142

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ECONOMÍA

Cálculo del Coeficiente de Gini, Caso de México, años 1994, 1996, 1998 y 2000

E N S A Y O

QUE PARA OBTENER EL TITULO DE LICENCIADO EN ECONOMIA

P R E S E N T A :

FRANCISCO REYNALDO RUIZ NEGRETE

ASESOR: MTRO. CARLOS MARTINEZ FAGUNDO

MÉXICO, D.F. DICIEMBRE DE 2002

INDICE

Introducción. 3

I. Relación Curva de Lorenz e Indice de Gini. 4

II. Definición geométrica del Índice de Gini. 8

III. Cálculo del Índice de Gini, mediante el ajuste de curvas. 11

3.1 Área de máxima desigualdad. 11
3.2 Área de concentración. 13

IV. Construcción de la Curva de Lorenz, utilizando el paquete Econometrics E-views, México años, 1994, 1996, 1998 Y 2000. 15

4.1 Método ordinario de mínimos cuadrados. 15
4.2 Paquete E-views. 18

V. Cálculo del coeficiente de Gini, caso de México, años 1994, 1996, 1998 Y 2000. 27

VI. Interpretación del Índice de Gini y de la Curva de Lorenz. 33

Conclusiones. 37

Bibliografía. 38

INTRODUCCION

El presente ensayo tiene la finalidad de enseñar el método para el cálculo del Índice de Gini y el de la Curva de Lorenz, como instrumentos estadísticos del análisis económico y, más específicamente para el estudio de la desigualdad social.
Los datos estadísticos son tomados de la Encuesta Nacional de Ingresos y Gastos de los Hogares, publicada por el (INEGI) Instituto Nacional de Estadística Geografía e Informática.
La información se agrupa por deciles, de **observaciones** (número de hogares) eje horizontal, y de la **variable** (ingreso) eje vertical, dando para cada uno frecuencias relativas, acumuladas de observaciones y de la variable.
Los cálculos que se hacen de los datos agrupados puede ser manual o mediante procedimientos electrónicos como el uso de un software de windows, en forma general lo es Excel y otro programa compatible E-views, este último es un programa de econometría para resolver problemas de cálculos del Indice de Gini y la Curva de Lorenz, y otros entornos de aplicacilón en el análisis financiero, previsiones macroeconómicas, previsiones de ventas y análisis del costo.
Es evidente que el desarrollo de la estadística y su aplicación mediante la econometría proporcionan un conjunto de conocimientos en las diversas actividades de nuestra sociedad.
Las técnicas desarrolladas proporcionan resultados más rápidos para su interpretación.

I. RELACION CURVA DE LORENZ E INDICE DE GINI

Para juzgar el problema de la desigualdad existente en una distribución que regularmente se encuentra bajo el influjo de juicios de valor, es objeto de discusiones que no pueden garantizar un acuerdo, por la diversidad de posturas teóricas, políticas y metodológicas irreconciliables, por tal motivo la técnica estadística que se usa para estudiar la desigualdad social establece en forma explícita e implícitamente el criterio de igualdad perfecta o democrática, es decir, a todos les debe corresponder la misma cantidad.
Conceptualmente Gini, como medida de desigualdad se construye sobre la base de los desvíos respecto a la norma democrática, cristalizada por la media aritmética, se puede entender como un resumen de las discrepancias, aunque no se exprese en términos de un promedio, es decir, se presenta e interpreta como una medida que reúne en un solo valor las comparaciones entre los valores de la variable que se desprenden del universo de las observaciones.
El coeficiente de Gini formaliza la norma democrática, estableciendo la distribución teórica, repartiendo por igual la variable entre todas las unidades.
Si la variable esta distribuida equitativamente con respecto a la norma democrática la proporción de la variable perteneciente a cada unidad $(\mathbf{q_i^t})$[1] debiese ser igual a la proporción que cada observación representa dentro del total $(\mathbf{p_i})$[2], por lo tanto:

$$\mathbf{q_i^t = \frac{X_i^t}{\Sigma X_i}}$$

$\mathbf{X_i^t}$ representa el valor que debiera poseer la **i-ésima** observación en el caso de repartición perfectamente democrática de la variable, además $\mathbf{\Sigma X_i^t = \Sigma X_i}$

$$\mathbf{p_i = \frac{i}{n}}$$

Con estas fórmulas se puede calcular las participaciones relativas en la variable y en las observaciones respectivamente.
Las frecuencias relativas acumuladas tanto de las observaciones $\mathbf{P_i}$ y de la variable $\mathbf{Q_i^t}$, van constituyendo los elementos para obtener la medida de concentración de Gini, esto es, el reflejo de las discrepancias o las desviaciones respecto de la norma democrática. La igualdad $\mathbf{p_i = q_i^t}$ necesariamente conduce a $\mathbf{P_i = Q_i^t}$, siendo consecuencia de construir una - -

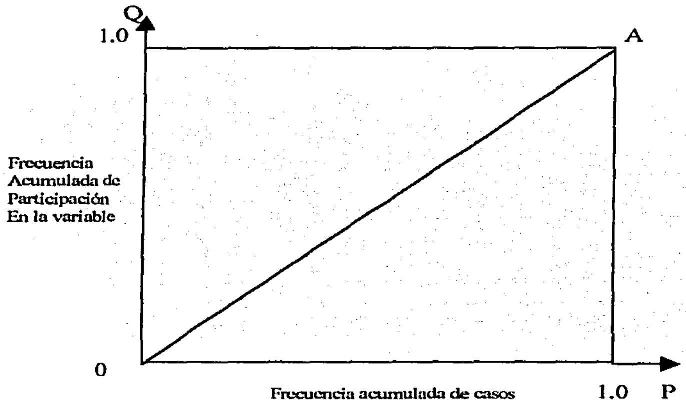
---

1 Con datos reales del INEGI y, específicamente de la Encuesta Ingreso Gasto de los Hogares, la participación relativa de la variable se calcula mediante la fórmula $q_i^t = \frac{X_i^t}{\Sigma X_i}$, en la que X_i^t es igual al ingreso corriente de cada decil y, la ΣX_i es
cada decil y, la ΣX_i es igual al ingreso corriente total de los hogares.

2 Para el caso de la participación relativa de las observaciones, se calcula mediante la fórmula $p_i = \frac{i}{n}$, en donde **i** es igual al número de hogares de cada decil y **n** es igual al total de hogares. El cálculo se realiza en el capítulo IV y V.

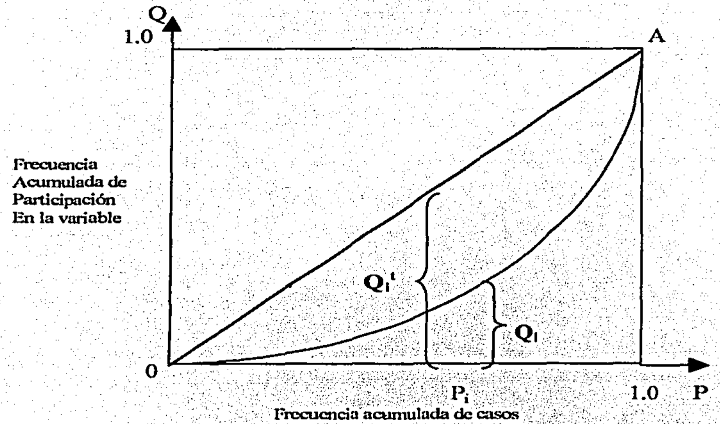
distribución teórica de frecuencias sustentado en la norma democrática. De ahí que la frecuencia relativa acumulada de la variable coincida con la frecuencia relativa acumulada de las observaciones, distribuyendo equitativamente un total, por ende, a una determinada proporción de los casos le corresponderá la misma proporción en la variable, es decir, por ejemplo el 0.25 de los casos acumulados le corresponde el 0.25 acumulado de la variable, al 0.50 de los casos acumulado le corresponde el 0.50 acumulado de la variable y así sucesivamente.

Representación de la distribución "equitativa"



El eje horizontal (P) representa la frecuencia acumulada de casos u observaciones y en el vertical (Q) la frecuencia relativa acumulada en el total de la variable. Por lo tanto, los valores máximos para ambas frecuencias relativas acumuladas son iguales a 1.
La diagonal del cuadrado formado por los puntos (0, 1.0, A Y 1.0), expresa de manera idealizada la distribución teórica construida sobre la base del principio de la repartición igualitaria, más conocida como línea de equidistribución o de igualdad perfecta.
Esto orienta la construcción del índice de Gini, que consiste en comparar dos distribuciones: La empírica y la que se obtiene de aplicar el principio de la repartición igualitaria (norma democrática).
Si representamos la distribución de los datos reales en la gráfica que se dibujó la distribución teórica, la línea resultante se encontrará siempre debajo de la recta de equidistribución, excepto en los valores extremos 0 y 1, a esta última línea es la que se conoce como curva de Lorenz.
La curva de Lorenz, se asemeja a una sección de parábola y se compara con una recta de distribución equitativa de 45°, cuanto más se aparta la curva de Lorenz de la recta de distribución equitativa, tanto mayor será el grado de desigualdad y de concentración del ingreso.

Gráfica de concentración



Para una proporción acumulada de unidades **(P_i)** mayor será la concentración en tanto mayor sea la distancia vertical **($d_i = P_i - Q_i$)** entre el punto ubicado sobre la curva de Lorenz y el situado sobre la línea de equidistribución. Al contrario, menor será la desigualdad en tanto menor sea la distancia entre ambos puntos.

Si se desea construir un índice de concentración éste debería basarse en la diferencia d_i:

$$\mathbf{d_i = P_i - Q_i}$$

Una observación i cualquiera será cero si coincide el "punto de equidistribución" con el "punto Lorenz". Para este caso

$$\mathbf{d_i = P_i - Q_i = P_i - P_i = 0}$$

ya que $\mathbf{Q_i = P_i}$

Un punto cualquiera i, contribuye a la concentración cuando $\mathbf{Q_i = 0}$, es decir, cuando la participación acumulada de la variable sea nula, corresponde una proporción acumulada de observaciones.
Luego **$d_i = P_i - Q_i = P_i = i / n$, ya que $Q_i = 0$.**

Los valores de las diferencias d_i estarán dentro del intervalo $0 \leq d_i \leq i / n$, esto implica que habrá tantos valores d_i como observaciones haya, es decir, como $i = 1, 2, \ldots, n$, habrá $d_1, d_2, \ldots, d_n$ diferencias, la **n-ésima** no contribuye a la desigualdad porque siempre es igual a cero:

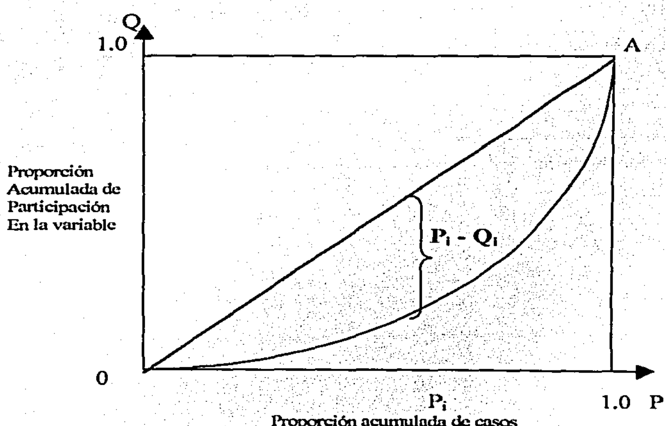
$$d_n = P_n - Q_n = 1 - 1 = 0.$$

En todos los casos excepto en la equidistribución se cumplirá que $P_i > Q_i$, por lo que las discrepancias serán positivas, por tanto su suma:

$$\sum_{i=1}^{n-1} d_i = \sum_{i=1}^{n-1} (P_i - Q_i)$$

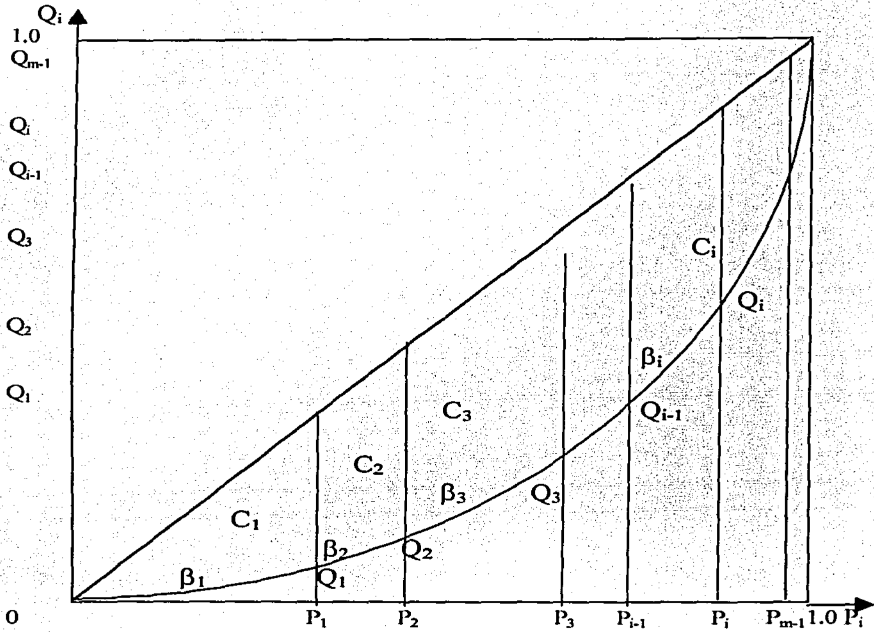
Estas discrepancias, gráficamente corresponden a las distancias verticales existentes entre la línea de equidistribución y la curva de Lorenz. La superficie delimitada por ambas líneas usualmente es conocida con el nombre de área de concentración, guardando una relación estrecha con las discrepancias d_i a partir de las cuales se construye el índice de Gini.

Área de concentración



II. DEFINICIÓN GEOMÉTRICA DEL INDICE DE GINI.

Para establecer un procedimiento adecuado de cálculo del coeficiente de Gini, hay que buscar la manera de cubrir el área de concentración (comprendida entre la línea de equidistribución y la curva de Lorenz) mediante la información contenida en la tabla de datos agrupados.



La fórmula de Gini tiene su explicación en el desarrollo geométrico, la cuál se evaluará analizando el área de concentración, usando un procedimiento indirecto que simplifica su cálculo, primero es necesario obtener el área:

donde $\beta > C$, $\Sigma\,(\beta_i - C_i) = (\beta - C)$ marcada los limites del eje de abscisas, la poligonal de Lorenz y la vertical imagen al eje de ordenadas levantada en el punto $P_m = 1.0$.

En la gráfica β_i simboliza el área debajo a la línea de equidistribución o de igualdad perfecta que representa a la parte del eje de las abscisas P_{i-1} y a los segmentos verticales de las ordenadas Q_i y Q_{i-1}. C_i representa el área de concentración que corresponde al **i-ésimo** intervalo, comprendida entre la línea de equidistribución y la poligonal de Lorenz.

De acuerdo a la explicación anterior el coeficiente de Gini asume la forma:

$$G = \frac{\frac{1}{2} - (\beta - C)}{\frac{1}{2}} = 1 - 2\,(\beta - C)$$

Iniciamos calculando el área $(\beta - C)$, la cual descomponemos en una serie de triángulos y rectángulos, tal como se aprecia en la gráfica. Al aplicar este principio de manera sistemática, se obtiene:

$$(\beta_1 - C_1) = \tfrac{1}{2} P_1\, Q_1$$

$$(\beta_2 - C_2) = (P_2 - P_1)\, Q_1 + \tfrac{1}{2}(P_2 - P_1)(Q_2 - Q_1) =$$

$$= 2\tfrac{1}{2}\,[\, P_2\, Q_1 - P_1\, Q_1 + \tfrac{1}{2}(P_2\, Q_2 - P_1\, Q_2 - P_2\, Q_1 + P_1\, Q_1)\,] =$$

$$= \tfrac{1}{2}\,[\, 2\,(P_2\, Q_1 - P_1\, Q_1) + (P_2\, Q_2 - P_1\, Q_2 - P_2\, Q_1 + P_1\, Q_1)\,] =$$

$$= \tfrac{1}{2}\,[\,(P_2 - P_1)\, Q_1 + (P_2 - P_1)\, Q_2\,] =$$

$$= \tfrac{1}{2}(P_2 - P_1)(Q_1 + Q_2) = \tfrac{1}{2}\, \mathbf{p_2}\, (\mathbf{Q_1 + Q_2})$$

$$(\beta_3 - C_3) = (P_3 - P_2)\, Q_2 + \tfrac{1}{2}(P_3 - P_2)(Q_3 - Q_2) =$$

$$= 2\tfrac{1}{2}\,[\, P_3\, Q_2 - P_2\, Q_2 + \tfrac{1}{2}(P_3\, Q_3 - P_2\, Q_3 - P_3\, Q_2 + P_2\, Q_2)\,] =$$

$$= \tfrac{1}{2}\,[\, 2\,(P_3\, Q_2 - P_2\, Q_2) + (P_3\, Q_3 - P_2\, Q_3 - P_3\, Q_2 + P_2\, Q_2)\,] =$$

$$= \tfrac{1}{2}\,[\,(P_3 - P_2)\, Q_2 + (P_3 - P_2)\, Q_3\,] =$$

$$= \tfrac{1}{2}(P_3 - P_2)(Q_2 + Q_3) = \tfrac{1}{2}\, \mathbf{p_3}\, (\mathbf{Q_2 + Q_3})$$

$$(\beta_i - C_i) = (P_i - P_{i-1})\, Q_{i-1} + \tfrac{1}{2}(P_i - P_{i-1})(Q_i - Q_{i-1}) =$$

$$= 2\tfrac{1}{2}\,[\, P_i\, Q_{i-1} - P_{i-1}\, Q_{i-1} + \tfrac{1}{2}(P_i\, Q_i - P_{i-1}\, Q_i - P_i\, Q_{i-1} + P_{i-1}\, Q_{i-1})\,] =$$

$$= \tfrac{1}{2}\,[2\,(P_i\, Q_{i-1} - P_{i-1}\, Q_{i-1}) + (P_i\, Q_i - P_{i-1}\, Q_i - P_i\, Q_{i-1} + P_{i-1}\, Q_{i-1})] =$$

$$= \tfrac{1}{2}\,[(P_i - P_{i-1})\, Q_{i-1} + (P_i - P_{i-1})\, Q_i] =$$

$$= \tfrac{1}{2}(P_i - P_{i-1})(Q_{i-1} + Q_i) = \tfrac{1}{2}\, \mathbf{p_i}\, (\mathbf{Q_{i-1} + Q_i})$$

$$(\beta_n - C_n) = (P_n - P_{n-1})\, Q_{n-1} + ½\,(P_n - P_{n-1})\,(Q_n - Q_{n-1}) =$$

$$= 2½\,[\, P_n Q_{n-1} - P_{n-1} Q_{n-1} + ½\,(P_n Q_n - P_{n-1} Q_n - P_n Q_{n-1} + P_{n-1} Q_{n-1})\,] =$$

$$= ½\,[\, 2\,(P_n Q_{n-1} - P_{n-1} Q_{n-1}) + (P_n Q_n - P_{n-1} Q_n - P_n Q_{n-1} + P_{n-1} Q_{n-1})] =$$

$$= ½\,[(P_n - P_{n-1})\, Q_{n-1} + (P_n - P_{n-1})Q_n] = ½\,(P_n - P_{n-1})\,(Q_{n-1} + Q_n) =$$

$$= ½\; p_n\,(Q_{n-1} + Q_n)$$

La expresión $(\beta - C) = ½\; p_i\,(Q_{i-1} + Q_i)$ la sustituimos en

$$G = \frac{½ - (\beta - C)}{½} = 1 - 2\,(\beta - C)$$

y obtenemos $G = 1 - 2\,[\, ½\; p_i\,(Q_{i-1} + Q_i)\,]$

$$G = 1 - \sum_{i=1}^{m} p_i\,(Q_{i-1} + Q_i)$$

Esta es la manera más usual para efectuar el cálculo del coeficiente de Gini para datos agregados que se obtiene de descomponer el área en una serie de triángulos y rectángulos que sumados convenientemente, da como resultado un procedimiento analítico que hace más sencillo el cálculo del índice de Gini.

III. CALCULO DEL INDICE DE GINI, MEDIANTE EL AJUSTE DE CURVAS

3.1 AREA DE MÁXIMA DESIGUALDAD

Otro camino para encontrar el índice de Gini es mediante el uso de la integral definida, la cuál se emplea en el cálculo de áreas bajo y entre curvas. Mediante este procedimiento se tiene que calcular el área de máxima desigualdad, área comprendida debajo de la línea diagonal que forma un ángulo de 45°, y el eje de las abscisas P_i y el de las ordenadas Q_i, como se muestra en la gráfica.
El área que se desea calcular esta determinada por $P_m = 0$ y $P_m = 1$, y para $Q_m = 0$ y $Q_m = 1$, los valores p_i y qi son iguales, es decir, la igualdad $p_i = q_i^t$ conduce a $P_i = Q_i^t$, en otras palabras, a cada proporción de P_i de casos le corresponde la misma proporción de Q_i en la variable, bajo estas condiciones se puede establecer que la función seria $f(P) = P$, esta área será máxima si los $\mathbf{Q_i}$ correspondientes a los primeros **(m – 1)** intervalos son iguales a cero. Luego:

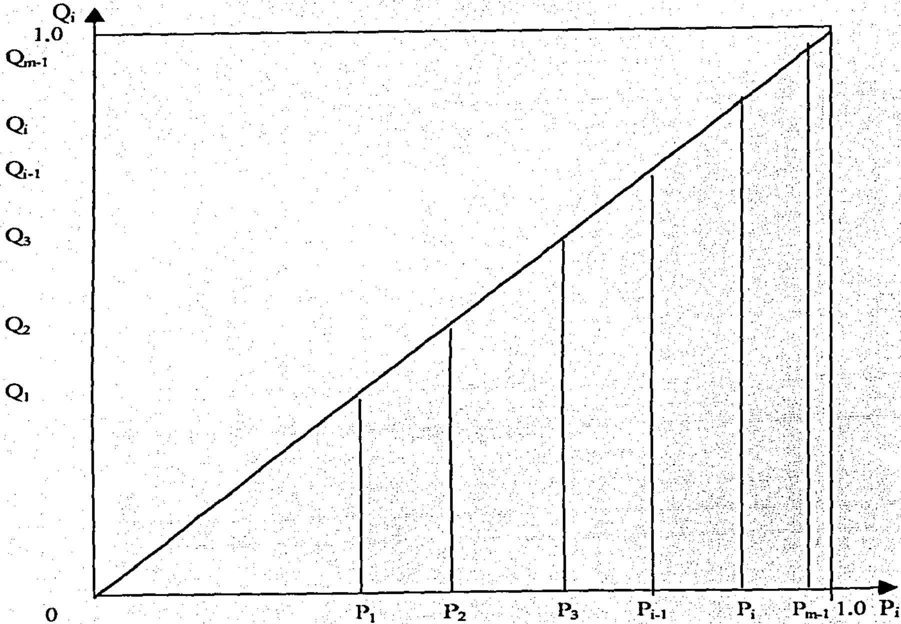
$$\int_0^1 (P_i - Q_i)\, d\, P_i = \int_0^1 P_i\, d\, P_i$$

Sustituyendo la solución será

$$\int_0^1 P_i\, d\, P_i = \frac{P_i^2}{2}\Big]_0^1 = \frac{1^2}{2} - \frac{0^2}{2} = \frac{1}{2} - 0 = 0.5$$

0.5 es el valor del área de máxima desigualdad.

Área de máxima desigualdad



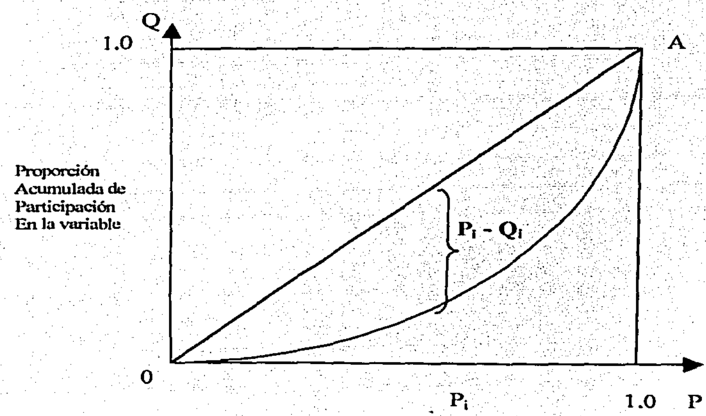
3.2 ÁREA DE CONCENTRACION

El área de concentración esta limitada gráficamente por la línea de distribución perfecta y la curva de Lorenz. Para calcular el área de concentración se aproxima la poligonal de Lorenz por una línea continua permitiendo obtener el área de concentración a través de:

$$\int_0^1 (P_i - Q_i)\, d\, P_i$$

Hay que hacer notar que no se conoce la función por eso es necesario realizar un proceso de ajuste de la curva o poligonal de Lorenz, como se muestra en el capítulo IV con la aplicación del método ordinario de mínimos cuadrados incluido en el paquete E-Views.

Área de concentración



Proporción acumulada de casos

Ya conociendo los datos de las dos integrales que respectivamente representan el área de máxima desigualdad y el del área de concentración, por lo tanto se puede encontrar a Gini, pero hay que recordar que el índice de Gini se define como una relación entre el área de concentración y el área de máxima desigualdad:

$$G = \frac{\int_0^1 (P_i - Q_i)\, d\, P_i}{\int_0^1 P_i\, d\, P_i}$$

Esta manera de calcular el índice de Gini es útil cuando es posible lograr un buen ajuste de una línea continua a los puntos del diagrama de concentración y, que al mismo tiempo, estos puntos sean suficientes como para obtener una estimación confiable de los parámetros de la ecuación.

IV. CONSTRUCCIÓN DE LA CURVA DE LORENZ UTILIZANDO EL PAQUETE ECONOMETRICS E-VIEWS, MEXICO, AÑOS 1994, 1996, 1998 Y 2000.

E-Views proporciona una manera rápida para el análisis de datos sofisticados, permitiendo desarrollar una relación estadística como una regresión, el problema general de hallar ecuaciones de curvas aproximantes que se ajusten a un conjunto de datos, es decir, se estima el valor de la variable Y a un valor dado de la variable X (método de mínimos cuadrados) y la curva resultante es una curva de regresión. E-views además es de gran utilidad para datos de análisis financieros, previsiones macroeconómicas, simulación, previsión de ventas y análisis del costo.

Lo que se pretende analizar es como este programa ahorra mucho tiempo y permite obtener resultados más fidedignos para el estudio de los problemas de la desigualdad. Los resultados de las ecuaciones para los cuatro años estudiados son los mismos, utilizando el paquete de E-Views, donde se incluye el método ordinario de mínimos cuadrados.

4.1 MÉTODO ORDINARIO DE MÍNIMOS CUADRADOS.

Con los datos de la tabla 1 columna 5 para el año 1994, se va construyendo la tabla 9 y, la función a utilizar es una parabólica o cuadrática que ajusta los datos, es la siguiente:

$$a_0 + a_1X + a_2X^2 = Y$$

deduciéndose las siguientes ecuaciones:

1. $a_0N + a_1\Sigma X + a_2\Sigma X^2 = \Sigma Y$
2. $a_0\Sigma X + a_1\Sigma X^2 + a_2\Sigma X^3 = \Sigma XY$
3. $a_0\Sigma X^2 + a_1\Sigma X^3 + a_2\Sigma X^4 = \Sigma X^2Y$

9. Tabla

$P_i = X$	$Q_i = Y$	X^2	X^3	X^4	XY	X^2Y
-5	0.0000	25	- 125	625	0.0000	0.0000
-4	0.0159	16	- 64	256	-0.0636	0.2544
-3	0.0435	9	- 27	81	-0.1305	0.3915
-2	0.0802	4	- 8	16	-0.1604	0.3208
-1	0.1266	1	- 1	1	-0.1266	0.1266
0	0.1833	0	0	0	0.0000	0.0000
1	0.2539	1	1	1	0.2539	0.2539
2	0.3413	4	8	16	0.6826	1.3652
3	0.4547	9	27	81	1.3641	4.0923
4	0.6158	16	64	256	2.4632	9.8528
5	1.0000	25	125	625	5.0000	25.0000
Σ 0	Σ 3.1152	Σ 110	Σ 0	Σ 1958	Σ 9.2827	Σ 41.6575

Las ecuaciones 1, 2 y 3 se simplifican porque la ΣX y ΣX^3 son cero, para las columnas 1 y 4, tabla 9. Sustituir los valores de la tabla en las ecuaciones anteriores:

1. $a_0 N + a_1 \Sigma X + a_2 \Sigma X^2 = \Sigma Y$ $\qquad 11a_0 + 0a_1 + 110a_2 = 3.1152$

2. $a_0 \Sigma X + a_1 \Sigma X^2 + a_2 \Sigma X^3 = \Sigma XY$ $\qquad 0a_0 + 110a_1 + 0a_2 = 9.2827$

3. $a_0 \Sigma X^2 + a_1 \Sigma X^3 + a_2 \Sigma X^4 = \Sigma X^2 Y$ $\qquad 110a_0 + 0a_1 + 1958a_2 = 41.6575$

1. $a_0 = \dfrac{3.1152 - 110a_2}{11}$ sustituir en la ecuación 3

$$110 \left[\frac{3.1152 - 110a_2}{11} \right] + 1958a_2 = 41.6575$$

$$\left[\frac{342.672 - 12100a_2}{11} \right] + 1958a_2 = 41.6575$$

$$31.152 - 1100a_2 + 1958a_2 = 41.6575$$

$$a_2 = \frac{41.6575 - 31.152}{858} = \frac{10.5055}{858} = 0.0122441$$

sustituir $a_2 = 0.0122441$ en la ecuación 1

$$a_0 = \frac{3.1152 - 110a_2}{11}$$

$$a_0 = \frac{3.1152 - 110\,(0.0122441)}{11} = \frac{3.1152 - 1.34685}{11} = \frac{1.768349}{11} = 0.160759$$

se despeja la ecuación 2. $\mathbf{0a_0} + 110a_1 + \mathbf{0a_2} = 9.2827$

$$a_1 = \frac{9.2827}{110} = 0.0843881$$

$$Y = a_0 + a_1X + a_2X^2$$

$$f(x) = 0.160759 + 0.0843881X + 0.0122441X^2$$

$$\int_a^b f(x)dx = f(b) - f(a)$$

$$\int_{-5}^{5} f(x)dx = \int_{-5}^{5} f(0.160759 + 0.0843881X + 0.0122441X^2)\,dx$$

$$\int_{-5}^{5} f(x)dx = (0.160759\,X + \frac{0.0843881X^2}{2} + \frac{0.0122441X^3}{3})\,]_{-5}^{5}$$

$$\int_{-5}^{5} f(x)dx = (0.160759\,X + \frac{0.0843881X^2}{2} + \frac{0.0122441X^3}{3}) - (0.160759\,X +$$
$$+ \frac{0.0843881X^2}{2} + \frac{0.0122441X^3}{3}) =$$

$$=[\,0.160759\,(5) + \frac{0.0843881\,(5^2)}{2} + \frac{0.0122441\,(5^3)}{3}\,] -$$

$$-[\,0.160759\,(-5) + \frac{0.0843881(-5^2)}{2} + \frac{0.0122441\,(-5^3)}{3}\,] =$$

$$= (0.803795 + 1.0548512 + 0.5101708) - (-0.803795 + 1.0548512 - 0.5101708) =$$

$= (0.803795 + \mathbf{1.0548512} + 0.5101708 + 0.803795 - \mathbf{1.0548512} + 0.5101708) =$

$= 2.6279316$

Como la escala de valores para P_i era de (0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 y 1.0), ahora para este caso paso de (-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5), entonces el valor del área de máxima desigualdad ya no es 0.5, sino pasa a 5, por lo tanto el Área de Concentración $= 5 - \int_{-5}^{5} f(x)dx = 5 - 2.6279316 = 2.3720684$

1994 $\quad \mathbf{G} = \frac{2.3720684}{5} = \mathbf{0.4744136}$

Realizando el mismo procedimiento de mínimos cuadrados para los años 1996, 1998 y 2000, el valor del índice será:

1996 $\quad \mathbf{G} = \frac{2.2680634}{5} = \mathbf{0.4536126}$

1998 $\quad \mathbf{G} = \frac{2.368501}{5} = \mathbf{0.4737002}$

2000 $\quad \mathbf{G} = \frac{2.3925974}{5} = \mathbf{0.47851948}$, aproximando el valor del índice es

$$\mathbf{G = 0.4811}$$

4.2 PAQUETE E-VIEWS

Una manera más rápida en la que se puede eliminar procesos tradicionales que están fuera de época, es la utilización de software como el econometrics E-Views, permiten tener los resultados más rápidos, como se anotan a continuación para los años 1994, 1996, 1998 y 2000. Tomando en consideración la distribución de frecuencias acumuladas de la variable (con los que se construye la curva de Lorenz), que son los datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH), tercer trimestre 1994, 1996, 1998 y 2000, México INEGI, en el que la variable OBS (es el total de hogares ordenados por deciles $\mathbf{P_i}$) son las frecuencias acumuladas de las observaciones, hay que aclarar que por el método de mínimos cuadrados se le tiene que dar un rango de –5 a +5 y, FACUM es el ingreso corriente total trimestral o las frecuencias acumuladas de participación en la variable $\mathbf{Q_i}$, es decir, son los datos del que se obtienen en el capítulo V de la tabla 1, 2, 3, 4 columna 5.

Tabla E1

<table>
<tr><th colspan="6">México año 1994</th></tr>
<tr><th>obs</th><th>OBS
X</th><th>FACUM
Actual (Y)</th><th>Fitted
Ajustado</th><th>Residual</th><th>Residual Plot</th></tr>
<tr><td>1990</td><td>-5.000000</td><td>0.000000</td><td>0.04492</td><td>-0.044922</td><td>| .* | . |</td></tr>
<tr><td>1991</td><td>-4.000000</td><td>0.015900</td><td>0.01911</td><td>-0.003212</td><td>| . * . |</td></tr>
<tr><td>1992</td><td>-3.000000</td><td>0.043500</td><td>0.01779</td><td>0.025709</td><td>| . |* . |</td></tr>
<tr><td>1993</td><td>-2.000000</td><td>0.080200</td><td>0.04096</td><td>0.039241</td><td>| . | * . |</td></tr>
<tr><td>1994</td><td>-1.000000</td><td>0.126600</td><td>0.08861</td><td>0.037986</td><td>| . | * . |</td></tr>
<tr><td>1995</td><td>0.000000</td><td>0.183300</td><td>0.16076</td><td>0.022542</td><td>| . |* . |</td></tr>
<tr><td>1996</td><td>1.000000</td><td>0.253900</td><td>0.25739</td><td>-0.003491</td><td>| . * . |</td></tr>
<tr><td>1997</td><td>2.000000</td><td>0.341300</td><td>0.37851</td><td>-0.037211</td><td>| . * | . |</td></tr>
<tr><td>1998</td><td>3.000000</td><td>0.454700</td><td>0.52412</td><td>-0.069420</td><td>| *. | . |</td></tr>
<tr><td>1999</td><td>4.000000</td><td>0.615800</td><td>0.69422</td><td>-0.078418</td><td>| *. | . |</td></tr>
<tr><td>2000</td><td>5.000000</td><td>1.000000</td><td>0.88880</td><td>0.111197</td><td>| . | . *|</td></tr>
</table>

México 1994

Dependent Variable: FACUM
Method: Least Squares
Date: 04/20/02 Time: 14:26
Sample: 1990 2000
Included observations: 11
FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.160758	0.028312	5.678073	0.0005
C(2)	0.084388	0.005927	14.23875	0.0000
C(3)	0.012244	0.002122	5.769884	0.0004

R-squared	0.967218	Mean dependent var	0.283200
Adjusted R-squared	0.959022	S.D. dependent var	0.307065
S.E. of regression	0.062159	Akaike info criterion	-2.491232
Sum squared resid	0.030910	Schwarz criterion	-2.382715
Log likelihood	16.70178	Durbin-Watson stat	1.355087

Estimation Command:
=====================
LS FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Estimation Equation:
=====================
FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Substituted Coefficients:
=====================
FACUM=0.1607582751+0.08438818182*OBS+0.01224417249*(OBS)^2

Curva de Lorenz
México 1994

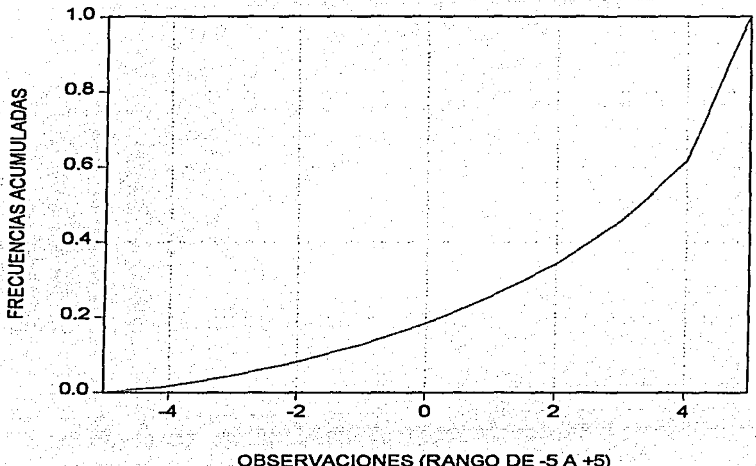


Tabla E2

<table>
<tr><th colspan="6">México año 1996</th></tr>
<tr><th>obs</th><th>OBS
X</th><th>FACUM
Actual (Y)</th><th>Fitted
Ajustado</th><th>Residual</th><th>Residual Plot</th></tr>
<tr><td>1990</td><td>-5.000000</td><td>0.000000</td><td>0.04139</td><td>-0.041388</td><td>| .* | . |</td></tr>
<tr><td>1991</td><td>-4.000000</td><td>0.017900</td><td>0.02102</td><td>-0.003121</td><td>| . * . |</td></tr>
<tr><td>1992</td><td>-3.000000</td><td>0.047900</td><td>0.02421</td><td>0.023695</td><td>| . |* . |</td></tr>
<tr><td>1993</td><td>-2.000000</td><td>0.087300</td><td>0.05094</td><td>0.036359</td><td>| . | * . |</td></tr>
<tr><td>1994</td><td>-1.000000</td><td>0.136300</td><td>0.10123</td><td>0.035071</td><td>| . | * . |</td></tr>
<tr><td>1995</td><td>0.000000</td><td>0.196000</td><td>0.17507</td><td>0.020931</td><td>| . |* . |</td></tr>
<tr><td>1996</td><td>1.000000</td><td>0.269200</td><td>0.27246</td><td>-0.003260</td><td>| . * . |</td></tr>
<tr><td>1997</td><td>2.000000</td><td>0.358800</td><td>0.39340</td><td>-0.034603</td><td>| . * | . |</td></tr>
<tr><td>1998</td><td>3.000000</td><td>0.473700</td><td>0.53790</td><td>-0.064198</td><td>|*. | . |</td></tr>
<tr><td>1999</td><td>4.000000</td><td>0.634000</td><td>0.70594</td><td>-0.071944</td><td>|*. | . |</td></tr>
<tr><td>2000</td><td>5.000000</td><td>1.000000</td><td>0.89754</td><td>0.102457</td><td>| . | . *|</td></tr>
</table>

México 1996

Dependent Variable: FACUM
Method: Least Squares
Date: 04/20/02 Time: 15:08
Sample: 1990 2000
Included observations: 11
FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.175069	0.026104	6.706535	0.0002
C(2)	0.085615	0.005464	15.66769	0.0000
C(3)	0.011776	0.001957	6.018569	0.0003

R-squared	0.972385	Mean dependent var	0.292827
Adjusted R-squared	0.965481	S.D. dependent var	0.308473
S.E. of regression	0.057312	Akaike info criterion	-2.653622
Sum squared resid	0.026277	Schwarz criterion	-2.545105
Log likelihood	17.59492	Durbin-Watson stat	1.349650

Estimation Command:
=====================
LS FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Estimation Equation:
=====================
FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Substituted Coefficients:
=====================
FACUM=0.1750685315+0.08561545455*OBS+0.01177587413*(OBS)^2

Curva de Lorenz
México 1996

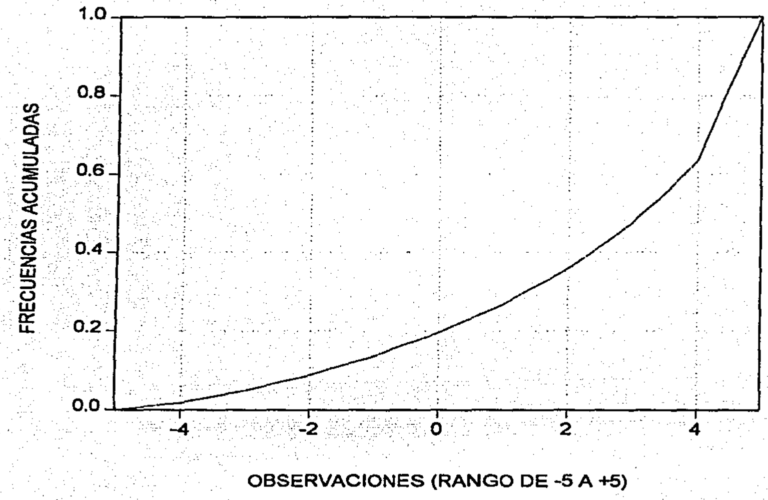


Tabla E3

México año 1998					
obs	OBS X	FACUM Actual (Y)	Fitted Ajustado	Residual	Residual Plot
1990	-5.000000	0.00000	0.04340	-0.04340	| .* | . |
1991	-4.000000	0.01500	0.01785	-0.00285	| . * . |
1992	-3.000000	0.04160	0.01683	0.02477	| . | * . |
1993	-2.000000	0.07790	0.04034	0.03756	| . | *. |
1994	-1.000000	0.12470	0.08838	0.03632	| . | *. |
1995	0.000000	0.18290	0.16094	0.02196	| . |* . |
1996	1.000000	0.25500	0.25804	-0.00304	| . * . |
1997	2.000000	0.34430	0.37967	-0.03537	| . * | . |
1998	3.000000	0.45920	0.52582	-0.06662	| * | . |
1999	4.000000	0.61890	0.69651	-0.07761	| *. | . |
2000	5.000000	1.00000	0.89172	0.10828	| . | . *|

México 1998				
Dependent Variable: FACUM				
Method: Least Squares				
Date: 04/20/02 Time: 15:21				
Sample: 1990 2000				
Included observations: 11				
FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.160944	0.027500	5.852412	0.0004
C(2)	0.084832	0.005757	14.73606	0.0000
C(3)	0.012265	0.002061	5.950129	0.0003
R-squared	0.969296	Mean dependent var		0.283591
Adjusted R-squared	0.961620	S.D. dependent var		0.308194
S.E. of regression	0.060377	Akaike info criterion		-2.549406
Sum squared resid	0.029163	Schwarz criterion		-2.440889
Log likelihood	17.02173	Durbin-Watson stat		1.375006

Estimation Command:
=====================
LS FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Estimation Equation:
=====================
FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Substituted Coefficients:
=====================
FACUM=0.1609440559+0.08483181818*OBS+0.01226468531*(OBS)^2

Curva de Lorenz
México 1998

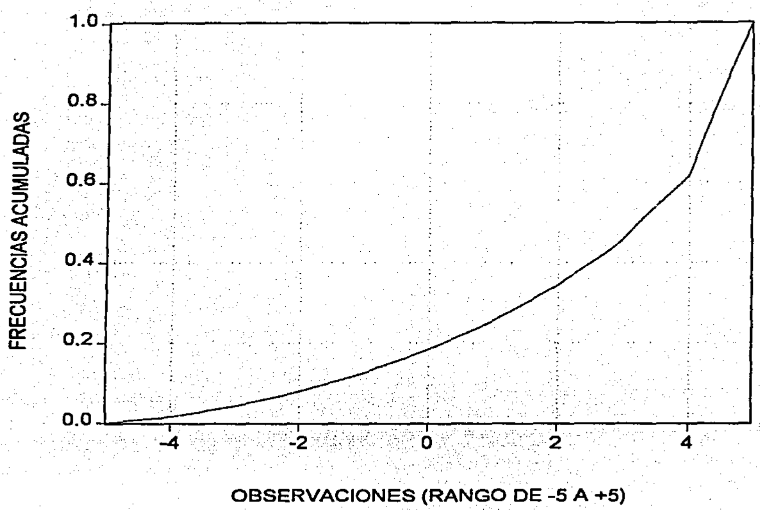


Tabla E4

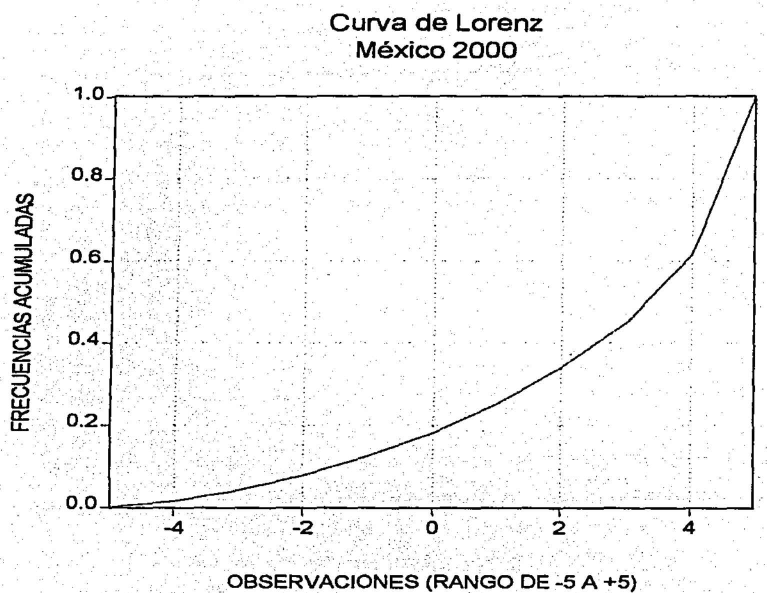
<table>
<tr><th colspan="6">México año 2000</th></tr>
<tr><th>obs</th><th>OBS
X</th><th>FACUM
Actual (Y)</th><th>Fitted
Ajustado</th><th>Residual</th><th>Residual Plot</th></tr>
<tr><td>1990</td><td>-5.000000</td><td>0.00000</td><td>0.04494</td><td>-0.04494</td><td>| .* | . |</td></tr>
<tr><td>1991</td><td>-4.000000</td><td>0.01520</td><td>0.01814</td><td>-0.00294</td><td>| . * . |</td></tr>
<tr><td>1992</td><td>-3.000000</td><td>0.04160</td><td>0.01603</td><td>0.02557</td><td>| . | * . |</td></tr>
<tr><td>1993</td><td>-2.000000</td><td>0.07760</td><td>0.03861</td><td>0.03899</td><td>| . | * . |</td></tr>
<tr><td>1994</td><td>-1.000000</td><td>0.12350</td><td>0.08588</td><td>0.03762</td><td>| . | * . |</td></tr>
<tr><td>1995</td><td>0.000000</td><td>0.18050</td><td>0.15785</td><td>0.02265</td><td>| . |* . |</td></tr>
<tr><td>1996</td><td>1.000000</td><td>0.25130</td><td>0.25451</td><td>-0.00321</td><td>| . * . |</td></tr>
<tr><td>1997</td><td>2.000000</td><td>0.33970</td><td>0.37587</td><td>-0.03617</td><td>| .* | . |</td></tr>
<tr><td>1998</td><td>3.000000</td><td>0.45210</td><td>0.52192</td><td>-0.06982</td><td>| *. | . |</td></tr>
<tr><td>1999</td><td>4.000000</td><td>0.61300</td><td>0.69266</td><td>-0.07966</td><td>| *. | . |</td></tr>
<tr><td>2000</td><td>5.000000</td><td>1.00000</td><td>0.88810</td><td>0.11190</td><td>| . | . *|</td></tr>
</table>

México 2000				
Dependent Variable: FACUM Method: Least Squares Date: 04/20/02 Time: 15:35 Sample: 1990 2000 Included observations: 11 FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.157851	0.028441	5.550118	0.0005
C(2)	0.084315	0.005954	14.16202	0.0000
C(3)	0.012347	0.002132	5.791853	0.0004
R-squared	0.966957	Mean dependent var		0.281318
Adjusted R-squared	0.958696	S.D. dependent var		0.307244
S.E. of regression	0.062442	Akaike info criterion		-2.482150
Sum squared resid	0.031192	Schwarz criterion		-2.373633
Log likelihood	16.65182	Durbin-Watson stat		1.367781

Estimation Command:
=====================
LS FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Estimation Equation:
=====================
FACUM=C(1)+C(2)*OBS+C(3)*(OBS)^2

Substituted Coefficients:
=====================
FACUM=0.1578508159+0.08431545455*OBS+0.0123467366*(OBS)^2


Curva de Lorenz
México 2000
1.0
0.8
0.6
0.4
0.2
0.0
FRECUENCIAS ACUMULADAS
-4
-2
0
2
4
OBSERVACIONES (RANGO DE -5 A +5)

V. CALCULO DEL COEFICIENTE DE GINI, CASO DE MÉXICO, AÑOS 1994, 1996, 1998 y 2000.

Para la aplicación de la fórmula se requiere de la información que normalmente se incluye en una tabla de distribución de frecuencias, tabla 1, 2 , 3 y 4. Apliquemos la siguiente fórmula

$$G = 1 - \sum_{i=1}^{m} p_i \ (Q_i + Q_{i-1})$$

construyendo la tabla de distribución de frecuencias para calcular el índice de Gini y la construcción de la curva de Lorenz, con los datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH), tercer trimestre 1994, 1996, 1998 y 2000, México INEGI, en el que la variable (P_i) es el total de hogares ordenados por deciles o las frecuencias acumuladas de las observaciones y, Q_i es el ingreso corriente total trimestral o las frecuencias acumuladas de participación en la variable.
Lo que se pretende analizar es la concentración del ingreso y el grado de desigualdad, para los años 1994, 1996, 1998 y 2000, el caso de México.

Tabla 1
Frecuencias de la ENIGH, México 1994.

Deciles de hogares	p_i Frecuencias relativas de las observaciones	q_i Frecuencias relativas de la variable	P_i Frecuencias acumuladas de las observaciones	Q_i Frecuencias acumuladas de la variable	$Q_i + Q_{i-1}$	$p_i\ (Q_i + Q_{i-1})$	$P_i - Q_i$ Distancias verticales entre la línea de equidistribución y la curva de Lorenz	$p_i\ (P_i - Q_i)$ Área de concentración
I	0.1	0.0159	0.1	0.0159	0.0159	0.00159	0.0841	0.00841
II	0.1	0.0276	0.2	0.0435	0.0594	0.00594	0.1565	0.01565
III	0.1	0.0367	0.3	0.0802	0.1237	0.01237	0.2198	0.02198
IV	0.1	0.0464	0.4	0.1266	0.2068	0.02068	0.2734	0.02734
V	0.1	0.0567	0.5	0.1833	0.3099	0.03099	0.3167	0.03167
VI	0.1	0.0706	0.6	0.2539	0.4372	0.04372	0.3461	0.03461
VII	0.1	0.0874	0.7	0.3413	0.5952	0.05952	0.3587	0.03587
VIII	0.1	0.1134	0.8	0.4547	0.7960	0.07960	0.3453	0.03453
IX	0.1	0.1611	0.9	0.6158	1.0705	0.10705	0.2842	0.02842
X	0.1	0.3842	1.0	1.0000	1.6158	0.16158	0.0000	0.00000
Total						**0.52304**	2.3848	**0.23848**

Aplicando la fórmula anterior el coeficiente de Gini es:

$$\mathbf{G = 1 - 0.52304 = 0.47696 = 0.\ 4770}$$

En forma tabular **el área de concentración** total de la columna 9 $[p_i\ (P_i - Q_i)]$ es igual a **0.23848** resultado igual seria si se resolviera la siguiente integral $\int_0^1 (P_i - Q_i)\, d\, P_i$

Comparando el valor del área de concentración hecho mediante el ajuste de la curva y el método de mínimos cuadrados (0.23720684), con el del método de datos tabulares tabla 1 sumatoria columna 9 (0.23848), la diferencia es muy pequeña, es de 127 cienmilésimos.

El área de máxima desigualdad esta formada por el triángulo que se encuentra debajo de la curva de distribución equitativa, es decir, del triángulo **0PQ**, cuyo resultado se obtiene mediante la solución de la integral

$$\int_0^1 (P_i - Q_i)\, d\, P_i = \int_0^1 P_i\, d\, P_i$$

$$\int_0^1 P_i\, d\, P_i = \frac{P_i^2}{2}\Big]_0^1 = \frac{1^2}{2} - \frac{0^2}{2} = \frac{1}{2} - 0 = 0.5$$

El índice de Gini es $\mathbf{G} = \dfrac{\textbf{Área de concentración}}{\textbf{Área de máxima desigualdad}}$, sustituyendo seria $\mathbf{G} = \dfrac{\mathbf{0.23848}}{\mathbf{0.5}} =$

$= \mathbf{0.47696}$

Ahora si comparamos el valor del índice de Gini obtenido por el ajuste de curvas y el método de mínimos cuadrados (0.4744136) con los datos tabulares (0.47696), la diferencia es de 254 cienmilésimos, insignificante la comparación.
Por cualquiera de los dos procedimientos se llega al mismo resultado.

Tabla 2
Frecuencias de la ENIGH, México 1996.

Deciles de hogares	p_i Frecuencias relativas de las observaciones	q_i Frecuencias relativas de la variable	P_i Frecuencias acumuladas de las observaciones	Q_i Frecuencias acumuladas de la variable	$Q_i + Q_{i-1}$	$p_i\ (Q_i + Q_{i-1})$	$P_i - Q_i$ Distancias verticales entre la línea de equidistribución y la curva de Lorenz	$p_i\ (P_i - Q_i)$ Área de concentración
I	0.1	0.0179	0.1	0.0179	0.0179	0.00179	0.0821	0.00821
II	0.1	0.0300	0.2	0.0479	0.0658	0.00658	0.1521	0.01521
III	0.1	0.0394	0.3	0.0873	0.1352	0.01352	0.2127	0.02127
IV	0.1	0.0490	0.4	0.1363	0.2236	0.02236	0.2637	0.02637
V	0.1	0.0597	0.5	0.1960	0.3323	0.03323	0.3040	0.03040
VI	0.1	0.0732	0.6	0.2692	0.4652	0.04652	0.3308	0.03308
VII	0.1	0.0896	0.7	0.3588	0.6280	0.06280	0.3412	0.03412
VIII	0.1	0.1149	0.8	0.4737	0.8325	0.08325	0.3263	0.03263
IX	0.1	0.1603	0.9	0.6340	1.1077	0.11077	0.2660	0.02660
X	0.1	0.3660	1.0	1.0000	1.6340	0.16340	0.0000	0.00000
Total						0.54422	2.2789	.22789

Aplicando el mismo procedimiento anterior, para el año 1996 Gini será igual a:

$$\mathbf{G = 1 - 0.54422 = 0.45578 = 0.4558}$$

Aplicando el procedimiento tabular y la forma mediante,

El índice de Gini es, $\mathbf{G} = \dfrac{\textbf{Área de concentración}}{\textbf{Área de máxima desigualdad}}$, sustituyendo seria $\mathbf{G} = \dfrac{\mathbf{0.22789}}{\mathbf{0.5}} =$

$= \mathbf{0.45578}$

Por cualquiera de los procedimientos se llega a un resultado igual.

Tabla 3
Frecuencias de la ENIGH, México 1998.

Deciles de hogares	p_i Frecuencias relativas de las observaciones	q_i Frecuencias relativas de la variable	P_i Frecuencias acumuladas de las observaciones	Q_i Frecuencias acumuladas de la variable	$Q_i + Q_{i-1}$	$p_i\ (Q_i + Q_{i-1})$	$P_i - Q_i$ Distancias verticales entre la línea de equidistribución y la curva de Lorenz	$p_i\ (P_i - Q_i)$ Área de concentración
I	0.1	0.0150	0.1	0.0150	0.0150	0.00150	0.0850	0.00850
II	0.1	0.0266	0.2	0.0416	0.0566	0.00566	0.1584	0.01584
III	0.1	0.0363	0.3	0.0779	0.1195	0.01195	0.2221	0.02221
IV	0.1	0.0468	0.4	0.1247	0.2026	0.02026	0.2753	0.02753
V	0.1	0.0582	0.5	0.1829	0.3076	0.03076	0.3171	0.03171
VI	0.1	0.0721	0.6	0.2550	0.4379	0.04379	0.3450	0.03450
VII	0.1	0.0893	0.7	0.3443	0.5993	0.05993	0.3557	0.03557
VIII	0.1	0.1149	0.8	0.4592	0.8035	0.08035	0.3408	0.03408
IX	0.1	0.1597	0.9	0.6189	1.0781	0.10781	0.2811	0.02811
X	0.1	0.3811	1.0	1.0000	1.6189	0.16189	0.0000	0.00000
Total						**0.52394**	**2.3805**	**0.23805**

Aplicando el mismo procedimiento anterior para el año 1998 Gini será igual a:

$$\mathbf{G = 1 - 0.52394 = 0.47606 = 0.4761}$$

El índice de Gini es $\mathbf{G} = \dfrac{\textbf{Área de concentración}}{\textbf{Área de máxima desigualdad}}$, sustituyendo seria $\mathbf{G} = \dfrac{\mathbf{0.23805}}{\mathbf{0.5}} =$

= 0.4761

Tabla 4
Frecuencias de la ENIGH, México 2000.

Deciles de hogares	p_i Frecuencias relativas de las observaciones	q_i Frecuencias relativas de la variable	P_i Frecuencias acumuladas de las observaciones	Q_i Frecuencias acumuladas de la variable	$Q_i + Q_{i-1}$	$p_i\ (Q_i + Q_{i-1})$	$P_i - Q_i$ Distancias verticales entre la línea de equidistribución y la curva de Lorenz	$p_i\ (P_i - Q_i)$ Área de concentración
I	0.1	0.0152	0.1	0.0152	0.0152	0.00152	0.0848	0.00848
II	0.1	0.0264	0.2	0.0416	0.0568	0.00568	0.1584	0.01584
III	0.1	0.0360	0.3	0.0776	0.1192	0.01192	0.2224	0.02224
IV	0.1	0.0459	0.4	0.1235	0.2011	0.02011	0.2765	0.02765
V	0.1	0.0570	0.5	0.1805	0.3040	0.03040	0.3195	0.03195
VI	0.1	0.0708	0.6	0.2513	0.4318	0.04318	0.3487	0.03487
VII	0.1	0.0884	0.7	0.3397	0.5910	0.05910	0.3603	0.03603
VIII	0.1	0.1124	0.8	0.4521	0.7918	0.07918	0.3479	0.03479
IX	0.1	0.1609	0.9	0.6130	1.0651	0.10651	0.2870	0.02870
X	0.1	0.3870	1.0	1.0000	1.6130	0.16130	0.0000	0.00000
Total						**0.51890**	**2.4055**	**0.24055**

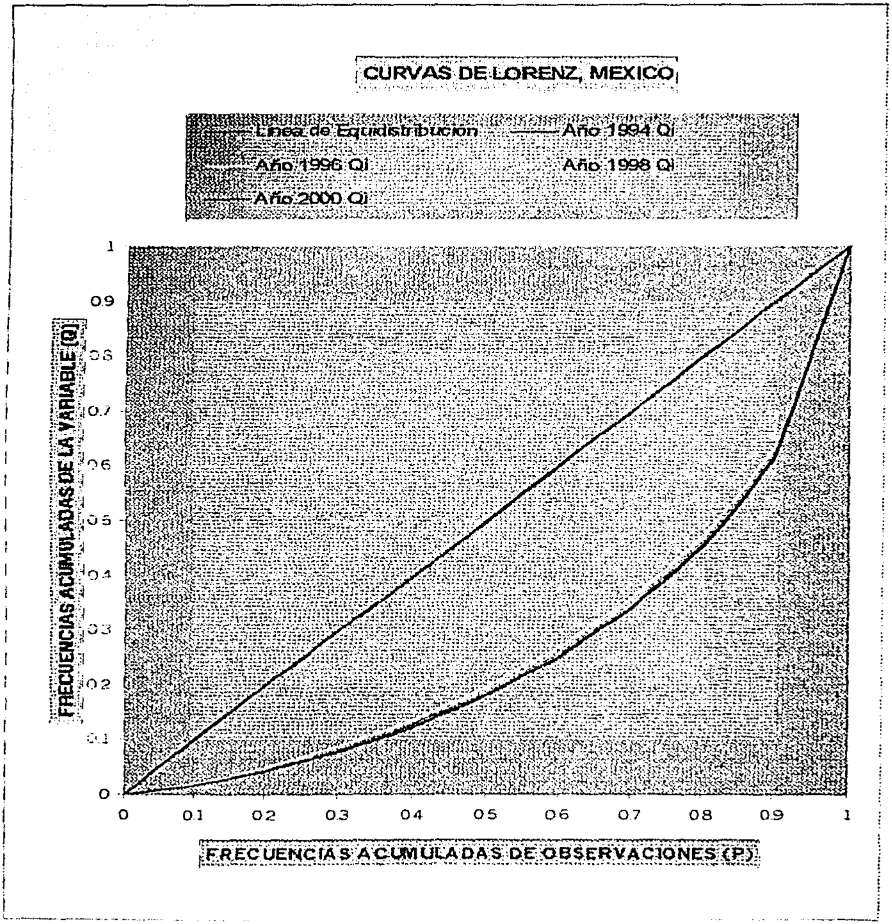
Aplicando el mismo procedimiento anterior para el año 2000 Gini será igual a:

$$\mathbf{G = 1 - 0.5189 = 0.4811}$$

El índice de Gini es $\mathbf{G} = \dfrac{\textbf{Área de concentración}}{\textbf{Área de máxima desigualdad}}$, sustituyendo sería $\mathbf{G} = \dfrac{\mathbf{0.24055}}{\mathbf{0.5}} =$

$\mathbf{= 0.4811}$

Por cualquiera de los procedimientos se llega al mismo resultado.
La siguiente gráfica muestra la curva de Lorenz para los cuatro años, 1994, 1996, 1998 y 2000.


CURVAS DE LORENZ, MEXICO
Linea de Equidistribucion
Año 1994 Qi
Año 1996 Qi
Año 1998 Qi
Año 2000 Qi
FRECUENCIAS ACUMULADAS DE LA VARIABLE (Q)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
FRECUENCIAS ACUMULADAS DE OBSERVACIONES (P)

VI. INTERPRETACIÓN DEL INDICE DE GINI Y DE LA CURVA DE LORENZ

Analizando las curvas de Lorenz para los años señalados en la gráfica anterior, observamos que para el año de 1996 las distancias verticales **(P_i - Q_i)** son menores con respecto al año 2000, 1998 y 1994, es decir, la desigualdad y la concentración fue menor que para el año 2000, 1998 y 1994, diferencia mínima pero se observa en la gráfica, y para estos tres años las distancias verticales **(P_i - Q_i)** fueron mayores, lo que implica por un lado mayor desigualdad y por otro mayor concentración del ingreso.
También hay que hacer notar, que entre más se aleje la curva de Lorenz de la línea de igualdad perfecta, reflejará mayor desigualdad y mayor concentración del ingreso, si esto se da en sentido contrario, habrá menor desigualdad y, menor concentración del ingreso, como lo ilustra la curva de Lorenz en color amarillo para el año 1996.
Desde el punto de vista numérico, el coeficiente Gini tenderá a cero y, serán mejores en términos de igualdad las condiciones de la distribución del ingreso, se aproxima a una realidad hipotética, es decir, de "igualdad perfecta"; tenderá a 1, su limite superior a medida que la distribución se hace más concentrada, esto es, mayor será el grado de concentración y, por tanto, de desigualdad. El coeficiente de Gini para los años 1994, 1996, 1998 y 2000 son respectivamente G = 0.4770, G = 0.4558, G = 0.4761, G = 0.4811 y de acuerdo a la interpretación anterior Gini para el año 1996 mejora en términos de igualdad y de la distribución del ingreso que para los años 1994, 1998 Y 2000 que reflejan más concentración y desigualdad en la distribución del ingreso. Lo anterior se explica tomando el valor de la columna siete de cualquiera de las tablas de frecuencias de la ENIGH (1, 2 3 y 4), México 1994, 1996, 1998 o 2000, supongamos que

$$\sum_{i=1}^{m} p_i\ (Q_i + Q_{i-1})$$

sea menor o casi igual a 1 (0.9999), sustituyendo en la fórmula

$$G = 1 - \sum_{i=1}^{m} p_i\ (Q_i + Q_{i-1})$$

G = 1 – 0.9999 = 0.0001, Gini tiende a cero, por lo tanto la distribución del ingreso reflejará una realidad hipotética de "igualdad perfecta", si el valor de la columna siete es mayor o casi igual a 0 (0.0001), el coeficiente Gini tenderá a 1, es decir, **G = 1 – 0.0001 = 0.9999** reflejará mayor concentración del ingreso y mayor desigualdad.

También es importante considerar los datos de la tabla 6, ayudan a fundamentar lo dicho anteriormente ya que el porcentaje de participación del 20% más rico es menor para el año 1996 (52.6) que para los años 1994 (54.5), 1998 (54.1) y 2000 (54.8), es decir, la concentración del ingreso para el año 1996 es menor y la desigualdad es menor, ya que la participación del 80% más pobre es de 47.4 para el año 1996, es mayor su participación con respecto a los otros años. [3]
Para las economías capitalistas la medida de distribución del ingreso que más frecuentemente se usa es el coeficiente Gini, la base teórica en la que sustenta es en la comparación entre lo que seria una realidad hipotética extrema de "igualdad perfecta", en la que todos los individuos tuvieran el mismo ingreso, y la realidad concreta y sensible de la sociedad, para un periodo dado.
Por su generalidad, medir la desigualdad y de la participación del ingreso con este coeficiente no da cuenta de los rasgos determinantes de la desigualdad, esto es, un mismo coeficiente puede corresponder a características marcadamente diferentes, es decir, no se puede saber si los pobres se están haciendo más pobres o los ricos se hacen más ricos, lo que se puede deducir es que las proporciones detalladas de ingreso indican si los deciles correspondientes de la distribución han ganado o perdido en relación con otros deciles, por ejemplo, la proporción recibida por el decil más pobre puede caer de 8 a 6 %, pero el nivel absoluto de ingreso de esas familias puede al mismo tiempo haberse duplicado.

La medición de la desigualdad y de la distribución funcional del ingreso se complementan, mediante la medición estadística de la distribución del ingreso por niveles, enriqueciendo gradualmente las posibilidades del análisis, esto es, el ordenamiento al que se hace alusión se expresa en términos de deciles de población, hogares o familias, así como una variedad de subdivisiones, como el 10% más pobre, el 80% más pobre, el 20% más rico, esta forma de agrupar facilita una caracterización más precisa de la distribución y su cotejo con otras situaciones, o con otra finalidad, relacionando las características de la distribución del ingreso con las estructuras productivas.

3 Paschoal Rossetti, considera que si el porcentaje de participación del 20% más rico esta por encima del 50% y el 80% más pobre este por abajo del 50% hay una desigualdad alta, habrá desigualdad moderada cuando el 20% más rico este entre un 49.1% hasta un 40% y el 80% más pobre este entre un 50.1% hasta 60% y, considera que hay desigualdad baja cuando el 20% más rico su participación este entre 39.9% y menos, el 80% más pobre debe tener una participación de un 60% y más.

Tabla 5
Estructura de la distribución del ingreso por estratos, México

Estrato de la población económicamente activa	1 9 9 4		1 9 9 6		1 9 9 8		2 0 0 0	
	Porcentaje de participación en el ingreso nacional	Valores acumulados	Porcentaje de participación en el ingreso nacional	Valores acumulados	Porcentaje de participación en el ingreso nacional	Valores acumulados	Porcentaje de participación en el ingreso nacional	Valores acumulados
10 % más pobre	1.6	1.6	1.8	1.8	1.5	1.5	1.5	1.5
10%	2.8	4.4	3.0	4.8	2.7	4.2	2.7	4.2
10%	3.7	8.1	3.9	8.7	3.6	7.8	3.6	7.8
10%	4.6	12.7	4.9	13.6	4.7	12.5	4.6	12.4
10%	5.7	18.4	6.0	19.6	5.8	18.3	5.7	18.1
10%	7.1	25.5	7.3	26.9	7.2	25.5	7.1	25.2
10%	8.7	34.2	9.0	35.9	8.9	34.4	8.8	34.0
10%	11.3	45.5	11.5	47.4	11.5	45.9	11.2	45.2
10%	16.1	61.6	16.0	63.4	16.0	61.9	16.1	61.3
10% más rico	38.4	100.0	36.6	100.0	38.1	100.0	38.7	100.0

Tabla 6
Porcentaje de participación en el ingreso nacional, México.

Por ciento de participación	1994	1996	1998	2000
20% más rico	54.5	52.6	54.1	54.8
80% más pobres	45.5	47.4	45.9	45.2

Tabla 7
Porcentaje concentrado de participación de los salarios mínimos y de los hogares, en el ingreso nacional, México

Año	Concentrado de salarios mínimos		Concentrado de hogares	
	20 % más rico	80 % más pobres	20 % más rico	80 % más pobres
1994	63.73	36.27	27.99	72.01
1996	59.22	40.78	25.52	74.48
1998	62.04	37.96	26.70	73.30
2000	67.73	32.27	31.86	68.14

Tabla 8
Porcentaje de participación de los salarios mínimos y de los hogares, en el ingreso nacional, México

Salarios mínimos	1994		1996		1998		2000	
	Porcentaje de participación en el ingreso nacional	Porcentaje de participación de los hogares	Porcentaje de participación en el ingreso nacional	Porcentaje de participación de los hogares	Porcentaje de participación en el ingreso nacional	Porcentaje de participación de los hogares	Porcentaje de participación en el ingreso nacional	Porcentaje de participación de los hogares
0.00 A 0.50	0.02	0.37	0.02	0.48	0.04	0.79	El rango va de 0.00 a 1.00	El rango va de 0.00 a 1.00
0.51 A 1.00	0.35	3.12	0.31	2.50	0.44	3.75	0.28	3.00
1.01 A 1.50	0.96	5.30	1.12	5.65	1.18	6.21	0.80	4.86
1.51 A 2.00	1.78	7.08	2.20	7.84	2.16	8.20	1.58	6.89
2.01 A 3.00	5.88	16.33	6.71	16.93	5.93	15.51	4.41	13.49
3.01 A 4.00	7.34	14.37	8.66	15.56	7.39	13.70	5.99	13.06
4.01 A 5.00	6.83	10.38	8.15	11.20	7.21	10.33	6.50	10.95
5.01 A 6.00	7.06	8.75	7.5	8.50	7.40	8.67	6.24	8.52
6.01 A 7.00	6.05	6.31	6.11	5.82	6.21	6.14	6.45	7.38
7.01 A 8.00	5.19	4.69	5.82	4.84	6.13	5.20	5.42	5.36
8.01 Y Más	58.54	23.30	53.40	20.68.	55.91	21.50	62.31	26.50

En la tabla 5 el decil I, el 10 % más pobre percibía 1.6 % de los ingresos de los hogares en 1994, aumenta su participación en 1996 a 1.8 % gano 0.2 %, pero para 1998 y 2000 bajo su participación a 1.5 %, es decir, perdió 0.1 % con respecto a 1994 y 0.3 con respecto a 1996. Para el decil X el 10 % más rico, también disminuyo su participación, del año 1994 de 38.4 % paso a 36.6 % en 1996 y para 1998 no recupero la posición que tenia en el 94, llegando a 38.1%, siendo mayor la participación en el 2000 llegando a 8.7%.

CONCLUSIONES

Si la participación del 20% más rico es menor como lo es para el año de 1996 con respecto a los años 1994, 1998 y 2000, entonces los sectores más favorecidos son los deciles de menores ingresos, es decir, la desigualdad y concentración del ingreso para 1996 es menor, se puede ver como la curva de Lorenz se acerca más a la curva de equidistribución, o las distancias verticales entre la curva de Lorenz y la línea de igualdad perfecta son menores y la gráfica comparativa de las curvas de Lorenz a simple vista lo refleja, ya que se puede ver como la curva de Lorenz para los años 1994, 1998 y 2000 es lo contrario, se alejan de la línea de igualdad perfecta, interpretándose que hay mayor concentración del ingreso y, por otro lado mayor desigualdad social, aunque la diferencia en el coeficiente de Gini no sea muy marcada el valor del índice tiende más a cero para el año 1996 y, las condiciones de la distribución reflejan menos concentración del ingreso, siendo mejores en términos de igualdad y lo corrobora la curva de Lorenz.

El coeficiente de Gini por su generalidad de medir la desigualdad, no da cuenta de los rasgos determinantes de esa desigualdad, porque un mismo coeficiente puede corresponder a características marcadamente diferentes, lo que sí puede indicar es como algunos deciles han ganado o perdido en la distribución en relación con otros deciles, es decir, no se puede saber si los ricos se están haciendo más ricos o los pobres se están haciendo más pobres.

BIBLIOGRAFÍA

1. CORTES, FERNANDO Y RUBALCABA, ROSA MARIA. TÉCNICAS ESTADÍSTICAS PARA EL ESTUDIO DE LA DESIGUALDAD SOCIAL. ED. EL COLEGIO DE MEXICO, MEXICO 1982.

2. MURRAY R. SPIEGEL. ESTADÍSTICA. 2/a EDICIÓN. ED. MC GRAW HILL. 1995.

3. PASCHOAL ROSSETTI, JOSE. INTRODUCCION A LA ECONOMIA, ED. HARLA, MEXICO 1994.

4. SAMUELSON, PAUL. ECONOMIA. ED. MC GRAW HILL

5. ENCUESTA NACIONAL DE INGRESO Y GASTO DE LOS HOGARES. INEGI, MÉXICO TERCER TRIMESTRE 1994.

6. ENCUESTA NACIONAL DE INGRESO Y GASTO DE LOS HOGARES. INEGI, MÉXICO TERCER TRIMESTRE 1996.

7. ENCUESTA NACIONAL DE INGRESO Y GASTO DE LOS HOGARES. INEGI, MÉXICO TERCER TRIMESTRE 1998.

8. ENCUESTA NACIONAL DE INGRESO Y GASTO DE LOS HOGARES. INEGI, MÉXICO TERCER TRIMESTRE 2000.