Psicothema 2008. Vol. 20, nº 2, pp. 311-316
www.psicothema.com
ISSN 0214 - 9915 CODEN PSOTEG

Determinación de la máxima varianza para el cálculo del factor de imprecisión sobre la escala de medida, y extensión a diferentes tipos de muestreo

José Antonio Martínez García y Laura Martínez Caro
Universidad Politécnica de Cartagena

La precisión de las estimaciones tiene que ser adecuadamente descrita en la investigación mediante encuesta, donde las escalas de medida ordinales y de intervalo son comúnmente utilizadas. En relación a la estimación de valores medios poblacionales, los errores absoluto y relativo están en función de esas escalas de medida. Este trabajo discute algunas de las asunciones en las que se fundamenta el «Factor de Imprecisión sobre la Escala de Medida —FIEM—». Este índice es una herramienta para evaluar el grado de imprecisión de las estimaciones, independientemente del rango de la escala de medida considerado. Específicamente, proponemos un nuevo método para determinar la varianza más desfavorable, el cual es consistente con la asunción de normalidad en la población, a diferencia del método original basado en una distribución bimodal. Este método reduce el valor de la varianza más desfavorable, y es fácilmente calculado a partir de la función de distribución normal estándar. Además, se muestra la relación de FIEM con otros tipos de muestreo probabilístico, como el muestreo estratificado y por conglomerados.

Determining the most unfavourable variance to calculate the Measurement Scale Imprecision Factor, and extension to other types of sampling methods. The precision of estimates must be adequately reported in survey research, where ordinal and interval measurement scales are commonly used. Regarding mean estimate, absolute and relative errors exist as a function of the measurement scales. This manuscript discusses some assumptions underlying the development of the Measurement Scale Imprecision Factor —MSIF—, a tool to assess the degree of imprecision of estimates, regardless of the scale rank considered. Specifically, we propose a new method for determining the most unfavourable variance, which is consistent with the normal distribution assumption, unlike the original assumption based on the bimodal distribution. This method reduces the value of the most unfavourable variance, which is easily computed using the cumulative normal standard distribution function. In addition, we show the relationship between MSIF and other types of probabilistic sampling methods, such as stratified and cluster sampling.

En un reciente artículo, Martínez y Martínez (2006) presentaron el «Factor de Imprecisión sobre la Escala de Medida» (*FIEM*) como herramienta para evaluar el grado de imprecisión de las estimaciones de valores medios en un muestreo aleatorio simple, independientemente del rango de escala de medida considerado. El desarrollo de este índice está justificado por la identificación de algunos errores comúnmente cometidos en la investigación mediante encuesta, relativos a la información suministrada sobre la precisión de las estimaciones. El índice propuesto, *FIEM*, puede ser utilizado para comparar los errores absolutos y relativos cuando se utilizan diferentes unidades de medida.

Fecha recepción: 9-4-07 • Fecha aceptación: 25-9-07
Correspondencia: José Antonio Martínez García
Facultad de Ciencias de la Empresa
Universidad Politécnica de Cartagena
30203 Cartagena (Spain)
E-mail: josean.martinez@upct.es

Parte de la génesis de este índice está basada en la estimación de la máxima varianza de la escala utilizada en la medición de la variable en cuestión. Sin embargo, esa estimación de la varianza más desfavorable está fundamentada en la asunción de que la variable sigue una distribución bimodal en «U» en la población. El objetivo de este trabajo es proponer una nueva forma de determinación de la varianza más desfavorable, concordante con la asunción de normalidad en la distribución poblacional. El cálculo de esa varianza más desfavorable permite relacionar *FIEM* con otros tipos de muestreo probabilístico, como el muestreo estratificado y por conglomerados.

La necesidad de *FIEM*

En muchas ocasiones, cuando no existe ningún estudio anterior que proporcione información sobre la variabilidad de la distribución de los datos poblacionales, se suele determinar el tamaño de muestra necesario antes de recoger los datos utilizando un criterio de error máximo admisible y situación más desfavorable en cuanto a dispersión de la distribución. De esta forma, el investigador se asegura de que con esa muestra, como mucho, va a cometer ese error fijado, aunque en realidad, el error cometido suele estar por debajo de ese límite, ya que la variabilidad muestral normalmente tiene un valor menor que el valor prefijado de dispersión poblacional.

Las fórmulas para manejar estos aspectos metodológicos son ampliamente conocidas y pueden consultarse en cualquier manual básico (por ejemplo, Levy y Lemeshow, 1999). De esta forma, si se quiere estimar la media de una característica de la población que se asume que sigue una distribución normal, como, por ejemplo, el nivel de motivación o la satisfacción de un consumidor, se puede calcular el tamaño de muestra óptimo en función de un error máximo admisible, ya sea absoluto (1) o relativo (2).

$$n \geq \frac{z^2 N \sigma_y^2}{z^2 \sigma_y^2 + (N-1)\varepsilon_{A(\bar{y})}^2} \tag{1}$$

$$n \geq \frac{z^2 N V_y^2}{z^2 V_y^2 + (N-1)\varepsilon_{R(\bar{y})}^2} \tag{2}$$

donde n= tamaño de la muestra; N= tamaño de la población; z= coeficiente de fiabilidad para un nivel de confianza α correspondiente al valor crítico de una distribución normal estándar; $\varepsilon_{A(\bar{y})}$ es el error absoluto; $\varepsilon_{R(\bar{y})}$ es el error relativo; σ_y^2 es la varianza poblacional y V_y^2 es el coeficiente de variación.

Conviene recordar, asimismo, que si el investigador pretende conocer una característica poblacional que es un valor de probabilidad binomial p, como la proporción de personas que desarrollan una enfermedad o la proporción de consumidores que comprarán un nuevo producto, puede valerse de expresiones similares para obtener el tamaño de muestra óptimo (3) y (4).

$$n \geq \frac{z^2 N\, pq}{z^2 pq + (N-1)\varepsilon_{A(p)}^2} \tag{3}$$

$$n \geq \frac{z^2 N \frac{q}{p}}{z^2 \frac{q}{p} + (N-1)\varepsilon_{R(p)}^2} \tag{4}$$

siendo q la proporción de individuos que no poseen la característica de interés (por lo que la varianza de p es pq).

Cuando no es posible recoger el número de cuestionarios planeado, o simplemente se realiza un número de encuestas al amparo de algún criterio de coste, es práctica común recalcular el error muestral con las expresiones (1), (2), (3) o (4), según los objetivos de la investigación. Es decir, se despeja n o ε en función del caso pertinente.

Es bastante común, asimismo, que las mediciones de conceptos abstractos en ciencias sociales (como la motivación o la satisfacción comentadas anteriormente) se instrumentalicen a través de escalas de medida ordinales (por ejemplo, Likert o diferencial semántico) y que se analicen éstas como escalas de naturaleza «métrica», es decir, como variables cuantitativas con carácter no discreto. De esta forma, son estudiados los valores medios de esas variables, con su correspondiente intervalo de confianza, pudiendo compararse la imprecisión de las estimaciones entre investigaciones que utilicen distintas escalas de medida ordinales.

Sin embargo, tal y como explican Martínez y Martínez (2006), un gran número de investigaciones suelen cometer dos equivocaciones fundamentales relacionadas con las consideraciones anteriores. En primer lugar, se produce la confusión de tomar como caso más desfavorable de la varianza la situación correspondiente al cálculo de tamaño muestral para el caso de proporciones (ecuación 3); en segundo lugar, se fija un error relativo máximo admisible sin tener en consideración el rango de la escala de medida utilizada. La primera equivocación hace que la varianza máxima siempre sea el producto del valor pq, es decir, 0.25 en el caso más desfavorable, cuando la varianza más desfavorable dependerá del rango de escala utilizado, y, por tanto, tendrá un valor distinto dependiendo de la longitud de la escala ordinal. La segunda confusión se traduce en que una investigación con el mismo error relativo máximo que otra pueda producir estimaciones más imprecisas si el rango de escala es menor. Es decir, no es lo mismo equivocarse en un 5% de la media en una escala sobre rango de 1 a 5, que en el mismo valor de una escala de 1 a 10. En el primer caso, la estimación es más imprecisa que en el segundo.

Para mitigar estas confusiones Martínez y Martínez (2006) proponen el factor de imprecisión sobre la escala de medida (*FIEM*), que se obtiene de la ecuación (5):

$$FIEM = 100\left(\frac{\varepsilon_A}{R_E}\right) \tag{5}$$

donde R_E es el rango de la escala de medida (diferencia entre los valores máximo y mínimo) y ε_A es el error absoluto de estimación. Así, si se quiere que la estimación de dos medias de una variable medida en escalas de diferente rango sea igual de imprecisa, basta con fijar *FIEM* al valor requerido de imprecisión y calcular ε_A, el cual será utilizado para determinar el tamaño de muestra óptimo.

El valor de *FIEM* es invariante frente a las escalas de medida, por lo que es una representación del grado de imprecisión de la estimación que permite comparar entre estudios que utilicen diferentes rangos de escalas.

FIEM debe utilizarse también para completar la información que proveen los intervalos de confianza una vez recogido los datos, ya que la amplitud del intervalo de confianza viene determinada por dos veces el error absoluto de estimación ($2\varepsilon_A$). De nuevo así, y ya recogidos los datos, se puede obtener un índice de imprecisión (*FIEM*) que permite comparar entre estimaciones realizadas con escalas de distinto rango.

Martínez y Martínez (2006), asimismo, proponen que se determine la máxima varianza bajo la asunción de que los valores de la población están igualmente repartidos en los dos extremos del intervalo de medición. La distribución sería entonces lo más dispersa posible, por lo que la varianza máxima sería $(R_E / 2)^2$, dependiente, por tanto, del rango de la escala. Si, tal y como hemos comentado, se recoge un número de encuestas determinado y luego se recalcula el error muestral máximo cometido, la máxima varianza servirá para realizar ese cálculo, del cual se obtendría posteriormente el valor de *FIEM*.

Determinación de la varianza en el caso más desfavorable

El desarrollo de FIEM se basa en la asunción de normalidad en la distribución poblacional, ya que las fórmulas anteriormente descritas están fundamentadas en esa asunción. Cuando la población no es normal, existen otros métodos para hallar los intervalos de confianza de los parámetros estimados, como la aplicación de la desigualdad de Chebychev o métodos no paramétricos, como los basados en el remuestreo, por ejemplo, el *boostrapping.* No obstante, por las propiedades del Teorema Central del Límite, se pueden obtener intervalos de confianza aproximados cuando la muestra es grande (>30), por lo que los errores muestrales (absoluto y relativo) o el tamaño de la muestra óptimo calculados en base a las fórmulas descritas serían una buena aproximación a medida que la muestra crece.

Sin embargo, si la asunción de normalidad en la población se mantiene, rápidamente se deduce que la determinación de la máxima varianza poblacional en el caso más desfavorable, es decir, en el que su valor es máximo, no es correcta. Si las dos mitades de los valores poblacionales estuvieran situados en los dos extremos de la escala, la distribución sería bimodal en forma de «U», y por tanto, alejada de la normalidad (figura 1).

La distribución normal tiene forma simétrica en torno a la media. Si consideramos que los valores poblacionales de la variable en cuestión están en el mismo rango que la escala de medida utilizada, la varianza máxima de todas las distribuciones normales que pudieran encontrarse entre los dos extremos del intervalo de la escala vendría determinada cuando la media poblacional estuviese exactamente en el valor medio del rango de escala considerado. Además, por las propiedades de la distribución normal, el 95% del área bajo la curva debería estar entre la media más/menos 1,96 desviaciones típicas (figura 2).

Por tanto, el único requerimiento necesario para el razonamiento anterior es que el 5% del área restante estuviera concentrada de forma equitativa en los dos extremos de la escala considerada. Bajo este supuesto, que de forma heurística acomodaría los extremos de la curva asintótica a los dos valores fijos extremos, la determinación de la máxima varianza σ^2 se computa resolviendo la siguiente expresión:

$$\sigma^2 = \left[\frac{1}{F(x)_{1-(\alpha/2)}} \frac{R_E}{2} \right]^2 = \left[\frac{R_E}{2z} \right]^2 \qquad (6)$$

donde $F(x)_{1-(\alpha/2)}$ es el inverso del valor de la función de distribución normal estándar para el nivel de confianza (normalmente 0,05), es decir, el coeficiente de fiabilidad z, y R_E es el rango de la escala de medida. De esta forma, aplicando la expresión 6, se puede obtener el valor de la varianza más desfavorable de una población normal. En la tabla 1 se muestran la nueva varianza calculada, junto con los errores máximos relativo y absoluto, así como el valor de *FIEM* resultante.

La demostración de la idoneidad de los valores de la máxima varianza puede realizarse resolviendo la integral de la función de densidad normal (7):

$$F(x) = \int_{-\infty}^{b} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\ )^2}{2\sigma^2}} dx \qquad (7)$$

siendo b el extremo superior del rango de la escala, la media de la distribución, y σ^2 la varianza más desfavorable. Esta integral puede resolverse de forma aproximada con Excel (DISTR.NORM(b; ; σ; VERDADERO)). Para cada rango de escala, $F(x)$ = 0,975, lo que está en consonancia con las consideraciones realizadas para la derivación de la máxima varianza.

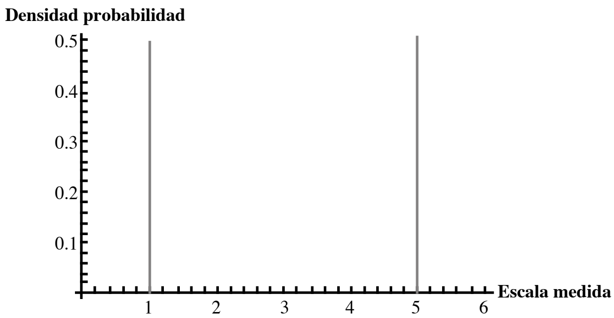

Figura 1. *Distribución bimodal en forma de «U» para valores poblacionales entre 1 y 5 (escala ordinal de cinco opciones de respuesta)*

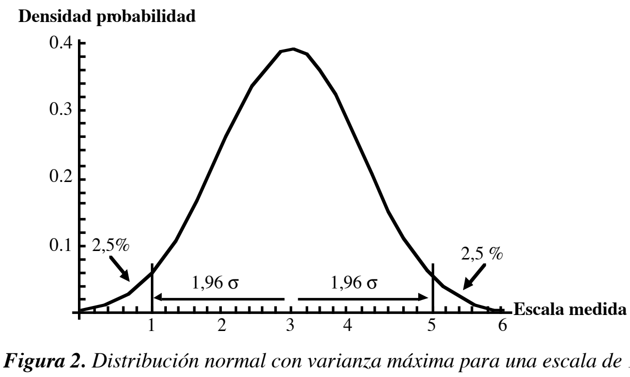

Figura 2. *Distribución normal con varianza máxima para una escala de 1 a 5 (nivel de confianza del 95%)*

Extensión a otros tipos de muestreo probabilístico

La preocupación por el cálculo del tamaño muestral y por informar de la imprecisión de las estimaciones es obviamente extensible a otros tipos de muestreo probabilístico. Vamos a discutir el papel de *FIEM* en dos de los muestreos más utilizados, el muestreo estratificado con afijación proporcional y el muestreo por conglomerados en una fase.

Muestreo estratificado

– *Obtención de una muestra aleatoria simple de cada estrato*: éste es el caso particular donde se quieren obtener unas es-

Tabla 1

Estadísticos e índices derivados de poblaciones normales para una muestra de 137 individuos sobre una población de 352 sujetos

	Varianza máxima σ^2	Rango R_E	Error absoluto ε_A	Error relativo ε_R	*FIEM*
Escala de 1 a 5	1.04	4	0.133	4.45%	3.33%
Escala de 1 a 7	2.34	6	0.200	5.00%	3.33%
Escala de 1 a 9	4.16	8	0.267	5.33%	3.33%
Escala de 1 a 10	5.27	9	0.300	5.46%	3.33%
Escala de 0 a 10	6.50	10	0.334	6.67%	3.33%

Fuente: corrección sobre los resultados de Martínez y Martínez (2006) timaciones con un error máximo prefijado a priori por cada estrato (no tiene porqué ser el mismo en cada uno de ellos). El tamaño de muestra resultante es la suma de las muestras de cada estrato. Al utilizarse un muestreo aleatorio simple en cada grupo, las circunstancias son idénticas a las discutidas en los párrafos anteriores. A la hora de recalcular el error cometido para la media de toda la muestra, tendríamos que tener en cuenta la siguiente expresión (8) en el caso de afijación proporcional (Levy y Lemeshow, 1999):

$$n \geq \frac{z^2 \frac{N}{1+\gamma} V_y^2}{z^2 \frac{V_y^2}{1+\gamma} + N\varepsilon_R^2} \tag{8}$$

Donde γ es el ratio entre las componentes de la varianza entre grupos (σ^2_{by}) y la varianza intragrupos (σ^2_{wy}), siendo la varianza total $\sigma^2_y = \sigma^2_{by} + \sigma^2_{wy}$. Dado que en cada estrato el caso más desfavorable de la varianza corresponde con una distribución normal centrada en el punto medio de la escala (véase figura 2), no habría diferencias entre estratos, por lo que σ^2_{by} sería cero, y por lo tanto $\gamma = 0$, con lo que la expresión (8) se reduciría de nuevo a la expresión (2). Como vemos, la asunción de varianza máxima en cada estrato se correspondería con el mismo caso discutido sobre muestreo aleatorio simple.

– *Obtención del tamaño total de muestra y posterior afijación*: en este caso, de nuevo $\sigma^2_y = \sigma^2_{by} + \sigma^2_{wy}$. Sin embargo, existirían múltiples combinaciones de las dos componentes de la varianza para determinar σ^2_y. Si no se disponen de datos anteriores, no se puede conocer el valor de γ, y, por tanto, la opción más aconsejable vuelve a ser la utilización de la expresión (2). Después, se determinaría el tamaño de muestra en cada estrato como una simple proporción del tamaño poblacional de éstos (véase Levy y Lemeshow, 1999). Sin embargo, conviene resaltar que la posterior afijación proporcional no daría como resultado que las estimaciones por estrato fueran igual de imprecisas, ya que aunque la función de muestreo (*n/N*) sea la misma en cada estrato, la imprecisión de las estimaciones no depende del resultado de ese cociente, sino, entre otros factores, de los valores absolutos que toman *n* y *N*.

Muestreo por conglomerados en una fase

Ahora el interés se centra en la obtención del número de conglomerados (*m*) necesarios para estimar la media poblacional con un límite máximo de error (9)

$$m = \frac{Mz^2 V_{1y}^2}{z^2 V_{1y}^2 + (M-1)\varepsilon_R^2} \tag{9}$$

Donde *M* es el número de conglomerados en la población y V^2_{1y} es el ratio entre la varianza de la distribución de los conglomerados totales (σ^2_{1y}) y la media poblacional. El valor de σ^2_{1y} será mayor cuanto mayor sea la homogeneidad dentro de cada conglomerado y exista más heterogeneidad entre conglomerados, quedando su valor limitado por la varianza máxima poblacional σ^2_y calculada tal y como propone esta investigación.

Por último, si por las circunstancias de la investigación se hubiera elegido uno o varios conglomerados de forma aleatoria para estimar la media total de la población y se quisiera conocer el grado de imprecisión de ésta, habría que obtener el error absoluto de la siguiente expresión (10), el cual directamente determinaría el valor de *FIEM* como hemos visto anteriormente:

$$\varepsilon_A = z\left(\frac{1}{\sqrt{m}}\right)\hat{\sigma}_{1y}\sqrt{\frac{M-m}{M-1}} \tag{10}$$

siendo $\hat{\sigma}_{1y}$ la varianza de la distribución de la característica Y sobre todos los conglomerados, la cual se obtiene a partir de (11):

$$\hat{\sigma}_{1y} = \left[\frac{\sum_{i=1}^{m}(y_i - \bar{y})^2}{m-1}\right]^{1/2} \sqrt{\frac{M-1}{M}} \tag{11}$$

Ilustración práctica

A continuación ilustramos con un simple ejemplo práctico algunas de las principales consideraciones debatidas en este artículo, en concreto las referidas al muestreo aleatorio simple y estratificado. En 2002 se realizó un estudio sobre los usuarios de piscinas de verano del Ayuntamiento de Cartagena, con el objetivo de evaluar su actitud hacia el servicio, concretamente su nivel de satisfacción. La población total (*N*) ascendía a 1.893 usuarios repartidos en tres piscinas de la siguiente forma: 634 en la piscina 1 (N_1), 809 en la piscina 2 (N_2), y 450 en la piscina 3 (N_3). La satisfacción se midió en una escala Likert de 1 a 5. Se seleccionó de forma aleatoria una muestra de individuos sin atender a ningún criterio prefijado, simplemente que el ratio de respuesta de cada piscina fuera cercano al 10% para asegurar cierta representatividad. De este modo se recogieron 58, 72 y 41 cuestionarios (n_1, n_2 y n_3, respectivamente), es decir, una muestra total (*n*) de 171 sujetos. Asimismo, no existían datos anteriores similares que pudieran aportar información acerca de la variabilidad de la población.

Caso 1

La situación que se plantea ahora es la de recalcular el error muestral una vez recogidos los datos, que es lo que normalmente aparece en la ficha técnica de numerosas investigaciones. Aplicando (6) la máxima varianza sería 1.04, y el error absoluto máximo resultaría en 0.146 tras aplicar (1). Por tanto, a través de (5) se obtiene un valor de *FIEM* de 3.648%.

Sin embargo, es importante informar sobre la imprecisión una vez analizados los datos, ya que ésta será posiblemente menor que la que debe indicarse en la ficha técnica. En este caso el nivel de satisfacción medio ($\bar{y}$) es 4.708, con varianza de 0.339, y con un intervalo de confianza al 95% de (4.623 ; 4.793). De esta forma el error absoluto de estimación es de 0.083, siendo el valor de *FIEM* de 2.081%, que, como preveíamos, es bastante menor.

Por último, si en un principio se hubiera calculado el tamaño muestral partiendo de un valor de *FIEM* determinado, por ejemplo, el de 2.081%, como no existían datos anteriores que pudieran dar una idea de la variabilidad de la distribución, el caso más desfavorable de la varianza hubiera producido un tamaño muestral de 443 sujetos.

Caso 2

El interés se centraría ahora en calcular la imprecisión de las estimaciones en cada estrato, que en este caso correspondería con cada piscina. Para ello bastaría con recalcular *FIEM* a partir del error de estimación de la media de cada estrato, que resultaría en los siguientes valores: 2.015, 3.993 y 3.996%. Como se puede apreciar, las estimaciones de las medias en los estratos 2 y 3 son más imprecisas que la estimación de la media total de la satisfacción de los usuarios. Si se quisiera, por ejemplo, calcular el tamaño de muestra mínimo para que la estimación en cada estrato fuera como mucho un 2.081% de imprecisa (en analogía con el caso anterior), el tamaño de la muestra total sería de 892 sujetos, que es el resultado de sumar los tres tamaños de muestra calculados a partir de (1) tomando N_1, N_2 y N_3 como valores poblacionales, respectivamente. De esta forma, n_1= 302, n_2= 337 y n_3=253. Como es fácilmente comprobable, el tamaño de muestra necesario en cada estrato para conseguir estimaciones igual de imprecisas en el caso más desfavorable de la varianza no es proporcional al tamaño real del estrato.

Una hoja de cálculo en Excel con la implementación de este ejemplo puede descargarse en www.upct.es/~gim/inv_jose.htm, seleccionando la página personal de los autores de este artículo. En ese archivo puede comprobarse cómo cambiarían los errores absoluto y relativo al modificar el rango de la escala, así como la independencia de *FIEM* ante esos cambios.

Conclusión, limitaciones y futuras investigaciones

El Factor de Imprecisión sobre la Escala de Medida (*FIEM*) es una sencilla opción para informar de manera más adecuada sobre la imprecisión de las estimaciones de valores medios en la investigación mediante encuesta. En este breve trabajo hemos mostrado la forma de establecer este índice a través de la determinación de la varianza más desfavorable cuando realmente se plantea normalidad en la población, así como el papel que juega *FIEM* en otros tipos de muestreo, como el muestreo estratificado con afijación proporcional y el muestro por conglomerados en una fase.

El problema de las unidades de medida en ciencias sociales y su significado es un tema ampliamente discutido en la literatura (por ejemplo, Cohen, Cohen, Aiken, y West, 1999). Precisamente Cohen et al. (1999) proponen el índice *POMP* (Percent of Maximum Possible Score - porcentaje de la máxima puntuación posible), como forma de unificar y comparar resultados utilizando distintas escalas de medición. El fundamento de *FIEM* y *POMP* es similar, aunque se refieren a conceptos distintos; imprecisión de la estimación frente a unificación de métrica, siendo la relación entre ambos no proporcional.

La utilización de *FIEM* tiene sentido cuando los datos están medidos dentro un rango, es decir, con escalas ordinales o de intervalo. En muchas ocasiones se asume que esas escalas son arbitrarias y que las unidades de medida no tienen sentido, reflejando sus puntuaciones las variaciones de un constructo latente subyacente (véase Hernández, Espejo y González-Romá, 2006). Esa es la forma habitual de proceder de ciertas metodologías, como los modelos de ecuaciones estructurales, por ejemplo. Para los análisis más básicos, como son los relacionados con la estimación de valores medios, las escalas sí que deben de tenerse en cuenta, porque el valor medio estimado diverge en precisión dependiendo del rango de medición.

El debate sobre el carácter o no discreto de las escalas ordinales está todavía bajo discusión (Ato y López, 1996), existiendo diferencias importantes en cuanto a su aplicación en análisis estadísticos (por ejemplo, Coenders y Saris, 1995; Jöreskog y Sörbom, 2001; Vermunt y Magidson, 2005). En este sentido, *FIEM* trata las escalas ordinales como continuas, y, por tanto, está sujeto a la misma crítica derivada de ese tratamiento no discreto.

FIEM asume que no existe heterogeneidad no observable en los datos. Es decir, no hay mezcla de distribuciones normales. Teniendo en cuenta la importancia actual que tiene el estudio de la heterogeneidad (por ejemplo, Allenby, Arora, y Ginter, 1998; Lubke y Muthén, 2005), y el desarrollo de metodologías que muestran cómo la asunción de homogeneidad es cuestionable (por ejemplo, Vermunt y Magidson, 2005), futuras investigaciones podrían tratar de estudiar la posible determinación de este índice en situaciones de heterogeneidad. No obstante, dado que *FIEM* viene determinado por el error máximo absoluto, y éste es establecido a priori a través de la máxima varianza de la escala, resultaría bastante complejo su análisis en situaciones de heterogeneidad, dado que esa heterogeneidad no observable viene determinada a posteriori, es decir, a través de los datos empíricos.

Finalmente, y completando las reflexiones de Martínez y Martínez (2006), podemos proponer las recomendaciones generales para la adecuada información sobre la imprecisión de las estimaciones sobre medias cuando se utilizan escalas acotadas por un rango:

1. Si se planifica el tamaño de muestra en función de un cierto error máximo, se deberían seguir los siguientes pasos:

   – Si no existe información anterior sobre la variabilidad poblacional, calcular la varianza más desfavorable tal y como propone este artículo.
   – Calcular el error máximo admisible a partir de *FIEM*.
   – Determinar el tamaño muestral.

2. Si, por el contrario, se recoge un número de cuestionarios determinado bajo otros criterios (coste, por ejemplo), o el número de cuestionarios válidos finales no coincide con lo previamente planificado, entonces:

   – Calcular la varianza para el caso más desfavorable.
   – Recalcular *FIEM* a partir del cálculo del error máximo admisible para el tamaño de muestra recogido.

3. Una vez que se han recogido los datos, el investigador debe:

   – Mostrar los intervalos de confianza, y, por ende, el error de estimación de la media. Recalcular *FIEM*, que será probablemente menor que el indicado como máximo para el caso más desfavorable.
   – Utilizar ese valor como criterio en la determinación del tamaño muestral de sucesivas investigaciones realizadas sobre la misma población, independientemente de la escala de medida que se utilice después.

De nuevo incidimos en la conveniencia de indicar el valor de *FIEM* en las fichas técnicas de los estudios como una representación del grado de imprecisión de las estimaciones que permita la comparación entre diferentes investigaciones.

Como concluyen Martínez y Martínez (2006), las decisiones sobre el tamaño de muestra óptimo pueden resultar mucho más complejas si se añaden otros factores al diseño de la investigación (hipótesis planteadas, potencia estadística, tamaños de efecto, etc.) (Ares, 2004; Manzano y Braña, 2003). Aun así, el investigador debe informar claramente acerca del error e imprecisión de sus estimaciones, siendo *FIEM* un valor universal que, además, facilitará el metaanálisis.

Referencias

Allenby, G.M., Arora, N., y Ginter, J.L. (1998). On the heterogeneity of demand. *Journal of Marketing Research, 35*, 384-389.

Ares, V.M. (2004). Visión algebraica unificada para el cálculo del tamaño de la muestra. *Metodología de Encuestas, 6*(1), 53-59.

Ato, M., y López, J.J. (1996). *Análisis estadístico para datos categóricos*. Madrid: Síntesis.

Coenders, G., y Saris, W.E. (1995). Categorization and measurement quality. The choice between pearson and polychoric correlations. En Saris, W.E., y Münnich, Á. (eds.): *The Multitrait-Multimethod Approach to Evaluate Measurement Instruments* (pp. 125-144). Budapest, Eötvös University Press.

Cohen, P., Cohen, J., Aiken, L., y West, S. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research, 34*(3), 315-346.

Hernández, A., Espejo, B., y González- Romá, V. (2006). The functioning of central categories middle level and sometimes in graded response scales: Does the label matter? *Psicothema*, 18(2), 300-306.

Jöreskog, K.G., y Sörbom, D. (2001). *LISREL 8.50*. Chicago: Scientific Software International.

Levy, P.S., y Lemeshow, S. (1999). *Sampling of populations: Methods and applications* (3ª ed). Wiley series in probability and statistics. Survey Methodology Section.

Lubke, G., y Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10*, 21-39.

Manzano, V., y Braña, T. (2003). Análisis de datos y técnicas de muestreo. En Levy, J.P., y Varela, J. (dir.): *Análisis multivariante para las Ciencias Sociales* (pp. 91-143), Madrid: Pearson Educación.

Martínez, J.A., y Martínez, L. (2006). El Factor de Imprecisión sobre la Escala de Medida (*FIEM*) en la estimación de medias en un muestreo aleatorio simple. *Investigación y Márketing, 92*, 66-70.

Vermunt, J.K., y Magidson, J. (2005). *Latent GOLD 4.0 User's Guide*. Belmont, Massachusetts: Statistical Innovations Inc.