ESTUDIO TEMPORAL DE LA DISTRIBUCIÓN DE LA RIQUEZA MEDIANTE CURVAS DE LORENZ

FEDRIANI MARTEL, Eugenio M.

Departamento de Economía y Empresa
Universidad Pablo de Olavide
correo-e: efedmar@dee.upo.es

MARTÍN CARABALLO, Ana M.

Departamento de Economía y Empresa
Universidad Pablo de Olavide
correo-e: ammarcar@dee.upo.es

RESUMEN

En este trabajo se pretende utilizar conceptos básicos relacionados con el Análisis Matemático (interpolación e integración de funciones en dos variables) para realizar un estudio de la distribución de la riqueza en un periodo de tiempo en diferentes naciones. Aprovechando datos macroeconómicos como el PIB y las curvas de Lorenz de cada país en distintos instantes de tiempo, se construye un indicador que permite agrupar las naciones según la evolución temporal de la distribución de su riqueza. La sencillez de cálculo del indicador utilizando programas de computación simbólica, así como la fácil interpretación económica del mismo, parecen sugerir que su uso podría ser útil para analizar el problema de la distribución de la riqueza a lo largo del tiempo en regiones que hayan sido objeto de actuaciones similares.

Según esto, el objetivo de este trabajo es ver la evolución de las desigualdades (según la distribución de la riqueza) de varias regiones, o bien de una única población, a lo largo de un periodo de tiempo; para ello, se construirá un indicador que utilizará entre otros datos las curvas de Lorenz de las regiones del estudio en distintos instantes de tiempo.

Palabras clave: Riqueza, curvas de Lorenz, interpolación, indicadores.

Área temática: Aspectos cuantitativos del fenómeno económico.

1. INTRODUCCIÓN

La pobreza, la desigualdad y el nivel de vida tienden a ser fenómenos más estables que, por ejemplo, la renta anual de un grupo de individuos; así, una caída de la renta anual (en cierto instante del tiempo), no implica necesariamente una caída en el nivel de vida. Por ello, a menudo es útil estudiar la evolución de la renta en amplios periodos de tiempo, aunque se puede comprobar que pueden presentarse algunos problemas en la lectura de los resultados cuando solo se analiza la renta (Fedriani y Martín, 2004). Tal vez una de las formas más objetivas de solventar las deficiencias de la clasificación sea analizar conjuntamente renta y desigualdad.

Las curvas de Lorenz[1] permiten comparar distribuciones y establecer ordenaciones entre ellas. En resumen, un grupo es más desigual que otro cuando su correspondiente curva de Lorenz queda siempre por debajo para toda la proporción de individuos que componen el grupo. La ordenación obtenida mediante las curvas de Lorenz es parcial y se puede hacer siempre que las curvas no se corten; cuando se cortan es imposible decir, con un mínimo de rigor científico y a partir de las representaciones gráficas, cuál de las distribuciones que se están comparando es más desigual en la distribución de la riqueza.

Para resolver el problema de la comparabilidad de distintas curvas de Lorenz, diversos autores han construido algunos indicadores que sintetizan en un único número la información dada por ellas, para de esta forma poder dar una ordenación total de las distribuciones que se están comparando. Un ejemplo de indicador de este tipo es el índice de Gini.

En este trabajo, el objetivo final consiste en ver la evolución de la pobreza y las desigualdades de varias regiones, o bien de una única población, a lo largo de un periodo de tiempo; para ello, se construirá un indicador utilizando la metodología que a continuación se detalla.

---

[1] Para ver distintos métodos de construcción de la Curva de Lorenz de una determinada población pueden consultarse Gupta (1984), Kakwani y Podder (1973) y Ortega, Martín, Fernández, Ladoux y García (1991), entre otros.

2. CONSTRUCCIÓN DEL INDICADOR DE DISTRIBUCIÓN TEMPORAL DE LA RIQUEZA

Suponemos conocidos los datos para cada una de las unidades de análisis del estudio y para varios instantes de tiempo del periodo que se quiere analizar. De momento, fijemos una de dichas unidades de análisis; posteriormente repetiremos este proceso para las otras unidades de análisis.

En primer lugar, se calcula la curva de Lorenz asociada a los datos que se tienen para cada uno de los instantes de tiempo $t_0, t_1,\ldots, t_k$. A cada una de las citadas curvas las representamos por $L_t(x)$ para $t \in \{t_0, t_1,\ldots, t_k\}$, con $x \in [0,1]$ (por ser $L_t(x)$ una curva de Lorenz).

A continuación, elegimos $\Pi_X \equiv \{0 = x_0 < x_1 < \ldots < x_n = 1\}$, una partición del intervalo [0,1] tal que $x_i = i \cdot h$ para $i=1,\ldots, k$ y $h=1/k$ (esto es, por simplicidad y ventajas de cálculo, elegimos h constante). Normalmente Π_X viene recomendado por la forma de calcular las curvas de Lorenz.

Fijado el valor en $x_j \in \Pi_X$ para cada instante de tiempo t_i, se tienen los puntos $(L_{t_i}(x_j), t_i)$ para $i=1,\ldots, k$ y $j=1,\ldots, n$, donde $\Pi_T \equiv \{0 = t_0 < t_1 < \ldots < t_{k-1} < t_k = T\}$ es una partición del intervalo [0,T], que normalmente vendrá determinada por la disponibilidad de los datos, según se ha visto.

El siguiente paso será calcular una función de interpolación de tales puntos. Para ello, será necesario utilizar la interpolación en varias variables, concretamente la interpolación en dos variables y en una malla rectangular $[0,1]\times[0,T]$. Consecuentemente, se planteará un problema de interpolación de Lagrange en la malla rectangular $[0,1]\times[0,T]$.

A la función de interpolación obtenida utilizando la interpolación de Lagrange para los puntos $(x_j, t_i, L_{t_i}(x_j))$ la representaremos por $L(x,t)$ (que es un polinomio en las variables x y t de grado n para la variable x y de grado k para la variable t, luego es un polinomio de grado $n+k$), donde $x \in [0,1]$ y $t \in [0,T]$.

El procedimiento anterior, obviamente, deberá repetirse para cuantas unidades de análisis compongan el estudio.

A partir de las funciones $L(x,t)$, se va a definir un indicador que se denotará *IDT* y nos permitirá comparar la situación de pobreza y desigualdad en distintas zonas.

<u>Definición 2.1</u>: *En las condiciones anteriores y con la notación recién establecida,*

$$IDT = T - \int_0^1 \int_0^T L(x,t)\, dt\, dx\,.$$

El indicador *IDT* se puede interpretar como una generalización del índice de Gini; éste nos indicaba el área comprendida entre la diagonal del cuadrado $[0,1]\times[0,1]$ y la curva de Lorenz y el indicador *IDT* nos va a dar el volumen comprendido entre la "diagonal" del cubo $[0,1]\times[0,\mathrm{T}]\times[0,1]$ y la superficie formada por las curvas.

2.1. Interpretación y propiedades del indicador *IDT*

En este apartado se hará una descripción más profunda de la interpretación del indicador *IDT* definido en la apartado anterior. Además, se mostrarán algunas de las propiedades que verifica el indicador *IDT*.

La curva de Lorenz en un cierto instante de tiempo *t* para una distribución dada se define como una función

$$\begin{aligned} L_t : [0,1] &\to [0,1] \\ x &\mapsto L_t(x). \end{aligned}$$

Como se ha dicho, a partir de la curva de Lorenz se definen varios indicadores de desigualdad (veáse por ejemplo Zheng (1997)). Un ejemplo de indicador de desigualdad construido a partir de la curva de Lorenz es el índice de Gini, que calcula la desigualdad de cierta distribución en un determinado instante de tiempo. Nuestro objetivo al definir el indicador *IDT* es diferente: calculadas las curvas de Lorenz para una misma distribución en instantes de tiempo distintos, nos proponemos ver cómo han evolucionado tales curvas; es decir, ver cómo ha evolucionado la desigualdad a lo largo de un periodo de tiempo. De esta forma, haciendo esto para varias distribuciones (o zonas) podremos elaborar una clasificación de las zonas según las desigualdades en el periodo de tiempo [0,T].

De esta forma, el indicador *IDT* geométricamente es el volumen comprendido entre el plano que pasa por el punto *(1,T,1)* y contiene al eje *OY* y la superficie $L_j(x,t)$ (que se calcula a partir de las curvas de Lorenz de una cierta zona en instantes de tiempo distintos). Por tanto, cuanto menor sea tal volumen, menos desigualdades existirán en el periodo de tiempo estudiado.

Cuando el número de zonas geográficas que forman parte del estudio que se realiza es reducido, sucede que el indicador *IDT* se comporta de forma similar a la media de los índices de Gini en cada instante de tiempo para cada una de las zonas. Por tanto, la clasificación de las zonas geográficas del estudio que se obtiene utilizando el indicador *IDT* es equivalente al que se obtiene si clasificamos tales zonas geográficas según los valores de la media aritmética de los índices de Gini de cada zona para todos los instantes de tiempo estudiados. No obstante, conviene avisar de la dificultad de encontrar bases de datos con los índices de Gini de diferentes regiones. Por otra parte, la ventaja más importante del uso del indicador *IDT* frente a la media aritmética, por ejemplo, es que permite la realización de un análisis temporal comparativo incluso cuando no se tengan datos simultáneos en todas las unidades de análisis. Esto es, cada unidad de análisis (dentro del periodo temporal que se tomará en consideración para el estudio) puede tener distinto número de datos, éstos pueden estar desigualmente distribuidos (diferente espaciamiento o un espaciamiento no uniforme), pueden incluirse nuevos datos conforme se vayan obteniendo, etc.; es decir, el proceso explicado es independiente de la irregularidad temporal de los datos con los que se cuenta.

Es necesario comentar también que la principal desventaja del indicador *IDT* es su coste computacional, ya que para su cálculo es necesario calcular las curvas de Lorenz para varios instantes de tiempo, además de calcular una función de interpolación polinómica en dos variables.

Proposición 2.1: *El indicador IDT es invariante ante cambios de escala.*

Demostración:

Las curvas de Lorenz son invariantes ante cambios de escala (Lorenz, 1905), por lo que los puntos $(x_j, t_i, L_{ti}(x_j))$ se mantienen constantes ante cambios de escala en los valores de la función $L(x,t)$. Por tanto, la función de interpolación calculada también se mantiene constante y consiguientemente el indicador *IDT* en cualquier unidad de análisis. □

Proposición 2.2: *El indicador IDT es independiente del tamaño de la población en la que se realiza el estudio.*

Demostración:

La curva de Lorenz en cada instante de tiempo t, $L_t(x)$, es independiente del tamaño de la población, por lo que la función de interpolación construida para el cálculo del indicador *IDT* en los puntos de la forma $(x_j, t_i, L_{ti}(x_j))$ para $i=1,\ldots,k$ y $j=1,\ldots,n$ es independiente del número de individuos de cada población. Por tanto, también se verifica que el indicador *IDT* es independiente del tamaño de la población en la que se realiza el estudio. □

De un modo similar, el indicador *IDT* permanece invariante si el número de individuos en cada nivel de la variable estudiada se ve alterado en la misma proporción en cada uno de los percentiles que se han utilizado en el cálculo de las curvas de Lorenz.

Proposición 2.3: *El indicador IDT verifica el Axioma de Transferencia Débil.*

Demostración:

Intuitivamente, a partir de la descripción geométrica de la curva de Lorenz, se puede observar que una transferencia de una unidad de análisis rica a una más pobre en un cierto instante de tiempo t eleva toda la curva de Lorenz entre los correspondientes percentiles; como consecuencia también se elevará la curva $L(x,t)$ y, por tanto, el valor del indicador *IDT* se verá reducido. □

Proposición 2.4: *El indicador IDT es continuo respecto de las variables que intervienen en su cálculo.*

Demostración:

Se deduce del proceso de selección del polinomio de interpolación (Hammerlin, 1991) y de propiedades ya utilizadas en la demostración de proposiciones similares a ésta. □

2.2. Un ejemplo

A continuación, se presenta un ejemplo para una mayor claridad y comprensión del método descrito anteriormente para el cálculo del indicador *IDT*. En el ejemplo se dispone de los datos del producto interior bruto (PIB) para los años 1991, 1996 y 2000 para cuarenta y una comarcas de la Comunidad Autónoma de Cataluña recogidos de los Anuarios Económico Comarcales publicados por Caixa Catalunya (Caixa Catalunya (1991), Caixa Catalunya (1996), Caixa Catalunya (2000)). Estas comarcas harán de individuos en nuestro análisis.

El PIB en cada uno de los años anteriormente indicados está dado en millones de pesetas corrientes. En primer lugar, para calcular las curvas de Lorenz, los datos se agrupan en nueve intervalos de amplitud constante 0,3, obteniéndose lo siguiente:

PIB (millones de pesetas corrientes)	**Número de Comarcas**		
	1991	**1996**	**2000**
1-1,3	6	0	0
1,3-1,6	16	5	4
1,6-1,9	14	14	13
1,9-2,2	2	15	14
2,2-2,5	3	2	4
2,5-2,8	0	4	5
2,8-3,1	0	0	0
3,1-3,4	0	0	0
3,4-3,7	0	1	1

Fuente: Anuario económico Comarcal. Caixa Catalunya.

A partir de tales datos se calculan las Curvas de Lorenz para cada uno de los años 1991, 1996 y 2000 utilizando la hoja de cálculo EXCEL[2] para realizar los cálculos previos. Realizados tales cálculos, y a partir de los datos obtenidos y utilizando el programa de cálculo simbólico MATHEMATICA[3], calculamos las curvas de Lorenz mediante interpolación de funciones. Las órdenes necesarias para calcular lo anterior en MATHEMATICA se describen a continuación:

```
<<NumericalMath`SplineFit`
<<Graphics`Spline`
```

[2] EXCEL© Copyright Microsoft Corporation 1985-2004.

[3] MATHEMATICA© Copyright 1988-2004 Wolfram Research, Inc.

Estas dos primeras sirven para cargar los paquetes específicos; además, para cada año se tiene algo del tipo:

```
datosaño91={{14.63,10.49},{53.66,45.78},{87.80,83.04},
      {92.68,89.28},{100,100}};
curva91=SplineFit[datosaño91,Cubic];
```

Si se quiere dibujar la curva de Lorenz calculada, utilizaremos la orden:

```
ParametricPlot[curva91[u],{u,0,4},PlotRange→All,Compiled→False]
;

datosaño96={{0,0},{12.2,8.91},{40.34,39.03},{84.93,76.83},
      {87.80,82.61},{97.56,95.64},{100,100}};
curva96=SplineFit[datosaño96,Cubic];
ParametricPlot[curva96[u],{u,0,4},PlotRange→All,Compiled→False]
;

datosaño00={{0,0},{9.76,6.95},{41.46,34.21},{75.61,68.60},
      {85.37,79.87},{97.56,95.75},{100,100}};
curva00=SplineFit[datosaño00,Cubic];
ParametricPlot[curva00[u],{u,0,4},PlotRange→All,Compiled→False]
;
```

A continuación, se calcula el polinomio de interpolación de Lagrange utilizando las siguientes órdenes:

En primer lugar se definen los intervalos; en el ejemplo son [0,1] y [0,2]:

```
a=0
b=1
c=0
d=2
```

A continuación se indica el número de nodos para cada uno de los intervalos anteriores; en el ejemplo son 3 y 2, respectivamente:

```
n=3
m=2
hx=(b-a)/n
ht=(d-c)/m
```

Se calculan los nodos para cada uno de los intervalos con las siguientes órdenes:

```
nodosx=Table[a+i*hx,{i,0,3}]
nodost=Table[c+j*ht,{j,0,2}]
nodos=Table[{nodosx[[i]],nodost[[j]]},{i,1,n+1},{j,1,m+1}]
```

Los polinomios de Lagrange para las variables *x* y *t* se calculan con la orden:

```
For[i=1,i≤n+1,i++,For[j=1,j≤m+1,j++,l[i,j]=
      Product[(x-nodosx[[r]])/(nodosx[[i]]-nodosx[[r]]),
      {r,1,i-1}]*Product[(x-nodosx[[r]])/(nodosx[[i]]-
      nodosx[[r]]),{r,i+1,n+1}]*Product[(t-nodost[[s]])/
      (nodost[[j]]-nodost[[s]]),{s,1,j-1}]*Product[(t-
      nodost[[s]])/(nodost[[j]]-nodost[[s]]),{s,j+1,m+1}]]]
```

A continuación se calcula el polinomio de interpolación de Lagrange:

```
curvas={curva91,curva96,curva00}
valores=Table[Part[curvas[[k]][nodosx[[i]]],2],{k,1,3},
      {i,1,n+1}]
traspvalores=Transpose[valores]
L[x_,t_]:=Sum[traspvalores[[i,k]]*l[[i,k]],{i,1,n+1},{k,1,3}]
polilagr=Table[l[i,k],{i,1,n+1},{k,1,3}]
```

Una vez calculado el polinomio de interpolación de Lagrange se procede a calcular el valor del indicador *IDT*:

$$IDT = T - \int_0^1 \int_0^2 L(x,t)\, dt\, dx = 12'3816\,.$$

El ejemplo que se ha descrito, como se indicó anteriormente, es solamente descriptivo del método que se ha propuesto en este trabajo para el estudio de la evolución de la renta y desigualdades en un cierto periodo de tiempo. Así, si se realizan estos mismos cálculos para varias regiones o zonas geográficas, podemos comparar la evolución de la renta y las desigualdades en el periodo de tiempo en el que se realiza el estudio para las distintas zonas.

3. REFERENCIAS BIBLIOGRÁFICAS

- CAIXA CATALUNYA (1991). "Anuario económico comarcal". Caixa Catalunya.
- CAIXA CATALUNYA (1996). "Anuario económico comarcal". Caixa Catalunya.
- CAIXA CATALUNYA (2000). "Anuario económico comarcal". Caixa Catalunya.
- FEDRIANI, E.M. y MARTÍN, A.M. (2004) "Estudio comparativo de la riqueza internacional mediante normas funcionales". XVIII Reunión anual ASEPELT-España.
- GUPTA, M.R. (1984). "Funcional form for estimating the Lorenz curve". Econometrica, 52, pp. 1313-1314.
- HAMMERLIN, G. (1986). "Numerical Mathematics". Springer-Verlag.
- KAKWANI, N.C. y PODDER, N. (1973). "On the estimation of Lorenz curves from grouped observations". International Economic Review, 14, 2, pp. 278-291.
- LORENZ, M.C. (1905). "Methods of measuring the concentration of wealth". Publications of the American Statistical Association, 9, pp. 209-219.
- ORTEGA, P.; MARTÍN, G.; FERNÁNDEZ, A.; LADOUX, M. y GARCÍA, A. (1991). "A new functional form for estimating the Lorenz curve". The review of Income and Wealth, 27, 4, pp. 447-452.
- ZHENG, B. (1997). "Agregate poverty measures". Journal of Economic Surveys, 11, 2, pp. 123-162.