

## Distribución Normal ¿Es tan frecuente como parece?

### *Normal Distribution: Is it as frequent as it seems?*

#### **Sr. Editor:**

Un gran número de artículos científicos no indican la distribución de las variables que estudian, dejándolas sobreentendidas como normales. Al no describir la distribución de las variables y hacerlo mediante medias (promedios) y desviaciones estándar, se asume una distribución normal. Si bien muchas variables tienden a la normalidad, es improbable que todos los estudios publicados tengan esta característica, sobre todo cuando el número de observaciones descritas es pequeño.

La descripción individual de cada variable es el primer paso en el análisis de datos, en cualquier estudio cuantitativo<sup>1</sup>. Ésta refleja la habilidad de los autores para caracterizar los datos recolectados, condensar la información y presentarla de manera simplificada a los lectores, de diversas formas: porcentajes, frecuencias, medias, medianas, desviaciones estándar, rangos o percentiles.

Entre las mencionadas, la media (o promedio) es una medida útil, debido a su simpleza para entender la tendencia central de una distribución. Lamentablemente, es afectada de manera importante por los valores extremos de distribuciones sesgadas. Por otro lado, la desviación estándar permite entender fácilmente la dispersión de las observaciones, pero al ser un valor único no puede describir adecuadamente a las curvas asimétricas<sup>1,2</sup>. Debido a ello, en algunas ocasiones se requieren otros métodos para describir la información.

Identificar el tipo de variable y su distribución permite elegir adecuadamente el camino a seguir: usar pruebas paramétricas o no paramétricas. Los métodos estadísticos paramétricos requieren de supuestos teóricos, como la normalidad y homogeneidad de varianzas, que si no se cumplen pueden invalidar los resultados.

Si bien los análisis no-paramétricos y/o de libre distribución<sup>3</sup> tienen menor poder estadístico, son los únicos disponibles para tratar variables categóricas, pues ninguna prueba paramétrica es aplicable en tal situación. Así mismo, si el tamaño de muestra es muy pequeño, puede no haber otra opción que usar estas pruebas. Las pruebas no paramétricas requieren cumplir menos suposiciones de los datos, por lo que son las indicadas para este tipo de variables<sup>4</sup>.

La falta de normalidad puede afectar la homogeneidad de varianzas, sobre todo cuando hay

mucha diferencia entre los grupos, recomendándose aplicar transformaciones para estabilizar la varianza y normalizar la respuesta<sup>3,5</sup>. Es importante estudiar la distribución de los residuos evaluando su simetría, aplicando la transformación de datos sólo si esta es asimétrica. Cuando la heterogeneidad de varianzas es muy irregular, es decir, cuando algunos grupos poseen mayor variabilidad que otros, sin que exista una relación entre medias y varianzas, se recomienda usar métodos no paramétricos<sup>3</sup>.

Los análisis realizados con presunciones incorrectas pueden llegar a resultados e interpretaciones erróneos, que afectan la calidad de la investigación, por lo que creemos que explorar esta línea de investigación será provechoso para valorar algunas de las limitaciones de los estudios y mejorar la calidad de la publicación científica regional.

Es fundamental identificar el tipo y distribución de las variables, resumirlas con la media y desviación estándar sólo cuando corresponda, y no asumir que todas tienen una distribución normal. La distribución normal es rara en algunas especialidades médicas<sup>6</sup>, mientras que la distribución no normal se encuentra frecuentemente entre la información a analizar, debiendo ser trabajada como tal y no ignorada.

*Jorge Osada, José Luis Rojas, Lupe Vidal  
Unidad de Epidemiología Clínica, Facultad de  
Medicina Alberto Hurtado, Universidad Peruana  
Cayetano Heredia. Lima, Perú.*

### Referencias

1. Whitley E, Ball J. Statistics review 1: presenting and summarising data. *Crit Care* 2002; 6 (1): 66-71.
2. Altman DG, Bland JM. Presentation of numerical data. *BMJ* 1996; 312 (7030): 572.
3. Calzadilla J, Guerra W, Torres V. Uso y abuso de las transformaciones matemáticas: Aplicaciones en modelos de análisis de varianza. *Rev Cubana Cienc Agri* 2002; 32 (2): 103-7.
4. Manikandan S. Data transformation. *Pharmacol Pharmacother* 2010; 1 (2): 126-7.
5. Siegel S, Castellan NJ. Estadística no paramétrica: aplicada a las ciencias de la conducta. 4ta edición. México: Trillas; 1995 (reimp 2007). p. 55-8.
6. Altman DG, Bland JM. Statistics notes: the normal distribution. *BMJ* 1995; 310 (6975): 298.

Los autores declaramos no tener conflictos de intereses en el presente documento.

Correspondencia a: Dr. Jorge Osada Lij. Ca. Alfredo Salazar 472 depto. "201", Lima 27, Lima, Perú. Tel: 51-1-992304701  
E-mail: j\_osada@hotmail.com.