

ERROR ESTÁNDAR DE MEDIDA Y LA PUNTUACIÓN VERDADERA DE LOS TESTS PSICOLOGICOS: CÁLCULO MEDIANTE UN MÓDULO EN VISUAL BASIC

Sergio Alexis Dominguez Lara*

Universidad Inca Garcilaso de la Vega

RESUMEN

Este trabajo presenta un módulo en lenguaje Visual Basic para la estimación de intervalos de confianza para la puntuación verdadera de los tests psicológicos. La teoría clásica de los tests postula que la puntuación observada de una persona evaluada mediante un tests psicológico está conformada por la puntuación verdadera y un margen de error. En tal sentido, para poder realizar una aproximación a la *puntuación verdadera* de una persona a partir de la puntuación observada, el uso del error estándar de medida es una estrategia habitual. El uso adecuado de este método y su falta de implementación en los paquetes estadísticos tradicionales justifica la construcción del software como una herramienta para el evaluador, para que pueda tomar decisiones más adecuadas en base a sus hallazgos en los tests psicológicos. Se discute el uso y las limitaciones de la herramienta.

Palabras clave: error estándar de medida, puntuación verdadera, confiabilidad, software.

THE STANDARD ERROR OF MEASUREMENT AND THE TRUE SCORE OF PSYCHOLOGICAL TESTS: CALCULATION USING A VISUAL BASIC MODULE

ABSTRACT

This paper presents a module in Visual basic to calculate the confidence interval for the true score of psychological tests. The classical test theory postulates that a person's observed score assessed by psychological tests consists of the true score and an error. In this regard, in order to make an approach to a person's true score from the observed score, the use of the standard error of measurement is a common strategy. Proper use of this method and its failure to implement in traditional statistical packages justifies the construction of software as a tool for the evaluator, for to take more appropriate decisions on the basis of their findings in psychological tests. The use and limitations of the tool was discussed.

Keywords: standard error of measurement, true score, reliability, software.

ERRO PADRÃO DE MEDIDA E ESCORE VERDADEIRO TESTES PSICOLÓGICOS: CÁLCULOS USANDO UM MÓDULO EM VISUAL BASIC

RESUMO

Este trabalho apresenta um módulo em linguagem Visual Basic para a estimativa de intervalos de confiança para os verdadeiros testes psicológicos de pontuação. Teoria teste clássico postula que a pontuação observada de uma pessoa avaliada por testes psicológicos consiste na pontuação verdadeira e uma margem de erro. A este respeito, a fim de fazer uma aproximação do valor verdadeiro de uma pessoa a partir do valor observado, usando o erro padrão de medição é uma estratégia comum. O uso adequado deste método e sua não implementação em pacotes estatísticos tradicionais justifica a construção de software como uma ferramenta para a avaliação, para que possa tomar as melhores decisões com base em suas descobertas em testes psicológicos. São discutidos o uso e as limitações da ferramenta.

Palavras-chave: erro padrão da medição, pontuação verdade, a confiabilidade do software.

*Docente investigador de la Facultad de Psicología y Trabajo Social de la Universidad Inca Garcilaso de la Vega. Magíster en Psicología. Candidato a Doctor por la Facultad de Psicología de la Universidad Nacional Mayor de San Marcos. E-mail: sdominguezmpcs@gmail.com

La revisión y el control del error asociado a la medición por medio de los tests psicológicos es uno de los objetivos fundamentales de la psicometría. En tal sentido, cuantificar e informar las aproximaciones a dicho error es fundamental, sobre todo cuando los instrumentos de evaluación psicológica son usados como apoyo a la toma de decisiones (Gempp, 2006). Para la delimitación del error en la medición se usa un elemento clave tanto en el proceso de construcción un instrumento, como en la decisión de usarlo: la confiabilidad del test.

La teoría clásica de los tests postula que toda *puntuación observada* de una persona evaluada mediante un tests psicológico está conformada por la *puntuación verdadera* y un *margen de error* (Muñiz, 2001). Aunque parezca sencillo, el objetivo principal que se mencionaba anteriormente, de cuantificar e informar el error, no se puede hacer por medio de una simple observación, tiene que aislarse por medio de otros métodos, y es ahí donde la confiabilidad brinda su aporte.

En una situación hipotética en la cual pudiéramos evaluar a una persona con un determinado test infinitas veces, el promedio de todas las aplicaciones sería una aproximación a la puntuación verdadera, y cualquier discrepancia observada entre las aplicaciones es el error en dicha medición (Gempp, 2006). Entonces, se entiende al error de medición como las discrepancias que se observan tras someter a una persona a n aplicaciones de un determinado test psicológico.

El error estándar de medida hace referencia a la desviación estándar de la distribución de las puntuaciones del error (Gempp, 2006, Cortada de Kohan, 1999), y puede calcularse por medio de la Ecuación (1):

$$\text{EEM} = \text{SD} \sqrt{1 - \text{confiabilidad}} \quad (1)$$

Para ejemplificar este y los demás procedimientos se tendrá estos datos de base:

Puntuación observada del sujeto: 12
Media del grupo: 10

SD: 2

Confiabilidad: 0.90

En tal sentido, el EEM calculado por medio de la Ecuación (1) sería de 0.63.

Para poder realizar una aproximación a la *puntuación verdadera* de una persona a partir de la *puntuación observada* en alguna evaluación, el uso del error estándar de medida es una estrategia habitual. Gempp (2006) describe dos aproximaciones, la tradicional y la basada en la regresión lineal.

La primera se basa en la estimación del intervalo de confianza de la *puntuación verdadera* solo en función de la *puntuación observada* en el test psicológico, y puede apreciarse por medio de las Ecuaciones (2a) y (2b).

$$\begin{aligned} \text{Limite inferior} &= x - z (\text{EEM}) & (2a) \\ \text{Limite superior} &= x + z (\text{EEM}) & (2b) \end{aligned}$$

Donde:

x = puntuación observada

z = valor de la distribución normal asociado

EEM = Error estándar de medida

Poniendo un ejemplo con los datos citados anteriormente, el intervalo de confianza va de 10.76 a 13.24. Esto quiere decir que hay un 95% de confianza de que ese intervalo incluya la *puntuación verdadera* de la persona, o en otras palabras, la oscilación probable de dicha *puntuación verdadera*.

Por otro lado, el segundo método, el que utiliza la regresión lineal, se enfoca en un intervalo de confianza utilizando la estimación de la *puntuación verdadera* como punto central, lo cual se aprecia en la Ecuación (3), mas como todo método de regresión tiene asociado un error de predicción, el cual se conoce como *error estándar de estimación* (EEE), el cual se puede observar en la Ecuación (4).

$$v = \text{confiabilidad} (x - X) + X \quad (3)$$

Donde:

x = puntuación observada

X = promedio de las puntuaciones del test en la muestra

$$EEE = EEM \sqrt{\text{confiabilidad}} \quad (4)$$

Realizando los cálculos respectivos con los datos brindados anteriormente, la puntuación verdadera estimada es de 11.8 y el EEE es de .6. Aunque recogiendo las recomendaciones de Morales-Vallejo (2007), la *puntuación verdadera* sería la *puntuación más probable*, es decir, no es algo fijo. En tal sentido, el *intervalo de confianza de la puntuación verdadera* en función de este método queda definido por las Ecuaciones (5a) y (5b), y va de 10.62 a 12.98.

$$\text{Limite inferior} = v - z (EEE) \quad (5a)$$

$$\text{Limite superior} = v + z (EEE) \quad (5b)$$

Donde:

v = puntuación verdadera estimada (Ecuación 3)

z = valor de la distribución normal asociado

EEM = Error estándar de estimación

Para una revisión más exhaustiva de estos métodos, así como de sus fundamentos teóricos, el lector interesado puede revisar los manuscritos de Aiken (2003), Gempp (2006), Morales-Vallejo (2007) y Cortada de Kohan (1999). Con relación a

los intervalos de confianza, la mención de los mismos es una práctica exigida por los estándares de reportes de investigación de la American Psychological Association (2010).

El objetivo del presente reporte es presentar un programa en lenguaje Visual Basic 6.0, compatible con todas las versiones de Windows, el cual sirva como método para el cálculo de los intervalos de confianza para la estimación de la puntuación verdadera siguiendo el esquema conceptual presentado anteriormente.

Elementos de la herramienta

El programa propuesto¹ para el cálculo de los intervalos de confianza para la estimación de la puntuación verdadera presenta cuatro secciones (Figura 1). La primera de ella se refiere al ingreso de los datos (la confiabilidad del instrumento, el promedio de la muestra, la desviación estándar y la puntuación observada).

En la segunda sección está orientada al cálculo del EEM y del EEE, como paso previo al cálculo de los intervalos de confianza. La tercera sección se presenta la posibilidad de estimar el intervalo de confianza mediante el método tradicional, y la cuarta y última sección está enfocada en el método por medio de la regresión lineal (incluye la

¹La herramienta está disponible al usuario sin costo escribiendo al autor del trabajo.

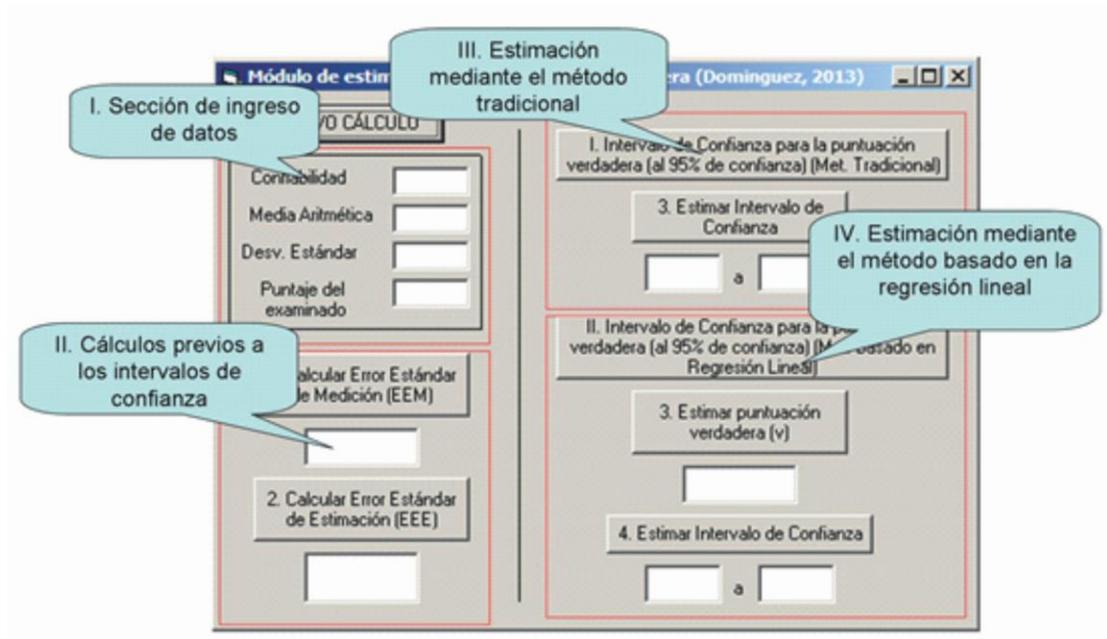


Figura 1. Presentación del módulo.

Aplicación de la herramienta

Para ejemplificar la aplicación de la herramienta, se utilizan datos provenientes del ejemplo

anterior. Luego de ello, se procede a calcular el intervalo de confianza mediante cada uno de los métodos indicados anteriormente, el tradicional y el basado en la regresión lineal.

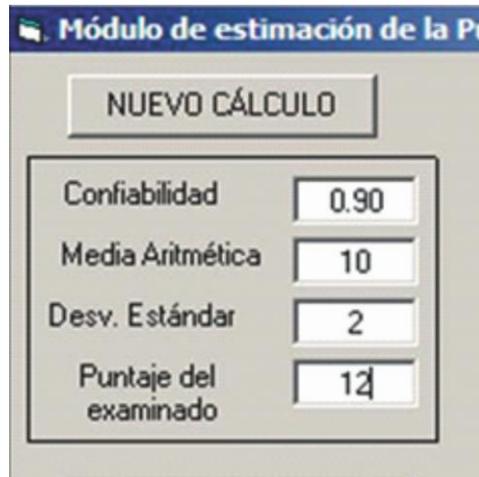


Figura 2. Vista de los datos colocados en el programa.

Se procede al cálculo del intervalo de confianza por medio del método tradicional, realizando

previamente el EEM.

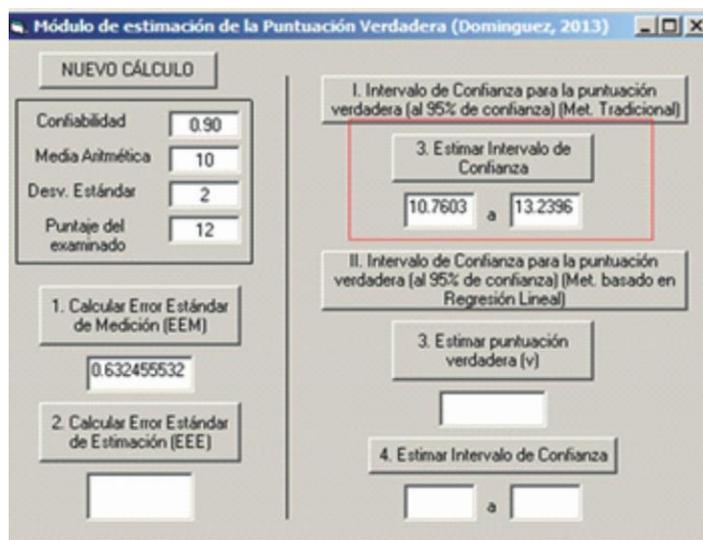


Figura 3. Vista de los resultados en el programa (método tradicional).

Por otro lado, se presenta el cálculo del intervalo de confianza mediante el método basado en la regresión lineal.

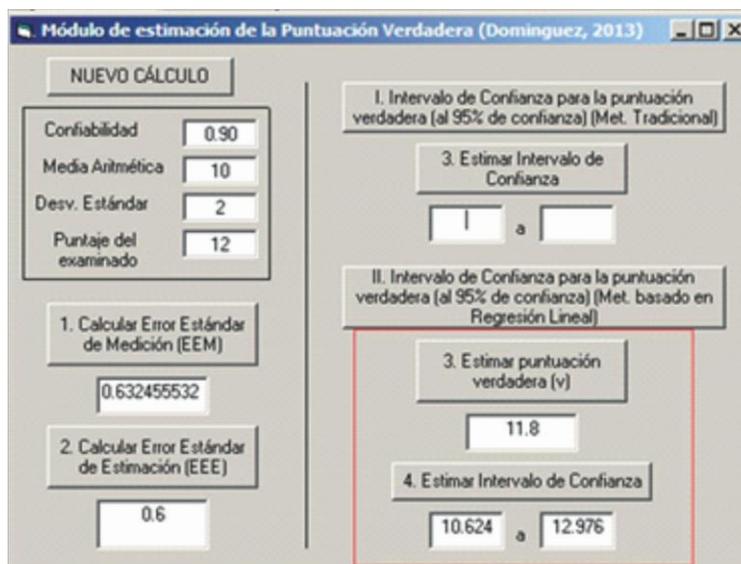


Figura 4. Vista de los resultados en el programa (método basado en la regresión lineal).

El presente módulo puede ser empleado en diversos campos de la evaluación psicológica, ya sea esta en consultorio o en investigación. Es así que se recomienda citar el EEM en los estudios de

validación de instrumentos, así como en el análisis e informes de tests psicológicos, para así alertar al usuario de las posibles variaciones de las puntuaciones propias de las evaluaciones. Por

otro lado, el reporte de los intervalos de confianza resulta necesario para que el usuario tenga presente esos límites al momento de emitir algún reporte o dictamen.

Limitaciones

Una de las limitaciones principales de la herramienta es que el cálculo de la confiabilidad (por el método de consistencia interna) debe hacerse por otros medios, sea este por medio de cálculo manual, o por intermedio de otros paquetes estadísticos.

REFERENCIAS

Aiken, L. (2003). *Tests psicológicos y evaluación* (11ma Edic.). México D.F.: Pearson Educación.

American Psychological Association (2010). *Manual de Publicaciones de la American Psychological Association* (3era Edic.). México D.F.: Manual Moderno.

Cortada de Kohan, N. (1999). *Teorías Psicométricas y Construcción de Tests*. Buenos Aires: Lugar Editorial.

Gempp, R. 2006. El error estándar de medida y la puntuación verdadera de los tests psicológicos: algunas recomendaciones prácticas. *Terapia Psicológica*, 24(2), 117-130.

Morales-Vallejo, P. (2007). *La fiabilidad de los tests y escalas*. Recuperado desde www.upcomillas.es/personal/peter/estadisticabasica/Fiabilidad.pdf, el 13.03.13, 8:00 p.m.

Muñiz, J. (2001). *Teoría Clásica de los Tests*. Madrid: Pirámide.

Recibido: 29 de agosto del 2013

Aceptado: 30 de octubre del 2013

