

ICS 67.240

XX XX



中华人民共和国国家标准

GB/T XXXX—202X/ISO 29842:2011

感官分析 方法学 平衡不完全区组设计

Sensory analysis - Methodology –Balanced incomplete block designs

(ISO 29842: 2011, IDT)

(征求意见稿)

202X- - 发布

202X - - 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	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 平衡不完全区组设计原理	1
5 数据分析	3
5.1 总则	3
5.2 评分数据的方差分析	3
5.3 顺序数据的 Friedman 秩和分析	5
6 在感官评价中的应用	6
附录 A（资料性附录）不完全区组设计目录	7
附录 B（资料性附录）针对评分数据的平衡不完全区组设计示例	15
附录 C（资料性附录）针对顺序数据的平衡不完全区组设计示例	17
参考文献	19

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 29842: 2011 《感官分析 方法学 平衡不完全区组设计》。

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 10221—2012 感官分析 术语 (ISO 5492: 2008, MOD)

——GB/T 3358.1—2009 统计学词汇及符号 第 1 部分：一般统计术语与用于概率的术语 (ISO 3534-1: 2006, IDT)

本标准与 ISO 29842: 2011 相比，订正了原文的错误，修正了原本中概念表述不够准确的部分，主要变化如下：

——将 3.2 “重复 (repetition)” 的定义，与我国已颁布的等同采用 ISO 3534-3 的 GB/T 3358.3—2009 《统计学词汇及符号 第 3 部分：实验设计》中的术语相统一。

——在 4 “平衡不完全区组设计原理” 中将 “ λ ” 的定义改为 “每个样品对被评价的次数”。

——在表 2 中将 “均方 (MS)” 这列中的内容分别改为 “ $MS_S = SS/v_S$ ” 和 “ $MS_E = SE/v_E$ ”。

与 ISO 29842: 2011 相比，本标准作了必要的编辑性修改，例如：

——4 “平衡不完全区组设计原理” 中第二段内容在该节倒数第三段又重复出现，为了本节内容简洁明了，合并此两段内容。

本标准修正了 ISO 29842: 2011 纳入国际标准修正案的内容，如下：

——将附录 A 中 “y)” 这行内容改为 “6 of 9 BIB: 9 个样品中的 12 个六元组 ($t = 9, k = 6, b = 12, r = 8, \lambda = 5$)。”

——将 “表 A.8 9 个样品中的 18 个六元组” 改为 “表 A.8 9 个样品中的 12 个六元组”。

——将 “表 A.12 的第 13 行” 改为以下内容：

13		×		×	×	×			×	×
----	--	---	--	---	---	---	--	--	---	---

本标准由全国感官分析标准化技术委员会 (SAC/TC566) 提出并归口。

本标准起草单位：

本标准主要起草人：

感官分析 方法学 平衡不完全区组设计

1 范围

本标准规定了在感官描述分析和喜好测试中应用平衡不完全区组设计的方法。

本标准适用于在一个独立实验中被测试样品数超过评价员能给出可靠结果的样品评价数量。

本标准还指出了平衡不完全区组设计的基本特征，及其在感官评价中的应用导则。

2 规范性引用文件

以下文件对于本标准的应用必不可少。凡是注明日期的引用文件，仅注明日期的版本适用于本标准。凡是未注明日期的引用文件，其最新版本（包括所有的修改单）适用于本标准。

ISO 3534-1 统计学词汇及符号 第1部分：一般统计术语与用于概率的术语（Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability）

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

3 术语和定义

ISO 5492、ISO 3534-1以及以下术语和定义适用于本标准。

3.1

区组设计 block design

在多样品感官分析研究中，评价员评价所有或部分样品的实验方案。

3.2

重复 repetition

对给定的处理实施多于一次的实验。

4 平衡不完全区组设计原理

平衡不完全区组（BIB）设计适用于感官评价中样品总数大于在感觉和心理未疲劳时所能评价的样品数。在平衡不完全区组设计中，每位评价员的一次独立实验仅评价总样品集的部分样品子集。平衡不完全区组设计例子如表1所示。

表1 具有5个样品和10个区组/评价员的BIB设计

区组 (评价员)	测试样品				
	1	2	3	4	5
1	×	×	×	—	—
2	×	×	—	×	—
3	×	×	—	—	×
4	×	—	×	×	—
5	×	—	×	—	×
6	×	—	—	×	×
7	—	×	×	×	—
8	—	×	×	—	×
9	—	×	—	×	×
10	—	—	×	×	×

BIB设计中最常用的符号如下。

t 测试样品总量；

k 评价员在一次评价中评价的样品数量 ($k < t$)；

b 单次重复BIB设计中的区组 (通常是评价员) 总数；

r 单次重复BIB设计中每个测试样品被评价次数；

λ 每个样品对被评价的次数；

p 基本BIB设计的重复次数。

在基本 (单次重复) BIB设计中, 每个评价员会评价 t 个样品中的 k 个样品 ($k < t$)。选择评价员评价的 k 个子集样品, 使得在单次重复BIB设计中, 每个样品被评价的次数相同, 并且每一样品对被评价的次数也相同。完成单次重复BIB设计所需的区组 (评价员) 数量由 b 表示, 每个样品最终被评价的次数由 r 表示, 每个样品对被评价的次数由 λ 表示。

整个BIB设计可能需要重复几次才能达到研究所需的足够精确度。基本BIB设计的重复次数用 p 表示。那么, 区组 (通常是评价员) 的总数是 $p \times b$, 每个样品的评价总数是 $p \times r$ 。每对样品被一起观察的总次数是 $p \times \lambda$ 。

BIB设计中关系所有样品的 r 和 λ 的常数值, 对于在本设计中所收集的评价数据具有重要的统计意义。常数 r 体现了评估所有样品平均值的精度。常数 λ 体现了所有的成对比较任意样品对的灵敏度。

5 数据分析

5.1 总则

采用平衡不完全区组设计能够获得两种类型的数据。当评价员使用标度对正在评价的属性或印象进行感知强度评分时，可获得评分数据或得分值。当评价员根据正在评价的属性由最低到最高（或反之亦然）对样品进行排序时，可获得顺序数据。对于评分数据和顺序数据采用不同的数据分析方法。

5.2 评分数据的方差分析

方差分析（ANOVA）用于分析从BIB设计获得的评分数据。BIB设计的方差分析模型中，其差异来源与随机（完全）区组设计中的差异来源相同。在这两种情况下，总差异被划分为区组（通常是评价员）、处理（通常是样品）和误差的单独影响。

因为每位评价员仅评价测试样品总数的子集，所以需要比随机（完全）区组设计更复杂的公式来计算BIB设计的ANOVA平方和。感官分析师应确保用于执行分析的程序能够处理BIB设计。在许多计算机统计软件包中，ANOVA程序仅适用于完全设计，即每位评价员评价所有测试样品的研究。对于不完全的设计，例如BIB设计，需要广义线性模型（GLM）程序或混合模型程序。

用于分析BIB数据的ANOVA形式取决于该设计的实施方式。

当实验形式为表1中示例的单个重复设计方式，其方差分析表如表2所示。

表2 平衡不完全区组设计ANOVA表 (单次重复)

差异来源	自由度 (DF)	平方和 (SS)	均方 (MS)	<i>F</i>
总和	$v_T = t \times r - 1$	S_T		
评价员	$v_B = b - 1$	S_B		
样品 (根据评价员进行调整)	$v_S = t - 1$	S_S	$MS_S = S_S / v_S$	MS_S / MS_E
误差	$v_E = t \times r - t - b + 1$	S_E	$MS_E = S_E / v_E$	

如果表2中的*F*统计量超过在相应自由度下的上 α 临界值对应的*F*值，则平均评分相同的零假设不成立。如果*F*统计量是显著的，则应采用多重比较程序，例如费希尔LSD（最小显著性差异，*L*），以确定哪些样品间存在显著差异。适用于此单次重复的BIB设计的费希尔LSD（*L*）公式如下：

$$L = t_{\alpha/2, v_E} \sqrt{\frac{2MS_E}{r}} \sqrt{\frac{k(t-1)}{(k-1)t}}$$

其中

t, *k*和*r*如第4条所定义；

MS_E 是ANOVA表中的误差均方；

ν_E 是ANOVA表中的误差自由度；

$t_{\alpha/2, \nu_E}$ 是具有 ν_E 自由度的 t 分布的上 $\alpha/2$ 临界值。

在 F 统计量和费希尔LSD的显著性评价中，应使用相同的 α 值。

为达到研究所需要的足够精度，需重复 p 次BIB设计。如果分配给每位评价员的区组总数太多，以至于无法完成全部评价，则 $p \times b$ 位评价员中的每一位评价员仅需要评价一个含有 k 个样品的区组。在每个区组内， k 个样品的评价顺序应随机进行。该设计的ANOVA表见表3。

表3 平衡不完全区组设计ANOVA表

(由 $p \times b$ 位评价员中每一位对含有 k 个样品的单个区组进行评价而构成的 p 次重复实验)

差异来源	自由度 (DF)	平方和 (SS)	均方 (MS)	F
总和	$\nu_T = t \times p \times r - 1$	S_T		
区组 (评价员)	$\nu_B = p \times b - 1$	S_B		
样品 (根据评价员进行调整)	$\nu_S = t - 1$	S_S	$MS_S = S_S / \nu_S$	MS_S / MS_E
误差	$\nu_E = t \times p \times r - t - p \times b + 1$	S_E	$MS_E = S_E / \nu_E$	

如果表3中的 F 统计量超过在相应自由度的 F 临界值，则平均评分相同的零假设不成立。如果 F 统计量是显著的，则应采用多重比较程序，例如费希尔 LSD (L)，以确定哪些样品间存在显著差异。适用于本BIB设计的费希尔 LSD (L) 公式如下：

$$L = t_{\alpha/2, \nu_E} \sqrt{\frac{2MS_E}{pr}} \sqrt{\frac{k(t-1)}{(k-1)t}}$$

其中

t , k 和 r 如第4条所定义；

MS_E 是ANOVA表中的误差均方；

ν_E 是ANOVA表中的误差自由度；

$t_{\alpha/2, \nu_E}$ 是具有 ν_E 自由度的 t 分布的上 $\alpha/2$ 临界值。

在 F 统计量和费希尔LSD的显著性评价中，应使用相同的 α 值。

如果每个评价员评价BIB设计中的所有 b 个区组，那么“评价员效应”和“评价员与样品间”交互作用可以从总差异中划分出来（参见表4）。当基本BIB设计的区组总数很小（例如， $b \leq 6$ ）时，这种方法尤其适用。区组以随机顺序呈现给评价员。在每个区组内，评价样品的顺序也应随机进行。在任何一种方式下，都考虑了来自评价员效应的差异，评价员与样品间的交互作用也取代了表2和表3中的误差项。

表4 平衡不完全区组设计ANOVA表

(由 p 位评价员中每一位对含有 k 个样品的 b 个区组进行评价而构成的 p 次重复评价)

差异来源	自由度 (DF)	平方和 (SS)	均方 (MS)	F
总和	$v_T = t \times p \times r - 1$	S_T		
评价员	$v_P = p - 1$	S_P		
区组 (每位评价员内)	$v_{B(P)} = p \times (b - 1)$	$S_{B(P)}$		
样品 (根据评价员调整)	$v_S = t - 1$	S_S	$MS_S = S_S / v_S$	$MS_S / MS_{A \times S}$
评价员×样品	$v_{A \times S} = (p - 1)(t - 1)$	$S_{A \times S}$	$MS_{A \times S} = S_{A \times S} / v_{A \times S}$	
残差	$v_E = p \times (t \times r - t - b + 1)$	S_E	$MS_E = S_E / v_E$	

如果表4中的 F 统计量超过在相应自由度的 F 临界值, 则平均评分相同的零假设不成立。如果 F 统计量是显著的, 则应采用多重比较程序, 例如费希尔LSD (L), 以确定哪些样品间存在显著差异。适用于本BIB设计的费希尔 LSD (L) 公式如下:

$$L = t_{\alpha/2, v_{A \times S}} \sqrt{\frac{2MS_{A \times S}}{pr}} \sqrt{\frac{k(t-1)}{(k-1)t}}$$

其中:

t, k, p 和 r 如第4条所定义;

$MS_{A \times S}$ 是ANOVA表中评价员×样品交互作用的均方值;

$v_{A \times S}$ 是ANOVA表中评价员×样品交互作用的自由度;

$t_{\alpha/2, v_{A \times S}}$ 是具有 $v_{A \times S}$ 自由度的 t 分布的上 $\alpha/2$ 临界值。

在 F 统计量和费希尔LSD的显著性评价中, 应使用相同的 α 值。

5.3 顺序数据的Friedman秩和分析¹⁾

应使用Friedman统计量去分析BIB设计产生的顺序数据。Friedman检验统计量 F_{test} 的公式如下:

$$F_{test} = \frac{12}{p\lambda t(k+1)} \sum_{j=1}^t R_j^2 - \frac{3(k+1)pr^2}{\lambda}$$

其中:

t, k, r, λ 和 p 如第4条所定义;

R_j 是第 j 个样品的秩和 (参考文献[8])。

对于 $t=3 \dots 6, k=2 \dots 5$ 和 $p=1 \dots 7$ (参考文献[9])的选定组合, 可以使用 F_{test} 临界值表。但是, 在大多数感官研究中, 区组的总数超过了临界值表中的数值。对于这些情况, 如果 F_{test} 值超过具有 $(t-1)$ 自由度的 χ^2 分布的上 α 临界值, 则实验拒绝样品间无差异的假设。

1) 有多种统计方法可用于分析从BIB设计中获得的顺序数据。鼓励有兴趣的读者阅读有关该主题的统计文献。因Friedman方法的强大统计能力和计算便捷性, 这里对其进行了详细讨论。

如果 χ^2 统计量显著，则应执行多重比较程序以确定哪些样品间存在显著差异。适用于本BIB设计的费希尔 LSD (L) 公式如下：

$$L = z_{\alpha/2} \sqrt{\frac{p(k+1)(rk-r+\lambda)}{6}}$$

其中：

p, k, r 和 λ 如第4条所定义；

$z_{\alpha/2}$ 是标准正态分布的上 $\alpha/2$ 临界值。

在 F_{test} 统计量和费希尔LSD的显著性评价中，应使用相同的 α 值。

6 在感官评价中的应用

在一个独立实验中，评价员所评价的样品数 k 不得超过评价员能给出可靠评分结果的样品评价数量。 k 的值取决于若干因素，包括样品感官特征的总强度、残余效应程度（例如余味延迟），以及评价员所需要评价的属性个数。感官分析师应限制评价员在一次评价实验中所要评价的样品数量，以避免感官疲劳和情绪烦躁。

感官疲劳和情绪烦躁限制了评价员开展可靠评价的评价样品数量。但是，感官分析师应尽一切努力让每位评价员尽可能多地进行评价，以避免情境效应的影响。由于每位评价员仅评价样品总数的子集，因此评价员不能发现存在于所有样品间的差异。这可能会导致评价员放大对所评价样品的评分差异。评价员在一次评价中所评价的样品数比例越大，他们接触到的感官差异范围就越大，因此情境效应对研究结果所产生偏见的可能性就越小。

应根据测试所需的精度水平，来确定每个样品被评价的总次数 $r \times p$ 。应采用相同的标准来确定完全区组设计中的评价总数。例如，如果按照标准要求，在完全区组研究中需要使用描述性分析小组对每个样品进行12次评价，并且在单次重复BIB设计中每个测试样的重复评价次数为 $r = 3$ ，则应对BIB设计重复 $p = 4$ 次以达到所需的评价总数。

附录A

(资料性附录)

不完全区组设计目录

以下BIB设计适用于 $t = 3 \dots 10$ 和 $k = 2 \dots t - 1$ 或 $k = 2 \dots 6$ ，以较小者为准。在确定的 t 值下，通过与任何大小 k 的区组的所有可能组合构建BIB设计。然而，有时也可以构建比所有可能组合更少的区组的BIB设计。对于 t 和 k 的任何值，目录给出了区组数最小的BIB设计。

- a) 2 of 3 BIB: 3个样品中所有可能的二元组 ($t = 3, k = 2, b = 3, r = 2, \lambda = 1$)。
- b) 2 of 4 BIB: 4个样品中所有可能的二元组 ($t = 4, k = 2, b = 6, r = 3, \lambda = 1$)。
- c) 3 of 4 BIB: 4个样品中所有可能的三元组 ($t = 4, k = 3, b = 4, r = 3, \lambda = 2$)。
- d) 2 of 5 BIB: 5个样品中所有可能的二元组 ($t = 5, k = 2, b = 10, r = 4, \lambda = 1$)。
- e) 3 of 5 BIB: 5个样品中所有可能的三元组 ($t = 5, k = 3, b = 10, r = 6, \lambda = 3$)。
- f) 4 of 5 BIB: 5个样品中所有可能的四元组 ($t = 5, k = 4, b = 5, r = 4, \lambda = 3$)。
- g) 2 of 6 BIB: 6个样品中所有可能的二元组 ($t = 6, k = 2, b = 15, r = 5, \lambda = 1$)。
- h) 3 of 6 BIB: 6个样品中的10个三元组 ($t = 6, k = 3, b = 10, r = 5, \lambda = 2$)。

见表A.1。

表 A.1 6 个样品中的 10 个三元组

区组	样品					
	1	2	3	4	5	6
1	×	×			×	
2	×	×				×
3	×		×	×		
4	×		×			×
5	×			×	×	
6		×	×	×		
7		×	×		×	
8		×		×		×
9			×		×	×
10				×	×	×

- i) 4 of 6 BIB: 6个样品中所有可能的四元组 ($t = 6, k = 4, b = 15, r = 10, \lambda = 6$)。
- j) 5 of 6 BIB: 6个样品中所有可能的五元组 ($t = 6, k = 5, b = 6, r = 5, \lambda = 4$)。
- k) 2 of 7 BIB: 7个样品中所有可能的二元组 ($t = 7, k = 2, b = 21, r = 6, \lambda = 1$)。
- l) 3 of 7 BIB: 7个样品中的7个三元组 ($t = 7, k = 3, b = 7, r = 3, \lambda = 1$)。

见表A.2。

表 A.2 7 个样品中的 7 个三元组

区组	样品						
	1	2	3	4	5	6	7
1	×	×			×		
2	×		×			×	
3	×			×			×
4		×	×				×
5		×		×		×	
6			×	×	×		
7					×	×	×

m) 4 of 7 BIB: 7个样品中的14个四元组 (t = 7, k = 4, b = 14, r = 8, λ = 4)。

见表A.3。

表A.3 7个样品中的14个四元组

区组	样品						
	1	2	3	4	5	6	7
1	×	×	×		×		
2	×	×	×				×
3	×	×		×		×	
4	×	×				×	×
5	×		×	×	×		
6	×		×	×		×	
7	×			×	×		×
8	×				×	×	×
9		×	×	×			×
10		×	×		×	×	
11		×		×	×	×	
12		×		×	×		×
13			×	×		×	×
14			×		×	×	×

n) 5 of 7 BIB: 7个样品中所有可能的五元组 (t = 7, k = 5, b = 21, r = 15, λ = 10)。

o) 6 of 7 BIB: 7个样品中所有可能的六元组 (t = 7, k = 6, b = 7, r = 6, λ = 5)。

p) 2 of 8 BIB: 8个样品中所有可能的二元组 (t = 8, k = 2, b = 28, r = 7, λ = 1)。

q) 3 of 8 BIB: 8个样品中所有可能的三元组 (t = 8, k = 3, b = 56, r = 21, λ = 6)。

r) 4 of 8 BIB: 8个样品中的14个四元组 (t = 8, k = 4, b = 14, r = 7, λ = 3)。

见表A.4。

表A.4 8个样品中的14个四元组

区组	样品							
	1	2	3	4	5	6	7	8
1	×	×	×	×				
2	×	×			×	×		
3	×	×					×	×
4	×		×		×		×	
5	×		×			×		×
6	×			×	×			×
7	×			×		×	×	
8		×	×		×			×
9		×	×			×	×	
10		×		×	×		×	
11		×		×		×		×
12			×	×	×	×		
13			×	×			×	×
14					×	×	×	×

- s) 5 of 8 BIB: 8个样品中所有可能的五元组 ($t = 8, k = 5, b = 56, r = 35, \lambda = 20$)。
- t) 6 of 8 BIB: 8个样品中所有可能的六元组 ($t = 8, k = 6, b = 28, r = 21, \lambda = 15$)。
- u) 2 of 9 BIB: 9个样品中所有可能的二元组 ($t = 9, k = 2, b = 36, r = 8, \lambda = 1$)。
- v) 3 of 9 BIB: 9个样品中的12个三元组 ($t = 9, k = 3, b = 12, r = 4, \lambda = 1$)。

见表 A.5。

表 A.5 9 个样品中的 12 个三元组

区组	样品								
	1	2	3	4	5	6	7	8	9
1	×	×	×						
2	×			×			×		
3	×				×				×
4	×					×		×	
5		×		×					×
6		×			×			×	
7		×				×	×		
8			×	×				×	
9			×		×		×		
10			×			×			×
11				×	×	×			
12							×	×	×

w) 4 of 9 BIB: 9个样品中的18个四元组 ($t = 9, k = 4, b = 18, r = 8, \lambda = 3$)。

见表A.6。

表 A.6 9 个样品中的 18 个四元组

区组	样品								
	1	2	3	4	5	6	7	8	9
1	×	×	×	×					
2	×	×		×					×
3	×	×			×		×		
4	×		×			×		×	
5	×		×					×	×
6	×			×		×	×		
7	×				×	×			×
8	×				×		×	×	
9		×	×		×	×			
10		×	×			×	×		
11		×				×		×	×
12		×		×	×			×	
13		×					×	×	×
14			×	×	×			×	
15			×	×			×		×
16			×		×		×		×
17				×	×	×			×
18				×		×	×	×	

x) 5 of 9 BIB: 9个样品中的18个五元组 ($t = 9, k = 5, b = 18, r = 10, \lambda = 5$)。

见表A.7。

表A.7 9个样品中的18个五元组

区组	样品								
	1	2	3	4	5	6	7	8	9
1	×	×	×		×				×
2	×	×	×				×	×	
3	×	×		×		×		×	
4	×	×			×	×		×	
5	×	×				×	×		×
6	×		×	×	×	×			
7	×		×	×	×		×		
8	×		×			×	×		×

9	×			×	×			×	×
10	×			×			×	×	×
11		×	×	×		×			×
12		×	×	×			×	×	
13		×	×		×			×	×
14		×		×	×	×	×		
15		×		×	×		×		×
16			×	×		×		×	×
17			×		×	×	×	×	
18					×	×	×	×	×

y) 6 of 9 BIB: 9个样品中的12个六元组 ($t = 9, k = 6, b = 12, r = 8, \lambda = 5$)。

见表A.8。

表A.8 9个样品中的12个六元组

区组	样品								
	1	2	3	4	5	6	7	8	9
1	×	×		×	×		×	×	
2	×	×			×	×	×		×
3	×	×	×	×	×	×			
4	×	×	×				×	×	×
5	×	×		×		×		×	×
6	×		×	×	×			×	×
7	×		×	×		×	×		×
8	×		×		×	×	×	×	
9		×	×	×	×		×		×
10		×	×	×		×	×	×	
11		×	×		×	×		×	×
12				×	×	×	×	×	×

z) 2 of 10 BIB: 10个样品中所有可能的二元组 ($t = 10, k = 2, b = 45, r = 9, \lambda = 1$)。

aa) 3 of 10 BIB: 10个样品中的30个三元组 ($t = 10, k = 3, b = 30, r = 9, \lambda = 2$)。

见表 A.9。

表 A.9 10 个样品中的 30 个三元组

区组	样品									
	1	2	3	4	5	6	7	8	9	10
1	×	×	×							
2	×	×		×						
3	×		×		×					
4	×			×		×				

5	×				×		×			
6	×					×		×		
7	×						×		×	
8	×							×		×
9	×								×	×
10		×	×			×				
11		×		×						×
12		×			×			×		
13		×			×				×	
14		×				×	×			
15		×					×		×	
16		×						×		×
17			×	×			×			
18			×	×				×		
19			×		×	×				
20			×				×			×
21			×					×	×	
22			×						×	×
23				×	×				×	
24				×	×					×
25				×		×			×	
26				×			×	×		
27					×	×				×
28					×		×	×		
29						×	×			×
30						×		×	×	

bb) 4 of 10 BIB: 10个样品中的15个四元组 ($t = 10, k = 4, b = 15, r = 6, \lambda = 2$)。

见表 A.10。

表 A.10 10 个样品中的 15 个四元组

区组	样品									
	1	2	3	4	5	6	7	8	9	10
1	×	×	×	×						
2	×	×			×	×				
3	×		×				×	×		
4	×			×					×	×
5	×				×		×		×	
6	×					×		×		×
7		×	×			×			×	
8		×		×			×			×

9		×			×			×		×
10		×					×	×	×	
11			×	×	×			×		
12			×		×				×	×
13			×			×	×			×
14				×	×	×	×			
15				×		×		×	×	

cc) 5 of 10 BIB: 10个样品中的18个五元组 ($t = 10, k = 5, b = 18, r = 9, \lambda = 4$)。

见表A.11。

表A.11 10个样品中的18个五元组

区组	样品									
	1	2	3	4	5	6	7	8	9	10
1	×	×	×	×	×					
2	×	×	×			×	×			
3	×	×		×		×			×	
4	×	×			×		×	×		
5	×		×			×		×	×	
6	×		×				×	×		×
7	×			×	×	×				×
8	×			×				×	×	×
9	×				×		×		×	×
10		×	×	×				×		×
11		×	×		×				×	×
12		×		×			×	×	×	
13		×			×	×		×		×
14		×				×	×		×	×
15			×	×	×		×		×	
16			×	×		×	×			×
17			×		×	×		×	×	
18				×	×	×	×	×		

dd) 6 of 10 BIB: 10个样品中的15个六元组 ($t = 10, k = 6, b = 15, r = 9, \lambda = 5$)。

见表A.12。

表 A.12 10 个样品中的 15 个六元组

区组	样品									
	1	2	3	4	5	6	7	8	9	10
1	×	×	×		×		×			×
2	×	×	×					×	×	×
3	×	×		×	×			×	×	
4	×	×		×		×	×	×		
5	×	×				×	×		×	×
6	×		×	×	×	×				×
7	×		×	×		×	×		×	
8	×		×		×	×		×	×	
9	×			×	×		×	×		×
10		×	×	×	×		×		×	
11		×	×	×		×		×		×
12		×	×		×	×	×	×		
13		×		×	×	×			×	×
14			×	×			×	×	×	×
15					×	×	×	×	×	×

附录B

(资料性附录)

针对评分数据的平衡不完全区组设计示例

B.1 问题和状况

芥末工厂的质量控制经理定期筛选要添加到参考样品池中的成品样品。定期需要新的参考样品，因为旧的样品会随着时间而改变不再合适。该操作还用于从池中清除任何可能已经恶化的当前参考样品。

B.2 实验设计

评价来自六个批次样品的总体异味。由15名受过训练的评价员进行评价，使用10点类项标度，0表示无异味，9表示极端异味。评价员对六个样品中的四个进行评价。每个评价员评价的四个样品由表B.1中列出的BIB设计确定。15个评价员中的每一位随机地从设计中分配了四个样品的一个区组。每个区组内的样品的呈现顺序是随机的。

表 B.1 评分数据的 BIB 设计
 $t = 6, k = 4, r = 10, b = 15, \lambda = 6$

区组 (评价员)	样品					
	1	2	3	4	5	6
1	×	×	×	×		
2	×			×	×	×
3		×	×		×	×
4	×	×	×		×	
5	×	×		×		×
6			×	×	×	×
7	×	×	×			×
8	×		×	×	×	
9		×		×	×	×
10	×	×		×	×	
11	×		×		×	×
12		×	×	×		×
13	×	×			×	×
14	×		×	×		×
15		×	×	×	×	

B.3 结果

异味强度得分见表B.2。通过能够执行BIB设计ANOVA的程序来分析数据。所得的ANOVA表见

表B.3。样品的F统计量非常显著 ($p < 0.0001$)，表明样品中的异味强度存在明显差异。将LSD多重比较程序应用于样品的平均评分以确定哪些样品间具有显著不同的异味强度。表B.4中的结果表明样品1具有比所有其它样品显著更高的异味。其余样品之间没有显著差异。

表B.2 六种芥末的异味强度等级

评价员	样品					
	1	2	3	4	5	6
1	6	1	1	2		
2	6			1	3	3
3		4	2		5	2
4	7	2	3		2	
5	3	5		1		1
6			1	1	3	2
7	7	4	4			3
8	2		1	1	1	
9		2		2	2	3
10	4	2		2	5	
11	5		3		1	1
12		3	2	1		2
13	4	2			1	1
14	5		2	2		1
15		2	4	5	3	

表 B.3 平衡不完全区组 ANOVA 评分数据表：六种芥末的异味强度

差异来源	平方和(SS)	自由度(DF)	均方(MS)	F	p
总计	156.60	59			
评价员 (区组)	38.58	14			
样品 (处理, 按区组调整)	64.08	5	12.82	9.60	< 0.0001
误差	53.42	40	1.34		

表 B.4 六种芥末的异味强度的调整平均数

样品	1	2	3	4	5	6
调整平均数	5.0 A	2.5 B	2.2 B	2.0 B	2.6 B	1.9 B

注：没有共同字母的平均值在5%的显著性水平上有显著差异 ($L = 1.1$)。

附录C

(资料性附录)

针对顺序数据的平衡不完全区组设计示例

C.1 问题和状况

辣椒酱制造商想要评价15种辣椒的辣味。制造商希望对在其产品中使用的品种的数量进行限制。

C.2 实验设计

由于辣椒的辣度很高，因此选择了平衡不完全区组设计。每个评价员评价 $t = 15$ 个样品中的 $k = 3$ 个样品。基本BIB设计由 $b = 35$ 个区组组成，其中每个样品被评价 $r = 7$ 次，并且每个样品对会被评价一次 ($\lambda = 1$)。基本设计重复 $p = 3$ 次，以获得每个样品的总共 $p \times r = 21$ 个评价。

在招募中随机选择一组105名评价员来参加测试。通过区分不同辣度的能力进行评价员筛选，并接受了排序法使用的指导。由于缺乏对感官测试的广泛培训，评价员被要求将三个样品按照1 = 最辣；2 = 中等辣；3 = 最不辣进行排序。

C.3 结果分析

为了使结果更容易分析，研究中的顺序数据按表C.1所示。给定辣椒品种的排名总和是对应于该品种列中所有数字的总和。通过计算Friedman检验统计量 F_{test} 的值，来确定品种间的辣度是否存在差异。 $F_{test} = 68.53$ 的值超过了 χ^2 的上5%临界值($\chi^2_{14,0.05} = 23.69$)，其中自由度为 $(t - 1) = 14$ ，那么可得样品集确实存在显著差异。计算顺序数据的95%LSD多重比较统计量以确定哪个品种存在显著不同 ($L = 11$)。多重比较的结果见表C.2

表 C.1 在 BIB 示例中获得的顺序数据结果：辣椒的辣味

区组 (评价员)	样品或种类 ^a														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	2	3												
2				3			1			2					
3					1					3					2
4						3					2		1		
5							3		1					2	
6	3			2	1										
7		2						3		1					
8			3										2	1	
9						3			2						1
10							2				1	3			
...
101	2													3	1

102		1		3		2									
103			1					2			3				
104					1				2			3			
105							3			2			1		
Rank sum 排名总和	35	45	54	43	28	37	55	42	37	50	49	50	34	42	29

^a.反馈: 1 =最辣; 2 =中等辣; 3 =最不辣。

C.2 BIB 顺序数据结果和统计分析: 辣椒的辣味

样品或种类	秩和				
5	28	a			
15	29	a			
13	34	a	b		
1	35	a	b		
6	37	a	b		
9	37	a	b		
14	42		b	c	
8	42		b	c	
4	43		b	c	
2	45			c	d
11	49			c	d
10	50			c	d
12	50			c	d
3	54				d
7	55				d

注: 没有共同字母的秩和在5%显著性水平上显著不同 ($L_{rank} = 11$)。

参 考 文 献

- [1] GB/T 3358.3 统计学词汇及符号 第3部分：实验设计
 - [2] ISO 6564 Sensory analysis — Methodology — Flavour profile methods
 - [3] ISO 6658 Sensory analysis — Methodology — General guidance
 - [4] ISO 8586-1 Sensory analysis — General guidance for the selection, training and monitoring of assessors — Part 1: Selected assessors
 - [5] ISO 8586-2 Sensory analysis — General guidance for the selection, training and monitoring of assessors — Part 2: Expert sensory assessors
 - [6] ISO 8587 Sensory analysis — Methodology — Ranking
 - [7] ISO 11036 Sensory analysis — Methodology — Texture profile
 - [8] ISO 13299 Sensory analysis — Methodology — General guidance for establishing a sensory profile
 - [9] COCHRAN, W.G., COX, G.M. Experimental designs, 2nd edition. New York, NY: Wiley, 1957. p611.
 - [10] GACULA, M.C. JR, SINGH, J. Statistical methods in food and consumer research, 2nd edition. Boston, MA: Elsevier/Academic Press, 2009. p853.
-