

Appendix 3 Sampling Method, Estimation Method, and Sampling Errors of Estimates

1. Sampling method

The sampling method is a stratified two-stage sampling using the EDs (enumeration districts) of the Population Census as the first-stage sampling units and the dwelling units* as the second-stage sampling units.

* A dwelling unit is a structure or a part of a structure intended for habitation of a household, such as a detached house, an apartment, and a living quarter in a dormitory.

(1) Sampling of EDs (First-stage sampling)

For each region*, all EDs are classified into strata according to their characteristic** and in each stratum, the sample EDs are selected with the probability proportion to the weight (roughly the number of households in the ED divided by 15). The number of the sample EDs is about 2,900 every month.

The following EDs are excluded from the sampling.

- (a) Reformatory institutions
- (b) Areas resided exclusively by the personnel of the Self-Defence Forces
- (c) Areas resided exclusively by the personnel of foreign armed forces
- (d) Water districts

The data for the inmates of reformatory institutions and the personnel of the Self-Defence Forces are reported by the authorities in charge and are added to the tabulation, whereas in the “Detailed Tabulation”, those are excluded from the tabulation.

* Hokkaido, Tohoku, Southern-Kanto, “Northern-Kanto, Koshin”, Hokuriku, Tokai, Kinki, Chugoku, Shikoku, “Kyusyu, Okinawa”

(note) Stratification and sampling are applied to 11 regions composed of Kyusyu (excluding Okinawa), Okinawa and other 9 regions.

** Stratification is applied by the industrial and status composition of employed persons and other characteristics in each ED, based on the Population Census. The design of the sampling frame is revised every five years, and the 2005 Census-based design is being introduced from May 2008 to August 2009, step by step.

(2) Sampling of dwelling units (Second-stage sampling)

For each sample ED, about 15 dwelling units are selected from all dwelling units in the ED by systematic sampling with a constant sampling interval.

All households living in the selected dwelling units, which are approximately 40,000 households in total, are surveyed.

(3) Sample rotation and structure of sub-samples

In order to improve accuracy of the estimation on both monthly and annual figures and their changes, the following sample rotation system is applied.

- (a) A sample ED remains in a sample for four consecutive months, leaves the sample for the following eight months, and joins the sample again for the same four months in the following year.
- (b) For each ED, two sets of dwelling units are selected. In the first year of enumeration for the sample ED, the households in the sample dwelling units in the first set are surveyed for the first two consecutive months, and then replaced by the households in the dwelling units of the other set. In the second year, the dwelling units of the first set enter the sample again, and are replaced by those of the other set in the same way as in the first year.
- (c) Under this system, one fourth of the sample EDs and half of the sample households are replaced every month. Three fourths of sample EDs are common from month to month and a half from year to year.

In this system, the whole sample is divided into the following eight sub-samples, each of which forms an independent random sample of the universe. These sub-samples are used in calculation of sampling errors.

The eight sub-samples are represented by A1, A2, B1, B2, C1, C2, D1 and D2.

- A i EDs beginning in January, May and September
- B i EDs beginning in February, June and October
- C i EDs beginning in March, July and November
- D i EDs beginning in April, August and December
- $i=1$ EDs of the first year
- $i=2$ EDs of the second year

Two sets of dwelling units (set of A2 and C2 or set of B2 and D2, corresponding to the second month in second year) are surveyed of the Special Questionnaire.

2. Estimation Method

(1) Outline of the estimation

The estimation method for the monthly figures of whole Japan is the ratio estimation with a benchmark of the population by age (15 age groups*), sex and area (2 areas: metropolitan area consisting of seven largest cities**, and the rest of the country) estimated from the Population Estimates based Population Census(refer to (3)).

* A change in the method of estimation by 15 age groups (15-19, ..., 80-84, 85-) used in the Labour Force Survey was introduced in 2007.

**Ku-area of Tokyo, Yokohama-shi, Nagoya-shi, Kyoto-shi, Osaka-shi, Kobe-shi and Kitakyusyu-shi

The basic formula of the estimation is as follows.

[Ex. the case for employed person]
Ratio estimate of employed person
=Linear estimate of employed person
× $\frac{\text{Benchmark population}}{\text{Linear estimate of total population}}$

(note) A linear estimation is to estimate the universe by multiplying the enumerated population by the reciprocal of the sampling ratio.

The quarterly and annual averages are the arithmetic mean of the monthly figures.

(2) Procedure of the estimation

The estimate of the population having characteristics X is obtained by summing the ratio estimates \hat{X} s for the sets of sex, age and area concerned.

\hat{X} for a set of sex, age and area is obtained as follows:

$$\hat{X} = \frac{\sum_{i=1}^L \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{\omega_i}{\omega_{ij}} \cdot f_{ij} \cdot x_{ij} \cdot \frac{P}{\sum_{i=1}^L \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{\omega_i}{\omega_{ij}} \cdot f_{ij} \cdot P_{ij}}}{\sum_{i=1}^L \sum_{j=1}^{m_i} x_{ij} \cdot F_i \cdot \frac{P}{\sum_{i=1}^L \sum_{j=1}^{m_i} P_{ij} \cdot F_i}}$$

i (=1, 2,, L) : each section by region and stratum

j (=1, 2,, m_i) : each sample ED in each section

x_{ij} : Enumerated population (for the sex and

age concerned) with characteristic X in j -th sample ED in i -th section

ω_{ij} : Weight for j -th sample ED in i -th section

f_{ij} : Reciprocal of the sampling ratio of dwelling units within j -th sample ED in i -th section (= ω_{ij})

ω_i : Total of the weights for all the EDs in i -th section

m_i : Number of the sample EDs in i -th section

F_i (= ω_i / m_i) : Multiplier for linear estimation for i -th section

P : Benchmark population (for the set of sex, age and area concerned)

P_{ij} : Enumerated population (for the sex and age concerned) in j -th sample ED in i -th section

$\frac{P}{\sum_{i=1}^L \sum_{j=1}^{m_i} P_{ij} \cdot F_i}$: Multiplier for ratio estimation (for the set of sex, age and area concerned)

The estimation method for the 10 regions is a parallel to that for the whole Japan.

For each item, adjustment is made so that the figures for 10 regions add up to the figure for the whole Japan.

(3) Benchmark population and its revision

“Population Estimates” released by the Statistics Bureau are used as the benchmark population for calculating the survey results.

These population estimates are calculated by using the Population Census counts as the base and by adding the other monthly data on population changes after the Census*.

Using the latest results of the Population Census as the base, the Population Estimates are revised every five years according to the release of new results of the Census conducted quinquennially. Therefore, the benchmark population for calculating the results of the Labour Force Survey revised every five years.

* A change in the method of computing the Japanese number of emigrants and immigrants using “the number beyond overseas length-of-stay 91 day” since Jan.-Mar. 2007.

The benchmark population for the estimation was revised to the 2005 Census-based population estimates in Jan.-Mar. 2007.

Reference Tab.1 Difference in level of the benchmark revisions (Detailed Tabulation)

(10 thousands, point)

	Population of 15 years old or more	Employed person	Unemployed person	Not in labour force
2007 (Bmk 2005 Census)	+ 5	- 3	-1	+ 8

(4) Estimation method for “Detailed Tabulation”

The quarterly and annual averages are the arithmetic mean of the monthly figures.

The estimation method for the monthly figures of whole Japan adopt the ratio estimation with the benchmark of population by age (5 age groups), sex and labour force status (Employed person, Unemployed person and Not in labour force) which estimated in Labour Force Survey*.

The basic formula of the estimation is as follows.

[Ex. the case for employed person on column A in the Special Questionnaire]

Estimate of column A in the Special Questionnaire

= Linear estimate of column A in the Special Questionnaire

× $\frac{\text{Employed person which estimated in Labour Force Survey}^*}{\text{Employed person which estimated in Detailed Tabulation}}$

Multiplier for linear estimation in Detailed Tabulation is same as Labour Force Survey*.

* Results from the Basic Questionnaire

3. Sampling error for “Detailed Tabulation”

The magnitude of sampling error varies by the size of estimates, the kind of item and the reference period. Standard errors for estimates are shown in the following tables.

Tab. 1 Standard error for annual average for the whole Japan (Detailed Tabulation)

Estimates (10 thousand persons)	Standard error (10 thousand persons)	Relative standard error (%)
5000	19.6	0.4
2000	11.9	0.6
1000	8.1	0.8
500	5.5	1.1
200	3.4	1.7
100	2.3	2.3
50	1.6	3.1
20	0.9	4.7
10	0.6	6.5

Tab. 2 Standard error for quarter average for the whole Japan (Detailed Tabulation)

Estimates (10 thousand persons)	Standard error (10 thousand persons)	Relative standard error (%)
5000	39.3	0.8
2000	23.7	1.2
1000	16.2	1.6
500	11.0	2.2
200	6.6	3.3
100	4.5	4.5
50	3.1	6.2
20	1.9	9.3
10	1.3	12.7