SAS Data Testing of Poverty Level and Child Dental Care

Ashley Rayhill

Portfolio Project – Option 2

CSU – Global Campus

MIS 500

December 1, 2019

Project Overview

For this project I have been able to combine a variety of the knowledge that I have gained over the last few weeks. I decided to do my research and overall final project in the SAS program, focusing on child dental care based on the poverty level. The SAS University Edition is a great free software that anyone can use to learn statistics and quantitative methods programming. SAS offers an interactive community with tutorials, support, and forums that help users have a consistent and easy-going experience. SAS is also a very sought out skill that many employers search for (). By preparing my statistical analyses within the SAS University Edition program I am able to learn and practice basic SAS skills from start to finish.

The SAS example code I used for this project analyzes child dental care amongst poverty level. I selected this code example because I thought it would be interesting to analyze, however instead of analyzing the 2016 data I wanted to analyze the updated 2017 version. I was able to use data from the Medical Expenditure Panel Survey (MEPS), which is a great resource for finding data gathered on individuals and their medical providers. Because I decided to use a different data set than the one in the example code, I had to update the example code to contain the right file, along with the updated year for the variables. I analyzed the MEPS HC-201 full year consolidated data file because it contains MEPS survey data of demographics, health status, access to care, employment, and quality of care which are variables that I could use to analyze.

Modified Code

The below code was extracted from Emily Mitchell’s summary table examples in her github profile (Mitchell, 2017). All modifications and notes I have made to the code are in red font to indicate that I have changed it from its original version. The figures below (1 through 3) contain the output tables from the modified code. Once the data was successfully uploaded and the tables were created, summary statistics and T-test statistic code were added to produce additional analyses.

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\* Accessibility and quality of care, 2017

/\*

/\* Children with dental care

/\*

/\* Example SAS code to replicate number and percentage of children with dental

/\* care, by poverty status

/\*

/\* Input file: C:\MEPS\h201.ssp (2017 full-year consolidated)

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

ods graphics off;

/\* Load FYC file \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*Update filename and location to match the location for my computer\*/

FILENAME h201 "/folders/myfolders/Porfolio/h201.ssp";

proc xcopy in = h201 out = WORK IMPORT;

run;

/\* Define variables \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\* - For 1996-2007, AGELAST must be created from AGEyyX, AGE42X, AGE31X \*/

/\* - For 1996, use 'POVCAT' instead of 'POVCAT96' \*/

/\* Update file name to h201 and any 2016 variables to be 2017 variables\*/

/\* AGELAST data has been set to include eligible patients older than 1 year old and less than 18 years old \*/

/\* DVTOT17 is the number of dental care visits. This has been set to be greater than 0 \*/

/\* Child Dental is created to contain data with children in the age range who have visited the dentist \*/

/\* Proc format is run on the MEPS data to map the child\_dental data and the POVCAT data from the h201 data file. These mappings will create values to test the dental visits for children and the different poverty levels \*/

data MEPS;

SET h201;

/\* Children receiving dental care \*/

child\_2to17 = (1 < AGELAST & AGELAST < 18)\*1;

child\_dental = ((DVTOT17 > 0) & (child\_2to17 = 1))\*1;

run;

proc format;

value child\_dental

1 = "One or more dental visits"

0 = "No dental visits in past year";

value POVCAT

1 = "Negative or poor"

2 = "Near-poor"

3 = "Low income"

4 = "Middle income"

5 = "High income";

run;

/\* Calculate estimates using survey procedures \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\* Survey procedures are calculated with the MEPS data to prepare the number of strata, clusters, observations, and the sum of the weight by using the proc surveryfreq formula. This will help create the tables to estimate the population totals, proportions, and the standard errors. \*/

ods output CrossTabs = out;

proc surveyfreq data = MEPS missing;

FORMAT child\_dental child\_dental. POVCAT17 POVCAT.;

STRATA VARSTR;

CLUSTER VARPSU;

WEIGHT PERWT17F;

TABLES child\_2to17\*POVCAT17\*child\_dental / row;

run;

proc print data = out noobs label;

where child\_2to17 = 1 and child\_dental ne . and POVCAT17 ne .;

var child\_dental POVCAT17 WgtFreq StdDev RowPercent RowStdErr;

run;

A screenshot of a cell phone

Description automatically generated

Figure 1. The SURVEYREQ Procedure

A screenshot of a cell phone

Description automatically generated

Figure 2. Table of POVCAT by child\_dental for Zero Dental Visits

A screenshot of a cell phone

Description automatically generated Figure 3. Table of POCAT17 by child\_dental for One or More Dental Visits

/\* Produce summary statistics for POVCAT data and child\_dental data \*/

/\* Summary statistics include the mean, standard deviation, minimum value, maximum value, and number of values being analyzed \*/

/\* Proc means data formula is used on the H201 data set and the MEPS dataset to analyze statistics for the poverty category variables and for the value child\_dental that was created earlier \*/

ods noproctitle;

ods graphics / imagemap=on;

proc means data=WORK.H201 chartype mean std min max n vardef=df;

var POVCAT17;

run;

ods noproctitle;

ods graphics / imagemap=on;

proc means data=WORK.MEPS chartype mean std min max n vardef=df;

var child\_dental;

run;

A screenshot of a cell phone

Description automatically generated

Figure 4. Summary Statistics of POVCAT17 and child\_dental

/\* Test for normality \*/

/\* Normality test is used to determine if the data sample was taken from a normally distributed population which is required to run a valid t test \*/

proc univariate data=WORK.MEPS normal mu0=0;

ods select TestsForNormality;

class child\_dental;

var POVCAT17;

run;

/\* t test \*/

/\* Two-tailed test is ran on the poverty categorical data based on the child dental data. This test is used for null-hypothesis testing and for testing to see if there is statistical significance. \*/

proc ttest data=WORK.MEPS sides=2 h0=0 plots(showh0);

class child\_dental;

var POVCAT17;

run;

A screenshot of a cell phone

Description automatically generated

Figure 5. T-test Statistics

A screenshot of a cell phone

Description automatically generatedA picture containing text

Description automatically generated Figure 6. Distribution of POVCAT17

Figure 7. Q-Q Plots of POVCAT17

Breakdown of the Analysis

For this statistical analysis I wanted to analyze the statistics of children who had dental care in the year 2017 by the different poverty levels. To test to see if there was a relationship between poverty level and child dental care I did hypothesis testing.

HO: There is no significant relationship between the poverty level and the children who received dental care in 2017.

HA: This is a significant relationship between the poverty level and the children who received dental care in 2017.

To analyze the statistics produced earlier we need to focus on the Pr>F values. In Figure 5. We can see that the Pr>F value for this data is 0.0004, which is less than 0.05. Since it is less than 0.05 we will want to focus on the Satterthwaite results listed in Figure 5. The T-value for Satterthwaite has a value of 8.74 and a p-value less than 0.0001. Since the p-value is less than 0.05, we will reject the null hypothesis. We can say there is a significant relationship between the poverty levels and children who receive dental care in 2017.

References

MEPS (2019). Medical Expenditure Panel Survey. Retrieved from

<https://meps.ahrq.gov/mepsweb/data_stats/download_data_files_detail.jsp?cboPufNumber=HC-201>

Mitchell, E. 2017. Analyzing MEPS data using SAS. Retrieved from [https://github.com/HHS- AHRQ/MEPS/blob/master/SAS/summary\_tables\_examples/care1\_child\_dental.sas](https://github.com/HHS-    AHRQ/MEPS/blob/master/SAS/summary_tables_examples/care1_child_dental.sas)

Purdue Edu, 2010. Two Sample t-test with SAS. Retrieved from <https://www.stat.purdue.edu/~tqin/system101/method/method_two_t_sas.htm>

SAS [Computer Software]. (2019). Retrieved from: <https://www.sas.com/en_us/software/university-edition/download-software.html>.