

Para tener en cuenta antes de solicitar apoyo estadístico

Álvaro Reyes P.

Oficina de Apoyo a la Investigación Clínica, HCUCh.

Cuando un investigador clínico se acerca a solicitar ayuda estadística, son tres los elementos esenciales que se consideran para el ajuste de una solución: las variables en estudio, el tamaño muestral y el diseño o tipo de estudio.

Con respecto a las variables, la primera pregunta es cómo se operacionaliza o en otras palabras, explicar cómo se miden dichas variables. Esta definición nos permitirá entonces conocer el nivel de medición, es decir, si el fenómeno en estudio se puede clasificar en dimensiones o categorías, o si además es posible establecer un orden o jerarquía entre las dimensiones, o si en el mejor de los casos es factible que la variable pueda ser cuantificada numéricamente. Me refiero a clasificar las variables en: nominales, ordinales o continuas. Consideremos que las herramientas estadísticas son múltiples y el tipo de análisis estadístico a utilizar se relaciona directamente con el tipo de variables estudiadas. En términos generales, las variables continuas permiten análisis algo más sofisticados y, por otra parte, cuesta menos encontrar una solución estadística a variables de los tipos nominales o continuas.

Podemos comprobar con facilidad que cuando trabajamos con un mayor tamaño muestral, éste nos

aporta mayor información y en consecuencia la variabilidad tiende a disminuir; lo que es bastante deseable en estadística. Todas las pruebas o dójimas estadísticas se basan en supuestos o asumen que los datos poseen ciertas características, de tal modo que podemos otorgarle mayor crédito a las conclusiones que se derivan de dichos test en la medida que estemos conscientes de que los datos cumplen determinadas condiciones. Por ello surgen los métodos no paramétricos, los que requieren de menos supuestos a expensas de conclusiones menos ambiciosas. En general, un tamaño muestral reducido o variables de tipo ordinales son razones para decidir por algún análisis no paramétrico.

Del diseño de investigación dependerá la profundidad y alcance de los resultados y conclusiones, por lo que con seguridad diseños de estudio con mayor relación de causalidad entre las variables, requerirán análisis estadísticos más sofisticados.

Lejos la consulta más frecuente que se hace a un estadístico es sobre el tamaño muestral necesario para el estudio y frente a eso permítanme la siguiente reflexión: puesto que para obtener una referencia sobre cuántos pacientes debiésemos ingresar a nuestro estudio, debemos recurrir a una ecuación matemática, distinta de acuerdo al

diseño del estudio o al tipo de variables, y como en cualquier ecuación, debemos ingresar valores para despejar la incógnita que constituye el “n”. Los primeros valores que de manera estándar se utilizan son: el característico p-value o magnitud del error Tipo I del 5% y el 80% de potencia estadística. En segundo lugar y más importante que los anteriores, es encontrar una magnitud que provenga idealmente de un trabajo publicado, lo más similar posible, que nos indique por ejemplo, que el porcentaje de éxito o fracaso en un grupo versus el otro fue de $X\%$, o que la diferencia promedio encontrada entre un tratamiento u otro fue de Y , o que la prevalencia de determinada cualidad fue de $Z\%$. A todo lo anterior, debemos agregar que estas magnitudes poseen fluctuaciones o variabilidad, lo que va a reflejarse en la desviación estándar o error estándar, magnitudes que también serán necesarias para el cálculo.

En la actualidad es cada vez menos probable abordar un tema de investigación sin que exista algún antecedente previo o que ningún investigador lo hubiese abordado con anterioridad. Sin embargo, cuando los investigadores reconocen no encontrar la información solicitada, la solución estaría en hacer un estudio piloto con un tamaño muestra reducido, y con los resultados obtenidos realizar la estimación. En ocasiones ocurre que el tipo de estudio planteado es exploratorio o descriptivo, donde el cálculo del tamaño muestral no está justificado.

Por eso es que ante la frecuente pregunta del tamaño muestral, la respuesta automática de cualquier estadístico será con una nueva pregunta acerca de si el investigador conoce alguna referencia bibliográfica de donde poder obtener los valores para el cálculo.

A pesar de mi corta experiencia como bioestadística, podría expresar la siguiente hipótesis: la calidad en el registro de la información y la construcción

de la base de datos es directamente proporcional a la experiencia y calidad del investigador. No puede ser un hecho casual que aquél que conoce y domina el tema investigado es capaz de depurar la información del mismo modo en que identifica y precisa la pregunta de investigación. Por el contrario, cuando el investigador busca a tientas, su proceder errático se manifiesta en el registro de información redundante, poco precisa y sin justificación. Estadísticamente el trabajo se transforma más en un tedioso problema de manejo de base de datos que de profundidad y fineza en el análisis. Bastaría una simple mirada exploratoria de los datos para darse cuenta de que en gran parte de los registros no existe variabilidad que amerite la pregunta de presencia de diferencias significativas.

Muchos investigadores se sorprenden cuando ven que invertimos más tiempo en estructurar una base de datos adecuada que en el análisis mismo. Lamentablemente aún los programas estadísticos solo interpretan códigos o números y no letras, palabras y frases, y menos aún combinaciones de números y letras.

En el tiempo que me ha correspondido trabajar como estadístico de la OAIC, he podido rescatar estas ideas que pudiesen ser útiles al momento de solicitar asesoría. En conclusión, no está demás insistir en que ahorraremos tiempo y evitaremos problemas en la medida en que hagamos las consultas pertinentes desde el momento en que proyectamos la investigación.

CORRESPONDENCIA

KLGO. Álvaro Reyes Ponce
Oficina de Apoyo a la Investigación Clínica
Hospital Clínico de la Universidad de Chile
Santos Dumont 999, Independencia, Santiago
Fono: 978 8535
E-mail: areyes@redclinicauchile.cl / alvaro.reyes@ucv.cl

