

## FORMACIÓN CONTINUADA – METODOLOGÍA Y TÉCNICAS

# Coeficiente de correlación intraclase

J.A. Martínez Pérez<sup>a,\*</sup> y P.S. Pérez Martín<sup>b</sup>

<sup>a</sup> Miembro de la Comisión Nacional de Calidad de Semergen, Madrid, España

<sup>b</sup> Medicina de Familia, Centro Salud Palma Norte, Madrid, España

Recibido el 3 de octubre de 2022; aceptado el 20 de noviembre de 2022

Disponible en Internet el 26 de diciembre de 2022



### PALABRAS CLAVE

Coeficiente de correlación intraclase;  
Concordancia;  
Variables cuantitativas;  
Análisis de la varianza;  
Medidas repetidas

**Resumen** El coeficiente de correlación intraclase permite medir la concordancia general entre dos o más mediciones que implican variables de carácter cuantitativo, obtenidas estas con diferentes instrumentos de medida o evaluadores. Se basa en un modelo de análisis de varianza con medidas repetidas. Su uso solo es posible si existe normalidad de las distribuciones de las variables, igualdad de varianzas e independencia entre los errores producidos por los observadores.

© 2022 Sociedad Española de Médicos de Atención Primaria (SEMERGEN). Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

### KEYWORDS

Intraclass correlation coefficient;  
Concordance;  
Quantitative variables;  
Analysis of variance;  
Repeated measures

### Intraclass correlation coefficient

**Abstract** The intraclass correlation coefficient makes it possible to measure the general concordance between two or more measurements that involve variables of a quantitative nature, obtained with different measuring instruments or evaluators. It is based on an analysis of variance model with repeated measures. Its use is only possible if there is normality of the distributions of the variables, equality of variances and independence between the errors produced by the observers.

© 2022 Sociedad Española de Médicos de Atención Primaria (SEMERGEN). Published by Elsevier España, S.L.U. All rights reserved.

\* Autor para correspondencia.

Correo electrónico: [jmartinezp@semergen.es](mailto:jmartinezp@semergen.es)  
(J.A. Martínez Pérez).

En el ámbito clínico en ocasiones es necesario valorar la fiabilidad de las mediciones u observaciones realizadas. En este sentido podemos diferenciar dos circunstancias: a) valorar la fiabilidad de los resultados obtenidos cuando repetimos la medición con el mismo instrumento y en condiciones idénticas, y b) evaluar esta fiabilidad cuando obtenemos los

**Tabla 1** Medidas consecutivas de la presión arterial sistólica en 10 pacientes, realizadas por dos médicos con el mismo esfigmomanómetro

| Paciente | Médico A<br>(mmHg) | Médico B<br>(mmHg) |
|----------|--------------------|--------------------|
| 1        | 135                | 140                |
| 2        | 140                | 145                |
| 3        | 130                | 135                |
| 4        | 145                | 150                |
| 5        | 140                | 145                |
| 6        | 150                | 160                |
| 7        | 140                | 145                |
| 8        | 135                | 140                |
| 9        | 140                | 145                |
| 10       | 135                | 145                |

Fuente: Prieto et al.<sup>1</sup>.

resultados con diferentes instrumentos de medida o con diferentes observadores<sup>1</sup>.

Cuando se mide, siempre habrá un cierto error que dependerá de forma aleatoria de variables como el instrumento, la manipulación, el sujeto evaluado, etc. Cualquier medición es el resultado de sumar el valor real de aquello que queremos medir y un error aleatorio.

Un instrumento fiable será aquel cuyo error aleatorio sea lo suficientemente pequeño como para que nos permita considerar que el resultado obtenido no es atribuible a errores en la medición<sup>2</sup>.

En el caso de variables cualitativas, los índices kappa y kappa ponderado han demostrado ser los estadísticos más adecuados para estimar la fiabilidad en distintas mediciones<sup>3</sup>, y cuando se trata de variables cuantitativas continuas se debe utilizar para hacer esta evaluación el coeficiente de correlación intraclase (CCI)<sup>4</sup>.

En numerosas ocasiones se cometen errores al valorar la fiabilidad de las mediciones repetidas en casos de variables cuantitativas, al recurrir al uso del coeficiente de correlación de Pearson ( $r$ )<sup>5</sup>. Como podemos observar en la [tabla 1](#), en donde se comparan dos mediciones consecutivas de la presión arterial sistólica en 10 pacientes, realizadas con el mismo esfigmomanómetro por dos médicos diferentes, la  $r$  es excelente (0,95); sin embargo, la medición del médico B es constantemente superior a la del médico A. Esto se debe a que este coeficiente  $r$  lo que evidencia es la intensidad de la asociación lineal entre dos variables, es decir, su grado de asociación, pero no nos indica a qué se debe la diferencia sistemática entre las mismas, a favor de una de ellas, es decir, su grado de concordancia.

Por este motivo, el uso del coeficiente de Pearson puede constituir una fuente de error en la medición, ya que omite en el cálculo la variabilidad intrasujeto e intersujeto, exponiendo al investigador a errores sistemáticos en sus interpretaciones<sup>5</sup>.

Con el objeto de soslayar este obstáculo, se utiliza el CCI, el cual se basa en el modelo de análisis de la variancia con medidas repetidas<sup>6</sup>. Mide el grado de concordancia entre medidas cuya variancia puede ser debida a<sup>5</sup>:

- Variabilidad intersujeto debida a la diferencias entre estos ( $\sigma^2_s$ ).
- Variabilidad intrasujeto, debida a las diferencias de las mediciones de un mismo sujeto ( $\sigma^2_j$ ).
- Variabilidad residual, es inexplicable (aleatoria), va unida a los errores de medición ( $\sigma^2_e$ ).

El CCI es el cociente entre la variancia de los sujetos y la variancia total (variancia intersujeto, más variancia intrasujeto, más variancia residual):

$$CCI = (\sigma^2_s) / (\sigma^2_s) + (\sigma^2_j) + (\sigma^2_e)$$

Se define como la proporción de variabilidad total debida a la variabilidad de los sujetos<sup>1</sup>.

Al aplicar el método se pueden obtener valores comprendidos entre 0 y 1, donde 0 significa ausencia de concordancia y 1, concordancia o fiabilidad absoluta. De manera convencional se admiten los siguientes valores<sup>7</sup>:

- < 0,40: Pobre.
- 0,40-0,59: Suficiente.
- 0,60-0,74: Bueno.
- 0,75-1: Excelente.

Para que se pueda usar el CCI es necesario que se cumplan las siguientes condiciones: que exista normalidad de las distribuciones de las variables, homocedasticidad (igualdad de varianzas) e independencia entre los errores producidos por los observadores<sup>1</sup>.

Una limitación del CCI es que depende de la variabilidad de los valores observados. Si los pacientes varían poco en sus mediciones (muestra homogénea), el CCI tiende a ser bajo, ya que compara la variancia entre pacientes con la variancia total observada. Si la muestra es heterogénea, el CCI propende a ser mayor<sup>1</sup>.

Como alternativa al CCI se puede utilizar el método de Bland y Altman, que consiste en representar gráficamente, en un diagrama de dispersión, la media de las dos mediciones como la mejor estimación del verdadero valor, frente a la diferencia absoluta entre los dos valores. El gráfico incluye, además, una línea horizontal en la diferencia media y dos líneas, llamadas límites de concordancia, a una distancia de 1,96 desviaciones estándar por arriba y por debajo de la primera. Si las diferencias entre los pares de observaciones siguen aproximadamente una distribución normal y los valores tienden a ser estables en todo el rango de medición, se espera que el 95% de esas diferencias caigan dentro de los límites de concordancia. Esto permite valorar gráficamente, de forma sencilla, el grado de acuerdo entre los dos métodos<sup>8</sup>.

## Cálculo del CCI en SPSS

Seleccionar en los menús: Analizar → Escala → Análisis de fiabilidad → (aquí elegimos las variables que vamos a valorar) → Estadísticos → Coeficiente de correlación intraclase → Modelo, Tipo ([fig. 1](#)) → Continuar → Aceptar: obtención de los valores del coeficiente de correlación intraclase ([fig. 2](#)).

**Análisis de fiabilidad: Estadísticos**

**Descriptivos para**

☐ Elemento

☐ Escala

☐ Escala si se elimina el elemento

**Resúmenes**

☒ Medias

☒ Varianzas

☐ Covarianzas

☐ Correlaciones

**Inter-elementos**

☐ Correlaciones

☐ Covarianzas

**Tabla de ANOVA**

☐ Ninguno

☒ Prueba F

☐ Chi-cuadrado de Friedman

☐ Chi-cuadrado de Cochran

☐ T-cuadrado de Hotelling

☐ Prueba de aditividad de Tukey

☒ Coeficiente de correlación intraclass

Modelo: **Mixto de dos factores**

Tipo: **Consistencia**

Intervalo de confianza: 95 %

Valor de prueba: 0

**Continuar** **Cancelar** **Ayuda**

**Figura 1** SPSS: desplegable de modelo y de tipo.

Modelo: mixto de dos factores; aleatorio de dos factores; aleatorio de un factor.

Tipo: acuerdo absoluto; consistencia.

| Coeficiente de correlación intraclass |                                     |                               |                 |                                |     |     |      |
|---------------------------------------|-------------------------------------|-------------------------------|-----------------|--------------------------------|-----|-----|------|
|                                       | Correlación intraclass <sup>b</sup> | 95% de intervalo de confianza |                 | Prueba F con valor verdadero 0 |     |     |      |
|                                       |                                     | Límite inferior               | Límite superior | Valor                          | gl1 | gl2 | Sig  |
| Medidas únicas                        | ,894 <sup>a</sup>                   | ,722                          | ,963            | 18,190                         | 14  | 14  | ,000 |
| Medidas promedio                      | ,944 <sup>c</sup>                   | ,838                          | ,981            | 18,190                         | 14  | 14  | ,000 |

Modelo de dos factores de efectos mixtos donde los efectos de personas son aleatorios y los efectos de medidas son fijos.

a. El estimador es el mismo, esté presente o no el efecto de interacción.

b. Coeficientes de correlación intraclass de tipo A que utilizan una definición de acuerdo absoluto.

c. Esta estimación se calcula suponiendo que el efecto de interacción está ausente, porque de lo contrario no se puede estimar.

**Figura 2** Resultados finales del coeficiente de correlación intraclass (CCI).

## Elección del modelo

En primer lugar debemos distinguir entre modelo de efectos aleatorios y modelo de efectos fijos. En el primero los observadores implicados en la medición, por ejemplo los médicos A y B, representan una «muestra» aleatoria de la población de posibles observadores (p.ej., otros médicos del centro de salud) que en el futuro harán uso del instrumento evaluado.

En el modelo de efectos fijos, los observadores que intervienen en el estudio son los únicos que participan en el mismo.

En este módulo tenemos que elegir una de las tres siguientes opciones:

- *Mixto de dos factores*. Se usa cuando los médicos son de efectos fijos y los pacientes son de efectos aleatorios (el más utilizado).
- *Aleatorio de dos factores*. Se emplea cuando tanto los sujetos como los evaluadores son de efectos aleatorios.
- *Aleatorio de un factor*. Esta opción se usa cuando los efectos de las personas son aleatorios y no tenemos en cuenta la influencia de los evaluadores. Por ejemplo, tomamos dos medidas sin importar el orden en el cual se han realizado, bien porque lo desconozcamos o bien porque seamos totalmente conscientes de que no influyen en el resultado de nuestro experimento<sup>2</sup>. Es el menos usado.

## Elección del tipo

Hay que elegir una de las dos siguientes alternativas:

- *Consistencia*. Se emplea en el caso de que, aunque el promedio de los valores de las medidas de los dos instrumentos estén asociados linealmente, existen diferencias constantes entre ellos en la medida persona a persona.
- *Acuerdo absoluto*. Se utiliza cuando los dos instrumentos miden los mismos valores.

## Interpretación de los resultados obtenidos

En la [figura 2](#) podemos observar que al aplicar el programa informático obtenemos el CCI, con su intervalo de confianza del 95%, y también el estadístico de prueba Anova para medidas repetidas (F) con su significación estadística. Este estadístico F es útil para valorar si existe sesgo (error sistemático) en la medición realizada<sup>7</sup>.

Los resultados figuran en función de si en el estudio solo participaba un observador (medidas únicas) o más de uno, en cuyo caso se expone la media de todos los observadores (medidas promedio). La mayor parte de las veces

nos interesan estas últimas, aplicando las medidas individuales solo en aquellas situaciones en que queremos saber cuán preciso es un evaluador único al evaluar los valores obtenidos en su medición.

Como vemos en el ejemplo, el CCI logrado es de 0,94 y su rango se sitúa entre el 0,83 y el 0,98, lo cual nos indica que la concordancia en la medición entre los dos instrumentos evaluados es excelente.

## Consideraciones éticas

El trabajo no implicó el uso de sujetos humanos.

## Financiación

El trabajo no contó con apoyo financiero.

## Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

## Bibliografía

1. Prieto L, Lamarca R, Casado A. La evaluación de la fiabilidad en las observaciones clínicas; el coeficiente de correlación intraclase. *Med Clin (Barc)*. 1998;110:142–5.
2. Rodríguez, M. Tipos de índice de correlación intraclase (ICC) en SPSS [consultado 22 Sep 2022]. Disponible en: <https://www.cursoestadisticaonline.com > Blog>.
3. Tang W, Hu J, Zhang H, Wu P, He H. Kappa coefficient: a popular measure of rater agreement. *Shanghai Arch Psychiatry*. 2015;27:62–7, <http://dx.doi.org/10.11919/j.issn.1002-0829.215010>.
4. Sorana D, Bolboacă. Medical diagnostic tests: A review of test anatomy, phases, and statistical treatment of data. *Comput Math Methods Med*. 2019, <http://dx.doi.org/10.1155/2019/1891569>, 1891569.
5. Correa-Rojas J. Coeficiente de Correlación Intraclase: Aplicaciones para estimar la estabilidad temporal de un instrumento de medida. *Cienc. Psicol*. 2021;15:e-2318, <http://dx.doi.org/10.22235/cp.v15i2.2318>.
6. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–63, <http://dx.doi.org/10.1016/j.jcm.2016.02.012>.
7. Liljequist D, Elfving B, Skavberg Roaldsen K. Intraclass correlation — A discussion and demonstration of basic features. *PLoS One*. 2019;14:e0219854, <http://dx.doi.org/10.1371/journal.pone.0219854>.
8. Epidat: programa para análisis epidemiológico de datos. Versión 4.2, julio 2016. Consellería de Sanidade, Xunta de Galicia, España; Organización Panamericana de la Salud (OPS-OMS); Universidad CES, Colombia.