

Costumbres alimenticias de los hogares franceses según categoría profesional y cantidad de hijos.

SOLUCIÓN

Exploración inicial de los datos

R1: Trabajador manual 5- Empleado 5- Trabajador manual 4

R2: Trabajador manual 2- Trabajador manual 3- Empleado 3

R3: Todas se encuentran medidas en la misma unidad, pero los valores tanto de las medias como de las desviaciones son diferentes.

R4: El consumo de pan está altamente correlacionado con el consumo de leche ($r = 0.86$).

R5: El consumo de carne presenta correlaciones altas con los consumos de: aves ($r = 0.98$), frutas ($r = 0.96$) y legumbres ($r = 0.88$).

R6: El consumo de vino presenta una correlación casi nula ($r = 0.01$) con el consumo de leche; correlaciones negativas y relativamente bajas (r comprendidos entre -0.46 y -0.49) con las legumbres, frutas, carnes y aves; y una correlación positiva pero baja ($r = 0.30$) con el consumo de pan.

Aplicación del método del ACP

Primera parte

R1: Los consumos son homogéneos en la unidad de medida (francos) pero no lo son en sus distribuciones. Para que la variabilidad diferente de cada uno de los consumos no afecte las comparaciones se utiliza como matriz de inercia la matriz de correlación.

R2: La inercia total vale 7. Recordemos que a partir de la matriz de correlación la inercia total es igual al número de variables. En este caso esa inercia se descompuso en 7 ejes. Por tanto la inercia media (el valor propio medio) es igual a 1. Una primera aproximación sería tomar aquellos ejes con valor propio superior al promedio, es decir los dos primeros (valores propios de 4.33 y 1.83 respectivamente). Por otra parte, luego del segundo eje se produce un salto (pasamos de un valor propio de 1.83 a uno de 0.63) que es seguido de un decrecimiento regular. Y, los dos primeros ejes acumulan 88.05% de la inercia total.

Si luego observamos la calidad de representación de los individuos, concluiríamos que se justifica analizar el tercer eje para estudiar un individuo que quedó mal representado por el primer plano (ejecutivo con dos hijos).

Y, por último si analizamos las correlaciones de las variables con los ejes; podemos ver que la variable que está insuficientemente representada por el primer plano es el vino. Esta variable es la que tiene mayor correlación con el tercer eje.

R3: Las familias más contributivas al primer eje son:

Ejecutivo 5 (39.1)

Trabajador manual (17.1)

Por otra parte se puede observar (en el cuadro de los individuos, columna DISTO) que ejecutivo 5 es el punto que está más alejado del origen, es decir que su consumo difiere más del consumo medio.

R4: Las familias mal representadas por el primer eje, con sus respectivos cosenos cuadrados son:

Ejecutivo 2 (0.00)

Trabajador manual 5 (0.01)

Empleado 4 (0.10)

R5: Las oposiciones mayores se dan entre las siguientes familias (con sus respectivas coordenadas):

Trabajador manual 2 (0.00)

Trabajador manual 3 (2.13)

Y

Ejecutivo 5 (-4.51)

Empleado 5 (-2.04)

R6: Trabajador manual 2 y trabajador manual 3 se encuentran próximos en el plano. Esa proximidad se encuentra en los valores originales de algunas variables. Trabajador manual 2 y trabajador manual 3 presentan ambas los más bajos consumos de frutas, carnes y aves.

Segunda parte

R1: Sí, es una buena combinación lineal de las variables originales ya que las mismas presentan correlaciones altas con él. Los consumos de legumbres, frutas, carnes y aves tienen correlaciones mayores a 0.90.

R2: El consumo de vino presenta la correlación más baja (0.43). Eso se puede detectar gráficamente porque el ángulo que forma el vector con el primer eje es importante.

R3: La variable mejor correlacionada con el segundo eje es el consumo de pan (0.84).

R4: Las variables cuyos vectores forman ángulos agudos pequeños presentan altas correlaciones. En ese caso podemos pensar en grupos de variables con correlaciones altas entre ellas y menores con el resto. Así podemos agrupar los consumos de:

- Carnes, frutas y aves (también las legumbres)

- Leche y pan

- Vino

R5: Las correlaciones bajas se dan entre las variables de los distintos grupos mencionados arriba; particularmente el vino presenta correlaciones bajas con las restantes.

R6: Observando el gráfico vemos que los extremos de los vectores (de las variables) se encuentran próximos al círculo de correlación; ello indica que están bien representadas en el plano. La variable que está peor representada por el plano es el consumo de vino cuyo vector forma un ángulo con el plano con un coseno de 0.78.

Tercera parte

R1: Todas las de 5 hijos (trabajador manual, empleado y ejecutivo)

R2: Ejecutivo 3, ejecutivo 4, ejecutivo 5, empleado 5.

R3: El primer eje opone las familias con consumos fuertes (izquierda) con las que tienen consumos menores (derecha). El segundo eje opone las familias que tienen altos consumos de leche y pan con las de menores consumos de esos productos. Se puede observar que los consumos globales aumentan en las familias con mayor cantidad de hijos.

R4.-Se observa un consumo global mayor al aumentar la categoría socio profesional. Si fijamos la cantidad de hijos; observamos que: a igual cantidad de hijos, los ejecutivos tienen consumos mayores de frutas, verduras, aves y carnes que los trabajadores manuales y los empleados se encuentran entre ellos.