

CARTAS AL DIRECTOR

Supuesto de máxima indeterminación: ¿error absoluto o error relativo en el cálculo del tamaño de la muestra?

Sr. Director:

En la carta al director de Suárez Gil P y Alonso JC¹ se realiza una crítica al uso del error absoluto (e_0) en la fórmula clásica del cálculo del tamaño de muestra necesario (n) para estimar una proporción (P) poblacional:

$$n = \frac{Z_{\alpha}^2 pq}{e_0^2}$$

proponiendo en su lugar la utilización del error relativo (e_r). La fórmula propuesta:

$$n = \frac{Z_{\alpha}^2 1-p}{e_r^2}$$

nos parece aceptable como alternativa, aunque creemos que la «falacia» del supuesto de máxima indeterminación denunciada por los autores en realidad no existe y el principal problema reside en la interpretación de cada uno de estos dos tipos de «errores». Precisamente por ello, en contra de lo que se afirma en la carta, y tal como ya defendimos en otra ocasión², parece más consistente utilizar el error absoluto para el cálculo del tamaño muestral.

El siguiente ejemplo permite aclarar estos conceptos. Si un investigador utiliza la fórmula propuesta en la carta para estimar una proporción de 0,6 con un e_r de 0,1, la n calculada será de 256, mientras que con el mismo e_r si la proporción es de 0,4, $n = 576$. Nótese cómo ambas proporciones se encuentran igual de cerca (o lejos) de la proporción 0,5 y son complementarias ($0,4 + 0,6 = 1$): estimar que una proporción de hombres es de 0,4, es idéntico que estimar que la proporción de mujeres es de 0,6, por lo tanto el tamaño de muestra calculado debería ser exactamente igual, teniendo en cuenta además que la varianza para una $P = 0,4$ es la misma que para una $P = 0,6$. Lo que ocurre en realidad es que idénticos errores relativos corresponden a distintos errores absolutos. Mientras que en la proporción de 0,4 el $e_0 = 0,04$, para $P = 0,6$ el $e_0 = 0,06$. En la primera situación se estima con mayor precisión y por consiguiente se requiere mayor tamaño de muestra. Utilizar el error absoluto produce resultados consistentes (« n » es idéntica para una $P = 0,6$ o para una $P = 0,4$).

La fórmula propuesta por los autores y la clásica son exactamente equivalentes (como es de esperar, ya que la primera es una simple transformación de la segunda) siempre que los errores absoluto y relativo se calculen el uno en función del otro. A igualdad de error *absoluto*, tanto con la fórmula clásica como la «transformada» el tamaño muestral máximo corresponde a la situación de «máxima indeterminación» ($P = 1-P = 0,5$). Por otra parte, el tamaño muestral calculado para dos proporciones simétricas (por ejemplo 0,4 y 0,6) es exactamente el mismo, como es de esperar.

Está claro en cambio que, si se fija de antemano el error *relativo*, entonces el error absoluto (la precisión de la estimación), varía con el valor de P . A igualdad de error *relativo* (pongamos 0,1, igual que en el ejemplo propuesto en la carta) una $P = 0,25$ se estima con una precisión (*error absoluto*) de 2,5 unidades

porcentuales, mientras que para una $P = 0,5$ la precisión será de 5 unidades porcentuales. Puesto que el tamaño de la muestra depende de la precisión requerida (elevada al cuadrado), no es nada de extrañar que una estimación el doble de precisa requiera casi cuatro veces más casos, tal como resulta del ejemplo en la carta. Todo esto, lejos de demostrar una «falacia» del supuesto de máxima indeterminación o la «ilegitimidad» del uso del error absoluto, pone de manifiesto justo lo contrario: ya que el objetivo final es estimar P mediante su intervalo de confianza, parece más coherente establecer a priori el «recorrido» de este intervalo, es decir utilizar el «error absoluto» en el cálculo del tamaño muestral necesario.

En conclusión, el problema no es la aplicación de una u otra fórmula, sino el hecho de no diferenciar claramente entre la utilización de una precisión absoluta o relativa: el uso, en este sentido, del término «error» para definir la «precisión» de una estimación no facilita la comprensión³. Por lo tanto, no compartimos en absoluto la afirmación de los autores según la cual «... *no es legítimo fijar el error absoluto en situación de máxima ignorancia sobre P ...*», puesto que e_r es una función de e_o y de P , y cualquiera de estos valores es objeto de una elección arbitraria por parte del investigador, el cual puede desconocer totalmente el valor de P en la población y, sin embargo, saber perfectamente con qué precisión (absoluta) quiere realizar la estimación.

J. Marrugat, J. Vila y M. Pavesi
Institut Municipal d'Investigació Mèdica (IMIN)

Bibliografía

1. Suárez-Gil P, Alonso JC. Sobre el supuesto de máxima indeterminación, el tamaño muestral y otras consideraciones sobre muestreo. Gac Sanit 1999;13:243-6.
2. Réplica a la carta de García C, Almenara J. Determinación del tamaño de muestra en variables cualitativas en las que se desconoce el valor del parámetro. Med Clin (Barc) 1999;112: 797-8.
3. Marrugat J, Vila JS, Pavesi M, Sanz F. Estimación del tamaño de la muestra en la investigación clínica y epidemiológica. Med Clin (Barc) 1998;111:267-76.