

Desigualdades educativas en la España del siglo XX. El caso de Castilla y León¹.

El presente artículo contiene dos partes diferenciadas: en la primera se articula una propuesta sobre la medición de la variación, que ya apareciera en un anterior artículo (Escobar, 1998), aunque en este caso se concentrará en la desigualdad educativa y se planteará como novedad un modo de representación gráfica del promedio, desviación y simetría de las variables con valores limitados, como es el caso de los años de educación formal que reciben los ciudadanos de un país. En la segunda parte, con una muestra, extraída del último censo de población (1991), de más de 2.000.000 de españoles con más de 25 años, se realiza una descripción y diagnóstico de la evolución de las desigualdades educativas en nuestro país en el presente siglo, fijando la atención en los datos de Castilla y León y en las diferencias de género.

Los estudios empíricos sobre la diversidad son uno de los aspectos más controvertidos, entre los que en mi opinión el más importante afecta a la medición de la desigualdad, polémica inacabada que sigue suscitando el interés de economistas y sociólogos. Este tema es central en el estudio de la Sociedad, como demuestra el que desde las concepciones jerárquicas de Platón al discurso sobre el origen de la desigualdad de Rousseau ya existiera la preocupación por la diversidad en el pensamiento presociológico. Posteriormente en los clásicos la preocupación por el estudio de la divisiones en la Sociedad es obvio y puede ser sintetizado con los conceptos de clase en el caso de Marx, status en el de Weber y solidaridad mecánica/orgánica en el caso de la obra de Durkheim. Nisbet (1966) señala el status como una de las cinco ideas-clave de la tradición sociológica. Las similitudes y desigualdades dentro y entre sociedades deben ser estudiadas por los sociólogos y metodológicamente hay un reto importante en proporcionar herramientas que permitan medir lo diverso en la sociedad, pues así como hay distintos modos de calibrar los promedios, las medidas de la diversidad son mucho más numerosas y de mucha más dificultad en su interpretación.

¹ Para la elaboración de este artículo se ha contado con la inestimable colaboración de José Martínez, quien se ha encargado de una parte importante del análisis de datos. En cualquier caso, sólo al autor son atribuibles los errores presentes en el presente texto.

Pero si importante para el estudio de la sociedad es encontrar instrumentos que nos permitan el estudio de la diversidad, no lo es menos desde un punto de vista estrictamente metodológico²: a menudo se confunde la estadística con la ciencia de los porcentajes y las medias, porque con mucha frecuencia sus estudios divulgativos sólo emplean estos estadísticos para contribuir a la sencillez y comprensibilidad de lo que se quiere presentar. Sin embargo, tan importante para dar una correcta imagen de la sociedad es ofrecer los promedios de las variables como la forma en como están distribuidas. Y, de modo similar, la estadística avanzada está construida básicamente a partir del concepto de varianza. En general, la mayor parte de las técnicas que sirven para la explicación causal entre variables utilizan la expresión “porcentaje de varianza explicado”

Para poder mantener las tesis sobre desigualdades educativas que aparecen en la segunda parte de este artículo, se explica un coeficiente de variación relativa que, a nuestro entender, supera las limitaciones que tiene el de Pearson; aunque sólo pueda aplicarse a variables acotadas, es decir, aquellas que tienen valores mínimos y máximos definidos.

Por todos es bien conocido que una de las medidas más utilizadas en análisis complejos de la estadística es la varianza, que es el promedio de las desviaciones al cuadrado de los valores de una variable con relación a la media.

Son propiedades de la varianza las siguientes: a) siempre es un número positivo, b) se puede obtener mediante la sustracción entre los cuadrados de la media cuadrática y la media aritmética, c) da un valor nulo cuando se aplica a una constante, d) al añadir una constante a una determinada variable (imaginemos que a todos los ciudadanos se le da (o quita) un año más de educación), la varianza no sufre cambios, y e) al multiplicarla por una constante (en el caso de que a todos se de o quite un determinado porcentaje de educación sobre el que ya se posee), la nueva varianza se ve multiplicada por dicha constante al cuadrado.

Dos características sobresalen en este estadístico: en primer lugar, que la varianza crece considerablemente cuando una variable tiene valores muy alejados de la media y, en segundo lugar, que las unidades en la que está expresada esta magnitud son unidades cuadradas de los valores originales de la variable. Por esta razón, resulta de interés reducir el estadístico a unidades reales, no cuadráticas, para lo que se emplea la desviación típica, que se obtiene con la mera raíz cuadrada de la varianza.

² Ejemplos de la preocupación por la medición de la diversidad lo encontramos en Atkinson (1970), Blalock (1991), Cortés y Rubalcaba (1984), Coulter (1984 y 1989), Esteben y Ray (1993), Osberg (1991) y Theil (1967).

No obstante, siguen planteándose dos problemas: primero, resulta difícil precisar si la variable está bien o mal distribuida o, dicho de otra manera, se carece de elementos de comparación para evaluar el grado en el que se dispersan las variables en estudio —¿es mucho que la desviación típica de una determinada variable sea igual a 5?—; y, segundo, resulta imposible poder comparar la dispersión entre dos variables medidas en magnitudes distintas —si además de estudiar la educación, se estudiara una determinada opinión y ésta tuviese una desviación de 3, ¿se podría afirmar que esta segunda variable está menos dispersa que la primera?

Este problema se solventa con las medidas relativas —o normalizadas según Weisberg (1986)— de variación. Son medidas relativas aquellas que carecen de unidades de medición. Una de las más conocidas es el llamado coeficiente de variación de Pearson, que se obtiene mediante el cociente entre la desviación típica y la media. Véase un ejemplo de su aplicación. Ante la pregunta de qué variable presenta mayor variación en una población: la edad o el número de hijos, el problema es que ambas están medidas con distintas escalas. Las unidades de la primera son años, las de la segunda personas. Tanto las medias como las varianzas —y claramente también las desviaciones típicas— serían incomparables. Pero si se calculan los coeficientes de variación las magnitudes resultantes podrían compararse.

Entre las propiedades de este coeficiente de variación, pueden considerarse en este contexto las siguientes: a) adopta el valor cero cuando la variable es constante, b) podría adoptar valores negativos en el caso de que la media lo fuera, c) cuando los valores de la media son muy bajos, el valor del coeficiente crece de modo excesivo, d) cuando se suma una constante a la variable (todos reciben la misma cantidad del valor en cuestión), el CV disminuye, e) al aumentar (o disminuir) en un determinado porcentaje las cantidades, el CV no experimenta cambio alguno, y f) el valor máximo que puede adoptar este coeficiente en el caso de la variable tenga valores positivos es $\sqrt{n-1}$.

Por estas características, es conveniente no utilizar el coeficiente de variación cuando la variable tiene valores negativos o cuando la media es próxima a cero. En realidad, este coeficiente tiene como supuesto de uso el que la variable de trabajo se encuentre semiacotada en el límite inferior (con el valor 0), en cuyo caso la situación de mayor variación sería aquella en la que todos los sujetos menos 1 tienen el valor 0 y el sujeto restante tiene el valor $n \times \bar{x}$.

Una primera medida que se puede adoptar para evitar que este cociente deje de tener el inconveniente de presentar unas cotas distintas según la distribución sería la de dividirlo por $\sqrt{n-1}$, que es su valor máximo. De esta manera, el coeficiente de variación ajustado presentaría la siguiente fórmula:

$$CV_{aj} = \frac{s}{x\sqrt{n-1}} \quad (1.)$$

Aun con todo, adolece esta medida de un supuesto muy débil, sobre todo, cuando el n es alto y el valor que puede adoptar una variable es limitado. Un claro ejemplo precisará mejor lo anteriormente dicho: en una encuesta con 1000 sujetos, con una edad media de 35 años y una desviación típica de 20 años; para el cálculo del coeficiente de variación máximo estaríamos suponiendo que habría 999 sujetos de 0 años y un sujeto con 999×35 años, es decir, 3465, edad obviamente inimaginable de que pueda ser cumplida por persona alguna.

Otra medida de variación relativa sería resultado de la división de la desviación típica entre la máxima posible con la condición de que la variable esté acotada por ambos lados, o dicho de otra forma, en el supuesto de que se conociera el rango de la distribución, en cuyo caso la situación más desfavorable sería aquella en la que el cincuenta por ciento de los casos tuviera el valor mínimo de la distribución y el otro cincuenta por ciento el valor máximo (Weisberg; 53). La media en estos casos sería:

$$\bar{X} = \frac{X_{\max} - X_{\min}}{2} \quad (2.)$$

Y la varianza correspondiente a la máxima polarización se ajustaría a la siguiente expresión:

$$Var_{\max} = \frac{(X_{\max} - X_{\min})^2}{4} \quad (3.)$$

Sin embargo, este cálculo de la desviación máxima adolece de un defecto: siempre supone la media constante, aun cuando en la distribución se presente una media distinta del punto equidistante entre los valores máximos y mínimos.

Por ello, también la medida recién mencionada tiende a sobrevalorar la varianza máxima de las distribuciones. La propuesta de este artículo es la de una varianza máxima que calculándose sobre valores acotados tenga en cuenta como dato la media de la muestra, es decir, se aboga por un coeficiente que reúna las dos características deseables de las anteriores medidas.

Ante una variable acotada con más de dos valores, la situación de mayor variación correspondería a aquella en la que sólo hubiera precisamente dos valores coincidiendo con los extremos.

Sabiendo que esta variable tiene una media \bar{X} , puede deducirse fácilmente las proporciones de los casos de los valores mínimo y máximo:

$$\left. \begin{aligned} p_{\min} + p_{\max} &= 1 \\ \bar{X} &= p_{\min} X_{\min} + p_{\max} X_{\max} \end{aligned} \right\} \quad (4.)$$

Y con las proporciones y los valores conocidos ya puede calcularse la varianza máxima:

$$Var_{\max} = (X_{\max} - \bar{X})^2 p_{\max} + (X_{\min} - \bar{X})^2 p_{\min} \quad (5.)$$

Fórmula que, empleando el álgebra básica, puede simplificarse del siguiente modo:

$$Var_{\max} = (X_{\max} - \bar{X})(\bar{X} - X_{\min}) \quad (6.)$$

Conocida la varianza máxima, es fácil obtener también la máxima desviación típica, obteniendo la raíz cuadrada de la primera. A partir de aquí puede calcularse un par de medidas interrelacionadas: Una proporción de varianza acotada, que sea el cociente entre la varianza empírica y la máxima condicionada a la media empírica y un coeficiente de variación acotado, que sea la razón entre las respectivas desviaciones típicas. Sus fórmulas serían las siguientes:

$$PV_a = \frac{s^2}{(\bar{X} - X_{\min})(X_{\max} - \bar{X})} \quad (7.)$$

$$CV_a = \frac{s}{\sqrt{(\bar{X} - X_{\min})(X_{\max} - \bar{X})}} = \sqrt{PV_a}$$

Ambas medidas presentan la deseable característica de poseer un rango entre 0 y 1. El valor mínimo se encuentra en 0 cuando la variable tiene una varianza de 0 y el valor máximo en 1 cuando la variable sólo presenta dos valores y éstos coinciden con los extremos de la distribución. Además, pueden obtenerse la una a partir de la otra como fácilmente puede comprobarse en la última fórmula.

Una interesante aplicación de estas medidas, especialmente la segunda por presentar generalmente valores mayores es la comparación de la variabilidad entre medidas con escalas distintas, al igual que ocurría con el coeficiente de variación. Sin embargo, a diferencia de éste, que puede tomar valores superiores al 100%, esta nueva medida siempre estará comprendida entre 0 y 1, o si se quiere, en términos porcentuales, entre 0 y 100.

Además, poseen como interesante característica el hecho de que una transformación lineal de la variable no altera este coeficiente, propiedad que no cumplen el resto de medidas de dispersión aplicables a variables numéricas. Quiere ello decir que si se suma o se multiplica a una determinada variable por una constante, el coeficiente obtenido con la variable transformada debe ser idéntico al calculado con la variable inicial. Ello no ocurre con la desviación típica (o varianza), pues se ve alterada al multiplicar la distribución por una constante, ni con el coeficiente de variación, que sufre modificación al añadirle una determinada constante. Esta propiedad es especialmente interesante en el campo de las ciencias sociales donde la inmensa mayoría de las escalas son arbitrarias -¿por qué medimos la ideología de 1 a 10 y no de 0 a 20, por ejemplo?, por lo que necesitamos medidas que resistan las decisiones sobre los valores de la escala. El caso más claro sería el de las escalas de Likert —normalmente medidas de 1 a 5— en las que es frecuente invertir los items utilizando la fórmula $X_i' = 6 - X_i$. Si se utiliza el coeficiente de variación, el resultado es distinto según utilicemos un invertido o no.

A continuación se expone en la tabla 1 ocho ejemplos de distribuciones simuladas de ideología con sus correspondientes medias, desviaciones típicas, porcentajes de varianza y coeficientes de variación ajustados. En los dos primeros lugares, tendríamos la situación en la que todos los sujetos de nuestro estudio tuviesen la misma ideología (4 en la distribución A y 3 en la distribución B), en ambos casos la varianza sería 0 y por tanto, aunque la desviación máxima —es decir, el denominador de la última fórmula, fuese de 9 u 8, el PV_a seguirá siendo de 0. Ahora bien, si la media de la variable es 4 y los sujetos están ubicados —sólo y necesariamente— a partes iguales en la extrema izquierda y en la extrema derecha; entonces la varianza es máxima y, por tanto, nuestro coeficiente es del 100%. Una segunda situación que debiera arrojar la misma cantidad de porcentaje de desviación empírica sería aquella en la que no siendo la media igual a 4, sólo existen valores extremos (Distribución D). En tal caso también el porcentaje de variación es del 100%: todos los sujetos están en los extremos de la distribución.

Situaciones menos extremas se presentan en las cuatro siguientes distribuciones: En la E y en la F, se han generado sendas tablas en las que se han situado el 50% de los sujetos en la media o en el valor extremo inferior. En esos casos, el porcentaje de variación es del 50% y del 75% respectivamente. En cambio, si además de los casos extremos, existen valores intermedios, este nuevo coeficiente disminuirá en proporción siempre inferior al número de sujetos que no tengan valores extremos. Comparándola con la distribución E en la distribución G se contempla cómo, habiendo salido de los extremos un 46% de casos, el coeficiente sólo disminuye 21 puntos porcentuales y del mismo modo la distribución H tiene un 25% de variación y sin embargo sólo un 50% de los casos están ubicados en la media y un 18% tienen valores extremos.

Tabla 1.- Distribuciones simuladas de ideología

Distribución A	x_i	f_i	$x_i f_i$	$(x_i - \mu)^2 f_i$	Distribución B	x_i	f_i	$x_i f_i$	$(x_i - \mu)^2 f_i$
Centro	4	1.00	4.00	0	Cent.-izquierda	3	1.00	3.00	0.00
			$\mu = 4.00$	$s^2 = 0.00$				$\mu = 3.00$	$s^2 = 0.00$
$PV_a =$	0%				$PV_a =$	0%			
$CV_a =$	0%				$CV_a =$	0%			
Distribución C	x_i	f_i	$x_i f_i$	$(x_i - \mu)^2 f_i$	Distribución D	x_i	f_i	$x_i f_i$	$(x_i - \mu)^2 f_i$
Ext. Izquierda.	1	0.50	0.50	4.5	Ext. Izquierda.	1	0.67	0.67	2.67
Ext. Derecha	7	0.50	3.50	4.5	Ext. Derecha	7	0.33	2.33	5.33
			$\mu = 4.00$	$s^2 = 9.00$				$\mu = 3.00$	$s^2 = 8.00$
$PV_a =$	100%				$PV_a =$	100%			
$CV_a =$	100%				$CV_a =$	100%			
Distribución E	x_i	f_i	$x_i f_i$	$(x_i - \mu)^2 f_i$	Distribución F	x_i	f_i	$x_i f_i$	$(x_i - \mu)^2 f_i$
Ext. Izquierda.	1	0.25	0.25	2.25	Ext. Izquierda.	1	0.50	0.50	2.00
Centro	4	0.50	2.00	0	Cent.-izquierda	3	0.25	0.75	0.00
Ext. Derecha	7	0.25	1.75	2.25	Ext. Derecha	7	0.25	1.75	4.00
			$\mu = 4.00$	$s^2 = 4.50$				$\mu = 3.00$	$s^2 = 6.00$
$PV_a =$	50%				$PV_a =$	75%			
$CV_a =$	71%				$CV_a =$	87%			
Distribución G	x_i	f_i	$x_i f_i$	$(x_i - \mu)^2 f_i$	Distribución H	x_i	f_i	$x_i f_i$	$(x_i - \mu)^2 f_i$
Ext. Izquierda.	1	0.02	0.02	0.22	Ext. Izquierda.	1	0.06	0.06	0.24
Cent.-izquierda	2	0.23	0.45	0.90	Izquierda	2	0.31	0.62	0.31
Centro	4	0.50	2.00	0.00	Cent.-izquierda	3	0.50	1.50	0.00
Cent.-derecha	6	0.23	1.35	0.90	Centro	6	0.09	0.54	0.81
Ext. Derecha	7	0.02	0.17	0.22	Ext. Derecha	7	0.04	0.28	0.64
			$\mu = 4.00$	$s^2 = 2.25$				$\mu = 3.00$	$s^2 = 2.00$
$PV_a =$	25%				$PV_a =$	25%			
$CV_a =$	50%				$CV_a =$	50%			

Para representar estas variables con límite, tanto su promedio, como su variación y su simetría, se propone el siguiente método:

1) Utilización de un diagrama de dos dimensiones en el que el eje de abscisas represente las distintas variables (comparaciones de muestras dependientes) o los distintos grupos (comparaciones de muestras independientes) y en el que el de ordenadas contemple todo el rango posible de las variables limitadas que se desean representar. Habrá en este eje, por tanto, tantas escalas como rangos distintos tengan las variables que se deseen representar.

2) Empleo de un símbolo que represente la media de la variable del que saldrá una línea de longitud (h) proporcional al porcentaje del rango total posible de la variable que represente el CV_a , esto es

$$h = CV_a (X_{\max} - X_{\min}) \tag{8.}$$

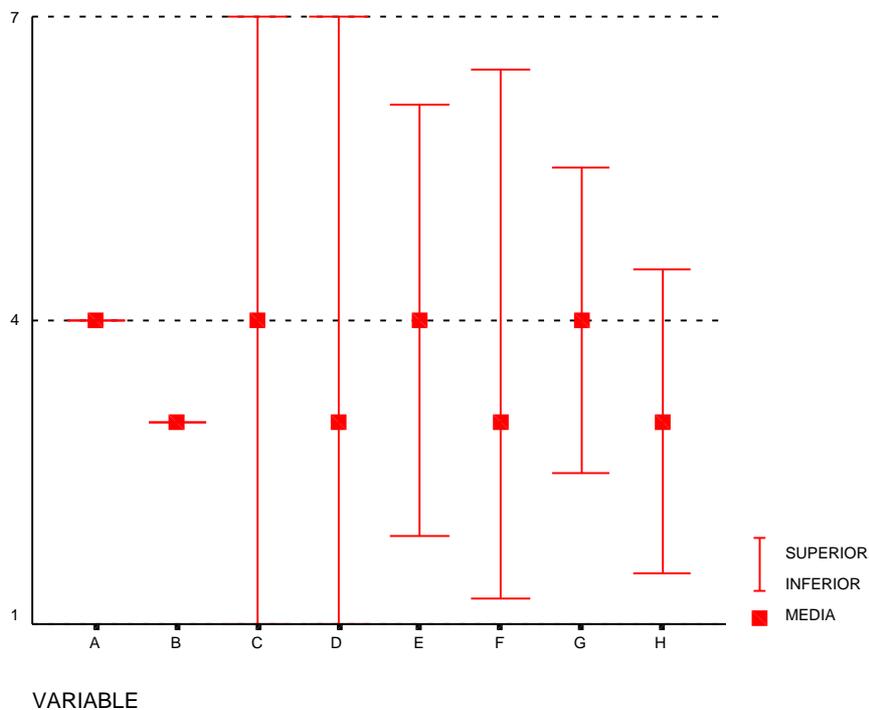
3) Esta línea de longitud h será simétrica con la media, salvo que llegue al límite de la variable. En este caso se desplazará en toda su longitud hacia el extremo opuesto lo suficiente para que no sobrepase los límites de la escala. Y, en el caso de que la mediana no coincida con la media aritmética, se propone que los límites superior e inferior de la línea vertical vengan dados, siempre y cuando no excedan los límites naturales de la variable, por la siguiente expresión, :

$$l_i = \bar{X} - CV_a \frac{(X_{\max} - X_{\min})}{2} \left(1 - \frac{2(\bar{X} - Med)}{(X_{\max} - X_{\min})} \right) \tag{9.}$$

$$l_{i+1} = \bar{X} + CV_a \frac{(X_{\max} - X_{\min})}{2} \left(1 + \frac{2(\bar{X} - Med)}{(X_{\max} - X_{\min})} \right)$$

Como demostración de esta técnica de representación de las variables con límite, se expone a continuación el gráfico correspondiente a las siete distribuciones ficticias de la ideología expuestas en la tabla 1.

Gráfico 1.- Gráficos de siete distribuciones ficticias de la ideología



Desigualdades educativas

Aunque las posturas radicales aboguen por igualdad de resultados, la igualdad de oportunidades se concibe como la posesión de medios externos que garanticen el pleno desarrollo de las capacidades de todas las personas que integran una sociedad. Uno de los medios más críticos para el logro del bienestar económico y mental de las ciudadanas y ciudadanos es la educación. Por lo que, en consecuencia, las políticas educativas son esenciales para el cumplimiento de los valores propuestos por los defensores de la igualdad de oportunidades. El argumento es el siguiente: si se da a todos los sujetos la educación máxima que permitan sus capacidades intelectuales, entonces se estará ante la sociedad más justa posible.

Dos son las posibles objeciones que se pueden hacer a este planteamiento: la primera es que ya existe de partida diferencias en las capacidades intelectuales que no vienen dadas por el esfuerzo del sujeto, sino por circunstancias sociales o biológicas adscritas y la segunda es hasta qué punto pueden juzgarse las capacidades intelectuales. ¿Somos todos capaces de conseguir un título universitario, sólo los que posean un determinado coeficiente intelectual, o es una cuestión de distribución estadística, de modo que sólo los mejores puedan alcanzar las potencialidades educativas que le permita obtener mejores puestos de trabajo y mayores recursos económicos?

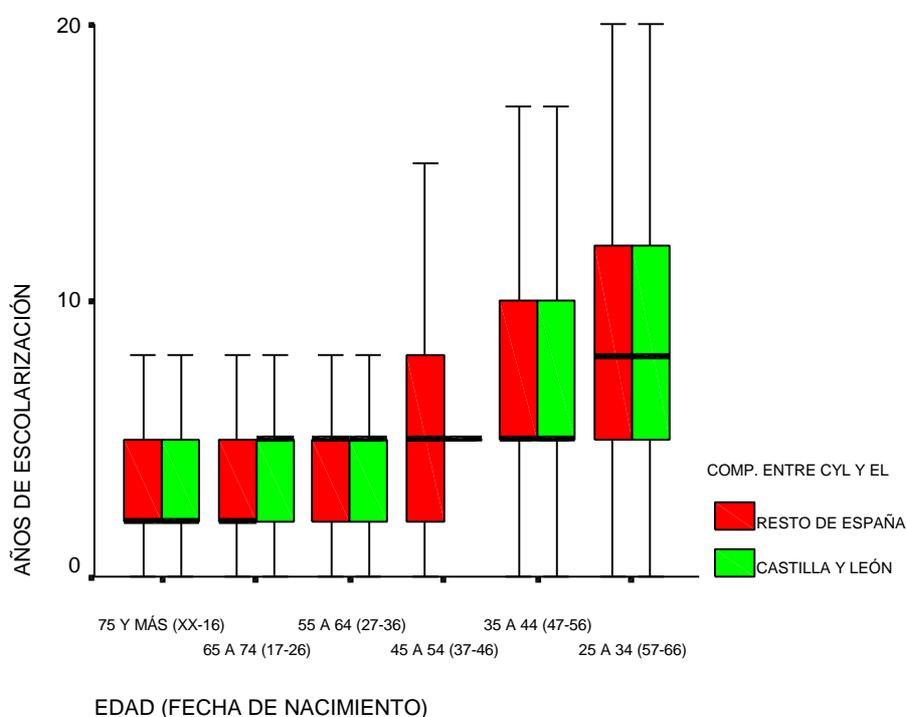
Tabla 2.- Distribución del nivel educativo de Castilla y León y del resto de España.

		% de COMP. ENTRE CYL Y EL RESTO DE ESPAÑA (RESIDENCIA)		
		AMBITO		Total
		RESTO DE ESPAÑA	CASTILLA Y LEÓN	
ESTUDIOS REALIZADOS	NO SABE LEER NI ESCRIBIR	4.5%	1.7%	4.3%
	SIN ESTUDIOS	26.3%	21.5%	26.0%
	ESTUD PRIMARIOS 6 5 CURSOS DE EGB O EQUIV.	34.5%	45.7%	35.3%
	BACH.ELEMENTAL,GRAD.ESCOLAR,EGB O EQUIV	14.3%	11.9%	14.1%
	FP.1er GRADO, OFICIALIA INDUSTRIAL	2.5%	2.1%	2.5%
	FP.2º GRADO,MAESTRIA INDUSTRIAL	2.0%	2.0%	2.0%
	BACH.SUPERIOR, BUP	5.8%	5.6%	5.8%
	OTRAS TITULACIONES MEDIAS	1.6%	1.1%	1.5%
	ARQUITEC.E INGENIERO TECNICOS	.9%	.7%	.9%
	DIPLOMADO DE ESCUELA UNIVERSITARIAS	2.7%	3.5%	2.8%
	ARQUITEC.O INGENIERO TEC.	.4%	.2%	.3%
	LICENCIADO UNIVERSIT.	3.3%	3.2%	3.3%
	ESTUD.SUP.NO UNIVERSIT.	.6%	.4%	.6%
	DOCTORADO	.3%	.2%	.3%
	ESTUD.DE POSGRADO O ESPECIALIEN LICENCIATURAS	.4%	.2%	.3%
Total		100.0%	100.0%	100.0%

Fuente: INE. Censo de Población Española. 1991. Elaboración propia

En cualquier caso es digno de atención el examen en cualquier sociedad de las desigualdades educativas que, al igual que las económicas, constituyen uno de los ejes de división y fragmentación social. En el presente artículo, se van a estudiar considerando la variable años de educación, extraída a partir de la clasificación efectuada por el censo de la población española de 1991³ y teniendo en cuenta sólo a la población mayor de 25 años pues es la única que ha podido completar su ciclo educativo. Considerando una muestra del 10% de las 24.776.040 personas con estas características, se obtuvo la distribución de la educación en España y, en concreto, en Castilla y León que se presenta en la tabla 2.

Gráfico 2.- Medianas y cuartiles de años de escolaridad según generación y región



Fuente: INE. Censo de Población Española. 1991. Elaboración propia

³ Esta clasificación consta de las siguientes categorías: No sabe leer ni escribir (0); Sin estudios (2); Estudios primarios o 5 cursos aprobados de EGB o equivalentes (5); Bachiller Elemental, Graduado Escolar, EGB completa o equivalentes (8); Formación Profesional 1er grado, Oficialía Industrial (10); Formación Profesional en 2º grado, Maestría Industrial (12); Bachiller Superior, BUP (12); Otras Titulaciones medias (12); Arquitecto e Ingeniero Técnico y Diplomado Escuelas Técnicas Superiores (15); Diplomado de Escuelas Universitarias y Diplomado (15); Arquitecto o Ingeniero Superior (17); Licenciado Universitario (16); Titulaciones de Estudios Superiores no universitarios (15); Doctorado (20) y Titulaciones de Estudios de Postgrado o Especialización para Licenciados (20). Entre paréntesis se ha incluido el valor numérico en años atribuido a cada una de estas categorías formales del sistema educativo español para convertirla en años de educación, según criterio ya seguido por Carabaña (1992).

Destaca claramente cómo el nivel de analfabetismo en la región de Castilla y León es inferior al del resto de España; sin embargo es menos evidente cómo el porcentaje de personas con estudios primarios completados es superior en el conjunto del resto de regiones españolas (31.1% frente a 34.7%). En conjunto, la media es ligeramente mayor (una décima) en las nueve provincias castellano y leonesas que en las restantes 41. De todos modos, para una justa comparación es necesario controlar por una variable crucial en el análisis de los niveles educativos como es la generación⁴, pues las oportunidades educativas son muy distintas entre las personas según hayan nacido a principios de siglo o si lo han hecho en su segunda mitad. La escolarización ha sido claramente creciente a lo largo del tiempo en España.

Tabla 3.- Promedios y desviaciones de años de escolarización en España y Castilla y León.

		AÑOS DE ESCOLARIZACIÓN							
		COMP. ENTRE CYL Y EL RESTO DE ESPAÑA (RESIDENCIA)							
		RESTO DE ESPAÑA				CASTILLA Y LEÓN			
		MEDIA	Med.	D.T.	CVa	MEDIA	Med.	D.T.	CVa
EDAD (FECHA DE NACIMIENTO)	25 A 34 (57-66)	9.0	8	4.2	.43	9.4	8	4.2	.42
	35 A 44 (47-56)	7.2	5	4.4	.45	7.5	5	4.1	.43
	45 A 54 (37-46)	5.6	5	4.0	.45	5.9	5	3.6	.40
	55 A 64 (27-36)	4.5	5	3.6	.43	4.8	5	3.2	.37
	65 A 74 (17-26)	4.0	2	3.4	.42	4.3	5	3.0	.36
	75 Y MÁS (XX-16)	3.3	2	3.2	.43	3.7	2	2.8	.36

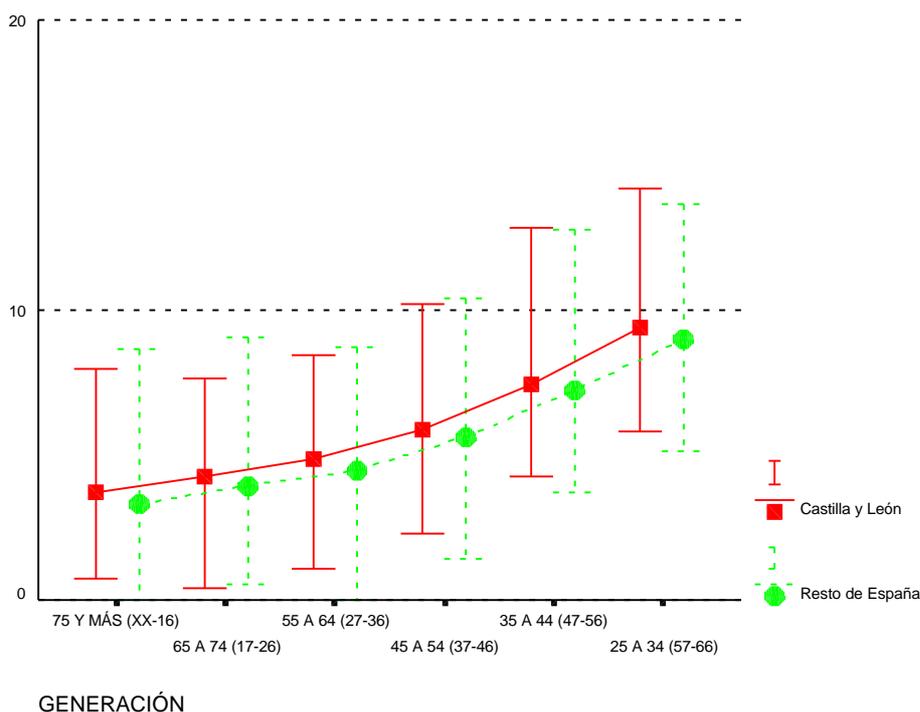
Fuente. Censo de la Población Española. 1991. Elaboración propia

⁴ Dos son los principales sesgos de un estudio temporal a partir de datos de corte transversal. Por un lado, la mortalidad. No todos los efectivos de una cohorte llegan vivos al momento presente, por lo que podría ser arriesgado extrapolar los resultados obtenidos con los vivos a los muertos. Por otro lado, la movilidad espacial, pues tomamos la población que reside en cada zona en el momento de realización del censo. Ambas objeciones son relevantes en la medida en que estén correlacionadas con el nivel educativo. La mortalidad no lo está con el nivel de estudios pues éste afecta a la morbilidad pero no a la edad de defunción. En cambio, las migraciones sí lo están muy positivamente. Por tanto, habrá que tener en cuenta también el efecto de las migraciones sobre la evolución educativa. En este contexto sólo se hará con los movimientos de población extrarregionales de Castilla y León.

Para los residentes, el número de años de escolarización alcanzado es muy similar en cada generación tanto en Castilla y León, como en el resto de España, con una muy leve superioridad en Castilla y León apenas reflejada en un gráfico exploratorio⁵ (gráfico 2) de las distribuciones, donde aparece netamente que la mediana de la educación ha ascendido de ser a principios de siglo de 2 años hasta alcanzar la primaria completa (8 años) en los nacidos entre 1957 y 1966. En cambio, si concentramos la atención en las medias, puede observarse que las correspondientes a Castilla y León son ligeramente más altas que las del resto de España, prácticamente en todas las generaciones analizables del presente siglo.

En el gráfico 2 se ve cómo la desigualdad intercuartílica se acrecienta con el paso del tiempo, como también lo atestiguan las desviaciones típicas presentes en la tabla 3, donde se ofrecen medias, desviaciones típicas y coeficientes de variación clasificados por región y generación.

Gráfico 3.- Medias y Cva de años de educación según generación y región



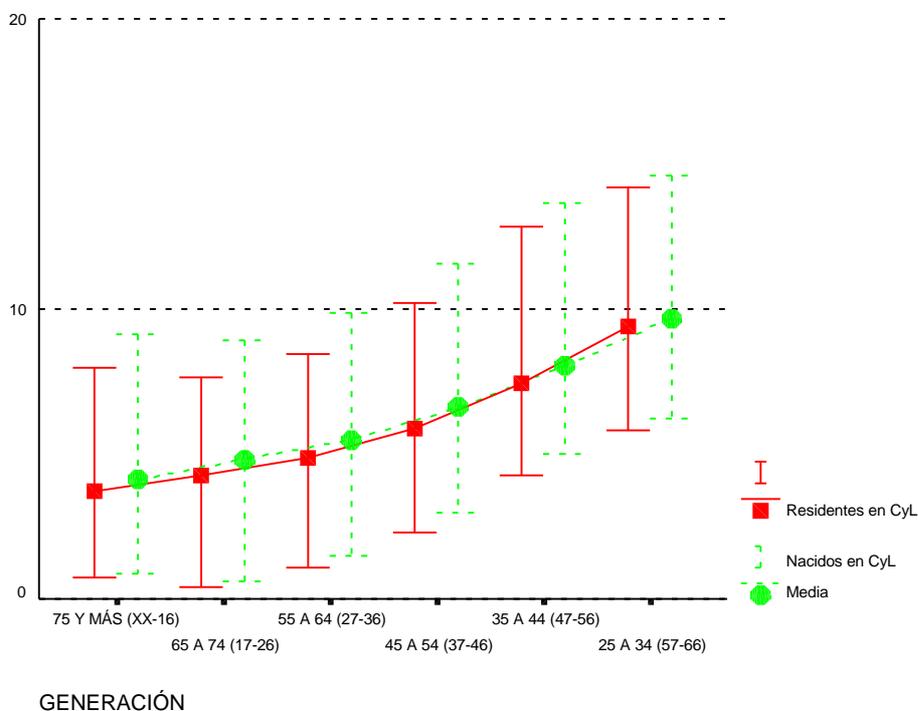
Fuente. Censo de la Población Española. 1991. Elaboración propia

⁵ Una de las deficiencias de este tipo de gráficos queda reflejada en la generación de personas de Castilla y León nacidas entre 1937 y 1946, pues al tener el primer cuartil y el tercero en estudios primarios incompletos (la antigua primaria) no aparece ninguna desviación en los datos por ser el rango intercuartílico igual a cero.

Sin embargo, la desigualdad relativa medida a través del coeficiente de variación acotado es claramente inferior en Castilla y León, con tendencia a incrementarse, de forma que casi se iguala en las últimas generaciones con el resto de España. Viéndolo con datos numéricos, se advierte que el coeficiente de variación acotado da valores bastante constantes a lo largo del periodo considerado, mientras que la desviación típica tiende a aumentar. Este incremento ha sido estudiado y demostrado con un conjunto amplio de países en estudios anteriores (Ram, 1980) y es consecuencia de la misma escala de medición que hace que a medida que los valores de la media se sitúan en el centro de la escala, la varianza aumente, de modo que si se hace un cruce gráfico entre media y varianza la forma de la relación aparece en forma de parábola invertida.

Si en lugar de considerar a los residentes en Castilla y León se promedian los estudios de los nacidos en ella, se observa claramente que hay siempre una media superior para éstos que para los residentes, lo que indica que los emigrantes tienen una educación mayor que los que se quedan en el interior de la región. También es mayor la dispersión tanto absoluta como relativa y, del mismo modo que converge la media, la desviación es cada vez menor a medida que las generaciones son más jóvenes.

Gráfico 4.- Medias y Cva de escolarización según generación y nacimiento en Cy L

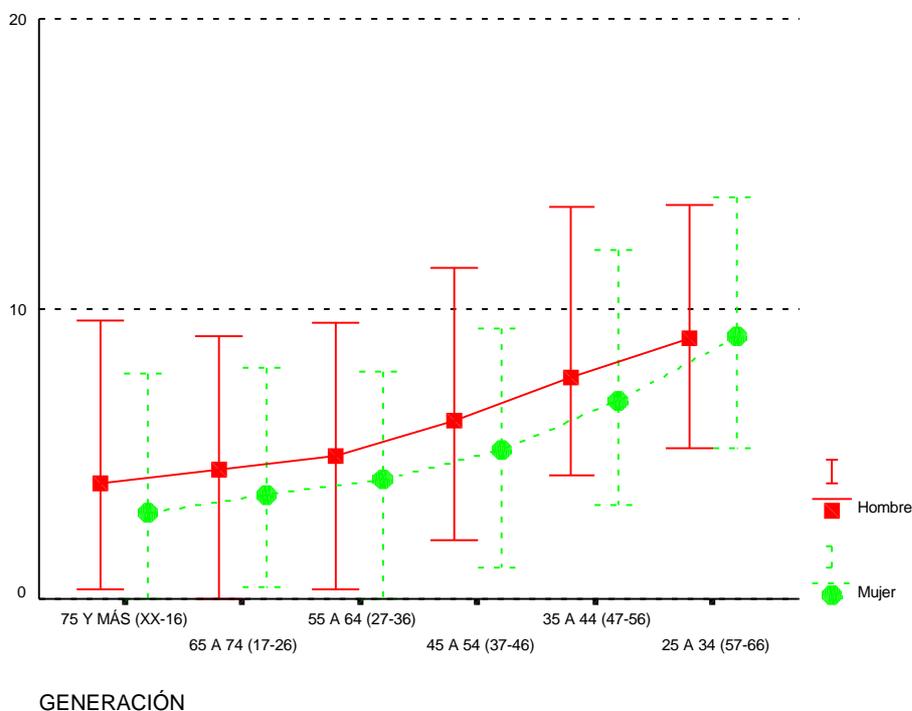


Fuente. Censo de la Población Española. 1991. Elaboración propia

Obviamente, el sexo es una variable ineludible para el estudio de la desigualdad educativa. En primer lugar, porque es una de las variables sociológicas clásicas para el estudio de la igualdad de oportunidades, además de la generación (o edad) que ya ha sido contemplada, la clase social (que no es objeto del presente artículo) y de la etnia (que en este país no puede ser analizada de momento) y, en segundo lugar, porque ha experimentado una gran evolución hacia la convergencia claramente perceptible a lo largo del presente siglo.

En el gráfico 6 se expone la evolución para el conjunto de España de las diferencias de género en educación. Es fácilmente apreciable que las primeras generaciones del siglo se caracterizan por una diferencia de un año de educación entre sexos, que se ve igualada por la generación que corresponde terminar sus estudios al comienzo de la democracia. El gráfico muestra cómo la desigualdad educativa ha sido mayor entre los varones. Como había pocas mujeres en el sistema superior de enseñanza, su heterogeneidad era bastante menor que la de los varones, caracterizada además por desviaciones hacia los valores bajos de la escala; pero a partir de las nacidas en el 47, la desigualdad entre las mujeres empieza a aumentar y alcanzar en la generación del 57 valores similares al de los hombres.

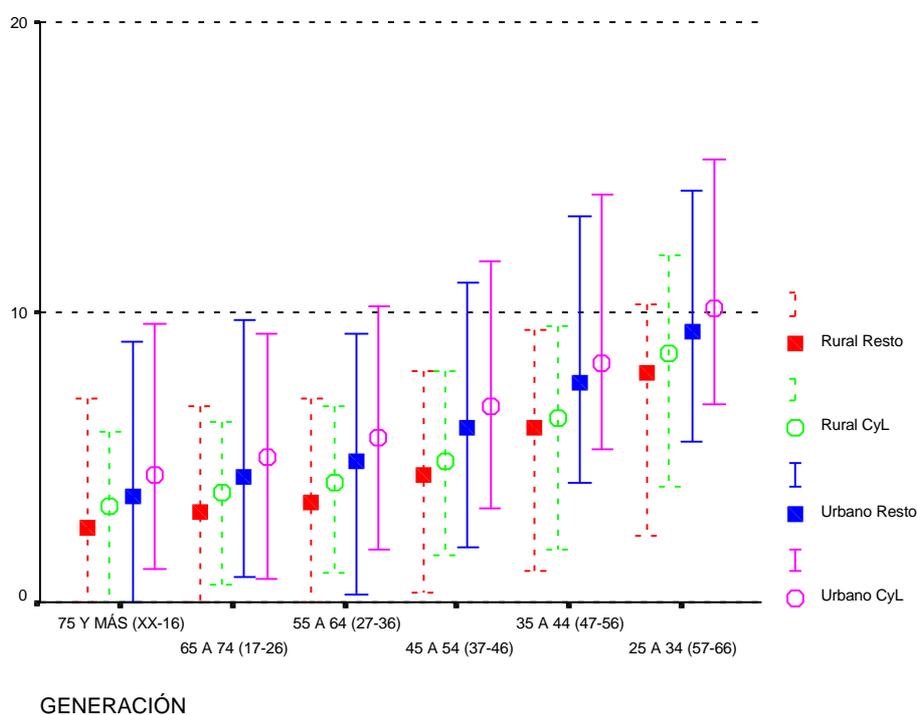
Gráfico 5.- Medias y Cva de la escolarización según generación y género.



Fuente. Censo de la Población Española. 1991. Elaboración propia

También deben ser tenidas en cuenta las desigualdades espaciales. En primer lugar se atenderá a las diferencias de hábitat, habiendo distinguido para ello, los municipios, por un lado, menores de 10.000 habitantes y, por el otro, los que tienen esta cantidad o superior de residentes. Los datos muestran algo tan evidente como que las personas con mayores estudios se concentran en los núcleos urbanos y que tanto en las ciudades como en los centros rurales, la media castellano y leonesa es superior a la existente en el resto de poblaciones españolas (gráfico 6).

Gráfico 6.- Medias y CV_a de escolarización según hábitat y región

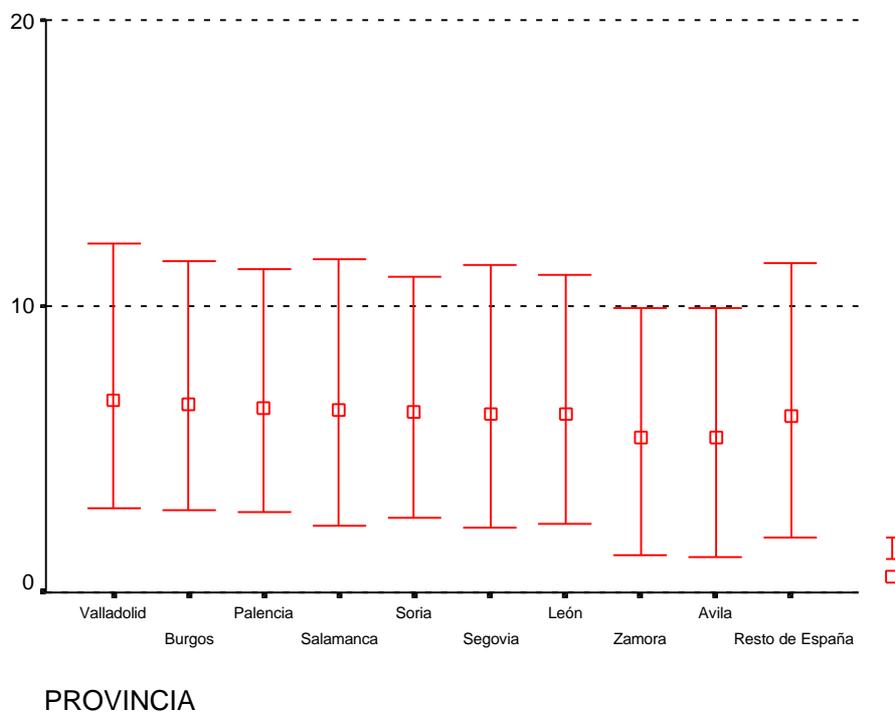


Fuente. Censo de la Población Española. 1991. Elaboración propia.

Y, en relación con las desigualdades han de resaltarse del mismo gráfico las dos siguientes pautas: primero que en los núcleos urbanos hay mayores desigualdades educativas que en los rurales y, segundo, la baja dispersión en Castilla y León del nivel educativo rural en las generaciones nacidas antes del 1946, especialmente las que tienen menos de 75 años, lo que puede ser explicado por una emigración diferencial, esto es, sólo se han quedado en los núcleos rurales de esta región quienes tenían bajos niveles educativos.

En segundo lugar, también es preciso tener en cuenta las desigualdades interprovinciales. Según se muestra en el gráfico 7, éstas son pequeñas. Destacan las provincias situadas en el eje Valladolid-Palencia-Burgos que son las tres con mayor promedio de años de educación entre sus residentes, mientras Ávila y Zamora, son las únicas que tienen una media inferior en medio año al conjunto de España. Y, centrado la atención a las desigualdades sobresalen dos hechos: uno la pocas diferencias interprovinciales, que se mueven en el rango desde Salamanca (46%) hasta Soria (42%) y dos, que la dispersión educativa es mayor fuera de la región (48%) que en cada una de sus provincias.

Gráfico 7.- Medias y CV_a de escolarización según provincias de Castilla y León



Fuente. Censo de la Población Española. 1991. Elaboración propia.

Resumiendo, el nivel educativo en Castilla y León ha sido secularmente mayor que en el resto de España, al tiempo que las desigualdades han sido menores, aunque con tendencia a converger. Y el nivel educativo y la desigualdad de las mujeres se ha incrementado más que los de los varones. Sin embargo, en su conjunto, no se aprecia una disminución de la desigualdad educativa; sino todo lo contrario. Medida en términos absolutos, la variación de los años que las personas han permanecido en el sistema escolar ha aumentado al tiempo que ha crecido la media. Pero, en este artículo se demuestra que, si se controla por la media, las desigualdades educativas no experimentan grandes cambios a lo largo de las seis primeras generaciones decenales del siglo, ni en Castilla y León, ni en el resto de España.

La pregunta que podría formularse es si estas desigualdades pueden considerarse tolerables para la garantía de la igualdad de oportunidades. Difícil respuesta tiene ésta pues hay que buscar el referente por el que hay que medir la igualdad óptima. Si ésta significa igual educación máxima para todos, es evidente que aún nos encontramos lejos de conseguir ese ideal; pero incluso, si el ideal es una distribución estadísticamente normal ($CV_a = 33\%$)⁶ en el reparto de los años de formación, también quedaba bastante por alcanzar a la generación nacida entre el 57 y el 66. Es posible que con los datos de la próxima generación, de los que dispondremos en un par de años, nos estemos aproximando, pero de lo que no cabe ninguna duda es que siguen siendo necesarias políticas que incrementen la formación y educación de las personas supuestamente menos desfavorecidas.

BIBLIOGRAFIA

- Atkinson, A.B. (1970) On the Measurement of Inequality, *Journal of Economic Theory* 2:244-263.
- Blalock, H.M. (1991) *Understanding Social Inequality*, Newbury Park, Sage.
- Bossert, W. y A. Pfingsten (1990) "Intermediate Inequality: Concepts, Indices, and Welfare Implications." *Mathematical Social Sciences*, 19: 117-134.
- Cortés, F. y Rubalcaba, R.M. (1984) *Técnicas estadísticas para el estudio de la desigualdad social*, México: El Colegio de México.
- Coulter, P.B. (1984) "Distinguishing Inequality and Concentration." *Political Methodology* 10: 323-335.

⁶ Véase Escobar (1998, 23)

- Coulter, P.B. (1989) *Measuring Inequality. A Methodological handbook*, Boulder: Westview Press.
- Escobar, M. (1998) "Desviación, desigualdad, polarización: medidas de la diversidad social", en *Revista Española de Investigaciones Sociológicas*, 82 9-36.
- Esteban, JM y D. Ray (1996) "Polarización y conflicto" *Perspectivas teóricas y comparadas de la igualdad. (II Simposio sobre Igualdad y Distribución de la Renta y Riqueza)*. Madrid: Fundación Argentaria.
- Goerlich, F.J. (1998) *Desigualdad, diversidad y convergencia: (algunos) instrumentos de medida*. Valencia: Instituto Valenciano de Investigaciones Económicas.
- Jacobson, H.I. (1969) "The Maximum Variance of Restricted Unimodal Distributions", en *Ann. Math. Statist.*, 10 1746-1752.
- Johnson, N.L. y Rogers, C.A. (1951) "Inequalities on moments of unimodal distributions" en *Ann. Math Statist.*, 22 433-439.
- Mas, M, F. Pérez, E. Uriel y L. Serrano (1995) *Capital Humano. Series Históricas, 1964-1992*, Fundación Bancaja.
- Nisbet, R. (1990) *La formación del pensamiento sociológico* Buenos Aires: Amorrortu.
- Osberg, Lars (ed.) (1991) *Economic Inequality and Poverty*, Armonk: Sharpe.
- Ram, R. (1990) "Educational Expansion and Scholing Inequality: International Evidence and Some Implications." *The Review of Economics and Statistics* 266-273.
- Rayner, S.C.W. (1975) "Variance Bounds", en *Indian Journal of Statistics*, 37 135-138.
- Ruiz Castillo, J.(1993) "Distribución personal de la renta: medición empírica y juicios de valor." *I Simposio sobre Igualdad y Distribución de la Renta y Riqueza*. Madrid: Fundación Argentaria.
- Theil, H. (1967) *Economics and Information Theory*. Amsterdam: North Holland Publishing Co.
- Weisberg, H.F. (1986) *Central Tendency and Variability*, Newbury Park: Sage

APÉNDICE

		AÑOS DE ESCOLARIZACIÓN												
		COMP. ENTRE CYL Y EL RESTO DE ESPAÑA (RESIDENCIA)								TOTAL				
		RESTO DE ESPAÑA				CASTILLA Y LEÓN								
		MEDIA	MED.	D.T.	CASOS	MEDIA	MED.	D.T.	CASOS	MEDIA	MED.	D.T.	CASOS	
VARON	EDAD (FECHA DE NACIMIENTO)	25 A 34 (57-66)	9.0	8	4.2	280686	9.1	8	4.1	19868	9.0	8	4.2	300555
		35 A 44 (47-56)	7.7	5	4.5	229776	7.7	5	4.3	15778	7.7	5	4.5	245554
		45 A 54 (37-46)	6.2	5	4.3	192678	6.2	5	3.9	13197	6.2	5	4.3	205874
		55 A 64 (27-36)	4.9	5	4.0	193017	5.1	5	3.4	15693	4.9	5	4.0	208710
		65 A 74 (17-26)	4.5	2	3.8	129041	4.5	5	3.2	11683	4.5	5	3.8	140724
	75 Y MÁS (XX-16)	4.0	2	3.8	73028	4.1	5	3.1	7537	4.0	2	3.7	80566	
MUJER	EDAD (FECHA DE NACIMIENTO)	25 A 34 (57-66)	9.0	8	4.3	276784	9.7	8	4.3	18690	9.1	8	4.3	295473
		35 A 44 (47-56)	6.8	5	4.2	231768	7.2	5	4.0	14258	6.8	5	4.2	246026
		45 A 54 (37-46)	5.1	5	3.6	198396	5.6	5	3.3	13493	5.1	5	3.6	211889
		55 A 64 (27-36)	4.0	5	3.2	209389	4.6	5	2.9	16747	4.1	5	3.2	226136
		65 A 74 (17-26)	3.5	2	2.9	163085	4.1	5	2.7	13741	3.6	2	2.9	176826
	75 Y MÁS (XX-16)	2.9	2	2.8	127533	3.5	2	2.6	11739	3.0	2	2.8	139272	

		AÑOS DE ESCOLARIZACIÓN												
		COMP. ENTRE CYL Y EL RESTO DE ESPAÑA (NACIMIENTO)								TOTAL				
		RESTO DE ESPAÑA				CASTILLA Y LEÓN								
		MEDIA	MED.	D.T.	CASOS	MEDIA	MED.	D.T.	CASOS	MEDIA	MED.	D.T.	CASOS	
VARON	EDAD (FECHA DE NACIMIENTO)	25 A 34 (57-66)	9.0	8	4.2	286463	9.7	8	4.1	14092	9.0	8	4.2	300555
		35 A 44 (47-56)	7.6	5	4.5	226667	8.4	8	4.4	18887	7.7	5	4.5	245554
		45 A 54 (37-46)	6.1	5	4.3	188869	7.2	5	4.4	17005	6.2	5	4.3	205874
		55 A 64 (27-36)	4.9	5	3.9	193322	6.0	5	4.1	15387	4.9	5	4.0	208710
		65 A 74 (17-26)	4.4	2	3.7	130511	5.4	5	4.1	10213	4.5	5	3.8	140724
	75 Y MÁS (XX-16)	3.9	2	3.7	73798	4.9	5	3.9	6768	4.0	2	3.7	80566	
MUJER	EDAD (FECHA DE NACIMIENTO)	25 A 34 (57-66)	9.0	8	4.3	279310	9.7	8	4.2	16163	9.1	8	4.3	295473
		35 A 44 (47-56)	6.8	5	4.2	225296	7.6	5	4.1	20730	6.8	5	4.2	246026
		45 A 54 (37-46)	5.1	5	3.6	193203	6.1	5	3.6	18686	5.1	5	3.6	211889
		55 A 64 (27-36)	4.0	5	3.2	208862	5.0	5	3.2	17274	4.1	5	3.2	226136
		65 A 74 (17-26)	3.5	2	2.9	163844	4.4	5	3.0	12982	3.6	2	2.9	176826
	75 Y MÁS (XX-16)	2.9	2	2.7	127990	3.6	2	2.8	11282	3.0	2	2.8	139272	