

**Taller # 8
Econometría 06216**

Profesores: Julio César Alonso C. – Ana Isabel Gallego

Monitora: Valentina Gatti – Ángela Bermúdez

Notas:

- Recuerde que únicamente tres preguntas, seleccionadas al azar, serán calificadas.
- Este taller deberá subirse a la plataforma Moodle hasta las 7:10 del martes 23 de marzo de 2010. **Sólo se calificaran talleres en formato pdf. Cualquier otro formato no será tenido en cuenta.**

INSTRUCCIONES:

- Este taller debe ser escrito en computador.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.
- Este taller es un trabajo en pareja. Por tanto el taller debe reflejar **únicamente** el trabajo de la pareja.
- Si bien no es necesario reportar todos los números decimales, sí lo es hacer los cálculos con **todos** ellos.

La empresa petrolera LORTEPOCE desea elaborar un modelo de regresión lineal múltiple para analizar la demanda del barril de gasolina. Para realizar el análisis el presidente de la empresa le pide al jefe de estudios económicos que construya la información necesaria y que estime el modelo.

A su vez, el jefe de estudios económicos le solicita a su asistente que construya la base de datos para realizar la estimación. El asistente empieza a construir las variables necesarias para el estudio recordado los conceptos básicos de su curso de microeconomía, y recuerda vagamente que la demanda de gasolina es una función del precio del bien, el precio de los bienes sustitutos, el precio de los bienes complementarios, el ingreso y los gustos o preferencias, entre otras variables.

La primera variable que construye se la hace llegar el jefe de producción de la empresa: Y_i es el número de barriles de gasolina.

Las dos siguientes variables que construye se las hace llegar el jefe de mercadeo de la empresa: $P1_i$ es el precio por litro de gasolina (en miles de pesos) y $P2_i$ es el precio por litro del gas (en miles de pesos).

La tercera y última variable que consigue construir es I_i el ingreso de los consumidores de gasolina en pesos del 2000. Esta variable se la hizo llegar el jefe de recursos humanos de la empresa, quien advierte que en la construcción del ingreso existe un error de cómputo al digitar la encuesta.

Finalmente, la base de datos es un corte transversal y se encuentra disponible en el archivo T8-01-10.xls (**Hoja1**).

1. Empleando la información anterior, responda las siguientes preguntas:

- a) Escriba el modelo que debería estimar el jefe de estudios económicos.
- b) Interprete a priori los coeficientes del modelo propuesto por el director del estudio y determine cuál es el signo esperado.
- c) Estime el modelo propuesto por el director del estudio y reporte los resultados en una tabla.

2. Continuando con el ejercicio anterior.

- a) Determine si existen síntomas (sospechas) de algún problema econométrico. Justifique.
- b) Confirme (formalmente) la existencia del problema econométrico.

3. En caso de haber encontrado formalmente un problema econométrico corríjalo (si es posible) por los dos métodos vistos en clase MCP y White. Reporte los resultados del modelo corregido en la tabla 1.

Un candidato a la presidencia de la república del triangulo tienen como uno de sus ejes de actuación mejorar las condiciones de los trabajadores en el mercado de trabajo. Para lo cual está interesado en una ecuación sobre la determinación del salario en su país. Para estimar la ecuación de determinación de salarios cuenta con la información disponible en el archivo T8-01-10.xls (Hoja 2).

4. Teniendo en cuenta el enunciado anterior y los datos del archivo T8-01-10.xls (**Hoja2**):

- a).Escriba la ecuación de salarios que debe estimar teniendo en cuenta el enunciado y **los datos suministrados**.
- b).Estime el modelo y reporte los resultados en una tabla (Tabla 2).

5. Continuando con el enunciado anterior

- a). Efectúe el análisis gráfico de los errores estimados. ¿Qué tipo de problema puede intuir a partir de este análisis? Explique.
- b) Realice las pruebas que considere necesarias para determinar la existencia o no de un problema econométrico en el modelo.

6. Continuando con el ejercicio anterior:

- a).De haber encontrado un problema econométrico, corrija por el método que considere apropiado. Muestre como efectuó su corrección y reporte el modelo estimado en la misma tabla empleada anteriormente (Tabla 2).

Taller # 8
Respuestas sugeridas
Economía 06216

Profesor: Julio Cesar Alonso - Ana Isabel Gállego
Maitor: Ángel María Bernabé - Valentina Gatti

Notas:

- Recuerde que únicamente tres preguntas, seleccionadas a la azar, serán calificadas.
- Este taller deberá subirse a la plataforma Moodle hasta las 7:10 de mañana del 23 de marzo de 2010. **Sólo se calificaran talleres en formato pdf. Cualquier otro formato no será tenido en cuenta.**

INSTRUCCIONES

- Este taller debe ser escrito en computador.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.
- Este taller es un trabajo en pareja. Por tanto el taller debe referirse únicamente al trabajo de la pareja.
- Si bien no es necesario reportar todos los números decimales, sí debe hacer los cálculos con todos los decimales.

La empresa petrolera LO FIED desea elaborar un modelo de regresión lineal múltiple para analizar la demanda de barril de gasolina. Para realizar el análisis el presidente de la empresa pide al jefe de estudios económicos que construya la información necesaria y que estimen el modelo.

A su vez, el jefe de estudios económicos solicita a su asistente que construya la base de datos para realizar la estimación. El asistente en consecuencia construye las variables necesarias para el estudio recordando los conceptos básicos de su curso de microeconomía y recuerda vagamente que la demanda de gasolina es una función de precio de gasolina, precio de bienes sustitutos, precio de bienes complementarios, el ingreso y las gustos o preferencias, entre otras variables.

La primera variable que construye se la hace llegar el jefe de producción de la empresa Y_i es el número de barriles de gasolina.

Las dos siguientes variables que construye se las hace llegar el jefe de mercadeo de la empresa: $P1_i$ es el precio por litro de gasolina (en miles de pesos) y $P2_i$ es el precio por litro del gas (en miles de pesos).

La tercera y última variable que consigue construir es I_i el ingreso de los consumidores de gasolina en pesos de 2000. Esta variable se la hizo llegar el jefe de recursos humanos de la empresa quien advierte que en la construcción del ingreso existe un error de código al digitar la encuesta.

Finalmente la base de datos es un corte transversal y se encuentra disponible en el archivo T8-01-10.xls (**Hojas**).

1. En base a la información anterior, responda las siguientes preguntas:

- a) Escriba el modelo que debería estimar el jefe de estudios económicos.

Nota: que la variable ingreso tiene un error de medición y que por tanto no se debe incluir en el modelo. Si la variable incluye los errores serán sesgados e inconsistentes. Por tanto se debe excluir la variable ingreso del modelo.

Ahora bien, se supone que la demanda de gasolina por barriles es una función de precio por litro de gasolina en miles de pesos y de precio por litro de gas en miles de pesos. El director del estudio escribió el siguiente modelo:

$$Y_i = \alpha_1 + \alpha_2 P1_i + \alpha_3 P2_i + \mu_i$$

Donde, Y_i es el número de barriles de gasolina, $P1_i$ es el precio por litro de gasolina (en miles de pesos), $P2_i$ es el precio por litro del gas (en miles de pesos) y μ_i es el respectivo término de error.

- b) Interprete a priori los coeficientes del modelo propuesto por el director del estudio y determine cuáles el signo esperado.

α_1 Se espera que el intercepto sea positivo pues indica la cantidad demandada de gasolina cuando el precio de la gasolina y su sustituto es cero.

α_2 Se espera signo negativo para $P1_i$ porque indica el efecto en el que incrementa la cantidad demandada de gasolina ante incrementos de 1000 pesos en el precio por litro de ésta.

α_3 Se espera signo positivo en $P2_i$ porque la gasolina y el gas son bienes sustitutos. α_3 indica el efecto en el que se incrementa la demanda de gasolina ante incrementos de mil pesos en el precio del gas.

- c) Estime el modelo propuesto por el director del estudio y reporte los resultados en una tabla.

Tabla 1. Estimaciones del modelo 1 y corrección (white).

	Modelo 1	Corrección White
	Variable Dependiente: Y	Variable Dependiente: Y
	Estadístico entre paréntesis	Estadístico entre paréntesis
	M.C	M.C
Constante	39.9987 * (1.71)	39.9987 * (1.86)
P1	-0.1964 *** (-12.41)	-0.1964 *** (-9.19)
P2	0.2981 *** (9.30)	0.2981 *** (7.66)

R ²	0.8184	0.8184
R ² - Ajustado	0.8107	0.8107
F	105.94 ***	98.80 ***
Nº de Observaciones	50	50

* Nivel de significancia 90%
 ** Nivel de significancia 95%
 *** Nivel de significancia 99%
 M O Mínimo Cuadrados Ordinarios

2. Continuando con el ejercicio anterior.

- a) Determinar si existen intenciones (sospechas) de algún problema económico. Justifique.

Noten que los datos son de corte transversal y por lo tanto pueden sospechar la presencia de un problema de heteroscedasticidad. Además pueden hacer algunas pruebas informales a través del análisis de los siguientes gráficos:

Gráfico 1. Residuos vs (n) muestra

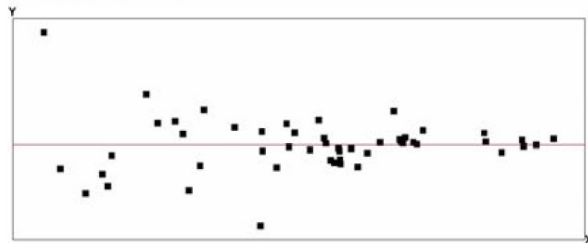


Gráfico 2. Residuos vs P1.

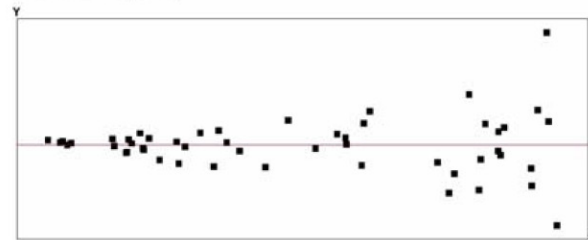
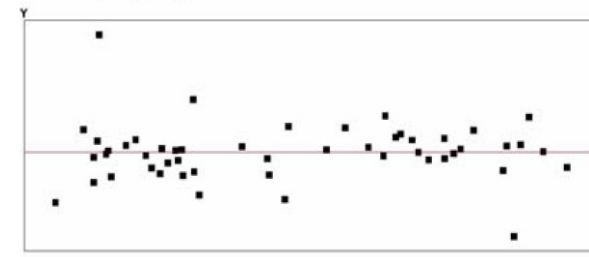


Gráfico 3. Residuos vs P2.



Pueden sospechar de la existencia de un problema de heteroscedasticidad relacionado con la variable independiente P1, precio de la gasolina (Gráfico 2).

- b) Confirmar (formalmente) la existencia del problema económico.

Noten que deben realizar las pruebas formales para detectar la presencia de un problema de heteroscedasticidad en el modelo.

i. Prueba de Goldfeld-Quandt

La hipótesis nula es la presencia de homocedasticidad frente a la hipótesis alternativa de heteroscedasticidad

$$H_0: \sigma_i^2 = \sigma^2$$

$$H_a: \sigma_i^2 = \sigma^2 P1_i^2$$

Para realizar esta prueba se deben seguir los siguientes pasos:

- Ordenar la variable o un conjunto de variables, para la ordenación de las observaciones (P1).
- Determinar el número de datos que se van a eliminar y en este caso, el número de observaciones adecuado para retirar es 8. Por lo tanto dividimos los datos restantes en dos grupos de 21 datos cada uno y se corre una regresión para cada grupo.
- Calcular el estadístico GQ, teniendo en cuenta los errores obtenidos en cada una de las regresiones de los puntos anteriores.
- Finalmente la decisión se toma así el F crítico es menor al estadístico GQ.

	P1
SSE1	0.00366934
SSE2	0.09049835
F(GQ)	24.66338633
F _{(60-8-2)(3), (50-8-6) α=0.01}	2.205018
Decisión	Rechazo al %

Al 1% de significancia la prueba muestra la presencia de un problema de heteroscedasticidad de la forma $\sigma^2 P1_i^2$

ii. Prueba de Breusch-Pagan

En esta prueba se consideran la hipótesis de una relación entre la varianza de error y un grupo de variables con o sin evidencia de heteroscedasticidad, versus la hipótesis nula de la no existencia de esta relación con o sin evidencia de heteroscedasticidad. Es decir,

$$H_0: \sigma_i^2 = \sigma^2$$

$$H_a: \sigma_i^2 = \sigma^2 F(\gamma - \delta Z_i)$$

Al realizar todo el procedimiento el resultado obtenido para el BP es 38.47, con un p-valor de 0.0000, lo que da como resultado que se puede rechazar la hipótesis nula con un 99% de confianza. Sin embargo, vale la pena ver si con la variable P1 en particular se presenta el problema de heteroscedasticidad, los resultados se muestran a continuación.

La regresión original da como resultado que el SSE es 0.09910971 por lo tanto $\sigma^2 = 0.0019821942$. Ahora corremos la regresión para P1. Tenemos que BP=33.2544303 que comparado con el estadístico chi cuadrado con un grado de libertad (6.63), nos permite rechazar con un 99% de confianza la nula de heteroscedasticidad.

iii. Prueba de White

Este se puede considerar como el más general de todos. De manera similar, tenemos que en esta prueba se busca contrastar la hipótesis nula de heteroscedasticidad versus la hipótesis alternativa de heteroscedasticidad. Por tanto, sus planteamientos son:

$$\sigma_i^2 = \sigma^2$$

$$H_a: \sigma_i^2 = \sigma^2 P1_i^2$$

Después de realizar la regresión el W_α debe ser igual a 17.58 que comparado con los niveles de significancia al 90%, 95% y 99%. Permite rechazar la hipótesis nula al 99%.

$$x_{4,0.1}^2 = 7.77944$$

$$x_{4,0.05}^2 = 9.48$$

$$x_{4,0.01}^2 = 13.2767$$

De las tres pruebas anteriores se puede concluir que existe un problema de heteroscedasticidad causado por la variable P1.

3. En caso de haber encontrado formalmente un problema económico crítico corríjalo (si es posible) por los dos métodos vistos en clase MCP y White. Reporte los resultados de manera corregida en la tabla 1.

i. Corrección por Mínimos Cuadrados Ponderados MCP

Noten que este método también puede usarse debido a que conocen la naturaleza de la heteroscedasticidad.

$$\frac{Y_i}{P1_i} = \frac{\beta_0}{P1_i} + \beta_1 + \beta_2 \frac{P2_i}{P1_i} + \frac{\mu_i}{P1_i}$$

$$VAR(\mu_i) = \sigma^2 P1_i^2$$

Entonces,

$$VAR\left(\frac{\mu_i}{P1_i}\right) = \frac{1}{P1_i^2} VAR(\mu_i) = \frac{1}{P1_i^2} \sigma^2 P1_i^2 = \sigma^2$$

Para probar si efectivamente se corrigió el problema hacen una inspección de los síntomas y encuentran que el problema parece haberse agravado.

Además, tenemos un estadístico de Breusch-Pagan para el modelo generado de 6.915921 con un p-valor asociado de 0.03, con lo que rechazamos la hipótesis nula de heteroscedasticidad al 95% de confianza.

Al realizar las pruebas de Goldfeld y Q para la variable $\frac{1}{P1_i}$, encontramos que el estadístico es 3.38, superior a 2.20502, por lo que se rechaza la nula de heteroscedasticidad al 99% de confianza.

También realizan la prueba de Breusch-Pagan con la variable $\frac{1}{P1_i}$, y obtienen un estadístico igual a 6.3799, superior a 3.84 pero inferior a 6.63, con lo que se concluye que con esta prueba también hay indicios de heteroscedasticidad al 95% de confianza. Para terminar, se realiza el test de White y se encuentra que el estadístico es $50 * 0.13 = 6.54$ que comparado con los valores críticos (al 90% 7.78), no permite rechazar la nula de heteroscedasticidad.

Encuentran entonces que el problema no se corrigió por el método MCP.

ii. Corrección de White

Los resultados se presentan en la tabla 1.

Un candidato a la presidencia de la república de Trángulo tienen como uno de sus ejes de actuación mejorar las condiciones de los trabajadores en el mercado de trabajo. Para lo cual está interesado en una ecuación sobre la detención de la salariedad en su país. Para este fin la ecuación de detención de salarios cuenta con la información disponible en el archivo T8-01-10.xls (Hoja 2).

4. Teniendo en cuenta el enunciado anterior y los datos del archivo T8-01-10.xls (Hoja 2):

- a). Escriba la ecuación de salarios que debe estimar teniendo en cuenta el enunciado y los datos suministrados.

En la base de datos suministrada en el archivo T8-01-10.xls (Hoja 2), disponen de las siguientes variables:

- La variable dependiente: Son los sueldos o salarios por empleado.
- Las variables independientes son:

PF es el precio del producto final a costo de factores
 U desempleo del país
 M Nivel de importaciones.

La variable M Nivel de importaciones no se incluye en el modelo en la estimación por dos razones, primero porque es una variable irrelevante y segundo porque tiene un error de medición al no estar toda la serie completa en el archivo suministrado.

Por lo tanto, la especificación correcta del modelo sería:

$$W_t = \alpha_0 + \alpha_1 PF_t + \alpha_2 U_t + \varepsilon_t$$

Donde ε_t es el respectivo término de error.

b). Estimé el modelo y reporte los resultados en una tabla (Tabla 2).

Tabla 2 Estimaciones de modelo y corrección por MG.

	Modelo 1	Corrección MG
Variable Dependiente: W	Variable Dependiente: W	Variable Dependiente: W
Estadístico t entre paréntesis	Estadístico t entre paréntesis	Estadístico t entre paréntesis
M.O	M.O	M.O
Constante	-2.4781 (-0.213)	-2.4781 (N/A)
PF	0.1841 *** (44.06)	-0.1938 *** (19.91)
U	-0.0194 *** (-0.32)	0.2981 *** (-0.959)
R ²	0.9986	0.9939
R ² -Ajustado	0.9985	0.9936
DW	0.9173	1.5192
Nº de Observaciones	36	35

* Nivel de significancia 90%

** Nivel de significancia 95%

*** Nivel de significancia 99%

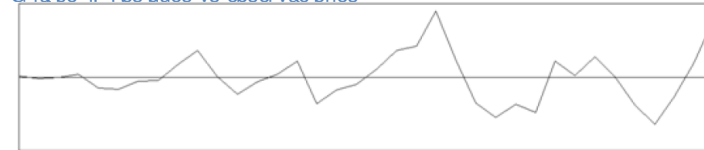
M.O Mínimos Cuadrados Ordinarios

5. Continuando con el enunciado anterior:

a). Efectúe el análisis gráfico de los errores estimados. ¿Qué tipo de problema puede intuirse a partir de este análisis? Explique.

Puede ser posible que exista un problema de autocorrelación debido a que los datos con los que estamos trabajando son una serie de tiempo. Para el análisis intuitivo realizamos el gráfico de los residuos frente a la muestra.

Gráfico 1. Residuos vs observaciones



Según lo observado en el gráfico 1, parece existir un problema de autocorrelación pues los errores tienden a seguir un comportamiento definido. Particularmente, puede ser autocorrelación positiva pues el signo de los residuos tiende a mantenerse el mismo por períodos.

b) Realice las pruebas que considere necesarias para determinar la existencia o no de un problema económico en el modelo.

i. Prueba de Durbin Watson

La hipótesis nula que sostiene esta prueba es la de no autocorrelación de primer orden, versus la hipótesis alternativa de autocorrelación de primer orden.

$$H_0: \rho = 0$$

$$H_0: \rho \neq 0$$

Primero podemos obtener el DW que arroja EasyReg, el cual da un valor de 0,9173 permitiendo intuir la existencia de autocorrelación positiva dado que este número es menor que 2.

Después vamos a usar el estadístico DW con el intervalo $d_u < DW < 4 - d_u$. En este caso, a un nivel de significancia del 1%, los valores críticos son $d_l = 1.153$ y $d_u = 1.376$. Primero se debe analizar si el estadístico no cae en las zonas de indeterminación, pues no tendrían sentido la prueba, es decir, $d_l < DW < d_u$ y $4 - d_u < DW < 4 - d_l$.

Dado que el valor calculado está por fuera de estos intervalos, la prueba aplica, adicionalmente al compararlo con el intervalo $d_u < DW < 4 - d_u$, el estadístico calculado está por debajo, entonces rechazamos la hipótesis nula de no autocorrelación de primer orden, es decir, aceptamos la existencia de este problema.

Por otra parte, para saber el signo de la autocorrelación, podemos contrastar la hipótesis alternativa de no autocorrelación positiva versus la hipótesis alternativa de autocorrelación positiva. Para ello, es necesario comparar el DW dentro de los siguientes intervalos, para este caso sería $0 < DW < 1.153$, intervalo en el que se encuentra el estadístico calculado, por lo que rechazamos la hipótesis nula de no autocorrelación positiva, es decir, aceptamos la existencia de un problema de autocorrelación positiva.

ii. Prueba de Pachs

La hipótesis nula que sostiene esta prueba es la de no autocorrelación de primer orden, versus la hipótesis alternativa de autocorrelación de primer orden.

$$H_0: \rho = 0$$

$$H_0: \rho \neq 0$$

Tenemos la siguiente información:

$$K = 11$$

$$N_+ = 20$$

$$N_- = 16$$

Debemos obtener los valores de:

$$E[k] = 18.78$$

$$var[k] = 8.52$$

Cambiamos el estadístico RA

$$RA = \frac{k - E[k]}{\sqrt{var[k]}} = -2.6643$$

Si deseamos contrastar a un nivel de significancia del 5% y el 1%, tendríamos los siguientes valores 1.96 y 2.575 respectivamente, por lo que podemos rechazar la hipótesis nula con una confianza del 99%. Es decir, que hay un problema de autocorrelación.

iii Prueba de Box-Pierce

La hipótesis nula es la no existencia de autocorrelación, versus la hipótesis alternativa de la existencia de la misma. De esta manera, los diez primeros rezagos que se muestran en el Cuadro 1, nos permiten rechazar la hipótesis nula, con lo cual se concluye la existencia de este problema dentro del modelo.

Cuadro 1. Prueba de Box-Pierce

Q(p)	Valor	p-valor	t-crítico 5% significancia	Decision
1	7,59	0,00588	3,84	rechazar
2	7,64	0,02188	5,99	rechazar
3	14,66	0,00213	7,81	rechazar
4	21,44	0,00026	9,49	rechazar
5	23,75	0,00024	11,07	rechazar
6	23,77	0,00058	12,59	rechazar
7	24,88	0,0008	14,07	rechazar
8	27,82	0,00051	15,51	rechazar
9	27,84	0,00102	16,92	rechazar
10	28,86	0,00131	18,31	rechazar

6. Continuando con el ejercicio anterior:

- De haber encontrado un problema econométrico, corríjalo por el método que considere apropiado. Muestre con efectividad su corrección y reporte el modelo estimado en la misma tabla en la página anterior (Tabla 2).

De acuerdo con las pruebas realizadas y teniendo en cuenta el análisis gráfico, se llega a la conclusión de que existe un problema de autocorrelación positiva en este modelo. Dado que la prueba de Durbin-Watson nos permitió concluir sobre la existencia de AR1, podemos utilizar el método de los Mínimos Cuadrados Generalizados. Para ello, rezagamos cada una de las variables un período, las multiplicamos por el coeficiente de correlación y se las restamos a las variables no rezagadas, de esta manera se obtiene el modelo (2).

$$W_t^* = \alpha_0^* + \alpha_1^* PF_t^* + \alpha_2^* U_t^* + \varepsilon_t^*$$

En donde,

$$\varepsilon_t^* = \varepsilon_t + \rho\varepsilon_{t-1}; \quad W_t^* = W_t + \rho W_{t-1}; \quad \alpha_0^* = \alpha_0(1 - \rho) \quad ; \quad PF_t^* = PF_t + \rho PF_{t-1} \quad \text{y}$$

$$U_t^* = U_t + \rho U_{t-1}$$

Primero se encontró que $\rho = 0.6257$, el cual resulta después de estimar el modelo con los rezagos (LAG). Realizados estos pasos del método de diferencias generalizadas (MG), los errores ya no presentarían el problema de autocorrelación, los resultados se reportan en la Tabla 2.