

Paper 204-31

**Implementing a Multiple Comparison Test for Proportions
in a 2xc Crosstabulation in SAS®**

Alan C. Elliott, Joan S. Reisch
Center for Biostatistics and Clinical Science
UT Southwestern Medical Center, Dallas

ABSTRACT

This presentation illustrates a method for creating a multiple comparison test for proportions in a 2 x C cross tabulation or contingency table using the FREQ procedure and a SAS macro called compprop. This test is analogous to a Tukey-type multiple comparison test as used with one-way analysis of variance.

INTRODUCTION

A common statistical analysis for count data is to compare the proportions of a binary outcome of interest for two or more groups. The data for this analysis are displayed in a table with 2 rows and c columns where c is the number of groups. In SAS, the analysis is usually performed using the FREQ procedure. The output of the FREQ procedure, when requested by the /CHISQ option, is a table that provides a test of significance using a Chi-square analysis testing the null hypothesis that there is no difference in the proportions experienced by the groups.

If the Chi-square test results in a p-value of 0.05 or smaller, the results are deemed significant, the null hypothesis is rejected, and it is concluded that there is a significant difference between at least two of the groups. In the case when there are only two groups being compared (a 2 x 2 table), it is easy to see which group has a larger proportion than the other. However, in the case where there are three or more groups (2 x c table), there is no follow-up test available in SAS to identify the significant differences between pairs of groups. We propose a procedure that can be readily implemented in SAS and will provide the comparisons.

This procedure is similar to multiple comparisons tests following one-way analysis of variance when an overall significant difference is found. A multiple comparison test (Tukey, Student Newman Keuls, etc.) is utilized to determine whether a significant mean difference exists between each pair of groups. In an analogous way pairs of proportions in a 2 x c frequency table can be compared once the overall Chi-square for the contingency or cross tabulation table is found significant.

Techniques discussed in Zar, 1999 illustrate how such a test can be implemented using a Tukey-type of multiple comparison procedure.

METHOD

Using the SAS FREQ procedure and a SAS macro named compprop, this multiple comparison test is implemented in the following example. Suppose the data in Table 1 are analyzed using the FREQ procedure. The proportions of interest in this table are those reported as the column percents for variable B, groups, which are 36.78, 39.81, 20.00 and 36.00.

The Chi-square statistic for this table (provided by the FREQ procedure with a /Chisq option) is 8.99 with $p = 0.03$. These results lead to rejection of the null hypothesis that all proportions are equal. However, the question remains, which pair or pairs of groups have significantly different proportions?

Table of a by b					
a	b				Total
Frequency	1	2	3	4	
Percent					
Row Pct					
Col Pct	1	2	3	4	Total
1	32	43	16	9	100
	10.67	14.33	5.33	3.00	33.33
	32.00	43.00	16.00	9.00	
	36.78	39.81	20.00	36.00	
2	55	65	64	16	200
	18.33	21.67	21.33	5.33	66.67
	27.50	32.50	32.00	8.00	
	63.22	60.19	80.00	64.00	
Total	87	108	80	25	300
	29.00	36.00	26.67	8.33	100.00

Table 1, Data for proportions example

PROGRAM OUTPUT

The COMPROP macro calculates the necessary values for the multiple comparison procedure and outputs a standard table of results. The first step in making all the comparisons is to transform each proportion as follows:

$$p'_i = \frac{1}{2} \left[\arcsin \sqrt{\frac{X_i}{n_i + 1}} + \arcsin \sqrt{\frac{X_i + 1}{n_i + 1}} \right]$$

where X_i is the observed count for category i and n_i is the sample size for that category. As an example, for the proportion 39.81 (category 2) the p'_i is

$$p'_2 = \frac{1}{2} \left[\arcsin \sqrt{\frac{43}{108 + 1}} + \arcsin \sqrt{\frac{43 + 1}{108 + 1}} \right] = 39.185$$

The resulting p'_i values for the four categories are

Category (B)	1	2	3	4
Calculated p'_i	37.43	39.18	26.83	37.19

The transformed values are then ranked in numerical order from smallest to largest and all differences are determined with the largest difference first (2 vs. 3) followed by the largest vs. the second smallest, etc. A standard error (SE) is computed for each comparison and its formula is given next.

$$SE = \sqrt{\frac{410.35}{n_A + 0.5} + \frac{410.35}{n_B + 0.5}}$$

Where n_A is the sample size in one of the groups in the comparison and n_B is the sample size in the other group. As an example the standard error for the comparison of groups 2 and 3 is:

$$SE = \sqrt{\frac{410.35}{n_2 + 0.5} + \frac{410.35}{n_3 + 0.5}} = \sqrt{\frac{410.35}{108 + 0.5} + \frac{410.35}{80 + 0.5}} = 2.979$$

The second column in the output Table 2 (displayed below) shows the difference between the two compared p_i values. For the first comparison the difference is 12.35. The appropriate SE's are shown in the third column.

The fourth column of the table is the "q" statistic which is the difference divided by the appropriate SE. The column labeled q(.05) contains the critical values for the q distribution for $\alpha = 0.05$, degrees of freedom, $v, = \infty$ and for $k =$ the number of groups found in Table B.5 of Zar (1999).

If the value of the calculated q statistic exceeds the q(.05) critical value, the program places the message "Reject" in the Conclusion column. Otherwise, the program places "Accept" in that column. Notice that several comparisons in the example table are listed as "Do not test." This happens because a non-significant comparison between groups 2 and 4 overlaps the comparison of groups 2 and 1. If groups 2 and 4 are not significant, any comparison within it cannot logically be significant and therefore is not tested.

The program listed below yields the results in the output shown in Table 2. The results are interpreted to indicate that at the 0.05 level, there is a significant difference between the proportions of groups 2 and 3 (indicated by the "Reject") conclusion. All other comparisons are not significant.

IMPLEMENTING THE COMPPROP PROGRAM

To implement the SAS code to perform this test, the following changes may need to be made. In the macro assignment statement

```
%let ncols=4;
```

change the 4 to the number of groups you are comparing. Following the DATALINES statement, enter your data as the counts from the two rows of your 2 x c cross tabulation table. For this example, the following data are the observed frequencies from Table 1:

```
DATALINES;
32 43 16 9
55 65 64 16
```

Your data must consist of only two rows of data. After making these changes in the SAS code, submit the program in SAS to create the output in Table 2.

Table 2. Sample program output for COMPPROP

Tukey Style Multiple Comparisons of Proportions						
Compare	Diff	SE	q	q(.05)	Conclude	
2 vs 3	12.35	2.98	4.15	3.633	Reject	
2 vs 4	1.99	4.46	0.45	3.633	Accept	
2 vs 1	Do not test					
1 vs 3	10.6	3.13	3.39	3.633	Accept	
1 vs 4	Do not test					
4 vs 3	Do not test					

Reference: Biostatistical Analysis, Fourth Edition, Jerrold Zar, 1999, p 564.

The code for the COMPROP program, which consists of the compprop macro and a main program are shown below:

```

*****
* SAS® macro to perform multiple comparisons on proportions *
*****;
%macro compprop(n=);
proc freq;
  weight wt;
  tables a*b /chisq out=freqout outpct;
run;
*Select variables from data set freqout;
data propmc;set freqout;
if a=1;
* Apply an arcsine transformation to the proportions;
pi180=180/3.141;
n=(count*100)/pct_col;
p=0.5*(arsin(sqrt(count/(n+1)))*pi180 + arsin(sqrt((count+1)/(n+1)))*pi180);
proc sort data=propmc;by p;
proc transpose data=propmc out=propmcl;var n p b;
*Separate out pieces of the data and recombine into the needed data set;
data tmp;set propmcl;
if _N_=1;
%do i=1 %to &ncols;
  n&i=col&i;
%end;
data tmp2;set propmcl;
if _N_=2;
data tmp3;set propmcl;
if _N_=3;
%do i=1 %to &ncols;
  lab&i=col&i;
%end;
drop coll-col&ncols;
data propmc3;merge tmp tmp2 tmp3;
run;
* The array q05() contains the table lookup values for the q statistic with infinity
degrees of freedom for k groups (1 to 10);
data propmc2;set propmc3;
array p(*) coll-col&ncols;
array na(*) n1-n&ncols;
array lab(*) lab1-lab&ncols;
array q05(10);
q05(1)=.;
q05(2)=2.772;
q05(3)=3.314;
q05(4)=3.633;
q05(5)=3.585;
q05(6)=4.030;

```

```

q05(7)=4.170;
q05(8)=4.286;
q05(9)=4.387;
q05(10)=4.474;
icompare=&n;
qcrit=q05(icompare);
* print out the multiple comparison results table;
msg='          ';
file print;
put '          Tukey Style Multiple Comparisons of Proportions';
put '          ';
put '          Compare      Diff      SE      q      q(.05)      Conclude';
put '          -----';
iskiprest=0;
* as the table is constructed, determine the correct Conclude message;
do i=icompare to 1 by -1;
do j=1 to i;
if i ne j then do;
se=round(sqrt((410.35/(na(i)+.5))+(410.35/(na(j)+.5))),.01);
diff=round(p(i)-p(j),.01);
q=round((p(i)-p(j))/se,.01);
iskip=0;
* Do not test if last one in series was Accept;
if msg='Accept' and j ne 1 then do
msg='Do not test';iskip=1;
put @5 lab(i) 'vs' @10 lab(j) @15 msg;
msg='Accept';
end;
* Do not test last one if previous one was Accept;
if iskip=0 and iskiprest=1 then do
msg='Do not test';iskip=1;
put @5 lab(i) 'vs' @10 lab(j) @15 msg;
end;
if iskip=0 then do;
if q gt qcrit then msg='Reject';else msg='Accept';
put @5 lab(i) 'vs' @10 lab(j) @ 15 diff @25 se @35 q @45 qcrit @55 msg ;
end;
if j=1 and (msg='Accept' or msg='Do no test')then iskiprest=1;
retain msg;
end;
end;
end;
put '          ';
put '          Reference: Biostatistical Analysis, Fourth Edition, Jerrold Zar, 1999, p
564.';
run;
%mend compprop;

*****
* MAIN PROGRAM *
* 1. Enter a value for ncols (number of columns in a 2 x c table) *
* 2. Enter data -- two rows with ncols numbers each row *
* This procedure can handle up to 10 columns *
*****;
* Set number of proportions/ncols (in r x c table);
*****
* Enter NCOLS BELOW *
* NCOLS can be from 2 to 10; *
* Enter data following the datalines *
* statement. *
* The number of data items in the data *
* must match the value of ncols *
*****;
%let ncols=4;
options ps=50;
data;

```

```

do a = 1 to 2;
  do b = 1 to &ncols;
    input wt @@;
    output;
  end;
end;
datalines;
32 43 16 9
55 65 64 16
;
run;
title "Chi-Square Analysis for a 2x&ncols table";
*****
*invoke the compprop macro *
*****;
%compprop(n=&ncols);
run;

```

CONCLUSION

The COMPPROP multiple comparisons procedure for a 2 x c contingency table analysis provides a post hoc test when the Chi-square test is significant which makes interpretation of results easier and clearer. This code can be modified to include other significance levels (other than 0.05) by changing the tabled values in the code to those appropriate for another significance level. For example, these values are found in table B.5 of the Zar (1999) text.

REFERENCES

Zar, J.H. (1999). *Biostatistical Analysis*, Fourth Edition, Prentice Hall, p 564. and App64.

CONTACT INFORMATION

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

Alan C. Elliott
Center for Biostatistics and Clinical Science
UT Southwestern Medical Center Dallas
5323 Harry Hines Blvd
Dallas TX 75390-8822
Phone (214) 648-2712
Fax (214) 648-7673
alan.elliott@utsouthwestern.edu

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 USA registration. Other brand and product names are trademarks of their respective companies.