

# ANÁLISIS ESTADÍSTICO Y SU REPORTE EN CIENCIAS DEL COMPORTAMIENTO

Dr. Guillermo Yáber-Oltra

## RESUMEN

Un análisis estadístico y su reporte apropiado, son elementos fundamentales en la investigación científica derivada de trabajos de grado, tesis o proyectos de investigación en Psicología. El enfoque predominante en el uso de la estadística aplicada en Psicología se denomina "prueba de significación de la hipótesis nula". Este enfoque ha sido objeto de severas críticas desde que comenzó su uso y los esfuerzos para "des-institucionalizarlo" no han tenido buenos resultados. Se presentan consideraciones para enriquecer el análisis estadístico que se realiza en la investigación psicológica. El uso del poder para estimar el tamaño de la muestra, el empleo de técnicas de análisis exploratorio de los datos, el reporte tanto de la estimación por intervalos de los parámetros poblacionales como del valor de probabilidad "p" encontrado y el tamaño del efecto, son algunas de las recomendaciones analizadas y sugeridas para mejorar el análisis estadístico y su reporte en la investigación en las ciencias del comportamiento.

Palabras clave: estimación, poder, efecto, hipótesis, análisis, estadística, método.

## INTRODUCCIÓN

La investigación científica en Psicología con enfoque empírico, emplea métodos estadísticos que le sirvan de fundamento para realizar descripciones y formular inferencias acerca de las variables y sus relaciones en la ciencias del comportamiento

y sociales. Aún si el enfoque del investigador o estudiante es “cualitativo”, esto no obsta para que potencialmente se empleen algunos métodos estadísticos. En un texto de introducción a la estadística aplicada en ciencias sociales se señala que “El campo de la estadística es un conjunto de procedimientos para reunir, medir, clasificar, codificar, computar, analizar y resumir información numérica adquirida sistemáticamente” (Ritchey, 2002, p.2).

La estadística se comenzó a utilizar en Psicología casi desde el inicio de esta disciplina, a finales del siglo XIX, y desde ese entonces, se viene utilizando para el análisis de datos en casi todas las áreas de la Psicología.

Los métodos estadísticos empleados en Psicología frecuentemente se agrupan bajo la denominación de estadística descriptiva e inferencial. El enfoque vigente para el uso de la estadística inferencial en ciencias del comportamiento se denomina “prueba de significación de la hipótesis nula” (NHST, por sus siglas en inglés), que se deriva del híbrido de dos perspectivas teóricas, la de Fisher y la de Neyman-Pearson (Kline, 2004, p.6).

Esta tradición de uso, ha sido objeto de severas críticas (por ejemplo Cohen, 1994) y eventualmente llevó a la Asociación de Psicólogos Americana (APA), ha formalizar un grupo de tarea para analizarlas. Los resultados de su trabajo fueron publicados (Wilkinson & the Task Force on Statistical Inference, 1999) y algunas de sus recomendaciones fueron acogidas por la más reciente versión del manual de publicaciones de la APA (APA, 2001).

A pesar de ello, algunos aspectos de estos lineamientos no han sido suficientemente “socializados” en la comunidad científica de estudiantes, profesionales e investigadores en Psicología. Por ello, nuestro propósito es revisar estas recomendaciones y proporcionar sugerencias para mejorar los reportes de investigaciones, trabajos de grado y tesis en Psicología y otras ciencias del comportamiento. Para ello, revisaremos los elementos que se incluyen en un plan de investigación.

La estructura de una investigación en ciencias del comportamiento es el plan general que se va a llevar a cabo para responder a la pregunta que hemos formulado. El plan incluye la especificación del tipo de investigación, el diseño, la especificación de la población y la muestra, el consentimiento informado de los participantes, el poder deseado y el procedimiento para recolectar y analizar la información. A continuación se considerarán cada uno de estos aspectos en relación con el análisis estadístico y su reporte.

## **TIPO DE ESTUDIO Y DISEÑO**

El tipo de investigación científica y su diseño, demarcan el ámbito de aplicación de las técnicas estadísticas. En los artículos científicos enviados a consideración de publicación, no se estila explicitar el tipo de estudio y su diseño. Sin embargo, esto sí es una práctica recomendable en los trabajos de licenciatura y de postgrado. En los trabajos de grado, es conveniente especificar el tipo de estudio y su diseño atendiendo a la manipulación o no de variables, el contexto donde se realiza el estudio y el número de veces en que se toman las medidas de las variables. Es de notar que el término “diseño” es aplicable mayormente a los estudios experimentales y debe en lo posible evitarse su uso en los estudios observacionales donde este término no parece apropiado.

En lo que concierne a la manipulación de las variables, una primera recomendación consiste en indicar si el estudio es experimental u observacional. Experimental si el plan contempla la manipulación de una o más variables independientes y observacional si esta manipulación no ocurre. Otros autores como Kerlinger y Lee (2002), sugieren la distinción entre experimental y no experimental. Sin embargo, la expresión no experimental describe vagamente el tipo de estudio y por ello se sugiere evitar su empleo para describirlo.

Es posible plantear estudios multivariados donde una variable es de tratamiento y otra es de clasificación. En estos casos es recomendable indicar la naturaleza híbrida del proyecto de investigación. En los estudios experimentales, cuando no se cumple con la triple condición de selección al azar de los sujetos de la población, su asignación al azar a las condiciones de tratamiento y asignación al azar de los tratamientos a las condiciones, se suele denominar a la investigación, cuasi-experimental. En los estudios experimentales las variables reciben la denominación de independientes (las que se manipulan) y dependientes (en las que observamos los efectos).

En los estudios observacionales, el segundo tipo de investigación, las variantes más frecuentes son los estudios descriptivos, de caso o campo, correlacionales y causal-comparativo. En los estudios descriptivos medimos una o más variables en una población o una muestra y analizamos los resultados. En los estudios de caso y campo medimos múltiples variables para pocos sujetos o en una unidad social. En los estudios correlacionales describimos y analizamos la covariación entre dos o más variables. Finalmente en los estudios de naturaleza causal-comparativa, exploramos posibles relaciones funcionales “causales” entre variables. Menos frecuentes aunque igualmente importantes son los estudios longitudinales o evolutivos, donde se describen variaciones de fenómenos con el transcurrir del tiempo o históricos donde se someten a prueba hipótesis indagando hechos ocurridos en un lapso de tiempo.

En los estudios observacionales no se recomienda emplear las etiquetas de variables independientes y dependientes. En su lugar es preferible indicar una expresión genérica (variable 1, variable 2..., variable n) o indicar que se trata de variables predictoras y criterio (en el caso de diseño que involucren el uso de técnicas de regresión) o variables exógenas y endógenas para el caso de estudios que emplean el análisis de ruta. Otras denominaciones existen para algunos estudios específicos como puede ocurrir en el caso del modelaje de ecuaciones estructurales (SEM, por sus siglas en inglés).

El contexto donde se realiza el estudio ayuda a describir el proyecto. El estudio será de campo o de laboratorio de acuerdo al ambiente donde éste se lleve a cabo.

Finalmente, se sugiere indicar si el estudio es de corte transversal o longitudinal atendiendo al número de veces en que se toman las medidas de las variables: estudios transversales cuando se toma una vez la medida y longitudinales cuando estas mediciones se realizan en varias oportunidades. Esta distinción transversal-longitudinal no aplica en algunos diseños experimentales por lo que se debe clarificar bien el tipo de estudio antes de emplear estos descriptores para estos diseños.

Otra forma de clasificar los tipos de estudio es simplemente indicando si el estudio es descriptivo, correlacional o explicativo, de acuerdo a si el trabajo intenta describir variables, relacionarlas o estudiar sus posibles relaciones causales. El tipo explicativo en esta clasificación incluye tanto estudios del tipo causal-comparativo como experimental.

En última instancia, cuando se trate de etiquetar un tipo de investigación, es preferible citar la fuente de donde se tomó el ejemplo, para que el lector pueda verificar la taxonomía empleada por los investigadores.

Las ayudas gráficas para ilustrar el tipo de investigación y el enunciado de la hipótesis estadística, suelen incluirse en los trabajos de grado, aunque se omiten generalmente en los reportes enviados a consideración de publicación de revistas científicas. En estas últimas se estila declarar el objetivo u objetivos del estudio en la sección de introducción del reporte.

La correcta especificación del tipo de investigación, contribuye a clarificar las técnicas estadísticas que se utilizarán para el análisis de los resultados obtenidos. El diseño experimental y las variantes de estudios observacionales limitan los análisis estadísticos que se pueden realizar por lo que éste es un aspecto crítico a tomar en cuenta al realizar investigaciones en ciencias del comportamiento.



El plan general de investigación incluye apartados de población y muestra, materiales e instrumentos y procedimiento. Estos aspectos imponen condiciones a los métodos estadísticos que podemos emplear en los proyectos de investigación. A continuación revisaremos cada uno de ellos.

## **POBLACIÓN Y MUESTRA**

Un criterio fundamental para el uso de métodos estadísticos es la definición de la población bajo estudio, la muestra que se va a recolectar y la unidad de análisis. Señalar el tipo de muestreo (aleatorio o no aleatorio) en el reporte, ayuda en la comprensión de los análisis estadísticos que se mostrarán en la sección de resultados y en la discusión. En este apartado, los autores también deberían reportar los procedimientos empleados para obtener el consentimiento informado de los humanos participantes en el estudio y del tratamiento apropiado de los animales no humanos cuando sea el caso. Para ello están disponibles los lineamientos del código de ética de la disciplina (APA, 2002), los lineamientos éticos para la conducta ética en el cuidado y uso de animales no humanos (APA, 1996) y aquéllos que puedan existir en la escuela, facultad o institución donde se realiza el estudio (por ejemplo, Escuela de Psicología UCAB, 2002).

En esta sección del reporte de investigación, se recomienda incluir información sobre las decisiones que llevaron a determinar el tamaño de la muestra. Para determinar el tamaño de la muestra es necesario especificar el tamaño del efecto (ES), el valor del nivel de significación alfa ( $\alpha$ ) y el poder ( $\beta$ ). En términos generales se acepta un nivel alfa de .05 y un poder del 80%. Es necesario especificar el tamaño del efecto que se espera obtener y a partir de estos tres datos, calcular el tamaño de la muestra. Este cálculo suele ser muy engorroso, sin embargo ya existen programas de cómputos (Borenstein, Rothstein y Cohen, 2001; Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A., 2007.) y módulos del paquete estadístico para ciencias sociales (SPSS), que permiten el cálculo del poder antes de iniciar el estudio.

Suponga que desea realizar un estudio de regresión múltiple con tres predictores, un tamaño del efecto (ES) medio .15, un alfa de .05 y un poder de .95 (siguiendo a Cohen, 1992), y uno desea averiguar el tamaño de la muestra. Para calcularlo se puede emplear por ejemplo la distribución gratuita del programa G\*Power 3.0 (Institut für experimentelle psychologie, 2007). Este programa contempla el análisis del poder a priori y a posteriori. Al aplicar el análisis a priori se observa que el resultado es N=119 sujetos. Este programa también proporciona los recursos para el análisis del poder luego de finalizado el estudio.

## **MATERIALES E INSTRUMENTOS**

Las técnicas de recolección de la información en ciencias del comportamiento incluyen la observación directa y la autoobservación, la encuesta (entrevistas y cuestionarios), tests y escalas. En todos estos casos, es necesario explicitar las variables que se miden con las diferentes técnicas y proporcionar la información de sus propiedades psicométricas (especialmente confiabilidad y validez). En el caso de instrumentación de laboratorio, se sugiere describir la instrumentación de tal modo que permita a otros investigadores replicar el proceso de medición. Los análisis estadísticos que son posibles de realizar dependen del nivel de medida de las variables por lo que su especificación es un aspecto crítico en los proyectos y reportes de una investigación.

## **PROCEDIMIENTO**

El procedimiento incluye un protocolo detallado de cómo se van a recolectar y analizar los datos. Es crítico describir con la mayor precisión posible las actividades que se realizarán para la recolección y el análisis pues esto permite garantizar futuras repeticiones del estudio. Se recomienda disponer condiciones para eliminar los sesgos de los experimentadores y anticipar las causas de posibles pérdidas de participantes en el estudio. En este apartado también puede incorporarse una mención a las técnicas de análisis de la información que se presentarán en los resultados (psicométrica, descriptiva, de supuestos e inferencial). En este apartado también puede indicarse el nivel de significación ( $\alpha$ ), el poder ( $\beta$ ) y el nivel aspirado como tamaño del efecto (ES).

## **RESULTADOS**

La sección de resultados de un reporte de investigación debe incluir el análisis psicométrico de los instrumentos empleados en el estudio cuando sea pertinente, un análisis descriptivo empleando preferiblemente análisis de datos exploratorios (Tukey, 1977) y el análisis inferencial. Es conveniente reportar los resultados de la estimación puntual y por intervalos de los parámetros poblacionales bajo estudio. Seguidamente es pertinente una verificación de los supuestos que son necesarios antes de aplicar pruebas de estadística inferencial así como realizar las transformaciones de los datos cuando esto sea necesario. Seguidamente se presentarán los resultados de la prueba de hipótesis que incluye el valor de  $p$  obtenido, el tamaño del efecto que se haya calculado y el poder resultante.

## **ANÁLISIS PSICOMÉTRICO**

Un aspecto fundamental de las medidas empleadas en las investigaciones en Psicología es que éstas sean confiables y válidas. Cuando se emplean instrumentos de recolección de información como escalas, cuestionarios o tests es importante reportar sus propiedades psicométricas. En el caso que se calculen coeficientes de consistencia interna en el estudio, los resultados deben presentarse y analizarse. Uno de los coeficientes más empleados es el coeficiente alfa de Cronbach, que de acuerdo a su autor, "( $\alpha$ ) es una fórmula general del cual un caso especial es el coeficiente de equivalencia de Kuder-Richardson" (Cronbach, 1951, p.297). Una práctica generalmente aceptada es la que señala que el límite inferior para aceptar este coeficiente es de .70 aunque podría ser de .60 en la investigación exploratoria (Hair, Anderson, Tatham y Black, 1999, p. 105).

Además de reportar el valor del coeficiente, es recomendable reportar el intervalo de confianza donde posiblemente se encuentra el parámetro poblacional. El cálculo a mano del intervalo de confianza para este coeficiente ha resultado hasta ahora engorroso, y no existe acuerdo unánime sobre la mejor forma de estimarlos, sin embargo ya se encuentra disponible un módulo (Ledezma,s.f.) para un programa gratuito de estadística (Young, 2002), que calcula el valor del intervalo de confianza para este coeficiente. Por otra parte, también se ha reportado códigos de cálculos para el coeficiente alfa para aplicarse en programas estadísticos como SAS o SPSS (Duhachek y Iacobucci, 2004). Reportar el intervalo de confianza para el coeficiente alfa facilita los estudios meta-analíticos posteriores y proporciona una estimación más precisa que su estimación puntual.

## **ANÁLISIS DESCRIPTIVO**

El análisis descriptivo de los datos recolectados es de gran ayuda antes de aplicar métodos de estadística inferencial. Por una parte se deben identificar, analizar y tomar decisiones sobre los datos omitidos (Switzer y Roth, 2002) y por otra parte analizar los puntajes extremos y casos influyentes. El tratamiento de los puntajes extremos (Roth y Switzer, 2002) es en la actualidad una tarea que se facilita con las técnicas del análisis exploratorio de datos (Tukey, 1977).

Los gráficos de caja y bigote "Box-plot" y la representación de tallo y hoja "Steam and leaf" (Riethey, 2002, cap. 3) construidos a mano o mediante el uso de programas estadísticos como el SPSS, proporcionan información visual acerca de la forma de la distribución de los datos, así como de los puntajes extremos o "outliers". Está más allá del alcance de este documento analizar el tratamiento de los puntajes extremos.

Es suficiente indicar que es necesario identificarlos con el gráfico de caja y bigote o el recurso de tallo y hoja, asegurarse que no son el resultado de un error de transcripción y emplear criterios aceptados para su aceptación o eliminación de la base de datos antes de proceder a los análisis de inferencia estadística.

Las formas convencionales de mostrar la información para la descripción de los datos como las tablas que indican las medidas de tendencia central y de variabilidad de los datos constituyen, hoy como ayer, recursos valiosos para que los lectores puedan apreciar de una forma más completa, los resultados de un estudio científico en ciencias del comportamiento.

## **VERIFICACIÓN DE SUPUESTOS**

La verificación de los supuestos que permite el uso de las pruebas estadísticas (por ejemplo normalidad), suele omitirse de los artículos de investigación, pero esto no invalida el hecho de que es necesario realizarlos y reportarlos para el caso de los trabajos de grado y tesis. La verificación de los supuestos y el uso de transformaciones en los datos antes de su análisis, es una práctica sana en el momento de realizar el análisis de los resultados de una investigación.

## **ANÁLISIS INFERENCIAL**

El propósito fundamental en toda investigación en Psicología basada en datos cuantitativos y con enfoque empírico es responder a las siguientes preguntas (Kirk, 2001): (a) ¿El efecto que observamos o las relaciones que encontramos pueden atribuirse a errores de muestreo?. Si el resultado no puede atribuirse a errores de muestreo, ¿que tan grande es el efecto?. Y la tercera pregunta ¿El efecto tendrá significación práctica?. La respuesta a la primera pregunta se obtiene al verificar el valor de probabilidad  $p$  obtenido al calcular la prueba de significación estadística. En Psicología se suele asumir a priori una valor  $p$  de .05. Si el valor  $p$  calculado es menor al establecido a priori, se afirma que los resultados no pueden atribuirse a errores en el muestreo. Este único resultado es insuficiente para responder a las otras dos preguntas que además parecen de mayor importancia. Para responder a la segunda, y a la tercera se reporta y analiza el resultado del cálculo del tamaño del efecto para el cual existen dos familias de índices: índices de diferencias de medias normalizadas y medidas de asociación (Kline, 2004, p.97) y se calcula la magnitud del poder resultante.

Tomemos el caso de un trabajo de grado (Brito y Ciccoti, 2007). En el estudio, se realizó un análisis de regresión múltiple con una medida de estrés como criterio y



satisfacción, nivel jerárquico, sexo, antigüedad y edad como predictores en trabajadores del servicio privado de salud en Caracas-Venezuela. El valor “p” asociado al valor de “F” resultó menor que .05, con un coeficiente de determinación ajustado de .247 y dos predictores (Satisfacción y antigüedad) para cuyas pruebas “t” el valor de p también resultó menor a un alfa de .05. El cálculo del tamaño del efecto (ES) a través de  $f^2$  (Cohen, 1992, p. 157) se calcula con la fórmula:  $f^2 = R^2/1-R^2$ . El tamaño del efecto encontrado es de  $f^2=.377$ , un tamaño de efecto grande (Cohen, 1992, p.157). Al realizar el análisis a posteriori del poder, empleando el programa de cálculo G\*power, el valor de éste es  $\beta=.99$ . Tenemos entonces un efecto (ES) grande, para valores de  $\alpha$  y  $\beta$ , aceptables en ciencias del comportamiento. Dos resultados resaltan, el efecto del estrés aumenta con el tiempo en la organización y covaría inversamente con la satisfacción. Con estos resultados la organización puede instrumentar una intervención para aumentar la insatisfacción y prevenir los efectos del estrés que pueden llevar a los empleados al agotamiento (burnout).

En términos de reporte, las tablas de correlación entre predictores y criterio, la tabla de los coeficientes de regresión y la información relativa al tamaño del efecto y poder resultante constituyen informaciones que rutinariamente deberían incluirse en los reportes de investigación.

## **CONCLUSIONES**

El análisis estadístico y su reporte es un ingrediente indispensable en todo trabajo de grado o de investigación con enfoque empírico en Psicología y otras ciencias del comportamiento que utilicen estos métodos para el análisis de los resultados. Con el tiempo se han incorporado un conjunto de recursos computacionales, muchos de ellos gratuitos, que permiten realizar análisis que en el pasado eran tediosos, engorrosos y con altas posibilidades de error. Hoy en día disponemos de programas para analizar el poder de una prueba, construir intervalos de confianza para parámetros poblacionales, gráficos que permite analizar puntajes extremos y casos influyentes entre otras posibilidades descritas en este documento. Los cursos de enseñanza de la estadística para ciencias del comportamiento y sociales y cursos de metodología de la investigación pueden ahora explotar al máximo estas facilidades para ayudar a los futuros investigadores en su esfuerzo por describir, controlar, predecir y explicar el comportamiento más allá del enfoque de “prueba de significación de la hipótesis nula” (Kline, 2004).

## REFERENCIAS

- APA (1996). Guidelines for ethical conduct in the care and use of animals. Washington: Autor.
- APA (2002/2001). Manual del estilo de publicaciones de la American Psychological Association. (2ª ed.). México: El Manual Moderno.
- APA (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, 57, 1060-1073.
- Borenstein, M., Rothstein, H. y Cohen, J., (2001). Power and precision. [programa informático]. Disponible en línea en URL: <http://power-analysis.com>.
- Brito, F. y Ciccotti, A. (2007). Efectos de la satisfacción laboral y nivel jerárquico sobre el estrés laboral en una empresa del sector salud de caracas. Trabajo de grado de licenciatura en relaciones industriales. Caracas: UCAB.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 1, 155-159.
- Cohen, J. (1994). The earth is round ( $p < .05$ ). *American Psychologist*, 49, 997-1003.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Duhachek, A. y Iacobucci, C. (2004). Alpha's standard error (ASE): an accurate and precise confidence interval estimate. *Journal of Applied Psychology*, 5, 792-808.
- Escuela de Psicología UCAB (2002). Contribuciones a la deontología de la investigación en Psicología . Caracas: UCAB
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Hair, j., Anderson, R., Tatham, R. y Black, W. (1999). Análisis multivariante. Madrid: Pearson-Prentice Hall.

- Institut für experimentelle psychologie (2007). G\*Power 3 [programa informático]. Disponible en línea en URL: <http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower>.
- Kerlinger, F. y Lee, H. (2002). Investigación del comportamiento: métodos de investigación en ciencias sociales. (4ª ed.). México: Mc Graw-Hill.
- Kirk, R. (2001). Promoting good statistical practices: some suggestions. *Educational and Psychological Measurement*, 61, 2, 213-218.
- Kline, (2004). *Beyond significance testing*. Washington: APA.
- Ledesma, R. (s.f.). Cálculo de intervalos de confianza para el coeficiente Alfa de Cronbach en ViSta. Recuperado el 15 de Octubre de 2007 desde el URL:  
<http://www.mdp.edu.ar/psicologia/vista/Using-ALPHA-CI.PDF>
- Ritchey, F. (2002). *Estadística para las ciencias sociales*. México: Mc-Graw Hill.
- Roth, P. y Switzer, F. (2002). Outliers and influential cases: Handling those discordant contaminated maverick rogues. En Steven Rogelber (ed.). *Handbook of research methods in industrial and organizational psychology*. Victoria-Australia: Blackwell.
- Switzer, F. y Roth, P. (2002). Coping with missing data. En Steven Rogelber (ed.). *Handbook of research methods in industrial and organizational psychology*. Victoria-Australia: Blackwell.
- Tukey, J. W. (1977). *Exploratory data analysis*. New York: Addison-Wesley.
- Young, F. (2002). ViSta, the Visual Statistic System [programa informático]. Disponible en línea en URL: <http://forrest.psych.unc.edu/research/index.html>.
- Wilkinson, L. & the Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604.