

### Práctica de Aula PA-H.2:

Una agencia de viajes de Chicago quiere analizar si hay diferencias significativas entre las familias en la elección del destino de vacaciones, más o menos alejados de su lugar de residencia, en función del número de hijos pequeños en la familia. Para ello dispone de una muestra de 200 familias de esta ciudad entrevistadas en el año 2007 y se especifica el siguiente modelo<sup>1</sup>:

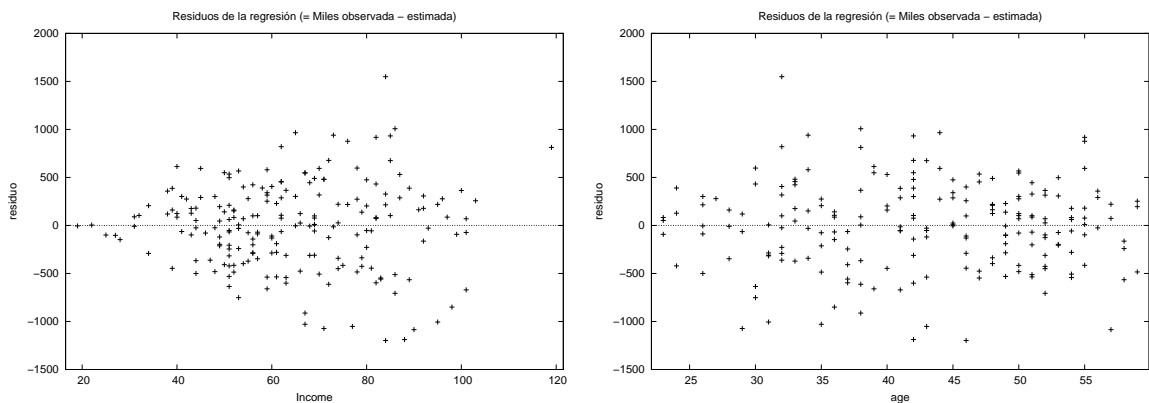
$$Miles_i = \beta_1 + \beta_2 Income_i + \beta_3 age_i + \beta_4 kids_i + u_i \quad i = 1, \dots, 200 \quad (1)$$

donde *Miles* son las millas recorridas por una familia en las vacaciones de ese año, *Income* es la renta familiar anual en miles de dólares, *age* es la edad media de los adultos en la familia y *kids* el número de hijos menores de 16 años existentes en la familia.

Una primera estimación del modelo por MCO produce los siguientes resultados:

$$\begin{aligned} \widehat{Miles}_i &= -391,55 + 14,201 Income_i + 15,741 age_i - 81,826 kids_i & (2) \\ (\widehat{des}(\hat{\beta}_{MCO})) & \quad (169,8) \quad (1,80) \quad (3,757) \quad (27,13) \\ R^2 &= 0,340605 \quad SCR = 40099000 \end{aligned}$$

Figura 1: Gráfico de residuos MCO sobre la variable *Income* y sobre la variable *age*



1. ¿Qué te sugieren los gráficos recogidos en la Figura 1? Comenta detalladamente cada uno de ellos.

Después de agrupar las observaciones de todas las variables en dos grupos en función de un ordenamiento decreciente de la variable *Income* y estimar el modelo (1) anterior por MCO separadamente para cada grupo, se obtienen las siguientes resultados:

<sup>1</sup>Fichero vacation.dat. Recogido en Hill, R. C., Griffiths, W. e. y G. C. Judge (2001), *Undergraduate Econometrics*, 2ª edn., John Wiley and Sons, Inc.

Primera submuestra: Estimaciones MCO utilizando las 80 observaciones 1–80

| Variable dependiente: Miles       |              |              |                 |         |
|-----------------------------------|--------------|--------------|-----------------|---------|
| Variable                          | Coefficiente | Desv. típica | Estadístico $t$ | valor p |
| const                             | -129,22      | 615,6100     | -0,2099         | 0,8343  |
| Income                            | 13,1490      | 6,1456       | 2,1396          | 0,0356  |
| age                               | 13,3666      | 7,5921       | 1,7606          | 0,0823  |
| kids                              | -114,1800    | 52,9888      | -2,1549         | 0,0343  |
| Suma de cuadrados de los residuos |              |              | 2,42765e+07     |         |
| $R^2$                             |              |              | 0,116112        |         |

Segunda submuestra: Estimaciones MCO utilizando las 80 observaciones 121–200

| Variable dependiente: Miles       |              |              |                 |         |
|-----------------------------------|--------------|--------------|-----------------|---------|
| Variable                          | Coefficiente | Desv. típica | Estadístico $t$ | valor p |
| const                             | -339,64      | 220,1600     | -1,5427         | 0,1271  |
| Income                            | 9,6880       | 4,0104       | 2,4157          | 0,0181  |
| age                               | 18,6511      | 3,8740       | 4,8143          | 0,0000  |
| kids                              | -66,0260     | 29,8963      | -2,2085         | 0,0302  |
| Suma de cuadrados de los residuos |              |              | 7,04816e+06     |         |
| $R^2$                             |              |              | 0,308962        |         |

- Realiza un contraste para verificar si lo que sugieren los gráficos es estadísticamente significativo. Debes señalar claramente todos los elementos del contraste incluidas la hipótesis nula y la alternativa.
- Si el contraste realizado te diera que rechazas la hipótesis nula, ¿qué cambiarías de los resultados presentados en (2) si no quisieras cambiar el método de estimación de los coeficientes? ¿Por qué y para qué lo harías? Explica detalladamente.

Se ha utilizado un método de estimación alternativo a MCO para mejorar en términos de eficiencia la estimación de los coeficientes  $\beta$ . Utilizando el software gretl se han obtenido los siguientes resultados:

Estimaciones MC.Ponderados utilizando las 200 observaciones 1–200

Variable dependiente: Miles

Variable utilizada como ponderación:  $\frac{1}{Income}$

| Variable | Coefficiente | Desv. típica | Estadístico $t$ | valor p |
|----------|--------------|--------------|-----------------|---------|
| const    | -408,37      | 145,7170     | -2,8025         | 0,0056  |
| Income   | 13,9705      | 1,6482       | 8,4762          | 0,0000  |
| age      | 16,3483      | 3,4222       | 4,7771          | 0,0000  |
| kids     | -78,3630     | 24,7355      | -3,1680         | 0,0018  |

Estadísticos basados en los datos ponderados:

|  |           |
|--|-----------|
| Suma de cuadrados de los residuos                    | 580616,00 |
| Desviación típica de los residuos ( $\hat{\sigma}$ ) | 54,4272   |
| $R^2$  | 0,3907    |
| $\bar{R}^2$ corregido                                | 0,3813    |
| $F(3, 196)$  | 41,8975   |

Estadísticos basados en los datos originales:

|  |             |
|--|-------------|
| Media de la var. dependiente                         | 1054,2300   |
| D.T. de la variable dependiente                      | 552,7990    |
| Suma de cuadrados de los residuos                    | 4,01134e+07 |
| Desviación típica de los residuos ( $\hat{\sigma}$ ) | 452,3940    |

4. Completa las siguientes expresiones sobre el término de perturbación del modelo y el método de estimación utilizado en la obtención de estos resultados.

$$\begin{array}{ccc}
 E(u_i) = & E(u_i^2) = & E(u_i u_j) = \\
 & \left[ & \right] \\
 \underbrace{E(uu')}_{(\dots \times \dots)} = & & 
 \end{array}$$

Criterio de estimación:.....  $SCR = \sum_{i=\dots}^{i=\dots} (Y_i^* - \hat{\beta}_1 X_{1i}^* - \hat{\beta}_2 X_{2i}^* - \hat{\beta}_3 X_{3i}^* - \hat{\beta}_4 X_{4i}^*)^2$

$Y_i^* = \dots$ ;       $X_{1i}^* = \dots$ ;       $X_{2i}^* = \dots$ ;

$X_{3i}^* = \dots$ ;       $X_{4i}^* = \dots$ ;

$$\hat{\beta}_{\dots} = \left[ \begin{array}{c} \vdots \\ \vdots \\ \vdots \\ \vdots \end{array} \right]^{-1} \left[ \begin{array}{c} \vdots \\ \vdots \\ \vdots \\ \vdots \end{array} \right]$$

5. Si tuvieras que contrastar  $H_0 : \beta_2 = 10$  ¿Cómo lo harías? Razona y explica tu respuesta.