

Medida de acuerdo entre dos clasificaciones de los docentes universitarios*

Agreement measure in two faculty classifications

Luis Nava Puente** y Surendra P. Sinha***

Recibido: 16-10-07 / Aceptado: 11-12-07. Códigos JEL: C1, C12

Resumen

El objetivo de esta investigación es evaluar el nivel de acuerdo en la clasificación de los docentes universitarios según los componentes: estudio (*CICPI*) y escalafón (*CICP2*), propuestos por Sinha, Ramoni, Orlandoni, Torres y Figueroa (2007). Se usan dos esquemas de categorización, con tres y cuatro categorías, respectivamente. Se propone el uso del coeficiente kappa condicional ajustado para medir el nivel de acuerdo para cada categoría. Este coeficiente cumple con la propiedad de aditividad. Se demuestra que el coeficiente kappa de Cohen puede ser aplicado en situaciones no clásicas. Resultados obtenidos por Sinha, et. al. (2007), son confirmados en forma complementaria en esta investigación.

Palabras clave: Índice de capacitación profesional, coeficiente Kappa, coeficiente Kappa condicional, coeficiente Kappa condicional ajustado, Análisis de Correspondencia.

Abstract

The objective of this investigation is to evaluate the agreement level in the classification of the university faculty according to the components: study (*CICPI*) and promotion (*CICP2*) proposed by Sinha, Ramoni, Orlandoni, Torres y Figueroa (2007). Two categorization schemes are used, with three and four classification categories, respectively. We propose the use of the conditional adjusted kappa coefficient to measure the agreement level for each classification category. This coefficient fulfills the additivity property. It is demonstrated that the coefficient kappa of Cohen can be applied in non classic situations. Results obtained for Sinha, et. al. (2007), is confirmed complementary form in this investigation.

Key words: Faculty capability index, Kappa coefficient, conditional Kappa coefficient, conditional adjusted Kappa coefficient, correspondent analysis.

* Este artículo corresponde a un resultado parcial de la tesis doctoral "Modelación GSK para acuerdo entre evaluadores usando mínimos cuadrados ponderados y su comparación con otros métodos alternos: LOG-LINEAL y GEE", desarrollada por Luis Nava Puente y tutorada por el Doctor Surendra Sinha.

** Universidad de Los Andes, Escuela de Estadística. E-mail: navalu@ula.ve

*** Universidad de Los Andes, Instituto de Estadística Aplicada y Computación. E-mail: sinha32@yahoo.com

1. Introducción

Las medidas de acuerdo entre observadores, metodologías, procedimientos, test de diagnósticos, etc., han representado un tema de gran interés y aplicación en diversos campos, aun cuando su desarrollo se ha dado fundamentalmente en disciplinas como medicina, psicología, sociología y educación. Diversos índices, denominados también coeficientes, se han propuesto para medir acuerdo, entre ellos destaca el coeficiente kappa de Cohen (1960). Este coeficiente es de uso común, dada su facilidad de cálculo e interpretación. Cuando se desea conocer el comportamiento del acuerdo por categoría de clasificación, denominado también acuerdo parcial, se usa el índice kappa condicional propuesto por Coleman (1966) y Light (1969). Sin embargo, este coeficiente no cumple con la propiedad de aditividad, propiedad que sí satisface el coeficiente propuesto en este artículo, llamado coeficiente de kappa condicional ajustado.

Al considerar los componentes: estudio (*CICP1*) y escalafón (*CICP2*) para evaluar el Riesgo Académico Institucional, definido por Sinha *et. al.* (2007), se esperaría que los dos componentes clasifiquen a menudo a los docentes en la misma categoría, de forma tal que exista un acuerdo alto entre ellas. Un acuerdo bajo en la clasificación de los docentes mediante estos componentes, podría considerarse como un indicador de alto Riesgo Académico Institucional.

2. Objetivos

El objetivo general de este trabajo es proponer el uso del coeficiente kappa condicional ajustado y aplicarlo en la medición del acuerdo entre los componentes: estudio y escalafón, en la clasificación de los docentes de la Universidad de Los Andes. Se definen por lo tanto, los siguientes objetivos específicos:

- a. Medir el acuerdo global en la clasificación de los docentes de la Universidad de Los Andes a partir de los componentes estudio y escalafón, usando el coeficiente kappa.

- b. Definir el coeficiente kappa condicional ajustado propuesto en este artículo.
- c. Medir los niveles de acuerdo por categoría de clasificación, usando el índice propuesto en (b).

2.1. Coeficiente Kappa

Sinha *et. al.* (2007) introducen el concepto de Índice de Capacidad Profesional (ICP), el cual viene dado por la suma de dos componentes: el componente de estudio (*CICP1*), basado en el máximo nivel de estudio alcanzado por los docentes en su trayectoria académica, y el componente de escalafón (*CICP2*), basado en el ascenso de los profesores. Según Sinha *et. al.* (2007) los componentes (*CICP1*) y (*CICP2*) se definen de la siguiente manera:

$$CICP1_i = P_i(1 + W_{1i}) \quad (1)$$

donde P_i representa la permanencia del profesor i en la institución medido en años, W_{1i} representa una valoración del mérito del nivel máximo del estudio alcanzado por el profesor i (0, 1.5, 2, 4 según sea licenciatura, especialización, maestría o doctorado, respectivamente), y

$$CICP2_i = (P_i - Ab_i) \left(1 + W_{2i} \left(1 - \frac{Rasc_i}{P_i} \right) \right) \quad (2)$$

donde Ab_i representa el tiempo de abandono (medido en años) del trabajo de ascenso por parte del profesor i (función de los tiempos reglamentarios para el ascenso), W_{2i} es un coeficiente de ponderación que valora el mérito alcanzado por el profesor i al realizar su trabajo de ascenso. Los valores posibles de W_{2i} en orden ascendente según el escalafón son: 2/15, 6/15, 10/15, 15/15. $Rasc_i$ representa el retraso acumulado (medido en años) en los distintos ascensos.

Se puede observar que la experiencia profesional es un elemento presente en ambos componentes, cada uno de ellos es también una función de

la misma variable P_i que indica la permanencia del profesor. Es de interés, medir el grado de acuerdo en la clasificación de los docentes usando estos dos componentes.

El coeficiente κ de Cohen, propuesto por Cohen (1960), para variables de escala nominal, es un índice de acuerdo de uso común dada su facilidad de cálculo e interpretación. Puede ser aplicado a tablas de clasificación cruzada para dos o más evaluadores. Es un estadístico que condensa y proporciona en un solo valor, toda la información relevante acerca del acuerdo y está definido por

$$\kappa = \frac{\pi_0 - \pi_c}{1 - \pi_c} \quad (3)$$

donde π_0 representa la probabilidad de que ambos evaluadores acuerden y π_c la probabilidad de acuerdo por casualidad.

Dado que las variables estudio (*CICP1*) y escalafón (*CICP2*) son continuas, para el cálculo del coeficiente kappa de Cohen es necesario obtener una versión categorizada de las mismas. Se usan dos esquemas de categorización: con tres y cuatro categorías, respectivamente. Las categorías se forman utilizando en el primer caso cuatro valores; el mínimo, el máximo y los percentiles $(\frac{1}{3})100$ ($P_{(\frac{1}{3})100}$) y $(\frac{2}{3})100$ ($P_{(\frac{2}{3})100}$), mientras que en el segundo se usan cinco valores; el mínimo, el máximo y los cuartiles. En ambos casos, estos valores están asociados con los componentes: estudio (*CICP1*) y escalafón (*CICP2*), del Índice de Capacidad Profesional (*ICP*). Las categorías para cada esquema de categorización, se muestran en el cuadro 1.

En el cuadro 2 se muestra el coeficiente kappa de Cohen estimado ($\hat{\kappa}$), el cual es obtenido mediante la expresión

$$\hat{\kappa} = \frac{P_o - P_c}{1 - P_c} \quad (4)$$

donde P_o representa la proporción de casos en los que los dos componentes acuerdan, proporción de acuerdo observado y, P_c la proporción de casos en la que los mismos acuerdan por casualidad, el cual es estima-

do asumiendo independencia en las clasificaciones realizadas por los dos componentes, es decir, $P_c = \sum_{i=1}^r P_{i+} P_{+i}$, donde r representa el número

de categorías en la tabla de contingencia (tablas A1 y A2 del apéndice), P_{i+} es la proporción de elementos clasificados en la categoría i de C1COD y P_{+i} la proporción de elementos clasificados en la categoría i de C2COD.

Cuadro 1. Esquemas de Categorización

| (1.1): Esquema de Tres Categorías (EC3) | | |
|---|--|---|
| Categoría | C1COD | C2COD |
| 1 | Docentes con nivel 1 de C1COD ($\text{Min}(\text{C1CP1}) \leq \text{C1CP1} < P_{(1/3)}^{100}(\text{C1CP1})$) | Docentes con nivel 1 de C2COD ($\text{Min}(\text{C1CP2}) \leq \text{C1CP2} \leq P_{(1/3)}^{100}(\text{C1CP2})$). |
| 2 | Docentes con nivel 2 de C1COD ($P_{(2/3)}^{100}(\text{C1CP1}) \leq \text{C1CP1} < P_{(2/3)}^{100}(\text{C1CP1})$) | Docentes con nivel 2 de C2COD ($P_{(1/3)}^{100}(\text{C1CP2}) \leq \text{C1CP2} < P_{(2/3)}^{100}(\text{C1CP2})$). |
| 3 | Docentes con nivel 3 de C1COD ($P_{(2/3)}^{100}(\text{C1CP1}) \leq \text{C1CP1} < \text{Max}(\text{C1CP1})$) | Docentes con nivel 3 de C2COD ($P_{(2/3)}^{100}(\text{C1CP2}) \leq \text{C1CP2} \leq \text{Max}(\text{C1CP2})$). |
| (1.2): Esquema de Cuatro Categorías (EC4) | | |
| Categoría | C1COD | C2COD |
| 1 | Docentes con nivel 1 de C1COD ($\text{Min}(\text{C1CP1}) \leq \text{C1CP1} < Q_1(\text{C1CP1})$) | Docentes con nivel 1 de C21COD ($\text{Min}(\text{C1CP2}) \leq \text{C1CP2} < Q_1(\text{C1CP2})$). |
| 2 | Docentes con nivel 2 de C1COD ($Q_1(\text{C1CP1}) \leq \text{C1CP1} < Q_2(\text{C1CP1})$) | Docentes con nivel 2 de C2COD ($Q_1(\text{C1CP2}) \leq \text{C1CP2} < Q_2(\text{C1CP2})$). |
| 3 | Docentes con nivel 3 de C1COD ($Q_2(\text{C1CP1}) \leq \text{C1CP1} < Q_3(\text{C1CP1})$) | Docentes con nivel 3 de C2COD ($Q_2(\text{C1CP2}) \leq \text{C1CP2} < Q_3(\text{C1CP2})$). |
| 4 | Docentes con nivel 4 de C1COD ($Q_3(\text{C1CP1}) \leq \text{C1CP1} \leq \text{Max}(\text{C1CP1})$) | Docentes con nivel 4 de C2COD ($Q_3 \leq \text{C1CP2} \leq \text{Max}(\text{C1CP2})$). |

C1COD y C2COD son la versión categorizada de C1CP1 y C1CP2. Min(v), Max(v), Q(v) y Pq(v) indican los valores de mínimo, máximo, cuartil r-ésimo y el percentil q-ésimo de la variable v (C1CP1 y C1CP2).

Podemos observar en la tabla 2 que el índice de acuerdo corregido por chance es mayor cuando se usa el esquema EC3, 0.4730. Usando la escala de valoración para $\hat{\kappa}$, propuesta por Landis y Koch (1977), podemos indicar que existe un acuerdo moderado entre *CICOD* y *C2COD* en la clasificación de los docentes, cuando usamos EC3 y un acuerdo regular cuando se usa EC4.

Ahora bien, la expresión de $\hat{\kappa}$ está dada por

$$\hat{\kappa} = \frac{n \sum_i n_{ii} - \sum_i n_{i+} n_{+i}}{n^2 - \sum_i n_{i+} n_{+i}} = \frac{\sum_i p_{ii} - \sum_i p_{i+} p_{+i}}{1 - \sum_i p_{i+} p_{+i}} = \frac{p_o - p_c}{1 - p_c} \quad (5)$$

donde n_{ij} es el elemento de la tabla de contingencia correspondiente a la fila i y la columna j , n_{i+} el número de elementos clasificados en la fila i , n_{+i} el número de elementos clasificados en la columna i y n el número total de elementos a clasificar.

De aquí podemos observar que $\hat{\kappa}$ puede expresarse como

$$\hat{\kappa} = \frac{p_o}{1 - p_c} - \frac{p_c}{1 - p_c} = \hat{\kappa}_o - \hat{\kappa}_c \quad (6)$$

donde $\hat{\kappa}_o$ representa el índice de acuerdo observado y $\hat{\kappa}_c$ el índice de acuerdo por chance o casualidad. En el cuadro 2 se muestran estos valores para los dos esquemas de categorización. En ambos casos, el acuerdo observado es aproximadamente dos veces más grande que el acuerdo esperado por casualidad.

Cuadro 2. Coeficiente Kappa de Cohen estimado

| Número de categorías | $\hat{\kappa}$ | $\hat{\kappa}_o$ | $\hat{\kappa}_c$ |
|----------------------|----------------|------------------|------------------|
| EC3 | 0.4730 | 0.9730 | 0.5000 |
| EC4 | 0.3408 | 0.6741 | 0.3333 |

Ahora bien, sea $\{m_j | (n_{ii}, F_k)\}; j = 1, \dots, k$; una partición numérica de n_{ii} tal que entre los n_{ii} acuerdos alcanzados existen m_j casos que poseen el atributo que caracteriza el k-ésimo nivel del factor F sujeto a la restricción que

$$\sum_{j=1}^k [m_j | (n_{ii}, F_k)] = n_{ii} ; i = 1, \dots, 3.$$

Los valores porcentuales correspondientes se pueden denotar por $p_j | (n_{ii}, F_k) = ([m_j | (n_{ii}, F_k)] / n_{ii}) \cdot (100)$. En el cuadro 3 se muestran las distintas partes de las particiones numéricas $\{m_j | (n_{ii}, F_k)\}$ y los valores porcentuales correspondiente a los factores Estado Civil, Categoría según el ascenso alcanzado, Estudios y Facultad, que caracterizan a los docentes de acuerdo a los esquemas de clasificación EC3 y EC4.

El cuadro 3 usa las siguientes abreviaturas para los nombres de las facultades, el escalafón profesoral y Estudios:

Arq = Arquitectura; Cie = Ciencias; Der = Derecho; FACES = Economía; Far = Farmacia; For = Forestal; Hum = Humanidades; Ing = Ingeniería; Med = Medicina; NURR = Núcleo Trujillo; NUT = Núcleo Táchira; Odo = Odontología; Vice = Vicerrectorado.

Instr = Instructor; Asis = Asistente; Agre = Agregado; Asoc = Asociado; Titular = Titular.

Lic = Licenciado; Esp = Especialización; Maes = Maestría; Doct = Doctorado.

Cuando se usa EC3, aquellos docentes que poseen valores de *CICP1* y *CICP2* inferiores al percentil $P_{(\frac{1}{3})100}$, son solteros, asistentes, de dedicación exclusiva, licenciados y pertenecen a las facultades de arquitectura, ingeniería, humanidades y Núcleo Táchira.

Los docentes que poseen valores de *CICP1* y *CICP2* superiores al percentil $P_{(\frac{2}{3})100}$ se caracterizan por ser casados, titulares, de dedicación exclusiva, doctores y pertenecen a las facultades de ciencias y Núcleo Trujillo.

Cuadro 3. Distribución de frecuencias absolutas y proporciones de acuerdos por categorías según esquemas de clasificación EC3 y EC4

| Factor | Nivel | Tres Categorías: (EC3) | | | | | | Cuatro Categorías: (EC4) | | | | | | | |
|--------------|------------|------------------------|-------|-------------|-------|-------------|-------|--------------------------|-------|-------------|-------|-------------|-------|-------------|-------|
| | | Categoría 1 | | Categoría 2 | | Categoría 3 | | Categoría 1 | | Categoría 2 | | Categoría 3 | | Categoría 4 | |
| | | Frec | % | Frec | % | Frec | % | Frec | % | Frec | % | Frec | % | Frec | % |
| Estado Civil | Casado | 172 | 37.89 | 206 | 69.36 | 323 | 78.02 | 108 | 34.95 | 76 | 63.87 | 124 | 70.45 | 237 | 77.96 |
| | Soltero | 275 | 60.57 | 75 | 25.25 | 63 | 15.22 | 197 | 63.75 | 40 | 33.61 | 41 | 23.30 | 47 | 15.46 |
| | Divorciado | 7 | 1.54 | 16 | 5.39 | 28 | 6.76 | 4 | 1.29 | 3 | 2.52 | 11 | 6.25 | 20 | 6.58 |
| Categoría | Instr | 105 | 23.13 | 4 | 1.35 | 0 | 0 | 100 | 32.36 | 2 | 1.68 | 5 | 2.84 | 0 | 0 |
| | Asis | 264 | 58.15 | 39 | 13.13 | 0 | 0 | 185 | 59.87 | 31 | 26.05 | 27 | 15.34 | 0 | 0 |
| | Agre | 85 | 18.72 | 112 | 37.71 | 31 | 7.49 | 24 | 7.77 | 58 | 48.74 | 0 | 0 | 13 | 4.28 |
| | Asoc | 0 | 0 | 134 | 45.12 | 93 | 22.46 | 0 | 0 | 26 | 21.85 | 118 | 67.05 | 64 | 21.05 |
| Estudios | Titular | 0 | 0 | 8 | 2.69 | 290 | 70.05 | 0 | 0 | 2 | 1.68 | 26 | 14.77 | 227 | 74.67 |
| | Lic | 290 | 63.88 | 26 | 8.75 | 0 | 0 | 238 | 77.02 | 9 | 7.56 | 0 | 0 | 0 | 0 |
| | Esp | 36 | 7.93 | 50 | 16.84 | 30 | 7.25 | 19 | 6.15 | 37 | 31.09 | 31 | 17.61 | 13 | 4.28 |
| | Maes | 117 | 25.77 | 199 | 67.00 | 142 | 34.30 | 48 | 15.53 | 71 | 59.66 | 108 | 61.36 | 101 | 33.22 |
| | Doct | 11 | 2.42 | 22 | 7.41 | 242 | 58.45 | 4 | 1.29 | 2 | 1.68 | 37 | 21.02 | 190 | 62.50 |
| | Arg | 52 | 11.45 | 17 | 5.72 | 18 | 4.35 | 42 | 13.59 | 9 | 7.56 | 7 | 3.98 | 9 | 2.96 |
| | Cie | 29 | 6.39 | 19 | 6.40 | 83 | 20.05 | 16 | 5.18 | 11 | 9.24 | 12 | 6.82 | 69 | 22.70 |
| | Der | 29 | 6.39 | 23 | 7.74 | 10 | 2.42 | 18 | 5.83 | 6 | 5.04 | 16 | 9.09 | 6 | 1.97 |
| | Faces | 31 | 6.83 | 21 | 7.07 | 17 | 4.11 | 19 | 6.15 | 8 | 6.72 | 13 | 7.39 | 10 | 3.29 |
| | Far | 32 | 7.05 | 14 | 4.71 | 12 | 2.90 | 21 | 6.80 | 4 | 3.36 | 9 | 5.11 | 7 | 2.30 |
| Facultad | For | 24 | 2.29 | 20 | 6.73 | 29 | 7.00 | 16 | 5.18 | 11 | 9.24 | 10 | 5.68 | 25 | 8.22 |
| | Hum | 45 | 9.91 | 30 | 10.10 | 59 | 14.25 | 31 | 10.03 | 9 | 7.56 | 26 | 14.77 | 43 | 14.14 |
| | Ing | 58 | 12.78 | 32 | 10.77 | 50 | 12.08 | 46 | 14.89 | 8 | 6.72 | 20 | 11.36 | 38 | 12.50 |
| | Med | 39 | 8.59 | 54 | 18.18 | 39 | 9.42 | 21 | 6.80 | 29 | 24.37 | 21 | 11.93 | 26 | 8.55 |
| | NURR | 28 | 6.17 | 41 | 13.80 | 64 | 15.46 | 13 | 4.21 | 11 | 9.24 | 28 | 16.91 | 47 | 15.46 |
| | NUT | 45 | 9.91 | 12 | 4.04 | 23 | 5.56 | 35 | 11.33 | 5 | 4.20 | 3 | 1.70 | 18 | 5.92 |
| | Odo | 40 | 8.81 | 12 | 4.04 | 4 | 0.97 | 31 | 10.03 | 8 | 6.72 | 9 | 5.11 | 0 | 0 |
| | Vice | 2 | 0.44 | 2 | 0.67 | 6 | 1.45 | 0 | 0 | 0 | 0 | 2 | 1.14 | 6 | 1.97 |

Frec y % indican m_j (n_{ij} , $F_{i\cdot}$) y p_j (n_{ij} , $F_{i\cdot}$) respectivamente. Fuente: Cálculos propios

Por otro lado, si se usa el esquema EC4, aquellos docentes que poseen valores de *CICP1* y *CICP2* inferiores al cuartil 1, son solteros, asistentes, de dedicación exclusiva, licenciados y pertenecen a las facultades de arquitectura, humanidades e ingeniería. Los docentes que poseen valores de *CICP1* y *CICP2* superiores al cuartil 3 se caracterizan por ser casados, titulares, de dedicación exclusiva, doctores y pertenecen a las facultades de ciencias, humanidades y Núcleo Trujillo.

2.2. Coeficiente Kappa condicional ajustado

Como se puede notar, el coeficiente kappa de Cohen nos permite medir acuerdo en la tabla de contingencia entera, es decir permite obtener acuerdo agregado como un solo valor del resumen global a través de todas las categorías. De esta forma no permite calcular el acuerdo en cada categoría separadamente. Es claro que el valor $\hat{\kappa}$ obtenido no implica que el acuerdo se comporte de la misma forma en todas las categorías, un coeficiente agregado alto por ejemplo, no significa que haya niveles altos de acuerdo en todas las categorías de clasificación.

Una posible medida para resolver esta situación es la propuesta por Coleman (1966) y Light (1969), citados por von Eye y Young (2005, p. 8) y denominada kappa condicional o parcial, definida para la categoría de evaluación i , como:

$$\kappa_i = \frac{(\pi_{ii}/\pi_{i+}) - \pi_{+i}}{1 - \pi_{+i}} \quad (7)$$

donde π_{ii} representa la probabilidad de que ambos componentes usen la categoría i , π_{i+} y π_{+i} son las probabilidades marginales asociadas con la fila y columna i de la tabla de clasificación. Nótese que,

$$\sum_{i=1}^r \kappa_i = \sum_{i=1}^r \frac{\pi_{ii} - \pi_{+i}}{\pi_{i+} - \pi_{+i}} = \sum_{i=1}^r \frac{\pi_{ii} - \pi_{+i}\pi_{+i}}{\pi_{i+} - \pi_{+i}\pi_{+i}} \neq \frac{\sum_{i=1}^r \pi_{ii} - \sum_{i=1}^r \pi_{+i}\pi_{+i}}{1 - \sum_{i=1}^r \pi_{+i}\pi_{+i}} = \kappa \quad (8)$$

Por lo tanto, este coeficiente no cumple con la propiedad de aditividad, es decir, su suma sobre todas las categorías no reproduce el kappa global.

Se propone en este artículo el uso del coeficiente kappa condicional ajustado, el cual para la categoría i se define como:

$$\kappa_i^* = \frac{(\pi_{ii} - \pi_{i+} \pi_{+i})}{1 - \sum_{i=1}^r \pi_{i+} \pi_{+i}} \tag{9}$$

Puede observarse que,

$$\kappa = \frac{\sum_{i=1}^I \pi_{ii} - \sum_{i=1}^I \pi_{i+} \pi_{+i}}{1 - \sum_{i=1}^I \pi_{i+} \pi_{+i}} = \frac{\sum_{i=1}^I (\pi_{ii} - \pi_{i+} \pi_{+i})}{1 - \sum_{i=1}^I \pi_{i+} \pi_{+i}} = \frac{(\pi_{11} - \pi_{1+} \pi_{+1}) + (\pi_{22} - \pi_{2+} \pi_{+2}) + \dots + (\pi_{II} - \pi_{I+} \pi_{+I})}{1 - \sum_{i=1}^I \pi_{i+} \pi_{+i}} \tag{10}$$

$$= \frac{(\pi_{11} - \pi_{1+} \pi_{+1})}{1 - \sum_{i=1}^I \pi_{i+} \pi_{+i}} + \frac{(\pi_{22} - \pi_{2+} \pi_{+2})}{1 - \sum_{i=1}^I \pi_{i+} \pi_{+i}} + \dots + \frac{(\pi_{II} - \pi_{I+} \pi_{+I})}{1 - \sum_{i=1}^I \pi_{i+} \pi_{+i}} = \kappa_1^* + \kappa_2^* + \dots + \kappa_I^* = \sum_{i=1}^I \kappa_i^*$$

De lo expuesto arriba puede verificarse que κ_i^* es una medida que cumple con la propiedad de aditividad y, por lo tanto, permite determinar porcentualmente el aporte de cada categoría de evaluación en el acuerdo global. Es una medida apropiada para efectos de comparación.

En cuadro 4 se presentan los valores estimados del kappa condicional ajustado, los del acuerdo observado (κ_{io}^*) y esperado por casualidad (κ_{ic}^*) para cada categoría, así como el aporte de cada categoría al acuerdo global.

Cuadro 4. Coeficiente Kappa Condicional Ajustado

| i | $\hat{\kappa}_i^*$ | $\hat{\kappa}_{io}^*$ | $\hat{\kappa}_{ic}^*$ | Aporte(%) | i | $\hat{\kappa}_i^*$ | $\hat{\kappa}_{io}^*$ | $\hat{\kappa}_{ic}^*$ | Aporte(%) |
|-------|--------------------|-----------------------|-----------------------|-----------|-----|--------------------|-----------------------|-----------------------|-----------|
| 1 | 0,2128 | 0,3791 | 0,1662 | 44,989 | 1 | 0,1463 | 0,2294 | 0,0831 | 42,92 |
| 2 | 0,0812 | 0,2481 | 0,1669 | 17,167 | 2 | 0,0050 | 0,0884 | 0,0833 | 1,47 |
| 3 | 0,179 | 0,3458 | 0,1669 | 37,844 | 3 | 0,0471 | 0,1307 | 0,0835 | 13,83 |
| | | | | | 4 | 0,1424 | 0,2257 | 0,0833 | 41,78 |
| Total | 0,473 | 0,973 | 0,5 | | | 0,3408 | 0,6742 | 0,3332 | 100,00 |

El acuerdo parcial asociado con la categoría 1, para los dos esquemas de categorización es el mayor, seguido por el nivel de acuerdo asociado con la última categoría.

Podemos observar que en cuanto al nivel de acuerdo observado, para todas las categorías, el mismo es mayor cuando usamos el EC3. Cuando se usa el EC4, para las categorías 1 y 4 el grado de acuerdo observado es aproximadamente tres veces más grande que el esperado por casualidad. Para el caso de tres categorías, para las categorías 1 y 3 el grado de acuerdo observado es aproximadamente dos veces más grande que el esperado por casualidad. Para aquellas categorías en las que se dan los valores de coeficiente kappa más pequeños, se puede observar que la diferencia entre el nivel de acuerdo observado y el esperado por casualidad es también menor.

Ahora bien, si se desea probar la hipótesis $H_0 : F(\pi) = X\beta = 0$, donde $\pi' = (\pi_{11}, \pi_{12}, \dots, \pi_{II})$ es el vector de probabilidades asociado con la tabla de contingencia, la metodología de mínimos cuadrado ponderados puede ser usada, considerando como la matriz de diseño, la matriz nula, $X = 0$, (Grizzle, Starmer y Koch, 1969). El estadístico de prueba para esta hipótesis es el dado por

$$W = F(p)S^{-1}F(p) \quad (11)$$

el cual bajo H_0 sigue una distribución asintótica Chi-cuadrado con u grados de libertad y donde $p = (p_{11}, p_{12}, \dots, p_{II})$ representa el vector de proporciones muestral, el cual estima a p . Su valor para el esquema EC3 es $W = 7882031$. Para 3 grados de libertad, este valor indica que el modelo es significativo ($p < 0.001$).

Para probar la significación estadística de κ_i^* , $i = 1, 2, 3$, debe usarse el estadístico

$$Z_i = \frac{\hat{k}_i - f_i(\pi)}{(S_{ii})^{1/2}} \sim N(0,1) \quad (12)$$

El valor de Z para las tres categorías se muestra en la tabla 5. En todos los casos, los valores sugieren significación estadística ($p < 0.001$). Esta tabla muestra además los límites superior e inferior de intervalos de confianza del 99% para κ_i^* .

Cuadro 5. Valores de Z para el Coeficiente Kappa Condicional Ajustado. Bajo el esquema de categorización EC3.

| Categoría | Z calculado | Intervalo de Confianza del 99% | |
|-----------|---------------|--------------------------------|-----------------|
| | | Limite inferior | Limite Superior |
| 1 | 22,35 | 0,188 | 0,237 |
| 2 | 7,24 | 0,052 | 0,11 |
| 3 | 21,64 | 0,158 | 0,2 |

Probar la hipótesis $H_0 : \kappa_1^* = \kappa_2^* = \kappa_3^*$ es equivalente a probar las hipótesis $H_{01} : L_1 F(P) = 0$ y $H_{02} : L_2 F(P) = 0$, donde $F(\pi)' = (\kappa_1^*, \kappa_2^*, \kappa_3^*)$, $L_1 = (2 \ -1 \ -1)$ y $L_2 = (0 \ 1 \ -1)$. Rechazar al menos una hipótesis de estas, es equivalente a rechazar $H_0 : \kappa_1^* = \kappa_2^* = \kappa_3^*$.

Los valores de Z asociados con las hipótesis H_{01} y H_{02} son obtenidos usando la relación

$$Z_i = \frac{L_i' F(P)}{(L_i' S L_i)^{1/2}} \sim N(0,1) \tag{13}$$

Los valores $Z_1 = 13,54$ y $Z_2 = 12,18$ permiten rechazar H_{01} y H_{02} ($p < 0.001$), y de esta forma, se rechaza H_0 . Esto indica que existen diferencias significativas entre los κ_i^* .

3. Conclusiones

La valoración del coeficiente Kappa de Cohen mediante la escala de Landis y Koch (1977) indica que el acuerdo se encuentra entre los niveles regular y moderado, para el uso de cuatro y tres categorías respectivamente.

Para ambos esquemas de categorización, el nivel de acuerdo observado es aproximadamente dos veces más grande que el esperado por casualidad.

Al calcular el coeficiente kappa condicional ajustado, puede notarse que la categoría 1, para ambos esquemas es la que mayor aporte hace al acuerdo general, 45.01% para el esquema de tres categorías y 42,92% para el esquema de cuatro categorías.

El valor del coeficiente kappa para los dos esquemas de categorización utilizados en este estudio varía entre los niveles regular y moderado sin que sea excelente, indica que existe una proporción importante de profesores tal que aun cuando un nivel alto de estudio sea alcanzado por ellos, esto no se corresponde similarmente con una mejor ubicación de ellos en el escalafón y viceversa. Esta situación señala la presencia de un tipo de riesgo académico institucional y confirma de manera complementaria una conclusión similar reportada por Sinha *et. al.* (2007) acerca de los profesores de 4 de las 13 facultades de la Universidad de Los Andes.

Para el esquema de categorización de tres categorías, aquellos docentes con valores de $CICP1$ y $CICP2$ inferiores al percentil $P_{(1/3)100}$ se caracterizan por ser solteros, asistentes, de dedicación exclusiva, licenciados y pertenecen a las facultades de arquitectura, ingeniería, humanidades y Núcleo Táchira. Los docentes que poseen valores de $CICP1$ y $CICP2$ superiores al percentil $P_{(2/3)100}$ se caracterizan por ser casados, titulares, de dedicación exclusiva, doctores y pertenecen a las facultades de ciencias y Núcleo Trujillo.

Para el esquema correspondiente a cuatro categorías, aquellos docentes que poseen valores de $CICP1$ y $CICP2$ por debajo del cuartil 1 son solteros, asistentes, de dedicación exclusiva, licenciados y pertenecen a las facultades de arquitectura, humanidades e ingeniería. Los docentes que poseen valores de $CICP1$ y $CICP2$ superiores al cuartil 3 se caracterizan por ser casados, titulares, de dedicación exclusiva, doctores y pertenecen a las facultades de ciencias, humanidades y núcleo Trujillo.

Dados los resultados obtenidos en este trabajo, podemos indicar que el coeficiente kappa de Cohen puede ser aplicado en situaciones no clásicas tal como se hizo en este trabajo con el objeto de medir el grado de acuerdo entre dos indicadores ($CICP1$ y $CICP2$) para clasificar profesores de una institución. El coeficiente de kappa de Cohen ha sido usado más comúnmente en forma clásica en sociología y medicina para medir el acuerdo entre los evaluadores.

4. Referencias

- Cicchetti, Domenic V., y Feinstein, Alvan R. (1990). "High agreement but low Kappa: II. Resolving the Paradoxes." *Journal of Clinical Epidemiology*, 43, 6, pp. 551-558.
- Díaz M., Luis G. (2002). *Estadística Multivariada: inferencia y métodos*. Bogotá: Universidad Nacional de Colombia. Facultad de Ciencias, Departamento de Estadística.
- Feinstein, Alvan R., y Cicchetti, Domenic V. (1990). "High agreement but low Kappa: I. The problems of two Paradoxes." *Journal of Clinical Epidemiology*, 43, 6, pp. 543-549.
- Fleiss, J. L., Levin, Bruce, & Paik, Myunghee Cho (2003). *Statistical methods for rates and proportions*. Third edition. New York: Wiley.
- Grizzle, James E., Starmer, C. Frank, y Koch, Gary G. (1969). "Analysis of categorical data by linear models." *Biometrics*, 25, pp. 489-504.
- Landis, R. J., y Koch, G. G. (1977). "The measurement of observer agreement for categorical data." *Biometrics*, 33, pp.159-174.
- Stokes, Maura E., Davis, Charles S., y Koch, Gary G. (2000). *Categorical data analysis using the SAS system*. Second edition. North Carolina: SAS Institute Inc..
- Sinha, Surendra P., Ramoni, Josefa, Orlandoni, Giampaolo, Torres, Elizabeth y Figueroa, Miguel (2007). "Conceptuación y Análisis Descriptivo del Riesgo Académico Institucional en las Universidades Nacionales Venezolanas: El Caso de la ULA." Mérida [Venezuela]: Universidad de Los Andes. (aceptado para su publicación en *EDUCERE*).
- Von Eye, Alexander, y Young Mun, Eun (2005). *Analyzing rater agreement: Manifest variable methods*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

5. Apéndice

Cuadro A1. Contingencia para el esquema de categorización EC3

*(C1COD * C2COD)*

| Categoría | 1 | 2 | 3 |
|-----------|-----|-----|-----|
| 1 | 454 | 124 | 19 |
| 2 | 137 | 297 | 166 |
| 3 | 8 | 177 | 414 |

Cuadro A2. Contingencia para el esquema de categorización EC4

*(C1COD * C2COD)*

| Categoría | 1 | 2 | 3 | 4 |
|-----------|-----|-----|-----|-----|
| 1 | 309 | 118 | 20 | 1 |
| 2 | 131 | 119 | 112 | 87 |
| 3 | 9 | 208 | 176 | 57 |
| 4 | 0 | 4 | 141 | 304 |