

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.

(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. Three sub-samples are used in calculation of sampling errors.

The eight sub-samples are represent 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

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 (13 age groups), sex and area (2 areas: metropolitan area consisting of seven largest cities, and the rest of the country) estimated from the Population Census and other sources.

The basic formula of the estimation is as follows. (ex. the case for employed persons)

$$\begin{aligned} & \text{Ratio estimate of employed persons} \\ & = \text{Linear estimate of employed persons} \\ & \quad \times \frac{\text{Benchmark population}}{\text{Linear estimate of total population}} \end{aligned}$$

(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 means 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 of sex, age and area is obtained as follows :

$$\begin{aligned} \hat{X} &= \frac{L}{i=1} \frac{1}{m_i} \frac{m_i}{j=1} \frac{\omega_i}{\omega_{ij}} \cdot f_{ij} \cdot x_{ij} \frac{P}{\frac{1}{m_i} \frac{m_i}{j=1} \frac{\omega_i}{\omega_{ij}} \cdot f_{ij} \cdot P_{ij}} \\ &= \frac{L}{i=1} \frac{m_i}{j=1} x_{ij} \cdot F_i \frac{P}{P_{ij} \cdot F_i} \end{aligned}$$

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 first-stage sampling of 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 first-stage sampling of 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

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

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

“The current 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.

The benchmark population for the estimation has been revised to the 2000 Census-based population estimates. So the figures for 2002 include some methodological component that shifts the total population aged 15 or more downward by approximately 60,000.

3. Sampling error

Though magnitude of standard errors varies with the size of estimates, the kind of item and the month and the year of survey, average figures of standard errors which are calculated by using the sub-samples are shown in the following tables.

Standard error for annual average for the whole Japan

Estimate (10 thousand persons)	Standard error (10 thousand persons)	Ratio of standard error (%)
5000	13.5	0.3
2000	8.4	0.4
1000	5.8	0.6
500	4.1	0.8
200	2.5	1.3
100	1.8	1.8
50	1.2	2.5
20	0.8	3.8
10	0.5	5.3

The formula used in calculating the ratio of standard error for annual average of estimates is as follows:

$$\sqrt{\frac{1}{8(8-1)} \sum_{i=1}^8 (\bar{X}_i - \bar{X})^2} / \bar{X}$$

\bar{X}_i and \bar{X} respectively refer to the annual average of estimates with characteristic X in the i -th sub-sample ($i=1, 2, \dots, 8$) and whole sample.

Standard error for monthly results for the whole Japan

Estimate (10 thousand persons)	Standard error (10 thousand persons)	Ratio of standard error (%)
5000	26.1	0.5
2000	16.7	0.8
1000	11.9	1.2
500	8.5	1.7
200	5.5	2.7
100	3.9	3.9
50	2.8	5.6
20	1.8	8.9
10	1.3	12.7

The formula used in calculating the ratio of standard error for monthly estimate is as follows :

$$\sqrt{\frac{1}{8(8-1)} \sum_{i=1}^8 (\hat{X}_i - \hat{X})^2} / \hat{X}$$

\hat{X}_i and \hat{X} respectively refer to the estimates with characteristic X in the i -th sub-sample ($i=1, 2, \dots, 8$) and whole sample.

4. Seasonal adjustment Methods

The seasonal adjustment procedure for the labour force series is the Census Method (X-11) programme with the control limit for extremes set at 9.8 ~ 9.9 , and using standard options for others.

The seasonally adjusted figures released monthly are calculated by using the predicted seasonal factors, which are estimated through seasonal adjustments of the data up to December of the previous year. Before the release of January results, the seasonal adjustments are computed retroactively by adding new data for twelve months of the previous year.

The *reference table 1* in this report shows the results of seasonal adjustments using the data up to December of 2005.