Estimación de la distribución estadística de la tasa global de fecundidad

Milenka Linneth Argote Cusi

Centro Nacional de Prevención y Atención al VIH/SIDA e ITS

Resumen

El método de remuestreo se aplicó para generar la distribución estadística de la tasa global de fecundidad con base en los datos de fecundidad de la Encuesta Nacional de Demografía y Salud de Bolivia de 1998, que tiene un diseño estratificado y bietápico. El test Kolmogorov Smirnov de la distribución estadística de la TGF generada en mil réplicas de la muestra nos indica que no existe la evidencia suficiente para rechazar la normalidad de la distribución por muestreo. Resulta que la tasa global de fecundidad es un estimador sesgado; sin embargo, el remuestreo reduce el sesgo (el coeficiente de variación es mucho mayor al sesgo estandarizado). Si bien la estimación del intervalo de confianza de la TGF bajo el supuesto de normalidad incluye con alta probabilidad el valor del parámetro poblacional, la técnica de remuestreo permitió encontrar intervalos menos sesgados.

Palabras clave: técnica de remuestreo, evaluación del sesgo,estimación de la tasa global de fecundidad, Bolivia, encuestas por muestreo.

Abstract

Estimation of the statistical distribution of the global rate of fecundity

Bootstrap method has been applied to generate statistical distribution of global fertility rate (GFR). In this research the demography and health national survey of Bolivia in 1998 was used as it was the most recent fertility survey available up to February 2005, this survey has a stratified and bietapic design. According Kolmogorov Smirnov test of global fertility rate statistical distribution there is not enough evidence to reject normality of resampling distribution. The global fertility rate is a biased estimator but standardized biased is lower than the coefficient of variation then it is consistent. The confidence intervals are consistent and show a convergent tendency. In spite of the GFR confidence interval under normal assumption includes with high probability the parameter value, bootstrap method allows finding more accurate estimations. This method is useful to evaluate sampling error and the bias of estimations.

Key words: bootstrap, bias evaluation, estimation, global fertility rate, Bolivia.

Introducción

Las poblaciones son sistemas complejos cuyo estudio es de gran interés en las ciencias sociales. Los sistemas complejos se caracterizan por la cantidad de elementos y sus relaciones, los cuales se hallan en continuo

intercambio en el tiempo para dar como resultado un todo mayor que la suma de sus partes. Una forma de estudiar la complejidad de las poblaciones es mediante la construcción de indicadores. Cuando un indicador se calcula de la población total, recibe el nombre de parámetro, mientras que si se obtiene de una muestra, se denomina estadística (Efron y Tibshirani, 1993). 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. Los errores no muestrales en una encuesta de fecundidad se deben a la falta de cobertura de todas las mujeres seleccionadas, errores en la formulación de las preguntas y en el registro de las respuestas, confusión en la interpretación de las preguntas, problemas de memoria y errores de codificación o de procesamiento. El error de muestreo, que se mide a través del error estándar, es una medida de la variación del estimador de un parámetro en todas las posibles muestras (Cochran, 1977). En la práctica no es posible obtener todas las posibles muestras de una población, para solucionar este problema la estadística paramétrica ha construido una base teórica fundamental. La ley de los números grandes y el teorema del límite central son dos teoremas básicos de la inferencia estadística tradicional que nos permiten suponer una distribución normal de estimadores de totales y de medias, pero ¿cuál es la distribución estadística de un estimador de razón, como la tasa global de fecundidad? ¿Cómo puedo estimar el error de muestreo de un estimador diferente a un total o media?

Inferencia paramétrica vs inferencia no paramétrica

La inferencia paramétrica nos permite construir distribuciones para promedios o proporciones con base en el teorema del límite central (TLC) y la ley de los números grandes (LNG), que asumen una distribución normal del estimador cuando el tamaño de muestra crece infinitamente. En el caso de una tasa (el cociente del número de eventos ocurridos entre el tiempo de exposición al riesgo) se trata de un estimador más complejo que un promedio o un total.

Según la teoría revisada, no existen fórmulas para estimar de forma analítica los intervalos de confianza de una tasa, pero sí se han utilizado otros métodos de estimación de varianzas para funciones no lineales (series de Taylor). Por otro lado, a ello se añade que se dispone de una muestra con diseño complejo

para la estimación de la TGF. Considerando las características peculiares del estimador, nace la interrogante de ¿cuál es la distribución estadística por muestreo de la TGF?

El presente trabajo plantea la hipótesis alternativa de que la distribución de la TGF, un estimador peculiar, es diferente de la normal frente a la hipótesis nula de que sí es normal, asumiendo que la TGF se asemeja a una razón de medias y por lo tanto el TLC y la LNG son aplicables. Se recurre a la técnica de remuestreo, una técnica estadística no paramétrica, para generar la distribución estadística por muestreo de la característica de interés y, a partir de ella, evaluar si tiene un comportamiento normal (Cochran, 1977).

La coherencia y la insesgabilidad de las estimaciones son un tema crucial en la inferencia estadística. 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 y ésta es la situación a la que frecuentemente está expuesto el investigador social. Pero ¿a que se debe el interés en el estimador de la tasa global de fecundidad?

Importancia de la fecundidad

El impacto de la fecundidad en la estructura social y económica de la población hacen de esta variable demográfica una prioridad en materia de políticas de población. Ya que la fecundidad es un componente importante de la dinámica demográfica, es esencial contar con adecuadas medidas de fecundidad, interpretaciones válidas de las tendencias y diferenciales, y razonables conjeturas acerca de su futura dirección (Campbell, 1983). Cabe aclarar que el impacto de la fecundidad no siempre es el de mayor relevancia en diferentes poblaciones. Puede ocurrir que en contextos que se ven atacados por enfermedades, como en África, la mortalidad sea el fenómeno demográfico que define el crecimiento poblacional. Por otro lado, en la actualidad, el volumen de las migraciones ha aumentado tanto en los últimos decenios que en países como México tienen un fuerte impacto en la estructura poblacional. En Bolivia, el impacto de la fecundidad en el crecimiento poblacional es mayor que otros fenómenos demográficos. En un ejercicio de proyección de población bajo diferentes escenarios de fecundidad, mortalidad y migración, las variaciones de la fecundidad modifican considerablemente la estructura poblacional. Para analizar los niveles y las tendencias de la fecundidad en Bolivia se deben considerar sus

características históricas, que se reflejan en una población mayoritaria indígena con bajos niveles de educación, elevado analfabetismo y condiciones sanitarias precarias, sobre todo en las áreas rurales. Un análisis minucioso a niveles más desagregados lleva a pensar que esta estabilidad habría resultado de la cancelación de dos tendencias de sentido contrario: una declinación de la fecundidad en las áreas urbanas y una elevación en las rurales (Carafa *et al*., 1983). En 2000, la TGF se reduce a cuatro y en 2003 se estima una TGF preliminar de 3.8 por mujer, pero aún no se alcanza la fecundidad deseada de 2.5 hijos por mujer. Si bien la fecundidad total ha disminuido en 2003 en el área rural aún se tiene una TGF elevada de 5.5 hijos por mujer. Debido a este comportamiento, según la Cepal, Bolivia pasa del grupo de 'transición incipiente' a 'transición moderada', en el cual las tasas de natalidad y mortalidad aún son altas comparadas con el resto de los países de América Latina.

Por lo tanto, es de gran interés la estimación de indicadores de la fecundidad que den cuenta de su comportamiento. Las medidas de la fecundidad son numerosas tanto por el interés en grupos poblacionales específicos como por la disponibilidad de la información. Las tres fuentes principales de información de la fecundidad son el registro civil, los censos y las encuestas. En algunos países el registro civil es de buena calidad, pero en los países en desarrollo, como Bolivia, éstos son menos confiables, por lo cual se acude a las encuestas retrospectivas para subsanar las deficiencias.

Las medidas más usadas de la fecundidad son las tasas específicas de fecundidad (TEF) por grupos quinquenales de edad, definidas por el cociente de los nacimientos ocurridos a mujeres de un grupo de edad $Nac_{x,x+5}$ entre los años-persona vividos en exposición al riesgo de las mujeres $Temujeres_{x,x+5}$ del mismo grupo de edad, y la TGF (sumatoria de las tasas específicas de fecundidad multiplicado por cinco) que representa el número de hijos por mujer al final de su vida reproductiva, bajo el supuesto que a lo largo de su vida tendrá la fecundidad presente. También podemos decir que la TGF es una combinación lineal de las TEF o, desde el punto de vista estadístico, que se trata de una combinación lineal de razones.

$$TEF_{x,x+5}^{t,t+1} = \frac{Nac_{x,x+5}^{t,t+1}}{Temujeres_{x,x+5}^{t,t+1}} \qquad (1)$$

$$TGF^{t,t+1} = 5 * \sum_{i=1}^{7} TEF_i^{t,t+1} \qquad (2)$$

Datos y métodos

La presente investigación ha utilizado los datos de la Encuesta Nacional de Demografía y Salud de 1998 (Endsa, 1998), que forma parte del programa de Encuestas de Demografía y Salud (DHS) que Macro Internacional Inc. ejecuta en varios países en desarrollo. La Endsa 98 tiene una muestra probabilística nacional, la cual es estratificada y bietápica. La estratificación se realizó a nivel de diferentes subdivisiones geográficas: regiones geográficas (altiplano, valle, llanos), por departamentos dentro de cada región y por grado de marginación de los municipios dentro de cada departamento, según sus niveles de pobreza y zona de residencia (urbano-rural). El trabajo de campo se inició el 23 de marzo de 1998 en la región de los Llanos y el 26 en las otras dos regiones; concluyó el 15 de septiembre. En una primera etapa, las denominadas áreas de enumeración censal fueron consideradas como las unidades primarias de muestreo (UPM) de las cuales se seleccionaron 823 en todo el país. En una segunda etapa, los hogares particulares listados en las UPM seleccionadas fueron establecidos como las unidades secundarias de muestreo (USM). 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).

Para el manejo de las bases de datos y la implementación del algoritmo[1] de remuestreo se utilizó el programa estadístico Stata versión 8.0, toda vez que está orientado a encuestas por muestreo y también nos provee funciones para la inferencia no paramétrica.

Estimación de la tasa global de fecundidad

Partamos de la definición teórica de una tasa: es el cociente del número de eventos ocurridos entre el tiempo de exposición de los individuos a experimentar el evento. Haciendo una analogía, esta definición se acerca a un proceso *poisson* en el cual medimos, por ejemplo, el número de misiles que caen en determinada área o el número de veces que llega un bus a un paradero en determinado tiempo, etc. La relación de unidades enteras y un valor continuo, establecida a través de una razón, es lo que le brinda la complejidad de representación e interpretación a una tasa. En Demografía existen varios métodos de estimación de la TGF, entre ellos los directos y los indirectos (Campbell, 1983). El método de

[1] Secuencia de pasos para obtener un resultado.

estimación de la TGF que se aplicó se describe de forma sintética en las figuras 1 y 2.

Uno de los principales factores que había que controlar era la clasificación correcta de los nacimientos y el tiempo de exposición aportado al denominador, por grupos quinquenales de edad de la madre al momento del nacimiento del hijo x. Ya que el tiempo de exposición es una variable continua se puede dar el caso que el evento se realice en los límites de los intervalos de los grupos quinquenales de edad, aportando un tiempo de exposición al grupo anterior y otro tanto al grupo actual.

El tiempo de exposición se mide en meses y una vez controlada la clasificación por grupos quinquenales de edad en los tres últimos años (en los últimos 36 meses) anteriores a la encuesta, se pondera la base para tener como resultado una tabla con las siguientes columnas: el grupo quinquenal (categorías de uno a siete), los nacimientos en cada grupo y el tiempo de exposición aportado por las mujeres en cada grupo de edad. A partir de dicha tabla se realizó el cálculo de las tasas específicas de fecundidad y de la tasa global de fecundidad.

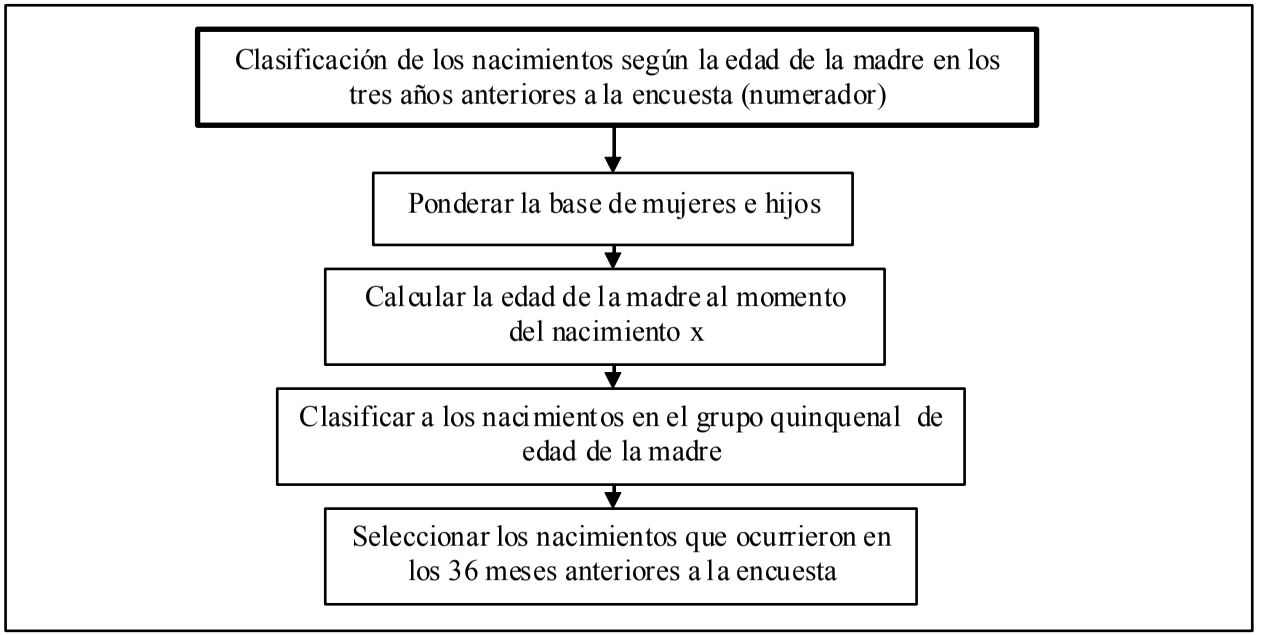
El procedimiento para el cálculo de la TGF se constituye en un módulo independiente e íntegro en sí mismo, que posteriormente se retoma (se hace referencia) en el programa de ejecución del remuestreo.

Antecedentes de la estimación de varianzas

Se han propuesto varios métodos de estimación de varianzas en la literatura para funciones más complejas que los totales y las medias. El más usado es el método de las series de Taylor que se aplica a estimadores definidos por funciones no lineales, pero como se calculan derivadas puede resultar tedioso y complicado de aplicar (Sul *et al*., 1989). Por otro lado, los métodos de grupos aleatorios utilizan submuestras, tratando en lo posible de mantener el diseño de la muestra original. Precisamente, la forma de obtener el tamaño de dichas submuestras es un problema en diseños complejos (Korn y Graubard, 1999).

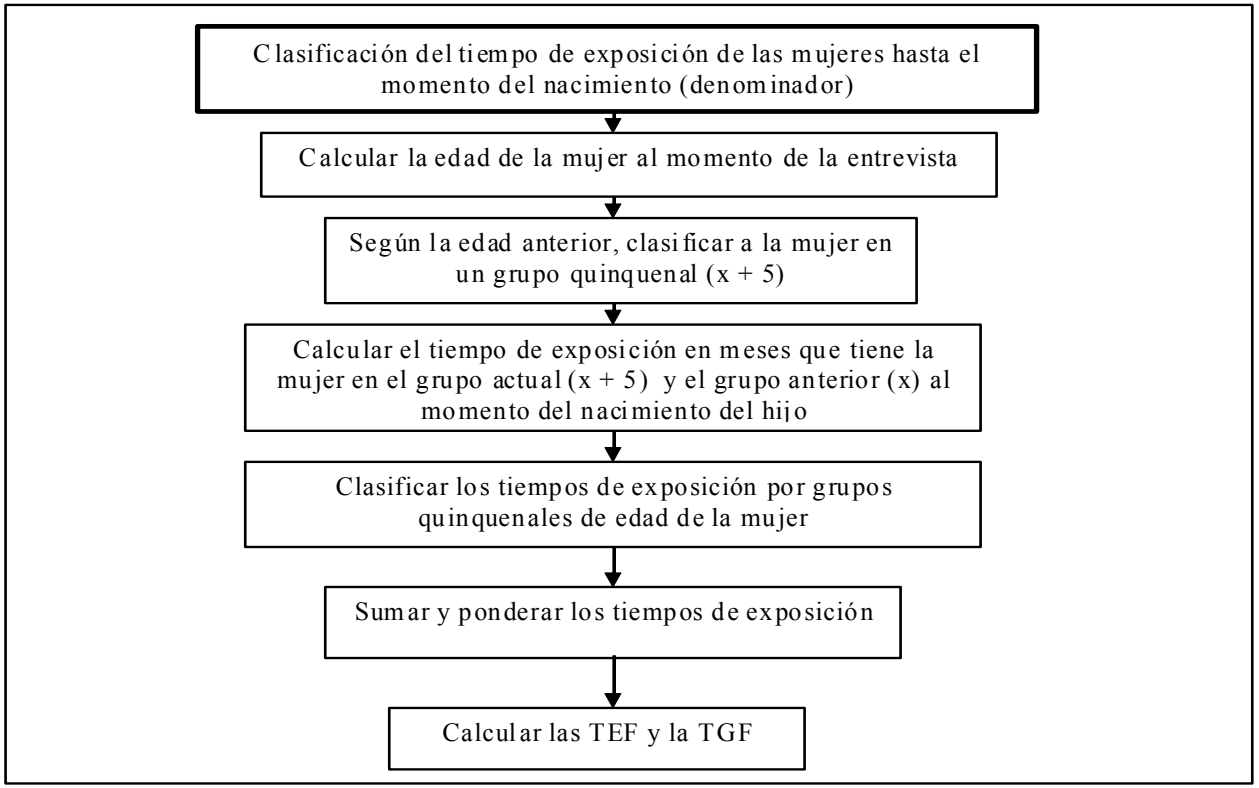
Una nueva alternativa se presenta con los métodos de remuestreo y réplicas. La ventaja de estos métodos es que conserva juntas las unidades de observación dentro de una unidad primaria mientras construye las réplicas, lo cual preserva la dependencia entre las unidades de observación dentro la misma unidad primaria (Setter, 1992a).

FIGURA 1
CÁLCULO DEL NUMERADOR DE LA TGF



Fuente: elaboración propia.

FIGURA 2
CÁLCULO DEL DENOMINADOR DE LA TGF



Fuente: elaboración propia.

Por ello son aplicables a diferentes diseños estratificados, polietápicos y de probabilidades que tienen una estructura de datos que no está idénticamente distribuida (Lohr, 2000). El método de remuestreo estándar, muestreo aleatorio simple con reemplazo, fue planteado en 1979 por Efron y Tibshirani; el interés era sobre todo estudiar el comportamiento del error estándar y del sesgo en función de la variabilidad muestral. Estos investigadores desarrollaron los conceptos básicos de la técnica con base en el principio *plug-in*.[2] En 1992, Sitter aplica el remuestreo con reemplazo (BWR) para muestras aleatorias estratificadas y observa que el estimador de la varianza, en comparación con la fórmula de Cochran (1977) para el mismo modelo, era sesgado. Para solucionar este problema propuso introducir un factor de corrección que ajusta el estimador, pero sólo para el caso de una muestra con un estrato. En este caso la varianza estimada con remuestreo es consistente e insesgada.

Los avances en el desarrollo de la técnica se relacionan con las propiedades deseables de los estimadores (insesgabilidad y consistencia) que se distorsionan en muestras complejas. En esta línea se han propuesto métodos de remuestreo para estimación de la varianza e intervalos de confianza donde el muestreo es sin reemplazo. Estos son: Jackknife, remuestreo sin reemplazo (BWO *bootstrap without replacement*), remuestreo con reemplazo (BWR b*ootstrap with replacement),* remuestreo con reemplazo mejorado (MMB *mirror match bootstrap*) y remuestreo con reescalamiento.

$\hat{\theta} = t(\hat{F})$

El método de Jackknife nos permite calcular estimadores de varianza que tienen consistencia asintótica para funciones no lineales de medias en diseños multietápicos en los cuales las unidades primarias de muestreo (UPM) son seleccionadas con reemplazo. Sin embargo, cuando la UPM es seleccionada sin

[2] Tomemos en cuenta una muestra aleatoria de tamaño n con una distribución de probabilidades denominada F .

$$F \rightarrow (x_1, x_2, ..., x_n)$$

La función de distribución empírica $\hat{F}$ asigna a cada realización de la muestra una probabilidad igual a 1/*n* a cada valor de

$x_i, i = 1,2,...,n$.

Cuando la distribución de probabilidades F es conocida, en el caso de un censo, encontrar la varianza σ^2 no es difícil. Usualmente no contamos con un censo, entonces recurrimos a la inferencia estadística que nos permite inferir propiedades de F a partir de una muestra aleatoria X. "Entonces θ es un parámetro de , mientras que $\hat{\theta}$ es una estadística basada en X. De esta manera, el estimador *plug-in* del parámetro

$\theta = t(F)$ está definido como .

reemplazo, Jackknife sólo se puede implementar para diseños estratificados (Efron, 1982). Otro método consiste en reescalar (ajustar el ponderador) el método estándar cuando el estimador es una función no lineal de las medias (Rao y Wu, 1988); en este método, el algoritmo se aplica con un tamaño de muestra seleccionado m_k no necesariamente igual a n_h (tamaño de muestra en el estrato h) y se reescala los valores de la remuestra apropiadamente para tener estimadores insesgados en el caso lineal. Como se tiene que reescalar cada dato en cada remuestra, este proceso puede ser complicado en muestras más complejas. Mientras BWO y BWR sólo son aplicables a diseños simples, la versión de remuestreo con reescalamiento se extiende para diseños más complejos para funciones de medias y son computacionalmente más intensivos y de difícil uso.

Aplicaciones de los métodos de remuestreo

Las investigaciones realizadas por los autores mencionados anteriormente han utilizado poblaciones finitas hipotéticas para experimentar con los diferentes modelos de remuestreo. Se ha puesto interés en el estudio de estimadores que son funciones lineales de medias; sin embargo, también se han realizado simulaciones para el caso de estimadores de razón, coeficientes de regresión, coeficientes de correlación y la mediana. Se concluye que el MMB y BWO tienen un buen desempeño en comparación con otros métodos (Sitter, 1992b). En 1994 se aplica el remuestreo con reemplazo mejorado (*mirror-mattch*, MMB) a un diseño estratificado y de tres etapas para la estimación de una razón con resultados aceptables. A pesar de que son necesarias más investigaciones, se tiene evidencia de que la estimación de los intervalos de confianza es apropiada según el principio *plug-in* (Robb, 1994). En Australia se aplicaron varios métodos de remuestreo para un estimador de razón (Y: ingreso por la venta de un producto; X: cantidad de ese producto) donde se encontraron elevados sesgos para el estimador que pueden ser atribuibles al tipo de estimador seleccionado, aunque también sugieren aplicar métodos de remuestreo más sofisticados para la corrección del sesgo (Davidson y MacKinnon, 1993).

Los métodos de remuestreo se han aplicado sobre todo en el área económica. Un ejemplo importante es el interés de estudiar la significancia estadística de los cambios en los indicadores de desigualdad y bienestar a través del índice de Gini. Si las encuestas de ingresos se basaran siempre en los mismos hogares, las variaciones temporales en los indicadores de desigualdad y bienestar verdaderamente reflejarían cambios en la distribución del ingreso. En realidad,

es más frecuente encontrar encuestas que se aplican a hogares diferentes periodo a periodo. Por ello, las diferencias en estos indicadores podrían atribuirse simplemente al hecho de que la muestra cambió y no a variaciones reales en la desigualdad del ingreso. Por ejemplo, el coeficiente de Gini computado en el año *t* puede ser superior al del año *t*-1, simplemente por fenómenos muestrales (independientemente de que haya cambiado la distribución del ingreso o no), por lo que la conclusión de que la distribución se ha vuelto más desigual no es necesariamente correcta (Gasparini y Sosa, 1998).

Características del modelo de remuestreo utilizado

Los resultados teóricos de la técnica de remuestreo fueron desarrollados para áreas estadísticas distintas de las encuestas por muestreo. La extensión del método de remuestreo a muestras complejas es reciente y uno de sus posibles usos es la estimación de distribución estadística por muestreo de un estimador de razón como la TGF.

En el presente trabajo $\hat{\tau}$ es un estimador de la tasa global de fecundidad de Bolivia en 1998. La única información disponible proviene de una muestra que tiene una estructura de datos jerárquicos, lo cual aumenta la complejidad de las estimaciones. El objeto de interés consiste en obtener una medida de dispersión para $F(\tau)$ y un intervalo de confianza para la estimación puntual de τ. La evaluación analítica de estas estadísticas requiere conocer la distribución $F(\tau)$. Lo cierto es que no se conoce la distribución por muestreo de *F* y que la derivación del error estándar (es) y el intervalo de confianza son analíticamente complejos; el método de remuestreo, como se mencionó, nos permite aproximar $F(\tau)$ utilizando la distribución empírica de la muestra $\hat{F}(\tau)$.

Se aplica el remuestreo con reemplazo considerando en una primera etapa las UPM y en una segunda etapa las USM. El remuestreo con reemplazo es menos eficiente que el muestreo sin reemplazo, pero se utiliza debido a la facilidad que brinda para elegir y analizar las muestras. El remuestreo se puede aplicar en muestreos por conglomerados en dos etapas, Robb (1994) encuentra que la contribución de la varianza en la segunda etapa será despreciable en comparación a la primera etapa (Lohr, 2000) cosa que no ocurre en la presente aplicación, como se explicará en detalle al presentar los resultados.

El modelo de remuestreo que aplica el presente trabajo se asemeja al remuestreo con reemplazo planteado por McCarthy y Snowden en 1985, con la diferencia de que aquí el tamaño m_h de las submuestras dentro de los estratos

es igual al tamaño n_h de la muestra en los estratos. El método consiste en obtener una muestra aleatoria simple con reemplazo de tamaño n_h independientemente en cada UPM (unidad censal), para luego, en la segunda etapa, hacer lo mismo con las USM (hogares).

Los ponderadores se aplican directamente en la estimación puntual de la TGF a partir de la muestra aleatoria simple de mujeres y sus hijos, siguiendo el diseño de la muestra; los pesos de muestreo poseen la información para determinar los errores estándar de las estimaciones, toda vez que en ellos se refleja el diseño de la muestra (Lohr, 2000).

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 (figura 3). Se aplica el método de percentiles para el cálculo de los intervalos de confianza al 95 por ciento porque la distribución estadística muestral de la TGF obtenida por remuestreo se asemeja a la distribución normal según la prueba Kolmogorov Smirnov[3] (0.901438 > 0.05). Una buena estimación del intervalo de confianza se obtiene en 1000 réplicas (Efron y Tibshirani, 1993), por lo cual los resultados y las conclusiones se presentan tomando en cuenta esta referencia.

Supuestos

X_n

Un primer supuesto sobre el cual se sustenta la técnica de remuestreo se refiere al hecho de considerar la muestra como si fuera la población total. Es decir, se considera que la distribución de los datos observados (muestra) representa una buena estimación de la distribución de la población. Dicho supuesto se fundamenta en la teoría asintótica que se basa en el análisis de la convergencia en probabilidad y la convergencia en la distribución. Sea la variable aleatoria X_n, esta variable converge en probabilidad con una constante *c* si cuando el tamaño de la muestra tiende a infinito, la probabilidad de que la diferencia entre el valor muestral y el verdadero sea mayor a un valor permisible ε tiende a cero, para cualquier $\varepsilon > 0$. Esto significa que cada vez es menos probable que sea diferente de *c,* a medida que *n*, el tamaño de la muestra, aumenta.

[3] El test Kolmogorov Smirnov (KS) de una muestra es una prueba de bondad de ajuste. Mide el grado de semejanza entre la distribución proveniente de los datos observados y una distribución teórica (Siegel, 1956: 47).

FIGURA 3
REMUESTREO PARA MUESTRAS ESTRATIFICADAS

1. Obtener una muestra con reemplazo $\{y^*_{hi}\}_{i=1}^{n_h}$ de la muestra original $\{y_{hi}\}_{i=1}^{n_h}$ independientemente en cada estrato.
2. Calcular $s^* = \{y^*_{hi} : h = 1,2,...,L; i = 1,2,...,n_h\}$ donde $\theta^* = \theta(s^*)$
3. Repetir el paso 1 un gran número de veces, B, para obtener $\hat{\theta}^*_1, \theta^*_2, ..., \theta^*_B$.
4. Estimar la var$(\hat{\theta}) = E(\theta^*(b) - \theta^*(\circ))^2$ donde $\hat{\theta}^*(\circ) = \sum_{b=1}^{B} \theta^*(b)/B$
5. Estimar el error $s\hat{e}_B = \left\{ \sum_{b=1}^{B} [\hat{\theta}^*(b) - \hat{\theta}^*(\circ)]^2 / (B-1) \right\}^{1/2}$
6. Estimar el sesgo $\hat{bias}_B = \hat{\theta}(\circ) - t(\hat{F})$
7. Estimar el intervalos de confianza al 95%

$\hat{\theta}_{\text{inf}} = \hat{\theta}^{*(\alpha)} = 2.5$ avo percentil de la distribución de $\hat{\theta}^*$

$\hat{\theta}_{\text{sup}} = \hat{\theta}^{*(1-\alpha)} = 97.5$ avo percentil de la distribución de $\hat{\theta}^*$

Fuente: Sitter (1992a).

Por otro lado, X_n converge en distribución a una variable aleatoria X con función de distribución acumulada $F(X)$, si la diferencia entre la distribución muestral y la distribución verdadera tiende a cero a medida que aumenta el tamaño de la muestra, para todos los puntos de continuidad de $F(X)$. Estos conceptos básicos de la teoría asintótica (Davidson y MacKinnon, 1993) están relacionados directamente con el principio *plug-in* (Efron y Tibshirani, 1993).

En una segunda instancia se supone que el estimador de la TGF que se utiliza en esta investigación es bueno. Para Chou (1977), existen infinitos estimadores de un parámetro; probar con todos ellos y encontrar el mejor es imposible. En este caso, la función utilizada para la estimación de la tasa global de fecundidad se considera un "mejor" estimador, toda vez que aplica la definición teórica de una tasa, es decir, considera en el denominador el tiempo de exposición al riesgo de las mujeres en edad reproductiva en lugar de la población promedio de mujeres en edad reproductiva que se utiliza comúnmente.

En relación a la muestra, se supone que la muestra básica que se utiliza para el remuestreo es representativa. Este concepto es crucial en todos los razonamientos estadísticos y se entiende como el hecho de que la muestra debe parecerse a la población. Es difícil asegurar la representatividad en muestras pequeñas que no han sido obtenidas con procedimientos aleatorios. Sin embargo, la ley de los grandes números nos permite esperar muestras representativas (Méndez, 2004). La muestra de la Endsa que procede de un muestreo estratificado y bietápico tiene una elevada probabilidad de ser representativa.

Resultados

Estimación paramétrica de los intervalos de confianza

Como punto de referencia se aplicó en primer lugar el método de inferencia estadística tradicional. La construcción de los estimadores implica la estimación de totales a los diferentes niveles de desagregación. Finalmente, la complejidad se resume a la sumatoria de la multiplicación de los factores de ponderación de los diferentes niveles, por la TGF ($\hat{t}$) al nivel más bajo.

En este ejercicio se ha construido una tabla que contiene información del código que identifica unívocamente a cada mujer, el ponderador, el grupo quinquenal al que pertenece, el tiempo de exposición en el grupo respectivo y el número de hijos en el grupo. Una vez anualizado el tiempo de exposición (te) se utilizó la función de estimación de razones en muestras complejas que posee

el Stata para la estimación de las tasas específicas de fecundidad definidas por el cociente hijos/te en cada grupo de edad. Posteriormente se calcula la TGF a través de una combinación lineal de las tasas específicas de fecundidad (considerando las covarianzas entre las TEF), se estima el error estándar y los intervalos de confianza bajo el supuesto de normalidad de la distribución estadística de la TGF. Para 1998 se estima una TGF de 4.243 hijos por mujer, dato que se acerca a la estimación oficial del INE de Bolivia (4.2), y el intervalo de confianza de 95 por ciento es [4.108, 4.379] (tabla 1). ¿Qué tan acertadas son estas estimaciones considerando que se obtienen de una muestra de la población total y que se asume una distribución teórica, la normal, para la estimación de los intervalos de confianza? *A priori* no se conoce la distribución estadística de la TGF.

Estimación no paramétrica: remuestreo aleatorio simple

Antes de la aplicación del remuestreo considerando el diseño complejo de la muestra se experimentó con muestreo aleatorio simple (MAS) para evaluar el comportamiento de los datos en relación a la teoría de Efron y Tibshirani (1993). Ejecutando el algoritmo de estimación de las TEF y la TGF que se describió en la sección de métodos se obtiene una estimación puntual inicial de $\hat{t} = t(\hat{F})$ = 4.228 de la TGF (no considera las covarianzas entre las TEF). Para 1 000 réplicas obtenemos una estimación $\hat{t} = t(\hat{F})$ = 4.327 y un intervalo de confianza de 95 por ciento de [4.187, 4.474] (tabla 1). El intervalo es coherente e incluye al valor estimado $\hat{t}$ (la TGF oficial en 1998, está más cerca del límite inferior). A medida que el número de réplicas aumenta, la distribución por muestreo se asemeja más a la normal mientras el *ês* disminuye de forma leve.

En la tabla 1 podemos observar que los coeficientes de variación (CV) son mayores al cociente entre el sesgo y el error estándar (con remuestreo aleatorio simple) aunque no por mucho, esto nos indica que el sesgo de $\hat{t}$ es consistente porque se estaría cumpliendo la siguiente desigualdad llamada sesgo de $\hat{t}$ estandarizado (Raj, 1968):

$$\frac{[E(\hat{t}) - t]}{[V(\hat{t})]^{1/2}} \leq CV(\hat{t}) \qquad (3)$$

Donde

(4)

Remuestreo en una etapa del diseño de la muestra

El tercer paso en la experimentación con remuestreo toma en cuenta solamente una etapa, esto en parte para verificar si realmente el error estándar aportado en una segunda etapa puede ser considerado despreciable, como lo indica la teoría basada en la series de Taylor. Se realizaron varios experimentos con diferentes números de réplicas (200, 400, 1000) considerando las UPM y se encontró una mayor estabilidad de la distribución por muestreo de $\hat{t}$ asemejándose a la normal (figura 4a). La prueba Kolmogorov Smirnov para 1000 réplicas (sig = 0.901438 > 0.05) nos indica que no podemos rechazar la hipótesis de la normalidad de la distribución estadística de la tasa global de fecundidad de Bolivia en 1998.

Cabe hacer notar que esta variabilidad se mantiene en determinados límites, para nuestro caso entre 0.05525 y 0.05882. También hay que notar que el valor del error estándar en 1 000 réplicas es menor que el error estándar con muestreo aleatorio simple (tabla 1). Esto se explica porque la varianza dentro de los estratos es menor que la varianza global a efectos del diseño.

$$CV(\hat{t}) = \frac{\hat{e}s}{|\bar{t}|}$$

Respecto al comportamiento del estimador de la TGF en función del número de réplicas, podemos observar que se tiene una tendencia hacia la convergencia. Para menos de 400 réplicas la TGF estimada es más fluctuante, mientras que para B mayor o igual que 400, el estimador adquiere mayor estabilidad. Este fenómeno comprueba la teoría asintótica detrás del método de remuestreo.

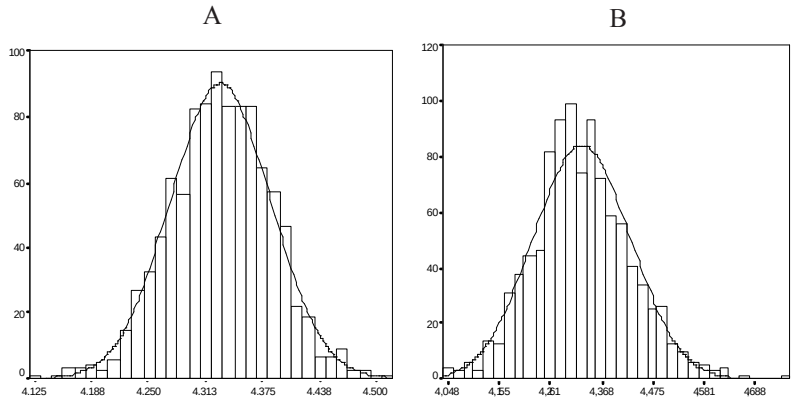
El remuestreo en una etapa del diseño, como era de esperarse, nos permite estimar una TGF sesgada, es decir, el valor esperado es diferente del parámetro que en este caso es 4.2, según el INE. Pero el sesgo estandarizado es mayor que el coeficiente de variación, lo cual nos indica que el sesgo no es consistente según la ecuación 4. De otra manera se puede interpretar también que el intervalo de confianza en una etapa del diseño de la muestra no estaría incluyendo toda la variabilidad del estimador. En tal sentido, aunque la teoría nos decía que la varianza en una segunda etapa es despreciable, fue necesario experimentar en una segunda etapa del diseño de la muestra.

TABLA 1
ERROR ESTÁNDAR (ES), MEDIA, SESGO, COEFICIENTE DE VARIACIÓN (CV) E INTERVALOS DE CONFIANZA (IC) ESTIMADOS POR INFERENCIA PARAMÉTRICA (1) Y NO PARAMÉTRICA (2, 3, 4)

						Ic 95%			
Es	Media	Parámetro 1998	Sesgo	Sesgo/es	CV	Inf.	Sup.	KS-Z	Sig.
Estimación paramétrica de la TGF									
0.069	4.24	4.230	0.013	0.192	1.630	4.108	4.379	N.A.	N.A.
Remuestreo aleatorio simple									
0.073	4.327	4.230	0.097	1.334	1.678	4.187	4.474	0.856	0.456
Remuestreo en una etapa del diseño de la muestra									
0.055	4.329	4.230	0.099	1.792	1.276	4.224	4.437	0.570	0.901
Remuestreo en dos etapas del diseño de la muestra									
0.102	4.327	4.230	0.097	0.952	2.364	4.172	4.499	0.995	0.275

Fuente: cálculos propios realizados en Stata, con base en la Endesa 1998.
Estimaciones a 1000 réplicas, NA = no aplica.
Se consideró como parámetro al dato oficial del INE en 1998.

FIGURA 4
DISTRIBUCIÓN ESTADÍSTICA DE LA TASA GLOBAL DE FECUNDIDAD PARA EL TOTAL DE LA POBLACIÓN GENERADA EN MIL RÉPLICAS DEL: A) REMUESTREO EN UNA ETAPA Y B) REMUESTREO EN DOS ETAPAS



Fuente: elaboración propia con base en los datos del remuestreo.

Remuestreo en dos etapas del diseño de la muestra

En primer lugar, se observa que la distribución tiene un comportamiento normal (figura 4b). La media de la distribución no varía, se mantiene constante en 4.32 hijos por mujer para 1998, valor que es mayor a la estimación paramétrica (método 1 de la figura 5b). El error estándar casi se duplica en comparación con el método anterior, ya que además de la variabilidad en las UPM se considera la variabilidad en las USM, por ello en este caso no podremos despreciar el error estándar en una segunda etapa del diseño de la muestra.

El error estándar bajo el supuesto de normalidad (método 1) no difiere en gran medida del remuestreo aleatorio simple (figura 5a), disminuye en el remuestreo en una etapa, pero aumenta considerando el diseño completo de la muestra (0.102).

Según la desigualdad del sesgo estandarizado (ecuación 4), el remuestreo en una segunda etapa nos brinda estimaciones más precisas y consistentes, dado que el coeficiente de variación es mucho mayor al sesgo estandarizado. En la figura 6a podemos observar que en los métodos uno, dos y cuatro se cumple

la desigualdad mientras que con el remuestreo en una etapa la probabilidad de que el intervalo incluyera los valores reales de la TGF era menor, este problema se resuelve considerando el diseño completo de la muestra en el remuestreo.

Los intervalos de confianza de 95 por ciento, estimados a partir de los diferentes métodos, son coherentes dado que incluyen las diferentes estimaciones de la TGF (figura 6b). Aproximadamente la amplitud del intervalo de confianza es la misma para los diferentes métodos a excepción del tercero (remuestreo en una etapa) que es menor. Considerando que se acepta la normalidad de la distribución estadística por muestreo de la TGF, se esperaría que los estadísticos estimados sean similares por cualquier método (menos el tercero); sin embargo, la mayor diferencia que se encuentra radica en la centralidad de la media por remuestreo en los diferentes intervalos estimados. La media por remuestreo (4.32) es superior a la estimación puntual por inferencia paramétrica (4.24), la cual se acerca al límite inferior de las estimaciones por remuestreo. Considerando el parámetro de la TGF a nivel nacional en 1998 de 4.2 hijos por mujer se obtuvieron sesgos mayores pero consistentes.

Cabe resaltar que, en 2002, el Instituto Nacional de Estadística de Bolivia publicó nuevas estimaciones de la TGF para los quinquenios 1990-1995, 1995-2000, en cuya confección tomó en cuenta todas las fuentes de información disponibles acerca de la fecundidad hasta la fecha (diversas encuestas y el censo de 2001). Para el periodo 1995-2000 se obtuvo una TGF de 4.32 hijos por mujer, justamente la media de la distribución estadística estimada por remuestreo en el presente trabajo. Es decir, si bien los intervalos de confianza varían en función del error estándar que a su vez depende del diseño de la muestra, la media por muestreo se mantiene constante y se puede considerar como una estimación de menor sesgo o que tiene una mayor probabilidad de acercarse al valor real.

El error estándar de una estimación estadística, usando un diseño multietápico como el usado para la Endsa 1998, es más complejo que el error estándar basado en el muestreo al azar simple y tiende a ser mayor que el error estándar producido por una muestra al azar simple. El incremento en el error estándar debido al uso de un diseño multietápico es conocido como el efecto del diseño y se define como la razón entre la varianza de la estimación con el diseño actualmente usado y la varianza de la estimación que resultaría si se usara una muestra al azar simple. Cuando toma el valor de uno, indicará que el diseño utilizado es tan eficiente (proporciona varianzas mínimas) como uno simple al azar, y mientras que un valor mayor a uno que el diseño utilizado produce una varianza mayor a la que se obtendría con una muestra simple al azar (Cepep, 2004).