

# Cálculo de la potencia estadística en estudios de asociación mediante sintaxis en SPSS

Javier Santabárbara-Serrano<sup>1</sup> 

Artículo metodológico. Recibido: 14/10/2020. Aceptado: 26/01/2021. Publicación avanzada: 26/02/2021. Publicado: 01/07/2021

## Resumen

**INTRODUCCIÓN.** El estudio de la asociación entre dos variables se expresa de forma analítica con el estadístico denominado coeficiente de correlación. Con frecuencia, encontramos asociaciones de magnitud relevante pero que no alcanzan la significación estadística, planteándonos entonces si el contraste de hipótesis tiene la suficiente potencia estadística para detectar diferencias cuando realmente existen. Desafortunadamente, SPSS no ofrece el cálculo de la potencia estadística para los estudios de asociación de variables cuantitativas. Por ello, se plantea el cálculo de la potencia estadística mediante sintaxis de SPSS en el escenario de estudios de asociación a través de un caso práctico.

**MÉTODO.** Se proporciona un fichero de sintaxis de SPSS para calcular la potencia estadística en estudios de asociación.

**DISCUSIÓN.** El uso del fichero de sintaxis será de utilidad para aquellos investigadores familiarizados con SPSS que necesiten realizar el cálculo de la potencia estadística en estudios de asociación.

## Palabras clave

Correlación; Potencia estadística; Sintaxis SPSS..

## Datos del autor

<sup>1</sup> Universidad de Zaragoza, España. Departamento de Microbiología, Pediatría, Radiología y Salud Pública. Facultad de Medicina. Contacto para la correspondencia: [jsantabarbara@unizar.es](mailto:jsantabarbara@unizar.es)

## Referencia recomendada

Santabárbara-Serrano, J. (2021). Cálculo del tamaño de muestra necesario para estimar el coeficiente de correlación de Pearson mediante sintaxis en SPSS. *REIRE Revista d'Innovació i Recerca en Educació*, 14(2), 1–6. <https://doi.org/10.1344/reire2021.14.232580>

© 2021 El autor. Este artículo es de acceso abierto sujeto a la licencia Reconocimiento 4.0 Internacional de Creative Commons, la cual permite utilizar, distribuir y reproducir por cualquier medio sin restricciones siempre que se cite adecuadamente la obra original. Para ver una copia de esta licencia, visite <https://creativecommons.org/licenses/by/4.0/>



#### Títol (català)

Càlcul de la potència estadística en estudis d'associació mitjançant sintaxi en SPSS

#### Resum

**INTRODUCCIÓ.** L'estudi de l'associació entre dues variables s'expressa de forma analítica amb l'estadístic anomenat *coeficient de correlació*. Sovint trobem associacions de magnitud rellevant però que no arriben a la significació estadística, i aleshores ens plantejem si el contrast d'hipòtesis té prou potència estadística per detectar diferències quan realment existeixen. Desafortunadament, SPSS no ofereix el càlcul de la potència estadística per als estudis d'associació de variables quantitatives. Per aquest motiu, es planteja el càlcul de la potència estadística mitjançant sintaxi de SPSS en l'escenari d'estudis d'associació a través d'un cas pràctic..

**MÈTODE.** Es proporciona un fitxer de sintaxi de SPSS per calcular la potència estadística en estudis d'associació.

**DISCUSSIÓ.** L'ús del fitxer de sintaxi serà d'utilitat per als investigadors familiaritzats amb SPSS que necessitin fer el càlcul de la potència estadística en estudis d'associació.

#### Paraules clau

Correlació; Potència estadística; Sintaxi SPSS.

#### Title (English)

Calculation of statistical power in association studies using SPSS syntax

#### Abstract

**INTRODUCTION.** The study of the association between two variables is expressed analytically with the statistic known as the correlation coefficient. Frequently, we find associations that are relevant but do not reach statistical significance. We therefore wonder whether the hypothesis test has sufficient statistical power to detect differences when they actually exist. Unfortunately, SPSS does not calculate the power for association studies of quantitative variables. For this reason, we propose the use of SPSS syntax to calculate statistical power in the scenario of association studies, and present a practical case.

**METHODS.** An SPSS syntax file is provided to calculate statistical power in association studies.

**DISCUSSION.** The use of the syntax file will be useful for researchers familiar with SPSS who need to calculate statistical power in association studies.

#### Keywords

Correlation; Statistical power; SPSS syntax.

## 1. Introducción: importancia del cálculo de la potencia estadística

El presente trabajo se focaliza en los estudios de asociación entre variables cuantitativas. El lector interesado en la aplicación de esta técnica estadística puede consultar Reguant-Álvarez et al. (2018), así como Santabárbara (2019). A modo de recordatorio, en los estudios de asociación, se pone a prueba la hipótesis nula ( $H_0$ ) de no relación entre las variables (coeficiente de correlación poblacional igual a cero:  $\rho = 0$ ) frente a la alternativa de existencia de asociación ( $H_1: \rho \neq 0$ ). Así, en el análisis del contraste de hipótesis, la atención principal está puesta en  $\alpha$ , la probabilidad de cometer un error de tipo I –rechazar la hipótesis nula cuando es verdadera–. Lamentablemente, se le ha prestado poca atención a  $\beta$ , la probabilidad de cometer un error de tipo II –no rechazar la hipótesis nula cuando es falsa–. Hay una razón que justifica esta diferencia: para un contraste dado,  $\alpha$  es un solo número asignado por el investigador antes de realizar el contraste –es una medida del riesgo aceptable de rechazar una hipótesis nula verdadera–; por otra parte,  $\beta$  puede adoptar uno de muchos valores. Supóngase que se pretende probar la hipótesis nula de que un parámetro de alguna población (p. e.,  $\rho$ ) es igual a algún valor específico (p. e., 0). Si  $H_0$  ( $\rho = 0$ ) es falsa y no se rechaza, se comete el error de tipo II. Si el valor supuesto del parámetro no es el valor correcto, el valor de  $\beta$  (la probabilidad de cometer un error de tipo II) depende de los siguientes factores: 1) valor correcto del parámetro de interés, 2) valor supuesto del parámetro, 3) valor de  $\alpha$ , y 4) tamaño de la muestra ( $n$ ) (Baldi y Moore, 2013; Daniel y Cross, 2013).

Para un contraste de hipótesis dado, es interesante saber qué tanto control se tiene sobre el error de tipo II. Si  $H_0$  es falsa, ¿cuál es la probabilidad de rechazarla? La información para contestar esta pregunta radica en la *potencia* del contraste, designada como  $1 - \beta$ . La cantidad  $1 - \beta$  es la probabilidad de rechazar la hipótesis nula cuando es falsa; esta puede calcularse para cualquier valor del parámetro respecto al cual se prueba la hipótesis (p. e.,  $\rho$ ). Por lo tanto,  $1 - \beta$  es la probabilidad de tomar la acción correcta cuando  $H_0$  es falsa (Baldi y Moore, 2013; Daniel y Cross, 2013). Habitualmente, se fija la potencia estadística en el 80 % ( $1 - \beta = 0,80$ ). Esto es, asumimos una probabilidad del 20 % de no rechazar la hipótesis nula cuando esta es en realidad falsa ( $\beta = 0,20$ ) (Glantz, 2006).

Supongamos que al analizar nuestros datos no existe suficiente evidencia muestral para rechazar la hipótesis nula de asociación entre dos variables ( $H_0: \rho = 0$ ), dado que  $p \geq \alpha$ . En este escenario, si el cálculo de la potencia arroja un valor superior a 0,80, esto implicaría un riesgo pequeño de incurrir en un error tipo II y, por tanto, estaríamos prácticamente seguros de la decisión tomada –no rechazar la hipótesis de ausencia de correlación entre las variables–. Ahora bien, si obtuviéramos un valor inferior a 0,80 para la potencia implicaría ampliar excesivamente la muestra, haciéndonos dudar de la decisión tomada. Siendo, por tanto, la potencia estadística un índice de la validez de nuestros resultados estadísticos, de ahí su relevancia en investigación (Triola, 2013). Máxime cuando se ha reportado que más de la mitad de los trabajos científicos concluyen erróneamente la no existencia de asociación cuando ésta no alcanza la significación estadística (Amrhein et al., 2019).

## 2. Justificación y objetivos del presente trabajo

Lamentablemente, SPSS no ofrece el cálculo de la potencia estadística para el coeficiente de correlación de Pearson. Luego, se plantea como objetivo principal proporcionar mediante sintaxis de SPSS la estimación el cálculo de la potencia estadística en un estudio de asociación, en el que se plantea el uso del coeficiente de correlación de Pearson, a través de un caso práctico que complementa el trabajo de Santabárbara et al. (2019) sobre la asociación de dos variables cuantitativas.

### 3. Potencia estadística para el coeficiente de correlación lineal de Pearson

#### 3.1. Definición y expresión matemática

Si en la práctica nos encontramos un coeficiente de correlación no significativo, cabe preguntarse por la potencia del contraste para detectar que el coeficiente de correlación poblacional es distinto de cero. Así, conocido el número de sujetos de la muestra, el valor del coeficiente de correlación que consideraremos de relevancia científica ( $\rho$ ) y fijada la probabilidad de cometer un error de tipo I ( $\alpha$ ), siendo el contraste bilateral, obtenemos la siguiente expresión para la potencia (Gerstman, 2013; Glantz, 2006):

$$1 - \beta = \Phi \left( Z < -z_{1-\alpha/2} + \frac{\frac{1}{2} \ln \left( \frac{1+|\rho|}{1-|\rho|} \right)}{\sqrt{\frac{1}{n-3}}} \right)$$

#### 3.2. Fichero de sintaxis de SPSS

Se proporciona el archivo de sintaxis nominado *potencia\_correlacion.sps* para tal objetivo –basado en la expresión matemática original de Glantz (2006) y Gerstman (2013)–, cuyo contenido se muestra en la Figura 1.

##### Figura 1

*Archivo de sintaxis potencia\_correlacion.sps*

---

```
* Introducimos el coeficiente de correlación (r),
el nivel de significación (sig) y el tamaño de la muestra (n).

data list free / r sig n.
begin data.
0,20 0,05 41
end data.

* Transformamos r en el valor z: fz.
compute fz = 0.5*ln((1+abs(r))/(1-abs(r))).

* Calculamos el valor crítico de la curva normal para el nivel de
significación: critz_sig.
compute critz_sig = abs(idf.normal(sig/2,0,1)).

* Calculamos la potencia (potencia).
compute potencia = cdf.normal(-critz_sig + (fz/sqrt(1/(n-3))), 0, 1).

formats r (f10.2) sig potencia (f10.2) / n (f8).

* Mostramos la potencia.
list r sig n potencia.
```

---

### 3.3. Caso práctico

Para la presentación del cálculo de la potencia estadística para el coeficiente de correlación de Pearson, seguiremos el ejemplo de Santabárbara et al. (2019) en el que se evaluó la existencia de una relación entre la actitud hacia la estadística de los estudiantes de un Máster en Medicina y su calificación final en Bioestadística. Sobre una muestra de 41 estudiantes, obtuvieron un coeficiente de correlación de 0,11 entre el “valor” que estos estudiantes otorgaban a la estadística en su desarrollo profesional y la calificación en Bioestadística, no alcanzando dicha asociación la significación estadística ( $p = 0,244$ ). Luego, suponiendo que un coeficiente de correlación de 0,2 o mayor podría considerarse como relevante científicamente, cabe preguntarse, ¿qué probabilidad tendría el contraste de hipótesis para detectar que existe asociación entre ambas variables (*potencia*:  $1 - \beta$ ) fijado el nivel de significación ( $\alpha$ ) en el 5 %?

En el fichero de sintaxis se introducirá el valor del *coeficiente de correlación de Pearson de relevancia científica* (0,20), el nivel de significación (en tanto por uno: 0,05) y el tamaño de la muestra (41), según se disponen en la Figura 2 (izda., ver flecha). A continuación, para seleccionar el texto del archivo de sintaxis, elegiremos el procedimiento *Editar* y, dentro de él, *Seleccionar todo* (Figura 2, dcha.):

**Figura 2**

*Archivo de sintaxis potencia\_correlacion.sps*



Una vez seleccionada la totalidad del texto, se pulsa el botón *[Play]* (recuadro rojo de la Figura 2, izda.) para ejecutar el programa e inmediatamente aparecerá en el *Visor de resultados* el intervalo de confianza buscado (Tabla 1):

**Tabla 1**

*Potencia estadística para el coeficiente de correlación lineal de Pearson*

$r$	$sig$	$n$	<i>potencia</i>
,20	,05	,41	,24

Luego, únicamente se dispone de una probabilidad de 0,24 (notablemente inferior a 0,80) de detectar que la correlación es estadísticamente significativa con nuestros datos muestrales del ejemplo (Santabàrbara et al., 2019). Nos planteamos, por tanto, la limitación de disponer de un insuficiente tamaño muestral para llevar a cabo dicho estudio.

## 4. Conclusión

El fichero de sintaxis de SPSS proporcionado en este trabajo (disponible en material suplementario) supone una ayuda metodológica a los investigadores durante la fase de análisis de datos cuando éste concluya que el coeficiente de correlación calculado no alcance la significación estadística y deba encontrarse una explicación a este hallazgo.

## Archivo complementario

Sintaxis SPSS para el cálculo de la potencia estadística para el coeficiente de correlación de Pearson: <https://revistes.ub.edu/index.php/REIRE/rt/suppFiles/32580/0>

## Referencias

- Amrhein, V., Greenland, S., y McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, 567, 305–307. <https://doi.org/10.1038/d41586-019-00857-9>
- Baldi, B., y Moore, D. S. (2013). *The practice of statistics in the life sciences* (3.ª ed.). McMillan Learning.
- Daniel, W. W., y Cross, C. L. (2013). *Biostatistics. A foundation for analysis in the health sciences* (10.ª ed.). John Wiley & Sons.
- Gerstman, B. (2015). *Basic Biostatistics: Statistics for Public Health Practice* (2.ª ed.). Jones & Bartlett.
- Glantz, A. G., (2006). *Bioestadística* (6.ª ed.). McGraw-Hill Interamericana.
- Reguant-Álvarez, M., Vilà-Baños, R., y Torrado-Fonseca, M. (2018). La relación entre dos variables según la escala de medición con SPSS. *REIRE Revista d'Innovació i Recerca en Educació*, 11(2), 45–60. <https://doi.org/10.1344/reire2018.11.221733>
- Santabàrbara, J. (2019). Cálculo del intervalo de confianza para los coeficientes de correlación mediante sintaxis en SPSS. *REIRE Revista d'Innovació i Recerca en Educació*, 12(2), 1–14. <https://doi.org/10.1344/reire2019.12.228245>
- Santabàrbara, J., Montenegro, S., y López-Antón, R. (2019). Conocimientos básicos en bioestadística y epidemiología y actitudes hacia la estadística en estudiantes de posgrado de medicina. *FEM Revista de la Fundació de Educació Mèdica*, 22(3), 145–152. <https://doi.org/10.33588/fem.223.993>
- Triola, M. F. (2013). *Estadística* (11.ª ed.). Pearson Educación.