The Sample Strategy for an UNICEF Survey in Croatia in 1996

Ksenija Dumičić  
*Faculty of Economics – Zagreb, Department of Statistics*  
*Kennedyjev trg 6*  
*10000 Zagreb, Croatia*  
kdumicic@oliver.efzg.hr

Srdan Dumičić  
*PULS – Market, Media and Public Opinion Research Agency*  
*Šubićeva 3*  
*10000 Zagreb, Croatia*  
sdumicic@puls.hr

1. Introduction

The paper describes the sample strategy (sample design and estimation method) applied in the large scale Multiindicator Survey in Health and Nutrition Sectors which was carried out in Croatia in 1996. The sampling methodology was proposed and agreed by representatives of the UNICEF (Monitoring Progress, UNICEF, 1995; Household Survey Manual, WHO, 1994), PULS - Croatian research agency (Dumičić et al.,1996) and the project leader from Clinic for Infant Diseases- Zagreb.

The target population was defined as "households with children under age 5". The respondents had to be mothers (or caretakers). A stratified random sample of clusters was proposed. A "cluster" is defined as a randomly selected population group of a size likely to include a specific number of children under age 5. In this survey, cluster sampling is a two-stage process involving: first, selecting a kind of area units, here called segments, in which clusters should be located; and second, identifying groups of households where interviews should be conducted. A “household” was defined as a group of persons who live and eat together. Since the post-war situation effects different areas of the country to varying degrees, stratification was based on areas of Croatia, to express these differences both before and during the survey period.

2. Stratified multi-stage sampling design

Since only a rather outdated sampling frame for selecting children up to five was available, a very simple sampling strategy seemed to be quite useful. Stratification according to the degree of peace before the survey period was applied, because this seemed to be important to influence the value of the health and nutrition variables to measure, so, the following three strata were formulated: "the part of the country which was free during the war"; strata "front line areas", and strata "liberated areas". To facilitate future comparisons, all the results sought are also given for the six traditional geographical regions of Croatia. The allocation of units over the three strata, as well as over the six regions is approximately proportionate to size, since the data from the Census’91 were out of date.

The sample of 14800 households was selected to reach approximately 2000 children up to 5. The number of households to be visited in Croatia was calculated according to Census’91, taking a household of average size as 3.1 persons, and a percentage of the population 1-year of age of 1.1%. According to Census’91 in Croatia there was 5.85% children under 5. To compensate for expected non-responses 10% of over sampling was proposed for the first two strata, and 15% for strata "liberated areas". The data source for the first two strata was the Census ‘91, and for the third, recent figures from the Office for Refugees was used. All sampling fractions amount to about 0.01, with the exception of the strata "liberated areas" where the fraction was about 0.02.
In the first two strata the following two-stage sampling approach was applied: segment selection stage (segments are primary sampling units, a "segment" is defined as a kind of an area unit created using a list of addresses of heads of households); and cluster selection stage (the cluster size determined was unique and it was "40 households", according to UNICEF experience this size offers an acceptable ratio between the cost and the efficiency of the results). At each stage units were sampled systematically. The segments were of different sizes with 120-250 households, and were selected by PPS selection. From each segment chosen clusters of 40 households were included in the sample. The proportional allocation of units over strata at each stage was used.

In the third "liberated" strata segments were selected randomly, but only in the settlements recognised as suitable for interviewing (and with children) and were chosen bearing in mind the fact that the proportion of rural to non-rural citizens had to be maintained. During the second stage, random walk method was applied in each selected segment to find 40 households. Random starting points were determined according to the situation in partially destroyed settlements.

3. Estimating the parameters

Estimation formulae for parameters assumed that in each of the strata a simple random sample was chosen and not a cluster sample. So, standard errors of estimates were lower than in cases the clusters of elements chosen should be taken into consideration. When estimating proportions and arithmetic means only simple unbiased estimators were recommended and calculated. Estimates appeared to be quite accurate. Weighting for stratified estimates was examined, but not reported because of poor reliability of the sampling frame. All the results are comparable to the results of the same kind of surveys from other countries. All reported proportions in Croatia as a whole at the 95% confidence level would be accurate within two percentage points (94%-96%). The coefficients of variation had very low values. Thus in the strata "liberated areas" e.g. for the variable "Have you ever breast-fed you child?" and the answer modality "yes", the proportion of breast-fed children was estimated at 0.912 (91.2% of children were breast-fed), with standard error of 0.0199, or coefficient of variation of 2.18%. This actual coefficient had the highest value in relation to others.

REFERENCES


RÉSUMÉ

Cet article vise à exposer le plan de sondage, la méthode d’échantillonnage et d’estimation, appliquée dans La multiindicateur enquête de grande envergure de la santé et de la nutrition, qui a été créée en Croatie en 1996. Le plan d’échantillonnage stratifié en trois étapes a été utilisé. La stratification respectée le niveau de paix. Un échantillon de 14800 ménages a été sélectionné avec 2000 des jeunes enfants jusqu'à 5 année. Dans les strates une procédure a plusieurs étapes a été appliquée: sélection des segments et sélection des groupes de ménages comme des unités fondamentales de sélection. Ce qui suppose qu’il en résulte un échantillon de probabilité. Les estimateurs étaient assez précis. Une pondération n’était pas recommandée dans le cadre de l’échantillon disponible qui se trouvait être démodée a cause de la migration générée de façon continu par la guerre en Croatie.