

# โปรแกรมสำเร็จรูปทางการวัดผลและวิจัยการศึกษา : IRT-BAY

โปรแกรมสำเร็จรูปมากมายหลายร้อยโปรแกรม แต่โปรแกรมทางการวัดผลและวิจัยการศึกษา มีอยู่เพียงไม่กี่โปรแกรมเท่านั้น โดยส่วนมากแล้วจะเป็นแกรมที่มาจากต่างประเทศ แต่โปรแกรมที่จะนำเสนอในที่นี้คือโปรแกรม IRT-BAY ซึ่งเป็นโปรแกรมที่เขียนโดยคนไทย ซึ่งก็คือท่านรองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี และรองศาสตราจารย์ ดร.คณิต ไข่มุก ซึ่งเป็นหนึ่งในไม่กี่โปรแกรมทางการวัดผลและวิจัยการศึกษาที่เขียนโดยคนไทย

## คุณลักษณะของโปรแกรม

1. เป็นโปรแกรมที่ใช้กับระบบปฏิบัติการ DOS ด้วยเนื้อที่เพียง 270 MB
2. ใช้คำนวณหาคุณภาพของแบบทดสอบในกรณีที่ตรวจให้คะแนนแบบผิดให้ 0 คะแนน ถูกให้ 1 คะแนน
3. วิเคราะห์ข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)
4. สามารถประมาณค่าความสามารถของผู้สอบ (q) ค่าอำนาจจำแนก (a)

ค่าความยาก (b) และค่าการเดา (c)

5. ใช้วิธีประมาณค่าพารามิเตอร์แบบเบย์เซียน (Bayesian)
6. สามารถวิเคราะห์ข้อสอบรายข้อตามทฤษฎีมาตรฐานเดิม (Classical Test Theory) โดยจะ

แสดงค่าความยาก (p) ค่าอำนาจจำแนก (Rbis, Rpbis) และความเชื่อมั่น

(KR-20)

7. แสดงโค้งสารสนเทศของข้อสอบ (Item Information Function)

## คุณลักษณะของข้อมูลนำเข้า

ข้อมูลที่นำเข้าวิเคราะห์จะประกอบด้วย 2 ส่วนคือ

ส่วนแรกจะอยู่ในบรรทัดแรก จะเป็นตัวเลขของข้อสอบ

ส่วนที่สองตั้งแต่บรรทัดที่สองเป็นต้นไป จะเป็นคำตอบของผู้สอบ

## ตัวอย่าง

แบบทดสอบฉบับหนึ่งมี 30 ข้อ แต่ละข้อมี 5 ตัวเลือก สอบกับกลุ่มตัวอย่าง 340 คน  
รูปแบบการป้อนข้อมูลมีดังนี้

```

Key 342514324122544323211215554415
001 353254244342525134531354521225
002 342514324122544233441424234315
003 342435351122225555221255225211
.....ข้อมูลตั้งแต่คนที่ 004 จนถึง 337.....
338 542434354342534323533322331324
339 342534324422525515411234551415
340 342414324122544345311215154415

```

ในตัวอย่าง บรรทัดแรกจะเป็นบรรทัดที่ป้อนเฉลย ตัวเลขที่ปรากฏเป็นรหัสที่แทน  
ตัวเลือกทั้ง 5 ตัวเลือก คือตัวเลข 1 แทนตัวเลือก ก. ตัวเลข 2 แทนตัวเลือก ข. ตัวเลข 3 แทน  
ตัวเลือก ค. ตัวเลข 4 แทนตัวเลือก ง. และตัวเลข 5 แทนตัวเลือก จ.

ตั้งแต่บรรทัดที่สองเป็นต้นมา จะเป็นคำตอบในแต่ละข้อของกลุ่มตัวอย่างจำนวน 340 คน

สังเกตจำนวนของรหัส ID เราใส่ใน 3 สดมภ์ (Column) คือสดมภ์ที่ 1 ถึง 3 และจะวางเว้นในสดมภ์  
ที่ 4 เริ่มต้นข้อสอบข้อที่ 1 ในสดมภ์ที่ 5 ดังนั้นเราสามารถเขียนเป็นรูปแบบของ  
ข้อมูลที่ป้อน ซึ่งต้องใช้ในการวิเคราะห์ด้วยโปรแกรม IRT-BAY ได้ว่า

(A3,5X,30f1.0)

A3 หมายถึง จำนวนสดมภ์ของรหัส ID

5X หมายถึง สดมภ์เริ่มต้นข้อสอบข้อที่ 1

30f1.0 หมายถึง จำนวนข้อสอบที่ป้อนทั้งหมด 30 ข้อ

ถ้าหากข้อมูลที่มีรูปแบบแตกต่างไปจากนี้ก็สามารถเขียนใหม่ได้โดยเปลี่ยนเฉพาะตัวเลข เช่น  
จำนวนสดมภ์ที่ป้อน ID มี 10 สดมภ์ เริ่มต้นข้อสอบข้อที่ 1 ในสดมภ์ที่ 11 ซึ่งมี

ข้อสอบจำนวน 50 ข้อ สามารถเขียนเป็นรูปแบบของข้อมูลที่ป้อนได้ว่า (A10,11X,50f1.0)

จากตัวอย่างข้อสอบ 30 ข้อสอบ แต่ละข้อมี 5 ตัวเลือก สอบกับกลุ่มตัวอย่าง 340 คน บันทึกข้อมูล  
ไว้ในแฟ้มที่ชื่อว่า nc3new.ans

### การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล สามารถเลือกวิเคราะห์ได้ 1 พารามิเตอร์ หรือ 2 พารามิเตอร์ หรือ 3 พารา  
มิเตอร์ โดยการพิมพ์คำสั่งดังนี้

```
D:\IRT_BAY>B1
```

สำหรับวิเคราะห์ 1 พารามิเตอร์คือความยาก (b)

```
D:\IRT_BAY>B2
```

สำหรับวิเคราะห์ 2 พารามิเตอร์คือความยาก (b) และอำนาจจำแนก (a)

D:\IRT\_BAY>B3

สำหรับวิเคราะห์ 3 พารามิเตอร์คือความยาก (b) อำนาจจำแนก (a) และการเดา (c)

เมื่อเลือกวิเคราะห์ที่พารามิเตอร์ก็ตาม โปรแกรมจะถามชื่อแฟ้มข้อมูล (Name of data file : ) ให้พิมพ์ชื่อแฟ้มข้อมูลลงไปดังนี้

Name of data file : nc3new.ans

เมื่อคลิกปุ่ม Enter แล้วโปรแกรมจะถามถึงรูปแบบของข้อมูลที่ป้อน (Format of data as (A4,6X,60f1.0) : ) ให้ป้อนรูปแบบของข้อมูลที่ป้อนลงไป ดังนี้

Format of data as (A4,6X,60f1.0) : (A3,5X,30f1.0)

เมื่อคลิกปุ่ม Enter แล้วโปรแกรมจะถามจำนวนกลุ่มตัวอย่างที่ป้อนและจำนวนข้อสอบที่จะวิเคราะห์ ในที่นี้เรามีกลุ่มตัวอย่างจำนวน 340 คน และข้อสอบที่จะวิเคราะห์จำนวน 30 ข้อ เราสามารถป้อนข้อมูลได้ดังนี้

Input number of subject,item : 340,30

เมื่อคลิกปุ่ม Enter โปรแกรมจะทำการประมวลผลทันที

### ผลลัพธ์ที่ได้จากการวิเคราะห์

ผลลัพธ์ที่ได้จากการวิเคราะห์จะถูกเก็บอยู่ใน 4 แฟ้ม คือ

- \*.CTT จะเก็บข้อมูลผลการวิเคราะห์ข้อสอบรายข้อตามทฤษฎี Classical Test Theory โดยจะแสดงคะแนนดิบของกลุ่มตัวอย่างแต่ละคน คะแนนเฉลี่ย คะแนนเบี่ยงเบนมาตรฐาน ค่าความยาก (p) ค่าอำนาจจำแนก (Rbis, Rpbis) และค่าความเชื่อมั่น (KR-20)
  - \*.ABL จะเก็บค่าพารามิเตอร์ความสามารถของผู้สอบแต่ละคน ความสามารถเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของความสามารถของผู้สอบ
  - \*.BAY จะเก็บค่าพารามิเตอร์ของข้อสอบ คือค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c)
  - \*.GRP จะเก็บโครงสร้างสารสนเทศของข้อสอบรายข้อ (Item Information Function)
- เครื่องหมาย \* คือชื่อแฟ้ม ซึ่งจะแปรเปลี่ยนไปตามจำนวนพารามิเตอร์ที่วิเคราะห์ ดังนี้

กรณีวิเคราะห์ 1 พารามิเตอร์ การวิเคราะห์จะใช้คำสั่ง B1 ผลการวิเคราะห์จะแสดงอยู่ในแฟ้มชื่อ B1.CTT, B1.ABL, B1.BAY และ B1.GRP

กรณีวิเคราะห์ 2 พารามิเตอร์ การวิเคราะห์จะใช้คำสั่ง B2 ผลการวิเคราะห์จะแสดงอยู่ในแฟ้มชื่อ B2.CTT, B2.ABL, B2.BAY และ B2.GRP

กรณีวิเคราะห์ 3 พารามิเตอร์ การวิเคราะห์จะใช้คำสั่ง B3 ผลการวิเคราะห์จะแสดงอยู่ในแฟ้มชื่อ B3.CTT, B3.ABL, B3.BAY และ B3.GRP

### เกณฑ์การพิจารณาข้อสอบรายข้อ

ผลการวิเคราะห์ตามทฤษฎีมาตรฐานเดิม (Classical Test Theory) ในโปรแกรมนี้จะแสดงค่าความยากและอำนาจจำแนกเฉพาะตัวถูกเท่านั้น ซึ่งในทางปฏิบัติจะต้องวิเคราะห์หาคุณภาพทั้งตัวถูกและตัวลวงเพื่อจะได้ปรับปรุงให้เป็นข้อสอบที่มีคุณภาพ แต่อย่างไรก็ตามโปรแกรมนี้เน้นวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) ซึ่งจะไม่สามารถวิเคราะห์รายตัวเลือกได้อยู่แล้ว ดังนั้นการดูผลการวิเคราะห์ตามทฤษฎีมาตรฐานเดิมจึงเป็นการดูตัวอย่างหยาบ คือดูเฉพาะตัวถูกเท่านั้น

ค่า P ที่ปรากฏจะบ่งบอกถึงระดับความยากของข้อสอบ ถ้าค่า P สูงแสดงว่าข้อสอบยาก ถ้าค่า P ต่ำแสดงว่าข้อสอบง่าย โดยเกณฑ์ของข้อสอบที่มีคุณภาพ ค่า P จะมีค่าอยู่ระหว่าง .20 ถึง .80

ค่า Rbis และ Rpbis จะบ่งบอกถึงค่าอำนาจจำแนกของข้อสอบ ถ้ามีค่าสูงแสดงว่าข้อสอบสามารถจำแนกบุคคลออกเป็นกลุ่มเก่งและกลุ่มอ่อนได้มาก ถ้ามีค่าต่ำแสดงว่าข้อสอบสามารถจำแนกบุคคลออกเป็นกลุ่มเก่งและกลุ่มอ่อนได้น้อย โดยปกติค่าอำนาจจำแนกของข้อสอบที่มีคุณภาพจะมีค่าอยู่ระหว่าง .20 ถึง 1.00

ผลการวิเคราะห์หาคุณภาพรายข้อตามทฤษฎีการตอบสนองข้อสอบ จะพิจารณาอยู่ 3 ตัวคือ

ค่า a หมายถึงค่าอำนาจจำแนก ถ้า a มีค่ามาก (เป็นบวก) แสดงว่าข้อสอบจำแนกคนเก่งและคนอ่อนได้มาก ถ้า a มีค่าน้อย (เป็นลบ) แสดงว่าข้อสอบจำแนกคนเก่งและคนอ่อนได้น้อย โดยปกติข้อสอบที่มีคุณภาพควรมีค่า a ตั้งแต่ .30 ขึ้นไป

ค่า b หมายถึงค่าความยาก ถ้า b มีค่ามาก (เป็นบวก) แสดงว่าข้อสอบยาก ถ้า b มีค่าน้อย (เป็นลบ) แสดงว่าข้อสอบง่าย โดยปกติข้อสอบที่มีคุณภาพควรมีค่า b อยู่ระหว่าง -2.0 ถึง 2.0

ค่า c หมายถึงค่าการเดา ถ้าค่า c มีค่ามากแสดงว่าข้อสอบสามารถเดาได้มาก ถ้าค่า c น้อยแสดงว่าข้อสอบสามารถเดาได้น้อย โดยปกติข้อสอบที่มีคุณภาพควรมีค่า c ต่ำกว่า .30

