INFERENCIA ESTADISTICA, NIVELES DE PRECISION Y DISEÑO MUESTRAL

Jacinto Rodríguez Osuna, María Luisa Ferreras y Adoración Núñez
Centro de Investigaciones Sociológicas

1. INTRODUCCION

En el campo de la investigación empírica es frecuente el recurso a la encuesta muestral cuando no se dispone de fuentes secundarias fiables referentes al tema que interesa estudiar. En estos casos la encuesta se presenta, casi siempre, como la única alternativa posible, ya que la realización de un censo resultaría inviable no sólo por sus elevados costes, sino también por lo prolongado del proceso de recogida y tratamiento de la información referida a todo el universo.

Por otra parte, con la encuesta se puede obtener mayor exactitud que con el censo y, además, los niveles de precisión que sean necesarios para los fines de la investigación.

La mayor exactitud[1] suele provenir del empleo de personal más cualificado en la recogida de la información mediante encuesta, debido a que se requiere un número muy inferior de agentes entrevistadores que de agentes

[1] Un ejemplo de lo que se viene diciendo es la Encuesta de Población Activa, cuyas estimaciones sobre actividad, ocupación y desempleo son más exactas y precisas que las que se derivan del censo. Esto es así porque la recogida de la información se hace por personal cualificado, hecho posible cuando el número de sujetos a entrevistar es limitado.

censales y, en consecuencia, se puede controlar más el proceso y seleccionar y preparar mejor al personal. Los niveles de precisión requeridos están en función de los fines de la investigación, y su consecución es un problema técnico que se traduce en la aplicación del diseño muestral adecuado a cada universo objeto de investigación.

Para poder garantizar la precisión de las estimaciones que han de dar paso a la inferencia estadística, «la muestra ha de constar de un número suficiente de elementos, elegidos al azar, tal que proporcione una seguridad estadística de que los resultados que se obtengan de ella puedan, dentro de los límites estimados, representar realmente al universo»[2]. Quiere esto decir que para que se pueda hacer inferencia estadística es necesario que se trate de muestreos aleatorios, adecuadamente dimensionados.

Se ha de trabajar con muestreos aleatorios porque son los únicos en los que se puede fijar el nivel de confianza y calcular los errores de muestreo[3]. El nivel de confianza expresa la probabilidad de acertar en la estimación, y los errores de muestreo indican la bondad de la misma. Cuando se trabaja con datos que provienen de encuesta, porcentajes, medias, totales, etc., no basta con ofrecer los resultados, sino que, además, es imprescindible conocer el nivel de confianza con que se está operando y la precisión de las estimaciones. Estos cálculos son posibles cuando se realizan muestreos aleatorios, a los que se aplica la teoría de probabilidades.

Las muestras han de estar adecuadamente dimensionadas porque el nivel de confianza y la precisión de las estimaciones guardan estrecha relación con el tamaño muestral. Un mayor nivel de confianza garantiza una mayor probabilidad de acertar, pero sólo se consigue aumentando el número de elementos de la muestra. Por su parte, la precisión de las estimaciones guarda relación inversa con el error muestral. Cuanto menor sea éste, menor es la dispersión de la distribución del estimador y, consiguientemente, la precisión es mayor.

Para conseguir esta menor dispersión es necesario aumentar convenientemente el número de elementos de la muestra hasta obtener los niveles de precisión requeridos. ¿Cuál es este nivel y cuál el nivel de confianza con que se debe trabajar? Son preguntas a las que se intentará responder conforme se aborden los correspondientes epígrafes.

Independientemente de lo anterior, hay otra serie de factores que afectan al diseño muestral y que inciden en el tamaño de la muestra. Se hace referencia a la estructura del universo y a los niveles de desagregación de las estimaciones. La estructura del universo, en la medida en que sea

[2] Restituto SIERRA BRAVO, *Técnicas de Investigación Social. Teoría y ejercicios,* Madrid, Paraninfo, 1985, p. 181.

[3] Existen muestreos no aleatorios cuyos resultados, en la práctica, son muy aceptables. Sin embargo, estadísticamente, no se puede calcular la precisión de sus estimaciones.

conocida, puede permitir aproximarse a la varianza de las variables que se pretende estimar y a la composición del mismo. Así se podrá afinar más en el cálculo del tamaño de la muestra y utilizar el tipo de diseño más adecuado. Por su parte, los niveles de desagregación de los resultados obligarán a dimensionar la muestra de forma que las estimaciones sean válidas no sólo a nivel global, sino también para determinadas subpoblaciones que interesa investigar.

El trabajo que se presenta a continuación aborda estas cuestiones en un intento de explicar cómo afecta cada una de ellas al diseño muestral y a la precisión de las estimaciones. Para mayor claridad, las explicaciones teóricas irán acompañadas de un ejemplo: el diseño de una hipotética muestra a funcionarios civiles en el que se introducirán sucesivamente los distintos factores que se analizan en el texto.

2. NIVEL DE CONFIANZA

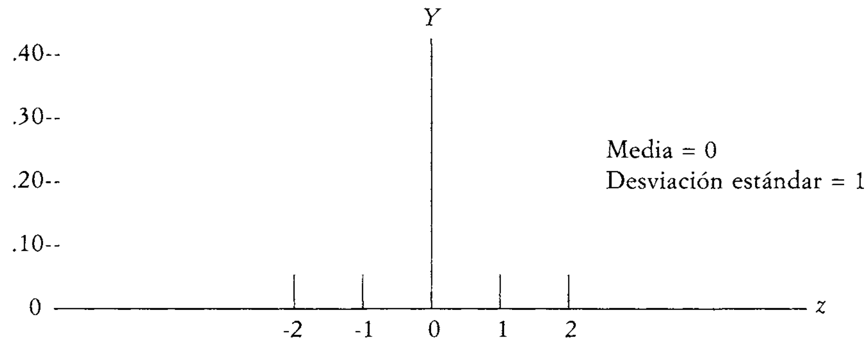
El nivel de confianza y el nivel de precisión son conceptos estrechamente relacionados que acotan la validez de la estimación. El primero hace referencia a la probabilidad de acertar, y el segundo refleja los errores de muestreo, es decir, la distribución del estimador. Dicho de otra forma, el primero indica la probabilidad de que el segundo, los errores de muestreo, no rebasen determinados límites. Cuando se dice, por ejemplo, que para un nivel de confianza del 95,5 por 100 el error de muestreo es de ±2 por 100, lo que se está expresando es que hay una probabilidad del 95,5 por 100 de que el valor real que se trata de estimar se encuentre dentro de los límites definidos por la estimación y el error de muestreo[4]. Indica también que existe la probabilidad del 1-0,955, es decir, del 4,5 por 100, de que la estimación rebase el error señalado.

Aunque la precisión de las estimaciones será objeto del siguiente epígrafe, se la menciona aquí por su relación con el nivel de confianza y porque el fundamento de uno y otro concepto hay que buscarlo en la teoría de probabilidades y, más en concreto, en la función de probabilidades que sigue una distribución normal[5], que se define por su media y su desviación estándar, cuya representación gráfica aparece en la página siguiente:

En el eje de abcisas se encuentra cada uno de los valores de la distribución, y en el de ordenadas, la probabilidad de obtenerlos. El área de la curva es igual a la unidad, ya que en la curva se encuentran el cien por cien

[4] Los límites de la estimación se delimitan añadiendo a la misma ± el error de muestreo multiplicado por el número de desviaciones estándar que se toma y que indica el nivel de confianza con que se está operando (véase cuadro 1).

[5] La distribución normal no es la única posible pero sí la más frecuente. Por eso aquí sólo se habla de este tipo de distribución de frecuencias.



de los casos. Por otra parte, al ser simétrica respecto a la media, cada una de las áreas a la derecha o a la izquierda de la misma vale 0,5.

Tal como se observa en el gráfico, en la distribución normal la mayor probabilidad se agrupa en torno a la media, de tal forma que entre ésta y ± una desviación estándar se da el 0,6826 de probabilidad; entre la media y ± dos desviaciones estándar, el 0,9544, etc. Más en concreto, con la ayuda de tablas se puede calcular el área de la curva desde la media hasta la desviación estándar. Los resultados para diferentes valores se reflejan en el cuadro 1. Traducido al campo que nos ocupa, quiere decir que si se acota la probabilidad en 0,5 desviaciones, hay un 38,30 por 100 de garantía de que el error absoluto no sea superior al 0,5×error de muestreo[6].

CUADRO 1

Probabilidad y errores de muestreo

Desviación	*Probabilidad derecha media*	*Error total (dcha.-izqda.)*	*Absoluto*
0,5	0,1915	0,3830	0,5 × e
1,0	0,3413	0,6826	1,0 × e
1,5	0,4332	0,8664	1,5 × e
2,0	0,4772	0,9544	2,0 × e
2,5	0,4938	0,9876	2,5 × e
3,0	0,4987	0,9974	3,0 × e

FUENTE: Elaboración propia.
NOTA: Por *e* se designa el error de muestreo.

[6] El cálculo del error de muestreo es muy sencillo si se parte del muestreo aleatorio simple. En este caso, tratándose de universos grandes, el error de las proporciones se calcula con la fórmula $e = \sqrt{\frac{P\,(1-P)}{n}}$

También quiere decir que hay un (1–0,3830=0,6170) 61,70 por 100 de probabilidad de que el error supere el cálculo anterior. Dado que el riesgo de equivocarse es elevadísimo, generalmente se opera en Ciencias Sociales con un nivel de confianza igual a dos desviaciones, lo que corresponde a una probabilidad del 95,44 por 100. Esto significa que de 100 muestras que se diseñaran adecuadamente, la probabilidad de que la estimación no tenga un error absoluto superior a 2×error de muestreo es del 95,44 por 100.

Antes de seguir adelante se quiere señalar que la teoría anterior, basada en la distribución normal, es aplicable a las muestras en general, bien porque el universo del que se extraen se distribuye normalmente, bien porque, según el teorema del límite central, a medida que aumenta el tamaño de la muestra, la distribución de la media de una muestra aleatoria, extraída prácticamente de cualquier población, tiene una distribución aproximadamente normal. En consecuencia, en muestras grandes[7] es de aplicación la distribución normal, por lo que se puede proceder a la inferencia estadística, dado que se puede conocer la distribución y calcular los errores que corresponden a diferentes niveles de confianza.

La incorporación del nivel de confianza al cálculo de la muestra se realiza incluyendo en la fórmula general un parámetro *k* que hace referencia a la desviación estándar[8]. A partir de aquí la fórmula del cálculo del tamaño de la muestra, con el nivel de confianza incorporado, para el muestreo aleatorio simple[9] es:

$$n = \frac{NK^2p(1-p)}{(N-1)e^2 + K^2p(1-p)}$$

Dicha fórmula se simplifica si se trata de universos grandes[10]:

$$n = \frac{K^2p(1-p)}{e^2}$$

[7] Según Wonnacott, «cuando el tamaño de la muestra *n* es de aproximadamente 10 ó 20, la distribución de la media generalmente es casi normal». Thomas H. WONNACOTT y Ronald J. WONNACOTT, *Introducción a la estadística,* México, Limusa, 1979, p. 143.

[8] Este parámetro designado aquí como *k* es denominado de diferentes maneras, según los autores.

[9] Las fórmulas varían en función del tipo de muestreo que se utilice y, también, del parámetro que se desee estimar. Aquí se trata de estimar proporciones.

[10] Son universos grandes, a efectos de los cálculos, aquellos en los que el porcentaje que representa la muestra con respecto al universo sea pequeño, generalmente inferior al 5 por 100. Como señala Cochran, «el factor CPF (Corrección de Poblaciones Finitas) se puede ignorar siempre y cuando la fracción de muestreo no exceda el 5 por 100 y para muchos propósitos, aun si es tan alto como el 10 por 100». William G. COCHRAN, *Técnicas de muestreo,* México, Compañía Editorial Continental, 1977, p. 49.

Aplicando esta última fórmula a distintos niveles de confianza, para un mismo error muestral, los resultados son éstos:

— Nivel de confianza *k=1* y probabilidad de 68,26 por 100:

$$n = \frac{1^2 \cdot 0{,}50 \cdot 0{,}50}{0{,}02^2} = 625 \text{ entrevistas}$$

— Nivel de confianza *k=2* probabilidad de 95,44 por 100:

$$n = \frac{2^2 \cdot 0{,}50 \cdot 0{,}50}{0{,}02^2} = 2.500 \text{ entrevistas}$$

La incidencia del nivel de confianza sobre el tamaño de la muestra es elevada, ya que, como se observa en el ejemplo anterior, el pasar de una probabilidad —garantía de acertar— del 68,26 al 95,44 por 100 implica cuadruplicar el tamaño de la muestra.

3. TAMAÑO DE LA MUESTRA Y PRECISION DE LAS ESTIMACIONES

Tamaño de la muestra y precisión de las estimaciones son conceptos inseparables. Si aumenta el tamaño de la muestra, *caeteris paribus*, lo hace el nivel de precisión, y se quiere aumentar la precisión de las estimaciones es necesario modificar el tamaño de la muestra. Dicho de otra forma, tamaño de la muestra y precisión de las estimaciones guardan estrecha relación y cualquier modificación de una de estas variables se refleja, además en la misma dirección, en la otra, aunque en distinta proporción, como se verá más abajo.

La precisión hace referencia a la concentración de los valores estimados en torno al valor que se trata de estimar, de tal forma que la distancia entre el valor a estimar y el valor estimado sea pequeña. Dicho en términos estadísticos, la precisión refleja la escasa dispersión de la distribución del estimador en el muestreo y se da cuando la magnitud de las desviaciones respecto a la media, obtenida por la reiteración del procedimiento de muestreo, es pequeña. La precisión se mide por la varianza del estimador o su raíz cuadrada, la desviación estándar, que dan cuenta de la dispersión de la distribución. De ahí que a la desviación estándar se la llame error de muestreo, concepto fundamental para determinar la mayor o menor precisión de las estimaciones y, consecuentemente, el tamaño de la muestra que corresponde a los diferentes niveles de precisión. En concreto, y para el caso de muestreo aleatorio simple y estimación de proporciones, la fórmula que liga el nivel de precisión con el tamaño de la muestra es la [1] y, si se parte de universos grandes[11], la [2]. Estas fórmulas son las utilizadas al hablar del nivel de confianza. Allí se prestaba atención al parámetro *K* y aquí al *e* (error de muestreo). Aplicando ambas fórmulas a nuestra muestra de funcionarios, el tamaño de la misma que se deduce para un error de 0,02, una proporción 0,5-0,5 y *k* igual a 2 es, en el primer caso:

$$n = \frac{2^2 \cdot 0{,}50 \cdot 0{,}50}{212.429 \cdot 0{,}02^2 + 4 \cdot 0{,}5 \cdot 0{,}5} = 2.470 \text{ entrevistas}$$

y en el segundo:

$$n = \frac{2^2 \cdot 0{,}50 \cdot 0{,}50}{0{,}02^2} = 2.500 \text{ entrevistas}$$

Ambos resultados son casi iguales porque se está trabajando con un universo grande[12] y de ahí que la eliminación de la CPF no afecta, prácticamente, al tamaño muestral. Si en lugar de un error del 0,02 se necesitara operar con un error sensiblemente inferior, del 0,01, por ejemplo, el tamaño de la muestra debería ser 10.000, es decir, cuatro veces mayor. Tal como dice Azorín Poch, en el caso del muestreo aleatorio simple, «para obtener un error mitad (o sea, doble precisión) haría falta una muestra no doble, sino cuatro veces mayor»[13].

Los cálculos anteriores permiten establecer la relación entre nivel de precisión (error de muestreo) y tamaño de la muestra, quedando patente que el aumento de la precisión repercute considerablemente en el tamaño de la muestra. De ahí que la determinación del nivel de precisión sea, en definitiva, uno de los requisitos imprescindibles para poder calcular el tamaño de la muestra.

El nivel de precisión requerido depende, en primer lugar, de la finalidad del estudio, ya que no es lo mismo estimar la aceptación por la población de un determinado producto comercial que el IPC. En el primer caso, una deficiente estimación puede repercutir negativamente en la política comercial de una empresa, mientras que una estimación poco fiable del IPC tiene consecuencias sobre la inflación, los salarios y, en general, afecta

[11] Véase nota 10.

[12] Véase la nota 10. Independientemente de este planteamiento, se suelen considerar universos grandes a aquellos que constan de más de 100.000 elementos.

[13] F. AZORÍN POCH, *Curso de muestreo y aplicaciones,* Madrid, Aguilar, 1972, p. 100.

a la política económica del país. En consecuencia, el nivel de precisión —error permitido— deberá guardar relación con la importancia del fenómeno que se desee estimar.

Independientemente de lo anterior, la precisión está relacionada con el valor numérico de la estimación. En nuestro ejemplo de funcionarios se ha estimado que el 61 por 100 ha ingresado en la Administración por oposición y el 18 por 100 por contrato y oposición restringida, con un error de muestreo del 0,02. Este error, expresado en términos relativos, arroja los siguientes resultados: $\frac{0{,}02}{0{,}61} = 0{,}033$ y $\frac{0{,}02}{0{,}18} = 0{,}11$. En el primer caso, el error representa el 3 por 100 de la estimación, mientras que en el segundo supone el 11 por 100. A partir de aquí se puede profundizar en la relación precisión-tamaño de la muestra, con un nuevo planteamiento en el que se sustituyen los errores de muestreo por errores relativos —coeficiente de variación—, estableciendo la relación entre los mismos y el tamaño muestral. Con este planteamiento es posible ajustarse más a los intereses de una investigación rigurosa.

El punto de partida para este enfoque es el coeficiente de variación que expresa, en términos relativos, la desviación típica ponderándola para el valor medio a que hace referencia. Su expresión matemática es: $CV = \frac{\sigma}{\bar{x}}$ lo que hace que, una vez conocido el valor de la desviación típica y del parámetro, los resultados se puedan expresar en términos relativos comparables. En el campo de la estadística inferencial se utiliza para expresar los errores de muestreo en porcentajes de la estimación. Su expresión matemática aquí es la misma, sustituyendo el error por la desviación (*e* por σ) y dando a la $\bar{x}$ el valor que corresponda, según se trate de aplicarlo a medias, totales o proporciones.

El coeficiente de variación permite cuantificar mejor la precisión de las estimaciones. Un coeficiente de variación elevado indica un bajo grado de precisión, sin que se pueda delimitar el listón a partir del cual la estimación no es válida. En este sentido, el propio INE, que hasta hace poco decía: «toda cifra que vaya afectada de un coeficiente de variación superior al 10 por 100 se debe acoger con las debidas precauciones y no depositar en ella más confianza de la que merece»[14], ahora dice: «toda estimación con un error de muestreo elevado debe ser tomada con reservas; aunque debe ser el usuario el que, de acuerdo con el grado de fiabilidad que precise, determine si un dato con un cierto error de muestreo le es útil o no para la toma de decisiones»[15]. Parece, por tanto, que el 10 por 100 se puede utilizar como punto de referencia, pero no como una barrera que establez-

[14] INE, *Población activa,* enero-marzo 1984, p. XXII.
[15] INE, *Encuesta de población activa,* abril-junio 1989, p. XIX.

ca qué estimaciones son precisas y cuáles no. Es la finalidad de la investigación, en definitiva, la que va a permitir fijar qué errores son permisibles y cuáles no para cada caso en concreto.

Dadas las ventajas que puede ofrecer el coeficiente de variación, sobre todo cuando se trata de estimar poblaciones marginales[16], con una determinada precisión, se incluye, a continuación, una fórmula alternativa a la expuesta con anterioridad. Se obtiene mediante una transformación de la fórmula del cálculo de los errores de muestreo:

$$e_{\hat{p}} = K \sqrt{\frac{N-n}{N-1} \cdot \frac{p(1-p)}{n}}$$

Si se dividen los dos términos por la proporción, *P*, el resultado es éste:

$$\frac{e_{\hat{p}}}{p} = \frac{K \sqrt{\frac{N-n}{N-1} \cdot \frac{p(1-p)}{n}}}{p}$$

y prescindiendo de la corrección para poblaciones finitas, dicha fórmula se transforma en:

$$CV = K \sqrt{\frac{(1-p)}{np}}$$

Despejando *n* se llega a la fórmula [3], que es la que aquí interesa, ya que permite el cálculo directo del tamaño de la muestra fijada la proporción[17] y el nivel de precisión[18]:

$$n = \frac{K^2(1-p)}{CV^2 p} \qquad [3]$$

Aplicando esta fórmula al cálculo de tamaño de la muestra, para estimar, con niveles de precisión del 0,05 y 0,10, el porcentaje de funcionarios

[16] Aquí se llaman poblaciones marginales a las que representan porcentajes muy bajos de los universos en estudio.

[17] La proporción es el parámetro a estimar y, por tanto, se desconoce. Por estudios pilotos o fuentes secundarias se pueden establecer aproximaciones al valor que se trata de estimar.

[18] El nivel de precisión lo fija el investigador en función de requerimientos técnicos y, muchas veces, también económicos, que imponen limitaciones al desarrollo de una muestra grande.

que han entrado en la administración por oposición —el 61 por 100 según datos conocidos— y por contrato y oposición restringida —el 18 por 100—, se obtienen los siguientes resultados:

Ingreso por oposición

CV=0,05 y k=2

$$n = \frac{4 \cdot 0{,}39}{0{,}05^2 \cdot 0{,}61} = 1.040 \text{ entrevistas}$$

y con un coeficiente de variación del 0,10 y k=2:

$$n = \frac{4 \cdot 0{,}39}{0{,}10^2 \cdot 0{,}61} = 256 \text{ entrevistas}$$

Ingreso por contrato y oposición restringida

CV=0,05 y k=2

$$n = \frac{4 \cdot 0{,}82}{0{,}05^2 \cdot 0{,}18} = 7.289 \text{ entrevistas}$$

Para un coeficiente del 0,10 y k=2:

$$n = \frac{4 \cdot 0{,}82}{0{,}10^2 \cdot 0{,}18} = 1.822 \text{ entrevistas}$$

El número de entrevistas necesarias para estimar distintas proporciones, con un nivel de precisión dado, varía considerablemente según el valor de la proporción que se trate de estimar. También varía en función de la precisión que se desee obtener.

4. LA HOMOGENEIDAD O HETEROGENEIDAD DEL UNIVERSO

Otro factor que puede afectar de forma relevante al tamaño de la muestra es la varianza poblacional, representada en nuestra fórmula [2] por P(1–P)[19]. El término se refiere a la distribución del universo, a la homogeneidad o heterogeneidad del mismo, en relación a una variable concreta. Así, si existen dos poblaciones compuestas por españoles y franceses y, en la primera, hay 80 españoles y 80 franceses y, en la segunda, 40 y 120, respectivamente, las proporciones son, en el primer caso, de 0,50-0,50, mientras que en el segundo son de 0,25-0,75. Consiguientemente, la probabilidad de extraer aleatoriamente de un listado, con los nombres de todo el universo, a un francés o a un español variará de acuerdo con las proporciones anteriormente señaladas. En el primer caso, la varianza será 0,50×0,50=0,250, mientras que en el segundo será 0,25×0,75=0,1875. Por tanto, se da mayor varianza en la primera población porque el universo es más heterogéneo. En consecuencia, el número de entrevistas necesario para conseguir un determinado nivel de precisión en la estimación habrá de ser mayor, tal como se deduce de la fórmula [2], cuya aplicación en este caso, para K=2 y un error del 0,02, da estos resultados:

— Proporciones 0,50-0,50:

$$n = \frac{2^2 \cdot 0{,}50 \cdot 0{,}50}{0{,}02^2} = 2.500 \text{ entrevistas}$$

— Proporciones 0,25-0,75:

$$n = \frac{2^2 \cdot 0{,}25 \cdot 0{,}75}{0{,}02^2} = 1.875 \text{ entrevistas}$$

La varianza máxima se da cuando los valores de la proporción son de 0,50-0,50, y a esta varianza corresponde el mayor número de entrevistas. A medida que los valores de la proporción se alejen de 0,50, el valor de la varianza será menor dado que la población se hace más uniforme. Por eso, cuando la proporción tienda a 0 o a 1 será necesario un número menor de entrevistas para alcanzar una determinada precisión.

Lo anterior pone de relieve la importancia de conocer la homogeneidad o heterogeneidad del universo para hacer el diseño muestral y dimensionarlo adecuadamente. En la realidad, se puede hacer una aproximación a la varianza del universo, en primer lugar, a través de datos secundarios. Esto ocurre siempre que se repiten, periódicamente, un determinado tipo de estudios que van facilitando el conocimiento del universo de que se trate. Es más, distintas estadísticas y bancos de datos pueden facilitar una aproximación que sea suficiente para mejorar el diseño muestral.

[19] P(1–P) es la varianza poblacional en el caso de que se trate de proporciones.

Si lo anterior no existiera, si el universo fuera desconocido, en los aspectos que interesa investigar, cabe el recurso a la encuesta piloto. En estos casos bastan pocas entrevistas, 60 u 80, estratégicamente distribuidas, para conseguir los objetivos de aproximación al conocimiento de la homogeneidad o heterogeneidad del universo.

Hasta aquí se ha hablado del cálculo de la varianza de una variable para diseñar la encuesta. Sin embargo, como dice Pulido San Román: «generalmente, una investigación mediante encuesta no se referirá a una única cuestión, sino que, como posteriormente veremos, recogerá datos sobre los múltiples aspectos que puede incluir un cuestionario. De esta forma, los resultados de cada pregunta exigirán un tamaño de muestra diferente para una misma fiabilidad. Ante la complejidad del problema, en la práctica se opta, generalmente, por determinar un tamaño de muestra tal que garantice una fiabilidad dada»[20]. Por eso, y en estas situaciones, se parte de la varianza máxima $0{,}50 \times 0{,}50$, lo que exige un mayor tamaño de la muestra pero da más garantías de precisión ante un universo desconocido que se pretende conocer a través de estimaciones mediante muestreo.

5. EL TIPO DE MUESTREO

Hasta aquí se ha hablado del muestreo aleatorio simple, estableciendo, para este tipo de muestreo, la relación entre nivel de precisión y tamaño de la muestra. En la realidad, sin embargo, hay otros muestreos, más utilizados, a los que es necesario hacer alusión.

El muestreo estratificado de elementos[21] se basa en la estratificación del universo en un número reducido de categorías de tal forma que cada una de ellas reúna a las unidades que, al menos en determinados aspectos, sean similares entre sí y distintas del resto[22]. La utilización de este sistema de muestreo suele reducir el error muestral para un tamaño dado de la muestra o, expresado en función de la precisión, es necesario un número menor de entrevistas para un error de muestreo dado. Esto sucede cuando la suma ponderada de las varianzas de los diferentes estratos es inferior a la varianza del universo en su conjunto.

Volviendo sobre el ejemplo de los funcionarios, se puede calcular el tamaño de la muestra necesario para estimar el porcentaje que ha ingresa-

---

[20] A. PULIDO SAN ROMÁN, *Estadística y técnicas de investigación social,* Madrid, Ediciones Pirámide, 1976, p. 184.

[21] Se hace referencia al muestreo estratificado de elementos para distinguirlo del muestreo estratificado de conglomerados. En el primer caso, el muestreo opera sobre individuos, mientras que en el segundo sobre conglomerados de individuos.

[22] El universo funcionarios se puede estratificar en cinco categorías: grupos A, B, C, D y E.

do en la administración por oposición, a partir de la fórmula [2], para un error de muestreo del 0,02 y una proporción que se supone del 0,61. En este caso, con el muestreo aleatorio simple, el resultado es el siguiente: $n = \dfrac{2^2 \cdot 0{,}61 \cdot 0{,}39}{0{,}02^2} = 2.379$ elementos. Si se utiliza el muestreo estratificado y se tienen en cuenta las varianzas por estrato habría que emplear una nueva fórmula, derivada de la [2]:

$$n_h = \frac{K^2 \cdot \Sigma[W_h \cdot P_h \cdot (1-P_h)]}{e^2} \qquad [4]$$

y que, aplicada al ejemplo de los funcionarios, daría los resultados del cuadro 2.

CUADRO 2

Cálculo del tamaño de la muestra de funcionarios, teniendo en cuenta la varianza de cada estrato

(Proporción de funcionarios que ingresaron por oposición)

Grupos	*Proporción ingreso por oposición*	*Varianza*	W_h	$W_h \times P_h\ (1-P_h)$
A	0,67	0,2211	0,123	0,0271953
B	0,72	0,2016	0,125	0,0252
C	0,68	0,2176	0,201	0,0437376
D	0,56	0,2464	0,478	0,1177792
E	0,42	0,2436	0,073	0,0177828
TOTAL	0,61			0,2316949

Siendo k=2 y e=0,02, $n_h = \dfrac{2 \cdot 0{,}2316949}{0{,}02^2} = 2.317$

NOTA: Se ha calculado para una afijación proporcional.

La varianza en este caso es de 0,2316949, inferior a la que resulta del muestreo aleatorio simple, que es de 0,2379. Por eso, el número de entrevistas necesario es menor, y el efecto de diseño[23] es de $\dfrac{0{,}2316949}{0{,}2379} = 0{,}97391$, lo que significa que hay, en este caso, una reducción de varianza del 0,026,

[23] El efecto de diseño se define como «la razón de la varianza de la estimación obtenida a partir de la muestra más compleja a la varianza de la estimación obtenida a partir de una muestra aleatoria simple del mismo número de unidades». William G. COCHRAN, *Técnicas de muestreo,* México, Editorial Continental, 1984, p. 119.

entre el diseño estratificado y el aleatorio simple. Dicho de otra forma, si se parte de una muestra de 2.500 entrevistas y este número se divide por el «efecto de diseño» $\frac{2.500}{0,97391} = 2.567$, se obtiene el número necesario de entrevistas para obtener, con el muestreo aleatorio simple, el mismo error muestral. En este supuesto habría que ampliar la muestra en 67 entrevistas.

En este caso, la ganancia con la estratificación es muy pequeña; no obstante, puede ser de importancia en los casos en que se trabaje con universos formados por categorías poblacionales homogéneas en su interior, bajo determinados aspectos, y heterogéneas entre sí. Volviendo sobre el ejemplo anterior, si la distribución de frecuencias por estratos fuera la que aparece en el cuadro 3, la varianza total se reduciría hasta 0,218 y, en consecuencia, el número de entrevistas necesario para obtener la precisión 0,02 (error de muestreo) sería de 2.180, es decir, de 135 entrevistas menos que las que se deducen en el cuadro 2. Por su parte, el efecto de diseño sería $\frac{0,2180000}{0,2379} = 0,91635$, con lo que el número de entrevistas necesario para obtener una precisión de 0,02 de error de muestreo sería de 2.728, en el supuesto de operar sobre una muestra de 2.500 entrevistas, obtenida por muestreo estratificado.

CUADRO 3

Cálculo del tamaño de la muestra de funcionarios, teniendo en cuenta la varianza de cada estrato

(Proporción de funcionarios que ingresaron por oposición. Las proporciones se han modificado)

Grupos	*Proporción ingreso por oposición*	*Varianza*	W_h	$W_h \times P_h (1-P_h)$
A	0,70	0,2100	0,123	0,02583
B	0,80	0,1600	0,125	0,02000
C	0,75	0,1875	0,201	0,0376875
D	0,51	0,2499	0,478	0,1194522
E	0,29	0,2059	0,073	0,0150307
TOTAL	0,61			0,2180000

Las explicaciones anteriores responden a casos en los que mediante la estratificación se obtienen varianzas menores. Sin embargo, esto no siempre sucede así (véase cuadro 4), por lo que, en estos casos, la estratificación no supone ganancias en precisión[24]. Tal como se observa en el cuadro 4, la suma ponderada de varianzas es de 0,1480105, superior al 0,1476 que se obtiene tratando al universo sin estratificar (0,18 · 0,82). El efecto de diseño es de $\frac{0,1480105}{0,1476} = 1,003$, mayor que la unidad y, por tanto,con el muestreo aleatorio simple se obtiene la misma precisión con un número algo menor de entrevistas.

CUADRO 4

Cálculo del tamaño de la muestra de funcionarios, teniendo en cuenta la varianza de cada estrato

(Proporción de funcionarios que ingresaron por contrato y oposición restringida)

Grupos	*Proporción ingreso por contrato/restring.*	*Varianza*	W_h	$W_h \times (1-P_h)$
A	0,13	0,1131	0,123	0,0139113
B	0,15	0,1275	0,125	0,0159375
C	0,13	0,1131	0,201	0,0227331
D	0,23	0,1771	0,478	0,0846538
E	0,18	0,1476	0,073	0,0107748
TOTAL	0,18			0,1480105

El muestreo por conglomerados es el más utilizado en muestreos dirigidos a la población en general, puesto que permite actuar con marcos incompletos. Sin embargo, desde el punto de vista de la precisión, es el que ofrece peores resultados, ya que aquí los errores de muestreo suelen ser mayores que en el muestreo aleatorio simple. La razón de ello hay que buscarla en la estructura de este muestreo en que las unidades muestrales están formadas por conglomerados, cuya composición puede ser muy variada (véanse tablas del Apéndice).

Los errores de la estimación en el muestreo por conglomerados son mayores que en el muestreo aleatorio simple y se puede analizar en comparación con los que se derivarían del mismo, que es el que sirve de prototipo, a través de la fórmula [5] $1+(M-1)\rho$[25], que recoge las variables que definen el «efecto del diseño».

[24] Aunque la estratificación no suponga ganancias en precisión, puede ser útil introducirla porque simplifica el proceso de muestreo.

[25] M=tamaño del conglomerado, ρ=coeficiente de correlación que mide la homogeneidad o la heterogeneidad del conglomerado.

Para que el muestreo por conglomerados sea igual de preciso que el muestreo aleatorio simple, $(M-1)\rho$ tiene que ser igual a 0, y esto sólo se consigue si ρ es igual a 0 o M es igual a 1. El primer caso, ρ igual a 0, significa que no hay correlación, que la homogeneidad del conglomerado es 0 y, por tanto, el conglomerado es *heterogéneo.* En consecuencia, la varianza del conglomerado es elevada y la varianza interconglomerados baja. Cada uno de los conglomerados es en sí mismo una muestra del universo, tanto más representativa cuanto mayor sea la varianza interna.

La situación anterior se suele dar rara vez, ya que, habitualmente, ρ es mayor que 0 y el tamaño del conglomerado mayor que 1, lo que supone un efecto de diseño superior a 1. Lo importante en este supuesto es que el tamaño del conglomerado no sea muy grande para que el efecto del diseño no se dispare.

En la realidad, y según los casos, suelen predominar unas u otras situaciones. Si los conglomerados tienden a la homogeneidad interna es conveniente extraer pocos elementos de cada conglomerado y muchos conglomerados, ya que la varianza interconglomerados es elevada. En caso contrario, si los conglomerados son heterogéneos, bastaría con elegir un número reducido de los mismos. En cualquiera de los casos, se puede decir que el diseño por conglomerados es el que menos contribuye a la precisión, aunque, a veces, sea el único que se pueda aplicar.

6. DESAGREGACION DE LOS RESULTADOS

Es frecuente, en muestreo, recurrir a diseños que permitan desagregar los resultados, de forma que las estimaciones se refieran no sólo al universo tomado globalmente, sino también a diferentes subdivisiones o dominios del mismo. Trasladando este planteamiento al ejemplo de funcionarios, se

CUADRO 5

Errores de muestreo por estratos con afijación proporcional

Grupos	*Peso*	*Proporción*	*N.º entrevistas*	*Error muestral*
A	0,122	0,13	308	0,019
B	0,125	0,15	312	0,020
C	0,201	0,13	502	0,015
D	0,478	0,23	1.195	0,012
E	0,073	0,18	183	0,028
TOTAL	1,000	0,18	2.500	0,007

NOTA: La proporción se refiere al porcentaje de funcionarios que han ingresado por contrato y oposición restringida.

puede desear hacer estimaciones referidas solamente al universo funcionarios, o también a cada uno de los cinco grupos en que se ha estratificado. En este último caso, no es posible operar con la muestra diseñada para el universo en su conjunto, ya que los errores muestrales, por dominios, serían elevados y, en consecuencia, las estimaciones poco precisas (véase cuadro 5).

A pesar de que la proporción es muy pequeña en el ejemplo y, por tanto, también las respectivas varianzas, el error de muestreo para cada grupo es relativamente importante, sobre todo si se tiene en cuenta que los cálculos están hechos para un nivel de confianza del 68,26 por 100. Si se opera con el nivel habitual del 95,46 por 100, los errores hay que multiplicarlos por 2 y el nivel de precisión se hace muy pequeño, sobre todo en las submuestras con un número de entrevistas inferior a 500. En consecuencia, habría que partir de un diseño que tuviera en cuenta la necesidad de desagregar los resultados y los niveles de precisión requeridos en cada estimación.

Para resolver esta situación se puede acudir a diferentes técnicas bien diferenciadas. La primera consiste en fijar el nivel de precisión para la categoría menos frecuente, en este caso funcionarios del grupo E; hallar el tamaño de la muestra para conseguir esta precisión y, posteriormente, hacer la afijación proporcional en todos los grupos. En el ejemplo de funcionarios del grupo E, si se parte de una proporción del 0,50-0,50[26] y un error de muestreo del 0,015 para un nivel de confianza del 68,26 por 100, el tamaño de la muestra sería de $n = \frac{P(1-P)}{e^2} = 1.111$ entrevistas. Aplicando la misma tasa de muestreo al conjunto de la muestra, habría que realizar 15.219 entrevistas, cuya distribución sería la siguiente: grupo A, 1.872 entrevistas; grupo B, 1.902 entrevistas; grupo C, 3.059 entrevistas; grupo D, 7.275 entrevistas, y grupo E, 1.111 entrevistas. La solución tomada garantiza los niveles de precisión, pero a costa de un número muy elevado de entrevistas, innecesario para los objetivos de la investigación, y con unos costes desproporcionados.

La segunda técnica, y la más habitual, consiste en recurrir a una afijación no proporcional, tratando a cada dominio como si fuera un universo diferente. Dado que el volumen de población, en universos grandes[27], no modifica el cálculo del tamaño de la muestra, se puede aplicar el mismo tamaño muestral a los diferentes dominios. En nuestro caso se aplicarían 1.111 entrevistas a cada grupo, lo que haría un total de 5.555 entrevistas. La solución es correcta, pero parece necesario hacer algunas anotaciones. En primer lugar, el número de 1.111 entrevistas se puede disminuir siem-

[26] La proporción 0,50-0,50 da la varianza máxima y de ahí que, para que el diseño sirva para cualquier tipo de distribución de la variable, sea necesario partir de este supuesto.

[27] Véase nota 10.

pre que la varianza sea menor que la supuesta[28] y el nivel de precisión requerido también (mayor error de muestreo). En segundo lugar, se puede trabajar con distintos niveles de precisión por dominio, en función de los intereses de la investigación. También puede ocurrir que las varianzas sean distintas, lo que modificaría el cálculo del tamaño muestral. Por último, hay que señalar que en multitud de ocasiones se recurre a muestras cargadas, a partir de diseños generales. Se trata de aplicar la afijación no proporcional sólo a determinados dominios que hay que analizar en profundidad.

En la muestra a funcionarios puede interesar estudiar todo el colectivo y, además, de forma especial, a los funcionarios del grupo A, etc. En este supuesto se puede partir de un diseño general de 2.500 entrevistas, por ejemplo, y aumentar solamente el número de entrevistas del grupo A hasta 1.000. Ello permite estudiar el universo funcionarios y el dominio o la categoría funcionarios de titulación superior, con un alto nivel de precisión.

Queda señalar, finalmente, que siempre que se utilizan diseños muestrales que modifican la afijación proporcional es necesario recurrir a la ponderación si se desea tabular la información de las diferentes submuestras conjuntamente. De lo contrario pueden aparecer resultados anómalos que provienen de agregar información procedente de encuestas en las que se han aplicado distintas tasas de muestreo.

Hasta aquí se ha hablado de los diferentes factores que se han de tener en cuenta, a la hora del diseño muestral, para calcular el tamaño de la muestra que permita alcanzar los adecuados niveles de precisión, como paso previo a la inferencia estadística. De todos ellos, el que ofrece mayores dificultades es la varianza, bien porque su cálculo, *ex ante,* resulta muy difícil, bien porque la muestra hay que dimensionarla para hacer diferentes estimaciones, con varianzas muy dispares[29]. El resto de los factores son conocidos, y de ahí que se puedan y se deban hacer estudios muy completos en la fase de preparación de la muestra.

Realizada la encuesta, se conocen los valores de las estimaciones, y con ellos se puede recomponer la varianza para cada variable. Quiere esto decir que, *ex post,* se conocen todos los factores que permiten acotar los niveles de precisión de cada estimación y que, por tanto, se puede, y se debe, descender del dato genérico, errores de muestreo de la encuesta, al dato concreto, cálculo del error de muestreo de cada estimación. Esto es imprescindible si se quiere valorar la precisión de cada estimación como paso necesario para extrapolar la estimación al universo del cual se ha extraído.

[28] Recuérdese que la varianza no se suele conocer a la hora del diseño, aunque se pueden hacer aproximaciones a la misma.

[29] En estos casos se supone, a efectos de cálculo muestral, varianza máxima.

APENDICE

A lo largo del artículo se han desarrollado, partiendo en cada caso de los mismos datos, tres tipos de muestreo (muestreo aleatorio simple, estratificado y por conglomerados) que afectan al tamaño y precisión, tal y como se ha indicado. Se resumen, a continuación, las características de cada tipo de muestreo y su relación con el nivel de precisión. Los datos comunes a los tres diseños son los siguientes:

Ambito: Nacional.
Universo: Funcionarios civiles adscritos a la Administración Central.
Tamaño: Varía según el diseño.
Afijación: Proporcional.

1. *Muestreo aleatorio simple*

Procedimiento: Se realiza en una sola etapa (monoetápico), con igual probabilidad y selección aleatoria.

Número de entrevistas: Para un error de muestreo de 0,02 y K=1, 625 entrevistas. Siendo el error de muestreo el mismo y K=2, el número de entrevistas es de 2.500. Este diseño sirve de punto de referencia para calcular el «efecto del diseño».

2. *Muestreo estratificado*

Procedimiento: El universo se estratifica en los cinco grupos que distingue la propia Administración: Grupo A (Cuerpo Técnico), Grupo B (Cuerpo de Gestión), Grupo C (Administrativos), Grupo D (Auxiliares) y Grupo E (Subalternos).

Una vez realizada la estratificación se hace el estudio de varianza de cada estrato (cuadro 2) y se procede al cálculo del tamaño de la muestra siguiendo la fórmula [4].

Número de entrevistas: El resultado es, para un error de 0,02 y K=2, de 2.315 entrevistas.

En este caso, en el que se han podido estudiar las varianzas de cada estrato, se consigue un efecto de diseño de 0,926. Por tanto, para la misma precisión son necesarias 185 entrevistas menos.

3. *Muestreo por conglomerados*

Procedimiento: Frente a los diseños anteriores, cuyas unidades muestrales son los individuos, aquí se trata de seleccionar conglomerados formados por conjuntos de individuos (funcionarios). Estos conglomerados son Unidades Administrativas (Direcciones Generales, Subsecretarías, Organismos Autónomos, etc.) integradas en los distintos Departamentos Ministeriales.

El procedimiento, también monoetápico, ha sido el siguiente:

1.° Se han agrupado los Ministerios en cinco estratos, como puede verse en la tabla I.

2.° La Administración se compone de 454 conglomerados, de los que *a priori* se estima que han de formar parte de la muestra 30. Se realiza una afijación proporcional para elegir los conglomerados que por cada Ministerio han de ser seleccionados y posteriormente se eligen éstos de forma aleatoria simple (véase tabla II).

3.° En los conglomerados seleccionados se entrevistan a todos los individuos que los componen.

El cálculo de los errores resulta, en este caso, prácticamente imposible por desconocer para cada Unidad Administrativa las proporciones reales de ingreso en la Administración por oposición y por contrato y oposición restringida (véanse cuadros 2 y 4). Debido al desconocimiento de estos datos, se ha intentado analizar la homogeneidad a partir de la composición de esas Unidades Administrativas. Para ello se ha estudiado la distribución de los grupos en cada conglomerado (véase tabla III) en un intento de descubrir la homogeneidad o heterogeneidad desde este aspecto. Tal como se observa en la tabla III, los conglomerados son muy dispares y, mientras unos aparecen como más heterogéneos en su composición, otros son más homogéneos. Al no cumplirse en todos la condición de heterogeneidad, el estudio por conglomerados ofrece aquí menos precisión que el aleatorio simple (efecto de diseño mayor que 1).

Ante esta situación, si se quiere realizar un diseño por conglomerados habría que hacer, después de la selección aleatoria de los mismos, una segunda etapa en la que se seleccionaran los individuos (funcionarios) por grupos, con afijación proporcional al número de funcionarios que tiene cada uno. En este caso se está volviendo al muestreo estratificado pero haciéndolo bietápico; primero se eligen los conglomerados y, en la segunda fase, los individuos.

TABLA I

Agrupación de Ministerios. Números de unidades administrativas de cada uno y número de conglomerados que se seleccionan

Estrato	*Ministerio*	*Unidades administrativas*	*Conglomerados*
1	Asuntos Exteriores	35	3
	Administraciones Públicas	19	1
	Relaciones con las Cortes	33	2
2	Agricultura	29	2
	Economía	31	2
	Industria	23	1
	Obras Públicas	25	1
	Transportes	33	3
3	Interior	34	2
	Justicia	18	1
	Defensa	39	3
4	Universidad	33	2
	Cultura	19	1
	Educación	30	2
5	Sanidad	19	1
	Trabajo	27	2
	Seguridad Social	7	1
Total conglomerados		454	Seleccionados 30

FUENTE: Subdirección General de Proceso de Datos de la Administración Pública. Elaboración propia.

TABLA II

Unidades administrativas seleccionadas y número de funcionarios en cada una de ellas

Estrato	*Unidad administrativa seleccionada*	*N.º funcionarios*
1	Secretaría Estado CC. Europeas	88
	Subsecretaría Exteriores	105
	Instituto Cooperación Iberoamericana	111
	Dirección General Organización Puestos Trabajo	161
	Dirección General Medios Comunicación Social	77
	Dirección General Servicios	161
2	Agencia Nacional del Tabaco	154
	Secretaría General Pesca Marítima	126
	Loterías y Apuestas	159
	Dirección General de Seguros	160
	Dirección General Servicios Industria	137
	Confederación Hidrográfica Duero	118
	Dirección General Infraestructura Transporte	189
	Secretaría General Técnica Transportes	137
	Dirección General Telecomunicaciones	136
3	Delegación Gobierno Madrid	161
	Dirección General Política Interior	98
	Dirección General Servicio Jurídico Estado	143
	Dirección General Personal (Defensa)	118
	Dirección General Armamento y Material	84
	Secretaría Estado de la Defensa	53
4	Universidad Autónoma País Vasco	243
	Universidad Valladolid	248
	Dirección General Libro y Bibliotecas	263
	Secretaría General Técnica Educación	212
	Dirección General Promoción Educativa	163
5	Dirección General Salud Alimentaria	151
	Dirección General Personal (Trabajo)	159
	Dirección General Acción Social	132
	Instituto Social Marina	121
	Total funcionarios a entrevistar	4.368

FUENTE: Subdirección General de Proceso de Datos de la Administración Pública. Elaboración Propia.

TABLA III

Distribución de funcionarios por grupos en los quince primeros conglomerados seleccionados

	GRUPOS				
	A	*B*	*C*	*D*	*E*
Conglomerados en el estrato 1					
1.1	25	7	6	50	—
	28,41	7,97	6,82	56,82	—
	–2	–1	–14	–4	–9
1.2	43	3	10	48	1
	40,95	2,86	9,52	45,71	0,95
	+16	–5	–10	–6	–8
1.3	29	2	19	47	14
	26,13	1,8	17,12	42,34	12,61
	+2	–6	–1	–7	+5
1.4	34	19	34	73	1
	21,12	11,8	21,12	45,34	0,62
	+7	+11	+14	+19	–8
1.5	8	10	23	31	5
	10,39	12,99	29,87	40,36	6,49
	–19	+2	+3	–23	–4
1.6	22	5	26	76	32
	13,66	3,11	16,15	47,20	19,88
	–5	–3	+6	+22	+23
TOTAL	161	46	118	325	53
%	23	7	17	46	7
$\bar{x}$	27	8	20	54	9

TABLA III (Continuación)

Distribución de funcionarios por grupos en los quince primeros conglomerados seleccionados

	GRUPOS				
	A	*B*	*C*	*D*	*E*
Conglomerados en el estrato 2					
2.1	19	57	43	26	9
	12,33	37,01	27,92	16,88	5,84
	–15	+25	+13	–17	+2
2.2	20	9	41	44	12
	15,87	7,14	32,53	34,92	9,52
	–14	–23	+11	+1	+5
2.3	5	19	42	80	13
	3,14	11,94	26,41	50,31	8,17
	–29	–13	+12	+37	+6
2.4	42	32	20	64	2
	26,25	20	12,5	40	1,25
	+8	—	–10	+21	–5
2.5	24	10	36	43	24
	17,51	7,29	26,27	31,38	17,51
	–10	–22	+6	—	–17
2.6	29	43	16	29	1
	24,57	36,44	13,55	24,57	0,58
	–5	+11	–14	–14	–6
2.7	72	63	18	35	1
	38	33,33	9,52	18,51	0,5
	+38	+31	–12	+8	–6
2.8	44	15	31	45	2
	32,11	10,94	22,62	32,84	1,45
	+10	–17	+1	+2	–5
2.9	48	42	20	24	2
	35,29	30,88	14,7	17,64	1,47
	+14	+10	–10	+19	–5
TOTAL	303	290	267	390	66
%	23	22	20	30	5
$\bar{x}$	34	32	30	43	7

NOTAS: 1.ª fila, número de funcionarios.
2.ª fila, porcentaje horizontal.
3.ª fila, diferencia entre el número de funcionarios del conglomerado y la media del estrato.

FUENTE: Subdirección General de Proceso de Datos de la Administración Pública. Elaboración propia.

TEXTOS CLASICOS