

# Econometría



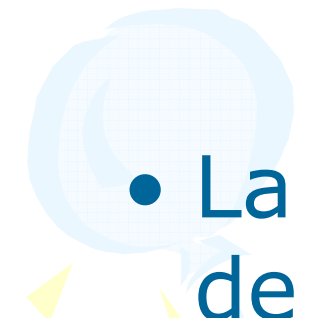

## **Errores de Especificación**



# Omisión de variables relevantes

- En general,  $\hat{b}_r$ , es un estimador sesgado, siendo el sesgo:

$$(X_r' X_r)^{-1} X_r' X_s \beta_s$$

- La matriz de varianzas y covarianzas de  $\hat{b}_r$ , es menor o igual que la de  $\hat{\beta}_r$
  - El estimador de  $\hat{\sigma}_\varepsilon^2$  de  $\hat{\sigma}_u^2$  es sesgado, siendo el sesgo positivo.
- 
- 

# Ejemplo: Datos encuesta de ocupación hotelera (2003:9)

	Empleo (en miles)	Número de viajeros (miles)	Estancia media	Ocupación	Es isla?
Andalucía	28,4	11.902,5	3,1	51,9	0,0
Aragón	3,6	1.848,0	2,1	37,2	0,0
Asturias (Principado de)	2,4	1.088,2	2,3	32,9	0,0
Balears (Illes)	25,9	6.716,0	7,2	69,9	1,0
Canarias	27,2	4.875,7	7,8	68,4	1,0
Cantabria	2,0	933,8	2,4	39,8	0,0
Castilla y León	6,2	3.647,6	1,7	34,5	0,0
Castilla-La Mancha	2,8	1.805,1	1,7	30,8	0,0
Cataluña	23,5	10.771,7	3,4	55,0	0,0
Comunidad Valenciana	13,4	5.579,7	3,9	59,8	0,0
Extremadura	2,2	1.000,7	1,7	30,7	0,0
Galicia	6,3	3.040,5	2,1	32,9	0,0
Madrid (Comunidad de)	10,7	5.748,9	2,1	50,5	0,0
Murcia (Región de)	2,0	882,5	3,0	48,7	0,0
Navarra (Comunidad Foral de)	1,1	557,7	2,0	35,3	0,0
País Vasco	3,2	1.540,6	1,9	45,3	0,0
Rioja (La)	0,7	446,2	1,8	43,3	0,0
	161,8	62.385,5	50,3	766,8	2,0

# Ejemplo: Estimación empleo en hoteles

## Resumen

### Estadísticas de la regresión

Coefficiente de	0,9970249
Coefficiente de	0,99405865
R <sup>2</sup> ajustado	0,91576449
Error típico	1,2079915
Observaciones	17

### ANÁLISIS DE VARIANZA

	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>	<i>Media cuadrada</i>	<i>F</i>	<i>Valor crítico de F</i>
Regresión	4	3173,93259	793,483147	543,763373	1,901E-13
Residuos	13	18,970165	1,45924346		
Total	17	3192,90275			

	<i>Coefficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>
Intercepción	0	#N/A	#N/A	#N/A	#N/A
Número de vi	0,0022267	0,00010245	21,7337712	1,3273E-11	0,00200536
Estancia med	2,00740779	0,81523144	2,4623778	0,02854105	0,24620736
Ocupación	-0,11792908	0,0456325	-2,58432202	0,02267013	-0,21651211
Es isla?	6,66733777	3,14750797	2,11829099	0,05398809	-0,13243978

# Ejemplo: Estimación empleo en hoteles con variables omitidas

## Resumen

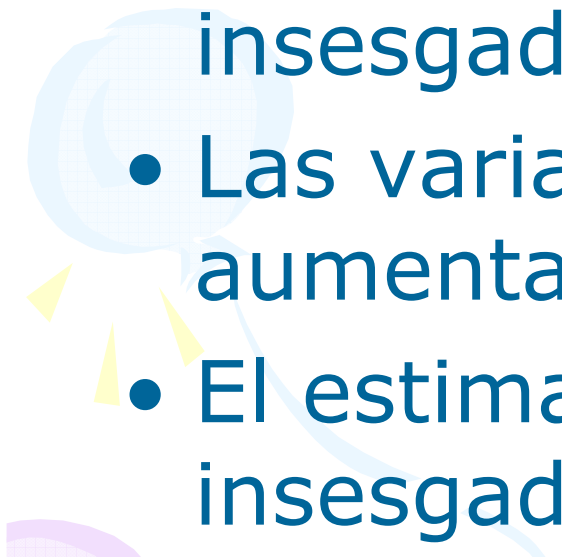
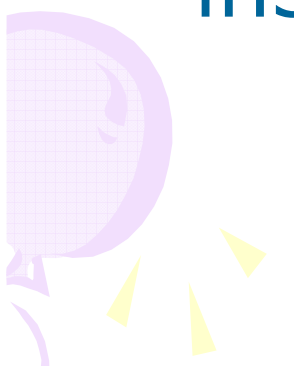
<i>Estadísticas de la regresión</i>	
Coefficiente de	0,88214439
Coefficiente de	0,77817873
R <sup>2</sup> ajustado	0,6750614
Error típico	7,11263339
Observaciones	17

0,9970249
0,99405865
0,91576449
1,2079915
17

	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>
Intercepción	0	#N/A	#N/A			
Estancia med	5,1442306	4,72424438	1,08890019	2,00740779	0,81523144	2,4623778
Ocupación	-0,09703113	0,26862374	-0,36121577	-0,11792908	0,0456325	-2,58432202
Es isla?	-5,52014375	18,2359796	-0,30270618	6,66733777	3,14750797	2,11829099



# Inclusión de variables relevantes

- Las estimaciones de los parámetros asociados a variables relevantes son insesgados  $E(\hat{b}_r) = \beta_r$
  - Las varianzas de los estimadores aumentan
  - El estimador de  $\hat{\sigma}_u^2$  sigue siendo insesgado
- 
- 

# Ejemplo: inclusión de variables irrelevantes

## Resumen

<i>Estadísticas de la regresión</i>		
Coefficiente de	0,99703495	0,9970249
Coefficiente de	0,99407869	0,99405865
R <sup>2</sup> ajustado	0,90877159	0,91576449
Error típico	1,25519475	1,2079915
Observaciones	17	17

	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>
Intercepción	0	#N/A	#N/A			
Número de vi	0,00221777	0,00011532	19,2314523	0,0022267	0,00010245	21,7337712
Estancia med	2,03156431	0,85552456	2,37464173	2,00740779	0,81523144	2,4623778
Ocupación	-0,11112205	0,05821449	-1,90883833	-0,11792908	0,0456325	-2,58432202
Es isla?	6,58181225	3,29791395	1,99575014	6,66733777	3,14750797	2,11829099
Dispone de in	-0,0081169	0,04027312	-0,20154633			

# Forma funcional incorrecta

Relación cierta  $Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \dots + \beta_k + u_i$

Relación que estimamos  $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

Test RESET

- estimar la regresión simple  $Y_i$  de sobre  $X_i$ , calcular el coeficiente de determinación  $R_o^2$ , y generar los valores ajustados  $\hat{Y}_i$
- estimar la regresión de  $Y_i$  sobre  $1, X_i, \hat{Y}_i^2, \dots, \hat{Y}_i^k$  y calcular el coeficiente de determinación,  $R_1^2$  ;

- calcular el estadístico  $F$

$$F = \frac{(R_1^2 - R_o^2) / (k - 2)}{(1 - R_1^2) / (n - k - 1)}$$

- y rechazar la hipótesis de forma funcional correcta si  $F > c$ , en donde  $c$  es el valor crítico talque  $Prob(F_{k-2, n-k-1} > c) = \alpha$ .



# Ejemplo: RESET

	Y	X	Y <sup>2</sup>	Y <sup>3</sup>	Y_estimado
Andalucía	28,4	11.902,5	941,024948	28867,0036	30,6761299
Aragón	3,6	1.848,0	23,3505915	112,835778	4,83224497
Asturias (Prin	2,4	1.088,2	8,29016451	23,869577	2,87926458
Balears (Illes)	25,9	6.716,0	300,84473	5218,11458	17,3448762
Canarias	27,2	4.875,7	159,133177	2007,4332	12,6147999
Cantabria	2,0	933,8	6,16334532	15,3011739	2,48260857
Castilla y León	6,2	3.647,6	89,4557526	846,081912	9,45810513
Castilla-La M	2,8	1.805,1	22,2987775	105,298353	4,72215814
Cataluña	23,5	10.771,7	771,157537	21414,8315	27,7697234
Comunidad V	13,4	5.579,7	208,06174	3001,15439	14,4243454
Extremadura	2,2	1.000,7	7,04611209	18,7035619	2,65445137
Galicia	6,3	3.040,5	62,3719439	492,588105	7,89759102
Madrid (Comu	10,7	5.748,9	220,792458	3280,77428	14,8590867
Murcia (Regió	2,0	882,5	5,52547256	12,9883547	2,35063237
Navarra (Com	1,1	557,7	2,29730203	3,48198698	1,51568533
País Vasco	3,2	1.540,6	16,3385074	66,0417491	4,04209195
Rioja (La)	0,7	446,2	1,51114991	1,85763901	1,22928837

$$F = \frac{(R_1^2 - R_0^2)/(k-2)}{(1-R_1^2)/(n-k-1)}$$

F 0,01593049  
 c= F(1,14)  
 alpha 0,05 4,6

se admite la hipótesis de que es la forma funcional correcta

	Coefficientes	Error típico	Estadístico t
Intercepción	0,08226884	1,7140869	0,04799572
X	0,00257038	0,0003421	7,51281864