



**Abstract ampliado**

## RESUMEN AMPLIADO

**Título: Gasto familiar a nivel regional: modelización y análisis (periodo 2006-2016)**

**Autores y e-mail de todos ellos:**

**M<sup>a</sup> Isabel López Rodríguez**

**Maja Barac**

**[Maria.I.Lopez@uv.es](mailto:Maria.I.Lopez@uv.es)**

**[Maja.Barac@uv.es](mailto:Maja.Barac@uv.es)**

**Departamento: Economía Aplicada**

**Universidad: València**

**Área Temática:** *Movimientos de la población, desigualdad, pobreza y crecimiento inclusivo*

**Resumen:**

**Introducción:** Resulta evidente la importancia del análisis y modelización de la renta, y consecuentemente de la distribución de los gastos, a la hora de tomar decisiones en términos de política económica, tanto a nivel nacional como regional. Dichas políticas deberían tener como finalidad principal conseguir un aumento de la igualdad acompañada de un crecimiento económico.

Para ello, es necesario identificar los modelos que explicarían tanto los ingresos como los gastos familiares y realizar un análisis de los niveles de desigualdad. Conseguido este objetivo se dispondría de herramientas importantes que permitirían a los gobiernos tomar las medidas oportunas para lograr, en principio, una disminución de la desigualdad tanto dentro de las diferentes comunidades autónomas (CC.AA.) como entre ellas. Esta reflexión ha generado en los últimos años estudios que han abordado tanto la obtención de la modelización de la renta como la interpretación económica de los parámetros del modelo propuesto, enmarcándose el presente trabajo en este contexto. En efecto, dado que se validó el modelo gamma biparamétrico para los gastos familiares en las distintas CC.AA. en las décadas de los 80 y 90 y habiendo estado inmersos en una fuerte crisis económica, que comenzó en torno al año 2007, se pretende dar respuesta a la siguiente cuestión: ¿sigue siendo dicho modelo válido para explicar los gastos de los hogares españoles a lo largo del periodo 2006-2016?

**Datos y Metodología:** Aunque son múltiples las propuestas respecto a los modelos considerados, centrándose en el caso de España, existe cierta unanimidad en cuanto a la base de datos considerada. En efecto en las dos últimas décadas del siglo XX se hizo uso mayoritariamente de los datos proporcionados por las Encuestas de Presupuestos Familiares (E.P.F.) de 1980-81 y 1990-91 y a partir del presente siglo de los datos obtenidos en las E.P.F. con base 2006, sucesoras de las anteriores y que con algunas variantes tenían el mismo objetivo que sus predecesoras pero, a diferencia de aquellas, se realizan con carácter anual.

Conviene en este punto resaltar que a cada dato muestral  $I_i/G_i$  correspondiente al Ingreso/Gasto anual del hogar elevado poblacionalmente le corresponde un factor de



elevación  $F_i$  (corresponde al número de hogares de la población representados por dicho dato), y que dicho factor procede del tipo de muestreo utilizado. Éste, que se engloba dentro del diseño general de la Encuesta General de Población (E.G.P.), se caracteriza por ser un diseño muestral complejo, en el sentido de que combina diversos tipos de muestreo (conglomerados, bietápico, estratificado, sistemático,...).

En cuanto al modelo propuesto para explicar la renta no ha existido tanta unanimidad. La controversia surgida en torno al modelo gamma biparamétrico, propuesto por Salem y Mount, rechazado por Dagum, validado por Rojo, Lafuente y López, origina que uno de los objetivos del presente estudio sea decantar la balanza en uno u otro sentido.

Dado que el periodo de tiempo considerado del que se dispone información abarca diez años (desde 2006 a 2016), en caso de que la respuesta a la pregunta planteada sobre la validez del modelo gamma biparamétrico fuese afirmativa, se estaría en disposición, adicionalmente, de analizar cómo ha evolucionado la economía de las diferentes CC.AA.

La elección de dicho modelo (que proviene del sistema generador de Pearson), radica básicamente en la interpretación económica de sus parámetros. En efecto, siendo un modelo de tipo continuo, con función de densidad:

$$f(x) = \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} \quad \forall x \in ]0, \infty[, \alpha > 0, \lambda > 0$$

( $\Gamma(\alpha)$ ) es la función gamma de Euler que obedece a cualquiera de las siguientes expresiones alternativas:  $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$  ó  $\Gamma(\alpha) = \lambda^\alpha \int_0^\infty y^{\alpha-1} e^{-\lambda y} dy$  )

depende de los parámetros  $\alpha$  y  $\lambda$ , de manera que al presentar el primero de ellos una fuerte correlación con el índice de Gini, posibilita la realización de análisis de desigualdad. El otro parámetro,  $\lambda$ , es un factor de escala que permite realizar estudios comparativos entre países/regiones con diferente unidad monetaria.

Algunas medidas de interés del modelo propuesto, y que serán utilizados en el presente estudio, son:

- Media:  $\mu = \frac{\alpha}{\lambda}$
- Varianza:  $\sigma^2 = \frac{\alpha}{\lambda^2}$
- Mediana:  $M_e = \frac{3\alpha - 1}{3\lambda}$

De dichas expresiones, se deduce que todas guardan una relación directa con  $\alpha$  e inversa con  $\lambda$ , de manera que consideradas dos distribuciones Gamma con idéntico parámetro  $\alpha$ , aquella que posea un valor de mayor  $\lambda$  tendrá una media, varianza y mediana de menor magnitud.

Respecto a la metodología utilizada a lo largo del estudio cabe indicar:

- El test de adherencia usado, para analizar la bondad del ajuste, será el de Kolmogorov-Smirnov.
- El método de estimación utilizado para estimar los parámetros  $\alpha$  y  $\lambda$  será el de máxima-verosimilitud.

Cabe indicar que el uso de este método de estimación lleva a la conclusión de que las expresiones de los estimadores máximo-verosímiles de  $\alpha$  y  $\lambda$  no son deducibles de forma analítica.



En efecto, para su obtención es precisa la resolución del siguiente sistema de ecuaciones:

$$\begin{cases} \hat{\lambda} = \frac{\hat{\alpha}}{\bar{X}} \\ \ln(\hat{\alpha}) - \Psi(\hat{\alpha}) = \ln\left(\frac{\bar{X}}{\bar{\tilde{X}}}\right) \end{cases}$$

Siendo para ello necesario el uso de técnicas propias del cálculo numérico. En este caso se ha optado por el uso combinado de los métodos de Newton-Raphson y Gauss-Laguerre, así como la elaboración de programas ad-hoc que permitan tanto la resolución del sistema como la manipulación de los datos y los cálculos necesarios para la obtención de los resultados que se irán presentando.

Además, se pueden aplicar a  $\hat{\alpha}$  y  $\hat{\lambda}$  las propiedades correspondientes a los estimadores máximo-verosímiles, siendo para este caso fundamental el uso de la propiedad de invarianza.

Por otra parte, tal y como puede comprobarse de las ecuaciones del sistema, es necesaria la obtención tanto de la media geométrica como la aritmética de los datos, medias que responden a la siguiente expresión:  $\tilde{X} = \prod_i D_i^{F_i/F_T}$  y  $\bar{X} = \sum_i D_i \cdot \frac{F_i}{F_T}$  (con  $F_T = \sum_i F_i$ )

siendo  $F_i$  el factor de elevación correspondiente i-ésimo valor, esto es, si  $D_i$  es un dato muestral (gasto) con factor de elevación  $F_i$ , esto se traduce en que dicho dato representa a  $F_i$  hogares.

**Análisis y resultados:** Una vez sentado el objetivo fundamental del trabajo (validación del modelo gamma biparamétrico como explicativo de los hogares a lo largo del período 2006-2016 y posterior análisis de la evolución de la desigualdad) y expuestas tanto la naturaleza de la información empírica como la metodología utilizada, el trabajo se amplía con dos nuevas vertientes, una de carácter descriptivo y otra de tipo inferencial. Concretamente:

- La obtención del gasto medio, necesario para la estimación de los parámetros del modelo permite realizar un estudio descriptivo en el siguiente sentido: resulta interesante conocer qué CC.AA. se encuentran entre las mejor situadas a nivel de gasto medio y cuáles en peor posición, esto es, las más favorecidas y las más “pobres”. Los resultados muestran dos clusters bien diferenciados que se mantienen a lo largo de los 11 años estudiados. Entre las mejor situadas cabe mencionar Madrid, País Vasco, Navarra y Cataluña y entre las peor posicionadas Extremadura, Canarias, Castilla y León, Castilla la Mancha y la Comunidad Valenciana.

Por otra parte, realizado un estudio longitudinal se observa una disminución del gasto en la casi totalidad de las CC.AA., de forma que el modelo de regresión para cada comunidad permite conocer la elasticidad para cada año y comunidad. Mientras que para el primero de los estudios descriptivos, de naturaleza transversal, no ha sido necesaria la deflactación de los datos, para el análisis longitudinal se ha procedido a deflactar a euros constantes del 2016.

- Tras la validación del modelo gamma biparamétrico (de los 204 contrastes de adherencia, tan sólo en 7 casos éste no se ha superado, lo que supone un porcentaje global de ajuste del 97%) y un análisis de la evolución de la desigualdad, que indica que ésta ha disminuido en una gran mayoría de las



CC.AA., a lo largo del periodo de tiempo considerado; el estudio inferencial se completa analizando:

⇒ Si el gasto medio obtenido a nivel poblacional (calculado teniendo en cuenta que  $\mu = \frac{\alpha}{\lambda}$ ) representa bien a los hogares. Este punto se

considera, pues parece relevante conocer hasta qué punto el gasto medio autonómico es aplicable y válido para todos los hogares de la comunidad autónoma a la que estos pertenecen.

Para ello se hace uso del coeficiente de variación de Pearson, que además permite realizar un estudio comparativo, tanto a nivel longitudinal como transversal, de la evolución de la variabilidad. Dicho estudio concluye que la variabilidad tiende a disminuir a lo largo del tiempo, pudiendo destacarse a modo de ejemplo Baleares, País Vasco y Madrid como las CC.AA. con menor variabilidad, situándose en el punto opuesto Extremadura, Castilla y León, Castilla la Mancha y Galicia.

⇒ Si los umbrales marcados por la mediana (obtenida teniendo en cuenta que  $M_e = \frac{3\alpha - 1}{3\lambda}$ ) son similares para todas las CC.AA. y se ha

mantenido a lo largo de los 11 años objeto de estudio. Los primeros resultados visibilizan la posición de Madrid y Navarra como el de las CC.AA. en las que el 50% de los hogares que menos gastan son las que realizan un gasto más elevado y Extremadura se encuentra en el caso contrario. Para llevar a cabo este análisis se ha trabajado, evidentemente, con los datos deflactados.

Tanto para la estimación del gasto medio poblacional ( $\mu$ ) como el de la mediana poblacional, atendiendo a las expresiones expuestas, ha sido de necesaria aplicación la propiedad de invarianza de los estimadores máximo-verosímiles.

**Discusión:** Se establecen, por último, como futuras líneas de investigación:

- El análisis de la evolución del gasto en los 12 ítems considerados por el INE (“Alimentos y bebidas no alcohólicas”, “Bebidas alcohólicas, tabaco y narcóticos”, “Artículos de vestir y calzado”, “Vivienda, agua, electricidad, gas y otros combustibles”, “Mobiliario, equipamiento del hogar y gasto corriente en conservación de la vivienda”, “Salud”, “Transporte”, “Comunicaciones”, “Ocio, espectáculos y cultura”, “Enseñanza”, “Hoteles, cafés y restaurantes”, “Otros bienes y servicios”)
- Estudiar si las diferencias detectados, en términos de desigualdad, son significativas, identificando clusters de CC.AA. con niveles de igualdad semejantes y situaciones de dominancia y alternancia entre CC.AA. con diferentes niveles de desigualdad.

En este sentido el análisis se completaría con un estudio de la evolución de las diferentes situaciones detectadas, esto es analizando si los clusters mantienen la misma composición a lo largo del tiempo y si las situaciones de dominancia/alternancia se mantienen en todo el periodo.

**Palabras Clave:** *modelización gasto familiar, desigualdad regional, índice Gini, comparativa CC.AA.*

**Clasificación JEL:** C51, E01, R15