Argote Cusi, Milenka Linneth
Estimación de la distribución estadística de la tasa global de fecundidad
Papeles de Población, vol. 13, núm. 54, octubre-diciembre, 2007, pp. 87-113
Universidad Autónoma del Estado de México
Toluca, México

Available in: http://www.redalyc.org/articulo.oa?id=11205405
Estimation of the statistical distribution of the total fertility rate

Milenka Linneth Argote Cusi
Centro Nacional de Prevención y Atención al VIH/SIDA e ITS

Abstract
Estimation of the statistical distribution of the total fertility rate

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.

Introduction

Populations are complex systems whose study is of great interest for social sciences. The complex systems are characterized by the quantity of elements and their relations, which are in continuous interchange in time...
to produce a whole larger than the addition of its parts. One form of studying the complexity of the populations is by means of the reconstruction of indicators. When an indicator is calculated from the total population, it receives the name of parameter, whereas if it is obtained from a sample, it is called statistics (Efron and Tibshirani, 1993). Generally, the parameters of total population are not known due to the cost of their obtaining, instead, samples are drawn to, from which we make estimations of the parameters. According to the theory of sampling, since it is only possible to obtain a sample of all the possibilities from the total population, the estimations we make from it are subject to sampling errors and non-sampling ones. Non-sampling errors in a fertility survey are due to the lack of coverage of all of the selected women, errors in the formulation of questions and in the registration of the answers, confusion in the interpretation of the questions, memory problems and codification or processing errors. Sampling error, which is measured by means of standard error, is a measure of the variation of the estimator of a parameter in all of the possible samples (Cochran, 1977). In practice, it is not possible to obtain all the possible samples of a population, in order to solve this problem parametric statistics has constructed a fundamental theoretical base. Large numbers law and the central limit theory are two basic theorems of the traditional statistic inference that allow us to suppose a normal distribution of estimators of totals and means, yet, what is the statistic distribution of an estimator of ratio, such as the global fertility rate? How can I estimate the sampling error of an estimator different from a total or a mean?

**Parametric inference vs. non-parametric inference**

Parametric inference allows us to construct distribution for averages or proportions with base on the central limit theorem (TLC) and the large numbers law (LNG), which assume a normal distribution of the estimator when the size of the sample grows infinitely. In the case of a rate (the quotient of the number of events occurred between time of risk exposure), it is a more complex estimator than an average or a total.

According to the revised theory, there are not formulae to estimate in an analytical form the intervals of a rate, however, other methods of variance estimation have been used for non-linear functions (Taylor series). On the other
side, we can add that there is a sample with complex design to estimate TFR. Considering the peculiar characteristics of the estimator, the following question arises: what is the statistical distribution by sampling of TFR?

The present work states the alternative hypothesis that the TFR distribution, a peculiar estimator, is different from the normal before the void hypothesis that it is indeed normal, assuming that TGF is similar to a ratio of means, hence, TLC and LNG are applicable. We draw to the bootstrapping technique, a non-parametrical statistic technique, to generate the statistic distribution by sampling of the characteristic of interest and, from it, evaluate whether it has a normal behavior (Cochran, 1977).

The coherence and un-biasing of estimation are a crucial team in statistic inference. If the estimations that are made from the total population samples cannot be adequately interpreted from the statistical point of view, affirmations that are not true for total population can be produced and this is the situation which frequently the social researcher is exposed to. Nonetheless, what is due the interest in the estimator of the global fertility rate to?

**Importance of fertility**

The impact of fertility on the population’s social and economic structure makes this demographic variable a priority in the sphere of population policies. Since fertility is an important component of demographic dynamics, it is essential to have adequate fertility measures, valid interpretations of the trends and differentials, and reasonable conjectures on their future direction (Campbell, 1983). It is worth mentioning that the impact of fertility is not always the one with the most relevance in different populations. It can occur that in contexts suffering a disease, such as in Africa, mortality is the demographic phenomenon that defines population’s growth. On the other side, nowadays, the volume of migrations has increased so much in the last decades that in countries such as Mexico, they have a heavy impact on demographic structure. In Bolivia, the impact of fertility on demographic growth is heavier than other population-related phenomena. In an exercise of population projection in different sets of fertility, mortality and migration, fertility’s variations considerably modify population’s structure. To analyze the levels and tendencies of fertility in Bolivia, its historic characteristics must be considered, which are reflected on a largely indigenous population with poor schooling levels, widely illiterate and precarious health conditions, mainly in
rural areas. A thorough analysis at more disaggregated levels leads to think that this stability would have turned out into the cancelation of two opposed tendencies: a declination of fertility in urban areas and elevation in the rural ones (Carafa et al., 1983). In 2000, TFR is reduced to four and in 2003 a preliminary TFR of 3.8 by woman is estimated. Even if total fertility has decreased in 2003 in the rural area there is a TFR of 5.5 children per woman. Due to this behavior, according to CEPAL, Bolivia changes from ‘incipient transition’ to ‘moderate transition’, where birth and mortality rates are still high compared to the rest of Latin American countries.

Hence, it is of great interest to estimate fertility’s indicators that give an account of its behavior. Fertility measures are numerous because of the interest in specific demographic groups as well as in the availability of information. The three main sources of information are civil registry, censuses and surveys. In some countries civil registry is reliable, however in countries such as Bolivia, it is not so reliable, so surveys are drawn to in order to correct deficiencies.

The measures most used by fertility are specific fertility rates (TEF) by quinquennial age groups, defined by the quotient of births occurred at women of a group age \( \text{Nac}_{x,x+5} \) divided by the years-person lived in exposure to women’s risk \( \text{Temujeres}_{x,x+5} \) of the same age group, and TFR (the addition of the specific fertility rates multiplied by five) that represents the number of children by woman at the end of their reproductive life, under the supposition that along their life they fertility will be present. We can also say that TFR is a linear combination of TEF or, from the statistical point of view, it is a linear combination of ratios.

\[
TEF_{x,x+5}^{t,t+1} = \frac{\text{Nac}_{x,x+5}^{t,t+1}}{\text{Temujeres}_{x,x+5}^{t,t+1}}
\]  

(1) \[
TGF^{t,t+1} = 5 \sum_{i=1}^{7} TEF_i^{t,t+1}
\]  

(2)

\[
TGF^{t,t+1} = 5 \sum_{i=1}^{7} TEF_i^{t,t+1}
\]  

(2)
Data and methods

The present research has used data from the 1998 National Survey on Demography and Health (Encuesta Nacional de Demografía y Salud de 1998, Endsa), which is part of the program of Surveys on Demography and Health (Encuestas de Demografía y Salud, DHS) that Macro International Inc. carries out in several developing countries. Endsa 98 has a probabilistic national sample, which is stratified and bi-phased. The stratification was carried out at the level of different geographic subdivisions: geographic regions (highland, valley, plains), by departments in each region and marginalization degree of the municipalities in each department, according to their poverty levels and abiding zone (urban-rural). Fieldwork began on March 23rd 1998, in the region of Los Llanos and on the 26th in other two regions; it ended on September 15th. In a first phase, the divisions called areas of censal enumeration were considered as the primary units of sampling (UPM) out of which 823 were selected in the country. In a second phase, the households listed in the selected UPM were established as the secondary units of sampling (USM). For this work’s ends, the units of analysis are women in fertile ages and their children’s births in the selected households (Endsa, 1998).

To manage databases and implement the resampling algorithm1 the statistical program Stata version 8.0 was used, as it is oriented to surveys by sampling and also offers functions for non-parametric inference.

Estimation of total fertility rate

Let us begin from the theoretical definition of rate: the quotient of the number of occurred events divided by the time of the individuals’ exposure to experience the event. Making and analogy, this definition is similar to a poisson process, where we measure for instance, the number of missiles that hit an area or the number of times a bus reaches a bus stop in determinate time, etc. The relation of entire units and a continuous value, established through a ratio, is what gives a rate the complexity of representation and interpretation. In Demography there are several methods to estimate TFR, among them the direct and indirect ones

1 Sequence of steps to obtain a result
Estimation of the statistical distribution of the total fertility rate / M. Argote

(Campbell, 1983). The methods of TFR estimation applied are summarized in figures 1 and 2.

One of the main factors which had to be controlled was the right classification of births and time of exposure in the denominator, by quinquennial groups of the mother’s age at the time of x children’s birth. Since time of exposure is a continuous variable, it is possible that the event occurs in the limits of the intervals of the quinquennial age groups, providing a time of exposure to adjacent groups.

Exposure time is measured in months and once the classification by quinquennial age groups in the three last years (in the last 36 months) prior to the survey is controlled, the base is pondered to have as a result a tabulation with the following columns: quinquennial group (categories one to seven), births in each group and exposure time women in each age group provide. From said tabulation the calculations of the specific fertility rates and the global fertility rate were made.

The procedure to calculate TFR becomes an independent and integral module on its own, which is later retaken (referred to) in the resampling execution program.

Background of variance estimation

Several methods to estimate variances have been put forward in the literature for more complex functions than the total and medians. The most used is the method for Taylor series that is applied to estimators defined by non-linear functions, however, since several derivatives are calculated it can be tedious and complex to apply (Sul et al., 1989). On the other side, the methods of random groups use subsamples, trying as much as possible to preserve the design in the original sample. Indeed, the way to obtain the size of said subsamples is a problem in complex designs (Korn and Graubard, 1999).

A new alternative is given by the methods of resampling and replicas. The advantage of these methods is that the units of observation are kept together inside a primary unit while it the replicas are constructed, which preserves dependency between the observation units within the same primary unit (Setter, 1992a).

Because of this they are applicable to different stratified, poly-phased and probabilistic designs that have a data distribution that is not identically distributed
FIGURE 1
CALCULATION OF TFR’S NUMERATOR

Classification of births according to the age of the mother in the three previous surveys (numerator)

Ponder the base of women and children

Calculate the age of the mother at the time of birth

Classify births in the mother’s quinquennial group

Select the births occurred in the 36 months previous to the survey

Source: own elaboration
Classification of exposure time of women until birth (denominator)

Calculate the age of women at the time of the interview

According to the previous age, classify women in a quinquennial group (x + 5)

Calculate the exposure time in months the woman has in the current group (x + 5) and in the previous group (x) at the time of the child’s birth

Classify the exposure times by women’s quinquennial age groups

Add and ponder the exposure times

Calculate TEF and TFR

Source: Own elaboration
inside a primary unit while it the replicas are constructed, which preserves dependency between the observation units within the same primary unit (Setter, 1992a).

Because of this they are applicable to different stratified, poly-phased and probabilistic designs that have a data distribution that is not identically distributed (Lohr, 2000). The standard resampling method, a method of simple random sampling with replacement, was proposed by Efron and Tibshirani in 1979; the interest was mainly in studying standard error and bias in function of sampling variability. These researchers developed the basic concepts of this technique from the plug-in principle.2 In 1992, Sitter applies bootstrap with replacement (BWR) for stratified random samples and observes that the estimator of variance, compared to Cochran’s formula (1977) for the same model, was biased. To solve this problem Sitter proposed introducing a correction factor that adjusts the estimator, yet only for the case of a sample with stratum. In this case the estimated variance with resampling is consistent and unbiased.

The advances in the technique’s development are related to the desirable priorities of the estimators (un-biasing and consistency) which are distorted in complex samples. In this line, resampling methods to estimate variance and confidence intervals where the sampling occurs without replacement have been proposed; namely: Jackknife, bootstrap without replacement (BWO), Bootstrap

2 Let us take into account a random sample of size n with a distribution of probabilities $F$.

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

The function of empirical distribution $\hat{F}$ assigns each realization of the sample a probability $1/n$ for each value of $x_i, i = 1, 2, ..., n$.

When the distribution of probabilities $F$ is known, in the case of a census, finding the variance $\sigma^2$ is not difficult. We do not usually have a census; then we draw to statistical inference that enables us to infer properties of $F$ from a random sample $X$. “Then $\theta$ is a parameter of $F$, while $\hat{\theta}$ is a statistics based on $X$. Hence the plug-in estimator of the parameter

$\theta = t(F)$ is defined as $\hat{\theta} = t(\hat{F})$
stratified designs (Efron, 1982). Another method consists in re-scaling (adjust the mediator) the standard method when the estimator is a non-linear function of the means (Rao and Wu, 1988); in this method, the algorithm is applied with a selected sample size $m_k$ not necessarily the size of $n_h$ (size of the sample in stratum $h$) and the values of the bootstrap are appropriately re-scaled so as to have unbiased estimators in the linear case. Since each datum has to be re-scaled in each bootstrap, this process can be complicated in more complex samples. Whereas BWO and BWR are only applicable to simple designs, the version of bootstrapping with re-scaling is extended to more complex designs for functions of means and which are more intensive and more difficult to use in a computing manner.

**Applications of bootstrapping methods**

The researches carried out by the aforementioned authors have used hypothetical finite populations to experiment with the different bootstrapping methods. Attention has been paid to the study of estimations which are linear functions of means; nevertheless, there have been also simulations of the case of estimators of ratio, regression coefficients, correlation coefficients and median. It is concluded that MMB and BWO have an acceptable performance compared to other methods (Sitter, 1992b). In 1994 MMB is applied to a 3-phase stratified design to estimate a ratio with acceptable results. Despite more research is necessary, there is evidence that the estimation of the confidence intervals is appropriated according to the *plug-in* principle (Robb, 1994). In Australia several bootstrap methods were applied for a ratio estimator ($Y$: income from the sale of a product; $X$: quantity of this product) where high biases that can be attributed to the sort of selected estimator were found, however, it is also suggested applying more sophisticated bootstrapping methods for bias correction (Davidson and MacKinnon, 1993).

Bootstrapping methods have been mainly applied in the economics’ field. An important example is the interest in studying the statistical significance of the changes in the indicators of inequality and welfare through Gini’s index. If the income surveys were always based on the same households, temporary variations in the inequality and welfare indicators would really reflect changes in income distribution. In reality, it is more frequent to find surveys applied to
Bootstrapping methods have been mainly applied in the economics’ field. An important example is the interest in studying the statistical significance of the changes in the indicators of inequality and welfare through Gini’s index. If the income surveys were always based on the same households, temporary variations in the inequality and welfare indicators would really reflect changes in income distribution. In reality, it is more frequent to find surveys applied to different households from one period to another. Thus, the differences in these indicators could be simply attributed to the fact that the sample changed, and not to real variations in inequality of income. For instance, Gini’s coefficient computed in year $t$ can be superior to that of the year $t-1$, simply due to sampling phenomena (whether there had been any change in income distribution), so the conclusion that the distribution has become more inequitable is not necessarily correct (Gasparini and Sosa, 1998).

**Characteristics of the bootstrapping model utilized**

The theoretical results from the bootstrapping technique were developed for different statistical areas of the surveys by sampling. The extension of bootstrap to complex samples is recent and one of its possible uses is the estimation of statistical distribution by sampling an estimator of ratio such as TFR.

In the present work $\tau$ is an estimator of Bolivia’s total fertility rate in 1998. The only information available comes from a sample with the structure of hierarchical data, which increases the complexity of the estimations. The object of interest consists in obtaining a measure of dispersion for $F(\tau)$ and a confidence interval for the punctual estimation of $\tau$. The analytical evaluation of these statistics requires being aware of $F(\tau)$ distribution. It is true indeed that the distribution by sampling of $F$ is not known and the derivation of the standard error(s) and the confidence interval are analytically complex; the bootstrap method, as it has been stated, allows us to approximate $F(\tau)$ using the empirical distribution of the sample $F(\tau)$.

Bootstrap with replacement is applied considering in the first phase UPM and in the second USM. Bootstrap is less efficient than bootstrap with replacement, nevertheless it is used as it provides facility to choose and analyze the samples. Bootstrap can be applied into sampling by conglomerate in two phases, Robb (1994) finds the contribution of the variance in the second phase will be
The bootstrap model applied in this work is similar to the resampling with replacement stated by McCarthy and Snowden en 1985, this time differenced by the fact that the $m_h$ subsamples’ size in the strata is the same as the size $n_h$ of the sample in the strata. The method consists in obtaining a simple random sample with replacement of size $n_h$ independently in each UPM (censal unit), then, in the second phase, making the same with USM (households).

The mediators are directly applied in the punctual estimation of TFR from the simple random sample of women and children, following the sample’s design; sampling weights have the information to determine standard errors, as the sample’s design is reflected upon them (Lohr, 2000).

Variance, standard error and bias are automatically calculated from the distribution by sampling of TFR, as the bootstrap technique indicates (figure 3). The percentile method is applied to calculate the confidence intervals at 95 percent, for the sampling statistical distribution of TFR obtained by resampling is similar to normal distribution according to the Kolmogorov Smirnov test (0.901438 > 0.05). A good estimation of the confidence interval is obtained in 1000 replicas (Efron and Tibshirani, 1993), therefore the results and conclusions are presented taking this reference into account.

**Suppositions**

A first supposition on which the bootstrap technique is supported refers to the fact of considering the sample as though it was the total population. This to say, it is considered that the distribution of the observed data (sample) represents a good estimation of the population’s distribution. Said supposition is founded on the asymptotic theory that is based on the analysis of convergence in probability and convergence in distribution. Be $X_n$ the random variable, said variable converges in probability with a constant $c$ when the size of the sample tends to infinite, the probability that the difference between the sampling value and the real one will be greater than a permissible value $\varepsilon$ tends to zero, for any $\varepsilon > 0$. This means that progressively there is a lower probability for $X_n$ to be different from $c$, to the extent $n$, the sample’s size, increases.

---

1 Kolmogorov Smirnov test (KS) of a sample is a test of goodness of fit. It measures the degree of similarity between the distribution from the observed data and a theoretical distribution (Siegel, 1956: 47).
1. Obtain a sample with replacement \( y_{i}^* \) of the original sample \( y_{i} \) independently in each stratum.

2. Calculate \( h_{ni} \) where \( h_{ni} = \theta^*(e^*) \).

3. Repeat step 1 a large number of times, \( B \), to obtain \( \theta_1^*, \theta_2^*, ..., \theta_B^* \).

4. Estimate the variance

\[
\text{var}(\hat{\theta}) = \frac{1}{B-1} \sum_{b=1}^{B} (\hat{\theta}_b^* - \hat{\theta})^2
\]

5. Estimate the error

\[
s_{\theta}^2 = \frac{1}{B-1} \sum_{b=1}^{B} (\theta^*_b - \hat{\theta})^2
\]

6. Estimate bias

\[
\text{bias} = \hat{\theta} - \theta
\]

7. Estimate the confidence intervals at 95%

\[
\text{inf}_{\alpha/2} = \hat{\theta} - z_{1-\alpha/2} \cdot \frac{s_{\theta}}{\sqrt{B}} \\
\text{sup}_{\alpha/2} = \hat{\theta} + z_{1-\alpha/2} \cdot \frac{s_{\theta}}{\sqrt{B}}
\]

Separately, \( X_n \) converges in a distribution to a random variable \( X \) with accumulated distribution function \( F(X) \), if the difference between the sampling distribution and the real distribution tends to zero to the extent the size of the sampling does, for all the points of continuity of \( F(X) \). These basic concepts from the asymptotic theory (Davidson and MacKinnon, 1993) are directly related to the plug-in principle (Efron and Tibshirani, 1993).

In a second instance it is supposed that the estimator of TFR used in this research is good. For Chou (1977), there are infinite estimators of a parameter; trying all of them out and finding the best is impossible. In this case, the function used to estimate the total fertility rate we considered a “better” estimator, since the theoretical definition of a rate is applied, this is to say, it considers that in the denominator the time of risk exposure of women in reproductive ages instead of the average population of women in reproductive ages that is commonly used.

In relation to the sample, it is supposed that the basic sample that is used to bootstrapping is representative. This concept is crucial in all statistical reasoning and it is understood as the fact that the sample should be similar to the population. It is difficult to secure the representativeness in small samples which have not been obtained with random processes. Nonetheless, the law of the large numbers allows us to expect representative samples (Méndez, 2004). The Endsa sample which comes from stratified, bi-phased sampling has a high probability of being representative.

**Results**

*Parametric estimation of the confidence intervals*

As a reference point in the first place the method of traditional statistical inference was applied. The construction of the estimators implies the estimation of totals to the different levels of disaggregation. Finally, the complexity is summarized as the addition of the multiplication of the pondering factors of the different levels, by the TFR (\( \tau \)) at the lowest level.

In this exercise there is a table that contains information of the code that univocally identifies each woman, the compensator, the quinquennial group they belong to, the exposure time in the respective group and the number of children in the group. Once the time of exposure is annualized (te) the function of estimation of rates in complex samples that Stata has to estimate specific fertility rates defined by the quotient of children/te in each age group was used.
Later TFR is calculated through a linear combination of the specific fertility rates (considering the covariances between TEF), the standard error and the confidence intervals are calculated under the supposition of a normal statistical distribution of TFR. For 1998 we consider a TFR of 4.243 children per woman, a datum close to Bolivia’s INE estimation (4.2), and the confidence interval of 95 percent is [4.108, 4.379] (table 1). How certain are these estimations considering that they are obtained from a sample of the total population and that a theoretical distribution, the normal, is assumed to estimate the confidence intervals? TFR’s statistical distribution is not known a priori.

Non-parametrical estimation: simple random bootstrapping

Before the application of bootstrapping considering the complex design of the sample, we experimented with simple random sampling (MAS) to evaluate the behavior of data in relation to the theory by Efron and Tibshirani (1993). By means of the algorithm of TEF and TFR estimation described in the section on the methods we obtained an initial punctual estimation of $t = t(F) = 4.228$ of TFR (it does not consider the covarianes between TEF). For 1000 replicas we obtain an estimation $t = t(F) = 4.327$ and a confidence interval of 95 percent of [4.187, 4.474] (table 1). The interval is coherent and includes the estimated value $t$ (the official TFR in 1998, is closer to the inferior limit). To the extent that the number of replicas increases, the distribution by sampling approaches the normal while $\hat{\varepsilon}$ slightly decreases.

In table 1 we see that the coefficients of variation (CV) are greater than the quotient between the bias and the standard error (with simple random bootstrapping), although not much, this indicates us that the bias of $\hat{t}$ is consistent as the following inequality, called standardized bias of $\hat{t}$, would be proven (Raj, 1968):

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

(3)
where

\[ CV(t) = \frac{\hat{\sigma}}{|t|} \]  

\( (4) \)

**Bootstrap in a phase of the sample design**

The third phase in the experimentation with bootstrap takes into account a phase; this is partly to verify if actually the standard error from the second phase can be negligible, such as it is indicated in the theory based on Taylor series. Several experiments were run with different numbers of replicas (200, 400, 1000) considering UPM and greater stability of the distribution by sampling of \( \hat{\tau} \) was found, being similar to the normal one (figure 4a). The Kolmogorov Smirnov test for 1000 replicas (sig = 0.901438 > 0.05) indicates us that we cannot reject the hypothesis of the normality of the statistical distribution of the total fertility rate of Bolivia in 1998.

It is necessary to point out that this variability is kept in determinate limit, for our case between 0.05525 and 0.05882. We also need to mention that the value of the standard error in 1000 replicas is lower than the standard error with simple random sampling (table 1). This is so because within the strata the variance is lower than the global variance for the design’s ends.

In respect to the behavior of the TFR estimator in function of the number of replicas, we see that there is a tendency toward convergence. For less than 400 replicas the TFR is more fluctuant, whereas for \( B \) greater or equal to 400, the estimator becomes more stable. This phenomenon proves the asymptotic theory in the background of the bootstrap method.

Bootstrapping in a phase of the design, as it was expected, allows us to estimate a biased TFR, this is to say, the expected value is different from the parameter that in this case is 4.2, according to INE. Nevertheless, the standardized bias is greater than the variation coefficient, which indicates that bias is not consistent according to equation 4. Otherwise, this can be interpreted as the confidence interval in a phase of the sample design would not be including the complete variability of the estimator. In such sense, in spite theory states that the variance in the second phase is negligible, it was necessary to experiment in a second phase of the sample design.
<table>
<thead>
<tr>
<th>Source: own calculation carried out in Stata, based in Endesa 1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>The official 1998 INE datum was considered as a parameter</td>
</tr>
</tbody>
</table>

**TABLE I**

<table>
<thead>
<tr>
<th>Standard Error (ES), Median, Bias, Variation Coefficient (CV) and Confidence Intervals (IC) Estimated by Parametric (1) and Non-Parametric (2, 3, 4) Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Es</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Parametric estimation of TFR</td>
</tr>
<tr>
<td>Simple random bootstrap</td>
</tr>
<tr>
<td>Single bootstrap in the sample design</td>
</tr>
<tr>
<td>Two-phase bootstrap in the sample design</td>
</tr>
</tbody>
</table>
**Figure 4**

Statistical distribution of TFR for the total population generated in a thousand replicas of A) bootstrap in a phase and B) bootstrap in two phases

Source: own elaboration based on data from bootstrap.

**Bootstrapping in two phases of the sample design**

In the first place, it is observed that distribution has a normal behavior (figure 4b). The distribution’s mean does not vary; it remains constant in 4.32 children per women for 1998, a value that is greater than the parametric estimation (method 1 of the figure 5b). The standard error is almost doubled in comparison with the previous method, since besides the variability in UPM, variability in USM is also considered because of this, in this case we will not be able to disregard the standard error in the second phase of the sample design.

The standard error under the supposition of normality (method 1) does not greatly differ from simple random bootstrap (figure 5), it decreases in single-phased bootstrapping, yet it increases considering the complete design of the sample (0.102).

According to the inequality of the standardized bias (equation 4), bootstrap in a second phase provide us with more precise and consistent estimations, given the variation coefficient is much larger than the standardized bias. In figure 6a we see that in the methods one, two and four the inequality is proven, while with bootstrapping in a single phase the probability that the interval would include the
real values of TFR was lower, this problem is solved considering the whole design of the sample in bootstrapping.

The confidence intervals of 95 percent, estimated from the different methods, are coherent since they include the different estimations of TFR (figure 6b). Approximately the amplitude of the confidence interval is the same for all of the different methods, with the exception of the third (bootstrap in a single phase) which is lower. Considering that the normality of the statistical distribution by sampling of TFR is accepted, it would be expected that the estimated statistics were similar by means of any method (but the third); nonetheless, the greater difference is to be found on the centrality of the mean by bootstrapping in the different intervals estimated. The mean by bootstrapping (4.32) is above the punctual estimation by parametric inference (4.24), which is closer to the inferior limit of the estimations by resampling. Considering that TFR’s parameter at national level in 1998 is 4.2 children per woman greater but consistent biases were obtained.

It is worth mentioning that, in 2002, the National Institute of Statistics of Bolivia (Instituto Nacional de Estadística de Bolivia) published new estimations of TFR for the 1990-1995 and 1995-2000 quinquenniums, whose confection took into account all of the information sources available on fertility thus far (diverse surveys and the 2001 census). For the 1995-2000 period a TFR of 4.32 children per woman was obtained, precisely the mean of the statistical distribution obtained by bootstrapping in the present work. This is to say, even if the confidence intervals vary in function of the standard error that at the time depends on the sample design, the mean by sampling remains constant and it can be considered as a estimation of lower bias or that it has a greater probability to approach the real value.

The standard error of a statistical estimation, using a multi-phased design, such as that used for Endsa 1998, is more complex that the standard error based on simple random sampling and tends to be greater that the standard error produced by a simple random sample. The increment in the standard error from the use of a multi-phased design is known as the effect of design and it is defined as the reason between the variance of the estimation with the currently used design and variance of the estimation that would be produced if a simple random sample was used. When it takes the value of one, indicates that the design is as efficient (produces minimal variances) as a simple at random; while a value greater than one, indicates that the design utilized produces a greater variance than that which would be obtained with a simple random sample (Cepep, 2004).
FIGURE 5
COMPARISON OF THE MAIN STATISTICS FROM THE APPLICATION OF PARAMETRIC AND NON-PARAMETRIC ESTIMATION

Source: own calculation based on the statistical distribution generated by bootstrapping.
FIGURE 6
STANDARDIZED BIAS AND CONFIDENCE INTERVALS ESTIMATED BY DIFFERENT METHODS

A

Source: own calculation based on the statistical distribution generated by bootstrapping.
In our case an Edis of 1.18 is obtained, proving that it is an efficient sample.

$$EDIS = \frac{VAR_{complex}(\hat{r})}{VAR_{simple}(\hat{r})}$$

(5)

**Bootstrapping by residence place**

As it was seen, the precision of an estimator through bootstrapping depends on the number of replicas, of the similarity of the bootstrapping with the original sample, of the parameter available for the calculation of bias (in our case 4.2) and $\hat{c}$. In order to contribute to explain $\hat{c}$ variations in a bi-stratified phased design it is interesting to analyze by urban/rural place of residence (figure 6).

It was already mentioned in the introduction that the differences in TFR by residence place in Bolivia are very important, which can influence on an increment or decrement of the total variance, moreover, the analysis by residence place allows us to verify the potential of bootstrapping by subgroups.

According to Kolmogorov-Smirnov test, the hypothesis of the normality of distribution for both the rural and urban areas cannot be rejected at a significance level of 0.1. Notwithstanding, we have to distinguish that according to statistical evidence the variance in the rural area is greater. This fact is also reflected on the differences of bias; the estimator’s bias $\hat{c}_{rural}$ is lower in the rural area than in the urban one. As bias depends on the extracted sample, the number of cases in each group can be influencing on the differences that are observed (7422 cases in the urban area and 3765 cases in the rural area). This warns us about the attention we should pay to the affirmations when samples of the total population are being worked with, since the changes can be caused by sampling variations and not necessarily by changes in the estimator of the population.

The amplitude of the confidence interval of the rural area is greater than that of the urban area, which indicates a greater dispersion of data from the rural area. This situation was expected, as fertility in urban Bolivia, based on Endsa 1998, is circa three children per woman, conversely, in the rural areas fertility rates are more heterogeneous and higher (approximately six children per woman).
Discussion

In regards to the research’s central topic: What is TFR’s statistical distribution? It is concluded that the hypothesis of the normality of its distribution cannot be rejected. Kolmogorov-Smirnov test indicates us that based on data from the Endsa-1998 sample there is not enough evidence to reject the void hypothesis stated. Hence, it is adequate to apply the law of the large numbers and the theorem of central limit in the calculation of the confidence intervals of TFR.

In the present research the supposition of normality and the technique of bootstrap have been applied to estimate the confidence interval of TFR’s estimator. Even if the results show us that when the number of replicas tends to infinite sampling distribution of TFR is similar to the normal, in table I and II we can see there are differences in the estimators’ precision.
### TABLE II

**TFR estimations, standard error (SE), bias and confidence interval (IC) by place of residence, Bolivia, 1998**

<table>
<thead>
<tr>
<th>Place of Residence</th>
<th>Median</th>
<th>Parameter 1998</th>
<th>Bias</th>
<th>Bias/es</th>
<th>CV</th>
<th>IC 95%</th>
<th>Inf.</th>
<th>Sup.</th>
<th>KS-Z</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>4.327</td>
<td>4.230</td>
<td>0.097</td>
<td>0.952</td>
<td>2.364</td>
<td>4.172</td>
<td>4.499</td>
<td>0.995</td>
<td>0.275</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>3.457</td>
<td>3.340</td>
<td>0.117</td>
<td>0.965</td>
<td>3.501</td>
<td>3.263</td>
<td>3.645</td>
<td>0.589</td>
<td>0.879</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>6.482</td>
<td>6.412</td>
<td>0.069</td>
<td>0.390</td>
<td>2.735</td>
<td>6.179</td>
<td>6.790</td>
<td>0.723</td>
<td>0.673</td>
<td></td>
</tr>
</tbody>
</table>

Source: own elaboration based on data from bootstrapping.
Under the supposition of normality of distribution we found that $\hat{\tau}_n = 4.243$, $\hat{\varepsilon}_n = 0.069$ and a confidence interval of 95 percent $[4.108, 4.379]$. Bootstrap method in two phases (1000 replicas) gives us a punctual estimation of $\hat{\tau} = 4.327$ with a confidence interval of 95 percent $[4.172, 4.499]$. Even if the estimation of the confidence interval of TFR under the supposition of normality includes with high probability the value of the demographic parameter, bootstrapping technique allows us to find more precise confidence intervals. In the case of simple random bootstrapping, such as the bootstrapping of the complex sample, confidence intervals include the official estimations and those from TFR bootstrapping; therefore, the confidence interval estimated by means of bootstrapping is coherent.

In this case the variance in a second phase cannot be considered negligible as the theory indicated; bootstrap in two phases increases the variability of data in as much as twice of a phase’s variability. This increment is coherent since it approaches the variability of simple random bootstrapping ($Edis = 1.18$), which is a good measure of the efficiency of the sample and that it is correctly reproduced.

The standard error estimated by bootstrapping the complex sample (0.102) has a value greater than the traditional estimation (0.069), reflecting the effect of a sample’s design. The importance of the sample’s design on the statistical inference is verified. So, when statistical model are constructed for samples, the use of functions oriented to complex samples is recommended, such as those of Stata, something that is not standard practice.

We found that TFR is a biased estimator for the sample of Endsa 1998 in Bolivia. This situation was expectable in accordance with the theory of ratio estimation. Based upon the theorems of parametric inference, we can only calculate unbiased indicators for any function of the means, however there are not formulae to calculate unbiased estimators of TFR, unless by means of approximations. Bias indicates us how far our estimation is from the demographic parameter since the fact of considering a sample and a technique of bootstrap allows us to calculate it automatically. The results show us a standardized bias much lower than the coefficient of variation (see figure 6a), which is an important characteristic to evaluate the consistency of estimations.

Finally, we managed to verify that normality of distribution is not rejected for subgroups either, both in the urban and rural areas in the present case, and so we can advance a step on the generalization of the theorem of the central limit
and the law of the large numbers for the estimator of the total fertility rate from a complex sample. Nonetheless, more experimentation with the technique is necessary for other populations with a different fertility pattern.

**Bibliography**


