

Consideraciones sobre la medida de la concentración por Gini

ROMERO GARCÍA, JOSÉ ENRIQUE
GAMERO ROJAS, JAVIER
Universidad de Sevilla

Introducción

A finales del siglo XIX y principios del XX en diversos países europeos (y en Estados Unidos) se estaban desarrollando rápidamente las primeras técnicas de la Estadística Descriptiva. Hasta cierto punto en este desarrollo hubo una “especialización” en diferentes países, así por ejemplo en Inglaterra se estaba desarrollando los campos aplicados a la psicología, biología, ciencias sociales, etc.; en Francia y en Rusia, por otra parte, existían considerables aportaciones con un enfoque más teórico y en Italia hubo una escuela de estadísticos aplicados al campo de la economía y concretamente al campo de las cantidades “macro” de los Estados. Ni que decir tiene que esta descripción esquemática no tiene vocación de ser exhaustiva en absoluto.

En este contexto es donde aparecen figuras como las de Pareto, Benini, Gini,... que estudian, entre otras cosas, la distribución y concentración de magnitudes relevantes para la macroeconomía (rentas, riqueza de individuos, tamaño de familias, créditos,...). Corrado Gini estudió ampliamente estos puntos y, como algunos otros estadísticos italianos, lo hizo fundamentalmente desde un punto de vista aplicado y, casi podríamos decir empírico. Gini, en el fondo, mantuvo una cierta desconfianza sobre el papel de las Matemáticas en la aplicación de la Estadística. Reproducimos a continuación algunas frases de la edición inglesa *The Contributions of Italy to Modern Statistical Methods* por Corrado Gini.

"for Statistics they [Mathematics] are no more than a means".

"This does not mean, however, that it is always convenient to use this method[Mathematics] ; the statistical material at our disposal may be too rough to allow of the application of exquisite methods, or for the purpose of the research it may appear useless to seek precision beyond a certain limit"

"His motto [of a statistician], instead of being represented by the words "Statistics with Mathematics," should be represented by "Statistics with the least mathematical means possible."

"I mean to say that there are wide fields in which... the statistical analysis does not need any instruments of calculation higher than simple arithmetic, and where it would be not only useless but harmful to have recourse to refined mathematical methods which, while they would not increase appreciably the fruitfulness of the research, would reduce the number of readers".

"Mathematics [they are not], however, sufficient to ensure success, which above all depends upon a kind of statistical intuition helped by wide experience".

Algunas de estas frases son muy interesantes en más de un sentido, pero no es nuestro propósito aquí el extendernos sobre este asunto, sino sólo dar una muestra de la mentalidad de Gini sobre la investigación analítica en la estadística.

El índice δ de Gini

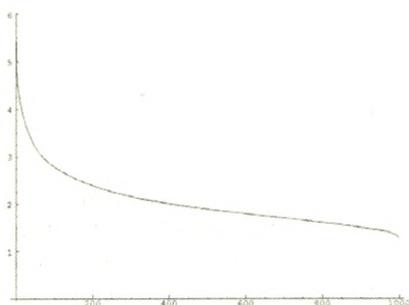
Siguiendo las ideas de Pareto y considerando las relaciones empíricas que observó con los datos disponibles de reparto de rentas en diferentes territorios y periodos de tiempo, Gini vio que, aproximadamente, se verificaba la siguiente relación:

$$1 - P_i = (1 - Q_i)^\delta \quad (1)$$

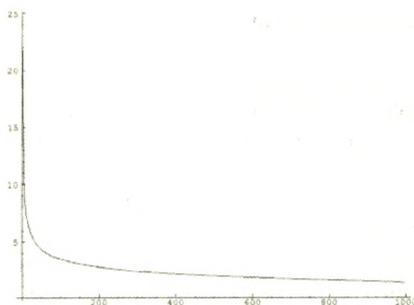
En donde, usando la notación más establecida, P_i es la frecuencia relativa acumulada en el individuo i -ésimo y Q_i es la cantidad relativa acumulada en ese mismo individuo i -ésimo (supuesto que los individuos están ordenados de menor a mayor renta).

Si esta relación se mantuviese cierta para todo " i " con un mismo δ , naturalmente este valor podría ser un indicativo de concentración de rentas en individuos. Con esta suposición, Gini calcula su coeficiente de concentración δ como un "valor medio" entre todos los "deltas" δ_i que se irán obteniendo para los diferentes individuos o clases " i ". Entendemos que este hallar el "valor medio" es un "ajuste" de la distribución observada al modelo (1)

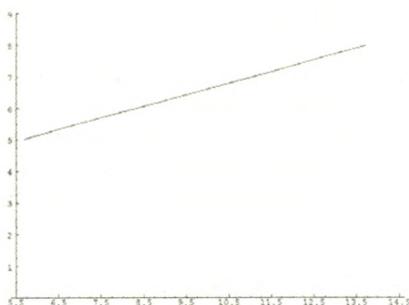
Este enfoque presenta el lógico problema de que no es adecuado cuando la relación (1) no es correcta, en el sentido de que para cada " i " hubiese valores "delta" claramente distintos, y peor aún, cuando los valores delta cambian según un patrón al crecer " i ". En realidad, sólo si la distribución de las rentas es de tipo Pareto, los valores delta serían teóricamente iguales. Con distribuciones tipo LogNormal o Gamma, por ejemplo, los deltas disminuyen sistemáticamente según crece " i ". En otros casos, como el propio Gini observará en la distribución del número de hijos por familia, los deltas crecen sistemáticamente. A continuación incluimos unos gráficos con los valores δ_i para diferentes supuestos distribucionales y ejemplos.



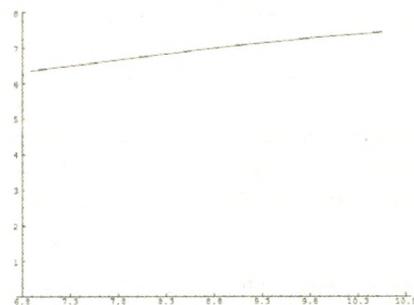
Distribución LN(0'5,1)



Distribución g(4,1)



Ejemplo de Gini pág.16



Ejemplo de Gini pág. 18

Los dos últimos gráficos corresponden a los ejemplos analizados por Gini en su tratado *Indici di concentrazione e di dipendenza* en las páginas 16 y 18. Puede observarse que en ninguno de los cuatro casos el valor δ es estable y por tanto no puede hablarse de un único δ que represente a la distribución.

En el caso de una distribución tipo Pareto el gráfico hubiese sido aproximadamente horizontal (con las posibles fluctuaciones muestrales, en su caso), por tanto la no horizontalidad viene a reflejar la diferencia de la distribución analizada con la Pareto. Así, veríamos que el segundo ejemplo de Gini es el más próximo a un caso paretiano.

Por tanto, en un amplio abanico de posibilidades realistas, no se puede hablar de “un” valor de δ que ligue P_i y Q_i , y por consiguiente el propósito de este índice queda desvirtuado, excepto en distribuciones paretianas.

Aunque apoyado en la distribución de Pareto, este índice δ tiene, sin embargo, a ojos de Gini, varias ventajas sobre el índice δ de Pareto. En lo que sigue exponemos el texto original de Gini y posteriormente un comentario nuestro encabezado por una “síntesis” entrecomillada.

“Ci pare che il nostro indice di concentrazione δ presenti alcuni vantaggi di fronte all'indice di distribuzione α del Pareto.

L'indice α è atto a ritrarre la distribuzione dei redditi solo al di sopra di un dato limite. Non si misura, dunque, mediante esso, la disuguaglianza di distribuzione di tutta la ricchezza, ma solo di una sua parte.

L'indice δ , al contrario, si adatta ugualmente alle seriazioni di redditi complete (che danno cioè tutti i redditi da 0 all'insù) che alle incomplete (che danno solo i redditi al di sopra di un dato limite);

“El índice δ sólo se puede aplicar a rentas por encima de cierto punto (debido a las características de la distribución de Pareto), mientras que δ puede aplicarse a la distribución completa de la renta”. Este comentario es criticable en el sentido de lo que hemos señalado anteriormente: δ es plenamente satisfactorio si la distribución es paretiana, pero si lo es, aparece el mismo problema de límite inferior de rentas que caracteriza a esta distribución. Es decir, la justificación plena de δ reside en la propia distribución de tipo Pareto.

(a) L'indice δ ha un campo di validità più esteso di quello dell'indice α .

“El índice α se aplica en un caso más restringido que el índice δ ”. En realidad Gini aplica δ a casos donde no es del todo satisfactorio (cuando los deltas no son constantes para todo i) y eso sería equivalente a aplicar α a distribuciones que no son exactamente paretianas. Como hemos apuntado anteriormente, en el fondo el procedimiento de cálculo “promedio” de Gini para δ es similar a “ajustar” α a distribuciones que no sean tipo Pareto.

(b) L'indice δ è molto più sensibile dell'indice α alle differenze di distribuzione dei redditi. Notevoli differenze di distribuzione dei redditi rimangono appena avvertite dai valori di α , specialmente se i valori di α sono bassi. Di qui era sorta presso molti l'idea, come vedemmo infondata, che la distribuzione della ricchezza fosse pressochè uguale in tutti i paesi e in tutti i tempi ;

“El índice δ es más sensible que α en la medición de la concentración”. Es decir, se afirma que $|\delta'(\alpha)| > 1$. Es interesante observar que δ y α son "autoinversos", o sea que $\delta(\alpha)$ es funcionalmente igual a $\alpha(\delta)$ (o, lo que es lo mismo, que esas funciones son simétricas respecto a la primera bisectriz). Por tanto, en sí, α y δ , globalmente, son igualmente sensibles, pero, y en esto Gini está en lo cierto, en el tramo habitualmente observado ($\delta > 2$), δ es más “sensible” que α .

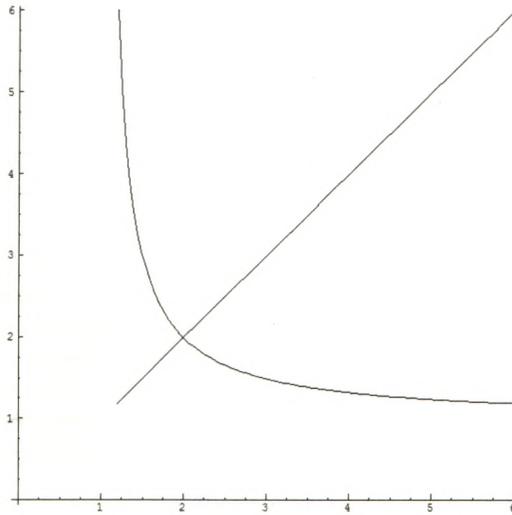


Gráfico de $d(a)$ y $a(d)$

Sobre este mismo punto, conviene señalar que Gini encuentra una refutación de la "ley" que Pareto creyó encontrar sobre la distribución de la riqueza. Pareto observó que la distribución de la riqueza en diferentes territorios y épocas tendía a seguir una distribución paretiana con un parámetro α similar; a ese tal valor α le quiso dar un significado de "ley" universal.

Gini refuta eso señalando que los valores "similares" de α corresponden a valores bastante distintos de α debido a la relación funcional antes señalada entre ambos parámetros. A su vez, valores algo distintos de δ implican concentraciones muy distintas. Citamos dos párrafos de estadístico italiano al respecto:

"Il Pareto, infatti, tratando le seriación dei redditi con metodo matematico, giunse alla conclusione que la distribuzione dei redditi globali è pressochè identica per tutti gli Stati e per tutti i tempi....."

"Notiamo però come, a differenze relativamente piccole nel valore dell'indice di concentrazione, corrispondano grande differenze nella distribuzione della ricchezza.

Se è $\delta=2$, $1/m$ del reddito tassato sarà posseduto da $1/m^2$ dei redditeri censito; per esempio $1/9$ dei censito possederà $1/3$ del reddito; $1/16$ ne possederà $1/4$.

Se è invece $\delta=3$, $1/m$ del reddito tassato sarà posseduto da $1/m^3$ dei redditeri censito; e così $1/2$ del reddito si troverà in mano non più di $1/4$, ma di $1/8$ di censito; $1/3$ del reddito in mano non di $1/9$, ma di $1/27$ dei censito; $1/4$ del reddito in mano non di $1/16$, ma di $1/64$ dei censiti."

Proseguimos con la exposición del resto de puntos sobre la comparativa de Gini entre su coeficiente y el de Pareto:

(c) L'influenza perturbatrice dell'evasione si fa sentire, come abbiamo mostrato,

meno sull'indice δ che sull'indice α .

- (d) L'indice δ ha un significato preciso. Esso indica l'exponente al quale bisogna innalzare una frazione dei redditi accertati per ottenere la frazione dei censito che li possiede. Quando δ crece, aumenta quindi la concentrazione dei redditi; nessun equivoco è possibile."

"El índice δ tiene una interpretación sencilla e intuitiva". Pero α , añadimos nosotros, por supuesto también la tiene: nos da precisamente la distribución y los cuantiles de la renta en forma sencilla.

A través de estos breves comentarios a las apreciaciones de Gini sobre su propuesta de índice en comparación con la medida de concentración de Pareto, podemos observar que no está fundamentada una significativa ventaja de δ sobre α . Aquí conviene traer a colación lo señalado en la sección 1 sobre la relativa desconfianza del estadístico veneciano respecto a la matemática teórica. Posiblemente si Gini hubiese profundizado en el análisis matemático de su coeficiente δ , se hubiese convencido de que no aportaba ventajas significativas y quizás, entrando ahora en lo hipotético, hubiese acelerado su adhesión definitiva al coeficiente que lleva su nombre.

Diferencias entre los conceptos y mediciones de la concentración entre Gini y Pareto

Pareto (1848-1923), aunque nacido en París y de madre francesa, hizo sus estudios y formación en Italia y puede considerarse perteneciente culturalmente a este país. Pareto fue de una generación anterior a Gini (36 años mayor) y es natural que éste tuviese como referencia relativa el trabajo de aquél. En *Indici di concentrazione e di dipendenza* Gini señala la disparidad entre la idea de "concentración" de Pareto y la suya.

"Altrettanto non si può dire dell'indice α , sulla cui interpretazione vi fu e vi è dissenso fondamentale. Mentre il Pareto infatti ritiene che il crescere di α indichi aumento di disuguaglianza nella distribuzione, il Benini ritiene al contrario che esso indichi diminuzione di disuguaglianza. È da avvertire che il dissenso dipende dal diverso significato che i due illustri statistici danno alle espressioni: maggiore o minore disuguaglianza della distribuzione. Poichè, per il Pareto, la disuguaglianza aumenta quando diminuisce la percentuale dei censiti con reddito superiore ad x , mentre, per il Benini, in tal caso, la disuguaglianza diminuisce.

Vi è certamente in tutte le definizioni qualche cosa di arbitrario. Ma dobbiamo guardarci, a scanso di equivoci, dall'attribuire ai termini un significato in contrasto con quello che viene loro comunemente attribuito.

Prendendo le parole nel loro significato etimologico e corrente, dobbiamo dire che la concentrazione della ricchezza aumenta e la sua disuguaglianza si fa più forte quando diminuisce la frazione dei censito al di sopra di un dato reddito che possiedono una data parte dei redditi accertati, o viceversa quando aumenta la parte dei redditi accertati posseduta da una data frazione di censito al di sopra di un dato reddito. Ora, in tal caso, aumenta il valore di δ e diminuisce corrispondentemente, in teoria, il valore di $\alpha = \frac{\delta}{\delta - 1}$.

È necessario dunque concludere che l'interpretazione del Benini, secondo il quale il diminuire di α indica un aumento di disuguaglianza nella distribuzione, ricorponde

meglio che quella del Pareto al significato che comunemente si attribuisce alla espressione <disuguaglianza nella distribuzione>.

Nei fatto però non avviene che ad un aumento di δ corresponda una diminuzione di α . Per la mancanza di coincidenza fra δ e $\frac{\alpha}{\alpha-1}$, può darsi che al creceré di δ corrisponda un aumento di α . In Hamburgo, per esempio, il valore di δ è aumentato dal 1883 al 1895, mentre è aumentato pure il valore di α ."

Es interesante señalar que en la visión de Pareto hay mucha concentración (en rentas) cuando una gran cantidad de personas tienen rentas pequeñas y un reducido número de personas tienen rentas mucho más mayores. Aquí, el economista italiano interpreta la concentración como "concentración de rentas parecidas" (dispersión relativa pequeña). En términos "modernos", siguiendo a Gini, esa situación se consideraría como "concentración pequeña" (dispersión relativa pequeña). Es decir, en términos de Gini "concentración" y dispersión (relativa) son conceptos paralelos, no opuestos,

Conclusiones

En sus trabajos sobre medición de la concentración distribucional de diversas magnitudes (rentas, riqueza, número de hijos,...) Gini fue elaborando a lo largo de los años una serie de coeficientes para determinar numéricamente tal cantidad. En estos sucesivos coeficientes podemos ver antecesores y "parientes" de los índices hoy corrientemente usados. Nos hemos centrado en este trabajo en el índice δ , que está especialmente relacionado con los trabajos de su antecesor Pareto e indirectamente relacionado con lo que conocemos como Curva de Lorenz. Gini se refiere al economista y matemático italo-francés en varias ocasiones para comparar sus descubrimientos y progresos en el campo. Así asevera, en su momento, que su coeficiente δ es superior al coeficiente α de Pareto y lo utiliza para refutar la "ley de Pareto" de igualdad de concentración entre los Estados. De alguna forma, también polemiza con las ideas de su antecesor en lo referente a la idea misma de "concentración" razonando sobre el significado de α y δ , como hemos observado.

Sin embargo, hemos podido señalar que el coeficiente δ no poseía todas las virtudes que Corrado Gini creía ver en él y, sólo hipotéticamente, hemos puesto en relación este exceso de confianza en la relativa displicencia del italiano hacia el método matemático.

Aunque su estudio se sale de los límites impuestos a este pequeño trabajo, cabe señalar que Gini, consciente de algunas limitaciones de su coeficiente δ , propuso simultáneamente otras variantes para diferentes aplicaciones. De esta manera nació el coeficiente τ , o el "coeficiente complejo" (x, ε) . Todo ello formó parte del proceso que llevó a Gini y a otros a construir las medidas de concentración actualmente en uso (y que siguen en proceso de evolución).

Bibliografia

- GINI, C. (1913), "Sulla misura della concentrazione e della variabilità dei caratteri", Atti del R. Istituto Veneto di Scienze, Lettere ed Arti.
- GINI, C. (1922), "Indici di concentrazione e di dipendenza", Biblioteca dell'Economista.
- GINI, C. (1926) "The Contributions of Italy to Modern Statistical Methods". Journal of the Royal Statistical Society, Vol. 89, No. 4. (Jul., 1926), pp. 703-724.