



对应分析在品牌定位研究中的应用解析

内容提要: 对应分析是研究定性变量间相互关系并将之以坐标图的方式进行体现的有效方法, 本文结合具体案例系统介绍了该方法的原理、方法及其在市场研究中的应用。

关键词: 对应分析 惯量 品牌定位

前言

对应分析(Correspondence Analysis)是研究定性变量(Nominal Variable)间相互关系的有效方法, 通过对交叉列表(Cross-table)结构的研究, 揭示变量不同水平间的对应关系, 是市场研究中经常用到的统计技术。

一、基本原理

假定某产品共有 p 个品牌, 形象评价用语 n 个, 以 X_{ij} 表示“认为第 i 个品牌具有第 j 个形象”的人数, 以 X_{it} 表示评价第 i 个品牌的总人数, X_{jt} 表示回答第 j 个形象的总人数 ($i=1,2,\dots,p; j=1,2,\dots,n$), 见表 1。

表 1 对应分析用交叉表

属性、水平		形象空间				
		1	2	...	n	total
品牌	1	X_{11}	X_{12}		X_{1n}	X_{1t}
	2	X_{21}	X_{22}		X_{2n}	X_{2t}
	...					
	p	X_{p1}	X_{p2}		X_{pn}	X_{pt}
	total	X_{t1}	X_{t2}		X_{tn}	N

表 1 中, 每个 X_{ij} 都代表着某个品牌拥有某形象的人数, 因而这一人数是品牌与形象的‘对应’点, 体现品牌与形象的相互“选择”关系, 以矩阵 X 表示, 有:

$$X = X_{(p \times n)}$$

以 N 表示调查总人数, 有:

$$N = \sum_{i=1}^p X_{it} = \sum_{j=1}^n X_{jt} = \sum_{i=1}^p \sum_{j=1}^n X_{ij}$$

以 $P_{ij}(j=1,2,\dots,n)$ 表示形象空间构成比, 则有:

$$P_{ij} = \frac{X_{ij}}{N} \times 100\%$$

以 $P_{it}(i=1,2,\dots,p)$ 表示各品牌构成比, 则有:

$$P_{it} = \frac{X_{it}}{N} \times 100\%$$



假定各品牌无形形象特征，则各品牌在每一具体形象上的构成相同。称在此假定下第 i 品牌在第 j 形象上的人数为理论数（常用 T 表示），计算公式为：

$$T_{ij} = \frac{X_{i\cdot} \times X_{\cdot j}}{N} \quad (i=1,2,\dots,p; j=1,2,\dots,n)$$

若假定成立， X_{ij} 与 T_{ij} 应非常接近，若 X_{ij} 、 T_{ij} 差距较大，则品牌与形象间存在对应关系，用 D_{ij} 表示差值的大小，公式为：

$$D_{ij} = X_{ij} - T_{ij}$$

由于 D_{ij} 受 $X_{i\cdot}$ 及 $X_{\cdot j}$ 数值大小的影响，不利于不同 D_{ij} 间的度量，为此对其进行标准化，公式为：

$$Z_{ij} = \frac{X_{ij} - T_{ij}}{\sqrt{X_{i\cdot} X_{\cdot j}}} \quad (1)$$

常称 Z_{ij} 为标准化变异变量。从公式的推倒过程可见， Z_{ij} 体现着品牌与形象空间可能的对应关系，以矩阵 Z 表示，有：

$$Z = Z_{(p \times n)}$$

以 $F=(f_1, f_2, \dots, f_p)'$ 表示品牌的形象维度(dimension，相当于偏好因子，假定 $p \leq n$)，

$Z = (Z_1, Z_2, \dots, Z_n)'$ 表示形象向量，理论证明，存在系数矩阵 $B=B_{(p \times n)}$ ， F 、 Z 、 B 间存在如下关系：

$$F = B_{(p \times n)} \times Z \quad (2)$$

利用统计软件，计算 $Z_{(p \times n)}$ 的特征根、特征向量，通过贡献率的比较确定需截取的维度，形成对应分析图。

二、应用实例

受某家电企业的委托，方舟公司在全国 10 大城市进行了入户研究，样本量为 3000 重点检测 5 个空调品牌的形象特征，包括 4 个具体空调品牌与 1 个消费者心目中的“理想品牌”，形象空间包括热情的少年、青春的少女、成熟的职业女性等 8 个形象指标。

1、基础资料的整理

对应分析需要将品牌指标与形象指标按交叉列表的方式整理，本例整理结果见表 2.1。

表 2.1 10 城市调研基础数据

品牌	形 象 空 间								
	热情的少年	青春的少女	成熟的职业女性	淳朴的工人	稳重的大嫂	星级酒店的服务员	有身份的主管	博学的教授	Active Margin
品牌 A	543	342	453	609	261	360	243	183	2994
品牌 B	245	785	630	597	311	233	108	69	2978
品牌 C	300	200	489	740	365	324	327	228	2973
品牌 D	401	396	395	693	350	309	263	143	2950



理想品牌	147	117	410	726	366	447	329	420	2962
Active Margin	1636	1840	2377	3365	1653	1673	1270	1043	14857

2、行、列特征的描述

首先对行的特征进行探讨，将表 2 中各数据除以该行的合计值，其大小表示各品牌在形象空间上的表现程度，其中每列的合计除以总合计称为聚集度，常用 Mass 表示，计算结果见表 2.2.1。

从表 2.2.1 可见，品牌 A 主要体现“淳朴的工人、成熟的女性”、品牌 B 主要在“青春的少女、成熟的职业女性”，其余类推。

同理对列的特征进行探讨（结果见表 2.2.2），从表 2.2.2 可见，“热情的少年”的主要体现在品牌 A 上，“青春的少女”主要体现在品牌 B 上，其余类推。

将行的特征、列的特征结合在一起，综合反映品牌与形象的关系，构成了对应分析的主要内容。

表 2.2.1 各行的基本特征(Row Profiles)

品牌	形象空间								
	热情的少年	青春的少女	成熟的职业女性	淳朴的工人	稳重的大嫂	星级酒店的服务员	有身份的主管	博学的教授	Active Margin
品牌 A	0.18	0.11	0.15	0.20	0.09	0.12	0.08	0.06	1.00
品牌 B	0.08	0.26	0.21	0.20	0.10	0.08	0.04	0.02	1.00
品牌 C	0.10	0.07	0.16	0.25	0.12	0.11	0.11	0.08	1.00
品牌 D	0.14	0.13	0.13	0.23	0.12	0.10	0.09	0.05	1.00
理想品牌	0.05	0.04	0.14	0.25	0.12	0.15	0.11	0.14	1.00
Mass	0.11	0.12	0.16	0.23	0.11	0.11	0.09	0.07	

表 2.2.2 各列的基本特征(Column Profiles)

品牌	形象空间								
	热情的少年	青春的少女	成熟的职业女性	淳朴的工人	稳重的大嫂	星级酒店的服务员	有身份的主管	博学的教授	Mass
品牌 A	0.33	0.19	0.19	0.18	0.16	0.22	0.19	0.18	0.20
品牌 B	0.15	0.43	0.27	0.18	0.19	0.14	0.09	0.07	0.20
品牌 C	0.18	0.11	0.21	0.22	0.22	0.19	0.26	0.22	0.20
品牌 D	0.25	0.22	0.17	0.21	0.21	0.18	0.21	0.14	0.20
理想品牌	0.09	0.06	0.17	0.22	0.22	0.27	0.26	0.40	0.20
Active Margin	1	1	1	1	1	1	1	1	

3、计算各对应点标准化变异量、矩阵

将表 2.1 中的数据代入公式 (1)，计算 Z 值，并计算各行、列的极差 R，见表 2.3。从品牌的角度，品牌 B 的形象最清晰 (R=0.26)，主要表现在“青春的少女”上，其次是理想品牌的表现 (R=0.23)；从形象空间的角度，品牌突出的是“青春的少女” (R=0.29) 与“博学的教授” (R=0.20)。

表 2.3 各对应点标准化变异值表 (Z 矩阵)



品牌	热情的少年	青春的少女	成熟的职业女性	淳朴的工人	稳重的大嫂	星级酒店的服务员	有身份的主管	博学的教授	极差 R
品牌 A	0.10	-0.01	-0.01	-0.02	-0.03	0.01	-0.01	-0.02	0.13
品牌 B	-0.04	0.18	0.06	-0.02	-0.01	-0.05	-0.08	-0.08	0.26
品牌 C	-0.01	-0.07	0.01	0.02	0.02	0.00	0.04	0.01	0.11
品牌 D	0.03	0.01	-0.03	0.01	0.01	-0.01	0.01	-0.04	0.07
理想品牌	-0.08	-0.11	-0.02	0.02	0.02	0.05	0.04	0.12	0.23
极差 R	0.18	0.29	0.09	0.05	0.05	0.10	0.11	0.20	

4、计算惯量，确定维度

惯量(inertia)实际上就是常说的特征根值(Eigenvalue)，表示相应维度对各类别的解释量，维度的数量最大等于“行变量数-1”与“列变量数-1”中的较小者，本例最多可以产生 4 个维度。从计算结果（表 2.4）可见，第一维度的解释量达 75%，前 2 个维度的解释量已达 95%。

选取几个维度对结果进行分析，需结合实际情况，一般解释量累计比例达 85%以上即可获得较好的分析效果，故本例取两个维度即可。对应分析中与惯量相联系的另一个概念是奇异值(Singular Value)，它等于惯量的平方根，表示交叉表中每个对应点的行分值与列分值的相关系数。

表 2.4 各维度的惯量、奇异值

Dimension	Singular Value	Inertia	Proportion	Cumulative
1	0.290	0.084	0.750	0.750
2	0.150	0.022	0.200	0.950
3	0.064	0.004	0.037	0.987
4	0.039	0.002	0.013	1.000
Total		0.112	1.000	1.000

5、计算行、列在各维度的得分（投影）

针对每一个惯量，由公式（2）分别计算行（品牌）、列（形象空间）在相应维度上得分，计算结果见表 7。

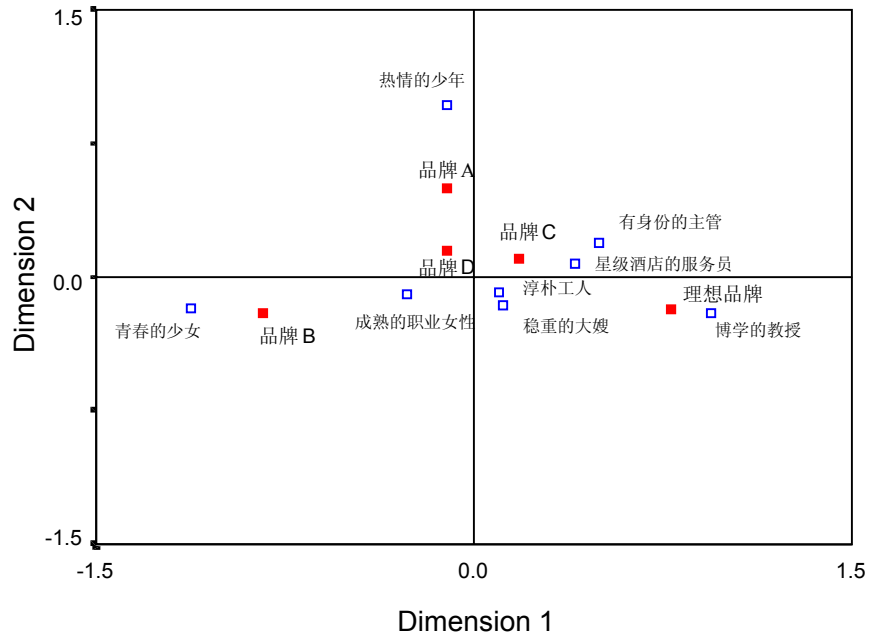
表 2.5 各类别在每个维度上的得分（投影）

品牌	Score		形象空间	Score	
	Dim1	Dim2		Dim1	Dim2
品牌 A	-0.050	0.577	热情的少年	-0.181	1.030
品牌 B	-0.890	-0.411	青春的少女	-1.142	-0.275
品牌 C	0.305	0.017	成熟的职业女性	-0.248	-0.195
品牌 D	-0.104	0.244	淳朴的工人	0.134	-0.049
理想品牌	0.742	-0.430	稳重的大嫂	0.119	-0.174
			星级酒店的服务员	0.357	0.003
			有身份的主管	0.566	0.127
			博学的教授	0.979	-0.414

6、对应分析图（定位图）

对应分析最精彩之处是将品牌指标与形象指标直观地表达在同一张图形上。将表 2.5 反映在 2 维坐标系下即得对应分析图（见图 1），从图中可以非常直观地反映出品牌 A 是“热情的少年”，品牌 B 是“青春的少女”，品牌 C 象是“淳朴的工人、稳重的大嫂”，品牌 D 是“成熟的职业女性”，而理想的品牌是“博学的教授”。

图 1 Correspondence Analysis



7、补充

由于品牌与形象指标在同一个坐标系下，可以借助欧氏距离公式从数量的角度度量品牌与形象指标间的密切程度，计算结果见表 2.7。表 2.7 实际上是表 2.2.1、表 2.2.2 的综合体现，从中可见，品牌 A 的形象主要是“热情的少年”，品牌 B 的形象主要是“青春的少女”，品牌 C 的形象主要是“淳朴的工人、稳重的大嫂、星级酒店的服务员、有身份的主管”，品牌 D 的形象主要是“成熟的职业女性”，理想品牌的形象主要是“博学的教授”。

另一方面，以欧氏距离可以度量品牌间的关系，如：品牌 A、B、C、D 与理想品牌的距离分别为 1.28、1.63、0.62、1.08，由此可见品牌 C 最接近理想品牌。

表 2.7 各品牌与各形象间的距离

品牌	热情的少年	青春的少女	成熟的职业女性	淳朴的工人	稳重的大嫂	星级酒店的服务员	有身份的主管	博学的教授
品牌 A	0.47	1.38	0.80	0.65	0.77	0.70	0.76	1.43
品牌 B	1.61	0.29	0.68	1.09	1.04	1.31	1.55	1.87
品牌 C	1.12	1.48	0.59	0.18	0.27	0.05	0.28	0.80
品牌 D	0.79	1.16	0.46	0.38	0.47	0.52	0.68	1.27
理想品牌	1.73	1.89	1.02	0.72	0.67	0.58	0.58	0.24

三、讨论与建议

1、对应分析的优点



对应分析的优点比较明显，主要体现在：

- (1) 定性指标的类别越多，越容易刻画相互间的关系；
- (2) 提供了将定性变量按定量方法处理的途径，从数量的角度揭示交叉列表行列变量间关系的方法；
- (3) 对应分析图将不同属性的指标反映在同一坐标系下，为品牌定位研究提供了一个有效的方法。

2、局限与建议

主要体现在：

- (1) 不能用于假设检验，并且品牌间、品牌与形象指标间的距离是一个相对关系，其距离大小无实际意义，如：品牌 A 与“热情的少年”距离为 0.5，品牌 B 与“热情的少年”距离是 1.0，并不表示品牌 A 在这一形象上是品牌 B 的 2 倍。
- (2) 要求样本随机性获得，对拦截访问资料、重点调查资料等非随机资料，进行对应分析时需根据配额计算调整系数。
- (3) 当对应点有 0 出现，尤其某类样本数量占样本总量的比例较小时，出现分析偏性的可能性较大，并且受极端值的影响较大。

（特别声明：按公司项目保密规定，本文所引用数据已进行一定处理，重点体现分析思路和方法）