응용통계연구 제 **20**권 *1*호, *2007*년, *pp. 23–38*

요인분석과 군집분석을 통한 세분화 및 전략방향 제시: 특수법인 사례를 중심으로

조용준[1)] 김영화[2)]

요 약

세분화는 크게 목적변수의 유무에 따라 분석방법이 달라지게 된다. 본 논문은 특수법인의 경영지표를 바탕으로 목적변수가 존재하지 않을 경우의 세분화를 통해 전략방향을 도출하는 사례연구를 제안하고자 한다. 군집분석을 통한 세분화의 경우, 많은 변수를 사용하여 분류를 하게 되면 군집별 특성화가 어렵게 된다. 따라서 군집의 특성을 잘 반영할 수 있는 대표적 요인변수를 요인분석을 통해 추출하고 이 대표요인을 바탕으로 2단계 군집분석을 통한 세분화를 고려하였다. 이를 통해 총 6개의 세분화 군집을 도출하고 각 군집별 강점요인을 강화하고 약점요인을 보완하는 방향으로 전략방향을 설정하여 제안하고자 한다.

주요용어: 2단계 군집분석, 세분화, 요인분석.

1. 서론

대량의 데이터를 보다 효과적으로 구분하여 특성이 비슷한 집단으로 분류하는 대표적인 방법인 세분화(segmentation)는 경영전략 측면에서도 효과적인 관리와 전략수립을 위한 필수적 방법이다. 왜냐하면 세분화를 통해서 각 세분화 집단별 특성을 파악할 수 있고 특성을 반영한 차별적 관리가 가능해지기 때문이다. 최근의 경영기법 중에서 세분화를 통한 관리는 그 효율성으로 인해 기업의 마케팅 분야에서 널리 사용되고 있다. 이것의 대표적인 예가 고객을 세분화하여 충성고객, 우수고객, 일반고객, 이탈고객 등으로 분류하여 차별화된 혜택제공과 마케팅 방법을 달리 하는 것이다.

이러한 세분화를 위한 데이터마이닝 기법은 목적변수의 유무에 따라 달라진다(Berry와 Gordon, 1997). 목적변수가 존재할 때 사용되는 알고리즘에는 신경망의 다층퍼셉트론(MLP), 방사형 기저함수(RBF)와 의사결정나무분석의 C5.0, CART, CHAID, QUEST, 로지스틱 회귀분석, 다중회귀분석, 판별분석 등이 있다(조용준과 허준, 2006). 목적변수가 존재하지 않을 때 사용되는 알고리즘은 군집분석으로 한정되는데 실무에서 그 활용도가 매우 높다고 할 수 있다. 즉, 특정 목적이 없을 경우나, 특정 목적에 대한 기준이나 변수가 없는 경우에 세분화를 수행할 수 있고 이러한 세분화를 통해 특정 목적의 기준을 제시하여 이를 통

1) (138-730) 서울시 송파구 신천동 11-6, 수협 수산경제연구원, 수석연구원
E-mail: cyj66@chol.net
2) (156-756) 서울시 동작구 흑석동 221, 중앙대학교 자연과학대학 통계학과, 부교수
E-mail: gogators@cau.ac.kr

해 목표변수로 활용할 수 있다는 점에서 중요하다고 할 수 있다(조용준과 허준, 2006). 데이터 마이닝 분야에서의 군집분석 방법은 K-평균(means) 군집분석으로 대표되는(Selim과 Ismail, 1984; Kövesi 등,2001) 최적화 방법과 계층적 군집화 방법(Hand, 1981), 자율학습에 의한 자기조직화 지도(SOM : self-organization maps)의 코호넨 네트워크(Kohonen, 1982), BIRCH(Balanced Iterative Reducing and Clustering using Hierarchies: Zhang 등, 1996) 알고리즘을 발전시킨 2단계 군집(2 Step Cluster) 분석(Chiu 등, 2001) 등이 있다.

본 연구는 특수법인의 사례에 대하여 세분화모형을 도출하고 이에 대한 각 세분화 집단별 전략방향을 설정하는 것을 목적으로 한다. 또한 데이터의 특성상 목표변수가 존재하지 않았을 때의 세분화방법을 사용한다. 즉, 변수의 특성상 목표변수가 존재하지 않기 때문에 군집분석을 통하여 데이터를 세분화하였으며 군집분석의 방법은 범주형 변수와 연속형 변수를 모두 사용할 수 있는 2단계 군집분석방법을 사용하였다. 이러한 군집분석을 통해 세분화를 수행할 경우 분류된 각 군집의 특성을 산출하여 군집마다의 차이점을 제시하여야 하는데 이를 군집별 특성화(characterizing)로 명명하도록 한다. 이러한 군집별 특성화를 하는 경우 변수의 수가 많게 되면 각 변수마다의 군집 기준이 많아져 각 군집별 특성의 도출이 매우 힘들게 된다. 이를 위하여 군집의 특성을 잘 반영할 수 있는 대표적 요인변수를 추출하여 이 대표변수를 통해 군집분류를 수행하는 것이 일반적 방법이다. 이러한 대표적 변수의 추출 또는 축약을 위해 사용되는 방법이 요인분석이며 이를 통해 군집분류의 특성을 반영할 대표변수를 추출하고 이를 바탕으로 군집분석을 통해 세분화하는 방법을 통해 전략방향을 제시하고자 한다. 이러한 방법의 분석은 실무에서 많이 응용되고 있으며, 실무분야에서의 연구도 활발히 이루어지고 있다(최호현과 김선범, 2006; 이종상, 2002; 남영우와 성은영, 2001; 임종호, 1993).

2. 실증분석을 위한 설계

2.1. 분석 프로세스

전술한 바와 같이 본 연구에서는 군집분석 방법으로 2단계 군집분석을 사용하였는데, 이는 서론에서 언급한 바와 같이 2단계 군집분석이 범주형 변수와 연속형 변수 모두를 사용하여 군집분석을 할 수 있는 장점을 가지고 있기 때문이다. 군집분석의 대표적 방법론은 K-means 군집분석과 계층적 군집분석인데, 이 방법들은 연속형 변수만을 사용할 수 있다. 왜냐하면 이 두 가지 방법은 거리의 유사성을 측정하여 분류하기 때문에 범주형 변수의 사용이 타당하지 않기 때문이다. 범주형 변수를 더미(dummy)변수로 처리하여 사용할 수 있기는 하지만 2단계 군집분석은 모형기반 거리척도를 사용하기 때문에 이러한 가공변수 없이 사용할 수 있다. 데이터 마이닝에서 사용하는 코호넨 네트워크(Kohonen Network)는 범주형 변수로도 군집분석이 가능하지만, 정방행렬을 어떻게 구성하는가에 따라 결과가 달라지기 때문에 본 연구에서는 사용하지 않았고 2단계 군집분석방법을 사용하여 특수법인을 분류군집화하였다. 또한 실제 분석에서는 SPSS v12.0 을 사용하였다.

본 연구에서는 29개 연속형 변수와 3개의 범주형 변수를 사용하여 세분화를 하였다. 서론에서 언급한 바와 같이 너무 많은 변수를 사용하여 군집별 특성을 도출할 경우, 각 군집

별 특성의 파악이 흩어져 군집별 변별력을 찾아내기 힘들며 또한 너무 많은 차원으로 인해 분석결과의 해석이 곤란할 것으로 판단된다. 따라서 이러한 변수를 축약하여 대표성을 갖는 소수의 통합변수의 산출이 필요하므로 변수축약에 사용되는 대표적인 통계분석 방법인 요인분석(factor analysis)을 통하여 29개의 연속형 변수 중 서로 특성이 비슷한 변수들끼리 통합하여 새로운 변수를 창출하였다.

연구의 프로세스는 총 32개의 변수를 데이터 마이닝을 통해 정제 및 가공한 후, 요인분석을 통해 변수를 축약하였다. 또한 축약된 통합적 변수를 사용하여 2단계 군집분석을 통하여 각 군집별 특성을 도출하였다.

2.2. 자료의 이해

본 연구에서 사용된 데이터는 국내 94개 특수법인의 경영지표와 관련된 실제 데이터이다. 또한 분석에 사용된 변수는 일반현황, 사업실적, 재무현황, 직원성과, 경영상태현황, 비율지표 등이다. 세부적으로 사용된 변수의 내용은 표 2.1과 같으며 표 2.2는 세부적으로 사용된 변수의 기술통계량을 요약한 것이다. 표 2.3 은 범주형 자료인 특수법인의 지역분포, 업종구분, 경영상태 현황에 대한 빈도표이다.

표 2.1: 사용변수 기술서

대분류	세 부 변 수
일반 현황	지역, 법인구분, 주주수, 위판장수, 상호점포수, 기타사업장수(계), 중매인수, 임직원수
사업실적	사업실적(백만), 지도실적(백만), 경제실적(백만), 금융실적(백만)
재무현황	자산총계(백만), 예치금실적, 부채계(백만), 자본계(백만), 출자금(백만), 대손충당금(백만), 비용(백만), 수익(백만), 손익(백만)
직원성과	임직원당 순이익(천원), 임직원당 비용(백만), 임직원당 사업실적(백만), 임직원당 수익(백만)
경영상태현황	종합현황
비율지표 (파생데이터)	출자금대비손익비율, 비용대비수익률, 총자본순이익률, 자기자본순이익률, 자기자본비율, 수지비율

표 2.2: 기술통계량

변수명	최소값	최대값	평균	표준편차
주주수	25	10254	1805.46	2037.87
위판장수	0	12	1.79	2.08
상호점포수	0	2	14.04	3.91
기타사업장계	0	20	2.78	3.77
중매인수	0	227	21.98	35.68
임직원수	13	445	71.84	63.28
자산총계(백만)	3260.29	619743.46	114449.09	115262.93
예치금실적(자산일부)	66.79	190780.15	30267.43	30506.20
부채계(백만)	3012.95	697830.77	115817.39	120123.68
자본계(백만)	-89725.36	19408.96	-1368.30	13850.30
출자금(백만)	242.56	17409.63	2866.73	2489.99
대손충당금(백만)	26.50	23633.92	2502.76	2989.30
사업실적(백만)	6034	670416	100955.87	100813.39
지도실적(백만)	6	1225	188.94	228.61
경제실적(백만)	0	335221	34242.34	44493.57
금융실적(백만)	0	393511	66526.16	70825.03
비용(백만)	1626.74	150438.27	18814.06	20600.01
수익(백만)	1670.71	154858.65	19358.78	21082.27
손익(백만)	-1711.69	4420.38	544.72	835.21
출자금대비손익비율	-124.61	96.79	18.83	32.32
비용대비수익률	87.43	128.90	103.39	5.12
총자본순이익률	-3.34	2.72	0.50	0.87
자기자본순이익률	-302.04	2906.79	66.82	399.25
자기자본비율	-23.05	14.45	-0.36	7.43
수지비율	77.58	114.38	96.95	4.71
임직원당 순이익(천원)	-35660.29	31731.71	7051.84	11557.42
임직원당 사업실적(백만)	262.35	2884.83	1323.47	481.17
임직원당 비용(백만)	66.51	764.23	245.98	114.53
임직원당 수익(백만)	67.59	793.56	253.03	115.95

표 2.3: 빈도표

변수	속성	빈도	퍼센트
지역분포	서울/경기	8	8.51
	강원	9	9.58
	충청	8	8.51
	전라	26	27.66
	경상	36	38.30
	제주	7	7.45
업종구분	지역	72	76.60
	업종/가공	22	23.40
경영상태	정상	41	43.16
	부실우려	44	46.32
	부실	9	9.47

3. 실증분석 결과

3.1. 요인분석 산출결과

본 연구에서는 요인분석의 요인추출방법으로 가장 대표적인 주성분 추출방법을 사용하였으며 요인회전은 하지 않았다. 표 3.1은 요인분석결과 산출된 요인별 고유값과 각 요인개수별로 설명된 총분산이다.

표 3.1: 요인개수별 설명된 총분산

성분	고유값	% 분산	% 누적
1	11.933512	41.1500415	41.15004
2	5.14800695	17.7517481	58.90179
3	2.83155609	9.76398653	68.66578
4	2.30449695	7.94654122	76.61232
5	1.60408872	5.53134043	82.14366
⋮	⋮	⋮	⋮
25	0.00315278	0.01087167	100

고유값이 '1'이상인 요인은 총 5개 요인이었으며 이는 29개의 변수를 5개의 요인변수로 축약했을 경우 82.14%를 설명할 수 있다는 의미이다. 이에 따라 5개의 요인으로 변수를 축약하였을 경우 각 5개 요인에 가장 영향을 많이 미치는 변수를 요인별로 정리하여 최종 요인의 수를 결정하였다. 다음 표 3.2는 5개의 요인으로 축약했을 때의 각 변수별 성분행렬이

표 3.2: 요인성분행렬

변수명	1요인	2요인	3요인	4요인	5요인
주주수	0.622886	-0.1219	-0.08955	-0.13791	0.528836
위판장수	0.543056	-0.10221	0.151849	-0.49801	0.378888
상호점포수	0.843167	0.022742	-0.15008	-0.1786	-0.17338
기타사업장수계	0.620005	-0.16066	0.125512	-0.35419	0.413727
중매인수	0.512506	-0.17451	-0.04636	-0.24456	0.509293
임직원수	0.904133	0.062906	0.036689	-0.24682	-0.20065
자산총계(백만)	0.928184	-0.05238	-0.24438	0.027869	-0.22519
예치금실적(자산일부)	0.789725	-0.09156	-0.51035	0.110228	-0.13002
부채계(백만)	0.910605	-0.07958	-0.32526	0.055182	-0.19041
자본계(백만)	-0.17328	0.254319	0.78719	-0.24666	-0.2226
출자금(백만)	0.908357	0.013486	0.171782	-0.15599	0.063728
대손충당금(백만)	0.614476	-0.2456	-0.5119	0.267923	0.061825
사업실적(백만)	0.977522	-0.00409	0.028262	-0.0381	-0.13894
지도실적(백만)	0.745015	0.104953	0.392972	-0.04645	-0.25492
경제실적(백만)	0.768567	0.063936	0.456583	-0.12028	0.041211
금융실적(백만)	0.90617	-0.04634	-0.24788	0.021434	-0.22281
비용(백만)	0.941774	0.028771	0.215452	0.054979	-0.03139
수익(백만)	0.940916	0.057875	0.215231	0.052554	-0.03649
손익(백만)	0.522124	0.751241	0.118818	-0.02946	-0.14696
출자금대비손익비율	0.082081	0.916598	-0.16092	0.055174	0.033674
비용대비수익률	-0.05124	0.917294	-0.24801	-0.08445	0.075137
총자본순이익률	0.0589	0.899019	-0.03245	0.1112	0.201698
자기자본순이익률	0.068868	0.043526	0.121604	-0.07395	0.344496
자기자본비율	-0.03332	0.144512	0.665969	-0.07264	-0.23415
수지비율	0.032458	-0.94189	0.234591	0.086037	-0.06742
임직원당 순이익(천)	0.069111	0.958941	-0.05279	0.121184	0.119339
임직원당 사업실적(백만)	0.462157	-0.09749	0.066957	0.573824	0.040887
임직원당 비용(백만)	0.400647	-0.05395	0.379393	0.765664	0.238082
임직원당수익(백만)	0.402626	0.042294	0.369481	0.76836	0.247059

다. 이를 바탕으로 각 요인에 영향을 많이 미치는 주변수를 도출하도록 하였다.

표 3.3은 요인 성분행렬을 바탕으로 각 요인별 주요 영향 변수를 추출하고 선택된 변수의 특성에 따라 각 요인별 특성을 변수화시킨 것이다.

표 3.3에서 알 수 있는 것처럼 5요인은 많은 변수들 가운데 자기자본순이익률만 영향

표 3.3: 요인별 소속변수 및 요인명

요인	소 속 변 수	요인명
1요인	사업실적(백만), 비용(백만), 수익(백만), 자산총계(백만), 부채계(백만), 출자금(백만), 금융실적(백만), 임직원수, 중매인수, 상호점포수, 예치금실적(자산일부), 위판장수, 손익(백만), 경제실적(백만), 지도실적(백만), 주주수, 기타사업장계, 대손충당금(백만)	법인규모
2요인	임직원당 순이익(천), 비용대비수익률, 총자본순이익률, 출자금대비손익비율, 손익(백만)	수익률 성과
3요인	자본계(백만), 자기자본비율	자본건전성
4요인	임직원당 수익(백만), 임직원당 비용(백만), 임직원당 사업실적(백만)	임직원 성과
5요인	자기자본순이익률	자본건전성

력이 큰 것으로 나타났으며 3요인과 속성면에서 비슷하므로 5요인은 3요인에 포함하여 총 4개의 요인으로 결정하였다. 따라서 최종 결정된 4개 요인은 법인규모, 수익률 성과, 자본건전성, 임직원 성과이다. 다음 표 3.4는 최종 결정된 요인 및 각 요인에 소속된 변수이다. 여기서 법인규모의 요인변수가 높다는 것은 이러한 변수의 값이 크다는 것으로 의미하므로 법인의 규모가 크다고 할 수 있다.

표 3.4: 최종요인별 소속변수 및 요인명

요인	소 속 변 수	요인명
1요인	사업실적(백만), 비용(백만), 수익(백만), 자산총계(백만), 부채계(백만), 출자금(백만), 금융실적(백만), 임직원수, 중매인수, 상호점포수, 예치금실적(자산일부), 위판장수, 손익(백만), 경제실적(백만), 지도실적(백만), 주주수, 기타사업장계, 대손충당금(백만)	법인규모
2요인	임직원당 순이익(천), 비용대비수익률, 총자본순이익률, 출자금대비손익비율, 손익(백만)	수익률 성과
3요인	자본계(백만), 자기자본비율, 자기자본순이익률	자본건전성
4요인	임직원당 수익(백만), 임직원당 비용(백만), 임직원당 사업실적(백만)	임직원 성과

표 3.5는 요인분석에서의 계수행렬이다. 이 계수행렬은 각 변수를 표준화하였을 때 각 요인별 계수를 해당 변수에 곱하여 합하면(가중합) 해당 데이터 값의 요인성분값을 산출하게 된다. 즉, 향후에 군집분석을 통해 군집으로 분류하여 특성을 도출하였을 때, 특수법인

의 현 상태를 반영하여 군집간 이동을 하여야 하는데, 군집분석에서 산출된 모형에 요인변수의 점수를 입력해야한다. 즉, 원데이터를 사용할 수 없게 된다. 따라서 새롭게 바뀐 데이터 값을 표준화하고 요인별 변수별 계수값을 가중합하면 요인변수의 값을 산출할 수 있게 된다. 따라서 향후 특수법인의 군집별 변동을 산출하기 위한 필수 데이터가 된다.

표 3.5: 요인 계수 행렬

변수명	1요인	2요인	3요인	4요인	5요인
주주수	0.052196	-0.02368	-0.03163	-0.05984	0.32968
위판장수	0.045507	-0.01986	0.053627	-0.2161	0.236201
상호점포수	0.070655	0.004418	-0.053	-0.077	5-0.10809
기타사업장수계	0.051955	-0.03121	0.044326	-0.1537	0.25792
어촌계수	0.042947	-0.0339	-0.01637	-0.10612	0.317497
임직원수	0.075764	0.012219	0.012957	-0.1071	-0.12509
자산총계(백만)	0.07778	-0.01017	-0.08631	0.012093	-0.14038
예치금실적(자산일부)	0.066177	-0.01779	-0.18024	0.047832	-0.08106
부채계(백만)	0.076307	-0.01546	-0.11487	0.023945	-0.1187
자본계(백만)	-0.01452	0.049401	0.278006	-0.10704	-0.13877
출자금(백만)	0.076118	0.00262	0.060667	-0.06769	0.039729
대손충당금(백만)	0.051492	-0.04771	-0.18078	0.116261	0.038542
사업실적(백만)	0.081914	-0.00079	0.009981	-0.01653	-0.08662
지도실적(백만)	0.062431	0.020387	0.138783	-0.02016	-0.15892
경제실적(백만)	0.064404	0.01242	0.161248	-0.05219	0.025692
금융실적(백만)	0.075935	-0.009	-0.08754	0.009301	-0.1389
비용(백만)	0.078918	0.005589	0.07609	0.023857	-0.01957
수익(백만)	0.078847	0.011242	0.076012	0.022805	-0.02275
손익(백만)	0.043753	0.145928	0.041962	-0.01278	-0.09162
출자금대비손익비율	0.006878	0.178049	-0.05683	0.023942	0.020993
비용대비수익률	-0.00429	0.178184	-0.08759	-0.03664	0.046841
총자본순이익률	0.004936	0.174634	-0.01146	0.048253	0.12574
자기자본순이익률	0.005771	0.008455	0.042946	-0.03209	0.214761
자기자본비율	-0.00279	0.028071	0.235195	-0.03152	-0.14597
수지비율	0.00272	-0.18296	0.082849	0.037334	-0.04203
임직원당 순이익(천원)	0.005791	0.186274	-0.01864	0.052586	0.074397
임직원당 사업실적(백만)	0.038728	-0.01894	0.023647	0.249002	0.025489
임직원당 비용(백만)	0.033573	-0.01048	0.133988	0.332248	0.148422
임직원당 수익(백만)	0.033739	0.008216	0.130487	0.333418	0.154018

3.2. 군집분석 산출결과

요인분석을 통해 산출된 4개의 요인변수와 범주형 변수인 업종구분, 지역, 부실여부의 변수를 입력변수로 하여 2단계 군집분석을 수행하였다. 군집의 개수는 6개로 한정하였다. 군집의 적절한 개수에 대한 연구(허명회와 이용구, 2004)가 있었지만, 가장 최적의 군집의 수에 대한 정확한 기준으로 받아들여지는 것은 없다. 즉, 연구자가 데이터의 특성 및 목적에 따라 최적이라고 판단되는 군집의 수를 제시하는 것이 일반적이다. 본 연구에서는 군집의 개수를 4~10개 정도로 만들어 각 군집의 특성을 산출하였을 때, 각 군집끼리의 특성이 가장 이질적이고 배타적인 특징을 나타내는 것을 군집의 개수로 선정하였다. 그 결과 가장 적절한 군집의 개수가 6인 것으로 판단되었으며 따라서 6개의 군집으로 분류하였다. 일반적으로 군집을 산출할 때, 각 변수별 척도를 통일하기 위하여 변수를 표준화하여 분석을 수행하는 것이 일반적이다(박성현 등, 2002). 하지만 본 연구의 사용된 연속형 변수는 요인변수의 요인점수로서 이는 이미 표준화되어 있으므로 원 요인점수를 입력변수로 사용하였다. 범주형 변수로는 법인구분, 지역, 경영상태의 변수를 입력변수로 사용하였다. 산출된 군집분석의 결과가 다음의 표 3.6~3.10 이다. 표 3.6은 산출된 군집의 프로파일로 각 군집의 크기를 나타낸다. 군집1과 군집5가 가장 큰 비율을 차지하고 있고 전반적으로 안정적 비율을 나타내는 것으로 나타났다.

표 3.6: 군집 프로파일

군집	N	법인 중 %
군집1	29	30.85106
군집2	11	11.70213
군집3	12	12.76596
군집4	9	9.574468
군집5	24	25.53191
군집6	9	9.574468

표 3.7은 요인변수에 대한 각 군집별 평균점수이다. 평균점수가 0을 기준으로 0보다 크면 평균보다 높음을, 그렇지 않으면 낮음을 나타낸다. 요인점수는 각 변수의 원 데이터 값의 평균이 아니므로 크기를 나타내는 기준이 필요하다. 따라서 7점 리커드척도를 사용하여 그 정도를 −1에서 −0.714사이는 '매우 낮음', −0.714에서 −0.423사이는 '낮음', −0.423에서 −0.143사이는 '다소 낮음', −0.143에서 0.143사이는 '보통', 0.143에서 0.423사이는 '다소 높음', 0.423에서 0.714사이는 '높음', 0.714에서 1사이는 '매우 높음' 으로 구분하였다.

표 3.7: 특수법인 군집별 특성

군집	법인규모	수익률 성과	자본건전성	임직원성과
군집1	-0.00646	-0.01881	-0.05871	-0.22415
군집2	-0.7428	-0.02377	-0.36657	0.06038
군집3	0.764703	0.291646	0.091523	-0.2725
군집4	0.280753	-0.76364	-1.39448	0.754286
군집5	-0.06211	0.104538	0.405229	-0.24548
군집6	-0.20605	0.185665	0.829053	0.91212

요인점수는 각 변수의 원 데이터 값의 평균이 아니므로 크기를 나타내는 기준이 필요하다. 따라서 7점 리커드척도를 사용하여 그 정도를 -1에서 -0.714사이는 '매우 낮음', -0.714에서 -0.423사이는 '낮음', -0.423에서 -0.143사이는 '다소 낮음', -0.143에서 0.143사이는 '보통', 0.143에서 0.423사이는 '다소 높음', 0.423에서 0.714사이는 '높음', 0.714에서 1사이는 '매우 높음' 으로 구분하였다. 표 3.8은 군집별 업종구분에 대한 교차표이다. 군집별 특성을 보면 군집2와 군집6은 업종/가공의 특성을 전적으로 나타내고 있으며 나머지 군집은 지역별 특성을 강하게 보이고 있다.

표 3.8: 군집별 법인구분

군집	지역		업종/가공	
	빈도	열 %	빈도	열 %
군집1	29	40.23	0	0.0
군집2	0	0.00	11	50.00
군집3	11	15.28	1	4.55
군집4	8	11.11	1	4.55
군집5	24	33.33	0	0.00
군집6	0	0.00	9	40.91
Total	72	100	22	100

표 3.9의 군집별 경영상태에 대한 교차표를 보면 군집1, 2는 부실우려법인의 특성을 지니고 있고 군집4는 부실의 특성을 지니고 있다. 군집5, 6은 정상의 특성을 지니고 있고 군집3은 정상과 부실우려의 특성을 지니고 있음을 알 수 있다.

표 3.9: 군집별 경영상태

군집	정상		부실우려		부실	
	빈도	열%	빈도	열%	빈도	열%
군집1	0	0.00	29	65.91	0	0.00
군집2	0	0.00	11	25.00	0	0.00
군집3	8	19.51	4	9.09	0	0.00
군집4	0	0.00	0	0.00	9	100.00
군집5	24	58.54	0	0.00	0	0.00
군집6	9	21.95	0	0.00	0	0.00
Total	41	100	44	100	9	100

표 3.10: 군집별 지역구분

군집	서울/경기		강원		충청		전라		경상		제주	
	빈도	열%	빈도	열%	빈도	열%	빈도	열%	빈도	열%	빈도	열%
군집1	0	0.00	3	33.33	3	37.50	13	50.00	10	27.78	0	0.00
군집2	1	12.50	0	0.00	0	0.00	6	23.08	4	11.11	0	0.00
군집3	6	75.00	0	0.00	0	0.00	0	0.00	0	0.00	6	85.72
군집4	1	12.50	3	33.33	1	12.50	4	15.39	0	0.00	0	0.00
군집5	0	0.00	3	33.33	4	50.00	3	11.54	14	38.89	0	0.00
군집6	0	0.00	0	0.00	0	0.00	0	0.00	8	22.22	1	14.29
Total	8	100	9	100	8	100	26	100	36	100	7	100

표 3.10은 군집별 지역분포에 대한 교차표이다. 군집별 특성을 보면 서울/경기는 군집3에 많이 소속되어 있고 강원도는 군집1, 군집4, 군집5에 같은 비율로 소속되어 있다. 충청도는 군집5와 군집1에 대부분 소속되어 있고 전라도는 군집1에 절반이 소속되어 있다. 경상도는 군집5, 군집1, 군집6의 순으로 많이 소속되어 있고 제주도는 군집3에 주로 소속되어 있다. 서울/경기지역과 제주도지역을 제외하고 군집별 특이한 분포를 나타내고 있지 못하다.

3.3. 군집별 특성할당

이상과 같은 군집분석 결과는 각 군집별 특성을 정성적 표현으로 종합 정리하여 통합적 군집특성화를 제시하여야 한다. 따라서 다음의 표 3.11은 표 3.7부터 표 3.10까지의 군집별 산출결과를 특성화한 것이다. 결과적으로 총 6개의 군집별 특성에 따라 군집명을 할당하였다. 하지만 표현된 군집명칭은 군집을 위한 구분이지 각 군집별 특수법인의 특성을 정확히 반영한 것은 아니므로 군집특성을 반영하여 각 특수법인을 구분하여야 한다.

표 3.11: 특수법인 군집별 특성화

구분	특 성	군집명칭
군집 1	타군집에 비해 법인규모, 법인수익률 성과, 자본건전성 모두 보통(평균적)이고 임직원 성과는 다소 떨어지며 전부 지역구분 법인	중형 일반법인
군집 2	타군집에 비해 법인규모는 매우 작고 자본건전성은 다소 떨어지며 법인수익률과 임직원 성과는 보통(평균적)이고 전부 업종/가공구분 법인	소형 일반법인
군집 3	타군집에 비해 법인규모는 매우 크고 법인수익률은 다소 높고 자본건전성은 보통이나 임직원 성과는 다소 떨어지고 주로 지역구분 법인이고 서울/경기, 제주도 지역의 법인이 주류	대형 수익성 중점 법인
군집 4	타군집에 비해 법인규모는 다소 크고 법인수익률은 매우 낮으며 자본건전성은 다소 낮으나 임직원 성과는 매우 높으며 주로 지역구분 법인	중대형 부실 임직원 성과중점 법인
군집 5	타군집에 비해 법인규모는 보통(평균적)이고 법인수익률은 보통이상이고 자본건전성은 매우 높고 임직원 성과는 다소 떨어지는 지역구분 법인	중형 자본건전성 중점 법인
군집 6	타군집에 비해 법인규모는 다소 작으나 법인수익률은 다소 높고 자본건정성은 매우 좋으며 임직원 성과도 매우 높고 전부 업종/가공구분 법인	중소형 우량법인

4. 각 군집별 전략방향

각 군집별 특성을 바탕으로 특수법인 전략방향을 제시하면 표 4.1과 같다. 기본 전략방향은 강점요인과 약점요인으로 분류하고 강점요인은 강화하고 약점요인을 우선 개선하는 것을 기본 전략방향으로 정하였다. 또한 특성이 없는 요인에 대해서는 지속적 개선하는 방향으로 제시하였다. 실제 각 군집별 세부적 경영전략 실행방안을 제시하는 것이 이러한 연구의 최종 결론이 되어야 하나, 이를 위해서는 각 기업별 경영전략차원의 진단과 문제제기

가 필요하므로 본 연구에서는 기본적인 방향을 제시하는 것으로 결론을 맺기로 한다.

표 4.1: 군집별 전략방향

구분	강점	약점	전략방향
군집 1		임직원 성과	임직원 성과향상을 위한 우선 개선방안 마련 및 법인수익률과 자본건전성 상승을 위한 점진적 개선방안 마련
군집 2		자본건전성	자본건전성확보를 위한 중장기적 구조조정안 마련 및 법인수익률 개선과 임직원 성과 독려를 통한 자본건전성 확보 방안 마련
군집 3	법인수익률	임직원 성과	임직원 성과를 독려하여 법인수익률을 더욱 증진하고 점진적 자본건전성확보 방안 마련
군집 4	임직원성과	법인수익률 자본건전성	임직원성과를 지속적으로 유지하며 법인수익률 개선을 위한 방안 마련 및 자본건전성 확보를 위한 초단기적 구조조정방안 마련이 시급
군집 5	자본건전성	임직원성과	자본건전성이 높아 임직원 성과노력 측면이 다소 부족하므로 임직원에 대한 개혁마인드 조성 및 성과 독려 방안 마련
군집 6	자본건전성 법인수익률 임직원성과		경쟁우위를 유지하기 위한 지속적인 개혁 추진 방안 마련

5. 결론

본 논문은 94개 특수법인을 대상으로 하는 실증분석으로서 요인분석과 군집분석을 통하여 세분화 분석을 수행하여 경영전략 방향을 수립하는 사례연구이다. 본 연구에서는 특수법인의 경영지표관련 데이터를 사용하여 변수의 수가 상대적으로 많을 때, 세분화를 위한 대표적 분류속성을 요인분석을 통해 도출하였다. 이렇게 산출된 대표적 속성을 바탕으로 목표변수가 존재하지 않았을 때의 세분화 방법으로 군집분석을 통해 세분화 군집을 분류하고 각 군집별 특성화를 통해 군집의 특성을 도출하였다. 또한 각 군집별 특성에 맞는 전략방향을 제시하여 관리의 효율성을 높이는 방향을 제시하였다.

분류를 위한 대표속성 도출을 위해 요인분석의 주성분추출방법에 의하여 대표요인을 도출하였으며 요인회전은 하지 않았다. 또한 세분화를 위하여 사용된 방법으로는 범주형변수와 연속형변수 모두를 사용할 수 있는 장점을 갖고 있는 2단계 군집분석을 사용하였다. 도출된 군집의 특성화를 통해 중형 일반법인, 소형 일반법인, 대형 수익성 중점 법인, 중대형 부실 임직원 성과중점 법인, 중형 자본건전성 중점 법인, 중소형 우량법인으로 군집명칭

을 부여하였다. 또한 각 군집별 특성을 바탕으로 특수법인 전략방향을 제시하였다. 기본적 전략방향은 강점요인과 약점요인으로 분류하고 강점요인은 강화하고 약점요인을 먼저 개선하는 것을 기본 방향으로 정하였다. 또한 특성이 없는 요인에 대해서는 지속적으로 개선하는 방향을 제시하였다.

향후 이러한 전략방향을 바탕으로 실질적 실행방안이 개발되어야 하며, 또한 이러한 전략이 실행되었을 때의 관리효과에 대한 검증이 필요하다. 그러나 이러한 것들은 기업별 중장기적 관점의 접근이 필요하며 수행된 결과를 바탕으로 효율성에 따라 전략방향의 수정이 필요하다. 따라서 이러한 부분은 차후연구과제로 남긴다.

본 연구에서는 경영 효율성을 높이기 위해 일반적으로 사용되는 세분화를 적용하는데 있어 사용되는 변수가 많을 때를 고려하여 변수축약을 통해 대표속성을 도출하고, 이를 바탕으로 세분화를 수행하여 관리의 효율성을 높이는 방안을 실제 데이터를 분석하여 제시한 것에 큰 의의가 있다고 판단된다.

참고문헌

남영우, 성은영 (2001). 인자분석과 군집분석에 의한 세계도시의 유형화, <한국도시지리학회지>, **4**, 1–12.

박성현, 조신섭, 김성수 (2002). <한글 SPSS>, SPSS 아카데미.

이종상 (2002). 지역유형구분을 위한 요인점수의 군집분석, <대한국토도시계획학회지>, **37**, 191–199.

임종호 (1993). 석회암 풍화산물에 대한 군집분석과 인자분석, <지리학연구>, **22**, 73–90.

조용준, 허준 (2006). 고객가치모형별 마케팅전략: 백화점 화장품 고객을 중심으로, *Journal of the Korean Data Analysis Society*, **8**, 335–348.

최호현, 김선범 (2006). 요인분석과 군집분석을 이용한 용도지역의 특성과 유형분류, <한국도시지리학회지>, **9**, 127–136.

허명회, 이용구 (2004). K-평균 군집화의 재현성 평가와 응용, <응용통계연구>, **17**, 135–144.

Berry, M. J. A. and Gordon, L. (1997). *Data Mining Techniques*, John Wiley & Sons, U.S.A.

Chiu, T., Fang, D., Chen, J., Wang, Y. and Jeris, C. (2001). A robust and scalable clustering algorithm for mixed type attributes in large database environment, *Proceedings of the seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 263–268, San Francisco, California.

Hand, D. J. (1981). *Discrimination and Classification*, John Wiley & Sons, New York.

Kohonen, T. (1982). Self-organized formation of topologically collect feature maps, *Biological Cybernetics*, **43**, 59–69.

Kövesi, B., Boucher, J. M. and Saoudi, S. (2001). Stochastic K-means algorithm for vector quantization, *Pattern Recognition Letters*, **22**, 603–610.

Selim, S. Z. and Ismail, M. A. (1984), K-means type algorithms: a generalized convergence theorem and characterization of local optimality, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 81–87.

Zhang, T., Ramakrishnan, R. and Livny, M. (1996). BIRCH: an efficient data clustering method for very large databases, *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data*, 103–114, Montreal, Canada.

[2006년 9월 접수, 2006년 10월 채택]

A Strategy Through Segmentation Using Factor and Cluster Analysis: focusing on corporations having a special status

Yong Jun Cho[1)] Yeong-Hwa Kim[2)]

ABSTRACT

Corporations adopt a segmentation depends on the existence of target variables, in general. In this paper, for the case of no target variables, a strategy through segmentation is proposed for corporations having a special status based on the management index. In case of segmentation using cluster analysis, however, if one classify according to many variables then he will be in face of difficulties in characterizing. Therefore, after extracting representative factors by factor analysis, a segmentation method through 2 step cluster analysis is employed on the basis of these representative factors. As a result, six segmentation groups are found and the resulting strategy is proposed which strengthens prominent factors and makes up defective factors for each group.

Keywords: 2 step cluster, factor analysis, segmentation.

1) Senior Researcher, National Federation of Fisheries Cooperatives Fisheries Economic Institute, Seoul 138-730, Korea
E-mail: cyj66@chol.com
2) Associate Professor, Department of Statistics, Chung-Ang University, Seoul 156-756, Korea
E-mail: gogators@cau.ac.kr