Econometría I (EC402)
Clase #16 - Extensiones del MCRL

Prof. Andrés M. Castaño

Ingeniería Comercial
Universidad Católica del Norte
Miercoles 10 de octubre de 2013

Regresión a traves del origen

- $Y_i = \beta_2 X_i + \mu_i$.
- A veces la teoría que sirve de base requiere o exige que el término intersección esté ausente.
- Hipótesis del ingreso permanente de M. Friedman, Teoría del Análisis de Costos, Teoría de Crecimiento de la Oferta Monetaria.
- Cómo se estiman este tipo de modelos sin intercepto? qué problemas presentan?

MCO sin intercepto

- MCO sin intercepción:

$$\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$$

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum X_i^2}$$

$$\hat{\sigma}^2 = \frac{\sum \hat{\mu}_i^2}{n-1}$$

MCO con intercepto

- MCO con intercepto:

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

$$\hat{\sigma}^2 = \frac{\sum \hat{\mu}_i^2}{n-2}$$

Diferencias

- En el modelo sin términos de intersección sólo se usan sumas sencillas de cuadrados y productos cruzados, en el con intersección se utilizan sumas ajustadas de la media y productos cruzados.
- Los grados de libertad son diferentes, por qué?
- Bajo el marco tradicional $\sum \hat{\mu}_i$ es siempre cero en el modelo con intersección. Sin intersección no es necesario que lo sea.
- El r^2 siempre positivo antes, ahora podría tomar valores negativos. r^2 puede no ser apropiado en modelos de regresión a través del origen.

Diferencias

- Se puede calcular lo que se conoce como el r^2 simple:

$$r_s^2 = \frac{(\sum X_i Y_i)^2}{\sum X_i^2 Y_i^2}$$

- Se trata de un r cuadrado que no es corregido por la media. No es comparable con el r cuadrado tradicional.
- A menos de que la teoría sea demasiado fuerte, se aconseja seguir apegado al modelo convencional.
- Si comenzamos del modelo tradicional con intercepto y este es no significativo, tendríamos de manera práctica una regresión a través del origen. Si se insiste desde el principio en correrlo sin intercepción se está cometiendo un error de especificación.

Ejemplo

En el ejercicio 5.5 introdujimos la *recta característica* del análisis de inversión, la cual puede ser escrita como

$$Y_i = \alpha_i + \beta_i X_i + u_i \qquad \textbf{(6.1.10)}$$

donde Y_i = tasa anual de ganancia (%) sobre "Afuture Fund"
X_i = tasa anual de ganancia (%) sobre el portafolio de mercado
β_i = coeficiente de la pendiente, conocida también como el coeficiente **beta** en la teoría del portafolio, y
α_i = la intersección

Ejemplo

$$\hat{Y}_i = \underset{(0.1916)}{1.0899 X_i} \qquad r^2 \text{ simple} = 0.7825 \qquad \textbf{(6.1.12)}$$

$$t = (5.6884)$$

$$\hat{Y}_i = \underset{(7.6886)}{1.2797} + \underset{(0.2383)}{1.0691 X_i} \qquad \textbf{(6.1.13)}$$

$$t = (0.1664) \qquad (4.4860) \qquad r^2 = 0.7155$$

Regresión sobre variables estandarizadas

- Las unidades en que la variable dependiente e independiente se expresa influye sobre la interpretación de los coeficientes de regresión.

- Esto se puede evitar si ambas variables se estandarizan.

$$Y_i^* = \frac{Y_i - \bar{Y}}{S_Y}$$

$$X_i^* = \frac{X_i - \bar{X}}{S_X}$$

- Para una variable estandarizada la media siempre es cero y la varianza es igual a 1.

Regresión sobre variables estandarizadas

- Efectuar la regresión (los coeficientes son conocidos como coeficientes beta):

$$
\begin{aligned}
Y_i^* &= \beta_1^* + \beta_2^* X_i^* + \mu_i^* \\
&= \beta_2^* X_i^* + \mu_i^*
\end{aligned}
$$

  Nos lleva a una regresión a través del origen. Las interpretaciones serán tomadas como el cambio en Beta desviaciones estándar que provoca un cambio en una desviación estándar de X.

Ventajas del modelo tradicional versus el modelo estandarizado

- En el escenario de regresión múltiple el coeficiente estandarizado es una forma de medir la fuerza relativa de las distintas regresoras.

- Relación entre los coeficientes tradicionales y el coeficiente beta:

$$\hat{\beta}_2^* = \hat{\beta}_2 \frac{S_x}{S_y}$$

Formas funcionales de los modelos de regresión

- Modelo log-log.
- Modelos semilogarítmicos (log-lin; lin-log).
- Modelos recíprocos.
- Modelo logarítmico recíproco.

Cómo medir elasticidad: Modelo log-log

- Modelo de regresión exponencial:

$$Y_i = \beta_1 X_i^{\beta_2} e^{\mu_i}$$

$$lnY_i = \alpha + \beta_2 lnX_i + \mu_i$$

- Esta ecuación se podría estimar igual por mínimos cuadrados ordinarios, Porqué?

- Bajos este marco el coeficiente β_2 mide la elasticidad de Y con respecto a X (esto es, el cambio porcentual en Y ante un cambio pequeño cambio porcentual en X.

- Características especiales: 1. El coeficiente de elasticidad entre Y y X permanece constante (modelo de elasticidad constante), 2. $\hat{\beta}_1$ es un estimador sesgado $\hat{\beta}_1 = antilog(\hat{\alpha})$. En términos prácticos muchas veces no es necesario preocuparse por obtener este estimador insesgado.

Ejemplo: elasticidad gasto en bienes durables respecto al gasto en consumo personal

$$\begin{aligned}
\ln \text{GASBD}_t &= -9.6971 + 1.9056 \ln \text{GCPERT}_t \\
ee &= (0.4341) \quad (0.0514) \\
t &= (-22.3370)^* \quad (37.0962)^* \quad r^2 = 0.9849
\end{aligned}$$

Cómo medir tasas de crecimiento: Modelo log-lin

- A veces el interés es encontrar la tasa de crecimiento de ciertas variables económicas (población, oferta monetaria, empleo, productividad)

- $lnY_t = \beta_1 + \beta_2 t + \mu_t$

- β_2 representa un cambio relativo en Y dado un cambio absoluto en el valor del regresor (t en este caso)

- Si se multiplica el cambio relativo en Y por 100, nos da el cambio porcentual (tasa de crecimiento) en Y ocasionado por un cambio absoluto en X $(100 * \beta_2)$

- Modelo de crecimiento vs modelo de tendencia lineal

Ejemplo: tasa de crecimiento del gasto en servicios

UN EJEMPLO ILUSTRATIVO:
LA TASA DE CRECIMIENTO
DEL GASTO EN SERVICIOS

Para ilustrar el modelo de crecimiento (6.6.6), considere los datos sobre el gasto en servicios proporcionados en la tabla 6.3. Los resultados de la regresión son los siguientes:

$$\begin{aligned} \ln \text{GES}_t &= \quad 7.7890 + 0.00743t \\ \text{ee} &= \quad (0.0023) \quad (0.00017) \\ t &= (3\,387.619)^* \quad (44.2826)^* \quad r^2 = 0.9894 \end{aligned} \tag{6.6.8}$$

Nota: GES significa gasto en servicios y el asterisco (*) denota que el valor p es extremadamente pequeño.

La interpretación de la ec. (6.6.8) es que durante un periodo de un trimestre (del primero al tercero de 1993), el gasto en servicios se incrementó a una tasa (trimestral) de 0.743%. Aproximadamente esto es igual a un crecimiento anual de 2.97%. Puesto que 7.7890 = log de GES al comienzo del periodo de análisis, si se toma su antilogaritmo se tiene 2.41390 (billones de dólares), como el valor inicial de GES (es decir, el valor al final del último trimestre de 1992). La recta de regresión obtenida mediante la ec. (6.6.8) se ilustra en la figura 6.4.

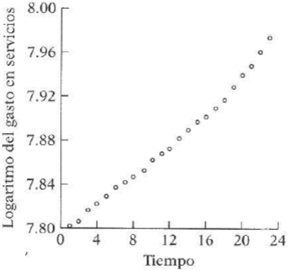


FIGURA 6.4

Modelo lin-log

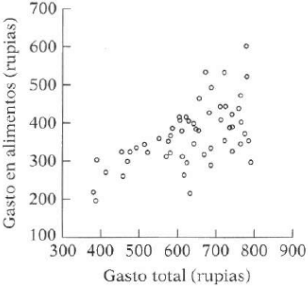
- A veces estamos interesados en saber el cambio absoluto en Y dado un cambio relativo en X (cambio porcentual).
- $Y_i = \beta_1 + \beta_2 lnX_i + \mu_i$
- $\beta_2 = \frac{cambio absoluto en Y}{Cambio relativo en X}$
- Para una buena interpretación, el valor del coeficiente de pendiente se debe dividir por 100.

Ejemplo: Gasto vs gasto en alimentos

UN EJEMPLO ILUSTRATIVO

Como ejemplo del modelo lin-log, revísese el ejemplo sobre gasto alimenticio en India, ejemplo 3.2. Ahí se ajustó un modelo lineal en las variables, como una primera aproximación. Pero si se grafican los datos, se obtiene la gráfica de la figura 6.5. Tal y como esta figura sugiere, el gasto alimenticio se incrementa en forma más lenta, conforme el gasto total aumenta, lo cual quizá proporcione sustento a la ley de Engels. Los resultados de ajustar el modelo lin-log a los datos son los siguientes:

$$\begin{aligned} \text{GASAL}_i &= -1\,283.912 \qquad + 257.2700 \ln \text{GASTOT}_i \\ t &= \quad (-4.3848)^* \qquad (5.6625)^* \qquad r^2 = 0.3769 \end{aligned} \tag{6.6.14}$$

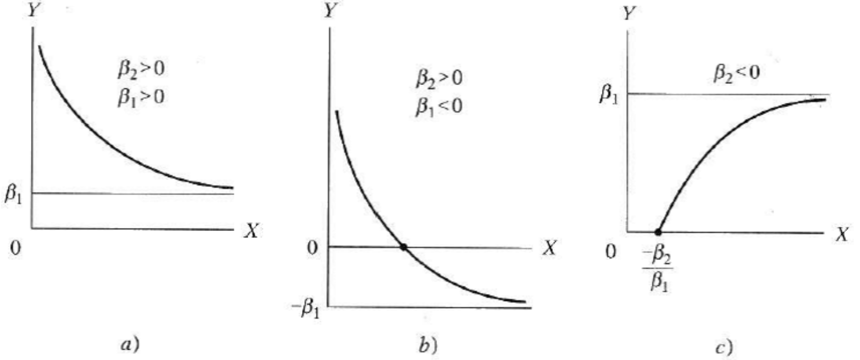


Modelos recíprocos

- $Y_i = \beta_1 + \beta_2 \frac{1}{X_i} + \mu_i$

- A medida que X aumenta indefinidamente, el término $\beta_2 \frac{1}{X_i}$ se acerca a cero, y Y se aproxima al valor límite o asintótico β_1.

Comportamiento del modelo reciproco



Ejemplo: Mortalidad infantil vs PIB per capita

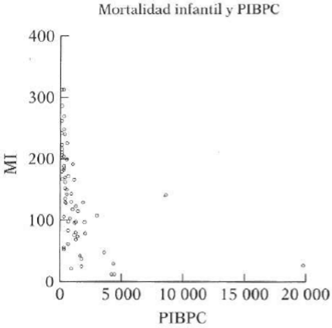


FIGURA 6.7 Relación entre la mortalidad infantil y el PIB *per cápita*, en 66 países.

Si se trata de ajustar el modelo recíproco (6.7.1), se obtienen los siguientes resultados:

$$\begin{aligned} \mathrm{MI}_i &= 81.79436 + 27\,237.17 \left(\frac{1}{\mathrm{PIBPC}_i} \right) \\ \mathrm{ee} &= (10.8321) \quad (3\,759.999) \end{aligned} \qquad \textbf{(6.7.2)}$$

Ejemplo tomado del taller de la clase anterior

Figura 1: **Resultados estimación ecuación de salarios a nivel comunal en chile, modelo normal.**

Source	SS	df	MS
Model	3651743.77	5	730348.754
Residual	1546443.08	324	4772.97246
Total	5198186.84	329	15799.96

Number of obs =	330
F(5, 324) =	153.02
Prob > F =	0.0000
R-squared =	0.7025
Adj R-squared =	0.6979
Root MSE =	69.087

sxh	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
esc	105.5124	5.123256	20.59	0.000	95.43334	115.5914
exper	-19.88702	18.01306	-1.10	0.270	-55.32435	15.55031
exper2	.8039632	.341007	2.36	0.019	.1330957	1.474831
populationc	-.0296257	.0215128	-1.38	0.169	-.071948	.0126967
tdocc	-8.294676	1.154945	-7.18	0.000	-10.56682	-6.022538
_cons	-698.3422	192.0285	-3.64	0.000	-1076.122	-320.562

Ejemplo tomado del taller de la clase anterior

Figura 2: **Resultados estimación ecuación de salarios a nivel comunal en chile, modelo con coeficientes estandarizados.**

sxh	Coef.	Std. Err.	t	P>|t|	Beta
esc	105.5124	5.123256	20.59	0.000	1.088288
exper	-19.88702	18.01306	-1.10	0.270	-.2620462
exper2	.8039632	.341007	2.36	0.019	.5773316
populationc	-.0296257	.0215128	-1.38	0.169	-.0527535
tdocc	-8.294676	1.154945	-7.18	0.000	-.2261092
_cons	-698.3422	192.0285	-3.64	0.000	.

Ejemplo tomado del taller de la clase anterior

Figura 3: **Resultados estimación ecuación de salarios a nivel comunal en chile, modelo log-log.**

Source	SS	df	MS
Model	21.487408	5	4.2974816
Residual	8.90832332	323	.027579948
Total	30.3957313	328	.092669912

Number of obs	=	329
F(5, 323)	=	155.82
Prob > F	=	0.0000
R-squared	=	0.7069
Adj R-squared	=	0.7024
Root MSE	=	.16607

lsxh	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lesc	2.462647	.1275451	19.31	0.000	2.211723	2.713571
lexper	.6490439	1.103312	0.59	0.557	-1.52154	2.819628
lexper2	.1950443	.6698844	0.29	0.771	-1.122843	1.512932
lpopulationc	-.0236082	.0109441	-2.16	0.032	-.0451389	-.0020776
ltdocc	-.211646	.0290758	-7.28	0.000	-.2688479	-.1544441
_cons	-2.49949	1.487045	-1.68	0.094	-5.425008	.4260269

Ejemplo tomado del taller de la clase anterior

Figura 4: **Resultados estimación ecuación de salarios a nivel comunal en chile, modelo log-lin.**

Source	SS	df	MS
Model	22.3168594	5	4.46337187
Residual	8.38104873	324	.025867434
Total	30.6979081	329	.093306712

Number of obs	=	330
F(5, 324)	=	172.55
Prob > F	=	0.0000
R-squared	=	0.7270
Adj R-squared	=	0.7228
Root MSE	=	.16083

lsxh	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
esc	.2510428	.0119269	21.05	0.000	.2275789	.2745068
exper	.038609	.0419343	0.92	0.358	-.0438889	.121107
exper2	.0002923	.0007939	0.37	0.713	-.0012695	.0018541
populationc	-.0000172	.0000501	-0.34	0.732	-.0001157	.0000814
tdocc	-.0200203	.0026887	-7.45	0.000	-.0253098	-.0147307
_cons	2.475142	.4470417	5.54	0.000	1.595671	3.354612

Ejemplo tomado del taller de la clase anterior

Figura 5: **Resultados estimación ecuación de salarios a nivel comunal en chile, modelo lin-log.**

Source	SS	df	MS
Model	3266170.85	5	653234.169
Residual	1889550.17	323	5850.00053
Total	5155721.02	328	15718.6616

Number of obs =	329
F(5, 323) =	111.66
Prob > F =	0.0000
R-squared =	0.6335
Adj R-squared =	0.6278
Root MSE =	76.485

sxh	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lesc	935.795	58.74152	15.93	0.000	820.2307	1051.359
lexper	2.475876	508.1353	0.00	0.996	-997.1969	1002.149
lexper2	159.5012	308.5184	0.52	0.606	-447.4581	766.4605
lpopulationc	-12.3712	5.040346	-2.45	0.015	-22.28725	-2.455145
ltdocc	-91.14955	13.39102	-6.81	0.000	-117.4942	-64.80492
_cons	-2556.622	684.8658	-3.73	0.000	-3903.983	-1209.262