Mantel-Haenszel Differential Item Functioning (DIF) Analyses Using SAS

In this document I explain how to use SAS to examine differential item functioning (DIF) using the Mantel-Haenszel (MH) method discussed in Chapter 16.

As explained in that chapter, DIF can be uniform or nonuniform. In uniform DIF, only the difficulty parameter (intercept or threshold) of the item differs across groups.

In non-uniform DIF, the discrimination (or slope) of the item differs across groups. Of course, items may display both uniform and nonuniform DIF (differences across groups in both difficulty and discrimination).

The MH method tests only for uniform DIF. Methods based on logistic regression test for both uniform and nonuniform DIF. LR methods are described in a separate document.

The data for this example are from a 31-item multiple choice test for which gender DIF was assessed. Data were available from 500 male and 803 female students. The data are available in the file "dif.sas7bdat." There are no missing data. For these data, males are coded as 0 and females are coded as 1.

Basic DIF testing requires that each item be tested separately. With a large number of items, this can be tedious. Because of this, I created a mini-program called a "macro" that uses SAS syntax to automate the process of running the analysis sequentially for all items and collating the results of these analyses.

In the rest of this document I first discuss the MH analysis for a single item. I then present and discuss the macro for the MH procedure. Finally, I present and discuss a program to combine and format the results of the macro-based analyses. Please refer to Chapter 16 for more detail on the MH procedures for detecting DIF.

Mantel-Haenszel Procedure for a Single Item

The Mantel-Haenszel procedure can be obtained from **proc freq** using the syntax below in which I use the generic variable names "item," "group," and "totalscore" for the item being tested for DIF, the grouping variable, and the total score used for matching.

proc freq data=XXX;
 tables *total**item*group/**cmh1 noprint** ;
run;

The option **cmh1** requests the statistics the MH chi-square value and its *p*-value, and the MH odds ratio (alpha).

The option **noprint** prevents the full chi-square table from being printed in the results. This can be useful in analyses such as this in which there is an item by group table for each score point in the total score.

The MH Odds Ratio (α_{MH})

On page 484 of the text I explain that the odds of a correct answer are calculated as the proportion of correct answers divided by the proportion of incorrect answers for a particular group. The odds ratio is the odds for the reference group divided by the odds for the focal group.

Confusingly, however, in SAS the odds are always calculated as the proportion for the lower category divided by the proportion for the higher category. Because incorrect answers are typically coded as 0 and correct answers as 1, in SAS the odds are calculated in the opposite way from that in the text: as the proportion of *incorrect* answers divided by the proportion of *correct* answers, or the odds of an *incorrect* answer.

The good news is that, because SAS provides the inverse of the value we want, we can simply invert the value obtained from SAS to get the usual DIF α_{MH} value of correct proportion/incorrect proportion.

A related point is that, in SAS the odds ratio α_{MH} is calculated as the odds for the group with the lower code divided by the odds for the group with the higher code. Because by convention values of α_{MH} for DIF analyses are calculated as the odds for the reference group divided by the odds of the focal group, I have coded the reference group (males) as 0 and the focal group (females) as 1.

The table below shows the output for item 1. The MH common odds ratio, or α_{MH} value, is .743. As noted previously, this is the odds of $p(in$correct$)/p($correct$)$ for the reference group (males, coded 0) divided by the corresponding odds for the focal group (females, coded 1). So, the odds of getting the item *incorrect* in the reference group (males) was .743 times less than the odds for the focal group (females).

We can invert this value by simply taking $1/\alpha_{MH}$, or $1/.743$, which yields ≈ 1.35. This inverted value indicates that the odds of getting the item *correct* for the focal group (males) is 1.35 times the odds for the focal group (females). Thus, males have greater odds, or a greater chance, of getting this item right, after controlling for overall ability.

Common Odds Ratio and Relative Risks				
Statistic	**Method**	**Value**	**95% Confidence Limits**	
Odds Ratio	Mantel-Haenszel	0.7435	0.5437	1.0166
	Logit **	0.7680	0.5587	1.0559
Relative Risk (Column 1)	Mantel-Haenszel	0.8385	0.6919	1.0162
	Logit **	0.9166	0.7660	1.0967
Relative Risk (Column 2)	Mantel-Haenszel	1.1203	0.9999	1.2553
	Logit **	1.2680	1.1574	1.3892

As noted in the text (see page 485), the value of α_{MH} is sometimes converted to the so-called delta (Δ_{MH}) metric by taking the natural log of α_{MH} and multiplying this by -2.35. Converting the inverted α_{MH} value of 1.35 to the delta metric yields $\Delta_{MH} \approx -.70$. Although higher values of α_{MH} indicate that the item is *easier* for the reference group, Δ_{MH} has a somewhat different interpretation. The delta value is zero if there is no DIF. It is positive if the item is easier for the focal group and negative if the item is easier for the reference group. Thus, values expressed in the delta metric will always be consistent with those expressed in the alpha metric in identifying the group for which the item is easier.

The MH Chi-square Value (χ^2_{MH})

Another aspect of DIF testing using Mantel-Haenszel is the MH chi-square (χ^2_{MH}) and its associated significance test. Both α_{MH} and Δ_{MH} can be thought of effect size metrics. The MH chi-square provides a statistical test of the equality of odds ratios in the reference and focal groups. Running the syntax above also yields the MH chi-square value, which is shown below for item 1.

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)				
Statistic	**Alternative Hypothesis**	**DF**	**Value**	**Prob**
1	**Nonzero Correlation**	1	3.4292	0.0641

The MH chi-square value of 3.492 is not statistically significant at the .05 level, indicating that the DIF is not significant. Using the ETS classification system for DIF, item 1 would be classified into level A because the absolute value of Δ_{MH} is less than 1 and χ^2_{MH} is not statistically significant. Level A items are considered to have negligible, if any, DIF.

Automating the process

As noted previously, DIF must be tested for each item, which can become tedious when there are many items on a scale. I created the macro below to automate the process in SAS. In the subsequent text I explain how it works.

options nodsnferr mcompilenote=all mlogic mprint;

The **options** command specifies that:

- error messages should be ignored and processing should continue if errors are found (**nodsnferr**)
- notes for any errors and info on the macro's instructions should be displayed in the log (**mcompilenote=all**)
- execution of the macro commands should be displayed in the log (**mlogic)**
- the commands generated by the macro should be displayed in the log (**mprint**)

```
%macro MH;
%do item= 1 %to 31;
proc freq data=XXX noprint;
 tables total*I&item*group
 /cmh1 noprint;
 Output out=temp cmh1;
run;
data temp; set temp;
item=&item;
keep item _CMHCOR_ P_CMHCOR _MHOR_;
Run;
data MH; set MH temp;
run;
proc datasets; delete temp ;
run;
%end;
%mend;
%MH;
```

The command **%macro MH** creates a macro named MH that will carry out the sequence of commands following it. The macros is ended by the command **%mend**;.

The symbol "%" is used to identify macros commands. The sequence of commands begins with a do loop (**%do item= 1 %to** 31). This part of the command "item = 1 %to 31" specifies a looping variable (item) that begins at 1 and ends at 31. Thus, the commands within the loop will be run 31 times. Note that there must be a % before the word "to."

The **%do** command is followed by the **proc freq** commands used previously to obtain MH statistics for item 1. Here, the name of the variable "item" is replaced by "I**&item.**" This indicates that the letter "I" should be concatenated with the current value of the looping variable "item." Recall that the 31 items are named i1, i2, i3, and so on. The specification "I**&item"** will create these names sequentially as the loop iterates from 1 to 31.

The subcommand "**Output out=temp cmh1;"** has been added to the **proc freq** commands. This creates a temporary datafile ("temp") that contains the statistics specified by cmh1 (the MH chi-square value, its *p*-value, and the MH odds ratio (alpha)).

The next set of commands is:
data temp; **set** temp;
item**=&item**;
keep item _CMHCOR_ P_CMHCOR _MHOR_;

These commands cause the macro to create a temporary file ("temp") at each iteration of the loop (i.e., for each iteration of the loop), create an item number from the current value of the looping index "&item," and keep the item number, the MH chi-square value (_CMNCOR_), the *p*-value for the chi-square (P_CMHCOR), and the MH odds ratio (_MHOR_).

Finally, the macro creates a new dataset called "MH" that contains the saved values from each "temp" file. It then deletes the temp datasets using the **dataset** command. The result is that the saved values for each item are in the dataset MH and the temp files for each item are deleted.

The command sequence ends with the call to the MH macro, "**%MH**." This causes the macro to begin execution.

Output from the macro

Running the syntax above will produce a new SAS file that contains the item number the MH chi-square value (_CMNCOR_), the *p*-value for the chi-square (P_CMHCOR), and the MH odds ratio (_MHOR_), as shown below:

	CMH Nonzero Correlation	P-value for CMH Nonzero Correlation	Mantel-Haenszel Adjusted Odds Ratio	item
1	0.6332502643	0.4261655961	1.1069148984	2
2	15.749915828	0.0000722912	1.6884252315	3
3	0.4393202854	0.5074507072	0.9167722084	4
4	6.8878576432	0.0086783216	0.6894547901	5
5	0.7021883177	0.4020493633	1.1123296676	6
6	7.4125385403	0.0064770868	0.6717194204	7
7	3.0769737527	0.0794081514	0.7146561841	8
8	0.3118536281	0.5765449464	1.0762937314	9
9	0.6925711811	0.4052912012	1.1176396622	10
10	3.2372012691	0.0719835064	1.2509476936	11
11	14.14990189	0.0001688059	0.5047466821	12
12	7.7156742568	0.005474341	1.577278158	13

The file contains the three MH statistics and the item number for each of the 31 items.

Renaming the MH values and calculating delta and the inverse of alpha

The next set of commands renames the MH values _CMHCOR_ , P_CMHCOR, and _MHOR_ , takes the inverse of the alpha value, and calculates the ETS delta value. (Recall that by default SAS computes the odds an incorrect rather than a correct response. Taking the inverse of the resulting alpha value produces the odds of a correct response.)

data MH; **set** MH;
rename _CMHCOR_ = chi P_CMHCOR=prob _MHOR_=alpha;
data MH; **set** MH;
alpha_inv = 1/alpha;
delta = **-2.35*log**(alpha_inv);
run;

Saving values into an rtf file

Finally, the last set of commands creates an rtf file that contains a table with the item number, the MH chi-square, the p-value for chi-square, the MH common odds ratio (alpha), the value of delta, and the inverse of the alpha value.

Enclosing the relevant commands between the **ods rft** and **ods rtf close** commands causes the rtf file to be created.

```
ods rtf file="XXX.rtf";
proc print; format alpha delta chi prob F9.5;
id item; var chi prob alpha alpha_inv delta ;
run;
ods rtf close;
```

The first eight rows of the resulting rtf file are shown below:

item	alpha	delta	chi	prob
1	0.74348	0.69657	3.42918	0.06405
2	1.10691	-0.23871	0.63325	0.42617
3	1.68843	-1.23092	15.74992	0.00007
4	0.91677	0.20421	0.43932	0.50745
5	0.68945	0.87386	6.88786	0.00868
6	1.11233	-0.25017	0.70219	0.40205
7	0.67172	0.93510	7.41254	0.00648
8	0.71466	0.78949	3.07697	0.07941