

Instructions for Applying Statistical Testing to ACS Data

What statements can you make in a report, presentation, etc. based on ACS results?

- The 2004 ACS estimate for the percent of grandparents living with and responsible for grandchildren is 65.1 for Arkansas.
- The 2004 estimate for Nebraska is 64.8 percent for the same characteristic.
- The 2003 estimate for Arkansas for the same characteristic is 61.2 percent.

When reporting these results, what comparisons can be stated as facts?

- Can you say “The ACS found a larger percent of grandparents living with and responsible for grandchildren in Arkansas than in Nebraska in 2004”?
- What about “for Arkansas, the 2004 estimate is higher than the 2003 estimate”?

It is Census policy to make these statements only if the comparisons are “statistically significant” at the 90 percent confidence level (or to note that the apparent difference is not statistically significant). This means that there is less than a 10 percent chance that the difference between the two estimates would be as large or larger by random chance alone. The determination of statistical significance takes into account the difference between the two estimates as well as the standard errors of both estimates. Users may choose to apply a confidence level different from 90 percent to their tests of statistical significance.

Users including ACS estimates in written reports and other documents need to be careful when making comparisons between two (or more) ACS estimates, between an ACS estimate and a fixed value, or between an ACS estimate and a value from another survey or other source. If users want to say

- “Estimate X is bigger than estimate Y.”
- “Estimate X this year is larger than estimate X last year.”
- “Estimate X is smaller than that value in Census 2000.”
- “State Z has the highest value of this characteristic.”

or any similar statement, users should first do the appropriate statistical test and determine whether the difference is statistically significant.

This document provides some basic instructions for performing the statistical testing, as well as obtaining the ACS standard errors needed to do statistical tests.

Obtaining Standard Errors

Where your standard errors come from, and whether they are readily available or you have to calculate them yourself, depends on where your ACS data is coming from. If the estimate of interest is published on American FactFinder (AFF), then AFF should also be the source of your standard errors. Possible sources for your data, and where to get standard errors:

1. ACS data from published tables on American FactFinder

All ACS estimates from tables on AFF include either the 90 percent margin of error or 90 percent confidence bounds. The margin of error is the maximum difference between the estimate and the upper and lower confidence bounds. Most tables on AFF containing 2005 or later ACS data display the margin of error.

Use the margin of error to calculate the standard error (dropping the “+/-” from the displayed value first) as:

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

where $Z = 1.65$ for 2005 and earlier years and 1.645 for 2006 and later years.

If confidence bounds are provided instead (as with most ACS data 2004 and earlier), calculate the margin of error first before calculating the standard error:

$$\text{Margin of Error} = \max(\text{upper bound} - \text{estimate}, \text{estimate} - \text{lower bound})$$

All published ACS estimates use 1.65 (for 2005 and previous years) or 1.645 (for 2006 and later years) to calculate 90 percent margins of error and confidence bounds. Other surveys may use other values.

2. ACS public-use microdata sample (PUMS) tabulations

Using the methods described in the 2006 Accuracy of the PUMS documentation, available at <http://www.census.gov/acs/www/Downloads/2006/AccuracyPUMS.pdf>, users can calculate Standard errors for tabulations from the 2006 ACS PUMS files. Users can find accuracy documents for previous year’s PUMS with appropriate instructions on the left hand side of this webpage: <http://www.census.gov/acs/www/Products/PUMS/codelist2006.html>.

NOTE: ACS PUMS design factors *should not* be used to calculate standard errors of full ACS sample estimates, such as those on AFF. Census 2000 design factors *should not* be used to calculate standard errors for *any* ACS estimate.

Obtaining Standard Errors for Derived Estimates

Once you have obtained standard errors for your basic estimates, there may be situations where you create derived estimates, such as percentages or differences, that also require standard errors.

All methods in this section are approximations and users should be cautious in using them. They may be overestimates or underestimates of the estimate’s standard error, and may not match direct calculations of standard errors or calculations obtained through other methods.

- Sum or Difference of Estimates

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

- 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 $P = \frac{A}{B}$.

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

If the value under the square root sign is negative, then instead use

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

If $P = 1$ then use

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

If $Q = 100\% \times P$ (a percent instead of a proportion), then $SE(Q) = 100\% \times SE(P)$.

- Means and Other Ratios

If the estimate is a ratio but the numerator is not a subset of the denominator, such as persons per household or per capita income, then

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

- Products

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

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

Users may combine these procedures for complicated estimates. For example, if the desired estimate is $P = \frac{A + B + C}{D + E}$, then SE(A+B+C) and SE(D+E) can be estimated first, and then those results used to calculate SE(P).

For examples of these formulas, please see the 2006 Accuracy of the Data document at <http://www.census.gov/acs/www/Downloads/2006/Accuracy.pdf>.

Instructions for Statistical Testing

Once standard errors have been obtained, doing the statistical test to determine significance is not difficult. For two estimates, A and B, with standard errors SE(A) and SE(B), let

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

If $Z < -1.645$ or $Z > 1.645$, then the difference between A and B is significant at the 90 percent confidence level. Otherwise, the difference is not significant.

This method can be used for any types of estimates: 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 the Census 2000 SF1), use zero for the standard error for that estimate in the above equation for Z.

This is the method used in determining statistical significance for the ACS Ranking Tables published on AFF. Note that your determination of statistical significance may not match the Ranking Table's result for the same pair of estimates, because the significance tests for the Ranking Tables are made using unrounded standard errors. Standard errors obtained from the rounded margins of error or confidence bounds are unlikely to match the unrounded standard error, and so statistical tests may differ.

Using the rule of thumb of overlapping confidence intervals does not constitute a valid significance test and users are discouraged from using that method.