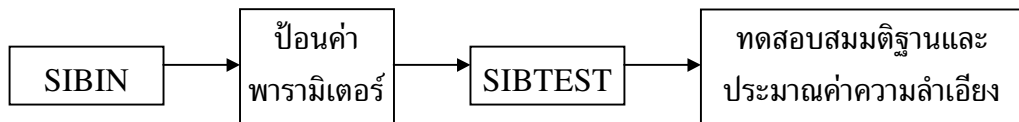


การวิเคราะห์ความลำเอียงด้วยโปรแกรมสำเร็จรูป SIBTEST

นัตรศิริ ปิยะพิมลสิทธิ์*

SIBTEST เป็นโปรแกรมสำเร็จรูปทางสถิติที่เขียนขึ้นด้วยภาษา Fortran การพัฒนา SIBTEST อยู่บนพื้นฐานของทฤษฎี IRT ในการทดสอบความลำเอียงของข้อสอบ SIBTEST เป็นโปรแกรมในการคำนวณความลำเอียงของข้อสอบของกลุ่มตัวอย่าง 2 กลุ่มระหว่างกลุ่มอ้างอิง (Reference group) และกลุ่มสนใจ (Focal group) โดยโปรแกรมจะทำการทดสอบสมมติฐานทางสถิติและประมาณค่าพารามิเตอร์

SIBIN เป็นอีกโปรแกรมหนึ่งที่มาพร้อมกับ SIBTEST เป็นโปรแกรมสำหรับให้ผู้ใช้ป้อนค่าพารามิเตอร์เพื่อใช้ในการประมวลผลร่วมกับ SIBTEST ดังแผนภูมิ



ในขั้นแรกจะใช้ SIBIN ในการตั้งค่าพารามิเตอร์ แล้วใช้ SIBTEST ในการคำนวณค่าความลำเอียงและทดสอบสมมติฐานทางสถิติ โดย SIBTEST จะป้อนค่าพารามิเตอร์ที่ถูกเก็บโดย SIBIN มาใช้ในการรันโปรแกรม ผลลัพธ์ที่ได้จากการรัน SIBTEST แต่ละครั้งจะแตกต่างกันไปขึ้นอยู่กับค่าพารามิเตอร์ที่ตั้งใน SIBIN

สิ่งจำเป็นในการวิเคราะห์

ในการใช้ SIBTEST จะต้องเตรียมดังนี้

1. เตรียมผลคะแนนที่ได้จากการสอบของกลุ่มอ้างอิงและกลุ่มสนใจโดยข้อสอบแต่ละข้อจะต้องตรวจให้คะแนนแบบ 0 - 1
2. โปรแกรม SIBIN กินเนื้อที่ประมาณ 78 K และ SIBTEST กินเนื้อที่ประมาณ 210 K

กลุ่มอ้างอิงและกลุ่มสนใจ

กลุ่มอ้างอิง (Reference Group) หมายถึง กลุ่มของผู้สอบที่เชื่อว่าเมื่อทำข้อสอบที่สงสัยว่าจะลำเอียงแล้วได้เปรียบ

กลุ่มสนใจ (Focal Group) หมายถึง กลุ่มของผู้สอบที่เชื่อว่าเมื่อทำข้อสอบที่สงสัยว่าจะลำเอียงแล้วเสียเปรียบ

*กศ.ม. (การวัดผลการศึกษา) <http://www.watpon.com>

ข้อจำกัดของ SIBTEST

SIBTEST สามารถวิเคราะห์ความลำเอียงของข้อสอบได้สูงสุด 80 ข้อ และจำนวนกลุ่มตัวอย่างในกลุ่มอ้างอิงและกลุ่มสนใจไม่เกินกลุ่มละ 3,000 คน

ตัวอย่างข้อมูล

สมมติเราต้องการวิเคราะห์ความลำเอียงระหว่างเพศของแบบทดสอบเลือกตอบวิชาพลศึกษา จำนวน 15 ข้อ เพื่อจะศึกษาดูว่าข้อสอบในแบบทดสอบวิชาพลศึกษานี้มีความลำเอียงต่อกลุ่มเพศหญิงหรือไม่ กลุ่มตัวอย่างเป็นนักเรียนจำนวน 200 คน โดย 100 คนแรกเป็นกลุ่มนักเรียนเพศชายซึ่งกำหนดให้เป็นกลุ่มอ้างอิง (Reference Group) และอีก 100 คนเป็นกลุ่มนักเรียนเพศหญิงซึ่งกำหนดให้เป็นกลุ่มสนใจ (Focal Group) นำผลการสอบมาป้อนเก็บไว้เป็น 2 แฟ้ม โดยกลุ่มอ้างอิงเก็บไว้ในแฟ้มชื่อ ur.dat และกลุ่มสนใจเก็บไว้ในแฟ้มชื่อ uf.dat

ขั้นตอนการป้อนข้อมูล

ในเอกสารฉบับนี้จะนำเสนอวิธีการป้อนข้อมูลโดยใช้ Editor เช่น Q-Edit หรือ Edit ที่มากับ DOS หรืออาจจะใช้ Word Processor ตัวอื่น ๆ เช่น Word Perfect, เวิร์ดราวิที หรือเวิร์ดจิวา 1.53 ก็ได้

ผู้ใช้สามารถป้อนข้อมูลได้ 2 ลักษณะคือ

ลักษณะแรก การป้อนคะแนนของข้อสอบแต่ละข้อจะป้อนติดต่อกัน คะแนนของข้อสอบข้อที่ 1 จะอยู่ในสดมภ์ที่ 1 ข้อสอบข้อที่ 2 จะอยู่ในสดมภ์ที่ 2 ข้อสอบข้อที่ 3 จะอยู่ในสดมภ์ที่ 3 จนกระทั่งถึงข้อสอบข้อที่ 15 อยู่ในสดมภ์ที่ 15 เช่น

สดมภ์ที่	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
คนที่ 1	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0
คนที่ 2	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1
:															
คนที่ 100	1	0	0	1	1	0	0	1	1	1	0	1	1	0	1

ลักษณะที่สอง การป้อนคะแนนของข้อสอบแต่ละข้อจะมีช่องว่างหรือจุลภาคคั่นระหว่างข้อ

คนที่ 1	0 0 0 0 1 0 0 0 0 0 1 0 0 1 0
คนที่ 2	1, 1 1 1 1 1, 1 1 1, 1 0 0 0 1
:	
คนที่ 100	1 0, 0 1 1 0 0 1 1 1 0, 1 1 0 1

ในเอกสารฉบับนี้จะใช้วิธีการป้อนข้อมูลตามลักษณะแรก ข้อมูลของกลุ่มอ้างอิงทั้งหมดถูกเก็บลงในแฟ้ม ur.dat

```
000010000010010
111111111100001
110001000000011
.....
.....
```

และข้อมูลของกลุ่มสนใจทั้งหมดถูกเก็บลงในแฟ้ม uf.dat

```
110100111001010
111101101101110
111111111101100
.....
.....
```

เราได้เพิ่มข้อมูลของกลุ่มอ้างอิง (ur.dat) และข้อมูลของกลุ่มสนใจ (uf.dat) แล้ว ก่อนจะทำการวิเคราะห์ความลำเอียง เราจำเป็นต้องตั้งค่าพารามิเตอร์ก่อน โดยใช้โปรแกรม SIBIN

การตั้งค่าพารามิเตอร์ใน SIBIN

ต้องใส่พารามิเตอร์ทั้งหมดที่ SIBIN ต้องการ ซึ่งจะใส่ในขณะรัน SIBIN โดยมีขั้นตอนดังนี้

```
C:SIBTEST>sibin
```

สั่งรันโปรแกรม SIBIN โดยพิมพ์ SIBIN แล้วกด ENTER โปรแกรมจะเริ่มทำงานโดยจะถามคำถาม และผู้ใช้จะต้องป้อนค่าพารามิเตอร์ต่าง ๆ ต่อไปนี้

```
This program uses user-supplied information, obtained in
an interactive session, to create the input parameter file
needed for running SIBTEST, the program that calculates the
Simultaneous Item Bias (SIB) hypothesis testing statistic
and the estimator of the amount of test bias/DIF.
```

Enter NAME for the INPUT PARAMETER FILE you are creating
that will be used to run SIBTEST.

Enter O if you wish to use the default name : SIB.INPUT

inchk.dat

โปรแกรมให้เรที่ตั้งชื่อแฟ้มที่ต้องการเก็บค่าพารามิเตอร์ ในที่นี้เราให้โปรแกรมเก็บค่าพารามิเตอร์ไว้ในแฟ้มชื่อ inchk.dat

Enter a TWO-LINE TITLE for the run(s) of SIBTEST that you
are creating an input parameter file for. Each line
may be up to 70 characters long.

Enter O if you wish the default title below :

***** SIMULTANEOUS ITEM BIAS ESTIMATION AND HYPOTHESIS TESTING *****

R. Shealy and W. Stout, programming assistance by L. Roussos

Sample Run of SIBIN and SIBTEST

Sunday, November 15, 1997

ใส่หัวข้อหรือข้อความที่จะเป็นการเตือนความทรงจำ จำนวน 2 บรรทัด

Enter the TOTAL NUMBER OF ITEMS that are
on the dataset you want to analyze.
Cannot be greater than 80

15

ใส่จำนวนข้อสอบที่จะวิเคราะห์ความลำเอียง โปรแกรมวิเคราะห์ได้ไม่เกิน 80 ข้อ ในที่นี้เรามีข้อสอบที่จะวิเคราะห์ 15 ข้อ

Enter NAME OF FILE containing REFERENCE GROUP test scores.

ur.dat

ใส่ชื่อแฟ้มข้อมูลที่เก็บผลคะแนนของกลุ่มอ้างอิง ในที่นี้เราเก็บไว้ในแฟ้มชื่อ ur.dat

Enter NAME OF FILE containing FOCAL GROUP test scores.

uf.dat

ใส่ชื่อแฟ้มข้อมูลที่เก็บผลคะแนนของกลุ่มสนใจ ในที่นี้เราเก็บไว้ในแฟ้มชื่อ uf.dat

Enter 1 if there are no spaces between the item scores

i.e., no spaces between the ones and zeroes.

Enter 2 if there is a space or comma between each item score.

1

ใส่รูปแบบการเก็บข้อมูล ใส่ 1 เมื่อไม่มีช่องว่างระหว่างข้อ และใส่ 2 เมื่อมีช่องว่างหรือจุลภาคคั่นระหว่างข้อ ในที่นี้เราป้อนข้อมูลแบบไม่มีช่องว่างระหว่างข้อ ใส่เลข 1

Enter NAME OF FILE to which the output of SIBTEST is to be written.

Enter 0 if you wish to use the default name : SIB.OUTPUT

outchk.dat

ใส่ชื่อแฟ้มข้อมูลที่ต้องการเก็บผลลัพธ์ที่ได้จากการประมวล ในที่นี้เราต้องการเก็บผลลัพธ์ไว้ในแฟ้มชื่อ outchk.dat

Enter MINIMUM REQUIRED NUMBER OF EXAMINEES per
statistic calculation cell.

This number must be an integer equal to 2 or more.

NOTE : Simulation studies have indicated that it is

desirable to include as many examinees as possible

in the statistical calculations, especially iff one

of the group sizes is small (e.g., less than 500).

So, the minimum value of 2 is recommended.

As a general rule of thumb, this number should be no larger

than the number of examinees in the smaller group (usually

the Focal group) divided by the number of items on the test.

2

ใส่จำนวนต่ำสุดของจำนวนผู้สอบที่ใช้ในการคำนวณ โดยปกติจะใส่ค่า 2

Enter ESTIMATE OF GUESSING on the test.

This is an estimate of the probability that an examinee of

very low ability will guess the right answer to an item on

the test (averages over all the items, if you think it's different for different items).

For a multiple-choice item with K options, a good estimate would be $1/K$ or $1/(K+1)$.

0.2

ใส่ค่าความน่าจะเป็นในการตอบถูก ในที่นี้ข้อสอบมี 5 ตัวเลือก = $1/5 = 0.2$

The final phase of the setting up of the input parameter file is the specification of the valid and suspect subtest items for the different runs of SIBTEST on your dataset. THERE ARE 3 OPTIONS :

Enter 1 to set up an automatic single-item DIF/bias study of all the items in a specified set of items. This set is often the entire test, but need not be. On each run of SIBTEST, on item will be tested for DIF/bias with the others serving as the valid subtest. Both SIB and Mantel-Haenszel statistics will be calculated.

Enter 2 to set up a study of a specified set of suspect items. Suspect subtests will be automatically set up for either all possible subsets of these items or for only each suspect item alone. The valid subtest paired with all these suspect subtests is either all the remaining items or some user-specified subset of the remaining items.

Enter 3 to set up a particular sequence of runs that you would like SIBTEST to do. For each run you will be asked to specify exactly which items are on the suspect subtest, which items are on the valid subtest, and, if necessary, which items are to be ignored.

1

ใส่เลขลำดับขั้นที่ต้องการให้โปรแกรมรัน

ใส่เลข 1 เมื่อจะคำนวณดัชนีความลำเอียงของข้อสอบที่สงสัยรายข้อพร้อมแสดงความลำเอียงของเมนเทลเฮนเซล ในข้อสอบที่สงสัย

ใส่เลข 2 เมื่อจะคำนวณดัชนีความลำเอียงของข้อสอบที่สงสัยรายข้อพร้อมกับแสดงความลำเอียงของข้อสอบที่สงสัยรวมทุกข้อ

ใส่เลข 3 เครื่องจะถามข้อที่ต้องการวิเคราะห์ความลำเอียงโดยจะถามทุก ๆ วัน ตั้งแต่รันครั้งที่ 1 เครื่องจะถามให้ใส่ข้อสอบที่ต้องการวิเคราะห์ 1 ข้อ รันครั้งที่ 2 เครื่องจะถามให้ใส่ข้อสอบที่ต้องการวิเคราะห์ 2 ข้อ โดยจำนวนข้อที่ป้อนจะเพิ่มขึ้นในแต่ละรัน

Enter NUMBER OF ITEMS suspected of bias DIF/DTF.

NOTE : If you are going to set up suspect subtests for every possible subset of the suspect items, then the number of suspect items is limited to 9.

15

ใส่จำนวนข้อสอบที่สงสัยว่าจะลำเอียง ในที่นี้สงสัย 15 ข้อ

Enter 1 to set up suspect subtests for every possible subset of the suspect items

Enter 2 to set up suspect subtests containing only single suspect items.

1

ใส่เลข 1 เพื่อจะจัดการกับข้อสอบที่สงสัยทุกข้อ

ใส่เลข 2 เพื่อจะจัดการกับข้อสอบเฉพาะข้อสอบที่สงสัย

Note : number of statistic calculation runs to be made = 15

Enter 0 to get the ABBREVIATED OUTPUT file which includes

- list off input parameter values common to all the runs
- the information written to the terminal screen, namely :
 - the run number
 - the bias/DIFF/DTF estimates for each run
 - the bias/DIF/DTF statistics for each run
 - and the p-value for each of the statistics

Note : This option does NOT include listings of the suspect and valid subtests for each run

Enter 1 to get the STANDARD OUTPUT file, which includes

- all of the abbreviated output plus, for each run:
 - listings of the suspect and valid subtests
 - and proportions of Reference and Focal examinees not used

in the calculation of bias/DIF/DTF statistics and estimates.

Enter 2 to get the EXPANDED OUTPUT, which includes all the standard output plus, for each run, detailed statistic calculation results including the average suspect subtest scores for the Reference and Focal groups for each statistic calculation cell.

2

ตั้งขนาดของผลลัพธ์ที่จะแสดงออกมาทางแฟ้มข้อมูล

ใส่ 0 แสดงผลลัพธ์แบบ ABBREVIATED OUTPUT

- จะแสดงจำนวนข้อสอบที่ลำเอียง
- ค่าความลำเอียงในแต่ละรัน
- การทดสอบสมมติฐานในแต่ละรัน
- แสดงว่า p-value

ใส่ 1 แสดงผลลัพธ์แบบ STANDARD OUTPUT

- จะให้ค่าเหมือนใส่ 0
- แสดงข้อที่สงสัยว่าจะลำเอียง
- แสดงสัดส่วนของกลุ่มอ้างอิงและกลุ่มสนใจที่ไม่ได้ใช้ในการคำนวณ

ใส่ 2 จะแสดงผลลัพธ์แบบ EXPANDED OUTPUT

- จะให้ค่าเหมือนใส่ 1
- แสดงคะแนนจริงที่ปรับแก้แล้วและก่อนปรับแก้

Enter 0 if you want the DEFAULT POOLED WEIGHTING.

Enter 1 if you want FOCAL-GROUP WEIGHTING IN ADDITION TO the POOLED WEIGHTING.

For each of the 15 runs of SIBTEST, the statistic SIB-uni will be calculated. By default, the statistic will be weighted by the pooled count of Ref. and Foc. grp. examinees in each cell. You may request that weighting by the Focal group also be used.

0

ใส่ 0 เมื่อให้น้ำหนักของกลุ่มอ้างอิงและกลุ่มสนใจเท่ากัน

ใส่ 1 เมื่อให้กลุ่มสนใจมีน้ำหนักมากกว่า

Enter 0 to specify and SINGLE TYPE OF P-VALUE for all the runs
 Enter 1 to SPECIFY the P-VALUE TYPE FOR EACH RUN individually
 Along with the SIB-uni z-statistic that will be calculated
 for each run, a p-value will also be calculated.
 This p-value will be calculated for one of three
 bias/DIF/DTF hypothesis tests :

- bias/DIF/DTF against either the Foc. or Ref. groups.
 (a two-tailed test for bias/DIF)
- bias/DIF/DTF against the Foc. grp. (a one-tailed test)
- bias/DIF/DTF against the Ref. grp. (a one-tailed test)

NOTE : The Mantel-Haenszel chi-square statistic is restricted,
 by definition, to the alternative hypothesis of
 bias/DIF against either the Focal or Reference groups.

0

ใส่ 0 เมื่อแสดงค่า p-value ในทุกรันและหลังจากรันทั้งหมดแล้ว
 ใส่ 1 เมื่อแสดงค่า p-value ในแต่ละรัน

Enter e to get p-value for bias/DIF/DTF AGAINST EITHER GROUP
 (to get the TWO-TAILED p-value)
 Enter f to get p-value for bias/DIF/DTF AGAINST FOCAL GROUP
 Enter r to get p-value for bias/DIF/DTF AGAINST REF. GROUP.

e

ใส่ e เมื่อแสดงค่า p-value ในกลุ่มทั้ง 2
 ใส่ f เมื่อแสดงค่า p-value ในกลุ่มสนใจ
 ใส่ r เมื่อแสดงค่า p-value ในกลุ่มอ้างอิง

You have now created the file : inchk.dat
 This file contains all the input parameters needed to run SIBTEST.
 To run SIBTEST, simply type that word and press the Return key.
 The first thing SIBTEST will ask you for is the name of the file
 you just created, inchk.dat

การรันโปรแกรม

พิมพ์ SIBTEST แล้วกด ENTER แล้วเครื่องจะถามชื่อแฟ้มที่เก็บพารามิเตอร์ ให้พิมพ์ชื่อแฟ้มลงไป ในที่นี้ให้ชื่อว่า inchk.dat แล้วกด ENTER โปรแกรมจะทำการประมวลผล

ผลลัพธ์ที่ได้จากการรัน SIBTEST

ผลลัพธ์ที่ได้จะแสดงรายละเอียดของข้อมูลในแต่ละวัน แต่รายละเอียดที่แสดงนี้จะมากหรือน้อยขึ้นอยู่กับข้อกำหนดพารามิเตอร์ผลลัพธ์ ผู้ใช้สามารถเลือกระดับของผลลัพธ์เป็น 0 หรือ 1 หรือ 2 ในการให้รายละเอียดสำหรับข้อมูลที่ต้องการในแต่ละวัน โปรแกรม SIBTEST จะรันโดยใช้พารามิเตอร์ที่อยู่ใน inchk.dat เมื่อประมวลผลแล้วจะเก็บผลลัพธ์ไว้ในไฟล์ชื่อ outchk.dat

ผลลัพธ์ชุดแรกจะแสดงรายละเอียดเกี่ยวกับผลลัพธ์ที่รันทั้งหมด ข้อมูลส่วนนี้จะแสดงรายละเอียดของผลลัพธ์ทั้งหมด ดังนี้

- ชื่อไฟล์ที่ใส่พารามิเตอร์
- จำนวนข้อสอบในแบบทดสอบ
- ชื่อไฟล์ที่เก็บข้อมูลของกลุ่มอ้างอิง
- ชื่อไฟล์ที่เก็บข้อมูลของกลุ่มสนใจ
- จำนวนต่ำสุดที่ต้องการของผู้สอบที่ใช้ในการคำนวณสถิติ
- ประมาณค่าการเดาของแบบทดสอบ
- จำนวนผู้สอบในกลุ่มอ้างอิง
- จำนวนผู้สอบในกลุ่มสนใจ

ถ้าเลือกผลลัพธ์แบบ STANDARD OUTPUT จะแสดงข้อมูลเพิ่มเติมดังนี้

- จำนวนวัน
- ข้อสอบที่สงสัยจะลำเอียง
- ข้อสอบที่ Valid

ถ้าเลือกผลลัพธ์แบบ EXPANDED OUTPUT จะแสดงข้อมูลเพิ่มเติมอีกดังนี้

- จำนวนผู้สอบในกลุ่มอ้างอิงที่ได้คะแนนในชุดของข้อสอบที่ Valid (NR)
- จำนวนผู้สอบในกลุ่มสนใจที่ได้คะแนนในชุดของข้อสอบที่ Valid (NF)
- ค่าเฉลี่ยของสัดส่วนของข้อสอบที่สงสัยลำเอียงของกลุ่มอ้างอิง (ybar-R)
- ค่าเฉลี่ยของสัดส่วนของข้อสอบที่สงสัยลำเอียงของกลุ่มสนใจ (ybar-F)
- ค่าเฉลี่ยที่ถูกปรับแก้แล้วของสัดส่วนของข้อสอบที่สงสัยลำเอียงของกลุ่มอ้างอิง (Adj. ybar-R)
- ค่าเฉลี่ยที่ถูกปรับแก้แล้วของสัดส่วนของข้อสอบที่สงสัยลำเอียงของกลุ่มสนใจ (Adj. ybar-F)
- ความแตกต่างของผลจากกลุ่มสนใจที่ปรับแก้แล้วกับกลุ่มอ้างอิงที่ปรับแก้แล้ว (D)

เมื่อ SIBTEST เป็นสถิติบ่งชี้ความลำเอียง ค่า D สามารถบอกโอกาสที่จะลำเอียงในระดับชั้นคะแนนนั้น ๆ ในการใช้ค่า D ควรจะแปลความหมายอย่างระมัดระวังมาก ๆ เพราะค่า D เพียงเล็กน้อยก็สามารถบ่งชี้ถึงความลำเอียงที่อยู่ภายในข้อสอบนั้น ๆ (local bias) และไม่สัมพันธ์กับการทดสอบทางสถิติ

- ความแตกต่าง (D) คุณด้วยน้ำหนักของกลุ่มทั้งสองเป็นการแจกแจงทางสถิติที่เรียกว่า เบต้ายูนิ (Beta-uni) (D^*wt) ถ้าผู้ใช้ให้น้ำหนักแก่กลุ่มสนใจ ในผลลัพธ์จะแสดงเป็นค่า D^*wt-p และ D^*wt-F

จากนั้นในผลลัพธ์แบบ EXPANDED OUTPUT จะให้ค่าในแต่ละวันดังนี้

- จำนวนที่เป็นไปได้ในแต่ละเซล
- สัดส่วนของจำนวนที่ถูกนำไปใช้

ในตอนท้ายของผลลัพธ์แบบนี้ซึ่งในผลลัพธ์แบบ STANDARD ด้วยคือ

- สัดส่วนของผู้สอบกลุ่มอ้างอิงที่ถูกคัดออก
- สัดส่วนของผู้สอบกลุ่มสนใจที่ถูกคัดออก
- ประมาณค่าความลำเอียง (Beta-uni) การทดสอบสมมติฐานทางสถิติด้วยสถิติ Z ค่าสถิติ p-value

ในกรณีที่เรากำหนดน้ำหนักให้กับกลุ่มสนใจ ในผลลัพธ์จะปรากฏค่า Beta-uni Focal, Beta-uni Pooled, Sib-uni Focal, และ Sib-uni Pooled

ผลสุดท้าย รายละเอียดที่เสนอมาททั้งหมดจะสรุปรวมในตอนท้าย

ในกรณีเลือกผลลัพธ์แบบ ABBREVIATED จะได้ดังนี้

- จำนวนวัน
- ประมาณค่าความลำเอียง bias/DIF/DTF ในแต่ละวัน/ข้อ
- ประมาณค่าสถิติในแต่ละวัน (SIB-uni Z-statistics)
- และค่า p ในแต่ละวัน (SIB-uni p-value)

ในข้อสอบที่สงสัยนั้นจะมีการทดสอบสมมติฐาน Mantel-Haenszel และแสดงออกมาเป็นค่า p และค่าความลำเอียง การทดสอบสมมติฐานของเมนเทลเฮนเซลนั้นจะใช้ไคสแควร์มี df เป็น 1

ตัวอย่างผลลัพธ์ที่ได้จากการวิเคราะห์

จากการวิเคราะห์ข้อสอบ 15 ข้อเพื่อวิเคราะห์ความลำเอียงของข้อสอบวิชาพลศึกษาระหว่างกลุ่มเพศชาย (กลุ่มอ้างอิง) และกลุ่มเพศหญิง (กลุ่มสนใจ) ผลลัพธ์ที่ได้จากการวิเคราะห์ทั้งหมดมีดังนี้

```
name of input parameter file = inchk.dat
number of items on test = 15
name of file for Ref. grp. scores = ur.dat
name of file for Focal grp. scores = uf.dat
```

minimum no. of examinees per statistic calculation cell = 2
 estimate of guessing on the test = .20
 number of runs for this data set = 15
 number of examinees in Reference group = 100
 number of examinees in Focal group = 100

แสดงชื่อแฟ้มข้อมูลที่เก็บค่าพารามิเตอร์ = inchk.dat
 แสดงจำนวนข้อสอบในแบบทดสอบ = 15
 แสดงชื่อแฟ้มข้อมูลที่เก็บข้อมูลของกลุ่มอ้างอิง = ur.dat
 แสดงชื่อแฟ้มข้อมูลที่เก็บข้อมูลของกลุ่มสนใจ = uf.dat
 จำนวนต่ำสุดของผู้สอบในการคำนวณทางสถิติ = 2
 ประมวลค่าการเดาของข้อสอบ = 0.2
 จำนวนครั้งที่รันข้อมูล = 15
 จำนวนผู้สอบในกลุ่มอ้างอิง = 100
 จำนวนผู้สอบในกลุ่มสนใจ = 100

OUTPUT FOR RUN NUMBER 1					OUTPUT FOR RUN NUMBER 1				
Suspect subtest items:									
1									
Valid subtest items:									
2 3 4 5 6 7 8 9 10 11									
12 13 14 15									
Valid									
Subtest					Adj.	Adj.			
Score	NR	NF	ybar-R	ybar-F	ybar-R	ybar-F	D	D*wt	
0	0	0	.0000	.0000	.0000	.0000	.0000	.0000	
1	0	0	.0000	.0000	.0000	.0000	.0000	.0000	
2	1	2	1.0000	.0000	1.0000	.0000	.0000	.0000	
3	2	8	.0000	.5000	.0000	.5205	-.5205	-.0283	
4	4	14	.5000	.5000	.4135	.5288	-.1153	-.0113	
5	10	6	.5000	.6667	.4662	.6579	-.1917	-.0167	
6	7	9	.7143	.4444	.6746	.4531	.2215	.0193	

7	9	11	.7778	.7273	.7555	.7982	-.0427	-.0046
8	18	9	.8889	1.0000	.8751	1.0000	-.1249	-.0183
9	10	10	.9000	1.0000	.8892	.9757	-.0865	-.0094
10	11	12	1.0000	.7500	1.0000	.7418	.2582	.0323
11	9	10	.8889	.9000	.8889	.9168	-.0279	-.0029
12	8	7	1.0000	1.0000	.9769	1.0000	-.0231	-.0019
13	8	1	1.0000	1.0000	1.0000	1.0000	.0000	.0000
14	3	1	1.0000	1.0000	1.0000	1.0000	.0000	.0000

no. of possible usable cells = 13
proportion of cells used = .769
proportion of Ref. grp. examinees eliminated = .120
proportion of Focal grp. examinees eliminated = .040

Beta-uni statistic	SIB-uni		Mantel-Haenszel Results			
	z	p-value for SIB-uni DIF against either Ref. or Foc. grp.	Chi sqr.	p-value for either Ref. or Foc. grp. (D-DIF)	Delta	MH alpha
-.042	-.721	.471	.00	.962	.170	.93

จำนวนครั้งที่รับ

ข้อสอบที่สงสัยว่าจะลำเอียง

ข้อสอบที่คงไว้ (Valid)

คะแนนของข้อสอบที่คงไว้ (Valid Subtest Score)

จำนวนคนในกลุ่มอ้างอิงที่ได้คะแนนในชั้นนั้น (NR)

จำนวนคนในกลุ่มสนใจที่ได้คะแนนในชั้นนั้น (NF)

คะแนนเฉลี่ยของกลุ่มอ้างอิง (ybar-R)

คะแนนเฉลี่ยของกลุ่มสนใจ (ybar-F)

คะแนนเฉลี่ยของกลุ่มอ้างอิงที่ขจัดความคลาดเคลื่อนออกแล้ว (Adj. ybar-R)

คะแนนเฉลี่ยของกลุ่มสนใจที่ขจัดความคลาดเคลื่อนออกแล้ว (Adj. ybar-F)

คะแนนความแตกต่างต่างของ Adj. ybar-R กับ Adj. ybar-F

ค่าประมาณความลำเอียง (Beta-uni) (D*wt)

จำนวนที่เป็นไปได้ในแต่ละเซลล์

สัดส่วนของเซลล์ที่ถูกใช้

สัดส่วนของผู้สอบในกลุ่มอ้างอิงที่ไม่ได้นำมาวิเคราะห์
 สัดส่วนของผู้สอบในกลุ่มสนใจที่ไม่ได้นำมาวิเคราะห์
 ดัชนีค่าความลำเอียงของข้อสอบที่สงสัย (Beta-uni)
 ค่าสถิติ Z (SIB-uni Z statistic)
 ค่า p-value (ค่า p-value จะต้องน้อยกว่าหรือเท่ากับ 0.05 จึงจะถือว่าลำเอียง)
 การทดสอบไคสแควร์ของค่าความลำเอียงแบบเมนเทลเฮนเซล
 ค่า p-value (ค่า p-value จะต้องน้อยกว่าหรือเท่ากับ 0.05 จึงจะถือว่าลำเอียง)
 ดัชนีความลำเอียงแบบเมนเทลเฮนเซลที่แปลงเป็นค่ามาตรฐาน (Delta DIF)
 ดัชนีความลำเอียงแบบเมนเทลเฮนเซล (MH alpha)

ผลสรุปในตอนท้ายจะแสดงค่าสถิติในแต่ละรัน โดยจะสรุปมาเฉพาะสถิติที่สำคัญดังนี้

SUMMARY OF THE RUNS							
p-value notation:							
R denotes p-value for test of bias/DIF/DTF against Ref. group							
F denotes p-value for test of bias/DIF/DTF against Foc. group							
E denotes p-value for test of bias/DIF/DTF against either the Ref. or Foc. group.							
NOTE: M-H Chi-square p-value is restricted, by definition, to type E.							
run	SIB-uni		Chi	Mantel-Haenszel			
no.	Beta-uni	z-statistic	p-value	sqr.	p	Delta	MH
					value	(D-DIF)	alpha
1	-.042	-.721	.471 E	.00	.962 E	.17	.93
2	-.034	-.519	.603 E	.18	.667 E	.60	.77
3	-.019	-.282	.778 E	.03	.862 E	-.02	1.01
4	-.061	-.996	.319 E	.82	.366 E	1.16	.61
5	.032	.456	.649 E	.01	.943 E	-.20	1.09
6	.021	.297	.766 E	.01	.920 E	.06	.98
7	.019	.256	.798 E	.01	.925 E	.06	.98
8	-.038	-.498	.618 E	2.22	.136 E	1.21	.60
9	-.031	-.500	.617 E	.14	.707 E	.55	.79
10	.033	.591	.554 E	.30	.582 E	-.94	1.49

11	.009	.160	.873 E	.02	.876 E	.34	.87
12	-.021	-.297	.767 E	.00	.966 E	.22	.91
13	-.047	-.661	.508 E	.55	.457 E	.75	.73
14	-.155	-2.158	.031 E	2.96	.085 E	1.53	.52
15	.454	7.062	.000 E	31.03	.000 E	-4.24	6.08

ดัชนีความลำเอียง (Beta-uni) (เป็นค่าที่นำมาใช้เขียนในรายงานการวิจัย)

ค่าสถิติ Z

ค่า p-value (ค่า P-value จะต้องน้อยกว่าหรือเท่ากับ 0.5 จึงจะถือว่าลำเอียง)

การทดสอบไคสแควร์ของดัชนีความลำเอียงแบบ MH

ค่า p-value (ค่า P-value จะต้องน้อยกว่าหรือเท่ากับ 0.5 จึงจะถือว่าลำเอียง)

ดัชนีความลำเอียง MH ที่แปลงเป็นค่ามาตรฐาน (Delta DIF)

ดัชนีความลำเอียง MH (MH alpha) (เป็นค่าที่นำมาใช้เขียนในรายงานการวิจัย)

การนำผลการวิเคราะห์ไปเขียนในรายงานการวิจัย

ในรายงานการวิจัยที่นำเสนอเกี่ยวกับการวิเคราะห์ความลำเอียงส่วนมากจะเป็นการเปรียบเทียบวิธีการวิเคราะห์ความลำเอียงหลาย ๆ วิธี หรือวิเคราะห์ความลำเอียงของข้อสอบที่มีความแตกต่างกันของกลุ่มตัวอย่าง ในกรณีที่เราวิเคราะห์ด้วยโปรแกรม SIBTEST นอกจากจะได้ค่าดัชนีความลำเอียงของ SIBTEST แล้ว ยังได้ค่าดัชนีความลำเอียงของเมนเทลเฮนเซล (Mantel Haenszel) อีกด้วย

จากผลลัพธ์ในการวิเคราะห์ความลำเอียงของข้อสอบ 15 ข้อ สามารถนำมาเขียนรายงานวิจัย โดยนำค่าดัชนีความลำเอียง (Beta-uni) มาใส่ตาราง

ตาราง 1 ค่าความลำเอียงของข้อสอบของแบบทดสอบวิชาพลศึกษา ที่วิเคราะห์ด้วยวิธี SIBTEST
จำแนกตามเพศ เมื่อใช้เพศชายเป็นกลุ่มได้เปรียบ

ข้อที่	ดัชนีความลำเอียง (Beta-uni)
1	-0.042
2	-0.034
3	-0.019
4	-0.061
5	0.032
6	0.021
7	0.019
8	-0.038
9	-0.031
10	0.033
11	0.009
12	-0.021
13	-0.047
14	-0.155
15	0.454

จากตาราง 1 พบว่า กลุ่มเพศหญิงทำแบบทดสอบวิชาพลศึกษาแล้วมีข้อสอบที่ลำเอียง โดยมีค่าดัชนีความลำเอียงอยู่ระหว่าง -0.155 ถึง 0.454 แต่มีข้อสอบที่ลำเอียงอย่างมีนัยสำคัญทางสถิติ 2 ข้อคือข้อ 14 และ 15 โดยข้อที่ 14 กลุ่มเพศชายเสียเปรียบกลุ่มเพศหญิง และในข้อที่ 15 กลุ่มเพศหญิงเสียเปรียบกลุ่มเพศชาย นอกนั้นเป็นข้อสอบที่ไม่ลำเอียง

หนังสืออ้างอิง

- เรวดี อินทสระ. ผลการตรวจสอบความลำเอียงของข้อสอบต่อการศึกษาความเที่ยงตรงเชิงพยากรณ์ของแบบทดสอบคัดเลือกที่คิดคะแนนต่างกัน. ปรินญานิพนธ์ กศ.ด. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ, 2539.
- อังคณา สายยศ. การศึกษาผลการวิเคราะห์ความลำเอียงของแบบทดสอบความเข้าใจในการอ่านตามลักษณะการใช้ภาษาพูดในครอบครัว. ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ, 2539.
- Stout William and Louis Roussos. SIBTEST USER MANUAL. Department of Statistics, University of Illinois, 1992.