
국민 체력 주성분 및 인자 분석

학과 : 수학과 (통계학과 복수전공)

학번 : 2017540035

이름 : 이혁종

차례

1. 서론

1.1 연구목적

1.2 문헌연구

1.3 데이터 설명

1.4 분석 방법

1.5 결과 활용 및 기대 효과

2. 본론

2.1 분석 방법 소개

2.2 데이터 분석 및 결과 설명

3. 결론

3.1 분석 결과 요약

3.2 분석의 장점 및 한계점 설명

3.3 추가 연구사항 제안

4. 참고문헌

1.1 연구 목적

1.2 문헌 연구

인간이 살아가는데 공부, 명예, 부도 중요하지만 모두 건강이 뒷받침해 줘야 한다. 여기서의 건강은 크게 보면 암과 같은 질병에 걸리지 않는 것을 의미한다. 이런 질병을 예방하기 위해서는 면역력이 높아야 되며 면역력을 높이기 위해서는 많은 요인이 있지만, 그중 체력이 중요한 역할을 한다. 많은 연구가[2], [3], [4] 이를 뒷받침 한다. 즉, 어느 정도의 체력이 있어야 건강한 삶을 유지 할 수 있으며 체력이 부족하면 건강함을 잃는다.

하지만 [1]에 의하면 종로로학원하늘교육이 분석한 결과 고등학생의 경우 1등급(80~100점) 비율이 2016년 5.9%에서 2018년 5.5%로 떨어졌다. 초등학생도 같은 기간 6.1%에서 5.9%로 하락했다. 단 중학생은 9.6%로 변화가 없었다. 바로 아래인 2등급(60~79점)도 초등학생은 42.3% → 40.3%, 중학생은 45.2% → 43.6%, 고등학생은 38.1% → 35.7%로 떨어졌다. 반면 최하위 등급인 4등급(20~39점)과 5등급(0~19점) 비율은 오히려 증가했다. 4등급의 경우 초등학생은 5.7% → 7.4%, 중학생은 6.9% → 8.2%, 고등학생은 11.8% → 13.6%로 늘었다. 5등급도 초중고교생 모두 0.1~0.4%포인트씩 증가했다. 즉 학생들의 운동량이 과거에 비해 절대적으로 부족하여 체력이 떨어졌고, [1], [6] 에 의하면 주 원인은 컴퓨터와 스마트폰 사용시간의 증가로 인한 운동시간 부족이라고 생각되고 있다.

따라서 [7]과 같이 청소년들의 체력문제는 심각한 문제이므로 이를 해결 하고자 하는 노력이 늘어나고 있다.

이 문제는 비단 청소년 집단만의 문제는 아니다. 따라서 본 보고서에는 대한민국 국민 체력을 성별/연령별로 분석하고 인자 분석을 통해 각 측정 항목간의 관계를 밝힘으로 국민 체력 증진에 도움을 주고자 한다.

1.3 데이터 설명

데이터는 한국스포츠정책과학원에 (<https://www.sports.re.kr/>) 등록된 2007년부터 2017년 까지 2년마다 시행된 국민 체력 실태 조사이다.

측변 변수는 연도별로 살짝 다르나 신체에 대한 변수와 체력에 대한 변수의 틀은 같다.

다음은 연도별 관측 표본수이다.

연령대 (대)	0	10	20	30	40	50	60
표본수 (명)	1,169	804	3,778	4,035	4,114	3,872	948
연도	2007	2009	2013	2015	2017	-	-
표본수 (명)	4,632	3,783	3,674	4,034	4,292	-	-
성별	남	여	-	-	-	-	-
표본수 (명)	8,654	8,087	-	-	-	-	-

변수명

2017	연령(만) 성별 신장 체중 BMI 윗몸일으키기 악력(D) 악력(ND) 제자리멀리뛰기 20m 왕복오래달리기 좌전굴 10m 왕복달리기 체지방률						
2015	연령(만) 성별 신장 체중 BMI 윗몸일으키기 악력(D) 악력(ND) 제자리멀리뛰기 20m 왕복오래달리기 50m 달리기 Skyndex 체지방률 이완혈압 수축혈압						
2013	나이 성별 신장 체중 BMI 윗몸일으키기 악력(D) 악력(ND) 제자리멀리뛰기 20m 왕복오래달리기 50m 달리기 Skyndex 체지방률						
2009	연령 성별 신장 체중 BMI계산 윗몸일으키기 악력 제자리멀리 왕복달리기20 달리기50 장좌체전굴 이완혈압 수축혈압						
2007	연령 성별 신장 체중 BMI 달리기50 멀리뛰기 팔굽혀펴 윗몸일으 이완혈압 수축혈압체지방률						

1.4 분석방법

1. 주성분 분석

p 개 변수로 이루어진 다변량 데이터를 p 보다 적은 n 개의 변수로 다시 설명하고자 하는 방법이다. 이때 새로운 변수 n 개는 사실 새로운 것이 아니고 p 개의 변수들의 선형결합으로 이루어진다. 이렇게 차원 축소를 통하여 다변량 데이터를 보다 적은 변수로 설명할 수 있으며 추가적인 분석 시 계산량을 줄일 수 있고 새로운 해석이 가능하다.

2. 인자분석

주성분 분석과는 다른 방법으로 인자분석은 데이터를 관측 불가능한 인자로 설명하는 것이 목적이다. 즉, 실제 데이터가 아닌 변수들 간의 내재하고 있는 공통의 구조를 파악한다. 주성분 분석과 큰 차이점은 주성분 분석에서는 관측된 변수들의 선형결합으로 주성분이 표현되고 반대로 인자분석에서는 공통인자들의 선형결합으로 관측된 변수들이 표현된다는 점이다.

1.5 결과 활용 및 기대 효과

기초 체력의 문제가 심각한 요즘 문제를 해결을 위해 국가적으로 노력하고 있고 각 지자체에서 특히 심혈을 기울이고 있는 만큼 해당 정책의 방향성과 근거를 제시할 수 있다.

2.1 분석 방법 소개

1. 기초 통계량

데이터가 많을수록 변수가 많을수록 데이터를 한눈에 보는 것이 중요하다. 따라서 진행할 집단별, 변수별 모든 평균을 계산한다.

1.1 5개의 연도별 집단의 변수를 파악하고 각 변수의 표본평균을 계산한다

1.2 1.1을 바탕으로 데이터를 합쳐 모든 데이터의 변수별 평균을 구한다.

1.3 성별로 변수별 평균을 구한다.

1.4 연령대별 변수별 평균을 구한다.

2. 주성분 분석

1.4의 분석방법에서 설명한 주성분 분석이다.

분석은 표본 공분산 행렬과 표본 상관계수 행렬 두 가지 모두를 가지고 분석을 진행한 후 두 결과를 비교하고 설명이 잘되는 분석을 채택한다. 전체 데이터의 80% 이상의 설명이 가능한 주성분 까지 분석한다.

2.1 전체 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 주성분 분석

2.2 여성 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 주성분 분석

2.3 남성 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 주성분 분석

3. 인자분석

1.4의 분석방법에서 설명한 인자 분석이다.

계산 방법으로는 주성분 방법과 최대우도 방법을 사용하고 추가로 인자적재 회전인 varimax 회전을 실시하고 모든 결과를 비교하여 설명이 잘되는 인자 분석 방법을 채택한다. 또한 heywood상황에는 공통성을 1로 지정한다.

3.1 전체 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 인자분석

3.2 여성 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 인자분석

3.3 남성 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 인자분석

2.2 데이터 분석 및 결과 설명 *자세한 결과 분석해석은 3.1에서

1. 기초 통계량

1.1 연도별 표본평균

1.2 전체 표본평균

1.3 성별 표본평균

1.4 연령별 표본평균

구 분	전체	2007	2009	2013	2015	2017
표본	20388	4632	3783	3674	4034	4292
연령	39.14	29.59	39.85	54.25	38.67	41.08
신장	163.13	152.78	165.49	165.37	167.70	165.93
체중	62.34	51.99	64.95	64.03	66.46	65.80
BMI	23.08	21.42	23.63	23.31	23.53	23.78
20m 달리기	28.93	-	28.62	28.47	30.06	28.53
50m 달리기	10.28	10.38	9.81	10.98	9.96	-
제자리멀리뛰기	166.48	151.15	170.98	165.70	178.07	168.71
윗몸일으키기	28.96	25.27	29.83	27.92	31.02	31.10
이완혈압	76.02	76.05	75.19	-	76.77	-
수축혈압	125.56	122.82	125.55	-	127.90	-
체지방률	21.45	18.45	-	20.86	20.34	26.23
악력	34.04	-	33.60	32.61	35.73	34.08

구 분	남	여	0	10	20	30	40	50	60
연령	36.5	37.6	8.1	14.5	24.4	34.5	44.5	54.3	61.9
신장	168.1	156.8	128.1	153.9	168.5	167.6	165.2	162.6	161.5
체중	68.4	55.1	28.7	49.5	64.4	66.6	65.3	64.3	63.3
BMI	23.8	22.2	17.3	20.3	22.5	23.6	23.8	24.3	24.2
50m 달리기	9.1	11.1	11.2	9.5	9.0	9.6	10.1	11.2	12.4
제자리멀리뛰기	192.3	139.1	115.0	162.8	188.7	178.8	169.3	151.5	136.2
윗몸일으키기	35.6	22.3	22.1	33.4	37.8	32.2	28.0	21.1	17.8
이완혈압	77.3	74.6	70.6	72.3	74.7	76.0	76.9	78.2	76.7
수축혈압	128.6	122.2	109.8	117.6	125.1	126.1	126.7	128.8	127.7
체지방률	17.6	26.0	13.3	17.3	20.1	21.8	22.7	23.7	24.6
악력	42.5	25.1	-	33.5	35.0	36.0	34.5	32.1	29.9
20m 달리기	36.4	20.4	-	35.8	37.9	31.3	27.0	21.9	17.1
10m 달리기	11.7	14.1	-	12.2	11.9	12.4	12.9	13.7	14.6

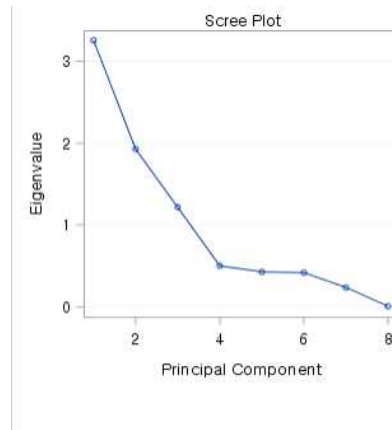
분석을 위한 기초 통계량으로 전체 / 그룹별 변수들의 평균은 위와 같다.

2. 주성분 분석

2.1 전체 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 주성분 분석

1) 표본 상관관계수 행렬을 이용한 주성분 분석

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	3.25595276	1.32836658	0.4070	0.4070
2	1.92758617	0.71013662	0.2409	0.6479
3	1.21744955	0.71510322	0.1522	0.8001
4	0.50234633	0.07290444	0.0628	0.8629
5	0.42944188	0.00971083	0.0537	0.9166
6	0.41973105	0.18186349	0.0525	0.9691
7	0.23786755	0.22824284	0.0297	0.9988
8	0.00962472		0.0012	1.0000



Eigenvectors									
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
height	신장	0.479848	-.045346	-.188885	-.075886	-.249890	0.571074	-.380591	0.439128
weight	체중	0.503913	0.181189	-.256465	-.158724	-.082296	-.073350	-.164406	-.763605
BMI	BMI	0.365273	0.402968	-.245616	-.150373	0.133655	-.608055	0.107918	0.471183
iwan	이완혈압	0.222609	0.222782	0.666326	0.329264	-.548930	-.201538	-.080349	0.003349
su	수축혈압	0.305287	0.175513	0.573771	-.225003	0.645910	0.267864	0.083337	-.016114
bodyfat	체지방율	-.047939	0.574535	-.249352	0.653029	0.149967	0.330160	0.213990	-.041251
jump	멀리뛰기	0.402418	-.394814	-.046569	0.086390	-.151998	0.090038	0.800822	0.004876
situp	윗몸일으	0.274593	-.486134	-.025926	0.594546	0.385748	-.263216	-.340539	0.008011

총 8개의 변수에 대하여 주성분 분석을 진행 한 결과 3개의 변수만으로 8개의 변수의 데이터의 약 80% 설명할 수 있음을 알 수 있다.

3개의 주성분은 다음과 같이 원래의 8개의 변수에 대한 선형변환으로 얻어진다.

$$P_1 = 0.479848x_1 + 0.503913x_2 + 0.365273x_3 + 0.222609x_4 + 0.305287x_5 - 0.047939x_6 \\ + 0.402418x_7 + 0.274593x_8$$

$$P_2 = -0.045346x_1 + 0.181189x_2 + 0.402968x_3 + 0.222782x_4 + 0.175513x_5 + 0.574535x_6 \\ - 0.394814x_7 - 0.486134x_8$$

$$P_3 = -0.188885x_1 - 0.256465x_2 - 0.245616x_3 + 0.666326x_4 + 0.573771x_5 - 0.249352x_6 \\ - 0.046569x_7 - 0.025926x_8$$

x_i 들은 순서대로 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기이다.

각 주성분은 다음과 같이 해석된다.

첫 번째 주성분은 체지방율을 제외한 대부분 변수의 영향을 고르게 받는다.

두 번째 주성분은 첫 번째 성분보다 더 확실한 특징을 가지고 있다.

신장에는 거의 영향을 받지 않는 변수이며 체중, BMI, 이완혈압, 수축혈압, 체지방률에는 양의 영향을 받으나 멀리뛰기, 윗몸일으키기에는 음의 영향을 받는다.

즉, 신체적 능력을 신체기능의 능력과 체력적 능력으로 나누어 볼 때 두 번째 성분은 신체기능에는 양의 상관성, 체력적 능력과는 음의 상관성이 있는 변수임을 알 수 있다.

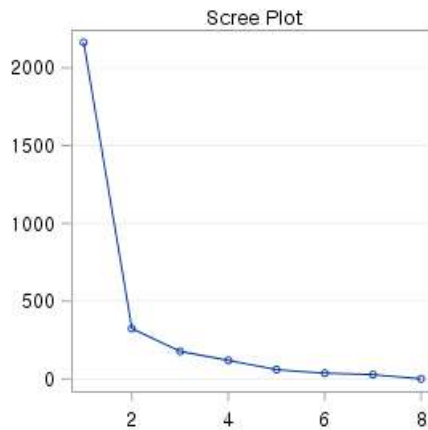
세 번째 주성분 역시 확실한 특징이 있다.

체력적 능력에는 영향을 거의 받지 않는다.

대신 신체기능에는 영향을 많이 받으며 특히 혈압에는 양의 영향을 신장, 체중, BMI, 체지방률에는 음의 영향을 받는다.

2) 표본 공분산 행렬을 이용한 주성분 분석

Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	2164.30415	1840.28275	0.7441	0.7441
2	324.02140	146.70549	0.1114	0.8555
3	177.31592	58.67565	0.0610	0.9165
4	118.64027	59.30098	0.0408	0.9573
5	59.33929	22.57115	0.0204	0.9777
6	36.76814	8.98841	0.0126	0.9903
7	27.77974	27.36979	0.0096	0.9999
8	0.40994		0.0001	1.0000



Eigenvectors									
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
height	신장	0.201349	0.357331	-.391657	0.159172	-.064429	-.603675	0.493272	0.202854
weight	체중	0.182851	0.530569	-.507942	0.154650	0.052206	0.229053	-.490010	-.328390
BMI	BMI	0.016704	0.123099	-.098850	0.015028	0.033093	0.246795	-.251920	0.921468
iwan	미완혈압	0.030826	0.313763	0.340363	-.084044	0.870464	-.141342	-.002349	0.001361
su	수축혈압	0.082597	0.616025	0.630002	0.031382	-.455784	0.076119	0.046994	-.007892
bodyfat	체지방률	-.060811	0.146331	-.174533	-.035523	0.128102	0.689784	0.670171	-.042715
jump	멀리뛰기	0.929640	-.191654	0.065218	-.290252	-.001164	0.099345	0.025834	0.000977
situp	윗몸일으	0.223707	-.198948	0.177098	0.925724	0.100848	0.103732	0.033146	0.004081

상관계수를 이용한 주성분 분석보다 적은 개수의 주성분으로 더 많은 데이터를 설명할 수 있다.

표본 공분산 행렬을 이용한 분석에서는 두 개의 주성분 만으로 8개 변수의 데이터를 약 86% 설명 할 수 있다.

각 주성분은 다음과 같이 해석된다.

첫 번째 주성분은 멀리뛰기에 영향을 크게 받는다.

두 번째 주성분은 체중과 혈압에 영향을 크게 받는다.

결과 비교

표본 공분산 행렬을 이용하면 2개의 주성분으로 8개 변수의 86%를 설명가능하고, 표본 상관계수 행렬을 이용하면 4개의 주성분으로 8개의 변수의 80%의 설명이 가능하다.

다만 변수별 단위가 모두 다르고 실제 주성분이 설명하는 것을 보면 표본 상관계수를 이용한 주성분 분석은 명확히 해당 주성분의 의미를 파악 할 수 있다.

2.2 여성 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 주성분 분석

1) 표본 공분산 행렬을 이용한 주성분 분석

Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	895.821868	579.289467	0.5652	0.5652
2	316.532401	166.859413	0.1997	0.7649
3	149.672988	33.135687	0.0944	0.8594
4	116.537302	60.586586	0.0735	0.9329
5	55.950715	21.843462	0.0353	0.9682
6	34.107253	18.037434	0.0215	0.9897
7	16.069819	15.877819	0.0101	0.9999
8	0.192000		0.0001	1.0000



Eigenvectors									
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
height	신장	0.122124	0.250949	0.498679	0.270505	-.040216	-.692658	0.270336	0.213969
weight	체중	0.027671	0.369134	0.588847	0.214933	0.052082	0.351787	-.451551	-.373727
BMI	BMI	-.017688	0.097591	0.127041	0.022763	0.031119	0.314754	-.242929	0.902470
iwan	이완혈압	0.006556	0.392948	-.238347	-.108700	0.868720	-.140570	-.049890	-.000428
su	수축혈압	0.014080	0.778827	-.422098	0.014961	-.455072	0.042185	0.077199	-.003143
bodyfat	체지방률	-.029827	0.101000	0.224933	0.029766	0.148921	0.509529	0.809662	-.008818
jump	멀리뛰기	0.945730	-.008042	0.036016	-.314946	-.022168	0.067448	-.000961	-.000739
situp	윗몸일으	0.297441	-.141405	-.318701	0.876354	0.101329	0.108126	-.001740	0.002206

세 개의 주성분을 이용하여 8개 변수의 86%의 데이터를 설명할 수 있다.

각 주성분은 다음과 같이 해석 된다.

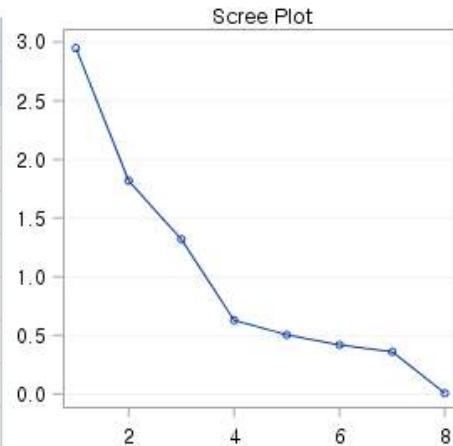
첫 번째 주성분은 멀리뛰기에 가장 큰 영향을 받는다.

두 번째 주성분은 수축 혈압에 가장 큰 영향을 받으며 멀리뛰기에는 영향을 받지 않는다.

세 번째 주성분은 신장과 체중에는 양의 영향을 받으며 혈압에는 음의 영향을 받는다.

2) 표본 상관계수 행렬을 이용한 주성분 분석

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	2.94715693	1.12968624	0.3684	0.3684
2	1.81747068	0.49920225	0.2272	0.5956
3	1.31826843	0.68998341	0.1648	0.7604
4	0.62828503	0.12392339	0.0785	0.8389
5	0.50436164	0.08614146	0.0630	0.9019
6	0.41822018	0.05875187	0.0523	0.9542
7	0.35946832	0.35269952	0.0449	0.9992
8	0.00676879		0.0008	1.0000



Eigenvectors									
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
height	신장	0.341936	0.413879	-.200176	-.570442	0.325606	0.270988	-.065444	0.403204
weight	체중	0.540354	0.114843	-.221363	-.153157	-.192910	-.055256	-.185897	-.739962
BMI	BMI	0.495072	-.150440	-.148126	0.218292	-.504249	-.286177	-.191706	0.538297
iwan	미완혈압	0.269731	0.036538	0.658404	0.210714	0.393846	-.068165	-.536892	-.000644
su	수축혈압	0.314938	0.080435	0.609018	-.123568	-.296091	0.185600	0.621295	-.006901
bodyfat	체지방률	0.393340	-.168125	-.287867	0.554867	0.480925	0.234873	0.373896	-.004437
jump	멀리뛰기	-.028095	0.640748	-.024764	0.183072	0.146936	-.680156	0.265120	-.005676
situp	윗몸일으	-.136213	0.588434	-.033286	0.449917	-.326857	0.533613	-.200152	0.004339

4개의 주성분을 이용하여 8개 변수의 84%의 데이터를 설명할 수 있다.

각 주성분은 다음과 같이 해석된다.

첫 번째 주성분은 신체기능에는 양의 영향을 받으며 체력적 능력에는 영향을 거의 받지 않거나 약간의 음의 영향을 받는다.

두 번째 주성분은 혈압에는 영향을 받지 않으며 체력적 능력에는 양의 영향을 크게 받는다.

세 번째 주성분은 혈압에 양의 영향을 크게 받으며 나머지 변수에는 음의 영향을 받고 특히 체력적인 요인에는 영향을 거의 받지 않는다.

네 번째 주성분은 신장에는 음의 영향을 크게 받으며 체지방률에는 양의 영향을 크게 받는다.

결과 비교

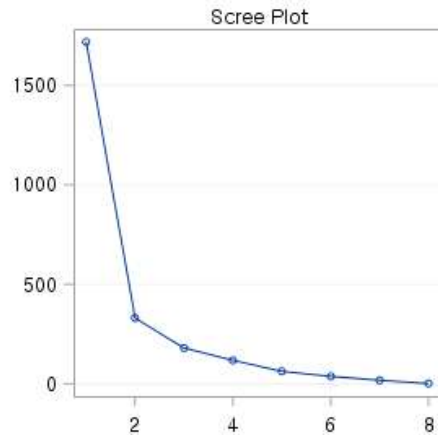
표본 공분산 행렬을 이용하면 3개의 주성분으로 8개 변수의 86%를 설명가능하고, 표본 상관계수 행렬을 이용하면 4개의 주성분으로 8개의 변수의 84%의 설명이 가능하다.

다만 변수별 단위가 모두 다르고 실제 주성분이 설명하는 것을 보면 표본 상관계수를 이용한 주성분 분석은 명확히 해당 주성분의 의미를 파악 할 수 있다.

2.3 남성 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 주성분 분석

1) 표본 공분산 행렬을 이용한 주성분 분석

Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	1719.47506	1389.39655	0.6984	0.6984
2	330.07851	150.82873	0.1341	0.8325
3	179.24978	62.40426	0.0728	0.9053
4	116.84551	54.81358	0.0475	0.9528
5	62.03194	25.58188	0.0252	0.9780
6	36.45005	19.07661	0.0148	0.9928
7	17.37344	16.96021	0.0071	0.9998
8	0.41323		0.0002	1.0000



Eigenvectors									
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
height	신장	0.229583	0.427549	-.285529	0.101902	-.084875	-.743793	0.267425	0.201579
weight	체중	0.199029	0.624637	-.404784	0.123988	0.054519	0.356966	-.395690	-.322529
BMI	BMI	0.021207	0.132398	-.076710	0.016248	0.035926	0.296967	-.177549	0.924591
iwan	이완혈압	0.025707	0.245708	0.395952	-.091275	0.870665	-.123376	-.024069	-.000145
su	수축혈압	0.089183	0.462271	0.748723	0.034612	-.459330	0.070146	0.024913	-.006629
bodyfat	체지방율	0.000577	0.191057	-.132163	-.039764	0.089179	0.447945	0.857572	-.020298
jump	멀리뛰기	0.929431	-.257394	0.055356	-.243575	0.006901	0.085985	0.008279	-.001882
situp	윗몸일으	0.186474	-.180777	0.104134	0.950547	0.107202	0.054470	0.060757	0.003552

두 개의 주성분을 이용하여 8개 변수의 83%의 데이터를 설명할 수 있다.

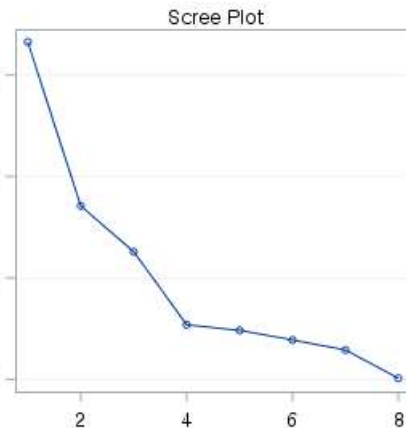
각 주성분은 다음과 같이 해석 된다.

첫 번째 주성분은 멀리뛰기에 가장 큰 영향을 받는다.

두 번째 주성분은 체중에 가장 큰 영향을 받으며 멀리뛰기에는 음의 영향을 받는다.

2) 표본 상관계수 행렬을 이용한 주성분 분석

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	3.32864943	1.61873477	0.4161	0.4161
2	1.70991466	0.45198599	0.2137	0.6298
3	1.25792867	0.72036209	0.1572	0.7871
4	0.53756658	0.05666614	0.0672	0.8543
5	0.48090044	0.09337850	0.0601	0.9144
6	0.38752194	0.09940158	0.0484	0.9628
7	0.28812036	0.27872244	0.0360	0.9988
8	0.00939792		0.0012	1.0000



Eigenvectors									
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
height	신장	0.464858	0.179718	-.110699	-.403737	-.218059	0.129365	-.562754	0.442010
weight	체중	0.520815	-.071057	-.186010	-.106667	-.005136	-.267036	-.140370	-.765955
BMI	BMI	0.434876	-.287528	-.188421	0.189756	0.202903	-.536122	0.332699	0.466279
iwan	미완혈압	0.201439	-.141757	0.695077	0.327737	-.569717	-.142580	-.062262	0.000620
su	수축혈압	0.278934	-.016023	0.619830	-.259706	0.649647	0.206064	0.075001	-.013150
bodyfat	체지방률	0.299635	-.465792	-.218271	0.367092	-.016457	0.713234	0.042740	-.007543
jump	멀리뛰기	0.305759	0.530303	-.041154	-.141316	-.287713	0.230235	0.683817	-.015274
situp	윗몸일으	0.147955	0.601202	-.031403	0.678252	0.285631	-.005255	-.271995	0.007765

4 개의 주성분을 이용하여 8개 변수의 83%의 데이터를 설명할 수 있다.

각 주성분은 다음과 같이 해석 된다.

첫 번째 주성분은 신장, 체중, BMI에 영향을 가장 많이 받으며 다른 변수도 골고루 영향을 받는다.

두 번째 주성분은 멀리뛰기와 윗몸일으키기에 양의 영향을 크게 받으며 다른 변수는 대체적으로 음의 영향을 받는다.

세 번째 주성분은 혈압에 양의 영향을 크게 받으며 나머지 변수에는 음의 영향을 받고 특히 체력적인 요인에는 영향을 거의 받지 않는다.

네 번째 주성분은 신장에는 음의 영향을 크게 받으며 체지방률에는 양의 영향을 크게 받는다.

결과 비교

표본 공분산 행렬을 이용하면 2개의 주성분으로 8개 변수의 83%를 설명가능하고, 표본 상관계수 행렬을 이용하면 4개의 주성분으로 8개의 변수의 83%의 설명이 가능하다.

다만 변수별 단위가 모두 다르고 실제 주성분이 설명하는 것을 보면 표본 상관계수를 이용한 주성분 분석은 명확히 해당 주성분의 의미를 파악 할 수 있다.

주성분 분석 결과 요약

전체 집단의 주성분 분석 결과와 성별로 나눈 두 집단의 주성분 분석 결과가 유사하다. 가장 두드러지는 특징은 표본 공분산 행렬과 표본 상관계수 행렬을 각각 사용할 때 둘의 차이점이 같게 나타났다.

표본 공분산 행렬을 사용한경우 2개 정도의 주성분으로 전체 데이터의 85% 정도를 설명하였다. 하지만 각각의 주성분의 의미가 불분명하여 해석이 합리적이지 못하다.

반대로 표본 상관계수 행렬을 사용할때에는 4개 정도의 주성분으로 82%정도의 데이터를 설명 할 수 있는데 이때는 각각의 주성분의 역할이 분명하게 드러났다.

또한 각각의 주성분에서 체력적인 요인과 신체 능력부분의 요인이 확실하게 분리 되는 것을 볼 수 있었다.

따라서 본 데이터의 주성분 분석을 진행하는 데에는 표본 상관계수 행렬을 이용하는 것이 합리적이다.

3. 인자 분석

3.1 전체 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 인자분석

1) 주성분 방법

Eigenvalues of the Correlation Matrix: Total = 8 Average = 1					Factor Pattern				
	Eigenvalue	Difference	Proportion	Cumulative			Factor1	Factor2	Factor3
1	3.25595276	1.32836658	0.4070	0.4070	weight	체중	0.90927	0.25156	-0.28298
2	1.92758617	0.71013662	0.2409	0.6479	height	신장	0.86585	-0.06296	-0.20841
3	1.21744955	0.71510322	0.1522	0.8001	jump	멀리뛰기	0.72613	-0.54815	-0.05138
4	0.50234633	0.07290444	0.0628	0.8629	BMI	BMI	0.65911	0.55947	-0.27101
5	0.42944188	0.00971083	0.0537	0.9166	bodyfat	체지방률	-0.08650	0.79767	-0.27513
6	0.41973105	0.18186349	0.0525	0.9691	situp	윗몸일으	0.49548	-0.67494	-0.02861
7	0.23786755	0.22824284	0.0297	0.9988	iwan	이완혈압	0.40168	0.30931	0.73521
8	0.00962472		0.0012	1.0000	su	수축혈압	0.55087	0.24368	0.63309

Final Communality Estimates: Total = 6.400988							
height	weight	BMI	iwan	su	bodyfat	jump	situp
0.79709446	0.97013603	0.82087520	0.79755305	0.76363437	0.71945618	0.83037814	0.70186105

3개의 인자를 이용하여 8개 변수 데이터의 80%를 설명 할 수 있다.

각 인자는 다음과 같다.

첫 번째 인자는 체중, 신장, 멀리뛰기, BMI에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 체지방률, BMI에 대한 인자적재값이 가장 크게 나타났고, 멀리뛰기와 윗몸 일으키기에 대한 인자적재값은 음수로 크게 나타났다.

세 번째 인자는 혈압에 관한 인자적재값이 양수로 큰 값을 가진다.

2) 주성분 방법 varimax회전 추가

Orthogonal Transformation Matrix			
	1	2	3
1	0.79684	0.46216	0.38916
2	0.37216	-0.88286	0.28644
3	-0.47596	0.08341	0.87550

Rotated Factor Pattern				
		Factor1	Factor2	Factor3
weight	체중	0.95285	0.17454	0.17816
BMI	BMI	0.86241	-0.21193	0.17949
height	신장	0.76571	0.43836	0.13646
situp	윗몸일으	0.15725	0.82248	-0.02555
jump	멀리뛰기	0.39907	0.81525	0.08059
bodyfat	체지방율	0.35888	-0.76716	-0.04606
iwan	미완혈압	0.08526	-0.02611	0.88860
su	수축혈압	0.22832	0.09226	0.83845

Final Communality Estimates: Total = 6.400988							
height	weight	BMI	iwan	su	bodyfat	jump	situp
0.79709446	0.97013603	0.82087520	0.79755305	0.76363437	0.71945618	0.83037814	0.70186105

앞의 결과에 varimax 회전을 추가 하였다.

각 인자는 다음과 같다.

첫 번째 인자는 체중, 신장, BMI에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 멀리뛰기와 윗몸일으키기에 대한 인자적재값이 가장 크게 나타났다.

세 번째 인자는 혈압에 관한 인자적재값이 양수로 큰 값을 가진다.

3) 최대우도 방법 heywood

Preliminary Eigenvalues: Total = 108.452216 Average = 13.5565271					Factor Pattern			
	Eigenvalue	Difference	Proportion	Cumulative			Factor1	Factor2
1	93.5257035	80.0942898	0.8624	0.8624	BMI	BMI	0.89722	0.44158
2	13.4314137	11.6655486	0.1238	0.9862	bodyfat	체지방율	0.43968	-0.11403
3	1.7658651	0.6849675	0.0163	1.0025	height	신장	0.00000	0.97334
4	1.0808976	1.1022952	0.0100	1.0125	weight	체중	0.50627	0.86238
5	-0.0213976	0.2755852	-0.0002	1.0123	jump	멀리뛰기	-0.13867	0.66266
6	-0.2969828	0.0615184	-0.0027	1.0095	situp	윗몸일으	-0.21185	0.41329
7	-0.3585012	0.3162808	-0.0033	1.0062	su	수축혈압	0.20033	0.30769
8	-0.6747820		-0.0062	1.0000	iwan	이완혈압	0.17236	0.17417

각 인자는 다음과 같다.

첫 번째 인자는 BMI에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 신장, 체중, 멀리뛰기에 대한 인자적재값이 가장 크게 나타났다.

4) 최대우도 방법 heywood varimax 회전 추가

Rotated Factor Pattern			
		Factor1	Factor2
BMI	BMI	0.99831	-0.05807
weight	체중	0.86558	0.50076
bodyfat	체지방율	0.32634	-0.31595
su	수축혈압	0.32597	0.16895
iwan	이완혈압	0.23582	0.06658
height	신장	0.47980	0.84687
jump	멀리뛰기	0.20600	0.64491
situp	윗몸일으	0.01940	0.46402

Orthogonal Transformation Matrix		
	1	2
1	0.87006	-0.49294
2	0.49294	0.87006

앞의 결과에 varimax 회전을 추가 하였다.

각 인자는 다음과 같다.

첫 번째 인자는 BMI에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 신장에 대한 인자적재값이 가장 크게 나타났다.

결과 비교

주성분방법과 최대우도 방법을 비교했을 때, 주성분 방법으로 구한 인자 적재값들이 해당 인자의 의미를 최대우도 방법보다 설명을 잘하는 경향을 보였다.

주성분 방법을 사용한 인자분석은 각 인자별로 어떤 의미를 갖는지가 분명했으며 varimax 회전을 추가하면 그 의미가 더욱 분명해 졌다.

3.2 여성 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 인자분석

1) 주성분 방법

Eigenvalues of the Correlation Matrix: Total = 8 Average = 1					Factor Pattern					
	Eigenvalue	Difference	Proportion	Cumulative			Factor1	Factor2	Factor3	Factor4
1	2.94715693	1.12968624	0.3684	0.3684	weight	체중	0.92764	0.15482	-0.25416	-0.12140
2	1.81747068	0.49920225	0.2272	0.5956	BMI	BMI	0.84990	-0.20281	-0.17007	0.17303
3	1.31826843	0.68998341	0.1648	0.7604	bodyfat	체지방률	0.67526	-0.22665	-0.33052	0.43981
4	0.62828503	0.12392339	0.0785	0.8389	height	신장	0.58701	0.55796	-0.22983	-0.45216
5	0.50436164	0.08614146	0.0630	0.9019	jump	멀리뛰기	-0.04823	0.86382	-0.02843	0.14511
6	0.41822018	0.05875187	0.0523	0.9542	situp	윗몸일으	-0.23384	0.79329	-0.03822	0.35662
7	0.35946832	0.35269952	0.0449	0.9992	iwan	미완혈압	0.46305	0.04926	0.75595	0.16702
8	0.00676879		0.0008	1.0000	su	수축혈압	0.54066	0.10844	0.69925	-0.09795

Final Communality Estimates: Total = 6.711181							
height	weight	BMI	iwan	su	bodyfat	jump	situp
0.91317605	0.96382507	0.82233416	0.81620501	0.80261831	0.81002216	0.77037008	0.81263023

1개의 인자를 이용하여 8개 변수 데이터의 84%를 설명 할 수 있다.

각 인자는 다음과 같다.

첫 번째 인자는 체중, BMI, 체지방률, 신장에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 멀리뛰기, 윗몸일으키기에 대한 인자적재값이 크게 나타났다.

세 번째 인자는 혈압에 관한 인자적재값이 양수로 큰 값을 가진다.

네 번째 인자는 체지방률에 대한 인자적재값의 양수로 큰 값을 가지며 반대로 신장에 대한 인자 적재값은 음수로 큰 값을 갖는다.

2) 주성분 방법 varimax회전 추가

Orthogonal Transformation Matrix					Rotated Factor Pattern					
	1	2	3	4			Factor1	Factor2	Factor3	Factor4
1	0.71902	0.44009	-0.13201	0.52145	bodyfat	체지방율	0.89607	0.00711	-0.07662	0.03398
2	-0.19417	0.08223	0.88192	0.42161	BMI	BMI	0.80949	0.21484	-0.20576	0.28029
3	-0.37000	0.89257	-0.04322	-0.25407	iwan	미완혈압	0.13643	0.89153	0.02488	-0.04625
4	0.55535	0.05363	0.45047	-0.69698	su	수축혈압	0.05458	0.86573	-0.05008	0.21826
					situp	윗몸일으	-0.10998	-0.05267	0.89279	-0.02633
					jump	멀리뛰기	-0.11130	0.03220	0.83478	0.24512
					height	신장	0.14766	0.07482	0.22084	0.91488
					weight	체중	0.66355	0.18761	-0.02961	0.69818

앞의 결과에 varimax 회전을 추가 하였다.

각 인자는 다음과 같다.

첫 번째 인자는 체중, BMI에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 혈압에 관한 인자적재값이 양수로 큰 값을 가진다.

세 번째 인자는 멀리뛰기, 윗몸일으키기에 대한 인자적재값이 크게 나타났다

네 번째 인자는 신장과 체중에 대한 인자적재값이 크게 나타났다.

3) 최대우도 방법 heywood

Preliminary Eigenvalues: Total = 150.132208 Average = 18.766526					Factor Pattern			
	Eigenvalue	Difference	Proportion	Cumulative			Factor1	Factor2
1	125.622784	102.094232	0.8367	0.8367	BMI	BMI	0.97143	0.23733
2	23.528552	22.229728	0.1567	0.9935	bodyfat	체지방률	0.53336	0.24381
3	1.298823	0.283223	0.0087	1.0021	su	수축혈압	0.26751	0.21923
4	1.015600	0.987979	0.0068	1.0089	iwan	미완혈압	0.20639	0.11962
5	0.027621	0.336716	0.0002	1.0091	situp	윗몸일으	-0.30612	0.16441
6	-0.309095	0.080724	-0.0021	1.0070	height	신장	0.00000	1.00000
7	-0.389819	0.272439	-0.0026	1.0044	weight	체중	0.69674	0.70864
8	-0.662258		-0.0044	1.0000	jump	멀리뛰기	-0.24716	0.33538

Prior Communality Estimates: SMC							
height	weight	BMI	iwan	su	bodyfat	jump	situp
0.96015625	0.98769238	0.97716844	0.35297371	0.40473631	0.37626058	0.38620969	0.34500592

각 인자는 다음과 같다.

첫 번째 인자는 BMI, 체지방률, 체중에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 신장, 체중에 대한 인자적재값이 가장 크게 나타났다.

4) 최대우도 방법 heywood varimax 회전 추가

Orthogonal Transformation Matrix		
	1	2
1	0.83117	-0.55602
2	0.55602	0.83117

Rotated Factor Pattern			
		Factor1	Factor2
weight	체중	0.97313	0.20160
BMI	BMI	0.93938	-0.34287
bodyfat	체지방율	0.57887	-0.09391
su	수축혈압	0.34424	0.03348
iwan	이완혈압	0.23806	-0.01533
height	신장	0.55602	0.83117
jump	멀리뛰기	-0.01895	0.41618
situp	윗몸일으	-0.16302	0.30686

앞의 결과에 varimax 회전을 추가 하였다.

각 인자는 다음과 같다.

첫 번째 인자는 체중과 BMI에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 신장에 대한 인자적재값이 가장 크게 나타났다.

3.3 남성 데이터에서 신장, 체중, BMI, 이완/수축혈압, 체지방, 윗몸일으키기, 팔굽혀펴기 변수의 인자분석

1) 주성분 방법

Eigenvalues of the Correlation Matrix: Total = 8 Average = 1					Factor Pattern					
	Eigenvalue	Difference	Proportion	Cumulative			Factor1	Factor2	Factor3	Factor4
1	3.32864943	1.61873477	0.4161	0.4161	weight	체중	0.95020	-0.09292	-0.20862	-0.07821
2	1.70991466	0.45198599	0.2137	0.6298	height	신장	0.84811	0.23501	-0.12416	-0.29602
3	1.25792867	0.72036209	0.1572	0.7871	BMI	BMI	0.79341	-0.37598	-0.21133	0.13913
4	0.53756658	0.05666614	0.0672	0.8543	situp	윗몸일으	0.26994	0.78615	-0.03522	0.49729
5	0.48090044	0.09337850	0.0601	0.9144	jump	멀리뛰기	0.55784	0.69344	-0.04616	-0.10361
6	0.38752194	0.09940158	0.0484	0.9628	bodyfat	체지방률	0.54667	-0.60909	-0.24481	0.26915
7	0.28812036	0.27872244	0.0360	0.9988	iwan	이완혈압	0.36752	-0.18537	0.77958	0.24029
8	0.00939792		0.0012	1.0000	su	수축혈압	0.50890	-0.02095	0.69519	-0.19041

Final Commuality Estimates: Total = 6.834059							
height	weight	BMI	iwan	su	bodyfat	jump	situp
0.87756635	0.96116323	0.83488243	0.83491547	0.77896264	0.80220936	0.80492084	0.93943902

1개의 인자를 이용하여 8개 변수 데이터의 85%를 설명 할 수 있다.

각 인자는 다음과 같다.

첫 번째 인자는 체중, 신장, BMI에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 멀리뛰기, 윗몸일으키기에 대한 인자적재값이 크게 나타났고 반대로 체지방률에 대한 인자적재값은 음수로 크게 나타났다.

세 번째 인자는 혈압에 관한 인자적재값이 양수로 큰 값을 가진다.

네 번째 인자는 윗몸일으키기에 대한 인자적재값이 양수로 큰 값을 가진다.

2) 주성분 방법 varimax회전 추가

Orthogonal Transformation Matrix				
	1	2	3	4
1	0.64546	0.65073	0.34547	0.20146
2	-0.56687	0.40849	-0.12145	0.70501
3	-0.34488	-0.13890	0.92757	-0.03703
4	0.37829	-0.62482	0.07419	0.67897

Rotated Factor Pattern					
		Factor1	Factor2	Factor3	Factor4
bodyfat	체지방률	0.88437	-0.02724	0.05573	-0.12747
BMI	BMI	0.85076	0.30513	0.13407	-0.00294
weight	체중	0.70835	0.65821	0.14024	0.08055
height	신장	0.34505	0.85009	0.12733	0.14016
jump	멀리뛰기	-0.05630	0.71742	0.05800	0.53263
iwan	미완혈압	0.16433	-0.09499	0.89043	0.07764
su	수축혈압	0.02857	0.34501	0.80907	-0.06728
situp	윗몸일으	-0.07115	0.19097	0.00200	0.94758

앞의 결과에 varimax 회전을 추가 하였다.

각 인자는 다음과 같다.

첫 번째 인자는 체지방률, BMI, 체중에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 체중과 신장에 관한 인자적재값이 양수로 큰 값을 가진다.

세 번째 인자는 혈압에 대한 인자적재값이 크게 나타났다

네 번째 인자는 멀리뛰기와 윗몸일으키기에 대한 인자적재값이 크게 나타났다.

3) 최대우도 방법 heywood

Preliminary Eigenvalues: Total = 110.077688 Average = 13.759711					Factor Pattern			
	Eigenvalue	Difference	Proportion	Cumulative			Factor1	Factor2
1	97.1893040	84.9475954	0.8829	0.8829	BMI	BMI	0.88296	0.46944
2	12.2417086	11.1517874	0.1112	0.9941	bodyfat	체지방률	0.52613	0.30174
3	1.0899212	0.1833647	0.0099	1.0040	height	신장	0.00000	1.00000
4	0.9065565	0.9831099	0.0082	1.0123	weight	체중	0.52773	0.83983
5	-0.0765534	0.2139923	-0.0007	1.0116	jump	멀리뛰기	-0.13183	0.60326
6	-0.2905457	0.0319663	-0.0026	1.0089	su	수축혈압	0.12750	0.32673
7	-0.3225120	0.3376789	-0.0029	1.0060	situp	윗몸일으	-0.15350	0.28477
8	-0.6601909		-0.0060	1.0000	iwan	미완혈압	0.13768	0.16291

Prior Communality Estimates: SMC							
height	weight	BMI	iwan	su	bodyfat	jump	situp
0.95541082	0.98407971	0.95926073	0.25855369	0.32410133	0.43441293	0.54290853	0.33660876

각 인자는 다음과 같다.

첫 번째 인자는 BMI, 체지방률에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 신장, 체중에 대한 인자적재값이 가장 크게 나타났다.

4) 최대우도 방법 heywood varimax 회전 추가

Rotated Factor Pattern			
		Factor1	Factor2
BMI	BMI	0.99950	0.03177
weight	체중	0.84409	0.52088
bodyfat	체지방율	0.60528	0.03869
su	수축혈압	0.25856	0.23697
iwan	이완혈압	0.19542	0.08546
height	신장	0.44115	0.89743
jump	멀리뛰기	0.14782	0.59954
situp	윗몸일으	-0.01213	0.32329

Orthogonal Transformation Matrix		
	1	2
1	0.89743	-0.44115
2	0.44115	0.89743

앞의 결과에 varimax 회전을 추가 하였다.

각 인자는 다음과 같다.

첫 번째 인자는 체중과 BMI에 대한 인자적재값이 크게 나타났다.

두 번째 인자는 신장에 대한 인자적재값이 가장 크게 나타났다.

3.1 분석 결과 요약

1. 주성분 분석

전체 그룹, 여성 그룹, 남성 그룹으로 총 세가지의 집단으로 나누어 분석을 실시하였다. 각 집단 별로는 표본 공분산 행렬과 표본 상관계수 행렬을 이용한 분석을 실시하여 총 6가지의 분석을 실시 하였다.

1. 전체집단

전체 집단은 분석 결과 표본 공분산 행렬을 이용한 분석이 표본 상관계수를 이용한 분석보다 더 적은 주성분으로 대부분의 데이터를 설명 하였다. 각 주성분의 의미가 불분명하여 해석이 합리적으로 되지 않았다. 데이터의 측정 스케일이 다른점을 고려해 보면 표본 상관계수를 이용한 분석이 합리적으로 나올것으로 예상 되었으며 결과 역시 표본 상관계수를 이용한 분석이 각 주성분이 갖는의미가 명확했다.

따라서 전체 집단의 표본 상관계수 행렬을 이용한 분석결과를 채택하는 것이 합리적이며 각 주성분의 의미는 다음과 같다.

첫 번째 주성분은 체지방률을 제외한 변수들의 가중평균을 의미한다. 두 번째 주성분은 BMI와 체지방률에 높은 영향을 받는다. 따라서 두 번째 주성분은 8개의 변수중에서 지방을 의미하는 BMI와 체지방률에 대한 주성분이라고 해석할 수 있다. 세 번째 주성분은 이완 혈압과 수축 혈압에 많은 영향을 받는 이는 혈압에 대한 주성분이라고 해석하는 것이 합리적이다.

따라서 전체 집단의 8개의 변수로 주성분 분석을 하면 세 개의주성분으로 전체 데이터의 80%를 설명 할 수 있으며 각각의 주성분이 확실한 의미가 있으므로 타당한 주성분 분석이다.

2. 여성 / 남성 집단

여성 집단과 남성 집단 모두 표본공분산 행렬과 표본 상관계수 행렬을 모두 이용하여 분석을 실시 하였다. 두 집단 모두 전체 집단과 마찬가지로 표본 공분산 행렬을 이용한 분석이 표본 상관계수를 이용한 분석 보다 더 적은 주성분으로 더 많은 설명력을 보이지만 의미가 불분명하고 이는 변수별 스케일이 다르기 때문이다.

따라서 여성 / 남성 집단의 표본 상관계수 행렬을 이용한 분석 결과를 채택하는 것이 합리적이며 각 주성분과 여성 / 남성 집단의 차이는 아래와 같다.

여성 집단의 경우 4개의 주성분을 이용하여 8개 변수의 84%의 데이터를 설명할 수 있으며 첫 번째 주성분은 신장, 체중, BMI, 혈압, 체지방률에 큰 영향을 받으며 멀리뛰기와 윗몸 일으키기에는 거의 영향을 받지 않거나 음의 영향을 받는다. 즉, 첫 번째 주성분은 신체 기능적 변수에 큰 영향을 받으며 체력적인 변수에는 거의 영향을 받지 않는다. 두 번째 주성분은 첫 번째 주성분과는 반대로 체력적인 변수에 영향을 크게 받는다. 세 번째 변수는 혈압에 영향을 크게 받는다.

여성 집단의 분석 결과 8개의 변수를 혈압과 신장, 체중, BMI, 체지방률과 멀리뛰기, 윗몸 일으키기의 세 개의 변수 집단으로 나누어져 각각의 주성분으로 설명됨을 알 수 있다.

남성 집단의 경우 4개의 주성분을 이용하여 8개 변수의 83%의 데이터를 설명할 수 있으며 여성 집단과 비슷하게 첫 번째 주성분은 신체 기능적 변수에, 두 번째 주성분은 체력적인 변수에 세 번째 변수는 혈압에 영향을 크게 받는 것으로 나타났다.

여성과 남성 모두 비슷하게 주성분 분석이 되었으며 이는 전체 집단의 주성분 분석 결과와 유사한 경향을 보였다.

2. 인자분석

전체 그룹, 여성 그룹, 남성 그룹으로 총 세가지의 집단으로 나누어 분석을 실시하였다. 각 집단 별로는 주성분 방법, 최대 우도 방법을 사용 하였으며 각각 varimax 회전을 실시 하였다. 따라서 총 12가지의 분석을 실시 하였다.

1. 전체집단

총 4가지의 분석 중 주성분 방법이 최대우도의 방법보다 인자는 많지만 각각의 인자의 의미가 분명하였다. 또한 varimax 회전을 사용하면 그 의미가 더욱 분명하다. 결과는 아래와 같다.

첫 번째 인자는 체중, 신장, BMI에 대한 인자적재값이 크게 나타나 신체 특성과 관련된 인자로 해석 가능하다. 두 번째 인자는 멀리뛰기와 윗몸일으키기에 대한 인자적재값이 가장 크게 나타 체력적 특성과 관련된 인자로 해석하고 세 번째 인자는 이완 / 수축 혈압에 대한 인자적재 값이 크게 나타나 혈압에 관한 인자로 볼 수 있다.

2. 여성 / 남성 집단

전체 집단과 같이 주성분 방법에 varimax 회전을 사용한 분석이 가장 좋은 결과를 산출 했으며 각각의 인자의 의미는 전체집단의 인자와 같게 나타났다.

3.2 분석의 장점 및 한계점 설명

주성분 분석은 적은 변수로 전체 데이터의 80% 정도를 설명하고 각각의 주성분도 특징이 명확하게 분석이 잘 되었고, 인자 분석 역시 각각의 인자가 의미하는 바가 명확했다.

다만, 데이터가 완전 최신이 아니고 연도별 변수가 통일 되지 않아서 분석에 어려움이 있고 더 많은 분석이 불가능 하였다.

3.3 추가 연구사항 제안

최신의 데이터가 있고 변수가 더 다양하게 있으며 특히 청소년의 체력 문제가 대두되는 현 상황에서 청소년들의 데이터가 더 많이 구축된다면 청소년의 기초 체력 증진을 위한 정책을 세우는데 더 도움이 많이 되는 분석이 가능할 것이다.

4 참고문헌

[1] 기사 학생체력 떨어지고 비만 늘어 “스마트폰 사용 등 운동 부족”

<https://www.joongang.co.kr/article/23461261#home>

[2]논문 [운동이 면역력에 미치는 효과분석]

<https://synapse.koreamed.org/upload/SynapseData/PDFData/0078in/in-5-117.pdf>

[3]기사 면역력 증진… 기본으로 돌아가라

https://health.chosun.com/site/data/html_dir/2016/09/29/2016092901284.html

[4]면역력을증가시키는운동

<https://www.koreascience.or.kr/article/JAKO200776242312440.pdf>

[5] 건강증진정보

http://www.wellness.ac.kr/bbs/board.php?bo_table=report_youth&wr_id=1

[6] 학생 체력 해마다 저하…전자기기 사용↑·운동량↓ 원인

<https://www.news1.kr/articles/?3616219>

[7] “즐기면서 체력 증진”…학교 체육의 변화

<https://news.kbs.co.kr/news/view.do?ncd=5298215>