



EL OCASO DE LA SIGNIFICANCIA ESTADÍSTICA

En los últimos meses posiblemente se hayan encontrado con comentarios relacionados con la propuesta, por parte de algunas publicaciones científicas, de abandonar la utilización de la significancia estadística¹ como criterio de evaluación de hipótesis (ver Amrhein et al. 2006; Halsey 2019). Las referencias mencionadas resumen las posturas que, desde hace ya unos años, en distintas publicaciones, y con distinta profundidad, plantean la necesidad de cambiar el procedimiento establecido para distinguir entre resultados relevantes e irrelevantes, a partir, generalmente, del umbral dado por una probabilidad del 5%.

Este proceso de inferencia por el cual se decide sobre la validez de una hipótesis representa una de las herramientas más elegantes e importantes con las que cuenta la ciencia. La posición y amplitud de la distribución de un estadístico de muestreo y su relación con la veracidad de una determinada afirmación proveen un medio tajante de tomar una decisión y seguir provisionalmente adelante por el camino. Tal vez el problema con la significancia estadística esté en su mala utilización e interpretación, no en el concepto en sí mismo. Pero lo que anuncian estos comentarios, y creo que lamentablemente, es que se avecinan cambios en un futuro próximo.

Los críticos del uso tradicional de la significancia estadística mencionan que se trata de un problema de la percepción de la mente humana y la utilidad de operar con categorías o etiquetas. De este modo, la evaluación estadística etiqueta como relevantes o irrelevantes a las hipótesis consideradas por el investigador. Manifiestan preocupación por la omisión de los resultados “negativos” y la detección de falsos conflictos entre distintos estudios. Comentan también que la selección de los resultados por parte de los cuerpos editoriales en base a su significancia estadística involucra un sesgo y abogan por la eliminación de esas prácticas. Un amplio resumen de las recomendaciones está disponible en Wasserstein et al. (2019).

El abandono del término “estadísticamente significativo” y su reemplazo por alternativas menos afines al uso de etiquetas resulta en afirmaciones indirectas que en el fondo expresan la misma antigua semántica, pero ahora edulcorada para las nuevas recomendaciones. Por ejemplo, podemos encontrar en publicaciones que suscriben a este cambio... el índice A es 0.78. Es menor al 5% de los valores obtenidos con el modelo nulo ($p=0.04$), lo que significa que el “factor en cuestión” juega un importante rol en la estructuración de estas comunidades de “organismos” a una escala regional (i.e. entre-hábitats). Otra... si “factores determinísticos” están poco representados o si no existe una clara dominancia de estos “factores”, se espera que el valor observado del índice B se sitúe en un rango dentro del 5% y el 95% de la distribución correspondiente a la hipótesis nula. Creo que es posible expresar esas mismas afirmaciones recurriendo a los términos usuales en forma más clara y directa sin cambiar su significado.

¹Hay varias traducciones utilizadas en la literatura en castellano para referirse al concepto de “statistical significance”. Aunque “significancia” no es un término aceptado por la Real Academia española su utilización es frecuente. Por ejemplo, la búsqueda de “significancia” en <https://scielo.org/> devuelve más de 2 000 resultados que corresponden a empleos del término.

Me permito, en las líneas sucesivas, aburrirlos con un desordenado inventario de puntos relacionados a estas cuestiones y avanzar en consideraciones que apuntan a descomprimir y poner en otro contexto este material.

Uno de los puntos más importantes a considerar es que la utilización de la significancia estadística para la evaluación de hipótesis es simplemente una herramienta. Los científicos hacen ciencia usando distintas herramientas; pero la estadística aunque frecuentemente utilizada, no es necesariamente obligatoria. Como los distintos oficios cuentan con sus herramientas para generar sus productos, la ciencia recurre a herramientas propias y prestadas para generar conocimiento. Sin embargo no todas son indispensables. La esencia primaria de la actividad científica esta basada fundamentalmente en la realización de observaciones honestas (con todo lo que ello implica) y el empleo de una lógica impecable (MacArthur 1973). Con esto no quiero desestimar la utilidad de la estadística en particular ni de las evaluaciones cuantitativas en general; pero me parece oportuno señalarlo ya que la frecuencia del uso de herramientas estadísticas varía considerablemente con la disciplina de estudio. En estudios ecológicos y evolutivos pareciera ser preponderante, mientras que en estudios taxonómicos tradicionales o en estudios generales sobre distribución y conservación su empleo no es tan prominente.

Muchas de estas propuestas de cambio surgen periódicamente, a veces como respuesta orgánica de grupos pertenecientes a sociedades científicas, y otras veces como planteos individuales; que a veces ponen sobre la mesa temas que ya fueron expuestos en el pasado, pero que ahora al proponer una revisión para una audiencia renovada consideran que puede promover un cambio positivo. “Vino viejo en botella nueva” es una frase que puede resumir adecuadamente esta recurrencia temática y que Oliver Pearson solía utilizar frecuentemente para desinflar ideas pretendidamente nóveles.

Es posible que, en alguna instancia, en las distintas disciplinas que convocan a nuestros lectores, el uso inapropiado de la significancia estadística haya generado conflictos. Particularmente, uno desarrollado en ecología de comunidades durante principios la década de los '80 generó un arduo y prolongado debate, y que por aquellos años resultó en un gasto considerable de papel (Strong et al. 1984).

Para resumirlo, la evaluación de la estructura de comunidades de distintos grupos taxonómicos ofrecía evidencia clara de la existencia de procesos de competencia o ninguna señal de ella, dependiendo del tipo de modelo y del conjunto de datos analizados. El conflicto sobre el hallazgo o no de resultados relevantes dependía (entre muchas otras cosas) de la capacidad del modelo de generar distribuciones efectivamente nulas (sin influencia del proceso en juego) y de un “gran amigo” del tamaño de la muestra: el poder estadístico.

El seguimiento de ese debate (ocurrido ya hace más de 40 años y en una disciplina no convencional de la ciencia) nos enseña a reportar enfáticamente aquellos resultados significativos correspondientes a hipótesis nulas adecuadamente especificadas y a ponderar los resultados no significativos en relación al poder estadístico asociado a las evaluaciones realizadas. Muchas de las recomendaciones expresadas por Wasserstein et al. (2019) pueden darse por cumplidas atendiendo los lineamientos y críticas detalladas en los argumentos de simples ecólogos en algunos de los capítulos de Strong et al. (1984). En este punto parece no haber nada nuevo bajo el sol.

En muchas publicaciones científicas, principalmente en las de mayor impacto, el uso absoluto de la estadística ha vuelto la lectura de los artículos en una actividad fatigosa y ex-

tremadamente técnica, en la que el lector puede sorprenderse tratando de seguir el hilo de los procedimientos realizados pero desatendiendo el foco de las principales ideas evaluadas en el trabajo. Otras publicaciones, de mayor impacto aún, han relegado los extensos aspectos metodológicos a las secciones suplementarias al final del texto del artículo o en documentos suplementarios con amplia extensión y detalle. En general las contribuciones publicadas en *Mastozoología Neotropical* carecen de estos problemas, pero es un aspecto que debe considerarse para manejar adecuadamente la información contenida en las contribuciones y balancear adecuadamente su distribución entre el texto principal y el suplementario.

Con relación a lo anterior, es claro que no hay una única forma de evaluar estadísticamente una hipótesis; que no hay una única forma de analizar los datos. El umbral de probabilidad convencional del 5% puede además resultar inadecuado para algunos investigadores. La posibilidad de acceso a los datos en los que se basan las afirmaciones expresadas en una publicación y a las sentencias u órdenes (“scripts”) utilizados para resumirlos y analizarlos, es una opción cada vez más frecuente en algunas publicaciones científicas. En algunas es incluso un requisito obligatorio. En un futuro próximo, sino ya mismo, es posible que cada lector acceda a la información original, y utilizando los procedimientos empleados por los autores u otros que considere más pertinentes, establezca sus propias conclusiones sobre las ideas evaluadas.

En el transcurso de los primeros meses del 2020 estableceremos algunos nuevos lineamientos de las normas editoriales que permitan una mayor accesibilidad a la información contenida en las contribuciones publicadas en *Mastozoología Neotropical*; en principio para el acceso a secuencias de ADN, el registro de nuevas especies en Zoobank, el almacenamiento y resguardo de datos originales y la comunicación de los resultados de los resúmenes y de las evaluaciones estadísticas. Mantendremos el uso tradicional de los términos utilizados en la comunicación de evaluaciones estadísticas, solicitando información accesoria, siempre que sea posible, y que permita considerar otros criterios para la interpretación de los resultados. Creemos que estos nuevos requerimientos permitirán una mayor visibilidad, y una mayor exposición al examen por pares, para las investigaciones de aquellos autores que decidan comunicar sus hallazgos en nuestra revista.

Marcelo Kittlein

Grupo Ecología y Genética de Poblaciones de Mamíferos,
Instituto de Investigaciones Marinas y Costeras (IIMyC),
Facultad de Ciencias Exactas y Naturales,
Universidad Nacional de Mar del Plata - CONICET,
Mar del Plata, Argentina.
<kittlein@gmail.com>

LITERATURA CITADA

- AMRHEIN, V., S. GREENLAND, & B. McSHANE. 2019. Retire statistical significance. *Nature* 567, 305-307. <https://doi.org/10.1038/d41586-019-00857-9>
- HALSEY, L. G. 2019. The reign of the p-value is over: what alternative analyses could we employ to fill the power vacuum? *Biology letters* 15:20190174. <https://doi.org/10.1098/rsbl.2019.0174>
- MACARTHUR, R. H. 1973. *Geographical ecology: patterns in the distribution of species*. Princeton University Press.
- STRONG JR, D. R., D. SIMBERLOFF, L. G.ABELE, & A. B. THISTLE (EDS.). 1984. *Ecological communities: conceptual issues and the evidence*. Princeton University Press.
- WASSERSTEIN, R. L., A. L. SCHIRM, & N. A. LAZAR. 2019. Moving to a World Beyond “p <0.05”. *The American Statistician* 73:1-19. <https://doi.org/10.1080/00031305.2019.1583913>