

OPTIMIZACIÓN DEL ERROR MUESTRAL EN ENCUESTAS A POBLACIÓN GENERAL: CRITERIOS DE ESTRATIFICACIÓN Y CÁLCULO DEL ERROR CON “WESVAR COMPLEX SAMPLES”

Manuel Trujillo Carmona

Instituto de Estudios Sociales Avanzados de Andalucía (CSIC)

RESUMEN

Cuando se realizan encuestas, el error muestral juega un papel crucial, ya que gran parte de los objetivos de la investigación suele consistir en comprobar si existen diferencias entre poblaciones o estimar determinadas magnitudes en la población. Por tanto, es muy importante rebajar este error en la fase del diseño de la muestra. Cuando la encuesta se realiza a la población general, se suele usar como unidad primaria de muestreo la sección censal. En estos casos uno de los procedimientos más útiles para reducir el error consiste en estratificar las secciones censales mediante variables que suponemos que tienen relación con las que queremos estudiar. La teoría estadística nos dice que esta estratificación permite, sin aumento de costes, disminuir el error. Sin embargo, aunque los argumentos teóricos avalan esta reducción, de nada nos sirve si no se sabe calcular dicho error. A menudo la situación es aún peor, si se utilizan programas estadísticos diseñados para tratar solo con muestras aleatorias simples. En este trabajo se presentan dos encuestas a población general en Andalucía donde se utilizó la estratificación de secciones. En una de ellas (calidad de vida de los andaluces) se utilizó una estratificación general, que usaba las principales variables sociodemográficas disponibles para las secciones censales. En la otra (intención de voto), se utilizó el voto en las anteriores elecciones. Para cada una de estas encuestas se describe el procedimiento de cálculo del error muestral utilizando el programa *Wesvar Complex Samples* y se simula cuál sería dicho error si no se hubiese utilizado la estratificación. Además, se compara este error con el que se obtendría con un muestreo aleatorio simple, lo que nos permite hallar el efecto de diseño, y se hace una estimación del número de encuestas que se pueden ahorrar mediante este procedimiento si se fija un error máximo.

Palabras clave: muestreo estratificado, error muestral, muestras complejas, jackknife.

Introducción

El objetivo de una encuesta es obtener información sobre la población objetivo del estudio con sólo dirigirse a una porción de la misma, con el coste de que dicha información no es exacta, sino que siempre existirán determinadas diferencias entre los resultados de la encuesta y los valores reales. Estas diferencias, medibles a través del error cuadrático medio, se pueden descomponer desde el punto de vista estadístico en dos factores: el sesgo y el error muestral (Cochran, 1980).

El error muestral está determinado en buena medida por el tamaño de la muestra, pero no es éste el único factor: un buen diseño puede reducir apreciablemente el error sin variar el número de entrevistas que se realizan. De hecho, la preocupación por un buen diseño muestral es constante en la planificación de una encuesta, pues recoge o es sensible a las características de ésta y de las variables que se pretenden estimar (Artés y Rueda, 1999). Igualmente, una decisión incorrecta al respecto puede generar una reacción en cadena, durante el proceso de la encuesta, que disminuya la credibilidad en las conclusiones finales (Manzano, 2000).

Una clase particular de encuestas es la que tiene como universo de referencia a toda la población de un lugar concreto, o al menos a un subgrupo muy amplio de la misma. En este tipo de encuestas, llamadas usualmente de población general, se requiere un diseño muestral complejo, especialmente cuando la unidad territorial donde se realiza la encuesta es grande. La dispersión de la muestra que conllevaría el uso de un muestreo aleatorio simple acarrearía unos costes desproporcionados y una gran dificultad en la realización del trabajo de campo, por lo que es necesario recurrir al muestreo por conglomerados. Los conglomerados o unidades primarias de muestreo (UMP) que se suelen utilizar son las secciones censales, ya que éstas son unidades de un tamaño conocido y regular y bien identificadas. Además de estas unidades se pueden conocer numerosas características, derivadas de Censos de Población y otras fuentes.

En este tipo de encuestas se puede aprovechar la información de que se dispone previamente sobre las secciones censales para plantear diseños estratificados que, agrupando en estratos las secciones censales más parecidas entre sí para las variables que se estudian, consigan una reducción del error muestral.

Un problema que se plantea tras realizar una encuesta con una muestra compleja, estratificada o no, es cómo medir el error muestral que se ha obtenido para cada estadístico, ya que los programas estadísticos más usuales solo obtienen errores para muestras aleatorias simples, y la expresión exacta de la varianza para diseños complejos a menudo es muy difícil de encontrar. Una alternativa se basa en la utilización de métodos de replicación, como el Jackknife. Este método tiene la ventaja de que ya está implementado en el programa *Wesvar Complex Samples*¹ versión 3.0, no siendo necesario para el usuario la construcción de rutinas para utilizarlo.

¹ *Wesvar Complex Samples* es propiedad de Westat Inc. (www.westat.com)

En este artículo se analizan, en primer lugar, las ventajas teóricas de la estratificación, y qué propiedades son deseables en los estratos que se utilizan, y se presentan dos diseños muestrales para encuestas a población general en Andalucía. Posteriormente se presenta un procedimiento con el que se puede medir el error muestral en este tipo de diseños, mediante el programa *Wesvar*, se calculan los errores muestrales para algunas preguntas de las dos encuestas, se simulan cuales serían dichos errores si no se hubiesen utilizado estratificaciones y se comparan los errores obtenidos con los que se obtendrían con muestreos aleatorios simples, lo que nos permite hallar los efectos de diseño. Por último, se hacen estimaciones del número de encuestas que se pueden ahorrar mediante este procedimiento si se fija un error máximo.

La estratificación

El objetivo que se persigue con la estratificación de la muestra es, como ya se ha dicho, aumentar su precisión mediante la reducción, en la medida de lo posible, del error muestral de los estimadores. Para ello, los estimadores de la variable de interés en cada estrato deben ser lo más diferentes que sea posible.

Por ejemplo, para la estimación de la media aritmética, la relación entre las varianzas de los estimadores cuando se utiliza un muestreo sin estratificar (V_{mas}) y estratificado con afijación proporcional (V_{est}), y con una población que se pueda considerar infinita, sería

$$V_{mas} = V_{est} + \frac{1}{n} \sum W_h (\bar{Y}_h - \bar{Y})^2,$$

siendo \bar{Y}_h el estimador obtenido en el estrato h , \bar{Y} el estimador para la población, y W_h el peso relativo del estrato h . (Cochran, 1980).

En consecuencia, nuestro objetivo será conseguir estratos con estimadores lo más distintos que sea posible. Ello se consigue idealmente dividiendo a la población (o los conglomerados, si el muestreo es de este tipo) según los valores de la variable que se quiere estudiar. Si no se dispone de información previa sobre dicha variable, entonces la ganancia en precisión se obtendrá si se utilizan unas variables cuya relación con ella sea lo suficientemente estrecha como para que en los estratos resultantes se obtengan estimadores alejados entre sí. La fuerza de esta relación será lo que determine la cuantía de la reducción del error.

En lo que sigue, nos situamos en las encuestas a población general, donde se utilizan como conglomerados las secciones censales. Sin embargo, toda la argumentación es válida para cualquier muestreo en donde se tenga alguna información previa sobre cómo se agrupa la población.

Algunas de las estratificaciones que se han usado tradicionalmente para este tipo de encuestas utilizan como estratos la provincia y el tamaño del municipio. Aunque siempre hay que tener en cuenta los objetivos del estudio, éstas no suelen ser las únicas variables que tienen relación con las características de la población que se

estudian. Por tanto, utilizar sólo estas características como estratos no es lo más eficiente, como se demuestra más adelante al simular los resultados que se obtendrían al utilizar esta estratificación y compararla con otra más compleja.

Por tanto, si se realiza una muestra que utiliza las secciones censales como unidades de muestreo primarias, es preferible buscar entre los datos disponibles aquellas variables que más relación tengan con lo que se quiere estudiar, y utilizarlas como variables de estratificación.

Una objeción que se suele poner a este tipo de muestreo es que, en numerosas ocasiones, una encuesta no pretende medir una sola variable, sino muchas, y por tanto no estaría claro cuál debería ser la variable por la cual se debe estratificar. Sin embargo, casi siempre todas esas variables estarán relacionadas con la posición social en general del individuo —¿Qué cuestión de las que se investiga en sociología no guarda alguna relación con la edad, el estatus económico o el nivel de estudios del sujeto?—. Por tanto, una estimación de la posición social media de los individuos en las secciones censales puede ser una buena variable para estratificar si no se tiene otra información, en estudios de este tipo.

Además de la mejora de los niveles de error, existe otra ventaja adicional del tipo de estratificación que aquí se propone, relacionado con el trabajo de campo: al existir una clasificación de las unidades geográficas, que se puede incorporar al fichero de datos, es posible estudiar las características de las zonas en donde se presente algún comportamiento especial: tasas de no-respuesta alta, sesgos en estimadores cuyo valor se conoce a priori, o cualquier otra incidencia. Al saberse en qué tipo de zonas ocurren estos problemas, se pueden prever sus soluciones en muestreos posteriores. Por ejemplo, si se obtienen bajas tasas de respuesta en zonas con una determinada tipología urbanística y sociológica —como viviendas unifamiliares de familias con estatus alto—, es posible diseñar estrategias para aumentar la respuesta en las zonas de esa clase antes de empezar el muestreo.

Un último problema es cómo reunir la información proporcionada por varias variables de forma que sirva para realizar la estratificación. La técnica más utilizada, cuando se forman estratos de dos o hasta tres variables, consiste en dividir cada variable por separado en grupos (normalmente pocos), y después obtener los estratos finales mediante la combinación de los grupos de todas las variables.

Sin embargo, para utilizar en la estratificación un número de variables que no sea muy reducido, es imposible pensar en recurrir a estratos cruzados de las variables, ya que si el número de éstas utilizado en la estratificación es superior a dos o tres, obtendríamos un número muy alto de estratos cruzados, de forma que podrían no contener ni siquiera una unidad muestral. En cambio, es muy útil utilizar el análisis de *cluster* o de conglomerados, ya que mediante este análisis se obtienen grupos (de secciones, en nuestro caso) que maximizan la heterogeneidad inter-grupos y minimizan la heterogeneidad intra-grupos.

Ejemplos de estratificaciones

Sin entrar en muchos detalles sobre los procedimientos de estratificación, ya que no es éste nuestro objetivo en estos momentos, sí referiremos en términos generales cómo se han realizado éstos.

Caso 1: Encuesta para la estimación del comportamiento electoral

En este caso, se trata de una encuesta cuyo objetivo principal fue medir el comportamiento electoral de los ciudadanos ante unas hipotéticas elecciones en Andalucía.

Dado que están disponibles los resultados obtenidos en anteriores elecciones en cada sección, estamos en el caso más favorable, ya que se puede utilizar como variable de estratificación una variable que, si no es la misma que queremos medir (ya que el voto puede cambiar entre elecciones), sí cuenta con una alta relación con la variable de interés (Espinosa, 2000; Luque, 2000).

En realidad, el comportamiento electoral en cada sección no es una variable simple, sino compleja, ya que para resumirlo es necesario hacer uso de varias variables. Aunque se podrían definir indicadores complejos que resumiesen dicho comportamiento, lo más sencillo es tomar como variables la tasa de abstención sobre el total de electores y las proporciones de voto a los principales partidos (PSOE, PP, PA, IU-CA), sobre votos emitidos².

Dado que no se dispone de una variable simple que ordene las secciones —y, por tanto, se pueda utilizar como criterio único de estratificación—, los estratos deben formarse combinando las variables de forma que sean lo más homogéneos posible en el conjunto de ellas. Estos grupos se pueden obtener mediante análisis de conglomerados. Utilizando esta técnica, se obtuvieron 10 conglomerados.

En conjunto, la estratificación se realiza en dos fases. En primer lugar, se realiza una estratificación provincial, con afijación proporcional al número de escaños que le corresponden a cada provincia en la cámara autonómica, repartiendo así el número de encuestas según su importancia electoral y no poblacional. Dentro de cada provincia, la afijación de la estratificación por grupos electorales es proporcional a su población mayor de 18 años, y la selección de las secciones dentro de cada estrato *provincia-grupo electoral* es proporcional también a su población mayor de 18 años.

² También habría otras, como los votos en blanco, nulos, o a pequeños partidos, pero dado que a dichos comportamientos se les concede poca importancia en el estudio, se prescinde de ellos al realizar la estratificación

Caso 2: Encuesta sobre la calidad de vida en Andalucía

Así como en el caso anterior estaba claro cuál era la variable a estudiar, en este caso esto no estaba tan claro, dado que en la calidad de vida se pueden incluir numerosos factores. Por ello, para optimizar el error de muestreo, se decidió hacer una estratificación que contemplara todas las variables disponibles que tuvieran alguna relación con la calidad de vida. Entre otras, se encuentra el tipo de hábitat, junto con 30 variables socioeconómicas obtenidas del censo de población y viviendas de 1991.

Con este gran conjunto de variables no es aconsejable utilizar directamente el análisis de conglomerados. En su lugar, se dividieron las secciones por tipos de hábitat y después, en cada grupo, se realizó un análisis de componentes principales con las variables censales. Con los principales factores obtenidos se realizó un análisis de conglomerados. Como resultado, se obtuvieron 28 grupos de secciones³.

Métodos de replicación para el cálculo del error de muestreo: el método JK_n

La idea fundamental de la replicación es facilitar el cálculo de la varianza seleccionando un conjunto de submuestras replicadas en lugar de una sola muestra. Cada submuestra debe ser seleccionada independientemente y con un sistema de selección idéntico. Entonces se calcula un estimador para cada submuestra, y la varianza muestral del estimador total puede ser estimada en base a la variabilidad de los estimadores de las submuestras. Las submuestras se llaman réplicas y los datos estadísticos calculados de las réplicas se llaman estimaciones replicadas.

Existen diversos métodos de replicación, entre los que se incluyen los métodos Jackknife, replicación de equilibrio repetido (*Balanced Repeated Replication*), y métodos Bootstrap. La elección del método depende del diseño muestral en gran medida y también de las posibilidades que nos ofrezca el software disponible.

En cuanto al software, existen varias posibilidades, como lo son: Stata, Sudaan, SAS y Wesvar (Cohen, 1997). En el caso de SAS (a partir de la versión 8), no se utiliza la replicación, sino otro método alternativo, la linealización (An, Watts y Stokes, 1999).

En nuestro caso, el software utilizado ha sido *Wesvar*, disponible en el Instituto de Estudios Sociales Avanzados. Además, este software tiene las ventajas de que está presentado en un entorno bastante amigable y fácil de usar, y el proceso de conversión de ficheros procedentes de otros paquetes, como SPSS, es bastante sencillo y rápido.

³ Ejemplos de estratificaciones similares se pueden encontrar en Lozares y López, 1990 y en Trujillo, 1998.

Entre los métodos de estimación que presenta Wesvar, el único que permite estimar las varianzas para diseños donde el número de conglomerados por estrato es variable es la variante del Jackknife llamada JK_n.

En el método JK_n, para formar cada réplica se excluye una sola unidad muestral primaria (UMP) de un solo estrato. Por tanto, el número de réplicas que será necesario realizar para cada estrato será igual al número de UMP que existan en el estrato.

Si para cada estrato h tenemos n_h UMP, también será n_h el número de réplicas que formaremos para dicho estrato. Si llamamos u_{hi} al estimador de U para el estrato h y la réplica i -ésima, la varianza es estimada por

$$v(\bar{u}) = \sum_h^L \left(\frac{(n_h - 1)}{n_h} \sum_i^{n_h} (u_{hi} - \bar{u})^2 \right) \quad (\text{Lee, Forthofer y Lorimor, 1989}).$$

Resultados

Verificamos, a continuación, la bondad del uso de la estratificación a propósito, calculando el error de muestreo obtenido en algunas preguntas de las dos encuestas expuestas anteriormente.

Además del error típico, se muestra también el efecto del diseño, que es el cociente entre la varianza obtenida con el muestreo realizado y la varianza que se obtendría con un muestreo aleatorio simple del mismo tamaño muestral (Kish, 1965).

También se indica el tamaño de muestra efectivo: una estimación del número de encuestas que se habrían necesitado con un muestreo aleatorio para obtener el mismo error típico obtenido. De esta manera se conoce la ganancia o pérdida en encuestas que se obtiene con los distintos diseños muestrales y también una estimación del dinero que se gana con este muestreo, si se multiplica el número de encuestas útiles realizado de más por el coste de realización de cada encuesta.

Procedimiento de utilización de Wesvar

Abreviadamente, el proceso que es necesario seguir en la utilización del paquete es el siguiente:

En primer lugar, se selecciona la fuente de datos. De ésta se seleccionan las variables que son de interés, tanto para calcular estimadores, como para servir de variables de control. También se seleccionan las variables necesarias para especificar el diseño muestral y, caso de que exista, es posible especificar una variable de ponderación general de la muestra.

El segundo paso consiste en especificar el diseño de la muestra, mediante la identificación de las variables que definen los estratos y los conglomerados.

En el diseño de la muestra también se puede especificar la existencia de UMP autorrepresentadas, es decir, cuya presencia en la muestra está garantizada, bien por su tamaño, o bien porque sea de especial interés su inclusión. Igualmente, se puede incluir un coeficiente de corrección para el caso de que la población de UMP no pueda considerarse infinita.

Finalmente, Wesvar ofrece una serie de facilidades para el perfeccionamiento de la muestra obtenida a posteriori: la postestratificación, y el ajuste por ponderaciones —*raking*— (Morganstein y otros, 1998).

Resultados para el caso 1

En este caso, vamos a utilizar como estratos las clases cruzadas de provincias y grupos de conglomerados (Cochran, 1980). En el método JK_n se requiere que en todos los estratos existan al menos dos UMP, por lo que en los casos en que solo haya uno, se unirán las correspondientes a distintas provincias.

Como la afijación provincial no es proporcional al tamaño de la población, se introduce la ponderación correspondiente para restituir la proporcionalidad.

Para comparar los resultados con los que se obtendrían en caso de que no se hubiese realizado la estratificación, calcularemos los mismos resultados utilizando como estrato únicamente la provincia, y también simularemos los resultados que se obtendrían si se hubiese utilizado un muestreo aleatorio simple.

Al modelo con estratificación provincia - grupo electoral, lo denotamos con *Dis. A*, y al modelo con estratificación únicamente provincial, *Dis. B*. Los resultados se muestran en las tablas 1 y 2.

Hemos elegido para verificar los resultados en la encuesta la variable de mayor interés, es decir, la intención de voto, y la variable más cercana a la que se utilizó para realizar la estratificación, el recuerdo de voto.

Tabla 1: *Resultados para la variable "intención de voto".*

	Error típico			Efecto del diseño		Tamaño efectivo de muestra		
	M.A.S.	Dis. A	Dis. B	Dis. A	Dis. B	M.A.S.	Dis. A	Dis. B
IU-CA	0,411	0,419	0,435	1,037	1,114	244	235	219
PA	0,342	0,373	0,392	1,188	1,313	159	134	121
PP	0,642	0,724	0,791	1,273	1,521	702	552	461
PSOE	0,796	0,908	1,020	1,300	1,641	1387	1067	845
Otros	0,114	0,126	0,126	1,211	1,215	18	15	15
En blanco	0,203	0,225	0,218	1,232	1,165	57	46	49
No votará	0,353	0,392	0,394	1,236	1,25	169	137	135
Indeciso	0,573	0,731	0,724	1,625	1,598	544	335	340
Total						3280	2521	2186

Una primera acotación que se puede realizar en este momento es la visualización de la desviación en que se incurre si se calculan los errores muestrales asumiendo un muestreo aleatorio simple, lo que, aunque es a todas luces equivocado, aún hoy se realiza muy frecuentemente. Estas desviaciones, como se ve, pueden llegar a ser de más del 50%.

Los resultados son semejantes en ambas preguntas: tanto el error típico como el efecto del diseño son menores para el diseño A en casi todas las respuestas de ambas variables, siendo las diferencias mayores en la pregunta de recuerdo de voto, al ser ésta la más relacionada con la variable de estratificación.

La prueba definitiva de la utilidad de la estratificación se obtiene comparando las diferencias de tamaño efectivo de muestra. Para la pregunta de mayor interés las encuestas efectivas que se han realizado de más son 335, más de un 10% de las realizadas, y que hubiese supuesto la realización de 407 encuestas reales más para obtener el mismo resultado. Teniendo en cuenta el coste medio por encuesta, que en este caso podría ser de unas 2.000 pts, nos encontramos con que hemos podido ahorrar unas 800.000 pts., cifra nada despreciable.

Tabla 2: Resultados para la variable "recuerdo de voto".

	Error típico			Efecto del diseño		Tamaño efectivo de muestra		
	M.A.S.	Dis. A	Dis. B	Dis. A	Dis. B	M.A.S.	Dis. A	Dis. B
IU-CA	0,426	0,442	0,472	1,076	1,232	259	241	210
PA	0,291	0,304	0,312	1,089	1,147	119	109	104
PP	0,614	0,681	0,737	1,231	1,442	631	513	438
PSOE	0,767	0,744	0,850	0,940	1,225	1207	1284	985
Otros	0,145	0,145	0,147	0,996	1,021	30	30	29
Abstención	0,546	0,586	0,587	1,153	1,159	450	390	388
Sin edad	0,460	0,451	0,451	0,963	0,962	306	318	318
Total						3002	2885	2472

Resultados para el caso 2

En este caso, los estratos son las clases cruzadas entre provincias y grupos de secciones, y las UMP son las secciones. Al igual que en el caso anterior, en los casos en que no existan dos UMP en un estrato, es necesario colapsar los estratos. Para comparar, utilizaremos los resultados que aparecen si se supone que la estratificación realizada ha sido la que utiliza la provincia y el tamaño de hábitat, de frecuente utilización. Éste será el diseño B.

Tabla 3: Resultados para la variable "satisfacción con la vivienda".

	Error típico			Efecto del diseño		Tamaño efectivo de muestra		
	M.A.S.	Dis. A	Dis. B	Dis. A	Dis. B	M.A.S.	Dis. A	Dis. B
Muy satisfecho	0,886	1,208	1,305	1,860	2,170	562	301	259
Bastante satisfecho	1,015	1,275	1,361	1,579	1,798	1384	876	767
Poco satisfecho	0,455	0,543	0,543	1,422	1,421	121	85	85
Ni satisfecho ni insatisfecho	0,444	0,507	0,522	1,303	1,383	121	93	87
Bastante insatisfecho	0,445	0,520	0,527	1,364	1,398	119	87	85
Muy insatisfecho	0,270	0,291	0,317	1,163	1,387	44	38	32
Total						2351	1480	1315

Tabla 4: Resultados para la variable "¿Cómo se siente con su vida en general?".

	Error típico			Efecto del diseño		Tamaño efectivo de muestra		
	M.A.S.	Dis. A	Dis. B	Dis. A	Dis. B	M.A.S.	Dis. A	Dis. B
Muy a disgusto	0,175	0,198	0,198	1,273	1,266	17	13	13
A disgusto	0,280	0,335	0,339	1,434	1,470	43	30	29
Más bien insatisfecho	0,454	0,483	0,481	1,133	1,126	125	110	111
Ni satisfecho ni insatisfecho	0,613	0,700	0,700	1,305	1,305	233	179	179
Más bien satisfecho	0,894	1,069	1,154	1,430	1,666	615	430	369
A gusto	0,998	1,122	1,162	1,263	1,352	935	741	691
Muy a gusto	0,760	0,947	1,015	1,554	1,787	393	253	220
Neutral	0,159	0,224	0,223	1,991	1,960	14	7	7
Nunca ha pensado en ello	0,119	0,114	0,111	0,919	0,872	9	10	10
Total						2384	1773	1629

Para ilustrar los resultados, se han escogido dos variables de las muchas que contiene el cuestionario, que no coinciden además con ninguna de las variables que se han utilizado en la estratificación. A pesar de ello, existe una innegable relación entre estas variables, de tipo subjetivo, y las características generales de las personas que responden, que son las variables que nos han servido para la estratificación. Por tanto, es de esperar que los resultados sean mejores con el diseño estratificado que sin estratificar.

Como se muestra en las tablas 3 y 4, los errores típicos son casi siempre más bajos en el diseño A que en el diseño B y, por tanto, los efectos de diseño menores. Ello lleva a una mayor muestra efectiva en el caso del diseño A, y por tanto a un ahorro cuantificable si se planifica la muestra con un error a priori.

Este ahorro es, en el primer caso, de 165 entrevistas, y en el segundo de 144. Es decir, cerca del 10% de la muestra efectiva. En el caso menos bueno de los dos, el del ahorro de 144 encuestas, habríamos necesitado 211 encuestas reales. Con un coste aproximado de 2.500 pts por encuesta, el ahorro se podría evaluar en unas 500.000 pts.

Conclusiones

En este artículo hemos tratado de exponer y comprobar los beneficios de la realización de estratificaciones específicas para las muestras, cuando se realizan encuestas a población general. Para ello, en primer lugar, se han señalado las ventajas teóricas que se obtienen de la realización de un muestreo estratificado; la teoría nos indica que las variables de estratificación deben ser lo más cercanas posible a las variables que son objeto de estudio. Además, se puede destacar otro tipo de ventajas, relacionadas con la facilidad de control del trabajo de campo, derivadas de un conocimiento más cercano de la población a la que se dirige.

En segundo lugar, se ha expuesto un método para comprobar en qué grado la disminución del error muestral que la teoría estadística nos demuestra aparece en una encuesta ya realizada: el cálculo del error muestral mediante replicación, con ayuda del programa Wesvar.

En tercer lugar, se han aplicado estas técnicas a dos encuestas realizadas por el Instituto de Estudios Sociales Avanzados de Andalucía, y efectivamente se ha comprobado la reducción de error que conlleva la utilización de este tipo de estratificación, en un caso frente a la ausencia de estratificación, y en otro frente a una estratificación más simple, compuesta por provincia y tamaño municipal. Además, se presentan otros indicadores, como el número efectivo de encuestas que se gana con la utilización de estos muestreos.

Una primera conclusión que se obtiene de los resultados es la necesidad de utilizar estimadores del error muestral que tengan en cuenta el diseño de la muestra, ya que si se calculan suponiendo que se hubiera realizado un muestreo aleatorio simple, la magnitud obtenida puede ser hasta un 50% de la real.

Una vez asumida la necesidad de utilizar estimadores realistas del error muestral, es inexcusable su reducción si ello es posible: la realización de una encuesta es

un proceso caro y muy delicado, por tanto es necesario incorporar toda la información que se tenga disponible, incluso desde el momento previo a la realización del muestreo, ya que de ello se deriva una ganancia en costes, evaluable para nuestros ejemplos en más de medio millón de pesetas, y en calidad de la encuesta.

Referencias

- An, A., Watts, D. y Stokes, M. (1999). Procedimientos SAS para el análisis de datos de encuestas. *Estadístico de encuestas*, 41: 10-13.
- Artés, E.M. y Rueda, M.M. (1999) Aportaciones al muestreo sucesivo. *Metodología de Encuestas*, 1(1) , 19-28.
- Cochran, W. G. (1980). *Técnicas de muestreo*. CECSA: México, D.F.
- Cohen, S.B. (1997). An evaluation of alternative PC-based software packages developed for the analysis of complex survey data. *American Statistician*, 51: 285-293.
- Espinosa, L. (2000) Distintas alternativas sobre la capacidad predictiva de las encuestas preelectorales. *Metodología de Encuestas*, 2(1), 148-150.
- Kish, L. (1965). *Survey sampling*. New York: John Wiley & Sons.
- Lee, E. S., Forthofer, R. N. y Lorimor, R. J. (1989). *Analyzing complex survey data*. Sage University Paper series on Quantitative applications in the Social Sciences, series nº 07-071. Beverly Hills: Sage Pubns.
- Lozares Colina, C. y López Roldán, P. (1990). Enquesta de la Regió Metropolitana de Barcelona 1990. Document de treball número 90/1. Construcció de la mostra estratificada. Barcelona: Institut d'Estudis Metropolitans de Barcelona.
- Luque, E. (2000) Sesgos e interpretaciones, la asignatura pendiente de las encuestas pre-electorales. *Metodología de Encuestas*, 2(1), 150-153.
- Manzano, V. (2000) Editorial. *Metodología de Encuestas*, 2(1) i-iv.
- Morganstein, D., Brick, J.M., Broene, M. y Nixon, M. G. (1998). El método de replicación para la estimación de errores de muestreo. Seminario Internacional de Euskadi, 37. Vitoria: Eustat.
- Trujillo Carmona, M. (1998). *Construcción y descripción de una tipología de secciones en Andalucía*. Documento de trabajo. Córdoba: Instituto de Estudios Sociales Avanzados de Andalucía.