

第3章 抽出詳細集計の抽出方法及び推定値の精度

標本の抽出方法

抽出詳細集計では、母集団を次の二つの層に分け世帯の抽出を行った。

- (1) 一般世帯及び30人未満の施設等の世帯（自衛隊の営舎内居住者及び矯正施設の入所者を除く。）（以下「標本層世帯」という。）
- (2) 30人以上の施設等の世帯（自衛隊の営舎内居住者及び矯正施設の入所者を含む。）（以下「しっ皆層世帯」という。）

このうち、標本層世帯の抽出は、市区町村別結果で一定の精度を確保するため、市区町村の人口規模に応じて標本の規模を定め、世帯を単位として無作為系統抽出を行い抽出した。しっ皆層世帯については、全ての世帯を抽出した。市区町村の人口規模別の平均抽出率は、表1のとおり。

表1 全国の抽出率及び市区町村の人口規模別平均抽出率

	抽出率					
	抽出率(%)	n分の1				
全国	9.9	10.1				

市区町村の人口規模	抽出率					
	市町村		特別区		政令指定都市の区	
	抽出率(%)	n分の1	抽出率(%)	n分の1	抽出率(%)	n分の1
50万以上	5.3	19.0	5.3	19.0	5.3	19.0
30万以上50万未満	5.0	20.2	8.5	11.8	5.0	20.2
20万以上30万未満	7.2	13.8	12.2	8.2	7.2	13.8
10万以上20万未満	13.5	7.4	21.7	4.6	13.5	7.4
5万以上10万未満	10.6	9.5	35.7	2.8	23.8	4.2
3万以上 5万未満	16.5	6.1	-	-	34.2	2.9
2万以上 3万未満	22.8	4.4	-	-	-	-
1万以上 2万未満	37.2	2.7	-	-	-	-
5千以上 1万未満	54.2	1.8	-	-	-	-
2千以上 5千未満	74.7	1.3	-	-	-	-
2千未満	85.5	1.2	-	-	-	-

結果の推定方法

抽出層世帯の結果は、当該市区町村の抽出率の逆数を乗率として線形推定により集計した。また、しっ皆層世帯については、集計値をそのまま用いた。

したがって、ある市区町村のある属性を有する人口又は世帯数の推定結果を得るための計算式は、次のとおりとなる。

$$\hat{X}_k = \left(\frac{N_{k1}}{n_{k1}} \cdot \sum_{i1} x_{i1} \right) + \sum_{i2} x_{i2}$$

また、全国又は都道府県における推計結果を得るための計算式は、次のとおりとなる。

$$\hat{X} = \sum_k \hat{X}_k = \sum_k \left\{ \left(\frac{N_{k1}}{n_{k1}} \cdot \sum_{i1} x_{i1} \right) + \sum_{i2} x_{i2} \right\}$$

ここで、

- k : 市区町村
- $i1$: 標本層世帯の抽出人員 (又は世帯)
- $i2$: しっ皆層世帯の人員 (又は世帯)
- \hat{X}_k : k 市区町村における推定値
- N_{k1} : k 市区町村における標本層世帯の総人員(又は世帯数)
- n_{k1} : k 市区町村における標本層世帯の総抽出人員(又は世帯数)
- x_{i1} : 標本層世帯の抽出人員ごと(又は世帯ごと)の当該属性
(当該属性に該当する場合は1, しない場合は0。以下同じ。)
- x_{i2} : しっ皆層世帯の人員ごと(又は世帯ごと)の当該属性
- \hat{X} : 全国又は都道府県の推定値

を示す。

なお、結果表章に当たっては、1の位の数字を四捨五入して10単位としたので、総数と内訳を合計した数値とは必ずしも一致しない。

推定値の精度

抽出詳細集計による結果は、標本によって得られた推定値であるため、標本誤差を含んでおり、全数集計すれば得られるはずの値とは必ずしも一致しない。

標本誤差は推定値の大きさや集計項目の種類によって異なるため、ここでは推定値の大きさに対する標準誤差率の目安(既定の標本数を単純無作為抽出したと仮定)を便宜、表2に示す。

この標準誤差率は、次の式により計算した。

(市区町村)

$$CV(\hat{X}_k) = \frac{1}{N_k \cdot \hat{p}_k} \sqrt{N_{k1} \cdot (N_{k1} - n_{k1}) \cdot \frac{\hat{p}_{k1}(1 - \hat{p}_{k1})}{n_{k1}}}$$

(全国又は都道府県)

$$CV(\hat{X}) = \frac{1}{N \cdot \hat{p}} \cdot \sqrt{\sum_k N_{k1} \cdot (N_{k1} - n_{k1}) \cdot \frac{\hat{p}_{k1}(1 - \hat{p}_{k1})}{n_{k1}}}$$

ここで、

- N_k : k 市区町村における総人員(標本層世帯及びしつ皆層世帯の人員の合計)
- \hat{p}_k : k 市区町村における総人員に占める推定値の比率 (= \hat{X}_k/N_k)
- \hat{p}_{k1} : k 市区町村における標本層世帯に占める推定値の比率 (= \hat{X}_{k1}/N_{k1})
- N : 全国又は都道府県における総人員(標本層世帯及びしつ皆層世帯の人員の合計)
- \hat{p} : 全国又は都道府県における総人員に占める推定値の比率 (= \hat{X}/N)
- \hat{X}_{k1} : k 市区町村における標本層世帯の推定値

$$\hat{X}_{k1} = \left(\frac{N_{k1}}{n_{k1}} \cdot \sum_{i1} x_{i1} \right)$$

を示す。

標準誤差率は、全数集計すれば得られるはずの値(真の値)の存在範囲を示す目安となるものである。すなわち、推定値を中心として、その前後に、標準誤差(推定値の大きさ×標準誤差率)の幅をとれば、その区間内に真の値があることが約68%の確率で期待され、また、標準誤差の2倍の幅をとれば、その区間内に真の値があることが約95%の確率で期待される。

例えば、全国の300,000という結果数値は、真の値が300,000±300,000(推定値の大きさ)×0.00623(標準誤差率)、すなわち、298,131から301,869までの間にあることが約68%の確率で期待され、300,000±300,000×0.00623×2、すなわち、296,262から303,738までの間にあることが約95%の確率で期待される。

なお、これらは、属性別人口及び世帯数に対して適用され、1世帯当たり人員など、平均の推定値には適用できない。

Chapter3 METHOD OF SAMPLE TABULATION AND RELIABILITY OF ESTIMATES

Sample Design

- (a) Private households and institutional households with less than 30 persons were sampled.
- (b) Institutional households with 30 persons or more, residents in camps of Self-Defense Forces, inmates of reformatory institutions, etc. were complete count.

The sampling ratios used for (a) were determined by the sampling size according to respective population size of municipality (*shi,ku,machi,mura*) to take into consideration of the accuracy. The results were shown in the following table.

Table1. Average Sampling Ratio for (a)

	Sampling Ratio	
	%	1/n
Japan	9.9	10.1

Population Size	Sampling Ratio					
	<i>Shi,Machi, or Mura</i>		<i>Ku (Tokyo)</i>		<i>Ku (20 Major Cities)</i>	
	%	1/n	%	1/n	%	1/n
500,000 inhabitants or more	5.3	19.0	5.3	19.0	5.3	19.0
300,000 to 499,999 inhabitants	5.0	20.2	8.5	11.8	5.0	20.2
200,000 to 299,999 inhabitants	7.2	13.8	12.2	8.2	7.2	13.8
100,000 to 199,999 inhabitants	13.5	7.4	21.7	4.6	13.5	7.4
50,000 to 99,999 inhabitants	10.6	9.5	35.7	2.8	23.8	4.2
30,000 to 49,999 inhabitants	16.5	6.1	—	—	34.2	2.9
20,000 to 29,999 inhabitants	22.8	4.4	—	—	—	—
10,000 to 19,999 inhabitants	37.2	2.7	—	—	—	—
5,000 to 9,999 inhabitants	54.2	1.8	—	—	—	—
2,000 to 4,999 inhabitants	74.7	1.3	—	—	—	—
under 2,000 inhabitants	85.5	1.2	—	—	—	—

Method of Estimation

The results of the households (a) were estimated as the product of the sample count by the reciprocal of the sampling ratio of each municipality.

However, the result of (b) was obtained by complete count.

An estimate \hat{X}_k or \hat{X} , that is, the number of persons or households for private households only having the characteristics concerned in an area is obtained as follows.

(Municipality)

$$\hat{X}_k = \left(\frac{N_{k1}}{n_{k1}} \cdot \sum_{i1} x_{i1} \right) + \sum_{i2} x_{i2}$$

(Japan or prefectures)

$$\hat{X} = \sum_k \hat{X}_k = \sum_k \left\{ \left(\frac{N_{k1}}{n_{k1}} \cdot \sum_{i1} x_{i1} \right) + \sum_{i2} x_{i2} \right\}$$

Where

- k : k -th municipality in the area concerned
- $i1$: A person or a household of (a)
- $i2$: A person or a household of (b)
- \hat{X}_k : An estimate of k -th municipality
- N_{k1} : Total population or household of (a) in the k -th municipality
- n_{k1} : Number of sample persons or households of (a) in the k -th municipality
- x_{i1} : Number of persons or households of (a) having the characteristics concerned counted for the k -th municipality
- x_{i2} : Number of persons or households of (b) having the characteristics concerned counted for the k -th municipality
- \hat{X} : Number of sample persons or households of (a) in Japan or prefectures

In addition, figures given in the tables may not necessarily add up to the total due to rounding.

Sampling Error of Estimate

The figures obtained from the sample tabulation mentioned above are not consistent with those obtained from the complete count due to sampling error.

Though magnitude of sampling errors varies with the size of estimates and the kind of items, the ratios of standard error by size of estimates calculated under the assumption of simple random sampling of household members are shown in the table below for convenience.

The ratios of standard error have been calculated by the following formula.

(Municipality)

$$CV(\hat{X}_k) = \frac{1}{N_k \cdot \hat{p}_k} \sqrt{N_{k1} \cdot (N_{k1} - n_{k1}) \cdot \frac{\hat{p}_{k1}(1 - \hat{p}_{k1})}{n_{k1}}}$$

(Japan or prefectures)

$$CV(\hat{X}) = \frac{1}{N \cdot \hat{p}} \cdot \sqrt{\sum_k N_{k1} \cdot (N_{k1} - n_{k1}) \cdot \frac{\hat{p}_{k1}(1 - \hat{p}_{k1})}{n_{k1}}}$$

Where

- N_k : Total population or household in the k -th municipality
- \hat{p}_k : Ratio of estimate to the total population or household in the k -th municipality (= \hat{X}_k/N_k)
- \hat{p}_{k1} : Ratio of estimate to the total population or household of (a) in the k -th municipality (= \hat{X}_{k1}/N_{k1})
- N : Total people or total household in Japan or prefectures

- \hat{p} : Ratio of estimate to the total population or household in Japan or prefectures(= \hat{X}/N)
- \hat{X}_{k1} : Number of sample persons or households of (a) in the k -th municipality

$$\hat{X}_{k1} = \left(\frac{N_{k1}}{n_{k1}} \cdot \sum_{i1} x_{i1} \right)$$

表2 推定値(人口及び世帯)の大きさ別標準誤差率(全国の結果)

Table2 Ratio of Standard Error by Size of Estimates (Japan)

推定値の大きさ Size of estimates	標準誤差率 Ratio of Standard Error	
	人口 Population	世帯 Household number
	10,000,000	0.00104
8,000,000	0.00117	0.00113
6,000,000	0.00136	0.00134
5,000,000	0.00150	0.00148
4,000,000	0.00168	0.00167
3,000,000	0.00195	0.00195
2,000,000	0.00240	0.00241
1,000,000	0.00340	0.00345
800,000	0.00381	0.00386
600,000	0.00440	0.00447
400,000	0.00539	0.00548
300,000	0.00623	0.00633
200,000	0.00763	0.00776
150,000	0.00881	0.00897
100,000	0.01080	0.01099
80,000	0.01207	0.01229
60,000	0.01394	0.01419
40,000	0.01708	0.01739
30,000	0.01972	0.02008
20,000	0.02415	0.02460
15,000	0.02789	0.02840
10,000	0.03416	0.03479
8,000	0.03819	0.03889
6,000	0.04410	0.04491
4,000	0.05401	0.05500
3,000	0.06237	0.06351
2,000	0.07638	0.07779
1,500	0.08820	0.08982
1,000	0.10802	0.11001
800	0.12077	0.12300
600	0.13946	0.14203
400	0.17080	0.17394
300	0.19722	0.20085
200	0.24155	0.24600
150	0.27892	0.28405
100	0.34160	0.34789