

Diseño muestral de la Encuesta sobre integración social y necesidades sociales

Manuel Trujillo Carmona
Instituto de Estudios Sociales Avanzados del CSIC



FUNDACIÓN FOESSA
FOMENTO DE ESTUDIOS SOCIALES
Y DE SOCIOLOGÍA APLICADA

1. Objetivos

El objetivo de la encuesta es cuantificar y analizar las condiciones de vida y la exclusión social en España, obteniendo datos representativos a nivel nacional y a nivel de algunas comunidades autónomas. Esto se pretende con dos perspectivas: una, estudiar a los hogares que están en situación de exclusión social, o con gran número de problemas, y por otro lado relacionar la situación de estos hogares con el resto de la población.

2. Unidades de encuestación y universo

El universo (población de referencia) son los hogares que se encuentran en España salvo Ceuta y Melilla. La unidad de encuestación son los hogares.

3. Procedimiento general de muestreo

La encuesta se realiza de forma presencial, en hogares. El procedimiento de selección de los hogares es bietápico, con una primera selección de secciones censales y una segunda selección mediante rutas aleatorias.

La encuesta contiene dos cuestionarios, uno de filtro y uno segundo el principal, con el grueso de las preguntas. El segundo cuestionario se pregunta solo a las personas que cumplen determinadas condiciones del primer cuestionario, y a una parte de los que no las cumplen. Este procedimiento sirve para garantizar que se obtenga un número suficiente de hogares con indicios de exclusión social.

4. Clasificación de los hogares

Dado que queremos estudiar en especial a los hogares con exclusión social, o que acumulen problemas, necesitaremos un número mínimo de estos para poder analizarlos. Para ello se ha diseñado un cuestionario de filtro, con una serie de preguntas que revelarían problemas que, si bien no son determinantes de la exclusión social, sí son indicios de una mala situación familiar. De esta manera se clasifican a priori los hogares en hogares con indicios de exclusión y sin indicios.

5. Selección de las secciones censales

La selección de las secciones censales donde se ha realizado la encuesta no ha sido proporcional a la población, sino que ha sido proporcional a la población estimada con indicios de exclusión social. Esta estimación es la misma utilizada en las encuestas FOESSA de 2008 y 2011, y su cálculo se puede consultar en el volumen del VI Informe sobre exclusión y desarrollo social en España 2008. Esta estimación se basó en los datos del Censo de Población y Viviendas de 2001 a nivel de sección censal.

Esto se hace así para conseguir hogares con problemas de exclusión social con mayor facilidad. La desproporcionalidad en la selección de las secciones censales se corrige a posteriori con ponderación.

6. Tamaño muestral

El objetivo de determinar las características de los hogares con problemas en sus condiciones de vida determina que tenemos una necesidad de un nivel máximo de error para los hogares que van a contestar el cuestionario segundo (cuestionario principal). Por tanto, en principio se fijará un tamaño muestral para este segundo cuestionario.

En determinadas comunidades autónomas se va a realizar una muestra específica. En estas comunidades la muestra total diseñada es de 650 cuestionarios, previéndose 400 para hogares con algún indicio de inclusión y 250 para hogares sin indicios.

En el resto de comunidades autónomas se calculará la muestra que les correspondería para alcanzar estos tamaños a nivel de NUT-1.

Con estos tamaños se garantiza un nivel de error, si se tratara de un muestreo aleatorio simple, del $\pm 5\%$ al nivel de confianza del 95% para la población con algún indicio de inclusión, y ese mismo nivel de error para los hogares en general, para las comunidades autónomas con muestra propia y para cada NUT-1.

La muestra total diseñada de España es de 8.769 cuestionarios, con 6.073 con indicios de exclusión y 2.696 sin indicios.

7. Procedimiento de aplicación de los cuestionarios

Como se ha dicho, la muestra está diseñada para conseguir un error máximo para los hogares con algún indicio de exclusión. Por tanto, necesitamos un número mínimo de hogares con esas características. Para ello, cuando se comienza un contacto con un hogar primero se realiza el cuestionario de filtro. El hogar se clasifica según si responde afirmativamente a alguna pregunta del cuestionario o no. En cada sección hay unas cuotas máximas para cada tipo de hogar. Si una vez realizado el cuestionario de filtro y el hogar clasificado en una de las dos categorías de hogares, la cuota correspondiente está libre, se realiza el cuestionario principal¹.

Dado que no se sabe a priori cuantos cuestionarios de filtro serían necesarios para conseguir la muestra requerida de cuestionarios con indicios de exclusión, se usó la Encuesta de Condiciones de Vida del INE en su oleada de 2011, para realizar una estimación a nivel de comunidad autónoma (Tabla 1). Con esa estimación se calculó el número de cuestionarios filtro que a priori serían necesarios para completar la muestra. Por secciones se impusieron las cuotas de forma proporcional al número estimado de hogares con exclusión procedente de la estimación realizada con el Censo de 2001. En cada sección se realizarán 18 cuestionarios de filtro.

El diseño preveía que a mitad de campo se revisarían estos números, de forma que si durante el trabajo de campo aparecía que estas estimaciones no eran correctas se pudiera modificar el diseño.

¹ En un principio solo se diseñaron cuotas máximas para los hogares sin indicios de exclusión, pero dado que aparecieron más hogares con indicios de exclusión de los previstos, en algunas rutas hubo que incluir también cuotas máximas para éstos.

Tabla 1. Tamaños muestrales diseñados

NUT-1	Comunidad autónoma	Cuestionario principal			Proporción estimada de hogares con indicios de exclusión	Cuestionarios filtro
		Total	Con indicios de exclusión	Sin indicios de exclusión		
1	Asturias	650	450	200	30,3%	2.556
	Cantabria	44	31	13	32,6%	126
	Galicia	650	450	200	36,1%	1.512
2	Aragón	650	450	200	28,9%	2.232
	Navarra	650	450	200	26,0%	2.988
	País Vasco	650	450	200	29,1%	2.106
	Rioja (La)	24	17	7	29,9%	72
3	Madrid	650	450	200	27,4%	1.854
4	Castilla la Mancha	149	104	45	34,4%	360
	Castilla León	650	450	200	25,7%	2.466
	Extremadura	650	450	200	34,6%	2.106
5	Baleares (Illes)	650	450	200	37,6%	918
	Cataluña	650	450	200	30,6%	1.458
	C. Valenciana	650	450	200	28,7%	2.268
6	Andalucía	650	450	200	43,9%	918
	Murcia	102	71	31	35,4%	216
7	Canarias	650	450	200	41,0%	1.044
	Total	8.769	6.073	2.696	33,1%	25.200

Fuente: Elaboración propia a partir de los datos del INE y de la EINFOESSA 2013

8. Selección del hogar

Una vez seleccionada la unidad primaria de muestreo, la sección censal, la selección de los hogares se realiza mediante rutas aleatorias dentro de la sección, definida por los callejeros del INE. En la selección de los hogares solo se sigue como norma las cuotas de categorías de hogar explicadas anteriormente.

9. Revisión de las estimaciones durante el trabajo de campo

Durante el trabajo de campo se comprobó que para la mayoría de las comunidades autónomas las estimaciones de hogares con indicios de exclusión eran muy inferiores a lo encontrado. Una de las explicaciones posibles es que la situación de los hogares había cambiado mucho de 2.011 a 2.013, en especial en la cantidad de hogares afectados por el desempleo. También puede influir la distinta metodología de selección de los hogares del INE, o diferencias en la forma de realización de la pregunta, en especial en el caso de los ingresos. Por ello se revisaron las estimaciones de cuestionarios de filtro necesarios para completar las cuotas y hubo que poner unas cuotas máximas para hogares sin indicios de exclusión, que inicialmente no se habían previsto.

Al necesitar menos cuestionarios de filtro, también se evidenció que en algunas comunidades autónomas no sería necesario realizar el trabajo de campo en todas las secciones

seleccionadas en principio, por lo que se eliminaron de la muestra prevista algunas secciones siempre cuidando de que se mantuviera la representatividad en la muestra que finalmente se realizaría.

Como resultado de estos ajustes, el número de cuestionarios filtro realizado ha sido bastante menor del previsto. El número de cuestionarios principales a hogares de cada tipo sí se ha ajustado aproximadamente a lo previsto.

Tabla 2. Proporción estimada y encontrada de hogares con indicios de exclusión

	% estimada	% encontrada
Andalucía	43,9%	71,5%
Aragón	28,9%	50,4%
Asturias	30,3%	47,8%
Islas Baleares	37,6%	54,7%
Islas Canarias	41,0%	68,2%
Cantabria	32,6%	51,3%
Castilla la Mancha	34,4%	60,1%
Castilla y León	25,7%	60,4%
Cataluña	30,6%	58,4%
Com. Valenciana	28,7%	65,8%
Extremadura	34,6%	56,3%
Galicia	36,1%	46,0%
Madrid	27,4%	54,4%
Murcia	35,4%	60,6%
Navarra	26,0%	53,3%
País Vasco	29,1%	47,7%
La Rioja	29,9%	46,7%

Fuente: Elaboración propia a partir de datos del INE y de la EINFOESSA 2013

10. Resultados del trabajo de campo

La EINSFOESSA es una encuesta dirigida a los hogares de España, en la que se ha recogido información de todas las personas que residen en cada uno de los hogares. Por tanto, existen dos niveles de análisis principal, el de los hogares, y el de la población.

La muestra de hogares ha resultado conformada por 8.776 encuestas, que han aportado información sobre 24.775 personas, lo que devuelve unos márgenes de error del $\pm 0,6\%$ para la información de la población y del $\pm 1,0\%$ para la información de los hogares. En ambos casos, el margen de error es óptimo, la información es estadísticamente representativa del conjunto del universo, y los intervalos de confianza permiten un análisis detallado de los niveles de integración y exclusión social.

Tabla 3. Muestra final realizada

	Cuestionarios filtro			Cuestionarios principales		
	Total	Con indicios de exclusión	Sin indicios de exclusión	Total	Con indicios de exclusión	Sin indicios de exclusión
Andalucía	1178	842	336	650	447	262
Aragón	1679	846	833	650	452	237
Asturias	1503	718	785	650	445	247
Islas Baleares	889	486	403	650	446	228
Islas Canarias	968	660	308	650	450	228
Cantabria	117	60	57	45	31	8
Castilla la Mancha	351	211	140	150	105	55
Castilla y León	1547	934	613	646	452	227
Cataluña	1301	760	541	650	450	253
Com. Valenciana	1540	1014	526	650	450	306
Extremadura	1289	726	563	650	447	256
Galicia	1212	557	655	650	437	221
Madrid	1352	735	617	650	452	275
Murcia	203	123	80	104	72	34
Navarra	1403	748	655	650	448	270
País Vasco	1747	834	913	656	450	265
La Rioja	75	35	40	25	18	9
Total	18.354	10.289	8.065	8.776	6.052	3.381

Fuente: EINFOESSA 2013

Tabla 4. Universo, muestra y margen de error de la EINSFOESSA en España

	Universo Población	Universo Hogares	Muestra Población	Muestra Hogares	Margen de error Población	Margen de error Hogares
Andalucía	8.390.624	3.155.417	1.958	650	± 2,2	± 3,8
Aragón	1.334.588	493.430	1.774	650	± 2,3	± 3,8
Asturias	1.063.241	395.019	1.767	650	± 2,3	± 3,8
Islas Baleares	1.115.255	421.931	1.786	650	± 2,3	± 3,8
Islas Canarias	2.110.039	811.682	2.014	650	± 2,2	± 3,8
Cantabria	*	*	112	45	*	*
Castilla-La Mancha	*	*	452	150	*	*
Castilla y León	2.507.574	923.263	1.620	646	± 2,4	± 3,9
Cataluña	7.451.281	2.743.449	1.843	650	± 2,3	± 3,8
Comunidad Valenciana	4.968.093	1.899.282	1.805	650	± 2,3	± 3,8
Extremadura	1.098.749	408.876	1.736	650	± 2,4	± 3,8
Galicia	2.753.960	1.028.316	1.958	650	± 2,2	± 3,8
Comunidad de Madrid	6.388.735	2.410.501	1.965	650	± 2,2	± 3,8
Región de Murcia	*	*	274	104	*	*
Navarra	636.652	233.660	1.870	650	± 2,3	± 3,8
País Vasco	2.170.163	794.313	1.759	656	± 2,3	± 3,8
La Rioja	*	*	82	25	*	*
ESPAÑA	47.129.783	17.440.800	24.775	8776	± 0,6	± 1,0

Fuente: Elaboración propia a partir de datos del INE y de la EINSFOESSA 2013

(*): CCAA sin representatividad territorial

El proceso de realización de las encuestas ha transcurrido entre el 22 de abril y el 12 de agosto de 2013. La información referida a las actividades económicas, así como la de pensiones y prestaciones económicas está referida a lo acontecido durante el año 2012.

La encuesta ha sido realizada de manera presencial, mediante una entrevista personal, y las respuestas se han recogido en un cuestionario estructurado y precodificado en el 99% de las variables. La encuesta ha sido administrada por entrevistadores previamente formados para la realización de la tarea, que han trabajado en un total de 1.051 rutas repartidas en 530 municipios de España. El trabajo de recogida de datos se ha desarrollado de manera satisfactoria, alcanzando los objetivos de tamaño y distribución muestral propuestos.

11. Ajustes posteriores al trabajo de campo

Una vez realizado el trabajo de campo se ha realizado la reponderación de la muestra. En primer lugar, se han calculado unos pesos debidos al diseño no proporcional por comunidad autónoma y probabilidad de selección de la muestra.

En segundo lugar se ha ajustado mediante calibración las características de los hogares obtenidos a las que se conocen por la Encuesta de Población Activa del INE, ya que por el tipo de muestreo el resultado podría no ser exacto. El ajuste ha sido por las variables:

- Tamaño de hogar (1, 2, 3, 4, 5 o más miembros).
- Situación laboral (todos los miembros inactivos, todos en paro, al menos uno trabaja).
- Nivel de estudios más alto de los miembros del hogar con menos de 65 años (hasta primaria o superior a primaria, o no se aplica).

Nacionalidad (todos los miembros comunitarios, todos extracomunitarios, mixtos).



FUNDACIÓN FOESSA
FOMENTO DE ESTUDIOS SOCIALES
Y DE SOCIOLOGIA APLICADA



Caritas