Tema 7: MEDIDAS DE CONCENTRACIÓN

1. Concepto.

Las medidas o índices de concentración tienen como objetivo fundamental cuantificar el grado de desigualdad en el reparto o distribución de una magnitud económica (rentas, negocio, beneficios, etc...), entre un número determinado de “unidades” (individuos, familias, empresas, etc...).

La forma de la distribución de una magnitud, representada por una variable estadística, ya se ha estudiado a través de diversas medidas de posición, dispersión, asimetría y apuntamiento. Lo que ahora nos interesa es la mayor o menor equidad en el reparto de la suma total observada de una magnitud entre los integrantes del conjunto perceptor de dicha suma. Para ello, deberemos recoger de cada elemento perteneciente al conjunto perceptor, la información de la cuantía individual recibida en el reparto. La dificultad reside en que, en muchas ocasiones, esa información viene agrupada en clases y, por tanto, el estudio de la concentración no se podrá hacer con la precisión debida.

Es evidente que las dos situaciones extremas que podemos considerar, respecto a la equidad en el reparto, son:

- **Mínima concentración o máxima igualdad**: cuando a todos los integrantes del conjunto perceptor se les asigna la misma cantidad en el reparto del monto total.

- **Máxima concentración o mínima igualdad**: cuando un único perceptor recibe la suma total a repartir y los demás no perciben nada.

Estas dos situaciones deberán estar claramente identificadas por las medidas de concentración que se definan y que, asimismo, deberán graduar las situaciones intermedias, entre las que se encontrará la mayoría de los casos estudiados en la práctica.

2. Índice de Gini.

Para la obtención de este índice de concentración se debe ordenar, previamente, el conjunto de perceptores de menor a mayor cuantía obtenida en el reparto de la magnitud distribuida. Si el número de perceptores es N y representamos por V_i el valor que le corresponde al perceptor i-ésimo, la ordenación supondrá que:

$$V_1 \leq V_2 \leq \cdots \leq V_i \leq \cdots \leq V_{N-1} \leq V_N$$

A continuación se obtendrán las cantidades acumuladas, tanto del número de perceptores como del valor percibido, para determinar, finalmente, las proporciones correspondientes a estas cantidades acumuladas.

Todo lo anterior se puede exponer en el siguiente cuadro:

percep-tores	Valor obtenido V_i	número acu-mulado de perceptores	valor acumulado U_i	$p_i = \frac{i}{N} \times 100$	$q_i = \frac{U_i}{U_N} \times 100$
1º	V_1	1	$U_1 = V_1$	p_1	q_1
2º	V_2	2	$U_2 = V_1+V_2$	p_2	q_2
...	...	...	...	...	...
iº	V_i	i	$U_i = V_1+...+V_i$	p_i	q_i
...	...	...	...	...	...
(N-1)º	V_{N-1}	N-1	$U_{N-1} = V_1+...+V_{N-1}$	p_{N-1}	q_{N-1}
Nº	V_N	N	$U_N = V_1+...+V_N$	$p_N = 100$	$q_N = 100$

Las dos últimas columnas nos describen de qué forma se ha llevado a cabo la distribución de la cuantía total entre los perceptores. Si el reparto hubiese sido igualitario, todos percibiendo la misma cantidad, se cumpliría que $p_i = q_i \ \forall\ i = 1, ... ,N-1$; mientras que en el caso de que un único perceptor obtuviera la cuantía total, caso de máxima concentración, se tendría que $q_i = 0 \ \forall\ i = 1, ... ,N-1$. En cualquier otro caso intermedio, que serán los considerados en la práctica, se dará que $p_i > q_i$, pudiéndose establecer que si estos valores están muy próximos la concentración será muy baja, y si están muy alejados entre sí, la concentración será elevada. Cuanto mayor sea la diferencia $(p_i - q_i)$ mayor será la concentración producida en el reparto, pudiendo, así, identificar en qué sectores o grupos del conjunto perceptor se da mayor concentración, por el mayor valor de esas diferencias.

El índice de concentración de Gini se define como:

$$\boxed{I_G = \frac{\sum_{i=1}^{N-1}(p_i - q_i)}{\sum_{i=1}^{N-1} p_i} = 1 - \frac{\sum_{i=1}^{N-1} q_i}{\sum_{i=1}^{N-1} p_i}} \quad \textbf{[7.1]}$$

Cumpliendo que:

- Caso de **mínima concentración o máxima igualdad** ($p_i = q_i \ \forall\ i = 1, ... ,N-1$): $\mathbf{I_G = 0}$

- Caso de **máxima concentración o mínima igualdad** ($q_i = 0 \ \forall\ i = 1, ... ,N-1$): $\mathbf{I_G = 1}$

- Casos intermedios: cuanta mayor concentración se detecte, más cerca de 1 se situará el valor del índice, y cuanta mayor igualdad haya en el reparto más cerca de 0 estará.

Por tanto:

$$\boxed{0 \le I_G \le 1}$$

Pero, como se ha expuesto antes, la información de la distribución nos puede venir agrupada en clases o intervalos. En este caso los datos deberán ordenarse según lo percibido por los integrantes de las clases y no por el valor que corresponda a toda la clase. Si se pudieran determinar las marcas de clase, éstas deberán estar ordenadas de menos a más. Las proporciones acumuladas de los valores se calcularán sobre los valores totales percibidos por las clases. El cuadro quedará de la siguiente forma:

clases	marcas de clase	frecuencias	valor obtenido V_i	frecuencias acumuladas	valor acumulado U_i	$p_i = \frac{N_i}{N} \times 100$	$q_i = \frac{U_i}{U} \times 100$
C_1	x_1	n_1	$V_1 = x_1 \cdot n_1$	N_1	$U_1 = V_1$	p_1	q_1
C_2	x_2	n_2	$V_2 = x_2 \cdot n_2$	N_2	$U_2 = V_1 + V_2$	p_2	q_2
...	...	...	...	...	...	...	...
C_i	x_i	n_i	$V_i = x_i \cdot n_i$	N_i	$U_i = V_1 + ... + V_i$	p_i	q_i
...	...	...	...	...	...	...	...
C_{n-1}	x_{n-1}	n_{n-1}	$V_{n-1} = x_{n-1} \cdot n_{n-1}$	N_{n-1}	$U_{n-1} = V_1 + ... + V_{n-1}$	p_{n-1}	q_{n-1}
C_n	x_n	n_n	$V_n = x_n \cdot n_n$	$N_n = N$	$U_n = V_1 + ... + V_n = U$	$p_n = 100$	$q_n = 100$
		N	U				

De tal forma que en la clase C_1 estarán incluidos los perceptores con menor cuantía asignada en el reparto, mientras que en la última clase estarán los que les haya correspondido mayor cantidad. Esta ordenación implica que las marcas de clase cumplirán que: $x_1 \le x_2 \le \cdots \le x_i \le \cdots \le x_{n-1} \le x_n$. Sin embargo, los valores V_i no tendrán que verificar una ordenación como ésa, ya que sus cuantías dependerán de las frecuencias (n_i) que correspondan a cada clase.

Para este caso el índice de concentración de Gini se definirá así:

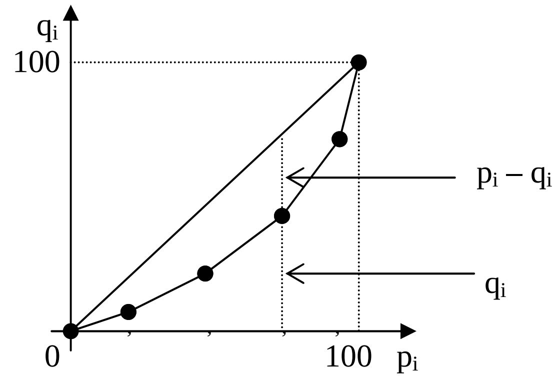
$$\boxed{I_G = \frac{\sum_{i=1}^{n-1}(p_i - q_i)}{\sum_{i=1}^{n-1} p_i} = 1 - \frac{\sum_{i=1}^{n-1} q_i}{\sum_{i=1}^{n-1} p_i}} \quad \textbf{[7.2]}$$

Manteniéndose la misma relación entre sus valores y los distintos casos de concentración que se describieron anteriormente. Lo único a señalar es que el valor de este último índice **[7.2]** será una aproximación del anterior **[7.1]**, dado que éste se calcularía con la información de todas las cuantías correspondientes a cada uno de los integrantes del grupo perceptor, con las proporciones acumuladas uno a uno, mientras que el otro utiliza las proporciones acumuladas por las clases o intervalos que recogen la información de los diferentes perceptores incluidos en ellas.

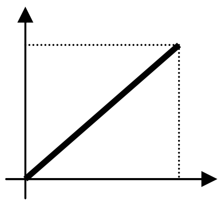
3. Curva de Lorenz.

La curva de Lorenz o curva de concentración es una gráfica que se deduce a partir de la información suministrada para el cálculo del índice de Gini y que, por tanto, refleja la mayor o menor concentración en la distribución de una magnitud. Como expondremos a continuación, existe una relación directa entre el índice de Gini y la forma de la curva de Lorenz, suponiendo ésta última una información adicional muy interesante sobre la forma en que se ha llevado a cabo el reparto de la cuantía total.

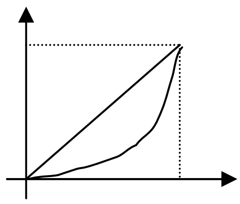
Para su representación gráfica llevaremos al eje de abscisas de un sistema de ejes cartesianos los valores de las proporciones acumuladas del número de perceptores "p_i", y en el eje de ordenadas situaremos las proporciones acumuladas de los valores obtenidos en el reparto "q_i". Ambas proporciones varían entre 0 y 100, y como sabemos cumplen que: $p_i \geq q_i$, de tal forma que si trazamos la diagonal, del cuadrado formado, que parte del origen de coordenadas, los puntos (p_i , q_i) se mantendrán en o por debajo de dicha diagonal. La curva de Lorenz es la resultante de unir todos los puntos (p_i , q_i), obteniéndose una gráfica, en forma de línea quebrada, que cuanto mayor sea el número de puntos más se aproximará a ser una curva, como la que se describe en el siguiente gráfico, con cuatro puntos intermedios (n=5 y, por tanto, $p_5 = q_5 = 100$):



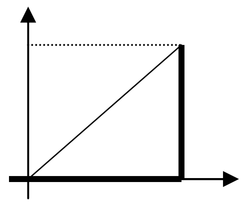
De tal forma que en el caso de mínima concentración la curva de Lorenz se confundirá con la diagonal. Luego, cuanta mayor concentración se establezca en el reparto, mayores serán las diferencias (p_i - q_i) y, por tanto, más alejada de la diagonal aparecerá la curva de Lorenz. En el otro caso extremo de máxima concentración la curva de Lorenz quedará formada por el propio eje de abscisas y la vertical trazada por el punto de abscisa igual a 100. Tal como se refleja en los siguientes gráficos:



mínima concentración
$\forall i = 1, ... ,N\text{-}1 \;\; p_i = q_i$
($I_G = 0$)



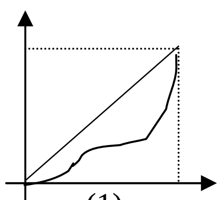
caso intermedio
$\exists\, i \,/\;\; p_i > q_i$
($0 < I_G < 1$)



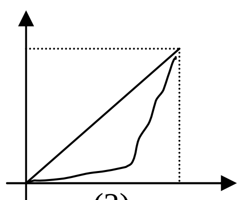
máxima concentración
$\forall i = 1, ... ,N\text{-}1 \;\; q_i = 0$
($I_G = 1$)

Como vemos, existe una relación entre el índice de Gini y la curva de Lorenz: a mayor alejamiento de la diagonal, por parte de la curva de Lorenz, mayor valor tomará el índice de concentración de Gini.

Mientras que el índice de Gini nos da un valor indicativo del nivel de concentración producido en el reparto, la curva de Lorenz nos describe gráficamente ese fenómeno, pudiendo identificar para que grupos de perceptores se acentúa la concentración y para cuáles de ellos se aminora. Como se puede distinguir en estos dos gráficos:



(1)



(2)

En el primer caso la concentración es de mayor importancia para los últimos perceptores, al ser en la segunda parte de la curva donde se separa más de la diagonal. Mientras que en el segundo caso la concentración es más amplia en el primer grupo de perceptores, a partir del cual la curva tiende con mayor pendiente a unirse a la diagonal.

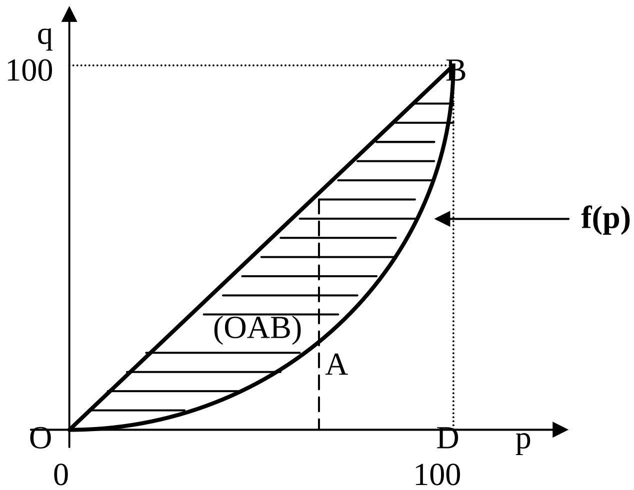
Además se puede determinar una relación más directa entre la curva de Lorenz y el valor del índice de Gini. Si le damos carácter continuo a las valores de q_i y de p_i la expresión del índice se podría igualar a:

$$I_G = \frac{\sum_i (p_i - q_i)}{\sum_i p_i} = \frac{\sum_i (p_i - q_i)\cdot \Delta p_i}{\sum_i p_i \cdot \Delta p_i}$$

Donde Δp_i representa una diferencia constante entre dos valores consecutivos de p_i. Si hacemos tender esa diferencia a cero, se obtendrá que:

$$\lim_{\Delta p_i \to 0} I_G = \frac{\lim_{\Delta p_i \to 0} \sum_i (p_i - q_i)\cdot \Delta p_i}{\lim_{\Delta p_i \to 0} \sum_i p_i \cdot \Delta p_i} = \frac{\text{Área (OAB)}}{\text{Área(ODB)}}$$

Siendo las áreas resultantes las que aparecen en el siguiente gráfico, donde la curva de Lorenz se supone continua y representada por la función f(p).



Es decir, que el valor del índice de Gini se puede aproximar por el cociente entre el área de la figura comprendida entre la diagonal y la curva de Lorenz y el área del triángulo (ODB) que forma la diagonal con el eje de abscisas y la vertical por el valor de p igual a 100. Esta aproximación será más precisa cuando sean menores las diferencias entre valores consecutivos de p_i, o, lo que es lo mismo, cuando sea mayor el número de perceptores considerados en el cálculo del índice de concentración. Si suponemos que la curva de Lorenz es continua y que las proporciones se determinan, no entre 0 y 100, sino entre 0 y 1, sin multiplicar por 100 los cocientes respectivos, se puede comprobar la siguiente relación:

$$I_G = \frac{\text{Área(OAB)}}{\text{Área(ODB)}} = \frac{\frac{1}{2} - \int_0^1 f(p)dp}{\frac{1}{2}} = 1 - 2\cdot \int_0^1 f(p)dp \quad 0 \le p \le 1$$

Que es una expresión teórica, bastante utilizada, del índice de Gini y la curva de Lorenz.

EJERCICIOS PROPUESTOS

1. En una determinada provincia, las rentas percibidas se han distribuido así:

Renta (en 10^5 pts.)	Nº de personas
2 — 10	10.000
10 — 30	25.000
30 — 50	40.000
50 — 70	20.000
70 — 100	5.000

1º: Determine el índice de concentración de Gini.
2º: Dibuje la curva de Lorenz.
3º: ¿Qué conclusiones se sacan de los resultados anteriores?.

2. Dos empresas **A** y **B** emplean a mil trabajadores cada una, clasificados en tres categorías: **I**, **II** y **III**. En un mes determinado han distribuido una misma masa salarial de cien millones de pesetas, de la siguiente forma:

EMPRESA A

categorías	**salarios**	**empleados**
I	50.000 - 70.000	500
II	70.000 - 80.000	400
III	150.000 - 650.000	100

EMPRESA B

categorías	**salarios**	**empleados**
I	10.000 - 50.000	500
II	100.000 - 200.000	400
III	200.000 - 300.000	100

1º: Calcule el índice de concentración de Gini para cada una de las empresas y dibuje en un mismo gráfico las dos curvas de Lorenz.

2º: A la vista de los resultados del apartado anterior critique y compare la equidad en el reparto de la masa salarial entre los empleados de las dos empresas.

3. En dos regiones diferentes se determinaron las siguientes distribuciones de la renta:

Región A

Renta en 10^6 pts.	Nº de personas
0,5 - 1,5	345
1,5 - 2,5	225
2,5 - 4,5	182
4,5 - 6,5	56
6,5 - 10	32

Región B

Renta en 10^6 pts.	Nº de personas
0,5 - 1,5	583
1,5 - 2,5	435
2,5 - 4,5	194
4,5 - 6,5	221
6,5 - 10	67

1º: Compruébese que el índice de concentración de Gini no depende de los niveles de renta, sino del número de individuos incluidos en cada nivel.

2º: Determínese la concentración de renta para el conjunto de las dos regiones.

4. Discuta la veracidad o falsedad de las siguientes proposiciones:

1) Puede obtenerse un valor aproximado del índice de Gini a partir de la curva de Lorenz.

2) Las curvas de Lorenz de dos distribuciones distintas no pueden cortarse.

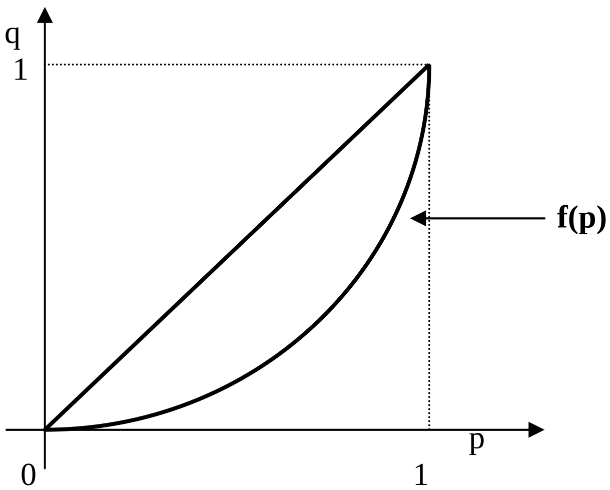
3) Si la distribución I tiene un índice de Gini: G_1=0,28 y la distribución II tiene G_2=0,30 la distribución (I + II) tiene por índice de Gini: G_{1+2} = 0,58.

5. Una sociedad tiene repartidas 50.000 acciones entre sus mil accionistas de la siguiente forma: 800 accionistas tienen 5 acciones cada uno, 120 accionistas tienen 50 acciones cada uno, 60 accionistas tienen 200 acciones cada uno y 20 accionistas poseen 1400 acciones cada uno.

1º: Obtenga el número medio de títulos por accionista y una medida del grado de concentración de los títulos en manos de los accionistas.

2º: A la vista del resultado anterior comente la equidad en el reparto del total de acciones entre los accionistas y la representatividad del número medio de títulos.

6. En el siguiente gráfico están representados en el eje de abscisas las proporciones acumuladas del nº de individuos entre los que se reparte cierta magnitud y en el eje de ordenadas las proporciones acumuladas de los valores repartidos. La función f(p) representa una estimación de la curva de Lorenz.



1º: Defina el Índice de concentración de Gini.
2º: Enuncie su relación con la curva de Lorenz.
3º: Calcule el valor del Índice de Gini si la curva de Lorenz es:

$$f(p) = p^2 \quad \text{para} \quad 0 \leq p \leq 1$$

7. En 1994 en España se estimaron los siguientes porcentajes que de la renta familiar disponible total correspondieron a los distintos hogares, representados por su porcentaje respecto al nº total de ellos y ordenados de menor a mayor renta percibida:

	Porcentaje de los hogares	**Porcentaje de la renta familiar disponible**
	20	**7,51**
	20	**11,63**
	20	**16,21**
	20	**21,42**
	20	**43,23**
TOTAL :	**100**	**100**

1º: Calcule el índice de concentración de Gini.
2º: Dibuje la curva de Lorenz.
3º: ¿Qué conclusiones se obtienen de los resultados anteriores?.

8. En cierta ciudad, dividida en 20 distritos municipales, se ha calculado el número de establecimientos comerciales, obteniéndose que hay un total de 60000 locales. Si se ordenan los distritos de menor a mayor número de establecimientos y se clasifican en 5 grupos, los resultados se pueden presentar de la siguiente forma:

Grupos de distritos	Número de distritos	Número medio de establecimientos por distrito
Grupo I	3	600
Grupo II	5	1500
Grupo III	5	2460
Grupo IV	6	5000
Grupo V	1	8400

Analice el grado de concentración en la distribución de los establecimientos comerciales entre los diferentes distritos municipales, por medio del índice de Gini y la curva de Lorenz.