

Taller 2 - Econometría I

Nicolás Ronderos PhD.

En todos los casos muestre sus salidas de STATA y comente sus resultados

1. Descargue la base de datos wage1.
 - a. Calcule estadísticas descriptivas de wage,educ,exper,tenure y numdep. Muestre sus resultados. Calcule el histograma de frecuencias de cada una de las variables. Interprete sus resultados.
 - b. Tabule y calcule las estadísticas descriptivas de las variables nonwhite, female y married. Interprete sus resultados.
 - c. Calcule el número de mujeres no blancas en la muestra.
 - d. Calcule el número de hombres blancos y casados de la muestra.
 - e. Calcule el número de mujeres blancas y solteras en la muestra.
 - f. Calcule la covarianza y la correlación entre wage,educ,exper,tenure y numdep.
2. Una pequeña empresa contrata a un consultor para predecir el valor de las ventas semanales de su producto si gasto en publicidad semanal aumenta a \$ 750. El consultor lleva un registro de cuánto gastó la empresa en publicidad por semana y las ventas semanales correspondientes durante los últimos seis meses. El consultor escribe: "Durante los últimos seis meses, el gasto semanal promedio en publicidad ha sido de \$ 500 y las ventas semanales promedio han sido de \$ 10,000. Según los resultados de una regresión lineal simple, predigo que las ventas serán de \$ 12,000 si se gastan \$ 750 por semana en publicidad".
 - a) ¿Cuál es la regresión simple estimada utilizada por el consultor para hacer esta predicción?
 - b) Dibuje un gráfico de la línea de regresión estimada. Busque los valores semanales promedio en el gráfico.
3. Un vendedor de gaseosas en los juegos de baseball de la Universidad de Costa Rica observa que se venden más refrescos cuanto más cálida es la temperatura en el momento del juego. Basado en 32

juegos en casa que cubren cinco años, el proveedor estima que la relación entre las ventas de refrescos y la temperatura es $\hat{y} = -240 + 8x$, donde y es el número de refrescos que vende y x la temperatura en grados Fahrenheit.

- a. Interprete los parámetros estimados. ¿Tienen sentido las estimaciones? ¿Por qué o por qué no?
- b. En un día en el que se pronostique que la temperatura a la hora del juego será 80F, prediga cuántos refrescos venderá el vendedor.
- c. ¿Por debajo de qué temperatura son cero las ventas previstas?
- d. Dibuje un gráfico de la línea de regresión estimada.

4. El modelo de valoración de activos de capital (CAPM) es un modelo importante en el campo de las finanzas. Explica las variaciones en la tasa de rendimiento de un activo en función de la tasa de rendimiento de un portafolio que consta de todas las acciones que cotizan en bolsa, lo que se denomina cartera de mercado. Generalmente, la tasa de rendimiento de cualquier inversión se mide en relación con su costo de oportunidad, que es el rendimiento de un activo libre de riesgo. La diferencia resultante se denomina prima de riesgo, ya que es la recompensa o castigo por realizar una inversión arriesgada. El CAPM dice que la prima de riesgo del activo j es proporcional a la prima de riesgo de la cartera de mercado. Es decir,

$$r_j - r_f = \beta_j(r_m - r_f)$$

Donde r_j y r_f son los rendimientos del activo j y la tasa libre de riesgo, respectivamente, r_m es el rendimiento del portafolio de mercado y β_j es el valor "beta" del activo j -ésimo. La beta de una acción es importante para los inversores, ya que revela la volatilidad de la acción. Mide la sensibilidad del retorno del activo j a la variación en todo el mercado de valores. Como tal, los valores de beta inferiores a 1 indican que la acción es "defensiva" ya que su variación es menor que el del mercado. Una beta mayor que 1 indica una "acción agresiva". Los inversores normalmente quieren conocer una estimación de la beta de una acción antes de comprarla. El modelo CAPM que se muestra arriba es el "modelo económico" en este caso. El "modelo econométrico" se obtiene al incluir una constante en el modelo (aunque la teoría lo diga debe ser cero) y un término de error. Esto es:

$$r_j - r_f = \beta_j(r_m - r_f) + u$$

a. Explique por qué el modelo econométrico anterior es un modelo de regresión simple como los discutidos.

b. En el archivo de datos `capm4.dat` hay datos sobre los rendimientos mensuales de seis empresas (Microsoft, GE, GM, IBM, Disney y Mobil-Exxon), la tasa de rendimiento de la cartera de mercado (MKT) y la tasa de rendimiento sobre el activo libre de riesgo (RISKFREE). Las 132 observaciones abarcan desde enero de 1998 hasta diciembre de 2008. Estime el modelo CAPM para cada empresa y comente sus valores beta estimados. ¿Qué firma parece más agresiva? ¿Qué empresa parece estar más a la defensiva?

c. Estime el modelo para cada empresa bajo el supuesto de que $\alpha_j = 0$. ¿Cambian mucho las estimaciones de los valores beta?

5. El archivo `stockton4.dat` contiene datos sobre 15,009 casas vendidas en Stockton, California, durante el período 1996–1998.

a. Grafique el precio de venta de las casas en función del área habitable para todas las casas en la muestra.

b. Estime el modelo de regresión: $sprice = \beta_0 + \beta_1 livarea + u$ para todas las casas en la muestra. Interprete las estimaciones y dibuje un boceto de la línea ajustada.

c. Estime el modelo cuadrático: $sprice = \beta_0 + \beta_1 livarea^2 + u$ para todas las casas en la muestra. ¿Cuál es el efecto marginal de 100 pies cuadrados adicionales de área habitable en una casa con 1500 pies cuadrados de área habitable?

d. En el mismo gráfico, trace las líneas ajustadas de los modelos lineal y cuadrático. ¿Cuál parece ajustarse mejor a los datos? Compare la suma de los residuos al cuadrado (SSE) de ambos modelos. ¿Cuál es menor?

e. Estime el modelo de regresión en (c) utilizando solo casas en lotes grandes. Repita la estimación para casas que no están en lotes grandes. Interprete las estimaciones. ¿Cómo se comparan los resultados?

f. Grafique el precio de venta de las casas en función de la edad de la casa (AGE). Estime el modelo lineal: $sprice = \beta_0 + \beta_1 age + u$ e interprete los coeficientes estimados. Luego, repita el ejercicio con el modelo log-lineal: $\ln(sprice) = \beta_0 + \beta_1 age + u$. Basándose en los gráficos y el ajuste visual de las líneas de regresión estimadas, ¿cuál de estos dos modelos preferiría? Explique.