Revista de la Facultad de Ciencias Universidad Nacional de Colombia, Sede Medellín
V 10 N°1 enero-junio de 2021 • ISSN-e 2357-5749 • Artículo Investigación • Páginas 87 a 100
DOI: https://doi.org/ 10.15446/revfaccienc.v10n1.84225

ÍNDICE PARA COMPARAR METODOLOGÍAS DE ESTIMACIÓN POR INTERVALOS PARA LAS TASAS ESTANDARIZADAS DIRECTAS[a]

INDEX FOR COMPARING INTERVAL ESTIMATION METHODOLOGIES FOR DIRECT STANDARDIZED RATES

HENRY HUMBERTO OROZCO QUICENO[b], JUAN CARLOS SALAZAR URIBE[c], JUAN CARLOS CORREA MORALES[d]

Recibido para revisar 22-04-2020, aceptado 27-06-2020, versión final 17-12-2020.
Artículo Investigación

RESUMEN: Las tasas estandarizadas se utilizan para comparar un evento en una población de estudio con una población estándar controlando los efectos confusores y además permiten dimensionar el evento de acuerdo a un mismo estándar. En este trabajo se evalúan con un estudio de simulación y por medio de un nuevo índice, los desempeños de varias metodologías para la estimación de las tasas estandarizadas a través de intervalos de confianza, algunas basadas en la distribución normal, Poisson y gamma. Este nuevo índice permite considerar tanto el nivel de cobertura como la amplitud del intervalo y evaluar la calidad de cada uno de los intervalos evaluados conjuntamente. Se identifican los métodos de intervalos de confianza con mejor desempeño de acuerdo al índice propuesto, entre estos se destacan los métodos gamma, el de momentos y bootstrap.

PALABRAS CLAVE: Tasas estandarizadas; epidemioloía; confusión; intervalos de confianza; bioestadística.

ABSTRACT: The standardized rates are used to compare an event in a study population with a standard population controlling the confounding effects and also allow the event to be dimensioned according to the same standard. In this paper, the performance of several methodologies for the estimation of standardized rates through confidence intervals, some based on the normal distribution, Poisson and gamma, are evaluated with a simulation study and by means of a new index. This new index allows us to consider both the level of coverage and the amplitude of the interval and evaluate the quality of each of the intervals evaluated jointly. The methods of confidence intervals with better performance according to the proposed index are identified, among these are the gamma, moments and bootstrap methods.

KEYWORDS: Standardized rates; epidemiology; confussion; confidence intervals; biostatistics.

---

[a]Henry Humberto Orozco-Quiceno, H. H., Salazar-Uribe, J. C. & Correa-Morales, J. C. (2021). Índice para comparar metodologías de estimación por intervalos para las tasas estandarizadas directas *Rev. Fac. Cienc.*, 10 (**1**), 87–100. DOI: https:// 10.15446/revfaccienc.v10n1.84225

[b]M. Sc. en Estadística, Escuela de estadística, Universidad Nacional Sede Medellin.
[c]Profesor Titular Escuela de Estadística, Universidad Nacional Sede Medellin.
[d]Profesor Asociado, Escuela de Estadística, Universidad Nacional Sede Medellin.

1. INTRODUCCIÓN

Las tasas estandarizadas son utilizadas para comparar un evento en diferentes poblaciones, controlar los efectos de otras variables y la composición de las poblaciones. permitiendo dimensionar el evento de acuerdo a un mismo estándar o una misma población, para identificar su magnitud o riesgo de acuerdo a los factores de cada grupo poblacional. Para su estimación a través de intervalos de confianza, existen diferentes metodologías, basadas principalmente en la distribución Poisson (Dobson *et al.*, 1991) y metodologías basadas en las distribuciones normal, gamma, beta y además se utilizan metodologías bayesianas. Todas ellas son de gran utilidad para la estimación de las tasas estandarizadas directas que generan intervalos de diferentes calidades, tanto a nivel de cobertura como en amplitud.

En este trabajo se comparan, usando un estudio de simulación por medio de un índice propuesto, algunas metodologías para la estimación de tasas estandarizadas directas por medio de intervalos de confianza, partiendo de propuestas que han tenido un buen desempeño, según los resultados de estudios previos, como son las metodologías para intervalos de confianza por aproximación bootstrap (DiCiccio & Efron, 1996), metodología basada en los momentos (Dobson *et al.*, 1991), métodos basados en la distribución gamma (Fay & Feuer, 1997). En esta comparación se utiliza un índice propuesto basado en una propuesta hecha por Correa & Sierra (2003), para la diferencia de proporciones, que permite medir la calidad de los intervalos en términos de cobertura y amplitud.

La existencia de muchas metodologías para estimar intervalos de confianza para las tasas estandarizadas de mortalidad y morbilidad (Keung & Filardo, 2008), con múltiples coberturas y desempeños de acuerdo a diferentes escenarios, hace que cada método no sea universal ni se pueda aplicar siempre, además cuentan con niveles de amplitud y coberturas variables (Correa & Sierra, 2003), por lo que se debe tener en cuenta las características de cada escenario y las limitaciones de cada método. Aplicando el índice propuesto, se pretende caracterizar las diferentes metodologías para la estimación por medio de intervalos de confianza para las tasas estandarizadas directas identificando escenarios donde cada intervalo muestra un mejor desempeño.

2. TASAS ESTANDARIZADAS DIRECTAS

En epidemiología se utilizan las tasas para medir la presencia de un evento en una población de interés, las cuales podrán ser muertes por una causa dada, mortalidad, la incidencia o prevalencia, o morbilidad. La estandarizacion permite comparar el evento de interes en diferentes poblaciones, controlar los efectos de otras variables y la composición de las poblaciones. Se pueden comparar las tasas de mortalidad dada una causa específica, tasas de incidencia o prevalencia de una enfermedad estratificadas por características demográficas como edad o género, las cuales se comparan con una población estándar o global, por ejemplo la de un país para comparar el evento en sus subregiones o la población mundial para comparar regiones o paises, para la cual la Organización Mundial de la Salud (OMS) ha propuesto estructuras de población global que se actualizan regularmente (Ahmad *et al.*, 2001). Adicionalmente, para una mejor interpretación, este se multiplica por una potencia de 10 con valores como 1.000, 10.000 y 100.000 entre otros. Esto permite evidenciar la presencia del evento en diferentes poblaciones y cuantificar su magnitud eliminando los efectos de factores externos que distorsionan o confunden, dando lugar a una comparación más clara y comprensible. Para su estimación se consideran dos formas: directa e indirecta. En las tasas directamente estandarizadas (DSR), se parte del cálculo de los pesos del evento de interés en la población bajo estudio y se calculan las tasas aplicando esos pesos a una población estándar (Woodward, 2005). La estructura de los datos son como se ilustra la Tabla 1, para la estandarización indirecta se calcula la tasa global y se multiplica por el número de eventos en la población bajo estudio.

Tabla 1: Estructura de datos para el cálculo de las tasas estandarizadas directas

Grupos de edad	Eventos	Población estudio	Población estándar
$Grupo_1$	λ_1	p_1	p_1^s
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Grupo_k$	λ_k	p_k	p_k^s

La tasa se estima con la siguiente expresión

$$\mathrm{DSR}_x = \sum_{i=1}^{k} w_i \lambda_i \tag{1}$$

Donde $w_i = \frac{p_i^s}{p_i}$ son los pesos de la población estándar del grupo de edad i, y λ_i el promedio de eventos en la población bajo estudio en un periodo de tiempo del grupo de edad i, y k representa el número de grupos edad en que se subdivide la población en estudio.

La estandarización directa, Woodward (2005) la define como la sumatoria de las tasas de eventos en las edades específicas en la población bajo estudio divido por el tamaño de la población estándar. Para evitar números pequeños - que a veces son difíciles de interpretar - estos usualmente se multiplican por 1.000, 10.000 o 100.000, entre otros. Para este cálculo, como se presenta a continuación, Woodward (2005), lo denota de la siguiente manera: Sean e_i número de eventos del i-ésimo grupo de edad de la población estudio, p_i es el tamaño del i-ésimo grupo de edad de la población estudio, p_i^s es el tamaño del grupo de edad de la población estándar y $p^s = \sum_i p_i^s$ es el tamaño total de la población estándar. Entonces la estandarización directa de la tasa (dsr) del evento por mil personas es:

$$dsr = 1000 \sum_i \left(\frac{e_i}{p_i} \right) \frac{p_i^{(s)}}{p^{(s)}}, i = 1, \ldots, k \tag{2}$$

Si se asume que el número de eventos observados e_i tienen una distribución Poisson, entonces el error estándar de la tasa directamente estandarizada se define así:

$$se\,(dsr) = \frac{1000}{p^{(s)}}\left\{\sqrt{\sum_i e_i\left(\frac{p_i^{(s)}}{p_i}\right)^2}\right\} \tag{3}$$

Un IC para la tasa directamente estandarizada, con una confianza del 95 %, es

$$dsr \pm 1{,}96se\,(dsr) \tag{4}$$

Según lo definido por Dobson *et al.* (1991), igual que Swift (2010), plantea que las tasas estandarizadas de mortalidad son sumas ponderadas de los parámetros Poisson, esto lo muestra a través de un estudio de simulación y aplicando un ejemplo de las tasas de mortalidad por cáncer con los datos de la OMS del proyecto 8 MONICA (Ahmad *et al.*, 2001). También Woodward (2005), considera que los eventos se distribuyen Poisson, estos coinciden con Keung & Filardo (2008), que realiza un estudio partiendo de este planteamiento y compara 20 métodos para estimar las tasas por medio de IC y con Swift (Swift, 2010), donde en su estudio para comparar los métodos que mejor desempeño tuvieron del trabajo de Keung & Filardo (2008), plantea, dados X_i, $i = 1, 2, 3, \ldots, k$, variables aleatorias que representan el número de eventos ocurridos en el grupo i, los $X_i, \ldots, X_k$ son independientes y se distribuyen Poisson con media $n_i\lambda_i$, donde λ_i es la tasa de eventos y n_i es el número de personas año por cada grupo i, si se define $\theta = n_i\lambda_i$ entonces $X_i \sim$Poisson (θ_i). Por lo tanto las tasas estandarizadas se definen así:

$$R = \sum_{i=1}^{k}\frac{c_i}{n_i}\theta_i \tag{5}$$

$$= \sum_{i=1}^{k} w_i\theta_i \tag{6}$$

Donde $w_i = \frac{c_i}{n_i}$. Los c_i son el número de personas año de la población estándar, constantes conocidas las cuales se usan como se había mencionado antes, llevada a una potencia de 10 para mejorar su interpretación, $i = 1, 2, 3, \ldots, k$.

Las tasas estandarizadas directas estimadas ajustadas por edad:

$$\hat{R} = \sum_{i=1}^{k} w_i x_i \tag{7}$$

donde los x_i son el número de eventos en $i = 1, 2, 3, \ldots, k$ (k grupos de edad independientes), $w_i = \dfrac{c_i}{p_i}$, c es una constante conocida de la población estándar mundial (Ahmad *et al.*, 2001) y p_i número de años persona en la población bajo estudio.

Interesa estimar un IC para el parámetro θ, usando diferentes metodologías, donde la calidad del intervalo está dada por las propiedades de sus extremos X_L (extremo inferior), X_U (extremo superior) y de que su amplitud sea mínima con probabilidad de cobertura lo más cerca al nivel nominal. Partiendo de los métodos seleccionados que mejor desempeño presentaron en los trabajos de Keung & Filardo (2008) y Swift (2010), los cuales se presentan a continuación, para estimar los intervalos de confianza para las tasas estandarizadas de morbi-mortalidad, se pretende comparar estos métodos, midiendo su desempeño, por medio de un índice que se propone y que tiene en cuenta, de manera simultánea, la amplitud y el nivel de cobertura del intervalo.

2.1. Intervalos de confianza bootstrap para las tasas estandarizadas

Estos IC se obtienen usando la metodología propuesta por DiCiccio (1991), Efron (1993), DiCiccio & Efron (1995) y DiCiccio & Efron (1996), basada en el método bootstrap. Para los intervalos de confianza, Swift (1995), propuso el siguiente modelo para las tasas estandarizadas. Con un nivel de confianza de $(1-\alpha)\times 100\,\%$, se estiman los intervalos de confianza para θ así:

$$L = \hat{\theta} + \frac{z_0 - z_{1-\frac{\alpha}{2}}}{\left(1 - a\left[z_0 - z_{1-\frac{\alpha}{2}}\right]\right)^2}\hat{\sigma} \tag{8}$$

y

$$U = \hat{\theta} + \frac{z_0 - z_{1-\frac{\alpha}{2}}}{\left(1 - a\left[z_0 + z_{1-\frac{\alpha}{2}}\right]\right)^2}\hat{\sigma} \tag{9}$$

donde $z_{1-\frac{\alpha}{2}} = \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$ es el $\left(1-\frac{\alpha}{2}\right)$ percentil de la función de distribución normal.

2.2. Método de igualación de momentos (Moments matching method)

Este método fue propuesto por Dobson *et al.* (1991) para la obtención de los límites de confianza aproximados para la suma ponderada de los parámetros Poisson como funciones lineales de los límites de confianza para un único parámetro Poisson (citado en Swift (2010)), para lo cual propone: $X = \sum_{i=1}^{k} x_i$ haciendo coincidir el primer y el segundo momento obteniendo los extremos (X_L, X_U) que representa el intervalo de confianza de la suma de los parámetros Poisson $\left(\sum_{i=1}^{k}\theta_i\right)$ donde un intervalo de confianza aproximado de los parámetros para la suma de los pesos Poisson w_i del parámetro θ son (Keung & Filardo, 2008)

$$L = \hat{\theta} + \sqrt{\frac{\hat{V}}{X}}\,(X_L - X) \tag{10}$$

$$U = \hat{\theta} + \sqrt{\frac{\hat{V}}{X}}\,(X_U - X) \tag{11}$$

Existen muchos métodos que se utilizan para construir intervalos de confianza Poisson para calcular X_L, X_U. Swift (2010) recopila y simula 10 métodos diferentes donde los mejores desempeños son: $(M1)$ intervalo de confianza exacto, Johnson *et al*., 2005 (citado en Swift (2010)), sobre la base de la relación entre la distribución Poisson y la chi-cuadrado, se construye el intervalo de confianza para la suma ponderada de los parámetros Poisson como:

$$X_L = \frac{1}{2}\chi^2_{\alpha/2}(2X) \tag{12}$$

$$X_U = \frac{1}{2}\chi^2_{1-\alpha/2}(2(X+1)) \tag{13}$$

donde $\chi^2_q(d)$ es el cuantil de la distribución chi-cuadrado con d grados de libertad.

El $(M8)$ aproximación del intervalo de confianza del mid-p (Dobson *et al.*, 1991).

$$X_L = \left(X+\frac{1}{2}\right)\left[1-\frac{1}{9\left(X+\frac{1}{2}\right)}-\frac{Z_{1-\alpha/2}}{\sqrt[3]{X+\frac{1}{2}}}\right]^{-3} \tag{14}$$

$$X_U = \left(X+\frac{1}{2}\right)\left[1-\frac{1}{9\left(X+\frac{1}{2}\right)}+\frac{Z_{1-\alpha/2}}{\sqrt[3]{X+\frac{1}{2}}}\right]^{-3} \tag{15}$$

2.3. Métodos basados en la distribución gamma

Fay & Feuer (1997) proponen una aproximación para los intervalos de confianza para las tasas estandarizadas basada en la distribución gamma. Los límites de la gamma son equivalentes a la forma de la distribución chi-cuadrado, (Keung & Filardo, 2008; Swift, 2010) simulan como método $(G1)$, el intervalo de confianza con extremos X_L, X_U respectivamente del $(1-\alpha)\times 100\%$, así:

$$L = \frac{\hat{V}}{2\hat{\theta}}\left(\chi^2\right)^{-1}_{d_0}(\alpha/2), \text{ con } d_0 = \left(\frac{2\hat{\theta}^2}{\hat{V}}\right) \tag{16}$$

$$U = \frac{\hat{V}+W^{*2}}{2\left(\hat{\theta}+W^*\right)}\left(\chi^2\right)^{-1}_{d_1}\left(1-\frac{\alpha}{2}\right), \text{ con } d_1 = \left(\frac{2\left(\hat{\theta}+W^*\right)^2}{\hat{V}+W^{*2}}\right) \tag{17}$$

La $\left(\chi^2\right)^{-1}_{d}$ es la inversa de la función de probabilidad chi-cuadrado con d grados de libertad. Existe un número infinito de opciones para W lo que resulta en diferentes límites de confianza superiores; Fay & Feuer (1997) seleccionan $W^* = max(w_1, w_2, w_3, \ldots, w_k)$ y muestran por simulación que estos intervalos son conservadores.

Tiwari *et al.* (2006) propuso una modificación al método G1 presentado en Swift (2010) basado en la corrección de continuidad, al asumir la distribución uniforme para todos los grupos. Para el intervalo de confianza de $(1-\alpha)100\%$ con extremos (L, U^*) donde L esta dado en el método G1 y U^* es:

$$U^* = \frac{\hat{V}^*}{2\hat{\theta}^*}\left(\chi^2\right)^{-1}_{d_2}\left(1-\frac{\alpha}{2}\right), \text{ con } d_2 = \left(\frac{2\hat{\theta}^{*2}}{\hat{V}^*}\right) \tag{18}$$

y $\hat{\theta}^* = \hat{\theta} + \frac{1}{k}\sum_{i=1}^{k} W_i$ y $\hat{V}^* = \hat{V} + \frac{1}{k}\sum_{i=1}^{k} W_i^2$. El cual es denominado por Swift (2010) como G4, aunque Tiwari *et al.* (2006) consideró el límite superior U con $W^* = \frac{1}{k}\sum_{i=1}^{k} W_i$, estos dos intervalos son muy similares y aunque existe un número infinito de opciones para W^*, este lo propuso Tiwari *et al.* (2006) con base en la idea de corrección por continuidad.

2.4. Método Bayesiano con distribuciones conjugadas Poisson y gamma

En el siguiente método para la estimación de la tasa bruta o global propuesta por Ross (2003), se utiliza la distribución de probabilidad Poisson,

$$f_{X|\lambda}(x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, 2, \ldots, \tag{19}$$

donde x es la medición del número de eventos y λ es el parámetro de la distribución. La distribución gamma es la a priori (no informativa) conjugada para la distribución de probabilidad Poisson. La distribución gamma tiene la siguiente forma:

$$f_\lambda(x) = \frac{\lambda^{\alpha-1}e^{-\lambda/\beta}}{\beta^\alpha}\Gamma(\alpha), \quad x > 0 \tag{20}$$

donde $\Gamma(\alpha)$ es la función gamma, donde para α un entero positivo $\Gamma(\alpha) = (\alpha-1)!$. Una forma de la distribución gamma puede ser definida con $\alpha = 1$ y $\beta = k_\beta$, donde k_β es algún valor muy grande. Con estos parámetros, la distribución a priori de las tasas corresponde a:

$$f_\lambda(x) = \frac{e^{-\lambda/k_\beta}}{k_\beta} \tag{21}$$

Para la estimación de los intervalos de confianza para las tasas basadas en la distribución Poisson, para una a priori no informativa se propone la distribución posteriori así:

$$f_{\lambda|x}(\lambda|x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad \lambda \in [0, \infty) \tag{22}$$

Cuando es una distribución gamma con parámetros $(x+1, 1)$ y con un α, el intervalo de confianza en términos de λ como lo propone para las tasas brutas Ross (2003) es el intervalo $[a, b]$ tal que

$$Pr\{\lambda \in [a,b]\,|x\} = 1/x! \int_a^b \lambda^x e^{-\lambda} d\lambda \tag{23}$$

para $1-\alpha$ asumiendo $0 \leq a \leq b$. El intervalo de confianza para la tasa es

$$[a/N, b/N] \tag{24}$$

Partiendo de lo propuesto anteriormente por Ross (2003), se propone para el caso de las tasas estandarizadas directas, considerando $N = w$ como el factor de estandarización según la población global y realizando la sumas de forma eurística de las tasas grupales para hallar los intervalos con extremos, inferior,

$$\sum_{i=1}^{k} (aw_i) \tag{25}$$

y superior,

$$\sum_{i=1}^{k}(bw_i) \tag{26}$$

donde w_i corresponde a los pesos de la población global sobre la población en estudio en cada uno de los k grupos.

Otra forma de ponderar las sumas de los extremos de los intervalos bayesianos, se basa en la propuesta de Ross (2003),

$$\left(\sum_{i=1}^{k} aw_i, \sum_{i=1}^{k} bw_i\right) \tag{27}$$

Se propone un método bootstrap para estimar las medianas en cada grupo de edad y luego se realizan las sumas obteniendo los extremos para los intervalos de confianza para la tasa estandarizada directa R.

3. ÍNDICE PROPUESTO PARA MEDIR LA BONDAD DE LOS INTERVALOS DE CONFIANZA

Una de las ideas más relevantes de este trabajo es la propuesta de un índice que incorpora la longitud del intervalo con el nivel de confianza real y que permite comparar intervalos de confianza para tasas estandarizadas. Se partirá del índice propuesto para la diferencia de proporciones por Correa & Sierra (2003), donde se propone que para evaluar los intervalos de confianza se deben considerar la precisión indicada por la longitud del intervalo y la probabilidad de cobertura, $P(L_{inf} \leq \pi_1 - \pi_2 \leq L_{sup})$. Estos dos criterios no se pueden analizar por separado porque de poco sirve un intervalo con probabilidad de cobertura alta si su longitud es muy grande o un intervalo con una longitud muy pequeña pero con probabilidad de cobertura muy baja. Idealmente se desea que los intervalos sean angostos y tengan probabilidad de cobertura muy cercana al nivel de confianza nominal.

Correa & Sierra (2003), proponen un índice para las proporciones donde se tiene en cuenta tanto la amplitud del intervalo como el nivel de cobertura por lo que coberturas superiores al nivel nominal y amplitudes mínimas generan un índice mayor, y amplitudes superiores o coberturas menores al nominal generan índices menores. Este índice busca penalizar los IC con mayores amplitudes y niveles de coberturas menores al nominal. El índice propuesto por Correa & Sierra (2003) para la diferencia de proporciones está definido de la siguiente forma,

$$I = \frac{2-LPI}{2} \times \frac{NR}{NN} \tag{28}$$

Donde LPI es la longitud promedio del intervalo , NR es el nivel de confianza real y NN es el nivel de confianza nominal. Este índice es útil en el caso de las diferencias de proporciones, por que la longitud del intervalo siempre está entre cero y dos.

Partiendo del índice anterior se propone un nuevo índice para evaluar los intervalos de confianza para estimar las tasa estandarizadas directas. Para el cual en vez de $I = \dfrac{2-LPI}{2}$, para la amplitud de un intervalo, partiendo del hecho que la amplitud de los intervalos de confianza de las tasas estandarizadas no esta acotada entre cero y 2 como en el caso de las diferencias de proporciones, se propone hacer una transformación de la amplitud del intervalo de confianza utilizando la función de probabilidad acumulada de una normal evaluadas en $\bar{\theta}$ (pues por el Teorema del Límite Central, la distribución asintótica de $\bar{\theta}$ es $N(0,1)$, y se calcula $1-W$ donde W está dado por el valor absoluto de las diferencias de la función de probabilidad acumulada de $\bar{\theta}_s$ menos la función de probabilidad acumulada de la distribución normal de $\bar{\theta}_i$ así,

$$W = \left|F\left(\bar{\theta}_s\right) - F\left(\bar{\theta}_i\right)\right| \tag{29}$$

Donde F es la f.d.a de una N(0,1),

$$F\left(\bar{\theta}\right) = \int_{-\infty}^{\theta} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz \tag{30}$$

Como en el índice de Correa & Sierra (2003), la segunda componente $\dfrac{NR}{NN}$, en esta propuesta, sólo se considera el NR ya que cuando el NR supera al NN genera factores mayores a 1 y multiplicado por el factor de amplitud el índice pierde consistencia, quedando el índice propuesto en este trabajo de la siguiente forma:

$$I = (1-W) \times NR \tag{31}$$

Este índice tiene la propiedad de que intervalos pequeños generan unos valores mayores en la diferencia de la amplitud del intervalo de confianza y al multiplicarlo por los niveles reales generan un mayor valor en el índice. Por lo contrario, para intervalos muy amplios se penaliza el índice generando valores menores. También cuando los niveles reales son superiores al nivel nominal aportan a un mejor índice y en el sentido contrario generan un menor valor en este factor disminuyendo el índice, lo cual permite evaluar la calidad de los intervalos de confianza para las tasas estandarizadas directas como lo plantea Correa & Sierra (2003), tanto en precisión como en probabilidad de cobertura.

4. ESTUDIO DE SIMULACIÓN

Para la comparación via simulación de los métodos considerados para estimar por intervalos de confianza las tasas estandarizadas directas de un evento dado, se realizaron 1000 simulaciones. Partiendo de un λ, que representa la tasa promedio de un evento dado, se generan con una distribución Poisson con parámetro λ, eventos al azar para cada simulación, con estas se estiman los intervalos de confianza por cada método.

Tabla 2: Estructura de datos que se usó para las simulaciones de las tasas estandarizadas directas

Grupos de edad	media de eventos	población estudio	población estándar
$edad_{25-29}$	λ_1	p_1	ps_1
$edad_{30-34}$	λ_2	p_2	ps_2
$edad_{35-39}$	λ_3	p_3	ps_3
$edad_{40-44}$	λ_4	p_4	ps_4
$edad_{45-49}$	λ_5	p_5	ps_5
$edad_{50-54}$	λ_6	p_6	ps_6
$edad_{55-59}$	λ_7	p_7	ps_7
$edad_{60-64}$	λ_8	p_8	ps_8

Como se presenta en la Tabla 2 se agruparon los datos en quinquenios desde los 25 años a los 60 años, para un k de 8 grupos de edad, partiendo del hecho de que en epidemiología no siempre se realizan estudios sobre el total de la población de 0 a 100 años, comúnmente se desarrollan estudios con subgrupos poblacionales divididos por grupos de edad o por edad y género o también es el caso que se realizan estudios que abarcan todos los grupos poblaciones desde 0 a 100 años. Se tomó como población estándar, la población SEGI (Ahmad *et al.*, 2001)propuesta por la OMS, la cual define las proporciones de los tamaños de población estándar por cada grupo de edad por quinquenios desde los 0 años hasta los 85 años y más, pero para las simulaciones se tomaron las proporciones de los grupos de edad de 25 a 65 equivalentes a los grupos de la problación estudio simulada. Se repite este proceso y se cálculan 1000 índices, en cada una de los tres escenarios propuestos, para evaluar el comportamiento de los diferentes métodos en la estimación de las tasas estandarizadas directas por intervalos de confianza, bootstrap, momentos, gamma y dos intervalos bayesianos partiendo de una a priori no informativa gamma.

Con los índices por cada uno de los métodos y en los tres escenarios propuestos, se realiza un análisis descriptivo e inferencial para evaluar las posibles diferencias entre ellos, adicionalmente con simulaciones de 100 intervalos de confianza por cada método y en los tres escenarios se realiza el análisis descriptivo de los intervalos de manera convencional, evaluando su amplitud y nivel de cobertura, por último se ilustran los métodos usando datos de cáncer de mama del año 2015 en mujeres mayores a 20 años comparando la tasas de mortalidad en tres regiones del país: Antioquia, Valle del Cauca y la ciudad de Bogotá; se tomó la información de registros de defunción publicados por el DANE (2017).

5. RESULTADOS

En las simulaciones se evidenció que los intervalos de confianza para las tasas estandarizadas directas por los métodos bootstrap, momentos y gamma son más consistentes al tamaño de las tasas, y son similares entre sí. Lo que no sucede con los métodos bayesianos, en los cuales para el método bayesiano heurístico, las simulaciones tienen un comportamiento similar, homogéneo, según las Tabla 3. Basados en la media y mediana se evidencia el comportamiento de los métodos en las tasas simuladas. Se resalta el método gamma que representa el intervalo de confianza con mayor valor en promedio del índice, lo que indica un mejor desempeño.

Tabla 3: Valores descriptivos para simulación de tasas estandarizadas directas

Método	N	Media	Mediana	Desviación	Mín	Máx
Bootstrap	1000	0,949	0,950	0,015	0,900	0,990
Gamma	1000	0,952	0,950	0,015	0,900	0,990
Momentos	1000	0,951	0,950	0,015	0,900	0,990
Bayes	1000	0,940	0,942	0,014	0,886	0,969
Bayes bootstrap	1000	0,932	0,935	0,017	0,852	0,967

Partiendo de un análisis inferencial para establecer si los índices de cada método de los intervalos de confianza para las tasas estandarizadas directas tienen diferencias, se realizó un ANOVA de una vía, el cual permitió explorar si existen diferencias significativas entre los métodos en el escenario simulado, correspondientes a una tasa de eventos promedio. De acuerdo al box-plot asociado a los métodos, parecen existir diferencias estadísticas importantes. Esto se corrobora con el ANOVA de una vía a través de componentes múltiples donde las pruebas presentan valores de significancia menores a 0,0001.

Tabla 4: Resultados de los métodos considerados basados en el test de Kruskal-Wallis

Test de Kruskal-Wallis		
Método	Chi-cuadrado*	Valor p
boostrap-momentos	77,5854	$<,0001$
bootstrap-gamma	19,1394	$<,0001$
bootstrap-bayesiano	152,9947	$<,0001$
gamma-momentos	31,7622	$<,0001$
gamma-bayesiano	277,4753	$<,0001$
momentos-bayesiano	232,5141	$<,0001$
bayesiano b-momentos	561,5546	$<,0001$
bayesiano b-bootstrap	462,4739	$<,0001$
bayesiano b-gamma	610,3963	$<,0001$
bayesiano b-bayesiano	126,9932	$<,0001$

*bayesiano b: bayesiano bootstrap, ** grados de libertad=1

Tabla 5: Resultados de los métodos considerados basados en el test de Kruskal-Wallis

Test de Kruskal-Wallis		
Método	Chi-cuadrado*	Valor p
Simulación 2	1003,1595	$<,0001$

*grados de libertad = 4

En las Tablas 4 y 5, se presentan los resultados de las medias de los IC para simulaciones de 100 intervalos, la cobertura y la amplitud de los intervalos para una tasa, partiendo de esta información se corrobora lo identificado en el análisis anterior donde se describen cada uno de los índices resultado de los métodos para estimar por intervalos las tasas estandarizadas directas. Con respecto a la cobertura se identifican que los intervalos con la metodología bayesiana y la bayesiana bootstrap presentan buenos niveles de cobertura, también presenta un buen desempeño en cobertura el método gamma el cual tiene unos valores superiores a los métodos de momentos y bootstrap los cuales tienen resultados muy consistentes

Tabla 6: Promedio de Intervalos de confianza por método

	Límite inferior	Límite superior	Cobertura	Amplitud
Bootstrap	15.460	25.851	0.950	10.391
Momentos	15.536	25.928	0.951	10.392
Gamma	15.710	26.084	0.952	10.374
Bayes	14.957	27.749	0.974	12.792
Bayes bootstrap	14.381	26.987	0.965	12.606

Con respecto a la amplitud de los intervalos para los casos bayesiano y bayesiano bootstrap, estos tienen una amplitud mayor en los tres escenarios (ver Tabla 6), pero como se especificó anteriormente tenían valores de coberturas superiores lo cual hace que genere un índice menor y no sean competidores con los otros tres métodos. Para el caso de los métodos boostrap, momentos y gamma, de los cuales este último registra valores menores de amplitud de intervalo que complementados con unos valores mayores en cobertura con respecto al bootstrap y al método de los momentos genera valores mayores del índice en los tres escenarios, representando el método con mejor desempeño. Los métodos bootstrap y momentos tienen unos valores de amplitud y cobertura conservadores en los diferentes escenarios pero son ligeramente menores al método gamma. Los métodos bayesianos entre sí comparten unos desempeños similares tanto en amplitud como en cobertura.

6. CONCLUSIONES

Consecuente con lo reportado en diferentes estudios, el índice propuesto mostró el resultado de los desempeños de cada uno de los métodos para estimar por intervalos de confianza las tasas estadarizadas directas. Según lo reportado por Swift (2010) y Keung & Filardo (2008), de los métodos analizados el que mejor desempeño presentó fue el gamma, igual resultado se obtuvo en el presente trabajo, donde lo analizado con el índice y en el análisis convencional de los intervalos por cada uno de los métodos, genera un mayor valor para los intervalos de confianza del método gamma, aunque separadamente se evidenció que los bayesianos contaban con un nivel de cobertura mayor pero igualmente un mayor valor de amplitud lo que hace que estos métodos sean poco conservadores.

Por otra parte para los métodos bootstrap y momentos, el índice presenta unos resultados para los intervalos de confianza de las tasas estandarizadas más conservadores, donde tanto los niveles de cobertura y amplitud tienen valores cercanos al método gamma, convirtiéndose en buenos competidores con el intervalo de confianza del método gamma y superiores a los métodos bayesianos.

El índice propuesto presentó un desempeño consistente y captura la calidad de los intervalos de confianza de las metodologías evaluadas permitiendo identificar los mejores desempeños tanto en cobertura como amplitud conjuntamente.

Referencias

Ahmad, O., Boschi-pinto, C., Lopez, A., Christopher J., Rafael L. & Mie, I. (2001). Age standardization of rates: a new WHO standard GPE Discussion. *Paper Series*, 31, 1-14.

Correa, J. & Sierra, E. (2003). Intervalos de confianza para la comparación de dos proporciones. *Revista Colombiana de Estadística*, 26, 1, 61-75.

DANE. (2017). Estadísticas vitales, [Consultada en julio de 2017].

Diciccio, T. (1991). *More Accurate Confidence Intervals in Exponential Families*. Stanford University.

Diciccio, T. & Efron, B. (1995). *Bootstrap confidence intervals*. Stanford University.

DiCiccio, T. & Efron B. (1996). Bootstrap Confidence Intervals. *Statistics In Medicine*, 11, 189-212.

Dobson, A., Kuulasmaa, K., Eberle, E. & Scherer, J. (1991). Confidence Intervals For Weighted Sums Of Poisson Parameters. *Statistics In Medicine*, 10, 457-462.

Efron, B. (1993). *An Introduction to the Bootstrap*. Chapman y Hall/CRC

Fay, M. & Feuer, E. (1997). Confidence Intervals For Directly Standardized Rates: A Method Based On The Gamma Distribution. *Statistics In Medicine*, 16, 791-801.

Johnson, N., Kemp, A.& Kotz, S. (2005). *Univariate Discrete Distributions*. John Wiley y Sons, New York

Keung, H. & Filardo Zheng, G. (2008). Confidence interval estimating procedures for standardized incidence rates. *Computational Statistics and Data Analysis*, 52, 3501-3516.

Ross, T. (2003). Acurate confidence intervals for binomial proportion and Poisson rate estimation. *Computers an Biology and Medicine*, 10, 1080

Swift, M. (1995). Simple confidence intervals for standardized rates based on the approximate bootstrap method. *Statistics in Medicine*, 14, 1875-1888.

Swift, M. (2010). A simulation study comparing methods for calculating confidence intervals for directly standardized rates. *Computational Statistics and Data Analysis*, 54, 1103-1108.

Tiwari, R., Clegg, L. & Zou, Z. (2006). Efficient interval estimation for age-adjusted cancer rates. *Statistical Methods in Medical Research*, 15, 547-567.

Woodward, M. (2005). *Epidemiology study design and data analysis*. Florida: Chapman y Hall/CRC