Resumen capítulo 2 Wooldridge:

68- Principal problema de la regresión simple: supuesto de *ceteris paribus* (4) se rompe porque existen variables no contempladas correlacionadas con la variable dependiente y la independiente. Otra ventaja del análisis de regresión múltiple es que puede incorporar relaciones con formas funcionales muy generales.

69- Ejemplo de regresión múltiple con el salario por hora. como en la ecuación (3.1) aparece la experiencia de manera explícita, se podrá medir el efecto de la educación sobre el salario, manteniendo constante la experiencia. Con un análisis de regresión simple —en el cual exper forma parte del término del error— hay que suponer que la experiencia no está correlacionada con la educación, un supuesto cuestionable.

70- Permite generalizar relaciones funcionales entre las variables. Interpretación de los beta cuando hay una elevado al cuadrado (derivada parcial). Supuesto fundamental: . Esta condición significa que, para cualesquiera valores de x1 y x2 en la población, el promedio del efecto de los factores no observables es igual a cero. Como en la regresión simple, la parte importante de este supuesto es que el valor esperado de u es el mismo para todas las combinaciones de x1 y x2; que este valor común es cero no es ningún supuesto siempre que el intercepto se incluya en el modelo. Esto significa que los otros factores que afectan wage no están relacionados en promedio con educ y exper. Insesgadez.

Tabla

Descripción generada automáticamente

75- el intercepto es el valor predicho para y cuando y . Algunas veces es interesante hacer y , otras no tiene sentido.

77- El poder del análisis de regresión múltiple es que permite hacer en un ambiente no experimental, lo que en las ciencias naturales puede hacerse con experimentos controlados de laboratorio: mantener constantes otros factores.

81- Un hecho importante acerca de R2 es que nunca disminuye y, en general, aumenta cuando se agrega otra variable independiente a la regresión. Este hecho algebraico se debe a que, por definición, la suma de los residuales cuadrados nunca aumenta cuando se añaden regresores al modelo.

El hecho de que la R2 nunca disminuya cuando se agrega cualquier variable a la regresión hace de R2 un instrumento poco confiable para decidir si agregar una o varias variables al modelo. El factor que debe determinar si una variable explicativa pertenece a un modelo es si ésta tiene un efecto parcial distinto de cero sobre y en la población. En el capítulo 4, cuando se estudie la inferencia estadística, se mostrará cómo probar esta hipótesis. Se verá también que, usada de manera adecuada, la R2 permite probar si un grupo de variables es importante para explicar y. Por ahora, la R2 se usará como una medida de la bondad de ajuste de un modelo determinado.

83-En general, que la R2 sea pequeña indica que es difícil predecir con mucha exactitud resultados individuales para y. Pero no significa que los estimadores sean sesgados, es decir, puede interpetarse un impacto ceteris paribus.

84- Supuestos.

1. Linealidad de los parámetros
2. Muestreo aleatorio de tamaño n.
3. No hay colinealidad perfecta (no puede haber una relación lineal entre las variables dependientes). Imposibilita el análisis por MCO.
4. Media condicional cero. Se rompe cuando la relación funcional está mal especificada (si nos olvidamos algún término cuadrático o logarítmico). También si se omite algún factor (sesgo por variable omitida).
5. Homocedasticidad (para la mínima varianza del estimador). La heterocedasticidad invalida los errores estándar.

89- Incluir variables irrelevantes no afecta la insesgadez de los estimadores, sino su varianza.

Tabla

Descripción generada automáticamente

92- ejemplo de sesgo por variable omitida en rendimientos educación.

93- Hay que recordar que la correlación entre una sola variable explicativa y el error, por lo general, da como resultado que todos los estimadores de MCO sean sesgados. La única excepción a esto es cuando x1 y x2 tampoco están correlacionadas.

97- La varianza del estimador aumenta con la colinealidad. Pero disminuye a mayor muestra.



98- Un alto grado de correlación entre ciertas variables independientes puede ser irrelevante respecto a qué tan bien pueden estimarse otros parámetros del modelo.

99- no se puede especificar cuánta correlación entre las variables explicativas es “demasiada”.

100- trade-off sesgo-varianza.

El costo de incluir una variable irrelevante en el modelo es una varianza mayor del estimador. Con muestras grandes es siempre conveniente incorporar a la variable irrelevante.