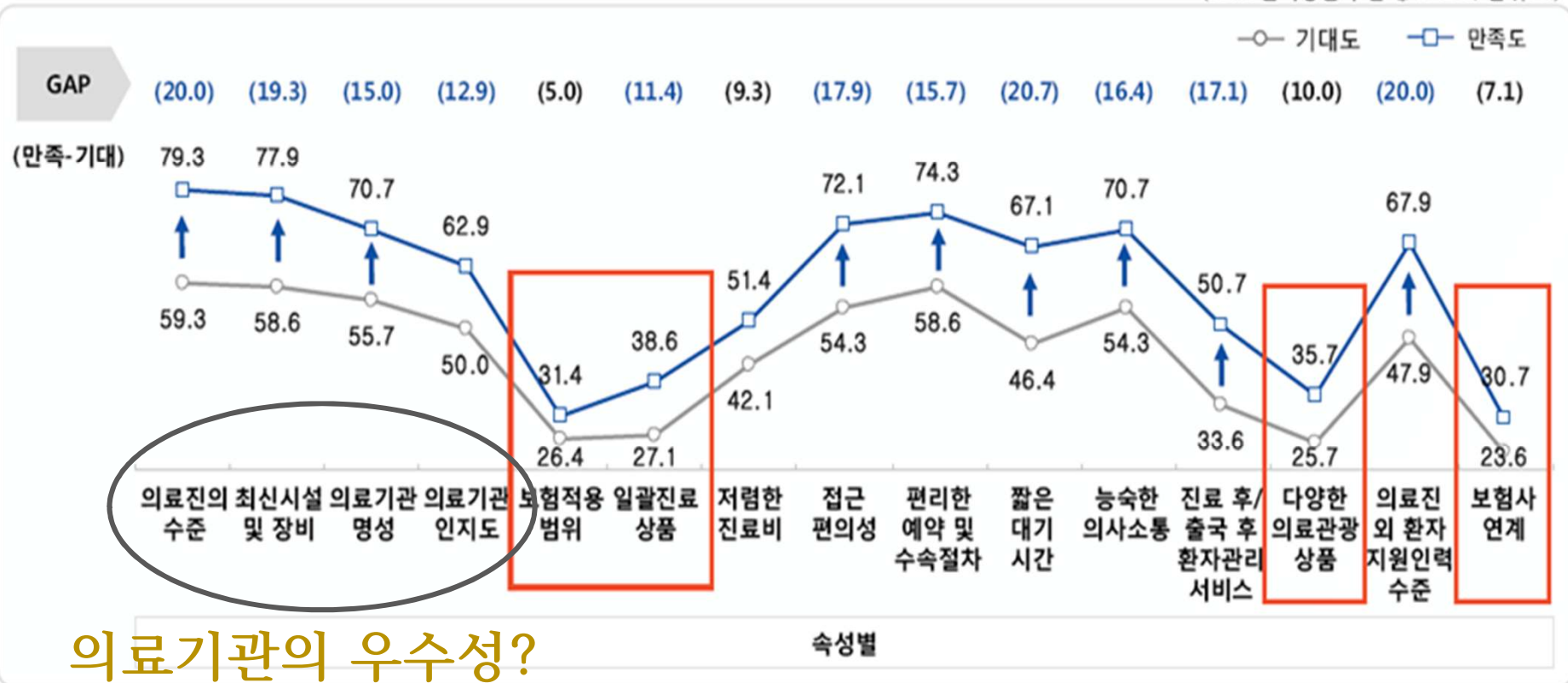


요인분석과 신뢰도분석

사례: 한국 의료관광 서비스 만족도 조사

한국 의료서비스 속성별 기대 및 만족도 (긍정평가 기준)

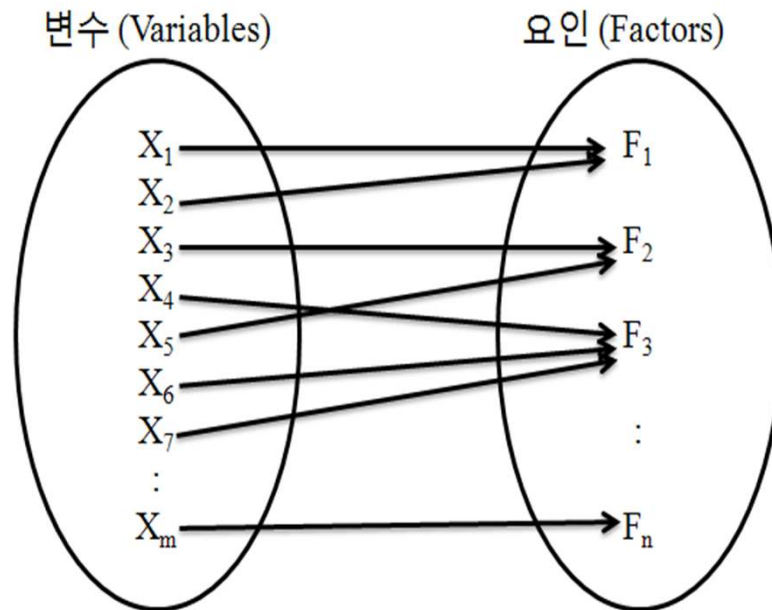
(Base: 한국경험자 전체, n=140, 단위: %)



(긍정평가%= 만족도 및 기대도 평가 중 4점, 5점 응답자의 비율을 말함)

요인분석의 개요

◆ 변수와 요인 (다수의 변수를 소수의 요인으로)



- 예) 차량의 선택시 고려사항: 외형, 인테리어, 연비, 가격
- 외형, 인테리어 → 디자인
- 연비, 가격 → 경제성

요인분석의 개요

- ◆ 차종(K5)에 대해 두 사람이 다음의 4가지 평가기준을 적용하여 평가한 후 10점 만점으로 점수를 부여하였다고 가정

- 차종의 만족도 평가

차종 \ 변수	외 형	인테리어	연 비	가 격
응답자 1	6.3	7.8	5.0	7.2
응답자 2	5.0	6.4	5.4	7.6

- 요인에 따른 만족도 점수 (요인점수)

차종 \ 요인	디자인	경제성
응답자 1	8	7
응답자 2	6	8

- 평가항목별 요인부하량

차종 \ 요인	디자인	경제성
외형	0.7	0.1
인테리어	0.8	0.2
연비	0.1	0.6
가격	0.2	0.8

요인분석의 개요

- ◆ 응답자 1의 외형 = (K5의 요인부하량×K5의 요인점수)의 합 = $8 \times 0.7 + 7 \times 0.1 = 6.3$
- ◆ 응답자 1의 인테리어 = $8 \times 0.8 + 7 \times 0.2 = 7.8$
- ◆ 응답자 1의 연비 = $8 \times 0.1 + 7 \times 0.6 = 5.0$
- ◆ 응답자 1의 가격 = $8 \times 0.2 + 7 \times 0.8 = 7.2$

- ◆ 응답자 2의 경우?

- ◆ 요인분석의 특징
 - 독립변수와 종속변수가 따로 없음
 - 모집단의 특성에 대한 추정을 하지 않음
 - 요인분석의 결과로 모수의 추정이나, 가설검정 등은 이루어지지 않으나 요인분석의 과정에서 일부 가설검정이 포함됨

요인분석의 목적

1. 요약을 통한 자료구조의 발견

- 다수의 변수를 상관관계를 고려하여 소수의 요인으로 압축하여 나타내므로 다수의 변수들 중 어떤 변수들이 비슷한 개념을 나타내고 있는지를 요인과 요인으로 묶여진 변수들을 통해 알 수 있음

2. 자료의 축소

- 여러 변수들이 비슷한 개념을 측정하고 있는 경우 요인분석을 통해 새로운 요인을 찾아내면 여러 변수들을 측정할 필요없이 소수의 요인만을 측정하여 대체할 수 있게 됨

3. 측정도구의 타당성 검토

- 많은 연구에서 새로운 평가항목을 개발하거나 여러 변수들로 구성된 구성개념간의 관계를 분석하고자 할 때 측정의 도구로 쓰인 변수들이 동일한 개념을 측정하는가의 여부를 확인하는데 사용됨

4. 다른 분석을 위한 데이터 생성

- 요인분석을 통해 얻어진 요인들을 회귀분석 등에 다시 활용하고자 하는 목적으로도 사용할 수 있음 (다중공선성을 제거한 값으로 변환하여 회귀분석)

요인분석의 기본 요건과 종류

◆ 요건

- 변수의 척도 : 비율척도나 등간척도로 구성될 것
- 표본의 크기 : 100개 이상 (200개 이상 권장)일 것, 변수의 개수보다 5배 이상일 것 변수간의 관계 (10배 이상 권장)
- 일정 수준이상의 상관관계를 가지는 변수들이 있을 것

◆ 요인분석의 종류

- 탐색적인 요인분석 (Exploratory Factor Analysis): 요인에 대한 기존의 가설이나 연구결과가 없는 경우에 사용되며 주어진 변수에 어떤 요인이 존재하는지를 처음 밝히는 연구에 적용되는 분석방법
- 확인적 요인분석 (Confirmatory Factor Analysis): 기존의 가설이나 연구결과가 존재하며 도출된 요인을 통해 수립된 가설을 모형화 하기 위해 사용되는 분석방법

요인분석의 용어

- ◆ 요인부하량 (Factor loading): 각 변수와 해당 요인간의 상관계수, 요인 적재량이라고도 함
- ◆ 요인점수 (Factor score): 변수들의 값을 변환하여 해당 요인의 값을 추정하여 변환한 값
- ◆ 고유치 (Eigenvalue): 한 요인에 대한 요인부하량의 제곱 (요인과 요인에 해당하는 변수와의 상관계수의 제곱합)으로 계산됨. 한 요인에 의해 설명되는 분산량을 나타내는 값. 값이 크면 요인이 해당 변수를 잘 설명한다는 의미. 잠재근 (Latent root)라고도 함
- ◆ 공통 분산(Communality): 요인에 의해 설명될 수 있는 변수의 분산의 백분율
- ◆ 스크리도표 (Scree plot): 요인 중 고유치가 높은 순으로 정렬하여 요인 (X축)과 고유치(Y축)를 그린 그림

요인분석의 순서

1. 연구문제의 확인
2. 요인분석의 가능성 확인
3. 요인추출방식 선택
4. 요인 수의 결정
5. 요인 회전방법 선택
6. 요인확정과 요인이름 부여
7. 타 연구에 활용 (Optional)

요인분석의 순서

1. 연구문제의 확인

- 탐색적인 요인분석 (Exploratory Factor Analysis)의 대상인지 확인적 요인분석 (Confirmatory Factor Analysis)의 대상인지 결정

2. 요인분석의 가능성 확인

- a. 상관행렬: 일정 수준 이상이어야 함
- b. Bartlett's test of sphericity를 확인 (가설 검정)
 - ❖ H_0 : 상관행렬이 단위행렬임 (변수간 상관관계가 없음)
 - ❖ H_1 : 상관행렬이 단위행렬이 아님 (변수간 상관관계가 있는 경우가 있음)
- c. Kaiser-Meyer-Olkin(KMO)의 MSA(Measure of Sampling Adequacy)를 확인
 - ❖ 표본이 적합한지를 요인분석에 적합한지를 나타내는 수치

MSA	해석
> 0.9	아주 적합
> 0.7	상당히 적합
> 0.6	다소 적합
> 0.5	어느 정도 적합
< 0.5	부적합

요인분석의 순서

d. Communality 값을 살펴봄.

❖ Communality는 추출된 요인이 각 변수를 설명해 주는 정도를 의미하는 것으로 보통 0.5 이상의 값을 가지는 것을 권장

e. Scree plot을 그려봄

❖ 크리도표 상에 고유값의 감소가 확연히 줄어드는 구간이 있는지 확인함. 스크리 도표 상에 확연하게 줄어드는 구간이 존재하면 그 직전까지의 요인 수로 분석을 고려해 볼 수 있음

3. 요인추출방식의 선택

●PCA(Principle Component Analysis), CFA(Common Factor Analysis), ML(Maximum Likelihood)와 GLS(Generalized Least Square) 등이 있다. 주로 쓰이는 방식은 PCA와 CFA임

요인분석의 순서

4. 요인 수의 결정

- 사전연구 기준: 사전 연구의 결과에서 나온 요인 수를 이용
- 고유치(eigen value) 기준: 특정 고유치를 넘은 요인의 수로 정하는 방법. 보통 고유치가 1이 넘는 경우를 많이 사용함. (하나의 개별 요인이 최소한 하나의 변수가 가지는 분산을 설명할 수는 있어야 한다는 생각이나 절대적인 기준은 아님)

Total Variance Explained

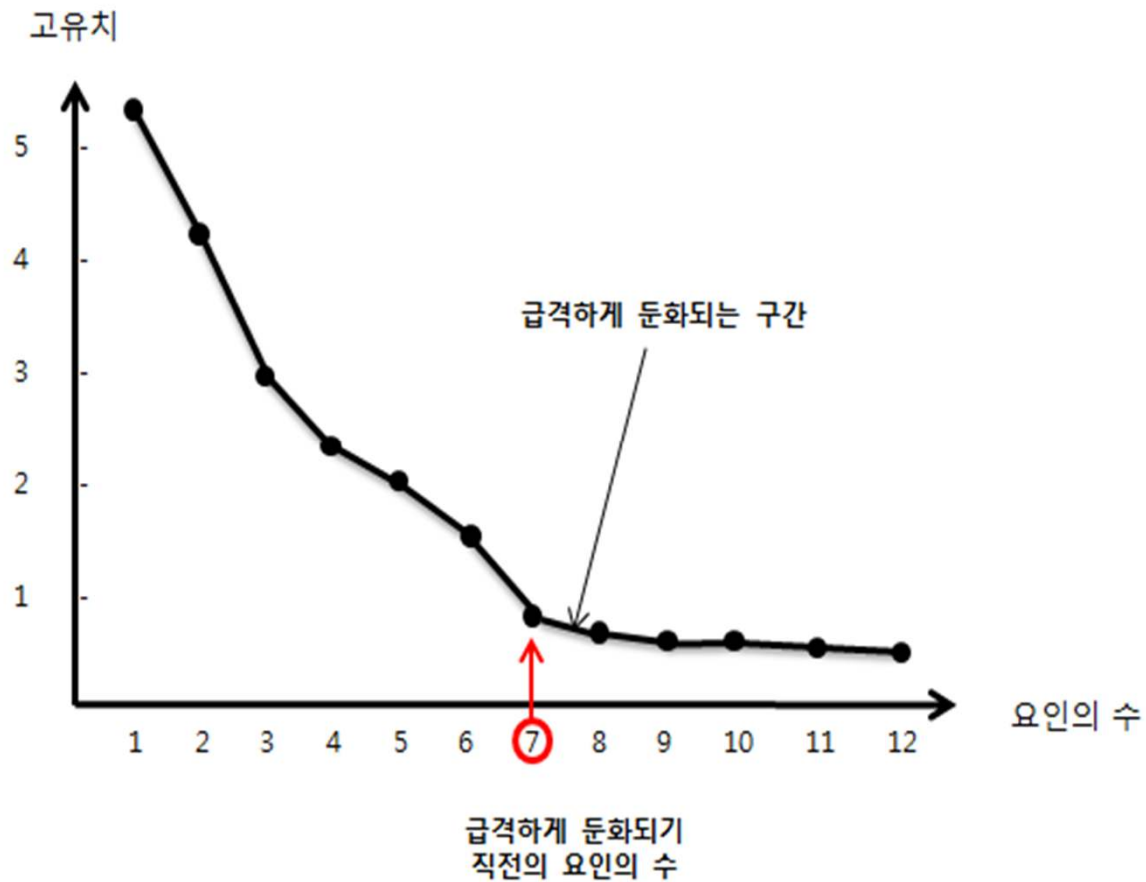
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.832	38.319	38.319	3.832	38.319	38.319
2	2.393	23.928	62.247	2.393	23.928	62.247
3	1.824	18.237	80.484	1.824	18.237	80.484
4	.757	7.571	88.055			
5	.621	6.213	94.268			
6	.288	2.878	97.146			
7	.163	1.633	98.779			
8	.074	.744	99.523			
9	.033	.331	99.854			
10	.015	.146	100.000			

Extraction Method: Principal Component Analysis.

- 스크리 도표 기준: 고유치의 감소가 급격히 둔화되기 직전까지의 후보 요인만을 요인으로 정하는 방법
- 요인 수 기준: 연구자가 원하는 요인의 수를 지정하는 방법
- 누적 분산비율 기준: 총분산의 누적비율이 일정한 값 이상을 만족시키는 요인들만을 요인으로 추출하는 방법

요인분석의 순서

- 참고: Scree Plot (스크리 도표)

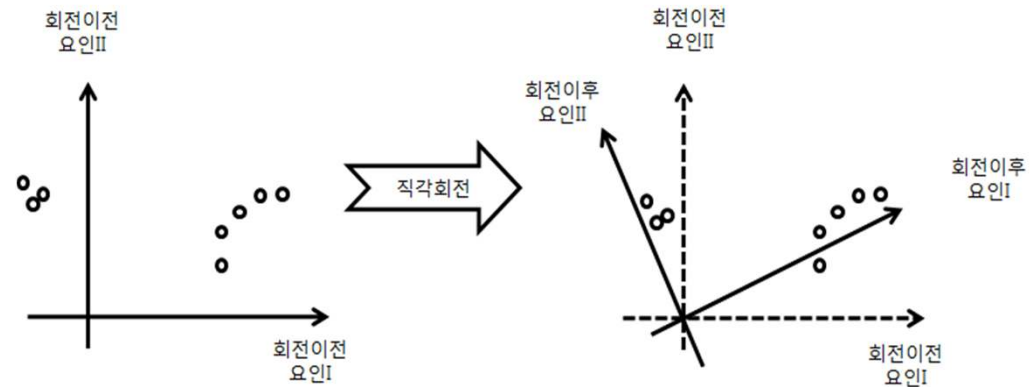


요인분석의 순서

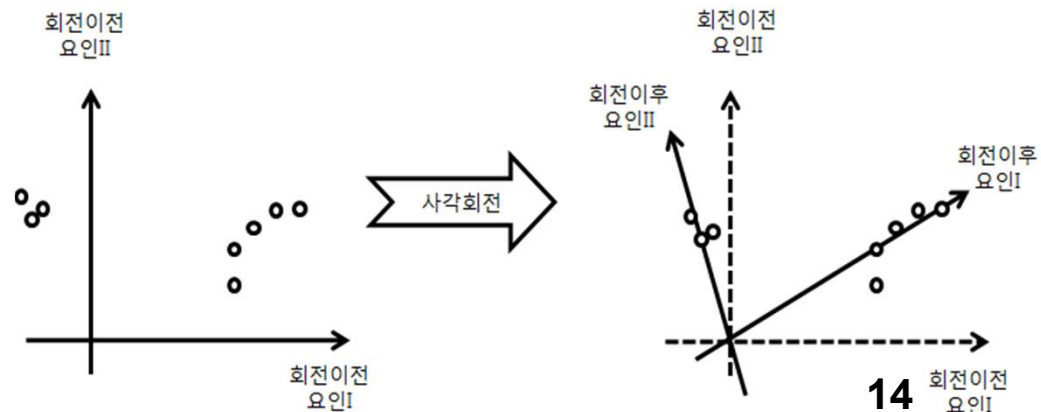
5. 요인 회전방법의 선택

- 요인을 나타내는 축을 회전하여 해당 요인의 상관계수는 최대화 하고 동시에 다른 요인들과의 상관계수는 최소화하는 방법
- 회전 방식은 크게 직각회전(Orthogonal rotation)과 사각회전 (Oblique rotation)으로 나뉨

- 직각회전: 요인의 축이 서로 직각을 유지하면서 회전, 요인간 독립임, 요인간 다중공선성 없음



- 사각회전 : 요인의 축이 서로 직각이 아님, 요인간 독립아님, 요인간 다중공선성 있음, 요인부하량이 최대가 되게 함



요인분석의 순서

6. 요인확정과 요인이름 부여

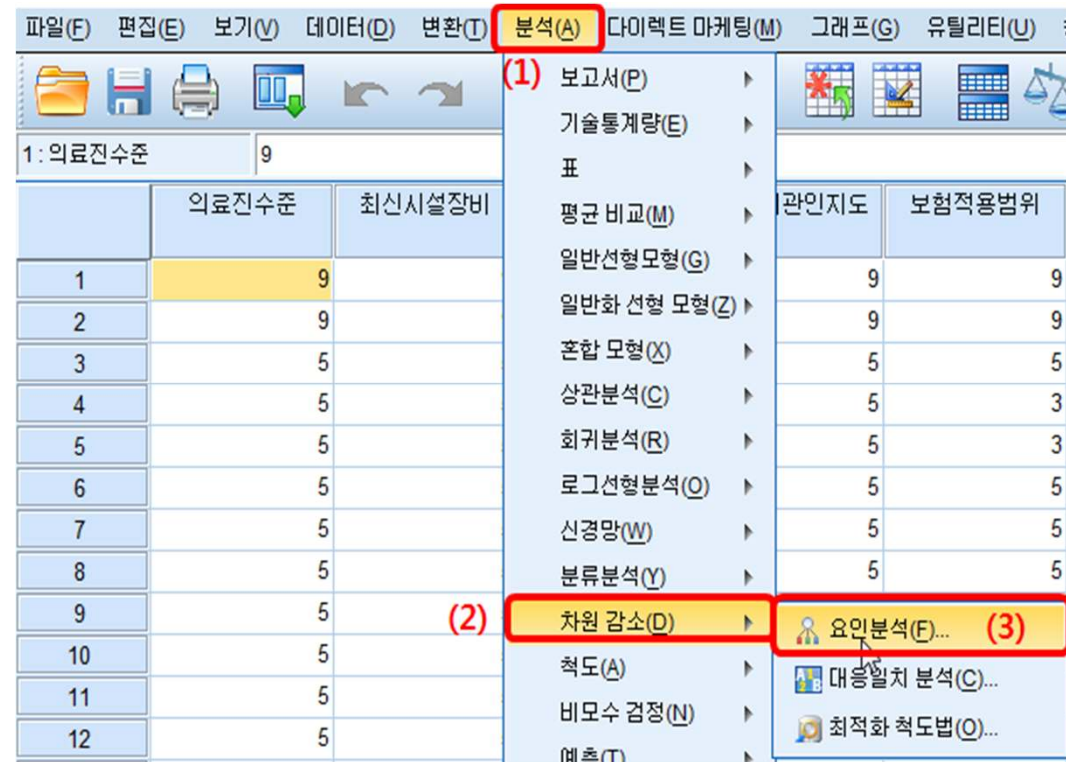
- 요인과 변수간의 요인부하량을 보고 높은 요인부하량을 가진 요인과 변수를 묶으면 됨
- 하나의 요인으로 묶인 변수들을 보고 변수들을 공통적으로 잘 설명할 수 있는 요인의 이름을 명명하면 됨

7. 타 연구에 활용 (Optional)

- 변수의 수를 줄이고 싶은 경우 요인으로 묶은 후에 요인점수를 이용하여 회귀분석 등에 활용할 수 있음
- 특히, 직각회전의 경우 요인점수간 다중공선성이 완전히 제거되는 장점이 있음

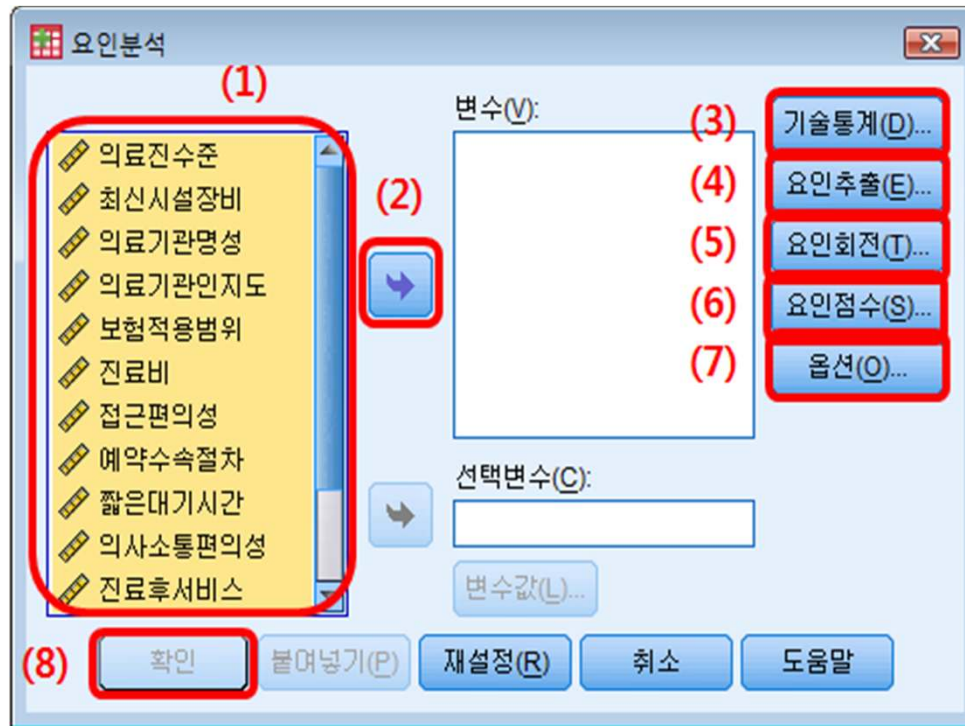
SPSS 이용과 보고서 작성

- ◆ 총 15개의 변수로 이루어진 국내 의료서비스 현황 및 만족도 조사 데이터를이용 (factor.xls)
- ◆ 총 15개의 변수를 요인분석을 통해 5개의 요인 나눈 결과 ‘일괄진료상품’은 다른 변수와 묶이지 않고 단독으로 존재함
- ◆ ‘일괄진료상품’을 제외한 14개의 변수를 이용하여 4개의 요인을 찾는 요인분석을 실시



SPSS 이용과 보고서 작성

◆ 변수선택

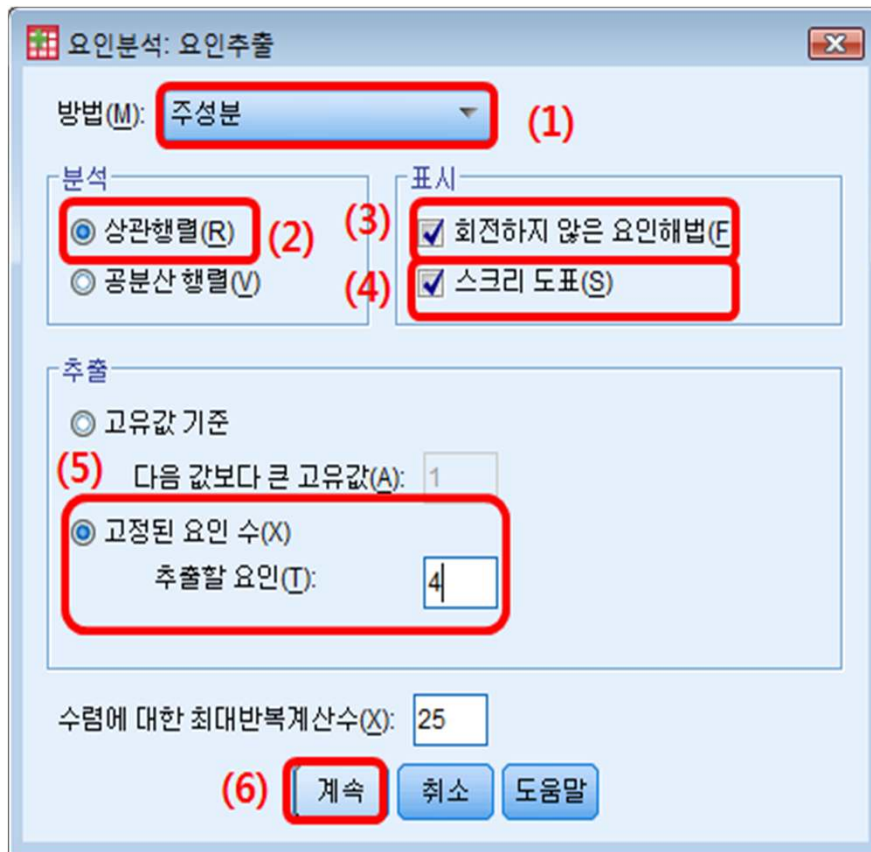


• 기술통계량 출력항목 선택



SPSS 이용과 보고서 작성

◆ 요인추출 방법과 요인수의 결정



요인분석: 요인추출

방법(M): 주성분 (1)

분석

☒ 상관행렬(R) (2) (3)

☐ 공분산 행렬(V) (4)

표시

☒ 회전하지 않은 요인해법(F)

☒ 스크리 도표(S)

추출

☐ 고유값 기준

(5) 다음 값보다 큰 고유값(A): 1

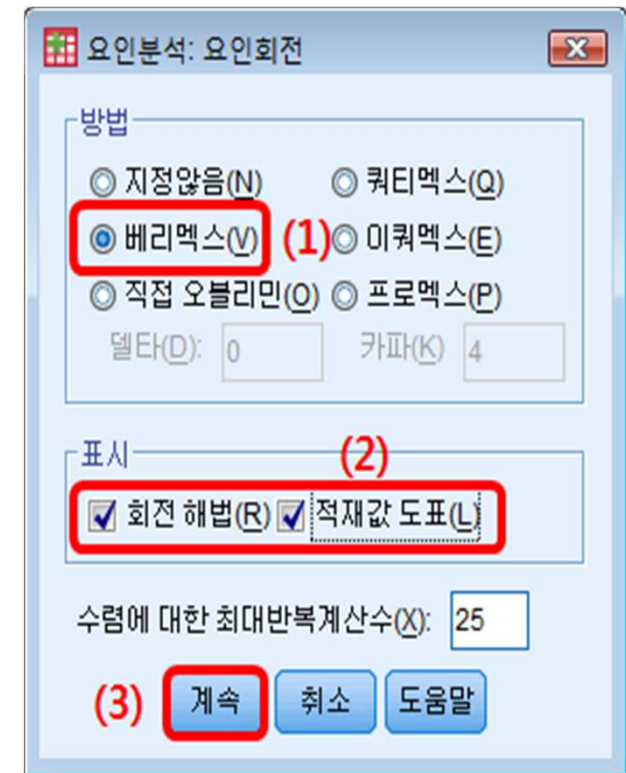
☒ 고정된 요인 수(X)

추출할 요인(I): 4

수렴에 대한 최대반복계산수(X): 25

(6)

• 요인회전 방법의 선택



요인분석: 요인회전

방법

☐ 지정없음(N)

☒ 베리맥스(V) (1)

☐ 쿼티맥스(Q)

☐ 이쿼맥스(E)

☐ 직접 오블리민(O)

☐ 프로맥스(P)

델타(D): 0 카파(K): 4

표시 (2)

☒ 회전 해법(R)

☒ 적재값 도표(L)

수렴에 대한 최대반복계산수(X): 25

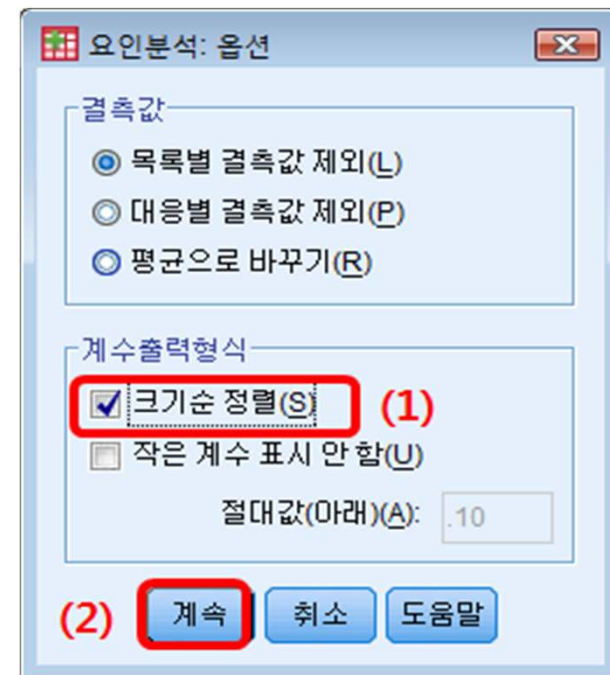
(3)

SPSS 이용과 보고서 작성

◆ 요인점수의 변수 저장



• 요인분석의 옵션 지정



SPSS 이용과 보고서 작성: 연구문제

◆ 연구 문제

- 한국의 의료관광상품을 평가하기 위해 5점 Likert scale로 설문을 구성함. 14개의 변수를 요인으로 재구성하여 변수의 구조를 이해하려고 함
- 연구문제 1. 요인분석이 가능한가?
- 연구문제 2. 어떤 변수가 요인으로 묶이는가? 변수의 구조는 어떠한가?

◆ 연구 대상

- 한국 관광공사(2008)의 설문 조사에서 한국 의료관광경험자 중 한국 의료서비스 이용 후 만족도 설문 항목에 응답한 140명을 대상으로 함

◆ 변수

변수명	코딩명	변수설명	척도
의료진수준, 최신시설장비, 의료기관명성, 의료기관인지도, 보험적용범위, 진료비, 접근편의성, 예약수속절차, 짧은대기시간, 의사소통편의성, 진료후서비스, 의료관광상품의다양성, 환자지원인력의 수준, 보험사연계	변수명과 동일	한국의 관광상품 평가 항목	등간 척도

◆ 분석 방법

- 요인분석을 실시, 요인분석의 가능성을 알기 위해 Kaiser-Meyer-Olkin 측도를 통해 표준형성 적절성 검정과, Bartlett의 구형성 검정을 실시. 요인분석이 가능하다고 판단되면 요인을 도출하고 변수와 해당요인에 적절한 요인명을 부여

SPSS 이용과 보고서 작성: 분석결과

- ◆ 요인분석의 가능성을 살펴보기 위해 다음의 값들의 확인이 필요하다.
 - Kaiser-Meyer-Olkin Measure: 0.938로 높게 나타나서 요인분석에 아주 적합하다.
 - Bartlett의 구형성 검정 : 다음의 가설을 검정한다.
 - ❖ H_0 : 모집단의 상관행렬은 단위행렬이다.
 - ❖ H_1 : 모집단의 상관행렬은 단위행렬이 아니다.
 - ❖ Bartlett의 구형성 검정의 p -value가 0.00이하로 나타나서 귀무가설을 기각한다. 상관행렬이 단위행렬이 아니므로 변수간의 상관관계가 존재하고, 이는 변수들이 요인으로 묶일 수 있다는 가능성을 보여주고 있다.
 - Communality 의 권장수준 만족: 추출된 communality가 0.8이상으로 권장수준인 0.5를 만족시키고 있다.
- ◆ **종합적으로 볼 때 요인분석이 가능하다.**

SPSS 이용과 보고서 작성: 분석결과

◆ 회전된 성분행렬을 통한 요인추출

변 수	성분			
	1	2	3	4
의료기관 명성	0.817	0.339	0.275	0.210
의료진 수준	0.803	0.263	0.258	0.317
최신 시설장비	0.786	0.289	0.330	0.306
의료기관 인지도	0.627	0.355	0.463	0.329
보험사 연계	0.313	0.787	0.374	0.152
보험 적용범위	0.391	0.780	0.112	0.377
의료관광 상품의 다양성	0.224	0.738	0.510	0.116
진료비	0.349	0.624	0.175	0.593
진료 후 서비스	0.217	0.611	0.575	0.252
환자 지원인력의 수준	0.488	0.299	0.733	0.141
짧은 대기시간	0.416	0.370	0.657	0.390
예약 수속절차	0.330	0.353	0.625	0.479
접근 편의성	0.534	0.219	0.287	0.674
의사소통 편의성	0.452	0.307	0.455	0.588

SPSS 이용과 보고서 작성: 분석결과

◆ 추출된 요인의 이름 부여

변수	도출된 요인
의료기관 명성	의료의 질
의료진 수준	
최신 시설장비	
의료기관 인지도	
보험사 연계	서비스의 경제성
보험 적용범위	
의료관광 상품의 다양성	
진료비	
진료 후 서비스	
환자 지원인력의 수준	행정지원
짧은 대기시간	
예약 수속절차	
접근 편의성	이용편의
의사소통 편의성	

신뢰도 분석 (RELIABILITY ANALYSIS)

- ◆ 신뢰도: 하나의 대상을 반복 측정하거나 유사한 도구로 여러 번 측정하는 경우 일관성 있는 결과가 산출되는 정도
- 통계적으로는 반복 측정된 측정값의 분산을 의미하며 신뢰도가 높은 경우 분산이 낮아서 일관성이 높은 결과가 산출됨을 의미함
 1. 재측정 신뢰도: 동일한 설문을 동일한 사람에게 반복하여 제시하고 얻은 결과의 일관성을 측정하는 방법
 2. 반분 신뢰도: 동일한 개념을 측정하는 여러 개의 문항을 만든 후에 측정문항을 무작위로 반분하여 두 집단에게 배포한 후 응답간의 상관관계를 분석하여 신뢰도를 구하는 방법
 3. 크론바흐 알파: 각 문항별 반분신뢰도를 평균한 값이다. 일반적으로 설문의 신뢰도를 구할 때 가장 많이 사용되는 분석방법. 구하는 공식은 다음과 같음

$$\text{Cronbach's } \alpha = \frac{N}{N-1} \left(1 - \sum_{i=1}^N \frac{\sigma_i^2}{\sigma_{total}^2} \right)$$

N 은 문항 수, σ_{total}^2 은 총분산, σ_i^2 는 각 문항의 분산

- ❖ 계수가 0.5이상이면 신뢰도가 있는 것으로 간주되고 0.7 이상이면 신뢰도가 높은 것으로 간주

크론바흐 알파의 계산

◆ 요인분석의 설문 응답 140부의 크론바흐 알파를 계산

	평균	표준편차
문항1	4.41	2.405
문항2	4.71	2.123
문항3	4.96	2.448
문항4	4.71	2.535
문항5	5.03	2.202
Total	23.82	10.506

$$Cronbach's \alpha = \frac{5}{5-1} \left\{ 1 - \frac{2.405^2 + 2.123^2 + 2.448^2 + 2.535^2 + 2.202^2}{10.506^2} \right\} = 0.937898$$

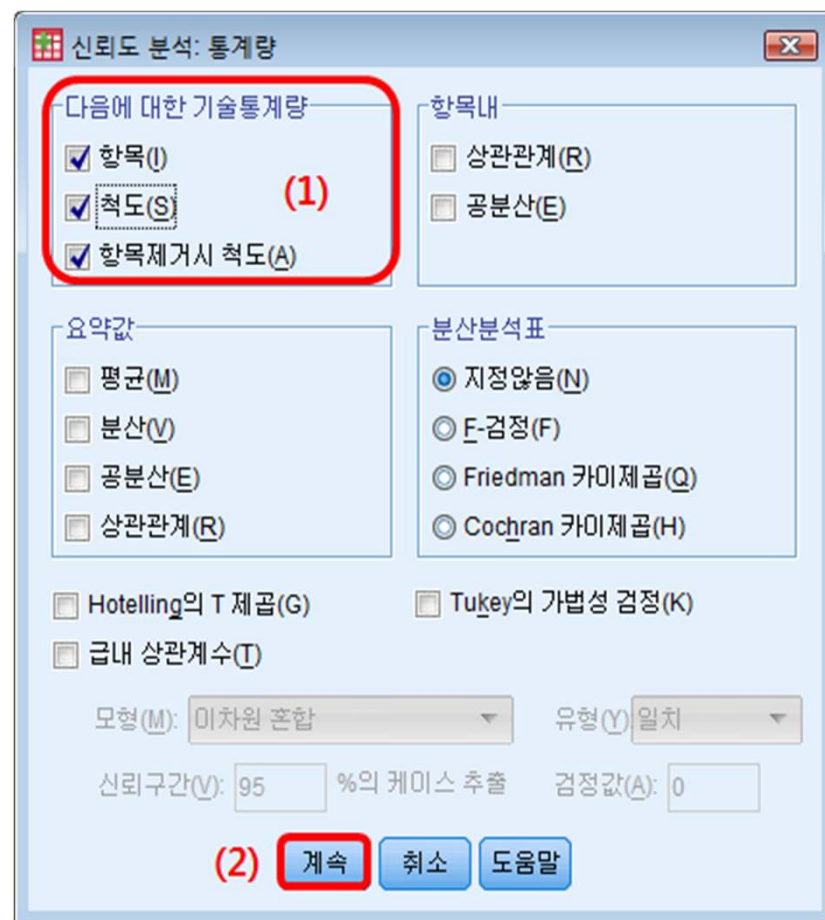
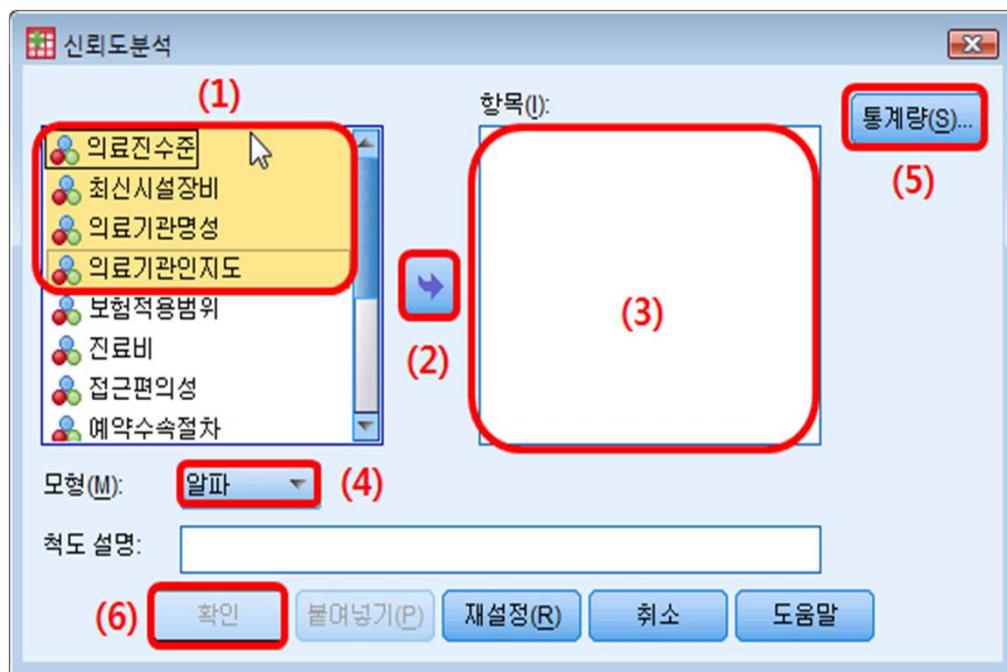
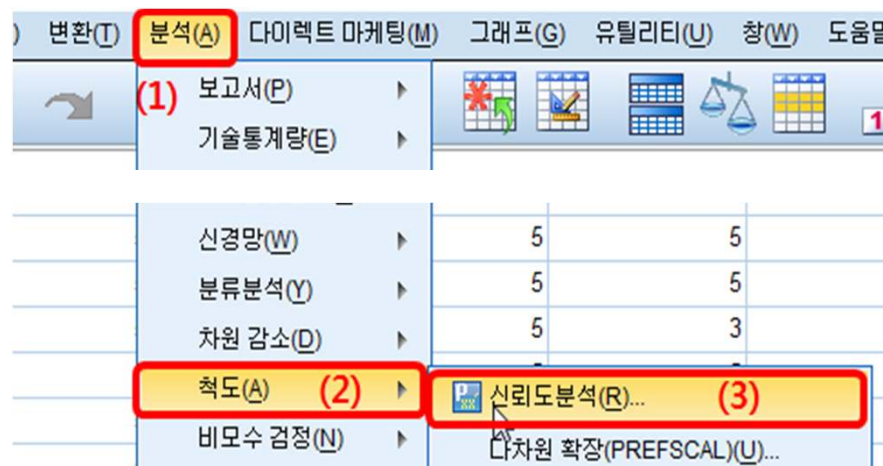
◆ 신뢰도 분석은 요인분석과 함께 사용됨

- 요인분석에 사용된 변수들이 하나의 요인으로 묶인 경우 변수들 사이에는 일관성이 존재해야 함
- 따라서 요인으로 묶인 변수들의 크론바흐 알파값이 권장 수준이상인지 확인하여 변수간의 설문 신뢰도가 있는지를 확인하는 절차가 필요함

신뢰도의 향상

- ◆ 신뢰도가 낮은 경우: 일관성이 낮은 문항을 제거하고 나머지 문항의 일관성이 향상됨.
- 문항의 제거 시 반드시 필요한 문항이었는지 제거되어도 무방한지 살펴보아야 함
- 만일 제거가 힘들다면 유사한 내용을 묻는 다른 문항으로 대체하거나 관련 연구에서 제시된 문항으로 대체하거나 응답자가 설문 내용을 잘 이해하고 답한 것인지 살펴보아야 함

SPSS 이용과 보고서 작성



SPSS 이용과 보고서 작성

요 인	변 수	Cronbach's α
의료의 질	의료기관 명성	.949
	의료진 수준	
	최신 시설장비	
	의료기관 인지도	
서비스의 경제성	보험사 연계	.938
	보험 적용범위	
	의료관광 상품의 다양성	
	진료비	
	진료 후 서비스	
행정지원	환자 지원인력의 수준	.924
	짧은 대기시간	
	예약 수속절차	
이용편의	접근 편의성	.877
	의사소통 편의성	

제거항목	항목이 삭제된 경우 Cronbach 알파
의료진의 수준	.935
최신시설 및 장비	.926
의료기관의 명성	.926
의료기관 인지도	.945