

### El análisis de datos en la investigación en psiquiatría y salud mental

**Pedro Monterrey<sup>1</sup>**  
**Carlos Gómez-Restrepo<sup>2</sup>**

#### **Resumen**

El análisis estadístico de los datos es parte fundamental para una investigación. Está constituido por diferentes componentes, que van desde la limpieza y validación de los datos hasta la realización de inferencias en respuesta al problema de investigación. Este artículo pretende mostrar de manera sistemática cada uno de estos componentes, resaltando sus elementos constitutivos.

**Palabras clave:** estadística, análisis de los datos, investigación en psiquiatría, metodología de investigación

**Title:** Data Analysis in Psychiatry and Mental Health Research

#### **Abstract**

Statistical analysis of data constitutes a crucial part in research. It has different components ranking from cleaning and validating data to the realization of data inferences in response to specific research problems. This article presents systematically the constitutive elements of each of these components.

**Key words:** Statistics, data analyses, psychiatric research, research methodology

---

<sup>1</sup> Matemático, PhD en matemáticas. Profesor asociado del Departamento de Epidemiología Clínica y Bioestadística, Pontificia Universidad Javeriana. Bogotá, Colombia.

<sup>2</sup> Médico psiquiatra. MSc en epidemiología clínica. Psicoanalista. Coordinador de la Especialidad en Psiquiatría de Enlace. Profesor asociado del Departamento de Epidemiología Clínica y Bioestadística y del Departamento de Psiquiatría y Salud Mental. Pontificia Universidad Javeriana, Bogotá, Colombia.

## Introducción

“En 1903 H.G. Welles formuló la hipótesis de que el pensamiento estadístico algún día sería tan necesario, para ser buen ciudadano, como la habilidad de leer y escribir” (1). La afirmación ha llegado a ser cierta, y lamentablemente el lenguaje estadístico domina en el campo de las investigaciones, llegando a desvirtuar el proceso de investigación al que debió potenciar (2,3). Muchos investigadores pierden de vista el problema que pretenden abordar, y no logran formularlo bien porque desde el momento de su identificación comienzan a pensar en términos estadísticos, términos que en ocasiones son superficialmente comprendidos, y que se vuelven equívocamente lo primordial al estructurar la pregunta de investigación. Lo que debía ser consecuencia de la identificación del problema y pregunta de investigación se vuelve parte del mismo.

El mal uso y abuso de la estadística es complejo, y se presenta tanto en el momento de identificar el problema de investigación como al analizar los datos e interpretarlos. La cuestión ha llegado al punto de que en muchas revistas científicas de prestigio se desestimula el uso de las pruebas de hipótesis en los artículos que publican, lo que es una forma de prohibición (4). Las críticas al dominio de la estadística en la investigación, y el rechazo que se desarrolla contra ella pueden resumirse en la afirmación de Rozeboom

“el folklorismo estadístico, herencia de un pasado primitivo, continúa dominando la escena local” (5).

Buena parte de esa situación se deriva de los mecanismos de enseñanza de la estadística, es decir, es responsabilidad de los propios estadísticos, aunque en apariencia recae en los investigadores. Los avances en computación y el acceso a programas de análisis de datos han puesto a disposición del investigador técnicas de análisis, en ocasiones complejas desde el punto de vista matemático, que son aplicadas sin mayor crítica. Así, en muchos casos el investigador se deja llevar por los sistemas de cómputo, y su labor se reduce a seleccionar una técnica estadística “adecuada para su problema”, dentro de un recetario de opciones que provee un mecanismo de entrada de datos a una caja negra, de donde se extraen “conclusiones” a partir de algún esquema de análisis, por lo general impersonal y basado en algún algoritmo mental que conduce a una decisión.

En consecuencia, el análisis estadístico se ha convertido en una especie de ritual (6,7) en el que la capacidad de análisis del investigador se suplanta por el uso de algoritmos de decisión, derivados de procedimientos estadísticos, que hacen perder el contacto con los datos y más gravemente aún, que los sustituyen a los ojos de los investigadores. Esta es posiblemente una de las causas para sugerir el uso de intervalos de confianza, en

vez de las pruebas de hipótesis (8), y posiblemente uno de los errores que han conducido al rechazo de la estadística en general y, en especial de las pruebas de hipótesis.

Independientemente de las críticas, el beneficio y en muchas ocasiones la necesidad de la estadística en la investigación es indiscutible. La necesidad de criterios para el análisis de los datos que sean objetivos, y que permitan estructurar adecuadamente las conclusiones de la investigación es un hecho. El problema está en que el investigador debe ver la estadística como un complemento de su disciplina y, utilizar sus métodos y procedimientos adecuados a su problema de investigación.

El objetivo del presente artículo es presentar secuencialmente cómo las diferentes componentes de la estadística pueden ayudar al investigador en el diseño del estudio y el análisis de sus datos, mostrando de manera especial cómo obtener y fundamentar sus conclusiones de investigación, desmitificando el exceso de fe en los intervalos de confianza como alternativa a las pruebas de hipótesis, y señalando cómo deben ser aplicadas para disminuir la posibilidad de errores, haciendo énfasis en una mirada integral a los datos como única opción de análisis.

### **La estadística en la investigación**

Lo primero a determinar es el problema de investigación,

siendo éste “una brecha entre una realidad observada y una realidad esperada o deseable” (9), cuya estructuración metodológica no tiene nada que ver con la estadística; en correspondencia, tampoco tiene sentido pensar en términos estadísticos al determinar la pregunta de investigación, los objetivos y las hipótesis. Estos componentes sólo se derivan del problema, siendo necesario ubicarse en el marco que brinda el área del conocimiento en que el problema se inscribe y, seguir las reglas derivadas de la metodología de la investigación.

Una vez estructurado el problema de investigación se piensa cómo utilizar la estadística para abordarlo; concretamente, el primer momento debe ser determinar qué tipo de estudio resuelve la pregunta que se tiene, para posteriormente evaluar el muestreo y cuántos individuos se requieren para responder la pregunta (poder del estudio). Esta decisión no es inmediata, y es consecuencia de la obtención y estructuración de la siguiente información:

- Determinación de las variables (mediciones) involucradas en el problema.
- Identificación del tipo de medidas estadísticas que representan los efectos que se deseen estudiar.
- Establecer los rangos de normalidad y los criterios de relevancia clínica de las mediciones.

### *Determinación de las variables involucradas en el problema*

El logro de los objetivos demanda la realización de un conjunto de mediciones o determinaciones. Ellas serán las variables a analizar en el proyecto; tienen diferentes funciones en la investigación, y en términos generales se corresponden con las denominadas variables dependientes, independientes y de confusión (9).

Un ejemplo sería la determinación de factores de riesgo para la presencia o no de trastorno distímico. Para ello se podría iniciar con un grupo de pacientes con distimia (casos) y otros sin este trastorno (controles), e indagar si variables como género, edad, tipo de crianza, antecedentes familiares de depresión, etc., son factores de riesgo o protección para esta entidad. Otro problema sería la determinación de dosis óptima de un antipsicótico atípico para pacientes con esquizofrenia crónica.

### *Identificación del tipo de medidas estadísticas que representan los efectos que se desee estudiar*

Derivados del problema existen parámetros o valores poblacionales que representan las relaciones del problema a estudiar; su estimación y análisis constituyen interés central de la investigación que se diseña. Para la determinación de esos parámetros es importante traducir el problema de investigación al lenguaje estadístico. Por ejemplo, en la Encuesta Nacional

de Salud Mental se podría pretender conocer el promedio de edad para el inicio de trastornos fóbicos; en este caso el promedio poblacional sería el parámetro de interés. Un estudio de casos y controles pretendería establecer la estimación del valor poblacional del OR (*odds ratio* o razón de disparidad) como medida de asociación entre desarrollo de esquizofrenia en la adolescencia e influencia de la madre durante el embarazo.

### *Establecimiento de los rangos de normalidad y los criterios de relevancia clínica de las mediciones*

Una vez determinadas las variables se debe buscar la siguiente información:

- Unidades de medición (por ejemplo, puntajes de una escala, peso en kilogramos, etc.).
- Rango de valores que caracterizan su comportamiento y la normalidad (por ejemplo, puntaje por encima del cual se identifica depresión en la escala de Hamilton o de Zung, etc.).
- Rango de valores que pudieran caracterizar la relevancia y/o irrelevancia clínica de los resultados.

Respecto al punto segundo (rango de valores que caracterizan su comportamiento y la normalidad) podríamos evaluar una nueva molécula antidepressiva (medicamento P), cuyo costo es dos veces más que el mejor medicamento usado (medicamento X), y tiene los mismos efectos secun-

darios y de seguridad que éste. ¿Cuál debería ser la mínima diferencia clínicamente relevante para evaluar la eficacia del nuevo antidepresivo, si conocemos que el antidepresivo X utilizado (medicamento control) recupera al 70% de los pacientes que lo toman? Una diferencia por debajo del 5% podría ser irrelevante clínicamente, porque el beneficio no superaría los costos.

El problema de investigación determina el tipo de estudio y éste, la estrategia de muestreo; de aquí se obtiene el tamaño de muestra o la cantidad de individuos a evaluar. Para esta determinación se tiene en cuenta: la diferencia que se pretende encontrar, el tipo de estudio y el procesamiento estadístico que se va a realizar, es decir, no hay un tamaño de muestra único; cada tipo de análisis estadístico demanda su tamaño de muestra adecuado.

Ejemplos de ello serían para el estudio de caso y controles sobre distimia y factores de riesgo: buscar OR mayores de 1,5; o en un ensayo clínico de antidepresivos buscar diferencias mayores de 5 puntos en la escala de Hamilton.

Son múltiples los criterios que se pueden emplear para determinar el tamaño de la muestra; pudiera haber tantos como criterios de análisis, pero hay dos de uso bastante generalizado: construir un intervalo de confianza, o hacer una prueba de hipótesis con ciertos patrones de calidad (4,10).

Los intervalos de confianza se construyen mediante un procedi-

miento que con una probabilidad dada asegura disponer de una estimación adecuada. Un intervalo de confianza con una confiabilidad del 95% es un rango de valores obtenidos a partir de la aplicación de un procedimiento de cálculo, en el que en el 95% de los casos se pueden obtener rangos de valores entre los que se encuentra el parámetro.

El concepto es simple, pero su sustento es poco natural para el investigador y por ello, en muchos casos, existen malas interpretaciones de lo que es el intervalo. Lo interesante en los intervalos de confianza es que el diámetro del intervalo (la diferencia entre el mayor y el menor valor) es directamente proporcional a la incertidumbre con la que los datos permiten obtener buenas informaciones sobre el parámetro que se estima.

La última información es consecuencia de que el diámetro es directamente proporcional a la variabilidad en el problema; por ello, al determinar el tamaño de muestra adecuado para utilizar los intervalos de confianza el problema sería: ¿cuál es el tamaño de muestra necesario para disponer de un intervalo de confianza con un diámetro determinado?

Por ejemplo, para estimar el OR entre la presencia de distimia y los antecedentes familiares de depresión mayor se pudiera desear construir un intervalo de confianza del 95%, cuya amplitud no sea mayor que 1,5. Otro ejemplo podría ser la estimación de la prevalencia de agorafobia con un intervalo de

confianza del 95% y una amplitud menor del 5%.

Las pruebas de hipótesis tan cuestionadas en los últimos tiempos tienen un entorno teórico confuso. Existe una profusa literatura explicando sus ventajas y, sobre todo sus problemas; los detalles no serán abordados en este trabajo, pero pueden ser vistos en otros (4). Partiendo de dos hipótesis: la denominada hipótesis nula,  $H_0$ , y su alternativa,  $H_A$ , se identifican los dos errores en el problema; el error de tipo I o rechazar  $H_0$  siendo cierta y el error de tipo II o aceptar  $H_0$  siendo falsa. La cuantía del error de tipo I se identifica con la letra  $\alpha$ . El error de tipo II es más complejo de manejar, pues normalmente es una familia de errores. El tamaño de muestra se calcula estableciendo los dos errores y, un valor a partir del cual es grave cometer el error de tipo II.

Imagine en un estudio de caso y controles para medir si el buen trato (controles) y mal trato (casos) de los hijos se relacionan con el maltrato del padre durante su infancia. En ese caso la hipótesis  $H_0$  sería que la diferencia de proporciones en maltrato del padre durante su infancia en los dos grupos fuera igual a cero, pero la alternativa sería  $H_A$ :  $p_1 - p_2 > 0$ . Para la determinación del tamaño de muestra se prefija el error de tipo I (por ejemplo, 0,05), el error de tipo II (por ejemplo, 0,20), y se determina un valor de la diferencia entre las proporciones que se comparan, a partir del cual ese

error de tipo II tenga consecuencias muy desfavorables, lo que es determinado en general por la diferencia clínicamente relevante.

### **El análisis de los datos**

El análisis de los datos debe permitir determinar los niveles de evidencia que ellos contienen a favor o en contra de la validez de la pregunta de investigación. Este ejercicio de análisis no debe ser hecho siguiendo esquemas rígidos, sino que debe adaptarse a las características de las mediciones. Algunos pasos que pudieran ayudar a organizar un análisis de tipo estándar son:

- Limpieza de los datos.
- Descripción de los datos.
- Identificación de las hipótesis estadísticas y realización de las inferencias.

#### *Limpieza de los datos*

Limpiar los datos, en primera instancia es estar seguros de que no hay inconsistencias en ellos. El primer punto debe ser chequear la digitación; hay que evitar trabajar con datos en los que hay problemas de digitación.

Un segundo momento en la validación es la búsqueda de consistencia entre respuestas a preguntas o determinaciones diferentes. Para ello se deben elaborar criterios de consistencia entre los valores de las variables; se listan los individuos con valores discordantes, y

se procede a analizar si son errores en la toma de la información, en la digitación o si es que los individuos son así, es decir, individuos con atributos poco usuales o extremos. Por ejemplo, un individuo con un peso extremo en un estudio descriptivo de la asociación de sobrepeso con depresión mayor.

Un tercer momento es la identificación de individuos con valores extremos o valores aberrantes, los denominados *outliers*. Esos individuos pudieran representar patrones poblacionales interesantes, y necesitan criterios de análisis especiales; en el caso de que sus datos tengan esta característica debe auxiliarse de un especialista en estadística, para decidir la mejor forma de proceder.

#### *Descripción de los datos*

Esta parte complementa la visión de los datos que comenzó a obtener en el punto anterior; acá debe usar medidas de resumen y de dispersión, así como gráficos que le permitan conocer cómo se comportan sus datos, cuál es el efecto de los individuos extremos, determinar las características de la variabilidad que tuvieron sus datos, identificar cuáles son las causas y cómo se entiende esa variabilidad. Las nociones fundamentales de la estadística descriptiva son útiles en este punto, pero es recomendable utilizar también criterios del análisis de datos exploratorio, por la potencia de esas técnicas para res-

ponder a las preguntas anteriores, y su capacidad de permitir construir hipótesis secundarias a partir de los datos (11).

#### *Identificación de las hipótesis estadísticas y realización de las inferencias*

Como están identificados los parámetros y las medidas de efecto, el primer paso debe ser construir los intervalos de confianza, con ellos se tiene una estimación de cuáles son sus valores y, calculando el diámetro del intervalo se pueden saber los niveles de incertidumbre o de calidad de esas estimaciones.

Aunque algunas publicaciones están en contra de las pruebas de hipótesis estas no deben ser desechadas por su características de objetividad y porque no siempre existen intervalos de confianza adecuados para el análisis (12,13). Por otra parte proponer la alternativa de los intervalos de confianza no es la solución al problema si éstos no se utilizan adecuadamente (14) pues son parte de la teoría de la pruebas de hipótesis (13).

Para hacer que las pruebas de hipótesis sean útiles y también que el análisis con los intervalos de confianza sea relevante es necesario tener en cuenta la relevancia clínica de los resultados, tal y como fue analizado en el punto “rango de valores que pudieran caracterizar la relevancia y/o irrelevancia clínica de los resultados”; por ejemplo, un medicamento nuevo puede intere-

Tabla 1. Criterios para la interpretación de los valores P

Valor P	Criterio de análisis del valor P
$0,1 < P < 1,0$	Débil evidencia contra la hipótesis nula
$0,01 < P < 0,1$	A medida que P disminuye aumenta la evidencia contra la hipótesis nula
$0,001 < P < 0,01$	
$0,0001 < P < 0,001$	A medida que P disminuye hay fuerte evidencia contra la hipótesis nula

sar solamente si es mejor que otro. Por tanto las hipótesis sería:  $H_0: p_1 - p_2 \leq 0$  y la hipótesis alternativa  $H_0: p_1 - p_2 > 0$ . Sin embargo, si por razones de costo y eficiencia el nuevo producto sólo es recomendable si el porcentaje de éxitos aumenta en un 15%, entonces las hipótesis deberían ser  $H_0: p_1 - p_2 \leq 0,15$  y la hipótesis alternativa  $H_0: p_1 - p_2 > 0,15$ .

Es recomendable calcular el valor P, que se debe publicar explícitamente; para su interpretación puede fijar un umbral con el fin de identificar valores pequeños, pero no debe utilizar universalmente 0,05, pues los niveles de riesgo varían. Es más aconsejable utilizar las consideraciones de Sterne y Smith (15), y emplear el criterio que aparece en la Tabla 1, para interpretar el valor P. No es conveniente ni aceptable publicar expresiones como NS, \*, \*\*,  $P < 0,05$ , etc.; se debe publicar el valor obtenido e interpretarlo. Presentando el valor P en la publicación, el lector puede apreciar la evidencia que se obtuvo, pues el valor P es una medida de la evidencia que arrojan los datos contra  $H_0$ .

Es muy importante tener en cuenta los resultados que indiquen efecto como aquellos que no son relevantes, pues todos los estudios finalmente podrían integrarse en un metaanálisis y ser analizados en el contexto de los resultados de otros estudios. El punto es, que aunque no haya efectos, es decir, cuando  $H_0$  sea cierta, por efectos del azar se va a cumplir que un 5% de los estudios que se realicen tendrán valor P por debajo de 0,05 aun cuando no existan diferencias, y no existan sesgos en su conducción.

### Referencias

1. Pagano M Gauvreau K. Fundamentos de bioestadística. 2ª ed. Internacional México: Thompson; 2001.
2. Bacallao J. El fantasma de los paradigmas. Rev Habanera Cienc Méd. 2002;1(2):33-43.
3. Silva LC. Los peligros del lenguaje estadístico. En Locutora J. (editor). De la idea a la palabra. Cómo preparar, elaborar y difundir una comunicación científica. Barcelona: Publicaciones Permyer; 2003. p.149-58.
4. Monterrey P, Gómez-Restrepo C. Aplicación de las pruebas de hipótesis en la investigación en salud. ¿Estamos en lo correcto? Universitas Médica. 2007;48(3):193-206.

5. Rozeboom W. The fallacy of the null hypothesis significance test. *Psychological Bulletin*. 1960;57(5):416-28.
6. Cohen J. The earth is round ( $p < 0.05$ ). *Am Psychol*. 1994;49(12):997-1003.
7. Gigerenzer G. We need statistical thinking, not rituals. *BBS*. 1998;21(2):199-200.
8. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *BMJ*. 1986;292(6522):746-50.
9. Polit DF, Hungler BP. Investigación científica en ciencias de la salud. Principios y métodos. México: McGraw-Hill Interamericana, HealthCare Group; 2000.
10. Castañeda, GF. Una mirada a los intervalos de confianza en investigación. *Rev Colomb Psiquiatr*. 2004;33(2):193-201.
11. Silva LC. Cultura estadística e investigación científica en el campo de la salud: una mirada crítica. Madrid: Díaz de Santos; 1997.
12. Fleiss JL. Confidence intervals vs. significance tests: quantitative interpretation. *Am J Public Health*. 1986;76(5):587-8.
13. Moran JL, Solomon PJ. A farewell to P values? *Crit Care Resusc*. 2004;6(2):130-7.
14. Natrella M. The relation between confidence intervals and tests of significance. *The American Statistician*. 1960;14(1):20-22, 38.
15. Sterne JA, Smith GD. Sifting the evidence-what's wrong with significance tests? *BMJ*. 2001;322(7280):226-31

*Recibido para evaluación:* 2 de febrero de 2008  
*Aceptado para publicación:* 7 de abril de 2008

Correspondencia  
Pedro Monterrey  
*Departamento de Epidemiología Clínica y Bioestadística*  
*Hospital Universitario San Ignacio*  
*Carrera 7 No. 40 -62, piso 2*  
*Bogotá, Colombia*  
*pmonterrey@javeriana.edu.co*