



Artículo de revisión bibliográfica. Revista Killkana Técnica. Vol. 7, No. 3, pp. 37-46, septiembre-diciembre, 2023.
ISSN 2528-8024. ISSN Elect. 2588-0888. Universidad Católica de Cuenca

La potencia estadística de los estudios con prueba de hipótesis

Statistical power of hypothesis testing studies



Estefanía Patiño-Ramón*

Fundación BRAINS. Cuenca-Ecuador
eestefaniap7@gmail.com

Carla Larrea Eyzaguirre

Universidad Privada Franz Tamayo, La Paz-El Alto

David Calderón

Estudios y Análisis Consultores Asociados Eaconsul Cía. Ltda. Cuenca-Ecuador



DOI

Resumen

Objetivo: El objetivo de la presente revisión es conocer la importancia de la potencia estadística en el muestreo de grupos paralelos para la comprobación de hipótesis con sus respectivos porcentajes de confiabilidad. **Metodología:** Se ejecutó la búsqueda en bases científicas por medio de artículos en revistas indexadas y libros de metodología de la investigación en idioma español e inglés. **Conclusiones:** El cálculo de la potencia estadística es un paso fundamental para asegurar la confiabilidad y el diseño de un estudio junto con el tamaño muestral al controlar los errores tipo I y tipo II desde el momento de la planificación hasta la finalización del estudio ofreciendo no sólo credibilidad sino también la garantía de los resultados obtenidos.

Palabras clave: *cálculo de potencia, estadística, potencia estadística, tamaño muestral.*

Abstract

Aim: The objective of this review is to know the importance of statistical power in the sampling of parallel groups for hypothesis testing with their respective reliability percentages. **Methodology:** A search was carried out in scientific databases by means of articles in indexed journals and research methodology books in Spanish and English. **Conclusions:** The calculation of statistical power is a fundamental step to ensure the reliability and design of a study together with the sample size by controlling type I and type II errors from the moment of planning to the end of the study, offering not only credibility but also the guarantee of the results obtained.

Key words: *power calculation, statistics, statistical power, sample size.*

I. INTRODUCCIÓN

En ciencias de la salud, las intervenciones deben ser basadas en evidencia, para lo cual el método científico es la herramienta que permite comprobar la validez de una intervención, para lo cual la misma tiene una serie de pasos que deben ser realizados rigurosamente, de manera sistemática, ordenada y con visión en el método inductivo[1]

El poder estadístico de un estudio está relacionado a la probabilidad de rechazar la hipótesis del investigador, cuando ésta debiera ser aceptada[2]. El tamaño de la muestra tiene mucho que ver al momento de plantearnos una estrategia para responder a una pregunta de investigación [2]

Usualmente, la potencia estadística se estima al momento de la planificación del estudio con el fin de determinar el tamaño de la muestra y posteriormente para la publicación e interpretación de los resultados obtenidos. Además, permite establecer la viabilidad del estudio en base a las hipótesis planteadas por el investigador sobre los resultados y los recursos [3],[4]

Cuando tenemos una investigación que pretende comprobar una hipótesis, vamos a enfrentarnos al sistema de hipótesis[2]

Hipótesis nula: No hay diferencia entre los tratamientos.

Hipótesis alternativa: Existe diferencia entre los tratamientos.

Al momento de realizar el estudio nosotros podemos concluir que los grupos son equivalentes (aceptamos la hipótesis nula) o podemos concluir que los grupos son estadísticamente diferentes (aceptamos la hipótesis alterna). Al tomar esta decisión estamos a expensas de cometer error tipo I (cuando rechazamos la hipótesis nula incorrectamente) o error tipo II (cuando se acepta la hipótesis nula incorrectamente)[2],[5].

Un conocimiento pobre del poder estadístico relacionado al tamaño de muestra conduce a un error en la comunicación con el bioestadístico, lo que se traduce en un diseño de investigación muy pobre [6].

El Dr. Staffa en 2020 propone 5 pasos para plantear adecuadamente el poder estadístico de un estudio y el tamaño de la muestra[6] :

- Paso 1.- Definir la variable dependiente, los grupos a comparar y formulación de la hipótesis.
- Paso 2.- Definir la diferencia que se espera encontrar entre el grupo 1 y el grupo 2.
- Paso 3.- Definir la prueba estadística apropiada para evaluar la hipótesis que se plantea.
- Paso 4.- Desarrollar el cálculo del tamaño de muestra.
- Paso 5.- Redactar el tamaño de muestra y el poder del estudio considerando el error Alpha o la significación estadística de la prueba estadística que se va a utilizar.

El objetivo de la presente revisión es conocer la importancia de la potencia estadística en el muestreo de grupos paralelos para la comprobación de hipótesis con sus respectivos porcentajes de confiabilidad.

II. ESTADO DEL ARTE

El tamaño adecuado de la muestra es un punto clave en el desarrollo de la investigación científica, debido a que permite garantizar diferencias clínicamente significativas y una potencia adecuada en aquello que se está investigando. [7],[8]

Es así, que mientras mayor sea el tamaño de la muestra, mayor será la potencia y del mismo modo, la probabilidad de significación estadística. Esto debido a que el aumento del tamaño de la muestra reduce el error estándar. [9] Una muestra ideal permite conseguir una potencia estadística suficiente y a su vez optimizar la capacidad de predicción de una enfermedad [7], [10],[11].

La potencia del estudio depende de la magnitud de la asociación o diferencia entre los grupos, la mutabilidad de la variable respuesta, el valor de significancia estadística y el tamaño de la muestra. [12],[13] Finalmente, mientras se tenga un buen número de participantes en una investigación, mayor será su potencia estadística. Por el contrario, aquellos estudios que cuentan con una muestra muy pequeña predicen conclusiones falsamente negativas. [12]

En cuanto a la literatura revisada se pudo encontrar que los artículos que narran sobre el cálculo del tamaño muestral casi siempre enfatizan en el diseño descriptivo, sin embargo, existen hasta 4 escenarios del muestreo a saber y que están descritos a continuación [14]:

Opción muestra de un grupo

Para aquellos estudios descriptivos en donde se plantea una investigación que implica como muestra un solo grupo conocido de estudio y que es representativo[14].

$$n = \frac{Z^2 * P * (1-P) * N}{Z^2 * P * (1-P) + E^2 * (N-1)}$$

Opción grupos para variable cualitativa

Se emplea la fórmula siguiente para investigaciones analíticas con dos muestras, es decir, dos grupos[14].

$$n = \left(\frac{Z\alpha \sqrt{2p * q} + Z\beta \sqrt{(p2 * q2) + (p1 * q1)}}{p1 - p2} \right)^2$$

Opción grupos para variable cuantitativa

En estudios que miden variables cuantitativas y que pretenden comparar dos grupos[14].

$$n = \frac{2(Z\alpha + Z\beta)^2 * \delta^2}{e^2}$$

El uso de grupos para variables cuantitativas nos permite trabajar bajo la hipótesis nula, con un planteamiento unilateral. Es decir, se cambia $Z\alpha/2$ por $Z\alpha$. Así también, el potencial uso de variables cuantitativas en la comparación de medias, nos facilita el cálculo de la probabilidad de confirmar que dos tratamientos son iguales (o que uno no es inferior al otro) cuando esta afirmación es verdadera.[8],[15]

En tal sentido, es necesario conocer la importancia de calcular tanto la potencia estadística como el tamaño muestral de un estudio, de ahí que existen varios softwares que ejecutan dichos cálculos como el OpenEpi, sin embargo, en este artículo se brinda una explicación matemática de la fórmula genérica para determinar la potencia estadística.

Muestreo de comparación de proporciones

Partiendo de la fórmula para muestreo de comparación de proporciones, presentamos esta

propuesta de despeje de fórmula para hallar el valor Z de la potencia de un estudio, que luego tendrá que ser ubicado en la tabla estadística Z, para determinar su equivalente en el porcentaje de potencia [14].

$$n = \left(\frac{Z \alpha \sqrt{2p * q} + Z\beta \sqrt{(p2 * q2) + (p1 * q1)}}{p1 - p2} \right)^2$$

$$\sqrt{n} = \frac{Z \alpha \sqrt{2p * q} + Z\beta \sqrt{(p2 * q2) + (p1 * q1)}}{p1 - p2}$$

$$\sqrt{n} (p1 - p2) = Z \alpha \sqrt{2p * q} + Z\beta \sqrt{(p2 * q2) + (p1 * q1)}$$

$$\sqrt{n} (p1 - p2) - (Z \alpha \sqrt{2p * q}) = Z\beta \sqrt{(p2 * q2) + (p1 * q1)}$$

$$\frac{\sqrt{n} (p1 - p2) - (Z \alpha \sqrt{2p * q})}{\sqrt{(p2 * q2) + (p1 * q1)}} = Z\beta$$

Donde:

- Z α = 1.96 constante de confiabilidad para error tipo I (95% confiabilidad)
- Z β = Valor Z del error tipo II (queremos hallar)
- n = tamaño de muestra por grupo
- p1 = probabilidad de éxito en el grupo 1
- q1 = probabilidad de fracaso en el grupo 1
- p2 = probabilidad de éxito en el grupo 2
- q2 = probabilidad de fracaso en el grupo 2
- p = promedio de p1 y p2
- q = promedio de q1 y q2

Aplicación de la fórmula en el Software OpenEpi versión 3.01

Para la aplicación de la fórmula de muestreo se emplea el Software OpenEpi versión 3.01, en donde se selecciona el tipo de cálculo que desea efectuarse y se introducen los datos para obtener los resultados de forma inmediata como se puede observar en la Figura 6. [16].

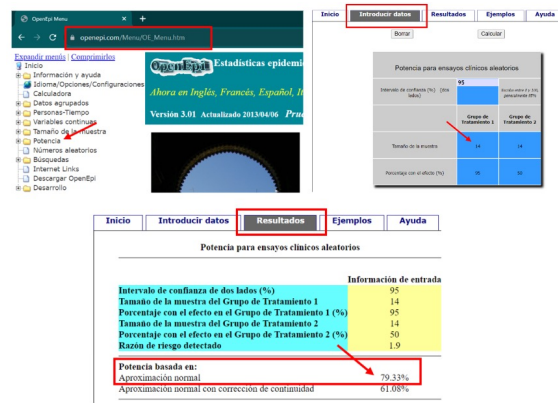


Fig. 1 Aplicación de la fórmula en el Software OpenEpi versión 3.01 [16].

Fórmula de muestreo para la comparación de dos grupos con variable cuantitativa

$$n = \frac{2(Z\alpha + Z\beta)^2 * \delta^2}{e^2}$$

$$\frac{n * e^2}{\delta^2} = 2(Z\alpha + Z\beta)^2$$

$$\frac{n * e^2}{2 * \delta^2} = (Z\alpha + Z\beta)^2$$

$$\sqrt{\frac{n * e^2}{2 * \delta^2}} = Z\alpha + Z\beta$$

$$\sqrt{\frac{n * e^2}{2 * \delta^2}} - Z\alpha = Z\beta$$

Donde:

- Z α = 1.96 constante de confiabilidad para error tipo I (95% confiabilidad)
- Z β = Valor Z del error tipo II (queremos hallar)
- n = tamaño de muestra por grupo
- δ = varianza de la variable
- e = Diferencia entre promedios (grupo 1 con grupo 2)

DISCUSIÓN

La potencia estadística desempeña un papel crucial en la investigación, actuando como un radar para detectar señales significativas en los datos. No obstante, esta herramienta puede tener limitaciones, especialmente en relación con el tamaño de la muestra, lo que aumenta el riesgo de errores de tipo II, donde una muestra pequeña o baja potencia puede ocultar una señal real [17]. Confiar en resultados incorrectos podría llevar a conclusiones equivocadas sobre la existencia o no de un efecto. [18]

Para mejorar la potencia estadística, se puede aumentar el tamaño de la muestra, expandiendo así nuestro “radar” para capturar más señales y mejorar la capacidad del estudio para detectar efectos reales. Además, ajustar el nivel de significancia ofrece flexibilidad en la interpretación de los resultados, controlando la probabilidad de errores tipo I y tipo II [17].

El análisis de potencia se ha convertido en un tema importante en los procesos de investigación y publicación. Sin embargo, el cálculo de la potencia requiere amplios conocimientos estadísticos, existiendo poco personal con habilidades de programación y los programas comerciales suelen ser demasiado costosos para su uso en la práctica [19],[20]. Los estudios con poca potencia no proporcionan estimaciones precisas y, por lo tanto, informan inadecuadamente sobre el efecto del tratamiento, dificultando las decisiones o juicios basados en evidencia [19]. Por ello es crucial que una investigación tenga suficiente potencia para mostrar diferencias clínicamente significativas.

Los cálculos de potencia ayudan a determinar si un estudio es factible en base a supuestos a priori sobre los resultados del estudio y los recursos disponibles, realizando un balance entre la probabilidad de observar el efecto real

y las probabilidades de error tipo I (falso positivo) y tipo II (falso negativo) [3],[20].

Se necesitan conocer cuatro componentes básicos a tomar en cuenta: alfa, beta, diferencia clínicamente relevante mínima y variabilidad. Estos componentes determinan la potencia estadística y el riesgo de errores tipo I y tipo II [20]. Especificar el poder estadístico deseado, típicamente 80% o 90%, que representa la probabilidad de rechazar correctamente la hipótesis nula cuando ésta es falsa [7].

Los estudios con tamaños de muestra muy grandes pueden producir hallazgos estadísticamente significativos con tamaños de efecto pequeños que pueden no ser clínicamente relevantes. Es beneficioso que los cuatro componentes se presenten claramente en los estudios analíticos. [21] No incluir estos elementos puede limitar la capacidad de otros investigadores para replicar los hallazgos del estudio y generar dificultades a la hora de interpretar los resultados del estudio [17].

El reporte deficiente de los cálculos a priori es otro problema, a pesar de ser esenciales [20]. Sin embargo, las fórmulas tradicionales solo existen para diseños simples con baja potencia, como prueba t o ANOVA de un factor. Los diseños factoriales y de bloques aleatorios son más potentes, pero no tienen fórmulas, por lo que no se pueden justificar a pesar de requerir menos muestra. Esto crea un dilema para el investigador [10],[20].

La potencia está directamente relacionada con el tamaño de muestra, alfa, tamaño del efecto. Reportar estos elementos permite una evaluación completa del estudio desde lo estadístico y clínico [22]. El tamaño de la muestra es un aspecto crítico en los estudios científicos, pues afecta directamente el diseño del estudio y la hipótesis planteada [20]. Un tamaño de muestra incorrecto puede llevar a resultados inadecuados. Existen fórmulas y programas

estadísticos para calcular el tamaño de muestra necesario, en función de la potencia deseada, el alfa, la desviación estándar y el tamaño del efecto esperado[22]. El software G*Power es recomendable para el cálculo del tamaño de la muestra y la potencia para diversos métodos estadísticos (F, t, χ^2 , Z y pruebas exactas), ya que es fácil de usar y gratuito. Proporciona calculadoras de tamaño del efecto y opciones gráficas [19],[23].

Un estudio con potenciación adecuada tiene más probabilidades de generar resultados confiables. La derivación de la fórmula de muestra facilita la comprensión de cómo realizar ajustes específicos en el tamaño de la muestra para garantizar que el estudio sea lo suficientemente robusto como para identificar efectos reales. Sin embargo, es importante reconocer las limitaciones y supuestos inherentes a la derivación de estas fórmulas, como la distribución normal de los datos y la homogeneidad de las varianzas. La interpretación de los resultados debe realizarse con precaución, teniendo en cuenta estas consideraciones[22].

En la práctica actual, la complejidad de las derivaciones manuales puede superarse mediante el uso de herramientas computacionales. Esto no solo facilita la implementación eficiente de la fórmula, sino que también permite explorar escenarios diversos y aplicar nuevos métodos estadísticos a medida que evoluciona la investigación.[24] Los investigadores quieren usar diseños de estudio más avanzados que permitan detectar efectos utilizando menos muestra. Pero como se explicó, no hay fórmulas matemáticas para calcular cuánta muestra se necesita en estos diseños complejos. Otra estrategia efectiva implica estimar los tamaños de efecto antes del estudio. Calcular la magnitud esperada de la diferencia entre grupos proporciona una guía para determinar cuántos participantes

son necesarios para detectar esa diferencia de manera confiable [20].

Otro punto importante es la valoración de libros tradicionales de metodología, como el de Hernández Sampieri: “Metodología de la investigación”, que aún es un texto muy empleado en nuestro medio, éste puede ser cuestionado en el contexto de la investigación contemporánea debido a diversas razones fundamentales. Estos textos, aunque proporcionan una base sólida en metodología de investigación, a menudo presentan limitaciones significativas que afectan su aplicabilidad y eficacia en el diseño de estudios actuales [22]. Un aspecto crítico que a menudo se trata superficialmente en estos libros es la potenciación estadística.[25] A medida que la complejidad de los análisis estadísticos aumenta, comprender y aplicar adecuadamente la potenciación se ha vuelto esencial. Sin embargo, estos recursos tradicionales pueden no proporcionar la profundidad necesaria en este tema crítico, dejando a los investigadores con una comprensión insuficiente [10],[26].

Hoy en día es esencial un enfoque más integrado que aborde conceptos avanzados como la potenciación en el contexto de diseños de investigación modernos. La colaboración con expertos en estadísticas avanzadas y metodologías contemporáneas puede ser clave para desarrollar recursos educativos más adaptados a las complejidades actuales de la investigación[19].

CONCLUSIONES Y RECOMENDACIONES

En la actualidad las investigaciones cada vez toman mayor valor científico y por ende los métodos para la estimación del tamaño muestral y con ello la potencia estadística deben ser idóneos para reducir la tasa general de errores en los datos, aumentando así los valores de

confiabilidad de este tipo de investigaciones. Adicionalmente, un éxito en la planificación y desarrollo de la metodología contribuye a la replicación de los hallazgos obtenidos para su posterior difusión.

AGRADECIMIENTOS

A los profesores Ebingen Villavicencio-Caparó y Silvio Castellanos; por su colaboración y revisión del artículo.

REFERENCIAS

- [1] E. V. Caparó, K. C. Leon, E. M. V. Leon, J. S. Heredia, and A. C. Duffaut, "Pasos para la planificación de una investigación clínica," *Odontol. Act.*, vol. 1, no. 1, pp. 72–75, Jan. 2016.
- [2] S. Jones, S. Carley, and M. Harrison, "An introduction to power and sample size estimation," *Emerg. Med. J.*, vol. 20, no. 5, p. 453, Sep. 2003.
- [3] S. A. J. Schmidt, S. Lo, and L. M. Holles-tein, "Research Techniques Made Simple: Sample Size Estimation and Power Calculation," *J. Invest. Dermatol.*, vol. 138, no. 8, Aug. 2018, doi: 10.1016/j.jid.2018.06.165.
- [4] L. D. Case and W. T. Ambrosius, "Power and sample size," *Methods Mol. Biol.*, vol. 404, 2007, doi: 10.1007/978-1-59745-530-5_19.
- [5] J. Cohen, "Statistical Power Analysis for the Behavioral Sciences," *The SAGE Encyclopedia of Research Design*, 1969, doi: 10.2307/2529115.
- [6] D. Zurakowski, "Statistical power and sample size calculations: A primer for pediatric surgeons," *J. Pediatr. Surg.*, vol. 55, no. 7, Jul. 2020, doi: 10.1016/j.jpedsurg.2019.05.007.
- [7] J. In et al., "Tips for troublesome sample-size calculation," *Korean J. Anesthesiol.*, vol. 73, no. 2, Apr. 2020, doi: 10.4097/kja.19497.
- [8] Ö. Kemal, "Power Analysis and Sample Size, When and Why?," *Turkish Archives of Otorhinolaryngology*, 2020, Accessed: Dec. 09, 2023. [Online]. Available: <https://pdfs.semanticscholar.org/41fe/af39462564456dd5f6b1a739049469e03934.pdf>
- [9] S. K. Mohanasundari and M. Sonia, "The Relationship between components of sample size Estimate and Sample Size," *Asian Journal of Nursing Education and Research*, 2022, doi: 10.52711/2349-2996.2022.00066.
- [10] C. C. Serdar, M. Cihan, D. Yücel, and M. A. Serdar, "Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies," *Biochem. Med.*, vol. 31, no. 1, pp. 0–0, Feb. 2021.
- [11] E. P. Hong and J. W. Park, "Sample Size and Statistical Power Calculation in Genetic Association Studies," *Genomics Inform.*, vol. 10, no. 2, p. 117, Jun. 2012.
- [12] J. M. A. Pallàs and J. J. Villa, *Métodos de investigación clínica y epidemiológica*. Elsevier Health Sciences, 2019.
- [13] D. Zurakowski, "Statistical power and sample size calculations for time-to-event analysis," *J. Thorac. Cardiovasc. Surg.*, vol. 166, no. 6, Dec. 2023, doi: 10.1016/j.jtcvs.2022.09.023.
- [14] E. V. Caparó, "EL TAMAÑO MUESTRAL PARA LA TESIS. ¿CUÁNTAS PERSONAS DEBO ENCUESTAR?," *Odontol. Act.*, vol. 2, no. 1, pp. 59–62, 2017.

- [15] G. Hickey, S. Grant, J. Dunning, and M. Siepe, “Statistical primer: sample size and power calculations—why, when and how?†,” *Eur. J. Cardiothorac. Surg.*, 2018, Accessed: Dec. 09, 2023. [Online]. Available: <https://pdfs.semanticscholar.org/81db/c78db0fbceae-10d428395845804c15168127.pdf>
- [16] K. M. Sullivan, A. G. Dean, and R. A. Mir, “OpenEpi - Toolkit Shell for Developing New Applications.” <https://www.openepi.com/Power/PowerRCT.htm> (accessed Dec. 19, 2023).
- [17] S. J. Peterson and S. Foley, “Clinician’s Guide to Understanding Effect Size, Alpha Level, Power, and Sample Size,” *Nutr. Clin. Pract.*, vol. 36, no. 3, Jun. 2021, doi: 10.1002/ncp.10674.
- [18] Z. Haile, “Power Analysis and Exploratory Research,” *J. Hum. Lact.*, 2023, doi: 10.1177/08903344231195625.
- [19] H. Kang, “Sample size determination and power analysis using the G*Power software,” *J. Educ. Eval. Health Prof.*, vol. 18, 2021, doi: 10.3352/jeehp.2021.18.17.
- [20] M. Noordzij, F. W. Dekker, C. Zoccali, and K. J. Jager, “Sample size calculations,” *Nephron Clin. Pract.*, vol. 118, no. 4, 2011, doi: 10.1159/000322830.
- [21] K. Fitzner and E. Heckinger, “Sample Size Calculation and Power Analysis: A Quick Review,” *Diabetes Educ.*, 2010, doi: 10.1177/0145721710380791.
- [22] B. N. Gaskill and J. P. Garner, “Power to the People: Power, Negative Results and Sample Size,” *J. Am. Assoc. Lab. Anim. Sci.*, vol. 59, no. 1, p. 9, Jan. 2020.
- [23] C. Y. Peng, H. Long, and S. Abaci, “Power Analysis Software for Educational Researchers,” *J. Exp. Educ.*, 2012, doi: 10.1080/00220973.2011.647115.
- [24] E. McCrum-Gardner, “Sample size and power calculations made simple,” *Int. J. Ther. Rehabil.*, 2010, doi: 10.12968/IJTR.2010.17.1.45988.
- [25] R. H. Sampieri, *METODOLOGÍA DE LA INVESTIGACIÓN: LAS RUTAS CUANTITATIVA, CUALITATIVA Y MIXTA*. McGraw-Hill Interamericana, 2018.
- [26] G. Shieh, “Effect size, statistical power, and sample size for assessing interactions between categorical and continuous variables,” *Br. J. Math. Stat. Psychol.*, 2018, doi: 10.1111/bmsp.12147.

Recibido: 22/12/2023

Aceptado: 30/12/2023

