

## G\*POWER 3: SOFTWARE PARA EL ANÁLISIS DE LA POTENCIA ESTADÍSTICA

La potencia (*statistical power*) de una prueba estadística es la probabilidad de que la hipótesis nula ( $H_0$ ) sea rechazada cuando esta es falsa, en otras palabras, es la probabilidad de obtener resultados estadísticamente significativos (Cohen, 1992).

**A**unque el concepto de potencia estadística es bastante conocido en la investigación experimental, son muy pocos los trabajos en los que se toma en cuenta, lo que indica una desatención importante por parte de los investigadores. Algunos estudios muestran el escaso interés de gran parte de autores por controlar la potencia, así como la omisión de exigencias al respecto por parte de las revistas especializadas (Bono & Arnau, 1995; Valera, Sánchez, Marín & Velandrino, 1998). Asimismo, se observa un sobreuso y mal uso de la técnica de comprobación de la hipótesis nula en el grueso de la investigación psicológica (Krantz, 1999; Jones & Sommerlund, 2007). Otro error común es buscar solamente saber si la relación entre las variables que se estudian es "estadísticamente significativa", desconociendo la potencia que puedan tener sus resultados (Barker-Baussel, R. & Li, 2002).

Algunos autores (Barker-Baussel, R. & Li, 2002) consideran que la inferencia estadística cuenta con dos pilares: La significancia estadística y la potencia estadística, siendo este último el más débil, en cuanto a uso y dominio de esta noción se refiere. Según Bono & Arnau (1995), ambos pilares hacen parte de dos corrientes teóricas diferentes en estadística, donde se enseñan temáticas que tiene la potencia estadística con el nivel de significancia ( $\alpha$ ), tamaño de la muestra ( $n$ ), y tamaño del efecto (TE). Cuando los valores de cualquiera de los cuatro son modificados, los otros tres cambian y cuando tres de las cuatro variables son modificadas, la cuarta resulta determinada.

## TAMAÑO DEL EFECTO (TE)

El tamaño del efecto es el grado de discrepancia entre la  $H_0$  y la  $H_1$ , en otras palabras, es el grado en el que la hipótesis nula es falsa. El tamaño del efecto es un indicador del distanciamiento de la evidencia experimental frente a la  $H_0$ . Por ejemplo, en una prueba t de diferencia de medias, cuanto más se aleje el índice t de 0, mayor será el tamaño del efecto y mayores serán las probabilidades de rechazar la hipótesis nula ( $H_0$ = no hay diferencias, es decir, un índice igual a 0). Cuando el tamaño del efecto es muy bajo, aumenta el riesgo de cometer error tipo II si se rechaza la hipótesis nula, esto depende del nivel de significancia con el que se trabaje.

Cuando se realizan análisis de potencia es necesario especificar el TE, lo cual no es sencillo, ya que, muchas veces, no se tiene conciencia de la magnitud de los fenómenos. La dificultad para determinar el tamaño del efecto se debe, principalmente, al sobreuso de las pruebas de significación, donde solo se presta atención a la significancia, la cual nos dice si hay diferencias significativas o no; sin embargo, poco informa sobre la magnitud real del fenómeno bajo estudio (Cohen, 1992a). Para conocer la importancia de un TE determinado, es necesario remitirse a investigaciones anteriores del tema que se esté abordando y/o adoptar ciertos criterios, como, por ejemplo, los aportados por Cohen (1992b). Este mismo autor advierte sobre los riesgos de excederse en las convencionalidades y propone un uso razonable de las mismas.

## NIVEL DE SIGNIFICANCIA ( $\alpha$ )

El nivel de significancia ( $\alpha$ ) se refiere a la probabilidad de rechazar la hipótesis nula siendo cierta. Generalmente se fija en un valor convencional de .05, el cual podría ser mayor o menor dependiendo del tipo de estudio que se realice (por ejemplo, en un estudio exploratorio es recomendable fijar valores de  $\alpha$  un poco mayores). Cuando el valor de  $\alpha$  es muy estricto (por ejemplo,  $\alpha=0,001$ ), se reduce el riesgo de incurrir en error tipo I; no obstante, se aumenta la probabilidad de incurrir en error tipo II, denotada como  $\beta$ , reduciendo la potencia (ya que esta última es el complemento de  $\beta$ ). Teniendo esto en mente, es importante ponderar los riesgos de cometer ambos tipos de error, tratando de mantener un balance.

## TAMAÑO MUESTRAL (N)

El tamaño muestral (N) también afecta la potencia. Valores grandes de N en las investigaciones aumentan la potencia, puesto que se reduce el error aleatorio y hay una mayor probabilidad de que el fenómeno bajo estudio se presente (Arnau & Bono, 1995). Las conclusiones que se deduzcan del estudio de una muestra, dependen de qué tan representativa sea, así como de qué tanto se acerque al tamaño de la población.

## G\*POWER 3

Dada la relación que hay entre estos conceptos y la importancia de cada uno de ellos en el campo de la investigación social, es importante contar con herramientas que permitan realizar estudios que incluyan análisis de las cuatro importantes variables descritas anteriormente. Este tipo de análisis son conocidos como estudios de potencia.

*G\*Power* es un software de uso libre, creado por investigadores de la Universidad Heinrich Heine de Düsseldorf en Alemania, el cual permite realizar varios estudios de potencia para los test estadísticos más usados (Universitat Dusseldorf, 2016):

- A priori
- Compromiso
- Criterio
- Post-hoc
- Sensibilidad

Faul, Erdfelder, Lang y Buchner (2007) mencionan que cualquiera de estos análisis en *G\*Power* 3 debe hacerse siguiendo cuatro pasos: 1) Seleccionar el test estadístico, 2) escoger uno de los cinco tipos de análisis de potencia, 3) proporcionar los valores de entrada requeri-

dos para el análisis, y 4) seleccionar el botón “calculate” y obtener los resultados.

A continuación, se ofrecerá un ejemplo de cómo realizar un análisis paso a paso de un estudio a priori. En este tipo de estudios se calcula el tamaño muestral en función de un nivel  $\alpha$  preestablecido, un nivel  $1 - \beta$  y el tamaño del efecto que se busca encontrar.

EJEMPLO DE CASO

En una investigación, se tiene como objetivo general encontrar diferencias en la puntuación en una prueba de matemáticas entre estudiantes que hicieron parte de un método de enseñanza piloto (grupo A) y otros que siguieron la instrucción tradicional (grupo B). Se trabajará con un nivel  $\alpha=0,05$ , una potencia esperada de 0,90 y un tamaño del efecto grande ( $\geq 0,8$ ). Los investigadores quieren conocer el tamaño de muestra requerido para encontrar diferencias significativas que den cuenta de que el método A es mucho más efectivo que el B. La prueba estadística que se utilizará será la  $t$  diferencia de medias en muestras independientes.

Para realizar este análisis, se debe iniciar el programa, el cual muestra una ventana tal como se observa en la Figura 1.

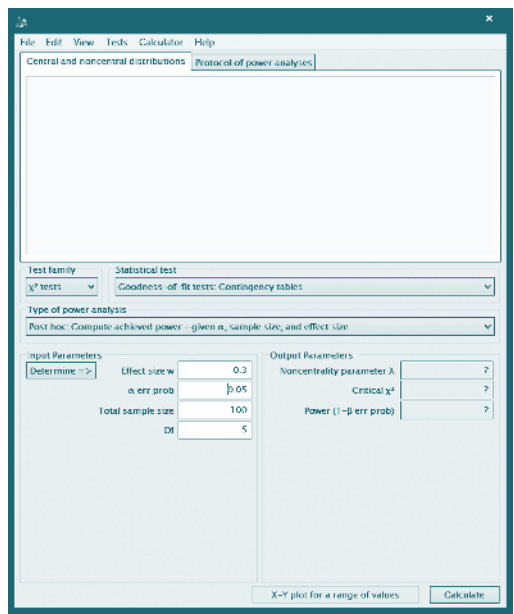


Figura 1. Ventana principal del G\*Power 3.

En primer lugar, se escoge el test estadístico que se utilizará. Para este caso, se trata de una prueba t. Primero se selecciona la familia del test (Figura 2) y, después, el test específico planeado para nuestro trabajo (Figura 3).

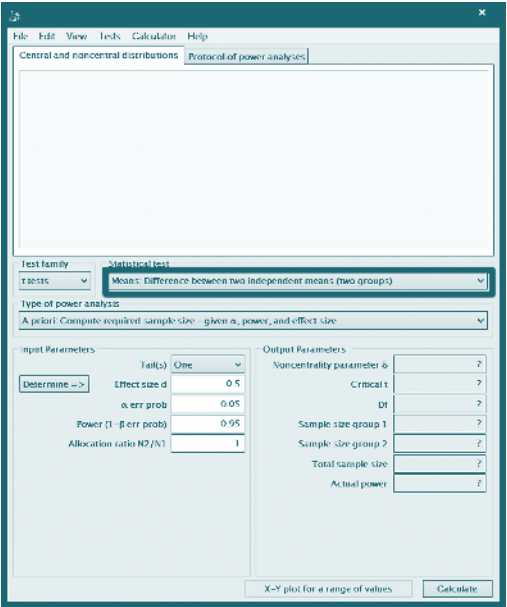


Figura 2. Selección de la familia de test.

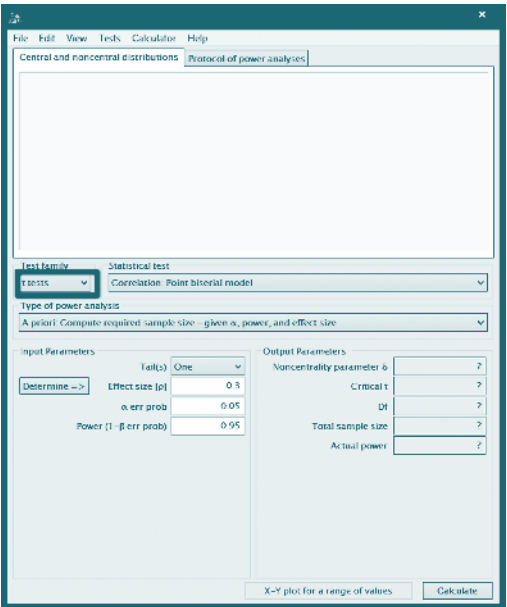


Figura 3. Selección del test. .

Después, en el siguiente menú desplegable, se escoge el estudio de potencia que se desea realizar. En este caso, se trata de un estudio *a priori* (Figura 4).

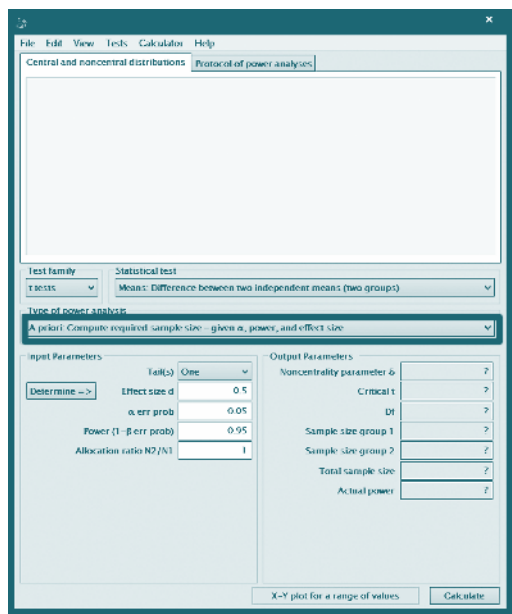


Figura 4. Selección del tipo de análisis de potencia – A priori.

En seguida, verificamos si los valores que aparecen por defecto en los parámetros de entrada corresponden con los que se trabajarán en la investigación, de lo contrario se realizan los cambios pertinentes (Figura 5). También se incluye la razón entre un grupo y el otro, en este caso, dicha razón será de 1:1. Cabe recordar que los valores que ingresemos en este paso, dependen de la naturaleza de nuestro estudio y de las características que los investigadores crean que deban tener sus resultados.

En este tipo de análisis a priori, se pueden modificar los parámetros, para ver cómo cambia el tamaño de muestra requerido en función de ellos. Por ejemplo, si cambiamos el nivel  $\alpha$  a un nivel más restrictivo, seguramente aumentará el tamaño de muestra requerido para evitar el error tipo I. Finalmente, se escoge la opción "Calculate" que está en la parte inferior derecha de la ventana (Figura 6).

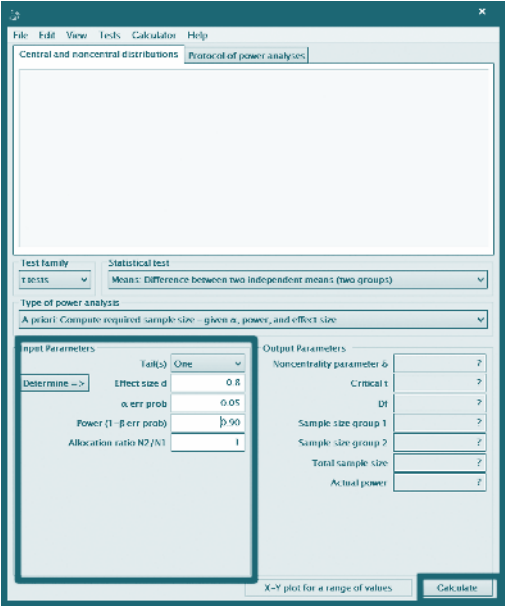


Figura 5. Identificación del módulo de parámetros de entrada.

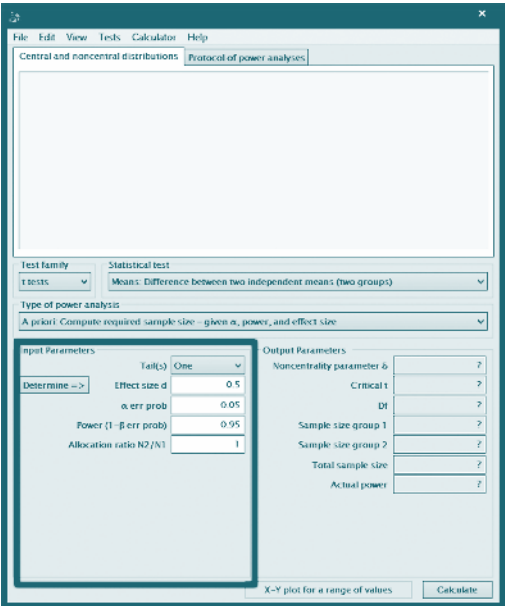


Figura 6. Definición de los parámetros de entrada y botón para calcular lo de salida.

Después de esto, obtenemos los parámetros de salida con su respectiva gráfica, tal como se observa en la Figura 7. El tamaño de muestra que se requiere, dados los niveles de error aceptados y el tamaño de efecto esperado, es de 28 estudiantes para cada grupo.

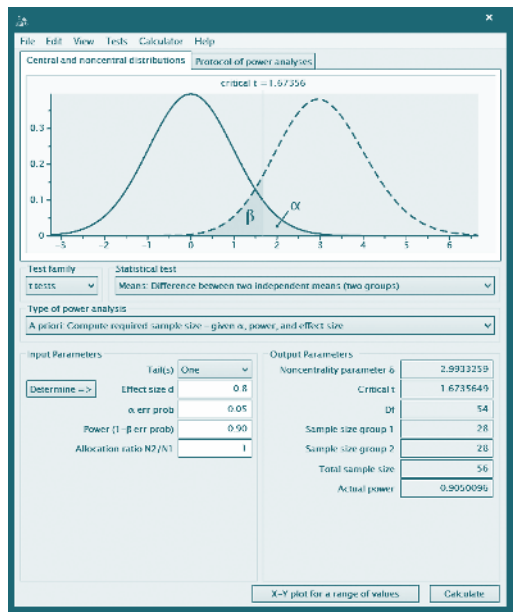


Figura 7. Output del análisis a priori según los parámetros especificados.

Además de esta información, podemos obtener distintos gráficos para modelar la relación entre los distintos parámetros de entrada con el de salida (Figuras 8, 9 y 10).

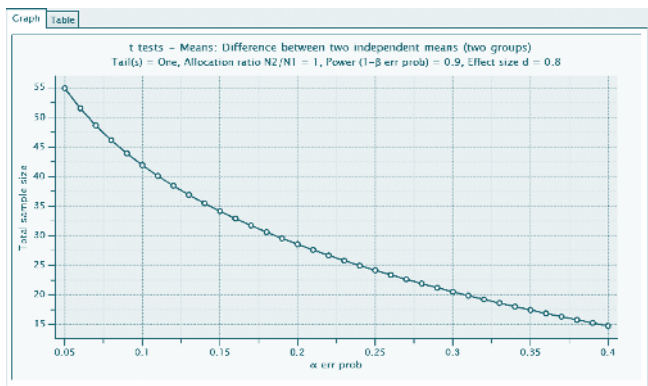


Figura 8. Relación entre n y potencia.



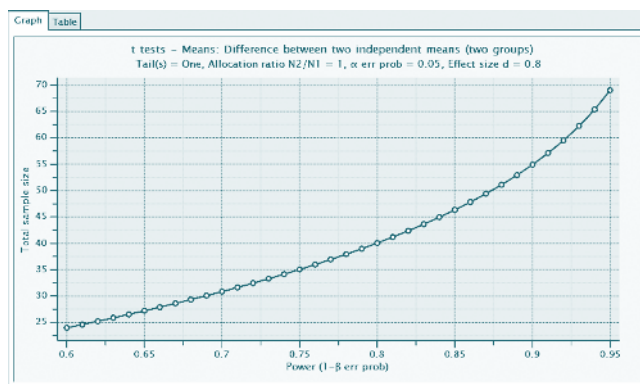


Figura 9. Relación entre  $n$  y  $\alpha$ .

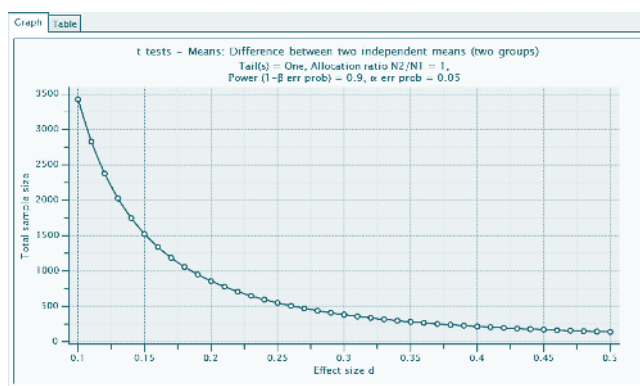


Figura 10. Relación entre  $n$  y TE.

En este ejemplo se aprecia la utilidad del programa y la facilidad para hacer uso de él. G\*Power 3 ofrece distintos tipos de análisis para una gran variedad de test estadísticos, lo cual pone al alcance de la mano una herramienta bastante útil para los investigadores de diferentes disciplinas.

## REFERENCIAS

- Barker-Bausse, R. & Li, Y. F. (2002). *Power Analysis for Experimental Research*. A practical guide for the biological, medical and social sciences. Nueva York: Cambridge University Press.
- Bono, R. & Arnau, J. (1995). Consideraciones generales en torno a los estudios de potencia. *Anales de psicología*, 11(2), 193-202.
- Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science*, 1(3), 98-101.
- Cohen, J. (1992a). Cosas que he aprendido (hasta ahora). *Anales de Psicología*, 8(1-2), 3-18.
- Cohen, J. (1992b). A power primer. *Psychological Bulletin*, 112 (1), 155-159.
- 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(2), 175-191. DOI: 10.3758/BF03193146
- Jones, A. & Sommerlund, B. (2007). A critical discussion of null hypothesis significance testing and statistical power analysis within psychological research. *Nordic Psychology*, 59(3), 223-230.
- Kraemer, H. C. (1985). A strategy to teach the concept and application of power of statistical tests. *Journal of Educational Statistics*, 10(3), 173-195.
- Krantz, D. (1999). The null hypothesis testing controversy in psychology. *Journal of the American Statistical Association*, 94(448), 1372-1381.
- Universitat Dusseldorf. (2016). G\*Power 3.1 Manual. Recuperado de: [http://www.gpower.hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche\\_Fakultaet/Psychologie/AAP/gpower/GPowerManual.pdf](http://www.gpower.hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/gpower/GPowerManual.pdf)
- Valera, A., Sánchez, J., Marín, F. & Velandrino, A. P. (1998). Potencia estadística de la revista de Psicología General y Aplicada. *Revista de Psicología General y Aplicada (1990-1992)*, 51(2), 233-246.

