Universidad Católica del Norte

Taller de Econometría.

Profesor:

Andrés Castaño

Alumnos:

THIARA RIVERA

JULIO ORELLANA

KAREN CASTILLO

BARBARA AVALOS

JOSEFA PELLEJERO

MARIANA NADAL

CATALINA ANGUITA

- Antofagasta, mayo de 2013 -

* 5.1

(a) Verdadero. La teoría está basada en variables en una distribución normal.

Puesto que los estimadores  (beta1) y (beta2)  son combinaciones lineales del error u, el cual se supone que se distribuye normalmente bajo CLRM, estos estimadores son también distribuidos normalmente.

(b) Verdadero. Siempre que E(u)=0 los estimadores OLS son insesgados

Se requieren supuestos no probables  para establecer sesgo

(c) Verdadero. En el caso que Eq(1) en aplicación 3ª sección 3A.1 será ausente, Este tema será discutido a profundidad en el capt. 5 sección 6.1( ver foto) para sacar ecuación.

(d) Verdadero. El valor de P es el nivel menor  de significancia en que la hipótesis nula puede ser rechazada. Los términos de nivel de significancia y tamaño de este texto son sinónimos

(e) verdadero. Esto deriva de Eq. (1) en aplicación  3A sección 3A.1 ( ver foto)

(f) Falso. Todo lo que podemos decir es que los datos disponibles no permiten rechazar la hipótesis nula

(g) Falso. Una mayor ( sigma cuadrado ) puede ser contrabalanceada por una mayor (sumatoria de los xi cuadrado) según ecuación (3.3.1)

(h) Falso. La media condicional de una variable aleatoria depende de los valores condicionados tomados por otra variable. Solo si  dos variables son independientes, las medias condicional e incondicional pueden ser los mismos.

(i) verdadero, según la ecuación (3.1.7) beta 1 = sumatoria de los xi cuadradado \*sumatoria de los yi menos sumatoria de los xi \* sumatoria de los xiyi / n (sumatoria de xi cuadrado) menos ( sumatoria xi) cuadrado

Esto es igual a = la media de Y - beta 2 muestral por la media de X

(i) Verdadero. Esto es obvio a partir de Eq (3.1.7)

(j) Verdadero. Refiere de Eq (3.5.2) Si X no tiene influencia en Y, beta2 será cero, en el caso que  sumatoria de los yi cuadrado es igual a la sumatoria de los ui muestrales al cuadrado

* 5.2

ANOVA Tabla para el gasto en alimentos en India

Fuente de variación

Debido a la regresión

Debido a residuos

TSS

F= 139023… =31.1013 con df= y 53 respectivamente

Bajo la hipótesis que no hay relación entre gasto de alimentos y gasto total, El valor de p para obtener el valor de F es casi cero. Sugiriendo que uno puede rechazar la hipótesis nula firmemente.

* 5.3

(a) 𝑆𝑎𝑙𝑚𝑒𝑑𝑖𝑜=0.7437+0.6416𝐸𝑠𝑐𝑜𝑙𝑎𝑟𝑖𝑑𝑎𝑑

Ee = (0.8355) (0.64169.6536=0.0664)

𝑡=(0.74370.8355=0.8797) (9.6536) 𝑟2=0.8944 𝑛=13

(b) En promedio el salario medio aumenta en un 64 centavos (0.64 dólares) cuando la escolaridad varia marginalmente, es decir aumenta un año.

(c) Con n=13 y grados de libertad=11, el valor t=9.6536, obtener dicho valor es muy improbable con un p-valor de casi 0, por lo que se podría rechazar la hipótesis nula.

(d) Fuente de Variación Suma de los Cuadrados Grados de Libertad

De la Regresión 74.9389 1

De los Residuos 8.8454 11

F=93.1929, el p-valor de este estadístico F es 0.000001, tan pequeño que podríamos rechazar la hipótesis nula.

e) Sin haber conocido previamente el 𝑟2, se podría haber obtenido, ya que se dio el valor 𝑡, el cual se relaciona con este así: 𝑟2=𝑡2[𝑡2+(𝑛−2)] , ya que tenemos que el valor t es 9.6536, queda: 𝑟2=9.65362(9.65362+11)≈0.8944

* 5.4

Verbalmente, bajo la hipótesis de que no hay una correlación entre las dos variables. Por lo tanto, si podemos demostrar que la covarianza entre dos variables es cero, entonces la correlación debe ser cero.

* 5.5

(a) usar la prueba t para probar la hipótesis de que el verdadero coeficiente de pendiente es 1.

t=𝛽2̂−1𝑠𝑒(𝛽2̂ )=1,0598−10,0728=0,821

Para 238 grados de libertad la prueba t no es significante a α=10%

(b) 𝑡=0,72640,3001=2,4205 que es significativa a 2%.

* 5.6

𝛽2̂ esta normalmente distribuido bajo el supuesto de normalidad (y como esta normalmente distribuido es continua) además la probabilidad de que una variable aleatoria continua tome un valor específico es cero da lo mismo si la igualdad es fuerte o no.

* 5.7

Bajo la hipótesis que B2=0, obtenemos

t==

por que ==

pero desde r= B2

t=

* 5.8

(a) Hay una asociación positiva en la FRLP entre 1972 y 1968, lo es cual es bastante obvio debido a que desde la Segunda Guerra Mundial ha aumentando la población de mujeres.

(b) Utilizando la prueba t de una cola = -1.7542. El valor t de una cola a es α=5% ya que el valor t estimado es significativo, a este nivel de significación, podemos rechazar la hipótesis de que el coeficiente de la pendiente cierto es 1 o mayor.

(c) La media FRLP es:0.2033+0.6560(0.58)≅ 0.5838. Para establecer un intervalo de confianza del 95% de este valor pronóstico, utilice la siguiente fórmula: 0.5838+-2.11(sea del valor medio previsto). Donde 2,11 es el valor t crítico 5%.

(d) Si no conocemos los datos reales, no seremos capaces de responder a cabalidad esta pregunta, ya que necesitamos los valores de los residuos para trazar y poder obtener el grafico de probabilidad normal.

* 5.9

(a)

(b) Salario= 12129.37 + 3.3076Gasto

ee=(1197.351)(0.3117) r2= 0.6968

(c) Si el salario por alumno aumenta en un dólar, el salario medio aumentara en 3.31 dólares. El interceptor en este caso no tiene una definición económica.

(d) El intervalo de confianza de un 95% para betados es: 3,3076 + -2 (0,3117) = 2.6843,931. Basado en este intevarlo de confianza INTERVALO de confianza que no se posee suficiente información para rechazar la hipótesis nula de que cierto coeficiente de la pendiente es 3.

(e) Los valores medios e individuales del pronóstico son el mismo, a saber, 12129,37 +3,3076 (5000) = 28.667. El error estándar del valor medio previsión es 520.5117 dólares y el error estándar de la previsión individual es 2,382.337.

Los intervalos de confianza son:

Predicción media:

28667 + -2 (520,5157)

(27.626 dólares, $ 29.708)

Predicción individual:

28 667 + -2 (2382.337)

(23,902 dólares, 33,432 dólares)

Como era de esperar, este último intervalo es más ancha que la antigua

(f) Si hacemos un histograma de los residuos esta se aproxima a una curva normal.

* 5.10.-

La tabla ANOVA para el sector de negocios es:

|  |  |  |  |
| --- | --- | --- | --- |
| Fuente de Variación | Suma de los Cuadrados | Grados de Libertad | Varianzas |
| De la Regresión | 38685.997 | 1 | 38685.997 |
| De los Residuos | 4934.138 | 37 | 133.355 |
| Total | 43620.135 |  |  |

El valor F es:

Bajo la hipótesis nula de que no existe relación entre la productividad y la compensación salarial real en el sector de negocios, el valor F sigue la distribución F con 1 y 37 grados de libertad en el numerador y denominador respectivamente. La probabilidad de obtener este valor F es 0,0000, es decir, prácticamente 0. Así, podemos rechazar la hipótesis nula.

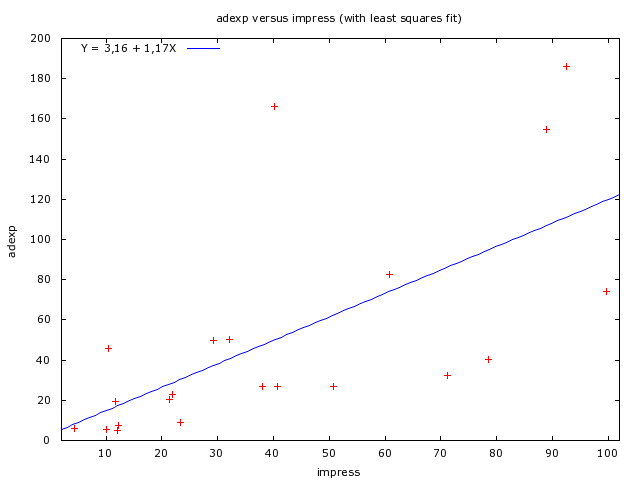
Para el sector de comercio no agrícola, la taba ANOVA es:

|  |  |  |  |
| --- | --- | --- | --- |
| Fuente de Variación | Suma de los Cuadrados | Grados de Libertad | Varianzas |
| De la Regresión | 37887.445 | 1 | 37887.445 |
| De los Residuos | 5221.585 | 37 | 141.129 |
| Total | 43109.04 |  |  |

STC: 43059.04 ; SRC:5221.585 ; SEC: 37837.455

Bajo la hipótesis nula que el coeficiente , la pendiente verdadera, es 0, el valor F es:

Si la hipótesis nula fuera cierta, sería prácticamente imposible obtener tal valor, por lo que se rechaza.

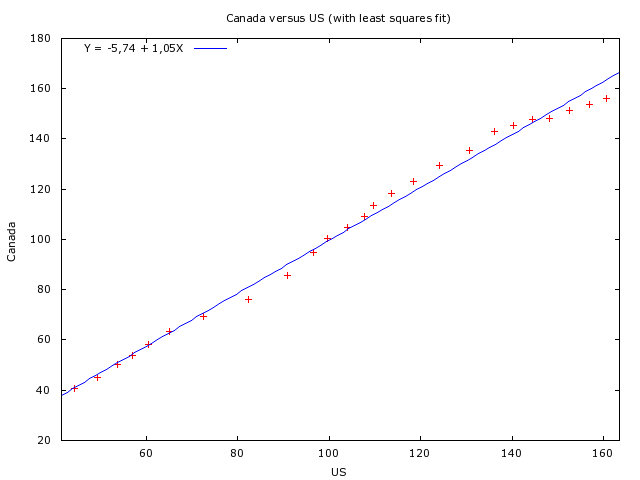
* 5.11.- a)El grafico muestra la relación entre las impresiones y el gasto publicitario

Ambas variables son no lineales, inicialmente al aumentar el gasto publicitario, también lo hacen las impresiones, pero luego decrecen.

b) No sería apropiado ajustar el modelo a un modelo de regresión dos variables lineales, pues ninguna de las variables lo es y no se ve una relación lineal, para este modelo podría ajustarse mejor un modelo cuadrático, donde Y representaría las impresiones y el gasto publicitario, así: , un modelo aun lineal en los parámetros pero cuadrático en la variable, aun así, aun no contamos con las herramientas para analizar este tipo de modelo de regresión.

c) Ajustando el modelo a una regresión bivariante, los resultados obtenidos son:

Ee: (7.08948) (0.0971203)

* 5.12

a) El grafico muestra que ambos índices son directamente proporcionales, si uno sube, el otro también.

b) VARIABLE COEFFICIENT STDERROR T STAT P-VALUE

const 6,21955 1,93254 3,218 0,00381 \*\*\*

Canada 0,941239 0,0173566 54,229 <0,00001 \*\*\*

El modelo sería:

C) Al ver los resultados de la regresión en gretl, la regresión de las dos variables es obviamente positiva. Sin necesidad de realizar pruebas se rechaza la hipótesis nula señalada, pues es estadístico t obtenido es: 54,229, y el valor p es prácticamente 0. Aunque la relación de ambos IPC es claramente positiva, no se puede condicionar el uno al otro, pues no hay teoría económica que lo respalde, lo cual es lo principal para darle sentido a una regresión econométrica.

* **5.13**

(A) Las dos regresiones son las siguientes

=0.150

=0.868

(B) El estadístico Jarque-Bera (JB) para la ecuación del precio del oro (Goldprice) es de 4.751, con un valor de p 0,093. El estadístico de JB para la ecuación NYSEIndex es 1,218 con un valor de p 0,544. A nivel de 5% de significancia, en ambos casos, no rechazamos la hipótesis de normalidad.

(C) Dado que el coeficiente de la pendiente de la regresión Goldprice no es estadísticamente distinto de cero, no tiene sentido saber si es diferente de 1.

(D) y (E) Utilizando el procedimiento de la prueba t, se obtiene:

Dado que este valor t supera el valor del t crítico 2.160, se rechaza la hipótesis nula. El coeficiente estimado es en realidad mayor que 1. Para este período de la muestra, la inversión en el mercado de valores fue probablemente una cobertura contra la inflación. Sin duda fue una mejor protección contra la inflación que la inversión en oro.

* **5.14**

(A) Ninguno parece ser mejor que los otros. Todos los resultados estadísticos son muy similares. Cada coeficiente de la pendiente es estadísticamente significativo al 99% de confianza.

(B) El constantemente alto no se puede utilizar para decidir qué agregado monetario es mejor. Sin embargo, esto no sugiere que no exista diferencia con respecto a que ecuación usar.

(C) No se puede deducir la respuesta con los resultados de la regresión. Pero últimamente el Banco de la reserva federal parece tener como objetivo la medida M2.

* **5.15**

Modelo de curva de indiferencia

=++

Ahora, pase y el parámetro de pendiente la intersección. Pero esto sigue siendo un modelo de regresión lineal, así como también los parámetros.

Los resultados de regresiones son las siguientes:

=3.2827+1.1009

=0.6935

El coeficiente de "pendiente" es estadísticamente significativo en el coeficiente de confianza 92%. La tasa marginal de sustitución (RMS) de Y para X es:

* **5.16**

a) ***Yᵢ = β₁ + β₂Xᵢ + uᵢ***

En donde ***Y*** es la tasa de cambio del día y ***X*** es la paridad del poder de compra (PPC). Si se mantiene el PPC, se esperaría que el intercepto sea 0 y la pendiente sea 1, cumpliendo así la teoría que con una moneda debe ser posible comprar la misma canasta de bienes.

b) Teniendo en cuenta los datos de la tabla 5.9, se obtendría que:

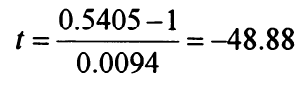
***Ŷᵢ = 24,6338 + 0,5405Xᵢ***

Ee(βᵢ)= (19,507) (0,0094)

*t* = (1,2628) (57,10161)

r² = 0,9917

Para probar la hipótesis nula que ***β₂ = 1***, se usa la prueba t-student, lo que nos entrega el siguiente resultado (t calculado):



El t calculado tiene un nivel altamente significativo, por lo que rechazamos la hipótesis nula. Es más, la pendiente es menor que 1.

c) Si el Economista continúa o no publicando el índice del Bigmac es irrelevante, aunque se pudo concluir que los resultados obtenidos no apoyan la teoría PPC.

* **5.17**

a) Teniendo los datos del ejercicio 2.16 y generando una regresión lineal de la forma ***Yᵢ = β₁ + β₂Xᵢ + uᵢ***, en donde ***Y*** representa los puntajes masculinos de matemáticas y ***X*** representa los puntajes femeninos, se obtiene la siguiente regresión:

***Ŷᵢ = 175,975 + 0,714Xᵢ***

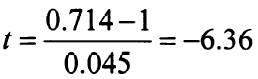
Ee(βᵢ)= (20,635) (0,045)

*t* = (8,528) (15,706)

r² = 0,918

b) Con un valor p cercano a 0, como 0,5970, el estadístico Jarque-Bera es de 1,0317. No se puede rechazar el supuesto de normalidad de los residuos estimados.

c) H₀: ***β₂ = 1*** ; es decir, que exista una correspondencia uno a uno entre los puntajes de matemáticas femeninos y los masculinos.



Es por esto, que con un 99% de confianza podemos rechazar la hipótesis nula.

d) La tabla ANOVA muestra la variabilidad explicada por el modelo más la variabilidad no explicada (residual), esto permitirá contrastar la significancia del modelo.

|  |  |  |  |
| --- | --- | --- | --- |
| Fuente de variación | Suma de cuadrados | Grados de libertad | Varianzas |
| Explicada | 948,193 | 1 | 948,193 |
| Residual | 87,782 | 22 | 3,990 |
| Total | 1071,975 | 23 |  |

Bajo la hipótesis nula que ***β₂ = 0***, según la tabla F, el estadístico F es 264,665; el valor p de haber obtenido ese valor F es casi 0, generando así el rechazo de la hipótesis nula.

* **5.18**

Ahora con los resultados de los puntajes de habilidad verbal masculinos (Y) y femeninos (X), se obtienen los siguientes resultados de la regresión lineal:

a) ***Ŷᵢ = 148,135 + 0,673Xᵢ***

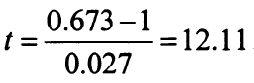
Ee(βᵢ)= (11,653) (0,027)

*t* = (12,713) (25,102)

r² = 0,966

b) Con un valor p de 0,5372, el estadístico Jarque-Bera es de 1,243. Por consiguiente podemos rechazar la hipótesis nula de no-normalidad.

c) H₀: ***β₂ = 1***, se obtiene el t calculado:



El t crítico al 5% de significancia es 2,074; es por esto que rechazamos la hipótesis nula de que la pendiente es igual a 1.

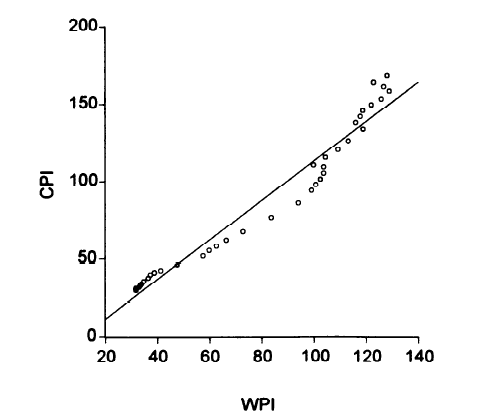
d) Tabla ANOVA:

|  |  |  |  |
| --- | --- | --- | --- |
| Fuente de variación | Suma de cuadrados | Grados de libertad | Varianzas |
| Explicada | 3157,586 | 1 | 3157,586 |
| Residual | 110,247 | 22 |  |
| Total | 32367,833 | 23 |  |

***β₂ = 0***, según la tabla F, el valor F es 630,131; el valor p de haber obtenido ese valor F también es casi 0, entonces también se rechaza la hipótesis nula, es decir, se rechaza la hipótesis que no haya relación entre las variables.

* **5.19**

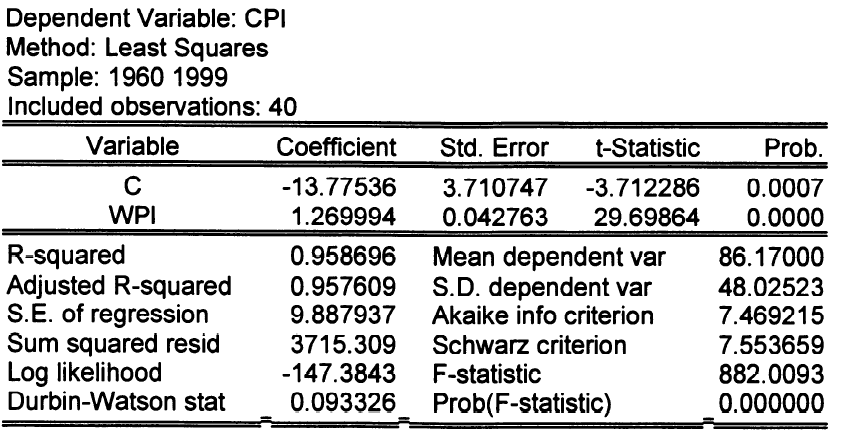
a)

****

Se estima una regresión lineal, es decir, una relación directa entre las variables IPC Y IPM, como se muestra en el gráfico expuesto.

b) El IPC representa los precios que pagan los consumidores, y el IPM representa el precio que pagan los productores. Es por esto que tratamos al IPC como la variable regresada y el IPM como la variable regresora.

c)



El valor t de la pendiente es 29,6986, bajo la hipotesis nula que no existe relacion entre los indices. El valor p de obtener ese valor t es casi 0, entonces podemos sugerir el rechazo de la hipótesis nula.

d) Con un valor p de 0,8456; el estadístico Jarque-Bera es de 0,3335. Entonces no es posible rechazar la normalidad en los erroes, pero es posible decir que los errores se distribuyen de manera simétrica.