UNIVERSIDAD NACIONAL DE COLOMBIA

# Comparación de diferentes metodologías para la estimación por intervalos para las tasas estandarizadas

# Contrast of different methodologies for interval estimation for standardized rates

**Henry Humberto Orozco Quiceno**

Universidad Nacional de Colombia
Facultad de Ciencias, Escuela de estadística
Medellin, Colombia
2018

# Comparación de diferentes metodologías para la estimación por intervalos para las tasas estandarizadas

**Henry Humberto Orozco Quiceno**

Tesis presentada como requisito parcial para optar al título de:
**Magister en Estadística**

Director:
Juan Carlos Salazar Uribe, Ph.D.

Línea de Investigación:
Bioestadística
Grupo de Investigación:
Grupo de investigación en estadística Universidad Nacional, Sede Medellin

Universidad Nacional de Colombia
Facultad de ciencias, escuela de estadística
Medellín, Colombia
2018

## Dedicatoria

A mis Hijos

A mis padres que desde niño sembraron este sueño.

# Agradecimientos

Agradezco en el alma a mi familia, en especial a mis hijos Karen, Leslie y Camilo, son la razón de este sueño, a Yesenia Grajales, mi novia, que ha sido un apoyo incondicional en todo este proceso, recuerdo el día que pasábamos junto a la universidad y decidimos que me presentaría a ver como era el examen y poder prepararlo para el próximo año y así convertir en una meta ese sueño aplazado y lejano de cursar la maestría.
También quiero agradecer y de forma muy especial a mi director de tesis Juan Carlos Salazar, por su dedicación, ideas, aportes y orientaciones, mas que el tutor es un amigo y compañero.
A la Universidad Nacional Sede Medellin por esta gran oportunidad, la escuela de estadística y sus profesores en especial a los profesores Juan Carlos Correa por sus aportes, Isabel Cristina Ramiréz Guevara, Norman Diego Giraldo, Fredy Hernandez Barajas por todo lo aportado durante mi formación.
Agradezco a todas las personas que aportaron su granito de arena, con un comentario, con una frase de motivación, a mis amigos que no los acompañé en este tiempo en las carreras de aventura, cosa que extrañé, la nostalgia que da ver la bici empolvada en el techo.
Agradezco a las instituciones que han apoyado este proceso a Metrosalud por el apoyo y el tiempo, De los Andes Cooperativa, por los recursos aportados. Por último agradezco a todas las personas, compañeros que tuvieron que ver conmigo en este proyecto.

# Resumen

Las tasas estandarizadas se utilizan para comparar un evento en una población de estudio con una población estándar. Permiten controlar los efectos confusores de otras variables y dimensionar el evento de acuerdo a un mismo estándar. En este trabajo se evalúan varias metodologías para la estimación de las tasas estandarizadas a través de intervalos de confianza, algunas basadas en la distribución normal, Poisson y gamma. Con un estudio de simulación se evalúan los desempeños de los intervalos de confianza de las metodologías boostrap, momentos, gamma y bayesiana, en tres escenarios de tamaños de tasas diferentes, comparándolas a través de un índice que 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, de momentos y bootstrap.

**Palabras clave:** ***Tasas estandarizadas, epidemiología, confusión, intervalos de confianza, bioestadística***.

# Abstract

Standardized rates are used to compare an event in a study population with a standard population. They allow to control the confounding effects of other variables and to size the event according to the same standard. In this paper several methodologies are evaluated for the estimation, through confidence intervals, of the standardized rates. Some of them based on the normal distribution, Poisson and gamma. By using a simulation study, with three different values for the rates, the performance of the moments, gamma, bootstrap and Bayesian confidence intervals are evaluated. The performance of each interval is measure through a index that takes into account both the coverage level and the amplitude of the interval and it also permits to evaluate the quality of each of the considered intervals. The 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, confusion, confidence intervals, biostatistics***

# Contenido

**Agradecimientos** **v**

**Resumen** **vii**

**1. Introducción** **2**
1.1. Planteamiento del problema . . . . . 3
1.2. Objetivos . . . . . 3
1.2.1. General . . . . . 3
1.2.2. Específicos . . . . . 3
1.3. Metodología propuesta . . . . . 4

**2. Tasas estandarizadas y métodos para estimación por intervalos de confianza** **5**
2.1. Estado del arte . . . . . 5
2.2. Marco teórico . . . . . 8
2.2.1. Intervalos . . . . . 8
2.2.2. Tasas estandarizadas de mortalidad o de morbilidad . . . . . 9
2.2.3. Intervalos de confianza bootstrap para las tasas estandarizadas . . . . 16
2.2.4. Método de igualación de momentos (Moments matching method) . . 16
2.2.5. Métodos basados en la distribución gamma . . . . . 17
2.2.6. Método Bayesiano con distribuciones conjugadas Poisson y gamma . 18

**3. Índice propuesto para medir la bondad de los intervalos de confianza** **20**

**4. Simulación, Resultados y discusión** **23**
4.1. Estudio de simulación . . . . . 23
4.2. Resultados . . . . . 24
4.2.1. Análisis descriptivo . . . . . 24
4.2.2. Análisis inferencial . . . . . 27
4.2.3. Análisis convencional de intervalos de confianza . . . . . 34
4.3. Aplicación: Datos de cáncer de cérvix . . . . . 41
4.4. Discusión . . . . . 43

**5. Conclusiones, recomendaciones y direcciones futuras** **45**
5.1. Conclusiones . . . . . 45
5.2. Recomendaciones . . . . . 46
5.3. Direcciones futuras . . . . . 46

**A. Anexo: Programas en R** **49**

**B. Anexo: programas en SAS** **54**

# Lista de Tablas

**2-1**. Estructura de datos para el cálculo de las tasas estandarizadas directas . . . 10

**4-1**. Estructura de datos que se usó para las simulaciones de las tasas estandarizadas directas . . . . . 24
**4-2**. Valores descriptivos para simulación de tasas pequeñas . . . . . 26
**4-3**. Valores descriptivos para simulación de tasas medias . . . . . 26
**4-4**. Valores descriptivos para simulación de tasas grandes . . . . . 26
**4-5**. Resultados de los métodos considerados basados en el test de Kruskal-Wallis 27
**4-6**. Resultados comparaciones múltiples de los métodos considerados basados en el test de Mann-Withney . . . . . 28
**4-7**. Resultados comparaciones múltiples de cada método en las simulaciones basados en el test de Kruskal-Wallis . . . . . 34
**4-8**. Resultados comparaciones múltiples de las simulaciones por cada método basados en el test de Mann-Withney . . . . . 34
**4-9**. Promedio de Intervalos de confianza por método y por escenario . . . . . 35
**4-10**.Muertes por cáncer de cérvix y distribución de la población en Colombia en el año 2015 por departamento . . . . . 42
**4-11**.Tasas estandarizadas directas (DSR) de mortalidad por cáncer de mama por departamento e intervalos de confianza por método . . . . . 42
**4-12**.Índice por cada método y regiones . . . . . 42

# Lista de Figuras

**3-1**. Ilustración del cálculo de $w$ del índice . . . . . 21

**4-1**. Índice por método cuando se asumen tasas pequeñas basado en 1000 simulaciones . . . . . 25
**4-2**. Índice por método cuando se asumen tasas medias basado en 1000 simulaciones 25
**4-3**. Índice por método cuando se asumen tasas grandes basado en 1000 simulaciones 26
**4-4**. Box-plot de los índices por método basados en 1000 simulaciones con tasas pequeñas (simulación 1) . . . . . 29
**4-5**. Box-plot de los índices por método basados en 1000 simulaciones con tasas medianas (simulación 2) . . . . . 29
**4-6**. Box-plot de los índices por método basados en 1000 simulaciones con tasas grandes (simulación 1) . . . . . 30
**4-7**. Box-plot de los índices por simulación basados en 1000 simulaciones para el método bootstrap . . . . . 31
**4-8**. Box-plot de los índices por simulación basados en 1000 simulaciones para el método de los momentos . . . . . 31
**4-9**. Box-plot de los índices por simulación basados en 1000 simulaciones para el método gamma . . . . . 32
**4-10**.Box-plot de los índices por simulación basados en 1000 simulaciones para el método bayesiano . . . . . 32
**4-11**.Box-plot de los índices por simulación basados en 1000 simulaciones para el método bayesiano bootstrap . . . . . 33
**4-12**.Cobertura método bootstrap tasas pequeñas . . . . . 36
**4-13**.Cobertura método momentos tasas pequeñas . . . . . 36
**4-14**.Cobertura método gamma tasas pequeñas . . . . . 36
**4-15**.Cobertura método bayesiano tasas pequeñas . . . . . 37
**4-16**.Cobertura método bayesiano bootstrap tasas pequeñas . . . . . 37
**4-17**.Cobertura método bootstrap tasas medias . . . . . 37
**4-18**.Cobertura método de momentos tasas medias . . . . . 38
**4-19**.Cobertura método gamma tasas medias . . . . . 38
**4-20**.Cobertura método bayesiano tasas medias . . . . . 38

**4-21**.Cobertura método bayesiano bootstrap tasas medias . . . . . 39
**4-22**.Cobertura método bootstrap tasas grandes . . . . . 39
**4-23**.Cobertura método de momentos tasas grandes . . . . . 39
**4-24**.Cobertura método gamma tasas grandes . . . . . 40
**4-25**.Cobertura método bayesiano tasas grandes . . . . . 40
**4-26**.Cobertura método bayesiano bootstrap tasas grandes . . . . . 40

# Capítulo 1

# Introducción

Las tasas estandarizadas son utilizadas en epidemiología para comparar un evento 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, permitiendo dimensionar el evento de acuerdo a un mismo estándar o a una misma población, para identificar su magnitud o riesgo de acuerdo a los factores de cada grupo poblacional. Su estimación es de gran importancia y como método se considera el cálculo de los intervalos de confianza, una herramienta que aporta validez a la estimación y un acercamiento al verdadero valor del parámetro de interés.

Para la estimación de las tasas estandarizadas de mortalidad o morbilidad a través de intervalos de confianza, existen diferentes metodologías, basadas principalmente en la distribución Poisson [Dobson A, 1991], donde se parte desde un solo dato [Correa, 2007], por ejemplo, el número de muertes maternas en la ciudad de Medellín en el año 2015, el cual tiene una distribución Poisson ya que refleja la frecuencia con que se presenta un evento en un periodo determinado, [Woodward, 2005]. También se presentan muestras pequeñas o grandes, para las cuales existen otras 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 y generan intervalos de diferentes calidades, tanto a nivel de cobertura como en amplitud.

En este trabajo se comparan usando un estudio de simulación, algunas metodologías para la estimación de tasas estandarizadas directas por medio de intervalos de confianza, para lo cual se parte de algunas 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 T, 1996], metodología basada en los momentos [Dobson A, 1991], métodos basados en la distribución gamma [Fay M, 1997], y un método bayesiano con distri-

buciones conjugadas Poisson y gamma para las tasas aplicadas en epidemiología, [Ross, 2003]. Esta comparación se basará en un nuevo índice que se propone (el cual combina cobertura y amplitud, que constituye el aporte más original de este trabajo) que permite medir la calidad de los intervalos en términos de cobertura y amplitud, además de comparar la variabilidad de los IC. Este índice se basa en una propuesta hecha por [Correa J, 2003], para la diferencia de proporciones.

Este trabajo está organizado de la siguiente forma: en el capítulo 2 se exponen el estado del arte, las definiciones de tasas, tipos, formas y los métodos para su estimación; en el capítulo 3 se describe el índice propuesto, el capítulo 4 contiene el estudio de simulación, resultados, la aplicación con datos de cáncer de mama, comparando varias regiones de Colombia en el año 2015 y luego se desarrolla la discusión sobre los resultados hallados; por último en el capítulo 5 se presentan algunas conclusiones, recomendaciones y direcciones futuras.

## 1.1. Planteamiento del problema

La existencia de muchas metodologías para estimar intervalos de confianza para las tasas estandarizadas directas de mortalidad o morbilidad [Keung H, 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 J, 2003]. Por lo que se deben tener en cuenta las características de cada escenario y las limitaciones de cada método.
Por lo anterior surge la pregunta: ¿Cuáles podrían ser las mejores metodologías para estimar las tasas estandarizadas directas en epidemiología por medio de intervalos de confianza?.

## 1.2. Objetivos

### 1.2.1. General

Comparar vía simulación con la ayuda de un índice que compara tanto la cobertura como la amplitud, diferentes metodologías para la estimación de tasas estandarizadas directas por medio de intervalos de confianza

### 1.2.2. Específicos

1. Medir la calidad de intervalos de confianza para las tasas estandarizadas directas bajo simulación aplicando un índice propuesto, que tiene en cuenta la longitud del intervalo y el nivel de confianza real.

2. Caracterizar las diferentes metodologías para la estimación de intervalos de confianza para las tasas estandarizadas directas identificando los diferentes escenarios donde cada intervalo muestra un mejor desempeño.

3. Ilustrar con un ejemplo de datos reales la estimación de los intervalos de confianza con las diferentes metodologías que se comparan.

## 1.3. Metodología propuesta

Revisión bibliográfica del estado del arte de las metodologás que actualmente se plantean para la estimación de los intervalos de confianza para las tasas estandarizadas directas de mortalidad o morbilidad.

Para comparar las diferentes metodologías evaluadas para la estimación de las tasas estandarizadas directamente de un evento dado, se propone como metodología realizar simulaciones usando R para estimar los intervalos de confianza en tres escenarios representando tres tamaños de tasas diferentes, uno representando una tasa del orden de 1 sobre 100 mil la cual denominaremos tasas pequeñas, otro escenario del orden de 1 sobre 10 mil que denominaremos tasas medianas y por ultimo la del orden de 1 sobre mil que denominaremos tasas grandes, esta selección se realizó de manera arbitraria con el fin de representar tres tamaños de tasas diferentes, para cada una de las metodologías. además se comparan los diferentes intervalos generados en la simulación aplicando un índice propuesto que evidencia la calidad del intervalo teniendo en cuenta la amplitud y el nivel de cobertura, ponderando la amplitud y el nivel de cobertura en un índice que penaliza los intervalos de mayor amplitud y niveles de coberturas inferiores al nominal.

# Capítulo 2

# Tasas estandarizadas y métodos para estimación por intervalos de confianza

En este capítulo se exponen los métodos y avances teóricos que se han desarrollado al respecto de las tasas, tipos y las formas de obtención, se exponen los conceptos de intervalos y los métodos para la estimación de las tasas estandarizadas directas.

## 2.1. Estado del arte

[Dobson A, 1991] plantea que las tasas estandarizadas de mortalidad son una aplicación de las sumas de pesos ponderados de las tasas medias de eventos Poisson, si el número de eventos es grande, es adecuado utilizar una aproximación a la normal pero no lo es tanto para la tasa de eventos pequeños, se presenta un método para calcular intervalos de confianza para la suma ponderada de los parámetros de Poisson como funciones lineales de los límites de confianza para una única media Poisson; este método lo ilustra con datos de las tasas de infarto de miocardio obtenidos como parte del proyecto WHO MONICA en Augsburg, Alemania [Dobson A, 1991].

[Swift, 1995] en este artículo, presentó expresiones simples para los intervalos de confianza de las tasas estandarizadas indirectas y directas, basadas en el método aproximado bootstrap (ABC) de DiCiccio y Efron. Para las tasas indirectas, el método ABC lo comparó con los métodos exactos; para tasas estandarizadas directas no contó con un método exacto disponible, por lo tanto, comparó el ABC con otros procedimientos simples para intervalos de confianza, el método estándar normal, log-normal y un método reciente propuesto por [Dobson A, 1991]. Los estudios de simulación de las propiedades de cobertura de estos métodos para las tasas estandarizadas directas mostraron que el método ABC funciona mejor, con probabilidades de cola equilibradas cerca de sus valores esperados.

En [Fay, 1999], consideró los intervalos de confianza exactos para estimadores del riesgo relativo, calculados con tasas directamente estandarizadas, DSR, asumiendo un modelo multiplicador de Poisson denominados intervalos DSR exactos. Estos intervalos relativos son mas estrechos para el riesgo relativo que algunos intervalos exactos estándar, sin embargo son más difíciles de calcular. Introdujo una aproximación para estimar estos intervalos usando la distribución F inversa. Fay mostró que la aproximación es equivalente a ambos tipos de intervalos exactos cuando el estándar es proporcional a ambas poblaciones y es asintóticamente equivalente al método normal con una transformación logarítmica a medida que los recuentos y años-persona aumentan a la misma velocidad. Comparó su aproximación con el intervalo DSR exacto, con los resultados ajustando algunos casos con poblaciones no proporcionales, la aproximación fue un poco más conservadora que los intervalos DSR exactos, comparado con otras aproximaciones conocidas de la misma manera mostró que estos otros intervalos son a menudo más liberales que los intervalos DSR exactos.

[Ross, 2003] desarrolló una propuesta para la estimación de los intervalos de confianza para las tasas basados en la distribución Poisson, utilizando la distribución gamma como la apriori (no informativa) conjugada para la distribución de probabilidad Poisson, obteniendo una aposteriori gamma con parámetros $(x + 1, 1)$ para hallar el intervalo de predicción para la tasa bruta o global $\hat{R}$ así: $[a/A, b/A]$ donde $0 < a < b$ es aplicable para tasas globales y específicas, $a$ hace refencia a los eventos y $A$ a los años personas bajo estudio, no se tuvo en cuenta las tasas estandarizadas.

[Tiwari R, 2006] planteó que las tasas de cáncer ajustadas por edad se definen como el promedio ponderado de las tasas de cáncer específicas por edad, donde las ponderaciones son positivas, conocidas y normalizadas, de modo que su suma es 1.[Fay M, 1997] desarrolló un intervalo de confianza para un solo ajuste por edad para la tasa basada en la aproximación gamma. [Fay, 1999] usó las aproximaciones gamma para construir un intervalo basado en la distribución F para la relación de dos tasas ajustadas por edad. [Tiwari R, 2006] propone modificaciones de los intervalos gamma y F, lleva a cabo un estudio de simulación para mostrar que estos intervalos gamma y F modificados son más eficientes que los intervalos gamma y F, respectivamente, en el sentido de que los intervalos propuestos tienen probabilidades de cobertura empírica inferiores o iguales a sus contrapartes, y que también retienen el nivel nominal. También proporcionó los intervalos de confianza normal y beta para una tasa única ajustada por edad, pero se muestra que son ligeramente liberales. Finalmente, para comparar dos tasas correlacionadas ajustadas por edad, los intervalos de confianza para la diferencia y para la proporción de las dos tasas ajustadas por edad se derivan incorporando la correlación entre las dos tasas.

[Hon Keung T, 2008] compararon los desempeños de procedimientos de construcción de in-

tervalos de confianza comunes para la tasa de incidencia o mortalidad directa estandarizada. Utilizó el método basado en momentos coincidentes, el método basado en la distribución gamma y el método basado en la distribución beta. Investigó, a través de simulaciones Monte Carlo basadas en datos de un estudio epidemiológico de enfermedades cardiovasculares y escenarios generados aleatoriamente sus desempeños. Estos procedimientos de estimación de intervalos, así como algunos procedimientos mixtos desarrollados para el presente estudio, se evaluaron en términos de la probabilidad de cobertura y la duración esperada. De los cuales presentaron mejor desempeño un intervalo basado en el método de los momentos que denominaron M8 y los basados en la distribución gamma.

[Swift, 2010] realizó un estudio de simulación que proporciona un informe suplementario del rendimiento del intervalo de confianza para las tasas estandarizadas directas obtenidas por el método aproximado bootstrap (ABC), presentó una comparación gráfica de la probabilidad de cobertura y la razón de las probabilidades de no cobertura de derecha a izquierda, en función de la varianza en los pesos utilizados para cada punto de simulación, para el método ABC, así como tres de los procedimientos recomendados por [Hon Keung T, 2008], momentos M8 y Gamma G1 y G4 (denominados así por[Hon Keung T, 2008]). Mostró que los intervalos ABC son buenos competidores en comparación con los otros tres intervalos de confianza.

[Krishnamoorthy et al., 2016] consideran los procedimientos de estimación de intervalos para la relación de dos medias de Poisson, el promedio ponderado de las tasas de Poisson, utilizado para comparar las tasas de incidencia de una enfermedad en un grupo de tratamiento y grupo de control. Proponen un intervalo de confianza (IC) basado en el enfoque MLS (the modified large sample) y lo comparan con los IC de Cox y los IC de puntuación de Wilson. Con este estudio indican que los IC MLS son comparables con IC de Cox, y ofrece mejoras en algunos casos. También proporcionan IC MLS para una suma ponderada de los medias de Poisson y lo comparan con los IC aproximados disponibles. Con este trabajo [Krishnamoorthy et al., 2016] demuestran que el método MLS produce IC simples y satisfactorios para el problema relacionado con las tasas de Poisson.

Las tasas directamente estandarizadas son una herramienta integral para presentar las tasas de enfermedades que dependen en gran medida de la edad, como el cáncer. Así lo afirman [Fay and Kim, 2017]. Estadísticamente, estas tasas se modelan como una suma ponderada de variables aleatorias de Poisson, este es un problema estadístico difícil, porque hay $k$ variables de Poisson observadas y $k$ medias desconocidas. El intervalo de confianza gamma, se ha demostrado mediante simulaciones en trabajos previos,[Swift, 2010],[Keung H, 2008], que tiene una cobertura nominal mínima en todos los escenarios simulados, pero puede ser demasiado conservador. Las modificaciones anteriores a ese método tienen una cobertura promedio más cercana a la nominal, pero no alcanzan la cobertura nominal limitada en

todas las situaciones. Además, esas modificaciones no son intervalos centrales, y la tasa de error de cobertura superior puede ser sustancialmente más de la mitad del error nominal. Por lo anterior [Fay and Kim, 2017] aplicaron una modificación que denominaron mid-p al intervalo de confianza gamma. Los métodos típicos de mid-p renuncian a la cobertura garantizada para obtener una cobertura que a veces es más alta y a veces más baja que la tasa de cobertura nominal, según los valores de los parámetros. El intervalo gamma mid-p no tiene cobertura garantizada en todas las situaciones; sin embargo, en las situaciones (no excepcionales) en que el método gamma es demasiado conservador, el intervalo gamma mid-p a menudo tiene al menos una cobertura nominal. El intervalo gamma mid-p es especialmente apropiado cuando se quiere un intervalo central, ya que las simulaciones muestran que en muchas situaciones las tasas de error de cobertura superior e inferior son en promedio menores o iguales a la mitad de la tasa de error nominal.

Por último a manera de ejemplo de la aplicación de estas metodologías se presenta un caso de estudio donde se comparan tasas de un evento en poblaciones diferentes. Con el objetivo de determinar la incidencia mundial de cáncer de pene [Eduardo et al., 2017] realizó una revisión exhaustiva y un metanálisis de estudios de observación, con lo cual estimó la tasa mundial de cáncer de pene, se compararon la tasas estandarizadas por continentes y construyó una tasa global, la cual dio un resultado pequeño como evidencia de que el cáncer de pene es una enfermedad poco frecuente y no evidenció diferencias significativas con respecto a la distribución por continentes, utilizó la tasa estandarizada directamente estimadas por intervalos para comparar la incidencia y el riesgo en diferentes continentes a través de tasas en las regiones.

## 2.2. Marco teórico

### 2.2.1. Intervalos

Los intervalos estadísticos dependen de la aplicación, esta determina cual intervalo usar basados en la necesidad del problema, de acuerdo a esto existen diferentes tipos: intervalos de confianza, intervalos de tolerancia, de credibilidad y de predicción. Los más utilizados son los intervalos de confianza, estos toman diferentes muestras aleatorias de una población las cuales producen cada una un rango que es susceptible de contener el parámetro de interés, se extraen $n$ muestras de las cuales se define el porcentaje de cobertura del total de intervalos que contiene el parámetro, lo cual se conoce como el nivel de confianza. Los intervalos de credibilidad utilizan métodos bayesianos a través de la distribución conjugada posterior y se halla el rango de probabilidad donde se contiene el parámetro de interés, se obtiene utilizando los cuantiles de la distribución aposteriori. Los intervalos de predicción son intervalos que contienen el valor de la variable dependiente para una nueva observación dado los valores de una muestra de las variables independientes. Los intervalos de tolerancia

permiten estimar con un nivel de confianza en que rango se encuentra una proporción de la población [William Meeker, 2017].
Los intervalos de confianza algunas veces se estiman partiendo de la distribución de la variable en estudio. En muchos casos se aplican aproximaciones a la distribución normal basados en la teoría del teorema central del límite, pero existen intervalos propios de cada distribución que son de mejor calidad, expresada por el nivel de cobertura, amplitud y simetría, diferencias que son relativamente mínimas y que en la práctica son difíciles de implementar, por lo cual en la literatura y en las aplicaciones de software encontramos métodos que no son técnicamente aceptables, pero su uso masivo se debe a la facilidad en la implementación [William Meeker, 2017].

En el caso de las tasas estandarizadas directas de un evento dado, por ejemplo, mortalidad o morbilidad se evidencia en la literatura el uso de intervalos basados en la distribución normal, pero también existen diferentes propuestas partiendo de que estos eventos se distribuyen Poisson y de la relación existente de esta con las distribuciones gamma, chicuadrado, exponencial, beta, aplicaciones bootstrap, teoría de momentos y teorema del límite central. En el siguiente apartado se evalúan algunos métodos, que según trabajos previos como los de [Swift, 2010] y [Keung H, 2008] que han tenido mejores desempeños, para lo cual se pretende medir su desempeño con un índice propuesto, basado en una propuesta de [Correa J, 2003] para proporciones, que tiene en cuenta el nivel de cobertura y la amplitud del intervalo.

### 2.2.2. Tasas estandarizadas de mortalidad o de morbilidad

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 o la incidencia o prevalencia de una enfermedad, morbilidad. Estas se expresan como,

$$Tasa_x = \frac{E}{N} \times C \tag{2-1}$$

donde $x$ es el evento de interés, $E$ el número de eventos y $N$ la población en estudio, por una constante $C$ que representa un valor ponderado de la población que ajusta el resultado, 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. Por ejemplo, si en una población de 44 mil habitantes se presenta una defunción por muerte en incidentes de tránsito, se calcula la razón de 1/44,000 lo cual da un número poco comprensible 0.000025; entonces se multiplica por 100 mil y se interpreta como que en el periodo dado se presentaron 2.5 muertes por incidentes de tránsito por cada 100 mil habitantes en la población especificada. Estas tasas se clasifican de acuerdo a la forma de calcularlas en: tasas crudas, específicas por subcategorías, que comúnmente son edad o género, y las tasas estandarizadas que pueden ser directas o

indirectas [Schoenbach, 1999].

Las tasas crudas expresan la presencia del evento en una población sin tener en cuenta su composición o categoría. Puesto que los eventos difieren de un grupo poblacional a otro, por ejemplo, la mortalidad por una causa dada es diferente en cada grupo de edad o en cada género, su frecuencia cambia de un grupo a otro, por esto se estiman las tasas específicas por edad o género, las cuales permiten identificar la frecuencia de presencia del evento en el grupo poblacional específico. Partiendo del hecho de que algunas enfermedades son propias de algunos grupos de edad o género, por ejemplo son diferentes las causas de muerte entre niños, jóvenes y ancianos o entre hombres y mujeres, existen morbilidades propias de cada uno de estos grupos [Schoenbach, 1999].

Para obtener las tasas estandarizadas, se parte de las tasas específicas, calculando los pesos en cada grupo y comparándolos con una población estándar (una población mayor de referencia). Por ejemplo, la población global de un país para comparar el evento en sus subregiones o la población mundial para la cual la Organización Mundial de la Salud (OMS) ha propuesto estructuras de población global(SEGI) que se actualizan regularmente [Ahmad et al., 2001]. Esto permite evidenciar la presencia del evento en diferentes poblaciones y cuantificar su magnitud eliminando los efectos de factores extraños 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, 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]. Por ejemplo, para medir la tasa de mortalidad por cáncer de mama en las ciudades de Medellín y Cali se calculan las tasas específicas por grupos de edad en cada ciudad y esos pesos se le aplican a la población colombiana, lo que permite una mejor comparabilidad entre las dos ciudades. La estructura de los datos son como se ilustra la tabla 2.1,

**Tabla 2-1**: Estructura de datos para el cálculo de las tasas estandarizadas directas

| Grupos de edad | Eventos | Población estudio | Población estándar |
|---|---|---|---|
| $Grupo_1$ | $\lambda_1$ | $p_1$ | $ps_1$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $Grupo_k$ | $\lambda_k$ | $p_k$ | $ps_k$ |

La tasa se estima con la siguiente expresión

$$\text{Tasa estandarizada}_x = \sum_{i=1}^{k} w_i \lambda_i \tag{2-2}$$

Donde $w_i$ son los pesos de la población estándar del grupo de edad $i$, y $\lambda_i$ la tasa específica

en la población bajo estudio en un periodo de tiempo del grupo de edad $i$, $\lambda_i = \frac{x_i}{p_1}$ y $k$ representa el número de grupos edad en que se subdivide la población en estudio.

**Tasas de eventos directamente estandarizadas**

La estandarización directa [Woodward, 2005], se define como el número esperado de eventos en la población estándar, por la tasa de eventos en la edad específica en la población bajo estudio divido por el tamaño de la población estándar. Para evitar números pequeños 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 del evento por mil personas, denotada por dsr, es:

$$dsr = 1000 \sum_i \left(\frac{e_i}{p_i}\right) \frac{p_i^{(s)}}{p^{(s)}}, i = 1, ..., k \tag{2-3}$$

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:

$$se\,(dsr) = \frac{1000}{p^{(s)}} \left\{ \sqrt{\sum_i e_i \left(\frac{p_i^{(s)}}{p_i}\right)^2} \right\} \tag{2-4}$$

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

$$dsr \pm 1{,}96 se\,(dsr) \tag{2-5}$$

**Tasas de eventos indirectamente estandarizadas**

[Woodward, 2005] desarrolla el cálculo de la tasa indirectamente estandarizada(isr) así: suponiendo que $e = \sum_i e_i$ son el número total de eventos en la población estudio, $e_i^{(s)}$, es el número de eventos en el $i$-simo grupo de edad de la población estándar, $s^{(s)} = \sum e_i^{(s)}$ es el número total de eventos en la población estándar, $\rho_i^{(s)} = e_i^{(s)}/p_i^{(s)}$ es la tasa de eventos en el $i$-simo grupo de edad para la población estándar y $\rho^{(s)} = e^{(s)}/p^{(s)}$ es el total de la tasa de eventos en la población estándar, entonces, el número esperado de eventos en la población estudio es:

$$E = \sum \rho_i^{(s)} p_i \tag{2-6}$$

$$= \sum \left(\frac{e_i^{(s)}}{p_i^{(s)}}\right) p_i \tag{2-7}$$

por definición la tasa de eventos esperados(SER) es:

$$SER = \frac{\text{observaciones de número de eventos}}{\text{número de eventos esperados}} = \frac{e}{E} \tag{2-8}$$

asumiendo una distribución Poisson para el número de eventos observados $e$, el error estándar de la tasa de eventos esperados es:

$$se\,(SER) = \frac{\sqrt{e}}{E} \tag{2-9}$$

como con la tasa estandarizada directa, la tasa indirecta se expresa por mil personas, así:

$$isr = 1000 \times SER \times \rho^{(s)} \tag{2-10}$$

$$= 1000 \times SER \times \frac{e^{(S)}}{p^{(s)}} \tag{2-11}$$

con un error estándar dado por,

$$se\,(isr) = 1000\rho^{(s)} se\,(SER) \tag{2-12}$$

$$= 1000\rho^{(s)}\,(s)\,\frac{\sqrt{e}}{E} \tag{2-13}$$

$$= \frac{1000e^{(s)}\sqrt{e}}{p\,(s)\,E} \tag{2-14}$$

Por lo tanto para medir ya sea la incidencia o prevalencia de un evento (mortalidad o morbilidad) se estudia estimando las tasas estandarizadas directas basadas en las sumas ponderadas de las tasas específicas por edad o grupos de edad más comúnmente, para las cuales se considera que los eventos por cada grupo de edad son independientes entre los grupos y se distribuyen Poisson, [Woodward, 2005]; dado esto se consideran, para construir los intervalos de confianza, la suma ponderada de los parámetros Poisson, [Swift, 2010]. contrario al caso de la estanzarización indirecta en la cual se realizan las sumas de eventos por cada grupo de edad y se pondera por una constante que esta dada por la tasa de eventos de la población estándar.

Según lo definido por [Dobson A, 1991], igual que [Swift, 2010], plantea que las tasas estandarizadas directas 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 [Dobson A, 1991]. También [Woodward, 2005], considera que los eventos se distribuyen Poisson, estos coinciden con [Keung H, 2008], que realiza un estudio partiendo de este planteamiento y compara 20 métodos para estimar las tasas por medio de IC y con [Swift, 2010], donde en su estudio para comparar los métodos que mejor desempeño tuvieron del trabajo de [Keung H, 2008],

plantea, dados $X_i, i = 1, 2, 3, ..., k$, variables aleatorias que representan el número de eventos ocurridos en el grupo $i$,los $X_i ... X_k$ son independientes y se distribuyen Poisson con media $n_i\lambda_i$, donde $\lambda_i$ es la tasa de eventos y $n_i$ es el número de personas año por cada grupo $i$, si se define $\theta_i = n_i\lambda_i$ entonces $X_i \sim$Poisson $(\theta_i)$. Por lo tanto las tasas estandarizadas se definen así:

$$Y = \sum_{i=1}^{k} \frac{c_i}{n_i}\theta_i \tag{2-15}$$

$$= \sum_{i=1}^{k} w_i\theta_i \tag{2-16}$$

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, $i = 1, 2, 3, ..., k$. El estimador de máxima verosimilitud de $Y$ es:

$$\hat{Y} = \sum_{i=1}^{k} w_i\hat{\theta}_i \tag{2-17}$$

De donde se tiene que el valor esperado de $Y$,

$$E(Y) = \sum_{i=1}^{k} w_i\theta_i \tag{2-18}$$

la varianza de $Y$ es

$$Var(Y) = \sum_{i=1}^{k} w_i^2\theta_i \tag{2-19}$$

y el estimador de la varianza de $Y$, $Var(Y)$, es $\hat{V}$ dado por

$$\hat{V} = \sum_{i=1}^{k} w_i^2\hat{\theta}_i \tag{2-20}$$

Interesa estimar un IC para el parámetro $\theta$, 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 H, 2008]y [Swift, 2010] , los cuales se presentan a continuación y se adiciona un método bayesiano para estimar los intervalos de confianza para las tasas estandarizadas de morbi-mortalidad. Se pretende comparar estos métodos midiendo su desempeño aplicando un índice propuesto por nosotros que tiene en cuenta tanto la amplitud como el nivel de cobertura del intervalo en la evaluación de su desempeño.

Las tasas estandarizadas directas estimadas ajustadas por edad: $\hat{R} = \sum_{i=1}^{k} w_i x_i$ donde los $x_i$ son el número de eventos en $i = 1, 2, 3, ..., k$ (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. Dado $R = \sum_{i=1}^{k} w_i \theta_i$, la real tasa ajustada por edad poblacional, donde $\theta = P_i \lambda_i$, $\lambda_i$ es la verdadera pero desconocida tasa de eventos específica por edad.

Dado que $X_i \sim$ Poisson $(\theta = P_i y_i)$. La MLE para $\theta_i$ en $k$ clases se obtiene de la siguiente manera:

$$f(x_i) = \frac{\theta^{x_i} e^{-\theta_i}}{x_i!} \qquad , x_i = 0, 1, 2..., i = 1, ..., k \tag{2-21}$$

$$logf(x_i) = x_i log\theta_i - \theta_i + logx_i! \tag{2-22}$$

$$\frac{\partial logf(x_i)}{\partial \theta_i} = \frac{x_i}{\theta_i} - 1 \tag{2-23}$$

$$\Rightarrow \frac{\partial logf}{\partial \theta_i} = 0 \tag{2-24}$$

Implica que $\dfrac{x_i}{\theta_i - 1} = 0 \Rightarrow x_i = \theta_i$ De esto se concluye que el MLE de $\theta_i$ es $x_i$, es decir $\hat{\theta}_i = x_i$, entonces el MLE para $\hat{R}$ esta dada por:

$$\hat{R} = \sum_{i=1}^{k} w_i \hat{\theta}_i \tag{2-25}$$

$$= \sum_{i=1}^{k} w_i \hat{x}_i \tag{2-26}$$

Ya que $\hat{R} \sim$ Poisson compuesta al ser una combinación lineal de v.a. Poisson independientes con media y varianza: $\hat{R} \sim$Poisson$\left(\theta_1 + \theta_2 + ...\theta_k = \sum_{i=1}^{k} \theta_i\right)$

$$E\left(\hat{R}\right) = E\left(\sum_{i=1}^{k} w_i X_i\right) \tag{2-27}$$

$$= \sum_{i=1}^{k} E(w_i X_i) \tag{2-28}$$

$$= \sum_{i=1}^{k} w_i E(X_i) \tag{2-29}$$

$$= \sum_{i=1}^{k} w_i \theta_i \tag{2-30}$$

Ya que los $X_i \sim \text{Poisson}(\theta)$ entonces $E\left(X_i\right) = \theta_i$

$$V\left(\hat{R}\right) = V\sum_{i=1}^{k} w_i X_i \tag{2-31}$$

$$= \sum_{i=1}^{k} V\left(w_i X_i\right) \tag{2-32}$$

ya que los $X_i$s son v.a. independientes, entonces

$$V\left(\hat{R}\right) = \sum_{i=1}^{k} w_i^2 V\left(X_i\right) \tag{2-33}$$

como

$$[V\left(X_i\right) = \theta_i \tag{2-34}$$

entonces

$$V\left(\hat{R}\right) = \sum_{i=1}^{k} w_i^2 \theta_i \tag{2-35}$$

Un estimador de $E\left(\hat{R}\right)$ es,

$$\widehat{E\left(\hat{R}\right)} = \sum_{i=1}^{k} w_i \hat{\theta}_i \tag{2-36}$$

$$= \sum_{i=1}^{k} w_i x_i \tag{2-37}$$

y un estimador de $V\left(\hat{R}\right)$ es,

$$\widehat{V\left(\hat{R}\right)} = \sum_{i=1}^{k} w_i^2 \hat{\theta}_i \tag{2-38}$$

$$= \sum_{i=1}^{k} w_i^2 x_i \tag{2-39}$$

### 2.2.3. Intervalos de confianza bootstrap para las tasas estandarizadas

Estos IC se obtienen usando la metodología propuesta por [Diciccio, 1991], [Efron, 1993], [Diciccio T, 1995] y [DiCiccio T, 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 $(1-\alpha)\,100\,\%$ se estiman los intervalos de confianza para $\theta$, así:

$$L = \hat{\theta} + \frac{z_0 - z_{1-\frac{\alpha}{2}}}{\left(1 - a\left[z_0 - z_{1-\frac{\alpha}{2}}\right]\right)^2}\hat{\sigma} \tag{2-40}$$

y

$$U = \hat{\theta} + \frac{z_0 - z_{1-\frac{\alpha}{2}}}{\left(1 - a\left[z_0 + z_{1-\frac{\alpha}{2}}\right]\right)^2}\hat{\sigma} \tag{2-41}$$

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.

### 2.2.4. Método de igualación de momentos (Moments matching method)

Este método fue propuesto por [Dobson A, 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 $\theta$ son: [Keung H, 2008]

$$L = \hat{\theta} + \sqrt{\frac{\hat{V}}{X}}\,(X_L - X) \tag{2-42}$$

$$U = \hat{\theta} + \sqrt{\frac{\hat{V}}{X}}\,(X_U - X) \tag{2-43}$$

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 de mejores desempeños son: $(M1)$ intervalo de confianza exacto, [Johnson N, 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:

$$X_L = \frac{1}{2}\chi^2_{\alpha/2}\left(2X\right) \tag{2-44}$$

$$X_U = \frac{1}{2}\chi^2_{1-\alpha/2}\left(2\left(X+1\right)\right) \tag{2-45}$$

donde $\chi^2_q\left(d\right)$ es el cuantil de la distribución chi-cuadrado con $d$ grados de libertad.
Y el (M8) aproximación del intervalo de confianza del mid-p [Dobson A, 1991].

$$X_L = \left(X+\frac{1}{2}\right)\left[1-\frac{1}{9\left(X+\frac{1}{2}\right)}-\frac{Z_{1-\alpha/2}}{\sqrt[3]{X+\frac{1}{2}}}\right]^{-3} \tag{2-46}$$

$$X_U = \left(X+\frac{1}{2}\right)\left[1-\frac{1}{9\left(X+\frac{1}{2}\right)}+\frac{Z_{1-\alpha/2}}{\sqrt[3]{X+\frac{1}{2}}}\right]^{-3} \tag{2-47}$$

### 2.2.5. Métodos basados en la distribución gamma

[Fay M, 1997] propuso 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 H, 2008], [Swift, 2010] simulan como método $(G1)$, el intervalo de confianza con extremos $X_L, X_U$ respectivamente del $(1-\alpha)\,100\,\%$, así:

$$L = \frac{\hat{V}}{2\hat{\theta}}\left(\chi^2\right)^{-1}_{\left(\frac{2\hat{\theta}^2}{\hat{V}}\right)}\left(\alpha/2\right) \tag{2-48}$$

$$U = \frac{\hat{V}+W^{*2}}{2\left(\hat{\theta}+W^*\right)}\left(\chi^2\right)^{-1}_{\left(\frac{2\left(\hat{\theta}+W^*\right)^2}{\hat{V}+W^{*2}}\right)}\left(1-\frac{\alpha}{2}\right) \tag{2-49}$$

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 M, 1997], selecciona $W^* = max\left(w_1, w_2, w_3 ... w_k\right)$ y mostró por simulación que estos intervalos son conservadores.

[Tiwari R, 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:

$$U^* = \frac{\hat{V}^*}{2\hat{\theta}^*}\left(\chi^2\right)^{-1}_{\left(\frac{2\hat{\theta}^{*2}}{\hat{V}^*}\right)}\left(1-\frac{\alpha}{2}\right) \tag{2-50}$$

Con $\hat{\theta}^* = \hat{\theta} + \frac{1}{k}\sum_{i=1}^{k} W_i$ 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 R, 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 R, 2006] en base a la idea de corrección de continuidad.

### 2.2.6. Método Bayesiano con distribuciones conjugadas Poisson y gamma

En el siguiente método para la estimación de la tasa bruta o global propuesto por [Ross, 2003], utiliza la distribución de probabilidad Poisson,

$$f_{X|\lambda}(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad , x = 0, 1, 2, .... \tag{2-51}$$

Donde $x$ es la medición del número de eventos y $\lambda$ es el parámetro de la distribución. La distribución gamma es la apriori (no informativa) conjugada para la distribución de probabilidad Poisson. La distribución gamma tiene la siguiente forma:

$$f_\lambda(x) = \left\{ \frac{\lambda^{\alpha-1}e^{-\lambda/\beta}}{\beta^\alpha}\Gamma(\alpha) \quad x > 0 \right. \tag{2-52}$$

Donde $\Gamma(\alpha)$ es la función gamma, para el cual $\alpha$ es un entero positivo y se puede expresar $\Gamma(\alpha) = (\alpha-1)!$, una forma de la distribución gamma puede ser definida con $\alpha = 1$ y $\beta = K_\beta$, $K_\beta$ algún valor muy grande. Con estos parámetros para la distribución apriori de las tasas corresponde a:

$$f_\lambda(x) = \left\{ \frac{e^{-\lambda/k_\beta}}{k_\beta} \right. \tag{2-53}$$

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

$$f_{\lambda|x}(\lambda|x) = \left\{ \frac{\lambda^x e^{-\lambda}}{x!} \quad \lambda \in [0, \infty) \right. \tag{2-54}$$

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

$$Pr\left\{\lambda \in [a,b]\,|X\right\} = 1/x! \int_a^b \lambda^x e^{-\lambda} d\lambda \tag{2-55}$$

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

$$\hat{R} = [a/N, b/N] \tag{2-56}$$

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,

$$\hat{R} = \left[\sum_{i=1}^{k} (a * w_i)\right] \tag{2-57}$$

y superior

$$\hat{R} = \left[\sum_{i=1}^{k} (b * w_i)\right] \tag{2-58}$$

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

Como otra forma para ponderar las sumas de los extremos para los intervalos de confianza bayesianos basados en lo propuesto por [Ross, 2003] para

$$\hat{R} = \left(\sum_{i=1}^{k} a * w_i, \sum_{i=1}^{k} b * w_i\right) \tag{2-59}$$

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 $\hat{R}$.

# Capítulo 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 J, 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, $P\left(L_{inf} \leq \pi_1 - \pi_2 \leq L_{sup}\right)$. Estos dos criterios no los podemos 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 J, 2003], propone 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 J, 2003] para la diferencia de las proporciones está definido de la siguiente forma

$$I = \frac{2 - LPI}{2} \times \frac{NR}{NN} \tag{3-1}$$

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 TLC, la distribución asintótica de $\bar{\theta}$ es $N(0,1)$ para $\bar{\theta}, (\theta_i, \theta_s)$, y se calcula $1-W$ donde $W$ esta 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,

$$W = \left|F\left(\bar{\theta}_s\right) - F\left(\bar{\theta}_i\right)\right| \tag{3-2}$$

Note que,

$$\begin{array}{c} (\theta_i^1, \theta_s^1) \\ (\theta_i^2, \theta_s^2) \\ \vdots \\ (\theta_i^{1000}, \theta_s^{1000}) \\ \hline \left(\bar{\theta}_i, \bar{\theta}_s\right) \end{array}$$

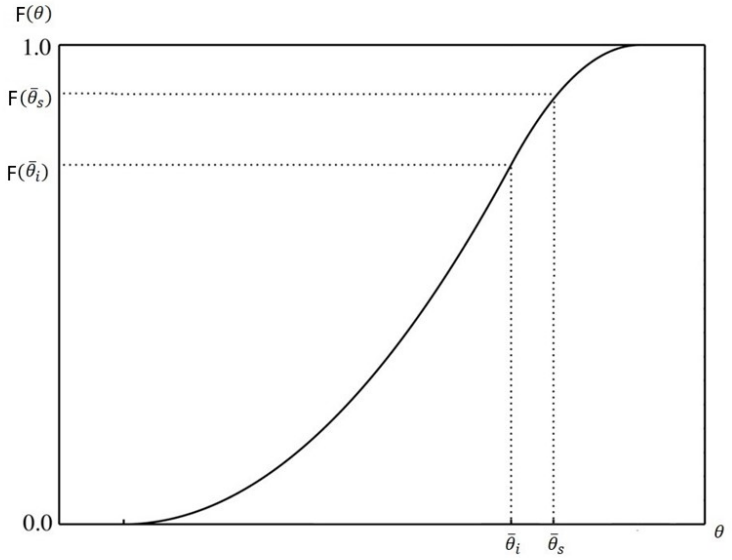


**Figura 3-1**: Ilustración del cálculo de $w$ del índice

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

$$F\left(\bar{\theta}\right) = \int_{-\infty}^{\theta} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz \tag{3-3}$$

Como en el índice de [Correa J, 2003] la segunda componente $\frac{NR}{NN}$ para nuestra propuesta solo 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 esta tesis de la siguiente forma,

$$I = (1 - W) \times \ NR \tag{3-4}$$

Obteniéndose el índice para el cual, 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 J, 2003], tanto en precisión como en probabilidad de cobertura.

# Capítulo 4

# Simulación, Resultados y discusión

En este capítulo se presenta el estudio de simulación, los resultados obtenidos y una aplicación con datos reales de cáncer de mama en varias regiones de Colombia del año 2015.

## 4.1. 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 $\lambda$, que representa la tasa promedio de un evento dado, se generan con una distribución Poisson $\lambda$, eventos al azar para cada simulación, con estas se estiman los intervalos de confianza por cada método, para tres escenarios de simulaciones con tamaños de tasas de eventos diferentes, se simulan una tasa de eventos en una proporción de 1 sobre un denominador de 100 mil, otra en una proporción de 1 sobre 10 mil y por ultimo una tasa de 1 sobre mil en la población bajo estudio, que representan tasas pequeñas en el caso de eventos raros, otro caso para tasas medias y un tercer caso con tasas grandes haciendo referencia a eventos más comunes, esta selección se hizo de forma arbitraria para evaluar diferentes tamaños de tasas.

Como se presenta en la tabla 4-1 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 quiquenios 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

**Tabla 4-1**: Estructura de datos que se usó para las simulaciones de las tasas estandarizadas directas

| Grupos de edad | media de eventos | población estudio | población estándar |
|---|---|---|---|
| $edad_{25-29}$ | $\lambda_1$ | $p_1$ | $ps_1$ |
| $edad_{30-34}$ | $\lambda_2$ | $p_2$ | $ps_2$ |
| $edad_{35-39}$ | $\lambda_3$ | $p_3$ | $ps_3$ |
| $edad_{40-44}$ | $\lambda_4$ | $p_4$ | $ps_4$ |
| $edad_{45-49}$ | $\lambda_5$ | $p_5$ | $ps_5$ |
| $edad_{50-54}$ | $\lambda_6$ | $p_6$ | $ps_6$ |
| $edad_{55-59}$ | $\lambda_7$ | $p_7$ | $ps_7$ |
| $edad_{60-64}$ | $\lambda_8$ | $p_8$ | $ps_8$ |

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 apriori 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, 2017a]

## 4.2. Resultados

### 4.2.1. Análisis descriptivo

En las figuras 4-1, 4-2 y 4-3 se observa el comportamiento del índice por cada uno de los métodos para los escenarios de tasas pequeñas representando un evento de poca frecuencia, tasas medias y tasas grandes para eventos con una mayor frecuencia. Como se había citado antes para ilustrar una enfermedad escasa o una más frecuente. 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 eurístico, a medida que aumenta la tasa este tiende a generar un mayor índice y

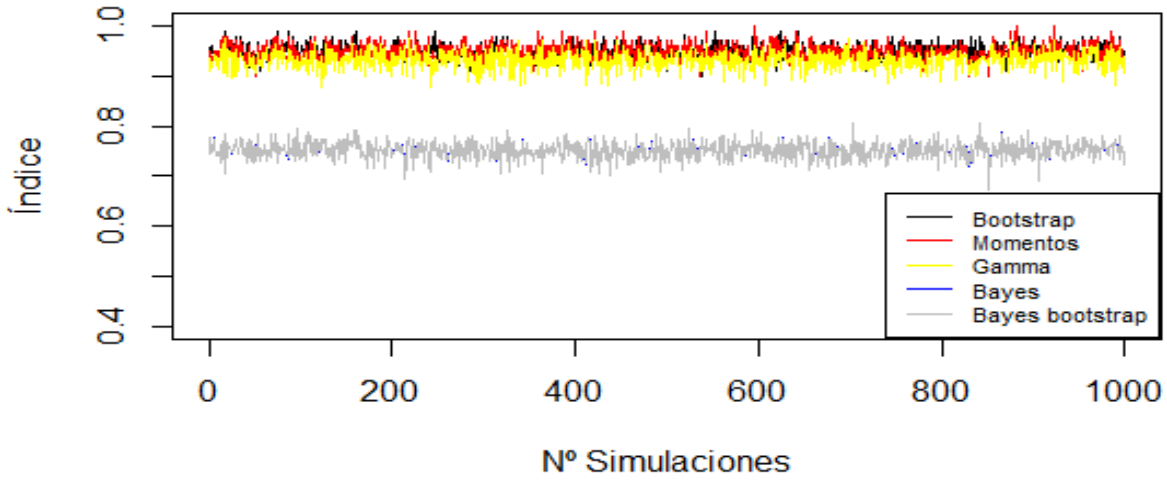


**Figura 4-1**: Índice por método cuando se asumen tasas pequeñas basado en 1000 simulaciones

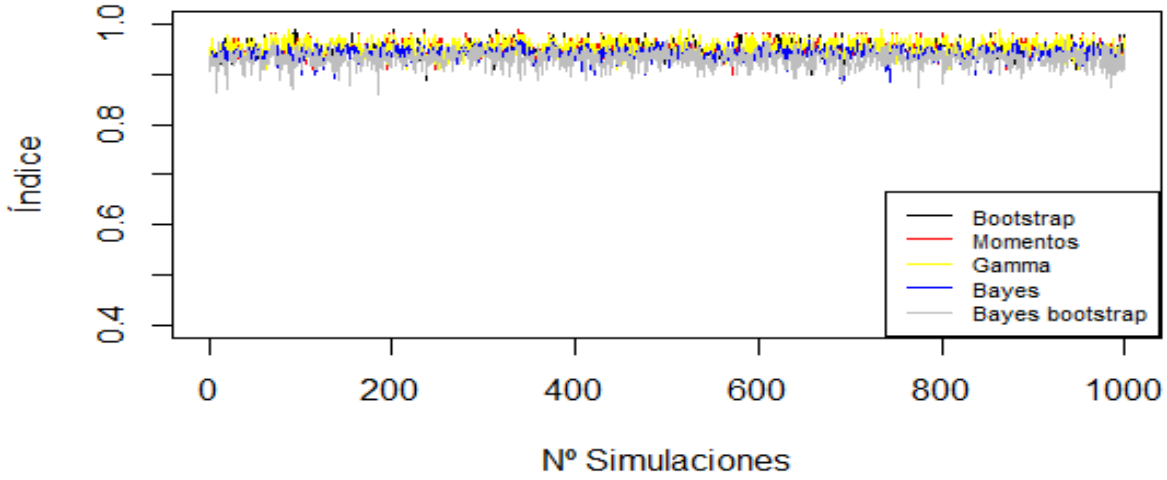


**Figura 4-2**: Índice por método cuando se asumen tasas medias basado en 1000 simulaciones

el bayesiano bootstrap igual que en el caso anterior genera valores pequeños del índice para tasas de eventos escasos, presentan un mayor índice en las tasas medias y en el caso de las tasas grandes vuelve a perder cobertura, lo que genera un índice de menor valor en este escenario.

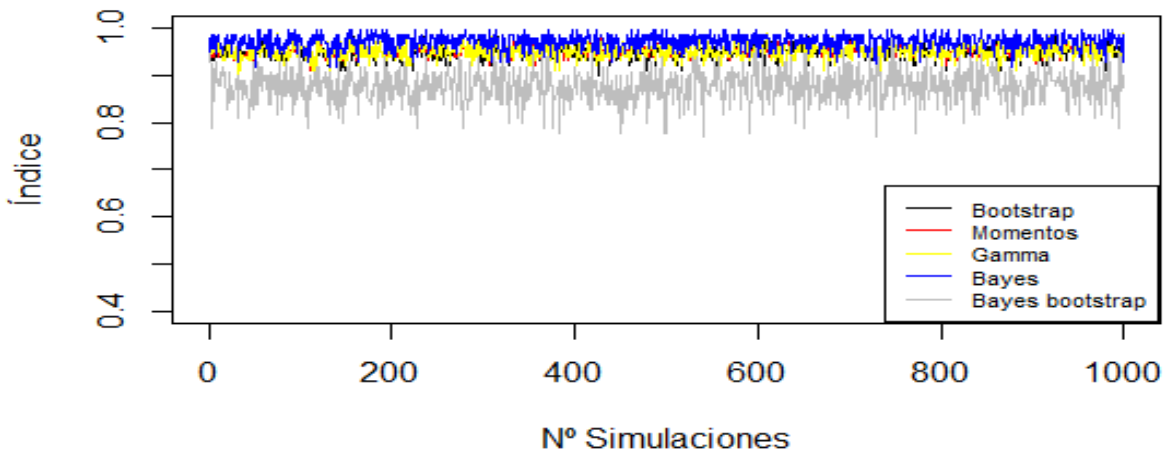


**Figura 4-3**: Índice por método cuando se asumen tasas grandes basado en 1000 simulaciones

**Tabla 4-2**: Valores descriptivos para simulación de tasas pequeñas

| Método | **N** | **Media** | **Mediana** | **Desviación** | **Mínimo** | **Máximo** |
|---|---|---|---|---|---|---|
| Bootstrap | 1000 | 0.949 | 0.949 | 0.016 | 0.899 | 0.990 |
| Gamma | 1000 | 0.931 | 0.933 | 0.018 | 0.876 | 0.976 |
| Momentos | 1000 | 0.947 | 0.950 | 0.018 | 0.880 | 1.000 |
| Bayes | 1000 | 0.752 | 0.752 | 0.017 | 0.675 | 0.806 |
| Bayes bootstrap | 1000 | 0.751 | 0.752 | 0.017 | 0.674 | 0.806 |

**Tabla 4-3**: Valores descriptivos para simulación de tasas medias

| Método | **N** | **Media** | **Mediana** | **Desviación** | **Mínimo** | **Máximo** |
|---|---|---|---|---|---|---|
| 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 |

**Tabla 4-4**: Valores descriptivos para simulación de tasas grandes

| Método | **N** | **Media** | **Mediana** | **Desviación** | **Mínimo** | **Máximo** |
|---|---|---|---|---|---|---|
| Bootstrap | 1000 | 0.947 | 0.950 | 0.015 | 0.900 | 0.990 |
| Gamma | 1000 | 0.950 | 0.950 | 0.015 | 0.900 | 0.990 |
| Momentos | 1000 | 0.949 | 0.950 | 0.015 | 0.900 | 0.990 |
| Bayes | 1000 | 0.971 | 0.976 | 0.016 | 0.907 | 0.997 |
| Bayes bootstrap | 1000 | 0.876 | 0.878 | 0.034 | 0.768 | 0.977 |

Las tres simulaciones tienen un comportamiento similar, homogéneo, según las tablas 4-2,4-3 y 4-4. Basados en la media y mediana se evidencia el comportamiento de los métodos en cada uno de los escenarios simulados, se observa un cambio importante en los métodos bayesianos cuando se incrementa el número de eventos con respecto a la población, los métodos bootstrap, momentos y gamma tienen un comportamiento muy consistente en los escenarios simulados. Se resalta el método gamma en los tres escenarios, para este representa el intervalo de confianza con mayor valor en promedio del índice, lo que indica un mejor desempeño.

### 4.2.2. Análisis inferencial

Partiendo de un análisis inferencial para establecer si los índices de cada método de estimación 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 cada uno de los escenarios simulados, correspondientes a una tasa de eventos pequeñas, medias y grandes, de acuerdo a los box-plots asociados a los métodos, figuras 4-4, 4-5 y 4-6, 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 4-5**: Resultados de los métodos considerados basados en el test de Kruskal-Wallis

| **Test de Kruskal-Wallis** | | |
|---|---|---|
| **Método** | **Chi-cuadrado*** | **Valor p** |
| **Simulación 1** | 3795,2181 | < ,0001 |
| **Simulación 2** | 1003,1595 | < ,0001 |
| **Simulación 3** | 47,6696 | < ,0001 |
| *grados de libertad = 4 | | |

En la tabla 4.5 se presentan los resultados del test de Kruskal-Wallis como prueba no paramétrica para corroborar las diferencias entre los métodos en cada una de las simulaciones, se infiere con valores p menores a 0.0001, que existen diferencias significativas entre todos los métodos para estimar por intervalos de confianza las tasas estandarizadas directas, en los tres escenarios evaluados, tasa de eventos baja, mediana y una tasa que representa un evento común o equivalente a tasas mayores. lo que permite inferir que los métodos para estimar por intervalos de confianza las tasas estandarizadas directas evaluados tiene diferencias significativas.

También se resaltan las diferencias observadas en las figuras 4-4 a la 4-6, entre los métodos bayesianos y frecuentistas, y alguna similaridad entre sí, o sea los métodos bootstrap, mo-

**Tabla 4-6**: Resultados comparaciones múltiples de los métodos considerados basados en el test de Mann-Withney

| | **Simulación 1** | | **Simulación 2** | | **Simulación 3** | |
|---|---|---|---|---|---|---|
| **Método** | **Chi-cuadrado** | **Valor p** | **Chi-cuadrado** | **Valor p** | **Chi-cuadrado** | **Valor p** |
| **boostrap-momentos** | 3.5573 | 0.0593 | 77.5854 | < ,0001 | 47.6696 | < ,0001 |
| **bootstrap-gamma** | 473.8538 | < ,0001 | 19.1394 | < ,0001 | 58.0125 | < ,0001 |
| **bootstrap-bayesiano** | 1499.2764 | < ,0001 | 152.9947 | < ,0001 | 7.106.866 | < ,0001 |
| **gamma-momentos** | 345.8728 | < ,0001 | 31.7622 | < ,0001 | 0.4285 | 0.5127 |
| **gamma-bayesiano** | 1499.2537 | < ,0001 | 277.4753 | < ,0001 | 624.1013 | < ,0001 |
| **momentos-bayesiano** | 1501.4823 | < ,0001 | 232.5141 | < ,0001 | 625.6995 | < ,0001 |
| **bayesiano b-momentos** | 1501.4823 | < ,0001 | 561.5546 | < ,0001 | 1420.5647 | < ,0001 |
| **bayesiano b-bootstrap** | 1499.2763 | < ,0001 | 462.4739 | < ,0001 | 1398.1150 | < ,0001 |
| **bayesiano b-gamma** | 1499.2537 | < ,0001 | 610.3963 | < ,0001 | 1423.7848 | < ,0001 |
| **bayesiano b-bayesiano** | 0.1167 | 0.7326 | 126.9932 | < ,0001 | 1471.8808 | < ,0001 |
| ***bayesiano b: bayesiano bootstrap, ** grados de libertad = 1** | | | | | | |

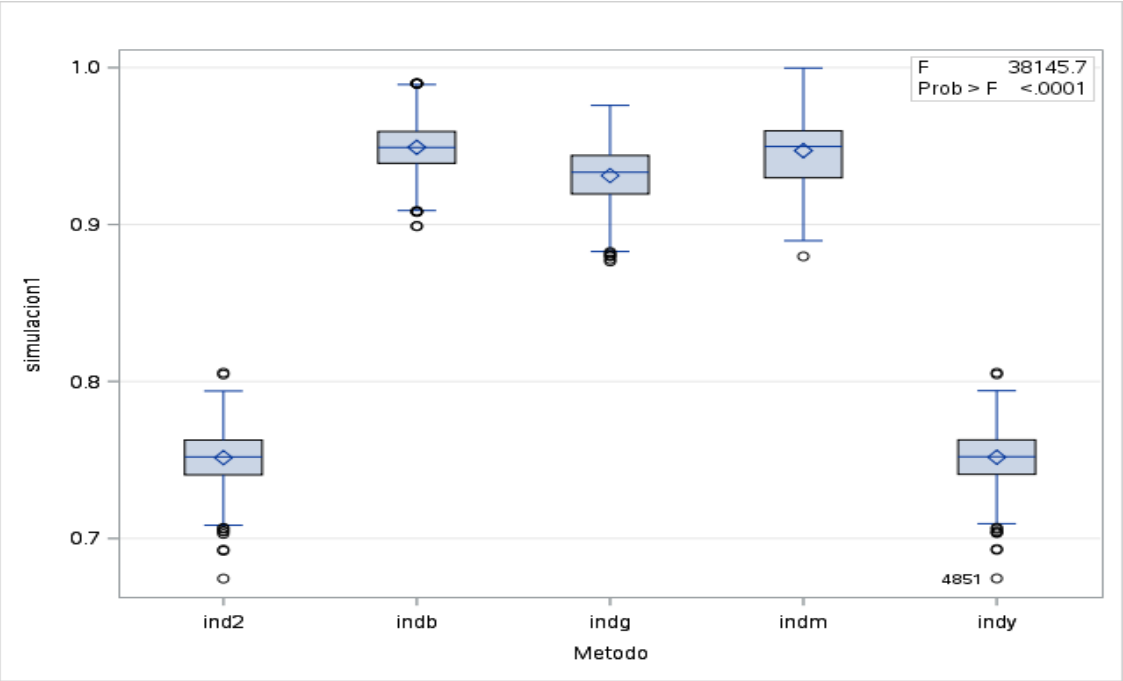


**Figura 4-4**: Box-plot de los índices por método basados en 1000 simulaciones con tasas pequeñas (simulación 1)

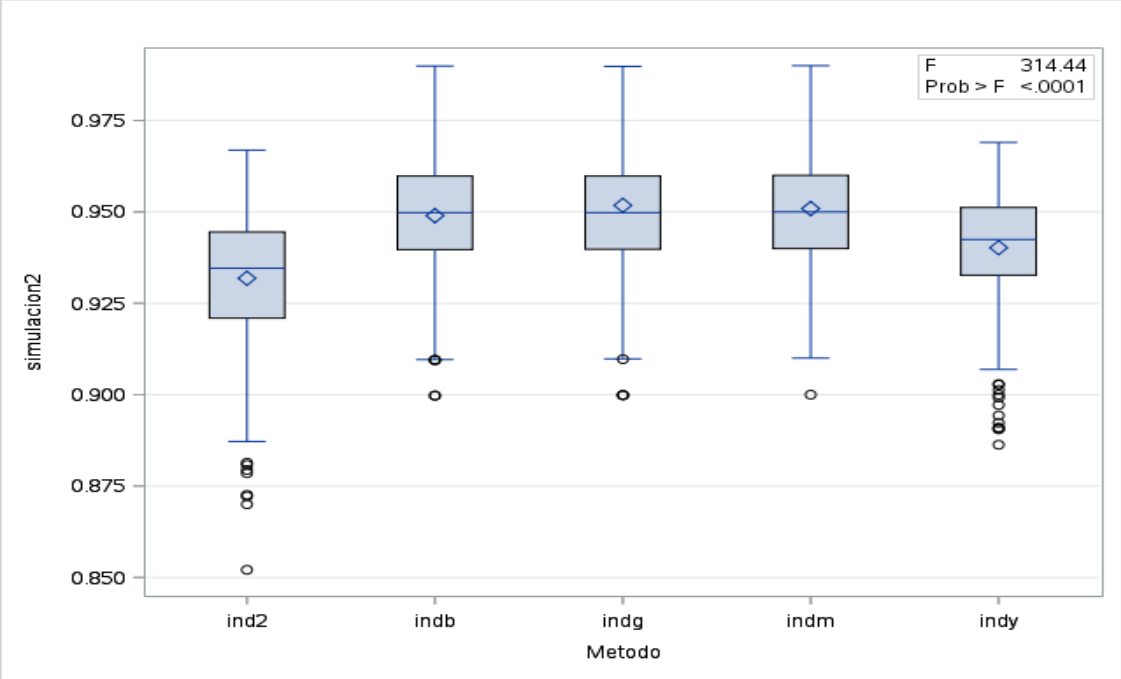


**Figura 4-5**: Box-plot de los índices por método basados en 1000 simulaciones con tasas medianas (simulación 2)

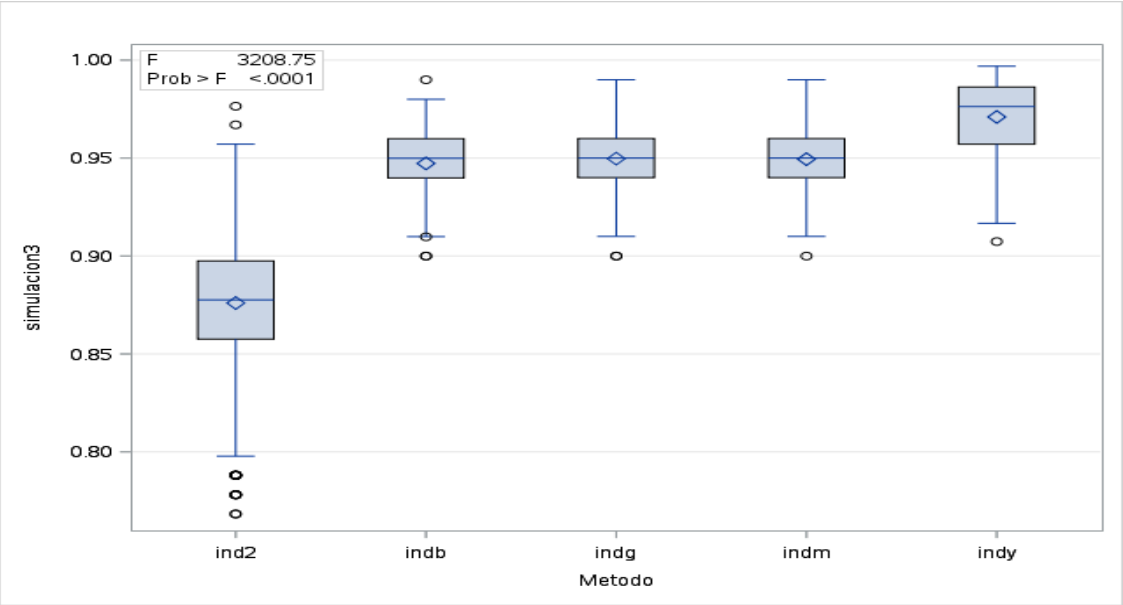


**Figura 4-6**: Box-plot de los índices por método basados en 1000 simulaciones con tasas grandes (simulación 1)

mentos y gamma tienen pocas diferencias en los escenarios simulados. Los bayesianos entre sí conservan un mismo patrón, solo difieren en el escenario de tasas de eventos grandes. Para aclarar si estas diferencias son significativas o no, se realiza un análisis de comparaciones múltiples de los métodos considerados basados en el test de Mann-Withney que permite identificar si existen diferencias entre sí en cada uno de los métodos.

Se realiza un análisis de comparaciones múltiples de los métodos considerados basados en el test de main-witheney para encontrar si existen diferencias entre los métodos uno a uno, los resultados del Test, tabla 4-6, para el caso de la simulación 1 correspondiente a la tasa de eventos pequeñas y para los métodos boostrap con momentos y bayesiano con bayesiano boostrap, con valores P: de 0.0593 y 0.7326 respectivamente, no se evidencia diferencias significativas de estos métodos entre sí. Para el caso del escenario de tasas de eventos medianos se reporta evidencia de que existen diferencias entre cada uno de los métodos entre sí valores $p < 0{,}0001$. Por último para las tasas de eventos comunes (simulación 3), se evidencian diferencias significativas entre cada uno de los métodos, valores $p < 0{,}0001$, excepto entre el método gamma y momentos, valor p: 0.5127 lo cual permite inferir que no existe diferencias entre los métodos.

Adicionalmente, se evaluó si existen diferencias en cada método en los tres escenarios, tasas pequeñas, medianas y grandes, para lo cual se observa en las figuras 4-7 a la 4-11 en los método bootstrap y de los momentos que no existen diferencias en el desempeño en los tres escenarios evaluados, para el método gamma se visualiza un cambio de la simulación 1 con

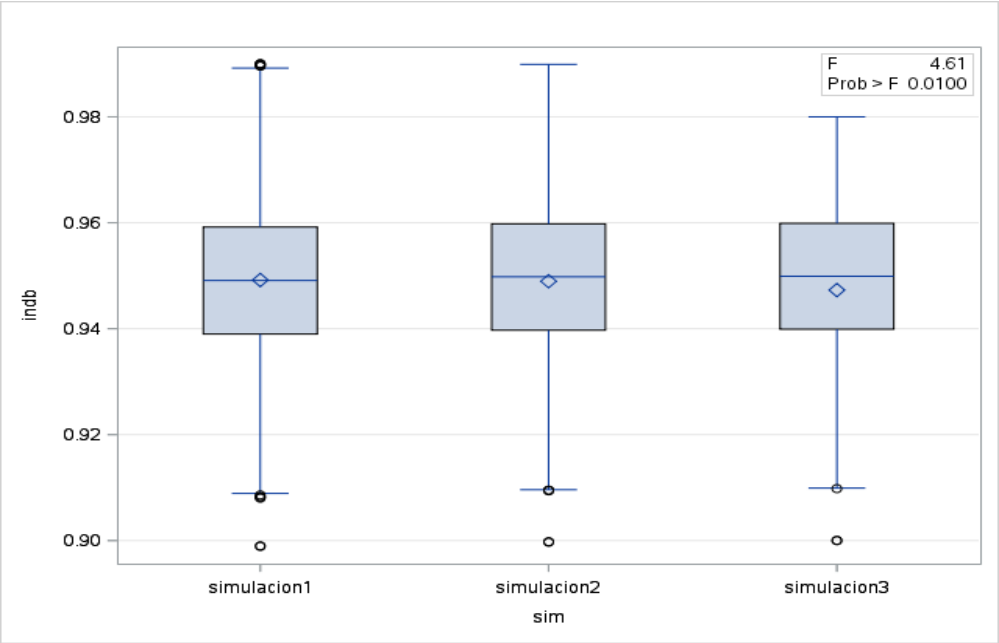


**Figura 4-7**: Box-plot de los índices por simulación basados en 1000 simulaciones para el método bootstrap

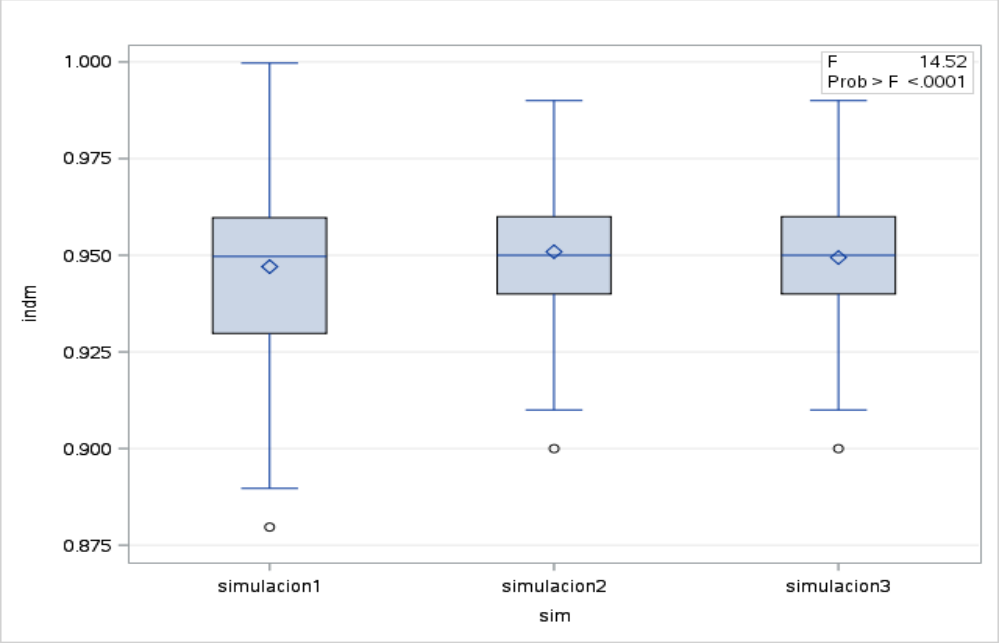


**Figura 4-8**: Box-plot de los índices por simulación basados en 1000 simulaciones para el método de los momentos

respecto a las otras 2, en caso de los métodos bayesianos si se evidencia cambios importantes entres las simulaciones evaluadas.

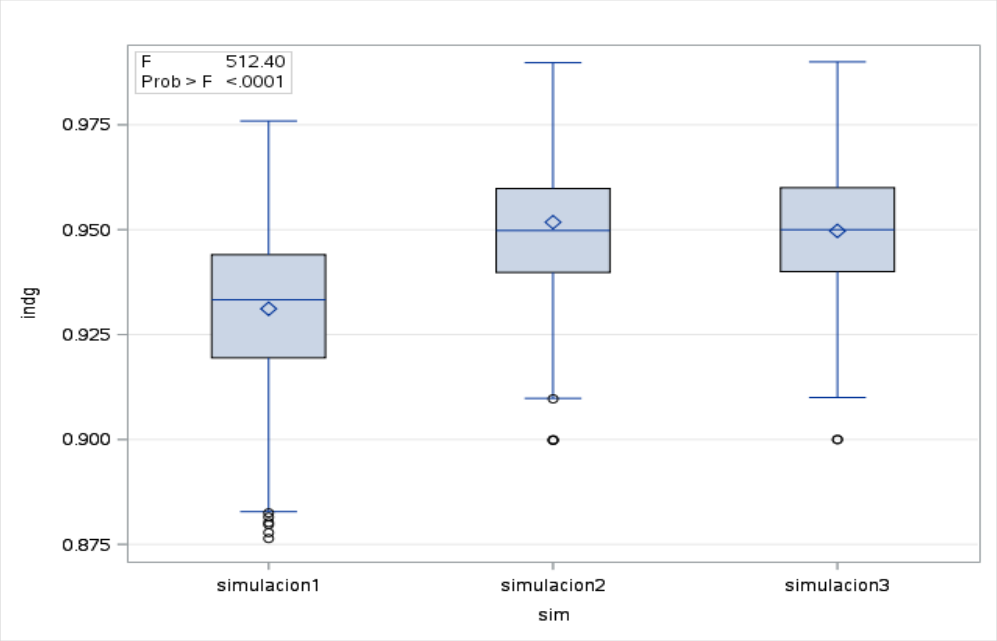


**Figura 4-9**: Box-plot de los índices por simulación basados en 1000 simulaciones para el método gamma

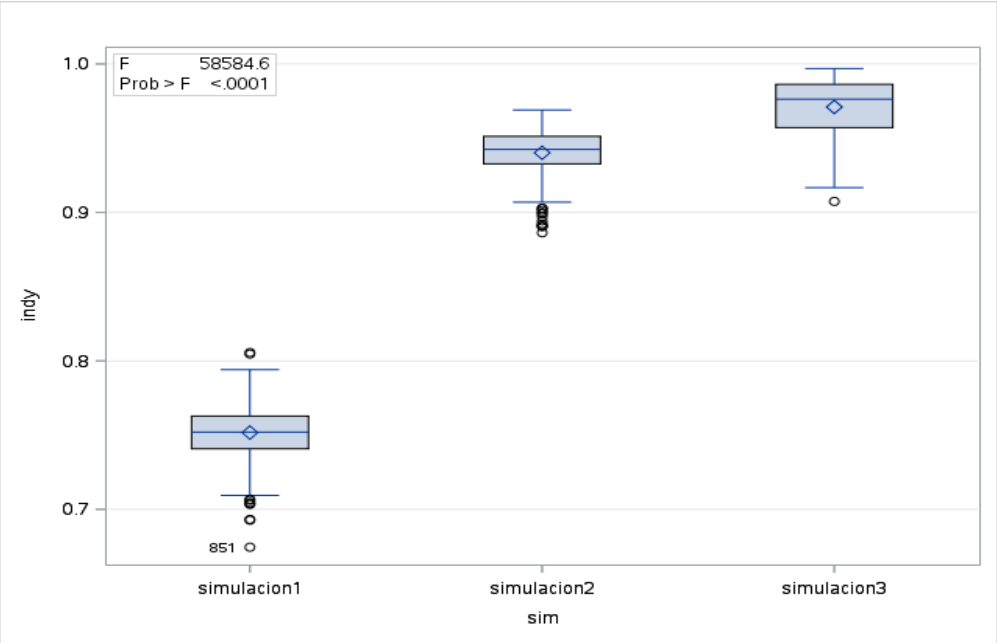


**Figura 4-10**: Box-plot de los índices por simulación basados en 1000 simulaciones para el método bayesiano

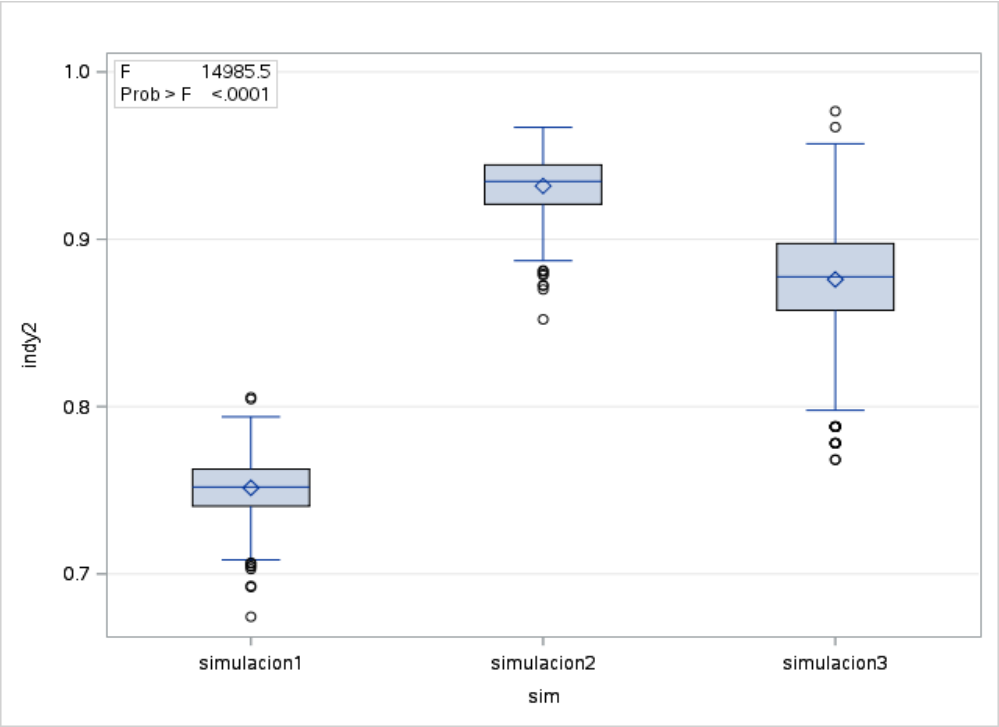


**Figura 4-11**: Box-plot de los índices por simulación basados en 1000 simulaciones para el método bayesiano bootstrap

También se realiza un análisis inferencial para identificar si existen diferencias significativas de cada método en las tres simulaciones con el test de Kruskal-Wallis, información presentada en la tabla 4-7, la cual evidencia que existe diferencia en cada método en los tres escenarios osea que cada método para estimar los IC para las tasas estandarizadas difieren de una simulación a otra. También se realizaron comparaciones múltiples de cada método entre los escenarios uno a uno, tabla 4-8, en la cual se evidencia con el test de Mann-Withney que en casi todos existen diferencias significativas en cada método entre escenarios uno a uno, solo para el caso del bootstrap se evidencia que no hay diferencias entre al simulación 2 y 3.

**Tabla 4-7**: Resultados comparaciones múltiples de cada método en las simulaciones basados en el test de Kruskal-Wallis

| **Test de Kruskal-Wallis** | | | | | |
|---|---|---|---|---|---|
| **Escenarios** | **Bootstrap** | **Momentos** | **Gamma** | **Bayesiano** | **Bayesiano bootstrap** |
| **Valor p** | 0,0004 | < ,0001 | 0,0005 | < ,0001 | < ,0001 |

**Tabla 4-8**: Resultados comparaciones múltiples de las simulaciones por cada método basados en el test de Mann-Withney

| **Escenarios** | **Bootstrap** | **Momentos** | **Gamma** | **Bayesiano** | **Bayesiano bootstrap** |
|---|---|---|---|---|---|
| **simulacion1-sumilacion2** | < ,0001 | < ,0001 | < ,0001 | < ,0001 | < ,0001 |
| **simulacion1-sumilacion3** | 0,0110 | < ,0001 | < ,0001 | < ,0001 | < ,0001 |
| **simulacion2-sumilacion3** | 0,1186 | < ,0001 | 0,0005 | < ,0001 | < ,0001 |

### 4.2.3. Análisis convencional de intervalos de confianza

En la tabla 4-9 se presentan los resultados de las medias de los IC para simulaciones de 100 intervalos, la cobertura y la amplitud de los intervalos para cada uno de los escenarios propuestos, tasa baja de eventos, media y grande, 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 en los escenarios de tasas de eventos baja y mediana, pero en amplitud son los que presentan mayor valor, dato que contribuye a un menor valor del índice, para el caso de tasas grandes la cobertura del intervalo de confianza bayesiano bootstrap tiene un promedio muy inferior a los otros siendo este dato el que afecta el índice en este caso. En cobertura presenta un buen desempeño el método gamma el cual tiene unos valores superiores a los métodos de momentos y bootstrap, además presenta menor amplitud proporcionado un valor del índice superior, lo que ayuda a identificar este como el método de mejor desempeño.

Con respecto a la amplitud de los intervalos para los casos bayesiano y bayesiano bootstrap, estos tienen una amplitud mayor en los tres escenarios, pero como se especificó anteriormente tenían valores de coberturas inferiores 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 presentan 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, aclarando que son sensibles al tamaño de tasas, el bayesiano a medida que aumenta la tasa aumenta su desempeño, y el bayesiano boostrap por lo contrario tiene un buen desempeño en las tasas medias, pero en las pequeñas y más que todo en las tasas más grandes pierde cobertura pasando de un 96 % a un 87 % .

**Tabla 4-9**: Promedio de Intervalos de confianza por método y por escenario

| IC simulación tasas de eventos pequeñas | | | | |
|---|---|---|---|---|
| | Límite inferior | límite superior | Cobertura | Amplitud |
| Bootstrap | 1.155 | 5.271 | 0.950 | 4.116 |
| Momentos | 1.350 | 5.481 | 0.947 | 4.131 |
| Gamma | 1.528 | 5.589 | 0.948 | 4.061 |
| Bayes | 1.363 | 6.576 | 0.962 | 5.213 |
| Bayes bootstrap | 1.346 | 6.520 | 0.962 | 5.174 |
| IC simulación tasas de eventos medias | | | | |
| 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 |
| IC simulación tasas de eventos grandes | | | | |
| Bootstrap | 185.794 | 223.302 | 0.947 | 37.508 |
| Momentos | 186.136 | 223.645 | 0.949 | 37.509 |
| Gamma | 186.378 | 223.880 | 0.950 | 37.502 |
| Bayes | 185.355 | 225.626 | 0.974 | 40.271 |
| Bayes bootstrap | 178.340 | 217.983 | 0.878 | 39.643 |

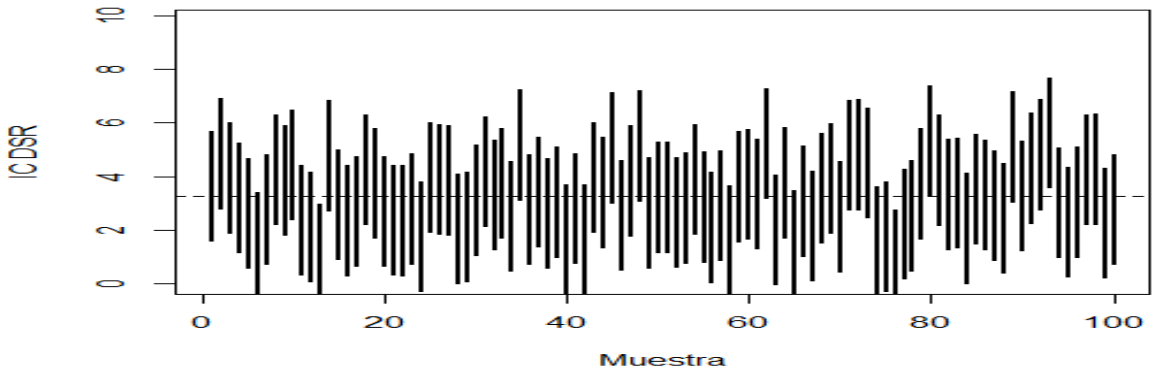


**Figura 4-12**: Cobertura método bootstrap tasas pequeñas

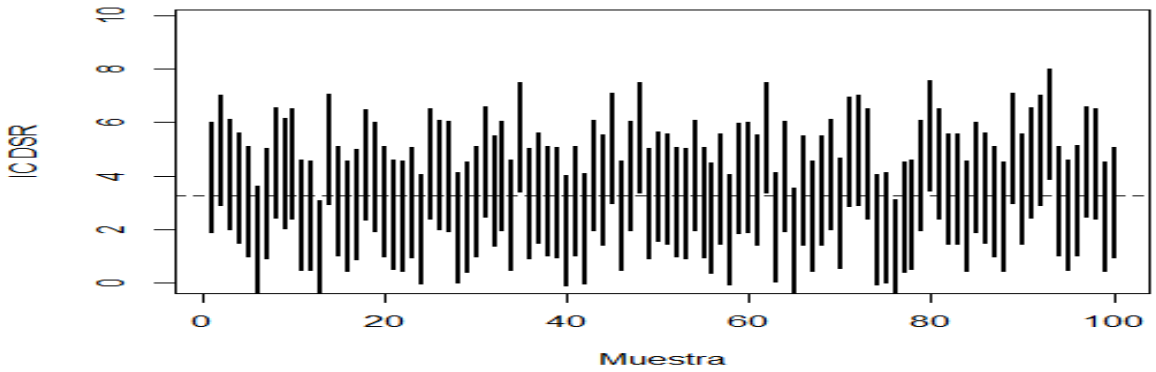


**Figura 4-13**: Cobertura método momentos tasas pequeñas

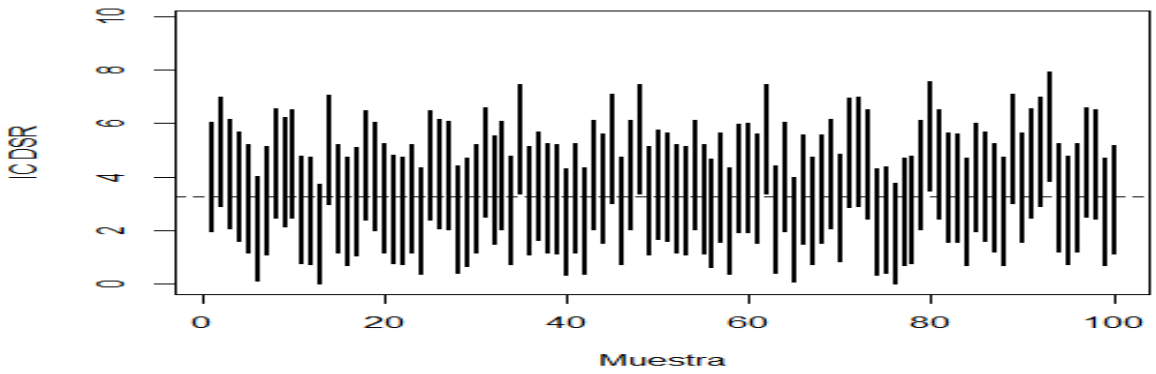


**Figura 4-14**: Cobertura método gamma tasas pequeñas

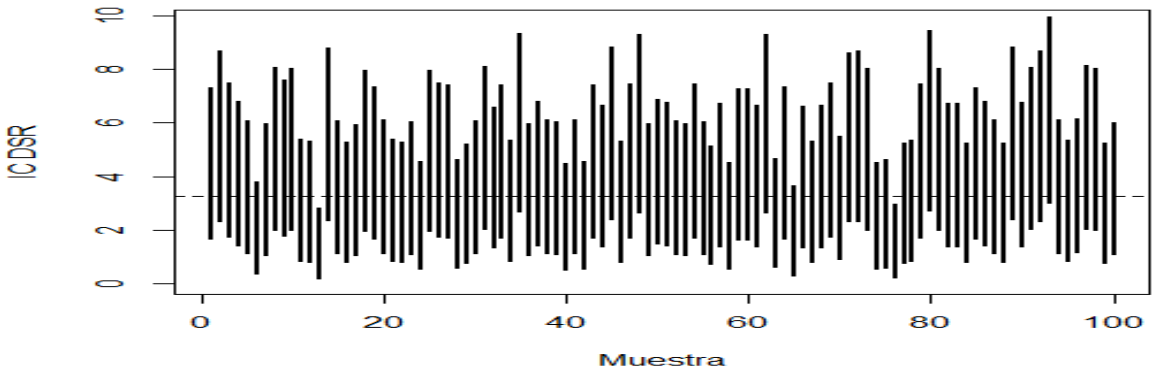


**Figura 4-15**: Cobertura método bayesiano tasas pequeñas

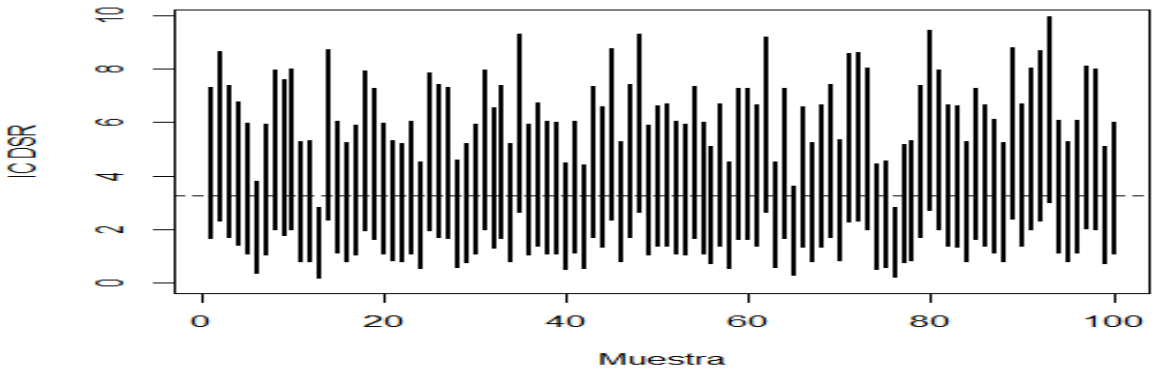


**Figura 4-16**: Cobertura método bayesiano bootstrap tasas pequeñas

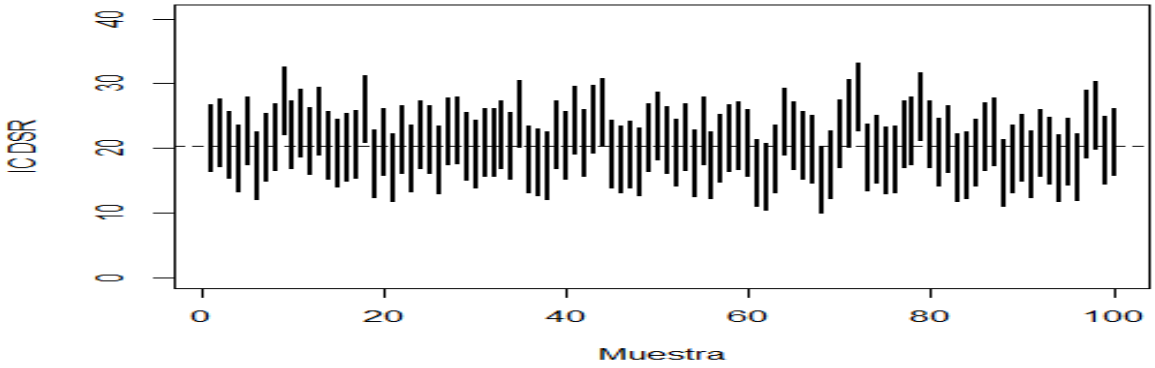


**Figura 4-17**: Cobertura método bootstrap tasas medias

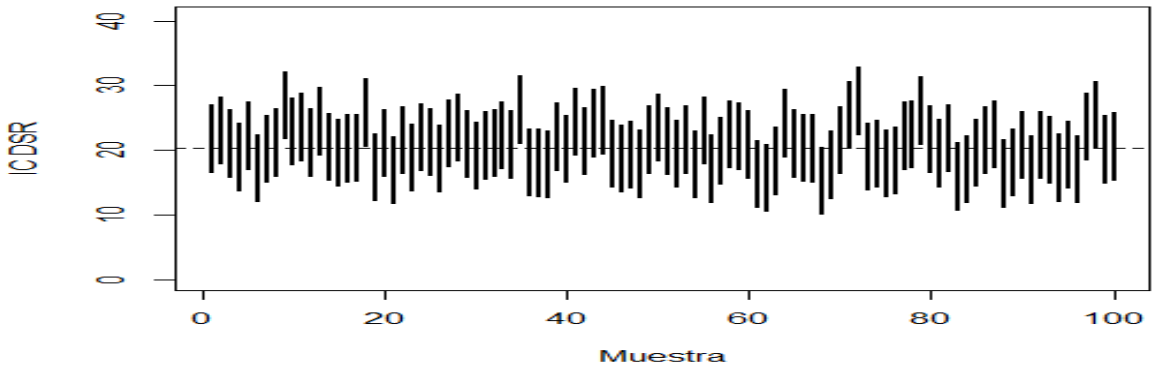


**Figura 4-18**: Cobertura método de momentos tasas medias

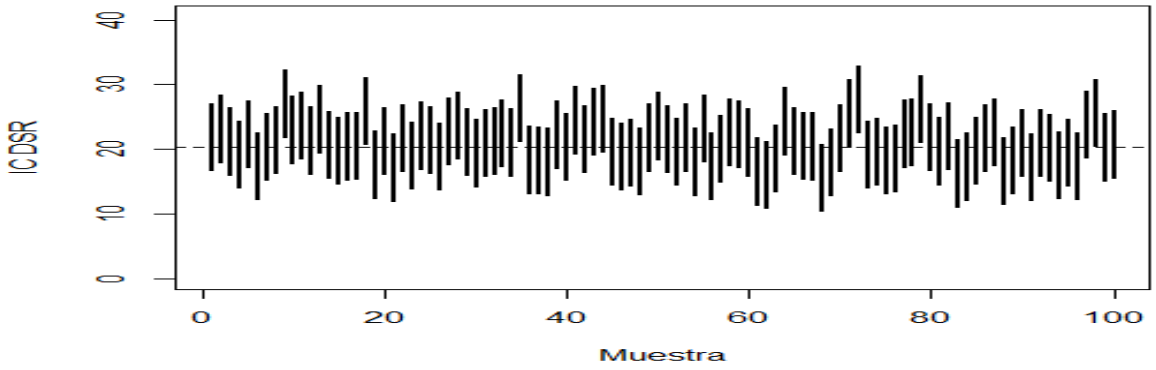


**Figura 4-19**: Cobertura método gamma tasas medias

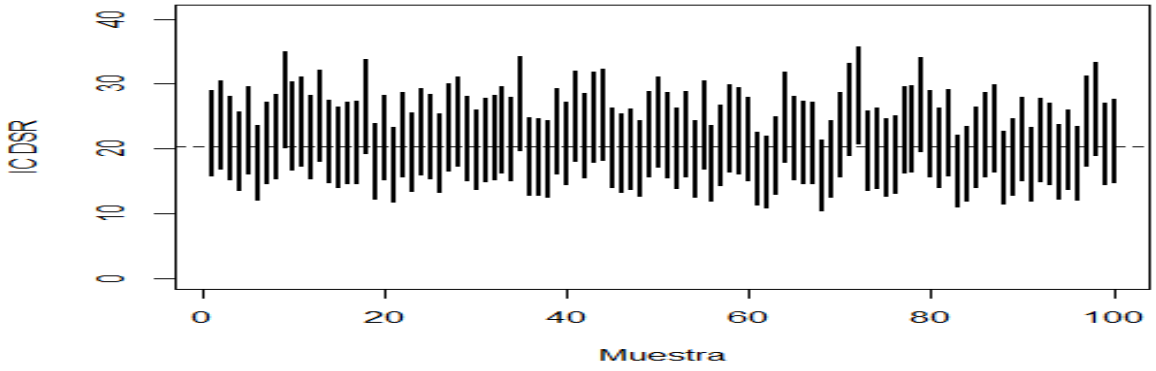


**Figura 4-20**: Cobertura método bayesiano tasas medias

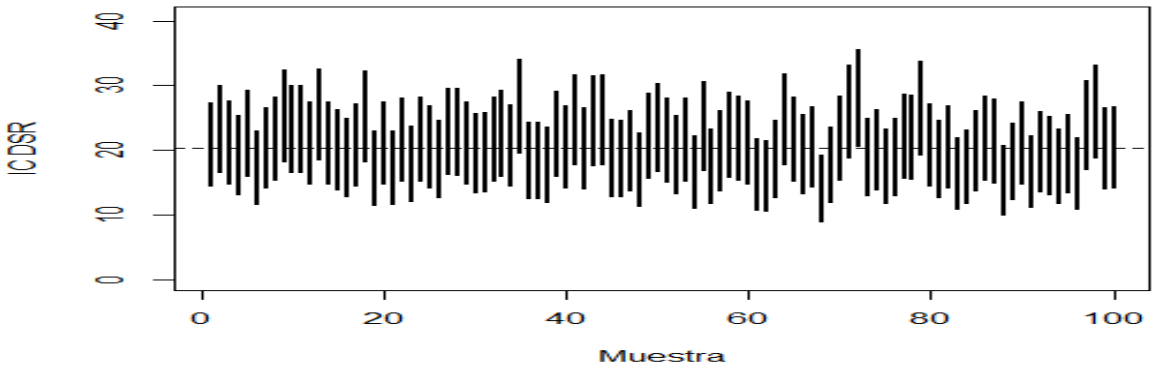


**Figura 4-21**: Cobertura método bayesiano bootstrap tasas medias

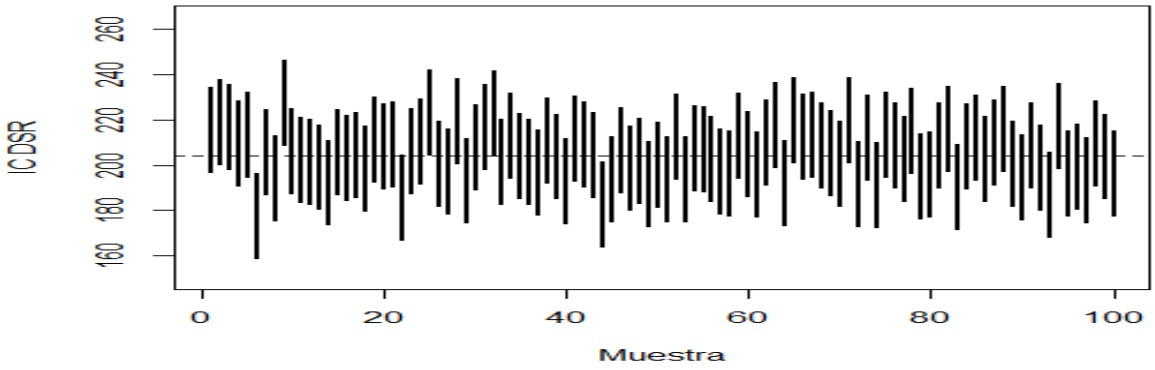


**Figura 4-22**: Cobertura método bootstrap tasas grandes

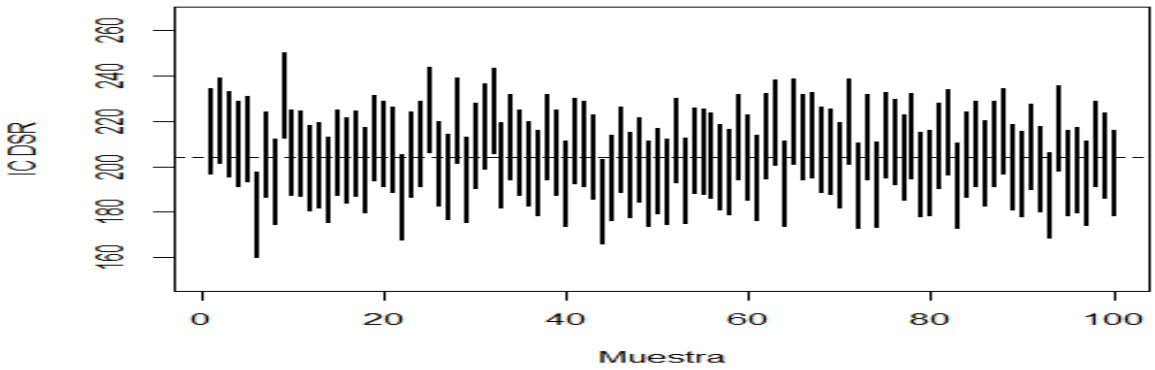


**Figura 4-23**: Cobertura método de momentos tasas grandes

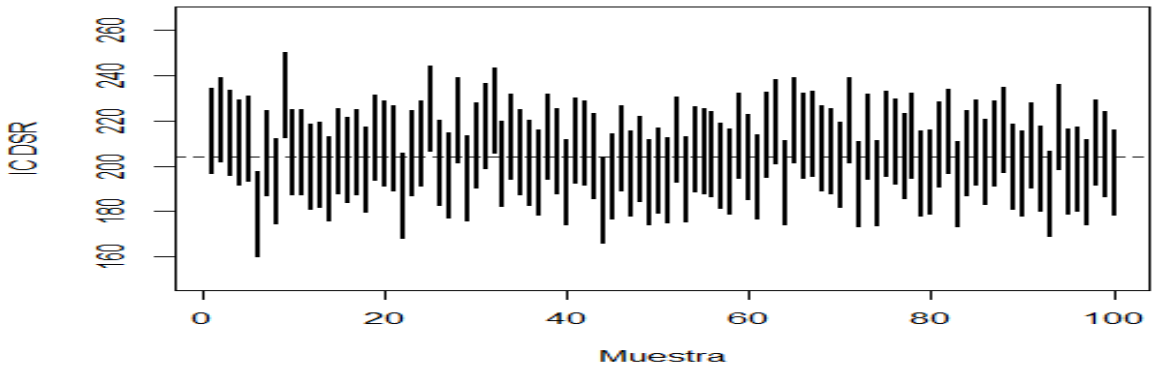


**Figura 4-24**: Cobertura método gamma tasas grandes

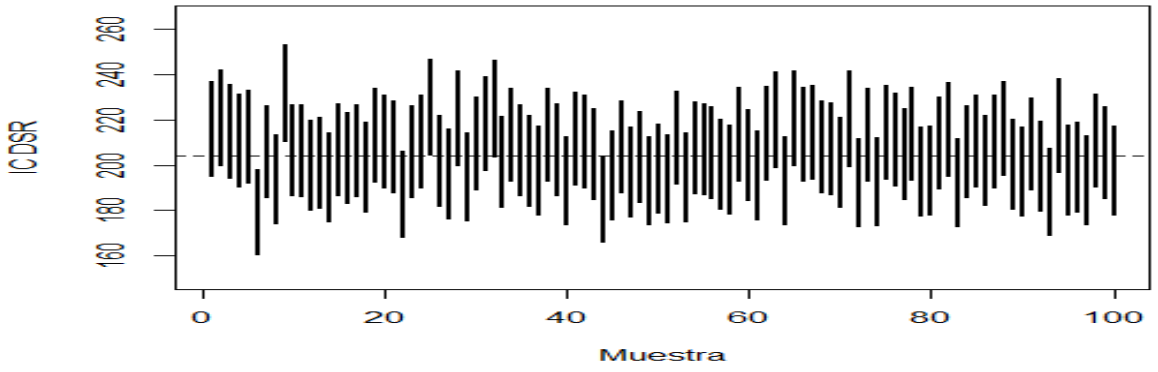


**Figura 4-25**: Cobertura método bayesiano tasas grandes

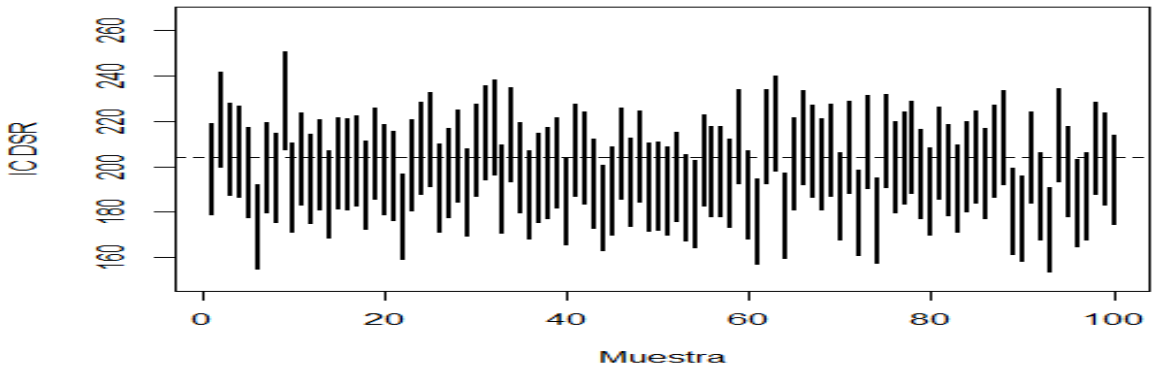


**Figura 4-26**: Cobertura método bayesiano bootstrap tasas grandes

En las figuras 4-12 a la 4-26 se visualizan 100 simulaciones de los intervalos de confianza para la estimación de las tasas estandarizadas directas para los tres escenarios evaluados, en las cuales se pueden observar tanto la simetría de los intervalos como las coberturas. Se observa como los métodos bootstrap y de los momentos, figura 12 y 13, para el escenario de tasas pequeñas son pocos simétricos y presentan límites inferiores menores que cero, lo cual evidencia una dificultad de estos IC en el caso de tasas pequeñas. Con respecto a las coberturas se observa cada uno de los niveles especificados anteriormente y su consistencia en cada uno de los escenarios.

## 4.3. Aplicación: Datos de cáncer de cérvix

A manera de ejemplo se presentan las tasas estandarizadas directas de mortalidad de cáncer de cérvix en mujeres mayores de 20 años distribuidas en quinquenios en Colombia en los departamentos: Antioquia, Valle del Cauca y la ciudad de Bogotá, esta selección de departamentos y ciudades fue propuesta por ser las tres principales regiones del país y a fin de facilitar la ilustración.

Para los datos, tabla 4-10, se tomaron los reportes de muertes por cáncer de cérvix reportados en los registros de defunción del Departamento Administrativo Nacional de Estadística [DANE, 2017a] igualmente se consultaron las poblaciones proyectadas para el año 2015 según el censo del año 2005 para cada una de las poblaciones de interés, también se tomó como población estándar el total de la población colombiana femenina por grupos de edad quinquenales, proyectada para el 2015 con el censo 2005 [DANE, 2017b], con la cual se calcularon las proporciones con respecto a la población total de cada grupo de edad en la población de mujeres, para tomarla como referencia en la estandarización de las tasas de mortalidad por cáncer de cérvix de cada uno de los departamentos y así tener comparabilidad entre las regiones especificadas, para identificar indirectamente (atendiendo al hecho de que las tasas se utilizan como una medida del riesgo) en cual de las tres poblaciones existe un mayor riesgo en las mujeres de morir por cáncer de cérvix.

En la tabla 4-11 se presentan las tasas estandarizadas directas de mortalidad por cáncer de cérvix en mujeres mayores de 20 años en Colombia en los departamentos de Antioquia, Valle del Cauca y la ciudad de Bogotá con los intervalos de confianza calculados aplicando los métodos bootstrap, momentos, gamma, bayesiano y bayesiano bootstrap, se observan los resultados de los intervalos calculados con los diferentes métodos en los cuales se visualizan

**Tabla 4-10**: Muertes por cáncer de cérvix y distribución de la población en Colombia en el año 2015 por departamento

| | Defunciones por tumor maligno de la mama | | | población por región y nacional de mujeres mayores de 20 años | | | |
|---|---|---|---|---|---|---|---|
| Grupo edad | Antioquia | Bogotá | Valle del cauca | Antioquia | Bogotá | Valle del cauca | población estándar |
| De 20 a 24 | 1 | 1 | 1 | 283575 | 331072 | 199015 | 0,086 |
| De 25 a 29 | 5 | 2 | 4 | 274616 | 316585 | 191119 | 0,080 |
| De 30 a 34 | 4 | 10 | 9 | 250025 | 344416 | 182506 | 0,074 |
| De 35 a 39 | 11 | 19 | 14 | 227416 | 321621 | 168715 | 0,068 |
| De 40 a 44 | 23 | 38 | 23 | 204587 | 281441 | 155689 | 0,061 |
| De 45 a 49 | 34 | 55 | 38 | 215569 | 274714 | 159005 | 0,062 |
| De 50 a 54 | 53 | 82 | 54 | 207941 | 263777 | 152044 | 0,058 |
| De 55 a 59 | 66 | 76 | 52 | 174994 | 222154 | 126696 | 0,048 |
| De 60 a 64 | 41 | 53 | 49 | 135918 | 171404 | 100361 | 0,037 |
| De 65 a 69 | 40 | 75 | 43 | 102445 | 128566 | 75501 | 0,029 |
| De 70 a 74 | 41 | 51 | 34 | 72464 | 90087 | 55067 | 0,021 |
| De 75 a 79 | 32 | 41 | 39 | 53880 | 62014 | 44555 | 0,016 |
| De 80 y más | 66 | 82 | 51 | 57573 | 63400 | 45561 | 0,017 |
| Total | 417 | 586 | 411 | 3301422 | 4068770 | 2378331 | 0.655 |

**Tabla 4-11**: Tasas estandarizadas directas (DSR) de mortalidad por cáncer de mama por departamento e intervalos de confianza por método

| Metodo | Antioquia | Bogotá | Valle del Cauca |
|---|---|---|---|
| DSR | 0.1241 | 0.1363 | 0.1612 |
| Bootstrap | (0.1121,0.1346) | (0.1260,0.1465) | (0.1473,0.1768) |
| Momentos | (0.1132,0.1357) | (0.1264,0.1469) | (0.1471,0.1765) |
| Gamma | (0.1131,0.1356) | (0.1263,0.1467) | (0.1469,0.1763) |
| Bayesiano | (0.0903,0.1687) | (0.1044,0.1770) | (0.1164,0.2206) |
| Bayesiano b | (0.0883,0.1659) | (0.1046,0.1777) | (0.1189,0.2255) |

**Tabla 4-12**: Índice por cada método y regiones

| Región | Bootstrap | Momentos | Gamma | Bayesiano | Bayesiano b |
|---|---|---|---|---|---|
| Antioquia | 0.928455 | 0.919484 | 0.918740 | 0.975673 | 0.967291 |
| Bogotá | 0.937999 | 0.927843 | 0.928010 | 0.977633 | 0.972323 |
| Valle del Cauca | 0.918577 | 0.948849 | 0.948512 | 0.975382 | 0.969613 |

diferencias en cada unos de los intervalos de confianza de las poblaciones, en particular para los métodos bayesiano y bayesiano bootstrap presentan una amplitud mucho mayor de los intervalos. Este caso corresponde al escenario simulado para tasas grandes donde se evidencia que el método bayesiano bootstrap pierde cobertura, en el caso del ejemplo los intervalos bayesianos presentan una mayor amplitud pero como tienen un nivel de cobertura muy superior no alcanza a penalizar el índice, como se visualiza en la tabla 4-12 la cual presenta los índices por método en las tres regiones, y se evidencia que estos métodos tienen un mayor valor.
Se identifica que la población en la cual se presentan más casos de defunciones en mujeres mayores de 20 años por cáncer de cérvix es en el Valle del Cauca, seguido de Bogotá y por último Antioquia es el que tiene la menor tasa de mortalidad por esta causa en mujeres de las edades especificadas.

## 4.4. Discusión

Para la estimación de las tasas estandarizadas directas a través de intervalos de confianza en un evento dado se han desarrollado diferentes métodos que generan intervalos de diferentes calidades, algunos con formas más complejas que dificultan su utilización además de la complejidad intrínseca de las tasas estandarizadas directas, las cuales en el proceso de estandarización dificultan la aplicación de diferentes métodos, partiendo desde los pesos por los cuales se ponderan cada uno los $k$ grupos, el tamaño de las poblaciones y el tamaño de la tasa de eventos, dependiendo si es un evento poco frecuente en la población estudio o si es un evento muy numeroso o común.

Es importante tener en cuenta el proceso de estandarización, en el caso directo a diferencia del indirecto donde se modela el evento y luego se estandariza, como lo representa [Woodward, 2005], en la siguiente expresión:

$$SER = \frac{\text{observaciones de número de eventos}}{\text{número de eventos esperados}} = \frac{e}{E}$$

$e$ es el número de eventos en la población en estudio y E es la tasa de la población estándar por lo que se define de manera global la tasa estandarizada, en la tasa estanzarizada directa se realiza una suma ponderada de los eventos por los pesos de cada uno de los $k$ grupos.

Partiendo del supuesto que los eventos en estudio tienen una distribución Poisson, existen diferentes apriori no informativas que permiten modelar un evento de esta naturaleza. Pero se presentan dificultades al momento de estimar las tasas estandarizadas directas las cuales son sumas de eventos que se distribuyen Poisson por los pesos de la razón de la población estándar y la población en estudio.

Los intervalos de confianza resultado de los métodos analizados presentan unos desempeños que son identificados por el índice propuesto siendo consecuente con el resultado concluido al analizar las coberturas y amplitudes por separado de los intervalos de confianza en cada método en los tres escenarios simulados. Aunque el índice puede en algunos casos presentar una falacia en el sentido que el intervalo tenga un nivel de cobertura superior y un tamaño de amplitud medio y si está comparado con otros métodos que también tienen cobertura buenas y amplitud mínimas pueden entrar a clasificar métodos con buen desempeños con los resultados de los índices ya que cualquiera de los factores pueden ayudar a generar un mayor valor, esta situación se presenta cuando las diferencias en coberturas y amplitudes entre los métodos son mínimas.

Según lo propuesto por [Fay and Kim, 2017], es consecuente con los resultados de los intervalos de confianza evaluados en el presente trabajo y tratado por [Fay and Kim, 2017] donde obtienen buenas propiedades para el método gamma, partiendo de este plantea una propuesta para mejorar su desempeño, con el cual [Nelson, 2017], construye una librería de R "dsrtest"la cual considera una prueba de hipótesis nula simple sobre una tasa directamente estandarizada y obtiene el intervalo de confianza correspondiente, usando un método seleccionado, teniendo en cuenta algunas opciones como por ejemplo: el IC basado en la gamma, en el método bootstrap, una asintotica y una basada en la distribución beta. Aunque la librería no considera el escenario cuando en un estudio se tiene como población objeto de estudio un subgrupo de la población y asume que siempre las ponderaciones de la población estándar suman 1 y no siempre necesariamente suman 1, se puede realizar estudios de subconjuntos de la población, como lo simulado en el presente trabajo donde se simula un evento para una población de 25 a 60 años y como se evidencia en el ejemplo de comparación de tasas estandarizadas directamente sobre el evento las mujeres con cáncer de cérvix mayores de 20 años donde no se incluye el total de la población sino un subconjunto por lo tanto no se puede asumir que el factor de estandarización siempre suma 1.

# Capítulo 5

# Conclusiones, recomendaciones y direcciones futuras

## 5.1. 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 H, 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, adicionalmente estos últimos presentan variaciones en la calidad de los intervalos al cambiar el tamaño del número de eventos, o sea al cambio en los tamaños de tasas de eventos presentados en la población estudiada, para tasas pequeñas presentan un menor índice, para tasa medianas se incrementa y luego para tasas grandes vuelve y pierden cobertura.

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 y en el caso de las tasas pequeñas alcanzan a superar a este último, pero en los escenarios de tasas medianas y grandes están por debajo, en general estos tres métodos no varían tanto en los diferentes escenarios evaluados, 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 representa la calidad de los intervalos de confianza de las metodologías evaluadas permitiendo identificar los mejores desem-

peños tanto en cobertura como amplitud conjuntamente. Evaluando la cobertura y amplitud de los intervalos de confianza se encuentra que el método que mejor desempeño presentó en los tres escenarios evaluados es el gamma, conclusión que es confirmada por el índice que dio como resultados mayores valores para el método gamma seguido del de momentos y bootstrap, por último el bayesiano y bayesiano boostrap son los de menor índice penalizado por su amplitud.

## 5.2. Recomendaciones

Con base en los resultados del estudio de simulación se recomienda utilizar para la estimación por intervalos de las tasas estandarizadas directas el método gamma, el cual proporciona intervalos de mejor calidad y no varía sustancialmente su desempeño en los diferentes escenarios considerados.

## 5.3. Direcciones futuras

Explorar los métodos bayesianos y posibles formas para la suma ponderada de las tasas estandarizadas, que permita identificar la distribución aposteriori que se ajuste más a las caraterísticas de las tasas estandarizadas directas.

# Bibliografía


[Ahmad et al., 2001] Ahmad, O. B., Boschi-pinto, C., and Lopez, A. D. (2001). Age standardization of rates: a new WHO standard. (31):1–14.

[Correa, 2007] Correa, J. (2007). Estimación por intervalo del parámetro de la distribución de poisson con una sola observación. *Revista Colombiana de Estadśtica*, 30(1):69–75.

[Correa J, 2003] Correa J, S. E. (2003). Intervalos de confianza para la comparación de dos proporciones. *Revista Colombiana de Estadística*, 26(1):61–75.

[DANE, 2017a] DANE (2017a). Defunciones no fetales 2010.

[DANE, 2017b] DANE (2017b). Proyecciones de Población.

[Diciccio, 1991] Diciccio, T. (1991). More accurate confidence intervals in exponential families. Technical Report 368, Stanford University, Stanford, California.

[Diciccio T, 1995] Diciccio T, E. B. (1995). Bootstrap confidence intervals. Technical Report 473, Stanford University, Stanford, California.

[DiCiccio T, 1996] DiCiccio T, E. B. (1996). Bootstrap confidence intervals. *Statistical Science*, 11:189–212.

[Dobson A, 1991] Dobson A, Kuulasmaa K, E. E. S. J. (1991). Confidence intervals for weighted sums of poisson parameters. *Statistics In Medicine*, 10:457–462.

[Eduardo et al., 2017] Eduardo, C., Cardona, M., and García-perdomo, H. A. (2017). Incidence of penile cancer worldwide : systematic review and meta-analysis. pages 1–10.

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

[Fay, 1999] Fay, M. P. (1999). Approximate confidence intervals for rate ratios from directly standardized rates with sparse data. *Communications in Statistics - Theory and Methods*, 28(9):2141–2160.

[Fay and Kim, 2017] Fay, M. P. and Kim, S. (2017). Confidence intervals for directly standardized rates using mid-p gamma intervals. *Biometrical Journal*, 59(2).

[Fay M, 1997] Fay M, F. E. (1997). Confidence intervals for directly standardized rates: A method based on the gamma distribution. *Statistics In Medicine*, 16:791–801.

[Hon Keung T, 2008] Hon Keung T, Filardoa g, Z. g. (2008). Confidence interval estimating procedures for standardized incidence rate. *Computational Statistics and Data Analysis*, 52(7):3501–3516.

[Johnson N, 2005] Johnson N, K. A. K. S. (2005). *Univariate Discrete Distributions*. John Wiley y Sons, New York.

[Keung H, 2008] Keung H, Filardo, Z. G. (2008). Confidence interval estimating procedures for standardized incidence rates. *Computational Statistics and Data Analysis*, 52:3501–3516.

[Krishnamoorthy et al., 2016] Krishnamoorthy, K., Peng, J., and Zhang, D. (2016). Modified large sample confidence intervals for poisson distributions: Ratio, weighted average, and product of means. *Communications in Statistics - Theory and Methods*, 45(1):83–97.

[Nelson, 2017] Nelson, M. (2017). Package dsrTest. (31):1–14.

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

[Schoenbach, 1999] Schoenbach, V. J. (1999). Estandarización de tasas y razones.

[Swift, 1995] Swift, M. (1995). Simple confidence intervals for standardized rates based on the approximate bootstrap method. *Statistics in Medicine*, 14:875–1888.

[Swift, 2010] 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, 2006] Tiwari R, Clegg L, Z. Z. (2006). Efficient interval estimation for age-adjusted cancer rates. *Statistical Methods in Medical Research*, 15:547–567.

[William Meeker, 2017] William Meeker, Gerald Hahn, L. E. (2017). *Statistical Intervals: A Guide for Practitioners and Researchers*. Wiley.

[Woodward, 2005] Woodward, M. (2005). *Epidemiology study design and data analisis*. Florida: Chapman y Hall/crcl.

# Apéndice A

# Anexo: Programas en R

```
library(bootstrap)
library(ggplot2)
#FUNCION PARA IC BAYES#
icbayesxi<-function(e,pi,psi,nivel){
ei<-c(rpois(8,lambdaei)) #muestra de eventos
w<-c(psi/pi) #pesos por grupo de edad i
as<-matrix(nrow=8,ncol=2)
for(i in 1:8){
as[i,]<-qgamma(c((1-nivel)/2,1-(1-nivel)/2),
c(e[i]+0.5,e[i]+0.5),c(1,1))
}
RLby<-100*(sum(w*as[,1]));RUby<-100*(sum(w*as[,2]))
return(icby<-c(RLby,RUby))
}
#FUNCION PARA BAYES CON BOOTSTRAP#
icbayesxi2<-function(e,pi,psi,nivel){
ei<-c(rpois(8,lambdaei)) #muestra de eventos
w<-c(psi/pi) #pesos por grupo de edad i
as<-matrix(nrow=8,ncol=2)
for(i in 1:8){
as[i,]<-qgamma(c((1-nivel)/2,1-(1-nivel)/2),
c(e[i]+0.5,e[i]+0.5),c(1,1))
}
RLby<-100*((w*as[,1]));RUby<-100*((w*as[,2]))
return(icby<-c(RLby,RUby))
}
#INICIO DE VARIABLES
#poblacion en estudio por grupos de edad quinquenales de 25 y mas
#lambdaei<-c(3,3,5,6,5,7,9,8)
```

```
#lambdaei<-c(35,36,11,32,52,70,134,178)
#lambdaei<-c(3570,3670,1130,3230,5240,7030,13450,17830)
#poblacion en estudio
pi<-c(1960736,1725481,1391860,1322185,979958,883643,894299,841838)
# Poblacion mundial OMS 2000-2025
psi<-c(7.93,7.61,7.15,659,6.04,5.37,4.55,11.95)
# tasa estandariza poblacional
dsrp<-100*sum(lambdaei*psi/pi)
tind<-NULL;ev<-NULL;tmdsr<-NULL;medi<-matrix(nrow = 20, ncol = 12)
ind<-NULL;vmdsr<-NULL;vmdsr<-NULL;g<-1000;indb<-NULL;indm<-NULL
indg<-NULL;wrb<-NULL;wrm<-NULL;wrg<-NULL;wrby<-NULL;vcnrb<-NULL
vcnrm<-NULL;vcnrg<-NULL;wrby2<-NULL;vcnby2<-NULL;vcnby<-NULL
indby2<-NULL;indby<-NULL;amb<-NULL;amm<-NULL;amg<-NULL;ambay<-NULL
meanIC<-matrix(nrow = g,ncol = 5)
#ciclo para 1000 simulaciones de \'indices
for(l in 1:g){
#muestra de eventos seudoaleatorio Poisson
ei<-matrix(rpois(8000,lambdaei), ncol =8, byrow = T)
dsr<-NULL;ssei<-NULL; aux<-NULL;sei<-matrix(nrow = n, ncol = 8);
w<-matrix(nrow = n, ncol = 8);des<-NULL;ssei<-NULL;n<-1000
as<-matrix(nrow = n, ncol = 2);icby<-matrix(nrow = n, ncol = 2)
icby2<-matrix(nrow = n, ncol = 2)
#ciclo para 1000 simulaciones de IC
for ( i in 1:n) {
w<-c(psi/pi) #pesos por grupo de edad i
dsr[i]<-100*(sum(w*ei[i,])) # tasa estandarizada
des[i]<-sqrt(sum(c((w^2)*ei[i,])))#desviacion
ssei[i]<-sum(ei[i,])#sumatoria de eventos
icby[i,] <-icbayesxi(ei[i,],pi,psi,0.95)# a, b ic bayesianos
by2<-matrix(c(icbayesxi2(ei[i,],pi,psi,0.95)),
nrow = 8, ncol = 2,byrow=F)
B<-200
azar<-matrix(NA,B,8)
medianas<-matrix(NA,B,2)
for(k in 1:B){
azar[k,]<-sample(1:8,replace=T)
azar1<-by2[azar[k,],]
medianas[k,]<-apply(azar1,2,sum)
}
icby2[i,]<-apply(medianas, 2, median)
}
#simulacion bootstrap
```

```
z<-NULL; a<-NULL;x <- dsr
for (i in 1:n) {theta <- function(x){mean(x)}
results <- bcanon(x,100,theta)
z[i]<-results$z0 #Z-0
a[i]<-results$acc #a
}
#BOOTSTRAP
RLb<-dsr+((z-qnorm(0.975))/((1-a*(z-0.975))^2))*sd(dsr)#lim inferior
RUb<-dsr+((z+qnorm(0.975))/((1-a*(z+0.975))^2))*sd(dsr)#lim superior
#MOMENTOS
#X inferior metodo momentos
XL<-(ssei+0.5)*((1-(1/(9*(ssei+0.5)))-
((qnorm(0.975))/(3*sqrt(ssei+0.5)))))^3
#X superior momentos
XU<-(ssei+0.5)*((1-(1/(9*(ssei+0.5)))+
((qnorm(0.975))/(3*sqrt(ssei+0.5)))))^3
RLm<-dsr+(sqrt((sd(dsr))^2/ssei))*(XL-ssei)# limite inferior momentos
RUm<-dsr+(sqrt((sd(dsr))^2/ssei))*(XU-ssei)# limite superior momentos
#GAMA
# grados de libertad limite inferior
gll<-(2*dsr^2)/(sd(dsr))^2
# grados de libertad limite superior
glu<-(2*(dsr+(max(w)))^2)/(((sd(dsr))^2)+(max(w))^2)
pchi2<-(qchisq(0.025,gll))^-1 # chi cuadrado limite inferior
pchi2u<-(qchisq(0.975,glu))^-1 # chi cuadrado limite inferior
RLG<-((sd(dsr))^2/(2*dsr))/pchi2 #Limite infeior gama
RUG<-((sd(dsr))^2/(2*dsr))/pchi2u #limite superior gama
IC<-matrix(round(c(RLb,RUb,RLm,RUm,RLG,RUG,icby,icby2,dsr),4),
ncol = 11, nrow = n, byrow = F,
colnames(c("RLb","RUb","RLm","RUm","RLG","RUG","dsr")))
# conteo de coberturas
cnrb<-0;cnrm<-0;cnrg<-0; cnby<-0;cnrby2<-0;cnrby<-0
for (h in 1:n) {if(RLb[h]<mean(dsrp))if(mean(dsrp)<RUb[h])
cnrb<-cnrb+1 }
for (h in 1:n) {if(RLm[h]<mean(dsrp))if(mean(dsrp)<RUm[h])
cnrm<-cnrm+1 }
for (h in 1:n) {if(RLG[h]<mean(dsrp))if(mean(dsrp)<RUG[h])
cnrg<-cnrg+1 }
for (h in 1:n) {if(icby[h,1]<mean(dsrp))if(mean(dsrp)<icby[h,2])
cnrby<-cnrby+1 }
for (h in 1:n) {if(icby2[h,1]<mean(dsrp))if(mean(dsrp)<icby2[h,2])
cnrby2<-cnrby2+1 }
```

```
#amplitud intervalos
meanIC[l,]<-c(mean(abs(IC[,1]-IC[,2])),mean(abs(IC[,3]-IC[,4])),
mean(abs(IC[,5]-IC[,6])),mean(abs(IC[,7]-IC[,8])),
mean(abs(IC[,9]-IC[,10])))
#calculos de pesos de densidades limites superior-inferioir
#densidad limite inf bootstrap
fxlsrb<-pnorm(IC[,1], mean =mean(IC[,1]),sd =sd(IC[,1]), log =F)
#densidad limite sup bootstrap
fxlurb<-pnorm(IC[,2], mean =mean(IC[,2]),sd =sd(IC[,2]), log =F)
dfxrb<-fxlsrb-fxlurb # diferencia bootstrap
wrb[l]<-mean((1-abs(dfxrb))) # 1 -diferencia bootstrap
#densidad limite inf momentos
fxlsrm<-pnorm(IC[,3], mean =mean(IC[,3]),sd =sd(IC[,3]), log =F)
#densidad limite sup momentos
fxlurm<-pnorm(IC[,4], mean =mean(IC[,4]),sd =sd(IC[,4]), log =F)
dfxrm<-(fxlurm)-(fxlsrm) # diferencia momentos
wrm[l]<-mean((1-abs(dfxrm))) # 1 -diferencia bootstrap
#densidad limite inf gama
fxlsrg<-pnorm(IC[,5], mean =mean(IC[,5]),sd =sd(IC[,5]), log =F)
#densidad limite sup gama
fxlurg<-pnorm(IC[,6], mean =mean(IC[,6]),sd =sd(IC[,6]), log =F)
dfxrg<-fxlurg-fxlsrg # diferencia gama
wrg[l]<-mean(1-abs(dfxrg)) # 1 -diferencia gama
#densidad limite inf gama
fxlsrby<-pnorm(IC[,7], mean =mean(IC[,7]),sd =sd(IC[,7]), log =F)
#densidad limite sup gama
fxlurby<-pnorm(IC[,8], mean =mean(IC[,8]),sd =sd(IC[,8]), log =F)
dfxrby<-fxlurby-fxlsrby # diferencia bayes
wrby[l]<-mean(1-abs(dfxrby)) # 1 -diferencia bayes
#densidad limite inf gama
fxlsrby2<-pnorm(IC[,7], mean =mean(IC[,7]),sd =sd(IC[,7]), log =F)
#densidad limite sup gama
fxlurby2<-pnorm(IC[,8], mean =mean(IC[,8]),sd =sd(IC[,8]), log =F)
dfxrby2<-fxlurby2-fxlsrby2 # diferencia bayes boostrap
wrby2[l]<-mean(1-abs(dfxrby2)) # 1 -diferencia bayes bootstrap
vcnrb[l]<-(cnrb/(n));vcnrm[l]<-(cnrm/(n))
vcnrg[l]<-(cnrg/(n));vcnby[l]<-(cnrby/(n))
vcnby2[l]<-(cnrby2/(n))
indb[l]<-wrb[l]*vcnrb[l] #indice bootstrap
indm[l]<-wrm[l]*vcnrm[l]  #indice momentos
indg[l]<-wrg[l]*vcnrg[l]   #indice gama
indby[l]<-wrby[l]*vcnby[l]  #indice bayes
```

```
indby2[l]<-wrby2[l]*vcnby2[l]   #indice bayes
}
#consolidacion de resultados
mamp<-round(apply(meanIC, 2, mean),3)
df<-matrix(round(c(indb,indm,indg,indby,indby2,vcnrb,vcnrm,
vcnrg,vcnby,vcnby2,wrb,wrm,wrg,wrby,wrby2),3),
ncol = 15,nrow = g)
dfe<-matrix(c(c(rep("indb",l),rep("indm",l),rep("indg",l),
rep("indy",l),rep("ind2y",l)),
round(c(indb,indm,indg,indby,indby2),4)),
ncol = 2, nrow = (5*l), byrow = F )
write.table(dfe, "misdatossas.txt", sep = ",",
quote = F, row.names = F)
dft<-data.frame(df)
par(mfrow=c(1,1))
plot(seq(1:length(dft$X1)),dft$X1,type = "l",
col=1,ylim=c(0.850, 1.05))
lines(seq(1:length(dft$X1)),dft$X2,col=2)
lines(seq(1:length(dft$X1)),dft$X3,col=3)
lines(seq(1:length(dft$X1)),dft$X4,col=4)
lines(seq(1:length(dft$X1)),dft$X5,col=5)
par(mfrow=c(1,1));l<-100
plot(range(1,l),range(0,2*dsrp),type = "n",
xlab = "Muestra",ylab = "IC DSR")
abline(h = dsrp, lwd = 1, lty = 2)
for (i in 1:l) {lines(rep(i, 2),c(IC[i,1],IC[i,2]),  lwd = 2)}
plot(range(1,l),range(0,2*dsrp),type = "n",
xlab = "Muestra",ylab = "IC DSR")
abline(h = dsrp, lwd = 1, lty = 2)
for (i in 1:l) {lines(rep(i, 2),c(IC[i,3],IC[i,4]),  lwd = 2)}
plot(range(1,l),range(0,2*dsrp),type = "n",
xlab = "Muestra",ylab = "IC DSR")
abline(h = dsrp, lwd = 1, lty = 2)
for (i in 1:l) {lines(rep(i, 2),c(IC[i,5],IC[i,6]),  lwd = 2)}
plot(range(1,l),range(0,2*dsrp),type = "n",
xlab = "Muestra",ylab = "IC DSR")
abline(h = dsrp, lwd = 1, lty = 2)
for (i in 1:l) {lines(rep(i, 2),c(IC[i,7],IC[i,8]),  lwd = 2)}
plot(range(1,l),range(0,2*dsrp),type = "n",
xlab = "Muestra",ylab = "IC DSR")
abline(h = dsrp, lwd = 1, lty = 2)
for (i in 1:l) {lines(rep(i, 2),c(IC[i,9],IC[i,10]),  lwd = 2)}
```

# Apéndice B

# Anexo: programas en SAS

```
proc freq data=henry.logistica;
table metodo;
title 'Frecuencias del ï¿½ndice de acuerdo al mï¿½todo';
run;

proc sort data=henry.logistica;
by metodo;
run;

*****Enfoque parametrico;

proc univariate data=henry.logistica normal;
by metodo;
var simulacion1;
title 'Descriptivo del ï¿½ndice de acuerdo al mï¿½todo';
run;

proc gchart data=henry.logistica;
by metodo;
vbar simulacion1 / space=0;
title 'Histogramas del ï¿½ndice de acuerdo al mï¿½todo';
run;


proc glm data=henry.logistica;
class metodo;
model simulacion1=metodo;
means metodo / cldiff bon hovtest=levene;
title 'ANOVA de una vï¿½a. Enfoque paramï¿½trico';
```

```
run;
quit;

*****Enfoque no parametrico;

proc npar1way data=henry.logistica wilcoxon;
class metodo;
var simulacion1;
title 'Kruskal-Wallis. Enfoque no-paramï¿½trico';
run;

****comparaciones multiples no parametricas;

proc npar1way data=henry.logistica wilcoxon;
where metodo in ('indb','indm');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;

proc npar1way data=henry.logistica wilcoxon;
where metodo in ('indb','indg');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;

proc npar1way data=henry.logistica wilcoxon;
where metodo in ('indb','indy');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;

proc npar1way data=henry.logistica wilcoxon;
where metodo in ('indg','indm');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;

proc npar1way data=henry.logistica wilcoxon;
```

```
where metodo in ('indg','indy');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;

proc npar1way data=henry.logistica wilcoxon;
where metodo in ('indy','indm');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;

proc npar1way data=henry.logistica wilcoxon;
where metodo in ('ind2','indm');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;
proc npar1way data=henry.logistica wilcoxon;
where metodo in ('ind2','indb');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;
proc npar1way data=henry.logistica wilcoxon;
where metodo in ('ind2','indg');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;
proc npar1way data=henry.logistica wilcoxon;
where metodo in ('ind2','indy');
class metodo;
var simulacion1;
title 'Mann-Whitney. Enfoque no-paramï¿½trico';
run;
proc means data=henry.logistica n mean median std min max ndec=3;
by metodo;
var simulacion1;
run;
```