

# เอกสารประกอบ : การฝึกปฏิบัติการใช้โปรแกรมวิเคราะห์ข้อสอบ

## 1. การติดตั้งโปรแกรม

- Down Load โปรแกรมวิเคราะห์ข้อสอบ [Setup Items Analysis มหาวิทยาลัย เทคโนโลยีสุรนารี.exe]  
จาก <http://www.sut.ac.th/ces/>
- Pass word 1J813MH13YK8M7

## 2. การเตรียมข้อมูลสำหรับการวิเคราะห์

ตัวอย่าง : การวิเคราะห์ข้อสอบ วิชา 100000

- จำนวนข้อสอบ 50 ข้อ
- จำนวนผู้เข้าสอบ 120 คน
- จำนวนตัวเลือก 5 ตัวเลือก

2.1 File ที่อ่านจากเครื่องตรวจกระดาษคำตอบ (M100000)

2.2 File เฉลย (M100000k)

2.3 File ที่จะนำเข้าโปรแกรมวิเคราะห์ข้อสอบ (M100000\_1) ซึ่งมีเฉลยข้อสอบอยู่บรรทัดบนสุด


ในการจัดเตรียมข้อมูลตาม 2.3 นั้น ใช้โปรแกรม Text Editor (EditPlus) สำหรับสร้าง และแก้ไขข้อความ

## 3. การนำเข้าข้อมูลเพื่อวิเคราะห์

**ข้อกำหนดการใช้โปรแกรมวิเคราะห์ข้อสอบ**

- 1) ไฟล์ข้อมูลที่นำมาวิเคราะห์ต้องเป็นไฟล์ที่ได้จากเครื่องตรวจข้อสอบของบริษัทคอนโทรล ดาต้า
- 2) จำนวนผู้สอบมากกว่าหรือเท่ากับ 12 คน
- 3) มีบรรทัดของเฉลยอยู่บรรทัดบนสุดเพียงบรรทัดเดียว
- 4) ชื่อ PATH และชื่อไฟล์ของข้อมูลที่นำมาวิเคราะห์ห้ามมีช่องว่าง

**ขั้นตอน**

- 1) เปิดโปรแกรมวิเคราะห์ข้อสอบ
- 2) นำข้อมูลเข้าระบบ
- 3) กำหนดชื่อไฟล์,  
ตำแหน่งเริ่มต้นเลขที่ (41) ความยาว (8)  
ตำแหน่งเริ่มต้นข้อสอบ (60) จำนวนข้อ (50)  
จำนวนตัวเลือก (5) ➡ 

#### 4. การประมวลผลข้อมูล

จะมีให้เลือกวิเคราะห์แบบอิงกลุ่ม และอิงเกณฑ์

- การประเมินแบบอิงกลุ่ม : ตัดสินความสามารถของผู้เรียนโดยเทียบกับความสามารถของกลุ่ม  
การวิเคราะห์ข้อสอบแบบอิงกลุ่ม : วิเคราะห์ความยาก (p) และอำนาจจำแนก (r)
- การประเมินแบบอิงเกณฑ์ : ตัดสินความสามารถของผู้เรียนเทียบกับเกณฑ์ที่กำหนด  
การวิเคราะห์ข้อสอบแบบอิงเกณฑ์ : ไม่เน้นหาค่าความยากของข้อสอบ จะเน้นที่ค่าอำนาจจำแนก  
ผู้ผ่าน-ไม่ผ่านเกณฑ์

#### 5. รายงานค่าสถิติต่าง ๆ

##### 5.1 รายงานการวิเคราะห์ข้อสอบในรูปแบบของค่าสถิติที่เป็นตัวเลข

- 5.1.1 รายงานรายคน
- 5.1.2 รายงานรายข้อ (Item analysis 1, 2)
- 5.1.3 รายงาน Test Score Distribution
- 5.1.4 รายงาน Relative Frequency
- 5.1.5 รายงานสรุปค่าสถิติ (Summary)
- 5.1.6 รายงานค่าการกระจาย

##### 5.2 รายงานการวิเคราะห์ข้อสอบในรูปแบบข้อความ

- 5.2.1 รายงานการวิเคราะห์รายข้อ.
  - 1) วิเคราะห์ค่าความยากง่ายและค่าอำนาจจำแนกแบบรายข้อ.
  - 2) กราฟ
  - 3) รายงานสรุปค่าเฉลี่ย
  - 4) รายงานสรุปคุณภาพของข้อสอบ.
- 5.2.2 รายงานการกระจายของค่า p
- 5.2.3 รายงานการกระจายของค่า r
- 5.2.4 รายงานการกระจายของค่า delta

## การวิเคราะห์ข้อสอบ

การวิเคราะห์ข้อสอบ มี 2 แบบ คือ

1. การวิเคราะห์ข้อสอบรายข้อ : เพื่อดูคุณสมบัติของข้อสอบแต่ละข้อในด้านต่าง ๆ คือ
  - ระดับความยาก (Level of difficulty หรือ p )
  - อำนาจจำแนก (Power of discrimination หรือ r )
  - ประสิทธิภาพของตัวเลือกตัวลวง
2. การวิเคราะห์ข้อสอบทั้งฉบับ : เพื่อตรวจสอบคุณภาพของเครื่องมือวัด (ข้อสอบ)

### การวิเคราะห์ข้อสอบรายข้อ (Item Analysis)

- **ระดับความยาก (P)** หมายถึง ข้อสอบแต่ละข้อนั้นมีระดับความยากมากน้อยเพียงใด เหมาะแก่การนำไปใช้ต่อหรือไม่ ซึ่งจะคิดจากสัดส่วนของจำนวนผู้ตอบข้อสอบข้อนั้นได้ถูกต้องกับจำนวนผู้เข้าสอบทั้งหมด โดยจะมีค่าอยู่ระหว่าง 0 ถึง 1  
เกณฑ์ในการพิจารณา

ระดับความยากของข้อสอบ (p)	ความหมาย
> 0.95	เป็นข้อสอบที่ง่ายมาก
0.81 - 0.94	เป็นข้อสอบที่ง่าย
0.20 - 0.80	เป็นข้อสอบที่ดี
< 0.20	เป็นข้อสอบที่ยากถึงยากเกินไป

ถ้าค่า p ยิ่งมาก แสดงว่าข้อสอบนั้น ยิ่งง่าย แต่ถ้าค่า p ยิ่งน้อย แสดงว่าข้อสอบนั้นยิ่งยาก

ดังนั้น ระดับความยากของข้อสอบ (p) ควรมีค่าระหว่าง 0.2 - 0.8 ซึ่งถือว่ามีความยากง่ายพอเหมาะ

ข้อสอบที่ระดับความยาก < 0.2 หรือ > 0.8 ควรตัดทิ้ง

- **อำนาจจำแนก (r)** หมายถึง ข้อสอบแต่ละข้อนั้น มีอำนาจจำแนกผู้ที่เรียนเก่งและเรียนอ่อนออกจากกันได้มากเพียงใด ข้อสอบที่มีอำนาจจำแนกดี คือ ข้อสอบที่ผู้เข้าสอบในกลุ่มเก่งตอบถูกมากกว่าผู้เข้าสอบในกลุ่มอ่อน  
ค่า r จะมีค่าระหว่าง + 1.00 ถึง -1.00

เกณฑ์ในการพิจารณา

อำนาจจำแนกของข้อสอบ (r)	ความหมาย
$\geq 0.40$	เป็นข้อสอบที่มีอำนาจจำแนกดีมาก
0.30 - 0.39	เป็นข้อสอบที่มีอำนาจจำแนกดี
0.20 - 0.29	เป็นข้อสอบที่มีอำนาจจำแนกพอใช้
< 0.20	เป็นข้อสอบที่มีอำนาจจำแนกต่ำ
0	เป็นข้อสอบที่ไม่มีอำนาจจำแนกเลย

ถ้าค่า r เป็นบวก แสดงว่าข้อสอบข้อนั้น (หรือตัวเลือกนั้น) คนเก่งตอบถูกมากกว่าคนอ่อน แต่ถ้ามีค่าติดลบ แสดงว่าข้อสอบข้อนั้น (หรือตัวเลือกนั้น) คนอ่อนตอบถูกมากกว่าคนเก่ง ค่าอำนาจจำแนก (r) ควรมีค่ามากกว่า 0.20 ขึ้นไป จึงจะเป็นข้อสอบที่มีอำนาจจำแนกอยู่ในเกณฑ์ใช้ได้ ค่า r ต่ำกว่า 0.20 ควรตัดทิ้ง

### ■ ประสิทธิภาพของตัวเลือกตัวลง

นอกจากตัวเลือกที่ถูกต้องแล้ว ยังต้องการทราบว่าตัวลงของสอบแต่ละข้อนั้น มีลักษณะที่ดีหรือไม่ ถ้าไม่ดีจะได้หาทางปรับปรุงให้ดีขึ้นต่อไป โดยดูจากค่า  $r$  ข้อสอบที่มีคุณภาพดีจะมีลักษณะของตัวเลือกและตัวลงดังนี้

- ตัวเลือกที่ถูกต้อง คนในกลุ่มเก่งต้องเลือกมากกว่ากลุ่มอ่อน
- มีการกระจายของผู้เลือกตอบตามตัวลงต่าง ๆ
- คนในกลุ่มอ่อนต้องเลือกตัวลงมากกว่าคนในกลุ่มเก่ง (ค่า  $r$  ของตัวลงจะติดลบ)

#### ตัวอย่าง : ข้อที่ 1

P = ค่าความยาก     $r$  = อำนาจจำแนก    \* = คำตอบที่ถูกต้อง

ข้อสอบ	ตัวเลือก	กลุ่มสูง	กลุ่มต่ำ	P	r	ผลการวิเคราะห์
1	1	2	3	0.079	0.031	คำตอบลง ค่าอำนาจจำแนกต่ำ
	*2	28	18	0.719	0.313	ข้อที่ค่อนข้างง่าย ค่าอำนาจการจำแนกดี
	3	2	10	0.188	0.250	คำตอบลง ค่าอำนาจการจำแนกพอใช้
	4	0	1	0.013	0.031	คำตอบลง ค่าอำนาจการจำแนกต่ำ
	5	0	0	0.000	0.000	คำตอบลง ไม่มีค่าอำนาจการจำแนก
	oth	0	0	0.000	0.000	

วิเคราะห์ : เป็นข้อสอบที่มีคุณภาพดี ถ้าดูข้อถูก \*2 ค่า  $P = 0.719$  แสดงว่าค่อนข้างไปทางง่าย ค่า  $r = 0.313$  ดี แสดงว่าแยกคนเก่งกับคนอ่อนได้ ควรเก็บไว้

#### ตัวอย่าง : ข้อที่ 5

ข้อสอบ	ตัวเลือก	กลุ่มสูง	กลุ่มต่ำ	P	r	ผลการวิเคราะห์
5	1	0	3	0.047	0.094	คำตอบลง ค่าอำนาจจำแนกต่ำ
	2	0	0	0.000	0.000	คำตอบลง ไม่มีค่าอำนาจจำแนก
	*3	32	29	0.953	0.094	ข้อที่ง่ายมาก ค่าอำนาจการจำแนกต่ำ
	4	0	0	0.000	0.000	คำตอบลง ไม่มีค่าอำนาจจำแนก
	5	0	0	0.000	0.000	คำตอบลง ไม่มีค่าอำนาจจำแนก
	oth	0	0	0.000	0.000	

วิเคราะห์ : ข้อสอบคุณภาพไม่ดี ถ้าดูข้อถูก \*3 ค่า  $P = 0.953$  แสดงว่าข้อสอบที่ง่ายมาก ค่า  $r = 0.094$  อำนาจจำแนกต่ำ ควรตัดทิ้ง

## การวิเคราะห์ข้อสอบทั้งชุด (Analysis of total test)

### SUMMARY

11/25/10

Type	Mean	Min	Median	Max	Stddev	Var
TEST SCORES	30.767	19.0000	31.000	42.000	5.237	27.429
DIFF. INDEX(p)	0.610	0.1410	0.657	0.953	0.208	0.043
DELTA	11.068	0.0000	11.150	17.400	3.315	10.988
DISC. INDEX(d)	0.263	-0.0940	0.250	0.625	0.164	0.027
BISERIAL(r)	0.281	-0.1190	0.261	0.895	0.188	0.035
POINT-BISERIAL RPB	0.238	-0.1000	0.217	0.466	0.131	0.017

#### KUDER-RICHARDSON RELIABILITY STATISTICS

KR20 = 0.666      SEM20 = 3.025  
 KR21 = 0.580      SEM21 = 3.394

#### CRONBACH ALPHA RELIABILITY STATISTICS

ALPHA = 0.666      SEM-ALPHA = 3.025

#### SPLIT-HALF RELIABILITY STATISTICS

RTT = 0.673      SEMTT = 2.997

[จากรายงานการวิเคราะห์ข้อสอบในรูปแบบของค่าสถิติที่เป็นตัวเลข: รายงานสรุปค่าสถิติ (Summary)]

## การพิจารณาผลการวิเคราะห์ข้อสอบทั้งชุด

1. พิจารณาระดับความยากเฉลี่ยของข้อสอบ (p) โดยดูจากค่า DIFF. INDEX
2. พิจารณาค่าอำนาจจำแนกเฉลี่ยของข้อสอบ (r) โดยดูจากค่า DISC. INDEX
3. พิจารณาค่าความเที่ยงของข้อสอบ (Test reliability) หมายถึง ความคงเส้นคงวาของค่าที่วัดได้ เมื่อวัดซ้ำหลาย ๆ ครั้ง โดยใช้ความเที่ยงแบบสอดคล้องภายใน ซึ่งเป็นความสอดคล้องกันระหว่างคะแนนรายข้อ ค่าความเที่ยงมีค่าระหว่าง 0 – 1 ถ้าค่าความเที่ยงเข้าใกล้ 1 แสดงว่าข้อสอบนั้นมีค่าความเที่ยงสูง คะแนนที่นักศึกษาสอบได้จะมีความคลาดเคลื่อนน้อย ถ้าค่าความเที่ยงเข้าใกล้ 0 แสดงว่าข้อสอบนั้นมีค่าความเที่ยงต่ำ คะแนนที่นักศึกษาสอบได้จะมีความคลาดเคลื่อนมาก โดยดูจากค่า KR20
4. พิจารณาค่าความคลาดเคลื่อนมาตรฐานในการวัด (Standard error of measurement : SEM) โดยดูจากค่า SEM20 เป็นค่าที่แสดงให้เห็นว่า การวัดนั้นมีความคลาดเคลื่อนมากน้อยเพียงใด เมื่อคิดเป็นคะแนนสอบครั้งนั้น ถ้ามีความเที่ยงสูง ค่าความคลาดเคลื่อนมาตรฐานในการวัดจะต่ำ

## สาเหตุที่ข้อสอบมีคุณภาพไม่ดี

1. มีความบกพร่องที่ตัวข้อสอบเอง เช่น คำถามไม่ชัดเจน หรือเฉลยผิด
2. ออกข้อสอบไม่ตรงตามวัตถุประสงค์ และเนื้อหาวิชาที่กำหนดไว้ เช่น ถามเรื่องที่ไม่สำคัญ หรือตัวเลขที่ไม่จำเป็นต้องจำ
3. ความบกพร่องที่การเรียนการสอน ไม่เน้นเรื่องที่สำคัญและต้องรู้ แต่นำเรื่องนั้นมาออกข้อสอบ