Estimating Total Respondents Using Ratio Estimator Approach

Krishna Kumar, Shamayla Durrin and Denise Chang

Introduction

In this document, we use survey data from the 2022 American Community Survey (ACS) obtained through IPUMS. We apply the ratio estimator approach of Laplace to estimate the total number of respondents in each state based on the number of those with doctoral degrees. We compare these estimates to the actual number of respondents and discuss any differences.

How to Obtain the Data

- Go to https://www.ipums.org/
- Go to IPUMS USA
- Click 'Get Data'
- Click 'Select samples' and choose only '2022 ACS'
- Go to 'Household' -> 'Geographic' and click the + on 'STATEICP'
- Go to 'Person' -> 'Demographic' and click the + on 'SEX'
- Go to 'Person' -> 'Education' and click the + on 'EDUC'
- Click view cart
- Create data extract
- Change data format to .csv
- Submit extract
- You will then get an email when the data is ready to download
- Download the data in your directory and read it using the code below.

Overview of the Ratio Estimators Approach

The Ratio Estimator Approach calculates the ratio of two means.

The ratio estimator is a method used to estimate the total size or characteristic of a population based on a sample. By calculating the ratio between a known variable and a target variable (for example, the ratio of doctoral degree holders to the total number of respondents in a sample), this ratio is then applied to a subgroup to estimate its total size. The method assumes that the ratio is relatively stable across different subgroups. For example, if 10% of California respondents have a doctoral degree, we can apply this ratio to other states to estimate their total respondents from the number with doctoral degrees.

This approach dates back to Quetelet and Laplace, who used similar methods to estimate populations based on partial data. A well-known variant is the capture-recapture method, which is commonly used in ecology. In this technique, a sample is captured, marked, and then released back into the population. Later, a second sample is captured, and the proportion of marked individuals is used to estimate the total population size. The ratio estimator is powerful when direct measurement of the whole population is impractical, allowing for accurate inference from smaller, representative samples.

Finding Total Respondents In State Using The Ratio Estimator

We will use the ratio estimator approach by first calculating the proportion of respondents with doctoral degrees to the total respondents in California. This ratio will be used as a benchmark for all other states. By applying this ratio to the number of doctoral degree holders in each state, we can estimate the total number of respondents for each state. This approach assumes that the proportion of doctoral degree holders is similar across states, allowing us to scale up from the known subset to estimate the full population.

- First we count the total number of respondents with a doctoral degree

- Then we calculate the total number of respondents in California

- Then we do total number of respondents with doctoral/total number of respondents in california

- Then we find the estimated total using our ratio: we do total number of doctoral divided by our california ratio

- Then we compare the estimated values with the actual values

Table 1: Comparison of Estimated and Actual Respondents by State

State ICP	Number of Doctoral Respondents	Estimated Total Respondents	Actual Total Respondents	Difference
1	600	37042.708	37369	-326.2917
2	165	10186.745	14523	-4336.2552
3	2014	124340.024	73077	51263.0243
4	244	15064.035	14077	987.0347
5	177	10927.599	10401	526.5990
6	131	8087.658	6860	1227.6580
11	152	9384.153	9641	-256.8472
12	1438	88779.024	93166	-4386.9757
13	2829	174656.370	203891	-29234.6302
14	1620	100015.312	132605	-32589.6875
21	1457	89952.043	128046	-38093.9566
22	620	38277.465	69843	-31565.5347
23	991	61182.207	101512	-40329.7934
24	1213	74888.009	120666	-45777.9913
25	513	31671.516	61967	-30295.4844
31	258	15928.365	33586	-17657.6354
32	321	19817.849	29940	-10122.1510
33	572	35314.049	58984	-23669.9514
34	621	38339.203	64551	-26211.7969
35	153	9445.891	19989	-10543.1094
36	60	3704.271	8107	-4402.7292
37	71	4383.387	9296	-4912.6128

State ICP	Number of Doctoral Respondents	Estimated Total Respondents	Actual Total Respondents	Difference
40	1531	94520.644	88761	5759.6441
41	460	28399.410	51580	-23180.5903
42	251	15496.200	31288	-15791.8003
43	2731	168606.061	217799	-49192.9392
44	1451	89581.616	109349	-19767.3837
45	450	27782.031	45040	-17257.9688
46	263	16237.054	29796	-13558.9462
47	1421	87729.481	109230	-21500.5191
48	647	39944.387	54651	-14706.6128
49	3216	198548.917	292919	-94370.0833
51	448	27658.556	46605	-18946.4444
52	1608	99274.458	62442	36832.4583
53	281	17348.335	39445	-22096.6649
54	841	51921.530	72374	-20452.4705
56	159	9816.318	18135	-8318.6823
61	896	55317.111	74153	-18835.8889
62	1031	63651.720	59841	3810.7205
63	175	10804.123	19884	-9079.8767
64	113	6976.377	11116	-4139.6233
65	282	17410.073	30749	-13338.9271
66	350	21608.247	20243	1365.2465
67	428	26423.799	35537	-9113.2014

4

State ICP	Number of Doctoral Respondents	Estimated Total Respondents	Actual Total Respondents	Difference
68	72	4445.125	5962	-1516.8750
71	6336	391171.000	391171	0.0000
72	647	39944.387	43708	-3763.6128
73	1195	73776.727	80818	-7041.2726
81	51	3148.630	6972	-3823.3698
82	214	13211.899	14995	-1783.1007
98	311	19200.470	6718	12482.4705

Why are the Actual values and Estimated Values Different?

- The ratio is based on one state's data (California) so it must not uniformly apply to all states, as each state has different educational systems and each state has different demographics of people.

- The reason the estimates of total respondents differ from the actual numbers is due to our assumption that the ratio of respondents with doctoral degrees is constant across all states. In reality, this is not the case, as many states have varying levels of resources that impact higher education. Some states may lack access to higher education institutions, leading to a lower proportion of doctoral degree holders. Factors such as income disparities, state education policies, and local economic conditions can significantly influence the number of individuals pursuing advanced degrees. Consequently, states with fewer resources or less emphasis on higher education will have lower ratios of doctoral degree holders, causing our estimates to be off.