Resumen Capítulo 2:

Capítulo 6 se tratan algunos otros temas sobre el análisis de regresión, incluyendo cuestiones como temas avanzados sobre formas funcionales, escalamiento de datos, predicción y bondad de ajuste.

22- Para establecer un modelo que “explique y en términos de x” hay que tomar en consideración tres aspectos. Primero, dado que entre las variables nunca existe una relación exacta, ¿cómo pueden tenerse en cuenta otros factores que afecten a y? Segundo, ¿cuál es la relación funcional entre y y x? Y, tercero, ¿cómo se puede estar seguro de que la relación entre y y x sea una relación ceteris paribus entre y y x (si es ese el objetivo buscado)?

23- a y se le conoce como la variable dependiente, la variable explicada, la variable de respuesta, la variable predicha o el regresando; a x se le conoce como la variable independiente, la variable explicativa, la variable de control, la variable predictora o el regresor. La variable u, llamada término de error, o perturbación en la relación, representa factores distintos a x que afectan a y. Por tanto, el cambio en y es simplemente multiplicado por el cambio en x. Esto significa que es el parámetro de la pendiente en la relación entre y y x, cuando todos los demás factores en u permanecen constantes; este parámetro es de interés primordial en la economía aplicada. El parámetro del intercepto , algunas veces llamado término constante, tiene también su utilidad, aunque es raro que tenga una importancia central en el análisis. La **linealidad** de la ecuación (2.1) implica que todo cambio de x en una unidad tiene siempre el mismo efecto sobre y, sin importar el valor inicial de x.

24- Ejemplo de ecuación sencilla para el salario (salario vs educación). Tema de no linealidad (rendimientos crecientes a la educación).

Problema de causalidad: *ceteris paribus*. Supuesto implícito de que el error (variable no observable) no está relacionada con la variable explicativa. La correlación no necesariamente es lineal, sino que puede ser una función, el supuesto crucial es que el valor promedio de u no depende del valor de x.

Supuesto de que la media del error es cero (2.5). Es necesario que exista la constante.

25- Ejemplo de salario, educación y capacidades innatas. El promedio de la capacidad debe ser el mismo sin importar el nivel de educación.

31- El nombre “mínimos cuadrados ordinarios” proviene del hecho de que estas estimaciones minimizan la suma de los residuales cuadrados.

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

32- Explica el por qué es al cuadrado: fácil de minimizar y porque es positivo.

Función de regresión muestral (FRM), funciones con sombrero (valores ajustados).

34- Ejemplo con salarios y educación: interpretación de coeficientes.

En algunos casos el análisis de regresión no se usa para determinar causalidad sino sólo para ver si dos variables

están relacionadas negativa o positivamente, de manera muy parecida a como se usa el análisis de correlación común.

40- es el cociente de la variación explicada entre la variación total; por tanto, se interpreta como la proporción de la variación muestral de y que es explicada por x.

41- Ahora vale la pena hacer hincapié en que una R-cuadrada muy baja no necesariamente significa que la ecuación de regresión de MCO sea inútil. Por ahora, hay que estar conscientes de que usar R-cuadrada como principal medida del éxito de un análisis econométrico puede acarrear problemas.

Hay que saber que las estimaciones de MCO cambian de maneras totalmente esperadas cuando cambian las unidades de medición de la variable dependiente o de la independiente.

42- Cuando se multiplica la variable dependiente por una constante c —lo que significa multiplicar cada valor de la muestra por c—, entonces las estimaciones de MCO del intercepto y de la pendiente también son multiplicadas por c. En general, si la variable independiente se divide o se multiplica por una constante distinta de cero, c, entonces el coeficiente de la pendiente de MCO se multiplica por c o se divide entre c, respectivamente. En general, cambiar sólo las unidades de medición de la variable independiente no afecta el intercepto. La bondad de ajuste del modelo no depende de las unidades de medición de las variables.

44- Expresión log-lineal para rendimientos de educación (Semielasticidad). Tabla de interpretación en la página 46.

47- Supuestos de MCO:

**Insesgamiento de los estimadores MCO**

1. Linealidad de los parámetros.
2. Muestreo aleatorio. Muestra aleatoria de tamaño n.
3. Variación muestral de la variable explicativa: no todos los valores muestrales son iguales.
4. Media condicional cero. Para todo valor de la variable explicativa, el valor esperado del error es cero. . Cuando se usa regresión simple, si u contiene factores que afectan a y y que están correlacionados con x, puede obtenerse una correlación espuria: esto significa que se encuentra una relación entre y y x que en realidad se debe a factores no observados que afectan a y y que resultan estar correlacionados con x.

**Mínima varianza**

1. Homocedasticidad. El error tiene la misma varianza para cualquier valor de la variable explicativa. .

53- Heterocedasticidad en la ecuación del salario: oportunidades.

58- **Regresión a través del origen**: estimador sesgado.