

UN ANÁLISIS SOBRE ALGUNAS MEDIDAS DE DISTANCIA ENTRE DISTRIBUCIONES DE RENTA

MATILDE LAFUENTE LECHUGA

Dpto. Métodos Cuantitativos para la Economía

Universidad de Murcia

I. INTRODUCCIÓN.

En la teoría de la medición de la desigualdad es de gran interés el poder comparar dos distribuciones de renta y medir el grado de desigualdad de renta entre ellas. Como es bien conocido, el método más simple para comparar dos distribuciones es clasificarlas por sus curvas de Lorenz a través del criterio de dominancia de Lorenz. Pero la ordenación generada por la comparación de dichas curvas es solamente parcial, ya que si dos curvas se intersecan este criterio de dominancia no será válido, en cuyo caso puede usarse alguna de las muchas medidas de desigualdad conocidas y clasificar las distribuciones de acuerdo con ellas. Ambos métodos calculan la desigualdad dentro de las respectivas poblaciones y las comparan.

En este trabajo estamos interesados en evaluar la desigualdad de renta entre dos poblaciones dadas; es decir, medir la distancia que pueda existir entre dos distribuciones de renta. Para este fin han sido propuestas en diferentes estudios [Dagum (1980), Shorrocks (1982), Ebert (1984), Chakravarty y Dutta (1987), Chakravarty (1990)] distintas medidas de distancia entre dos distribuciones de renta.

El problema de determinar la distancia entre distribuciones de renta no es nuevo. Las medidas de desigualdad absoluta son un caso especial de medidas de distancia; es cuando una distribución de renta se compara con la distribución igualitaria de la misma renta total.

Comenzamos introduciendo el concepto de función de distancia ética, para analizar, posteriormente, las distintas medidas de distancia aludidas y proponer una nueva la cual será estudiada detenidamente.

Para poder entender algunas de las medidas de distancia propuestas en los trabajos mencionados anteriormente, necesitamos del concepto de renta equivalente introducido por Atkinson (1970), Kolm (1976a,b) y Sen (1973), x_e , que es el nivel de renta per cápita que si se asignara a todos los individuos de la población permitiría alcanzar el nivel de bienestar de la distribución actual. Esta renta equivalente, x_e , vendrá dada implícitamente por

$$W^n(x_e, I^n) = W^n(x)$$

donde $\mathbf{1}^n$ es el n -vector de unos; $\mathbf{x} \in \mathbb{R}_+^n$ es una distribución de renta de tamaño n y $W^n: \mathbb{R}_+^n \rightarrow \mathbb{R}^1$ es una función de bienestar social (FBS) que satisface las siguientes propiedades:

- i) Para todo $n \in \mathbb{N}$, W^n es continua.
- ii) Para todo $n \in \mathbb{N}$, W^n es creciente a lo largo de la línea de igualdad.
- iii) Cada superficie de nivel de W^n , $n \in \mathbb{N}$, cruza la línea de igualdad.

es decir, que la FBS es regular (El concepto de FBS regular puede verse en, por ejemplo, Chakravarty, 1990, página 44).

Al ser W^n regular, \mathbf{x}_e se puede expresar explícitamente como $\mathbf{x}_e = \beta^n(\mathbf{x})$. Esta función, β^n es continua y es una cardinalización específica de W^n , es decir

$$W^n(x) \geq W^n(y) \iff \beta^n(x) \geq \beta^n(y) \iff x_e \geq y_e$$

Así, un perfil es preferido a otro, con el mismo tamaño de población, si y solamente si, su renta equivalente es mayor.

II. FUNCIÓN DE DISTANCIA

El conjunto de las distribuciones de renta sobre una población de tamaño n , viene dado por \mathbb{R}_+^n . Consecuentemente, el conjunto de todas las posibles distribuciones de renta es

$$R_+ = \bigcup_{n \in \mathbb{N}} \mathbb{R}_+^n$$

Consideramos cualquier función escalar no negativa d definida sobre $\mathbb{R}_+ \times \mathbb{R}_+$, $d: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+^1$. Definimos

$$\forall m, n \in \mathbb{N} \quad d_{m,n}: \mathbb{R}_+^m \times \mathbb{R}_+^n \rightarrow \mathbb{R}_+^1$$

Así, la función $d_{m,n}$ es la restricción de d sobre el dominio $\mathbb{R}_+^m \times \mathbb{R}_+^n$.

Definición 1:

La función $d: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+^1$ es una **función de distancia ética** si y solamente si " $m, n \in \mathbb{N}$ y

para cualquier $\mathbf{x} \in R_+^m, \mathbf{y} \in R_+^n$, se satisfacen las siguientes propiedades:

- i) $d_{m,n}$ es continua.
- ii) Reflexiva: $d_{m,n}(\mathbf{x}, \mathbf{y}) = 0$ si y solamente si las rentas equivalentes de \mathbf{x} e \mathbf{y} son iguales para cualquier FBS regular.
- iii) Simétrica: $d_{m,n}(\mathbf{x}, \mathbf{y}) = d_{n,m}(\mathbf{y}, \mathbf{x})$
- iv) Desigualdad triangular:

$$d_{m,n}(x, y) \leq d_{m,k}(x, z) + d_{k,n}(z, y) \quad \forall z \in R_+^k, k \in N$$

- v) Principio de población:

$$d_{mk, nl}(x^{(k)}, y^{(l)}) = d_{m,n}(x, y), \quad \forall k, l \in N$$

$$\text{donde } x^{(k)} = (\overbrace{x_1, \dots, x_1}^k, \dots, \overbrace{x_m, \dots, x_m}^k)$$

Las propiedades ii) a iv) son las usuales de una medida de distancia; la v), que es paralela al principio de población de Dalton, nos dice que la función de distancia es independiente del tamaño de las poblaciones.

III. DISTINTAS MEDIDAS DE DISTANCIA

Dagum (1980) propuso dos índices a los que llama razones de distancia económica que reflejan el grado de riqueza o bienestar de una población relativa a otra.

El primer índice de Dagum, d^0 se define como sigue. Sean dos distribuciones de renta $\mathbf{x} \in R^m$ e $\mathbf{y} \in R^n$ y suponemos que \mathbf{x} tiene una media más alta. Dagum selecciona \mathbf{x} como una clase de distribuciones de referencia y examina las rentas en la otra distribución. Si \mathbf{x} e \mathbf{y} contienen valores no comunes, entonces d^0 se define como:

$$d_{m,n}^0(x, y) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \rho(x_i - y_j)$$

donde $\rho(\cdot)$ es una función indicador con la propiedad

$$p(z) = \begin{cases} 1 & \text{si } z > 0 \\ 0 & \text{si } z < 0 \end{cases}$$

d^0 tendrá, en general, una discontinuidad de salto correspondiente a cualquier par de distribuciones que, o tienen una media común, o contienen al menos un valor de renta común.

Shorrocks (1982) demostró que d^0 viola las propiedades reflexiva y desigualdad triangular que toda función de distancia debe de verificar.

Como una alternativa a d^0 , Dagum propuso un segundo índice, que en términos de nuestra notación anterior viene dado por

$$\begin{aligned} d_{m,n}^1(x, y) &= \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (x_i - y_j) r(x_i - y_j) \\ &= \frac{1}{2}(\mathbf{m}_x - \mathbf{m}_y) + \frac{1}{2mn} \sum_{i=1}^m \sum_{j=1}^n |x_i - x_j| \end{aligned}$$

Shorrocks observó que como medida de distancia económica d^1 es una mejora de d^0 . Satisface la propiedad de desigualdad triangular pero la reflexiva sigue sin cumplirse puesto que

$$d_{m,m}^1(x, x) = \frac{1}{2m^2} \sum_{i=1}^m \sum_{j=1}^m |x_i - x_j| = \mathbf{m}_x G(x)$$

donde $G(\mathbf{x})$ es el valor del índice de Gini para la distribución \mathbf{x} , que será positivo a menos que $x_i = \mu_x$, para todo i .

Así, d^0 y d^1 tienen un gran número de inconvenientes para ser tratadas como medidas de distancia económicas.

Shorrocks argumentó que para reflejar el grado de riqueza de una población relativa

a otra se deberían comparar los niveles medios de bienestar de las dos poblaciones usando alguna FBS. Propuso como distancia económica entre dos distribuciones el valor absoluto de la diferencia entre las rentas equivalentes de cada distribución:

$$d_{m,m}(x, y) = |\mathbf{b}(x) - \mathbf{b}(y)| \quad \forall x, y \in R_+^m$$

donde, β es la función de renta equivalente. Esta medida de distancia está bien definida y es continua para todo par de distribuciones \mathbf{x} e \mathbf{y} bajo las condiciones usuales impuestas sobre la FBS. Es también no negativa y satisface las propiedades que deben acompañar a toda función de distancia.

Ebert (1984), siguiendo la propuesta de Dagum, presenta una clase de índices estadísticos de distancia entre dos distribuciones de renta. Estos índices se basan en funciones de distribución y la clase se obtiene por una aproximación axiomática que está completamente caracterizada por las propiedades asociadas normalmente con una función de distancia y por un conjunto de axiomas adicionales que exponemos a continuación. (Ebert realiza una sustitución en la propiedad reflexiva señalada en la definición 1, por una más fuerte: dos distribuciones son idénticas si y solo si las inversas de sus funciones de distribución coinciden).

Axioma 1: Homogeneidad lineal

$$\forall m, n \in N, \forall x \in R_+^m, \forall y \in R_+^n, \text{ y para todo escalar } l > 0$$

$$d_{m,n}(lx, ly) = l d_{m,n}(x, y)$$

Axioma 2: Invarianza por traslaciones

$$\forall m, n \in N, \forall x \in R_+^m, \forall y \in R_+^n$$

$$d_{m,n}(x + \mathbf{a} l^m, y + \mathbf{a} l^n) = d_{m,n}(x, y)$$

donde \mathbf{a} es cualquier escalar tal que $x + \mathbf{a} l^m \in R_+^m, y + \mathbf{a} l^n \in R_+^n$

Estos axiomas implican que las medidas de distancia son medidas absolutas. El primero de

ellos nos indica que $d_{m,n}$ varía en la misma proporción en que lo hace \mathbf{x} e \mathbf{y} ; el segundo afirma que algunos tipos de modificaciones simultáneas de \mathbf{x} e \mathbf{y} no afectan a la medida de distancia.

De este modo la clase de medidas de distancia propuesta por Ebert, tiene la forma:

$$d_{m,n}^r(x, y) = \left[\int_0^1 / F_x^{-1}(v) - F_y^{-1}(v) / dv \right]^{1/r}, \quad r \geq 1$$

siendo las distribuciones de renta \mathbf{x} e \mathbf{y} variables aleatorias no negativas cuyas funciones de distribución vienen representadas por F_x y F_y , respectivamente.

En el caso discreto, y denotando por \mathbf{x}' e \mathbf{y}' cualquier permutación de las variables \mathbf{x} e \mathbf{y} , de forma que se verifique que $x'_i \leq x_{i+1}'$ e $y'_i \leq y_{i+1}'$, la clase de índices propuesto por Ebert, para $m=n$, es

$$d_{m,m}^r(x, y) = \left[\sum_{i=1}^m \frac{1}{m} |x_{i'} - y_{i'}| \right]^{1/r}$$

La propiedad reflexiva de una función de distancia para esta clase de índices nos diría que la función de distancia toma el valor cero si y solamente si las funciones de distribución inversas coinciden. Sin embargo, las rentas equivalentes pueden ser iguales incluso si las funciones de distribución inversas no coinciden. Pero como se está interesado en medir la distancia entre el bienestar medio de dos poblaciones, la propiedad reflexiva parece ser una condición totalmente razonable.

$d_{m,n}^r(\mathbf{x}, \mathbf{y})$ es la r -ésima media de las diferencias absolutas de las funciones de distribución inversas. Estas medidas poseen las siguientes propiedades adicionales:

- i) $d_{m,m}^r(x + \mathbf{a} \cdot I_m, x) = \mathbf{a}, \quad \forall \mathbf{a} \in R, \quad \forall x \in R^m$
- ii) $d_{m,m}^r(x, 0 \cdot I_m) = \mathbf{m}_x$, con \mathbf{m}_x la media de $x \quad \forall x \in R^m$
- iii) $d_{m,m}^r(\mathbf{m}_1 \cdot I_m, \mathbf{m}_2 \cdot I_m) = |\mathbf{m}_1 - \mathbf{m}_2| \quad \forall \mathbf{m}_1, \mathbf{m}_2 \in R$

Cada medida de distancia $d_{m,n}^r(\mathbf{x}, \mathbf{y})$ se puede usar como una medida de desigualdad. Por ejemplo, podemos usar $d_{m,n}^r(\mathbf{x}, \mu_{\mathbf{x}}, 1_m)$ el r -ésimo momento central de \mathbf{x} , para medir la desigualdad en una población. Tomando $r=1$ y $r=2$, Ebert obtiene dos medidas de desigualdad absolutas bien conocidas:

i) La desviación de la media de \mathbf{x} :

$$d_{m,m}^1(x, \mathbf{m}_x, 1_m) = \frac{1}{m} \sum_{i=1}^m |x_i - \mathbf{m}_x|$$

ii) La desviación estandar de \mathbf{x} :

$$d_{m,m}^2(x, \mathbf{m}_x, 1_m) = \left[\frac{1}{m} \sum_{i=1}^m |x_i - \mathbf{m}_x|^2 \right]^{\frac{1}{2}}$$

IV. UNA NUEVA MEDIDA DE DISTANCIA.

Continuando la línea seguida por Ebert, proponemos una clase de índices estadísticos de distancia entre dos distribuciones de renta. Esta clase de medidas se estudian en el caso discreto y se basan en diferencias de las variables respecto a sus respectivas medias.

Sean $\mathbf{x}=(x_1, x_2, \dots, x_n)$ e $\mathbf{y}=(y_1, y_2, \dots, y_n)$ dos distribuciones de renta de igual tamaño ordenadas de forma creciente, es decir:

$$x_i \leq x_{i+1} \quad \forall i = 1, 2, \dots, n-1$$

$$y_j \leq y_{j+1} \quad \forall j = 1, 2, \dots, n-1$$

La clase de medidas de distancia que proponemos tiene la siguiente expresión:

$$d_{n,n}^p(x, y) = \left[\sum_{i=1}^n \frac{1}{n} \left| |x_i - \mathbf{m}_x|^p - |y_i - \mathbf{m}_y|^p \right| \right]^{\frac{1}{p}}$$

donde μ_x y μ_y son las medias de las distribuciones de renta \mathbf{x} e \mathbf{y} respectivamente.

Esta familia de medidas satisface todas las propiedades requeridas para ser una función de distancia ética, salvo quizás la reflexiva, que para esta clase de índices nos diría que la medida de distancia toma el valor cero si y solamente si las diferencias de las

variables respecto a sus medias respectivas son iguales.

Como comprobamos posteriormente nuestra función de distancia verifica igualmente los axiomas adicionales propuestos por Ebert, lo que implica que las medidas de distancia son medidas absolutas.

Axioma 1: Homogeneidad lineal

$\forall n \in N, \forall x, y \in R_+^n$ y para todo escalar $l > 0$

$$d_{n,n}^p(lx, ly) = \left[\sum_{i=1}^n \frac{l}{n} \left| \frac{l}{l} x_i - \frac{l}{l} m_x \right|^p - \frac{l}{l} y_i - \frac{l}{l} m_y \right]^{\frac{l}{p}} = l d_{n,n}^p(x, y)$$

Axioma 2: Invarianza por traslaciones

$\forall n \in N, \forall x, y \in R_+^n$

$$d_{n,n}^p(x + a l^n, y + a l^n) = \left[\sum_{i=1}^n \frac{l}{n} \left| \frac{l}{l} x_i + \frac{l}{l} a - \left(\frac{l}{l} m_x + \frac{l}{l} a \right) \right|^p - \frac{l}{l} y_i + \frac{l}{l} a - \left(\frac{l}{l} m_y + \frac{l}{l} a \right) \right]^{\frac{l}{p}} = d_{n,n}^p(x, y)$$

donde a es cualquier escalar tal que

$$x + a l^n \in R_+^n, \quad y + a l^n \in R_+^n$$

De esta forma cada índice de distancia $d_{n,n}^p(x, y)$ se puede usar como una medida de desigualdad absoluta. Al igual que la clase de medidas definidas por Ebert, si $p=1$ y $p=2$, usando $d_{n,n}^p(x, \mu_x 1^n)$ obtenemos dos medidas de desigualdad absolutas muy utilizadas:

a) La desviación media de \mathbf{x} :

$$d_{n,n}^1(x, m_x l^n) = \sum_{i=1}^n \frac{l}{n} |x_i - m_x|$$

b) La desviación estandard de \mathbf{x} :

$$d_{n,n}^2(x, m_x l^n) = \left[\sum_{i=1}^n \frac{l}{n} |x_i - m_x|^2 \right]^{\frac{l}{2}}$$

Finalmente vamos a aplicar al estudio de la distancia entre dos distribuciones de renta la métrica generalizada de Minkowski, cuya expresión es para dos distribuciones de renta $x=(x_1, x_2, \dots, x_n)$, $y=(y_1, y_2, \dots, y_n)$ de igual tamaño y ordenadas de forma creciente:

$$d_{n,n}^p(x, y) = K \left[\sum_{i=1}^n w_i (x_i - y_i)^p \right]^{\frac{l}{p}} \quad p \geq 1$$

donde K es una constante y w_i un factor de ponderación.

Esta métrica verifica las propiedades exigidas en la definición 1 para ser una función de distancia ética, así como los axiomas de homogeneidad lineal e invarianza por traslaciones propuestos por Ebert:

a) Homogeneidad lineal:

$$d_{n,n}^p(Ix, Iy) = K \left[\sum_{i=1}^n w_i (I x_i - I y_i)^p \right]^{\frac{1}{p}} = I d_{n,n}^p(x, y)$$

$$\forall n \in N, \forall I > 0, \forall x, y \in R_+^n$$

b) Invarianza por traslaciones:

$$d_{n,n}^p(x + a I^n, y + a I^n) = K \left[\sum_{i=1}^n w_i (x_i + a - (y_i + a))^p \right]^{\frac{1}{p}} = d_{n,n}^p(x, y)$$

$$\forall n \in N, \forall x, y \in R_+^n$$

Además, para $p=1$, $w_i=i$ y $K=2/\{n(n-1)\}$, obtenemos

$$d_{n,n}^1(x, \mathbf{m}_x I^n) = \frac{2}{n(n-1)} \sum_{i=1}^n i (x_i - \mathbf{m}_x)$$

que es precisamente el índice de desigualdad absoluto de Gini

BIBLIOGRAFÍA

- ATKINSON, A.B. (1970): "On the Masurement of Inequality". **Journal of Economics Theory**, 2; 244-263.
- CASAS, J.M. y F.J. CALLEALTA (1989): " Distancia entre distribuciones de renta: contratación y aplicación". **Actas de las XIV Jornadas Hispano-Lusas de Matemáticas**; 789-793.
- CHAKRAVARTY, S.R. (1990): "**Ethical social index numbers**". Springer-Verlag Berlin. Heidelberg.
- CHAKRAVARTY, S.R. y B. DUTTA (1987): "A note on measures of distance between income distributions". **Journal of Economic Theory**, 41; 185-188.
- DAGUM, C. (1980): "Inequality measures between income distributions with applications". **Econometrica**, 48 (7); 1791-1803.
- EBERT, U. (1984): "Measures of distance between income distributions". **Journal of Economic Theory**, 32; 266-274.
- HARDY, G.H., J.E. LITTLEWOOD y G. POLYA (1934): "**Inequalities**". Cambridge University Press.
- KOLM, S.C. (1976a): "Unequal inequalities. I". **Journal of Economic Theory**, 13; 416-442.
- KOLM, S.C. (1976b): "Unequal inequality. II". **Journal of Economic Theory**, 13; 82-111.
- RUIZ-CASTILLO, J. (1986): "Problemas conceptuales en la medición de la desigualdad". **Hacienda Pública Española**, 101; 17-31.
- SEN, A. (1973): "**On Economic Inequality**". Oxford University Press.
- SHORROCKS, A.F. (1982): "On the distance between income distributions". **Econometrica**, 50 (5); 1337-1339.
- SILBER, J. y Z.M. BERREBI (1988): "Distance functions and the comparison of development levels". **Economics Letters**, 27; 195-200.
- ZUBIRI, I. (1985): "Una introducción al problema de la medición de la desigualdad". **Hacienda Pública Española**, 92; 291-317.