Technical Appendix

This appendix addresses the most frequently encountered mathematical operations when using this book. The first segment involves manipulation of the data. The second segment addresses the reliability of the estimates. For more information on manipulating the data, please consult a mathematics or statistics textbook. For more information on calculating the reliability of the estimates, please consult the technical documentation for the March 2015 Survey at http://www2.census.gov/programs-surveys/cps/techdocs/cpsmar15.pdf.

Manipulating the data

Converting a percentage of a population to a count of units

First, divide the percentage by 100. Then multiply that decimal by the total population.

Example: How many aged units 65 or older have income from Veterans' benefits?

In Table 2.A1, 5.0 percent of aged units 65 or older had income from Veterans' benefits, and there was a total of 34,614,000 aged units 65 or older. Dividing the percentage by 100 yields: 5.0 ÷ 100 = 0.050. Then multiply: 0.050 × 34,614,000 = 1,730,700.

Combining two percentage distributions

First, convert each percentage to a count of units. Then add the two counts of interest. Finally, divide by the sum of the two total populations.

Example: What percentage of aged units 65–74 had total money income of $15,000–$19,999?

In Table 3.A1, 7.5 percent of aged units 65–69 and 9.7 percent of aged units 70–74 had total money income of $15,000–$19,999. There were a total of 11,056,000 aged units 65–69 and 8,053,000 aged units 70–74.

First, find the number of aged units with total money income of $15,000–$19,999:

0.075 × 11,056,000 = 829,200 aged units 65–69 had total money income of $15,000–$19,999

0.097 × 8,053,000 = 781,141 aged units 70–74 had total money income of $15,000–$19,999

829,200 + 781,141 = 1,610,341 aged units 65–74 had total money income of $15,000–$19,999

Second, find the total population:

11,056,000 (aged units 65–69) + 8,053,000 (aged units 70–74) = 19,109,000 aged units 65–74

Finally, divide the population of interest by the total population:

1,610,341 ÷ 19,109,000 = 0.084 or 8.4 percent of aged units 65–74 had total money income of $15,000–$19,999.

Note: This procedure cannot be used on medians or some means presented in this publication.

Estimating a particular percentile limit

This is also known as getting a cumulative distribution from a frequency distribution. Add percentages in the frequency distribution (column) until you exceed the percentile limit you want. Then interpolate within that last interval to estimate your desired percentile (see example below).

Example: What was the Social Security income cutoff for the bottom decile (10 percent) of beneficiary aged units 65 or older?

In Table 5.A1, get the total percent (cumulative distribution) by adding up the percents in the aged units 65 or older column until you exceed 10 percent. Because 8,000–8,999 is the first row to exceed 10 percent total, the 10-percent limit is between $8,000 and $8,999.

Next look at the total percent immediately lower than 10 percent (here it's 8.4). So, 10 − 8.4 = 1.6 means that you need 1.6 percentage points more of the population. There are 2.8 percentage points in the 8,000–8,999 category. Take the proportion 1.6/2.8 (what you need/what you have) and multiply it by 1,000 (the total number of dollars for the row category). (1.6/2.8) × 1,000 = $571. Add 571 to 8,000 (the bottom dollar for the row). The bottom decile limit is 8,571.

Social Security (dollars)	Percent	Social Security (dollars)	Total percent
1–499	0.1	< 500	0.1
500–999	0.2	< 1,000	0.3
1,000–1,499	0.3	< 1,500	0.6
1,500–1,999	0.3	< 2,000	0.9
2,000–2,499	0.4	< 2,500	1.3
2,500–2,999	0.3	< 3,000	1.6
3,000–3,499	0.5	< 3,500	2.1
3,500–3,999	0.5	< 4,000	2.6
4,000–4,499	0.3	< 4,500	2.9
4,500–4,999	0.6	< 5,000	3.5
5,000–5,999	0.9	< 6,000	4.4
6,000–6,999	1.7	< 7,000	6.1
7,000–7,999	2.3	< 8,000	8.4
8,000–8,999	2.8	< 9,000	11.2

Reliability of the Estimates

Because the figures in this report are based on a sample of the older population, all reported statistics (counts, percentages, and medians) are only estimates of population parameters and may deviate somewhat from their true values—that is, from the values that would have been obtained from a complete census using the same questionnaires, instructions, and interviewers.

The standard error is primarily a measure of sampling variability—that is, it measures the variations that occur by chance because a sample rather than the entire population is surveyed. As calculated for this report, the standard error also partly measures the effect of response and enumeration errors but does not measure systematic biases in the data. The chances are about 68 out of 100 that an estimate for the sample would differ from a complete census figure by less than the standard error. The chances are about 95 out of 100 that the difference would be less than twice the standard error.

Standard Error of Estimated Percentages

The reliability of an estimated percentage, computed by using sample data for both numerator and denominator, depends on both the size of the percentage and the size of the total on which the percentage is based. The approximate standard error S_x of an estimated percentage can be obtained using the formula

$$s_{x,p} = \sqrt{\frac{b}{x} p(100 - p)}$$

Here *x* is the total number of persons, families, or households (the base of the percentage), *p* is the percentage, and *b* is the parameter from the following table associated with the characteristic in the numerator of the percentage.

Characteristics	Total or white	Black	Asian	Hispanic
Below poverty level	2,441	2,441	2,441	2,441
All income levels	1,526	1,747	1,747	1,747
People by family income	3,047	3,488	3,488	3,488

Use of this formula in calculating the standard error of a single percentage is illustrated as follows:

> An estimated 44.7 percent of units aged 65 or older had total money income of $30,000 or more in 2014 (Table 3.A1). Because the base of this percentage is approximately 34,614,000—the number of units aged 65 or older—the standard error of the estimated 44.7 percent is approximately 0.3 percent. The chances are 68 out of 100 that the estimate would have shown a figure that differed from one resulting from a complete census by less than 0.3 percent. The chances are 95 out of 100 that the estimate would have shown a figure differing from one after a complete census by less than 0.6 percent—that is, this 95 percent confidence interval would range from 44.1 percent to 45.3 percent.

For a difference between two sample estimates, the standard error is approximately equal to the square root of the sum of the squares of the standard errors of each estimate considered separately. This formula will represent the actual standard error quite accurately for the difference between separate and uncorrelated characteristics. If, however, there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error.

A comparison of the difference in the percentage of units aged 62 to 64 and 65 or older who had total money income of $30,000 or more in 2014 illustrates how to calculate the standard error of a difference between two percentages:

> 44.7 percent of the 34,614,000 units aged 65 or older and 57.5 percent of the 7,673,000 units aged 62 to 64 had total money income of $30,000 or more in 2014 (Table 3.A1)—a difference of 12.8 percentage points. The standard errors of those percentages are 0.3 and 0.7, respectively. The standard error of the estimated difference of 12.8 percentage points is about
>
> $$0.8 = \sqrt{(0.3)^2 + (0.7)^2}$$
>
> The chances are 68 out of 100 that the difference is between 12.0 and 13.6 percentage points and 95 out of 100 that it is between 11.2 and 14.4 percentage points. Because the confidence interval around the difference does not include zero, there is a statistically significant difference between the proportions of units who are aged 62 to 64 and those who are aged 65 or older with income of $30,000 or more.

Confidence Limits of Medians

The sampling variability of an estimated median depends on the distribution as well as on the size of the base. Confidence limits of a median based on sample data may be estimated as follows: (1) using the appropriate base, the standard error of a 50 percent characteristic is determined; (2) the standard error determined in step 1 is added to and subtracted from 50 percent; and (3) the confidence interval around the median corresponding to the two points estimated in step 2 is then read from the distribution of the characteristic. A two-standard-error confidence limit may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error. This procedure may be illustrated as follows:

> The median total money income of the estimated 34,614,000 units aged 65 or older was $30,193 in 2014 (Table 3.A1). The standard error of 50 percent of those units expressed as a percentage is about 0.33 percent. As interest usually centers on the confidence interval for the median at the two-standard-error level, it is necessary to add and subtract twice the standard error obtained in step 1 from 50 percent. This procedure yields limits of approximately 49.3 percent and 50.7 percent. By interpolation, 49.3 percent of units aged 65 or older had total money income below $29,106, and 50.7 percent had total money income below $30,764. Thus, the chances are about 95 out of 100 that the census would have shown the median to be greater than $29,106 but less than $30,764.