

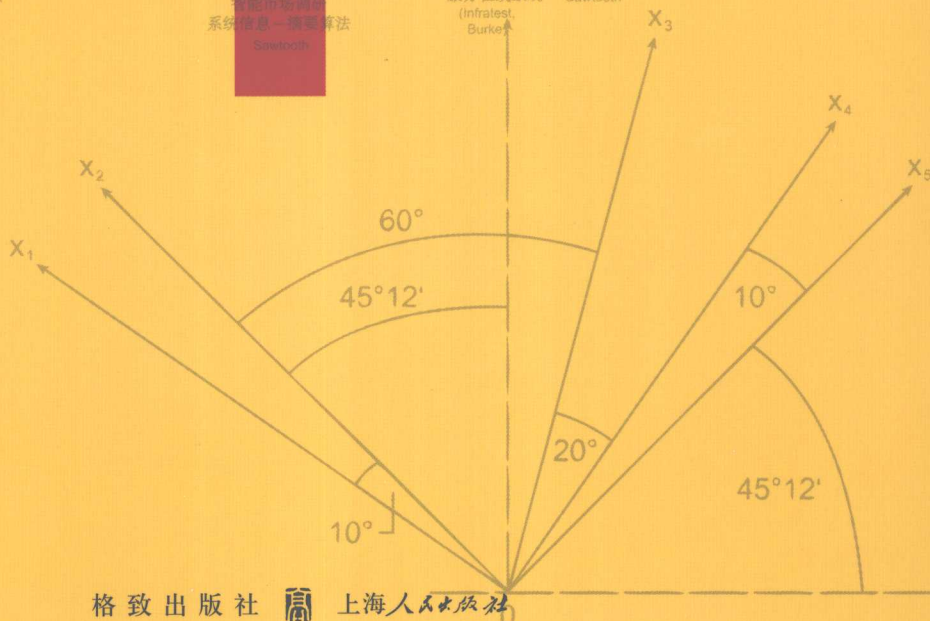
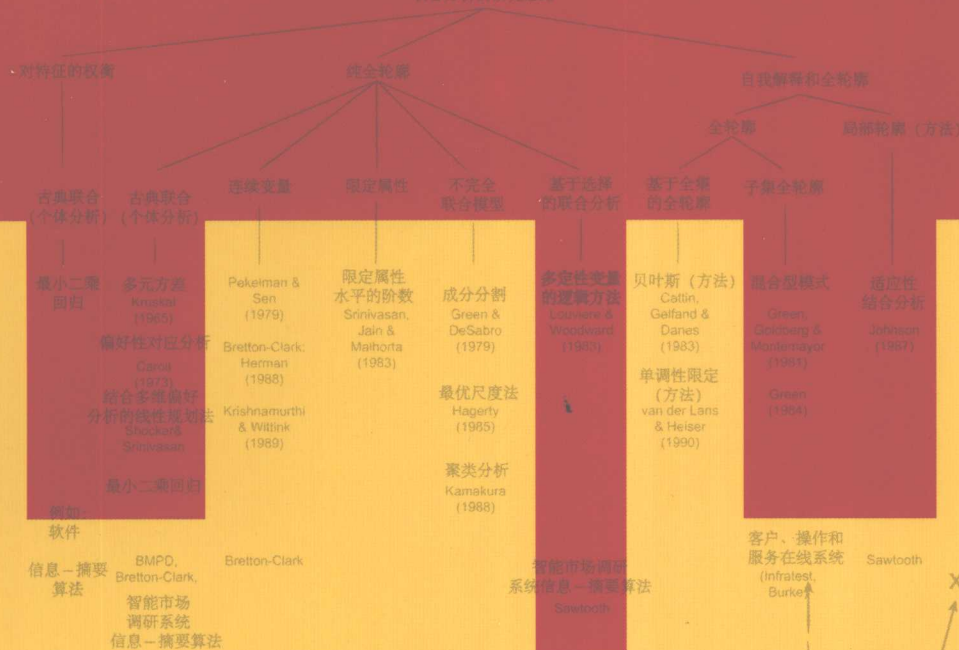
MULTIVARIATE STATISTICAL ANALYSIS

[德] 克劳斯·巴克豪斯 [德] 本德·埃里克森 [德] 伍尔夫·普林克
[中] 王煦逸 [德] 儒尔夫·威伯 —— 著

多元统计分析方法

用SPSS工具

联合分析的研究思路



MULTIVARIATE STATISTICAL ANALYSIS

[德] 克劳斯·巴克豪斯 [德] 本德·埃里克森 [德] 伍尔夫·普林克
[中] 王煦逸 [德] 儒尔夫·威伯 —— 著

多元统计分析方法

用SPSS工具

图书在版编目(CIP)数据

多元统计分析方法:用 SPSS 工具/(德)巴克豪斯
(Backhaus, K.)等著. —上海:格致出版社:上海人民
出版社,2009

书名原文:Multivariate Analysemethoden

ISBN 978-7-5432-1571-9

I. 多… II. 巴… III. 多元分析:统计分析-软件包,
SPSS IV. 0212.4

中国版本图书馆 CIP 数据核字(2009)第 020774 号

责任编辑 麻俊生

美术编辑 路 静

多元统计分析方法

——用 SPSS 工具

[德]克劳斯·巴克豪斯 [德]本德·埃里克森

[德]伍尔夫·普林克 [中]王煦逸 [德]儒尔夫·威伯 著

出版 世纪出版集团 格致出版社
www.ewen.cc www.hibooks.cn
上海人民出版社
(200001 上海福建中路193号24层)



编辑部热线 021-63914988

市场部热线 021-63914081

发行 世纪出版集团发行中心
印刷 上海图宇印刷有限公司
开本 787×1092 毫米 1/16
印张 25.75
插页 1
字数 579,000
版次 2009 年 4 月第 1 版
印次 2009 年 4 月第 1 次印刷
ISBN 978-7-5432-1571-9/F·152
定价 50.00 元

With financial support of the Institute for Business
to Business Marketing (IAS) and the Förderkreis für
Industriegütermarketing e. V. (FKI) at Westphalian
Wilhelms-University Muenster (WWU), Thyssen-
Krupp AG and CDHK Tongji University Shanghai

本书的出版得到
明斯特大学 B2B 营销研究所、投资品市场学基金会、
蒂森克虏伯股份公司和
同济大学中德学院
的资助

中文版前言

本书《多元统计分析方法》是德国该领域的经典权威著作,它受到各个运用领域师生与读者的推崇,特别在经济管理、工程学、自然科学、医学、心理学和社会学等领域得到广泛的运用,因此我们决定在其他语言市场中推广本书,这是我们从事《多元统计分析方法》(中文版)出版工作的背景。在本书中文版的出版工作中,同济大学中德学院蒂森克虏伯管理控制讲座教授王煦逸博士加入了我们的作者团队,从事了大量的修订工作,为此本书德方作者表示衷心的感谢。德方作者团队由于语言关系无法直接从事中文版的修订工作,我们很高兴王教授作为该领域资深专家能够与我们合作完成本项工作。

和德语版一样,本书的中文版涉及了当前多元统计分析的所有基本方法,在修订过程中,我们保持了本书德语版的特色:在涉及最低限度数学知识的前提下,我们运用通俗易懂的语言,并结合一个具体案例来介绍所有的多元统计方法。同时本书详细解释了统计软件SPSS 15.0中文版的相应运用步骤,使读者能够更进一步地深入理解有关内容,从而使本书具有很强的运用可操作性。本书涉及以下内容:

1. 回归分析
2. 时间序列分析
3. 方差分析
4. 判别分析
5. 逻辑回归
6. 联列表和相合性分析
7. 因子分析
8. 聚类分析
9. 联合分析

通晓德语的读者还可以通过我们的官方网站 www.multivariate.de 得到下列更进一步高等多元统计分析方法的德语阐释:

10. 非线性回归
11. 结构方程模型
12. 确认型因子分析
13. 神经网络
14. 多维排序

15. 高维联列表

16. 基于选择的联合分析

通过互联网不断更新书籍内容的新观念在德语区已经被广泛接受,我们期望这个观念也会得到中国读者的响应,我们将在官网上开辟中文读者服务区,为读者提供优质的服务。

本书的出版工作得到了我们学术助理的大力协助,在此我们对于来自各个大学的他们表示衷心的感谢,特别是明斯特大学的 Harald Neun 先生为本书中文版的出版担任了德方的协调工作,同时还要感谢来自同济大学中德学院在明斯特大学的交流学生史雯婷和杨新罡同学,她们为本书的出版投入了大量的精力。

我们还要感谢给予中文版出版工作大力支持的同济大学中德学院企业管理系褚慧、李曼和数学系李玉瑜同学,她们的努力使本书的中文版更加完善。最后还要感谢上海世纪出版集团格致出版社麻俊生先生,他为本书的顺利出版作出了重要的贡献。

对于虽然这里没有提到名字但对于本书的出版作出贡献的各界人士,我们在此表示衷心的感谢,当然我们对于本书内容的正确性负全部责任,书中有不妥之处,欢迎广大读者批评指正。

克劳斯·巴克豪斯

本德·埃里克森

伍尔夫·普林克

王煦逸

儒尔夫·威伯

2008 年秋于明斯特、马格德堡、柏林、上海和特里尔

Vorwort

zur chinesischen Auflage

Das Buch „Multivariate Analysemethoden“ ist in Deutschland ein Klassiker. Es ist disziplinübergreifend äußerst beliebt bei Studierenden und Dozenten verschiedenster Fachrichtungen, insbesondere bei Wirtschaftswissenschaftlern, Ingenieuren, Naturwissenschaftlern, Mediziner, Psychologen und Soziologen. Vor diesem Hintergrund haben wir uns entschieden, das Buch auch für andere Märkte und in anderen Sprachen zugänglich zu machen. Hiermit legen wir die chinesische Version des Buches „Multivariate Analysemethoden“ vor. Das Autorenteam der chinesischen Auflage ist durch unseren Kollegen Prof. Dr. Wang Xuyi, Lehrstuhlinhaber für ThyssenKrupp-Lehrstuhl für Controlling CDHK Tongji Universität, ergänzt worden, dem die deutschen Autoren für sein großes Engagement bezüglich der chinesischen Version ganz außerordentlich Dank sagen. Da die deutschen Autoren selbst nicht in der Lage sind, die chinesische Auflage zu beurteilen, freuen wir uns umso mehr, mit Herrn Kollegen Wang einen kompetenten Partner für die chinesische Ausgabe gefunden zu haben.

Wie die deutsche Ausgabe auch, umfasst die chinesische Version die Verfahren der multivariaten Analysemethoden, die heute den Standard der multivariaten Analyse bilden. Dabei sind wir dem Motto der deutschen Ausgabe auch bei der chinesischen Version treu geblieben: Wir versuchen so weit wie möglich bei geringst möglichen Anforderungen an mathematische Vorkenntnisse eine allgemeinverständliche Darstellung der Verfahren anhand eines für alle Methoden verwendeten Fallbeispiels. Die konsequente Anwendungsorientierung dokumentiert sich auch darin, dass durch die Erläuterung der Fallbeispiele unter Verwendung von SPSS 15.0 für Windows chinesische Version eine vollständige Nachvollziehbarkeit der Verfahren durch den Leser gewährleistet ist. Im Einzelnen behandeln wir in diesem Buch die folgenden „grundlegenden Verfahren der multivariaten Analyse“:

1. Regressionsanalyse
2. Zeitreihenanalyse
3. Varianzanalyse
4. Diskriminanzanalyse
5. Logistische Regression
6. Kreuztabellierung und Kontingenzanalyse
7. Faktorenanalyse

8. Clusteranalyse
9. Conjoint-Analyse

Für den Leser, der Deutsch versteht, bieten wir darüber hinaus für die nachstehend genannten „komplexen Verfahren der multivariaten Analyse“ ein deutschsprachiges Manuskript unter www.multivariate.de an:

10. Nichtlineare Regression
11. Strukturgleichungsmodelle
12. Konfirmatorische Faktorenanalyse
13. Neuronale Netze
14. Multidimensionale Skalierung
15. Korrespondenzanalyse
16. Auswahlbasierte Conjoint-Analyse

Wir hoffen sehr, dass das im deutschsprachigen Bereich sehr gut angenommene Konzept auch in China Akzeptanz findet. Wir werden ein Link für chinesischen Leser einrichten, damit sie gute Unterstützung von uns bekommen können.

Dieses Buch hätte nicht vorgelegt werden können, ohne dass eine Reihe von Mitarbeiterinnen und Mitarbeitern sich über das normale Maß hinaus für dieses Buch engagiert hätten. Dank für vielfältige und umfassende Hilfe gebührt den Mitarbeitern an verschiedenen Standorten. In Münster hat sich insbesondere Herr Dipl.-Ing. Harald Neun um die Koordination der Erstellung der chinesischen Auflage gekümmert. Aus Shanghai haben sich darüber hinaus in Münster die Gaststudentinnen Shi Wenting und Yang Xingang unermüdlich in das Projekt eingebracht.

Unser Dank gilt auch für die Studentinnen von Cdhk, die die Veröffentlichung der chinesischen Version unterstützt haben, sie sind Frau Chu Hui, Frau Li Man, Frau Li Yuyu. Durch ihr Engagement hat sie dazu beigetragen, das Buch in kurzer Zeit auf dem chinesischen Markt gebracht werden kann. Wir danken auch Herr Ma Junsheng, Mitarbeiter von Truth & Wisdom Press, der die Redakteurarbeit dieses Buches übernommen hat. Ohne sein Engagement kann das Buch nicht auf dem Markt kommen.

Wir bedanken uns bei allen namentlich und nicht namentlich genannten Helfern. Selbstverständlich gehen alle eventuellen Mängel zu unseren Lasten.

Münster, Magdeburg, Berlin, Shanghai, Trier im Herbst 2008

Klaus Backhaus
Bernd Erichson
Wulff Plinke
Xuyi Wang
Rolf Weiber

目 录

导论	1
0.1 撰写本书的目的	1
0.2 数据和测度	3
0.3 多元分析方法的分类	5
0.3.1 结构检验的方法	6
0.3.2 结构发现的方法	9
0.3.3 小结	11
0.4 SPSS 的使用	12
0.4.1 数据	12
0.4.2 简单的统计量和图形	20
0.4.3 命令语言	25
0.4.4 SPSS 系统	32
参考文献	33
第 1 章 回归分析	35
1.1 问题的提出	35
1.2 分析过程	38
1.2.1 建模	38
1.2.2 估计回归函数	39
1.2.3 检验回归函数	46
1.2.4 检验回归系数	52
1.2.5 检验模型前提	55
1.3 案例	64
1.3.1 组回归分析	64
1.3.2 逐步回归分析	72
1.3.3 SPSS 命令	76

1.4 应用建议	77
1.5 数学附录	77
参考文献	79

第2章 时间序列分析 81

2.1 问题的提出	81
2.2 分析过程	84
2.2.1 绘制时间序列	84
2.2.2 建模	85
2.2.3 估计模型	85
2.2.4 进行预测	86
2.2.5 检验预测有效性	89
2.2.6 非线性趋势模型	90
2.2.7 考虑结构性变化	95
2.2.8 考虑周期性波动	96
2.3 使用 SPSS	97
2.3.1 线性回归	97
2.3.2 曲线拟合	98
2.4 案例:人造黄油市场的时间序列分析	100
2.4.1 外推模型	100
2.4.2 结构模型	102
2.4.3 总结	105
2.5 应用建议	106
2.6 数学附录	107
参考文献	107

第3章 方差分析 109

3.1 问题的提出	109
3.2 分析过程	110
3.2.1 单因素方差分析	110
3.2.2 双因素方差分析	115
3.2.3 方差分析推广精选	122
3.3 案例	124

3.3.1 问题的提出	124
3.3.2 结果	125
3.3.3 SPSS 命令	127
3.4 应用建议	127
参考文献	129

第4章 判别分析 131

4.1 问题的提出	131
4.2 分析过程	132
4.2.1 定义组	133
4.2.2 建立判别函数	133
4.2.3 估计判别函数	135
4.2.4 检验判别函数	145
4.2.5 检验判别变量	149
4.2.6 将新元素分类	151
4.3 案例	157
4.3.1 问题的提出	157
4.3.2 结果	158
4.3.3 逐步判别分析	168
4.3.4 SPSS 命令	169
4.4 应用建议	171
4.5 数学附录	171
参考文献	176

第5章 逻辑回归 178

5.1 问题的提出	178
5.1.1 逻辑回归的基本思想	178
5.1.2 建立逻辑回归方程	179
5.2 分析过程	182
5.2.1 建模	182
5.2.2 估计逻辑回归函数	184
5.2.3 解释回归系数	186
5.2.4 检验模型整体	189

5.2.5 检验特征变量	198
5.3 案例	199
5.3.1 问题的提出	199
5.3.2 结果	200
5.3.3 SPSS 命令	210
5.4 应用建议	211
5.5 数学附录	212
参考文献	215

第6章 联列表和相合性分析 217

6.1 问题的提出	217
6.2 分析过程	219
6.2.1 编制联列表	220
6.2.2 解释结果	221
6.2.3 检验相关关系	223
6.3 案例	228
6.3.1 问题的提出	228
6.3.2 结果	230
6.3.3 SPSS 命令	232
6.4 应用建议	233
参考文献	234

第7章 因子分析 235

7.1 问题的提出	235
7.2 分析过程	240
7.2.1 变量选择和相关矩阵的计算	240
7.2.2 因子的提取	246
7.2.3 累积方差贡献率的确定	254
7.2.4 因子的数量	258
7.2.5 因子阐释	260
7.2.6 因子得分的确定	262
7.2.7 因子分析的小结	264
7.3 案例	266

7.3.1 问题的提出	266
7.3.2 结果	267
7.3.3 SPSS 命令	276
7.4 应用建议	277
7.4.1 使用因子分析时的问题	277
7.4.2 进行因子分析的建议	281
7.5 数学附录	282
参考文献	284

第8章 聚类分析

285

8.1 问题的提出	285
8.2 分析过程	286
8.2.1 相似性的确定	286
8.2.2 合并算法的选择	298
8.2.3 类数的确定	313
8.3 案例	317
8.3.1 问题的提出	317
8.3.2 结果	318
8.3.3 SPSS 命令	325
8.4 应用建议	326
8.4.1 进行聚类分析之前的考虑	326
8.4.2 聚类分析使用说明	328
参考文献	330

第9章 联合分析

331

9.1 问题的提出	331
9.2 分析过程	334
9.2.1 属性和属性水平	334
9.2.2 受测设计	335
9.2.3 受测体的评价	339
9.2.4 效用值的估计	340
9.2.5 效用值的聚集	345
9.3 案例	348

9.3.1 问题的提出	348
9.3.2 分析结果	351
9.3.3 SPSS 命令	362
9.4 运用建议	369
9.4.1 经典的联合分析的实施	369
9.4.2 选择的联合分析方法的应用	369
9.5 数学附录	372
参考文献	373
附录 1 t -分布表	375
附录 2 F -分布表	377
附录 3 按照 Cochran 的 c -分布表	387
附录 4 χ^2 -分布表	389
附录 5 χ^2 -分布的百分率值—自由度为 ν 的分布	391
附录 6 Durbin-Watson-检验表(置信度 0.95)	393
附录 7 q -值表	395

导 论

0.1 撰写本书的目的

多元统计分析方法作为实践科学中经验研究的基础之一,如今得到越来越迅速的发展,在研究方法上也不断的推陈出新,这些方法涉及众多新的应用领域。现在又开发出了新的或更好的电脑程序,它们为统计方法的实际应用提供了保障。但是在使用这些方法时,一些使用者可能会感到有些障碍,这主要是由以下原因造成的:

- (1) 对数学描述的不理解;
- (2) 对电脑程序运用有一定的畏惧;
- (3) 对方法及其运用的可能性缺乏认识和理解。

这造成了在专业上需要运用统计方法的人士与统计方法专家之间的鸿沟,目前还没有足够的专业文献能够消除这些障碍。

本书的作者以克服这一障碍为目的,也因此形成了本书如下的特点:

(1) 在描述方法时,特别注意方法的可理解性。在所有章节中,首先考虑到的是,没有数学背景读者对其的理解,而不是首先进行详细的数学说明。当然,对方法数学内容的基本特点也有阐述,以便让那些在此方法上有过研究的人能够对该方法的基本功能、运用可能性及局限性有更深入的理解。

(2) 通过详细的例子来说明方法,使读者更容易理解。通过这些例子,读者更容易了解和领会方法的使用过程。

(3) 另外,在方法的描述上,作者把一个例子运用到多个方法中,这样可以更简单地处理问题,同时方法的结果也具有可比较性。例子的原始数据可以通过网址 www.multivariate.de 得到。

这些例子来自于市场营销领域,对其的描述可以保证让每个读者理解所提出的问题,并可以将这些专业的应用问题引申到别的领域。

(4) 加工数据资料范围的确定考虑了各个方法的程序计算花费,使得支出在可以接受的范围内。这些程序必须在计算机支持下才能运行,因此方法描述,既包括方法的基本理念,同时也涉及作为辅助工具的计算机程序的使用。目前存在一系列的程序软件包,这些软件包不仅允许计算机专家使用多元统计分析的方法,而且随着个人电脑不断普及和功能的不断完善,以及使用界面构造的简单明了,使程序的使用也变得越来越简单,因此对于实际问题的研究,有关专家可以不依赖于电脑专家而运用统计软件包。

本书中的大多数例子通过程序软件包或程序系统来计算。这个程序软件包就是 SPSS

(Statistical Package for the Social Sciences),程序系统将应用程序集合置于通用使用界面下。SPSS 有很广泛的使用领域,特别是在高等教育领域,同时也广泛运用于生产实践中。它可支持许多企业系统的大型计算机、工作站和个人电脑。

(5) 本书具有“工具书”的特点。该书所选择的叙述方法使读者在任何情况下都可以理解方法运用的步骤。本书给出了事例所依据的所有原始数据,都可以在网址 [www. multi-variate. de](http://www.multivariate.de) 找到,并且可以通过该网址订购大量的案例。对电脑程序的语法命令有单独的叙述,读者可以亲自操作,并很快会了解到,在使用计算机的情况下,这些方法的运用很简单,同时读者也可以将自己得到的结果与该书中所给出的结果进行比较。

(6) 在每种方法中,计算机所计算出来的结果将分别根据相应的程序表达来解释和说明。这样新使用这些方法的读者可以很快在自己的结果中找到阐释方向。

(7) 本书尤其重视对每种方法结果实质内容的阐释。在方法中,阐明了结果出发点的主观操纵可能性,并指出了操纵空间。这样,方法使用者可以在详细描述方法必要细节的基础上,区分结果的主观和客观的决定因素,这清楚地指出,方法的使用者在对其结果进行阐释时所要担负的责任。

总结一下本书所叙述的特点,可以看出,本书既适合于那些新使用这些方法的初学者,也适合那些必须研究这些方法结果的人,可以给予他们必要的帮助。此书中介绍的每种方法都简单易懂,读者在阅读时并不受每章顺序的限制。

对于阅读此书中文版的读者,在此要明确指出:我们在本书例子中所涉及的产品是在德国市场上的典型产品。主要是关于三种黄油品牌,即:Delikado、holländische Markenbutter 和 Weihnachtsbutter,以及八种人造黄油品牌:Becel、Du darfst、Rama、Homa、Flora Soft、SB、Sanella 和 Botteram。对黄油品牌和人造黄油品牌在化学上可以如下区分:黄油是以动物脂肪(牛奶)为主,而人造黄油品牌是以植物脂肪为主。在人造黄油品牌中,品牌 Flora Soft、Becel 和 Du darfst 与其余品牌的区别在于:这些人造黄油品牌有减肥目的,因此具有较低的卡路里含量,如表 1 所示。

表 1 黄油品牌特色

黄油品牌		
品 牌	描 写	购买行为
Delicado	一种奶油状的黄油,以成本价格出售。	顾客在购买该产品时不一定重视产品重量。
Holländische Markenbutter	一种中等质量的黄油,其价格相对较便宜。	每天都需要食用的产品。
Weihnachtsbutter	一般在圣诞节时以冷冻状态上市。	仅仅在圣诞节时食用。
人造黄油品牌		
品 牌	描 写	购买行为
Becel	一种卡路里含量较低的人造黄油,不适合烤和煎。	一般作为健康食品而被购买。

(续表)

人造黄油品牌		
品 牌	描 写	购买行为
Du darfst	Becel 品牌的直接竞争产品。	一般作为健康食品而被购买。
Rama	是“普通”人造黄油品牌中的市场领先者。	既可用于煎和烤,也可以用于涂抹面包。
Homa	Rama 品牌的直接竞争产品。	既可用于煎和烤,也可以用于涂抹面包。
Flora Soft	具有可涂抹性的一种人造黄油。	作为面包涂抹的一种产品,可用于烹饪、烤和煎。
SB	一种向日葵味的人造黄油,中等价位。	对向日葵味的偏好。
Sanella	一种仅仅用于烧烤而使用的人造黄油。	一般用于准备烧烤食物而购买。
Botteram	一种传统的和最早上市的人造黄油。	既可用于烤和煎,又可以用于涂抹面包。

下面将对多元统计分析方法进行简单的概述。因为在后面需要按照每种方法对数据资料的要求进行分类,因此在这里先提前给出一些注意事项。由于考虑到有部分读者是初学者,因此我们有必要对此作出简要说明。^①

0.2 数据和测度

用于多元分析的“原始资料”是(已存在的或还需调查得到的)数据。数据的质量由测量的种类和方式决定,测量的结果即为数据。测量意味着:根据特定的规则将对象的属性用数字表达出来。

本质上,用于研究属性的种类决定了人们能在多高的程度上测量属性水平,即:人们在多高的程度上用数字表达这些水平。例如,一个人的体形大小可以很容易用数字来表达,但是他的智力、动机或健康状况就很难用数字来表达了。

每个属性的水平来自于“标杆”,这个标杆称为测度。根据对象属性用数字表达的种类和方式不同,可以将测度尺度分为以下四种测度水准:

- (1) 名义测度;
- (2) 序数测度;
- (3) 区间测度;
- (4) 比率测度。

测度水准既取决于数据的信息含量,又取决于计算方法的可用性。以下简单说明一下

^① 相应的统计基础参见:例如:Bleymüller/Gehlert/Gülicher(2004)或 Fahrmeir/Künstler/Pigeot/Tutz(2007)。