



reseñas educativas // education review

editores: gustavo e. fischman gene v. glass melissa cast-brede

revista de reseñas de libros, de acceso libre y multi-lingüe

7 de marzo de 2011

ISSN 1094-5296

Fernández-Cano, Antonio y Fernández-Guerrero, Inés M^a (2009). *Crítica y alternativas a la significación estadística en el contraste de hipótesis*. Colección Cuadernos de Estadística (# 37). Madrid: Editorial La Muralla.

195 páginas

ISBN: 978-84-7133-786-3

Reseñado por Teresa Lara-Moreno

Universidad de Granada. Granada.

Racionalidad de la obra

Esta monografía, *Crítica y alternativas a la significación estadística en el contraste de hipótesis*, presenta uno de los tópicos más relevantes de la metodología estadística actual con fuertes implicaciones en la evaluación de las ciencias sociales y biomédicas. Se considera una cuestión bastante controvertida cual es el uso de tests de significación estadística para el contraste de hipótesis.

Muy pocos tópicos metodológicos habrán generado tanta controversia como el del uso e interpretación de los tests de SE. Si algún tema metodológico ha sido considerado, indagado y cuestionado profundamente durante los últimos cincuenta años, éste ha sido el uso de la significación estadística para el análisis de datos de investigaciones en las diversas



Citación: Lara-Moreno, T. Reseña de Fernández-Cano, A. y Fernández-Guerrero, I. M^a (2009). *Crítica y alternativas a la significación estadística en el contraste de hipótesis*. Colección Cuadernos de Estadística (# 37). Madrid: Editorial La Muralla. *Reseñas Educativas*, 14 Recuperado [fecha] de

<http://www.edrev.info/reviews/revs274.pdf>

Esta obra está disponible en edición digital en: <http://digital.casalini.it/9788471339195>

ciencias sociales (psicológicas, económicas, políticas, sociológicas), en ciencias de la educación, e incluso biomédicas (sobre todo en medicina social y epidemiología). En el contexto pedagógico español, la cuestión se considera por primera vez en el manual de Fernández-Cano (1995, pp. 189-205) y posteriormente éste autor publica un artículo en la revista *Bordón* Fernández-Cano (2001). El presente estudio continúa, ahonda y sistematiza aquellas propuestas, pretendiendo suscitar abundante discusión útil sobre este tema, que ha causado más confusión y controversia que cualquier otro aspecto del análisis de datos contemporáneo

La sana práctica científica de reevaluar periódicamente los métodos disponibles para generar y certificar el conocimiento, ha determinado que la cuestión de la significación estadística aparezca de manera virulenta y con carácter cíclico. Aquí, al par que se critica el uso habitual de esta técnica analítica, se ofrecen técnicas alternativas que la complementan.

Por ello, estamos tan necesitados de monografías metodológicas como ésta que tratan de superar los compendios (i.e. Morrison y Hankel, 1970; Harlow, Mulaik y Steiger, 1997), que han considerado por yuxtaposición aportes diversos al tópico de la SE. Se trata de ofrecer una nueva visión, más compacta e integrada de las críticas (pros y contras) y alternativas analíticas compiladas en Kline (2005). No obstante, el tópico de la significación estadística se percibe como una cuestión plena de controversia y es probable que continúe así en el futuro.

Estructura de la monografía

El libro consta de dos partes. En la primera parte (Capítulos 1 a 5), el capítulo 1 introduce una serie de conceptos clave: hipótesis y sus tipos, teoría, las dos aproximaciones a la Significación Estadística (Fisher *versus* Neyman-Pearson), la dualidad de situaciones en el contraste de la hipótesis nula (H_0) de apoyar *versus* no apoyar su rechazo y diversas acepciones metonímicas de la Significación Estadística. El 2º capítulo se centra en la improcedencia de la Significación Estadística para el contraste de hipótesis ofreciendo una serie de razones: el uso de inferencias lógicas incorrectas en la determinación de la Significación Estadística, la improcedencia como algoritmo de decisión, la supeditación al tamaño muestral, la desconsideración de la idea de potencia del contraste, la irrelevancia como detonador de la magnitud del efecto, que posibilita una generalización limitada, la arbitrariedad de la marca $p = 0.05$ y su relativización según la disciplina en que se aplica, la lateralidad genera ambigüedad en la inferencia, la asociación de su uso continuado a prácticas y supuestos investigadores cuestionables así como la proclividad de la Significación Estadística a múltiples e incorrectas interpretaciones y falsas concepciones.

El capítulo 3 explora las consecuencias del uso masivo de la significación estadística; en concreto que la Significación Estadística no permite una falsación rigurosa de teorías, la desconsideración de la significación sustantiva y la generación del sesgo pro estudios positivos.

En el capítulo 4 se indagan los orígenes de tan ubicua práctica y su persistencia a lo largo del tiempo. Fisher se presenta como el inductor inicial aunque el poderoso diseminador fueron modelos analíticos de origen norteamericano auspiciados por potentes asociaciones profesionales (APA y AERA). El capítulo prosigue exponiendo una serie de enconadas razones de defensa de la Significación Estadística. Principalmente su utilidad tras el uso de técnicas analíticas robustas.

La primera parte termina con una relación de consideraciones éticas en torno a la Significación Estadística. La segunda parte consta de seis capítulos (5 a 10), y en ella se ofrecen algunos métodos alternativos que la complementan y mejoran el contraste de hipótesis haciéndolo más ajustado tal constaste. El capítulo 5 se centra en la magnitud del efecto distinguiendo dos acepciones: tamaños del efecto y estimadores de la magnitud de la asociación mediante porcentaje de varianza explicada. La sección dedicada al tamaño del efecto es extensa: discute su racionalidad, expone procedimientos y técnicas de cálculo, que abarcan las propuestas seminales de Glass (1976) a

las más actualizadas y compiladas por Grissom y Kim (2005), relaciona estadísticos inferenciales y tamaños del efecto y se reconocen las limitaciones del tamaño del efecto.

El capítulo 6 está dedicado a la estimación interval enfatizando el uso de diagramas de caja y bigotes y su interpretación. También se confronta los intervalos de confianza *versus* tests de Significación Estadística y se examinan los problemas afines a los intervalos de confianza y las resistencias potenciales al empleo de los mismos. Termina este capítulo con la presentación de la estimación interval basada en la no-centralidad.

Todo el capítulo 7 está dedicado a la inferencia bayesiana aplicada a análisis de datos en situaciones de un diagnóstico y experimentales. Se completa este bloque con el uso de la inferencia bayesiana en la determinación de un intervalo de confianza.

Una singular novedad de esta monografía es todo el capítulo 8 dedicado a la Significación Clínica como una aproximación analítica complementaria: cuantitativa y cualitativa al par. El bloque se inicia al conceptualizar la idea de significación clínica y en oposición a la significación estadística. Después, se presentan tres procedimientos para determinar la significación clínica: validación social, evaluación clínica e índices afines y evaluación diagnóstica. Se reconocen al término unas críticas a los procedimientos de significación clínica.

El capítulo 9 se centra en la replicación y el metanálisis. Se indica cuándo realizar una replicación y tipos de estas. El cuerpo central de este bloque se dedica a la entronización del metanálisis para la síntesis de estudios replicados. Se expone la relación entre metaanálisis y significación estadística y se ofrece la secuencia metodológica de éste. También se cuestiona el uso del metaanálisis para el contraste de hipótesis. Termina el capítulo apuntando algunas técnicas de replicabilidad: *bootstrap* ante todo.

El epílogo (capítulo 10) se centra en la necesidad de reflexionar sobre los datos, se proponen pautas para la reflexión sobre los datos en el contraste de hipótesis e incluso una cierta logística que facilite reflexionar sobre estos. Se ofrecen un procedimiento general para el contraste de hipótesis y falsación de teorías y que se completa con una serie de comentarios finales.

La obra termina con una extensa relación de referencias, bibliográficas inexorablemente escritas en lengua inglesa y de procedencia de USA. Sin embargo, se echa en falta que los autores hubiesen dedicado algún capítulo a análisis gráfico-visual, análisis exploratorio de datos y enfatizar las técnicas robustas, aunque el número 2 de esa colección ya publicó otra monografía al respecto.

Resumiendo, esta monografía nos ofrece una novedosa visión panorámica, más compacta e integrada de las críticas (pros y contras) y alternativas analíticas a la significación estadística, que aún no disponíamos en lengua española. Se adjuntan abundantes tablas y cuadros que ejemplifican las propuestas sin entrar en las minuciosidades de la teoría estadística, en su lenguaje, desarrollos y fórmulas quintaesenciadas.

Dado que esta monografía es eminentemente metodológica, no puramente estadística y dirigida a investigadores en ciencias sociales, médicas y de la educación, los autores, en palabras suyas, han optado por usar un lenguaje directo, llano, uniformando notaciones, comprensible a una mayor audiencia, sin las florituras de demostraciones matemáticas que lo habrían hecho, si no incomprensible, sí al menos tedioso.

Los autores nos advierten que, ante la diversidad y complejidad de las alternativas disponibles y tropelías antiéticas factibles, podríamos caer en el escepticismo epistemológico y la desconfianza analítica. Esta desconfianza es bien fácil de suscitar, pues para acceder al conocimiento no existe un camino algorítmico, mecánico, “regio”, tal como Pitágoras le dijo al faraón. No hay una técnica estándar para analizar hipótesis.

Tampoco es intención de los autores, según confiesan, que, tras la lectura de este texto, el lector extraiga una visión de la Estadística como un arma pernicioso y, menos aún, una actitud

negativa al estudio de esta disciplina, que le haga desembocar en lo que se ha venido a denominar como “estadistocofobia”. Nada de eso pretenden; antes bien, la Estadística se concibe como un repertorio instrumental y procedimental, no sólo circunscrito al análisis de datos sino también a otros momentos y enfoques de la investigación. Tan delicado instrumental, puesto al servicio del conocimiento y del estudio de la realidad social, médica y educativa, debe ser tratado reflexiva y flexiblemente, y no como procedimiento rígido, mecánico y ritualista.

Esta monografía ha querido alertar de una disfunción, casi en fase de metástasis, que ha ocasionado el mal uso de una de sus múltiples herramientas (la SE). La información estadística y sus técnicas alternativas no son cosa baladí, pueden significar la diferencia entre la vida y la muerte; por ejemplo, cuando un paciente con cáncer tiene que decidir si se somete a un tratamiento médico doloroso basándose en cierta probabilidad de éxito o cuando un jurado decidirá sobre la culpabilidad de un acusado basándose en evidencia del ADN.

Lo que necesitamos enseñar a nuestros alumnos (investigadores noveles), no es tanto SE, sino pensamiento estadístico: como generar hipótesis audaces, derivar predicciones precisas, planificar experimentos que minimicen el error real (y no insertarlo en una razón F), analizar los datos separadamente, antes que agregarlos automáticamente, y reflexionar sobre ellos en profundidad. Los métodos estadísticos deberán guiar y disciplinar nuestro pensamiento pero nunca determinarlo; porque como manifestaba hace ya más de cien años d. Santiago Ramón y Cajal (1898/1999; p. 124):

No basta examinar; hay que contemplar; impregnemos de emoción y simpatía las cosas observadas; hagámoslas nuestras, tanto por el corazón como por la inteligencia. Sólo así nos entregarán su secreto. Porque el entusiasmo acrecienta y afina nuestra capacidad perceptiva.

Referencias

- Fernández Cano, A. (1995). *Métodos para evaluar la investigación en Psicopedagogía*. Madrid: Síntesis.
- Fernández Cano, A. (2001). Una reflexión metodológica sobre la significación estadística en investigación educativa. *Bordón. Revista de Pedagogía*, 53(4), 499-507.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5 (10), 3-8.
- Grissom, R. J. Y Kim, J. J. (2005). *Effects sizes for research: A broad practical approach*. Mahwah, NJ: Lawrence Erlbaum.
- Harlow, L. L.; Mulaik, S. A. Y Steiger, J. H. (Eds.) (1997). *What if there were no significance tests*. Mahwah, NJ: Lawrence Erlbaum.
- Kline, R.B. (2005). *Beyond significance testing. Reforming data analysis methods in behavioral research*. Washington, DC: APA.
- Morrison, D. E. Y Hankel, R. E. (Eds.) (1970). *The significance test controversy: A reader*. Hawthorne, NY: De Gruyter Aldine.
- Ramon y Cajal, S. (1999). *Reglas y consejos sobre investigación científica. Los tónicos de la voluntad* (15ª ed.). Madrid: Espasa-Austral. Original publicado en 1898.

Acerca de los autores: **Antonio Fernández-Cano** es doctor en Filosofía y Ciencias de la Educación. Trabaja actualmente como catedrático de la Universidad de Granada (España) en la Facultad de Ciencias de la Educación. Sus áreas de interés son las metodologías de la investigación educativa con especialización en análisis de datos, evaluación de programas y de la investigación y cienciometría pedagógica. Sus direcciones de contacto son de correo electrónico: afcانو@ugr.es y de página en red: <http://www.ugr.es/local/afcانو>

Inés M^a Fernández-Guerrero es licenciada en Medicina y Cirugía y especializada en Medicina Preventiva y Salud Pública. Trabaja actualmente como médica clínica en el Servicio de Urgencias del Hospital “Virgen de las Nieves” del Servicio Andaluz de Salud en Granada (España). Sus áreas de interés son la dimensión social de la medicina y el impacto de la investigación médica sobre la práctica clínica. Su dirección de correo electrónico es: inesmariafernandez@yahoo.es

Acerca de la autora de la reseña: Teresa Lara-Moreno es doctora en Pedagogía. Trabaja actualmente como profesora asociada en el área de Teoría e Historia de la Educación de la Facultad de Ciencias de la Educación de la Universidad de Granada (España). Sus áreas de interés son el estudio de las instituciones contemporáneas de educación, valores y educación, el análisis de políticas y prácticas educativas y el impacto de la investigación educativa sobre la práctica docente. Su dirección de correo electrónico es: tlara@ugr.es

El copyright es retenido por el/la autor/a quien otorga el derecho de primera publicación a
Reseñas Educativas/Education Review
<http://edrev.info>



Editores

Gustavo E. Fischman

fischman@edrev.info

Gene V Glass

glass@edrev.info

Melissa Cast-Brede

cast-brede@edrev.info