

第六章 差别试验

差别试验要求品评员评定两个或两个以上的样品中是否存在感官差异（或偏爱其一）。它是感官分析中经常使用的两类方法之一。它是让品评员回答两种样品之间是否存在不同，一般不允许“无差异”的回答（即强迫选择），即品评员未能觉察出样品之间的差异。差别试验的结果分析是以每一类别的评定员数量为基础的。例如：有多少人回答样品 A，多少人回答样品 B，多少人回答正确。结果的解释基于频率和比率的统计学原理，根据能够正确挑选出产品差别的品评员的比率来推算出两种产品间是否存在差异。

差别试验的应用很广。有些情况下，研究者的目的在于确定两种样品是否不同。而有些情况下，研究者的目的是区分两种样品是否相似。以上这两种情况可通过选择合适的试验敏感参数，如 α 、 β 、P。

α ，也叫 α -风险，它的定义是错误的估计两者之间的差别存在的可能性。也叫第 I 类错误。

β ，也叫 β -风险，它的定义是错误估计两者之间的差异不存在的可能性。也叫第 II 类错误。

P，是指能分辨出的差异的人数比例。

在以寻找差异为目的的差别试验中，只需要考虑 α 值，而 β 值和 P 值通常不需要考虑。在以寻找相似性为目的的差别试验中，试验者要考虑合适的 P 值，然后确定一个较小的 β 值， α 值可以大一些。而某些情况下，试验者要综合考虑 α 、 β 、P 值，这样才能保证参与评定的人数在可能的范围之内。

如果样品间的差别非常大，以至很明显，则差别试验是无效的。当样品间的差别很微小时，差别试验是有效的。在试验中需要注意样品外表、形态、温度和数量等所引起的误差。差别试验中常用的方法有：成对比较检验法、二-三点检验法、三点检验法、“A”-“非 A”检验法、五中取二检验法、选择试验法、配偶试验法。

第一节 成对比较检验法（paired comparison test）

以随机顺序同时出示两个样品给评价员，要求评价员对这两个样品进行比较，判定整个

样品或者某些特征强度顺序的一种评价方法称为成对比较检验法或者两点检验法。成对比较试验有两种形式，一种叫做差别成对比较（双边检验），也叫简单差别试验和异同试验，另一种叫定向成对比较法（单边检验）。决定采取哪种形式的检验，取决于研究的目的。如果感官评价员已经知道两种产品在某一特定感官属性上存在差别，那么就应采用定向成对比较试验。如果感官评价员不知道样品间何种感官属性不同，那么就应采用差别成对比较试验。

一 方法特点

以下分别介绍定向成对比较法和差别成对比较法的特点，并对试验中的注意事项进行阐述。

1. 定向成对比较法

在定向成对比较试验中，受试者每次得到 2 个（一对）样品，组织者要求回答这些样品在某一特性方面是否存在差异，比如在甜度、酸度、红色度、易碎度等。两个样品同时呈送给评价员，要求评价员识别出在这一指定的感官属性上程度较高的样品。

- 1) 试验中，样品有两种可能的呈送顺序（AB、BA），且呈送顺序应该具有随机性，评价员先收到样品 A 或样品 B 的概率应相等。
- 2) 评价员必须清楚地理解感官专业人员所指定的特定属性的含义。评价员不仅应在识别指定的感官属性方面受过专门训练，而且在如何执行评分单所描述的任务方面也应受过训练。
- 3) 该检验是单向的。定向成对比较检验的对立假设是：如果感官评价员能够根据指定的感官属性区别样品，那么在指定方面程度较高的样品，由于高于另一样品，因此被选择的概率较高。该检验结果可给出样品间指定属性存在差别的方向。
- 4) 感官专业人员必须保证两个样品只在单一的所指定的感官方面有所不同，否则此检验法则不适用。比如，增加蛋糕中的糖加量，会使蛋糕变的比较甜，但同时会改变蛋糕的色泽和质地。在这种情况下，定向成对比较法并不是一种很好的区别检验方法。

2 差别成对比较法

评价员每次得到 2 个（1 对）样品，被要求回答样品是相同还是不同。在呈送给评价员的样品中，相同和不相同的样品数是一样的。通过比较观察的频率和期望的频率，根据 χ^2

分布检验分析结果。

- 1) 差别成对比较试验中，样品有四种可能的呈送顺序（AA、BB、AB、BA）。这些顺序应在评价员中交叉进行随机处理，每种顺序出现的次数相同。
- 2) 评价员的任务是比较两个样品，并判断它们是相同还是相似。这种工作比较容易进行。评价员只需熟悉评价的感官特性，可以理解评分单中所描述的任务，但他们不需要接受评价特定感官属性的训练。一般要求 20~50 名品评人员来进行试验，最多可以用 200 人，或者 100 人。试验人员要么都接受过培训，要么都没接受过培训，但在同一个试验中，参评人员不能既有受过培训的也有没受过培训的。
- 3) 该检验是双边的。差别成对比较检验的对立假设规定，样品之间可觉察出不同，而且品评员可正确指出样品间是相同或不相同的概率大于 50%。此检验只表明评价员可辨别两种样品，并不表明某种感官属性方向性的差别。
- 4) 当试验的目的是要确定产品之间是否存在感官上的差异，而产品由于供应不足而不能同时呈送 2 个或多个样品时，选取此试验较好。

3. 成对比较法试验的注意事项

- 1) 成对比较检验法是最简便也是应用最广泛的感官检验方法，它常被应用于食品的风味检验。如偏爱检验。在偏爱检验中，一般应了解两种样品间，哪一种更受欢迎。此方法也常被用于训练评价员，在其筛选、考核、培训中是常用的方法。
- 2) 进行成对比较检验时，从一开始就应分清是差别成对比较还是定向成对比较。如果检验目的只是关心两个样品是否不同，则是差别成对比较，如果想具体知道样品的特性，比如哪一个更好，更受欢迎，则是定向成对比较。
- 3) 成对比较检验法具有强制性。在成对比较检验法中有可能会出现“无差异”的结果，通常这是不允许的，因而要求评价员“强制选择”，以促进鉴评员仔细观察分析，从而得出正确结论。尽管两者反差不强烈，但没有给你下“没有差异”结论的权力，故必须下一个结论。在评价员中可能会出现“无差异”的反应，有这类人员时用强制选择可以增加得出有效结论的机会，即“显著结果的机会”。这个方法的缺点是鼓励人们去猜测，不利于评定人员忠诚地去纪录“无差异”的结果，出现这种情况时，实际上是相当于减少了鉴评员的人数。因此要对评价员进行培训，以增强对样品的鉴别能力。减少这种错误的发生。
- 4) 因为该检验方法容易操作，因此没有受过培训的人都可以参加，但是他必须熟悉要

评价的感官特性。如果要评价的是某项特殊特性，则要使用受过培训的人员。因为这种检验方法猜对的可能是 50%，因此需要参加的人员的人数要多一点，从表 6.3 可以知道，如果参加人员的人数是 15 人，要达到 $\alpha = 0.01$ 水平下的显著差异，必须有 13 人同时同意才行。如果参加人数是 60 人，只要有 40 人意见一致就可以达到 $\alpha = 0.01$ 的显著水平。

二 问答表的设计和做法

问答表的设计应和产品特性及试验目的相结合。一般常用的问答表如 6-1、6-2、6-3、6-4、6-5 所示。呈送给受试者两个带有编号的样品，要使组合形式 AB 和 BA 数目相等，并随机呈送，要求受试者从左到右尝试样品，然后填写问卷。

表 6-1 差别成对比较检验问答表示例

<p>异同试验</p> <p>姓名：_____ 日期：_____</p> <p>样品类型：_____</p>
<p>试验指令：</p> <ol style="list-style-type: none"> 1. 从左到右品尝你面前的两个样品。 2. 确定两个样品是相同还是不同。 3. 在以下相应的答案前面划 \checkmark
<p>_____两个样品相同</p> <p>_____两个样品不同</p>
<p>评语：</p>

表 6-2 差别成对比较检验常用问卷示例

<p>日期：_____</p>

姓名：_____

检验开始前请用清水漱口。两组成对比较试验中各有两个样品需要评价，请按照呈送的顺序品尝各组中的编码样品，从左至右，由第一组开始。将全部样品摄入口中，请勿再次品尝。回答各组中的样品是相同还是不同？圈出相应的词。在两种样品品尝之间请用清水漱口，并吐出所有的样品和水。然后进行下一组的试验，重复品尝程序。

组别

- | | | |
|----|----|----|
| 1. | 相同 | 不同 |
| 2. | 相同 | 不同 |

表 6-3 定向成对比较调查问卷示例

日期：_____

姓名：_____

检验开始前，用清水漱口。分别对两组定向成对比较试验中的两个样品进行评价。请按照样品呈送程序品尝各组中的编码样品，从左向右，由第一组开始。将全部样品放入口中，请勿再次品尝。在每一对中圈出较甜样品的代码。在品尝一种样品后，即品尝下一个样品前，应用清水漱口，并吐出所有的样品和水。然后，进行下一组品尝，重复品尝程序。

组例

- | | | |
|----|-------|-------|
| 1. | _____ | _____ |
| 2. | _____ | _____ |

表 6-4 定向成对比较试验问答表示例

定向成对比较试验

姓名：_____

日期：_____

试验指令：在你面前有 2 个样品，从左到右依次品尝这 2 个样品，在你认为甜的样品编号上画圈。你可以猜测，但必须有所选择。

111

123

三 结果分析与判断

根据 A、B 两个样品的特性强度的差异大小，确定检验是差别成对比较还是定向成对比

较。如果样品 A 的特性强度（或被偏爱）明显优于 B，换句话说，参加检验的评价员，作出样品 A 比样品 B 的特性强度大（或被偏爱）的判断概率大于作出样品 B 比样品 A 的特性强度大（或被偏爱）的判断概率，即 $P_A > 1/2$ 。例如，两种饮料 A 和 B，其中饮料 A 明显甜于样品 B，则该检验是定向成对比较；如果这两种样品有显著差别，但没有理由认为 A 或 B 的特性强度大于对方或被偏爱，则该检验是差别成对比较（双边检验）。

- (1) 对于单边检验，统计有效回答表的正解数，此正解数与表 6-5 中相应的某显著性水平的数相比较，若大于或等于表中的数，则说明在此显著水平上，样品间有显著性差异，或认为样品 A 的特性强度大于样品 B 的特性强度（或样品 A 更受偏爱）
- (2) 对于双边检验，统计有效回答表的正解数，此正解数与表 6-6 中相应的某显著性水平的数相比较，若大于或等于表中的数，则说明在此显著水平上，样品间有显著性差异，或认为样品 A 的特性强度大于样品 B 的特性强度（或样品 A 更受偏爱）
- (3) 表中 n 值大于 100 时，答案最少数按以下公式计算，取最接近的整数值。

$$X = \frac{n+1}{2} + K\sqrt{n}$$

式中，K 值为：

显著水平	5%	1%	0.1%
单边检验 K 值	0.82	1.16	1.55
双边检验 K 值	0.98	1.29	1.65

表 6-5 二-三点检验和成对比较检验（单边）法检验表

答案 数目 (n)	显著水平			答案 数目 (n)	显著水平			答案 数目 (n)	显著水平		
	5%	1%	0.1%		5%	1%	0.1%		5%	1%	0.1%

7	7 7 -	24	17 19 20	41	27 29 31
8	8 8 -	25	18 19 21	42	27 29 32
9	9 9 -	26	18 20 22	43	28 30 32
10	10 10 10	27	19 20 22	44	28 31 33
11	9 10 11	28	19 21 23	45	29 31 34
12	10 11 12	29	20 22 24	46	30 32 34
13	10 12 13	30	20 22 24	47	30 32 35
14	11 12 13	31	21 23 25	48	31 33 35
15	12 13 14	32	22 24 26	49	32 34 36
16	12 14 15	33	22 24 26	50	32 34 37
17	13 14 16	34	23 25 27	60	37 40 43
18	13 15 16	35	23 25 27	70	43 46 49
19	14 15 17	36	24 26 28	80	48 51 55
20	15 16 18	37	24 27 29	90	54 57 61
21	15 17 18	38	25 27 29	100	59 63 66
22	16 17 19	39	26 28 30		
23	16 18 20	40	26 28 31		

表 6-6 成对比较检验（双边）法检验表

答案 数目 (n)	显著水平			答案 数目 (n)	显著水平			答案 数目 (n)	显著水平		
	5%	1%	0.1%		5%	1%	0.1%		5%	1%	0.1%
7	7	-	-	24	18	19	21	41	28	30	32
8	8	8	-	25	18	20	21	42	28	30	32
9	8	9	-	26	19	20	22	43	29	31	33
10	9	10	-	27	20	21	23	44	29	31	34
11	10	11	11	28	20	22	23	45	30	32	34

12	10 11 12	29	21 22 24	46	31 33 35
13	11 12 13	30	21 23 25	47	31 33 36
14	12 13 14	31	22 24 25	48	32 34 36
15	12 13 14	32	23 24 26	49	32 34 37
16	13 14 15	33	23 25 27	50	33 35 37
17	13 15 16	34	24 25 27	60	39 41 44
18	14 15 16	35	24 26 28	70	44 47 50
19	15 16 17	36	25 27 29	80	50 52 56
20	15 17 18	37	25 27 29	90	55 58 61
21	16 17 19	38	26 28 30	100	61 64 67
22	17 18 19	39	27 28 31		
23	17 19 20	40	27 29 31		

实例分析：

例 1 成对比较法 饮料的甜度

某饮料厂生产有四种饮料，编号分别为“798”、“379”、“527”和“806”。其中，两种编号为“798”和“379”的饮料，其中一个略甜，但两者都有可能使评价员感到更甜。编号为“527”和“806”的两种饮料，其中“527”配方明显较甜。请通过成对比较试验来确定哪种样品更甜，您更喜欢哪种样品。

试验设计与分析

在统计学分析中，在得出某一结论之前，应事先选定某一显著性水平。所谓显著性水平，是当原假设是真而被拒绝的概率（或这种概率的最大值），也可看作为得出这一结论所犯错误的可能性。在感官分析中，通常选定 5% 的显著性水平可认为是足够的。原假设一般是这样：两种样品之间在特性强度上没有差别（或对其中之一没有偏爱）。应当注意：原假设可能在“5% 的水平”上被拒绝，而在“1% 的水平”上不被拒绝。如果原假设在“1% 水平”上被拒绝，则在“5% 水平”上更被拒绝。因此，对 5% 的水平用“显著”一词表示，而对 1% 用“非常显著”一词表示。本例选择 5% 显著水平（ $\alpha \leq 0.05$ ）。

两种饮料编号为“798”和“379”，其中一个略甜，但两者都有可能使评价员感到更甜。

属双边检验。编号为“527”和“806”的两种饮料，其中“527”配方明显较甜，属单边检验。调查问卷如表 6-7 所示。

表 6-7 成对比较试验调查问卷

姓名： _____

产品： _____

日期： _____

(1) 请评价您面前的两个样品，两个样品中哪个更 _____ 甜。

(2) 两个样品中，您更喜欢的是 _____ 。

(3) 请说出您的选择理由： _____ 。

结果分析：

共有 30 名优选评价员参加鉴评，统计结果如下：

(1) 18 人认为“798”更甜，12 人选择“379”更甜。

(2) 22 人回答更喜欢“379”，8 人回答更喜欢“798”。

(3) 22 人认为“527”更甜，8 人回答“806”更甜。

(4) 23 人回答更喜欢“527”，7 人回答更喜欢“806”。

(1)、(2) 属双边检验。查表“798”和“379”两种饮料甜度无明显差异（接受原假设），“379”饮料更受欢迎。

(3)、(4) 属单边检验。查表“527”比“806”更甜（拒绝原假设），“527”饮料更受欢迎。

例 2 定向成对比较检验 啤酒的苦味

问题：某啤酒酿造商得到的市场报告称，他们酿造的啤酒 A 不够苦。该厂又使用了更多的酒花酿制了啤酒 B。

项目目标：生产一种苦味更重一些的啤酒，但不要太重。

试验目标：对啤酒 A 和 B 进行对比，看两者之间是否在苦味上存在虽然很小但却是显著的差异。

试验设计：选用方向性差异（成对比较）试验，为了确保试验的有效性，将 α 设为 1%，否定假设是 H_0 ：A 的苦味与 B 的苦味相同；备择假设是 H_a ：B 的苦味 > A 的苦味；因此检验

是单边检验。两种啤酒分别被标有 452 和 603，试验由 40 人参加。问卷类似表 6-1，表 6-2。

试验的问题是：哪一个样品更苦？

样品筛选：试验之前由一小型品评小组进行品尝，以确保除了苦味之外，两种样品之间其他的差异非常小。

分析结果：有 26 人选择样品 B，从表 6-5 中可知， $\alpha=1\%$ 对应的临界值时 28，因此两种样品之间不存在显著差异。

注意事项：在确定成对比较试验是单边的还是双边时，关键的一点是看备择假设是单边还是双边的。当试验目的为了确定某项改进措施或处理方法的效果时，通常使用单边检验。表 6-8 是一些单边和双边的常见例子。

表 6-8 单边和双边的常见例子

单边	双边
确认试验啤酒比较苦	确定哪一个啤酒更苦
确认试验产品更受欢迎	确定哪一个产品更受欢迎
$A > B$ 或 $B > A$	备择假设为样品 $A \neq$ 样品 B ，而不是样品 $A > B$

例 3 差别成对比较检验-----甜橙风味试验

问题：某饮品公司一直使用一种含有转基因成分的甜橙香味物质，但欧洲市场最近规定，转基因成分需要在食品成分表中标出。为了防止消费者的抵触情绪，某公司决定使用一种不含转基因成分的甜橙香味物质，但初步试验表明，不含转基因成分的物质可能甜橙香味没有原来浓，现在研究人员想知道这两种香味物质的甜橙香气是否有所差别。

项目目标：研究开发一种具有甜橙香气特征的产品。

试验目标：测量两种风味物质赋予产品甜橙香味特征的相对能力，即两种甜橙风味是否不同。

试验设计：因为不同的人对甜橙风味会有不同的看法，因此需要参加试验的人数要多一些，并且不一定需要培训。试验由 45 人参加，将 α 设为 5%。否定假设是 H_0 ；样品 A 的甜橙风味=样品 B 的甜橙风味，备择假设是 H_a ：样品 A 的甜橙风味 \neq 样品 B 的甜橙风味，因为只关心是否有所不同，所以这个检验是双边的。样品分别被标为 793（原产品）和 743（新产品），问卷如表 6-9 所示。

表 6-9 试验调查问卷

方向性差别试验	
姓名： _____	日期： _____
样品类型： _____	
研究特性： _____	
<p>试验说明：</p> <p>从左到右品尝每对样品，然后作出你的判断。</p> <p>如果是没有明显的差异，可以猜一个答案，如果猜不出来，也可以做“无差异”的判断。</p>	
<p>试验组样品</p> <p>793 734</p>	<p>哪一个更具有甜橙风味</p> <p>_____</p>
<p>建议： _____</p> <p>_____</p>	

样品筛选： 试验之前对两种样品进行品尝，以确定它们的风味确实相似。

分析结果： 有 32 人认为样品 793 的甜橙风味更强。当 $\alpha = 0.05$ 的临界值时 30，因此认为两种样品之间存在显著差异。

解释结果： 为了保持原有市场，建议慎重使用该不含转基因成分的风味物质，因此从试验可以看出它的甜橙的风味不如原产品的浓，应继续试验，寻找合适的替代品。

第二节 二-三点检验法 (duo-trial test)

二-三点检验法由 Peryam 和 Swartz 于 1950 年提出的方法。先提供给评价员一个对照样品，接着提供两个样品，其中一个与对照样品相同或者相似。要求评价员在熟悉对照样品后，从后者提供的两个样品中挑选出与对照样品相同的样品，这种方法，也被称为一-二点检验法。二-三点试验的目的是区别两个同类样品是否存在感官差异，但差异的方向不能被检验指明。即感官评价员只能知道样品可觉察到差别，而不知道样品在何种性质上存在差别。

二-三点检验法有两种形式：一种叫做固定参照模式；另一种叫做平衡参照模式。在固定参照模式中，总是以正常生产为参照样；而在平衡参照模型中，正常生产的样品和要进行检验的样品被随机用做参照样品。如果参评人员是受过培训的，他们对参照样品很熟悉的情况下，使用固定参照模式；当参评人员对两种样品都不熟悉，而他们没有接受过培训时，使用平衡参照模式。在平衡参照模式中，一般来说，参加评定的人员可以没有专家，但要求人数较多，其中选定评定员通常 20 人，临时参与的可以多达 30 人，即 50 人之多。

一、 方法特点

1. 此方法是常用的三点检验法的一种替代法。在样品相对地具有浓厚的味道，强烈的气味或者其他冲动效应时，会使人的敏感性受到抑制，这时才使用这种方法。
2. 该方法比较简单，容易理解。但从统计学上来讲不如三点检验法具有说服力，精度较差（猜对率为 1/2），故此方法常用于风味较强、刺激较烈和产生余味持久的产品检验，以降低鉴评次数，避免味觉和嗅觉疲劳。另外，外观有明显差别的样品不适宜此法。
3. 二-三点检验法也具有强制性。该试验中已经确定知道两个样品是不同的，这样，当两样品区别不大时，不必像三点检验法去猜测。然而，差异不大的情况依然是存在的。当区别的确不大时，评价员必须去猜测，哪一个是特别一些的，这样，他的正确答复的机会是一半。为了提高全组的准确性，二-三点检验法要求有 25 组样品。如果这项检验非常重要，样品组数应当增加，在正常情况下，其组数一般不超过 50 个。
4. 这种方法在做品尝时，要特别强调漱口。在样品的风味很强烈的情况下，在做第二试验之前，都必须彻底地洗漱口腔，不得有残留物和残留味的存在。做完自己的样品后，如果后面还有一批同类的样品检验，最好是稍微离开现场一定时间，或回到品尝室饮用一些白开水等净水。
5. 固定参照二-三点检验中，样品有两种可能的呈送顺序，如 $R_A B A$ 、 $R_A A B$ ，应在所有的评价员中交叉平衡。而在平衡参照二-三点检验中，样品有四种可能的呈送顺序，如 $R_A B A$ 、 $R_A A B$ 、 $R_B A B$ 、 $R_B B A$ ，一般的评价员得到一种样品类型作为参照，而另一半的评价员得到另一种样品类型作为参照。样品在所有的评价员中交叉平衡。当评价员对两种样品都不熟悉，或者没有足够的数量时，可运用平衡参照二-三点试验。

二、 问答表设计与做法

二-三点检验虽然有两种形式，从评价员角度来讲，这两种检验的形式是一致的，只是

所使用的作为参照物的样品是不同的。二-三检验问答卷的一般形式如表 6-10 所示。

表 6-10 二-三检验问答卷的一般形式

二-三点检验		
姓名：_____	日期：_____	
试验指令：在你面前有 3 个样品，其中一个标明“参照”，另外两个标有编号。从左向右依次品尝 3 个样品，先是参照样，然后是两个样品。品尝之后，请在与参照相同的那个样品的编号上划圈。你可以多次品尝，但必须有答案。		
参照	321	586

三、结果分析与判断

有效鉴评表数为 n ，回答正确的表数为 R ，查表 6-5 中为 n 的一行的数值，若 R 小于其中所有数，则说明在 5% 水平，两样品间无显著差异，若 R 大于或等于其中某数，说明在此数所对应的显著水平上，两样品间有显著差异。

例如，某饮料厂为降低饮料成品的异味，在加工中添加了某种除味剂，为了了解除味剂的效果，运用二-三点检验法进行试验，由 41 名评价员进行检查，其中有 20 名接受到的对照样品是未经去味的制品，令 21 名接受到的对照样品是经去味处理的制品，共得到 40 张有效答案，其中有 28 张回答正确，查表 6-5 中 $n=41$ 一栏，知 $27(5\%) < 28 < 29(1\%)$ ，则在 5% 显著水平，两样品间有显著差异，即去除异味效果显著。

例 4 平衡参照二-三点试验

问题：一个产品香味开发人员要知道两种赋予面巾纸香味的方法（直接加到面巾上面和加到面巾纸盒里）是否会使产品香气的浓度和香气品质有所不同。

项目目标：确定两种加香方法是否会使面巾纸在正常存放时间之后有所不同。

试验目标：确定两种产品在存放 3 个月之后是否在香气上存在不同。

实验设计：样品在同一天准备，使用完全相同的香味物质和相同的面巾纸，只是赋予香味的方法不同，将两种样品放在相同的条件下存放 3 个月。有 50 人参加实验，样品编号及排组

情况参照三点试验，两种样品各自被用做参照样 25 次。准备工作表及试验答卷见表 6-12、表 6-13 所示。

结果分析：在进行试验的 50 人中，有 33 人做出了正确的选择。根据表 6-5，在 5% 显著水平下，临界值是 32，所以说两种产品的香味之间有差别。

表 6-12 面巾纸二-三点检验准备工作表

样品准备工作表		
日期：_____		
编号：_____		
样品类型：面巾纸		
实验类型：二-三点检验（平衡参照模型）		
产品情况	含有 2 个 A 的号码	含有 2 个 B 的号码
A：新产品	959 257	448
B：原产品（对比）	723	539 661
呈送容器标记情况	号码顺序	代表类型
小组		
1	AAB	R-257-723
2	BBA	R -661-448
3	ABA	R- 723-257
4	BAB	R -448-661
5	BAA	R -723-257
6	ABB	R -661-448
7	AAB	R -959-723
8	BBA	R -539-448
9	ABA	R -723-959
10	BAB	R -448-539
11	BAA	R -723-959
12	ABB	R -448-661
R 为参照，将以上顺序依次重复，直到 50 组。		

表 6-13 面巾纸二-三点检验问答表

二- 三点检验		
品评员编号： _____	日期： _____	
样品：面巾纸		
试验指令：		
<ul style="list-style-type: none"> b) 请将杯子盖子拿掉，从左到右依次闻您面前的样品。 c) 最左边的是参照样。确定哪一个带有编号的样品的香味同参照样相同。 d) 在您认为相同的编号上划圈。 		
如果您认为带有编号的两个样品非常相近，没有什么区别，您也必须在其中选择一个。		
参照	539	448

解释结果：感官分析人员可以告知那位香味研究人员，通过二-三点检验方法，在给定的香气成分、纸张和存放期下，这两种产品在香味上有差别。要知道具体哪一种香气存在更长的时间，还需要做描述性分析。

例 5 固定参照二-三点试验

问题：一个奶粉生产商现在有两个奶粉包装的供应商，A 是他们已经使用多年的产品，B 是一种新产品，可以延长货架期。他想知道这两种包装对奶粉风味的影响是否不同。而且这个奶粉生产商觉得有必要在奶粉风味稍有改变和奶粉货架期的延长上做一些平衡，也就是说，他愿意为延长货架期而冒奶粉风味可能发生变化的风险。

项目目标：确定奶粉包装的改变是否会在储存一段时间后使得奶粉的风味有所改变。

实验目标：两种奶粉包装的奶粉在室温存放 10 周之后冲泡，在风味上是否有所差异。

评定人数的确定：将 α 值定为 5%，评定人员是 90 人。

试验设计：对于这个实验来说，固定模型的二-三点检验更合适一些，因为品评人员对该公司的产品，用 A 种奶粉包装的奶粉，非常熟悉。为了节省时间，试验可以分为三组，每组 30 人，同时进行。以 A 为参照，每组都要熟悉 $30 \times 2 = 60$ 个 A 和 30 个 B。

结果分析：在 3 组中，分别有 17，18，19 个人做出了正确选择。根据附表 6-5，当参评人数是 30， α 值为 5% 时，临界值是 20。然而从整个大组来看做出了正确选择的人数是 54，从附表得出的临界值是 54。这两个结果有些出入。但要知道，30 并不是该实验要求的参评

人数，查看结果还要依据真正的参评人数，90人。

解释结果：如果将3个小组合并起来考虑，在 α 值为5%的水平下，A和B是存在差异的。下面需要确定哪一种产品更好，可以检查评价者是否写下了关于两种产品之间不同的评语。如果没有。将样品送给描述分析小组。如果经过描述检验后，仍不能确定哪一个产品好与另外一个产品，可以进行消费者试验。再最终确定哪一种包装的奶粉更被接受。

第三节 三点检验法 (triangle test)

三点检验是差别检验当中最常用的一种方法，是由美国的(本格逊) Bengtson 及其同事首先提出的。在检验中，同时提供三个编码样品，其中有两个是相同的，另外一个样品与其他两个样品不同，要求品评员挑选出其中不同于其他两个样品的检验方法，也称为三角试验法。三点检验法可使感官专业人员确定两个样品间是否有可觉察的差别，但不能表明差别的方向。三点检验法常被应用在以下几个方面：1) 确定产品的差异是否来自成分、工艺、包装和储存期的改变；2) 确定两种产品之间是否存在整体差异；3) 筛选和培训检验人员，以锻炼其发现产品差别的能力。

一、 方法特点

1. 在感官评定中，三点检验法是一种专门的方法，用于两种产品的样品间的差异分析，而且适合于样品间细微差别的鉴定，如品质管制和仿制产品。其差别可能与样品的所有特征，或者与样品的某一特征有关。
2. 三点检验试验中，每次随机呈送给评价员3个样品，其中2个样品是一样的，一个样品则不同。并要求在所有的评价员间交叉平衡。为了使三个样品的排列次序和出现次数的几率相等，这两种样品可能的组合是：BAA、ABA、AAB、ABB、BAB 和 BBA。在实验中，组合在六组中出现的几率也应是相等的，当评价员人数不足六的倍数时，可舍去多余样品组，或向每个评价员提供六组样品做重复检验。
3. 对三点检验的无差异假设规定：当样品间没有可觉察的差别时，做出正确选择的概率时1/3。因此，在试验中此法的猜对率为1/3，这要比成对比较法和二-三点法的1/2猜对率准确度低得多。
4. 在食品三点检试验中，所有评价员都应基本上具有同等的鉴别能力和水平，并且因

食品的种类不同, 评定员也应该是各具专业之所长。参与评价的人数多少要因任务而异, 可以在 5 人到上百人的很大范围内变动, 并要求做差异显著性测定。三点检验通常要求品评人员在 20~40 之间, 而如果试验目的是检验两种产品是否相似时 (是否可以相互替换), 要求的参评人员人数则为 50~100。

5. 食品三点检验法要求的技术比较严格, 每项检验的主持人都要亲自参与评定。为使检验取得理想的效果, 主持人最好组织一次预备试验, 以便熟悉可能出现的问题, 以及先了解一下原料的情况。但要防止预备试验对后续的正常检验起诱导作用。
6. 在三点检验中, 评价组的主持人只允许其小组出现以下两种结果。第一种, 根据“强迫选择”的特殊要求, 必须让评价员指明样品之一与另 2 个样品不同。第二种, 根据实际, 对于的确没有差别的样品, 允许打上“无差别”字样。这两点在显著性测定表上查找差异水平时, 都是要考虑到的。
7. 评价员进行检验时, 每次都必须从左到右的顺序品尝样品。评价过程中, 允许评价员重新检验已经做过的那个样品。评价员找出与其他两个样品不同的一个样品或者相似的样品, 然后对结果进行统计分析。
8. 三点检验法是比较复杂的, 即使是有经验的评价员也会感到不是很容易。如当其中某一对被认为是相同的时候, 也还得用另一样品的特征去证明。这样反复的互证, 是看起来尚可而做起来难的事情。为了判断正确, 不能让评价员知道其排列的顺序。所以样品的排序者不能参加评价。
9. 当参加评价的工作人员数目不是很多的时, 可选择此法。

二、 问答表设计与做法

在问答表的设计中, 通常要求评价员指出不同的样品或者相似的样品。对于评价员必须告知该批检验的目的, 提示要简单明了, 不能有暗示。常用的三角检验法问答表如 6-14 所示。

表 6-14 三点检验问答表的一般形式

三点检验		
姓名: _____	日期: _____	
试验指令:		
在你面前有 3 个带有编号的样品, 其中有两个是一样, 而另一个和其他两个不同。请从左到右依次品尝 3 个样品, 然后在与其他两个样品不同的那一个样品的编号上划圈。你可以多次品尝, 但不能没有答案。		
624	801	129

三、 结果分析与判断

(1) 按三点试验法要求统计回答正确的问答表数, 查表可得出两个样品间有无差异。

例如, 36 张有效鉴评表, 由 21 张正确地选择出单个样品, 查表 6-15 中 n=36 栏。由于 21 大于 1% 显著水平的临界值 20, 小于 0.1% 显著水平的临界值 22, 则说明在 1% 显著水平, 两样品间有差异。

当有效鉴评表大于 100 时 (n>100 时), 表明在差异的鉴评最少数为 $0.4714z\sqrt{n} + \frac{(2n+3)}{6}$ 的近似整数; 若回答正确的鉴评表数大于或等于这个最少数, 则说明两样品间有差异。

式中 z 值为:

显著水平	5%	1%	1.0%
Z 值	1.64	2.33	3.10

表 6-15 三点检验法检验表

答案 数目 (n)	显著水平			答案 数目 (n)	显著水平			答案 数目 (n)	显著水平		
	5%	1%	0.1%		5%	1%	0.1%		5%	1%	0.1%
4	4	-	-	33	17	18	21	62	28	31	33
5	4	5	-	34	17	19	21	63	29	31	34
6	5	6	-	35	17	19	22	64	29	32	34
7	5	6	7	36	18	20	22	65	30	32	35
8	6	7	8	37	18	20	22	66	30	32	35
9	6	7	8	38	19	21	23	67	30	33	36
10	7	8	9	39	19	21	23	68	31	33	36

11	7	8	10	40	19	21	24	69	31	34	36
12	8	9	10	41	20	22	24	70	32	34	37
13	8	9	10	42	20	22	25	71	32	34	37
14	9	10	11	43	21	23	25	72	32	35	38
15	9	10	12	44	21	23	25	73	33	35	38
16	9	11	12	45	22	24	26	74	33	36	39
17	10	11	13	46	22	24	26	75	34	36	39
18	11	12	13	47	23	24	27	76	34	36	39
19	11	12	14	48	23	25	27	77	34	37	40
20	11	13	14	49	23	25	28	78	35	37	40
21	12	13	15	50	24	26	28	79	35	38	41
22	12	14	16	51	24	26	29	80	35	38	41
23	13	15	16	52	24	27	29	82	36	39	42
24	14	16	18	53	25	27	29	84	37	40	43
25	15	16	18	54	25	27	30	86	38	40	44
26	15	17	19	55	26	28	30	88	38	41	44
27	15	17	19	56	26	28	31	90	39	42	45
28	16	18	20	57	26	29	31	92	40	43	46
29	15	16	18	58	27	29	32	94	41	44	47
30	15	17	19	59	27	29	32	96	42	44	48
31	15	17	19	60	28	30	33	98	42	45	49
32	16	18	20	61	28	30	33	100	43	46	49

例 5 茶叶试验

问题：现有 2 种茶叶，一种是原产品，一种使用一批新种植的品种，感官检验人员想知道这两种产品之间是否存在差异。

项目目标：两种产品之间是否存在差异。

试验目标：检验两种产品之间的总体差异性。

试验设计：因为试验目的是检验两种产品之间的差异，我们将 α 值设为 5%，有 12 个品评人员参加检验，因为每个所需的样品是 3 个，所以一共准备了 36 个样品，新产品和原产品各 18 个，按表 6-16 安排实验。试验中使用随机号码。

表 6-16 茶叶差异试验准备工作表

样品准备工作表	
日期：_____	
编号：_____	

样品类型： _____				
试验类型： _____				
产品情况	含有 2 个 A 的号码使用情况		含有 2 个 B 的号码使用情况	
A: 新产品	533	681	576	
B: 原产品 (对比)	298	885	372	
呈送容器标记情况	号码顺序		代表类型	
小组				
1	533	681	298	AAB
2	576	886	372	ABB
3	885	372	576	BBA
4	298	681	533	BAA
5	533	298	681	ABA
6	885	576	372	BAB
7	533	681	298	AAB
8	576	885	372	ABB
9	885	372	576	BBA
10	298	681	533	BAA
11	533	298	681	ABA
12	885	576	372	BAB
样品准备程序:				
1. 两种产品各准备 18 个, 分 2 组 (A 和 B) 放置, 不要混淆。				
2. 按照上表的编号, 每个号码各准备 6 个, 将两种产品分别标好。即新产品 (A) 中标有 533、681 和 298 号码的样品个数分别为 6 个, 原产品 (B) 中标有 576、885 和 372 的样品个数也分别是 6 个。				
3. 将标记好的样品按照上表进行组合, 每份相应的小组号码和样品号码也写在文答卷上, 呈送给品评人员。				

试验结果: 将 12 份答好的问答卷回收, 按照上表核对答案, 统计答对的人数。经核对, 在该试验中, 共有 9 人做出了正确选择。根据表 6-15 在 $\alpha = 5\%$, $n=12$ 时, 对应的临界值是 8, 所以这两种产品之间存在差异的。

结论：这两种茶叶（新产品和原产品）是存在差异的。

第四节 “A”—非“A”检验法（“A”—非“A”test）

在感官评定人员先熟悉样品“A”以后，再将一系列样品呈送给这些检验人员，样品中有“A”，也有“非A”。要求参评人员对每个样品做出判断，那些是“A”，那些是非“A”。这种检验方法被成为“A”—“非A”检验法。这种是与否的检验法，也称为单项刺激检验。此试验适用于确定原料、加工、处理、包装和贮藏等各环节的不同所造成的两种产品之间存在的细微的感官差别，特别适用于检验具有不同外观或后味样品的差异检验，也适用于确定鉴评员对产品某一种特性的灵敏性。

一、 方法特点

1. 此检验本质上是一种顺序成对差别检验或简单差别检验。评价员先评价第一个样品，然后再评价第二个样品，要求评价员指明这些样品感觉上是相同还是不同。此试验的结果只能表明评价员可觉察到样品的差异，但无法知道样品品质差异的方向。
2. 此试验中，样品有4种可能的呈送顺序，如AA、BB、AB、BA。这些顺序要能够在评价员之间交叉随机化。在呈送给评价员的样品中，分发给每个品评员的样品数应相同，但样品“A”的数目与样品“非A”的数目不必相同。每次试验中，每个样品要被呈送20~50次。每个品评者可以只接受一个样品，也可以接受2个样品，一个“A”，一个“非A”，还可以连续品评10个样品。每次评定的样品数量视检验人员的生理疲劳程度而定，受检验的样品数量不能太多，应以评品人数较多来达到可靠的目的。
3. 评价员必须经过训练，使之能够理解评分表所描述的任务，但他们不需要接受特定感官方面的评价训练。通常需要10~50名品评人员参加试验，他们要经过一定的训练，做到对样品“A”和“非A”比较熟悉。
4. 需要强调的一点是，参加检验评定的人员一定要对样品“A”和“非A”非常熟悉，否则，没有标准或参照，结果将失去意义。
5. 检验中，每次样品出示的时间间隔很重要，一般是相隔2到5分钟。

二 问答表设计与做法

A-非A检验法问答表的一般形式如表 6-17 所示。

表 6-17 A-非 A 检验法问答表的一般形式

A-非 A 检验			
姓名: _____		日期: _____	
样品: _____			
试验指令: 1. 在试验之前对样品“A”和“非 A”进行熟悉, 记住它们的口味 2. 从左到右依次品尝样品, 在品尝完每一个样品之后, 在其编码后面相对应位置上打√ 注意: 在你所得到的样品中, “A”和“非 A”的数量是相同的。			
样品顺序号	编号	该样品是	
		“A”	“非 A”
1	591		
2	304		
3	547		
4	743		
5	568		
6	198		

三 结果分析与判断

对鉴评表进行统计, 并汇入表 6-18 中, 并进行结果分析。表中 n_{11} 为样品本身是“A”, 鉴评员也认为是“A”的回答总数; n_{22} 为样品本身是“非 A”, 鉴评员也认为是“非 A”的回答总数; n_{21} 为样品本身是“A”, 而鉴评员认为是“非 A”的回答总数, n_{12} 为样品本身是“非 A”, 而鉴评员认为是“A”的回答总数。 $n_{1\cdot}$ 、 $n_{2\cdot}$ 为第 1、2 行回答数之和, $n_{\cdot 1}$ 、 $n_{\cdot 2}$ 为第 1、2 列回答数之和, n 为所有回答数, 然后用 χ^2 检验来进行解释。结果参看表 6-18。

表 6-18 结果统计表

判 别	判 别 数	样 品		累 计
		“A”	非“A”	
	判为“A”的回答数	n_{11}	n_{12}	$n_{1\cdot}$
	判为“非 A”的回答数	n_{21}	n_{22}	$n_{2\cdot}$

累计	$n \cdot 1$	$n \cdot 2$	n
----	-------------	-------------	-----

假设鉴评员的判断与样品本身的特性无关。

当回答总数为 $n \leq 40$ 或 n_{ij} ($i=1, 2; j=1, 2$) ≤ 5 时, χ^2 的统计量为:

$$\chi^2 = \frac{[n_{11} \times n_{22} - n_{12} \times n_{21} - n/2]^2 \times n}{n_{\cdot 1} \times n_{\cdot 2} \times n_{1 \cdot} \times n_{2 \cdot}}$$

当回答总数是 $n > 40$ 和 $n_{ij} > 5$ 时, χ^2 的统计量为:

$$\chi^2 = \frac{[n_{11} \times n_{22} - n_{12} \times n_{21}]^2 \times n}{n_{\cdot 1} \times n_{\cdot 2} \times n_{1 \cdot} \times n_{2 \cdot}}$$

将 χ^2 统计量与 χ^2 分布临界值(附表 1.)比较:

当 $\chi^2 \geq 3.84$, 为 5% 显著水平;

当 $\chi^2 \geq 6.63$, 为 1% 显著水平。

因此, 在此选择的显著水平上拒绝原假设, 即鉴评员的判断与样品特性相关, 即认为样品“ A ”与“非 A”有显著差异。

当 $\chi^2 < 3.84$, 为 5% 显著水平;

当 $\chi^2 < 6.63$, 为 1% 显著水平。

因此, 在此选择的显著水平上接受原假设, 即认为鉴评员的判断与样品本身特性无关, 即认为样品“ A ”与“非 A”无显著性差异。

例 6 冷藏与室温保藏

30 位鉴评员判定某种食品经过冷藏(“ A ”)和室温储藏(“非 A”)后, 两者的差异关系。每位鉴评员评价两个“ A ”和三个“非 A”, 则统计结果如表 6-19。

表 6-19 结果统计表

判 别	样 品 数	“ A ”	“非 A”	累计
--------	-------------	-------	-------	----

判别	“A”	40	40	80
评价数	“非 A”	20	50	70
累计		60	90	150

由于 $n=150>40$, $n_{ij}>5$, 则

$$\begin{aligned}
 \chi^2 &= \frac{[n_{11} \times n_{22} - n_{12} \times n_{21}]^2 \times n}{n_{\cdot 1} \times n_{\cdot 2} \times n_{1 \cdot} \times n_{2 \cdot}} \\
 &= \frac{(40 \times 50 - 20 \times 40)^2 \times 150}{60 \times 90 \times 80 \times 70} \\
 &= 7.14
 \end{aligned}$$

因为 $\chi^2 = 7.14 > 6.63$, 所以在 1% 显著水平上有显著差异。

例 7 一种新型鲜味剂与味精

问题: 一名产品开发人员正在研究用一种鲜味剂来替换某调味料中目前用量为 0.5% 的味精成分。前期试验表明, 0.1% 的该鲜味剂相当于 0.5% 的味精, 但是如果一次品尝的样品超过 1 个小时, 由于该鲜味剂鲜味的后味、其他味道和口感等因素, 就会让人感觉到某些异样。该开发人员知道, 含有这种新型鲜味剂和味精的调味料是否能够被识别出来。

项目目标: 确定 0.1% 的该鲜味剂能否代替 0.5% 的味精。

试验目标: 直接比较这两种鲜味物质, 并减少味道的延迟和覆盖效应。

试验设计: 分别将鲜味剂和味精配制成 0.1% 和 0.5% 的溶液, 将鲜味剂溶液设为 “A”, 将味精溶液设为 “非 A”。由 20 人参加品评, 每人得到 10 个样品, 每个样品品尝一次, 然后回答是 “A” 还是 “非 A”, 在品尝下一个样品之前用清水漱口, 并等待 1min, 问答卷见表 6-20。

表 6-20 试验调查问卷表

“A” - “非 A” 检验
姓名: _____ 日期: _____
样品: 鲜味调味品

试验指令：

1. 在试验之前，对样品“A”和“非 A”进行熟悉，记住它们的口味。
 2. 从左到右依次品尝样品，在品尝完每一个样品之后，在其编号后面的相应位置中打√
- 注意：在你所得到的样品中，“A”和“非 A”的数量是相同的。

样品顺序号	编号	该样品是	
		“A”	“非 A”
1	591	_____	_____
2	304	_____	_____
3	547	_____	_____
4	743	_____	_____
5	568	_____	_____
6	198	_____	_____
7	974	_____	_____
8	552	_____	_____
9	687	_____	_____
10	303	_____	_____

分析结果：得到的结果如表 6-21 所示

表 6-21 试验结果统计

回答情况	样品真实情况		
	“A”	“非 A”	总计
“A”	60	35	95
“非 A”	40	65	105
总计	100	100	200

$$x^2 = \frac{(60 \times 65 - 40 \times 35)^2 \times 200}{100 \times 100 \times 95 \times 105}$$

$$x^2 = \frac{125000}{9975} = 12.53$$

设 $\alpha = 0.05$ ，由 x^2 分布表。 $df = 1$ ，（一共有 2 种样品）得到 $x^2 = 3.84$ ， $12.53 > 3.84$ ，所以，0.1% 的鲜味剂和 0.5% 味精溶液之间存在显著差异。

结果的解释：通过试验，可以告诉该研究人员，0.1% 的甜味剂和 0.5% 的味精溶液是不同的，它能够被识别出来，如果想搞清楚如何不同，可以进一步做描述分析的感官试验。

例 8 两种新型鲜味剂与味精

问题：一名产品开发人员正在研究用两种甜味剂来替换某调味料中目前用量为 0.5% 的味精成分。前期试验表明，0.1% 的鲜味剂甲和 0.1% 的鲜味剂乙相当于 0.5% 的味精，但是如果一次品尝的样品超过 1 个小时，由于该两种鲜味剂鲜味的后味、其他味道和口感等因素，就会让人感觉到某些异样。该开发人员知道，含有这两种新型鲜味剂和味精的调味料是否能够被识别出来。

项目目标：确定 0.1% 鲜味剂甲和 0.1% 鲜味剂乙能否代替 0.5% 的味精。

试验目标：检验味精“A”、鲜味剂“(非 A)₁”、鲜味剂“(非 A)₂”三者之间在甜味上是否有显著性差异。

试验设计：分别将鲜味剂甲和乙与味精配制成 0.1%、0.1% 和 0.5% 的溶液，将味精溶液设为“A”，分别将鲜味剂甲溶液和鲜味剂乙溶液设为“(非 A)₁”和“(非 A)₂”。由 20 人参加品评，每人得到 10 个样品，每个样品品尝一次，然后回答是“A”还是“非 A”，在品尝下一个样品之前用清水漱口，并等待 1min，问答卷见表 6-22。

表 6-22 试验调查问卷表

A- (非 A) ₁ 、(非 A) ₂ 检验	
姓名：_____	日期：_____
样品：鲜味调味品	
试验指令：	
3. 在试验之前，对样品“A”和“(非 A) ₁ ”、“(非 A) ₂ ”进行熟悉，记住它们的口味。	
4. 从左到右依次品尝样品，在品尝完每一个样品之后，在其编号后面的相应位置中打√	

注意：在你所得到的样品中，“A”、“(非 A)₁”和“(非 A)₂”三者的数量是相同的。

样品顺序号	编号	该样品是		
		“A”	“(非 A) ₁ ”	“(非 A) ₂ ”
1	591	_____	_____	_____
2	304	_____	_____	_____
3	547	_____	_____	_____
4	743	_____	_____	_____
5	568	_____	_____	_____
6	198	_____	_____	_____
7	974	_____	_____	_____
8	552	_____	_____	_____
9	687	_____	_____	_____
10	303	_____	_____	_____

分析结果：得到的结果如表 6-23 所示

表 6-23 试验结果统计

回答情况	样品真实情况			
	“A”	“非 A”		总计
		“(非 A) ₁ ”	“(非 A) ₂ ”	
“A”	60	45	40	145
“非 A”	40	55	40	135
总计	100	100	80	280

本例中因为 $n_{..}$ 大于 40 和 n_{ij} 大于 5，所以用如下公式：

$$\chi^2 = \frac{[n_{11} \times n_{22} - n_{12} \times n_{21}]^2 \times n_{..}}{n_{.1} \times n_{.2} \times n_{1.} \times n_{2.}}$$

把试验结果统计表中的各值代入上式，得 $x^2=4.65$ ，4.65 小于 $df=2$ ； $\alpha=0.05$ 的对应临界值 5.99(见附表 1.)，因此得出结论：认为味精、鲜味剂“(非 A)₁”、鲜味剂“(非 A)₂”三者之间在鲜味上无显著差异。

第五节 五中取二检验法 (two out of five test)

同时提供给评价员五个以随机顺序排列的样品，其中两个是同一类型，另三个是另一种类型。要求评价员将这些样品按类型分成两组的一种检验方法称为五中取二检验法。该方法在测定上更为经济，统计学上更具有可靠性，但在评定过程中容易出现感官疲劳。

一、 方法特点

1. 此检验方法可识别出两样品间的细微感官差异。从统计学上讲，在这个试验中单纯猜中的概率是 1/10，而不是三点试验的 1/3，二-三检验的 1/2。统计上更具有可靠性。
2. 人数不是要求很多，通常只需 10 人左右或稍多一些。当评价员人数少于十个时，多用此方法。
3. 由于要从 5 个样品中挑出 2 个相同的产品，这个试验易受感官疲劳和记忆效果的影响，并且需用样品量较大。一般只用于视觉、听觉、触觉方面的试验，而不是用来进行味道的检验。
4. 在每次评定试验中，样品的呈送有一个排列顺序，其可能的组合有 20 个，如：AAABB、ABABA、BBBAA、BABAB、AABAB、BAABA、BBABA、ABBAB、ABAAB、ABBAA、BABBA、BAABB、BAAAB、BABAA、ABBBA、ABABB、AABBA、BBAAA、BBAAB、AABBB。

二、 回答表设计与做法

在五中取二检验法试验中，一般常用的问答表如表 6-24 所示。

表 6-24 五中取二检验问答表

五中取二检验

姓名：_____

日期：_____

试验指令：

- 按以下的顺序观察或感觉样品，其中有 2 个样品是同一种类型的，另外 3 个样品是另外一种类型。
- 测试之后，请你认为相同的两种样品的编码后面划√

编号	评语
862 _____	_____
568 _____	_____
689 _____	_____
368 _____	_____
542 _____	_____

三、 结果分析与判断

根据试验中正确作答的人数，查表得出五中取二试验正确回答人数的临界值，最后作比较。假设有效鉴评表数为 n ，回答正确的鉴评表数为 k ，查表 6-25 中 n 栏的数值。若 k 小于这一数值，则说明在 5% 显著水平两种样品间无差异。若 k 大于或等于这一数值，则说明在 5% 显著水平两种样品有显著差异。

表 6-25 五中取二检验法检验表 ($\alpha = 5\%$)

评价员数 (n)	正 答 最 少 数 (k)	评价员数 (n)	正 答 最 少 数 (k)	评价员数 (n)	正 答 最 少 数 (k)
9	4	23	6	37	8
10	4	24	6	38	8
11	4	25	6	39	8
12	4	26	6	40	8
13	4	27	6	41	8
14	4	28	7	42	9
15	5	29	7	43	9
16	5	30	7	44	9
17	5	31	7	45	9

18	5	32	7	46	9
19	5	33	7	47	9
20	5	34	7	48	9
21	6	35	8	49	10
22	6	36	8	50	10

例 9 原料稳定性差异检验

某食品厂为了检查原料质量的稳定性，把两批原料分别添加入某产品中，运用五中取二试验对添加不同批次的原料的两个产品进行检验。由十名评价员进行检验，其中有三名鉴别员正确地判断了五个样品的两种类型。查表 $n=10$ 一栏得到正答最少数为 4，大于 3，说明这两批原料的质量无差别。

例 10 纺织品粗糙程度的比较

问题：一纺织品供应商想用一种聚酯/尼龙混合品代替目前的聚酯织品。但是有人反映说该替代品手感粗糙、刮手。

项目目标：确定该聚酯/尼龙混合品是否真得很粗糙，需要改进。

试验目标：测定两种纺织品手感的差异。

试验设计：因为在该实验不涉及品尝，只是触觉，所以适合用五中取二检验法进行试验。一般来说，由 12 人组成的评定小组就足以发现产品之间的非常小的差别。从上面 20 个组合中，任意选取 12 个组合，将样品分别放在一张纸板后面，品评人员可以摸到样品，但不能看到，每个样品的纸板前标有该样品的随机编号（三位随机数字表），然后让评定者回答，哪两个样品相同，而不同于其他 3 个样品。问答卷如表 6-26 所示

表 6-26 纺织品粗糙程度的比较问答表

五中取二检验	
姓名：_____	日期：_____
试验指令：	
3. 按以下的顺序用手指或手掌感觉样品，其中有 2 个样品是同一种类型的，另外 3 个样品是另外一种类型。	

4. 测试之后，请你认为相同的两种样品的编码后面划√

编号

评语

862_____

568_____

689_____

368_____

542_____

结果分析：在 12 参评人员中，有 5 人做出了正确的选择。查表 6-25，回答正确人数的临界值是 4。说明产品之间有显著性差异。

解释结果：应该告知该生产商，这两种产品之间存在着显著的差异。

例 11 饼干生产中的油脂

问题：饼干生产中为了节省成本，要用一种氢化植物油替换现有配方中的另一种起酥油。替换之后，饼干表面的光泽有所降低。

项目目标：市场部想在产品进行消费者之前知道，用这两种配方制成的产品是否存在视觉上的差异。

试验目标：确定这两种饼干是否在外观上存在统计学上的差异。

试验设计：筛选 10 名参评人员，以确定他们在视力上和对颜色的识别上没有差异。将样品放在白瓷盘中，以白色作为背景，在白炽灯光下进行试验。

分析结果：在 10 名受试者当中，有 5 人正确选出了相同的两个样品，查表 6-25，正确回答的临界值为 4，说明这两种产品之间存在显著差异。

解释结果：可以告知有关人员，替换后的效果不够理想。

第六节 选择试验法

从三个以上的样品中，选择一个最喜欢或最不喜欢的样品的检验方法。它常用于嗜好

调查。

一、 方法特点

1. 试验简单易懂，不复杂，技术要求低。
2. 不适用于一些味道很浓或延缓时间较长的样品，这种方法在做品尝时，要特别强调漱口，在做第二试验之前，都必须彻底地洗漱口腔，不得有残留物和残留味的存在。
3. 对评定员没有硬性规定要求必须经过培训，一般在 5 人以上，多则 100 人以上。
4. 常用于嗜好调查，出示样品的顺序是随机的。方法比较简单。

二、 问答表设计与做法

常用的选择试验法调查问答如表 6-27 所示。

表 6-27 选择试验法调查问答表示例

选择试验法		
姓名： _____	日期： _____	
试验指令： 1. 从左到右依次品尝样品。 2. 品尝之后，请在你最喜欢的样品号码上画圈。		
256	387	583

三、 结果分析与判断

(1) 求数个样品间有无差异，根据检验判断结果，用如下公式求值

$$x_0^2 = \sum_{i=1}^m \frac{(x_i - \frac{n}{m})^2}{\frac{n}{m}}$$

式中 m——样品数

n——有效鉴评表数

x_i —— m 个样品中，最喜好其中某个样品的人数

查 χ^2 分布表 (见附表 1.)，若 $x_0^2 \geq \chi^2(f, \alpha)$ (f 为自由度， $f = m-1$ ， α 为显著水平)，

说明 m 个样品在 α 显著水平存在差异，若 $x_0^2 < \chi^2(f, \alpha)$ ，说明 m 个样品在 α 显著水平不存在差异。

(2) 求被多数人判断为最好的样品与其他样品间是否存在差异，根据 χ^2 检验判断结果，

用如下公式求 x_0^2 值

$$x_0^2 = (x_i - \frac{n}{m})^2 \frac{m^2}{(m-1)n}$$

查 χ^2 分布表 (见附表)，若 $x_0^2 < \chi^2(1, \alpha)$ ，说明此样品与其它样品之间在 α 水平存在差异。

否则，无差异。

例 12 产品比较

某生产厂家把自己生产的商品 A，与市场上销售的三个同类商品 X，Y，Z 进行比较。

由 80 位鉴赏员进行评价，并选出最好的一个产品来，结果如下表所示：

商品	A	X	Y	Z	合计
认为某商品最好的人员	26	32	16	6	80

求四个商品间的喜好度有无差异。

$$x_0^2 = \sum_{i=1}^m \frac{(x_i - \frac{n}{m})^2}{\frac{n}{m}} = \frac{n}{m} \sum_{i=1}^m (x_i - \frac{n}{m})^2$$
$$= \frac{4}{80} \times \left\{ \left(26 - \frac{80}{4} \right)^2 + \left(32 - \frac{80}{4} \right)^2 + \left(16 - \frac{80}{4} \right)^2 + \left(6 - \frac{80}{4} \right)^2 \right\}$$

$$= 19.6$$

$$f = 4 - 1 = 3$$

查 χ^2 分布表知：

$$\chi^2(3, 0.05) = 7.8 < x_0^2 = 19.6$$

$$\chi^2(3, 0.01) = 1.34 < x_0^2 = 19.6$$

所以，结论为四个商品间的喜好度在 1% 显著水平有显著性差异。

求多数人判断为最好的商品与其他商品间是否有差异。

$$x_0^2 = \left(x_i - \frac{n}{m}\right)^2 \frac{m^2}{(m-n)n} = \left(32 - \frac{80}{4}\right)^2 \times \frac{4^2}{(4-1) \times 80} = 9.6$$

查 x^2 分布表知：

$$x^2(1, 0.05) = 3.84 < x_0^2 = 9.6$$

$$x^2(1, 0.01) = 6.63 < x_0^2 = 9.6$$

所以，结论为被多数人判断为最好的商品 X 与其他商品间存在极显著差异，但与商品 A 相

比，由于 $\left(32 - \frac{58}{2}\right)^2 \times \frac{2^2}{(2-1) \times 58} = 0.62$ 远远小于 $x^2(1, 0.05)$ 。故可以认为无差异。

第七节 配偶试验法

把两组试样逐个取出各组的样品进行两两归类的方法叫做配偶试验法。

一、 方法特点

1. 此方法可应用于检验鉴别能力，也可用于识别样品间的差异。
2. 检验前，两族种样品的顺序必须是随机的，但样品的数目可不尽相同，如 A 组有 m 个样品，B 组中可有 m 个样品，也可由 m+1 或 m+2 个样品，但配对数只能是 m 对。

二、 回答表的设计与做法

配偶试验法问答表的一般形式如下表 6-28 所示。

表 6-28 配偶试验法问答表的一般形式

配偶试验法	
姓名： _____	日期： _____

试验指令：

1. 有两组样品，要求从左到右依次品尝。
2. 品尝之后，归类样品

A 组	B 组
256	658
583	456
596	369
154	489

归类结果：

_____和_____

_____和_____

_____和_____

_____和_____

三、 结果分析与判断

统计出正确的配对数平均值，即 \bar{S}_0 ，然后根据以下情况查表 6-29 或 表 6-30 中的相应值，得出有无差异的结论。

- (1) m 对样品重复配对时，(即由两个以上鉴评员进行配对时)，若 \bar{S}_0 大于或等于表 6-30 中的相应值，说明在 5%显著水平样品间有差异。

表 6-29 配偶试验检验表 ($\alpha=5\%$)

n	S	n	S	n	S	n	S
1	4.00	6	1.83	11	1.64	20	1.43
2	3.00	7	1.86	12	1.58	25	1.36
3	2.33	8	1.75	13	1.54	30	1.33
4	2.25	9	1.67	14	1.52		
5	1.90	10	1.60	15	1.50		

注：此表为 m 个和 m 个样品配对时的检验表。适用范围：m ≥ 4 重复次数 n。

表 6-30 配偶法检验表 (α = 5%)

m		
	m+1	m+2
3	3	3
4	3	3
5	3	3
6 以上	4	3

注：此表为 m 个和(m+1)或 (m+2) 个样品配对时的检验表

(2) m 个样品与 m 个或(m+1)或 (m+2) 个样品配对时，若 \bar{S}_0 大于或等于表 6-30 中 n=1 栏或表 6-31 中的相应值，说明在 5% 显著水平样品间有差异，或者说鉴别员在此显著水平有识别能力。

例 13 加工食品的配偶试验

由四名评价员通过外观，对八种不同加工方法的食物进行配偶试验。结果如下表 6-31 所示。

表 6-31 试验结果统计表

样品 \ 评价员	A	B	C	D	E	F	G	H
1	B	C	E	D	A	F	G	B
2	A	B	C	E	D	F	G	H
3	A	B	F	C	E	D	H	C
4	B	F	C	D	E	C	A	H

四个人的平均正确配偶数 $\bar{S}_0 = \frac{3+6+3+4}{4} = 4$ ，查表 6-29 中 n=4 栏， $S=2.25 < \bar{S}_0=4$ ，说明这八个产品在 5% 显著水平有显著差异，或这四个评价员有识别能力。

例 14 评价员训练试验

向某个评价员提供砂糖、食盐、酒石酸、硫酸奎宁、谷氨酸钠五种味道的稀释溶液分别为 400, 130, 50, 6.4, 50mg/100mL 和两杯蒸馏水, 共七杯试样。结果如下: 甜——食盐、咸——砂糖、酸——酒石酸、苦——硫酸奎宁、鲜——蒸馏水, 即该评价员判断出两种味道的试样, 即 $\bar{S}_0=2$ 而查表 6-30 中 $m=5$, $(m+2)$ 栏的临界值为 $3 > \bar{S}_0=2$ 说明该鉴别员在 5% 显著水平无判断味道的能力。