

Intervalos de confianza, pes y farolas

I. García-Doval

Servicio de Dermatología. Complejo Hospitalario de Pontevedra. España.

Sr. Director:

El grupo de los hospitales General Universitario y San Juan de Alicante ha venido publicando en *Actas* una serie de trabajos bibliométricos muy útiles, laboriosos e interesantes^{1,2}. Con la única intención de mejorarlos señalo un defecto que se reitera en los mismos, y que sin ser grave, podría indicar una debilidad metodológica en el proceso de edición de *Actas*.

La situación habitual en investigación es que estudiemos una muestra y queramos obtener resultados aplicables a la población de la que la muestra ha surgido. Como partimos de una muestra hay un error aleatorio asociado a nuestra medición, que pretendemos cuantificar mediante la estimación estadística. La forma más frecuente de hacerlo es mediante el empleo de intervalos de confianza, con los que tenemos el valor obtenido en nuestra muestra y un rango de valores entre los que es probable que se encuentre el valor real de la población estudiada.

Pero excepcionalmente tenemos resultados no de una muestra, sino de toda la población. En este contexto la estimación no tiene sentido: ya sabemos cuál es el resultado en la población

sin ningún género de dudas. Ésta es la situación de los artículos referidos. Si, por ejemplo, se estudian todos los artículos de *Actas* del trienio 2003-2005 y se quieren comparar con todos los del trienio 2000-2002, no es necesaria la estimación, y los intervalos de confianza aportados en el artículo son innecesarios y falsos (parten de la base de que estamos estudiando una muestra de una población mucho mayor): el intervalo de confianza real sólo incluye el valor encontrado.

Lo mismo se puede decir del contraste de hipótesis: el valor de p nos indica la probabilidad de que encontremos los resultados de nuestra muestra siendo cierta la hipótesis nula (generalmente de ausencia de diferencia) y considerando la existencia de una variabilidad aleatoria asociada al muestreo. Es aplicable cuando se estudian muestras. En este caso se estudian las poblaciones completas y no tiene sentido utilizar contraste de hipótesis (todas sus p «serían» cero cuando hay alguna diferencia y uno cuando no la hay). Este caso sirve, además, de ejemplo del valor real de la p : todas las diferencias que ustedes encuentran son estadísticamente significativas ($p = 0$), la dificultad está en sa-

ber si lo son desde el punto de vista «clínico», algo que la estadística no puede decir.

Creo que los trabajos mejorarían si se eliminasen las p y los intervalos de confianza, y al describir los datos se añadiesen, en su lugar, medidas de la dispersión (DE, percentiles o rangos).

Como anécdota jocosa, me viene a la cabeza el excelente humor inglés de Castle³, quien en su libro de introducción a la estadística decía que los médicos solemos utilizar la estadística como los borrachos las farolas: para apoyarnos, no para iluminarnos. Un cordial saludo.

Bibliografía

1. Belinchón I, Miralles J, Ramos JM. Análisis de la producción científica de la revista *Actas Dermo-sifiliográficas* en el trienio 2003-2005. *Actas Dermosifiliogr.* 2006;97:428-35.
2. Ramos JM, Belinchón I, Miralles J, Ballester R, Sánchez-Yus E. Análisis de la producción científica nacional e internacional de los dermatólogos españoles (1988-2000). *Actas Dermosifiliogr.* 2006;97:436-43.
3. Castle WM. *Statistics in operation*. Edinburgh: Churchill Livingstone; 1979. p. 4.