

基于 R 语言的 PanelCheck 数据分析手册

版本 3.0

李嘉伟

PanelCheck 官网 <https://www.panelcheck.com>

目录

第一部分 前言	6
第 1 章 原始数据格式说明	6
1.1 英文模板	6
1.2 中文模板	6
第 2 章 R 包说明	7
第 3 章 成果总览	8
第 4 章 绘图说明	8
第 5 章 PanelCheck 分析流程	9
第二部分 单变量 (Univariate)	10
第 6 章 线图 (Line Plots)	10
6.1 单一样本 (Sample_line_plot_sp)	10
6.1.1 函数介绍	10
6.1.2 可视化分析	11
6.2 单一样本-评价者 (Sample_line_plot_sp_as)	12
6.2.1 函数介绍	12
6.2.2 可视化分析	13
6.3 单一样本-评价者-重复次序 (Sample_line_plot_sp_as_re)	14
6.3.1 函数介绍	14
6.3.2 可视化分析	16
6.4 总结	16
第 7 章 均值-标准差图 (Mean & STD Plots)	16
7.1 基于特定评估者 (Mean_STD_plot_ass)	17
7.1.1 函数介绍	17
7.1.2 可视化分析	18
7.2 基于特定感官属性 (Mean_STD_plot_attr)	18
7.2.1 函数介绍	18
7.2.2 可视化分析	19

第 8 章 相关关系图 (Correlation Plots)	20
8.1 相关关系图 (Corr_plot_as_attr)	20
8.1.1 函数介绍	20
8.1.2 可视化分析	22
第 9 章 轮廓图 (Profile Plots)	22
9.1 轮廓图 (Profile_plot_attr)	22
9.1.1 函数介绍	22
9.1.2 可视化分析	25
第 10 章 蛋壳图 (Eggshell Plots)	26
10.1 蛋壳图 (Eggshell_plot_attr)	26
10.1.1 函数介绍	26
10.1.2 可视化分析	29
第 11 章 F&P 图 (F & P Plots)	29
11.1 F 值图 (F_plot)	29
11.1.1 函数介绍	29
11.1.2 可视化分析	32
11.2 P 值图 (P_plot)	32
11.2.1 函数介绍	32
11.2.2 可视化分析	35
第 12 章 均方误差图 (MSE Plots)	35
12.1 均方误差图 (MSE_plot)	35
12.1.1 函数介绍	35
12.1.2 可视化分析	38
第 13 章 P 值-均方误差图 (P-MSE Plots)	38
13.1 基于给定评估者 (P_MSE_ass)	38
13.1.1 函数介绍	38
13.1.2 可视化分析	39
13.2 基于给定感官属性 (P_MSE_attr)	40
13.2.1 函数介绍	40
13.2.2 可视化分析	41
第三部分 多变量 (Multivariate)	42
第 14 章 Tucker-1 方法绘图 (Tucker-1 Plots)	42
14.1 共同得分 (Tucker1_common_score)	42
14.1.1 函数介绍	42
14.1.2 可视化分析	44
14.2 基于评估者的相关载荷图 (Tucker1_corrload_ass)	45
14.2.1 函数介绍	45
14.2.2 可视化分析	46
14.3 基于感官属性的相关载荷图 (Tucker1_corrload_attr)	46
14.3.1 函数介绍	46

14.3.2 可视化分析	48
第 15 章 曼哈顿图 (Manhattan Plots)	49
15.1 曼哈顿图 (Manhattan_plot)	49
15.1.1 函数介绍	49
15.1.2 可视化分析	50
第四部分 一致性 (Consensus)	51
第 16 章 评估者权重 (Assessor Weights)	51
16.1 评估者权重图 (Ass_weight_plot)	51
16.1.1 函数介绍	51
16.1.2 可视化分析	52
第 17 章 一致性数据 (Consensus data)	53
17.1 一致性数据 (Consen_data)	53
17.1.1 函数介绍	53
17.1.2 可视化分析	54
第 18 章 主成分得分 (PCA Scores)	54
18.1 主成分得分 (PCA_scores)	54
18.1.1 函数介绍	54
18.1.2 可视化分析	56
第 19 章 主成分载荷 (PCA Loadings)	57
19.1 主成分载荷 (PCA_loadings)	57
19.1.1 函数介绍	57
19.1.2 可视化分析	58
第 20 章 主成分相关载荷 (PCA Correlation Loadings)	59
20.1 主成分相关载荷 (PCA_cor_loading)	59
20.1.1 函数介绍	59
20.1.2 可视化分析	60
第 21 章 双标图 (Bi-Plot)	61
21.1 双标图 (Biplot)	61
21.1.1 函数介绍	61
21.1.2 可视化分析	62
第 22 章 主成分解释方差 (PCA Explained Variance)	63
22.1 主成分解释方差 (PCA_exp_var)	63
22.1.1 函数介绍	63
22.1.2 可视化分析	64
第 23 章 蜘蛛网图 (Spiderweb Plot)	64
23.1 蜘蛛网图 (Spiderweb_plot)	64
23.1.1 函数介绍	64
23.1.2 可视化分析	66

第五部分 全局分析 (Overall)	67
第 24 章 双因素方差分析-只评估 1 次 (2-way ANOVA(1rep))	67
24.1 评估者主效应 (ANOVA_2way_1rep_ass_effect)	67
24.1.1 函数介绍	67
24.1.2 可视化分析	68
24.2 产品主效应 (ANOVA_2way_1rep_pro_effect)	69
24.2.1 函数介绍	69
24.2.2 可视化分析	70
24.3 样本均值检验 (ANOVA2_1rep_Sample_mean_LSD)	70
24.3.1 函数介绍	70
24.3.2 可视化分析	71
第 25 章 双因素方差分析 (2-way ANOVA)	71
25.1 评估者主效应 (ANOVA_2way_ass_effect)	72
25.1.1 函数介绍	72
25.1.2 可视化分析	73
25.2 产品主效应 (ANOVA_2way_pro_effect)	73
25.2.1 函数介绍	73
25.2.2 可视化分析	74
25.3 评估者 * 产品交互效应 (ANOVA_2way_asspro_inter)	74
25.3.1 函数介绍	74
25.3.2 可视化分析	75
25.4 样本均值检验 (ANOVA2_Sample_mean_LSD)	76
25.4.1 函数介绍	76
25.4.2 可视化分析	77
第 26 章 三因素方差分析 (3-way ANOVA)	77
26.1 评估者主效应 (ANOVA_3way_ass_effect)	77
26.1.1 函数介绍	77
26.1.2 可视化分析	78
26.2 产品主效应 (ANOVA_3way_pro_effect)	78
26.2.1 函数介绍	78
26.2.2 可视化分析	79
26.3 重复评估主效应 (ANOVA_3way_rep_effect)	80
26.3.1 函数介绍	80
26.3.2 可视化分析	81
26.4 评估者 * 产品交互效应 (ANOVA_3way_asspro_inter)	81
26.4.1 函数介绍	81
26.4.2 可视化分析	82
26.5 评估者 * 重复评估交互效应 (ANOVA_3way_assrep_inter)	82
26.5.1 函数介绍	82
26.5.2 可视化分析	83
26.6 产品 * 重复评估交互效应 (ANOVA_3way_prorep_inter)	83
26.6.1 函数介绍	83
26.6.2 可视化分析	84
26.7 样本均值检验 (ANOVA3_Sample_mean_LSD)	85

26.7.1 函数介绍	85
26.7.2 可视化分析	86
参考文献	86

第一部分 前言

PanelCheck 软件是由挪威研究理事会以及工业和研究伙伴于 2005 年 1 月至 2008 年 12 月资助, Nofima 研究所和丹麦科技大学信息与数学建模实验室所持续开发的软件, 可用于进行食品感官分析。笔者基于 R 语言对该软件的所有统计方法以及图表进行复刻, 并在此基础上赋予一定的灵活性。如需下载该软件, 可点击[此处](#)进入 PanelCheck 官网下载。

第 1 章 原始数据格式说明

1.1 英文模板

同使用软件一样, 利用 R 语言进行分析对导入英文模板数据格式有一定要求:

数据要求

- 1、导入数据与同级文件夹下 sample.csv 中的数据格式保持一致
- 2、指定的三列列名为 Assessor, Sample, Replicate, 之后为各个感官属性得分, 每个感官属性为一列
- 3、指定三列的顺序最好与上述顺序一致。若 Replicate 只有 1 次, 该列可以没有。
- 4、导入数据格式为最好为 csv 数据格式, 如果是其他 Excel 格式可以导出为 csv 格式
- 5、如果是其他数据格式如 txt 格式等, 需要使用其他 R 函数如 read.table()。
- 6、原始数据中的感官属性评分打分范围默认为 1~10, 有些绘图函数需要指定 score 参数。

1.2 中文模板

中文模板与英文模板仅仅在于指定的 非感官属性列名 不一致:

数据要求

- 1、导入数据与同级文件夹下 `dlfxEvaluateTable.csv` 中的数据格式保持一致
- 2、指定的三列列名为 `序号`、`样品信息`、`次数`，之后为各个感官属性得分，每个感官属性为一列
- 3、指定三列的顺序最好与上述顺序一致。若 `次数` 只有 1 次，该列可以没有。
- 4、导入数据格式为最好为 `csv` 数据格式，如果是其他 Excel 格式可以导出为 csv 格式
- 5、如果是其他数据格式如 txt 格式等，需要使用其他 R 函数如 `read.table()`。
- 6、原始数据中的感官属性评分打分范围默认为 `1~10`，有些绘图函数需要指定 `score` 参数。

第 2 章 R 包说明

在本次 R 语言代码图表实现中，共使用了以下 R 包：

绘图

- LIB `ggplot2` —— 主要绘图系统
- LIB `ggrepel` —— 用于防止图形文本标签重叠
- LIB `ggsci` —— 用于 SCI 论文样式配色
- LIB `ggpubr` —— 用于 `ggplot` 绘图排版
- LIB `ggforce` —— 用于 `ggplot2` 扩展绘图
- LIB `fmsb` —— 用于绘制蜘蛛网图
- LIB `grid` —— 用于手动构建图例
- LIB `gtable` —— 用于手动构建图例

数据处理与分析

- LIB `reshape2` —— 数据整合与重构
- LIB `stringr` —— 字符串处理
- LIB `broom` —— 用于统计结果格式转换
- LIB `FactoMineR` —— 用于主成分分析

笔者写的函数中，都设定了 R 包检测程序，即如果运行该函数的 R 软件中没有所要求的 R 包，运行该函数时，程序会自动运行安装。

如果您的 R 软件版本过低，或与对应 R 包所要求的版本不匹配，自动安装程序可能会无效并且报错。在该情况下，笔者建议更新您的 R 软件，并且推荐手动安装。比如安装 `ggplot2` 时，只需运行以下代码：
`install.packages("ggplot2")`。

第 3 章 成果总览

在复刻 PanelCheck 软件绘图时，笔者一共写了 39 个函数，总代码量 6000 余行。除此之外，在汇总文件夹中，`Output_panelcheck()` 函数支持本地一次性输出所有基本图表（默认参数）；`Output_html_panelcheck()` 函数支持一次性输出网页格式图表（默认参数）。单独每个函数的参数最少只有 1 个，最多达到 12 个参数。不过读者不用担心函数数量过多或函数参数过多而影响应用：

1、因为虽然函数数量较多，但是每个函数的命名都与 PanelCheck 软件上命名保持一致，读者可以轻松找到自己所需要方法绘图的函数

2、每个函数的参数在每个函数的代码文件中都有解释，并且所有函数的参数设置都具有高度相似性，也就是只要搞清楚了几个函数的参数含义，基本上所有函数参数都可以理解。

接下来，笔者将按照 PanelCheck 方法分类，依次介绍每个统计方法的数据可视化函数使用以及图表理解方法。在此开始之前，我们执行以下代码导入原始数据，在没有特别说明的情况下，我们使用 `sample.csv` 数据进行分析示例。

```
1 # 导入原始数据
2 # 引号内部是本地文件路径
3 # 反斜杠 "\" 在 R 语言里是转义字符，所以在文件路径中使用斜杠 "/"
4 data <- read.csv("E:/R 语言应用/食品感官分析-PanelCheck/说明手册/sample.csv")
```

第 4 章 绘图说明

原始数据存在多种情形：

- `Replicate/次数` 评估 1 次与 `Replicate/次数` 评估多次
- `感官属性` 的数量也会对某些分析方法产生影响。

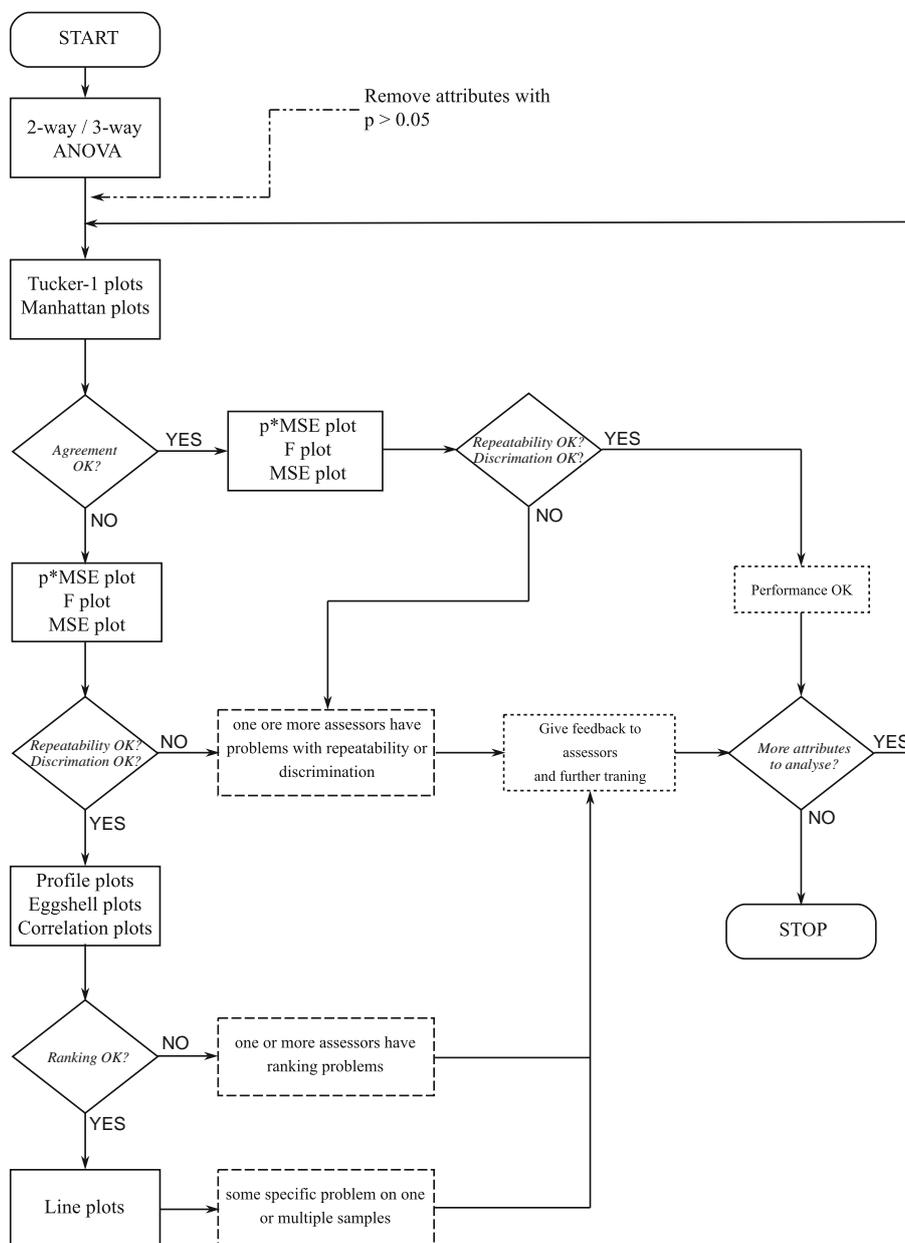
1、当评估一次时，`线图 (Line Plots)` 中的 `F plot`, `P Plot`, `MSE Plot`, `P*MSE Plot` 方法不能使用，原因在于这些绘图基于给定属性与评估者下的单因素方差分析，一次评估下无法进行方差分析；

2、当评估多次时，则不能使用 `One-way-ANOVA(1rep)`；

3、当原始数据中感官属性数量少于 3 个时，则不能使用 `主成分分析方法 (PCA)`，即无法使用 `Consensus 分析`；少于 2 个时，不能绘制 `曼哈顿图`。

第 5 章 PanelCheck 分析流程

以下为 PanelCheck 面板分析流程图：



第二部分 单变量 (Univariate)

在单变量分析中有多种数据可视化方法，笔者接下来将一一介绍。

第 6 章 线图 (Line Plots)

线图 (Line Plots) 的主要功能是让研究人员初步查看自己所得的实验数据，大致对数据有基本的了解，为后面利用统计方法分析作基础。线图包括了三种图形，分别是：

- 基于每个单独样本的感官属性、评估者与重复次序的线段散点图——单一样本 (`Sample_line_plot_sp`)
- 基于每个单独样本、评估者的感官属性与重复次序的线段散点图——单一样本-评价者 (`Sample_line_plot_sp_as`)
- 基于每个单独样本、评估者、重复次序的感官属性线段散点图——单一样本-评价者-重复次序 (`Sample_line_plot_sp_as_re`)

6.1 单一样本 (`Sample_line_plot_sp`)

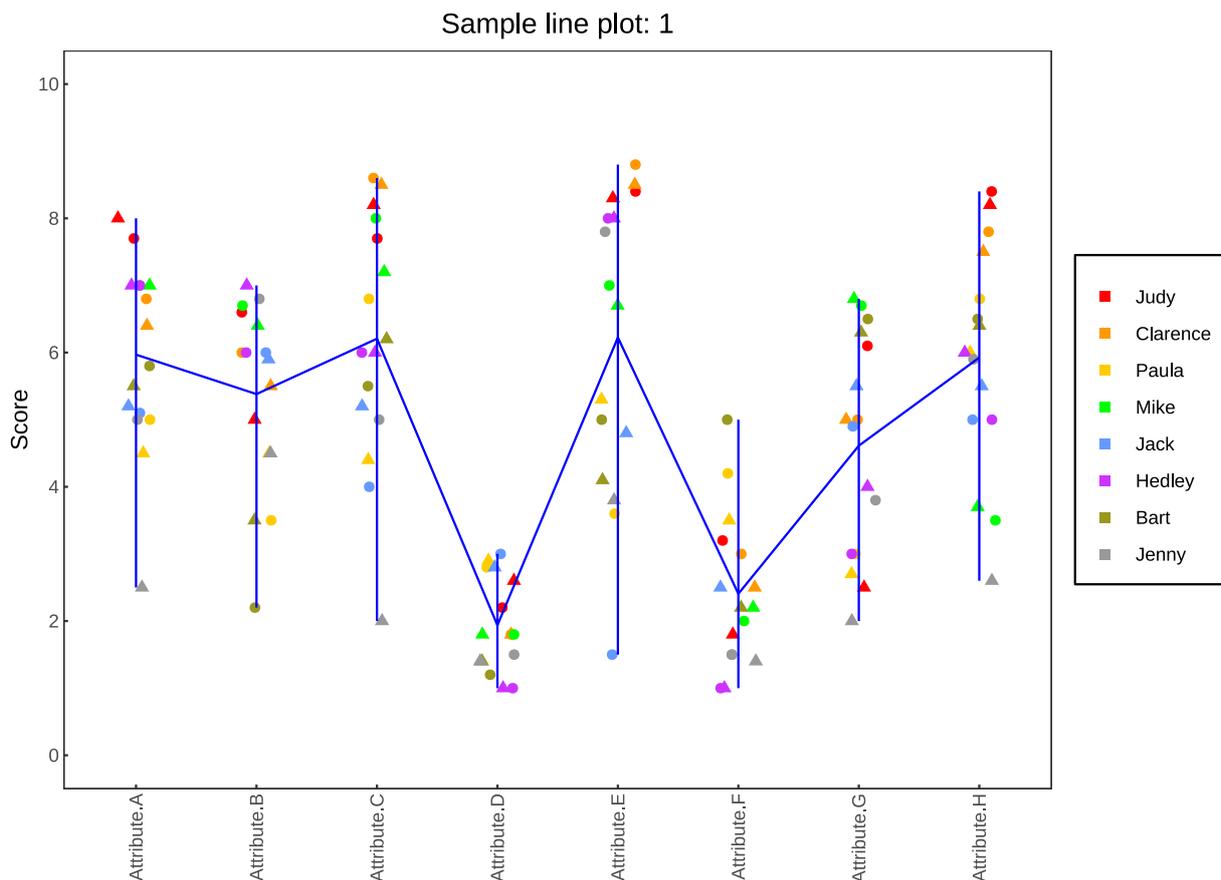
6.1.1 函数介绍

函数 `Sample_line_plot_sp()` 可以实现绘制单一样本 (Sample) 图。该函数储存于 `Sample_line_plot_sp.R` 同名文件中，并且后续所有函数都存于其同名 R 源文件中。为了调用该函数，需要首先运行该 R 文件，见如下示例：

```

1 # 运行函数源文件
2 #source 函数的参数为文件路径，"./" 代表同级文件路径，后面接着为子路径
3 source("./Univariate-code/Sample_line_plot_sp.R",encoding="utf-8")
4 # 调用该函数
5 Sample_line_plot_sp(data=data,sample="1",angle=90,color="blue",lang="en",
6                       score=10)

```



参数介绍

data: 原始数据

sample: 需要绘制的样本名

angle: 当属性过多时用于调整 x 轴刻度标签的文本角度, 默认为 90 度

color: 直线与折线的颜色, 默认为蓝色

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

score: 数据中评分尺度的最大值, 默认为 10

6.1.2 可视化分析

上述示例是对数据 sample.csv 中的样本 **1** 进行数据可视化。

• **散点:**

- 图中的散点是针对样本 **1** 中, 对应属性、对应评估者、对应重复次序的评估得分
- 不同的 **颜色** 的散点代表不同的评估者

- 不同的 **形状** 的散点代表不同的重复次序
- 每个属性的散点之所以不在同一条直线上是因为防止多个散点重合，影响研究人员观察，所以设定了微小的扰动，更有利于视觉分析
- **直线:**
 - 每个属性的 **竖线** 的长度代表该属性得分最高与最低的差值，很容易看出来
 - 连接每个属性的 **折线** 上的点代表该样本该属性下的平均评估得分

图表分析

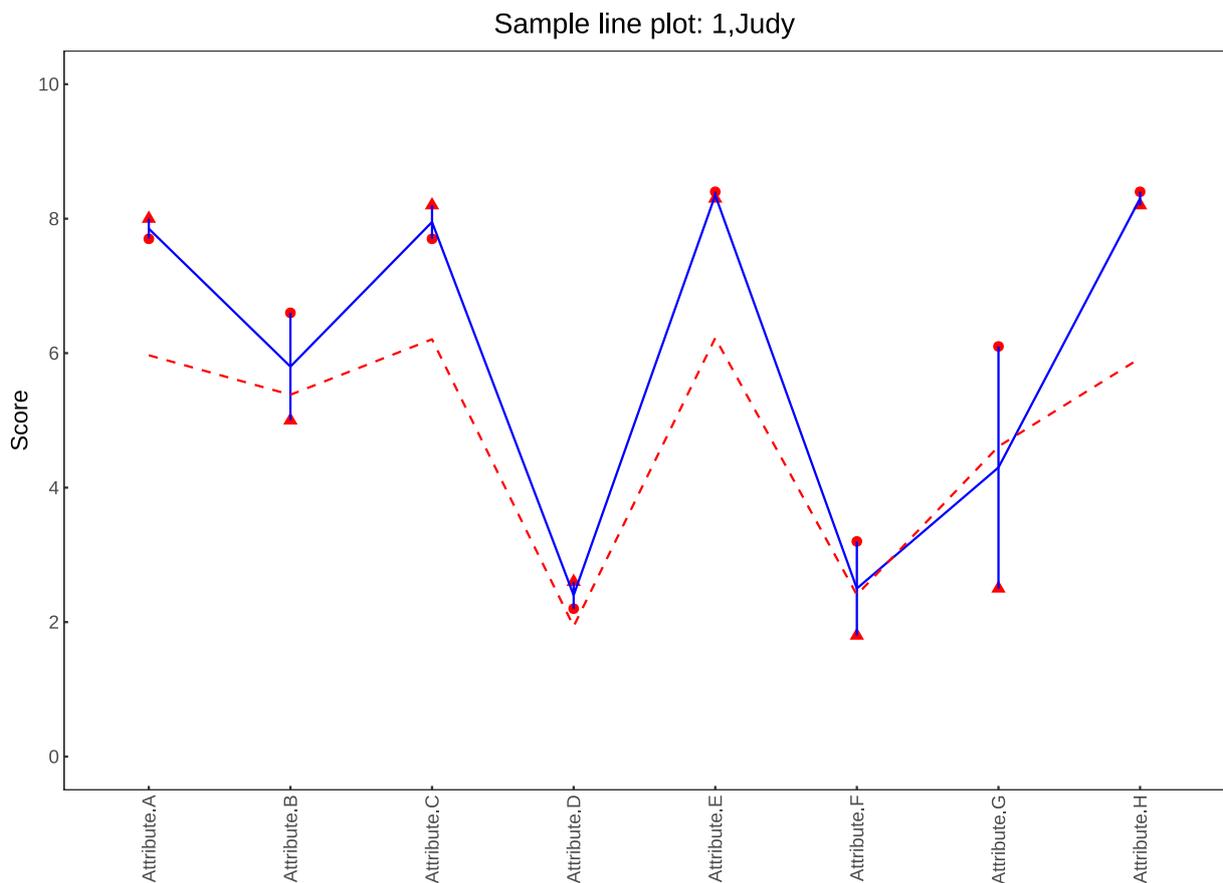
从该类图中，我们可以大致观测出该样本下各感官属性的评估得分是否有显著差异？各评估者之间的评估得分在某感官属性上是否存在差异？重复评估会让评估者的得分呈现一个怎样的变化？这些问题我们都从图中分析出来，从而便于我们初步了解大致情况。

6.2 单一样本-评价者 (`Sample_line_plot_sp_as`)

6.2.1 函数介绍

函数 `Sample_line_plot_sp_as()` 可以实现绘制单一样本-评价者 (Sample+Assessor) 图。该函数储存与 `Sample_line_plot_sp_as.R` 同名文件中。为了调用该函数，需要首先运行该 R 文件，见如下示例：

```
1 # 运行函数源文件
2 source("../Univariate-code/Sample_line_plot_sp_as.R", encoding="utf-8")
3 # 调用该函数
4 Sample_line_plot_sp_as(data=data, sample="1", assessor="Judy", angle=90,
5                       color=c("red", "blue"), score=10, lang="en")
```



参数介绍

data: 原始数据

sample: 需要绘制的样本名

assessor: 需要绘制的评估者名

angle: 当属性过多时用于调整 x 轴刻度标签的文本角度, 默认为 90 度

color: 直线与虚线颜色设置, 默认为红色与蓝色

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

score: 数据中评分尺度的最大值, 默认为 10

6.2.2 可视化分析

上述示例是对数据 `sample.csv` 中的样本 `1` 与评估者 `Judy` 进行数据可视化。该类图本质上是对 `Sample_line_plot_sp()` 函数图形的更细致分析。

- 散点:

- 图中的散点是针对样本 1 与 Judy 的评分中, 对应属性、对应重复次序的评估得分
- 散点颜色与 `Sample_line_plot_sp()` 函数图形的该评估者颜色一致, 更有利于两幅图对比
- 不同的形状的散点代表不同的重复次序
- 直线:
 - 每个属性的竖实线连接该属性下两个表示评估得分的散点
 - 连接每个属性的实折线上的点代表该样本该评估者的该属性下的两个重复次序得分的平均评估得分
 - 红色 (默认参数) 的虚折线表示该样本下每个属性的平均评估得分, 即所有重复次序, 所有评估者的平均得分

图表分析

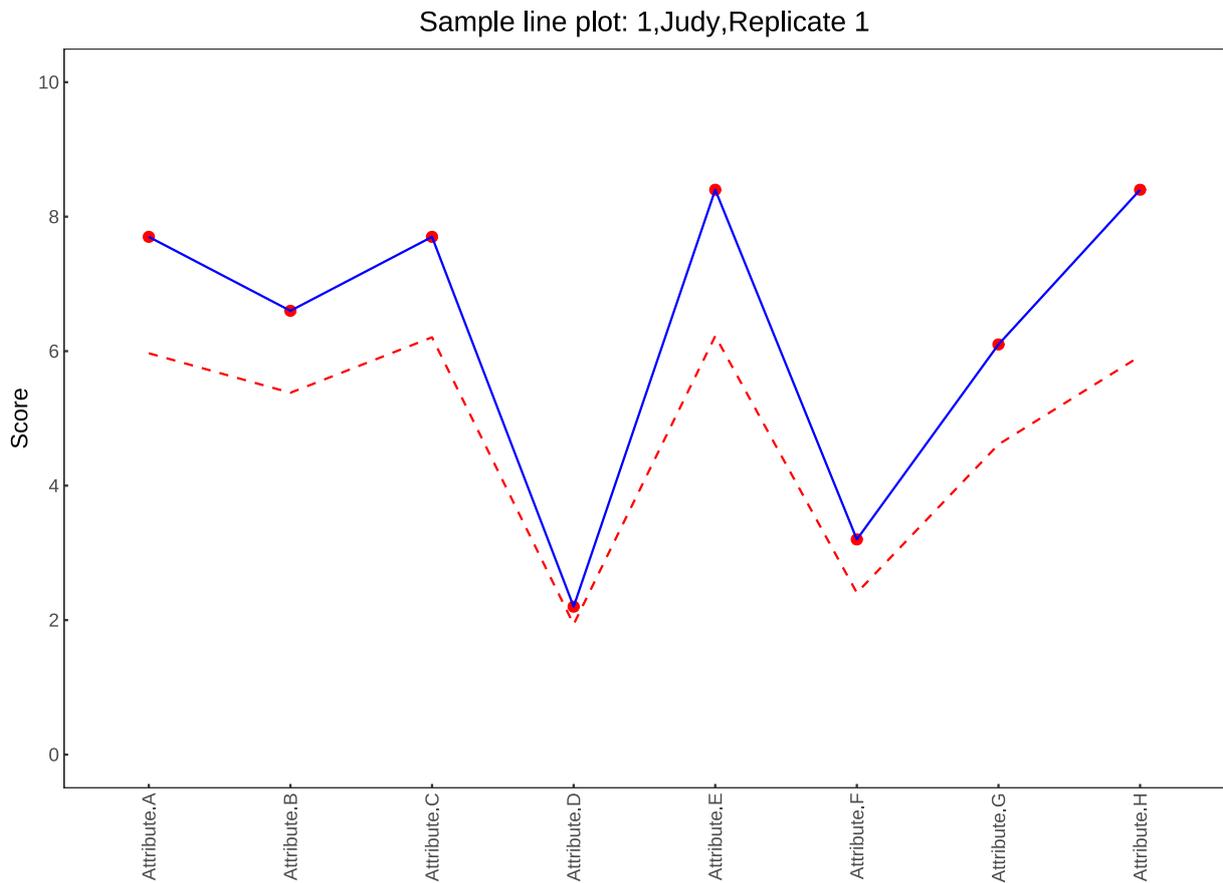
从该类图中, 我们可以观测出该样本下该评估者在各感官属性的评估得分是否有显著差异? 该评估者对该样本在某感官属性上重复评分具有多大差异? 与其他属性相比差异大还是小? 整体上差异如何? 与该样本下的平均得分相比是否有显著差异? 这些问题我们都可以从图中大致观察出来, 从而便于我们后面做更为科学的统计分析。

6.3 单一样本-评价者-重复次序 (`Sample_line_plot_sp_as_re`)

6.3.1 函数介绍

函数 `Sample_line_plot_sp_as_re()` 可以实现绘制单一样本-评价者-重复次序 (Sample+Assessor+Replicate) 图。该函数储存于同名文件中。为了调用该函数, 需要首先运行该 R 文件, 见如下示例:

```
1 # 运行函数源文件
2 source("../Univariate-code/Sample_line_plot_sp_as_re.R", encoding="utf-8")
3 # 调用该函数
4 Sample_line_plot_sp_as_re(data=data, sample="1", assessor="Judy", replicate="1",
5                             angle=90, color=c("red", "blue"), score=10,
6                             lang="en")
```

**参数介绍**

data: 原始数据

sample: 需要绘制的样本名

assessor: 需要绘制的评估者名

replicate: 需要绘制的重复次序

angle: 当属性过多时用于调整 x 轴刻度标签的文本角度, 默认为 90 度

color: 直线与虚线颜色设置, 默认为红色与蓝色

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

score: 数据中评分尺度的最大值, 默认为 10

6.3.2 可视化分析

上述示例是对数据 `sample.csv` 中的样本 1、评估者 Judy 和其第 1 次观测数据进行数据可视化。该类图本质上也是对 `Sample_line_plot_sp()` 函数图形的更细致分析。

- 散点:
 - 图中的散点是针对样本 1 与 Judy 的评分中，对应属性、对应重复次序的评估得分
 - 散点颜色依旧与 `Sample_line_plot_sp()` 函数图形的该评估者颜色一致，更有利于两幅图对比
 - 由于是对应单独重复次序的评分，所以每个属性只有 1 个散点，代表着该重复次序的评分
- 直线:
 - 蓝色（默认参数）的实折线连接每个属性上的点，可以让研究人员观察出波动情况
 - 红色（默认参数）的虚折线表示该样本下每个属性的平均评估得分，即所有重复次序，所有评估者的平均得分，因而可以发现这与上述 `Sample_line_plot_sp_as()` 函数图形的红色虚线一模一样。

图表分析

从该类图中，我们可以观测出该样品，该评估者，特定重复次序下的感官属性评分。根据这些评分与该样本下的平均得分，我们可以做对比，比较差异是否显著，或者说哪些属性上有显著差异，这又能说明什么，从而便于我们后面作针对性分析。

6.4 总结

线图 (Line Plots) 针对每个样本，每个评估者，每个重复次序都能细致的观测出各感官属性的评估得分，并与整体样本平均得分作比较。所以，由上述说明可见，其所能绘制的图表是数量很庞大的。假如数据中样本有 S 个，评估者有 A 个，重复评估次数为 R 次，则上述图表能够输出：

$$\text{Sample_line_plot_sp}() \rightarrow S \text{ 张图}$$

$$\text{Sample_line_plot_sp_as}() \rightarrow S * A \text{ 张图}$$

$$\text{Sample_line_plot_sp_as_re}() \rightarrow S * A * R \text{ 张图}$$

所以在具体实际分析中一般可以先查看 `Sample_line_plot_sp` 图，再根据实际情况，有选择的更细致的查看下一级的绘图。

第 7 章 均值-标准差图 (Mean & STD Plots)

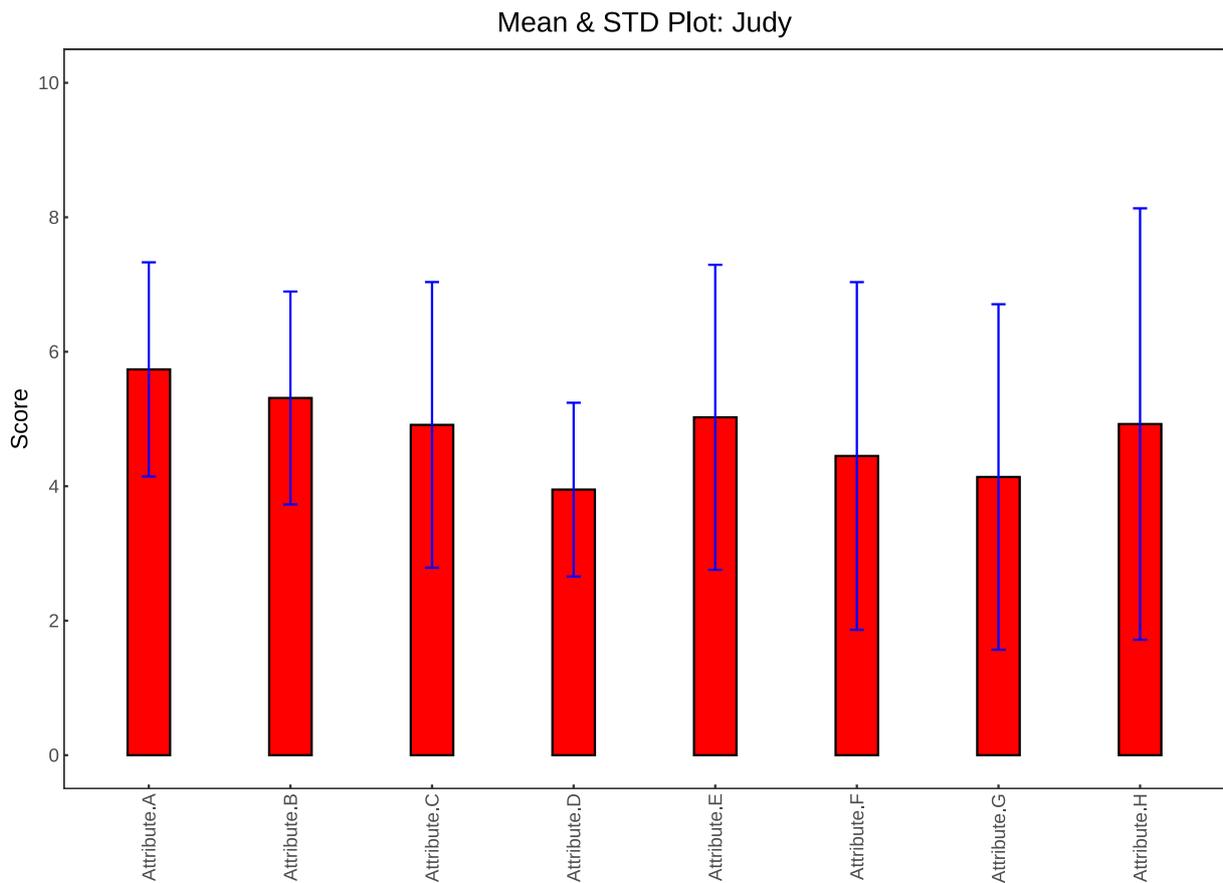
均值-标准差图 (Mean & STD Plots) 是分别对各个感官属性、各个评估者的评分的均值与标准差进行可视化，便于研究人员直观的分析各个感官属性、各个评估者的评分情况。笔者依照其针对对象不同，将其分为两类，一类是基于特定评估者，一类是基于特定感官属性。因而定义了两个函数：`Mean_STD_plot_ass` 和 `Mean_STD_plot_attr`。

7.1 基于特定评估者 (`Mean_STD_plot_ass`)

7.1.1 函数介绍

该函数定义与其同名 R 源文件中，同样在使用该函数时，需要首先利用 `source` 函数运行该 R 文件，然后才能调用该函数，见如下示例：

```
1 # 运行函数源文件
2 source("../Univariate-code/Mean_STD_plot_ass.R", encoding="utf-8")
3 # 调用该函数
4 Mean_STD_plot_ass(data=data, assessor="Judy", angle=90, color=c("red", "blue"),
5                   score=10, lang="en")
```



参数介绍

data: 原始数据

assessor: 需要绘制的评估者名

angle: 当属性过多时用于调整 x 轴刻度标签的文本角度, 默认为 90 度

color: 直线与虚线颜色设置, 默认为红色与蓝色。

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

score: 数据中评分尺度的最大值, 默认为 10

7.1.2 可视化分析

上述示例是根据原始数据, 基于特定评估者 (`Judy`), 对其各个感官属性的平均得分以及标准差进行可视化。

- 该类图的柱子高度代表该评估者的各个感官属性的 `平均得分`, 高度越高, 平均得分越高
- 以每个属性的柱子的上边界为中心, 上下等距延伸的为 `误差线`, 上下延伸的距离长度代表其标准差。

图表分析

从该类图中我们可以直观的观察出特定评估者的各个感官属性的平均得分, 并可以互相作对比。同时, 误差线的长度能反映该感官属性的评分波动情况。可能有些感官属性平均而言评分很高, 但是波动较大, 或者说评分较低, 但是波动较小, 这些可能是我们需要特别关注的。平均得分只能反应其平均水平, 标准差能反应所有评分离平均水平的离散程度, 两个指标需要综合考虑。

7.2 基于特定感官属性 (`Mean_STD_plot_attr`)

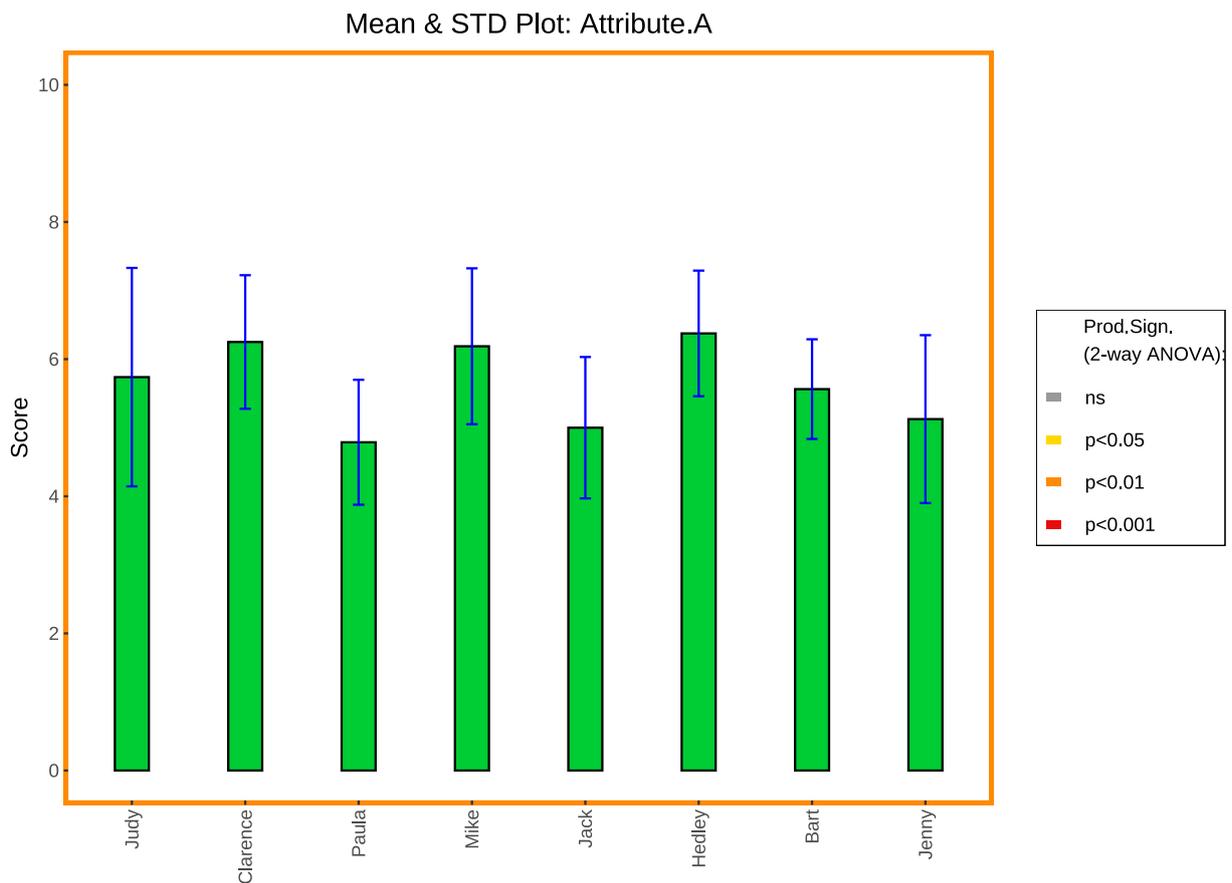
7.2.1 函数介绍

该函数定义与其同名 R 源文件中, 同样在使用该函数时, 需要首先利用 `source` 函数运行该 R 文件, 然后才能调用该函数, 见如下示例:

```

1 # 运行函数源文件
2 source("./Univariate-code/Mean_STD_plot_attr.R", encoding="utf-8")
3 # 调用该函数
4 Mean_STD_plot_attr(data=data, attribute="Attribute.A", angle=90,
5                   color=c("#00cc33", "blue"), score=10, lang="en")

```



注意读者可能发现，在 `sample.csv` 文件中感官属性是 `Attribute A`，中间是空格；而在调用 R 函数给参数赋值时，感官属性中间是“.”。这是因为在将 `sample.csv` 导入到 R 语言中时，感官属性作为列名，而 R 语言命名规范中不能出现空格，所以自动用“.”替代空格。

参数介绍

data: 原始数据

attribute: 需要绘制的感官属性名称

angle: 当属性过多时用于调整 x 轴刻度标签的文本角度，默认为 90 度

color: 直线与虚线颜色设置，默认为暗绿色与蓝色。

lang: 指定中英文模板，默认为“en”:
 _____”cn”: 数据为中文模板
 _____”en”: 数据为英文模板

score: 数据中评分尺度的最大值，默认为 10

7.2.2 可视化分析

本节示例整体上与基于评估者的均值-标准差图类似。不同的是：

- 本节示例是基于特定感官属性，对各个评估者对该属性打分情况进行可视化
- 同样，每根柱子代表每个评估者对该感官属性打分的均值，误差线代表打分情况的标准差
- 值得特别注意的是，您会发现本节示例中，图形面板边框有颜色标记，结合右边图例，您大概可以看出与双因素方差分析有关。该边框颜色标记是以样本与评估者为因素，因变量为感官属性评估得分进行的双因素方差分析：

$$Attribute = Assessor + Sample,$$

- 如本节示例绘图右边图例所示：
 - 灰色代表样本因素的检验 **不显著** (not significant)
 - 黄色代表样本因素的检验 **P 值小于 0.05**
 - 橙色代表样本因素的检验 **P 值小于 0.01**
 - 红色代表样本因素的检验 **P 值小于 0.001**

事实上，包括本节示例以及后面所有基于特定感官属性进行分析的可视化图表都会有边框颜色标记以及右边图例。并且都是样本因素与评估者因素对特定感官属性评估得分的双因素方差分析（无交互作用）。双因素方差分析检验中因素显著意味着该因素会显著影响因变量，即该因素下不同的水平下，因变量会显著不同。反之不显著则因素的变化对因变量影响不大。在实际研究中，我们更希望看到因素是显著的。同时，因素是否显著也取决于我们使用的检验水平，一般是以检验水平 α 为 0.05，在不同的研究领域中，检验水平也可以设置为 0.01, 0.001 等。P 值越小，该因素越显著；P 值小于所给定检验水平 α 值，即可认为该因素显著。

图表分析

从该类图中我们可以直观的观察出特定感官属性的各个评估者的平均得分与其对应的离散程度。这也与上节基于特定评估者的均值-标准差图类似，这里不在赘述。更为需要注意的是，该感官属性的双因素方差检验是否显著显得尤为重要，因为后面需要统计方法都需要以此为基础，这也可以从 PanelCheck 流程图可以看出。如果该感官属性的双因素方差检验结果显著，则意味着不同的样本之间，该感官属性的平均评分有着显著差异，反之，则没有差异。如果检验该属性没有差异，研究人员可以选择将该属性从数据中剔除，不进行下一步分析。

第 8 章 相关关系图 (Correlation Plots)

相关关系图是为了分析对于在各个感官属性评分中，各评估人员评分情况与整体平均值是否保持一致。

8.1 相关关系图 (`Corr_plot_as_attr`)

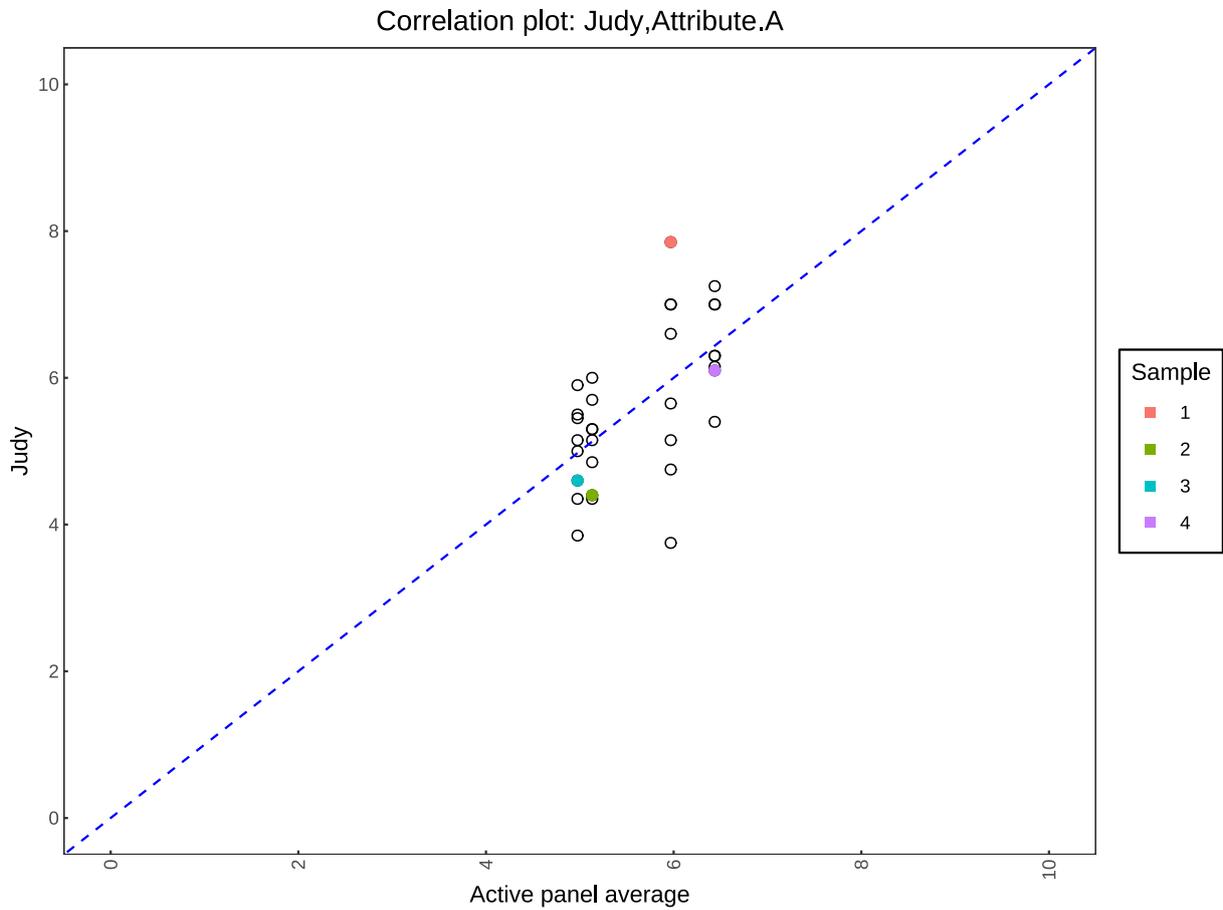
8.1.1 函数介绍

调用函数前请先运行同名函数源文件 `Corr_plot_as_attr`。请看如下示例：

```
1 # 运行函数源文件
2 source("../Univariate-code/Corr_plot_as_attr.R", encoding="utf-8")
```

```

3 # 调用该函数
4 Corr_plot_as_attr(data=data, assessor="Judy", attribute="Attribute.A", angle=90,
5                   color="blue", score=10, lang="en")
    
```



参数介绍

data: 原始数据

assessor: 需要利用颜色填充的评估人员名称

attribute: 需要绘制的感官属性名称

angle: 当属性过多时用于调整 x 轴刻度标签的文本角度, 默认为 90 度

color: 虚线颜色设置, 默认为蓝色。

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

score: 数据中评分尺度的最大值, 默认为 10

8.1.2 可视化分析

上述示例首先是基于特定感官属性 `Attribute.A`，以该属性评分下各个样本的总体平均得分为 x 轴，以该属性下每个样本和评估者两两搭配的平均得分为 y 轴，被填充颜色的散点则是在图中指定的评估者 `Judy` 所代表的在该属性下的各个样本的平均评估得分。

- 本节示例中蓝色虚线对角线是 $y = x$ 直线。如果评估人员完全同意该属性的整体样本平均值，散点将落在蓝色虚线上

图表分析

该类图中可以直观的观测出每个评估人员对该属性下的每个样本评分均值是否与整体均值相差不大，或者说相差大不大。如果相差很小，则对应填充颜色散点会落在蓝色虚线上，否则则会偏离。与蓝色虚线距离越远，则说明该评估人员与所有评估人员的均值显著不一致，各评估人员评分方差比较大，也意味着各评估人员意见显著不一致，需要后面进一步分析。反之，距离越近则说明各评估人员意见比较一致。当然这都是粗略的从图中分析。

第 9 章 轮廓图 (Profile Plots)

轮廓图是基于给定感官属性对各个样本下的各个评估人员的评分进行可视化。以各样本名作为 x 轴，以各类评分均值为 y 轴。PanelCheck 对轮廓图提供了多个绘图选项，包括 x 轴标签的展示顺序，是计算重复次序均值作为样本下评估人员评分还是各重复次序的评分都绘制出来等。

9.1 轮廓图 (`Profile_plot_attr`)

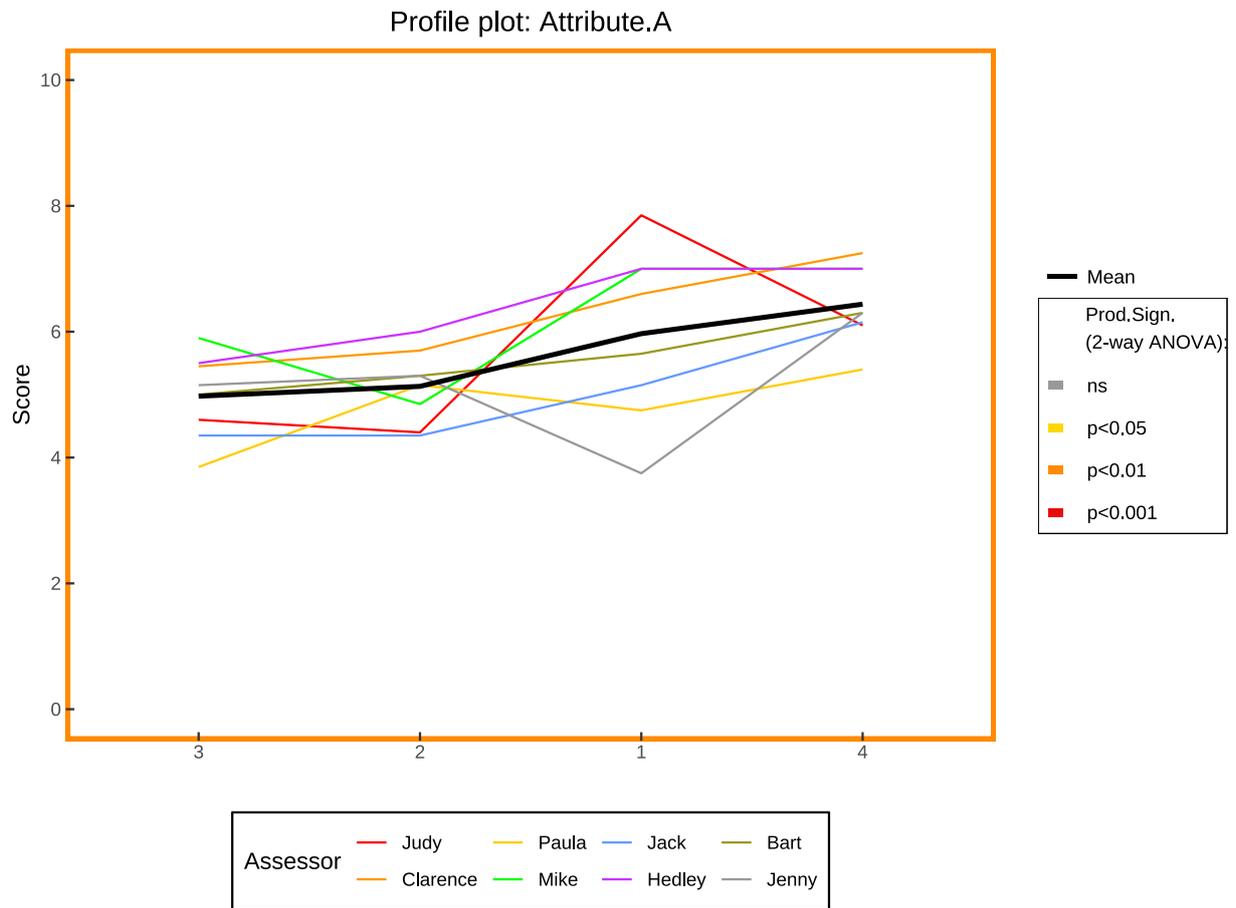
9.1.1 函数介绍

函数 `Profile_plot_attr` 可以实现轮廓图，并且涵盖了 PanelCheck 的所有可调控参数，同样使用之前，先要运行其同名 R 源文件。

```

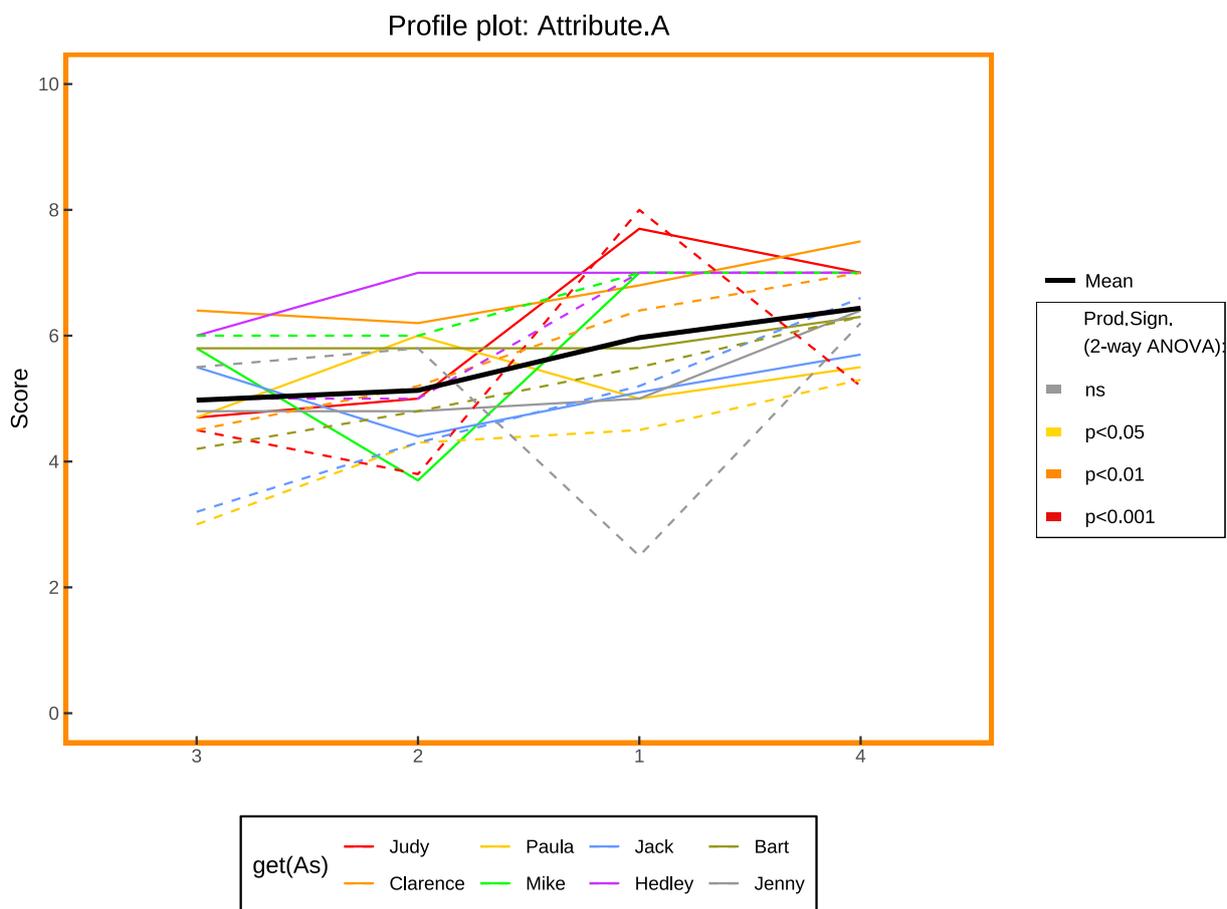
1 # 运行函数源文件
2 source("../Univariate-code/Profile_plot_attr.R",encoding="utf-8")
3 # 调用该函数
4 # 展示样本重复次序均值评分
5 Profile_plot_attr(data=data,attribute="Attribute.A",angle=0,
6                   color="black",logic_sort=TRUE,show="sample_averages",
7                   score=10,lang="en")

```



```

1 # 运行函数源文件
2 source("../Univariate-code/Profile_plot_attr.R",encoding="utf-8")
3 # 调用该函数
4 # 展示样本所有重复次序评分
5 Profile_plot_attr(data=data,attribute="Attribute.A",angle=0,
6                   color="black",logic_sort=TRUE,show="sample_replicates",
7                   lang="en",score=10)
    
```



参数介绍

data: 原始数据

attribute: 需要绘制的感官属性名称

angle: 当属性过多时用于调整 x 轴刻度标签的文本角度, 默认为 0 度

color: 均值线的颜色, 默认为黑色

logic_sort: 逻辑值, x 轴是否按照样本得分均值由低到高顺序显示, 默认为 TRUE

show: 展示各重复次序平均值还是每个重复观测值, 默认为 "sample_averages"

—— "sample_averages" —— 展示各重复次序平均值

—— "sample_replicates" —— 展示每个重复观测值

lang: 指定中英文模板, 默认为 "en":

—— "cn": 数据为中文模板

—— "en": 数据为英文模板

score: 数据中评分尺度的最大值, 默认为 10

9.1.2 可视化分析

本节示例绘制了两张图,一张图是参数show="sample_averages",另一张图是参数show="sample_replicates"。该参数的设置是基于 PanelCheck 用户界面右下角所示可选择选项,同样 logic_sort 参数也是如此,都是复刻了右下角可修改选项:

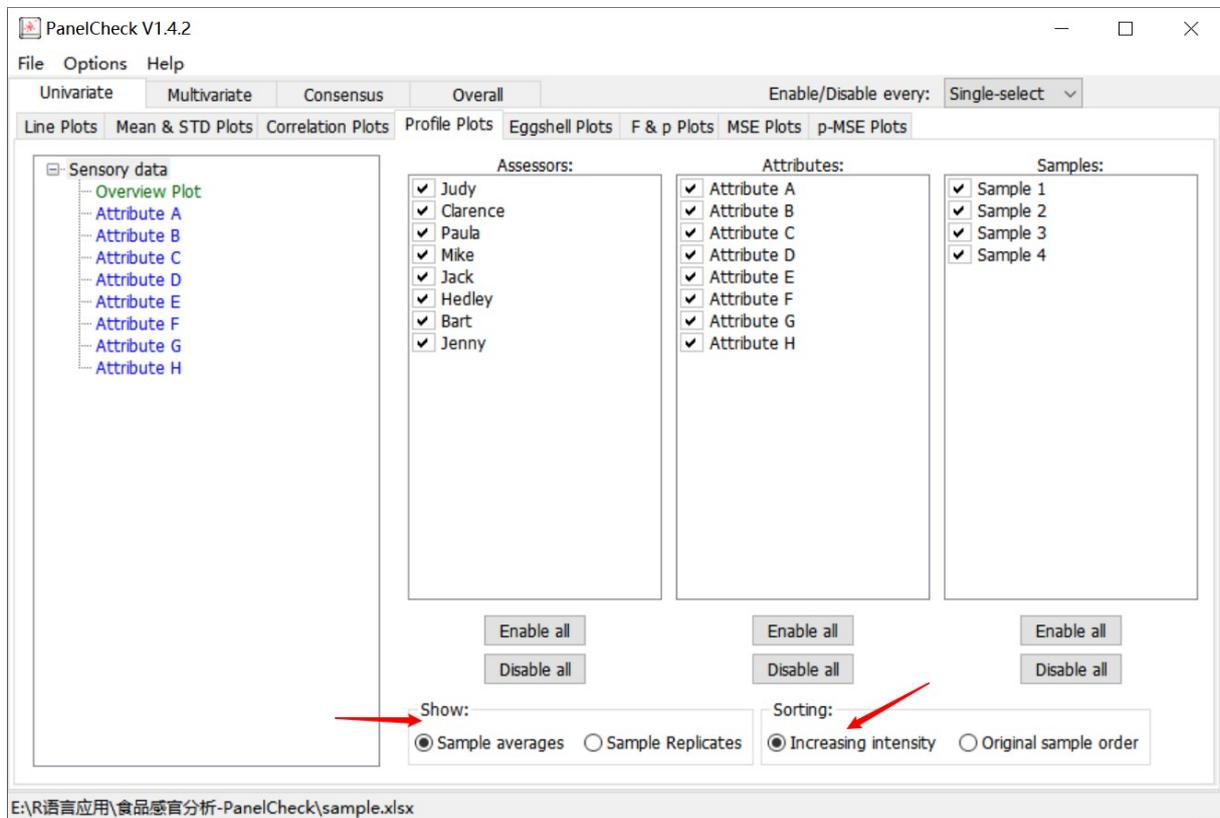


图 1: PanelCheck 软件 Profile Plots 绘图界面

在本节示例两张图中:

- 绘图面板边框的颜色标记是基于给定属性的双因素方差分析结果, 具体细节请看7.2.2.
- 黑色直线是在给定属性下, 连接每个样本的各评估人员评分均值点
- 其他颜色标记的直线如绘图下边图例所示, 为给定属性下各评估人员在各样本下的评估得分均值
- 注意第二张绘图比第一张绘图多了等同于评估人员数量条直线, 且每种颜色都有直线和虚线两条, 这是因为第二种图并没有计算所有重复次序的均值, 而是将其以不同线型绘制出来而已, 表示的是各重复次序下的评分情况

! 注意, 如果重复次序过多, 不建议使用 show="sample_replicates", 因为这会导致绘图中线条数量过多, 极大影响直观视觉分析。

图表分析

轮廓图 (Profile Plots) 可以让研究人员大致观测给定感官属性下各样本的评分情况, 直观观察各样本评分是否一致? 呈现怎样的波动? 是否有某个评估人员的评分显著影响了样本均值评分? 这些我们都可以大致分

析出来。

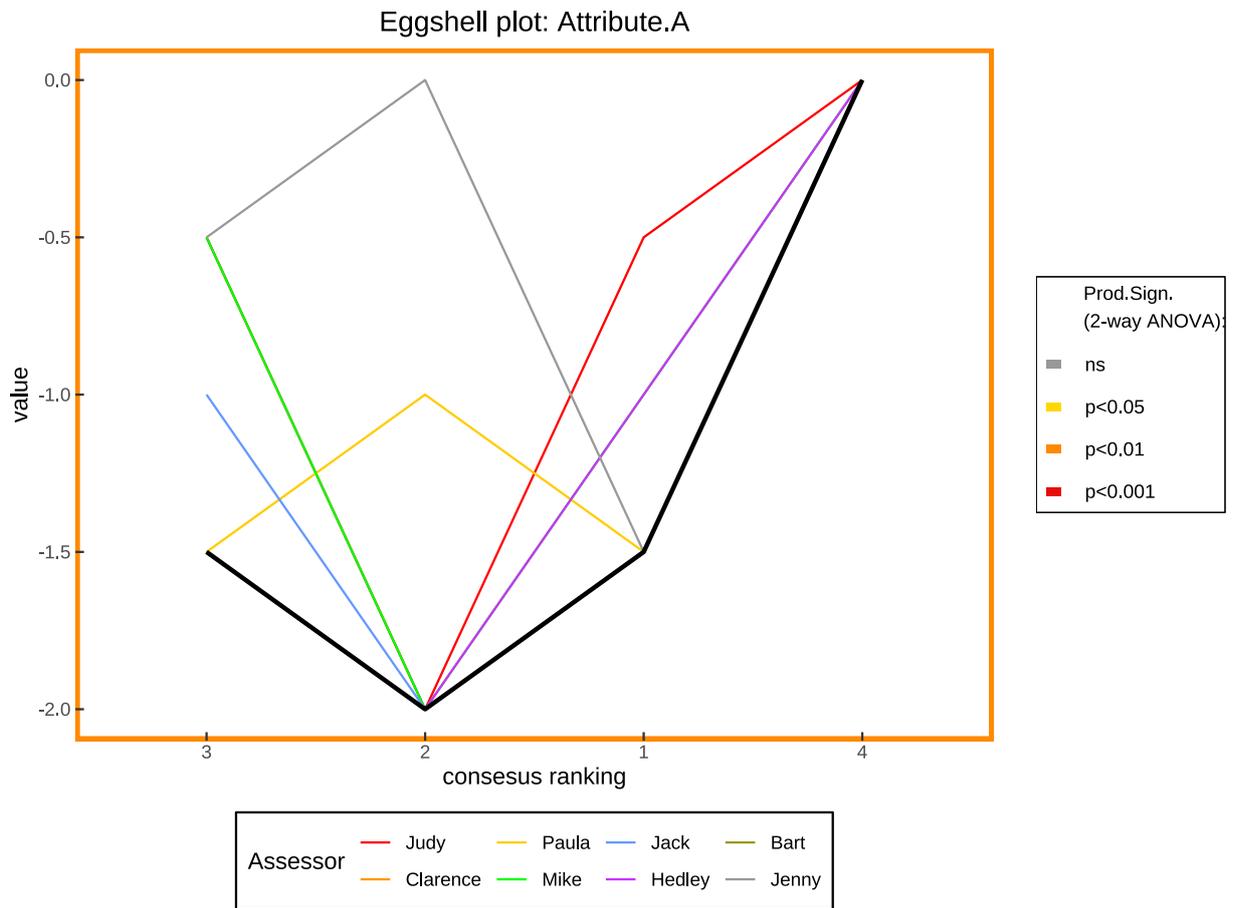
第 10 章 蛋壳图 (Eggshell Plots)

蛋壳图显示了每个评估员如何根据一致性 (Consensus) 对所有测试样本进行排名。黑色 (默认参数) 粗线代表一致性线 (一致性样本排名), 其余每一条线代表一个特定评估者的样本排名。一名评估者的线越接近一致性线, 该评估者就越同意该一致性。最左边的样本被一致性排序为最低 (低属性强度), 反之亦然。

10.1 蛋壳图 (Eggshell_plot_attr)

10.1.1 函数介绍

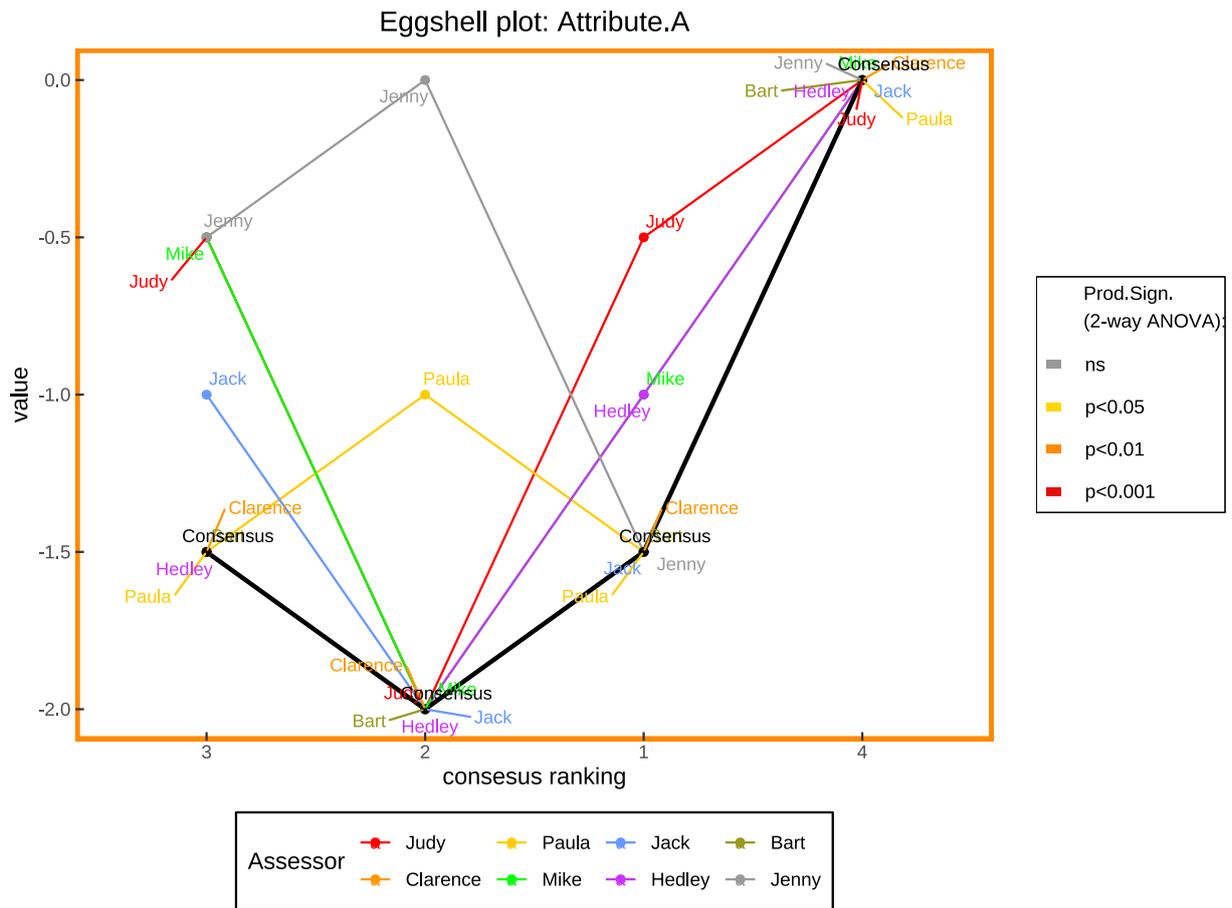
```
1 # 运行函数源文件
2 source("../Univariate-code/Eggshell_plot_attr.R", encoding="utf-8")
3 # 调用该函数
4 # 不显示散点以及文本标签
5 Eggshell_plot_attr(data=data, attribute="Attribute.A", linetype="solid",
6                   text=FALSE, text_size=3, angle=0, color="black", lang="en")
```



```

1 # 运行函数源文件
2 source("../Univariate-code/Eggshell_plot_attr.R",encoding="utf-8")
3 # 调用该函数
4 # 显示散点以及文本标签
5 Eggshell_plot_attr(data=data,attribute="Attribute.A",linetype="solid",
6                     text=TRUE,text_size=3,angle=0,color="black",
7                     lang="en")

```



参数介绍

data: 原始数据

attribute: 需要绘制的感官属性名称

linetype: 绘制的线型, 默认线型为实线"solid"

———可选值有虚线"dashed", 长虚线"longdash", 点虚线"dotdash", 点"dotted", "twodash"

text: 逻辑值, 是否显示评估者文本标签, 默认值为 FALSE

———便于当线段重合时分辨某评估者属于哪段折线

textsize: 若 text 为 TRUE, 则可以通过此参数调整文本标签字号大小

———数值越大, 字体越大, 默认值为 3

angle: 设置 x 轴刻度文本的显示角度, 在文本较长时非常有用, 默认为 0 度 (水平)

color: 一致性排名 (Consensus) 颜色, 默认为黑色

lang: 指定中英文模板, 默认为"en":

———"cn": 数据为中文模板

———"en": 数据为英文模板

10.1.2 可视化分析

蛋壳图基于这样的假设，即所有样本都通过感官等级测试或感官特征数据排名（其中等级是根据评估员的分数计算的）为每个评估者排序。当用于感官分析时，蛋壳图直观地显示了每个评估员相对于一个属性的一致性样本排名的累积样本排名。由于一致性样本排名计算比较复杂，这里不再叙述，感兴趣的话请阅读参考文献 [1]。

- 第一张绘图与 PanelCheck 软件图表基本一致。不同在于，软件中使用紫色（不可修改）粗线代表一致性排名情况，而笔者写的函数中可以修改粗线颜色，默认为黑色粗线
- 由于蛋壳图很容易使得多个评估人员的线条重合，会很大干扰视觉直观分析。尽管 PanelCheck 使用不轮廓图同线型如虚线来区分重合的线条，在重合线条很多的情况下，作用不大
- 所以修改 text 参数，绘制显示散点与文本标签的蛋壳图可以让研究人员很容易找到自己想要查看的评估人员的累计样本排名。

图表分析

蛋壳图提供了被测样本中评价者之间的关系如何变化的指示。通过这种方式，评审组组长可以在一张图中获得所有评审员的图形化概览，从而可以很容易地识别出那些在排名上与一致性排名不同的评审员。也就是说，如果所有评估者累计样本排名与黑色（默认参数）粗线很接近，那么说明评估者对该感官属性一致性就好，反之若某评估者的累计排名明显不同于一致性排名，则可以很容易、快速地识别出该评估者在该属性上存在问题。

第 11 章 F&P 图 (F & P Plots)

F 值图与 P 值图是对给定感官属性与评估者下，对样本进行单因素方差分析 (One-way ANOVA)，然后分别对各自的 F 统计量的值，与检验 P 值进行可视化。

11.1 F 值图 (F_plot)

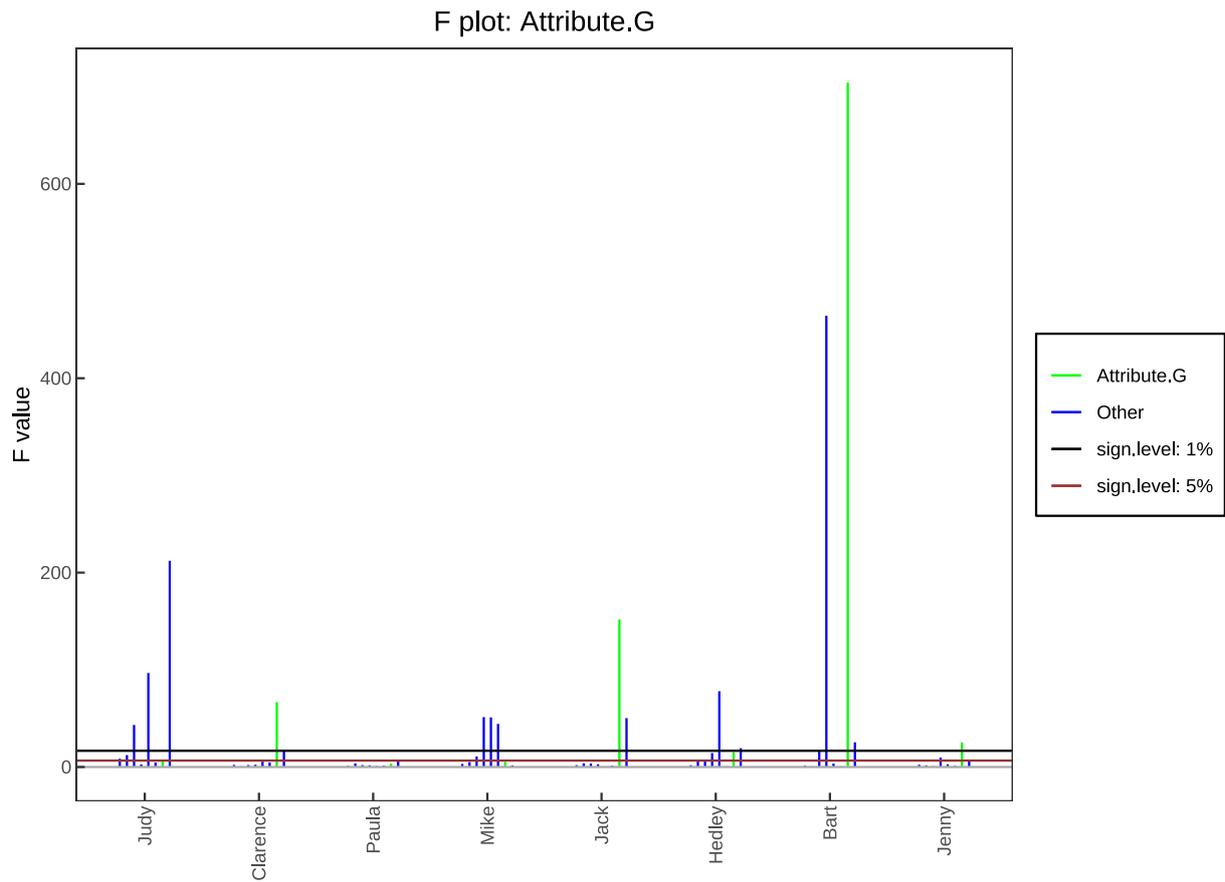
11.1.1 函数介绍

函数 `F_plot` 可以轻松绘制基于给定属性下（按照评估者分类）或者给定评估者下（按照感官属性分类）的 F 值图，见如下示例：

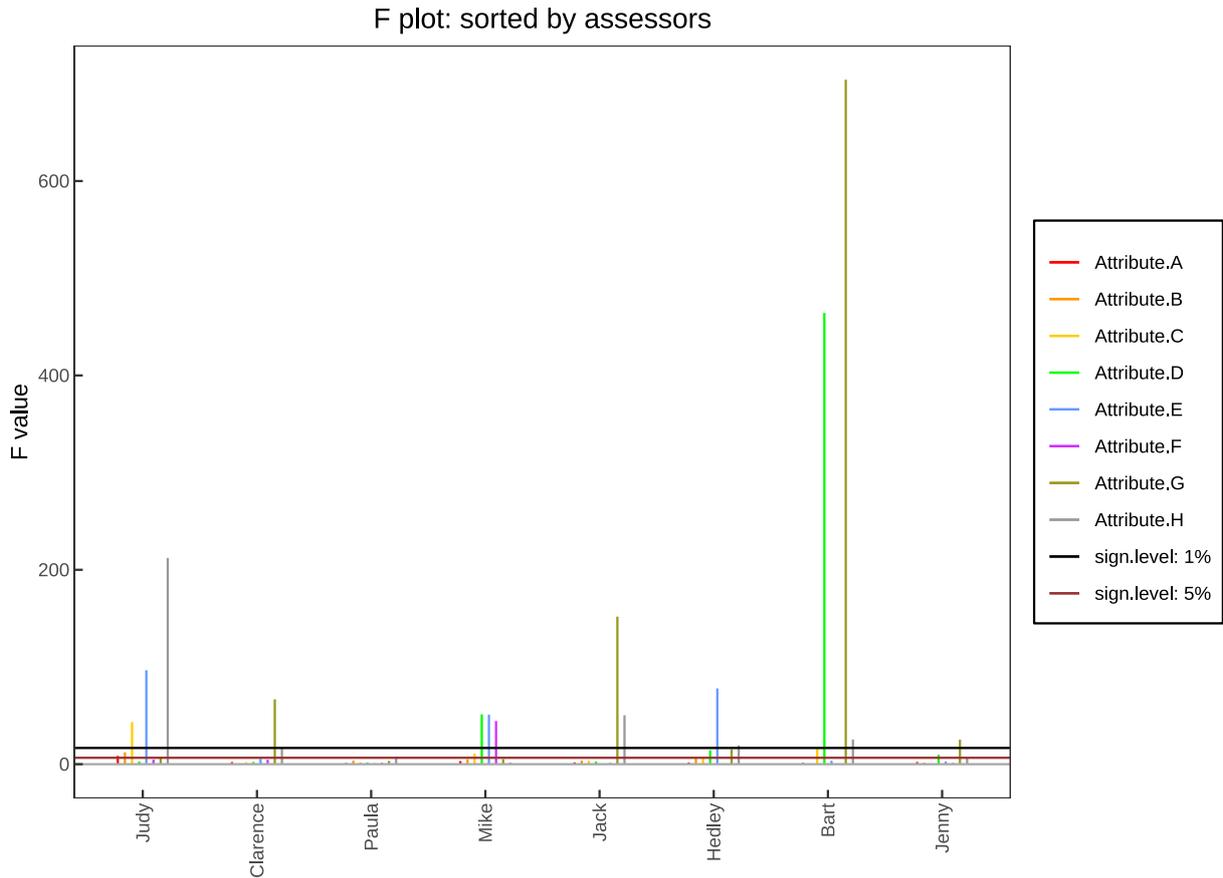
```

1 # 运行函数源文件
2 source("../Univariate-code/F_plot.R", encoding="utf-8")
3 # 调用该函数
4 # 绘制突出显示给定感官属性下的 F 值图
5 F_plot(data=data, attribute="Attribute.G", by="assessor",
6         vline_color=c("green", "blue"), hline_color=c("black", "#993333"),
7         angle=90, lang="en")

```



```
1 # 运行函数源文件
2 source("../Univariate-code/F_plot.R",encoding="utf-8")
3 # 调用该函数
4 # 绘制突出显示所有感官属性下的 F 值图
5 F_plot(data=data,attribute="all",by="assessor",
6         hline_color=c("black","#993333"),angle=90,lang="en")
```



参数介绍

data: 原始数据

attribute: 当值为属性名时, 输出单独属性的 F plot; 当值为"all" 时, 输出所有属性的 F plot

assessor: 当值为评估值名称时, 输出该评估者的 F plot; 当值为"all" 时, 输出所有评估者的 F plot

angle: 设置 x 轴刻度文本的显示角度, 在文本较长时非常有用, 默认为 90 度 (垂直)

by: 指定 F plot 是按照评估者分类绘制, 还是按照属性分类绘制:

————— 默认值为"assessor"(评估者), 如果按照属性分类, 则为"attribute"

————— 当 by 为"assessor" 时, 需要指定 attribute 参数;

————— 当 by 为"attribute" 时, 需要指定 assessor 参数

vline_color: 当且仅当 attribute 参数不为"all" 时可以设置:

————— 表示特定属性与其他属性的竖线颜色, 有默认值

hline_color: 设置检验水平 0.01 与 0.05 的水平线颜色, 有默认值

lang: 指定中英文模板, 默认为"en":

————— "cn": 数据为中文模板

————— "en": 数据为英文模板

11.1.2 可视化分析

F 值图会首先绘制所有属性与所有评估者对于样本的两两单因素方差分析 F 值，并以竖线的高度来表现 F 值的大小，然后根据给定的属性（按评估者分类，作为 x 轴）或给定的评估者（按属性分类，作为 x 轴）[分类由 *by* 参数决定]，用不同颜色标记它们。同时，根据 F 检验的 *上侧 0.05 分位数* 值和 *上侧 0.01 分位数* 值，绘制不同颜色水平线。

图表分析

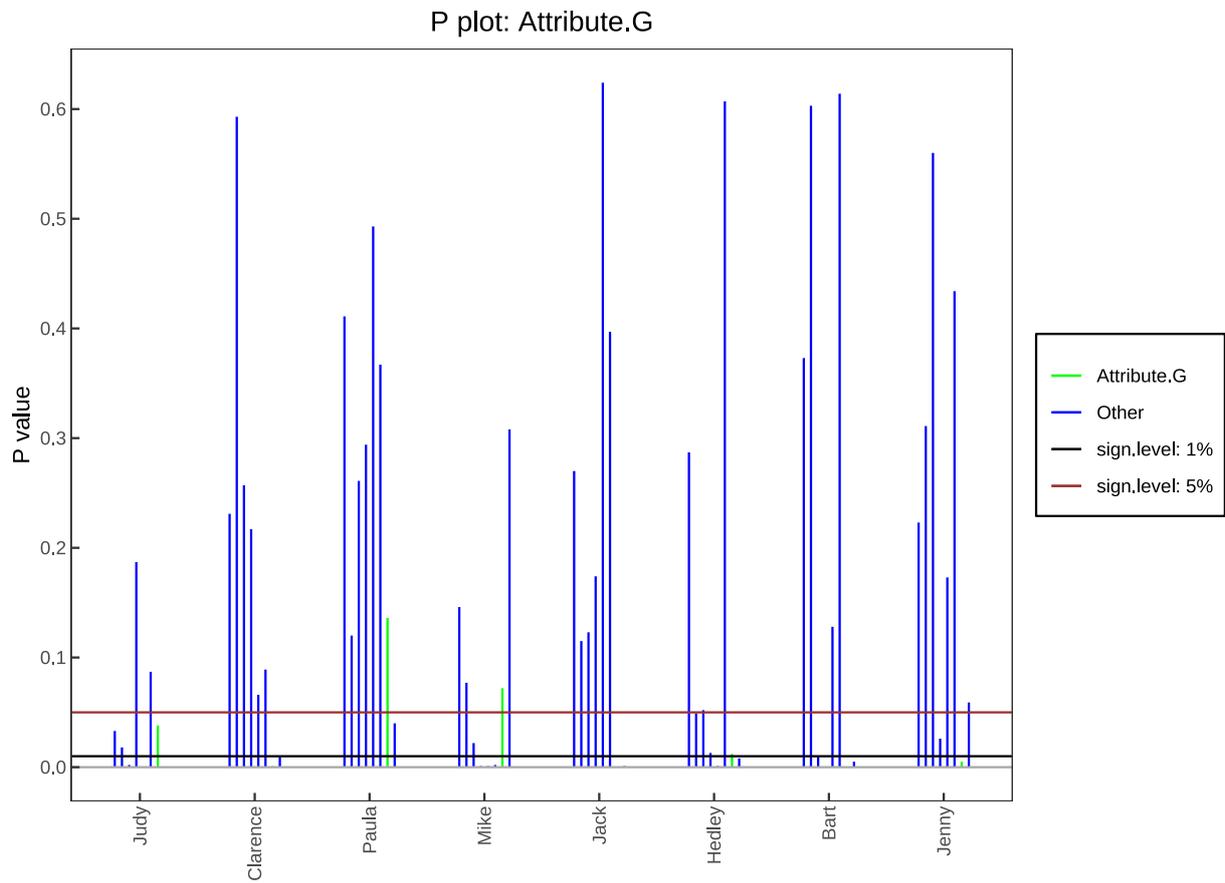
F 值图是以统计检验中的单因素方差分析为基础，根据 F 值的大小来衡量评估人员对样本的区分能力，并且 F 值越高越好。而两条分位数的水平线其实等价于给定 F 检验的两个检验水平：0.01 和 0.05。如果 F 值大于某个水平线，意味着在该检验水平下，该评估人员根据该感官属性是可以显著区分该样本的，反之，则没有充分证据表明能够区分。

11.2 P 值图 (*P*_plot)

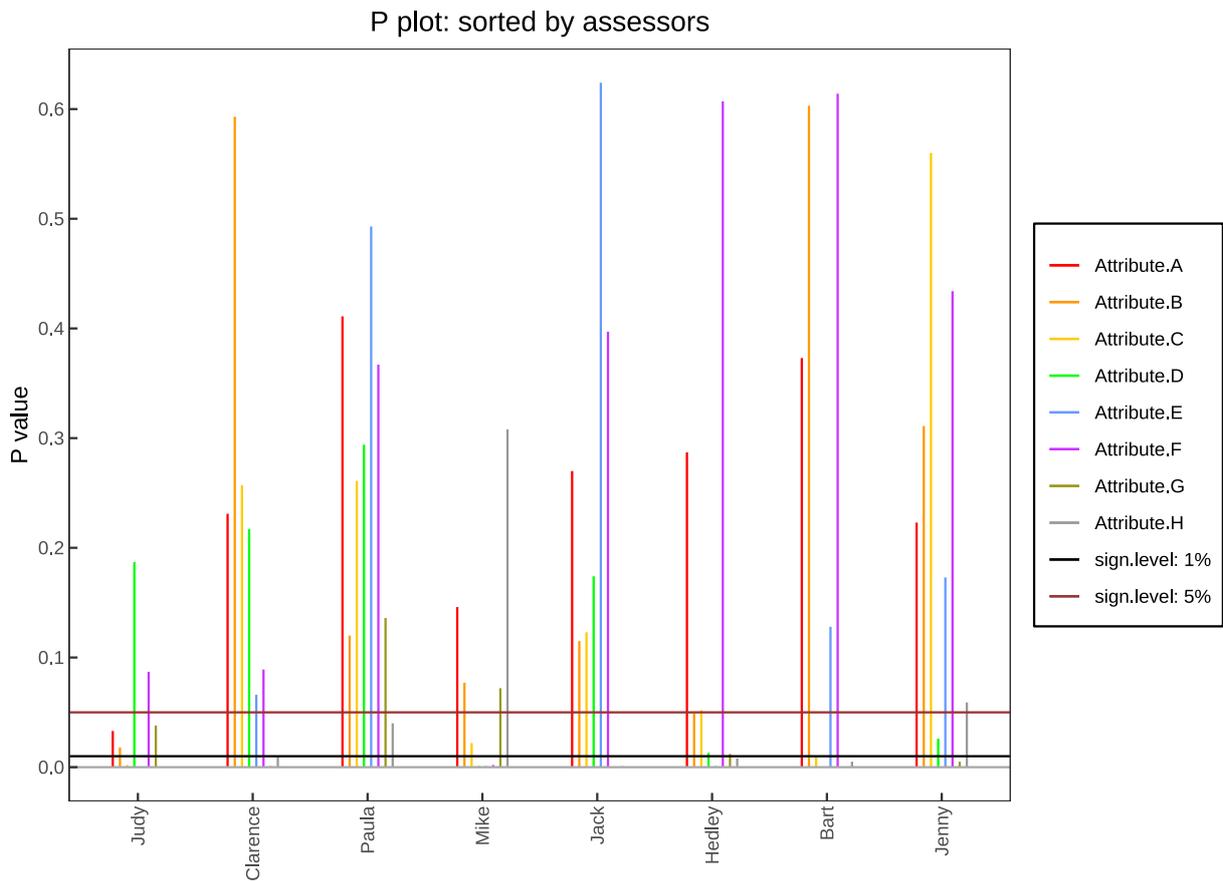
11.2.1 函数介绍

与 F 值图类似，我们使用 *P*_plot 函数绘制 P 值图，见如下示例：

```
1 # 运行函数源文件
2 source("../Univariate-code/P_plot.R",encoding="utf-8")
3 # 调用该函数
4 # 绘制突出显示给定感官属性下的 P 值图
5 P_plot(data=data,attribute="Attribute.G",by="assessor",
6         vline_color=c("green","blue"),hline_color=c("black","#993333"),
7         angle=90,lang="en")
```



```
1 # 运行函数源文件
2 source("../Univariate-code/P_plot.R",encoding="utf-8")
3 # 调用该函数
4 # 绘制突出显示所有感官属性下的 P 值图
5 P_plot(data=data,attribute="all",by="assessor",
6         hline_color=c("black","#993333"),angle=90,lang="en")
```



参数介绍

data: 原始数据

attribute: 当值为属性名时, 输出单独属性的 P plot; 当值为"all" 时, 输出所有属性的 P plot

assessor: 当值为评估值名称时, 输出该评估者的 P plot; 当值为"all" 时, 输出所有评估者的 P plot

angle: 设置 x 轴刻度文本的显示角度, 在文本较长时非常有用, 默认为 90 度 (垂直)

by: 指定 P plot 是按照评估者分类绘制, 还是按照属性分类绘制:

————— 默认值为"assessor"(评估者), 如果按照属性分类, 则为"attribute"

————— 当 by 为"assessor" 时, 需要指定 attribute 参数;

————— 当 by 为"attribute" 时, 需要指定 assessor 参数

vline_color: 当且仅当 attribute 参数不为"all" 时可以设置:

————— 表示特定属性与其他属性的竖线颜色, 有默认值

hline_color: 设置检验水平 0.01 与 0.05 的水平线颜色, 有默认值

lang: 指定中英文模板, 默认为"en":

————— "cn": 数据为中文模板

————— "en": 数据为英文模板

11.2.2 可视化分析

P 值图绘制方法与 F 值图几乎完全一致，不同的是其利用 F 检验的 P 值替代 F 值进行可视化，具体细节请看上节11.1.2。

图表分析

P 值图也是以统计检验中的单因素方差分析为基础，根据 P 值的大小来衡量评估人员对样本的区分能力，但是 P 值是越小越好。可以看到两条水平线分别位于 P 值为 0.01 和 0.05 处。如果 P 值小于某个水平线，意味着在该检验水平下，该评估人员根据该感官属性是可以显著区分该样本的，反之，则没有充分证据表明能够区分。

第 12 章 均方误差图 (MSE Plots)

均方误差图与上一章11F&P 值图应用的统计理论完全一致，即使用了单因素方差分析。不同于利用 F 值于 P 值进行可视化，均方误差图使用的是均方误差统计量 **MSE** (Mean Square Error) 进行可视化，并且没有所谓的显著性水平线。

12.1 均方误差图 (**MSE_plot**)

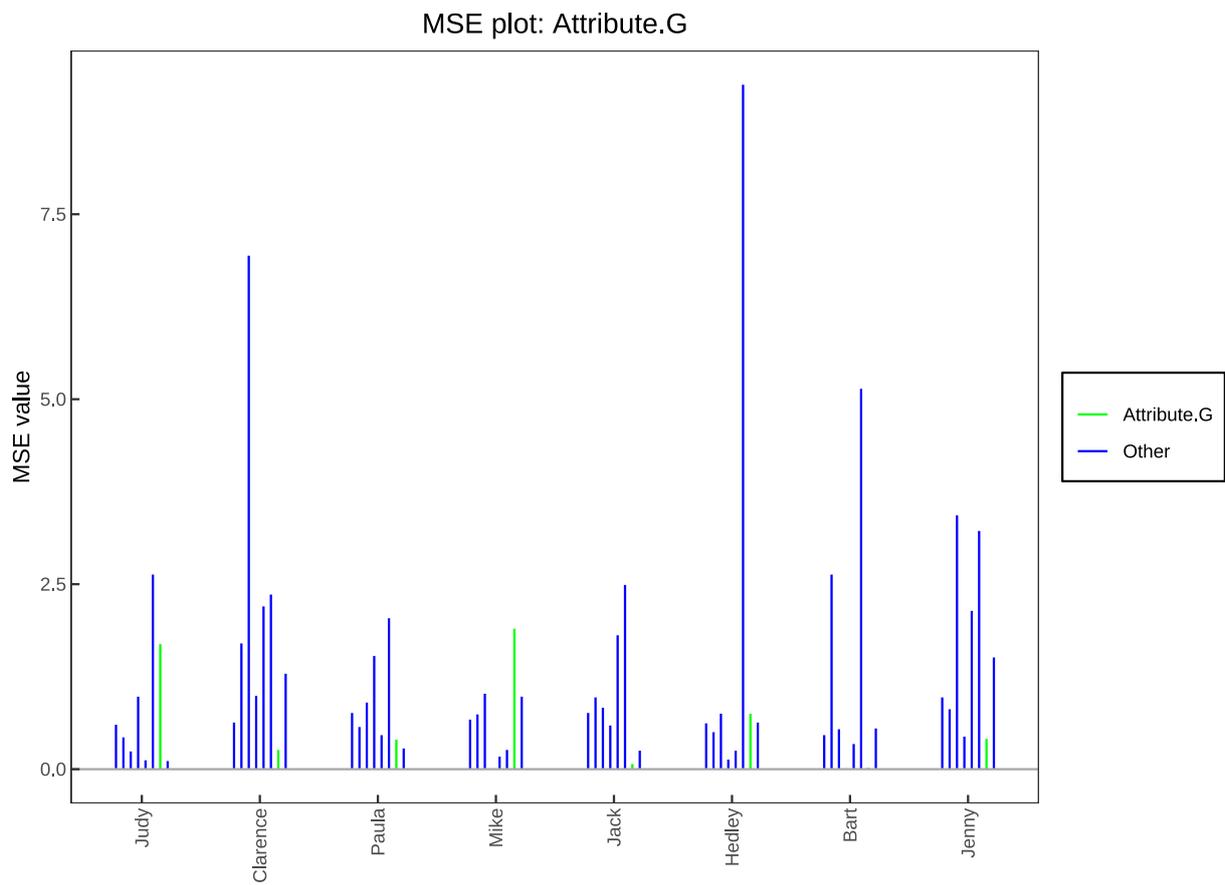
12.1.1 函数介绍

与 F 值与 P 值图类似，使用函数 **MSE_plot** 进行绘制，见如下示例：

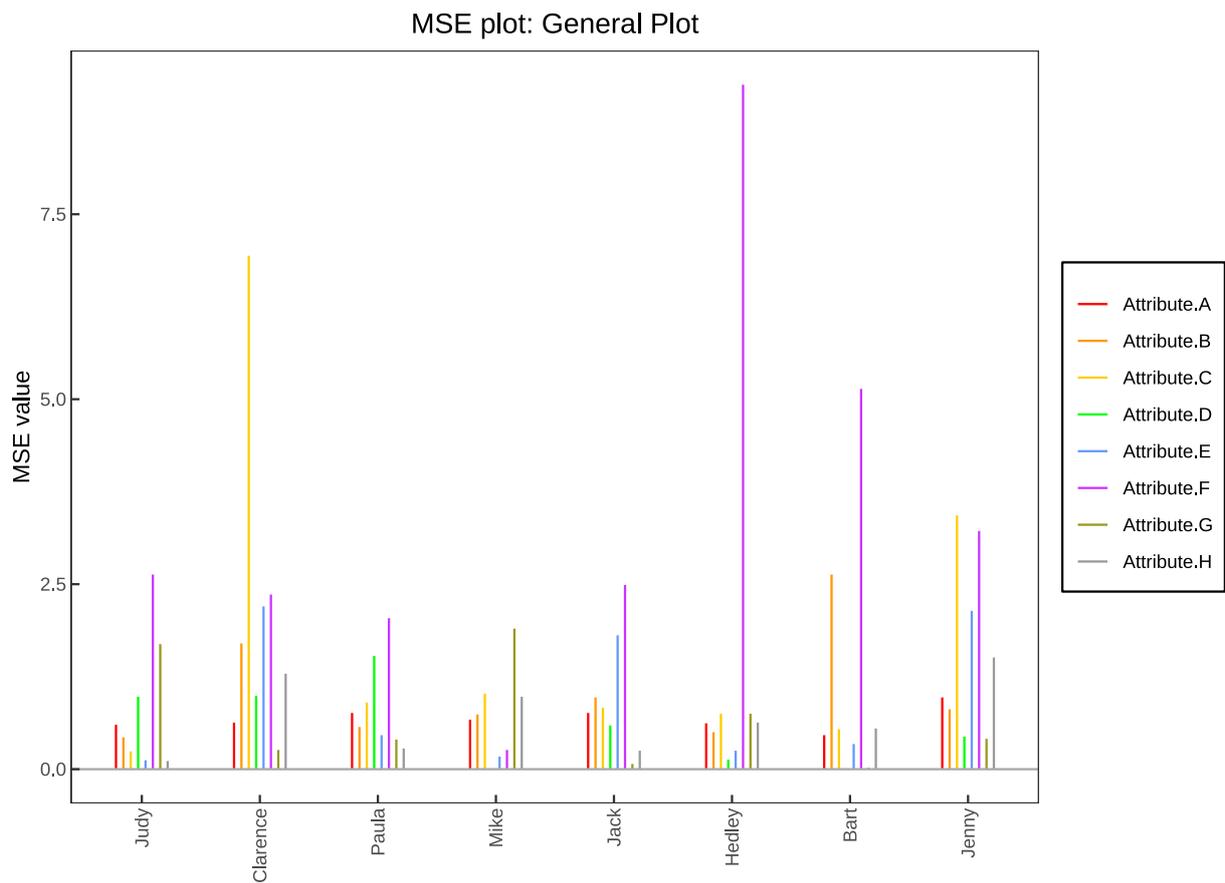
```

1 # 运行函数源文件
2 source("../Univariate-code/MSE_plot.R", encoding="utf-8")
3 # 调用该函数
4 MSE_plot(data=data, attribute="Attribute.G", by="assessor",
5           vline_color=c("green", "blue"), angle=90, lang="en")

```



```
1 # 运行函数源文件
2 source("../Univariate-code/MSE_plot.R",encoding="utf-8")
3 # 调用该函数
4 MSE_plot(data=data,attribute="all",by="assessor",angle=90,lang="en")
```



参数介绍

data: 原始数据

attribute: 当值为属性名时, 输出单独属性的 MSE plot; 当值为"all" 时, 输出所有属性的 MSE plot

assessor: 当值为评估值名称时, 输出该评估者的 MSE plot; 当值为"all" 时, 输出所有评估者的 MSE plot

angle: 设置 x 轴刻度文本的显示角度, 在文本较长时非常有用, 默认为 90 度 (垂直)

by: 指定 MSE plot 是按照评估者分类绘制, 还是按照属性分类绘制:

————— 默认值为"assessor"(评估者), 如果按照属性分类, 则为"attribute"

————— 当 by 为"assessor" 时, 需要指定 attribute 参数;

————— 当 by 为"attribute" 时, 需要指定 assessor 参数

vline_color: 当且仅当 attribute 参数不为"all" 时可以设置:

————— 表示特定属性与其他属性的竖线颜色, 有默认值

lang: 指定中英文模板, 默认为"en":

————— "cn": 数据为中文模板

————— "en": 数据为英文模板

12.1.2 可视化分析

图表分析

MSE 值图绘制的是单因素方差分析的均方误差, 参数基本与 F&P 值图一致, 绘制分析方法基本一致。MSE 值可以用来衡量重复评估是否是效果不错的。并且 MSE 值越低越好。MSE 值越低, 说明重复评估的评分差距越小, 评估得分越可信。

第 13 章 P 值-均方误差图 (P-MSE Plots)

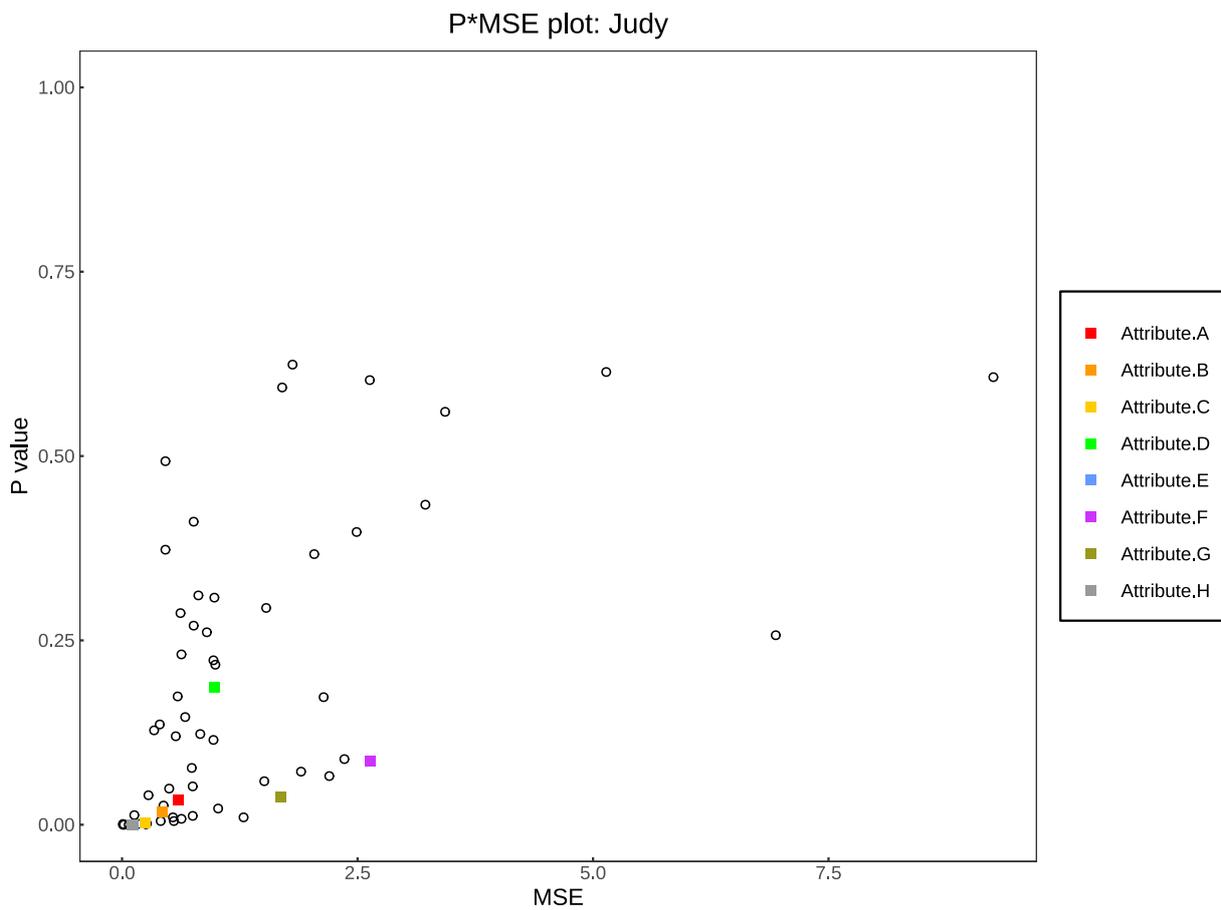
P 值-均方误差图是结合了单因素方差分析 P 值与 MSE 值, 以 MSE 值为 x 轴, P 值为 y 轴, 绘制散点图。

13.1 基于给定评估者 (P_MSE_ass)

13.1.1 函数介绍

函数 `P_MSE_ass` 可以实现绘制突出颜色标记给定评估者的 P 值-均方误差图。见如下示例:

```
1 # 运行函数源文件
2 source("../Univariate-code/P_MSE_ass.R", encoding="utf-8")
3 # 调用该函数
4 P_MSE_ass(data=data, assessor="Judy", lang="en")
```



参数介绍

data: 原始数据

assessor: 需要绘制的评估者名称

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

13.1.2 可视化分析

P 值-均方误差会绘制所有感官属性与所有评估者两两基于对样本的单因素方差分析的 P 值 (作为 y 轴), MSE 值 (作为 x 轴), 然后对于给定的评估者的散点进行颜色标记。

图表分析

P 值和 MSE 值分别是分辨率 (检测样本之间差异的能力) 和重复性 (检测重复评估效果) 的衡量标准, P 值和 MSE 值都越小, 评估者的表现就越好。因而, P 值-均方误差图可以快速、轻松地检测出哪些评估员在哪一属性上表现不佳。

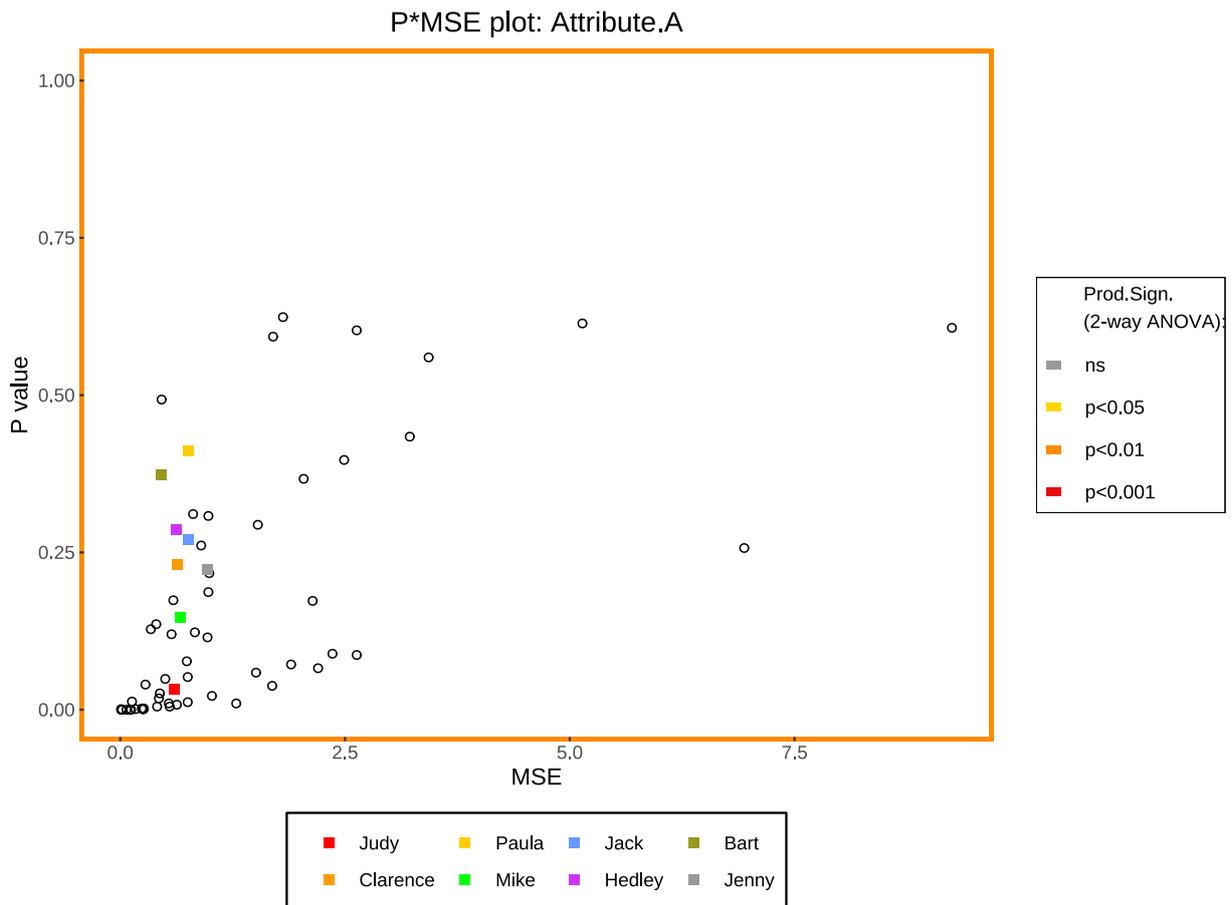
13.2 基于给定感官属性 (P_MSE_attr)

13.2.1 函数介绍

函数 P_MSE_attr 则可以实现绘制突出颜色标记给定感官属性的 P 值-均方误差图。见如下示例：

```

1 # 运行函数源文件
2 source("../Univariate-code/P_MSE_attr.R", encoding="utf-8")
3 # 调用该函数
4 P_MSE_attr(data=data, attribute="Attribute.A", lang="en")
    
```



参数介绍

data: 原始数据

attribute: 需要绘制的感官属性名称

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

13.2.2 可视化分析

P 值-均方误差会绘制所有感官属性与所有评估者两两基于对样本的单因素方差分析的 P 值 (作为 y 轴), MSE 值 (作为 x 轴), 然后对于给定的感官属性的散点进行颜色标记。关于绘图面板的边框颜色显著性标记, 前面已有介绍, 请看[7.2.2](#)

图表分析

P 值和 MSE 值分别是分辨率 (检测样本之间差异的能力) 和重复性 (检测重复评估效果) 的衡量标准, P 值和 MSE 值都越小, 评估者的表现就越好。因而, P 值-均方误差图可以快速、轻松地检测出哪些评估员在哪一属性上表现不佳。

第三部分 多变量 (Multivariate)

第 14 章 Tucker-1 方法绘图 (Tucker-1 Plots)

Tucker-1 是有一种多变量分析方法，其本质是对特定格式的展开数据矩阵 $X_{i,av}$ 所进行的主成分分析方法 (PCA)，想深入了解该方法可以参考 [3, 4]。 $X_{i,av}$ 表示维度为 $J \times K$ 的一个评估者的矩阵，其中样本得分基于各重复的平均值 (可修改)，因此在索引中用 “av” 表示，这意味着这个展开矩阵的维数是 $J \times (i \times K)$ (J 为样本数， i 为评估者数量， K 为感官属性数量)。

14.1 共同得分 (Tucker1_common_score)

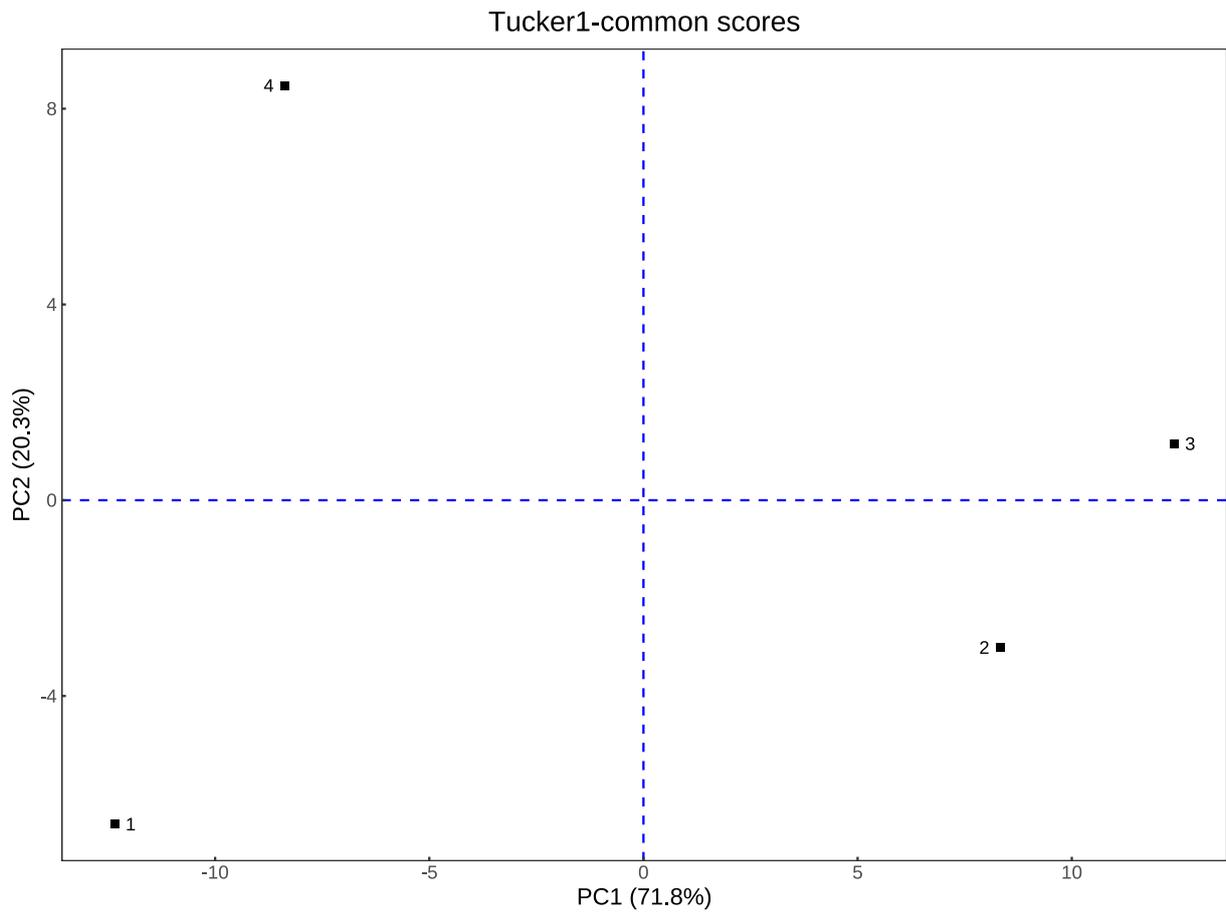
14.1.1 函数介绍

Tucker1_common_score 函数可以绘制 Tucker-1 方法的共同得分图，见如下示例：

```

1 # 运行函数源文件
2 source("../Multivariate-code/Tucker1_common_score.R",encoding="utf-8")
3 # 调用该函数
4 Tucker1_common_score(data=data,pc_x=1,pc_y=2,scale=FALSE,
5                       model_on="sample_averages",color_point="black",
6                       equal_axis=FALSE,color_line="blue",color_label="black",
7                       label_size=3,lang="en")

```



参数介绍

data: 原始数据

pc_x: x 轴的主成分, 例如 1 表示第一主成分。默认为 1, 表示第一主成分

pc_y: y 轴的主成分, 例如 2 表示第二主成分。默认为 2, 表示第二主成分

scale: 逻辑值, 进行主成分分析时, 是否首先对样本矩阵标准化, 默认为 FALSE, 即不标准化

model_on: 模型样本矩阵形式, 默认为 "sample_averages", 可选值有:

—————"sample_averages": 表示为按照重复评估得分均值构造样本矩阵进行主成分分析;

—————"sample_replicates": 表示所有重复评估得分都参与构造样本矩阵进行主成分分析

equal_axis: 逻辑值, 图形中 x 轴与 y 轴距离尺度是否相等, 默认为 FALSE, 即不相等

color_point: 图中的样本散点颜色, 默认为黑色

color_line: 图中的虚线颜色, 默认为蓝色

color_label: 图中散点的文本标签颜色, 默认为黑色

label_size: 图中的文本标签字体大小, 默认值为 3, 数字越大, 字体越大

lang: 指定中英文模板, 默认为 "en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

14.1.2 可视化分析

上述示例中, 散点表示的是四个样本所代表的第一主成分与第二主成分的共同得分值。由参数介绍可以知道, 研究人员可以修改 (*pc_x*) 与 (*pc_y*) 参数自由控制想要绘制的主成分, 而这个主成分最大值为样本的数量。

需要注意的是, 就本示例而言, 在 PanelCheck 软件中, 每个坐标轴可以自由选择 4 个主成分, 但是在笔者的函数中, 每个坐标轴只能选择 1 到 3, 三个主成分。这是因为在主成分分析中, 原本是有等同于样本数量个主成分存在, 但是每个主成分的 **方差贡献率** 是不同的, 有大小之分, 某个主成分方差贡献率越高, 说明其很好的捕捉了数据的变异性。PanelCheck 不会剔除方差贡献率为 0 的主成分, 仍然将其绘制出来, 而笔者的函数会自动剔除贡献率为 0 的主成分, 并且对于研究人员而言, 方差贡献率为 0 的主成分基本没有什么研究意义, 可以忽略。而本示例中的四个主成分中, 刚好第四主成分方差贡献率为 0, 所以就没有考虑。

图表分析

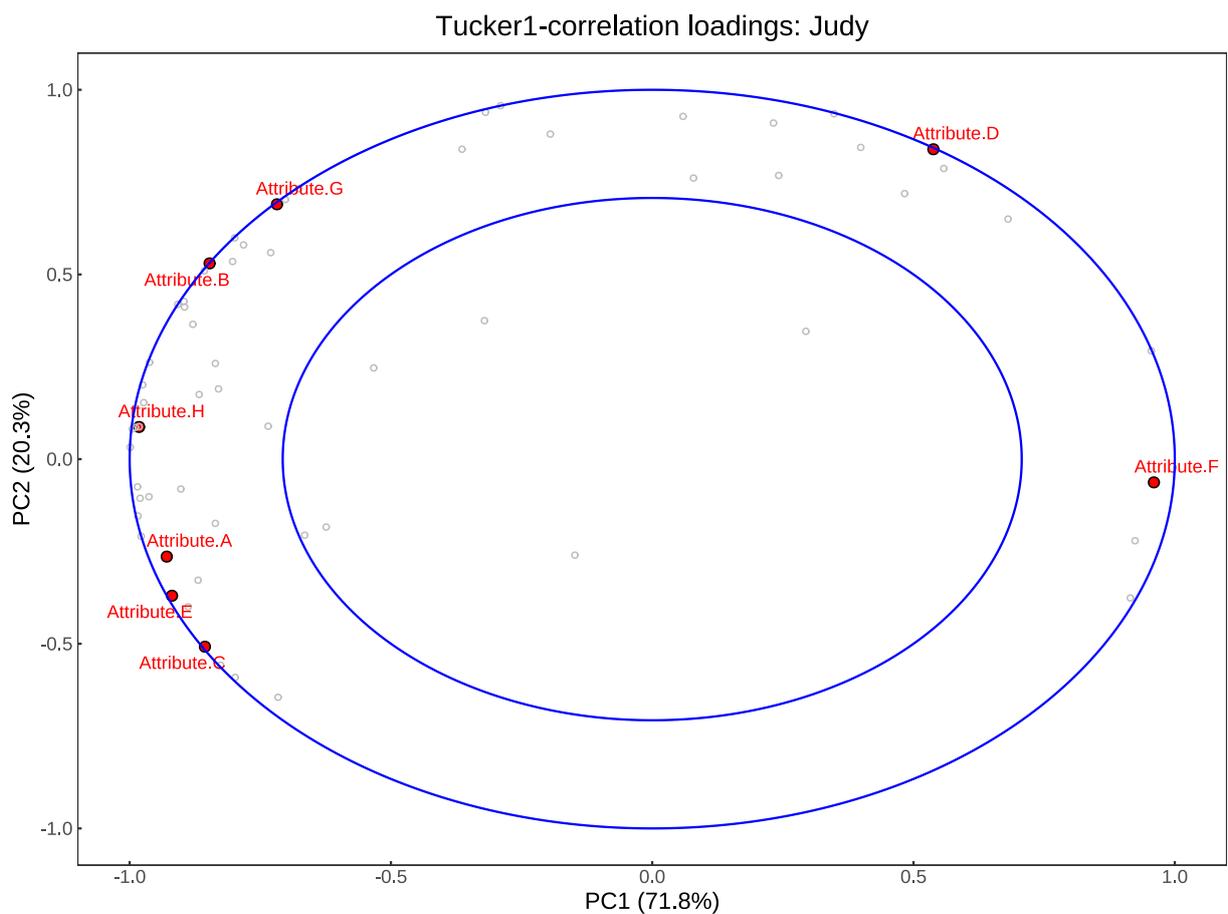
共同得分图显示了被测试的 J 个样本之间的相互关系, 即它显示了样本之间沿着找到的主成分的相似和不同之处。这个曲线图没有给出关于评估员或整体性能的直接信息, 但它是一个有价值的可视化工具, 可以帮助研究人员粗略而快速地调查面板是否能通过考虑所解释的差异来区分样本。如果前几个 (通常是两个) 主成分 (PC) 的解释方差相对较高, 则数据中存在较大的变异性, 这可能表明该小组在样本之间有很好的区分能力。注意, Tucker-1 共同得分图的解释方差通常在前几个 PC 中比普通一致性平均矩阵上的 PCA 的解释方差要低一些。这是因为 Tucker-1 分析基于更多变量, 因此数据中存在更多噪声。

14.2 基于评估者的相关载荷图 (*Tucker1_corrload_ass*)

14.2.1 函数介绍

我们使用 *Tucker1_corrload_ass* 绘制基于评估者的相关载荷图，见如下示例：

```
1 # 运行函数源文件
2 source("../Multivariate-code/Tucker1_corrload_ass.R",encoding="utf-8")
3 # 调用该函数
4 Tucker1_corrload_ass(data=data,assessor="Judy",pc_x=1,pc_y=2,scale=FALSE,
5                       model_on="sample_averages",equal_axis=FALSE,lang="en",
6                       color_circle="blue",color_ass="red",text_size=3)
```



参数介绍

data: 原始数据

assessor: 评估者名称

pc_x: x 轴的主成分, 例如 1 表示第一主成分。默认为 1, 表示第一主成分

pc_y: y 轴的主成分, 例如 2 表示第二主成分。默认为 2, 表示第二主成分

scale: 逻辑值, 进行主成分分析时, 是否首先对样本矩阵标准化, 默认为 FALSE, 即不标准化

model_on: 模型样本矩阵形式, 默认为 "sample_averages", 可选值有:

—————"sample_averages": 表示为按照重复评估得分均值构造样本矩阵进行主成分分析;

—————"sample_replicates": 表示所有重复评估得分都参与构造样本矩阵进行主成分分析

equal_axis: 逻辑值, 图形中 x 轴与 y 轴距离尺度是否相等, 默认为 FALSE, 即不相等

color_circle: 图中的两个圆圈的线条颜色, 默认为蓝色

color_ass: 图中的该评估者的属性散点颜色与文本标签颜色, 默认为红色

text_size: 图中的文本标签字体大小, 默认值为 3, 数字越大, 字体越大。

lang: 指定中英文模板, 默认为 "en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

14.2.2 可视化分析

相关载荷图提供了每个评估员和整个感官面板的性能信息。该图包含 $i \times K$ 个点, 每个点代表一个评估者-属性组合。通过突出显示不同的点, 无论是一个评估员的点, 还是一个属性的点, 都可以直观地显示单个评估员或整个评审团的表现, 点在图中的位置反应了关于个人或评估小组整体的表现如何。

图表分析

特定评估人员的感官属性所包含的噪音越多, 点看起来就越接近原点, 即绘图的中间。评估者的属性包含的信息越系统, 越完整, 它看起来就越接近外椭圆 (100% 解释了该属性的方差)。内椭圆表示 50% 的可解释方差, 可视为一个属性至少应该被视为足够好的多少可解释差异的经验下限。建议也参考更高的 PC, 因为有些评估员可能在 PC1 和 PC2 之外的其他维度上有很大的系统差异。

14.3 基于感官属性的相关载荷图 (*Tucker1_corrload_attr*)

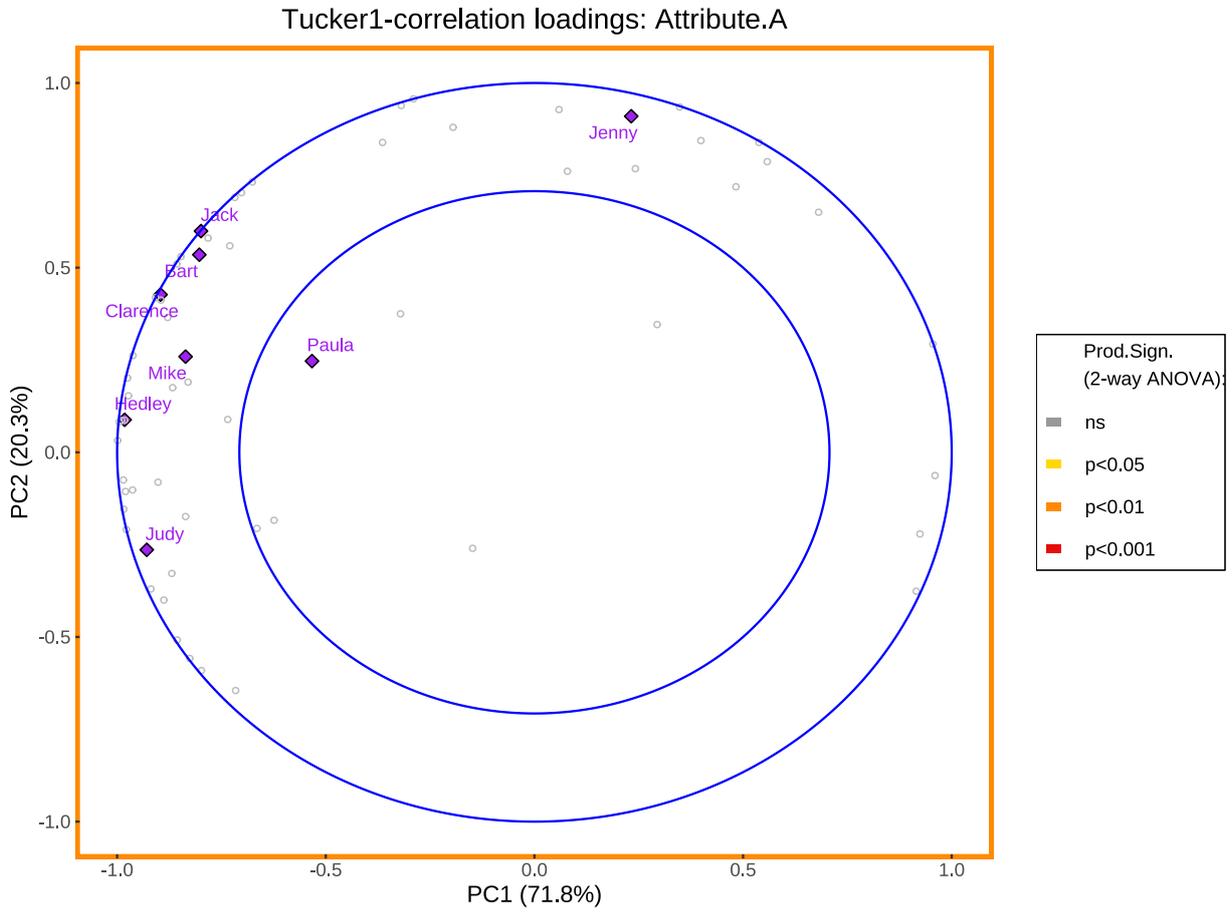
14.3.1 函数介绍

与上一小节的相关载荷图基本类似, 不同于是以感官属性为绘图角度, 见如下示例:

```

1 # 运行函数源文件
2 source("../Multivariate-code/Tucker1_corrload_attr.R",encoding="utf-8")
3 # 调用该函数
4 Tucker1_corrload_attr(data=data,attribute="Attribute.A",pc_x=1,pc_y=2,
5                       scale=FALSE,model_on="sample_averages",
6                       equal_axis=FALSE,color_circle="blue",color_attr="purple",
7                       text_size=3,lang="en")

```



参数介绍

data: 原始数据

attribute: 感官属性名称

pc_x: x 轴的主成分, 例如 1 表示第一主成分。默认为 1, 表示第一主成分

pc_y: y 轴的主成分, 例如 2 表示第二主成分。默认为 2, 表示第二主成分

scale: 逻辑值, 进行主成分分析时, 是否首先对样本矩阵标准化, 默认为 FALSE, 即不标准化

model_on: 模型样本矩阵形式, 默认为 "sample_averages", 可选值有:

_____ "sample_averages": 表示为按照重复评估得分均值构造样本矩阵进行主成分分析;

_____ "sample_replicates": 表示所有重复评估得分都参与构造样本矩阵进行主成分分析

equal_axis: 逻辑值, 图形中 x 轴与 y 轴距离尺度是否相等, 默认为 FALSE, 即不相等

color_circle: 图中的两个圆圈的线条颜色, 默认为蓝色

color_attr: 图中的该属性的所有评估者散点颜色与文本标签颜色, 默认为紫色

text_size: 图中的文本标签字体大小, 默认值为 3, 数字越大, 字体越大。

lang: 指定中英文模板, 默认为 "en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

14.3.2 可视化分析

相关载荷图提供了每个评估员和整个感官面板的性能信息。关于图形面板边框颜色标记与右边图例, 请看 7.2.2。该图包含 $i \times K$ 个点, 每个点代表一个评估者-属性组合。通过突出显示不同的点, 无论是一个评估员的点, 还是一个属性的点, 都可以直观地显示单个评估员或整个评审团的表现, 点在图中的位置反应了关于个人或评估小组整体的表现如何。关于更多统计方面细节, 可以参考 [5]。

图表分析

特定评估人员的感官属性所包含的噪音越多, 点看起来就越接近原点, 即绘图的中间。评估者的属性包含的信息越系统, 越完整, 它看起来就越接近外椭圆 (100% 解释了该属性的方差)。内椭圆表示 50% 的可解释方差, 可视为一个属性至少应该被视为足够好的多少可解释差异的经验下限。建议也参考更高的 PC, 因为有些评估员可能在 PC1 和 PC2 之外的其他维度上有很大的系统差异。

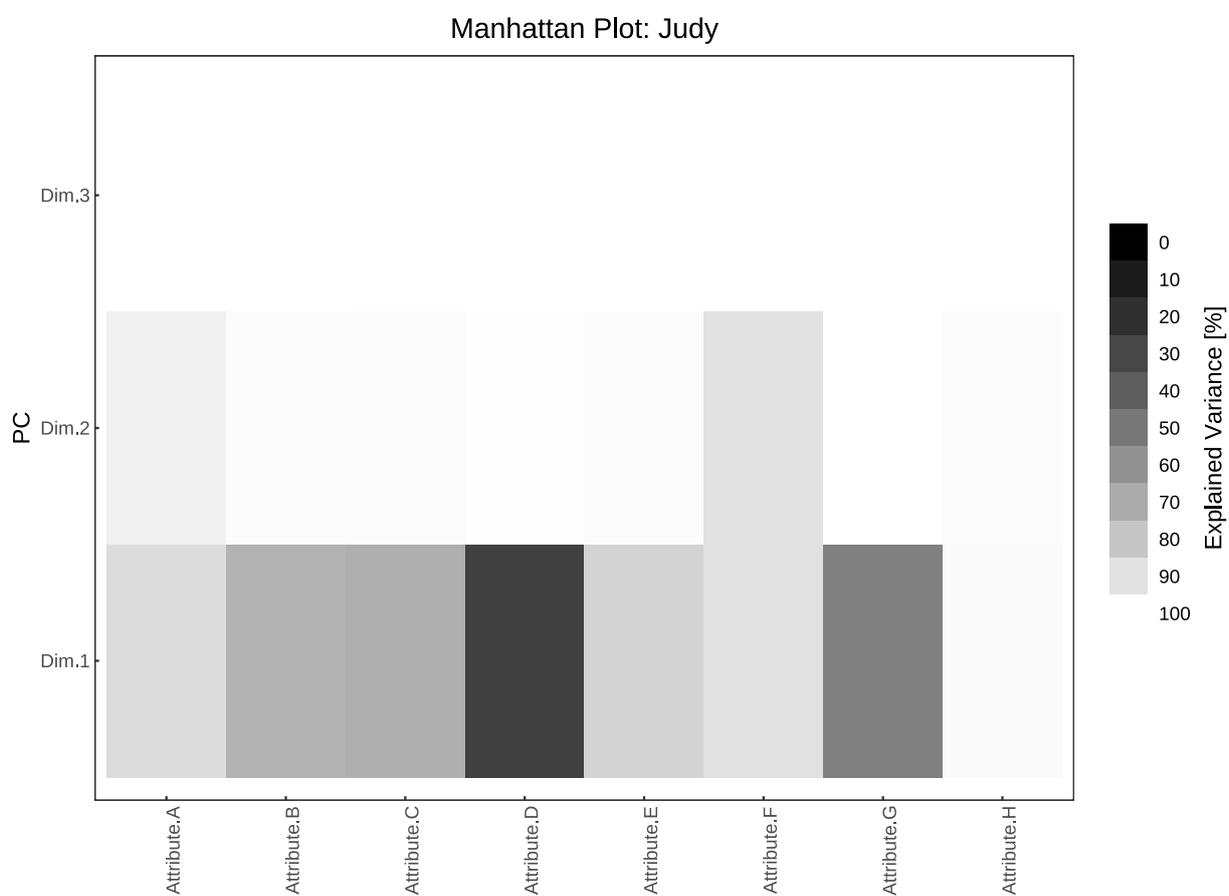
第 15 章 曼哈顿图 (Manhattan Plots)

15.1 曼哈顿图 (Manhattan_plot)

15.1.1 函数介绍

我们使用 `Manhattan_plot` 函数轻松实现曼哈顿图，请看如下示例：

```
1 # 运行函数源文件
2 source("../Multivariate-code/Manhattan_plot.R", encoding="utf-8")
3 # 调用该函数
4 Manhattan_plot(data=data, by="attribute", assessor="Judy", scale=FALSE,
5               color_high="white", color_low="black", angle=90, lang="en")
```



参数介绍

data: 原始数据

by: 指定是按照属性分类可视化 (x 轴为感官属性), 还是按照评估者分类 (x 轴为评估者), 默认为感官属性:
 _____”attribute”: 按照感官属性可视化
 _____”assessor”: 按照评估者可视化

assessor: 评估者名称, 当且仅当 by=”assessor” 时设置此参数

attribute: 感官属性名, 当且仅当 by=”attribute” 时设置此参数

scale: 逻辑值, 是否将得分矩阵标准化, 然后在进行 PCA, 默认为 FALSE, 即不进行标准化

color_high: 设置解释方差 100

color_low: 设置解释方差为 0 时的颜色映射, 默认为黑色

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为”en”:
 _____”cn”: 数据为中文模板
 _____”en”: 数据为英文模板

15.1.2 可视化分析

曼哈顿图是对数据矩阵 $X_{i,av}$ 应用 PCA (主成分分析), 然后在图中可视化每个属性的解释方差。假如面板 (原始数据) 中由 $i = 10$ 名评估员组成, 则在给定 $K = 9$ 个属性的情况下, 解释方差的数量将为 $10 * 9 = 90$ 个。注意, 与 PanelCheck 不同, 笔者将曼哈顿图的方差贡献率的累计方向作了个调换——由从上到下变为从下到上, 个人认为这更适合一般人的视觉观感。

图表分析

曼哈顿图以灰色阴影 (默认参数) 可视化地显示了主成分 (垂直轴) 解释每个属性和每个评估者的变异性有多大。深色色调表示只解释了一小部分差异, 而浅色色调则相反。极值点是 “黑色” (0% 解释方差) 和 “白色” (100% 解释方差)。一般来说, PC1 的颜色会变深, 然后从下到上增加一个主成分, 颜色就会变浅, 因为显示的解释差异在每个 PC 上都是累积的。PC3 处的解释方差是 PC1、PC2 和 PC3 的解释方差之和。曼哈顿图区块中特定评估者-属性组合的色调越浅, 给出的系统变化就越多。



注意, 在多变量绘图部分中, PanelCheck 提供了非常灵活的可以让用户操控的选项, 类似于图 1。由于篇幅限制, 所展示示例只是函数功能很微小的一部分, 读者需要自行修改函数的所有参数, 体会其对于 PanelCheck 的近乎完美的复刻, 同时, 也要观察两者的不同之处。

第四部分 一致性 (Consensus)

本部分内容是基于将原始数据转化为特定的一致性数据 (Consensus data), 然后进行主成分分析 (PCA), 对 PCA 结果绘制不同的图表。PanelCheck 将该部分内容按数据处理方式不同, 分为:

- **Original**: 对原始 Consensus data 直接进行主成分分析;
- **Standardized**: 将原始 Consensus data 进行标准化处理, 再进行主成分分析;
 - 如果某些评估员没有使用相同尺度量表, 或者在一个或多个属性的跨度上与其他评估员意见不一致, 数据标准化会很有用。通常, 通过应用标准化可以改善结果, 但不一定比不处理更准确。
- **STATIS**: 利用 STATIS 方法处理 Consensus data, 再进行主成分分析;
 - 与 PCA 一样, STATIS 方法将数据集中的变化分解为主成分, 并根据评价者对样本的评价和样本 (或属性) 的平均得分在主成分空间中对评价者进行监控;
 - STATIS 方法也有两种情形, 一个是基于协方差矩阵 (Covariance), 一个是基于相关矩阵 (Correlation);

关于 STATIS 方法, 其数据处理过程较为复杂, 这里不再叙述, 想了解具体细节可以参考:

- <https://personal.utdallas.edu/herve/Abdi-Statist2007-pretty.pdf>.

由于仅仅是三种数据处理方式不同, 绘图基本都是类似的, 所以笔者将按照不同的绘图函数依次对其介绍, 每个函数中的 `type` 参数, 可以指定三种不同的数据处理方式。

第 16 章 评估者权重 (Assessor Weights)

16.1 评估者权重图 (Ass_weight_plot)

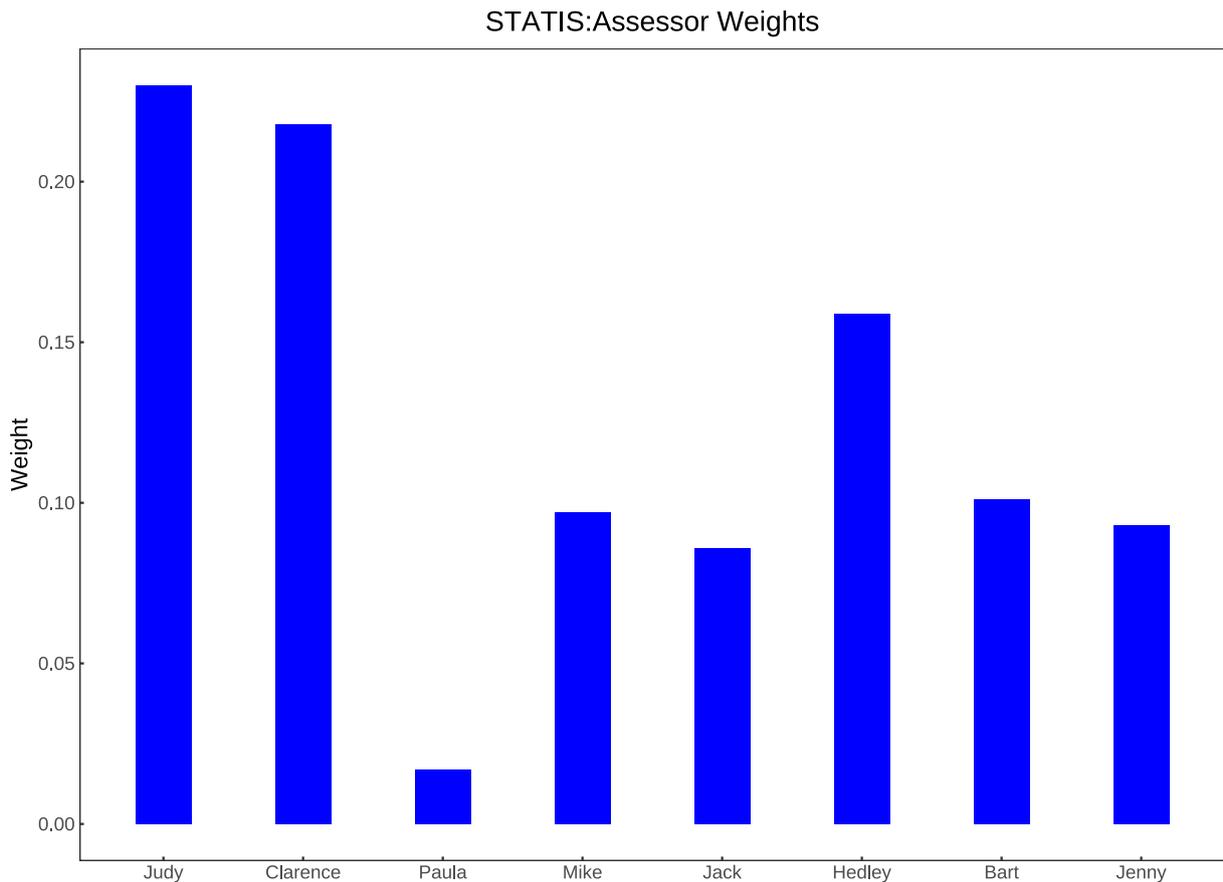
16.1.1 函数介绍

注意! 该类图只存在于 STATIS 方法中! 在 STATIS 方法中, 有一个很重要的参数叫做 评估者权重, 它是计算 STATIS 方法下的 Consensus data 的关键参数。本函数 `Ass_weight_plot` 的目的就是将权重值以条形图的形式可视化。见如下示例:

```

1 # 运行函数源文件
2 source("../Consensus-code/Ass_weight_plot.R", encoding="utf-8")
3 # 调用该函数
4 Ass_weight_plot(data=data, based_on="cov", color="blue", lang="en")

```



参数介绍

data: 原始数据

based_on: 指定基于什么矩阵进行 STATIS 分析, 默认值为"cov":

—————"cov": 指定进行基于协方差 (Covariance) 的 STATIS 分析

—————"cor": 指定进行基于相关性 (Correlation) 的 STATIS 分析

color: 图中的柱子颜色, 默认为蓝色

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

16.1.2 可视化分析

图表分析

评估者权重图很直观的展示了各评估者的权重大小, 并作了比较。在 **STATIS** 方法中, 评估者的权重越大, 说明该评估者在所有评估者中更具有代表性, 一致性。

! 从下一小节开始, 所有函数都有 `type` 参数来指定三种处理方法中的一个!

第 17 章 一致性数据 (Consensus data)

17.1 一致性数据 (Consensus data)

17.1.1 函数介绍

该函数根据指定的三种处理方法中一个用来将原始数据处理为进行主成分分析的一致性矩阵 (Consensus data), 见如下示例:

```

1 # 运行函数源文件
2 source("../Consensus-code/Consen_data.R", encoding="utf-8")
3 # 调用该函数
4 Consen_data(data=data, type="original", based_on="cov", lang="en")

```

Samples	Attribute.A	Attribute.B	Attribute.C	Attribute.D	Attribute.E	Attribute.F	Attribute.G	Attribute.H
Sample 1	5.97	5.38	6.21	1.94	6.22	2.41	4.61	5.92
Sample 2	5.13	4.58	4.79	3.22	2.92	4.57	1.97	2.28
Sample 3	4.97	4.62	2.35	2.96	2.62	5.53	1.93	1.77
Sample 4	6.44	6.42	4.94	4.25	4.42	4.16	6.21	5.71

参数介绍

data: 原始数据

type: 设置对原始数据进行处理的方法, 默认值为"original":

———"original": 对原始数据进行简单平均构造共性矩阵

———"standardized": 对原始数据进行标准化再平均, 构造共性矩阵

———"STATIS": 利用 STATIS 方法对原始数据处理, 构造共性矩阵

based_on: 当且仅当 *type*="STATIS" 时, 参数有效, 默认值为"cov":

———"cov": 指定进行基于协方差 (Covariance) 的 STATIS 分析

———"cor": 指定进行基于相关性 (Correlation) 的 STATIS 分析

lang: 指定中英文模板, 默认为"en":

———"cn": 数据为中文模板

———"en": 数据为英文模板

17.1.2 可视化分析

该函数只是用于让研究人员查看经过指定方法处理之后的 Consensus data。

第 18 章 主成分得分 (PCA Scores)

18.1 主成分得分 (PCA_scores)

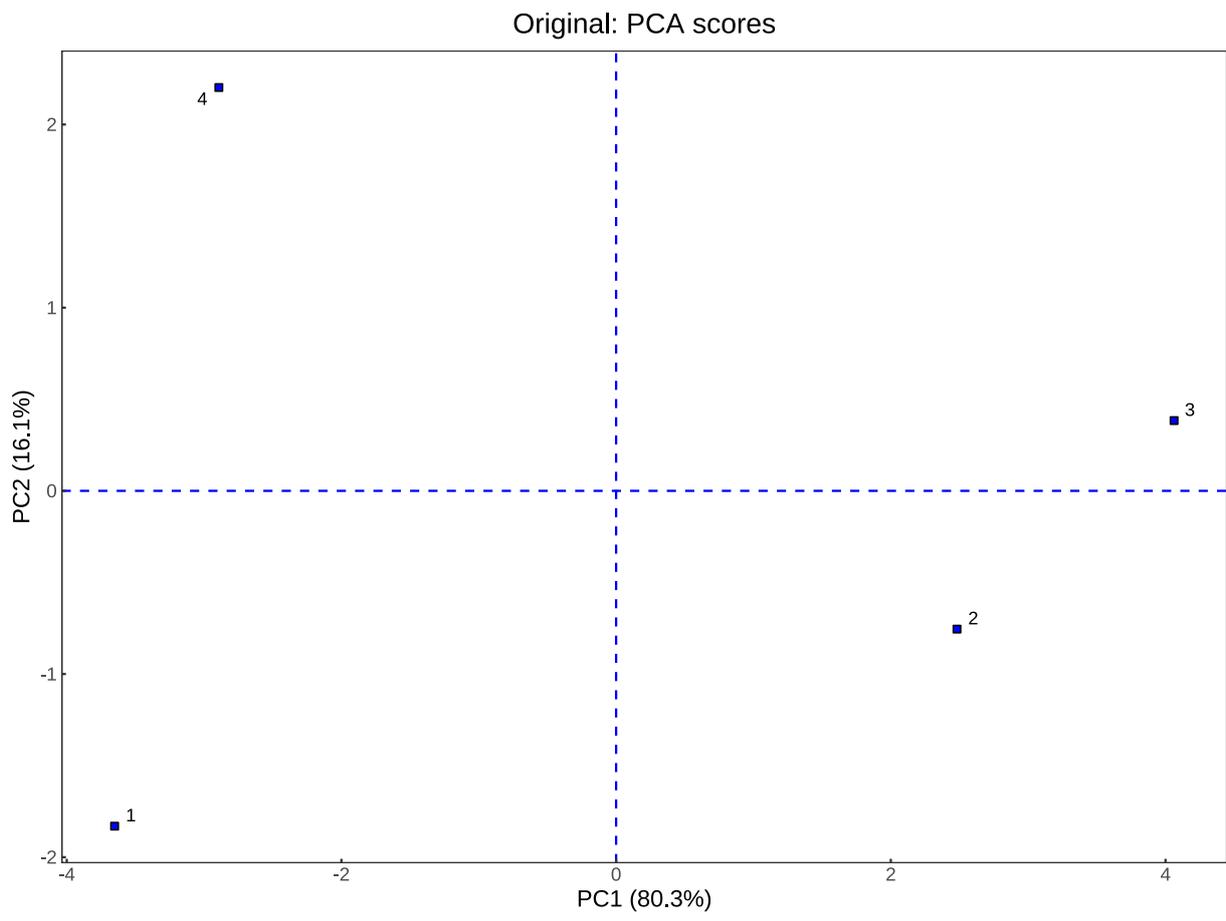
18.1.1 函数介绍

该函数对进行主成分分析 (PCA) 所得出的主成分得分进行可视化, 以两个主成分得分分别作为 x 轴与 y 轴, 绘制所有样本。请见如下示例:

```

1 # 运行函数源文件
2 source("../Consensus-code/PCA_scores.R",encoding="utf-8")
3 # 调用该函数
4 PCA_scores(data=data,pc_x=1,pc_y=2,type="original",based_on="cov",
5             equal_axis=FALSE,color_line="blue",lang="en",
6             color_text="black",color_pt="blue",text_size=3)

```



参数介绍

data: 原始数据

pc_x: x 轴的主成分, 例如 1 表示第一主成分。默认为 1, 表示第一主成分

pc_y: y 轴的主成分, 例如 2 表示第二主成分。默认为 2, 表示第二主成分

type: 设置对原始数据进行处理的方法, 默认值为 "original":

—— "original": 对原始数据进行简单平均构造共性矩阵

—— "standardized": 对原始数据进行标准化再平均, 构造共性矩阵

—— "STATIS": 利用 STATIS 方法对原始数据处理, 构造共性矩阵

based_on: 当且仅当 *type*="STATIS" 时, 参数有效, 默认值为 "cov":

—— "cov": 指定进行基于协方差 (Covariance) 的 STATIS 分析

—— "cor": 指定进行基于相关性 (Correlation) 的 STATIS 分析

equal_axis: 逻辑值, 图形中 x 轴与 y 轴距离尺度是否相等, 默认为 FALSE, 即不相等

color_pt: 图中的样本散点颜色, 默认为蓝色

color_line: 图中的虚线颜色, 默认为蓝色

olor_text: 图中散点的文本标签颜色, 默认为黑色

text_size: 图中的文本标签字体大小, 默认值为 3, 数字越大, 字体越大

lang: 指定中英文模板, 默认为 "en":

—— "cn": 数据为中文模板

—— "en": 数据为英文模板

18.1.2 可视化分析

PCA 得分图与 14.1 及其相似, 两者都是使用了主成分分析方法, 不同点在于前期的构造数据矩阵的方法不同, 所以分析方法与之基本类似, 限于篇幅这里不再重复。

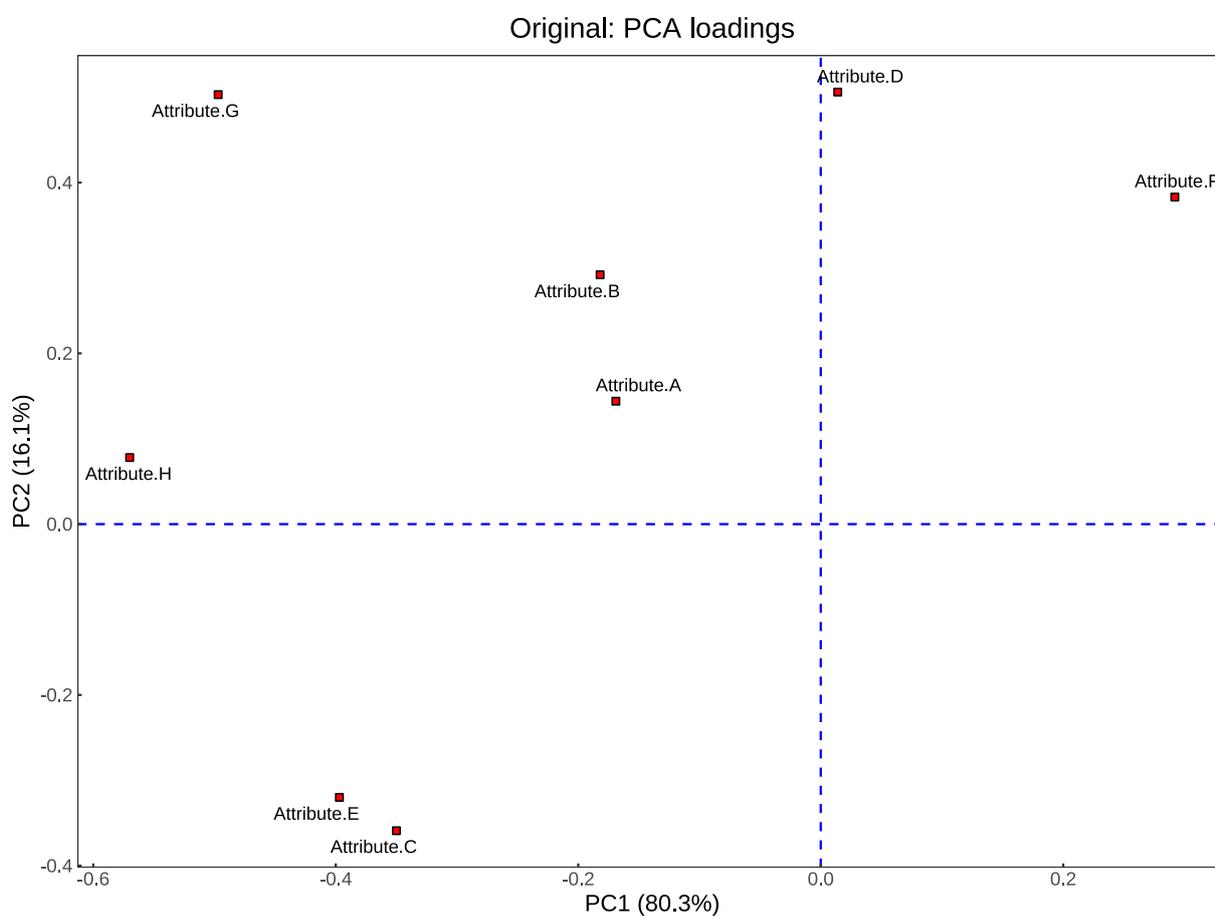
第 19 章 主成分载荷 (PCA Loadings)

19.1 主成分载荷 (PCA_loadings)

19.1.1 函数介绍

与绘制样本主成分不同，该函数利用主成分作为 x 轴与 y 轴，绘制所有感官属性的主成分载荷，见如下示例：

```
1 # 运行函数源文件
2 source("../Consensus-code/PCA_loadings.R", encoding="utf-8")
3 # 调用该函数
4 PCA_loadings(data=data, pc_x=1, pc_y=2, type="original", based_on="cov",
5               equal_axis=FALSE, color_line="blue", lang="en",
6               color_text="black", color_pt="red", text_size=3)
```



参数介绍

data: 原始数据

pc_x: x 轴的主成分, 例如 1 表示第一主成分。默认为 1, 表示第一主成分

pc_y: y 轴的主成分, 例如 2 表示第二主成分。默认为 2, 表示第二主成分

type: 设置对原始数据进行处理的方法, 默认值为 "original":

—— "original": 对原始数据进行简单平均构造共性矩阵

—— "standardized": 对原始数据进行标准化再平均, 构造共性矩阵

—— "STATIS": 利用 STATIS 方法对原始数据处理, 构造共性矩阵

based_on: 当且仅当 *type*="STATIS" 时, 参数有效, 默认值为 "cov":

—— "cov": 指定进行基于协方差 (Covariance) 的 STATIS 分析

—— "cor": 指定进行基于相关性 (Correlation) 的 STATIS 分析

equal_axis: 逻辑值, 图形中 x 轴与 y 轴距离尺度是否相等, 默认为 FALSE, 即不相等

color_pt: 图中的样本散点颜色, 默认为红色

color_line: 图中的虚线颜色, 默认为蓝色

olor_text: 图中散点的文本标签颜色, 默认为黑色

text_size: 图中的文本标签字体大小, 默认值为 3, 数字越大, 字体越大

lang: 指定中英文模板, 默认为 "en":

—— "cn": 数据为中文模板

—— "en": 数据为英文模板

19.1.2 可视化分析

图表分析

主成分载荷的绝对值越大, 对于主成份的影响就越大。而这种影响则可以通过载荷所代表的点到原点之间的距离来进行衡量。如果某个点在进行主成份分析之后位置在原点附近, 则说明这个该特征属性的波动对于样本之间的区别贡献不大, 反之, 贡献显著。

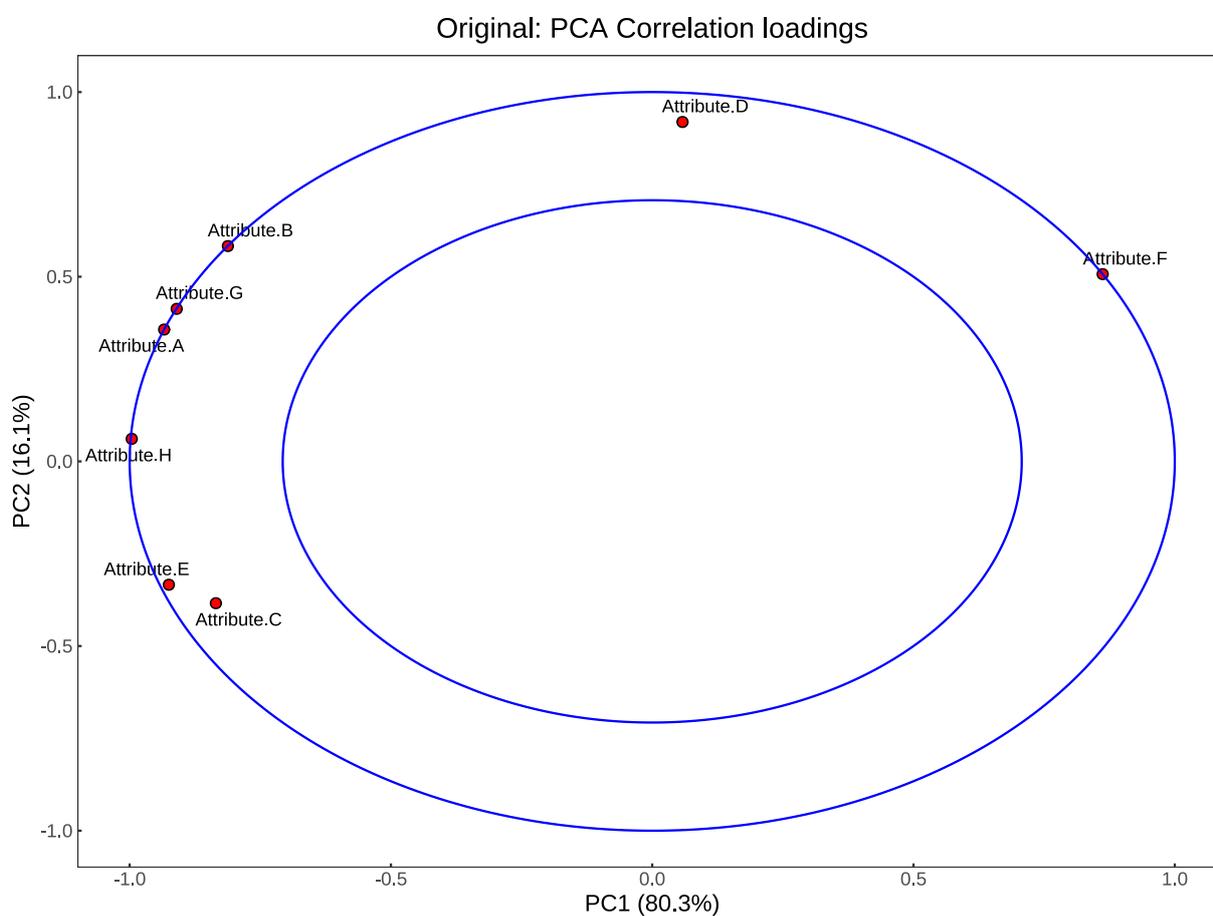
第 20 章 主成分相关载荷 (PCA Correlation Loadings)

20.1 主成分相关载荷 (PCA_cor_loading)

20.1.1 函数介绍

相关载荷与 Tucker-1 方法14.2的相关载荷除了前期数据处理方法不一样之外，其他基本一致，见如下示例：

```
1 # 运行函数源文件
2 source("../Consensus-code/PCA_cor_loading.R", encoding="utf-8")
3 # 调用该函数
4 PCA_cor_loading(data=data, pc_x=1, pc_y=2, type="original", based_on="cov",
5                 equal_axis=FALSE, color_line="blue", lang="en",
6                 color_text="black", color_pt="red", text_size=3)
```



参数介绍

data: 原始数据

pc_x: x 轴的主成分, 例如 1 表示第一主成分。默认为 1, 表示第一主成分

pc_y: y 轴的主成分, 例如 2 表示第二主成分。默认为 2, 表示第二主成分

type: 设置对原始数据进行处理的方法, 默认值为"original":

—————"original": 对原始数据进行简单平均构造共性矩阵

—————"standardized": 对原始数据进行标准化再平均, 构造共性矩阵

—————"STATIS": 利用 STATIS 方法对原始数据处理, 构造共性矩阵

based_on: 当且仅当 *type*="STATIS" 时, 参数有效, 默认值为"cov":

—————"cov": 指定进行基于协方差 (Covariance) 的 STATIS 分析

—————"cor": 指定进行基于相关性 (Correlation) 的 STATIS 分析

equal_axis: 逻辑值, 图形中 x 轴与 y 轴距离尺度是否相等, 默认为 FALSE, 即不相等

color_pt: 图中的样本散点颜色, 默认为红色

color_line: 图中的虚线颜色, 默认为蓝色

olor_text: 图中散点的文本标签颜色, 默认为黑色

text_size: 图中的文本标签字体大小, 默认值为 3, 数字越大, 字体越大

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

20.1.2 可视化分析

本节所示主成分相关载荷图显示了各感官属性在两个主成分下的相关载荷系数, 分析方法与 14.2 基本一致。

图表分析

特定评估人员的感官属性所包含的噪音越多, 点看起来就越接近原点, 即绘图的中间。评估者的属性包含的信息越系统, 越完整, 它看起来就越接近外椭圆 (100% 解释了该属性的方差)。内椭圆表示 50% 的可解释方差, 可视为一个属性至少应该被视为足够好的多少可解释差异的经验下限。

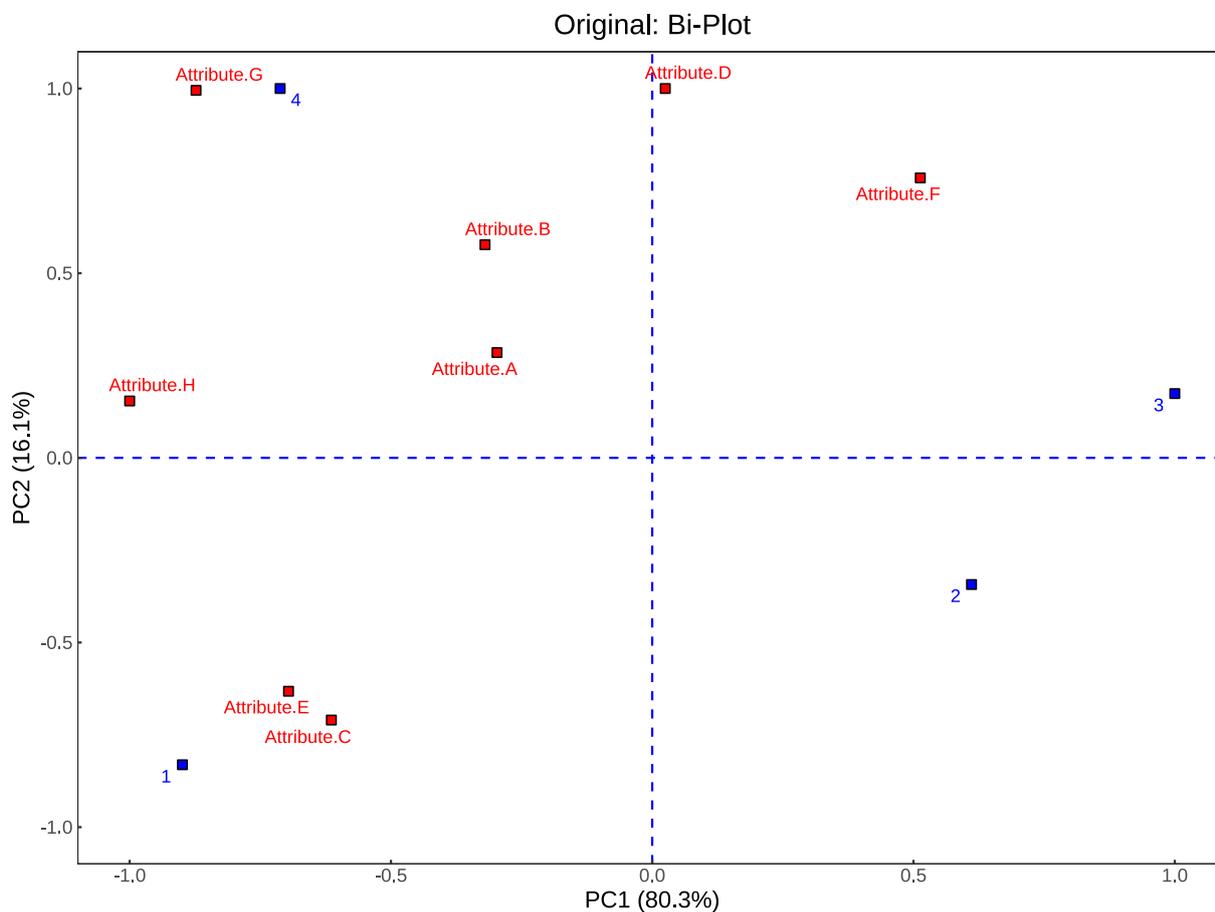
第 21 章 双标图 (Bi-Plot)

21.1 双标图 (Biplot)

21.1.1 函数介绍

双标图 (Bi-plot) 将各感官属性与各样本绘制在同一绘图面板上, 可以更直观的进行整体分析, 见如下示例:

```
1 # 运行函数源文件
2 source("../Consensus-code/Biplot.R",encoding="utf-8")
3 # 调用该函数
4 Biplot(data=data,pc_x=1,pc_y=2,type="original",based_on="cov",
5         equal_axis=FALSE,color_line="blue",lang="en",
6         color_text=c("red","blue"),color_pt=c("red","blue"),
7         text_size=3)
```



参数介绍

data: 原始数据

pc_x: x 轴的主成分, 例如 1 表示第一主成分。默认为 1, 表示第一主成分

pc_y: y 轴的主成分, 例如 2 表示第二主成分。默认为 2, 表示第二主成分

type: 设置对原始数据进行处理的方法, 默认值为 "original":

—— "original": 对原始数据进行简单平均构造共性矩阵

—— "standardized": 对原始数据进行标准化再平均, 构造共性矩阵

—— "STATIS": 利用 STATIS 方法对原始数据处理, 构造共性矩阵

based_on: 当且仅当 type="STATIS" 时, 参数有效, 默认值为 "cov":

—— "cov": 指定进行基于协方差 (Covariance) 的 STATIS 分析

—— "cor": 指定进行基于相关性 (Correlation) 的 STATIS 分析

equal_axis: 逻辑值, 图形中 x 轴与 y 轴距离尺度是否相等, 默认为 FALSE, 即不相等

color_pt: 图中的样本散点颜色, 默认为红色和蓝色

color_line: 图中的圆圈颜色, 默认为蓝色

olor_text: 图中散点的文本标签颜色, 默认为红色和蓝色

text_size: 图中的文本标签字体大小, 默认值为 3, 数字越大, 字体越大

lang: 指定中英文模板, 默认为 "en":

—— "cn": 数据为中文模板

—— "en": 数据为英文模板

21.1.2 可视化分析

图表分析

使用双标图可在一个图形中评估指定两个主成分的数据结构和载荷。在双标图中, 点之间的距离, 反映它们对应的样本之间的差异大小。两点相距较远, 对应样本差异大; 两点相距较近, 对应样本差异小, 存在相似性。利用样本点的相似性, 我们可以直观的将样本分类。

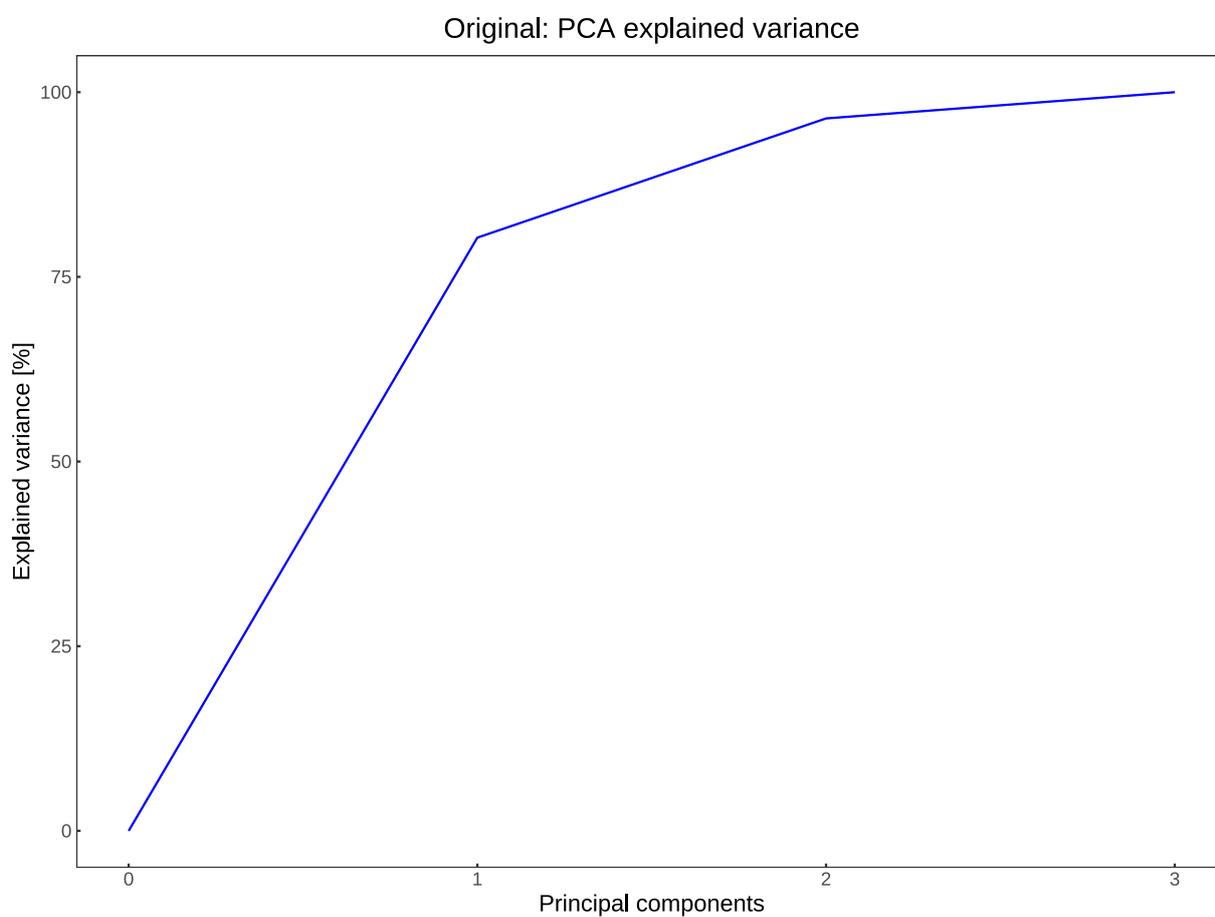
第 22 章 主成分解释方差 (PCA Explained Variance)

22.1 主成分解释方差 (PCA_exp_var)

22.1.1 函数介绍

该函数只是为了将各个主成分的解释方差，即方差贡献率以累计方差贡献率的形式进行可视化，见如下示例：

```
1 # 运行函数源文件
2 source("../Consensus-code/PCA_exp_var.R", encoding="utf-8")
3 # 调用该函数
4 PCA_exp_var(data=data, type="original", based_on="cov",
5             color_line="blue", lang="en")
```



参数介绍

data: 原始数据

type: 设置对原始数据进行处理的方法, 默认值为"original":

—————"original": 对原始数据进行简单平均构造共性矩阵

—————"standardized": 对原始数据进行标准化再平均, 构造共性矩阵

—————"STATIS": 利用 STATIS 方法对原始数据处理, 构造共性矩阵

based_on: 当且仅当 *type*="STATIS" 时, 参数有效, 默认值为"cov":

—————"cov": 指定进行基于协方差 (Covariance) 的 STATIS 分析

—————"cor": 指定进行基于相关性 (Correlation) 的 STATIS 分析

color_line: 实线的颜色, 默认为蓝色

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

22.1.2 可视化分析

该图也叫碎石图 (Scree Plot), 可以用于让研究人员判断选取主成分个数。

图表分析

该图可以直观的让研究人员理解哪些主成分的方差贡献率较大, 并且方差贡献率较大的主成分所在的直线一般上升陡峭。除此之外, 最重要的是, 在应用主成分分析时, 研究人员更想知道应该取前几个主成分以达到降维的目的, 常用的标准有: 一个按累计贡献率达到 70% 或 80% 来确定主成分个数, 另一个是按大于特征值均值的特征值所代表的主成分的个数。其实还可以根据观看碎石图的拐点在哪, 拐点处即可当作选取的分界点, 即选取拐点之前的主成分。

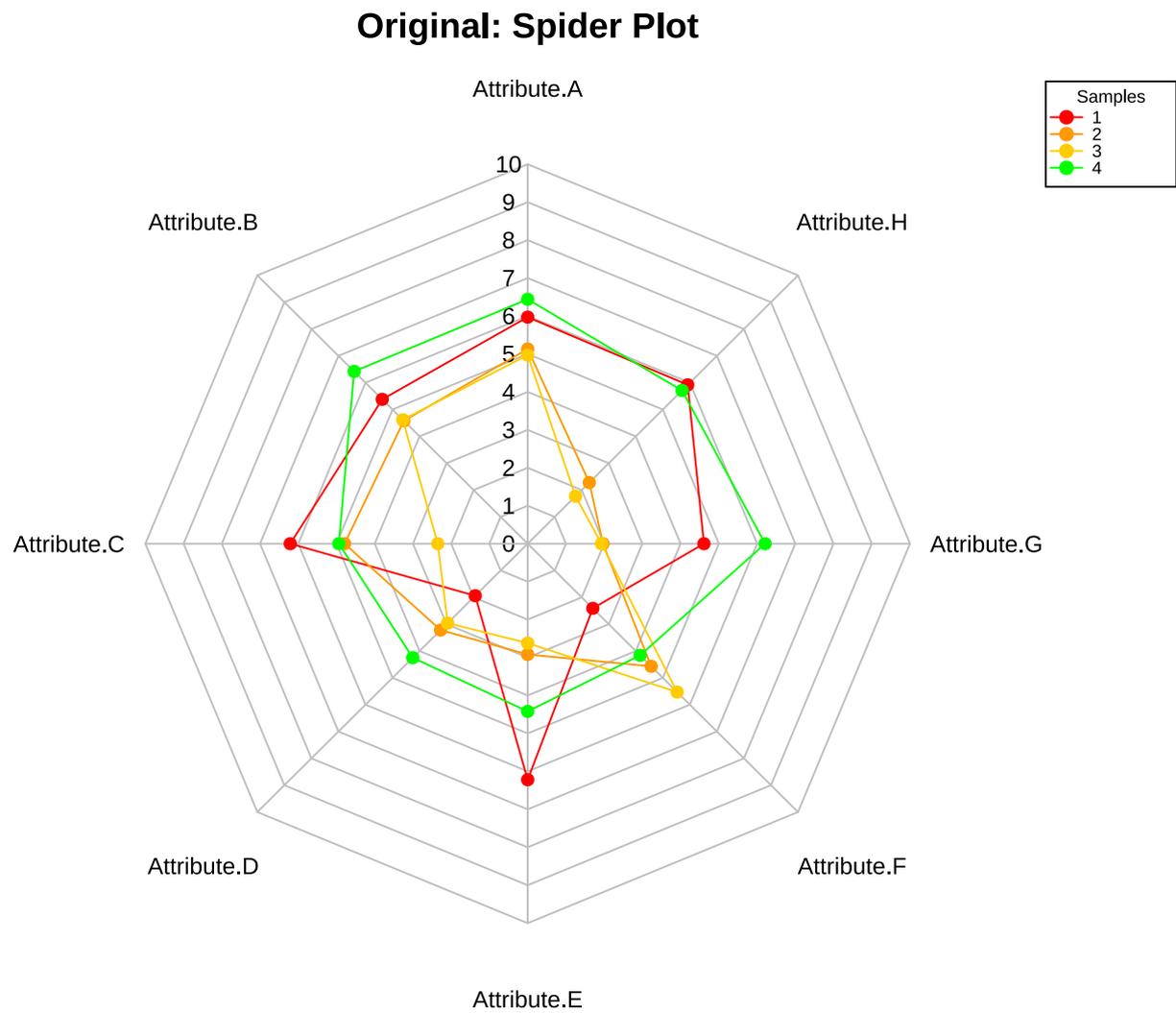
第 23 章 蜘蛛网图 (Spiderweb Plot)

23.1 蜘蛛网图 (Spiderweb_plot)

23.1.1 函数介绍

蜘蛛网图也叫雷达图, 用于展示多维有序数据。见如下示例:

```
1 # 运行函数源文件
2 source("../Consensus-code/Spiderweb_plot.R",encoding="utf-8")
3 # 调用该函数
4 Spiderweb_plot(data=data,score=10,type="original",based_on="cov",lang="en")
```



参数介绍

data: 原始数据

score: 原始数据评分尺度的最大值, 比如按 1~10 评分尺度评分, 则该参数值为 10, 默认值为 10

type: 设置对原始数据进行处理的方法, 默认值为"original":

———"original": 对原始数据进行简单平均构造共性矩阵

———"standardized": 对原始数据进行标准化再平均, 构造共性矩阵

———"STATIS": 利用 STATIS 方法对原始数据处理, 构造共性矩阵

based_on: 当且仅当 *type*="STATIS" 时, 参数有效, 默认值为"cov":

———"cov": 指定进行基于协方差 (Covariance) 的 STATIS 分析

———"cor": 指定进行基于相关性 (Correlation) 的 STATIS 分析

lang: 指定中英文模板, 默认为"en":

———"cn": 数据为中文模板

———"en": 数据为英文模板

23.1.2 可视化分析

如本节示例所示, 蜘蛛网图 (也叫雷达图) 展示的是一致性数据 (*Consensus data*), 即 *Consen_data* 函数所得出的数据。

图表分析

雷达图将多个维度的数据量映射到坐标轴上, 这些坐标轴起始于同一个圆心点, 通常结束于圆周边缘, 将同一组的点使用线连接起来就称为雷达图。它可以将多维数据进行展示, 但是点的相对位置和坐标轴之间的夹角是没有任何信息量的。在坐标轴设置恰当的情况下雷达图所围面积能表现出一些信息量。每一个维度的数据都分别对应一个坐标轴, 这些坐标轴具有相同的圆心, 以相同的间距沿着径向排列, 并且各个坐标轴的刻度相同。连接各个坐标轴的网格线通常只作为辅助元素。将各个坐标轴上的数据点用线连接起来就形成了一个多边形。坐标轴、点、线、多边形共同组成了雷达图。雷达图还可以展示出数据集中各个变量的权重高低情况, 非常适用于展示性能数据。

雷达图也有一些主要缺点:

- 1、如果雷达图上多边形过多会使可读性下降, 使整体图形过于混乱。特别是有颜色填充的多边形的情况, 上层会遮挡覆盖下层多边形;
- 2、如果变量过多, 也会造成可读性下降, 因为一个变量对应一个坐标轴, 这样会使坐标轴过于密集, 使图表给人感觉很复杂。所以最佳实践就是尽可能控制变量的数量使雷达图保持简单清晰。

第五部分 全局分析 (Overall)

在全局分析 (Overall)部分中, 与 PanelCheck 一致, 笔者将依次介绍:

- 只评估一次的双因素方差分析 (2-way ANOVA(1rep));
- 一般双因素方差分析 (2-way ANOVA);
- 三因素方差分析 (3-way ANOVA)。

第 24 章 双因素方差分析-只评估 1 次 (2-way ANOVA(1rep))

该方法仅仅适用于每个评估人员评估每个感官属性只评估一次的情形。由于前面所使用的原始数据的重复评估两次, 所以并不能使用本方法。所以笔者为了能够进行该方法的示例, 将原始数据 (sample.csv) 删去第 2 次重复评估的评分, 重新导入该数据:

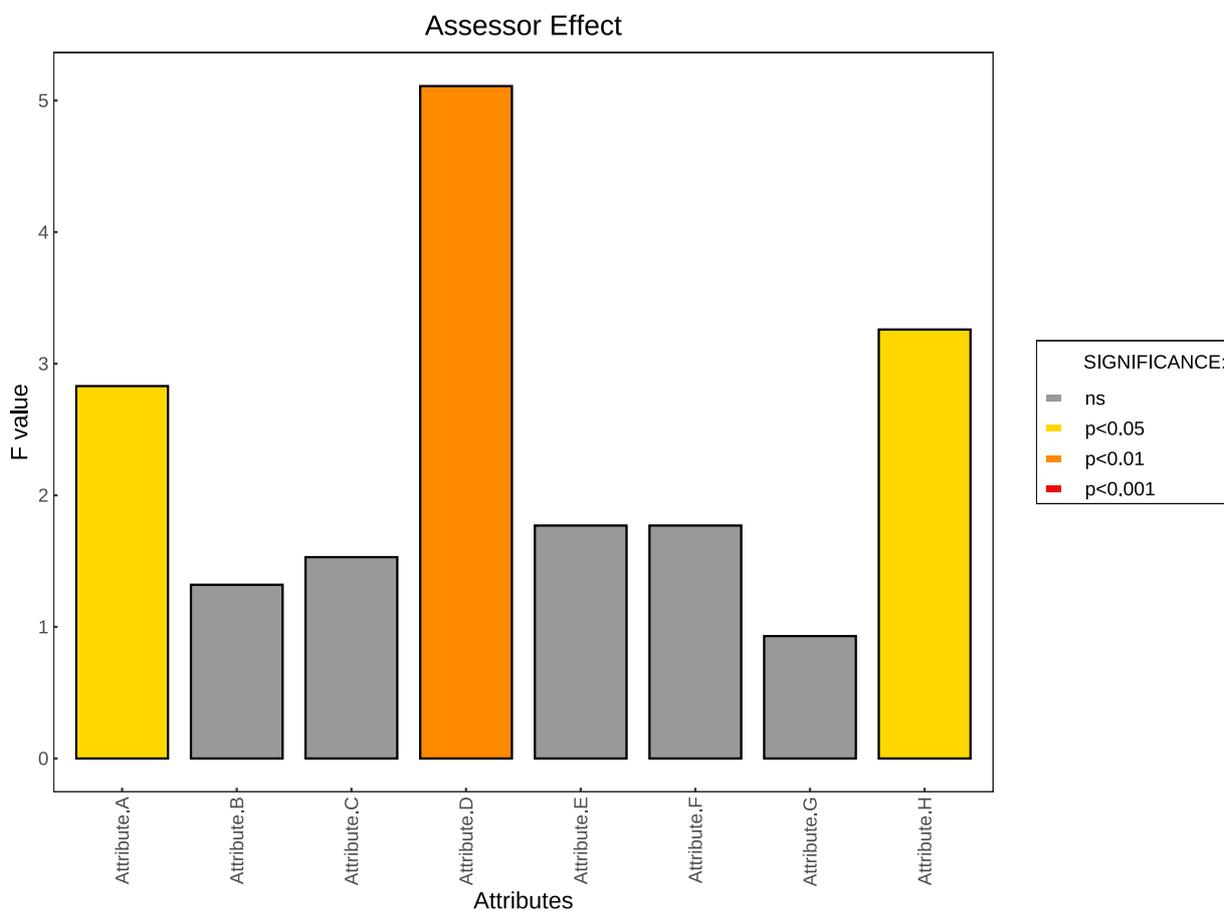
```
1 # 导入修改后数据
2 data<-read.csv("sample-1rep.csv")
```

24.1 评估者主效应 (ANOVA_2way_1rep_ass_effect)

24.1.1 函数介绍

将各感官属性得分作为因变量, 评估者与产品 (即样品) 作为自变量进行无交互作用的双因素方差分析。该函数对各属性分析结果中的评估者变量的 F 值作为其评估者效应进行可视化, 见如下示例:

```
1 # 运行函数源文件
2 source("../Overall-code/2_way_ANOVA(1_rep)/ANOVA_2way_1rep_ass_effect.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA_2way_1rep_ass_effect(data=data,angle=90,lang="en")
```

**参数介绍**

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

24.1.2 可视化分析

与前面的 F 值图 (11.1) 比较类似, 都是以 F 值为纵坐标。从上面示例可以看出, 每个柱子高度都代表该感官属性所代表的双因素方差分析的评估者 F 检验的 F 值大小; 柱子的填充颜色代表各感官属性评估者变量的显著性程度, 参考 7.2.2。

图表分析

参考 7.2.2, 对于评估者效应显著的感官属性, 说明不同的评估人员对该感官属性有显著的区别能力, 反之, 则区分能力不明显。

24.2 产品主效应 (ANOVA_2way_1rep_pro_effect)

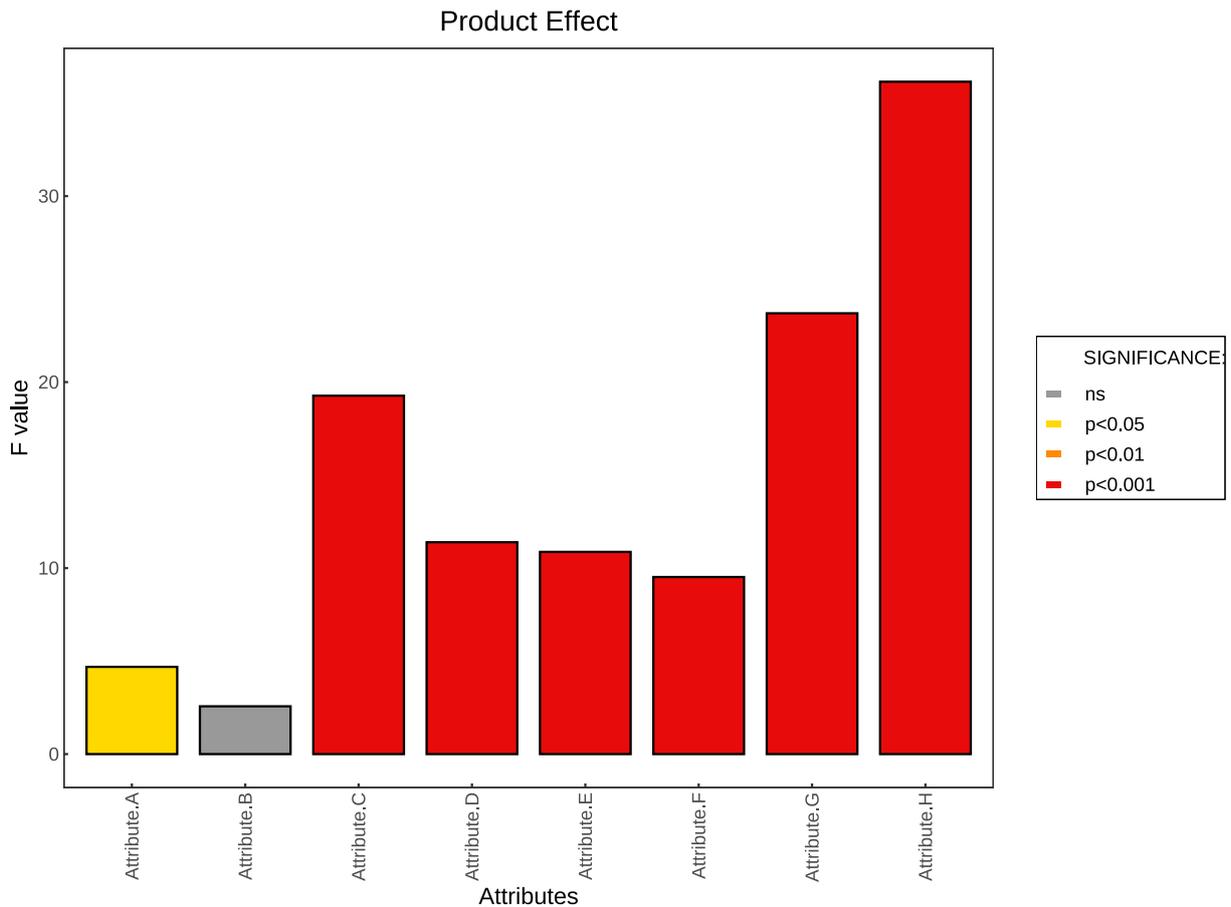
24.2.1 函数介绍

与评估者效应类似，仅仅是换了一个产品即样品角度，见如下示例：

```

1 # 运行函数源文件
2 source("../Overall-code/2_way_ANOVA(1_rep)/ANOVA_2way_1rep_pro_effect.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA_2way_1rep_pro_effect(data=data,angle=90,lang="en")

```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度，默认为 90 度

lang: 指定中英文模板，默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

24.2.2 可视化分析

与评估者效应类似，从上面示例可以看出，每个柱子高度都代表该感官属性所代表的双因素方差分析产品 (Sample)F 检验的 F 值大小；柱子的填充颜色代表各感官属性产品变量 (Sample) 的显著性程度。

图表分析

对于产品效应显著的感官属性，说明不同的产品，其该感官属性评分有显著差异，反之，则差异不明显。

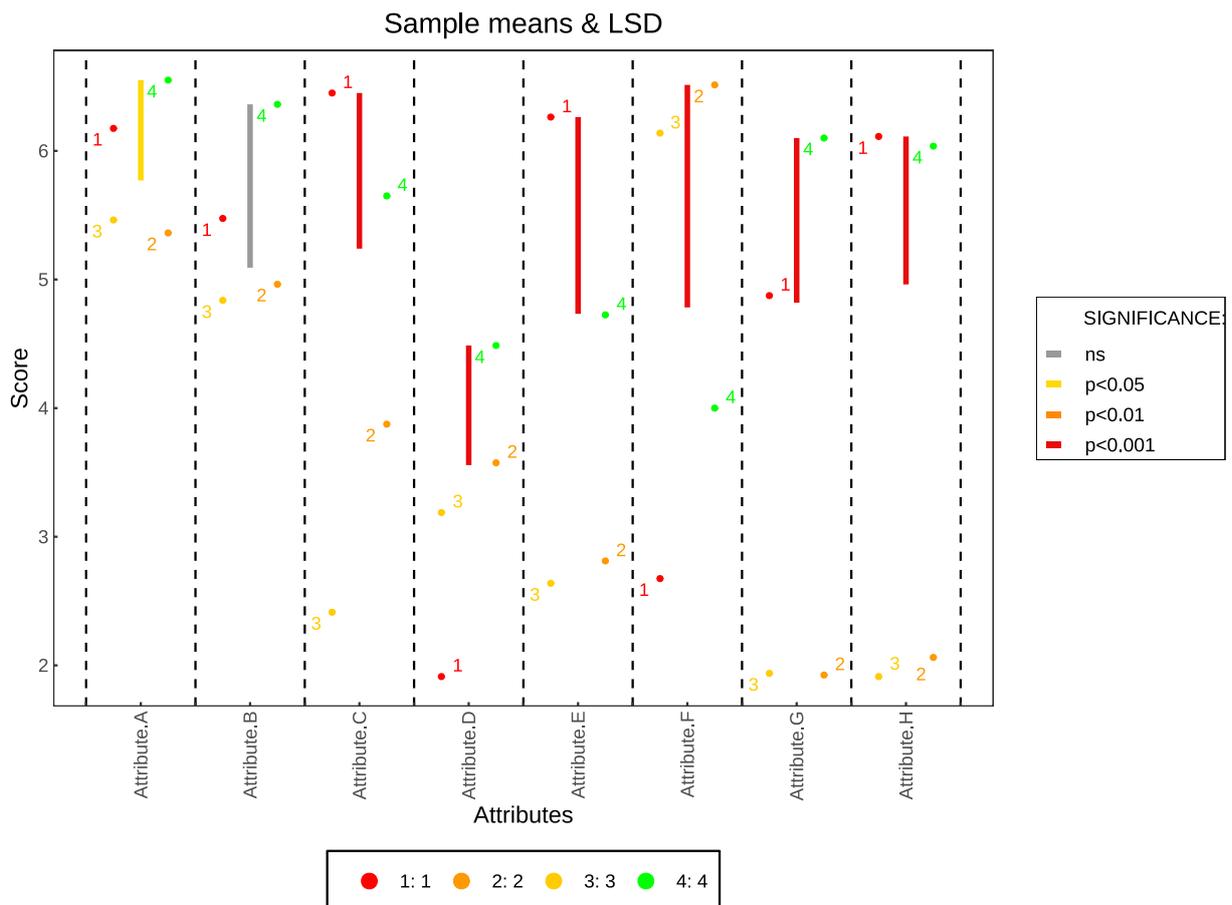
24.3 样本均值检验 (ANOVA2_1rep_Sample_mean_LSD)

24.3.1 函数介绍

该函数绘制每个样本的各个感官属性平均得分以及其 LSD 成对检验区间。需要注意的是，参数 (p.adj) 可以指定对 LSD 检验的 P 值修正方法。请见如下示例：

```

1 # 运行函数源文件
2 source("../Overall-code/2_way_ANOVA(1_rep)/ANOVA2_1rep_Sample_mean_LSD.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA2_1rep_Sample_mean_LSD(data=data,angle=90,p.adj="none",lang="en")
    
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

p.adj: 指定 LSD 多重比较的 p 值修正, 默认为"none":

—————"none": 即不进行 P 值修正

—————"bonferroni": 即进行 bonferroni 修正

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

24.3.2 可视化分析

如上示例, 图形元素过多, 包括:

- **颜色标记散点**: 散点表示对应感官属性下各个样本的平均评估得分, 颜色填充用以区分不同的样品;
- **LSD 线条**: 该线条长度表示 LSD 两两成对检验的均值区间, 颜色填充表示对应属性的样本显著性, 参考 7.2.2。
- **图例**: 绘图面板下方的图例中, 冒号 (:) 前面的数字代表第几个样本, 冒号后面的代表样本名称。由于原始数据中, 样本名以数字代替, 所以才会显示上述的图例情形。而图形中散点的 **数字标签** 代表的是第几个样本, 并不是样本名称。

图表分析

样本均值-LSD 图简单地描绘了样本平均值, 包括每个属性的 LSD 条。LSD 条应该用作每个属性的动态测量棒, 即两个样本的对应感官属性的平均得分差值超过该 LSD 条长度, 则两样本在该属性下具有显著差异。**p.adj** 参数可以指定对 LSD 检验进行 bonferroni P 值修正, 一般是更有说服力的。在每个属性中将样本放置在栏的每一侧只是为了增强可读性。

第 25 章 双因素方差分析 (2-way ANOVA)

由于下面的双因素方差分析与三因素方差分析适用于多重复评估的实验数据, 所以我们重新导入原始数据 **sample.csv**:

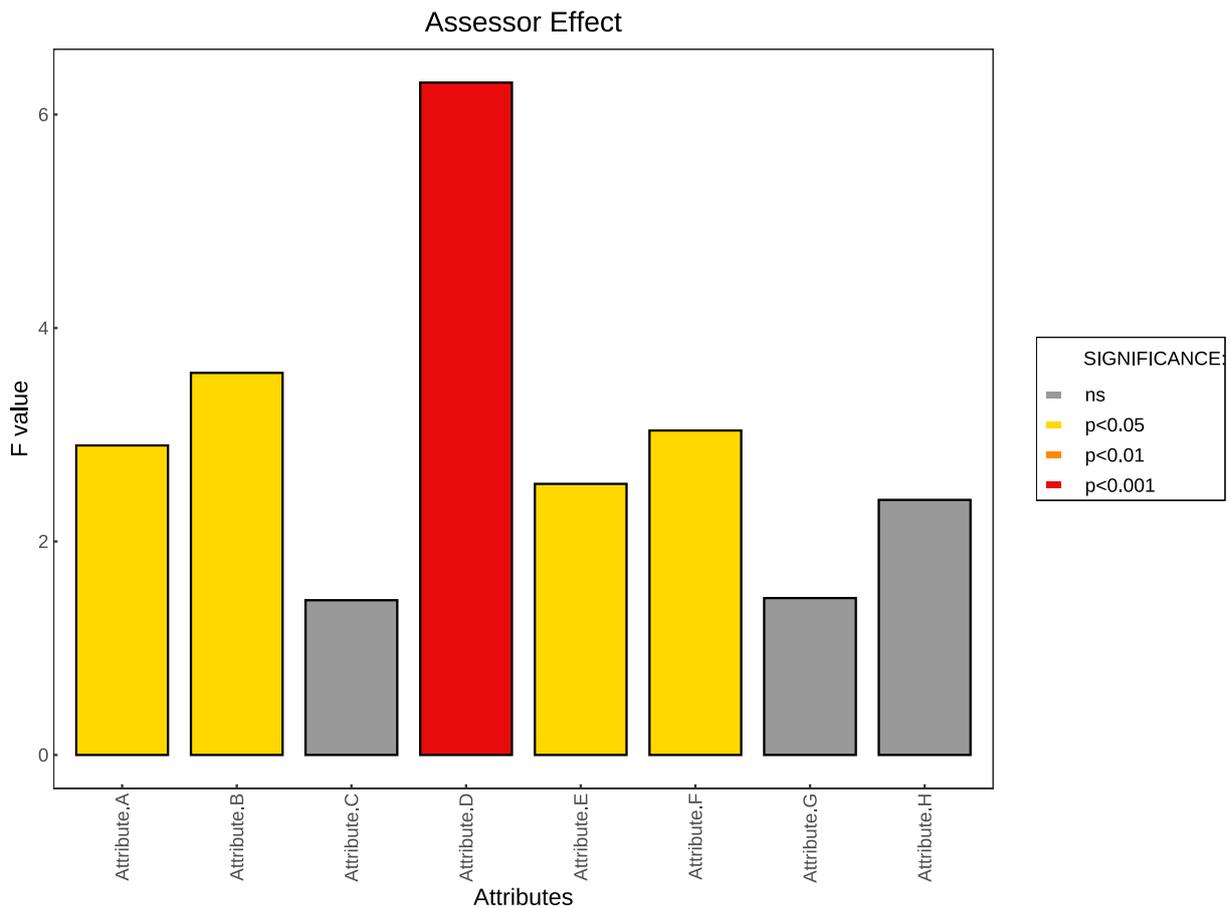
```
1 # 导入原始的多重复评估数据
2 data<-read.csv("sample.csv")
```

25.1 评估者主效应 (*ANOVA_2way_ass_effect*)

25.1.1 函数介绍

该函数绘图与只评估一次的双因素方差分析的评估者效应基本一致，见如下示例：

```
1 # 运行函数源文件
2 source("../Overall-code/2_way_ANOVA/ANOVA_2way_ass_effect.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA_2way_ass_effect(data=data,angle=90,lang="en")
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度，默认为 90 度

lang: 指定中英文模板，默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

25.1.2 可视化分析

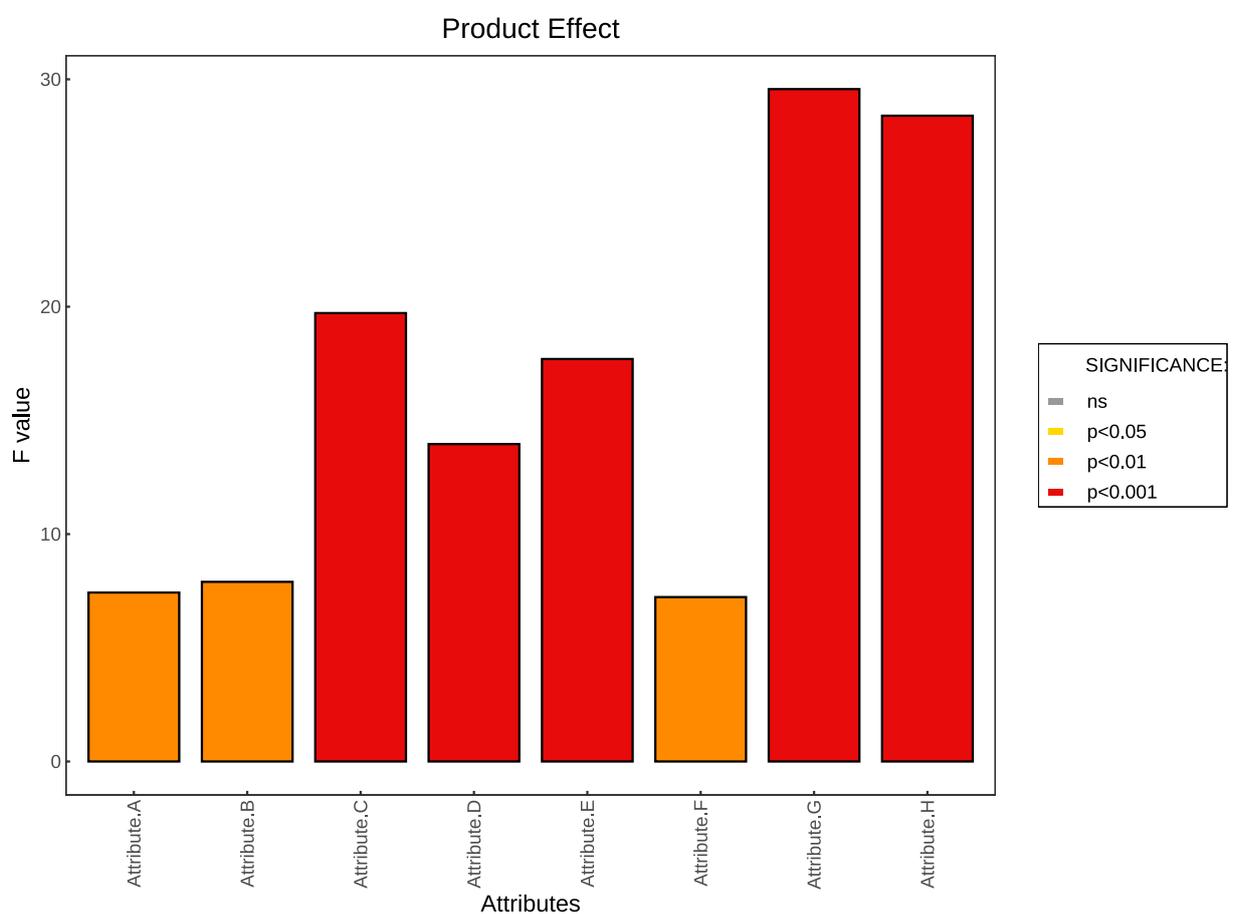
对于评估者效应分析参考24.1.2。

25.2 产品主效应 (ANOVA_2way_pro_effect)

25.2.1 函数介绍

与上一节的产品效应基本一致，见如下示例：

```
1 # 运行函数源文件
2 source("../Overall-code/2_way_ANOVA/ANOVA_2way_pro_effect.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA_2way_pro_effect(data=data,angle=90,lang="en")
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

25.2.2 可视化分析

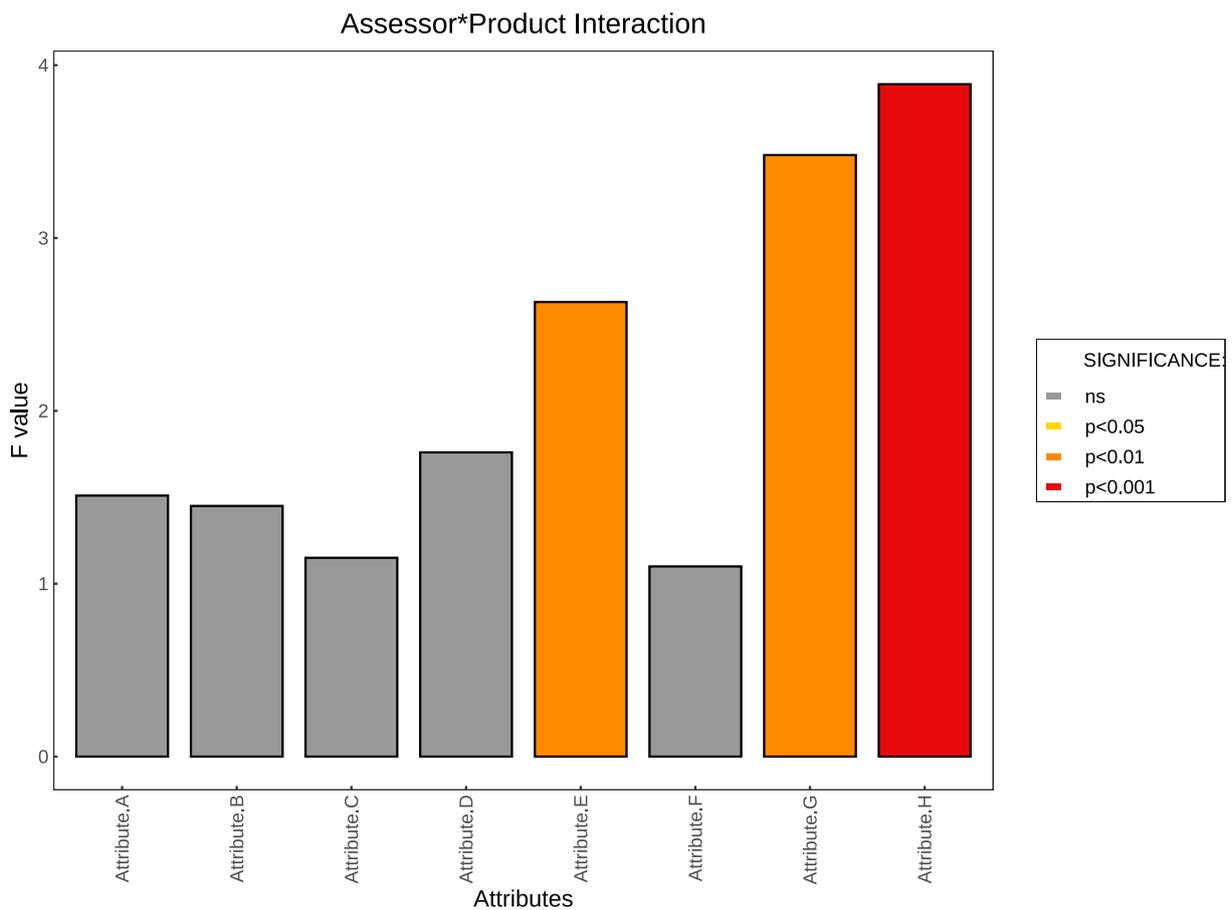
分析请参考[24.2.2](#)。

25.3 评估者 * 产品交互效应 (`ANOVA_2way_asspro_inter`)

25.3.1 函数介绍

交互效应的绘图与单变量主效应完全类似, 不同在于分析时, 评估者和产品的联合效应是否会显著影响感官属性评估得分。请见如下示例:

```
1 # 运行函数源文件
2 source("../Overall-code/2_way_ANOVA/ANOVA_2way_asspro_inter.R", encoding="utf-8")
3 # 调用该函数
4 ANOVA_2way_asspro_inter(data=data, angle=90, lang="en")
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

25.3.2 可视化分析

交互效应与单变量效应分析是基本一致的, 只不过交互效应是考虑了两个变量的联合影响。

图表分析

交互效应是反映两个或两个以上自变量相互依赖、相互制约, 共同对因变量的变化发生影响。换句话说, 如果一个自变量对因变量的影响效应会因另一个自变量的水平不同而有所不同, 则我们说这两个变量之间具有交互效应。交互效应显著, 则说明评估者对感官属性评分的影响程度, 会因产品的不同而显著不同。

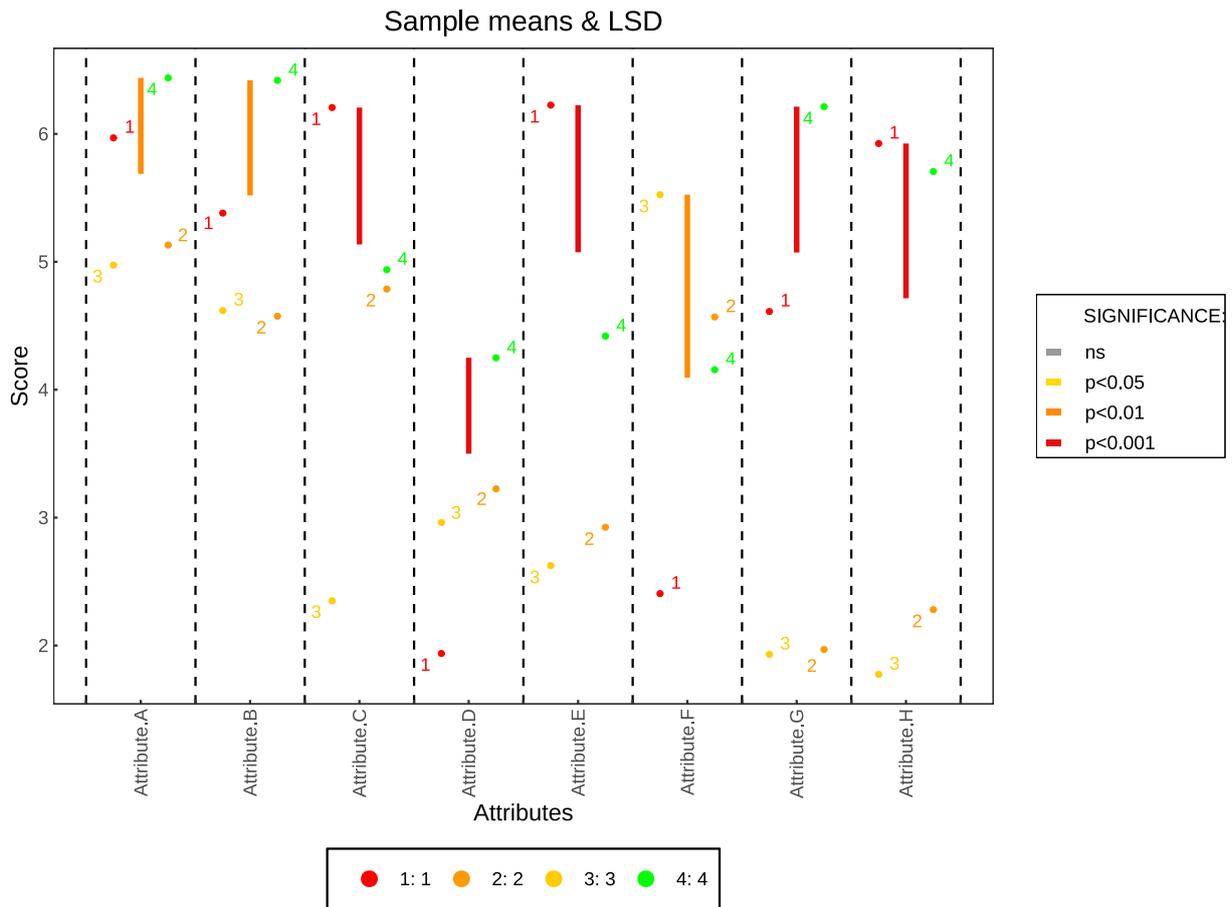
25.4 样本均值检验 (ANOVA2_Sample_mean_LSD)

25.4.1 函数介绍

均值检验绘图与上述完全一致，见如下示例：

```

1 # 运行函数源文件
2 source("../Overall-code/2_way_ANOVA/ANOVA2_Sample_mean_LSD.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA2_Sample_mean_LSD(data=data,angle=90,p.adj="none",lang="en")
    
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

p.adj: 指定 LSD 多重比较的 p 值修正, 默认为"none":

—————"none": 即不进行 P 值修正

—————"bonferroni": 即进行 bonferroni 修正

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

25.4.2 可视化分析

可视化分析与24.3.2完全一致。

第 26 章 三因素方差分析 (3-way ANOVA)

三因素方差分析与双因素方差分析的区别在于考虑了重复评估 (**Replicate**) 对因变量感官属性评分的影响。

26.1 评估者主效应 (**ANOVA_3way_ass_effect**)

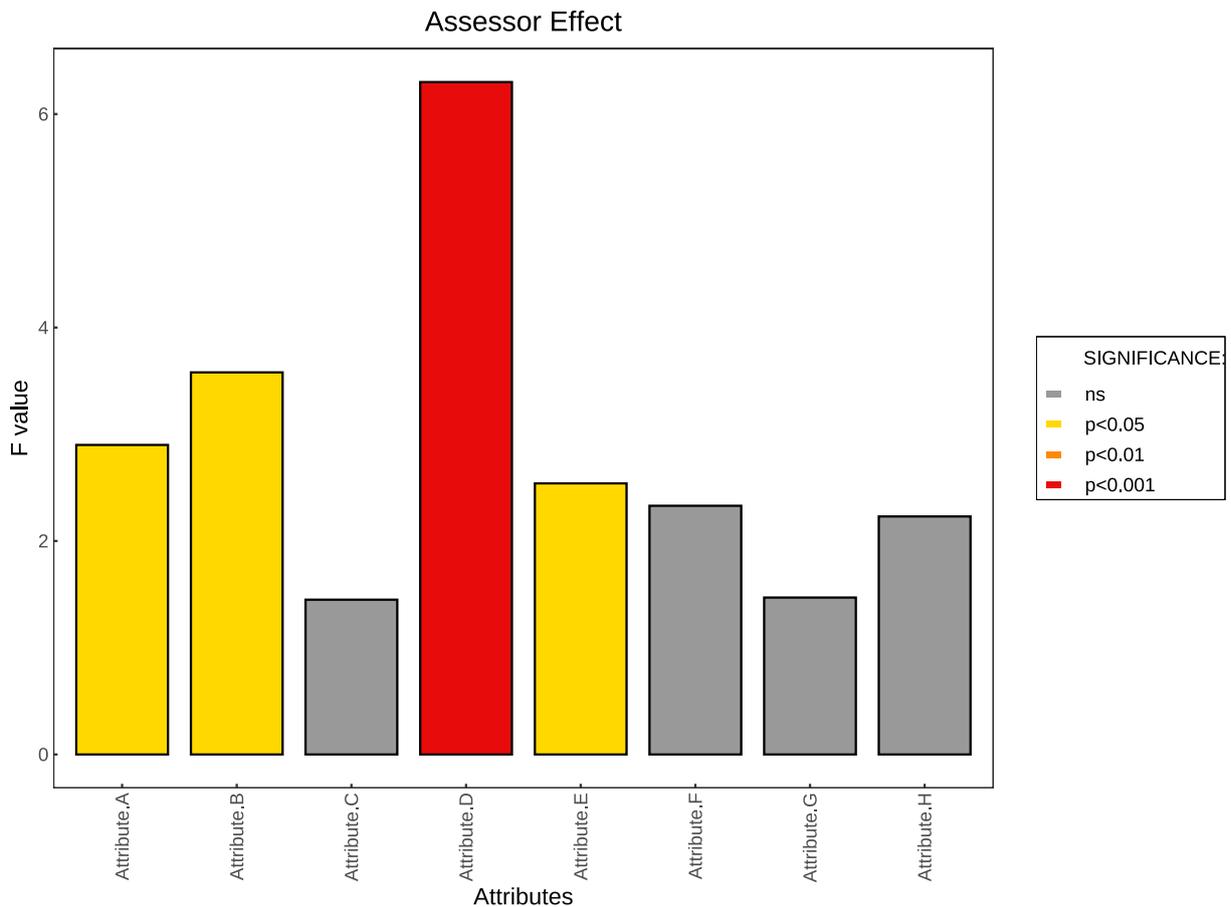
26.1.1 函数介绍

该效应与双因素方差分析评估者主效应类似, 请见如下示例:

```

1 # 运行函数源文件
2 source("../Overall-code/3_way_ANOVA/ANOVA_3way_ass_effect.R", encoding="utf-8")
3 # 调用该函数
4 ANOVA_3way_ass_effect(data=data, angle=90, lang="en")

```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

26.1.2 可视化分析

图表分析方式与上述评估者效应一致, 只是由于考虑了更多自变量而导致数值不同而已。请参考[24.1.2](#)。

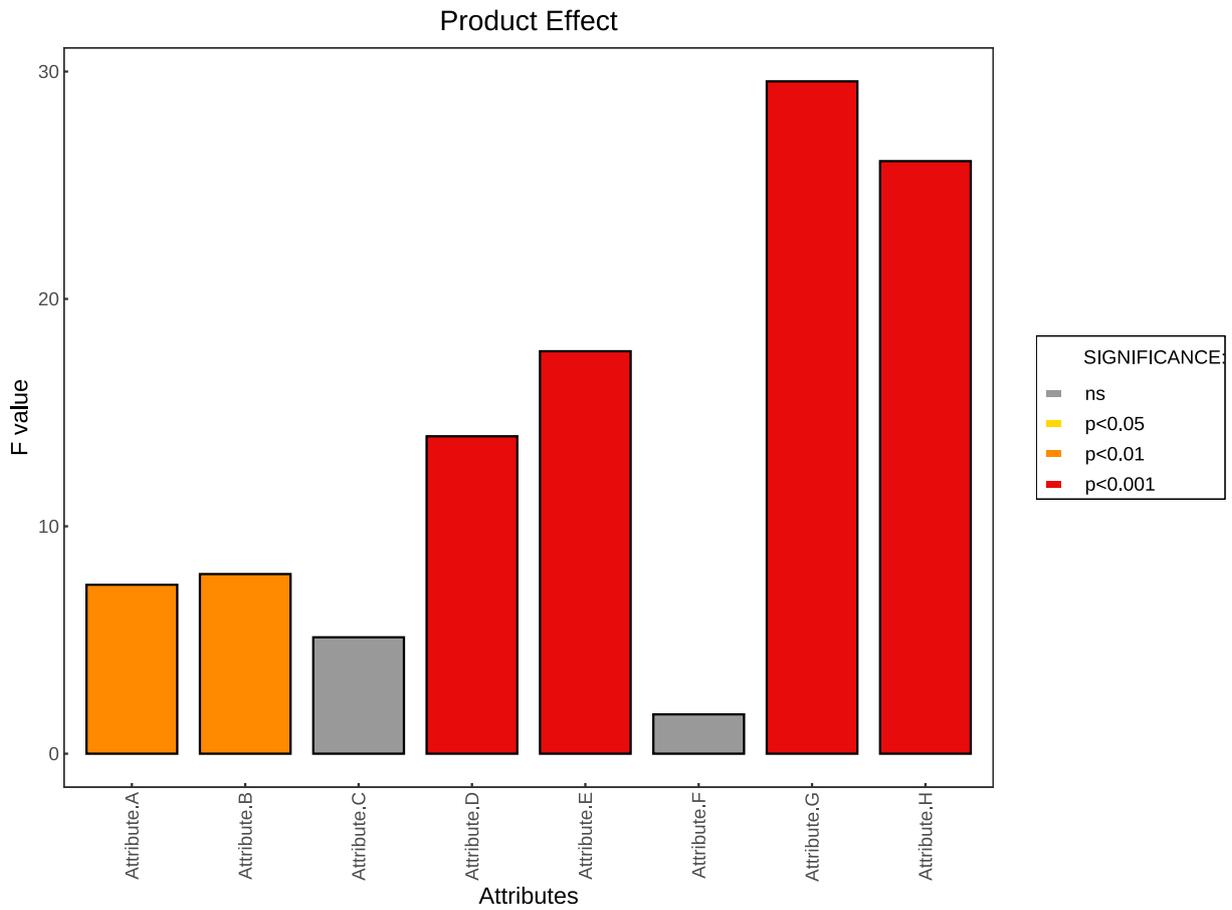
26.2 产品主效应 (ANOVA_3way_pro_effect)

26.2.1 函数介绍

请见如下示例:

```
1 # 运行函数源文件
2 source("../Overall-code/3_way_ANOVA/ANOVA_3way_pro_effect.R", encoding="utf-8")
```

```
3 # 调用该函数  
4 ANOVA_3way_pro_effect(data=data,angle=90,lang="en")
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

26.2.2 可视化分析

分析请参考[24.2.2](#)。

26.3 重复评估主效应 (ANOVA_3way_rep_effect)

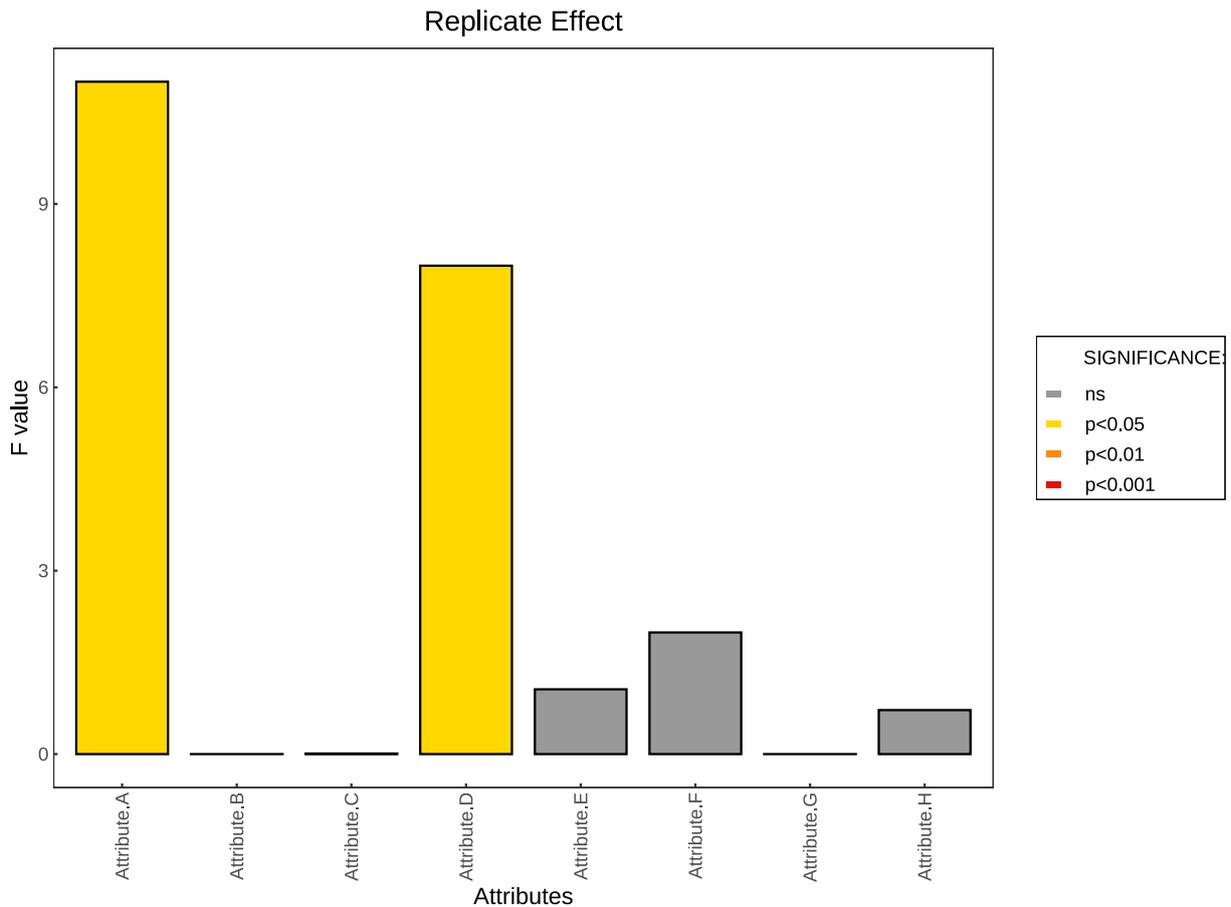
26.3.1 函数介绍

重复评估主效应就是考虑该因素对感官属性评分是否有显著影响，请见如下示例：

```

1 # 运行函数源文件
2 source("../Overall-code/3_way_ANOVA/ANOVA_3way_rep_effect.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA_3way_rep_effect(data=data,angle=90,lang="en")

```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度，默认为 90 度

lang: 指定中英文模板，默认为"en":

—————"cn": 数据为中文模板

—————"en": 数据为英文模板

26.3.2 可视化分析

图表分析

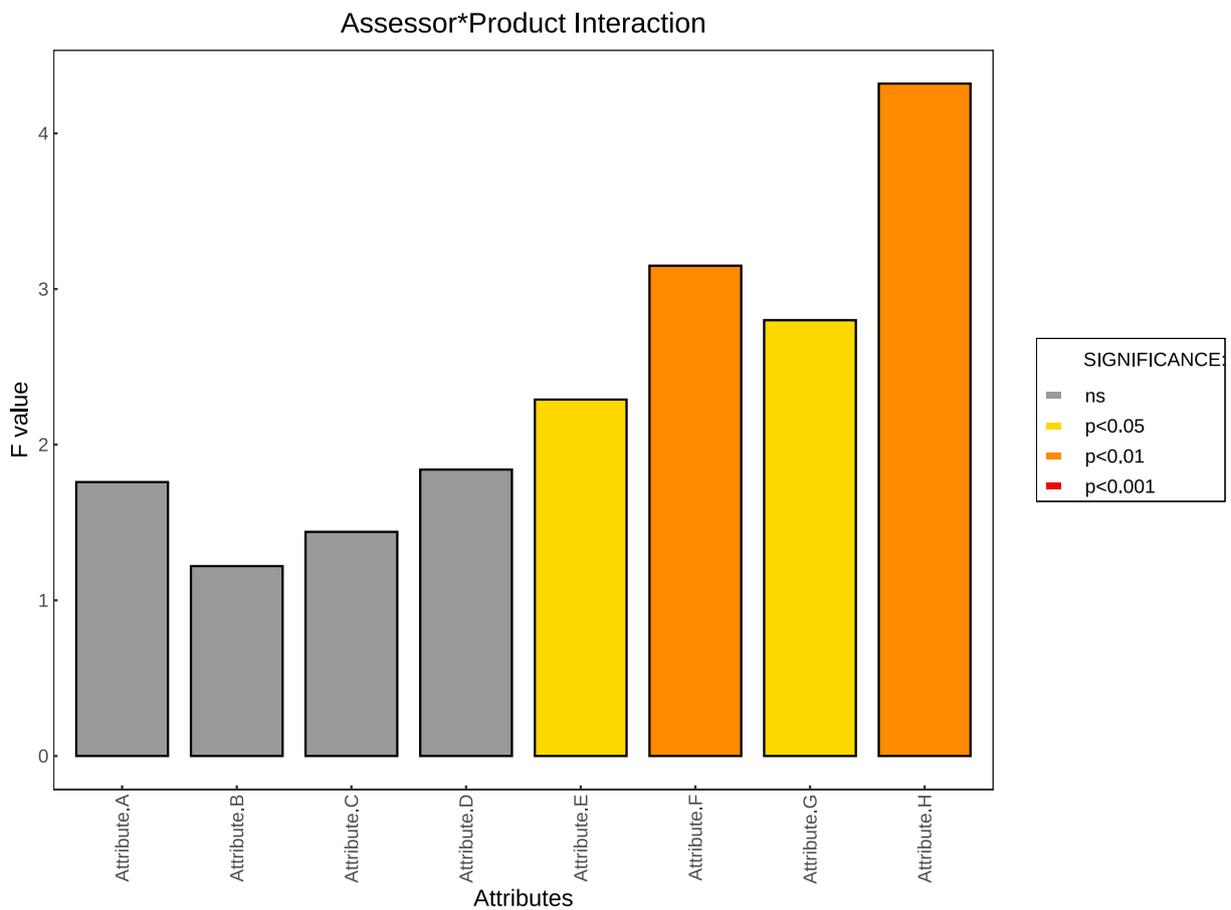
参考7.2.2, 对于重复评估主效应显著的感官属性, 说明重复评估会使得该感官属性的评分有显著差别, 反之, 则差别不明显。

26.4 评估者 * 产品交互效应 (ANOVA_3way_asspro_inter)

26.4.1 函数介绍

请看如下示例:

```
1 # 运行函数源文件
2 source("../Overall-code/3_way_ANOVA/ANOVA_3way_asspro_inter.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA_3way_asspro_inter(data=data,angle=90,lang="en")
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

26.4.2 可视化分析

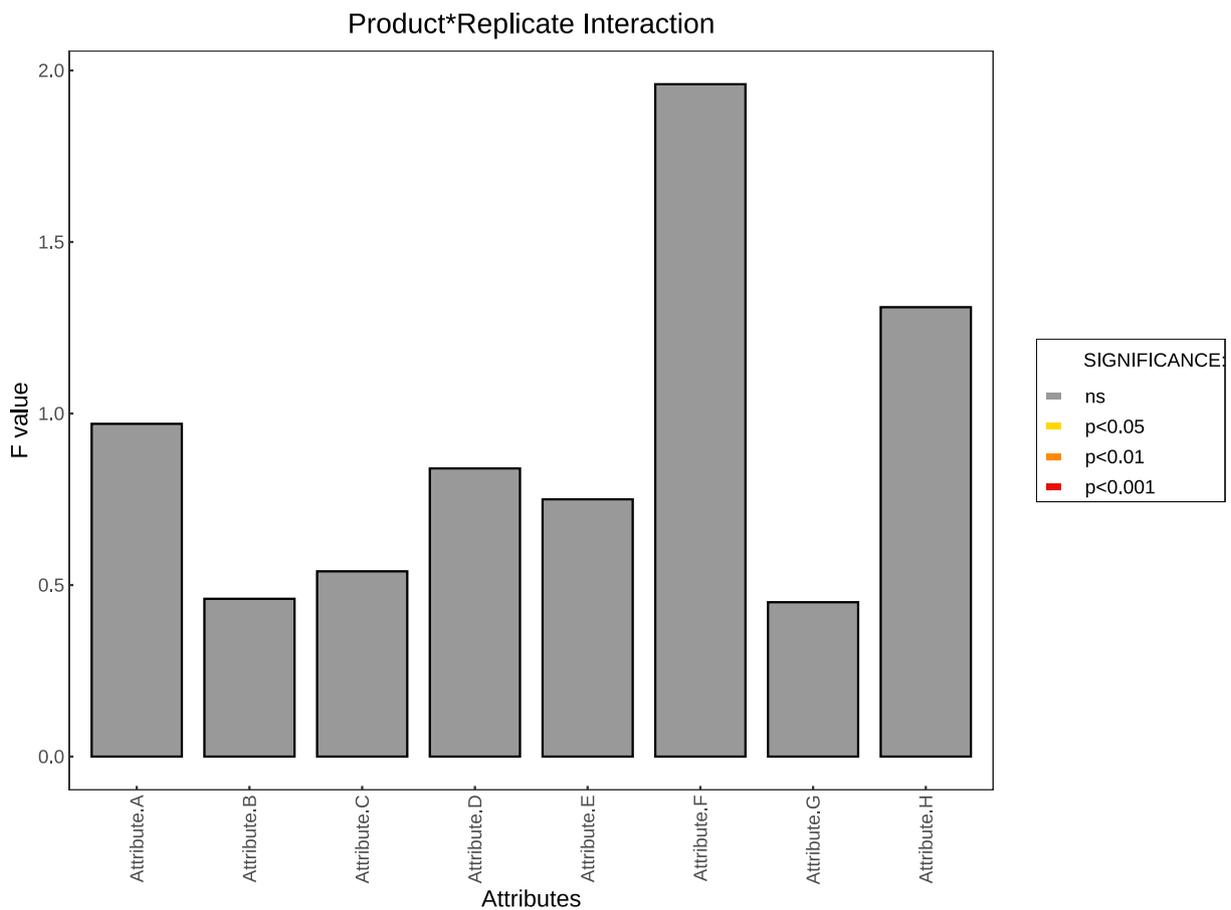
与主效应类似, 交互效应考虑了两个变量的联合影响, 可以参考[25.3.2](#)。

26.5 评估者 * 重复评估交互效应 (`ANOVA_3way_assrep_inter`)

26.5.1 函数介绍

请见如下示例:

```
1 # 运行函数源文件
2 source("../Overall-code/3_way_ANOVA/ANOVA_3way_assrep_inter.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA_3way_assrep_inter(data=data,angle=90,lang="en")
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

26.5.2 可视化分析

与主效应类似, 交互效应考虑了两个变量的联合影响, 可以参考[25.3.2](#)。

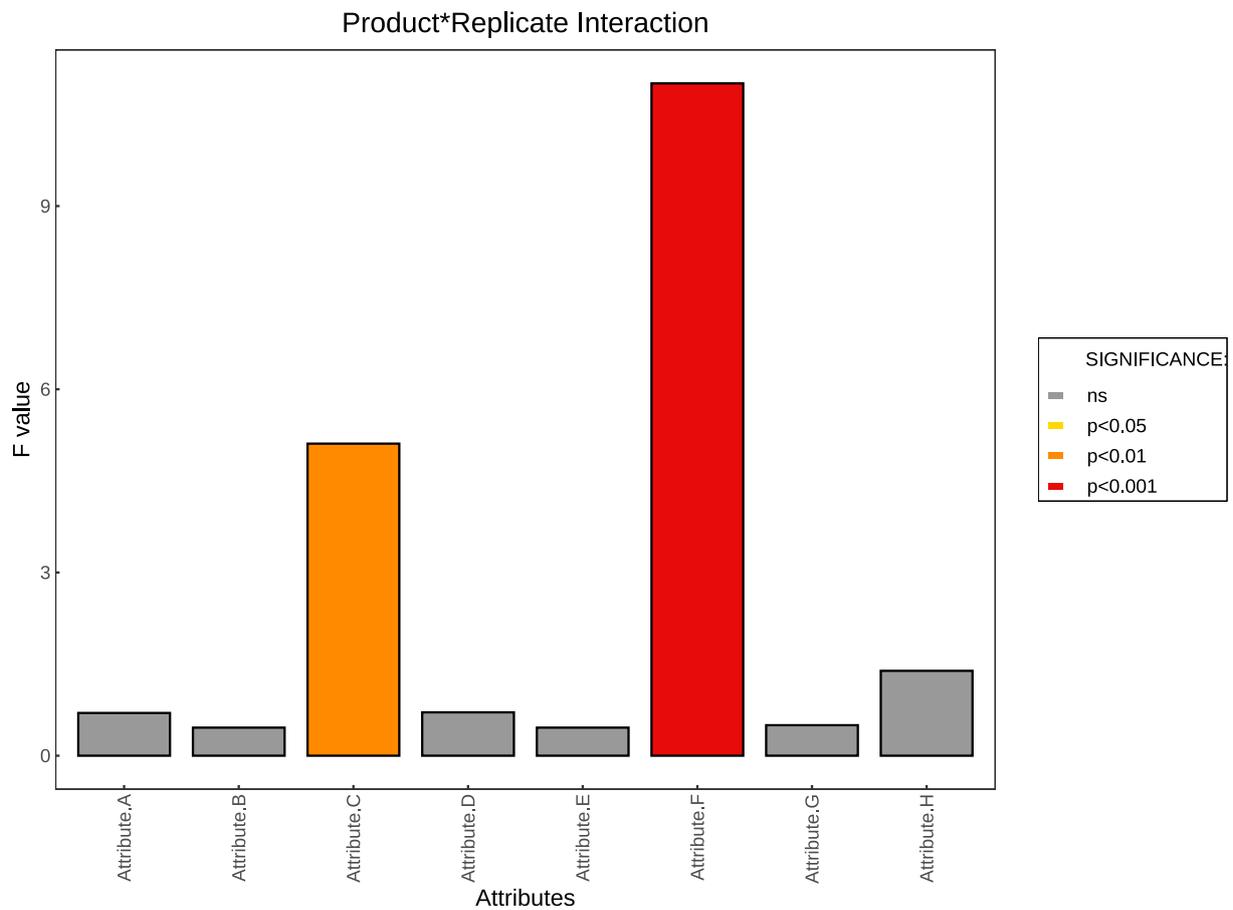
26.6 产品 * 重复评估交互效应 (*ANOVA_3way_prorep_inter*)

26.6.1 函数介绍

请见如下示例:

```
1 # 运行函数源文件
2 source("../Overall-code/3_way_ANOVA/ANOVA_3way_prorep_inter.R", encoding="utf-8")
```

```
3 # 调用该函数  
4 ANOVA_3way_prorep_inter(data=data, angle=90, lang="en")
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

lang: 指定中英文模板, 默认为"en":

_____ "cn": 数据为中文模板

_____ "en": 数据为英文模板

26.6.2 可视化分析

与主效应类似, 交互效应考虑了两个变量的联合影响, 可以参考[25.3.2](#)。

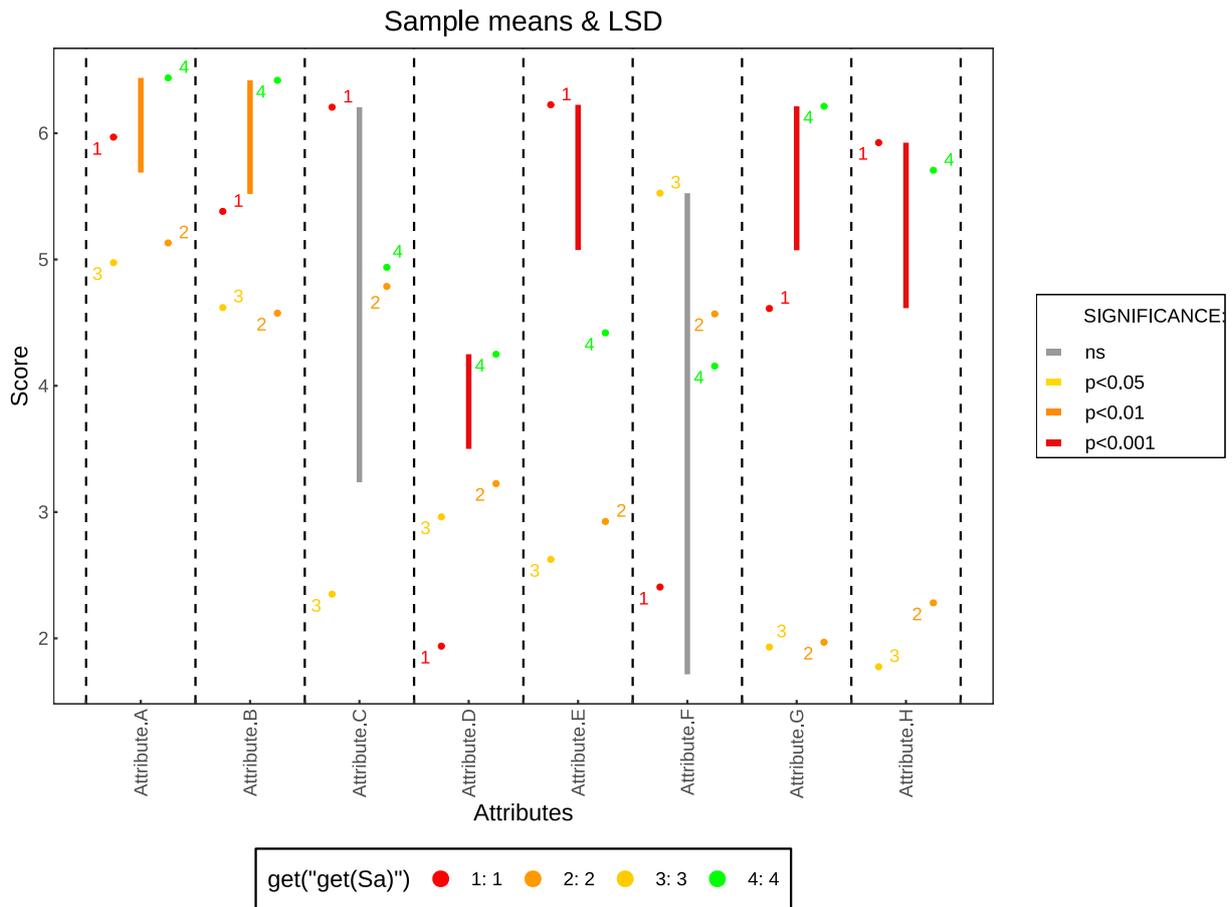
26.7 样本均值检验 (ANOVA3_Sample_mean_LSD)

26.7.1 函数介绍

请见如下示例:

```

1 # 运行函数源文件
2 source("../Overall-code/3_way_ANOVA/ANOVA3_Sample_mean_LSD.R",encoding="utf-8")
3 # 调用该函数
4 ANOVA3_Sample_mean_LSD(data=data,angle=90,p.adj="none",lang="en")
    
```



参数介绍

data: 原始数据

angle: x 轴文本标签显示角度, 默认为 90 度

p.adj: 指定 LSD 多重比较的 p 值修正, 默认为"none":

———"none": 即不进行 P 值修正

———"bonferroni": 即进行 bonferroni 修正

lang: 指定中英文模板, 默认为"en":

———"cn": 数据为中文模板

———"en": 数据为英文模板

26.7.2 可视化分析

可视化分析与24.3.2完全一致。

参考文献

- [1] Oliver Tomic, Asgeir Nilsen, Magni Martens, Tormod Næs, Visualization of sensory profiling data for performance monitoring, LWT 40 (2007) 262–269
- [2] Oliver Tomic, Giorgio Luciano, Asgeir Nilsen, Grethe Hyldig, Kirsten Lorensen, Tormod Næs, Analysing sensory panel performance in a proficiency test using the PanelCheck software, Eur Food Res Technol (2010) 230:497-511
- [3] Tucker LR (1964) The extension of factor analysis to three-dimensional matrices. In: Frederiksen N, Gulliksen H (eds) Contributions to mathematical psychology. Holt, Rinehart & Winston, New York
- [4] Tucker LR (1966) Some mathematical notes on three-mode factor analysis. Psychometrika 31:279-311
- [5] Dahl T, Tomic O, Wold JP, Næs T (2008) Some new tools for visualizing multi-way sensory data. Food Qual Prefer 19:103-113
- [6] P.B. Brockhoff, Statistical testing of individual differences in sensory profiling, Food Quality and Preference 14 (2003) 425-43.