

Comparación y evaluación de la distribución del estimador de la tasa global de fecundidad de Bolivia

Milena Linneth ARGOTE-CUSI

Corpotalentos (ONG)

Resumen

Se aplicó el método de remuestreo para generar la distribución estadística de la TGF de Bolivia en 1998 y 2003 con el objetivo de evaluar diferencias del estimador en ambos momentos. Se consideraron las muestras de las ENDSAS para 1998 y 2003 de Bolivia que tienen diseño estratificado y bietápico. Las distribuciones resultan positivas para las pruebas de normalidad; sin embargo, tienen sus propias características de asimetría y curtosis. El método permite generar la distribución estadística empírica para el estimador, lo cual no es factible en la realidad. También permite evaluar la precisión del estimador toda vez que el algoritmo de replicación puede ser modificado y aplicado a otras poblaciones u otros indicadores.

Palabras clave: distribución estadística, técnica de remuestreo, tasa global de fecundidad, Bolivia.

Abstract

Comparison and evaluation of the distribution of the estimator of the total fertility rate of Bolivia

The resampling method was applied to generate the statistical distribution of TFR of Bolivia in 1998 and 2003 so as to evaluate differences in the estimator at both times. Samples of ENDSAS for 1998 and 2003 of Bolivia, which have a stratified and two-stage design, were considered. The distributions turnout to be positive for the normality tests; nonetheless, they have their own characteristics of asymmetry and kurtosis. The method allows generating the empirical statistical distribution for the estimator, which is not feasible in the reality. It also allows evaluating the precision of the estimator as the replication algorithm can be modified and applied to other populations or other indicators.

Key words: statistical distribution, resampling technique, total fertility rate, Bolivia.

Introducción

En la investigación social nos aproximamos a la realidad de las poblaciones mediante indicadores. Cuando un indicador se calcula a partir de la población total, recibe el nombre de parámetro, mientras que si se obtiene de una muestra, se denomina estadística (Lohr, 2000). Generalmente, los parámetros de la población total no son conocidos

debido al costo que implica obtenerlos; en su lugar se dispone de muestras a partir de las cuales hacemos estimaciones de los parámetros. Según la teoría de muestreo, como sólo es posible obtener una muestra de todas las posibles de la población total, las estimaciones que hagamos a partir de ella están sujetas a errores muestrales y no muestrales. En la práctica no es factible obtener todas las posibles muestras de una población. Para solucionar este problema, la estadística paramétrica ha construido una base teórica fundamental que incluye la ley de los números grandes (LNG) y el teorema del límite central (TLC). Éstos nos permiten hacer inferencias para la población total. Un supuesto importante de esta teoría es la normalidad de la distribución estadística del fenómeno en estudio.

Si tenemos la posibilidad de contar con dos muestras (a, b) en dos momentos ($t, t + 5$) de la misma población, resulta interesante aplicar el método de remuestreo y evaluar si la diferencia del estimador en el tiempo es significativa o no, ya que de no ser así subyace la duda de que los cambios se hayan debido a variaciones muestrales, lo cual afectaría de manera no adecuada nuestras decisiones para cambiar esa realidad. Partiendo del trabajo de Argote (2007), el objetivo del presente estudio es verificar la normalidad de la distribución del la TGF en 2003 y posteriormente verificar si la diferencia entre 1998 y 2003 es significativa o no.

Esta investigación se enmarca en la búsqueda de una adecuada y rigurosa interpretación de estimadores a partir de muestras. Si las estimaciones que se realizan a partir de muestras de la población total no se interpretan de forma adecuada, desde el punto de vista estadístico, se pueden hacer afirmaciones que no son válidas para la población total. Por ello es crucial una evaluación de los datos, así como el pleno conocimiento del método que los generó (diseño de la muestra) antes de seleccionar los métodos estadísticos adecuados a utilizar para su análisis.

Acorde con la exposición anterior, nuestro marco teórico se compone de tres pilares fundamentales: la importancia de la replicación en la estadística, las características de la distribución normal y métodos paramétricos para la comparación de dos muestras.

La replicación en la experimentación científica

Según Bayarri y Martínez (1997), las réplicas en la experimentación científica son innatas a la estadística; sin embargo, pocos estudios se han encontrado sobre su modelización sistemática y sobre el hecho de sacar el máximo provecho a la información que puede proporcionar la replicación.

En este sentido, la presente investigación se constituye en un ejemplo de la utilidad de la replicación a partir del remuestreo y del análisis de los datos generados en apoyo a la evaluación de la diferencia de medias de dos muestras.

En este caso, la replicación a través del remuestreo (*bootstrapping*) tiene como primer objetivo encontrar la distribución estadística del estimador de la TGF para Bolivia en 2003 y, con los resultados de 1998 (Argote-Cusi, 2007), comparar ambas distribuciones y contrastar las diferencias. El objetivo final es analizar las siguientes cuestiones: la generalización de las conclusiones a otras poblaciones de origen, el estudio de la variabilidad del estimador en función de la evolución de la población (en este caso se consideran dos puntos en el tiempo) y, por otro lado, el estudio del sesgo del estimador ocasionado por las ausencia o deficiencias del diseño.

Respecto a la primera cuestión, de acuerdo con el trabajo de Argote (2007) y los resultados presentados en el presente artículo, podemos afirmar que se espera con alta probabilidad que el estimador de la TGF de otras poblaciones diferentes a la boliviana tengan un comportamiento semejante a la normal, pero con sus propias diferencias, no tan evidentes de percibir en muestras grandes, pero que al final existen y reflejan la heterogeneidad de comportamientos. Este hecho era de esperarse, toda vez que los datos reflejan un comportamiento no estacionario de las poblaciones (Bayarri y Martínez, 1997), pues los datos de la fecundidad en una población presentan correlación y ésta es una característica necesaria de considerar al analizar las estimaciones. Además, hay que recordar que la TGF considera un supuesto importante: que las tasas específicas de fecundidad (TEF) se mantienen constantes en el tiempo, cosa que no es real. Estos factores deben ser considerados en cualquier estudio que analice los datos de la fecundidad de las poblaciones con base en muestras, para poder formular a partir de ellos un modelo predictivo.

Para analizar la variabilidad del estimador se dispone de dos muestras en el tiempo 1998 y 2003, y aplicando los test necesarios (*t*-Student, *t*-Welch) se observó que ambas tienen varianza distinta, lo cual es una consideración importante a la hora de construir un modelo predictivo (Núñez-Anton y Zimmerman, 2001) que considere la no estacionalidad de las varianzas.

Ya que la experimentación o replicación del estudio no se realiza en condiciones iniciales iguales (si bien las muestras se toman a la misma población de mujeres y tienen el mismo diseño, se consideraron más unidades primarias de muestreo en 2003 que en 1998, así como también es probable que se hayan entrevistado a las mismas mujeres en algunos casos,

ya que las mujeres en 1998, en el grupo 15 a 19 años, después de cinco años pueden ser elegibles para la muestra de 2003), por lo cual se trataría de una muestra combinada transversal y longitudinal. Estos factores, entre otros, se reflejan en el sesgo del estimador y en las características propias de cada distribución estadística por muestreo.

Cabe destacar que la replicación de experimentos permite fácilmente la aplicación de la modelización bayesiana. Acorde con Bayarri y Martínez (1997), la flexibilidad del modelo jerárquico bayesiano permite una fácil modelización de las relaciones que surgen en la replicación. La discrepancia entre el parámetro poblacional y el estimador presentan relaciones en el ámbito temporal, espacial o poblacional. Esta última implicación de la presente investigación merecería ser estudiada más detalladamente como parte de la elaboración de un modelo predictivo.

La distribución normal

La distribución normal es una de las más utilizadas en estudios que modelan la dinámica social de las poblaciones y está relacionada con los fenómenos naturales.¹ Sobre todo cuando se cuenta con gran cantidad de datos se asume una distribución normal (basados en la LNG y el TLC) de los datos en estudio, aunque sería importante verificar si es correcto asumir directamente este supuesto, sobre todo cuando se trata de muestras pequeñas. El presente trabajo no asume de antemano la distribución normal de la TGF, sino que la encuentra y la evalúa mediante pruebas estadísticas. Tenemos como antecedente que Argote (2007) verificó que la distribución estadística de la TGF de Bolivia para 1998 tenía un comportamiento normal aplicando el test Kolmogorov Smirnov (sig. 0.790) con media 4.3, error estándar de 0.102 e intervalo de confianza a 95 por ciento [4.172, 4.499]. Cinco años después (2003) la TGF de la población del mismo país, de la cual se cuenta con una muestra obtenida con el mismo diseño muestral que en 1998, ¿también tendrá una distribución normal?

¹ La importancia de la distribución normal se debe principalmente a que hay muchas variables asociadas a fenómenos naturales que siguen el modelo de la normal: caracteres morfológicos de individuos; caracteres fisiológicos, como el efecto de un fármaco; caracteres sociológicos, como el consumo de cierto producto por un mismo grupo de individuos; caracteres psicológicos, como el cociente intelectual; el nivel de ruido en telecomunicaciones; errores cometidos al medir ciertas magnitudes, y valores estadísticos muestrales como la media.

Para nuestro caso en particular, los parámetros que definen la distribución normal son su media y su desviación estándar (μ, σ) cuya función de densidad viene dada por la ecuación 1:

$$(1) \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad -\infty < x < +\infty$$

Si logramos experimentar con las variaciones de la media y la desviación estándar (μ, σ) de la distribución normal, nos daremos cuenta de que en realidad podemos reconstruir una familia de distribuciones que varían en la centralidad y la forma de la también llamada campana de Gauss. Sin embargo, mediante una transformación, todas se pueden representar por la distribución normal estándar con media 0 y desviación 1, para la cual existen tablas con base en la distribución de densidad de probabilidad que nos permiten encontrar la probabilidad de que la variable aleatoria que sigue este comportamiento se encuentre entre dos valores determinados, la otra opción sería integrar bajo la función de densidad (ecuación 1).

Dentro de las características más importantes de la distribución normal o campana de Gauss se incluye que la media, la mediana y la moda tienen el mismo valor, reflejando la simetría de la distribución de densidad. La forma de la campana está determinada por las variaciones de la media y de la desviación estándar; a diferentes valores de μ , la campana se desplaza a lo largo del eje horizontal. Por otra parte, cuanto mayor sea el valor de σ , más se dispersarán los datos en torno a la media y la curva será más aplana (Díaz y Fernández, 2001).

Es recomendable no solamente hacer una evaluación subjetiva de una distribución estadística para determinar si ésta tiene una distribución normal o no. El test Kolmogorov-Smirnov (K-S test) es el más utilizado para contrastar la hipótesis de normalidad de una distribución. Según la teoría, para muestras de más de 30 casos se asume el comportamiento normal, pero aunque dispongamos de muestras muy grandes, seguramente no llegaremos a contar con una distribución normal perfecta, por lo cual una valoración de las características de la distribución no está de más y nos ayudaría a comprender mejor el comportamiento del fenómeno demográfico en estudio, que en este caso es la fecundidad.

La asimetría se puede valorar con la diferencia entre la media, moda y la mediana, pero contamos también con el coeficiente de asimetría de Fisher (véase ecuación 2).

$$(2) \quad \gamma_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}} \quad \begin{array}{ll} \gamma_1 = 0, \text{ simétrica} \\ \gamma_1 < 0, \text{ asimétrica — a — izquierda} \\ \gamma_1 > 0, \text{ asimétrica — a — derecha} \end{array}$$

Si el coeficiente de Fisher es igual a cero, la campana es totalmente simétrica, mas si es menor a cero es asimétrica a la izquierda, o mayor a cero, asimétrica a la derecha.

Una segunda pregunta que podemos tener es si la campana es más o menos achatada respecto al modelo normal, para lo cual se cuenta con el coeficiente de aplastamiento o curtosis de Fisher:

$$(3) \quad \gamma_2 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} - 3 \quad \begin{array}{ll} \gamma_2 = 0, \text{ mesocúrtica} \\ \gamma_2 < 0, \text{ platicúrtica} \\ \gamma_2 > 0, \text{ leptocúrtica} \end{array}$$

Donde: mesocúrtica indica que la curva es tan achatada como la normal; platicúrtica, más achatada que la normal, y leptocúrtica, más apuntada que la normal.

Ya que se asume (H_0) la normalidad de la distribución de la TGF, se hizo uso de la prueba K-S para verificar la hipótesis nula. Pero, además, en esta investigación se evaluaron las condiciones que permitan aplicar el test *t*-Student para la comparación de medias de la TGF, toda vez que uno de los supuestos de esta prueba es que ambas muestras presentan un comportamiento normal.

Métodos paramétricos para la comparación de dos muestras

Un buen hábito antes de realizar el ejercicio de comparación de dos medias es conocer claramente las características de las muestras como el tipo de diseño utilizado para la obtención de los datos, el tamaño de la muestras, si se trata de muestras dependientes o pareadas y el tipo de variable, en virtud de que de ello dependen los métodos estadísticos a aplicar de forma adecuada (Díaz y Fernández, 2001).

Existen métodos paramétricos y no paramétricos para la comparación de muestras dependiendo de si el comportamiento de las distribuciones cumple con los supuestos o no. Para los fines del presente trabajo profundizaremos en el primero de los métodos, ya que las distribuciones a comparar tienen un comportamiento normal bajo el test K-S en 1998 (sig. 0.790) y 2003 (sig. 0.686).

En la literatura se encuentra que el tipo de prueba a aplicar para contrastar la hipótesis de investigación depende de las características de los datos que son nuestra evidencia. Para comparar dos muestras, éstas pueden ser independientes o dependientes entre sí. Son independientes cuando se toma a diferentes individuos la misma variable de estudio, mientras que las muestras dependientes se caracterizan porque contienen variables que se toman a los mismos individuos más de una vez. Este segundo diseño es propio de los estudios de casos control y de los estudios longitudinales.

Para las muestras independientes es necesario que se cumpla que ambas distribuciones tengan un comportamiento normal e igual varianza, de tal modo que sea viable aplicar el test *t*-Student para la comparación de dos medias (Lohr, 2000). El valor de contrastación de la prueba viene dado por la ecuación 4.

$$(4) \quad t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{(n-1)\hat{S}_1^2 + (m-1)\hat{S}_2^2}{n+m-2} \sqrt{\frac{1}{n} + \frac{1}{m}}}}$$

Donde: \bar{X} e \bar{Y} son el valor medio en cada uno de los grupos y \hat{S}_1^2, \hat{S}_2^2 las cuasi varianzas muestrales. Otra manera de contrastar la hipótesis nula es por medio de los intervalos de confianza, que vienen dados por:

$$(5) \quad (\bar{X} - \bar{Y}) \pm t_{0.975}^{n+m-2} \sqrt{\frac{(n-1)\hat{S}_1^2 + (m-1)\hat{S}_2^2}{n+m-2}} \sqrt{\frac{1}{n} + \frac{1}{m}}$$

donde $t_{0.975}^{n+m-2}$

denota el valor que, según la distribución t -Student, con $n + m - 2$ grados de libertad deja a su derecha el 2.5 por ciento de los datos.

Puede ocurrir que las poblaciones tengan diferente variabilidad, en cuyo caso es recomendable aplicar la prueba F -Snedecor (Lohr, 2000) para contrastar qué tan significativa es la diferencia entre ellas:

$$(6) \quad F = \frac{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}{\frac{1}{m-1} \sum_{i=1}^m (Y_i - \bar{Y})^2} = \frac{\hat{S}_1^2}{\hat{S}_2^2}$$

Donde se espera que la razón de las varianzas siga una distribución F -Snedecor con parámetros $(n - 1)$ y $(m - 1)$. Si la prueba es significativa, entonces es conveniente aplicar una corrección al test t -Student, llamada test t -Welch:

$$(7) \quad t = \frac{(\bar{X} - \bar{Y})}{\sqrt{\frac{\hat{S}_1^2}{n} + \frac{\hat{S}_2^2}{m}}}$$

Donde el número de grados de libertad dependerá de las varianzas según la siguiente fórmula:

$$(8) \quad f = \frac{\left(\frac{\hat{S}_1^2}{n} + \frac{\hat{S}_2^2}{m} \right)}{\frac{1}{n+1} \left(\frac{\hat{S}_1^2}{n} \right)^2 + \frac{1}{m+1} \left(\frac{\hat{S}_2^2}{m} \right)^2} - 2$$

En este caso, el intervalo de confianza estaría dado por:

$$(9) \quad (\bar{X} - \bar{Y}) \pm t_{0.975}^f \sqrt{\frac{\hat{S}_1^2}{n} + \frac{\hat{S}_2^2}{m}}$$

Los conceptos básicos sobre la distribución normal y los métodos paramétricos para la comparación de medias que aquí se mencionaron se utilizaron para el análisis y el testeo de la hipótesis en cuestión. Como se mencionó anteriormente, se tienen métodos no paramétricos para la comparación de medias como el test de Wilconson, por lo que invito a revisar esa literatura a los investigadores que trabajen con muestras menores a 30 casos.

Datos y métodos

Se utilizaron los datos de la Encuesta Nacional de Demografía y Salud de 2003 ($m = 17\,654$). La de 1998 tenía un tamaño muestral igual a 11 187, con un diseño estratificado y bietápico a nivel nacional. Para efectos de este trabajo, las unidades de análisis son las mujeres en edad fértil y los nacimientos de sus hijos localizados en los hogares seleccionados (ENDSA, 1998 y 2003).

Para la estimación de la TGF se utilizó la definición teórica clásica, cuyo numerador es el número de nacimientos correspondientes a las mujeres de entre 15 y 49 años en los tres últimos años anteriores al momento de la entrevista, y el denominador es la sumatoria de tiempos de exposición, de las madres antes de experimentar el nacimiento de x hijo. El proceso de clasificación del tiempo de exposición en el denominador de cada grupo quinquenal cuidó la correcta aportación de tiempo de exposición considerando la edad exacta de la mujer en el momento de experimentar el nacimiento de un hijo, ya que podría ocurrir el caso que una mujer aporte un tiempo de exposición a dos grupos quinquenales adyacentes.

Se aplicó el remuestreo con reemplazo considerando las unidades primarias de muestreo (UPM) (en 1998 se tomaron 823 UPM y en 2003 se contó con 1000 UPM) y luego, dentro de éstas, las unidades secundarias de muestreo (USM u hogares). La varianza, el error estándar y el sesgo se calculan automáticamente a partir de la distribución por muestreo de la TGF, como indica la técnica del remuestreo. Una buena estimación del intervalo de confianza se obtiene en 1 000 réplicas (Efron y Tibshirani,

1993), por lo cual los resultados y las conclusiones se presentan tomando en cuenta esta referencia.²

Se utilizó el test Kolmogorov-Smirnov para probar si la distribución estadística de la TGF en 2003 estimada por métodos no paramétricos tenía un comportamiento normal. También se calcularon el coeficiente de asimetría y de curtosis de cada distribución en los dos momentos. En virtud de que las varianzas eran diferentes, se utilizó la prueba *F*-Snedecor para evaluar su significancia estadística, por lo tanto, no se pudo aplicar de forma clásica el test *t*-Student para la comparación de medias, sino una modificación al mismo, llamado test de Welch, que nos permitió evaluar la hipótesis nula (H_0 : la diferencia de medias de la TGF en 1998 y 2003 es cero).

Resultados

Un primer análisis exploratorio de las muestras originales nos reveló que se trata de muestras aleatorias con un diseño estratificado y bietápico, en dos momentos en el tiempo a diferentes individuos, por lo cual se trata de muestras independientes no pareadas. Por su parte, la tasa global de fecundidad, calculada según la definición teórica de una tasa, constituye una variable continua.

Aplicando el método de remuestreo con reemplazo a la muestra original de la ENDSA 2003 se obtuvo la distribución estadística de la TGF en 1000 réplicas. Según la prueba Kolmogorov-Smirnov, es muy probable que la hipótesis nula sea verdadera, por lo cual no hay evidencia suficiente para rechazar la normalidad de su distribución. Junto a los hallazgos de Argote (2007) acerca de la distribución estadística de 1998 podemos decir que ambas tienen comportamiento semejante a la normal (sig 0.790 *versus* 0.686) según el test K-S.

Respecto a la simetría de las distribuciones, podemos observar en el cuadro 1 que en ambos casos, si redondeamos la media, mediana y moda a un decimal, tenemos una TGF de 4.3 para 1998 y 3.8 para 2003. Las variaciones en ambas reflejarían leves inclinaciones a la derecha o a la izquierda de la normal, que no son significativas según los datos. Por tanto, en el eje x podemos imaginar la distribución de 2003 más cerca del cero que la de 1998 (véase gráfica 1). La desviación estándar en 2003 es 12 por ciento menor que en 1998, lo cual significaría que tenemos una distribución

² El detalle de las características y los supuestos bajo los cuales se aplicó el método de remuestreo se encuentran en el artículo de Argote-Cusi, 2007.

más apuntada en 2003 que 1998; esto refleja que en 2003 los valores del estimador están más cerca de la media; sin embargo, esta curva tiene más irregularidades que la de 1998.

CUADRO 1
CARACTERÍSTICAS DE LA DISTRIBUCIÓN ESTADÍSTICA
DE LA TGF TOTAL Y POR LUGAR DE RESIDENCIA DE BOLIVIA
EN 1998 Y 2003

Estadístico	1998	2003
N	1000	1000
Media	4.326	3.835
Mediana	4.327	3.839
Moda*	4.333	3.810
Desviación estándar	0.102	0.090
Varianza	0.010	0.008
Pr (skewness)	0.216	0.506
Pr (kurtosis)	0.616	0.933
Kolmogorov-Smirnov test	0.651	0.790
Sig. 2-tailed	0.790	0.686

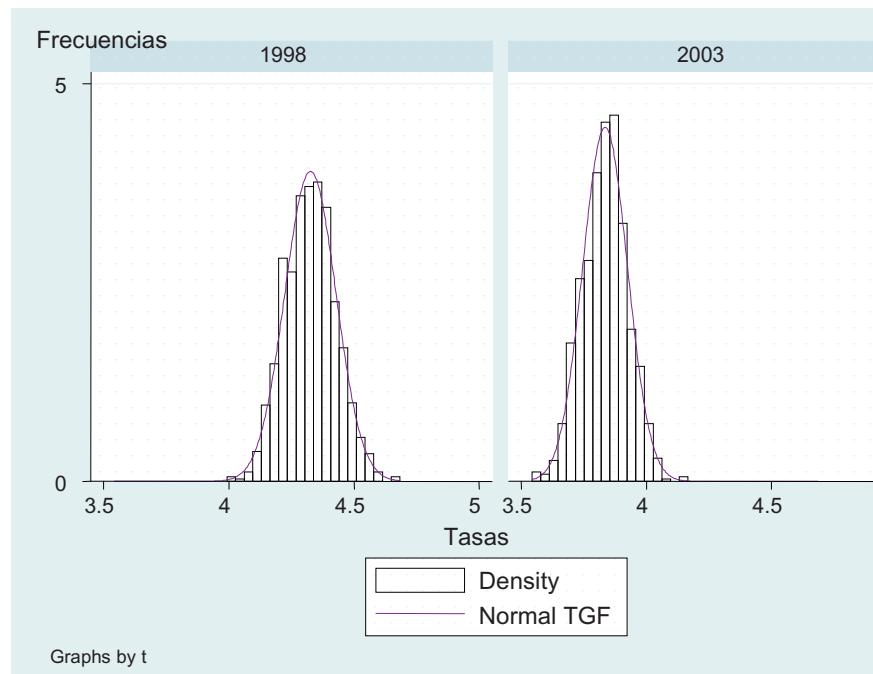
* Existen múltiples modas, se muestra el menor valor.

Fuente: cálculos propios con base en la distribución estadística de la TGF de 1998 y 2003.

Si el coeficiente de curtosis y el de asimetría son iguales a cero, significa que tenemos una distribución normal perfecta. En el cuadro 1 podemos apreciar que la distribución de 2003 es aproximadamente el doble de asimétrica a la derecha y 30 por ciento más alta que 1998 (platicúrtica), en comparación con la normal estándar. A pesar de estas variaciones de las características de la muestra, la bondad de la distribución normal nos permite contar con buenas estimaciones de la TGF a través de la media de la distribución; así mismo, nos da la posibilidad de encontrar el sesgo y los intervalos de confianza del estimador de forma automática (cuadro 2).

Un hallazgo que llama la atención es que, para la réplica 1000, el sesgo de 1998 es mayor a 2003, es decir, que considerando la muestra original para 1998 se espera tener un estimador más alejado del parámetro poblacional que en 2003. Y de hecho, comparando la estimación oficial del INE de Bolivia, halló una TGF de 4.2 para 1998, mientras que encontramos una estimación de 4.3 por este método. En cambio, las estimaciones son muy semejantes para 2003 (3.81 *versus* 3.83). Esto deja en evidencia la importancia de contar con un valor del sesgo, que nos habla de la precisión de nuestras estimaciones, el cual se estima automáticamente a partir de la distribución por remuestreo.

GRÁFICA 1
DISTRIBUCIÓN ESTADÍSTICA DE LA TASA GLOBAL DE FECUNDIDAD
TOTAL DE BOLIVIA EN 1998 Y 2003



Fuente: construcción propia con base en la distribución estadística de la TGF de 1998 y 2003.

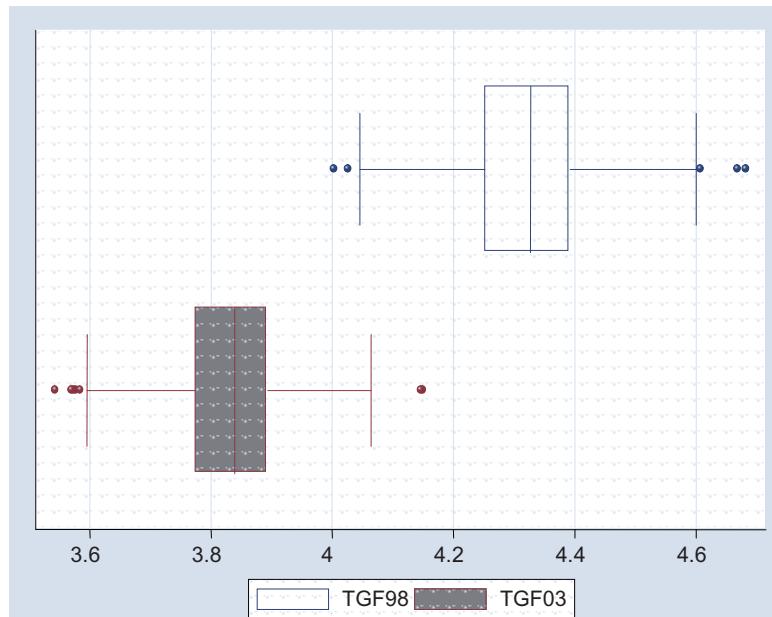
CUADRO 2
MEDIA, VARIANZA E INTERVALOS DE CONFIANZA DEL
ESTIMADOR DE LA TASA GLOBAL DE FECUNDIDAD PARA 1998 Y
2003 EN LA RÉPLICA 1000, SEGÚN EL DISEÑO COMPLEJO DE LA
MUESTRA

Año	Es	Media	Sesgo	Sesgo/es	CV	Remuestreo en dos etapas del diseño de la muestra	
						Ic 95%	Superior
1998	0.102	4.327	0.097	0.952	2.364	4.172	4.499
2003	0.090	3.835	-0.003	-0.033	2.336	3.693	3.984

Fuente: cálculos propios con base en la distribución estadística de la TGF de 1998 y 2003.

El diagrama de caja nos muestra claramente una menor variabilidad de la TGF en 2003 (ancho de los rangos intercuartilicos). Cinco años después, la TGF se ubica más cerca del cero en el eje X, es decir, la TGF disminuye en 11 por ciento, con base en lo cual podemos decir que tiene decrementos anuales de 2.2 por ciento, aproximadamente.

GRÁFICA 2
DIAGRAMA DE CAJA DE LA DISTRIBUCIÓN DE LA TASA GLOBAL
DE FECUNDIDAD TOTAL DE BOLIVIA EN 1998 Y 2003



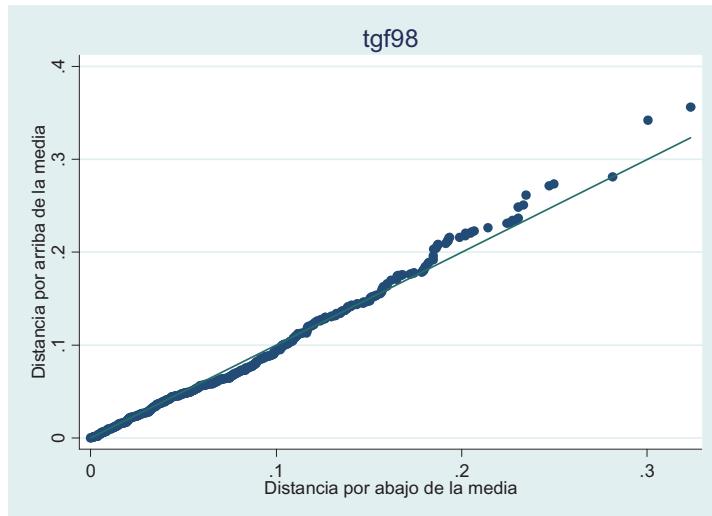
Fuente: construcción propia con base en la distribución estadística de la TGF de 1998 y 2003.

Se observa que el estimador de la TGF es sesgado, ya que la diferencia entre la media y los valores del estimador es diferente de cero. Para 1998, el valor medio del sesgo es de -0.000000151, y para 2003, 0.0000000412, o sea, es muy pequeño y tiende a cero, lo cual refleja la coherencia de las estimaciones aquí realizadas. La negatividad del sesgo se interpretaría como que el estimador es mayor a la media, y ser positivo, lo contrario. Es decir, que en 1998 hay más valores por encima de la media (véase A del grafico 3). Para 2003, un sesgo positivo indica que la media es mayor al estimador en general y eso también se refleja en la curva B del mismo gráfico. Si logramos integrar el área bajo la curva respecto a la media podemos observar que existe mayor sesgo en la curva de 2003 que en 1998, pero éste tiende a cero.

El grafico 4 nos permite apreciar cómo se dispersan los 1000 valores del estimador de la TGF en cada réplica (A) y en cada valor del estimador (B), que demuestran que el método de remuestreo es una técnica robusta para la generación de datos empíricos que no se podrían hacer en la realidad.

GRÁFICA 3
CURVA DE SIMETRÍA DE LA DISTRIBUCIÓN ESTADÍSTICA
DE LA TGF DE BOLIVIA EN 1998 Y 2003

A

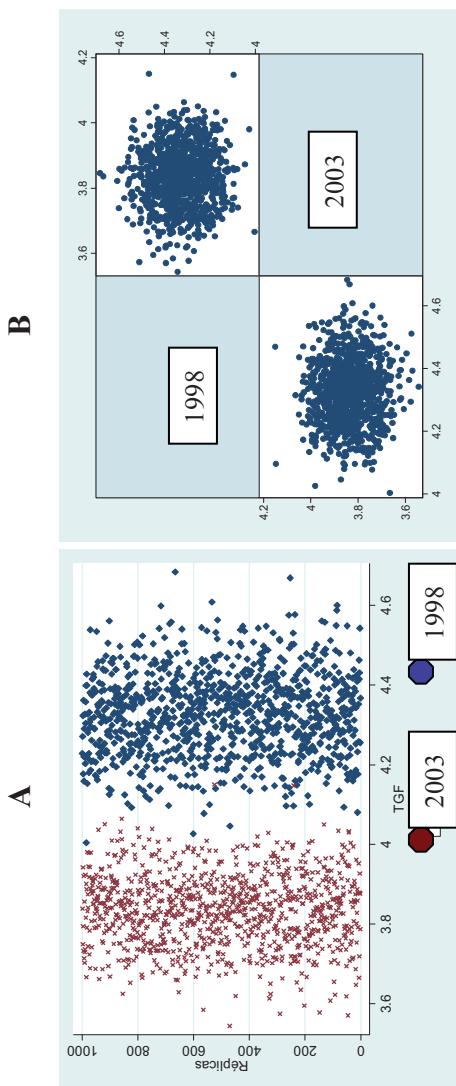


B



Fuente: construcción propia con base en la distribución estadística de la TGF de 1998 y 2003.

GRÁFICA 4
DIAGRAMAS DE DISPERSIÓN DE LOS VALORES QUE TOMA EL ESTIMADOR EN 1000
RÉPLICAS DE LA MUESTRA ORIGINAL, TGF DE BOLIVIA EN 1998 Y 2003



Fuente: construcción propia con base en la distribución estadística de la TGF de 1998 y 2003.

CUADRO 3
PRUEBA PARA LAS VARIANZAS: *F*-SNEDECOR

Ho: $sd(tgf98) = sd(TGF03)$
Grados de libertad = 999 999
Cola inferior = $F_L = 1/F_{obs.} = 0.767$
Cola superior = $F_U = F_{obs.} = 1.304$
Ha: $sd(1) < sd(2)$ $P > F_{obs.} = 1.0000$
Ha: $sd(1) = sd(2)$ $P < F_L + P > F_U = 0.0000$
Ha: $sd(1) > sd(2)$ $P > F_{obs.} = 0.0000$

CUADRO 4
PRUEBA PARA LAS MEDIAS: TEST DE WELCH

Ho: $mean(x) - mean(y) = diff = 0$
Grados de libertad = 1965.81
diff 0.491585
t 114.33
Ha: $diff < 0$ $P < t = 1.0000$
Ha: $diff = 0$ $P > t = 0.0000$
Ha: $diff > 0$ $P > t = 0.0000$

La varianza en 2003 es menor en 20 por ciento de la varianza en 1998, este hecho también se puede verificar en el diagrama de cajas (gráfico 2), donde la distancia entre extremos es menor en 2003, es decir, los datos en 2003 están más concentrados alrededor de la media. Según la prueba *F*-Snedecor (véase cuadro 3), se rechaza la hipótesis de igual variabilidad entre ambas muestras (sig 0.0000). En virtud de lo anterior, el presente estudio constituye una comparación de muestras independientes con distintas varianzas, donde el test de Welch es apropiado.

Al evaluar la hipótesis de nula: si las TGF de fecundidad en ambos momentos son iguales, el test de Welch nos indica que es poco probable que esta hipótesis sea verdadera, es decir, son diferentes; además, en 1998 es mayor a 2003. Por lo tanto, la diferencia de medias no se debe a variaciones muestrales, si no a las características propias de la población boliviana.

Discusión

El aporte del presente trabajo radica en la utilización de métodos computacionales que permiten hacer el remuestreo para encontrar la

distribución estadística de un indicador demográfico y a partir de ella hacer inferencias para la población total desconocida en el marco de la estadística no paramétrica. Recordemos que como esta tarea no se puede realizar en la práctica, se adopta el teorema del límite central y la ley de los grandes números, que asume una distribución normal del estimador. Ahora bien, las distribuciones de 1998 y 2003 de la TGF tienen comportamiento normal, pero con sus características propias (1998, *skewness*: 0.216 y *curtosis*: 0.616; 2003, *skewness*: 0.506 y *curtosis*: 0.933), ya que la distribución encontrada pertenece al caso particular del estimador de la TGF para la población de mujeres en edad fértil de Bolivia en 1998 y 2003. En este sentido, si se aplica este método para otras poblaciones, se esperaría que la distribución por muestreo del indicador también se asemejara a la normal, según los teoremas de la estadística inferencial, pero serían diferentes entre ellas debido a las características intrínsecas de la muestra de determinada población. A la luz de la teoría general de sistemas, este hecho es de esperarse, pues los sistemas complejos no lineales estarían representados por un conjunto de funciones que pertenecen a una misma familia, por lo cual el problema siguiente a investigar entonces no sería cuál es la distribución estadística de la TGF para X país, población o muestra, sino, ¿cuál es el impacto en utilizar una u otra distribución (paramétrica o no paramétrica) de determinada población, en el tema de proyecciones de población o de modelos de predicción?

Mas al haber un ejercicio estadístico de la comparación de dos medias a partir de dos muestras, el fin del mismo en el ámbito demográfico se resalta. Generalmente, cuando se realiza la evaluación del estado de la población, los tomadores de decisión ponen énfasis en los números absolutos y comparan las diferencias entre ellos de forma automática, pero pocas veces se realiza un análisis riguroso de las diferencias bajo determinada fuente de información, sobre todo para controlar el hecho de no hacer afirmaciones erróneas que se deben a las características propias de los datos o a variaciones muestrales. A partir de la estadística y de métodos computacionales, aquí se llevó a cabo un experimento científico que nos permitió verificar de forma rigurosa nuestra hipótesis. Cabe resaltar en este sentido que se ha construido un algoritmo que se puede aplicar a otras muestras y responder a otro tipo de preguntas de investigación. También se puede utilizar el mismo algoritmo para otros indicadores demográficos para los que nos interese respaldar de forma científica los cambios del sistema social en estudio. Así mismo, dicho algoritmo puede formar parte de otro algoritmo mayor que nos permita hacer un análisis comparativo para

diversos años de los cambios en las tendencias de determinado indicador, controlando el sesgo de los errores muestrales.

Otro aporte del presente trabajo es su utilidad para el análisis de la precisión del estimador. Al tratar con sistemas complejos ellos se dirigen por comportamientos no lineales representados por conjuntos de números determinados. En Demografía cabe preguntarnos: ¿Es realmente significativa la diferencia de una TGF de x en t_1 y z en t_2 siendo que $(z-x)$ tiende a 0? ¿Qué significa esta diferencia? ¿Podemos redondear?

Sin duda, el impacto de la variación del número se podría apreciar más claramente si contáramos con el numerador y denominador de esa tasa, en los cuales se contabilizan números enteros y en ese caso podremos ver si una diferencia de más o menos 100 nacimientos, por ejemplo, entre uno y otro numerador es una diferencia significativa, y lo mismo para el denominador. Al ser el método de remuestreo iterativo y controlable a través del algoritmo es posible hacer un seguimiento en retrospectiva para evaluar este tipo de impactos.

¿Es factible hablar de la precisión del estimador de la TGF?

El concepto de estimación está ligado a la no exactitud, por lo tanto, sí es factible hablar de la precisión de los estimadores demográficos (Dysert, 2006). Precisión es el grado en el que una estimación se aleja del valor real o parámetro poblacional y está ligado a las características de la fuente de datos, su procesamiento y el método de estimación, entre otros. De esta manera, la precisión de un estimador está dada por un rango de datos que tienen cierta probabilidad de ocurrencia, representada por una distribución probabilística. Idealmente, la distribución probabilística más conocida, asociada a varios estimadores, es la distribución normal; sin embargo, en los sistemas sociales es factible tener otras distribuciones o tener características asimétricas que modifican la probabilidad real de subestimación (50 por ciento) o sobreestimación (50 por ciento).

En este sentido, las características propias de las distribuciones por muestreo de la TGF obtenidas para 1998 y 2003 no hacen más que precisar las probabilidades de ocurrencia del estimador. Si el coeficiente curtosis y el de asimetría son iguales a cero, significa que tenemos una distribución normal perfecta, con 50 por ciento de probabilidad de sobreestimar el valor, así como de subestimarlo. En esta lógica, ambas distribuciones son asimétricas a la derecha: 1998, 22 por ciento, y 2003, 51 por ciento. La distribución de 2003 es aproximadamente el doble de asimétrica a la derecha y 30 por

ciento más concentrada que la de 1998 (plasticúrtica), en comparación con la normal estándar. Lo anterior implica que la probabilidad de sobreestimar el valor medio de la TGF en 2003 es dos veces mayor que en 1998.

También se observa que el intervalo de confianza es 11 por ciento menor en 2003 que en 1998, lo cual disminuye nuestro rango de incertidumbre, que a su vez puede deberse a la mayor cantidad de unidades primarias de muestreo consideradas, es decir, a una mayor información de la población.

Los conceptos mencionados brindan una nueva forma de ver las estimaciones o resaltar la incertidumbre que presentan las mismas. No se puede considerar a las estimaciones como simples números exactos, cuando están asociados a una probabilidad de ocurrencia, como lo demuestra esta investigación sobre el estimador de la tasa global de fecundidad.

Para mostrar la compleja representación de los estimadores tenemos la siguiente definición según Dysert, 2006: “Una estimación es el valor esperado de una compleja ecuación de elementos probabilísticos sujetos a variación aleatoria dentro rangos definidos”. En este contexto, los científicos sociales no podemos más que estar conscientes de la variabilidad, la precisión e incertidumbre en que se ven envueltas las estimaciones y procurar el uso de nuevas técnicas y metodologías para manejar estas características, como lo ofrece la estadística no paramétrica y los métodos computacionales.³

Conclusiones

En el tema del modelado de sistemas complejos, encontrar la distribución estadística que logre representar con mayor probabilidad el comportamiento del sistema real en estudio es un reto. De lograrse esto, se ganaría un peldaño más en la representatividad del modelo en cuestión, lo cual aumentaría también la probabilidad de poder realizar mejores proyecciones hacia el futuro y se conformaría como un mejor insumo para la planificación a mediano y largo plazos. Por ello, analizar las características de la distribución estadística del fenómeno estudiado es un proceso que no podemos dar por sobreentendido.

La robustez del método de remuestreo radica en que podemos encontrar la distribución estadística empírica acorde con la realidad del estimador,

³ Dysert (2006) menciona en su artículo los métodos de contingencia y de análisis de riesgo que se aplican a la estimación en el área de la ingeniería industrial, los cuales son susceptibles de ser aplicados en el ámbito demográfico.

y así mismo, estimar el valor medio, varianza, sesgo e intervalos de confianza a partir de esa misma distribución, sin hacer supuestos *a priori* que podrían no ser certeros, considerando el tema de la precisión. Aunque las distribuciones encontradas tienen comportamiento normal, presentan asimetrías y curtosis, lo cual modifica las probabilidades de sobre o subestimación respecto al valor mas probable, por lo que el método nos brinda una mayor riqueza de análisis de precisión del estimador respecto de asumir por *default* el tan usado supuesto de la normalidad.

Además, el método nos permite hacer otro tipo de evaluaciones de carácter retrospectivo en la construcción del estimador para evaluar si pequeños cambios en el estimador implican pequeños cambios en el numerador y denominador del mismo.

Aun queda la interrogante de cómo se comporta este estimador para subgrupos más pequeños, como el área urbana o la rural e inclusive para menores tamaños muestrales, casos en los cuales, según la teoría del muestreo, los métodos estadísticos son mas ingeniosos, ya que no se cuenta con tamaños de muestras grandes que reduzcan la incertidumbre. Así mismo queda pendiente analizar a mayor profundidad el impacto que tienen las pequeñas variaciones de la TGF en el numerador y denominador del mismo.

En el marco de las aplicaciones de las distribuciones *gaussianas* para modelar poblaciones, estos hallazgos pueden constituirse en un camino en el análisis de las distribuciones que mejor representan indicadores demográficos, que a su vez mejoren la probabilidad de certeza de las proyecciones poblacionales y, así mismo, la planificación del futuro. ¿Será que sí se puede planificar mejor bajo condiciones iniciales más acordes a la realidad?

Bibliografía

ARGOTE-CUSI, Milenka, 2007, “Estimación de la distribución estadística de la tasa global de fecundidad”, en *Papeles de Población*, Nueva Época, año 13 núm. 54, octubre-diciembre, Toluca.

BAGAJEWICZ, M., 2005, “On a new definition of a stochastic-based accuracy concept of data reconciliation-based estimators”, en *Computer Aided Chemical Engineering*, vol. 20, part 2.

BAGAJEWICZ, Miguel y Nguyen DUYQUANG, 2006, *Stochastic-based accuracy of data reconciliation estimators for linear systems*, University of Oklahoma.

- BAYARRI M.J. y M. A. MARTÍNEZ, 1997, “Diseño y análisis bayesianos de réplicas en la experimentación científica”, en *QUÉSTIÓ*, vol. 21, núm. 1 Departamento de Estadística e I. O. de la Universidad de Valencia.
- CHAMBERS, R. y D. DORFMAN, 1994, “Robust simple survey inference via bootstrapping and bias correction: the case of ratio estimator”, en *Southampton Statistical Sciences*, Research Institute, University of Southampton.
- DIAS, Tiago, N. ROMA y L. SOUSA, s/f, *Two level scalable motion estimation architecture with fractional-pixel accuracy and efficient data re-usage*, Instituto Superior Técnico/INESC-ID, Lisboa.
- DÍAZ, Pértegas y Pita FERNÁNDEZ, 2001, *Unidad de epidemiología clínica y bioestadística*, Complejo Hospitalario Juan Canalejo. A. Coruña. Cad. Aten Primaria.
- DYSERT, Larry, 2006, “Is estimate accuracy an oxymoron?”, en *AACE International Transactions*, Conquest Consulting Group, Vancouver.
- EFRON, B. y R. TIBSHIRANI, 1993, *An introduction to the bootstrap*, Chapman & Hall, Nueva York.
- ENDSA, 1998. *Informe final*, Instituto Nacional de Estadística de Bolivia.
- ENDSA, 2003, *Informe final*, Instituto Nacional de Estadística de Bolivia.
- FERRE, Alberto y ROMERO, Rafael, 1994. “Sensibilidad frente a datos anómalos de tres estimadores de efectos de dispersión con datos no necesariamente replicados”, en *Estadística Española*, vol. 36, núm. 135, Departamento de Estadística e I.O.,
- GOTTLIEB, Alex, 2001, *Asymptotic accuracy of the jackknife variance estimator for certain smooth statistics*, en scholar.google.com.
- LOHR, Sharon, 2000, *Muestreo: diseño y análisis*, Internacional Thomson-Paraninfo, México.
- NÚÑEZ-ANTON V. y D. ZIMMERMAN, 2001, “Modelización de datos longitudinales con estructuras de covarianza con estacionarias: modelos de coeficientes aleatorios frente a modelos alternativos”, en *QUÉSTIÓ*, vol. 25, núm. 2.
- PERREA, Manuel, 1999, “Tiempos de reacción y psicología cognitiva: dos procedimientos para evitar el sesgo debido al tamaño muestral”, en *Psicológica*, núm. 20, Departamento de Metodología. Facultad de Psicología, Valencia.
- RAJ, Des, 1968, *Sampling theory*, Mac Graw Hill.
- SITTER, R. 1992b, “Comparing three bootstrap methods for survey data”, en *The Canadian Journal of Statistics*, núm. 20.

Milenka Linneth ARGOTE CUSI

Maestra en Estudios de Población por Flacso-Méjico. Es investigadora en la institución Corpotalentos ONG. Publicaciones recientes: 2009, *Ánalisis macroeconómico del SIDA en tres países latinoamericanos: México, Brasil y Argentina 1995-2005*, en The Latin American Studies Association (LASA), del 11 al 14 de junio en Rio de Janeiro-Brasil; 2008, *Ánalisis*

comparativo de la distribución estadística de la tasa global de fecundidad de Bolivia en 1998 y 2003, en IX Reunión Nacional de Demografía, Mérida-Yucatán, 8 al 11 de octubre; 2007, “Estimación de la distribución estadística de la tasa global de fecundidad”, en *Papeles de Población*, Nueva Época año 13 núm. 54, octubre-diciembre, Toluca.

Correo electrónico: milenkalineth@yahoo.com.mx