

CARTA AL EDITOR

COMO UN ERROR EN EL USO DE LOS DATOS ALTERA LAS CONCLUSIONES

How a misuse of the data alters the conclusions

En el trabajo comentado a continuación, Camus (1995) revisa críticamente las publicaciones de la Revista Chilena de Historia Natural durante el bienio 1992-1993, identifica los errores estadísticos que se presentan, los reúne en ocho tipos de errores principales y efectúa comentarios y sugerencias de inapreciable utilidad. Lamentablemente comete un error en la confección de la Tabla 1 (ver página 15), al calcular las frecuencias, lo que a mi parecer podría afectar los juicios de Camus (1995). Por ello, aquí se pretende corregir el error, enmendar la apreciación de Camus (1995) o destacar mi propia conclusión y resaltar por último el deseo de Camus (1995) en cuanto a mejorar

la calidad de las publicaciones de la revista analizada.

Camus (1995) señala en su análisis la existencia de ocho errores estadísticos frecuentes y aporta la Tabla 1, en la cual figuran en la primera columna los tipos de error, en la segunda columna agrega el número de artículos conteniendo el error, y en la tercera incluye las frecuencias en porcentaje. En la leyenda de esta Tabla señala textual: "Frecuencia de ocurrencia de ocho tipos de errores estadísticos, asociados a la inferencia basada en pruebas de significancia, encontrados en 46 artículos de investigación...". En realidad, lo que Camus (1995) incluyó en la tercera columna es la Frecuencia Relativa (Zar, 1996) y no la Frecuencia de Ocurrencia, ya que Camus (1995) dividió los valores de la segunda columna por la suma total de errores (93) y luego lo llevó a porcentajes, es decir, usó frecuencia relativa, como se ve en la Tabla 1 que sigue:

TABLA 1

Tabla 1 de Camus (1995) : "Frecuencia de ocurrencia de 8 tipos de errores estadísticos, asociados a la inferencia basada en pruebas de significancia, encontrados en 46 artículos de investigación publicados en Revista Chilena de Historia Natural durante 1992 y 1993.

Los tipos de error se citan en la tabla en forma resumida (ver detalles en el Texto [Camus, 1995]) y no están ordenados de acuerdo a su importancia".

Table 1 by Camus (1995): "Occurrence frequency of eight types of statistical errors, related with inferences based on significance tests, found in 46 research articles published in Revista Chilena de Historia Natural during 1992 and 1993. Errors are listed below in summarized way (see details in the text [Camus, 1995]) and are not ranked according their importance".

TIPO DE ERROR	NUMERO DE ARTICULOS	
	CONTENIENDO EL ERROR	FRECUENCIA (%)
Indicar sólo N.S. sin probabilidad	16	17,2
Indicar cercanía a la significancia	0	0
No dar estimación de diferencia real	8	8,6
ANDEVA con información insuficiente	10	10,8
Probabilidad = tamaño del efecto	1	1,1
Notación deficiente o insuficiente	27	29,0
Interpretación confusa o contradictoria	17	18,3
Sin análisis estadístico	14	15,1

Tomada de Camus (1995), página 15.

Infiriendo que lo pertinente para el estudio de Camus (1995), en esta parte de su trabajo, es la frecuencia y no la frecuencia relativa en esta tabla, corresponde corregir los valores de la tercera columna y no el

encabezado de la misma. Por lo tanto, aquí sólo deben dividirse los valores de la segunda columna por 46 (número total de muestras) y multiplicar por 100, resultando la siguiente Tabla 2:

TABLA 2

Tabla 1 de Camus (1995) corregida. "Frecuencia de ocurrencia de los 8 los tipos de errores estadísticos encontrados por Camus en la publicaciones de 1992 y 1993 de la Revista Chilena de Historia Natural".

Table 1 by Camus (1995) corrected. "Occurrence frequency of 8 types of statistical errors found by Camus in the publications of 1992 and 1993 of the Revista Chilena de Historia Natural".

TIPO DE ERROR	NUMERO DE ARTICULOS	
	CONTENIENDO EL ERROR	FRECUENCIA (%)
Indicar sólo N.S. sin probabilidad	16	34,8
Indicar cercanía a la significancia	0	0
No dar estimación de diferencia real	8	17,4
ANDEVA con información insuficiente	10	21,7
Probabilidad = tamaño del efecto	1	2,2
Notación deficiente o insuficiente	27	58,7
Interpretación confusa o contradictoria	17	37,0
Sin análisis estadístico	14	30,4

Modificada de Camus (1995), página 15.

Durante el análisis de su Tabla 1, Camus (1995) sólo destaca dos aspectos interesantes, por un lado los tipos de errores más frecuentes y por otro lado los tipos de errores menos frecuentes. Este análisis somero de Camus (1995) podría deberse a que su Tabla 1 (ver Tabla 1 de este trabajo) presenta sólo valores pequeños, siendo el más alto de un 29%. Al observar la Tabla 2, vemos valores superiores al 50%, lo que es un panorama distinto al apreciado por Camus (1995) y que es analizado a continuación por conjuntos de tipos de errores, en los cuales se agrupan los tipos de errores por cercanía en sus valores de frecuencia de ocurrencia en porcentaje.

En el primer conjunto sólo se incluye el tipo de error de mayor frecuencia, el cual se comete en el 58,7% de los trabajos publicados. Esto es indicio de que este tipo de error (Notación deficiente o insuficiente) se comete prácticamente en tres de cada cinco

trabajos publicados. Emitir juicios sobre este porcentaje es absolutamente subjetivo, pero me atrevo a pecar afirmando que es grave. El tipo de error puede deberse a descuido, pero su frecuencia es superior al 50%, lo que indica que es un descuido casi generalizado entre los investigadores analizados.

Un segundo conjunto de tipos de errores que destacan en esta revisión de Camus (1995) son "Interpretación confusa o contradictoria", "Indicar sólo N.S. sin probabilidad", y "Sin análisis estadístico" con valores de 37,0%, 34,8% y 30,4% respectivamente, lo que indica que en aproximadamente uno de cada tres artículos se comete alguno de estos tipos de errores. Esto resulta grave si pensamos que en cada publicación efectuada al menos uno de estos tipos de errores estará presente.

Un tercer conjunto de tipos de errores son "ANDEVA con información insuficiente" y "No dar estimación de diferencia

real” los que alcanzan a 21,7 % y 17,4 % respectivamente. Esto implica que estos tipos de errores se cometen aproximadamente en uno de cada cinco publicaciones.

En el cuarto conjunto he agrupado aquellos tipos de errores de menor frecuencia, que son “Indicar cercanía a la significancia” y “Probabilidad = tamaño del efecto” los que alcanzan valores de 0% y 2,2% respectivamente.

Lo que para Camus (1995) “...podría estar relacionado simplemente con descuido o errores de tipeo...”, para mí es grave, ya que más de la mitad de los trabajos que contienen información cuantitativa efectúan una notación deficiente o insuficiente. Es más, los resultados de la tabla analizada indican que por cada publicación existirán al menos dos tipos de errores presentes, en aquellos trabajos que usen herramientas estadísticas. Esta reciente afirmación se basa en que la frecuencia de ocurrencia acumulada, para los ocho tipos de errores, alcanza al 202,2%. Camus (1995) deja entrever que la notación deficiente puede estar relacionada con los otros tipos de errores que le siguen en aparición “...sugiriendo que la notación podría estar reflejando un manejo o conocimiento poco adecuado del significado de los elementos de la prueba (e.g., estadígrafos...)”. Esto es en realidad un mal uso de las herramientas estadísticas. Las altas frecuencias de ocurrencia de los tipos de errores agrupados

en los conjuntos 2 y 3, que son “Interpretación confusa o contradictoria” (37%), “Indicar sólo N.S. sin probabilidad” (34,8%), “Sin análisis estadístico” (30,4%), “ANDEVA con información insuficiente” (21,7%) y “No dar estimación de diferencia real” (17,4%) indican que el conocimiento o uso de la estadística es deficiente, por parte de los autores responsables de las publicaciones analizadas por Camus (1995).

Por la inapreciable utilidad del Editorial de Camus (1995), creo conveniente destacar la frase final de este: “La responsabilidad de mejorar la calidad del análisis de datos recae principalmente en los autores de un trabajo, pero también sobre los revisores y editores, que debieran involucrarse íntimamente en el proceso”.

LITERATURA CITADA

- CAMUS, PA (1995) Uso e importancia de la inferencia estadística: algunos problemas frecuentes detectados en la Revista Chilena de Historia Natural. Revista Chilena de Historia Natural 68: 12-18.
- ZAR, JH (1996) Biostatistical analysis. Prentice-Hall International, Inc., New Jersey, U.S.A. 662 pp. + 256 pp. Apéndice + x.

PEDRO A. VERGARA

Programa de Magíster en Ciencias,
mención: Producción, Manejo y
Conservación de Recursos Naturales
Universidad de Los Lagos, Casilla 933, Osorno, Chile