

Content

吉
如
禮

1.Kappa 简介

2.Kappa 测试流程

3.Kappa 测试执行步骤

4.Kappa 判别标准

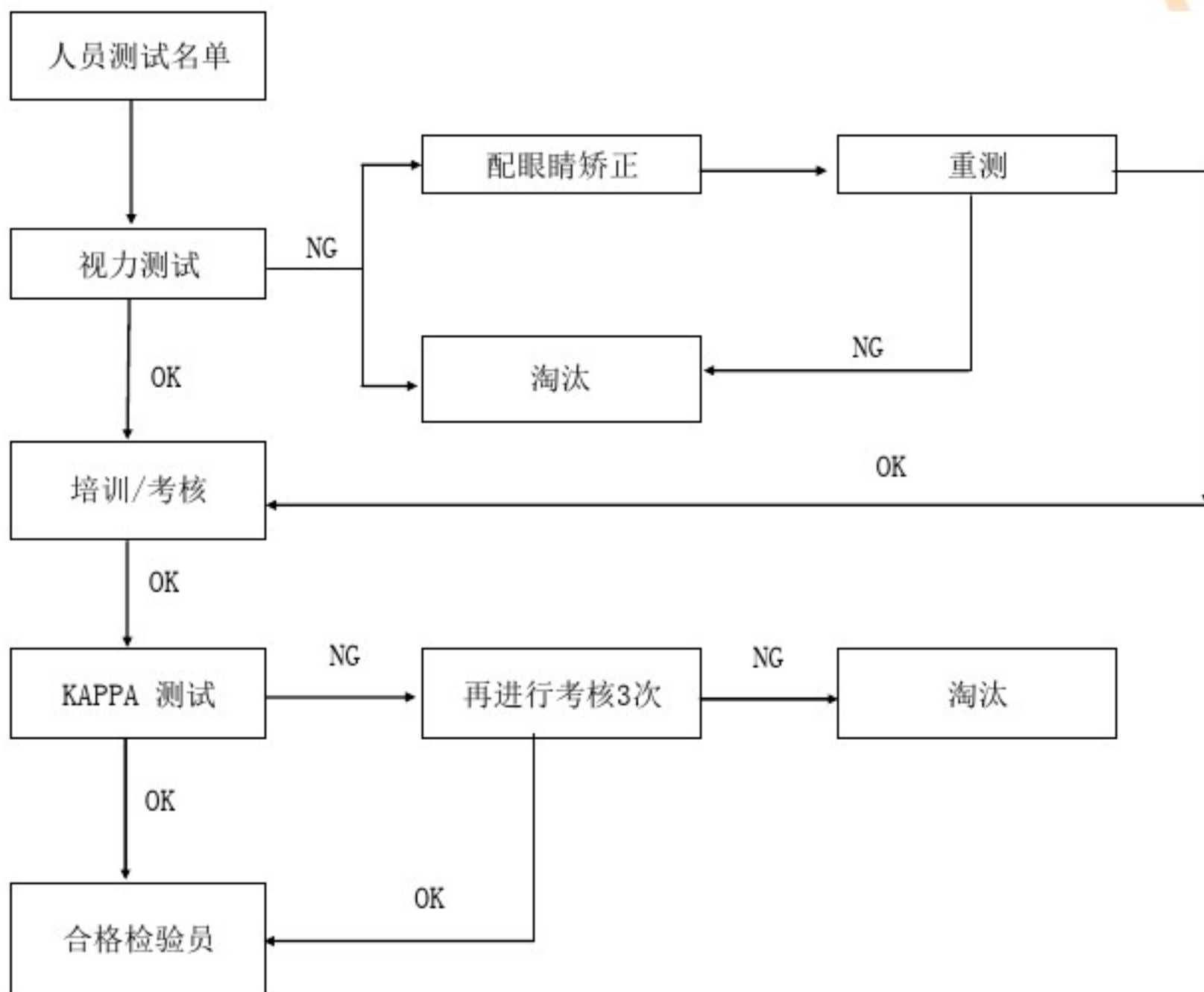


Kappa简介:

- **Kappa**，中文为**卡帕**，是度量测验结果一致程度的统计量。
- ◆ 在计数型测量系统中研究一个测量员重复两次（或测量结果与标准之间的一致性）测量结果的一致性，或者两个测量员的测量结果之间的一致性

KAPPA 测试流程

吉祥



KAPPA 测试执行步骤

吉祥慶

- 1. 测试点及人员确定 (工站, 名单)
- 2. 检验人员视力检查 (矫正视力1.0上)
- 3. 产品检验标准培训 (培训记录)
- 4. 测试样品收集
- 5. Kappa测试 (kappa表)
- 6. Kappa值计算 (交叉表)



◆ 测试样品收集

在选择测试样品时，要考虑以下几方面：

如果只有两个类别：良品和次品，

◆ 至少应该选择**20个良品和20个次品**

◆ 最多可选择 **50个良品和50个次品**

◆ 尽量保持大约**50%的良品和50%的次品**

◆ 选择不同程度的良品和次品

如果样品类别超过2种，其中一类是优良，其它类别是不同种类的缺陷方式

- ◆至少应该选择大约50%的良品和每种缺陷方式最少为10%的产品。比如选择40个样品，其中良品约需20个，其中一类不良缺陷至少需要4个。
- ◆这些类别应该互相排斥，否则它们应该合并起来



KAPPA Test



- ◆ 让每个测试员至少两次判定同一样品
- ◆ 为每个测试员建立独立的Kappa表，计算他们的Kappa值



Kappa表例

吉野

	First Mea.	Second Mea.	First Mea.	Second Mea.	First Mea.	Second Mea.	
Part	Rater A	Rater A	Rater B	Rater B	Rater C	Rater C	
1	Good	Good	Good	Good	Good	Good	
2	Bad	Bad	Good	Bad	Bad	Bad	
3	Good	Good	Good	Good	Good	Good	
4	Good	Bad	Good	Good	Good	Good	
5	Bad	Bad	Bad	Bad	Bad	Bad	
6	Good	Good	Good	Good	Good	Good	
7	Bad	Bad	Bad	Bad	Bad	Bad	
8	Good	Good	Bad	Good	Good	Bad	
9	Good	Good	Good	Good	Good	Good	
10	Bad	Bad	Bad	Bad	Bad	Bad	
11	Good	Good	Good	Good	Good	Good	
12	Good	Good	Good	Bad	Good	Good	
13	Bad	Bad	Bad	Bad	Bad	Bad	
14	Good	Good	Bad	Good	Good	Good	
15	Good	Good	Good	Good	Good	Good	
16	Bad	Good	Good	Good	Good	Good	
17	Bad	Bad	Bad	Good	Bad	Good	
18	Good	Good	Good	Good	Good	Good	
19	Bad	Bad	Bad	Bad	Bad	Bad	
20	Bad	Good	Good	Bad	Bad	Bad	



第一格代表检验员A在第一次和第二次测量中判定为良品的次数

第二格代表检验员A在第一次测量中判定一个样品为次品，在第二次测量中判定为良品的次数

交叉表

		Rater A First Measure		
		Good	Bad	
Rater A Second Measure	Good	10	2	12
	Bad	1	7	8
		11	9	

边格的数字代表行和列的总和

第三格代表检验员A在第一次测量中判定一个样品为良品，在第二次测量中判定为次品的次数

第四格代表检验A在第一次测量和第二次测量中判定一个样品为次品的次数

交叉表-比例

		Rater A First Measure		
		Good	Bad	
Rater A Second Measure	Good	10	2	12
	Bad	1	7	8
		11	9	

下表代表上表的数据，其中每个单元格用总数的百分比来表示

		Rater A First Measure		
		Good	Bad	
Rater A Second Measure	Good	0.5	0.1	0.6
	Bad	0.05	0.35	0.4
		0.55	0.45	

代表10/20

检验员A比例

由行和列的总和计算而得

计算检验员A的Kappa值

$$K = \frac{P_{observed} - P_{chance}}{1 - P_{chance}}$$

$P_{observed}$

- 检验员两次判断一致比率=检验员A两次一致判定为良品的比率+检验员A两次一致判定为次品的比率

P_{chance}

- 预期偶然达成一致的比率=（检验员A第一次判定为良品的比率*第二次判定为良品的比率）+（检验员A第一次判定为次品的比率*第二次判定为次品的比率）

对Kappa的定义将有所不同，取决于我们是在定义检验员内部Kappa，还是在定义检验员之间的Kappa

计算A的Kappa

吉
如
意

		Rater A First Measure		
		Good	Bad	
Rater A Second Measure	Good	0.5	0.1	0.6
	Bad	0.05	0.35	0.4
		$P_{observed} = 0.55$	$P_{chance} = 0.45$	

$$K = \frac{P_{observed} - P_{chance}}{1 - P_{chance}}$$

$P_{observed}$ 等于上表对角线上概率的总和:

$$P_{observed} = (0.500 + 0.350) = 0.850$$

P_{chance} 等于每个分类概率乘积的总和:

$$P_{chance} = (0.600 * 0.55) + (0.400 * 0.45) = 0.51$$

那么 $K_{rater A} = (0.85 - 0.51) / (1 - 0.51) = 0.693$

如果要比较检验员 A 测量结果与标准值之间的一致性，则只需把second Measure改为标准值即可

Minitab工具

吉祥如意

- **Kappa**计算也可以通过软件**Minitab**来实现
- 路径： 打开**Minitab** 软件—打开工作表—统计—质量工具—属性—一致性分析.下表为范例.



kappa test



范例

吉祥

1. 打开工作表，选择kappa test excl.

The screenshot displays the Minitab software interface. A 'Open Worksheet' dialog box is centered on the screen, showing a file list with 'KAPPA test' selected. The dialog box includes fields for 'Search Range' (桌面), 'File Name' (KAPPA test), and 'File Type' (所有 (*.*)). The background shows the Minitab main window with a spreadsheet grid and a menu bar. The taskbar at the bottom shows the Start button and several open applications including 'KAPPA', 'Microsoft PowerPoint', and 'Minitab - 无标题'.

吉祥

吉祥

吉祥

2. 点击统计—质量工具—属性—一致性分析

The screenshot shows the Minitab software interface. The 'Stat' menu is open, and the path 'Stat > Quality Tools > Attribute > Attribute Consistency' is highlighted. The spreadsheet data is as follows:

	C1-I 检验员	C2 样品	C3-I 判定值
1	A	1	Good
2	A	1	Good
3	B	1	Good
4	B	1	Good
5	C	1	Good
6	C	1	Good
7	A	2	Bad
8	A	2	Bad
9	B	2	Good
10	B	2	Bad
11	C	2	Bad
12	C	2	Bad
13	A	3	Good
14	A	3	Good
15	B	3	Good
16	B	3	Good
17	C	3	Good
18	C	3	Good
19	A	4	Good
20	A	4	Bad
21	B	4	Good
22	B	4	Good
23	C	4	Good
24	C	4	Good
25	A	5	Bad
26	A	5	Bad
27	B	5	Bad
28	B	5	Bad
29	C	5	Bad
30	C	5	Bad
31	A	6	Good
32	A	6	Good
33	B	6	Good

Statistical menu options visible:

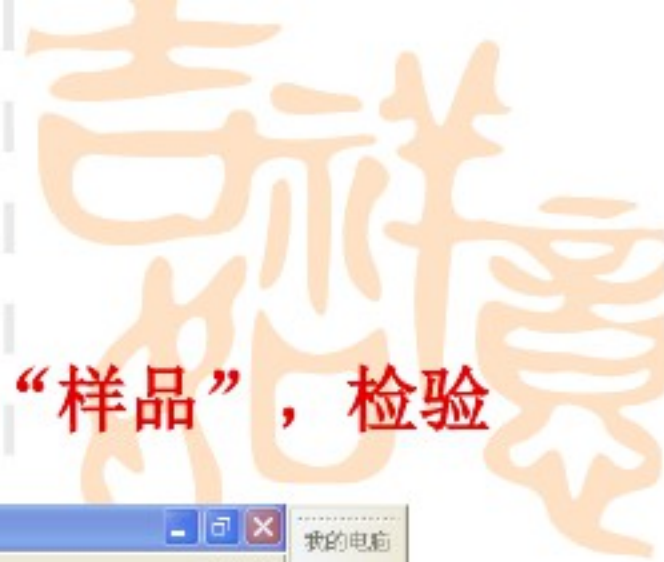
- 基本统计量 (Basic Statistics)
- 回归 (Regression)
- 方差分析 (ANOVA)
- DOE (DOE)
- 控制图 (Control Charts)
- 质量工具 (Quality Tools) - Selected
- 可靠性/生存 (Reliability/Survival)
- 多变量 (Multivariate)
- 时间序列 (Time Series)
- 表格 (Tables)
- 非参数 (Nonparametric)
- EDA (EDA)
- 功效和样本数量 (Power and Sample Size)

Quality Tools sub-menu options:

- 运行图 (Run Chart)
- Pareto 图 (Pareto Chart)
- 因果 (Cause-and-Effect)
- 个体分布标识 (Individual Distribution Identification)
- Johnson 变换 (Johnson Transformation)
- 能力分析 (Capability Analysis)
- Capability Sixpack (Capability Sixpack)
- 量具研究 (Gage Study)
- 属性一致性分析 (Attribute Consistency) - Selected
- 按属性抽样验收 (Attribute Sampling Acceptance)
- 按变量抽样验收 (Variable Sampling Acceptance)
- 多变异图 (Multivariate Chart)
- 对称图 (Symmetry Chart)

Taskbar: 开始 | 收件箱 - Microso... | XAPPA | B0313395 - Yindo... | Microsoft PowerP... | Minitab - 无标题... | 9:30





3. 输入数据：属性列（T）为“判定值”，样本（L）为“样品”，检验员（A）为“检验员”

The screenshot shows the Minitab software interface with a data table and a dialog box for '属性一致性分析' (Attribute Consistency Analysis).

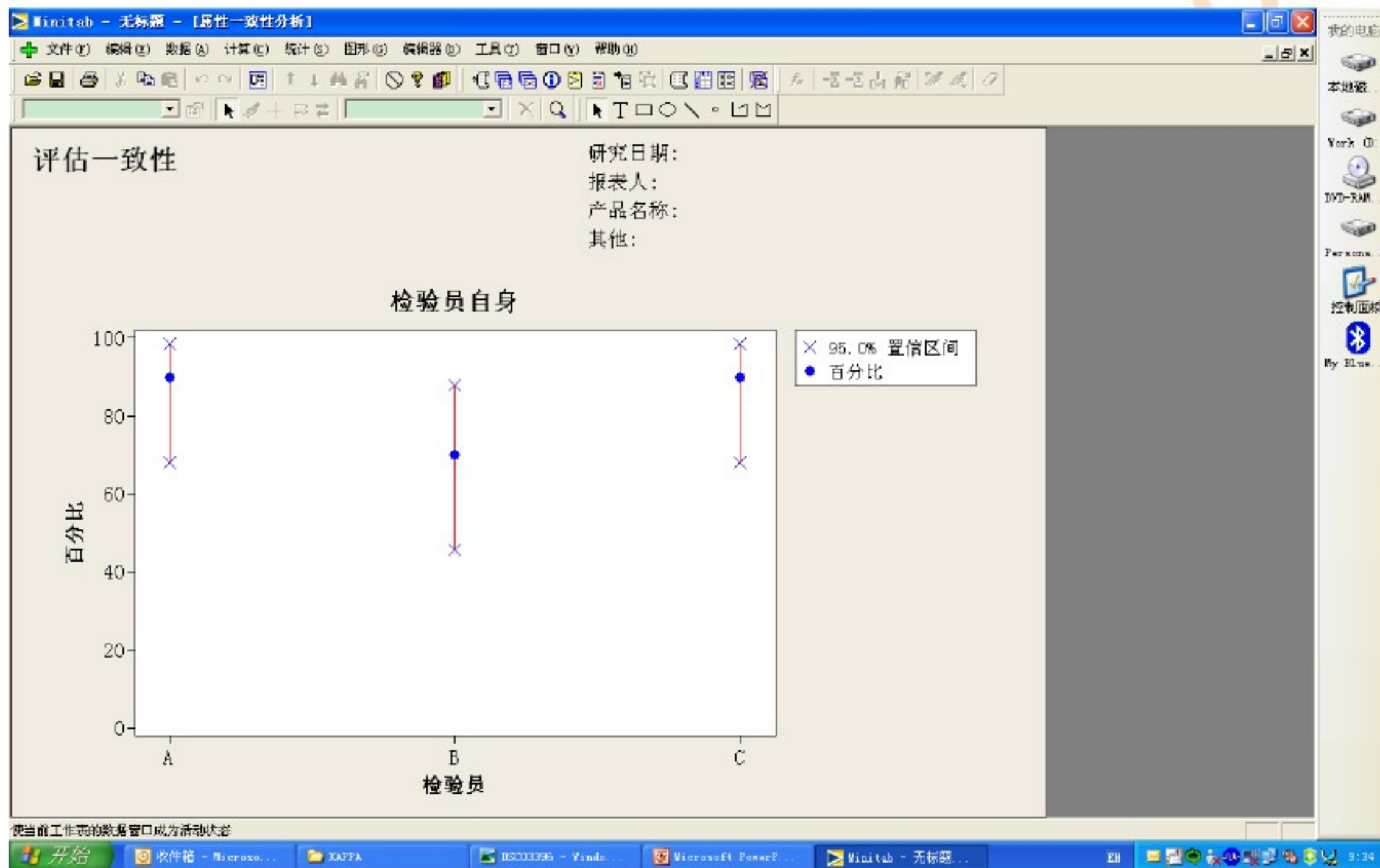
	C1-T	C2	C3-T	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
	检验员	样品	判定值																
1	A		1 Good																
2	A		1 Good																
3	B		1 Good																
4	B		1 Good																
5	C		1 Good																
6	C		1 Good																
7	A		2 Bad																
8	A		2 Bad																
9	B		2 Good																
10	B		2 Bad																
11	C		2 Bad																
12	C		2 Bad																
13	A		3 Good																
14	A		3 Good																
15	B		3 Good																
16	B		3 Good																
17	C		3 Good																
18	C		3 Good																
19	A		4 Good																
20	A		4 Bad																
21	B		4 Good																
22	B		4 Good																
23	C		4 Good																
24	C		4 Good																
25	A		5 Bad																
26	A		5 Bad																
27	B		5 Bad																
28	B		5 Bad																
29	C		5 Bad																
30	C		5 Bad																
31	A		6 Good																
32	A		6 Good																
33	B		6 Good																

The dialog box '属性一致性分析' (Attribute Consistency Analysis) is open, showing the following settings:

- 数据排列方式 (Data Arrangement):
 - 属性列 (T): 判定值 (Attribute Column)
 - 样本 (L): 样品 (Sample)
 - 检验员 (A): 检验员 (Inspector)
- 多列 (M): (For each inspector, enter the trials)
- 已知标准/属性 (K): (Optional)
- 属性数据的类别是有序的 (O): (Optional)



4. “确定”后可以得到如下结果：检验员自身的一致性的95%置信区间



结果: Sheet1

判定值 的属性一致性分析

检验员自身

评估一致性

检验员	# 检验数	# 相符数	百分比	95 % 置信区间
A	20	18	90.00	(68.30, 98.77)
B	20	14	70.00	(45.72, 88.11)
C	20	18	90.00	(68.30, 98.77)

相符数: 检验员在多个试验之间, 他/她自身标准一致。

检验员自身的kappa

Fleiss Kappa 统计量

检验员	响应	Kappa	Kappa 标准误	Z	P (与 > 0)
A	Bad	0.791667	0.223607	3.54044	0.0002
	Good	0.791667	0.223607	3.54044	0.0002
B	Bad	0.375000	0.223607	1.67705	0.0468
	Good	0.375000	0.223607	1.67705	0.0468
C	Bad	0.791667	0.223607	3.54044	0.0002
	Good	0.791667	0.223607	3.54044	0.0002

Cohen 的 Kappa 统计量

检验员	响应	Kappa	Kappa 标准误	Z	P (与 > 0)
A	Bad	0.791667	0.223607	3.54044	0.0002
	Good	0.791667	0.223607	3.54044	0.0002
B	Bad	0.375000	0.223607	1.67705	0.0468
	Good	0.375000	0.223607	1.67705	0.0468
C	Bad	0.791667	0.223607	3.54044	0.0002
	Good	0.791667	0.223607	3.54044	0.0002

检验员	响应	Kappa	Kappa 标准误	Z	P(与 > 0)
A	Bad	0.791667	0.223607	3.54044	0.0002
	Good	0.791667	0.223607	3.54044	0.0002
B	Bad	0.375000	0.223607	1.67705	0.0468
	Good	0.375000	0.223607	1.67705	0.0468
C	Bad	0.791667	0.223607	3.54044	0.0002
	Good	0.791667	0.223607	3.54044	0.0002

检验员之间

评估一致性

# 检验数	# 相符数	百分比	95 % 置信区间
20	13	65.00	(40.78, 84.61)

相符数: 所有检验员的评估一致。

检验员之间的 kappa

Fleiss Kappa 统计量

响应	Kappa	Kappa 标准误	Z	P(与 > 0)
Bad	0.694444	0.0577350	12.0281	0.0000
Good	0.694444	0.0577350	12.0281	0.0000

Cohen 的 Kappa 统计量

要计算 Kappa, 必须具有两个检验员, 每个检验员进行一个试验。

属性一致性分析

Kappa量测能力评价指标

	判断基准 (良好)	判断基准 (考虑)	判断基准 (不足)
判断指标	90% ↑	75~90%	75% ↓

上面的判断基准根据Project的目标有可能变更。

举个例，不良率0.1%→0.001% 改善课题的情况下检出率必须为100%。（我们的目标VLRR为100DPPM）

满足指数40% → 60% 改善课题时，如果检出率是75%以上就可以被选择。

基本上，达不到100%时，必须调查其原因。

KAPPA测试频率

吉祥慶

- ---定期进行kappa测试
- >测试周期：每月/1次
- >受测对象：品保部QC & 制造单位全检人员



- 属性型测量系统的Fleiss Kappa分析

- 标签：[属性型](#)

- [测量系统分析](#)

- [fleiss](#)

- [kappa](#)

- [杂谈](#)

- 分类：[质量工具与方法](#)

- Kappa分析是属性型测量系统分析的典型方法，在AIAG的MSA手册中推荐的是Cohen Kappa分析法。Cohen方法一个最大的缺陷是只能按每个评价者相同的评价序列进行比较（例如：有两个评价者各测量了3轮，则对于同一个被测对象，只能A1与B1比较、A2与B2比较、A3与B3比较，A1的结果不能与B2和B3进行比较，A2的结果不能与B1和B3进行比较，A3的结果不能与B1和B2进行比较）。而Fleiss Kappa分析法则克服了这种缺陷，Fleiss方法是将测量同一对象的所有测量结果进行两两比较而得到结果一致的次数（例如：A1=0、A2=1、B1=1、B2=1，则结果一致的情形有：A2B1、A2B2、B1A2、B1B2、B2A2、B2B1，即结果一致的次数为6次。这里在比较时是两两互相比较，即“你”与“我”比较后，“我”还要与“你”比较，也就是两两之间实际比较了两次），依此方法得到一致性的比较结果，在此基础上计算Kappa值。

若要全面解释Fleiss Kappa方法，需要很长的篇幅及图表来说明。以下是一个示例，感兴趣的读者可以体会Fleiss Kappa是如何计算出来的。

分类	0=Not OK 1=OK							A×A	B×B	A×B	A1×R	A2×R	A3×R	B1×R	B2×R	B3×R	
	检验对象 编号	参考 值R	评价者 - 试验数														
			A1	A2	A3	B1	B2										B3
1	1	1	1	1	1	1	1	6	6	30	2	2	2	2	2	2	
2	1	1	1	1	1	1	1	6	6	30	2	2	2	2	2	2	
3	0	1	0	0	0	0	0	2	6	20	0	2	2	2	2	2	
4	0	0	0	0	0	1	0	6	2	20	2	2	2	2	0	2	
5	0	0	0	1	0	0	0	2	6	20	2	2	0	2	2	2	
观察到的成对一致性组合数 n^*								22	26	120	8	10	8	10	8	10	
可能的成对一致性组合数 N^*								30	30	150	10	10	10	10	10	10	
观察到成对一致性组合数的概率 $P_{obs} = n^*/N^*$								0.73	0.87	0.80	0.80	1.00	0.80	1.00	0.80	1.00	
观察到的每个分类的评价数 n_k ($k=0, \dots, N_c$)								7	8	15	5	6	5	6	5	6	
总评价次数N								15	15	30	10	10	10	10	10	10	
观察到的每个分类的评价数占比 n_k/N								0.467	0.533	0.5	0.5	0.6	0.5	0.6	0.5	0.6	
								0.533	0.467	0.5	0.5	0.4	0.5	0.4	0.5	0.4	



随机一致性评估的期望概率			0.502	0.502	0.5	0.5	0.52	0.5	0.52	0.5	0.52	
$P_{exp} = \sum_k (n_k/N)^2$			0.502	0.502	0.5	0.5	0.52	0.5	0.52	0.5	0.52	
Kappa: $(P_{obs} - P_{exp}) / (1 - P_{exp}) = k$			0.464	0.732	0.600	0.6	1	0.6	1	0.6	1	
每个评价者比对于参考值的Kappa						0.733			0.867			
所有评价者比对于参考值的Kappa						0.800						
最小Kappa值决定最终结果												
评价者姓名	代码	评价者 Kappa	≥0.9	≥0.7-0.9	<0.7	评价者比对参考值的Kappa	≥0.9	≥0.7-0.9	<0.7			
	A	0.464			×	0.733			×			
	B	0.732		×		0.867			×			
所有评价者之间的Kappa						所有评价者比对参考值的Kappa						
所有人		0.600			×	0.800					×	
最终结果			所有结果中的最小值						Kappa =		0.464	
Kappa ≥ 0.9		能力满足要求	0.9 > Kappa ≥ 0.7			能力有条件满足要求	Kappa < 0.7	×	能力不满足要求			



習字

THANKS!

