

MÉTODOS DE ESTIMACIÓN PARA ÁREAS PEQUEÑAS Y UNA APLICACIÓN A LA PREVALENCIA ANTICONCEPTIVA*

Alfredo Aliaga
Macro Internacional Inc.

RESUMEN

Durante la implementación del programa internacional de encuestas de demografía y salud (EDS) ha aumentado la demanda de información sobre áreas geográficas o grupos de población no contemplados en el diseño de esas encuestas. En el presente documento se pasa revista a los principales obstáculos técnicos que deben enfrentar las diversas metodologías de estimación para satisfacer esa demanda y se propone un proceso de estimación alternativo, derivado de la metodología de regresión.

* Documento presentado en la Conferencia conjunta IAOS-AFSA, Comisión Económica para África de las Naciones Unidas, Addis Abeba, Etiopía.

**ESTIMATION METHODS FOR SMALL AREAS
AND THEIR APPLICATION IN MEASURING THE USE
OF CONTRACEPTIVES**

ABSTRACT

During the implementation of the Demographic and Health Survey (DHS) Programme, demand has been growing for information on geographic areas and population groups that have not been defined in the design of these surveys. This article reviews the main technical obstacles that the various estimation methodologies will have to overcome in order to meet that demand and proposes an alternative regression-based estimation methodology.

INTRODUCCIÓN

Aunque se pueden obtener *estimaciones directas* de las características sociodemográficas a nivel nacional en todas las encuestas por muestreo de las Encuestas Demográficas y de Salud (EDS), generalmente no es posible lograr *estimaciones directas para áreas pequeñas*, ya que el tamaño de las muestras correspondientes a esas áreas generalmente no es suficiente para que aquellas sean representativas —es decir, el tamaño adoptado para la muestra nacional es insuficiente para realizar estimaciones para los dominios pequeños. Del mismo modo, el nivel de clasificación utilizado para el diseño de las muestras nacionales no es adecuado para las áreas pequeñas. Sin embargo, en los últimos tiempos, los directores de programas y los encargados de la formulación de políticas de los países en desarrollo han manifestado interés en la estimación de indicadores sociodemográficos básicos para las áreas pequeñas a fin de definir objetivos, asignar recursos y supervisar el funcionamiento de los programas de atención de salud y de planificación familiar.

La realización de encuestas nacionales que también sean representativas de las áreas geográficas de nivel inferior es posible, pero no puede ser viable desde el punto de vista de los costos. Además, los errores no vinculados con el muestreo suelen aumentar a medida que crece el tamaño de la muestra. Una solución alternativa es utilizar las *estimaciones indirectas* para las áreas pequeñas. En el presente estudio se pasa revista a los diversos enfoques utilizados en las estimaciones de las áreas pequeñas y de los principales problemas que presentan.

1. FUENTES DE DATOS

Censos de población. Los censos de población —realizados generalmente cada 10 años— proporcionan los valores de base para obtener algunos indicadores sociodemográficos. En los países en que se realizan las EDS, se puede utilizar la información censal para determinar las características básicas de las áreas pequeñas y luego aplicar diversos procedimientos para elaborar las estimaciones estadísticas correspondientes a las áreas pequeñas seleccionadas. La información censal también puede utilizarse para determinar los factores de ponderación censal para las características no recopiladas en el censo y como antecedente en los procedimientos de modelización.

Registros administrativos. Los registros administrativos, como las estadísticas vitales o los registros de salud, son una fuente importante de información para las estadísticas de áreas pequeñas. Sin embargo, el programa de las EDS no los ha utilizado ampliamente porque en muchos casos no es fácil acceder a esos registros en los países en que se realizan las encuestas. Además, los registros administrativos se realizan más bien con fines programáticos que estadísticos, de modo que las definiciones individuales de las fuentes generalmente no son compatibles y no pueden combinarse con facilidad con los datos que poseen definiciones estandarizadas, como los datos de las EDS. Así, el uso de registros administrativos como fuente de información para las estadísticas de áreas pequeñas puede ser limitado.

Grandes encuestas de hogares. A menudo, las grandes encuestas de hogares están a cargo de las oficinas de estadística nacionales de los países en que se realizan las EDS. Cuando se dispone de datos correspondientes a las primeras, éstos pueden proporcionar información de base sobre las estadísticas de áreas pequeñas.

Las EDS. El proyecto de Encuestas de Demografía y Salud (EDS) proporciona información sobre algunos indicadores sociodemográficos y de salud para la población total, para las poblaciones urbanas y rurales y para algunas poblaciones regionales importantes. Así, las EDS, en forma individual o combinadas con otras bases de datos, permiten obtener los datos sociodemográficos y de salud necesarios para las estimaciones de áreas pequeñas.

2. PROBLEMAS FUNDAMENTALES DE LAS ESTIMACIONES DE ÁREAS PEQUEÑAS

Cualquier encuesta por muestreo se diseña para obtener estimaciones sobre dominios de gran tamaño. Aunque los problemas relacionados con las áreas pequeñas son críticos, resultaron menos polémicos para los dominios principales. Puede encontrarse un buen análisis sobre este tema en *Small Areas Statistics: An International Symposium* (Platek y otros, 1987).

Definiciones y conceptos. A veces no existe una definición común o un mismo conjunto de definiciones y conceptos en las diversas fuentes de datos. Por ejemplo, puede haber diferencias en las condiciones que deben reunir los encuestados para cada fuente de datos, en los procedimientos utilizados para la recopilación de datos y en la manera en que el método de estimación se aplica a la base de datos. Tal disparidad hace difícil combinar las fuentes de datos. Estos problemas son más graves en los países en desarrollo.

Cambios en las políticas de los gobiernos. Los cambios en las políticas de los gobiernos pueden incidir en la producción de datos de las fuentes. Estos cambios se verán reflejados en la población que abarca la fuente, el grado de cobertura de diversos temas y la calidad de los datos obtenidos, asociada al tiempo y los recursos asignados a la recopilación de datos.

Problemas de los procedimientos de estimación. Los procedimientos de estimación están supeditados a la situación relativa a los factores mencionados. El tipo de estimación dependerá de las condiciones de la información del país de que se trate. En muchos países en desarrollo es recomendable recurrir a procedimientos sencillos en lugar de utilizar métodos complejos. Para estos últimos se requiere personal muy capacitado, una mayor asignación de recursos y un procesamiento de datos intensivo.

Privacidad. Como existen diferencias entre las fuentes y es preciso combinar información de varias bases de datos, a menudo hay que analizar registros individuales o grupos de registros. El problema de la privacidad y sus consecuencias es una preocupación legítima del público en general y los investigadores deben establecer procedimientos claros para asegurar la privacidad de la información.

Tamaño. El censo de población es la mejor fuente para la estimación de una estadística básica referida a un dominio, cualquiera sea la magnitud de éste. Sin embargo, para estadísticas más complejas,

recopiladas en estudios limitados —y no en censos— no pueden realizarse estimaciones de estas estadísticas para todos los casos en razón de las limitaciones del número de casos, ya que puede existir carencia de datos o muy pocos datos. Las encuestas por muestreo pueden desempeñar un papel importante en la complementación de esa información. Sin embargo, no permiten obtener estimaciones para dominios de cualquier tamaño. Generalmente, la muestra para una encuesta se diseña con el fin de calcular estimaciones para dominios de gran magnitud. La definición del área pequeña está relacionada con los objetivos del diseño de la muestra y con su tamaño.

En general, el programa EDS proporciona estimaciones de los parámetros demográficos básicos correspondientes al país en su conjunto, a las áreas urbanas y rurales como subgrupos distintos y a las principales regiones del país. Las áreas de menor tamaño que las regiones principales apenas están representadas en las EDS y el tamaño de la muestra termina siendo nulo o muy pequeño. En el programa de EDS, el término *área pequeña* se limita al nivel geográfico inmediato inferior —generalmente una provincia o un distrito— cuyos datos no se publican. Actualmente, el programa de EDS no prevé realizar estimaciones para las áreas pequeñas.

Estimación. La estimación del dominio forma parte de todas las encuestas por muestreo. El tipo de procedimiento de estimación depende del diseño de la muestra. Como es difícil medir el sesgo, los procedimientos que permiten obtener estimaciones no sesgadas son más aceptables que los procedimientos que dan por resultado estimaciones sesgadas. Algunos autores prefieren los procedimientos de estimación consistentes porque cuando el tamaño de la muestra llega a ser suficientemente grande, el valor de la estimación es más confiable para el dominio, es decir, el valor del sesgo tiende a ser pequeño. Algunos autores tratan de utilizar esa consistencia para asegurar que las propiedades de la muestra sean convenientes, al menos cuando ésta es grande.

Otro método de clasificación general de los procedimientos de estimación se basa en el nivel de probabilidad muestral del diseño de la muestra o en el uso de datos auxiliares de otras fuentes. Se puede hacer una *estimación directa* usando los datos de la encuesta, pero sólo las unidades correspondientes al dominio del área pequeña. Otro tipo de estimación es una modificación de la estimación directa, denominada *estimación directa modificada*. En ella se utilizan datos de la encuesta principal —de otros dominios, tanto para la variable de estudio como

para la auxiliar— pero se conservan las propiedades del diseño, como la consistencia y la ausencia de sesgo. Finalmente, existe un número creciente de investigaciones sobre las *estimaciones indirectas (o modelización)* que utilizan datos externos —tanto para las variables de estudio como las auxiliares— que no pertenecen al dominio o al período correspondientes al estudio, con o sin consideración de las propiedades de diseño de la encuesta.

Evaluación. Una forma ampliamente aceptada de evaluar cualquier estimación es determinar su error relativo. Este valor se calcula como el cociente entre el error estándar de la estimación y el valor de ésta. Un problema fundamental para cualquier encuesta por muestreo es decidir si es más conveniente evaluar la estimación de un gran número de variables usando estimadores de diseño o la de un número reducido de estimaciones indirectas realizadas porque el tamaño de la muestra era pequeño.

Los procedimientos que permiten obtener estimaciones consistentes presentan otros problemas. Las características de estas estimaciones, calculadas a partir de una muestra de gran tamaño, son adecuadas, pero no ocurre lo mismo cuando el tamaño de la muestra es pequeño. Por otra parte, una de las dificultades más importantes radica en obtener medidas confiables de calidad —errores de muestreo o errores cuadráticos medios— cuando se usan estos procedimientos de estimación para una encuesta determinada.

Un problema fundamental en ciertos procedimientos de modelización para el cálculo de las estimaciones es que el modelo no sea válido para algunas áreas pequeñas. Puede ocurrir algo semejante con las estimaciones directas: si bien la muestra puede tener un tamaño razonable, es posible que el valor de la estimación no sea confiable debido a las variaciones muestrales.

Otra dificultad en la estimación de áreas pequeñas está vinculada con el uso de información correspondiente a distintos períodos. Si el investigador quiere analizar los cambios en el tiempo de una determinada variable habrá ciertas dudas sobre la confiabilidad de esa estimación. Este problema es menos evidente cuando sólo se requiere una estimación puntual en un momento determinado.

Idealmente, el objetivo es desarrollar un procedimiento razonable de estimación que sea apropiado a todas las estimaciones posibles del dominio según un procedimiento de calidad determinado. Al respecto, se sugiere realizar un estudio de validación para medir la eficacia de cada valor directo en comparación con un valor indirecto bajo varias

condiciones. Sin embargo, la realización de un estudio de esta naturaleza exige un esfuerzo muy grande.

Diseño muestral. El diseño muestral satisface las necesidades de los dominios principales. Sin embargo, se debe considerar la posibilidad de contar con enfoques apropiados para otros dominios en los que la estimación se considera posible en un futuro cercano. Los procedimientos de estimación para los dominios planificados están bien establecidos, pero resultan más difíciles para los dominios no planificados. Para resolver este problema es preciso encontrar fórmulas de transacción en el diseño de la muestra. Las soluciones parciales comprenden modificaciones en las estrategias de afijación del tamaño muestral, del nivel de agrupación y de la estratificación de la muestra.

Si se conoce el área pequeña de antemano, la afijación de la muestra se puede modificar mediante el diseño de la muestra —y tratar el área como dominio planificado— aumentando el tamaño de ésta. (Esto supone disponer de suficientes recursos humanos y financieros.) Si mediante el estudio sólo se desean obtener estimaciones nacionales, la afijación proporcional de la muestra puede ser suficiente. Sin embargo, si la muestra se utilizara para obtener estimaciones provinciales o de distrito, entonces se podrían obtener tamaños de muestra comparables entre provincias fijando un tamaño de muestra igual para cada una a fin de lograr el mismo nivel de confiabilidad. Si se aumenta el tamaño de las muestras de las provincias según las necesidades, se obtendrá un tamaño muestral mayor al que correspondería al país en su conjunto mediante una afijación proporcional. Podría producirse una situación similar si primero se examina una estimación provincial y luego una estimación por distritos de una misma provincia. En este caso, el diseño de la muestra tendrá que conciliar la confiabilidad de los diversos niveles del dominio. En general, será necesario encontrar una solución de compromiso entre distintos tipos de afijación: proporcional o igualitaria, óptima o proporcional, óptima o igualitaria, etc.

Otro factor que puede utilizarse para mejorar las estimaciones de áreas pequeñas es el nivel de clasificación. En general, el nivel apropiado para las estimaciones de los dominios nacionales y subnacionales no es adecuado para la representación de las áreas pequeñas. Teniendo en cuenta la magnitud de los recursos disponibles para la encuesta, una solución de compromiso es disminuir el nivel de clasificación en las primeras etapas del diseño de la muestra. En el programa EDS, por lo general, los niveles de clasificación utilizados

son adecuados para los niveles nacionales, subnacionales y regionales. Sin embargo, para mejorar el tamaño muestral para las áreas pequeñas, debe incorporarse al diseño de la muestra un número mayor de unidades más pequeñas en la primera etapa.

La estratificación de las muestras puede utilizarse para determinar áreas pequeñas más representativas disminuyendo el tamaño de los estratos. Si al mismo tiempo se disminuye el nivel de clasificación, se mejorarán las estimaciones para los dominios no planificados.

Redefinición de las áreas pequeñas. Si no se logra mejorar las estimaciones del dominio pequeño aumentando el nivel de la afijación o disminuyendo el nivel de clasificación, una solución intermedia sería redefinir el dominio colapsando áreas pequeñas cercanas entre sí y realizando una afijación de tamaño muestral suficiente en ese nivel. De este modo se mejoraría el proceso de estimación confiando en que el área pequeña redefinida permitirá obtener información similar a la del área original. Este enfoque, conveniente en la práctica, puede ser objetable, pues supondría que la condición de homogeneidad del dominio redefinido también sería válida para el dominio pequeño original.

Si el país se propone modificar los límites de los dominios administrativos, de modo que también se vean afectados los límites de las áreas pequeñas en un futuro cercano, podría plantearse un problema de redefinición. Esa posibilidad puede tenerse en cuenta en el diseño, redefiniendo áreas unitarias estables —no afectadas por el cambio— y seleccionables, que más adelante se podrían agrupar para redefinir el dominio pequeño.

Aunque los problemas que acabamos de enumerar no son exhaustivos o exclusivos, pueden interactuar entre sí impidiendo, de hecho, una solución al problema de la redefinición de las áreas pequeñas —especialmente si el tamaño de la muestra no se diseñó inicialmente para realizar estimaciones de dominios pequeños. Si en el proceso de diseño no se tuvo en cuenta la necesidad de realizar una estimación de área pequeña, en una etapa posterior del análisis puede ser necesario desarrollar procedimientos de estimación para los dominios pequeños.

3. PROCEDIMIENTOS DE ESTIMACIÓN

Lo que sigue es una presentación de algunos enfoques utilizados para la estimación de dominios pequeños. No es intención del autor pasar revista exhaustivamente a los procedimientos de estimación de áreas pequeñas. Una buena descripción de tales procedimientos puede encontrarse en *Small area statistics: An international symposium* (Platek y otros, 1987).

Estimación directa. Como ya se ha analizado, la estimación directa de áreas pequeñas se basa en los datos de encuestas provenientes solamente del dominio del área pequeña. Un procedimiento aceptable en la mayoría de los casos se basa en el diseño, especialmente si aquel es prácticamente insesgado. Sin embargo, ese tipo de diseño no es satisfactorio cuando el tamaño de la muestra es pequeño, es decir, cuando los errores de muestreo de las estimaciones son grandes. Es preciso destacar enfáticamente que la estimación directa *sólo debe ser utilizada* cuando el tamaño de muestra es suficientemente grande como para permitir una estimación confiable.

La fórmula general para la estimación del cociente de una estimación directa viene dada por:

$$r = \frac{\sum_i w_i y_i}{\sum_i w_i x_i} \quad [1]$$

donde

w_i es el factor de ponderación de diseño para la unidad i ;
 y_i es el valor de y para la unidad i ; y
 x_i es el valor de x para la unidad i .

Si X es una variable de enumeración, el estimador del cociente se reduce a la media ponderada de Y , y puede expresarse como:

$$r = \frac{\sum_i w_i y_i}{\sum_i w_i} \quad [2]$$

Cuando se conoce el tamaño N de la población, la estimación directa de una expansión total viene dada por:

$$\hat{Y} = N \frac{\sum_i w_i y_i}{\sum_i w_i} \quad [3]$$

En general, si se conoce el valor total de la variable X para todos los estratos, podemos definir el valor total ampliado como:

$$\hat{Y} = \sum_h X_h r_h \quad [4]$$

donde

X_h es el valor total conocido de la variable auxiliar en los estratos h .

Puede obtenerse una estimación directa de la regresión como estimador aproximado de diseño no sesgado. Viene dada por:

$$Y_{reg.a}^* = Y_a^* + \beta_a^* (X_a - X_a^*) \quad [5]$$

donde:

$Y_{reg.a}^*$ es el valor total expandido predicho en el dominio pequeño a ;

β_a es el coeficiente de la regresión en el dominio pequeño a ;

donde β_a se calcula como:

$$\beta_a^* = \frac{\sum_{i \in s_a} v_i^{-1} \omega_i Y_i X_i}{\sum_{i \in s_a} v_i^{-1} \omega_i X_i X_i} \quad [6]$$

donde el valor del subíndice i corresponde a la i -ésima unidad en el dominio de la muestra a ;

X_a es el valor total X para el dominio pequeño a ;

Y_a^* es el valor de expansión de Y en el dominio pequeño a ; y

X_a^* es el valor de expansión de X en el dominio pequeño a .

Se puede obtener un estimador directo aproximado modificado sustituyendo el valor del coeficiente para un valor calculado para toda la muestra, que podemos llamar $Y_{\text{reg-mo},i}^*$

Estimación indirecta. Este tipo de estimación está relacionado con el procedimiento de estimación sintético, que se basa en la hipótesis de que el área pequeña es similar en algún sentido a otra área, a menudo un área más grande que la contiene. Aunque este tipo de estimación tendría una varianza más pequeña, puede resultar sesgada si no se cumple la hipótesis.

$$Y_i = \sum_j \left(\frac{N_{ij}}{N_i} \times y_j \right) \quad [7]$$

donde

Y_i = el valor total estimado o el valor del cociente para la i -ésima área pequeña;

y_j = el valor total observado o el valor del cociente para la j -ésima categoría auxiliar;

N_{ij} = el número estimado de unidades en la j -ésima categoría auxiliar, para la i -ésima área, tal como se observa en la fuente externa;

N_i = el número estimado de unidades en la i -ésima área, tal como se observa en la fuente externa, y

N_{ij}/N_i = las ponderaciones de ajuste.

Entre los métodos indirectos para obtener estimaciones de área pequeña cabe mencionar los que se basan en el uso de un modelo y, en especial, los que utilizan una regresión. Generalmente, la ventaja de este último enfoque es que logra cierta estabilidad en el modelo usando información cercana o circundante al dominio pequeño.

Aliaga y Muhuri (1994) han desarrollado un procedimiento de modelización (con varias características deseables) que permite estimar la tasa de prevalencia del uso de anticonceptivos como variable dependiente relacionada con once variables independientes:

- Cada variable es categórica a nivel individual (dos o más categorías).
- Cada variable fue transformada en una variable binaria; la dicotomización se basó en la *maximización de la capacidad de discriminación sobre el uso de anticonceptivos*. Se creó una variable continua (proporciones o porcentajes) para la unidad del nivel inmediato superior (conglomerado) para cada variable, *es decir, cada una tenía la misma escala y era continua*. Alrededor de cada dominio del área pequeña, se crearon grupos con áreas vecinas, en forma circular, de modo que cada grupo tuviera alrededor de 30 conglomerados. Cada grupo comprende de dos a cinco dominios pequeños, *es decir, que si se circunda el dominio pequeño en forma circular se obtendrá información sobre todas las áreas vecinas que puedan tener un comportamiento similar*.
- Se crearon modelos de regresión a nivel de conglomerado para cada grupo que tuviera *estabilidad* (cada grupo tenía alrededor de 30 conglomerados).
- La distribución del cociente entre la media muestral y el error de muestreo estimado se puede aproximar mediante la distribución normal.
- Se estimó la tasa de prevalencia para cada grupo que circundaba el área de dominio pequeño usando los valores pronosticados correspondientes, a nivel de conglomerado.
- La estimación final para el dominio pequeño se obtenía promediando las estimaciones de todos los grupos.

Cualquiera sea el modelo utilizado, pueden plantearse varios problemas o restricciones. Por ejemplo, la estimación se basa en las hipótesis del modelo, que pueden no cumplirse. Asimismo, la estimación se debe restringir a determinados tipos de variables.

Estimadores combinados. Cuando se han calculado al menos dos tipos de estimaciones distintas, es posible también calcular una estimación que sea una combinación lineal de ambas. Por ejemplo, si hay una estimación directa y otra indirecta, la nueva estimación se puede calcular como sigue:

$$r_c = \alpha r_d + (1 - \alpha) r_i \quad [7]$$

donde

α tiene un valor entre 0 y 1;
 r_c es el estimador combinado;
 r_d es el estimador directo; y
 r_i es el estimador indirecto.

El valor de α puede considerarse un factor de ponderación asignado al estimador directo. Por lo tanto, el término $(1 - \alpha)$ es la ponderación para la estimación indirecta. Entre las distintas soluciones posibles, una de ellas se obtiene cuando α es proporcional a la precisión de la estimación directa (la inversa del error de muestreo), es decir, $(1 - \alpha)$ es proporcional a la precisión de la estimación indirecta. La estimación combinada siempre será relativamente más cercana a la que tiene más precisión y tendrá un valor intermedio entre ambas.

Otros procedimientos. Hay varios otros procedimientos de estimación basados en principios estadísticos, como el enfoque bayesiano, cuya forma es similar al estimador combinado, siendo α un valor probabilístico. Si se dispone de información *a priori* para estimar una distribución de ese tipo (α), puede calcularse una estimación razonable para el dominio pequeño. Aunque este principio es bastante interesante, una de las dificultades principales es la necesidad de establecer una distribución *a priori*. El uso del enfoque bayesiano mediante el muestreo de Gibbs (generando una secuencia de Gibbs de variables aleatorias de distribuciones condicionales, que exige un uso intensivo de sistemas computarizados) puede servir para superar este problema empírico generando la distribución *a posteriori* de la estimación. El muestreo de Gibbs plantea problemas importantes en cuanto a la puesta en práctica y la comparación de las diversas maneras de extraer información de la secuencia de Gibbs.

4. APLICACIONES

Cualquier procedimiento de estimación que se base en el diseño de la muestra es más aceptable que un procedimiento que no tenga en cuenta el diseño. De éstos, cualquier procedimiento que se base en el diseño y sea insesgado (o aproximadamente insesgado) es más aceptable que un procedimiento sesgado. Para la estimación directa se debe utilizar, siempre que sea posible, un procedimiento basado en el diseño e insesgado (o aproximadamente insesgado). Sin embargo, a medida que el

nivel del dominio se hace más pequeño, este tipo de enfoque pierde fiabilidad. Aunque existen restricciones para los dominios pequeños, el procedimiento de estimación directa da una idea general de la estimación para esos dominios críticos (un tamaño de muestra pequeño, es decir, un número reducido de conglomerados).

En las EDS, para cualquier dominio que tenga 30 o más conglomerados, el procedimiento de estimación directa permitirá obtener una estimación homogénea y fiable y la probabilidad de que sea inadecuada será baja (5%). Para los dominios que tienen entre 10 y 30 conglomerados, el procedimiento de estimación directa puede ser útil si se combina con otros resultados. Para los dominios de 10 conglomerados o menos, el procedimiento de estimación directa es muy poco confiable. Tales estimaciones no deben utilizarse directamente, sino que se pueden combinar con otros tipos de estimación.

En el caso de la estimación provincial de la República Dominicana (véase el cuadro 1), y la estimación distrital de Kenia (véase el cuadro 2), se determinó que:

1. En las estimaciones del Distrito Nacional y la provincia de Santiago en la República Dominicana y del distrito de Nairobi en Kenia, usando alrededor de 30 conglomerados o más, se obtuvieron valores absolutamente coherentes utilizando distintos enfoques. Por lo tanto, para la prevalencia del uso de anticonceptivos la estimación directa debe usarse sin más.

2. Para cada una de las provincias de la República Dominicana y cada uno de los distritos de Kenia, en que las muestras tienen entre 10 y 20 conglomerados, el enfoque de la regresión arroja resultados absolutamente coherentes con las estimaciones combinadas. Por lo tanto, cuando la diferencia absoluta entre la regresión y las estimaciones combinadas es menor que el 10% de la regresión, la alternativa razonable recomendada es la estimación por regresión. Quizás deba usarse la estimación combinada cuando la diferencia supere el 10%.

3. Para cada una de las provincias de la República Dominicana y cada uno de los distritos de Kenia, con una muestra de 10 conglomerados o menos, el valor medio entre la regresión y las estimaciones sintéticas permitiría obtener una estimación más razonable.

Cuadro 1
**ESTIMACIONES DIRECTA, INDIRECTA, POR REGRESIÓN Y COMBINADA
 DE LA PREVALENCIA DEL USO DE ANTICONCEPTIVOS POR LAS
 MUJERES CASADAS: REPÚBLICA DOMINICANA, 1991**

Región y provincia	Estimación directa ^a			Estimación por regresión		
	Valor	Error muestral	Estimación indirecta ^b	Valor	Error muestral	Estimación combinada
Región 0	60.7					
Distrito Nacional	60.7	2.0	60.6	61.7	1.8	61.2
Región I	50.8					
Peravia	48.7	5.8	51.5	54.6	3.0	52.6
San Cristóbal	55.9	3.7	53.7	52.3	2.6	53.8
Monte Plata	43.1	3.9	51.4	52.3	3.5	47.9
Región II	61.0					
Santiago	62.1	3.5	62.1	60.7	4.4	61.4
Puerto Plata	62.8	5.3	62.0	58.6	2.1	59.8
La Vega	59.4	4.4	61.5	58.0	2.0	58.4
Españillat	56.2	9.1	62.0	57.3	3.5	57.0
Monseñor Nouel	61.7	1.5	61.2	56.7	2.0	59.6
Región III	57.4					
Salcedo	53.4	5.6	57.1	56.1	3.4	55.1
Duarte	57.2	4.5	56.1	55.3	2.3	55.9
María T. Sánchez	65.9	7.5	57.8	57.9	1.6	59.3
Samaná	44.6	10.2	55.0	57.9	1.6	56.1
Sánchez Ramírez	56.7	8.2	58.5	53.7	4.3	54.7
Región IV	47.1					
Barahona	51.9	2.7	47.9	50.0	3.3	51.0
Pedernales	48.3	9.6	45.1	50.0	2.4	49.7
Bahoruco	38.8	6.5	47.7	43.4	4.8	41.4
Independencia	44.8	9.1	46.2	49.1	2.5	48.2
Región V	50.6					
La Romana	44.0	7.9	50.1	50.1	4.0	48.0
La Altagracia	57.5	3.8	51.4	49.8	4.1	53.8
El Seybo	46.5	9.7	51.0	49.5	3.4	48.7
San P. de Marcorís	54.2	3.5	52.3	50.8	3.3	52.5
Hato Mayor	58.0	7.0	51.9	49.6	2.7	51.9
Región VI	39.7					
San Juan	34.1	8.2	39.5	41.5	6.6	38.2
Azua	52.3	4.5	39.0	52.7	3.2	52.5
Elías Piña	32.1	5.9	39.9	40.5	8.0	35.7
Región VII	58.6					
Valverde	53.9	2.4	59.1	58.4	2.9	55.9
Santiago Rodríguez	73.6	7.4	62.7	55.1	4.9	62.5
Dajabón	51.8	5.6	56.5	55.8	6.3	53.7
Montecristi	62.8	4.4	56.3	59.3	2.6	60.6

^a Datos de la EDS de 1991.

^b Datos de la EDS de 1991 y datos de la Encuesta de Hogares Ampliada de 1991.

Cuadro 2
ESTIMACIONES DIRECTA, INDIRECTA, POR REGRESIÓN Y COMBINADA
DE LA PREVALENCIA DEL USO DE ANTICONCEPTIVOS POR LAS
MUJERES CASADAS: KENIA, 1992

Región y provincia	Estimación directa ^a			Estimación por regresión		
	Valor	Error muestral	Estimación indirecta ^b	Valor	Error muestral	Estimación combinada
Nairobi						
Nairobi	33.5	2.1	34.9	38.5	3.6	35.3
Central	39.5					
Kiambu	37.3		38.1	37.9	3.2	
Kirinyaga	54.2	3.2	38.5	44.4	2.0	48.2
Muranga	32.1	5.4	39.2	38.8	2.5	36.7
Nyandarua	39.2		38.4	40.4	3.2	
Nyeri	40.8	3.2	39.0	43.3	2.8	42.1
Costera	18.1					
Kilifi	10.8	1.7	17.2	15.2	2.6	12.5
Kwale	16.2		17.0	17.3	4.3	
Mombasa	24.5	3.6	17.0	15.0	2.0	18.4
Taita	30.7		18.1	29.0	1.9	
Oriental	40.2					
Embu	47.2		40.0	42.1	2.4	
Kitui	41.3		39.6	41.6	2.7	
Machakos	40.4	2.9	40.0	41.0	1.7	40.8
Meru	36.3	3.5	39.9	43.4	3.8	39.7
Nyanza	13.8					
Kisii	21.5	2.6	13.5	16.9	1.4	18.5
Kisumu	17.8	2.8	13.3	14.8	1.3	15.8
Siaya	8.5	2.2	13.7	10.1	1.5	9.5
Nyanza meridional	6.1	1.4	13.2	11.2	1.0	9.1
Valle del Rift	29.6					
Baringo	12.5		29.4	27.1	3.1	
Elgeyo Marak	16.7		29.4	15.0	2.8	
Kajiado	52.0		28.8	41.5	4.6	
Kericho	24.1	2.6	29.1	19.3	1.6	21.1
Laikipia	68.5		29.4	45.0	4.5	
Nakuru	47.2	5.4	29.2	34.0	5.2	40.5
Nandi	16.7		29.3	18.0	2.2	
Narok	22.6		28.8	23.3	4.5	
Trans Nzoia	28.9		29.2	15.1	2.5	
Uasin Gishu	13.4	2.5	29.2	15.7	2.3	14.6
Pokot Occidental	0.0		29.1	n.d.	n.d.	
Occidental	13.7					
Bungoma	9.4	1.5	13.2	11.8	1.8	10.5
Busia	16.1		13.2	11.7	2.2	
Kakamega	14.9	2.0	n.d.	14.4	1.7	14.6

^a Datos de la EDS de 1989.

^b Datos de la EDS de 1989 y censo de población de 1990.

Nota: Los errores de muestreo de las estimaciones directas sólo se calcularon para los distritos de más de 9 unidades primarias de muestreo (UPM).

n.d.: no se dispone de datos.

5. CONCLUSIONES

Diversos estudios indican que no existe un enfoque único para obtener estimaciones para los dominios pequeños cuyos resultados sean satisfactorios en todos los casos. Las propiedades de una estimación determinada dependerán del tamaño y la homogeneidad del dominio de área pequeña, de la disponibilidad de información externa, del carácter continuo o discreto de la variable de interés, de su idoneidad para explicar cualquier relación hipotética y del nivel de capacidad para discriminar las ocurrencias de un evento. Un enfoque adecuado para una situación determinada puede no serlo para otra.

Siempre puede considerarse la posibilidad de utilizar otro tipo de estimaciones (en forma individual o combinada). Para los dominios que tienen un número crítico de conglomerados, la estimación definitiva sólo debe darse después de realizar una evaluación del tipo de variable y de las condiciones de entorno del dominio pequeño.

Los procedimientos combinados de estimación deben utilizarse después de realizar una evaluación cuidadosa y no en forma automática, es decir, usando criterios puramente matemáticos. La información disponible en el sector público puede haberse utilizado con fines muy diversos. Por lo tanto, la combinación de distintas fuentes de información debe hacerse con prudencia para evitar controversias.

BIBLIOGRAFÍA

- Aliaga, Alfredo y Thanh Le (1991), "Methodology for small area estimation with DHS samples", *Institute for Resource Development. Demographic and Health Surveys World Conference; proceedings*, vol. 1, Columbia, Maryland, IRD/Macro International Inc.
- Aliaga, Alfredo y Pradip K. Muhuri (1994), "Methods of estimating contraceptive prevalence rates for small areas: applications in the Dominican Republic and Kenya", DHS Methodological Reports, N° 3, Calverton, Maryland, Macro International Inc.
- González, M.E. (1973), "Use and evaluation of synthetic estimates", *Proceedings of the American Statistical Association (Social Statistics Section)*, Washington, D.C., Asociación Americana de Estadística.
- Platek, R. y otros (comps.) (1987), *Small area statistics: An international symposium*, Nueva York, Wiley & Sons.