

# Econometría I (EC402)

## Clase #18 - MCRL con variables dicotomas

Prof. Andrés M. Castaño

Ingeniería Comercial  
Universidad Católica del Norte  
Lunes 11 de noviembre de 2013

## Comentarios generales...

- Las variables explicativas pueden no ser de escala de razón, de escala de intervalo, o de escala ordinal.
- Las variables cualitativas que tienden a ser importantes son: sexo, raza, color, religión, nacionalidad, región geográfica, trastornos políticos y afiliación a un partido.
- Modelos ANOVA: una regresada cuantitativa y unas regresoras cualitativas

# Ejemplo: Difiere el salario promedio anual de los maestros de escuelas públicas por área geográfica?

51 estados clasificados en tres regiones: región noreste y norte-centro (21 estados), región sur (17 estados) y región oeste (13 estados)

**TABLA 9.1** SALARIO PROMEDIO DE LOS MAESTROS DE ESCUELAS PÚBLICAS, POR ESTADO, 1986

Salario	Gasto	$D_2$	$D_3$	Salario	Gasto	$D_2$	$D_3$
19 583	3 346	1	0	22 795	3 366	0	1
20 263	3 114	1	0	21 570	2 920	0	1
20 325	3 554	1	0	22 080	2 880	0	1
26 800	4 642	1	0	22 250	3 731	0	1
29 470	4 669	1	0	20 940	2 853	0	1
26 610	4 888	1	0	21 800	2 533	0	1
30 678	5 710	1	0	22 934	2 729	0	1
27 170	5 536	1	0	18 443	2 305	0	1
25 853	4 168	1	0	19 538	2 642	0	1
24 500	3 547	1	0	20 460	3 124	0	1
24 274	3 159	1	0	21 419	2 752	0	1
27 170	3 621	1	0	25 160	3 429	0	1
30 168	3 782	1	0	22 482	3 847	0	0
26 525	4 247	1	0	20 969	2 509	0	0
27 360	3 982	1	0	27 224	5 440	0	0
21 690	3 568	1	0	25 892	4 042	0	0
21 974	3 155	1	0	22 644	3 402	0	0
20 816	3 059	1	0	24 640	2 829	0	0
18 095	2 967	1	0	22 341	2 297	0	0
20 939	3 285	1	0	25 610	2 932	0	0
22 644	3 914	1	0	26 015	3 705	0	0
24 624	4 517	0	1	25 788	4 123	0	0
27 188	4 349	0	1	29 132	3 608	0	0
33 990	5 020	0	1	41 480	8 349	0	0
23 382	3 594	0	1	25 845	3 766	0	0
20 627	2 821	0	1				

Nota:  $D_2$  = 1 para estados del Noreste y Norte-centro; 0 de otra región.

$D_3$  = 1 para estados del Sur; 0 de otra región.

Fuente: National Education Association, como se informó en *Albuquerque Tribune*, 7 de noviembre de 1986.

# Especificación



$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \mu_i$$

- ▶  $Y_i$  = Salario promedio de los maestros de las escuelas públicas en el estado  $i$
- ▶  $D_{2i}$  = es 1 si el estado está en el Noroeste o Norte-centro, 0 para otra región del país
- ▶  $D_{3i}$  = es 1 si el estado está en el Sur, 0 para otra región del país

- Cuál es el salario medio de los maestros de escuelas públicas en la región Noreste y Norte-centro?

$$E(Y_i \mid D_{2i} = 1, D_{3i} = 0) = \beta_1 + \beta_2$$

- Cuál es el salario medio de los maestros de escuelas públicas del Sur?

$$E(Y_i \mid D_{2i} = 0, D_{3i} = 1) = \beta_1 + \beta_3$$

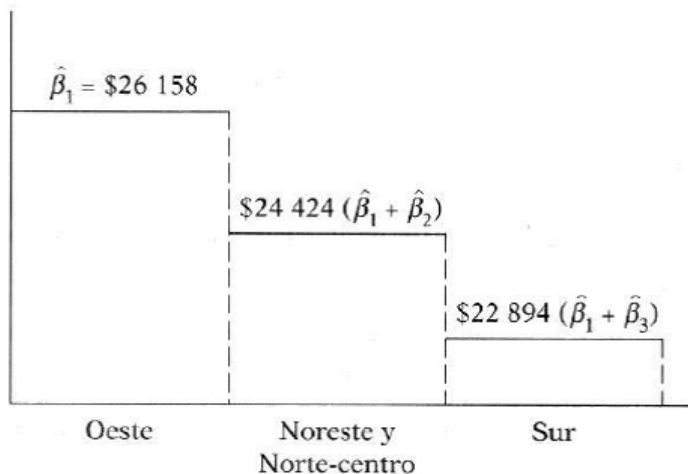
- Cuál es el salario medio de los maestros de escuelas públicas del Oeste?

$$E(Y_i \mid D_{2i} = 0, D_{3i} = 0) = \beta_1$$

## Ejemplo

$$\begin{aligned}\hat{Y}_i &= 26\,158.62 - 1\,734.473D_{2i} - 3\,264.615D_{3i} \\ ee &= (1\,128.523) \quad (1\,435.953) \quad (1\,499.615) \\ t &= (23.1759) \quad (-1.2078) \quad (-2.1776) \\ &\quad (0.0000)^* \quad (0.2330)^* \quad (0.0349)^* \quad R^2 = 0.0901\end{aligned}\tag{9.2.5}$$

## Ejemplo



# Precaución en el uso de variables dicótomas

- Si una variable cualitativa tiene  $m$  categorías, sólo hay que agregar  $(m-1)$  variables dicótomas  $\implies$  Colinealidad perfecta (trampa de la variable dicótoma).
- La categoría a la cual no se asigna variable dicótoma se conoce como categoría de comparación, base, de control, de referencia u omitida. Todas las comparaciones se hacen con base en la categoría de comparación.
- El valor de la intersección ( $\beta_1$ ) representa el valor medio de la categoría de comparación.
- Los coeficientes anexos a las variables dicótomas se conocen como coeficientes de intersección diferencial. Indican en qué medida el valor de la intersección que se le asigna 1 varía del coeficiente de intersección de la categoría de comparación.
- Se pueden introducir tantas variables dicotómicas como categorías cuando se omite de la regresión la intersección
- Qué criterio es mejor:  $(m-1)$  variables dicótomas o sacar el intercepto  $\implies$  Planteamiento de Kennedy.

## Ejemplo sin intersección

$$\begin{aligned}\hat{Y}_i &= 26\,158.62D_{1i} + 24\,424.14D_{2i} + 22\,894D_{3i} \\ ee &= (1\,128.523) \quad (887.9170) \quad (986.8645) \\ t &= (23.1795)^* \quad (27.5072)^* \quad (23.1987)^* \\ &\quad R^2 = 0.0901\end{aligned}$$

$\beta_1$  = salario medio de los maestros en el Oeste

$\beta_2$  = salario medio de los maestros en el Noreste y Norte-centro

$\beta_3$  = salario medio de los maestros en el Sur



# Modelos ANOVA con dos variables cualitativas

## EJEMPLO 9.2

SALARIOS POR HORA EN RELACIÓN CON EL ESTADO CIVIL Y LA REGIÓN DE RESIDENCIA

De una muestra de 528 personas tomada en mayo de 1985, se obtuvieron los siguientes resultados de regresión:<sup>8</sup>

$$\begin{aligned}\hat{Y}_i &= 8.8148 + 1.0997D_{2i} - 1.6729D_{3i} \\ ee &= (0.4015) \quad (0.4642) \quad (0.4854) \\ t &= (21.9528) \quad (2.3688) \quad (-3.4462) \quad (9.3.1) \\ &\quad (0.0000)^* \quad (0.0182)^* \quad (0.0006)^* \\ R^2 &= 0.0322\end{aligned}$$

donde  $Y$  = salario por hora (\$)

$D_2$  = estado civil; 1 si es caso, 0 en otro caso

$D_3$  = región de residencia; 1 si es del Sur, 0 en otro caso

# Modelos Ancova: mezcla de cuantitativas con cualitativas

- Los modelos ANOVA no son tan comunes en economía: Modelos de análisis de Covarianza.

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 X_i + u_i \quad (9.4.1)$$

donde  $Y_i$  = salario anual promedio de los maestros de escuelas públicas en el estado (\$)

$X_i$  = gasto en escuelas públicas por alumno (\$)

$D_{2i}$  = 1 si el estado es del Noreste o Norte-centro; 0 en otro caso

$D_{3i}$  = 1 si el estado es del Sur; 0 en otro caso

# Modelos Ancova: mezcla de cuantitativas con cualitativas

## SALARIO DE LOS MAESTROS RESPECTO A LA REGIÓN Y AL GASTO EN ESCUELAS PÚBLICAS POR ALUMNO

De los datos mostrados en la tabla 9.1, los resultados del modelo (9.4.1) son los siguientes:

$$Y_i = 13\,269.11 - 1\,673.514D_{2i} - 1\,144.157D_{3i} + 3.2889X_i$$
$$ee = (1\,395.056) \quad (801.1703) \quad (861.1182) \quad (0.3176)$$
$$t = (9.5115)^* \quad (-2.0889)^* \quad (-1.3286)^{**} \quad (10.3539)^*$$
$$R^2 = 0.7266$$

# Modelos Ancova: mezcla de cuantitativas con cualitativas

