Estudios de Economía Aplicada
Nº 10, 1998. Págs. 165-180

Una escala para la clasificación de variables socio-económicas en función del Coeficiente de Discriminación de Ivanovic

MIGUEL UCETA, S.
Universidad de Zaragoza

Esta versión incluye todas las correcciones sugeridas por el evaluador, las cuales me han parecido oportunas y por las que le quedo muy agradecido.

RESUMEN

El objetivo de este artículo es presentar una escala, dentro del intervalo de variación del coeficiente de Ivanovic, que permita clasificar el grado de discriminación de variables socio-económicas a la hora de efectuar comparaciones entre diversas zonas geográficas. Para construir la misma se han tomado como patrón los valores que alcanza dicho coeficiente cuando la variable se distribuye en las mencionadas zonas según los términos de algunas sucesiones recurrentes elementales. La gradación que se propone no es la única que se puede construir, pero pensamos que alcanza el objetivo perseguido.

Palabras Clave: Indicadores socio-económicos, discriminación, concentración, sucesiones recurrentes.

ABSTRACT

The objective of this article is to present a scale within the variation range of the Ivanovic coefficient which will permit to classify the discrimination degree of the socio-economic variables when making comparisons between different geographical areas. To be able to do it, we have used as a pattern the values reached by said coefficient when the variable is distributed among the geographical zones according to the terms of some elemental recurrent successions. Gradation here proposed is not the only one which can be made, but we think it reaches the objective pursued.

Key words: Socio-economic indicators, discrimination, concentration, recurrent successions.

Artículo recibido en enero de 1998. Revisado en mayo de 1998.

1. Introducción

A la hora de seleccionar variables para efectuar comparaciones entre distintas zonas geográficas (países, provincias, áreas comerciales, etc.) es habitual utilizar coeficientes que permitan medir el grado de discriminación que dichas variables presentan en el conjunto de países observados. Es interesante hacer notar que una misma variable puede tener un escaso poder discriminante en un conjunto de zonas observadas y no obstante tener un alto grado de discriminación en otro conjunto. De la misma forma se puede afirmar que, para el mismo conjunto de zonas, una misma variable puede ser muy discriminante en un periodo y no discriminar prácticamente nada en otro.

Para medir el poder de discriminación de una variable, uno de los coeficientes que puede utilizarse es el propuesto por Ivanovic (1974) que viene dado por la siguiente expresión:

$$CD_i = 2/(m.(m-1)).\sum_{j,l>j}^{k_i} m_{ji}.m_{li} \left| (X_{ji} - X_{li})/\mu_i \right|$$

Donde: X_i = Variable a estudiar
m = Número de zonas geográficas
X_{ji} = Valor de la variable X_i en la zona j
m_{ji} = Número de zonas con valor X_{ji}
μ_i = Media de X_i
k_i = Número de valores diferentes de X_i

En Zarzosa Espina (1994), se demuestra que dicho coeficiente está acotado entre 0 y 2, tomando el valor 0 cuando todos los valores de la variable coinciden (nulo poder discriminante) y siendo su límite 2 cuando el número de zonas tiende a infinito y existe una única zona con valor de la variable distinto de cero y todas las demás con valor nulo (máximo poder discriminante).

Uno de los problemas que puede aparecer, en el contexto que estamos considerando, es el tener que seleccionar la variable que presenta mayor poder discriminante entre varias. La solución no es complicada puesto que si halláramos el coeficiente de discriminación de cada una de ellas, en el mismo conjunto de zonas geográficas, podríamos ordenarlas sin mayor dificultad y elegir aquella o aquellas que más nos interesaran.

Para efectuar dicha selección no sólo debe tenerse en cuenta el "poder discriminante" sino también la "cantidad de información" que aporta la variable en cuestión a la finalidad buscada. En este sentido, el propio Ivanovic desarrolla el índice de "cantidad de información global" perfeccionado por lo que Zarzosa Espina denomina el índice de "cantidad de información global Ivanovic-Pena" (1996) que combina, en una misma medida, el coeficiente de discriminación de Ivanovic (1974) y el factor corrector de Pena Trapero (1977), que evita la duplicidad de la Información.

No obstante puede ocurrir que se nos plantee el caso de tener que estudiar una variable aislada, como puede ser el caso de la distribución geográfica de las ventas de una empresa o del Indice de Capacidad de Compra, ver Miguel Uceta (1996), del Indice Turístico por áreas comerciales, ver Chasco-Miguel (1997) a), etc. y nos interese saber si su poder discriminante es bajo, medio o alto, independientemente de la cantidad de información, ya que no hay que seleccionarla entre un conjunto de variables, para ello necesitaremos una escala de medida que nos permita clasificar la variable en función del CD.

El objetivo del presente trabajo es presentar una escala de medida que permita clasificar las variables y resolver por tanto el problema considerado. Para su construcción utilizaremos como patrones variables que se distribuyan, en el conjunto de zonas observadas, según los términos de algunas sucesiones recurrentes elementales. Es evidente que las variables de tipo territorial no suelen distribuirse de este modo pero pueden servirnos como modelos teóricos que nos permitan construir los intervalos en los que quede dividida la escala de clasificación.

2. El C.D. de Ivanovic en variables distribuidas según una progresión aritmética no decreciente.

Supongamos que la variable objeto de estudio se distribuye entre las m zonas consideradas ($m > 1$ y finito), según una progresión aritmética de diferencia $d \geq 0$ siendo $x_1 > 0$ el primer término de la progresión. Como x_1 es un número real no nulo, no perdemos generalidad si suponemos que $d = t.x_1$, donde $t \geq 0$. En este caso se tiene:

$$x_1 = x_1$$

$$x_2 = x_1 + d = x_1 + t.\, x_1 = x_1(1 + t)$$

....

$$x_m = x_1 + (m - 1).d = x_1(1 + (m - 1).t)$$

$$\mu = (x_1 + \ldots\ldots + x_m)/m = x_1(2 + (m - 1).t)/2$$

Si $t = 0$ entonces $\mu = x_1$ y $CD = 0$, ya que todos los términos coinciden.

Si $t \neq 0$ todos los valores que toma la variable son distintos ($m_{ji} = 1$ para todo j) y al tratarse de una única variable podemos eliminar el subíndice correspondiente a la misma, por tanto la expresión del coeficiente es:

$$CD = 2/(m.(m-1)).\sum_{j,l>j}^{m} \left|(X_j - X_l)/\mu\right| = \frac{2.\,2}{m.(m-1).\, x_1(2+(m-1).t)}[\,P(d)]$$

con $j = 1,\ldots,m$

con $[\,P(d)] = [d+2d+\ldots\ldots+(m-1)d+d+2d+\ldots\ldots+(m-2)d+\ldots\ldots\ldots+d+2d+d] > 0$

Sustituyendo $d = t.x_1$ y operando:

$$CD = \frac{4.t}{m.(m-1).(2+(m-1).t)}[\,(m-1)+2.(m-2)+3.(m-3)+\ldots\ldots+(m-1).(m-(m-1))] =$$

$$CD = \frac{4.t}{m.(m-1).(2+(m-1).t)}[\,m.(1+2+\ldots\ldots+(m-1)) - (1^2+2^2+\ldots.+(m-1)^2)] =$$

donde: $(1+2+\ldots\ldots+(m-1)) = (1+m-1).(m-1)/2 = m.(m-1)/2$

$(1^2+2^2+\ldots.+(m-1)^2) = (m-1).m.(2(m-1)+1)/6 = m.(m-1).(2m-1)/6$

y sustituyendo en la expresión anterior y extrayendo factor común m.(m-1)/2 queda:

$$CD = \frac{4.t.m.(m-1)}{2.m.(m-1).(2+(m-1).t)}[\,m-(2m-1)/3] = \frac{2.t}{(2+(m-1).t)}[\,m-(2m-1)/3]$$

$$\boxed{CD = \frac{2.t.(m+1)}{3.(2+(m-1).t)}}$$

Como casos particulares se obtienen:

a) Si t tiende a 0 como m finito $\Rightarrow$ CD tiende a 0 hecho lógico puesto que la variable tendería a tener el mismo valor (x_1) en todas las zonas.

b) Si t = 1 $\Rightarrow$ CD = 2/3.

c) Si t = 2 $\Rightarrow$ CD = 2(m+1)/3m.

d) Si t tiende a infinito $\Rightarrow$ CD tiende a 2(m+1)/3(m-1).
   d.1) Si m=2 $\Rightarrow$ CD tiende a 2, evidente ya que prácticamente todo el valor de la variable se acumula en uno de las dos zonas.
   d.2) Si m=3 $\Rightarrow$ CD tiende a 4/3.
   d.3) Si m=5 $\Rightarrow$ CD tiende a 1.
   Se observa que al crecer m, CD disminuye siendo su cota inferior 2/3.
   Por ejemplo, para m=1000, CD = 0,6680013 $\approx$ 2/3.

3. El C.D. de Ivanovic en variables distribuidas según una progresión geométrica.

Si suponemos que la variable se distribuye entre las $m > 1$ zonas geográficas según una progresión geométrica de razón $r \geq 0$ siendo $x_1 > 0$ el primer término de la progresión, tenemos:

$x_1 = x_1$
$x_2 = x_1.r$
....
$x_m = x_1.r^{m-1}$

Si $r = 1$ entonces $\mu = x_1$ y CD = 0, ya que todos los valores de la variable son iguales.

Si $r \neq 1$:

$$\mu = \frac{(x_1 + \ldots\ldots + x_m)}{m} = \frac{x_1.(r^m - 1)}{m.(r-1)}$$

que sustituidas en la expresión del coeficiente nos da:

$$CD = \frac{2.m.(r-1)}{x_1.(r^m-1).m.(m-1)} \left| \left[\sum_{j,l>j}^{m-1} (x_1.r^l - x_1.r^j) \right] \right| =$$

con $j = 0,....,(m-1)$

$$CD = \frac{2.m.(r-1).\ x_1}{x_1.(r^m-1).m.(m-1)} \left| [(m-1)r^{m-1} - r^{m-2} -...- 1 + (m-2)r^{m-2} - r^{m-3} -...-1 +......+ r - 1] \right|$$

y utilizando la expresión general de la suma de los términos de una progresión geométrica de razón r y operando:

$$= \frac{2.(r-1)}{(r^m-1).(m-1)} \left| [(m-1)r^{m-1} - (r^{m-1}-1)/(r-1) + (m-2)r^{m-2} - (r^{m-2}-1)/(r-1) +......+ r - (r-1)/(r-1)] \right|$$

expresión que se puede representar como:

$$CD = \frac{2.(r-1)}{(r^m-1).(m-1)} \left| \left[\sum_{j=1}^{m-1} j.r^j - \sum_{j=1}^{m-1} (r^j-1)/(r-1) \right] \right|$$

y simplificando por $r - 1 \neq 0$

$$\boxed{CD = \frac{2}{(r^m-1).(m-1)} \left| \left[\sum_{j=1}^{m-1} j.r^j(r-1) - \sum_{j=1}^{m-1} (r^j-1) \right] \right|}$$

Como casos particulares se obtienen:

a) si r tiende a infinito:

Expresando el CD en función de las potencias de r, ordenadas de mayor a menor:

$$CD = \frac{2}{(r^m-1).(m-1)} \left| [(m-1).r^{m-1}.r + P(r)] \right|$$

con P(r) polinomio de grado menor o igual que m, por tanto:

CD tiende a $\frac{2.(m-1)}{(m-1)}$ = 2, que es el máximo valor que puede alcanzar el coeficiente.

b) Si r tiende a 1:

$$CD = \lim_{r \to 1} \frac{2.(r-1)}{(r^m-1).(m-1)} \left| \left[\sum_{j=1}^{m-1} j.r^j - \sum_{j=1}^{m-1} (r^j-1)/(r-1) \right] \right| =$$

$$CD = \lim_{r \to 1} \frac{2}{(r^{m-1}+r^{m-2}+\ldots\ldots+1).(m-1)} \left| \left[\sum_{j=1}^{m-1} j.r^j - \sum_{j=1}^{m-1} (r^{j-1}+r^{j-2}+\ldots\ldots+1) \right] \right| =$$

$$CD = \frac{2}{m.(m-1)} \left| \left[\sum_{j=1}^{m-1} j - \sum_{j=1}^{m-1} j \right] \right| = 0$$, como cabía esperar al ser $x_1 \approx x_2 \approx \ldots\ldots \approx x_m$.

c) Si r tiende a 0:

$$CD = \lim_{r \to 0} \frac{2.(r-1)}{(r^m-1).(m-1)} \left| \left[\sum_{j=1}^{m-1} j.r^j - \sum_{j=1}^{m-1} (r^j-1)/(r-1) \right] \right| =$$

$$CD = \frac{2}{(m-1)} \left| \left[- \sum_{j=1}^{m-1} (+1) \right] \right| = \frac{2\,|-(m-1)|}{(m-1)} = 2$$

hecho lógico puesto que hay un valor distinto de cero y todos los demás prácticamente nulos.

4. Calculo del Coeficiente de Ivanovic para variables distribuidas según sucesiones recurrentes elementales

Basados en los resultados que hemos obtenido en los apartados anteriores se ha procedido a calcular el coeficiente de Ivanovic para variables distribuidas según distintas progresiones aritméticas y geométricas, tomando distintos valores para d y r, considerando en primer lugar el número de términos (zonas consideradas) igual a cincuenta, que es el número de provincias españolas, excluidas Ceuta y Melilla y posteriormente m = 17, que es el número de comunidades autónomas . Se han calculado también los totales relativos acumulados de los valores de la variable, ordenados de menor a mayor, y las frecuencias relativas acumuladas expresándose todo ello en %.

a) *Variables distribuidas según progresiones aritméticas:*

En la tabla 1 se muestran los resultados obtenidos con valores de CD aumentando de 0,125 en 0,125. La primera columna representa la frecuencias relativas acumuladas en %, escaladas de 10 en 10. El resto de columnas representan los totales acumulados en % de los valores de la variable, para los distintos valores del CD. Se muestran también las columnas correspondientes a:

CD = 2/3, que coincide con una progresión en la que $d = x_1$, es decir t = 1.

CD = 102/150, que se obtiene para t = 2.

CD = 102/147, que es el valor que se obtiene con una progresión aritmética de 50 términos, cuando t tiende a infinito.

En dicha tabla puede observarse como para un CD = 0 la discriminación es nula, ya que el porcentaje del total acumulado de la variable coincide con el porcentaje de las frecuencias relativas acumuladas, en cambio si nos desplazamos por la misma fila hacia la derecha la discriminación aumenta ya que el mismo porcentaje de frecuencias relativas tiene cada vez menor porcentaje de total acumulado, es decir, aumenta la desigualdad. Como curiosidad cabe comentar como para CD = 0,68 a un 50% de las frecuencias relativas acumuladas les corresponde el 25% de los valores acumulados de la variable, mientras que en el otro 50% se distribuye el 75% restante.

Tabla 1

CD	0,0000	0,1250	0,2500	0,3750	0,5000	0,6667 = 2/3	0,6750	0,6800 = 102/150	0,6939 = 102/147
F. relativas acumu.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.
10,00%	10,00%	8,35%	6,69%	5,04%	3,38%	1,18%	1,07%	1,00%	0,82%
20,00%	20,00%	17,06%	14,12%	11,18%	8,24%	4,31%	4,12%	4,00%	3,67%
30,00%	30,00%	26,14%	22,28%	18,42%	14,56%	9,41%	9,15%	9,00%	8,57%
40,00%	40,00%	35,59%	31,18%	26,76%	22,35%	16,47%	16,18%	16,00%	15,51%
50,00%	50,00%	45,40%	40,81%	36,21%	31,62%	25,49%	25,18%	25,00%	24,49%
60,00%	60,00%	55,59%	51,18%	46,76%	42,35%	36,47%	36,18%	36,00%	35,51%
70,00%	70,00%	66,14%	62,28%	58,42%	54,56%	49,41%	49,15%	49,00%	48,57%
80,00%	80,00%	77,06%	74,12%	71,18%	68,24%	64,31%	64,12%	64,00%	63,67%
90,00%	90,00%	88,35%	86,69%	85,04%	83,38%	81,18%	81,07%	81,00%	80,82%
100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

b) *Variables distribuidas según progresiones geométricas:*

En la tablas 2 y 3, construidas de forma análoga a la anterior, se muestran los resultados obtenidos con valores de CD aumentando de 0,125 en 0,125[1]. En este caso se puede recorrer todo el intervalo de variación del coeficiente de Ivanovic.

Tabla 2

CD	0,0000	0,1250	0,2500	0,3750	0,5000	0,6750	0,7500	1,0000
F. relativas acumu.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum
10,00%	10,00%	8,43%	7,00%	5,72%	4,58%	3,20%	2,69%	1,31%
20,00%	20,00%	17,17%	14,54%	12,12%	9,91%	7,16%	6,11%	3,17%
30,00%	30,00%	26,24%	22,66%	19,29%	16,12%	12,06%	10,46%	5,80%
40,00%	40,00%	35,65%	31,41%	27,31%	23,36%	18,12%	16,00%	9,54%
50,00%	50,00%	45,41%	40,83%	36,29%	31,79%	25,63%	23,05%	14,84%
60,00%	60,00%	55,54%	50,98%	46,33%	41,61%	34,91%	32,02%	22,36%
70,00%	70,00%	66,05%	61,90%	57,57%	53,05%	46,40%	43,44%	33,02%
80,00%	80,00%	76,95%	73,67%	70,16%	66,38%	60,62%	57,97%	48,14%
90,00%	90,00%	88,26%	86,35%	84,24%	81,91%	78,22%	76,46%	69,59%
100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

1. Por razones de espacio se ha omitido la columna correspondiente a C D = 0,8750

Tabla 3

CD 1,1250	1,1250	1,2500	1,3750	1,5000	1,6250	1,7500	1,8750	2,0000
F. relativas acumu.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum
10,00%	0,81%	0,44%	0,20%	0,06%	0,01%	0,00%	0,00%	0,00%
20,00%	2,05%	1,17%	0,56%	0,19%	0,03%	0,00%	0,00%	0,00%
30,00%	3,92%	2,38%	1,21%	0,46%	0,10%	0,00%	0,00%	0,00%
40,00%	6,76%	4,36%	2,42%	1,04%	0,27%	0,02%	0,00%	0,00%
50,00%	11,07%	7,63%	4,63%	2,26%	0,72%	0,09%	0,00%	0,00%
60,00%	17,61%	13,00%	8,68%	4,87%	1,95%	0,35%	0,00%	0,00%
70,00%	27,52%	21,86%	16,08%	10,39%	5,22%	1,44%	0,05%	0,00%
80,00%	42,57%	36,44%	29,65%	22,13%	13,96%	5,93%	0,66%	0,00%
90,00%	65,39%	60,45%	54,50%	47,06%	37,37%	24,35%	8,10%	0,00%
100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

En estas tablas también se aprecia como para un CD = 0 la discriminación es nula y si nos desplazamos por la misma fila hacia la derecha la discriminación aumenta, ya que el mismo porcentaje de frecuencias relativas tiene cada vez menor porcentaje del total acumulado de los valores de la variable, es decir aumenta la desigualdad. Observando detenidamente las tablas puede comprobarse que:

- Para un CD = 0,675 a un 50% de las frecuencias relativas acumuladas les corresponde casi el 25% de los valores acumulados de la variable, mientras que en el otro 50% se distribuye el 75% restante.

- Para un CD = 1 a un 50% de las frecuencias relativas acumuladas les corresponde aproximadamente el 15% de los valores acumulados de la variable, a un 80% menos del 50% y el último tramo (10% de las zonas consideradas) se atribuye más del 30% de los valores de la variable.

- Para CD = 1,5 esta desigualdad es mucho más acusada puesto que más del 50% de los valores de la variable se concentran en un 10% de zonas y además el 50% de zonas no alcanza ni siquiera el 2,5% de los valores acumulados de la variable.

A continuación se presentan los resultados que se obtienen considerando 17 zonas (número de comunidades autónomas españolas) y siguiendo un procedimiento análogo al desarrollado en los párrafos anteriores. Se han calculado las frecuencias

ÍNDICE

relativas acumuladas para todas las zonas, es por ello por lo que las tablas presentan 17 filas de datos.

En la tabla 4 se muestran los valores obtenidos utilizando progresiones aritméticas. Como en el caso de las provincias también se presentan las columnas correspondientes a los casos particulares t = 1 (CD = 2/3); t = 2 (CD = 36/51) y cuando t tiende a infinito (CD = 36/48).

Tabla 4

CD	0,0000	0,1250	0,2500	0,3750	0,5000	0,6667 = 2/3	0,6750	0,7059 = 36/51	0,7500 = 36/48
F. relativas acumu.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.
5,88%	5,88%	4,90%	3,92%	2,94%	1,96%	0,65%	0,59%	0,35%	0,00%
11,76%	11,76%	9,93%	8,09%	6,25%	4,41%	1,96%	1,84%	1,38%	0,74%
17,65%	17,65%	15,07%	12,50%	9,93%	7,35%	3,92%	3,75%	3,11%	2,21%
23,53%	23,53%	20,34%	17,16%	13,97%	10,78%	6,54%	6,32%	5,54%	4,41%
29,41%	29,41%	25,74%	22,06%	18,38%	14,71%	9,80%	9,56%	8,65%	7,35%
35,29%	35,29%	31,25%	27,21%	23,16%	19,12%	13,73%	13,46%	12,46%	11,03%
41,18%	41,18%	36,89%	32,60%	28,31%	24,02%	18,30%	18,01%	16,96%	15,44%
47,06%	47,06%	42,65%	38,24%	33,82%	29,41%	23,53%	23,24%	22,15%	20,59%
52,94%	52,94%	48,53%	44,12%	39,71%	35,29%	29,41%	29,12%	28,03%	26,47%
58,82%	58,82%	54,53%	50,25%	45,96%	41,67%	35,95%	35,66%	34,60%	33,09%
64,71%	64,71%	60,66%	56,62%	52,57%	48,53%	43,14%	42,87%	41,87%	40,44%
70,59%	70,59%	66,91%	63,24%	59,56%	55,88%	50,98%	50,74%	49,83%	48,53%
76,47%	76,47%	73,28%	70,10%	66,91%	63,73%	59,48%	59,26%	58,48%	57,35%
82,35%	82,35%	79,78%	77,21%	74,63%	72,06%	68,63%	68,46%	67,82%	66,91%
88,24%	88,24%	86,40%	84,56%	82,72%	80,88%	78,43%	78,31%	77,85%	77,21%
94,12%	94,12%	93,14%	92,16%	91,18%	90,20%	88,89%	88,82%	88,58%	88,24%
100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

A la vista de los resultados obtenidos, comentarios análogos a los efectuados anteriormente son igualmente válidos.

Las tablas 5 y 6 muestran los valores obtenidos para m = 17 utilizando progresiones geométricas.

ÍNDICE

Tabla 5

CD	0,0000	0,1250	0,2500	0,3750	0,5000	0,6750	0,7500	1,0000
F. relativas acumu. acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum	Total acum
5,88%	5,88%	4,95%	4,12%	3,37%	2,70%	1,90%	1,61%	0,81%
11,76%	11,76%	10,01%	8,41%	6,96%	5,65%	4,05%	3,45%	1,79%
17,65%	17,65%	15,17%	12,88%	10,78%	8,86%	6,47%	5,55%	2,99%
23,53%	23,53%	20,44%	17,55%	14,86%	12,36%	9,20%	7,97%	4,44%
29,41%	29,41%	25,83%	22,42%	19,20%	16,17%	12,28%	10,73%	6,21%
35,29%	35,29%	31,32%	27,50%	23,83%	20,33%	15,75%	13,90%	8,36%
41,18%	41,18%	36,93%	32,79%	28,76%	24,87%	19,66%	17,53%	10,97%
47,06%	47,06%	42,66%	38,31%	34,02%	29,81%	24,07%	21,68%	14,15%
52,94%	52,94%	48,52%	44,07%	39,63%	35,20%	29,05%	26,44%	18,01%
58,82%	58,82%	54,49%	50,08%	45,60%	41,07%	34,66%	31,89%	22,70%
64,71%	64,71%	60,59%	56,34%	51,97%	47,47%	40,98%	38,13%	28,40%
70,59%	70,59%	66,82%	62,88%	58,76%	54,45%	48,12%	45,28%	35,33%
76,47%	76,47%	73,18%	69,69%	65,99%	62,06%	56,16%	53,47%	43,76%
82,35%	82,35%	79,68%	76,80%	73,70%	70,36%	65,23%	62,85%	54,01%
88,24%	88,24%	86,31%	84,21%	81,91%	79,40%	75,46%	73,60%	66,46%
94,12%	94,12%	93,08%	91,94%	90,67%	89,25%	86,99%	85,90%	81,60%
100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tabla 6

CD	1,1250	1,2500	1,3750	1,5000	1,6250	1,7500	1,8750	2,0000
F. relativas acumu.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum.	Total acum
5,88%	0,52%	0,30%	0,14%	0,05%	0,01%	0,00%	0,00%	0,00%
11,76%	1,17%	0,69%	0,35%	0,13%	0,03%	0,00%	0,00%	0,00%
17,65%	2,00%	1,21%	0,63%	0,25%	0,06%	0,01%	0,00%	0,00%
23,53%	3,04%	1,89%	1,01%	0,43%	0,11%	0,01%	0,00%	0,00%
29,41%	4,35%	2,78%	1,56%	0,69%	0,20%	0,02%	0,00%	0,00%
35,29%	6,00%	3,96%	2,31%	1,09%	0,35%	0,05%	0,00%	0,00%
41,18%	8,08%	5,51%	3,35%	1,68%	0,59%	0,10%	0,00%	0,00%
47,06%	10,70%	7,56%	4,81%	2,57%	1,00%	0,19%	0,01%	0,00%
52,94%	14,01%	10,25%	6,83%	3,89%	1,67%	0,39%	0,02%	0,00%
58,82%	18,18%	13,79%	9,63%	5,88%	2,80%	0,78%	0,05%	0,00%
64,71%	23,44%	18,45%	13,54%	8,84%	4,67%	1,56%	0,14%	0,00%
70,59%	30,06%	24,59%	18,96%	13,28%	7,79%	3,12%	0,41%	0,00%
76,47%	38,41%	32,67%	26,50%	19,91%	12,99%	6,25%	1,23%	0,00%
82,35%	48,93%	43,30%	36,99%	29,82%	21,64%	12,50%	3,70%	0,00%
88,24%	62,20%	57,30%	51,57%	44,65%	36,05%	25,00%	11,11%	0,00%
94,12%	78,92%	75,74%	71,83%	66,83%	60,04%	50,00%	33,33%	0,00%
100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

ÍNDICE

Puede comprobarse como el comportamiento es muy semejante al obtenido en las tablas 2 y 3 teniendo en cuenta que los valores no son directamente comparables, puesto que los tramos considerados por filas no coinciden pero entendemos que no presentan dificultad para analizar la secuencia de la distribución de la variable según los valores crecientes del coeficiente de discriminación.

Si en lugar de utilizar progresiones aritméticas o geométricas como patrón se emplean otro tipo de sucesiones recurrentes monótonas, los resultados que se obtienen son muy parecidos. Del mismo modo haciendo variar el número de zonas observadas se consiguen análogos resultados. Razones de espacio nos aconsejan a no reiterar la exposición de las tablas obtenidas. Únicamente a título de ejemplo presentamos los valores obtenidos para $m = 50$ si la variable se distribuye siguiendo la sucesión de Fibonacci: $x_1 = x_2 = 1$; $x_{n+1} = x_n + x_{n-1}$ para $n \geq 2$. A efectos comparativos se incluye la columna correspondiente a CD = 1,8750 de la tabla 3.

Tabla 7

CD	1,8679	1,8750
Frecuencias relativas acumuladas	Total acumulado con variable distribuida según sucesión de Fibonacci	Total acumulado con variable distribuida según una progresión geométrica
10,00%	0,00%	0,00%
20,00%	0,00%	0,00%
30,00%	0,00%	0,00%
40,00%	0,00%	0,00%
50,00%	0,00%	0,00%
60,00%	0,01%	0,00%
70,00%	0,07%	0,05%
80,00%	0,81%	0,66%
90,00%	9,02%	8,10%
100,00%	100,00%	100,00%

Puede verse como los resultados obtenidos son muy semejantes, apreciándose un alto grado de concentración de la variable ya que un 10% de zonas (último tramo) representa más del 90% del valor de la variable.

5. Una posible escala de graduación del Coeficiente de Ivanovic

Como resumen de los resultados expuestos proponemos una escala, que obviamente no es la única que se puede elaborar pero que entendemos sirve para resolver el problema que planteábamos al principio de la exposición y que no es otro que el clasificar el grado de discri minación de una variable socio-económica medido por el coeficiente de Ivanovic. Para la construcción de la escala, que en definitiva es establecer una jerarquía, se han tenido en cuenta las siguientes consideraciones:

a) El número de tramos que se ha escogido para realizar la escala es cinco, análogo al empleado en el Proceso Analítico Jerárquico (A.H.P.), ver Saaty (1.988), que es una de las técnicas utilizadas en Teoría de la Decisión.
b) Se ha procedido a comparar el valor del CD con el nivel de concentración de los valores de la variable, considerando un poder de discriminación débil si aproximadamente el primer 50% de las zonas, ordenadas de menor a mayor según los valores de la variable, acumula más del 40% de los mismos, moderado si acumula entre un 25 y un 40%, fuerte entre un 15 y un 25%, muy fuerte entre un 2 y un 15% y extremadamente fuerte si acumulaba menos de un 2%.

Tomando como patrones las tablas incluidas en los apartados anteriores y teniendo en cuenta las consideraciones anteriormente mencionadas, presentamos la escala que aparece en la tabla 8, como instrumento para clasificar el poder de discriminación de una variable socio-económica.

Tabla 8

Valor de CD	Poder de discriminación
Menor que 1/4	Débil
Entre 1/4 y 2/3	Moderado
Entre 2/3 y 1	Fuerte
Entre 1 y 3/2	Muy fuerte
Mayor que 3/2	Extremadamente fuerte

6. Aplicación práctica de la escala propuesta

Presentamos en este apartado dos ejemplos en los que se aplica la escala propuesta anteriormente. El primero de ellos hace referencia a los Indices de Capacidad de Compra para 1.995 (ICC95) correspondientes a bienes de tipo 2 propuestos por el autor en Miguel Uceta (1996). En el mencionado trabajo se calculaba el Coeficiente de Ivanovic para los ICC95, obteniéndose un valor de 0,9816 si se tomaban como zonas geográficas las provincias y de 1,0700 si las zonas consideradas eran las comunidades autónomas. Vemos que ambos valores están próximos a uno y por tanto el poder de discriminación puede considerarse como fuerte, aunque en el caso del obtenido para las comunidades autónomas puede incluso considerarse como muy fuerte, según la escala anterior. Hemos de hacer notar que este ejemplo corrobora la afirmación expuesta por Zarzosa Espina (1994) "Una variable, con un índice de concentración dado, es más o menos discriminante en el conjunto de m zonas en función del valor de m: A medida que m aumenta, el poder discriminante de la variable disminuye".

El segundo ejemplo se refiere a la variable que mide el porcentaje de nacimientos sobre el total de la población. El estudio se ha realizado con los datos recogidos en el Anuario Estadístico del I.N.E. de 1993 y se refieren a la población de hecho por provincias según el Censo de 1991 y los nacimientos registrados el mismo año. En la tabla 9 se muestra el CD para esta variable así como la concentración por tramos de la misma. A la vista del resultado obtenido CD = 0,2320, aplicando la escala propuesta, podemos concluir que el poder de discriminación es débil, como cabía esperar por la propia definición de la variable. A efectos meramente comparativos se incluye también la columna correspondiente a CD = 0,2500, de la tabla 2. Puede observarse como las cifras son muy semejantes y están en línea con los comentarios realizados a lo largo del presente trabajo.

Tabla 9

CD	0,2320	0,2500
F. relativas acumuladas	Total acumulado	Total acumulado
10,00%	7,35%	7,00%
20,00%	15,37%	14,54%
30,00%	23,64%	22,66%
40,00%	32,15%	31,41%
50,00%	41,42%	40,83%
60,00%	51,32%	50,98%
70,00%	62,06%	61,90%
80,00%	73,79%	73,67%
90,00%	86,58%	86,35%
100,00%	100,00%	100,00%

7. Conclusiones

En el presente artículo se han calculado las expresiones para la obtención del coeficiente de Ivanovic para variables socio-económicas cuya distribución por zonas geográficas coincida con los términos de una progresión aritmética o una progresión geométrica de términos positivos. Estudiándose algunos casos particulares.

En el caso de las progresiones aritméticas se ha encontrado una expresión que reduce notablemente el cálculo del mencionado coeficiente:

$$CD = \frac{2.t.(m+1)}{3.(2+(m-1).t)} \quad ; \; t = d/x_1 \; ; \; d = \text{diferencia de la progresión} \; ;$$

x_1 = primer término

m = número de zonas geográficas

Se ha propuesto una escala, que obviamente no es la única que puede construirse, para poder clasificar una variable socio-económica según su poder de discriminación en zonas geográficas.

Finalmente se ha utilizado dicha escala para la clasificación de algunas variables concretas.

Bibliografía

CHASCO YRIGOYEN C. y MIGUEL UCETA S. (1997): Análisis de las áreas comerciales de Andalucía. Actas del I Congreso de Ciencia Regional de Andalucía. Jerez de la Frontera.

CHASCO YRIGOYEN C. y MIGUEL UCETA S. (1997): Indicadores socio-económicos en el Anuario Comercial de España. Actas de la XXIII Reunión de Estudios Regionales. Valencia.

I.N.E. (1995): España. Anuario Estadístico. Madrid

IVANOVIC, B. (1974): Comment établir une liste des indicateurs de developpment. Revue de Statistique Apliquée. Vol. XXII, nº 2, París.

MIGUEL UCETA S. (1996): El Mercado Potencial. su delimitación y cuantificación. Actas de la X reunión de ASEPELT-ESPAÑA. Albacete.

PENA TRAPERO J.B. (1977): Problemas de la medición del bienestar y conceptos afines. (Una aplicación al caso español). Madrid, Instituto Nacional de Estadística.

SAATY, T.L. (1988) The Analitic Hierarchy Process. RSW Publications.

ZARZOSA ESPINA P. (1994): El criterio de discriminación en la selección de indicadores de bienestar. Análisis del coeficiente de discriminación de Ivanovic. Estudios de Economía Aplicada, 2, pág. 169-185.

ZARZOSA ESPINA P. (1996): Aproximación a la medición del bienestar social. Valladolid. Secretariado de publicaciones. Universidad, págs. 158 y sgtes.