



## **Consideraciones ante el uso de la prueba de Shapiro-Wilk cuando se trabaja con muestras pequeñas**

**Considerations when using the  
Shapiro-Wilk test when working  
with small samples**

10.20960/angiologia.00568

03/01/2024

## **Consideraciones ante el uso de la prueba de Shapiro-Wilk cuando se trabaja con muestras pequeñas**

### ***Considerations when using the Shapiro-Wilk test when working with small samples***

Señor director:

En la publicación de Caicedo Valdés y colaboradores (1), se trabajó con muestras de tamaño pequeño (entre 2 a 8 sujetos, según el grupo). Para poder establecer si los datos seguían o no una distribución normal utilizaron la prueba de Shapiro-Wilk, su resultado implicó que usaran las prueba t de Student y la prueba de Pearson para sus análisis de datos. El uso de ambas pruebas es el resultado de tener una distribución normal de los datos.

Aquí creemos que es importante realizar la siguiente aclaración: la prueba de Shapiro-Wilk (SW) es altamente sensible al tamaño de la muestra: por ejemplo, podría indicar que no existe una distribución normal en una muestra muy grande, ya que puede detectar desviaciones no importantes de la normalidad y, a partir de ellas, reportar un valor de  $p$  muy pequeño, lo que indicaría una errónea no normalidad en la distribución de los datos. Por el contrario, cuando se tiene una muestra de tamaño pequeño, esta prueba puede no considerar desviaciones importantes, lo que implica que no siempre puede discriminar ante la no normalidad (2,3). Es decir, cuando el tamaño muestral es pequeño, la prueba de SW puede señalar que la distribución es normal, aun cuando esta no lo sea realmente.

Dado lo anterior, el uso de la prueba de SW siempre debe contrastarse con alguna representación gráfica, como Q-Q *plot* (los puntos deben estar sobre la recta) o un histograma (las barras deben seguir la distribución de la curva normal por superposición) (3,4), cuando el tamaño muestral sea pequeño; se considera un tamaño muestral pequeño cuando este es inferior a 30 unidades (5).

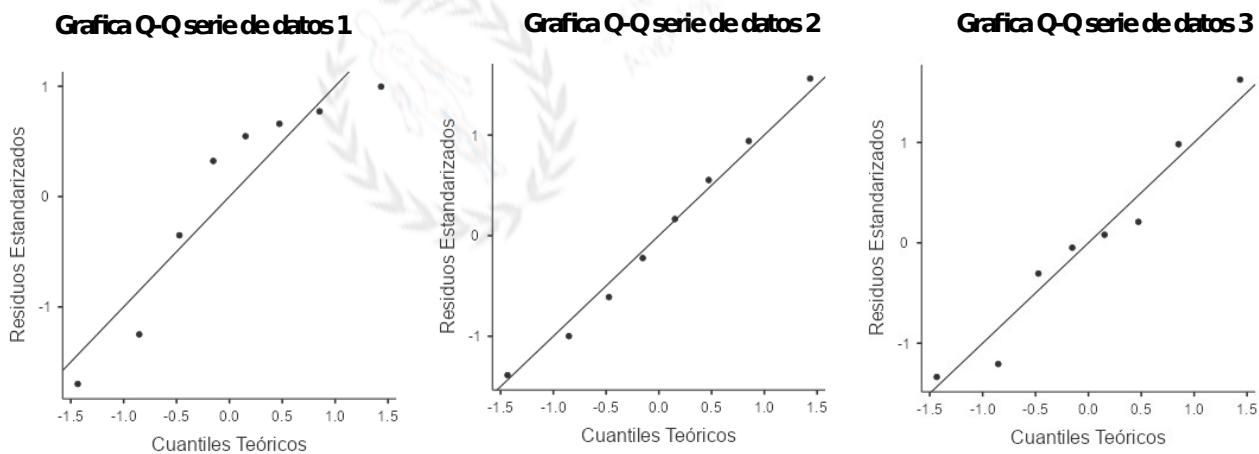
En la figura 1 podemos ver un ejemplo en el que tenemos tres series de datos con un  $n = 8$ . En todos ellos, al aplicar la prueba de Shapiro-Wilk, nos indica que la distribución es normal ( $p > 0,05$ ), pero cuando contrastamos estos resultados con su correspondiente grafica Q-Q *plot*, solo la serie 2 efectivamente la tiene.

En el área de ciencias de la salud es muy común trabajar con muestras de tamaño pequeño, y dado que el tipo de distribución va a determinar las pruebas estadísticas a utilizar, siempre es recomendable contrastar los resultados obtenidos en la prueba de Shapiro Wilk con representaciones gráficas para evitar potenciales errores en el análisis de datos y en sus respectivas conclusiones.

**Serie de datos 1**  
60; 160; 170; 180; 120; 150; 165; 80  
Valor  $p$  de Shapiro-Wilk: 0,134

**Serie de datos 2**  
55;25;45;40;50;35;30;63  
Valor  $p$  de Shapiro-Wilk: 0,976

**Serie de datos 3**  
150;180;190;250;170;80;300;70  
Valor  $p$  de Shapiro-Wilk: 0,741



**Figura 1.** Comparación entre p-valor y grafico Q-Q *plot* para determinar la normalidad de una serie de datos de  $n = 8$ .

Ángel Roco-Videla<sup>1</sup>, Sergio Vladimir Flores<sup>2</sup>, Mariela Olguin-Barraza<sup>3</sup>, Raúl Aguilera-Eguía<sup>4</sup>

<sup>1</sup>Facultad de Salud y Ciencias Sociales. Universidad de las Américas. Santiago, Chile. <sup>2</sup>Facultad de Ciencias de la Salud. Universidad Autónoma de Chile. Santiago, Chile. <sup>3</sup>Facultad de Ciencias de Salud. Programa de Magíster en Ciencias Químico-Biológicas. Universidad Bernardo O'Higgins. Santiago, Chile. <sup>4</sup>Departamento de Salud Pública. Facultad de Medicina. Universidad Católica de la Santísima Concepción. Concepción, Chile

*Conflictos de interés: los autores declaran no tener conflictos de interés.*

## BIBLIOGRAFÍA

1. Caicedo Valdés D, Devesa Peleteiro P, Pazos González A, Romero Vázquez L, Martínez Pérez M, Arce Vázquez VM, et al. Evaluación de la reperfusión tras angiogénesis terapéutica con fotopletismografía en un modelo murino de isquemia de miembros inferiores. *Angiología* 2019;71(4):127-34. DOI: [10.20960/angiologia.00002](https://doi.org/10.20960/angiologia.00002)
2. Sainani KL. Dealing with non-normal data. *PM R* 2012;4(12):1001-5. DOI: [10.1016/j.pmrj.2012.10.013](https://doi.org/10.1016/j.pmrj.2012.10.013)
3. Pedrosa I, Juarros-Basterretxea J, Robles-Fernández A, Basteiro J, García-Cueto E. Pruebas de bondad de ajuste en distribuciones simétricas, ¿qué estadístico utilizar? *Univ Psychol* 2014;14(1):245-54. DOI: [10.11144/javeriana.upsy14-1.pbad](https://doi.org/10.11144/javeriana.upsy14-1.pbad)
4. Razali NM, Wah, YB. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics* 2011;2(1):21-33. Disponible en <https://www.nrc.gov/docs/ML1714/ML17143A100.pdf>
5. Hogg R, Tanis E, Zimmerman D. Probability and statistical inference. 10.<sup>a</sup> edición. Upper Saddle River (NJ), Estados Unidos: Pearson; 2019.