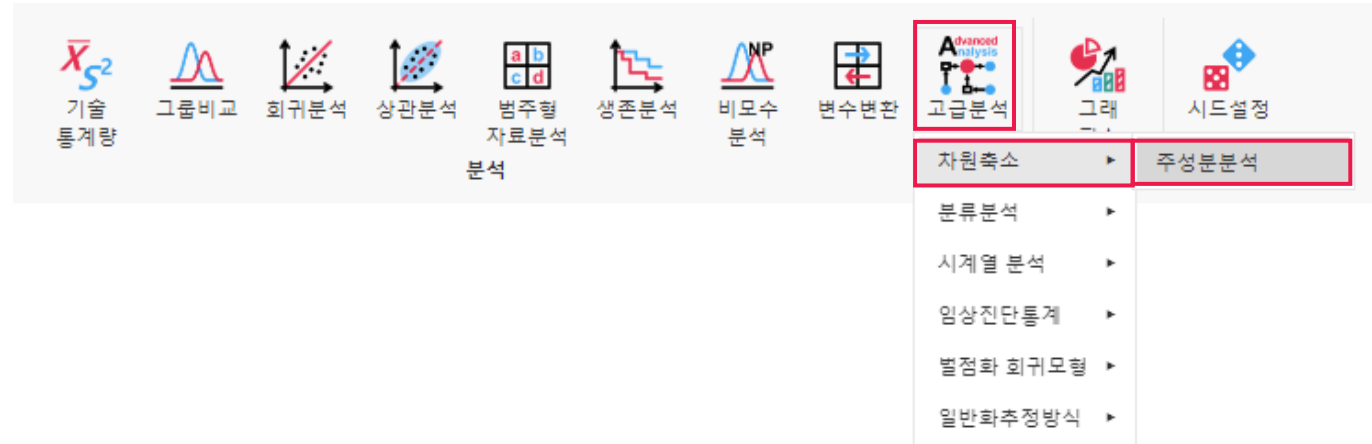


주성분분석

메뉴 호출하기

- 고급분석 > 차원축소 > 주성분분석



주성분분석(Principal Component Analysis, PCA)은 데이터들의 주성분을 찾는 모듈입니다. 서로 상관성이 높은 여러 변수들의 선형조합으로 만든 새로운 변수들로 요약 및 축소하는 분석방법입니다. 요인 분석과 달리, 목표 변수를 고려하여 잘 예측하고 분류하기 위해서 기존 변수들의 선형 결합으로 이루어진 주성분을 찾는 분석 방법입니다.

• 변수설정 탭

주성분분석

변수설정 출력옵션

① 입력 데이터 형식 (데이터 외의 경우 작업기록 기능에서 제외)

☒ 데이터
 ☐ 공분산행렬
 ☐ 상관행렬
 ☐ 비유사행렬

데이터

전체 변수

id
bweight
lowbw
gestwks
preterm
matage
hyp
sex

② 선택변수(2개이상필수)

>

<

③ ☒ 분산의 표준화 (Scaling to have unit variance)

도움말 재설정

확인 취소

메뉴 요소	설명
① 입력 데이터 형식	<p>데이터와 상관행렬, 공분산행렬, 비유사행렬 네 가지 중 하나를 택합니다.</p> <ul style="list-style-type: none"> 데이터 (Default) : 엑셀 스프레드 시트에 있는 데이터에서 변수를 택하여 분석하고자 할 때 선택합니다. 상관행렬 : 데이터가 상관행렬인 경우에 선택합니다. 모든 데이터가 입력 값으로 활용됩니다. 공분산행렬 : 데이터가 공분산행렬인 경우에 선택합니다. 모든 데이터가 입력 값으로 활용됩니다. 비유사행렬 : 데이터가 비유사행렬인 경우에 선택합니다. 모든 데이터가 입력 값으로 활용됩니다.
② 선택변수	<p>[입력 데이터 형식]이 '데이터'인 경우에만 선택변수를 지정할 수 있습니다. 주성분분석에 사용할 변수를 선택합니다. 적어도 두 개 이상의 양적변수를 선택해야 합니다.</p>
③ 분산의 표준화	<p>[입력 데이터 형식]에서 '데이터' 혹은 '공분산행렬'를 선택할 경우 활성화됩니다. 선택된 각 변수의 분산을 1로 표준화 합니다.</p> <p>선택변수의 측정단위가 다르거나, 변수 간의 변동성을 비교하는 방법이 불분명할 때 표준화를 실시하면 더 바람직한 결과 해석이 가능합니다.</p>

- 출력옵션 탭

주성분분석

변수설정 출력옵션

저장

① ☒ 주성분점수 (Coordinates of Observations)
주성분의 수

② ☒ 부하행렬 (Loadings/Coordinates of Variables)
저장경로

그래프

③ ☐ 스크리도표

④ ☒ 행렬도 (Biplot)
차원의 수

메뉴 요소	설명
1. 주성분점수 (Coordinates of Observations)	<p>주성분의 표준편차, 분산의 기여도, 누적 기여율과 각 주성분의 고유값, 고유벡터를 출력합니다. 각 주성분이 변수명으로 저장됩니다.</p> <ul style="list-style-type: none"> 주성분의 수 : [주성분점수]를 선택할 경우 활성화됩니다. 출력할 주성분의 수를 정할 수 있습니다. 자연수만 입력이 가능하며, Default는 2입니다.
2. 부하행렬(Loadings/Coordinates of Variables)	<p>주성분 식의 가중치를 확인합니다. 부하행렬을 선택할 경우 저장경로가 활성화됩니다.</p> <ul style="list-style-type: none"> 저장경로 : [부하행렬]을 선택할 경우 활성화됩니다. 부하행렬을 지정한 저장경로에 저장합니다.
3. 스크리도표	<p>스크리 도표를 출력합니다. 스크리 도표로 각 주성분의 비중을 파악하여 적절한 주성분의 개수를 정할 수 있습니다.</p>
4. 행렬도(Biplot)	<p>행렬도를 출력합니다. 행렬도는 주성분분석으로 찾은 주성분과 원데이터의 변수 간의 상관관계를 보여줍니다.</p> <ul style="list-style-type: none"> 차원의 수 : '행렬도 (Biplot)'를 선택하면 활성화됩니다. 행렬도의 차원을 지정합니다. 2 이상의 자연수만 입력이 가능하며 Default는 2입니다.