7. UNDERSTANDING ERROR AND DETERMINING STATISTICAL SIGNIFICANCE

Sources of Error

The data in American Community Survey (ACS) products are estimates of the actual figures that would have been obtained if the entire population—rather than the chosen ACS sample—had been interviewed using the same methodology. All estimates produced from sample surveys have uncertainty associated with them as a result of being based on a sample of the population rather than the full population. This uncertainty—called sampling error—means that estimates derived from the ACS will likely differ from the values that would have been obtained if the entire population had been included in the survey, as well as from values that would have been obtained had a different set of sample units been selected for the survey.

Sampling error is the difference between an estimate based on a sample and the corresponding value that would be obtained if the estimate were based on the entire population. Measures of the magnitude of sampling error reflect the variation in the estimates over all possible samples that could have been selected from the population using the same sampling methodology. The margin of error is the measure of the magnitude of sampling error provided with all published ACS estimates.

In addition to sampling error, data users should recognize that other types of error—called nonsampling error—might also be introduced during any of the complex operations used to collect and process ACS data. Nonsampling error can result from problems in the sampling frame or survey questionnaires, mistakes in how the data are reported or coded, issues related to data processing or weighting, or problems related to interviewer bias or nonresponse bias. Nonresponse bias results when survey respondents differ in meaningful ways from non-respondents. Nonsampling error may affect ACS data by increasing the variability of the estimates or introducing bias into ACS results. The U.S. Census Bureau tries to minimize nonsampling error through extensive research and evaluation of sampling techniques, questionnaire design, and data collection and processing procedures.

Nonsampling error is very difficult to measure directly, but the Census Bureau provides a number of indirect measures to help inform users about the quality of ACS estimates. The section on "Measures of Nonsampling Error" includes a more detailed description of the different types of nonsampling error in the ACS and measures of ACS data quality. More information on ACS data quality measures for the nation and individual states is available on the Census Bureau's Web page on Sample Size and Data Quality.[59]

Measures of Sampling Error

Margins of Error and Confidence Intervals

A margin of error (MOE) describes the precision of an ACS estimate at a given level of confidence. The confidence level associated with the MOE indicates the likelihood that the ACS sample estimate is within a certain range (the MOE) of the population value. The MOEs for published ACS estimates are provided at a 90 percent confidence level. From these MOEs, data users can easily calculate 90 percent confidence intervals that define a range expected to contain the *true* or population value of an estimate 90 percent of the time. For example, in the Data Profile for Selected Social Characteristics (Table DP02) for Colorado, a portion of which is shown in Table 7.1, data from the 2015 ACS 1-year estimates indicate that there were 564,757 one-person households in the state in 2015 with an MOE of 10,127. By adding and subtracting the MOE from the point estimate, we can calculate the 90 percent confidence interval for that estimate:

$$564,757 - 10,127 = 554,630 = \text{Lower bound of the interval}$$

$$564,757 + 10,127 = 574,884 = \text{Upper bound of the interval}$$

---

[59] U.S. Census Bureau, American Community Survey, Sample Size and Data Quality, <www.census.gov/acs/www/methodology /sample-size-and-data-quality/>.

Table 7.1. Sample Estimates and Margins of Error in American FactFinder: 2015

Subject	Colorado			
	Estimate	Margin of Error	Percent	Percent Margin of Error
HOUSEHOLDS BY TYPE				
Total households	2,074,735	+/-7,548	2,074,735	(X)
Family households (families)	1,331,861	+/-11,075	64.2%	+/-0.5
With own children of the householder under 18 years	597,501	+/-8,075	28.8%	+/-0.4
Married-couple family	1,038,040	+/-9,389	50.0%	+/-0.4
With own children of the householder under 18 years	435,028	+/-7,158	21.0%	+/-0.3
Male householder, no wife present, family	93,024	+/-5,026	4.5%	+/-0.2
With own children of the householder under 18 years	48,969	+/-4,185	2.4%	+/-0.2
Female householder, no husband present, family	200,797	+/-6,582	9.7%	+/-0.3
With own children of the householder under 18 years	113,504	+/-4,699	5.5%	+/-0.2
Nonfamily households	742,874	+/-10,127	35.8%	+/-0.5
Householder living alone	564,757	+/-10,127	27.2%	+/-0.5
65 years and over	182,959	+/-5,249	8.8%	+/-0.3
Households with one or more people under 18 years	657,324	+/-8,777	31.7%	+/-0.4
Households with one or more people 65 years and over	497,903	+/-4,124	24.0%	+/-0.2

Source: U.S. Census Bureau, American FactFinder, Table DP02: Selected Social Characteristics in the United States.

Therefore, we can be 90 percent confident that the true number of one-person households in Colorado in 2015 falls somewhere between 554,630 and 574,884. Put another way, if the ACS were independently conducted 100 times in 2015, sampling theory suggests that 90 times the estimate of one-person households in Colorado would fall in the given confidence interval. Estimates with smaller MOEs—relative to the value of the estimate—will have narrower confidence intervals indicating that the estimate is more precise and has less sampling error associated with it.

> TIP: When constructing confidence intervals from MOEs, data users should be aware of any "natural" limits on the upper and lower bounds. For example, if a population estimate is near zero, the calculated value of the lower confidence bound may be less than zero. However, a negative number of people does not make sense, so the lower confidence bound should be reported as zero instead.
>
> For other estimates, such as income, negative values may be valid. Another natural limit would be 100 percent for the upper confidence bound of a percent estimate. Data users should always keep the context and meaning of an estimate in mind when creating and interpreting confidence intervals.

Standard Errors and Coefficients of Variation

A standard error (SE) measures the variability of an estimate due to sampling and provides the basis for calculating the MOE. The SE provides a quantitative measure of the extent to which an estimate derived from a sample can be expected to deviate from the value for the full population. SEs are needed to calculate coefficients of variation and to conduct tests of statistical significance. Data users can easily calculate the SE of an ACS estimate by dividing the positive value of its MOE by 1.645 as shown below:[60]

$$\text{SE} = \frac{\text{MOE}}{1.645} \tag{1}$$

Using the data in Table 7.1, the SE for the number of one-person households in Colorado in 2015 would be:

$$\frac{10,127}{1.645} = 6,156$$

---

[60] Data users working with ACS 1-year estimates for 2005 or earlier should divide the MOE by the value 1.65 as that was the value used to derive the published MOE from the SE in those years.

The SE for an estimate depends on the underlying variability in the population for that characteristic and the sample size used for the survey. In general, the larger the sample size, the smaller the SE of the estimates produced from the sample data. This relationship between sample size and SE is the reason that ACS estimates for less populous areas are only published using multiple years of data. Combining data from multiple ACS 1-year files increases sample size and helps to reduce SEs.

Coefficients of variation are another useful measure of sampling error. A coefficient of variation (CV) measures the *relative* amount of sampling error that is associated with a sample estimate. The CV is calculated as the ratio of the SE for an estimate to the estimate itself $(\widehat{X})$ and is usually expressed as a percent:

$$CV = \frac{SE}{\widehat{X}} * 100 \tag{2}$$

A small CV indicates that the SE is small relative to the estimate, and a data user can be more confident that the estimate is close to the population value. The CV is also an indicator of the reliability of an estimate. When the SE of an estimate is close to the value of the estimate, the CV will be larger, indicating that the estimate has a large amount of sampling error associated with it and is not very reliable. For the example of one-person households in Colorado, the CV would be calculated as:[61]

$$\frac{6,156}{564,757} * 100 = 0.011 * 100 = 1.1\%$$

A CV of 1.1 percent indicates that the ACS estimate of one-person households in Colorado has a relatively small amount of sampling error and is quite reliable. Data users often find it easier to interpret and compare CVs across a series of ACS estimates than to interpret and compare SEs.

There are no hard-and-fast rules for determining an acceptable range of error in ACS estimates. Instead, data users must evaluate each application to determine the level of precision that is needed for an ACS estimate to be useful. For more information, visit the Census Bureau's Web page on Data Suppression.[62]

Determining Statistical Significance

One of the most important uses of ACS data is to make comparisons between estimates—across different geographic areas, different time periods, or different population subgroups. Data users may also want to compare ACS estimates with data from past decennial censuses. For any comparisons based on ACS data, it is important to take into account the sampling error associated with each estimate through the use of a statistical test for significance. This test shows whether the observed difference between estimates likely represents a true difference that exists within the full population (is statistically significant) or instead has occurred by chance because of sampling (is not statistically significant). Statistical significance means that there is strong statistical evidence that a true difference exists within the full population. Data users should not rely on overlapping confidence intervals as a test for statistical significance because this method will not always provide an accurate result.

---

[61] The examples provided in this section use unrounded values in their calculations, but the values displayed are rounded to three decimal places. For percentages that use a numerator or denominator from a prior example, the unrounded value is used, although the rounded value is displayed.

[62] U.S. Census Bureau, American Community Survey, Data Suppression, <www.census.gov/programs-surveys/acs/technical-documentation/data-suppression.html>.

When comparing two ACS estimates, a test for significance can be carried out by making several calculations using the estimates and their corresponding SEs. These calculations are straightforward given the published MOEs available for ACS estimates in American FactFinder (AFF) and many other Census Bureau data products. The steps to test for a statistically significant difference between two ACS estimates are as follows:

1. Calculate the SEs for the two ACS estimates using formula (1).
2. Square the resulting SE for each estimate.
3. Sum the squared SEs.
4. Calculate the square root of the sum of the squared SEs.
5. Divide the difference between the two ACS estimates by the square root of the sum of the squared SEs.
6. Compare the absolute value of the result from Step 5 with the critical value for the desired level of confidence (1.645 for 90 percent, 1.960 for 95 percent, or 2.576 for 99 percent).
7. If the absolute value of the result from Step 5 is greater than the critical value, then the difference between the two estimates can be considered statistically significant, at the level of confidence corresponding to the critical value selected in Step 6.

Algebraically, the significance test can be expressed as follows:

$$\text{If} \quad \left| \frac{\hat{X}_1 - \hat{X}_2}{\sqrt{\left[SE(\hat{X}_1)\right]^2 + \left[SE(\hat{X}_2)\right]^2}} \right| > Z_{CL} \tag{3}$$

then the difference between estimates $\hat{X}_1$ and $\hat{X}_2$ is statistically significant at the specified confidence level (CL)

where $\hat{X}_1$ and $\hat{X}_2$ are the estimates being compared
SE_1 is the SE for estimate $\hat{X}_1$
SE_2 is the SE for estimate $\hat{X}_2$

Z_{CL} is the critical value for the desired confidence level (1.645 for 90 percent, 1.960 for 95 percent, and 2.576 for 99 percent).

The example below shows how to determine if the difference in the estimated percentage of householders age 65 or older who live alone between Florida (estimated percentage = 12.6, MOE = 0.2) and Arizona (estimated percentage = 10.5, MOE = 0.3) is statistically significant, based on 2015 ACS data. Using formula (1) above, first calculate the corresponding standard errors for Florida (0.122) and Arizona (0.182) by dividing the MOEs by 1.645. Then, using formula (3) above, calculate the test value as follows:

$$\left| \frac{\widehat{X}_1 - \widehat{X}_2}{\sqrt{\left[SE(\widehat{X}_1) \right]^2 + \left[SE(\widehat{X}_2) \right]^2}} \right| = \left| \frac{12.6 - 10.5}{\sqrt{(0.122)^2 + (0.182)^2}} \right| = \left| \frac{2.1}{\sqrt{0.015 + 0.033}} \right| =$$

$$\left| \frac{2.1}{\sqrt{0.048}} \right| = \left| \frac{2.1}{0.219} \right| = 9.581$$

Since the test value (9.581) is greater than the critical value for a confidence level of 90 percent (1.645), the difference in the percentages is statistically significant at a 90 percent confidence level. A rough interpretation of the result is that the user can be 90 percent certain that a difference exists between the percentage of householders aged 65 or older who live alone in Florida and in Arizona.

By contrast, if the corresponding estimate for Indiana (estimated percentage = 10.8, MOE = 0.2, SE = 0.122) were compared with the estimate for Arizona, formula (3) would yield:

$$\left| \frac{\widehat{X}_1 - \widehat{X}_2}{\sqrt{\left[SE(\widehat{X}_1) \right]^2 + \left[SE(\widehat{X}_2) \right]^2}} \right| = \left| \frac{10.8 - 10.5}{\sqrt{(0.122)^2 + (0.182)^2}} \right| = \left| \frac{0.3}{\sqrt{0.015 + 0.033}} \right| =$$

$$\left| \frac{0.3}{\sqrt{0.048}} \right| = \left| \frac{0.3}{0.219} \right| = 1.369$$

Since the test value (1.369) is less than the critical value for a confidence level of 90 percent (1.645), the difference in percentages is not statistically significant. A rough interpretation of the result is that the user cannot be certain to any sufficient degree that the observed difference in the estimates between Indiana and Arizona was not due to chance.

The Census Bureau has produced a Statistical Testing Tool to make it easier for ACS data users to conduct tests of statistical significance when comparing ACS estimates.[63] This tool consists of an Excel spreadsheet that will automatically calculate statistical significance when data users are comparing two ACS estimates or multiple estimates. Data users simply need to download ACS data from the Census Bureau's AFF Web site and insert the estimate and MOE into the correct columns and cells in the spreadsheet. The results are calculated automatically. The result "Yes" indicates that estimates are statistically different and the result "No" indicates the estimates are *not* statistically different.[64]

Comparisons Within the Same Time Period

Comparisons involving two estimates from the same time period (e.g., from the same year or the same 5-year period) are straightforward and can be carried out as described in the previous section as long as the areas or groups are nonoverlapping (e.g., comparing estimates for two different counties, or for two different age groups). On the other hand, if the comparison involves a large area or group and a subset of the area or group (e.g., comparing an estimate for a state with the corresponding estimate for a county within the state, or comparing an estimate for all females with the corresponding estimate for African American females) then the two estimates may not be independent. In these cases, the data user may need to use a different approach that accounts for the correlation between the estimates in performing the statistical test of significance.

---

[63] U.S. Census Bureau, American Community Survey (ACS), Statistical Testing Tool, <www.census.gov/programs-surveys/acs/guidance /statistical-testing-tool.html>.

[64] This tool only conducts statistical testing on the estimates keyed in by the data user for comparison within the spreadsheet, and it does not adjust the MOE when making multiple comparisons, nor incorporate a Bonferroni correction or any other method in the results of the statistical testing.

Making Comparisons With Overlapping Multiyear Estimates

TIP: When comparing estimates from two multiyear periods, ideally comparisons should be based on nonoverlapping periods (e.g., comparing estimates from 2006–2010 with estimates from 2011–2015).

The comparison of two estimates for different, but overlapping periods is challenging since the difference is driven by the nonoverlapping years. For example, when comparing the 2010–2014 ACS 5-year estimates with the 2011–2015 ACS 5-year estimates, data for 2011 through 2014 are included in both estimates. Their contribution is subtracted out when the estimate of differences is calculated. While the interpretation of this difference is difficult, these comparisons can be made with caution. Under most circumstances, the estimate of difference should not be interpreted as a reflection of change between the last 2 years.

The use of MOEs for assessing the reliability of change over time is complicated when change is being evaluated using multiyear estimates. From a technical standpoint, change over time is best evaluated with multiyear estimates that do not overlap. At the same time, many data users will not want to wait until 2019 (when 2014–2018 ACS 5-year data will be available) to evaluate change since 2009–2013. Users who need to compare two 5-year estimates of the same geography that overlap in sample years must use a different formula to calculate the SE of this difference. To account for the sample overlap, use the following approximation to the SE:

$$SE\left(\hat{X}_1 - \hat{X}_2\right) \cong \sqrt{(1-C)}\sqrt{\left[SE(\hat{X}_1)\right]^2 + \left[SE(\hat{X}_2)\right]^2}$$

(4)

where C is the fraction of overlapping years. For example, the periods 2009–2013 and 2010–2014 overlap for 4 out of 5 years, so C=4/5=0.8. If the periods do not overlap, such as 2006–2010 and 2011–2015, then C=0.

With this SE, data users can test for the statistical significance of the difference between the two estimates using formula (3) with the modification from formula (4):

$$\sqrt{(1-C)}\sqrt{\left[SE(\hat{X}_1)\right]^2 + \left[SE(\hat{X}_2)\right]^2} \text{ for } \sqrt{\left[SE(\hat{X}_1)\right]^2 + \left[SE(\hat{X}_2)\right]^2}$$

Substitute in the denominator of formula (3) for the significance test.

The example below shows how to test for a statistically significant difference between the estimated percentages of householders aged 65 or older who lived alone in Holmes County, Florida, in 2010–2014 (estimated percentage = 13.1, MOE = 2.3, SE = 1.398) and in 2006–2010 (estimated percentage = 13.6, MOE = 2.3, SE = 1.398). In this example, only 1 of the 5 years overlaps, so C = 1/5=0.2. Using formula (3) with the modification from formula (4) yields:

$$\left|\frac{\hat{X}_1 - \hat{X}_2}{\sqrt{(1-C)}\sqrt{\left[SE(\hat{X}_1)\right]^2 + \left[SE(\hat{X}_2)\right]^2}}\right| = \left|\frac{13.1 - 13.6}{\sqrt{(1-0.2)}\sqrt{(1.398)^2 + (1.398)^2}}\right| =$$

$$= \left|\frac{-0.5}{\sqrt{0.8}\sqrt{1.955 + 1.955}}\right| = \left|\frac{-0.5}{0.894 * 1.977}\right| = \left|\frac{-0.5}{1.769}\right| = 0.283$$

Since the test value (0.283) is less than the critical value for a confidence level of 90 percent (1.645), the difference in the percentages is not statistically significant at a 90 percent confidence level.

Custom (User-Derived) Estimates

In some cases, data users will need to construct custom ACS estimates by combining data across multiple geographic areas or population subgroups, or it may be necessary to derive a new percentage, proportion, or ratio from published ACS data. For example, one way to address the issue of unreliable estimates for individual census tracts or block groups is to aggregate geographic areas, yielding larger samples and more reliable estimates. In such cases, additional calculations are needed to produce MOEs and SEs, and to conduct tests of statistical significance for the derived estimates. The section on "Calculating Measures of Error for Derived Estimates" provides detailed instructions on how to make these calculations.

Advanced users who are aggregating ACS estimates can use the Census Bureau's Variance Replicate Tables to produce MOEs for selected ACS 5-year Detailed Tables.[65] Users can calculate MOEs for aggregated data by using the variance replicates. Unlike available approximation formulas, this method results in an exact MOE by incorporating the covariance. More information about the Variance Replicate Tables is available in the section on "Calculating Measures of Error for Derived Estimates."

Some advanced data users will also want to construct custom ACS estimates from the Census Bureau's Public Use Microdata Sample (PUMS) files. Separate instructions for calculating SEs and conducting significance tests for PUMS estimates are available on the Census Bureau's PUMS Technical Documentation Web page.[66]

It is important for data users to remember that the error measures and statistical tests described in this section do not tell us about the magnitude of nonsampling errors. More information about those types of errors is available in the section on "Measures of Nonsampling Error."

Additional Background Information and Tools

Sample Size and Data Quality
<www.census.gov/acs/www/methodology/sample-size-and-data-quality/>
This Web page describes the steps the Census Bureau takes to ensure that ACS data are accurate and reliable. It also includes several measures of ACS data quality for the nation and states.

Statistical Testing Tool
<www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>
The Statistical Testing Tool is a spreadsheet that tests whether ACS estimates are statistically different from one another. Simply copy or download ACS estimates and their MOEs into the tool to get instant results of statistical tests.

Variance Replicate Tables
<www.census.gov/programs-surveys/acs/data/variance-tables.html>
Variance replicate estimate tables include estimates, MOEs, and 80 variance replicates for selected ACS 5-year Detailed Tables. The tables are intended for advanced users who are aggregating ACS data within a table or across geographic areas. Unlike available approximation formulas, this method results in an exact MOE by incorporating the covariance.

---

[65] U.S. Census Bureau, American Community Survey (ACS), Variance Replicate Tables, <www.census.gov/programs-surveys/acs/data/variance-tables.html>.
[66] U.S. Census Bureau, American Community Survey (ACS), PUMS Technical Documentation, <www.census.gov/programs-surveys/acs/technical-documentation/pums/documentation.html>.