

因子分析方法在营销调研中的应用

陈寿雨

(浙江越秀外国语学院 浙江 绍兴 321000)

【摘要】因子分析是一种常用的多元统计分析方法。本文以因子分析方法在营销调研中的应用为例,介绍应用因子分析方法的七个步骤:确定因子分析的目的、构造变量之间的相关系数矩阵、确定因子分析的方法、旋转因子、解释因子、计算因子得分和判断模型拟合度等,并举例说明因子分析在市场细分、产品、价格、渠道、促销等营销调研方面若干应用场景。

【关键词】因子分析;多元统计;营销调研

The application of factor analysis method in marketing research

Shouyu Chen

(Zhejiang Yuexiu University of Foreign Languages, Shaoxing, Zhejiang, 321000)

【Abstract】Factor analysis is a commonly used multivariate statistical analysis method. Taking the application of factor analysis method in marketing research as an example, this paper introduces seven steps of applying factor analysis method: Determine the purpose of factor analysis, construct the correlation coefficient matrix between variables, determine the method of factor analysis, spin factor, explanatory factor, calculate factor score and determine the fit degree of the model, etc., and illustrate some application scenarios of factor analysis in marketing research such as market segmentation, product, price, channel and promotion.

【Key words】Factor analysis; Multivariate statistics; Marketing research

因子分析是一种应用广泛的多元统计分析方法。本文通过一个具体的营销调研例子,介绍因子分析应用的流程。应用因子分析方法不需要很高深的数学基础,因为涉及到模型计算方面的部分都可以由软件来完成。从应用的角度来说,使用者只需理解因子分析的基本原理、实现步骤及其需要注意的问题,并能对软件输出结果进行正确的解释。

1 因子分析的含义与类型

因子分析是指一类多元统计分析法,这类方法的主要目标是用少数几个潜在或设想的变量(潜变量)代替多个观测变量。如一项围绕学生对就读学校的满意度研究,问卷内容一般会涉及学生对学校的教职工、硬件设施、校园环境、管理水平等等方面的评价,收集大量与学校满意度相关的一系列变量的数据。在这些变量中,一些变量具有较高的相关性,如与硬件设施相关的变量。通过观察变量之间的相关系数矩阵,可能就会发现其中一些变量之间的相关系数较大,而与其他变量之间的相关系数较小。因子分析方法可以找出用于解释这些高度相关变量的少数几个因子。例如,学生对有关学校硬件设施一系列问题的回答很可能受学校对硬件设施投入状况的影响,而对学校教职员相关问题的回答则受师资教学水平和职工服务意识的影响,因此可以把“硬件设施”和“教职员工”

作为两个(潜在的)因子分别来代替两组相关变量用于解释学生对学校满意度的高低。

因子分析可分为探索性因子分析(EFA)和验证性因子分析(CFA)。探索性因子分析用于研究者事先不知道在数据背后存在几个什么样的因子,该类因子分析用于探索用最小数量的因子来解释观测变量之间的协方差。验证性因子分析基于现有的理论和经验已提出了相关的因子,该类因子分析将用样本数据来验证所提出的这些因子的假设是否得到支持。EFA和CFA也不是完全分离的。研究者可以把样本数据分成两半,一半用于探索性因子分析来找出潜在的因子,再用另一半样本数据进行验证性因子分析,用于检验这些因子是否得到支持。

2 进行因子分析的步骤

第一步:确定因子分析的目的。进行因子分析首先要明确因子分析的目的。例如,营销调研人员为了了解顾客对洗发水的偏好,围绕顾客对洗发水在去屑、头发柔顺等方面提出6个问题(表1),并试图通过因子分析的方法缩减变量数,归纳顾客购买洗发水的利益诉求。

第二步:构造变量之间的相关系数矩阵。因子分析不是根据原始的变量数据直接进行分析的,它需要先把数据转为变量之间的相关系数或协方差形式(表2)。

表1 顾客对洗发水偏好的调查

Q1	洗发水去屑效果很重要
Q2	洗发水使头发更柔顺很重要
Q3	洗发水使头发有光泽很重要
Q4	洗发水去油功能很重要
Q5	洗发水洁净功能很重要
Q6	洗发水气味很重要

表2 变量的相关系数矩阵

相关系数	去屑	柔顺	光泽	去油	洁净	气味
去屑	1					
柔顺	-0.163	1				
光泽	0.002	0.474	1			
去油	0.845	-0.218	-0.149	1		
洁净	0.840	-0.118	0.125	0.729	1	
气味	0.017	0.521	0.646	-0.017	0.243	1

通过观察相关系数,可以直观的判断数据是否合适用来进行因子分析。从表2中可以看出,一些变量之间存在较强的线性相关关系,说明数据适合进行因子分析。另外,因子分析对样本容量的大小有一定的要求,通常样本容量要达到变量个数的4-5倍。如果在实际调研过程中发现样本容量不足的情况,则对因子分析的结果要十分谨慎对待。同时也有正规的方法来检验数据是否适合进行因子分析。KMO和巴特利特检验就是正规的检验方法。KMO用来检验因子分析的样本充足度,KMO值越大,说明样本数据越适合因子分析,一般要求其值在0.5以上,本例KMO值为0.639,符合要求。巴特利特检验是根据相关系数矩阵计算得出,其值越大就越有证据拒绝变量之间不相关的假设,从而适合进行因子分析,本例中巴特利特球形检验的卡方统计量为99.484,p值近似

为0,从而拒绝变量之间不相关的假设,可以得出数据适用进行因子分析的结论。

第三步:确定因子分析的方法。因子分析有多种实现方法,提取主成分是其中较为常用的一种方法。主成分分析基于变量的协方差矩阵(数据标准化之后,等价于相关系数矩阵),能够用最小数量的主成分(公因子)来最大化的解释变量之间的协方差矩阵。在本例中,总共提取2个公因子,累积解释全部方差的79.844%(表3)。

第四步:旋转因子。因子分析产生因子矩阵,又叫因子模式矩阵。因子矩阵的系数表明标准化之后的变量如何用因子来表示。这些系数叫因子载荷(表4中“因子载荷”栏),等于因子和变量之间的简单相关系数。若某个变量与一个因子之间的因子载荷系数绝对值越大,说明两者的线性相关性强。虽然初始的因子载荷能够反映变量与因子之间的关系,但有时很难用因子来解释这些变量,因为因子很可能与许多变量相关。通过因子旋转,能够使因子载荷矩阵更容易解释。因子旋转的方法可以分为两大类,分别是正交旋转和斜交旋转。其中常用的是方差极大法,是一种正交旋转方法,旋转后使各因子互相独立,并使得用因子来解释变量变得容易,也更方便对因子命名。斜交旋转时轴不再保持垂直,旋转后的因子是相关的。表4中的“旋转后的因子载荷”栏是用方差极大法进行正交旋转之后得到的因子载荷。

第五步:解释因子。通过解读旋转后的因子载荷矩阵,发现因子1与洗发水的去屑、去油、洁净等特征有关,可以把因子1称为功能因素;因子2与洗发水的柔顺、光泽、气味等特征有关,可以把该因子称为美观因素。因子载荷图可以直观显示两个因子与变量之间的关系(图1)。变量洁净、去屑、去油位于因子1(图中为因子F2)的末端;变量柔顺、光泽、气味位于因子2(图

表3 用主成分分析法提取的公因子

成分	初始特征值			提取载荷平方和			旋转载荷平方和		
	总计	方差(%)	累积(%)	总计	方差(%)	累积(%)	总计	方差(%)	累积(%)
1	2.667	44.448	44.448	2.667	44.448	44.448	2.655	44.256	44.256
2	2.124	35.396	79.844	2.124	35.396	79.844	2.135	35.589	79.844
3	0.542	9.030	88.874						
4	0.360	6.004	94.879						
5	0.212	3.534	98.412						
6	0.095	1.588	100.000						

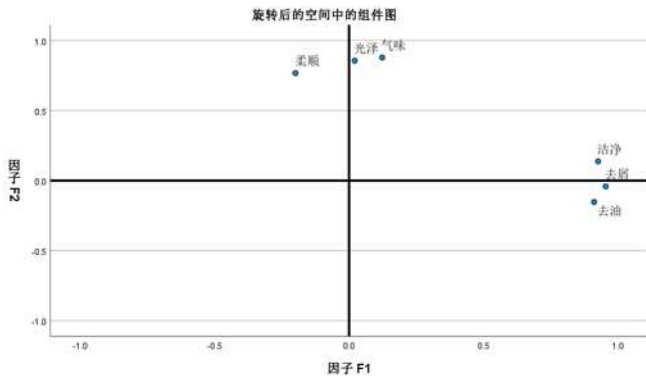


图1 因子载荷图

表4 相关因子系数矩阵

变量	因子载荷		旋转后的因子载荷		因子得分系数	
	因子1	因子2	因子1	因子2	因子1	因子2
去屑	0.950	0.098	0.955	-0.041	0.359	-0.006
柔顺	-0.309	0.730	-0.200	0.767	-0.065	0.357
光泽	-0.104	0.850	0.021	0.856	0.020	0.401
去油	0.924	-0.018	0.911	-0.152	0.341	-0.059
洁净	0.896	0.271	0.926	0.138	0.351	0.077
气味	-0.006	0.887	0.123	0.878	0.059	0.413

中为因子F1)的末端。由此可以发现顾客购买洗发水有两大利益诉求：功能利益和美观利益。

第六步：计算因子得分。因子分析提取公因子（即因子1、因子2）后，可以计算每个公因子的得分，用于后续的统计分析，如可以把公因子作为解释变量放入到多元回归模型中，可以降低多元回归模型的多重共线性问题。因子得分就是原始变量的线性组合，其中权重为因子得分系数矩阵（表4中“因子得分系数”栏）

第七步：判断模型拟合度。因子分析的效果要从模型的拟合度大小来判断。拟合度是基于原始变量相关系数矩阵（实测相关性，即表2）与由因子载荷推导出来的原始变量相关系数矩阵（再生相关系数）进行比较，求两者之差，即残差。如果残差越小，说明因子分析模型的拟合度越高；反之拟合度就越低，需要重新检查模型。在本例中，只有7个残差绝对值大于0.05，表明因子分析模型拟合良好（表5）。

3 因子分析在营销调研中的应用场景

因子分析可以用于概括潜在顾客的不同特征，从

表5 实测相关性与再生相关性的比较

再生相关系数	去屑	柔顺	光泽	去油	洁净	气味
去屑	0.913					
柔顺	-0.222	0.628				
光泽	-0.016	0.652	0.733			
去油	0.876	-0.299	-0.112	0.854		
洁净	0.879	-0.079	0.137	0.823	0.877	
气味	0.081	0.649	0.754	-0.022	0.235	0.786
残差	去屑	柔顺	光泽	去油	洁净	气味
去屑						
柔顺	0.059					
光泽	0.018	-0.179				
去油	-0.031	0.081	-0.037			
洁净	-0.038	-0.039	-0.012	-0.094		
气味	-0.064	-0.128	-0.108	0.004	0.008	

而对顾客进行细分，方便企业选择目标市场。例如围绕顾客对手机的偏好调查之后，可以根据顾客对手机的品牌、系统、配置、价位和特色功能的重视程度，将他们划分为不同类型的顾客群体。

在产品调研中，因子分析可以用来概括和梳理顾客对各种产品属性的偏好，从而可以更有针对性的开发新产品。在价格调研中，可以发现不同价格敏感性顾客的特征差异，从而可以更深入的了解不同类型顾客的特征。在渠道调研中，通过因子分析可以发现顾客购买产品的主要渠道，从而使销售渠道布局更有针对性。在促销调研中，了解目标顾客的媒体使用偏好，从而找出能够更易接触目标顾客的广告宣传媒介。

参考文献：

- [1] 纳雷希·马尔霍特拉. 营销调研：应用导向[M]. 中国人民大学出版社，2020.
- [2] 金在温，查尔斯·W·米勒. 因子分析——统计方法与应用问题[M]. 格致出版社，2016.

作者简介：

陈寿雨（1977.05- ），男，汉族，浙江苍南人，副教授，研究方向：统计应用、市场调查。