El mito del número 30

Oscar Federico Nave Herrera*
fnave@digi.usac.edu.gt

"Los números redondos son siempre falsos".
Samuel Johnson

Con frecuencia se escucha la aseveración de que una muestra mayor de 30 puede ser considerada grande y suficiente para análisis estadísticos inferenciales; así también, en los textos de estadística se hace una división sobre estimaciones o pruebas de hipótesis para muestras grandes o pequeñas, sin una definición precisa de ello, tomando como único criterio si la varianza poblacional se conoce o no.

En esta ocasión quiero referirme a este tema con el objeto de aclarar por qué este número, si bien tiene fundamentación estadística, no es ni debe ser considerado como elemento decisorio en el momento de aplicar estadística inferencial.

Primero hay que determinar de dónde viene el mito del número 30. Para ello es necesario recurrir al teorema del límite central, que puede reducirse a la siguiente expresión: "Si de una población infinita se toma una muestra lo suficientemente grande (que tienda a infinito), el promedio de la variable de interés en dicha muestra tiende a ser igual al promedio de la población; además, la distribución de probabilidad para los promedios

Apuntes perplejos

Medir, confundir

Alfonso Chacón Rodríguez*
alchacon@itcr.ac.cr

Dijo Descartes: "Aquellos que buscan el camino directo a la verdad, no deben molestarse con ningún objeto del cual no se posea una certeza igual a la de las demostraciones de la aritmética y la geometría". Es este quizás el punto de arranque de lo que siglos más tarde se llamaría positivismo. De repente, solo lo medible alcanzará estatus de verdad. Es una revolución intelectual que profundizarán Bacon, Kant, Comte y muchos otros filósofos posteriores: deberemos separar el problema en su componente más básico y medible, y a partir de ahí construir la única verdad objetiva. (Pero el buen viejo del Sócrates platónico, seamos sinceros, ya había orientado un poco el camino por esa ruta miles de años atrás, cuando dijo en el diálogo Eutifrón que solo podía conocerse aquello que podía definirse de manera clara y precisa). Fue una revolución, de resultados tremendos. El auge de las ciencias naturales y exactas, de la tecnología moderna, se asienta sobre este rotundo pilar. Y sus éxitos pronto harían que otras disciplinas del pensamiento humano ansiaran ese estatus casi inconmovible. ¿Si teníamos una ley de Newton para la mecánica, por qué no una ley similar para el comportamiento social, económico, humano? La adopción acelerada de las técnicas mensurables por parte de las llamadas ciencias sociales (ya hasta ciencias del derecho he visto en el nombre de algunas carreras universitarias), llegaría así al extremo de que la misma filosofía mire hoy hacia su hija menor –la antiguamente llamada filosofía natural– para pedir consejo y guía sobre cómo ejercer su oficio de manera rigurosa. Que un filósofo tan connotado como Daniel Dennet, por ejemplo, prácticamente reniegue de lo que no es medible puede parecer el triunfo máximo de las ciencias naturales como último repositorio de lo que es cognoscible (acabando entonces con todo aquello que Aristóteles llamó metafísica).

Y sin embargo (un gran sin embargo), ¿es tan cierto que solo aquello cuantificable puede llamarse conocimiento? Sócrates afirma que, para usar un término, hay que conocerlo, y eso implica primero saber su significado. Es decir, que la validez de una preposición queda circunscrita a la autoridad lingüística, si se quiere. Los positivistas y las ramas filosóficas paralelas como el empirismo, llevarán esta propuesta más allá, al afirmar que solo lo demostrable por la vía de la ciencia, puede afirmarse que sea conocimiento. Pero veamos: en un tenor igual al de Sócrates, Azorín diría después, mucho después, que escribir con metáforas es hacer trampa. ¡Ay, que la respuesta de Francisco Umbral a Azorín bien habría valido hacérsela entender al bueno de Sócrates a su vez: el lenguaje no es más que metáforas! ¿Es posible hablar de amor solo porque conocemos la definición del término? No tengo por seguro que Dafnis y Cloe portaran consigo un diccionario para acompañar sus escarceos pastoriles. Y en todo caso, Wittgenstein también aclaró un poco el terreno al sentar el límite último de cualquier lenguaje como sustituto o significante (mas no la cosa misma de la que se habla), al postular que sobre aquello de lo que no puede hablarse, más vale callar. Y esto incluye cualquier lenguaje, incluyendo el lenguaje lógico matemático. Es decir: que sí, podemos describir la realidad mediante las matemáticas, pero no necesariamente toda la realidad (Gödel aquí tendría mucho que aportar).

Pero la moda de lo medible, de lo enumerable, no se detiene ante verdades tan evidentes si ya se ha entronizado implacablemente. Hemos de numerar, clasificar, como buenos discípulos aristotélicos y cartesianos. Midamos el desarrollo de un país: con el PIB, el Gini, y si eso no basta creemos un índice de felicidad (queda para otros el dilucidar qué pretendemos dar a entender con esto de país desarrollado o feliz). Y es que últimamente, parece, lo importante es entrar en el *Top Ten*. Incluso en aquello más sagrado que nos predicó el pensamiento positivista: el conocimiento racional y científico. De repente, el prestigio en esos claustros que llamamos universidad –ese anticuado concepto europeo y medieval de repositorio del saber que hoy se pretende replantear en términos de construcción de empleo y generación de riqueza— debe apegarse a las mismas normas que deciden la canción, película o tuit más popular. El problema, que de obvio pareciera se vuelve que puedan ser calculados en todas las posibles muestras del mismo tamaño, tiende a ser normal, con un valor central o promedio exactamente igual al verdadero valor del promedio poblacional y una desviación estándar que depende del tamaño de la muestra y que se denomina error estándar".

Este teorema es considerado el fundamento de la inferencia estadística; sin embargo, queda en forma ambigua el tamaño de la muestra que garantice esos supuestos, ya que no aclara lo que es "suficientemente grande" (la acotación sobre la tendencia a infinito no tiene un sentido práctico); solamente se hace la generalización de que mientras más grande sea la muestra, más cerca se llegarán a estimar los verdaderos valores poblacionales.

En segundo término, y a mi criterio, la verdadera razón de la existencia del mito del número 30 está en los conceptos de distribuciones de probabilidad aplicados para "muestras grandes" y para "muestras pequeñas". Para poder hacer inferencias se necesita conocer la distribución de probabilidad de la distribución muestral de la variable y así generalizar los resultados de la muestra a la población. A partir de esto, cuando se habla de promedios (y solamente en ese caso tendría aplicación el mito), se ha asumido que si el número de muestra es menor de 30, la distribución muestral sigue una distribución t de Student, mientas que si el número de muestra es mayor, sigue una distribución normal estándar (Z); en realidad, el concepto estadístico es más complicado ya que no tiene que ver tanto con la muestra sino con la estimación de la varianza o variación de los datos en la población y aquí necesariamente se debe reconocer el papel que han jugado los textos de estadística, como tercer elemento que ha contribuido al mito del número 30.

Cuando se hace una investigación, se obtiene un cálculo de la variabilidad de la muestra, pero nunca se puede llegar a conocer la variabilidad poblacional; por lo tanto, la estimación o inferencia está sujeta al error estándar y la distribución de probabilidad muestral para la media sigue una distribución de t de Student y no una distribución normal estándar Z, la cual solo sería aplicable si la variación poblacional fuera conocida.

Curiosamente resulta que la distribución t a medida que la muestra va aumentando se va pareciendo cada vez más a la distribución Z, a tal grado que a partir de 30 ambas son sumamente parecidas ("parecido" no es "igual"). Esto es pocas veces reconocido en los textos, como lo indica el estadístico mexicano Ernesto Cervantes López: "Obviamente, en la práctica siempre se desconoce el valor de μ y casi siempre el valor de σ, de manera que esto es solo el respaldo teórico de toda la inferencia estadística" (μ y σ son las representaciones de la media y desviación estándar poblacionales). Otros autores indican que considerar que muestras mayores a 30 tienen una distribución normal es simplemente una regla práctica; por lo tanto, no se puede pretender que esas muestras, definidas a priori, sean representativas.

En los libros de texto de estadística se encuentran las denominadas "tablas", que invariablemente corresponden a distribuciones de probabilidad; estas tablas originalmente fueron de gran ayuda considerando que no se contaba con las facilidades computacionales para lograr tener los algoritmos de cualquier distribución de probabilidad. Pero estas tablas presentan limitaciones, ya que por razones de espacio solamente muestran algunos valores en función de probabilidades prefijadas o números de muestra (grados de libertad), también establecidos a gusto del autor; así, resultó más sencillo decir que si una muestra es superior a 30 la distribución t puede aproximarse a la distribución Z, con el equívoco de llamarla una "muestra grande".

En la enseñanza de la estadística se ha acostumbrado desarrollar los problemas de libro con base en suposiciones: "se tiene una variable supuestamente normal"; "suponga que la varianza poblacional es conocida"; "suponga que se tiene un número de muestra grande". Como consecuencia, es común encontrar investigaciones con análisis de resultados en los que no se han comprobado algunos supuestos que deben ser considerados como requisito. No está demás enfatizar que el investigador no trabaja con ejemplos de libro y que los supuestos son recursos teóricos; los datos de una investigación son elementos que deben reflejar realidades, por lo que deben sustentarse en principios estadísticos consistentes, lo cual únicamente se logra por medio de una muestra bien calculada y un sólido diseño de muestreo.

No hay que pensar que una muestra es buena porque refleja cierto porcentaje de la población, ni mucho menos que es buena simplemente porque supera el mito del número 30.

** Oscar Federico Nave Herrera es químico biólogo y estadístico. Forma parte del Programa de Asesoría Estadística para Investigación de la Dirección General de Investigación de la Universidad de San Carlos de Guatemala.

invisible, es que no parece convincente clasificar una compleja institución enfocada a la docencia, la extensión comunitaria, la investigación y la producción de saber, siguiendo un método que al final no parece más serio que contar el número de "me gusta" que recibe un video en las redes sociales. Máxime si existen una decena o más de índices distintos que pululan perorando sobre cuál universidad, investigador o publicación científica es mejor. Porque si por un lado es cuestionable en el buen sentido matemático –ese del que muchos matemáticos, físicos, químicos, biólogos e ingenieros parecen últimamente andar escasos—, que un solo número escalar permita evaluar un fenómeno multidimensional, lo cierto es que la realidad cognoscible es mucho más que el resultado de un indicador. Y que en todo caso, si fuera posible de veras obtener ese número mágico que resuelva que aquella universidad o revista es mejor que esta otra, como buenos científicos al menos habría que esperar que la solución ofrecida fuese comprobable, algo de lo que, desgraciadamente, carecen todos estos índices. Pero ello, de nuevo, no detiene el uso de las métricas mencionadas y, cegados ante su aparente cientificidad, asignamos recursos y prestigio sin preguntarnos la cuestión básica que cualquiera que se diga científico debería plantearse a la hora de considerar un resultado: ¿es esto verificable? Porque para poner las cosas claras, ¿nos es precisamente la falta de posibilidades de verificación lo que separa a la astrología de la astronomía? Es decir, ¿no son al final estos índices –de ranking, impacto, cómo se llamen– tan útiles como decidir sobre la calidad de mi investigación a punta de horóscopos?

*Profesor de la Escuela de Ingeniería Electrónica del Instituto Tecnológico de Costa Rica. Ingeniero en electrónica. Tiene una maestría en literatura inglesa y un doctorado en ingeniería con orientación electrónica.