

ANALFABETISMO.—EL PROBLEMA ESTADISTICO

Las estadísticas relativas a analfabetismo se consiguen directamente por medio de los censos de población o por medio de estimaciones obtenidas de muestras de esta población, y por métodos indirectos.

En todos los casos, la tasa de analfabetismo puede ser calculada con relación a la población total o con relación a la población de una edad determinada.

Lo primero que tendremos que hacer antes de medir el analfabetismo por cualquiera de esos procedimientos es concretar la definición de analfabeto, siendo esta cuestión muy importante, pues hay necesidad de uniformar definiciones y procedimientos, con el fin de que los datos recogidos sean comparables con los de otros países. Algo se ha hecho en América. Reuniones de representantes de países americanos y un Comité especial del Instituto Interamericano de Estadística, han llegado a un acuerdo sobre la definición de *alfabetismo* para fines estadísticos, como «la facultad de leer y escribir un párrafo sencillo» en una lengua y sobre la inclusión de la pregunta «¿Sabe usted leer y escribir?» en la cédula censal.

Esta definición ha sido criticada por el grupo especialista en analfabetismo del «Mysore Seminar» como inadecuado desde un punto de vista cultural, y ha sugerido la definición siguiente: «Alfabetismo es la aptitud para leer y escribir con comprensión y facilidad un párrafo sencillo», y, además, las condiciones que ha de cumplir una persona para ser verdaderamente «letrada» (1):

- a) Hablar y comprender su lengua clara y fácilmente.
- b) Leer materias necesarias para su vivir diario.
- c) Expresar sus ideas escribiendo.
- d) Hacer un cálculo aritmético sencillo.
- e) Tener algún conocimiento de la Historia, cultura e Instituciones de su comunidad y país.
- f) Tener algún concepto sobre su relación con la comunidad mundial.

Medida del analfabetismo por medio de los censos de población.—Hasta la fecha, en España se ha medido el analfabetismo por medio de los censos de población, que se realizan cada diez años, habiendo sido efectuado el último en 1950.

La publicación «Hechos y cifras», de la U. N. E. S. C. O., de julio

(1) Literacy: condición de letrado o instruído.

de 1952, presenta una tabla con los datos numéricos relativos a analfabetismo en 86 países y territorios; estos datos proceden de las estimaciones o censos más recientemente efectuados en el curso de los veinticinco últimos años. Los resultados de la mayor parte de los censos efectuados alrededor de 1950 no los publica, por no tenerlos en la fecha de la publicación a su disposición. De España tiene los relativos al censo de 1940.

La tabla consta de los datos siguientes:

País	Año del censo o de la estimación	Edad de la población	Porcentaje de analfabetos
------	----------------------------------	----------------------	---------------------------

De los 86 países que figuran, 74 han medido el analfabetismo por censo y 12 por estimación, y con respecto a las edades, consideran:

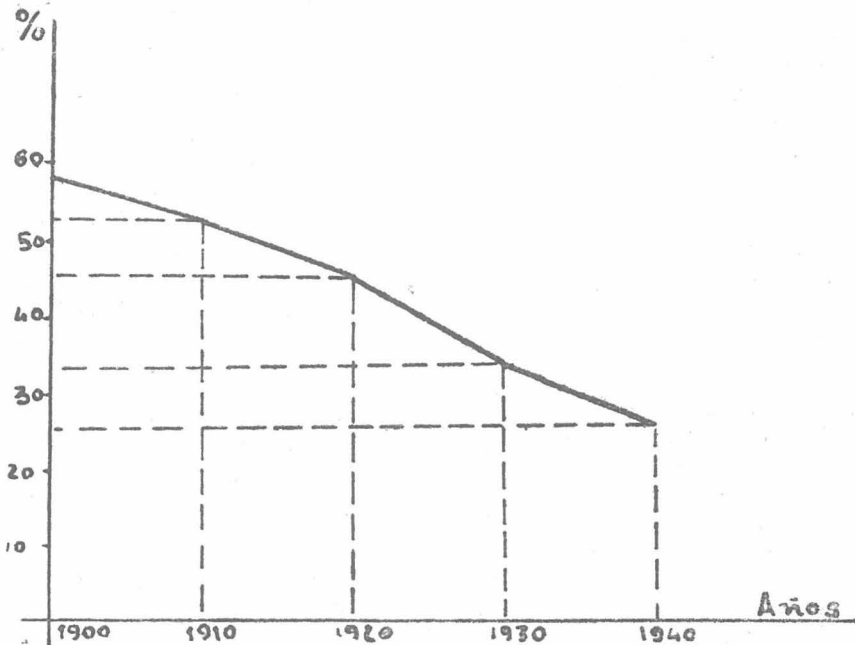
Todas las edades	16 países.		
5 años y más	1	»	
6 » »	1	»	
7 » »	8	»	
9 » »	1	»	
10 » »	43	»	entre ellos, España.
11 » »	1	»	
14 » »	2	»	
15 » »	5	»	
Ninguna indicación	8	»	

86

En España, en los distintos censos se han obtenido las siguientes tasas de analfabetismo, es decir, porcentajes de analfabetos de diez años y más, con respecto a la población total de diez años y más:

Año	Porcentaje
1900	58,8 %
1910	52,4 %
1920	44,3 %
1930	33,2 %
1940	23,2 %

Con relación a algunos de los resultados sobre analfabetismo obtenidos en España en el censo de 1950, que no se ha publicado todavía, gracias a la autorización del Ilmo. Sr. D. Adolfo Melón, jefe del Servicio de Estadísticas Demográficas del I. N. E., puedo incluir una nota sobre resultados, proporcionada por el jefe de la Sección de Censos del citado Servicio.



NOTA DEL SERVICIO DE ESTADISTICAS DEMOGRAFICAS

«Resultados.—Por lo que se refiere a la distribución geográfica del analfabetismo, resulta que en España hay tres zonas perfectamente diferenciadas: Primera, de escaso analfabetismo, con coeficientes tan buenos como los de los países más cultos, en todo caso inferiores al 5 por 100; ésta es la mitad superior de España, con la exclusión de Galicia, Asturias, Cataluña y zonas próximas a ellas de otras provincias. Segunda, de analfabetismo medio, con coeficientes no muy superiores al 10 por 100, integrada por las provincias de Galicia, Asturias y Cataluña, excluidas de la zona anterior, y Extremadura, Castilla la Nueva y Levante. Tercera, zona de incultura, que en algunos casos presenta coeficientes superiores al 40 por 100, en Andalucía.

Por generaciones, el coeficiente varía en forma extraordinaria: desde el 69 por 100 que da la generación escolar de 1846-1850, disminuye hasta resultar inferior al 10 por 100, el de las últimas generaciones escolares. Esto explica el aparente escaso éxito de todas las campañas contra el analfabetismo, pues al actuar directamente sobre los escolares, como es

lógico, no modifican el coeficiente general lo suficiente, ya que éste sólo bajará hasta donde sea debido cuando hayan desaparecido las masas de analfabetos que forman en las generaciones antiguas.

Por sexo, la mujer da coeficientes aproximadamente dobles que el varón.

A título de ejemplo:

Provincia de Alava.—Zona urbana: Varones, 0,66 %; mujeres, 1,47 %. Zona intermedia: Varones, 1,11 %; mujeres, 1,49 %. Zona rural: Varones, 2,66 %; mujeres, 3,11 %.

Provincia de Segovia.—Zona urbana: Varones, 3,23 %; mujeres, 5,07 %. Zona intermedia: Varones, 2,93 %; mujeres, 4,71 %. Zona rural: Varones, 2,18 %; mujeres, 4,69 %.

Y así en las demás.

En el censo de 1950, y por primera vez, se ha medido el analfabetismo por separado en las zonas urbanas (núcleos de más de 10.000 habitantes), intermedias (núcleos de 2.000 a 10.000) y rurales (núcleos de menos de 2.000 h.), con resultados aparentemente sorprendentes, pero de fácil explicación.

En las regiones de escaso analfabetismo supera la zona intermedia, y en algunos casos la urbana, a la rural. Esta última es la de menores coeficientes. Así pasa en toda Castilla la Vieja. Y los motivos son los siguientes:

a) Los pueblos pequeños no crecen o crecen muy poco. Su escuela y personal docente fué y sigue siendo suficiente para atender la población escolar. No hay problema.

b) Los pueblos que crecieron no consiguieron aumentar en la misma proporción sus medios de enseñanza, resultando insuficientes. Así ocurre con las capitales, en las que la insuficiencia la sufren las clases menos dotadas económicamente, situadas, en general, en suburbios, que en gran parte se clasificaron en zona intermedia.

c) En los suburbios de las grandes capitales residen los inmigrantes de menor nivel económico procedentes de otras provincias, y éstos, en general, son del grupo más inculdo de las poblaciones de procedencia.

Una vez se disponga de la totalidad de los coeficientes españoles será posible fijar con precisión las zonas geográficas, económicas y demográficas en las que debe atacarse este problema, que, al parecer, no son las generalmente tenidas como tales, por falta de una información concreta.»

Estimación por medio de muestras.—Una muestra es una parte o subconjunto que seleccionamos de un conjunto al cual llamamos *población*.

Consideramos dos clases de muestras, *muestras opináticas* y *muestras probabilísticas*.

En la selección de los elementos de una muestra opinática intervienen

la opinión o juicio de una persona, que puede ser el entrevistador o el grupo de personas que dirige el trabajo, u otras. Estas muestras tienen el defecto fundamental de que su precisión es puramente subjetiva, y el error de muestreo no puede controlarse ni calcularse.

En la selección de los elementos de una muestra probabilística o aleatoria se sigue un procedimiento automático o por sorteo. La precisión de los resultados obtenidos tiene una base objetiva, y en estas muestras puede controlarse y calcularse el error de muestreo.

Está, por lo tanto, claro que cuando se trata de tomar decisiones importantes a partir de la información muestral, debemos tomar muestras probabilísticas. El folleto de la U. N. E. S. C. O., mencionado anteriormente, cita la estimación del analfabetismo hecha en Estados Unidos en 1947, efectuada por medio de una muestra probabilística en 42 Estados y en el distrito de Columbia.

De los distintos tipos de muestreo probabilístico consideramos los dos más importantes: el muestreo irrestrictamente aleatorio y estratificado.

Muestreo irrestrictamente aleatorio.—Se saca de una población de tamaño v :

$$X_1, X_2, \dots, X_v, \text{ una muestra}$$

$$x_1, x_2, \dots, x_n,$$

por medio de un sorteo (tablas de números aleatorios).

En toda muestra hay que tener en cuenta dos problemas fundamentales:

1.º Problema de estimación de las características de la población a partir de la muestra.

2.º Grado de confianza que podemos depositar en estas estimaciones.

Designemos por ω la proporción o tanto por 1 de sujetos de una característica determinada; por ejemplo, tanto por 1 de analfabetos de

la población que estimamos por $p = \frac{c}{n}$ siendo c el número de sujetos

que presentan esa característica en la muestra, en nuestro caso:

c = número de analfabetos.

n = número de individuos de la muestra.

El error de muestreo de este estimador viene dado por

$$D[p] = \sqrt{\frac{v-n}{v-1} \frac{\omega \varphi}{n}} \quad \text{y} \quad \sqrt{\frac{v-n}{v-1} \frac{pq}{n}}$$

donde

$$\varphi = 1 - \omega \quad ; \quad q = 1 - p$$

Si la distribución del estimador es normal (basta que v y n sean grandes), sabemos que:

$$P[|p - \omega| < E] = P_k$$

donde E es el error máximo permisible:

$$E = k D[p]$$

teniendo K y P_k (coeficiente de confianza) tabulados:

K	P_k
1	0,68268
2	0,95450
3	0,99730
.....

Por ejemplo, si tenemos $K = 2$, $P_k = 0,95$

$$E = 2 D[p]$$

luego:

$$P[|p - \omega| < E] = 0,95$$

de donde obtenemos el *intervalo de confianza*

$$(p - E, p + E)$$

intervalo aleatorio que tiene una probabilidad de 0,95 de comprender al parámetro ω .

Expresado en forma de porcentaje, podemos decir que hay una probabilidad igual a 0,95 de que los límites del intervalo

$$(p - E) \cdot 100 \quad \text{y} \quad (p + E) \cdot 100$$

comprendan a $\omega \cdot 100$ y tomaremos como estimación del porcentaje de analfabetos: $p \cdot 100$ punto medio del intervalo anterior.

Muestreo estratificado aleatorio.—Se considera dividida la población en subpoblaciones, clases o grupos lo más homogéneos posibles, que denominamos *estratos*. En nuestro caso podían ser las provincias. Se toma una muestra irrestrictamente aleatoria de cada uno de dichos estratos.

Siendo m el número de estratos (número de provincias), v_i el tamaño del estrato i -ésimo en la población (número de habitantes de la

provincia i) y p_i la proporción o tanto por 1 de analfabetos del estrato 1

$$p_i = \frac{c_i}{n_i} \quad ; \quad q_i = 1 - p_i$$

siendo c_i = número de analfabetos en la muestra del estrato i

El estimador de ω es

$$\hat{p} = \frac{1}{v} \sum_{i=1}^m v_i p_i$$

El error de muestreo es

$$D[\hat{p}] \approx \frac{1}{v} \sqrt{\sum_{i=1}^m v_i^2 \frac{v_i - n_i}{v_i - 1} \frac{p_i q_i}{n_i}}$$

como antes, el intervalo de confianza será

$$(p - E, p + E)$$

y el error máximo permisible $E = K D[\hat{p}]$

y podemos estimar el porcentaje de analfabetos

$$p \cdot 100$$

con un coeficiente de confianza P_k .

POBLACIONES INFINITAS

En poblaciones infinitas, todas estas fórmulas quedan muy simplificadas; en nuestro caso, si medimos el analfabetismo en toda España, la población española en este caso puede considerarse como muy grande y aplicar las fórmulas de las poblaciones infinitas.

Para el error de muestreo de la proporción en el muestreo irrestrictamente aleatorio, tendríamos

$$D[p] = \sqrt{\frac{\omega \varphi}{n}} \approx \sqrt{\frac{p q}{n}}$$

y en el estratificado

$$D[\hat{p}] \approx \sqrt{\sum_{i=1}^m v_i^2 \frac{p_i q_i}{n_i}}$$

si los estratos son muy grandes.

Métodos indirectos de medir el analfabetismo.—En algunos países, una pregunta sobre el número de años de asistencia escolar ha sustituido a la medida directa del analfabetismo.

En el Seminario de Río de 1949 se propuso que aquellos que no han pasado el primer año de escuela deben ser considerados como analfabetos; los que han completado un año deben ser considerados como semi-analfabetos, y los que tienen dos años escolares, como alfabetos imperfectos, mientras que los que han pasado del tercer año deberán ser considerados como alfabetos.

Un informe del Bureau de Censos, Wáshington, EE. UU., considerando los datos sobre enseñanza y alfabetismo, indican que toda la población de catorce años y más, con cinco o más años de escuela, pueden ser considerados como alfabetos. Los porcentajes de incapaces para leer y escribir que tenían menos de cinco años de escuela eran: para aquellos con cuatro, tres, dos, uno y cero años de escuela, 5 por 100, 19 por 100, 46 por 100, 67 por 100 y 80 por 100, respectivamente.

Este procedimiento de medir el analfabetismo por el número de años de enseñanza escolar completa tiene el inconveniente de que individuos que no han recibido enseñanza formal (en la escuela) pueden ser altamente letrados, y al contrario: entre los alfabetos habrá alguno que, faltándole por completo la práctica, pierde tal facultad que poseía y llegue a ser analfabeto.

Otros ejemplos de medidas indirectas:

Francia medía el analfabetismo del año 1832 a 1906 por los esposos que no sabían firmar el acta matrimonial. Desde 1907, los que no sabían leer ni escribir según su propia declaración al contraer matrimonio. Después de 1924, por los datos de los mozos incorporados a filas, hasta 1936, que se hizo por medio de un censo.

Como se ve, en todos los casos de medida indirecta, lo que obtenemos es una estimación bastante burda del grado de analfabetismo, y siempre es preferible, sobre todo si se quiere medir con mayor exactitud, alguno de los procedimientos directos.

SEBASTIÁN FERRER MARTÍN