

UCLA

UCLA Previously Published Works

Title

[Uncertainty of mortality estimators and hypothesis testing: the case of Latin America and the Caribbean, 1850-2010].

Permalink

<https://escholarship.org/uc/item/58p1s215>

Journal

Notas de poblacion, 44(104)

ISSN

0303-1829

Authors

Palloni, Alberto
Beltrán-Sánchez, Hiram
Aguirre, Guido Pinto

Publication Date

2017

Peer reviewed



HHS Public Access

Author manuscript

Notas Poblacion. Author manuscript; available in PMC 2019 January 11.

Published in final edited form as:

Notas Poblacion. 2017 ; 44(104): 13–32.

Incertidumbre de los estimadores de mortalidad y pruebas de hipótesis: el caso de América Latina y el Caribe, 1850–2010

[Uncertainty of mortality estimators and hypothesis testing: the case of Latin America and the Caribbean, 1850–2010]

Alberto Palloni¹, Hiram Beltrán-Sánchez², and Guido Pinto Aguirre³

¹Center for Demography of Health and Aging (CDHA), Universidad de Wisconsin-Madison, y RAND Labor and Population, Corporación RAND

²Department of Community Health Sciences y California Center for Population Research (CCPR), Universidad de California en Los Ángeles (UCLA)

³Center for Demography of Health and Aging (CDHA), Universidad de Wisconsin-Madison

Resumen

Proponemos un procedimiento simple para dar cuenta de la incertidumbre que se produce al tener disponible una multiplicidad de estimadores de indicadores de mortalidad adulta en los análisis estadísticos. Consideramos situaciones donde hay estimadores alternativos del mismo parámetro de población, cada uno depende de un conjunto de supuestos que pueden superponerse, y algunos, o todos, podrían caracterizar erróneamente el parámetro objetivo. La incertidumbre surge como resultado de la sensibilidad variable de los estimadores a las violaciones de supuestos o la falta de información sobre las condiciones que generan los datos utilizados para calcular los estimadores. El procedimiento que aquí se propone permite al investigador utilizar todos los estimadores (plausibles), en lugar de tener que elegir solo uno que ex ante se considera “el mejor o el correcto”. Esto se logra asignando a cada estimador un puntaje de precisión que depende de: i) errores conocidos debidos a la violación de supuestos en los que se basa el estimador, y ii) probabilidades (estimadas) de que los supuestos se violen en un caso particular. Las subsiguientes inferencias sobre determinantes o tendencias de la mortalidad pueden entonces basarse en todos los estimadores y conducir a pruebas de hipótesis más robustas y conservadoras. Si bien la aplicación que utilizamos como ilustración corresponde a la mortalidad, la metodología puede aplicarse a cualquier tipo de parámetro demográfico.

Abstract

We propose a simple procedure to address the uncertainty that arises when multiple estimators of adult mortality indicators are available in statistical analyses. We consider situations in which there are alternative estimators for the same population parameter, each one depending on a set of potentially overlapping assumptions, and some or all potentially characterizing the target parameter erroneously. Uncertainty arises because of the varying sensitivity of estimators to assumption violations or lack of information about how estimators have been calculated. The

proposed procedure allows researchers to use all of the (plausible) estimators, instead of having to choose only one that, ex ante, is considered “the best or right one”. This is achieved by assigning a precision score to each estimator depending on: (i) known errors attributable to violation of the assumptions on which the estimator is based, and (ii) (estimated) probability that the assumptions are violated in one particular case. The ensuing inferences on mortality determinants or trends can now be based on all estimators, leading to more robust and conservative hypotheses tests. Notwithstanding its use for mortality in this article, the methodology can be applied to any type of demographic parameter.

Palabras clave

Mortalidad; América Latina; esperanza de vida; incertidumbre; *bootstrap*

Keywords

mortality; Latin America; life expectancy; uncertainty; bootstrap

Introducción

En el estudio de la mortalidad de las poblaciones con datos defectuosos es habitual trabajar con varios estimadores de los mismos parámetros. Como norma, estos parámetros son estadísticas resumidas de patrones de supervivencia, como la esperanza de vida a determinadas edades. Mientras que algunos estimadores se obtienen tras aplicar métodos que ajustan por falta de integridad del registro de defunciones (población), otros se calculan mediante la combinación de estadísticas vitales (no ajustadas) y factores de ajuste de fuentes exógenas, unos son resultados de la combinación de estimadores independientes de la mortalidad adulta e infantil basados en supuestos sobre modelos de mortalidad y, finalmente, otros pueden ser el producto de cálculos de terceros que dependen de reglas computacionales de divulgación total, parcial o nula. Como consecuencia, hay un espacio que contiene estimadores $\{k_i\}$ para un único país-año i o punto de observación y, al menos superficialmente, todos ellos pueden ser igualmente plausibles. Una práctica común consiste en elegir uno de los estimadores disponibles utilizando reglas más o menos explícitas y, en la mayoría de los casos, sin considerar formalmente la incertidumbre que conlleva esa elección. Por lo general, el trabajo del investigador termina cuando un valor único de entre los disponibles en el espacio $\{k_i\}$ se presenta como una opción adecuada y se utiliza en análisis posteriores. Esta práctica podría subestimar la incertidumbre de las inferencias que se basan en los estimadores escogidos. No obstante, esto no necesariamente debe ser así. Nos proponemos mostrar que, en determinadas condiciones, se pueden considerar todos los estimadores disponibles y asociar a cada uno un puntaje de incertidumbre que puede incorporarse explícitamente a los análisis de los datos. Primero formulamos un esbozo del procedimiento para después mostrar su uso en una aplicación simple.

Cabe aclarar que la idea de calcular estimadores de parámetros de población y de simultáneamente dar cuenta de la incertidumbre no es nueva. Investigaciones recientes ofrecen plataformas alternativas para producir previsiones demográficas en virtud de un

marco bayesiano que maneja explícitamente fuentes de incertidumbre (Alkema y otros, 2012; Raftery y otros, 2012; Gerland y otros, 2014). Nuestro aporte va en el mismo sentido, pero difiere en que las fuentes de incertidumbre que consideramos hacen referencia explícita a la concordancia (o la falta de ella) entre los observables y los supuestos en los que se basa el cálculo de los estimadores.

A. Formulación del problema

Asumamos que tenemos un estimador de un parámetro de población $\tilde{\theta}_i$ que se calculó utilizando un conjunto de datos D_i para el país-año i^4 . El cálculo de la estimación solo depende de un supuesto, a , que puede ser concordante con los procesos que generaron D_i ($a = 1$) o discordante ($a = 0$). Hay una función de error, $M(\tilde{\theta}_i)$, que define la magnitud del error del estimador cuando se viola el supuesto. $M(\cdot)$ es arbitrario y puede elegirse entre diversas medidas. En esta ocasión elegimos el error absoluto, por ejemplo, $M(\tilde{\theta}_i) = |\tilde{\theta}_i - \theta_i|$, y consideramos que el estimador es preciso cuando la medida de error, $M(\tilde{\theta}_i)$, se mantiene por debajo de un valor umbral arbitrario ε . La cantidad de interés es la probabilidad de que $M(\tilde{\theta}_i) < \varepsilon$, a saber,

$$P(M(\tilde{\theta}_i) < \varepsilon | D_i) = P(M(\tilde{\theta}_i) < \varepsilon | A = 0, D_i)P(A = 0 | D_i) + \quad (2.1) \\ + P(M(\tilde{\theta}_i) < \varepsilon | A = 1, D_i)P(A = 1 | D_i).$$

En este caso simple, la precisión de la estimación solo depende de un supuesto,

$$P(M(\tilde{\theta}_i) < \varepsilon | A = 1) \sim 1. \text{ Si esto es así}$$

$$P(M(\tilde{\theta}_i) < \varepsilon | D_i) \sim P(M(\tilde{\theta}_i) < \varepsilon | A = 0, D_i)P(A = 0 | D_i) + P(A = 1 | D_i). \quad (2.2)$$

Se podría asignar valores a los dos términos desconocidos en la expresión (2.1), calcular $P(M(\tilde{\theta}_i) < \varepsilon | D_i)$ y utilizar esta cantidad explícitamente como medida de incertidumbre en cualquier análisis que incluya a $\tilde{\theta}_i$. La expresión (2.1), no obstante, está excesivamente simplificada. En la práctica, la precisión de los estimadores que nos interesan para este análisis depende de la concordancia entre las condiciones empíricas que generan los datos D_i y múltiples supuestos asociados a un método de cálculo. En el caso más general, procedemos de la siguiente manera: supongamos que elegimos el método j para calcular un estimador $\tilde{\theta}_{ij}$ y que el cálculo involucra un conjunto de supuestos $A_j = \{a_1, a_2, \dots, a_r\}$, cada uno de los cuales puede ser violado (o no) por condiciones empíricas contenidas en los datos D_i . Sea Φ_{jk} un vector de unos y ceros que representa la configuración k th de supuestos contenida en el conjunto de todas las configuraciones posibles de supuestos $R = 2^r$. Solo una de estas configuraciones representa el caso de concordancia perfecta entre supuestos y

⁴En la mayoría de los casos, un único dato puntual consiste en una unidad de observación (país) y un período de tiempo (año). Para evitar confusiones y a menos que se indique lo contrario, utilizaremos un único subíndice i para referirnos tanto a la unidad de observación como al período de tiempo.

observables. Para mayor conveniencia, permitamos que esta sea la última en la lista o la configuración R^{th} , de modo que $\Phi_j R = (1, 1, \dots, 1)$. Este vector representa condiciones que producen un valor óptimo de $\tilde{\Theta}_{ij}$ dados los datos Di . Para completar la descripción, definimos medidas de probabilidad para cada vector Φ_{jk} y reescribimos la expresión anterior de una manera más general:

$$p(\tilde{\Theta}_{ij}) = P(M(\tilde{\Theta}_{ij}) < \varepsilon | Di) = \sum_{n=1}^R P(M(\tilde{\Theta}_{ij}) < \varepsilon | snj = 1, Di) P(snj = 1 | Di) \quad (2.3)$$

donde $snj = 1$ cuando el vector de supuestos que caracteriza a las condiciones empíricas contenidas en Di es Φ_{jn} y 0 de lo contrario y $P(snj = 1 | Di)$ es la probabilidad de que las condiciones representadas por el vector Φ_{jn} sean concordantes con Di . El lado derecho de la expresión (2.3) es el producto de un vector fila ($1 \times R$) ρ_{1j} de probabilidades condicionales de error para cada configuración de supuestos y un vector columna ($R \times 1$) ρ_{2j} de probabilidades de ocurrencia de configuraciones R .

Esta definición de la configuración de supuestos requiere dos simplificaciones. Primero, asumimos que una clasificación 0/1 para evaluar la concordancia entre supuestos y datos es suficiente, aunque podría ocurrir que la naturaleza o magnitud, y no solo la existencia o no existencia de la discordancia, también sea relevante. Segundo, en la expresión (2.3) asumimos que el número de supuestos es pequeño. Cuando los métodos de estimación dependen de más de dos o tres supuestos, el valor de R puede aumentar muy rápido y la evaluación de la expresión (2.3) se hace difícil. Si bien estas dos simplificaciones se pueden evitar (véase la sección E), en adelante mantendremos una configuración como la que se ha definido anteriormente, ya que simplifica la descripción del procedimiento y la discusión de la ilustración.

B. Incertidumbre de los estimadores de cabalidad relativa del registro de defunciones

Los estimadores de mortalidad adulta en países con estadísticas deficientes se calculan mediante dos tipos de técnicas: las que ajustan los datos observados para producir un estimador objetivo (como la esperanza de vida a los 5 años) o las que obtienen directamente un valor del parámetro que, a su vez, puede utilizarse para calcular factores de ajuste. Ambas técnicas son idénticas desde el punto de vista de la estimación y asumiremos que generan una estimación de la cabalidad relativa del registro de defunciones⁵. Existen aproximadamente entre 10 y 15 técnicas para estimar la cabalidad relativa de las estadísticas de mortalidad adulta en los países con estadísticas vitales deficientes y empadronamientos censales. Si bien algunas de estas técnicas son variantes de otras y comparten supuestos, la mayoría se basan en un pequeño conjunto de supuestos únicos que no se superponen y que,

⁵La integridad relativa hace referencia al error neto de una tasa de mortalidad observada, por ejemplo, el producto de errores tanto en el numerador, o el número de eventos (defunciones) observados, como en el denominador, o los años de exposición de la población observados. Se define como el cociente de M_x^0/M_x^T donde M_x^0 es la tasa de mortalidad observada en el grupo etario $(x, x + 1)$ y M_x^T es la tasa de mortalidad verdadera, desconocida.

bajo ciertas condiciones, pueden hacer que sean más (menos) atractivas en relación con las técnicas alternativas. El comportamiento de estos estimadores es una función de la concordancia entre las condiciones empíricas que producen los datos sobre los que se aplican y los supuestos invocados para su cálculo. Para establecer un vínculo entre las condiciones empíricas que producen datos observados, los supuestos necesarios para el cálculo de estimadores y la magnitud de los errores provocados por la desviación de los supuestos respecto de las condiciones empíricas, usamos simulaciones numéricas que nos permiten definir la matriz del conjunto de vectores ρ_{1j} para cada técnica seleccionada $j = 1, \dots, L$.

1. Simulaciones numéricas

En total, consideraremos 13 técnicas distintas para estimar la cabalidad relativa del registro de defunciones. Estas se utilizaron para elaborar estimadores de mortalidad ajustadas para 19 países de América Latina y el Caribe durante el período 1850–2010 (Palloni, Pinto y Beltrán-Sánchez, 2015). Los siguientes son los supuestos clave en los que se basa el grupo de técnicas consideradas: 1) estabilidad de la población, 2) invariancia etaria de la cabalidad relativa de los censos y el registro de defunciones, 3) cabalidad de dos censos contiguos, 4) cierre a la migración y 5) ausencia de declaración inexacta de la edad. Algunas técnicas apelan a todos estos supuestos, mientras que otras solo necesitan algunos de ellos. La variabilidad del comportamiento de los estimadores depende de cuán sensible es cada técnica a la violación de los supuestos en los que se basa. Solo dos de estas técnicas se basan en el supuesto de estabilidad y todas se basan en los supuestos 2 a 5. Para los métodos que no requieren estabilidad, el error producido por falta de estabilidad siempre se fija en 0. En adelante, nos referimos a estos supuestos como el conjunto $A = a1, a2, \dots, a5$.

Simulamos casi 6.000 poblaciones con un número elevado de condiciones de mortalidad, fecundidad y migración, algunas de las cuales violan uno o más de los supuestos en los que se basa el cálculo de las estimadores de mortalidad (Palloni, Pinto y Beltrán-Sánchez, 2015). Para cada población simulada, calculamos estimadores de niveles de mortalidad utilizando las técnicas seleccionadas y medidas de errores en cada caso. Para simplificar la evaluación numérica de los errores y “poblar” los vectores ρ_{1j} usamos dos métodos simplificados: **Clases de desviaciones de los supuestos**: si bien las poblaciones en el conjunto simulado pueden violar un determinado supuesto de diversas maneras (en términos de dirección y magnitud de las desviaciones), nuestra aplicación solo da cuenta de si los datos se acercan de manera aproximada a determinado supuesto o no.

- i. **Clases de desviaciones de los supuestos**: si bien las poblaciones en el conjunto simulado pueden violar un determinado supuesto de diversas maneras (en términos de dirección y magnitud de las desviaciones), nuestra aplicación solo da cuenta de si los datos se acercan de manera aproximada a determinado supuesto o no.
- ii. **Muestras de configuración de supuestos**: un conjunto de cinco supuestos, cada uno de los cuales puede estar en dos estados (concordante o discordante), produce un total de $R = 2^5 = 32$ posibles “configuraciones de desviación”. Para reducir este espacio de configuraciones, solo consideramos poblaciones

simuladas en las que los supuestos 2, 3 y 5 nunca se cumplen, ya que es más probable que estas representen bien a las poblaciones reales de América Latina y el Caribe. Con estas simplificaciones solo necesitamos considerar cuatro configuraciones de supuestos⁶.

Cabe reiterar que con la adopción de estos métodos simplificados se busca facilitar ilustraciones del procedimiento y motivar su aplicación, pero no se pretende sugerir que conducen a un tratamiento óptimo del problema. Esto último requiere, sin duda, una reducción menos simplificada del espacio de configuración de violación de los supuestos.

2. Definición empírica de los vectores ρ_{1j}

La información que se extrae de la simulación y las simplificaciones (i) y (ii) nos permite calcular el vector ρ_{1j} . Primero fijamos el valor de ϵ en 0,05 y consideramos los errores en los estimadores observados en cada una de las cuatro configuraciones posibles de supuestos definidas anteriormente, y en cada caso calculamos la frecuencia de veces que $|\tilde{\theta}_{ij} - \theta| < 0,05$ para la técnica j y la población simulada i . A continuación, calculamos la probabilidad correspondiente $P(|\tilde{\theta}_{ij} - \theta| < 0,05)$. Por último, repetimos estos cálculos para cada una de las 13 técnicas consideradas.

3. Definición empírica de los vectores ρ_{2j}

El vector ρ_{2j} para una técnica j implica estar en posesión de información sobre el caso que se está estudiando. Esta información puede tener dos fuentes. La primera es conocimiento previo sobre las condiciones empíricas que generan los datos. Por ejemplo, es bien sabido que el descenso de la mortalidad en la mayoría de los países de América Latina y el Caribe comenzó, casi con seguridad, después de 1950 y, al menos en algunos países, poco después de la finalización del siglo XIX. En cualquiera de estos casos, el supuesto de estabilidad solo se cumplirá (aproximadamente) por un período corto después del comienzo del descenso de la mortalidad pero no en las décadas cercanas a la finalización del siglo XX. Así, esta información indica que, para algunos países, cualquier configuración de los supuestos, que incluya el supuesto de estabilidad ($a_1=1$) debe tener una probabilidad cercana a 0 para cualquier conjunto de datos observado después de, aproximadamente, 1970. Del mismo modo, sabemos que los censos en los países de América Latina y el Caribe están distorsionados por la exageración sistemática de la edad, y también se sospecha que estos errores disminuyen con el transcurso del tiempo. Así, a cualquier configuración de supuestos que incluya el supuesto de información de edad adulta confiable $a_5 = 1$ debe asignarse una probabilidad de 0 para todos los años anteriores a 1980 y de más de 0 para después de 1980.

La segunda fuente de información sobre las condiciones empíricas que generan los datos es la información derivada de pruebas de diagnóstico. Estas pruebas están diseñadas para determinar si los datos producen señales que podrían indicar la violación de uno o más supuestos. Por ejemplo, la sobreestimación de edad adulta conduce a signos reveladores, ya que obliga a que los patrones de mortalidad adulta disminuyan demasiado rápido en las edades mayores o generan razones de supervivencia intercensales extremadamente

⁶Centrarse únicamente en poblaciones en las que estos supuestos nunca se cumplen producirá estimaciones extremadamente conservadoras respecto de la precisión de las estimaciones.

anómalas. Cuando se aplican estas pruebas de diagnóstico, el investigador dispone de fuentes adicionales de información que pueden ser utilizadas para asignar probabilidades para definir el vector ρ_{2j} . La definición de los vectores ρ_{2j} puede llevarse a cabo mediante dos estrategias complementarias. La primera es la propia evaluación y juicio del investigador y la segunda es la recolección de juicios procedentes de un pequeño grupo de expertos. La utilización de valoraciones de expertos cuando hay incertidumbre es una estrategia muy conocida, de uso habitual en diversos campos de investigación y que tiene antecedentes en la previsión demográfica. En este artículo usamos un vector de probabilidades estimadas por los autores tras considerar una combinación de pruebas diseñadas para identificar la violación de supuestos y la valoración sobre las tendencias demográficas pasadas. Por ejemplo, para decidir sobre la probabilidad de que un determinado país-año se vea influenciado por la declaración inexacta de la edad, implementamos pruebas para verificar la coherencia entre el empadronamiento en dos censos y los decesos intercensales (Palloni, Pinto y Beltrán-Sánchez, 2015). De modo similar, para decidir sobre las probabilidades de desviación del supuesto de estabilidad usamos contrastes entre poblaciones estables de tablas modelo y la distribución observada de la población por edades quinquenales, así como también información sobre el comienzo del descenso de la fecundidad y la mortalidad.

Es necesario agregar que la asignación de probabilidades es una etapa crucial del procedimiento que proponemos. La ilustración que aquí se describe se toma muchas libertades al simplificar esta etapa. Esto tiene como propósito facilitar la descripción de la metodología, más que llegar a conclusiones definitivas. En aplicaciones futuras del método utilizaremos una base de datos más elaborada que incluirá la asignación alternativa de probabilidades de valoraciones de expertos.

C. Manejo alternativo de la incertidumbre de los estimadores

Los valores $p(\tilde{\theta}_{ij})$ asociados a cada estimación $\tilde{\theta}_{ij}$ para el dato puntual D_i son medidas de incertidumbre de las que queremos dar cuenta en el análisis. Hay tres maneras de hacerlo.

1. Estimadores simples

La solución más simple es elegir para cada dato puntual un estimador satisfactorio, es el argumento $\arg \max_{\{\theta_{ij}\}}(p(\theta_{ij}))$, esto es, el estimador con la más alta probabilidad de estar dentro del 5% del valor verdadero del parámetro de población. Esta estrategia es similar a la estrategia estándar utilizada en la mayoría de las investigaciones que se realizan en este campo, aunque en la práctica los investigadores no calculan $p(\theta_{ij})$, sino que asignan valores subjetivos y después eligen un estimador que se piensa tiene la mayor probabilidad de ser el “correcto”. El problema con esta estrategia es que toda la información disponible contenida en estimadores alternativos se descarta y nunca se considera la posibilidad de incertidumbre.

2. Estimadores ponderados

Una solución relativamente simple consiste en utilizar ponderaciones en forma proporcional a alguna función adecuadamente estandarizada de $p(\theta_{ij})$, $g(p(\theta_{ij}))$ y definir un estimador puntual del parámetro desconocido como el promedio ponderado $\hat{\theta} = \sum_{n=1}^R g(p(\tilde{\theta}_{ij}))\theta_{ij}$. Esta estrategia puede implementarse tanto en análisis simples (por ejemplo, estimación de

una tendencia de mortalidad en el tiempo) como en casos más complejos (por ejemplo, estimación de modelos para identificar determinantes de las tendencias de mortalidad). Si bien la estrategia es interesante por su simplicidad, en realidad, oculta la incertidumbre que subyace en los estimadores, a menos que se calcule y considere explícitamente $var(\hat{\theta})\theta$. Y aun en este caso tenemos que reconocer que la incertidumbre no solo depende de la varianza observada de $\tilde{\theta}_{ij}$, sino también de la variabilidad de $p(\theta_{ij})$ generada por los resultados de pruebas de diagnóstico o valoraciones de expertos.

Otro método para incorporar la incertidumbre asociada a los θ_{ij} es el uso de ponderaciones pero esta vez directamente en la estimación de un modelo en el que el investigador está interesado. Supongamos, por ejemplo, que queremos estimar en un país la pendiente de la tendencia temporal de la esperanza de vida en 100 años y que hay múltiples estimadores θ_{ij} por año, cada una asociada con valores $p(\theta_{ij})$. Se puede adaptar un modelo que asigna a cada estimación una ponderación proporcional a $p(\theta_{ij})$. De modo similar, si uno está interesado en las relaciones entre mortalidad a edades tempranas y adultas en un período de 50 años, como ocurre cuando se elaboran tablas de mortalidad modelo, se pueden usar ecuaciones estructurales con ponderación de observaciones

3. Estimadores de remuestreo *bootstrap*

Una solución más natural consiste en utilizar el método de remuestreo *bootstrap* que considera el conjunto de estimadores $\tilde{\theta}_{ij}$ para un punto determinado (país-año) como el resultado de una realización particular de una meta “población” de estimadores. Los resultados de nuestra simulación nos permiten conocer los detalles de la distribución de errores asociados a cada estimador en determinadas condiciones definidas por la violación de supuestos. Si la población estudiada pertenece al universo de poblaciones incluidas en la simulación, un estimador cualquiera del parámetro objetivo tiene una distribución de errores conocida a partir de la simulación. Esta información es contenida en las cantidades $p(\tilde{\theta}_{ij})$ que expresan la probabilidad de que el estimador correspondiente sea una medida “precisa” del parámetro objetivo en virtud de los datos disponibles⁷. De manera alternativa, se podría pensar en $p(\tilde{\theta}_{ij})$ como una cantidad proporcional al número de veces que $\tilde{\theta}_{ij}$ sería la “opción correcta” si calculáramos los estimadores para un determinado parámetro objetivo una gran cantidad de veces.

Para implementar el método *bootstrap* procedemos de la siguiente manera: para cada punto u observación, extraemos al azar un estimador $\tilde{\theta}_{ij}$ del conjunto de estimadores disponibles para ese punto u observación con una probabilidad proporcional a $p(\tilde{\theta}_{ij})$. Se repite el proceso para todos los puntos u observaciones y, cuando cada uno tiene un estimador, estimamos el modelo deseado. El procedimiento se repite N veces y producirá N estimadores de los parámetros del modelo que nos interesa. Por último, calculamos la media o la mediana de las estimadores y obtenemos los errores estándar.

La principal diferencia entre un procedimiento basado en ponderaciones de los estimadores disponibles y el procedimiento de remuestreo *bootstrap* es que el primero genera medidas de incertidumbre de los estimadores de interés que dependen del modelo (o de como los pesos

⁷Según nuestra definición, “precisa” significa que la estimación se encuentra dentro del 5% del valor verdadero.

escogidos se manejen en la estimación del modelo). En cambio la incertidumbre de los estimadores obtenidos por el *bootstrap* no dependen del modelo, puesto que los errores estándar solo dependen del procedimiento de remuestreo *bootstrap*.

D. Una aplicación: estimación de tendencias temporales para la esperanza de vida

En esta sección proponemos una aplicación simple que consiste en la estimación de tendencias temporales de la esperanza de vida al nacer en Cuba y Guatemala en el período 1900–2010. Cuba y Guatemala no podrían ser más diferentes en lo que a calidad de estadísticas vitales se refiere. En ambos casos, los estimadores de mortalidad para el período anterior a 1950 están sujetos a niveles más altos de incertidumbre que los correspondientes al período posterior a 1950. No obstante, la magnitud de la incertidumbre es mucho mayor en el caso de Guatemala que en el de Cuba durante el todo el período analizado. A continuación se describen las peculiaridades de los datos disponibles para ambos países.

1. Modelos concurrentes

Proponemos estimar dos modelos alternativos donde la variable dependiente es la esperanza de vida al nacer en el país i y el año t , $Ei(t)$, y la única variable independiente es una medida de tiempo (el número de años desde 1900). Los modelos, logístico y Box-Cox, se definen de la siguiente manera

$$Ei(t) = ai / (1 + \exp(\beta i(t - \gamma i))) \quad (5.1)$$

$$\frac{Ei(t)^{\lambda_i - 1}}{\lambda_i} = \theta_i + \varphi_i (t^{\delta_i} - 1) / \delta_i \quad (5.2)$$

donde el índice ‘i’ se refiere al país. En los dos casos los estimadores se calculan usando mínimos cuadrados no lineales sobre datos de Cuba y Guatemala para el período 1900–2010.

El primero es un modelo logístico de tres parámetros que asume la existencia de un límite para la esperanza de vida, así como un único punto de inflexión. El segundo es un modelo Box-Cox de cuatro parámetros que no hace suposiciones sobre los límites ni la curvatura de la trayectoria temporal. Para simplificar la aplicación, fijamos el valor de δ para que sea igual a 0, de modo que la escala de tiempo es logarítmica y solo deban estimarse tres parámetros. La principal distinción entre estos dos modelos es que el modelo logístico tiene implicaciones claras sobre la naturaleza de longevidad (existencia de un límite superior), así como sobre la naturaleza y estructura de las tendencias temporales pasadas; por ejemplo, la tasa de aumento de la esperanza de vida se eleva en las etapas iniciales del descenso para después disminuir tras un tiempo bien definido (el punto de inflexión). Por su parte, el modelo Box-Cox es más flexible y no impone una estructura *ex ante* sobre la trayectoria temporal. Nuestro propósito es mostrar que las inferencias sobre la naturaleza de las

tendencias temporales —y sus implicancias para el futuro de la esperanza de vida— no siempre son tan robustas como puede pensarse cuando el investigador ignora la incertidumbre de los valores de las esperanzas de vida que se utilizan para la estimación.

Es importante destacar que, como se formuló anteriormente, la tarea de estimación de los parámetros se enfrenta no a una sino a dos fuentes de incertidumbre. La primera está relacionada con los estimadores de esperanza de vida anual en cada país. La segunda es sobre la naturaleza del modelo correcto. Si bien el procedimiento que empleamos a continuación puede implementarse para lidiar con ambos tipos de incertidumbre, la aplicación que aquí discutimos solo está dirigida al primero⁸. Para evitar el dilema que implica la incertidumbre del modelo, estimaremos arbitrariamente los modelos que parecen más adecuados para cada país: el modelo (5.1) para Cuba y el modelo (5.2) para Guatemala.

2. Datos y métodos

a) Fuentes de datos—Usamos la base de datos Latin American Mortality Database (LAMBdA) creada para respaldar el estudio empírico de la historia de las tendencias de mortalidad en los países de América Latina tras la independencia⁹. Esta base de datos documenta el período entre 1848 y 2014 y contiene alrededor de 500 tablas de mortalidad. Los datos para calcular las tablas de mortalidad para el período 1930–2010 son tasas de mortalidad por edad ajustadas por falta de cabalidad del empadronamiento censal y el registro de defunciones, así como las exageraciones de la edad en edades adultas. Estas tablas de mortalidad están disponibles sobre una base anual. Las tablas de mortalidad para la mayoría de los países antes de 1930 se basan en la aplicación de métodos de población estable generalizada a las estadísticas vitales parciales o la información censal. Elegimos un máximo de diez estimadores diferentes de esperanza de vida al nacer por cada período de cinco años entre 1900 y 2015. No todos los países-años incluyen todos los estimadores y, además, los períodos previos a 1950 incluyen estimadores calculados con métodos que no se usan en el período posterior a 1950.

Para generar probabilidades de que los estimadores estén dentro del rango del 5% seguimos dos reglas. La primera regla se aplica a estimadores calculados con técnicas que solo se aplican a las edades por encima de los 5 años. Para estas (y para otras técnicas que no utilizamos en este artículo) contamos con estimadores precisos obtenidos mediante simulaciones masivas (Palloni, Pinto y Beltrán-Sánchez, 2015). La única dificultad de esta estrategia radica en que para calcular la esperanza de vida al nacer debemos utilizar estimadores de mortalidad por debajo de los 5 años tomadas de otras fuentes (Encuestas de Demografía y Salud, Encuesta Mundial de Fecundidad) para las cuales no contamos con orientación acerca de incertidumbre. Una solución primitiva a este problema es asignar al estimador de mortalidad en edades menores a 5 de cada país-año una probabilidad igual a la de los estimadores de mortalidad adulta, una regla que supondrá un tratamiento muy injusto para los estimadores de mortalidad infantil¹⁰.

⁸Véase una aplicación completa en Palloni y Beltrán-Sánchez (2016).

⁹Véase Latin American Mortality Database (LAMBdA) [en línea] <http://www.ssc.wisc.edu/cdha/latinmortality/>.

¹⁰Esto es porque con toda seguridad nuestros estimadores de mortalidad por debajo de los 5 años dependen de datos de múltiples encuestas (WFS y DHS) y procedimientos no paramétricos que proporcionan estimadores anuales. Los únicos errores que podrían

La segunda regla se aplica a algunos estimadores de terceros o de métodos que no se incluyeron y evaluaron en nuestras simulaciones. La mayoría se aplican al período previo a 1950. En estos casos, asignamos probabilidades de acuerdo con nuestra opinión sobre la calidad de los datos utilizados y la sensibilidad de los métodos empleados a violaciones de los dos supuestos clave en los que se basan, a saber: i) la precisión de los ajustes para corregir por migración y ii) la adecuación de los patrones modelos de mortalidad escogidos¹¹.

Para replicar los resultados que se obtendrían con una de las estrategias simples identificadas anteriormente, utilizamos las medianas del conjunto de estimadores para cada país-año. Esto implica que asumimos que el investigador conoce el conjunto completo de estimadores disponibles para cada país-año, pero opta por mantenerse escéptico respecto de la incertidumbre, o más bien ignorarla, y simplemente elige la mediana de cada conjunto como el mejor estimador.

b) Características de los datos de mortalidad para Cuba y Guatemala—En gráfico 1.A se presenta el conjunto completo de estimadores de esperanza de vida al nacer que consideramos en esta aplicación. Como ya se indicó, la dispersión de los valores es considerable en las primeras dos décadas del siglo XX y disminuye con el transcurso del tiempo. Esto coincide con que las estadísticas vitales antes de 1950 son defectuosas en casi todos los países de América Latina y el Caribe y el conjunto de métodos factibles para generar estimadores durante este período es más heterogéneo que en los años posteriores a 1950. Cabe destacar que si bien el rango de estimadores para el período más reciente se reduce en ambos países, se registran persistentes contrastes, ya que Guatemala se ubica en el extremo superior del espectro de incertidumbre. Debido a que existe una gama más amplia de métodos (y de supuestos utilizados) para estimar la mortalidad cuando las estadísticas básicas son menos confiables, obtenemos una mayor variedad de estimadores en el caso de Guatemala.

Mientras que en el gráfico 1.B se presentan los valores de mediana, mínimo y máximo de los estimadores para Cuba y Guatemala, en el gráfico 1.C se muestran los valores de mediana, mínimo y máximo de las probabilidades de precisión asociadas al conjunto de estimadores para cada año con información disponible¹². No resulta sorprendente constatar que las barras de máximo-mínimo, tanto para los estimadores de esperanza de vida como para las probabilidades de precisión, son más anchas para los primeros años del período considerado que para los períodos más recientes, y para Guatemala más que para Cuba. En particular, las probabilidades de precisión abarcan un amplio rango para el período previo a 1950, mientras que los valores para el período posterior a 1950 se encuentran muy concentrados en el rango de valor extremo 0,80–0,90, que es muy pequeño. Este contraste refleja la mayor

presentarse se asocian al muestreo de las encuestas de base y al algoritmo no paramétrico. Así, la incertidumbre de estos estimadores debe ser por fuerza menor que la de los estimadores de mortalidad adulta.

¹¹Sabemos que la asignación de probabilidades de precisión para estos métodos no es la mejor. En una situación ideal, deberíamos usar simulaciones y valoraciones de expertos para evaluar la concordancia de los supuestos de los métodos y las condiciones observadas. No obstante, para los fines de este artículo es suficiente utilizar una regla de asignación que se aproxime a lo que podría concluir cualquier valoración de expertos.

¹²Recordemos que en esta aplicación estas son probabilidades de que los estimadores se encuentren dentro de un rango de no más de un 5% del valor verdadero.

incertidumbre asociada a los estimadores para el primer período, cuando estos países no contaban con estadísticas vitales completas y precisas. Existen estimadores para este período y algunos se basan en supuestos que son difíciles de verificar. Por su parte, los estimadores correspondientes al período posterior a 1950 se basan en unos pocos métodos que utilizan estadísticas vitales razonablemente precisas y para los cuales podemos evaluar errores respecto de la violación de supuestos (mediante la simulación). Asimismo, los supuestos en los que se basan estos métodos son menos y se adjudican con más facilidad.

El resumen precedente de estadísticas descriptivas confirma que la variabilidad de los estimadores alternativos es mayor para los primeros años que para los últimos, las probabilidades de precisión son mayores en los años más recientes y, finalmente, confirman que existen contrastes entre un país y otro en la magnitud de la incertidumbre de los estimadores en ambos períodos. Esto es de esperarse porque la calidad de las estadísticas vitales ha ido mejorando con el tiempo, pero sigue siendo diferente de un país a otro. Además, los métodos para estimar la mortalidad en el primer período se basan en supuestos cuya validez es más difícil de probar que aquellos en que se basan los métodos usados en los años más recientes.

3. Estimación de los modelos

Para calcular estimadores de los parámetros de los modelos creamos con el método *bootstrap* N réplicas para cada país del conjunto de estimadores de esperanza de vida al nacer para cada país-año. El muestreo se realiza con reemplazo, con ponderaciones proporcionales a las probabilidades de errores que se asocian a cada estimador. A continuación, usamos cada muestra de *bootstrap* y calculamos estimadores de los errores estándar y parámetros de cada modelo y obtenemos una distribución de cada estimador (y su error estándar) para cada uno de los dos modelos. Estas distribuciones contienen toda la incertidumbre asociada a los estimadores de la esperanza de vida. Por último, comparamos resultados e inferencias con resultados e inferencias que se obtendrían cuando el investigador ignora el caso de Cuba.

En los gráficos 2A–2C se muestra la distribución de los estimadores de los tres parámetros de la función logística para Cuba. Debido a que la tendencia temporal en este país (véase el gráfico 1.A) experimenta una desaceleración en la tasa de aumento de la esperanza de vida, una función logística obtiene valores razonables del límite o umbral de la esperanza de vida, así como también del punto de inflexión (el año en que la pendiente de la primera derivada de la tendencia temporal cambia de positiva a negativa). La dispersión de valores en los gráficos 2A–2C es un reflejo de la incertidumbre asociada solo a los estimadores de esperanza de vida que tenemos disponibles. Las líneas verticales oscuras identifican el estimador del parámetro correspondiente que se obtiene utilizando la mediana de los estimadores de esperanzas de vida para cada año. Si bien la dispersión es bastante grande en todos los casos, los parámetros que parecen más influenciados por incertidumbre son alfa y beta. Su coeficiente de variación es de tamaño moderado e igual a 0,17 y 0,26, respectivamente.

a) El caso de Guatemala—En los gráficos 3.A y 3.B se muestran la constante y la pendiente de la tendencia temporal en un modelo Box-Cox¹³. En el gráfico 3.C se presenta la distribución de la esperanza de vida alrededor de 1980. Las distribuciones de los estimadores de alfa y beta aparecen muy distorsionadas con colas muy gruesas, una propiedad que no presagia nada bueno si uno opta por ignorar la incertidumbre y, en su lugar, utiliza estimadores simples (sus valores se identifican con líneas verticales oscuras y gruesas). Asimismo, hay que tener en cuenta que los valores predichos de esperanza de vida para 1980 implican una variación no trivial y tienen una media (y una moda) alrededor de 51–52. Esto contrasta con el valor predicho igual a 48 que se obtiene cuando el modelo Box-Cox se estima usando estimadores simples de esperanzas de vida para cada año (esto es la mediana de los estimadores para cada año).

El principal mensaje que se desprende de estos ejemplos es el siguiente: en condiciones que se aplican en forma bastante general a los países con información defectuosa, la estimación, incluso de los modelos más simples sobre la base de estrategias simples, esconderá incertidumbre que debería aparecer de manera prominente en todas las inferencias. De este modo, por ejemplo, cualquier predicción de esperanza de vida que se base en una estimación simple de tendencias temporales incluye un importante componente estocástico que se ocultará a los usuarios y ofrecerá una falsa idea de precisión de las predicciones o pronósticos.

Hay que tener en cuenta que nada de lo que hemos elaborado alude a la incertidumbre que el investigador pueda tener con respecto a los modelos que son más apropiados. La combinación de incertidumbre de los estimadores de esperanzas de vida para cada año con incertidumbre sobre el modelo genera una situación compleja que, de seguro, amplifica los efectos que tiene la existencia de valores alternativos de la esperanza de vida para cada año.

E. Conclusiones

El propósito de este artículo es simple: mostrar las consecuencias que tiene considerar explícitamente la incertidumbre de estimadores alternativos de parámetros demográficos calculados con técnicas que se basan en supuestos de precisión variable. Describimos un procedimiento simple que se basa en simulaciones para cuantificar sesgos de estimadores de parámetros demográficos cuando se violan combinaciones de supuestos. Además, recurrimos a estimadores de las probabilidades de que distintas configuraciones de supuestos concuerden (o no) con los datos observados. En principio, estas probabilidades empíricas pueden obtenerse de dos fuentes: i) pruebas de diagnóstico que proveen indicadores cuyos valores pueden traducirse en aproximaciones de probabilidades de que se cumplan supuestos individuales (o una combinación de supuestos), y ii) valoraciones de expertos sobre las condiciones empíricas que generan los datos a los que se aplican métodos para obtener estimadores del parámetro demográfico y su concordancia con la configuración de supuestos.

¹³Recordemos que en el modelo Box-Cox aquí estimado, la variable independiente es el logaritmo del número de años transcurridos desde 1900, en tanto que la escala de la variable dependiente se identifica a partir de los datos como el tercer parámetro del modelo Box-Cox.

En las aplicaciones descritas en este artículo nos basamos en simulaciones para generar información sobre la magnitud de los errores en que se incurre cuando se violan supuestos de un conjunto de técnicas para estimar la mortalidad. Utilizamos probabilidades de concordancia entre supuestos y datos basadas en nuestra opinión y en pruebas de diagnóstico en lugar de basarlas en juicios probabilísticos emitidos por un grupo de expertos. También incluimos una segunda clase de estimaciones demográficas: aquellas para las cuales no existe una evaluación numérica de los sesgos debido a desviaciones de los supuestos. Estas incluyen estimaciones de terceros totalmente documentadas, incluidas la descripción e identificación precisa de los datos y las técnicas utilizadas para generar estimadores.

Para considerar la incertidumbre de los estimadores de mortalidad propusimos un procedimiento simple de remuestreo *bootstrap* que requiere conocer las probabilidades de precisión de cada estimador dada la configuración de supuestos y las probabilidades de que una configuración de supuestos sea concordante (o discordante) con los datos.

Si bien son bastante simples, las aplicaciones que aquí se describen muestran que la incertidumbre en torno a los estimadores de mortalidad ejerce una influencia no trivial en estimadores de parámetros de modelos que definen trayectorias temporales de la esperanza de vida, así como también para los valores que resultan proyecciones o pronósticos futuros. Además de esto, la comparabilidad entre países puede verse muy comprometida, sobre todo cuando la magnitud de la incertidumbre difiere mucho en los casos que se comparan. Por último, debido a que la sensibilidad a la incertidumbre del parámetro de un modelo (en nuestro caso, de modelos relacionados con tendencias temporales de esperanza de vida) puede diferir, las inferencias sobre la naturaleza de los procesos que el analista está tratando de entender pueden resultar bastante confusas.

En los trabajos que se realicen en el futuro se podrán tomar tres caminos. Primero, en nuestra aplicación definimos arbitrariamente la precisión estableciendo un valor umbral de $\alpha < 0,05$. Esto es innecesario. De hecho, el conocimiento *ex ante* de errores debido a la violación de supuestos producidos por el estudio de simulación genera información completa sobre la distribución de errores para cada método, no solo una estimación puntual de este. Así, se puede calcular cualquier parámetro de las distribuciones de errores del estimador, incluyendo la media, centiles, así como también medidas de alcance y asimetría. Para hacer esto explícito, uno puede redefinir los vectores ρ_{1j} e incluir información sobre todos los parámetros de la distribución de errores como se considere necesario.

En segundo lugar, nos basamos en un método simplificado mediante el cual la concordancia entre supuestos y datos es captada por completo por una variable binaria. Esto puede generalizarse a costa de incrementar la complejidad de la descripción y el cálculo. Por ejemplo, la falta de estabilidad puede representarse con dos regímenes no estables: uno donde solo la mortalidad se reduce con el tiempo y otro donde tanto la mortalidad como la fecundidad se reducen siguiendo un régimen que reproduce, con alta fidelidad, la experiencia de las poblaciones estudiadas. De modo similar, puede haber más de un tipo de desviación de lo observado con respecto al supuesto de invariancia etaria en la cabalidad de los registros censales: uno con un aumento lineal de la cabalidad y otro con una función escalonada que aumenta de manera lineal después de los 25 años. De este modo los

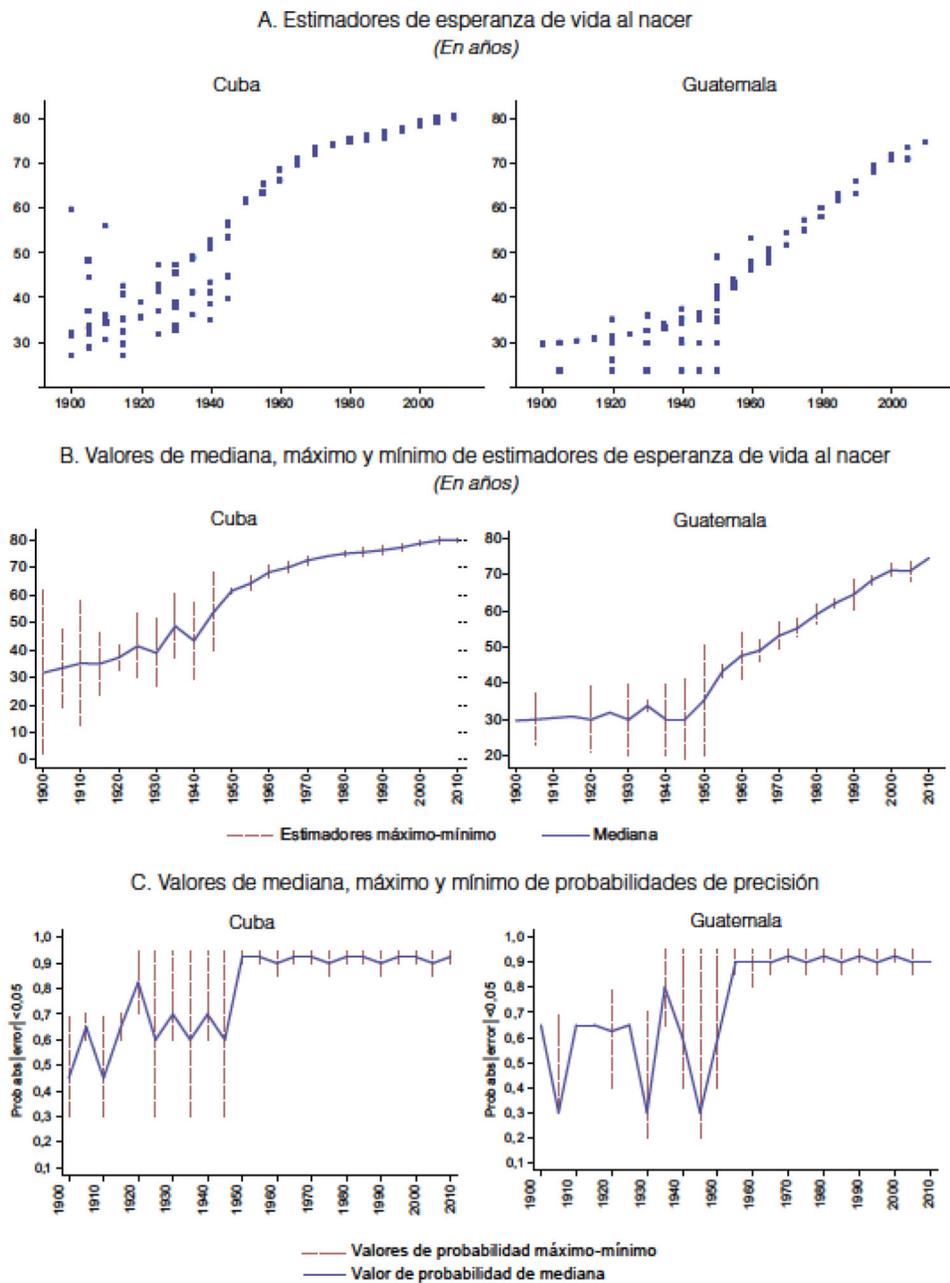
supuestos pueden no solamente ser clasificados como adecuados o no adecuados, como hicimos en este artículo, sino también caracterizarse por el grado o tipo de desviación con respecto a lo observado y los errores asociados en cada caso podrían calcularse directamente a partir de poblaciones simuladas. Si bien esto es posible, conlleva un problema: inevitablemente aumentará el conjunto de configuraciones posibles a las que debe asignarse una probabilidad. No es tan difícil asignar una probabilidad de que una población se desvíe o no de un estado estable. Más complicado es que la desviación del estado estable ocurra debido a un descenso combinado de la mortalidad y la fecundidad.

En tercer lugar, los estimadores de parámetros de modelos son sensibles a desviaciones de *configuraciones* completas de supuestos con respecto a datos empíricos, y no solo a las desviaciones de supuestos considerados separadamente. En cualquier aplicación concreta es importante caracterizar lo más completamente posible todo el espacio de configuraciones de supuestos. Debido a que estos espacios pueden llegar a ser muy grandes y complejos, el investigador debe reducirlos e incluir solo aquellos que sean plausibles, realistas o pertinentes respecto del caso empírico estudiado. No obstante, incluso tras una exitosa reducción del espacio, uno se debe enfrentar al problema de que la mayor complejidad de las configuraciones complica la asignación de probabilidades a cada configuración.

Por último, deberíamos analizar las condiciones que intensifican (atenúan) el papel de la incertidumbre de los parámetros demográficos. ¿Importan tanto, más o menos cuando se estiman las tendencias temporales de cada país? ¿Qué tan influyentes son cuando el investigador hace proyecciones a futuro de indicadores demográficos cuya tendencia temporal se ha estimado previamente? ¿Son más o menos importantes cuando se estiman medidas de tendencia central, de dispersión o distribuciones? ¿Cuánto más relevantes pueden ser cuando se estiman parámetros demográficos que dependen de cantidades más específicas, como las tasas de mortalidad por edad, donde los errores pueden tener una influencia considerablemente mayor?

Bibliografía

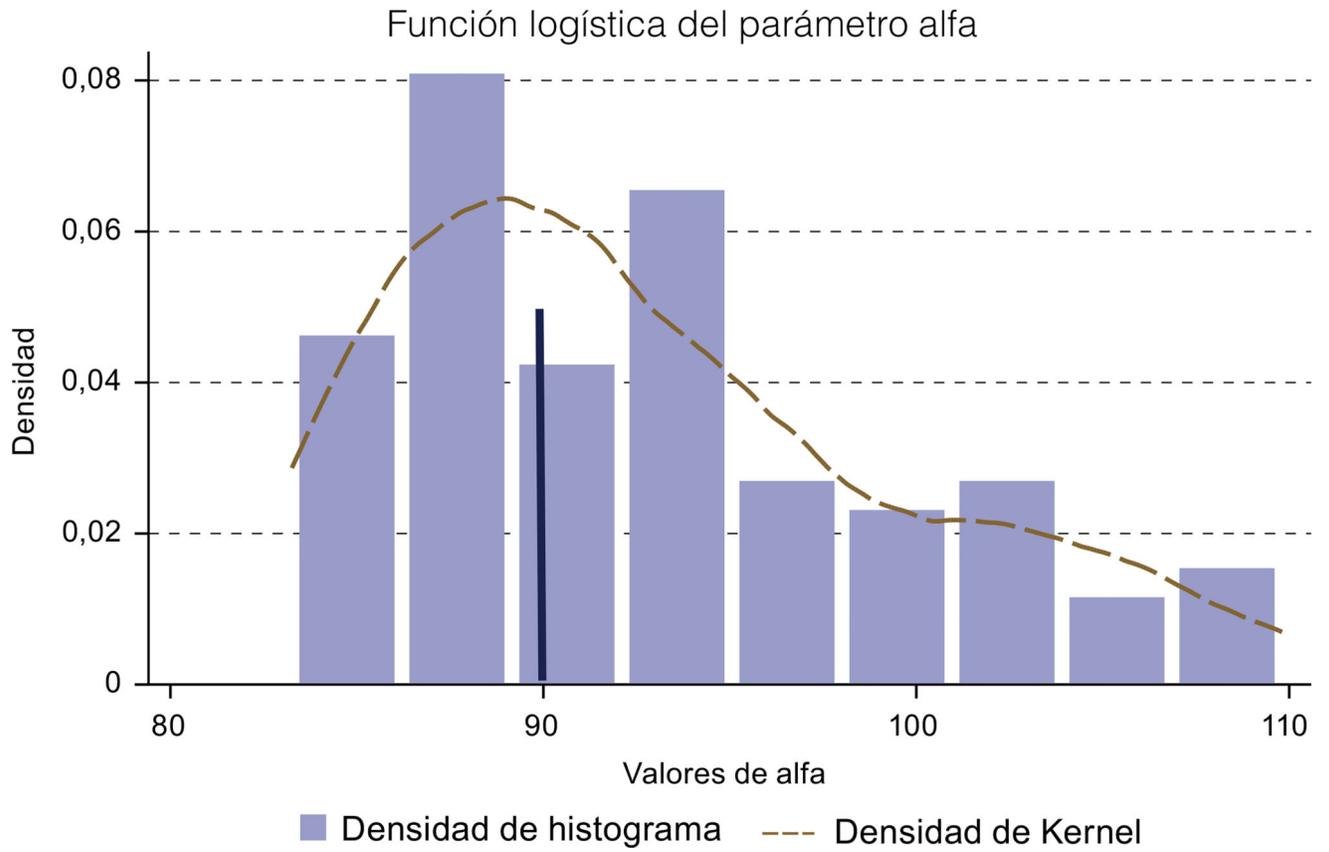
- Alkema, Ly; otros, y. Demographic Research. Vol. 26. Rostock: Max Planck Institute for Demographic Research (MPIDR); 2012. Estimating trends in the total fertility rate with uncertainty using imperfect data: examples from West Africa. abril
- Gerland, P, otros, y. Science. Vol. 346. Washington, D.C: Asociación Estadounidense para el Progreso de la Ciencia, octubre; 2014. World population stabilization unlikely this century.
- Palloni, A, Beltrán-Sánchez, yH. Estimation of mortality and the treatment of parameter and model uncertainty. documento presentado en la Reunión Anual 2016 de la Asociación Demográfica de los Estados Unidos de América; Washington, D.C: 2016. 31 de marzo a 2 de abril
- Palloni, A; Pinto, G; Beltrán-Sánchez, yH. Estimation of Life Tables 1850– 2010: Adjustments for Relative Completeness and Age Misreporting. Two Centuries of Mortality Decline in Latin America: From Hunger to Longevity. 2015. [en línea] http://www.ssc.wisc.edu/cdha/latinmortality/wp-content/uploads/book_Ch2.pdf
- Raftery, A, otros, y. Proceedings of the National Academy of Sciences. Vol. 109. Washington, D.C: Academia de Ciencias de los Estados Unidos, agosto; 2012. Bayesian probabilistic population projections for all countries.

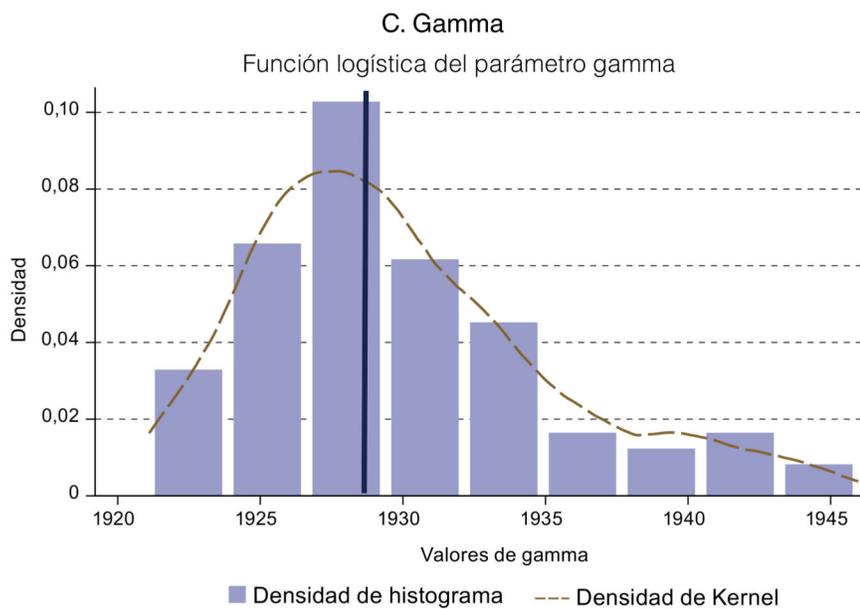
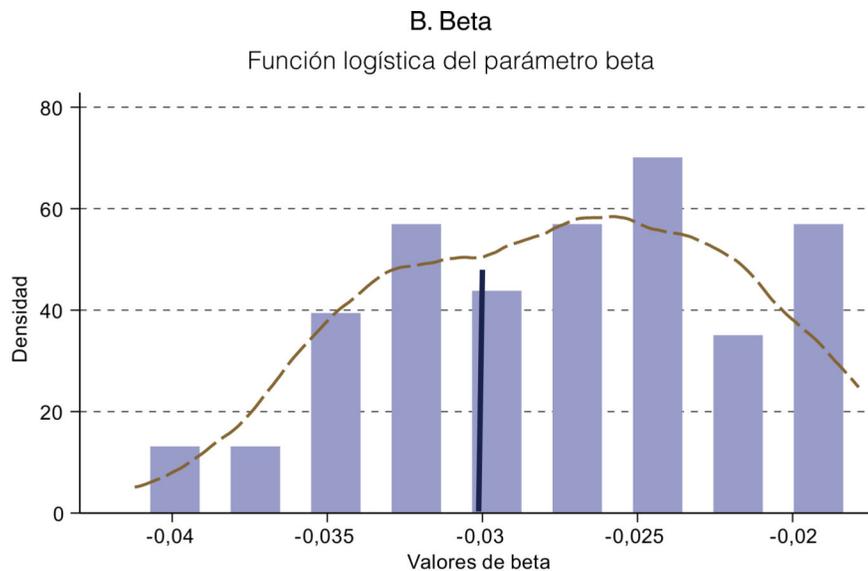


Fuente: Elaboración propia.

Gráfico 1.
Cuba y Guatemala: resumen de estadísticas de los datos

A. Alfa



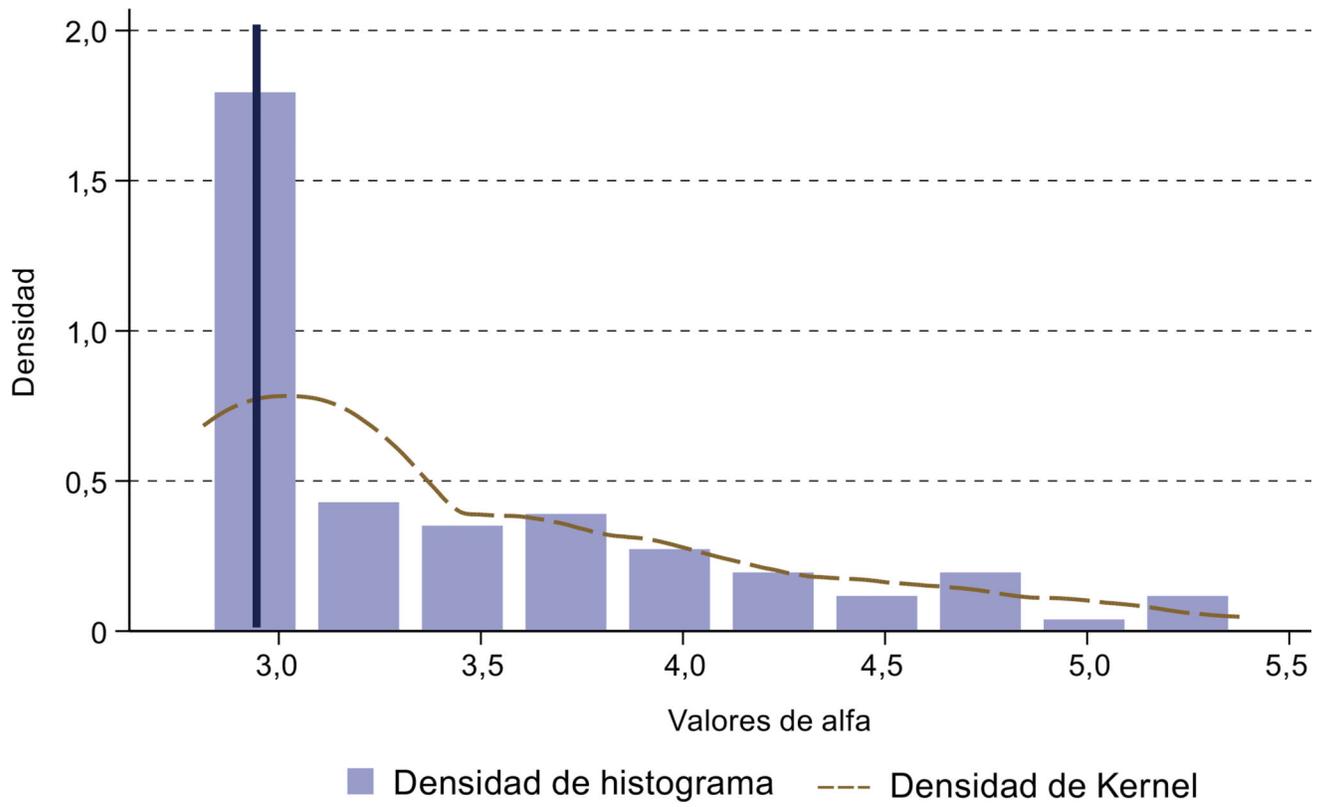


Fuente: Elaboración propia.

Gráfico 2.
Cuba: estimadores de parámetros para el modelo logístico

A. Alfa

Función Box-Cox del parámetro alfa

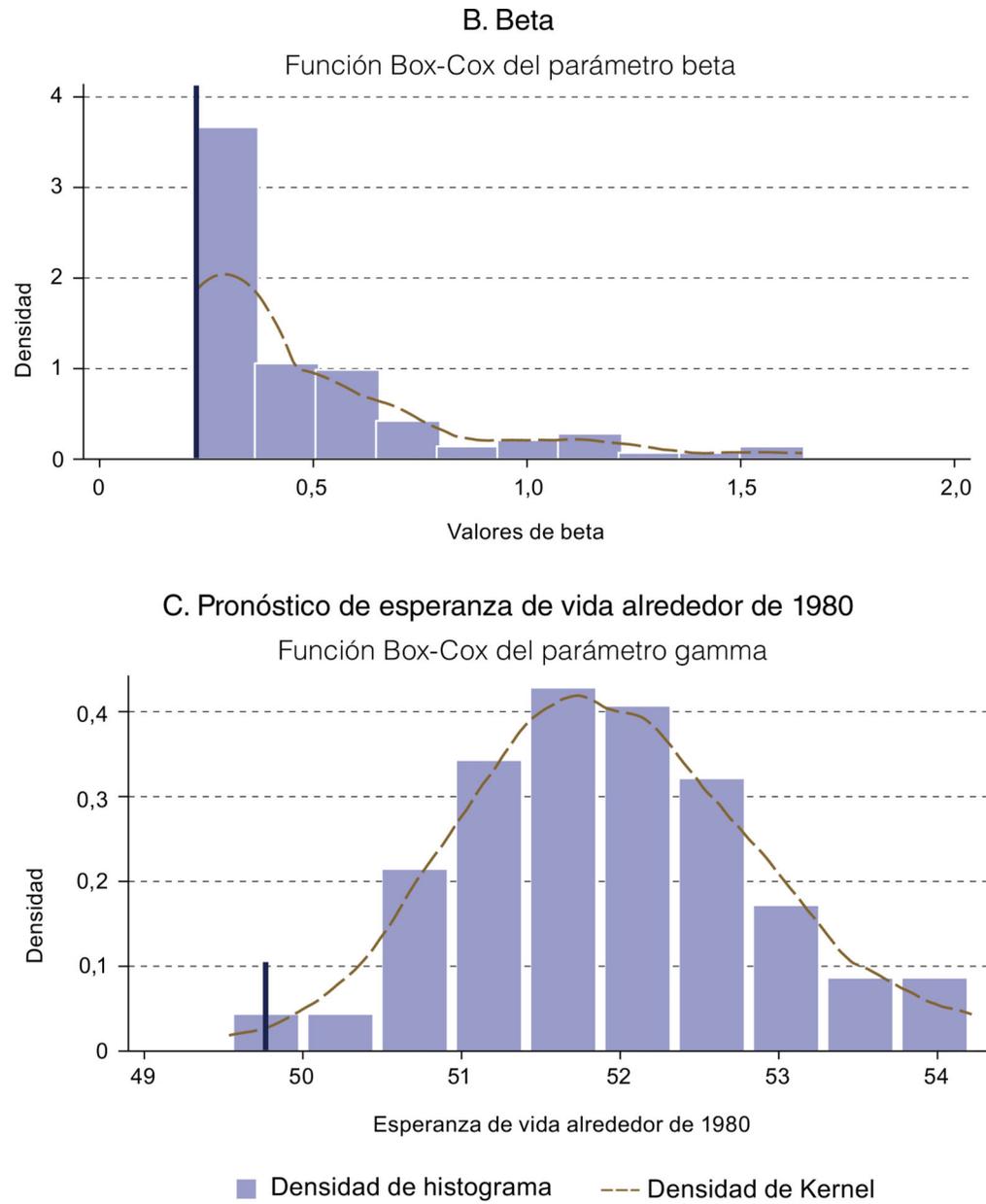


Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



Fuente: Elaboración propia.

Gráfico 3.
Guatemala: estimadores de parámetros para el modelo Box-Cox