



ECONOMETRIA UNPLUGGED



TITULO

ECONOMETRÍA UNPLUGGED (para principiantes)

AUTOR

Economista Luis Damian Jungnikel Matamoros¹

Correo. damianjungnikel@hotmail.com

¹ Profesor de matemáticas y base de datos en la Unidad Educativa CPE en Bachillerato. Anteriormente ayudante de investigación en el Instituto de Investigaciones Económicas y Políticas de la Universidad de Guayaquil.

RESUMEN

El uso de las tecnologías en el aula, y más aún en la Universidad, ha automatizado casi todos los procesos rigurosos e inclusive tediosos de matemática, estadística y en este caso de la econometría. No obstante, es recomendable que el estudiante de economía comprenda cada proceso que descansa en los resultados del software econométricos.

Este manual (no lo podemos considerar como libro) o formulario está pensado para ayudar al estudiante, en ausencia de software econométrico, o para ejercitar sus conocimientos de matemática, estadística y economía. Su nombre, UNPLUGGED deriva de desconcertado, o desenchufado, haciendo referencia a la capacidad de resolver problemas econométricos aún en ausencia del software especializado.

ABSTRACT

The use of technology in the classroom, and even more in the University, has automated most rigorous and even tedious mathematics, statistics and econometrics this case processes. However, it is recommended that the economics student understands each process that rests on the results of econometric software.

This manual (which cannot be considered as a book complete) or form is designed to help students in the absence of econometric software, or to exercise their knowledge of mathematics, statistics and economics. Its name UNPLUGGED its referring to the ability to solve econometric problems even in the absence of specialized software.

PALABRAS CLAVE

- Econometría
- Formulario econométrico
- Econometría introductoria
- Estadística Univeristaria
- Modelación matemática
- Economía Matemática
- Estadística económica.

CONTENIDOS

El siguiente trabajo extrae las fórmulas más utilizadas en los libros de econometría introductoria, y pueden ser un auxilio al estudiante en ausencia de programas econométricos. Sin embargo, advertimos que esta recopilación no pretende ser un libro de estudio, ni de texto, ya que, solo es una referencia rápida de fórmulas, para el alumno o el maestro.

Este texto consta un resumen de las principales fórmulas de la materia, pero esconde un afán simplista, y por tanto, aborda de manera superflua el contenido esencial que el estudiante debe revisar en la bibliografía usada y/o la indicada por sus maestros.

En la segunda parte (que saldrá próximamente) se utiliza los datos disponibles de la economía ecuatoriana, de más de cinco décadas, que ha sido recopilada de la principal fuente de información el Banco Central del Ecuador (BCE), y secundariamente, el Banco Mundial y la Comisión Económica para América Latina y el Caribe (CEPAL). Con esa información se contemplan varios modelos econométricos muy sencillos, que buscan que el estudiante se familiarice solo de forma introductoria a la práctica econométrica.

OBJETIVO.

El objetivo de este trabajo, es brindar al estudiante de Ciencias Económicas un breve resumen de la materia de Econometría. Hemos tratado de ser lo más práctico posible y no confundir con los procedimientos y fórmulas (y de hecho, casi siempre le pasa al estudiante inicial) con el análisis de hipótesis, agregando reglas muy sencillas.

Claro está, que en Econometría no es tan fácil dar un manual de que hacer en cada caso, ya que es una materia muy extensa, suponiendo entonces, no solo el estudio de la ciencia, sino también un poco de imaginación, o arte. Por lo tanto, este libro pretende ser introductorio, y no abarcará ningún aspecto teórico profundo, o complicado.

El estudiante puede imprimir el contenido de la primera parte, para tener a la mano las fórmulas y procedimientos para ejercicios econométricos UNPLUGGED.

DECALOGO DEL ECONOMETRISTA.

Por favor lea este decálogo antes de introducirse en el catálogo de fórmulas econométricas: Sigüientes "diez mandamientos de la econometría aplicada que extraemos del afamado libro de Gujarati (Gujarati & Porter, 2009).

1. Utilizarás la teoría económica aplicándola con sentido común.
2. El relevante es superior a la elegancia matemática.
3. No harás análisis econométrico ignorante y alejado del contexto.
4. Revisa la normalidad de los datos.
5. Utiliza la simplicidad en todo, más aun en lo probabilístico...
6. Se perseverante en los resultados que has encontrado.
- 7 Reduce los costos de investigas y recopilar datos.
8. La realidad supera la ficción de los textos de teoría.
9. la significancia práctica es superior a la significancia estadística.
10. Acepta las críticas, y defiende tus trabajos.

Es original de (Kennedy)

Nota acerca de la simbología

Debe notarse que en todas las ecuaciones de este trabajo:

- *El símbolo (Σ) denota sumatoria.
- *Los símbolos con énfasis (Λ) son estimados.
- *En donde los símbolos con énfasis ($\bar{}$) son medios o promedios.
- *Las letras mayúsculas corresponden a valores absolutos del muestreo.
- *Las letras minúsculas corresponden a desviaciones con respecto a la media.
- *Generalmente se denomina "Y" a la variable dependiente, endógena o explicada.
- *Se designa a X para indicar la variable o las variables explicativas, *exógenas o explicativas.

NO DEBERÁ TOMAR ESTE FORMULARIO COMO LIBRO DE TEXTO; YA QUE NO SE ENCONTRARÁN TEMAS O EXPLICACIONES AVANZADAS, PARA CUALQUIER CONSULTA REFIERASE A LA BIBLIOGRAFIA USADA.

INDICE TEMATICO

Contenido	
TITULO	i
ECONOMETRÍA UNPLUGGED (para principiantes)	i
AUTOR	i
RESUMEN	ii
ABSTRACT	ii
PALABRAS CLAVE	ii
CONTENIDOS	iii
OBJETIVO	iii
DECALOGO DEL ECONOMETRISTA	iv
Nota acerca de la simbología	iv
INDICE TEMATICO	v
INDICE DE ECUACIONES	vi
INDICE DE ILUSTRACIONES	¡Error! Marcador no definido.
PRIMERA PARTE:	1
FÓRMULAS PARA EJERCICIOS UNPLUGGED	1
A. Introducción a la econometría: repaso estadístico	2
A.1. Estadística descriptiva	2
A.2. Estadística inferencial	4
Breve Árbol de decisiones para resolver problemas estadísticos de probabilidad	6
Capítulo 1	7
1. Escogiendo un modelo y del tipo de regresión	7
1.1. La regresión	7
1.2. Supuestos de los mínimos cuadrados ordinarios	8
1.3. Modelos lineales	8
1.4. Ecuaciones normales para realizar una regresión	8
1.5. Modelos no lineales	9
1.6. Errores de las perturbaciones	10
1.7. Ecuaciones para hallar los parámetros	10
1.9. Varianza de errores y parámetros	11
1.10. Regresión a través del origen	12

1.11.	El uso de variables ficticias.....	13
1.12.	Modelos con variables dicótomas.....	13
1.13.	Regresión lineal por secciones.....	14
1.14.	Efecto multiplicativo de interacción.....	14
1.15.	Regresión semi logarítmicas.....	15
1.16.	Regresión con variable dicótoma endógena de decisión o ficticia	15
1.17.	El uso de ecuaciones simultáneas.....	15
1.18.	Método MCI.....	16
1.1.6.	Método de MC2E	16
Capítulo 2.....		18
2.	Pruebas básicas al realizar un modelo econométrico.....	18
2.1.	Prueba de significación de los parámetros.....	18
2.2.	La bondad del ajuste.....	18
2.3.	La prueba Global.....	20
2.4.	Correlación.....	20
2.5.	Multi-colinealidad o micro-numerosidad.....	21
2.6.	Prueba de Multi-colinealidad	21
2.7.	Heterocedasticidad.....	22
2.8.	Prueba de Heterocedasticidad.....	22
2.9.	Auto correlación.....	24
2.10.	Prueba de Auto correlación.....	24
2.11.	Modelo rezagado para corregir auto correlación.....	25
Capítulo 3.....		27
3.	Prediciendo el futuro.....	27
3.1.	Intervalos de confianza.....	27
3.2.	Prueba de Intervalos de confianza.....	27
Capítulo 4.....		30
4.	Pruebas definitivas para escoger el modelo.....	30
4.1.	Pruebas de la estructura del modelo	30
4.1.1.	Prueba de simultaneidad o de Hausman.....	30
4.1.2.	Prueba de exogeneidad.....	30
4.1.3.	Prueba de estabilidad estructural o CHOW	30
4.1.4.	Prueba MWD de la forma funcional del modelo.....	31
4.1.5.	Prueba reset de Ramsey.....	32
4.2.	Advertencia y conclusión.....	33
Bibliografía		33

INDICE DE ECUACIONES

Ecuación 1: Media aritmética.....	2
Ecuación 2: Desviaciones.....	2

Ecuación 3: Tabla de momentos	2
Ecuación 4: Media ponderada.	2
Ecuación 5: Media armónica.	3
Ecuación 6: Mediana.	3
Ecuación 7: Recorrido.	3
Ecuación 8: Desviación absoluta.	3
Ecuación 9: Varianza.	3
Ecuación 10: Desviación estándar.	3
Ecuación 11: Coeficiente de Pearson.	3
Ecuación 12: Coeficiente de asimetría de Fisher.	3
Ecuación 13: Coeficiente de curtosis.	3
Ecuación 14: Proporción muestral.	4
Ecuación 15: Característica.	4
Ecuación 16: Reglas básicas de probabilidad.	4
Ecuación 17: Teorema de Bayes.	4
Ecuación 18: Conteos.	4
Ecuación 19: Varianza probabilística.	5
Ecuación 20: Distribución binomial.	5
Ecuación 21: Distribución híper-geométrica.	5
Ecuación 22: Modelo de regresión lineal.	7
Ecuación 23: Interpretación del modelo lineal.	8
Ecuación 24: Reemplazo de variables por sumatorias.	8
Ecuación 25: Ecuaciones normales de la regresión lineal.	8
Ecuación 26: Regresión exponencial linealizada en modelo log-lin.	9
Ecuación 27: Regresión potencial linealizada modelo log-log.	9
Ecuación 28: Modelo potencial, Cobb-Douglas.	9
Ecuación 29: Regresión recíproca linealizada en inversa.	10
Ecuación 30: Errores de las perturbaciones. Una variable explicativa (una variable exógena o simple)	10
Ecuación 31: Encontrar los parámetros lineales de ecuación simple	10
Ecuación 32: Encontrar los parámetros lineales de ecuación múltiple.	10
Ecuación 33: Modelo para hallar perturbaciones.	11
Ecuación 34: Varianza y covarianza simple	11
Ecuación 35: Modelo para encontrar perturbaciones múltiples.	12
Ecuación 36: Varianza y covarianza múltiple.	12
Ecuación 37: Regresión con constante igual a cero.	12
Ecuación 38: Modelo con dicótomas en la constante.	13
Ecuación 39: Modelo con dicótoma en la pendiente.	13
Ecuación 40: Modelo con dicótoma en constante y pendiente.	14
Ecuación 41: Modelo de Umbral.	14
Ecuación 42: Modelo de interacción.	15
Ecuación 43: Modelo semi logarítmico	15
Ecuación 44: Modelo Potencial y con dicótoma.	15
Ecuación 45: Modelo de ecuaciones simultáneas	16
Ecuación 46: Valor Z o T calculado.	18
Ecuación 47: El problema de bondad de ajuste.	18
Ecuación 48: La bondad del ajuste Simple, desde la tabla ANOVA	19
Ecuación 49: La bondad del ajuste. Varias fórmulas.	19
Ecuación 50: La prueba global o Valor F	20
Ecuación 51: Ejemplo de modelo para corregir Multi-colinealidad.	21
Ecuación 52: valor FIV	22
Ecuación 53: Valor TOL.	22
Ecuación 54: Prueba de Park.	22
Ecuación 55: La razón en la prueba Golden Quant.	22
Ecuación 56: Valor Pi	23
Ecuación 57: Valor SEC.	23
Ecuación 58: Modelo para estimar Chi en la prueba White.	23
Ecuación 59: Valor d.	24
Ecuación 60: Valor p aproximado.	25
Ecuación 61: Modelo rezagado.	25

Ecuación 62: Calculo del primer dato rezagado.....	25
Ecuación 63: Regresion ajustada para disminuir auto correlación.....	25
Ecuación 64: Rangos de predicción.....	27
Ecuación 65: Rango de predicción de Chií	28
Ecuación 66: Predicción media.....	28
Ecuación 67: Varianza de predicción de Y.....	28
Ecuación 68: Rango de predicción individual.	28
Ecuación 69: Rango de predicción por algebra matricial.....	28
Ecuación 70: Prueba de exogeneidad.	30
Ecuación 71: Calculo de los grados de libertad, prueba CHOW	31
Ecuación 72: Suma de errores, prueba CHOW	31
Ecuación 73: Calculo de F o global en la prueba CHOW	31
Ecuación 74: T calculado en la prueba MWD.....	31
Ecuación 75: Modelo para prueba MWD.....	31
Ecuación 76: Calculo de F en la prueba RESET.	32

PRIMERA PARTE:

FÓRMULAS PARA EJERCICIOS UNPLUGGED.

A. Introducción a la econometría: repaso estadístico

A.1. Estadística descriptiva.

Estas fórmulas fueron extraídas de varios autores y principalmente de (Ruiz Muñoz., 2004) y que recomendamos revisarlo previamente a la revisión de este texto. Además fueron complementadas por otros autores que son inevitables recomendar: (Webster, 1998), (Mason & D, 2000).

La mayoría de los estadísticos pueden ser obtenidos directamente en Excel. Deberá seguir los siguientes pasos:

- 1.- Activar análisis de datos desde el menú inicio de Excel, seleccionar opciones, activar herramientas.
- 2.- Seleccionese datos, análisis de datos.
- 3.- Especifique los estadísticos que desea.

Ecuación 1: Media aritmética.

$$\bar{x} = \frac{\sum(x)}{N}$$

Ecuación 2: Desviaciones.

$$\sum(x - \bar{x}) = 0$$

Cuando la distribución se aproxima a la normalidad o momento cero.

Ecuación 3: Tabla de momentos

$$m^0 = \sum(X - \bar{X})^0 = 1 \text{ Donde m es el momento.}$$

$$m^1 = \frac{\sum(x - \bar{x})^1}{N} = 0$$

$$m^2 = \frac{\sum(x - \bar{x})^2}{N} = \text{véase varianza } \sigma^2$$

$$m^3 = \frac{\sum(x - \bar{x})^3}{N} = \text{véase asimetría}$$

$$m^4 = \frac{\sum(x - \bar{x})^4}{N} = \text{véase curtusis}$$

Ecuación 4: Media ponderada.

$$\bar{X} = \frac{\sum x w}{\sum w}$$

En donde w es la ponderación.

Ecuación 5: Media armónica.

$$X = \frac{N}{\sum_x \frac{1}{x} n}$$

Con la condición de que $X \neq 0$.

Ecuación 6: Mediana.

$$Me = \frac{x + \bar{x}}{2}$$

Ecuación 7: Recorrido.

$$Re = Max\ x - Min\ x.$$

Ecuación 8: Desviación absoluta.

$$s = \frac{|x - \bar{x}|}{N}$$

Ecuación 9: Varianza.

$$var = \frac{\sum(x - \bar{x})^2 n}{N} = s^2$$

Ecuación 10: Desviación estándar.

$$s = \sqrt[3]{var}$$

Ecuación 11: Coeficiente de Pearson.

$$CV = \frac{\sigma}{\bar{x}}$$

A menor CV menor dispersión de los datos.

Ecuación 12: Coeficiente de asimetría de Fisher.

$$Fisher = \frac{\frac{\sum(x - \bar{x})^3 n}{N}}{s^3}$$

A valores positivos es asimetría a la derecha, caso contrario a la izquierda.

Ecuación 13: Coeficiente de curtosis.

$$curtuosis = \frac{\frac{\sum(x - \bar{x})^4 n}{N}}{s^4}$$

A valores mayores a 3 es leptocúrtica (o elevada).

Menores a 3 platicúrtica (o achatada).

Igual o cerca a 3 es normal o mesocúrtica.

Ecuación 14: Proporción muestral.

$$P(X) = \frac{X}{n} = \bar{X} \text{ Donde } n \text{ es tamaño de la muestra y } P \text{ es probabilidad.}$$

Ecuación 15: Característica.

$$P(X) = \frac{X}{N} = u \text{ Donde } N \text{ es población.}$$

Ecuación 16: Reglas básicas de probabilidad.

Excluyente.

$$P(A \cup B) = P(A) + P(B)$$

$$P(A) = 1 - P(A')$$

No excluyente.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Independiente.

$$P(A \text{ y } B) = P(A) * P(B)$$

Condiciona independiente.

$$P(A|B) = P(B)$$

Conjunta.

$$P(BA) = P(B|A) * P(A)$$

$$P(B|A) = P(BA)/P(A)$$

Ecuación 17: Teorema de Bayes.

$$P(A|B) = \frac{P(A) * P(B|A_i)}{((P(A_1) * P(B_1|A_1)) + \dots + (P(A_n) * P(B_n|A_n)))}$$

Ecuación 18: Conteos.

$$m \text{ y } n \text{ formas} = m * n$$

$$m \text{ o } n \text{ formas (solo una de ellas)}$$

$$nPr = \frac{n!}{(n-r)!}$$

Donde P es permutación y r es la razón escogida.

$$nC_r = \frac{n!}{r!(n-r)!}$$

Donde C es combinación.

n^r , caso especial.

Ecuación 19: Varianza probabilística.

$$u = \sum(x * P(X)^2) = nP * (X)$$

$$\sigma^2 = \sum(X - u)^2 * P(X)$$

$$\sigma^2 = \sum X * P(X) - u^2$$

Ecuación 20: Distribución binomial.

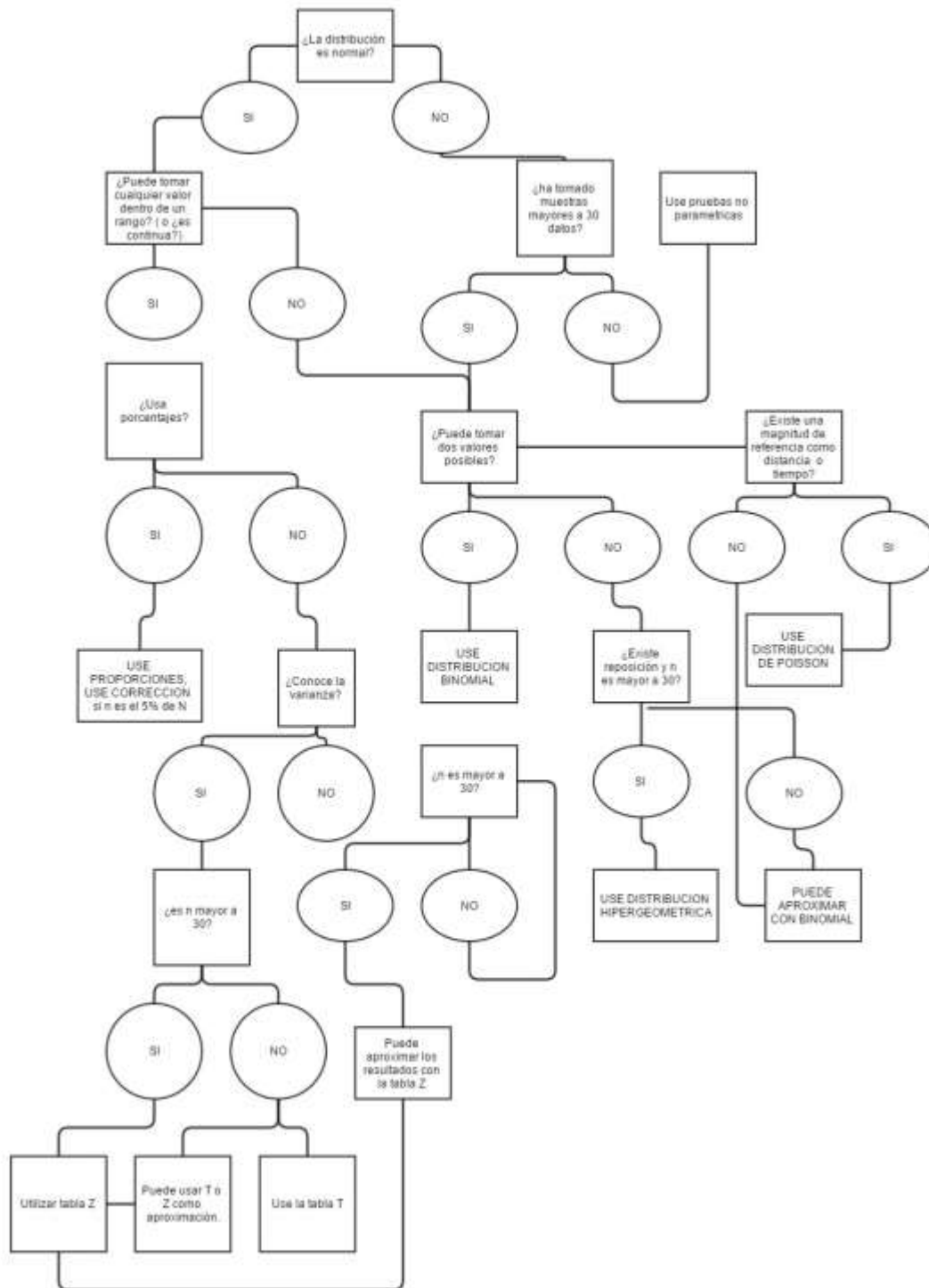
$$P(X) = nC_r * P(X)^r (X)(1 - P(X))^{1-r}$$

Ecuación 21: Distribución híper-geométrica.

$$P(X) = \frac{d^X * e^{-d}}{X!} \text{ Donde } d \text{ es "distancia", y, } e \text{ es el valor logarítmico.}$$

$$\sigma^2 = u$$

ANEXO de estadística: Breve Árbol de decisiones para resolver problemas estadísticos de probabilidad.



Elaborado por Eco. Luis Damian Jungnikel Matamoros.

1. Escogiendo un modelo y del tipo de regresión.

En esta primera parte, se han tomado como referencia dos autores que sin duda recopilan de forma amena y completa la economía introductoria y que han servido de base para la realización de este trabajo: (Gujarati & Porter, 2009), (Perez Lopez, 2006) y (Salvatore, 1998).

1.1. La regresión.

Salvatore define a la regresión como la hipótesis que se plantea entre la relación lineal (en los parámetros exclusivamente) entre una variable que se llamará explicada, endógena o dependiente, entre una o más variables de nombre explicativa, exógena, o independiente.

Lo cual plantea, entonces, la existencia de coeficientes que indican el efecto multiplicativo del comportamiento de variables en sus resultados, en el caso de ecuaciones lineales, y porcentuales en ecuaciones logarítmicas.

Ecuación 22: Modelo de regresión lineal.

$$Y = f(X_1, X_2, \dots, X_i)$$

$$Y = a + bX + u$$

Donde:

a es constante,

b es la pendiente,

X es exógena,

Y endógena, y

u son los errores.

El método más general para realizar la regresión es el de Mínimos Cuadrados Ordinarios MCO, pero tiene una restricción, debe cumplir con los supuestos la estimación del modelo no presente sesgos o viole los supuestos de los MCO.

Nuestras investigaciones económicas incluyen un término llamado perturbación aleatoria o término del error (lleva generalmente la letra e, u, v, etc.) y el MCO busca minimizar la diferencia de esos errores, elevándolos al cuadrado para darle mayor peso a dichos errores y obtener una regresión óptima. Por ello la regresión econométrica deberá cumplir los supuestos de los supuestos MCO.

1.2. Supuestos de los mínimos cuadrados ordinarios

- 1) Se supone que el error esta normalmente distribuido.
- 2) El error tiene media de cero.
- 3) Los términos del error no están correlacionados entre sí o con las variables.
- 4) Los errores tienen una desviación aproximadamente constante, y no muestran tendencias.
- 5) Las explicativas y los errores no deben estar relacionados.

1.3. Modelos lineales.

(Véase fórmulas para Plantear una regresión lineal, ecuaciones normales)

Ecuación 23: Interpretación del modelo lineal.

$\Delta Y = B_1 \Delta X$ Donde B^0 no se relaciona con la variación de X.

Es decir las variaciones de X generan variaciones en Y.

1.4. Ecuaciones normales para realizar una regresión.

Primero se plantea la ecuación. La constante es multiplicada por "n" (número de datos), las variables son reemplazadas por sus sumatorias.

Ecuación 24: Reemplazo de variables por sumatorias.

$$\sum Y = nB_0 + B_1 \sum X_1$$

Para obtener así la primera ecuación normal. Luego se multiplica toda la ecuación por la sumatoria de cada una de las variables explicativas de la primera ecuación normal. Así obtenemos las ecuaciones que sean necesarias, ya que si hay una variable explicativa, debemos obtener dos ecuaciones normales. Es decir el número de ecuaciones normales depende del número de variables explicativas más la explicada.

Las ecuaciones normales son válidas para todo tipo de ecuaciones, solo si ha sido linealizada primero.

Ecuación 25: Ecuaciones normales de la regresión lineal.

$$\sum YX_1 = B_0 \sum X_1 + B_1 \sum X_1^2$$

$$\sum YX_2 = B_0 \sum X_1^2 + B_1 \sum X_1 X_2$$

Entonces, reemplazamos en las tres ecuaciones) normales encontradas con sus respectivas sumatorias (valido solo para dos exógenas).

Nótese que B_0 se refiere a la contante.

Luego se utiliza el método para resolución de ecuaciones simultáneas que más le convenga.

1.5. Modelos no lineales.

Su construcción es muy simple y se aplican las mismas reglas de una regresión lineal, ya que lo que interesa, es en sí, que los parámetros (tales como coeficientes y pendientes) sean lineales. Es decir, que un modelo esté expresado en potencias, logaritmos, quebrados no afecta al modelo, y no quebranta ningún supuesto de los MCO, siempre y cuando no se alteren sus parámetros ni los errores.

Si va a realizar una regresión de forma "manual" o UNPLUGGED, deberá transformar los datos en la construcción de las tablas conforme lo pide la nueva ecuación.

Ecuación 26: Regresión exponencial linealizada en modelo log-lin.

(Expresión inicial)

$$Y_i = b_0 b_1^x e_i$$

(Expresión linealizada)

$$\log Y = \log b_0 + x_i \log b_1 + \text{Log} e_i ; \text{ Donde log significa logaritmo y lin es lineal}$$

(Interpretación)

$$\% \Delta Y = (100 B_1) \Delta X \text{ Es decir elasticidades.}$$

Ecuación 27: Regresión potencial linealizada modelo log-log.

(Expresión inicial)

$$Y_i = b_0 X_i^{b_1} e_i$$

(Expresión linealizada)

$$\log Y_i = \log b_0 + b_1 \log X_i + \log e_i$$

(Interpretación)

$$\% \Delta Y = B_1 \% \Delta X \text{ En elasticidades.}$$

Ecuación 28: Modelo potencial, Cobb-Douglas.

(Expresión inicial)

$$Y = B_0 X_1^{B_1} X_2^{B_2} u$$

(Expresión linealizada)

$$\log Y = \log b_0 + b_1 \log X_1 + b_2 \log X_2 + \log(u)$$

(Interpretación)

$$\% \Delta Y = B_j \% \Delta X \text{ En elasticidades.}$$

Ecuación 29: Regresión recíproca linealizada en inversa.

(Expresión linealizada)

$$\frac{Y}{X} = b_0 + B_1 \left(\frac{1}{X} \right)$$

1.6. Errores de las perturbaciones (Simple).

Ecuación 30: Errores de las perturbaciones. Una variable explicativa (una variable exógena o simple)

$$\sum \hat{u}_i^2 = \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \hat{B}_1 - \hat{B}_2 X_i)^2$$

Función objetivo → Minimizar.

Se busca minimizar el impacto de los errores en la regresión final.

1.7. Ecuaciones para hallar los parámetros (Simple).

Ecuación 31: Encontrar los parámetros lineales de ecuación simple

$$\hat{B}_2 = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2}$$

(A través de sumatorias)

$$\hat{B}_2 = \frac{\sum xy}{\sum x^2}$$

(A través de desviaciones)

$$\hat{B}_1 = \bar{Y} - B_2 \bar{X}$$

(A través de promedios).

1.8. Ecuaciones para hallar los parámetros (Múltiple)

Ecuación 32: Encontrar los parámetros lineales de ecuación múltiple.

$$\min \sum \hat{u}_i^2 = \sum (Y_i - \hat{B}_1 - B_2 X_2 - B_3 X_3)^2$$

$$\bar{Y} = \hat{B}_1 + \hat{B}_2 \bar{X}_2 + \hat{B}_3 \bar{X}_3$$

Básicamente son los mismos pasos de una ecuación simple, para obtener las ecuaciones normales complejas.

$$Y_i = n \hat{B}_1 + \hat{B}_2 \sum X_{2i} + \hat{B}_3 \sum X_{3i}$$

$$Y_i X_{2i} = \hat{B}_1 \sum X_{2i} + \hat{B}_2 \sum X_{2i}^2 + \hat{B}_3 \sum X_{2i} X_{3i}$$

$$Y_i X_{3i} = \hat{B}_1 \sum X_{3i} + \hat{B}_2 \sum X_{2i} X_{3i} + \hat{B}_3 \sum X_{2i}^2$$

$$\bar{B}_1 = \bar{Y} - \bar{B}_2 \bar{X}_2 - \bar{B}_3 \bar{X}_3$$

$$\hat{B}_2 = \frac{(\sum y_i x_{2i})(\sum x_{3i}^2) - (\sum y_i x_{3i})(\sum x_{2i} x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2}$$

$$\hat{B}_3 = \frac{(\sum y_i x_{3i})(\sum x_{2i}^2) - (\sum y_i x_{2i})(\sum x_{2i} x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2}$$

Nótese que minúsculas se refiere a desviaciones respecto a la media y que B_1 se refiere a la constante.

(Por algebra matricial de los parámetros de una regresión lineal, simple o múltiple).

$$y = X\hat{B} + \hat{u}$$

$$\begin{array}{cccc} y & x & B & u \\ n * 1 & n * k & k * 1 & n * 1 \end{array}$$

$$\hat{B} = \begin{array}{cc} (X'X)^{-1} & X' y \\ k * 1 & k * k \quad (k * n)(n * 1) \end{array}$$

En donde...

X' es matriz prima. X^{-1} es matriz inversa.

1.9. Varianza de errores y parámetros.

(Modelo simple.)

Hallar primero las perturbaciones.

Ecuación 33: Modelo para hallar perturbaciones.

$$\sum \hat{u}_i^2 = \sum y_i^2 - \hat{B}_2^2 \sum x_i^2$$

$$\sum \hat{u}_i^2 = \sum y_i^2 - \left(\frac{(\sum x_i y_i)^2}{\sum x_i^2} \right)$$

Ecuación 34: Varianza y covarianza simple

$$\hat{\sigma}^2 = \sum \hat{u}_i^2 / (n - 2)$$

$$\text{var}(\hat{B}_2) = \hat{\sigma}^2 / \sum x_i^2$$

$var(\widehat{B}_1) = \sum x_i^2 / n \sum x_i^2$ en donde B_1 es la constante.

$$cov(\widehat{B}_1, \widehat{B}_2) = -\bar{X} var(\widehat{B}_2) = -\bar{X} \left(\frac{\sigma^2}{\sum x_i^2} \right)$$

(Modelo múltiple o dos exógenas).

Hallar primero las perturbaciones.

Ecuación 35: Modelo para encontrar perturbaciones múltiples.

$$\sum u_i^2 = \sum y_i^2 - \widehat{B}_2 \sum y_i x_{2i} - \widehat{B}_3 \sum y_i x_{3i}$$

$$\widehat{\sigma}^2 = \sum \widehat{u}_i^2 / (n - 3)$$

Ecuación 36: Varianza y covarianza múltiple.

$$var(\widehat{B}_1) = \left(\frac{1}{n} + \frac{(\bar{X}_2^2 \sum x_{3i}^2 + \bar{X}_3^2 \sum x_{2i}^2 - 2\bar{X}_2 \bar{X}_3 \sum x_{2i} x_{3i})}{\sum x_{2i}^2 \sum x_{3i}^2 - (\sum x_{2i} x_{3i})^2} \right) * \sigma^2$$

En donde B_1 es la constante.

$$var(\widehat{B}_2) = \left(\frac{\sum x_{3i}^2}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2} \right) * \sigma^2$$

$$var(\widehat{B}_3) = \left(\frac{\sum x_{2i}^2}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2} \right) * \sigma^2$$

(Por algebra matricial.)

$$\sum \widehat{u}_i^2 = \sum y_i^2 - \widehat{B}_2^2 \sum x_i^2$$

$$\widehat{\sigma}^2 = \frac{\widehat{u}_i^2}{n-k} = \frac{\widehat{u} \widehat{u}}{n-k}$$

Donde K es el número de variables del modelo.

$$var - cov(\widehat{B}) = \sigma^2 (X'X)^{-1}$$

1.10. Regresión a través del origen

Ecuación 37: Regresión con constante igual a cero.

$$Y_i = \widehat{B}_2 X_i + \widehat{u}_i$$

$$\sigma^2 = \widehat{u}_i^2 / (n - 1)$$

$$\widehat{B}_2 = \sum X_i Y_i / \sum X_i^2 \quad \text{No existe } B_1 \text{ o constante.}$$

$$\hat{B}_2 = \sum x_i y_i / \sum x_i^2 \text{ (Por desviaciones).}$$

$$\text{var}(\hat{B}_2) = \sigma^2 / \sum X_i^2$$

$$\text{var}(\hat{B}_2) = \sigma^2 / \sum x_i^2 \text{ (Por desviaciones).}$$

$$r^2 = (X_i Y_i) / \sum X_i^2 \sum Y_i^2 \text{ (Coeficiente de determinación simple, aplicable).}$$

1.11. El uso de variables ficticias.

Cuando sea necesario, podemos agregar variables que describen situaciones que no se pueden analizar cuantitativamente, que generalmente llevan la letra D, es decir son cualitativas. Pueden expresar valores positivos entre el cero y el uno, y en algunos casos el menos uno,

Las variables dicótomas son variables cualitativas ficticias que solo pueden obtener valores fijos generalmente entre cero y uno e inclusive menos uno.

En el caso de estudio de variables endógenas con valores probabilísticos, se utilizaran variables ficticias con un rango variable entre cero y uno. Es así como investigador tiene la oportunidad de incluir fenómenos sociales, probabilísticos, y cualitativos, al modelo, además de comportamientos inusuales dados en épocas históricas por medio de las variables dicótomas, como efecto de guerra, comparación de resultados de modelos de planificación económica, cambio climático, etc.

1.12. Modelos con variables dicótomas.

Ecuación 38: Modelo con dicótomas en la constante.

$$Y = b_0 + b_1 X_1 + B_2 D + u \text{ (Para medir el efecto en la constante, donde D es dicótoma)}$$

Cuando D es cero, entonces:

$$Y = b_0 + B_1 X_1$$

Cuando D es uno, entonces:

$$Y = (B_0 + B_2) + B_1 X_1$$

Ecuación 39: Modelo con dicótoma en la pendiente.

$$Y = B_0 + B_1 X_1 + B_2 X * D + u \text{ (Para medir el efecto de la dicótoma en la pendiente).}$$

Cuando D es cero, entonces:

$$Y = B_0 + B_1 X_1$$

Cuando D es uno, entonces:

$$Y = B_0 + (B_1 + B_2) X_1$$

Ecuación 40: Modelo con dicótoma en constante y pendiente.

$$Y = B_0 + B_1X_1 + B_2D + B_3X_1 * D + u \quad (\text{Para medir el efecto de las dicótomos en la pendiente y la constante.})$$

Cuando D es cero, entonces:

$$Y = B_0 + B_1X_1$$

Cuando D es uno, entonces:

$$Y = (B_0 + B_2) + (B_1 + B_3)X_1$$

Estos modelos con variables dicótomos presentan el inconveniente de amplificar la probabilidad de ocurrencia de Multi-colinealidad, por tanto, se deben agregar menos variables que categorías, o aumentar el número de la muestra.

1.13. Regresión lineal por secciones.

Sirve para medir el efecto de medidas tomadas con la condición de un mínimo, por ejemplo: cuanto se incrementa la pendiente de tributación cuando se supera la base imponible.

Ecuación 41: Modelo de Umbral

$$Y = B_0 + B_1X_1 + B_2(X_1 - U)D + u$$

(En donde D es dicótoma y U es el Umbral y funciona como un nodo que conecta dos pendientes).

Cuando no se supera el Umbral, entonces:

$$Y = B_0 + B_1X_1$$

Cuando se supera U (o el Umbral) se puede interpretar:

$$Y = B_0 + (B_1 + B_2)X_1$$

O efecto total y se puede observar cómo se incrementó el efecto de la medida ya que:

$$Y = B_0 + B_1X_1 + B_2X_2 \quad (\text{En donde } X_2=(X_1-U))$$

1.14. Efecto multiplicativo de interacción

Cuando se desea medir como afectan dos variables cualitativas, dicótomas o estocásticas es necesaria realizar la siguiente manipulación:

Ecuación 42: Modelo de interacción.

$$Y = B_0 + B_1X_1 + B_2D_1 + B_3D_2 + B_4(D_1D_2) + u$$

En donde B4 es el efecto de interacción.

1.15. Regresión semi logarítmicas.

Ecuación 43: Modelo semi logarítmico

$$\text{Log}(Y) = B_0 + B_1D + u$$

Interpretación:

Se diferencia de los modelos anteriores porque su análisis no es por efecto multiplicativo, si no por elasticidades.

$(e^{B_1}-1)*100$ es decir cambio porcentuales.

1.16. Regresión con variable dicótoma endógena de decisión o ficticia

Ecuación 44: Modelo Potencial y con dicótoma.

Se utiliza un modelo potencial, preferentemente:

$$Y = e^{B_1D}$$

Si D=0 queda:

$$Y = e^{B_1D} = 1$$

Cuando D=1

$$Y = e^{B_1D} = e^{B_1}$$

1.17. El uso de ecuaciones simultáneas.

El problema que implica el uso de ecuaciones simultáneas está en determinar si la ecuación está identificada, con la finalidad de hallar coeficientes únicos como resultado, de coeficientes estructurales de los coeficientes reducidos.

Si una ecuación está perfectamente identificada se puede utilizar MCI o MC2E indiferentemente, en algunos casos se prefiere MCI por su sencillez, pero trae consigo la dificultad de encontrar las varianzas.

Si una ecuación este sobre identificado se puede utilizar MC2E.

El método de MC2E, para encontrar cual es la ecuación identificada es simple:

A. Se enumeran las variables endógenas y exógenas del sistema de ecuaciones,

B. De ello se toma una ecuación de referencia,

Donde A= (Las variables exógenas excluidas en la ecuación de referencia)

Y B= (las variables endógenas entradas menos uno)

A > B está sobre identificada,

A = B está identificada,

A < B está sub-identificada, o no tiene solución.

Si esta exactamente identificada se puede obtener coeficientes únicos como resultado.

En sobre identificada pudiesen hallarse más de una solución al sistema de ecuación.

En ecuación sub identificada no es posible una solución.

1.18. Método MCI.

Realice los siguientes pasos:

Ecuación 45: Modelo de ecuaciones simultáneas

1.-después de identificar, Obténgase las ecuaciones reducidas

$$Y_1 = B_0 + B_1 X_1 + u_1$$

$$Y_2 = B_2 + B_3 X_2 + u_2$$

2.- Aplíquese MCO para encontrar los parámetros de las ecuaciones reducidas

3.- Obténgase los nuevos coeficientes a partir de las reducidas (ahora estructurales)

$$B_4 = B_2 - B_5(B_0)$$

$$B_5 = B_4/B_2$$

4.- La ecuación final queda así:

$$Y_1 = B_4 + B_5 Y_2$$

1.1.6. Método de MC2E.

Realice los siguientes pasos:

1.-después de identificar, de la su identificada, agregase todas las demás exógenas que no entran el modelo.... (Variables instrumentales).

2.- Obténgase las estimadas.

3.- Haga una regresión con las estimadas, donde la endógena de la su identificada se transforma en exógena.

Queda a consideración del lector profundizar los pros y los contras de los modelos

Capítulo 2.

2. Pruebas básicas al realizar un modelo econométrico.

2.1. Prueba de significación de los parámetros.

Con el fin de probar si los coeficientes son o no diferentes a cero, debemos plantear la hipótesis nula de que los coeficientes son cero, y la hipótesis alternativa que no son cero, debemos encontrar la desviación estándar. Y aplicar el supuesto de normalidad donde la media muestral es cero.

Ecuación 46: Valor Z o T calculado.

$$t \text{ o } Z = \frac{\hat{B}_i - B_i}{\sigma(\hat{B}_i)}$$

En donde B_i es cada uno de los parámetros de la ecuación, y σ es la desviación estándar.

Regla:

Si T o Z calculada > T o Z de tabla (en valor absolutos) al nivel de significancia, es estadísticamente significativo, es decir, NO acepte que los coeficientes son cero. En otras palabras los coeficientes pasan la prueba. (No olvide que grafique para evitar equivocaciones).

En el Software Econométrico si el valor P (o probabilístico) es menor a 0,05 No acepte (o rechace) que el coeficiente es cero, es decir, los coeficientes pasan la prueba.

2.2. La bondad del ajuste.

Es necesario conocer que tan exacta es la regresión, es decir, que porcentaje de los datos se ajustan a la regresión para ello recurriremos al coeficiente de determinación, mejor conocido como R^2 . Este estadístico se puede obtener rápidamente por medio de la tabla ANOVA (o de análisis de varianzas).

La determinación de los datos es un problema de grado. Los valores que puede obtener el R^2 son de Cero a Uno.

Ecuación 47: El problema de bondad de ajuste

$$1 \geq R^2 \geq 0$$

Reglas:

Cuando $R^2 = 0$ los datos no se ajustan a la regresión, plantéese otro modelo.

Cuando $R^2 \leq 0,5$ la regresión tiene un ajuste débil.

Cuando $R^2 \geq 0,5$ la regresión tiene un ajuste mediano.

Cuando $R^2 \geq 0,8$ La regresión es fuerte (o más del 80% de los datos se ajustan).

Cuando $R^2 = 1,0$ La regresión es perfectamente ajustada.

Ecuación 48: La bondad del ajuste Simple, desde la tabla ANOVA

$$STC = SEC + SRC$$

En donde:

STC o suma de errores totales al cuadrado: $(Y - \bar{Y})^2 = y$

SEC= Suma de errores explicados: $(\hat{Y} - \bar{Y})$

SRC=Suma de errores no explicados: $(Y - \hat{Y}) = \hat{y}$

$$1 = \left(\frac{SEC}{STC}\right) + \left(\frac{SRC}{STC}\right)$$

$$R^2 = 1 - \left(\frac{SRC}{STC}\right)$$

Tabla ANOVA.

Tipo de error.	Cuadrados.	Grado de libertad.	
SEC o de regresión. No explicada.	$\hat{y}_i^2 = \tilde{B}_2^2 \sum x_j^2$	1	$\tilde{B}_2^2 \sum x_j^2$
SRC o residuos. Explicada	$\sum u_i^2$	n-2	varianza
STC o totales.	$\sum y_i^2$	n-1	$F = ((SEC/gl)/(SRC/gl))$

Ecuación 49: La bondad del ajuste. Varias fórmulas

(Por sumatorias)

$$R^2 = \frac{(\hat{Y} - \bar{Y})^2}{(Y - \bar{Y})^2} = SEC/STC$$

(Por desviaciones)

$$R^2 = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = \frac{\tilde{B}_2^2 \sum x_i^2}{\sum y_i^2} = \tilde{B}_2^2 \left(\frac{\sum x_i^2}{\sum y_i^2}\right) = \tilde{B}_2^2 \left(\frac{s_x^2}{s_y^2}\right) = (\sum y_i \hat{y}_i)^2 / ((\sum y_i^2)(\sum \hat{y}_i^2))$$

(Múltiple)

$$R^2 = (\tilde{B}_2 \sum y_i x_{2i} + \tilde{B}_3 \sum y_i x_{3i}) / \sum y_i^2 \text{ (Por desviaciones)}$$

(Ajustada)

En la bondad de ajuste se considera también el coeficiente de determinación ajustado, el cual “castiga” al coeficiente de determinación con el número de datos y el número de variables explicativas. Pocos datos o muchas variables, reducirán drásticamente el R^2 .

$$R^2 \text{ asjusted} = 1 - (1 - R^2) \frac{(n-1)}{(n-k)}$$

(Por matrices)

$$R^2 = \frac{\hat{B}'X'y - n\bar{Y}^2}{y'y - n\bar{Y}^2}$$

2.3. La prueba Global.

Es la proporción de la variación de la endógena o dependiente, explicada por las variables independientes en su conjunto, dicho de forma más simple, se busca comprobar la hipótesis que todos los coeficientes de la ecuación en su conjunto no sean cero.

Para ello se realiza la distribución F y se compara a la distribución F de tabla.

Nótese que se incluye anteriormente la tabla Anova (análisis de Varianza) que simplifica el desarrollo de la distribución F calculado.

Ecuación 50: La prueba global o Valor F

(Simple)

$$\text{valor } F = \frac{SEC}{\frac{SRC}{gl}} = \left(\frac{\hat{B}_2^2 \sum x_i^2}{\sum u_i^2} \right)_{n-2}$$

(Múltiple)

$$F = \frac{\frac{B_2 \sum y_1 x_{21} + B_3 \sum y_1 x_{31}}{\sum u_{12}^2}}{n-3} = \left(\frac{SEC}{gl} \right) / \left(\frac{SRC}{gl} \right)$$

(Por Matrices)

$$F = \left(\frac{B'X'y - nY^2}{k-1} \right) / \left(\frac{y'y - B'X'y}{n-k} \right)$$

Si F calculado > F de la tabla (en valor absoluto) no acepte o rechace que $B_j=0$; en el software econométrico no acepte o rechace que $B_j=0$ si el valor p es menor a 0,05.

2.4. Correlación.

Para realizar la prueba de que grado de relación, entre dos variables exógenas, se realiza la prueba de

matrices de correlación, siendo este un problema de grado parecido al del R_2 . (Véase reglas de la bondad de ajuste).

A diferencia del R^2 , se toma en cuenta el signo de la relación.

Positivo= relación directa.

Negativo=relación inversa.

Prueba de Correlación

2.5. Multi-colinealidad o micro-numerosidad.

Se refiere al problema antes mencionado, la relación fuerte de una variable explicativa con otra explicativa en la matriz de correlación, la cual es violatoria de los supuestos de los MCO.

En este caso específicamente la Multi-colinealidad ocurre cuando una variable explicativa está fuertemente relacionada con las otras variables explicativas, y para comprobarlo debemos realizar una regresión para cada variable explicativa excluyendo a la dependiente (o endógena o explicada).

Asumimos el valor del R^2 como el coeficiente de correlación múltiple entre las exógenas. Ya que es un problema de grado se acepta hipótesis de Multi-colinealidad si el valor supera el 0,7 (ese valor depende de la significancia determinada por el investigador).

Podemos además verificar este problema por medio de las varianzas, siempre que se cumplan estas condiciones:

- A. Si los coeficientes son muy pequeños y
- B. la varianza es muy grande.

Además en el análisis del Software contemplaremos lo siguiente:

- A. Una $R^2 > 0,7$.
- B. F probabilístico $< 0,05$
- C. Muchas de las variables tienen probabilidad mayor a 0,05, aceptándose que sean coeficiente cero.

Se puede resolver en unos casos agregando más datos, o eliminando las variables correlacionadas, otra forma es convertir la ecuación al dividirla por una variable, por ejemplo:

Ecuación 51: Ejemplo de modelo para corregir Multi-colinealidad

$$\frac{Y}{X_3} = B_1 \left(\frac{1}{X_3} \right) + B_2 \left(\frac{X_2}{X_3} \right) + B_3 + \left(\frac{u}{X_3} \right)$$

Sin embargo, buscamos la sencillez y la menor manipulación de los errores.

2.6. Prueba de Multi-colinealidad

Factor inflado de varianza y Tolerancia:

Se acepta Multi-colinealidad cuando FIV (Factor inflado de varianza) tiende a Diez. Es perfecta cuando TOL (Factor de tolerancia) tiende a cero.

Ecuación 52: valor FIV

$$FIV = \frac{1}{1-R_{23}^2}$$

Ecuación 53: Valor TOL

$$TOL = \frac{1}{FIV} = (1 - R^2 \text{ adjusted})$$

2.7. Heterocedasticidad.

Violación del supuesto del MCO en el que los errores no son estables, no son iguales, no son constantes. Caso contrario se define como homocedasticidad, cuando la varianza de los errores es constante.

En las pruebas (generalmente) se realiza una regresión con los errores, y de ahí se realiza la pruebas globales y de distribución chí cuadrada (estadística 1-2), y si hay evidencia de linealidad, se acepta Heterocedasticidad, ya que si los errores son iguales, su pendiente sería cero.

En su análisis gráfico, se descarta la Heterocedasticidad si los puntos muestrales de los errores son aleatorios. Pero esto es solo una ayuda, debido a que las percepciones visuales pueden engañar.

2.8. Prueba de Heterocedasticidad

Prueba de Park.

1.-realiza la regresión

Ecuación 54: Prueba de Park

$$\text{Log}(\sigma_i^2) = \text{log}(\sigma^2) + B \text{log}(X_i) + v_i$$

F calculado >F tabla Heterocedasticidad (en valor absoluto).

En el software, Si el valor probabilístico de la Chí cuadrada excede 0,05 acéptese Homocedasticidad.

Prueba de Golden Quant.

1.- Ordénese los valores en orden ascendente

2.-Elimine las observaciones centrales

3.-Ajusten dos regresiones para el primer grupo y para el segundo grupo

4.-Calcule la razón

Ecuación 55: La razón en la prueba Golden Quant

$$\text{razón} = SRC(2)/SRC(1)$$

- A. Si F calculado (o razón) > F tabla heterocedásticos (en valor absoluto).
- B. Si F calculado (o razón) < F tabla Homocedástica (en valor absoluto).
- C. En el software, Si el valor probabilístico de la Chií cuadrada excede 0,05 acéptese Homocedasticidad.

Prueba PGB

- 1.- Obténgase residuos.
- 2.- Obténgase la varianza como sumatoria de errores dividido para n.
- 3.- Constrúyase Pi.

Ecuación 56: Valor Pi

$$Pi = u_i^2 / \hat{\sigma}^2$$

- 4.-Construyase regresión con Pi como endógena.
 - 5.-Obtengase el SEC
- Ecuación 57: Valor SEC

$$SEC = 1/2(SRC)$$

- A. Chií calculado (SEC) >chií tabla (M-1 GL) heterocedástica
- B. si chií calculado (SEC) <chií tabla heterocedástica (en valor absoluto)
- C. En el software, Si el valor probabilístico de la Chií cuadrada excede 0,05 acéptese homocedasticidad

Prueba de heterocedasticidad de White

- 1.- estime

Ecuación 58: Modelo para estimar Chi en la prueba White

$$u_i^2 = B_0 + B_1X_1 + B_2X_2 + B_3X_1^2 + B_4X_2^2 + B_5X_1X_2 + v$$

- 2.- calcule la chií cuadrada de tabla.

Chií calculada = n * R² En donde los grados de libertad son =m-1 auxiliar (las veces que se “potencia”).

Regla:

- Si chií calculado > chií tabla (M-1 GL) heterocedástica.
- Si chií calculado < chií tabla heterocedástica.

En el software, Si el valor probabilístico de la Chií cuadrada excede 0,05 acéptese homocedasticidad

Prueba de contraste de Glesjer.
EN EL PROGRAMA:

En EVIEWS realice la siguiente ecuación:

- 1.-
 $ABS(RESID) = LOG(Y)$

Regla:

Recházese heterocedasticidad si valor P de F es mayor a 0,05.

2.9. Auto correlación.

La auto correlación se refiere a un problema de violación de los supuestos de los MCO y que ocurre generalmente en series de tiempo, los errores presentan tendencia en el tiempo, y supone la existencia de ciclos y/o tendencias, relaciones no lineales, la omisión de variables relevantes para ello recurrimos a una serie de pruebas.

Las formas de corregir el problema están en transformar los datos o recurriendo a el modelo Cochrane Rut, en el programa econométrico.

EN EL PROGRAMA:

Pasos:

- 1.- Realizar la regresión en EVIEWS.
- 2.- Verifíquese si existe auto-correlación.
- 3.-Realizar una nueva regresión agregando la variable dummy, es decir AR (1) o MA (1).

2.10. Prueba de Auto correlación

Prueba de Durbin Watson

- 1.-Calcúlese los residuos y residuos rezagados.
- 1.- Calcúlese la d.

Ecuación 59: Valor d

$$d = \frac{\sum(\hat{u} - \hat{u} \text{ rezagado})^2}{\sum \hat{u}^2}$$

Se debe proceder a estimar los residuos y los residuos rezagados.

Divida el rango "d" de tabla en cuatro segmentos.

se acepta hipótesis de auto correlación positiva.

si cae dentro de los dos segmentos de en medio, rechazese hipótesis de auto correlación.

se acepta hipótesis de autocorrelación negativa.

EN EL PROGRAMA:

En EViews revítese la tabla de salida de la ecuación estimada. Si el valor de Durbin Watson supera 2, sospéchese de Auto-correlación.

Ilustración 1: Prueba de auto correlación.

Para el análisis de la tabla d (Durbin Watson) debemos dividir el intervalo para cuatro, siendo los cuartos al extremo como zona de indecisión. Esto es válido para probar hipótesis de existencia de Auto correlación positiva y negativa.

Para el caso de auto correlación positiva solamente solo se considerara que el valor no sea menor que el rango, es decir si el valor cae dentro de rango es indeciso, entonces:

Si el valor d calculado > d tabla inferior recházese auto correlación positiva.

2.11. Modelo rezagado para corregir auto correlación.

1.- Aproxímese el valor p:

Ecuación 60: Valor p aproximado.

$$\tilde{p} = 1 - \left(\frac{d \text{ calculado}}{2} \right)$$

2.- modifíquese los datos:

Ecuación 61: Modelo rezagado

$$Y^* = (Y - p * Y \text{ rezagado}) \quad X^* = (X - p * X \text{ rezagado})$$

3.- el primer dato rezagado sería:

Ecuación 62: Calculo del primer dato rezagado.

$$\text{primera } Y \text{ rezgada} = Y * (\sqrt{1 - p}) \quad \text{primera } X \text{ rezgada} = X * (\sqrt{1 - p})$$

4.- Realice la siguiente regresión.

Ecuación 63: Regresion ajustada para disminuir auto correlación.

$$Y^* = B_0 + B_1 X^* + \varepsilon$$

5.- Compruébese la disminución de la ocurrencia de auto correlación.

Capítulo 3

3. Prediciendo el futuro.

3.1. Intervalos de confianza.

Los intervalos de confianza, sirven para suponer en cuanto podrían variar los valores estimados, es decir, el alcance de predecir de nuestro modelo, respecto a los verdaderos valores (dentro de la significancia elegida).

Entre ellos tenemos intervalos de confianza para los coeficientes, para la media, para las varianzas, para las predicciones, sean individuales o de la media.

3.2. Prueba de Intervalos de confianza.

Estimación de intervalos

- 1.- Primero calcúlese la varianza de los parámetros.
- 2.- Defínase la significancia, para obtener la t de tabla.
- 3.- Calcúlese los rangos de predicción.

Ecuación 64: Rangos de predicción.

$$\mathbf{rango inferior} = B_j - t(\sigma(B_j))$$

$$\mathbf{rango superior} = B_j + t(\sigma(B_j))$$

Intervalos de confianza para la varianza

- 1.- Calcules la varianza de los errores.

2.- Determinése el nivel de significancia para la Chií de tabla.

3.- Calcúlese los rangos de predicción.

Ecuación 65: Rango de predicción de Chií

$$\text{rango inferior} = (n - 2) \left(\frac{\sigma^2}{\text{chií}_{\frac{\alpha}{2}}} \right)$$

$$\text{rango superior} = (n - 2) \left(\frac{\sigma^2}{\text{chií}_{(1-\frac{\alpha}{2})}} \right)$$

Ecuación 66: Predicción media

$$\text{var}(\hat{Y}_0) = \sigma^2 \left(\left(\frac{1}{n} \right) + \left(\frac{(x_0 - \bar{x})^2}{\sum x_i^2} \right) \right)$$

$$\text{Rango inferior} = \hat{B}_1 + \hat{B}_2 X_0 - (t * \sigma(\hat{Y}_0))$$

$$\text{Rango superior} = \hat{B}_1 + \hat{B}_2 X_0 + (t * \sigma(\hat{Y}_0))$$

Predicción individual

1.- calcúlese la varianza de la predicción de Y.

Ecuación 67: Varianza de predicción de Y

$$\text{var}(Y_0 - \hat{Y}_0) = \sum (Y_0 - \hat{Y}_0)^2 = \sigma^2 \left(1 + \left(\frac{1}{n} \right) + \left(\frac{(x_0 - \bar{x})^2}{\sum x_i^2} \right) \right)$$

2.- calcúlese el rango de predicción.

Ecuación 68: Rango de predicción individual.

$$\text{rango inferior} = \hat{B}_1 + \hat{B}_2 X_0 - (t * \sigma(\hat{Y}))$$

$$\text{rango superior} = \hat{B}_1 + \hat{B}_2 X_0 + (t * \sigma(\hat{Y}))$$

Por algebra matricial.

Ecuación 69: Rango de predicción por algebra matricial.

$$\text{var}(\hat{Y}_0 | X_0) = \sigma^2 X_0' (X'X)^{-1} X_0 \quad \text{var}(Y_0 | X_0) = \sigma^2 (1 + X_0' (X'X)^{-1} X_0)$$

Capítulo 4

4. Pruebas definitivas para escoger el modelo.

4.1. Pruebas de la estructura del modelo

Debemos primero, en secuencia lógica, comprobar las pruebas anteriores, para comprobar si la estructura del y especificación del modelo es correcta.

4.1.1. Prueba de simultaneidad o de Hausman.

- 1.- En un sistema de ecuaciones, elijase una variable endógena y realice una regresión.
- 2.- Obtengase una regresión de la otra variable endógena y los residuos de la primera regresión.

Regla:

Si el coeficiente de la segunda regresión es cero, rechácese problema de simultaneidad.

4.1.2. Prueba de exogeneidad.

Se usa para determinar que variable es endógena y cual es exógena en un modelo de ecuación múltiple.

- 1.- Calcúlese la siguiente regresión:

Ecuación 70: Prueba de exogeneidad.

$$Y_1 = B_0 + B_1 Y_2 + B_2 Y_3 + B_3 X_1 + B_4 \hat{Y}_2 + B_5 \hat{Y}_3 + u$$

Regla:

Si no pasa la prueba F con valor P mayor a 0,5 o F calculado < F tabla, entonces Y_2 y Y_3 son endógenas.

4.1.3. Prueba de estabilidad estructural o CHOW

Para comprobar si conviene agregar ficticias a un modelo.

- 1.- Se estima regresión SRC (suma de residuos cuadrados) restringida:
- 2.- Se agrupan los datos en dos partes.
- 3.- Se obtiene SRC1 y SRC2. Los grados de libertad serán:

Ecuación 71: Calculo de los grados de libertad, prueba CHOW

$$\text{Grados de libertad} = (n_1 + n_2 - 2K)$$

4.- se obtiene SRC no restringida.

Ecuación 72: Suma de errores, prueba CHOW

$$SRCnr = SRC1 + SRC$$

5.- Calcúlese F

Ecuación 73: Calculo de F o global en la prueba CHOW

$$F = \left(\frac{SRCr - SRCnr}{k} \right) / \left(\frac{SRCnr}{n_1 + n_2 - k} \right)$$

Regla:

Recuérdese que se trabaja con los grados de libertad descritos anteriormente.

Si F calculado > F tabla rechazar hipótesis de estabilidad estructural.

Si F calculado < F tabla aceptar la hipótesis de estabilidad estructural.

4.1.4. Prueba MWD de la forma funcional del modelo.

Sirve para comprobar la conveniencia de usar un modelo logarítmico.

Pasos:

1.- Haga la ecuación lineal y obtenga las estimadas.

2.- Realice la ecuación logarítmica y obtenga las estimadas.

3.- Obtenga los logaritmos de las estimadas lineales y luego réstalas de los estimados de los logarítmicos, se le llamara Z.

4.- Obtenga la regresión:

Ecuación 74: T calculado en la prueba MWD.

$$Y \text{ para } t \text{ calculado} = B_1 X + B_2 Z$$

Regla:

Si t calculado > t tabla es lineal.

5.- obtenga el antilogaritmo de Z, llámelo Z₂.

6.- obtenga la regresión:

Ecuación 75: Modelo para prueba MWD.

$$\log(Y) \text{ para } t \text{ calculado} = B_3 X + B_4 Z_2$$

Regla:

Si t calculado $>$ t tabla es logarítmico

4.1.5. Prueba reset de Ramsey.

Sirve para comprobar la especificación del modelo.

1.- Obtenga la Y estimada.

2.- Introduzca estos valores en la ecuación, ya sea como potencias, o por interacción, depende de usted. A dicha ecuación la conocemos como nueva.

3.- Obtenga R^2 , de las dos ecuaciones.

Ecuación 76: Calculo de F en la prueba RESET.

$R_2^2 = R^2$ nueva; $R_1^2 = R^2$ vieja; $M_2 =$ parametros del modelo nuevo; $K_2 =$ regresoras nuevas.

$$F \text{ calculada} = \frac{(R_2^2 - R_1^2) / k_2}{(1 - R_2^2) / (n - m_2)}$$

Regla:

Si F calculado $>$ F tabla acéptese la hipótesis de que el modelo estuvo mal especificado.

En el software, acepte una buena especificación del modelo si el valor probabilístico de chií cuadrada es mayor a 0,05 o si la F es menor a 0,05.

Criterio de Akaike y Swartz.

En EVIEWS, se encuentran dichos estadísticos.

1.- Plantéese varios modelos, con diferentes combinaciones de variables, lineales y no lineales.

Regla:

Escójase el modelo que tiene R^2 ajustado más alto, un valor Durbin Watson más cercano a 2 y los valores de criterio Akaike y Swartz más bajos.

4.2. Advertencia y conclusión.

Existen diferentes formas de entretener y demostrar con la econometría su poder práctico en el estudiante, por ejemplo, al hacer ejercicios de datos de series de tiempo anteriores al periodo actual, con la finalidad de obtener estimadores, y comprobar en cuanto ha acertado la predicción de sus modelos. Esto es muy importante, ya que, se entenderá la importancia que tiene la econometría en la práctica.

Es decir, estaremos pendientes de la consistencia de los modelos que el estudiante pudo realizar, y comprobar que sin importar la elegancia matemática, de nada sirve un modelo si no puede predecir las tendencias, o mucho menos explicarlas.

Además, repetiremos lo que manifestó Gujarati: “si se torturan los datos, estos delatarán lo que uno desee escuchar”, por lo que en esta recopilación breve e introductoria de fórmulas más utilizadas de econometría se buscó la sencillez de la explicación, y se espera que se aplique así en la práctica (Gujarati & Porter, 2009). En otras palabras, cada ejercicio pretende liberar al estudiante de tendencias nocivas que lo alejan de la pluralidad de paradigmas, y que responden a diferentes puntos de vista de investigadores, y formará su propio punto de vista, de cómo funciona la economía real, apegando sus resultados matemáticos (modelos), estadísticos (tablas y gráficos), econométricos (validación de los supuestos implícitos en el modelo, y en los MCO) en contrapeso con la realidad social.

Este texto no pretendió ser un texto de consulta, sino más bien un formulario de ayuda, de la cual puede valerse tanto el estudiante como el profesor, al momento de exponer su clase o resolver ejercicios econométricos.

Nótese que los ejercicios econométricos pueden ser extensos para realizarse manualmente, pero vale la pena realizar al menos un ejercicio completo, para ejercitar la destreza matemática, estadística y más aún, el análisis realizado con datos calculados a “mano alzada” tienen mejor sabor.

Bibliografía

- Gujarati, D. N., & Porter, D. C. (2009). *Econometría*. Santa Fé: Mc Graw Hill.
- Kennedy, P. (s.f.). *Universidad Simón Fraser*.
- Keynes, J. M. (1936). *Teoría general del empleo, el interés y el dinero*. Cambridge: Cambridge.
- Mason, & D, R. (2000). *Estadística Comercial y Económica*.
- Perez Lopez, C. (2006). *Ejercicios resueltos de econometría: Paso a Paso*. Madrid, ESPAÑA: Paraninfo.
- Ruiz Muñoz., D. (2004). *Manual de Estadística*. EUMED.
- Salvatore, D. (1998). *Econometría*.
- Webster, A. L. (1998). *Estadística*.