INTRODUCCIÓN A LA ECONOMETRÍA

3er curso LE y LADE

Dpto. de Econometría y Estadística (EA3)

UPV—EHU

1 Introducción

1.1 Definiciones. Elementos de Econometría

Introducción: Definiciones

ECONOMETRÍA

- etimológica:
  $o\acute{\iota}\kappa o\varsigma$ [oíkos], 'hogar',
  y $\nu\acute{o}\mu o\varsigma$ [nómos], 'reglas'
  por lo tanto economía $\rightsquigarrow$ 'administración doméstica',
  $+$ $\mu\varepsilon\tau\rho\acute{\omega}$ [metró], 'medir'.
  Economía $+$ Medición
- aditiva:
  Ciencia social que aplica
  Teoría económica, Matemática e Inferencia Estadística
  en el análisis de fenómenos económicos (Goldberger(1964)).
- utilitaria: El arte del económetra $=$ definir el modelo apropiado $+$ encontrar el procedimiento estadístico óptimo
  $\rightsquigarrow$ económetra $\neq$ estadístico;
  $\cdots +$ sólida formación en Economía (Malinvaud(1963)).

Introducción: Definiciones

- sencilla: aplicación de métodos estadísticos a datos económicos (Maddala(1977)).
- concisa: determinación empírica de leyes económicas (Theil(1971)).
- AFG(2004): La econometría trata de
  - formular (o especificar),
  - cuantificar (o estimar),
  - validar (o contrastar),

  relaciones entre variables económicas.

Introducción: 3 Elementos:

- **TEORÍA ECONÓMICA:**
  se encarga de
  - (general:) análisis de la economía
  - (específico:) relaciones entre variables económicas
- **DATOS:**
  cuantificar NO es uno de los objetivos de la T^a Económica
- **ESTADÍSTICA:**
  proporciona la estructura básica de métodos de procesamiento de datos para:
  - (estimar:)
    cuantificar relaciones entre variables de forma apropiada.
  - (contrastar:)
    validar resultados de acuerdo con ciertos estándares ya establecidos.

1.2 Concepto y ejemplo de modelo: Desde el modelo económico al modelo econométrico.

Elemento 1: T^a Económica: modelo básico

- - **Caso:** gerente de empresa o director de ventas,
  - **Interés:** saber la relación entre sus ventas y sus precios.
- - **Lógica económica básica:** ventas como función del precio $\rightsquigarrow$ modelo económico básico:

$$\underset{ventas}{V} = f(\underset{\underset{(-)}{precio}}{p})$$

$f(\bullet)$ es una función genérica
(T^a Ec: $f(\bullet) =$ fn inversa $\rightsquigarrow$ ventas $\uparrow$ si precio $\downarrow$.)

Elemento 1: T^a Económica: vars adicionales

- **Lógica económica adicional:**
  ventas dependen de
  - situación empresas rivales (p.ej. precio competencia)
  - condiciones del mercado (p.ej. ciclo económico)
- **Modelo completo:**

$$\underset{\text{ventas}}{V} = f\Big(\underset{\substack{\text{precio}\\(-)}}{p}\;,\;\underset{\substack{\text{precio competencia}\\(+)}}{pc}\;,\;\underset{\substack{\text{ciclo}\\(+)}}{c}\Big)$$

- NOTA:
  modelo económico propuesto $\equiv$ resumen de ideas,
  pero nada nuevo para el gerente;
  él necesita un modelo específico para su empresa
  $\rightsquigarrow$ como sus ventas**responden** a **sus** precios.

Elemento 2: Datos:

- Información específica:
  el gerente tiene **información** sobre:
- - sus ventas y sus precios **(datos cuantitativos)**
  - precios de la competencia **(datos cuantitativos)**
  - momento cíclico **(datos cualitativos)**
- *p.ej.*:

fechas	Ventas	precios	p.comp	ciclo
ene 80	1725	12,37	11,23	alto
feb 80	1314	11,25	10,75	alto
abr 95	1234	13,57	14,5	bajo
⋮	⋮	⋮	⋮	⋮

y todo esto todos los meses hasta diciembre de 2004.

Elemento 2: Datos: modelo específico

- modelo específico para los datos existentes:

$$V_t = f(p_t, pc_t, c_t), \qquad t = 1980{,}1, \ldots, 2004{,}12$$

donde el subíndice t indica el periodo o momento de la relación.

- hasta ahora:
  - modelo económico: resumen de ideas generales sobre la relación
  - datos: o información específica sobre las diferentes variables

  - **¿Cómo poner juntos estos elementos. . . ????**

E2: ¿modelo (genérico) $+$ datos (especifico)?:

- **A:** supuestos sobre $f(\bullet)$; *p.ej.*: relaciones lineales.
  El modelo será entonces:

$$V_t = \beta_0 + \beta_1 p_t + \beta_2 pc_t + \beta_3 c_t, \qquad t = 1980{,}1, \ldots, 2004{,}12$$

- $\beta \ldots =$ parámetros o coeficientes :
  p.ej. β_1 responde a la pregunta:
  ¿ en cuánto cambian las ventas si el precio cambia en una unidad monetaria ?
  $\rightsquigarrow$ políticas de precios, decisiones de producciones etc. para la empresa.

- **B:** indicadores:
  asignar valores cuantitativos a variables cualitativas (como Ciclo): *p.ej.* sustituir por indicador tal como Índice de Producción Industrial.

E2: ¿Modelo $+$datos?: perturbaciones aleatorias

- Después de esto el modelo expresa una relación cuantitativa entre las variables:

$$\begin{aligned} 1725 &= \beta_0 + 12,37\beta_1 + 11,23\beta_2 + 101,7\beta_3 \qquad \text{(1980.Ene)} \\ 1314 &= \beta_0 + 11,25\beta_1 + 10,75\beta_2 + 97,3\beta_3 \qquad \text{(1980.Feb)} \\ \vdots &= \vdots \end{aligned}$$

- **NOTA:** . . . ¿ diferente relación para cada mes ??? . . .
- **C:** termino perturbación;

- volver al modelo *económico* genérico:
  $\Rightarrow$ comportamiento estable entre variables
  $\Rightarrow$ comportamiento “promedio” reflejado en los datos
  $\Rightarrow$ añadir termino u_t para cubrir pequeña discrepancias. . .

E2: ¿Modelo+datos?: interpretación

- El modelo econométrico será finalmente:

$$V_t = \underbrace{\beta_0 + \beta_1 p_t + \beta_2 pc_t + \beta_3 c_t}_{\text{(“influencias” importantes y sistemáticas)}} + \underbrace{u_t}_{\text{(término de perturbación aleatoria)}}$$

- Interpretación de u_t:
  - $\Rightarrow$ efectos que afectan las ventas ligeramente en cada periodo pero no están reflejados explícitamente en el modelo.
  - $\Rightarrow$ pequeñas discrepancias en los datos.
  - $\Rightarrow$ efectos no sistemáticos $\equiv$ más erráticos.
  - $\Rightarrow$ **variable aleatoria** con cierta ley de probabilidad (*p.ej.*: dn Normal).

Elemento 3: Estadística:

- Modelo contiene una variable aleatoria
  $\rightsquigarrow$ procedimientos **estadísticos** que garanticen buenos resultados:

$\Rightarrow$ para estimar valor numérico de los coeficientes,

$\Rightarrow$ para contrastar la validez de la relación,

- el modelo estimado
  - no será un modelo genérico
  - sino un modelo específico para la empresa
- ofrecerá al gerente
  información específica para tomar decisiones.

1.3 El Modelo Econométrico. La Perturbación o término de Error.

Características Básicas: notación datos

Modelo econométrico más general con K variables:

- para datos de series temporales:

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

- o, para datos transversales (sección cruzada):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_K X_{Ki} + u_i, \quad i = 1, 2, \ldots, N.$$

- o, para datos de panel:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \cdots + \beta_K X_{Kit} + u_{it}, \quad \begin{cases} i = 1, 2, \ldots, N; \\ t = 1, 2, \ldots, T. \end{cases}$$

Características Básicas: notación vars

- Y: la variable que queremos *explicar*:
  v dependiente, v explicada, v endógena o regresando.
- $X_1, X_2 \ldots X_K$: variables que *explican* la variable Y:
  v explicativas, v independientes, v exógenas o regresores.
- $\beta_k, (k = 1 \ldots K)$: *constantes desconocidas* que determinan la relación entre las variables:
  parámetros o intercepto y coeficientes.
  $\widehat{\beta}_k$ es el coeficiente *estimado* .
- u: variable que recoge *otros efectos no importantes* presentes en los datos:
  perturbación aleatoria o término de error.

Diferencias Básicas con el modelo económico

Presencia de una perturbación aleatoria que

- refleja el comportamiento errático:

$$Y_t = \underbrace{\beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt}}_{\text{parte sistemática}} + \underbrace{u_t}_{\text{parte no sistemática o aleatoria}} \qquad t = 1, 2 \ldots T.$$

- tiene media cero:

$$E(Y_t) = E(\beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt}) + \underbrace{E(u_t)}_{=0} \quad t = 1, 2 \ldots T.$$

- por lo tanto parte sistemática $\equiv$ media comportamiento de Y.
- otros supuestos sobre u (hipótesis básica, etc.)
  $\rightsquigarrow$ comportamiento probabilístico en casos diferentes
  $\rightsquigarrow$ herramientas estadísticas $\rightsquigarrow$ Métodos Econométricos.

Clasificación de modelos econométricos

Diferentes enfoques:

- considerando los tipos de datos:
  - modelos de Series temporales.
  - modelos de Sección cruzada.
- considerando el periodo de observación:
  - **M estático.:** Vars medidas en el mismo momento.
  - **M dinámico.:** Vars referidas a diferentes periodos:
    p.ej. $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{1,t-1} + \beta_3 X_{2,t-1} + u_t$
- considerando el número de relaciones:
  - **Modelos uniecuacionales:** una única relación o ecuación.
  - **Modelos de ecuaciones simultáneas:** más de una ecuación.

etc.

1.4 Etapas en la elaboración del modelo. Utilidad del modelo.

Etapas en la elaboración del modelo

0. **Selección.** Resumir la teoría de interés:
   - elegir la variable a explicar: Y.
   - elegir la relación general: $Y = f(X)$.
1. **Especificación.** Esbozar modelo econométrico consistente con la teoría:
   - elegir las variables explicativas: $X_1 \dots X_K$.
   - elegir la forma funcional: *p.ej.* $f(\cdot) \equiv$ lineal.
   - elegir el comportamiento probabilístico (distribución) de la perturbación aleatoria: u, *p.ej.* $u_t \sim$ iid$\mathcal{N}(0, \sigma^2)$.

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_K X_K + u.$$

Etapas en la elaboración del modelo

2. **Estimación.** Cuantificar parámetros desconocidos de acuerdo con los datos existentes:
   - encontrar datos para las variables:
     $Y_t, X_{1t}, \ldots, \ldots, X_{Kt}$ for $t = 1, \ldots, T$.
   - elegir el método estadístico apropiado, *p.ej.* MCO:

$$Y_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_{1t} + \cdots + \widehat{\beta}_K X_{Kt} + \widehat{u}_t, \quad t = 1, 2 \ldots T.$$

3. **Validación.** Evaluar si el modelo representa correctamente el problema inicial:
   - inferencia estadística sobre hipótesis.
   - modelo no apto $\rightsquigarrow$ volver a la fase de especificación.

Utilización del modelo econométrico

El modelo que haya pasado por todas la etapas previas puede entonces utilizarse para:

- **análisis económico:**
  - interpretación de coeficientes,
  - contraste de hipótesis,
  - etc.
- **predicción:**
  - **predicción series temporales:**
    predecir valores futuros de Y.
  - **en general:**
    responder a preguntas del tipo
    ¿ qué pasaría si...?

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

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

supone para cada observación:

$$Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \cdots + \beta_K X_{K1} + u_1$$

$$Y_2 = \beta_0 + \beta_1 X_{12} + \beta_2 X_{22} + \cdots + \beta_K X_{K2} + u_2$$

$$\ldots\ldots$$

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_K X_{Kt} + u_t$$

$$\ldots\ldots$$

$$Y_T = \beta_0 + \beta_1 X_{1T} + \beta_2 X_{2T} + \cdots + \beta_K X_{KT} + u_T$$

MRLG en forma matricial (1)

o de otra manera en forma matricial:

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

MRLG en forma matricial (2)

- es decir:

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

-

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

↰

2.2 Supuestos Básicos (Clásicos) . Interpretación.

Supuestos Básicos del MRLG (1)

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

     X_k explica $Y \Leftrightarrow X_k \in$ 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:
     - $Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \beta_3 \dfrac{1}{X_t} + u_t$
     - $Y_t = A X_{1t}^{\beta_1} X_{2t}^{\beta_2} \mathrm{e}^{u_t}$ (¿ Por qué ?)
     - ¿ y éste ? $\quad Y_t = \beta_0 + \beta_1 \dfrac{1}{X_t - \beta_2} + u_t$
     - ¿ y estos otros ? $\quad \ln Y_t = \beta_0 X_t^{\beta_1} u_t; \qquad Y_t = \beta_0 X_t^{\beta_1} + u_t$

       $Y_t = \beta_1 X_{1t} + \beta_2 X_{1t} X_{2t} + u_t; \quad Y_t = \beta_0 + \beta_1 X_{1t}^{X_{2t}} + u_t$

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 :
  - $Y_t = \beta_0 + \beta_1 X_t + \beta_2 \big(2X_t + 3\big) + u_t$
  - $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 \big(X_{1t} + X_{2t}\big) + u_t$
- <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:
   $\mathsf{E}\left(u_t\right) = 0 \quad \forall t$ (¿ no es obvio ?).
   b) Homocedástico:
   $\mathsf{Var}\left(u_t\right) = \mathsf{E}\left(u_t^2\right) = \sigma_u^2 \; (= \sigma^2)$ const ($\forall t$).
   c) Serialmente incorrelacionado:
   $\mathsf{Cov}\left(u_t, u_s\right) = \mathsf{E}\left(u_t u_s\right) = 0 \quad \forall t \neq s$.
   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:

$$\underset{(T\times T)}{\mathsf{E}(uu')} = \begin{bmatrix} \mathsf{E}(u_1^2) & \mathsf{E}(u_1u_2) & \dots & \mathsf{E}(u_1u_T) \\ \mathsf{E}(u_2u_1) & \mathsf{E}(u_2^2) & \dots & \mathsf{E}(u_2u_T) \\ \dots & \dots & \dots & \dots \\ \mathsf{E}(u_Tu_1) & \mathsf{E}(u_Tu_2) & \dots & \mathsf{E}(u_T^2) \end{bmatrix}$$

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

Supuestos Básicos en forma matricial (2)

- de manera más compacta:

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

- plus 4*d*:

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

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

MRLS: la FRP

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

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

- 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)

- $\widehat{\alpha}, \widehat{\beta} \quad \rightsquigarrow$ 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:

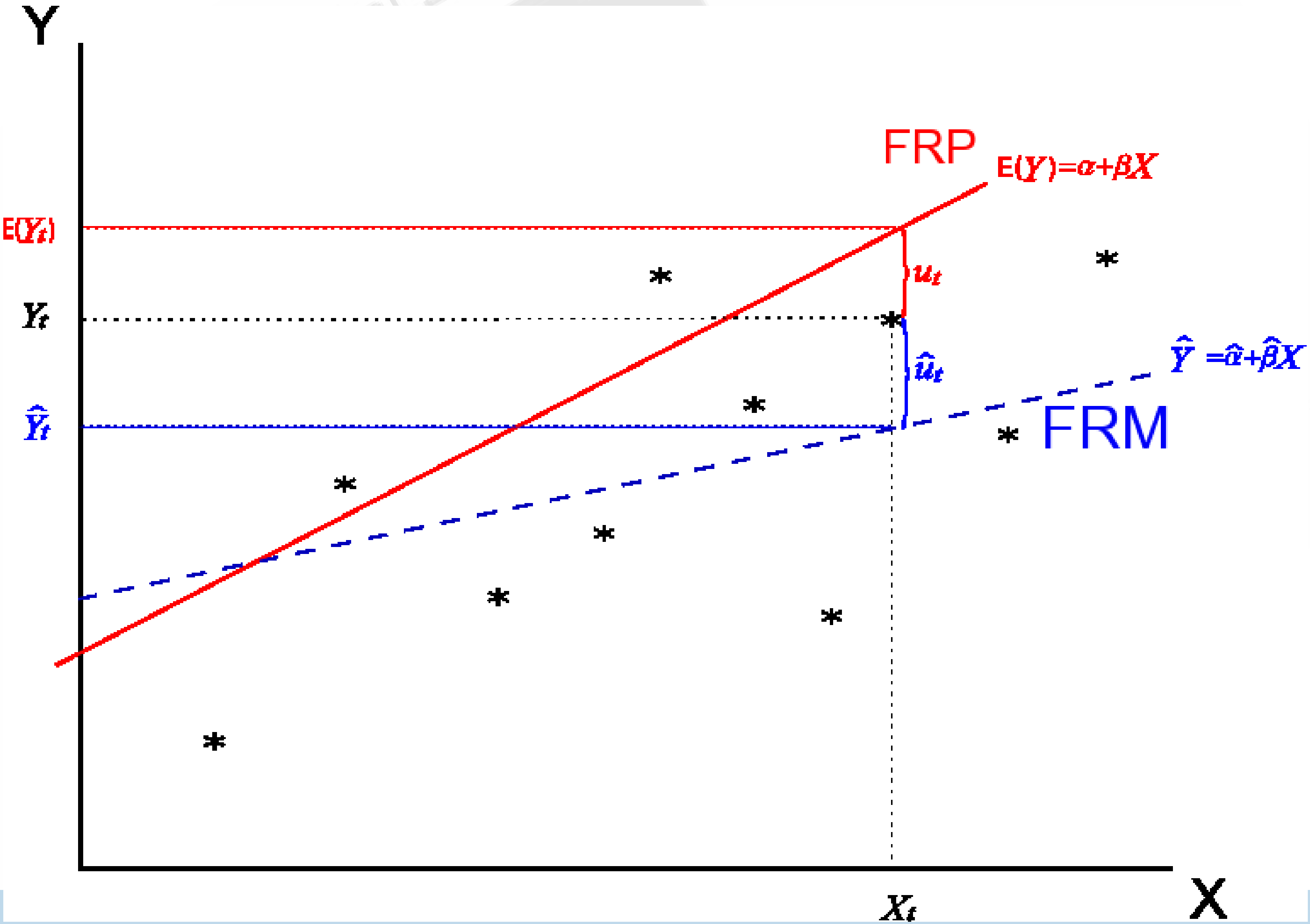
$$u_t = Y_t - \mathsf{E}\big(Y_t\big) = Y_t - \alpha - \beta X_t$$

- 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

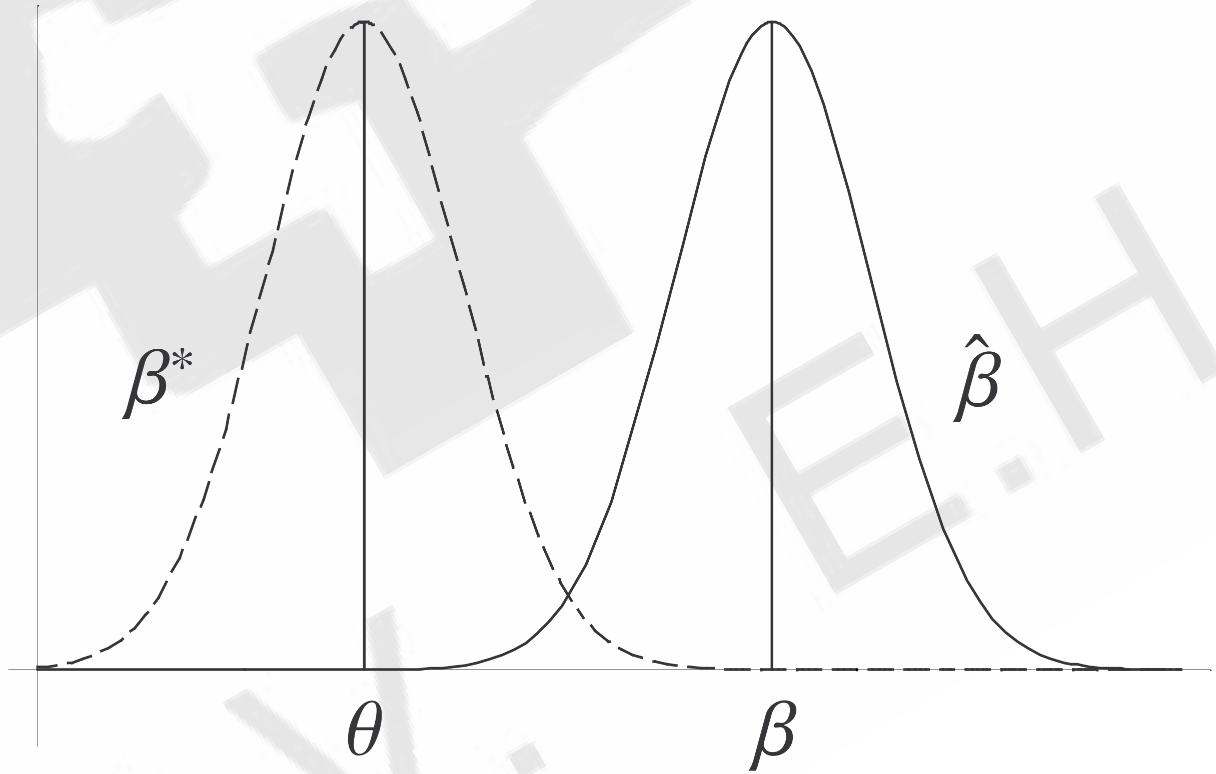


Estimación: Propiedades Deseadas (1)

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

Insesgadez:

$$\mathrm{E}\left(\widehat{\beta}\right) = \beta \quad \Leftrightarrow \quad \widehat{\beta} \text{ insesgado}$$

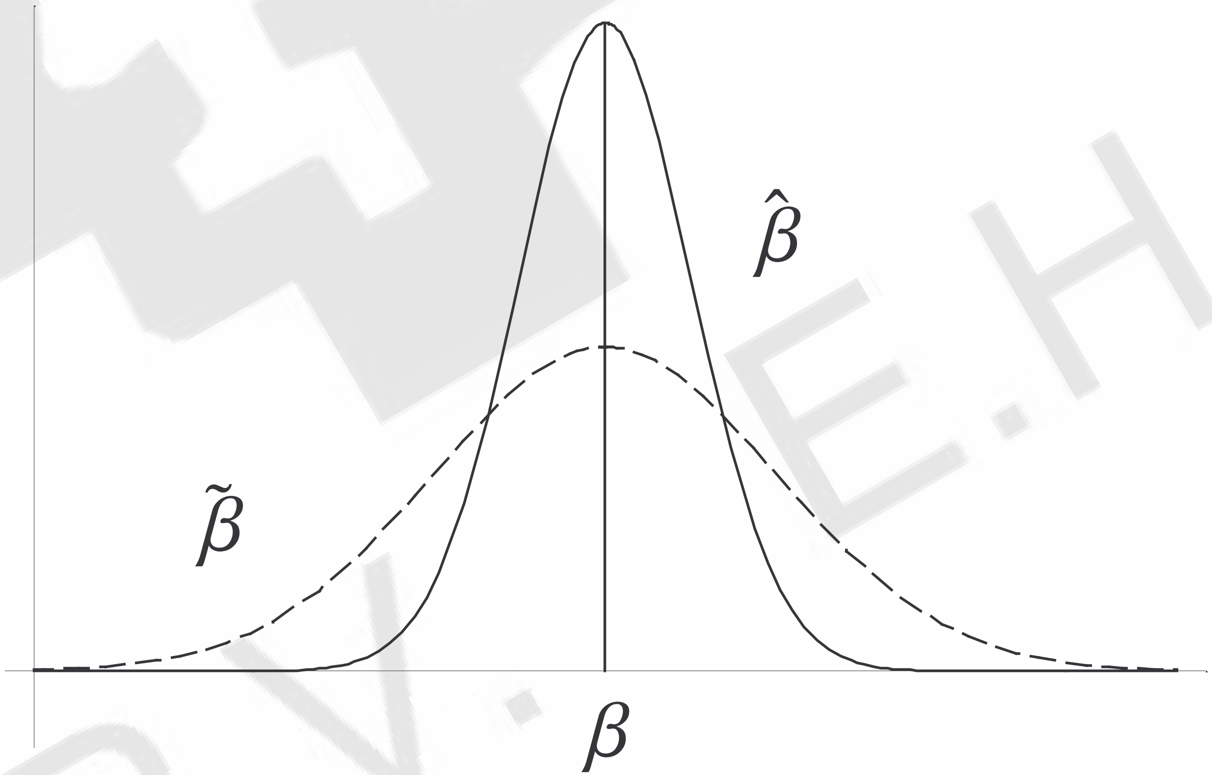


Estimación: Propiedades Deseadas (2)

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

Eficiencia relativa:

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



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^{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^{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:

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

¡ 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:

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

- 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

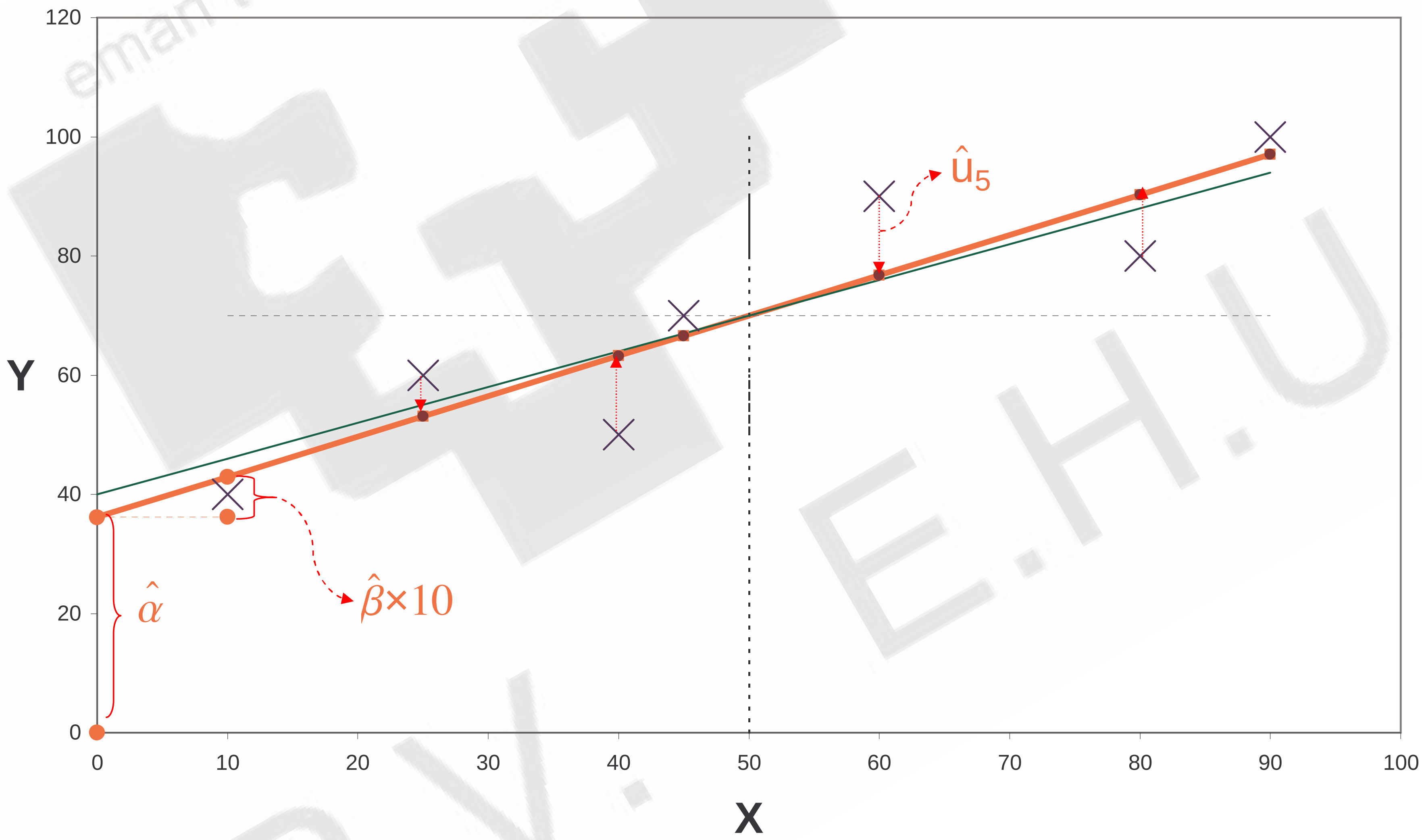
- Datos. . .
- Datos centrados o “en desviaciones”
  (desviaciones de las medias respectivas). . .
- Cuadrados y productos. . .

	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

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

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)

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

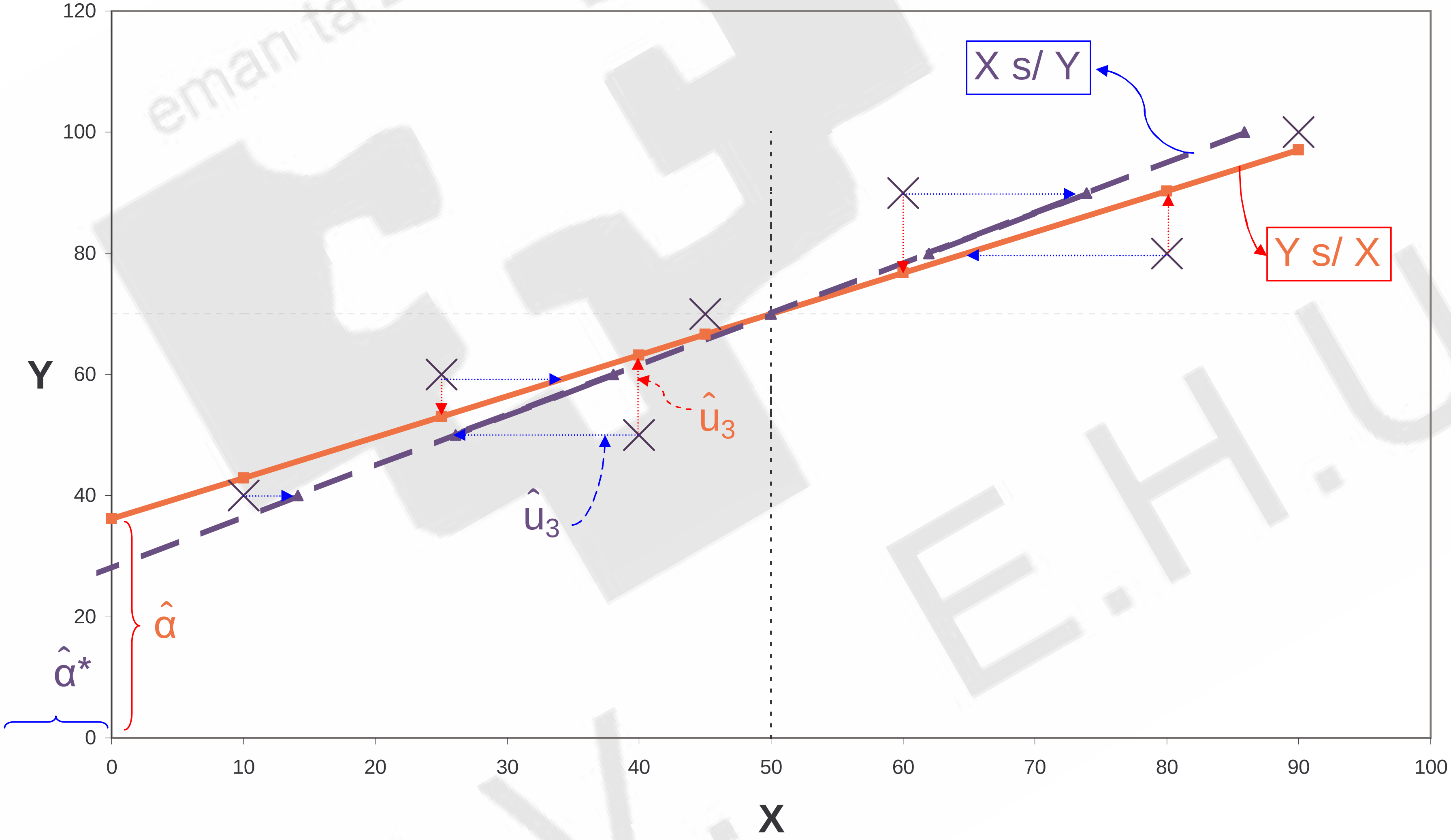
1. residuos suman cero: $\sum \widehat{u}_t = 0$
   Demo: directamente de las c.1^{er}o. $\square$
2. $\overline{\widehat{Y}} = \overline{Y}$
   Demo: por def.: $\widehat{u}_t = Y_t - \widehat{Y}_t \quad \rightsquigarrow \quad \overline{\widehat{Y}} = \overline{Y} - \overline{\widehat{u}}$,
   pero $\overline{\widehat{u}} = \frac{1}{T} \sum \widehat{u}_t = 0$ (de prop 1) $\rightsquigarrow \quad \overline{\widehat{Y}} = \overline{Y}$. $\square$
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) $\square$

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. $\square$
5. residuos ortogonales a la parte explicada de Y: $\sum \widehat{Y}_t \widehat{u}_t = 0$
   Demo: $\sum (\widehat{\alpha} + \widehat{\beta} X_t)\, \widehat{u}_t =$

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

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:

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

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

- $$\boxed{\underset{\substack{\downarrow \\ (SCT)}}{\sum y_t^2} = \underset{\substack{\downarrow \\ (SCE)}}{\sum \widehat{y}_t^2} + \underset{\substack{\downarrow \\ (SCR)}}{\sum \widehat{u}_t^2}}$$

- 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.1^{er}o. (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

$$R^2 = \frac{\frac{1}{T}\sum \widehat{y}_t^2}{\frac{1}{T}\sum y_t^2} = \frac{\frac{1}{T}\sum (\widehat{\beta} x_t)^2}{\frac{1}{T}\sum y_t^2} = \frac{\widehat{\beta}^2 \frac{1}{T}\sum x_t^2}{\frac{1}{T}\sum y_t^2}$$

$$= \widehat{\beta}^2 \frac{\text{Var}\left(X\right)}{\text{Var}\left(Y\right)} = \frac{\text{Cov}\left(Y, X\right)^2}{\text{Var}\left(X\right)^2} \frac{\text{Var}\left(X\right)}{\text{Var}\left(Y\right)}$$

$$= \frac{\text{Cov}\left(Y, X\right)^2}{\text{Var}\left(X\right)\text{Var}\left(Y\right)}$$

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

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$\big(X, Y\big)$? . . . ą‘ Inténtalo !!)

2.3b MCO en el MRLG.

MRLG: la FRP

- Recuérdese: modelo con K variables explicativas :

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt} + u_t, \\ Y &= X\beta + u \end{aligned} \tag{2}$$

se llama MRLG.

- Función de Regresión de Población (FRP):
$E(u) = 0 \rightsquigarrow$ *parte sistemática* o FRP:

$$\begin{aligned} E(Y_t) &= \beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt} \\ E(Y) &= X\beta \end{aligned}$$

- Interpretación de los coeficientes:
  - $\beta_0 = E(Y_t | X_{1t} = X_{2t} = \cdots = X_{Kt} = 0)$: Valor esperado de Y_t cuando todas las variables explicativas son iguales a cero.
  - $\beta_k = \dfrac{\partial E(Y_t)}{\partial X_{kt}} \simeq \dfrac{\Delta E(Y_t)}{\Delta X_{kt}}, \quad k = 1 \dots K$: Incremento en el valor (esperado) Y_t cuando $X_k \uparrow$ una unidad (c.p.).

Función de Regresión Muestral (FRM)

- Objetivo del MRLG: obtener estimador $\widehat{\beta} = (\widehat{\beta}_0, \widehat{\beta}_1 \ldots, \widehat{\beta}_K)'$ del vector de parámetros desconocidos en (2).
  $\widehat{\beta} \rightsquigarrow$ modelo estimado, ajuste o FRM:

$$\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_{1t} + \cdots + \widehat{\beta}_K X_{Kt}$$

$$\widehat{Y} = X\widehat{\beta}$$

- Notas:
  - Perturbaciones en FRP:

$$u_t = Y_t - E(Y_t) = Y_t - \beta_0 - \beta_1 X_{1t} - \cdots - \beta_K X_{Kt}$$

$$u = Y - E(Y) = Y - X\beta$$

  - Residuos en FRM:

$$\widehat{u}_t = Y_t - \widehat{Y}_t = Y_t - \widehat{\beta}_0 - \widehat{\beta}_1 X_{1t} - \cdots - \widehat{\beta}_K X_{Kt}$$

$$\widehat{u} = Y - \widehat{Y} = Y - X\widehat{\beta}$$

- Residuos son a la FRM lo que las perturbaciones a la FRP.

Estimación: MCO

- aplicar ajuste Mínimo-Cuadrático al MRLG: $Y = X\beta + u$,
- o bien en forma de observación:

$$\min_{\beta_0 \dots \beta_K} \sum_{t=1}^{T} u_t^2 \text{ donde } u_t = Y_t - \beta_0 - \beta_1 X_{1t} - \cdots - \beta_K X_{Kt}$$

- o en forma matricial:

[recuérdese:

$$u' = (u_1, u_2, \dots, u_T) \qquad u = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_T \end{pmatrix}$$

entonces $u'u = u_1^2 + u_2^2 + \cdots + u_T^2 = \sum_{t=1}^{T} u_t^2$]

- es decir

$$\boxed{\min_{\beta} \; u'u \quad \text{donde} \quad u = Y - X\beta}$$

Nota: derivadas vectoriales

- Sea $u = u(\beta)$: derivs de $c\,u$ y $c\,u^2$ con respecto a β:

$$\frac{\partial}{\partial\beta}(c\,u) = c\,\frac{\partial u}{\partial\beta} \qquad \text{y} \qquad \frac{\partial}{\partial\beta}\,u^2 = 2\,u\frac{\partial u}{\partial\beta}$$

- Con vectores y matrices esto es bastante similar:

- La derivada de la combinación lineal $u'c$

$$\underset{(1\times n)}{u'}\;\underset{(n\times 1)}{c} \qquad \left(=\textstyle\sum_{i=1}^{n} c_i u_i, \text{ *es decir* ¡escalar!!}\right)$$

con respecto a $\underset{(k\times 1)}{\beta}$ es: $\frac{\partial(u'c)}{\partial\beta} = \frac{\partial u'}{\partial\beta}c$

- La derivada de la suma de cuadrados $u'u$

$$\underset{(1\times n)}{u'}\;\underset{(n\times 1)}{u} \qquad \left(=\textstyle\sum_{i=1}^{n} u_i^2, \text{ *es decir* ¡escalar!!}\right)$$

con respecto a $\underset{(k\times 1)}{\beta}$ es: $\frac{\partial(u'u)}{\partial\beta} = 2\,\frac{\partial u'}{\partial\beta}u$

c.1^{er}o. en forma matricial

$$\min_{\beta}(u'u) \quad \text{donde} \quad u = Y - X\beta$$

Primera derivadas de SC $u'u$ con respecto a β:

$$\begin{aligned}\frac{\partial u'u}{\partial\beta} &= 2\,\frac{\partial u'}{\partial\beta}u \\ &= 2\,\frac{\partial(Y' - \beta' X')}{\partial\beta}u \\ &= -2\,X'u\end{aligned}$$

en el mínimo:

$$\boxed{\text{c.}1^{\text{er}}\text{o.} \quad \underset{(K+1\times T)}{X'} \; \underset{(T\times 1)}{\hat{u}} = 0_{K+1}}$$

Estimación: Ecuaciones normales y EMC de β

Resolviendo las c.1^{er}o. obtenemos las ecuaciones normales:

$$X'(Y - X\widehat{\beta}) = 0 \;\Rightarrow$$

$$\boxed{\underset{(K{+}1\,\times\,1)}{X'Y} = \underset{(K{+}1\,\times\,K{+}1)}{X'X} \;\; \underset{(K{+}1\,\times\,1)}{\widehat{\beta}}} \tag{3}$$

De donde premultiplicando por $(X'X)^{-1}$ obtenemos el estimador MCO:

$$\boxed{\widehat{\beta}_{\text{MCO}} = (X'X)^{-1}X'Y}$$

Estimación: EMC de β (cont)

- donde $X'X$ es una matriz $[K{+}1 \times K{+}1]$: [¿ recuérdese X e Y ? $\longrightarrow$]

-

$$\underset{(K+1\,\times\,K+1)}{X'X} = \begin{bmatrix} T & \sum X_{1t} & \sum X_{2t} & \dots & \sum X_{Kt} \\ \sum X_{1t} & \sum X_{1t}^2 & \sum X_{1t}X_{2t} & \dots & \sum X_{1t}X_{Kt} \\ \dots & \dots & \dots & \dots & \dots \\ \sum X_{Kt} & \sum X_{Kt}X_{1t} & \sum X_{Kt}X_{2t} & \dots & \sum X_{Kt}^2 \end{bmatrix}$$

- y $X'Y$ y $\widehat{\beta}$ son vectores $[K{+}1 \times 1]$:

$$\underset{(K+1\,\times\,1)}{X'Y} = \begin{bmatrix} \sum Y_t \\ \sum X_{1t}Y_t \\ \dots \\ \sum X_{Kt}Y_t \end{bmatrix} \qquad \underset{(K+1\,\times\,1)}{\widehat{\beta}} = \begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \\ \dots \\ \widehat{\beta}_K \end{bmatrix}$$

Estimador MCO con datos centrados

Una forma alternativa de obtener el estimador MCO es

$$\boxed{\widehat{\beta}^{\star}_{\text{MCO}} = (x'x)^{-1}x'y}$$

para los coeficientes del modelo.

...junto con el intercepto estimado obtenido de la primera ecuación normal

$$\boxed{\widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1\overline{X}_1 - \cdots - \widehat{\beta}_K\overline{X}_K}$$

Nota: caso especial de $K = 1 \rightsquigarrow$ ¡ formulas idénticas a las del MRLS !! (Demuéstralo).

2.4b Propiedades de la FRM.

Propiedades de los residuos y la FRM (1)

$$\left.\begin{array}{c}\widehat{\beta} \\ \widehat{\beta}^{\star} \rightsquigarrow \widehat{\beta}_0\end{array}\right\} \rightsquigarrow \widehat{Y} = X\widehat{\beta} \rightsquigarrow \widehat{u} = Y - \widehat{Y}$$

1. residuos suman cero: $\sum \widehat{u}_t = 0$
   Demo: directamente de las c.1^{er}o.:

$$X'\widehat{u} = 0 \Rightarrow \begin{bmatrix} \sum_1^T \widehat{u}_t \\ \sum_1^T X_{1t}\widehat{u}_t \\ \sum_1^T X_{2t}\widehat{u}_t \\ \dots \\ \sum_1^T X_{Kt}\widehat{u}_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

$\square$

2. $\overline{\widehat{Y}} = \overline{Y}$
3. la FRM pasa a través del vector $(\overline{X}_1, \dots \overline{X}_K, \overline{Y})$:
   $\overline{Y} = \widehat{\beta}_0 + \widehat{\beta}_1\overline{X}_1 + \cdots + \widehat{\beta}_K\overline{X}_K$

Nota: Estas propiedades 1 a 3 son satisfechas si la regresión tiene intercepto; es decir, si X tiene una columna de "unos".

Propiedades de los residuos y la FRM (2)

4. residuos ortogonales a v. explicativas: $X'\widehat{u} = 0$

   Demo: directamente de las c.1^{er}o. (ver 1) o, alternativamente:

$$X'\widehat{u} = X'(Y - X\widehat{\beta}) = X'Y - X'X\widehat{\beta}$$
$$= X'Y - \underbrace{X'X(X'X)^{-1}}_{=I_{K+1}} X'Y = 0$$

$\square$

5. residuos ortogonales a la parte explicada de Y: $\widehat{Y}'\widehat{u} = 0$

   Demo: $\widehat{Y}'\widehat{u} = (X\widehat{\beta})'\widehat{u} = \widehat{\beta}' \underbrace{X'\widehat{u}}_{=0} = 0$ $\square$

2.5b Bondad de Ajuste:
Coeficiente de Determinación (R^2) y Estimación de la Varianza del Error.

Bondad de Ajuste: R^2 (repaso)

Recuérdese (lo mismo que antes pero ahora lo hacemos con vectores):

$$
\begin{aligned}
Y'Y &= (\widehat{Y}' + \widehat{u}')(\widehat{Y} + \widehat{u}) \\
&= \widehat{Y}'\widehat{Y} + \widehat{u}'\widehat{u} + 2\widehat{Y}'\widehat{u} \\
&= \widehat{Y}'\widehat{Y} + \widehat{u}'\widehat{u} \quad \text{(desde prop 5)}
\end{aligned}
$$

$$Y'Y - T\overline{Y}^2 = \widehat{Y}'\widehat{Y} - T\overline{\widehat{Y}}^2 + \widehat{u}'\widehat{u} \qquad \text{(de prop 2)}$$

$$\boxed{\underset{(SCT)}{\underset{\downarrow}{y'y}} = \underset{(SCE)}{\underset{\downarrow}{\widehat{y}'\widehat{y}}} + \underset{(SCR)}{\underset{\downarrow}{u'u}}}$$

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

$$0 \leq R^2 \leq 1$$

Bondad de Ajuste: R^2 (repaso cont)

Nota 1: R^2 mide la proporción de la variación de la variable dependiente explicada por la variación de (una combinación lineal de) las variables explicativas.

Nota 2:

$$\text{no intercepto} \Rightarrow \begin{cases} \nexists 1^a \text{ fila de c.}1^{\text{er}}\text{o.} \rightsquigarrow \begin{cases} \sum \widehat{u}_t \neq 0, \\ \overline{\widehat{Y}} \neq \overline{Y}, \end{cases} \\ \text{no válido} R^2 \text{ (¡ Recuérdese !)} \end{cases}$$

Estimación de $\mathrm{Var}(u_t)$

$$\sigma^2 = \mathsf{Var}(u_t) = \mathsf{E}(u_t^2) \simeq \frac{1}{T}\sum_{t=1}^{T} u_t^2$$

pero con residuos se tienen que satisfacer $K{+}1$ relaciones lineales en $X'\widehat{u} = 0$ así que perdemos $K{+}1$ grados de libertad:

$$\widehat{\sigma}^2 = \frac{1}{T{-}K{-}1}\sum_{t=1}^{T} \widehat{u}_t^2$$

Por lo tanto proponemos el siguiente estimador:

$$\boxed{\widehat{\sigma}^2 = \frac{\mathsf{RSS}}{T{-}K{-}1}}$$

el cual claramente es un estimador insesgado:

Demo:

$$\mathsf{E}(\widehat{\sigma}^2) = \frac{\mathsf{E}(\mathsf{SCR})\,(^*)}{T{-}K{-}1} = \frac{T{-}K{-}1}{T{-}K{-}1} = \sigma^2$$

☐ (* ver libro de texto)

2.6 Propiedades del Estimador de Mínimos Cuadrados en muestras finitas.
El Teorema de Gauss-Markov.

Propiedades del Estimador MCO (1)

El estimador $\widehat{\beta}_{\text{MCO}} = (X'X)^{-1}X'Y$ tiene las siguientes propiedades:

- **Lineal:** $\widehat{\beta}_{\text{MCO}}$ es una combinación lineal de perturbaciones:

$$\begin{aligned}\widehat{\beta} &= (X'X)^{-1}X'(X\beta + u)\\ &= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'u\\ &= \beta + (X'X)^{-1}X'u\\ &= \beta + \Gamma'u\end{aligned}$$

- **insesgado:** Dado queo $\text{E}(u) = 0$, $\widehat{\beta}_{\text{MCO}}$ es insesgado:

$$\begin{aligned}\text{E}(\widehat{\beta}) &= \text{E}(\beta + \Gamma'u)\\ &= \beta + \Gamma'\text{E}(u)\\ &= \beta\end{aligned}$$

Propiedades del Estimador MCO (2)

- **Varianza:** Recuérdese:

$$\text{Var}\big(u\big) = \sigma^2 I_T,$$
$$\widehat{\beta} = \beta + (X'X)^{-1}X'u,$$

$$
\begin{aligned}
\text{Var}\big(\widehat{\beta}\big) &= \text{E}\big((\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'\big) \\
&= \text{E}\big((X'X)^{-1}X'u\,u'X(X'X)^{-1}\big) \\
&= (X'X)^{-1}X'\,\text{E}\big(uu'\big)\,X(X'X)^{-1} \\
&= (X'X)^{-1}X'\,\sigma^2 I_T\,X(X'X)^{-1} \\
&= \sigma^2(X'X)^{-1}X'X(X'X)^{-1}
\end{aligned}
$$

$$\boxed{\text{Var}\big(\widehat{\beta}\big) = \sigma^2(X'X)^{-1}}$$

Propiedades del Estimador MCO (2cont)

$$\text{Var}(\widehat{\beta}) = \begin{bmatrix} \text{Var}(\widehat{\beta}_0) & \text{Cov}(\widehat{\beta}_0, \widehat{\beta}_1) & \dots & \text{Cov}(\widehat{\beta}_0, \widehat{\beta}_K) \\ \text{Cov}(\widehat{\beta}_1, \widehat{\beta}_0) & \text{Var}(\widehat{\beta}_1) & \dots & \text{Cov}(\widehat{\beta}_1, \widehat{\beta}_K) \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ \text{Cov}(\widehat{\beta}_K, \widehat{\beta}_0) & \text{Cov}(\widehat{\beta}_K, \widehat{\beta}_1) & \dots & \text{Var}(\widehat{\beta}_K) \end{bmatrix}$$

$$\sigma^2(X'X)^{-1} = \sigma^2 \begin{bmatrix} a_{00} & a_{00} & a_{01} & \dots & a_{0K} \\ a_{10} & a_{11} & a_{12} & \dots & a_{1K} \\ a_{20} & a_{21} & a_{22} & \dots & a_{2K} \\ \dots & \dots & \dots & \dots & \dots \\ a_{K0} & a_{K1} & a_{K2} & \dots & a_{KK} \end{bmatrix}$$

es decir a_{kk} es el $(k+1, k+1)$-elemento del matriz $(X'X)^{-1}$:

$$\text{Var}(\widehat{\beta}_k) = \sigma^2 a_{kk}$$

$$\text{Cov}(\widehat{\beta}_k, \widehat{\beta}_i) = \sigma^2 a_{ki}$$

Teorema Gauss-Markov

"Dados los supuestos básicos del MRLG, el estimador MCO es el de varianza mínima (el mejor) entre todos los estimadores lineales e insesgados"

$$\widehat{\beta}_{\text{MCO}} = \textbf{ELIO} = \textbf{E}_{\text{stimador}}\textbf{L}_{\text{ineal}}\textbf{I}_{\text{nsesgado}}\textbf{O}_{\text{ptimo}}$$

Demo:

Sea $\widetilde{\beta}$ algún otro estimador lineal insesgado:

$$\widetilde{\beta} = D'Y = D'(X\beta + u) = D'X\beta + D'u$$

$$\mathsf{E}(\widetilde{\beta}) = D'X\beta + D'\mathsf{E}(u) = D'X\beta = \beta \Rightarrow \boxed{D'X = I_K}$$

entonces $\widetilde{\beta} = \beta + D'u \quad \rightsquigarrow \quad \widetilde{\beta} - \beta = D'u$
y su varianza:

$$\text{Var}(\widetilde{\beta}) = E\left[(\widetilde{\beta} - \beta)(\widetilde{\beta} - \beta)'\right] = \mathsf{E}(D'u\,u'D)$$

$$= D'\,\mathsf{E}(uu')\,D = D'\,\sigma^2\,I_T\,D = \sigma^2\,D'D$$

Teorema Gauss-Markov (cont)

$$
\begin{aligned}
\text{Var}(\widetilde{\beta}) - \text{Var}(\widehat{\beta}) &= \sigma^2 D'D - \sigma^2 (X'X)^{-1} \\
&= \sigma^2 \left[D'D - (X'X)^{-1} \right] \\
&= \sigma^2 \left[D'D - D'X\,(X'X)^{-1}\,X'D \right] \\
&= \sigma^2 D' \underbrace{\left[I_T - X\,(X'X)^{-1}\,X' \right]}_{M} D \\
&= \sigma^2 D'\,(MM)\,D \\
&= \sigma^2 (D'M)(M'D) = D^{\star\prime} D^{\star} \\
&> 0
\end{aligned}
$$

Esto es, en particular **todas** las varianzas individuales serán mayores que las respectivas por MCO.

2.3c MCO: Expresiones útiles y Cronología.

Expresiones útiles para SC

- $$SCT = \sum (Y_t - \overline{Y})^2 = \sum Y_t^2 - T\overline{Y}^2 = Y'Y - T\overline{Y}^2$$

- $$SCE = \sum (\widehat{Y}_t - \overline{\widehat{Y}})^2 = \sum \widehat{Y}_t^2 - T\overline{\widehat{Y}}^2 = \sum \widehat{Y}_t^2 - T\overline{Y}^2 = \widehat{Y}'\widehat{Y} - T\overline{Y}^2$$
$$= (X\widehat{\beta})'(X\widehat{\beta}) - T\overline{Y}^2 = \widehat{\beta}'\underbrace{X'X\widehat{\beta}}_{X'Y} - T\overline{Y}^2 = \widehat{\beta}'X'Y - T\overline{Y}^2$$

- $$SCR = \sum \widehat{u}_t^2 = \widehat{u}'\widehat{u} \qquad = \sum Y_t^2 - \sum \widehat{Y}_t^2 = Y'Y - \widehat{\beta}'X'Y$$

Principales expresiones y Cronología

- $Y = X\beta + u$
- $(X'X)^{-1} \qquad X'Y$
- $\widehat{\beta} = (X'X)^{-1}X'Y$
- $ESS = \widehat{\beta}'X'Y - T\overline{Y}^2$ (¡ necesita $\overline{Y}$!)
- $TSS = Y'Y - T\overline{Y}^2$
- $RSS = Y'Y - \widehat{\beta}'X'Y$ (¡ no $\overline{Y}$!)
- $R^2 = \dfrac{ESS}{TSS} = 1 - \dfrac{RSS}{TSS}$
- $\widehat{\sigma}^2 = \dfrac{RSS}{T - K - 1}$
- $\widehat{\text{Var}(\widehat{\beta})} = \widehat{\sigma}^2 (X'X)^{-1}$

2.7a Omisión de variables relevantes.

Omisión de variables relevantes

- relación verdadera:

$$Y = X\beta + u = \left[X_I \;\middle|\; X_{II} \right] \left(\frac{\beta_I}{\beta_{II}} \right) + u$$

$$X = \left[\begin{array}{cccc|ccc} 1 & X_{11} & \dots & X_{K_1,1} & X_{K_1+1,1} & \dots & X_{K1} \\ 1 & X_{12} & \dots & X_{K_1,2} & X_{K_1+1,2} & \dots & X_{K2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & X_{1T} & \dots & X_{K_1,T} & X_{K_1+1,T} & \dots & X_{KT} \end{array} \right], \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{K_1} \\ \hline \beta_{K_1+1} \\ \vdots \\ \beta_K \end{pmatrix}.$$

$$Y = X_I\beta_I + X_{II}\beta_{II} + u$$

- relación estimada :

$$Y = X_I\beta_I + v \qquad \text{donde} \quad v = X_{II}\beta_{II} + u,$$

$$\text{entonces} \quad \mathrm{E}(v) \neq 0 \quad \rightsquigarrow \quad \mathrm{E}(\widehat{\beta}) \neq \beta.$$

es decir $\widehat{\beta}$ es sesgado.

Omisión de variables relevantes: consecuencias

Resumen:

- Estimador MCO de los coeficientes es *sesgado* (excepto si $x_I' x_{II} = 0$).
- Estimador MCO del intercepto es *siempre sesgado*.
- Estimador de varianza del error es *siempre sesgado*.

2.7b Multicolinealidad

Multicolinealidad Perfecta

Caso extremo:

- combinación lineal exacta:
  - $\sum_{k=0}^{K} \lambda_k X_{kt} = 0, \quad \lambda \neq 0, \quad X_{0t} = 1,$
  - $\exists\, X_i \mid X_i = \lambda_0^* + \sum_{\substack{k=1 \\ k \neq i}}^{K} \lambda_k^* X_{kt},$
  - $\exists\, X_i, X_j \mid \text{Corr}(X_i, X_j) = 1,$
  - $\exists\, X_i \mid$ regres aux X_i sobre $\{X_k\}_{\substack{k=1 \\ k \neq i}}^{K} \quad \rightsquigarrow \quad \mathrm{R}_i^2 = 1.$
- Problema:
  - $\mathrm{rk}\, X < K{+}1$, (X no es de rango completo)
  - $\rightsquigarrow \quad \det(X) = 0$
  - $\rightsquigarrow \quad \nexists (X'X)^{-1}$
  - $\rightsquigarrow$

¿ $\widehat{\beta}$?

Multicolinealidad perfecta: ejemplo

- Sea $X_{4t} = 2X_{1t} \quad \forall t$:

$$X_{4t} = 0 + 2X_{1t} + 0 \cdot X_{2t} + 0 \cdot X_{3t} + 0 \cdot X_{5t} + \cdots + 0 \cdot X_{Kt},$$

- ¿ no error ? $\Rightarrow$ regres aux X_4 sobre $\{X_k\}_{\substack{k=1\\k\neq 4}}^{K}$ $\rightsquigarrow$ ¡ $\mathrm{R}_4^2 = 1$!!

- Especificación del modelo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \cdots + u_t, t = 1, 2 \ldots, T,$$
$$X_{4t} = 2X_{1t},$$

- y sustituyendo en el modelo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 (2X_{1t}) + \cdots + u_t,$$
$$= \beta_0 + \underbrace{(\beta_1 + 2\beta_4)}_{\beta_1^\star} X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + u_t$$

- ahora tenemos un parámetro menos para estimar.

Multicolinealidad: contra-ejemplo

$$Y_t = \beta_0 + \beta_1^\star X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + u_t$$

- Sólo quedan K parámetros a estimar,
  pero β_1 y β_4 no se pueden estimar por separado:
  - únicamente podemos estimar una combinación lineal de ellos:
    $$\beta_1^\star = \beta_1 + 2\beta_4,$$
  - *es decir* ¡ efecto combinado de X_{1t} y X_{4t} sobre Y_t !!

- (Ejercicio: Pruébese con $X_{2t} - 3X_{3t} = 10, \quad \forall t$.)

- multicolinealidad $=$ *relaciones lineales*
  pero... ¿ qué ocurre si la relación *no es lineal* ? *p.ej.:*
  $$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{1t}^2 + u_t$$
  - X tiene rango completo por columnas $\rightsquigarrow$ no hay problema.

Multicolinealidad perfecta : consecuencias

- algunos parámetros no pueden estimarse por separado.
- algunas estimaciones son sólo c.l. de los parámetros.
- R^2 es correcto:
  recoge correctamente la proporción de (la varianza de) Y_t explicada por la regres
- Las predicciones de Y son todavía validas.

2.7c Multicolinealidad imperfecta

Multicolinealidad imperfecta

- Problema:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \cdots + u_t, t = 1, 2 \ldots, T,$$

$$X_{4t} = 2X_{1t} + v_t,$$

$v_t =$ discrepancia entre X_{4t} y $2\,X_{1t}$,

- relación aproximada :
- regresión auxiliar X_{4t} sobre el resto $\rightsquigarrow$ $\mathrm{R}^2 \approx 1$.
- es una cuestión de grado ($x'x$ no diagonal $\rightsquigarrow$ variables correlacionadas)
- Nota: si no se especifica perfecta/imperfecta quiere decir mc imperfecta.

Multicolinealidad: Síntomas

- Síntomas típicos :
  - alto R^2

    (grupo relevante de regresores)
  - pero variables parecen no relevantes individualmente

    (incapacidad de separar efectos de regresores).
- más formalmente:

$$\text{Var}\left(\widehat{\beta}^{\star}\right) = \sigma^2 (x'x)^{-1} = \frac{\sigma^2}{T}\text{Var}\left(X^{\star}\right)^{-1}$$

$$\Rightarrow \quad \text{Var}\left(\widehat{\beta}_k\right) = \frac{\sigma^2}{T\text{Var}\left(X_k\right)(1-\mathrm{R}_k^2)},$$

- de forma que en el ejemplo anterior $X_{4t} \approx 2X_{1t}$:
  - Corr$\left(X_4, X_1\right) \uparrow$
  - R_4^2 y R_1^2 $\upuparrows$
  - denominador $\downarrow$
  - varianzas $\upuparrows$

Multicolinealidad: Consecuencias

- Algunos coeficientes no son significativos, a pesar de que sus variables tienen un efecto importante sobre la variable dependiente.
- De todas maneras, Gauss-Markov
  $\Rightarrow$ estimadores lineales, insesgados y de varianza mínima,
  entonces *no es posible encontrar un ELI (más)Optimo.*
- R^2 es correcto:
  recoge correctamente la proporción de (la varianza de) Y_t
  explicado por la regresión.
- Predicciones de Y son todavía validas.

Multicolinealidad: Como detectar

- Pequeños cambios en los datos
  $\Rightarrow$ importantes cambios en las estimaciones
  (pueden afectar hasta sus signos).
- Estimaciones de los coeficientes
  no son significativas de forma individual...
- ...pero sí lo son de forma conjunta.
- Alto coeficiente de determinación R^2.
- Regresiones auxiliares entre regresores
  $\Rightarrow$ alto R^2_k.

Multicolinealidad: Algunas soluciones

La multicolinealidad no es un problema fácil de solucionar.
De todas formas, de

$$\text{Var}\big(\widehat{\beta}_k\big) = \frac{\sigma^2}{T\text{Var}\big(X_k\big)(1-\mathrm{R}_k^2)},$$

tenemos que para reducir la varianza podríamos:

- $\mathbf{T}\uparrow$: Incrementar el número de observaciones T .
  También, las diferencias entre regresores pueden incrementar.
- $\text{Var}\big(\mathbf{X}\big)\uparrow$: Incrementar dispersión de los datos; *p.ej.* estudio sobre la función del consumo:
  muestra de familias $\leftrightsquigarrow$ todas las rentas posibles.
- $\text{Var}\big(\mathbf{X}\big)\uparrow$: Incluir información adicional.
  p.ej. imponer restricciones sugeridas por T^a. Ec.
- $\sigma^{\mathbf{2}}\downarrow$: Añadir un nuevo regresor relevante todavía no incluido.
  También evitaría serios problemas de sesgo.
- $\mathbf{R_k^2}\downarrow$: Eliminar variables que pueden ser causa de la multicolinealidad.
  (Aunque tener cuidado de no omitir algún regresor relevante).

2.8 Estimador MCO bajo Restricciones.

MRLG bajo restricciones lineales (1)

- objetivos de capítulos previos:
  - Modelo Econométrico (MRLG), características y supuestos básicos. . .
  - pero. . . no hay conocimiento sobre los parámetros del modelo.
  - El Método Mínimo-Cuadrático de estimación de parámetros (MCO).
  - Propiedades de los estimadores resultantes .
- objetivos del presente capítulo:
  - Información *a priori* sobre los valores del parámetro (o c.l.) . . .
  - dada por
    - teoría económica,
    - otros trabajos empíricos,
    - experiencia propia, etc.
  - Modelo No Restringido $\Rightarrow$ MC Ordinarios.
  - Modelo Restringido $\Rightarrow$ MC Restringidos.
  - Comprobar, dado el modelo estimado, si la información es compatible con los datos disponibles.

MRLG bajo restricciones lineales: ejemplos

- función de producción con rendimientos de escala constantes: $\beta_K + \beta_L = 1$.
- demanda del producto como función del precio: $\beta = -1$ (por ejemplo).
- en el MRLG: supongamos que $\beta_2 = 0$ y $2\beta_3 = \beta_4 - 1$:
  - Modelo completo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_{Kt} X_{Kt} + u_t, \text{ con } \beta_2 = 0 \text{ y } 2\beta_3 + 1 = \beta_4;$$

  - Modelo transformado alternativo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + 0 X_{2t} + \beta_3 X_{3t} + (2\beta_3 + 1) X_{4t} + \cdots + \beta_{Kt} X_{Kt} + u_t$$
$$Y_t - X_{4t} = \beta_0 + \beta_1 X_{1t} + \beta_3 (X_{3t} + 2X_{4t}) + \cdots + \beta_K X_{Kt} + u_t$$
$$Y_t^* = \beta_0 + \beta_1 X_{1t} + \beta_3 Z_t + \cdots + \beta_K X_{Kt} + u_t$$

    donde $Y_t^* = Y_t - X_{4t}$ y $Z_t = X_{3t} + 2X_{4t}$.
  - Este modelo transformado:
    - se puede estimar por MCO:

$$\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_3, \widehat{\beta}_5, \ldots, \widehat{\beta}_K \text{ , junto con } \widehat{\beta}_2 = 0 \text{ y } \widehat{\beta}_4 = 2\widehat{\beta}_3 + 1 \text{ .}$$

    - tiene nueva variable endógena Y_t^* (no siempre es así: p.ej. si $\beta_2 = 0$ sólo) y nueva variable explicativa Z_t.

MRLG bajo restricciones lineales (2)

- El método de “transformación” es bueno solamente para casos sencillos.
- En general, q restricciones lineales (no redundantes) entre parámetros:

$$\begin{array}{c} 1 \\ \vdots \\ q \end{array} \begin{pmatrix} \diamond & \diamond & \diamond & \dots & \diamond \\ \vdots & & & & \\ \diamond & \diamond & \diamond & \dots & \diamond \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_K \end{pmatrix} = \begin{pmatrix} \diamond \\ \vdots \\ \diamond \end{pmatrix}$$

- para una matriz R y vector r dados,

$$\underset{(q \times K+1)}{R} \beta = \underset{(q \times 1)}{r}$$

- <1>ejemplo de caso no valido (¿ por qué ?):

$$\beta_3 = 0, \quad 2\beta_2 + 3\beta_4 = 1, \quad \beta_1 - 2\beta_4 = 3, \quad 6\beta_4 = 2 - 4\beta_2 + \beta_3$$

MRLG bajo restricciones lineales (2cont)

- Escribir el ejemplo anterior $\beta_2 = 0$ y $2\beta_3 = \beta_4 - 1$ ($q = 2$ restricciones) como en la formula general:

$$\underset{(2\times K+1)}{R}\;\; \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 2 & -1 & 0 & \dots & 0 \end{pmatrix} \underset{(K+1\times 1)}{\beta}\;\; \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_K \end{pmatrix} = \underset{(2\times 1)}{r}\;\; \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

- En general, escribimos el MRLG sujeto a q restricciones lineales como:

$$\underset{(T\times 1)}{Y} = \underset{(T\times K+1)}{X}\;\underset{(K+1\times 1)}{\beta} + \underset{(T\times 1)}{u},$$

$$\underset{(q\times K+1)}{R}\;\underset{(K+1\times 1)}{\beta} = \underset{(q\times 1)}{r}.$$

Estimación:MC Restringidos.

- Típico ejercicio de optimización:

$$\min_{\beta}(u'u) \quad \text{donde} \quad u = Y - X\beta,$$

sujeto a $R\,\beta = r$.

- Lagrangiano:

$$L(\beta,\lambda) = u'u - 2\lambda'(R\beta - r)$$

$$\min_{\beta,\lambda} L(\beta,\lambda).$$

- Primeras derivadas:

$$\frac{\partial L(\beta,\lambda)}{\partial \beta} = -2\,X'u - 2\,R'\lambda,$$

$$\frac{\partial L(\beta,\lambda)}{\partial \lambda} = -2\,(R\,\beta - r),$$

Estimación: MC Restringidos (cont).

- c.1^{er}o. $\rightsquigarrow$ ecuaciones normales:

$$X'\,\widehat{u}_R + R'\,\widehat{\lambda} = 0, \tag{4}$$

$$R\,\widehat{\beta}_R = r, \tag{5}$$

donde $\widehat{\beta}_R$ y $\widehat{\lambda}$ son valores de β, λ que satisfacen las c.1^{er}o. y los residuos

$$\widehat{u}_R = Y - X\,\widehat{\beta}_R. \tag{6}$$

- Despejando $\widehat{\beta}_R$ y $\widehat{\lambda}$:

$$\widehat{\lambda} = [R\,(X'X)^{-1}R']^{-1}(r - R\widehat{\beta}),$$

$$\widehat{\beta}_R = \widehat{\beta} + (X'X)^{-1}R'[R\,(X'X)^{-1}R']^{-1}(r - R\widehat{\beta})$$

$$= \widehat{\beta} + A(r - R\widehat{\beta}) = (I - AR)\,\widehat{\beta} + Ar \tag{7}$$

donde $A = (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}$.

Estimación MCR: características

- Expresión (7): $\widehat{\beta}_R = \widehat{\beta} + A(r - R\widehat{\beta}) \rightsquigarrow$
  - la estimación restringida $\widehat{\beta}_R$ puede obtenerse como una función de la estimación ordinaria (no restringida): $\widehat{\beta}$
  - $R\widehat{\beta} \simeq r \quad \Rightarrow \quad \widehat{\beta}_R$ (restringido) $\simeq \widehat{\beta}$ (no restringido) .
- <2>Ecuaciones normales (4): $X'\,\widehat{u}_R + R'\,\widehat{\lambda} = 0 \rightsquigarrow$
  - se satisfacen las restricciones (obvio).
  - $X'\widehat{u}_R \neq 0$, *es decir*:
    - suma de residuos restringidos no es cero,
    - residuos restringidos no ortogonales a las variables explicativas,
    - entonces, residuos restringidos no ortogonales a $\widehat{Y}_R$ ajustado.
  - SCT $\neq$ SCR$_R$ + SCE$_R$
    (compárese con caso ordinario y con ecuación transformada: ¿ R^2 ??).

Propiedades del estimador MCR (1)

Expresión (7) : $\widehat{\beta}_R = (I - AR)\,\widehat{\beta} + Ar \quad \rightsquigarrow$

1. **Lineal:** estimador MCR $\widehat{\beta}_R$ es c.l. del estimador MCO $\widehat{\beta}$, que es lineal, luego, a su vez, $\widehat{\beta}_R$ es también lineal

2. **Sesgo:** estimador MCR $\widehat{\beta}_R$ es $\begin{cases} \text{sesgado,} & \text{si } R\beta \neq r \text{ ,} \\ \text{insesgado,} & \text{si } R\beta = r \text{ cierto} \end{cases}$

   Demo:

$$\mathrm{E}\big(\widehat{\beta}_R\big) = (I - AR)\,\mathrm{E}\big(\widehat{\beta}\big) + Ar = (I - AR)\,\beta + Ar = \beta + A(r - R\beta).$$

3. **Matriz de Covarianzas:** $\mathrm{Var}\big(\widehat{\beta}_R\big) = (I - AR)\mathrm{Var}\big(\widehat{\beta}\big) = \sigma^2 (I - AR)(X'X)^{-1}$
   Demo: (ver apuntes o libro de texto)

Propiedades del estimador MCR (2)

4. **Varianza menor** que los estimadores MCO,
   aunque las restricciones no sean ciertas:

   Demo:

$$\text{Var}\big(\widehat{\beta}_R\big) = \text{Var}\big(\widehat{\beta}\big) - AR\,\text{Var}\big(\widehat{\beta}\big) = \text{Var}\big(\widehat{\beta}\big) - \text{(matriz psd)}.$$

$\square$

5. resultado sorprendente (aparentemente):
   - menos "incertidumbre" sobre parámetros $\rightsquigarrow$ más precisión en estimación. . .
   - pero. . . hacía un resultado erróneo (sesgado) si la restricción no es cierta.

Multicolinealidad frente a restricciones

Es importante distinguir claramente entre dos casos diferentes:

- relaciones lineales entre regresores
  (*es decir* multicolinealidad):
  $$\textit{p.ej.}\quad X_{4t} = 2X_{1t}$$
  $\Rightarrow$ falta información para estimaciones individuales.
- relaciones lineales entre coeficientes:
  $$\textit{p.ej.}\quad \beta_4 = 2\,\beta_1$$
  $\Rightarrow$ información extra sobre parámetros
  $\rightsquigarrow$ estimadores con varianza más pequeña.
- modelos a estimar respectivanmente:

$$Y_t = \beta_0 + (\underbrace{\beta_1 + 2\beta_4}_{\beta_1^*})X_{1t} + \beta_2 X_{2t} + + \cdots + u_t,$$

$\Rightarrow\ \widehat{\beta}_1^*$ pero ¿ $\widehat{\beta}_1, \widehat{\beta}_4$?

$$Y_t = \beta_0 + \beta_1(\underbrace{X_{1t} + 2X_{4t}}_{X_{1t}^*}) + \beta_2 X_{2t} + + \cdots + u_t,$$

$\Rightarrow\ \widehat{\beta}_1$ y $\widehat{\beta}_4 = 2\,\widehat{\beta}_1$

3 El Modelo de Regresión Lineal (II). Inferencia y Predicción.

3.1a Distribución del Estimador de Mínimo Cuadrados bajo el supuesto de Normalidad.

Estimador MCO bajo Normalidad

- Si $Y = X\beta + u$, donde $\boxed{u \sim \mathcal{N}(0, \sigma^2 I_T)}$, entonces (recuérdese) el estimador MCO:

$$\begin{aligned} \widehat{\beta}_{\text{MCO}} &= (X'X)^{-1}X'Y = \beta + (X'X)^{-1}X'u \\ &= \beta + \Gamma' u \quad \text{es lineal en perturbaciones.} \end{aligned}$$

- Por lo tanto, la misma distribución Multivariante Normal, con (recuérdese)

$$\begin{cases} \mathrm{E}(\widehat{\beta}) & = \beta, \\ \mathrm{Var}(\widehat{\beta}) & = \sigma^2 (X'X)^{-1}. \end{cases}$$

- Es decir:

$$\boxed{\widehat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1})}$$

Estimador MCO bajo Normalidad (casos)

Dado que $\boxed{\widehat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1})}$:

- Para el k-ésimo coeficiente:

$$\widehat{\beta}_k \sim \mathcal{N}(\beta_k, \sigma^2 a_{kk})$$

donde a_{kk} es el $(k+1)$-ésimo elemento diagonal de $(X'X)^{-1}$

- *por ejemplo:* $\widehat{\beta}_1 \sim \mathcal{N}(\beta_1, \sigma^2 a_{11})$,

$a_{11} = 2^o$ elemento diagonal.

- Para un conjunto de combinaciones lineales:

$$R\widehat{\beta} \sim \mathcal{N}(R\beta, \sigma^2 R(X'X)^{-1}R').$$

- Para un subvector de $\widehat{\beta}$: $R = [0_s \dots 0_s | I_s]$; entonces

$$\widehat{\beta}^s \sim \mathcal{N}(\beta^s, \sigma^2 A_{ss})$$

donde $\beta^s =$ subvector de β, $A_{ss} =$ submatriz de $(X'X)^{-1}$.

Estimador MCO bajo Normalidad (casos)2

- En particular, si $R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow$

  $R \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \beta^\star$ (sin intercepto):
- y

$$(X'X)^{-1} = \begin{pmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{pmatrix};$$

- entonces

$$\widehat{\beta}^\star \sim \mathcal{N}(\beta^\star, \sigma^2 \diamond)$$

Residuos MCO bajo Normalidad

- De forma similar, si $\boxed{u \sim \mathcal{N}(0, \sigma^2 I_T)}$,

  Entonces,

$$\boxed{\widehat{u} \sim \mathcal{N}(0, \sigma^2\, M)}$$

- En particular, para el 4^o residuo:

$$\widehat{u}_t \sim \mathcal{N}(0, \sigma^2\, m_{44})$$

  donde m_{44} es el 4^o elemento diagonal de la matriz M.

3.1b Contrastar Hipótesis: un Repaso.

Hipótesis y Contrastes (rep1)

- Punto de partida:

$$\left.\begin{array}{ll} Y & = X\beta + u \\ u & \sim \mathcal{N}(0, \sigma^2 I_T) \end{array}\right\} \left\{\begin{array}{l} \widehat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1}) \\ \widehat{u} \sim \mathcal{N}(0, \sigma^2 M) \end{array}\right.$$

- Hipótesis: “conjetura sobre la fn de dn de(l) parámetro(s)”.
  Por ejemplo:
  - en MRLS: $\widehat{\beta} \sim \mathcal{N}(\beta, v)$; sea $\beta = 2{,}5$.
  - en MRLG: $\widehat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1})$; sea $\beta_1 + \cdots + \beta_K = 1$.
  - en general: T^a Ec. $\rightsquigarrow$ hipótesis
    p.ej.: Fn Cobb-Couglas:

$$Y_t = e^{\beta_0} L_t^{\beta_1} K_t^{\beta_2} e^{u_t}$$

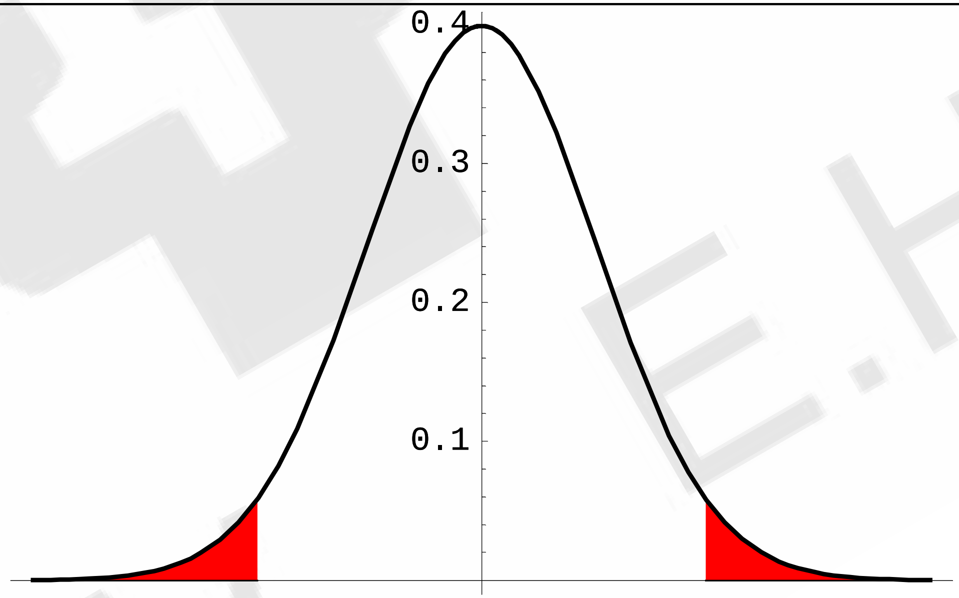
    con rendimientos de escala constantes : $\beta_1 + \beta_2 = 1$

- Contraste: “procedimiento para rechazar o aceptar la hipótesis”

Hipótesis y Contrastes (rep2)

	elementos	pasos
a)	hipótesis para contrastar (sobre estimador)	$H_0 : \ldots$ frente a $H_a : \ldots$ (disjunta)
b)	dn estimador	obtener estadístico de contraste con dn tabulada bajo H_0 :
c)	regla decisión	$\overbrace{\begin{array}{cc} \in \text{región crítica} & \notin \text{región crítica} \\ (\text{“grande”}) & (\text{“pequeño”}) \\ \Downarrow & \Downarrow \\ \text{Rechazar} & \text{no Rechazar} \end{array}}^{\text{estadístico calculado}}$

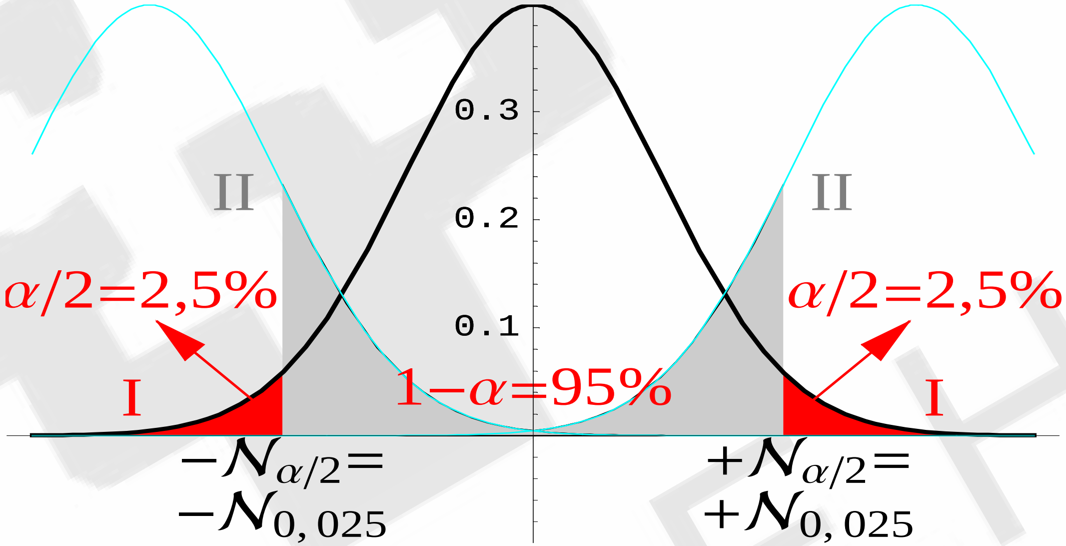
Ejemplo:

a)	$H_0 : \beta = 2{,}5$ frente a $H_a : \beta \neq 2{,}5$ (Var(β)=4)
b)	$\widehat{\beta} \sim \mathcal{N}(\beta, 4) \rightsquigarrow z = \dfrac{\widehat{\beta} - \beta}{2} \sim \mathcal{N}(0,1)$
c)	$z = \dfrac{\widehat{\beta} - 2{,}5}{2} \in$ 

Hipótesis y Contrastes: Región Crítica

nivel de significación $\alpha = 5\,\% = 0{,}05$

región crítica	región de confianza	región crítica

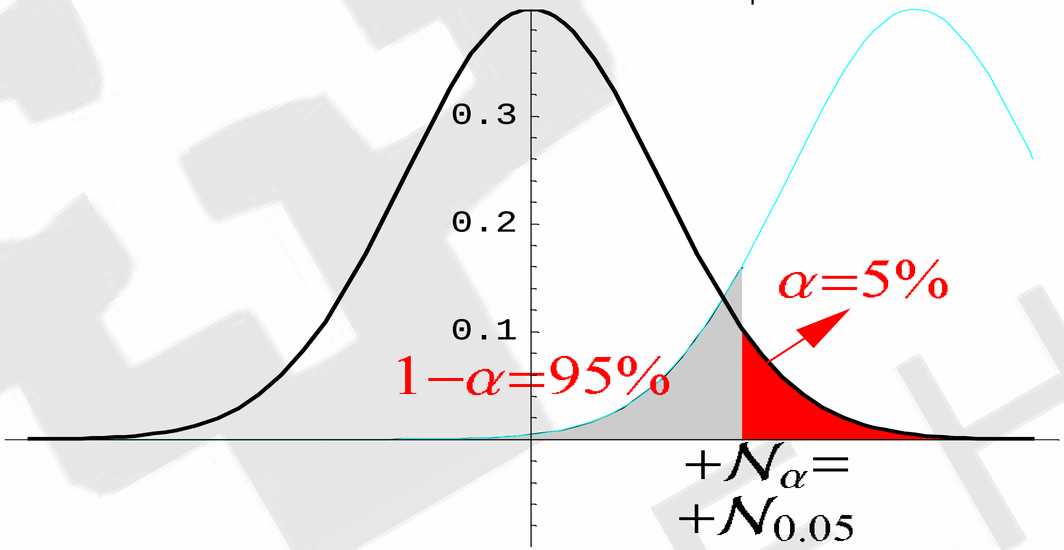


región crítica	región de confianza	región crítica

Hipótesis y Contrastes: Región crítica (una cola)

nivel de significación $\alpha = 5\,\% = 0{,}05$

región de confianza | región crítica



región de confianza | región crítica

Hipótesis y Contrastes: Distribuciones (repaso)

1. Def de χ^2 (ji-cuadrado):

$$\left.\begin{array}{c} Z_i \sim \text{iid}\,\mathcal{N}(0,1) \\ Z \sim \mathcal{N}(0, I_m) \end{array}\right\} Z'Z = \sum_{i=1}^{m} Z_i^2 \sim \chi^2(m) \qquad \begin{cases} \mathrm{E}\big(\chi^2(m)\big) = m \\ \mathrm{Var}\big(\chi^2(m)\big) = 2m \end{cases}$$

1b. $Z \sim \mathcal{N}(\mu, \Omega) \Rightarrow (Z-\mu)'\Omega^{-1}(Z-\mu) \sim \chi^2(m)$

2. Def de t (Student): $\left.\begin{array}{c} Z \sim \mathcal{N}(0,1), \quad W \sim \chi^2(m) \\ Z, W \text{ independientes} \end{array}\right\} \dfrac{Z}{\sqrt{W/m}} \sim t(m)$

3. Def de $\mathcal{F}$ (Snedecor): $\left.\begin{array}{c} V \sim \chi^2(n) \quad W \sim \chi^2(m) \\ V, W \text{ independientes} \end{array}\right\} \dfrac{V/n}{W/m} \sim \mathcal{F}_m^n$

4b. $n = 1 \Rightarrow \dfrac{Z^2}{W/m} \sim \mathcal{F}_m^1 \equiv t(m)^2$

Hipótesis y Contrastes: Resultado útil

De $\widehat{u} \sim \mathcal{N}(0, \sigma^2 M)$:

- $\dfrac{\text{SCR}}{\sigma^2} = \sum(\widehat{u}_t^2/\sigma^2) = \sum \mathcal{N}(0,1)^2\text{'s} \sim \chi^2(T\!-\!K\!-\!1)$
- Entonces: $\dfrac{\widehat{\sigma}^2}{\sigma^2} = \dfrac{\text{SCR}}{\sigma^2(T\!-\!K\!-\!1)} = \dfrac{\text{SCR}}{\sigma^2(T\!-\!K\!-\!1)} = \chi^2/\text{d.f.'s}$
  - $\dfrac{\text{expr}}{\sigma} \sim \mathcal{N}(0,1)$:
  - $\dfrac{\text{expr}}{\widehat{\sigma}} = \dfrac{\text{expr}/\sigma}{\widehat{\sigma}/\sigma} = \dfrac{\text{expr}/\sigma}{\sqrt{\widehat{\sigma}^2/\sigma^2}} = \dfrac{\mathcal{N}(0,1)}{\sqrt{\chi^2/\text{d.f.'s}}} = \boldsymbol{t}$
  - $\dfrac{\text{expr}}{\sigma^2} \sim \chi^2(n)$:
  - $\dfrac{\text{expr}}{\widehat{\sigma}^2} = \dfrac{\text{expr}/\sigma^2}{\widehat{\sigma}^2/\sigma^2} \Rightarrow \dfrac{\frac{\text{expr}}{\sigma^2}/n}{\widehat{\sigma}^2/\sigma^2} = \dfrac{\chi^2(n)/n}{\chi^2/\text{d.f.'s}} \sim \mathcal{F}$
- En resumen: $\sigma^2 \to \widehat{\sigma}^2 \quad \Rightarrow$ ¡ $\mathcal{N}(0,1) \to \boldsymbol{t}$!!
  ¡ $\chi^2 \to \mathcal{F}$!!

3.2a Contraste para Significación de un único parámetro. Intervalos de Confianza.

Contraste de Signif de un parámetro: dn

- Estandarizar $\widehat{\beta}_i \sim \mathcal{N}(\beta_i, \sigma^2 a_{ii})$

$$\frac{\widehat{\beta}_i - \beta_i}{\sqrt{\text{Var}(\widehat{\beta}_i)}} = \frac{\widehat{\beta}_i - \beta_i}{\sigma\sqrt{a_{ii}}} = \frac{\widehat{\beta}_i - \beta_i}{\sigma_{\widehat{\beta}_i}} \sim \mathcal{N}(0,1)$$

- cambiar σ por $\widehat{\sigma}$:

$$\frac{\widehat{\beta}_i - \beta_i}{\widehat{\sigma}\sqrt{a_{ii}}} = \frac{\widehat{\beta}_i - \beta_i}{\sqrt{\widehat{\text{Var}(\widehat{\beta}_i)}}} = \boxed{\frac{\widehat{\beta}_i - \beta_i}{S_{\widehat{\beta}i}} \sim t_{(T-K-1)}}$$

- Nota: $\sigma_{\widehat{\beta}i} \to S_{\widehat{\beta}i} \quad \Rightarrow \quad$ ¡ $\mathcal{N}(0,1) \to t$!!

Contraste de Signif de un parámetro: regla

- $$\boxed{\frac{\widehat{\beta}_i - \beta_i}{S_{\widehat{\beta}i}} \sim \boldsymbol{t}\,(T\!-\!K\!-\!1)}$$
- ¿ Qué Contraste ? $\begin{cases} H_0 : \beta_i = c & \text{(contraste informativo)} \\ H_0 : \beta_i = 0 & \text{(contraste de significatividad)} \end{cases}$
- **Recuérdese:** Hipótesis $\rightsquigarrow$ estadístico $\rightsquigarrow$ regla. . .
- Contraste de Significatividad:
  - Hipótesis: $H_0 : \beta_i = 0$ frente a $H_a : \beta_i \neq 0$
  - Estadístico: $t = \dfrac{\widehat{\beta}_i}{S_{\widehat{\beta}i}} \sim \boldsymbol{t}\,(T\!-\!K\!-\!1)$ bajo H_0 :
  - Regla: $|t| = |\dfrac{\widehat{\beta}_i}{S_{\widehat{\beta}i}}| > \boldsymbol{t}_{\alpha/2}(T\!-\!K\!-\!1) \quad \Rightarrow$ rechazar H_0 :

    $\Rightarrow$ β_i es (estadísticamente o significativamente) distinto de cero

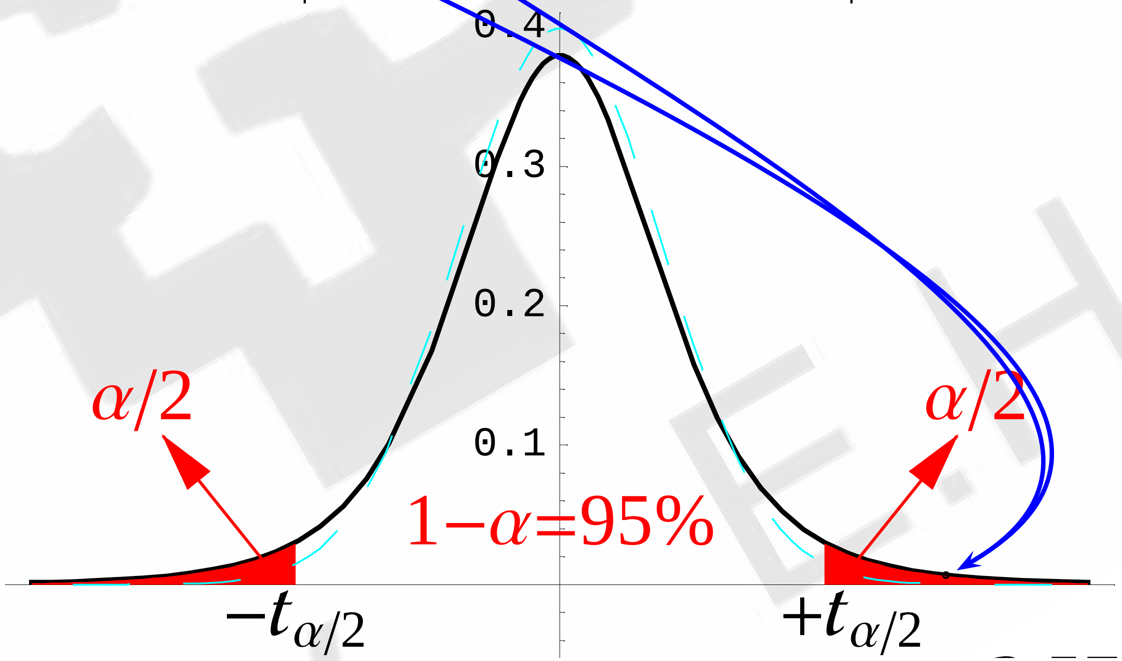
    $\Rightarrow$ X_i es una variable (estadísticamente) relevante o significativa.
- de forma similar para el contraste informativo $H_0 : \beta_i = c$ (Ejercicio: ¡ Inténtalo !!)

Contraste de Signif de un parámetro: regla (cont)

- Regla: $|t| = |\frac{\widehat{\beta}_i}{S_{\widehat{\beta}_i}}| > t_{\alpha/2(T-K-1)} \quad \Rightarrow \quad$ rechazar H_0 :
-

nivel de significación $\alpha = 5\,\% = 0{,}05$

región crítica	región de confianza	región crítica

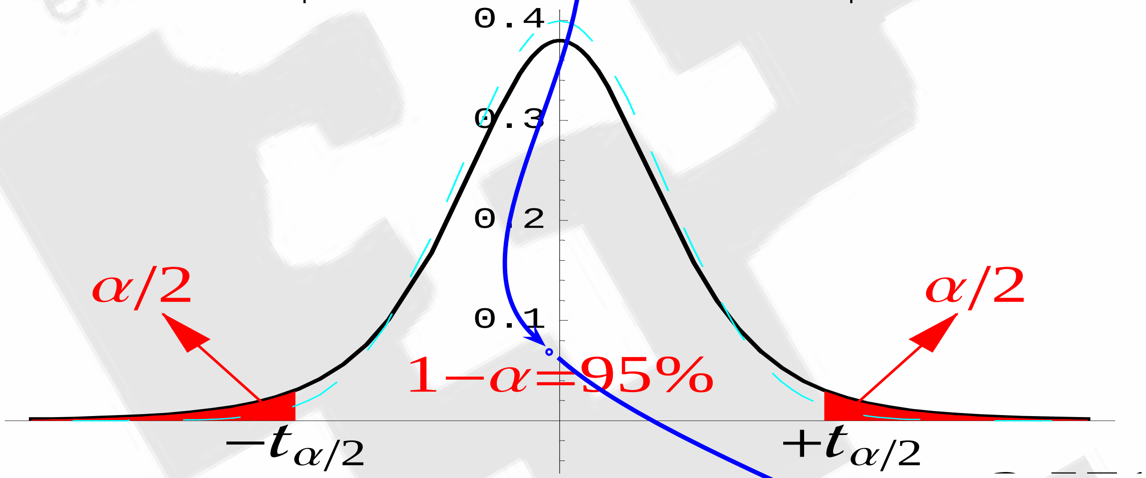


región crítica	región de confianza	región crítica

Intervalo de confianza para β_i

- Recuérdese que $\dfrac{\widehat{\beta}_i - \beta_i}{S_{\widehat{\beta}i}} \sim t_{(T-K-1)}$

región crítica | región de confianza | región crítica



región crítica	región de confianza	región crítica

es decir: $\Pr[-t_{\alpha/2} \leq \dfrac{\widehat{\beta}_i - \beta_i}{S_{\widehat{\beta}i}} \leq +t_{\alpha/2}] = 1 - \alpha$

$$\Pr[\underbrace{\widehat{\beta}_i - t_{\alpha/2}\, S_{\widehat{\beta}i} \leq \beta_i \leq \widehat{\beta}_i + t_{\alpha/2}\, S_{\widehat{\beta}i}}_{\text{IC}_{1-\alpha}(\beta_i)}] = 1 - \alpha$$

Intervalo de confianza para β_i (cont)

- Es decir:

$$\text{IC}_{1-\alpha}(\beta_i) = [\widehat{\beta}_i \pm t_{\alpha/2}\, S_{\widehat{\beta}i}]$$

- *p.ej.* para $\alpha = 5\,\%$, $T–K–1 = 25$, $\widehat{\beta}_i = 2{,}12$ y $S_{\widehat{\beta}i} = 0{,}08$:

$$\text{IC}_{95\,\%}(\beta_i) = [\widehat{\beta}_i \pm t_{2{,}5\,\%}(25)\, S_{\widehat{\beta}i}]$$
$$= [\widehat{\beta}_i \pm 2{,}06\, S_{\widehat{\beta}i}] = [2{,}12 \pm 2{,}06 \cdot 0{,}08] = [1{,}9552\,;\ 2{,}2848]$$

- contraste por medio del intervalo de confianza:

Hipótesis:	$H_0 : \beta_i = c$ frente a $H_a : \beta_i \neq c$
Intervalo:	$\text{IC}_{95\,\%}(\beta_i)$
Regla:	Rechazar H_0 : si $c \notin \text{CI}_{95\,\%}(\beta_i)$, con significación del 5 %

- *p.ej.* ¿ $H_0 : \beta_i = 0$? $\Rightarrow$ Rechazar $\Rightarrow$ β_i es significativo (al nivel del 5 %).

Contraste de única Combinación Lineal

- Supongamos MRLG restringido con 1 restricción ($q = 1$):
  $R\beta = r$ pero ahora más sencillo. . .
  $R = d'$ (cualquier fila de $K+1$ valores $d_0, d_1, \ldots, +d_K$) y
  $r = c$ (cualquier valor escalar):
- Sea $H_0 : \nu = d'\beta = d_0\beta_0 + d_1\beta_1 + \cdots + d_K\beta_K = c$
  es decir,
  un contraste informativo sobre el valor c que toma una única combinación lineal ν de los parámetros.

Contraste de única Combinación Lineal: Ejemplo

- Sea la fn de Cobb-Douglas linealizada

$$\log Y_t = \alpha + \beta_L \log L_t + \beta_K \log K_t + u_t$$

- $d' = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$ y $c = 1$:

$$\begin{aligned} H_0 : \nu = d'\beta &= \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} \alpha \\ \beta_L \\ \beta_K \end{pmatrix} & &= \beta_L + \beta_K \\ &= c & &= 1 \end{aligned}$$

es decir, $H_0 : \beta_L + \beta_K = 1$;
el contraste de la hipótesis de rendimientos de escala constantes.

Contraste de única Combinación Lineal: dn

- Dado que $\widehat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1})$, tenemos que

$$d'\widehat{\beta} \sim \mathcal{N}(d'\beta, \sigma^2 d'(X'X)^{-1} d)$$

$$\widehat{\nu} \sim \mathcal{N}(\nu, \text{Var}(\widehat{\nu}))$$

donde $\text{Var}(\widehat{\nu}) = \sigma^2 \sum_{i,j=0}^{K} d_i d_j a_{ij}$

- Como antes, estandarizando $\widehat{\nu}$

$$\frac{\widehat{\nu} - \nu}{\sqrt{\text{Var}(\widehat{\nu})}} \sim \mathcal{N}(0,1)$$

- Por lo tanto (recuérdese que $\sigma \to \widehat{\sigma}$):

$$\Rightarrow \quad \boxed{\frac{\widehat{\nu} - \nu}{S_{\widehat{\nu}}} \sim t_{(T-K-1)}}$$

donde $S_{\widehat{\nu}} = \widehat{\sigma} \sqrt{\sum_{i,j=0}^{K} d_i d_j a_{ij}}$.

Contraste de única Combinación Lineal: regla

- $\boxed{\dfrac{\widehat{\nu}-\nu}{S_{\widehat{\nu}}} \sim \boldsymbol{t}_{(T–K–1)}}$
- ¿ Qué Contraste ? $\Big\{ H_0 : \nu (= d'\beta) = c$ (contraste informativo)
- **Recuérdese:** Hipótesis $\rightsquigarrow$ estadístico $\rightsquigarrow$ regla. . .
- Contraste de una combinación lineal:
  - Hipótesis: $H_0 : \nu = c$ frente a $H_a : \nu \neq c$
  - Estadístico:

$$t = \frac{\widehat{\nu}-c}{S_{\widehat{\nu}}} \sim \boldsymbol{t}_{(T–K–1)} \text{ bajo } H_0 :$$

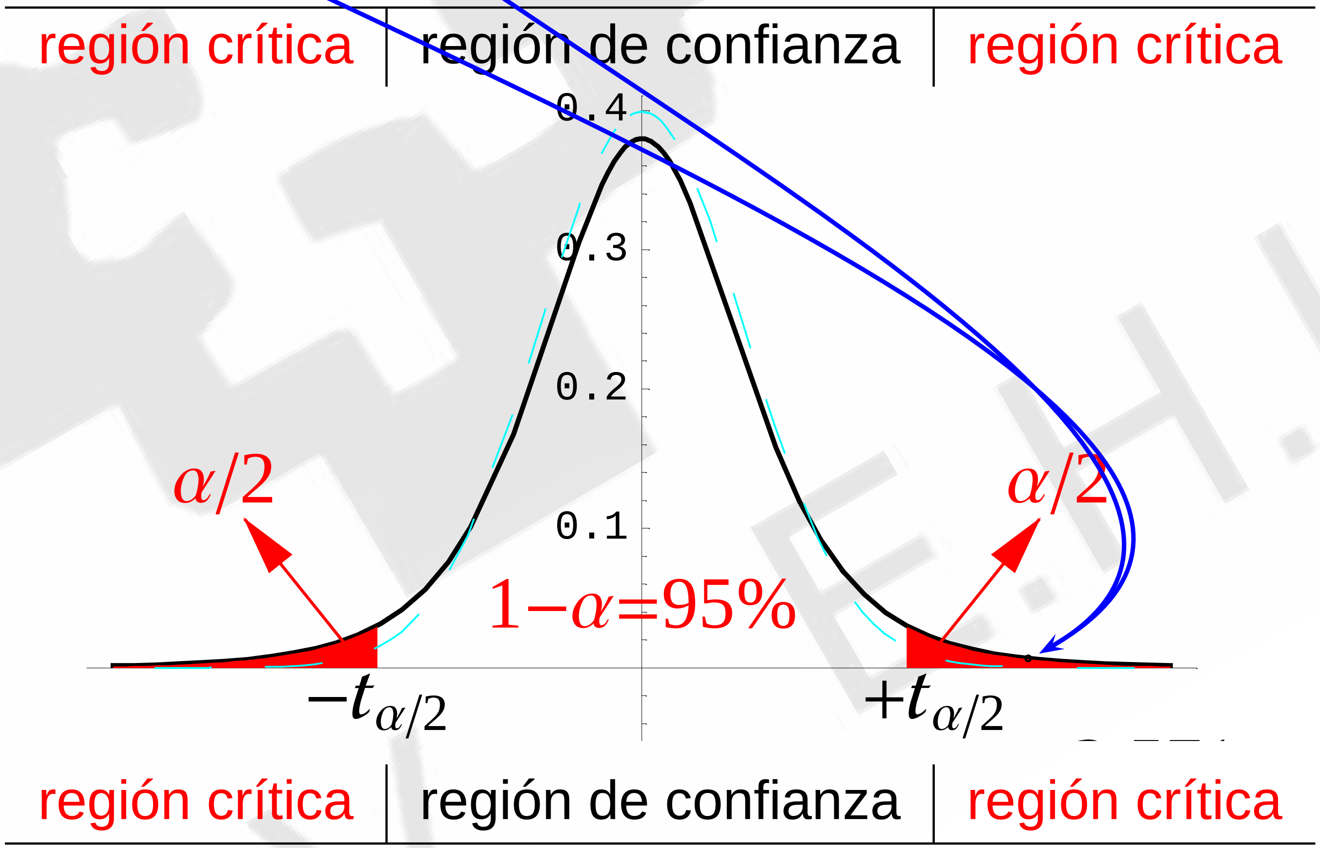
  - Regla: $\quad |t| > \boldsymbol{t}_{\alpha/2}(T–K–1) \quad \Rightarrow \quad$ rechazar H_0 :

    $\Rightarrow$ valor de la combinación lineal no es correcta.
  - *compárese* con el contraste de un único parámetro β_k, ¿ algunas similitudes ?.

Contraste de única Combinación Lineal: regla (c

- Regla: $|t| = \left| \dfrac{\widehat{\nu} - c}{S_{\widehat{\nu}}} \right| > t_{\alpha/2(T-K-1)} \quad \Rightarrow \quad$ rechaza H_0 :
-

nivel de significación $\alpha = 5\,\% = 0{,}05$



Contraste de única Combinación Lineal: Ejemplo

- En la fn de Cobb-Douglas linealizada:
  $\widehat{\log Y}_t = \widehat{\alpha} + \widehat{\beta}_L \log L_t + \widehat{\beta}_K \log K_t, \quad T = 53;$
- $\widehat{\log Y}_t = 2{,}10 + 0{,}67 \log L_t + 0{,}32 \log K_t, \quad \widehat{\sigma}^2 = 4$;

$$(X'X)^{-1} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & -1 \\ 0 & -1 & 7 \end{pmatrix}$$

- Contrastar la H_0 : rendimientos de escala constantes al nivel de significación del $\alpha = 5\,\%$:

Contraste de única Combinación Lineal: Ej (cont

- Hipótesis: $H_0 : \beta_L + \beta_K = 1$ frente a $H_a : \beta_L + \beta_K \neq 1$
- Estadístico:
  - $$\begin{aligned}\widehat{\nu} &= \widehat{\beta}_L + \widehat{\beta}_K \\ &= 0{,}67 + 0{,}27 = 0{,}89\end{aligned}$$
  - $$\begin{aligned}S_{\widehat{\nu}} &= \sqrt{\widehat{\text{Var}}(\widehat{\beta}_L) + \widehat{\text{Var}}(\widehat{\beta}_K) + 2\widehat{\text{Cov}}(\widehat{\beta}_L, \widehat{\beta}_K)} \\ &= \widehat{\sigma}\sqrt{a_{11} + a_{22} + 2a_{12}} \\ &= 2\sqrt{4 + 7 + 2(-1)} = 2\sqrt{9} = 6\end{aligned}$$
  - $$\begin{aligned}t &= \frac{\widehat{\nu} - 1}{S_{\widehat{\nu}}} \\ &= \frac{0{,}89 - 1}{6} = \frac{-0{,}11}{6} = -0{,}018.\end{aligned}$$
- Regla: $|t| = 0{,}018 < t_{0{,}025}(50) = 2{,}01 \quad \Rightarrow$ no rechazar H_0 :

$\Rightarrow$ “rendimientos de escala constantes” es apoyado por datos.

3.2b Contraste de Significatividad Conjunta.

Contraste de Significatividad Conjunta: dn

- $H_0 : \beta_1 = \beta_2 = \cdots = \beta_K = 0$
- $H_0 : \beta^\star = \mathbf{0}$
- $$\widehat{\beta}^\star \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \begin{bmatrix} a_{00} & a_{01} & \dots & a_{0K} \\ a_{10} & a_{11} & \dots & a_{1K} \\ \vdots & \vdots & \dots & \vdots \\ a_{K0} & a_{K1} & \dots & a_{KK} \end{bmatrix}\right) \sim \mathcal{N}(\mathbf{0}, \sigma^2 (x'x)^{-1})$$
- Estandarizar y construir Suma Cuadrados:
  - $$\frac{\widehat{\beta}^{\star\prime} x' x \widehat{\beta}^\star}{\sigma^2} \sim \chi^2(K) \quad \text{bajo } H_0 :$$
- Entonces (recuérdese el cambio $\sigma^2 \to \widehat{\sigma}^2$):

$$\boxed{F = \frac{\widehat{\beta}^{\star\prime} x' x \widehat{\beta}^\star / K}{\widehat{\sigma}^2} \sim \mathcal{F}^{K}_{T-K-1}}$$

Contraste de Significatividad Conjunta: regla

- $\boxed{F = \frac{\widehat{\beta}^{\star\prime} x' x \widehat{\beta}^{\star} / K}{\widehat{\sigma}^2} \sim \mathcal{F}^{K}_{T\text{–}K\text{–}1}}$ bajo H_0 :
- Contraste de significatividad conjunta: $\left\{ H_0 : \beta^\star = 0 \right.$
- **Recuérdese:** Hipótesis $\rightsquigarrow$ estadístico $\rightsquigarrow$ regla. . .
  - Hipótesis: $H_0 : \beta^\star = 0$ frente a $H_a : \beta^\star \neq 0$ (*es decir* $\exists \beta_i \neq 0$)
  - Estadístico:

$$F = \frac{\widehat{\beta}^{\star\prime}\, x' x\, \widehat{\beta}^{\star} / K}{\widehat{\sigma}^2} = \frac{\widehat{y}'\widehat{y} / K}{\widehat{u}'\widehat{u} / (T\text{–}K\text{–}1)} = \frac{\text{SCE} / K}{\text{SCR} / (T\text{–}K\text{–}1)}$$

$$= \frac{(\text{SCE}/\text{SCT}) / K}{(\text{SCR}/\text{SCT}) / (T\text{–}K\text{–}1)} = \frac{R^2 / K}{(1 - R^2) / (T\text{–}K\text{–}1)} \sim \mathcal{F}^{K}_{T\text{–}K\text{–}1} \text{ bajo } H_0 :$$

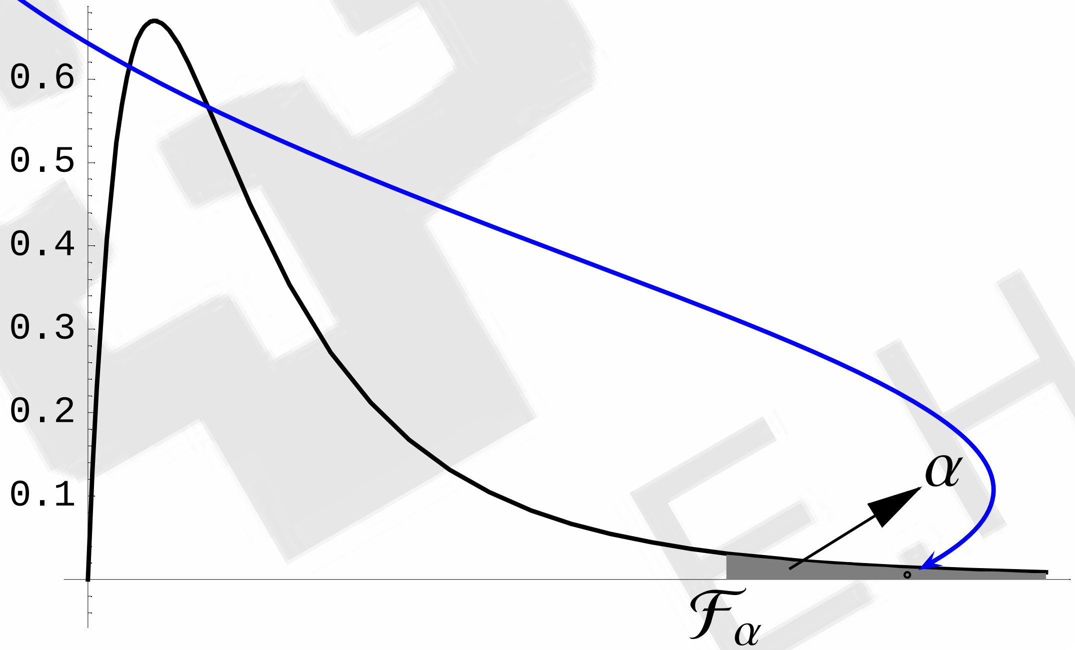
  - Regla: $F > \mathcal{F}_\alpha(K, T\text{–}K\text{–}1) \quad \Rightarrow$ rechazar H_0 :
    $\Rightarrow$ todos los coefs son conjuntamente significativos (diferentes de cero)
    $\Rightarrow$ la regresión es (estadísticamente) relevante.

Contraste de Significatividad Conjunta: regla (co

- Regla: $F > \mathcal{F}_\alpha(K, T\text{–}K\text{–}1) \quad \Rightarrow \quad$ rechazar H_0 :
-

nivel de significación $\alpha = 5\,\% = 0{,}05$

región de confianza	región crítica



región de confianza	región crítica

Contraste de Significatividad Conjunta: Ejemplo

- En el ejemplo anterior (fn de Cobb-Douglas linealizada:)

$$\widehat{\log Y}_t = \widehat{\alpha} + \widehat{\beta}_L \log L_t + \widehat{\beta}_K \log K_t, \quad T = 53;$$
$$\widehat{\log Y}_t = 2{,}10 + 0{,}67 \log L_t + 0{,}32 \log K_t, \quad \widehat{\sigma}^2 = 4; R^2 = 0{,}88$$

- Contrastar significatividad conjunta

con el $\alpha = 5\,\%$ de significación:

-

$$F = \frac{R^2 / K}{(1 - R^2)/(T{-}K{-}1)}$$
$$= \frac{0{,}88/2}{(1 - 0{,}88)/(50)} = \frac{0{,}44}{0{,}024} = 183{,}33 > \mathcal{F}_{0{,}05}(2, 50) = 3{,}19$$

$\Rightarrow$ β_K y β_L son conjuntamente significativos
$\Rightarrow$ regresión es relevante (estadísticamente).

3.3 Contraste General para Restricciones Lineales.

Contraste de Restricciones Lineales: Ejemplo 1

- Recuérdese el MRLG sujeto a q restricciones lineales:

$$\underset{(T\times 1)}{Y} = \underset{(T\times K{+}1)}{X}\ \underset{(K{+}1\times 1)}{\beta} + \underset{(T\times 1)}{u},$$

$$H_0: \underset{(q\times K{+}1)}{R}\ \underset{(K{+}1\times 1)}{\beta} = \underset{(q\times 1)}{r}.$$

- contrastes previos $\equiv$ casos especiales de R.L.s:
  1. Sea el MRLG con
     $q=1$, $R = \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \end{bmatrix}$ y $r = 0$:

$$H_0 : R\beta = \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \end{bmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_K \end{pmatrix} = \beta_2$$

$$= r \qquad\qquad = 0$$

es decir, $H_0 : \beta_2 = 0$;

el contraste de significatividad individual de X_2 .

Contraste de Restricciones Lineales: Ejemplo 2

- $H_0: \underset{(q\times K+1)}{R} \; \underset{(K+1\times 1)}{\beta} = \underset{(q\times 1)}{r}$.

3. Sean $q=2$ restricciones tales que

$$R = \begin{bmatrix} 0 & 2 & 3 & 0 & \dots & 0 \\ 1 & 0 & 0 & -2 & \dots & 0 \end{bmatrix} \text{ y } r = \begin{bmatrix} 5 \\ 3 \end{bmatrix}:$$

-

$$H_0: R\beta = \begin{bmatrix} 0 & 2 & 3 & 0 & \dots & 0 \\ 1 & 0 & 0 & -2 & \dots & 0 \end{bmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_K \end{pmatrix} = \begin{bmatrix} 2\beta_1 + 3\beta_2 \\ \beta_0 - 2\beta_3 \end{bmatrix}$$

$$= r = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

es decir, el MRLG bajo $H_0: \begin{cases} 2\beta_1 + 3\beta_2 = 5 \\ \beta_0 - 2\beta_3 = 3 \end{cases}$

Contraste de Restricciones Lineales: Ejemplo 3

- $H_0: \underset{(q \times K+1)}{R} \underset{(K+1 \times 1)}{\beta} = \underset{(q \times 1)}{r}$.

2. Sean $q = K$ restricciones tales que

$$R = \begin{bmatrix} \mathbf{0} & \vert & \mathbf{I}_K \end{bmatrix} = \left(\begin{array}{c|cccc} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{array}\right) \text{ y } r = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

- $$H_0 : R\beta = \begin{bmatrix} \mathbf{0}_K & \vert & \mathbf{I}_K \end{bmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta^\star \end{pmatrix} = \beta^\star$$

$$= r \qquad = \mathbf{0}$$

es decir, $H_0 : \beta^\star = \mathbf{0}$;

el contraste de significatividad conjunta de la regresión.

Contrastes de Restricciones Lineales: dn

- . . . entonces, ¿ podemos tener un estadístico de contraste general que cubra todas las hipótesis de la forma

$$H_0 : \underset{(q \times K+1)}{R} \underset{(K+1 \times 1)}{\beta} = \underset{(q \times 1)}{r} \quad ?$$

- Dado que $\quad \widehat{\beta} \sim \mathcal{N}(\beta, \sigma^2(X'X)^{-1}), \quad$ tenemos que

$$R\widehat{\beta} \sim \mathcal{N}(R\beta, \sigma^2 R(X'X)^{-1}R')$$

- Como antes, estandarizando $R\widehat{\beta}$ y escribiendo la SC,

$$\frac{(R\widehat{\beta} - R\beta)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - R\beta)}{\sigma^2} \sim \chi^2(q)$$

- Por lo tanto (recouérdese que cambiar $\sigma^2 \to \widehat{\sigma}^2$):

$$\boxed{\frac{(R\widehat{\beta} - R\beta)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - R\beta)/q}{\widehat{\sigma}^2} \sim \mathcal{F}^q_{T-K-1}}$$

Contraste General de Restric Lineales: regla

- ¿ Qué Contraste ? $\left\{ H_0 : R\beta = r \right.$
- **Recuérdese:** Hipótesis $\rightsquigarrow$ estadístico $\rightsquigarrow$ regla...
- Contraste de restricciones lineales:
  - Hipótesis: $H_0 : R\beta = r$ frente a $H_a : R\beta \neq r$
  - Estadístico:

$$F = \frac{(R\widehat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - r)/q}{\widehat{\sigma}^2} \sim \mathcal{F}^q_{T\text{–}K\text{–}1} \text{ bajo } H_0 :$$

  - Regla: $F > \mathcal{F}_\alpha(q, T\text{–}K\text{–}1) \quad \Rightarrow \quad$ rechazar H_0 :

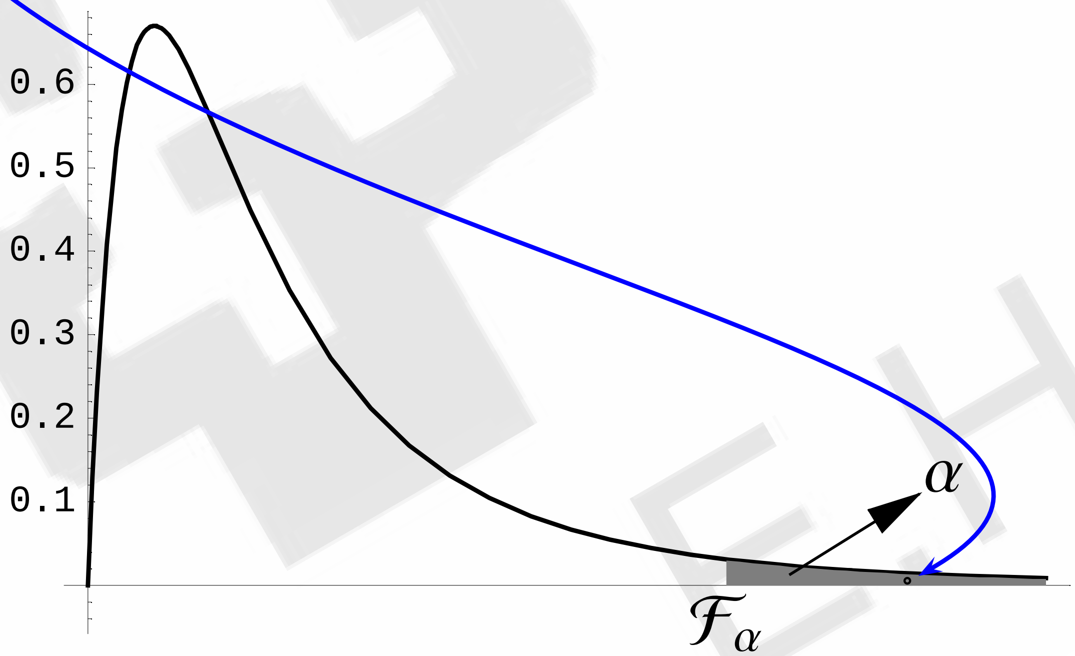
    $\Rightarrow \quad$ restricciones lineales no son (conjuntamente) ciertas.

Contraste General de Restric Lineales: regla (co

- Regla: $F > \mathcal{F}_\alpha(q, T–K–1) \quad \Rightarrow \quad$ rechazar H_0 :
-

nivel de significación $\alpha = 5\,\% = 0{,}05$

región de confianza	región critica



región de confianza	región critica

3.4 Contrastes basados en la Suma de Cuadrados Residuales.

Contraste General de Restric Lineales: regla 2

- Hipótesis: $H_0 : R\beta = r$ frente a $H_a : R\beta \neq r$
- Estadístico:

$$F = \frac{(R\widehat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - r)/q}{\widehat{\sigma}^2}$$

- utilizando el resultado $\widehat{\beta}_R = (I - AR)\widehat{\beta} + Ar$, el numerador es la diferencia entre SCs:

$$F = \frac{(\text{SCR}_R - \text{SCR})/q}{\text{SCR}/(T\!-\!K\!-\!1)} \sim \mathcal{F}^{q}_{T\text{-}K\text{-}1} \text{ bajo } H_0 :$$

- Regla: $F > \mathcal{F}_\alpha(q, T\!-\!K\!-\!1) \quad \Rightarrow$ rechazar H_0 :

  $\Rightarrow$ restricciones lineales no son (conjuntamente) ciertas.

Contraste General de Restric Lineales: Resumen

- Hipótesis: $H_0 : R\beta = r$ frente a $H_a : R\beta \neq r$
- Estadístico:

$$F = \frac{(R\widehat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - r)/q}{\widehat{\sigma}^2}$$

$$= \frac{(\text{SCR}_R - \text{SCR})/q}{\text{SCR}/(T\text{–}K\text{–}1)} \sim \mathcal{F}^q_{T\text{–}K\text{–}1} \text{ bajo } H_0 :$$

- Regla: $F > \mathcal{F}_\alpha(q, T\text{–}K\text{–}1) \quad \Rightarrow$ rechazar H_0 :
  $\Rightarrow$ restricciones lineales no son ciertas (conjuntamente).
- Nota: la forma SC necesita estimar la regresión dos veces: una con restricciones y otra sin ellas.
- y, por supuesto, también puede utilizarse para contrastar la significatividad individual, la significatividad conjunta, restricciones informativas, etc.

Contraste basado en SC: Ej Cobb-Douglas

- Hipótesis: $H_0 : \beta_L + \beta_K = 1$ frente a $H_a : \beta_L + \beta_K \neq 1$
- Estadístico:
  - $$\widehat{\nu} = \widehat{\beta}_L + \widehat{\beta}_K$$
    $$= 0{,}67 + 0{,}27 = 0{,}89$$
  - $$S_{\widehat{\nu}} = \sqrt{\widehat{\text{Var}(\widehat{\beta}_L)} + \widehat{\text{Var}(\widehat{\beta}_K)} + 2\widehat{\text{Cov}(\widehat{\beta}_L, \widehat{\beta}_K)}}$$
    $$= \widehat{\sigma}\sqrt{a_{11} + a_{22} + 2a_{12}}$$
    $$= 2\sqrt{4 + 7 + 2(-1)} = 2\sqrt{9} = 6$$
  - $$t = \frac{\widehat{\nu} - 1}{S_{\widehat{\nu}}}$$
    $$= \frac{0{,}89 - 1}{6} = \frac{-0{,}11}{6} = -0{,}018.$$
- Regla: $|t| = 0{,}018 < t_{0{,}025}(50) = 2{,}01 \quad \Rightarrow$ no rechazar H_0 :

$\Rightarrow$ la hipótesis de "rendimientos de escala constantes.es apoyada por los datos.

Contraste basado en SC: Ej Cobb-Douglas (2)

- Alternativamente, utilizar forma SC para calcular este ratio t:
  sin restricciones:
  $\log Y = \alpha + \beta_L \log L + \beta_K \log K + u, \quad \rightsquigarrow \quad \text{SCR} = 200$
- restringido: $\log Y = \alpha + \beta_L \log L + (1 - \beta_L) \log K + u$

$$\log(Y/K) = \alpha + \beta_L \log(L/K) + u, \quad \rightsquigarrow \quad \text{SCR}_R = 200{,}001296$$

- $$F = \frac{(\text{SCR}_R - \text{SCR})/q}{\text{SCR}/(T{-}K{-}1)}$$

$$= \frac{(200{,}001296 - 200)/1}{200/50} = \frac{,001296}{4} = 0{,}000324$$

$$< \mathcal{F}_{0,05}(1,50) = 4{,}04$$

- o (recuérdese que $t\,(m) = \sqrt{\mathcal{F}\,(1,m)}$)

$$t = \sqrt{F} = \sqrt{0{,}000324} = 0{,}018$$

$$< t_{0,05}(50) = 2{,}01$$

Contraste General: Ejemplo 2

- MRLG con $q = 2$, $R = \begin{bmatrix} 0 & 2 & 3 & 0 & \dots & 0 \\ 1 & 0 & 0 & -2 & \dots & 0 \end{bmatrix}$ y $r = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$:

$$R\widehat{\beta} = \begin{bmatrix} d_1'\widehat{\beta} \\ d_2'\widehat{\beta} \end{bmatrix} = \begin{bmatrix} 2\widehat{\beta}_1 + 3\widehat{\beta}_2 \\ \widehat{\beta}_0 - 2\widehat{\beta}_3 \end{bmatrix}$$

-

$$R(X'X)^{-1}R' = \begin{bmatrix} d_1'(X'X)^{-1}d_1 & d_1'(X'X)^{-1}d_2 \\ d_2'(X'X)^{-1}d_1 & d_2'(X'X)^{-1}d_2 \end{bmatrix}$$

$$= \begin{bmatrix} 4a_{11} + 9a_{22} + 12a_{12} & 2a_{10} - 4a_{13} + 3a_{02} - 6a_{23} \\ & a_{00} + 4a_{33} - 4a_{03} \end{bmatrix}$$

- Por lo tanto $F =$

$$\frac{\begin{bmatrix} 2\beta_1 + 3\beta_2 - 5 & \beta_0 - 2\beta_3 - 3 \end{bmatrix} \begin{bmatrix} 4a_{11} + 9a_{22} + 12a_{12} & 2a_{10} - 4a_{13} + 3a_{02} - 6a_{23} \\ & a_{00} + 4a_{33} - 4a_{03} \end{bmatrix}^{-1} \begin{bmatrix} 2\beta_1 + 3\beta_2 - 5 \\ \beta_0 - 2\beta_3 - 3 \end{bmatrix} /2}{\widehat{\sigma}^2}$$

$\sim \mathcal{F}^2_{T-K-1}$ bajo H_0 :

- *es decir*, un estadístico "F" para contrastar dos restricciones lineales conjuntamente.

Contraste General: Ejemplo 2

- Alternativamente (más fácil), utilizar forma SC para calcular este estadístico F:

$$H_0 : \begin{cases} 2\beta_1 + 3\beta_2 = 5 \\ \beta_0 - 2\beta_3 = 3 \end{cases}$$

$$\beta_1 = \frac{5 - 3\beta_2}{2}, \qquad \beta_0 = 3 + 2\beta_3$$

- sin restricciones:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \cdots + u \ \rightsquigarrow \text{SCR}$$

- restringido:

$$Y = (3 + 2\beta_3) + (2{,}5 - 1{,}5\beta_2) X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \cdots + u$$

$$\underbrace{Y - 3 - 2{,}5X_1}_{Y^*} = \beta_2 (\underbrace{X_2 - 1{,}5X_1}_{X_2^*}) + \beta_3 (\underbrace{X_3 + 2}_{X_3^*}) + \beta_4 X_4 \cdots + u$$

$$Y^* = \beta_2 X_2^* + \beta_3 X_3^* + \beta_4 X_4 \cdots + u \ \rightsquigarrow \text{SCR}_R$$

- y $F = \frac{(\text{SCR}_R - \text{SCR})/q}{\text{SCR}/(T-K-1)}$, etc.

Contraste General: Ejemplo 3

- MRLG con $q = K$, $R = \begin{bmatrix} \mathbf{0}_K & \vline & \mathbf{I}_K \end{bmatrix}$ y $r = \mathbf{0}_K$:

$$R\widehat{\beta} \quad \rightsquigarrow \text{selecciona} \quad \beta^\star$$

$$R(X'X)^{-1}R' \quad \rightsquigarrow \text{selecciona} \quad \begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0K} \\ a_{10} & a_{11} & \cdots & a_{1K} \\ \vdots & \vdots & \cdots & \vdots \\ a_{K0} & a_{K1} & \cdots & a_{KK} \end{bmatrix} = (x'x)^{-1}$$

- Por lo tanto:

$$F = \frac{(\widehat{\beta}^\star - 0)'[(x'x)^{-1}]^{-1}(\widehat{\beta}^\star - 0)/K}{\widehat{\sigma}^2} = \frac{\widehat{\beta}^{\star\prime}\, x'x\, \widehat{\beta}^\star/K}{\widehat{\sigma}^2}$$

- *es decir*, el estadístico“F” habitual para contrastar la significación conjunta de la regresión.

Contraste General: Ejemplo 3

- Alternativamente, utilizar forma SC para calcular esta F:

$$\text{sin restricciones: } Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_K X_K + u \quad \rightsquigarrow \quad \text{SCR}$$

$$\text{restringido: } Y = \beta_0 + u \quad \rightsquigarrow \quad \text{SCR}_R = \text{SCT}$$

- Estadístico:

$$F = \frac{(\text{SCR}_R - \text{SCR})/q}{\text{SCR}/(T{-}K{-}1)} = \frac{(\text{SCT} - \text{SCR})/K}{\text{SCR}/(T{-}K{-}1)} = \frac{\text{SCE}/K}{\text{SCR}/(T{-}K{-}1)}$$

$$= \frac{R^2/K}{(1 - R^2)/(T{-}K{-}1)}$$

obteniendo la misma formula que antes.

3.5 Predicción por Punto e Intervalo de Predicción.

Predicción

- Capítulos anteriores: **Especificación, Estimación y Validación**.
- Este capítulo: Fase final: **Uso $=$ Predicción**.
- Punto de partida: modelo apropiado para describir el comportamiento de la variable Y:

  Especificación del Modelo

  $\Downarrow$

  Estimación de los Parámetros

  $\Downarrow$

  Validación y Contraste de Hipótesis

-

$\Downarrow$

resultado no satisfactorio	resultado satisfactorio
$\Downarrow$	$\Downarrow$
re-especificar modelo	ir a la fase de predicción

Concepto

- Series temporales: predicción (de valores futuros)
  $\Rightarrow$ Previsión
- Sección cruzada: predicción (de los valores no observados)
  $\Rightarrow$ Simulación
- En general: predicción $\Rightarrow$ respuesta a preguntas del tipo... "¿ y si... ?",
  es decir ¿ qué valor tomaría Y *si* $X = X_p$ *?*

Elementos Básicos

- Modelo o FRP:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt} + u_t$$
$$Y_t = X_t' \beta + u_t, \quad t = 1, \ldots, T.$$

- Modelo estimado o FRM:

$$\widehat{Y}_t = X_t' \widehat{\beta}, \quad t = 1, \ldots, T. \tag{8}$$

- Observación a predecir: con subíndice $p =$ (normalmente $p \notin [1, T]$):

$$Y_p = X_p' \beta + u_p. \tag{9}$$

- Perturbaciones aleatorias u_p:

$$\mathsf{E}(u_p) = 0, \quad \mathsf{E}(u_p^2) = \sigma^2, \quad \mathsf{E}(u_p u_s) = 0 \quad \forall\, s \neq p.$$

- Valor conocido del vector X_p' .

Predicción por Punto

- Sustituyendo en FRM (8):

$$\boxed{\widehat{Y}_p = X_p'\,\widehat{\beta}.} \tag{10}$$

es decir, valor numérico como aproximación al valor desconocido.

Error de Predicción

- El error cometido (al tomar $\widehat{Y}_p$ en vez del verdadero Y_p) es

$$e_p = Y_p - \widehat{Y}_p,$$

- que puede expresarse como:
- una función de las dos fuentes de error introducidas en la predicción.
- Bajo normalidad:

$$(\widehat{\beta} - \beta) \sim \mathcal{N}(0, \sigma^2 (X'X)^{-1}), \quad \text{y} \quad u_p \sim \mathcal{N}(0, \sigma^2),$$

- de modo que

$$\boxed{e_p \sim \mathcal{N}(0, \sigma_e^2)},$$

- donde la varianza del error de predicción es:

$$\sigma_e^2 = X_p' \underbrace{\text{Var}(\widehat{\beta})}_{\sigma^2(X'X)^{-1}} X_p + \underbrace{\text{Var}(u_p)}_{\sigma^2} + \underbrace{\text{Cov}(\widehat{\beta}, u_p)}_{0}$$

$$= \sigma^2 (1 + X_p'(X'X)^{-1} X_p).$$

Intervalo de Predicción

- Error de predicción estandarizado:

$$\frac{e_p - 0}{\sigma_e} = \frac{e_p}{\sigma\sqrt{1 + X_p'(X'X)^{-1}X_p}} \sim \mathcal{N}(0,1),$$

- Recuérdese como al cambiar $\sigma \to \widehat{\sigma} \quad \Rightarrow \quad$ ¡ $\mathcal{N}(0,1) \to \boldsymbol{t}$!!, luego

$$\frac{e_p}{\widehat{\sigma}_e} = \frac{e_p}{\widehat{\sigma}\sqrt{1 + X_p'(X'X)^{-1}X_p}} \sim \boldsymbol{t}_{(T-K-1)}.$$

- Por lo tanto:

$$Pr(-\boldsymbol{t}_{\alpha/2} \leq \frac{e_p}{\widehat{\sigma}_e} \leq \boldsymbol{t}_{\alpha/2}) = 1 - \alpha,$$

- y despejando Y_p :

$$Pr(\widehat{Y}_p - \widehat{\sigma}_e \boldsymbol{t}_{\alpha/2} \leq Y_p \leq \widehat{Y}_p + \widehat{\sigma}_e \boldsymbol{t}_{\alpha/2}) = 1 - \alpha.$$

- Luego, el intervalo de confianza $(1-\alpha)$ para Y_p es:

$$IC(Y_p)_{(1-\alpha)} = \left[\widehat{Y}_p \pm \widehat{\sigma}_e \boldsymbol{t}_{\alpha/2}\right],$$

el cual mide la precisión de la predicción por punto.

Predicción: Ejemplo

- en el ejemplo anterior (fn Cobb-Douglas linealizada :)

$$\widehat{\log Y}_t = \widehat{\alpha} + \widehat{\beta}_L \log L_t + \widehat{\beta}_K \log K_t, \quad T = 53;$$

$$\widehat{\log Y}_t = 2{,}10 + 0{,}67 \log L_t + 0{,}32 \log K_t, \quad \widehat{\sigma}^2 = 4$$

- "¿ Qué valor tomaría Y_p si $\log L_p = 2{,}5; \log K_p = 2{,}0$?":
- $X_p' = \begin{bmatrix} 1 & 2{,}5 & 2{,}0 \end{bmatrix}$
-

$$\widehat{\log Y_p} = X_p' \widehat{\beta} = \begin{bmatrix} 1 & 2{,}5 & 2{,}0 \end{bmatrix} \begin{bmatrix} 2{,}10 \\ 0{,}67 \\ 0{,}32 \end{bmatrix}$$

$$= 2{,}10 + 0{,}67 \cdot 2{,}5 + 0{,}32 \cdot 2{,}0 = 4{,}42$$

Predicción: Ejemplo

- Construir un IC del $95\,\%$ para el verdadero Y_p:

$$\begin{aligned}
\widehat{\sigma}_e^2 &= \sigma^2(1 + X_p'(X'X)^{-1}X_p) \\
&= 4\left(1 + \begin{bmatrix}1 & 2{,}5 & 2{,}0\end{bmatrix}\begin{pmatrix}2 & 0 & 0\\ 0 & 4 & -1\\ 0 & -1 & 7\end{pmatrix}\begin{bmatrix}1\\ 2{,}5\\ 2{,}0\end{bmatrix}\right) \\
&= 4\left(1 + \begin{bmatrix}2 & 8 & 11{,}5\end{bmatrix}\begin{bmatrix}1\\ 2{,}5\\ 2{,}0\end{bmatrix}\right) \\
&= 4\,(1 + 45) = 4 \cdot 46 = 184
\end{aligned}$$

-

$$\begin{aligned}
CI(\log Y_p)_{0{,}95} &= \left[\widehat{\log Y_p} \pm \widehat{\sigma}_e\, t_{0{,}025}(50)\right] \\
&= \left[4{,}42 \pm \sqrt{184} \cdot 2{,}01\right] \\
&= [4{,}42 \pm 27{,}25] \\
&= [-22{,}84 \quad ; \quad 31{,}68]
\end{aligned}$$

4.1 Variables Ficticias. Definición y uso en el MRLG.

Variables Ficticias: Definición

- Var explicativa cualitativa $\rightsquigarrow$ sub-muestras $T_1, T_2, \ldots$
  según categoría o características
- ejemplos:
  - vars puramente cualitativas:
    - difs individuales: sexo, raza, estado civil, etc.
    - difs temporales: estación, guerra/paz, etc.
    - difs espaciales: países, regiones, urbano/rural, etc.
  - vars cuantitativas por tramos: renta, edad, etc.
- Acordar: no podemos utilizar vars cualitativas…
  entonces substituir por **vars ficticias**…
- Def. de Variable Ficticia:

$$D_{jt} = \begin{cases} 1, & \text{si } t \in \text{categoria } j \text{ ;} \\ 0, & \text{si no.} \end{cases}$$

$$\Rightarrow D_{jt} = \mathcal{I}(t \in T_j)$$

1 VC con 2 categorías

$$\text{Consumo} = f([\text{cte}],\ \text{renta},\ \text{sexo})$$

$\text{Consumo} \rightarrow Y_t$, $[\text{cte}] \rightarrow [1]$, $\text{renta} \rightarrow R_t$, $\text{sexo} \rightarrow Masc,\ Fem$

$S_{1t} = \mathcal{I}(t \in M) \quad S_{2t} = \mathcal{I}(t \in F)$

Muestra:

t	Y	cte	R	S
1	Y_1	1	R_1	M
2	Y_2	1	R_2	F
3	Y_3	1	R_3	F
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
T	Y_T	1	R_T	M

¿ X ?

$\Rightarrow$

t	Y	cte	R	S_1	S_2
1	Y_1	1	R_1	1	0
2	Y_2	1	R_2	0	1
3	Y_3	1	R_3	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
T	Y_T	1	R_T	1	0

X

En principio: substituir VC por

tantas VFs como categorías haya.

Trampa de Var Ficticias: 1 var cualitativa

- Modelo: $Y_t = \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \gamma_2 S_{2t} + u_t$
- Problema (Trampa VF):
  X es una matriz $(T \times 4)$, pero

  $S_1 + S_2 = [1]$ (c.l. exacta) $\Rightarrow \mathrm{rk}(X) = 3 < 4$ (*es decir* MC perfecta)

- $\Rightarrow \det(X'X) = 0$
  $\Rightarrow$ ¡ $(X'X)^{-1}$ no existe !! y
  ¡ no se puede calcular $\widehat{\beta}$!!

- Solución General: eliminar UNA de las cols. que causa el problema:[1] o S_1 o S_2 .
- (POSIBLE Solución: eliminar intercepto . . . pero. . .

Solución: VF sin categoría

SOLUCIÓN MÁS HABITUAL: eliminar categoría: *p.ej.* F (S_2):

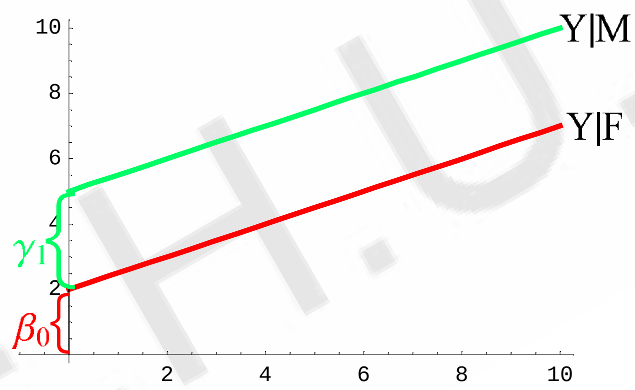
- Modelo a estimar:

$$
\begin{aligned}
Y_t &= \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \cancel{\gamma_2} S_{2t} + u_t \\
&= \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + u_t
\end{aligned}
$$

- <2>Modelos Submuestras:

$$
\begin{aligned}
\mathsf{E}\big(Y_t | S = F\big) &= \beta_0 + \beta_1 R_t \\
\mathsf{E}\big(Y_t | S = M\big) &= \beta_0 + \beta_1 R_t + \gamma_1
\end{aligned}
\quad \Rightarrow
$$

sin categoría F

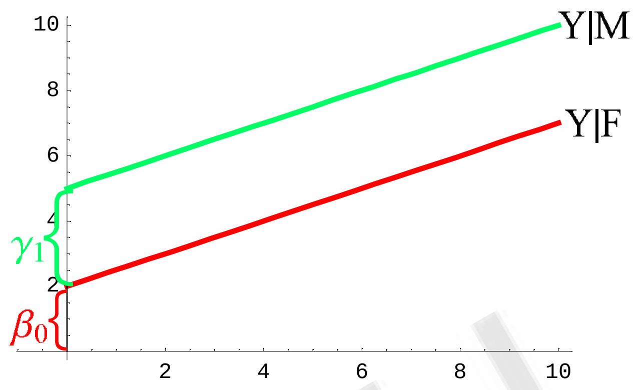


- Interpretación Coeficientes:

$$\mathsf{E}\big(Y_t | R_t = 0, S = F\big) = \beta_0$$

$$\mathsf{E}\big(Y_t | S = M\big) - \mathsf{E}\big(Y_t | S = F\big) = \gamma_1$$

Interpretación Coeficientes



$$\mathsf{E}\big(Y_t|S=M\big) - \mathsf{E}\big(Y_t|S=F\big) = \gamma_1$$

$$\mathsf{E}\big(Y_t|R_t=0, S=F\big) = \beta_0$$

- <3>es decir,

  $\beta_0 =$ consumo esperado Mujeres (base) si $R_t = 0$.

  $\gamma_1 =$ dif. consumo esperado de Hombres (frente a base $=$ Mujeres).

  $\beta_1 = \Delta$ consumo si $\Delta R_t = 1$ (*c.p.*).

Recuérdese: Este caso sólo significa diferentes interceptos para cada categoría.
Nota: Eliminar una categoría $\rightsquigarrow$ transformarla en base de referencia.

Contrastes Habituales con 1 VC

Hipótesis: variable cualitativa (Sexo) no significante
(no afecta el Consumo)
es decir $Masc$ y Fem mismo Consumo:

- Modelo Sin Restringir

$$Y_t = \beta_0 + \beta_1 R_t + \gamma S_{1t} + u_t$$

- Hipótesis: $H_0 : \gamma = 0$ frente a $H_a : \gamma \neq 0$
- Modelo Restringido:

$$Y_t = \beta_0 + \beta_1 R_t + u_t$$

- Véase el habitual Estadístico t (o Estadístico F basado en SCR).

1 VC con 2 cats + 1 VC con 3 cats

$$\begin{array}{ccccccc}
\text{Consumo} = & f([\text{cte}], & \text{renta}, & \text{sexo}, & & \text{territorio CAV}) \\
\downarrow & \downarrow & \downarrow & \swarrow \ \searrow & & \swarrow \downarrow \searrow \\
Y_t & [1] & R_t & Masc \quad Fem & & A \quad B \quad G
\end{array}$$

$$S_{1t} = \mathcal{I}(t \in M) \qquad T_{1t} = \mathcal{I}(t \in A)$$
$$S_{2t} = \mathcal{I}(t \in F) \qquad T_{2t} = \mathcal{I}(t \in B)$$
$$T_{3t} = \mathcal{I}(t \in G)$$

Muestra:

t	Y	cte	R	S	T
1	Y_1	1	R_1	M	B
2	Y_2	1	R_2	F	G
3	Y_3	1	R_3	F	B
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
T	Y_T	1	R_T	M	A
		¿ X ?			

$\Rightarrow$

t	Y	cte	R	S_1	S_2	T_1	T_2	T_3
1	Y_1	1	R_1	1	0	0	1	0
2	Y_2	1	R_2	0	1	0	0	1
3	Y_3	1	R_3	0	1	0	1	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$			$\vdots$	
T	Y_T	1	R_T	1	0	1	0	0
				X				

Recuérdese: En principio, sustituir var cualitativa por tantas vars Ficticias como categorías haya.

Trampa de Var Ficticias: 2 vars cualitativas

- Modelo: $Y_t = \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \gamma_2 S_{2t} + \delta_1 T_{1t} + \delta_2 T_{2t} + \delta_3 T_{3t} + u_t$
- Problema (Trampa VF):
  X es una $(T \times 7)$ matriz, pero

$$S_1 + S_2 = T_1 + T_2 + T_3 = [1]$$

(2 c.l. exactas) $\Rightarrow \mathrm{rk}(X) = 5 < 7$ (*es decir* MC perfecta)

- $\Rightarrow \det(X'X) = 0$
  $\Rightarrow$ ¡ $(X'X)^{-1}$ no existe !! y
  ¡ no se puede calcular $\widehat{\beta}$!!

- Solución General: eliminar UNA de las col's que causa el problema: [1] o (S_1 o S_2) o (T_1 o T_2 o T_3).

Solución: VF sin combinación de categorías

SOLUCIÓN MÁS HABITUAL:

eliminar última categoría de cada VF: S_2 y T_3:

- Modelo a estimar:

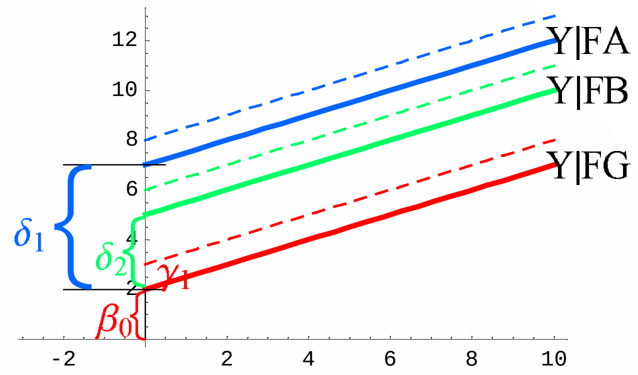
$$
\begin{aligned}
Y_t &= \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \cancel{\gamma_2} S_{2t} + \delta_1 T_{1t} + \delta_2 T_{2t} + \cancel{\delta_3} T_{3t} + u_t \\
&= \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \delta_1 T_{1t} + \delta_2 T_{2t} + u_t
\end{aligned}
$$

- <6>Modelos Submuestras:

	$S = M$	$S = F$	$M - F$
$T = A$	$\beta_0 + \beta_1 R_t + \gamma_1 + \delta_1$	$\beta_0 + \beta_1 R_t + \delta_1$	γ_1
$T = B$	$\beta_0 + \beta_1 R_t + \gamma_1 + \delta_2$	$\beta_0 + \beta_1 R_t + \delta_2$	γ_1
$T = G$	$\beta_0 + \beta_1 R_t + \gamma_1$	$\beta_0 + \beta_1 R_t$	γ_1
$A - G$	δ_1	δ_1	
$B - G$	δ_2	δ_2	
$A - B$	$\delta_1 - \delta_2$	$\delta_1 - \delta_2$	

Coeficiente Interpretación

sin categorías F ni G



$$\mathsf{E}\big(Y_t|S=M\big) - \mathsf{E}\big(Y_t|S=F\big) = \gamma_1$$
$$\mathsf{E}\big(Y_t|T=A\big) - \mathsf{E}\big(Y_t|T=G\big) = \delta_1$$
$$\mathsf{E}\big(Y_t|T=B\big) - \mathsf{E}\big(Y_t|T=G\big) = \delta_2$$
$$\mathsf{E}\big(Y_t|R_t=0, S=F, T=G\big) = \beta_0$$

- <5>es decir,

$\beta_0 =$ consumo esperado Mujeres G (base) if $R_t = 0.$

$\gamma_1 =$ dif. consumo esperado Hombres frente a Mujeres .

$\delta_1 =$ dif. consumo esperado A frente a $G.$

$\delta_2 =$ dif. consumo esperado B frente a $G.$

$\beta_1 = \Delta$ consumo si $\Delta R_t = 1$ (*c.p.*).

Recuérdese: Este caso sólo significa diferentes interceptos para cada categoría.
Recuérdese: Eliminar una (combinación de) categoría(s)
$\rightsquigarrow$ transformarla en base de referencia.

Contrastes habituales con 2 VCs

Hipótesis: Variable Sexo no afecta Consumo
(pero territorio de residencia puede que sí)

- Modelo sin Restringir:

$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \delta_1 T_{1t} + \delta_2 T_{2t} + u_t$$

($\gamma_1 =$ dif. esperada Consumo $Masc$ frente a Fem)

- Hipótesis: $H_0 : \gamma_1 = 0$ frente a $H_a : \gamma_1 \neq 0$
- Modelo Restringido:

$$Y_t = \beta_0 + \beta_1 R_t + \delta_1 T_{1t} + \delta_2 T_{2t} + u_t$$

- Usar el habitual t Estadístico (o F Estadístico basado en SCR)

Otros Contrastes habituales con 2 VCs

- Modelo sin Restringir (sin S_2 ni T_3):

$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \delta_1 T_{1t} + \delta_2 T_{2t} + u_t$$

  - Recuérdese: γ_1 es dif. Consumo esperado $Masc$ frente a Fem (base)
    δ_1 y δ_2 son dif. Consumo esp. de A y B frente a G (base)

- Hipótesis: Mismo Consumo para todos
  (independientemente del Sexo y Territorio):
  - $H_0 : \gamma_1 = \delta_1 = \delta_2 = 0$
  - Modelo Restringido:

$$Y_t = \beta_0 + \beta_1 R_t + u_t$$

- Hipótesis: Territorio de Residencia no afecta Consumo
  (pero $Masc$ frente a Fem puede que si):
  - $H_0 : \delta_1 = \delta_2 = 0$
  - Modelo Restringido:

$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + u_t$$

Otros Contrastes habituales con 2 VCs

- Modelo sin Restringir (sin S_2 ni T_3):

$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \delta_1 T_{1t} + \delta_2 T_{2t} + u_t$$

  - Recuérdese: δ_1 y δ_2 son dif. Consumo esperados de A y B frente a G (base)

- Hipótesis: Residentes del mismo sexo en A y B tienen el mismo nivel de consumo (pero G puede ser diferente):
  - $H_0 : \delta_1 = \delta_2$ frente a $H_a : \delta_1 \neq \delta_2$
  - Modelo Restringido:

$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \delta(\underbrace{T_{1t} + T_{2t}}_{1-T_{3t}}) + u_t$$

- Hipótesis: Residentes del mismo sexo en B y G tienen el mismo nivel de consumo (pero A puede ser diferente):
  - $H_0 : \delta_2 = 0$ frente a $H_a : \delta_2 \neq 0$
  - Modelo Restringido:

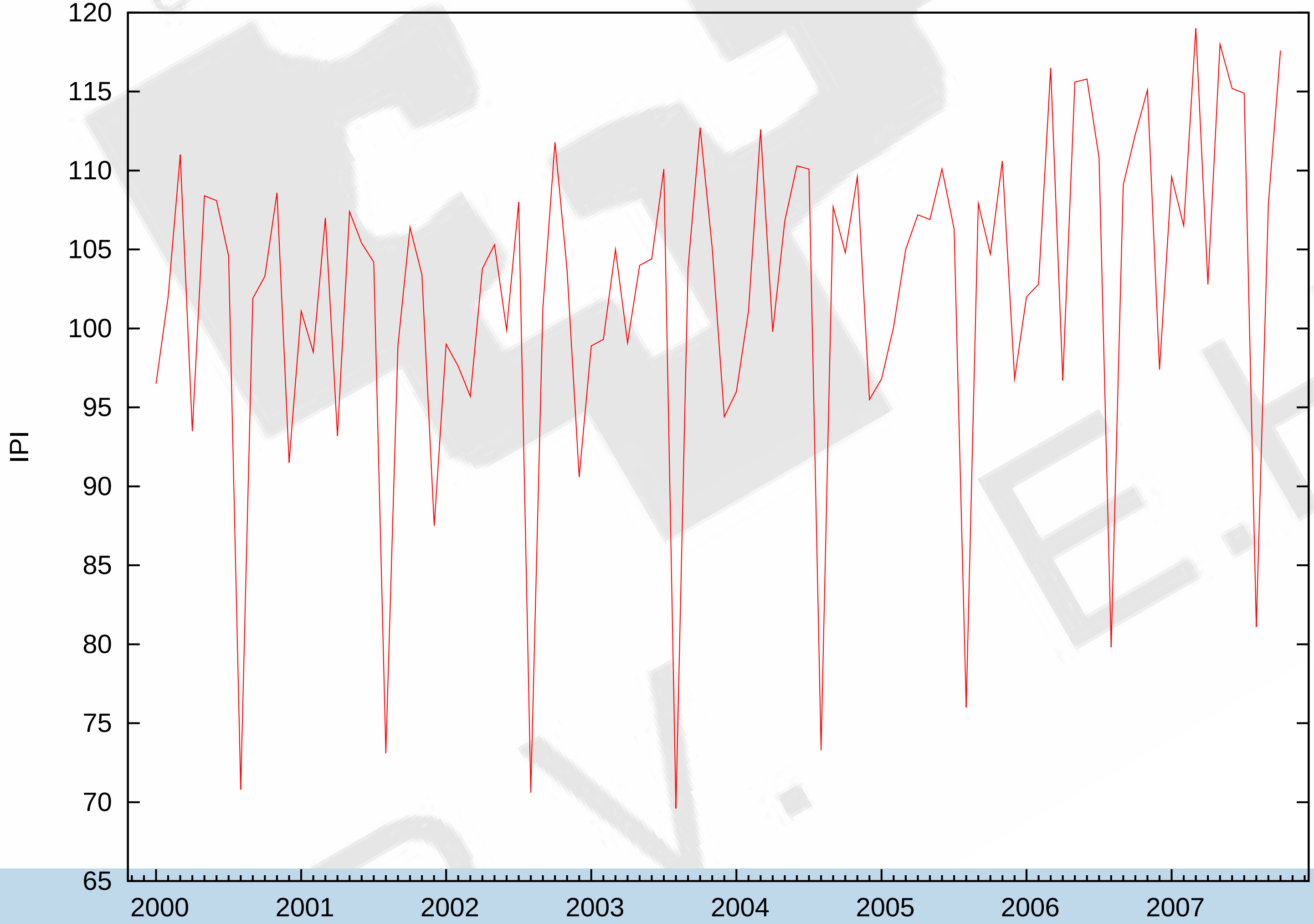
$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 S_{1t} + \delta_1 T_{1t} + u_t$$

4.2 Efecto estacional

Efecto estacional

- Efecto estacional:
-
- Var estacional $\rightsquigarrow$ submuestras $T_1, T_2, \ldots$

según estaciones/meses



Variables Ficticias Estacionales: Definición

- Def. Variable Ficticia Estacional:

$$D_{jt} = \begin{cases} 1, & \text{en } t \in \text{estación } j = 1, 2, 3, 4, \ldots; \\ 0, & \text{sino.} \end{cases}$$

- *p.ej.* para datos trimestrales:

date (t)	IPI_t	X_t	D_{1t}	D_{2t}	D_{3t}	D_{4t}
1975.1	·	·	1	0	0	0
1975.2	·	·	0	1	0	0
1975.3	·	·	0	0	1	0
1975.4	·	·	0	0	0	1
1976.1	·	·	1	0	0	0
1976.2	·	·	0	1	0	0
1976.3	·	·	0	0	1	0
1976.4	·	·	0	0	0	1
1977.1	·	·	1	0	0	0
$\vdots$	$\vdots$	$\vdots$		$\ldots$		
2000.1	·	·	1	0	0	0
2000.2	·	·	0	1	0	0
2000.3	·	·	0	0	1	0
2000.4	·	·	0	0	0	1
2001.1	·	·	1	0	0	0
$\ldots$	$\ldots$	$\ldots$		$\ldots$		

Variables Ficticias Estacionales: Definición (2)

- Modelo a estimar:

$$
\begin{aligned}
IPI_t &= \beta_0 + \beta_1 X_t + \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \cancel{\gamma_4} D_{4t} + u_t \\
&= \beta_0 + \beta_1 X_t + \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + u_t
\end{aligned}
$$

- ¿ Interpretación de los parámetros γ ?
- <1>¿ Qué ocurre si los datos son observaciones mensuales (como de hecho en el ejemplo IPI) ?

fecha (t)	IPI_t	X_t	D_{1t}	D_{2t}	D_{3t}	D_{4t}
1975.ene	·	·	1	0	0	0
1975.feb	·	·	1	0	0	0
1975.mar	·	·	1	0	0	0
1975.abr	·	·	0	1	0	0
1975.may	·	·	0	1	0	0
1975.jun	·	·	0	1	0	0
1975.jul	·	·	0	0	1	0
1975.ago	·	·	0	0	1	0
1975.sep	·	·	0	0	1	0
1975.oct	·	·	0	0	0	1
1975.nov	·	·	0	0	0	1
1975.dic	·	·	0	0	0	1
1976.ene	·	·	1	0	0	0
1976.feb	·	·	1	0	0	0
1976.mar	·	·	1	0	0	0
...	...	...	...			

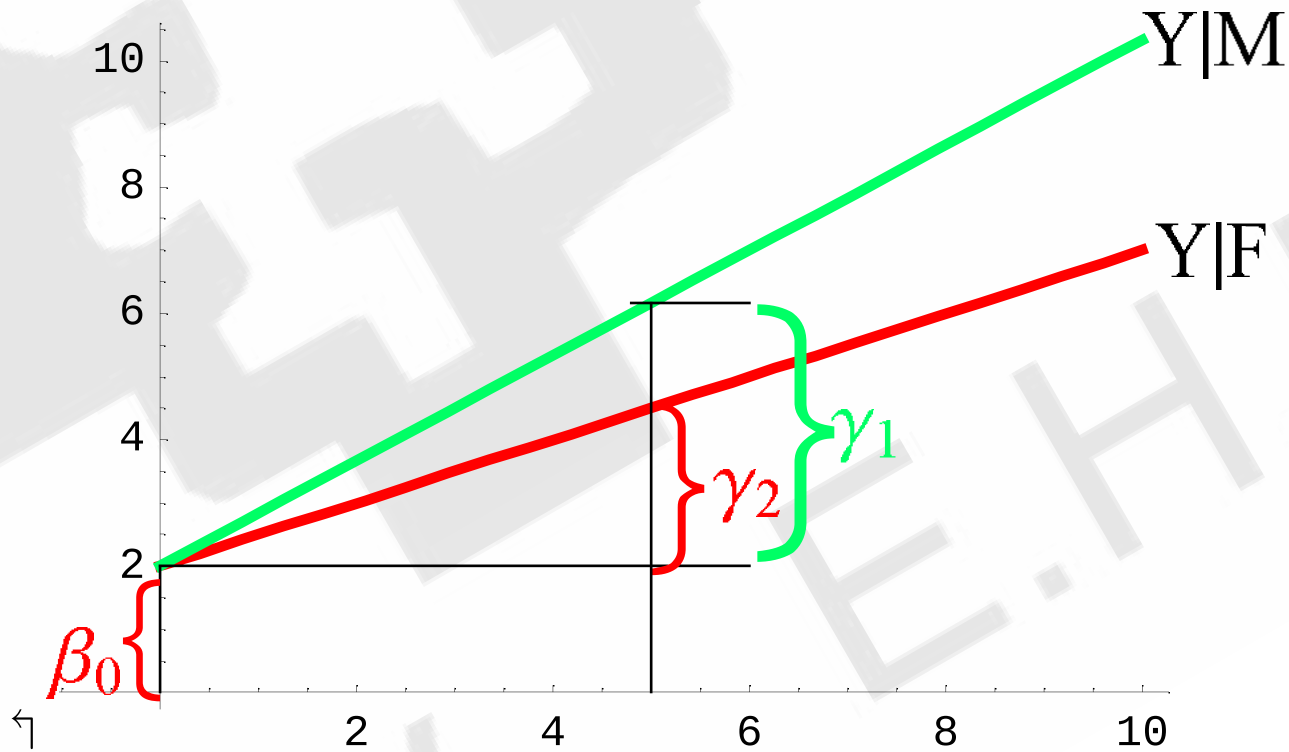
ó

$D_{1t}\ldots$											$\ldots D_{12t}$
1	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0
...											

4.3 Interacción entre VFs y Vars cuantitativas

Interacción entre VFs y Vars cuantitativas

En vez de diferentes *interceptos*, necesitamos diferentes pendientes para cada categoría:



es decir, diferente respuesta "Y" para el mismo "X"

Trampa Var Ficticias: interacción

- <1>Matriz X:

cte	R	$R \times S_1$	$R \times S_2$
1	R_1	$R_1 \times 1$	$R_1 \times 0$
1	R_2	$R_2 \times 0$	$R_2 \times 1$
1	R_3	$R_3 \times 0$	$R_3 \times 1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	R_T	$R_T \times 1$	$R_T \times 0$

$\Rightarrow$

cte	R	RS_1	RS_2
1	R_1	R_1	0
1	R_2	0	R_2
1	R_3	0	R_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	R_T	R_T	0

- Modelo:

$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 R_t S_{1t} + \gamma_2 R_t S_{2t} + u_t$$

- Problema (trampa VF): X es matriz $T \times 4$, pero

$$RS_1 + RS_2 = R \Rightarrow \text{rk}(X) = 3 < 4 \quad \text{(MC exacta)}$$

- Solución General: eliminar UNA de las col's que causan el problema: R ó RS_1 ó RS_2 .

Solución: Interacción sin categoría

- SOLUCIÓN MÁS HABITUAL:

  eliminar última categoría de la VF: F (RS_2):

- Modelo a estimar:

$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 R_t S_{1t} + u_t$$

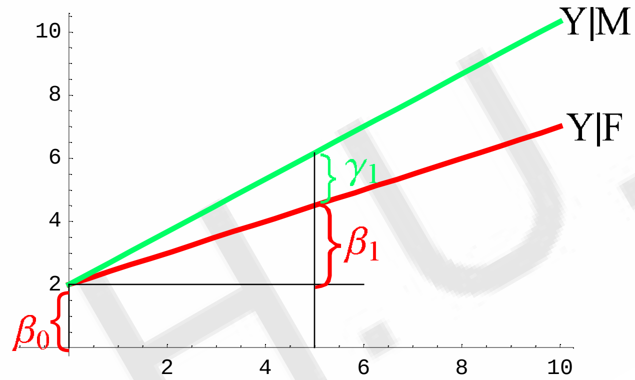
no $R\,S_2$

- <2>Modelos Submuestras:

$$\mathrm{E}\big(Y_t | S = F\big) = \beta_0 + \beta_1 R_t$$

$$\mathrm{E}\big(Y_t | S = M\big) = \beta_0 + \underbrace{(\beta_1 + \gamma_1)}_{\beta_1^*} R_t$$

$\Rightarrow$



- Interpretación de los coeficientes:

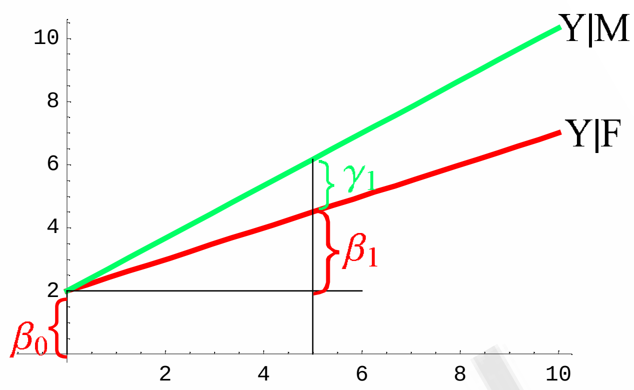
$$\mathrm{E}\big(Y_t | R_t = 0\big) = \beta_0$$

$$\frac{\partial \mathrm{E}\big(Y_t | S = F\big)}{\partial R_t} = \beta_1$$

$$\frac{\partial \mathrm{E}\big(Y_t | S = M\big)}{\partial R_t} = \beta_1 + \gamma_1$$

Interpretación de los coeficientes

no $R\,S_2$



$$\mathsf{E}\left(Y_t|R_t=0\right)=\beta_0$$

$$\frac{\partial \mathsf{E}\left(Y_t|S=F\right)}{\partial R_t}=\beta_1$$

$$\frac{\partial \mathsf{E}\left(Y_t|S=M\right)}{\partial R_t}=\beta_1+\gamma_1$$

- <3>es decir,

  $\beta_0=$ consumo esperado si $R_t=0$.

  $\beta_1=\Delta$ consumo Mujeres si $\Delta R_t=1$ (*c.p.*).

  $\gamma_1=$ dif Δ consumo para Hombres (frente a base $=$ Mujer).

Recuérdese: Este caso significa diferentes pendientes para cada categoría.
Recuérdese una vez más: Eliminar una (combinación de) categoría(s)
$\rightsquigarrow$ transformarla en base de referencia.

Contrastes Habituales con Interacción

Hipótesis: M y F Consumo igual
o la variable Sexo no afecta Consumo:

- Modelo sin Restringir:

$$Y_t = \beta_0 + \beta_1 R_t + \gamma_1 R_t S_{1t} + u_t$$

- Hipótesis: $H_0 : \gamma_1 = 0$ frente a $H_a : \gamma_1 \neq 0$
- Modelo Restringido:

$$Y_t = \beta_0 + \beta_1 R_t + u_t$$

- Usar el habitual Estadístico t (o Estadístico F basado en SCR)

Fin

FIN