

Análisis de componentes principales

Práctica

Pedro López-Roldán

Departamento de Sociología

Centro de Estudios Sociológicos sobre la Vida Cotidiana y el Trabajo (QUIT)

Instituto de Estudios del Trabajo (IET)

Universidad Autónoma de Barcelona

Pedro.Lopez.Roldan@uab.cat



Sandra Fachelli

Departamento de Sociología

Universidad de Barcelona

Grupo de Investigación en Educación y Trabajo (GRET)

Universidad Autónoma de Barcelona

Sandra.Fachelli@ub.edu



Exploración del 1er Modelo de Estratificación

Reproducción lo más fiel posible del modelo que el Arquitecto Torres había hecho para el Censo de Población y Vivienda de 1991

Material Docente
Práctica ACP
Análisis de Componentes
Principales: un ejemplo de la
Ciudad Autónoma de
Buenos Aires.



Sandra Fachelli
Departamento de Sociología y
Análisis de las Organizaciones
Universidad de Barcelona
Grupo de Sociología Empírica
Sandra.Fachelli@ub.edu

Grupo de Investigación en Educación
y Trabajo (GRET)
Universidad Autónoma de Barcelona



Pedro López-Roldán
Departamento de Sociología
Centro de Estudios Sociológicos sobre
la Vida Cotidiana y el Trabajo (QUIT)
Instituto de Estudios del Trabajo (IET)
Universidad Autónoma de Barcelona
Pedro.Lopez.Roldan@uab.cat



Link: <http://ddd.uab.cat/record/169929> o
<http://pagines.uab.cat/plopez/sites/pagines.uab.cat/plopez/files/ACP-Torres.pdf>

ACP: Un ejemplo sobre la Ciudad Autónoma de Buenos Aires. 1ra Parte

Los autores del Material Docente han recibido el encargo de elaborar una muestra estratificada utilizando el Censo de Población y Vivienda de 2010 de Argentina.

La primera parte del análisis consiste en reproducir para 2010 un análisis que el Arquitecto Torres realizó para la Ciudad de Bs. As. con el Censo de 1991.

Ejercicio 1:

Se pretende que reproduzcas el análisis seleccionado las variables definidas por el modelo que se presenta a continuación. Realiza el ACP e interpreta los resultados.

Si el análisis arroja varios factores interpreta los dos primeros.

Seleccione la base: • [CABA-Censo2010Torres.sav](#)

<http://pagines.uab.cat/plopez/sites/pagines.uab.cat.plopez/files/CABA-Censo2010Torres.sav>

Si necesita corroborar el análisis seleccione la sintaxis • [ACP-Torres.sps](#)

<http://pagines.uab.cat/plopez/sites/pagines.uab.cat.plopez/files/ACP-Torres.sps>

Material Docente
Práctica ACP
Análisis de Componentes
Principales: un ejemplo de la
Ciudad Autónoma de
Buenos Aires.



Sandra Fachelli
Departamento de Sociología y
Análisis de las Organizaciones
Universidad de Barcelona
Grupo de Sociología Empírica
Sandra.Fachelli@ub.edu
Grupo de Investigación en Educación
y Trabajo (GRET)
Universidad Autónoma de Barcelona



Pedro López-Roldán
Departamento de Sociología
Centro de Estudios Sociológicos sobre
la Vida Cuotidiana y el Trabajo (QUIT)
Instituto de Estudios del Trabajo (IET)
Universidad Autónoma de Barcelona
Pedro.Lopez.Roldan@uab.cat



1) Selecciona las siguientes variables y realiza el ACP

| Reproducción Modelo Torres | |
|--|-------------------------------------|
| A. Hacinamiento (más de dos personas por cuarto) | |
| H12_Hasta0.50personas1* | H12_1.501.99personas1 |
| H12_0.510.99personas1 | H12_2.003.00personas1 |
| H12_1.001.49personas1 | H12_Másde3.00personas1 |
| B. Tipo de tenencia de vivienda | D. Origen |
| V9_Propietariodelaviviendaydelterreno1 | P3_Argentina1* |
| V9_Propietariosólordelavivienda1 | P3_Otropaís1 |
| V9_Inquilino1 | E. Tipo de vivienda |
| V9_Ocupanteporpréstamo1 | V2_Casa1* |
| V9_Ocupanteportrabajo1 | V2_Rancho1 |
| V9_Otrasituación1* | V2_Casilla1 |
| C. Educación | V2_Departamento1 |
| P5_Nuncaasistió1 | V2_Piezaeninquilinato1 |
| P6_Primary1 | V2_Piezaenhotelfamiliaropensión1 |
| P6_Secundario1 | V2_Localnoconstruidoparahabitación1 |
| P6_Superiornouniversitario1 | V2_Viviendamóvil1 |
| P6_Universitario1 | V2_Personasviviendoenlacalle1 |
| * Variables eliminadas para evitar combinación lineal perfecta | |

Variable H12:
Escala con nivel
de hacinamiento
del hogar

Hacinamiento
extremo: 3 personas
por cuarto
(sin contar baño y
cocina)

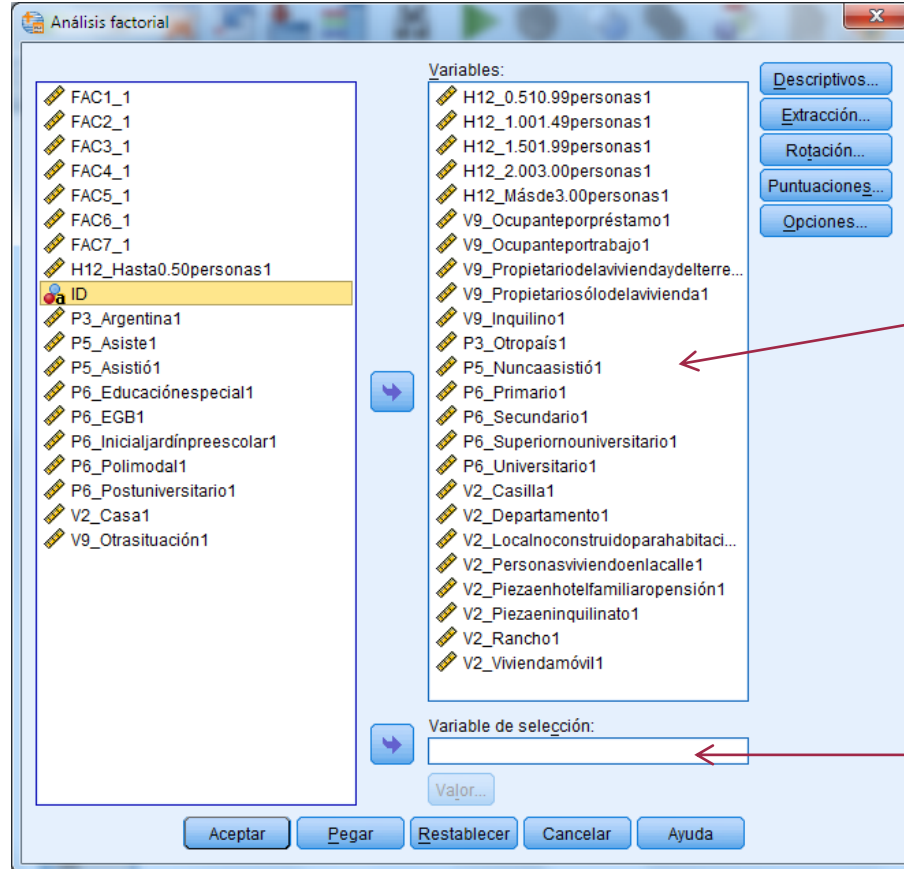
Atención
con esta
consigna:

Casas muy
precarias
(tipo barraca)

Inquilinato: vivienda
precaria donde
viven varias familias
alquilando
habitaciones.

Análisis Factorial con SPSS

Analizar... /
Reducción de Dimensiones... /
Factor...



Variables del análisis

Variable para la selección de casos, según el valor de un número entero
Ej: variable provincia=3

Análisis Factorial con SPSS

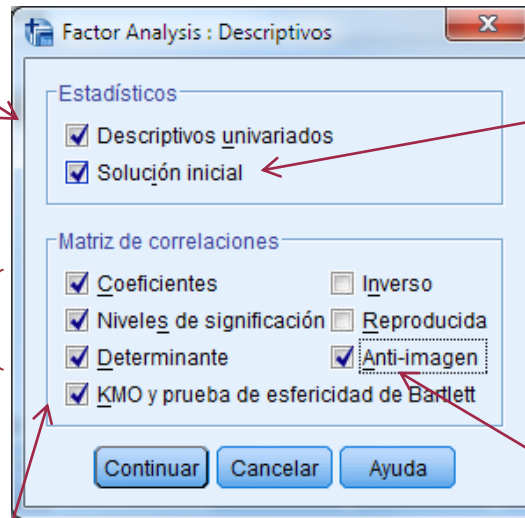
Descriptivos

Media, desviación típica y número de **casos** válidos para cada variable

Matriz de correlaciones con la significación y el determinante (tiene que ser próximo a cero pero no cero porque eso sería combinación lineal perfecta y genera una indeterminación en el cálculo)

La **medida de adecuación muestral** de Kaiser-Meyer-Olkin contrasta si las correlaciones parciales entre las variables son pequeñas.

La **prueba de esfericidad de Bartlett** contrasta si la matriz de correlaciones es una matriz identidad, que indicaría que el modelo factorial es inadecuado



La **solución inicial** muestra las communalidades iniciales, los autovalores y el porcentaje de varianza explicada

Matriz de correlaciones **anti-imagen** contiene los negativos de los coeficientes de correlación (covarianza) parcial. La mayoría de los elementos no diagonales deben ser pequeños. La diagonal muestra la **medida de adecuación muestral** de cada variable

Análisis Factorial con SPSS

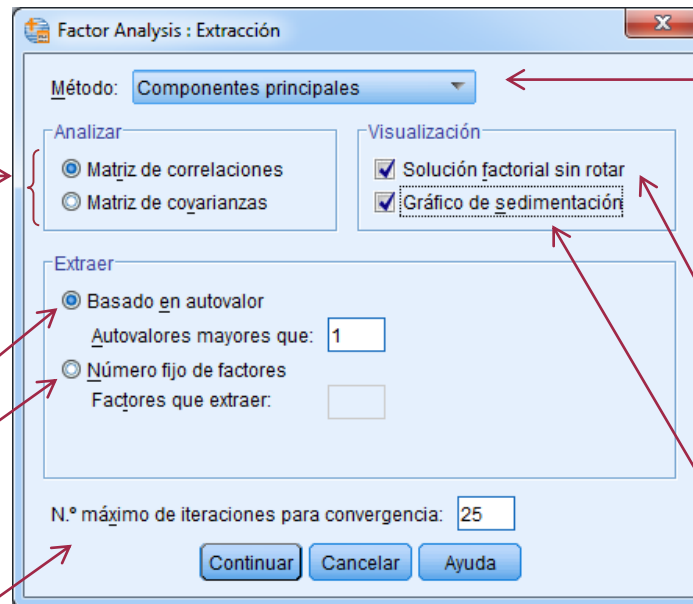
Extracción

Permite elegir la matriz de **Correlaciones** (variables con escalas distintas) o de **Covarianzas** (para varios grupos con varianza distintas para cada variable)

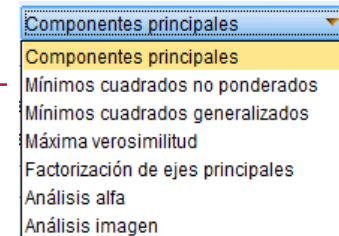
Nº de factores retenidos

- a) autovalores > 1 u otro valor
- b) Se especifica el nº

Permite especificar el número máximo de pasos que el algoritmo puede seguir para estimar la solución



Métodos de análisis factorial



Muestra las saturaciones factoriales sin rotar, las comunalidades y los autovalores de la solución factorial

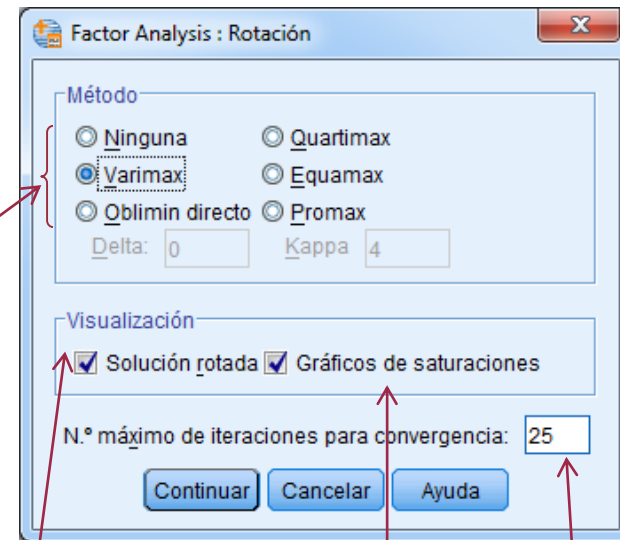
Gráfico de la varianza que se asocia a cada factor.

Análisis Factorial con SPSS

Rotación

Selección del método de rotación factorial

- **Método varimax.** Rotación ortogonal que minimiza el número de variables que tienen saturaciones altas en cada factor. Simplifica la interpretación de los factores.
- **Criterio Oblimin directo.** Rotación oblicua (no ortogonal). Si delta es igual a cero las soluciones son las más oblicuas. A medida que delta se va haciendo más negativo, los factores son menos oblicuos. Se puede introducir un número menor o igual que 0,8.
- **Método quartimax.** Rotación que minimiza el número de factores necesarios para explicar cada variable. Simplifica la interpretación de las variables observadas.
- **Método equamax.** Método de rotación que es combinación del método varimax, que simplifica los factores, y el método quartimax, que simplifica las variables. Se minimiza tanto el número de variables que saturan alto en un factor como el número de factores necesarios para explicar una variable.
- **Rotación Promax.** Rotación oblicua que permite que los factores estén correlacionados. Esta rotación se puede calcular más rápidamente que una rotación oblimin directa, por lo que es útil para conjuntos de datos grandes.

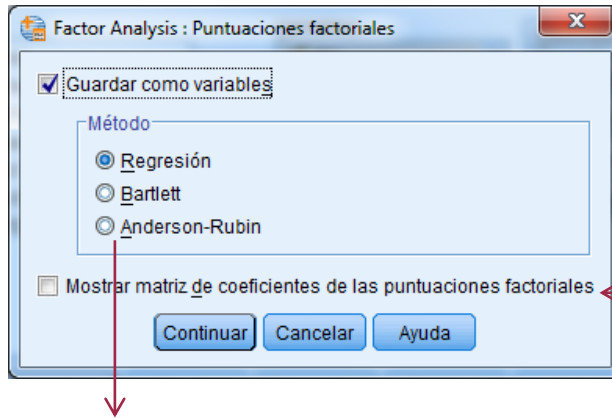


Para rotaciones **ortogonales:** matriz de configuración rotada y matriz de transformación de factor.
Para **oblicuas:** matrices de correlaciones de factor, estructura y patrón

En 2 o 3 dimensiones, según el nº retenido. Si se pide rotación es de la solución rotada
Nº máximo de pasos que el algoritmo puede seguir para llevar a cabo la rotación

Análisis Factorial con SPSS

Puntuaciones factoriales



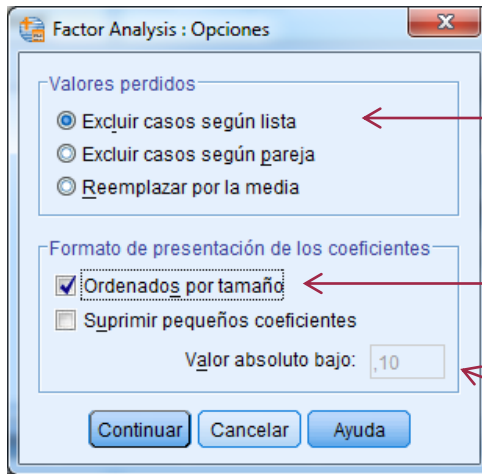
Muestra los coeficientes por los que se multiplican las variables para obtener puntuaciones factoriales. Muestra las correlaciones entre las puntuaciones factoriales

Método alternativo:

- **Método de regresión.** Las puntuaciones tienen media 0 y una varianza igual al cuadrado de la correlación múltiple entre las puntuaciones factoriales estimadas y los valores factoriales verdaderos. Las puntuaciones pueden correlacionarse incluso si los factores son ortogonales.
- **Puntuaciones de Bartlett.** Las puntuaciones tienen media 0. Se minimiza la suma de cuadrados de los factores únicos sobre el rango de las variables.
- **Método de Anderson-Rubin.** Es una modificación del método de Bartlett, que asegura la ortogonalidad de los factores estimados. Las puntuaciones resultantes tienen una media 0, una desviación típica de 1 y no correlacionan entre sí.

Análisis Factorial con SPSS

Opciones



Tratamiento de valores perdidos

- a) Excluye todos los perdidos de cada variable
- b) Excluye los perdidos de las parejas de variables que no puede calcular las correlaciones
- c) Sustituye el valor perdido por la media

Los coeficientes se ordenan por tamaño

Se suprimen los coeficientes cuyos valores absolutos sean menores que el valor especificado

Ejercicio 2

¿Puedes mejorar el modelo?

Analiza la matriz anti-imagen para valorar la eliminación de variables con valores bajos teniendo en cuenta:

- a) Eliminar variables con valores de la “Medida de adecuación muestral” por debajo de 0,5
- b) Que no desmejore el total de varianza explicada
- c) Que no aumente la cantidad de factores

Práctica Adicional:

Selecciona algunas variables con criterio teórico sobre algún tema que le interese de la base del INE-Censo 2001 Barcelona y realice un ACP.

La base está disponible en:

<http://pagines.uab.cat/plopez/sites/pagines.uab.cat/plopez/files/Barcelona.SAV>

Recuerda:

- a) No utilizar el 100% de las categorías de una variable, elimine al menos una para que no haya combinación lineal perfecta.
- b) Analiza la matriz anti-imagen para valorar la eliminación de variables con valores bajos.