Instructions for Applying Statistical Testing to American Community Survey Data

This document provides instructions on how to carry out statistical testing using American Community Survey (ACS) estimates. Worked examples are provided in a separate document located at: https://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html.

Obtaining ACS Data

Data users may obtain ACS data from data.census.gov. Data users may wish to consult the data.census.gov training provided through the Census Academy, which is located at: https://www.census.gov/data/academy/topics/data-census-gov.html.

See the sections "Creating Estimates and MOEs Using Microdata" and "Additional Methods to Obtain ACS Data" below for additional ways to obtain ACS data.

Basic Statistical Test

The test of statistical significance takes into account the difference between the two estimates as well as the standard errors of both estimates. For two estimates, A and B, with standard errors SE(A) and SE(B), let

$$Z = \frac{A - B}{\sqrt{SE(A)^2 + SE(B)^2}}$$

The difference between A and B is significant at the 90 percent confidence level if:

$$Z < -1.645$$
or
$$Z > 1.645$$

Otherwise, the difference is not significant at the 90 percent confidence level.

This means that there is less than a 10 percent chance that the difference between these two estimates would be as large or larger by random chance alone.

Converting ACS Margins of Errors to Standard Errors

All published ACS estimates on data.census.gov include the 90 percent margin of error (MOE). Use the MOE to calculate the standard error (SE) as:

$$\text{Standard Error} = \text{Margin of Error} / 1.645$$

Note that some estimates are controlled to the official Population Estimates. Data users will see five asterisks or stars (*****) in place of a number for the MOE. For these estimates, set both the MOE and SE to zero (0).

MOEs which have two or three stars or asterisks (** or ***) indicate that a statistical test is not appropriate.

Note also that for ACS data from 2005 and earlier, 1.65 should be used instead of 1.645 when converting the MOE to the SE.

Statistical Testing Tool

The Census Bureau provides a tool for data users to carry out statistical testing. Data users may compare two estimates to each other as well as comparing multiple estimates to one another. Data users do not need to convert MOEs to SEs. The tool handles that automatically.

The statistical testing tool may be found at: https://www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html

Notes on Carrying Out Statistical Testing

1. The Z-score listed above in the basic statistical test is the method used in determining statistical significance for the ACS Comparison Profiles published on data.census.gov.

   Note that data.census.gov publishes rounded estimates and MOEs. However, the statistical testing for the Comparison Profiles is calculated using unrounded SEs. Therefore, statistical results using the published estimates may not match the published significance results.

2. Users may choose to apply a confidence level different from 90 percent to their tests of statistical significance. For example, if $Z < -1.96$ or $Z > 1.96$, then the difference between A and B is significant at the 95 percent confidence level.

   How to convert the MOE from a 90 percent confidence level to a 95 percent confidence level may be found in the webinar entitled "Using American Community Survey (ACS) Estimates and Margins of Error" located at: https://www.census.gov/programs-surveys/acs/guidance/training-presentations/acs-moe.html

3. The Z-score test can be used for any types of estimates, such as counts, percentages, proportions, means, medians, etc. It can be used for comparing across years, or across surveys. If one of the estimates is a fixed value or comes from a source without sampling error (such as a count from the 2010 Census), use zero for the standard error for that estimate in the Z-score equation.

4. Making and interpreting comparisons between ACS single-year and multiyear estimates is very difficult, and is not advised.

5. Using the rule of thumb of overlapping confidence intervals does not constitute a valid significance test and users should not use that method for determining whether estimates are statistically different from one another.

Approximating Standard Errors for Derived Estimates

Data users combining published estimates to create derived estimates will need to approximate the SE.

The approximation formulas listed below will also work for approximating MOEs. Data users do not need to convert the MOEs obtained from data.census.gov to SEs before using these formulas. Instead, simply replace the SEs in the formulas with the appropriate MOEs. If you multiply both sides of any of the SE formula below by 1.645, and then distribute the 1.645 appropriately within the square root, all of the SEs may be converted to MOEs.

Note that these approximations do not consider the correlation or covariance between the basic estimates. They may be overestimates or underestimates of the derived estimate's SE and MOE depending on whether the two basic estimates are highly correlated in either the positive or negative direction. As a result, the approximations are not expected to match MOEs published on data.census.gov.

1. **Sum or Difference of Estimates**

$$SE(A + B + \cdots) = SE(A - B - \cdots) = \sqrt{SE(A)^2 + SE(B)^2 + \cdots}$$

As the number of basic estimates involved in the sum or difference increases, the results of this formula become increasingly different from the standard error derived directly from the ACS microdata. Care should be taken to work with the fewest number of basic estimates as possible. If there are estimates involved in the sum that are controlled in the weighting then the approximate standard error can be tremendously different.

2. **Proportions and Percents**

Here we define a proportion as a ratio where the numerator is a subset of the denominator, for example the proportion of persons 25 and over with a high school diploma or higher.

Let P = A / B. Note that the proportion will range from zero to one.

$$SE(P) = \frac{1}{B}\sqrt{SE(A)^2 - P^2 \times SE(B)^2}$$

If the value under the square root sign is negative, then instead use the formula below. Note that this is the same formula used for a ratio.

$$SE(P) = \frac{1}{B}\sqrt{SE(A)^2 + P^2 \times SE(B)^2}$$

For the special case where the proportion, P, equals 1 use

$$SE(P) = \frac{SE(A)}{B}$$

Finally, to calculate a percentage, which ranges from 0% to 100%, simply multiply the proportion by 100. That is, PCT = 100% × P. To calculate the SE of the percent, multiply the SE of the proportion by 100%. That is, SE(PCT) = 100% × SE(P).

3. Means and Other Ratios

Ratios are similar to proportions, however, the numerator is not a subset of the denominator. The calculation of a mean may also be considered a ratio. Some example of ratios are persons per household and per capita income. In addition, use this formula for calculating the SE of percent change.

$$SE\left(\frac{A}{B}\right) = \frac{1}{B}\sqrt{SE(A)^2 + \left(\frac{A}{B}\right)^2 \times SE(B)^2}$$

4. Products

For a product of two estimates - for example if users want to estimate a numerator by multiplying a percent by its denominator - the standard error can be approximated as

$$SE(A \times B) = \sqrt{A^2 \times [SE(B)]^2 + B^2 \times [SE(A)]^2}$$

5. Using Multiple Approximations

Data users may need to use two or more approximations to obtain the SE or MOE for their estimate.

For example, data users may need to combine several estimates to create their desired numerator and denominator before calculating a percent.

For Example, let P = (A + B + C) / (D + E).

Before calculating the SE(P), you would first have to find SE(A+B+C) and SE(D+E). You would then use those SEs to calculate SE(P).

Calculating Standard Errors Using Variance Replicate Estimates Tables

For advanced users who want to obtain more accurate SEs for derived estimates, the ACS provides Variance Replicate Estimate (VRE) tables. These are augmented versions of published 5-year ACS detailed tables, for selected summary levels. Users may calculate SEs and MOEs of derived estimates with replicate estimates from the VRE tables using the same replicate variance methodology as the published ACS MOEs. This replicate method incorporates the covariance between estimates that the approximation methods described above miss.

VRE tables are located at: https://www.census.gov/programs-surveys/acs/data/variance-tables.html. The technical documentation may be found at: https://www.census.gov/programs-surveys/acs/technical-documentation/variance-tables.html.

A webinar on how to calculate MOEs using the replicate estimates in the VRE tables is located at: https://www.census.gov/data/academy/webinars/2020/calculating-margins-of-error-acs.html.

Creating Estimates and MOEs Using Microdata

Data users may wish to create estimates that are not available in an ACS data product. They may do so by using the ACS Public Use Microdata Sample (PUMS). The PUMS files are a sample of the full ACS microdata with additional disclosure avoidance measures applied. Data users may calculate their own estimates for individual characteristics. In addition, MOEs and SEs may be calculated using either the replicate variance method or using a generalized variance formula (GVF) using design factors.

Information on how to use PUMS data may be found on the PUMS technical documentation page: http://www.census.gov/programs-surveys/acs/technical-documentation/pums/documentation.html.

Data users may also use the free online Microdata Analysis Tool (MDAT) tool to calculate PUMS estimates. The tool is located at: https://data.census.gov/mdat/.

Additional Methods to Obtain ACS Data

In addition to data.census.gov, there are other places to obtain ACS data.

1. **Application Programming Interface**

   ACS data are also available directly through the application programming interface (API). More information may be found at: https://www.census.gov/data/developers.html.

   Training on the API is available through Census Academy. Data users may find the webinar "Demystifying the Census API", located at: https://www.census.gov/data/academy/webinars/2020/demystifying-the-census-api.html. As well as "How to Extract Data from the Census API" located at: https://www.census.gov/data/academy/data-gems/2018/api.html

2. ACS Summary Files

The ACS Summary File is a set of comma-delimited text files that contain the same data as the ACS detailed tables. Summary File data may be found here: https://www.census.gov/programs-surveys/acs/data/summary-file.html.

Documentation for how to access and use the Summary Files is available at: https://www.census.gov/programs-surveys/acs/technical-documentation/summary-file-documentation.html

3. Census Bureau Apps

The Census Bureau has released several apps which allow data users an easy method to obtain relevant data for their needs. Apps such as Census Business Builder, My Congressional District and My Tribal Area provide a select set of characteristics relevant to the target audience.

Data may be viewed in the app or downloaded. A list of apps may be found at: https://www.census.gov/data/data-tools.html

4. ACS Data on the FTP Site

Data for the ACS Summary Files, PUMS files and other data may be obtained via the appropriate Census FTP site. A list of the locations for various data are available at: https://www.census.gov/programs-surveys/acs/data/data-via-ftp.html