Universidad Francisco Marroquín Facultad de Ciencias Económicas

Econometría I Repaso Primer Parcial (vale por 2 hojas de trabajo)

Nombre: David Corzo Carné: 20190432

**Serie I:** Responda de forma clara y concisa las preguntas que se le presentan a continuación.

* 1. **¿Cuál es la diferencia entre una relación estocástica y una determinista?, presente un ejemplo**

El modelo es determinístico cuando se tiene precisión y certeza de los valores de los parámetros dados por el modelo mientras que el modelo es estocástico cuando los parámetros son variables aleatorias y lo que se intenta estimar se obtiene sin certidumbre. Ejemplos de relaciones determinísticas pueden ser leyes de termodinámica, leyes de enfriamiento de Newton, que siempre se cumplen con exactitud. Ejemplos de relaciones estocásticas son modelos poblacionales que se pueden predecir sin certidumbre, por ejemplo el hecho que hay personas naciendo se puede tomar en cuenta, pero a la misma vez se mueren y son variables aleatorias que no se pueden determinar con certidumbre.

* 1. **¿Por qué se hacen pruebas de significancia estadística en un modelo de regresión?**

Se utilizan para probar que las variables en cuestión explican el modelo bien y probar estadísticamente que son relevantes al modelo. Si no son relevantes las variables se retiran del modelo.

* 1. **“Un modelo econométrico con un R2 elevado, en donde sus coeficientes son altamente significativos, puede llevarnos a establecer causalidad entre la variable dependiente y las independientes” (verdadero o falso, justifique claramente su respuesta)**

Falso, un R^2 alto establece correlación alta, juntamente con los coeficientes altamente significativos. Pero aun que se puede defender que haya causalidad entre la variable dependiente e independiente, correlación no es causalidad.

* 1. **¿Mencione dos de las tres diferentes razones que se exponen para justificar el término de error en un análisis econométrico?**

1. Sólo incluir las variables significativas en la estimación del modelo de regresión.
2. Siempre las variables tienen error, ninguna es 100% exacta.
   1. **Explique con sus palabras ¿Qué características tiene un modelo de regresión lineal estimado por el método de mínimos cuadrados ordinarios?**

Un modelo de regresión lineal tiene una o varias variables independientes, una variable dependiente. El modelo de regresión lineal busca trazar una recta utilizando datos históricos, esta recta está compuesta por coeficientes betas y el intercepto inicial, el objetivo es intentar explicar la variable dependiente utilizando lo que sabemos de las dependientes. Se calcula R^2 junto con otras variables como error para ver qué tan bien explica el modelo los datos y qué tanto fallan.

**Serie II:** Resuelva los problemas que se le presentan a continuación, deje constancia de todo su procedimiento. No olvide escribir su respuesta final en lapicero para tener derecho a revisión.

1. **El siguiente modelo es una versión simplificada del modelo de regresión múltiple utilizado por Biddle y Hamermesh (1990) para estudiar el intercambio entre tiempo dedicado al sueño y dedicado al trabajo, así como ver otros factores que afectan el sueño:**



En donde la variable sleep mide la cantidad total de minutos que duerme la persona por la noche durante una semana, totwrk mide la cantidad total de minutos trabajados por semana, educ mide los años de educación formal y age representa la edad medida en años.

1. **Si los adultos intercambian sueño por trabajo, ¿cuál es el signo esperado para ?, explique**

Debería de ser negativo por que mientras más se duerme menos trabajo puede hacer, y por ello cuando se trabaja más se disminuye el sueño y cuando se duerme más se disminuye el trabajo.

1. **¿Qué signos cree que tendrán y ?, explique**

Deberían de ser negativos, con más edad se tiende a dormir más y mientras más educación implica que debe de tener menos tiempo de sueño, cuando se estudia y se incrementa la edad se disminuye el sueño por lo que para calcular el sueño se debe descontar la educación y la edad y por lo tanto deberían de ser negativos.

1. **Luego de estimar el modelo presentado se obtienen los siguientes resultados (los errores estándar se presentan en paréntesis):**



(2,000) (0.084) (6.47) (1.09)

n=20, ,

1. **Interprete cada uno de los coeficientes según el contexto del problema**

es el intercepto y se obtiene cuando todas las variables son cero, en este caso se dormiría 3,638.25 minutos a la semana. Los coeficientes que acompañan al resto de las variables significan que por cada unidad de cada una incrementa proporcionalmente al coeficiente que la acompañe. La expresión denota que por cada minuto que se trabaje la cantidad que se duerme disminuye en . La expresión denota que por cada minuto de educación disminuye el sueño en proporción de 11.13. La expresión implica que por cada unidad adicional de edad se incrementa 2.20 el sueño de las personas.

1. **¿Son significativos los coeficientes?, sea α=10%**

El t objetivo debe de ser mayor al t crítico, por lo tanto, cada coeficiente es significativo.

1. **Estime el R2 de la regresión e interprete el resultado**

Tener un cercano a 100% implica que el modelo explica bien el casi 100% de los datos, tener en este caso uno de 50.48% implica que el modelo no explica casi la mitad de los datos y por lo tanto no es tan preciso como como se quisiera, por lo tanto, no es significativo.

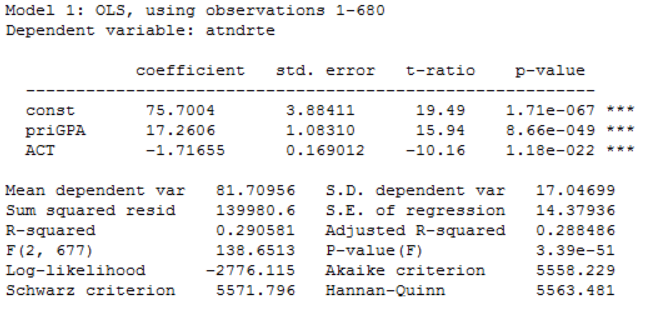
1. **Si una persona trabaja cinco horas más a la semana, ¿cuántos minutos se predice que disminuya *sleep*?**

Se duerme aproximadamente 44 minutos menos a la semana.

1. **En la tabla 1 se presentan los resultados de la estimación econométrica del siguiente modelo:**

En donde ***atndrte*** es el porcentaje de asistencia a clases del estudiante, ***priGPA*** es el promedio universitario acumulado del estudiante y ***ACT***es la calificación en el examen de admisión a la universidad.

**Tabla 1**

****

1.0831

75.700

-10.16

1. Complete la tabla 1
2. Interprete el significado de los coeficientes de la regresión.

denota el valor que surgiría si todas las demás variables fueran cero, en este modelo significa que si tiene cero en todas las demás variables asistiría 75.7%. La expresión denota que por cada incremento en unidades en el promedio general acumulado la asistencia aumento en proporciones de 17.2606. La expresión implica que por cada unidad extra en la calificación del examen de admisión la asistencia decaerá en proporciones de -1.71655.

1. ¿Son significativos los coeficientes?, utilice el valor p y un α=5%

Si son significativos puesto a que tiene tres asteriscos, por lo que implica que el valor-p es significativo incluso si significancia fuera 0.001, con significancia 0.05 también es significativo.

1. Estime la variación total de la variable dependiente
2. Un problema de interés para los funcionarios de salud (y para otros) es determinar los efectos que el fumar durante el embarazo tiene sobre la salud infantil. Una medida de la salud infantil es el peso al nacer; un peso demasiado bajo puede ubicar al niño en riesgo de contraer varias enfermedades. Ya que es probable que otros factores que afectan el peso al nacer estén correlacionados con fumar, deben considerarse. Por ejemplo, un nivel de ingresos más alto en general da como resultado el acceso a mejores cuidados prenatales y a una mejor nutrición de la madre. Una ecuación que reconoce estos factores es:



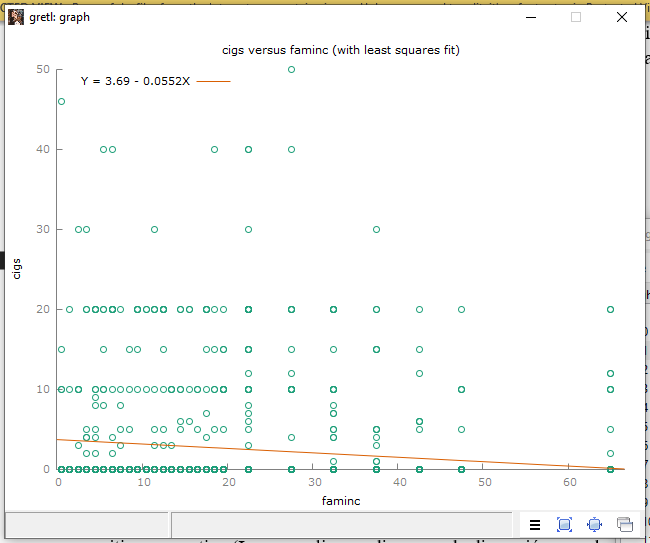
1. **¿Cuál es el signo más probable para B2?**

debería de ser positivo puesto a que mientras ocurren aumentos en el ingreso familiar se tiene más accesibilidad a comodidades como los de mejores cuidados prenatales y una mejor nutrición de la madre.

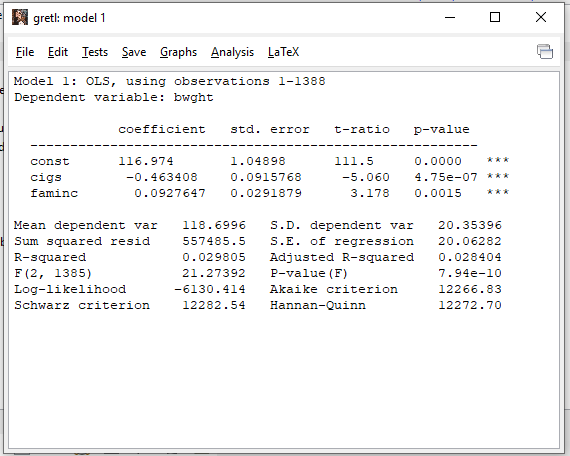
1. **¿Cree que cigs y faminc estén correlacionados? Explique por qué la correlación puede ser positiva o negativa (Luego realice un diagrama de dispersión usando Gretl).**

Cigs y faminc si debiesen de seguir una correlación inversa pues si tiene mayor ingreso usualmente se tiene más conciencia y educación sobre los efectos de fumar mientras las madres están embarazadas.

En la siguiente gráfica se aprecia que en efecto hay una correlación inversa.

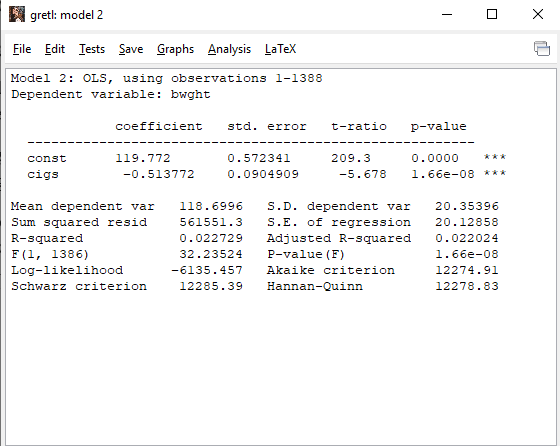
****

1. **Ahora, calcule la ecuación con y sin faminc utilizando los datos del archivo BWGHT.RAW. Dé los resultados en forma de ecuación incluyendo el tamaño de la muestra y la Rcuadrada (Utilice Gretl).**

****

1. **Explique sus resultados enfocándose en si el añadir faminc modifica de manera sustancial el efecto esperado de cigs sobre bwght (Compre los resultados de sus modelos, analizando R2 y la significancia de sus coeficientes. Evalúe si estos aumentaron o disminuyeron).**

La variable faminc no es significativa para el modelo puesto a que no impacta de manera significativa los resultados obtenidos de cigs y bwight, de hecho, la cigs cambia en proporciones de 0.05 y la constante cambian en .

****

1. Cargue la base de datos disponible en MiU en formato Gretl y estime un modelo en donde explique el consumo de pollo en función de las variables independientes proporcionadas (es posible construir distintos modelos, presente aquel que según su conocimiento es el “mejor”). A partir de los resultados obtenidos complete la siguiente tabla (vea la hoja en Excel disponible en MiU en donde se provee la explicación de las variables para interpretar con precisión los coeficientes estimados):

**Explique en términos generales los pasos que siguió para llegar al mejor modelo (cuando utilice más de una variable independiente es importante ver el comportamiento del R2 ajustado y no el R2). No olvide mostrar los conocimientos adquiridos en los cursos de economía para explicar la lógica detrás del desarrollo del modelo y en la interpretación de coeficientes.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable Independiente (nombre)** | **Coeficiente** | **Es significativo al 10% (1 si es cierto, 0 si es falso)** | | **Interprete el valor del coeficiente** | |
| **Constante** | 38.6472 | 1 | | Si todas las demás variables con cero, se consumiría 38.6472. | |
| X1 (ingreso) | 0.0108762 | 1 | | Por cada aumento en el ingreso de los consumidores el consumo de pollo aumenta en proporciones de 0.01. | |
| X2 (precio del pollo) | -0.541084 | 1 | | Por cada centavo adicional agregado a el precio del pollo, el consumo del pollo disminuye en proporcion de 0.541. | |
| X3 (precio del cerdo) | 0.174055 | 1 | | Por cada centavo adicional agregado al precio del cerdo, el consumo de pollo aumenta en proporciones de 0.17. | |
| X4 (precio de la carne) | N/A | 0 | | No se toma en cuenta puesto a que el coheficiente no es significativo. | |
| **Escriba el resultado del siguiente estadístico** | | | | | **Interprete el significado del R2 ajustado** | |
| R2 ajustado | | | 0.926651 | | El modelo explica bien 92% de los datos, es bastante significativo puesto a que R^2 es un valor cercano a 1. | |