

## Laboratorio 2: Análisis de Componentes Principales

KEVIN STEVEN GARCÍA<sup>a</sup>, ALEJANDRO VARGAS<sup>b</sup>, ALEJANDRO SOTO<sup>c</sup>

### 1. Introducción

En este informe se llevará a cabo el análisis de componente principales a una base de datos correspondiente a la “Encuesta de Presupuestos Familiares del año 1990/91” en España. Se cuenta con 51 observaciones y 9 variables. Las observaciones son las provincias españolas más Ceuta y Melilla, que aparecen unidas como una única provincia, y las variables son:  $X_1$  = alimentación,  $X_2$  = vestido y calzado,  $X_3$  = vivienda,  $X_4$  = mobiliario doméstico,  $X_5$  = gastos sanitarios,  $X_6$  = transporte,  $X_7$  = enseñanza y cultura,  $X_8$  = turismo y ocio,  $X_9$  = otros gastos. En este análisis se tomará en cuenta conceptos como el análisis exploratorio de los datos, la nube de individuos y de variables (circulo de correlaciones), la representación simultanea, la construcción e interpretación del índice, los cosenos cuadrados y las contribuciones y las relaciones de transición.

### 2. Análisis exploratorio

En el análisis exploratorio vamos a realizar un análisis descriptivo y gráfico para cada variable y se obtendrá la matriz de correlaciones para posteriormente realizar comparaciones con los resultados del ACP.

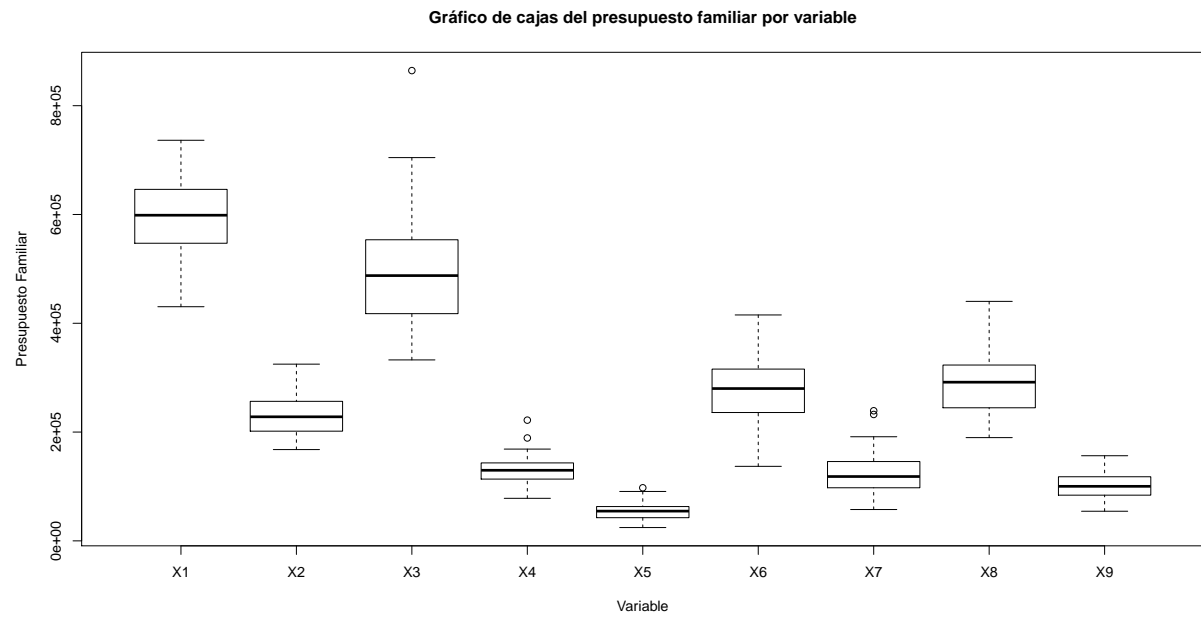
- Estadísticas descriptivas

| Descriptivas    | Alimentación | Vestido y Calzado | Vivienda | Mobiliario doméstico | Gastos sanitarios | Transporte | Enseñanza y cultura | Turismo y ocio | Otros gastos |
|-----------------|--------------|-------------------|----------|----------------------|-------------------|------------|---------------------|----------------|--------------|
| Mínimo          | 430442       | 167814            | 332662   | 78217                | 24476             | 136992     | 57607               | 189811         | 54442        |
| 1st. Cuartil    | 547131       | 201655            | 417691   | 113550               | 42610             | 235941     | 97603               | 244666         | 84019        |
| Mediana         | 598669       | 228072            | 487651   | 129768               | 54684             | 280023     | 118269              | 291708         | 100255       |
| Media           | 600781       | 232043            | 499956   | 129688               | 55104             | 278112     | 125310              | 294558         | 102225       |
| 3rd. Cuartil    | 646183       | 256498            | 553526   | 143317               | 63087             | 315771     | 145831              | 323239         | 117851       |
| Máximo          | 736441       | 324877            | 864553   | 221954               | 97595             | 415313     | 239187              | 440275         | 156493       |
| Desviación Est. | 66004.47     | 37922.05          | 107438   | 26666.27             | 16702.17          | 60014.13   | 40173.34            | 60612.32       | 23548.8      |

<sup>a</sup>Código: 1533173. E-mail: kevin.chica@correounivalle.edu.co

<sup>b</sup>Código: 1525953. E-mail: jose.alejandro.vargas@correounivalle.edu.co

<sup>c</sup>Código: 1532457. E-mail: asotomurillo@gmail.com

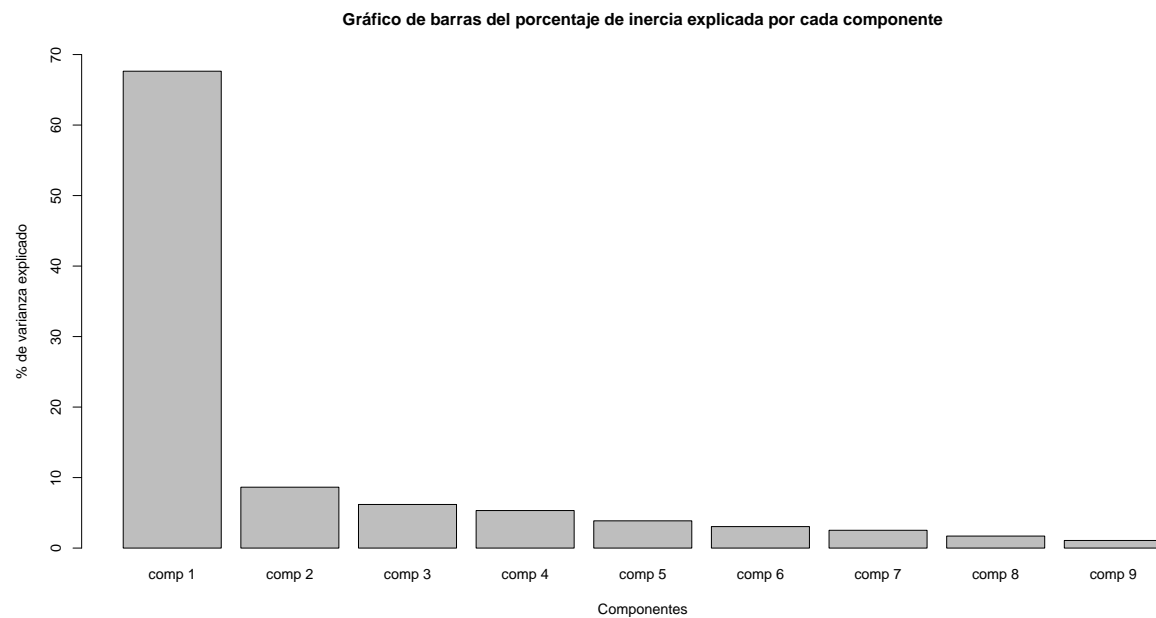


- Matriz de correlaciones

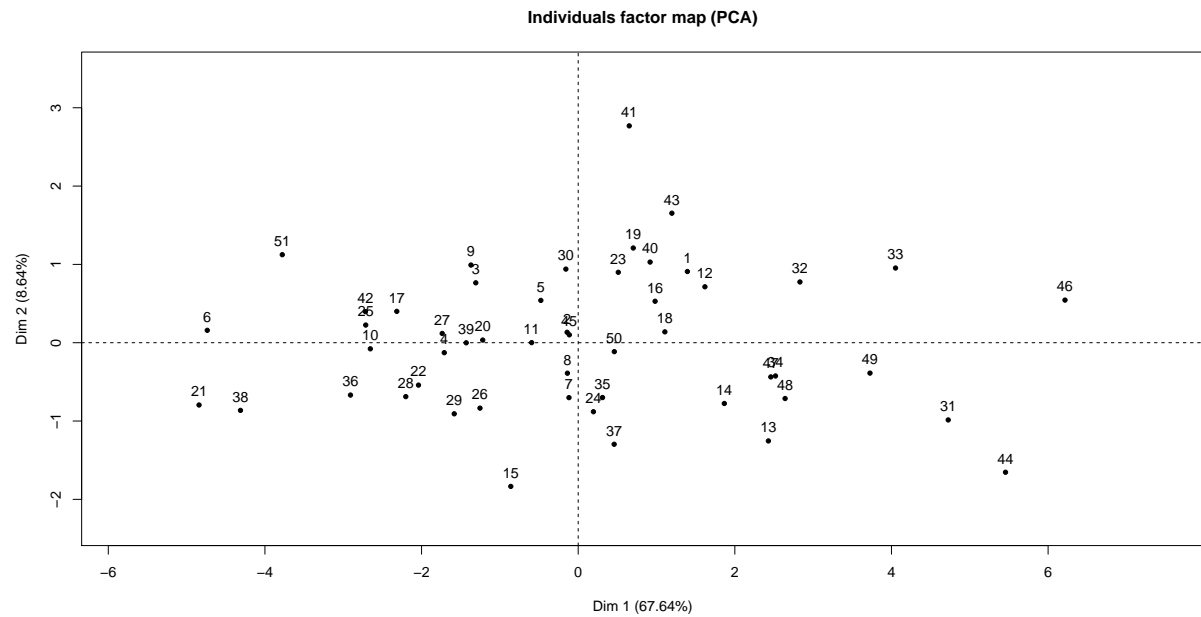
|    | X1       | X2       | X3       | X4       | X5       | X6       | X7       | X8       | X9       |
|----|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| X1 | 1        | 0.531518 | 0.475464 | 0.516663 | 0.474304 | 0.527562 | 0.588739 | 0.438100 | 0.483185 |
| X2 | 0.531518 | 1        | 0.549165 | 0.633367 | 0.505310 | 0.587491 | 0.545900 | 0.462608 | 0.532391 |
| X3 | 0.475464 | 0.549165 | 1        | 0.698918 | 0.670385 | 0.691356 | 0.833379 | 0.810211 | 0.541736 |
| X4 | 0.516663 | 0.633367 | 0.698918 | 1        | 0.674083 | 0.727259 | 0.755607 | 0.756133 | 0.635271 |
| X5 | 0.474304 | 0.505311 | 0.670385 | 0.674083 | 1        | 0.729750 | 0.812126 | 0.705834 | 0.590650 |
| X6 | 0.527562 | 0.587492 | 0.691356 | 0.727259 | 0.729750 | 1        | 0.744703 | 0.726837 | 0.713231 |
| X7 | 0.588739 | 0.545900 | 0.833379 | 0.755607 | 0.812126 | 0.744703 | 1        | 0.844991 | 0.562388 |
| X8 | 0.438100 | 0.462608 | 0.810211 | 0.756133 | 0.705834 | 0.726837 | 0.844991 | 1        | 0.582652 |
| X9 | 0.483185 | 0.532391 | 0.541736 | 0.635271 | 0.590650 | 0.713231 | 0.562388 | 0.582652 | 1        |

### 3. Porcentaje de Inercia y ejes seleccionados

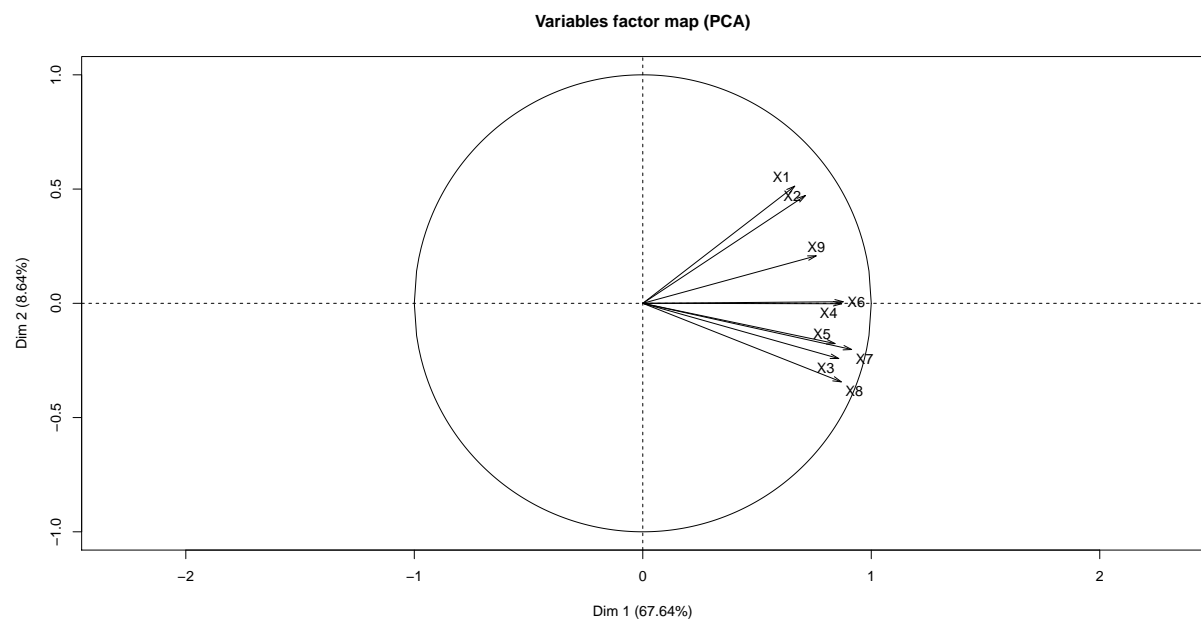
|        | Valor propio | Porcentaje de varianza | Porcentaje de varianza acumulada |
|--------|--------------|------------------------|----------------------------------|
| comp 1 | 6.08732572   | 67.636952              | 67.63695                         |
| comp 2 | 0.77720888   | 8.635654               | 76.27261                         |
| comp 3 | 0.55666880   | 6.185209               | 82.45782                         |
| comp 4 | 0.47914455   | 5.323828               | 87.78164                         |
| comp 5 | 0.34768511   | 3.863168               | 91.64481                         |
| comp 6 | 0.27435674   | 3.048408               | 94.69322                         |
| comp 7 | 0.22712456   | 2.523606               | 97.21683                         |
| comp 8 | 0.15329446   | 1.703272               | 98.92010                         |
| comp 9 | 0.09719117   | 1.079902               | 100.00000                        |



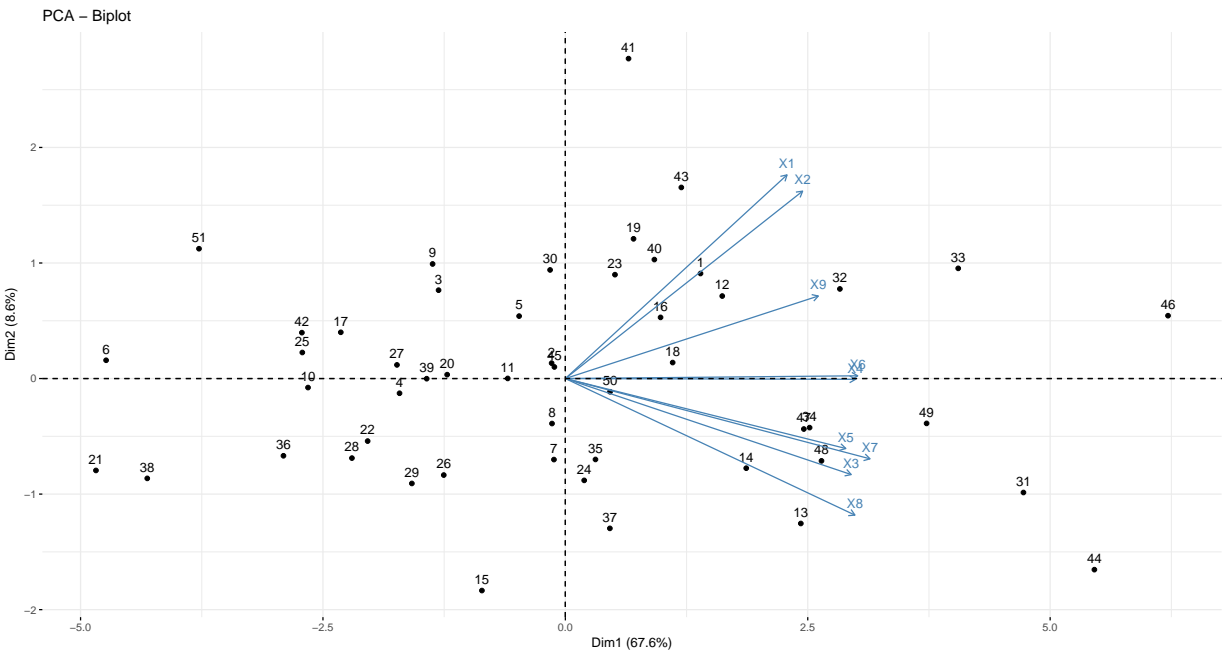
## 4. Nube de individuos



## 5. Circulo de correlaciones



6. Representación simultanea



7. Contribuciones y cosenos cuadrados

Los cosenos cuadrados y las contribuciones de las variables en las tres primeras dimensiones o componentes son:

| Variable | Cosenos      | Cuadrados    | Contribuciones |              |
|----------|--------------|--------------|----------------|--------------|
|          | Componente 1 | Componente 2 | Componente 1   | Componente 2 |
| X1       | 0.4423566    | 0.2628765    | 7.266847       | 33.82314     |
| X2       | 0.5068016    | 0.2226187    | 8.325522       | 28.64335     |
| X3       | 0.7356329    | 0.0584314    | 12.084665      | 7.518112     |
| X4       | 0.7579939    | 0.0000035    | 12.452002      | 0.0004571    |
| X5       | 0.7082081    | 0.0310161    | 11.634141      | 3.990710     |
| X6       | 0.7696271    | 0.00004785   | 12.643107      | 0.0061565    |
| X7       | 0.8356907    | 0.0408603    | 13.728371      | 5.257316     |
| X8       | 0.7549465    | 0.1181842    | 12.401940      | 15.20623     |
| X9       | 0.5760683    | 0.0431702    | 9.463405       | 5.554519     |

8. Construcción Índice

9. Relaciones de transición

10. Varianza de las componentes

11. Conclusiones

??