Revista Latinoamericana de Metodología de la Investigación Social.
N°26. Año 13. Octubre 2023 - Marzo 2024. Argentina. ISSN 1853-6190. Pp. 9-30.

La lógica de los coeficientes

The logic of coefficients

Hugo Darío Echevarría

Resumen

En este artículo, muestro que los coeficientes de asociación y de confiabilidad, se basan en los mismos principios y, basándome en ellos, propongo algunos coeficientes alternativos al *Q* de Yule, al Alfa de Cronbach (*Alfa*) y al Coeficiente N° 20 de Kuder y Richardson (*KR20*). En primer lugar, me refiero a un modo intuitivo de analizar tablas de contingencia de 2 x 2. En segundo lugar, introduzco dos coeficientes de asociación y los comparo con el coeficiente *Q* de Yule.En tercer lugar, planteo algunos coeficientes de confiabilidad (*CC1 y CC1.2*), más simples que *Alfa* y *KR20,* mostrando que *CC1* no depende del número de ítems del instrumento de recolección de datos, lo que si sucede con *Alfa*. Este con algunos datos adoptó valores negativos, lo que no sucede con los *CC* y el análisis mostró que reflejan la confiabilidad de los datos considerados. *Alfa* en una tabla con datos muy inconsistentes registró un valor alto, en cambio los *CC* resultaron bajos, sobre todo *CC1*, que en general, se mostró más adecuado que *Alfa*.

Palabras clave: Coeficientes de asociación; *Q* de Yule; confiabilidad; coeficiente *Alfa* de Cronbach; coeficiente N° 20 de Kuder y Richardson.

Abstract

In this article, I show that the association and reliability coefficients are based on the same principles, although their formulas may seem very different, and based on them, I propose some alternative coefficients to Yule's *Q*, Cronbach's Alpha (*Alpha*) and Kuder and Richardson's Coefficient (*KR20*). First, I introduce an intuitive way of analyzing 2 x 2 contingency tables. Second, I introduce two association coefficients and compare them with the Yule's *Q* coefficient; one for nominal scales and the other for ordinal scales. Thirdly, I propose some reliability coefficients (CC), based on the internal consistency approach, that have simpler formulas than those commonly used, and I compare them with Alpha and KR20 coefficients. Two of them turned out to be appropriate: CC1 and CC1.2 for interval and dichotomous scales respectively. Besides, CC1 does not depend on the number of items of the data collection instrument, which does happen with Alpha. Coefficient Alpha with some data adopted negative values due to a negative average covariance between items, which does not happen with the CC, that is, they have the advantage of not requiring this assumption; and the analysis showed that they reflect the reliability of the data considered. Coefficient Alpha in a table with very inconsistent data showed a high value, whereas CC were low, especially CC1, which in general was more appropriate than *Alpha*.

Keywords: Association coefficients; reliability; Yule's Q coefficient; Cronbach coefficient Alpha; coefficient N° 20 of Kuder and Richardson.

1. Introducción[1]

Un aspecto importante de muchas investigaciones cuantitativas es el análisis de las relaciones entre variables y, según Mayntz, et al. (1975) existen tres grandes formas de hacerlo: el cruzamiento, el coeficiente de correlación y la prueba de significación o prueba de hipótesis. El *cruzamiento* representa una forma imprecisa (aunque a veces sumamente útil), en tanto que el coeficiente da como resultado una magnitud exacta, es decir, puede calcularse con la cantidad de decimales que se desee. No obstante, por un lado, existe una gran cantidad de coeficientes, y para su elección se deben tener en cuenta, fundamentalmente, las variables que se están considerando, la distribución empírica hallada y los objetivos del investigador. Pero, por otro lado, sus ecuaciones como norma general son difíciles de entender, lo que supone a su vez, que es complicado analizar cuáles responden mejor a la finalidad para las que se han creado. Es frecuente que se presenten complejas argumentaciones, que a veces, lejos de resolver los problemas acentúan la falta de acuerdo; si bien son sumamente necesarias, pues permiten avanzar en las diferencias existentes entre las formas de cálculo propuestas (véase por ejemplo, Cronbach, 2004; Dominguez-Lara, 2016).[2] Por esto, en este trabajo introduzco algunas ecuaciones más simples para algunos coeficientes y, al mismo tiempo, con análisis sencillos, esto es, que sólo requieren conocimientos elementales de matemática, muestro que efectivamente pueden cumplir con la finalidad para la que fueron concebidos. En primer lugar propongo un coeficiente creado para ser usado con escalas nominales, en segundo lugar, una alternativa al coeficiente *Q* de Yule y, puesto que la confiabilidad es en última instancia "una especie de correlación con un rango posible de 0 a 1,00"[3] (Cronbach, 2004: 394), presento algunas alternativas al coeficiente *Alfa* de Cronbach, distintas a las que se han planteado, que también se destacan por su simplicidad y por lo tanto, por la facilidad para entender el objetivo al que responden.

En la primera parte de este artículo, trato las dos primeras formas que consideran Mayntz, et al. (1975), con ejemplos de coeficientes de asociación. Me interesa sobre todo mostrar como se construyen los coeficientes, por ello, considero el coeficiente *Q* de Yule junto a otros que son más simples aún (*E1* y *E2*). Además, los comparo con el cruzamiento, un modo intuitivo de indagar la relación entre variables. En otro lugar (Echevarría, 2019a y b) hice algunas consideraciones sobre la forma de interpretar el coeficiente de correlación y la relación que existe entre este y la prueba de significación[4]. Aquí me refiero exclusivamente a la estadística descriptiva, es decir, solamente a estadísticos orientados a reducir los datos de una muestra (o de una población si se toman todas las unidades de análisis que la integran), como así también a procedimientos para analizar una de las propiedades psicométricas -la confiabilidad- en relación a los datos recolectados en una situación específica. Esto último, lo hago en la segunda parte, aunque limitándome al enfoque de la consistencia interna. Introduzco algunos coeficientes mostrando como, en última instancia, si bien tienen fórmulas diferentes, sus fundamentos son compartidos con los coeficientes de asociación.

---

[1] Versión ampliada y corregida de Análisis de la asociación entre variables categóricas (2017). En Michelini, D., Pérez Zabala, G., y Galetto, N. (Editores). Violencia: problemas y abordajes. XXII Jornadas Internacionales Interdisciplinarias. Río Cuarto: Ediciones del ICALA y de Echevarría, H. La confiabilidad con el enfoque de la consistencia interna. Trabajo presentado en la XXV Jornadas Interdisciplinarias de la Fundación ICALA (Intercambio cultural alemán-latinoamericano). Conflictos sociales y convivencia democrática. Río Cuarto, 29 y 30 de octubre de 2020. Río Cuarto. Agradezco a María Inés Valsecchi su generosa ayuda en la traducción de algunos textos de inglés.

[2] Después de presentar todos los argumentos en favor del Coeficiente *H*, Dominguez-Lara (2016: 91) sostiene que, similar a lo que ocurre con *Alfa* "parece que su magnitud está influida por la cantidad de ítems que existen (...) No obstante, son necesarios estudios de simulación que reflejen diferentes condiciones experimentales para reafirmar este argumento, y desarrollos posteriores que permitan lograr un ajuste apropiado de su magnitud en función del número de ítems a fin de no sobredimensionar su valor".

[3] "A kind of correlation with a possible range from 0 to 1.00" (traducción propia).

[4] Cuando las variables son categóricas suelen mencionarse como coeficientes de asociación, y cuando son numéricas, como coeficientes de correlación. Lo planteado aquí vale tanto para ambos, aunque considero sólo coeficientes a asociación.

2. El cruzamiento

Comenzaré partiendo de un ejemplo real,[5] en el que se indagó el grado en que la *comprensión lectora* de textos de lógica y matemática, se relaciona al *rendimiento* de los alumnos de primer año del nivel universitario, tomando como indicador del último las notas del primer parcial (Echevarría, 2008). El Cuadro 1 contiene en filas la variable *comprensión lectora de matemáticas*, medida en este caso con la que, en ese trabajo, se llamó prueba N° 2. Se crearon dos grupos, uno integrado por todos los alumnos cuyo puntaje en esa prueba resultó menor o igual a 8 (la mediana) y otro, con los alumnos que superaron esa cantidad.

En columnas se incluyó el resultado del parcial, y las categorías fueron menor a 4 (no lo aprobaron, *nota baja*) y mayor o igual a 4 (los que aprobaron, *nota alta*). La definición de lo que es una *nota alta* puede ser cuestionable, pero útil a los efectos de explicar los conceptos que trato aquí. Además, los porcentajes se calcularon en el sentido de la variable independiente, aplicando la "*regla de causa y efecto*" (Zeizel, 1974: 37).

Cuadro 1. Cantidad de alumnos según comprensión lectora y primer parcial de Matemática

	Primer parcial de Matemática (variable dependiente)		
Comprensión lectora (variable independiente)	Menor a 4 (notas bajas)	Mayor o igual a 4 (notas altas)	Total
Menor o igual que la mediana (baja)	61 (a) 45,2%	74 (b) 54,8%	135,00 100,00%
Mayor que la mediana (alta)	5 (c) 6,4%	73 (d) 93,6%	78,00 100,00%

Fuente: Elaboración propia con datos recolectados por el autor junto a otros proporcionados por docentes de la FCE, UNRC (Facultad de Ciencias Económicas, Universidad Nacional de Río Cuarto, ingresantes año 1999).

Ahora bien, dada la hipótesis *a mayor comprensión, es más probable observar un mayor rendimiento,* ¿qué resultados permiten sostenerla? Nótese que contiene dos variables, y tal vez el lector advierta que, según ella, deben observarse más casos en las celdas *a* (*menor comprensión* y *notas bajas,* Cuadro 1 y Cuadro 2) y *d* (*mayor comprensión* y *notas altas*). También es esperable que sea bajo el número de alumnos en las celdas *b* (*menor comprensión* y *notas altas*) y *c* (*mayor comprensión* y *notas bajas*). Se puede decir que los primeros (celdas *a* y *d*) son los *casos corroboratorios* de esta hipótesis, esto es, lo que debiera observarse si fuera plausible, o lo que es lo mismo, quienes *comprenden mejor* deberían obtener *notas altas* y viceversa.

Pero respecto del Cuadro 1, estos datos, ¿corroboran o refutan la hipótesis? Se ve, por ejemplo, que la celda *d* cuenta con un mayor porcentaje de casos que la *b* (93,6% y 54,8% respectivamente), lo que está a favor de la hipótesis planteada. Recíprocamente, la celda *a*, cuenta con un porcentaje de casos superior a la celda *c* (45,2% y 6,4% respectivamente).

En el Cuadro 1 se dio un cierto *cruzamiento* entre los porcentajes altos de las celdas *a* y *d*, y los bajos de las celdas *c* y *b*, esto es, entre los casos *corroboratorios* (los primeros) y los *refutatorios* (los últimos), lo que estaría indicando que la hipótesis se ha corroborado. No obstante, los casos refutatorios no están ausentes, entonces, ¿se debe aceptar o no la hipótesis? Precisamente, el *cruzamiento* no da una cifra exacta: un investigador con estos datos podría concluir que la hipótesis se corroboró, mientras que otro más exigente, podría considerar que, dado el número de casos refutatorios que de todos modos se registraron, se dio una refutación. Todas las formas de decidir presentan limitaciones, el *cruzamiento* tiene la dificultad de no ofrecer un número preciso que indique la fuerza de la relación estudiada, lo que se logra con los coeficientes de asociación y de correlación.

[5] Con algunos datos perdidos, se ingresaron datos ficticios para que el ejemplo refleje más claramente lo que quiero mostrar aquí.

Cuadro 2. Combinaciones posibles de las categorías de las variables comprensión lectora y notas del parcial de matemáticas

	Notas del parcial	
Comprensión lectora	Menor a 4	Mayor o igual a 4
Menor o igual que la mediana (menor comprensión)	Celda a: alumnos de menor comprensión y notas bajas	Celda b: alumnos de menor comprensión y notas más altas
Mayor a la mediana (mayor comprensión)	Celda c: alumnos de mayor comprensión y notas bajas	Celda d: alumnos de mayor comprensión y notas más altas

Fuente: Elaboración propia.

3. Coeficientes de asociación

Los coeficientes, por ejemplo, los de correlación o asociación y los de confiabilidad, se basan en una función matemática muy simple:

$$y = x_1/x_2.$$

Es decir, en una función que consiste en dividir una variable sobre otra. Un rápido análisis de la misma permite advertir que:

> Si se mantiene constante x_2, a medida que aumenta x_1, también lo hace y.
>
> Si se mantiene constante x_1, a medida que aumenta x_2, disminuye y. O lo que es lo mismo, a medida que disminuye x_2, aumenta y.

En otros términos, y varía en función directa de x_1 y en función inversa de x_2. Por ejemplo, en la Figura 1, tomando sólo valores positivos para simplificar, puede verse la función:

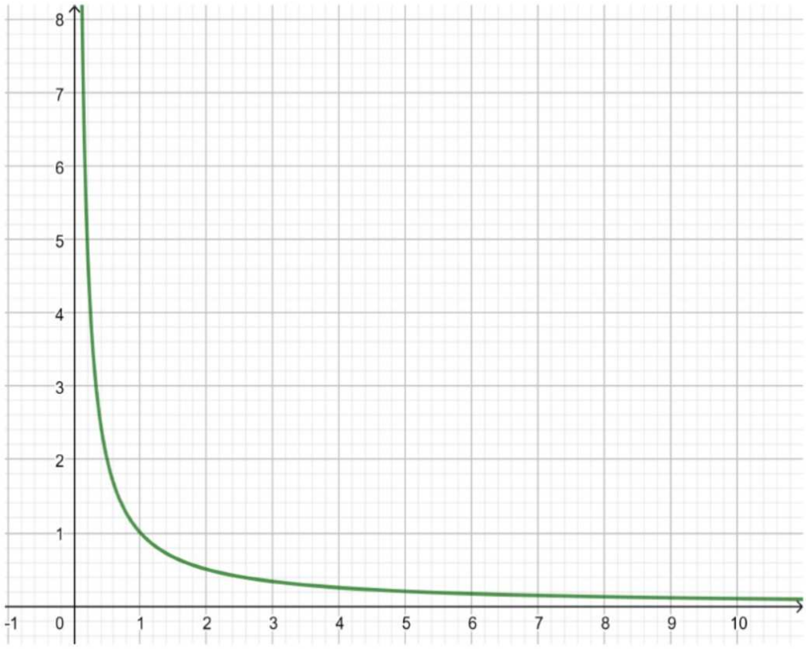
$$y = 1/x.$$

Se advierte que a medida que x se aproxima a *1*, también lo hace y; a medida que x aumenta, más pequeño se torna y, es decir, a medida que aumenta x, la función tiende a *0*. Recíprocamente, cuando más pequeño es x (x más se aproxima a cero), mayor es y.

Como sostuve más arriba, otra forma de analizar la relación entre las variables consiste en calcular un *coeficiente de asociación o correlación*. Existe una gran cantidad de ellos, el que se seleccione depende fundamentalmente del tipo de variables que se están considerando, de la distribución empírica observada, de los objetivos del estudio, e incluso, de los que habitualmente se utilizan en el área de investigación considerada (esto último facilita comparar los resultados hallados con los de otras investigaciones).

Acá trato el caso más simple, es decir, con dos variables categóricas dicotómicas. Uno de ellos es el coeficiente *Q de Yule*, que para algunos se debe aplicar sólo si ambas variables son ordinales (por ejemplo, Mayntz et al., 1975 y Marradi, 2007). Briones (2002) lo considera pertinente cuando las variables son nominales, mientras que para Cazau “se limita a dos variables nominales, con dos categorías cada una” (2006: 61) e igual idea sostiene De la Fuente (2011).

Figura 1. Gráfico de la función matemática *y=1/x*



Fuente: Elaboración propia.

Tanto para la abscisa (coordenada horizontal o eje *x*), como para la ordenada (coordenada vertical o eje *y*) se incluyó sólo la parte positiva, por ser la que interesa aquí.

Pero antes de tratarlo, presento un coeficiente muy simple *que lo propongo como un modo de explicar este concepto*. El mismo consiste simplemente en *dividir los casos corroboratorios sobre el total de casos*. Por razones de comodidad y dada la función educativa con que en un primer momento lo concebí, lo llamé *E1.*

El principio básico de los coeficientes de asociación (y de correlación) es cuantificar de algún modo la variación conjunta de ambas variables y ponerla en relación a la variación total posible, pero, como dije más arriba, no hay una única forma de hacerlo.

Para el ejemplo que estaba analizando (Cuadro 1), *E1* se calcula del siguiente modo (las letras representan la cantidad de casos de las celdas de un cuadro de dos variables con dos categorías cada una, recordar el Cuadro 2):

$$E1 = a+d/a+b+c+d$$

$$E1 = 61+73/(61+74+5+73) = 0.63$$[6]

La idea de este coeficiente es lograr una cifra que nos indique el grado de asociación entre las variables, en una escala que va de *0* a *1*. Pero, ¿se comporta como es deseable?, esto es, ¿es capaz de diferenciar entre situaciones en que, cuando la relación entre las variables es débil, el coeficiente da un número bajo (más cerca de *0*), cuando es media, un número cercano a *0,50* y si es alta, se aproxima a *1*? (este es el principio general de los coeficientes para escalas nominales, aunque no todos varían entre *0* y *1*).

[6] Si lo multiplicamos por 100, obtenemos el porcentaje de *casos corroboratorios*, pero no es esta la finalidad de los coeficientes.

4. Asociación entre variables nominales

Por otro lado, ¿*E1* puede ser aplicado a todas las hipótesis que relacionan variables categóricas? Siendo *V1* y *V2* dos variables categóricas cualesquiera, algunas hipótesis pueden ser de las formas siguientes:

V1 se relaciona con *V2*,

Cuanto mayor sea *V1* más probable es hallar valores altos en *V2*,

Cuanto mayor sea *V1* más probable es hallar valores bajos en *V2*.

La primera es típica de las variables nominales, y simplemente postula una relación entre variables, pero no dice nada de su dirección, mientras que las dos últimas suponen que ambas son al menos ordinales. Podría pensarse que el coeficiente *E1* no sería el más adecuado para hipótesis del segundo y tercer tipo. Para analizarlo presento un ejemplo (Cuadro 3). Se trata datos de una muestra de ingresantes a la FCE de la UNRC (año 1999) y se vinculan las variables *género* y *residencia*. Puede verse que ambas son dicotómicas y que existen diferencias en el tipo de residencia que eligieron los estudiantes según el género: mientras que las mujeres se inclinaron hacia una residencia familiar en un 70,6%, un 66,4% de los varones eligió esa modalidad. Pero, ¿qué tan fuerte es la relación entre las variables? (si es que puede considerarse que esta existe).

A la derecha del Cuadro 3 pueden verse tres coeficientes que dan una respuesta a la pregunta y que, dado que ofrecen valores muy distintos, generan un nuevo interrogante: ¿cuál debe considerarse?

Pero antes de continuar es necesario explicar los otros dos coeficientes. *Q* es el conocido *Q* de Yule (mencionado en un sinnúmero de publicaciones, lo tomé de Marradi, 2007); a *E2* lo he elaborado, originalmente, con una función didáctica y es una versión simplificada del *Q* de Yule:

$$Q = (a * d) - (b * c) / (a * d) + (b * c)$$

$$E2 = (a + d) - (b + c) / (a + d + b + c).$$

Cuadro 3. Cantidad de casos según género y tipo de residencia, y coeficientes de asociación

	Residencia		
Género	Familiar	Independiente	Total
Femenino	115 70,6 %	48 29,4 %	163 100,0%
Masculino	79 66,4 %	40 33,6 %	119 100,0%

E1 = 0,55
E2 = 0,10
Q = 0,10

Fuente: Elaboración propia con datos proporcionados por el Centro de Cómputos de la FCE, UNRC (ingresantes año 1999).

Q se basa en multiplicar los casos corroboratorios (celdas *a* y *d* en un cuadro de 2 x 2, recordar los Cuadros 1 y 2), restarle los casos refutatorios también multiplicados entre sí (celdas *b* y *c*), y luego dividir esta cifra por la suma de ambas multiplicaciones. *E2* como dije es una variación del mismo. A los casos corroboratorios se les restan los refutatorios y a este resultado se lo divide por el número total de casos. En definitiva, conceptualmente, es así:

E2 = (casos corroboratorios – casos refutatorios) / total de casos.

Total de casos, también puede considerarse como:

Total de casos = casos corroboratorios + casos refutatorios,

por lo tanto, *E2* también puede pensarse como:

E2= (casos corroboratorios–casos refutatorios)/(casos corroboratorios+casos refutatorios).

El Cuadro 4 contiene las mismas variables que el Cuadro 3, pero con datos ficticios que indican una fuerte relación entre ellas. Se observa que algo más del 70% de las *mujeres* prefieren una residencia de tipo *familiar*, y en los *varones* esta cifra no llega al 23%. Recíprocamente, la residencia *independiente*, es preferida en un 29,5% por las *mujeres*, y en un 77,6% por los *varones*. Ante esto, es lógicamente esperable que los coeficientes sean mayores que los del Cuadro 3, y se observa que esto sucede con los tres coeficientes (ver la última columna del Cuadro 4).

Cuadro 4. Cantidad de casos según género y tipo de residencia y coeficientes de asociación

	Residencia		
	Familiar	Independiente	Total
Femenino	115 70,5%	48 29,5%	163 100%
Masculino	35 22,4%	121 77,6%	156 100%

E1 = 0,74
E2 = 0,48
Q = 0,78

Fuente: Elaboración propia en base a datos ficticios.

No obstante, debe advertirse que *E1* tiene una diferencia con los anteriores: un valor próximo a *0* indica una asociación entre las variables de sentido opuesto al postulado; por ejemplo, con los datos del Cuadro 6, se obtiene un coeficiente de *0. Es interpretable*, pues según se aproxima a *0,* a *0.50,* o *1* nos indica una asociación inversa, inexistente o positiva. No obstante, *tiene la dificultad de que su escala es diferente a las consideradas por los coeficientes, por lo que su uso podría generar confusión*, si bien puede ser de utilidad para ayudar a pensar sobre el modo que estos se construyen e interpretan.

Si las *mujeres* hubieran preferido en mayor medida la residencia *independiente* y los varones una *familiar*, podríamos tener datos como los del Cuadro 5, esto es, coherentes con esta situación. Se nota algo diferente a los cuadros anteriores: *E2* y *Q* dan resultados negativos, mientras que *E1*, da un valor bajo, pero positivo. Comienzo por lo que tal vez sea más simple: *E1* varía entre *0* y *1*, ya que en definitiva es una ecuación de la forma $E1=x_1/x_2$, donde x_1 son todos los casos corroboratorios y x_2 el total de ellos. Obviamente, no puede haber mayor número de casos corroboratorios que el total, por lo tanto, variará de 0 a 1 y será siempre positivo (tanto el numerador como el denominador no pueden ser negativos).

Cuadro 5. Cantidad de casos según género y tipo de residencia y coeficientes de asociación

	Residencia		
	Familiar	Independiente	Total
Femenino	35 23.3%	115 76.7%	150 100%
Masculino	121 71,6%	48 28,4%	169 100%

E1 = 0.26
E2 = -0.48
Q = -0.78

Fuente: Elaboración propia en base a datos ficticios.

En el ejemplo que acabo de mostrar, si enunciamos la siguiente hipótesis: *los varones prefieren una residencia independiente en mayor proporción que las mujeres y viceversa*; entonces se esperaría mayor cantidad de casos en las celdas *femenino y familiar* (celda a), y *masculino e independiente* (celda d). *Q* también resulta interpretable, pues un resultado negativo indica que se halló una relación opuesta a la esperada (como en el Cuadro 5 o el 6), lo que sugiere que también se puede aplicar con variables nominales, y lo mismo podemos decir de *E2*, dejando de lado la posible existencia de casos particulares en que esto no suceda.

Cuadro 6. Cantidad de casos según género y tipo de residencia y coeficientes de asociación

	Residencia		
	Familiar	Independiente	Total
Femenino	0 0%	115 100%	150 100%
Masculino	121 100%	0 0%	169 100%

E1 = 0.00
E2 = -1.00
Q = -1.00

Fuente: Elaboración propia en base a datos ficticios.

5. Variables ordinales categóricas

Los cuadros 7 al 10 contienen datos hipotéticos. En filas se ubica la variable independiente y sus categorías son *B* y *A* (por ejemplo, podría ser un nivel *Bajo* y *Alto* en *comprensión lectora*). En columnas la segunda variable (por ejemplo, *rendimiento Bajo* y *Alto*). Se obtuvieron porcentajes en fila y a la derecha, se muestran los valores de *E1*, *E2* y *Q*.

En el Cuadro 7 se da una fuerte relación positiva entre las variables, en el Cuadro 8, no existe relación, mientras que en el Cuadro 9, se da una relación fuerte pero negativa y en el Cuadro 10, una relación positiva perfecta (todos los casos son corroboratorios).

Cuadro 7. Cantidad de casos y coeficientes con dos variables hipotéticas

	B	A	T
B	185 89,4%	22 10,6%	207 100%
A	25 13,5%	160 86,5%	185 100%

E1 = 0,88
E2 = 0,76
Q = 0,96

Fuente: Elaboración propia en base a datos ficticios.

Cuadro 8. Cantidad de casos y coeficientes con dos variables hipotéticas

	B	A	T
B	95 49,2%	98 50,8%	193 100%
A	85 49,1%	88 50,9%	173 100%

E1 = 0,50
E2 = 0,00
Q = 0,00

Fuente: Elaboración propia en base a datos ficticios.

Cuadro 9. Cantidad de casos y coeficientes con dos variables hipotéticas

	B	A	T
B	23 11,4%	179 88,6%	202 100%
A	169 84,9%	30 15,1%	199 100%

E1 = 0,13
E2 = -0,74
Q = -0,96

Fuente: Elaboración propia en base a datos ficticios.

Cuadro 10. Cantidad de casos y coeficientes con dos variables hipotéticas

	B	A	T
B	115 100%	0 0%	115 100%
A	0 0%	121 100%	121 100%

E1 = 1,00
E2 = 1,00
Q = 1,00

Fuente: Elaboración propia en base a datos ficticios.

Se observa que cuantos más casos *corroboratorios* se presenten, los coeficientes tenderán a *1* (Cuadro 7). Recíprocamente, cuantos más casos *refutatorios* hay, tanto *Q* como *E2* se aproximan a *-1*, lo que implica una fuerte asociación entre las variables, pero de sentido inverso al postulado por la hipótesis (Cuadro 9). Si la relación es perfecta llegan al *1* (Cuadro 10), y cuanto más se equilibran -los casos corroboratorios y refutatorios-, *E2* y *Q* más se aproximarán a *0*, lo que implica ausencia de relación (Cuadro 8). En cambio, como dije más arriba, con *E1*, sucede algo diferente: cuando la relación es fuerte, pero de sentido opuesto, se aproxima a *0*; si es muy baja o inexistente, a *0.50;* y si es fuerte, pero en sentido directo, tenderá a *1*.

Por otro lado, *E2* tiene una ventaja sobre *Q*. En el Cuadro 11, se ve que *Q* da un resultado muy alto cuando esto no debiera suceder. En efecto, si bien la primera fila no contiene casos refutatorios (todos están en la categoría *B* en las dos variables, esto es, *bajo*), no obstante, un número importante de casos refutatorios se aprecian en la celda *c* (casos *A* en fila y *B* en columna), por lo tanto, la relación no puede ser perfecta como indica *Q*, sino que debiera ser muy débil y negativa, como efectivamente ofrece *E2*. Esto se debe a que *Q*, al multiplicar las celdas *b* y *c*, se obtiene *0*, pues la primera no contiene casos, y todo número multiplicado por *0* da esta cantidad. *E2,* en lugar de multiplicar suma, por ello esto no sucede. En general, *Q* no es aplicable si alguna de sus celdas contiene un 0, pues puede dar resultados engañosos. También se ve en todas las tablas presentadas hasta aquí que en general *E2* y *Q* dan resultados similares, pero el primero es menor sobre todo cuando la relación entre las variables es débil, tema que no puedo profundizar aquí.

Cuadro 11. Cantidad de casos y coeficientes con dos variables hipotéticas

	B	A	T
B	115 100%	0 0%	115 100%
A	300 71,3%	121 28,7%	421 100%

E1 = 0,44
E2 = -0,12
Q = 1,00

Fuente: Elaboración propia en base a datos ficticios.

6. La confiabilidad con el enfoque de la consistencia interna

Es mucho lo que se ha discutido sobre la forma de evaluar la aplicación de los IRDs (instrumentos de recolección de datos), pero más allá de la polémica, su inclusión se considera imprescindible, o al menos importante, en las utilizaciones que se hacen de ellos: "Cada vez que se administra un instrumento de medición debe calcularse la confiabilidad, al igual que evaluarse la evidencia sobre la validez" (Hernández Sampieri et al., 2014: 208).

Un momento importante en esta polémica se dio cuando la APA, la AERA y el NCME (American Psychological Association, American Educational Reserach Association y National Council on Measurement in Education) propusieron las famosas normas técnicas para evaluar los test psicológicos. En ese entonces, se aceptaba que los requisitos que debían observarse al usarlos, para hacerlo de un modo científica y éticamente válido, eran la validez, la confiabilidad y la objetividad (Múrat, 1983).

El concepto de *objetividad* es, ciertamente, muy controversial, y a veces suele mencionarse la confiabilidad entre examinadores (Olaz, 2008), lo que aumenta la confusión. Considero más conveniente hablar de *concordancia ínterobservadores*, pues no supone el sentido fuerte que el primero podría connotar (Echevarría, 2019a). De todos modos, este tema encierra una serie de complejas sutilezas, imposibles de tratar en el espacio de este trabajo. Lo más importante para este artículo, es que, en la actualidad, estos requisitos se postulan para toda utilización de un IRD (Prieto y Delgado, 2010), aunque generalmente se incluyen sólo los dos primeros.

Mientras que la *validez* se relaciona a la evidencia brindada de que el IRD midió efectivamente lo que se pretendió medir en un uso determinado, la *confiabilidad* se refiere a la exactitud con que la prueba lo logró.

En lo que resta de este artículo, presentaré algunos procedimientos de cálculo de la confiabilidad, limitándome a una forma específica de hacerlo: el "enfoque de la consistencia interna" (Nunnally y Bernstein, 1995: 281). Trataré de mostrar que, en última instancia, en los coeficientes de confiabilidad, subyace la sencilla función matemática mencionada más arriba: ($y = x_1/x_2$). Lo haré del modo más simple posible, aunque *supongo conocido el concepto de varianza*. Me referiré a las propiedades psicométricas relacionadas a la aplicación de los IRDs, luego introduciré algunos coeficientes alternativos al *Alfa* de Cronbach (*Alfa*) y al Coeficiente N° 20 de Kuder y Richardson (*KR20*), aunque más simples que estos, lo que facilita advertir el modo en que logran los objetivos para los que fueron creados. Los analizaré mediante el uso de datos ficticios que representan situaciones claramente contrastantes, precisamente para que sus diferencias sean más fáciles de apreciar.

7. Las propiedades psicométricas relacionadas a la aplicación de los IRDs

Creo que el término *propiedades psicométricas* no es el más adecuado para mencionar la validez del proceso de medición u observación, ya que da la impresión de que sólo deben tenerse en cuenta al considerar los test psicométricos; de todos modos, dado que no hallé otro mejor, he decidido adoptarlo. Además, *validez* se utiliza a veces con dos sentidos diferentes: por un lado, el mencionado más arriba (grado de seguridad de que el IRD midió efectivamente lo que se pretendió medir) y, por otro lado, para mencionar la validez de todo el proceso de medición, que involucra, al menos, también a la *confiabilidad* y en algunas circunstancias particulares a la *concordancia ínterobservadores.*

Tabla 1. Propiedades psicométricas de los IRDs

Validez	Contenido	
	Criterio	Concurrente
		Predictiva
Confiabilidad	Medida de estabilidad (test-retest)	
	Formas alternativas o paralelas	
	Mitades partidas	
	Medidas de consistencia interna	

Fuente: Elaboración propia.

Para delimitar el tema que trato, en la Tabla 1 presento *los modos más comunes de considerar estos dos conceptos*.[7] En relación al tópico de este artículo, puede verse que existen cuatro grandes formas de operacionalizar la confiabilidad; como dije más arriba, me referiré a una de ellas: las medidas de consistencia interna.

8. Coeficientes de confiabilidad con el enfoque de la consistencia interna

Siempre que se mide alguna variable, se dan dos tipos de variabilidad: la intrínseca del fenómeno y la debida al error de medición.[8] Los *coeficientes de confiabilidad* pretenden cuantificar la segunda variabilidad en relación a la variabilidad total y, específicamente, los de *consistencia interna*, se basan en estimar la consistencia entre los ítems de un IRD que miden el mismo constructo o al menos una dimensión determinada del mismo.

Antes de continuar, una breve explicación de las Tablas 3 a la 7 y la 9, que muestran algunos ejemplos con datos ficticios, que podrían haber surgido de aplicar un IRD con 5 ítems a 10 sujetos. Se ubicaron los sujetos en filas y los ítems en columnas. Tomando la Tabla 3, podría ser, por ejemplo, una prueba de aritmética que tiene 5 subpruebas (ítems) y se obtuvieron puntajes del 1 al 4. O también, cinco afirmaciones relacionadas a la actitud hacia la desigualdad de género, con una escala de valoración de 4 opciones, en que 4 significa totalmente de acuerdo con cada afirmación y 1 totalmente en desacuerdo, por mencionar sólo dos casos posibles (la Tabla 2 contiene todas las abreviaturas que también se explican más adelante).

Difícilmente se pueden llegar a observar en la realidad estos datos, pues las tablas han sido construidas de modo que representen situaciones muy diferentes, para que se pueda apreciar con nitidez cómo funcionan los coeficientes. Además, prescindo de considerar algunas características importantes que debe tener un IRD. Por ejemplo, con resultados como en la Tabla 5, de nada serviría haber recolectados esos datos, dada la falta de discriminación que habría presentado el IRD.

Tabla 2. Abreviaturas usadas

CC1, CC2 y CC3	Coeficientes de confiabilidad
Alfa	Alfa de Cronbach
KR20	Coeficiente N° 20 de Kuder y Richardson
VC	Varianzas en columnas
VF	Varianzas en filas
S*VC*	Sumatoria de las varianzas en columnas
MeVC	Media o promedio de las varianzas en columnas
SVF	Sumatoria de las varianzas en filas
MeVF	Media (promedio) de las varianzas en filas
Vtot	Varianza del puntaje total
*p*q*	Producto de *p* por *q* (*p* = probabilidad o frecuencia de respuestas correctas y *q* = *1-p*)
Ít	Ítem
k	Número de ítems de una prueba

Fuente: Elaboración propia

[7] Además, hay que tener presente que el término *validez* también se usa para referirse a otras propiedades metodológicas. En otro lugar presenté un modelo para analizar todo el proceso inferencial en las ciencias sociales (Echevarría, 2019a; 2019b). El segundo representa una versión sintética del primero. Estoy adoptando aquí una postura ingenua respecto de estos conceptos para poder referirme al tema específico de la confiabilidad en el espacio disponible.

[8] También tenemos la varianza del total que depende de ambas. *Alfa*, como muestro más adelante, se basa en ella, y la considero en el coeficiente *CC3* que introduzco más adelante e implícitamente en los otros que presento.

Utilizando la función mencionada más arriba, voy a definir un coeficiente de confiabilidad, basado en dividir la variabilidad intersujeto entre la variabilidad intersujeto más la variabilidad intrasujeto, tomando la varianza como indicadora de variabilidad. Por ejemplo, la Tabla 3 contiene las varianzas en columna (fila VC), es decir, la variabilidad entre sujetos para un mismo ítem. Por otro lado, la columna VF, muestra todas las varianzas en fila, o sea, de los valores obtenidos por cada sujeto o caso en los distintos ítems. *Si los valores son iguales o similares en filas, entonces se observa coherencia entre los ítems para un mismo sujeto; en este caso, el IRD presenta consistencia interna porque arroja resultados similares (varianza baja) en cada caso*. Esta tabla fue construida con valores muy coherentes al interior de cada sujeto, esto es, con poca variabilidad intrasujetos. El primer caso muestra los siguientes valores: 1, 2, 1, 1, 1. Excepto este y el 9, los restantes son más coherentes aún, pues todos los casos tienen el mismo puntaje en los cinco ítems.

Si adoptamos los supuestos de que todos los ítems miden el mismo constructo, en el mismo sentido y con un grado de dificultad similar entre los ítems, tenemos una variabilidad debida a errores en el proceso de medición (en filas) y una variabilidad inherente al fenómeno estudiado, es decir, entre los distintos sujetos (en columna).

El primer coeficiente que presento se basa en dividir la media o promedio de la variabilidad intersujetos (la intrínseca al fenómeno estudiado, en columnas), sobre esta misma variabilidad intersujetos, más la variabilidad intrasujetos (el error de medición), también considerando la media. Si esta última es baja en relación a la primera, el IRD será muy confiable. También supone una escala intervalar o de razón en cada uno de los ítems (porque los puntajes de los ítems se suman para obtener el puntaje total y por la fórmula que se usa para calcular la varianza) y su fórmula es la siguiente:

$$CC1 = MeVC / (MeVC + MeVF),$$

Donde:

CC1: Coeficiente de confiabilidad 1,

MeVC: Media de las varianzas en columnas (intersujetos),

MeVF: Media de las varianzas en filas (intrasujetos).

Como ya dije, la columna *VF* contiene la varianza de las respuestas de cada sujeto. Como es lógico, dada la consistencia de sus respuestas, la varianza fue baja. El primer sujeto y el 9 tienen una varianza de 0,2 y los restantes de 0 puesto que tuvieron el mismo puntaje en todos los ítems (fueron totalmente consistentes, no hubo variabilidad en las respuestas de cada sujeto a los distintos ítems). Por otro lado, la fila *VC* muestra las varianzas para cada ítem, entre los distintos sujetos (la varianza de cada columna). *Es la varianza debida al fenómeno estudiado. En este ejemplo hipotético, relacionada a la diferencia de conocimientos en aritmética entre los distintos sujetos.*

La celda *SVC* muestra la suma de todas las varianzas en este sentido, es decir, es la suma de todas la varianzas de los ítems (en la Tabla 3 se redondearon algunas cifras, y es igual a 1,611+1,38+1,611+1,611+1,82=8,03); y la celda *MeVC* esta cifra divida por 5, es decir, la media (8,03/5). Recíprocamente, la celda *SVF* muestra la suma de todas las varianzas de cada sujeto (o en fila) que suman 0,4 (0,2+0,2). Al poner datos muy similares en fila, estas varianzas resultaron bajas y, por lo tanto, la suma también lo es. Además, la celda *MeVF*, muestra esta cantidad dividida entre 10 (que es la cantidad de filas o de casos), o sea, la celda contiene la media o promedio de las varianzas de los puntajes al interior de cada sujeto. La celda *CC1* muestra el coeficiente obtenido que, aplicando la fórmula vista más arriba, se calculó del siguiente modo (más adelante me refiero a los otros coeficientes):

$$CC1 = 1{,}607 / (1{,}607+0{,}04) = 0{,}976.$$

Como no podía ser de otro modo, dados los datos, resultó alto. El coeficiente tiene la forma:

$CC1 = x_1/(x_1+x_2)$.

Fue definido de modo que a medida que aumenta x_2 respecto de x_1, el coeficiente disminuye, pues como mostré más arriba, en una función de ese tipo, cuando aumenta el denominador en relación al numerador, disminuye el resultado de la división. En otros términos, al aumentar la variabilidad intrasujetos, disminuye el coeficiente. Por ejemplo, la Tabla 5, contiene datos que reflejan mucha variabilidad -inconsistencia- intrasujetos y muy poca intersujetos, por lo que el coeficiente resultó muy bajo.

Recíprocamente, cuando disminuye x_2 en relación a x_1, aumenta el coeficiente, pues en este caso el denominador tiende al numerador, con lo que el coeficiente tiende a *1*, resultando que nunca puede superar esa cantidad. En el caso extremo de que $x_2 = 0$, si bien es prácticamente imposible llegar a obtener este resultado (estaríamos ante un IRD totalmente confiable), reemplazando tenemos:

$$CC1 = x_1/(x_1+0) = x_1/x_1 = 1.$$

Además, nunca podrá ser menor a *0*, pues se basa en varianzas, que por sus propiedades no pueden ser negativas.

Con los mismos principios, también podemos definir otros coeficientes del siguiente modo:

$$CC2 = SVC / (SVC + SVF),$$

$$CC3 = Vtot / (Vtot + SVF),$$

Donde:

SVC = Sumatoria de las varianzas en columna (entre sujetos),

SVF = Sumatoria de las varianzas en fila (intrasujetos),

Vtot = Varianza del puntaje total.[9]

Tabla 3. Puntajes de 10 casos a 5 Ítems (Ít) y coeficientes

Caso	Ít 1	Ít 2	Ít 3	Ít 4	Ít 5	VF	Total
1	1	2	1	1	1	0,2	6
2	4	4	4	4	4	0	20
3	2	2	2	2	2	0	10
4	4	4	4	4	4	0	20
5	1	1	1	1	1	0	5
6	2	2	2	2	2	0	10
7	3	3	3	3	3	0	15
8	4	4	4	4	4	0	20
9	3	3	3	3	4	0,2	16
10	1	1	1	1	1	0	5
VC	1,61	1,38	1,61	1,61	1,82		39,34

SVC = 8,03 MeVC = 1,61 Vtot = 39,34	SVF = 0,40 MeVF = 0,04 CC1 = 0,976	CC2 = 0,953 CC3 = 0,99 Alfa = 0,995

Fuente: Elaboración propia en base a datos ficticios.

[9] Estrictamente hablando todos los coeficientes vistos hasta ahora requieren escalas numéricas (intervalares o de razón).

En *CC2*, en el numerador se ubica la suma de las varianzas de los ítems (ubicados en columnas), y en el denominador esta suma más la suma de las varianzas en fila (entre los ítems al interior de cada sujeto). En *CC3* se reemplaza la suma de las varianzas en columnas por la varianza del puntaje total.

Tabla 4. Puntajes de 10 casos a 5 Ítems (Ít) y coeficientes

Caso	Ít 1	Ít 2	Ít 3	Ít 4	Ít 5	VF	Total
1	1	3	2	4	1	1,7	11
2	4	2	3	4	1	1,7	14
3	4	2	3	2	1	1,3	12
4	4	4	4	4	4	0	20
5	1	3	4	1	2	1,7	11
6	4	1	3	2	5	2,5	15
7	1	3	4	3	2	1,3	13
8	1	2	3	4	4	1,7	14
9	1	2	2	3	4	1,3	12
10	2	2	1	3	4	1,3	12
VC	2,233	0,711	0,989	1,111	2,4		7,1556

SVC	7,444	SVF	14,5	CC2	0,339
MeVC	1,489	MeVF	1,45	CC3	0,330
Vtot	7,156	CC1	0,507	Alfa	-0,050

Fuente: Elaboración propia en base a datos ficticios.

Tabla 5. Puntajes de 10 casos a 5 Ítems (Ít) y coeficientes

Caso	Ít 1	Ít 2	Ít 3	Ít 4	Ít 5	VF	Total
1	1	2	3	4	5	2,5	15
2	1	2	3	4	5	2,5	15
3	1	2	3	4	5	2,5	15
4	1	2	3	4	5	2,5	15
5	1	2	3	4	5	2,5	15
6	1	2	3	4	5	2,5	15
7	1	2	3	4	5	2,5	15
8	1	2	3	4	5	2,5	15
9	1	2	3	4	5	2,5	15
10	2	3	4	5	4	1,3	18
VC	0,1	0,1	0,1	0,1	0,1		0,9

SVC = 0,5 MeVC = 0,1 Vtot = 0,9	SVF = 23,8 MeVF = 2,38 CC1 = 0,04	CC2 = 0,021 CC3 = 0,036 Alfa = 0,556

Fuente: Elaboración propia en base a datos ficticios.

Tabla 6. Puntajes de 10 casos a 5 Ítems (Ít) y coeficientes

Caso	Ít 1	Ít 2	Ít 3	Ít 4	Ít 5	VF	Total
1	1	2	3	4	5	2,5	15
2	1	2	3	4	5	2,5	15
3	1	2	3	4	5	2,5	15
4	1	2	3	4	5	2,5	15
5	1	2	3	4	5	2,5	15
6	1	2	3	4	5	2,5	15
7	1	2	3	4	5	2,5	15
8	1	2	3	4	5	2,5	15
9	1	2	3	4	5	2,5	15
10	1	1	1	1	1	0	5
VC	0	0,1	0,4	0,9	1,6		10

SVC	3
MeVC	0,6
Vtot	10
SVC/Vtot	0,3

SVF	22,5
MeVF	2,25
CC1	0,21

CC2	0,118
CC3	0,308
Alfa	0,875

Fuente: Elaboración propia en base a datos ficticios.

9. Coeficiente *Alfa* de Cronbach

Cuando los datos no son dicotómicos como en las Tablas 3 a la 7 y la 9, el coeficiente que más se utiliza es el *Alfa* de Cronbach (*Alfa*), cuya fórmula es menos intuitiva (tomada de Nunnally y Bernstein, 1995: 261):

*Alfa = [k/(k-1)] * [1 – (SVC / Vtot)].*[10]

Donde:

k = Número de ítems,

SVC = Sumatoria de las varianzas en columna (de los ítems) y

Vtot = Varianza del puntaje total.

[10] Existe otra forma de calcularlo, equivalente a esta, basada en las correlaciones de los ítems. Además, utilizo aquí las mismas abreviaturas de los coeficientes anteriores (Tabla 2).

Tabla 7. Puntajes de 10 casos a 5 Ítems (Ít) y coeficientes

Caso	Ít 1	Ít 2	Ít 3	Ít 4	Ít 5	VF	Total
1	1	2	2	4	5	2,7	14
2	1	2	2	4	5	2,7	14
3	1	2	2	4	5	2,7	14
4	1	2	3	4	5	2,5	15
5	1	2	3	4	5	2,5	15
6	4	3	2	2	1	1,3	12
7	4	3	2	2	1	1,3	12
8	4	3	2	2	1	1,3	12
9	4	3	3	2	1	1,3	13
10	4	3	3	2	1	1,3	13
VC	2,50	0,28	0,27	1,11	4,44		1,38

SVC = 8,6 MeVC = 1,72 Vtot = 1,38	SVF = 19,6 MeVF = 1,96 CC1 = 0,47	CC2 = 0,305 CC3 = 0,413 Alfa = -6,552

Fuente: Elaboración propia en base a datos ficticios.

Como puede verse, el *Alfa* depende también del número de ítems, de modo que cuando menos son estos, más lo aumenta. La Tabla 8 muestra algunos resultados de la operación *k/(k-1)*, para distintos números de ítems, comenzando por el menor posible (no se puede calcular la consistencia interna si no tenemos al menos dos ítems). Cuando mayor sea el resultado de esta división, mayor será el coeficiente. Si son sólo dos ítems (*k*=2), entonces *[1 – (SVC / Vtot)]* se multiplica por 2, es decir que duplica el valor del coeficiente. Si son tres ítems, se multiplica por 1,5, etc.

Para la Tabla 3, tenemos:

*Alfa = [5/(5-1)] * [1-(8.033/39.34)] = 1.25 * (1-0,204)] = 0.995.*

En este coeficiente, la parte clave de la fórmula es *SVC/Vtot.* El Cuadro 12 muestra este cociente junto a otros estadísticos calculados para las Tablas 3 a 6 (columna *SVC/Vtot*, se ordenó por esta columna para facilitar su lectura, en algunos se observan pequeñas diferencias con los incluidos en estas tablas por los redondeos realizados). Puede verse que *Alfa* está en relación inversa a este cociente, es decir, cuando menor es el resultado de *SVC/Vtot,* más alto resulta *Alfa*. En la ecuación de *Alfa* incluida más arriba, este cociente se le resta a uno, por lo tanto, cuando mayor sea, menor será el resultado de esta resta, y luego, cuando se multiplica por *k/(k-1)*, el resultado seguirá siendo bajo, aunque un poco más alto (como vimos en el párrafo anterior). O sea que *Alfa* se basa en cuantificar la variabilidad de los ítems en relación a la variabilidad total, agregando una corrección que depende del número de ítems.

¿Qué sucede con el *CC1* cuando existe más variabilidad (inconsistencias) intrasujetos (entre los ítems para un mismo caso)? La Tabla 4 incluida más arriba, muestra un ejemplo de este tipo. Si el coeficiente refleja lo opuesto de las contradicciones (variabilidad o varianza) en las respuestas de cada sujeto, es esperable que adquiera un valor medio cuando estas no son demasiadas, como en esta Tabla, y muy bajo con los datos de la Tabla 5, como efectivamente sucede. Esta última fue construida de modo que se dé prácticamente la máxima contradicción intrasujetos: excepto el 10, con puntajes del 1 al 5 todos presentan respuestas distintas a todos los ítems. También se aprecia una mínima contradicción entre sujetos (para cada ítem, 9 de ellos dieron la misma respuesta, se introdujo un caso con puntajes diferentes, pues si todos tuvieran el mismo total, la varianza del total *-Vtot-* daría cero, y *Alfa* no se podría calcular por tener cero en la división *SVC / Vtot*).

Tabla 8. Resultados de la división k/(k-1), (desde 2 a 10 ítems)

k	k-1	k/(k-1)
2	1	2,00
3	2	1,50
4	3	1,33
5	4	1,25
6	5	1,20
7	6	1,17
8	7	1,14
9	8	1,13
10	9	1,11

Fuente: Elaboración propia.

Cuadro 12. Algunos estadísticos de las Tablas 3 a 6

Tabla	SVC	SVF	Vtot	SVC/Vtot	1-(SVC/Vtot)	Alfa
3	8,03	0,4	39,34	0,20	0,80	0,99
5	0,5	23,8	0,9	0,56	0,44	0,56
4	7,44	14,5	7,16	1,04	-0,04	-0,05
6	8,6	19,6	1,38	6,23	-5,23	-6,54

Fuente: Elaboración propia.

En la Tabla 5, se nota que *Alfa* es demasiado alto, pues son datos que muestran la máxima contradicción intrasujetos posible y da un valor de 0,556, lo cual no es deseable, mientras que los *CC* dan una cifra más baja, sobre todo *CC1* cuyo valor es 0,04.

Incluso, *si por ejemplo al caso 10 le damos el valor de 1 en todas las respuestas* (ver Tabla 6), *Alfa* se eleva a 0,875. Según Campo-Arias y Oviedo (2008: 834), "para los más liberales, la consistencia interna de una escala se considera aceptable cuando se encuentra entre 0,70 y 0,90 (…). Otros más conservadores sugieren que (…) es adecuada si el coeficiente alcanza valores entre 0,80 y 0,90". Así, con esos datos y con un solo caso consistente, *Alfa* sería aceptable, lo cual resulta inadmisible dado que 9 casos contienen datos totalmente inconsistentes. En cambio, *CC1*, *CC2* y *CC3* son iguales a *0,21*; *0,12* y *0,31* respectivamente, lo cual resulta más coherente (redondeando los dos últimos).

La Tabla 7, en cambio, muestra una situación en que se da una alta contradicción en filas (intrasujetos), pero con mayor variabilidad entre ítems y una varianza del total muy baja. Además, la parte superior (hasta el caso 5) presenta puntajes bajos en los dos primeros ítems y altos en los dos últimos, sucediendo lo opuesto con los sujetos 6 al 10. Esto hace que *Alfa* sea menor que cero, lo que se debe a que los datos violan un supuesto del modelo: presentan una covarianza promedio entre ítems negativa (lo que también se observa en la Tabla 4). Pero puede verse que esto no sucede con *CC1, CC2 y CC3*, por lo que tienen la ventaja de no requerir este supuesto.

No obstante, los coeficientes *CC2* y *CC3* tienen la dificultad de que dependen del número de casos como se aprecia en la Tabla 9, que es la Tabla 5 duplicada (es decir, incluye los casos 11 al 20 con los mismos datos que los del 1 al 10). Esto es así, porque consideran la sumatoria de las varianzas en filas, que no son otra cosa que las varianzas de cada uno de los casos, entonces, al aumentar estos aumenta esta sumatoria y al estar en el denominador, disminuyen estos coeficientes. Esto no sucede con *Alfa*, ni con *CC1.* Este último, al tomar los promedios de las varianzas en filas y columnas, evita estar influido por el número de casos y de ítems (en la Tabla 9 se da la misma proporción de contradicciones en fila y las mismas varianzas en columnas que en la Tabla 5 y tanto *Alfa* y como *CC1* mantienen sus valores).

10. Coeficientes de consistencia interna para escalas dicotómicas

Es posible generalizar estos coeficientes para datos dicotómicos, *para lo cual se debe modificar la fórmula de calcular la varianza*. Para el caso anterior (prueba de conocimientos de matemática), podría contener, por ejemplo, 5 problemas puntuándose con *1* si un problema (ítem) dado es respondido en forma correcta y *0* en el caso contrario. El Coeficiente *CC1.2* tiene la misma fórmula del *CC1,* pero utilizando el procedimiento de cálculo de la varianza para datos dicotómicos, al igual que el coeficiente *KR20* que es análogo al *Alfa.* Las Tablas 10 y 11 contienen datos dicotómicos, en la primera, al ser muy consistentes al interior de cada sujeto, los coeficientes resultaron altos, a la inversa de la Tabla 11 que presenta muchas contradicciones en filas. En esta última se observa que, como es esperable, ambos dan un valor muy bajo.

11. Consideraciones finales

La idea central de todos los *coeficientes* de asociación y correlación, es cuantificar la relación entre la variación conjunta sobre la variabilidad total posible. Los que presenté aquí, creados para variables nominales y ordinales consisten en calcular las coincidencias entre categorías en cada caso sobre el total de casos. Al calcular estos coeficientes, con un número representamos el grado de asociación entre al menos dos variables y en algunos, también su dirección, lo cual también se aprecia en el cruzamiento, aunque de un modo impreciso.

Comencé con un coeficiente muy simple, *E1*, que, si bien ofrece un resultado interpretable, tiene la dificultad de que su escala es distinta a las que usan generalmente los coeficientes, lo que puede dar una idea confusa de lo hallado.

E2 resulta más simple que *Q,* ya que en definitiva se obtiene de sumar los casos corroboratorios, luego restarle los refutatorios y dividir esta cantidad por el total de casos. Además, tiene sobre *Q* la ventaja de permitir que alguna celda no contenga casos con la seguridad de no distorsionar su resultado. En general *E2* dio un valor más bajo que *Q*, aunque para explicar estas diferencias se necesitan métodos matemáticos más avanzados que los que puedo utilizar aquí, por lo que sería interesante estudiar estas diferencias en trabajos futuros, incluso, comparándolo con otros coeficientes.

En relación a los coeficientes de confiabilidad, el valor de los *CC* depende de la consistencia de las respuestas de los sujetos a los distintos ítems, es decir, tratan de cuantificar esta consistencia. Esta variabilidad intrasujetos se incluyó en el denominador, de modo que cuanto más aumenta en relación a la variabilidad entre los ítems, más disminuye la magnitud de estos coeficientes. En los ejemplos presentados, se notó que reflejan la confiabilidad de los datos. Además, *CC1* y *CC1.2* no dependen del número de ítems del IRD, lo que si sucede con *Alfa* y *KR20* (Domínguez-Lara, 2016), y *Alfa* para algunas tablas adoptó valores negativos por tener una covarianza promedio entre ítems negativa, es decir, por violar uno de sus supuestos, lo que no sucede con *CC1,* o sea que tiene la ventaja de no requerir este supuesto.

Tabla 9. Puntajes de 20 casos a 5 Ítems y coeficientes (Tabla 5 duplicada)

Caso	Ít 1	Ít 2	Ít 3	Ít 4	Ít 5	VF	Total
1	1	2	3	4	5	2,5	15
2	1	2	3	4	5	2,5	15
3	1	2	3	4	5	2,5	15
4	1	2	3	4	5	2,5	15
5	1	2	3	4	5	2,5	15
6	1	2	3	4	5	2,5	15
7	1	2	3	4	5	2,5	15
8	1	2	3	4	5	2,5	15
9	1	2	3	4	5	2,5	15
10	2	3	4	5	4	1,3	18
11	1	2	3	4	5	2,5	15
12	1	2	3	4	5	2,5	15
13	1	2	3	4	5	2,5	15
14	1	2	3	4	5	2,5	15
15	1	2	3	4	5	2,5	15
16	1	2	3	4	5	2,5	15
17	1	2	3	4	5	2,5	15
18	1	2	3	4	5	2,5	15
19	1	2	3	4	5	2,5	15
20	2	3	4	5	4	1,3	18
VC	0,1	0,1	0,1	0,1	0,1		0,9

SVC	0,474
MeVC	0,09
Vtot	0,853
SVC/Vtot	0,56

SVF	47,6
MeVF	2,38
CC1	0,04

CC2	0,010
CC3	0,018
Alfa	0,556

Fuente: Elaboración propia.

Otro aspecto importante que se vio, es que en una de las tablas *Alfa* dio un valor demasiado alto dadas las contradicciones intrasujetos que contiene, incluso si se le introducen a un solo caso datos consistentes, daría un valor más alto aún, lo cual es inadmisible. En cambio, *CC1*, *CC2* y *CC3* dieron valores bajos, lo que se supone que debe ocurrir, dada esa hipotética situación.

CC1, al igual que *Alfa*, tampoco depende del número de casos, y como *CC1.2* se basa en los mismos principios, es esperable que con este suceda lo mismo, por lo que sería interesante indagar en trabajos futuros si esto es así. También quedan otros problemas por resolver, como por ejemplo, establecer procedimientos de cálculo de los intervalos de confianza para los coeficientes de asociación presentados, como así también para los de confiabilidad, cuando se pretende generalizar de la muestra a la población, ya que, como dije al principio, aquí me referí exclusivamente a la estadística descriptiva, es decir, solamente a procedimientos orientados a caracterizar una muestra o a indagar una de las propiedades psicométricas de los datos recolectados.

Tabla 10. Respuestas de 10 casos a 5 Ítems y coeficientes

Caso	Ít. 1	Ít. 2	Ít. 3	Ít. 4	Ít. 5	p*q	Total
1	1	1	0	0	0	0,24	2
2	1	1	1	1	1	0	5
3	1	1	1	1	1	0	5
4	1	1	1	1	1	0	5
5	1	1	1	1	1	0	5
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
p	0,5	0,5	0,4	0,4	0,4		
q	0,5	0,5	0,6	0,6	0,6		
p*q	0,25	0,25	0,24	0,24	0,24		

SVC = 1,22 MeVC = 0,24 Vtot = 6,18	SVF = 0,24 MeVF = 0,02	CC1.2 = 0,91 KR20 = 1,00

Fuente: Elaboración propia en base a datos ficticios.

Tabla 11. Respuestas de 10 casos a 5 Items y coeficientes

Caso	Ít. 1	Ít. 2	Ít. 3	Ít. 4	Ít. 5	p*q	Total
1	1	1	0	0	0	0,24	2
2	1	1	0	0	0	0,24	2
3	1	1	0	0	0	0,24	2
4	1	1	0	0	0	0,24	2
5	1	1	0	0	0	0,24	2
6	1	1	0	0	0	0,24	2
7	1	1	0	0	0	0,24	2
8	1	1	0	0	0	0,24	2
9	1	1	0	0	0	0,24	2
10	1	1	0	0	1	0,24	3
p	1	1	0	0	0,1		
q	0	0	1	1	0,9		
p*q	0	0	0	0	0,09		

SVC = 0,09 MeVC = 0,02 Vtot = 0,10	SVF = 2,40 MeVF = 0,24	CC1.2 = 0,07 KR20 = 0,13

Fuente: Elaboración propia en base a datos ficticios.

12. Bibliografía

BRIONES, G. (2002). *Metodología de la investigación cuantitativa en ciencias sociales.* Instituto Colombiano para el fomento de la Educación superior. Recuperado de: https://metodoinvestigacion.files.wordpress.com/2008/02/metodologia-de-la-investigacion-guillermo-briones.pdf

CAMPO-ARIAS, A. y OVIEDO, H. (2008). Propiedades psicométricas de una escala: la consistencia Interna. *Rev. Salud Pública*, 10 (5), 831-839.

CAZAU, P. (2006). *Fundamentos de Estadística.* Buenos Aires: Universidad de Buenos Aires. Recuperado de: http://www.listinet.com/bibliografia-comuna/Cdu311-6247.pdf

CRONBACH, L. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64, 391–418.

DE LA FUENTE, S. (2011). *Tablas de contingencia*. Facultad de Ciencias Económicas y Empresariales. Universidad Autónoma de Madrid. Recuperado de: http://www.estadistica.net/ECONOMETRIA/CUALITATIVAS/CONTINGENCIA/tablas-contingencia.pdf

DOMINGUEZ-LARA, S. (2016). Evaluación de la confiabilidad del constructo mediante el coeficiente H: breve revisión conceptual y aplicaciones. *Psychologia: Avances de la Disciplina*, 10 (2), 87-94.

ECHEVARRÍA, H. (2008). *Comprensión lectora de lógica y matemática en alumnos universitarios.* Río Cuarto: EFUNARC.

__________ (2019a). *Métodos de investigación e inferencias en Ciencias Sociales. Una propuesta para analizar su validez*. Río Cuarto: UniRío Editora. Recuperado de: http://www.unirioeditora.com.ar/?s=ECHEVARR%C3%8CA

__________ (2019b). Análisis de la validez de las inferencias en Ciencias Sociales. En A. Bono y M. Aguilera (Comps.), *Notas sobre investigación en Humanidades. Actas de las Jornadas de Investigación de la Facultad de Ciencias Humanas 2018.* Río Cuarto: UniRío Editora. (pp. 518-539). Recuperado de: http://www.unirioeditora.com.ar/producto/notas-investigacion-humanidades

HERNÁNDEZ SAMPIERI, R.; FERNÁNDEZ COLLADO, C. y BAPTISTA LUCIO, M. (2014). *Metodología de la investigación*. México: Mc Graw Hill.

MARRADI, A. (2007). El análisis bivariable. En A. Marradi, A. Archenti y J. Piovani (eds.), *Metodología de las ciencias sociales* (pp. 247-277). Buenos Aires: Emecé.

MAYNTZ, R., HOLM, K. y HÜBNER, P. (1975). *Introducción a los métodos de la sociología empírica.* Madrid: Alianza.

MÚRAT, F. (1983). *La evaluación del comportamiento*. Córdoba: Universidad Nacional de Córdoba.

NUNNALLY, J. y BERNSTEIN, I. (1995). *Teoría Psicométrica*. Buenos Aires: Mc Graw Hill.

OLAZ, F. (2008). Confiabilidad. En S. Tornimbeni, E. Pérez y F. Olaz (comps.), *Introducción a la psicometría* (pp. 71-99). Buenos Aires: Paidós.

PRIETO, G. y DELGADO, A. (2010). Fiabilidad y Validez. *Papeles del Psicólogo*, 31 (1), 67-74.

ZEIZEL, H. (1974). *Dígalo con números*. México: Fondo de Cultura Económica.

Autor.

Hugo Darío Echevarría

Ex docente, actualmente jubilado, de la Universidad Nacional de Río Cuarto y de la Universidad Nacional de Villa María, Argentina.

Doctor en Ciencias Sociales y Magíster en Epistemología y Metodología Científica por la Universidad Nacional de Río Cuarto, Argentina.

E-Mail: hechevarria2007@hotmail.com

Citado.

ECHEVARRIA, Hugo (2023). La lógica de los coeficientes. *Revista Latinoamericana de Metodología de la Investigación Social – ReLMIS,* N°26, Año 13, pp. 9-30.

Plazos.

Recibido: 03/03/2021. Aceptado: 10/02/2022.