

Estimación paramétrica de la confiabilidad y diferencias confiables

Confidence interval methods and its reliability

Sr. Editor:

Es común entre los profesionales de la salud elegir instrumentos que se adecúen a las características del grupo a evaluar, que tengan un marco teórico sólido y propiedades psicométricas adecuadas, es decir, que las inferencias que se realizarán sean válidas y sus puntajes confiables. En este sentido, el WHOQoL diseñado por la Organización Mundial de la Salud (OMS) para la evaluación de la calidad de vida (CV), incluyendo su versión breve (WHOQoL-Bref) estudiada por Urzúa y Caqueo-Urizar¹, no debe escapar de esta pretensión, especialmente si se usará como línea base para programas de intervención y estudios longitudinales de CV.

Urzúa y Caqueo-Urizar brindan un aporte significativo al estudiar psicométricamente el WHOQoL-Bref en Latinoamérica, lo cual abre las puertas para que otros investigadores de la región puedan tomarlo como base para realizar sus propios estudios. Sin embargo, dos aspectos fueron omitidos: la discusión rigurosa de los coeficientes de confiabilidad y el impacto de sus magnitudes en la interpretación de los perfiles de calidad de vida. Estos aspectos serán desarrollados en la presente carta.

En el análisis de la confiabilidad por consistencia interna, el coeficiente alfa (α) indica la proporción de error de medición presente al aplicar un instrumento en determinado grupo², por lo cual puede ser un criterio útil para seleccionar un test psicológico, entre otros, considerando ciertos estándares de adecuación del coeficiente α . Evidentemente, los criterios para evaluar un coeficiente α varían según el autor de referencia, pero parece llegarse al consenso que, aun para fines de investigación, se exige un valor mínimo de 0,70 para considerarlo adecuado³. En tal sentido, Urzúa y Caqueo-Urizar reportaron los siguientes coeficientes: 0,71 para la dimensión Salud Física; 0,75 para la dimensión Salud Psicológica; 0,63 para la dimensión Social y 0,76 para la dimensión Ambiental. Entre ellos, uno se encuentra por debajo del límite, otro al límite, y dos aparentemente elevados. Sin embargo, para generalizar

estas estimaciones a las poblaciones de referencia, es necesario calcular intervalos de confianza. El intervalo de confianza (IC) para el coeficiente α es un aspecto que debe reportarse a fin de tener un estimador probabilístico del "verdadero" valor α en la población⁴. Esta información paramétrica puede servir para evaluar la significación de los cálculos estadísticos obtenidos en la muestra (en este caso, el coeficiente α). Cabe mencionar que se espera que el límite inferior del IC para el coeficiente α sea mayor que 0,70³. Mediante el programa ICalfa⁵ se realizaron los cálculos de los IC (nivel de confianza de 95%) para cada α bajo el método Fisher⁶, obteniéndose lo siguiente: para Salud Física: 0,688-0,731; para Salud Psicológica: 0,730-0,768; para la dimensión Social: 0,603-0,656; y para la dimensión Ambiental: 0,741-0,778. Estos resultados indican que las dimensiones Salud Psicológica y Ambiental mantienen un nivel de confiabilidad poblacionalmente superior a 0,70, mientras que la dimensión Social puede alcanzar niveles debajo de 0,70, ya que tiene todo su IC debajo de 0,70.

Aunque la mayoría de los coeficientes α supera el valor de 0,70, para aplicaciones clínicas y de intervención individual, magnitudes altas de confiabilidad son la regla más que la excepción, especialmente si se planea realizar comparaciones en base a perfiles^{3,7}. Por lo tanto, los resultados de las comparaciones grupales reportadas están involucrando magnitudes de error de medición que pueden limitar las interpretaciones (por ejemplo, en las diferencias entre perfiles). Para corroborar esto, y basándonos en el trabajo de Charter⁸, se puede usar una prueba estadística de la diferencia confiable entre dos puntajes dentro de un mismo instrumento; esto es, determinar si la diferencia (D) entre dos puntajes es lo suficientemente grande que la producida por error de medición. El procedimiento se respalda en la siguiente formulación:

$$Z_{conf} = \frac{D}{\sqrt{EEM_1^2 + EEM_2^2}}$$

Donde se requiere del error estándar de medición ($EEM = DE_x \sqrt{1 - r_{xx}}$), la confiabilidad (r_{xx} en este caso los coeficientes α) y la dispersión (DE) de los puntajes. Modificando la ecuación anterior, se obtiene la fórmula que permitirá determinar

la diferencia confiable necesaria en los niveles de confianza de 80% ($z = 1,64$), 95% ($z = 1,96$) y 99% ($z = 2,58$):

$$D = Z \sqrt{EEM_1^2 + EEM_2^2}$$

A fin de ver el impacto de los coeficientes de confiabilidad sobre la configuración de un perfil, como ejemplo se considerarán dos escalas (Dimensión Ambiental y Social) que corresponden con el mayor y menor coeficiente α reportado, y el grupo de varones de 20 a 29 años. Asumiendo que los puntajes de ambas escalas se estandarizan a puntajes T ($M = 50$ y $DE = 10$), el error estándar de medición de ambos factores es 4,899 y 6,083, respectivamente. Finalmente, la diferencia confiable requerida entre los factores, en los niveles 80%, 95% y 99% es, respectivamente, 13, 15 y 20 puntos. Para ejemplificar, se considerará el caso más favorable, es decir, el perfil de un sujeto en el cual todos los puntajes se acercan al promedio del grupo de pertenencia, y siendo la Dimensión Ambiental y Social las escogidas, sus puntajes serán 59 y 68. Ante ello, se desea establecer si la diferencia entre ambos puntajes es lo suficientemente grande como para poder interpretarse en el marco de un perfil integrado de calidad de vida. Eligiendo un nivel de confianza de 95%, se requiere una diferencia mínima de 15 puntos para establecer que la misma es confiable y, considerando que la actual es de 9 puntos, esta puede ser interpretada como un efecto del error de medición y no requerir mayor atención. Si el investigador requiere alta precisión, se requerirá una diferencia mínima de 20 puntos (99% de confianza). Sin embargo, con confiabilidades de 0,90 o más, las discrepancias mínimas confiables serán menores debido al reducido error de medición.

*Sergio Alexis Domínguez Lara,
César Merino Soto, Jhonatan S. Navarro Loli
Instituto de Investigación de Psicología,
Universidad de San Martín de Porres.*

Referencias

1. Urzúa A, Caqueo-Urizar A. Estructura factorial y valores de referencia del WHOQoLBref en población adulta chilena. *Rev Med Chile* 2013; 141: 1547-54.
2. Ledesma R. Análisis de consistencia interna mediante Alfa de Cronbach: un programa basado en gráficos dinámicos. *Psico-USF* 2002; 7 (2): 143-52.
3. Merino C, Navarro J, García W. Revisión de la consistencia interna del Inventario de Inteligencia Emocional de Bar-On, EQ-I: YV. *Revista Peruana de Psicología y Trabajo Social* 2014; 3 (1): 141-54.
4. Fan X, Thompson B. Confidence intervals about score reliability coefficients, please: an EPM guidelines editorial. *Educ Psychol Meas* 2001; 61: 517-31.
5. Domínguez S, Merino C. ¿Por qué es importante reportar los intervalos de confianza del coeficiente alfa de Cronbach? *Revista Latinoamericana de Ciencias Sociales, niñez y Juventud* 2015; 13 (2): 1326-8.
6. Romano JL, Kromrey JD, Owens CM, Scott HM. Confidence interval methods for coefficient alpha on the basis of discrete, ordinal response items: Which one, if any, is the best? *J Exp Educ* 2011; 79 (4): 382-403.
7. Nunnally JC, Bernstein IJ. *Teoría Psicométrica*. 3ª ed. México D.F.: McGraw-Hill Latinamericana; 1995.
8. Charter AR. Formulas for reliable and abnormal differences in raw test scores. *Percept Motor Skill* 1996; 83 (3): 1017-8.

Correspondencia a:

Sergio Alexis Domínguez Lara
Av. Tomás Marsano 242 (5° piso), Lima 34-Perú.
sdominguezmpcs@gmail.com

Nota: Se ofreció posibilidad de réplica a los autores aludidos, sin recibir su respuesta.