

中国标准出版社第一编辑室 编

# 中国食品工业 标准汇编

## 感官分析方法卷



中国标准出版社



# 中国食品工业标准汇编

## 感官分析方法卷

中国标准出版社第一编辑室 编

中国标准出版社

2007

# 中国食品工业标准汇编

## 感官分析方法卷

中国标准出版社第一编辑室编

### 图书在版编目 (CIP) 数据

中国食品工业标准汇编·感官分析方法卷/中国标准出版社第一编辑室编. —北京: 中国标准出版社, 2007  
ISBN 978-7-5066-4542-3

I. 中… II. 中… III. ①食品工业-标准-汇编-中国  
②食品感官评价-标准-汇编-中国 IV. TS207.2

中国版本图书馆 CIP 数据核字 (2007) 第 093272 号

中国标准出版社出版发行  
北京复兴门外三里河北街 16 号

邮政编码: 100045

网址 [www.spc.net.cn](http://www.spc.net.cn)

电话: 68523946 68517548

中国标准出版社秦皇岛印刷厂印刷

各地新华书店经销

\*

开本 880×1230 1/16 印张 18.25 字数 546 千字

2007 年 7 月第一版 2007 年 7 月第一次印刷

\*

定价 78.00 元

如有印装差错 由本社发行中心调换

版权专有 侵权必究

举报电话: (010)68533533

## 编 者 的 话

《中国食品工业标准汇编》是我国食品标准化方面的一套大型丛书,按行业分类分别立卷,由中国标准出版社陆续出版。本汇编为丛书的一卷。

感官分析技术是近代新兴的一门学科和技术,感官分析标准是重要的食品基础性标准。感官分析标准及技术应用于食品新产品开发和产品改进,对于加快我国食品产品进入国际化循环具有重大的意义。为解决生产、科研、检验、监督等部门缺少标准和标准收集不全的实际困难,为指导企业生产,提高产品质量,我们将2007年6月底以前发布的感官分析有关标准汇编成册,以满足全国食品企业、各级食品产品质量监督检验机构、各级食品卫生监督检验机构等部门的需要。

本汇编主要内容包括两个部分:基础标准和分析方法标准、相关标准。本汇编共收集与感官分析相关的国家标准20项。本汇编每个部分的标准按标准号由小到大编排。

本汇编收集的标准的属性已在目录上标明(GB或GB/T),年号用四位数字表示。鉴于部分国家标准是在标准清理整顿前出版的,现尚未修订,故正文部分仍保留原样,读者在使用这些标准时,其属性以目录上标明的为准(标准正文“引用标准”中标准的属性请读者注意查对)。

鉴于本汇编所收集的标准发布年代不尽相同,所用计量单位、符号在本汇编出版时未作改动。

本汇编可供食品生产、科研、销售单位的技术人员,各级食品监督、检验机构的人员、各管理部门的相关人员使用,也可供大专院校有关专业的师生参考。

本汇编由中国标准出版社第一编辑室选编。

编 者

2007年6月

# 目 录

## 一、基础标准和分析方法标准

GB/T 10220—1988	感官分析方法总论	3
GB/T 10221—1998	感官分析 术语	17
GB/T 12310—1990	感官分析方法 成对比较检验	37
GB/T 12311—1990	感官分析方法 三点检验	44
GB/T 12312—1990	感官分析 味觉敏感度的测定	50
GB/T 12313—1990	感官分析方法 风味剖面检验	56
GB/T 12314—1990	感官分析方法 不能直接感官分析的样品制备准则	64
GB/T 12315—1990	感官分析方法 排序法	67
GB/T 12316—1990	感官分析方法 “A”-“非 A”检验	83
GB/T 13868—1992	感官分析——建立感官分析实验室的一般导则	93
GB/T 14195—1993	感官分析 选拔与培训感官分析优选评价员导则	102
GB/T 15549—1995	感官分析 方法学 检测和识别气味方面评价员的入门和培训	129
GB/T 16291—1996	感官分析 专家的选拔、培训和管理导则	144
GB/T 16860—1997	感官分析方法 质地剖面检验	155
GB/T 16861—1997	感官分析 通过多元分析方法鉴定和选择用于建立感官剖面的描述词	170
GB/T 17321—1998	感官分析方法 二、三点检验	193
GB/T 19547—2004	感官分析 方法学 量值估计法	201

## 二、相关标准

GB/T 3358.1—1993	统计学术语 第一部分 一般统计术语	225
GB/T 3358.2—1993	统计学术语 第二部分 统计质量控制术语	255
GB/T 3358.3—1993	统计学术语 第三部分 试验设计术语	276

注：本汇编收集的标准的属性已在目录上标明(GB或GB/T)，年号用四位数字表示。鉴于部分国家标准是在标准清理整顿前出版的，现尚未修订，故正文部分仍保留原样，读者在使用这些标准时，其属性以目录上标明的为准(标准正文“引用标准”中标准的属性请读者注意查对)。

# 一、基础标准和分析方法标准



# 中华人民共和国国家标准

UDC 620.1  
:612.8  
:001.4  
GB 10220—88

## 感官分析方法总论

Sensory analysis—Methodology—  
General guidance

本标准参照采用国际标准 ISO 6658—1985《感官分析——方法学——总论》。

### 1 主题内容与适用范围

本标准给出了感官分析方法学的一般性导则。它将指导使用者选择适宜的感官分析方法。

本标准规定了感官分析的一般要求,介绍了各种检验方法以及对检验结果进行分析的统计学技术。

本标准适用于在实验室内对样品进行感官分析,特别适用于食品的感官分析。

本标准一般不适用于消费者偏爱检验。对也适用于偏爱检验的方法在相应的条款中指出。

### 2 引用标准

GB 10221.1~10221.4 感官分析术语

GB 3358 统计学名词及符号

### 3 术语

本标准所用感官分析术语见 GB 10221.1~10221.4。有关统计学术语见附录 A 及 GB 3358。

### 4 一般要求

#### 4.1 评价员

##### 4.1.1 评价员的种类

实验室内感官分析的评价员与消费者偏爱检验的评价员是两类不同的评价员。前者需要专门的选择与培训,后者只要求评价员的代表性。本标准中所涉及的评价员主要指前者。

实验室内感官分析的评价员有初级评价员、优选评价员、专家三种。

##### 4.1.2 评价员的数量

所需要的评价员的数量与所要求的结果的精度、检验的方法、评价员水平等因素有关。一般来讲,要求的精度越高,方法的功效越低,评价员水平越低,需要的评价员的数量越多。本标准中第5章对每种检验方法都给出了所需要的各类评价员的最少数量。考虑到实际中评价员可能缺席的情况,因此评价员数量应超过所要求的评价员的数目。一般多出50%。

##### 4.1.3 评价员的基本条件和要求

检验的目的不同对评价员的要求也不完全相同。基本条件是:

- a. 身体健康,不能有任何感觉方面的缺陷。
- b. 各评价员之间及评价员本人要有一致的和正常的敏感性。
- c. 具有从事感官分析的兴趣。
- d. 个人卫生条件较好,无明显个人气味。

国家技术监督局1988-12-22批准

1989-07-01实施



e. 具有所检验产品的专业知识并对所检验的产品无偏见。

为保证评价质量,要求评价员在感官分析期间具有正常的生理状态。为此对评价员有相应的要求,例如要求评价员不能饥饿或过饱,在检验前1 h内不抽烟、不吃东西,但可以喝水。评价员不能使用有气味的化妆品,身体不适时不能参加检验。

#### 4.1.4 评价员的选择与培训

应该使用在实际检验过程中可能遇到的各种样品以及使用第5章介绍的各类检验方法选择和培训评价员。用于选择和培训评价员的检验方法应与评价员将要使用的检验方法相一致。应该让评价员使用同一方法进行多次检验,根据正确回答的比例判断其水平。对评价员应定期考核。

敏感性检验常被用于选择与培训评价员。这些方法大致可分为:

- a. 阈检验:用于确定评价员的不同的阈值,例如,刺激阈、识别阈、差别阈和最大阈。
- b. 稀释检验:用于确定可感觉到的混入食品中的其他物质的最低量。

### 4.2 检验的物理条件

#### 4.2.1 环境

感官分析应在专门的检验室内进行。应给评价员创造一个安静的不受干扰的环境。检验室应与样品制备室分开。室内应保持舒适的温度与通风。避免无关的气味污染检验环境。检验室空间不宜太小,以免评价员有压抑的感觉。座位应舒适。

应限制音响,特别是应尽量避免能使评价员分心的谈话和其他干扰,应控制光的色彩和强度。

#### 4.2.2 器具与用水

与样品接触的容器应适合所盛样品。容器表面无吸收性并对检验结果无影响。应尽量使用已规定的标准化的容器。

应保证供水质量。为某些特殊目的,可使用蒸馏水,矿泉水,过滤水、凉开水等。

### 4.3 被检样品

#### 4.3.1 抽样

应按有关抽样标准抽样。在无抽样标准情况下有关方面应协商一致,要使被抽检的样品具有代表性,以保证抽样结果的合理性。

#### 4.3.2 样品的制备

样品的制备方法应根据样品本身的情况以及所关心的问题而定。例如对于正常情况是热吃的食品就应按通常方法制备并趁热检验。片状产品检验时不应将其均匀化。应尽可能使分给每个评价员的同种产品具有一致性。

有时评价那些不适于直接品尝的产品,检验时应使用某种载体。

对风味作差别检验时应掩蔽其他特性,以避免可能存在的交互作用。

对同种样品的制备方法应一致。例如,相同的温度,相同的煮沸时间,相同的加水量,相同的烹调方法等。

样品制备过程应保持食品的风味。不受外来气味和味道的影响。

#### 4.3.3 样品的分发

样品应编码,例如用随机的三位数字编码,并随机地分发给评价员,避免因样品分发次序的不同影响评价员的判断。

为防止产生感官疲劳和适应性,一次评价样品的数目不宜过多。具体数目将取决于检验的性质及样品的类型。评价样品时要有一定时间间隔,应根据具体情况选择适宜的检验时间。一般选择上午或下午的中间时间,因为这时评价员敏感性较高。

## 5 检验方法

### 5.1 方法的选择

在选择适宜的检验方法之前,首先要明确检验的目的。一般有两类不同的目的,一类主要是描述产品,另一类主要是区分两种或多种产品。第二类目的包括:确定差别;确定差别的大小;确定差别的方向;确定差别的影响。

当检验目的确定后,为了选择适宜的检验方法,还要考虑到置信度、样品的性质以及评价员等因素。

## 5.2 检验方法的分类

除在4.1.4中介绍的敏感性检验外,常用的检验方法可分为以下三类:

- 差别检验 用以确定两种产品之间是否存在感官差别。
- 标度和类别检验 用于估计差别的顺序或大小,或者样品应归属的类别或等级。
- 分析或描述性检验 用于识别存在于某样品中的特殊感官指标。该检验也可以是定量的。

## 5.3 差别检验

### 5.3.1 差别检验的类型

- 成对比较检验;
- 三点检验;
- 二-三点检验;
- 五中取二检验;
- “A”-“非A”检验。

### 5.3.2 成对比较检验

#### 5.3.2.1 应用

成对比较检验可用于:

- 确定两种样品之间是否存在某种差别,差别的方向如何。
- 确定是否偏爱两种样品中的某一种。
- 评价员的选择与培训。

这种检验方法的优点是简单且不易产生感官疲劳。

这种检验方法的缺点是,当比较的样品增多时,要求比较的数目立刻就会变得极大以至无法比较。

#### 5.3.2.2 评价员

所需要的评价员数目:

- 7个以上专家;或
- 20个以上优选评价员;或
- 30个以上初级评价员。

对于综合性研究,例如消费者偏爱检验,则需要视检验内容、要求而配备更多的评价员。

#### 5.3.2.3 做法

以确定的或随机的顺序将一对或多对样品分发给评价员。向评价员询问关于差别或偏爱的方向等问题。差别检验和偏爱检验的问题不应混在一起。

#### 5.3.2.4 结果的分析

见6.1.1。

### 5.3.3 三点检验

#### 5.3.3.1 应用

三点检验可用于以下场合:

- 确定两种样品之间细微的差别;
- 当能参加检验的评价员数量不多时;
- 选择和培训评价员。

该检验的缺点是:

- a. 用这种方法评价大量样品是不经济的;
- b. 用这种方法评价风味强烈的样品比成对比较检验更容易受到感官疲劳的影响;
- c. 要保证两种样品完全一样是很困难的。

5.3.3.2 评价员

所需要的评价员数目:

- 6个以上专家;或
- 15个以上优选评价员;或
- 25个以上初级评价员。

5.3.3.3 做法

向评价员提供一组三个已经编码的样品,其中两个样品是相同的,要求评价员挑出其中单个的样品。三个不同排列次序的样品组中,两种样品出现的次数应相等,它们是:

BAA	ABB
ABA	BAB
AAB	BBA

5.3.3.4 结果的分析

见6.1.2。

5.3.4 二-三点检验

5.3.4.1 应用

二-三点检验用于确定被检样品与对照样品之间是否存在感官差别。这种方法尤其适用于评价员很熟悉对照样品的情形。

如果被测样品有后味,这种检验方法就不如成对比较检验适宜。

5.3.4.2 评价员

需要20个以上初级评价员。

5.3.4.3 做法

首先向评价员提供已被识别的对照样品,接着提供两个已编码的样品,其中之一与对照样品相同。要求评价员识别出这一样品。

5.3.4.4 结果的分析

见6.1.3。

5.3.5 五中取二检验

5.3.5.1 应用

当仅可找到少量的(例如10个)优选评价员时可选用五中取二检验方法。

这种方法的优点是确定差别比用其他检验方法更节省(这种方法在统计学上功效高)。

这种检验方法的缺点与三点检验相同,而且更容易受到感官疲劳和记忆效果的影响。在利用视觉、听觉和触觉的感官分析中可使用该方法。

5.3.5.2 评价员

需要10个以上优选评价员。

5.3.5.3 做法

向评价员提供一组五个已编码的样品,其中两个是一种类型的,另外三个是一种类型,要求评价员将这些样品按类型分成两组。当评价员数目不足20时,样品出现的次序应随机地从以下20种不同的排序中挑选:

AAABB	BBBAA
AABAB	BBABA
ABAAB	BABBA

BAAAB	ABBBA
AABBA	BBAAB
ABABA	BABAB
BAAAB	ABBAB
ABBAA	BAABB
BABAA	ABABB
BBAAA	AABBB

5.3.5.4 结果的分析

见6.1.4。

5.3.6 “A”-“非A”检验

5.3.6.1 应用

“A”-“非A”检验主要用于评价那些具有各种不同外观或留有持久后味的样品。这种方法特别适用于无法取得完全类似样品的差别检验。

5.3.6.2 评价员

所需要的评价员数目：

20个以上优选评价员；或

30个以上初级评价员。

5.3.6.3 做法

首先将对照样品“A”反复提供给评价员，直到评价员可以识别它为止，然后每次随机给出一个可能是“A”或“非A”的样品，要求评价员辨别。提供样品应有适当的时间间隔，并且一次评价的样品不宜过多以免产生感官疲劳。

5.3.6.4 结果的分析

见6.1.5。

5.4 标度和类别检验方法

5.4.1 检验的类型

- a. 排序；
- b. 分类；
- c. 评估；
- d. 评分；
- e. 分等。

5.4.2 排序

5.4.2.1 应用

排序法具有广泛的用途。但是它的区别能力并不强，这种方法可用于：

- a. 筛选样品以便安排更精确的评价；
- b. 选择产品；
- c. 消费者接受检查及确定偏爱的顺序；
- d. 选择与培训评价员。

当评价少量样品(6个以下)的复杂特性(例如质量和风味)以及当评价大量样品(20个以上样品)的外观时这种方法是迅速有效的。

5.4.2.2 评价员

所需要的评价员数目：

2个以上专家；或

5个以上优选评价员；或

10个以上初级评价员(对于消费者检验需要100个以上评价员)。

5.4.2.3 做法

检验之前,评价员对被评价的指标和准则要有一致的理解。在检验中,每个评价员以事先确定的顺序检验编码的样品并安排一个初步的顺序作为结果,然后通过重新检验样品来检查和调整这个顺序。

5.4.2.4 结果的分析

见6.2.1。

5.4.3 分类

5.4.3.1 应用

在估价产品的缺陷等情况时可用分类法。

5.4.3.2 评价员

所需要的评价员数目:

3个以上专家;或

3个以上优选评价员。

5.4.3.3 做法

明确定义并使专家或优选评价员理解所使用的分类类别。每个评价员检查所有的样品并将其归于某一个类别中。

5.4.3.4 结果的分析

见6.2.2。

5.4.4 评估

5.4.4.1 应用

评估法可用以评价:

- a. 一个或多个指标的强度;
- b. 偏爱程度。

由于评估法能估计指标强度或偏爱的程度,所以这种检验方法比排序法能提供更多的信息。

5.4.4.2 评价员

a. 为确定指标的强度,需要:

1个以上的专家;或

5个以上优选评价员;或

20个以上初级评价员;

b. 为确定偏爱的程度,需要:

50个以上初级评价员(对两种样品);或

100个以上初级评价员(对三种以上样品)。

5.4.4.3 做法

明确定义并使评价员理解所使用的类别。标度可以是图示的或描述性的,可以是单极标度也可以是双极标度。

每个评价员检验所有的样品并将这些样品归于某一标度的位置上。如果把类别标示成数字,则不应将该数字视为评分结果。

5.4.4.4 结果的分析

见6.2.3。

5.4.5 评分

5.4.5.1 应用

评分法可用于评价一种或多种指标的强度。

#### 5.4.5.2 评价员

所需要的评价员数目：

- 1个以上专家；或
- 5个以上优选评价员；或
- 20个以上初级评价员。

#### 5.4.5.3 作法

首先清楚定义所使用的标度类型。标度可以是等距的也可以是比率的。检验时先由评价员分别评价样品指标，然后由检验的组织者按事先确定的规则在评价员评价的基础上给样品指标打分。

#### 5.4.5.4 结果的分析

见6.2.4。

#### 5.4.6 分等

##### 5.4.6.1 应用

这种检验方法主要用于产品质量评价。

##### 5.4.6.2 评价员

评价员的数目取决于所使用的具体分等的作法。

##### 5.4.6.3 做法

首先确定能代表产品质量的感官指标，然后清楚地定义所使用的标度。在满足等距标度或比率标度的条件下可使用综合评分法，在只满足顺序标度的条件下可使用综合评估法。综合评分法即首先对样品的各有关指标分别评分，再根据各指标对整个产品质量的重要程度确定的权数对各指标的分数加权平均得出对整个样品的评分结果。综合评估法即首先对样品的各有关指标分别评估，得出各指标评估类别的频数表。各指标的评估类别数应相等。然后根据各指标的权数对各指标相应类别的频数加权平均。其加权平均值最大的那个类别代表该样品的评估结果。可在评分或评估的基础上再划分等级。

##### 5.4.6.4 结果的分析

见6.2.5。

#### 5.5 分析或描述性检验

##### 5.5.1 检验的类型

这些检验方法可适用于一个或多个样品，以便同时定性和定量地表示一个或多个感官指标。可分为以下几类：

- a. 简单描述检验；
- b. 定量描述和感官剖面检验。

##### 5.5.2 简单描述检验

###### 5.5.2.1 应用

- a. 识别和描述某一特殊样品或许多样品的特殊指标；
- b. 将感觉到的特性指标建立一个序列。

这种检验方法可用于描述已经确定的差别。也可用于培训评价员。

###### 5.5.2.2 评价员

- a. 对特性指标的识别和描述：需要5个以上专家；
- b. 对所感觉到的特性指标确定一个序列：需要5个以上优选评价员。

###### 5.5.2.3 做法

这种检验可适用于一个或多个样品。当在一次评价会上呈现多个样品时，样品分发顺序可能对于检验结果产生某种影响。可通过使用不同的样品顺序重复进行检验估计出这种影响的大小。第一个出现的样品最好是对照样品。

每个评价员独立地评价样品并作记录，可以提供一张指标检查表，可先由评价小组负责人主持一次

讨论然后再评价。

#### 5.5.2.4 结果的解释

设计一张适合于样品的描述性词汇表。根据每一描述性词汇的使用频数得出评价结果。最好对评价结论作公开讨论。

#### 5.5.3 定量描述和感官剖面检验

##### 5.5.3.1 应用

这类检验方法可用于：

- a. 新产品的研制；
- b. 确定产品之间差别的性质；
- c. 质量控制；
- d. 提供与仪器检验数据相对比的感官数据。

##### 5.5.3.2 评价员

由5个以上优选评价员或专家组成评价小组。这些评价员都要经过该种方法的特殊培训。

##### 5.5.3.3 做法

用被检验的样品的各种特性预先进行一组试验，以便确定出其重要的感官特性。用这些试验结果设计出一张描述性词汇表并确定检验样品的程序。评价小组经过培训掌握方法，特别是学会如何使用这些术语词汇。在这一阶段提供一组纯化合物或自然产品的参比样是很有用的。这些参比样会产生出特殊的气味或风味或者具有特殊的质地或视觉特性。

在检验会议上，评价员对照词汇表检查样品。在强度标度上给每一出现的指标打分。要注意所感觉到的各因素的顺序，包括后味出现的顺序，并对气味和风味的整个印象打分。

##### 5.5.3.4 结果的解释

一种方式是先由评价员分别评价，然后评价小组负责人表列这些结果并组织讨论不同意见，如有必要还可对样品重新检查。根据讨论结果，评价小组对剖面形成一致的意见。另一种方式是不讨论或至多只有一个简短的讨论，得到的剖面是多少评价员评分的平均值。

处理这些结果没有简单的统计方法，但多变量分析技术可用来揭示产品之间和评价员之间是否有显著差异。

## 6 结果的分析

### 6.1 差别检验

5.3中描述的各种差别检验方法是用来确定两种样品A和B之间是否存在着可觉察的差别(或是否偏爱某一个)。分析是以每一类别的评价员数量为基础。例如有多少人偏爱样品A，多少人偏爱样品B，多少人回答的正确。对差别检验结果的解释主要是运用统计学的二项分布参数检验。

有两个共性问题：

#### a. 在差别检验中对“无差别”回答的处理：

在差别检验中，可能有的评价员没能觉察出两种样品之间的差别，因而产生“无差别”的回答，但是为了统计分析的需要，一般规定不允许“无差别”的回答(即强迫选择)。如果允许出现“无差别”的回答，那么有两种处理办法：

——忽略“无差别”的回答，即从评价小组的评价总数中减去这些数。

——将“无差别”的结果分配到其他类的回答中，即在成对比较检验和二-三点检验中将这种结果的各一半归于A和B类中，在三点检验中将这种结果的三分之一归于回答正确的类中，在五中取二检验中将这种结果的十分之一归于回答正确的类中。

#### b. 序贯方法

事先不固定检验次数，而是不断检查积累的结果，只要能作出是否有差别的判断，检验即可停止。序

贯方法的优点是所需要的检验次数和评价员数一般比较少。

### 6.1.1 成对比较检验

#### 6.1.1.1 统计解释

这种检验有两种形式。第一种形式关心的是两种样品之间具体差别的方向以及确定这种方向；第二种形式关心的是偏爱两种样品中的哪一种。

这种分析仅适用于由从样品 A 取出的一个样品与从样品 B 取出的一个样品组成的成对样品，即仅适用于被检验的成对样品是 AB 或 BA 而不是 AA 或 BB。

成对比较检验的原假设：这两种样品没有显著性差别，因而无法根据样品的特性强度或偏爱程度区别这两种样品。换句话说，每个参加检验的评价员作出样品 A 比样品 B 的特性强度大或样品 B 比样品 A 的特性强度大(或被偏爱)判断的概率是相等的，即  $P_A = P_B = \frac{1}{2}$ 。

根据作出样品 A(或样品 B)具有较大特性强度或被偏爱的判断数目对结果作出解释。结果的解释还取决于与原假设相反的备择假设。根据备择假设的性质确定检验是双边的还是单边的。

#### 6.1.1.2 双边检验

双边检验是只需要发现两种样品在特性强度上是否存在差别(强度检验)或者是否其中之一更被消费者偏爱(偏爱检验)。

备择假设：这两种样品有显著差别，因而可以区别这两种样品。换句话说，每个参加检验的评价员作出样品 A 比样品 B 的特性强度大或样品 B 比样品 A 的特性强度大(或被偏爱)判断的概率是不等的，即  $P_A \neq P_B (P_A > P_B \text{ 或 } P_A < P_B)$ 。

如果对某一种样品投票的人数不少于附录 B 表 B1 中第二列的数，则在 5% 的显著性水平上拒绝原假设，从而得出结论：两种样品之间有显著性差别。如果对样品 A 投票的人数多，则可得出结论，样品 A 的某种指标强度大于样品 B 的同种指标强度(或被明显偏爱)。

#### 6.1.1.3 单边检验

单边检验是希望发现某一指定样品，例如样品 A 比另一种样品 B 具有较大的强度(强度检验)，或者被偏爱(偏爱检验)。

备择假设：样品 A 的特性强度(或被偏爱)明显优于样品 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)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)判断的概率。即  $P_A > \frac{1}{2}$ 。

如果选择样品 A 的数目不少于附录 B 表 B1 中第四列的数，则在 5% 的显著性水平上拒绝原假设而接受备择假设。

例：在一个有 30 个评价员参加的检验中，20 个人偏爱 A，10 个人偏爱 B，并且没有理由认为 A 或 B 应被偏爱(即检验是双边的)。较大一组的人数(即 20)与附录 B 表 B1 中评价员数为 30 的第二列的数比较(即 21)。由于观测值(指偏爱 A 的人数 20)低于表中的值，所以原假设在 5% 显著性水平上不被拒绝，并且不可能得出这两种产品有哪一个更被偏爱的结论。

另一方面，如果有先验知识，A 应被偏爱，则该检验是单边的。偏爱 A 的评价员数与附录 B 表 B1 中评价员数为 30 的第四列的数(即 20)比较。由于从检验中得到的数等于表中的数，所以将以 5% 的显著性水平拒绝原假设。并且应得出对产品 A 有明显的偏爱的结论。

### 6.1.2 三点检验

原假设：不可能根据特性强度区别这两种样品。在这种情况下正确识别出单个样品的概率为  $P = \frac{1}{3}$ 。

备择假设：可以根据特性强度区别这两种样品。在这种情况下正确识别出单个样品的概率为  $P >$

$\frac{1}{3}$ 。

该检验是单边的。如果正确回答的数目大于或等于附录 B 表 B1 中第三列的相应的数，则以 5% 的显



著性水平拒绝原假设而接受备择假设。

6.1.3 二-三点检验

原假设：不可能区别这两种样品。在这种情况下，识别出与对照样品相同的样品的概率是  $P = \frac{1}{2}$ 。

备择假设：可以根据样品的特性强度区分这两种样品。在这种情况下正确识别出与对照样品相同的样品的概率为  $P > \frac{1}{2}$ 。

该检验是单边的。如果正确回答的数目大于或等于附录 B 表 B1 中第四列中的数，那么将以 5% 显著性水平拒绝原假设而接受备择假设。

6.1.4 五中取二检验

原假设：不可能区别这两种样品。在这种情况下能正确地将两种样品分开的概率是  $P = \frac{1}{10}$ 。

备择假设：可以根据样品的特性强度区分这两种样品。在这种情况下正确区别这两种样品的概率为  $P > \frac{1}{10}$ 。

该检验是单边的，将正确回答的数与附录 B 表 B1 的第五列中相应的数比较，如果正确回答的数目大于或等于附录 B 表 B1 中第五列中相应的数，则以 5% 的显著性水平拒绝原假设而接受备择假设。

6.1.5 “A”-“非 A”检验

可用列联表形式表示检验结果，然后作  $\chi^2$  检验。

原假设：评价员的判别(认为样品是“A”或“非 A”)与样品本身的特性(样品本身是“A”或“非 A”)相互独立。即：

$$n_{11}/n_{21} = n_{12}/n_{22}$$

其中： $n_{11}$ ——样品本身是“A”评价员也认为是“A”的回答数；

$n_{21}$ ——样品本身是“A”而评价员认为是“非 A”的回答数；

$n_{12}$ ——样品本身是“非 A”而评价员认为是“A”的回答数；

$n_{22}$ ——样品本身是“非 A”评价员也认为是“非 A”的回答数。

备择假设：评价员的评价与样品本身特性有关，即：

$$n_{11}/n_{21} \neq n_{12}/n_{22}$$

$\chi^2$  统计量如下：

$$\chi^2 = \sum_{i,j} \frac{(|E_{ij} - E_{ij}| - 0.5)^2}{E_{ij}} \dots\dots\dots (1)$$

式中： $E_{ij}$ ——各类观测数  $n_{ij}(i = 1, 2; j = 1, 2)$ ；

$E_{i.}$ ——等于  $n_{i0} \cdot n_{0j} / n_{00}$ 。

将  $\chi^2$  值与附录 B 表 B2 中对应自由度为 1 的表列值比较，若大于或等于表列值则拒绝原假设而接受备择假设。若小于表列值则接受原假设。

表 1 观测值

评价员判别 \ 样品特性	“A”	“非 A”	总计
“A”	$n_{11}$	$n_{12}$	$n_{10}$
“非 A”	$n_{21}$	$n_{22}$	$n_{20}$
总计	$n_{01}$	$n_{02}$	$n_{00}$

6.2 标度和类别检验

对于 5.4 中指出的各类检验方法。选择什么样的手段分析和解释数据取决于检验的目的以及所检验的样品是一种、两种还是多种。本节给出了所使用的统计方法的简要内容。