



Universidad de Jaén

Facultad de Ciencias Sociales
y Jurídicas

Trabajo Fin de Grado

ESTUDIO ECONOMÉTRICO DE LA DEMOGRAFÍA EN ANDALUCÍA

María Jesús Piqueras Rodríguez

Mayo, 2019

ÍNDICE

| | |
|---|----|
| RESUMEN..... | 2 |
| 1. INTRODUCCIÓN | 3 |
| 2. ASPECTOS TEÓRICOS SOBRE LA DEMOGRAFÍA..... | 3 |
| 3. FACTORES DETERMINANTES..... | 5 |
| 4. ANÁLISIS ECONOMETRICO..... | 8 |
| 4.1 Modelo 1..... | 9 |
| 4.2 Modelo 2..... | 10 |
| 4.2.1 Multicolinealidad | 12 |
| 4.3 Modelo 3..... | 13 |
| 4.3.1 Multicolinealidad | 15 |
| 4.3.2 Especificación..... | 15 |
| 4.3.3 No- linealidad..... | 16 |
| 4.4 Modelo 4..... | 18 |
| 4.4.1 Especificación..... | 19 |
| 4.4.2 Multicolinealidad | 20 |
| 4.4.3 Heterocedasticidad | 21 |
| 4.4.4 Normalidad de residuos..... | 24 |
| 4.4.5 Contraste de chow | 25 |
| 4.4.6 Autocorrelación | 26 |
| CONCLUSIONES..... | 30 |
| BIBLIOGRAFÍA | 32 |

RESUMEN

El objetivo de este trabajo de fin de grado es conocer la evolución que ha experimentado la población en Andalucía desde el año 1992 al año 2016 en función de una serie de factores demográficos. Para conseguir dicho objetivo se estudia la dinámica demográfica, como variable explicada, la Población, más las variables explicativas más significativas como son la natalidad, la fecundidad, la mortalidad y la migración. Además de estudiar cómo afectan las variables más significativas, se incluirá otra variable explicativa secundaria, como la incorporación de la mujer al mundo laboral, para así poder saber el efecto que tiene sobre la población.

Con dichos datos se hace un estudio econométrico detallado aplicando *las fases de la metodología Box-Jenkins y el programa econométrico de distribución libre Gretl*. Para terminar, se elaboran las conclusiones de los resultados obtenidos en el estudio.

ABSTRACT

The aim of this dissertation is to find out about the evolution that Andalusia population has experimented from 1992 to 2016, regarding a series of demographic factors. To achieve this purpose, the demographic dynamics is examined, such as the variable explained, the population, as well as the most significant explanatory variables such as birth, fertility, mortality and migration. In addition, a study of how the most significant variables and another secondary explanatory variable will be included, as the incorporation of women in the world of work. Thus, we can know the effect it has on the population. With this data, an econometric study is made, applying the different phases of Box-Jenkins methodology and Gretl econometric programme of free distribution. To conclude, there will be an elaboration of the conclusions of the data obtained during the research.

1. INTRODUCCIÓN

En los últimos años en España y más concretamente en Andalucía, se ha observado cómo la población es más elevada ahora que en el siglo anterior, aproximadamente hay unas 2.237.559 personas más.

Andalucía representa el 17.8% de población en España.

Dentro de este crecimiento y la alta representación poblacional que supone Andalucía para España también se ha observado que su población es madura y anciana. Por lo que la elaboración de este trabajo constará en estudiar qué factores y en qué proporción afectaran a ese aumento y características de la población. Por lo que explicaré de forma teórica las variables que influyen en la población y seguidamente se procederá a la estimación de un modelo econométrico del que me ayudo para explicar la evolución que ha sufrido Andalucía en los años dados anteriormente.

La motivación a la hora de elegir esta variable de estudio es por la evolución creciente que ha tenido la población en la actualidad en comparación con épocas anteriores, pero dicho incremento no ha sido tan significativo debido a que hoy día sigue habiendo problemas de población anciana y el que cada vez más disminuye el número medio de hijos.

Para este proceso me ayudaré con el software econométrico *Gretl* disponible en la web: <http://gretl.sourceforge.net>.

2. ASPECTOS TEÓRICOS SOBRE LA DEMOGRAFÍA

La Demografía se define como el estudio de la población, pero dicha palabra tiene más de un significado dependiendo del contexto en el que la usemos.

Debido a la variedad de significados ha sido difícil que se le reconozca como “ciencia”, pues uno de los requisitos que se necesita para que tal concepto sea reconocido como ciencia es que tiene que ser claro, o sea que no de pie a confusión. Por lo que es recomendable saber los diferentes usos que se le puede dar al término población, para así poder encuadrarlo en los distintos tipos de estudio (como puede ser un estudio demográfico, sociológico o antropológico).

No nos sirve para la demografía las siguientes formas de utilizar el término población:

1. La definición social común de población: debido a que nos lleva a confusión, principalmente por lo dicha ya anteriormente, no cabría diferencia entre un sociólogo, un demógrafo o un antropólogo.
2. El uso estadístico de la población. Debido a que son diversas observaciones que se hacen a un colectivo de referencias.

A pesar de lo dicho anteriormente, aún la demografía mantiene en sí misma dos maneras muy diferentes para entender el término población. Lo que viene siendo la población como stock (número de personas en un momento dado). Y la de estructura (clasificada en diferentes segmentos como puede ser la edad, el sexo, el dinero que ganan al año...). Por último, la población como colectivo, continuidad en el tiempo. A lo que se refiere con el término continuidad en el tiempo es cuando el estudio no se hace a segmentos en concreto o a una cantidad de personas en un momento o lugar determinado, sino que nos interesa más cómo viven las personas, sus comportamientos y cómo se repone dicha población después de un fallecimiento. En este término se ve a la población como sistemas reproductivos o sea que es lo mismo que decir que el estudio no es estático sino dinámico de la población en sí.

Lo que extraemos de estas definiciones es, que para estudiar el tema que nos interesa (la demografía), la población la población que tenemos que ver es un colectivo dinámico, lo que hemos definido anteriormente como sistema reproductivo, lo que viene siendo cómo nacen, viven y mueren las personas.

Las condiciones primarias que la demográfica estudia de la población son los momentos en que una vida empieza, termina y las características y conductas que condicionan la reproducción de tales vidas. También podemos encontrar condicionantes secundarios como pueden ser: la política, la sanidad, el mercado laboral...

Una vez desgranado la clase de población exacta que se utiliza en el estudio demográfico y definir la estructura demográfica, considero que una de las definiciones más completas y condensada es la descrita por HAUSER Y DUCAN. Viene a decir que la demografía es el estudio del número, distribución territorial y composición de la población, de cómo varía y qué variables provocan esa variación. Dichas variables se identifican como natalidad, mortalidad y migración.

3. FACTORES DETERMINANTES

Con lo dicho anteriormente, varios autores¹ llegan a la conclusión que para estudiar la demografía hay que tener en cuenta algunos factores que se clasifican de la siguiente manera:

- **Variable dependiente, también denominada como explicada o variable causal o endógena.** Este tipo de variable es aquella que el modelo explica. En nuestro caso dicha variable sería la Población ya que la demografía tiene como objeto el estudio del comportamiento de esta.
- **Variables independientes, explicativas o exógenas.** Tratan de explicar el comportamiento y características de la población (variable dependiente).

En el modelo que vamos a estudiar estas variables pueden ser clasificadas como variables primarias y secundarias.

Las primarias serían:

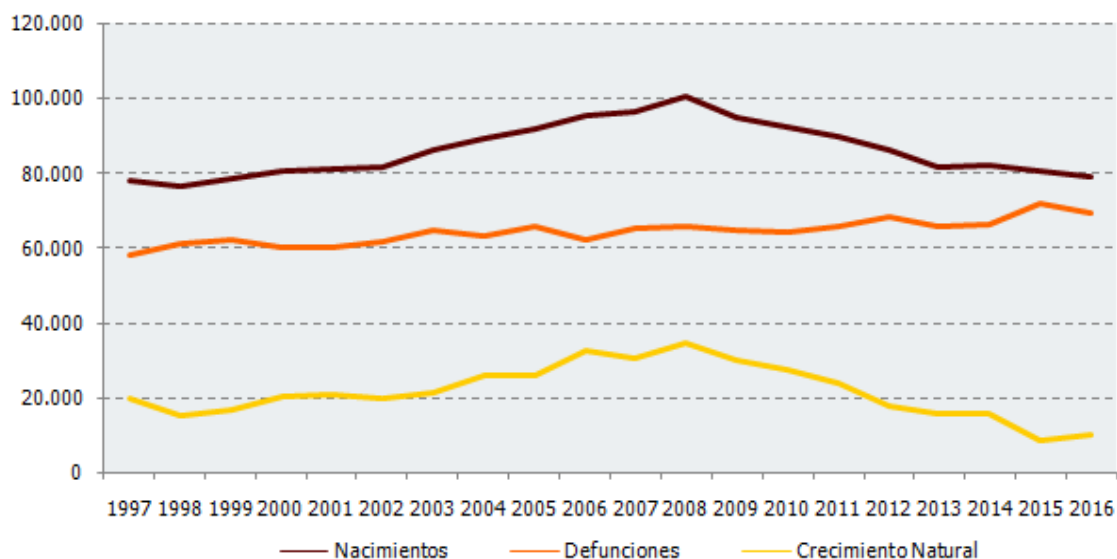
1º Natalidad: número de nacimientos de una población determinada durante un periodo de tiempo. Esta variable es importante ya que como hemos dicho anteriormente necesitamos saber el número de nacimientos y cómo varían en las distintas épocas.

2º Mortalidad: número de fallecimientos de una determinada población en un periodo establecido. También imprescindible para llegar al número de habitantes en un determinado lugar y tiempo.

3º Migraciones: calculada como la inmigración menos la emigración. Su definición sería la del número de personas que vienen de otros países a una determinada región o comunidad e incluso personas de la misma nación se trasladan a dicha región o comunidad (inmigración) menos el número de personas de una región o comunidad que se van a otro lugar ya sea del mismo país o distinto (emigración). Importante dicha variable para saber el flujo de personas en los distintos años y saber su evolución.

¹ Lucas Marín, A. y Romero, J. (2006), *Estructura social. La realidad de las sociedades avanzadas*, Madrid, Pearson "Población y Demografía", pp. 144-120.

Evolución de nacimientos, defunciones y crecimiento Natural.



Fuente: [Instituto de Estadística y Cartografía de Andalucía](#)

Con esta Gráfica nos damos cuenta del crecimiento que presentan sobre todo los nacimientos más o menos a partir del año 2004 y la bajada que vuelve a experimentar sobre el año 2013.

La variable independiente secundaria, la he incluido para ver si pueden explicar el crecimiento, disminución o estancamiento de la población, o sea si es o no significativas. Esta variable es:

La incorporación de la mujer al mundo laboral: he visto interesante estudiarla ya que cada vez son más las mujeres que deciden unirse al mundo laboral, y si esta variable puede influir en la natalidad y por consiguiente a la población.

Por otro lado, considero que hay que hacer mención a la variable de la fecundidad. Dicha variable se asemeja mucho con la natalidad, tanto que da pie a confusión, pero tiene ciertos matices que hacen que se diferencien.

Hay autores como los referidos anteriormente que defienden la teoría de incorporar a la natalidad en el estudio demográfico, pero hay otros estudios² que defienden la incorporación de la fecundidad a dicho estudio.

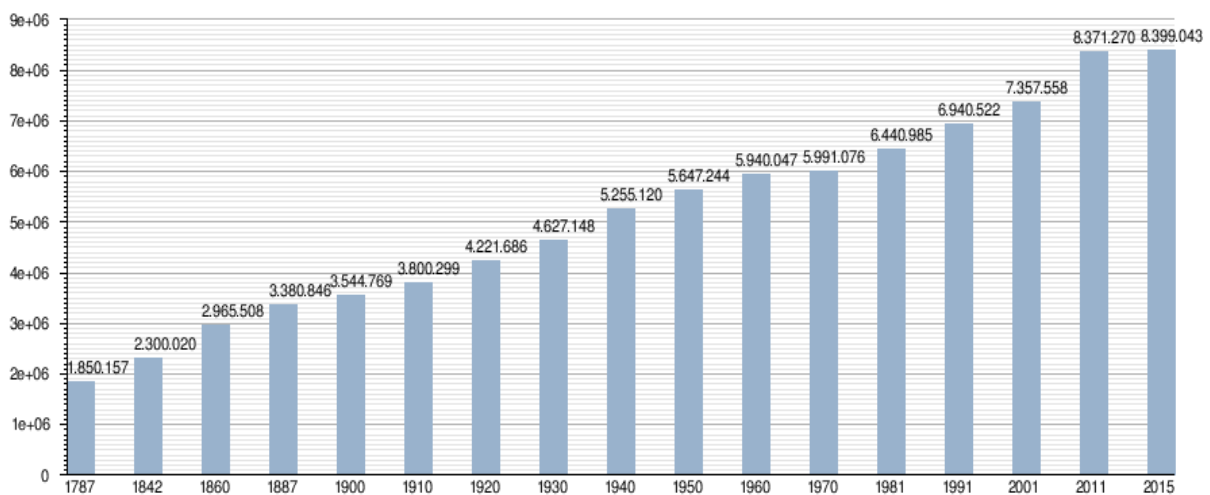
² INE

La diferencia reside en que la tasa de la natalidad, estudia los nacimientos ocurridos en una población. Mientras que fecundidad hace referencia a los nacimientos en relación con la población en la que tiene lugar, es decir, las mujeres en edad fértil.

Se considera que la preferencia de ciertos autores por la elección de la tasa de fecundidad en lugar de la tasa de natalidad se debe a que la natalidad es un estudio que hace una media de manera global, es decir no se centra en el estudio de nacimientos de mujeres que están edad de reproducirse, cosa contraria que hace la tasa de fecundidad que hace un estudio más concreto de los nacimientos en mujeres que están en su edad fértil, por lo que muchos autores clasifican este estudio como real y más lógico que la tasa de natalidad.

En la gráfica que se muestra a continuación, muestra lo dicho anteriormente de que la población andaluza crece desde épocas anteriores a la actualidad, pero lo hace cada vez más en proporciones inferiores. De ahí mi estudio de la incorporación de la mujer al mundo laborar, para ver si dicha variable influye en ese descenso continuado, ya que en años anteriores eran pocas mujeres las que trabajaban fuera de sus casas en comparación con la actualidad.

Evolución de la población en Andalucía



Fuente: [Instituto de Estadística de Andalucía](#)

Por último, exponer los datos³⁴ con los que vamos a trabajar:

³ INE

⁴ La Junta de Andalucía

| T. crec. población | T. Natalidad | T. Mortalidad | T. Saldo migratorio | T. Fecundidad | T. Empleo mujeres |
|--------------------|--------------|---------------|---------------------|---------------|-------------------|
| 4,76 | 12,57 | 7,80 | 2,17 | 49,35 | 20,31 |
| 4,17 | 12,24 | 8,08 | 2,17 | 47,67 | 19,60 |
| 3,57 | 11,50 | 7,93 | 2,15 | 44,46 | 18,91 |
| 3,04 | 11,14 | 8,11 | 2,16 | 42,80 | 19,72 |
| 2,61 | 10,78 | 8,18 | 2,15 | 41,20 | 20,83 |
| 2,8 | 10,86 | 8,07 | 2,13 | 41,26 | 21,43 |
| 2,16 | 10,59 | 8,45 | 2,17 | 40,14 | 22,04 |
| 2,3 | 10,84 | 8,56 | 2,24 | 41,00 | 23,44 |
| 2,78 | 11,04 | 8,27 | 2,31 | 41,65 | 25,07 |
| 2,89 | 11,05 | 8,17 | 6,43 | 41,67 | 26,56 |
| 2,71 | 11,01 | 8,30 | 10,67 | 41,57 | 28,31 |
| 2,82 | 11,42 | 8,60 | 11,19 | 43,18 | 29,89 |
| 3,4 | 11,64 | 8,24 | 14,15 | 44,09 | 31,42 |
| 3,32 | 11,76 | 8,45 | 13,99 | 44,64 | 34,06 |
| 4,15 | 12,03 | 7,88 | 11,10 | 45,83 | 35,84 |
| 3,83 | 11,93 | 8,10 | 12,56 | 45,61 | 37,31 |
| 4,25 | 12,28 | 8,03 | 7,31 | 47,28 | 37,13 |
| 3,62 | 11,48 | 7,85 | 4,02 | 44,62 | 34,74 |
| 3,34 | 11,10 | 7,76 | 3,41 | 43,64 | 35,52 |
| 2,86 | 10,72 | 7,87 | 2,62 | 42,62 | 34,42 |
| 2,15 | 10,30 | 8,15 | -0,32 | 41,39 | 33,64 |
| 1,88 | 9,71 | 7,83 | -2,39 | 39,52 | 32,38 |
| 1,89 | 9,80 | 7,90 | -0,61 | 40,43 | 32,65 |
| 1,03 | 9,60 | 8,57 | -0,53 | 40,20 | 34,25 |
| 1,19 | 9,431652126 | 8,24 | -0,59 | 40,03 | 34,79 |

Fuente: Elaboración propia a partir de datos del INE

4. ANÁLISIS ECONÓMETRICO

En esta sección aplicamos las siguientes fases para elaborar el modelo econométrico: Especificación, Estimación y Diagnóstico.

4.1. Modelo 1

En primer lugar, se plantea el modelo econométrico bajo la forma lineal:

$$Y = \beta_0 + \beta_1 \times_1 + \beta_2 \times_2 + \beta_3 \times_3 + \beta_4 \times_4 + u$$

Donde:

- Y: Tasa de crecimiento anual de la población (en ‰).
- X₁: Tasa de natalidad en Andalucía (en ‰).
- X₂: Tasa de mortalidad en Andalucía (en ‰).
- X₃: Tasa del saldo migratorio en Andalucía (en ‰).
- X₄: Tasa de mujeres empleadas en Andalucía (en ‰)

Para estimar los parámetros se ha utilizado el método de Mínimos Cuadrados Ordinarios (MCO) obteniendo los siguientes resultados:

Resultados de la estimación del Modelo 1.

| | Coeficiente | Desv. Típica | Estadísticot | valor p |
|---------------------|-------------|--------------|--------------|-------------|
| const | -0.00618563 | 0.0612213 | -0.1010 | 0.9205 |
| T. natalidad | 0.998024 | 0.00206116 | 484.2 | <0.0001 *** |
| T. mortalidad | -0.993904 | 0.00518327 | -191.8 | <0.0001 *** |
| T. saldo migratorio | 0.000232082 | 0.000372079 | 0.6237 | 0.5398 |
| T. empleo mujeres | -0.00064493 | 0.000205526 | -3.138 | 0.0052 *** |

Los asteriscos que aparecen en la última columna junto al número del p-valor correspondiente a cada variable, nos dicen el nivel de significación, por tanto, al aparecer 3 asteriscos nos viene a decir que la variable natalidad, mortalidad y tasa de empleo de la mujer son individualmente significativas con cualquier tipo de α (0.05, 0.01 y 0.1)

| | | | |
|------------------------|-----------|-----------------------|-----------|
| Media de la vble. dep. | 2.940.800 | D.T. de la vble. dep. | 0.935236 |
| Suma de cuad. residuos | 0.000474 | D.T. de la regresión | 0.004866 |
| R-cuadrado | 0.999977 | R-cuadrado corregido | 0.999973 |
| F(4, 20) | 221621.9 | Valor p (de F) | 3.76e-46 |
| Log-verosimilitud | 1.004.521 | Criterio de Akaike | -190.9043 |
| Criterio de Schwarz | -184.8099 | Crit. de Hannan-Quinn | -189.2139 |
| rho | 0.574909 | Durbin-Watson | 0.774949 |

Con los resultados obtenidos, el modelo de regresión ajustado sería:

$$\hat{Y} = -0.00618563 + 0.998024X_1 - 0.993904X_2 + 0.000232082X_3 - 0.00064493X_4$$

Al estimar el modelo con las variables nombradas al principio, donde explicamos mediante la teoría por qué dichas variables se supone que influyen en la población, el resultado no resulta coherente, ya que las migraciones es una variable importante y principal en el modelo y aquí nos sale como no significativa.

Debido a lo comentado en el párrafo anterior vamos a plantear otro modelo, incorporando la variable fecundidad. Dicha incorporación se va a realizar por la importancia que se le da a la fecundidad en el estudio de la demografía en muchos otros manuales⁵⁶ como la pequeña incisión que hacemos al principio.

4.2. Modelo 2

Para empezar, se plantea el modelo econométrico bajo la forma lineal:

$$Y = \beta_0 + \beta_1 \times_1 + \beta_2 \times_2 + \beta_3 \times_3 + \beta_4 \times_4 + \beta_5 \times_5 + u$$

Donde:

⁵ Publicación de las Naciones Unidad (2014), *Seria Manuales n°82*, "los datos demográficos, limitaciones y métodos de evaluación".

⁶ Hernández A., García J., García I. y Pardo C. (2006), Artículo especial. "Indicadores demográficos y situación actual de la fecundidad en España"

- Y: Tasa de crecimiento anual de la población (en ‰).
- X₁: Tasa de natalidad en Andalucía (en ‰).
- X₂: Tasa de fecundidad en Andalucía (en ‰).
- X₃: Tasa de mortalidad en Andalucía (en ‰).
- X₄: Tasa de saldo migratorio en Andalucía (en ‰)
- X₅: Tasa de mujeres empleadas en Andalucía (en ‰)

Para estimar los parámetros se ha utilizado el método de Mínimos Cuadrados Ordinarios (MCO) obteniendo los siguientes resultados:

Resultados de la estimación del Modelo 2.

| | Coeficiente | Desv. Típica | Estadísticot | valor p |
|---------------------|-------------|--------------|--------------|------------|
| const | 0.00109457 | 0.0394558 | 0.02774 | 0.9782 |
| T. natalidad | 101.599 | 0.00357994 | 283.8 | <0.0001*** |
| T. fecundidad | -0.00491098 | 0.000908691 | -5.404 | <0.0001*** |
| T. mortalidad | -0.994727 | 0.00334202 | -297.6 | <0.0001*** |
| T. saldo migratorio | -0.00073593 | 0.000299194 | -2.460 | 0.0237** |
| T. empleo mujeres | -8.68648e-5 | 0.000167890 | -0.5174 | 0.6109 |

Los asteriscos que aparecen en la última columna junto al número del p-valor correspondiente a cada variable, nos dicen el nivel de significación, por tanto, al aparecer 3 asteriscos nos viene a decir que la variable natalidad, fecundidad y mortalidad son individualmente significativos con cualquier tipo α (0.05, 0.01 y 0.1) y tasa de migración es significativa con α (0.05 y 0.1).

| | | | |
|------------------------|-----------|-----------------------|-----------|
| Media de la vble. dep. | 2.940.800 | D.T. de la vble. dep. | 0.935236 |
| Suma de cuad. residuos | 0.000187 | D.T. de la regresión | 0.003134 |
| R-cuadrado | 0.999991 | R-cuadrado corregido | 0.999989 |
| F(5, 19) | 427365.1 | Valor p (de F) | 2.74e-47 |
| Log-verosimilitud | 1.120.907 | Criterio de Akaike | -212.1815 |
| Criterio de Schwarz | -204.8682 | Crit. de Hannan-Quinn | -210.1531 |
| rho | 0.034127 | Durbin-Watson | 1.931.709 |

Con los resultados obtenidos, el modelo de regresión ajustado sería:

$$\hat{Y} = 0.00109457 + 101.599X_1 - 0.00491098X_2 - 0.994727X_3 - 0.00073593X_4 - 8.68648e-5X_5$$

Con estos resultados observamos que el problema de significación de la variable migración se ha corregido, pero ha surgido un nuevo problema de coherencia con el signo negativo que aparece en la tasa de fertilidad, no tiene sentido que la población disminuya cuando aumenta la fertilidad.

Para poder saber el porqué de ese signo a continuación realizamos el estudio de la colinealidad.

4.2.1. Multicolinealidad

Factores de inflación de varianza (VIF)

| | |
|---------------------|--------|
| T. natalidad | 22,752 |
| T. fecundidad | 13,999 |
| T. mortalidad | 1,731 |
| T. saldo migratorio | 5,35 |
| T. empleo mujeres | 2,879 |

Siendo $VIF(j) = 1/(1 - R(j)^2)$, donde $R(j)$ es el coeficiente de correlación múltiple entre la variable j y las demás variables independientes.

Como dos de los valores resultantes son superiores a 10, existe en este modelo 2 problemas de colinealidad con las variables que superan dicho valor. Por tanto, se afirma que hay problemas de multicolinealidad.

Para solucionar el problema tenemos que eliminar una de las dos variables que ocasionan el problema de colinealidad por lo que volvemos a plantear otro modelo diferente pero esta vez eliminando la natalidad, ya que en el primer modelo aparecía la natalidad y no estábamos conforme con los resultados, debido a que el saldo migratorio no nos salía significativo.

A continuación, se plantea un modelo donde aparece la tasa de fertilidad, variable que varios autores como he especificado anteriormente y otros,⁷ prefieren debido a que es más objetiva que la natalidad.

4.3. Modelo 3

En primer lugar, se plantea el modelo econométrico bajo la forma lineal:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + u$$

Donde:

- Y: Tasa de crecimiento anual de la población (en ‰).
- X1: Tasa de fecundidad en Andalucía (en ‰).
- X2: Tasa de mortalidad en Andalucía (en ‰).
- X3: Tasa del saldo migratorio en Andalucía (en ‰).
- X4: Tasa de mujeres empleadas en Andalucía (en ‰)

Para estimar los parámetros se ha utilizado el método de Mínimos Cuadrados Ordinarios (MCO) obteniendo los siguientes resultados:

⁷ Pérez Puchan Pedro. "La fecundidad como factor del crecimiento demográfico".

Resultados de la estimación del Modelo 3

| | coeficiente | Desv.Típica | Estadístico t | valor p |
|---------------------|-------------|-------------|---------------|------------|
| constante | 3.07178 | 2.40816 | 1.276 | 0.2167 |
| T. fecundidad | 0.234588 | 0.0213874 | 10.97 | <0.0001*** |
| T. mortalidad | -1.16685 | 0.208588 | -5.594 | <0.0001*** |
| T. saldo migratorio | 0.0668850 | 0.0114849 | 5.824 | <0.0001*** |
| T. empleo mujeres | -0.0357169 | 0.00707470 | -5.049 | <0.0001*** |

Debido a que en la última columna aparecen en todas las variables los tres asteriscos, podemos decir que todas ellas son individualmente significativas a cualquier tipo de alfa.

| | | | |
|------------------------|-----------|-----------------------|-----------|
| Media de la vble. dep. | 2.940.800 | D.T. de la vble. dep. | 0.935236 |
| Suma de cuad. residuos | 0.791430 | D.T. de la regresión | 0.198926 |
| R-cuadrado | 0.962298 | R-cuadrado corregido | 0.954758 |
| F(4, 20) | 1.276.206 | Valor p (de F) | 6.16e-14 |
| Log-verosimilitud | 7.686.406 | Criterio de Akaike | -5.372812 |
| Criterio de Schwarz | 0.721567 | Crit. de Hannan-Quinn | -3.682491 |
| rho | 0.729373 | Durbin-Watson | 0.572334 |

Con los resultados obtenidos, la expresión el modelo de regresión ajustado sería:

$$\hat{Y} = 3.07178 + 0.234588X_1 - 1.16685X_2 - 1.16685X_3 - 0.0357169X_4$$

En primer lugar, se puede decir que el modelo es globalmente significativo, ya que se rechaza la hipótesis nula en el contraste de significación global dado por:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1: \beta_j \neq 0 \text{ al menos para algún } j=1, \dots, 4$$

Al ser el p-valor igual a $6.16e^{-14}$, inferior a cualquier nivel de significación usual. En el modelo planteado podemos observar que todas las variables son individualmente significativas para cualquier α y tiene una alta capacidad explicativa del 96,22% considerando que supera la marca establecida del 70%, por tanto, se puede decir que el modelo 3 es fiable.

Es importante destacar que los signos que acompañan a las distintas variables son coherentes.

4.3.1. Multicolinealidad

Factores de inflación de varianza (VIF).

| | |
|--------------------|-------|
| T. Fecundidad | 1.925 |
| T. Mortalidad | 1.674 |
| T.Sald. migratorio | 1.958 |
| T. Empleo mujeres | 1.269 |

Siendo $VIF(j) = 1/(1 - R(j)^2)$, donde $R(j)$ es el coeficiente de correlación múltiple entre las variables j y las variables independientes.

Como ninguno de los resultados son superiores a 10, valor a partir del cual existiría problema de multicolinealidad, podemos afirmar que el modelo 3 no presenta problemas de multicolinealidad entre las variables.

4.3.2. Especificación

El siguiente paso sería ver si el modelo está bien o mal especificado, por lo que a continuación realizaremos el Test Reset de Ramsey cuya hipótesis sería:

H_0 : El modelo está bien especificado

H_1 : El modelo está mal especificado

➤ **Contraste de especificación RESET (cuadrados y cubos)**

Estadístico de contraste: $F = 5.375548$, con valor $p = P(F(2,18) > 5.37555) = 0.0148$

➤ **Contraste de especificación RESET (cuadrados sólo)**

Estadístico de contraste: $F = 8.824941$, con valor $p = P(F(1,19) > 8.82494) = 0.00785$

➤ **Contraste de especificación RESET (cubos sólo)**

Estadístico de contraste: $F = 7.189674$, con valor $p = P(F(1,19) > 7.18967) = 0.0148$

Rechazo H_0 si p -valor $< \alpha$. Para afirmar que el modelo está bien especificado, no puede rechazarse H_0 en ninguna de las tres variantes del test, de otro modo el modelo estará mal especificado.

En nuestro caso el p -valor en cada una de las variantes es menor que cualquier α , por lo que el modelo está mal especificado.

Esto podría significar que hay un error de omisión de variable relevante, que en este caso no sería debido a que los errores estándares son reducidos y según los estudios teóricos dichas variables son las relevantes para el estudio demográfico. Un segundo error que podría ser es, una forma funcional incorrecta. Para saber qué forma funcional es la correcta puede realizarse un test de no- linealidad.

4.3.3. No- linealidad

Cuya hipótesis sería:

H_0 : el modelo es lineal

H_1 : el modelo es cuadrático,

o bien

H_1 : el modelo es lineal en logaritmos

Regresión auxiliar para el contraste de no linealidad (términos al cuadrado)

MCO, usando las observaciones 1992-2016 ($T = 25$)

Variable dependiente: uhat

| | Coeficiente | Desv. Típica | Estadístico T | Valor p |
|------------------------|--------------|--------------|---------------|----------|
| const. | -9.71893 | 470.480 | -0.2066 | 0.8389 |
| T. fecundidad | 0.330706 | 0.644203 | 0.5134 | 0.6147 |
| T. mortalidad | 0.0668147 | 105.606 | 0.03209 | 0.9748 |
| T. saldo migratorio | 0.0668147 | 0.0347260 | 1.924 | 0.0723* |
| T. empleo mujeres | 0.0518949 | 0.116063 | 0.4471 | 0.6608 |
| sq_T. fecundidad | -0.00377535 | 0.00714767 | -0.5282 | 0.6046 |
| sq_T. mortalidad | -0.0165754 | 0.643460 | -0.02576 | 0.9798 |
| sq_T. saldo migratorio | -0.00576835 | 0.00220100 | -2.621 | 0.0185** |
| sq_T. empleo mujeres | -0.000803100 | 0.00204729 | -0.3923 | 0.7000 |

Estadístico de contraste: $TR^2 = 9.50481$, con valor $p = P(\text{Chi-cuadrado}(4) > 9.50481) = \underline{\underline{0.0496485}}$

La hipótesis nula se rechaza cuando $p\text{-valor} < \alpha$, por lo que en este caso al ser $p\text{-valor} 0.0496485 < 0.05$, implica que se rechaza H_0 . En consecuencia, el modelo lineal no está bien especificado y debería contener términos cuadráticos.

Pero para asegurarnos más en cual debería ser la forma correcta del modelo, no bastándonos con lo anteriormente, realizaremos el mismo análisis, pero para términos logarítmicos.

Regresión auxiliar para el contraste de no linealidad (términos logarítmicos)

MCO, usando las observaciones 1992-2016 ($T = 25$)

| | Coeficiente | Desv. Típica | Estadístico t | Valor p |
|---------------------|-------------|--------------|---------------|---------|
| const. | -151.795 | 140.666 | -1.079 | 0.2956 |
| T. fecundidad | -1.07586 | 0.645952 | -1.666 | 0.1141 |
| T. mortalidad | -0.562545 | 11.7768 | -0.04777 | 0.9625 |
| T. saldo migratorio | -0.0204740 | 0.0168045 | -1.218 | 0.2397 |

| | | | | |
|---------------------|-----------|----------|---------|--------|
| T. empleo mujeres | -0.137928 | 0.127328 | -1.083 | 0.2938 |
| l_T. fecundidad | 48.6720 | 28.9423 | 1.682 | 0.1109 |
| l_T. mortalidad | 5.23106 | 96.5396 | 0.05419 | 0.9574 |
| l_T. empleo mujeres | 3.82523 | 3.51292 | 1.089 | 0.2914 |

Estadístico de contraste: $TR^2 = 3.9931$, con valor $p = P(\text{Chi-cuadrado}(3) > 3.9931) = \underline{\underline{0.26221}}$

En este caso el p-valor correspondiente es 0.26221, lo que implica que no se rechaza H_0 . En consecuencia, con lo dicho anteriormente ni el modelo lineal ni el logarítmico es correcto, por lo que al nuevo modelo que se plantee hay que añadirle el término cuadrático que corresponda.

A pesar de que en el contraste de no-linealidad con términos cuadrados nos ha salido el p-valor rozando, pero no llegando, a $\alpha=0.05$, comprobaremos si realmente es o no, el añadir términos al cuadrado, el problema de la mal especificación del modelo. Ya que se descartaría por datos evidente que no puede ser ni lineal ni logarítmico.

4.4. Modelo 4

En este modelo después del análisis realizado en el modelo anterior vemos que añadir al modelo la tasa del saldo migratorio al cuadrado para intentar solucionar el tema del error funcional, es lo más correcto debido a que en el test de no linealidad dicha variable nos sale significativa, cosa que las demás variables cuando las elevamos al cuadrado nos aparecen como no significativas.

Por lo que el cambio de añadir la tasa de saldo migratorio al cuadrado nos da el siguiente modelo:

$$Y = \beta_0 + \beta_1 \times_1 + \beta_2 \times_2 + \beta_3 \times_3 + \beta_4 \times_3^2 + \beta_5 \times_4 + u$$

Resultados de la estimación del modelo 4.

| | <i>Coeficiente</i> | <i>Desv. Típica</i> | <i>Estadístico t</i> | <i>valor p</i> |
|------------------------|--------------------|---------------------|----------------------|----------------|
| constante | 2.97163 | 196.618 | 1.511 | 0.1471 |
| T. fecundidad | 0.221807 | 0.0178798 | 12.41 | <0.0001*** |
| T. mortalidad | -1.12124 | 0.170838 | -6.563 | <0.0001*** |
| T. saldo migratorio | 0.145505 | 0.0254821 | 5.710 | <0.0001*** |
| Sq T. saldo migratorio | -0.00620938 | 0.00187139 | -3.318 | 0.0036*** |
| T. empleo mujeres | -0.0289235 | 0.00612772 | -4.720 | 0.0001*** |

Todas las variables son significativas individualmente, con cualquier tipo de alfa, ya que son tres los asteriscos que acompañan a los últimos datos.

| | | | |
|------------------------|-----------|-----------------------|-----------|
| Media de la vble. dep. | 2.940.800 | D.T. de la vble. dep. | 0.935236 |
| Suma de cuad. residuos | 0.501081 | D.T. de la regresión | 0.162397 |
| R-cuadrado | 0.976130 | R-cuadrado corregido | 0.969848 |
| F(5, 19) | 1.553.950 | Valor p (de F) | 9.97e-15 |
| Log-verosimilitud | 1.339.983 | Criterio de Akaike | -14.79967 |
| Criterio de Schwarz | -7.486415 | Crit. de Hannan-Quinn | -12.77128 |
| rho | 0.524178 | Durbin-Watson | 0.939371 |

La ecuación del modelo ajustado sería:

$$\hat{Y} = 2.97163 + 0.221807 X_1 - 1.12124 X_2 + 0.145505 X_3 - 0.00620938 X_3^2 - 0.0289235 X_4$$

Este modelo guarda relación con los anteriores en que es globalmente significativo ya que su p-valor es $9.97e^{-15}$ es inferior a cualquier alfa. Su capacidad explicativa es elevada ya que presenta un 97.61 % muy superior al mínimo establecido del 70%.

4.4.1. Especificación

H₀: El modelo está bien especificado

H₁: El modelo está mal especificado

➤ **Contraste de especificación RESET (cuadrados y cubos)**

Estadístico de contraste: $F = 1.770633$, con valor $p = P(F(2,17) > 1.77063) = 0.2$

➤ **Contraste de especificación RESET (cuadrados sólo)**

Estadístico de contraste: $F = 2.282690$, con valor $p = P(F(1,18) > 2.28269) = 0.148$

➤ **Contraste de especificación RESET (cubos sólo)**

Estadístico de contraste: $F = 1.676064$, con valor $p = P(F(1,18) > 1.67606) = 0.212$

Rechazo H_0 si $p\text{-valor} < \alpha$. Para afirmar que el modelo está bien especificado, no puede rechazarse H_0 en ninguna de las tres variantes del test, de otro modo el modelo estará mal especificado.

En nuestro caso, el p -valor de cada uno de los contrastes realizados, es decir, de cuadrados y cubos, sólo cuadrados y solo cubos, es mayor que cualquier α que pudiésemos utilizar para hacer el estudio. En consecuencia, podemos decir que el modelo ya está bien especificado.

4.4.2. Multicolinealidad

Después de comprobar que el modelo está bien especificado procedemos a estudiar la colinealidad entre variables, revelándose los siguientes resultados:

Factores de inflación de varianza (VIF)

| | |
|------------------------|--------|
| T. fecundidad | 2.019 |
| T. mortalidad | 1.685 |
| T. saldo migratorio | 14.463 |
| sq_T. saldo migratorio | 13.851 |
| T. empleo mujeres | 1.429 |

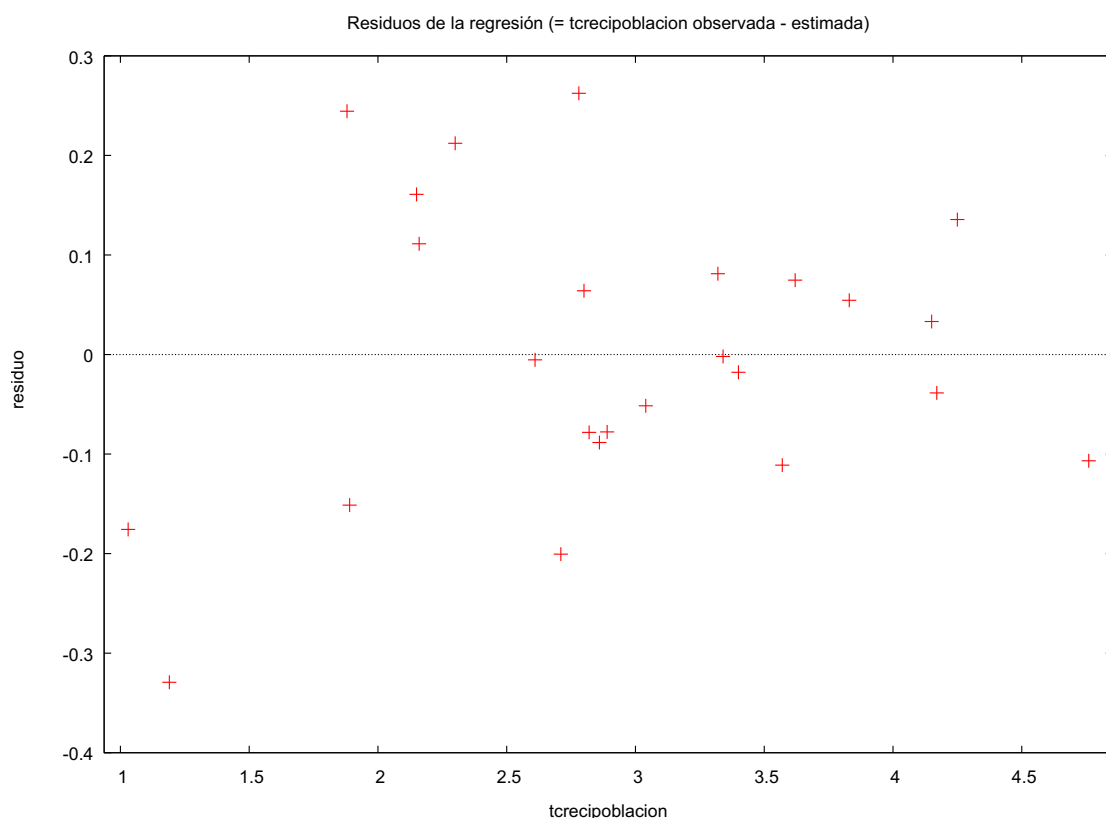
Siendo $VIF(j) = 1/(1 - R(j)^2)$, donde $R(j)$ es el coeficiente de correlación múltiple entre la variable j y las demás variables independientes.

Con los resultados obtenidos podemos observar que el problema que causaba anteriormente

la fertilidad al superar el valor 10 como índice máximo, ya se ha corregido. Ahora el problema sería la colinealidad que existe en el saldo migratorio, considero que es normal que supere el índice debido a que estamos hablando de la misma variable, pero una elevada al cuadrado, por lo que “ignoramos” dicho resultado y continuamos con el análisis.

4.4.3. Heterocedasticidad

En primer lugar, vamos a estudiar la posible presencia de heterocedasticidad a través de procedimientos subjetivos como son las gráficas. Para ello vamos a obtener una gráfica de residuos frente a la variable dependiente (en este caso la población).



A tenor de lo que muestra el gráfico, no se observa que los datos sigan ningún tipo de patrón que nos haga sospechar de la presencia de heterocedasticidad en el modelo. Pero para una comprobación más exhaustiva, vamos a proceder, en segundo lugar, a métodos objetivos, como son los contrastes de White, Breusch- Pagan y Koenker. La hipótesis que se plantea es la siguiente:

H_0 : El modelo es homocedástico

H₁: El modelo es heterocedástico

- Contraste de heterocedasticidad de White

MCO, usando las observaciones 1992-2016 (T = 25)

Variable dependiente: $uhat^2$

| | coeficiente | desviacion tipica | estadístico t | p-valor |
|------------------------|-----------------|----------------------|---------------|---------|
| constante | -714.692 | 303.712 | -2.353 | 0.0653 |
| T. fecundidad | 101.884 | 0.431715 | 2.360 | 0.0648 |
| T. mortalidad | 101.884 | 514.249 | 2.274 | 0.0721 |
| T. saldo migratorio | -116.839 | 0.416544 | -2.805 | 0.0378 |
| sq_T. saldo migratorio | 0.0612567 | 0.0324175 | 1.890 | 0.1174 |
| T. empleo mujeres | 0.244033 | 0.109240 | 2.234 | 0.0758 |
| sq_T. fecundidad | -0.00413471 | 0.00179710 | -2.301 | 0.0697 |
| X2_X3 | - 0.0757176 | 0.0334193 | -2.266 | 0.0728 |
| X2_X4 | 0.0207228 | 0.0104607 | 1.981 | 0.1044 |
| X2_X5 | - 0.00151644 | 0.000928766 | -1.633 | 0.1635 |
| X2_X6 | -0.00361063 | 0.00209351 | -1.725 | 0.1452 |
| Sq_T mortalidad | -0.509071 | 0.232794 | -2.187 | 0.0804 |
| X3_X4 | 0.0267777 | 0.0376471 | 0.7113 | 0.5087 |
| X3_X5 | -0.000299659 | 0.00325939 | -0.09194 | 0.9303 |
| X3_X6 | -0.00675085 | 0.00875795 | -0.7708 | 0.4756 |
| X4_X5 | 0.000262787 | 0.00119104 | 0.2206 | 0.8341 |
| X4_X6 | 0.00371268 | 0.00847039 | 0.4383 | 0.6795 |
| Sq_sq T. saldo migra. | -1,36E-02 | 3,76E-05 | -0.003604 | 0.9973 |
| X5_X6 | -2,33E+00 | 0.000571356 | -0.04072 | 0.9691 |

| | | | | |
|---------------------------|------------------|-------------|--------|--------|
| Sq_T empleo mujeres | - 0.000762609 | 0.000494831 | -1.541 | 0.1839 |
|---------------------------|------------------|-------------|--------|--------|

P-valor= **0.282337**

- Contraste de heterocedasticidad de Breusch-Pagan

MCO, usando las observaciones 1992-2016 (T = 25)

Variable dependiente: \hat{u}^2 escalado

| | Coefficiente | Desv. Típica | Estadístico t | Valor p |
|---------------------|--------------|--------------|---------------|---------|
| constante | -9.24986 | 138.047 | -0.6701 | 0.5109 |
| T. fecundidad | -0.0567881 | 0.125536 | -0.4524 | 0.6561 |
| T. mortalidad | 146.711 | 119.947 | 1.223 | 0.2362 |
| T. saldo migratorio | -0.276158 | 0.178912 | -1.544 | 0.1392 |
| sq_T saldmigratorio | 0.0112358 | 0.0131392 | 0.8551 | 0.4031 |
| T. empleo mujeres | 0.0521288 | 0.0430232 | 1.212 | 0.2405 |

P - valor= **0.127099**

- Contraste de heterocedasticidad de Koenker

MCO, usando las observaciones 1992-2016 (T = 25)

Variable dependiente: \hat{u}^2 escalado (variante robusta de Koenker)

| | Coefficiente | Desv. Típica | Estadístico t | P-valor |
|------------------------|--------------|--------------|---------------|---------|
| constante | -0.205440 | 0.276690 | -0.7425 | 0.4669 |
| T. fecundidad | -0.00113822 | 0.00251614 | -0.4524 | 0.6561 |
| T. mortalidad | 0.0294057 | 0.0240412 | 1.223 | 0.2362 |
| T. saldo migratorio | -0.00553510 | 0.00358597 | -1.544 | 0.1392 |
| sq_T. saldo migratorio | 0.000225202 | 0.000263352 | 0.8551 | 0.4031 |

| | | | | |
|-------------------|------------|-------------|-------|--------|
| T. empleo mujeres | 0.00104483 | 0.000862323 | 1.212 | 0.2405 |
|-------------------|------------|-------------|-------|--------|

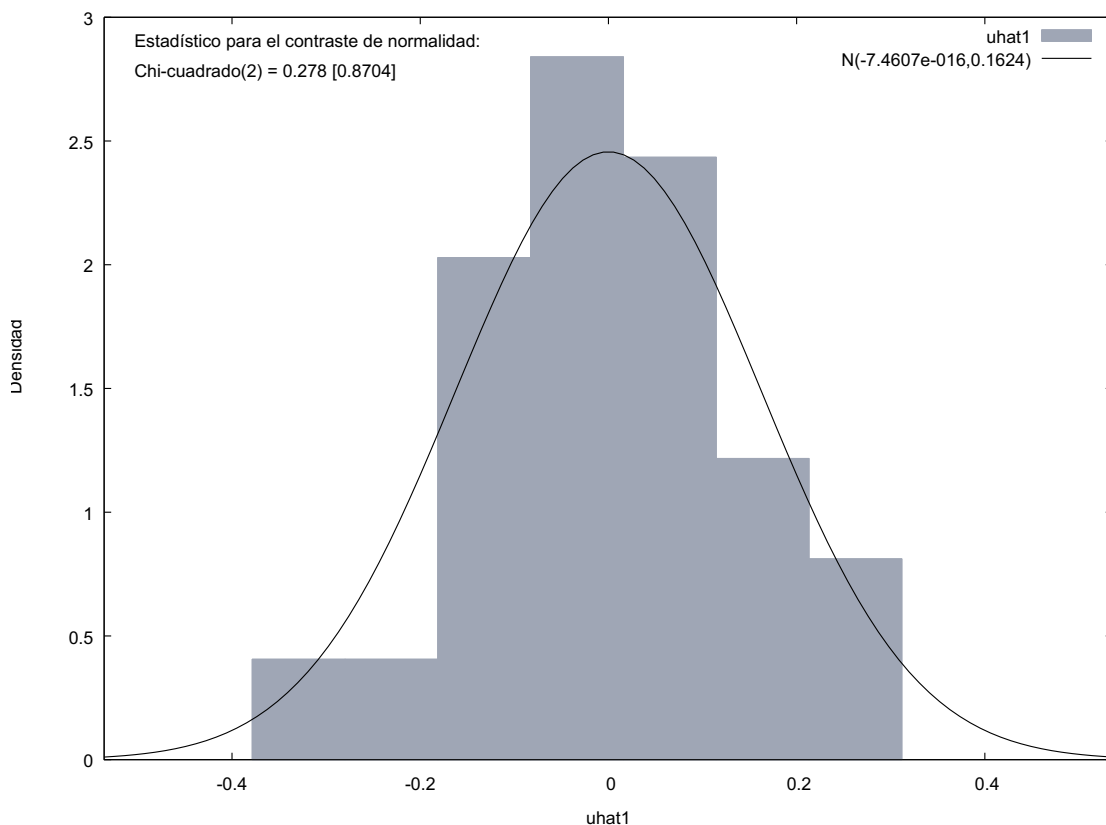
P- valor= **0.068526**

Después del estudio de los tres contrastes de heterocedasticidad del modelo, podemos concluir que todos los p. valores son superiores a los niveles de significación (1% y 5%). En nuestro estudio elegimos un $\alpha=0.05$, por lo que, siguiendo el criterio de rechazo, se rechaza H_0 si $p\text{-valor} < \alpha$ y en todos y cada uno de los contrastes planteados no se rechaza la hipótesis nula, lo que quiere decir que el modelo es homocedástico.

4.4.4. Normalidad de residuos

H_0 : El error se distribuye normalmente

H_1 : el error no se distribuye normalmente



En esta distribución de los residuos p-valor es 0.87044. como se puede ver, el p-valor es superior al nivel de significación del 5%, por tanto, no se rechaza la hipótesis nula, afirmando por tanto que los datos de este modelo se distribuyen normalmente.

Pero para observar la normalidad de los residuos son más precisos los contrastes de Doornik-Hansen, Shapiro-Wilk, Lilliefors y Jarque-Bera, cuyos p-vaes correspondientes son: 0.870436, 0.982898, 1 y 0.934906. Por lo que se corrobora que los datos del modelo se distribuyen normalmente ya que dichos contrastes siguen rechazando la hipótesis nula al ser sus p-valores superiores a $\alpha=0.05$.

4.4.5. Contraste de chow

Este contraste se utiliza cuando se sospecha que ha podido existir un cambio estructural en la muestra.

$$H_0: \beta_j = \beta'_j \quad j=0, \dots, k$$

H₁: Alguna pareja de coeficientes es distinta.

Regresión aumentada para el contraste de chow.

| | Coeficiente | Desv. Típica. | Estadístico t. | valor p |
|---------------------|-------------|---------------|----------------|------------------|
| const | 0.891468 | 1.85280 | 0.4811 | 0.6384 |
| T. fecundidad | 0.221053 | 0.0123321 | 17.93 | 1.51e-010 *** |
| T. mortalidad | -0.966851 | 0.195924 | -4.935 | 0.0003 *** |
| T.saldomigratorio | 0.00767016 | 0.0781462 | 0.09815 | 0.9233 |
| sq_T.saldomigrato~ | -0.00138717 | 0.00550510 | -0.2520 | 0.8050 |
| T. empleomujeres | 0.0260143 | 0.0207883 | 1.251 | 0.2328 |
| splitdum | -3.13708 | 2.62197 | -1.196 | 0.2529 |
| sd_T. fecundidad | 0.177273 | 0.0400832 | 4.423 | 0.0007 *** |
| sd_T. mortalidad | -0.297764 | 0.228164 | -1.305 | 0.2145 |
| sd_T.saldomigrato~ | -0.0292886 | 0.228164 | -0.3373 | 0.7413 |
| sd_sq_T. saldomigr~ | 0.00460762 | 0.00593427 | 0.7764 | 0.4514 |
| sd_T. empleomuj~ | -0.0826398 | 0.0298766 | -2.766 | 0.0160 ** |

| | | | |
|------------------------|-----------|-----------------------|-----------|
| Media de la vble. dep. | 2.940.800 | D.T. de la vble. dep. | 0.935236 |
| Suma de cuad. residuos | 0.086105 | D.T. de la regresión | 0.081385 |
| R-cuadrado | 0.995898 | R-cuadrado corregido | 0.992427 |
| F(11, 13) | 2.869.385 | Valor p (de F) | 1.22e-13 |
| Log-verosimilitud | 3.541.478 | Criterio de Akaike | -46.82955 |
| Criterio de Schwarz | -32.20304 | Crit. de Hannan-Quinn | -42.77278 |
| rho | -0.135854 | Durbin-Watson | 1.987666 |

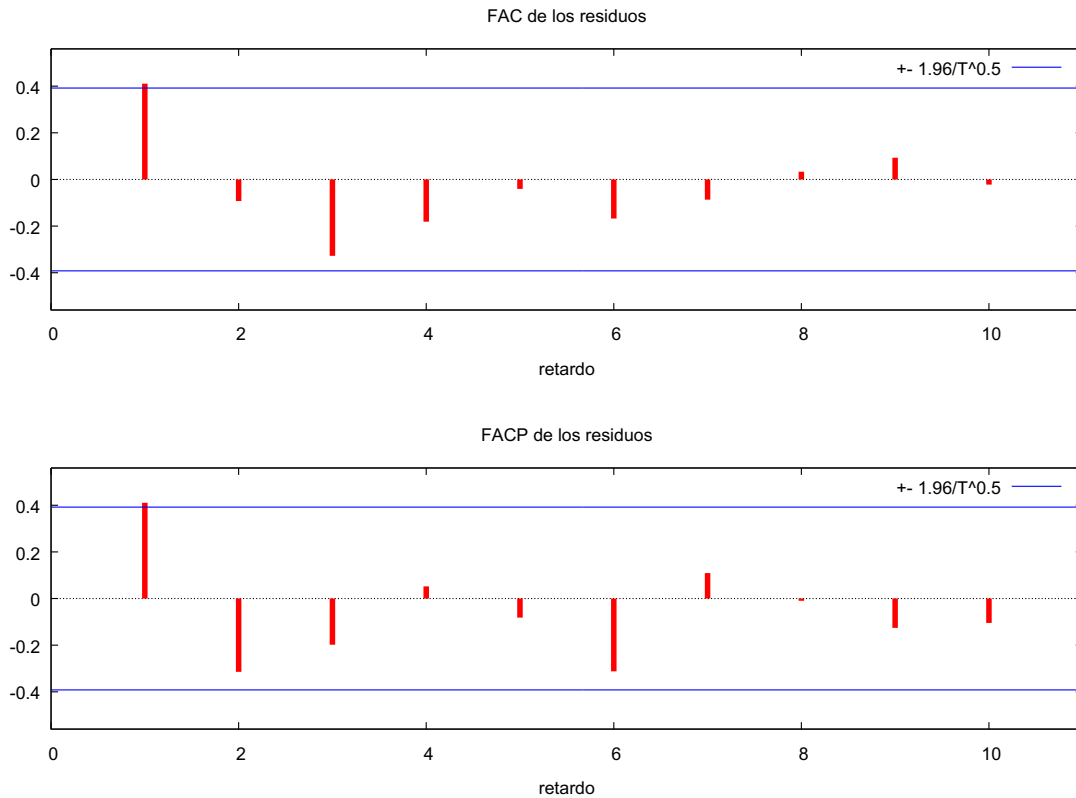
Contraste de Chow de cambio estructural en la observación 2004. $F(6, 13) = 10.442$ con **valor p 0.0002**.

Con el resultado obtenido decimos que al ser p-valor $(0.0002) < \alpha (0.05)$, se rechaza la hipótesis nula por lo que existe cambio estructural en la población a partir del año 2004.

4.4.6. Autocorrelación

A continuación, vamos a proceder a observar si el modelo presenta autocorrelación o no. Para ello primero vamos a utilizar el método subjetivo, el cual se basa en gráficos de residuos.

Para que las gráficas nos muestren la ausencia o presencia de autocorrelación en el modelo hay que obtener primero los residuos de dicho modelo, donde en las gráficas lleva el nombre (\hat{u}_t).



Para poder decir que gracias a las gráficas se afirma la presencia de autocorrelación en el modelo, tendrían que superarse las líneas azules, nivel que marca la confianza.

En este caso se aprecia que la primera barra supera las bandas azules de confianza, lo que indica que hay autocorrelación de orden 1.

Para hacer un estudio más objetivo de este problema, se hace uso de los procedimientos formales basados en los contrastes de Durbin-Watson, Wallis y por último Breusch-Godfrey. En nuestro caso al ser un problema de autocorrelación de orden 1 sólo podemos utilizar dos contrastes el de Durbin-Watson (ya que estudia la autocorrelación orden 1) y el de Breusch-Godfrey (cualquier autocorrelación), ya que el contraste de Wallis estudia solamente la autocorrelación de orden 4.

- Contraste de Durbin-Watson

$$u_t = \rho u_{t-1} + \varepsilon_t$$

↓

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Con el estadístico de contraste = 0.939371 y un p- valor = 2.24841e⁻⁰⁰⁵, menor que cualquier nivel de significación, pero más concretamente es menor que $\alpha=0.05$, se rechazaría la hipótesis nula, es decir, que existiría autocorrelación de orden 1 en nuestro modelo, tal y como nos decía el procedimiento subjetivo (las gráficas).

- Contraste de Breusch-Godfrey de autocorrelación de primer orden.

Para asegurarnos de que efectivamente existe autocorrelación de orden 1, vamos a realizar este segundo contraste, ya que por la información que recoge sus resultados son más fiables que en los del contraste anterior. El contraste a plantear sería el siguiente:

$$u_t = \phi_1 u_{t-1} + \varepsilon_t$$

↓

$$H_0: \phi_1 = 0$$

$$H_1: \phi_1 \neq 0$$

Como en este caso el estadístico de contraste es 7.695867 y su p-valor es 0.0125, se afirma que el p-valor sigue siendo menor al nivel de significación utilizado durante todo el estudio (5%) por lo que se rechazaría la hipótesis nula. Reafirmamos así que sí existe autocorrelación de orden 1 en nuestro modelo.

Después de haber agotado todos los recursos para ajustar la especificación del modelo la forma para eliminar los efectos de la estimación de un problema de autocorrelación consiste, evidentemente, en la utilización de Mínimos Cuadrados Generalizados Factibles (MCGF).

Por ello, para corregir la autocorrelación estimaremos el modelo mediante el método de Cochrane-Orcutt combinado con Hildreth-Lu.

- Método de Hildreth-Lu combinado con el ajuste fino de Cochrane-Orcutt.

Modelo 4: Resultado de la combinación de ambos.

rho = 0.9999

| | Coeficiente | Desv. Típica | Estadístico t | Valor p |
|------------------------|--------------|--------------|---------------|-----------|
| constante | -630.136 | 245.149 | -2.570 | 0.0193 |
| T. fecundidad | 0.223675 | 0.0178464 | 12.53 | 2.50e-010 |
| T. mortalidad | -1.01196 | 0.0592949 | -17.07 | 1.46e-012 |
| T. saldo migratorio | 0.0129830 | 0.0181466 | 0.7155 | 0.4835 |
| sq_T. saldo migratorio | -0.000655702 | 0.00100558 | -0.6521 | 0.5226 |

| | | | | |
|-------------------|-----------|-----------|-------|--------|
| T. empleo mujeres | 0.0344730 | 0.0215288 | 1.601 | 0.1267 |
|-------------------|-----------|-----------|-------|--------|

Estadísticos basados en los datos rho-diferenciados:

| | | | |
|------------------------|----------|-----------------------|----------|
| Media de la vble. dep. | 2.865000 | D.T. de la vble. dep. | 0.873389 |
| Suma de cuad. residuos | 0.109353 | D.T. de la regresión | 0.077943 |
| R-cuadrado | 0.994681 | R-cuadrado corregido | 0.993204 |
| F(5, 18) | 152.8782 | Valor p (de F) | 4.32e-14 |
| rho | 0.619684 | Durbin-Watson | 0.578840 |

Sin considerar la constante, el valor p más alto fue el de la variable 7 (sq_tsaldomigratorio)

Una vez hecho todo el análisis y llegar a un modelo correcto, el nuevo modelo de regresión ajustado sería:

$$\hat{Y} = -630.136 + 0.223675X_1 - 1.01196X_2 + 0.0129830X_3 - 0.000655702X_3^2 + 0.0344730X_4$$

Las interpretaciones a los resultados obtenidos serían:

$\hat{\beta}_1$ = indica que cuando la tasa de fecundidad en Andalucía aumenta en 1‰, la tasa de crecimiento de la población en Andalucía aumentará, en promedio, de 0.223675‰ manteniendo el resto de las variables constantes.

$\hat{\beta}_2$ = si la tasa de mortalidad en Andalucía aumenta en 1‰, la tasa de crecimiento de la población en Andalucía disminuirá, en promedio, 1.01196‰ manteniendo el resto de variables constantes.

Para la interpretación de la variación de la tasa de crecimiento de la población en Andalucía en función de la tasa de migración, se procedería a realizar la derivada y siempre dependerá del propio valor de partida de la tasa de migración. Dicha variación está dada por: $0.0129830X_3 - 2 * 0.000655702X_3$ estudia para qué valores de la tasa de migración el efecto sobre la población es negativo y para cuáles positivo.

$\hat{\beta}_4$ = si la tasa de mujeres empleadas en Andalucía aumenta en 1‰, la tasa de crecimiento de la población en Andalucía aumentará, en promedio, de 0.0344730‰, manteniendo el resto de variables constantes.

Y, por último, decir que el modelo tiene una alta capacidad explicativa de 99 % por lo que el modelo es fiable ya que se supera el mínimo establecido del 70%

5. CONCLUSIONES

La realización de este trabajo con el planteamiento de distintos modelos econométricos nos ha llevado hasta la obtención de un modelo que nos permite explicar la evolución de la tasa de crecimiento anual de la población en Andalucía en función de una serie de factores demográficos.

Gracias a esto, observamos que las variables que afectan al número de población de forma positiva, es decir, las variables que hacen que la población aumente son: la fertilidad, la incorporación de la mujer al mundo laboral y el saldo migratorio; con esta última se llega a la conclusión de que muchas personas procedentes del extranjero u otros lugares de la nación se establecen en Andalucía. Este hecho no resulta extraño, ya que muchos extranjeros vienen a Andalucía, más concretamente a la Costa del Sol, para pasar su vida de jubilados. Por otra parte, la variable que afectan de forma negativa al crecimiento de la población es la mortalidad, como es evidente.

Con respecto a la incorporación al mundo laboral de la mujer decir que, aunque influya de forma positiva a la población, lo hace de forma insignificante. Esto puede deberse al miedo que sienten las mujeres trabajadoras de poder perder su puesto de trabajo tras un embarazo, casos bastantes escuchados en los medios de comunicación. Se considera que dicho porcentaje insignificante aumente tras la lucha continuada de la protección de la mujer en España.

Por otro lado, se ha dejado a un lado la variable natalidad debido a que hace referencia a los nacimientos de forma global, o sea, nos da una visión más abstracta. En cambio, la fecundidad es un estudio más concreto ya que representa el número de hijos por mujer que se encuentra en edad fértil. Este grupo de edad se fija actualmente entre los 15 – 49 años.

El estudio de cada modelo no ha hecho sino mejorar la calidad de los resultados a la vez que su coherencia, puesto que en un primer momento a pesar de que los signos que acompañaban a los distintos resultados tenían sentido, la variable migración aparecía como no significativa, dicho resultado no tenía sentido, ya que distintos estudios sobre la demografía demuestran su importancia. En el segundo modelo añadimos la fecundidad, por su importancia anteriormente comentada, pero aparecía el problema de la colinealidad con la natalidad, por lo que en el tercer modelo eliminamos la natalidad, pero dicho modelo resultó mal especificado por un problema funcional que se solucionó elevando al cuadrado la tasa de migración, ya que en el análisis de no- linealidad que resultó significativa. En este último modelo nos encontramos con el problema de autocorrelación de orden 1, el cual se corrigió estimando el modelo por

mínimos cuadrados generalizados factibles, concretamente mediante el método de Cochrane-Orcutt combinado con Hildreth-Lu.

Sería muy interesante repetir este estudio de aquí a 5 años para poder observar si realmente se podría cambiar la influencia de la incorporación de la mujer al mundo laboral de aumentar de forma insignificante, en la actualidad, el número de nacimiento por mujer trabajadora a aumentar la media del número de hijo de forma significativa a pesar de estar trabajando.

6. BIBLIOGRAFÍA

Susino Arbucias, Joaquín (1993). *“Evolución reciente de la población y prevenciones”*. Ingrasa, Sevilla.

León Castillo, Luís A. (2015): *Análisis econométrico de la población demográfica*. Lambayeque, Perú.

Weeks, J. R. (1993). *Sociología de la población: introducción a los conceptos y cuestiones básicas*, Alianza, Madrid.

Lucas Marín, A. y Romero, J. (2006), *Estructura social. La realidad de las sociedades avanzadas*, Madrid, Pearson” Población y Demografía”.

Arias Veira, Pedro. (2002). *“Diferencias entre Tasa de fecundidad y Tasa de natalidad”* Artículo en el periódico la Voz De Galicia

INE: *“Demografía y Población”*. <https://www.ine.es/>

Instituto de Estadística y Cartografía de Andalucía.

<https://www.juntadeandalucia.es/institutodeestadisticaycartografia/padron/index.htm>

Rodríguez Saldaña, Mac. (1994): *“Política familiar y demografía”*, en BOROBIO, D. (Coord): *La familia en un mundo cambiante*, Publicaciones Universidad pontificia de Salamanca.

VARIOS AUTORES (1990): *“Mujer y demografía”*. *Ministerio de Asuntos Sociales*. Instituto de la mujer, Madrid.