# Appendix:
# Sampling Variability

Estimates based on sample data may differ from the figures that would have been obtained had all, rather than specified samples, of the records been used. These differences are termed sampling variability. The standard error is a measure of the sampling variability. About 68 percent of all possible probability samples selected with the same specifications will give estimates within one standard error of the figure obtained from a compilation of all records. Similarly, approximately 90 percent will give estimates within 1.645 standard errors, about 95 percent will give estimates within two standard errors, and about 99 percent will give estimates within two and one-half standard errors. The standard error of an estimate depends on the design elements such as the method of sampling, sample size, and the estimation process.

Because of the large number of data cells tabulated from the sample files, it is not practical to calculate the standard error for every possible cell. However, standard errors for a large number of cells were estimated. These estimates were used to fit regression curves to provide estimates of approximate standard errors associated with tabulated counts and proportions.

The tables showing the sampling variability provide a general order of magnitude for similar estimates from the various sample files. Table A-1 presents approximate standard errors for the estimated number of persons from the 1 percent and the 10 percent files. The reliability of an estimated percentage depends on both the size of the percentage and on the size of the total on which the percentage is based. Data in Table A-2 provide approximations of the standard errors of the estimated percentage of persons in the 1 percent and 10 percent files. The standard errors are expressed in percentage points, and the bases shown are in terms of inflated data.

**Table A-1.**
**Approximations of standard errors of estimated number of persons**

| Size of estimate (inflated) | Standard error |
|---|---|
| *1 percent file* | |
| 500 | 250 |
| 1,000 | 300 |
| 2,500 | 500 |
| 5,000 | 800 |
| 7,500 | 900 |
| 10,000 | 1,100 |
| 25,000 | 1,700 |
| 50,000 | 2,400 |
| 75,000 | 3,000 |
| 100,000 | 3,400 |
| 250,000 | 5,400 |
| 500,000 | 7,800 |
| 750,000 | 9,600 |
| 1,000,000 | 11,100 |
| 5,000,000 | 25,800 |
| 10,000,000 | 36,900 |
| 25,000,000 | 57,700 |
| 50,000,000 | 76,100 |
| 75,000,000 | 82,900 |
| *10 percent file* | |
| 100 | 30 |
| 500 | 70 |
| 1,000 | 100 |
| 5,000 | 225 |
| 10,000 | 300 |
| 50,000 | 700 |
| 100,000 | 1,000 |
| 500,000 | 2,200 |
| 1,000,000 | 3,200 |
| 2,000,000 | 4,300 |
| 3,000,000 | 5,300 |
| 5,000,000 | 6,500 |
| 10,000,000 | 8,500 |
| 20,000,000 | 9,300 |

**Table A-2.**
**Approximations of standard errors of estimated percentage of persons**

| Size of base (inflated) | 2 or 98 | 5 or 95 | 10 or 90 | 25 or 75 | 50 |
|---|---|---|---|---|---|
| *1 percent file* | | | | | |
| 1,000 | 4.7 | 7.3 | 10.1 | 14.5 | 16.8 |
| 10,000 | 1.5 | 2.3 | 3.2 | 4.6 | 5.3 |
| 50,000 | 0.7 | 1.0 | 1.4 | 2.1 | 2.4 |
| 100,000 | 0.5 | 0.7 | 1.0 | 1.5 | 1.7 |
| 500,000 | 0.2 | 0.3 | 0.4 | 0.7 | 0.8 |
| 1,000,000 | 0.1 | 0.2 | 0.3 | 0.5 | 0.5 |
| 5,000,000 | 0.1 | 0.1 | 0.1 | 0.2 | 0.2 |
| 10,000,000 | a | 0.1 | 0.1 | 0.2 | 0.2 |
| 50,000,000 | a | a | a | 0.1 | 0.1 |
| 100,000,000 | a | a | a | a | a |
| *10 percent file* | | | | | |
| 500 | 1.9 | 3.0 | 4.1 | 5.9 | 6.8 |
| 1,000 | 1.3 | 2.1 | 2.9 | 4.1 | 4.8 |
| 2,500 | 0.8 | 1.3 | 1.8 | 2.6 | 3.0 |
| 10,000 | 0.4 | 0.6 | 0.9 | 1.3 | 1.5 |
| 50,000 | 0.2 | 0.3 | 0.4 | 0.6 | 0.7 |
| 100,000 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |
| 500,000 | a | 0.1 | 0.1 | 0.2 | 0.2 |
| 1,000,000 | a | 0.1 | 0.1 | 0.1 | 0.2 |
| 5,000,000 | a | a | a | a | 0.1 |
| 10,000,000 | a | a | a | a | a |
| 50,000,000 | a | a | a | a | a |

a. Less than 0.05 percent.