CHAPTER 2

METHODOLOGY

2.1 Sample Design

A Stratified Two - Stage Sampling was adopted for the survey. Provinces were constituted strata. The primary and secondary sampling units were blocks for municipal areas / villages for non - municipal areas and private households / persons in the special households respectively.

Stratification

Provinces were constituted strata. There were altogether 76 strata. Each stratum was divided into two parts according to the type of local administration, namely municipal areas and non - municipal areas.

Selection of Primary Sampling Unit

The sample selection of blocks / villages were performed separately and independently in each part by using probability proportional to size - total number of households. The total sample blocks / villages was 2,050 from 109,966 blocks / villages.

The total number of sample blocks / villages selected for enumeration by region and type of local administration was as follows :

Region / Stratum	Total	Municipal Areas	Non - Municipal Areas
Bangkok Metropolis	450	450	-
Central (Excluding Bangkok Metropolis)	400	200	200
North	400	200	200
Northeast	400	200	200
South	400	200	200
Total	**2,050**	**1,250**	**800**

Selection of Secondary Sampling Unit

Private households were our ultimate sampling units. A new listing of private households were made for every sample block / village to serve as the sampling frame. In each sample block / village, a systematic sample of private households were selected with 40 sample households per block / village.

All special households located within the sample areas were included in the sample and the persons in the special household were systematically selected for the interviewing.

The total number of sample private households selected for enumeration by region and type of local administration was as follows :

Region / Stratum	Total	Municipal Areas	Non - Municipal Areas
Bangkok Metropolis	18,000	18,000	-
Central (Excluding Bangkok Metropolis)	16,000	8,000	8,000
North	16,000	8,000	8,000
Northeast	16,000	8,000	8,000
South	16,000	8,000	8,000
Total	**82,000**	**50,000**	**32,000**

2.2 Method of Estimation

The survey results were presented separately 2 parts. Part 1 were presented information of persons and part 2 were presented information for households.

The survey results were presented separately for the Bangkok Metropolis and the remaining 75 provinces were classified by region, municipal areas and non - municipal areas.

Let l = 1 , 2 , 3 , ... ,42 (age - sex group)

k = 1 , 2 , 3 , ... , m_{hij} (sample block / village)

j = 1 , 2 (type of local administration)

i = 1 , 2 , 3 , ... , A_h (province)

h = 1 , 2 , 3 , 4 , 5 (region)

PART 1 : INFORMATION OF PERSONS

1.1 Estimate of the Total Number of Persons with Characteristic X

1.1.1 Adjusted estimate of the total number of persons with characteristic X for the l^{th} age - sex group, j^{th} area , h^{th} region was based on the formula :

$$x''_{1hjl} = \frac{x'_{1hjl}}{y'_{1hjl}} Y_{1hjl} = r_{1hjl} Y_{1hjl} \qquad \text{..... (1)}$$

where x'_{1hjl} is the ordinary estimate of the total number of persons with characteristic X for the l^{th} age - sex group, j^{th} area , h^{th} region.

y'_{1hjl} is the ordinary estimate of the total population for the l^{th} age - sex group, j^{th} area, h^{th} region.

Y^{*}_{1hjl} is the estimate, based on the population projection of the total population for the the l^{th} age - sex group, j^{th} area, h^{th} region.

r_{1hjl} is the ratio of the estimate of the total number of persons with characteristic X to the estimate of the total population for the l^{th} age - sex group, j^{th} area, h^{th} region.

---

* Population Projections for Thailand 1990 - 2020, Human Resources Planning Division, National Economic and Social Development Board, The Eighth National Economic and Social Development Planning, March 1995.

The formula of the estimate from a stratified two - stage sampling was as follows :

i) $$x'_{1hjl} = \sum_{i=1}^{A_h} x'_{1hijl}$$ ..... (2)

where x'_{1hijl} is the ordinary estimate of the total number of persons with characteristic X for the l^{th} age - sex group, j^{th} area, i^{th} province, h^{th} region.

$$x'_{1hijl} = \frac{1}{m_{hij}} \sum_{k=1}^{m_{hij}} \frac{1}{P_{hijk}} \frac{N_{hijk}}{n_{hijk}} x_{1hijkl}$$

x_{1hijkl} is the total number of persons with characteristic X for the l^{th} age - sex group, k^{th} sample block / village, j^{th} area, i^{th} province, h^{th} region.

N_{hijk} is the total number of listing households in the k^{th} sample block / village, j^{th} area, i^{th} province, h^{th} region.

n_{hijk} is the total number of sample households in the k^{th} sample block / village, j^{th} area, i^{th} province, h^{th} region.

P_{hijk} is the probability of selection of the k^{th} sample block / village, j^{th} area, i^{th} province, h^{th} region.

m_{hij} is the total number of sample blocks / villages in the j^{th} area, i^{th} province, h^{th} region.

A_h is the total number of provinces in the h^{th} region and $\sum_{h=1}^{5} A_h = 76$

ii) $$y'_{1hjl} = \sum_{l=1}^{A_h} y'_{1hijl}$$ ..... (3)

where y'_{1hijl} is the ordinary estimate of the total population for the l^{th} age – sex group, j^{th} area, i^{th} province, h^{th} region.

$$y'_{1hijl} = \frac{1}{m_{hij}} \sum_{k=1}^{m_{hij}} \frac{1}{P_{hijk}} \frac{N_{hijk}}{n_{hijk}} y_{1hijkl}$$

y_{1hijkl} is the total number of the population enumerated for the l^{th} age - sex group, k^{th} sample block / village, j^{th} area, i^{th} province, h^{th} region.

1.1.2 Adjusted estimate of the total number of persons with characteristic X for the j^{th} area, h^{th} region was based on the formula :

$$x''_{1hj} = \sum_{l=1}^{34^{42}} x''_{1hjl} \qquad \text{..... (4)}$$

1.1.3 Adjusted estimate of the total number of persons with characteristic X for the l^{th} age - sex group, h^{th} region was based on the formula :

$$x''_{1hl} = \sum_{j=1}^{2} x''_{1hjl} \qquad \text{..... (5)}$$

1.1.4 Adjusted estimate of the total number of persons with characteristic X for the h^{th} region was based on the formula :

$$x''_{1h} = \sum_{j=1}^{2} x''_{1hj} = \sum_{l=1}^{34^{42}} x''_{1hl} \qquad \text{..... (6)}$$

1.1.5 Adjusted estimate of the total number of persons with characteristic X for the j^{th} area was based on the formula :

$$x''_{1j} = \sum_{h=1}^{5} x''_{1hj} \qquad \text{..... (7)}$$

1.1.6 Adjusted estimate of the total number of persons with characteristic X for the l^{th} age - sex group of the whole kingdom was based on the formula :

$$x''_{1l} = \sum_{h=1}^{5} x''_{1hl} \qquad \text{..... (8)}$$

1.1.7 Adjusted estimate of the total number of persons with characteristic X for the whole kingdom was based on the formula :

$$x''_1 = \sum_{h=1}^{5} x''_{1h} = \sum_{j=1}^{2} x''_{1j} = \sum_{l=1}^{34} x''_{1l} \qquad \text{..... (9)}$$

1.2 Estimate of Variance of the Total Number of Persons with Characteristic X

1.2.1 The estimate variance of x''_{1hjl} was

$$\hat{V}(x''_{1hjl}) = \left[\frac{Y_{1hjl}}{y'_{1hjl}}\right]^2 \sum_{i=1}^{A_h} \frac{1}{m_{hij}(m_{hij}-1)} \left[\sum_{k=1}^{m_{hij}} z'^2_{1hijkl} - m_{hij} z'^2_{1hijl}\right] \qquad \text{..... (10)}$$

where $z'_{1hijkl} = x'_{1hijkl} - r_{1hjl} y'_{1hijkl}$

$$z'_{1hijl} = x'_{1hijl} - r_{1hjl} y'_{1hijl}$$

$$x'_{1hijkl} = \frac{1}{P_{hijkl}} \frac{N_{hijk}}{n_{hijk}} x_{1hijkl}$$

$$y'_{1hijkl} = \frac{1}{P_{hijk}} \frac{N_{hijk}}{n_{hijk}} y_{1hijkl}$$

1.2.2 The estimate variance of x''_{1hj} was

$$\hat{V}(x''_{1hj}) = \sum_{l=1}^{42} V(x''_{1hjl}) \qquad \text{..... (11)}$$

1.2.3 The estimate variance of x''_{1hl} was

$$\hat{V}(x''_{1hl}) = \sum_{j=1}^{2} \hat{V}(x''_{1hjl}) \qquad \text{..... (12)}$$

1.2.4 The estimate variance of x''_{1h} was

$$\hat{V}(x''_{1h}) = \sum_{j=1}^{2} \hat{V}(x''_{1hj}) = \sum_{l=1}^{\overset{42}{34}} \hat{V}(x''_{1hl}) \quad \text{..... (13)}$$

1.2.5 The estimate variance of x''_{1j} was

$$\hat{V}(x''_{1j}) = \sum_{h=1}^{5} \hat{V}(x''_{1hj}) \quad \text{..... (14)}$$

1.2.6 The estimate variance of x''_{1l} was

$$\hat{V}(x''_{1l}) = \sum_{h=1}^{5} \hat{V}(x''_{1hl}) \quad \text{..... (15)}$$

1.2.7 The estimate variance of x''_{1} was

$$\hat{V}(x''_{1}) = \sum_{h=1}^{5} \hat{V}(x''_{1h}) = \sum_{j=1}^{2} \hat{V}(x''_{1j}) = \sum_{l=1}^{\overset{42}{34}} \hat{V}(x''_{1l}) \quad \text{..... (16)}$$

1.3 Coefficient of Variation (CV) of the Total Number of Persons with Characteristic X

1.3.1 The formula of CV of x''_{1hjl} was

$$CV(x''_{1hjl}) = \frac{\sqrt{\hat{V}(x''_{1hjl})}}{x''_{1hjl}} \times 100\% \quad \text{..... (17)}$$

1.3.2 The formula of CV of x''_{1hj} was

$$CV(x''_{1hj}) = \frac{\sqrt{\hat{V}(x''_{1hj})}}{x''_{1hj}} \times 100\% \quad \text{..... (18)}$$

1.3.3 The formula of CV of x''_{1hl} was

$$CV(x''_{1hl}) = \frac{\sqrt{\hat{V}(x''_{1hl})}}{x''_{1hl}} \times 100\% \quad \text{..... (19)}$$

1.3.4 The formula of CV of x''_{1h} was

$$CV(x''_{1h}) = \frac{\sqrt{\hat{V}(x''_{1h})}}{x''_{1h}} \times 100\% \quad \text{..... (20)}$$

1.3.5 The formula of CV of x''_{1j} was

$$CV(x''_{1j}) = \frac{\sqrt{\hat{V}(x''_{1j})}}{x''_{1j}} \times 100\% \quad \text{..... (21)}$$

1.3.6 The formula of CV of x''_{1l} was

$$CV(x''_{1l}) = \frac{\sqrt{\hat{V}(x''_{1l})}}{x''_{1l}} \times 100\% \quad \text{..... (22)}$$

1.3.7 The formula of CV of x''_{1} was

$$CV(x''_{1}) = \frac{\sqrt{\hat{V}(x''_{1})}}{x''_{1}} \times 100\% \quad \text{..... (23)}$$

PART 2 : INFORMATION OF HOUSEHOLDS

2.1 Estimate of the Total Number of Households with Characteristic X

2.1.1 Adjusted estimate of the total number of households with characteristic X for the j^{th} area, h^{th} region was based on the formula :

$$x''_{2hj} = \frac{x'_{2hj}}{y'_{2hj}} Y_{2hj} = r_{2hj} Y_{2hj} \quad \text{..... (24)}$$

where x'_{2hj} is the ordinary estimate of the total number of households with characteristic X for the j^{th} area, h^{th} region.

y'_{2hj} is the ordinary estimate of the total households for the j^{th} area, h^{th} region.

Y^{*}_{2hj} is the estimate, based on the population projection of the total households for the j^{th} area, h^{th} region.

r_{2hj} is the ratio of the estimate of the total number of households with characteristic X to the estimate of the total households for the j^{th} area, h^{th} region.

The formula of the estimate from a stratified two - stage sampling was as follows :

$i)$ $$x'_{2hij} = \sum_{i=1}^{A_h} x'_{2hij}$$ ..... (25)

where x'_{2hij} is the ordinary estimate of the total number of households with characteristic X for j^{th} area, i^{th} province, h^{th} region.

$$x'_{2hij} = \frac{1}{m_{hij}} \sum_{k=1}^{m_{hij}} \frac{1}{P_{hijk}} \frac{N_{hijk}}{n_{hijk}} x_{2hijk}$$

x_{2hijk} is the total number of households with characteristic X for the k^{th} sample block / village, j^{th} area, i^{th} province, h^{th} region.

$ii)$ $$y'_{2hj} = \sum_{i=1}^{A_h} y'_{2hij}$$ ..... (26)

---

* Population Projections for Thailand 1990 - 2020, Human Resources Planning Division, National Economic and Social Development Board, The Eighth National Economic and Social Development Planning, March 1995.

where y'_{2hij} is the ordinary estimate of the total households for the j^{th} area, i^{th} province, h^{th} region.

$$y'_{2hij} = \frac{1}{m_{hij}} \sum_{k=1}^{m_{hij}} \frac{1}{P_{hijk}} \frac{N_{hijk}}{n_{hijk}} y_{2hijk}$$

y_{2hijk} is the total number of the households enumerated for the k^{th} sample block / village, j^{th} area, i^{th} province, h^{th} region.

2.1.2 Adjusted estimate of the total number of households with characteristic X for the h^{th} region was based on the formula :

$$x''_{2h} = \sum_{j=1}^{2} x''_{2hj} \qquad \text{..... (27)}$$

2.1.3 Adjusted estimate of the total number of households with characteristic X for the j^{th} area was based on the formula :

$$x''_{2j} = \sum_{h=1}^{5} x''_{2hj} \qquad \text{..... (28)}$$

2.1.4 Adjusted estimate of the total number of households with characteristic X for the whole kingdom was based on the formula :

$$x''_{2} = \sum_{h=1}^{5} x''_{2h} = \sum_{j=1}^{2} x''_{2j} \qquad \text{..... (29)}$$

2.2 Estimate of Variance of the Total Number of Households with Characteristic X

2.2.1 The estimate variance of x''_{2hj} was

$$\hat{V}\left(x''_{2hj}\right) = \left[\frac{Y_{2hj}}{y'_{2hj}}\right]^2 \sum_{i=1}^{A_h} \frac{1}{m_{hij}\left(m_{hij}-1\right)} \left[\sum_{k=1}^{m_{hij}} z'^{2}_{2hijk} - m_{hij} z'^{2}_{2hij}\right] \qquad \text{..... (30)}$$

where $z'_{2hijk} = x'_{2hijk} - r_{2hj} y'_{2hijk}$

$$z'_{2hij} = x'_{2hij} - r_{2hj} y'_{2hij}$$

$$x'_{2hijk} = \frac{1}{P_{hijk}} \frac{N_{hijk}}{n_{hijk}} x_{2hijk}$$

$$y'_{2hijk} = \frac{1}{P_{hijk}} \frac{N_{hijk}}{n_{hijk}} y_{2hijk}$$

2.2.2 The estimate variance of x''_{2h} was

$$\hat{V}(x''_{2h}) = \sum_{j=1}^{2} \hat{V}(x''_{2hj}) \quad \text{..... (31)}$$

2.2.3 The estimate variance of x''_{2j} was

$$\hat{V}(x''_{2j}) = \sum_{h=1}^{5} \hat{V}(x''_{2hj}) \quad \text{..... (32)}$$

2.2.4 The estimate variance of x''_{2} was

$$\hat{V}(x''_{2}) = \sum_{h=1}^{5} \hat{V}(x''_{2h}) = \sum_{j=1}^{2} \hat{V}(x''_{2j}) \quad \text{..... (33)}$$

2.3 Coefficient of Variation (CV) of the Total Number of Households with Characteristic X

2.3.1 The formula of CV x''_{2hj} was

$$CV(x''_{2hj}) = \frac{\sqrt{\hat{V}(x''_{2hj})}}{x''_{2hj}} \times 100\% \quad \text{..... (34)}$$

2.3.2 The formula of CV x''_{2h} was

$$CV(x''_{2h}) = \frac{\sqrt{\hat{V}(x''_{2h})}}{x''_{2h}} \times 100\% \quad \text{..... (35)}$$

2.3.3 The formula of CV x''_{2j} was

$$CV(x''_{2j}) = \frac{\sqrt{\hat{V}(x''_{2j})}}{x''_{2j}} \times 100\,\% \qquad \text{..... (36)}$$

2.3.4 The formul $CV(x''_{2j}) = \frac{\sqrt{\hat{V}(x''_{2j})}}{x''_{2j}} \times 100\,\%$ a of CV x''_{2} was

$$CV(x''_{2}) = \frac{\sqrt{\hat{V}(x''_{2})}}{x''_{2}} \times 100\,\% \qquad \text{..... (37)}$$

2.3 Data Collection

The 2005-2006 survey of population change used time of survey for 1 year. The survey have listing before enumeration and total sample of households 82,000 household was interview by province of statistical official.

2.4 Data Processing

Manual editing and coding was carried out at provincial statistical offices (PSOs) in all provinces in Thailand. Then, the questionnaires were sent to the Central Office for data capture for the first round data collection. An Intelligent Character Recognition (ICR) solution was adopted for data capture. Data processing, including tabulation and analysis, was carried out at the central office. Data in the first round was present on population characteristics. After data collection in round 2-5, data entry on microcomputer by using FTP program carried out at provincial statistical offices in all provinces in Thailand. The data files were transfer to the central office for tabulation, analysis and report writing on the survey of population change.

2.5 In Round Figures

In the statistical tables, all absolute figures are independently round : hence the group total may not always be equal to the sum of the individual fiqures.