"I don't know" and Multiple Choice Analysis of Pre- and Post-Tests

Karen Spears
State Nutrition Specialist
Reno, Nevada
kspears@CABNR.UNR.edu

Mary Wilson
Area Extension Specialist
Las Vegas, Nevada
wilsonm@UNCE.unr.edu

University of Nevada, Reno Cooperative Extension

Abstract: Evaluation is an essential component of any Extension education program. One tool, the pre- and post-test, provides measurable evaluation data. Yet often the answer "I don't know" or all possible answers to a multiple choice question are not included in the repeated measure analysis. Because more than two answers are offered, the test of marginal homogeneity repeated measure analysis should be used, which takes into account all response sets when comparing a participant's pre- and post-test answers. This article provides an example of conducting and interpreting the test of marginal homogeneity.

Introduction

Pre- and post- tests are a common method to determine if an Extension program has achieved its knowledge outcomes. When constructing a pre-and post-test, the format of the questions should be considered, especially because the question format will affect which repeated measure analysis should be conducted. If the pre-and post- tests contain more than two levels or categorical responses, for example "I don't know" or multiple choice answers, then the test of marginal homogeneity should be used. The test of marginal homogeneity extends the McNemar test to those cases where the variable of interest assumes more than two nominal or ordinal values for $k \times k$ table. The test of marginal homogeneity is an extremely important test yet an underutilized statistical method. The basic concept for the marginal homogeneity test is illustrated in Table 1 (Uebersax, 2006).

Table 1.
Summarization of Pretest (Rows) and Posttest (Columns) for Question Where Three Potential Answers Are Available

Posttest Answers

P		1-Correct	2-False	3-"I don't know"	
r e t e s t	1	*P_{11}	P_{12}	P_{13}	**$p_1.$
	2	P_{12}	P_{22}	P_{23}	$p_2.$
	3	P_{13}	P_{32}	P_{33}	$p_3.$
		$p._1$	$p_2.$	$p_3.$	1.0

*P =Represents paired answers (e.g. answer 1 pretest and answer 1 postest = P11)
**p =Represents the total participants that provided a specific answer

Questionnaire Design: "I don't know"

It is common to ignore the "I don't know" answer. Just comparing the percentage of participants who correctly answer a question before and after a program does not include those who changed their answer to or from "I don't know." Therefore, the true the impact of an education program may be underestimated or overestimated.

Rockwell and Kohn introduced a post-then-pre test evaluation method (Rockwell and Kohn, 1989). This approach sought to reduce pretest bias due a participant's lack of prior exposed to a concept. They suggest that upon program completion, participants report both their prior and post knowledge/behavior simultaneously. However, participants may express what they feel the cooperative Extension educator wants to hear.

Offering "I don't know" as an answer is potentially a better means for addressing pretest bias. Education is defined as deliberately transmitting knowledge, skills, and values to another (Webster's New World Dictionary and Thesaurus, 1996). If a participant is unaware of an idea prior to the program, but acquires understanding; learning has occurred. This is the goal of Extension education.

Example: Calcium, It's Not Just Milk

The Nevada Calcium, It's Not Just Milk program will serve as an example of using the test of marginal homogeneity in Cooperative Extension. The Calcium, It's Not Just Milk program was initiated in 2000 and serves to increase the consumption calcium rich foods among middle school students. Educational methods have included direct instruction by the teacher or a guest dietitian, food tasting events, and poster displays. In 2006, 747 middle school students participated in the program. There were 8 students who lacked matching pre- and post-tests results or declining participation. The total sample size for analysis is 739.

Question 1 asked do "Young children need more calcium than teenagers?" It contained three possible answers: True (incorrect answer =1), False (correct =2), or I don't know=3. On the pre-test, 153 (20.7%) students answered the question correctly, and 406 (55%) answered the question correctly on the post-test.

Because this is a repeated measures question with three response levels in a one sample population, each pre-test response can be grouped with three of the possible post-test responses. Therefore there are nine response sets (Table 2).

Table 2.
Possible Combinations of Pre-Test and Post-Test Responses to Question 1 "Young Children Need More

Calcium Than Teenagers?"

Young Children Need More Calcium than Teenagers?				
	Post-test Frequency (with percentage)			**Total**
Pre-test Frequency	True =1	False =2	I don't know = 3	
True = 1	180 (24%)	219 (30%)	22 (3%)	421
False = 2	33 (4.5%)	115 (16%)	5 (1%)	153
I don't know =3	59 (8%)	72 (10%)	34 (5%)	165
Total	272	406	61	739

Pre vs. Post Response Sets (9 response sets)	
1:1	24% incorrect pre and post
1:2	30% incorrect pre than correct post
1:3	3% incorrect pre than I don't know post
2:1	16% correct pre and post
2:2	4% correct pre then incorrect post
2:3	1% correct pre then I don't know post
3:1	5% I don't know pre and post
3:2	8% I don't know pre then incorrect post
3:3	10% I don't know pre then correct post

Procedure

Computing the Test of Marginal Homogeneity in SAS®, requires three steps:

1. Input your dataset in pre -and post-test format for each subject.

2. Run SAS PRO CATMOD procedure.

3. Interpret the SAS output

Collaborating with a statistician is recommended for database construction and conducting the SAS PRO CATMOD procedure.

Step 1: Input Dataset

Data can be entered into SAS in a variety of ways. You can directly type your research data into SAS, but this requires some finagling column attributes. It is easier to create a Window Excel file (Table 3). If your research data is spread out into multiple Windows Excel spreadsheets, merge them before importing your data into SAS. SAS can actually merge files, but it is easier to conduct this in Excel. Using the SAS import wizard will assist in importing your data.

Table 3.
Window Excel File Exert From "Calcium, It's Not Just Milk" Results For Question 1 "Young Children Need More Calcium Than Teenagers?"

SUBJID	gender	age	Question 1 Pre-test (q1pre)	Question 1 Post-test (q1post)
1001	Male	12	3	1
1003	Male	13	1	1
1004	Female	13	2	2
1005	Male	13	1	2
1006	Female	13	1	1
1007	Male	13	1	1
1008	Female	13	1	3

Step 2: SAS Procedure

An example of the procedure is:

PROC CATMOD;

     WEIGHT SUBJID;
     RESPONSE MARGINALS;
     MODEL q1pre*q1post=_response_/freq;
     repeated time 2;

title2 'Test of Marginal Homogeneity';
quit;

Step 3: Interpret SAS Output

The SAS output will contain several sections of information (Table 4). The Analysis of Variance provides the degrees of freedom, chi-square value for intercept and time, and statistical value of Percent Chi Square. Because time was <0.0001, the test of marginal homogeneity is interpreted as having no marginal homogeneity between pre- and post- test. In other words, the distribution of answers pre-test differs significantly from the answers from the post-test in the same subjects.

Table 4.
SAS Test of Marginal Homogeneity Output from Calcium, It's Not Just Milk Result for Question 1 "Young Children Need More Calcium Than Teenagers?" (n=739)

```
                        Analysis of Variance

        Source        DF      Chi-Squire     Pr>ChiSq
        Intercept     2       16762619       <.0001
        Time          2       885726.0       <.0001
        Residual      0


                  Test of Marginal Homogeneity
                      The CATMOD procedure

              Analysis of Weighted Least Squares Estimates


Effect      Parameter    Estimate    Standard Error    Chi-Square    Pr>ChiSq
Intercept      1         0.4815       0.000278          3001601      <.0001
               2         0.3647       0.000260          1973445      <.0001
Time           3         0.1173       0.000251          218211.5     <.0001
               4        -0.1962       0.000218          810720.5     <.0001
```

Questionnaire Design: Multiple Choice

Multiple choice questions with one correct answer should also be an evaluated by the test of marginal homogeneity. The process and result interpretation are similar to what is conducted when "I don't know" is provided as an option to a question.

An example multiple choice questions with one correct answer from the Calcium, It's Not Just Milk program is Question 8: "What is the serving size of this item on this food label?" Possible answers were 1=1 cup (correct), 2=1/2 cup, 3=1 ½ cups, 4=2 cups. The test of marginal homogeneity can be used with ordinal data (Table 5).

The SAS Output is very similar, except there are 16 response levels.

Table 5.
Possible Combinations of Pre-Test and Post-Test Responses to Question 8: "What Is the Serving Size of this Item on this Food Label?" (n=733*)

What is the Serving Size of this Item on this Food Label?					
	Post-Test Frequency (with percentage)				**Total**
Pre-test Frequency	**1 Cup= 1**	**½ Cup = 2**	**1 ½ Cups =3**	**2 Cups = 4**	

1 Cup= 1	12	10	8	4	34
	(1.6%)	(1.4%)	(1.1%)	(0.5%)	
Â½ Cup = 2	499	15	34	19	567
	(68.3%)	(2.1%)	(4.7%)	(2.6%)	
1 Â½ Cups =3	46	8	19	7	80
	(6.3%)	(1.1%)	(2.6%)	(1.0%)	
2 Cups = 4	23	4	13	10	50
	(3.2%)	(0.5%)	(1.8%)	(1.4%)	
Total	580	37	74	40	731
*Frequency missing =16					

Pre vs. Post Response Sets (16 response sets)	
1:1	1.6% correct pre and correct post
1:2	1.4% correct pre than incorrect post
1:3	1.1% correct pre than incorrect post
1:4	0.5% correct pre than incorrect post
2:1	68.3% incorrect pre then correct post
2:2	2.1% incorrect pre then incorrect post
2:3	4.7% incorrect pre then incorrect post
2:4	2.6% incorrect pre then incorrect post
3:1	6.3% incorrect pre then correct post
3:2	1.1% incorrect pre then incorrect post
3:3	2.6% incorrect pre then incorrect post
3:4	1.0% incorrect pre then incorrect post
4:1	3.2% incorrect pre then correct post
4:2	0.5% incorrect pre then incorrect post
4:3	1.8% incorrect pre then incorrect post
4:4	1.4% incorrect pre then incorrect post
*total percentage >100 due to rounding	

Table 6.

SAS Test of Marginal Homogeneity Output From Calcium, It's Not Just Milk Results for Question 8: "What is the Serving Size of this Item on this Food Label?" (n=731)

Analysis of Variance

Source	DF	Chi-Squire	Pr>ChiSq
Intercept	3	52191971	<.0001
Time	3	5498919	<.0001
Residual	0		

The CATMOD procedure

Analysis of Weighted Least Squares Estimates

Effect	Parameter	Estimate	Standard Error	Chi-Square	Pr>ChiSq
Intercept	1	0.4290	0.000148	8398377	<.0001
	2	0.4222	0.000162	6823538	<.0001
Time	3	0.00947	0.000173	301164.2	<.0001
	4	-0.3906	0.000183	4551713	<.0001
	5	0.3772	0.000187	4076560	<.0001
	6	0.00227	0.000147	237.72	<.0001

A statistical significant increase was found answering Question 8 between pre- and-post tests. Only 4.7% (34/731) of participants answered the question correctly pre-test and correctly post-test, while 79.3% (580/731) answered the question incorrectly pre-test and correctly post-test (p<0.001).

Summary

The evaluation component is an important aspect in developing an education program. When a pre- and post-test is selected, the number of possible answers determines the statistical test that will be conducted. A statistician should be consulted early in project creation. Extension educators cannot ignore the" I don't know" or multiple choice questions when more than two levels are provided in the questionnaire. The test of marginal homogeneity is a statistical analytic test that will provide a more appropriate evaluation. This reinforces the need to begin with the-end-in-mind in program evaluation.

References

Agnes, M. (1996). *Webster's new world dictionary and thesaurus*, New York, NY: Simon & Schuster, Inc,

Rockwell, S. K., & Kohn, H. (1989). Post-Then-Pre Evaluation. *Journal of Extension* [On-line], 27(2) Article 2FEA5. Available at: http://www.joe.org/joe/1989summer/a5.php

Uebersax, J. (2006). McNemar test of marginal homogeneity. [On-line], Retrieved December 8, 2009 from: ht,tp://john-uebersax.com/stat/mcnemar.htm