

# 面向制造企业的信息化评价方法

王慧明<sup>1</sup>, 齐二石<sup>1</sup>, 王慧敏<sup>2</sup>

(1. 天津大学管理学院, 天津 300072; 2. 北京市经济管理干部学院, 北京 100102)

**[摘要]** 基于因子分析方法, 对天津市 51 家制造企业的信息化状况进行定量的综合评价, 并采用聚类分析方法对评价结果进行了分类, 找出了各类企业的优势和不足, 为企业信息化工作的开展提供了科学依据。

**[关键词]** 制造企业信息化; 因子分析; 聚类分析

**[中图分类号]** C939 **[文献标识码]** A **[文章编号]** 1009-1742(2004)08-0086-05

## 1 引言

制造企业信息化 (manufacturing enterprise informatization) 是制造企业将信息技术、自动化技术、现代管理技术与制造技术相结合, 改善制造企业的经营、管理、产品开发和生产等各个环节, 提高生产效率、产品质量和企业创新能力, 降低消耗, 带动产品设计方法和设计工具的创新, 从而全面提升企业的竞争力<sup>[1]</sup>。制造企业选择实施信息化是基于企业生存发展的内在动力和源于竞争对手、顾客等外部压力综合作用下的战略选择, 目的是创造和保持持久的竞争能力和竞争效果。制造企业信息化是信息技术环境下制造企业发展的必然选择。制造企业信息化是一项系统工程, 从属于信息化大系统。制造企业信息化建设过程, 实际上也是一个不断评价和完善的过程。制造企业信息化评价从表面上看是一个检验体系, 事实上更重要的是一个企业信息化建设的指导体系。有了这个体系就可以帮助企业有针对性地解决信息技术运行过程中的效率、成本、服务、技术创新等方面的问题。由于中国制造企业信息化尚处于初级阶段, 与信息化技术研究相比, 信息化评价研究相对滞后。

为了对制造企业信息化状况进行综合评价, 通常采用的是简单的加权合成法。这种方法的突出问

题是主观赋权, 人为地给定权数, 一方面会导致对某一个因素过高或过低的估计, 使评价结果不能反映企业的真实情况; 另一方面会诱使企业粉饰或片面追求权重较高的指标。而多元统计分析中的因子分析法在构造综合评价值时所涉及的权数都是从数学变换中伴随生成的, 不是人为确定的, 因此具有客观性<sup>[2]</sup>。笔者利用因子分析法, 以天津市制造企业为例, 对其信息化状况进行综合分析评价, 取得了较好的分析评价结果。

## 2 评价指标选取与调查数据整理

### 2.1 指标选取

制造企业信息化评价指标是用来衡量制造企业信息化实施状况的。在选取指标时, 应力求全面、真实地反映制造企业信息化的实施状况。体现制造业的基本特征, 并且指标的选取还应遵循以下原则: a. 综合性——选择独立性较强、能综合反映指标群特征的指标, 尽量避免选择意义相近的指标; b. 实用性和易操作性——各指标含义明确, 数据资料容易获得, 计算方法简明易懂; c. 可比性——保证指标口径的一致性, 便于横向、纵向比较。在征求了京津两地研究、管理和生产领域多位专家的意见后, 提出了制造企业信息化评价指标共 17 项, 如图 1 所示。

**[收稿日期]** 2003-12-13; **修回日期** 2004-02-03

**[基金项目]** “八六三” 高新技术研究发展计划资助项目 (2003AA4Z2040)

**[作者简介]** 王慧明 (1976-), 女, 内蒙古赤峰市人, 天津大学博士研究生

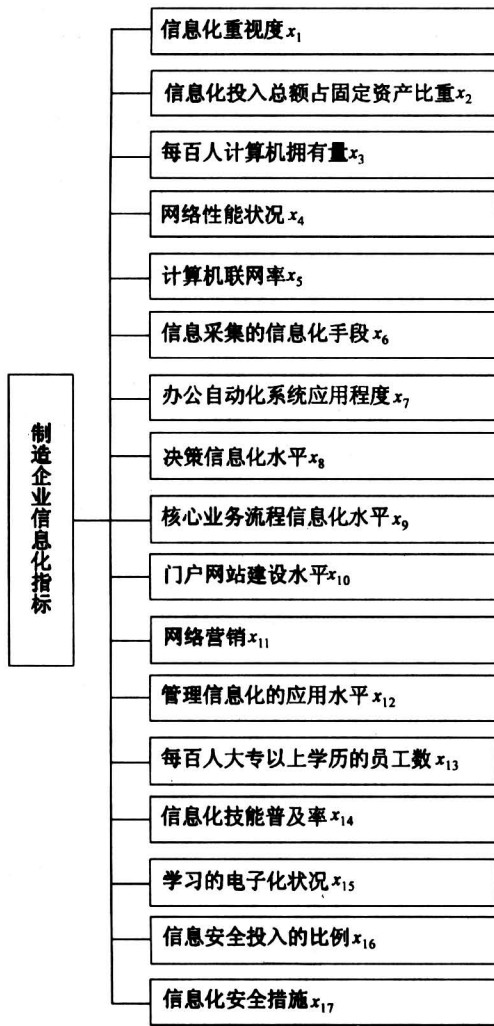


图 1 制造企业信息化评价指标

Fig.1 Evaluation architecture for manufacturing enterprise informatization

### 2.2 调查数据整理

将评价指标转化为问卷中的调查问题，形成调查问卷。调查问题的数量要合适，要便于调查结果的统计和处理，否则，问卷调查结果会很不理想，甚至徒劳无功<sup>[3]</sup>。笔者采用结构性问卷调查，调查表设计成选择题的形式，对选择题的备选答案做了精心设计。定性指标可量化为 1—100 分。为了使因子分析更加有效，企业个数要大于指标个数，且越大越好，但在实际计算中也不可能取太多企业，笔者对 2001 年天津市制造业中的食品加工及制造和饮料制造业、纺织业、普通机械、专用设备制造业的企业信息化状况进行了调查。共发出问卷 90 份，实际收回问卷 76 份。

为了有效地统计分析，需要对调查数据进行整

理，笔者对所回收的调查问卷做了以下工作：

1) 修正数据。部分调查问卷显示的数据有明显的错误。比如，有的企业填写时忽略了计量单位“万元”，错误地以“元”为单位填写，此类错误在数据整理中都须加以修正。

2) 删除部分样本。即删除明显不合理的样本。比如，有的调查企业参加培训的管理者人数大于该企业的管理人员总数，数据间显示矛盾，这是不合理的。为了避免这类样本对研究结果准确性造成影响，按照谨慎性原则将此类样本删除。

3) 缺失值处理。用 SPSS 数据分析软件进行数据分析时，需要对缺失值进行处理。在统计分析中笔者选用平均值替代缺失值。

整理后有效问卷有 51 份。按出资者分，内资企业 25 家，中外合资企业 16 家，外商独资企业 10 家；按企业规模分，大型企业 13 家，中型企业 24 家，小型企业 14 家。

## 3 企业信息化状况的因子分析

### 3.1 因子分析理论基础

20 世纪初期，K. Pearson 和 C. Spearman 等人关于智力的定义和测量工作，开始了因子分析的近代发展。因子分析是主成分分析的一种推广。因子分析是一种通过对样本相关阵的内部依赖关系的研究，将多个指标归结为少数几个不可观测的综合指标以简化观测系统，并最终对各样品施行定量化评价的多元统计分析方法<sup>[4]</sup>。它的基本思想是根据相关性的大小把变量分组，使得同组内的变量之间相关性较高，不同组之间的变量相关性较低，其关键在于正交因子模型的建立和分析应用。

设  $x_i (i = 1, 2, \dots, p)$  为评价指标随机变量，其中某些随机变量之间具有不同程度的相关性。假定每个变量已经过标准化，正交因子模型指的是：存在  $m (m \leq p)$  个综合指标（又称主因子） $f_i (i = 1, 2, \dots)$ ，使  $x_i$  可用它们的线性组合表示为<sup>[5]</sup>

$$X = AF + \varepsilon,$$

式中  $X = (x_1, x_2, \dots, x_p)'$  为可观测的随机向量； $F = (f_1, f_2, \dots, f_m)'$  为公因子； $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)'$  为特殊因子； $A$  为因子载荷矩阵， $A = (a_{ij})_{p \times m}$  ( $a_{ij}$  称为第  $i$  个指标在第  $j$  个主因子上的载荷)，且满足条件  $f_1, f_2, \dots, f_m$  不相关且方差皆为 1， $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$  不相关且方差不同； $\text{cov}(\varepsilon, F) = 0$ ，即  $F$  和  $\varepsilon$  不相关； $m = p$  时  $\varepsilon = 0$ 。

上述模型在具体应用时,可合理选定主因子个数  $m$ ,略去线性表达式中的  $\epsilon$ ,于是式中  $A$  的估计与计算对最终评价至关重要。

3.2 制造企业信息化综合得分计算

笔者就所涉及的实际问题,以主成分法具体确定主因子,计算有实际统计分析意义的因子载荷矩阵,展开相关讨论分析。

1) 对原始数据  $x_{ij}$  (第  $x_j$  个指标在第  $i$  个样品上的原始观测值) 进行标准化处理,以消除量纲对评价结果的影响,于是得到标准化后的矩阵  $X$ 。

2) 求出标准化后的样本相关系数矩阵  $R = (r_{jk})$ 。计算相关矩阵  $R$  的特征值  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{17} \geq 0$  (特征值依次为 4.539, 1.979, 1.571, 1.498, ...) 和对应的特征向量  $e_1, e_2, \dots, e_{17}$ 。

3) 使用累积方差贡献率来确定主因子个数  $m$ 。第  $j$  个因子的方差贡献率为  $\lambda_j/17$ 。贡献率可

衡量出每个主因子所占原始数据信息量的多少,贡献率越大,该因子相对越重要。同时以主因子的累积方差贡献率  $\sum \lambda_j/17$  来确定因子数  $m$ 。由于前 8 个主因子的累积方差贡献率为 89.16%,即能用前 8 个综合指标来反映原始数据表中的绝大部分信息,故抽出  $m=8$  个主因子。

4) 建立主成分因子分析的载荷矩阵为  $A = (\sqrt{\lambda_1}e_1, \sqrt{\lambda_2}e_2, \dots, \sqrt{\lambda_8}e_8)$ 。

5) 因子旋转。得到的初始因子载荷矩阵  $A$ ,由于因子载荷的大小相差不大,由此造成对因子的解释有困难,因此,为了得出较明确的分析结果,采用了方差最大正交旋转,使每个因子载荷向 0 和 1 两极分化,以达到载荷二次方的方差最大。旋转后得到正交载荷矩阵  $A^*$  (见表 1),完全符合简单结构的准则,而且每个主因子都有明确的含义。

表 1 评价指标的载荷值

Table 1 Factor score of evaluation indexes

评价指标	载 荷 值							
	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$	$f_7$	$f_8$
$x_1$	0.131	-0.048 7	0.064	0.898	-0.027 8	0.149	0.072	0.061 9
$x_2$	-0.181	0.020 1	-0.010 8	0.036 7	-0.018 3	0.006 76	-0.916	0.0014
$x_3$	0.012 9	0.907	0.162	0.062 6	-0.064 7	0.021 5	0.007 79	0.119
$x_4$	0.110	-0.105	0.051	0.016 1	-0.056 1	0.025	-0.036 4	0.048 3
$x_5$	0.321	0.352	0.102	0.022 2	-0.012 7	0.663	0.279	0.110
$x_6$	0.213	0.112	0.194	0.116	0.078 5	-0.113	0.042 7	0.886
$x_7$	0.132	0.299	0.856	0.227	0.193	0.019 3	0.259	-0.092 2
$x_8$	0.883	0.077 1	0.064	0.142	0.187	0.102	-0.094 9	0.055 6
$x_9$	0.810	0.226	0.131	0.133	-0.036 8	-0.007 3	-0.411	0.089
$x_{10}$	0.087 3	0.209	0.882	-0.098 3	0.122	0.107	-0.135	0.280
$x_{11}$	0.022 6	-0.189	0.019 5	0.120	0.041	0.884	-0.119	-0.173
$x_{12}$	0.873	-0.099 8	0.228	-0.023 1	0.394	0.167	0.094 4	0.128
$x_{13}$	0.116	0.872	0.219	0.006 9	0.329	-0.082 6	-0.022 1	0.019 3
$x_{14}$	-0.058 5	0.117	0.186	0.527	0.389	-0.018 4	-0.251	0.493
$x_{15}$	0.253	0.080 2	0.132	0.080 6	0.874	0.040 8	-0.004 4	0.086 3
$x_{16}$	0.128	0.324	-0.107	0.563	0.462	-0.052 2	-0.002 3	0.089 2
$x_{17}$	0.453	-0.023 7	0.597	0.036 7	-0.163	-0.040 0	-0.186	0.259

6) 采用回归方法计算各企业的主因子得分,将主因子表示为指标变量的线性组合:  $f_j =$

$$\sum_{k=1}^{17} \beta_{jk} x_k, (j = 1, 2, \dots, 8)$$

7) 计算各企业信息化状况的综合得分。企业

在各主因子上的得分乘以对应方差贡献率再求和,便得到综合得分(排序结果如表 2 所示):

$$F_{MEI} = 0.317 0 f_1 + 0.166 4 f_2 + 0.092 4 f_3 + 0.088 1 f_4 + 0.071 1 f_5 + 0.061 5 f_6 + 0.049 9 f_7 + 0.045 2 f_8,$$

通过综合得分  $F_{MEI}$  的大小可以对企业的信息化水平进行评价。综合得分多，意味着该企业的信

表 2 制造企业信息化评价

Table 2 Evaluation of manufacturing enterprise informatization

企业名称	$F_{MEI}$	企业名称	$F_{MEI}$	企业名称	$F_{MEI}$	企业名称	$F_{MEI}$	企业名称	$F_{MEI}$
企业 1	2.793 7	企业 12	0.259 5	企业 23	-0.195 0	企业 34	-0.377 6	企业 45	-0.561 4
企业 2	2.263 6	企业 13	0.230 3	企业 24	-0.200 1	企业 35	-0.382 2	企业 46	-0.623 2
企业 3	1.787 3	企业 14	0.214 3	企业 25	-0.216 2	企业 36	-0.396 2	企业 47	-0.664 5
企业 4	1.622 7	企业 15	0.199 8	企业 26	-0.223 9	企业 37	-0.442 1	企业 48	-0.809 7
企业 5	1.249 6	企业 16	0.121 7	企业 27	-0.296 6	企业 38	-0.452 2	企业 49	-0.828 3
企业 6	0.897 4	企业 17	-0.096 5	企业 28	-0.334 3	企业 39	-0.484 5	企业 50	-0.878 5
企业 7	0.838 3	企业 18	-0.130 2	企业 29	-0.345 6	企业 40	-0.484 5	企业 51	-0.904 5
企业 8	0.786 2	企业 19	-0.147 3	企业 30	-0.348 2	企业 41	-0.491 5		
企业 9	0.477 7	企业 20	-0.158 4	企业 31	-0.349 5	企业 42	-0.513 4		
企业 10	0.445 6	企业 21	-0.178 6	企业 32	-0.362 6	企业 43	-0.515 4		
企业 11	0.313 4	企业 22	-0.183 8	企业 33	-0.377 3	企业 44	-0.547 5		

虽然因子分析的计算过程比较复杂，但是通过运用 SPSS 统计软件中的因子分析 (FACTOR ANALYSIS) 模块可以十分方便地完成上述运算，为实际分析提供了有力的运算支持。

#### 4 信息化评价结果的聚类分析

聚类分析是将研究对象按其共性进行分类，以便系统地加以科学研究的一种有效方法，其目的在于辨识在某些特征上相似的事物，并把事物就这些特征划分成若干类，使同一类事物具有高度的共质性，而不同类事物具有高度的相异性<sup>[6]</sup>。聚类分析方法常用的有系统聚类法、动态聚类法、模糊聚类法，等等。笔者采用系统聚类法对上述综合得分进行最优分割，把样本按综合得分高低分成三类，再通过计算各类各指标的平均值来找出各企业的优势和差距，以便各企业制定提高信息化水平的策略。

系统聚类法是在聚类分析中应用最为广泛的一种方法，它的基本原理是<sup>[7]</sup>：首先将一定数量的样品或指标各自看成一类，再根据样品（或指标）的亲疏程度，将亲疏程度最高的两类进行合并，然后考虑合并后的类与其他类之间的亲疏程度，再进行合并。重复这一过程，直至将所有的样品（或指标）合并为一类。衡量亲疏程度的指标有两种，即距离和相似系数。其中，距离是将每个样品看成是  $m$  个变量对应的  $m$  维空间中的一个点，在该空间中所定义的距离越近，则亲密程度越高。系统聚类法中类间距离的常用计算方法主要有最短距离法、最长距离法、中间距离法、重心法和类平均法等。

1) 数据准备。把上述所得的综合得分记为

$F_{MEI1}, F_{MEI2}, \dots, F_{MEI51}$ 。

2) 聚类原理。选用聚类分析中的系统聚类法。采用欧氏距离  $d_{ij} = F_{MEIi} - F_{MEIj}$ ，表示样品  $i$  和样品  $j$  之间的距离，并采用系统聚类法中常使用的由 Sokal 和 Michener 提出的类平均距离公式

$$D_{pq}^2 = \frac{1}{n_p n_q} \sum_{\substack{i \in G_p \\ j \in G_q}} d_{ij}^2,$$

式中  $n_p$  与  $n_q$  为类  $G_p$  与  $G_q$  中元素的个数。 $d_{ij}$  为  $G_p$  中元素  $f_i$  与  $G_q$  中元素  $f_j$  间的欧氏距离。

3) 聚类结果。对  $F_{MEI}$  按上述原理进行聚类，用 SPSS 统计分析软件进行上述计算过程，可得聚类谱系图（略）。对该图进行分析，可将样品分成三类，各类所包含的样品如下：第一类包括企业 1—企业 4，第二类包括企业 5—企业 16，第三类包括企业 17—企业 51。

4) 分类结果分析。上述分类中，第一类企业的综合信息化水平高，其综合得分大于 1.5；第二类企业的综合信息化水平一般，其综合得分介于 0.15 与 1.5 之间；第三类企业的综合信息化水平低下，其综合得分小于 0.15。这三类企业相应的各指标的平均值列于表 3。

对上述计算结果进行分析可以得出以下结论：

1) 第一类企业  $x_4$  和  $x_{11}$  的平均值是所有三类企业中最差的，说明其综合信息化水平虽然最高，但它的平均网络性能状况和网络营销率却不佳。对其所含的 4 个企业在  $x_4$  上进行分析，发现除第 4 个

企业外,其他3个企业的Internet出口带宽都小于128 kB。在 $x_{11}$ 即网络营销率上进行分析,发现除第4个企业外,其他3个均未开展电子商务业务。

表3 各指标的平均值

Table 3 Average of indexes

指标	第一类企业	第二类企业	第三类企业
$x_1$	65.25	78.25	60.34
$x_2$	39.41	27.98	24.46
$x_3$	10.80	10.33	7.81
$x_4$	25.00	30.00	25.14
$x_5$	52.93	65.41	40.54
$x_6$	67.50	53.75	46.29
$x_7$	39.00	25.83	10.40
$x_8$	65.00	55.00	31.14
$x_9$	65.00	55.83	45.71
$x_{10}$	44.50	23.89	19.74
$x_{11}$	0.63	7.37	1.36
$x_{12}$	47.63	34.46	19.17
$x_{13}$	38.64	13.92	12.00
$x_{14}$	50.64	27.34	14.59
$x_{15}$	38.43	12.13	0.04
$x_{16}$	27.50	15.00	1.71
$x_{17}$	35.28	22.83	24.19

2) 第一类企业对各指标的平均值与第二类比较,可以看出两者各有千秋,但综合来看,第一类优于第二类。第一类企业对 $x_2, x_3, x_6, x_7, x_8, x_9, x_{10}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}$ 和 $x_{17}$ 的平均值都明显优于第二类企业相应的平均值,而前者对 $x_1, x_4, x_5, x_{11}$ 的平均值都稍逊于后者,因此从总体来看,前者综合水平高于后者。

3) 总体来看,除 $x_1, x_4, x_5, x_{11}$ 对第一类企业异常和 $x_{17}$ 对第二类企业异常外,其他各指标都呈现出从第一类企业向第三类企业递减的态势。

用系统聚类法划分出来的三类企业间的信息化水平的差异是明显的。各企业可根据上述分析结果制定对策,发挥原有优势,弥补不足,从而提高信息化水平。

## 5 结语

笔者综合运用多元统计方法的因子分析和聚类分析对天津市51家制造企业的信息化状况进行综合排序和分类。通过因子分析法对指标体系量化求解,最终得到制造企业信息化的综合得分;用系统聚类法对综合得分进行分析,得到各企业信息化状况的分布情况。笔者的研究结果为天津市制造企业制定信息化战略,分析企业信息化状况提供了科学依据,具有重要的理论与实际意义。

### 参考文献

- [1] 杨海成. 制造业信息化工程总体实施方案 [J]. 航空制造技术, 2002, 7 (7): 17~22
- [2] 何有世, 徐文芹. 因子分析法在工业企业经济效益综合评价中的应用 [J]. 数理统计与管理, 2003, 22 (1): 19~22
- [3] 李晶. 社会调查方法 [M]. 北京: 中国人民大学出版社, 2003
- [4] 王学仁, 王松桂. 实用多元统计分析 [M]. 上海: 上海科学技术出版社, 1990
- [5] 安希忠, 林秀梅. 实用多元统计方法 [M]. 吉林: 吉林科学技术出版社, 1992
- [6] Dilon W R, Goldstein M. Multivariate Analysis, Methods and Applications [M]. New York: Wiley, 1984
- [7] 柯惠新, 黄京华, 沈浩. 调查研究中的统计分析法 [M]. 北京: 北京广播学院出版社, 1992

## A Study on Evaluation Method of Manufacturing Enterprise Informatization

Wang Huiming<sup>1</sup>, Qi Ershi<sup>1</sup>, Wang Huimin<sup>2</sup>

(1. School of Management, Tianjin University, Tianjin 300072, China; 2. Department of Economy and Trade, Beijing Economic Management Institute, Beijing 100102, China)

[Abstract] Using the factor analysis approach, the paper makes quantitative analysis and comprehensive evaluation of the 51 manufacturing enterprises in Tianjin in terms of informatization. Then, the method of cluster analysis is adopted to classify enterprises in order to find out their advantages and disadvantages, and to provide a scientific basis for the informatization development in manufacturing enterprises.

[Key words] manufacturing enterprise informatization; factor analysis; cluster analysis