INTRODUCCIÓN A LA ECONOMETRÍA

3er curso LE y LADE

Tema 2a

Dpto. de Econometría y Estadística (EA3)

UPV—EHU

2 El Modelo de Regresión Lineal (I). Especificación y Estimación.

2.1 Especificación del Modelo de Regresión Lineal General (MRLG).

Especificación del MRLG (1)

- **Objetivo:** Cuantificar la relación entre:
  - una variable Y y
  - un conjunto de K variables explicativas $X_1, X_2, \ldots, X_K$,
  - a través de un modelo lineal.
- **Punto de partida:**
  - un modelo lineal : $Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_K X_K + u$,
  - una muestra de datos de tamaño T: $Y_t, X_{1t}, X_{2t}, \ldots, X_{Kt},\ t = 1 \ldots T$, donde

$$Y_t = t\text{-ésima obs de } Y,$$

$$X_{kt} = t\text{-ésima obs de } X_k,\ k = 1, 2 \ldots K.$$

Especificación del MRLG (2)

- MRLG:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt} + u_t,\ t = 1, 2 \ldots T,$$

  cuyos elementos son (recuérdese):
  - Y: variable dependiente ,
  - $X_k,\ k = 1 \ldots K$: variables explicativas ,
  - β_0: intercepto,
    $\beta_k,\ k = 1 \ldots K$: coeficientes (parámetros a estimar),
  - u: error o perturbación (aleatoria no observable), que permite capturar:
    - variables no incluidas en el modelo,
    - comportamiento aleatorio de los agentes económicos,
    - errores de medición, etc.

MRLG en forma de observación

El modelo

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt} + u_t, \; t = 1, 2 \dots T,$$

supone para cada observación:

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \cdots + \beta_K X_{K1} + u_1 \\
Y_2 &= \beta_0 + \beta_1 X_{12} + \beta_2 X_{22} + \cdots + \beta_K X_{K2} + u_2 \\
&\dots\dots \\
Y_t &= \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_K X_{Kt} + u_t \\
&\dots\dots \\
Y_T &= \beta_0 + \beta_1 X_{1T} + \beta_2 X_{2T} + \cdots + \beta_K X_{KT} + u_T
\end{aligned}$$

MRLG en forma matricial (1)

o de otra manera en forma matricial:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_t \\ \dots \\ Y_T \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \cdots + \beta_K X_{K1} \\ \beta_0 + \beta_1 X_{12} + \beta_2 X_{22} + \cdots + \beta_K X_{K2} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_K X_{Kt} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \beta_0 + \beta_1 X_{1T} + \beta_2 X_{2T} + \cdots + \beta_K X_{KT} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_t \\ \dots \\ u_T \end{bmatrix}$$

MRLG en forma matricial (2)

- es decir:

$$\underset{(T \times 1)}{\underset{Y}{\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_t \\ \dots \\ Y_T \end{bmatrix}}} = \underset{(T \times K{+}1)}{\underset{X}{\begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{K1} \\ 1 & X_{12} & X_{22} & \dots & X_{K2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_{1t} & X_{2t} & \dots & X_{Kt} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_{1T} & X_{2T} & \dots & X_{KT} \end{bmatrix}}} \underset{(K{+}1 \times 1)}{\underset{\beta}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{pmatrix}}} + \underset{(T \times 1)}{\underset{u}{\begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_t \\ \dots \\ u_T \end{bmatrix}}}$$

- $$\boxed{Y = X\beta + u.}$$

↰

2.2 Supuestos Básicos (Clásicos) . Interpretación.

Supuestos Básicos del MRLG (1)

1. Supuestos sobre la relación:
   - El modelo está correctamente especificado:

$$X_k \text{ explica } Y \Leftrightarrow X_k \in \text{modelo}.$$

2. Supuestos sobre los parámetros:
   - son constantes a lo largo de la muestra,
   - aparecen de forma lineal (*es decir* una constante más coeficientes)
     - $Y_t = \beta_0 + \beta_1 X_t + u_t$
   - Nota: pero las vars $Y, X_1, X_2, \ldots$ pueden ser transformaciones:
     - $Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \beta_3 \dfrac{1}{X_t} + u_t$
     - $Y_t = A X_{1t}^{\beta_1} X_{2t}^{\beta_2} e^{u_t}$ (¿Por qué?)
     - ¿y éste? $Y_t = \beta_0 + \beta_1 \dfrac{1}{X_t - \beta_2} + u_t$
     - ¿y estos otros? $\ln Y_t = \beta_0 X_t^{\beta_1} u_t; \quad Y_t = \beta_0 X_t^{\beta_1} + u_t$
       $Y_t = \beta_1 X_{1t} + \beta_2 X_{1t} X_{2t} + u_t; \quad Y_t = \beta_0 + \beta_1 X_{1t}^{X_{2t}} + u_t$

Supuestos Básicos del MRLG (2)

3. Supuestos sobre las variables explicativas:
   a) $X_1, \ldots, X_K$, son cuantitativas y fijas (*es decir* no aleatorias).
   b) $X_1, \ldots, X_K$, son linealmente independientes:
   - $\nexists X_k | X_k =$ comb. lin. de otros (¿por qué?)

- <2>Ejemplos de casos no válidos :
  - $Y_t = \beta_0 + \beta_1 X_t + \beta_2 (2X_t + 3) + u_t$
  - $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 (X_{1t} + X_{2t}) + u_t$
- <2>Ejemplos de casos válidos:
  - $Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + u_t$
  - $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{1t} X_{2t} + u_t$

Supuestos Básicos del MRLG (3)

4. Supuestos sobre el término de perturbación:
   a) Media cero:
   $\mathsf{E}(u_t) = 0 \quad \forall t$ (¿no es obvio?).
   b) Homocedástico:
   $\mathsf{Var}(u_t) = \mathsf{E}(u_t^2) = \sigma_u^2 \; (= \sigma^2)$ const ($\forall t$).
   c) Serialmente incorrelacionado:
   $\mathsf{Cov}(u_t, u_s) = \mathsf{E}(u_t u_s) = 0 \quad \forall t \neq s$.
   d) Distribución normal$^{(*)}$:
   $u_t \sim \mathcal{N} \quad \forall t$. (* añadido)

- Supuestos 4*a*–4*d* conjuntamente:

$$u_t \sim \text{iid}\, \mathcal{N}(0, \sigma_u^2)$$

Supuestos Básicos en forma matricial (1)

- de 4*a*: Vector de medias:

$$\underset{(T\times 1)}{\mathsf{E}(u)} = \begin{bmatrix} \mathsf{E}(u_1) \\ \mathsf{E}(u_2) \\ \vdots \\ \mathsf{E}(u_T) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0_T$$

- de 4*b* y 4*c*: Matriz de Covarianzas:

$$\underset{(T\times T)}{\mathsf{E}(uu')} = \begin{bmatrix} \mathsf{E}(u_1^2) & \mathsf{E}(u_1 u_2) & \ldots & \mathsf{E}(u_1 u_T) \\ \mathsf{E}(u_2 u_1) & \mathsf{E}(u_2^2) & \ldots & \mathsf{E}(u_2 u_T) \\ \ldots & \ldots & \ldots & \ldots \\ \mathsf{E}(u_T u_1) & \mathsf{E}(u_T u_2) & \ldots & \mathsf{E}(u_T^2) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_u^2 & 0 & \ldots & 0 \\ 0 & \sigma_u^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \sigma_u^2 \end{bmatrix} = \sigma_u^2 I_T$$

Supuestos Básicos en forma matricial (2)

- de manera más compacta:

$$\boxed{\underset{(T\times 1)}{u} \sim \big(\underset{(T\times 1)}{0} \,,\, \underset{(T\times T)}{\sigma_u^2 I_T} \big)}$$

- plus 4*d*:

$$\boxed{\underset{(T\times 1)}{u} \sim \mathcal{N}\big(\underset{(T\times 1)}{0} \,,\, \underset{(T\times T)}{\sigma_u^2 I_T} \big)}$$

2.3a Mínimos Cuadrados Ordinarios (MCO) en el Modelo de Regresión Lineal Simple (MRLS).

MRLS: la FRP

- Con $K = 1 \quad \rightsquigarrow \quad Y_t = \beta_0 + \beta_1 X_{1t} + u_t,$

$$\text{(MRLS):} \quad Y_t = \alpha + \beta X_t + u_t. \qquad (1)$$

- Función de Regresión Poblacional (FRP):
  $\mathsf{E}(u_t) = 0 \quad \rightsquigarrow \quad$ *parte sistemática* o FRP:

$$\mathsf{E}(Y_t) = \alpha + \beta X_t$$

- Interpretación de los parámetros:
  - $\alpha = \mathsf{E}(Y_t | X_t = 0)$: Valor esperado de Y_t cuando la variable explicativa es cero.
  - $\beta = \dfrac{\partial \mathsf{E}(Y_t)}{\partial X_t} \simeq \dfrac{\Delta \mathsf{E}(Y_t)}{\Delta X_t}$: Incremento en el valor (esperado) de Y_t cuando $X \uparrow$ una unidad (*c.p.*).
- Objetivo: obtener estimaciones $\widehat{\alpha}$, $\widehat{\beta}$ de los parámetros desconocidos α, β en (1).

Función de Regresión Muestral (FRM)

- $\widehat{\alpha}$, $\widehat{\beta} \quad \rightsquigarrow \quad$ modelo estimado o FRM:

$$\widehat{Y}_t = \widehat{\alpha} + \widehat{\beta} X_t$$

- Interpretación de las estimaciones:
  - $\widehat{\alpha} = (\widehat{Y}_t | X_t = 0)$: Valor estimado de Y_t cuando la variable explicativa es cero.
  - $\widehat{\beta} = \dfrac{\partial \widehat{Y}_t}{\partial X_t} \simeq \dfrac{\Delta \widehat{Y}_t}{\Delta X_t}$: Incremento estimado en Y_t cuando $X \uparrow$ una unidad (*c.p.*).
- Nótese la diferencia: un estimador (una formula) frente a una estimación (un número).

Perturbaciones frente a Residuos

- Perturbaciones en FRP:

$$u_t = Y_t - \mathsf{E}(Y_t) = Y_t - \alpha - \beta X_t$$

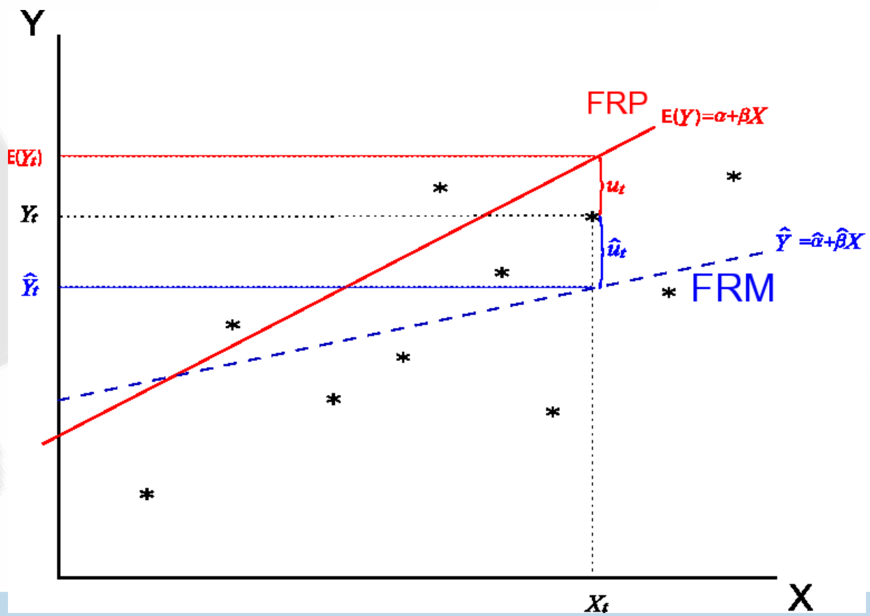
- Residuos en FRM:

$$\widehat{u}_t = Y_t - \widehat{Y}_t = Y_t - \widehat{\alpha} - \widehat{\beta} X_t$$

- Los Residuos son a la FRM

  lo que las perturbaciones son a la FRP.

MRLS: FRP y FRM

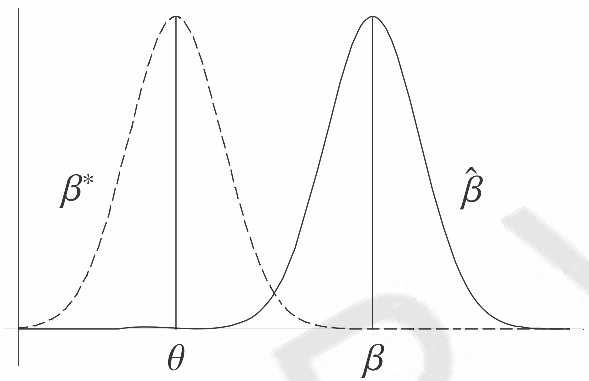


Estimación: Propiedades Deseadas (1)

Sea $\widehat{\beta}$ un estimador de $\beta\ldots$

Insesgadez:

$$\mathsf{E}(\widehat{\beta}) = \beta \quad \Leftrightarrow \quad \widehat{\beta} \text{ insesgado}$$

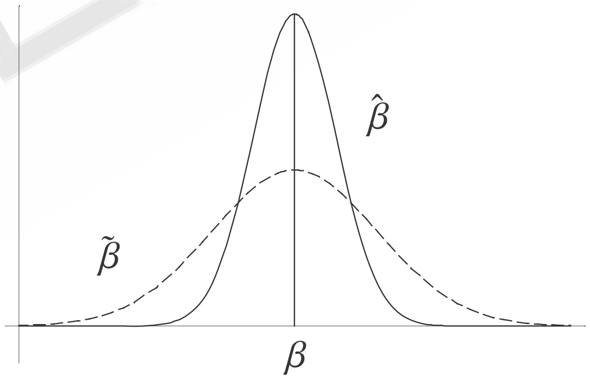


Estimación: Propiedades Deseadas (2)

Sean $\widehat{\beta}$ y $\widetilde{\beta}$ dos estimadores insesgados de $\beta\ldots$

Eficiencia relativa:

$$\mathsf{Var}(\widehat{\beta}) \leq \mathsf{Var}(\widetilde{\beta}) \quad \Leftrightarrow \quad \widehat{\beta} \text{ relativamente eficiente}$$



Estimación: criterio MCO

- <1>[] MRLS: $Y_t = \alpha + \beta X_t + u_t,$
- aplicar ajuste Mínimo-Cuadrático:

$$\min_{\alpha,\beta} \sum_{t=1}^{T} u_t^2 \quad \text{donde} \quad u_t = Y_t - \alpha - \beta X_t:$$

- Primeras derivadas:
  - $$\frac{\partial \sum u_t^2}{\partial \alpha} = 2\sum u_t \frac{\partial u_t}{\partial \alpha} = 2\sum u_t(-1)$$
  - $$\frac{\partial \sum u_t^2}{\partial \beta} = 2\sum u_t \frac{\partial u_t}{\partial \beta} = 2\sum u_t(-X_t)$$
- c.1$^{\text{er}}$o. (mínimo) $\Rightarrow$ primeras derivadas son cero:
  - $$\sum \widehat{u}_t = \sum (Y_t - \widehat{\alpha} - \widehat{\beta} X_t) = 0$$
  - $$\sum \widehat{u}_t X_t = \sum (Y_t X_t - \widehat{\alpha} X_t - \widehat{\beta} X_t^2) = 0$$

Estimación: Ecuaciones normales y EMC de α

- De las c.1$^{\text{er}}$o. anteriores:

$$\sum (Y_t - \widehat{\alpha} - \widehat{\beta} X_t) = 0$$
$$\sum (Y_t X_t - \widehat{\alpha} X_t - \widehat{\beta} X_t^2) = 0$$

- obtenemos las Ecuaciones Normales:

$$\left.\begin{array}{r} \sum Y_t = T\widehat{\alpha} + \widehat{\beta}\sum X_t \\ \sum Y_t X_t = \widehat{\alpha}\sum X_t + \widehat{\beta}\sum X_t^2 \end{array}\right\}$$

¡sistema de 2 ecuaciones con 2 incógnitas!!

- <1>Dividiendo la 1^a ec. normal por T:

$$\frac{1}{T}\sum Y_t = \frac{1}{T} T\widehat{\alpha} + \widehat{\beta}\frac{1}{T}\sum X_t$$

- Es decir:

$$\boxed{\widehat{\alpha}_{MCO} = \overline{Y} - \widehat{\beta}\,\overline{X}}$$

Estimación: Ecuaciones normales y EMC de β

- Sustituyendo $\widehat{\alpha}$ en la 2^a ec. normal:

$$\sum Y_t X_t = (\overline{Y} - \widehat{\beta}\overline{X})\sum X_t + \widehat{\beta}\sum X_t^2$$

- ...dividiendo por T y agrupando términos:

$$\frac{1}{T}\sum Y_t X_t = (\overline{Y} - \widehat{\beta}\overline{X})\frac{1}{T}\sum X_t + \widehat{\beta}\frac{1}{T}\sum X_t^2$$
$$\frac{1}{T}\sum Y_t X_t - \overline{Y}\,\overline{X} = \widehat{\beta}\left(\frac{1}{T}\sum X_t^2 - \overline{X}^2\right)$$

- ...y despejando la incógnita:

$$\widehat{\beta} = \frac{\frac{1}{T}\sum Y_t X_t - \overline{Y}\,\overline{X}}{\frac{1}{T}\sum X_t^2 - \overline{X}^2} = \frac{\frac{1}{T}\sum y_t x_t}{\frac{1}{T}\sum x_t^2} \quad \left[\begin{array}{c} \text{¿Por qué?} \\ \text{¿Por qué?} \end{array}\right] \quad \longrightarrow$$

- Es decir:

$$\boxed{\widehat{\beta}_{MCO} = \frac{\sum y_t x_t}{\sum x_t^2} = \frac{\text{Cov}(Y,X)}{\text{Var}(X)}}$$

Recuérdese: ¿varianzas y covarianzas?

- ¿varianza de los datos originales (no centrados)?

$$\text{Var}(X) = \frac{1}{T}\sum x_t^2 = \frac{1}{T}\sum (X_t - \overline{X})^2$$
$$= \frac{1}{T}\sum X_t^2 + \frac{1}{T}\sum \overline{X}^2 - \frac{2}{T}\overline{X}\sum X_t$$

$$\boxed{\frac{1}{T}\sum x_t^2 = \frac{1}{T}\sum X_t^2 - \overline{X}^2}$$

- ¿covarianza de los datos originales (no centrados)?

$$\text{Cov}(Y,X) = \frac{1}{T}\sum x_t y_t = \frac{1}{T}\sum (X_t - \overline{X})(Y_t - \overline{Y})$$
$$= \frac{1}{T}\sum X_t Y_t + \frac{1}{T}\sum \overline{X}\,\overline{Y} - \frac{1}{T}\overline{Y}\sum X_t - \frac{1}{T}\overline{X}\sum Y_t$$

$$\boxed{\frac{1}{T}\sum x_t y_t = \frac{1}{T}\sum X_t Y_t - \overline{X}\,\overline{Y}}$$

Ejemplo numérico : datos de prod de fresas

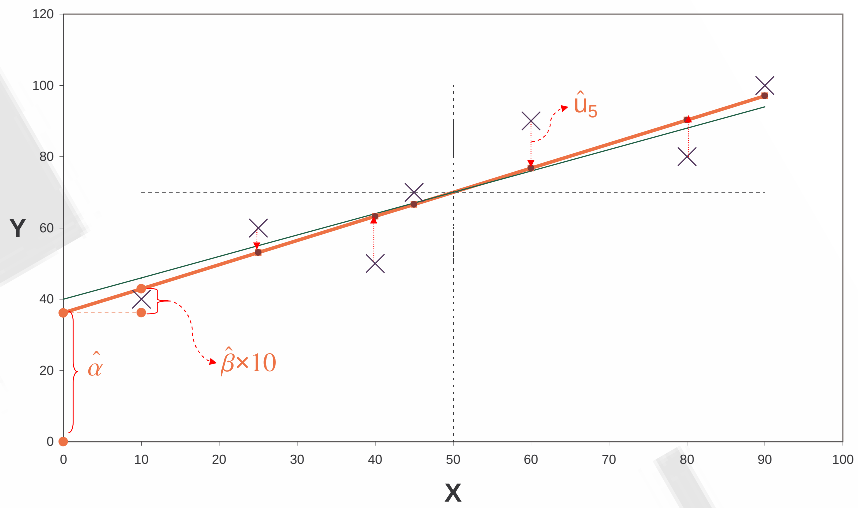
- Datos. . .
- Datos centrados o “en desviaciones” (desviaciones de las medias respectivas). . .
- Cuadrados y productos. . .

	Y	X	y	x	y^2	x^2	yx
	40	10	-30	-40	900	1600	1200
	60	25	-10	-25	100	625	250
	50	40	-20	-10	400	100	200
	70	45	0	-5	0	25	0
	90	60	20	10	400	100	200
	80	80	10	30	100	900	300
	100	90	30	40	900	1600	1200
Suma					2800	4950	3350
Media	70	50	0	0	400	707.14	478.57

$\widehat{\alpha} = 36{,}162 \; (= \overline{Y} - \widehat{\beta}\overline{X})$ $\qquad$ $\widehat{\beta} = 0{,}677 \; \left(= \dfrac{\text{Cov}(Y,X)}{\text{Var}(X)}\right)$

Se puede utilizar formulas basados en los datos originales. . . (Ejercicio: ¡Inténtalo!!

Ejemplo numérico: gráfico regresión fresas



2.4a Propiedades de la Función de Regresión Muestral.

Propiedades de los residuos y la FRM (1)

$$\widehat{\beta}_{\text{MCO}} \rightsquigarrow \widehat{\alpha}_{\text{MCO}} \rightsquigarrow \widehat{Y}_t = \widehat{\alpha} + \widehat{\beta}X_t \rightsquigarrow \widehat{u}_t = Y_t - \widehat{Y}_t$$

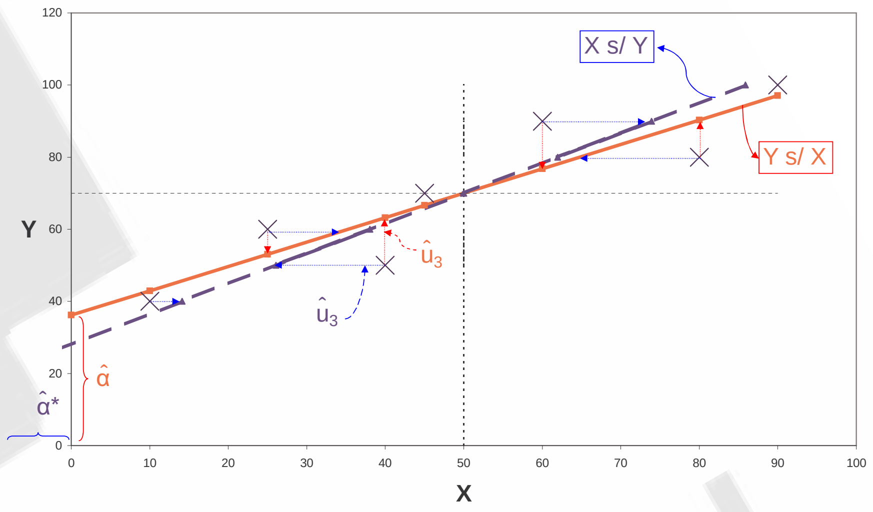
1. residuos suman cero: $\sum \widehat{u}_t = 0$
   Demo: directamente de las c.1^{er}o. □
2. $\overline{\widehat{Y}} = \overline{Y}$
   Demo: por def.: $\widehat{u}_t = Y_t - \widehat{Y}_t \rightsquigarrow \overline{\widehat{Y}} = \overline{Y} - \overline{\widehat{u}}$,
   pero $\overline{\widehat{u}} = \frac{1}{T}\sum \widehat{u}_t = 0$ (de prop 1) $\rightsquigarrow \overline{\widehat{Y}} = \overline{Y}$. □
3. la FRM pasa por el par de medias $(\overline{X}, \overline{Y})$:
   $\overline{Y} = \widehat{\alpha} + \widehat{\beta}\overline{X}$
   Demo: de $\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}$ (1^a ec. normal) □

Propiedades de los residuos y la FRM (2)

4. residuos ortogonales a v. expl. X: $\sum X_t \widehat{u}_t = 0$
   Demo: directamente de las c.1^{er}o.
5. residuos ortogonales a la parte explicada de Y: $\sum \widehat{Y}_t \widehat{u}_t = 0$
   Demo: $\sum (\widehat{\alpha} + \widehat{\beta} X_t)\, \widehat{u}_t =$

$$\widehat{\alpha} \underbrace{\sum \widehat{u}_t}_{=0 \text{ (de prop 1)}} + \widehat{\beta} \underbrace{\sum X_t \widehat{u}_t}_{=0 \text{ (de prop 4)}} = 0$$

Causalidad: Y sobre X frente a X sobre Y



Propiedades de los residuos y FRM (5)

8. <3> $\widehat{\alpha}_{\text{MCO}}$ y $\widehat{\beta}_{\text{MCO}}$ insesgados $\rightsquigarrow$ ¡valor esperado = valor verdadero!
9. <2>[]*Demo:*

▶

$$\widehat{\beta} = \frac{\sum y_t x_t}{\sum x_t^2}$$

$$\mathsf{E}(\widehat{\beta}) = \frac{1}{\sum x_t^2} \sum \underbrace{\mathsf{E}(y_t)}_{\beta x_t} x_t = \frac{1}{\sum x_t^2}\, \beta \sum x_t^2$$

$$\boxed{\mathsf{E}(\widehat{\beta}) = \beta}$$

▶

$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}$$

$$\mathsf{E}(\widehat{\alpha}) = \frac{1}{T} \sum \mathsf{E}(Y_t) - \mathsf{E}(\widehat{\beta})\overline{X}$$

$$= \frac{1}{T} \sum (\alpha + \beta X_t) - \beta\overline{X} = \alpha + \beta\overline{X} - \beta\overline{X}$$

$$\boxed{\mathsf{E}(\widehat{\alpha}) = \alpha}$$

2.5a Bondad de Ajuste: Coeficiente de Determinación (R^2).

Bondad de Ajuste: Coeficiente de determinación

- Descomposición de la Suma de Cuadrados:

$$\sum Y_t^2 = \sum(\widehat{Y}_t^2 + \widehat{u}_t^2 + 2\widehat{Y}_t\widehat{u}_t)$$
$$= \sum \widehat{Y}_t^2 + \sum \widehat{u}_t^2 \qquad \text{(de prop 5)}$$

- $$\sum Y_t^2 - T\overline{Y}^2 = \sum \widehat{Y}_t^2 - T\overline{\widehat{Y}}^2 + \sum \widehat{u}_t^2 \qquad \text{(de prop 2)}$$
- $$\boxed{\underset{(SCT)}{\sum y_t^2} = \underset{(SCE)}{\sum \widehat{y}_t^2} + \underset{(SCR)}{\sum \widehat{u}_t^2}}$$
- Definición de R^2:

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT}$$

$0 \leq R^2 \leq 1$ (¿Interpretación en términos de la varianza total??)

No intercepto $\rightsquigarrow$ inválido R^2

- <1>[] MRLS: $Y_t = \beta X_t + u_t$,
- aplicar el ajuste Mínimo-Cuadrático:

$$\min_{\beta} \sum_{t=1}^{T} u_t^2 \quad \text{donde} \quad u_t = Y_t - \beta X_t :$$

- Primeras derivadas:

$$\frac{\partial \sum u_t^2}{\partial \beta} = 2\sum u_t \frac{\partial u_t}{\partial \beta} = 2\sum u_t(-X_t)$$

- c.1ero. (mínimo) $\Rightarrow$ primera derivada $=$ cero:

$$\sum \widehat{u}_t X_t = \sum (Y_t X_t - \widehat{\beta} X_t^2) = 0$$

- !‘ $\nexists$ 1^a ecuación!! $\rightsquigarrow$ $\begin{cases} \sum \widehat{u}_t \neq 0, \\ \overline{\widehat{Y}} \neq \overline{Y}, \end{cases}$ $\rightsquigarrow$ R^2inválido (¿Por qué?)

Relación de R^2 con el coef de correlación

$$R^2 = \frac{\frac{1}{T}\sum \widehat{y}_t^2}{\frac{1}{T}\sum y_t^2} = \frac{\frac{1}{T}\sum(\widehat{\beta}x_t)^2}{\frac{1}{T}\sum y_t^2} = \frac{\widehat{\beta}^2\frac{1}{T}\sum x_t^2}{\frac{1}{T}\sum y_t^2}$$
$$= \widehat{\beta}^2 \frac{\text{Var}(X)}{\text{Var}(Y)} = \frac{\text{Cov}(Y,X)^2}{\text{Var}(X)^2}\frac{\text{Var}(X)}{\text{Var}(Y)}$$
$$= \frac{\text{Cov}(Y,X)^2}{\text{Var}(X)\text{Var}(Y)}$$

$R^2 = r_{X,Y}^2$

Ejemplo numérico: datos de prod fresas (cont)

- recuérdese datos y cálculos previos...
- hacer lo mismo para valores ajustados...
- ahora calcular R^2...

	y^2	$\widehat{Y}$	$\widehat{y}$	$\widehat{y}^2$	$\widehat{u}$	$\widehat{u}^2$
	900	42.92	-27.07	732.82	-2.92	8.58
	100	53.08	-16.91	286.25	6.91	47.87
	400	63.23	-6.76	45.80	-13.23	175.09
	0	66.61	-3.38	11.45	3.38	11.45
	400	76.76	6.76	45.80	13.23	175.09
	100	90.30	20.30	412.21	-10.30	106.15
	900	97.07	27.07	732.82	2.92	8.58
Media	400	70	0	323.88		
Suma	2800			2267.17		532.82
	SCT			SCE		SCR

$R^2 = 0{,}8097$ $\left(= \frac{\text{SCE}}{\text{SCT}} = 1 - \frac{\text{SCR}}{\text{SCT}}\right)$

(Ejercicio: ¿Cómo compara esto con Corr(X,Y)? ... ¡Inténtalo!!)