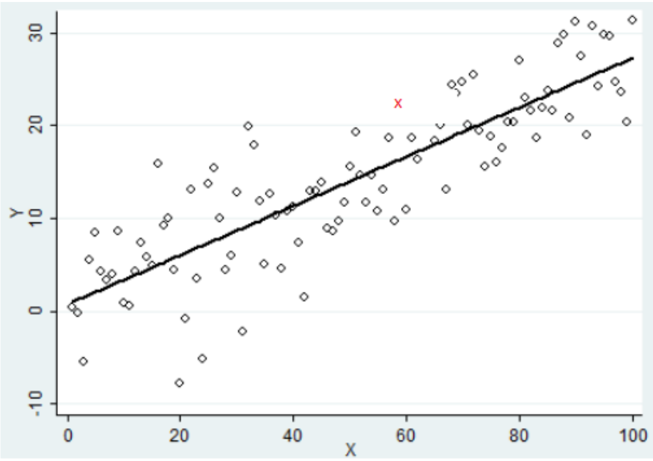
Material para las clases teóricas virtuales

Clase 2: Estimación del modelo lineal con 2 variables, método de mínimos cuadrados ordinarios.

1. Presentación en pdf para imprimir disponible en [este link]
2. Videos (son 4 videos en total):
   - **Video 1 – Regresión estimada y criterio de SRC**: [link al video 1]
     Se presenta la regresión estimada y se discute el criterio de minimización de la suma de los residuos cuadráticos.

     Preguntas guía para el video 1:

   a) En el gráfico a continuación se ilustra una nube de puntos y se marca con x una de las observaciones. Primero suponé que la línea continua representa la recta de regresión (la verdadera) $\alpha + \beta X_i$. ¿Qué representa la línea vertical entre el punto x y esa recta? Ahora suponé que la línea continua representa la recta de regresión estimada $\hat{\alpha} + \hat{\beta} X_i$. En este caso, ¿qué representa la línea vertical entre el punto x y esa recta? Relacionar esto con la siguiente frase: "podemos interpretar al error de estimación como una especie de estimador de μ_i" (ver página 5 en las filminas).



   b) Según el criterio de minimizar SRC ¿qué significa que la regresión estimada pase lo más cerca posible de los datos?

   c) Graficá un ejemplo donde SRC=0. ¿Podés pensar en algún ejemplo de la Economía que pueda representarse así?

   d) ¿Qué opinás sobre la posibilidad de utilizar la suma de los errores de estimación como medida de error global de estimación?

   - **Video 2 – Estimadores de MCO**: [link al video 2]
     Se plantea el problema de minimización de la SRC y se obtienen los estimadores de MCO a partir de las condiciones de primer orden.

     Preguntas guía para el video 2:

     a) Asegurate que podés obtener las ecuaciones (3) y (4) a partir de las CPO por tu cuenta.

b) Demostrá que (4´) y (4) son expresiones equivalentes. Te sugiero empezar por (4´), reemplazar por la definición de variables en desvíos y seguir con álgebra.

- **Video 3 – Propiedades algebraicas y R cuadrado**: link al video 3
  A partir de las condiciones de primer orden, se obtienen distintas propiedades de los estimadores de MCO. Se discute la intuición detrás de cada una. Se propone el R cuadrado como medida de bondad del ajuste de la regresión estimada por MCO y se discuten sus propiedades y utilidad.

  Preguntas guía para el video 3:
  a) A partir de las CPO, demostrá todas las propiedades algebraicas vistas en clase.
  b) Suponé un modelo sin ordenada al origen, es decir, $Y_i = \alpha + \beta X_i + \mu_i$ con $\alpha = 0$. Planteá las CPO de la minimización de la SRC para este modelo particular y explicá por qué en este caso no hay nada que garantice que la suma de los errores de estimación sea cero.
  c) Mostrá que la segunda propiedad algebraica implica que la covarianza entre los errores de estimación y la variable explicativa es cero.
  d) Usá la propiedad 4 (relación entre el coeficiente de regresión y el de correlación) para entender cómo los cambios en las unidades de medida afectan al coeficiente de regresión $\hat{\beta}$. Por ejemplo, pensá en ejercicios como este: tengo una muestra donde Y y X están medidas en pesos; estimo una regresión de Y en función de X por MCO y obtengo $\hat{\beta}$=2. Ahora reexpreso los valores de X en centavos. ¿Cuál es el nuevo valor de $\hat{\beta}$?
  e) Demostrá cómo se llega a la expresión de la descomposición de la STC de la página 27 de las filminas. ¿Por qué decimos que esta descomposición es válida si se estima por MCO?
  f) Mostrá que el R^2 vale 1 cuando hay una relación lineal exacta entre Y y X; mostrá que R^2 vale 0 si $\hat{\beta}$=0.
  g) Algo que no está en las filminas y te conviene estudiar/entender: en el modelo lineal con 2 variables, el R^2 es igual al cuadrado del coeficiente de correlación entre Y y X.
  h) ¿Por qué decimos que la regresión estimada por MCO maximiza el R^2?

- **Video 4 – Ejemplos empíricos**: link al video 4
  Se presentan dos aplicaciones: i) estimación de la relación entre rendimiento académico y tamaño de clase, y ii) estimación de la función de consumo Keynesiana.

  Preguntas guía para el video 4:
  a) A partir de la regresión estimada de la página 32, calculá el valor estimado del puntaje para un distrito escolar que tiene 25 alumnos por maestro en promedio.
  b) Suponé que se quiere reducir el tamaño de las clases para mejorar el desempeño. ¿Qué políticas se te ocurren?

c) Suponé una política que reduce el tamaño de clase de 25 a 15 alumnos por clase en un distrito escolar. Según las predicciones de nuestro modelo, ¿cuánto mejoraría el desempeño de los estudiantes? ¿Vale la pena implementar esta política? Atención: hay un error en las cuentas en el video (ver desde el minuto 6:20). Además, fijate que en lugar de hacer la cuentas como en el video, podrías directamente usar esta expresión $\Delta \widehat{Puntaje} = \hat{\beta}\Delta X_i$ reemplazando $\hat{\beta}$ por su valor estimado y ΔX_i por 10 (=25-15).

d) El R^2 del modelo da solo 0.049. ¿Cómo interpretás este número? ¿Esto significa que el modelo no sirve?

e) En el ejemplo de la función de consumo Keynesiana para la década de 1940 (ver página 44 de las filminas), dijimos que la guerra (años 1942-1945) afecta sistemáticamente al consumo, reduciendo su valor respecto al que habríamos esperado para esos niveles de ingreso. Entonces, si nuestro modelo lineal es $C_i = a + bI_i + \mu_i$, no es razonable pensar que E$(\mu_i) = 0$ para i=1942, 1943, 1944 y 1945. ¿Por qué? Además, fijate que en ese caso la regresión $a + bI_i$ no coincide con el valor esperado del consumo. Cuando veamos el modelo con más variables explicativas, esta discusión va a terminar de tener sentido.