

Modelo Econométrico

Trabajo Final

**El tiempo empleado en la realización de
trabajos Universitarios**

Cháve Jiménez, María

Gallardo Millán, Jose

Madrid Barrera, Antonio

Salas Laureano, Antonio

Índice

1. Introducción

- Tema
- Motivación y objetivos
- Estudios Previos
- Hipótesis de partida

2. Datos

- Metodología
- Definición de Variables

3. Análisis del modelo econométrico

- Resumen estadísticos descriptivos
 - Gráficos de cajas
 - Gráficos de Dispersión X-Y
- Planteamiento del modelo econométrico
 - Análisis del Modelo
 - Corrección del Modelo.

4. Resultados

- Normalidad
- Multicolinealidad
- Heterocedasticidad.
 - Método Grafico
 - Test de White
 - *Breuch – Pagan*

5. Conclusiones

6. Bibliografía

1-Tema

Ya llega la cuesta de enero, y desde octubre hemos estado pensando; tras aquellos maravillosos días de septiembre, entre apuntes y subrayadores, machacándonos a estudiar; en ponernos las pilas desde el principio. Yendo a clase desde el primer día, cogiendo apuntes con mucho esmero, y demás historias, que a medida que entraba el otoño iban cayendo en el olvido, resguardándonos del frío en la calle entre las cañas y el cachondeo.

Entonces, ¿qué nos pasa? Lo más seguro es que cada uno sea diferente, tenga sus problemas, ya sea la debilidad por perder el tiempo mirando a través de la ventana durante las clases, o la de pararse frente los apuntes esperando descifrar de entre letras y números el dibujo oculto. Aún así, a continuación presentaremos un modelo econométrico que a lo mejor resuelva nuestras dudas.

1.1 Motivo y Objetivo del trabajo.

Realizamos este trabajo, para saber si conociendo como empleamos nuestro tiempo, podremos olvidarnos de estresados días de enero, o los calurosos de agosto. Por ello, a través de las variables seleccionadas, podremos conocer para nuestro futuro universitario como llevar a cabo de la mejor manera nuestro trabajo de estudiar.

Todo empezó el día en que el profesor de econometría dijo: debéis elaborar un modelo econométrico. Desde aquel principio de octubre todos sabíamos que el trabajo tenía que acabarse, pero hasta el día de su finalización han surgido muchos problemas. Diferentes factores interponiéndose para su realización, como en cualquier otra asignatura. El ocio, la natural holgazanería del individuo, el nivel de importancia que le das, si tienes pareja o no...

Por lo que variables como estas, y otras que aluden a la media de la carrera, el sexo o la estructura del grupo, ¿influirán en el modo en el que utilizamos el tiempo para realizar un trabajo, antes de su fecha de entrega?

1.2 Estudio Previos

Obtener información previa referente al tema, ha sido un poco complicado, ya que, mayormente este estudio se basa en estereotipos sociales o también en la simple cotidianidad de nuestras vidas.

A pesar de ello, ha habido algunos puntos en la red, en la que hemos obtenido información de gran valor para nuestro estudio.

1.3 Hipótesis de partida.

-Plazo (en días): cuanto más tiempo tengamos hasta la entrega del trabajo, más tiempo tenemos para el comienzo de este. (Signo positivo del parámetro).

-Pareja: Los estudiantes con pareja tardaran más en empezar el trabajo que el estudiante que este soltero. (Signo positivo del parámetro).

-Sexo: La percepción de que las mujeres son más organizada que los hombres, nos hace pensar que se administraran mejor el tiempo que los hombres, aunque esto no es una regla fija. Por ello pensamos que el tiempo que tarda una mujer en empezar un trabajo será menor que el de un hombre. (Signo positivo del parámetro).

-Estructura: Al ser un trabajo hecho en grupo, hay más personas para ponerse de acuerdo para repartirse las tareas, empezar el trabajo, quedar para trabajar... (Signo positivo del parámetro)

-Actividades: Cuando un alumno tiene actividades extraescolares, tiene menos tiempo para realizar trabajos, ya que tiene bastante tiempo ocupado con otras obligaciones, aun así, los realiza, pero el tiempo hasta que comienza la realización aumentara. (Signo positivo del parámetro).

-Nivel de vago: cuanto más flojo sea el estudiante, más tiempo pasará hasta que este empiece el trabajo. (Signo positivo del parámetro).

-Nota media: Cuanto mayor sea la nota media que ha obtenido a lo largo de su carrera estudiantil, menor será el tiempo que transcurrirá hasta el comienzo del trabajo. (Signo positivo del parámetro).

-Grado de importancia: cuanto mayor sea la importancia del trabajo sobre la nota final de la asignatura, menor será el tiempo que transcurrirá desde que se manda hasta su comienzo. (Signo positivo del parámetro).

2. Datos

2.1 Metodología Empleada.

Hemos utilizado una muestra de 115 personas para llevar a cabo este estudio, Estas personas respondieron a una encuesta enviada a alumnos universitarios a través de GoogleDocs. Además, para recoger y realizar el estudio, utilizaremos los programas Excel y GRETL respectivamente.

2.2 Definición de las variables.

$$\text{Tiempo}_i = \beta_1 + \beta_2 \text{Plazo}_i + \beta_3 \text{Pareja}_i + \beta_4 \text{Sexo}_i + \beta_5 \text{Estructura}_i + \beta_6 \text{Actividades}_i + \beta_7 \text{Vagancia}_i + \beta_8 \text{Notamedia}_i + \beta_9 \text{Importancia}_i + u_i$$

Nombre de la variable	Tipo de Variable	Descripción
Tiempo	Endógena/Cuantitativa	Tiempo entre el día que mandan el trabajo, y en el que decides empezarlo.
Plazo	Explicativa/Cuantitativa	Tiempo desde que te lo manda, hasta el día de entrega
Pareja	Explicativa/dicotómica	Referida a si el individuo tiene o no pareja codificándose 1 como “sí” y 0 como “no”.
sexo	Explicativa/dicotómica	Referida a sexo del estudiante estableciendo 1 para “hombre” y 0 para “mujer”.
estructura	Explicativa/dicotómica	Hace referencia al número de personas que participan en el trabajo codificándola como 1 para “grupo” y 0 para “individual”.
Actividades	Explicativa/Cuantitativa	Horas dedicadas, de media al día, a realizar actividades no relacionadas con los estudios (clases de inglés, gimnasio, prácticas,...).
Vagancia	Explicativa/Cuantitativa	Nivel de holgazanería presentada por el estudiante, a partir de una autoevaluación en una escala de 1 a 10.
Nota media	Explicativa/Cuantitativa	Nota media actual de la carrera, media de 0 a 10
Grado de importancia	Explicativa/Cuantitativa	importancia del trabajo para la nota final en un rango de 1 a 5

Con esta tabla podemos ver como hay más estudiantes con pareja que realizan nuestro estudios, que los que se encuentran solteros, ya que el valor de la media de esta variable es superior a 0,5, por lo que en la tabulación se aproxima a 1= Pareja.

También se puede observar como el tiempo medio para comenzar un trabajo es de 15 días, y que el máximo en el que los estudiantes lo comienzan es de 32 días.

Un dato curioso que podemos observar es que la variable vagancia se sitúa en un valor de 5,23, por lo que podemos deducir que hay bastantes estudiantes que le cuesta arrancar en la elaboración de un trabajo, en nuestro estudio.

Se observa como las variables Tiempo, Actividad, Nota media e importancia presentan una asimetría positiva hacia la derecha, y el resto de variables presenta asimetría negativa hacia la izquierda.

3-Análisis del Modelo Económico

3.1 Resumen estadísticos Descriptivos.

Estadísticos principales, usando las observaciones 1 - 115

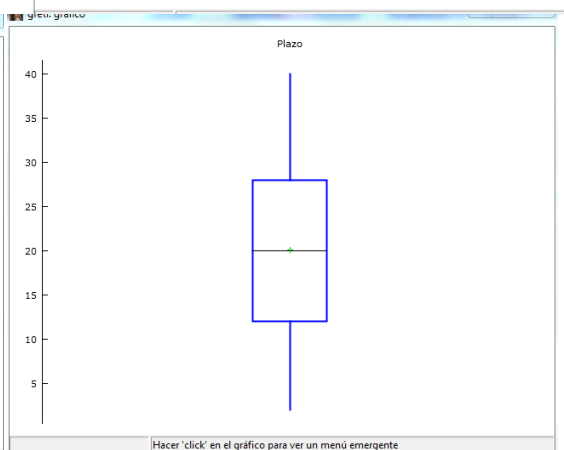
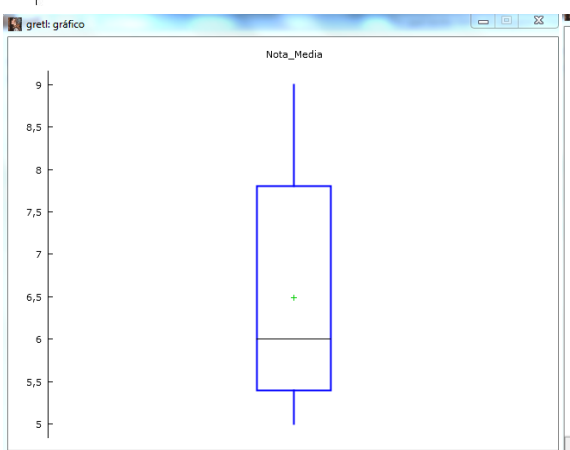
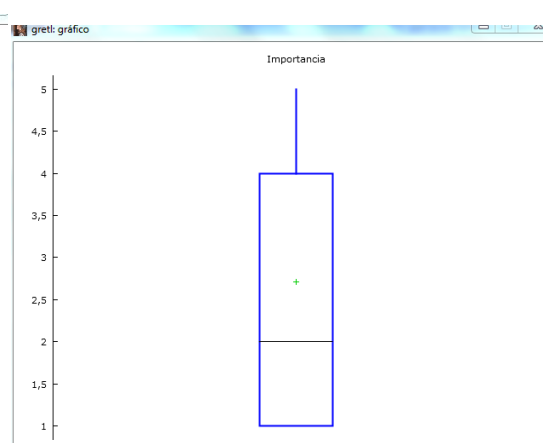
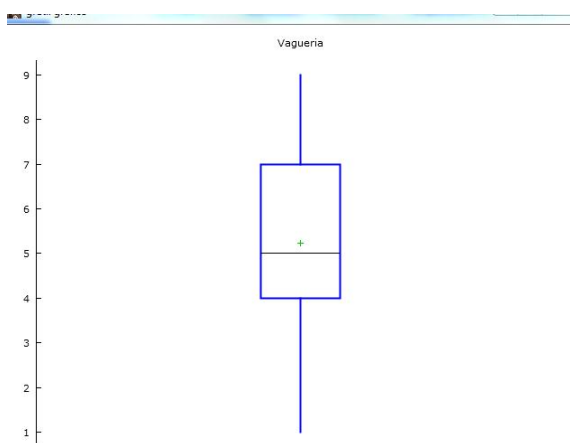
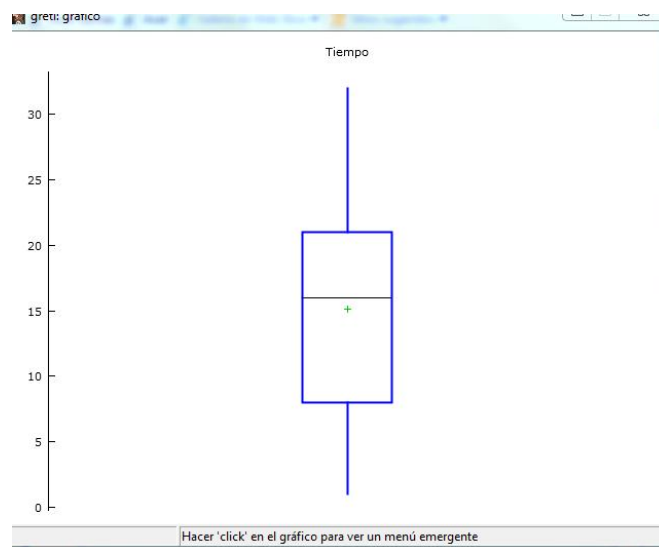
	Media	Mediana	Mínimo	Máximo
Tiempo	15,165	16,000	1,0000	32,000
Plazo	20,061	20,000	2,0000	40,000
Pareja	0,57391	1,0000	0,00000	1,0000
Sexo	0,53043	1,0000	0,00000	1,0000
Estructura	0,59130	1,0000	0,00000	1,0000
Actividades	2,1435	2,0000	0,00000	6,0000
Vagueria	5,2348	5,0000	1,0000	9,0000
Nota_Media	6,4913	6,0000	5,0000	9,0000
Importancia	2,7130	2,0000	1,0000	5,0000

	Desv. Típica.	C.V.	Asimetría	Exc. de curtosis
Tiempo	8,6012	0,56717	0,13669	-1,0069
Plazo	9,2918	0,46318	-0,16113	-0,99165
Pareja	0,49667	0,86541	-0,29894	-1,9106
Sexo	0,50126	0,94499	-0,12197	-1,9851
Estructura	0,49374	0,83501	-0,37146	-1,8620
Actividades	1,3777	0,64275	0,45954	-0,38082
Vagueria	2,0873	0,39873	-0,22701	-0,67529
Nota_Media	1,2105	0,18648	0,29746	-1,2829
Importancia	1,4374	0,52980	0,36929	-1,1950

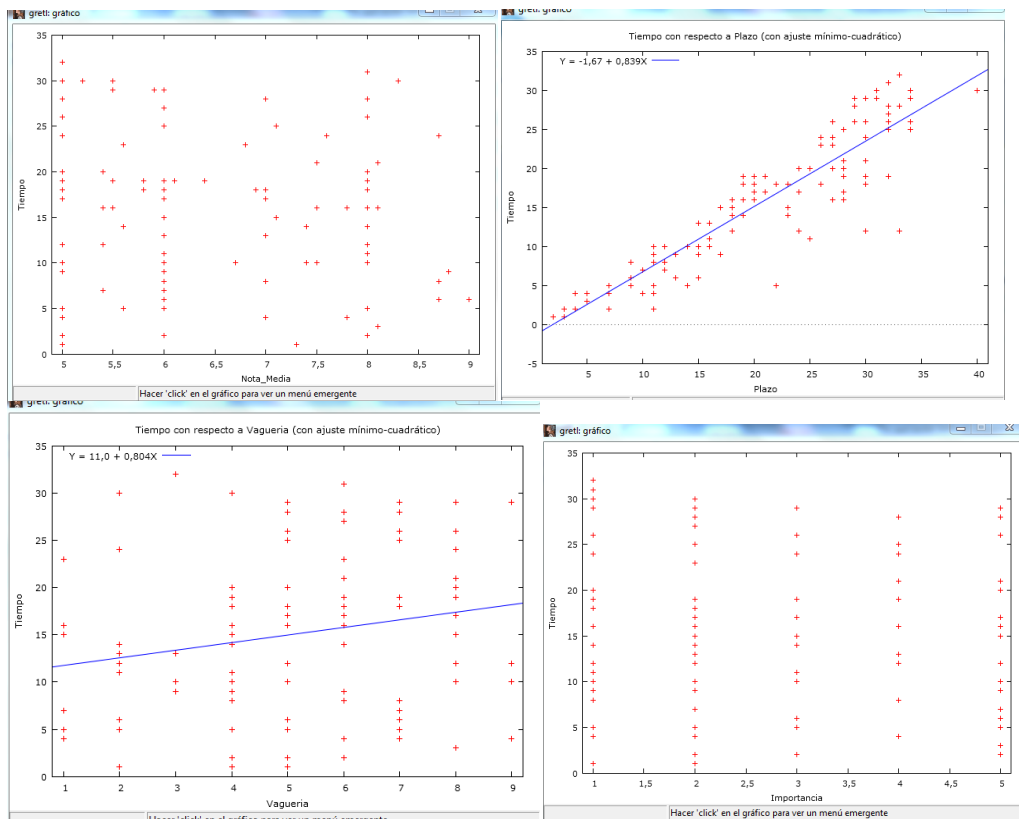
3.1.1 Gráficos de cajas

Los gráficos de cajas muestran las variables Vaguería, Tiempo, Plazo, Nota Media.

En estos gráficos se muestran los valores máximos y mínimos de las variables, además del primer y tercer cuartil, la mediana y la media. Con ello podremos observar como se distribuyen los datos a lo largo de la base de datos.



3.1.2 Gráficos de Dispersión X-Y



3.1.3 Matriz de Correlaciones

gretl: matriz de correlación

Coeficientes de correlación, usando las observaciones 1 - 124
valor crítico al 5% (a dos colas) = 0,1764 para n = 124

Plazo	Pareja	Sexo	Estructura
1,0000	-0,0445	0,0226	0,0566
	1,0000	0,0415	0,0420
		1,0000	-0,0304
			1,0000

Actividades	Vagueria	Nota_Media	Importancia
-0,0682	0,1512	0,0261	-0,0388
-0,1058	0,1176	-0,0777	0,0602
0,0425	0,0460	0,0535	0,0603
0,0187	0,0523	0,0590	0,0255
1,0000	0,1109	0,1123	0,1161
	1,0000	-0,0116	0,1419
		1,0000	-0,0209
			1,0000

Esta matriz muestra las relaciones existentes entre las variables del modelo, es decir cuánto dependen unas de otras. Los datos superiores al 0.8 denota una dependencia alta, es decir, una dependen mucho de otra.

Debemos de tener en cuenta que los valores altos también pueden significar la presencia de multicolinealidad, en nuestro caso, ningún dato supera este valor, por lo que en principio nuestro modelo no sufriría de multicolinealidad.

3.2 Planteamiento del Modelo Econométrico.

gretl: modelo 2

	Coefficiente	Desv. Típica	Estadístico t	Valor p	
const	4,72149	2,05679	2,296	0,0237	**
Plazo	0,835959	0,0332754	25,12	1,92e-046	***
Pareja	0,488099	0,633359	0,7707	0,4426	
Sexo	-0,495215	0,613365	-0,8074	0,4213	
Estructura	0,896051	0,617952	1,450	0,1500	
Actividades	-0,349730	0,229750	-1,522	0,1309	
Vagueria	0,268246	0,152162	1,763	0,0808	*
Nota_Media	-0,824950	0,256133	-3,221	0,0017	***
Importancia	-0,801020	0,216235	-3,704	0,0003	***
Media de la vble. dep.	15,16522	D.T. de la vble. dep.	8,601235		
Suma de cuad. residuos	1121,290	D.T. de la regresión	3,252416		
R-cuadrado	0,867049	R-cuadrado corregido	0,857015		
F(8, 106)	86,41077	Valor p (de F)	6,55e-43		
Log-verosimilitud	-294,1229	Criterio de Akaike	606,2457		
Criterio de Schwarz	630,9501	Crit. de Hannan-Quinn	616,2731		

Sin considerar la constante, el valor p más alto fue el de la variable 3 (Pareja)

Los datos anteriormente mostrados en la tabla respecto a las variables muestran que:

- La variable Plazo (Medido en días): es decir, si el plazo aumenta en un día, el tiempo hasta que se comenzó a realizar aumentará en 0,835959 días.

-La variable Relación: es decir, si los estudiantes tiene una relación, tardará 0,488099 días más que otro estudiante que no tenga pareja. (El parámetro tiene signo positivo).

-La variable Sexo: Si el estudiante es mujer, tardará 0,495215 días menos en empezar un trabajo que si el estudiante es hombre.

-La variable Estructura: si el trabajo debe realizarse en grupo, se empleara 0,896051 días más para empezar si fuera individual. (El parámetro tiene signo positivo).

-La variable Actividad: si aumentan en una unidad las horas al día dedicadas a actividades, el tiempo hasta que se empieza el trabajo disminuirá en 0,349730

días.

-La variable Vaguería: si la vaguería aumenta en un punto, el tiempo que tardará el estudiante en comenzar el trabajo aumentará en 0,268246 días.

-La variable Nota media: por cada punto que aumente la nota media aumenta en un punto, el tiempo hasta comenzar el trabajo disminuirá en 0,824950 días.

-La Grado de importancia: cuando aumenta la importancia en punto, el tiempo disminuirá en 0,801020 días.

El R^2 de este modelo sería de 0,867049, respecto al R^2 corregido podemos decir que su valor es 0,857015

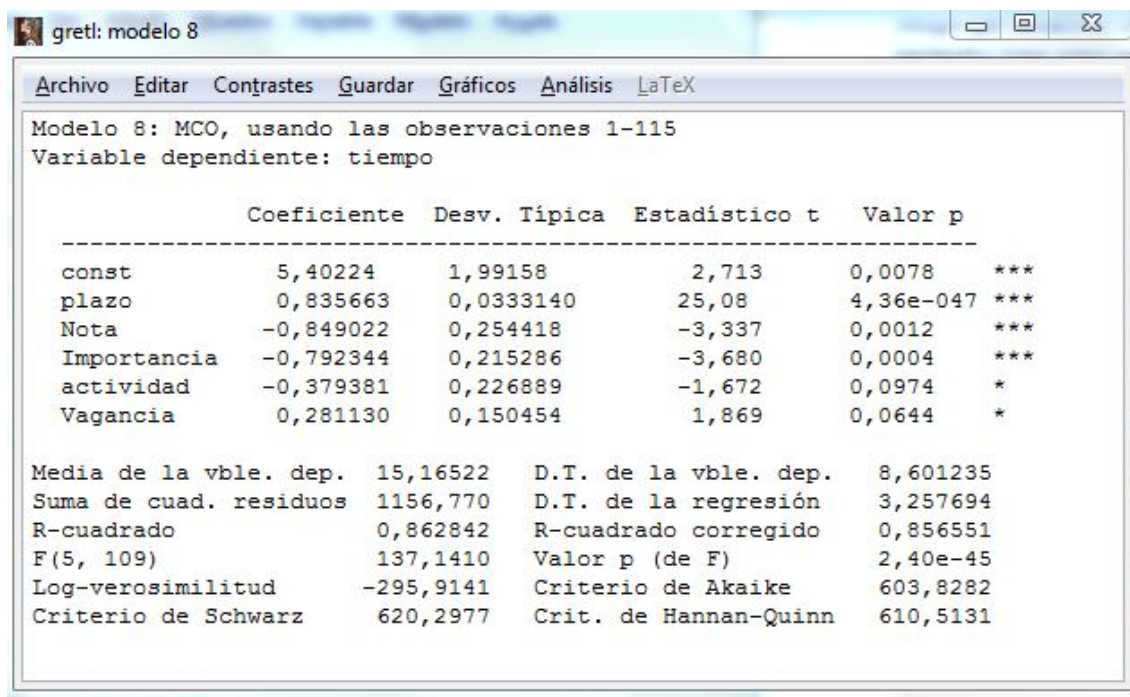
Se puede decir que las variables plazo, nota media e importancia son significativa, ello se puede ver en la tabla con la presencia de tres asteriscos en las mencionadas variables.

La vagancia en cambio solo es significativa para un nivel de significación de un 10, en este caso se puede observar la existencia de un solo asterisco.

En cambio, las variables pareja, sexo, estructura y actividades no son significativos individualmente, por ello no constan con de ningún asteriscos.

El p-valor de F es de $6,55e^{-43}$, por ello, podemos decir que el modelo es conjuntamente significativo.

3.2.2 Modelo corregido



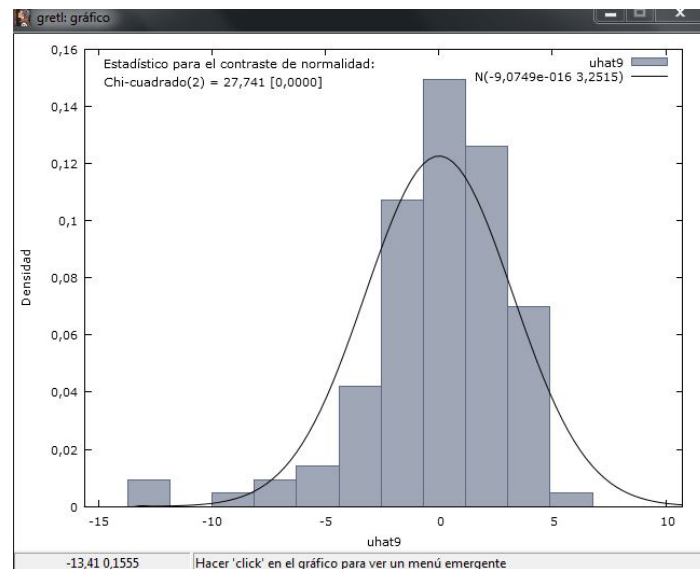
	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	5,40224	1,99158	2,713	0,0078	***
plazo	0,835663	0,0333140	25,08	4,36e-047	***
Nota	-0,849022	0,254418	-3,337	0,0012	***
Importancia	-0,792344	0,215286	-3,680	0,0004	***
actividad	-0,379381	0,226889	-1,672	0,0974	*
Vagancia	0,281130	0,150454	1,869	0,0644	*

Media de la vble. dep.	15,16522	D.T. de la vble. dep.	8,601235
Suma de cuad. residuos	1156,770	D.T. de la regresión	3,257694
R-cuadrado	0,862842	R-cuadrado corregido	0,856551
F(5, 109)	137,1410	Valor p (de F)	2,40e-45
Log-verosimilitud	-295,9141	Criterio de Akaike	603,8282
Criterio de Schwarz	620,2977	Crit. de Hannan-Quinn	610,5131

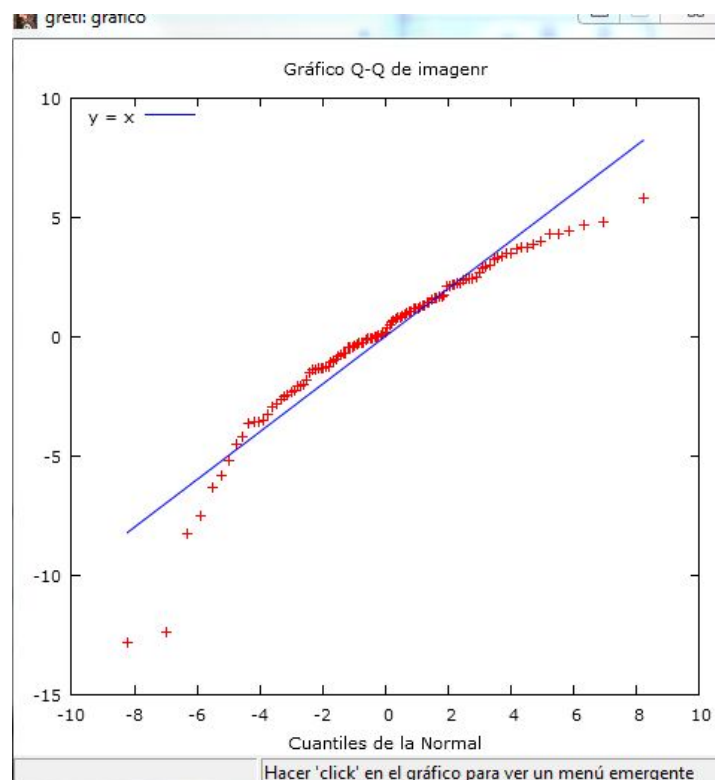
Tras realizar el modelo, sacando de las variables que no eran significativas en sí mismas, hemos podido observar que en sí, el modelo explica algo más con las variables que no son significativas, por lo que seguiremos trabajando con ellas en el resto de procesos de nuestro estudio.

4. Resultados

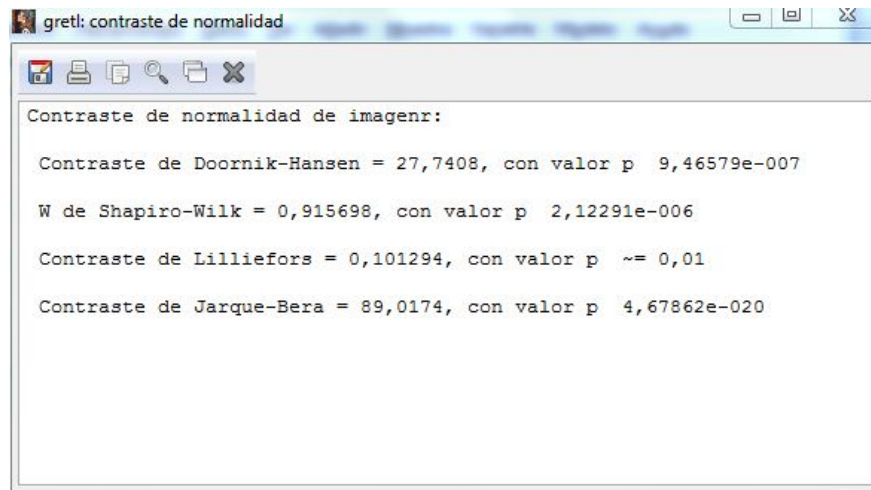
4.1 Normalidad



En la gráfica anterior, se puede observar la distribución de los residuos, en este caso podemos intuir que no actúa como la normal, dado que existe una asimetría hacia la izquierda y es leptocurtida.

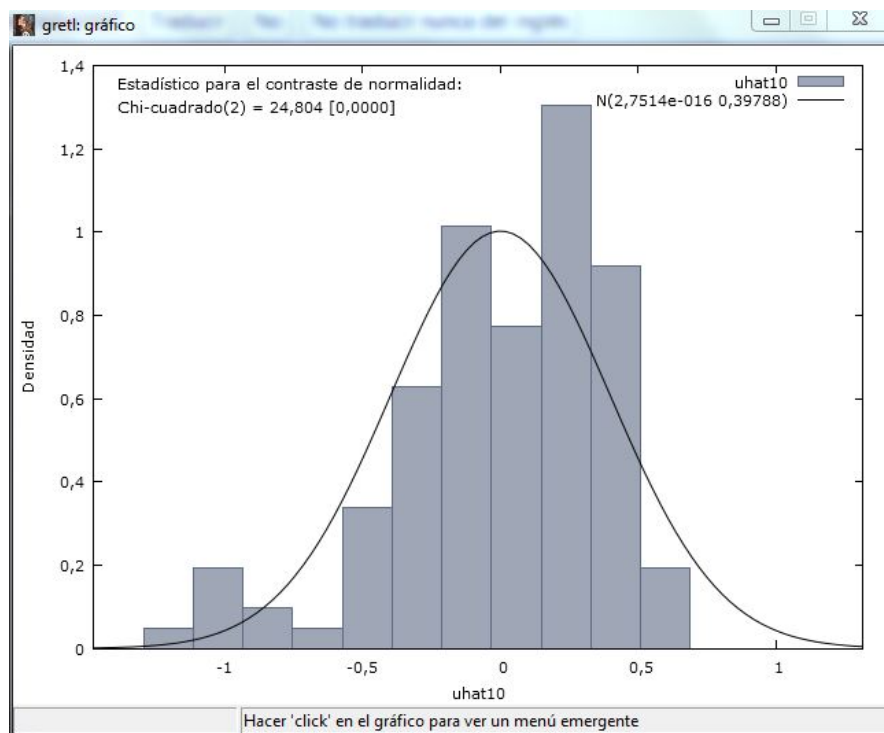


En el gráfico Q-Q, se puede observar que los residuos no se ajustan del todo a la diagonal, otro síntoma de falta de normalidad.

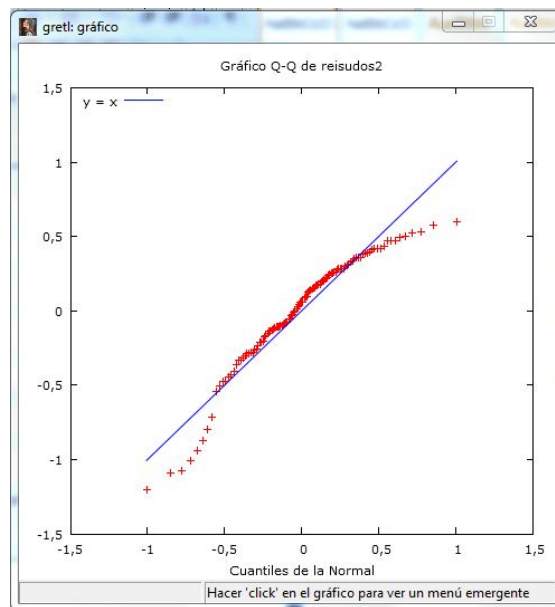


Tras hacer el contraste de normalidad, el P valor es de 4.67862-020, es decir, no podemos rechazar la hipótesis nula, No existe normalidad en el modelo.

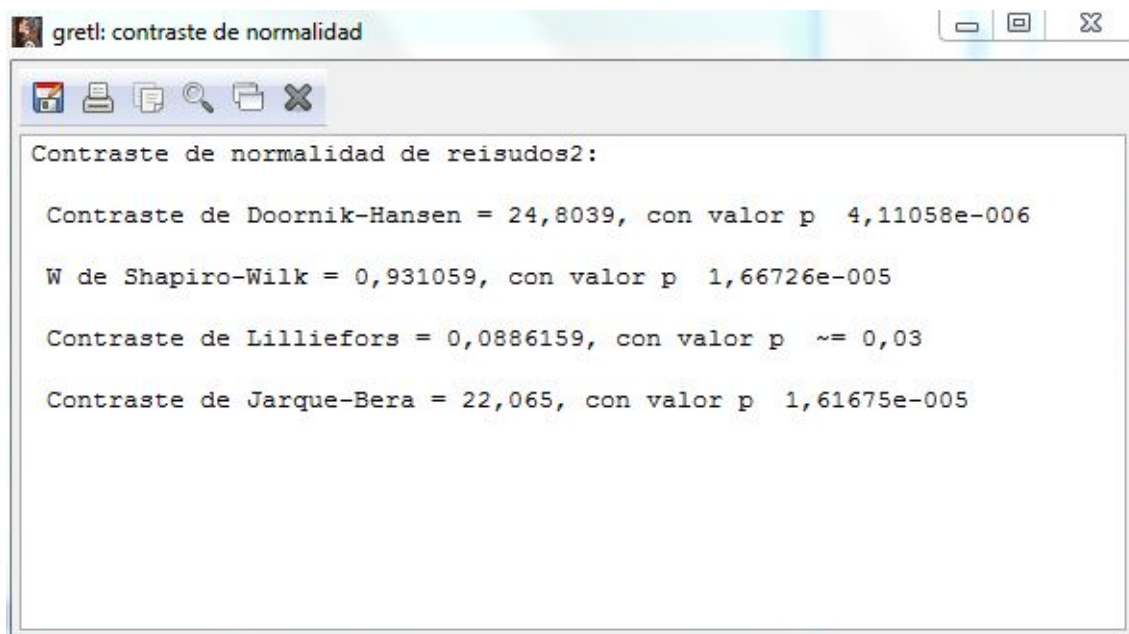
Para intentar solucionarlo, generaremos una nueva variable, esta vez trabajaremos con el logaritmo de la variable, realizaremos el modelo y obtendremos los mismos gráficos que anteriormente.



En este caso, podemos ver que la grafica ha mejorado, mejorando la asimetría, del mismo modo que en el caso anterior, realizaremos el grafico Q-Q

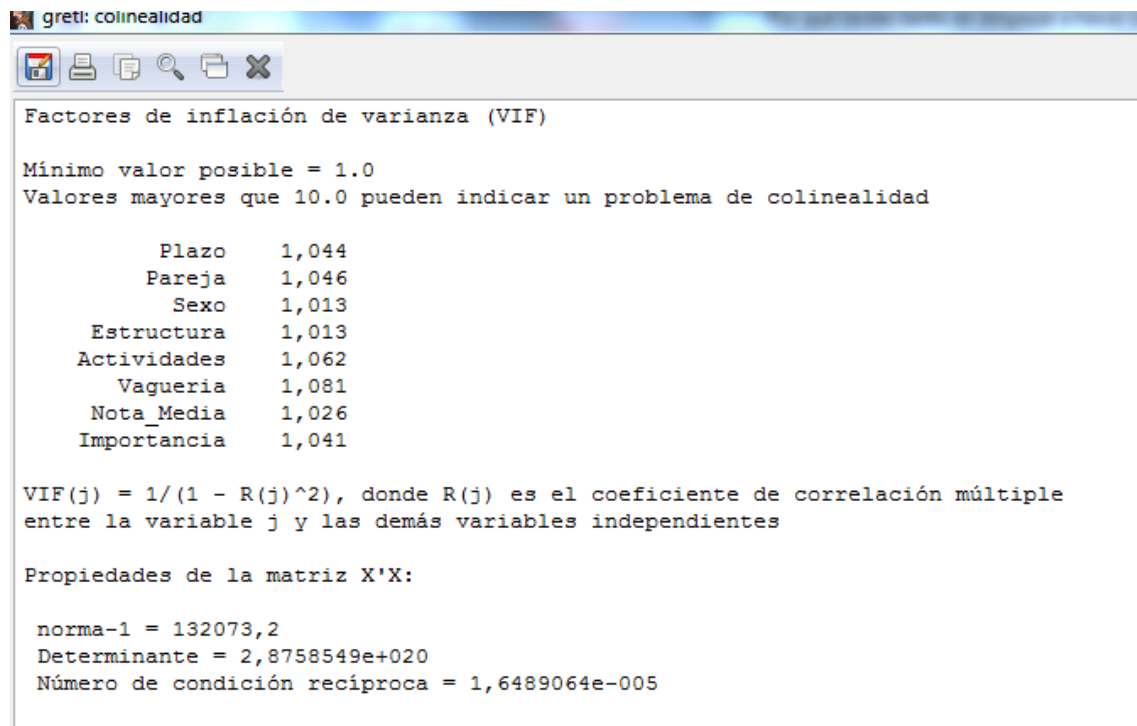


En este grafico podemos confirmar que la situación de los residuos han mejorado, aunque sigue habiendo observaciones muy separadas de la diagonal.



Tras realizar por segunda vez el contraste de normalidad, la prueba da una P valor de 1.61675-005, una vez más, no se puede rechazar H_0 , por lo que no existe normalidad en este modelo.

4.2 Multicolinealidad.



```
gretl: colinealidad

Factores de inflación de varianza (VIF)

Mínimo valor posible = 1.0
Valores mayores que 10.0 pueden indicar un problema de colinealidad

      Plazo      1,044
      Pareja     1,046
      Sexo       1,013
      Estructura 1,013
      Actividades 1,062
      Vagueria   1,081
      Nota_Media 1,026
      Importancia 1,041

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

Propiedades de la matriz X'X:

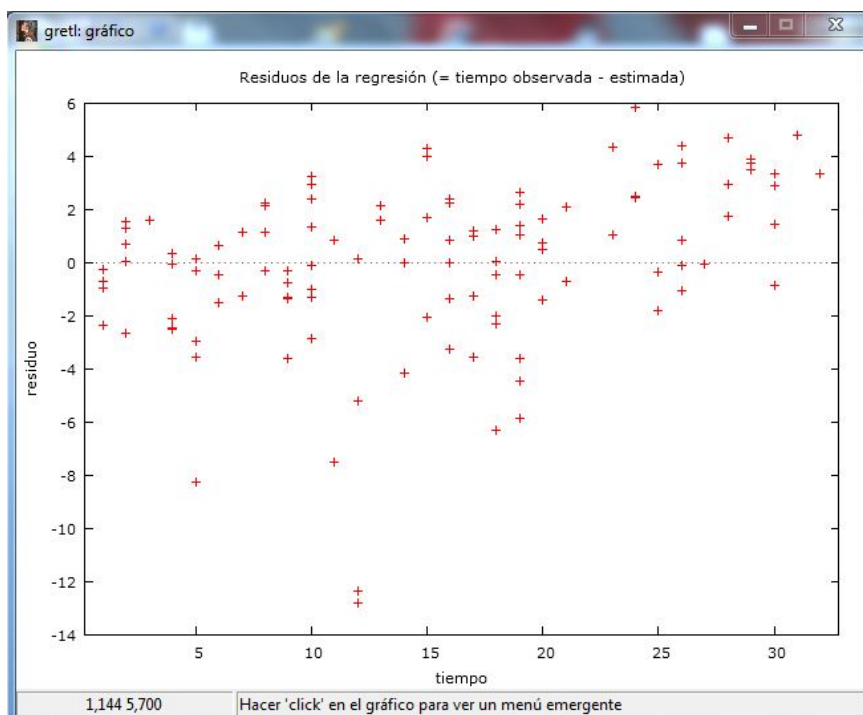
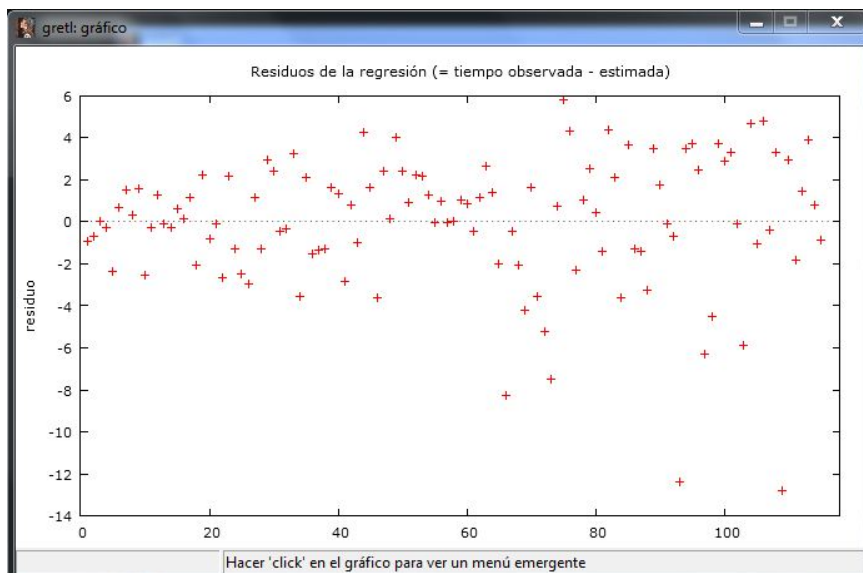
norma-1 = 132073,2
Determinante = 2,8758549e+020
Número de condición recíproca = 1,6489064e-005
```

Tras realizar un test de colinealidad, podemos observar que ninguna de nuestras variables da un valor superior a 10, ello denota la no existencia de colinealidad. .

4.3 Heterocedasticidad

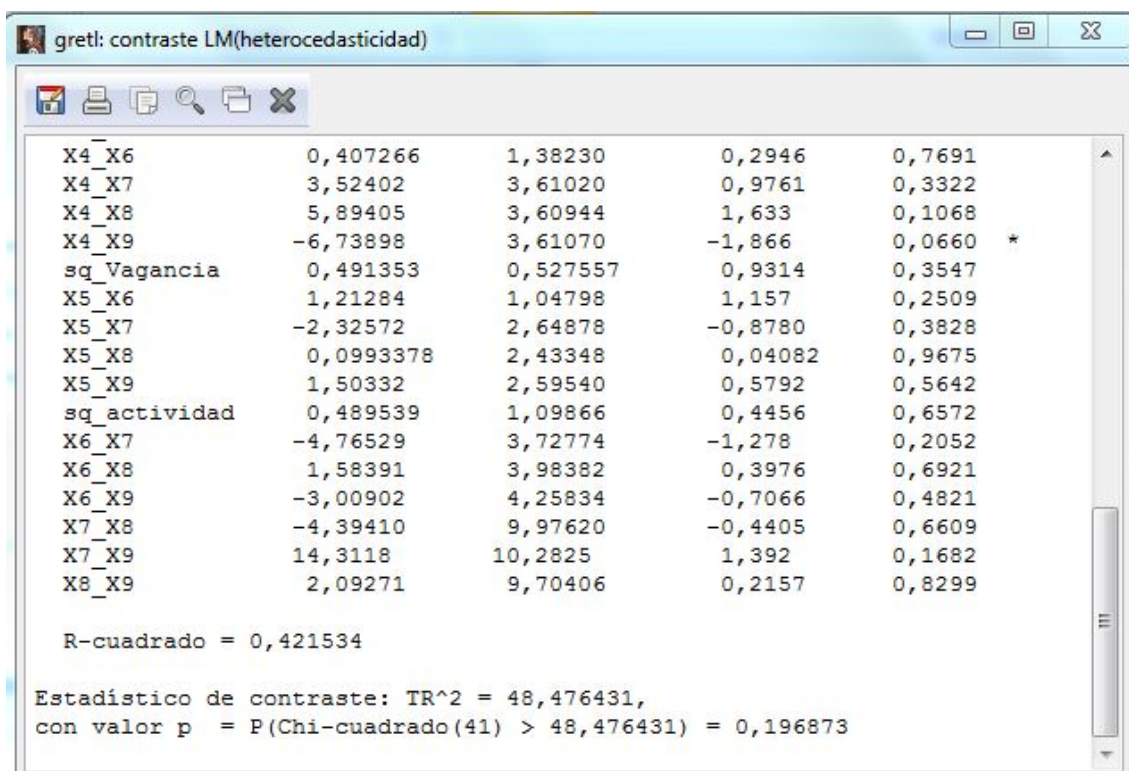
Método grafico

En un principio decidimos estudiar la heterocedasticidad mediante la grafica de los residuos y las variables, aunque dado que no conseguíamos información relevante y útil decidimos no mencionarlas.



Como se puede observar, los gráficos no muestran datos relevantes.

Test de White



X4_X6	0,407266	1,38230	0,2946	0,7691
X4_X7	3,52402	3,61020	0,9761	0,3322
X4_X8	5,89405	3,60944	1,633	0,1068
X4_X9	-6,73898	3,61070	-1,866	0,0660 *
sq_Vagancia	0,491353	0,527557	0,9314	0,3547
X5_X6	1,21284	1,04798	1,157	0,2509
X5_X7	-2,32572	2,64878	-0,8780	0,3828
X5_X8	0,0993378	2,43348	0,04082	0,9675
X5_X9	1,50332	2,59540	0,5792	0,5642
sq_actividad	0,489539	1,09866	0,4456	0,6572
X6_X7	-4,76529	3,72774	-1,278	0,2052
X6_X8	1,58391	3,98382	0,3976	0,6921
X6_X9	-3,00902	4,25834	-0,7066	0,4821
X7_X8	-4,39410	9,97620	-0,4405	0,6609
X7_X9	14,3118	10,2825	1,392	0,1682
X8_X9	2,09271	9,70406	0,2157	0,8299

R-cuadrado = 0,421534

Estadístico de contraste: $TR^2 = 48,476431$,
con valor p = $P(\text{Chi-cuadrado}(41) > 48,476431) = 0,196873$

Una vez realizado el test de White se puede observar un p valor = 0,196873.

Con este dato podemos confirmar la existencia de Homocedasticidad, es decir, rechazamos la H_0 dado que el P valor es mayor que 0.05.

Test Breusch-Pagan

```
Contraste de heterocedasticidad de White -  
Hipótesis nula: No hay heterocedasticidad  
Estadístico de contraste: LM = 48,4764  
con valor p =  $P(\text{Chi-cuadrado}(41) > 48,4764) = 0,196873$   
  
Contraste de heterocedasticidad de Breusch-Pagan -  
Hipótesis nula: No hay heterocedasticidad  
Estadístico de contraste: LM = 46,3489  
con valor p =  $P(\text{Chi-cuadrado}(8) > 46,3489) = 2,04017\text{e-}007$ 
```

Tras realizar la prueba de Breusch-Pagan, nos da un P valor de 2.04017-007 es decir, según esta prueba no podríamos rechazar H_0 , es decir existiría heterocedasticidad.

4.4 Autocorrelación

Nuestra base de datos no está compuesta por datos de corte transversal, es decir, los datos cogidos de una muestra aleatoria y una sola vez, por ello no podría existir la presencia de auto correlación, por ese motivo no llevamos a cabo ninguna prueba de contraste de Durbin-Watson.

5. Conclusiones

En cuanto a la variable sobre el nivel de holgazanería, sorprendentemente, el parámetro estimado que acompaña a dicha variable ha sido mucho menor de lo esperado: 0,268246. Esto puede ser por el carácter subjetivo del grado de pereza. Igualmente otras variables: tener pareja no da resultado, pudiendo deducir que puede darse el caso de la hipótesis, pero que más bien depende del individuo. O el mito de las mujeres también cae.

Por otro lado, retractándonos en lo dicho sobre la variable “nivel de vagancia”, podemos decir que el tiempo cada individuo lo usa para lo que lo quiere, cada uno con expectativas diferentes. Pudiendo, mejorar las notas a través del aumento de horas de estudio, o de la importancia dada (individualmente) a cada trabajo y al día a día en el transcurso de las clases. Por ello, ponte a estudiar, apaga la tele, cierra el whatsapp y ponte a hincar los codos.

6. Bibliografía

<http://trabajoseconometria.blogspot.es/1339542660/por-qu-tardamos-tanto-en-empezar-los-trabajos-/>