

中华人民共和国国家标准

GB/T XXXX—XXXX/ISO 16820:2019

感官分析 方法学 序贯分析

Sensory analysis—Methodology—Sequential analysis

(ISO 16820: 2019, IDT)

(征求意见稿)

2025 - ** - **发布

2025 - ** - **实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

| | |
|------------------|----|
| 前 言 | 3 |
| 1 范围 | 4 |
| 2 规范性引用文件 | 4 |
| 3 术语、定义和符号 | 4 |
| 4 原则 | 5 |
| 5 程序 | 5 |
| 附录 A 示例 | 6 |
| 参考文献 | 10 |

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件等同采用 ISO 16820:2019《感官分析 方法学 序贯分析》。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国感官分析标准化技术委员会（SAC/TC566）提出并归口。

本文件起草单位：中国标准化研究院、江苏大学、大连工业大学、浙江工商大学、深圳雾芯科技有限公司、四川省绵阳市丰谷酒业有限责任公司、首都师范大学、渤海大学、中国农业科学院茶叶研究所、上海家化联合股份有限公司、陕西科仪阳光检测技术服务有限公司。

本标准主要起草人：钟葵、赵镭、张迪、邹小波、史波林、刘登勇、邹国华、汪厚银、董秀萍、王茜茜、田师一、张明成、卢中明、黄建桂、许勇泉、吴琼、王娜、项雅科、周滢、屠锦娣。

感官分析 方法学 序贯分析

1 范围

本文件描述了一种适用于强迫选择感官差别检验数据统计的方法，包括三点检验、两-三点检验、三点强迫选择法（3-AFC）和两点强迫选择法（2-AFC）。

本文件适用于感官分析中判定产品间以及产品成分、加工、包装、处理或贮存条件发生变化时是否存在可感知差异。本文件同时适用于评价员的筛选、培训与管理。

注：每次检验结束后，依据试验结果进行判定：终止测试并判定是否存在差异，或继续开展测试。相较于传统的感官差别检验方法（预先设定好检验次数），序贯分析法通常在更少的试验次数下得出判定结论。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；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）

注：GB/T 10221—2021 感官分析 术语（ISO 5492:2008，IDT）

3 术语、定义和符号

ISO 5492 界定的以及下列术语和定义适用于本文件。

ISO和IEC维护的用于标准化的术语数据库网址如下：

——ISO在线浏览平台：<https://www.iso.org/obp>；

——IEC电工百科：<http://www.electropedia.org/>。

3.1

α -风险 alpha-risk, α -risk

当不存在可感知差异时，判定存在可感知差异的概率。

注：也称为第 I 类错误、显著性水平或假阳性率。

3.2

β -风险 beta-risk, β -risk

当存在可感知差异时，判定不存在可感知差异的概率。

注：也称为第 II 类错误或假阴性率。

3.3

敏感性 sensitivity

用于评价检验性能特征的通用术语。

注：从统计学角度，检验敏感性由 α 、 β 和 p_d 值共同确定。

3.4

符号

p_0 ：当不存在可感知差异时，正确回答数的比例。

p_1 ：当存在可感知差异时，正确回答数的比例。

p_d ：能检验出两个产品间存在可感知差异的评价员的比例。

4 原则

选定检验类型（如三点检验、成对比较检验等），检验敏感性由所选定的 α 、 β 和 p_d 值确定。

根据 α 、 β 、 p_0 和 p_1 值计算决策区间界值。每次测试结束后，将总正确回答数与决策区间界值进行比较，以判定后续操作：

- 停止测试，并判定存在差异；
- 停止测试，并判定不存在差异；
- 继续开展测试。

5 程序

5.1 绘制附录图 A.1 所示的序贯分析图，根据 α 、 β 、 p_0 和 p_1 确定决策区间界值，具体步骤如下：

- a) 根据研究人员愿意接受的假阳性和假阴性结果风险，设定 α 和 β 值。其中， α 是当不存在可感知差异（正确回答数的比例为 p_0 ）时，判定为存在可感知差异的概率； β 是当存在可感知差异（正确回答数的比例为 p_1 ， $p_1 > p_0$ ）时，判定为不存在可感知差异的概率。
- b) p_0 是当不存在可感知差异时的正确回答数比例（即猜对比例）。 p_0 值取决于所采用的差别检验方法：
 - 三点检验和 3-AFC 检验： $p_0 = 1/3$ ；
 - 成对比较检验和 2-AFC 检验： $p_0 = 1/2$ 。
- c) p_1 是存在可感知差异时的正确回答数的比例。 p_1 值取决于 p_d 值：
 - 三点检验和 3-AFC 测试： $p_1 = p_d + (1 - p_d)/3$ ；
 - 成对比较检验和 2-AFC 测试： $p_1 = p_d + (1 - p_d)/2$ ；

采用瑟斯通 δ 法评价 2 个产品间感官差异大小时，可使用参考文献[7]中提供的转换表（从 δ 到 p_d 和从 p_d 到 δ ），根据采用的检验方法选择对应的 p_d 值。

d) 决策区间界值的计算公式如下：

$$\text{下限值: } d_0 = \frac{\lg(\beta) - \lg(1 - \alpha) - n \times \lg(1 - p_1) + n \times \lg(1 - p_0)}{\lg(p_1) - \lg(p_0) - \lg(1 - p_1) + \lg(1 - p_0)}$$

$$\text{上限值: } d_1 = \frac{\lg(1 - \beta) - \lg(\alpha) - n \times \lg(1 - p_1) + n \times \lg(1 - p_0)}{\lg(p_1) - \lg(p_0) - \lg(1 - p_1) + \lg(1 - p_0)}$$

5.2 每次差别检验完成后，将总正确回答数（y轴）与检验次数（x轴）绘制成图：

- 若总正确回答数位于图中上限线和下限线之间，则继续进行检验；
- 若总正确回答数位于上限线的上部区域，终止测试，判定在 α 显著性水平上存在可感知差异；
- 若总正确回答数位于下限线的下部区域，终止测试，判定不存在显著可感知差异（即当实际存在可感知差异时，正确回答数比例达到 p_1 的可能性小于 $(1 - \beta)$ ）。

附录 A
(资料性)
示例

A.1 示例 1——多组三点检验的序贯分析——小组中 2 名受训者的接受或拒绝测试

A.1.1 背景

某感官分析师拟通过三点检验评估 2 名受训者的表现，以决定其是否入选评价小组。每位受训者接受多组三点检验，组间留有充分的间隔时间，以避免感官疲劳。

A.1.2 试验设计

采用序贯分析（如图 A.1 所示）确定接受或拒绝受训者所需的试验次数。为确定决策区间界值（即图 A.1 中的两条线），需对 α 、 β 、 p_0 和 p_1 四个参数分别赋值。三点检验中， $p_0 = 1/3$ （即猜对概率，对应 $p_d = 0$ ）。通常将最低可接受的检测率设定为 $p_d = 50\%$ ，此时： $p_1 = 0.5 + (1 - 0.5) \times 1/3 = 2/3$ 。

若需减少确定决策所需的试验次数，可适当降低最低可接受的检测率，如取 $p_d = 40\%$ ，此时： $p_1 = 0.4 + (1 - 0.4) \times 1/3 = 0.60$ 。

注：本示例中， p_d 不是评价员群体中具备区分能力的人员比例，而是单个评价员实际能区分的样品比例。

感官分析师最终选用的参数如下：

—— $\alpha = 0.05$ ，接受不合格受训者的概率；

—— $\beta = 0.10$ ，拒绝合格受训者的概率；

—— $p_0 = 1/3$ ，最大不可接受的水平（即三点检验中原假设对应的 p 值）；

—— $p_1 = 2/3$ ，最小可接受的水平（即 $p_d = 0.50$ 时，检出存在差异样本的概率）。

A.1.3 结果分析与解释

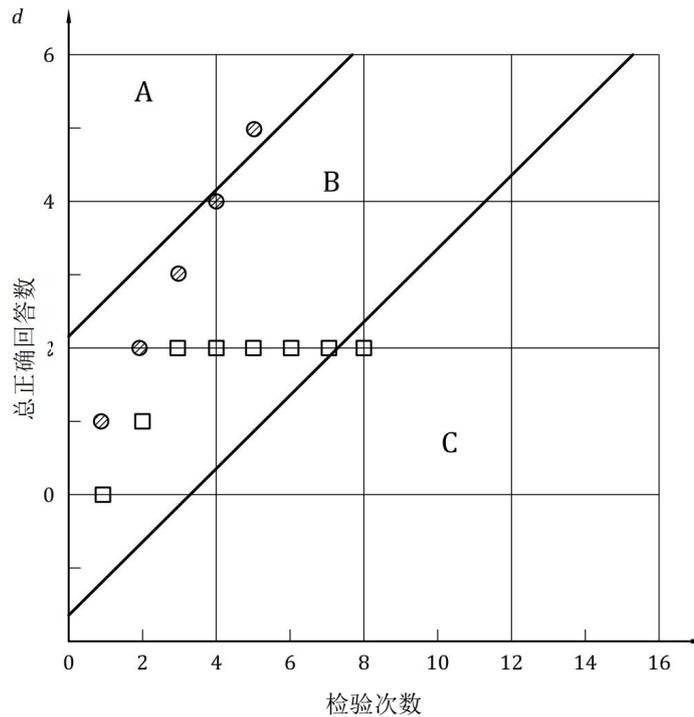
依据感官分析师设定的 α 、 β 、 p_0 和 p_1 参数值，采用 5.1 中的公式计算决策区间界值，决策区间下限值 (d_0) 和上限值 (d_1) 结果如下：

$$\text{下限值: } d_0 = \frac{\lg(\beta) - \lg(1 - \alpha) - n \times \lg(1 - p_1) + n \times \lg(1 - p_0)}{\lg(p_1) - \lg(p_0) - \lg(1 - p_1) + \lg(1 - p_0)} = \frac{\lg(0.1) - \lg(1 - 0.05) - n \times \lg(1 - 2/3) + n \times \lg(1 - 1/3)}{\lg(2/3) - \lg(1/3) - \lg(1 - 2/3) + \lg(1 - 1/3)} = -1.624 + 0.5n$$

$$\text{上限值: } d_1 = \frac{\lg(1 - \beta) - \lg(\alpha) - n \times \lg(1 - p_1) + n \times \lg(1 - p_0)}{\lg(p_1) - \lg(p_0) - \lg(1 - p_1) + \lg(1 - p_0)} = \frac{\lg(1 - 0.1) - \lg(0.05) - n \times \lg(1 - 2/3) + n \times \lg(1 - 1/3)}{\lg(2/3) - \lg(1/3) - \lg(1 - 2/3) + \lg(1 - 1/3)} = 2.085 + 0.5n$$

每次三点检验完成后，将结果记录于图 A.1 中。输入首次检验结果时，若回答正确，则输入点坐标为 $(x, y) = (1, 1)$ ；若回答错误，则输入点坐标为 $(x, y) = (1, 0)$ 。此后每完成一次检验，输入点坐标时， x 值每次增加 1；若回答正确， y 值相应增加 1；若回答错误，则 y 值保持不变。持续进行检验，直至所绘制的点触及或越过任一决策边界，并据此得出相应结论（即接受或拒绝该受训者）。

受训者 A 在所有检验中均回答正确，开展 5 次检验后即被接受。受训者 B 在首次三点检验中回答错误，第 2 次和第 3 次均回答正确，但随后每次检验均答错，在第 8 次检验后被拒绝。



标引序号说明:

- ⊙ 受训者 A;
- 受训者 B;
- n 检验次数;
- d 正确回答数;
- A 接受区;
- B 继续测试区;
- C 拒绝区。

注: 第 5 次检验后, 受训者 A 越过“继续测试区”, 进入“接受区”。第 8 次检验后, 受训者 B 越过“继续测试区”, 落入“拒绝区”。

图 A.1 三点检验中的序贯分析——示例 1: 两位受训者的选择

A.2 示例 2——多组两-三点检验的序贯分析——贮存期间牛肉饼热异味检测

A.2.1 背景

企业质量控制团队检测到冷藏 5 d 的牛肉饼经复热后产生了热异味。项目负责人拟设定一个该产品的合理冷藏保质期限。

A.2.2 实验设计

初步评价结果表明, 冷藏 5 d 的牛肉饼复热后有强烈热异味, 而冷藏 1 d 的牛肉饼则未检出热异味。感官分析师采用多组两-三点检验, 分别对冷藏 1 d、3 d 和 5 d 的牛肉饼样品与未经冷藏的新鲜烤制牛肉饼进行一一对比评价。

三组样品对 (新鲜对照样与冷藏 1 d 样品、新鲜对照样与冷藏 3 d 样品、新鲜对照样与冷藏 5 d 样

品) 分别独立开展两-三点检验, 三组试验均以新鲜烤制样品作为参比样。每位评价员完成评价后, 将结果与之前的正确回答数进行累加, 绘制总正确回答数图(图 A.2)。测试持续进行, 直至各贮存样品被判定为与新鲜烤制对照样相似或存在差异为止。

分析师选取的参数值如下:

- $\alpha = 0.10$, 为接受不合格牛肉饼的概率;
- $\beta = 0.10$, 为拒绝合格牛肉饼的概率;
- $p_0 = 0.50$, 为不存在可感知差异时正确回答数的比例(即两-三点检验的原假设概率值);
- $p_1 = 0.70$, 为 $p_d = 0.40$ 时, 判定贮存样品与新鲜对照样存在可感知差异的概率(即 $p_1 = 0.40 + (1 - 0.40) \times 0.50 = 0.70$)。

A.2.3 结果分析与解释

依据感官分析师设定的 α 、 β 、 p_0 和 p_1 参数值, 采用 5.1 中的公式计算决策区间上下界值, 决策区域下限值(d_0)和上限值(d_1)结果如下:

$$\text{下限值: } d_0 = \frac{\lg(\beta) - \lg(1-\alpha) - n \times \lg(1-p_1) + n \times \lg(1-p_0)}{\lg(p_1) - \lg(p_0) - \lg(1-p_1) + \lg(1-p_0)} = \frac{\lg(0.10) - \lg(1-0.10) - n \times \lg(1-0.70) + n \times \lg(1-0.50)}{\lg 0.70 - \lg 0.50 - \lg(1-0.70) + \lg(1-0.50)} = -2.59 + 0.6n$$

$$\text{上限值: } d_1 = \frac{\lg(1-\beta) - \lg(\alpha) - n \times \lg(1-p_1) + n \times \lg(1-p_0)}{\lg(p_1) - \lg(p_0) - \lg(1-p_1) + \lg(1-p_0)} = \frac{\lg(1-0.10) - \lg(0.10) - n \times \lg(1-0.70) + n \times \lg(1-0.50)}{\lg 0.70 - \lg 0.50 - \lg(1-0.70) + \lg(1-0.50)} = 2.59 + 0.6n$$

将测试结果绘制成图 A.2, 同时列出三组样品测试的总正确回答数(表 A.1)。

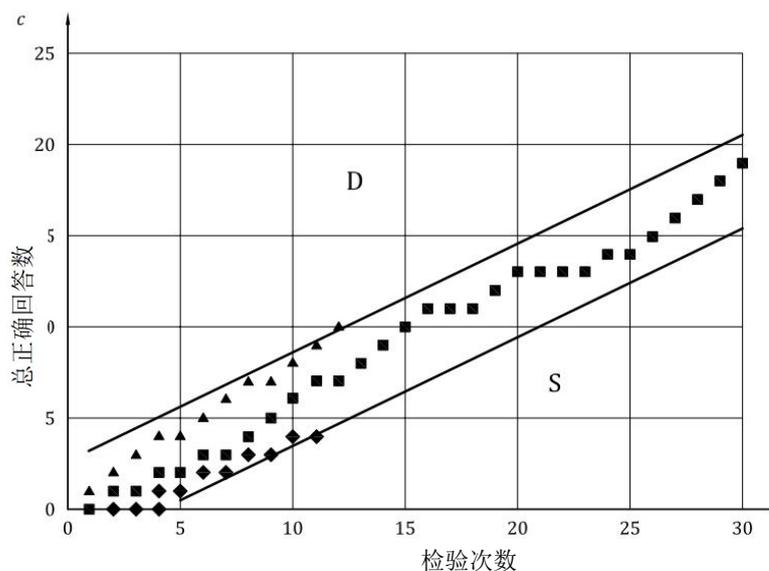
表 A.1 示例 2 检验结果

| 评价员 | 检验 A (冷藏 1 天) | | 检验 B (冷藏 3 天) | | 检验 C (冷藏 5 天) | |
|-----|---------------|--------|---------------|--------|---------------|--------|
| | 结果 | 总正确回答数 | 结果 | 总正确回答数 | 结果 | 总正确回答数 |
| 1 | I | 0 | I | 0 | C | 1 |
| 2 | I | 0 | C | 1 | C | 2 |
| 3 | I | 0 | I | 1 | C | 3 |
| 4 | C | 1 | C | 2 | C | 4 |
| 5 | I | 1 | I | 2 | I | 4 |
| 6 | C | 2 | C | 3 | C | 5 |
| 7 | I | 2 | I | 3 | C | 6 |
| 8 | C | 3 | C | 4 | C | 7 |
| 9 | I | 3 | C | 5 | I | 7 |
| 10 | C | 4 | C | 6 | C | 8 |
| 11 | I | 4 | C | 7 | C | 9 |
| 12 | | | I | 7 | C | 10 |
| 13 | | | C | 8 | | |
| 14 | | | C | 9 | | |
| 15 | | | C | 10 | | |
| 16 | | | C | 11 | | |
| 17 | | | I | 11 | | |
| 18 | | | I | 11 | | |
| 19 | | | C | 12 | | |
| 20 | | | C | 13 | | |

| | | | | | | |
|----|--|--|---|----|--|--|
| 21 | | | I | 13 | | |
| 22 | | | I | 13 | | |
| 23 | | | I | 13 | | |
| 24 | | | C | 14 | | |
| 25 | | | I | 14 | | |
| 26 | | | C | 15 | | |
| 27 | | | C | 16 | | |
| 28 | | | C | 17 | | |
| 29 | | | C | 18 | | |
| 30 | | | C | 19 | | |

注：I=不正确回答；C=正确回答。

贮存 1 d 的样品经 11 次成对比较检验后，判定与新鲜烤制对照样相似（不存在可感知差异）。贮存 5 d 的样品经 12 次检验后，判定与对照样品存在可感知差异。贮存 3 d 的样品经 30 次检验后，不能判定与对照样相似或存在可感知差异（见图 A.2）。感官分析师将上述结果反馈给项目负责人，并建议暂将 3 d 作为贮存期暂定指标，待后续继续试验以获得明确结论后，再确定贮存期限。



标引序号说明：

◆ 贮存 1 d 的样品；

■ 贮存 3 d 的样品；

▲ 贮存 5 d 的样品；

n 检验次数；

c 总正确回答数；

D 样品存在差异区域；

S 样品相似区域；

注：贮存 1 d 的样品与对照样在 11 次检验后判定为相似，贮存 5 d 的样品与对照样在 12 次检验后判定存在差异。存放 3 d 的样品在 30 次检验后判定处于“继续检验区”，因此在现有试验条件下尚未有明确结论，如需判定需继续增加检验次数。

图 A.2 两-三点检验中的序贯分析—示例 2：贮存期间牛肉饼的热异味

参 考 文 献

- [1] ISO 3534-1, Statistics—Vocabulary and symbols—Part 1: Probability and general statistical terms
 - [2] ISO 4120, Sensory analysis—Methodology—Triangular test
 - [3] ISO 5495, Sensory analysis—Methodology—Paired comparison test
 - [4] ISO 6658, Sensory analysis—Methodology—General guidance
 - [5] ISO 8586, Sensory analysis—General guidelines for the selection, training and monitoring of selected assessors and expert sensory assessors
 - [6] ISO 10399, Sensory analysis—Methodology—Duo-trio test
 - [7] Jesionka V., Rousseau B., Ennis J.M. Transitioning from proportion of discriminators to a more meaningful measure of sensory difference. *Food Quality and Preference*. 2014, 32 *pp.* 77–82.
-