

**Cortada de Kohan, Nuria ; Macbeth, Guillermo**

*El tamaño del efecto en la investigación psicológica*

*Effect dimension on psychological research*

Revista de Psicología Vol. 3 N° 5, 2007

Este documento está disponible en la Biblioteca Digital de la Universidad Católica Argentina, repositorio institucional desarrollado por la Biblioteca Central "San Benito Abad". Su objetivo es difundir y preservar la producción intelectual de la Institución.

La Biblioteca posee la autorización del autor para su divulgación en línea.

Cómo citar el documento:

Cortada de Kohan, N., Macbeth, G. (2007). El tamaño del efecto en la investigación psicológica [en línea]. *Revista de Psicología*, 3(5).

Disponible en: <http://bibliotecadigital.uca.edu.ar/repositorio/revistas/efecto-investigacion-psicologica-kohan-macbeth.pdf> [Fecha de consulta:....]

## **El tamaño del efecto en la investigación psicológica**

*Effect dimension on psychological research*

Nuria Cortada de Kohan  
Profesora Honoraria de la Universidad de Buenos Aires (UBA).

Guillermo Macbeth  
Becario postdoctoral del Consejo Nacional  
de Investigaciones Científicas y Técnicas (CONICET).  
Instituto de Investigaciones Psicológicas  
de la Universidad del Salvador (IIPUS)<sup>1</sup>

### **Resumen**

El propósito de este trabajo es centrarse en la pregunta más común que suele hacerse el investigador cuando comienza un estudio: ¿Con cuántos casos necesito trabajar? Los investigadores son cuidadosos con sus procedimientos pero a menudo no tienen en cuenta un tema fundamental de la estadística que les concierne directamente. Suponiendo que la hipótesis del investigador es cierta, es necesario determinar el tamaño de la muestra (n) que se necesita para dar a esta hipótesis una posibilidad razonable de demostrar que es correcta. El tamaño del efecto es el grado en que el fenómeno en cuestión se presenta en la población, es decir, el grado en que la hipótesis nula es falsa. Se presenta aquí el modo de representar el tamaño del efecto en una investigación según Jacob Cohen.

### **Abstract**

The aim of this work is to offer an account of the commonest question than the researcher in psychology usually asks at the beginning of every project: How many cases do I need? The researchers are careful with their procedures but often they do not consider a fundamental subject of statistics that concerns to them directly. Supposing that the hypothesis of the

---

<sup>1</sup> Marcelo T. de Alvear 1314 (C1058AAV)  
Ciudad Autónoma de Buenos Aires, Argentina  
Teléfono: (+5411)-4813-3404  
Correo electrónico: ncortada@psi.uba.ar

investigator is certain, it is necessary to determine the sample size ( $n$ ) that is needed to give to this hypothesis a reasonable possibility to demonstrate that it is, indeed, correct. The effect-size is the degree in which the effect under investigation appears in the population, that is to say, the degree in which the null hypothesis is false. It is presented in this paper the way to represent the effect-size according to Jacob Cohen.

*Palabras clave:* tamaño del efecto, potencia de una prueba, nivel de significación, número de casos.

*Key words:* effect-size, power analysis, significance level, number of cases.

### *Introducción*

El propósito de este trabajo se centra en la pregunta más común que suele hacerse el investigador cuando comienza un estudio: ¿Con cuántos casos necesito trabajar? (Kraemer & Thiemann, 1987). En general, no se suele apreciar debidamente que en los análisis estadísticos existen dos hipótesis a considerar: 1) la hipótesis nula que supone que la variable investigada no tiene efecto y son los datos los que tendrán la posibilidad de rechazar este supuesto (Grissom & Kim, 2005); y 2) si se puede rechazar la hipótesis nula entonces la hipótesis alternativa se acepta. Así, el centro de la estadística aplicada ha sido principalmente evaluar la hipótesis nula, mientras que a menudo la consideración de la hipótesis alternativa (es decir, la hipótesis del investigador) ha sido olvidada.

Generalmente, se puede contestar de inmediato la siguiente pregunta: Si la hipótesis nula es cierta, ¿cuál es la probabilidad de rechazarla? El error de rechazar la hipótesis nula cuando es verdadera se llama *error alfa* o *error de Tipo I* y depende del nivel de significación elegido por el investigador (generalmente es del 5 ó del 1 por ciento). Sin embargo, raramente se puede contestar la siguiente pregunta: Si la hipótesis alternativa es aceptada ¿en qué porcentaje se podría rechazar? Este es el error de *Tipo II* o *error beta*, es decir, es el error de rechazar la hipótesis alternativa siendo ésta verdadera (Cohen, 1994). El valor  $1 - \beta$  se llama potencia del test estadístico y siempre es de desear que este valor sea lo más grande posible (Cohen, 1992b).

El propósito de este trabajo es señalar qué debe hacerse para fijar la potencia de cualquier estudio antes de comenzar la investigación. Esto llevará a poder establecer la aceptación de la hipótesis alternativa.

### *El cálculo del tamaño del efecto y sus implicancias en estudios científicos*

Los investigadores generalmente son cuidadosos con los procedimientos de una investigación. Por ejemplo, los experimentalistas controlan la mayor parte de las variables, trabajan con muestras homogéneas, manipulan cuidadosamente las variables independientes, tratan de examinar a los participantes dentro de una franja de tiempo limitada, etc. A su vez, los que trabajan con investigaciones correlacionales definen con cuidado sus poblaciones y sus métodos de muestreo, establecen la confiabilidad y validez de sus instrumentos, hacen estudios piloto para cuidar todo el proceso, etc.

Sin embargo, a menudo los investigadores no tienen en cuenta un tema fundamental de la estadística que les concierne directamente. Suponiendo que la hipótesis del investigador es aceptada, es necesario determinar el tamaño de la muestra ( $n$ ) que se necesita para dar a esta hipótesis una posibilidad razonable de demostrar su aceptación. Pero para determinar la  $n$  se deben contestar algunos interrogantes muy importantes. Por ejemplo: suponiendo que la hipótesis alternativa es aceptada para el parámetro que interesa, ¿cuál es el valor mínimo en el que el grupo experimental debe diferenciarse de algún grupo control o valor de referencia para que el investigador pueda estar completamente satisfecho de que ha establecido la validez de la hipótesis?

Para poder responder a la pregunta anterior, el investigador debe haber llevado a cabo algún estudio preliminar a partir del cual habrá sacado los datos cuantitativos para poder estimar el parámetro, por ejemplo, la media aritmética, el porcentaje, el coeficiente de correlación, etc., así como también debe haber estimado la variabilidad de dicho valor estadístico. La diferencia cuantitativa expresada (por ejemplo, la diferencia entre las medias del grupo control y el grupo experimental) se divide por la variabilidad (por ejemplo, la desviación estándar) y el resultado es el tamaño del efecto (Cohen, 1988).

El tamaño del efecto es el grado en que el fenómeno en cuestión se presenta en la población, es decir, el grado en que la hipótesis nula es falsa.

Así, el investigador se encuentra en condiciones de responder a la siguiente pregunta esencial: en el supuesto caso en que la hipótesis es aceptada, ¿qué probabilidad se quiere asignar a esta investigación para obtener un resultado significativo (es decir, para poder rechazar la hipótesis nula)? Este interrogante se responde a partir de la potencia del test de significación (Cohen, 1978, 1992a). En este punto, se presenta un problema. Si se asigna un valor de po-

tencia casi máximo como, por ejemplo, un valor de potencia de 0,99 entonces el investigador se encontrará seguro al aceptar la hipótesis, sin embargo, el número de casos necesarios ( $n$ ) puede llegar a ser terriblemente grande, por ejemplo, alrededor de 150.000, lo cual imposibilita su estudio. En consecuencia, generalmente los investigadores se conforman con un valor de potencia de alrededor de 0,70 a 0,90. No obstante, siempre debe tenerse en cuenta que al trabajar con una potencia muy baja, menor que estos valores, es muy peligroso. Se comprende esta dificultad al establecer la relación entre los errores de Tipo II y I, es decir, estableciendo la relación  $\beta/\alpha$  (Cohen, 1990, 1994). Por ejemplo, si un investigador usa un  $\alpha = 0,05$  y una potencia de 0,40, entonces  $\beta = 1 - 0,40 = 0,60$ . La relación entre los errores de Tipo II y I ( $\beta/\alpha$ ) es de  $0,60/0,05 = 12$ . Esto implica que el error de Tipo I ( $\alpha$ ) es 12 veces más serio que el error de Tipo II ( $\beta$ ).

Así, el poder o potencia de una prueba estadística es la probabilidad de rechazar la hipótesis nula cuando la hipótesis nula es falsa y, por lo tanto, debe ser rechazada. El poder aumenta a medida que el error de Tipo II o beta disminuye, es decir, cuando la potencia es igual a  $1 - \text{beta}$ . La potencia es la probabilidad de no cometer el error de Tipo II, esto es, de no dejar de ver un efecto o una relación en donde realmente existe.

Rosenthal y Rosnow (1991) señalan que el análisis de la potencia tiene dos propósitos fundamentales: 1) el diseño de una investigación; y 2) la evaluación de investigaciones ya realizadas, como en el meta-análisis (Macbeth, Cortada de Kohan & Razumiejczyk, en prensa), dado que el conocimiento de investigaciones anteriores resulta fundamental para determinar el tamaño necesario de una muestra y lograr un nivel de significación alto para el tamaño del efecto que se desea.

Existen diferentes modos de representar el tamaño del efecto en una investigación. Uno de ellos es el propuesto por Jacob Cohen (1988, 1990).

Cuando una investigación supone la comparación entre dos grupos, es común indexar el tamaño del efecto examinando la diferencia entre sus medias. Tales diferencias, sin embargo, tienen poco sentido aparte de la escala de medición usada. Por esto es útil dividir la diferencia de las medias por la sigma ( $\sigma$ ) común a ambas, de modo que la diferencia se represente en unidades estándar. A estas diferencias estandarizadas Cohen las denominó  $d$ .

$$d = \frac{\overline{X}_1 - \overline{X}_2}{\sigma}$$

Cuando las  $\sigma$  no han sido informadas, una buena aproximación a  $d$  es la siguiente:

$$d = \frac{2t}{\sqrt{g \cdot l}}$$

Cohen ha proporcionado las tablas indispensables para siete pruebas estadísticas muy usadas, como  $t$ ,  $r$ ,  $\chi^2$ ,  $F$ , etc. A su vez, este autor informa los valores aproximados del tamaño de efecto para cada prueba según sea un tamaño del efecto pequeño ( $d = 0,20$ ), mediano ( $d = 0,50$ ) o grande ( $d = 0,80$ ). Informan las tablas, además, los tamaños de la potencia de 0,15 a 0,90 para cada tamaño el efecto, así como los valores para pruebas de tipo unilateral (un solo extremo) o bilaterales (dos extremos). Por ejemplo, con un efecto modesto de tipo mediano de 0,50 para la prueba  $t$  en cada grupo se necesita para una potencia de 0,80 (lo que es una beta de 0,20 y un alfa bilateral de 0,05) una muestra ( $n$ ) igual a 64 casos en cada grupo. Si alfa = 0,01, entonces  $n = 96$  casos en cada grupo.

**Tabla 1**

$\alpha = 0,01$  unilateral (prueba  $t$  para grupos independientes)

Potencia	0,70	0,80	0,90
$d = 0,20$	400	500	700
$d = 0,50$	69	84	110
$d = 0,80$	28	33	42

**Tabla 2**

$\alpha = 0,05$  unilateral (prueba  $t$  para muestras independientes)

Potencia	0,70	0,80	0,90
$d = 0,20$	250	300	450
$d = 0,50$	39	50	68
$d = 0,80$	15	20	27

**Tabla 3**

$\alpha = 0,01$  bilateral (prueba t para muestras independientes)

Potencia	0,70	0,80	0,90
d = 0,20	500	600	800
d = 0,50	80	96	120
d = 0,80	33	38	48

**Tabla 4**

$\alpha = 0,05$  bilateral (prueba t para muestras independientes)

Potencia	0,70	0,80	0,90
d = 0,20	300	400	600
d = 0,50	50	64	84
d = 0,80	21	26	34

Existen diversas variantes sobre el tamaño del efecto como, por ejemplo, la delta ( $\Delta$ ) de Glass (Clement, Mercier & Pastò, 2002; Rosenthal & Rosnow, 1991) que se define como la diferencia entre las medias de los grupos sobre las del grupo control, y la g de Hedges & Olkin (1985) que es la diferencia entre las medias, dividido por la s de ambos grupos combinados.

La importancia de la determinación del tamaño del efecto y del análisis de la potencia antes de comenzar una investigación ha determinado la existencia de una serie de hechos y condiciones que siempre deben tenerse en cuenta al diseñar un estudio, por ejemplo: 1) cuánto más fino es el nivel de significación, más grande debe ser el tamaño de la muestra (se necesitan más casos para una prueba a nivel alfa de  $p = 0,01$  que a nivel  $p = 0,05$ ); 2) las pruebas bilaterales (de dos extremos) necesitan muestras más grandes que las unilaterales; 3) cuánto más pequeño es el tamaño del efecto, más grande debe ser la muestra; 4) cuánto mayor sea la potencia, mayor será la muestra necesaria; 5) cuánto más pequeña es la muestra, menor es la potencia y mayor es la probabilidad del error de la estimación; 6) una muestra de 20 o menos casos tiene un gran riesgo de error, de modo que no es recomendable; y 7) nunca debe hacerse una encuesta de opinión con menos de 1000 casos.

## *Conclusiones*

Cuando el investigador quiere averiguar la  $n$  necesaria en un estudio con dos muestras independientes (control y experimental), debe pensar en la  $n$  como una función del tamaño del efecto ( $d$ ), del alfa (unilateral o bilateral) y de la potencia. Por ejemplo, si un investigador desea una potencia de 0,80 para detectar un tamaño del efecto de 0,50 ( $d$  mediano) con un alfa = 0,05 (bilateral) hallará en las tablas de Cohen la cantidad de casos necesarios, en este ejemplo,  $n = 64$  casos (en cada grupo). Si la prueba fuera unilateral, solo necesitaría  $n = 50$  casos en cada grupo. Si se necesita un tamaño del efecto más fino ( $d = 0,20$ ) en los ejemplos anteriores, necesitará en el primer ejemplo  $n = 400$  (para cada grupo) y, en el segundo ejemplo,  $n = 300$  (para cada grupo).

## Bibliografía

- CLEMENT, M., MERCIER, P. & PASTO, L. (2002). Sample Size, Confidence, and Contingency Judgement. *Canadian Journal of Experimental Psychology*, 56(2), 128-137.
- COHEN, J. (1978). Partialled Products are Interactions; Partialled Powers are Curve Components. *Psychological Bulletin*, 85(4), 858-866.
- (1988). *Statistical Power Analysis for the Behavioral Sciences (Rev.Ed.)*. New York: Academic Press.
- (1990). Things I Have Learned (So Far). *American Psychologist*, 45(12), 1004-1012.
- (1992a). A Power Primer. *Psychological Bulletin*, 112(1), 155-159.
- (1992b). Fuzzy Methodology. *Psychological Bulletin*, 112(3), 409-410.
- (1994). The Earth is Round ( $p < .05$ ). *American Psychologist*, 49(12), 997-1003.
- GRISSOM, R.J. & KIM, J.J. (2005). *Effect Sizes for Research*. London: LEA.
- HEDGES, L. & OLKIN, D. (1985). *Statistical Methods in Meta Analysis*. Ney York: Academic Press.
- KRAEMER, H.C. & THIEMANN, S. (1987). *How Many Subjects?* London: Sage.
- MACBETH, G., CORTADA DE KOHAN, N. & RAZUMIEJCZYK, E. (en prensa). El Meta-Análisis: La Integración de los Resultados Científicos. *Revista Evaluar*.
- ROSENTHAL, R. G. & ROSNOW, R. L. (1991). *Essentials of Behavioral Research. Research Methods and Data Analysis (2<sup>nd</sup> Edition)*. Boston, Mass.: McGraw Hill.