ANALISIS DEL CAMBIO EN LA CONCENTRACION A TRAVES DEL COEFICIENTE DE GINI

FERNANDO CORTÉS C.*
ROSA MA. RUBALCAVA**

INTRODUCCIÓN

AL ESTUDIAR EL MODO como se distribuye un total entre *n* observaciones (perceptores) se debe distinguir entre (i) el nivel de la variable (ii) el grado de desigualdad en su repartición (iii) su forma y (iv) el cambio o variación en los grados de concentración.

Para el análisis del nivel que ha alcanzado la variable se dispone de las medidas de tendencia central que para este fin suelen reducirse al promedio, la mediana y la moda.

Existe un gran número de índices estadísticos que permiten formarse una idea del grado de desigualdad existente en la distribución de una variable,[1] pero, sin duda el más conocido y utilizado es el coeficiente de Gini.

La forma de la concentración se puede estudiar geométricamente a través de la curva de Lorenz y aritméticamente mediante el coeficiente Gini intervalo.[2]

* Facultad Latinoamericana de Ciencias Sociales, sede México (FLACSO-México) y profesor de El Colegio de México.
** El Colegio de México.

[1] Por ejemplo la varianza de los logaritmos, el coeficiente de variabilidad, el índice entrópico de Theil, etc.

[2] Fernando Cortés y Rosa Ma. Rubalcava. "El coeficiente Gini-intervalo; la forma de la concentración a través de una descomposición del índice de Gini". *Demografía y Economía* Núm. 49, El Colegio de México.

Es bastante frecuente que al medir desigualdad a través del índice de Gini e interesarse por la variación que ha experimentado en el tiempo (o en comparar la concentración de una variable en diferentes espacios),* se recurra al simple expediente de establecer la diferencia o diferencias entre los valores calculados en cada momento lo que permite concluir en cuánto aumentó o disminuyó el grado de la concentración (o bien en qué lugar es más acentuada en comparación a otro tomado como base de la comparación).

Si bien la discrepancia entre dos o más valores del coeficiente de Gini entrega información que sintetiza el resultado de los procesos sociales que determinan la dinámica o la diferencial de la desigualdad, puede resultar de interés conocer los cambios (o diferencias) que llevan a la variación del coeficiente.

El problema que nos interesa abordar en este trabajo se puede plantear en los siguientes términos: habiéndose calculado varios coeficientes de Gini, uno para cada distribución de frecuencias ¿cómo se pueden comparar unos valores con otros?, o puesto de otra manera ¿cómo se puede descomponer la variación sufrida por el índice de Gini (ΔG)?, o bien ¿cómo ha cambiado la forma de la concentración?

A partir de esta interrogante central surgen naturalmente otras: ¿Qué intervalos de clase realizan el mayor aporte a ΔG? ¿Cuáles estratos perdieron y cuáles ganaron? ¿Con cuánto contribuyen a la formación del ΔG?

Las preguntas planteadas definen el campo problemático de este trabajo. Como las proposiciones que pretenden ser respuesta a estos interrogantes se construyen a partir del coeficiente Gini intervalo y considerando que no es muy conocido incluimos, a continuación, un breve resumen de sus características principales.**

Definición y principales características del coeficiente Gini intervalo

Dos o más conjuntos de observaciones pueden presentar el mismo nivel de desigualdad, lo que no impide que la forma de la concentración sea diferente.

Este hecho se presenta también cuando se trabaja sobre la base de distribuciones agregadas*** que originan dos poligonales (curvas) de Lorenz distintas que delimitan áreas de concentración idénticas. A pesar de la diferencia en la forma de la desigualdad los índices de Gini asumirán el mismo valor.

La construcción de un diagrama de concentración es el camino inmediato para percatarse de las distintas maneras en que se constituye el nivel de desigualdad global.

El área delimitada por la figura OACEG es igual a la que se encuentra en el interior de OBDFG, de manera que el coeficiente de Gini será el mismo para la

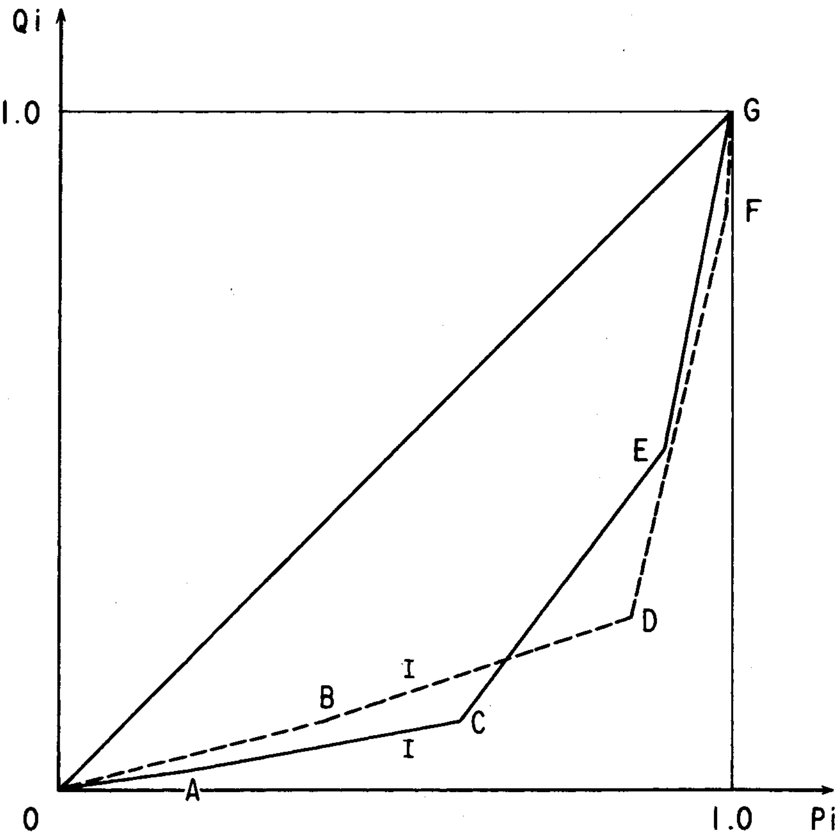
* En la exposición se enfatiza la comparación intertemporal, sin embargo las conclusiones se pueden extender fácilmente al caso de distribuciones situadas en distintos espacios.

** Un tratamiento detallado se encuentra en Cortés y Rubalcava, *op. cit.*

*** Aunque la desposición del índice de Gini que presentamos en este trabajo se aplica tanto al nivel agregado como al individual en la exposición nos referiremos principalmente a datos agrupados.

Gráfica 1

Forma de la concentración*



* **P es la proporción acumulada de las observaciones y Q del valor de la variable.**

distribución I que para la II. Ahora bien, el aporte a la desigualdad de los tramos bajos de II es mayor que el que hacen las categorías bajas de I y como el área total es igual en ambos casos, debe cumplirse que la contribución a la desigualdad de los tramos superiores de I debe ser mayor que la que hacen los superiores de II. Es decir, en las distribuciones representadas en la gráfica 1 se ve que es distinta la forma que conduce a un mismo nivel global de desigualdad.

¿Qué significan estas formas distintas? *Grosso modo* se puede interpretar la situación representada en la figura como que en la distribución I las unidades más pobres (vale decir, a las que les corresponde una menor proporción de la variable) reciben proporcionalmente más que las más desfavorecidas por la repartición que refleja la distribución II y que las más ricas en I tienen una participación en el total, menor que en II. El mismo nivel de desigualdad se conforma, en el caso de la distribución I por un "perjuicio" menor que en II de las unidades pobres y un "beneficio" también menor para las unidades ricas.

Lo anterior es sólo consecuencia del hecho de que se debe repartir un pastel de un tamaño dado y en tanto a algunas unidades les corresponda una tajada grande, otras necesariamente recibirán una parte más pequeña. Por ello si las observaciones más pobres de una distribución son menos perjudicadas que las más pobres de otra, se sigue inmediatamente que las más ricas no serán tan beneficiadas como las más favorecidas de la otra distribución.

Aun cuando baste construir el diagrama de concentración para distinguir la forma de la desigualdad, resulta ilustrativo abordar su estudio desde el punto de vista numérico (véase la gráfica 2).

Para ello aplicamos a cada segmento de la poligonal de Lorenz, la idea básica sobre la cual se construye el índice de Gini. En lugar de tomar toda el área de concentración (¶) la descomponemos en una serie de subáreas ($¶_1, ¶_2, \ldots, ¶_m$) tales que sumadas conformen el área total ¶ ($\sum_{i=1}^{m} ¶_i = ¶$). A partir de cada una de los $¶_i$ se propone un coeficiente de concentración que ayuda en el estudio de la forma de la desigualdad. Para ello se establece la relación entre cada $¶_I$ y su correspondiente subárea de concentración máxima A_i:

$$(1)\quad G_i = \frac{¶_i}{A_i} \qquad (i = 1, 2, \ldots, m)$$

obteniéndose así el coeficiente de desigualdad Gini-intervalo.

A partir de la definición de G_i y mediante operaciones algebraicas se llega a una fórmula simple para calcular el valor que asume G_i en cada intervalo.

$$(2)\quad G_i = 1 - \frac{Q_i + Q_{i-1}}{P_i + P_{i-1}} \qquad (i = 1, 2, \ldots, m)$$

Entre las propiedades del coeficiente Gini-intervalo hay que destacar que: (i) $0 \leq G_i \leq 1$, el valor *0* lo asume en el caso de equidistribución y el *1* cuando la concentración es máxima. (ii) Siempre se cumple que:

$$G_1 \geq G_2 \geq G_3 \geq \ldots \geq G_m$$

es decir que los valores que toma el coeficiente en una distribución de frecuencias son no crecientes.

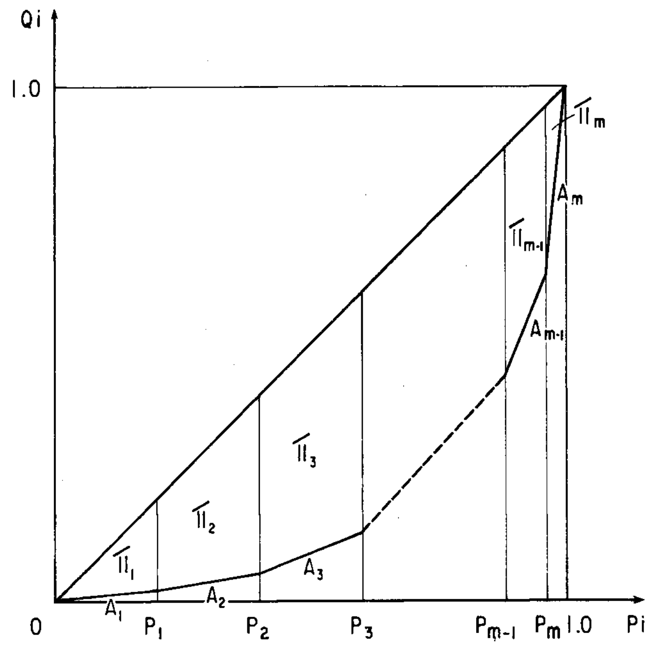
(iii) De la propiedad anterior tenemos como consecuencia que:

Si $G_1 = 0$ entonces $G_2 = G_3 = \ldots = G_m = 0$

Cuando hay distribución equitativa en el primer intervalo, ($G_1 = 0$), entonces se concluye que la variable se encuentra equitativamente distribuida entre todos los grupos. Ello es una consecuencia inmediata de la ordenación a que se someten las observaciones antes de construir la gráfica de concentración o calcular el índice de Gini. Hay que recordar que si el orden es de menor a mayor (como en el caso de las gráficas incluidas en este trabajo) entonces las observaciones más perjudicadas se encuentran en el primer intervalo de la distribución de frecuencias.
(iv) El coeficiente de concentración de Gini se puede expresar como una combinación lineal de los índices Gini-intervalo:

Gráfica 2

Coeficiente Gini-intervalo



(3) $G = \Sigma\, a_i\, G_i$

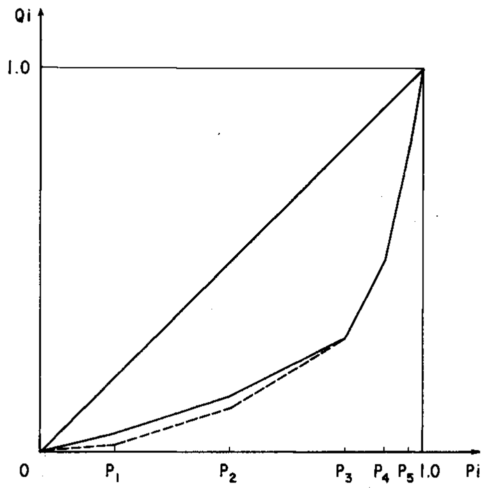
donde

(4) $a_i = (P_i - P_{i-1})(P_i + P_{i-1}) = p_i(P_i + P_{i-1})$

Se sabe que el coeficiente Gini-intervalo proporciona una buena vía para estudiar las modificaciones en la forma de la concentración, pero que no permite evaluar la ganancia o pérdida de las clases, o puesto en otros términos, que no entrega información para medir las contribuciones de los estratos a la variación del índice de Gini.* En efecto, cualquier redistribución entre, por ejemplo, dos intervalos de clase, sin modificación en los restantes altera los valores de los Gini-intervalos de toda la distribución de frecuencias y esta característica no corresponde a un índice que pretenda evaluar el aporte de cada estrato, lo que se puede apreciar fácilmente en una gráfica de concentración ya que una pérdida (o ganancia) de, por ejemplo, el primer intervalo en favor (o en contra de cualquier otro modifica la forma de la curva y por consiguiente varían algunos de los coeficientes Gini para cada intervalo. Ello es una consecuencia directa de la forma como se ha construido el diagrama de concentración, al definir el eje de las ordenadas** en términos de frecuencias acumuladas, por lo que cualquier modificación en las frecuencias precedentes se propaga a algunas de las consecuentes.***

Gráfica 3

Redistribución desde el primer intervalo en favor del tercero



* Cortés, F. y Rubalcava, R.M., *op. cit.*

** La preocupación se centra en el eje de las ordenadas porque, como se verá más adelante, para realizar las comparaciones se supone que la proporción de casos en cada intervalo de clase es el mismo en ambas distribuciones.

***Suponemos que la acumulación se hace desde los valores bajos a los altos de la variable. Es decir, que para un intervalo s cualquiera tenemos, por ejemplo, que:

$$Q_s = q_1 + q_2 + \ldots + q_s = Q_{s-1} + q_s$$

Para examinar esta idea desde el ángulo numérico supongamos que tenemos la siguiente distribución de frecuencias:

Cuadro 1

Distribución de frecuencias (hipotética) de una variable

p_i	q'_i	P_i	Q'_i
0.20	0.12	0.20	0.12
0.40	0.28	0.60	0.40
0.40	0.60	1.00	1.00
1.00	1.00		

p_i: proporción de observaciones en el intervalo *i*

q_i: proporción de la variable en el intervalo *i*

P_i: proporción acumulada de las observaciones en el intervalo *i*

Q'_i: proporción acumulada de las variables en el intervalo *i*

Supongamos que se produce un proceso de concentración en que el primer intervalo pierde 10% de su participación en la variable ganándolo el tercero. Por lo tanto la distribución original se modifica de la siguiente manera:

Cuadro 2

Distribución de frecuencias (hipotética) después de una transferencia desde el primer al tercer estratos

p_i	q_i	P_i	Q_i
0.20	0.02	0.20	0.02
0.40	0.28	0.60	0.30
0.40	0.70	1.00	1.00
1.00	1.00		

Los símbolos que encabezan las columnas del cuadro 2 se definen de la misma manera que los del cuadro 1.

El cálculo de los coeficientes Gini-intervalo requiere de una serie de operaciones intermedias que se resumen en el siguiente cuadro:

Cuadro 3

Coeficientes Gini y Gini-intervalo para las distribuciones original y modificada*

		Distribución original			Distribución modificada		
P_i	$P_i + P_{i-1}$	$Q'_i + Q'_{i-1}$	G'_i	$G'_i a_i$	$Q_i + Q_{i-1}$	G_i	$G_i a_i$
0.20	0.20	0.12	0.400	0.016	0.02	0.900	0.036
0.40	0.80	0.52	0.350	0.112	0.32	0.600	0.192
0.40	1.60	1.40	0.125	0.080	1.30	0.188	0.120
			G'=	0.208		G=	0.348

(*) Se entiende por distribución modificada aquella que se genera por una redistribución.

Como era de esperar, la medida de concentración Gini aumenta su valor debido a que la variable se distribuye más inequitativamente ahora que antes, sin embargo, lo que nos interesa destacar es que los valores de los coeficientes Gini-intervalo de la distribución modificada aumentan todos, así como también lo hacen las contribuciones que realizan al valor del Gini global. Sabemos que la redistribución sólo afectó al primero y tercer intervalos pero, a pesar de ello, también se modificaron los valores de G_2 y $G_2\ a_2$. Por lo tanto, *el coeficiente Gini intervalo no es un buen indicador de la contribución que hace cada intervalo de clase a la variación del índice de desigualdad de Gini.*

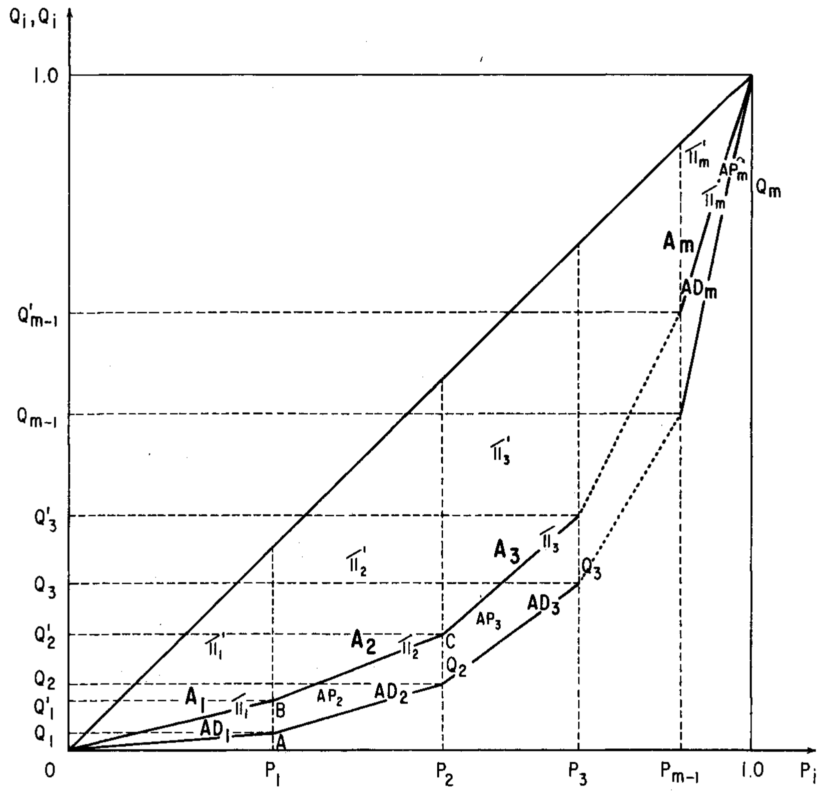
Antes de pasar a examinar una descomposición del coeficiente de Gini que se ajuste a nuestras preguntas centrales, hay que recordar que cuando destacamos el análisis del cambio en los niveles de desigualdad se debe considerar que la variación se puede originar en modificaciones en los valores de q_i con p_i constantes, cambios en los p_i con q_i constantes o en alteraciones de ambos simultáneamente.

En los dos primeros casos las distribuciones de frecuencia comparten algunos elementos lo que facilita su análisis comparativo; no acontece lo mismo en la tercera situación. En este trabajo sólo nos preocuparemos por los cambios que se deben a las modificaciones en los tamaños relativos del pastel y se supondrá constante la distribución porcentual de los perceptores. Este es el caso típico de la distribución del ingreso por "deciles" y, en general, de los cuadros estadísticos en que se controla (entre otros, por motivos de comparación) la distribución relativa de las unidades. Privilegiamos este punto de vista en razón de la disponibilidad de estadísticas. Difícilmente se tiene acceso a información que muestre las variaciones en los p_i dejando constantes los valores de q_i.

UNA DESCOMPOSICIÓN DEL ÁREA DE CONCENTRACIÓN: ÁREAS DE DESVENTAJA Y DE PROPAGACIÓN

Al producirse un cambio en la repartición de la variable tiene lugar un desplazamiento de la curva de Lorenz que origina un área $\Delta\P = \P - \P'$ en que $\P'$ es el área de concentración de la distribución primitiva y $\P$ el área de concentración de la distribución desplazada (*véase* gráfica 4). Sabemos que el área $\Delta\P$ puede ser un indicador del cambio en la forma de la concentración, de su nivel, o bien una consecuencia de ambos factores conjugados.

Gráfica 4



El área $\Delta\P$ se puede descomponer en :

$$\Delta\P = \Delta\P_1 + \Delta\P_2 + \ldots + \Delta\P_m = \sum_{i=1}^{m} \Delta\P_i$$

es decir, en una suma de subáreas, una para cada intervalo. Un cambio en la forma pero no en el nivel se manifiesta en que $\Delta\P = 0$, pero sus componentes no son todos nulos. Esto quiere decir que algunos $\Delta\P_i > 0$; otros $\Delta\P_i \leq 0$ de manera que sumadas hacen que $\Delta\P = 0$.

También $\Delta\P = 0$ en el caso trivial en que no ha habido modificaciones de ninguna especie, es decir, $\Delta\P_i = 0$ para todo $i = 1, 2, \ldots, m$.

El valor del área $\Delta\P$ será mayor que cero en aquellos casos en que aumenta el nivel y simultáneamente cambia la forma de la desigualdad.

Cada subárea de concentración formada por las diferencias entre las áreas de concentración de las curvas original y modificada ($\Delta\P_i$) se puede descomponer en dos partes: *a)* una que reflejaría el desplazamiento que habría tenido lugar en cada segmento de la poligonal de Lorenz bajo el supuesto de que no se hubiese modificado la participación de la clase estadística correspondiente y *b)* la parte restante se debe al cambio en la cuota (ya sea de ganancia o de pérdida) que obtuvo el estrato.

La primera de estas componentes recogería aquella parte de $\Delta\P_i$ que es consecuencia directa de la propagación del cambio en la pendiente de un segmento anterior al i-ésimo de la poligonal de Lorenz y que hemos denominado área de propagación $(A\ P)_i$, en tanto que la segunda muestra la modificación efectiva que ha experimentado la participación del intervalo de clase y que hemos llamado área de desventaja $(A\ D)i$.

De acuerdo con estas definiciones tenemos que:

$$(5) \quad \Delta\P_i = \mathrm{AP}_i + AD_i \quad (i = 1, 2, \ldots, m)$$

Una vez establecida la descomposición procederemos a examinar el modo de cálculo de las áreas involucradas en la fórmula (5). Las líneas discontinuas paralelas a la curva original de Lorenz en la gráfica 4 dividen el área $\Delta\P_i$ en los componentes AP_i y AD_i. Aquélla encerrada por los límites de intervalo y por las paralelas (por ejemplo el paralelogramo $ABCQ_2^*$) muestra lo que habría acontecido si la participación del intervalo no se hubiera modificado $(AP)_2$, y en consecuencia el área restante recoge los cambios debido a la participación de la clase en la distribución de la variable $(AD)_2$.

De la misma gráfica 4 se desprende que el cálculo de AP_i requiere que previamente se determine la ordenada del punto Q_i^* que se encuentra sobre la paralela a la poligonal original en cada intervalo que pasa por el punto de coordenadas (P_i, Q_i) de la curva de concentración modificada.

La pendiente de la recta sobre la cual se encuentra Q_i^* se obtiene a través de:

$$m_i = \frac{q_i}{p_i} \quad (i = 1, 2, \ldots, m)$$

expresión que corresponde a la razón de ventaja del i-ésimo intervalo. Para determinar la ordenada de Q_i^* basta con aplicar la ecuación de la recta cuando se conoce la pendiente y un punto por el que debe pasar (el punto de coordenadas P_i, Q_i). En este caso particular la ecuación asume la siguiente forma:

$$Q_i^* = m_i P_i + (Q_{i-1} - m_i P_{i-1}), \quad (i = 1, 2, \ldots, m)$$

Al agrupar convenientemente los términos y recordando que $p_i = p_i - p_{i-1}$, se puede rescribir como:

$$(6) \quad Q_i^* = q_i' + Q_{i-1} \qquad (i = 1, 2, \ldots, m)$$

La igualdad (6) nos dice que la ordenada del punto que divide a Δ¶ en las áreas de desventaja y propagación se obtiene como la suma de la participación acumulada del intervalo anterior, $(i-1)$, en la nueva poligonal y la parte que le correspondía a la i-ésima clase en la antigua distribución.

Entre las propiedades de Q_i^* queremos destacar que:

(i) $Q_1^* = q_1' + Q_0 = q_1'$ ya que $Q_0 = 0$

La ordenada del punto Q^* en la primera clase es igual a la parte relativa que le corresponde al primer intervalo en la distribución original.

(ii) $Q_m^* = q_m' + Q_{m-1} = 1 + (q_m' - q_m)$

En el último estrato la ordenada de Q^* se forma sumando a la unidad *el cambio* en las participaciones del último estrato en la nueva y antigua distribuciones. Cuando la curva modificada es producto de un proceso de concentración que favorece al último intervalo (aunque no necesariamente sólo al último) entonces $q_m' < q_m$ y el término $(q_m' - q_m)$ será negativo por lo cual Q^* será menor que la unidad. En el caso contrario $q_m' > q_m$ y Q_m^* será mayor que uno.

(iii) Si ambas distribuciones coinciden, es decir, si no se ha alterado la distribución de la variable entonces:

$$Q_i^* = Q_i = Q_i'$$

(iv) Cuando la parte que le corresponde al estrato i es la misma en ambas distribuciones $(q_i' = q_i)$ entonces:

$$Q_i^* = q_i + Q_{i-1} = Q_i$$

esto quiere decir que Q^* del i-ésimo estrato es igual a la proporción acumulada de la variable en la nueva distribución.

En el caso de que $q_i > q_i'$, se deriva que $Q_i > Q_i^*$ y si $q_i < q_i'$ se concluye que $Q_i < Q_i^*$.

Una vez que conocemos el modo de obtener la ordenada Q_i^* y sus principales propiedades, estamos en condiciones de establecer una fórmula que permita calcular el área de desventaja (AD). De la simple observación de la gráfica 4 se ve que $(AD)_i$ se puede obtener por diferencia de áreas de triángulos.

$$(AD)_i = \frac{(P_i - P_{i-1})(Q_i^* - Q_{i-1})}{2} - \frac{(P_i - P_{i-1})(Q_i - Q_{i-1})}{2}$$

$$(i = 1, 2, \ldots, m)$$

Después de algunos manejos algebraicos se llega a:

$$(7) \qquad (AD)_i = \frac{p_i\,(Q_i^* - Q_i)}{2} = \frac{p_i\,(q_i' + Q_{i-1} - Q_i)}{2} = \frac{p_i\,(q_i' - q_i)}{2}$$

$(AD)_i$ asume el valor cero siempre que $Q_i^* = Q_i$. Sabemos que esto ocurre en el caso en que coinciden plenamente la poligonal original y la modificada y también cuando no se alteró la parte proporcional de que se apropia el intervalo i. En otros términos, si al comparar dos distribuciones en el tiempo o en el espacio, la participación del estrato i es la misma $(q_i = q_i')$ entonces, como era de esperar, $(AD)_i = 0$. Si la cuota que le corresponde a la clase i en una y otra distribución es tal que $q_i > q_i'$ (que según las propiedades de Q_i^* conduce a que $Q_i > Q_i^*$), entonces $(AD)_i < 0$ márcándole una desventaja negativa, es decir, una mejor posición relativa en la distribución modificada que en la original. Ocurre exactamente lo contrario en la situación inversa, ya que al ubicarse en una situación desmedrada en la distribución nueva en relación a la antigua ($q_i < q_i'$ que implica $Q_i < Q_i^*$) $(AD)_i > 0$. El área de desventaja es positiva cuando el intervalo de clase i se encuentra en situación desmejorada.

Si se produce una redistribución entre dos clases manteniéndose constante la participación de las restantes, el área de desventaja será la misma (pero con signo opuesto) para ambos intervalos siempre que sus tamaños sean iguales. Supóngase que la transferencia Δq afecta a dos clases genéricas k y l de modo que

$$q_k = q_k' - \Delta q \text{ y } q_l = q_l' + \Delta q$$

De acuerdo con (7) tenemos que las áreas de desventajas antes de la transferencia serán:

$$(AD)_k = \frac{1}{2} \, p_k \, (q_k' - q_k) \text{ y } (AD)_l = \frac{1}{2} \, p_l \, (q_l' - q_l)$$

y una vez que se produce la redistribución son:

$$(AD)_k = \frac{1}{2} \, p_k \, (q_k' - q_k' + \Delta q) \text{ y } (AD)_l = \frac{1}{2} \, p_l \, (q_l' - q_l' - \Delta q)$$

Luego:

$$(AD)_k = \frac{1}{2} \, p_k \, \Delta q \quad \text{y} \quad (AD)_l = - \frac{1}{2} \, p_l \, \Delta q$$

Si $p_k = p_l$, se concluye que:

$$(AD)_k = (AD)_l$$

El interés también se puede centrar en el área de desventaja global, entendiéndose por ésta la suma de las áreas de cada intervalo de clase:

(8) $$A D = \sum_{i=1}^{m} AD_i = \frac{1}{2} \sum_{i=1}^{m} P_i \, (Q_i - Q_i')$$

Al reemplazar (7) en (8) se llega a:

(9) $$AD = \frac{1}{2} \sum_{i=1}^{m} (q_i' - q_i) \, p_i$$

El área de desventaja total es igual a la diferencia en las participaciones relativas de cada estrato en las distribuciones original y modificada, ponderada por el tamaño de los estratos. Se trata de una media ponderada de ganancias o pérdidas en que la nueva distribución se toma como base de comparación.

A primera vista parecería lógico esperar que cualquier redistribución de la variable debería llevar a que $A \ D = 0$. En efecto, una transferencia entre, por ejemplo, dos intervalos implica un aumento porcentual en la participación de uno y una baja de la misma magnitud en el otro. La pérdida relativa iguala a la ganancia de modo que la suma sería cero. Pero este razonamiento sólo es válido si $p_i = p$ para todo i (donde p es una constante) envuelto en el proceso redistributivo, en otros términos si las clases involucradas son de igual tamaño. Supongamos que la transferencia Δ_q se produce desde el primero al segundo intervalo* manteniéndose constante las participaciones de los restantes estratos. En consecuencia, al remplazar $q'_1 = q_1 + \Delta_q$ y $q'_2 = q_2 - \Delta_q$ y $q'_i = q_i$ $(i = 3, 4, \ldots, m)$ en (9) tenemos:

$$AD = \frac{1}{2} \left[(q_1 + \Delta q) - q_1\right] p_1 + \left[(q_2 - \Delta q) - q_2\right] p_2$$

$$AD = \frac{1}{2} (p_1 - p_2) \Delta q \tag{10}$$

Ahora bien, si $p_1 = p_2 = p$, entonces:

$$AD = 0.$$

Lo que significa que el área de desventaja global es nula. Este resultado se puede extender fácilmente a transferencias que afecten a más de dos intervalos de clase. Lo expuesto nos da pie para afirmar que las redistribuciones que involucran a dos o más intervalos de clase generarán una desventaja global nula siempre que los tamaños de los estratos involucrados sean iguales.

Un caso particular de interés es aquel en que todas las clases tienen el mismo peso relativo:

$$p_i = p \quad (i = 1, 2, \ldots, m)$$

ya que:

$$AD = \frac{1}{2} p \sum_{i=1}^{m} (q'_i - q_i) = 0$$

Si en una distribución de frecuencias los tamaños relativos de los estratos son iguales entonces el área de desventaja global siempre será igual a cero. Esto quiere decir que al analizar la distribución del ingreso por "deciles" el valor de *AD* será nulo.

A pesar de que hemos destacado la comparación en el tiempo, estos resultados son igualmente aplicables al estudio de distribuciones ubicadas en diferentes espacios.

De la ecuación (9) se desprende que el área de desventaja global será positiva siempre que

*Hacemos este supuesto por comodidad en las operaciones algebraicas, pero ello no se traduce en pérdida de generalidad del resultado que exponemos.

$$\sum_{i=1}^{m} q_i' \, p_i > \sum_{i=1}^{m} q_i \, p_i.$$

Sabemos que las transferencias implican que el aumento en la participación de un estrato significa la disminución en otro, por lo tanto si la redistribución se produjo entre las clases k y l, y se tiene que $q_k' > q_k$ entonces $q_l < q_l'$ en la misma proporción Δq. Esto permite concluir que $AD > 0$ siempre que el tamaño del estrato perjudicado sea mayor que el del beneficiado. Asumirá un valor menor que cero cuando el tamaño del beneficiado sea mayor que el del perjudicado.

Aun cuando este resultado se ejemplificó con dos intervalos se puede extender a transferencias que impliquen más de dos. *En conclusión el área de desventaja global será positiva si los tamaños relativos de las clases que han perdido son mayores que los de las que han ganado en el proceso redistributivo y será negativo en la situación contraria.*

Hay que notar que el signo del AD global depende básicamente de los pesos relativos de los estratos sin importar la distancia a que se encuentran ni el monto de la transferencia.

Un modo simple de obtener la fórmula que permita computar el área de propagación (AP) consiste en establecer una manera operativa para determinar el cambio en el área de concentración correspondiente al intervalo i ($\Delta\P_i$). De la gráfica 4 se desprende que:

$$\Delta\P_i = \P_i - \P_i' = \frac{1}{2}(P_i - P_{i-1}) \left[(P_i + P_{i-1}) - (Q_i + Q_{i-1})\right]$$
$$- \frac{1}{2}(P_i + P_{i-1}) \left[(P_i + P_{i-1}) - (Q_i' + Q_{i-1}')\right]$$

al factorizar y simplificar llegamos a:

$$(11) \qquad \Delta\P_i = \frac{1}{2}(P_i - P_{i-1}) \left[(Q_i' - Q_i) + (Q_{i-1}' - Q_{i-1})\right]$$
$$(i = 1, 2, \ldots, m)$$

$(AP)_i$ se puede determinar con cierta facilidad a partir de las ecuaciones (7) y (11). En efecto:

$$(AP)_i = \Delta\P_i - (AD)_i = \frac{1}{2}(P_i - P_{i-1})$$
$$\left[(Q_i' - Q_i) + (Q_{i-1}' - Q_{i-1}) - (Q_i^* - Q_i)\right] \; ; \; (i = 1, 2, \ldots, m)$$

Al sustiruir Q_i^* por su equivalente según la ecuación (6) y después de algunas operaciones algebraicas se llega a:

$$(12) \qquad (AP)_i = (P_i - P_{i-1})(Q'_{i-1} - Q_{i-1}) = p_i\,(Q'_{i-1} - Q_{i-1});$$
$$(i = 1, 2, \ldots, m)$$

Tanto de la gráfica 4, cuanto de la ecuación (12) se aprecia que el área de propagación del primer intervalo será siempre nula:

$$(AP)_1 = p_1 \ (0) \quad \text{ya que } Q'_0 = Q_0 = 0$$

y que a partir del segundo:

$$(AP)_i = 0 \qquad \text{si y sólo si } Q'_{i-1} = Q_{i-1} \text{ para } (i = 2, \ldots, m)$$

En palabras esto quiere decir que *el área de propagación del intervalo genérico i será cero si la participación acumulada hasta el estrato inmediatamente anterior es la misma tanto en la antigua como en la nueva distribución.* La consecuencia lógica de esta propiedad es que para cualquier $i > 2$, $(AP)_i = 0$ siempre que se produzcan transferencias que garanticen que $Q'_{i-1} = Q_{i-1}$.

Con el propósito de analizar un poco más esta idea reescribimos (12) en los términos de frecuencias simples en lugar de frecuencias acumuladas:

$$(13) \qquad (AP)_i = p_i \ \Big[(q'_1 - q_1) + (q'_2 - q_2) + \ldots + (q'_k - q_k) + \ldots + (q'_l - q_l) + \ldots + (q'_{i-1} - q_{i-1})\Big] \ ; \ (i = 1, 2, \ldots, m)$$

Esta igualdad servirá para estudiar, entre los muchos casos posibles, dos casos particulares en que $(AP)_i = 0$:

1. La nueva distribución se genera por transferencias que tienen lugar sólo entre los $(i - 1)$ primeros estratos.

Para profundizar en esta situación considérese que la nueva distribución se genera por una transferencia Δq que beneficia al intervalo l y perjudica al k en que k y $l \leq i - 1$. Simbólicamente esto quiere decir que:

$q'_i = q_i$ para todo i, excepto k y l y que $q_k = q'_k - \Delta q$;
$q_l = q'_l + \Delta q$

Al sustituir estas condiciones en la ecuación (13) se tiene:

$$(AP)_i = \ [(q'_k - (q'_k - \Delta q)) + (q'_l - (q'_l + \Delta q))] \ p_i; \ (i = 1, 2, \ldots, m)$$

$$(AP)_i = \ [(-\Delta q) + (\Delta q)] \ p_i = 0 \qquad (i = 1, 2, \ldots, m)$$

Este resultado es fácilmente generalizable al caso en que se producen transferencias que involucran uno o más pares de estratos considerados entre los $(i - 1)$ primeros.

Supóngase que la nueva distribución surge de un proceso redistributivo en que se extrae Δq a un intervalo y se entrega a los restantes $(i - 2)$. Sin pérdida de generalidad, considérese que el perjudicado fue el primero y los restantes, es decir el $2, 3, \ldots, i - 1$, fueron los beneficiados. Se tiene entonces que:

$$q_1 = q_1' - \Delta q \text{ y que } q_j = q_j' + b_j \, \Delta q \quad (j = 2, 3, \ldots, i - 1)$$

como el total Δq a repartir se agota en las clases beneficiadas debe cumplirse que $\sum_{j=2}^{i-1} b_j = 1$.

El reemplazo de estas condiciones en la ecuación (13) genera la expresión particular:

$$\begin{aligned}(AP)_i = [p_i \quad (q_1' - q_1' + \Delta q) + {} & (q_2' - q_2' - b_2 \, \Delta q) + \ldots + \\ & (q_k' - q_k' - b_k \, \Delta q) + \ldots + \\ & (q_l' - q_l' - b_l \, \Delta q) + \ldots + \\ & (q_{i-1}' - q_{i-1}' - b_{i-1} \, \Delta q)] \; ; \\ & (i = 1, 2, \ldots, m)\end{aligned}$$

$$(AP)_i = \left[\Delta q - \Delta q \sum_{j=2}^{i-1} b_j \right] p_i \qquad (i = 1, 2, \ldots, m)$$

Luego:

$$(AP)_i = 0 \qquad (i = 1, 2, \ldots, m)$$

La combinación de los casos estudiados permite sostener que si la nueva distribución de frecuencias se genera por transferencias que se encierran en los $(i - 1)$ primeros estratos, entonces el área de propagación del i-ésimo será nula.

2. La distribución modificada surge de transferencias entre los $(i - 1)$ primeros estratos y los $[m-(i - 1)]$ restantes, pero los flujos son encontrados y de igual magnitud.

Con el propósito de examinar esta idea se denominará grupo 1 al conjunto formado por los estratos 1 al $(i - 1)$ y grupo *2* al constituido por los restantes. Ahora bien, de la ecuación (13) se desprende que si la clase k (ubicada en el grupo 1) pierde Δq en favor de uno o más miembros del grupo 2 y la l (también perteneciente al grupo 1) gana Δq de los estratos que conforman el grupo 2, entonces el área de propagación que corresponde al i-ésimo intervalo de clase será nula. La demostración es idéntica a la presentada en el punto 1.

En consecuencia el área de propagación del intervalo genérico i será nula si ha habido transferencias que afectan a las $(i - 1)$ primeras clases o bien si las que ocurren entre los grupos 1 y 2 son de signo encontrado pero de la misma magnitud. También será cero si tienen lugar simultáneamente ambos tipos de redistribuciones.

Según la ecuación (12) el área de propagación correspondiente al intervalo i será positiva si $Q'_{i-1} > Q_{i-1}$ y negativa en el caso contrario. La condición $Q'_{i-1} > Q_{i-1}$ implica que los (i-1) primeros intervalos de clase tomados en conjunto han sido perjudicados por el proceso redistributivo. En efecto, la desigualdad nos dice que los primeros (i-1) intervalos (los más pobres) tenían una participación mayor en la distribución original que la que tienen en la distribución modificada. Cuando se cumple que $Q'_{i-1} < Q_{i-1}$ tenemos la situación exactamente opuesta.

De lo anterior se puede concluir que el área de propagación de los primeros intervalos tenderá a ser positiva cuando ha tenido lugar un proceso de concentración y negativa si se ha tendido hacia una repartición democrática de la variable. Se usa la forma *tenderá a ser* debido a que $Q'_{i-1} > Q_{i-1}$ no implica que $q'_j > q_j$ para todo $(j = 1, 2, \ldots, i - 1)$, por lo tanto a pesar de que se satisfaga la condición que afecta a las frecuencias acumuladas, algunos intervalos podrían resultar beneficiados.

Efecto desventaja, efecto propagación y Gini-intervalo

Una vez que se dispone de las fórmulas que permiten calcular las áreas de desventaja y de propagación se pueden establecer sus relaciones con el coeficiente Gini-intervalo.

Según la ecuación (5) tenemos que:

$$\Delta\P_i = (AD)_i + (AP)_i \qquad (i = 1, 2, \ldots, m)$$

en que $\Delta\P_i = \P_i - \P'_i$ donde $\P_i$ y $\P'_i$ simbolizan, respectivamente, las áreas de concentración de la distribución nueva y antigua.

Al sustituir $\Delta\P$ en la ecuación anterior se tiene:

$$\P_i = \P'_i + (AD)_i + (AP)_i \qquad (i = 1, 2, \ldots, m) \tag{14}$$

y al dividir (14) entre el área la máxima concentración del intervalo i se llega a:

$$\frac{\P_i}{A_i} = \frac{\P'_i}{A_i} + \frac{(AD)_i}{A_i} + \frac{(AP)_i}{A_i}$$

De acuerdo con la definición del coeficiente Gini-intervalo (*véase* ecuación 1) $G_i = \dfrac{\P_i}{A_i}$ y $G'_i = \dfrac{\P'_i}{A_i}$, es decir, el miembro de la izquierda y el primero de la derecha simbolizan el coeficiente Gini-intervalo de la distribución modificada y de la original respectivamente. A la relación entre las áreas de desventaja y de propagación y el área de máxima concentración que les corresponde se le denominará efecto desventaja y efecto propagación y se simbolizarán por $(ED)_i$ y $(EP)_i$. Luego:

$$(15) \qquad G_i = G_i' + (ED)_i + (EP)_i \qquad (i = 1, 2, \ldots, m).$$

Esta ecuación expresa que el coeficiente Gini-intervalo del intervalo genérico i de la nueva distribución de frecuencias se forma por la suma del coeficiente de la distribución original, el efecto desventaja y el efecto propagación.

Se sabe que el coeficiente Gini-intervalo asume valores dentro del intervalo cerrado cero, uno. Por lo tanto:

$$0 \leq G_i, G_i' \leq 1; \ (i = 1, 2, \ldots, m)$$

En la sección anterior se estableció que las áreas de desventaja y de propagación se considerarán positivas o negativas según la distribución se haya alejado o aproximado a la igualdad. En consecuencia al ponerlas en relación al área de máxima concentración de sus correspondientes intervalos se concluye inmediatamente que los valores absolutos de $(ED)_i$ y $(EP)_i$ no pueden ser mayores que la unidad:

$$\left|(ED)_i\right| \ ; \quad \left|(EP)_i\right| < 1 \ ; \ (i = 1, 2, \ldots, m).$$

Dada la relación entre las áreas de desventaja y de propagación y sus efectos, los valores que asuman estas últimas deben interpretarse de manera análoga a las primeras. Esto quiere decir que si el efecto desventaja del i-ésimo intervalo es:

a) *Igual a cero*, o bien coinciden ambas distribuciones o bien no se alteró la proporción de la variable de que se apropia el estrato i.

b) *Mayor que menos uno y menor que cero* $(-1 < (ED)_i < 0)$, entonces el i-ésimo intervalo tiene ahora una mejor posición relativa que antes.

c) *Mayor que cero, pero menor que uno* $[0 < (ED)_i < 1]$, luego la clase i empeoró su participación en la variable. Además sabemos que una transferencia entre dos estratos del mismo tamaño origina áreas de desventaja iguales, pero no ocurre lo mismo con el efecto desventaja. En efecto, como $(ED)_i = (AD)_i/A_i$, en la medida que las áreas de máxima concentración (A_i) sean diferentes los $(ED)_i$ serán distintos a pesar de que los $(AD)_i$ sean iguales.

Por otra parte tenemos que si el efecto propagación es:

a) Igual a cero $[(EP)_i = 0]$, la participación acumulada hasta el estrato $(i\text{-}1)$ es la misma en ambas distribuciones.

b) Negativo $[(EP)_i < 0]$, los primeros $(i\text{-}1)$ intervalos de clase han sido favorecidos en la nueva distribución y

c) Positivo $[0 < (EP_i) < 1]$, los primeros $(i\text{-}1)$ estratos resultaron perjudicados después de las modificaciones en la repartición.

Ahora bien, es fácil ver en la gráfica 4 que el área de máxima concentración en el i-ésimo intervalo se puede obtener a través de:

$$A_i = \frac{(P_i - P_{i-1})(P_i + P_{i-1})}{2} = \frac{p_i \; (P_i + P_{i-1})}{2}$$

Al reemplazar $(ED)_i$ y $(EP)_i$ de la igualdad (15) por las fórmulas de cálculo de sus correspondientes áreas (ecuaciones 7 y 12) y del área de máxima concentración A_i se tiene:

$$G_i = G_i' + \frac{\frac{1}{2} p_i \; (q_i' - q_i)}{\frac{1}{2} p_i \; (P_i + P_{i-1})} + \frac{p_i \; (Q_{i-1}' - Q_{i-1})}{\frac{1}{2} p_i (P_i + P_{i-1})} \quad (i = 1, 2, \ldots, m)$$

Después de simplificar se llega a la siguiente expresión:

$$(16) \qquad G_i = G_i' + \frac{(q_i' - q_i)}{P_i + P_{i-1}} + \frac{2\,(Q_{i-1}' - Q_{i-1})}{P_i + P_{i-1}} \quad (i = 1, 2, \ldots, m)$$

Para terminar esta sección mostraremos las potencialidades analíticas y limitaciones de la descomposición del coeficiente Gini-intervalo, para ello usaremos la formación que nos entregan las distribuciones de frecuencias de la sección 2 (cuadros 1 y 2). Recuérdese que la segunda distribución se generó a partir de la primera, bajo el supuesto de que tuvo lugar un proceso regresivo que condujo a que el primer estrato perdiese 10% que favoreció al último y que la participación del segundo se mantuvo constante.

El coeficiente de Gini para cada uno de los tres intervalos (G_i) de la nueva distribución se forma por la suma de los correspondientes coeficientes de la distribución previa y los efectos desventaja y propagación. La última columna del cuadro es igual a la suma de las tres inmediatamente anteriores. Hay que destacar que: *a)* El efecto desventaja del segundo intervalo de clase es igual a cero $[(ED)_2 = 0]$ resultado que refleja cabalmente lo sucedido. La participación de este estrato no se modificó. *b)* La transferencia de 10% que perjudicó a la clase 1 en beneficio de la 3 se manifiesta en un efecto desventaja positivo para la primera $[(ED)_1 = 0.500]$ y negativo para la última $[(ED_3 = -0.062)]$. La magnitud absoluta del efecto desventaja en uno y otro caso es sustancialmente distinta como resultado de las diferencias de tamaños de los estratos (que incide en el cálculo de las áreas de desventaja) y de las áreas de máxima concentración. *c)* El valor nulo del efecto propagación en el primer intervalo de clase sólo es una ilustración numérica del resultado algebraico que ya habíamos establecido. Además, como consecuencia del cambio en la participación del primer intervalo aparece efecto propagación en el segundo y tercero.

Cuadro 4

Efectos desventaja y propagación y coeficientes Gini-intervalo

P_i+P_{i-1}	$q_i'-q_i$	$2(Q_{i-1}'-Q_{i-1})$	$(ED)_i$	$(EP)_i$	G_i'	G_i
0.200	0.100	0.00	0.500	0.000	0.400	0.900
0.800	0.000	0.20	0.000	0.250	0.350	0.600
1.600	-0.100	0.20	-0.062	0.125	0.125	0.188

Un rápido vistazo a las dos últimas columnas del cuadro 4 muestra que:

$$G_i > G_i' \qquad \text{para} \quad (i = 1, 2, 3),$$

es decir, que al comparar la concentración por clases de la vieja y nueva distribución resultó que el coeficiente Gini-intervalo de la primera es siempre menor que el de la segunda. Al observar la composición de G_i línea por línea se concluye que las variaciones se deben a diferentes fuentes, en el primer estrato se origina en su totalidad en el efecto desventaja, en el segundo exclusivamente en el efecto propagación y en el tercero en una combinación de ambos que resulta de una propagación positiva mayor que una desventaja negativa (ganancia).

DESCOMPOSICIÓN DEL CAMBIO DEL COEFICIENTE DE GINI

Con las fórmulas (3) y (4) de la segunda sección podemos obtener una descomposición del cambio experimentado por el coeficiente de Gini (global). En efecto, si multiplicamos los dos miembros de (16) por a_i y sumamos, tendremos:

$$(17) \qquad G = G' + \sum_{i=1}^{m} \frac{(q_i' - q_i)a_i}{P_i + P_{i-1}} + 2 \sum_{i=1}^{m} \frac{(Q_{i-1}' - Q_{i-1})}{P_i + P_{i-1}} a_i ,$$

Esta ecuación se puede rescribir como:

$$(18) \qquad G - G' = ED + EP$$

donde $(G-G')$ es la variación experimentada por el índice de Gini (G simboliza al coeficiente de la nueva distribución y G' el de la antigua), ED y EP representan los efectos desventaja y propagación totales respectivamente y resultan ser una suma ponderada de sus correspondientes medidas en cada intervalo:

$$(19) \qquad ED = \sum_{i=1}^{m} \frac{(q_i' - q_i)\, a_i}{P_i + P_{i-1}}$$

$$(20) \qquad EP = 2 \sum_{i=1}^{m} \frac{(Q_{i-1}' - Q_{i-1})\, a_i}{(P_i + P_{i-1})}$$

Al sustituir a_i (*véase* ecuación 4) en (18) se llega a la siguiente fórmula de cálculo:

$$(21) \qquad \Delta G = G - G' = ED + EP =$$

$$\sum_{i=1}^{m} (q_i' - q_i) p_i + 2 \sum_{i=1}^{m} (Q_{i-1}' - Q_{i-1}) p_i =$$

$$\sum_{i=1}^{m} CED_i + 2 \sum_{i=1}^{m} CEP_i$$

donde CED_i y CEP_i denotan las contribuciones del intervalo i ya sea por efecto desventaja o propagación a la variación del índice de Gini.

La ecuación (21) nos entrega una respuesta a las interrogantes centrales que se plantearon en la introducción. Ya sabemos cómo ligar los valores de Gini calculados en dos distribuciones de frecuencias, también sabemos cómo se descompone ΔG, o cómo varía la forma de la concentración. Sin embargo aún no tenemos una medida que permita evaluar la contribución de cada intervalo a la variación del índice de Gini.

El efecto desventaja asociado a cada clase estadística (ED_i) mide, en esencia, el cambio que experimentó su participación en la distribución original con respecto a la nueva $[(q_i' - q_i)]$. Pero esta no es la única vía con que contribuye al valor del índice de Gini, puesto que su aporte también se hace presente a través del efecto propagación. Tanto su definición como su fórmula nos indican que este efecto acumula las "ganancias" y "pérdidas" de los estratos.

Con el propósito de llegar a una medida de la contribución de cada intervalo de clase a ΔG estableceremos, como un paso intermedio, la relación entre CEP_i y CED_i. Sabemos que:

$$CEP_i = 2\,(Q_{i-1}' - Q_{i-1})\, p_i \qquad (i = 1, 2, \ldots, m)$$

Desarrollando las frecuencias acumuladas y arreglando convenientemente los términos:

$$CEP_i = 2 \left[(q_1' - q_1) + (q_2' - q_2) + \ldots + (q_{i-1}' - q_{i-1})\right] p_i$$

$$(i = 1, 2, \ldots, m)$$

Por otra parte se tiene que:

$$CED_i = (q_i' - q_i)\, p_i \qquad (i = 1, 2, \ldots, m)$$

Al sustituir los términos encerrados por los paréntesis redondos de la expresión anterior en función de ED_i, se llega a:

$$CEP_i = 2 \left[\frac{CED_1}{p_1} + \frac{CED_2}{p_2} + \ldots + \frac{CED_{i-1}}{p_{i-1}} \right] p_i \quad (i = 1, 2, \ldots, m)$$

Luego:

$$(22) \qquad CEP_i = 2 \sum_{j=1}^{i-1} CED_j \; p_i/p_j \qquad (i = 1, 2, \ldots, m)$$

Esta ecuación muestra que CEP_i es una suma ponderada de las contribuciones de los efectos desventaja asociados a las $(i - 1)$ primeras clases.

No es difícil demostrar que:

$$(23) \qquad EP = \sum_{i=1}^{m} CEP_i = 2 \sum_{i=1}^{m} \sum_{j=1}^{i-1} CED_j \; p_i/p_j = 2 \sum_{i=1}^{m-1} CED_i \left[\frac{1 - P_i}{p_i} \right]$$

en que P_i simboliza las frecuencias relativas acumuladas de las observaciones en el estrato i. Por otra parte se tiene que:

$$(24) \qquad ED = \sum_{i=1}^{m} CED_i$$

Remplazando (23) y (24) en (21) se obtiene:

$$(25) \quad \Delta G = G - G' = ED + EP = \sum_{i=1}^{m} CED_i + \sum_{i=1}^{m-1} CEP_i$$

o bien:

$$(26) \quad \Delta G = \sum_{i=1}^{m} CED_i + 2 \sum_{i=1}^{m-1} CED_i \left[\frac{1 - P_i}{p_i} \right]$$

La ecuación (26) muestra la variación en la concentración (ΔG) en términos de las ventajas y desventajas que experimentan los intervalos de clase en relación a la distribución original. Al realizar las sumas indicadas en (26) se llega a:

$$(27)\quad \Delta G = 2 \sum_{i=1}^{m} CED_i \left[\frac{1}{2} + \frac{1 - P_i}{P_i}\right]$$

Si simbolizamos $C_i \Delta G$ a la contribución del intervalo i a la variación del índice de Gini entonces podemos rescribir (27) de la siguiente forma:

$$(28)\qquad \Delta G = \sum_{i=1}^{m} C_i\, \Delta G$$

Ahora bien, el aporte de cada intervalo a la variación del coeficiente de Gini dependerá de si ha sido perjudicado o beneficiado en el proceso redistributivo. Sabemos que si una clase pierde entonces $CED_i > 0$ y $CED_i < 0$ si gana. El signo de cada sumando de (27) o de (28) depende del signo de CED_i ya que los restantes términos son todos positivos. De lo anterior se concluye que si un intervalo resulta perjudicado en la nueva distribución (con respecto a la antigua) entonces su aporte a ΔG será positivo, es decir, contribuirá a que aumente el grado de concentración. Por el contrario si fuese beneficiado $C_i \Delta G$ será negativo, por lo que reducirá el valor de la variación del índice de Gini.

Con estos antecedentes surgen de inmediato algunas preguntas ¿Qué ocurre con ΔG si es el resultado de la suma de valores con signos encontrados?, o más específicamente ¿cómo es que Gini aumenta su valor cuando ha tenido lugar un proceso de concentración y lo disminuye cuando ha ocurrido lo inverso? Con el simple propósito de entender la mecánica de la operación de los factores que inciden en ΔG, supongamos que la nueva distribución se genera por una redistribución regresiva, se le extrae una cantidad a uno de los primeros intervalos y y se le entrega a uno ubicado en un rango más alto en la jerarquía. La pérdida del primero se manifestará en una $CED_i > 0$ y la ganancia del segundo en una $CED_i < 0$. Como los pesos de (27) tienen la característica de ser no crecientes, entonces le dará mayor importancia a la pérdida y menor a la ganancia. Estos dos movimientos encontrados conducirán a un $\Delta G > 0$ y por lo tanto a un mayor valor de G en la nueva distribución respecto a la antigua. Un proceso redistributivo de sentido inverso conduce, por las mismas razones, a un $\Delta G < 0$.

Retomemos el ejemplo para examinar la potencialidad de las ecuaciones de descomposición de ΔG. La evaluación del miembro derecho de (21) sólo requiere ponderar por p_i la segunda y tercera columnas del cuadro 4.

Los valores del coeficiente de Gini para ambas distribuciones son conocidos. En efecto, en el cuadro 3 se obtuvo como resultado $G' = 0.208$ y $G = 0.348$. Al sustituir todos estos valores en la ecuación (21) se concluye que:

$$0.348 - 0.208 = -0.020 + 0.160 = 0.140$$

Cuadro 5

Descomposición de la variación del coeficiente de Gini en los efectos desventaja y propagación totales

P_i	$q_i' - q_i$	$2(Q_{i-1}' - Q_i)$	$P_i(q_i' - q_i)$ (CED_i)	$2(Q_{i-1}' - Q_i - 1)\ P_i$ (CEP_i)
0.20	0.100	0.000	0.020	0.000
0.40	0.000	0.200	0.000	0.080
0.40	-0.100	0.200	-0.040	0.080
			-0.020	0.160

El aumento del nivel de concentración se debe fundamentalmente al efecto propagación que se contrarresta por un movimiento opuesto del efecto desventaja. La transferencia de 10% del estrato 1 al 3 provoca una ventaja en el 3 que es, en términos absolutos, igual al doble de la pérdida del 1 lo que induce un efecto neto negativo en la medida de desigualdad global ($ED = -0.020$).

Veamos ahora cómo se constituye el valor de ΔG en términos de las contribuciones de las clases. Al alimentar (27) con la información del cuadro 5 se tiene que:

$$\Delta G = 2(0.020)\left[\frac{1}{2} + \frac{(1-0.20)}{0.20}\right] + 2(0.000)\left[\frac{1}{2} - \frac{(1-0.60)}{0.40}\right] + \\ + 2(-0.040)\left[\frac{1}{2} - \frac{(1-1)}{0.40}\right]$$

$$\Delta G = 0.180 + 0.000 - 0.040 = 0.140$$

Esta ecuación nos dice que la mayor concentración en la distribución de la variable (0.140) se debe fundamentalmente a la pérdida sufrida por el primer estrato (0.180) contrarrestada por la ganancia del tercero (- 0.040) y que el segundo no realiza aporte alguno. La redistribución de 10% en perjuicio de la primera clase y en beneficio de la tercera se tradujo en una contribución de 0.180 y - 0.040 respectivamente, al cambio en el valor del coeficiente de Gini.

Una interpretación de Δ G y las condiciones Pigou-Dalton y de cambio relativo

A partir de la igualdad (27) y después de algunas simplificaciones se llega a:

$$(29) \qquad \Delta G = \sum_{i=1}^{m} (q_i' - q_i)(2 + p_i - 2\, P_i)$$

Esta ecuación no sólo sirve para evaluar la contribución que realizan los intervalos de clase al cambio en el valor del índice de Gini, sino que también proporciona una forma de interpretar $\Delta\, G$ que permite examinar por qué no cumple el criterio de cambio relativo y sí satisface el de Pigou Dalton.

En la descomposición de $\Delta\, G$ dada por (29) es posible distinguir dos partes, una cuyo valor no varía en el tiempo $(2 + p_i - 2P_i)$ porque hemos supuesto que la proporción de observaciones en cada estrato no se altera de una distribución a otra, y la parte $(q_i' - q_i)$, que se modifica en función del cambio en las participaciones relativas. Tomando en cuenta estas características de los elementos constituyentes de (29), podemos decir que la variación en el valor del índice de Gini depende, en esencia, de las modificaciones experimentadas por la cuota de que se apropia cada estrato. El valor de $\Delta\, G$ es una suma ponderada de la diferencia entre q_i y q_i', que refleja lo acontecido a las participaciones relativas de cada clase estadística. $\Delta\, G$ resulta de la comparación de las proporciones de la variable que corresponden a cada estrato en una y otra distribuciones.

Con el propósito de usar las ecuaciones de descomposición para analizar las bondades estadísticas del índice de Gini examinemos brevemente los criterios que deben satisfacer las "buenas medidas de desigualdad". *a)* Invarianza respecto a transformaciones proporcionales o cambios de escala. Si aplicamos un coeficiente de desigualdad a una distribución de frecuencias debemos obtener el mismo resultado que si lo usamos para medir la concentración de la misma variable pero transformada proporcionalmente. Por ejemplo, si medimos el ingreso en pesos y luego en centavos el índice de desigualdad debiera arrojar el mismo resultado. *b)* La condición Pigou-Dalton establece que si se realiza una transferencia extrayéndose una porción a una unidad que posee más, en favor de otra que tiene menos, entonces el indicador debiera marcar una caída en el nivel de concentración como reflejo de una mayor aproximación hacia la norma democrática. Si la redistribución fuese regresiva entonces el coeficiente debiera aumentar su valor. *c)* La condición de cambio relativo exige que los indicadores de desigualdad sean sensibles a las variaciones en el grado de desigualdad según el nivel en que se realizan las transferencias. El índice de concentración debe ser más "elástico" en el caso de redistribuciones que involucren a los niveles bajos. Esto quiere decir que si se transfiere la tierra afectando a una persona que posee grandes extensiones y se favorece a un minifundista, el coeficiente debiera marcar una caída mayor que si se hubiese favorecido al propietario de un predio de tamaño medio; o bien que debiera experimentar una caída mayor si la redistribución del ingreso se hace en contra de la burguesía, en favor del proletariado que si beneficia a las capas medias.

Es evidente que el índice de Gini satisface el criterio de invarianza a transformaciones proporcionales de la variable ya que sólo depende de proporciones.

Para examinar el índice de Gini desde el ángulo de los criterios restantes hay que considerar que si bien se han desarrollado para intercambios entre las unidades individuales no es difícil extenderlos para el caso de transferencias en distribuciones de frecuencias de datos agrupados bajo la condición de que todos los estratos sean del mismo tamaño.*

Al imponer esta última condición sobre (29) llegamos a:

$$\Delta G = \Sigma \, (q_i' - q_i) \, [2 + (1 - 2_i)p] \tag{30}$$

El elemento $(q_i' - q_i)$ refleja directamente el impacto de las transferencias: es positivo para el intervalo cuya participación decayó y negativo para el que la aumentó. La suma de estas diferencias siempre será igual a cero debido a que cuando tiene lugar una redistribución lo que ganan unos intervalos de clase iguala las pérdidas de los otros.

La otra parte de la fórmula $[2 + (1 - 2_i)] \, p$ origina una serie aritmética decreciente cuya función es la de atribuir pesos diferenciales a los cambios en las participaciones relativas. Supongamos que se pasa de una distribución a otra a través de una transferencia que consiste en extraer una parte Δq al estrato r y entregársela al intervalo s.

Es conveniente tener presente que los estratos son todos del mismo tamaño, que están ordenados de menor a mayor y que los subíndices r y s denotan la posición de los afectados.** En este caso los dos únicos términos diferentes de cero en el cambio del índice de concentración son los correspondientes a r y s de modo que ΔG se reduce a:

$$\Delta G = (q_r' - q_r) \; [2 + (1 - 2r) \, p] \; + \; (q_s' - q_s) \; [2 + (1 - 2s) \, p]$$

La redistribución implica que:

$$q_r = q_r' - \Delta q \;\; \text{y} \;\; q_s = q_s' + \Delta q$$

Sustituyendo estas igualdades en la ecuación anterior tenemos que:

$$\Delta G = \Delta q \;\; [2 + (1 - 2r)p] \; - \; \Delta q \;\; [2 + (1 - 2s)p]$$

y

$$\Delta G = 2p \, \Delta q \;\; (s - r) \tag{31}$$

* En la tercera sección se vio que el impacto de una transferencia también depende del tamaño de los intervalos de clase involucrados. Para evitar el efecto tamaño se supone que todos son iguales, lo que establece un paralelismo con el caso de las observaciones individuales.

** Denotamos por r al estrato que pierde y por s al que se ve favorecido por la pérdida de r sin importar cuál de los dos tiene el lugar de mayor jerarquía.

De (31) se concluye que:
a) si $r > s$ entonces ΔG será negativa ($\Delta G < 0$) y si *b)* $r < s$ entonces ΔG será positivo. La primera de estas conclusiones nos dice que si el estrato r se ubica por encima del s en la ordenación, entonces la transferencia originará una caída en el valor del índice de Gini. El valor del coeficiente de Gini disminuye cuando la redistribución es progresiva. La segunda conclusión extraída de (31) nos señala que si la transferencia benefició a un intervalo ubicado en un rango más alto en la jerarquía que el perjudicado, entonces ΔG será positivo, y el índice de Gini experimentará un alza. Si la redistribución es regresiva, entonces G aumenta. Por lo tanto el índice de concentración de Gini cumple con el criterio Pigou-Dalton.

La ecuación (31) también nos permite entender por qué Gini no satisface el criterio de cambio relativo. En efecto, dada una transferencia ΔG, la variación en G depende de la distancia a que se encuentran los intervalos involucrados y no intervienen para nada sus ubicaciones. Esto quiere decir que si la redistribución se realiza entre dos pares de clases separadas por la misma distancia, la variación en G será de igual tamaño. Por ejemplo ΔG será el mismo si la transferencia Δq es en favor del tercero y en contra del primero que si perjudica al cuarto y beneficia al sexto. Esto quiere decir que el coeficiente de concentración de Gini no cumple con el criterio de cambio relativo.

Una aplicación

Con el propósito de mostrar las bondades analíticas de los desarrollos presentados hemos tomado de una publicación del Banco de México la información sobre el ingreso familiar mensual.*

Sobre la base de esta información hemos calculado los coeficientes de Gini obteniendo:

$$G_{1968} = 0.522 \quad \text{y} \quad G_{1977} = 0.482^{**}$$

Por lo tanto:

$$\Delta G = 0.482 - 0.522 = -0.040$$

La distribución del ingreso es más equitativa en 1977 que en 1968. Hubo una tendencia hacia una repartición más democrática.

El cuadro (7) resume la aplicación de las ecuaciones de la sección anterior.

* J. Diez Canedo y Gabriel Vera. "Distribución del ingreso en México: 1977". Subdirección de Investigación Económica, Banco de México, S.A. Análisis Estructural, Cuaderno I, p. 26.

** Estas cifras presentan leves diferencias con el estudio del Banco de México, que pueden explicarse por redondeo.

Cuadro 6

Distribución del ingreso familiar mensual (en México) de 1968 y 1977, por deciles

Decil de familias	1968		1977	
	p_i	q_i'	p_i	q_i
I	0.10	0.0118	0.10	0.0117
II	0.10	0.0219	0.10	0.0237
III	0.10	0.0304	0.10	0.0341
IV	0.10	0.0409	0.10	0.0457
V	0.10	0.0507	0.10	0.0586
VI	0.10	0.0645	0.10	0.0727
VII	0.10	0.0834	0.10	0.0932
VIII	0.10	0.1124	0.10	0.1221
IX	0.10	0.1634	0.10	0.1726
X	0.10	0.4206	0.10	0.3656

La descomposición del cambio en el nivel de desigualdad presentado en las columnas (2), (3) y (4) corresponde a la aplicación de la ecuación (21). El efecto desventaja total es nulo debido a que los intervalos son todos de igual tamaño ($p_i = p = 0.10$) y por lo tanto la variación del valor de Gini es igual a la contribución del efecto de propagación.

La simple observación de la segunda columna muestra que el proceso redistributivo en el lapso considerado perjudicó al decil de los más ricos (décimo decil), que afectó levemente al de los más pobres (primer decil) y que los restantes fueron beneficiados.* Esta información aunque útil es incompleta ya que no nos permite cuantificar la contribución de cada intervalo a ΔG.

Al aplicar la ecuación (27) o (30) se obtienen los resultados de la columna (5) que permiten evaluar el aporte que realiza cada decil a la variación del coeficiente de Gini. Se puede apreciar que los deciles más beneficiados son los comprendidos entre el III y el VIII, ambos extremos incluidos, que las ganancias del II y el IX son relativamente menos importantes, que el I modificó apenas su participación en el ingreso y que sólo hubo un decil (el X) que sufrió una pérdida relativa de importancia. Todo el proceso de redistribución progresiva experimentada por la distribución del ingreso en México entre 1968 y 1977 se debió, en esencia, al deterioro del 10% más rico de la población.

* Recuérdese que estas afirmaciones se refieren a la forma como se modificó la repartición de la variable y que a partir de ellas nada se puede decir respecto a variaciones en el monto del ingreso.

Cuadro 7

Descomposición de la variación en el coeficiente de Gini para la distribución del ingreso familiar mensual (en México), entre 1968 y 1977

Deciles	Contribución del efecto desventaja	Contribución del efecto propagación	Efecto Total	Contribución de cada intervalo a ΔG
	(CED_i)	(CEP_i)	(4)-(2)+(3)	$(Ci\Delta G)$
(1)	(2)	(3)	(4)	(5)
I	0.0000	0.0000	0.0000	(0.0002)
II	-0.0002	-0.0000	-0.0002	(-0.0031)
III	-0.0004	-0.0003	-0.0007	(-0.0056)
IV	-0.0005	-0.0011	-0.0016	(-0.0062)
V	-0.0003	-0.0020	-0.0028	(-0.0087)
VI	-0.0008	-0.0036	-0.0044	(-0.0074)
VII	-0.0010	-0.0053	-0.0063	(-0.0069)
VIII	-0.0010	-0.0072	-0.0082	(-0.0049)
IX	-0.0009	-0.0092	-0.0101	(-0.0028)
X	0.0055	-0.0110	-0.0055	(-0.0055)
	0.0000	-0.0398	-0.0398	-0.0399=-0.040

$\Delta G = 0.0000 - 0.0398 = -0.0398 \approx -0.040$

CONCLUSIONES

Es bastante usual que al estudiar los cambios en los grados de concentración de la distribución de frecuencias de una variable, se recurra a comparar los valores alcanzados por el índice de Gini. Este procedimiento permite sustentar aseveraciones de tipo, según sea el caso, "la desigualdad aumentó", "la desigualdad disminuyó" o bien "la desigualdad se mantuvo".

Si el interés radica en un análisis un poco más detallado, por ejemplo, saber cuáles intervalos mejoraron su posición relativa y cuáles la empeoraron, es necesario estudiar directamente la distribución o bien recurrir a otras medidas tales como la variación de las razones de ventaja. Pero estos análisis no permiten evaluar el aporte que realizan los estratos al cambio experimentado por el coeficiente de concentración de Gini.

El propósito central de este trabajo es, el de examinar los componentes que inducen a la variación en el valor de G (ΔG) cuando se compara la distribución de una variable en dos puntos del tiempo o del espacio.

Bajo el supuesto de que los tamaños de los estratos son los mismos para ambas distribuciones se definieron los efectos desventaja y propagación. Al modificarse la participación relativa de los intervalos se producen dos cambios en el área de concentración. Uno se debe directamente a la alteración en la inclinación de los segmentos de la curva de Lorenz (que origina el efecto desventaja) mientras que el otro es producto de la transmisión del cambio de pendiente sobre algunos de los restantes segmentos de la poligonal (generándose de este modo el efecto propagación).

La suma de los efectos desventaja y de propagación correspondientes a cada clase resultó ser igual a su coeficiente Gini-intervalo.

Tomando pie en el hecho de que el índice de Gini es una combinación lineal de los coeficientes Gini-intervalo, se expresa el cambio de G como la suma de los efectos desventaja y de propagación totales. De esta manera se obtuvo una primera aproximación para determinar los elementos que integran ΔG.

Como el efecto de propagación correspondiente a cada intervalo sólo puede surgir como consecuencia de los cambios en las participaciones relativas de los estratos, fue posible establecer sus vínculos con los efectos desventaja. De esta manera se obtuvo una fórmula para evaluar la contribución que realiza cada intervalo a la variación en el valor del coeficiente de Gini.

También se derivó una ecuación para ΔG que permite demostrar que el coeficiente de Gini cumple con el criterio Pigou-Dalton, pero no el de cambio relativo.

Para concretar las ideas desarrolladas en el ámbito estadístico, se aplicaron las medidas desarrolladas, al estudio del cambio que experimentó la distribución del ingreso familiar en México, entre los años 1968 y 1977. Un breve análisis permitió establecer que los únicos deciles perjudicados fueron el primero y el último y que los restantes se beneficiaron.