Community Research Fellows Training Program Evaluation: Using SAS® to Analyze Pre- and Post-Test Data

Sarah A. Lyons and Melody S. Goodman

Washington University in St. Louis School of Medicine, St. Louis, MO

ABSTRACT

Background: The Community Research Fellows Training (CRFT) is an evidenced-based public health research methods course designed to increase community research literacy and enhance the infrastructure for Community Based Participatory Research (CBPR). The goals of the training are to: 1) adapt, implement, and evaluate a research methods training program for community members, 2) promote the role of underserved populations in research by enhancing the capacity for CBPR and 3) develop partnerships between academic researchers with community members in the St. Louis Greater Metropolitan area to address health disparities.

Methods: Multidisciplinary faculty conduct 12 didactic training sessions and 3 experiential workshops based on standard MPH curriculum. A comprehensive (formative/summative) mixed-method (quantitative/qualitative) evaluation approach is utilized to assess participant knowledge and satisfaction. As part of this program evaluation, participants complete, pre- and post-tests at each session. In order to streamline the grading process, a macro was created that allows for the input of participants answers and an answer key. The macro grades the tests, calculates the scores and plots the distribution of grades and the difference between pre and post-test.

Results: Use of this macro decreased time to grade pre/post-test and developed reports on participant and class progress; change in knowledge could be determined both numerically and visually. Graphs allow for the visualization of a particular participant’s success as well as the overall cohort.

Conclusions: The macro can be customized to fit the needs of the instructor as well as be generalized to more advanced types of tests.

Macro created using SAS® 9.3. Intended user level: Beginner.

INTRODUCTION

In order to enhance the infrastructure for community based participatory research to address health disparities in the St. Louis greater metropolitan area, the Program to Eliminate Cancer Disparities (PECaD) at the Siteman Cancer Center funded a pilot called the Community Research Fellows Training program. In this program, community members participate in a 15- week training program to increase their research literacy. The course is based on a Masters of Public Health curriculum, designed to expose the Fellows to the many facets of public health research. At the end of the program, they have the option of partnering with a researcher at the university and running their own pilot project that targets a specific need of their community. This program was inspired by the Community Alliance for Research Empowering Social Change (CARES) that had been previously implemented on Long Island, New York.

The ultimate goal of the training program is to connect academic researchers with passionate community members in order to serve the St. Louis Greater Metropolitan area and address health disparities. This partnership is anticipated to promote the role of underserved populations in research by enhancing the capacity of Community Based Participatory Research (CBPR).

This is the inaugural cohort of Fellows and the initiation of a second class will be dependent on the success of this initial program; thus program evaluation is critical. The program is evaluated in many ways, including an assessment of session-by-session knowledge. Each training session covers a particular public health topic and the fellows were administered a pre-test that focuses on the session’s learning objectives before the lecture begins. After the session, they were administered a post-test with the same number of questions, testing the same material. Prior to Session 9, the questions were not identical but covered the same content, as the goal was to capture overall knowledge, not improved skill at taking the test. Beginning with Session 9, the format was changed so that the questions on the pre-test were identical to those on the post-test, but in a different order. This was done to reduce the chance that a participant was receiving high scores due to chance alone.

We seek to investigate whether the sessions met their learning objectives and improved knowledge in our Fellows. For program evaluation purposes, improved session knowledge between pre-test and post-tests (indicated by higher post-test scores than pre-test scores) will indicate a successful training session. We hope to find significant enough improvement to allow continuation of the program and continue to foster the bonds between the community and academia. A continued partnership will enable community-academic partnerships to conduct beneficial research based on a community -driven research agenda.
MACRO OVERVIEW

Because of the volume of data to be analyzed, a macro was developed to grade the pre/post tests for each session in an efficient manner. The macro allows an Excel or text file to be read in with the Fellows' answers to the pre-test and post-test questions. The first eight sessions contained five questions on the pre-test and five on the post-test. Starting with Session 9, the number of questions on each test was increased to 10. The macro shown here can easily be adapted to accurately analyze tests with any number of questions. The macro grades the tests and calculates the scores for the pre-test and the post-test of each Fellow. The distributions of both tests are plotted and a Wilcoxon signed rank test is run to compare the averages. By outputting the results graphically, we hope to be able to make the results understandable to all, regardless of statistical background. The project team works closely with a community advisory board that needs to be able to understand program evaluation results in order to give proper feedback.

For the purposes of this paper, data from Training Session 8 (Community Based Participatory Research) will be used to demonstrate the steps of the macro. The pre-test and post-test each consisted of 5 questions, made up of varying combinations of True/False and multiple choice questions. The answers were coded as “1” if the participant selected the first choice (a), “2” for the second choice (b) and so on. If an answer is missing, this would be coded as “-9”.

<table>
<thead>
<tr>
<th>Fig.1. Sample Session 8 Pre-Test Question (multiple choice):</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. In a Community Based Participatory Research (CBPR) project who determines the health problem to be studied/analyzed?</td>
</tr>
<tr>
<td>a. University researchers</td>
</tr>
<tr>
<td>b. Community members</td>
</tr>
<tr>
<td>c. Government policy-makers</td>
</tr>
<tr>
<td>d. University researchers and community members together</td>
</tr>
</tbody>
</table>

In this multiple choice case (shown in Figure 1), the correct answer is “d. University researchers and community members together” and so that answer choice would be coded as “4”.

Figure 2 demonstrates a True/False question. If a Fellow selected “True”, the answer choice would be coded as “1” and if they selected “False”, that would be coded as a “2”.

<table>
<thead>
<tr>
<th>Fig.2. Sample Session 8 Pre-Test Question (True/False):</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. In a CBPR setting, researchers should begin their discussion with community members by asking questions rather than offering solutions.</td>
</tr>
<tr>
<td>a. True</td>
</tr>
<tr>
<td>b. False</td>
</tr>
</tbody>
</table>

INPUT

The macro presupposes a dataset containing a de-identified participant id and their answers to the 10 questions. The first 5 variables (Q1, Q2, Q3, Q4, Q5) correspond to the participant’s answers to the pre-test and the final 5 variables (Q6, Q7, Q8, Q9, Q10) correspond to the post-test answers. Figure 3 demonstrates the set-up of the Session 8 dataset, using real data from four Fellows.

<table>
<thead>
<tr>
<th>ID</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
</tr>
</thead>
<tbody>
<tr>
<td>2470</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>2474</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2525</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>2630</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>-9</td>
</tr>
</tbody>
</table>
The macro requires the input of a dataset similar to that in Figure 4, as well as answer values for each of the pre-test and post-test questions. An extension of this macro would be to have the answer key in a separate file and read it in. For this small example, it is just as efficient to manually type in the answer variables. The values for the answer key variables (a1, a2...a10) will be coded in the same fashion as the Fellows’ answer choices in the dataset ("1" for choice a, "2" for choice b and so on). For example, the correct answer to the question from Figure 2 is “True” and so the variable a2 would be given a value of 1.

```plaintext
%macro prepost (dataset, a1, a2, a3, a4, a5, a6, a7, a8, a9, a10);
  data grading;
    set &dataset;
    if q1 = &a1 then pt1=1; else pt1=0;
    if q2 = &a2 then pt2=1; else pt2=0;
    ...
  The macro grades each question individually and creates a dichotomous score variable (pt_). For each question that is correct, the scoring system assigns a value of 1 to pt_. This scoring system allows for missing values to be coded as incorrect and thus assigned a point value of 0.

  ...  grade1= 100*(sum(pt1,pt2,pt3,pt4,pt5)/5);
  grade2= 100*(sum(pt6,pt7,pt8,pt9,pt10)/5);
  diff= grade2-grade1;
  keep id grade1 grade2 diff;
```

The macro then calculates the average score (grade1, grade2) for each test (out of 100) as well as the difference (diff) defined as the post-test score minus the pre-test score. The final dataset will only keep the necessary variables of id, grade1, grade2, and diff variables to save on space and time. This dataset also provides a way to examine an individual participant’s grades.

**PROC UNIVARIATE**

**SUMMARY STATISTICS**

In order to analyze the pre-test and post-test scores, the Proc Univariate procedure was utilized. We were interested in the distribution of the individual grades. We used Proc Univariate’s histogram option to determine whether the pre-test and post-test grades were normally distributed. We also received summary statistics including the mean, median, quartiles and variance of each test. In order to show that the session was effective in increasing knowledge, we hope to see an increase in the mean and median of the scores between pre-test and post-test. We also hope to see a decrease in the variance between pre-test and post-test, indicating that scores became more clustered after the session.

```plaintext
  proc univariate data=grading;
    var grade1 grade2;
    histogram grade1 grade2/ normal href=80;
  run;
```

The histogram / normal option overlays the normal curve on top of the histogram, which allows for easy visualization of the distribution of the grades which will aid in determining which statistical test to perform on the data.

The inclusion of a horizontal reference line is another way to visualize the shift in scores (href=80). We chose the value of 80 as our reference line because that indicates a Fellow got 4 out of 5 questions correct and can therefore be assumed to have understood the majority of the material. This can be adapted to whatever value is deemed most appropriate.
We can see that neither distribution in Figure 4 is normal. Also, the post-test averages are left-skewed, which is encouraging as it indicates the mean shifted in a positive direction. We had hoped to see an increase in mean score and a decrease in variance between pre- test and post-tests. Proc Univariate’s summary statistics (displayed in Table 1 and Table 2.) indicate that the mean of the post-test increased from the pre-test’s 71% to 89.5%. There is also a decrease in variance, as the variance of the pre-test is 265.6 and this decreases to a value of 225.4 for the post-test.

<table>
<thead>
<tr>
<th>Table 1. Pre-Test Basic Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Test</strong></td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Mode</td>
</tr>
<tr>
<td>Std Deviation</td>
</tr>
<tr>
<td>Variance</td>
</tr>
<tr>
<td>Range</td>
</tr>
<tr>
<td>Interquartile Range</td>
</tr>
<tr>
<td>Percentile</td>
</tr>
<tr>
<td>95th percentile</td>
</tr>
<tr>
<td>75th percentile</td>
</tr>
<tr>
<td>50th percentile</td>
</tr>
<tr>
<td>25th percentile</td>
</tr>
<tr>
<td>5th percentile</td>
</tr>
</tbody>
</table>
**WILCOxon TEST**

In order to compare the significance of the difference between pre-test and post-test scores, a Wilcoxon signed rank test was used. A Wilcoxon signed rank test compares the median difference between pairs of observations. It is a non-parametric counterpart of the paired t-test. The data are not normally distributed (as seen in Figure 4), so using a t-test would not be appropriate in this case.

```
proc univariate data=grading;
var diff;
run;
```

Using Proc Univariate, we ran the Wilcoxon signed rank test on the difference variable (diff) calculated in the original data step. A significant difference was found (p< 0.0001). This indicates that the Fellows scored significantly higher on the post-test than the pre-test for this particular session. This is encouraging in terms of program evaluation as we can infer that the session information was useful and improved the Fellows knowledge of this particular public health topic.

**CONCLUSIONS**

This macro was extremely useful in reducing the time to grade and analyze pre-test and post-test for each of the program’s 12 sessions. The outputted data can be analyzed to determine the effectiveness of each session on increasing Fellows’ knowledge. Due to the structural nature of this particular macro, it can be easily customized to fit the needs of any instructor as well as be generalized to more advanced types of tests.

**POSSIBLE EXTENSIONS**

This macro can be extended in many ways that stretch beyond the use listed above. It can be simplified to quickly calculate scores and class averages from a single test. Additions of logic statements would allow for multiple correct answers. For lengthy tests, the macro could be adapted to read in the answer key as a separate file in order to save time from manually inputting many answer variables. The macro can further be adapted by adding virtually any SAS procedure that is deemed appropriate. For example, if the data was normally distributed, a t-test could be run on the difference variable by simply adding a Proc T-Test to the macro.

**REFERENCES**


**ACKNOWLEDGEMENTS**

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates US registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

I would like to thank the Community Research Fellows Training cohort for their continued enthusiasm.

The work of Dr. Goodman is supported by National Institutes of Health, National Cancer Institute grant U54CA153460 and Washington University School of Medicine Faculty Diversity Scholars Program.

**CONTACT INFORMATION**

Your comments and questions are valued and encouraged. Contact the author at:

Sarah A. Lyons
Washington University in St. Louis School of Medicine || Division of Public Health Sciences
660 S Euclid Avenue, Campus Box 8100, St. Louis, MO, 63110
Work Phone: 314-747-2568 | Email: lyonss@wudosis.wustl.edu
APPENDIX A: MACRO

%macro prepost (dataset, a1, a2, a3, a4, a5, a6, a7, a8, a9, a10);
    data grading;
        set &dataset;
        if q1 = &a1 then pt1=1;
        else pt1=0;
        if q2 = &a2 then pt2=1;
        else pt2=0;
        if q3 = &a3 then pt3=1;
        else pt3=0;
        if q4 = &a4 then pt4=1;
        else pt4=0;
        if q5 = &a5 then pt5=1;
        else pt5=0;
        if q6 = &a6 then pt6=1;
        else pt6=0;
        if q7 = &a7 then pt7=1;
        else pt7=0;
        if q8 = &a8 then pt8=1;
        else pt8=0;
        if q9 = &a9 then pt9=1;
        else pt9=0;
        if q10 = &a10 then pt10=1;
        else pt10=0;
        grade1= 100*(sum(pt1,pt2,pt3,pt4,pt5)/5);
        grade2= 100*(sum(pt6,pt7,pt8,pt9,pt10)/5);
        diff= grade2-grade1;
        keep id grade1 grade2 diff;
    run;
    proc univariate data=grading;
        var grade1 grade2;
        histogram grade1 grade2/ normal href=80;
    run;
    proc univariate data=grading;
        var diff;
    run;
%mend;

/*To call the macro, run the following code*/
%prepost(fellows, 4, 1, 3, 1, 2, 1, 2, 2, 3, 1);