EL TAMAÑO DE UNA MUESTRA EN UN ESTUDIO DE PREVALENCIA

Carlos Escalante Angulo
Docente, Investigador, Facultad de Optometria

RESUMÉN

La investigación epidemiológica moderna se realiza por medio de muestras representativas cuyos valores, con determinado nivel de confianza, se generalizan a la correspondiente población. Este artículo expone nociones básicas para calcular el tamaño de una muestra probabilística simple en estudios de prevalencia y como, a partir del tamaño muestral, calcular el parámetro en el caso de una proporción (p).

Por lo general, como se puede observar en los estudios que aparecen publicados en las revistas especializadas de la profesión optométrica, en lo principal el investigador está más interesado en la significación clínica de los resultados que en su posible significado epidemiológico, aunque ambos aspectos pueden ir de la mano. Este trabajo se propone describir las técnicas estadísticas para calcular el tamaño de una muestra probalística en estudios de prevalencia y el proceso de estimación por intervalo para determinar el parámetro poblacional.

Importancia de la Investigación

En cualquier disciplina de la salud, la investigación científica es indispensable como fuente primordial de conocimientos para mejorar la calidad de la vida individual y colectiva, incrementar la racionalidad administrativa para la toma de decisiones en salud pública y enriquecer la docencia universitaria para la formación de los profesionales de esas disciplinas.

La investigación clínica, metodológicamente correcta, puede aportar evidencias para mejorar procedimientos diagnósticos y terapéuticos y aquilatar la eficiencia de la profesión.

Por otra parte, la investigación epidemiológica, al proporcionar información fidedigna sobre la prevalencia, incidencia y factores condicionantes de la salud visual y ocular, sienta las bases teóricas y estadísticas para el diseño de política en salud y la programación de recursos en términos de costo – beneficio. Pero la investigación epidemiológica trasciende este horizonte económico – de suyo importante – y sus resultados son útiles para informar a la sociedad sobre la realidad de ésta en materia de salud visual y ocular y, de ésta manera, estimular una toma de conciencia colectiva sobre los problemas que afectan el bienestar físico, psicológico y social de la comunidad.

Dentro de este marco axiológico referencial, este artículo se propone describir nociones esenciales sobre como determinar el tamaño de una muestra probabilística simple en estudios de prevalencia y, a partir de ese tamaño, como estimar el parámetro poblacional en el caso de una proporción (p).

Muestra; función y tamaño

En la investigación epidemiológica, se toman muestras para determinar, por medio de los valores de éstas, los correspondientes valores de la población del estudio.
Tres factores (entre otros), ninguno de los cuales tienen relación directa con el tamaño de la población, determinan el tamaño de la muestra:

a) El nivel de significación es nivel de confianza que el investigador especifique, el cual generalmente es .95. Esto significa que al estimar una proporción como parámetro poblacional, hay 95% de probabilidad de que esta caiga dentro del intervalo definido por los valores.

$$P \pm 1.96 \sqrt{\frac{P \times q}{n}}$$

Inversamente, significa que esta dispuesto a asumir un error de 0.05, esto es, la probabilidad de que el parámetro (la proporción), caiga fuera del intervalo de confianza. De ser así, la mitad de 0.05 se presentará por arriba del intervalo y la otra mitad por debajo.

b) El mínimo error permisible a juicio del investigador, el cual generalmente es 3% y 2% del tamaño de la muestra.
c) La variación de la condición en la población: a mayor variación, mayor tamaño.

La función de una muestra es servir de medio para estimar el parámetro de la población: por ejemplo, la proporción de defectos refractivos no corregidos en la población del estudio se calcula utilizando la proporción encontrada en la muestra. Con este fin se utiliza la siguiente formula relativa a la estimación de proporciones con un nivel de confianza de 0.95.

$$P \pm 1.96 \sqrt{\underline{P \times q}} \times \sqrt{\underline{N - n}}$$

$$n \qquad\qquad N-1$$

La fórmula incluye el error estándar calculado con el factor de corrección cuando el tamaño muestral es superior al 10% de la población.

Supongamos que a falta de estudios previos, un estudio piloto exploratorio realizado nos muestra que alrededor del 20% puede padecer defectos refractivos no corregidos. En el estudio que proyectamos no deseamos incurrir en un error mayor del 3% en la estimación del parámetro poblacional. En consecuencia, la fórmula anterior quedará, reemplazando, sin el factor de corrección, así:

$$P \pm 1.96 \sqrt{\frac{0.20 \times 0.80}{n}} = 0.03,$$

En esta ecuación la única incógnita es n, el tamaño de la muestra, que, despejada, quedaría así:

$$n = \frac{(1.96)^2 \times 0.20 \times 0.80}{0.0009} = 683$$

Si suponemos, vía de ejemplo, que el tamaño de la población estudiada es 4.000 casos, una muestra de 683 es superior al 10% de la población y, en consecuencia, debe introducirse un factor de corrección para poblaciones finitas que, en este caso, será:

$$n^1 = \frac{n}{1 + \frac{n-1}{N}} = \frac{683}{1 + \frac{683 - 1}{4.000}} = 583$$

Así, 583 es una muestra de menor tamaño que la esperada.

Con esta muestra estaremos en condiciones de estimar la proporción de casos en la población que presentan defectos refractivos no corregidos.

Estimación del parámetro

Se recuerda que un valor muestral (por ejemplo, una proporción) permite estimar la proporción de casos en la población dentro de un intervalo de confianza. Siguiendo con nuestro ejemplo, deseamos estimar la proporción de casos con defectos refractivos no corregidos en una población de 4.000 casos. La fórmula siguiente, ya conocida, permite estimar el intervalo de confianza dentro del cual se ubica el parámetro, con un nivel de confianza de 95%.

$$P \pm 1.96 \sqrt{\frac{p \times q}{n}} \times \sqrt{\frac{N - n}{N - 1}}$$

Obsérvese en la fórmula anterior que el error estándar de la proporción es multiplicado por un factor de corrección si la muestra, como en nuestro ejemplo, es superior al 10% de la población.

Remplazando se tiene:

$$0.20 \pm 1.96 \sqrt{\frac{0.20 \times 0.80}{583}} \times \frac{4000 - 583}{4000 - 1} = 0.20 \pm 0.03 = .23 - .17$$

La proporción de casos de defectos refractivos sin corregir en la población se estima entre 23 y 17 por ciento.

No siempre el optómetra investigador opera con muestras representativas, esto es, con muestras que contienen un número suficiente de casos seleccionados al azar. Si tiene, por ejemplo 1.000 historias clínicas como la población de estudio, puede decidir seleccionar al azar una muestra de 200 casos para estimar la proporción de casos con defectos refractivos sin corregir con un nivel de confianza de 95%. Calcula la proporción de casos en la muestra que, por vía de ejemplo, es el 10% y aplica la fórmula:

$$0.10 \pm 1.96 \sqrt{\frac{.10 \times .90}{200}} = .14 - .06$$

El intervalo de confianza oscila entre 14% y 6% con un error de .05 (nivel de confianza .95) Deduzca el lector por si mismo que el aumentar el tamaño de la muestra disminuye correlativamente el intervalo de confianza y aumenta la precisión de la estimación.

Naturalmente, en la investigación optométrica de prevalencia no basta con estimar parámetros, es necesario considerar también la posible significación clínica de los resultados.

BIBLIOGRAFÍA RECOMENDADA

1. Alvin Lewis, Bioestadística. México: C.E.C.S.A., 1977, Capítulo 5.

2. B. Daw son – Saunders y Robert G. Trapp. Bioestadística Médica. México : Manual Moderno 1997, Capítulo 5 y 9.