Medir la desigualdad en educación

Contribución al IV Workshop de la REDE TISSE (Territórios Inteligentes e Sustentáveis no Âmbito Social e Educativo), Andorra, noviembre 2022

Francesc J. Hernàndez (Universitat de València)

Medir la desigualdad en educación

- Medir la desigualdad como comparación de cuantiles
- Medidas de la desigualdad que consideran todo el universo
- Idem con un solo par de datos (modelo parabólico y modelo exponencial)
- Aplicaciones educatives
- Ampliación: la Curva del Gran Gatsby
- Hacia un indicador unitario (variables absolutas y relativas)

Contribución al IV Workshop de la REDE TISSE (Territórios Inteligentes e Sustentáveis no Âmbito Social e Educativo), Andorra, noviembre 2022

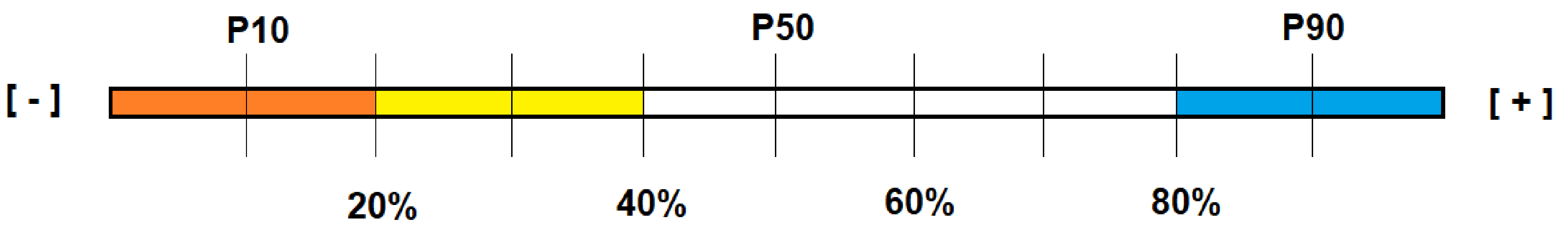
Medidas de la desigualdad como comparación de cuantiles

Cuantiles: partes de una muestra ordenada: cuartiles (4), quintiles (5), deciles (10), percentiles (100)

Unión Europea:

S80/S20

La relación entre el ingreso total recibido por el 20 % de la población con el ingreso más alto (quintil superior) y el recibido por el 20 % de la población con el ingreso más bajo (quintil más bajo). La renta debe entenderse como renta disponible equivalente. El indicador se basa en EU-SILC (estadísticas sobre ingresos, inclusión social y condiciones de vida).



)/S20) by sex [TESSI180]

OCDE/OECD:

S80/S20

P90/P10

La relación entre el valor del límite superior del noveno decil (es decir, el 10 % de las personas con ingresos más altos) y el del primer decil.

P90/P50

La relación entre el valor límite superior del noveno decil al ingreso medio.

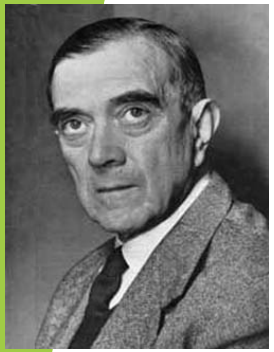
P50/P10

La relación del ingreso mediano al valor límite superior del primer decil.

Índice de Palma

El índice de Palma es la proporción de todos los ingresos recibidos por el 10% de las personas con los ingresos disponibles más altos dividida por la participación de todos los ingresos recibidos por el 40% de las personas con los ingresos disponibles más bajos.

https://data.oecd.org/inequality/income-inequality.htm



Medidas de la desigualdad que consideran todo el universo

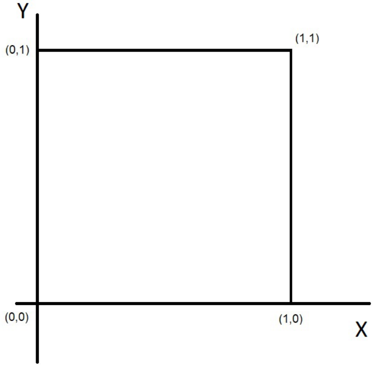
Índice de Gini (IGini) o Coeficiente de Gini (CGini) [CGini =100.IGini]

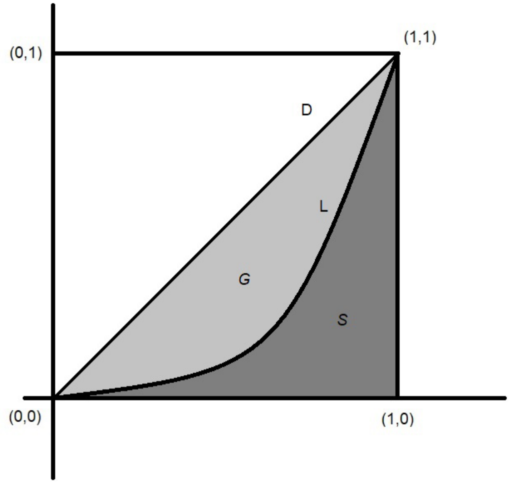
X: proporción de población sobre el total

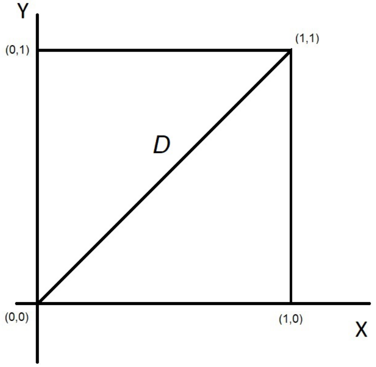
Y: proporción de ingresos sobre el total

Si x=0, entonces y=0

Si y=1, entonces x=1

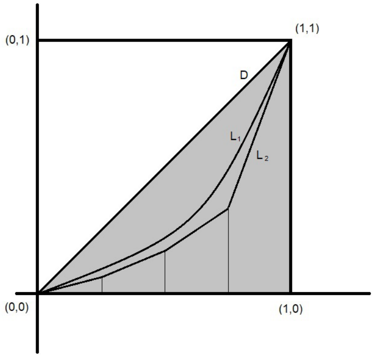


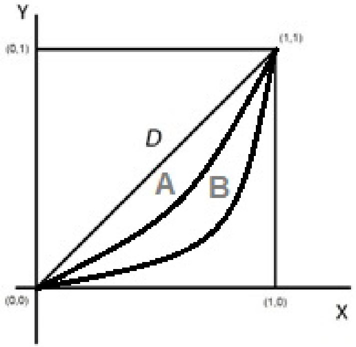




Plena igualdad

Los valores están sobre la diagonal creciente





Desigualdad

Una curva o línea quebrada (cuantiles) por debajo de la diagonal creciente

$$y = ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + fx + g$$

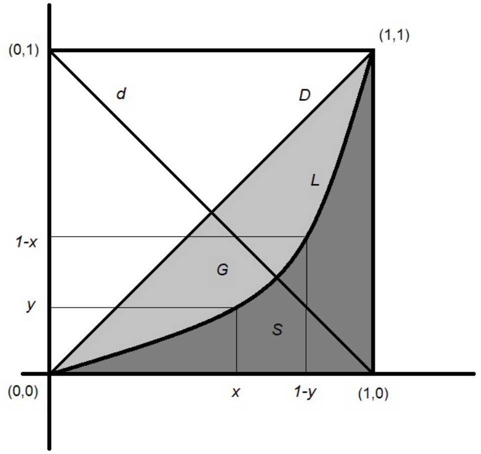
$$S = \frac{a}{7} + \frac{b}{6} + \frac{c}{5} + \frac{d}{4} + \frac{e}{3} + \frac{f}{2} + g$$

$$IGini = \frac{G}{G+S} = \frac{G}{\frac{1}{2}} = 2G \quad CGini = 200G$$

Medidas de la desigualdad que consideran todo el universo

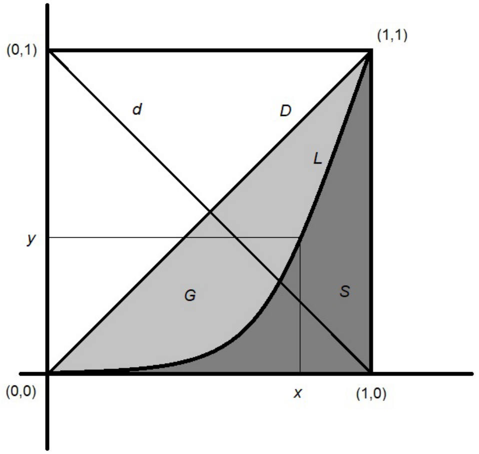
Índice de Gini (IGini) o Coeficiente de Gini (CGini) [CGini =100.IGini]

Con un dato (x,y), como disponemos también de (0,0) y (1,1), podemos suponer modelos según tipos de curvas



Curva parabólica

Condición de simetría respecto a la diagonal menor



Curva exponencial

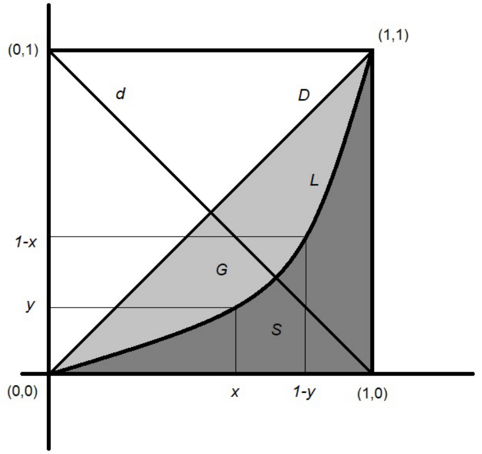
Sin condición de simetría respecto a la diagonal menor y con la forma $y = a^x$

En ambos casos la resolución pasa por: a) establecer la curva L, b) obtener su ecuación polinómica, c) calcular S mediante una integral definida (problema clásico de área bajo curva), d) calcular S, e) calcular el IGini o el CGini por las fórmulas anteriores

Medidas de la desigualdad que consideran todo el universo

Índice de Gini (IGini) o Coeficiente de Gini (CGini) [CGini =100.IGini]

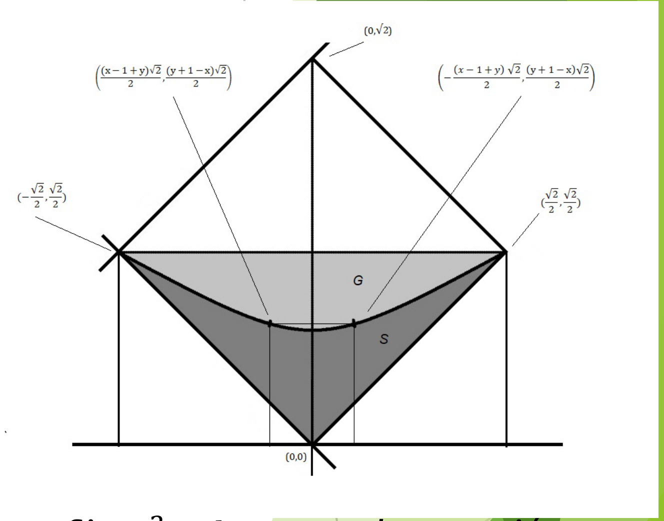
Curva parabólica



Curva parabólica

Condición de simetría respecto a la diagonal menor

Coordenadas originales (x, y)	Coordenadas después de la traslación y la rotación (x', y')
(0, 0)	$(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$
(0, 1)	$(0, \sqrt{2})$
(1, 0)	(0, 0)
(1, 1)	$(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$
(x, y)	$\left(\frac{(x-1+y)\sqrt{2}}{2}, \frac{(y+1-x)\sqrt{2}}{2}\right)$
$(1-y, 1-x)$	$\left(-\frac{(x-1+y)\sqrt{2}}{2}, \frac{(y+1-x)\sqrt{2}}{2}\right)$



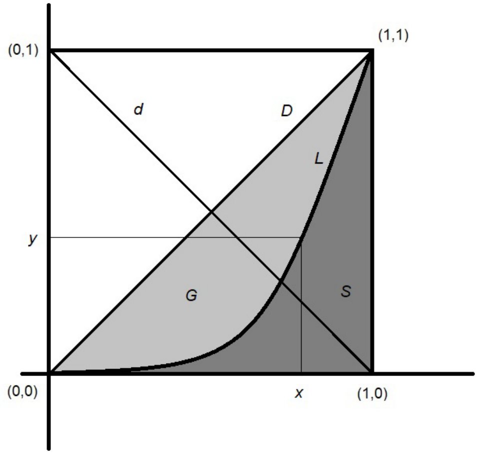
Si $ax^2 + bx + c$ es la ecuación de la curva, entonces

$$CGini = 200\left[1 - \left(\frac{a\sqrt{2}}{6} + b + c\sqrt{2}\right)\right]$$

Medidas de la desigualdad que consideran todo el universo

Índice de Gini (IGini) o Coeficiente de Gini (CGini) [CGini =100.IGini]

Curva exponencial



Curva exponencial

Sin condición de simetría respecto a la diagonal menor y con la forma $y = a^x$

1) Partimos de una fórmula simple:

$$y = (2^x - 1)^z$$

2) Calculamos z

$$z = \frac{\log y}{log(2^x - 1)}$$

3) Con el valor obtenido de z, asignamos valores desde 0,00 hasta 1,00 en la fórmula primera. Los representamos en unos ejes XY y le pedimos al programa que nos proporcione la ecuación polinómica de la línea de tendencia:

$$y = ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + fx + g$$

4) Se calcula la integral definida y se averigua S, lo que se puede simplificar como una suma de fracciones así:

$$S = \frac{a}{7} + \frac{b}{6} + \frac{c}{5} + \frac{d}{4} + \frac{e}{3} + \frac{f}{2} + g$$

5) Entonces:

$$CGini = 100 - 200S$$

Aplicaciones educativas

Estudio de los flujos de estudiantado en los centros educativos

(https://roderic.uv.es/handle/10550/77346)

Sea x la proporción de nuevos estudiantes E que acceden al curso 1º del año A respecto a todos los matriculados en ese curso 1º C y sea y la proporción de esos estudiantes E que son egresados del curso A+n, que denominaremos E', siendo n la duración normal del grado, respecto de los matriculados en el curso 1º C, de manera que:

$$x = \frac{E}{C}; y = \frac{E'}{C}; E \geq E'$$

Estudio del efecto free-rider en los Estudios Superiores **(Villar & Hernandez, en prensa)**

Sea Q la ecuación de la curva de la inversión disponible de los quintiles de ingresos de la población de un país y sea U la ecuación de la curva del gasto privado en Estudios Superiores. Teniendo en cuenta que la proporción mayor de gasto público en Estudios Superiores, se produciría un efecto free-rider si:

$$CGini_Q < CGini_U$$

Correlación de desigualdad social y desigualdad educativa

(https://ojs.uv.es/index.php/RASE/article/view/22027)

Dada una prueba estandarizada, tipo PISA, se puede calcular la correlación (R) entre la desigualdad social (CGini) y los resultados de rendimiento según cuartiles de índice socioeconómico.

Índices de desigualdad lingüística **(https://roderic.uv.es/handle/10550/76706)**

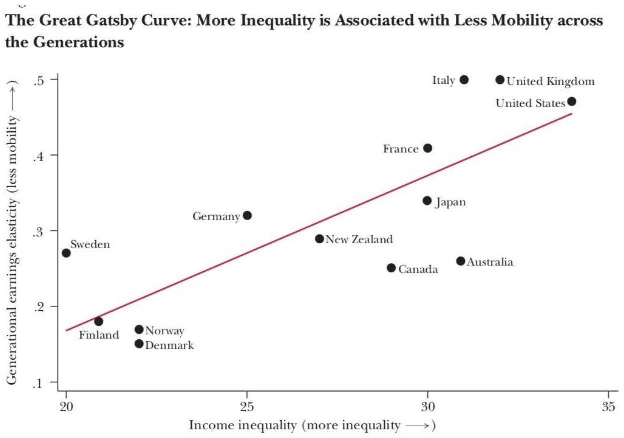
Sea x la tasa de competencia en una lengua minorizada en un determinado ámbito multilingüe y sea y la tasa de uso de esa misma lengua, se pueden comparar los CGini resultantes en diversos ámbitos.

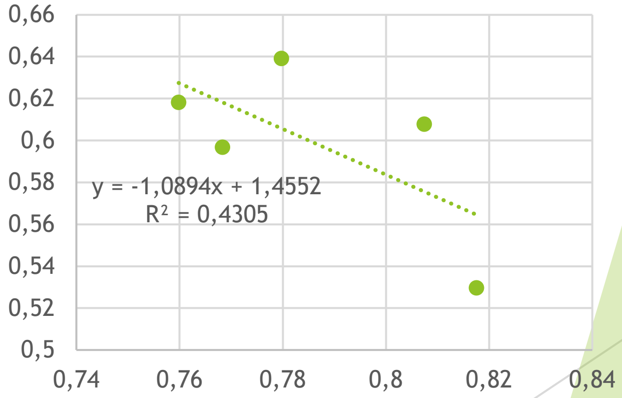
Ampliación: la Curva del Gran Gatsby

La Curva del Gran Gatsby fue popularizada por Alan Krueger, en una conferencia en el Center for American Progress, que había preparado con Reid Stevens y Miles Corak. Fue éste quien propuso el nombre, basándose en el conocido personaje de la novela de Scott Fitzgerald, al que se le hacía difícil ascender de clase social a pesar de disponer de dinero. La curva muestra la relación entre desigualdad económica e inmovilidad social. El descubrimiento del grupo de Krueger era que ambas variables –desigualdad e inmovilidad– presentaban una alta correlación para un grupo de países, cuya línea de tendencia recta presentaba una ecuación $y = 2{,}2x - 0{,}27$ y un valor de $R^2 = 0{,}76$, lo que significa un coeficiente de correlación de Pearson de $R = 0{,}87$.

«*Very illuminating –and disturbing*», escribió Paul Krugman (2012) de la Curva del Gran Gatsby. El célebre economista la define así:

> *En el eje horizontal se encuentra el coeficiente de Gini, una medida de la desigualdad. En el eje vertical existe la elasticidad intergeneracional de los ingresos: cuánto afecta a un aumento del 1 por ciento de los ingresos de tu padre a tus ingresos esperados; cuanto mayor es ese número, menor es la movilidad social.*



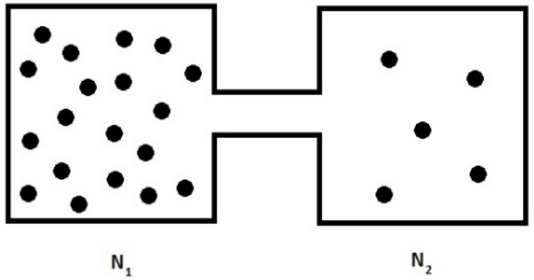


A la izquierda, la curva original de Krueger; a la derecha, una adaptación al estudio sociolingüístico del autor. Se trata de relacionar **desigualdad** e **inmovilidad**.

La Curva del Gran Gatsby puede utilizarse para relacionar desigualdades educativas y, por ejemplo, nivel educativo de los progenitores.

Hacia un indicador unitario (valores absolutos y relativos)

En los estudios educativos, se percibe la limitación de utilizar valores absolutos o valores relativos (frecuencias). ¿Se pueden combinar ambas en un **indicador unitario**? Esta cuestión se planteó en la física (termodinámica) y fue resuelta con la ecuación de Boltzmann. En su versión simplificada, podríamos afirmar: Si tenemos un gas en dos depósitos conectados N_1 y N_2 entonces podemos obtener una medida de su desorden (entropia) con la fórmula:



$$W = k \, ln\Omega \qquad \Omega = \frac{(N_1 + N_2)!}{N_1! \, N_2!}$$

En esta fórmula W es la medida de la entropia, k una constante (en el caso de Boltzmann, $ln\Omega$ es un número muy elevado y k es una constante muy baja: $1{,}380 \times 10^{-23}$), ln el logaritmo natural, N_1 y N_2 el número de moléculas en los depósitos y “!” el signo de factorial.

El logaritmo natural es el logaritmo que tiene como base el número e (que es una constante: 2,718281828459). De manera que: $\ln a = b$, si: $e^b = a$.

El factorial de un número x es el producto entero no negativo es el producto de todos los números inferiores o iguales a él hasta la unidad. Es decir:

$x! = x(x-1)(x-2)(x-3) \ldots (3)(2)(1)$

Las calculadores habituales no permiten calcular factoriales muy elevados. Para resolver este problema podemos:

a) Utilizar páginas o programa especializados. Por ejemplo: *https://www.wolframalpha.com/*

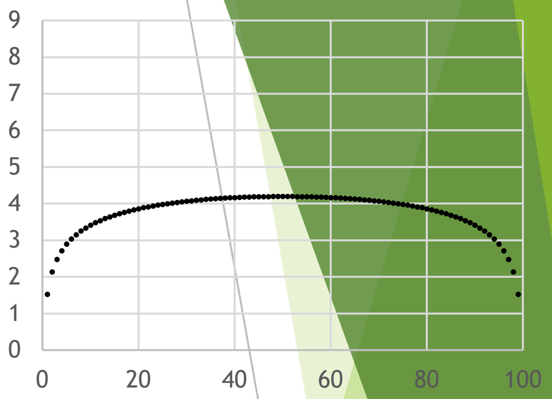
b) Usar la fórmula de Stirling: $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$

c) Como la representación XY de $N_1 + N_2$ y su ln proporciona una recta cuya inclinación depende de la proporción entre N_1 y N_2, entonces se puede calcular esta, buscar dos parejas de valores (x, y) más pequeños que cumplan la proporción, establecer sus factoriales, representarlos, calcular la ecuación de la línea que una esos dos puntos y substituir los valores para los datos originales.

Hacia un indicador unitario (valores absolutos y relativos)

En el caso de Boltzmann, $ln\Omega$ es un número muy elevado y k es una constante muy baja: $1{,}380 \times 10^{-23}$, pero esto no es así en las adaptaciones educativas. Por ello, podemos proceder a realizar una adaptación y (un logaritmo del logaritmo) y formular lo que denominaremos E (épsilon):

$$\mathrm{E} = \ln\left(\ln\frac{(N_1+N_2)!}{N_1!N_2!}\right)$$

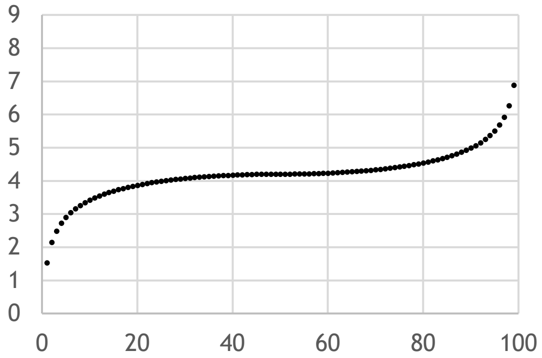


Ahora bien, como esta fórmula produce una representación simétrica (en el gráfico del ejemplo: $N_1 + N_2$=100), podemos redefinirla a partir de alguna clausula del tipo:

$$\mathbf{E} = \ln\left[\ln\frac{(N_1+N_2)!}{N_1!\,N_2!}\right] + \theta$$

$$\begin{cases} Si\ N_1 > N_2 \Longrightarrow \theta = 0 \\ Si\ N_1 < N_2 \Longrightarrow \theta = 2\left\{\ln\left[\ln\dfrac{(N_1+N_2)!}{\dfrac{(N_1+N_2)}{2}!\,\dfrac{(N_1+N_2)}{2}!}\right] - \ln\left[\ln\dfrac{(N_1+N_2)!}{N_1!\,N_2!}\right]\right\} \end{cases}$$

Lo que produce una representación del tipo:



En una investigación sobre competencias y usos lingüísticos (Hernández, TSC en prensa) se han obtenido correlaciones de $|R| = 0{,}850$ con las valores relativos y con las valores absolutas $|R| = 0{,}938$, lo que es un resultado muy satisfactorio.

¡Muchas gracias por su atención!

http://www.uv.es/fjhernan

francesc.j.hernandez@uv.es