

# IBM SPSS Conjoint 20



注意：使用本信息及其支持的产品之前，请阅读注意事项第 32 页码下的一般信息。

此版本适用于 IBM® SPSS® Statistics 20 及所有后续发布和修订，除非在新版本中另有说明。

Adobe 产品屏幕截图重印已获得 Adobe Systems Incorporated 的许可。

Microsoft 产品屏幕截图重印已获得 Microsoft Corporation 的许可。

受许可保护材料 – IBM 所有

**Copyright IBM Corporation 1989, 2011.**

美国政府用户受限权利 – 使用、复制或披露受与 IBM Corp. 签订的 GSA ADP Schedule Contract 的限制。

---

# 前言

IBM® SPSS® Statistics 是一种用于分析数据的综合系统。Conjoint 可选附加模块提供本手册中描述的其他分析方法。此 Conjoint 附加模块必须与 SPSS Statistics Core 系统一起使用，并已完全集成到了该系统中。

## 关于 IBM Business Analytics

IBM Business Analytics 软件提供决策者赖以提高业务绩效的完整、一致和准确的信息。包括业务智能、预测分析、财务状况和战略管理以及分析应用程序在内的一整套产品组合让您即刻、清楚地了解当前绩效并依此采取行动，以及能够预测未来的成果。结合丰富的行业解决方案、被证明的实践经验 and 专业的服务，无论公司规模大小，都能促使其获得最高的产能、自信自觉地做出决定并得到更好的成绩。

作为产品组合的一部分，IBM SPSS Predictive Analytics 软件帮助公司预测未来实践并采取积极行动，促使其获得更好的业务成果。全世界的商业政府和学术客户依赖 IBM SPSS 技术，因其具有竞争力的优势，能够吸引、留住和发展客户，同时减少欺诈和减轻风险。通过将 IBM SPSS 软件融入日常运营中，公司成为具有预测性的企业，能够引导和自觉做出决策，以满足业务目标，实现可观的竞争优势。欲知更多信息或联系代表，请访问 <http://www.ibm.com/spss>。

## 技术支持

我们提供有“技术支持”以维护客户。客户可就 IBM Corp. 产品使用或某一受支持硬件环境的安装帮助寻求技术支持。如要联系技术支持，请访问 IBM Corp. 网站，网址为 <http://www.ibm.com/support>。在请求协助时，请准备好您和您组织的 ID 以及支持协议。

## 针对学生的技术支持

如果您是使用任何学生版 IBM SPSS 软件产品的学生，请访问我们专为学生提供的在线教育解决方案 (<http://www.ibm.com/spss/rd/students/>) 页面。如果您是使用大学提供的 IBM SPSS 软件副本的学生，请联系所在大学的 IBM SPSS 产品协调员。

## 客户服务

如果对发货或帐户存在任何问题，请联系您当地的办事处。请先准备好您的序列号以供识别。

## 培训讲座

IBM Corp. 提供公开的以及现场的培训讲座。所有讲座都是以实践小组为特色的。讲座将定期在各大城市开展。有关这些讲座的更多信息，请前往 <http://www.ibm.com/software/analytics/spss/training>。

## 附加出版物

SPSS Statistics: 数据分析指南、SPSS Statistics: Statistical Procedures Companion 和 SPSS Statistics: Advanced Statistical Procedures Companion (由 Marija Norušis 编写, 并已由 Prentice Hall 出版) 作为建议的补充材料提供。这些出版物涵盖 SPSS Statistics Base 模块、Advanced Statistics 模块和 回归模块中的统计过程。无论您是刚开始从事数据分析工作, 还是已准备好使用高级应用程序, 这些书籍都将帮助您最有效地利用在 IBM® SPSS® Statistics 产品中找到的功能。有关其他信息, 包括出版物的内容和示例章节, 请参阅作者的网站: <http://www.norusis.com>

---

# 内容

## 1 联合分析简介 1

完全概要文件方法 . . . . .	1
正交数组 . . . . .	2
实验刺激源 . . . . .	2
收集和分析数据 . . . . .	2

## 部分 I: 用户指南

## 2 生成正交设计 4

定义正交设计的值 . . . . .	5
正交设计选项 . . . . .	6
ORTHOPLAN 命令附加功能 . . . . .	6

## 3 显示设计 7

显示设计标题 . . . . .	8
PLANCARDS 命令附加功能 . . . . .	8

## 4 运行联合分析 9

要求 . . . . .	9
指定计划文件和数据文件 . . . . .	9
指定如何记录数据 . . . . .	10
的可选子命令 . . . . .	10

## 部分 II: 示例

## 5 使用联合分析为地毯清洁用品偏好建模 14

生成正交设计 . . . . .	14
创建实验刺激源: 显示设计 . . . . .	17

运行分析 . . . . .	19
效用得分 . . . . .	21
系数 . . . . .	22
相对重要性 . . . . .	22
相关性 . . . . .	23
反向 . . . . .	23
运行模拟个案 . . . . .	23
模拟的偏好概率 . . . . .	24

## 附录

<b>A 样本文件</b>	<b>25</b>
<b>B 注意事项</b>	<b>32</b>
<b>参考书目</b>	<b>34</b>
<b>索引</b>	<b>36</b>

# 联合分析简介

联合分析是一种用于开发有效产品设计的市场研究工具。使用联合分析，研究人员可以回答以下问题，如：哪些产品属性对消费者重要及哪些产品属性对消费者不重要？消费者心中最喜欢及最不喜欢的产品属性级别有哪些？领先竞争对手的产品与我们现有或提出的产品的偏好市场份额是多少？

联合分析的的优点在于让被调查者以与消费者大概进行的相同方式进行选择—通过互相交易功能。

例如，假设您想预订航班。您可以坐在狭窄的座位或宽敞的座位上。如果这是唯一的考虑因素，您的选择会很清楚。您很可能喜欢宽敞的座位。或者假设您可以选择票价：225 美元或 800 美元。仅就价格而言，不考虑任何其他因素，您会更偏爱较低的价格。最后，假设您可以搭乘需要两个小时的直达航班或需要五个小时的停站航班。多数人会选择直达航班。

上述方法的缺点在于一次单个属性单独代表选项。联合分析表示属性集合定义的产品之间的选项。以下选项说明了这一点：您喜欢座位狭窄 225 美元的停站航班还是座位宽敞 800 美元的直达航班？如果舒适、价格和持续时间是相关属性，那么可能有八个产品：

产品	舒适	价格	持续时间
1	狭窄	\$225	2 小时
2	狭窄	\$225	5 小时
3	狭窄	\$800	2 小时
4	狭窄	\$800	5 小时
5	宽敞	\$225	2 小时
6	宽敞	\$225	5 小时
7	宽敞	\$800	2 小时
8	宽敞	\$800	5 小时

给定以上选项，产品 4 可能是最不喜欢的，而产品 5 可能是最喜欢的。被调查者对于其他产品的偏好由对于被调查者什么重要来隐式决定。

使用联合分析，您可以确定每个属性的相关重要性以及最喜欢每个属性的什么级别。如果最喜欢的产品由于某些原因并不可行（如价格），您会知道第二最喜欢的选项。如果您有被调查者的其他信息（如背景人口统计信息），您可能能够识别可以包装不同产品的细分市场。例如，商业旅行者和学生旅行者可能有不同的偏好，可以通过不同的产品来满足。

## 完全概要文件方法

联合使用**完全概要文件**（也称为**完全概念**）**方法**，其中被调查者根据偏好对一组概要文件或卡片进行评级、排序或评分。每个概要文件描述一个完整的产品或服务，对于感兴趣的所有因子（属性）由因子级别的不同组合构成。

## 正交数组

如果涉及多个因子并且每个因子有多个级别，完全概要文件方法的潜在问题很快就变得非常明显。所有可能的级别组合所得的概要文件总数变得太大，被调查者无法以一种有意义的方法评级或评分。要解决此问题，完全概要文件方法使用称为**分数因子设计**，表示所有可能的因子级别组合的适合分数。结果集合，称为**正交数组**，旨在获取每个因子级别的主效应。一个因子级别与另一个因子级别之间的交互假设为可以忽略。

生成正交设计过程用于生成正交数组，通常是联合分析的起始点。还允许您生成因子级别的组合，称为**保持个案**，由主体评估，但是不用于构建偏好模型。相反，他们用于查看模型的有效性。

## 实验刺激源

正交设计中的每个因子级别组合代表研究的产品的不同版本，应该以单个产品概要文件的形式展示给主体。这帮助被调查者只将重点集中在当前进行评估的一个产品。应该通过确保除功能的不同组合外所有概要文件的物理外观都类似，标准化刺激源。

使用显示设计过程简化产品概要文件的创建。进行生成正交设计过程生成的或由用户输入的设计，并生成一组即开即用格式的产品概要文件。

## 收集和分析数据

由于通常有大量主体间的偏好变异，许多联合分析将重点集中在单个主体。要归纳结果，选择目标总体的主体随机样本，以便可以检查小组结果。

联合研究中的样本大小变化极大。在一个报告 (Cattin 和 Wittink, 1982) 中，作者称商业联合研究中的样本大小的范围通常是从 100 到 1,000，其中 300 到 550 是最典型的范围。在另一项研究 (Akaah 和 Korgaonkar, 1988) 中，发现典型的是较小的样本大小 (小于 100)。通常样本应具有足够的大小以确保可靠性。

一旦选择了样本，研究人员将管理每个被调查者的概要文件或卡片集合。联合过程允许三种数据记录方法。在第一种方法中，要求主体将偏好得分分配到每个概要文件。当使用 Likert 刻度或当要求主体分配从 1 到 100 的数字以指示偏好时，这种方法是典型方法。在第二种方法中，要求主体将等级分配到从 1 到概要文件总数的每个概要文件。在第三种方法中，要求主体根据偏好对概要文件排序。使用最后一种方法，研究人员以每个主题给定的顺序记录概要文件编号。

使用联合过程进行数据分析 (仅通过命令语法可以获得)，得到每个因子级别的一个效用得分，称为**部分价值**。这些效用得分与回归系数类似，为每个因子级别提供了对偏好的定量度量，其值愈大相应的偏好愈强烈。以公共空单位表达部分价值，允许将其加在一起，对于任何因子级别组合给定总效用或总体偏好。部分价值然后构成一个模型，用于预测任何产品概要文件的偏好，包括概要文件，称为**模拟个案**，实际上在实验中并未提供。

从联合分析中获取的信息可以应用于各种市场研究问题。可以用于调查产品设计、市场份额、战略广告、成本效益分析和市场细分等领域。

尽管此手册的重点集中在市场研究应用，联合分析在几乎任何需要测量人们的观点或判断力的科学或商业领域都非常有用。



# 部分 I: 用户指南

# 生成正交设计

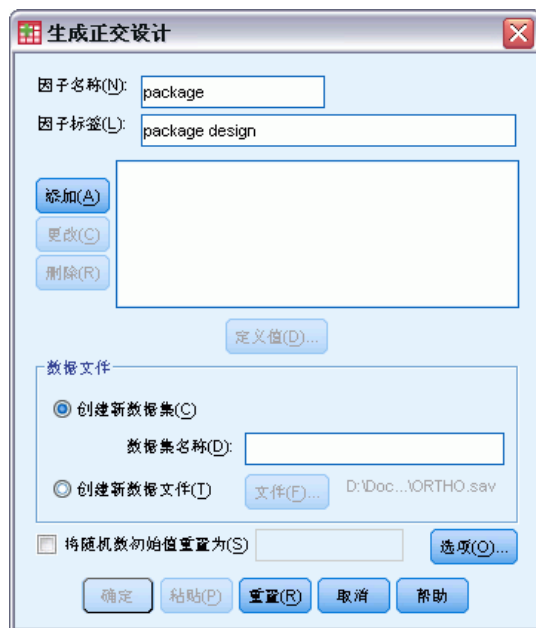
“生成正交设计”生成一个包含正交主效应设计的数据文件，允许多个因子的统计检验，而无需检验每个因子级别组合。可以使用显示设计过程显示此设计，可以通过其他过程（如联合过程）使用数据文件。

**示例。** 低价航线启动对确定构成产品的各种因子的潜在消费者的相对重要性感兴趣。价格明显是一个主要因子，但是其他因子有何重要，如座位大小、停站数量以及是否包括饮料/快餐服务？由于大量概要文件，一项要求被调查者对代表所有可能的因子组合的产品概要文件进行评级的调查并不合理。生成正交设计过程创建缩减的产品概要文件集合，该集合如果要包括在调查中，其大小足够小；而如果要评估每个因子的相对重要性，则又足够大。

## 生成正交设计

- ▶ 从菜单中选择：  
数据 > 正交设计 > 生成...

图片 2-1  
生成正交设计对话框



- ▶ 定义至少一个因子。在因子名称文本框中输入一个名称。因子名称可以是任何有效变量名称，status\_ 或 card\_ 除外。您还可以分配一个可选因子标签。
- ▶ 单击添加添加因子名称和可选标签。要删除一个因子，在列表中选择并单击删除。要修改因子名称或标签，在列表中选择，修改名称或标签，并单击更改。
- ▶ 通过选择因子并单击定义值定义每个因子的值。

**数据文件。**允许您控制正交设计的目标。您可以在当前会话中将设计保存到新数据集或外部数据文件中。

- **创建新数据集。**在当前会话中创建一个包含计划生成的因子和个案的新数据集。
- **新建数据文件。**创建一个包含计划生成的因子和个案的外部数据文件。缺省情况下，此数据文件命名为 ortho.sav，保存到当前目录中。单击文件指定一个不同的文件名称和目标。

**将随机数种子重置为。**将随机数种子重置为指定值。种子可以从 0 到 2,000,000,000 的任何整数值。在会话内，程序在每次您生成一组随机数时使用不同的种子，从而生成不同的结果。如果您想复制相同的随机数，您应该先设置种子值，然后生成您的第一个设计并在以后每次生成设计时将种子重置为相同值。

根据需要，您可以：

- 单击选项指定正交设计中最小个案数，选择保持个案。

## 定义正交设计的值

图片 2-2  
生成设计定义值对话框

值和标签 r		自动填充	
	值(V)	标签	从 1 至(M)
1(1):	1	A*	<input type="text"/>
2(2):	2	B*	<input type="text"/>
3(3):	3	C*	<input type="text"/>
4(4):	<input type="text"/>	<input type="text"/>	<input type="text"/>
5(5):	<input type="text"/>	<input type="text"/>	<input type="text"/>
6(6):	<input type="text"/>	<input type="text"/>	<input type="text"/>
7(7):	<input type="text"/>	<input type="text"/>	<input type="text"/>
8(8):	<input type="text"/>	<input type="text"/>	<input type="text"/>
9(9):	<input type="text"/>	<input type="text"/>	<input type="text"/>

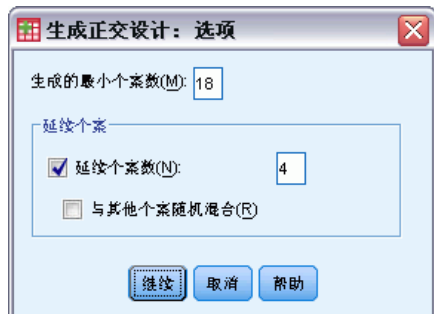
您必须将值分配给选定的一个或多个因子的每个级别。在值和标签之后显示因子名称。

输入每个因子值。您可以选择给这些值以描述性标记。如果您不将标签分配到值，对应于值的标签将自动分配（即，值 1 分配标签 1，值 3 分配标签 3，依此类推）。

**自动填充。**允许您用从 1 开始的连续值自动填充“值”框。输入最大值，并单击填充以填入值。

## 正交设计选项

图片 2-3  
生成正交设计选项对话框



**要生成的个案的最小个数。**指定计划的最少个案个数。选择一个正整数，该数小于等于可以从因子级别的所有可能组合中形成的个案总数。如果您不显式指定要生成的最小个案数，将生成正交计划所需的最小个案数。如果正交过程无法生成至少最小值所需的概要文件数，将生成适合指定因子和级别的最大数。注意设计不一定包括精确的指定个案数，而是正交设计中可能最小的个案数，将此值用作最小值。

**保持个案。**您可以定义主体评估的保持个案，但是不包括在联合分析中。

- **保持个案数。**除规则计划个案以外，还创建保持个案。保持个案由主体判断，但是当联合过程估计效用时不使用。您可以指定任何正整数，该数小于等于可以从因子级别的所有可能组合中形成的个案总数。保持个案从另一个随机计划生成，而非主效应实验计划。保持个案不重复实验概要文件或相互重复。缺省情况下，没有生成任何保持个案。
- **与其他个案随机混合。**将保持个案与试验个案随机混合。当取消选择此选项时，保持个案在实验个案之后单独出现。

## ORTHOPLAN 命令附加功能

使用命令语法语言还可以：

- 向活动数据集追加正交设计，而非新建一个设计。
- 在生成正交设计之前而非创建设计之后指定模拟个案。

请参见命令语法参考以获取完整的语法信息。

# 显示设计

显示设计过程允许您打印实验设计。您可以以草稿列表格式打印设计，或将其打印为您可以展示给联合研究中的主体的概要文件。此过程可以显示由“生成正交设计”过程创建的设计或在活动数据集中显示的任何设计。

## 显示正交设计

- ▶ 从菜单中选择：  
数据 > 正交设计 > 显示...

图片 3-1  
“显示设计”对话框



- ▶ 将一个或多个因子移至因子列表。
- ▶ 选择概要文件在输出中的显示格式。

**格式。** 可以选择以下格式选项中的一个或多个：

- **为试验者列出。** 以草稿格式显示设计，该设计将保持概要文件从试验概要文件中区分出来，并在试验概要文件和保持概要文件之后单独列出仿真概要文件。
- **主题概要文件。** 生成可以展示给主体的概要文件。此格式不区分保持概要文件，也不生成仿真概要文件。

根据需要，您可以：

- 单击标题定义概要文件的页眉和页脚。

## 显示设计标题

图片 3-2  
“显示设计：标题”对话框



**概要文件标题。** 输入最长可为 80 个字符的概要文件标题。在主对话框中，如果已选择为试验者列出，则标题出现在输出的顶部；如果已选择主题概要文件，则标题出现在每个新概要文件的顶部。对于主题概要文件，如果在标题的任何位置指定了特殊字符序列 )CARD，过程会将其替换为序列概要文件编号。未对为试验者列出转换此字符序列。

**概要文件页脚。** 输入最长可为 80 个字符的概要文件页脚。在主对话框中，如果已选择为试验者列出，则页脚出现在输出的底部；如果已选择主题概要文件，则标题出现在每个概要文件的底部。对于主题概要文件，如果在页脚的任何位置指定了特殊字符序列 )CARD，过程会将其替换为序列概要文件编号。未对为试验者列出转换此字符序列。

## PLANCARDS 命令附加功能

使用命令语法语言还可以：

- 将主题概要文件写入一个外部文件（使用 `OUTFILE` 子命令）。

请参见命令语法参考以获取完整的语法信息。

# 运行联合分析

联合过程尚不可使用图形用户界面。要获取联合分析，您必须将 `CONJOINT` 命令的命令语法输入到语法窗口，然后运行。

- 有关完整联合分析（包括生成和显示正交设计）上下文中 `CONJOINT` 命令的命令语法示例，请参见第 5 章。
- 有关完整命令语法信息，有关 `CONJOINT` 命令，请参见命令语法参考。

## 从语法窗口运行命令

从菜单中选择：

文件 > 新建 > 语法...

这样可以打开语法窗口。

- ▶ 输入 `CONJOINT` 命令的命令语法。
- ▶ 突出显示语法窗口中的命令，单击语法编辑器工具栏上的“运行”按钮（指向右边的三角形）。

参见 Core 系统用户指南获取有关在语法窗口中运行命令的更多信息。

## 要求

联合过程需要两个文件—一个数据文件和一个计划文件—以及数据的记录方式规格（例如，每个数据点是从 1 到 100 的偏好得分）。计划文件由主体评估的产品概要文件集合构成，应该使用[生成正交设计](#)过程生成。数据文件包含从主体收集的这些概要文件的偏好得分或等级。分别使用 `PLAN` 和 `DATA` 子命令指定计划和数据文件。使用 `SEQUENCE`、`RANK` 或 `SCORE` 子命令指定数据记录方法。以下命令语法显示最小规格：

```
CONJOINT PLAN='CPLAN.SAV' /DATA='RUGRANKS.SAV'  
/SEQUENCE=PREF1 TO PREF22.
```

## 指定计划文件和数据文件

`CONJOINT` 命令为指定计划文件和数据文件提供了大量选项。

- 您可以显式指定两个文件的文件名。例如：

```
CONJOINT PLAN='CPLAN.SAV' /DATA='RUGRANKS.SAV'
```

- 如果只指定了一个计划文件或一个数据文件，则 `CONJOINT` 命令会读取指定文件，并使用活动数据集作为另一个文件。例如，如果您指定一个数据文件，却忽略了一个计划文件（您不能两个都忽略），活动数据集将用作计划文件，如下例所示：

```
CONJOINT DATA='RUGRANKS.SAV'
```

- 您可以使用星号 (\*) 代替文件名以指示活动数据集，如下例所示：

```
CONJOINT PLAN='CPLAN.SAV' /DATA=*
```

活动数据集用作偏好数据。注意计划文件和数据文件不能都使用星号 (\*)。

## 指定如何记录数据

您必须指定偏好数据的记录方式。可以使用以下三种方式之一记录数据：按顺序、作为等级或作为偏好得分。这三种方式由 **SEQUENCE**、**RANK** 和 **SCORE** 子命令指示。您必须指定这些子命令的唯一一个作为 **CONJOINT** 命令的一部分。

### SEQUENCE 子命令

**SEQUENCE** 子命令指示数据按顺序记录，因此数据文件中的每个数据点是一个概要文件编号，以最喜欢的概要文件开始，以最不喜欢的概要文件结束。如果要求主体从最喜欢到最不喜欢对概要文件进行排序，这就是数据的记录方式。研究人员记录哪个概要文件编号是第一个，哪个概要文件编号是第二个，依此类推。

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/SEQUENCE=PREF1 TO PREF22.
```

- 变量 **PREF1** 包含正交计划中 22 个概要文件中最喜欢的概要文件的概要文件编号。变量 **PREF22** 包含计划中最不喜欢的概要文件的概要文件编号。

### RANK 子命令

**RANK** 子命令指示每个数据点是一个等级，以概要文件 1 的等级开始，然后是概要文件 2 的等级，依此类推。如果要求主体将等级分配到每个概要文件，从 1 到 n，其中 n 是概要文件数，这就是数据的记录方式。等级越低表明偏爱的程度越高。

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22.
```

- 变量 **RANK1** 包含正交计划中 22 个概要文件中概要文件 1 的等级。变量 **RANK22** 包含概要文件 22 的等级。

### SCORE 子命令

**SCORE** 子命令指示每个数据点是一个分配到概要文件的偏好得分，以概要文件 1 的得分开始，然后是概要文件 2 的得分，依此类推。例如，可能通过要求主体分配从 1 到 100 的数字以显示其偏爱概要文件的程度来生成此数据类型。得分越高表明偏爱的程度越高。

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/SCORE=SCORE1 TO SCORE22.
```

- 变量 **SCORE1** 包含概要文件 1 的得分，**SCORE22** 包含概要文件 22 的得分。

## 的可选子命令

**CONJOINT** 命令提供大量可选子命令，提供超越所需的其他控制和功能。



## SUBJECT 子命令

**SUBJECT** 子命令允许您从数据文件中指定一个变量，用作主体的标识。如果您不指定一个主体变量，**CONJOINT** 命令假定数据文件中的所有个案来自一个主体。以下示例指定文件 `rugranks.sav` 中的变量 `ID` 将用作主体标识。

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
/SUBJECT=ID.
```

## FACTORS 子命令

**FACTORS** 子命令允许您指定模型，描述因子与等级或得分之间的期望关系。如果您未为因子指定模型，**CONJOINT** 会假定一个离散模型。您可以指定以下四种模型之一：

**DISCRETE**。**DISCRETE** 模型指示因子级别是分类级别，对于因子与等级或得分之间的关系不作任何假定。这是缺省值。

**LINEAR**。**LINEAR** 模型指示因子与等级或得分之间的期望线性关系。您可以使用关键字 **MORE** 和 **LESS** 指定线性关系的期望方向。**MORE** 指示期望喜欢更高级别的因子，而 **LESS** 指示期望喜欢更低级别的因子。指定 **MORE** 或 **LESS** 不会影响效用的估计。只用于识别估计不匹配期望方向的主体。

**IDEAL**。**IDEAL** 模型指示因子与等级或得分之间的期望二次关系。假设因子有一个理想级别，且距此理想点（任意方向）的距离与减少偏好相关联。使用此模型描述的因子至少有三个级别。

**ANTI IDEAL**。**ANTI IDEAL** 模型指示因子与等级或得分之间的期望二次关系。假设因子有一个最坏级别，且距此理想点（任意方向）的距离与增加偏好相关联。使用此模型描述的因子至少有三个级别。

以下命令语法提供使用 **FACTORS** 子命令的示例：

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
/RANK=RANK1 TO RANK22 /SUBJECT=ID
/FACTORS=PACKAGE BRAND (DISCRETE) PRICE (LINEAR LESS)
SEAL (LINEAR MORE) MONEY (LINEAR MORE).
```

- 注意包装和品牌建模为离散。

## PRINT 子命令

**PRINT** 子命令允许您控制表格输出的内容。例如，如果您有大量主体，您可以选择将输出仅限于摘要结果，忽略每个主体题的详细输出，如下例所示：

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
/RANK=RANK1 TO RANK22 /SUBJECT=ID
/PRINT=SUMMARYONLY.
```

您还可以选择输出是否包括实验数据的分析，记录文件中包括的任何模拟个案结果，二者或二者皆不。模拟个案不依据主体评估，但是代表您感兴趣的产品概要文件。联合过程使用实验数据的分析，为每个模拟概要文件预测相关偏好。在下列示例中，不显示每个主体的详细输出，输出限于模拟的结果：

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
```

```
/RANK=RANK1 TO RANK22 /SUBJECT=ID  
/PRINT=SIMULATION SUMMARYONLY.
```

### **PLOT 子命令**

PLOT 子命令控制输出中是否包括图。就像表格输出（PRINT 子命令）一样，您可以控制输出是否限于摘要结果或包括每个主体的结果。缺省情况下，没有生成任何图。在下列示例中，输出包括所有可用图：

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22 /SUBJECT=ID  
/PLOT=ALL.
```

### **UTILITY 子命令**

UTILITY 子命令以 IBM® SPSS® Statistics 格式写入数据文件，包含每个主体的详细信息。包括 DISCRETE 因子的效用、LINEAR、IDEAL 和 ANTIIDEAL 因子的斜率和二次函数、回归常数和估计的偏好得分。这些值然后可以用于进一步分析或用于使用其他过程制作其他图。以下示例创建一个名为 rugutil.sav 的效用文件：

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22 /SUBJECT=ID  
/UTILITY='RUGUTIL.SAV'.
```

# 部分 II:

## 示例

# 使用联合分析为地毯清洁用品偏好建模

在一个常用的联合分析示例 (Green 和 Wind, 1973) 中，一家公司非常重视一种新型地毯清洁用品的市场营销，希望检验以下五种因素对消费者偏好的影响—包装设计、品牌名称、价格、优秀家用品标志和退货保证。包装设计有三个因子水平，每个因子水平因刷体位置而不同；有三个品牌名称 (K2R、Glory 和 Bissell)；有三个价格水平；最后两个因素各有两个级别 (有或无)。下表显示了在地毯清洁用品调查中所使用的变量以及这些变量的变量标签和变量值。

表 5-1  
地毯清洁用品调查中的变量

变量名	变量标签	值标签
package	package design	A*、B*、C*
brand	brand name	K2R、Glory、Bissell
price	price	\$1.19, \$1.39, \$1.59
seal	Good Housekeeping seal	no、yes
money	money-back guarantee	no、yes

可能有其他描绘地毯清洁用品的因子和因子级别，但是管理部门只对这些感兴趣。这在联合分析中是重点。您只想选择您认为对主体偏好 (因变量) 影响最大的那些因子 (自变量)。使用联合分析，您将基于这五个因子为客户偏好建立一个模型。

此示例使用以下数据文件中的信息：carpet\_prefs.sav 包含从主体收集来的数据，carpet\_plan.sav 包含要调查的产品概要文件，conjoint.sps 包含运行分析所需的命令语法。有关详细信息，请参阅附录 A 中的样本文件中的 IBM SPSS Conjoint 20。

## 生成正交设计

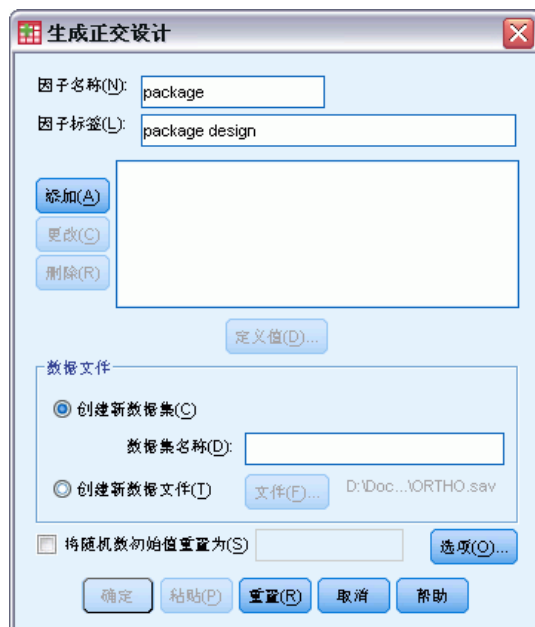
联合分析中的第一步是创建作为产品概要文件呈现给主体的因子级别组合。由于即使是少量因子和每个因子的若干级别也会导致无法管理的大量潜在产品概要文件，因此您需要生成一个代表子集，称为**正交数组**。

生成正交设计过程创建一个正交数组—也称为**正交设计**—，并在数据文件中存储信息。与大部分过程不同，在运行生成正交设计过程之前不需要活动数据集。如果您没有活动数据集，您可以选择创建一个，从您在对话框中选择的选项生成变量名称、变量标签和值标签。如果您已经有一个活动数据集，您可以将其替换或将正交设计保存为单独的数据文件。

创建正交设计：

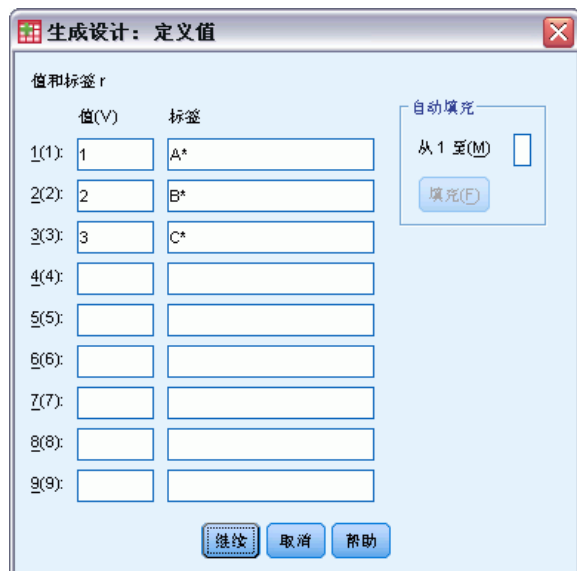
- ▶ 从菜单中选择：  
数据 > 正交设计 > 生成...

图片 5-1  
生成正交设计对话框



- ▶ 在因子名称文本框中输入 `package`，在因子标签文本框中输入 `package design`。
- ▶ 单击添加。  
这将创建一个标有 `package` ‘`package design`’ (?) 的项目。选择此项。
- ▶ 单击定义值。

图片 5-2  
生成设计定义值对话框



- ▶ 输入值 1、2 和 3 代表包装设计 A\*、B\* 和 C\*。同时输入标签 A\*、B\* 和 C\*。

- ▶ 单击**继续**。

您现在想为剩余因子 brand、price、seal 和 money 重复此过程。使用下表中的值和标签，包括您已为 package 输入的值。

因子名称	因子标签	值	标签
package	package design	1, 2, 3	A*、B*、C*
brand	brand name	1, 2, 3	K2R、Glory、Bissell
price	price	1.19, 1.39, 1.59	\$1.19, \$1.39, \$1.59
seal	Good Housekeeping seal	1, 2	no、yes
money	money-back guarantee	1, 2	no、yes

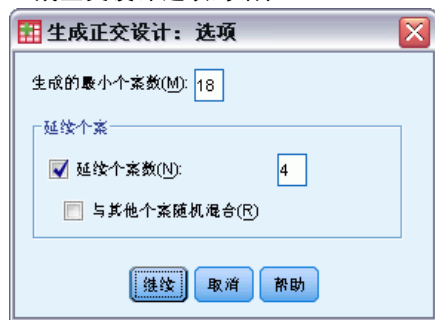
完成因子规格后：

- ▶ 在“数据文件”组中，保留缺省**创建新数据集**，并输入一个数据集名称。生成的设计将在当前会话中以指定的名称保存到新数据集。
- ▶ 选择将随机数种子重置为并输入值 2000000。

生成正交设计需要一组随机数。如果您想重复设计—在本例中为用于提供的个案研究的设计—您需要设置种子值，然后生成设计并在以后每次生成设计时将该值重置为相同值。使用种子值 2000000 生成用于此个案研究的设计。

- ▶ 单击**选项**。

图片 5-3  
生成正交设计选项对话框



- ▶ 在要生成的最小个案数文本框中，键入 18。

缺省情况下，将生成正交数组所需的最小个案数。过程确定需要受管理的个案数，以允许估计效用。您还可以按如上操作指定要生成的最小个案数。您可能想进行这一操作，因为缺省的最小个案数太小，根本没用，或者您有实验设计考虑，需要某个最小个案数。

- ▶ 选择保持个案数并键入 4。

保持个案由主体判断，但是联合分析不会用其来估计效用。它们用于查看估计效用的有效性。保持个案从另一个随机计划生成，而非实验正交计划。

- ▶ 在“生成正交设计：选项”对话框中单击**继续**。
- ▶ 在“生成正交设计”对话框中单击**确定**。

图片 5-4  
地毯清洁用品实例的正交设计

	包装	品牌	价格	奖章	退款	优选	STATUS_	CARD_
14	A*	Glory	\$1.19	是	否	11.00	设计	14
15	B*	K2R	\$1.39	是	是	1.00	设计	15
16	A*	K2R	\$1.19	否	否	15.00	设计	16
17	A*	Bissell	\$1.59	否	是	17.00	设计	17
18	B*	Bissell	\$1.19	否	否	6.00	设计	18
19	A*	Bissell	\$1.59	是	否	21.00	延续	19
20	C*	K2R	\$1.19	是	否	3.00	延续	20
21	A*	Glory	\$1.59	否	否	22.00	延续	21
22	A*	Bissell	\$1.19	否	否	18.00	延续	22

正交设计将显示在数据编辑器中，通过显示值标签而非实际数据值可以达到最佳查看效果。此操作可通过从“视图”菜单中选择值标签来完成。

数据文件中的变量是用于指定设计的因子。在设计中每个个案代表一个产品概要文件。注意，数据文件中出现两个附加变量 CARD\_ 和 STATUS\_。CARD\_ 将序列号分配给每个用于识别概要文件的概要文件。STATUS\_ 指示概要文件是否是实验设计（前 18 个个案）、保持个案（最后 4 个个案）或模拟个案（将在此个案研究的稍后一个主题中讨论）的一部分。

正交设计是数据分析所需的输入。因此，您想将设计保存为数据文件。为方便起见，当前设计已保存在 carpet\_plan.sav（正交设计也称为计划）。

## 创建实验刺激源：显示设计

一旦您创建了一个正交设计，您将想用其创建产品概要文件以供主体评级。您可以在单个表中获取概要文件列表或在单独表中显示每个概要文件。

显示正交设计：

- ▶ 从菜单中选择：  
数据 > 正交设计 > 显示...

图片 5-5  
“显示设计”对话框



- ▶ 选择因子的 package、brand、price、seal 和 money。
- ▶ 变量 STATUS\_ 和 CARD\_ 中包含的信息被自动包括在输出中，因此不需要选择。
- ▶ 在格式组中选择为试验者列出。这会使整个正交设计显示在单个表中。
- ▶ 单击确定。

图片 5-6  
显示正交设计：单个表布局

	卡标识	package design	brand name	price	Good Housekeepin g sale	money-back quarantee
1	1	A*	Glory	\$1.39	yes	no
2	2	B*	K2R	\$1.19	no	no
3	3	B*	Glory	\$1.39	no	yes
4	4	C*	Glory	\$1.59	no	no
5	5	C*	Bissell	\$1.39	no	no
6	6	A*	Bissell	\$1.39	no	no
7	7	B*	Bissell	\$1.59	yes	no
8	8	A*	K2R	\$1.59	no	yes
9	9	C*	K2R	\$1.39	no	no
10	10	C*	Glory	\$1.19	no	yes
11	11	C*	K2R	\$1.59	yes	no
12	12	B*	Glory	\$1.59	no	no
13	13	C*	Bissell	\$1.19	yes	yes
14	14	A*	Glory	\$1.19	yes	no
15	15	B*	K2R	\$1.39	yes	yes
16	16	A*	K2R	\$1.19	no	no
17	17	A*	Bissell	\$1.59	no	yes
18	18	B*	Bissell	\$1.19	no	no
19	19	A*	Bissell	\$1.59	yes	no
20	20	C*	K2R	\$1.19	yes	no
21	21	A*	Glory	\$1.59	no	no
22	22	A*	Bissell	\$1.19	no	no



输出和数据编辑器中显示的正交设计外观类似—每个概要文件一行，因子作为列。但注意，列标题是变量标签而非您在数据编辑器中看到的变量名称。同时注意保持个案与脚注相同。试验者会对此感兴趣，但是您当然不想让主体知道哪个个案是保持个案（如有）。

根据您创建和提供最终产品概要文件的方式，您可能想将此表保存为 HTML、Word/RTF、Excel 或 PowerPoint 文件。此操作可通过在浏览器中选择表，右键单击并选择导出轻松完成。此外，如果您正在使用导出的版本创建最终产品概要文件，请确保删去保持个案的脚注。

也许通过为每个产品概要文件生成单独的表可以更好地满足您的调查需要。此选项有助于导出到 PowerPoint，因为每个表（产品概要文件）都被置于单独的 PowerPoint 幻灯片上。

在单独表中显示每个概要文件：

- ▶ 单击“对话框调用”按钮，选择显示设计。
- ▶ 取消选择为试验者列出并选择主体概要文件。
- ▶ 单击确定。

图片 5-7

显示正交设计：多表布局

配置文件编号 1

Card ID	package design	brand name	price	Good Housekeeping seal	money-back guarantee
1	A*	Glory	\$1.39	yes	no

配置文件编号 2

Card ID	package design	brand name	price	Good Housekeeping seal	money-back guarantee
2	B*	K2R	\$1.19	no	no

每个产品概要文件的信息显示在单独表中。此外，保持个案与其余个案不相上下，因此使用单个表布局时无需删除保持个案的标识。

## 运行分析

您已生成正交设计并学习了如何显示相关产品概要文件。您现在准备学习如何运行联合分析。

图片 5-8  
地毯清洁用品实例的偏好数据

	ID	PREF1	PREF2	PREF3	PREF4	PREF5	PREF6	PREF7
1	1	13	15	1	20	14	7	11
2	2	15	7	18	2	12	3	11
3	3	2	18	14	16	22	13	20
4	4	13	10	20	14	2	18	16
5	5	13	18	2	10	20	15	9
6	6	15	2	3	12	18	7	20
7	7	13	7	15	18	2	3	10
8	8	15	7	13	4	6	16	8
9	9	20	9	10	11	4	5	13
10	10	8	21	19	17	4	11	12

从主体收集到的偏好数据存储于 carpet\_prefs.sav 中。数据由 10 个主体的响应构成，每个都由变量 ID 的唯一值标识。主体被要求按照从最喜欢到最不喜欢的顺序对 22 个产品特征进行排序。变量 PREF1 到 PREF22 包含相关产品概要文件的 ID，即来自 carpet\_plan.sav 的卡 ID。例如，主体 1 最喜欢概要文件 13，因此 PREF1 的值是 13。

数据分析任务需要使用命令语法—具体而言，需要使用 CONJOINT 命令。在文件 conjoint.sps 中提供有所需的命令语法。

```
CONJOINT PLAN='file specification'
  /DATA='file specification'
  /SEQUENCE=PREF1 TO PREF22
  /SUBJECT=ID
  /FACTORS=PACKAGE BRAND (DISCRETE)
  PRICE (LINEAR LESS)
  SEAL (LINEAR MORE) MONEY (LINEAR MORE)
  /PRINT=SUMMARYONLY.
```

- PLAN 子命令指定包含正交设计的文件—在本例中为 carpet\_plan.sav。
- DATA 子命令指定包含偏好数据的文件—在本例中为 carpet\_prefs.sav。如果您选择偏好数据作为活动数据集，您可以将文件规格替换为星号 (\*)（不加引号）。
- SEQUENCE 子命令指定偏好数据中的每个数据点都是一个概要文件编号，以最喜欢的概要文件开始，并以最不喜欢的概要文件结束。
- SUBJECT 子命令指定变量 ID 标识主体。
- FACTORS 子命令指定描述偏好数据与因子级别之间的期望关系的模型。指定的因子是指在 PLAN 子命令上命名的计划文件中定义的变量。
- 当因子级别为分类级别且未对级别和数据之间的关系作出任何假设时，使用关键字 DISCRETE。这是分别代表包装设计和品牌名称的因子 package 和 brand 的个案。如果因子未由 DISCRETE、LINEAR、IDEAL 或 ANTIIDEAL 其中之一标注或未包括在 FACTORS 子命令上，则假定为 DISCRETE。
- 关键字 LINEAR 用于其余因子，指示数据期望与因子线性相关。例如，偏好通常期望与价格线性相关。您还可以使用关键字 IDEAL 和 ANTIIDEAL 指定二次模型（本例中未使用）。

- 关键字 **MORE** 和 **LESS** (在 **LINEAR** 之后) 指示关系的期望方向。由于我们期望较低价格的较高偏好，因此关键字 **LESS** 用于 **price**。但是，如果我们期望优秀家庭用品认可标志和退款保证的较高偏好，则关键字 **MORE** 用于 **seal** 和 **money** (调用“将所有这些因子的级别对于 **no** 设为 1，对于 **yes** 设为 2”)。

指定 **MORE** 或 **LESS** 不会更改系数的符号或影响效用的估计。这些关键字只用于标识估计值不匹配期望方向的主体。类似地，选择 **IDEAL** 而非 **ANTIIDEAL**，反之亦然，不会影响系数或效用。

- **PRINT** 子命令指定输出只整体包含主体组的信息 (**SUMMARYONLY** 关键字)。每个主体的信息不单独显示。

尝试运行此命令语法。确保您已包含访问 **carpet\_prefs.sav** 和 **carpet\_plan.sav** 的有效路径。有关所有选项的完整说明，请参见命令语法参考中的 **CONJOINT** 命令。

## 效用得分

图片 5-9  
效用得分

		实用工具估计	标准误
package	A*	-2.233	.192
	B*	1.867	.192
	C*	.367	.192
brand	K2R	.367	.192
	Glory	-.350	.192
	Bissell	-.017	.192
price	\$1.19	-6.595	.988
	\$1.39	-7.703	1.154
	\$1.59	-8.811	1.320
seal	no	2.000	.287
	yes	4.000	.575
money	no	1.250	.287
	yes	2.500	.575
(常数)		12.870	1.282

此表显示效用（部分价值）得分及其每个因子级别的标准误。效用值越高表示偏好越强烈。正如期望的，价格和效用之间有一种逆关系，其中较高价格对应较低效用（负值愈大是指效用愈低）。正如期望的，认可标志和退款保证的存在对应较高的效用。

由于效用都在公共单元中表示，因此可以一起添加以给定任何组合的**总效用**。例如，包装设计为 B\*，品牌为 K2R，价格为 \$1.19 并且没有认可标志和退款保证的清洁用品的总效用是：

$$\begin{aligned} & \text{utility}(\text{package B*}) + \text{utility}(\text{K2R}) + \text{utility}(\$1.19) \\ & + \text{utility}(\text{no seal}) + \text{utility}(\text{no money-back}) + \text{constant} \end{aligned}$$

或

$$1.867 + 0.367 + (-6.595) + 2.000 + 1.250 + 12.870 = 11.759$$

如果清洁用品的包装设计为 C\*, 品牌为 Bissell, 价格为 \$1.59, 且拥有认可标志和退款保证, 那么总效用是:

$$0.367 + (-0.017) + (-8.811) + 4.000 + 2.500 + 12.870 = 10.909$$

## 系数

图片 5-10  
系数

	B 估计
price	-5.542
seal	2.000
money	1.250

此表显示那些指定为 LINEAR 的因子的线性回归系数 (对于 IDEAL 和 ANTIIDEAL 模型, 还有一个二次项)。通过将级别乘以系数来确定特定因子级别的效用。例如, 价格 \$1.19 的预测效用在效用表中列为 -6.595。这是由价格级别 1.19 乘以价格系数 -5.542 所得的值。

## 相对重要性

每个因子的效用值范围 (从最高到最低) 测量因子对于总体偏好的重要性。拥有较大效用范围的因子比拥有较小范围的因子更加重要。

图片 5-11  
重要性值

package	35.635
brand	14.911
price	29.410
seal	11.172
money	8.872

平均重要分数

此表测量每个因子的相对重要性, 称为**重要性**得分或值。通过分别取每个因子的效用范围并除以所有因子效用范围的和来计算该值。因此, 该值代表百分比, 并且具有总和是 100 的属性。注意, 计算是对每个主体单独进行的, 且随后会对所有主体取结果的平均值。

注意, 无论是否使用 SUBJECT 子命令, 总体或摘要效用和正交设计的回归系数都相同, 但重要性通常仍有所不同。对于没有使用 SUBJECT 子命令的摘要结果, 可以直接从摘要效用计算重要性, 就和处理单个主体一样。但是, 在使用 SUBJECT 子命令时, 单个主体的重要性取平均值, 这些平均重要性通常不会匹配使用摘要效用计算的重要性。

结果显示包装设计对总体偏好影响最大。这意味着包含最喜欢的包装的产品概要文件和那些包含最不喜欢的包装的概要文件之间的偏好差异很大。结果同时显示退款保证对于确定总体偏好最不重要。价格非常重要, 但是没有包装设计重要。可能这是因为价格范围并不太大。

## 相关性

图片 5-12  
相关系数

	值	Sig.
Person 的 R	.982	.000
Kendall 的 tau	.892	.000
Kendall 的 tau for Houldouts	.667	.087

此表显示两个统计量，Pearson 的 R 和 Kendall 的 tau，它们测量观察和估计偏好之间的相关性。

此表同时显示保持概要文件的 Kendall 的 tau。记住，保持概要文件（在当前的示例中是四个）由主体评级，但是联合分析过程不会用其估计效用。相反，联合分析过程会计算这些概要文件的观察和预测等级顺序之间的相关性，以检查效用的有效性。

在许多联合分析中，参数的数量与评级的概要文件数接近，这将人工夸大观察和估计得分之间的相关性。在这些个案中，保持概要文件的相关性也许能够更好地指示模型拟合。但是切记，保持概要文件始终生成较低的相关系数。

## 反向

在为 price、seal 和 money 指定 LINEAR 模型时，我们为变量值与该值偏好之间的线性关系选择一个期望的方向（LESS 或 MORE）。联合分析过程跟踪偏好显示与期望关系相反的主体数—例如，高价的较高偏好或退款保证的较低偏好。这些个案称为**反向**。

图片 5-13  
依据因子和主体的反向数

因子	price		3
	money		2
	seal		2
	brand		0
	package		0
主体	1	主体 1	1
	2	主体 2	2
	3	主体 3	0
	4	主体 4	0
	5	主体 5	0
	6	主体 6	1
	7	主体 7	0
	8	主体 8	0
	9	主体 9	1
	10	主体 10	2

此表显示每个因子和每个主体的反向数。例如，这些主体显示 price 的反向。即，他们偏好高价的产品概要文件。

## 运行模拟个案

联合分析的真正本领是能够预测主体不会进行评级的产品概要文件偏好。这些称为**模拟个案**。模拟个案和正交设计的概要文件以及所有保持概要文件一样都是计划的一部分。

输入模拟个案最简单的方法是从数据编辑器进行输入，使用在您生成实验设计时创建的值标签。

在计划文件中输入模拟个案：

- ▶ 在数据编辑器窗口的上新行上，选择一个单元格并从列表中选择所需值（可以通过从视图菜单中选择值标签显示值标签）。为所有变量（因子）重复操作。
- ▶ 为 STATUS\_ 变量的值选择模拟。
- ▶ 为 CARD\_ 变量输入用作标识的整数值。模拟个案应该使用与其他个案不同的编号。

图片 5-14  
地毯清洁用品数据包括模拟个案

包装	品牌	价格	奖章	退款	STATUS_	CARD_
A*	Bissell	\$1.59	否	是	设计	17
B*	Bissell	\$1.19	否	否	设计	18
A*	Bissell	\$1.59	是	否	延续	19
C*	K2R	\$1.19	是	否	延续	20
A*	Glory	\$1.59	否	否	延续	21
A*	Bissell	\$1.19	否	否	延续	22
C*	K2R	\$1.19	是	否	模拟	1
B*	Glory	\$1.19	否	否	模拟	1

该图显示地毯清洁用品研究的一部分计划文件，并添加了两个模拟个案。为方便起见，这些都包括在 carpet\_plan.sav 中。

模拟个案的分析由先前使用过的相同命令语法完成，即文件 conjoint.sps 中的语法。实际上，如果您运行了先前描述的语法，您会注意到输出也包括模拟个案的结果，因为他们都包括在 carpet\_plan.sav 中。

您可以选择和初始分析一起运行模拟—具体操作见此处—或稍后通过在计划文件中包括模拟个案并重新运行 CONJOINT 来运行模拟。有关更多信息，请参见命令语法参考中的 CONJOINT 命令。

## 模拟的偏好概率

图片 5-15  
模拟结果

卡片编号	ID	Maximum Utility <sup>a</sup>	Bradley-Terry-Luce	Logit
1	1	30.0%	43.1%	30.9%
2	2	70.0%	56.9%	69.1%

在三种不同的选择概率模型中，此表给定将每个模拟个案选为最喜欢个案的预测概率。**最大效用模型**确定的概率可由预测会选择概要文件的被调查者数量除以总的被调查者数量获得。对每个响应者而言，预测的选择只是具有最大总效用的概要。**BTL (Bradley-Terry-Luce) 模型**确定的概率是一个概要文件效用与所有模拟概要文件效用的比率，在所有被调查者之间进行平均。**logit 模型**类似于 BTL，不同的是它使用效用的自然对数而非效用本身。在此研究中的所有 10 个主体中，所有三个模型都指示偏好模拟概要文件 2。

# 样本文件

随产品一起安装的样本文件可以在安装目录的 Samples 子目录中找到。对于以下每种语言在“样本”子目录中有单独的文件夹：英语、法语、德语、意大利语、日语、韩语、波兰语、俄语、简体中文、西班牙语和繁体中文。

并非所有样本文件均提供此处的全部语言版本。如果样本文件未提供某种语言的版本，则相应语言文件夹中包含该样本文件的英语版本。

## 描述

以下是对在整个文档的各种示例中使用的样本文件的简要描述。

- **accidents.sav**。该假设数据文件涉及某保险公司，该公司正在研究给定区域内汽车事故的年龄和性别风险因子。每个个案对应一个年龄类别和性别类别的交叉分类。
- **adl.sav**。该假设数据文件涉及在确定针对脑卒中患者的建议治疗类型的优点方面的举措。医师将女性脑卒中患者随机分配到两组中的一组。第一组患者接受标准的物理治疗，而第二组患者则接受附加的情绪治疗。在进行治疗的三个月时间里，将为每个患者进行一般日常生活行为的能力评分并作为原始变量。
- **advert.sav**。该假设数据文件涉及某零售商在检查广告支出与销售业绩之间的关系方面的举措。为此，他们收集了过去的销售数据以及相关的广告成本。
- **aflatoxin.sav**。该假设数据文件涉及对谷物的黄曲霉毒素的检测，该毒素的浓度会因谷物产量的不同（不同谷物之间及同种谷物之间）而有较大变化。谷物加工机从 8 个谷物产量的每一个中收到 16 个样本并以十亿分之几 (PPB) 为单位来测量黄曲霉毒素的水平。
- **anorectic.sav**。在研究厌食/暴食行为的标准症状参照时，研究人员 (Van der Ham, Meulman, Van Strien, 和 Van Engeland, 1997) 对 55 名已知存在进食障碍的青少年进行了调查。其中每名患者每年都将进行四次检查，因此总观测数为 220。在每次观测期间，将对这些患者按 16 种症状逐项评分。但 71 号和 76 号患者的症状得分均在时间点 2 缺失，47 号患者的症状得分在时间点 3 缺失，因此有效观测数为 217。
- **bankloan.sav**。该假设数据文件涉及某银行在降低贷款拖欠率方面的举措。该文件包含 850 位过去和潜在客户的财务和人口统计信息。前 700 个个案是以前曾获得贷款的客户。剩下的 150 个个案是潜在客户，银行需要按高或低信用风险对他们进行分类。
- **bankloan\_binning.sav**。该假设数据文件包含 5,000 位过去客户的财务和人口统计信息。
- **behavior.sav**。在一个经典示例中 (Price 和 Bouffard, 1974)，52 名学生被要求以 10 分的标度对 15 种情况和 15 种行为的组合进行评价，该 10 分的标度介于 0 = 平均值在个人值之上，值被视为相异性。
- **behavior\_ini.sav**。该数据文件包含 behavior.sav 的二维解的初始配置。

- **brakes.sav**。该假设数据文件涉及某生产高性能汽车盘式制动器的工厂的质量控制。该数据文件包含对 8 台专用机床中每一台的 16 个盘式制动器的直径测量。盘式制动器的目标直径为 322 毫米。
- **breakfast.sav**。在一项经典研究中 (Green 和 Rao, 1972), 21 名 Wharton School MBA 学生及其配偶被要求按照喜好程度顺序对 15 种早餐食品进行评价, 从 1 = 他们的喜好根据六种不同的情况加以记录, 从“全部喜欢”到“只带饮料的快餐”。
- **breakfast-overall.sav**。该数据文件只包含早餐食品喜好的第一种情况, 即“全部喜欢”。
- **broadband\_1.sav**。该假设数据文件包含各地区订制了全国宽带服务的客户的数量。该数据文件包含 4 年期间 85 个地区每月的订户数量。
- **broadband\_2.sav**。该数据文件和 broadband\_1.sav 一样, 但包含另外三个月的数据。
- **car\_insurance\_claims.sav**。在别处被提出和分析的 (McCullagh 和 Nelder, 1989) 关于汽车损坏赔偿的数据集。平均理赔金额可以当作其具有 gamma 分布来建模, 通过使用逆联接函数将因变量的均值与投保者年龄、车辆类型和车龄的线性组合关联。提出理赔的数量可以作为尺度权重。
- **car\_sales.sav**。该数据文件包含假设销售估计值、订价以及各种品牌和型号的车辆的物理规格。订价和物理规格可以从 edmunds.com 和制造商处获得。
- **car\_sales\_uprepared.sav**。这是 car\_sales.sav 的修改版本, 不包含字段的任何已转换版本。
- **carpet.sav**。在一个常用示例中 (Green 和 Wind, 1973), 一家公司非常重视一种新型地毯清洁用品的市场营销, 希望检验以下五种因素对消费者偏好的影响—包装设计、品牌名称、价格、优秀家用品标志和退货保证。包装设计有三个因子水平, 每个因子水平因刷体位置而不同; 有三个品牌名称 (K2R、Glory 和 Bissell); 有三个价格水平; 最后两个因素各有两个级别 (有或无)。十名消费者对这些因素所定义的 22 个特征进行了排序。变量优选包含对每个特征的平均等级的排序。低排序与高偏好相对应。此变量反映了对每个特征的偏好的总体度量。
- **carpet\_prefs.sav**。该数据文件所基于的示例和在 carpet.sav 中所描述的一样, 但它还包含从 10 位消费者的每一位中收集到的实际排列顺序。消费者被要求按照从最喜欢到最不喜欢的顺序对 22 个产品特征进行排序。carpet\_plan.sav 中定义了变量 PREF1 到 PREF22 包含相关特征的标识符。
- **catalog.sav**。该数据文件包含某编目公司出售的三种产品的假设每月销售数据。同时还包括 5 个可能的预测变量的数据。
- **catalog\_seasfac.sav**。除添加了一组从“季节性分解”过程中计算出来的季节性因子和附带的日期变量外, 该数据文件和 catalog.sav 是相同的。
- **cellular.sav**。该假设数据文件涉及某便携式电话公司在减少客户流失方面的举措。客户流失倾向分被应用到帐户, 分数范围从 0 到 100。得到 50 分或更高分数的帐户可能会更换提供商。
- **ceramics.sav**。该假设数据文件涉及某制造商在确定新型优质合金是否比标准合金具有更高的耐热性方面的举措。每个个案代表对一种合金的单独检验; 个案中会记录合金的耐热极限。
- **cereal.sav**。该假设数据文件涉及一份 880 人参与的关于早餐喜好的民意调查, 该调查记录了参与者的年龄、性别、婚姻状况以及生活方式是否积极 (根据他们是否每周至少做两次运动)。每个个案代表一个单独的调查对象。



- **clothing\_defects.sav**。这是关于某服装厂的质量控制过程的假设数据文件。检验员要对工厂中每次大批量生产的服装进行抽样检测并清点不合格的服装的数量。
- **coffee.sav**。这是关于六种冰咖啡的认知品牌形象 (Kennedy, Riquier, 和 Sharp, 1996) 的数据文件。对于 23 种冰咖啡特征属性中的每种属性, 人们选择了由该属性所描述的所有品牌。为保密起见, 六种品牌用 AA、BB、CC、DD、EE 和 FF 来表示。
- **contacts.sav**。该假设数据文件涉及一组公司计算机销售代表的联系方式列表。根据这些销售代表所在的公司部门及其公司的秩来对每个联系方式进行分类。同时还记录了最近一次的销售量、最近一次销售距今的时间和所联系公司的规模。
- **creditpromo.sav**。该假设数据文件涉及某百货公司在评价最新信用卡促销的效果方面的举措。为此, 随机选择了 500 位持卡人。其中一半收到了宣传关于在接下来的三个月内降低消费利率的广告。另一半收到了标准的季节性广告。
- **customer\_dbase.sav**。该假设数据文件涉及某公司在使用数据仓库中的信息来为最有可能回应的客户提供特惠商品方面的举措。随机选择客户群的子集并为其提供特惠商品, 同时记录下他们的回应。
- **customer\_information.sav**。该假设数据文件包含客户邮寄信息, 如姓名和地址。
- **customer\_subset.sav**。来自 customer\_dbase.sav 的拥有 80 个个案的子集。
- **debate.sav**。该假设数据文件涉及在某政治辩论前后对该辩论的参与者所做的调查的成对回答。每个个案对应一个单独的调查对象。
- **debate\_aggregate.sav**。该假设数据文件分类汇总了 debate.sav 中的回答。每个个案对应一个辩论前后的偏好的交叉分类。
- **demo.sav**。这是关于购物客户数据库的假设数据文件, 用于寄出每月的商品。将记录客户对商品是否有回应以及各种人口统计信息。
- **demo\_cs\_1.sav**。该假设数据文件涉及某公司在汇编调查信息数据库方面的举措的第一步。每个个案对应不同的城市, 并记录地区、省、区和城市标识。
- **demo\_cs\_2.sav**。该假设数据文件涉及某公司在汇编调查信息数据库方面的举措的第二步。每个个案对应来自第一步中所选城市的不同的家庭单元, 并记录地区、省、区、市、子区和单元标识。还包括设计前两个阶段的抽样信息。
- **demo\_cs.sav**。该假设数据文件包含用复杂抽样设计收集的调查信息。每个个案对应不同的家庭单元, 并记录各种人口统计和抽样信息。
- **dmdata.sav**。该假设数据文件包含直销公司的人口统计学和购买信息。dmdata2.sav 包含收到测试邮件的联系人子集的信息, dmdata3.sav 包含未收到测试邮件的其余联系人的信息。
- **dietstudy.sav**。该假设数据文件包含对 “Stillman diet” (Rickman, Mitchell, Dingman, 和 Dalen, 1974) 的研究结果。每个个案对应一个单独的主体, 并记录其在实行饮食方案前后的体重 (磅) 以及甘油三酸酯的水平 (毫克/100 毫升)。
- **dvdplayer.sav**。这是关于开发新的 DVD 播放器的假设数据文件。营销团队用原型收集了焦点小组数据。每个个案对应一个单独的被调查用户, 并记录他们的人口统计信息及其对原型问题的回答。
- **german\_credit.sav**。该数据文件取自加州大学欧文分校的 Repository of Machine Learning Databases (Blake 和 Merz, 1998) 中的 “German credit” 数据集。
- **grocery\_1month.sav**。该假设数据文件是在数据文件 grocery\_coupons.sav 的基础上加上了每周购物 “累计”, 所以每个个案对应一个单独的客户。所以, 一些每周更改的变量消失了, 而且现在记录的消费金额是为期四周的研究过程中的消费金额之和。

- **grocery\_coupons.sav**。该假设数据文件包含由重视顾客购物习惯的杂货连锁店收集的调查数据。对每位顾客调查四周，每个个案对应一个单独的顾客周，并记录有关顾客购物地点和方式的信息（包括那一周里顾客在杂货上的消费金额）。
- **guttman.sav**。Bell (Bell, 1961) 创建了一个表，用来阐释可能的社会群体。Guttman (Guttman, 1968) 引用了该表的一部分，其中包括五个变量，用于描述以下七个理论社会群体的社会交往、对群体的归属感、成员的物理亲近度以及关系正式性：观众（比如在足球比赛现场的人们）、听众（比如在剧院或听课堂讲座的人们）、公众（比如报纸或电视观众）、组织群体（与观众类似但具有紧密的关系）、初级群体（关系密切）、次级群体（自发组织）及现代社区（因在物理上亲近而导致关系松散并需要专业化服务）。
- **health\_funding.sav**。该假设数据文件包含关于保健基金（每 100 人的金额）、发病率（每 10,000 人的比率）以及保健提供商拜访率（每 10,000 的比率）的数据。每个个案代表不同的城市。
- **hivassay.sav**。该假设数据文件涉及某药物实验室在开发用于检测 HIV 感染的快速化验方面的举措。化验结果为八个加深的红色阴影，如果有更深的阴影则表示感染的可能性很大。用 2,000 份血液样本来进行实验室试验，其中一半受到 HIV 感染而另一半没有受到感染。
- **hourlywagedata.sav**。该假设数据文件涉及在政府机关和医院工作的具有不同经验水平的护士的时薪。
- **insurance\_claims.sav**。该假设数据文件涉及某保险公司，该公司希望构建一个模型用于标记可疑的、具有潜在欺骗性的理赔。每个个案代表一次单独的理赔。
- **insure.sav**。该假设数据文件涉及某保险公司，该公司正在研究指示客户是否会根据 10 年的人寿保险合同提出理赔的风险因子。数据文件中的每个个案代表一副根据年龄和性别进行匹配的合同，其中一份记录了一次理赔而另一份则没有。
- **judges.sav**。该假设数据文件涉及经过训练的裁判（加上一个体操爱好者）对 300 次体操表演给出的分数。每行代表一次单独的表演；裁判们观看相同的表演。
- **kinship\_dat.sav**。Rosenberg 和 Kim (Rosenberg 和 Kim, 1975) 开始分析 15 个亲属关系项（伯母、兄弟、表兄妹、女儿、父亲、孙女、祖父、祖母、孙子、母亲、侄子或外甥、侄女或外甥女、姐妹、儿子和叔叔）。他们让四组大学生（两组女同学，两组男同学）根据相似程度将各项排序。他们让其中的两组同学（一组女同学，一组男同学）进行了两次排序，第二次排序和第一次排序采取的标准不同。这样，一共得到六组“源”。每个源对应一个  $15 \times 15$  的相似性矩阵，其单元格中的值等于源中的人数减去此源中对象被划分的次数。
- **kinship\_ini.sav**。该数据文件包含 kinship\_dat.sav 的三维解的初始配置。
- **kinship\_var.sav**。该数据文件包含自变量 gender、gener(ation) 和 degree (of separation)，这些变量可用于解释 kinship\_dat.sav 的解的维数。具体而言，它们可用来将解的空间限制为这些变量的线性组合。
- **marketvalues.sav**。该数据文件涉及 1999 - 2000 年间 Algonquin, Ill. 地区新的房屋开发中的住房销售。这些销售仅仅来自公众记录。
- **nhis2000\_subset.sav**。美国健康访问调查 (NHIS) 是针对美国全体公民的大型人口调查。该调查对美国的具有全国代表性的家庭样本进行了面对面的访问，并获取了每个家庭的成员的健康行为和健康状态的人口统计信息和观察数据。该数据文件包含取自 2000 年调查信息的子集。国家健康统计中心。2000 年美国健康访问调查。公用数据文件和文档。

ftp://ftp.cdc.gov/pub/Health\_Statistics/NCHS/Datasets/NHIS/2000/。2003 年发布。

- **ozone.sav**。这些数据包含了用来根据其余变量预测臭氧浓度的六个气象变量的 330 个观察值。在以前的研究人员中, (Breiman 和 Friedman(F), 1985) 和 (Hastie 和 Tibshirani, 1990) 发现了这些变量之间的非线性, 这妨碍了标准回归方法。
- **pain\_medication.sav**。该假设数据文件包含用于治疗慢性关节炎疼痛的抗炎药的临床试验结果。我们感兴趣的是该药见效的时间以及它和现有药物的比较。
- **patient\_los.sav**。该假设数据文件包含被医院确诊为疑似心肌梗塞(即 MI 或“心脏病发作”)的患者的治疗记录。每个个案对应一位单独的患者, 并记录与其住院期有关的一些变量。
- **patlos\_sample.sav**。该假设数据文件包含在治疗心肌梗塞(即 MI 或“心脏病发作”)期间收到溶解血栓剂的患者样本的治疗记录。每个个案对应一位单独的患者, 并记录与其住院期有关的一些变量。
- **poll\_cs.sav**。该假设数据文件涉及民意测验专家在确定正式立法前公众对法案的支持水平方面的举措。个案对应注册的选民。每个个案记录选民居住的县、镇、区。
- **poll\_cs\_sample.sav**。该假设数据文件包含在 poll\_cs.sav 中列出的选民的样本。该样本是根据 poll.csplan 中指定的设计来选取的, 而且该数据文件记录包含概率和样本权重。请注意, 由于该抽样计划使用与大小成正比(PPS)方法, 因此, 还有一个文件(poll\_jointprob.sav)包含联合选择概率。在选取了样本之后, 对应于选民人群统计信息及其对提交法案的意见的附加变量将被收集并添加到数据文件。
- **property\_assess.sav**。该假设数据文件涉及某县资产评估员在利用有限的资源不断更新资产价值评估方面的举措。个案对应过去一年中县里所出售的资产。数据文件中的每个个案记录资产所在的镇、最后评估资产的评估员、该次评估距今的时间、当时的估价以及资产的出售价格。
- **property\_assess\_cs.sav**。该假设数据文件涉及某州资产评估员在利用有限的资源不断更新资产价值评估方面的举措。个案对应该州的资产。数据文件中的每个个案记录资产所在的县、镇和区, 最后一次评估距今的时间以及当时的估价。
- **property\_assess\_cs\_sample.sav**。该假设数据文件包含在 property\_assess\_cs.sav 中列出的资产的样本。该样本是根据 property\_assess.csplan 中指定的设计来选取的, 而且该数据文件记录包含概率和样本权重。在选取了样本之后, 附加变量 Current value 将被收集并添加到数据文件。
- **recidivism.sav**。该假设数据文件涉及某政府执法机构在了解其管辖区域内的屡犯率方面的举措。每个个案对应先前的一名罪犯, 并记录其人口统计信息和第一次犯罪的详细资料; 如果在第一次被捕后两年内又第二次被捕, 则还将记录两次被捕间隔的时间。
- **recidivism\_cs\_sample.sav**。该假设数据文件涉及某政府执法机构在了解其管辖区域内的屡犯率方面的举措。每个个案对应应在 2003 年 6 月期间第一次被捕释放的先前的一名罪犯, 并记录其人口统计信息和第一次犯罪的详细资料, 及其第二次被捕的数据(如果发生在 2006 年 6 月底之前)。根据 recidivism\_cs.csplan 中指定的抽样计划从抽样部门选择罪犯; 该计划使用与大小成正比(PPS)方法, 因此, 还有一个文件(recidivism\_cs\_jointprob.sav)包含联合选择概率。
- **rfm\_transactions.sav**。此假设数据文件包含购买交易数据, 即每笔交易的购买日期、购买商品和消费金额。

- **salesperformance.sav**。这是关于评估两个新的销售培训课程的假设数据文件。60 名员工被分成 3 组且都接受标准的培训。另外，组 2 接受技术培训；组 3 接受实践教程。在培训课程结束时，对每名员工进行测验并记录他们的分数。数据文件中的每个个案代表一名单独的受训者，并记录其被分配到的组以及测验的分数。
- **satisf.sav**。该假设数据文件涉及某零售公司在 4 个商店位置所进行的满意度调查。总共对 582 位客户进行了调查，每个个案代表一位单独客户的回答。
- **screws.sav**。该数据文件包含关于螺钉、螺栓、螺母和图钉的特征的信息 (Hartigan, 1975)。
- **shampoo\_ph.sav**。这是关于某发制品厂的质量控制的假设数据文件。在规定的间隔对六批独立输出的产品进行检测并记录它们的 pH 值。目标范围是 4.5 - 5.5。
- **ships.sav**。在别处被提出和分析的 (McCullagh 等., 1989) 关于波浪对货船造成的损坏的数据集。在给定了船的类型、建造工期和服务期后，可以根据泊松比率发生来为事件计数建模。在因子交叉分类构成的表格中，每个单元格的分类汇总服务月数提供遇到风险的值。
- **site.sav**。该假设数据文件涉及某公司在为扩展业务而选择新址方面的举措。该公司聘请了两名顾问分别对选址进行评估，除了提供长期报告外，他们还要以“前景颇佳”、“前景良好”或“前景不佳”来对每个选址进行总结。
- **smokers.sav**。该数据文件摘自 1998 年全国家庭药物滥用调查并且是美国家庭的概率样本。(<http://dx.doi.org/10.3886/ICPSR02934>) 因此，分析该数据文件的第一步应该是对数据进行加权以反映总体趋势。
- **stocks.sav**。该假设数据文件包含某一年的股票价格和成交量。
- **stroke\_clean.sav**。该假设数据文件包含某医学数据库在经过“数据准备”选项中的过程清理后的状态。
- **stroke\_invalid.sav**。该假设数据文件包含某医学数据库的初始状态及一些数据输入错误。
- **stroke\_survival**。此假设数据文件涉及正在研究结束缺血性中风后复元计划的患者存活时间的研究人员面临着很多挑战。中风后，记录心肌梗塞、缺血性中风或出血性中风的发生及其时间。样本为左侧截短，因为只包含在中风后管理的复元计划结束后存活的患者。
- **stroke\_valid.sav**。该假设数据文件包含在使用“验证数据”过程检查值后，某医学数据库的状态。它仍包含潜在异常个案。
- **survey\_sample.sav**。此数据文件包含调查数据，包括人口统计学数据和各种态度测量。它基于 1998 NORC 综合社会调查的变量子集，但某些数据值已经过修改，并添加了其他虚拟变量以供演示用途。
- **telco.sav**。该假设数据文件涉及某电信公司在减少客户群中的客户流失方面的举措。每个个案对应一个单独的客户，并记录各类人口统计和服务用途信息。
- **telco\_extra.sav**。该数据文件与 telco.sav 数据文件类似，但删除了“tenure”和经对数转换的客户消费变量，代替它们的是标准化的对数转换客户消费变量。
- **telco\_missing.sav**。该数据文件是 telco.sav 数据文件的子集，但某些人口统计数据值已被缺失值替换。

- **testmarket.sav**。该假设数据文件涉及某快餐连锁店为其菜单添加新项目的计划。有三种可能的促销新产品的活动，所以会在多个随机选择的地点引入新的项目。在每个地点采用不同的促销方式，并记录新项目前四周的每周销售情况。每个个案对应单独地点的一周。
- **testmarket\_1month.sav**。该假设数据文件是在数据文件 testmarket.sav 的基础上加上了每周销售“累计”，所以每个个案对应一个单独的地点。所以，一些每周更改的变量消失了，而且现在记录的销售是为期四周的研究过程中的销售之和。
- **tree\_car.sav**。该假设数据文件包含人口统计和车辆购买价格数据。
- **tree\_credit.sav**。该假设数据文件包含人口统计和银行贷款历史数据。
- **tree\_missing\_data.sav** 该假设数据文件包含具有大量缺失值的人口统计和银行贷款历史数据。
- **tree\_score\_car.sav**。该假设数据文件包含人口统计和车辆购买价格数据。
- **tree\_textdata.sav**。这是一个只有两个变量的样本数据文件，主要打算在指定测量级别和值标签之前显示变量的默认状态。
- **tv-survey.sav**。该假设数据文件涉及由某电视演播室进行的一项关于是否要继续制作一档成功的节目的调查。906 位调查对象被问及他们在各种情况下是否会收看该节目。每行代表一位单独的调查对象；每列代表一种单独的情况。
- **ulcer\_recurrence.sav**。此文件包含某项研究的部分信息，该研究旨在比较两种用来防止溃疡复发的治疗的功效。它提供了区间数据的优秀示例并且已在别处被提出和分析 (Collett, 2003)。
- **ulcer\_recurrence\_recoded.sav**。该文件重新组织 ulcer\_recurrence.sav 中的信息以允许为研究的每个区间的事件概率建模而不是简单地研究结束事件概率建模。它已在别处被提出和分析 (Collett 等., 2003)。
- **verd1985.sav**。该数据文件涉及某项调查 (Verdegaal, 1985)。该调查记录了 15 个主体对 8 个变量的响应。需要处理的变量被分成 3 个集。数据集 1 包含 年龄 和 婚姻；数据集 2 包含 宠物 和 新闻；数据集 3 包含 音乐 和 居住。宠物被尺度化为多名义而年龄被尺度化为有序；所有其他变量都被尺度化为单名义。
- **virus.sav**。该假设数据文件涉及某因特网服务提供商 (ISP) 在确定病毒对其网络的影响方面的举措。他们从发现病毒到威胁得以遏制这段时间内跟踪其网络上受感染的电子邮件的流量的 (近似) 百分比。
- **wheeze\_steubenville.sav**。这是关于空气污染对儿童健康影响的纵向研究的一个子集 (Ware, Dockery, Spiro III, Speizer, 和 Ferris Jr., 1984)。这些数据包含儿童的气喘状况的重复二分类测量 (这些儿童来自 Steubenville, Ohio, 年龄为 7 到 10 岁)，以及母亲在研究的第一年中是否为吸烟者的固定记录。
- **workprog.sav**。该假设数据文件涉及一份尝试为弱势群体提供较好的工作的政府工作计划。文件后还有一个潜在计划参与者的样本，其中一些参与者是被随机选择来参加该计划的，而其他参与者则不是。每个个案代表一位单独的计划参与者。
- **worldsales.sav** 该假设数据文件包含按不同大洲和产品列出的销售收入。

# 注意事项

这些信息开发用于在全球提供的产品和服务。

IBM 可能在其他国家/地区中不提供在本文档中讨论的产品、服务或功能。请咨询您当地的 IBM 代表以了解有关您所在地区当前可用产品和服务的信息。任何对 IBM 产品、程序或服务的引用，并不意味着仅可使用这些 IBM 产品、程序或服务。作为替代，可以使用任何功能相当的产品、程序或服务，前提是不侵犯任何 IBM 知识产权。不过，用户应自行负责评估和验证任何非 IBM 产品、程序或服务的工作情况。

在本文档中介绍的主题可能涉及 IBM 的专利或申请中的专利。您获得该文档并不意味着授予您任何这些专利许可。您可以将书面许可查询函件发送至：

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY  
10504-1785, U. S. A.

要查询双字节字符集 (DBCS) 相关许可证信息，请联系所在国家/地区中的 IBM 知识产权部门，或者以书面形式将查询函件发送至：

Intellectual Property Licensing, Legal and Intellectual Property Law, IBM Japan  
Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 242-8502 Japan.

**以下段落不适用于英国或此类条款与当地法律不符的其他国家/地区：** INTERNATIONAL BUSINESS MACHINES 公司“按原样”提供本出版物，不保证任何明示或暗示，包括但不限于对非侵权性、适销性或对特定用途适用性的暗示担保。有些州/省不允许针对某些交易的明示或暗示免责条款，因此本声明可能不适用于您。

本信息可能含有技术误差或排版错误。此处的信息会定期进行更改；这些更改会体现在本出版物的新版本中。IBM 可能随时对本出版物中所述的产品和/或程序进行改进和/或更改，恕不另行通知。

在本信息材料中对任何非 IBM 网站的引用仅为了方便用户，并不以任何方式表明对这些网站的认可。这些网站上的材料并非本 IBM 产品材料的一部分，您对这些网站的使用需自担风险。

IBM 可以自认为适当并且不会对您构成任何约束的任何方式使用或分发您提供的任何信息。

如果本程序的受许可方试图了解有关程序的信息以启用：(i) 在独立创建的程序和其他程序（包括本程序）之间交换信息；(ii) 相互使用交换的信息，则应联系：

IBM Software Group, Attention:Licensing, 233 S. Wacker Dr., Chicago, IL 60606,  
USA.

此类信息的提供应遵照相关条款和条件，其中包括在某些情况下支付适当费用。

在本文档中介绍的受许可保护程序，及其所有受许可保护材料由 IBM 在双方签署的“IBM 客户协议”、“IBM 国际程序许可证协议”或任何其他等同协议下提供。

非 IBM 产品的相关信息来自这些产品的供应商，及其发布的公告或其他公开来源。IBM 并未测试过这些产品，无法确认有关非 IBM 产品的性能准确性、兼容性或任何其他声明。有关非 IBM 产品功能的问题应由这些产品的供应商负责。

本信息包含日常业务运营中使用的数据和报告的示例。为了尽可能详尽地对其进行说明，示例中包含了人员的姓名、公司、品牌和产品的名称。所有这些名称均为虚构，与真实商业企业使用的名称和地址的任何雷同纯属巧合。

如果您正在查阅此信息的软拷贝，照片和彩色插图可能不会显示。

## 商标

IBM、IBM 徽标、ibm.com 和 SPSS 是 IBM Corporation 的商标，在全球许多司法辖区注册。有关最新的 IBM 商标列表，请访问网页 <http://www.ibm.com/legal/copytrade.shtml>。

Adobe、Adobe 徽标、PostScript 和 PostScript 徽标是 Adobe Systems Incorporated 在美国和/或其他国家或地区的注册商标或商标。

Intel、Intel 徽标、Intel Inside、Intel Inside 徽标、Intel Centrino、Intel Centrino 徽标、Celeron、Intel Xeon、Intel SpeedStep、Itanium 和 Pentium 是 Intel Corporation 或其子公司在美国和其他国家或地区的商标或注册商标。

Java 和所有基于 Java 的商标和徽标是 Sun Microsystems, Inc. 在美国、其他国家或地区或两者的商标。

Linux 是 Linus Torvalds 在美国、其他国家或地区或两者的注册商标。

Microsoft、Windows、Windows NT 和 Windows 徽标是 Microsoft Corporation 在美国、其他国家或地区或两者的商标。

UNIX 是 The Open Group 在美国和其他国家或地区的注册商标。

此产品使用 Polar 工程咨询公司的 WinWrap Basic，版权 1993 年–2007 年，<http://www.winwrap.com>。

其他产品和服务名称可能是 IBM 或其他公司的商标。

Adobe 产品屏幕截图重印已获得 Adobe Systems Incorporated 的许可。

Microsoft 产品屏幕截图重印已获得 Microsoft Corporation 的许可。



---

# 参考书目

- Akaah, I. P., 和 P. K. Korgaonkar. 1988. A conjoint investigation of the relative importance of risk relievers in direct marketing. *Journal of Advertising Research*, 28:4, .
- Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.
- Blake, C. L., 和 C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mllearn/MLRepository.html>.
- Breiman, L., 和 J. H. Friedman(F). 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, .
- Cattin, P., 和 D. R. Wittink. 1982. Commercial use of conjoint analysis: A survey. *Journal of Marketing*, 46:3, .
- Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.
- Green, P. E., 和 V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.
- Green, P. E., 和 Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.
- Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, .
- Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.
- Hastie, T., 和 R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.
- Kennedy, R., C. Riquier, 和 B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, .
- McCullagh, P., 和 J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.
- Price, R. H., 和 D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, .
- Rickman, R., N. Mitchell, J. Dingman, 和 J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228, .
- Rosenberg, S., 和 M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, .
- Van der Ham, T., J. J. Meulman, D. C. Van Strien, 和 H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, .
- Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens* (in Dutch). Leiden: Department of Data Theory, University of Leiden.



Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, 和 B. G. Ferris Jr..  
1984. Passive smoking, gas cooking, and respiratory health of children living  
in six cities. *American Review of Respiratory Diseases*, 129, .

# 索引

- BTL (Bradley-Terry-Luce) 模型, 24
- )CARD
  - 在显示设计中, 8
- card\_ 变量
  - 在生成正交设计中, 4
- Kendall 的 tau, 23
- logit 模型, 24
- Pearson 的 R, 23
- status\_ 变量
  - 在生成正交设计中, 4
  
- 保持个案, 2
  - 在生成正交设计中, 6
  
- 列表格式
  - 在显示设计中, 7
  
- 单一概要文件格式
  - 在显示设计中, 7
  
- 反向, 23
- 反理想模型, 20
  
- 命令语法
  - CONJOINT 命令, 20
- 商标, 33
  
- 因子, 1, 14
- 因子级别, 1, 14
  
- 完全概要文件方法, 1
  
- 总效用, 21
  
- 效用得分, 2, 21
- 数据文件
  - 在生成正交设计中, 4
  
- 显示设计, 2, 7, 17
  - )CARD, 8
  - 保存概要文件, 8
  - 列表格式, 7
  - 单一概要文件格式, 7
  - 标题, 8
  - 页脚, 8
  
- 最大效用模型, 24
  
- 标题
  - 在显示设计中, 8
  
- 样本大小, 2
- 样本文件
  - 位置, 25
- 模拟个案, 2, 11, 23
  - 在生成正交设计中, 6
- 模拟结果
  - BTL (Bradley-Terry-Luce) 模型, 24
  - logit 模型, 24
  - 最大效用模型, 24
  
- 正交数组, 2
- 正交设计
  - 保持个案, 6
  - 显示, 7, 17
  - 最小个案, 6
  - 生成, 4, 14
  
- 法律注意事项, 32
  
- 理想模型, 20
  
- 生成正交设计, 2, 4, 14
  - 保持个案, 6
  - 定义因子名称、标签和值, 5
  - 数据文件, 4
  - 最小个案, 6
  - 模拟个案, 6
  - 随机数种子, 4
  
- 相关系数, 23
  
- 离散模型, 20
  
- 系数, 22
  
- 线性模型, 20
  
- 语法
  - CONJOINT 命令, 20
  
- 部分价值, 2
  
- 重要性值, 22
- 重要性得分, 22
  
- 随机数种子
  - 在生成正交设计中, 4
  
- 页脚
  - 在显示设计中, 8