

종합적과학기술력에 대한 통계적평가에서 나서는 몇가지 문제

박사 부교수 황 순 희

1. 서 론

경애하는 김정은동지께서는 다음과 같이 말씀하시였다.

《과학기술력은 국가의 가장 중요한 전략적자원이며 사회발전의 강력한 추동력입니다.》

나라의 전반적인 과학기술을 세계첨단수준에 올려세우고 과학기술의 주도적역할에 의하여 모든 부문을 급속히 발전시키자면 하루빨리 국가의 가장 중요한 전략적자원이며 사회발전의 강력한 추동력인 과학기술력을 비약적으로 높여나가야 한다.

종합적과학기술력을 세계선진수준에 올려세우자면 그에 대한 평가방법을 과학적으로 확립하고 그에 토대하여 다른 나라들과의 대비속에서 그 수준을 평가하고 그를 높여나가기 위한 과학기술적대책을 세워나가는것이 중요하다.

경쟁력일반에 대한 문제는 1980년대에 세계경제연단(WEF)과 국제관리발전연구소(IMD)를 비롯한 국제기구들과 일부 나라들에서 논의되기 시작하다가 그후 각양각색의 경쟁력에 관한 연구가 활발히 진행되면서 기업경쟁력과 산업경쟁력, 지역경쟁력, 국가경쟁력 등에 대하여 연구가 진행되였다.

세계적으로 해당 나라들의 과학기술발전수준과 잠재력, 과학기술의 활용능력에 대한 정성적 및 정량적평가가 세분화되어 진행되고있으며 그에 따라 평가내용과 방법, 수단도 부단히 발전하고있다.

선행연구에서는 국가 및 지역경쟁력의 관건적요소로서 과학기술경쟁력을 정의하였거나 경제발전의 요소로 취급한것이 보편적이였다.

어느 한 도서에서는 과학기술경쟁력은 국가의 과학기술의 총체, 과학기술실력이며 과학기술수준과 잠재력의 종합적구현이라고 정의하면서 그것은 국가나 지역의 총체적인 과학기술수준과 잠재력의 발전능력을 반영할뿐아니라 국제적인 경쟁력의 중요구성부분, 관건적인 요인으로 되어 경쟁에서 결정적인 작용을 놀며 인류사회가 지속적인 발전을 이룩하는데서 중요한 추동작용을 한다고 해설하였다.

그리고 과학기술경쟁력은 일정한 과학기술발전환경에서 한개 나라 혹은 지역이 과학연구와 기술개발, 기술갱신과 기술이전 등 과학기술활동을 통하여 반영되는 해당 국가 혹은 지역의 과학기술투입과 산출, 과학기술과 경제의 일체화정도 및 과학기술잠재력의 종합적인 수준을 반영한다고 하였다.

이러한 과학기술경쟁력은 경제발전을 추진시키고 경제력을 강화하며 사회의 지속적 발전을 추동하는 과학기술의 능력으로 된다는것이다.

선행연구에서는 종합적과학기술력에 대하여서도 연구되였다.

어느 한 도서에서는 종합적과학기술력은 나라의 모든 과학기술력량과 수단, 과학기술논문과 저작, 발명, 특허를 비롯하여 이룩된 과학기술성과, 과학기술투자 및 기술수출

입액, 첨단기술산업액 등 과학기술과 련관되어있는 모든 부문, 분야의 량적 및 질적능력의 총체라고 정의하였다.

도서에서는 종합적과학기술력이 경제발전을 추동하는 요인으로만이 아니라 경제를 포괄한 사회발전의 추동력으로 된다는데 대하여서와 고정무게에 의한 종합적과학기술력 평가와 가변무게에 의한 종합적과학기술력평가방법을 제기하였다.

선행연구에서는 종합적과학기술력에 대한 정확한 정의와 그것을 과학적으로 분석평가하기 위한 보다 편리한 방법론을 제기하지 못하였다.

론문에서는 종합적과학기술력에 대한 정의와 구성요소, 평가를 위한 지표체계를 설정한데 기초하여 주성분분석법과 무리분석법을 리용한 종합적과학기술력평가방법에 대하여 해설하였다.

2. 본 론

2.1. 종합적과학기술력평가지표체계

종합적과학기술력은 한 나라범위에서 규정되는 과학기술의 힘이다. 다시말하여 과학기술지식을 소유한 인재와 그에 의하여 창조되고 축적된 과학기술지식과 과학기술수단의 총체, 그것들이 결합되어 발휘되는 힘이 곧 과학기술력이다.

종합적과학기술력은 과학기술연구개발능력과 과학기술활용능력 그리고 과학기술지원능력을 포괄한다.

과학기술연구개발능력은 과학기술성과나 결과를 이룩할수 있도록 일정한 국가 혹은 지역적범위에서 축적되어있는 과학기술적잠재력이라고 할수 있다.

과학기술활용능력은 과학연구와 기술개발, 기술이전 등 과학기술활동을 통하여 이루어지는 과학기술활동의 결과물이라고 할수 있다.

과학기술지원능력은 과학기술연구개발능력과 활용능력을 높여주고 공고히 해주는 기초적인 능력이라고 할수 있다.

과학기술연구개발능력과 활용능력은 그것을 높여주고 뒤받침해주는 기초적인 능력에 의하여서도 좌우된다. 과학기술투자나 교육수준, 과학기술정책, 사회적환경 등은 과학기술연구개발능력과 활용능력제고를 위한 중요한 요인으로 된다.

과학기술력통계지표체계는 연구목적과 과학기술력의 구성요소에 따라 여러 측면을 포괄하여 설정할수 있다.

과학기술력통계지표체계는 과학기술연구개발능력을 특징짓는 지표, 과학기술활용능력을 특징짓는 지표, 과학기술지원능력을 특징짓는 지표들로 설정할수 있다.

과학기술연구개발능력을 특징짓는 지표에는 과학기술력량과 과학기술수단과 관련한 지표들이 속한다.

과학기술력량과 관련한 통계지표는 절대값지표로서 연구자(연구일군)수, 연구개발인원총수, 학위학직소유자수(박사수, 석사수)와 상대값지표로서 인구 만명당 연구개발인원총수, 인구 만명당 연구자수, 인구 만명당 학위학직소유자수(박사수, 석사수) 등으로 설정할수 있다.

과학기술수단과 관련한 지표로는 과학기술연구에 종사하는 연구 및 개발단위수, 현물형태별 과학기술수단의 수, 과학기술정보량과 같은 현물지표와 함께 과학기술부문의

고정재산총액과 그것이 전체 고정재산총액에서 차지하는 비중과 같은 화폐표시지표가 될 수 있다.

과학기술활용능력지표로는 연구개발 및 기술혁신활동을 통하여 창조된 직접적인 과학기술성과를 반영하는 지표와 과학기술성과가 반영된 경제적결과를 나타내는 지표로 설정할 수 있다.

직접적인 과학기술성과를 나타내는 지표에는 과학기술논문수, 발명 및 특허신청건수, 기술수출액과 상대값지표로서 연구일군 한사람당 발명 및 특허권수, 연구일군 한사람당 과학기술논문수 등이 속한다.

과학기술성과가 반영된 경제적결과를 나타내는 지표에는 첨단기술제품생산액과 그것이 국내총생산액에서 차지하는 비중, 경제발전에 대한 과학기술의 기여율과 같은 지표들이 포함될 수 있다.

과학기술지원능력은 과학기술력의 제고를 추동하는 전제조건들이다.

과학기술력의 제고를 추동하는 전제조건은 과학기술제도, 과학기술행정사업기구, 과학기술발전정책, 지적소유권보호체제 등 과학기술경쟁력을 형성하고 강화하는 사회적환경과 과학기술에 대한 자금지출, 교육발전수준과 같은것이다.

이러한 지표에는 과학기술력을 강화할수 있는 사회적환경과 관련한 지표와 과학기술 부문에 대한 투자액, 과학기술발전사업비, 교육사업비, 인구 만명당 대학생수와 졸업생수, 교육의 물질적토대와 같은 교육발전과 관련한 지표들이 포함될 수 있다.

2.2. 종합적과학기술력평가방법

2.2.1. 주성분분석을 리용한 종합적과학기술력평가방법

주성분분석법은 다변량자료를 요약하는 방법으로서 많은 변량의 값을 정보의 손실이 적도록 1개 또는 몇개의 무상관인 종합적지표로서 대표하는 방법이다.

주성분분석을 리용한 종합적과학기술력평가는 다음과 같은 절차를 거쳐 진행할 수 있다.

1) 종합적과학기술력평가를 위한 지표설정 및 계산값의 수집

표 1. 종합적과학기술력평가지표

번호	지 표	번호	지 표
1	X_1 : 만명당 연구자수(명)	7	X_7 : 첨단기술제품수입액
2	X_2 : 과학기술발전사업비가 국가예산 지출에서 차지하는 비중	8	X_8 : 연구자 1인당 과학기술론 총발표건수(건)
3	X_3 : 인구 1인당 연구개발비	9	X_9 : 세계적인 특허개발수에서 차지하는 비율
4	X_4 : 연구자 1인당 연구개발비	10	X_{10} : 특허개발수(건)
5	X_5 : 인구 백만명당 석사수(명)	11	X_{11} : 기술무역수지
6	X_6 : 인구 백만명당 박사수(명)	12	X_{12} : 첨단기술제품수출액

표 2. 종합적과학기술력 평가지표값

나라 지표	1	2	3	4	5	6	7
X_1	41.4	40.2	40.1	70	9.8	38.3	71.9
X_2	2.85	2.73	1.63	3.47	2.08	2.23	0.52
X_3	1 230.1	1 44.0	621.8	1 258.5	247.3	837.9	1.01
X_4	184. 4	344. 7	87.8	171. 1	141. 6	153. 0	0.151
X_5	2 129	2 421	3 884	615	1000	2 027	2 685
X_6	330	247	326	131	200	176	258
X_7	214 223	499 837	125 577	157 680	516 802	136 486	158 600
X_8	0.42	0.17	0.056	0.087	0.06	0.42	0.98
X_9	10.11	27.03	3.28	29.55	3.3	4.6	3
X_{10}	5 465	14 606	1 770	15 970	1 785	2 484	1 000
X_{11}	12 427.6	38 900	3 979.3	28 868	10 200	25 200	10 000
X_{12}	256 884	382 994	110 51	148 943	651 588	144 154	120 000

2) 자료행렬의 표준화

자료행렬을 표준화하게 되는것은 우의 통계자료의 측정단위가 서로 다르기때문이다. 그러므로 변수의 평균값이 1, 표준편차가 1인 분포에 따르도록 표준화하게 된다.

표 3. 표준화된 자료

나라 지표	1	2	3	4	5	6	7
X_1	-0.147 4	-0.204 1	-0.208 8	1.203 6	-1.6	-0.28	1.29
X_2	0.673 7	0.547 4	-0.610 5	1.326 3	-0.136 8	0.021 1	-1.778 9
X_3	0.777 5	1.169 5	-0.337 2	0.829 9	-1.025 3	0.058 2	-1.472 4
X_4	0.281 9	1.823 1	-0.646 9	0.154	-0.129 6	-0.020 0	-1.462 2
X_5	0.018 7	0.288 4	1.639 2	-1.379 2	-1.023 7	-0.075 4	0.532 1
X_6	1.229	0.116 8	1.175 4	-1.437 5	-0.512 9	-0.834 5	0.264 2
X_7	-0.255 7	1.395 1	-0.768	-0.582 5	1.493 1	-0.705 0	-0.577 2
X_8	0.125	-0.656 3	0.562 5	-0.915 6	-1.000	0.125 0	1.875 0
X_9	-0.123	1.321 9	-0.706 2	1.537 1	-0.704 5	-0.593 5	-0.730 1
X_{10}	-0.107 5	1.318 4	-0.683 8	1.531 1	-0.681 5	-0.572 5	-0.803 9
X_{11}	-0.481 6	1.614 1	-1.150 4	0.819 9	-0.657 9	0.529 5	-0.673 8
X_{12}	-0.011 8	0.623 6	-0.751 7	-0.555 7	1.977	-0.579 8	-0.701 6

3) 상관행렬의 구성과 그 고유값, 고유벡토르계산

$$\begin{aligned}
 R &= \begin{pmatrix} 1.0000 & -0.0955 & 0.0744 & -0.2980 & 0.0475 & -0.1325 & -0.6281 & 0.4804 & 0.3536 & 0.3283 & 0.1693 & -0.7664 \\ -0.0955 & 1.0000 & 0.8671 & 0.7300 & -0.5742 & -0.3562 & 0.1815 & -0.8246 & 0.7443 & 0.7637 & 0.6047 & 0.1692 \\ 0.0744 & 0.8671 & 1.0000 & 0.8306 & -0.1807 & -0.0683 & 0.0813 & -0.5522 & 0.7877 & 0.8017 & 0.7135 & -0.0857 \\ -0.2980 & 0.7300 & 0.8306 & 1.0000 & -0.2307 & -0.1209 & 0.5905 & -0.7058 & 0.7062 & 0.7211 & 0.7895 & 0.3859 \\ 0.0475 & -0.5742 & -0.1807 & -0.2307 & 1.0000 & 0.7817 & -0.3065 & 0.6606 & -0.4197 & -0.4232 & -0.3805 & -0.4396 \\ -0.1325 & -0.3562 & -0.0683 & -0.1209 & 0.7817 & 1.0000 & -0.0911 & 0.4877 & -0.3952 & -0.3946 & -0.5360 & -0.1236 \\ -0.6281 & 0.1815 & 0.0813 & 0.5905 & -0.3065 & -0.0911 & 1.0000 & -0.5787 & 0.2183 & 0.2250 & 0.3024 & 0.9233 \\ 0.4804 & -0.8246 & -0.552 & -0.7058 & 0.6606 & 0.4877 & -0.5787 & 1.0000 & -0.5647 & -0.5874 & -0.5040 & -0.6084 \\ 0.3536 & 0.7443 & 0.7877 & 0.7062 & -0.4197 & -0.3952 & 0.2183 & -0.5647 & 1.0000 & 0.9994 & 0.8174 & -0.0057 \\ 0.3283 & 0.7637 & 0.8017 & 0.7211 & -0.4232 & -0.3946 & 0.2250 & -0.5874 & 0.9994 & 1.0000 & 0.8181 & 0.0059 \\ 0.1693 & 0.6047 & 0.7135 & 0.7895 & -0.3805 & -0.5360 & 0.3024 & -0.5040 & 0.8174 & 0.8181 & 1.0000 & 0.0477 \\ -0.7664 & 0.1692 & -0.0857 & 0.3859 & -0.4396 & -0.1236 & 0.9233 & -0.6084 & -0.0057 & 0.0059 & 0.0477 & 1.0000 \end{pmatrix} \\
\lambda &= \begin{pmatrix} -0.000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.000 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.158 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.417 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.770 & 7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.594 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2.898 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.173 & 5 \end{pmatrix} \\
L &= \begin{pmatrix} 0.1152 & 0.5223 & -0.1117 & -0.3954 & -0.0628 & -0.0460 & -0.1424 & 0.3608 & -0.2823 & 0.1811 & -0.5240 & -0.0391 \\ -0.3169 & 0.2687 & 0.0638 & -0.0560 & 0.1923 & 0.6069 & -0.1058 & 0.1056 & 0.5086 & -0.0413 & -0.0631 & 0.3580 \\ 0.1803 & -0.4259 & -0.3702 & -0.4331 & -0.2167 & -0.1356 & -0.2536 & -0.0723 & 0.2641 & -0.3355 & -0.2042 & 0.3204 \\ 0.6282 & 0.3776 & 0.1997 & 0.2628 & 0.0568 & 0.0226 & -0.1880 & -0.2341 & -0.1095 & -0.3477 & 0.0675 & 0.3501 \\ -0.0790 & 0.1443 & -0.0459 & -0.2332 & -0.1804 & 0.3102 & 0.5640 & -0.2529 & -0.1604 & -0.5535 & -0.1071 & -0.2490 \\ -0.3125 & 0.1054 & 0.1388 & 0.1630 & 0.0881 & -0.3448 & -0.2600 & 0.4452 & 0.0771 & -0.6362 & 0.0344 & -0.2027 \\ -0.0302 & -0.0756 & -0.3037 & -0.2869 & 0.5563 & 0.1314 & -0.0128 & 0.1676 & -0.4684 & -0.1043 & 0.4418 & 0.1956 \\ 0.0610 & -0.3343 & 0.0186 & 0.2017 & -0.0410 & 0.5587 & -0.5064 & -0.0109 & -0.3081 & -0.0771 & -0.2247 & -0.3491 \\ 0.0460 & -0.3991 & 0.6614 & -0.1202 & 0.0425 & 0.0352 & 0.2653 & 0.3056 & -0.2012 & -0.0316 & -0.2366 & 0.3499 \\ -0.0748 & -0.0608 & -0.4877 & 0.5911 & -0.1641 & 0.0540 & 0.2833 & 0.2902 & -0.1731 & -0.0392 & -0.2280 & 0.3545 \\ -0.5819 & 0.0902 & 0.0987 & 0.0106 & -0.0732 & -0.1853 & -0.2420 & -0.5172 & -0.3729 & -0.0087 & -0.1594 & 0.3363 \\ -0.0655 & 0.1022 & 0.0865 & -0.1265 & -0.7234 & 0.1664 & -0.1322 & 0.2513 & -0.1562 & 0.0391 & 0.5325 & 0.1438 \end{pmatrix}
 \end{aligned}$$

4) 주성분의 개수결정

주성분분석에 의하여 새롭게 설정되는 지표들의 개수는 지표들의 루적기여률이 80~90% 이상을 보장하도록 결정한다.

$$\frac{\lambda_1 + \lambda_2 + \lambda_3}{p} \times 100 = \frac{6.1735 + 2.8982 + 1.594}{11} \times 100 = 97(\%)$$

루적기여률이 97%이므로 주성분의 개수를 3개로 취한다.

5) 지표행렬(새 인자행렬)의 구성

주성분의 개수를 3개로 취했으므로 계산에 리용되는 고유값과 고유벡토르는 다음과 같다.

$$\lambda_1 = 6.1735, \lambda_2 = 2.8982, \lambda_3 = 1.594$$

$$L = \begin{pmatrix} -0.0391 & -0.5240 & 0.1811 \\ 0.3580 & -0.0631 & -0.0413 \\ 0.3204 & -0.2042 & -0.3355 \\ 0.3501 & 0.0675 & -0.3477 \\ -0.2490 & -0.1071 & -0.5535 \\ -0.2027 & 0.0344 & -0.6362 \\ 0.1956 & 0.4418 & -0.1043 \\ -0.3491 & -0.2247 & -0.0771 \\ 0.3499 & -0.2366 & -0.0316 \\ 0.3545 & -0.2280 & -0.0392 \\ 0.3363 & -0.1594 & -0.0087 \\ 0.1438 & 0.5325 & 0.0391 \end{pmatrix}$$

따라서 $Z = XL$ 에 의하여 매개 나라들에 대한 지표행렬(새 인자행렬)을 계산하면 다음과 같다.

표 4. 새 인자행렬

나라 지표	1	2	3	4	5	6	7
Z_1	0.0039	3.1208	-2.444	2.9227	0.0077	-0.2187	-3.4189
Z_2	-0.0772	-0.0875	-0.0292	-1.7502	2.7352	-0.1249	-0.7234
Z_3	-1.1827	-1.1903	-1.6337	1.2289	2.0267	0.3002	0.4494

6) 주성분부하행렬

$$A = LA^{1/2} = \begin{pmatrix} -0.0391 & -0.5240 & 0.1811 \\ 0.3580 & -0.0631 & -0.0413 \\ 0.3204 & -0.2042 & -0.3355 \\ 0.3501 & 0.0675 & -0.3477 \\ -0.2490 & -0.1071 & -0.5535 \\ -0.2027 & 0.0344 & -0.6362 \\ 0.1956 & 0.4418 & -0.1043 \\ -0.3491 & -0.2247 & -0.0771 \\ 0.3499 & -0.2366 & -0.0316 \\ 0.3545 & -0.2280 & -0.0392 \\ 0.3363 & -0.1594 & -0.0087 \\ 0.1438 & 0.5325 & 0.0391 \end{pmatrix} \begin{pmatrix} \sqrt{6.1735} & & \\ & \sqrt{2.8982} & \\ & & \sqrt{1.594} \end{pmatrix} = \begin{pmatrix} -0.0971 & -0.8921 & 0.2286 \\ 0.8895 & -0.1074 & -0.0521 \\ 0.7961 & -0.3476 & -0.4236 \\ 0.8699 & 0.1149 & -0.4390 \\ -0.6187 & -0.1823 & -0.6988 \\ -0.5036 & 0.0586 & -0.8032 \\ 0.4860 & 0.7521 & -0.1317 \\ -0.8674 & -0.3825 & -0.0973 \\ 0.8694 & -0.4028 & -0.0399 \\ 0.8808 & -0.3881 & -0.0495 \\ 0.8356 & -0.2714 & -0.0110 \\ 0.3573 & 0.9065 & 0.0494 \end{pmatrix}$$

7) 결과평가

합성지표 $Z = \sum \frac{\lambda_k}{\sum \lambda_k} Z_k$ 에 의한 계산결과는 다음과 같다.

표 5. 주성분분석에 의한 평가결과

나라 지표	1	2	3	4	5	6	7
평가값	-0.193	1.66	-1.725	1.336 5	1.245 2	-0.161 6	-2.162 1
순위	5	1	6	2	3	4	7

평가결과를 보면 나라 2가 가장 앞선 순위를 차지하고있으며 그다음은 나라 4, 나라 5의 순위이다.

주성분부하행렬에서 보는것처럼 제1주성분은 연구개발비와 특허개발측면, 제2주성분은 첨단기술제품수출측면, 제3주성분은 연구자측면을 대표한다. 따라서 종합적과학기술력을 높이려면 과학기술부문에 대한 투자를 늘이고 특허개발과 첨단기술산업을 활성화하는 한편 연구개발력량을 보강하여야 한다는것을 보여주고있다.

과학기술력이 앞선것으로 평가된 나라 2와 나라 4는 연구개발비와 특허개발에서 다른 나라들보다 우세하다. 특히 나라 2는 연구자 한사람당 연구개발비가, 나라 4는 특허개발이 특별히 우세하며 나라 5는 첨단기술제품무역액이 다른 나라들에 비하여 특별히 크다.

주성분분석을 통하여 해당 나라들의 종합적과학기술력이 어느 수준에 있는가 하는것을 평가할수 있으며 그것을 높여나가려면 어디에 힘을 집중하여야 하겠는가 하는 과학적인 근거를 마련함으로써 과학기술전략을 세우는데 이바지할수 있다.

2.2.2. 무리분석법을 리용한 종합적과학기술력평가방법

일반적으로 무리분석법은 관측되는 많은 수의 현상, 대상들을 그것들의 류사성에 의하여 분류하는 통계적인 방법으로서 다양한 사회경제적현상들을 그 특성에 따라 분류하고 그에 기초하여 여러가지 결론과 대책을 세우는데 널리 리용된다.

무리분석법은 어떤 관찰집단을 그 특성이 서로 차이나는 여러개의 세분된 집단들로 분류하는 방법이다. 이때 해당한 개개의 세부집단에는 동종의 특성을 가지는 대상들이 포함되며 서로 다른 세부집단에 속하는 대상들은 그 특성에서 서로 차이나게 된다.

무리분석법의 특징은 분류되는 무리의 개수가 분석전에 규정되는것이 아니라 주어진 자료에 따라 그 수가 결정된다는것이다.

무리분석에서 적용되는 기본방법은 관찰값들사이의 거리 즉 통계수자들사이의 거리를 계산하고 그것을 기준으로 하여 집단을 형성하는 방법이다.

무리분석에서 주어진 관찰값(통계수자자료)들은 그것들사이의 비류사성(또는 류사성) 혹은 거리에 기초하여 분류된다. 이러한 거리들(류사성)은 단일 또는 다중파라미터들에 기초할수 있는데 여기서 매 파라미터는 관찰자료들을 일정한 집단으로 나누기 위한 규칙이나 조건들을 반영한다.

다차원공간에서 관찰값들사이의 거리를 계산하는 가장 단순한 방법은 유클리드거리를 계산하는것이다.

유클리드거리는 두 관찰값이나 대상들사이의 기하학적거리이다. 이것은 두개의 표본사이의 거리를 계산하는 가장 일반적인 방법이다.

이 방법은 표본의 크기에서의 변화에 기초하여 직접적으로 두개의 표본사이의 차이를 계산한다. 유클리드거리는 적절하게 표준화되어있는 자료묶음에서 보통 리용된다. 이

거리는 두 관찰값들사이의 매 차이의 두제곱합을 두제곱뿌리하는 방법으로 계산된다.

$$de_{ij} = \sqrt{\sum_{k=1}^n (X_{ik} - X_{jk})^2}$$

여기서 de_{ij} 는 두 관찰값들사이의 거리, X_{ik} 는 k 째 변수에 대한 i 째 관찰값의 크기, X_{jk} 는 k 째 변수에 대한 j 째 관찰값의 크기, n 은 변수의 개수이다.

유클리드거리가 작을수록 관찰값들은 보다 유사하다고 본다.

그러나 이 지표는 변수들의 측정단위에 따라 달라진다. 변수들이 서로 다른 척도에 기초하여 측정된다면 큰 값을 가지는 변수들은 작은 값을 가지는 변수들보다 거리를 더 크게 만들게 된다. 그러므로 변수들이 서로 다른 척도로 측정되어있는 경우에 변수값들을 표준화하여 계산하는것이 중요하다.

종합적과학기술력에 대한 평가를 다음과 같은 방법으로 진행할수 있다.

표 6. 종합적과학기술력 평가지표값(수자는 가상자료임)

나라	연구사수(만명)	첨단기술제품 생산액(만원)	과학기술 논문수(건)	국내총생산액가 운데서 특허액 비중(%)	과학기술투자 (원)
1	7 760	0.764 653 339	29.757 135 66	0.000 392 707	234 457 666.7
2	14 920	0.625 863 298	34.222 803 16	0.004 357 402	3 047 520 66
3	294 000	8.686 933 066	24.200 950 35	0.002 017 317	6 368 083 3
4	2 620	2.181 898 774	37.782 566 66	0.000 267 403	66 733 333.33
5	1 493 000	11.925 758 11	28.367 409 19	0.014 185 618	14 144 907 3
6	38 870	11.636 169 28	23.132 399 87	0.001 362 005	4 698 158 66
7	9 360	0.983 053 092	21.525 008 37	0.007 120 474	676 375 333.4
8	10	0.027 504 417	26.993 679 54	0.015 003 702	567 581 666.7
9	86 300	3.134 666 468	34.091 903 18	0.000 419 816	1 097 462 67
10	571960	10.363 695 33	22.877 777 33	0.001 436 87	11 198 605 0
11	6780	4.933 155 218	15.186 654 83	0.001 452 368	316 928 333.4
12	21 850	3.276 536 945	31.057 318 12	0.001 187 415	325 438 666.7
13	113 510	0.924 868 482	25.815 667 17	0.000 473 425	93 023 333.33
14	5 195 220	13.197 277 86	18.327 619 43	0.001 191 776	10 618 439 3
15	39 270	8.150 133 519	34.093 228 66	0.001 302 73	3 310 582 00
16	100 940	0.177 519 817	21.197 291 33	0.000 262 411	59 414 666.67
17	199 160	2.540 102 343	17.719 631 74	0.000 461 465	142 607 666.7
18	3 101 340	16.223 800 55	23.239 706 88	0.011 575 486	26 671 883 3
19	162 310	3.397 556 992	20.484 711 12	0.000 548 419	2 026 196 33

우의 자료를 가지고 여러가지 통계분석프로그램들을 리용하여 무리분석을 손쉽게 진행할수 있다.

표 6의 자료를 통계분석프로그램에 넣어 무리분석을 진행하면 다음과 같은 무리가 얻어진다. 이때 집단의 유사성을 밝히는데 리용된 거리계산방법은 유클리드거리계산방식이다.

무리분석의 결과에 따라 19개의 나라들을 구분하면 크게 3개의 집단으로 나누어지는데 나라 18이 하나의 집단, 나라 5, 10, 14가 또 하나의 집단 그리고 나머지가 다른 하나의 집단을 형성한다. 다시말하여 3개의 큰 무리가 형성된다.

매 무리속에 속하여있는 집단들은 서로 동일한 특성을 가지고있는 집단들로서 종합적과학기술력수준이 유사한 집단으로 된다.

무리분석에 의한 종합적과학기술력평가에서는 여기에 리용할수 있는 지표(관찰값)들을 무엇으로 설정하겠는가 하는 문제와 함께 얻어진 결과값을 가지고 과학기술력에 대한 어떤 평가와 결론을 내리겠는가 하는것이 기본으로 된다.

만일 두번째 무리에 속하여있는 나라 5가 세계적으로 볼 때 인구 만명당 연구일근수가 대단히 많고 과학기술발전수준이 높은것으로 하여 세계적으로 공인되고있는 나라라고 하자. 이 두번째 무리속에 속하여있는 나라들은 다 동일한 특성을 가진 나라들로서 종합적과학기술력이 높은 집단이라고 볼수 있다.

종합적과학기술력이 높은 두번째 집단에 포함되는 나라들의 지표값을 구체적으로 보면 국내총생산액가운데서 특허수입액이 차지하는 비중은 중간정도이고 과학기술논문수는 작았으나 연구사수와 과학기술에 대한 투자, 첨단기술제품생산액은 높은 나라들이다.

이로부터 종합적과학기술력을 높이기 위하여서는 과학기술력량을 강화하고 과학기술 부문에 대한 투자를 늘여 첨단기술제품을 많이 생산해내는데 선차적인 힘을 넣어야 한다는것을 알수 있다.

이와 같이 무리분석결과는 해당 나라들에서 어떤 대책을 세우는데 선차적인 힘을 넣으며 매 지표들을 어떻게 조절할 때 최대의 효과를 나타낼수 있는가에 근거하여 올바른 전략을 세우고 집행해나갈수 있게 한다.

무리분석법은 그 평가과정의 매우 간단한 우점이 있지만 그것을 과학기술력평가에 리용하는 경우에 일련의 제한성이 있다. 그것은 무리분석자체가 무리를 동일한 특성을 가지는 집단들로 나누어놓는 방법인것만큼 나누어진 집단을 놓고는 해당 집단의 수준이 높은지 낮은지 하는것을 알수 없다는것이다.

그러나 과학기술발전의 측면에서 널리 소개된 나라를 알고있다면 다시말하여 그 나라의 수준이 어느 수준(상급, 중급, 하급)에 있다는것을 알수 있다면 그러한 나라가 속한 무리의 수준을 판단할수 있으며 그 집단의 특성을 밝혀낼수 있다. 즉 무리분석법은 일정한 전제하에서만 과학기술력평가에 적용할수 있다.

3. 결 론

주성분분석법과 무리분석법을 리용한 종합적과학기술력에 대한 평가는 종합적과학기술력을 제고하기 위한 대책들을 과학적으로 세울수 있는 방도를 찾을수 있게 하는것으로 하여 다른 분석수법들에 비하여 우월하다고 평가할수 있다.

종합적과학기술력평가를 위하여서는 측정단위가 다른 자료들의 변환에 대한 여러가지 방법가운데서 가장 합리적인 방법을 선택하는 문제와 과학적인 자료들을 수집하기 위한 방법론적문제들을 비롯하여 해결하여야 할 문제들도 존재한다.

우리는 종합적과학기술력에 대한 통계적평가방법론을 더욱 완성하여 종합적과학기술력을 강화하기 위한 과학적인 방법론을 확립함으로써 나라의 자립적발전능력을 확대강화하기 위한 사업에서 획기적인 전진을 이룩하여야 한다.

실마리어 종합적과학기술력, 과학기술력평가방법