Leslie Bustos Díaz

Cristian Cordero Goncalvez

Camila Lara Henríquez

Lizette Molina Aban

Alexandra Paniagua Iturra

Kevin Sepúlveda Castillo

Maribel Valderrama Jara

5.1- Verdadero y falso

a) La prueba t de significancia estudiada en este capitulo requiere que las distribuciones muéstrales de los estimadores B1 y B2(gorro) sigan una distribución normal.

Verdadero, la prueba de significancia desarrolla un estadístico de prueba y examina su distribución muestral bajo la hipótesis nula. Este estadístico de prueba es conocido como normal, la t, la F o la ji cuadrada.

b) Aunque el termino perturbación en el MCRL no esté normalmente distribuido, los estimadores MCO continúan siendo insesgados.

Verdadero, los estimadores MCO deben cumplir con 3 características para ser un MELI, determinados en el teorema de Gauss-Markov, siempre y cuando la esperanza de ui sea igual a 0. Pero ninguna suposición con respecto a el insesgamiento de la variable aleatoria ui.

c) Si no hay intersección en el modelo de regresión, las ui (= u(gorro)) estimadas no sumarán cero.

Verdadero, si no se encuentra la intersección en el modelo, la ecuación expuesta en el capítulo de sumatoria no existirá.

d) El valor p y el tamaño de un estadístico de prueba tienen el mismo significado.

Verdadero, una vez que se ha calculado el estadístico de prueba, resulta sencillo obtener el valor de p. El valor de p da la probabilidad exacta de obtener el estadístico de prueba estimado bajo la hipótesis nula.

e) En un modelo de regresión que contenga la intersección, la suma de los residuos es siempre cero.

Verdadero, la igualar la ecuación a cero, en la cual se encuentra presente la intersección del modelo, se cumple que la sumatoria es cero, además de encontrar los estimadores.

f) Si una hipótesis nula no es rechazada, es verdadera.

Falso, no quiere decir que es verdadera, ya que esta hipótesis anula o refuta, con objetivo de apoyar una hipótesis alternativa y cuando se le es útil.

g) Entre mayor sea el valor de σ², mayor será la varianza de B2(gorro), dada en 3.3.1

Falso, el tamaño de σ² será mayor, cuando mayor sea la sumatoria de las xi^2, al menos que esta sumatoria se encuentre constante, puede cumplirse que a mayor varianza de B2, el σ² será mayor.

h) Las medias condicional e incondicional de una variable aleatoria significan lo mismo.

Falso, la media condicional de la variable aleatoria depende de los valores tomados por otra variable, se encuentra condicionada por otra variable. Si las dos variables son independientes entre sí, pueden significar lo mismo.

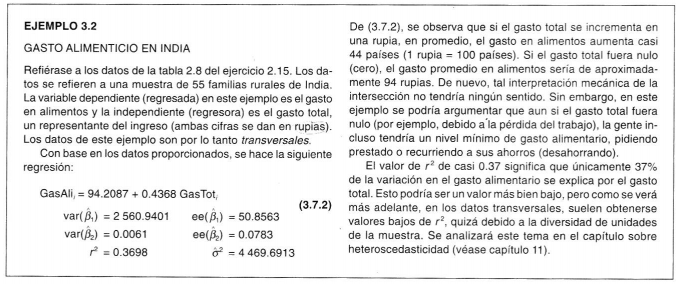
i) En una FRP de dos variables, si el coeficiente de la pendiente B2 es cero, la intersección B1 es un estimada por la media muestral Y(media).

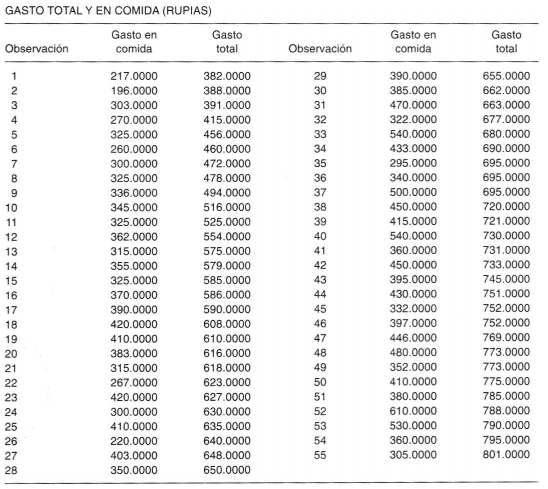
Verdadero, al B2 tener pendiente cero, la intersección B1 puede estimarse al restarle B2 a la media condicionada.

j) La varianza condicional, var (Yi|Xi)= σ² y la varianza incondicional de Y, var (Y)= σ²y, serían la misma si X no tuviera influencia en Y.

Verdadero, se hacen cero cuando X no tiene influencia en Y, la sumatoria de las y^2 cuando es igual a u^2, en este caso serian las mismas varianzas.

5.2- Constrúyase la tabla ANOVA a la manera de la tabla 5.4 para el modelo de regresión dado en (3.7.2) y pruébese la hipótesis de que no existe relación entre el gasto en alimentos y el gasto total en India.

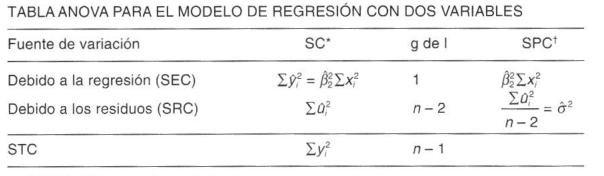


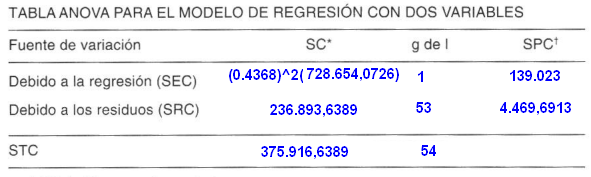


Y= Gasto de Alimentación.   
X= Gasto Total

= 728.654,0726  
**n**= 55 familias  
= 236.893,6389

Podemos construir la tabla ANOVA remplazando valores:





Fo= (139.023/4.469,6913)= 31,1013

R= Fo > F(1,53,0.01)   
Se rechaza la Hipótesis porque el F tabla es mayor al F calculado con un 0,01 por ciento de probabilidad. Es decir no existe relación entre el gasto de alimentación con el gasto Total

5.3-

a) Error estándar β2 = β2 estimado/ t = 0.6417/9.6536 = 0.0664.

tβ2 = β1/ ee = 0.7437/0.8355 = 0.8797.

b) ¿Cómo se interpreta el coeficiente 0.6416?

Por cada año adicional de escolaridad, el salario aumenta en un 0.64 dólares.

c) ¿Se rechazaría la hipótesis de que la educación no tiene efecto alguno sobre los salarios? ¿Cuál prueba se usaría? ¿Por qué? ¿Cuál es el valor p del estadístico de prueba?

La hipótesis de que la educación no tiene efecto alguno sobre los salarios se rechazaría, ya que la probabilidad de obtener el valor t (t=9.6536) es muy pequeña. Porque se encuentra por muy debajo de 0.001. Por lo tanto el P-value es casi cero.

d) Constrúyase la tabla ANOVA para este ejemplo y pruebe la hipótesis de que el coeficiente de la pendiente es 0. ¿Cuál prueba de utilizaría y por qué?

|  |  |  |  |
| --- | --- | --- | --- |
| Fuente de Variación | SC | G de l | SPC |
| SEC | 74.9389 | 1 | 74.9389 |
| SRC | 8.8454 | 11 | 0.8041 |
| STC | 83.7843 | 12 | - |

F= 74.9389/0.8041 = 93.1960

El valor p de F según la hipótesis de que no hay relación entre las dos variables es casi cero, es por esto que podemos rechazarla.

e) Supóngase que en la regresión que se acaba de dar, el valor R2 no se proporciona. ¿Se podría haber obtenido con base en otra información dada en la regresión?

R2 = t2/ [t2+(n-2)] = 9.65362/9.65362 + 11 = 0.8944.

5.4- Sea P^2 el verdadero coeficiente de correlación poblacional. Supóngase que se desea probar la hipótesis de que P^2 = 0. Explíquese verbalmente como se probaría esta hipótesis.

R: El resultado de la Hipótesis Nos indica que no existe correlación entre las 2 variables. Por lo que podemos decir que la covarianza entre las variables será cero, es decir su correlación es cero.

5.5-

5.6- La ecuación (5.3.5) puede ser escrita también como



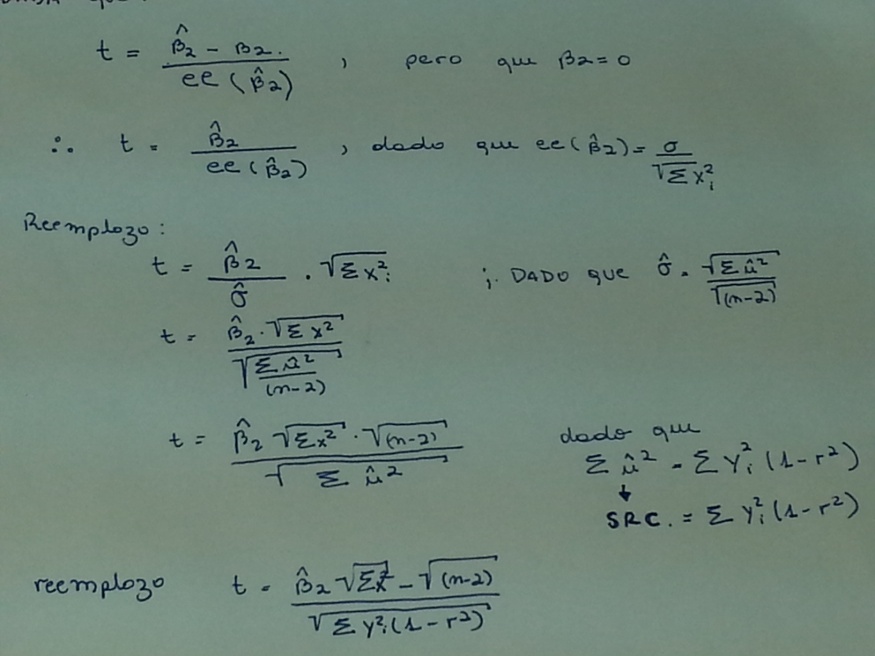
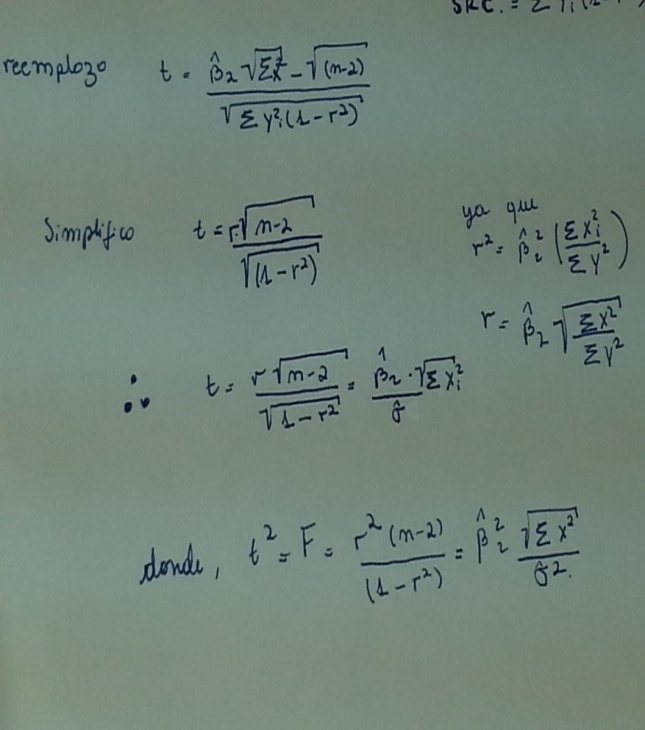
Es decir, la desigualdad débil puede ser reemplazada por la desigualdad fuerte. ¿Por qué?

No hay diferencia entre desigualdad débil y desigualdad fuerte, ya que según la teoría la probabilidad de que una variable aleatoria continua adquiera un valor específico es cero. Por tanto realizar la diferencia no tendría relevancia.

5.7

R.A. Fisher ha derivado la distribución muestral del coeficiente de correlación definido en (3.5.13).Si se supone que las variables X y Y tienen una distribución normal conjunta , es decir, si provienen de una distribución normal bivariada , entonces , bajo el supuesto de que el coeficiente de correlación poblacional p es cero , es posible demostrar que t = r √n-2/√1-r2 sigue la distribución *t* de Student con n-2 g de l. Demuéstrese que este valor t es idéntico al valor t dado en (5.3.2)bajo la hipótesis nula de que β2=0.Por lo tanto, establézcase que bajo la misma hipótesis nula F=t2.

Dado que:

5.8- Considérese el siguiente resultado de una regresión:

Yi=0,2033+0,6560Xi

ee= (0,0976)(0,1961)

r2=0,397 SRC=0,0544 SEC=0,0358

Donde Y=tasa de participación de la fuerza de trabajo(TPFT) de las mujeres en 1972 y X=TPFT de las mujeres en 1968. Los resultados de la regresión se obtuvieron de una muestra de 19 ciudades de Estados Unidos.

1. ¿Cómo se interpretaría esta regresión?

A medida que aumente la variable independiente Xi=TPFT de las mujeres de 1968, la variable dependiente TPFT de 1972 aumentara. Esto demuestra una relación directa entre ambas lo cual se considera debido a un aumento de la población femenina.

1. Pruébese la hipótesis H0:β2=1, contra H1: β2>1. ¿Qué prueba se utilizaría?, ¿por qué? ¿Cuáles son las suposiciones subyacentes de la(s) prueba(s) que se utilizará(n)?

T=(0,6560-1)/0,1961= -1,754207037

Comparar con t critico con un 0,05(5%) con 17 grado de libertad

1,754207037 > 1,740 Por ende se rechaza la Hipótesis nula es significativo. Es utilizada la prueba de significancia a través de su segundo criterio de decisión.

1. Supóngase que la TPFT para 1968 fue de 0,58(o 58%).Con base en los resultados de la regresión dado antes, ¿Cuál es la TPFT media en 1972? Establézcase un intervalo de confianza del 95% para la predicción de la media.

Media= 0,2033+0,6560(58%)=0,58378

Intervalo de confianza 95%:

0,58378+2,11< Valor<0,58378-2,1=2,69378<valor<-1,52622

Tomemos en cuenta que se desconoce el valor del error estándar.

1. ¿Cómo se comprobaría la hipótesis de que el termino de error en la regresión sobre la población esta normalmente distribuido? Muéstrense los cálculos necesarios.

No se puede determinar el siguiente ejercicio debido a la falta de datos.

5.9

La tabla 5.5 proporciona datos sobre el salario promedio de un maestro de escuela pública (el sueldo anual está en dólares) y el gasto en educación pública por alumnos (dólares ) para 1985 en los 50 estados y el distrito de Columbia en estados unidos.

A fin de averiguar si existe alguna relación entre el salario del maestro y el gasto por alumno en las escuelas públicas, se surgió el siguiente modelo: Sueldo = β0+β2 gasto+µi donde la variable sueldo es el salario del maestro y la variable gasto significa gasto por alumno.

a) Grafíquense los datos y trácese la recta de regresión

b) Supóngase con base en el inciso a) , que se decide estimar el modelo de regresión dado antes. Obténgase los estimados de los parámetros, sus errores estándar,r2 , la SRC y la SEC.

c) Interprétese la regresión ¿tiene sentido económico?

d) Establézcase un intervalo de confianza del 95% para β2 ,¿se rechazaría la hipótesis de que el verdadero coeficiente es 3.0?

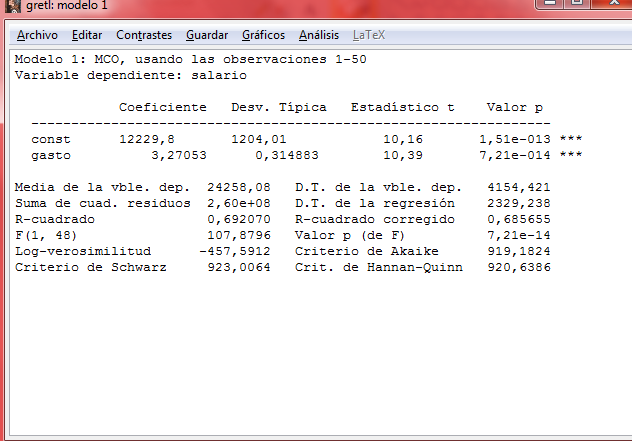
e) Obténgase el valor individual pronosticado y la media del sueldo , si el gasto por alumno es de $5000. También establézcanse intervalos de confianza para la cifra correspondiente al gasto dado antes.

f) ¿Cómo probaría la suposición de la normalidad del termino error?

j) Muestre la(s) prueba(s) utilizada(s)

Gráfico: de dispersión

B)



Modelo de regresión

Y=12229.8 + 3.27053 gasto

c) si el gasto por alumno es de un dólar. el salario medio se incrementa en alrededor de $ 3.27.y enfocándose en el término de intersección no tiene sentido económico viable

d) 3.27053 +2 (0.3117)≤β2 ≤ 3.27053 – 2 (0.3117)

No se rechaza la hipótesis nula y el coeficiente de la pendiente es igual = 3

5.10 Refiérase al ejercicio 3.20 para construir las tablas ANOVA y probar la hipótesis de que no existe ninguna relación entre la productividad y la compensación salarial real. ¿Es válido lo anterior para el sector de comercio y para el no agrícola?

|  |  |  |  |
| --- | --- | --- | --- |
| Fuente de Variación | SC | G de l | SPC |
| SEC | 38685.997 | 1 | 38685.997 |
| SRC | 4934.138 | 37 | 133.355 |
| STC | 43620.135 | 38 |  |

F=28685.997/133.355 = 290.0978

|  |  |  |  |
| --- | --- | --- | --- |
| Fuente de Variación | SC | G de l | SPC |
| SEC | 37887.455 | 1 | 37881.455 |
| SRC | 5221.585 | 37 | 141.129 |
| STC | 43109.04 | 38 |  |

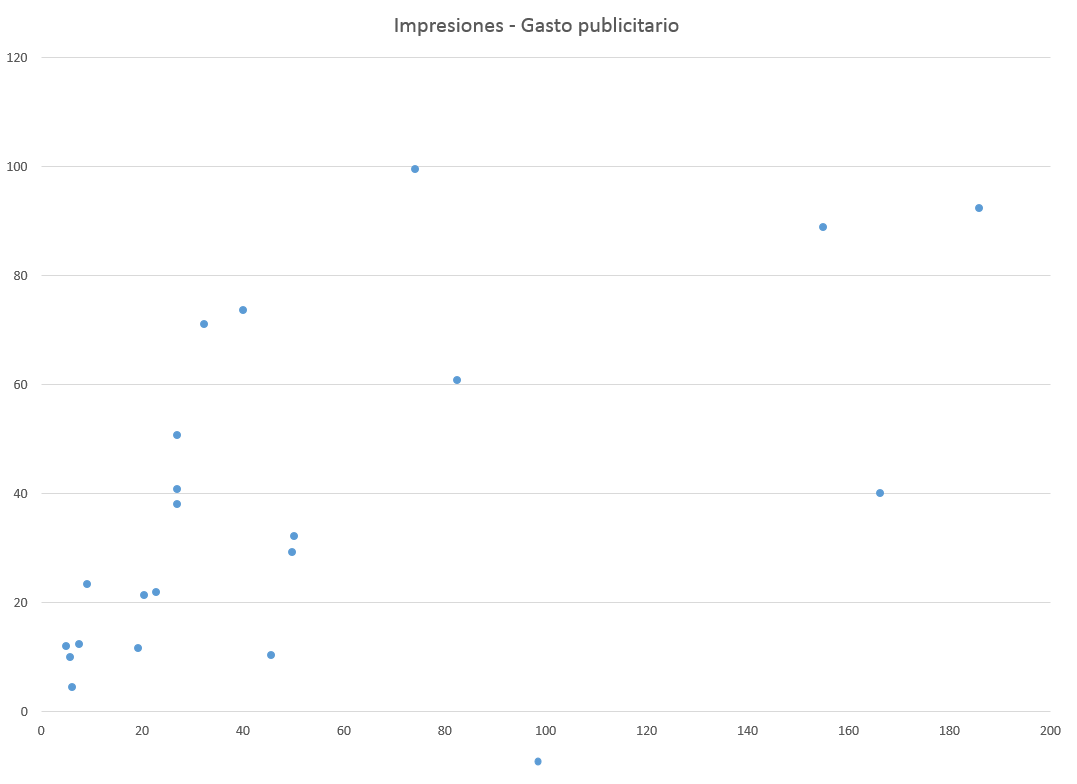
F = 37881.455/141.129 = 268.4172

Siguiendo la hipótesis de que no existe relación entre la productividad y la compensación salarial real. Y considerando que la distribución F está entre 1 numerador y 37 como denominador, la probabilidad es mucho menor a 0.01. Por lo tanto, nuestro F es casi cero, la hipótesis nula se rechaza.

5.11 Refiérase al ejercicio 1.7

a) Grafíquese los datos sobre impresiones en el eje vertical y el gasto publicitario en el horizontal, ¿Qué tipo de relación se observa?

|  |  |  |
| --- | --- | --- |
| Empresa | Impresiones,  Millones. | Gasto, Millones  de dólares de 1983 |
| Miller lite | 32,1 | 50,1 |
| Pepsi | 99,6 | 74,1 |
| Stroh's | 11,7 | 19,3 |
| Fed'l Express | 21,9 | 22,9 |
| Burger King | 60,8 | 82,4 |
| Coca Cola | 73,6 | 40,1 |
| McDonald's | 92,4 | 185,9 |
| MCI | 50,7 | 26,9 |
| Diet Cola | 21,4 | 20,4 |
| Ford | 40,1 | 166,2 |
| Levi's | 40,8 | 27 |
| Bud Lite | 10,4 | 45,6 |
| ATT/Bell | 88,9 | 154,9 |
| Calvin Klein | 12 | 5 |
| Wendy’s | 29,2 | 49,7 |
| Polaroid | 38 | 26,9 |
| Shasta | 10 | 5,7 |
| Meow Mix | 12,3 | 7,6 |
| Oscar Meyer | 23,4 | 9,2 |
| Crest | 71,1 | 32,4 |
| Kibbles 'N Blits | 4,4 | 6,1 |



Las 2 variables no son lineales. En el comienzo, a medida que aumenta el gasto en publicidad, las impresiones también lo hacen, pero poco a poco empiezan a disminuir.

b) ¿Sería apropiado ajustar un modelo de regresión de dos variables a los datos? ¿Por qué sí o porque no? Si la respuesta es negativa, ¿Qué tipo de modelo de regresión se ajustaría a los datos? ¿Se cuenta con las herramientas necesarias para ajustar dicho modelo?

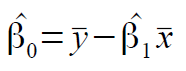
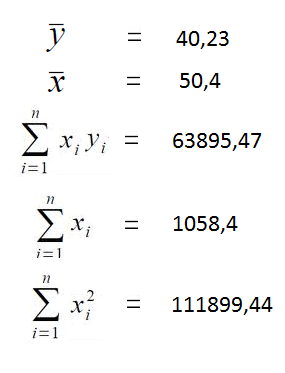
Como se puede ver, no sería apropiado adecuarse a un modelo de regresión lineal. No contamos con herramientas para encontrar un modelo adecuado.

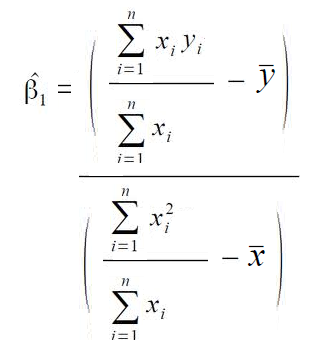
El siguiente modelo se puede ajustar a los datos:

Yi = β1 + β2 X2i + β3 X22i + ui.

Donde Y= impresiones; X= Gastos en Publicidad

Este es un modelo de regresión cuadrática con parámetros lineales.

c) Supóngase que no se grafican los datos y que simplemente se ajusta el modelo de regresión con dos variables a los datos. Obténgase los resultados usuales se la regresión.



Yi = 22.163 + 0.3631 Xi.

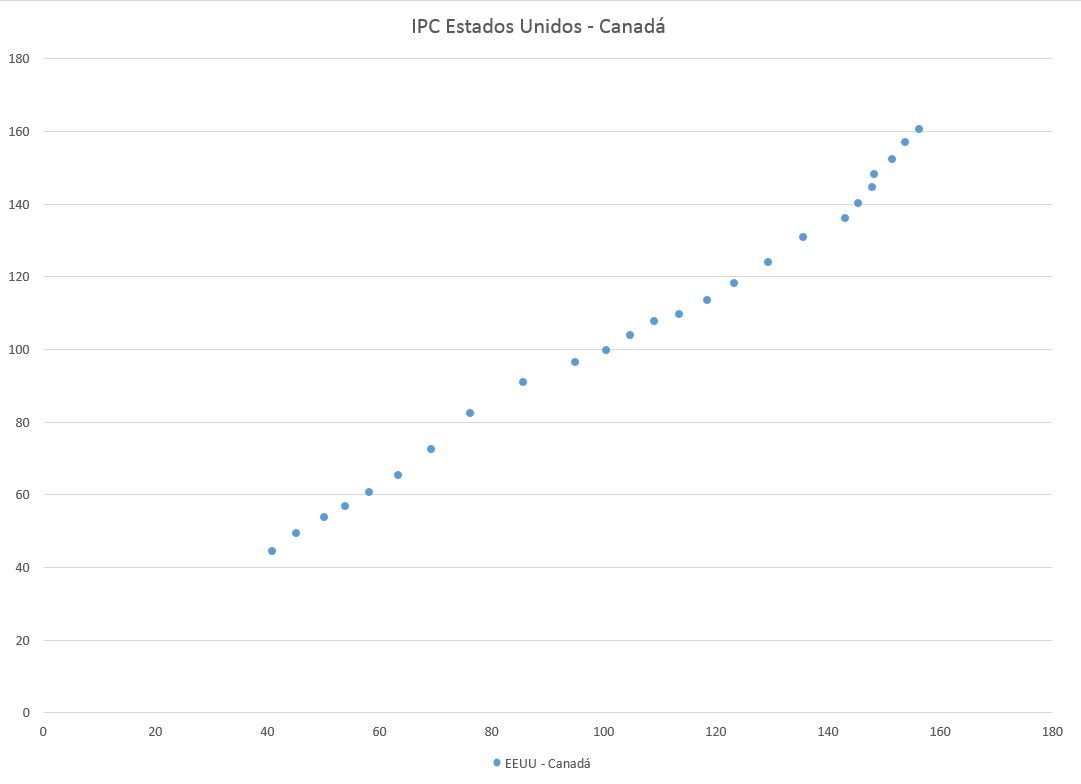
Ee= (7.089) (0.0971)

R2=0.424

5.12 Refiérase al ejercicio 1.1

a) Grafíquese el índice de precios al consumidor (IPC) estadounidense en relación con el canadiense ¿Qué revela la gráfica?

|  |  |
| --- | --- |
| Estados Unidos | Canadá |
| 44,4 | 40,8 |
| 49,3 | 45,2 |
| 53,8 | 50,1 |
| 56,9 | 53,9 |
| 60,6 | 58,1 |
| 65,2 | 63,3 |
| 72,6 | 69,2 |
| 82,4 | 76,1 |
| 90,9 | 85,6 |
| 96,5 | 94,9 |
| 99,6 | 100,4 |
| 103,9 | 104,7 |
| 107,6 | 109 |
| 109,6 | 113,5 |
| 113,6 | 118,4 |
| 118,3 | 123,2 |
| 124 | 129,3 |
| 130,7 | 135,5 |
| 136,2 | 143,1 |
| 140,3 | 145,3 |
| 144,5 | 147,9 |
| 148,2 | 148,2 |
| 152,4 | 151,4 |
| 156,9 | 153,8 |
| 160,5 | 156,3 |



La grafica realizada muestra que las tasas de inflación se comportan de la misma forma, es decir, sus comportamientos son similares y por lo tanto ambas variables de ambos países se mueven de manera conjunta, poseen una relación positiva.

b) Supóngase que se desea predecir el IPC de EEUU con base en el canadiense. Desarróllese un modelo apropiado.

Modelo 1: estimaciones MCO utilizando las 25 observaciones 1-25

Variable dependiente: y

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Variable* | *Coeficiente* | *Desv. típica* | *Estadístico t* | *valor p* |  |
| const | 6,21955 | 1,93254 | 3,2183 | 0,00381 | \*\*\* |
| x | 0,941239 | 0,0173566 | 54,2294 | <0,00001 | \*\*\* |

Media de la var. dependiente = 104,756

Desviación típica de la var. dependiente. = 36,5677

Suma de cuadrados de los residuos = 249,047

Desviación típica de los residuos = 3,29061

R2 = 0,99224

R2 corregido = 0,991902

Grados de libertad = 23

Log-verosimilitud = -64,208

Criterio de información de Akaike = 132,416

Criterio de información Bayesiano de Schwarz = 134,854

Criterio de Hannan-Quinn = 133,092

El modelo se ha desarrollado considerando los datos pertinentes de ambos países e insertándolos en un Excel para posteriormente obtener el modelo recién presentado mediante el programa Gretel.

c) Pruébese la hipótesis de que no existe relación entre ambos IPC. Utilícese α=5%. Si se rechaza la hipótesis nula, ¿significa que el IPC canadiense “condiciona” al IPC estadounidense? ¿Por qué sí o por qué no?

De acuerdo al modelo recién estimado se puede determinar que existe una relación positiva entre ambas variables es decir ambas aumentan o disminuyen de manera similar. Ante lo cual se puede desde ya rechazar la hipótesis nula de que no existe relación entre ambos IPC. Según el valor de P que entrega el modelo, se deja ver que la probabilidad de que se logre obtener el valor t es prácticamente cero. Además el valor P indica que la probabilidad de rechazar la hipótesis cuando esta es verdadera es demasiado baja, frente a lo cual se puede rechazar la hipótesis nula planteada en el apartado. No obstante, a pesar de que las variables se comportan de manera similar y claramente están relacionadas no se puede establecer que el IPC canadiense condicione al IPC estadounidense, ya que como se ha dicho en reiteradas ocasiones no se puede establecer si existe una relación de causalidad entre ambas variables a pesar de la relación que se pueda observar en el modelo.

5.13

a) Estímense las dos regresiones dadas en dichos ejercicios y obténganse los errores estándar .asi como los otros resultados usuales.

b) Pruébese la hipótesis de que las perturbaciones en los dos modelos de regresión están normalmente distribuidas

c) En la regresión del precio del oro .Pruébese la hipótesis de que B2 =1 es decir , que existe una relación uno a uno entre los precios del oro y el IPC (es decir el oro es una perfecta barrera.) ¿Cuál es la hipótesis nula que se está probando? ¿Cuál es el valor P?

d) Repitase el paso c) para la regresión del índice de la bolsa de New York (nyse) ¿La inversión en el mercado de valores representa una barrera pefecta contra la inflación? ¿Cuál es la hipótesis nula que se esta probando? ¿valor p?

e) entre el oro y el mercado de valores ¿Cuál inversión se elegirá? ¿en que se basaría la decisión?

a) Las dos regresiones son las siguientes:

Primera regresión=186,183+1,842x Sus respectivos errores estándar 125,403 para 186,183 y 1,215 para 1,842

Sus respectivos t 1,484para 186,183; 1,515 para 1,842. Donde r2=0,150

Segunda regresión= 102,060+2,129

Sus respectivos errores: 23,767 para 102,060; 0,230 para 2,129

Sus respectivos t= -4,294 para 102,060; 9,247 para 2,129.Donde r2=0,868

b) con los datos obtenidos no rechazamos ninguna de la hipótesis de normalidad

c) ya que el coeficiente de la pendiente en la regresión precio del oro no es estadísticamente diferente de cero, no tiene sentido para saber si es diferente de

(d) y (e) mediante el procedimiento habitual se obtiene.

T=2.129-1/0.230=4.91

este valor supera el valor del t critico de 2.160, por lo tanto rechazamos la hipótesis nula.y deducimos que el coeficiente estimado es en realidad > 1.

La elección de que inversión se elegiría, La inversión en el mercado de valores probablemente funciona como un escudo de protección contra la inflación sin duda fue una mejor protección contra la inflación que la inversión en oro.

5.14- a) Al observar la siguiente tabla podemos notar que todos los resultados son muy equivalente por lo que no se puede deducir cual es el mejor. Son estadísticamente significativos al 99%

b) No se puede deducir que agregado monetario es mejor a través de la resolución del r2.

c) No se pueden medir los resultados de la regresión de los datos de la tabla.