Quality Control and Industrial Statistics*, 5th ed., Richard D. Irwin, Inc., Homewood, Illinois, 1986.

10. Eagle, A. R., "A Method For Handling Errors in Testing and Measuring," *Industrial Quality Control*, March, 1954, pp. 10-14.

11. Ermer, Donald S., and E-Hok, Robin Yang, "Reliable Data is an Important Commodity," University of Wisconsin, Madison; published in *The Standard*, ASQ Newsletter of the Measurement Quality Division, Vol. 97-1, Winter 1997.

12. Foster, Lowell W., *GEO-METRICS: The Metric Application of Geometric Tolerancing*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1974.

13. Gillespie, Richard, *A History of the Hawthorne Experiments*, Cambridge University Press, New York, 1991.

14. Grubbs, F. E., "On Estimating Precision of Measuring Instruments and Product Variability," *Journal of the American Statistical Association*, Vol. 43, 1948, pp.243-264.

15. Grubbs, F. E., "Errors of Measurement, Precision, Accuracy and the Statistical Comparison of Measuring Instruments," *Technometrics*, Vol. 15, February 1973, pp. 53-66.

16. Gruska, G. F., and Heaphy, M. S., "Measurement Systems Analysis," TMI Conference Workshop, Detroit, Michigan, 1987.

17. *Guide to the Expression of Uncertainty in Measurement*, ISO/TAG 4/WG 3 (1995).

18. Hicks, Charles R., *Fundamental Concepts in the Design of Experiments*, Holt, Rinehart and Winston, New York, 1973.

19. Hahn, J. H. and Nelson, W., "A Problem in the Statistical Comparison of Measuring Devices," *Technometrics*, Vol. 12., No. 1, February 1970, pp. 95-102.

20. Hamada, Michael, and Weerahandi, Sam, "Measurement System Assessment Via Generalized Inference." *Journal of Quality Technology*, Vol. 32, No. 3, July 2000.

21. Heaphy, M. S., et. al., "Measurement System Parameters," Society of Manufacturing Engineering - IQ81-154, Detroit, Michigan, 1981.

22. *International Vocabulary of Basic and General Terms in Metrology* (abbr. VIM), ISO/IEC/OIML/BIPM (1993).

23. Ishikawa, Kaoru, *Guide to Quality Control*, Asian Productivity Organization, 1986

24. Jaech, J. L., "Further Tests of Significance for Grubbs' Estimators," *Biometrics*, 27, December, 1971, pp. 1097-1101.

25. James, P. D., "Graphical Displays of Gage R&R Data," AQC Transaction, ASQC, 1991.

26. Lipson, C., and Sheth, N. J., *Statistical Design & Analysis of Engineering Experiment*, McGraw-Hill, New York, 1973.

27. Maloney, C. J., and Rostogi, S. C., "Significance Test for Grubbs' Estimators," *Biometrics*, 26, December, 1970, pp. 671-676.

28. Mandel, J., "Repeatability and Reproducibility," *Journal of Quality Technology*, Vol.4, No.2, April, 1972, pp.74-85.

29. McCaslin, J. A., and Gruska, G. F., "Analysis of Attribute Gage Systems," ASQC Technical Conference Transactions, 1976, pp. 392-399.

30. *Measurement Systems Analysis Reference Manual*, Second Edition, Chrysler, Ford, GM, Second Printing, 1998.

31. *Minitab User's Guide 2: Data Analysis and Quality Tools, Release 13 for Windows*, Minitab, Inc., February 2000.

32. Montgomery, Douglas C., and Runger, George C., "Gauge Capability and Designed Experiments, Part I: Basic Methods; Part II: Experimental Design Models and Variance Component Estimation." *Quality Engineering*, Marcel Dekker, Inc., 1993.

33. Nelson, L., "Use of the Range to Estimate Variability," *Journal of Quality Technology*, Vol. 7, No.1, January, 1975, pp.46-48.

34. *Potential Failure Mode and Effects Analysis (FMEA) Reference Manual*, Third Edition, DaimlerChrysler, Ford, GM, 2001.

35. *Production Part Approval Process (PPAP)*, Third Edition, DaimlerChrysler, Ford, GM, 2000.

36. *Statistical Process Control Reference Manual(SPC)*, Chrysler, Ford, GM, Second Printing, 1995.

37. Thompson, W. A., Jr., "The Problem of Negative Estimates of Variance Components," *Annals of Mathematical Statistics*, No. 33, 1962, pp. 273-289.

38. Thompson, W. A., Jr., "Precision of Simultaneous Measurement Procedures", *American Statistical Association Journal*, June, 1963, pp. 474-479.

39. Traver, R. W., "Measuring Equipment Repeatability - The Rubber Ruler," 1962 ASQC Convention Transactions, pp. 25-32.

40. Tsai, Pingfang, "Variable Gauge Repeatability and Reproducibility Study Using the Analysis of Variance Method," *Quality Engineering*, Marcel Dekker, Inc., 1988.

41. Vardeman, Stephen B., and VanValkenburg, Enid S., "Two-Way Random-Effects Analyses and Gauge R&R Studies," *Technometrics*, Vol. 41, No. 3, August 1999.

42. Wheeler, D. J., "Evaluating the Measurement Process When the Testing is Destructive," TAPPI, 1990 Polymers Lamination & Coatings Conference Proceedings, pp. 805-807 (www.tappi.org; plc90905.PDF).

43. Wheeler, D. J., and Lyday, R. W., *Evaluating the Measurement Process*, SPC Press, Inc., Knoxville, Tennessee, 1989.

范例表格

允许读者复制本节中的表格

本节中的每个表格为 GRR 数据的收集和报告的可能格式
只要包含相同的信息，并能够得到相同的结果，不排斥采用其它格式的表格

量具重复性和再现性数据收集表

评价人 / 试验次数	零件										平均值
	1	2	3	4	5	6	7	8	9	10	
A 1											
2											
3											
平均值											$\overline{X}_a =$
极差											$\overline{R}_a =$
B 1											
2											
3											
平均值											$\overline{X}_b =$
极差											$\overline{R}_b =$
C 1											
2											
3											
平均值											$\overline{X}_c =$
极差											$\overline{R}_c =$
零件平均值											$\overline{\overline{X}} =$ $R_p =$
$(\overline{R}_a =\ ] + [\overline{R}_b =\ ] + [\overline{R}_c =\ ])$/[评价人数 =]=											$\overline{\overline{R}}$
$\overline{X}_{DIFF}$ = [最大值$\overline{X}$ =] − [最小值$\overline{X}$ =]=											$\overline{X}_{DIFF}$ =
$*UCL_R$ = [$\overline{\overline{R}}$ =] × [D_4 =]=											

* 2 次测量时 D_4=3.27，3 次测量时 D_4=2.58。UCL_R 代表个别 R 值的控制限。圈出那些超出控制限的点，查明其原因并采取纠正措施。重复以相同的评价人使用相同的量具读取这些值；或剔除这些数值，并从其余的观测值中重新平均和计算 $\overline{\overline{R}}$，以及其限值。

注：_____

量具重复性和再现性报告

零件编号和名称：　　量具名称：　　日期：

特性：　　量具编号：　　执行者：

规范：　　量具型式：

从数据表：$\overline{\overline{R}}=$　　$\overline{X}_{DIFF}=$　　$R_p=$

测量单元分析	总变差% (*TV*)
重复性 — 设备变差 (*EV*) $EV=\overline{\overline{R}}\times K_1$ $=$ _____ × _____ $=$ _____	$\%EV=100[EV/TV]$ $=100[$_____/_____$]$ $=$ _____%
再现性 — 评价人变差 (*AV*) $AV=\sqrt{(\overline{X}_{DIFF}\times K_2)^2-(EV^2/(nr))}$ $=\sqrt{(____\times____)^2-(____^2/(__\times__))}$ $=$ _____ n=零件　　r=试验次数	$\%AV=100[AV/TV]$ $=100[$_____/_____$]$ $=$ _____%
重复性和再现性 (*GRR*) $GRR=\sqrt{EV^2+AV^2}$ $=\sqrt{(____^2+____^2)}$ $=$ _____	$\%GRR=100[GRR/TV]$ $=100[$_____/_____$]$ $=$ _____%
零件变差 (*PV*) $PV=R_p\times K_3$ $=$ _____ × _____ $=$ _____	$\%PV=100[PV/TV]$ $=100[$_____/_____$]$ $=$ _____%
总变差 (*TV*) $TV=\sqrt{GRR^2+PV^2}$ $=\sqrt{(____^2+____^2)}$ $=$ _____	$ndc=1.41\left(PV/GRR\right)$ $=1.41($_____/_____$)$ $=$ _____

试验次数	K_1
2	0.8862
3	0.5908

评价人人数	2	3
K_2	0.7071	0.5231

零件数量	K_3
2	0.7071
3	0.5231
4	0.4467
5	0.4030
6	0.3742
7	0.3534
8	0.3375
9	0.3249
10	0.3146

使用于本表中的理论与常数的资料，请参考 *MSA 参考手册*，第三版。

索引

A

Accuracy 准确度 – 6, 48 (Also see Bias)
Additive Model 附加的模型 – 188
Analysis of Variance (*ANOVA*) 方差分析 – 117
Analytic Method 分析的方法 – 135
Appraiser Variation (*AV*) 评价人变差 – 7, 110, 188
Artifact 加工品 – 6, 41
ASTM (American Society for Testing and Materials) 美国试测与材料协会 – *iii*, 7, 31, 48, 54
Attribute 计数 – 125, 135, 149
Average 平均 – 99, 102, 110, 161

B

Bias 偏倚 – 6, 49, 51, 60, 85-91

C

Calibration 校准 – 25, 37, 38, 41
Calibration Standard 校准标准 – 41
Capability (Measurement System) 能力（测量系统） – 7, 8, 16
Cause and Effect 因果图 – 15, 65
Charts:
    Average 平均值图 – 102
    Error 误差图 – 108
    Range 极差图 – 45, 104
    Run 链图 – 105
Check Standard 检查标准 – 42
CMM 三坐标测量仪 – 12, 23, 57
Complexity 复杂 – 23
Consistency 一致性 – 8, 56
Consumer's Risk 顾客的风险 – 17
Control Charts 控制图 – 8, 47, 84, 88

D

Designed Experiment 实验设计 – 69
Destructive Testing 破坏性试验 – 143-144
Discrimination 分辨力 – 5, 43
Drift (See Stability) 漂移（见稳定性） – 6, 50

E

Ensemble 整体的 – 41
Environment 环境 – 36
Equipment Variation (EV) 设备变差 – 7, 52 (see Repeatability)
Error Charts 误差图 – 108
Ergonomics 人机工程学 – 37

F

Fixturing 夹具 – 36, 39
Flexibility 弹性 – 31
Flow Chart 流程图 – 24, 25
FMEA 失效模式及后果分析 – 11, 24
Funnel Experiment 漏斗实验 – 21

G

Gage 量具 – 5
Gage Repeatability (GR) 量具重复性 – Appendix D, 197
Gage Repeatability and Reproducibility (GRR) 量具重复性与再现性 – 7, 33, 55, 56, 63, 122, 155, 177
Geometric Dimensioning and Tolerancing (GD & T) 几何尺寸与公差 – 28, 159
Guide to the Expression of Uncertainty of Measurement (GUM) 测量不确定度呈现的指南 – 63

H

Histogram 直方图 – 85, 87, 109, 206

I

Interactions 相互作用 – 3, 4, 81, 106-7, 117-122, 187-190, 207

J, K

L

Linearity 线性 – 6, 16, 50, 51, 77, 85, 92-95, 152, 207
Long Term Capability 长期能力 – 207
Lower Control Limit (LCL) 下控制限 – 112
Lower Limit 下限 – 135, 136, 138, 171

M

Maintenance 维护保养 – 8, 24, 25, 29, 30, 32, 35, 37, 50, 53, 55, 69, 91
Master 基准 – 3, 23, 25, 29, 30, 32, 35, 41, 50, 53, 55, 69, 83, 85, 208
Master Standard 基准标准 – 62
Master Value 参考值 – 3, 42
Measurand 被测物 – 59, 61, 207
Measurement Lift Cycle 测量生命周期 – 24
Measurement System 测量系统 – iii, 3, 4, 5, 6, 8
Miss Rate 失误率 – 17, 130, 132

MRA (Mutual Recognition Arrangement) 互相一致的同意 - 9

N

NBS (National Bureau of Standards) 国家标准局 – 9
NIST (National Institute of Standards) 国家标准协会 – 9, 10, 32, 43
NMI (National Measurements Institute) 国家测量协会 – 9, 10, 43
Non-replicable 不可重复的 – 143, 144, 147, 148, 152, 207

O

Observed Process Variance 观察到的过程变差 – 19
Observed Value 观察到的数值 – 18, 108, 109
Operating Range 操作范围 – 6, 8, 50, 52, 73, 74, 87, 92, 94
Operational Definition 可操作的定义 – 5, 6, 10, 12, 13, 40, 43, 65

P

Performance (Measurement) 性能（测量）– 8, 16, 31, 40, 57-58, 77
Performance Curve 性能曲线 – 40, 135, 136, 138, 138-140, 169-174
P. i. S. M. O. E. A. Error Model 零件、仪器、标准、方法、作业者、环境、假设的误差模型 – Appendix F, 201
Plan-Do-Study-Act (PDCA) 计划-执行-研究-行动 – 66
Precision 精密度 – 7, 52, 59, 71, 207
Process Control 过程控制 – 5, 12, 13, 16, 18, 22, 24, 39, 47, 62, 73, 74, 96, 162, 207

R

Random 随机的 – 7, 16, 52, 56, 74, 92, 99, 118, 177
Randomization 随机化 – 117
Range 极差 – 46
Readability (see Discrimination or Resolution) 可读性（见分辨力或解析度） – 5, 43, 49, 50, 53
Reference Value 参考值 – 5, 6, 9, 10, 42, 43, 49, 83, 84, 85, 108, 135-138, 145, 169, 208
Repeatability 重复性 – 7, 48, 52, 56, 58, 60 (and bias), 84, 86, 96-117, 139, 145, 169, 208
Reproducibility 再现性 – 7, 16, 53, 54, 55, 96-117, 169, 179, 209
Replicable 可重复的 – 144, 209
Resolution 解析度 – 5, 13, 39, 43, 44, 46, 74, 96, 209
Resolution:
    Apparent 可视的 – 46, 205
    Effective 有效的 – 5, 13, 37, 39, 44, 46, 206
Risk Analysis Method 风险分析方法 – 125
R & R (See Gage Repeatability and Reproducibility or GRR) 重复性与再现性
Rule of Tens 十分之一规则 – see Ten to One Rule
Run Chart 链图 – 30, 105

S

Scatter Plot 散点图 – 106
Sensitivity 敏感度 – 5, 8, 13, 37, 39, 44, 46, 52, 86, 90, 209
Specification Limits 规范限值 – 13, 17
Stable 稳定的 – 3, 6, 8, 11, 12, 16, 21, 22, 30, 39, 42, 46, 77, 84, 88, 96, 147, 209
Standard 标准 – 5, 9, 11, 12, 30, 31, 34, 38, 41-42, 53, 62, 71-72
Statistical Control 统计学上的控制 – 3, 6, 13, 16, 18, 48, 104, 125
S. W. I. P. E. 标准、工作件、仪器、人 / 程序、环境 – 14

T

Target 目标 – 10, 18, 21
Ten to One Rule (10 to 1) 十分之一比例规则 – 43
Tolerance 公差 – 13, 23, 31, 39, 46, 73, 77, 116, 126, 210
Tolerance Range 公差范围 – 191, 192
Traceability 可追溯性 – 9, 10, 62
Transfer Standard 转换标准 – 41
True Value 真值 – 5, 6, 8, 10, 40, 43, 62, 163, 205
Type I Error 第一类误差 – 17
Type II Error 第二类误差 – 17

U

Uncertainty 不确定度 – 5, 8, 10, 41, 56, 58, 61-63
Uniformity 均一性 – 8, 16, 53, 56, 57, 58
Upper Control Limit (UCL) 上控制限 – 112, 177

V

Variation 变差 – 6, 7, 11, 13, 16, 18, 19, 24, 40, 45, 48, 50, 56, 65, 73
VIM (International Vocabulary of Basic and General Terms in Metrology) 国际基本和通用的度量衡术语词汇 – 9, 58

W

Whiskers Chart Whiskers 图 – 107, 146 (footnote)
Within-Part Variation 零件内部变差 – 53, 97, 123, 159
Working Standard 工作标准 – 4

X

X-Y Plots X-Y 图 – 110, 111

MSA 手册使用者反馈过程

为了持续不断改善的概念，本汽车工业测量系统分析 (MSA) 手册进行正式的周期性评审 / 修订的过程。为达到顾客满意，这种评审不仅每年要充分考虑任何汽车制造商应用的更改要求，而且还要为汽车工业界和使用者沟通，从手册使用者获得反馈意见以使本手册更有效和增加使用价值。为此，欢迎书面提出有关手册理解和易读性的正反两方面的建设性意见。请在下述适当位置填写您的意见，并指出手册具体的页码。按下述地址寄送您的意见：

你的姓名_____

代表_____

公司 / 部门名称

地址_____

_____

_____

电话____ (____) _____

请列出您的三大汽车顾客及其场所：

_____ _____

顾客地址

_____ _____

顾客地址

_____ _____

顾客地址

反馈意见（需要时可附加页）

_____

_____

_____

_____

_____

寄至：

北京品士质量管理顾问有限公司
北京市海淀区学院路蓟门饭店 3508 室　邮编：100088

电话：(8610) 6235 1900;　　传真：(8610) 6200 5456
电子邮件：info@plexus.com.cn;　网站：http://www.plexus.com.cn

请利用 www.aiag.org (or http://www.aiag.org/publications/quality/msa3.html) 以提出您对内容反馈的英文电子文件。