



中华人民共和国国家标准

GB/T 39625—2020/ISO 13299:2016

感官分析 方法学 建立感官剖面的导则

Sensory analysis—Methodology—
General guidance for establishing a sensory profile

(ISO 13299:2016, IDT)

2020-12-14 发布

2020-12-14 实施

国家市场监督管理总局 发布
国家标准化管理委员会

目 次

前言	I
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 检验的一般条件	3
4.1 实验室和设备	3
4.2 评价员	3
4.3 产品	4
4.4 样品	4
4.5 初步讨论	4
5 描述性方法的原则和主要特征	4
5.1 共识性感官剖面	4
5.2 对照偏离参比描述法(相对参比评分)	4
5.3 自由选择剖面	4
5.4 自选特性排序剖面	5
5.5 定量描述剖面	5
5.6 定性感官剖面	5
5.7 动态主导感官属性测试(TDS)	5
6 建立感官剖面的步骤	5
6.1 总则	5
6.2 测试准备	5
6.3 开展测试	7
6.4 结果统计与解释	7
6.5 测试报告	7
附录 A(资料性附录) 共识性感官剖面	9
附录 B(资料性附录) 对照参比描述法(相对参比评分)	11
附录 C(资料性附录) 自由选择剖面	13
附录 D(资料性附录) 自选特性排序剖面	15
附录 E(资料性附录) 定性感官剖面	17
附录 F(资料性附录) 定量描述剖面	19
附录 G(资料性附录) 动态主导感官属性测试(TDS)	25
附录 H(资料性附录) 评价小组数据的单变量分析	29
参考文献	38

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准使用翻译法等同采用 ISO 13299:2016《感官分析 方法学 建立感官剖面的导则》。

与本标准中规范性引用的国际文件有一致性对应关系的我国文件如下：

- GB/T 10220—2012 感官分析 方法学 总论(ISO 6658:2005, IDT)
- GB/T 10221—2012 感官分析 术语(ISO 5492:2008, MOD)
- GB/T 13868—2009 感官分析 建立感官分析实验室的一般导则(ISO 8589:2007, IDT)
- 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(idt ISO 5496:1992)
- GB/T 16291.1—2012 感官分析 选拔、培训与管理评价员一般导则 第 1 部分:优选评价员(ISO 8586-1:1993, MOD)
- GB/T 16291.2—2010 感官分析 选拔、培训和管理评价员一般导则 第 2 部分:专家评价员(ISO 8586-2:2008, IDT)
- 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(idt ISO 11035:1994)

本标准做了下列编辑性修改：

- 将原标准术语和定义 3.9 中的“ISO 5492:2008/修改单 1:—¹⁾”更改为“ISO 5492:2008/修改单 1:2016”，删去原标准中的脚注 1“出版中”，同时将标准中其余脚注的数字依次向前移动 1 位；
- 将原标准附录 H 中 H.2.2 最后一句“表 H.1 数据表明不显著”更改为“表 H.2 数据表明，产品因素影响显著”，纠正原文错误；
- 将原标准 H.4.1.1 中的“ $\alpha\delta_{ir}$ 是产品 \times 重复的互作效应”更改为“ $\alpha\delta_{ir}$ ——评价员 \times 重复的互作效应”，纠正原文错误；
- 将原标准 H 中表 H.5 中的“ $r-1=1$ ”更改为“ $s-1=1$ ”，纠正原文错误；
- 将原标准 H.4.1.2 中 F' 分母的自由度计算公式中，将“ $+MS_{a \times p \times s}$ ”更改为“ $-MS_{a \times p \times s}$ ”；
- 将原标准附录 H 中“ MS_{error} ”更改为“ MS_e ”；
- 将原标准附录 H 中表 H.9 中的“8.842”更改为“8.839”，“38.85”更改为“38.56”，纠正原文错误；
- 将原标准附录中表 H.11 中的“8.84”更改为“0.87”，“0.000 3”更改为“0.47”，“77.09”更改为“77.00”，“1.62”更改为“1.63”，纠正原文错误。

本标准由全国感官分析标准化技术委员会(SAC/TC 566)提出并归口。

本标准起草单位：中国标准化研究院、中国烟叶公司、中国烟草总公司郑州烟草研究院、内蒙古伊利实业集团股份有限公司、浙江工商大学。

本标准主要起草人：钟葵、赵镭、史波林、汪厚银、李锐、申玉军、苏玉芳、田师一、张璐璐、刘龙云。

感官分析 方法学

建立感官剖面的导则

1 范围

本标准给出了建立感官剖面的导则。

本标准适用于能通过视觉、嗅觉、味觉、触觉或听觉评价的所有产品或样品(例如食品、饮料、烟草制品、化妆品、纺织品、纸张、包装材质、空气或水样品等)的感官分析。

本标准也适用于人类认知和行为的研究。

感官剖面的一些应用如下:

- 用于开发或改进产品;
- 根据感官特性定义产品、制定生产标准或贸易标准;
- 定义用于产品货架期测试的“新鲜”参比样;
- 研究和改进产品的保质期;
- 将产品与参比样、市面上同类产品或开发中的产品进行比较;
- 描述产品的感知特性,并将其与仪器分析、理化性质和/或消费者接受性等因素相关联;
- 描述样品中臭味或异味的种类和强度(例如污染研究)。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

ISO 5492 感官分析 术语(Sensory analysis—Vocabulary)

ISO 5496 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(Sensory analysis—Methodology—Initiation and training of assessors in the detection and recognition of odours)

ISO 6658 感官分析 方法学 总论(Sensory analysis—Methodology—General guidance)

ISO 8586 感官分析 选拔、培训与管理评价员一般导则(Sensory analysis—General guidance for the selection, training and monitoring of assessors)

ISO 8589 感官分析 建立感官分析实验室的一般导则(Sensory analysis—General guidance for the design of test rooms)

ISO 11035 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(Sensory analysis—Identification and selection of descriptors for establishing a sensory profile by a multidimensional approach)

ISO 11136 感官分析方法 对受控区域的消费者开展喜好测试的一般导则(Sensory analysis—Methodology—General guidance for conducting hedonic tests with consumers in a controlled area)

3 术语和定义

ISO 5492 界定的以及下列术语和定义适用于本文件。

3.1

特性 attribute

样品具有的可被感知的特征。

[来源:ISO 5492:2008,1.3,有修改]

3.2

感官剖面 sensory profile

对样品感官特性(3.1)的描述,常同时赋予其特性强度值。

3.3

部分感官剖面 partial sensory profile

对某些选定的感官特性(3.1)的描述,常同时赋予其特性强度值。

示例: 气味剖面,风味剖面和质地剖面。

3.4

定量描述剖面 quantitative descriptive profile

对样品感官特性(3.1)及其强度值的描述。

[来源:ISO 5492:2008,4.24,有修改]

3.5

定性感官剖面 qualitative sensory profile

对样品感官特性(3.1)的描述,不包含强度值。

[来源:ISO 5492:2008,4.23,有修改]

3.6

共识性感官剖面 consensus sensory profile

感官评价小组对产品感官特性(3.1)进行集体讨论后得到的具有共识性的描述。

3.7

对照参比描述法(相对参比评分) deviation from reference method (relative-to-reference rating)

将所有样品与参比样进行对比评价的一种定量描述感官剖面(3.2)的方法。

3.8

自由选择感官剖面 free-choice sensory profile

每位评价员独立描述一组样品,为其选择感官特性(3.1)和评分的方法。

3.9

自选特性排序剖面 flash profile

评价员选择自己的描述词,并依据描述词强度对产品排序,从而确定产品特征的方法。

注: 此法是自由选择剖面的一种变体,区别在于自选特性排序剖面使用排序代替自由选择剖面中的评分。

[来源:ISO 5492:2008/修改单 1:2016,4.57]

3.10

动态主导感官属性测试 temporal dominance of sensations;TDS

要求每位评价员连续指出产品评价过程中不同时间呈现的主导感官特性的方法。

3.11

评价小组 sensory panel

参加感官分析的评价员组成的小组。

[来源:ISO 5492:2008,1.9,有修改]

3.12

小组组长 panel leader

管理小组活动并招募、培训和管理评价员的人员。

注：小组组长应负责设计和开展感官测试、分析和解释数据。

[来源：ISO 13300(所有部分)，有修改]

3.13

优选评价员 **selected assessor**

挑选出的具有开展感官测试能力的评价员。

[来源：ISO 5492:2008，1.6]

3.14

专家评价员 **expert sensory assessor**

具有高度的感官敏感性、足够的培训时长和丰富的感官分析经验，能够对各类产品做出一致的、可重复的感官评价的优选评价员(3.13)。

[来源：ISO 5492:2008，1.8]

4 检验的一般条件

4.1 实验室和设备

实验室应按照 ISO 6658 中的规定配备用于样品准备的设备。

开展感官剖面的实验室应符合 ISO 8589 中的规定。当需要开展关于剖面结果、待测产品和参比样等相关讨论时，房间内布局应满足评价员之间交流的需求。小组组长应确保在合适的环境条件下开展产品评价(例如适宜的照明条件)。

小组组长负责开展感官剖面工作，内容包括：

- 培训评价员；
- 维护评价小组；
- 开展测试。

小组组长应符合 ISO 13300-1 和 ISO 13300-2 规定的任职资格(例如熟悉招募和培训步骤)。

4.2 评价员

本标准采用优选评价员或专家评价员进行感官剖面建立。评价员的选拔、培训和管理应符合 ISO 8586 中的规定。

评价员的数量及培训内容取决于选用的剖面方法。通常，评价员水平越高，培训时间越长，评价结果的重复性和再现性就越高。结果解释以及产品间的差异发现也取决于评价员的数量和培训效果。

候选评价员的招募可通过面谈、发招募单或单独联系等方式。按照所需人数 2~3 倍的数量进行面试和筛选。应重点考察候选人的以下特点：

- 产品评价所需的健康状况；
- 兴趣和动机；
- 能投入的时间以及出勤保证；
- 及时性；
- 集中注意力的能力；
- 记忆能力；
- 如实沟通和感知表达的能力；
- 待测样品特征差异的辨别力；

——团队合作的能力。

评价员个体的感官敏锐性差异可通过组建 10 人或 10 人以上的评价小组进行组内平衡。

4.3 产品

待测产品及其制备条件应明确。

示例：冲泡速溶咖啡可以选择用水或用牛奶，加糖或不加糖。

4.4 样品

样品的准备和提供应符合 ISO 6658 中的规定。确保评价员不能从样品提供方式中得出有关样品特性的结论。必要时应使用有色眼镜或彩色光源遮蔽样品的外观差异。

样品的准备和分发应在恒温条件下进行，并保证操作规范。样品采用三位随机编码，提供顺序应采用适当的实验设计。

为提高测试结果的可靠性和有效性，每个样品或每组样品应重复测试不少于 2~3 次，并尽可能安排在不同测试日。重复次数取决于实验要求的精度、观测结果的离散度以及评价员熟悉样品后逐渐提升的辨别力。重复测试数据可以估算试验误差。同批次产品的重复检验结果反映了同一评价员评分值的离散程度，不同批次产品的重复检验结果反映了产品间的差异。应在测试方案中明确指出，哪些样品需重复检验，以及样品准备和测试的具体要求。

评价员完成所有样品评价前，样品信息均不得公开。

4.5 初步讨论

应确保评价员完全熟悉待测样品的每一个特性和 ISO 6658 中规定的原则和方法。有需要时应开展关于测试中存在的问题和样品性质的小组讨论。讨论会上，应提供一些典型样品进行讨论，小组组长应确保讨论内容不会对后续评价带来偏差。

5 描述性方法的原则和主要特征

5.1 共识性感官剖面

共识性感官剖面分析中，评价员通过分享个人观点，就产品的特性、特性出现顺序和特性强度达成一致性结果。

通常，共识性感官剖面使用的标度类项数较少，每种特性只有 1 个评分值（达成一致的评分）。若有评价员不同意小组的评分时，应如实记录在报告中。

5.2 对照偏离参比描述法（相对参比评分）

样品通常以成对形式提供。评价员依据通用特性列表中的每一个特性，直接对 2 个产品进行比较。评价员也可以先对 2 个产品评分后再比较评分值。2 个以上产品比较时，应在相同条件下将每个产品与参比样进行比较。

采用统计方法对样品和参比样间的差异进行分析。

5.3 自由选择剖面

自由选择剖面中，每位评价员使用他/她自己建立的特性列表而不是通用特性列表。

采用合适的多元分析方法(例如广义普鲁克分析)对结果进行解释,分析结果应以绘图方式呈现。

5.4 自选特性排序剖面

自选特性排序剖面是自由选择剖面的一种变体,是同时提供所有样品,对样品已选定的特性进行比较后再排序的一种方法。

采用合适的多元分析方法(例如广义普鲁克分析)对结果进行解释,分析结果常以绘图方式呈现。

5.5 定量描述剖面

定量描述剖面中,评价员依据通用特性列表中特性对样品进行评价,并对每一个特性强度进行评分。

建立定量描述剖面的技术较多,一些已注册商标¹⁾。结果中的特性强度评分采用单变量或多变量分析。

5.6 定性感官剖面

定性感官剖面中,评价员只依据提供的特性列表判断产品是否存在该特性,无需给出特性的感知强度。

与定量描述剖面相比,定性感官剖面特性列表中的词汇范围更广,产品的依赖性更小。小组培训时应集中在对大样本量的参比样的识别和记忆上。为便于记忆,参比样应稳定且不随时间变化。相比定量描述剖面,定性感官剖面所需评价员的数量和/或重复次数应更多。

定性感官剖面结果通常以每种特性出现的频次表示。

5.7 动态主导感官属性测试(TDS)

TDS是一种动态剖面,要求每位评价员连续记录在评价过程中占主导的感知特性。

评价员依据通用特性列表选择占主导的感知特性。同时,可对所选占主导的特性强度进行评分。

TDS的结果以每一时刻占主导的特性的占比表示,通常以曲线图形式呈现, x 轴设置为时间轴。待测产品不同特性的TDS曲线应合并到一张图中。

6 建立感官剖面的步骤

6.1 总则

本部分给出了建立感官剖面的通用步骤。每种方法的详细描述,请参阅相应附录。

6.2 测试准备

6.2.1 选择训练用样品

训练用样品应符合ISO 8586中的规定。

6.2.2 筛选评价员

评价员的筛选应符合ISO 8586和ISO 5496中的规定。

1) QDA®和Spectrum™是目前适合定量描述剖面分析的商业程序实例。这些信息仅为方便本标准的使用者而提供,并不构成对相关程序的认可。

6.2.3 选择最佳特性

对待测样品开展尽可能完整的描述性分析,以获取一组不重叠、独立、客观、明确和可参比的特性。此步骤非常重要,通常采用一种感官剖面或多种感官剖面共同完成。如需建立一个通用特性列表时,小组组长可采用表 1 中提供的 3 种方式中的任意一种或几种方法结合开展。

表 1 最佳特性的选择方法

序号	原 则	方 法	优 点	缺 点
1	现有采用: 采用已建立的描述词和参比样	采用参考文献和专家咨询的方式选择适宜的描述词(现有全部的或部分的描述词)。采用参比样训练评价员理解每个描述词的含义。必要时,可为描述词提供一个强度标度	借鉴了专家多年积累的丰富经验。其他评价小组可以解释该剖面,并与其他相关研究展开比较	对于某些特定样品,现有的特性描述词或参比样可能不够准确或恰当。新描述词开发期间,产品中本应被感知的特性可能会被遗漏
2	讨论开发: 使用评价小组,通过讨论的形式开发描述词	在经验丰富的小组组长指导下,优选评价员组建的评价小组通过圆桌讨论方式开发特性描述词。 讨论中会使用到参比样,参比样可由小组组长或测试方提供,或研讨时由评价员自行提供。本方法可与方法 1 结合使用	描述词的开发过程比方法 3 耗时少	建立的剖面仅适用于特定的评价小组和样品。无参比样时,其他评价小组不能对其进行解释
3	评价筛选: 使用评价小组,通过多轮次评价产品的形式生成描述词	按照 ISO 11035,评价员采用一组预先准备好的培训样品鉴别和选取合适的特性描述词,再进一步通过统计分析逐渐删减描述词数目	使用一个完全客观的选择和删减流程,最大限度地删除传统误解或先入为主的描述词,所选的描述词能最佳表达出评价员感知到的样品特性	建立的剖面仅适用于特定的评价小组和样品。无参比样时,其他评价小组不能对其进行解释。该过程相对耗时,且需要有一定经验,特别是数据分析方面的经验

建立特性描述词期间,可以要求评价员提供对产品的一个或多个的整体评价,例如:

- 整体的果味或辣感强度;
- 整体的风味强度。

注意:经过培训的感官评价小组开展的产品整体评价不应包括喜好测试,因其会带来结果偏差。经过培训的评价员能够避免(有意识或无意识)使用典型消费者的评价策略,从而对产品进行客观描述。如需开展喜好测试指导新产品开发,应按照 ISO 11136 中的规定进行。

6.2.4 确定评价顺序

对于某些产品(例如饮料),确定产品特性的感知顺序是感官剖面的重要环节。其他的某些产品,评价过程中产品特性的感知顺序会改变,例如与质地或物理性质相关的特征。

示例:一块融化的巧克力,一张湿润的面巾纸。

产品特性的评价顺序通常由特性的感知顺序决定。后味或后感,是产品被食用或使用后仍然可被感知的特性,通常放在最后评价。通过样品提供方式可以预知食物质地在口腔中的感知顺序,这一点类似于肤觉和织物感评价。小组组长通过设计样品提供方式(咀嚼或手动挤压)确定特性的感知顺序。

6.2.5 选择适宜的响应标度

感官剖面中需测定样品特性强度时,应谨慎选择响应标度。感官剖面采用的响应标度可以是数字的或语义的标度,连续的或不连续的标度,单极的或双极的标度,如 ISO 4121 中所述。

注:采用响应标度测定的数据可能会受到“末端效应”影响。评价员通常不会选择靠近标度两端的数字,避免后续样品的评分更加极端。通常,连续响应标度,如线性标度,不易受末端效应影响。

6.2.6 培训评价员

评价员的培训应按照 ISO 8586 中的规定进行,宜将培训意见反馈给评价员。

6.3 开展测试

6.3.1 回答表

使用纸质打印回答表,表中应注明评价所采用的标度;也可使用计算机或平板电脑进行数字化记录。回答表上应留有空白,要求评价员对感官特性提供意见或建议。

注:当一页纸或一个电脑屏幕上同时显示多个特性的评价记录时,可能会相互影响或曲解(例如光环效应,对一个特性正面或负面的评价可能会影响对下一个特性的评价)。

6.3.2 样品评价

应安排评价员在独立评价间开展样品评价。

根据使用的评价方法,样品可单样(一次提供一个样品,连续提供)或同时多个样品提供。多样同时提供有利于样品间的比较,但会导致剖面分析不独立。无论采用何种供样方式,评价员间的样品提供顺序都应遵循平衡或随机原则,以减少顺序效应的影响。

每位评价员和每个轮次中评价的样品数量应根据每轮次所需时长、产品性质、特性数量和预期差异进行调整。如果预计样品间的差异很小,且样品具有强烈或持久的味道时,可减少每个轮次中评价样品的数量。

6.4 结果统计与解释

进一步分析前,应对原始数据进行检查和图形化处理,旨在:

- 检查并纠正明显错误;
- 在适当情况下,可删除异常值;
- 获得对整体数据的一个初步印象;
- 发现对后续分析中具有价值的可能的数据趋势。

采用多种分析方法分析数据,如数据的平均值及标准偏差,以及箱形图、直方图和曲线图。分析的数据可以是全部的数据,或者是评价员个体数据和某个选定特性的数据。

数据处理取决于采用的统计方法。通常,数据解释包括三个步骤。第一步,重点关注评价员的表现和检查数据是否存在实验误差。第二步,关注每一个描述词,找到最能区分待测产品差别的描述词。这一步通常称为单变量分析(见附录 H)。第三步,评估第一阶段所有描述词的有效性,也称为多变量分析。这一步通常在描述词分类后进行,如视觉描述词、风味描述词、味觉描述词和质地描述词。

注:本标准无法对所有多变量分析技术一一展开叙述,附录中介绍不同剖面方法时有所提及。

6.5 测试报告

测试报告应包含以下信息:

- a) 开展测试的公司或实验室名称,小组组长和研究负责人姓名;

- b) 测试目的；
- c) 样品的完整信息；
- d) 测试日期、评价时间和持续时长；
- e) 测试操作过程(包括与本标准中方法不一致的任何条件):
 - 1) 注明采用本标准进行评价,包括质量控制流程；
 - 2) 详细的测试方案:样品提供顺序和方式,重复性的定义(相同/不同的日期,相同/不同的批次)；
 - 3) 评价员的数量和类型(优选评价员或专家评价员)；
 - 4) 待测特性描述词及定义、评价方案和参比样；
 - 5) 采用的响应标度；
 - 6) 评价期间提供给评价员的任何信息(例如产品的类型或品牌信息)；
- f) 测试结果,结果进行统计分析后以适宜形式(图形、图表和绘图等)呈现。如果与合作伙伴或赞助者有相关协议,报告内容可无需特别详细。

附录 A

(资料性附录)

共识性感官剖面

A.1 原则

共识性感官剖面中,评价员对产品展开评价,讨论产品的特性和/或强度,最终达成一致性的描述结果。该方法要求结果达成一致,而不是数据均值化。评价小组应通过大量培训来确定产品的特性,并对感知到的特性和强度得到一致性结果。

20 世纪 40 年代末 Arthur D Little 有限公司开发了风味剖面法(Flavor Profile Method, FPM), 20 世纪 60 年代通用食品公司的产品评价和质地技术小组开发了质地剖面法(Texture Profile Method, TPM)。这 2 种方法的重点均是结果达成共识。在此基础上,研究人员于 20 世纪 90 年代继续开发出 SpectrumTM法²⁾。FPM 和 TPM 的本质区别在于:FPM 只有一种方式达到共识性,而 TPM 有 2 种方式(见 A.6)。

A.2 感官特性

评价小组组长负责提供和维护参比样,评价员开发和定义整个小组使用的通用描述词。培训期间确定每个描述词的参比样和定义。FPM 中,评价员描述风味并估计强度,同时评价产品的“整体印象”(称为振幅)。TPM 使用标准化的术语,从物理性质和感官性质两方面描述产品的质地特征。

A.3 标度

最初,FPM 法采用 5 点标度评价特性的强度。

0:不存在;

) (:处于阈值水平;

1:轻微;

2:中等;

3:强烈。

后期对该标度进行了扩展,包括使用箭头、1/2 s 或加号及减号。TPM 法最初采用 13 点标度,也是 FPM 的扩展标度。如今,TPM 最常用的是 15 点标度,如 10 cm 或 15 cm 线性标尺。

A.4 评价员

建立共识性感官剖面的评价员具有以下特点:

——数量少(FPM 中为 4~6 位);

——接受过高强度培训。FPM 分析时会提供给评价员一整套参比样,参比样范围应包括待测产品全部的类型、成分和加工技术差异。TPM 的评价员应学习如何使用标度评分,并保证结果的再现性。

2) QDA[®] 和 SpectrumTM是目前适合定量描述剖面分析的商业程序实例。这些信息仅为方便本标准的使用者而提供,并不构成对相关程序的认可。

见参考文献[16]。

A.5 步骤

- 评价员独立对样品逐一评价,记录特性的强度和出现顺序。
 - 产品评价结束后,小组组长收集评价结果,带领评价小组成员展开讨论,就产品特性、强度和其他方面(如特性出现顺序)达成一致性结果。讨论过程通常持续几分钟或几十分钟,期间提供参比样有助于讨论达成一致性结果。评价小组达成一致性结果时,描述性感官“剖面”得以建立。
 - 结果未达成一致时,评价小组组长可决定提供产品开展新一轮评价或者记录评价员间的差异。
 - 重复该过程直至所有产品评价完毕。
 - 小组组长应避免出现资深评价员或主导性人物的观点影响其他评价员的情形。
- 以上是基本步骤,实际测试时会出现多种情况。

A.6 结果统计与解释

FPM 不采用统计分析,数据结果常以表格或图形呈现。TPM 的数据结果由小组协商一致或通过数据统计分析得出。最终报告中,数据以表格或图形呈现。

采用多元统计方法(如“树”形图或主成分分析)进行产品间相似性和差异性的分析。

附录 B (资料性附录)

对照参比描述法(相对参比评分)

B.1 原则

对照参比描述法,也称相对参比评分,由 Larson 和 Pangborn(1978)首次提出。该方法是通用描述性分析方法的变体,使用参比样评价所有样品(样品成对形式呈现)。该方法适用于有明确参比样或样品区分困难的情形。通常采用 2 种方式,一种是评价员直接比较样品与参比样并评分(参比样设在标度的中心点),另一种是评价员根据样品和参比样的评分差对样品进行比较。这种情况下,评价员通常不知道哪个样品是参比样。

B.2 感官特性

所有评价员采用通用特性列表,特性列表的建立步骤与定量描述分析一致。特性列表应在培训期间建立,可以使用现有描述词构成的“封闭列表”,或由评价小组重新建立。

B.3 标度

响应标度可以是结构式的或非结构式的标度。

样品与参比样直接比较时采用双极标度,参比样评分作为中心点。该情况下,对于给定特性,评价员对样品的评分通常是比参比样更高或更低(图 B.1)。



图 B.1 参比样作为中心参比的非结构式标度示例

当样品不直接与参比样对比评价时(评分后再比较差异),采用与定量描述分析类似的标度。对于给定特性,2 种标度非常相似(图 B.2)。

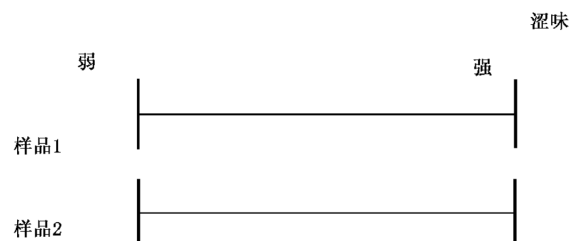


图 B.2 计算参比样和样品间差异的非结构式标度示例

B.4 评价员

通常根据待测产品,参照定量描述剖面进行评价员的筛选和培训。每次评价时采用相同的参比样,

能有效提高评价员结果的一致性以及多次评价时的评分稳定性。

B.5 步骤

培训过程类似于定量描述剖面。训练环节包括评价员熟悉产品类别的环节,开发和掌握特性列表的环节以及使用标度对特性强度评分的环节等。

评价时,评价员会收到一组 2 个样品。评价员和评价轮次之间的样品提供顺序需遵循平衡的原则。通常有以下 2 种方式:

- 直接比较(图 B.1)。2 个样品中,1 个样品为参比样,另 1 个样品与参比样进行比较。通常先评价参比样,评价过程中评价员可将待测样与参比样重复比较。
- 分别评分(图 B.2)。评价员可多次品尝样品,也可同时评价 2 个样品(如同时在面颊两侧涂抹不同化妆品)。评价结束后再对 2 个样品的差异进行分析。

以上 2 种方式中,均可将参比样设成样品(“盲参比样”,即未明确指出哪个样品是参比样)展开评价,此法常用于评价员可靠性的内部测定。

特定条件下,可开展 2 种以上样品间的比较,同时应注意:

- 所有样品与参比样间的比较都应在相同条件下;
 - 参比样筛选应慎重,确保与待测样品是同类型产品。
- 与定量描述剖面一样,应展开重复性评价。

B.6 结果统计与解释

对上述 2 种方法中样品和参比样之间的差异展开分析。

与中心点相比,样品评分低于参比样的通常为负值,高于参比样的为正值,采用经典单变量统计分析方法。样品成对比较时(其中 1 个是参比样),采用双尾配对 t 检验。样品间需相互比较时,只要满足上述条件,可采用单变量和多变量分析方法。该情况下,每对样品中参比样的分值用盲参比样的分值代替。

附录 C

(资料性附录)

自由选择剖面

C.1 原则

自由选择剖面(FCP)是每位评价员建立自己的特性列表,并采用预先确定的或自己选择的标度对特性强度量化的方法。

FCP 常用于 3 个或以上、具有多个特性的产品间的比较及差异分析。该法对建立产品的认知地图,确定产品的空间(例如用于市场研究)很有用,也可用于建立产品特性的描述词。

C.2 感官特性

评价员独自建立自己的感官特性描述词,且不需要对描述词含义进行解释。组织评价员开展产品特性的讨论有利于更好地描述产品,评价员使用自己的特性列表,讨论无需达成一致性结果。每位评价员的描述词无需一致,评价员自己能理解即可。要求评价员始终使用自己的特性列表完成对所有产品的评价。评价员可选择所有感官特性或部分感官特性,如形状、颜色、味道和气味等进行评价。有些评价员的特性描述词数量很少。对于特定特性的描述没有限定,评价员可以使用任何他/她想用的词汇或短语对特性进行描述(如“闻起来像露营的气味”)。

C.3 标度

评价员采用预先确定的或自己选择的标度对特性强度进行评分。允许评价员采用不同标度进行产品评价。要求评价员始终使用自己的特性列表完成对所有产品的评价,无需对评价员进行标度使用培训。

C.4 评价员

评价员应具有良好的视觉、嗅觉、味觉、触觉和听觉,事先未接受过某一产品的特定培训。要求筛选具有良好描述能力的评价员,能够使用较为全面、综合的词汇对具有多种特性的产品展开描述。无需对评价员展开培训。

评价员数量宜不少于 10 位。

C.5 步骤

第一步,评价员建立自己的产品特性列表,并选择预先确定的或自己想使用的标度。

第二步,在符合标准的测试条件下,评价员始终使用自己的特性列表完成所有待测产品的特性及强度评价。样品提供顺序遵循随机原则。每位评价员和每个样品均应不少于 2 次重复。

贮备网格法可作为 FCP 的替代方法。相比 FCP,贮备网格法生成产品特性的方式更加结构化。首先生成产品的一些简单特性,再通过一系列比较逐渐完善产品特性。具体实施时,先选择 2~4 个产品(通常 3 个)进行比较,获得产品的特性。然后再将上一轮中的 1 个产品与另外 2 个新产品进行比较。

要求评价员回答 2 个新产品与第 1 个产品的相似或不同之处。依此类推,3 个产品一组,测试照此持续进行。

C.6 结果统计与解释

FCP 得到的产品描述不能采用方差分析(ANOVA)、主成分分析或线性回归等方法统计分析。通常采用广义普鲁克分析法(GPA)对每位评价员的单次评价和重复评价数据展开统计,通过数据转换(旋转、平移和缩放)以获取具有一致性的结果。

附 录 D
(资料性附录)
自选特性排序剖面

D.1 原则

自选特性排序剖面由 Sieffermann 于 2000 年首次提出,结合了自由选择剖面和排序两种方法。评价员独立描述一组产品间的差异并生成产品特性。自选特性排序剖面通过感官映射图,在感官空间上将相似的产品分成一类。特性强度评价采用排序法,主要强调产品间的相对差异。

D.2 感官特性

要求每位评价员能独立对产品间的差异进行描述,并生成特性描述词。这是成功开展自选特性排序剖面的重要前提,不允许使用喜好类描述词汇。

评价员间围绕产品的特性展开讨论有利于更好描述产品。每位评价员应使用自己的特性列表,讨论结果无需达成一致。

D.3 标度

自选特性排序剖面不使用标度,只要求评价员按照特性强度对样品进行排序。

D.4 评价员

评价员应具有良好的视觉、嗅觉、味觉、触觉和听觉,之前未接受过针对某一产品的特定培训。应选用具有良好描述能力的优选评价员,能够使用综合全面的词汇对具有多个特性的产品展开描述,后期无需进一步培训。

评价员数量宜不少于 10 位。

D.5 步骤

首先进行 1~2 轮次的培训,指导评价员如何描述和辨别产品间的差异。通常此过程中,每位评价员会获得一组产品独立展开描述,通过样品分类找到产品间差异。后期评价员在小组组长协助下去除无关特性或喜好特性,并添加相关的特性。

接下来开展样品测试,根据特性强度对产品排序,依次完成所有特性的测试。宜同一轮次开展部分样品的重复评价或多个轮次中进行所有样品的重复评价。

与 FCP 不同,自选特性排序剖面是所有产品采用平衡区组或平衡不完全区组设计后同时提供给评价员。根据待测产品的数量设计样品提供的方式,样品数量宜不少于 5~6 个。待测样品数量超过 12 个时,应采用平衡不完全区组设计。

D.6 结果统计与解释

与 FCP 类似,采用多元统计方法对自选特性排序剖面的结果进行分析,如广义普鲁克分析(GPA)。评价员的可靠性可通过计算重复间的相关性进行考察(斯皮尔曼秩相关性)。

附录 E

(资料性附录)

定性感官剖面

E.1 原则

定性感官剖面是基于描述词选用频次的一种方法,常用于描述复杂的芳香类产品,如葡萄酒或奶酪。通常给评价员提供一个足够数量的描述词列表,评价员从表中选出与待测产品最相关的描述词。根据评价员的选择次数(引用次数)计算评分。足量的描述词能确保评价员选择最合适的词汇对感知到的特性进行精准描述。采用专家评价员(经过培训的评价员)展开测试。

E.2 感官特性

采用“封闭式”特性列表,由已有明确定义的现有描述词或专门针对该产品开发和改进的描述词组成。特性列表应在培训期间建立。列表中描述词的数量多于定量剖面特性列表中的描述词数量,通常为 100 个左右。特性主要集中在香气或气味上,应确保能与香气参比样精确对应。每个描述词都应有 1 个对应的标准香气参比样(商用参比样或天然产品)。

E.3 标度

定性感官剖面不使用标度。评价员根据感知的产品特性,在描述词列表中进行选择,标注为检验到或未检验到。

E.4 评价员

E.4.1 评价员数量

评价员的数量应高于定量描述中所需评价员数量,通常是定量剖面评价小组的 2 倍(16~30 位合格的评价员)。

E.4.2 资格和培训

培训通常分为 2 个阶段:常规培训和特定产品培训。常规培训期间(根据需要记忆的特性数量,进行 5~15 轮次或更多),评价员应熟悉特性列表,能从多个标准香气参比样中选出最符合术语定义的参比样。

常规培训期间,根据产品进行特性列表的开发及修正。评价员应对特性列表展开讨论并修订,删除不相关、不明确或冗余的描述词,补充有关联的描述词。

特定产品培训期间,评价员应通过嗅闻标准香气参比样熟悉待测产品。

E.5 步骤

经过培训的评价员拿到一整组产品,每种产品的品尝顺序应遵循平衡的原则。对于每种产品,要求

每位评价员从特性列表中选择与他们感知的香气或气味匹配的描述词。评价员能选择的描述词数量有限,特性列表中描述词的最大数量取决于:

- 1) 评价员能够使用最大数量的描述词表达对产品感知的可能性;
- 2) 人类对混合气体辨别能力的有限性。

定性感官剖面应重复测试。

E.6 结果统计与解释

完成所有数据收集后,对描述词的选用频次进行排序,确定每个产品最相关的描述词。一个产品或一个重复中,只有超过一定比例(通常为小组的 10%~15%)的评价员选用的描述词才被考虑用于后续的统计分析。以产品为横列,描述词为纵列,选用频次为指标建立产品的列联表。

结果可采用多种统计方法进行分析,常见的有:

- 对每个术语的选用频次(2 次重复)进行卡方(χ^2)分析,找出产品间特性“强度”的显著差异;
- 多元分析:对特性列表中术语的选用频次进行统计分析,表中所列描述词应是被 15%以上评价员(一个产品/重复中)选用的描述词。

附录 F

(资料性附录)

定量描述剖面

F.1 原则

定量描述剖面(QDP)是采用经过培训的评价小组,对产品所有感官特性进行描述并对强度量化的方法。

针对待测产品建立专属评价小组,使用与待测产品相同类型的产品进行培训,并生成产品的感官特性。

所得剖面是专属评价小组针对特定产品类别建立的剖面。无参比样时,其他评价小组不能对其解释。

F.2 感官特性

定量描述剖面应采用经过培训的评价员选用和定义产品的特性,并对指定的系列产品进行区分。

通常,第一次培训侧重生成产品的感官特性(如食品的外观、香气、味道、质地和后味等特性),并确保评价员能正确使用覆盖所有待测样品特性强度范围的标度。

通常,通过讨论达成一致或采用多元分析方法对于特性建立初期鉴定出的、数量较多的描述词进行删减,最终建立一个具有有限数量描述词的特性列表(见 ISO 11035)。

F.3 标度

确定样品特性后,选择合适的响应标度对样品特性强度进行测定。

定量感官剖面采用的响应标度可以是连续的标度(10 cm~15 cm)或不连续的标度。通常,标度两端标记为无(或非常弱)和强(或非常强)。

F.4 评价员

选用经过培训的评价员,采用与待测产品同类型的样品对通过选拔测试的培训员进行强化培训(见 ISO 8586)。

F.4.1 评价员资格

F.4.1.1 招募和选拔

合格的招募者应是待测产品的使用者或接受者。选拔测试中,他们应具备以下条件:

——辨别待测产品间差异的能力;

——记忆产品感官特性的能力。

评价员应能保证能参加测试项目完成所需的所有环节。

F.4.1.2 培训

针对特定产品展开培训。

优选评价员在小组组长指导下建立感官特性描述词,或对现有描述词进行修改完善,确保能对待测产品的所有感官特性展开描述。

培训样品应覆盖待测产品的所有感官特性及其强度的变化范围。此环节应重点关注培训样品的种类和偏差的代表性范围(所有产品类型,包括所有的典型产品以及与典型产品有差异的非典型产品)。

由评价小组制定每个特性的定义及标准化的评价步骤,并开展产品评分练习。

参比样由小组组长或测试方提供,也可由评价员在训练期间自行提供。通常,参比样与测试样品不同。

1 个产品类别的训练时间通常需 10 h~20 h 或更长。训练时间的长短取决于产品的特性以及分配给评价小组的任务。

评价小组组长应负责组织培训、提供产品、必要时介绍参比样以及协助产品评价活动的相关工作,但不应参与产品评价。

F.4.1.3 表现评估

培训结束时,小组组长可通过监测表现数据来改善评价小组和评价员的表现,考察培训效果和筛除表现不佳、不宜继续留用的评价员。

小组评价产品时,应进行重复性测试,通常重复 2~3 次。

采用统计分析考察整个评价小组和每位评价员的一致性和重复性。

F.4.2 评价员数量

评价员数量应不少于 8 位,宜 12~15 位。

F.5 步骤

——通常,待测样品采用单样依次提供的顺序。

——宜重复测定 2~3 次。

F.6 结果统计与解释

F.6.1 总则

核查后的数据经可采用多种统计方法进行结果分析。方差分析(ANOVA)是最常用的统计方法,双因素方差分析广泛应用于产品间以及评价员间的差异分析。其他差异来源,如“重复性”的差异同样需考虑。具体统计分析方法见附录 H。

特性间的交互作用可采用多元统计方法。同时分析不同产品对多个特性影响的统计方法较多,主成分分析(PCA)是最常用的方法,尤其适用于数据量大、背景相关信息较少的情况。PCA 可在方差—协方差矩阵或相关矩阵上进行。第 1 种方法中所有特性的权重并不相同,依据每种特性的差异而定。第 2 种方法并不考虑特性差异,删除差别小的特性后再赋予所有特性相同权重。特性被拒绝的临界概率通常是 0.05,变化范围通常为 0.10~0.50。

F.6.2 结果呈现

结果常以数字和图表形式呈现(如蛛网图、剖面曲线及多元分析的相关图表)。

F.7 方法

F.7.1 定量描述分析³⁾

定量描述分析(QDA[®])法是基于产品特性的测量和量化。大量参考文献和书籍均对 QDA 进行了介绍,见参考文献[9]、[27]、[28]。QDA 的建立为人类感觉相关的测量提供了良好实践和理论基础。随着人类实践和相关知识的不断进步,QDA 法也从最初演化到今天的 Tragon QDA[®]。

QDA 法应用广泛,包括消费者描述和区分产品的语言理解、绘制感知到的产品相似性和差异、成分替换、新品开发、对比评价、索赔证明和广告等用途。当与消费者的情感测量(例如喜欢、偏好、态度、情感、包装、使用)相关联时,可以使用 QDA 数据确定影响消费者选择行为的最关键感官特性,并理解消费者独特偏好的原因。开发人员基于这些信息开发满足目标消费者利益和需求的产品,市场营销人员也可通过 QDA 得到有关产品的描述词和感官特性,帮助他们更好地进行产品宣传。

F.7.1.1 感官特性

建立感官特性的描述词通常需要 8 h~12 h 的小组讨论。这是一个迭代过程,每次讨论都是在上一次讨论结果的基础上进行,直至所有描述词建立完成。部分描述词的建立步骤中可能包含观察行为,包括在家和/或其他的使用环境下。对建立的描述词进行初步测试后再开展描述词补充。描述词使用前、使用中和使用后,创建 30~40 个或更多的感官特性词汇充分地描述产品的感觉均属于正常情况。

F.7.1.2 标度

QDA 采用非结构化的或半结构化的线性标度对感知的特性及强度进行测量和量化。

通常采用非结构化的或半结构化(6 in/15 cm)的线性标尺,距离两端各 0.5 in 处标识为参照点,对感知的差异和强度进行测量和量化。这些等距标度在众多心理物理学出版物中均有介绍。

F.7.1.3 评价员

F.7.1.3.1 评价员数量

评价员数量宜为 12 位。

F.7.1.3.2 资格和培训

F.7.1.3.2.1 招募与选拔

候选者应是待测产品类型的爱好者,招募数量应大于产品评价所需的平均人数,按照感官敏锐性、语言表达流畅度以及参加测试项目完成所需所有环节的可能性筛选约 30 位候选人或消费者。

F.7.1.3.2.2 培训

主持人或小组组长负责培训工作。小组组长负责引导和推动小组展开讨论,制定小组活动时间表,协助小组建立待测产品的常用描述词。主持人不负责教学工作,主要观察消费者在典型消费者应用情境下的消费行为和对产品的看法。

候选评价员感官敏锐性测试包括 30 项测试项目,重复性测试结果采用差别检验。测试内容包括产

3) QDA[®] 和 Spectrum[™]是目前适合定量描述剖面分析的商业程序实例。这些信息仅为方便本标准的使用者而提供,并不构成对相关程序的认可。

品间的差异区分、感兴趣产品类型间的差异区分以及不同难度范围(从简单、适中到难区分)的差异区分。

F.7.1.3.2.3 表现评估

QDA 法提供了小组表现评估的统计方法,包括单个评价员在特性、重复性上的表现以及相对于整个评价小组的整体差异。

下一步分析重点是感知产品间的相似性和差异性。评价结果采用方差分析,详细介绍见参考文献[5]。

F.7.1.4 步骤

针对感兴趣的产品类别建立标准化的评价步骤。

评价步骤中应包括消费者对此类产品的典型使用行为模拟,可能需要多个评价步骤。

采用平衡区组设计,宜不少于 3 次重复。

F.7.1.5 结果统计与解释

结果采用方差分析(ANOVA)。

QDA 数据结果有多种呈现形式,最常见的是蛛网图(也称雷达图),统计指标有平均值和差异显著性分析。

F.7.2 SpectrumTM 法

F.7.2.1 原则

SpectrumTM法是基于描述性剖面的方法,包括使用参比样进行特性的测定及量化。SpectrumTM法起源于风味剖面 and 质地剖面,通过扩展标度、建立更多的特性参比样等方式,将方法适用范围扩展到了食品描述性评价之外,如肤觉和织物感评价。该方法从评价员筛选、小组组长选拔、小组培训、验证以及培训后小组维护等各个阶段均有明确的方法和步骤。这些实践操作可使描述性评价小组在多环节测定和多产品类别评价时获取可靠性高、再现性强的统计数据。

F.7.2.2 感官特性

SpectrumTM法是采用一个经过多种产品类型(食品、香水、肤觉、织物和纸张等)SpectrumTM训练的评价小组来获取项目或产品的感官特性。评价员按照给定的评价步骤,在有经验的小组组长指导下建立产品的特性描述词。每个特性有明确的定义和实物参比样,描述词应能够描述和区分同一类型的不同产品。建立的特性应是可感知、主要的、交叉不重叠的和单一的特性。

F.7.2.3 标度

SpectrumTM标度是一个 0~15 分的强度标度,每一分又分成 10 等份,共计 150 个分值。该标度在早期标度基础上进一步细化,使评价员能够使用更精确的评分值区分样品的差异。SpectrumTM标度属于通用标度,涵盖了食品或肤觉所能体验到的所有强度范围。采用绝对强度值进行标记,15 分并非真正终点,可以根据需要延长。大多数产品强度均在 0~15 分范围内,必要时也可采用超过 15 分的强度

4) QDA[®] 和 SpectrumTM是目前适合定量描述剖面分析的商业程序实例。这些信息仅为方便本标准的使用者而提供,并不构成对相关程序的认可。

值(例如极端的酸味,更高的油腻感)。为比较样品内和样品间的强度,培训中应开展特性和强度交互设计的相关训练。例如强度为5的甜味与强度为5的苦味或强度为5的薄荷味。Spectrum™法的通用标度是基于大样本规模的食物种类,涵盖的强度范围包括刚能觉察(阈值)、轻微、强和非常强。Spectrum™中,通用参比样的标度值是来自10个不同行业评价小组重复评价的结果。应定期核查参比样的标度值变化(因其对应的常规产品自身也在变化,见参考文献[4])。

F.7.2.4 评价员

F.7.2.4.1 资格

F.7.2.4.1.1 招募与选拔

采用多种措施从招募人员中筛选评价员。招募期间,针对健康问题、参与所有评价环节的可能性和对工作内容/时间的兴趣度、基础视力以及描述产品或情境的语言表达能力等方面,对潜在评价员进行筛查。根据项目特点,招募人数通常为所需评价员数量的3倍。通过预筛的人员进行面试以及敏锐性和技能的相关选拔。达到筛选标准,且敏锐性和组内互动优秀的人员选拔为评价员(宜18~20位)。

F.7.2.4.1.2 培训

培训过程主要包括一系列高强度的训练环节以及每个环节后紧接着的若干练习,以期增强对待测产品的熟悉度。例如1个培训方案包括28h的训练加45h的练习。培训首先对Spectrum™进行整体介绍,再采用相关领域的多种产品类别展开培训。培训和练习初期主要介绍整体方法、强度标度以及如何选择特性词汇/投票流程。后期的培训和练习主要集中在评价小组的待测项目等特定领域。

F.7.2.4.1.3 表现评估

培训结束时应考察评价小组的可靠性和一致性。验证实验可对评价小组整体和评价员个体的表现进行评估,帮助确定潜在的需再培训的方面。受训评价小组的数据应具有再现性,且能有效对产品进行区分。通过验证实验设计,使用样品和重复性评价对评价小组这两方面的表现进行测试。除培训结束时,宜定期对评价小组开展工作表现评估(例如每年一次)。小组验证的概念类似于其他科学仪器的校准。

F.7.2.4.1.4 评价员数量

Spectrum™评价小组通常由10~12位评价员组成。评价员数量可根据评价员的工作年限、表现评估和评价方法(例如一致性数据)进行调整,但宜不少于6位。

F.7.2.5 步骤

通过5个规范步骤构建产品品质空间感官特性的描述词。

- a) 参比样框架建立——评价员通过体验大量已有明确描述词的产品增加经验和扩展感兴趣的领域。体验的产品可包括市场上的产品和用于建立理想描述词的样品。
- b) 描述词生成/分组——评价员根据自己经验对产品进行描述。评价小组组长根据相似性对描述词进行分组,并与评价小组一起初步建立产品的描述词。
- c) 特性参比样——评价员体验已有明确描述词的参比样。
- d) 描述词优化——小组组长引导开展参比样适用性和术语冗余性的讨论。删去冗余、有偏差或可整合的描述词,建立特性描述词。
- e) 描述词验证——采用成对样品测试建立的描述词。描述词应能有效区分样品,该步骤也可用

于描述词或参比样优化。

F.7.2.6 结果统计与解释

Spectrum™评价小组的数据采用参数检验法进行分析。强度标度上有 151 个差异点,标准统计指标包括平均值、标准偏差、方差分析、平均距离和相关性等。频次属于非参数据,例如存在或不存在的特性。无论是个体数据还是共识性评价的平均值,均可采用多元分析(聚类分析、主成分分析/因子分析等)进行产品分类。

附录 G

(资料性附录)

动态主导感官属性测试(TDS)

G.1 原则

TDS 是一种感官描述性方法,要求评价员连续指出占主导地位的感觉。对于评价员,“主导感觉”被定义为给定时间引起注意的感觉,但并不表示这种感觉应是产品中非常强烈或最激烈的感觉。与时间强度法相比,TDS 考虑到随时间变化的感知多维性。

TDS 能够收集到整个品尝过程中不同时间点感受到的一系列感官特性。

每位评价员配备一个秒表,可在纸上记录下答案,但使用计算机记录会更加舒适,评价员也不必担心时间问题,能更专注于识别占主导的特性。收集的数据进一步分析后绘制成占主导特性随时间变化的曲线。

G.2 感官特性

提供给评价员一个特性列表,要求评价员从该列表中选择在整个评价过程中占主导的特性。因此,特性列表中的定义是评价员回答的关键要素。TDS 特性列表通常从特性生成时开始构建,如定量描述性分析,然后通过一致性或统计分析进行删减。

列表中描述词数量宜不超过 10~12 个。列表太长时,部分评价员无法保证顾及整个列表。

列表中特性的位置通常会影响特性的选择顺序,列表顶部的特性通常比底部的特性更早被选择。宜平衡各评价员的特性顺序,但对每位评价员来说,整个实验过程中特性在列表中的顺序应该始终保持一致(无论产品和重复过程如何变化)。

同一列表中可包含多种类型的特性(例如风味特性和质地特性)。由于不同种类特性间的强度水平差异较大,每种特性宜单独评价,或选择将不同种类的特性放在 1 个轮次中评价,以获取更多信息。

G.3 标度

TDS 最简便之处是不使用标度,评价时仅需在特性列表中选择感觉占主导的特性。

G.4 评价员

G.4.1 培训

不同于定量剖面的培训,TDS 在培训期间并不开展强度培训,主要侧重针对不同的感官特性进行训练,提高对主导感觉的判断力。优选评价员应为每个特性建立一致性定义。使用参比样可提高评价员对特性的识别能力。

培训期间(次数根据评价员而定,通常 1~4 次)应着重引导评价员熟悉 TDS 的特征(如使用计算机系统)。

当评价员熟悉了 TDS 方法后,所需的培训课程数量少于定量描述分析,通常只需 2~4 个轮次就可

适应特性列表。

G.4.2 评价员数量/重复次数

TDS 采用经过培训的评价员,评价员数量应多于定量剖面,如 16 位评价员,测试重复次数不少于 2 次,确保每个产品最终不少于 30 个评价结果。

G.5 步骤

- 将一口量大小的产品一次性提供给评价员,要求评价员将整个样品放入口腔内,同时开始评价(通常点击“开始”按钮)。
- 当评价员认为主导感觉已经发生改变时,他/她应在特性列表中找到新的占主导感觉。评价员重复该操作,直至感知过程结束。点击“停止”按钮,停止计时。
- 评价员可自主地给相同产品选择同一个特性,或者完全不选择列表中的一个或多个特定特性。
- 测试过程应明确指出何时开始和何时停止评价,确保收集数据的可靠性。

G.6 结果统计与解释

G.6.1 原始数据处理

感官分析师需确定每次评价期间占主导的特性名称和计时开始后所需评价时间。占主导的特性是指从被选择后一直到选择其他特性前的特性。

G.6.2 TDS 曲线

图 G.1 显示了甜味特性的 TDS 曲线绘制。曲线图绘制时需要考虑到每个特性,对于每个时间点,计算选择某一占主导特性的评价员(评价员 \times 重复)比例。将这些平滑(非迭代平滑样条)比例与时间进行对应作图,计算 TDS 曲线。1 个产品所有特性的 TDS 曲线合在一张图中呈现。

为了更深入展示 TDS 图形,绘制两条线。一条是“随机水平”线,是某一特性可以被随机感知到的主导率(P_0), P_0 值等于 $1/p$,其中 p 是特性的数量。另一条是“显著性水平”线,最小值。这个比例值通过二项式分布计算获取,通常认为是显著高于 P_0 值。

注:由于样品进入口腔后持续时间不同,因此评价员间感觉的时间尺度有差异。TDS 曲线绘制时应考虑到这一现象,对每位评价员的数据进行标准化处理。将开始评分到完成吞咽视为一个咀嚼期, x 轴显示从 $x=0$ (占主导的特性首次被感知)到 $x=100$ (吞咽),计算每个主导特性的持续时间占整个咀嚼期的比例。未标准化处理的曲线便于观察不同特性在持续时间上的差异,但由于所有评价并非同时结束,导致结果不容易理解。标准化处理后的曲线有利提升产品性能(更高的主导率,特别是刚开始时和结束前)。

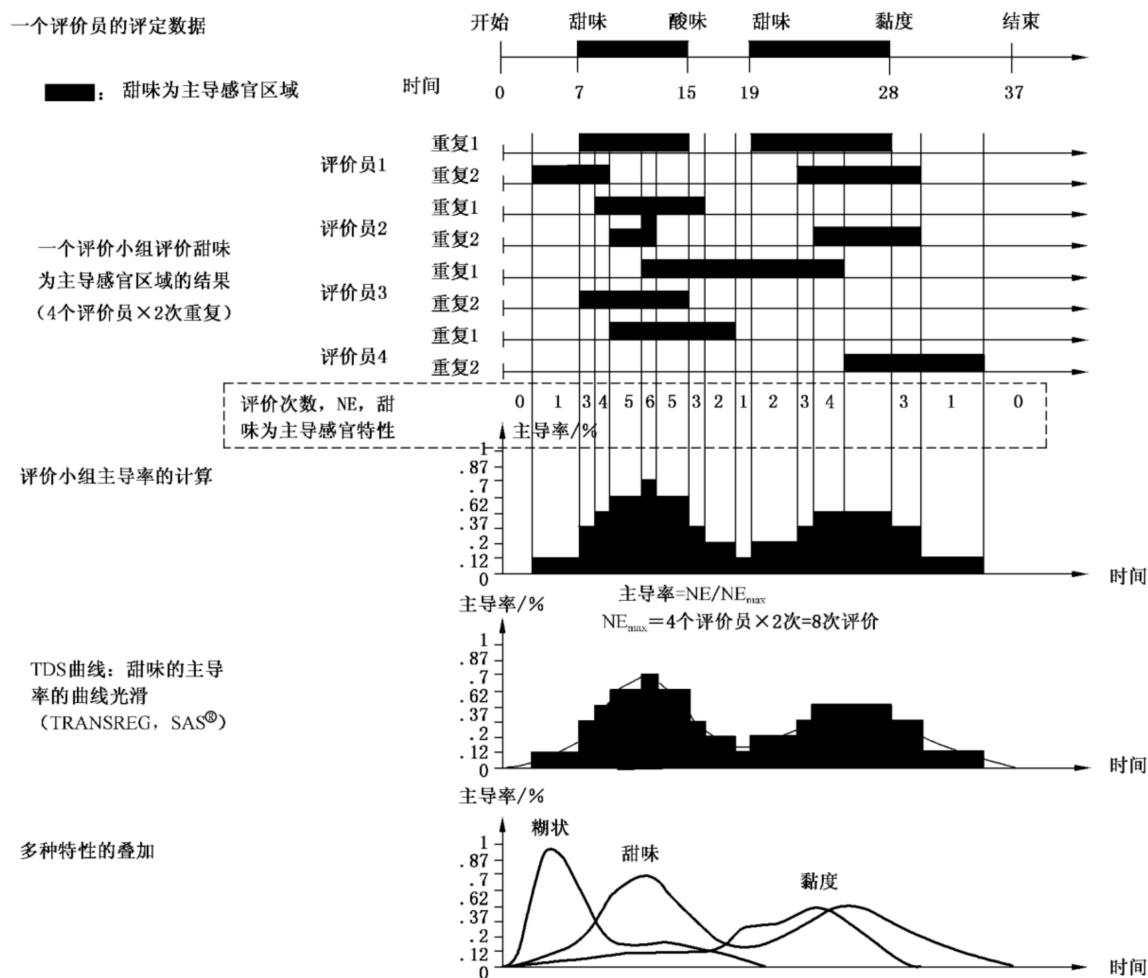


图 G.1 甜味特性的 TDS 示例(4 个评价员,2 次重复)

G.6.3 TDS 差异曲线

TDS 曲线在一个图上显示 p 个特性的曲线。为了比较 2 个产品,可以将这些主导率存在差异的特性曲线进行叠加。当差异达到显著时就会在图中绘出,凸显了产品间随时间变化时的特性差异。对 2 个二项式进行检验,可以得到随时间变化的特性差值曲线的显著性水平。

G.6.4 其他统计方法

也可使用其他统计方法进行结果分析,例如整个评价期间特性“停止”次数的方差分析,对占主导特性的持续时间开展主成分分析(PCA)或典型变量分析(CVA),阐明随时间变化的产品平均差异;或对不同产品随时间变化进行 PCA 分析,突出产品差异的时间动态变化。

G.7 其他方法

TDS 法通常采用经过培训的优选评价员,但也可采用“当地消费者”通过 TDS 计算机程序开展。为了便于消费者理解,特性列表须足够简单。首先应向消费者介绍感觉的时间概念,让他们明确特性的定义,并指导消费者学会使用 TDS 计算机程序。必要时给消费者提供几个体验样品,让消费者熟悉评

价方法和产品特性。消费者开展的 TDS 通常不会重复测试,因此消费者数量宜在 30 个以上。

根据参考文献[28],若要求评价员对所选特性的强度评分,会增加任务难度,应进行相应培训。为了不混淆 2 个不同的认知过程,占主导的特性(定性任务)选择和强度评分(定量任务),宜用按钮替代强度标度对强度进行表征。图 G.2 给出一个饼干响应的示例。

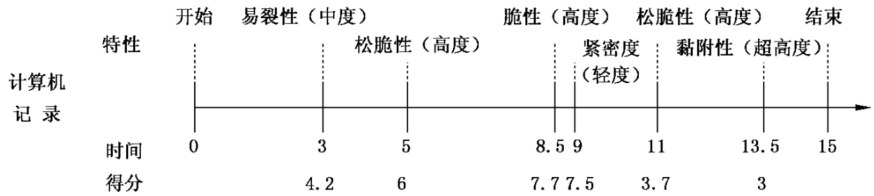


图 G.2 TDS 计算机记录示例

松脆特性占了 2 次主导地位,第一次的持续时间是 3.5 s(5 s~8.5 s,强度评分:4.2),第二次持续了 2.5 s(11 s~13.5 s,强度评分:3.7)。整个 TDS 期间,松脆特性占主导的总持续时间为 3.5 s+2.5 s=6 s。脆性强度的评分通过 2 个评分上的平均加权(持续时间)计算: $4.2 \times 3.5 / 6 + 3.7 \times 2.5 / 6 = 3.99 \approx 4$ 。

附录 H (资料性附录)

评价小组数据的单变量分析

H.1 总则

下述分析方法特别适用于采用定量描述剖面、Spectrum™和 QDA®⁵⁾法对产品进行评价。本附录提供了一种对产品之间重要特性进行比较的方法。

H.2 双因素方差分析(产品和评价员因素),无重复

H.2.1 ANOVA 模型

通常,基本统计模型如下:

$$y_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} \quad \dots\dots\dots (H.1)$$

式中:

y_{ij} ——第 i 个评价员($i:1$ 到 a)对第 j 个产品($j:1$ 到 p)的评分值;

μ ——所有评分值的平均值;

α_i ——评价员效应,假设为随机因素;

β_j ——产品效应,假设为固定因素;

$\alpha\beta_{ij}$ ——评价员×产品的互作效应。

试验设计通常称为随机完全区组设计(1 个评价员是 1 个区组),也称为内部评价员设计或重复测量设计。

H.2.2 示例

10 位评价员对 3 种产品展开评价,结果如表 H.1 所示。表 H.1 中数据进行双因素方差分析,结果如表 H.2 所示。

表 H.1 无重复的测试结果示例

评价员	产品 1	产品 2	产品 3
1	2	3	3
2	8	6	7
3	5	4	5
4	4	3	5
5	5	3	5
6	4	4	8
7	4	5	5
8	6	2	4
9	4	3	8
10	5	6	5
平均值	4.7	3.9	5.5

5) QDA®和 Spectrum™是目前适合定量描述剖面分析的商业程序实例。这些信息仅为方便本标准的使用者而提供,并不构成对相关程序的认可。

表 H.2 表 H.1 数据的双因素方差分析

方差来源	因素状态	自由度 (<i>df</i>)	平方和 (<i>SS</i>)	均方 (<i>MS</i>)	<i>F</i> 值	<i>p</i> 值
评价员	随机	$a-1=9$	34.300	$MS_a=3.811$	$MS_p/MS_{a \times p}=3.95$	0.038
产品	固定	$p-1=2$	12.800	$MS_p=6.400$		
评价员×产品	随机	$(a-1) \times (p-1)=18$	29.200	$MS_{a \times p}=1.622$		

根据评价员×产品的交互作用对产品因素进行检验。表 H.2 数据表明,产品因素影响显著。

注:假设评价员是固定因素,例如研究评价员在小组培训期间的表现。除非假设试验误差估算是评价员×产品的交互作用,否则不能对 2 个固定因素进行检验。这个假设下,可以根据试验误差对 2 个固定因素进行检验,即评价员×产品的交互作用。

H.3 双因素方差分析(产品和评价员因素),有重复

H.3.1 ANOVA 模型

通常,基本模型如下:

$$y_{ijr} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijr} \quad \dots\dots\dots (H.2)$$

式中:

- y_{ijr} ——第 i 个评价员($i:1 \sim a$)对第 j 个产品($j:1 \sim p$)在第 r 次重复中($r:1 \sim r$)的评分值;假定所有产品的重复次数相同;
- μ ——评分值的平均值;
- α_i ——评价员效应,假定为随机因素;
- β_j ——产品效应,假定为固定因素;
- $\alpha\beta_{ij}$ ——评价员×产品的互作效应;
- ϵ_{ijr} ——随机重复误差,是试验误差的估计值。

这种试验设计通常称为随机完全区组设计,每个试验单元不止一个观测值,由产品和评价员的交互作用形成。通常模型中重复是随机的,但如果是同一个评价轮次或同一个产品批次内的重复,则重复之间存在关联。事实上,随机完全区组设计中,不同产品评分应彼此独立,产品之间不存在关联。一个示例是单个轮次的品尝中,10 位评价员均需评价 4 个品种、每个品种 2 根胡萝卜样品。每次重复中,提供给每位评价员 1 根胡萝卜,因此每个品种需要 20 根胡萝卜,每位评价员需评价 $4 \times 2=8$ 根胡萝卜。

H.3.2 示例

表 H.1 的评价员对产品进行了 2 次评价,表 H.1 中只给出了重复 1 的结果。表 H.3 给出了 2 次重复的结果,双因素方差分析结果如表 H.4 所示。

表 H.3 研究结果示例(两次重复)

评价员	产品 1		产品 2		产品 3	
	重复	重复	重复	重复	重复	重复
	1	2	1	2	1	2
1	2	3	3	3	3	5
2	8	7	6	6	7	7
3	5	6	4	6	5	7
4	4	5	3	5	5	8
5	5	7	3	5	5	6
6	4	6	4	2	8	8
7	4	5	5	8	5	7
8	6	8	2	5	4	6
9	4	3	3	2	8	7
10	5	7	6	4	5	8

表 H.4 表 H.3 双因素方差分析

方差来源	因素状态	自由度 (<i>df</i>)	平方和 (<i>SS</i>)	均方 (<i>MS</i>)	<i>F</i> 值	<i>p</i> 值
评价员	随机	$a-1=9$	48.350	$MS_a=5.372$	$MS_a/MS_e=3.322$	0.006 3
产品	固定	$p-1=2$	38.033	$MS_p=19.017$	$MS_p/MS_{a \times p}=6.673$	0.006 8
评价员×产品	随机	$(a-1) \times (p-1)=18$	51.300	$MS_{a \times p}=2.850$	$MS_{a \times p}/MS_e=1.763$	0.082 4
误差		$(r-1) \times a \times p=30$	48.500	$MS_e=1.617$		
总计		$a \times p \times r-1=59$	186.183			

根据评价员×产品的交互作用检验产品因素,产品因素显著。

根据试验误差检验评价员因素和评价员×产品的交互作用。评价员因素显著,评价员×产品的交互作用不显著。

注:如果评价员是固定因素,例如研究评价员在小组培训期间的表现,则表 H.4 的 3 个因素均是根据误差进行检验。则产品因素的 *F* 值等于 11.761(*p* 值=0.000 17);另外 2 个因素的 *F* 值不变。

H.4 三因素方差分析(产品、评价员和重复因素)

H.4.1 重复因素与其他 2 个因素交互

H.4.1.1 ANOVA 模型

模型含有 8 个组分:

$$y_{ijr} = \mu + \alpha_i + \beta_j + \delta_r + \alpha\beta_{ij} + \alpha\delta_{ir} + \beta\delta_{jr} + \alpha\beta\delta_{ijr} \quad \dots\dots\dots (H.3)$$

式中:

y_{ijr} ——第 *i* 个评价员($i:1 \sim a$)对第 *j* 个产品($j:1 \sim p$)在第 *r* 次重复中($r:1 \sim r$)的评分值;

- α_i ——评价员效应,假设为随机因素;
- β_j ——产品效应,假设为固定因素;
- δ_r ——重复效应,假设为固定或随机因素均可,最常见的是随机因素;
- $\alpha\beta_{ij}$ ——评价员 \times 产品的互作效应;
- $\alpha\delta_{ir}$ ——评价员 \times 重复的互作效应;
- $\beta\delta_{jr}$ ——产品 \times 重复的互作效应;
- $\alpha\beta\delta_{ijr}$ ——评价员 \times 产品 \times 重复的互作效应。

该试验设计通常也称为两因素完全重复测量设计或三因素交叉设计。

当每个试验单元的观测值多于1个时,如评价员—产品—重复3个因素1组,则该模型设计还有第9个组分,为试验误差估计。

剖面研究中,重复因素通常是评价轮次。例如H.3中的重复1在星期二测试,重复2在同一周的星期四测试,即重复1为轮次1,重复2为轮次2。

H.4.1.2 示例

当评价员为随机因素,产品和轮次为固定因素时,表H.3数据进行三因素方差分析,结果如表H.5所示。

表 H.5 表 H.3 的三因素方差分析

方差来源	因素状态	自由度 (df)	平方和 (SS)	均方 (MS)	F 值	p 值
评价员	随机	$a-1=9$	48.350	$MS_a=5.372$	$MS_p/MS_{a \times p}=6.673$ $MS_s/MS_{a \times s}=8.405$	0.006 7 0.017 6
产品	固定	$p-1=2$	38.033	$MS_p=19.017$		
轮次	固定	$s-1=1$	16.017	$MS_s=16.017$		
评价员 \times 产品	随机	$(a-1) \times (p-1)=18$	51.300	$MS_{a \times p}=2.850$	$MS_{p \times s}/MS_{a \times p \times s}=0.787$	0.47
产品 \times 轮次	固定	$(p-1) \times (s-1)=2$	1.233	$MS_{p \times s}=0.617$		
评价员 \times 轮次	随机	$(a-1) \times (s-1)=9$	17.150	$MS_{a \times s}=1.906$		
评价员 \times 产品 \times 轮次	随机	$(a-1) \times (p-1) \times (s-1)=18$	14.100	$MS_{a \times p \times s}=0.783$		
总计		$a \times p \times s-1=59$				

根据评价员 \times 产品的交互作用检验产品因素,结果显著。根据评价员 \times 轮次的交互作用检验轮次因素,结果显著。根据评价员 \times 产品 \times 轮次的交互作用检验产品 \times 轮次的交互作用,结果不显著(参考文献[10],361-367)。

不能对其他因素进行检验。但是,如果假设评价员 \times 产品 \times 轮次的交互作用为可忽略的,则该部分可认为是试验误差的估计。这种假设下,评价员因素、评价员 \times 产品的交互作用和评价员 \times 轮次的交互作用均可根据评价员 \times 产品 \times 轮次的交互作用进行检验。

注 1: 轮次为随机因素时,该试验中包含2个随机因素(评价员和轮次)和1个固定因素(产品)。

- 根据评价员 \times 轮次的交互作用检验评价员因素; F 值等于 $5.372/1.906=2.818$ (p 值=0.069)。
- 根据评价员 \times 轮次的交互作用检验轮次因素; F 值等于 $16.017/1.906=8.403$ (p 值=0.017)。
- 根据评价员 \times 产品 \times 轮次的交互作用检验评价员 \times 产品的交互作用, F 值等于 $2.850/0.783=3.640$ (p 值=0.004 4)。
- 根据评价员 \times 产品 \times 轮次的交互作用检验产品 \times 轮次的交互作用,如表H.5所示(p 值=0.47)。
- 产品因素和评价员 \times 轮次交互作用不能被检验。对于产品因素而言,这非常困扰。但是,通过计算类似 F 值的 F' 值克服这个缺陷(参考文献[10],368-370):

$$F' = MS_p / (MS_{a \times p} + MS_{p \times s} - MS_{a \times p \times s}) = 19.017 / (2.850 + 0.617 - 0.783) = 19.017 / 2.684 = 7.085$$

F' 分母的自由度采用 Satterthwaite 近似法获得：

$$\frac{(MS_{a \times p} + MS_{p \times s} - MS_{a \times p \times s})^2}{\frac{MS_{a \times p}^2}{df_{a \times p}} + \frac{MS_{p \times s}^2}{df_{p \times s}} + \frac{MS_{a \times p \times s}^2}{df_{a \times p \times s}}}$$

对于这个例子：

$$df = (2.684\ 2) / (0.451 + 0.190 + 0.034) = 7.204 / 0.675 = 10.673$$

分子自由度为 2 和分母自由度为 11(最接近 10.67 的整数值) $F' = 7.085$, 概率等于 0.002 7。产品因素显著。

注 2：当 3 个主要因素均是固定因素时，除非假定评价员 \times 产品 \times 轮次的交互作用可忽略，否则不能检验因素的效应。这种假设下，所有 6 个因素(3 个主要因素和 3 个交互作用)的检验均根据评价员 \times 产品 \times 轮次的交互作用进行。

H.4.2 重复因素嵌套在产品因素中

这种类型的试验设计需仔细分析试验条件。以下有 3 个例子，例 1：烤肉的感官分析，重点关注猪肉品种的影响。共有 5 个品种，每个品种有 4 头猪。这 4 头猪(重复)嵌套在品种因素(主效应因素)中。例 2：啤酒生产商生产了一种新啤酒产品，有 5 种替代配方。每个配方有 3 个批次。批次因素嵌套在配方因素内。例 3：研究 6 种马铃薯的差异。每个品种种 3 行，随机分配在一块田中。马铃薯收获后，每行的马铃薯放在不同容器中。行因素嵌套在品种因素中。

这种试验设计称为嵌套因子设计。

嵌套因子的名称因方法而异，本节中称为批次。

H.4.2.1 方差分析模型

模型有 5 个组分：

$$y_{ijr} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \delta_r + \beta\delta_{jr} \dots\dots\dots (H.4)$$

式中：

y_{ijr} ——第 i 个评价员($i:1 \sim a$)对第 r 次批次中($r:1 \sim r$, 假设所有产品批次的数量相同)的第 j 个产品($j:1 \sim p$)的评分值；

α_i ——评价员效应，假设为随机因素；

β_j ——产品效应，假设为固定因素；

$\alpha\beta_{ij}$ ——评价员 \times 产品的互作效应；

δ_r ——批次效应，假设为随机因素，也可能是固定因素；

$\beta\delta_{jr}$ ——产品 \times 批次的互作效应。

当产品和批次都是固定因素时，唯一的随机因素是评价员。表 H.3 数据进行嵌套分析，结果如表 H.6 所示。

表 H.6 表 H.3 数据的嵌套方差分析(产品和批次为固定因素)

方差来源	因素状态	自由度 (df)	平方和 (SS)	均方 (MS)	F 值	p 值
评价员	随机	$a - 1 = 9$	48.350	$MS_a = 5.372$	$MS_p / MS_{a \times p} = 6.673$	0.006 8
产品	固定	$p - 1 = 2$	38.033	$MS_p = 19.017$		
评价员 \times 产品	随机	$(a - 1) \times (p - 1) = 18$	51.300	$MS_{a \times p} = 2.850$		
批次(产品)	固定	$(b - 1) \times p = 3$	17.250	$MS_b = 5.750$	$MS_b / MS_{a \times b} = 4.970$	0.007 1
评价员 \times 批次(产品)	随机	$(a - 1) \times (b - 1) \times p = 27$	31.250	$MS_{a \times b} = 1.157$		
总计		$a \times p \times b - 1 = 59$				

根据评价员×产品交互作用检验产品因素,根据嵌套在产品因素中的评价员×批次交互作用检验批次因素,结果显示这2个因素均显著。

注1:当批次为随机因素时,此时有2个随机因素(评价员和轮次)和1个固定因素(产品)。评价员因素、评价员×产品交互作用和批次因素是根据评价员×批次(产品)的交互作用进行检验。评价员 F 值等于 $5.372/1.157=4.643$ (p 值=0.096),评价员×产品的交互作用 F 值等于 $2.850/1.157=2.463$ (p 值=0.016 6)。批次 F 值已经在表 H.6 中计算出。

产品因素用准 F 或 F' 进行检验。

第一步:计算均方,通过以下计算获得:

$$MS_{\text{test},P} = MS_{b(p)} + MS_{a \times p} - MS_{a \times b} = 5.750 + 2.850 - 1.157 = 7.443$$

第二步:找到 $MS_{\text{test},P}$ 的 df 值,通过以下计算获得:

$$Den = \frac{MS_{\text{test},P}^2}{\frac{MS_{b(p)}^2}{df_{b(p)}} + \frac{MS_{a \times p}^2}{df_{a \times p}} + \frac{MS_{a \times b}^2}{df_{a \times b}}} = \frac{7.443^2}{\frac{5.75^2}{3} + \frac{2.85^2}{18} + \frac{1.157^2}{27}} = 4.808$$

第三步:计算 F' :

$$F' = \frac{MS_p}{MS_{\text{test},P}} = \frac{19.017}{7.443} = 2.555$$

最后一步:在 H_0 下找到 F' 的相关概率,分子 $df=2$ 、分母 $df=5$ (最接近 4.808 的整数)时值为 0.172。产品因素不显著。

注2:假设是3个固定因素时,所有的检验均针对评价员×批次(产品)的交互作用进行。评价员 F 值等于 4.64 (p 值=0.000 9);产品 F 值等于 16.44 (p 值=0.000 2),评价员×产品的交互作用 F 值等于 2.46 (p 值=0.017)。表 H.6 中已列出批次的 F 值。

H.5 随时间变化的评价数据分析

H.5.1 总则

H.4 中包含3个因素:评价员、产品和轮次(或批次)。本节中一样,只是用时间因素替代了轮次因素。与产品因素相比,轮次因素是次要因素,但本节中时间因素与产品因素一样重要,因为产品因素会随时间而变化。表 H.7 是1个研究数据示例,10位评价员对4种牙膏的清新特性强度进行评分,分值范围是0~20分。数据的总和如表 H.8 所示。

表 H.7 4种牙膏清新特性强度随时间变化的4次评价

P1				P2				P3				P4			
T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
8	12	9	8	11	10	2	2	9	10	7	7	6	3	3	1
14	17	15	15	18	17	14	14	16	18	14	15	15	18	17	14
16	17	12	12	15	14	10	9	14	10	8	9	13	15	9	6
15	19	14	14	13	17	10	10	5	15	10	10	10	13	8	7
9	8	6	6	7	9	5	5	7	9	3	4	6	7	4	2
20	20	16	16	19	19	16	16	19	17	15	15	19	17	15	12
11	14	11	10	10	10	7	8	11	14	9	8	10	15	9	8
15	19	13	13	15	17	7	6	12	15	9	10	13	14	10	7
12	12	5	4	10	10	4	3	10	10	2	2	9	9	3	8
10	10	7	7	13	8	3	2	13	11	5	6	11	9	5	2

表 H.8 表 H.7 数据的总和

	时间 1	时间 2	时间 3	时间 4	
产品 1	130	148	108	105	491
产品 2	131	131	78	75	415
产品 3	116	129	82	86	413
产品 4	112	120	83	67	382
	489	528	351	333	1 701

H.5.2 所有评价员对 4 种产品均进行评价:产品、时间和评价员三因素交互

表 H.7 中,第 1 行给出了评价员 1 的数据,第二行给出评价员 2 的数据,以此类推。

H.5.2.1 三因素方差分析

结果如表 H.9 中所示,检验步骤同表 H.5,但结果展示与表 H.5 略有差异。产品因素(固定因素)是根据评价员 \times 产品的交互作用进行检验,时间因素(固定因素)是根据时间 \times 评价员的交互作用进行检验,而时间 \times 产品交互作用是根据试验误差(例如产品 \times 时间 \times 评价员的交互作用)进行检验。

表 H.9 表 H.7 数据的三因素方差分析

方差来源	因素 状态	自由度 (df)	平方和 (SS)	均方 (MS)	F 值	p 值
评价员	随机	$a-1=9$	2 056.18	$MS_a=228.46$	$MS_p/MS_{a \times p}=8.839$	0.000 3
产品	固定	$p-1=3$	161.22	$MS_p=53.74$		
评价员 \times 产品	随机	$(a-1) \times (p-1)=27$	164.09	$MS_{a \times p}=6.08$	$MS_t/MS_{t \times a}=38.56$	0.000 0
时间	固定	$(t-1)=3$	716.12	$MS_t=238.71$		
时间 \times 产品	固定	$(t-1) \times (p-1)=9$	45.21	$MS_{t \times p}=5.02$	$MS_{t \times p}/MS_e=2.43$	0.016 8
时间 \times 评价员	随机	$(t-1) \times (a-1)=27$	167.19	$MS_{t \times a}=6.19$		
误差		$(a-1) \times (p-1) \times (t-1)=81$	167.23	$MS_e=2.06$		
总计		159	3 477.24			

H.5.2.2 重复测量的多元方差分析

三因素方差分析需要满足方差-协方差假设(也称球形假设)。如果用时间因素(T1、T2、T3 和 T4)作为因变量进行多元方差分析(MANOVA),则这个假设的限制性较小。关于表 H.10 的解释,请查阅相关统计书籍。

表 H.10 表 H.7 数据的多元方差分析

	测试	值	<i>F</i> 值	<i>df</i>	误差 <i>df</i>	<i>p</i> 值
时 间	Wilk	0.087 10	87.345 27	3	25.000 00	0.000 000
	Pillai	0.912 90	87.345 27	3	25.000 00	0.000 000
	Hotelling	10.481 43	87.345 27	3	25.000 00	0.000 000
	Roy	10.481 43	87.345 27	3	25.000 00	0.000 000
时间×产品	Wilk	0.491 16	2.299 36	9	60.994 04	0.026 935
	Pillai	0.597 36	2.237 62	9	81.000 00	0.027 499
	Hotelling	0.855 90	2.250 70	9	71.000 00	0.028 152
	Roy	0.486 99	4.382 87	3	27.000 00	0.012 273
时间×评价员	Wilk	0.175 62	2.220 86	27	73.655 21	0.003 738
	Pillai	1.202 13	2.005 92	27	81.000 00	0.008 874
	Hotelling	2.698 18	2.365 07	27	71.000 00	0.002 065
	Roy	1.621 9	4.866 57	9	27.000 00	0.000 640

多元方差分析结果表明,4 项检验中,时间×产品的交互作用均显著。

注:也可创建一个新变量代表时间的变化量,然后执行另一个 MANOVA:差异₁=时间 2—时间 1,差异₂=时间 3—时间 2,差异₃=时间 4—时间 3。这些新变化或差异变量可用于衡量每个时间间隔内发生的变化(见参考文献[8])。

H.5.3 每位评价员只评价 1 种产品:评价员因素嵌套在与时间因素交互的产品因素中

这种情况下,表 H.7 数据是 40 位评价员的评价结果,每种产品均由 10 位评价员进行评价。
表 H.7 数据的方差分析结果如表 H.11 和表 H.12 所示。

表 H.11 表 H.7 数据的嵌套方差分析

方差来源	因素 状态	自由度 (<i>df</i>)	平方和 (<i>SS</i>)	均方 (<i>MS</i>)	<i>F</i> 值	<i>p</i> 值
产品	固定	$p-1=3$	161.22	$MS_p=53.74$	$MS_p/MS_{a \times p}=0.87$	0.47
评价员(产品)	随机	$(a-1) \times p=36$	2 220.28	$MS_{a \times p}=61.67$		
时间	固定	$(t-1)=3$	716.12	$MS_t=238.71$	$MS_t/MS_e=77.00$	0.000 00
时间×产品	固定	$(t-1) \times (p-1)=9$	45.21	$MS_{t \times p}=5.02$	$MS_{t \times p}/MS_e=1.63$	
误差		108	334.43	$MS_e=3.10$		

表 H.6 中,批次因素是嵌套在产品因素中的一个固定因素,而本节中嵌套在产品因素中的评价员因素是一个随机因素,时间因素和时间×产品的交互作用的检验并不相同。

表 H.12 评价员嵌套在产品中时表 H.8 数据的多元方差分析

	检验	数值	F 值	df	误差 df	p 值
时间	Wilk	0.149 451	64.499 77	3	34.000 0	0.000 000
	Pillai	0.850 549	64.499 77	3	34.000 0	0.000 000
	Hotelling	5.691 156	64.499 77	3	34.000 0	0.000 000
	Roy	5.691 156	64.499 77	3	34.000 0	0.000 000
时间×产品	Wilk	0.636 200	1.880 93	9	82.897 7	0.066 100
	Pillai	0.401 348	1.853 34	9	108.000 0	0.066 810
	Hotelling	0.512 838	1.861 41	9	98.000 0	0.066 763
	Roy	0.339 206	4.070 47	3	36.000 0	0.013 742
注：MANOVA 以时间因素(T1、T2、T3 和 T4)作为因变量。						

Roy 检验结果表明,时间×产品交互作用显著。其余 3 项检验 p 值均接近显著性水平值。

参 考 文 献

通用

- [1] ISO 704 Terminology work—Principles and methods
- [2] ISO 4121 Sensory analysis—Guidelines for the use of quantitative response scales
- [3] ISO 13300-1 Sensory analysis—General guidance for the staff of a sensory evaluation laboratory—Part 1: Staff responsibilities
- [4] ISO 13300-2 Sensory analysis—General guidance for the staff of a sensory evaluation laboratory—Part 2: Recruitment and training of panel leaders
- [5] GACULA Jr. M. C. 1997. Descriptive Sensory Analysis in Practice. Food & Nutrition Press, 712 pp.
- [6] LAWLESS H. T. 2013. Quantitative Sensory Analysis. Psychophysics, Models and Intelligent Design. Wiley & Sons, pp.
- [7] LAWLESS H. T., HEYMANN H. 2010. Sensory Evaluation of Food. Principles and Practices. 2nd ed., Springer, 596 pp.
- [8] MEILGAARD M. C., CIVILLE G. V., CARR B. T. 2007. Sensory Evaluation Techniques. 4th ed., CRC Press, 448 pp.
- [9] STONE H., BLEIBAUM R. N., THOMAS H. T. 2012. Sensory Evaluation Practices, 4th ed., Academic Press, 438 pp.

统计分析

- [10] ABDI H., EDELMAN B., VALENTIN D., DOWLING W. G. 2009. Experimental Design and Analysis for Psychology, Oxford University Press, 538 pp.6
- [11] GACULA M., SINGH J., BI J., ALTAN S. 2009. Statistical Methods in Food and Consumer Science, 2nd ed, pp.
- [12] HOWELL D. 2002. Statistical Methods for Psychology. Duxbury, 5th ed, pp.
- [13] NAES T., BROCKHOFF P. B., TOMIC O. 2010. Statistics for Sensory and Consumer Science, Wiley & Sons, 287 pp.

共识性感官剖面

- [14] BRANDT M. A., SKINNER E. Z., COLEMAN J. A. Texture profile method. J. Food Sci. 1963, 28(4): 404-409.
- [15] CAIRNCROSS S. E., SJOSTROM L. B. Flavor profiles-A new approach to flavor problems. Food Technol. 1950, 4, 308-311.
- [16] CHAMBERS D. H., ESTEVE E., RETIVEAU A. Effect of milk pasteurization on flavor properties of seven commercially available French cheese types. J. Sens. Stud. 2010, 25(4): 494-511.

对照参比描述法

- [17] LARSON-POWERS N. M., PANGBORN R. M. Descriptive analysis of the sensory properties of beverage s and gelatine containing sucrose and synthetic sweeteners. J. Food Sci. 1978, 43(11): 47-51.
- [18] BOYLSTON T. D., REITMEIER C., MOY J., MOSHER G. A., TALADRIZ L. Sensory quality and nutrient composition of three Hawaiian fruits treated by X-irradiation. J. Food Qual. 2002, 25(5): 419-433.

自选特性排序剖面

[19] SIEFFERMANN J.M. 2000. Le profil Flash. Un outil rapide et innovant d'évaluation sensorielle descriptive, AGORAL 2000, XIIèmes rencontres, Montpellier, France In: Tec. & Doc., L'innovation; de l'idée au succès, 335-340.

[20] DELARUE J., SIEFFERMANN J.M. Sensory mapping using Flash profile-Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. Food Qual. Prefer. 2004, 15(4): 383-392.

自由选择剖面

[21] WILLIAMS A.A., LANGRON S.P. The use of free-choice profiling for the evaluation of commercial ports. J. Sci. Food Agric. 1984, 35(5): 558-568.

[22] GUÀRDIA M.D., AGUIAR A.P.S., CLARET A., ARNAU J., GUERRERO L. Sensory characterization of dry-cured ham using free-choice profiling. Food Qual. Prefer. 2010, 21(1): 148-155.

[23] PIGGOT J.R., WATSON M.P. A comparison of free-choice profiling and the repertory grid method in the flavor profiling of cider. J. Sens. Stud. 1992, 7(1): 133-145.

定性感官剖面

[24] CAMPO E., BALLESTER J., LANGLOIS J., DACREMONT C., VALENTIN D. Comparison of conventional descriptive analysis and a citation frequency-based descriptive method for odor profiling: An application to Burgundy Pinot noir wines. Food Qual. Prefer. 2010, 21(1): 44-55.

定量描述剖面

[25] FINDLAY C.J., CASTURA J.C., LESSCHAEVE I. Feedback calibration: A training method for descriptive panels. Food Qual. Prefer. 2007, 18(2): 321-328.

[26] LAWLESS L.J.R., CIVILLE G.V. Developing Lexicons: Review A. J. Sens. Stud. 2013, 28(4): 270-281.

定量描述分析

[27] STONE H., SIDEL J., OLIVER R.S., WOOLSEY A., SINGLETON R.C. Sensory evaluation by quantitative descriptive analysis. Food Technol. 1974, 28(1): 24-34.

[28] SEE ALSO STONE et al., 2012, 250-274 (op cit.).

光谱

[29] CIVILLE G.V., DUS C.A. Development of terminology to describe the handfeel properties of paper and fabrics. J. Sens. Stud. 1990, 5(1): 19-32.

[30] MUNOZ A.M., CIVILLE G.V. Universal, product and attribute specific scaling and the development of common lexicons in descriptive analysis. J. Sens. Stud. 1998, 139(1): 57-75.

[31] SEE ALSO MEILGAARD et al., 2007, 189-253 (op cit).

TDS

[32] PINEAU N., CORDELLE S., IMBERT A., ROGEAUX M., SCHLICH P. 2003. Dominance temporelle des sensations—Codage et analyse d'un nouveau type de données sensorielles. XXXVèmes Journées de Statistiques, Lyon, 2-6 Juin, 777-780.

[33] PINEAU N., SCHLICH P., CORDELLE S., MATHONNIERE C., ISSANCHOU S., IMBERT A., ROGEAUX M., ETIEVANT P., KOSTER P. Temporal dominance of sensations. Construction of the TDS curves and comparison with time-intensity. Food Qual. Prefer. 2009, 20(6): 450-455.

中 华 人 民 共 和 国

国 家 标 准

感官分析 方法学

建立感官剖面的导则

GB/T 39625—2020/ISO 13299:2016

*

中国标准出版社出版发行
北京市朝阳区和平里西街甲2号(100029)
北京市西城区三里河北街16号(100045)

网址: www.spc.org.cn

服务热线: 400-168-0010

2020年12月第一版

*

书号: 155066 · 1-66567

版权专有 侵权必究



GB/T 39625-2020