

371.26013

ร 376 ก

5.3

การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลอง
โลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซ์ิมัม ไลค์ลิตูด วิธีนิวตัน-ราฟสัน
และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน
ด้วยเทคนิคมอนติ คาร์โล

ปริญญาโท

ของ

รัตนา ศรีเหรียญ

14 ส.ค. 2540

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรการศึกษาคณะศึกษาศาสตร์ วิชาเอกการทดสอบและวัดผลการศึกษา

มิถุนายน 2539

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

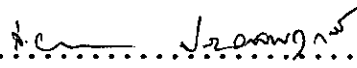
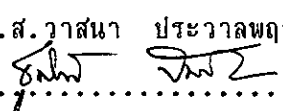
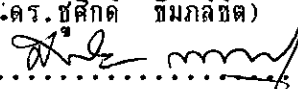
อ. 53170

ปริญญานิพนธ์ฉบับนี้ ได้รับเงินทุนช่วยเหลือการวิจัยจาก "ทุนสิรินธร" ซึ่งสมเด็จพระเทพรัตนราชสุดาฯ ได้ทรงกรุณาโปรดเกล้าพระราชทานมหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อพระราชทานแก่นิสิตในการทำวิจัยทางการศึกษา นับเป็นพระกรุณาธิคุณอย่างหาที่สุดมิได้

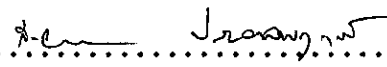
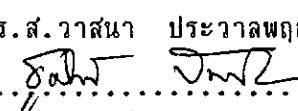
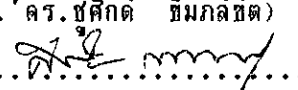
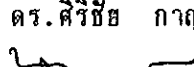
รัตนา ศรีเหรียญ

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญานิพนธ์ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาคุษฎีบัณฑิต สาขาการทดสอบและวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

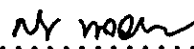
คณะกรรมการควบคุม

.....  ประธาน
(รศ.ดร.ส.วาสนา ประवालพฤษ์)
.....  กรรมการ
(อ.ดร.ชศักดิ์ ทัมภลิขิต)
.....  กรรมการ
(รศ.ดร.ศิริชัย กาญจนวาสี)

คณะกรรมการสอบ

.....  ประธาน
(รศ.ดร.ส.วาสนา ประवालพฤษ์)
.....  กรรมการ
(อ.ดร.ชศักดิ์ ทัมภลิขิต)
.....  กรรมการ
(รศ.ดร.ศิริชัย กาญจนวาสี)
.....  กรรมการ
(อ.ดร.บุญลือ ทองอยู่)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษาคุษฎีบัณฑิต สาขาการทดสอบและวัดผลการศึกษา ของมหาวิทยาลัย-ศรีนครินทรวิโรฒ

.....  คณบดีบัณฑิตวิทยาลัย
(ดร.ศิริยุภา พูลสุวรรณ)

วันที่...12... เดือน...มิถุนายน..... พ.ศ.2539.

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลงได้ด้วยความช่วยเหลือ แนะนำ ให้ข้อคิดเห็นและ
แก้ไขข้อบกพร่องต่าง ๆ อย่างดียิ่งจาก รองศาสตราจารย์ ดร.ส.วาสนา ประवालพฤษ
อาจารย์ ดร.ชูศักดิ์ ชัมภลิต และรองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี ผู้วิจัยรู้สึก
ซาบซึ้ง และขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้ ขอกราบขอบพระคุณ ดร.บุญลือ
ทองอยู่ รองเลขาธิการ ก.ค. ที่ได้กรุณาให้ข้อเสนอแนะเพิ่มเติม เพื่อให้ปริญญานิพนธ์ฉบับนี้
สมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ อาจารย์ชาญศักดิ์ เหลืองไตรรัตน์ ศูนย์สารสนเทศ สำนักงาน
ปลัดกระทรวง กระทรวงศึกษาธิการ ที่กรุณาเสนอข้อมูลผลการสอบแข่งขันเพื่อบรรจุบุคคล
เข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี 2537 วิชาภาษาไทย ของ อ.ก.ค.กรม
สามัญศึกษา เพื่อใช้ในการทำปริญญานิพนธ์ฉบับนี้

ขอขอบคุณ อาจารย์ปรีชา อรุณสวัสดิ์ และอาจารย์บุญเรือง ศรีเหรียญ ตลอดจน
พี่ ๆ เพื่อน ๆ น้อง ๆ ทั้งที่ทำงานและชาววัดพล ที่เป็นกำลังใจและให้ความช่วยเหลือด้าน
ต่าง ๆ ในการทำปริญญานิพนธ์ฉบับนี้

สุดท้ายนี้ ผู้วิจัยขอน้อมรำลึกถึงพระคุณบิดา มารดา ตลอดจนญาติพี่น้องที่ได้สนับสนุน
ผู้วิจัยตลอดมา และขอน้อมรำลึกถึงพระคุณของครู อาจารย์ ที่ให้การอบรมสั่งสอนและสนับสนุน
การศึกษาของผู้วิจัย ตั้งแต่อดีตถึงปัจจุบัน

รัตนา ศรีเหรียญ

สารบัญ

บทที่		หน้า
1	บทนำ	1
	ความเป็นมาและความสำคัญของปัญหา	1
	วัตถุประสงค์ในการวิจัย	5
	สมมุติฐานในการวิจัย	5
	ขอบเขตของการวิจัย	7
	ข้อตกลงเบื้องต้น	8
	คำนิยามศัพท์เฉพาะ	9
	ประโยชน์ที่คาดว่าจะได้รับ	11
2	เอกสารและงานวิจัยที่เกี่ยวข้อง	12
	แนวคิดของทฤษฎีการตอบข้อคำถาม	12
	ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม	13
	ลักษณะเด่นของทฤษฎีการตอบข้อคำถาม	15
	แบบจำลองที่ใช้ในทฤษฎีการตอบข้อคำถาม	16
	แบบจำลองปกติของออโจว์	17
	แบบจำลองโลจิสติก	20
	ดัชนีกำหนดคุณภาพและการเปรียบเทียบคุณภาพของแบบทดสอบ	23
	ค่าฟังก์ชันสารสนเทศของแบบทดสอบ	23
	ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ	25
	การเปรียบเทียบประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์	27
	วิธีประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก	
	3 พารามิเตอร์	28
	วิธีแมกซิมัม โลคัลลิสต์	29
	วิธีมาร์จินัล แมกซิมัม โลคัลลิสต์	39
	วิธีอีวาโลลิสต์	42

วิธีของเบส์	42
การศึกษาโดยใช้เทคนิควิมอนติ คาร์โล	53
ความเป็นมาของเทคนิควิมอนติ คาร์โล	53
ขั้นตอนของระเบียบเทคนิควิมอนติ คาร์โล	54
จุดเด่นของการใช้เทคนิควิมอนติ คาร์โล	59
งานวิจัยที่เกี่ยวข้อง	60
งานวิจัยที่เกี่ยวข้องในประเทศ	60
งานวิจัยที่เกี่ยวข้องต่างประเทศ	62
3 วิธีดำเนินการวิจัย	66
ประชากร	66
กลุ่มตัวอย่าง	66
เครื่องมือที่ใช้ในการวิจัย	66
สถิติที่ใช้ในการวิจัย	66
การวิเคราะห์ข้อมูล	68
4 ผลการวิเคราะห์ข้อมูล	81
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล	81
การเสนอผลการวิเคราะห์ข้อมูล	82
ผลการวิเคราะห์ข้อมูล	83
5 สรุป อภิปรายผล และข้อเสนอแนะ	512
วัตถุประสงค์ในการวิจัย	512
กลุ่มตัวอย่าง	512
เครื่องมือที่ใช้ในการวิจัย	512
การวิเคราะห์ข้อมูล	513
สรุปผลการวิเคราะห์ข้อมูล	515

บทที่

หน้า

อภิปรายผล	521
ข้อเสนอแนะ	524
บรรณานุกรม	526
ประวัติย่อของผู้วิจัย	531

บัญชีตาราง

ตาราง

หน้า

1	อนุพันธ์อันดับที่ 1 และอนุพันธ์อันดับที่ 2 สำหรับฟังก์ชันโลคัลลิสต์ ของค่าพารามิเตอร์ข้อสอบ (a,b,c) และความสามารถ ของผู้สอบ (e) ในแบบจำลองโลจิสติก 3 พารามิเตอร์	33
2	ค่าฟังก์ชันสารสนเทศสำหรับค่าพารามิเตอร์ข้อสอบ และ ค่าความสามารถของผู้เข้าสอบ ในแบบจำลองโลจิสติก 3 พารามิเตอร์	37
3	อนุพันธ์อันดับที่ 1 และอนุพันธ์อันดับที่ 2 ของ $\ln f(e,b,a,c u)$	51
4	สรุปความเหมือนและความแตกต่างของกระบวนการประมาณ ค่าพารามิเตอร์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวีเอสติก และวิธีของเบส์	52
5	ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.00 ของแบบทดสอบ	83
6	ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่า ด้วยวิธีแมกซิมัม โลคัลลิสต์ เมื่อขนาดของกลุ่มตัวอย่าง และจำนวนข้อสอบต่างกัน	86
7	ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่า ด้วยวิธีอีวีเอสติก เมื่อขนาดของกลุ่มตัวอย่างและจำนวน ข้อสอบต่างกัน	90
8	ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่า ด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวน ข้อสอบต่างกัน	94
9	ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณค่า ด้วยวิธีแมกซิมัม โลคัลลิสต์ เมื่อขนาดของกลุ่มตัวอย่าง และจำนวนข้อสอบต่างกัน	99

10	ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณ ค่าด้วยวิธีอีวริสติก เมื่อขนาดของกลุ่มตัวอย่างและ จำนวนข้อสอบต่างกัน	103
11	ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณ ค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและ จำนวนข้อสอบต่างกัน	108
12	ค่าสถิติพื้นฐานของค่าการเดาของข้อสอบที่ประมาณ ค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด เมื่อขนาดของกลุ่ม ตัวอย่าง และจำนวนข้อสอบต่างกัน	112
13	ค่าสถิติพื้นฐานของค่าการเดาของข้อสอบที่ประมาณ ค่าด้วยวิธีอีวริสติก เมื่อขนาดของกลุ่มตัวอย่าง และจำนวนข้อสอบต่างกัน	116
14	ค่าสถิติพื้นฐานของค่าการเดาของข้อสอบที่ประมาณ ค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง และจำนวนข้อสอบต่างกัน	120
15	ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณ ค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด เมื่อขนาดของกลุ่ม ตัวอย่าง และจำนวนข้อสอบต่างกัน	124
16	ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณ ค่าด้วยวิธีอีวริสติก เมื่อขนาดของกลุ่มตัวอย่างและ จำนวนข้อสอบต่างกัน	128
17	ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณ ค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและ จำนวนข้อสอบต่างกัน	132
18	เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน	136

19	เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่ประมาณค่าด้วยวิธีวิธีวิสถิติ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน	147
20	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีวิธีวิสถิติ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	148
21	เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน	149
22	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 60 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	150
23	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 70 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	151
24	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	152
25	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 90 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	153
26	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 100 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	154

27	เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวีรอสติก และ วิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน	155
28	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวีรอสติก และ วิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการ เปรียบเทียบพหุคูณ	159
29	เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาด ของกลุ่มตัวอย่างต่างกัน	163
30	เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีอีวีรอสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน	164
31	เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน	165
32	เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวีรอสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน	166
33	การทดสอบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวีรอสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	170
34	เปรียบเทียบความแตกต่างของค่าการเดาของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาด ของกลุ่มตัวอย่างต่างกัน	174

35	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 50 ข้อ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	175
36	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 60 ข้อ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	176
37	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 70 ข้อ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	177
38	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	178
39	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 90 ข้อ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	179
40	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 100 ข้อ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	180
41	เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบ ที่ประมาณค่าด้วยวิธีอีวริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน	173
42	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 50 ข้อ ที่ประมาณค่าด้วยวิธีอีวริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	174
43	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 60 ข้อ ที่ประมาณค่าด้วยวิธีอีวริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	175

44	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 70 ข้อ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	176
45	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	177
46	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 90 ข้อ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	178
47	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 100 ข้อ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	179
48	เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน	180
49	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 50 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	181
50	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 60 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	182
51	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 70 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	183
52	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	184

53	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 90 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน ด้วยการเปรียบเทียบพหุคูณ	185
54	เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม लाईलिस्ट วิธีอิวิริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน	186
55	การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม लाईलिस्ट วิธีอิวิริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ	190
56	เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม लाईलिस्ट เมื่อขนาดของ กลุ่มตัวอย่างต่างกัน	194
57	เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน	195
58	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน และข้อสอบมีจำนวน 50 ข้อ ด้วยการ เปรียบเทียบพหุคูณ	196
59	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน และข้อสอบมีจำนวน 60 ข้อ ด้วยการ เปรียบเทียบพหุคูณ	198

60	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน และข้อสอบมีจำนวน 70 ข้อ ด้วยการ เปรียบเทียบบุคคล	199
61	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน และข้อสอบมีจำนวน 80 ข้อ ด้วยการ เปรียบเทียบบุคคล	200
62	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน และข้อสอบมีจำนวน 100 ข้อ ด้วยการ เปรียบเทียบบุคคล	201
63	เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง ต่างกัน	202
64	เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอิวิริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน	203
65	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอิวิริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบบุคคล	207
66	เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์	211
67	เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีอิวิริสติก	212

68	เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีของเบส์	213
69	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่มีจำนวนต่างกัน ของกลุ่มตัวอย่างขนาด 2,000 คน ที่ประมาณค่าด้วยวิธีของเบส์ ด้วยการเปรียบเทียบพหุคูณ	214
70	เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์	215
71	การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ ด้วยการเปรียบเทียบพหุคูณ	219
72	เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อจำนวน ข้อสอบต่างกัน	224
73	เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีอีวีริสติก เมื่อจำนวนข้อสอบต่างกัน	225
74	เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน	226
75	เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน	227
76	การทดสอบความแตกต่างของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ด้วยการ เปรียบเทียบพหุคูณ	231

77	เปรียบเทียบความแตกต่างของค่าการเคาของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อจำนวน ข้อสอบต่างกัน	236
78	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อกลุ่มตัวอย่างมีจำนวน 200 คน ด้วยการเปรียบเทียบพหุคูณ	237
79	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อกลุ่มตัวอย่างมีจำนวน 600 คน ด้วยการเปรียบเทียบพหุคูณ	238
80	เปรียบเทียบความแตกต่างของค่าการเคาของข้อสอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน	239
81	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อกลุ่ม ตัวอย่างมีจำนวน 900 คน ด้วยการเปรียบเทียบพหุคูณ	240
82	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อกลุ่ม ตัวอย่างมีจำนวน 1,000 คน ด้วยการเปรียบเทียบพหุคูณ	241
83	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อกลุ่ม ตัวอย่างมีจำนวน 1,500 คน ด้วยการเปรียบเทียบพหุคูณ	242
84	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อกลุ่ม ตัวอย่างมีจำนวน 2,000 คน ด้วยการเปรียบเทียบพหุคูณ	243
85	เปรียบเทียบความแตกต่างของค่าการเคาของข้อสอบ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน	244

86	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อกลุ่ม ตัวอย่างมีจำนวน 2,000 คน ด้วยการเปรียบเทียบกับบุคคล	245
87	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อกลุ่ม ตัวอย่างมีจำนวน 2,500 คน ด้วยการเปรียบเทียบกับบุคคล	246
88	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่มีจำนวนต่างกัน ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อกลุ่ม ตัวอย่างมีจำนวน 3,000 คน ด้วยการเปรียบเทียบกับบุคคล	247
89	เปรียบเทียบความแตกต่างของค่าการเคาของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และ วิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน	248
90	การทดสอบความแตกต่างของค่าการเคาของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ด้วยการ เปรียบเทียบกับบุคคล	252
91	เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อจำนวนข้อสอบ ต่างกัน	257
92	เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอีวริสติก เมื่อจำนวนข้อสอบต่างกัน	258
93	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอีวริสติก เมื่อจำนวนข้อสอบต่างกันของ กลุ่มตัวอย่างจำนวน 200 คน ด้วยการเปรียบเทียบ บุคคล	259

94	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 300 คน ด้วยการเปรียบเทียบ พหุคูณ	260
95	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 400 คน ด้วยการเปรียบเทียบ พหุคูณ	261
96	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 500 คน ด้วยการเปรียบเทียบ พหุคูณ	262
97	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 600 คน ด้วยการเปรียบเทียบ พหุคูณ	263
98	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 700 คน ด้วยการเปรียบเทียบ พหุคูณ	264
99	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 800 คน ด้วยการเปรียบเทียบ พหุคูณ	265

100	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอีวีรอสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 900 คน ด้วยการเปรียบเทียบ พหุคูณ	266
101	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอีวีรอสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 1,000 คน ด้วยการเปรียบเทียบ พหุคูณ	267
102	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอีวีรอสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 1,500 คน ด้วยการเปรียบเทียบ พหุคูณ	268
103	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอีวีรอสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 2,000 คน ด้วยการเปรียบเทียบ พหุคูณ	269
104	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอีวีรอสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 2,500 คน ด้วยการเปรียบเทียบ พหุคูณ	270
105	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีอีวีรอสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 3,000 คน ด้วยการเปรียบเทียบ พหุคูณ	271
106	เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน	272

107	เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน	273
108	การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และ วิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ด้วยการเปรียบเทียบพหุคูณ	277

บัญชีภาพประกอบ

ภาพประกอบ

หน้า

1	เส้นโค้งลักษณะข้อสอบของแบบจำลองปกติของออไจว์	
	2 พารามิเตอร์	18
2	เส้นโค้งลักษณะข้อสอบของแบบจำลองปกติของออไจว์	
	3 พารามิเตอร์	19
3	เส้นโค้งลักษณะข้อสอบของแบบจำลองโลจิสติก 1 พารามิเตอร์	21
4	เส้นโค้งลักษณะข้อสอบของแบบจำลองโลจิสติก 2 พารามิเตอร์	22
5	เส้นโค้งลักษณะข้อสอบของแบบจำลองโลจิสติก 3 พารามิเตอร์	22
6	ค่าโอเกนขององค์ประกอบที่มีค่าเกิน 1.00 ของแบบทดสอบ	84
7	เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ	281
8	เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ	282
9	เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ	283
10	เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ	284

- 11 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลค์ลิฮูด
เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 285
- 12 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลค์ลิฮูด
เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 286
- 13 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อข้อสอบ
มีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 287
- 14 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อข้อสอบ
มีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 290
- 15 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อข้อสอบ
มีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 291
- 16 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อข้อสอบ
มีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 292

- 17 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวีรีสติก เมื่อข้อสอบ
มีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 293
- 18 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวีรีสติก เมื่อข้อสอบ
มีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 294
- 19 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบ
มีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 297
- 20 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบ
มีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 298
- 21 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบ
มีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 299
- 22 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบ
มีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับ
ความสามารถต่าง ๆ 300

- 23 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ 301
- 24 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ 302
- 25 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ 305
- 26 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 300 คน และข้อสอบมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ 306
- 27 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ 307
- 28 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 500 คน และข้อสอบมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ 308

- 29 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 600 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ 309
- 30 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ 310
- 31 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ 311
- 32 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ 312
- 33 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ 313
- 34 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ 314

- 35 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิสต์
เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด
ต่างกัน ๗ ระดับความสามารถต่าง ๆ 315
- 36 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิสต์
เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน และข้อสอบมีขนาด
ต่างกัน ๗ ระดับความสามารถต่าง ๆ 316
- 37 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิสต์
เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาด
ต่างกัน ๗ ระดับความสามารถต่าง ๆ 317
- 38 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวีรติค เมื่อกลุ่ม
ตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาดต่างกัน
๗ ระดับความสามารถต่าง ๆ 321
- 39 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวีรติค เมื่อกลุ่ม
ตัวอย่างมีขนาด 300 คน และข้อสอบมีขนาดต่างกัน
๗ ระดับความสามารถต่าง ๆ 322
- 40 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีอีวีรติค เมื่อกลุ่ม
ตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาดต่างกัน
๗ ระดับความสามารถต่าง ๆ 323

- 41 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อกลุ่มตัวอย่างมีขนาด 500 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 324
- 42 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อกลุ่มตัวอย่างมีขนาด 600 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 325
- 43 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 326
- 44 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 327
- 45 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อกลุ่มตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 328
- 46 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 329

- 47 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรสถิต เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 330
- 48 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรสถิต เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 331
- 49 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรสถิต เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 332
- 50 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรสถิต เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 333
- 51 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 337
- 52 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 300 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 338

- 53 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่ม
ตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 339
- 54 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่ม
ตัวอย่างมีขนาด 500 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 340
- 55 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่ม
ตัวอย่างมีขนาด 600 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 341
- 56 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่ม
ตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 342
- 57 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่ม
ตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 343
- 58 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจาก
การประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่ม
ตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 344

- 59 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 345
- 60 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 346
- 61 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 347
- 62 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 348
- 63 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 349
- 64 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซ์ิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 353

- 65 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิตูด
เมื่อข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 359
- 66 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิตูด
เมื่อข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 365
- 67 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิตูด
เมื่อข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 371
- 68 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิตูด
เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 377
- 69 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิตูด
เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 383
- 70 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อ
ข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 391

- 71 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรีสติก เมื่อ
ข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 397
- 72 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรีสติก เมื่อ
ข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 403
- 73 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรีสติก เมื่อ
ข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 409
- 74 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรีสติก เมื่อ
ข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 415
- 75 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีรีสติก เมื่อ
ข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 421
- 76 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อ
ข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 429

- 77 เปรียบเทียบปัจจัยประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อ
ข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 435
- 78 เปรียบเทียบปัจจัยประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อ
ข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 441
- 79 เปรียบเทียบปัจจัยประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อ
ข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 447
- 80 เปรียบเทียบปัจจัยประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อ
ข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 453
- 81 เปรียบเทียบปัจจัยประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อ
ข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ 459
- 82 เปรียบเทียบปัจจัยประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด
เมื่อกลุ่มตัวอย่างมีขนาด 200 คน ณ ระดับความสามารถต่าง ๆ 467
- 83 เปรียบเทียบปัจจัยประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
มาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด
เมื่อกลุ่มตัวอย่างมีขนาด 300 คน ณ ระดับความสามารถต่าง ๆ 468

[illegible]

[illegible]

[illegible]

[illegible]

- 120 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผล
 มาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อ
 กลุ่มตัวอย่างมีขนาด 3,000 คน ณ ระดับความสามารถต่าง ๆ 509

ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันแม้จะมีการพัฒนาทฤษฎีการทดสอบแบบดั้งเดิมและนำไปใช้อย่างแพร่หลายก็ตาม แต่ก็มียารงานการวิจัยและบทความที่กล่าวถึงทฤษฎีการทดสอบแบบดั้งเดิมว่าวัดได้ไม่ตรงกับความสามารถที่แท้จริง ผลที่ได้จึงมีความคลาดเคลื่อนมาก นอกจากนี้ ทฤษฎีการทดสอบแบบดั้งเดิมยังถูกตั้งข้อสงสัยเกี่ยวกับความน่าเชื่อถือของผลการวัด (Hambleton and Cook. 1977 : 14 - 15) ประการแรก คือ ค่าความยาก และค่าอำนาจจำแนกของข้อสอบจะแปรเปลี่ยนไปตามกลุ่มผู้สอบ ประการที่สอง การเปรียบเทียบความสามารถระหว่างบุคคลจะเปรียบเทียบกันได้ก็ต่อเมื่อใช้แบบทดสอบฉบับเดียวกันหรือแบบทดสอบคู่ขนาน ซึ่งการสร้างแบบทดสอบคู่ขนานเป็นสิ่งที่เป็นไปได้ยากในทางปฏิบัติ ประการที่สาม ความคลาดเคลื่อนมาตรฐานในการวัดของผู้สอบแต่ละคนซึ่งเดิมกำหนดว่าเท่ากัน แต่ลอร์ด และโนวิก (Lord and Novick. 1968 : 14) ได้แสดงให้เห็นว่า ความคลาดเคลื่อนมาตรฐานในการวัดของผู้สอบแต่ละคนนั้นไม่เท่ากัน ซึ่งจุดอ่อนเหล่านี้ทฤษฎีการทดสอบแบบดั้งเดิมไม่สามารถเสนอแนวทางในการแก้ไขจุดอ่อนเหล่านี้ได้อย่างชัดเจน

ดังนั้น นักวัดผลจึงได้พัฒนาทฤษฎีการทดสอบอื่น ๆ ที่เหมาะสมกว่ามาใช้ในการวัดและแก้ปัญหาจุดอ่อนดังกล่าว ซึ่งทฤษฎีใหม่ที่เกิดขึ้นมาและกำลังเป็นที่สนใจของนักวัดผลมากในขณะนี้ คือ ทฤษฎีคุณลักษณะภายใน (Latent Trait Theory) หรือทฤษฎีการตอบข้อคำถาม (Item Response Theory) หรือโค้งลักษณะข้อสอบ (Item Characteristic Curve) (สงวน ลักษณะ. 2525 : 47) โดยเชื่อกันว่าถ้านาทฤษฎีนี้มาใช้ให้ตรงกับข้อตกลงเบื้องต้นแล้ว จะสามารถแก้ปัญหาที่เกิดจากทฤษฎีการทดสอบแบบดั้งเดิมได้ (Hambleton and Cook. 1977 : 16 - 32) เนื่องจากทฤษฎีการตอบข้อคำถามจะอธิบายความสามารถที่แท้จริงของบุคคลและคุณสมบัติของข้อสอบ โดยอาศัยแบบจำลองหรือฟังก์ชันทางคณิตศาสตร์ซึ่งมีค่าพารามิเตอร์ความยากของข้อสอบกับความสามารถของบุคคล เป็นตัวกำหนดความน่าจะเป็นที่บุคคลจะตอบข้อสอบนั้นได้ถูก

อย่างไรก็ตาม ความเชื่อเกี่ยวกับความสัมพันธ์ระหว่างระดับความสามารถกับโอกาสที่จะตอบข้อสอบถูกมีหลายความเชื่อ เป็นผลให้มีแบบจำลอง (Model) ที่ใช้อธิบายความสัมพันธ์ดังกล่าวในทฤษฎีการตอบข้อคำถามหลายแบบด้วยกัน เริ่มตั้งแต่ปี ค.ศ. 1952

ลอร์ด (Lord) ได้คิดแบบจำลองนอร์มอลอวจ์ (Normal Ogive Model) ขึ้น ต่อมา ในปี ค.ศ. 1960 ราสช์ (Rasch) ได้คิดแบบจำลองขึ้น เรียกว่า แบบจำลองของราสช์ (Rasch Model) ซึ่งเป็นแบบจำลองที่ได้รับความนิยมมากในปัจจุบัน และเป็นแบบจำลองที่ง่ายที่สุดของทฤษฎีการตอบข้อคำถาม จากนั้นในปี ค.ศ. 1969 เบิร์นบอร์ม (Birnbaum) ได้ปรับแบบจำลองนอร์มอลอวจ์ เป็นแบบจำลองโลจิสติก (Logistic Model) ซึ่งแบบจำลองนี้แบ่งออกเป็น 3 แบบ คือ แบบจำลองโลจิสติก 1 พารามิเตอร์ (One-Parameter Logistic Model) แบบจำลองโลจิสติก 2 พารามิเตอร์ (Two-Parameter Logistic Model) และแบบจำลองโลจิสติก 3 พารามิเตอร์ (Three-Parameter Logistic Model) (Hambleton and Cook, 1977 : 80 - 82) และเนื่องจากแบบจำลองโลจิสติกนี้ได้รับการพัฒนาดัง 3 รูปแบบ จึงทำให้สามารถนำไปใช้กับสถานการณ์ต่าง ๆ ได้ตามความเหมาะสม โดยการวิเคราะห์ข้อสอบที่ต้องการความละเอียด แบบจำลองโลจิสติก 3 พารามิเตอร์ จะเป็นแบบจำลองที่สามารถอธิบายโอกาสที่ผู้เข้าสอบสามารถตอบข้อสอบได้ถูกต้องและครอบคลุมกว่า (Hambleton and Cook, 1977 : 16 - 32)

ในการนำทฤษฎีการตอบข้อคำถามมาใช้ นอกจากผู้ใช้จำเป็นต้องเลือกแบบจำลองที่เหมาะสมแล้ว การเลือกวิธีการประมาณค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้เข้าสอบให้เหมาะสมกับสภาพการวัดแต่ละครั้ง ก็เป็นสิ่งที่มีความจำเป็นเช่นกัน เพราะแต่ละวิธีมีข้อจำกัดในการประมาณค่าแตกต่างกัน สำหรับแบบจำลองโลจิสติกนั้นวิธีการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้เข้าสอบ 3 วิธีใหญ่ ๆ (Swaminathan and Gifford, 1986 : 589 - 601) คือ วิธีแมกซิมัม ไลค์ลิฮูด (Maximum Likelihood Estimation) ที่อธิบายโดยลอร์ด วิธีฮิวริสติก (Heuristic Estimation) ซึ่งมีฮูรี (Urry) เป็นผู้สนับสนุน และวิธีของเบส์ (Bayesian Estimation) ที่อธิบายโดยสวามินาธาน และกัฟฟอร์ด โดยวิธีประมาณค่าทั้งสามวิธีนี้มีแนวคิดพื้นฐานที่แตกต่างกัน สามารถแยกออกได้เป็น 2 แนว (สุชาดา กิระนันท์, 2525 : 26 - 27) แนวแรกเรียกว่า สถิตินอนเบส์เชียน (Non Bayesian Statistics) แนวคิดนี้จะพิจารณาค่าพารามิเตอร์ว่าเป็นค่าคงที่ ที่ไม่ทราบค่าและต้องการประมาณ วิธีการประมาณค่าพารามิเตอร์ที่อาศัยแนวคิดนี้ ได้แก่ วิธีแมกซิมัม ไลค์ลิฮูด และวิธีฮิวริสติก ส่วนอีกแนวหนึ่งเรียกว่า สถิติแบบเบส์เชียน (Bayesian Statistics) แนวคิดนี้จะพิจารณาว่า ค่าพารามิเตอร์เป็นตัวแปรสุ่มที่มีค่าที่เป็นไปได้อยู่ในช่วงหนึ่งหรือเซตหนึ่ง และมีการแจกแจงความน่าจะเป็นของตัวเอง ซึ่งจะเป็นสิ่งที่บอกว่าค่าพารามิเตอร์เหล่านั้นมีโอกาสเกิดขึ้นมากน้อยโดยการ

เปรียบเทียบเช่นไร การแจกแจงความน่าจะเป็นของค่าพารามิเตอร์นี้ เรียกว่า การแจกแจงเริ่มแรก (Prior Distribution) อันเป็นแนวคิดพื้นฐานของการประมาณค่าโดยวิธีของเบส์ นอกจากความแตกต่างในแนวคิดพื้นฐานแล้ว วิธีประมาณค่าทั้ง 3 วิธี ยังมีความแตกต่างกันในเรื่องเวลาที่ใช้ในการคำนวณ ค่าใช้จ่ายในการคำนวณ ความสะดวกในการคำนวณ ขนาดของกลุ่มผู้เข้าสอบ และจำนวนข้อสอบที่เหมาะสมในแต่ละวิธีอีกด้วย ซึ่งจะส่งผลให้วิธีการประมาณค่าพารามิเตอร์แต่ละวิธีมีจุดเด่น จุดด้อย หรือข้อจำกัดที่ต่างกัน และค่าพารามิเตอร์ที่ประมาณได้จากแต่ละวิธีก็จะมีค่าแตกต่างกันด้วย (วิชดา บัวคง. 2533 : 4) อย่างไรก็ตาม แม้วิธีประมาณค่าพารามิเตอร์จะมีความแตกต่างกันในหลาย ๆ ด้าน แต่ผู้พัฒนาวิธีการประมาณค่าแต่ละวิธีต่างก็ต้องการประมาณค่าความสามารถของบุคคลให้มีความแม่นยำมากที่สุด และดัชนีที่ความแม่นยำของการประมาณค่าความสามารถนี้เป็นแนวคิดที่สำคัญของทฤษฎีการตอบข้อคำถาม ที่หาได้จากฟังก์ชันสารสนเทศ (Information Function) ของข้อสอบและแบบทดสอบ นั้นเอง (Birnbbaum. 1968 : 429)

ค่าฟังก์ชันสารสนเทศนี้จะมีค่ามากหรือน้อยขึ้นอยู่กับค่าพารามิเตอร์ของข้อสอบ ซึ่งเป็นตัวกำหนดลักษณะของข้อสอบ และค่าฟังก์ชันสารสนเทศนี้จะมีค่าแตกต่างกัน ถ้าเราใช้ผลจากการวัดด้วยข้อสอบข้อเดียวกัน หรือแบบทดสอบฉบับเดียวกันประมาณค่าความสามารถในระดับที่แตกต่างกัน และ ณ ระดับความสามารถใดที่ค่าฟังก์ชันสารสนเทศที่ได้มีค่าสูง แสดงว่าช่วงการประมาณค่าจะแคบ ซึ่งหมายความว่า ความแม่นยำในการประมาณค่าความสามารถที่ระดับนั้นจะมีมาก ดังนั้น ถ้ามีแบบทดสอบหลายฉบับที่วัดความสามารถเดียวกันแล้ว จะสามารถเปรียบเทียบคุณภาพหรือประสิทธิภาพของแบบทดสอบเหล่านั้นได้ โดยการนำเอาค่าฟังก์ชันสารสนเทศ ณ ระดับความสามารถเดียวกันของแบบทดสอบที่ต้องการเปรียบเทียบมาหาค่าอัตราส่วนกัน เรียกว่า ค่าดัชนีประสิทธิภาพสัมพัทธ์ (Relative Efficiency) ของแบบทดสอบฉบับหนึ่งเทียบกับแบบทดสอบอีกฉบับหนึ่ง ณ ระดับความสามารถใด ๆ ซึ่งถ้าค่าดัชนีประสิทธิภาพสัมพัทธ์ดังกล่าวมีค่าเท่ากับ 1 แสดงว่า แบบทดสอบทั้งสองฉบับมีประสิทธิภาพในการประมาณค่าเท่ากัน วิธีการเปรียบเทียบประสิทธิภาพของแบบทดสอบนี้ใช้ได้กับการเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ต่าง ๆ ในแบบทดสอบแต่ละฉบับด้วย ดังนั้น การเปรียบเทียบคุณภาพของแบบทดสอบหลาย ๆ ฉบับ หรือการเปรียบเทียบคุณภาพของแบบทดสอบฉบับหนึ่งที่มีวิธีการประมาณค่าพารามิเตอร์หลาย ๆ วิธี โดยนำค่าฟังก์ชันสารสนเทศเป็นตัวกำหนดคุณภาพนั้น จึงเป็นการเปรียบเทียบคุณภาพโดยสรุปจากทุกแง่มุม เนื่องจากค่าฟังก์ชันสารสนเทศเป็นฟังก์ชันของค่าพารามิเตอร์ของข้อสอบที่มีลักษณะเป็น

ดัชนีผสม (Composite Index) ที่สร้างขึ้นจากดัชนีที่กำหนดคุณภาพของข้อสอบหลาย ๆ ดัชนีด้วยกัน ทำให้การเปรียบเทียบคุณภาพของแบบทดสอบหรือข้อสอบ ไม่จำเป็นต้องเปรียบเทียบคุณลักษณะที่จะประเด็นอีกต่อไป (วิชดา บัวคง. 2533 : 4 - 5)

จากที่กล่าวมาแล้ว จะเห็นได้ว่าการเลือกวิธีประมาณค่าพารามิเตอร์ที่มีประสิทธิภาพและเหมาะสมกับขนาดของกลุ่มผู้เข้าสอบหรือจำนวนข้อสอบจึงเป็นขั้นตอนที่จำเป็นสิ่ง ทั้งนี้เพื่อให้การนำทฤษฎีการตอบข้อคำถามไปใช้เป็นไปอย่างสมบูรณ์ตามความมุ่งหมาย เนื่องจากวิธีการประมาณค่าแต่ละวิธีมีความแตกต่างกันในหลาย ๆ ด้าน โดยเฉพาะอย่างยิ่งความแตกต่างในเรื่องแนวคิดทางสถิติที่ทำให้ค่าพารามิเตอร์ที่ประมาณได้จากแต่ละวิธีแตกต่างกัน เป็นผลให้ค่าฟังก์ชันสารสนเทศที่เป็นดัชนีผสมของค่าพารามิเตอร์แตกต่างกันด้วย แต่จากการศึกษางานวิจัยภายในประเทศเกี่ยวกับการนำทฤษฎีการตอบข้อคำถาม ในแบบจำลองโลจิสติก 3 พารามิเตอร์ไปประยุกต์ใช้ในการวิเคราะห์หาคุณภาพของข้อสอบ พบว่า ใช้การประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบด้วยวิธีแมกซิมัม ไลค์ลิสต์ เป็นส่วนใหญ่ ทั้ง ๆ ที่วิธีแมกซิมัม ไลค์ลิสต์มีข้อจำกัด คือ ไม่สามารถประมาณค่าความสามารถของผู้เข้าสอบที่ได้คะแนนเต็มหรือคะแนนศูนย์ได้ จึงจำเป็นต้องตัดผู้สอบเหล่านี้ออกไปในการประมาณค่า ทำให้เหลือผู้เข้าสอบที่มีระดับความสามารถใกล้เคียงกัน ดังนั้น ถ้าต้องการเปรียบเทียบความสามารถของผู้สอบเมื่อทำการทดสอบด้วยแบบทดสอบ 2 ฉบับใด ๆ แล้ว จะมีแนวโน้มว่าความสามารถของผู้สอบเฉลี่ยแล้วมีค่าไม่แตกต่างกัน จึงอาจสรุปได้ว่าแบบทดสอบ 2 ฉบับนั้น มีคุณภาพไม่แตกต่างกัน ซึ่งข้อสรุปดังกล่าวอาจเป็นข้อสรุปที่คลาดเคลื่อน อันเนื่องมาจากข้อจำกัดของวิธีการประมาณค่าแบบแมกซิมัม ไลค์ลิสต์ ก็ได้ (เสรี เปรมชาติโรจน์. 2531 : 90 - 91) ซึ่งสวามินาธาน และกิฟฟอร์ด (Swaminathan and Gifford. 1986 : 589 - 601) กล่าวว่า การประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ไม่มีข้อจำกัดเหมือนวิธีแมกซิมัม ไลค์ลิสต์ กล่าวคือ สามารถประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้เข้าสอบที่ทำข้อสอบถูกหรือผิดทุกข้อได้ โดยเฉพาะในกลุ่มตัวอย่างที่มีขนาดเล็กประมาณ 200 - 300 คน วิธีของเบส์สามารถประมาณค่าพารามิเตอร์ของข้อสอบได้ใกล้เคียงกว่าวิธีแมกซิมัม ไลค์ลิสต์ แต่อย่างไรก็ตามทั้งวิธีแมกซิมัม ไลค์ลิสต์ และวิธีของเบส์ต่างต้องอาศัยเครื่องคอมพิวเตอร์ขนาดใหญ่ในการคำนวณ จึงมีข้อจำกัดที่จะทำได้เฉพาะในสถานที่ซึ่งมีความพร้อมทั้งด้านบุคลากรและเครื่องคอมพิวเตอร์ แต่การประมาณค่าด้วยวิธีอิวิริสติก ซึ่งเป็นวิธีที่ลอร์ดเป็นผู้ริเริ่มไว้และมีอิริเป็นผู้สนับสนุนนั้น สามารถทำการประมาณค่าพารามิเตอร์ของข้อสอบได้ จากค่าสถิติที่ได้จากทฤษฎีการทดสอบแบบดั้งเดิม และสามารถ

ใช้เครื่องคำนวณขนาดเล็กได้ แม้ว่าการประมาณค่าด้วยวิธีนี้จะมีความคลาดเคลื่อนอยู่บ้าง แต่ชมิคท์ (Schmidt. 1977 : 613 - 620) ได้ทำการปรับแก้สูตรการประมาณค่า และยืนยันว่าวิธีอีวริสติกสามารถนำไปใช้ได้ดีเท่ากับวิธีแมกซิมัม โคล์ลิสต์ โดยเฉพาะในกลุ่มผู้เข้าสอบที่มีขนาดใหญ่

จากความแตกต่างของวิธีการประมาณค่าพารามิเตอร์ทั้ง 3 วิธี ทำให้ผู้วิจัยมีความสนใจที่จะเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซิมัม โคล์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยเทคนิควิธีมอนติ คาร์โล เพื่อเป็นการตรวจสอบและขยายความรู้ในเชิงทฤษฎีเกี่ยวกับวิธีการประมาณค่าพารามิเตอร์ว่าวิธีใดจะมีประสิทธิภาพมากที่สุด

วัตถุประสงค์ในการวิจัย

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซิมัม โคล์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยเทคนิควิธีมอนติ คาร์โล โดยมีวัตถุประสงค์เฉพาะของการวิจัย ดังนี้

1. เพื่อศึกษาค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน
2. เพื่อศึกษาค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน
3. เพื่อเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลจากการประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัม โคล์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

สมมติฐานในการวิจัย

สวามินาธาน และกิฟฟอร์ด (Swaminathan and Gifford. 1986 : 589 - 601) กล่าวว่า วิธีการประมาณค่าที่ใช้กันอยู่ในปัจจุบันในแบบจำลองโลจิสติก 3 พารามิเตอร์ คือ วิธีการแมกซิมัม โคล์ลิสต์ และวิธีอีวริสติก ยังมีข้อบกพร่อง เพราะไม่สามารถประมาณค่าอำนาจจำแนก และค่าการเดา ได้ดีพอ ถ้าไม่มีการกำหนดขอบเขตจำกัดของช่วงค่าอำนาจ

จำแนก และค่าการเคา ซึ่งข้อจำกัดเหล่านี้ส่งผลให้วิธีแมกซิมัมไลค์ลิฮูด และวิธีอีวีริสติก ไม่สามารถประมาณค่าความสามารถของผู้เข้าสอบที่ได้คะแนนเต็มหรือคะแนนศูนย์ได้ รวมทั้งไม่สามารถประมาณค่าพารามิเตอร์ของข้อสอบที่ทุกคนตอบถูกหมดหรือผิดหมดได้ ในขณะที่การประมาณค่าด้วยวิธีของเบส์นั้นสามารถแก้ปัญหาดังกล่าวได้ โดยการกำหนดการแจกแจงเริ่มแรกของค่าอำนาจจำแนก และค่าการเคา ทำให้การประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของบุคคลมีความถูกต้องมากยิ่งขึ้น ยิ่งกว่านั้นการประมาณค่าด้วยวิธีของเบส์ยังช่วยลดความซับซ้อนของกระบวนการประมาณค่าความสามารถของบุคคล เมื่อกลุ่มตัวอย่างผู้เข้าสอบและจำนวนข้อสอบมีขนาดเล็ก (สุชาดา กิระนันท์, 2525 : 100) และลอร์ด ได้ศึกษาวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ด้วยวิธีแมกซิมัมไลค์ลิฮูด ได้ข้อสรุปว่าการประมาณค่าพารามิเตอร์ในแบบจำลองโลจิสติก 3 พารามิเตอร์ ให้ความถูกต้องเพียงพอ นั้น ควรจะมีข้อสอบประมาณ 50 ข้อ และผู้เข้าสอบอย่างน้อย 1,000 คน (Lord, 1968 : 989 - 1020) ส่วนการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีริสติกนั้นแบบทดสอบควรประกอบด้วยข้อสอบอย่างน้อย 80 ข้อ (Urry, 1974 : 253 - 260)

นอกจากนี้ การประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัมไลค์ลิฮูด และวิธีของเบส์ จะทำการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้เข้าสอบด้วยกระบวนการหาค่าซ้ำ ๆ จนกว่าจะได้ค่าคงที่ (Convergence) แต่วิธีอีวีริสติกจะประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีการประมาณค่าอย่างง่าย ๆ โดยประมาณจากค่าความยาก และค่าอำนาจจำแนก ที่ได้จากการวิเคราะห์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม จึงทำให้ค่าพารามิเตอร์ของข้อสอบที่ได้ อาจมีความคลาดเคลื่อน (Schmidt, 1977 : 613 - 620) สำหรับการประมาณค่าความสามารถของบุคคลด้วยวิธีอีวีริสติกจะใช้วิธีแมกซิมัมไลค์ลิฮูด แบบมีเงื่อนไข (Hambleton and Swaminathan, 1985 : 81 - 88) ด้วยการหาค่าซ้ำ ๆ จนกว่าค่าความสามารถที่ประมาณจะลู่เข้าสู่ค่าคงที่เช่นกัน ดังนั้น อาจกล่าวได้ว่าวิธีประมาณค่าพารามิเตอร์ของข้อสอบทั้ง 3 วิธี น่าจะมีความคลาดเคลื่อนแตกต่างกัน กล่าวคือ วิธีอีวีริสติก น่าจะมีความคลาดเคลื่อนสูงที่สุด รองลงมา คือ วิธีแมกซิมัมไลค์ลิฮูด และวิธีของเบส์ ตามลำดับ สำหรับวิธีการประมาณค่าพารามิเตอร์ความสามารถของผู้เข้าสอบนั้น วิธีแมกซิมัมไลค์ลิฮูด และวิธีอีวีริสติก นั้น น่าจะได้ผลไม่แตกต่างกัน เพราะอาศัยวิธีแมกซิมัมไลค์ลิฮูด เหมือนกัน แต่วิธีของเบส์น่าจะประมาณค่าพารามิเตอร์ความสามารถของผู้เข้าสอบได้ดีกว่าวิธีอื่น เพราะมีการกำหนดการแจกแจงเริ่มแรกให้แก่ค่าอำนาจจำแนก และค่าการเคา

ด้วยเหตุที่แนวคิดและวิธีการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของ
ผู้เข้าสอบด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ มีความแตกต่างกัน น่าจะ
ทำให้ค่าพารามิเตอร์ที่ได้จากการประมาณค่าทั้ง 3 วิธี แตกต่างกันด้วย ดังนั้น ผู้วิจัยจึงได้ตั้ง
สมมติฐานในการวิจัยไว้ดังนี้

1. ค่าพารามิเตอร์ของข้อสอบที่วิเคราะห์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก
และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีค่าแตกต่างกัน
2. ค่าพารามิเตอร์ความสามารถของผู้สอบที่วิเคราะห์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธี
อีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีค่าแตกต่างกัน
3. ค่าพารามิเตอร์ของข้อสอบที่วิเคราะห์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก
และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน มีค่าแตกต่างกัน
4. ค่าพารามิเตอร์ความสามารถของผู้สอบที่วิเคราะห์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธี
อีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน มีค่าแตกต่างกัน
5. ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลจากการประมาณค่าพารามิเตอร์
ของข้อสอบด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่าง
ต่างกัน จะมีค่าไม่เท่ากับ 1
6. ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลจากการประมาณค่าพารามิเตอร์
ของข้อสอบด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน
จะมีค่าไม่เท่ากับ 1

ขอบเขตของการวิจัย

1. ประชากรที่ใช้ในการศึกษาครั้งนี้ เป็นคะแนนผลการสอบแข่งขันเพื่อบรรจุบุคคล
เข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี พ.ศ. 2537 วิชาภาษาไทย ของ อ.ก.ค.กรม
สามัญศึกษา ซึ่งมีผู้เข้าสอบแข่งขันจำนวน 15,258 คน
2. กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นคะแนนผลการสอบแข่งขันเพื่อบรรจุ
บุคคลเข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี พ.ศ. 2537 วิชาภาษาไทย ของ อ.ก.ค.
กรมสามัญศึกษา ซึ่งสุ่มออกเป็น 13 ขนาดกลุ่ม คือ ขนาดกลุ่ม 200 คน 300 คน 400 คน
500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน 1,500 คน 2,000 คน
2,500 คน และ 3,000 คน โดยกลุ่มตัวอย่างแต่ละกลุ่มตอบข้อสอบที่มีจำนวนต่างกัน 6 ขนาด
คือ ขนาดข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ 90 ข้อ และ 100 ข้อ ด้วยการทดลอง

เข้ากลุ่มละ 40 ครั้ง ด้วยเทคนิควิมมอนติ คาร์โล

3. ตัวแปรที่ใช้ในการศึกษาค้นคว้า

3.1 ตัวแปรอิสระ ได้แก่

3.1.1 วิธีการวิเคราะห์ข้อสอบ 3 วิธี คือ วิธีแมกซีมัม โคลลิฮูด วิธีอีวริสติก และวิธีของเบส์

3.1.2 ขนาดของกลุ่มตัวอย่าง 13 ขนาด คือ ขนาดกลุ่ม 200 คน 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน

3.1.3 ขนาดของข้อสอบ 6 ขนาด คือ ขนาดข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ 90 ข้อ และ 100 ข้อ

3.2 ตัวแปรตาม ได้แก่

3.2.1 ค่าพารามิเตอร์ของข้อสอบ

3.2.2 ค่าพารามิเตอร์ความสามารถของผู้สอบ

3.2.3 ค่าประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์

4. เครื่องมือที่ใช้ในการศึกษาค้นคว้านี้เป็นแบบทดสอบวิชาภาษาไทย แบบปรนัย ชนิด 4 ตัวเลือก จำนวน 100 ข้อ ซึ่ง อ.ก.ค.กรมสามัญศึกษา เป็นผู้สร้างตามหลักสูตรของ ก.ค. ที่ใช้ในการสอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู

ข้อตกลงเบื้องต้น

1. การวิจัยครั้งนี้ มุ่งศึกษาเฉพาะการประมาณค่าพารามิเตอร์ของข้อสอบและค่าพารามิเตอร์ความสามารถของผู้สอบ โดยใช้แบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซีมัม โคลลิฮูด วิธีอีวริสติก และวิธีของเบส์ เท่านั้น ไม่ได้ศึกษาเกี่ยวกับรายละเอียดของเนื้อหาวิชาภาษาไทย

2. การประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ จะศึกษาเฉพาะกรณีที่มีการกำหนดการแจกแจงเริ่มแรกของค่าอำนาจจำแนก (a) และค่าการเดา (c) เท่านั้น

3. ระดับความสามารถของผู้สอบ แบ่งออกเป็น 3 กลุ่ม คือ

3.1 $\theta > 1$ SD. หมายถึง กลุ่มตัวอย่างที่มีระดับความสามารถสูง

3.2 $\theta \pm 1$ SD. หมายถึง กลุ่มตัวอย่างที่มีระดับความสามารถปานกลาง

3.3 $\theta < -1$ SD. หมายถึง กลุ่มตัวอย่างที่มีระดับความสามารถต่ำ

4. การสอบในสถานที่และผู้ดำเนินการสอบต่างกัน ไม่มีผลต่อการตอบข้อคำถามของผู้สอบ

5. คะแนนที่ได้จากการสอบ เป็นคะแนนที่ผู้สอบทุกคนตั้งใจทำแบบทดสอบอย่างเต็มความสามารถ

คำนิยามศัพท์เฉพาะ

1. แบบทดสอบ หมายถึง แบบทดสอบวิชาภาษาไทยที่ใช้ในการสอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี พ.ศ. 2537 ของ อ.ก.ค.กรมสามัญศึกษา ซึ่งออกตามหลักสูตรที่ใช้ในการสอบแข่งขันที่ ก.ค. เป็นผู้กำหนด มีลักษณะเป็นแบบปรนัย ชนิด 4 ตัวเลือก จำนวน 100 ข้อ

2. ค่าฟังก์ชันสารสนเทศของแบบทดสอบ หมายถึง ความแม่นยำในการประมาณค่าความสามารถจริงด้วยคะแนนจากการตอบแบบทดสอบ ซึ่งมีค่าเท่ากับผลรวมอัตราส่วนกำลังสองของความถี่ของโด่งลักษณะข้อสอบ ต่อค่าความแปรปรวนของข้อสอบข้อนั้น ๗ ระดับความสามารถ ๑ ของข้อสอบทุก ๆ ข้อ ในแบบทดสอบ

3. ค่าประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์ หมายถึง อัตราส่วนระหว่างค่าฟังก์ชันสารสนเทศของแบบทดสอบภาษาไทย ที่เป็นผลมาจากวิธีประมาณค่าพารามิเตอร์วิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีเอสดี และวิธีของเบส ๗ ระดับความสามารถเดียวกัน ซึ่งก็คือค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ นั่นเอง

4. วิธีแมกซิมัม ไลค์ลิฮูด หมายถึง การประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้เข้าสอบ ในแบบจำลองโลจิสติก 3 พารามิเตอร์ ด้วยวิธีการประมาณค่าพารามิเตอร์ วิธีมาร์จินัล แมกซิมัม ไลค์ลิฮูด (Marginal Maximum Likelihood Procedure) โดยวิธีการนี้จะประมาณค่าพารามิเตอร์ของข้อสอบก่อน แล้วจึงทำการประมาณค่าความสามารถของผู้สอบ ด้วยโปรแกรม BILOG 3

5. วิธีอีวีเอสดี หรือ วิธีของฮูริ หมายถึง วิธีประมาณค่าพารามิเตอร์ของข้อสอบด้วยแบบจำลองโลจิสติก 3 พารามิเตอร์ โดยใช้ค่าประมาณจากค่าสถิติในทฤษฎีการทดสอบแบบดั้งเดิม และประมาณค่าความสามารถของผู้สอบโดยวิธีแมกซิมัม ไลค์ลิฮูด แบบมีเงื่อนไข (Conditional Maximum Likelihood Estimation) ด้วยโปรแกรมคอมไพเลอร์ภาษา TURBO PASCAL Version 7.0 ที่ผู้วิจัยเป็นผู้พัฒนา

6. วิธีของเบส หมายถึง วิธีประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้เข้าสอบในแบบจำลองโลจิสติก 3 พารามิเตอร์ ด้วยวิธีของเบส ที่มีการกำหนดการแจกแจงเริ่มแรกของค่าอำนาจจำแนก และค่าการเดา ด้วยโปรแกรมคอมพิวเตอร์ ภาษา TURBO PASCAL Version 7.0 ที่ผู้วิจัยเป็นผู้พัฒนา

7. ค่าพารามิเตอร์ หมายถึง ค่าที่ได้จากการวิเคราะห์แบบทดสอบวิชาภาษาไทย เป็นรายข้อ โดยใช้แบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซีมัม ไลค์ลิสต์ วิธีลิวิริสติก และวิธีของเบส ได้ค่าพารามิเตอร์ของข้อสอบ คือ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดาของข้อสอบ (c) และค่าพารามิเตอร์ความสามารถของผู้สอบ (๑) ด้วยเทคนิควิธีมอนติ คาร์โล จากการทดลองซ้ำกลุ่มละ 40 ครั้ง

8. ค่าความยากของข้อสอบ (b) หมายถึง ค่าที่แสดงถึงความสามารถที่จุดโค้งลักษณะข้อสอบมีความชันมากที่สุด (จุดเปลี่ยนโค้ง) ซึ่งเป็นระดับความสามารถที่จะตอบข้อนั้นถูกต้องด้วยความน่าจะเป็น 0.5 เมื่อไม่มีการเดา ค่าความยากของข้อสอบนี้จะมีค่าตั้งแต่ $-\alpha$ ถึง $+\alpha$ แต่ในทางปฏิบัติจะพิจารณาค่าระหว่าง -3 ถึง +3 (Warm. 1978 : 52 ; Lord. 1980 : 12) ค่า -3 แสดงว่า ข้อสอบข้อนี้น่ายากมาก ค่า +3 แสดงว่า ข้อสอบข้อนี้นั้นง่ายมาก

9. ค่าอำนาจจำแนกของข้อสอบ (a) หมายถึง ค่าที่เป็นสัดส่วนกับความชันของโค้ง หรือความชันของความน่าจะเป็นที่ผู้สอบซึ่งมีระดับความสามารถ ๑ จะตอบข้อสอบข้อนั้นถูก ณ จุดเปลี่ยนโค้ง ค่าอำนาจจำแนกนี้มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติจะพิจารณาค่าระหว่าง 0.5 ถึง 2.5 (Warm. 1978 : 52) ค่า 0 แสดงว่า ข้อสอบไม่มีอำนาจจำแนก ค่า 2.5 แสดงว่า ข้อสอบมีค่าอำนาจจำแนกสูง

10. ค่าการเดาของข้อสอบ (c) หมายถึง ค่าที่แสดงถึงความน่าจะเป็นที่ผู้สอบสามารถตอบข้อสอบข้อนั้นถูกโดยไม่มีความรู้ ($\phi = -\infty$) ค่าการเดาของข้อสอบมีค่าที่เป็นไปได้ตั้งแต่ 0 ถึง 1 ถ้าข้อสอบข้อนั้นไม่สามารถที่จะตอบถูกได้ด้วยการเดาแล้ว ค่าการเดาจะเท่ากับ 0 (Lord. 1980 : 12 - 14) ในทางปฏิบัติจะพิจารณาคัดเลือกข้อสอบที่มีค่าการเดาน้อยกว่า 0.30

11. ค่าความสามารถของผู้สอบ (๑) หมายถึง คะแนนจริงของผู้เข้าสอบที่ประมาณได้จากคะแนนรวมของการตอบแบบทดสอบ โดยแบบจำลองโลจิสติก 3 พารามิเตอร์

12. แบบจำลองโลจิสติก 3 พารามิเตอร์ หมายถึง แบบจำลองที่แสดงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบถูกต้องกับระดับความสามารถในรูปของฟังก์ชันการแจกแจงสะสม

แบบโลจิสต์ โดยแบบจำลองนี้ประกอบด้วยค่าพารามิเตอร์ของข้อสอบ 3 ค่า คือ ค่าความยาก ค่าอำนาจจำแนก และค่าการเดา

13. ขนาดของกลุ่มตัวอย่าง หมายถึง จำนวนผู้สอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี พ.ศ. 2537 ของ อ.ก.ค. กรมสามัญศึกษา ซึ่งได้รับการสุ่มจากประชากรเป็นกลุ่มตัวอย่างจำนวน 13 ขนาด คือ ขนาด 200 คน 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน

14. จำนวนข้อสอบต่างกัน หมายถึง จำนวนข้อสอบของแบบทดสอบวิชาภาษาไทยที่ใช้ในการสอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี พ.ศ. 2537 ของ อ.ก.ค.กรมสามัญศึกษา ซึ่งออกตามหลักสูตรที่ใช้ในการสอบแข่งขันที่ ก.ค. เป็นผู้กำหนด มีลักษณะเป็นแบบปรนัย ชนิด 4 ตัวเลือก จำนวน 100 ข้อ โดยผู้วิจัยได้แบ่งข้อสอบออกเป็น 3 เนื้อหาใหญ่ ๆ ตามตารางวิเคราะห์เนื้อหา แล้วจึงทำการสุ่มข้อสอบตามตารางวิเคราะห์เนื้อหาเป็น 6 ขนาด คือ ขนาดข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ 90 ข้อ และ 100 ข้อ

15. เทคนิควิซิมอนติ คาร์โล หมายถึง การใช้คอมพิวเตอร์ช่วยในการจำลองสถานการณ์ (Simulation) โดยใช้ตัวเลขสุ่ม (Random Number) ที่มีลักษณะการกระจายของความน่าจะเป็นแบบสม่ำเสมอ (Uniform Distribution) มาสร้างรูปแบบการตอบข้อสอบของกลุ่มตัวอย่างให้เหมือนสถานการณ์จริง และมีการทดลองซ้ำ เพื่อหาข้อสรุปเกี่ยวกับประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์

ประโยชน์ที่คาดว่าจะได้รับ

ผลจากการศึกษาค้นคว้าครั้งนี้ คาดว่ามีประโยชน์ดังต่อไปนี้

1. ผลจากการตรวจสอบทำให้ทราบว่าวิธีการประมาณค่าพารามิเตอร์วิธีใด จะมีประสิทธิภาพมากที่สุด เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน
2. เป็นการขยายความรู้ในเชิงทฤษฎีเกี่ยวกับประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ ด้วยวิธีแนกขั้วมิม ไลค์ลิสต์ วิธีวิวิธิตติ และวิธีของเบส์ เมื่อจำนวนข้อสอบและขนาดของกลุ่มตัวอย่างกัน
3. เป็นแนวทางในการศึกษาวิธีประมาณค่าพารามิเตอร์ในมิติอื่น ๆ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษารังนี้ ผู้วิจัยได้ศึกษาเนื้อหาต่าง ๆ
ดังต่อไปนี้

1. แนวคิดของทฤษฎีการตอบข้อคำถาม
2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม
3. ลักษณะเด่นของทฤษฎีการตอบข้อคำถาม
4. แบบจำลองที่ใช้ในทฤษฎีการตอบข้อคำถาม
5. ดัชนีกำหนดคุณภาพและการเปรียบเทียบคุณภาพของแบบทดสอบ
 - 5.1 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ
 - 5.2 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ
6. การเปรียบเทียบประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์
7. วิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์
 - 7.1 วิธีแมกซ์ิมัม ไลค์ลิฮูด
 - 7.2 วิธีลิวริสติก
 - 7.3 วิธีของเบส์
8. การศึกษาโดยใช้เทคนิควิธีมอนติ คาร์โล
 - 8.1 ความเป็นมาของเทคนิควิธีมอนติ คาร์โล
 - 8.2 ขั้นตอนของระเบียบเทคนิควิธีมอนติ คาร์โล
 - 8.3 จุดเด่นของการใช้เทคนิควิธีมอนติ คาร์โล
9. งานวิจัยที่เกี่ยวข้อง
 - 9.1 งานวิจัยที่เกี่ยวข้องภายในประเทศ
 - 9.2 งานวิจัยที่เกี่ยวข้องต่างประเทศ

1. แนวคิดของทฤษฎีการตอบข้อคำถาม

ทฤษฎีการตอบข้อคำถามเป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างลักษณะ (Trait) หรือความสามารถ (Ability) ที่มีอยู่ภายในตัวบุคคล ซึ่งไม่สามารถสังเกตได้ กับพฤติกรรม การตอบข้อคำถามของบุคคลนั้น ซึ่งเป็นสิ่งที่สังเกตได้โดยตรง (Lord and Novick, 1968)

: 358) ในรูปแบบจำลองทางคณิตศาสตร์โดยให้คะแนนที่ได้รับจากการตอบข้อสอบ y แทน พฤติกรรมการตอบข้อสอบ และ θ แทนคุณลักษณะหรือความสามารถภายในตัวบุคคล ฟังก์ชัน การถดถอยของ y บน θ จะเป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงถึงความสัมพันธ์ระหว่าง พฤติกรรมการตอบข้อสอบกับระดับความสามารถ ที่เรียกว่า ฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) (Lord and Novick. 1968 : 360) ซึ่งสามารถ กำหนดได้หลายรูปแบบ ทั้งนี้ขึ้นอยู่กับข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลที่ได้จากการทดสอบ (Hambleton and Cook. 1977 : 75)

2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม

แม้ทฤษฎีการตอบข้อคำถามจะประกอบด้วยแบบจำลองหลายแบบจำลอง และในแต่ละ แบบจำลองจะมีข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลจากการทดสอบและความสัมพันธ์ระหว่างตัวแปร ที่สังเกตได้กับตัวแปรที่สังเกตไม่ได้ก็ตาม แต่วางก็มีข้อตกลงเบื้องต้นที่สำคัญ คือ

2.1 มิติของลาเทนท์สเปซ (Dimensionality of Latent Space) กล่าวคือ ลักษณะหรือความสามารถที่เป็นตัวกำหนดพฤติกรรมการตอบข้อสอบข้อใดข้อหนึ่ง อาจมีได้หลาย ลักษณะ เรียกลักษณะเหล่านี้ว่า ลาเทนท์สเปซ (Latent Space) และจำนวนลักษณะทั้งหมด ในลาเทนท์สเปซ ก็คือ มิติของลาเทนท์สเปซ ในทฤษฎีการตอบข้อคำถามโดยทั่วไปมักจะถือว่า ลักษณะหรือความสามารถที่เป็นตัวกำหนดพฤติกรรมการตอบข้อสอบจะมีเพียงลักษณะเดียว และยอมรับกันว่าเป็นข้อตกลงเบื้องต้นที่สำคัญประการหนึ่งของแบบจำลองหลายแบบในทฤษฎีนี้ที่ เรียกว่า Unidimensionality Assumption ซึ่งข้อตกลงนี้ต้องการเพียงว่า มีความ สามารถ หรือคุณลักษณะที่เด่นและสำคัญ (Dominant Factor) เพียงลักษณะเดียวที่เป็นตัว กำหนดพฤติกรรมการตอบข้อสอบ การถือว่าลาเทนท์สเปซที่สนใจมีเพียงมิติเดียว จะช่วยให้ แบบจำลองของทฤษฎีการตอบข้อคำถามมีความซับซ้อนน้อยลง และง่ายต่อการแปลความหมาย ของคะแนนที่ได้จากแบบทดสอบ กล่าวคือ องค์ประกอบที่เด่นและสำคัญนั้นจะเป็นความสามารถ ของผู้เข้าสอบ (Hambleton and Swaminathan. 1985 : 16 - 17) และสามารถทำ การตรวจสอบความเป็นไปได้ของข้อตกลงนี้ด้วยวิธีการวิเคราะห์องค์ประกอบ (Hambleton and Cook. 1977 : 78)

ในกรณีที่ไม่สามารถทำการตรวจสอบความเป็นมิติเดียวโดยการวิเคราะห์องค์- ประกอบได้ อาจคาดคะเนความเป็นมิติเดียวของแบบทดสอบได้ดังนี้ (Warm. 1978 : 101)

2.1.1 แบบทดสอบที่มีลักษณะเหมือนกับจะวัดทักษะต่าง ๆ เพียงอย่างเดียว เช่น แบบทดสอบสะกดคำ คำศัพท์ อ่านเอาเรื่อง อุปมาอุปไมย แบบทดสอบวัดมิติสัมพันธ์ คณิตศาสตร์ และความเข้าใจในการใช้เครื่องมือ เป็นต้น

2.1.2 แบบทดสอบที่วัดความรู้เป็นส่วน ๆ แต่มีเนื้อหาเกี่ยวข้องกันมาก เช่น แบบทดสอบชีววิทยา ซึ่งต้องอาศัยความรู้เกี่ยวกับเคมี เซลล์ และพลังงาน

2.1.3 แบบทดสอบที่มีการวัดเนื้อหาที่สัมพันธ์กันในเชิงตรรกและเรียงลำดับต่อเนื่องกัน เช่น เหตุผลเชิงคณิตศาสตร์ อนุกรมเลข

2.2 ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึง พฤติกรรมการตอบข้อสอบต่าง ๆ ในแบบทดสอบของบุคคลใดบุคคลหนึ่งมีความเป็นอิสระเชิงสถิติ กล่าวคือ การตอบข้อสอบข้อหนึ่งไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ ในแบบทดสอบ และถ้าการตอบข้อสอบมีความเป็นอิสระแล้ว ความน่าจะเป็นร่วมของคะแนนรายชื่อของผู้เข้าสอบแต่ละคน คือ

$$\text{Prob} [U_1=u_1, U_2=u_2, \dots, U_n=u_n | \theta] = \prod_{i=1}^n P_i(\theta)^{u_i} Q_i(\theta)^{1-u_i}$$

เมื่อ U_i , $i = 1, 2, \dots, n$ เป็นผลการตอบข้อสอบ n ข้อ (เมื่อตอบถูก $u_i = 1$, เมื่อตอบผิด $u_i = 0$)

P_i = ความน่าจะเป็นของผู้สอบคนหนึ่งที่ทำข้อสอบข้อที่ i ถูก

$Q_i = 1 - P_i$

สำหรับการตรวจสอบความเป็นอิสระในการตอบข้อสอบนั้น แฮมเบิลตัน และสวามินาธาน (Hambleton and Swaminathan. 1985 : 24) กล่าวว่า ถ้าแบบทดสอบมีความเป็นมิติเดียวแล้ว แบบทดสอบจะมีความเป็นอิสระในการตอบข้อสอบด้วย

2.3 โค้งลักษณะข้อสอบ (Item Characteristic Curve) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบข้อนั้นถูก กับระดับความสามารถที่วัดได้ด้วยข้อสอบข้อนั้น โค้งลักษณะข้อสอบมีหลายรูปแบบ ทั้งนี้ขึ้นอยู่กับความเชื่อในแบบจำลองที่ใช้อธิบายความสัมพันธ์ดังกล่าว (Hambleton and Swaminathan. 1985 : 25) สำหรับข้อสอบที่มีการตรวจให้คะแนนเป็น 1 เมื่อตอบถูก และให้คะแนน 0

เมื่อตอบผิดนั้น ความน่าจะเป็น $[P_i(e)]$ ในการตอบถูกสามารถเขียนแทนได้ด้วยฟังก์ชันโค้งปกติสะสมหรือฟังก์ชันโลจิสติก (Lord. 1980 : 12 - 13)

เมื่อทำการเลือกแบบจำลอง หรือฟังก์ชันทางคณิตศาสตร์ที่ใช้แสดงความน่าจะเป็นในการตอบถูก และเก็บรวบรวมข้อมูลแล้ว อาจมีการตรวจสอบความเหมาะสมระหว่างแบบจำลองกับข้อมูลว่ามีความเหมาะสมเพียงใด การตรวจสอบแบบจำลองสามารถทำได้โดยตรง ถ้าเราสามารถวัดความสามารถได้อย่างถูกต้อง แต่เนื่องจากเราไม่สามารถวัดความสามารถได้ การตรวจสอบดังกล่าวจึงเป็นไปได้ยาก วิธีที่ทำได้ คือ การหาผลการทำนายจากแบบจำลองหลาย ๆ ครั้ง แล้วตรวจสอบกับข้อมูลที่วัดได้ว่า ผลการทำนายหลาย ๆ ครั้งนั้นถูกต้องโดยประมาณหรือไม่ ในทางปฏิบัติจะแทนพารามิเตอร์จริงด้วยค่าพารามิเตอร์ที่ประมาณได้ โดยหวังว่าจะเหมาะสมโดยประมาณกับข้อมูลที่วัดได้ เช่น ถ้าเลือกใช้แบบจำลองของราสส์ต้องตรวจสอบว่า ข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากัน และมีค่าการเดาใกล้เคียง หรือถ้าเลือกใช้แบบจำลองโลจิสติก แบบ 2 พารามิเตอร์ ต้องตรวจสอบว่า การตอบข้อสอบของผู้เข้าสอบมีการเดาน้อย (Hambleton and Swaminathan. 1985 : 157 - 161)

3. ลักษณะเด่นของทฤษฎีการตอบข้อคำถาม

ถ้าแบบจำลองการตอบข้อคำถามและข้อมูลจากผลการสอบมีความเหมาะสมกันแล้ว ทฤษฎีการตอบข้อคำถามจะมีความเหนือกว่า (Superiority) ทฤษฎีการทดสอบแบบดั้งเดิมดังนี้ (Hambleton and Swaminathan. 1985 : 10 - 13)

3.1 ความเป็นอิสระจากกลุ่มผู้สอบ (Sample-Free) หมายถึง ไม่ว่าจะนำข้อสอบไปใช้สอบกับบุคคลใด โค้งลักษณะข้อสอบจะคงเดิม

3.2 ความเป็นอิสระจากกลุ่มข้อสอบ (Item-Free) หมายถึง ในการประมาณค่าความสามารถ (e) ของผู้เข้าสอบ จะใช้ข้อสอบชุดใดก็ได้ จำนวนเท่าใดก็ได้ ซึ่งบางครั้งอาจใช้เพียง 3 - 5 ข้อ ก็สามารถประมาณค่าความสามารถที่แท้จริงของผู้เข้าสอบได้ ทั้งนี้ข้อสอบที่นำมาใช้ต้องได้รับการคัดเลือกจากคลังข้อสอบขนาดใหญ่ ที่ข้อสอบแต่ละข้อได้ระบุคุณลักษณะที่วัดไว้

3.3 ความไม่แปรเปลี่ยนของค่าพารามิเตอร์ (Invariant of Item Parameter) หมายถึง ไม่ว่าจะประมาณค่าพารามิเตอร์จากกลุ่มตัวอย่างที่มีความสามารถระดับใดก็ตาม ค่าพารามิเตอร์ที่ได้จะไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ทั้งนี้เพราะฟังก์ชันการตอบข้อสอบ ซึ่งเป็นฟังก์ชันที่อธิบายความสัมพันธ์ระหว่างพฤติกรรมการตอบข้อสอบกับระดับความ

สามารถนั้น เป็นฟังก์ชันการถดถอย (Regression Function) ของคะแนนจากการสอบบนความสามารถ ซึ่งตามทฤษฎีสถิติ ฟังก์ชันการถดถอยจะไม่แปรเปลี่ยนไปตามการแจกแจงความถี่ของตัวแปรทำนาย ดังนั้น ค่าพารามิเตอร์ต่าง ๆ ในสมการถดถอย ซึ่งเป็นตัวกำหนดรูปร่างของเส้นถดถอย จึงไม่เปลี่ยนแปลงด้วย (Lord. 1980 : 34)

เบเกอร์ (Baker. 1977 : 170) กล่าวว่า ความไม่แปรเปลี่ยนของค่าพารามิเตอร์ขึ้นอยู่กับเงื่อนไข 2 ประการ คือ ประการแรก ความสามารถ (Ability) ที่กล่าวถึง ต้องสามารถนิยามได้ชัดเจนและวัดได้ด้วยข้อสอบ ประการที่สอง ความสามารถที่วัดนั้นต้องมีความคงที่ภายในช่วงระยะเวลาหนึ่ง นอกจากนี้ ลอร์ด (Lord. 1980 : 36) ยังอธิบายเพิ่มเติมว่า การไม่แปรเปลี่ยนของค่าพารามิเตอร์นั้น มิได้หมายความว่า ค่าพารามิเตอร์ที่ประมาณได้โดยใช้กลุ่มตัวอย่างที่แตกต่างกันจะมีค่าเท่ากันเสมอ การที่ค่าพารามิเตอร์ของข้อสอบจะเท่ากันหรือไม่ ขึ้นอยู่กับเงื่อนไขในการทดสอบบางประการ เช่น ถ้าเลือกมาตรวัดที่มีจุดเริ่มต้นและหน่วยในการวัดเดียวกันแล้ว ค่าพารามิเตอร์ที่ประมาณได้จากกลุ่มตัวอย่างที่แตกต่างกันจะมีค่าเท่ากัน แต่ถ้าเลือกมาตรวัดที่มีจุดเริ่มต้นและหน่วยในการวัดแตกต่างกันแล้ว การไม่แปรเปลี่ยนของค่าพารามิเตอร์จะหมายความว่า ค่าพารามิเตอร์ที่ประมาณได้จากกลุ่มตัวอย่างที่ต่างกันของข้อสอบชุดหนึ่ง จะมีความสัมพันธ์กันในเชิงเส้นตรง

4. แบบจำลองที่ใช้ในทฤษฎีการตอบข้อคำถาม

ทฤษฎีการตอบข้อคำถามได้มีการพัฒนาแบบจำลองขึ้นมาหลายแบบ โดยแบบจำลองแต่ละแบบจะต่างกันในเรื่องของฟังก์ชันทางคณิตศาสตร์ และจำนวนพารามิเตอร์ที่กล่าวถึง ซึ่งสามารถจำแนกได้เป็น 3 ประเภทใหญ่ ๆ คือ (Hambleton and Swaminathan. 1985 : 35 - 52)

4.1 แบบจำลองที่ประยุกต์ใช้กับการให้คะแนนแบบ 2 ค่า (Binary Item) คือ ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน สำหรับแบบจำลองนี้สามารถแบ่งตามลำดับขั้นการพัฒนาเป็น 2 ระยะ คือ

4.1.1 แบบจำลองในระยะเริ่มแรก (ค.ศ. 1943 - 1968) เป็นแบบจำลองที่ไม่ค่อยจะมีความสอดคล้องกับข้อมูลที่ได้จากการสอบในสถานการณ์จริงมากนัก แต่ก็ เป็นแบบจำลองที่มีคุณค่าต่อการพัฒนาเป็นแบบจำลองที่เหมาะสมในระยะเวลาต่อมา ตัวอย่างแบบจำลองประเภทนี้ เช่น Guttman Perfect Scale, Latent Distance Model และ Linear Model

4.1.2 แบบจำลองที่ได้รับการพัฒนาขึ้นในระยะหลัง (ค.ศ. 1952 - 1982) เป็นแบบจำลองที่ได้รับการปรับปรุงให้มีความสอดคล้องกับข้อมูลที่ได้จากการสอบในสถานการณ์จริงมากขึ้น สะดวกในการนำไปปฏิบัติ และมีผู้สนใจนำไปใช้อย่างแพร่หลาย ได้แก่ One-, Two-, Three-Parameter Normal Ogive Model และ One-, Two-, Three-, Four-Parameter Logistic Model

4.2 แบบจำลองที่ประยุกต์ใช้กับการให้คะแนนที่มากกว่า 2 ค่า (Multichotomously Scored) แบบจำลองนี้มีจุดมุ่งหมายในการเพิ่มความแม่นยำในการประมาณค่าความสามารถของผู้เข้าสอบ โดยใช้สารสนเทศจากการตอบข้อสอบถูกหรือผิด ตัวอย่างแบบจำลองประเภทนี้ ได้แก่ Norminal Response Model, Grade Response Model และ Partial Credit Model

4.3 แบบจำลองที่ประยุกต์ใช้กับการให้คะแนนแบบต่อเนื่อง (Continuous) เป็นแบบจำลองที่ใช้ประมาณค่าความสามารถของผู้เข้าสอบทางจิตวิทยา หรือทางด้านที่เกี่ยวกับเจตคติ (Attitude) ซึ่งต้องมีการตอบข้อสอบในมาตราที่ต่อเนื่อง (Continuous Scale) ตัวอย่างแบบจำลองประเภทนี้ คือ Continuous Response Model ซึ่ง เซมิจิมา (Samejima. 1972) เป็นผู้พัฒนาขึ้น

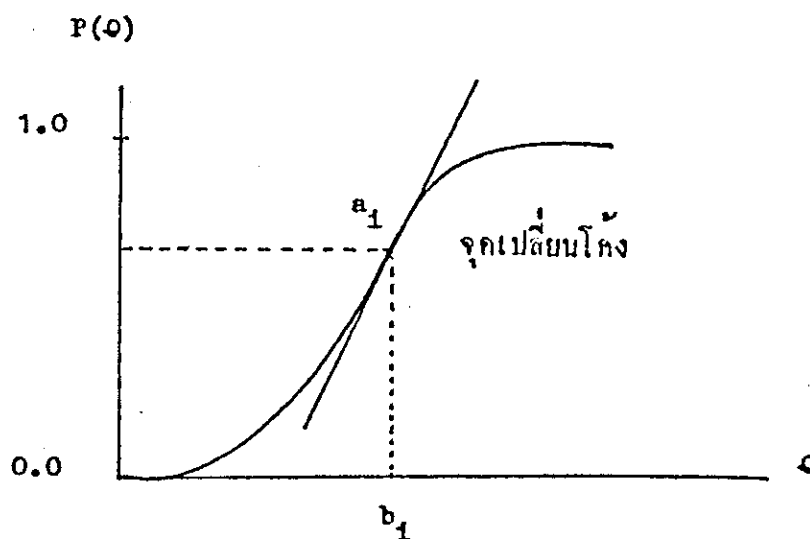
สำหรับการวิจัยครั้งนี้ ผู้วิจัยจะเสนอรายละเอียดเฉพาะแบบจำลองที่มีการประยุกต์ใช้กับการให้คะแนนแบบ 2 ค่า ที่ได้รับการพัฒนาขึ้นในระยะหลัง และเป็นที่ยอมรับในปัจจุบัน คือ แบบจำลองปกติของออโวก์ และแบบจำลองโลจิสติก เท่านั้น

แบบจำลองปกติของออโวก์ (Normal Ogive Model)

แบบจำลองนี้ได้รับการพัฒนาขึ้นในปี ค.ศ. 1944 โดย ลอเลย์ (Lawley) (Lord. 1986 : 369) และลอว์ได้ทำการศึกษาและปรับปรุงแบบจำลองนี้ พร้อมทั้งพัฒนาวิธีการประมาณค่าพารามิเตอร์ของแบบจำลอง โดยนำไปใช้กับข้อมูลที่ได้จากแบบทดสอบวัดผลสัมฤทธิ์ และแบบทดสอบวัดความถนัด (Hambleton and Swaminathan. 1985 : 35) โดยแบบจำลองนี้เชื่อว่า ความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบถูกกับระดับความสามารถจะอยู่ในรูปของฟังก์ชันการแจกแจงสะสมปกติ (Cumulative Normal Ogive) ที่แสดงได้ดังสมการ

$$P_1(e) = \int_{-\infty}^{a_1(e - b_1)} (2\pi)^{-1/2} \exp[-t^2/2] dt$$

สมการนี้จะมีค่าพารามิเตอร์ของข้อสอบเพียง 2 ค่า คือ a_1 และ b_1 จึงเรียกแบบจำลองที่อธิบายด้วยสมการนี้ว่า แบบจำลองปกติของออร์ไว 2 พารามิเตอร์ (Two-Parameter Normal Ogive Model) ในแบบจำลอง 2 พารามิเตอร์นี้ เส้นโค้งลักษณะข้อสอบจะมีลักษณะคล้ายรูปตัว S ดังภาพประกอบ 1 โดยที่โอกาสในการตอบข้อสอบถูกจะมีค่าเข้าใกล้ 0 เมื่อ e มีค่าน้อย ๆ และจะมีค่าเพิ่มขึ้นเรื่อย ๆ และจะเข้าใกล้ 1 เมื่อ e มีค่ามาก ๆ

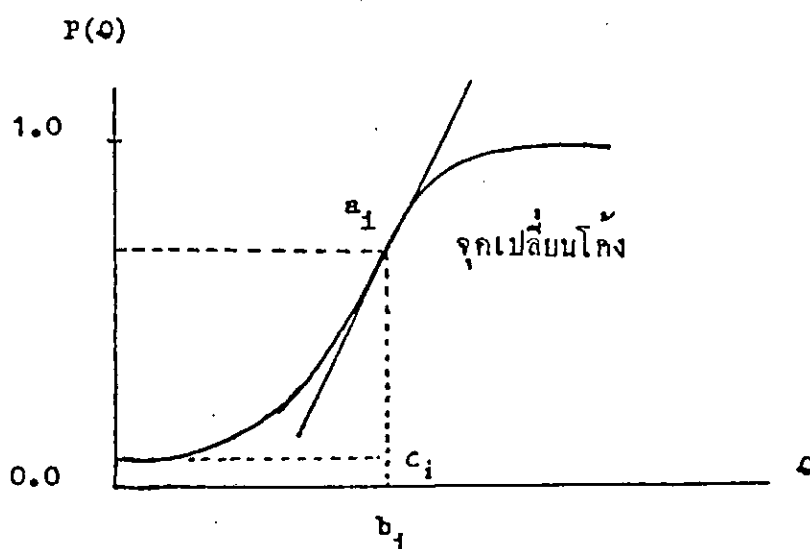


ภาพประกอบ 1 เส้นโค้งลักษณะข้อสอบของแบบจำลองปกติของออร์ไว 2 พารามิเตอร์

แต่ในกรณีที่ใช้ข้อสอบแบบเลือกตอบ ผู้เข้าสอบที่มีความสามารถเพียงเล็กน้อย อาจมีโอกาสดตอบข้อสอบถูกได้โดยการเดา ทำให้แบบจำลองที่มีพารามิเตอร์ 2 ตัว ไม่เหมาะที่จะใช้อธิบายสถานการณ์การสอบเช่นนี้ เพราะถ้า e มีค่าน้อยแล้ว ค่า $P_1(e)$ จะไม่เข้าใกล้ 0 อีกต่อไป แต่จะมีค่าเข้าใกล้ค่า ๆ หนึ่งที่มากกว่า 0 ดังนั้น จึงมีการเพิ่มค่าพารามิเตอร์ของข้อสอบเข้าไปอีก 1 ตัว คือ c_1 และสมการของแบบจำลอง 3 พารามิเตอร์นี้ คือ

$$P_1(\mathbf{e}) = c_1 + (1-c_1) \int_{-\infty}^{a_1(\mathbf{e}-b_1)} (2\pi)^{-1/2} \exp[-t^2/2] dt$$

เรียกแบบจำลองนี้ว่า แบบจำลองปกติของลอจิสต์ 3 พารามิเตอร์ (Three-Parameter Normal Ogive Model) และเส้นโค้งลักษณะข้อสอบจะมีลักษณะดังภาพประกอบ 2



ภาพประกอบ 2 เส้นโค้งลักษณะข้อสอบของแบบจำลองปกติของลอจิสต์ 3 พารามิเตอร์

อนึ่ง รูปร่างของโค้งลักษณะข้อสอบแต่ละข้อ จะแตกต่างกันไปขึ้นอยู่กับค่าพารามิเตอร์ของข้อสอบที่เป็นดัชนีบอกลักษณะของข้อสอบ โดยค่าพารามิเตอร์แต่ละตัวมีความหมาย ดังต่อไปนี้

ค่าอำนาจจำแนก (a_1) มีค่าเป็นส่วนโดยตรงกับความชันของเส้นโค้งลักษณะข้อสอบ ณ จุดเปลี่ยนโค้ง (ความชัน ณ จุดเปลี่ยนโค้ง มีค่าเท่ากับ $a_1(1-c_1)/\sqrt{2\pi}$) ค่า a_1 นี้ เป็นค่าพารามิเตอร์ที่บ่งชี้คุณภาพของข้อสอบในแง่สารสนเทศที่จะได้ในการประมาณค่าระดับความสามารถ ($\mathbf{e}$) โดยมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ในทางปฏิบัติค่า a_1 ของข้อสอบโดยทั่วไปจะมีค่าอยู่ระหว่าง 0.50 ถึง 2.50 (Warm. 1978 : 52)

ค่าความยาก (b_i) เป็นค่าระดับความสามารถ (θ) ระดับหนึ่งที่ทำให้เส้นโค้งลักษณะข้อสอบมีการเปลี่ยนโค้ง ณ จุดเปลี่ยนโค้งจะอยู่ที่ $(b, P(b))$ ค่า b_i นี้ เป็นค่าพารามิเตอร์ที่บอกตำแหน่งของเส้นโค้งลักษณะข้อสอบ ณ $\theta = b$ และถ้า c_i มีค่า = 0 โอกาสที่ผู้สอบจะตอบข้อสอบถูกต้องจะมีค่าเท่ากับ $(1-c_i)/2$ หรือเท่ากับ 0.5 ค่า b_i มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ ในทางปฏิบัติจะพิจารณาค่า b_i ที่อยู่ในช่วง -3 ถึง +3 (Warm. 1978 : 52 ; Lord. 1980 : 12)

ค่าการเดา (c_i) (Guessing Parameter or Pseudo-Chance Score Level) เป็นค่ากำกับเส้นโค้งที่ต่ำสุด (Lower Assymtotes) ของเส้นโค้งลักษณะข้อสอบ และเป็นค่าที่แสดงถึงค่าโอกาสที่บุคคลที่ไม่มีความสามารถเลย ($\theta = -\infty$) จะตอบข้อสอบถูกต้อง ซึ่ง c_i นี้ มีค่าอยู่ระหว่าง 0 ถึง 1 ถ้าข้อสอบข้อใดไม่สามารถจะตอบถูกต้องด้วยการเดาแล้ว ค่า c_i จะมีค่าเท่ากับ 0 (Lord. 1980 : 12 - 14)

ระดับความสามารถ (θ) หมายถึง คะแนนจริงของผู้สอบที่ประมาณได้จากคะแนนรวมของการตอบแบบทดสอบ โดยปรับให้เป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 และความสามารถนี้สามารถเปลี่ยนแปลงได้จากการเรียนรู้ (Hambleton and Swaminathan. 1985 : 54 - 55)

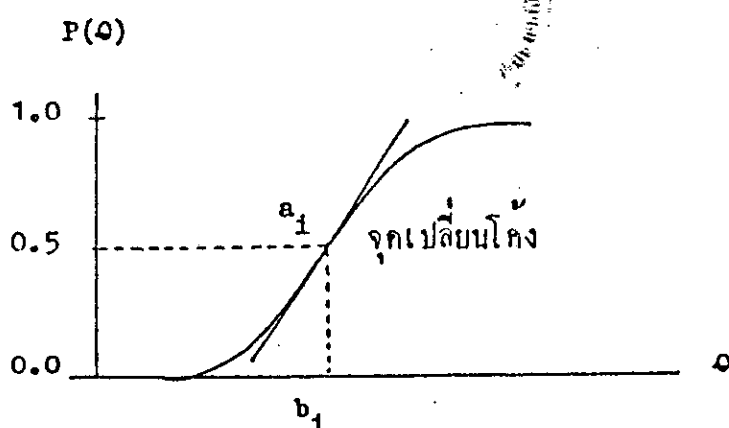
แบบจำลองโลจิสติก (Logistic Model)

แบบจำลองโลจิสติกเป็นแบบจำลองที่มีลักษณะใกล้เคียงกับแบบจำลองปกติของออโลว์ กล่าวคือ ในแบบจำลองโลจิสติก เส้นโค้งลักษณะข้อสอบของแบบจำลองปกติของออโลว์ และแบบจำลองโลจิสติกจะมีลักษณะใกล้เคียงกันมาก คือ เป็นรูปตัว S เนื่องจากค่าฟังก์ชันปกติของออโลว์กับค่าฟังก์ชันโลจิสติกจะแตกต่างกันเพียงเล็กน้อย โดยในแบบจำลองโลจิสติกนั้น ความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบถูกต้องกับระดับความสามารถจะอยู่ในรูปของฟังก์ชันการแจกแจงสะสมแบบโลจิสติก (Logistic Cumulation Distribution Function) ซึ่งมีความสะดวกในการประมาณค่ามากกว่า สำหรับค่าพารามิเตอร์ของข้อสอบ (a_i , b_i และ c_i) ในแบบจำลองโลจิสติกมีความหมายเช่นเดียวกับแบบจำลองปกติของออโลว์ แต่มีค่า D ซึ่งเป็นค่าองค์ประกอบของการปรับสเกลและมีค่าคงที่เท่ากับ 1.70 เข้ามาเกี่ยวข้อง แบบจำลองโลจิสติกนี้ได้รับการพัฒนาเพื่อให้เหมาะสมกับสถานการณ์การทดสอบ เช่น ในกรณีที่ผู้สอบตอบข้อสอบได้ด้วยการเดาปลายค่าสุดของโค้งลักษณะข้อสอบจะไม่เป็นศูนย์ แสดงว่า การทำข้อสอบนั้นมีการเดาเกิดขึ้น และในกรณีที่ผู้สอบทำข้อสอบด้วยความสะเพร่า

แบบจำลองนี้จะเพิ่มค่าความรอบคอบขึ้นอีกหนึ่งค่า ซึ่งจะเห็นได้จากปลายสูงสุดของโค้งลักษณะข้อสอบมีค่าน้อยกว่า 1 แบบจำลองโลจิสติกนี้แบ่งเป็น 4 แบบจำลองย่อย ๆ ตามพารามิเตอร์ดังนี้ (Hambleton and Swaminathan. 1985 : 34 - 49)

1. แบบจำลองที่ใช้พารามิเตอร์ตัวเดียว (One-Parameter Logistic Model) หรือราสช์โมเดล (Rasch Model) แบบจำลองนี้ถือว่า ข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากัน ($a_i = a$) และไม่มีการเดา ($c_i = 0$) สมการจะมีค่าพารามิเตอร์เพียง 1 ตัว คือ b_i เขียนเป็นสมการได้ดังนี้

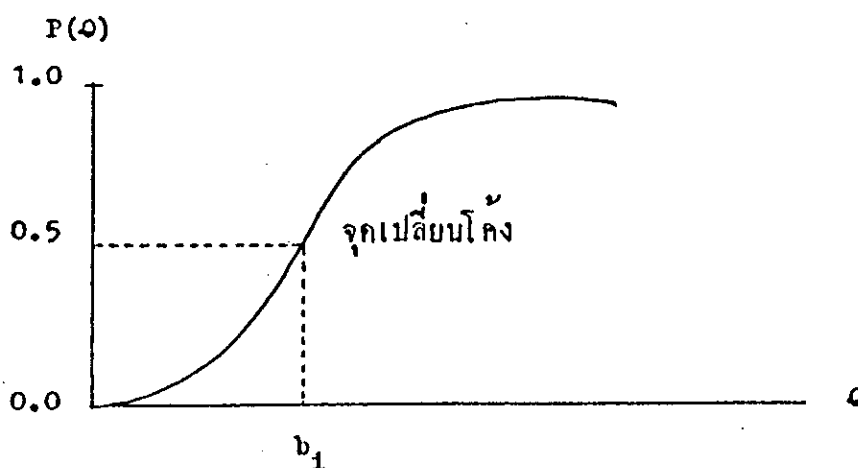
$$P_i(\theta_i) = 1 / (1 + e^{-(\theta_i - b_i)})$$



ภาพประกอบ 3 เส้นโค้งลักษณะข้อสอบของแบบจำลองโลจิสติก 1 พารามิเตอร์

2. แบบจำลองที่ใช้พารามิเตอร์ 2 ตัว (Two-Parameter Logistic Model) เบิร์นบอม (Birnbau) ได้เสนอแบบจำลองนี้ในปี ค.ศ. 1957 และพัฒนาโดยใช้พารามิเตอร์ 2 ตัว คือ ค่าความยาก (b_i) และค่าอำนาจจำแนก (a_i) ในกรณีที่ผู้สอบมีโอกาสตอบข้อสอบถูกต้องด้วยการเดาน้อยมาก โดยกำหนดให้ค่าพารามิเตอร์การเดา (c_i) = 0 เขียนเป็นสมการได้ดังนี้

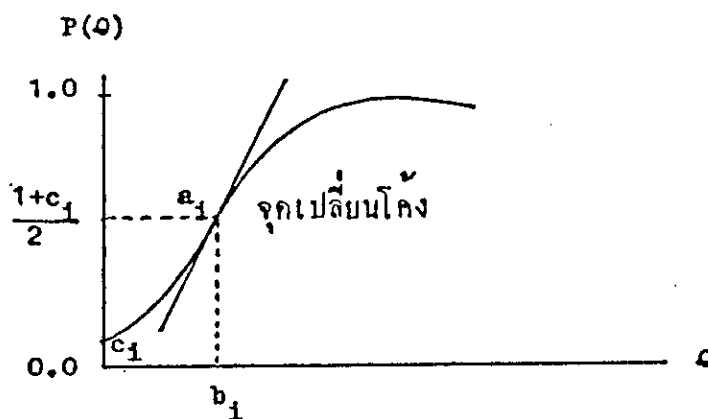
$$P_1(e_{-1}) = 1 / (1 + e^{-D a_1 (e_{-1} - b_1)})$$



ภาพประกอบ 4 เส้นโค้งลักษณะข้อสอบของแบบจำลองโลจิสติก 2 พารามิเตอร์

3. แบบจำลองที่ใช้พารามิเตอร์ 3 ตัว (Three-Parameter Logistic Model) เป็นแบบจำลองที่พัฒนามาจากแบบจำลองที่ใช้พารามิเตอร์ 2 ตัว เพื่อให้เหมาะสมกับแบบทดสอบที่มีอิทธิพลของการเดาเข้ามาเกี่ยวข้อง โดยเพิ่มพารามิเตอร์อีกหนึ่งตัว คือ ค่าการเดา เขียนเป็นสมการได้ดังนี้

$$P_1(e_{-1}) = c_1 + (1 - c_1) / (1 + e^{-D a_1 (e_{-1} - b_1)})$$



ภาพประกอบ 5 เส้นโค้งลักษณะข้อสอบของแบบจำลองโลจิสติก 3 พารามิเตอร์

4. แบบจำลองที่ใช้พารามิเตอร์ 4 ตัว (Four-Parameter Logistic Model) แบบจำลองนี้เสนอโดยแม็คโดนัลด์ (McDonald) ในปี ค.ศ. 1967 ซึ่ง ลอร์ด (Lord) และ บาร์ตัน (Barton) ได้พัฒนาในปี ค.ศ. 1981 โดยเพิ่มพารามิเตอร์ค่าความรอบคอบ (Careless ; γ_i) ในแบบจำลองที่ใช้พารามิเตอร์ 3 ตัว ซึ่งเป็นค่าที่แสดงให้เห็นถึงกรณีที่ผู้สอบที่มีความสามารถสูงตอบข้อสอบผิด อาจจะเป็นเพราะความไม่รอบคอบ ทำให้ผู้นั้นไม่เลือกคำตอบที่ถูกต้อง ซึ่งจะเห็นได้จากปลายสูงสุดของโค้งลักษณะข้อสอบต่ำกว่า 1 เขียนเป็นสมการได้ดังนี้ (Hambleton and Swaminathan. 1985 : 49)

$$P_i(\theta_i) = c_i + (\gamma_i - c_i) / (1 + e^{-D \cdot a_i(\theta_i - b_i)})$$

5. ดัชนีกำหนดคุณภาพและการเปรียบเทียบคุณภาพของแบบทดสอบ

ทฤษฎีการตอบข้อคำถามพิจารณาคุณภาพของแบบทดสอบจากดัชนีที่ใช้กำหนดคุณภาพ คือ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function) ซึ่งเป็นฟังก์ชันของค่าพารามิเตอร์ของข้อสอบที่เป็นดัชนีผสม โดยสร้างจากดัชนีที่ใช้กำหนดคุณภาพของข้อสอบหลาย ๆ ดัชนีของทุก ๆ ข้อ ในแบบทดสอบ โดยกำหนดว่า ณ ระดับความสามารถใดที่ค่าฟังก์ชันสารสนเทศของแบบทดสอบมีค่าสูง แสดงว่า แบบทดสอบนั้นมีคุณภาพดีมากในการประมาณค่าความสามารถที่ระดับนั้น และดัชนีที่ใช้เปรียบเทียบคุณภาพของแบบทดสอบ คือ ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ โดยการนำค่าฟังก์ชันสารสนเทศของแบบทดสอบมาหาอัตราส่วน ณ ระดับความสามารถเดียวกัน (วิชุดา บัวคง. 2533 : 32)

5.1 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ

สารสนเทศสามารถนิยามได้จากลักษณะความไม่แน่นอนของเหตุการณ์บางประการ ถ้าความไม่แน่นอนในการเกิดเหตุการณ์หนึ่งมีน้อย แสดงว่า มีสารสนเทศเกี่ยวกับเหตุการณ์นั้นมาก จากนั้นสามารถกล่าวเมื่อนำไปใช้ในทฤษฎีการตอบข้อคำถาม โดยพิจารณาจากกระบวนการอ้างอิงเชิงสถิติแล้ว ความไม่แน่นอนในการประมาณค่าพารามิเตอร์ในทางทฤษฎีจะมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$

ความไม่แน่นอนในการประมาณค่าพารามิเตอร์ ในทางทฤษฎีสามารถแสดงได้ในลักษณะของความแปรปรวนที่ได้จากการสุ่มกลุ่มตัวอย่างเพื่อประมาณค่าพารามิเตอร์ ถ้าเราประมาณค่าเฉลี่ยของประชากรจะได้ค่าความคลาดเคลื่อนมาตรฐานกำลังสองของค่าเฉลี่ยเป็น $\sigma_x^2 = \sigma_{x/n}^2$ ซึ่งถ้าความคลาดเคลื่อนมาตรฐานกำลังสองมีค่าน้อยเท่าไร ค่าสารสนเทศจะเพิ่มมากขึ้น ดังนั้น จึงกำหนดสารสนเทศเกี่ยวกับพารามิเตอร์ว่าเป็นสัดส่วนกลับของความ

คลาดเคลื่อนมาตรฐานกำลังสอง ดังสมการ

$$I_x = 1 / 6x^2$$

สารสนเทศเป็นสิ่งที่มีความสำคัญมากสำหรับทฤษฎีการตอบข้อคำถาม เพราะแบบทดสอบและสเกลต่าง ๆ พัฒนาขึ้นโดยใช้กระบวนการตอบข้อคำถามที่ประเมินได้จากความแน่นอนในการประมาณค่าความสามารถ (๑) จากแบบทดสอบ และสเกลต่าง ๆ โดยทั่วไปแล้วจะใช้ค่าความคลาดเคลื่อนมาตรฐานกำลังสองของความสามารถ (๑) หรือสารสนเทศเป็นตัววัดความแน่นอนในการประมาณค่าสารสนเทศนี้ ไม่จำเป็นต้องมีค่าคงที่ตลอดพิสัยของความสามารถ และค่าสารสนเทศ ณ ระดับความสามารถใด จะขึ้นอยู่กับจำนวนและคุณลักษณะของข้อสอบที่ใช้ในการประมาณค่าความสามารถ (๑) นั้น ซึ่งเบิร์นบอร์ม (Birnbaum, 1968) และลอร์ด (Lord, 1980) ได้นิยามฟังก์ชันสารสนเทศของแบบทดสอบว่าเป็นสัดส่วนกลับของความแปรปรวนอย่างสุ่มของการประมาณค่าความสามารถของผู้เข้าสอบแต่ละคนแบบความน่าจะเป็นสูงสุด เมื่อเทียบกับความสามารถรวมของผู้เข้าสอบทุกคน (Hulin, 1983 : 55) นั้นหมายความว่า เมื่อจำนวนข้อสอบเพิ่มขึ้น $1/\sqrt{I(๑)}$ ก็จะเป็นความคลาดเคลื่อนมาตรฐานของความสามารถ ซึ่งเขียนเป็นสมการได้ดังนี้

$$I(๑_i) = \frac{\sum_{j=1}^n [P_j(๑_i)]'^2}{P_i(๑_i) Q_i(๑_i)}$$

เมื่อ $[P_j(๑)]'$ เป็นค่าอนุพันธ์ของโค้งลักษณะข้อสอบ (ICC) ของข้อสอบข้อที่ j ณ ระดับความสามารถ ๑ และฟังก์ชันสารสนเทศของข้อสอบ มีสูตรดังนี้

$$I(๑_j, U_j) = \frac{[P_j(๑_j)]'^2}{P_j(๑_j) Q_j(๑_j)}$$

$$\text{เมื่อ } P_j(๑_j)' = Da_j Q_j(P_j - c_j) / (1 - c_j)$$

$$P_i(\theta_i) = c_i + (1 - c_i)/(1 + e^{-Da_i(\theta_i - b_i)})$$

$$Q_i(\theta_i) = (1 - c_i)/(1 + e^{Da_i(\theta_i - b_i)})$$

วาร์ม (Warm, 1978 : 77) กล่าวว่า ในการสอบทุกครั้งย่อมมีความคลาดเคลื่อนในการวัดเกิดขึ้น ที่เรียกว่า ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ (Standard Error of the Ability Estimates : SE(θ)) ซึ่งมีค่าเท่ากับรากที่สองของส่วนกลับของฟังก์ชันสารสนเทศของแบบทดสอบ ดังสมการ

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

ค่าฟังก์ชันสารสนเทศของแบบทดสอบนี้ จะมีค่าแปรเปลี่ยนไปตามสเกลความสามารถ และแปรผกผัน (Inverse) กับความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ดังนั้น ถ้าฟังก์ชันสารสนเทศมีค่ามากแล้วความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจะมีค่าน้อย นอกจากนี้ ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ($SE^2(\theta)$) ของผู้เข้าสอบทั้งหมด ยังมีความสัมพันธ์กับค่าความเชื่อมั่นของทฤษฎีการทดสอบแบบดั้งเดิมด้วย เมื่อคะแนนเป็นมาตรฐานและมีค่าความเบี่ยงเบนมาตรฐานในการวัดเป็น 1 ดังสมการ (Warm, 1978 : 77)

$$r_{xx} = 1 - \overline{SE^2(\theta)}$$

5.2 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ

ทฤษฎีการตอบข้อคำถามมีจุดมุ่งหมายในการสอบ คือ เมื่อใช้แบบทดสอบวัดความสามารถของผู้สอบแต่ละคนแล้ว ต้องการให้ผลการสอบนั้นสามารถประมาณค่าความสามารถที่มีอยู่ภายในตัวบุคคลที่ไม่สามารถสังเกตหรือวัดได้โดยตรง ดังนั้น แบบทดสอบที่ใช้จึงควรเป็นแบบทดสอบที่ให้ผลการสอบที่สามารถนำไปใช้ในการประมาณค่าความสามารถที่ต้องการวัดได้ถูกต้องแม่นยำมากที่สุด ซึ่งดัชนีที่ใช้ในการประมาณค่าความสามารถที่ต้องการวัดได้อย่างถูกต้องแม่นยำมากที่สุด ก็คือ การพิจารณาคุณภาพของแบบทดสอบ หรือค่าฟังก์ชันสารสนเทศ

ของแบบทดสอบ นั้นเอง โดยกำหนดว่า ๗ ระดับความสามารถที่ค่าฟังก์ชันสารสนเทศของแบบทดสอบมีค่าสูง แสดงว่า แบบทดสอบนั้นมีคุณภาพที่ดีในการประมาณค่าความสามารถที่ระดับนั้น

ค่าฟังก์ชันสารสนเทศนี้ นอกจากจะใช้ในการพิจารณาคุณภาพของแบบทดสอบแล้ว ยังสามารถนำมาใช้ในการเปรียบเทียบคุณภาพระหว่างแบบทดสอบ 2 ฉบับ ได้อีกด้วย (Hambleton and Swaminathan, 1985 :101) กล่าวคือ ถ้าเรามีแบบทดสอบหลายฉบับที่ใช้วัดความสามารถเดียวกันแล้ว เราสามารถเปรียบเทียบคุณภาพของแบบทดสอบเหล่านั้นได้ โดยการนำเอาค่าฟังก์ชันสารสนเทศของแบบทดสอบที่ต้องการเปรียบเทียบ 2 ฉบับ ๗ ระดับความสามารถเดียวกันมาหาอัตราส่วนกัน ค่าอัตราส่วนของค่าฟังก์ชันสารสนเทศของแบบทดสอบ 2 ฉบับ ที่นำมาเปรียบเทียบ ๗ ระดับความสามารถเดียวกัน ก็คือ ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ นั้นเอง ซึ่งสามารถเขียนเป็นสมการได้ดังนี้

$$RE(y, x) = \frac{I(e, y)}{I(e, x)}$$

นั่นคือ ค่าประสิทธิภาพสัมพัทธ์ของคะแนนจากแบบทดสอบ y กับคะแนนจากแบบทดสอบ x ก็คือ อัตราส่วนระหว่างค่าฟังก์ชันสารสนเทศของแบบทดสอบทั้ง 2 ฉบับ โดยที่คะแนน x และ y เป็นคะแนนที่ได้จากแบบทดสอบที่แตกต่างกัน 2 ฉบับ ๗ ระดับความสามารถเดียวกัน หรือ x และ y เป็นผลจากวิธีการให้คะแนนหรือวิธีการประมาณค่าพารามิเตอร์ที่แตกต่างกัน 2 วิธี ของแบบทดสอบฉบับเดียวกัน

จากนิยามของค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบดังกล่าว ค่าความสามารถใน $I(e, y)$ จึงเป็นค่าความสามารถเดียวกันกับใน $I(e, x)$ และค่าประสิทธิภาพสัมพัทธ์ของคะแนนจากแบบทดสอบ 2 ฉบับ จะมีค่าแปรเปลี่ยนไปตามระดับความสามารถด้วย (Lord, 1980 : 83) อย่างไรก็ตาม ค่าประสิทธิภาพสัมพัทธ์นี้ไม่ขึ้นอยู่กับสเกลที่ใช้วัดความสามารถ กล่าวคือ ค่าดัชนีนี้จะไม่แปรเปลี่ยนไปตามการแปลงสเกลการวัดความสามารถ (Lord, 1980 : 89)

การแปลความหมายของค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ๗ ระดับความสามารถใด สามารถแปลได้ดังนี้ (Warm, 1978 : 76)

ถ้าค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่ามากกว่า 1 แสดงว่า แบบทดสอบ y มีคุณภาพสูงกว่าแบบทดสอบ x ที่ระดับความสามารถนั้น

ถ้าค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่าน้อยกว่า 1 แสดงว่า แบบทดสอบ y มีคุณภาพต่ำกว่าแบบทดสอบ x ที่ระดับความสามารถนั้น และ

ถ้าค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่าเท่ากับ 1 แสดงว่า แบบทดสอบ y และแบบทดสอบ x มีคุณภาพเท่ากัน ที่ระดับความสามารถนั้น

ซึ่งแนวคิดและวิธีการหาคุณภาพของแบบทดสอบทั้ง 2 ประการ คือ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ และค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ดังกล่าวนี้นี้ สามารถนำไปใช้กับวิธีประมาณค่าพารามิเตอร์ทุกวิธีในแบบจำลองโลจิสติก (วิชุดา นีวงศ์. 2533 : 16)

6. การเปรียบเทียบประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์

เนื่องจากวิธีการประมาณค่าพารามิเตอร์มีหลายวิธี ดังนั้น จึงจำเป็นต้องมีการเปรียบเทียบประสิทธิภาพสัมพัทธ์ของวิธีการประมาณค่าพารามิเตอร์ 2 วิธีใด ๆ ว่า วิธีการประมาณค่าพารามิเตอร์วิธีใดจะดีกว่า โดยการหาอัตราส่วนระหว่างค่าความแปรปรวนของวิธีประมาณค่าพารามิเตอร์วิธีที่หนึ่ง กับค่าความแปรปรวนของวิธีประมาณค่าพารามิเตอร์วิธีที่สอง เช่น สมมติให้ M_1 และ M_2 เป็นวิธีประมาณค่าพารามิเตอร์วิธีที่หนึ่งและวิธีที่สอง โดยมี $V(M_1)$ และ $V(M_2)$ เป็นความแปรปรวนของวิธีประมาณค่าพารามิเตอร์วิธีที่ 1 และวิธีที่ 2 ตามลำดับแล้ว ถ้าต้องการเปรียบเทียบประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์ M_1 กับ M_2 จะได้เท่ากับ $V(M_1) / V(M_2)$ (สุชาดา กิระนันท์. 2525 : 42 - 43) แต่จากดัชนีกำหนดคุณภาพของแบบทดสอบ พบว่า ความแปรปรวนที่ได้จากการประมาณค่าพารามิเตอร์เป็นส่วนกลับกับค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Hulin. 1983 : 55; citing Birnbaum. 1968; Lord. 1980) ดังนั้น จะได้ว่า

$$\frac{V(M_1)}{V(M_2)} = \frac{I(\theta, M_2)}{I(\theta, M_1)}$$

ทำให้สามารถใช้การเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบฉบับเดียวกันที่เป็นผลมาจากวิธีการประมาณค่าพารามิเตอร์ 2 วิธีใด ๆ แทน การเปรียบเทียบประสิทธิภาพของวิธี

ประมาณค่าพารามิเตอร์ 2 วิธีใด ๆ ในแบบทดสอบฉบับเดียวกัน ($EFF(x,y)$) ได้ กล่าวคือ วิธีประมาณค่าพารามิเตอร์ x เทียบกับวิธีประมาณค่าพารามิเตอร์ y ในแบบทดสอบฉบับหนึ่ง มีค่าเท่ากับค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบฉบับนั้น ($RE(y,x)$) ที่เป็นผลมาจากวิธีประมาณค่าพารามิเตอร์ด้วยวิธี y เทียบกับวิธี x ณ ระดับความสามารถใด ๆ ดังสมการ

$$EFF(x,y) = \frac{V(o,x)}{V(o,y)} = \frac{I(o,y)}{I(o,x)} = RE(y,x)$$

ดังนั้น ถ้า $RE(y,x)$ มีค่ามากกว่า 1 แล้ว $EFF(x,y)$ ก็จะมีค่ามากกว่า 1 ด้วย นั่นหมายความว่า การประมาณค่าพารามิเตอร์ด้วยวิธี y มีประสิทธิภาพสูงกว่าวิธี x และการประมาณค่าพารามิเตอร์ด้วยวิธี x มีความแปรปรวนสูงกว่าวิธี y แต่ถ้า $RE(y,x)$ มีค่าน้อยกว่า 1 แล้ว $EFF(x,y)$ ก็จะมีค่าน้อยกว่า 1 ด้วย ซึ่งหมายความว่า การประมาณค่าพารามิเตอร์ด้วยวิธี y มีประสิทธิภาพต่ำกว่าวิธี x และการประมาณค่าพารามิเตอร์ด้วยวิธี x มีความแปรปรวนน้อยกว่าวิธี y จะเห็นได้ว่า ไม่ว่าจะใช้ค่าความแปรปรวนหรือค่าฟังก์ชันสารสนเทศของแบบทดสอบที่ได้จากวิธีประมาณค่าพารามิเตอร์ 2 วิธีใด ๆ มาทำการเปรียบเทียบเพื่อหาประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์ 2 วิธีใด ๆ แล้ว การแปลผลที่ได้จะมีความหมายเหมือนกัน สำหรับการวิจัยครั้งนี้ ผู้วิจัยใช้การเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบฉบับเดียวกันที่เป็นผลมาจากวิธีการประมาณค่าพารามิเตอร์วิธีเดียวกัน เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน แทนการใช้ค่าสัดส่วนของความแปรปรวน ในการเปรียบเทียบวิธีประมาณค่าพารามิเตอร์วิธีเดียวกัน

7. วิธีประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์

แบบจำลองแต่ละแบบของทฤษฎีการตอบข้อคำถามจะมีวิธีการประมาณค่าพารามิเตอร์แตกต่างกัน ซึ่งแฮมเบิลตัน และคูก์ (Hambleton and Cook. 1977 : 75 - 95) ได้สรุปวิธีประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ในทฤษฎีการตอบข้อคำถามว่ามี 3 วิธีใหญ่ ๆ คือ

- 7.1 วิธีแมกซิมัม ไลค์ลิฮูด (Maximum Likelihood Estimation)
- 7.2 วิธีฮิวริสติก (Heuristic Estimation)
- 7.3 วิธีของเบย์ส์ (Bayesian Estimation)

7.1 วิธีแมกซิมัม ไลค์ลิฮูด (Maximum Likelihood Estimation)

แนวคิดของวิธีแมกซิมัม ไลค์ลิฮูด มีว่า การประมาณค่าพารามิเตอร์ทำได้โดยอาศัยผลที่ได้จากตัวอย่างสุ่มที่เลือกมาจากการแจกแจงที่ทราบรูปแบบของฟังก์ชันความหนาแน่น แต่ไม่ทราบค่าพารามิเตอร์ จึงใช้ความน่าจะเป็นในการเลือกตัวอย่าง และวัดค่าได้ ($U_1 = u_1, U_2 = u_2, \dots, U_n = u_n$) มาพิจารณาค่าประมาณของค่าพารามิเตอร์ (สุชาดา กิระนันท์. 2525 : 100)

จากการนิยามความน่าจะเป็นของผู้เข้าสอบในการตอบข้อสอบข้อที่ i เมื่อ $U_i = 1$ สำหรับการตอบถูก และ $U_i = 0$ สำหรับการตอบผิด ซึ่งแสดงได้ดังสมการ

$$\begin{aligned} P(U_i | \theta, b, a, c) &= P(U_i = 1 | \theta, b, a, c) P(U_i = 0 | \theta, a, b, c) \\ &= P_i^{u_i} (1 - P_i)^{1-u_i} \\ &= P_i^{u_i} Q_i^{1-u_i} ; \quad Q_i = 1 - P_i \end{aligned}$$

ถ้าผู้เข้าสอบตอบข้อสอบ n ข้อ และแบบทดสอบมีลักษณะการวัดเพียงมิติเดียว (คือ มีความเป็นอิสระเฉพาะที่) แล้ว ความน่าจะเป็นของการตอบแสดงได้ด้วยฟังก์ชันความหนาแน่นร่วม (Joint Density Function) ดังสมการต่อไปนี้

$$\begin{aligned} P(U_1, U_2, \dots, U_n | \theta, b, a, c) &= P(U_1 | \theta, b, a, c) P(U_2 | \theta, a, b, c) \dots \\ &\quad P(U_n | \theta, a, b, c) \\ &= \prod_{i=1}^n P(U_i | \theta, b, a, c) \\ &= \prod_{i=1}^n P_i^{u_i} Q_i^{1-u_i} \end{aligned}$$

สมการดังกล่าวเป็นความน่าจะเป็นของการตอบข้อสอบ n ข้อ ที่สามารถวัดหรือสังเกตได้โดยที่ $U_1, U_2, \dots, U_n$ เป็นตัวแปรสุ่มที่มีค่าเฉพาะเป็น $u_1, u_2, \dots, u_n$ เมื่อ u_i มีค่า

เท่ากับ 1 หรือ 0 และเนื่องจากสมการนี้เป็นฟังก์ชันทางคณิตศาสตร์ของค่า θ, b, a, c ที่จะบอกว่า ตัวแปรสุ่มนี้มีโอกาสเกิดขึ้นเมื่อใด จึงเรียกฟังก์ชันนี้ว่าฟังก์ชันความน่าจะเป็นหรือ ฟังก์ชันไลค์ลิฮูด (Likelihood Function) ซึ่งแสดงได้ดังสมการ

$$L(u_1, u_2, \dots, u_n | \theta, b, a, c) = P_1^{u_1} Q_1^{1-u_1}$$

เมื่อ $u_i = 1$ ค่าของ Q_i ก็หมดไป เพราะ $Q_i^{1-u_i} = 1$ และเมื่อ $u_i = 0$ ค่าของ P_i จะหมดไป เพราะ $P_i^{u_i} = 1$ และฟังก์ชันไลค์ลิฮูดที่มีผู้สอบ N คน ตอบข้อสอบ n ข้อ มีสมการดังนี้

$$L(u | \theta, b, a, c) = L(u_1, u_2, \dots, u_n | \theta_1, \theta_2, \dots, \theta_n, \\ b_1, b_2, \dots, b_n, a_1, a_2, \dots, a_n, \\ c_1, c_2, \dots, c_n)$$

$$L(u | \theta, b, a, c) = \prod_{j=1}^n \prod_{i=1}^n L(u_{i,j} | \theta, b, a, c)$$

$$L(u | \theta, b, a, c) = \prod_{j=1}^n \prod_{i=1}^n P_{i,j}^{u_{i,j}} Q_{i,j}^{1-u_{i,j}}$$

$$\begin{aligned} \text{เมื่อ } u &= \text{เวกเตอร์ผลการตอบข้อสอบ } n \text{ ข้อ ของผู้เข้าสอบ } N \text{ คน} \\ u_{i,j} &= \text{เวกเตอร์ผลการตอบข้อสอบ } n \text{ ข้อ ของผู้เข้าสอบคนที่ } j \\ P_{i,j} &= P_i(\theta_j, b_i, a_i, c_i) \end{aligned}$$

การประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด คือ การหาค่าประมาณของค่าพารามิเตอร์ที่จะทำให้ฟังก์ชันไลค์ลิฮูดมีค่าสูงสุด ซึ่งโดยปกติจะทำการหาค่าพารามิเตอร์ที่ทำให้ลอการิทึม (Logarithm) ของฟังก์ชันไลค์ลิฮูดมีค่าสูงสุด ทั้งนี้เพราะค่าประมาณของพารามิเตอร์ที่ทำให้ฟังก์ชันทั้ง 2 ฟังก์ชันมีค่าสูงสุดเป็นค่าเดียวกัน แต่การหาค่าประมาณที่ทำให้ลอการิทึมของฟังก์ชันไลค์ลิฮูดมีค่าสูงสุดนั้นทำได้ง่ายกว่า ซึ่งลอการิทึมของฟังก์ชันไลค์ลิฮูด

ของตัวแปร e, b, a, c คือ

$$\ln L(u|e, b, a, c) = \sum_{j=1}^N \sum_{i=1}^n (u_{ij} \ln P_{ij} + (1 - u_{ij}) \ln Q_{ij})$$

และการหาค่าพารามิเตอร์ e, b, a, c ที่ทำให้ $\ln L(u|e, b, a, c)$ มีค่าสูงสุด ทำได้โดยกำหนดให้อนุพันธ์ของ $\ln L(u|e, b, a, c)$ มีค่าเท่ากับศูนย์ แล้วแก้สมการหาค่ารากของอนุพันธ์ ดังนี้

$$d \ln L(u|e, b, a, c) / d e_j = 0 \quad ; \quad j = 1, 2, \dots, N$$

$$d \ln L(u|e, b, a, c) / d b_i = 0 \quad ; \quad i = 1, 2, \dots, n$$

$$d \ln L(u|e, b, a, c) / d a_i = 0 \quad ; \quad i = 1, 2, \dots, n$$

$$d \ln L(u|e, b, a, c) / d c_i = 0 \quad ; \quad i = 1, 2, \dots, n$$

โดยทั่วไปสมการโลคัลลิสต์นี้ไม่เป็นสมการเส้นตรง ดังนั้น การหาค่า e_j, a_i, b_i, c_i ที่ทำให้ $\ln L(u|e, b, a, c)$ มีค่าสูงสุดจึงไม่สามารถหาได้ด้วยการใช้วิธีการอย่างง่าย แต่หาได้โดยใช้วิธีของนิวตัน ราฟสัน (Newton-Raphson Procedure) (Lord, 1980 : 179 - 191) ซึ่งเป็นการหาค่าประมาณโดยการทำซ้ำ ๆ (Iterative) จนมีค่าคงที่ โดยมีขั้นตอนดังต่อไปนี้

ขั้นที่ 1 กำหนดค่าเริ่มต้นสำหรับการประมาณค่าความสามารถ e_j และค่าพารามิเตอร์ a_i, b_i และ c_i

1. ค่าเริ่มต้นสำหรับการประมาณค่าความสามารถ e_j คือ

$$e_j^{(0)} = \ln (x_j / (n - x_j))$$

$$\text{เมื่อ } \ln = \text{Natural Logarithm}$$

$$x_j = \text{คะแนนสอบของคนที่ } j$$

$$n = \text{จำนวนข้อสอบ}$$

2. ค่าเริ่มต้นสำหรับการประมาณค่าพารามิเตอร์ของข้อสอบ a_i, b_i และ c_i อาจใช้ค่าประมาณจากวิธีวิวิธคติ หรือค่าอื่น ๆ ที่เหมาะสม

ขั้นที่ 2 ประมวลค่าความสามารถของผู้เข้าสอบแต่ละคน e_j โดยกำหนดให้ค่าพารามิเตอร์ของข้อสอบที่ประมวลได้ในครั้งก่อนเป็นค่าคงที่ ดังนี้

$$e_{j1}^{(m+1)} = e_{j1}^{(m)} - g(e_{j1}^{(m)})/h(e_{j1}^{(m)})$$

เมื่อ $e_{j1}^{(m)}, e_{j1}^{(m+1)}$ = ค่าประมวลความสามารถของผู้เข้าสอบคนที่ j ครั้งที่ m และครั้งที่ $m+1$

$$g(e_{j1}^{(m)}) = d \ln L(u|e, b, a, c)/de_{j1}; \text{ ประมวลที่ } e_{j1}^{(m)}$$

$$h(e_{j1}^{(m)}) = d^2 \ln L(u|e, b, a, c)/de_{j1}^2; \text{ ประมวลที่ } e_{j1}^{(m)}$$

ประมวลค่าซ้ำ (Iterative) จนกว่าค่าประมวลความสามารถ e_{j1} จะมีค่าเข้าใกล้ค่าใดค่าหนึ่ง (Convergence) คือ ค่าประมวลครั้งที่ $m+1$ และครั้งที่ m มีค่าแตกต่างกันน้อยกว่าค่าคงที่ที่กำหนดไว้ เช่น 0.001

ขั้นที่ 3 ประมวลค่าพารามิเตอร์ของข้อสอบแต่ละข้อ a_i, b_i และ c_i โดยกำหนดให้ค่าความสามารถของผู้เข้าสอบที่ประมวลได้ในครั้งก่อนเป็นค่าคงที่

$$x_{i1}^{(m+1)} = x_{i1}^{(m)} - g(x_{i1}^{(m)})/h(x_{i1}^{(m)})$$

เมื่อ $x_{i1}^{(m)}, x_{i1}^{(m+1)}$ = เวกเตอร์ของค่า a, b และ c ของข้อสอบข้อที่ i
= ค่าประมวลพารามิเตอร์ของข้อสอบข้อที่ i ครั้งที่ m และครั้งที่ $m+1$

$$g(x_{i1}^{(m)}) = d \ln L(u|e, b, a, c)/dx_{i1} \text{ ของค่า } x_{i1}^{(m)}$$

$$h(x_{i1}^{(m)}) = d^2 \ln L(u|e, b, a, c)/dx_{i1}^2 \text{ ของค่า } x_{i1}^{(m)}$$

และประมวลค่าซ้ำ (Iterative) จนกว่าค่าประมวลจะลู่เข้าใกล้ค่าใดค่าหนึ่ง (Convergence)

ขั้นที่ 4 ประมวลค่าซ้ำในขั้นที่ 2 และขั้นที่ 3 จนกว่าค่าประมาณ θ , a , b , และ c จะมีค่าคงที่ และมีความถูกต้องเพียงพอ หรือทำให้ $\ln L(\theta, b, a, c)$ มีค่าสูงที่สุด

ค่าอนุพันธ์อันดับที่ 1 และค่าอนุพันธ์อันดับที่ 2 ของค่าพารามิเตอร์ θ , a , b , c แสดงไว้ในตารางที่ 1 ส่วนค่าสุดท้ายที่ได้จากการประมวลจะเรียกว่า ค่าที่ได้จากการประมวลโดยวิธีแมกซิมัม ไลค์ลิฮูด และเพื่อให้ได้ค่าคงที่เร็วขึ้นในการประมวลค่าซ้ำแต่ละครั้ง ค่าความสามารถจะถูกปรับให้มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ซึ่งจะเป็นผลให้ค่า a , b และ c ต้องถูกปรับตามไปด้วย (Hambleton and Swaminathan. 1985 : 127 - 138)

ตาราง 1 อนุพันธ์อันดับที่ 1 และอนุพันธ์อันดับที่ 2 สำหรับฟังก์ชันไลค์ลิฮูดของค่าพารามิเตอร์ ข้อสอบ (a , b , c) และความสามารถของผู้เข้าสอบ (θ) ในแบบจำลองโลจิสติก 3 พารามิเตอร์

อนุพันธ์ (Derivative)	สัญลักษณ์แสดง (Expression)
$d \ln L / d \theta_i$	$D \sum_{i=1}^n \frac{a_i (P_{iA} - c_i) (u_{iA} - P_{iA})}{(1 - c_i) P_{iA}}$
$d \ln L / d a_i$	$\frac{D}{(1 - c_i)} \sum_{i=1}^n \frac{(\theta_i - b_i) (P_{iA} - c_i) (u_{iA} - P_{iA})}{P_{iA}}$
$d \ln L / d b_i$	$\frac{-D a_i}{(1 - c_i)} \sum_{i=1}^n \frac{(P_{iA} - c_i) (u_{iA} - P_{iA})}{P_{iA}}$

ตาราง 1 (ต่อ)

อนุพันธ์ (Derivative)	สัญลักษณ์แสดง (Expression)
$d \ln L / d c_i$	$\frac{1}{(1 - c_i)} \sum_{j=1}^n \frac{(u_{ij} - P_{ij})}{P_{ij}}$
$d^2 \ln L / d c_i^2$	$D^2 \sum_{j=1}^n \frac{a_i (P_{ij} - c_i) Q_{ij}}{(1 - c_i)^2 P_{ij}} \left[\frac{u_{ij} c_i - P_{ij}}{P_{ij}} \right]$
$d^2 \ln L / d a_i$	$\frac{D^2}{(1 - c_i)} \sum_{j=1}^n \left[\frac{(c_i - b_i)^2 (P_{ij} - c_i) Q_{ij}}{P_{ij}} \left(\frac{u_{ij} c_i - P_{ij}}{P_{ij}} \right) \right]$
$d^2 \ln L / d b_i$	$\frac{D^2 a_i}{(1 - c_i)^2} \sum_{j=1}^n \frac{(P_{ij} - c_i) Q_{ij}}{P_{ij}} \left[\frac{u_{ij} c_i - P_{ij}}{P_{ij}} \right]$
$d^2 \ln L / d c_i$	$\frac{1}{(1 - c_i)^2} \sum_{j=1}^n \left[\frac{u_{ij} - 1}{P_{ij}} \right] - \left[\frac{u_{ij} Q_{ij}}{P_{ij}} \right]$

$$P_{ij} = c_i + (1 - c_i) / (1 + \exp^{(-D a_i (c_j - b_i))})$$

$$u_{ij} = \text{ผลการตอบข้อที่ } i \text{ ของคนที่ } j$$

หมายเหตุ

1. เมื่อต้องการใช้แบบจำลองโลจิสติก 2 พารามิเตอร์ ให้กำหนดค่า c_1 เป็น 0
2. เมื่อต้องการใช้แบบจำลองโลจิสติก 1 พารามิเตอร์ ให้กำหนดค่า c_1 เป็น 0 และ a_1 เป็น 1

คุณสมบัติที่สำคัญของวิธีแมกซิมัม ไลค์ลิฮูด (Hambleton and Swaminathan.

1985 : 88 - 89)

1. มีความคงที่เมื่อจำนวนของกลุ่มผู้เข้าสอบและข้อสอบเพิ่มขึ้น การประมาณค่าก็จะมีค่าคงที่ไปสู่ค่าที่แท้จริง
2. ฟังก์ชันของค่าสถิติมีความเพียงพอของสารสนเทศทั้งหมดที่เกี่ยวข้องกับพารามิเตอร์
3. มีประสิทธิภาพ เพราะวิธีการประมาณค่าวิธีนี้มีความแปรปรวนน้อยที่สุด
4. มีลักษณะของการกระจายเข้าใกล้แบบปกติ

ในกรณีที่กลุ่มตัวอย่างผู้เข้าสอบและข้อสอบมีขนาดเล็กคุณสมบัติดังกล่าวข้างต้นจะเป็นจริงเมื่อทำการประมาณค่าความสามารถของบุคคล หรือประมาณค่าพารามิเตอร์ของข้อสอบเพียงอย่างใดอย่างหนึ่งเท่านั้น (Hambleton and Swaminathan. 1985 : 127 - 129) แต่ถ้าผู้เข้าสอบและจำนวนข้อสอบเพิ่มขึ้น การประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้เข้าสอบไปพร้อม ๆ กัน ค่าประมาณที่ได้จะมีค่าเข้าใกล้ค่าพารามิเตอร์จริง (Swaminathan and Gifford. 1982 : 350) ดังนั้น ในการวิจัยครั้งนี้จึงทำการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด เพราะเป็นการประมาณค่าพารามิเตอร์ของข้อสอบก่อน แล้วจึงประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ ด้วยโปรแกรม BILOG 3 ของมิลส์เลวี และบ็อค เนื่องจากผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ เมื่อจำนวนผู้สอบมีขนาดตั้งแต่ 200 คน ถึง 3,000 คน และข้อสอบมีขนาดตั้งแต่ 50 ข้อ ถึง 100 ข้อ

ข้อจำกัดและปัญหาของกระบวนการประมาณค่าพารามิเตอร์ของข้อสอบและ
ความสามารถของผู้เข้าสอบ ไปพร้อม ๆ กันของวิธีแมกซิมัม ไลค์ลิสต์ (Hambleton and
 Swaminathan. 1985 : 132)

1. ในการประมาณค่าพารามิเตอร์ในขั้นที่ 2 และ 3 โดยใช้ค่าอนุพันธ์
 อันดับที่ 2 จากตาราง 1 เป็นตัวหารในกระบวนการนิวตัน ราวสันนั้น มีโอกาสที่ค่าประมาณ
 ที่ได้จะไม่ลู่เข้าสู่ค่าคงที่ จึงอาจหลีกเลี่ยงปัญหานี้ได้โดยใช้ค่าสารสนเทศที่อยู่ในตาราง 2
 แทน

2. เมื่อสมการไลค์ลิสต์ไม่ใช่สมการเชิงเส้นตรง ทำให้สามารถหารากของ
 สมการที่ทำให้ฟังก์ชันไลค์ลิสต์มีค่าสูงสุดได้หลายค่า แต่ค่าเหล่านี้ไม่สามารถนำไปใช้หรือ
 ประกันได้ว่าเป็นค่าพารามิเตอร์ที่แท้จริง

3. บางครั้งค่าพารามิเตอร์หรือค่าที่ได้จากการประมาณไม่ตกอยู่ในขอบเขต
 ของค่าพารามิเตอร์ กล่าวคือ อาจมีค่าใดค่าหนึ่งอยู่ภายนอกขอบเขตที่ยอมรับได้ ในกรณีเช่นนี้
 ต้องมีการกำหนดขอบเขตจำกัดของค่าประมาณไว้ เพื่อให้ค่าประมาณที่ได้ไม่สูงหรือต่ำเกินไป
 นัก แต่การกระทำเช่นนี้เป็นจุดอ่อนของกระบวนการประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ โดย
 เฉพาะในแบบจำลอง 2 และ 3 พารามิเตอร์ จึงทำให้เกิดปัญหาตามมาเกี่ยวกับความเที่ยง
 ตรง (Validity) ของค่าที่ประมาณได้

4. เนื่องจากการประมาณค่าพารามิเตอร์ในแบบจำลอง 3 พารามิเตอร์
 มีสมการหลายสมการที่ต้องหารากที่ทำให้ฟังก์ชันไลค์ลิสต์มีค่าสูงสุด ด้วยวิธีของนิวตัน ราวสัน
 จึงจำเป็นต้องใช้เครื่องคอมพิวเตอร์ขนาดใหญ่ในการทำงาน

ตาราง 2 ค่าฟังก์ชันสารสนเทศสำหรับค่าพารามิเตอร์ข้อสอบและค่าความสามารถของผู้เข้าสอบในแบบจำลองโลจิสติก 3 พารามิเตอร์

พารามิเตอร์ (Parameter)	ฟังก์ชันสารสนเทศ (Information Function)
θ_i	$D^2 \sum_{j=1}^n \frac{a_i^2 (P_{ij} - c_i) Q_{ij} (c_i - P_{ij})}{P_{ij} (1 - c_i)^2}$
a_i	$\frac{D^2}{(1 - c_i)^2} \sum_{j=1}^n \frac{(\theta_i - b_i)^2 (P_{ij} - c_i)^2 Q_{ij}}{P_{ij}}$
b_i	$\frac{D^2 a_i^2}{(1 - c_i)^2} \sum_{j=1}^n \frac{(\theta_i - b_i) (P_{ij} - c_i)^2 Q_{ij}}{P_{ij}}$
c_i	$\frac{D}{(1 - c_i)^2} \sum_{j=1}^n \frac{(\theta_i - b_i) (P_{ij} - c_i) Q_{ij}}{P_{ij}}$

วิธีมาร์จินัล แมกซิมัม ไลค์ลิฮูด (Marginal Maximum Likelihood

Procedure) (Mislevy and Bock. 1990 : 6 - 8)

การประมาณค่าพารามิเตอร์ด้วยวิธีมาร์จินัล แมกซิมัม ไลค์ลิฮูด นี้ สามารถใช้ได้กับแบบทดสอบที่มีจำนวนข้อมากและน้อย วิธีการประมาณค่าเริ่มจากรูปแบบการตอบข้อสอบของผู้สอบ ดังนี้

$$X = (x_1, x_2, \dots, x_n)$$

และสำหรับผู้สอบที่มีความสามารถ θ ค่าความเป็นไปได้สามารถแสดงได้ดังนี้

$$P(X|\theta) = \prod_{i=1}^n P_i(\theta)^{x_i} [1-P_i(\theta)]^{1-x_i}$$

ค่าความน่าจะเป็นที่ได้เป็นความน่าจะเป็นที่มีเงื่อนไขว่าทราบค่าความสามารถ θ ซึ่งต่างจากความน่าจะเป็นของรูปแบบ x จากผู้ไม่ทราบค่าความสามารถ θ ซึ่งสุ่มมาจากประชากรที่มีการกระจายของค่าความสามารถ θ อยู่ในรูปฟังก์ชันความหนาแน่นต่อเนื่อง (Continuous Density $g(\theta)$) ที่อยู่ในรูปของความน่าจะเป็นที่ไม่มีเงื่อนไขและสามารถแสดงได้ดังนี้

$$P(X) = \int_{-\infty}^{+\infty} P(X|\theta) g(\theta) d\theta$$

ค่านี้เรียกว่า ความน่าจะเป็นโดยปราศจากเงื่อนไขของ X และเนื่องจากค่าความสามารถ θ ถูก Integrate ออกไป ค่านี้จึงเป็นฟังก์ชันของพารามิเตอร์ของข้อสอบเท่านั้น ในวิธีมาร์จินัล แมกซิมัม ไลค์ลิฮูด ค่าของพารามิเตอร์ของข้อสอบจะพิจารณาใช้ค่าที่ทำให้สมการข้างล่างมีค่าสูงสุด

$$\log L_M = \sum_{i=1}^S r_i \log \bar{P}(X_i)$$

เมื่อ $P(X_i)$ ประมาณได้จากสูตร Gaussian Quadrature

r_i คือ ความถี่ของรูปแบบการตอบข้อสอบรูปแบบ X_i จากจำนวนผู้สอบทั้งหมด (N) และ

S คือ จำนวนรูปแบบที่เป็นไปได้ทั้งหมดจากข้อมูลการตอบข้อสอบ

ในแบบจำลองโลจิสติก 3 พารามิเตอร์ ค่าสูงสุดของสมการดังกล่าวสำหรับข้อ i หาได้จากสมการไลค์ลิฮูด ดังนี้

$$\sum_{k=1}^s \frac{\bar{r}_{ik} - \bar{N}_k P_i(X_k)}{P_i(X_k)[1-P_i(X_k)]} \frac{d P_i(X_k)}{d \begin{bmatrix} c_i \\ a_i \\ b_i \end{bmatrix}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{เมื่อ } \bar{r}_{ik} = \sum_{j=1}^s r_{ij} X_{kj} P(X_j/X_k) A(X_k)/\bar{P}(X_j)$$

$$\bar{N}_k = \sum_{j=1}^s r_{ij} P(X_j/X_k) A(X_k)/\bar{P}(X_j)$$

เป็นค่าคาดหวังภายหลังของจำนวนข้อสอบที่ถูก และจำนวนข้อสอบที่ทำที่จุด X_k (X_{ij} คือ คะแนน 0, 1 สำหรับข้อสอบ i ในรูปแบบการตอบ l)

จากนั้นใช้ขั้นตอนการคำนวณ EM และวิธีการของ Newton - Gauss (Fisher-Scoring) แก้สมการ

ปัจจุบันในการคำนวณเพื่อประมาณค่าพารามิเตอร์ด้วยวิธีมาร์จินัล แมกซิมัม ไลค์ลิตูด สามารถทำได้โดยสะดวก เนื่องจาก มีสเลวี และ บ็อค (Mislevy and Bock. 1990 : 6 - 8) ได้พัฒนาโปรแกรมคำสั่งชื่อ BILOG 3 สำหรับคำนวณไว้เรียบร้อยแล้ว

7.2 วิธีฮิวริสติก (Heuristic Estimation)

วิธีฮิวริสติกเป็นวิธีประมาณค่าพารามิเตอร์ที่ฮูรีเป็นผู้สนับสนุน (Urry. 1974 : 256 - 260) ซึ่งแนวคิดของวิธีนี้มีอยู่ว่า ค่าพารามิเตอร์ความยาก และค่าพารามิเตอร์อำนาจจำแนกของทฤษฎีการตอบข้อคำถามสามารถประมาณ (Approximate) ได้จากค่าอำนาจจำแนกที่คำนวณได้จากสหสัมพันธ์ไบซีเรียล หรือสหสัมพันธ์พอยท์ไบซีเรียล และค่าความยากที่คำนวณได้จากสัดส่วนของจำนวนผู้ตอบถูกต้องผู้ตอบทั้งหมด ที่ได้จากทฤษฎีการทดสอบแบบดั้งเดิม ส่วนการนำวิธีฮิวริสติกไปใช้ให้มีประสิทธิภาพนั้น แบบทดสอบควรประกอบด้วยข้อสอบอย่างน้อย จำนวน 80 ข้อ และมีความเชื่อมั่น (KR-20) ประมาณ 0.90

สำหรับกระบวนการที่ใช้ในการประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีฮิวริสติกนี้ เป็นกระบวนการอย่างง่าย ไม่ต้องมีการทำซ้ำ โดยมีสมการดังนี้ (Urry.

1974 : 256 - 259; Schmidt. 1977 : 613 - 619; Warm. 1978 : 51 - 53;
Hambleton and Swaminathan. 1985 : 146; citing Jensema. 1976)

$$a_i = \frac{D_i(P_i, Q_i)^{1/2}}{[(KR-20)(1-C_i)^2 Y_i - D_i^2 P_i Q_i]^{1/2}} ; i=1,2,\dots,n \text{ ข้อ}$$

$$b_i = \frac{Y_i Z_i (1-C_i)(KR-20)^{1/2}}{D_i(P_i, Q_i)^{1/2}} ; i=1,2,\dots,n \text{ ข้อ}$$

c_i = สัดส่วนของผู้ตอบถูกที่อยู่ปลายล่างของโค้งที่เกิดจากการพล็อตจำนวน
คนตอบถูกสะสมในแต่ละระดับคะแนนรวม หลังจากหักคะแนนของข้อที่
กำลังหาค่าพารามิเตอร์ออกจากคะแนนรวมทั้งฉบับก่อน

เมื่อ a_i = ค่าอำนาจจำแนกของข้อที่ i

b_i = ค่าความยากของข้อที่ i

c_i = ค่าการเดาของข้อที่ i

C_i = ค่าการเดาของข้อที่ i ในทฤษฎีการทดสอบแบบดั้งเดิม (สัดส่วนกลับ
ของจำนวนตัวเลือกในข้อที่ i)

D_i = ค่าอำนาจจำแนกของข้อที่ i ในทฤษฎีการทดสอบแบบดั้งเดิม

$$(\text{Point Biserial Correlation} = \frac{\bar{X}_x - \bar{X}_c}{SD_x} \left[\frac{P_i}{Q_i} \right]^{1/2})$$

เมื่อ P_i' = ค่าความยากของข้อที่ i ในทฤษฎีการทดสอบแบบดั้งเดิม

P_i = $(P_i' - C_i) / (1 - C_i)$

Q_i' = $1 - P_i'$

Q_i = $1 - P_i$

Z_i = ค่าของคะแนนมาตรฐานที่มีพื้นที่ใต้โค้งปกติด้านขวามือเท่ากับ P_i

Y_i = ค่าพิกัดฉาก (Ordinate) ของโค้งที่มีการแจกแจงปกติ ณ จุด Z_i

KR-20 = ค่าความเชื่อมั่นของแบบทดสอบที่คำนวณโดยใช้สูตรครุเดอร์-ริชาร์ดสัน
สูตรที่ 20

ซึ่งค่าพารามิเตอร์ของข้อสอบที่ประมาณได้ด้วยวิธีอีวีเอสติดเกี่ยวข้องกับแบบจำลองโลจิสติก 3 พารามิเตอร์ โดยสามารถเขียนเป็นรูปแบบความสัมพันธ์เชิงคณิตศาสตร์ของค่าพารามิเตอร์ได้ดังนี้

$$P_i(\theta_{i,j}) = c_i + (1 - c_i) / (1 + e^{-D a_i (\theta_{i,j} - b_i)})$$

ส่วนการประมาณค่าความสามารถของผู้สอบใช้วิธีแมกซิมัม ไลค์ลิฮูด แบบมีเงื่อนไข (Conditional Maximum Likelihood Estimation of Ability) ซึ่งเป็นการประมาณค่าความสามารถหลังจากทราบค่าพารามิเตอร์ของข้อสอบแต่ละข้อแล้ว โดยใช้กระบวนการนิวตัน ราฟสัน และกำหนดค่าความสามารถเริ่มต้น สำหรับผู้สอบคนที่ a เป็น

$$\theta_{0,j} = \ln [r_{.j} / (n - r_{.j})] = \theta_0$$

$$\begin{aligned} \text{เมื่อ } r_{.j} &= \text{จำนวนข้อสอบที่ผู้เข้าสอบคนที่ } j \text{ ตอบถูก} \\ n &= \text{จำนวนข้อสอบที่ผู้เข้าสอบคนที่ } j \text{ ทำทั้งหมด} \\ \ln &= \text{Natural Logarithm} \end{aligned}$$

จะได้ค่า θ ของการทำซ้ำครั้งที่ $m+1$ ดังสมการ

$$\theta_{.j}^{(m+1)} = \theta_{.j}^{(m)} - h_{.j}^{(m)}$$

$$\text{โดยที่ } h_{.j}^{(m)} = \left[\frac{d \ln L}{d \theta_{.j}} \right] \div \left[\frac{d^2 \ln L}{d \theta_{.j}^2} \right]$$

$$\frac{d \ln L}{d \theta_{.j}} = \sum_{i=1}^n a_i (U_{i,j} - P_{i,j}) (P_{i,j} - c_i) / P_{i,j} (1 - c_i)$$

$$\frac{d^2 \ln L}{d \theta_{.j}^2} = \sum_{i=1}^n a_i^2 (P_{i,j} - c_i) (U_{i,j} - P_{i,j})^2 Q_{i,j} / P_{i,j}^2 (1 - c_i)^2$$

เมื่อ	$U_{i,j}$	=	ค่าผลการตอบ (ถูกเป็น 1 ผิดเป็น 0) ของผู้เข้าสอบคนที่ j
	$P_{i,j}$	=	โอกาสที่ผู้เข้าสอบคนที่ j ทำข้อสอบข้อ i ถูกต้อง
		=	$c_i + (1 - c_i) / (1 + e^{-D a_i (e_j - b_i)})$
	D	=	A Scaling Factor มีค่าเท่ากับ 1.7
	e_j	=	ค่าความสามารถของผู้เข้าสอบคนที่ j
	a_i	=	ค่าอำนาจจำแนกของข้อสอบข้อที่ i ที่ประมาณโดยวิธีอีวีรติค
	b_i	=	ค่าความยากของข้อสอบข้อที่ i ที่ประมาณโดยวิธีอีวีรติค
	c_i	=	ค่าการเดาของข้อสอบข้อที่ i ที่ประมาณโดยวิธีอีวีรติค

และเมื่อ $|h_{i,j}^{(m)}|$ มีค่าน้อยกว่าหรือเท่ากับ 0.001 จึงสิ้นสุดกระบวนการ (Hambleton and Swaminathan. 1985 : 81 - 88)

หมายเหตุ

1. หากต้องการทำการประมาณค่าความสามารถของผู้สอบในแบบจำลอง 2 พารามิเตอร์ ให้กำหนดค่า $c_i = 0$
2. หากต้องการทำการประมาณค่าความสามารถของผู้สอบในแบบจำลอง 1 พารามิเตอร์ ให้กำหนดค่า $c_i = 0$ และ $a_i = 1$

ในการคำนวณเพื่อประมาณค่าพารามิเตอร์ด้วยวิธีอีวีรติคนี้ ผู้วิจัยได้พัฒนาโปรแกรมคอมพิวเตอร์ TURBO PASCAL Version 7.0 สำหรับการคำนวณค่าพารามิเตอร์

7.3 วิธีของเบส์ (Bayesian Estimation)

เนื่องจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม โลคัลลิซูด มีข้อจำกัดและปัญหาเมื่อทำการประมาณค่าพารามิเตอร์ของข้อสอบ และค่าความสามารถของผู้เข้าสอบไปพร้อม ๆ กัน ดังกล่าวแล้ว วิธีของเบส์จึงอาจเป็นวิธีที่เหมาะสมกว่าทั้งนี้เพราะวิธีของเบส์มีแนวคิดบางประการที่ต่างออกไปจากแนวคิดของวิธีแมกซิมัมโลคัลลิซูด และได้รับการพัฒนาให้สามารถประมาณค่า e และ a, b, c ร่วมกันได้อย่างมีประสิทธิภาพ (วิศุตา บัวคง. 2533 : 47; อ้างอิงมาจาก Swaminathan and Gifford. 1982) กล่าวคือ การประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์มีแนวคิดที่ว่า ค่าความสามารถ e และค่าพารามิเตอร์ของข้อสอบ a, b และ c เป็นตัวแปรสุ่ม (Random Variable) จากการแจกแจงที่แสดงได้

ด้วยฟังก์ชันความหนาแน่นร่วม (Joint Density Function) $f(\theta, b, a, c)$ และเรียกฟังก์ชัน $f(\theta, b, a, c)$ นี้ว่า การแจกแจงเริ่มแรก (Prior Distribution) ของค่า θ, b, a และ c ซึ่งทำให้การใช้ $L(U|\theta, b, a, c)$ เพียงอย่างเดียวในการประมาณค่า θ, b, a และ c ถูกพิจารณาว่า เป็นการใช้อำนาจที่น้อยอย่างไม่ครบถ้วน เพราะยังมีการแจกแจงเริ่มต้นร่วมกับ $f(\theta, b, a, c)$ ที่ควรนำมาใช้ในการประมาณค่าพารามิเตอร์ด้วย

จากการนิยามความน่าจะเป็นของผู้เข้าสอบในการตอบข้อสอบข้อที่ i เมื่อ $U_i = 1$ สำหรับการตอบถูก และ $U_i = 0$ สำหรับการตอบผิด แสดงได้ดังสมการ

$$\begin{aligned} P(U_i | \theta, b, a, c) &= P(U_i = 1 | \theta, b, a, c) P(U_i = 0 | \theta, b, a, c) \\ &= P_i^{U_i} Q_i^{1-U_i} \quad ; \quad Q_i = 1 - P_i \end{aligned}$$

ถ้าผู้เข้าสอบตอบข้อสอบ n ข้อ และข้อสอบมีความเป็นอิสระเฉพาะที่แล้ว ความน่าจะเป็นของการตอบสามารถแสดงได้ด้วยฟังก์ชันความหนาแน่นร่วม ดังสมการ

$$P(U_1, U_2, \dots, U_n | \theta, b, a, c) = \prod_{i=1}^n P_i^{U_i} Q_i^{1-U_i}$$

จากสมการดังกล่าวข้างต้น เป็นความน่าจะเป็นของการตอบข้อสอบ n ข้อ ที่สามารถวัดหรือสังเกตได้ โดยที่ $U_1, U_2, \dots, U_n$ เป็นตัวแปรสุ่มที่มีค่าเฉพาะเป็น $u_1, u_2, \dots, u_n$ เมื่อ u_i มีค่าเท่ากับ 1 หรือ 0 และเนื่องจากสมการนี้เป็นฟังก์ชันทางคณิตศาสตร์ของค่า θ, b, a, c ที่จะบอกว่าตัวแปรสุ่มนี้มีโอกาสเกิดขึ้นเพียงใด จึงเรียกฟังก์ชันนี้ว่า ฟังก์ชันความน่าจะเป็น หรือฟังก์ชันไลค์ลิฮูด (Likelihood Function) ซึ่งแสดงได้ดังนี้

$$L(u_1, u_2, \dots, u_n | \theta, b, a, c) = \prod_{i=1}^n P_i^{u_i} Q_i^{1-u_i}$$

เมื่อ $u_i = 1$ ค่าของ Q_i จะหมดไป และเมื่อ $u_i = 0$ ค่าของ P_i ก็จะหมดไป และฟังก์ชันไลค์ลิฮูด ที่มีผู้เข้าสอบ N คน ตอบข้อสอบ n ข้อ มีสมการดังนี้

$$L(u|e, b, a, c) = \prod_{j=1}^m \prod_{i=1}^n P_{i,j}^{u_i} Q_{i,j}^{1-u_i}$$

$$\begin{aligned} \text{เมื่อ } u &= \text{เวกเตอร์ผลการตอบข้อสอบ } n \text{ ข้อ ของผู้เข้าสอบ } N \text{ คน} \\ P_{i,j} &= P_i(e_j, b_j, a_j, c_j) \end{aligned}$$

ดังนั้น ถ้าพิจารณาความน่าจะเป็นร่วมของการตอบข้อสอบ $P(U|e, b, a, c)$ แล้ว จะเห็นว่าการแจกแจงของตัวแปร U ขึ้นอยู่กับค่าความสามารถ e และค่าพารามิเตอร์ของข้อสอบ a, b และ c ถ้าค่า e, a, b และ c เปลี่ยนไป โอกาสที่ตัวแปร U มีค่าเท่ากับ u ก็จะเปลี่ยนไปด้วย ดังนั้น การทราบผลการตอบข้อสอบ u จึงน่าจะช่วยให้ทราบค่า e, a, b และ c ได้ดีขึ้นขึ้น ซึ่งอาจแสดงได้ด้วยการแจกแจงอย่างมีเงื่อนไขของค่า e, b, a และ c เมื่อทราบผลการตอบข้อสอบ $f(e, b, a, c|u)$ และเรียกฟังก์ชันนี้ว่าการแจกแจงภายหลัง (Posterior Distribution)

การแจกแจงภายหลังร่วมกันของค่าความสามารถ e และค่าพารามิเตอร์ของข้อสอบ a, b และ c เมื่อทราบผลการตอบข้อสอบ u ก็คือ

$$f(e, b, a, c|u) = L(u|e, b, a, c) f(e, b, a, c) / f(u)$$

$$\text{เมื่อ } f(u) = \text{การแจกแจงมาร์จินัล (Marginal) ของผลการตอบข้อสอบ}$$

$$f(e, b, a, c) = \text{การแจกแจงเริ่มแรกของค่า } e, b, a \text{ และ } c$$

$$f(u|e, b, a, c) = \text{ฟังก์ชันไลค์ลิฮูด}$$

$$f(e, b, a, c|u) = \text{การแจกแจงภายหลังของค่า } e, b, a \text{ และ } c \text{ เมื่อทราบผลการตอบข้อสอบ } u$$

ซึ่งการแจกแจงภายหลังนี้เป็นฟังก์ชันที่ใช้ในการประมาณค่า e, b, a และ c ด้วยวิธีของเบย์ โดยมีส่วนแตกต่างจากการประมาณด้วยวิธีแบบทิมมส์-ไลค์ลิฮูด คือ การแจกแจงเริ่มแรก $f(e, b, a, c)$ และการแจกแจงมาร์จินัล $f(u)$ โดยการแจกแจงมาร์จินัลนี้เป็นการแจกแจง

ที่ไม่ขึ้นอยู่กับค่า θ , b , a และ c จึงถือว่าเป็นค่าคงที่ในการประมาณค่า θ , b , a และ c

จากแนวคิดของการประมาณค่าด้วยวิธีของเบย์ดังกล่าว ทำให้สามารถจำแนกกระบวนการดำเนินการตามแนวคิดของเบย์ออกเป็น 2 กระบวนการ ดังนี้

1. กระบวนการกำหนดลักษณะของการแจกแจงเริ่มแรก มีอยู่ 2 ชั้น คือ (Swaminathan and Gifford 1986. : 589 - 601)

1.1 กำหนดให้การแจกแจงเริ่มแรกของค่าความสามารถ θ ค่าอำนาจจำแนก a ค่าความยาก b และค่าการเดา c เป็นอิสระต่อกัน ดังนี้

$$f(\theta, b, a, c) = f(\theta) f(b) f(a) f(c)$$

โดยกำหนดลักษณะการแจกแจงของ $f(\theta)$, $f(b)$, $f(a)$ และ $f(c)$ ไว้ดังนี้

1.1.1 การแจกแจงเริ่มแรกของค่าความสามารถ $f(\theta)$ มีข้อตกลงว่า ข้อสารสนเทศที่มีมาก่อนของค่าความสามารถของผู้เข้าสอบแต่ละคนไม่แตกต่างกัน สามารถใช้แทนกันได้ (Exchangeability) และค่าความสามารถเป็นตัวแปรสุ่มที่มีการแจกแจงเป็นปกติ (Normal Distribution)

$$f(\theta_i | \mu_-, \sigma_-^2) = N(\mu_-, \sigma_-^2)$$

เมื่อ $N(\mu_-, \sigma_-^2) =$ การแจกแจงปกติ (Normal Distribution) ที่มีค่าเฉลี่ยเท่ากับ μ_- และค่าความแปรปรวนเท่ากับ σ_-^2

1.1.2 การแจกแจงเริ่มแรกของค่าความยากของข้อสอบ $f(b)$ อาจใช้กระบวนการเดียวกันกับการกำหนดการแจกแจงเริ่มแรกของค่าความสามารถ คือ มีข้อตกลงว่า $f(b)$ มีการแจกแจงเป็นปกติ หรืออาจจะไม่กำหนดการแจกแจงเริ่มแรกไว้ก็ได้ (Swaminathan and Gifford. 1985 : 35)

1.1.3 การแจกแจงเริ่มแรกของค่าอำนาจจำแนกของข้อสอบ $f(a)$ ควรเป็นการแจกแจงแบบไคว์ (Chi Distribution) เนื่องจากค่าอำนาจจำแนกของข้อสอบโดยทั่วไปจะต้องเป็นค่าบวก และเป็นความชันของเส้นโค้งลักษณะของข้อสอบ ณ จุดเปลี่ยนโค้ง

$$f(a_i | v_1, w_1) \propto a_i^{v_1 - 1} \exp[-a_i^2 / 2w_1]$$

เมื่อ v_1 = Degrees of Freedom
 w_1 = Scale Parameter

1.1.4 การแจกแจงเริ่มแรกของค่าการเดาของข้อสอบ $f(c)$ ควรมีการแจกแจงแบบเบต้า (Beta Distribution) เนื่องจากค่าพารามิเตอร์ c_i มีขอบเขตอยู่ตั้งแต่ 0 - 1

$$f(c_i | s_1, t_1) \propto c_i^{s_1 - 1} (1 - c_i)^{t_1 - 1}$$

เมื่อ s_1, t_1 = Scale Parameter

1.2 กำหนดค่าที่เป็นตัวเลขของพารามิเตอร์สำหรับการแจกแจงเริ่มแรก

1.2.1 พารามิเตอร์ของการแจกแจงเริ่มแรกของค่าความสามารถ θ ได้แก่ μ_θ และ σ_θ^2 อาจกำหนดให้ $\mu_\theta = 0$ และ $\sigma_\theta^2 = 1$ ซึ่งจะทำให้การประมาณค่าความสามารถ θ มีความสะดวก และประมาณค่าได้รวดเร็วขึ้น (Swaminathan and Gifford. 1985 : 351 - 355)

1.2.2 พารามิเตอร์ของการแจกแจงเริ่มแรกของค่าความยาก b หากมีการกำหนดลักษณะการแจกแจงไว้ ก็ใช้ค่าเดียวกับพารามิเตอร์ของการแจกแจงเริ่มแรกของค่าความสามารถ θ ดังกล่าวแล้วข้างต้น

1.2.3 พารามิเตอร์ของการแจกแจงเริ่มแรกของค่าอำนาจจำแนก a ได้แก่ v_1 และ w_1 การกำหนดค่าของ v_1 และ w_1 ที่เหมาะสม อาจจะหาได้

จากการกำหนดพิสัย (Range) ของค่า a คือ ถ้าให้ H เป็นขีดจำกัดบนของพิสัย และให้ L เป็นขีดจำกัดล่างของพิสัย จะหาค่า v_1 และ w_1 ได้จากสูตร

$$v_1 = 1/2 (1 + Z_{(1/2)\alpha} ((H+L)/(H-L))^2)$$

$$w_1 = 1/2 ((H-L)/Z_{(1/2)\alpha})^2$$

เมื่อ $Z_{(1/2)\alpha}$ = ค่า Z ของการแจกแจงปกติมาตรฐาน Standard Normal Distribution) ที่ระดับนัยสำคัญ α

$$v_1 = \text{Degrees of Freedom}$$

นอกจากวิธีดังกล่าว อาจกำหนดให้ v_1 และ w_1 ของการแจกแจงเริ่มแรก $f(a)$ ของข้อสอบทุกข้อเท่ากัน คือ $v_1 = 10$ และ $w_1 = 0.1$ ซึ่งจะทำให้การประมาณค่าพารามิเตอร์มีความสะดวกยิ่งขึ้น (Swaminathan and Gifford. 1985 : 356 - 357) และการกำหนดเช่นนี้ก็ยังคงทำให้ค่าประมาณของพารามิเตอร์ a_1 ตกอยู่ในช่วงที่ยอมรับได้ คือ $0.40 < a_1 < 1.55$ ด้วยความเชื่อมั่น 99% (Swaminathan and Gifford. 1985 : 356)

1.2.4 พารามิเตอร์ของการแจกแจงเริ่มแรกของค่าการเดา c ได้แก่ s_1 และ t_1 การกำหนดค่าของ s_1 และ t_1 ที่เหมาะสมอาจหาได้จากการสังเกตสัดส่วนการตอบถูกของกลุ่มผู้เข้าสอบที่มีความสามารถในระดับต่ำมาก กล่าวคือ ถ้าให้ m แทนจำนวนผู้เข้าสอบที่มีระดับความสามารถต่ำมาก และให้ M แทน สัดส่วนการตอบถูกของผู้เข้าสอบกลุ่ม m แล้ว สามารถหาค่า s_1 และ t_1 ได้จากสูตร

$$s_1 = m M$$

$$\text{และ } t_1 = m (1 - M) - 2$$

นอกจากวิธีนี้ อาจกำหนดให้ s_1 และ t_1 ของการแจกแจงเริ่มแรก $f(c)$ ของข้อสอบทุกข้อเท่ากัน คือ $s_1 = 2$ และ $t_1 = 12$ ซึ่งจะทำให้การประมาณค่าพารามิเตอร์มีความ

สะดวกยิ่งขึ้น (Swaminathan and Gifford. 1986 : 589 - 601) และการกำหนด
เช่นนี้ ก็ยังคงทำให้ค่าประมาณของค่าพารามิเตอร์ c_i ตกอยู่ในช่วงที่ยอมรับได้ คือ 0.026
 $< c_i < 0.317$ ด้วยความเชื่อมั่น 99% (วิชุดา บัวคง. 2533 : 52; อ้างอิงมาจาก
Novick and Jackson. 1974 : 402)

2. กระบวนการประมาณค่าพารามิเตอร์ข้อสอบและความสามารถของผู้เข้า
สอบ

การประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ คือ การหาค่าประมาณ
 θ_j ($j = 1, 2, \dots, N$) และ b_i, a_i, c_i ($i = 1, 2, \dots, n$) ที่ทำให้ฟังก์ชันการ
แจกแจงภายหลัง $f(\theta, b, a, c|u)$ มีค่าสูงสุด ปกติจะหาค่า θ_j, b_i, a_i และ c_i ที่
ทำให้ $\ln f(\theta, b, a, c|u)$ มีค่าสูงสุด ถ้า $\ln f(\theta, b, a, c|u)$ เป็นฟังก์ชันที่หาอนุพันธ์
ได้ การหาค่า θ_j, b_i, a_i และ c_i จะหาได้จากการอนุพันธ์ของ $\ln f(\theta, b, a, c|u)$
และกำหนดให้อนุพันธ์ของ $\ln f(\theta, b, a, c|u)$ มีค่าเท่ากับศูนย์ แล้วจึงหาค่ารากของ
อนุพันธ์ของ $\ln f(\theta, b, a, c|u)$ ซึ่งแสดงได้ดังสมการต่อไปนี้

$$\begin{aligned} f(\theta, b, a, c|u) &= L(u|\theta, b, a, c) f(\theta) f(b) f(a) f(c)/f(u) \\ \ln f(\theta, b, a, c|u) &= \ln L(u|\theta, b, a, c) + \ln f(\theta) + \ln f(b) \\ &\quad + \ln f(a) + \ln f(c) + \text{constant} \\ d \ln f(\theta, b, a, c|u)/d \theta_j &= 0 \\ d \ln f(\theta, b, a, c|u)/d b_i &= 0 \\ d \ln f(\theta, b, a, c|u)/d a_i &= 0 \\ d \ln f(\theta, b, a, c|u)/d c_i &= 0 \end{aligned}$$

สมการอนุพันธ์ของ $\ln f(\theta, b, a, c|u) = 0$ นี้ เรียกว่า สมการโมดัล (Modal
Equation) ค่ารากของสมการโมดัล คือ ค่าความสามารถ θ และค่าพารามิเตอร์ของ
ข้อสอบ a, b, c ที่ทำให้ฟังก์ชันการแจกแจงภายหลังมีค่าสูงสุด อาจทำได้โดยใช้เทคนิคนิวตัน
ราฟสัน ซึ่งเป็นการหาค่าประมาณโดยการหาค่าซ้ำ (Iterative) ดังมีขั้นตอนดังต่อไปนี้

ขั้นที่ 1 กำหนดค่าเริ่มต้นสำหรับใช้ในการประมาณค่าความสามารถ θ_j
และค่าพารามิเตอร์ของข้อสอบ a_i, b_i และ c_i

1. ค่าเริ่มต้นสำหรับใช้ในการประมาณค่าความสามารถ e_j คือ

$$e_j^{(0)} = \ln (X_j / (n - X_j))$$

เมื่อ $\ln$ = Natural Logarithm

X_j = คะแนนสอบของผู้เข้าสอบคนที่ j

n = จำนวนข้อสอบ

2. ค่าเริ่มต้นสำหรับใช้ในการประมาณค่าพารามิเตอร์ของข้อสอบ a_j , b_j และ c_j คือ

$$a_j^{(0)} = R_j / (1 - R_j^2)^{1/2}$$

$$b_j^{(0)} = Z_j / R_j$$

$$c_j^{(0)} = 1 / m_j$$

เมื่อ R_j = Point - Biserial Correlation

Z_j = ค่า Z ของการแจกแจงปกติมาตรฐานที่พื้นที่ใต้โค้งปกติมาตรฐานด้านขวามีค่าเท่ากับ P_j

$$P_j = \frac{\sum_{i=1}^N U_{ij} / N}{N}$$

N = จำนวนผู้เข้าสอบทั้งหมด

m_j = จำนวนตัวเลือกในข้อสอบข้อที่ j

ขั้นที่ 2 ประมาณค่าความสามารถของผู้เข้าสอบแต่ละคน e_j โดยกำหนดให้ค่าพารามิเตอร์ของข้อสอบที่ประมาณค่าได้ในครั้งก่อนเป็นค่าคงที่

$$e_j^{(m+1)} = e_j^{(m)} - g(e_j^{(m)}) / h(e_j^{(m)})$$

เมื่อ $e_j^{(m)}$, $e_j^{(m+1)}$ = ค่าประมาณความสามารถของคนที่ j ครั้งที่ m และ $m+1$

$$g(e_j^{(m)}) = d \ln f(e, b, a, c | u) / d e_j$$

$$h(e_i^{(m)}) = d^2 \ln f(e, b, a, c|u) / d e_i^2$$

ประมาณค่าซ้ำ (Iterative) จนกว่าค่าประมาณความสามารถ e_i จะเข้าสู่ค่าคงที่ค่าใดค่าหนึ่ง (Convergence) คือ ค่าประมาณครั้งที่ $m+1$ และครั้งที่ m มีค่าแตกต่างกันน้อยกว่าค่าคงที่ที่กำหนดไว้ เช่น 0.001

ขั้นที่ 3 ประมาณค่าพารามิเตอร์ของข้อสอบแต่ละข้อ a_i , b_i และ c_i โดยกำหนดให้ค่าความสามารถของผู้เข้าสอบที่ประมาณค่าได้ในครั้งก่อนเป็นค่าคงที่

$$a_i^{(m+1)} = a_i^{(m)} - g(a_i^{(m)}) / h(a_i^{(m)})$$

$$b_i^{(m+1)} = b_i^{(m)} - g(b_i^{(m)}) / h(b_i^{(m)})$$

$$c_i^{(m+1)} = c_i^{(m)} - g(c_i^{(m)}) / h(c_i^{(m)})$$

เมื่อ $a_i^{(m)}, a_i^{(m+1)} =$ ค่าประมาณอำนาจจำแนกของข้อสอบข้อที่ i ครั้งที่ m และ $m+1$

$b_i^{(m)}, b_i^{(m+1)} =$ ค่าประมาณความยากของข้อสอบข้อที่ i ครั้งที่ m และ $m+1$

$c_i^{(m)}, c_i^{(m+1)} =$ ค่าประมาณการเดาของข้อสอบข้อที่ i ครั้งที่ m และ $m+1$

$$g(a_i^{(m)}) = d \ln f(e, b, a, c|u) / d a_i$$

$$h(a_i^{(m)}) = d^2 \ln f(e, b, a, c|u) / d a_i^2$$

$$g(b_i^{(m)}) = d \ln f(e, b, a, c|u) / d b_i$$

$$h(b_i^{(m)}) = d^2 \ln f(e, b, a, c|u) / d b_i^2$$

$$g(c_i^{(m)}) = d \ln f(e, b, a, c|u) / d c_i$$

$$h(c_i^{(m)}) = d^2 \ln f(e, b, a, c|u) / d c_i^2$$

ประมาณค่าซ้ำ (Iterative) จนกว่าค่าประมาณจะเข้าสู่ค่าคงที่ (Convergence)

สำหรับค่าของอนุพันธ์อันดับ 1 $g(.)$ และอนุพันธ์อันดับ 2 $h(.)$ มีส่วนประกอบสองส่วน คือ ส่วนที่ได้จากฟังก์ชันโลคัลลิสต์ และส่วนที่ได้จากการแจกแจงเริ่มแรก ซึ่งส่วนประกอบทั้งสองส่วนนี้แสดงไว้ในตาราง 3

ขั้นที่ 4 ประมาณค่าขั้นที่ 2 และขั้นที่ 3 จนกว่าค่าประมาณ θ_1, a_1, b_1 และ c_1 จะมีค่าคงที่ และมีความถูกต้องเพียงพอ หรือทำให้ $f(\theta, b, a, c|u)$ มีค่าสูงที่สุด

ในการคำนวณเพื่อประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์นี้ ผู้วิจัยได้พัฒนาโปรแกรมคอมพิวเตอร์ TURBO PASCAL Version 7.0 สำหรับการคำนวณค่าพารามิเตอร์

ตาราง 3 องค์พันธ์อันดับที่ 1 และองค์พันธ์อันดับที่ 2 ของ $\ln f(\theta, b, a, c|u)$

พารามิเตอร์ ที่มา	องค์พันธ์อันดับที่ 1	องค์พันธ์อันดับที่ 2
θ_1 Likelihood	$D \sum_{i=1}^n a_i (P_{i,1} - c_1) (U_{i,1} - P_{i,1}) / P_{i,1} (1 - c_1)$	$D^2 \sum_{i=1}^n [a_i^2 (P_{i,1} - c_1) Q_{i,1} / (1 - c_1)^2 P_{i,1}] [(U_{i,1} c_1 / P_{i,1}) - P_{i,1}]$
Prior	$- \theta_1$	-1
a_1 Likelihood	$D \sum_{i=1}^n (\theta_1 - b_i) (P_{i,1} - c_1) (U_{i,1} - P_{i,1}) / P_{i,1} (1 - c_1)$	$D^2 \sum_{i=1}^n (\theta_1 - b_i)^2 (P_{i,1} - c_1) (U_{i,1} c_1 - P_{i,1}^2) Q_{i,1} / P_{i,1} (1 - c_1)^2$
Prior	$+ (v_1 - 1) / a_1 - a_1 / w_1$	$- (v_1 - 1) / a_1 - 1 / w_1$
b_1 Likelihood	$D \sum_{i=1}^n a_i (P_{i,1} - c_1) (U_{i,1} - P_{i,1}) / P_{i,1} (1 - c_1)$	$D^2 \sum_{i=1}^n a_i^2 (P_{i,1} - c_1) (U_{i,1} c_1 - P_{i,1}^2) Q_{i,1} / P_{i,1}^2 (1 - c_1)^2$
Prior	$-$	$-$
c_1 Likelihood	$D \sum_{i=1}^n (U_{i,1} - P_{i,1}) / P_{i,1} (1 - c_1)$	$D^2 \sum_{i=1}^n [U_{i,1} (2P_{i,1} - 1) - P_{i,1}^2] / P_{i,1} (1 - c_1)^2$
Prior	$+ s_1 / c_1 - t_1 / (1 - c_1)$	$- s_1 / c_1 + t_1 / (1 - c_1)^2$

$$P_{i,1} = c_1 + (1 - c_1) / \{ (1 + \exp^{[-D a_i (\theta_1 - b_i)])} \}$$

จากการศึกษาวิธีประมาณค่าพารามิเตอร์ทั้ง 3 วิธี ดังกล่าวข้างต้น ผู้วิจัยได้สรุปความเหมือนและความแตกต่างกันของวิธีประมาณค่าพารามิเตอร์ทั้ง 3 วิธีไว้ ในตาราง 4 ดังนี้

ตาราง 4 สรุปความเหมือนและความแตกต่างของกระบวนการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีเอสดีค และวิธีของเบส์

การประมาณค่า	วิธีแมกซิมัม ไลค์ลิฮูด	วิธีอีวีเอสดีค	วิธีของเบส์
1. เป็นวิธีการประมาณค่าพารามิเตอร์ในแบบ แบบจำลองโลจิสติก 3 พารามิเตอร์	/	/	/
2. ใช้แนวคิดฟังก์ชันสารสนเทศของแบบทดสอบได้	/	/	/
3. ใช้การประมาณค่าแบบซ้ำ ๆ จนกว่าจะได้ค่า คงที่	/		/
4. ใช้ฟังก์ชันไลค์ลิฮูดในสมการประมาณค่า	/		/
5. ใช้ข้อมูลผลการตอบเพียงอย่างเดียวเพื่อ ประมาณค่าพารามิเตอร์	/		
6. ประมาณค่าพารามิเตอร์ของข้อสอบและความ สามารถของผู้เข้าสอบไปพร้อม ๆ กัน	/		/
7. ประมาณค่าความสามารถของบุคคลด้วยวิธี แมกซิมัม ไลค์ลิฮูด	/	/	
8. ประมาณค่า σ ด้วยวิธีแมกซิมัม ไลค์ลิฮูด อย่างมีเงื่อนไข		/	
9. ใช้แนวคิดทางสถิติแบบนอนเบส์เชียน	/	/	
10. ใช้แนวคิดทางสถิติแบบเบส์เชียน			/
11. มีการกำหนดขอบเขตของค่าพารามิเตอร์	/	/	

8. การศึกษาโดยใช้เทคนิคมอนติ คาร์โล (Monte Carlo Method)

8.1 ความเป็นมาของเทคนิคมอนติ คาร์โล

การจำลองสถานการณ์ (Simulation) เป็นวิธีการหนึ่งที่ใช้ในกระบวนการแก้ปัญหาในด้านต่าง ๆ ซึ่งมีมานานแล้ว แต่ที่ได้รับความสนใจและนำมาใช้ในการแก้ปัญหาในสาขาอาชีพต่าง ๆ อย่างแพร่หลายในปัจจุบันนี้ เนื่องจากความเจริญก้าวหน้าทางเทคโนโลยีคอมพิวเตอร์ สำหรับกระบวนการที่ใช้ในการจำลองสถานการณ์นั้นจะแบ่งออกเป็น 2 ส่วน คือ การสร้างแบบจำลองส่วนหนึ่ง และการนำเอาแบบจำลองนั้นไปใช้งานเชิงวิเคราะห์อีกส่วนหนึ่ง เพื่อประโยชน์ในการอธิบายพฤติกรรม และเพื่อการปรับปรุงการดำเนินงานของระบบงานจริง (ศิริจันทร์ ทองประเสริฐ. 2537 : 1)

ในการจำลองสถานการณ์ด้วยคอมพิวเตอร์จะต้องมีการคำนวณ ซึ่งมีข้อมูลทั้งที่เป็นข้อมูลนำเข้าและผลลัพธ์จากแบบจำลอง ดังนั้น การจัดเตรียมและการวิเคราะห์ข้อมูลรวมทั้งขั้นตอนต่าง ๆ ที่ใช้กับการจำลองสถานการณ์ จึงต้องอาศัยวิธีการต่าง ๆ ทางสถิติเข้ามาช่วย และวิธีการทางสถิติวิธีการหนึ่งซึ่งเป็นที่นิยมใช้กันมากและเกือบจะมีความจำเป็นในทุก ๆ การจำลองสถานการณ์ ก็คือ การสุ่มตัวอย่างด้วยเทคนิคมอนติ คาร์โล (Monte Carlo Sampling Technique) และเพื่อเน้นถึงความจำเป็นในการใช้เทคนิคดังกล่าว การจำลองสถานการณ์นี้จึงถูกเรียกว่า การจำลองสถานการณ์ด้วยเทคนิคมอนติ คาร์โล (Monte Carlo Simulation) ดังนั้น เทคนิคมอนติ คาร์โล จึงเป็นสาขาหนึ่งของคณิตศาสตร์ที่ใช้คอมพิวเตอร์ช่วยในการจำลองสถานการณ์ (Simulation) โดยอาศัยตัวเลขสุ่ม (Random Number) มาสร้างตัวแปรให้เหมือนกับสถานการณ์จริง และมีการทดลองซ้ำหลาย ๆ ครั้ง เพื่อให้ได้ค่าที่แน่นอนที่จะใช้เป็นข้อสรุปหรืออธิบายปรากฏการณ์ต่าง ๆ ในสถานการณ์จริง (Kerlinger. 1988 : 192 ; Krobe and Forge. 1980 : 242 ; Taha. 1988 : 13) หรือช่วยหาคำตอบในเรื่องราวต่าง ๆ ที่ยังไม่แน่ใจในผลที่จะเกิดขึ้น (Hammersley and Handscomb. 1964 : 2)

เทคนิคมอนติ คาร์โล มีการนำมาใช้นานแล้ว โดยในราวคริสต์ศตวรรษที่ 17 มีการนำเอาเทคนิคมอนติ คาร์โล ซึ่งในสมัยนั้นยังไม่ได้เรียกว่า มอนติ คาร์โล มาพัฒนาทฤษฎีความน่าจะเป็น (Probability Theory) (Barton. 1970 : 137) ประมาณปี ค.ศ. 1733 จอร์ส หลุยส์ เลคลเอร์ และบัพฟอง (Georges Louis Leclere and Comte de Buffon) ได้ทำการทดลองหาค่าพาย (π) โดยการโยนเข็มที่มีความยาว

ขนาด k หน่วย อย่างสุ่ม ลงมาบนพื้นราบที่มีเส้นขนานอยู่ โดยให้ระยะห่างระหว่างเส้นขนานแต่ละเส้นห่างกัน d หน่วย และกำหนดให้ $d > k$ จะได้ว่าความน่าจะเป็น (Probability : P) ที่เข็มจะตัดกับเส้นขนาน $P = 2k/\pi d$ ซึ่งถ้าความน่าจะเป็น (P) เป็นค่าสุ่ม ก็จะหาค่า π ได้ (Mihram. 1972 : 186; Rubinstein. 1981 : 11; บัญชา พนเจริญศักดิ์. 2527 : 700; เลิศลักษณ์ กลิ่นหอม. 2532 : 78) ในราวปี ค.ศ.1908 กอสเซต (W.S. Gosset) ได้ศึกษาการแจกแจงความถี่ของความสูงของนักโทษอาชญากรรมจำนวน 3,000 คน โดยเทียบกับการแจกแจงความถี่ของกลุ่มตัวอย่างที่สุ่มมาครั้งละ 4 คน จำนวน 750 กลุ่มตัวอย่าง ผลการศึกษา พบว่า การแจกแจงความถี่ทั้งสองมีลักษณะเหมือนกัน กอสเซตได้ตั้งชื่อการแจกแจงความถี่ที่ค้นพบนี้ว่า การแจกแจงค่าที (t-distribution) (Meyer. 1956 : 6-7; Spence. 1983 : 406) ทั้งหมดนี้ถือว่าเป็นจุดเริ่มต้นของการศึกษาที่ใช้เทคนิควิธีมอนติ คาร์โล (Monte Carlo Method)

สำหรับการพัฒนาระเบียบวิธีการมอนติ คาร์โล อย่างจริงจังนั้น เริ่มตั้งแต่ช่วงระหว่างสงครามโลกครั้งที่สอง คือ ในราวปี ค.ศ. 1944 อูลาม และวอน นิวแมน (Ulam and Von Neumann) เป็นผู้ตั้งชื่อ มอนติ คาร์โล ซึ่งเป็นชื่อรหัสลับของงานที่ทำใน ลอส อลามอส (Los Alamos) ซึ่งงานนี้เป็นงานที่เกี่ยวข้องกับการสร้างระเบิดปรมาณูในโครงการแมนฮัตตัน (Manhattan Project) โดยอูลาม (Ulam) ได้นำเทคนิควิธีมอนติ คาร์โล มาหาผลของการแพร่อย่างสุ่มของนิวตรอน (Neutron Diffusion) ในวัสดุเชื้อเพลิง ซึ่งเป็นการทดลองทางคณิตศาสตร์ เพื่อหาผลของคำตอบก่อนที่จะทำการทดลองจริง ซึ่งเป็นการลดอันตราย และช่วยประหยัดค่าใช้จ่ายก่อนการทดลองจริง (Spence. 1983 : 405; สมชาย สีนาน. 2528 : 29) หลังจากนั้นได้มีการนำเอาเทคนิควิธีมอนติ คาร์โล มาใช้อย่างกว้างขวางทั้งทางด้านฟิสิกส์ คณิตศาสตร์ สถิติ และการวิจัย เทคนิควิธีมอนติ คาร์โล นับว่ามีประโยชน์อย่างมากในการขยายความรู้ในเชิงทฤษฎี เช่น การนำมาศึกษาค่าความคลาดเคลื่อนทางสถิติ การเปรียบเทียบประสิทธิภาพของการทดสอบต่าง ๆ เป็นต้น

8.2 ขั้นตอนของระเบียบเทคนิควิธีมอนติ คาร์โล

หลักการสำคัญของเทคนิควิธีมอนติ คาร์โล (Monte Carlo Method) ก็คือการนำเอาตัวเลขสุ่ม (Random Number) มาประยุกต์ใช้ในการแก้ปัญหาต่าง ๆ ซึ่งมีขั้นตอนที่สำคัญ ดังนี้

1. สร้างตัวเลขสุ่ม (Generate Random Number) ในระยะแรก ๆ กระทำโดยการอาศัยเครื่องมือทางกายภาพ เช่น ลีธูเรต ลูกเต๋า ไพ่ กระดาษเขียนเบอร์ เป็นต้น เครื่องมือดังกล่าวใช้ได้เมื่อต้องการใช้ตัวเลขสุ่มจำนวนไม่มากนัก ต่อมาเมื่อมีความต้องการใช้ตัวเลขสุ่มจำนวนมาก ๆ ก็มีการหันมาใช้เครื่องมืออิเล็กทรอนิกส์ที่ให้แพร่หลายที่สุด คือ เครื่องสร้างตัวเลขแบบสุ่มที่สร้างขึ้นโดยบริษัทแรนด์ (RAND) ที่ได้ตัวเลขแบบสุ่มจากเครื่องกำเนิดพัลส์อิเล็กทรอนิกส์ (Electronic Pulse Generator) ซึ่งทำงานด้วยเสียง เครื่องดังกล่าวสามารถสร้างตัวเลขสุ่มได้เป็นล้านตัว ตัวเลขแบบสุ่มของบริษัท RAND นี้ มีการตีพิมพ์และลงไว้ในเทปแม่เหล็กเพื่อจำหน่าย

ในการสร้างหรือเลือกใช้ตัวเลขสุ่มกับเครื่องคอมพิวเตอร์นี้ มีปัญหา 2 ประการ คือ เป็นการยากที่จะทำให้คอมพิวเตอร์สามารถเรียกใช้ได้เมื่อมีความต้องการ และเป็นการยากที่จะทำให้เครื่องมือดังกล่าว สร้างตัวเลขสุ่มชุดเดิม เมื่อต้องการใช้เปรียบเทียบวิธีการต่าง ๆ ภายใต้เงื่อนไขของระบบเลขสุ่มชุดเดียวกัน หรือถ้าจะเก็บเลขสุ่มเหล่านี้ไว้ในหน่วยความจำหรือจานแม่เหล็ก ก็จะทำให้สูญเสียหน่วยความจำหรือเสียเวลาในการค้นหา ฉะนั้น การสร้างตัวเลขสุ่มในคอมพิวเตอร์จึงนิยมสร้างตัวเลขแบบสุ่มเทียม (Pseudo Random Number) โดยอาศัยสูตรทางคณิตศาสตร์ วิธีที่ใช้กันมากมี 2 วิธี คือ (ศิริจันทร์ ทองประเสริฐ. 2537 : 87 - 89)

วิธีที่ 1 วิธีส่วนกลางกำลังสอง (Midsquare Method) เป็นวิธีที่ใช้ในยุคแรก ๆ ของการสร้างตัวเลขสุ่มเทียม โดยมีขั้นตอนในการสร้างตัวเลขแบบสุ่ม ดังนี้ (ศิริจันทร์ ทองประเสริฐ. 2537 : 87 - 89)

1. เลือกตัวเลขขึ้นมาสี่หลัก
 2. ยกกำลังสองของตัวเลขนั้น ถ้าตัวเลขที่ได้ไม่ครบแปดหลักให้เติมศูนย์ข้างหน้าให้ครบแปดหลัก
 3. ใช้เลขสี่หลักกลางที่ได้ในขั้นที่ 2 เป็นตัวเลขแบบสุ่ม
 4. ยกกำลังสองของตัวเลขในข้อ 3
 5. ทำซ้ำในขั้นที่ 3 และ 4 จนได้จำนวนตัวเลขสุ่มตามต้องการ
- ตัวอย่าง สมมติว่าเลือกตัวเลขสุ่มตัวแรกเป็น $X_0 = 2712$ เป็นเลขสี่

หลักแรก

$$X_0 = 2712 \quad \text{จะได้} \quad X_0^2 = 07354944$$

$$X_1 = 3549 \quad \text{จะได้} \quad X_1^2 = 12595401$$

$$X_2 = 5954 \quad \text{จะได้} \quad X_2^2 = 35450116$$

$$X_3 = 4501 \quad \text{จะได้} \quad X_3^2 = 20259001$$

$$X_4 = 2590 \quad \text{จะได้} \quad X_4^2 = 06708100$$

$$X_5 = 7081 \quad \text{etc.}$$

ฉะนั้น ตัวเลขสุ่มที่ได้ คือ 2712, 3549, 5954, 4501, 2590, 7081, ... แต่การสร้างตัวเลขสุ่มโดยวิธีนี้ยังมีปัญหาสองประการ คือ ไม่ทราบช่วงของตัวเลขสุ่มที่จะหมุนกลับมาซ้ำเดิมอีก และถ้ากำหนดตัวเลขสุ่มตัวแรกไม่เหมาะสมจะทำให้ได้ตัวเลขสุ่มที่ไม่เป็นแบบสุ่ม เช่น ถ้ากำหนดตัวเลขสุ่มตัวแรกเป็น $X_0 = 4500$

$$X_0 = 4500 \quad \text{จะได้} \quad X_0^2 = 20250000$$

$$X_1 = 2500 \quad \text{จะได้} \quad X_1^2 = 06250000$$

$$X_2 = 2500 \quad \text{จะได้} \quad X_2^2 = 06250000$$

$$X_3 = 2500 \quad \text{จะได้} \quad X_3^2 = 06250000$$

$$X_4 = 2500 \quad \text{etc.}$$

ตัวเลขสุ่มที่ได้ คือ 4500, 2500, 2500, 2500, ... ซึ่งไม่เป็น

ตัวเลขแบบสุ่ม

วิธีที่ 2 วิธีเศษเหลือ (Congruent Method) วิธีเศษเหลือที่นิยมใช้กันมากที่สุด คือ เศษเหลือของผลคูณ (Multiplicative Congruential Method) ซึ่งมีสูตร ดังนี้

$$X_{i+1} = aX_i \pmod{m}$$

โดยที่ a และ m ต้องเป็นตัวเลขที่ไม่เป็นค่าลบ การสร้างตัวเลขแบบสุ่มนี้เริ่มต้นด้วยตัวเลขเริ่มต้น ตัวเลขตัวต่อไปจะได้จากการคูณด้วยค่าคงที่ a และหารด้วย m เศษเหลือจากการหาร คือ ตัวเลขที่ต้องการ

ตัวอย่าง สมมติว่าเลือก $a = 2$, $m = 10$ และ $X_0 = 1$

$$X_0 = 1$$

$$X_1 = 2 \times 1 \pmod{10} = 2$$

$$X_2 = 2 \times 2 \pmod{10} = 4$$

$$X_3 = 2 \times 4 \pmod{10} = 8$$

$$X_4 = 2 \times 8 \pmod{10} = 6$$

$$X_5 = 2 \times 6 \pmod{10} = 2$$

ตัวเลขสุ่มที่ได้ คือ 1, 2, 4, 8, 6, 2, ... ซึ่งจะเห็นว่าเลขสุ่มเริ่มวนกลับมาเลขเดิมแล้ว เพื่อไม่ให้เลขสุ่มวนกลับมาเลขเดิมเร็วเกินไป ในทางปฏิบัติเมื่อใช้คอมพิวเตอร์สร้างตัวเลขสุ่มโดยวิธีนี้ จึงนิยมกระทำดังนี้

1. เลือกตัวเลขสุ่มใด ๆ ที่น้อยกว่า 9 หลัก ใช้เป็นค่า X_0 โดยปกติถ้าเป็นคอมพิวเตอร์ฐานสอง X_0 จะให้เป็นเลขคี่ที่เป็นบวก และคอมพิวเตอร์ฐานสิบ X_0 จะให้เป็นเลขบวกที่หารด้วย 2 และ 5 ไม่ลงตัว

2. คูณเลขในข้อ 1 ด้วยค่า a ซึ่ง a นี้ ควรมีค่าไม่น้อยกว่า 5 หลัก โดยปกติค่า a จะหาได้จาก

$$a = 8T + 3 \dots\dots\dots \text{สำหรับคอมพิวเตอร์ฐานสอง}$$

$$a = 200T + Q \dots\dots\dots \text{สำหรับคอมพิวเตอร์ฐานสิบ}$$

เมื่อ T คือ เลขใด ๆ ที่เป็นบวก

Q คือ ค่าหนึ่งค่าใดต่อไปนี้ $\pm (3, 11, 13, 19, 21, 27, 29, 37, 53, 59, 61, 67, 69, 77, 83, \text{ หรือ } 91)$

3. คูณผลคูณในข้อ 2 ด้วย $1/m$ โดยปกติค่า m จะหาได้จาก

$$m = 2^b \dots\dots\dots \text{สำหรับคอมพิวเตอร์ฐานสอง}$$

$$m = 10^d \dots\dots\dots \text{สำหรับคอมพิวเตอร์ฐานสิบ}$$

เมื่อ b คือ จำนวน bites ใน 1 word

d คือ จำนวนหลักใน 1 word

4. เลือกใช้ตัวเลขหลังทศนิยมของเลขที่ได้ในข้อ 3 เป็นตัวเลขสุ่มที่
ต้องการ

2

5. ตัดจุดทศนิยมออกจากตัวเลขในข้อ 4 แล้วดำเนินการตามขั้นตอนที่
ต้องการ

6. ทำซ้ำในขั้นที่ 2 ถึง 5 จนกว่าจะได้จำนวนตัวเลขสุ่มมากเท่าที่
ต้องการ

ในปัจจุบันนี้ ได้มีโปรแกรมสำเร็จรูปสำหรับสร้างตัวเลขสุ่มในเครื่อง
คอมพิวเตอร์ เช่น ในภาษาเบสิก (BASIC) มีคำสั่งที่ใช้เรียกตัวเลขสุ่มได้ คือ RANDOMIZED
และ RND ส่วนในภาษาฟอร์แทรน (FORTRAN) ก็มีคำสั่งเรียกตัวเลขสุ่ม คือ RANDUM เป็นต้น

ในปี ค.ศ. 1987 เจโนสกี (Janosky. 1987 : 4367 - A) ได้
ศึกษาการใช้ไมโครคอมพิวเตอร์ในการสร้างตัวเลขสุ่มเทียม (Psuedo Random Number)
ที่มีการแจกแจงแบบยูนิฟอร์ม (Uniform Distribution) 3 วิธี คือ วิธีของไอบีเอ็ม-
แรนดุม (IBM-RANDU) วิธีของมอร์แกน (The Morgan. 1984) และวิธีของเลแมนและ
ไบเล่ย์ (The Leman and Bailey. 1968) และที่มีการแจกแจงแบบปกติ (Normal
Distribution) 3 วิธี คือ ใช้ทฤษฎีเซ็นทรัลลิมิต (Central Limit Theorm) วิธีของ
บอกซ์ และมุลเลอร์ (The Box and Muler Method. 1958) และวิธีโพลาร์ (The
Polar Method, Marsaglia and Bray. 1964) พร้อมทั้งศึกษาประสิทธิภาพระหว่างการ
แจกแจงแบบยูนิฟอร์มกับการแจกแจงแบบปกติ ผลการศึกษาพบว่า การสร้างตัวเลขสุ่มเทียม
ที่มีการแจกแจงแบบยูนิฟอร์มโดยใช้วิธีของเลแมนกับไบเล่ย์ให้ผลดีที่สุด ส่วนการแจกแจงปกติ
วิธีที่ดีที่สุดเป็นการผสมระหว่างการใช้อุทฤษฎีเซ็นทรัล ลิมิต กับวิธีของเลแมนกับไบเล่ย์

นอกจากวิธีการที่กล่าวมาแล้ว 2 วิธี ในปัจจุบันมีการใช้วิธีการอีกหลาย
อย่างในการสร้างตัวเลขแบบสุ่มเทียม รวมทั้งมีโปรแกรมสำเร็จรูปให้เลือกใช้ โดยผู้ใช้ไม่
ต้องสร้างโปรแกรมเอง แต่วิธีการและโปรแกรมที่ใช้ในการสร้างตัวเลขแบบสุ่มนั้น ต้องมี
คุณสมบัติของตัวเลขสุ่มที่ดี ดังนี้

1. ตัวเลขสุ่มที่ได้จะต้องมีลักษณะการกระจายความน่าจะเป็นแบบ
สม่ำเสมอ (Uniform Distribution)

2. ตัวเลขสุ่มที่ได้ต้องเป็นอิสระแก่กัน

3. อนุกรมของตัวเลขสุ่มที่ได้ต้องสามารถสร้างซ้ำเดิมได้

(Reproducible)

4. อนุกรมของตัวเลขสุ่มที่ได้ต้องไม่ซ้ำเดิมในช่วงที่ต้องการใช้ตัวเลขสุ่ม
5. ต้องใช้เวลาน้อยในการสร้างตัวเลขสุ่ม
6. ต้องใช้หน่วยความจำในคอมพิวเตอร์น้อย

2. การนำตัวเลขสุ่มมาประยุกต์ใช้กับปัญหาต่าง ๆ ขั้นตอนนี้เป็น การนำตัวเลขสุ่มไปสร้างตัวแปรตามลักษณะการแจกแจงของปัญหาที่จะศึกษา เพื่อเป็นข้อมูลของปัญหานั้น (Swarup and Mohan. 1982 : 666 - 674) เช่น สร้างตัวเลขสุ่มขึ้นมาจำนวนหนึ่ง แล้วนำเลขสุ่มนั้นไปสร้างเป็นคะแนนผลการสอบของผู้เรียน บางครั้งตัวแปรของปัญหาที่จะศึกษาไม่ได้สร้างจากตัวเลขสุ่มโดยตรง แต่มีขั้นตอนที่ต้องอาศัยตัวเลขสุ่มเป็นพื้นฐานก็ได้

3. ทำการทดลองซ้ำหลาย ๆ ครั้ง หลักการสำคัญประการหนึ่งของเทคนิควิธีมอนติ คาร์โล ก็คือ ต้องมีการทดลองซ้ำหลาย ๆ ครั้ง (Krope. 1980 : 242) เพื่อลดความคลาดเคลื่อนของค่าตอบที่จะได้ (Hammersley and Handscomb. 1964 : 5, 21) และสามารถสรุปเป็นความน่าจะเป็นของการเกิดเหตุการณ์ในปัญหานั้น ๆ

8.3 จุดเด่นของการใช้เทคนิควิธีมอนติ คาร์โล

เนื่องจากเทคนิควิธีมอนติ คาร์โล จะใช้ตัวเลขสุ่มเป็นพื้นฐานในการสร้างตัวแปรของปัญหา โดยอาศัยทฤษฎี สูตร หรือกฎเกณฑ์ต่าง ๆ ที่มีอยู่ และมีการทดลองซ้ำหลาย ๆ ครั้ง เพื่อลดความคลาดเคลื่อนต่าง ๆ จึงนับว่ามีประโยชน์ที่สำคัญ ดังนี้

1. เทคนิควิธีมอนติ คาร์โล สามารถควบคุมตัวแปรแทรกซ้อนและสามารถสังเกตได้อย่างสมบูรณ์ (Completely Controlled and Completely Observed) (Payne. 1988 : 3) นอกจากนี้ ยังสามารถทำการทดลองซ้ำภายใต้สภาพแวดล้อม (Context) เดิมหลาย ๆ ครั้งได้ ซึ่งในการทดลองจริงนั้นทำไม่ได้ เพราะไม่สามารถรักษาสภาพแวดล้อมให้เหมือนเดิมทุกอย่างได้ เมื่อเวลาเปลี่ยนไป

2. ในการใช้เทคนิควิธีมอนติ คาร์โล ถ้ามี ทฤษฎี สูตร หรือกฎเกณฑ์ต่าง ๆ ที่ถูกต้องรองรับในการสร้างตัวแปรของปัญหาในการทดลองแล้ว จะทำให้ผลที่ได้ถูกต้องแม่นยำกว่าเมื่อใช้ทดลองในสถานการณ์จริง ทั้งนี้เพราะสามารถลดตัวแปรแทรกซ้อนในเชิงจิตวิทยาได้

3. ทำให้สิ้นเปลืองเวลา แรงงาน และค่าใช้จ่ายน้อยกว่า เมื่อเทียบกับการทดลองในสถานการณ์จริง

ดังนั้น ในการวิจัยครั้งนี้ ผู้วิจัยจึงได้นำเทคนิควิธีมอนติ คาร์โล มาใช้ในการเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซ์ิม ไลค์ลิสต์ วิธีอีวีรอสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน โดยใช้โปรแกรมคอมพิวเตอร์ภาษา TURBO PASCAL Version 7.0 ที่ผู้วิจัยเป็นผู้พัฒนา

9. งานวิจัยที่เกี่ยวข้อง

9.1 งานวิจัยที่เกี่ยวข้องในประเทศ

วินัย วงศ์ฤทธิวัฒนา (2532 : 95) ได้เปรียบเทียบผลการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก แบบ 2 พารามิเตอร์ ระหว่างวิธีของเบส์กับวิธีแมกซ์ิม ไลค์ลิสต์ กลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ข้อมูล เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ที่ศึกษาอยู่ในโรงเรียนของอำเภอมนรมย์ จังหวัดชัยนาท ปีการศึกษา 2531 จำนวน 323 คน เครื่องมือที่ใช้เป็นแบบทดสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยได้สร้างขึ้นเอง การวิเคราะห์ข้อมูลใช้วิธีของเบส์และวิธีแมกซ์ิม ไลค์ลิสต์ ประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 2 พารามิเตอร์ การเปรียบเทียบผลการประมาณค่าพารามิเตอร์ใช้ค่าสถิติที่ (t-test) การศึกษาความสัมพันธ์ระหว่างค่าพารามิเตอร์ที่ประมาณค่าด้วยวิธีของเบส์กับค่าพารามิเตอร์ที่ประมาณค่าด้วยวิธีแมกซ์ิม ไลค์ลิสต์ ใช้ค่าสหสัมพันธ์แบบเพียร์สัน ผลการวิจัยพบว่า

1. ค่าอำนาจจำแนก (a) ที่ประมาณค่าด้วยวิธีของเบส์และวิธีแมกซ์ิม ไลค์ลิสต์ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ
2. ค่าความยาก (b) ที่ประมาณค่าด้วยวิธีของเบส์และวิธีแมกซ์ิม ไลค์ลิสต์ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05
3. ค่าอำนาจจำแนก (a) ที่ประมาณค่าด้วยวิธีของเบส์และวิธีแมกซ์ิม ไลค์ลิสต์ มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
4. ค่าความยาก (b) ที่ประมาณค่าด้วยวิธีของเบส์และวิธีแมกซ์ิม ไลค์ลิสต์ มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

วิชุดา บัวคง (2533 : 115) ได้เปรียบเทียบประสิทธิผลของวิธีการประมาณค่าพารามิเตอร์ข้อสอบและพารามิเตอร์ความสามารถของผู้เข้าสอบระหว่างวิธีแมกซิมัมไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ ของแบบจำลองโลจิสติก 3 พารามิเตอร์ โดยศึกษาจากแบบทดสอบวัดผลสัมฤทธิ์ วิชาภาษาไทย (ด้านการใช้ภาษา) และแบบทดสอบความถนัดด้านการใช้ภาษาไทย ซึ่งมีจำนวน 40 ข้อ และ 80 ข้อ ตามลำดับ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดกรมสามัญศึกษา จังหวัดจันทบุรี จำนวน 1,028 คน การวิเคราะห์ข้อมูลใช้โปรแกรม LOGIST 5 Version 2.5 และโปรแกรมภาษาฟอร์แทรน เพื่อประมาณค่าพารามิเตอร์ความยาก อำนาจจำแนก การเดาของข้อสอบ และความสามารถของผู้เข้าสอบตามวิธีประมาณค่าทั้ง 3 วิธี คำนวณค่าฟังก์ชันสารสนเทศของข้อสอบและแบบทดสอบ ค่าประสิทธิภาพสัมพัทธ์และความเที่ยงตรงเชิงเกณฑ์สัมพัทธ์ของแบบทดสอบ ผลการวิจัยพบว่า

1. แบบทดสอบวัดผลสัมฤทธิ์ที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัม ไลค์ลิฮูด มีประสิทธิภาพสูงที่สุด ในกลุ่มผู้เข้าสอบที่มีความสามารถสูง รองลงไป คือ แบบทดสอบที่ประมาณค่าด้วยวิธีของเบส์ และวิธีอีวีริสติก ตามลำดับ ส่วนในกลุ่มผู้เข้าสอบที่มีความสามารถปานกลางและต่ำ นั้น แบบทดสอบที่ประมาณค่าด้วยวิธีของเบส์ มีประสิทธิภาพสูงที่สุด รองลงไป คือ แบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด และวิธีอีวีริสติก ตามลำดับ สำหรับค่าความเที่ยงตรงตามสภาพที่เป็นผลจากการประมาณค่าความสามารถของผู้เข้าสอบทั้ง 3 วิธี แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยวิธีของเบส์ ให้ค่าความเที่ยงตรงตามสภาพสูงที่สุด รองลงไป คือ วิธีอีวีริสติก หรือวิธีแมกซิมัม ไลค์ลิฮูด
2. แบบทดสอบความถนัดที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัม ไลค์ลิฮูด มีประสิทธิภาพสูงที่สุด ในกลุ่มผู้เข้าสอบที่มีความสามารถสูง รองลงไปคือ แบบทดสอบที่ประมาณค่าด้วยวิธีของเบส์ และวิธีอีวีริสติก ตามลำดับ ส่วนในกลุ่มผู้เข้าสอบที่มีความสามารถปานกลาง แบบทดสอบที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบส์ มีประสิทธิภาพสูงที่สุด รองลงไป คือ แบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด และวิธีอีวีริสติก ตามลำดับ แต่ในกลุ่มผู้สอบที่มีความสามารถต่ำ แบบทดสอบที่ประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบส์ มีประสิทธิภาพสูงที่สุด รองลงไปคือ แบบทดสอบที่ประมาณค่าด้วยวิธีอีวีริสติก และวิธีแมกซิมัม ไลค์ลิฮูด ตามลำดับ สำหรับค่าความเที่ยงตรงเชิงทำนายที่เป็นผลจากการประมาณค่าความสามารถของผู้สอบทั้ง 3 วิธี แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยวิธีของเบส์ ให้ค่าความเที่ยงตรงเชิงทำนายสูงที่สุด รองลงไป คือ วิธีอีวีริสติก หรือวิธีแมกซิมัม ไลค์ลิฮูด ตามลำดับ

9.2 งานวิจัยที่เกี่ยวข้องต่างประเทศ

ลอร์ด (Lord and Novick. 1968 : 989 - 1020) ได้ศึกษาเกี่ยวกับการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ด้วยวิธีแมกซิมัมไลค์ลิฮูด โดยใช้ผลการสอบจากแบบทดสอบ SAT (Scholastic Aptitude Test) ฉบับภาษา (Verbal) ฟอร์ม MSA 3 ของผู้เข้าสอบ 2,862 คน พบว่า การประมาณค่าความสามารถและค่าพารามิเตอร์ของข้อสอบต้องตัดผู้เข้าสอบที่ตอบข้อสอบถูกทุกข้อ และข้อสอบที่ผู้เข้าสอบทุกคนตอบถูกออกจากการประมาณค่า เพราะการประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิฮูดไม่สามารถประมาณค่าความสามารถของผู้เข้าสอบที่ได้คะแนนเต็มหรือคะแนนศูนย์ และค่าพารามิเตอร์ของข้อสอบที่ผู้เข้าสอบทุกคนตอบถูกหรือตอบผิดได้ ในกรณีนี้ ลอร์ด กล่าวว่า อาจหลีกเลี่ยงปัญหาดังกล่าวได้โดยใช้ข้อสอบให้มากกว่า 20 ข้อ (Lord. 1980 : 51) นอกจากนี้ยังพบว่า การประมาณค่าอำนาจจำแนกของข้อสอบบางข้อได้ค่าประมาณที่ไม่ลู่อู่เข้าสู่ค่าคงที่ กล่าวคือ ไม่สามารถประมาณค่าที่ทำให้ค่าประมาณครั้งที่ $m+1$ และครั้งที่ m มีค่าแตกต่างกันน้อยกว่าค่าคงที่ (ขนาดน้อย ๆ) ที่กำหนดไว้ได้ อีกทั้งการประมาณค่าการเดาของข้อสอบบางข้อยังได้ค่าประมาณที่ไม่ถูกต้อง เช่น ได้ค่าเป็นลบหรือเป็นค่าบวกที่มากผิดปกติ อันเนื่องมาจากการประมาณค่าอำนาจจำแนกไม่ลู่อู่เข้าสู่ค่าคงที่ และความผิดปกติของค่าการเดานี้อาจมีผลทำให้การประมาณค่าพารามิเตอร์ของข้อสอบข้ออื่น ๆ หรือค่าพารามิเตอร์ตัวอื่นคลาดเคลื่อนไปด้วย ลอร์ดจึงแก้ปัญหาโดยกำหนดขีดจำกัดบน (Upper Limit) ของค่าอำนาจจำแนก a ไว้ที่ 2.0 และกำหนดพิสัย (Range) ของค่าการเดา c ให้มีค่าอยู่ในช่วง 0 ถึง 0.2 ซึ่งการกำหนดพิสัยของค่าพารามิเตอร์ดังกล่าว เป็นผลให้การประมาณค่าพารามิเตอร์ได้ค่าประมาณที่ถูกต้องและเหมาะสมยิ่งขึ้น และเมื่อลอร์ด (Hambleton and Swaminathan. 1985 : 480; citing Lord. 1975) ทำการศึกษาซ้ำโดยวิธีจำลองสถานการณ์ (Simulation) จากข้อมูลผลการตอบข้อสอบ 90 ข้อ ของผู้เข้าสอบ 2,995 คน พบว่า มีค่าประมาณอำนาจจำแนกของข้อสอบ 3 ข้อ ที่ไม่ลู่อู่เข้า (No Convergence) แต่เมื่อกำหนดขีดจำกัดบนของค่าอำนาจจำแนกของข้อสอบทั้ง 3 ข้อ ดังกล่าวให้เท่ากับ 1.75 แล้ว การประมาณค่าอำนาจจำแนกของข้อสอบทุกข้อ ได้ค่าประมาณที่มีความสอดคล้องกับค่าอำนาจจำแนกจริง ($r_{\text{true}} = 0.92$) นอกจากนี้ ลอร์ดยังได้พบว่า ถ้าข้อสอบค่อนข้างง่ายแล้วจะมีแนวโน้มที่จะประมาณค่าอำนาจจำแนกได้ต่ำกว่าค่าจริง แต่ถ้าข้อสอบค่อนข้างยากจะมีแนวโน้มที่จะประมาณค่าอำนาจจำแนกได้สูงกว่าค่าจริงเช่นกัน

ชมิคท์ (Schmidt. 1977 : 613 - 620) พบว่า การประมาณค่าพารามิเตอร์ด้วยวิธีวิธีวิสถิติมีความคลาดเคลื่อนอย่างเป็นระบบ กล่าวคือ การประมาณค่าอำนาจจำแนกได้ค่าประมาณต่ำกว่าค่าจริงเสมอ และการประมาณค่าความยากได้ค่าประมาณสูงกว่าค่าจริงเสมอ ซึ่งชมิคท์ได้แนะนำให้ใช้รากที่สองของค่าความเชื่อมั่นของแบบทดสอบที่คำนวณโดยสูตรคูเดอร์-ริชาร์ดสัน ที่ 20 เพื่อปรับให้ได้ค่าประมาณที่ถูกต้องยิ่งขึ้น ซึ่งค่าประมาณจากวิธีวิสถิติที่ปรับแก้แล้วสามารถนำไปใช้ได้

สวามินาธาน และกิฟฟอร์ด (Hulin. 1983 : 99 - 100; citing Swaminathan and Gifford. 1979) ได้ศึกษาความเที่ยงตรงของค่าประมาณที่ได้จากวิธีแมกซีมิ้มไลค์ลิสต์กับค่าพารามิเตอร์จริงของข้อสอบและผู้เข้าสอบ จากข้อมูลในสถานการณ์จำลอง โดยวิธีวิธีวิสถิติสหสัมพันธ์แบบเพียร์สันในการอธิบายความเที่ยงตรง ได้ข้อสรุปว่า ความสัมพันธ์หรือความเที่ยงตรงระหว่างค่าประมาณกับค่าพารามิเตอร์จะแปรตามจำนวนข้อสอบและจำนวนผู้เข้าสอบ ซึ่งผลที่ได้นี้สอดคล้องกับข้อสรุปของลอร์ด (Lord. 1968) ที่กล่าวว่า เพื่อให้การประมาณค่าพารามิเตอร์ในแบบจำลองโลจิสติก 3 พารามิเตอร์ มีความถูกต้องเพียงพอ นั้น ควรจะมีข้อสอบประมาณ 50 ข้อ และผู้เข้าสอบอย่างน้อย 1,000 คน และต่อมาในปี ค.ศ. 1982 สวามินาธาน และกิฟฟอร์ด (Swaminathan and Gifford. 1982 : 175 - 191) ได้ศึกษาเกี่ยวกับการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 1 พารามิเตอร์ หรือแบบจำลองของราสช์ (Rasch Model) ด้วยวิธีของเบส์ โดยการจำลองสถานการณ์ พบว่า วิธีของเบส์สามารถประมาณค่าได้อย่างถูกต้องและสามารถประมาณค่าพารามิเตอร์ของข้อสอบที่ผู้เข้าสอบทุกคนตอบถูกหรือตอบผิดหมดได้ รวมทั้งสามารถประมาณค่าความสามารถของผู้เข้าสอบที่ได้คะแนนเต็มหรือคะแนนศูนย์ได้ด้วย

ในปี ค.ศ. 1980 แม็คคินลีย์ และเร็คเคส (McKinley and Reckase. 1980) และปี ค.ศ. 1983 สวามินาธาน และกิฟฟอร์ด ได้นำกระบวนการของวิธีวิสถิติไปทดลองใช้ พบว่า ถ้าขนาดของผู้เข้าสอบ และจำนวนข้อสอบมีมากพอแล้ว ผลการประมาณค่าด้วยวิธีวิสถิติ จะส่งผลต่อการประมาณค่าความสามารถของผู้เข้าสอบให้มีความถูกต้องมากขึ้น จนสามารถนำไปเทียบกับผลการประมาณที่ได้จากวิธีแมกซีมิ้ม ไลค์ลิสต์ ต่อมา ในปี ค.ศ. 1985 สวามินาธาน และกิฟฟอร์ด ได้พัฒนาวิธีของเบส์เพิ่มเติม และนำไปใช้ประมาณค่าพารามิเตอร์ ในแบบจำลองโลจิสติก 2 พารามิเตอร์ โดยการจำลองสถานการณ์ 12 สถานการณ์ กล่าวคือ ให้มีผลการตอบข้อสอบ 15 ข้อ 25 ข้อ และ 35 ข้อ ของผู้เข้าสอบ 50 คน 100 คน และ 500 คน พบว่า ค่าสหสัมพันธ์ (r) ของค่าพารามิเตอร์ (θ, b, a)

และค่าประมาณ ($\hat{a}, \hat{b}, \hat{a}$) และค่าความแตกต่างระหว่างค่าเฉลี่ยกำลังสอง (Mean Squared Difference : MSD) ของค่าพารามิเตอร์ และค่าประมาณในสถานการณ์ต่าง ๆ ตั้งแต่ข้อสอบ 15 ข้อ ผู้เข้าสอบ 50 คน ถึงข้อสอบ 35 ข้อ ผู้เข้าสอบ 500 คน เป็นดังนี้ (Swaminathan and Gifford. 1985 : 349 - 364)

1. $r_{\hat{a}}^2$ มีค่าตั้งแต่ 0.934 ถึง 0.966 และ $MSD_{\hat{a}}$ มีค่าตั้งแต่ 0.129 ถึง 0.086
2. $r_{\hat{b}}^2$ มีค่าตั้งแต่ 0.990 ถึง 0.996 และ $MSD_{\hat{b}}$ มีค่าตั้งแต่ 0.041 ถึง 0.009
3. $r_{\hat{a}}^2$ มีค่าตั้งแต่ 0.706 ถึง 0.935 และ $MSD_{\hat{a}}$ มีค่าตั้งแต่ 0.041 ถึง 0.031

จากการศึกษาในครั้งนี้ ทำให้ได้ข้อสรุปว่า ถ้ากลุ่มตัวอย่างผู้เข้าสอบ และจำนวนข้อสอบ มีขนาดและจำนวนมากขึ้นแล้ว ค่าสหสัมพันธ์ที่ได้จะมีค่ามากขึ้นด้วย ยกเว้นในบางกลุ่มตัวอย่างที่ค่า $r_{\hat{a}}^2$ ลดลงเล็กน้อย ในขณะที่ค่าความแตกต่างระหว่างค่าเฉลี่ยกำลังสองของค่าพารามิเตอร์ และค่าประมาณจะมีค่าลดลง ซึ่งสวามินาธาน และกัฟฟอร์ด ยังคงพัฒนาวิธีของเบส์อย่างต่อเนื่อง จนกระทั่งปี ค.ศ.1986 ได้นำวิธีของเบส์ไปใช้ประมาณค่าพารามิเตอร์ในแบบจำลองโลจิสติก 3 พารามิเตอร์ เพื่อศึกษาประสิทธิภาพของวิธีประมาณโดยจำลองสถานการณ์ 6 สถานการณ์ คือ ให้มีผลการสอบจากข้อสอบ 25 ข้อ และ 30 ข้อของผู้เข้าสอบ 100 คน 200 คน และ 400 คน ได้ข้อค้นพบเกี่ยวกับค่าสหสัมพันธ์ของค่าพารามิเตอร์จริงกับค่าประมาณ และค่าความแตกต่างระหว่างค่าเฉลี่ยกำลังสองของค่าพารามิเตอร์และค่าประมาณ (Mean Squared Differences Between Estimates and True Values) ในสถานการณ์ที่มีข้อสอบ 25 ข้อ ของผู้เข้าสอบ 100 คน ถึงสถานการณ์ที่มีข้อสอบ 30 ข้อ ของผู้เข้าสอบ 400 คน ดังนี้ (Swaminathan and Gifford. 1986 : 589 - 601)

1. $r_{\hat{a}}^2$ มีค่าตั้งแต่ 0.930 ถึง 0.950 และ $MSD_{\hat{a}}$ มีค่าตั้งแต่ 0.139 ถึง 0.100
2. $r_{\hat{b}}^2$ มีค่าตั้งแต่ 0.947 ถึง 0.992 และ $MSD_{\hat{b}}$ มีค่าตั้งแต่ 0.176 ถึง 0.024
3. $r_{\hat{a}}^2$ มีค่าตั้งแต่ 0.584 ถึง 0.697 และ $MSD_{\hat{a}}$ มีค่าตั้งแต่ 0.083 ถึง 0.068

4. r_{c_2} มีค่าตั้งแต่ 0.637 ถึง 0.730 และ MSD_{c_2} มีค่าตั้งแต่ 0.035 ถึง 0.022

ฮูลิน และคนอื่น ๆ (Hulin and Others. 1983 : 101 - 103) ได้ใช้วิธีจำลองสถานการณ์ (Simulation) ศึกษาเกี่ยวกับความถูกต้องในการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก ด้วยวิธีแมกซิมัม ไลค์ลิฮูด พบว่า การประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ จะมีความถูกต้องเพียงพอกับแบบทดสอบมีข้อสอบอย่างน้อย 30 ข้อ และขนาดของกลุ่มตัวอย่างไม่น้อยกว่า 1,000 คน เพราะจะทำให้ค่าเฉลี่ยของความแตกต่างระหว่างโค้งลักษณะข้อสอบของค่าพารามิเตอร์จริงกับโค้งลักษณะข้อสอบของค่าประมาณ (Root Mean Square Error : RMSE) มีค่าต่ำ ($RMSE < 0.05$) และค่าสหสัมพันธ์ระหว่างค่าความสามารถจริง θ กับ ค่าประมาณความสามารถ $\hat{\theta}$ มีค่าสูง ($r_{\theta\hat{\theta}} > 0.860$)

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องจะพบว่า การศึกษาเพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยเทคนิควิชันมอนติ คาร์โล นั้น ยังไม่ปรากฏแน่ชัดว่ามีผู้ใดศึกษา ดังนั้น ผู้วิจัยจึงมีความประสงค์ที่จะศึกษาประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยเทคนิควิชันมอนติ คาร์โล ว่าจะให้ผลแตกต่างกันหรือไม่ อย่างไร

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยเพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยเทคนิคมอนติ คาร์โล มีขั้นตอนในการวิจัย ดังนี้

ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นคะแนนของผู้เข้าสอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี พ.ศ. 2537 ในวิชาภาษาไทย ของ อ.ก.ค.กรมสามัญศึกษา จำนวน 15,258 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นคะแนนของผู้เข้าสอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี พ.ศ. 2537 ในวิชาภาษาไทย ของ อ.ก.ค.กรมสามัญศึกษา ซึ่งสุ่มออกเป็น 13 ขนาด คือ ขนาดกลุ่ม 200 คน 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน โดยกลุ่มตัวอย่างแต่ละกลุ่มจะต้องตอบข้อสอบที่มีขนาดต่างกัน 6 ขนาด คือ ขนาดข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ 90 ข้อ และ 100 ข้อ โดยการทดลองซ้ำกลุ่มละ 40 ครั้ง ด้วยเทคนิคมอนติ คาร์โล (ค่าช χ^2 : 2534 : 85)

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบทดสอบวิชาภาษาไทย แบบปรนัย ชนิด 4 ตัวเลือก จำนวน 100 ข้อ ซึ่ง อ.ก.ค.กรมสามัญศึกษา เป็นผู้สร้างตามหลักสูตรของ ก.ค.ที่ใช้ในการสอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู

สถิติที่ใช้ในการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยได้ใช้ทั้งโปรแกรมคอมพิวเตอร์สำเร็จรูป และโปรแกรมคอมพิวเตอร์ ภาษา TURBO PASCAL Version 7.0 ที่ผู้วิจัยได้พัฒนาขึ้นใหม่ เพื่อวิเคราะห์

ข้อมูลดังต่อไปนี้

1. ใช้โปรแกรมคอมพิวเตอร์ ภาษา TURBO PASCAL Version 7.0 ที่ผู้วิจัยพัฒนาขึ้น เพื่อแปลงคะแนนการตอบข้อสอบเป็น 0 ในกรณีที่ตอบผิด และ 1 ในกรณีที่ตอบถูก และใช้ในเทคนิควิธียอนติ คาร์โล

2. ใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0 เพื่อวิเคราะห์องค์ประกอบของคะแนนผลการตอบแบบทดสอบวิชาภาษาไทย ในการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบ

3. ใช้โปรแกรม BILOG 3 เพื่อประมาณค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบ ด้วยวิธีมาร์กีนัล แมกซิมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

4. ใช้โปรแกรมคอมพิวเตอร์ ภาษา TURBO PASCAL Version 7.0 ที่ผู้วิจัยพัฒนาขึ้น เพื่อประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ด้วยวิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

5. ใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0 เพื่อเปรียบเทียบความแตกต่างของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ซึ่งประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน โดยใช้สถิติทดสอบแบบ Repeated Measure (AxS) Design (Keppel. 1973 : 401 - 402) ถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติแล้ว ทดสอบความแตกต่างรายคู่ด้วยวิธีการของเชฟเฟ (Scheffe's Test) (ลิวน ศาสยศ และอังคณา ศาสยศ. 2536 : 119)

6. ใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0 เพื่อเปรียบเทียบความแตกต่างของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน โดยใช้สถิติทดสอบแบบ Repeated Measure (AxS) Design (Keppel. 1973 : 401 - 402) ถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติแล้ว ทดสอบความแตกต่างรายคู่ด้วยวิธีการของเชฟเฟ (Scheffe's Test) (ลิวน ศาสยศ และอังคณา ศาสยศ. 2536 : 119)

7. ใช้โปรแกรมภาษา TURBO PASCAL Version 7.0 ที่ผู้วิจัยพัฒนาขึ้น เพื่อเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของวิธีการประมาณค่าพารามิเตอร์ของข้อสอบ ด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบ

ต่างกัน ๗ ระดับความสามารถต่าง ๆ

การวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ ผู้วิจัยทำการวิเคราะห์ข้อมูลด้วยเครื่องไมโครคอมพิวเตอร์ โดยใช้โปรแกรมสำเร็จรูป คือ โปรแกรม SPSS for WINDOWS Version 6.0 และโปรแกรม BILOG 3 รวมทั้งได้พัฒนาโปรแกรมคอมพิวเตอร์ ภาษา TURBO PASCAL Version 7.0 เพื่อใช้สำหรับเทคนิควิธีมอนติ คาร์โล และใช้ในการประมาณค่าพารามิเตอร์ด้วยวิธีวิวิธวิธี และวิธีของเบส์ รวมทั้งใช้ในการคำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบ และค่าสถิติต่าง ๆ ที่ใช้ในการทดสอบสมมติฐาน โดยมีขั้นตอนดังต่อไปนี้

1. การตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ ตามข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม

1.1 การตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว ผู้วิจัยได้ตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว โดยการนำคะแนนผลการสอบวิชาภาษาไทยมาทำการวิเคราะห์หองค์ประกอบ โดยใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0 ซึ่งในการวิเคราะห์หองค์ประกอบนี้ ใช้การสกัดองค์ประกอบวิธีองค์ประกอบหลัก (Principal Component) เพราะเป็นวิธีการลดข้อมูลหรือตัวแปรให้น้อยลง โดยอาศัยหลักความสัมพันธ์เชิงเส้นระหว่างตัวแปร (A Linear Combination of the Observed Data) ที่ใช้เป็นข้อมูล แต่ไม่มีการสมมุติเกี่ยวกับความสัมพันธ์เชิงสาเหตุและผลระหว่างองค์ประกอบและตัวแปร สำหรับการหมุนแกนใช้การหมุนแบบมุมฉากด้วยวิธีแวนนิแมกซ์ (Varimax) เพราะวิธีการนี้ เป็นวิธีที่พยายามลดจำนวนตัวแปรที่มีน้ำหนักองค์ประกอบมากบนแต่ละองค์ประกอบให้เหลือน้อยที่สุด ซึ่งจะช่วยให้การตีความหมายขององค์ประกอบได้ง่ายขึ้น (สุชาติ ประสิทธิ์-รัฐสินธุ์. 2533 : 20, 35) หลังจากนั้นจะพิจารณาค่าไอเกน (Eigen Value) ถ้าค่าไอเกนขององค์ประกอบที่ 1 ของแบบทดสอบวิชาภาษาไทย มีค่าสูงกว่าค่าไอเกนขององค์ประกอบที่ 2 มาก และค่าไอเกนขององค์ประกอบที่ 2 มีค่าสูงกว่าค่าไอเกนขององค์ประกอบถัดไปเพียงเล็กน้อย แสดงว่า แบบทดสอบวิชาภาษาไทยมีคุณสมบัติการวัดเพียงมิติเดียว (Lord. 1980 : 21)

1.2 การตรวจสอบคุณสมบัติความเป็นอิสระในการการตอบข้อสอบ ทำได้โดยการพิจารณาว่า ถ้าแบบทดสอบมีคุณสมบัติการวัดเพียงมิติเดียวแล้ว ก็จะมีคุณสมบัติความเป็นอิสระในการตอบข้อสอบด้วย (Hambleton and Swaminathan. 1985 : 24)

2. การสุ่มตัวอย่างแบบใส่คืน (Random Sampling With Replacement) เพื่อสุ่มตัวอย่างผลการสอบของผู้สอบออกเป็น 13 ขนาดกลุ่ม คือ ขนาดกลุ่ม 200 คน 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน โดยกลุ่มตัวอย่างแต่ละกลุ่มจะตอบข้อสอบที่มีขนาดต่างกัน 6 ขนาด คือ ขนาดข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ 90 ข้อ และ 100 ข้อ ซึ่งได้มาด้วยการสุ่มจากตารางวิเคราะห์เนื้อหา ที่แบ่งออกเป็น 3 เนื้อหาใหญ่ ๆ โดยการทดลองซ้ำกลุ่มละ 40 ครั้ง ด้วยเทคนิควิธีมอนติ คาร์โล

3. การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบด้วยวิธีวิธีสติคและวิธีของเบส์ ซึ่งในการวิจัยครั้งนี้ ผู้วิจัยได้พัฒนาโปรแกรมคอมพิวเตอร์ขึ้นเพื่อใช้ในการวิจัยครั้งนี้ 2 โปรแกรม คือ

3.1 โปรแกรม HEU-IRT เป็นโปรแกรมที่ใช้สำหรับการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบด้วยวิธีวิธีสติค

3.2 โปรแกรม BAY-IRT เป็นโปรแกรมที่ใช้สำหรับการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบด้วยวิธีของเบส์

ในการพัฒนาโปรแกรมคอมพิวเตอร์ทั้งสองโปรแกรมนี้นี้ มีขั้นตอนในการพัฒนาดังต่อไปนี้

ขั้นที่ 1 เลือกโปรแกรมแปลภาษา ปัจจุบันมีโปรแกรมแปลภาษาทั้งที่เป็นแบบ Compiler และ Interpreter ให้เลือกใช่มากมาย แต่ผู้วิจัยได้เลือกใช้โปรแกรมแปลภาษาของ TURBO PASCAL Version 7.0 ซึ่งเป็นแบบ Compiler ทั้งนี้เพราะเมื่อ Compiler เรียบร้อยแล้ว สามารถนำโปรแกรมที่ได้ไปใช้งานได้โดยไม่ต้องมีการ Compiler อีก ทำให้สะดวกต่อการใช้งาน และถ้ามีการปรับปรุงโปรแกรมที่ได้เพียงเล็กน้อยก็สามารถนำไป Run บนโปรแกรม WINDOWS ได้ ทำให้ขีดความสามารถในการคำนวณเพิ่มขึ้น

ขั้นที่ 2 การวิเคราะห์ระบบงานที่จะนำมาใช้กับโปรแกรมคอมพิวเตอร์
ขั้นตอนนี้เป็นงานที่เกี่ยวข้องกับการวิเคราะห์ข้อมูลทางสถิติ ซึ่งประกอบด้วย

1. ข้อมูลที่นำมาใช้ในการวิเคราะห์ ข้อมูลที่นำมาใช้ในการวิเคราะห์ต้องอยู่ในรูปของตัวอักษรที่ได้จากการตรวจกระดาษคำตอบด้วยเครื่อง OMR ซึ่งใน 1 เรคคอร์ดจะประกอบด้วยรหัสประจำตัวผู้เข้าสอบ และผลการตอบข้อสอบของผู้เข้าสอบ ซึ่งมีความยาวทั้งหมด 110 ตัวอักษร

2. ขั้นตอนการประมวลผล เป็นการแทนค่าของตัวแปรต่าง ๆ ในสมการที่นำมาใช้ เพื่อคำนวณหาค่าพารามิเตอร์ตามที่ต้องการ ซึ่งการแทนค่าดังกล่าวจะถูกกระทำด้วยเครื่องคอมพิวเตอร์ ทั้งนี้รวมถึงการกระทำซ้ำเพื่อประมาณค่าพารามิเตอร์ในรูปแบบต่าง ๆ

3. ผลลัพธ์ที่ได้จากการประมวลผล ผลลัพธ์ที่ได้จากการประมวลผลจะถูกเก็บอยู่ในแฟ้มข้อมูล และจะถูกพิมพ์ออกทางเครื่องคอมพิวเตอร์

ขั้นที่ 3 การเขียนโปรแกรม เป็นการเขียนโปรแกรมเพื่อสร้าง Source Code แล้วนำไป Compiler ในขั้นตอนนี้อยู่อาศัยเลือกใช้ Editor ของโปรแกรม TURBO PASCAL Version 7.0 เพราะสะดวกในการนำไป Compiler และการตรวจแก้ไข

ขั้นที่ 4 นำโปรแกรมที่เขียนไปใช้กับข้อมูลทดสอบ ขั้นตอนนี้เป็นการทดสอบความถูกต้องของการประมวลผล ซึ่งถ้าพบข้อผิดพลาดจะต้องดำเนินการแก้ไขให้ถูกต้องก่อนนำไปใช้จริง

ขั้นที่ 5 นำโปรแกรมที่พัฒนาแล้วไปใช้กับข้อมูลจริง

4. การประมาณค่าพารามิเตอร์ อำนาจจำแนก (a_i) ค่าความยาก (b_i) ค่าการเดา (c_i) ของข้อสอบ และค่าความสามารถของผู้เข้าสอบ (θ_j) โดยใช้วิธีประมาณค่าที่แตกต่างกัน 3 วิธี ดังนี้

4.1 การประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบจากผลการสอบที่มีจำนวนผู้สอบและจำนวนข้อสอบแตกต่างกัน ด้วยวิธีแมกซิมัม ไลค์ลิฮูด ซึ่งเป็น การวิเคราะห์ข้อสอบรายข้อตามแบบจำลองโลจิสติก 3 พารามิเตอร์ โดยสามารถเขียนเป็นรูปแบบความสัมพันธ์เชิงคณิตศาสตร์ของค่าพารามิเตอร์ได้ดังนี้

$$P_i(\theta_j) = c_i + (1 - c_i) / (1 + e^{-D a_i (\theta_j - b_i)}) ; (i=1, 2, \dots, n)$$

เมื่อ $P_i(\theta_j)$ = ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถ θ ตอบข้อสอบข้อที่ i ถูก

θ_j = ระดับความสามารถของผู้สอบคนที่ j

a_i = ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i = ค่าความยากของข้อสอบข้อที่ i

c_i = ค่าการเดาของข้อสอบข้อที่ i

D = A Scaling Factor มีค่าเท่ากับ 1.7

e = ค่าคงที่ มีค่าเท่ากับ 2.71828...

สำหรับการวิจัยครั้งนี้ ใช้โปรแกรมสำเร็จรูป BILOG 3 ของมิสเลวี และบ็อค (Mislevy and Bock. 1990 : 6 - 8) ซึ่งเป็นการประมาณค่าพารามิเตอร์ ด้วยวิธีมาร์จิ้นัล แมกซิมัม ไลค์ลิฮูด เพราะโปรแกรมนี้สามารถใช้ได้กับแบบทดสอบที่มีจำนวน ข้อมากและน้อย วิธีการประมาณค่าเริ่มจากรูปแบบการตอบข้อสอบของผู้สอบ ดังนี้

$$X = (x_1, x_2, \dots, x_n)$$

และสำหรับผู้สอบที่มีความสามารถ θ ค่าความเป็นไปได้สามารถแสดงได้ดังนี้

$$P(X|\theta) = \prod_{i=1}^n P_i(\theta)^{x_i} [1-P_i(\theta)]^{1-x_i}$$

ค่าความน่าจะเป็นที่ได้เป็นความน่าจะเป็นที่มีเงื่อนไขว่าทราบค่าความสามารถ θ ซึ่งต่างจากความน่าจะเป็นของรูปแบบ x จากผู้ไม่ทราบค่าความสามารถ θ ซึ่งสุมมาจากประชากรที่มีการกระจายของค่าความสามารถ θ อยู่ในรูปฟังก์ชันความหนาแน่นต่อเนื่อง (Continuous Density $g(\theta)$) ที่อยู่ในรูปของความน่าจะเป็นที่ไม่มีเงื่อนไข และสามารถแสดงได้ดังนี้

$$P(X) = \int_{-\alpha}^{+\alpha} P(X|\theta) g(\theta) d\theta$$

ค่านี้เรียกว่า ความน่าจะเป็นโดยปราศจากเงื่อนไขของ X และเนื่องจากค่าความสามารถ θ ถูก Integrate ออกไป ค่านี้จึงเป็นฟังก์ชันของพารามิเตอร์ของข้อสอบเท่านั้น ในวิธี มาร์จิ้นัล แมกซิมัม ไลค์ลิฮูด ค่าของพารามิเตอร์ของข้อสอบจะพิจารณาใช้ค่าที่ทำให้สมการข้างล่างมีค่าสูงสุด

$$\log L_M = \sum_{i=1}^S r_i \log \bar{P}(X_i)$$

เมื่อ $P(X_i)$ ประมาณได้จากสูตร Gaussian Quadrature

r_i คือ ความถี่ของรูปแบบการตอบข้อสอบรูปแบบ X_i จากจำนวน ผู้สอบทั้งหมด (N) และ

S คือ จำนวนรูปแบบที่เป็นไปได้ทั้งหมดจากข้อมูลการตอบข้อสอบ

ในแบบจำลองโลจิสติก 3 พารามิเตอร์ ค่าสูงสุดของสมการดังกล่าว สำหรับข้อ i หาได้จากสมการโลคัลไลต์ ดังนี้

$$\sum_{k=1}^s \frac{\bar{r}_{ik} - \bar{N}_k P_i(X_k)}{P_i(X_k) [1 - P_i(X_k)]} \frac{d P_i(X_k)}{d \begin{bmatrix} c_i \\ a_i \\ b_i \end{bmatrix}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{เมื่อ } \bar{r}_{ik} = \sum_{j=1}^s r_{ij} X_{1j} P(X_1/X_k) A(X_k) / P(X_1)$$

$$\bar{N}_k = \sum_{j=1}^s r_{1j} P(X_1/X_k) A(X_k) / \bar{P}(X_1)$$

เป็นค่าคาดหวังภายหลังของจำนวนข้อสอบที่ถูก และจำนวนข้อสอบที่ทำที่จุด X_k (x_{1k} คือ คะแนน 0, 1 สำหรับข้อสอบ i ในรูปแบบการตอบ 1)

จากนั้นใช้ขั้นตอนการคำนวณ EM และวิธีการของ Newton-Gauss (Fisher-Scoring) แก้สมการ

4.2 การประมาณค่าพารามิเตอร์ของข้อสอบ และค่าความสามารถของผู้สอบ จากผลการสอบที่มีจำนวนผู้สอบและจำนวนข้อสอบแตกต่างกันด้วยวิธีวิวิธวิธี โดยใช้โปรแกรมคอมพิวเตอร์ภาษา TURBO PASCAL Version 7.0 ซึ่งสามารถเขียนรูปแบบความสัมพันธ์เชิงคณิตศาสตร์ของค่าพารามิเตอร์ ได้ดังนี้ (Warm. 1978 : 51; Hambleton and Swaminathan. 1985 : 146; citing Jensema. 1976)

$$a_i = \frac{D_i(P_i Q_i)^{1/2}}{[(KR-20)(1-C_i)^2 Y_i - D_i^2 P_i Q_i]^{1/2}} \quad ; i=1,2,\dots,n \text{ ข้อ}$$

$$b_i = \frac{Y_i Z_i (1-C_i) (KR-20)^{1/2}}{D_i (P_i Q_i)^{1/2}} \quad ; i=1,2,\dots,n \text{ ข้อ}$$

- c_i = สัดส่วนของผู้ตอบถูกต้องอยู่ปลายล่างของโค้งที่เกิดจากการพล็อต
 จำนวนคนตอบถูกต้องในแต่ละระดับคะแนนรวม หลังจากหัก
 คะแนนของข้อที่กำลังหาค่าพารามิเตอร์ออกจากคะแนนรวมทั้ง
 ฉบับก่อน
 เมื่อ a_i = ค่าอำนาจจำแนกของข้อที่ i
 b_i = ค่าความยากของข้อที่ i
 c_i = ค่าการเดาของข้อที่ i
 C_i = ค่าการเดาของข้อที่ i ในทฤษฎีการทดสอบแบบดั้งเดิม
 (สัดส่วนกลับของจำนวนตัวเลือกในข้อที่ i)
 D_i = ค่าอำนาจจำแนกของข้อที่ i ในทฤษฎีการทดสอบแบบดั้งเดิม
 (Point Biserial Correlation = $[(\bar{X}_i - \bar{X}_0)/SD_x](P'_i/Q_i)^{1/2}$)
 P'_i = ค่าความยากของข้อที่ i ในทฤษฎีการทดสอบแบบดั้งเดิม
 P_i = $(P'_i - C_i)/(1 - C_i)$
 Q'_i = $1 - P'_i$
 Q_i = $1 - P_i$
 Z_i = ค่าของคะแนนมาตรฐานที่มีพื้นที่ใต้โค้งปกติด้านขวามือเท่ากับ P_i
 Y_i = ค่าพิกัดฉาก (Ordinate) ของโค้งที่มีการแจกแจงปกติ ณ จุด
 Z_i
 KR-20 = ค่าความเชื่อมั่นของแบบทดสอบที่คำนวณโดยใช้สูตรครุเดอร์-
 ริชาร์ดสัน สูตรที่ 20
 $\bar{X}_x$ = คะแนนเฉลี่ยของคะแนนรวมของกลุ่มที่ทำข้อสอบข้อ i ถูก
 $\bar{X}_0$ = คะแนนเฉลี่ยของคะแนนรวมของผู้สอบทุกคน
 SD_x = ค่าเบี่ยงเบนมาตรฐานของคะแนนรวมของผู้สอบทุกคน

หาค่าพารามิเตอร์ของข้อสอบที่ประมาณได้ด้วยวิธีวิธีวิสถิติเกี่ยวข้องกับ
 แบบจำลองโลจิสติก 3 พารามิเตอร์ โดยสามารถเขียนเป็นรูปแบบความสัมพันธ์เชิงคณิตศาสตร์
 ของค่าพารามิเตอร์ได้ดังนี้

$$P_i(e_j) = c_i + (1 - c_i) / (1 + e^{-D(\bar{X}_i - \bar{X}_0 - b_i)}); (i=1, 2, \dots, n)$$

ส่วนการประมาณค่าความสามารถของผู้สอบใช้วิธีแมกซิมัม โลคัลลิฮูด แบบมีเงื่อนไข (Conditional Maximum Likelihood Estimation of Ability) ซึ่งเป็นการประมาณค่าความสามารถหลังจากทราบค่าพารามิเตอร์ของข้อสอบแต่ละข้อแล้ว โดยใช้กระบวนการนิวตัน ราฟสัน และกำหนดค่าความสามารถเริ่มต้น สำหรับผู้สอบคนที่ a เป็น

$$e_{0,j} = \ln [r_j / (n - r_j)] = e_0$$

เมื่อ r_j = จำนวนข้อสอบที่ผู้เข้าสอบคนที่ j ตอบถูก
 n = จำนวนข้อสอบที่ผู้เข้าสอบคนที่ j ทำทั้งหมด
 $\ln$ = Natural Logarithm

จะได้ค่า e ของการทำซ้ำครั้งที่ $m+1$ ดังสมการ

$$e_j^{(m+1)} = e_j^{(m)} - h_j^{(m)}$$

โดยที่
$$h_j^{(m)} = \left| \frac{d \ln L}{d e_j} \right| - \left| \frac{d^2 \ln L}{d e_j^2} \right|$$

$$\frac{d \ln L}{d e_j} = \sum_{i=1}^n D \Sigma a_i (U_{i,j} - P_{i,j}) (P_{i,j} - c_i) / P_{i,j} (1 - c_i)$$

$$\frac{d^2 \ln L}{d e_j^2} = \sum_{i=1}^n D^2 \Sigma a_i^2 (P_{i,j} - c_i) (U_{i,j} - P_{i,j}) Q_{i,j} / P_{i,j} (1 - c_i)^2$$

เมื่อ $U_{i,j}$ = ค่าผลการตอบ (ถูกเป็น 1 ผิดเป็น 0) ของผู้เข้าสอบคนที่ j

$$P_{i,j} = \text{โอกาสที่ผู้เข้าสอบคนที่ } j \text{ ทำข้อสอบข้อ } i \text{ ถูกต้อง}$$

$$= (c_i + (1 - c_i) / (1 + \exp^{(-D a_i (e_j - b_i)}))$$

$$D = \text{A Scaling Factor มีค่าเท่ากับ 1.7}$$

$$e_j = \text{ค่าความสามารถของผู้เข้าสอบคนที่ } j$$

- a_i = ค่าอำนาจจำแนกของข้อสอบข้อที่ i ที่ประมาณโดยวิธีฮิวริสติก
 b_i = ค่าความยากของข้อสอบข้อที่ i ที่ประมาณโดยวิธีฮิวริสติก
 c_i = ค่าการเดาของข้อสอบข้อที่ i ที่ประมาณโดยวิธีฮิวริสติก

และเมื่อ $|h_i^{(m)}|$ มีค่าน้อยกว่าหรือเท่ากับ 0.001 จึงสิ้นสุดกระบวนการ (Hambleton and Swaminathan. 1985 : 81 - 88)

4.3 การประมาณค่าพารามิเตอร์ของข้อสอบ และค่าความสามารถของผู้เข้าสอบ จากผลการสอบจากผลการสอบที่มีจำนวนผู้สอบและจำนวนข้อสอบแตกต่างกัน ด้วยวิธีของเบส โดยใช้โปรแกรมคอมพิวเตอร์ ภาษา TURBO PASCAL Version 7.0 ซึ่งมีขั้นตอนดังต่อไปนี้

4.3.1 ค่าเริ่มต้นสำหรับใช้ในการประมาณค่าพารามิเตอร์ โดย

$$e^{(0)} = \ln \left[\frac{P_i}{1-P_i} \right]$$

$$a_i^{(0)} = R_i / (1 - R_i^2)^{1/2}$$

$$b_i^{(0)} = Z_i / R_i$$

$$c_i^{(0)} = 1 / m_i$$

$$P_i = \frac{\sum_{j=1}^n u_{i,j}}{N}$$

$$Q_i = \frac{\sum_{j=1}^n u_{i,j}}{n}$$

- เมื่อ
- $u_{i,j}$ = ผลการตอบข้อที่ i ของผู้สอบคนที่ j
 R_i = Point - Biserial Correlation
 Z_i = ค่า Z ของการแจกแจงปกติมาตรฐานที่พื้นที่ใต้โค้งปกติมาตรฐานด้านขวามือ มีค่าเท่ากับ P_i

$$P_i = \frac{\sum_{j=1}^n u_{i,j}}{N}$$

$$Q_i = \frac{\sum_{j=1}^n u_{i,j}}{n}$$

$$N = \text{จำนวนผู้เข้าสอบทั้งหมด}$$

$$n_i = \text{จำนวนตัวเลือกในข้อสอบข้อที่ } i$$

4.3.2 ประมวลค่าพารามิเตอร์ด้วยวิธีการหาค่าซ้ำ ๆ จนได้ค่าคงที่
(Iterative) ดังนี้

ขั้นที่ 1 ประมวลค่าความสามารถของผู้สอบแต่ละคน

$$e_{i,j}^{(m+1)} = e_{i,j}^{(m)} - g(e_{i,j}^{(m)}) / h(e_{i,j}^{(m)})$$

ขั้นที่ 2 ประมวลค่าอำนาจจำแนก (a_i) ค่าความยาก (b_i)
และค่าการเดา (c_i) ของข้อสอบ

$$a_i^{(m+1)} = a_i^{(m)} - g(a_i^{(m)}) / h(a_i^{(m)})$$

$$b_i^{(m+1)} = b_i^{(m)} - g(b_i^{(m)}) / h(b_i^{(m)})$$

$$c_i^{(m+1)} = c_i^{(m)} - g(c_i^{(m)}) / h(c_i^{(m)})$$

เมื่อ $e_{i,j}^{(m)} =$ ค่าความสามารถของคนที่ j ที่ได้จากการประมวล
ครั้งที่ m

$a_i^{(m)} =$ ค่าอำนาจจำแนกของข้อสอบข้อที่ i ที่ได้จากการ
ประมวลครั้งที่ m

$$b_i^{(m)} = \text{ค่าความยากของข้อสอบข้อที่ } i \text{ ที่ได้จากการประมาณครั้งที่ } m$$

$$c_i^{(m)} = \text{ค่าการเดาของข้อสอบข้อที่ } i \text{ ที่ได้จากการประมาณครั้งที่ } m$$

$$g[\theta_{i,j}^{(m)}] = \sum_{j=1}^n D \sum a_i (P_{i,j} - c_i) (U_{i,j} - P_{i,j}) / P_{i,j} (1 - c_i) - \theta_{i,j}$$

$$h[\theta_{i,j}^{(m)}] = \sum_{j=1}^n D^2 \sum [a_i^2 (P_{i,j} - c_i) Q_{i,j} / (1 - c_i)^2 P_{i,j}] [(U_{i,j} c_i / P_{i,j}) - P_{i,j} - 1]$$

$$g[a_{i,j}^{(m)}] = \sum_{j=1}^n D \sum (\theta_{i,j} - b_i) (P_{i,j} - c_i) (U_{i,j} - P_{i,j}) / P_{i,j} (1 - c_i) \\ + (v_i - 1) / a_i - a_i / w_i$$

$$h[a_{i,j}^{(m)}] = \sum_{j=1}^n D^2 \sum (\theta_{i,j} - b_i)^2 (P_{i,j} - c_i) (U_{i,j} c_i - P_{i,j}^2) Q_{i,j} / P_{i,j} (1 - c_i)^2 \\ - (v_i - 1) / a_i - 1 / w_i$$

$$g[b_{i,j}^{(m)}] = -D \sum a_i (P_{i,j} - c_i) (U_{i,j} - P_{i,j}) / P_{i,j} (1 - c_i)$$

$$h[b_{i,j}^{(m)}] = \sum_{j=1}^n D^2 \sum a_i^2 (P_{i,j} - c_i) (U_{i,j} c_i - P_{i,j}^2) Q_{i,j}^2 / P_{i,j} (1 - c_i)^2$$

$$g[c_{i,j}^{(m)}] = \sum_{j=1}^n D \sum (U_{i,j} - P_{i,j}) / P_{i,j} (1 - c_i) + s_i / c_i - t_i / (1 - c_i)$$

$$h[c_{i,j}^{(m)}] = \sum_{j=1}^n D^2 \sum [U_{i,j} (2P_{i,j} - 1) - P_{i,j}^2] / P_{i,j}^2 (1 - c_i)^2 \\ - s_i / c_i^2 + t_i / (1 - c_i)^2$$

เมื่อ	$U_{i,j}$	=	ผลการตอบข้อที่ i ของผู้สอบคนที่ j
	D	=	A Scaling Factor มีค่าเท่ากับ 1.7
	$P_{i,j}$	=	$c_i + (1 - c_i) / (1 + \exp(\frac{D a_i (e_{i,j} - b_i)})$
	v_i	=	Degrees of Freedom ของ a_i จาก Prior Distribution
	w_i	=	Scaling Factor ของ a_i จาก Prior Distribution
	s_i	=	ค่าพารามิเตอร์ตัวที่หนึ่งของ c_i จาก Prior Distribution
	t_i	=	ค่าพารามิเตอร์ตัวที่สองของ c_i จาก Prior Distribution

การประมาณค่าพารามิเตอร์จะดำเนินการตามขั้นที่หนึ่งและขั้นที่สองซ้ำ ๆ จนกว่าจะได้ค่าคงที่ค่าอำนาจจำแนก (a_i) ค่าความยาก (b_i) ค่าการเดา (c_i) และค่าความสามารถ ($e_{i,j}$) ที่ได้ในครั้งสุดท้าย คือ ค่าที่ประมาณได้จากวิธีของเบส์เมื่อได้มีการกำหนดการแจกแจงเริ่มแรกของค่าอำนาจจำแนกด้วยการแจกแจงแบบไคว์ (Chi Distribution) และกำหนดการแจกแจงเริ่มแรกของค่าการเดาด้วยการแจกแจงแบบเบต้า (Beta Distribution) (Swaminathan and Gifford. 1986 : 589 - 601)

5. การคำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบจากการประมาณค่าพารามิเตอร์แต่ละวิธี ณ ระดับความสามารถต่าง ๆ โดยการนำเอาค่าพารามิเตอร์ของข้อสอบในแบบทดสอบจากการประมาณค่าแต่ละวิธีมาคำนวณหาค่าฟังก์ชันสารสนเทศของข้อสอบในแบบทดสอบ ณ ระดับความสามารถต่าง ๆ 31 ระดับ คือ -3.0, -2.8, -2.6, -2.4, -2.2, -2.0, -1.8, -1.6, -1.4, -1.2, -1.0, -0.8, -0.6, -0.4, -0.2, 0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.6, 2.8 และ 3.0 โดยใช้สูตรในการคำนวณดังนี้

ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ระดับความสามารถ $e_{i,j}$ [$I(e_{i,j})$]

คือ

$$I(e_{i,j}) = \frac{\sum_{i=1}^n P_i(e_{i,j})^2}{P_i(e_{i,j}) Q_i(e_{i,j})}$$

เมื่อ i แทน ข้อสอบข้อที่ 1, 2, ..., n

$$P'_i(\theta_i) = Da_i Q_i (P_i - c_i) / (1 - c_i)$$

$$P_i(\theta_i) = c_i + (1 - c_i) / (1 + e^{-Da_i(\theta_i - b_i)})$$

$$Q_i(\theta_i) = (1 - c_i) / (1 + e^{Da_i(\theta_i - b_i)})$$

6. เปรียบเทียบความแตกต่างของค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบ ซึ่งประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน โดยใช้สถิติทดสอบแบบ Repeated Measure (AxS) Design (Keppel. 1973 : 401 - 402) ถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติแล้ว ทดสอบความแตกต่างรายคู่ ด้วยวิธีการของเชฟเฟ (Scheffe's Test) (ลิวน ศาสยศ และอังคณา ศาสยศ. 2536 : 119) โดยใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0

7. เปรียบเทียบความแตกต่างของค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบ ซึ่งประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน โดยใช้สถิติทดสอบแบบ Repeated Measure (AxS) Design (Keppel. 1973 : 401 - 402) ถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติแล้ว ทดสอบความแตกต่างรายคู่ด้วยวิธีการของเชฟเฟ (Scheffe's Test) (ลิวน ศาสยศ และอังคณา ศาสยศ. 2536 : 119) โดยใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0

8. การคำนวณค่าประสิทธิภาพสัมพัทธ์ของวิธีประมาณค่าพารามิเตอร์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ระหว่างวิธีประมาณค่าพารามิเตอร์วิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวริสติก และวิธีของเบส์ ณ ระดับความสามารถต่าง ๆ (θ_i) โดยนำเอาค่าฟังก์ชันสารสนเทศของแบบทดสอบที่ได้จากการประมาณค่าแต่ละวิธี เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ มาคำนวณ โดยใช้สูตรในการคำนวณดังนี้

$$RE(y, x) = \frac{I(\theta, y)}{I(\theta, x)}$$

9. ทดสอบค่าประสิทธิภาพสัมพัทธ์ของวิธีประมาณค่าพารามิเตอร์ ดังนี้

เมื่อ $RE(y, x) = 1$ แสดงว่า แบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวีรติค และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีคุณภาพเท่ากัน

$RE(y, x) > 1$ แสดงว่า แบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวีรติค และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีคุณภาพไม่เท่ากัน

$RE(y, x) < 1$ แสดงว่า แบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวีรติค และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีคุณภาพไม่เท่ากัน

บทที่ 4
ผลการวิเคราะห์ข้อมูล

การศึกษาเพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยเทคนิควิซมอนด์ คาร์โล ครั้งนี้ เพื่อให้เข้าใจตรงกันในการแปลผลการวิเคราะห์ข้อมูล ผู้วิจัยจึงได้กำหนดสัญลักษณ์และอักษรย่อต่าง ๆ ไว้ดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

a	แทน	ค่าอำนาจจำแนกของข้อสอบ
b	แทน	ค่าความยากของข้อสอบ
c	แทน	ค่าการเดาของข้อสอบ
e	แทน	ค่าความสามารถของผู้สอบ
$\bar{a}$	แทน	ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ
$\bar{b}$	แทน	ค่าความยากเฉลี่ยของข้อสอบ
$\bar{c}$	แทน	ค่าการเดาเฉลี่ยของข้อสอบ
$\bar{e}$	แทน	ค่าความสามารถเฉลี่ยของผู้สอบ
a_{SD}	แทน	ค่าความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบ
b_{SD}	แทน	ค่าความเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบ
c_{SD}	แทน	ค่าความเบี่ยงเบนมาตรฐานของค่าการเดาของข้อสอบ
e_{SD}	แทน	ค่าความเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบ
SS	แทน	ผลบวกกำลังสองของคะแนนจากแบบทดสอบ
MS_B	แทน	ความแปรปรวนของคะแนนจากแบบทดสอบระหว่างกลุ่ม
MS_W	แทน	ความแปรปรวนของคะแนนจากแบบทดสอบภายในกลุ่ม
df	แทน	ช่วงชั้นของความเป็นอิสระ
F-test	แทน	ค่าวิกฤตของ F

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษารั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูล ดังนี้

1. ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ

2. ค่าสถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ซึ่งประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

3. ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบซึ่งประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

4. ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบซึ่งประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน

5. ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่ได้จากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความสามารถต่าง ๆ

6. ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่ได้จากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ

7. ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่ได้จากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความสามารถต่าง ๆ

8. ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่ได้จากวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ

สำหรับข้อมูล โปรแกรมที่ใช้ในการวิเคราะห์ข้อมูล และผลการวิเคราะห์ข้อมูลอย่างละเอียด ผู้วิจัยได้สำเนาไว้ในแผ่นเทปแม่เหล็ก (Diskette) ผู้ที่สนใจสามารถติดต่อผู้วิจัยได้ที่สำนักงาน ก.ค. กระทรวงศึกษาธิการ

ผลการวิเคราะห์ข้อมูล

1. ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ

เนื่องจากข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม ได้กำหนดว่า แบบทดสอบต้องมีคุณสมบัติการวัดเพียงมิติเดียว และมีความเป็นอิสระในการตอบข้อสอบ ดังนั้นผู้วิจัยจึงนำเสนอผลการตอบแบบทดสอบมาตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว โดยการวิเคราะห์องค์ประกอบ ได้ผลดังปรากฏในตาราง 5

ตาราง 5 ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.00 ของแบบทดสอบ

องค์ประกอบ	ค่าไอเกน	องค์ประกอบ	ค่าไอเกน	องค์ประกอบ	ค่าไอเกน
1	5.786	2	1.887	3	1.494
4	1.364	5	1.253	6	1.234
7	1.161	8	1.155	9	1.148
10	1.139	11	1.124	12	1.116
13	1.106	14	1.094	15	1.090
16	1.088	17	1.080	18	1.077
19	1.073	20	1.062	21	1.059
22	1.057	23	1.048	24	1.045
25	1.040	26	1.035	27	1.030
28	1.026	29	1.020	30	1.017
31	1.014	32	1.008		

จากตาราง 5 ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบทดสอบ ด้วยวิธีการวิเคราะห์องค์ประกอบ พบว่า เมื่อนำแบบทดสอบมาวิเคราะห์องค์ประกอบ แบบทดสอบซึ่งประกอบด้วยข้อสอบ จำนวน 100 ข้อ มีองค์ประกอบที่มีค่าไอเกนเกิน 1.00 อยู่

ภาพประกอบ 6 ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.00 ของแบบทดสอบ

2. ค่าสถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิชูด วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

ภายหลังการตรวจสอบและพบว่า แบบทดสอบมีคุณสมบัติการวัดเพียงมิติเดียว และมีความเป็นอิสระในการตอบข้อสอบ ผู้วิจัยได้ทำการสุ่มตัวอย่างแบบใส่คืน เพื่อสุ่มตัวอย่างผลการตอบของผู้สอบออกเป็น 13 ขนาด โดยกลุ่มตัวอย่างแต่ละขนาดกลุ่มต้องตอบข้อสอบที่มีขนาดต่างกัน 6 ขนาด โดยการทดลองซ้ำกลุ่มละ 40 ครั้ง ด้วยเทคนิควิธีมอนติคาร์โล เพื่อประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบด้วยวิธีแมกซีมัม ไลค์ลิชูด วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน โดยผู้วิจัยได้แสดงค่าสถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ดังนี้

2.1 ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิชูด เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 6

ตาราง 6 ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแบบทรีแมตริกซ์
ไคลด์ริสต์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
200	50	-2.037 ถึง 3.061	0.325	1.143	26, 37, 45
	60	-1.808 ถึง 3.056	0.349	1.076	32, 44, 45
	70	-2.116 ถึง 3.047	0.344	1.031	7, 37, 50, 52, 63
	80	-2.100 ถึง 3.054	0.366	0.964	8, 42, 53, 56, 60, 72
	90	-2.244 ถึง 3.045	0.343	1.022	8, 47, 59, 62, 68, 75
	100	-2.155 ถึง 3.067	0.372	0.964	8, 42, 45, 51, 64, 65, 68, 90, 91
300	50	-2.839 ถึง 2.869	0.316	1.099	22, 34, 37
	60	-2.413 ถึง 2.961	0.340	1.051	27, 41, 44
	70	-2.319 ถึง 2.930	0.302	1.050	32, 47, 50, 51
	80	-2.216 ถึง 2.968	0.319	1.001	37, 53, 56, 60
	90	-2.243 ถึง 2.969	0.286	1.030	42, 59, 62, 68
	100	-2.350 ถึง 2.989	0.319	1.069	45, 65, 68, 75
400	50	-2.304 ถึง 2.506	0.333	0.970	33, 34, 37, 45, 46
	60	-5.415 ถึง 2.485	0.233	1.217	40, 41, 44, 45
	70	-3.338 ถึง 2.472	0.249	1.080	7, 46, 47, 50, 52, 63
	80	-1.743 ถึง 2.480	0.322	0.975	8, 52, 53, 72, 76
	90	-2.755 ถึง 2.460	0.316	0.932	8, 42, 58, 59, 62, 68, 82, 86
	100	-2.415 ถึง 2.704	0.346	0.958	8, 45, 48, 64, 65, 68, 75, 91, 96

ตาราง 6 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	s_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
500	50	-2.189 ถึง 2.606	0.328	1.078	22, 26, 34, 37
	60	-2.277 ถึง 2.640	0.294	1.095	32, 41, 44
	70	-2.244 ถึง 2.645	0.280	1.076	37, 47, 50, 52
	80	-2.515 ถึง 2.631	0.271	1.039	42, 53, 56, 60
	90	-2.539 ถึง 2.630	0.283	1.038	42, 47, 52, 59, 62, 68
	100	-2.465 ถึง 2.630	0.317	1.063	45, 51, 58, 65, 68, 75
600	50	-1.944 ถึง 2.718	0.336	1.076	26, 33, 34, 37
	60	-2.838 ถึง 2.674	0.305	1.138	40, 41, 44
	70	-2.946 ถึง 2.670	0.260	1.154	46, 47, 50, 52
	80	-2.563 ถึง 2.693	0.327	1.020	42, 52, 53, 56, 60
	90	-2.584 ถึง 2.680	0.323	1.036	47, 58, 59, 62, 68
	100	-2.615 ถึง 2.666	0.369	1.060	51, 64, 65, 68, 75
700	50	-2.244 ถึง 2.992	0.221	1.184	34, 37
	60	-2.306 ถึง 3.017	0.258	1.136	41, 44
	70	-2.168 ถึง 2.984	0.238	1.107	47, 50, 52
	80	-2.149 ถึง 2.988	0.280	1.041	37, 53, 56, 60
	90	-2.167 ถึง 3.014	0.253	1.062	42, 59, 62, 68
	100	-2.229 ถึง 2.968	0.284	1.092	45, 65, 68, 75
800	50	-2.298 ถึง 2.837	0.302	1.143	34, 37
	60	-2.462 ถึง 2.875	0.323	1.096	41, 44
	70	-3.590 ถึง 2.886	0.287	1.112	47, 50, 52, 63
	80	-2.027 ถึง 2.859	0.346	0.966	37, 53, 56, 60, 72
	90	-3.184 ถึง 2.847	0.262	1.091	59, 62, 68
	100	-2.770 ถึง 2.826	0.336	1.031	45, 65, 68, 75

ตาราง 6 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
900	50	-2.333 ถึง 2.775	0.380	1.039	22, 26, 33, 34, 37
	60	-3.459 ถึง 2.828	0.277	1.206	27, 41, 44
	70	-2.923 ถึง 2.806	0.260	1.165	32, 47, 50, 52
	80	-3.479 ถึง 2.811	0.262	1.145	37, 53, 56, 60
	90	-3.490 ถึง 2.817	0.254	1.145	42, 59, 62, 68
	100	-3.027 ถึง 2.829	0.303	1.141	45, 65, 68, 75
1,000	50	-2.069 ถึง 2.906	0.385	1.028	22, 25, 34, 37, 45
	60	-1.994 ถึง 2.894	0.392	0.994	27, 31, 41, 44, 55
	70	-2.166 ถึง 2.909	0.345	0.998	32, 36, 47, 50, 52, 63
	80	-3.783 ถึง 2.870	0.308	1.060	37, 41, 53, 56, 60
	90	-2.090 ถึง 2.879	0.320	0.984	42, 46, 59, 62, 68, 82
	100	-2.235 ถึง 2.838	0.342	1.031	45, 50, 75, 91
1,500	50	-2.930 ถึง 3.276	0.294	1.232	26, 34, 37
	60	-2.643 ถึง 3.282	0.307	1.190	32, 41, 44
	70	-2.356 ถึง 3.200	0.343	1.064	7, 32, 37, 47, 50, 52
	80	-2.216 ถึง 3.128	0.352	1.006	8, 37, 42, 53, 56, 60
	90	-2.430 ถึง 3.099	0.294	1.073	42, 47, 59, 62, 68
	100	-2.809 ถึง 3.094	0.321	1.129	45, 51, 65, 68, 75
2,000	50	-2.950 ถึง 2.966	0.242	1.232	34, 37
	60	-2.619 ถึง 2.978	0.278	1.158	41, 44
	70	-2.401 ถึง 2.956	0.244	1.143	47, 50, 52
	80	-2.327 ถึง 2.954	0.299	1.054	37, 53, 56, 60
	90	-2.351 ถึง 2.934	0.272	1.073	42, 59, 62, 68
	100	-2.453 ถึง 2.890	0.301	1.117	45, 65, 68, 75

ตาราง 6 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
2,500	50	-2.696 ถึง 2.989	0.325	1.157	26, 34, 37
	60	-1.930 ถึง 1.979	0.392	1.036	32, 41, 44, 55
	70	-3.349 ถึง 2.975	0.288	1.139	37, 47, 50, 52
	80	-2.262 ถึง 2.971	0.337	1.016	42, 53, 56, 60, 72
	90	-2.235 ถึง 2.977	0.324	1.024	47, 59, 62, 68, 82
	100	-2.473 ถึง 2.968	0.344	1.079	51, 65, 68, 75, 91
3,000	50	-3.259 ถึง 3.170	0.304	1.248	26, 34, 37
	60	-2.568 ถึง 3.235	0.325	1.194	32, 41, 44
	70	-3.570 ถึง 3.216	0.305	1.176	32, 37, 47, 50, 52
	80	-2.997 ถึง 3.181	0.329	1.085	37, 42, 53, 56, 60
	90	-2.571 ถึง 3.206	0.315	1.080	42, 47, 59, 62, 68
	100	-2.931 ถึง 3.198	0.337	1.139	45, 51, 65, 68, 75

จากตาราง 6 พบว่า ค่าอำนาจจำแนกของข้อสอบมีค่าต่ำสุดเท่ากับ -5.415 เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาด 60 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน และข้อสอบมีขนาด 60 ข้อ ค่าอำนาจจำแนกของข้อสอบจะมีค่าสูงสุดเท่ากับ 3.282

สำหรับค่าอำนาจจำแนกเฉลี่ยของข้อสอบมีค่าต่ำสุดเท่ากับ 0.221 เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาด 50 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน หรือ 2,500 คน และข้อสอบมีขนาด 60 ข้อ ค่าอำนาจจำแนกเฉลี่ยของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.392 ส่วนความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบมีค่าต่ำสุดเท่ากับ 0.932 เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาด 90 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน ข้อสอบมีขนาด 50 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่า

อำนาจจำแนกของข้อสอบจะมีค่าสูงสุดเท่ากับ 1.248

ในการประมาณค่าอำนาจจำแนกของข้อสอบด้วยวิธีแมกซีมัม ไลค์ลิสต์ นี้พบว่า มีข้อสอบที่ไม่เหมาะสมกับโมเดลน้อยที่สุด 2 ข้อ เมื่อกลุ่มตัวอย่างมีขนาด 700 คน 800 คน และ 2,000 คน ข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และ 400 คน ข้อสอบมีขนาด 100 ข้อ แล้วพบว่า มีข้อสอบที่ไม่เหมาะสมกับโมเดลมากที่สุด 9 ข้อ

2.2 ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธี
อิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 7

ตาราง 7 ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อ
ขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
200	50	-4.128 ถึง 3.259	0.320	1.393	-
	60	-6.459 ถึง 5.739	0.162	1.900	-
	70	-4.813 ถึง 9.752	0.409	1.917	-
	80	-8.665 ถึง 10.617	0.538	2.320	-
	90	-9.650 ถึง 4.696	0.080	2.063	-
	100	-12.909 ถึง 6.116	0.515	2.266	-
300	50	-5.096 ถึง 4.747	-0.044	1.757	-
	60	-5.104 ถึง 5.084	-0.071	1.651	-
	70	-3.796 ถึง 4.569	0.093	1.520	-
	80	-3.048 ถึง 5.607	0.263	1.646	-
	90	-3.204 ถึง 4.717	0.126	1.520	-
	100	-15.009 ถึง 3.157	0.160	2.237	-

ตาราง 7 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
400	50	-1.622 ถึง 3.000	0.286	2.383	-
	60	-1.750 ถึง 6.345	0.559	1.191	-
	70	-1.462 ถึง 1.669	0.509	0.511	-
	80	-1.776 ถึง 2.870	0.479	0.863	-
	90	-1.938 ถึง 3.869	0.425	1.024	-
	100	-7.638 ถึง 13.771	0.394	2.726	-
500	50	-1.483 ถึง 1.916	0.483	0.589	-
	60	-1.587 ถึง 1.882	0.379	0.665	-
	70	-1.592 ถึง 1.776	0.436	0.681	-
	80	-1.826 ถึง 1.878	0.497	0.715	-
	90	-1.816 ถึง 2.285	0.385	0.829	-
	100	-5.578 ถึง 3.994	0.049	1.645	-
600	50	-2.825 ถึง 1.768	0.095	1.181	-
	60	-1.980 ถึง 2.246	0.243	0.968	-
	70	-1.865 ถึง 2.146	0.366	0.885	-
	80	-2.823 ถึง 3.534	0.370	1.132	-
	90	-1.809 ถึง 2.011	0.449	0.888	-
	100	-8.909 ถึง 7.714	-0.010	2.693	-
700	50	-1.879 ถึง 2.192	0.268	0.939	-
	60	-1.900 ถึง 2.810	0.335	1.085	-
	70	-3.691 ถึง 6.267	0.329	1.026	-
	80	-1.869 ถึง 2.174	0.326	0.927	-
	90	-1.877 ถึง 2.713	0.413	0.878	-
	100	-19.439 ถึง 4.728	-0.010	2.693	-

ตาราง 7 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
800	50	-1.996 ถึง 9.755	0.574	1.565	-
	60	-1.987 ถึง 2.602	0.444	0.869	-
	70	-1.796 ถึง 2.655	0.521	0.696	-
	80	-2.604 ถึง 3.601	0.472	1.111	-
	90	-1.800 ถึง 2.918	0.614	0.868	-
	100	-4.176 ถึง 9.669	0.444	2.618	-
900	50	-1.981 ถึง 2.655	0.424	0.990	-
	60	-1.896 ถึง 2.982	0.566	0.935	-
	70	-1.544 ถึง 2.910	0.543	0.816	-
	80	-1.588 ถึง 2.830	0.576	0.847	-
	90	-1.960 ถึง 2.622	0.536	0.885	-
	100	-8.390 ถึง 7.295	0.498	2.103	-
1,000	50	-1.994 ถึง 2.674	0.560	0.998	-
	60	-1.993 ถึง 2.560	0.438	1.005	-
	70	-1.899 ถึง 2.550	0.409	0.926	-
	80	-1.992 ถึง 2.763	0.533	0.947	-
	90	-1.808 ถึง 4.192	0.535	0.913	-
	100	-4.195 ถึง 4.498	0.288	1.583	-
1,500	50	-1.857 ถึง 2.431	0.604	0.698	-
	60	-1.649 ถึง 2.373	0.603	0.708	-
	70	-1.819 ถึง 2.282	0.485	0.803	-
	80	-1.672 ถึง 2.718	0.615	0.776	-
	90	-1.645 ถึง 2.984	0.579	0.727	-
	100	-8.286 ถึง 12.866	0.123	2.381	-

ตาราง 7 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
2,000	50	-1.700 ถึง 2.126	0.558	0.728	-
	60	-1.990 ถึง 2.524	0.507	0.990	-
	70	-1.681 ถึง 2.450	0.531	0.770	-
	80	-1.917 ถึง 2.635	0.580	0.827	-
	90	-1.878 ถึง 2.956	0.530	0.872	-
	100	-8.241 ถึง 6.782	0.210	1.944	-
2,500	50	-1.769 ถึง 2.703	0.619	0.804	-
	60	-1.567 ถึง 2.620	0.507	0.785	-
	70	-1.747 ถึง 2.005	0.437	0.753	-
	80	-1.968 ถึง 2.540	0.524	0.832	-
	90	-1.746 ถึง 2.685	0.513	0.791	-
	100	-3.497 ถึง 5.875	0.439	1.699	-
3,000	50	-1.975 ถึง 2.828	0.448	0.952	-
	60	-1.983 ถึง 2.901	0.493	1.028	-
	70	-1.897 ถึง 2.590	0.491	0.914	-
	80	-1.898 ถึง 2.359	0.532	0.867	-
	90	-1.932 ถึง 2.559	0.552	0.841	-
	100	-2.806 ถึง 6.136	0.548	1.682	-

จากตาราง 7 พบว่า ค่าอำนาจจำแนกของข้อสอบมีค่าต่ำสุดเท่ากับ -19.439 เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาด 100 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาด 100 ข้อ ค่าอำนาจจำแนกของข้อสอบจะมีค่าสูงสุดเท่ากับ 13.771 สำหรับค่าอำนาจจำแนกเฉลี่ยของข้อสอบมีค่าต่ำสุดเท่ากับ -0.071 เมื่อ

กลุ่มตัวอย่างมีขนาด 300 คน และข้อสอบมีขนาด 60 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน และข้อสอบมีขนาด 50 ข้อ ค่าอำนาจจำแนกเฉลี่ยของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.619 ส่วนความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบมีค่าต่ำสุดเท่ากับ 0.511 เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาด 70 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 400 คน ข้อสอบมีขนาด 100 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบจะมีค่าสูงสุดเท่ากับ 2.726 และไม่พบว่าข้อสอบที่ไม่เหมาะสมกับโมเดลจากการประมาณค่าด้วยวิธีนี้

2.3 ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 8

ตาราง 8 ค่าสถิติพื้นฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
200	50	-0.786 ถึง 5.274	0.986	0.893	-
	60	-1.412 ถึง 2.752	0.842	0.529	-
	70	-0.932 ถึง 4.791	0.992	0.681	-
	80	-1.165 ถึง 3.846	0.837	0.558	-
	90	-1.161 ถึง 9.074	0.938	1.191	-
	100	-8.980 ถึง 4.728	0.824	1.142	-
300	50	-0.007 ถึง 1.319	0.814	0.237	-
	60	-0.051 ถึง 1.291	0.779	0.273	-
	70	-0.357 ถึง 1.326	0.849	0.319	-
	80	-1.325 ถึง 4.384	0.820	0.533	-
	90	-0.007 ถึง 3.733	0.897	0.614	-
	100	-1.352 ถึง 1.445	0.778	0.359	-

ตาราง 8 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
400	50	-1.503 ถึง 1.768	0.373	2.226	-
	60	-1.823 ถึง 1.305	0.707	0.880	-
	70	-1.047 ถึง 1.162	0.631	0.316	-
	80	-9.099 ถึง 2.387	0.562	1.209	-
	90	-1.009 ถึง 1.580	0.813	0.350	-
	100	-0.510 ถึง 2.530	0.841	0.370	-
500	50	-3.947 ถึง 2.559	0.605	0.824	-
	60	-0.576 ถึง 1.276	0.718	0.436	-
	70	-4.702 ถึง 1.279	0.630	0.850	-
	80	-3.185 ถึง 7.209	0.697	1.156	-
	90	-2.203 ถึง 5.887	0.839	0.897	-
	100	-3.205 ถึง 1.944	0.718	0.580	-
600	50	0.052 ถึง 1.684	0.699	0.325	-
	60	-1.631 ถึง 1.468	0.680	0.473	-
	70	-1.812 ถึง 1.844	0.570	0.391	-
	80	0.044 ถึง 6.626	0.643	0.734	-
	90	-0.437 ถึง 6.442	0.627	0.699	-
	100	-1.674 ถึง 2.864	0.607	0.530	-
700	50	-0.791 ถึง 1.098	0.609	0.366	-
	60	-0.860 ถึง 2.655	0.661	0.191	-
	70	-0.915 ถึง 2.962	0.795	0.196	-
	80	-0.803 ถึง 2.951	0.732	0.209	-
	90	-0.884 ถึง 2.778	0.783	0.223	-
	100	0.154 ถึง 1.205	0.716	0.219	-

ตาราง 8 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะกับโมเดล
800	50	0.195 ถึง 1.068	0.658	0.189	-
	60	-0.478 ถึง 2.796	0.904	0.171	-
	70	-0.553 ถึง 1.068	0.635	0.178	-
	80	-0.889 ถึง 2.554	0.595	0.213	-
	90	-1.217 ถึง 2.658	0.701	0.277	-
	100	-1.599 ถึง 2.968	0.698	0.282	-
900	50	0.004 ถึง 1.905	0.680	0.292	-
	60	-0.806 ถึง 1.261	0.624	0.304	-
	70	-1.276 ถึง 1.149	0.622	0.352	-
	80	-0.556 ถึง 12.800	0.806	1.388	-
	90	-1.067 ถึง 1.138	0.634	0.315	-
	100	-0.869 ถึง 1.174	0.639	0.279	-
1,000	50	-1.607 ถึง 1.240	0.541	0.591	-
	60	-2.151 ถึง 2.582	0.592	0.656	-
	70	-1.696 ถึง 3.623	0.788	0.621	-
	80	-1.952 ถึง 1.404	0.531	0.571	-
	90	-0.201 ถึง 1.330	0.689	0.368	-
	100	-0.767 ถึง 2.436	0.722	0.461	-
1,500	50	-1.779 ถึง 1.953	0.541	0.638	-
	60	-1.454 ถึง 1.433	0.548	0.460	-
	70	-2.275 ถึง 2.539	0.515	0.719	-
	80	-1.597 ถึง 1.449	0.515	0.539	-
	90	-1.441 ถึง 2.541	0.474	0.544	-
	100	-1.094 ถึง 3.104	0.421	0.452	-

ตาราง 8 (ต่อ)

จำนวนคน	จำนวนข้อ	a	$\bar{a}$	a_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
2,000	50	-0.026 ถึง 0.405	0.240	0.097	-
	60	-0.050 ถึง 0.407	0.238	0.101	-
	70	-0.049 ถึง 0.398	0.229	0.103	-
	80	-0.061 ถึง 0.395	0.228	0.100	-
	90	-1.228 ถึง 2.285	0.686	0.632	-
	100	-1.710 ถึง 3.482	0.708	0.821	-
2,500	50	-1.763 ถึง 2.608	0.585	0.908	-
	60	-0.999 ถึง 1.279	0.362	0.298	-
	70	-0.479 ถึง 1.051	0.490	0.280	-
	80	-1.466 ถึง 2.344	0.580	0.622	-
	90	-1.870 ถึง 2.416	0.656	0.758	-
	100	-1.589 ถึง 1.201	0.387	0.413	-
3,000	50	-1.799 ถึง 2.549	0.287	0.695	-
	60	-0.893 ถึง 1.007	0.436	0.279	-
	70	-1.010 ถึง 1.762	0.428	0.395	-
	80	-0.898 ถึง 0.930	0.415	0.252	-
	90	-1.105 ถึง 1.726	0.522	0.660	-
	100	-1.130 ถึง 1.305	0.371	0.319	-

จากตาราง 8 พบว่า ค่าอำนาจจำแนกของข้อสอบมีค่าต่ำสุดเท่ากับ -9.099 เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาด 80 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาด 80 ข้อ ค่าอำนาจจำแนกของข้อสอบจะมีค่าสูงสุดเท่ากับ 12.800 สำหรับค่าอำนาจจำแนกเฉลี่ยของข้อสอบมีค่าต่ำสุดเท่ากับ 0.228 เมื่อกลุ่ม

ตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด 80 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาด 70 ข้อ ค่าอำนาจจำแนกเฉลี่ยของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.992 ส่วนความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบมีค่าต่ำสุดเท่ากับ 0.100 เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด 80 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 900 คน ข้อสอบมีขนาด 80 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบจะมีค่าสูงสุดเท่ากับ 2.726 และไม่พบว่ามีข้อสอบที่ไม่เหมาะสมกับโมเดลจากการประมาณค่าด้วยวิธีนี้

2.4 ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัมไลค์ลิชูด เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 9

ตาราง 9 ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกทิมม ไลค์ลิสต์
เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{SD}	ข้อที่ไม่เหมาะกับโมเดล
200	50	-7.892 ถึง 4.678	-0.601	2.505	26, 37, 45
	60	-8.353 ถึง 4.064	-0.722	2.404	32, 44, 55
	70	-8.667 ถึง 4.997	-0.772	2.368	7, 37, 50, 52, 63
	80	-8.334 ถึง 4.401	-0.832	2.136	8, 42, 53, 56, 60, 72
	90	-8.058 ถึง 4.570	-0.719	2.142	8, 47, 59, 62, 68, 75
	100	-7.020 ถึง 4.359	-0.787	2.038	8, 42, 45, 51, 64, 65, 68, 90, 91
300	50	-6.384 ถึง 6.626	-0.591	2.422	22, 34, 37
	60	-6.039 ถึง 4.817	-0.711	2.269	27, 41, 44
	70	-6.043 ถึง 4.635	-0.658	2.245	32, 47, 50, 51
	80	-6.279 ถึง 4.423	-0.726	2.177	37, 53, 56, 60
	90	-5.888 ถึง 4.445	-0.620	2.174	42, 59, 62, 68
	100	-5.860 ถึง 4.640	-0.660	2.155	45, 65, 68, 75
400	50	-4.016 ถึง 5.266	-0.464	2.009	33, 34, 37, 45, 46
	60	-4.139 ถึง 12.848	-0.340	2.696	40, 41, 44, 55
	70	-4.173 ถึง 8.029	-0.399	2.469	7, 46, 47, 50, 52, 63
	80	-4.310 ถึง 10.157	-0.584	2.269	8, 52, 53, 72, 76
	90	-4.289 ถึง 6.335	-0.573	2.088	8, 42, 58, 59, 62, 68, 82, 86
	100	-4.278 ถึง 5.022	-0.650	2.030	8, 45, 48, 64, 65, 68, 75, 91, 96

ตาราง 9 (ต่อ)

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{sd}	ข้อที่ไม่เหมาะสมกับโมเดล
500	50	-5.389 ถึง 4.879	-0.604	2.304	22, 26, 34, 37
	60	-5.227 ถึง 5.462	-0.573	2.427	32, 41, 44
	70	-5.233 ถึง 5.619	-0.580	2.459	37, 47, 50, 52
	80	-5.176 ถึง 6.259	-0.564	2.402	42, 53, 56, 60
	90	-5.106 ถึง 5.914	-0.606	2.331	42, 47, 52, 59, 62, 68
	100	-7.862 ถึง 5.485	-0.723	2.419	45, 51, 58, 65, 68, 75
600	50	-4.004 ถึง 5.264	-0.535	2.339	26, 33, 34, 37
	60	-4.457 ถึง 7.438	-0.526	2.606	40, 41, 44
	70	-4.651 ถึง 7.600	-0.475	2.735	46, 47, 50, 52
	80	-4.593 ถึง 6.331	-0.698	2.363	42, 52, 53, 56, 60
	90	-5.825 ถึง 6.587	-0.710	2.419	47, 58, 59, 62, 68
	100	-6.325 ถึง 6.614	-0.820	2.525	51, 64, 65, 68, 75
700	50	-4.582 ถึง 5.379	-0.230	2.568	34, 37
	60	-4.358 ถึง 4.761	-0.479	2.359	41, 44
	70	-4.398 ถึง 4.693	-0.488	2.327	47, 50, 52
	80	-4.433 ถึง 5.185	-0.636	2.166	37, 53, 56, 60
	90	-4.649 ถึง 5.017	-0.565	2.235	42, 59, 62, 68
	100	-4.517 ถึง 5.024	-0.607	2.240	45, 65, 68, 75
800	50	-3.906 ถึง 6.193	-0.383	2.493	34, 37
	60	-4.078 ถึง 6.581	-0.524	2.437	41, 44
	70	-4.009 ถึง 9.065	-0.505	2.478	47, 50, 52, 63
	80	-4.034 ถึง 5.363	-0.702	2.135	37, 53, 56, 60, 72
	90	-5.126 ถึง 8.284	-0.477	2.576	59, 62, 68
	100	-4.816 ถึง 6.993	-0.627	2.339	45, 65, 68, 75

ตาราง 9 (ต่อ)

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
900	50	-5.072 ถึง 4.904	-0.786	2.207	22, 26, 33, 34, 37
	60	-4.981 ถึง 9.512	-0.600	2.814	27, 41, 44
	70	-4.931 ถึง 8.206	-0.591	2.797	32, 47, 50, 52
	80	-4.859 ถึง 9.570	-0.591	2.789	37, 53, 56, 60
	90	-5.613 ถึง 9.741	-0.565	2.817	42, 59, 62, 68
	100	-8.891 ถึง 8.285	-0.717	2.863	45, 65, 68, 75
1,000	50	-5.869 ถึง 4.170	-0.665	2.058	22, 25, 34, 37, 45
	60	-5.610 ถึง 4.645	-0.770	2.057	27, 31, 41, 44, 55
	70	-5.789 ถึง 4.930	-0.720	2.153	32, 36, 47, 50, 52, 63
	80	-5.440 ถึง 9.541	-0.647	2.375	37, 41, 53, 56, 60
	90	-5.445 ถึง 4.420	-0.672	2.151	42, 46, 59, 62, 68, 62
	100	-5.173 ถึง 4.781	-0.679	2.152	45, 50, 75, 91
1,500	50	-4.174 ถึง 7.837	-0.320	2.654	26, 34, 37
	60	-4.027 ถึง 7.279	-0.454	2.609	32, 41, 44
	70	-4.085 ถึง 5.793	-0.620	2.313	7, 32, 37, 50, 52
	80	-4.132 ถึง 5.409	-0.671	2.277	8, 37, 42, 53, 56, 60
	90	-5.120 ถึง 6.192	-0.528	2.450	42, 47, 59, 62, 68
	100	-5.181 ถึง 6.929	-0.576	2.454	45, 51, 65, 68, 75
2,000	50	-4.079 ถึง 7.633	-0.183	2.901	34, 37
	60	-5.604 ถึง 6.545	-0.415	2.771	41, 44
	70	-5.755 ถึง 6.619	-0.414	2.732	47, 50, 52
	80	-5.799 ถึง 5.686	-0.619	2.484	37, 53, 56, 60
	90	-5.857 ถึง 5.671	-0.567	2.514	42, 59, 62, 68
	100	-5.764 ถึง 6.056	-0.589	2.553	45, 65, 68, 75

ในการประมาณค่าความยากของข้อสอบด้วยวิธีแมกซีมัม ไลค์ลิสต์ นี้ พบว่า มีข้อสอบที่ไม่เหมาะกับโมเดลน้อยที่สุด 2 ข้อ เมื่อกลุ่มตัวอย่างมีขนาด 700 คน 800 คน และ 2,000 คน ข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และ 400 คน ข้อสอบมีขนาด 100 ข้อ แล้ว พบว่า มีข้อสอบที่ไม่เหมาะกับโมเดลมากที่สุดจำนวน 9 ข้อ

2.5 ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอีวีวิสต์ค เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 10

ตาราง 10 ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอีวีวิสต์ค เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{SD}	ข้อที่ไม่เหมาะกับโมเดล
200	50	-14.251 ถึง 44.765	2.628	8.820	-
	60	-8.896 ถึง 549.468	10.384	70.806	-
	70	-119.132 ถึง 6.896	-1.539	17.148	-
	80	-57.239 ถึง 9.150	0.561	6.917	-
	90	-168.842 ถึง 14.205	-1.005	19.048	-
	100	-46.459 ถึง 9.926	0.696	5.297	-
300	50	-8.409 ถึง 13.769	1.088	2.691	-
	60	-5.410 ถึง 10.598	1.192	2.101	-
	70	-2.653 ถึง 29.133	1.542	3.717	-
	80	-13.722 ถึง 10.403	1.017	2.338	-
	90	-12.758 ถึง 19.289	1.182	2.855	-
	100	-14.572 ถึง 14.033	1.173	2.542	-

ตาราง 10 (ต่อ)

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
400	50	-4.296 ถึง 15.194	1.062	2.408	-
	60	-5.559 ถึง 33.547	1.351	4.461	-
	70	-52.115 ถึง 14.015	0.260	6.737	-
	80	-6.345 ถึง 21.000	1.399	3.442	-
	90	-14.264 ถึง 50.792	1.678	6.113	-
	100	-21.302 ถึง 19.476	1.299	3.929	-
500	50	-17.137 ถึง 8.915	0.963	2.975	-
	60	-16.289 ถึง 8.661	1.019	2.661	-
	70	-8.401 ถึง 8.233	0.975	2.039	-
	80	-7.073 ถึง 7.786	1.050	1.630	-
	90	-7.030 ถึง 10.538	1.239	1.884	-
	100	-5.221 ถึง 17.292	1.377	2.239	-
600	50	-7.700 ถึง 5.969	1.115	1.666	-
	60	-6.414 ถึง 6.090	1.194	1.571	-
	70	-3.835 ถึง 19.883	1.488	2.623	-
	80	-2.998 ถึง 34.110	1.647	3.911	-
	90	-19.825 ถึง 8.404	1.031	2.637	37
	100	-11.645 ถึง 8.214	1.019	1.577	-
700	50	-7.679 ถึง 4.249	0.916	1.639	-
	60	-15.578 ถึง 4.323	0.826	2.398	-
	70	-1.753 ถึง 12.830	1.372	2.188	-
	80	-43.483 ถึง 4.709	0.538	5.108	-
	90	-18.738 ถึง 5.966	0.932	2.405	-
	100	-10.913 ถึง 5.374	1.019	1.577	-

ตาราง 10 (ต่อ)

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{so}	ข้อที่ไม่เหมาะสมกับโมเดล
800	50	-18.873 ถึง 6.111	0.672	3.254	-
	60	-4.745 ถึง 12.625	1.278	2.048	-
	70	-3.731 ถึง 7.100	1.197	1.460	-
	80	-3.122 ถึง 7.004	1.778	1.315	-
	90	-3.024 ถึง 8.688	1.261	1.522	-
	100	-2.199 ถึง 6.174	1.179	1.152	-
900	50	-4.648 ถึง 9.707	1.310	1.953	-
	60	-4.146 ถึง 11.190	1.323	1.910	-
	70	-3.144 ถึง 10.636	1.353	1.851	-
	80	-3.030 ถึง 35.762	1.710	4.248	-
	90	-2.914 ถึง 155.216	3.013	15.898	-
	100	-47.938 ถึง 21.583	0.995	5.541	-
1,000	50	-56.162 ถึง 9.703	-0.588	9.393	-
	60	-5.952 ถึง 19.827	1.383	3.006	-
	70	-5.144 ถึง 8.908	1.162	1.876	-
	80	-3.429 ถึง 9.009	1.163	1.710	-
	90	-3.175 ถึง 16.801	1.270	2.166	-
	100	-2.207 ถึง 71.144	1.782	7.096	-
1,500	50	-12.682 ถึง 7.418	0.776	2.279	37
	60	-6.460 ถึง 7.034	0.898	1.488	44
	70	-4.376 ถึง 19.873	1.179	2.612	51
	80	-13.576 ถึง 6.927	0.822	2.089	-
	90	-8.271 ถึง 7.143	0.834	1.636	-
	100	-7.208 ถึง 6.021	0.954	1.363	75

ตาราง 10 (ต่อ)

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{sd}	ข้อที่ไม่เหมาะสมกับโมเดล
2,000	50	-7.208 ถึง 25.083	1.357	3.809	-
	60	-4.000 ถึง 7.492	1.060	1.566	-
	70	-4.287 ถึง 7.380	1.038	1.585	-
	80	-4.454 ถึง 9.889	1.075	1.626	-
	90	-4.477 ถึง 18.071	1.186	2.174	-
	100	-3.845 ถึง 26.656	1.278	2.770	-
2,500	50	-2.582 ถึง 32.680	1.790	4.634	-
	60	-70.552 ถึง 5.574	-0.073	9.326	-
	70	-16.687 ถึง 5.960	0.864	2.505	-
	80	-11.128 ถึง 12.046	1.028	2.173	-
	90	-10.314 ถึง 18.172	1.140	2.459	-
	100	-7.282 ถึง 20.432	1.208	2.348	-
3,000	50	-9.074 ถึง 9.135	0.903	2.447	-
	60	-9.643 ถึง 5.273	0.877	1.956	-
	70	-9.900 ถึง 7.626	0.947	1.938	-
	80	-10.063 ถึง 18.622	1.082	2.562	-
	90	-11.187 ถึง 93.853	1.932	9.935	-
	100	-39.310 ถึง 4.292	0.553	4.262	-

จากตาราง 10 พบว่า ค่าความยากของข้อสอบมีค่าต่ำสุดเท่ากับ -168.848 เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาด 90 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาด 60 ข้อ ค่าความยากของข้อสอบจะมีค่าสูงสุดเท่ากับ 549.466 สำหรับค่าความยากเฉลี่ยของข้อสอบมีค่าต่ำสุดเท่ากับ -1.539 เมื่อกลุ่ม

ตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาด 70 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาด 60 ข้อ ค่าความยากเฉลี่ยของข้อสอบจะมีค่าสูงสุดเท่ากับ 10.384 ส่วนความเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบมีค่าต่ำสุดเท่ากับ 1.363 เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน และข้อสอบมีขนาด 100 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาด 60 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบจะมีค่าสูงสุดเท่ากับ 70.806

ในการประมาณค่าความยากของข้อสอบด้วยวิธีอีวริสติกนี้ พบว่า มีข้อสอบที่ไม่เหมาะสมกับโมเดล 1 ข้อ เมื่อกลุ่มตัวอย่างมีขนาด 600 คน ข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างขนาด 1,500 คน ข้อสอบมีขนาด 50 ข้อ 60 ข้อ 70 ข้อ และ 100 ข้อ

2.6 ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 11

ตาราง 11 ค่าสถิติพื้นฐานของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีของเบย์ เมื่อ
ขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{sd}	ข้อที่ไม่เหมาะสมกับโมเดล
200	50	-9.390 ถึง 4.076	-0.082	2.414	-
	60	-1.711 ถึง 3.484	0.503	1.115	-
	70	-6.005 ถึง 6.506	0.039	2.163	-
	80	-4.201 ถึง 6.948	0.136	2.135	-
	90	-11.676 ถึง 9.166	-0.182	2.262	-
	100	-6.460 ถึง 8.179	0.339	2.076	-
300	50	-6.408 ถึง 3.672	-0.035	1.241	-
	60	-4.613 ถึง 4.942	0.151	1.477	-
	70	-4.905 ถึง 3.554	0.017	1.325	-
	80	-9.510 ถึง 5.191	0.044	1.652	-
	90	-4.386 ถึง 5.110	-0.030	1.197	-
	100	-3.048 ถึง 9.746	0.221	1.468	-
400	50	-2.293 ถึง 11.568	0.286	2.033	-
	60	-2.202 ถึง 9.174	0.486	1.955	-
	70	-3.613 ถึง 8.193	0.246	2.602	-
	80	-2.688 ถึง 3.087	0.008	1.029	-
	90	-9.941 ถึง 6.120	-0.279	1.904	-
	100	-5.263 ถึง 11.397	0.076	1.701	-
500	50	-4.624 ถึง 3.590	-0.160	1.464	-
	60	-8.112 ถึง 9.220	-0.209	2.132	-
	70	-5.323 ถึง 4.086	-0.171	1.519	-
	80	-13.159 ถึง 5.946	-0.213	2.805	-
	90	-10.581 ถึง 7.525	-0.069	2.235	-
	100	-11.995 ถึง 5.088	-0.122	1.900	-

ตาราง 11 (ต่อ)

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
600	50	-1.433 ถึง 3.222	0.003	0.738	-
	60	-5.256 ถึง 7.457	0.017	1.536	-
	70	-3.519 ถึง 4.241	0.211	1.189	-
	80	-10.262 ถึง 3.448	-0.179	1.859	-
	90	-6.378 ถึง 2.919	-0.158	1.225	-
	100	-9.984 ถึง 8.625	0.039	1.906	-
700	50	-1.051 ถึง 2.574	0.160	0.540	-
	60	-4.209 ถึง 4.084	0.002	0.736	-
	70	-4.065 ถึง 1.213	-0.270	0.718	-
	80	-4.724 ถึง 5.849	0.196	0.987	-
	90	-5.071 ถึง 4.061	-0.181	0.591	-
	100	-9.001 ถึง 7.207	-0.025	1.736	-
800	50	-2.887 ถึง 2.444	0.031	0.724	-
	60	-4.034 ถึง 4.076	-0.063	0.398	-
	70	-4.475 ถึง 3.242	-0.136	0.658	-
	80	-5.071 ถึง 3.437	-0.224	0.543	-
	90	-4.895 ถึง 3.749	-0.190	0.537	-
	100	-4.765 ถึง 3.233	0.060	0.532	-
900	50	-3.329 ถึง 9.300	0.154	1.474	-
	60	-4.437 ถึง 4.248	0.026	0.933	-
	70	-3.773 ถึง 1.812	-0.016	0.701	-
	80	-2.428 ถึง 4.984	0.148	0.945	-
	90	-6.584 ถึง 3.500	-0.067	0.976	-
	100	-4.467 ถึง 1.985	-0.020	0.712	-

ตาราง 11 (ต่อ)

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
1,000	50	-5.337 ถึง 7.421	-0.011	1.789	-
	60	-12.224 ถึง 9.673	-0.405	2.704	-
	70	-5.536 ถึง 6.771	0.120	1.590	-
	80	-11.146 ถึง 5.398	-0.350	2.053	-
	90	-7.611 ถึง 8.338	0.027	1.826	-
	100	-14.784 ถึง 5.565	-0.387	2.444	-
1,500	50	-6.274 ถึง 9.243	0.493	2.509	-
	60	-8.358 ถึง 11.162	0.093	2.294	-
	70	-3.509 ถึง 7.395	-0.014	1.575	-
	80	-6.263 ถึง 6.277	0.121	1.811	-
	90	-6.413 ถึง 4.410	-0.236	1.669	-
	100	-5.504 ถึง 10.734	0.125	1.847	-
2,000	50	-6.974 ถึง 13.383	0.626	2.934	-
	60	-4.642 ถึง 7.108	0.304	1.718	-
	70	-9.022 ถึง 9.961	0.467	2.743	-
	80	-6.243 ถึง 5.919	0.329	1.853	-
	90	-8.357 ถึง 7.244	0.568	3.136	-
	100	-4.344 ถึง 4.744	0.091	1.789	-
2,500	50	-6.430 ถึง 5.210	0.139	1.795	-
	60	-5.496 ถึง 5.828	0.130	1.574	-
	70	-4.043 ถึง 3.290	-0.017	1.395	-
	80	-4.323 ถึง 3.142	0.026	1.285	-
	90	-7.683 ถึง 4.401	-0.095	2.106	-
	100	-2.067 ถึง 2.405	0.006	0.521	-

ตาราง 11 (ต่อ)

จำนวนคน	จำนวนข้อ	b	$\bar{b}$	b_{SD}	ข้อที่ไม่เหมาะสมกับโมเดล
3,000	50	-5.005 ถึง 8.883	0.541	2.196	-
	60	-4.480 ถึง 3.434	0.036	1.498	-
	70	-4.779 ถึง 2.716	-0.169	1.466	-
	80	-4.465 ถึง 3.084	-0.399	1.508	-
	90	-4.955 ถึง 5.843	0.086	2.541	-
	100	-5.472 ถึง 3.312	-0.052	1.328	-

จากตาราง 11 พบว่า ค่าความยากของข้อสอบมีค่าต่ำสุดเท่ากับ -14.784 เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาด 100 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด 50 ข้อ ค่าความยากของข้อสอบจะมีค่าสูงสุดเท่ากับ 13.383 สำหรับค่าความยากเฉลี่ยของข้อสอบมีค่าต่ำสุดเท่ากับ -0.405 เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาด 60 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด 50 ข้อ ค่าความยากเฉลี่ยของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.626 ส่วนความเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบมีค่าต่ำสุดเท่ากับ 0.398 เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาด 60 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด 90 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบจะมีค่าสูงสุดเท่ากับ 3.136

ในการประมาณค่าความยากของข้อสอบด้วยวิธีของเบสส์ ไม่พบว่า มีข้อสอบที่ไม่เหมาะสมกับโมเดล

2.7 ค่าสถิติพื้นฐานของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม
ไลค์ลิชต์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 12

ตาราง 12 ค่าสถิติพื้นฐานของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิชต์
เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{SD}
200	50	0.135 ถึง 0.328	0.247	0.032
	60	0.135 ถึง 0.328	0.231	0.033
	70	0.129 ถึง 0.321	0.224	0.032
	80	0.123 ถึง 0.314	0.219	0.030
	90	0.124 ถึง 0.343	0.229	0.032
	100	0.155 ถึง 0.500	0.252	0.042
300	50	0.137 ถึง 0.354	0.225	0.032
	60	0.127 ถึง 0.360	0.219	0.033
	70	0.115 ถึง 0.359	0.212	0.034
	80	0.115 ถึง 0.408	0.215	0.045
	90	0.117 ถึง 0.404	0.223	0.039
	100	0.114 ถึง 0.373	0.227	0.038
400	50	0.192 ถึง 0.465	0.243	0.053
	60	0.159 ถึง 0.500	0.233	0.054
	70	0.155 ถึง 0.483	0.232	0.053
	80	0.141 ถึง 0.441	0.220	0.045
	90	0.139 ถึง 0.456	0.234	0.051
	100	0.137 ถึง 0.500	0.239	0.052

ตาราง 12 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{SD}
500	50	0.134 ถึง 0.392	0.250	0.046
	60	0.130 ถึง 0.500	0.243	0.053
	70	0.128 ถึง 0.402	0.231	0.047
	80	0.129 ถึง 0.500	0.235	0.053
	90	0.129 ถึง 0.452	0.241	0.053
	100	0.126 ถึง 0.473	0.239	0.053
600	50	0.127 ถึง 0.500	0.253	0.053
	60	0.125 ถึง 0.500	0.240	0.049
	70	0.120 ถึง 0.500	0.233	0.049
	80	0.117 ถึง 0.500	0.226	0.049
	90	0.120 ถึง 0.500	0.224	0.047
	100	0.116 ถึง 0.424	0.223	0.045
700	50	0.132 ถึง 0.500	0.229	0.049
	60	0.125 ถึง 0.500	0.214	0.049
	70	0.122 ถึง 0.500	0.209	0.053
	80	0.121 ถึง 0.500	0.208	0.053
	90	0.122 ถึง 0.500	0.218	0.054
	100	0.123 ถึง 0.500	0.222	0.054
800	50	0.123 ถึง 0.425	0.220	0.046
	60	0.126 ถึง 0.426	0.210	0.043
	70	0.128 ถึง 0.426	0.210	0.046
	80	0.125 ถึง 0.427	0.204	0.042
	90	0.123 ถึง 0.426	0.207	0.044
	100	0.126 ถึง 0.427	0.207	0.042

ตาราง 12 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{SD}
900	50	0.129 ถึง 0.342	0.242	0.036
	60	0.131 ถึง 0.347	0.230	0.035
	70	0.126 ถึง 0.431	0.225	0.044
	80	0.125 ถึง 0.428	0.223	0.042
	90	0.124 ถึง 0.428	0.227	0.044
	100	0.121 ถึง 0.427	0.225	0.045
1,000	50	0.115 ถึง 0.399	0.216	0.052
	60	0.110 ถึง 0.354	0.202	0.041
	70	0.107 ถึง 0.459	0.203	0.055
	80	0.105 ถึง 0.450	0.199	0.050
	90	0.106 ถึง 0.367	0.213	0.047
	100	0.103 ถึง 0.479	0.216	0.052
1,500	50	0.120 ถึง 0.500	0.221	0.056
	60	0.129 ถึง 0.500	0.209	0.054
	70	0.124 ถึง 0.500	0.209	0.062
	80	0.120 ถึง 0.500	0.205	0.060
	90	0.122 ถึง 0.500	0.213	0.060
	100	0.116 ถึง 0.500	0.215	0.062
2,000	50	0.121 ถึง 0.500	0.209	0.062
	60	0.102 ถึง 0.500	0.198	0.061
	70	0.099 ถึง 0.500	0.196	0.060
	80	0.070 ถึง 0.500	0.191	0.060
	90	0.072 ถึง 0.500	0.201	0.060
	100	0.076 ถึง 0.500	0.206	0.062

ตาราง 12 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	C_{SD}
2,500	50	0.110 ถึง 0.500	0.202	0.063
	60	0.109 ถึง 0.500	0.189	0.061
	70	0.108 ถึง 0.500	0.192	0.062
	80	0.093 ถึง 0.500	0.191	0.061
	90	0.093 ถึง 0.500	0.196	0.062
	100	0.094 ถึง 0.500	0.201	0.065
3,000	50	0.111 ถึง 0.500	0.206	0.071
	60	0.092 ถึง 0.500	0.191	0.070
	70	0.292 ถึง 0.500	0.189	0.070
	80	0.078 ถึง 0.500	0.184	0.067
	90	0.079 ถึง 0.500	0.191	0.066
	100	0.084 ถึง 0.500	0.193	0.068

จากตาราง 12 พบว่า ค่าการเดาของข้อสอบมีค่าต่ำสุดเท่ากับ 0.070 เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน ข้อสอบมีขนาด 80 ข้อ และค่าการเดาของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.500 สำหรับค่าการเดาเฉลี่ยของข้อสอบมีค่าต่ำสุดเท่ากับ 0.184 เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาด 80 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 600 คน และข้อสอบมีขนาด 50 ข้อ ค่าการเดาเฉลี่ยของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.253 ส่วนความเบี่ยงเบนมาตรฐานของค่าการเดาของข้อสอบมีค่าต่ำสุดเท่ากับ 0.030 เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาด 80 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาด 50 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่าการเดาของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.071

2.8 ค่าสถิติพื้นฐานของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีวิธีสถิติ
เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 13

ตาราง 13 ค่าสถิติพื้นฐานของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีวิธีสถิติ เมื่อ
ขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{SD}
200	50	0.066 ถึง 0.361	0.126	0.069
	60	0.065 ถึง 0.361	0.126	0.069
	70	0.065 ถึง 0.361	0.126	0.076
	80	0.065 ถึง 0.448	0.130	0.081
	90	0.065 ถึง 0.672	0.152	0.104
	100	0.085 ถึง 0.537	0.175	0.132
300	50	0.044 ถึง 0.310	0.093	0.056
	60	0.044 ถึง 0.310	0.086	0.057
	70	0.044 ถึง 0.356	0.099	0.069
	80	0.044 ถึง 0.464	0.103	0.075
	90	0.044 ถึง 0.419	0.100	0.071
	100	0.057 ถึง 0.472	0.120	0.075
400	50	0.054 ถึง 0.339	0.108	0.068
	60	0.054 ถึง 0.506	0.126	0.082
	70	0.054 ถึง 0.344	0.118	0.074
	80	0.054 ถึง 0.677	0.107	0.089
	90	0.054 ถึง 0.344	0.103	0.061
	100	0.043 ถึง 0.548	0.090	0.068

ตาราง 13 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	C_{SD}
500	50	0.043 ถึง 0.333	0.107	0.072
	60	0.044 ถึง 0.294	0.099	0.050
	70	0.043 ถึง 0.276	0.100	0.056
	80	0.043 ถึง 0.323	0.096	0.059
	90	0.043 ถึง 0.488	0.084	0.064
	100	0.034 ถึง 0.395	0.085	0.069
600	50	0.049 ถึง 0.390	0.105	0.066
	60	0.049 ถึง 0.320	0.103	0.068
	70	0.049 ถึง 0.586	0.112	0.090
	80	0.049 ถึง 0.523	0.113	0.087
	90	0.049 ถึง 0.522	0.117	0.074
	100	0.048 ถึง 0.398	0.110	0.065
700	50	0.042 ถึง 0.261	0.085	0.046
	60	0.042 ถึง 0.282	0.090	0.060
	70	0.042 ถึง 0.358	0.098	0.067
	80	0.042 ถึง 0.261	0.086	0.052
	90	0.042 ถึง 0.422	0.093	0.071
	100	0.048 ถึง 0.312	0.110	0.065
800	50	0.067 ถึง 0.736	0.145	0.119
	60	0.067 ถึง 0.368	0.132	0.076
	70	0.067 ถึง 0.478	0.145	0.091
	80	0.067 ถึง 0.412	0.139	0.089
	90	0.067 ถึง 0.384	0.127	0.074
	100	0.080 ถึง 0.629	0.167	0.109

ตาราง 13 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{sd}
900	50	0.060 ถึง 0.476	0.142	0.103
	60	0.060 ถึง 0.576	0.134	0.218
	70	0.060 ถึง 0.393	0.141	0.115
	80	0.060 ถึง 0.576	0.134	0.098
	90	0.060 ถึง 0.469	0.130	0.103
	100	0.071 ถึง 0.444	0.156	0.125
1,000	50	0.015 ถึง 0.377	0.159	0.174
	60	0.094 ถึง 0.470	0.177	0.205
	70	0.011 ถึง 0.412	0.140	0.171
	80	0.011 ถึง 0.310	0.137	0.177
	90	0.013 ถึง 0.330	0.134	0.113
	100	0.098 ถึง 0.638	0.210	0.114
1,500	50	0.065 ถึง 0.403	0.132	0.073
	60	0.067 ถึง 0.453	0.133	0.083
	70	0.061 ถึง 0.402	0.180	0.138
	80	0.067 ถึง 0.418	0.138	0.104
	90	0.069 ถึง 0.433	0.151	0.096
	100	0.065 ถึง 0.444	0.123	0.078
2,000	50	0.053 ถึง 0.363	0.128	0.081
	60	0.056 ถึง 0.473	0.162	0.165
	70	0.041 ถึง 0.356	0.110	0.063
	80	0.058 ถึง 0.402	0.115	0.074
	90	0.060 ถึง 0.377	0.121	0.065
	100	0.088 ถึง 0.609	0.183	0.118

ตาราง 13 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{sd}
2,500	50	0.050 ถึง 0.292	0.111	0.061
	60	0.052 ถึง 0.361	0.142	0.116
	70	0.053 ถึง 0.432	0.152	0.110
	80	0.055 ถึง 0.358	0.111	0.069
	90	0.056 ถึง 0.451	0.136	0.104
	100	0.070 ถึง 0.484	0.138	0.089
3,000	50	0.064 ถึง 0.477	0.127	0.087
	60	0.065 ถึง 0.379	0.127	0.379
	70	0.069 ถึง 0.485	0.183	0.164
	80	0.067 ถึง 0.423	0.148	0.103
	90	0.015 ถึง 0.491	0.154	0.087
	100	0.080 ถึง 0.504	0.158	0.089

จากตาราง 13 พบว่า ค่าการเดาของข้อสอบมีค่าต่ำสุดเท่ากับ 0.011 เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน ข้อสอบมีขนาด 70 ข้อ และ 80 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาด 50 ข้อ ค่าการเดาของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.736 สำหรับค่าการเดาเฉลี่ยของข้อสอบมีค่าต่ำสุดเท่ากับ 0.085 เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาด 50 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาด 100 ข้อ ค่าการเดาเฉลี่ยของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.210 ส่วนความเบี่ยงเบนมาตรฐานของค่าการเดาของข้อสอบมีค่าต่ำสุดเท่ากับ 0.046 เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาด 50 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน ข้อสอบมีขนาด 60 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่าการเดาของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.379

2.9 ค่าสถิติพื้นฐานของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์
เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 14

ตาราง 14 ค่าสถิติพื้นฐานของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อ
ขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{sd}
200	50	0.004 ถึง 0.373	0.186	0.111
	60	0.003 ถึง 0.494	0.192	0.111
	70	0.002 ถึง 0.499	0.200	0.123
	80	0.004 ถึง 0.457	0.198	0.111
	90	0.012 ถึง 0.495	0.185	0.132
	100	0.005 ถึง 0.496	0.211	0.128
300	50	0.009 ถึง 0.457	0.232	0.088
	60	0.032 ถึง 0.476	0.237	0.090
	70	0.008 ถึง 0.442	0.221	0.107
	80	0.015 ถึง 0.408	0.227	0.081
	90	0.021 ถึง 0.470	0.230	0.093
	100	0.028 ถึง 0.458	0.228	0.092
400	50	0.031 ถึง 0.495	0.226	0.113
	60	0.031 ถึง 0.477	0.209	0.103
	70	0.019 ถึง 0.351	0.214	0.067
	80	0.026 ถึง 0.414	0.200	0.091
	90	0.005 ถึง 0.483	0.222	0.112
	100	0.014 ถึง 0.487	0.223	0.108

ตาราง 14 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{SD}
500	50	0.015 ถึง 0.434	0.216	0.098
	60	0.006 ถึง 0.482	0.225	0.131
	70	0.009 ถึง 0.464	0.227	0.119
	80	0.015 ถึง 0.496	0.228	0.101
	90	0.033 ถึง 0.477	0.216	0.110
	100	0.003 ถึง 0.475	0.220	0.097
600	50	0.072 ถึง 0.438	0.249	0.079
	60	0.070 ถึง 0.398	0.221	0.068
	70	0.117 ถึง 0.386	0.233	0.053
	80	0.035 ถึง 0.425	0.242	0.057
	90	0.094 ถึง 0.370	0.249	0.049
	100	0.100 ถึง 0.396	0.248	0.054
700	50	0.164 ถึง 0.327	0.239	0.037
	60	0.145 ถึง 0.440	0.246	0.042
	70	0.037 ถึง 0.339	0.235	0.049
	80	0.145 ถึง 0.405	0.245	0.040
	90	0.073 ถึง 0.369	0.239	0.044
	100	0.010 ถึง 0.325	0.245	0.043
800	50	0.149 ถึง 0.283	0.239	0.027
	60	0.120 ถึง 0.283	0.234	0.030
	70	0.004 ถึง 0.373	0.244	0.047
	80	0.054 ถึง 0.317	0.242	0.037
	90	0.013 ถึง 0.377	0.245	0.039
	100	0.055 ถึง 0.409	0.246	0.043

ตาราง 14 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{sd}
900	50	0.021 ถึง 0.325	0.233	0.064
	60	0.032 ถึง 0.372	0.243	0.063
	70	0.068 ถึง 0.401	0.242	0.065
	80	0.070 ถึง 0.375	0.234	0.069
	90	0.067 ถึง 0.381	0.246	0.060
	100	0.015 ถึง 0.395	0.237	0.061
1,000	50	0.058 ถึง 0.448	0.243	0.090
	60	0.070 ถึง 0.437	0.232	0.085
	70	0.050 ถึง 0.494	0.242	0.110
	80	0.006 ถึง 0.476	0.229	0.096
	90	0.002 ถึง 0.499	0.227	0.121
	100	0.042 ถึง 0.464	0.235	0.102
1,500	50	0.095 ถึง 0.460	0.252	0.094
	60	0.047 ถึง 0.451	0.246	0.092
	70	0.003 ถึง 0.457	0.242	0.096
	80	0.065 ถึง 0.497	0.257	0.100
	90	0.032 ถึง 0.484	0.244	0.097
	100	0.119 ถึง 0.416	0.249	0.073
2,000	50	0.007 ถึง 0.349	0.149	0.093
	60	0.008 ถึง 0.257	0.147	0.073
	70	0.010 ถึง 0.325	0.133	0.079
	80	0.007 ถึง 0.317	0.160	0.073
	90	0.020 ถึง 0.446	0.170	0.130
	100	0.081 ถึง 0.488	0.244	0.055

ตาราง 14 (ต่อ)

จำนวนคน	จำนวนข้อ	c	$\bar{c}$	c_{SD}
2,500	50	0.019 ถึง 0.487	0.296	0.101
	60	0.186 ถึง 0.337	0.253	0.035
	70	0.021 ถึง 0.434	0.251	0.066
	80	0.024 ถึง 0.347	0.243	0.043
	90	0.041 ถึง 0.498	0.158	0.129
	100	0.039 ถึง 0.380	0.246	0.049
3,000	50	0.030 ถึง 0.428	0.237	0.109
	60	0.173 ถึง 0.337	0.247	0.038
	70	0.018 ถึง 0.409	0.245	0.077
	80	0.011 ถึง 0.309	0.240	0.041
	90	0.019 ถึง 0.464	0.168	0.137
	100	0.010 ถึง 0.325	0.247	0.041

จากตาราง 14 พบว่า ค่าการเดาของข้อสอบมีค่าต่ำสุดเท่ากับ 0.002 เมื่อกลุ่มตัวอย่างมีขนาด 200 คน ข้อสอบมีขนาด 70 ข้อ และเมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน ข้อสอบมีขนาด 90 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 200 คน ข้อสอบมีขนาด 70 ข้อ และเมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน ข้อสอบมีขนาด 90 ข้อ ค่าการเดาของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.499 สำหรับค่าการเดาเฉลี่ยของข้อสอบมีค่าต่ำสุดเท่ากับ 0.133 เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด 70 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน และข้อสอบมีขนาด 50 ข้อ ค่าการเดาเฉลี่ยของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.296 ส่วนความเบี่ยงเบนมาตรฐานของค่าการเดาของข้อสอบมีค่าต่ำสุดเท่ากับ 0.027 เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาด 50 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาด 90 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่าการเดา

ของข้อสอบจะมีค่าสูงสุดเท่ากับ 0.379

2.10 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซีมัมไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 15

ตาราง 15 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซีมัมไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	θ	$\bar{\theta}$	θ_{SD}
200	50	-3.023 ถึง 2.693	0.000	1.003
	60	-3.165 ถึง 2.994	0.000	1.003
	70	-3.238 ถึง 3.063	0.000	1.003
	80	-3.396 ถึง 2.887	0.000	1.003
	90	-3.587 ถึง 3.060	0.000	1.003
	100	-3.491 ถึง 2.641	0.000	1.003
300	50	-3.161 ถึง 2.858	0.000	1.002
	60	-3.362 ถึง 3.120	0.000	1.002
	70	-3.520 ถึง 3.297	0.000	1.002
	80	-3.556 ถึง 3.372	0.000	1.002
	90	-3.436 ถึง 3.439	0.000	1.002
	100	-3.716 ถึง 3.566	0.000	1.002
400	50	-3.297 ถึง 3.029	0.000	1.001
	60	-3.524 ถึง 3.329	0.000	1.001
	70	-3.595 ถึง 3.444	0.000	1.001
	80	-3.723 ถึง 2.746	0.000	1.001
	90	-3.669 ถึง 2.841	0.000	1.001
	100	-4.042 ถึง 3.071	0.000	1.001

ตาราง 15 (ต่อ)

จำนวนคน	จำนวนข้อ	e	$\bar{e}$	e_{SD}
500	50	-2.907 ถึง 2.549	0.000	1.001
	60	-3.159 ถึง 2.868	0.000	1.001
	70	-3.252 ถึง 3.002	0.000	1.001
	80	-3.383 ถึง 3.151	0.000	1.001
	90	-3.457 ถึง 3.282	0.000	1.001
	100	-3.618 ถึง 3.197	0.000	1.001
600	50	-2.978 ถึง 2.608	0.000	1.001
	60	-3.145 ถึง 2.829	0.000	1.001
	70	-3.233 ถึง 2.953	0.000	1.001
	80	-1.166 ถึง 1.938	0.000	1.001
	90	-3.438 ถึง 3.228	0.000	1.001
	100	-3.567 ถึง 3.366	0.000	1.001
700	50	-3.168 ถึง 2.857	0.000	1.001
	60	-3.506 ถึง 3.248	0.000	1.001
	70	-3.549 ถึง 3.335	0.000	1.001
	80	-3.608 ถึง 3.426	0.000	1.001
	90	-3.679 ถึง 3.312	0.000	1.001
	100	-3.489 ถึง 3.673	0.000	1.001
800	50	-3.110 ถึง 2.771	0.000	1.001
	60	-3.298 ถึง 3.030	0.000	1.001
	70	-3.448 ถึง 3.233	0.000	1.001
	80	-3.580 ถึง 3.392	0.000	1.001
	90	-3.677 ถึง 3.229	0.000	1.001
	100	-3.634 ถึง 3.473	0.000	1.001

ตาราง 15 (ต่อ)

จำนวนคน	จำนวนข้อ	e	$\bar{e}$	e_{SD}
900	50	-2.741 ถึง 2.384	0.000	1.001
	60	-2.977 ถึง 2.645	0.000	1.001
	70	-3.092 ถึง 2.797	0.000	1.001
	80	-3.255 ถึง 2.989	0.000	1.001
	90	-3.314 ถึง 3.063	0.000	1.001
	100	-3.446 ถึง 3.209	0.000	1.001
1,000	50	-3.111 ถึง 2.794	0.000	1.001
	60	-3.321 ถึง 3.067	0.000	1.001
	70	-3.439 ถึง 3.221	0.000	1.001
	80	-3.535 ถึง 3.334	0.000	1.001
	90	-3.601 ถึง 3.431	0.000	1.000
	100	-3.709 ถึง 3.545	0.000	1.001
1,500	50	-3.056 ถึง 2.683	0.000	1.000
	60	-3.287 ถึง 2.985	0.000	1.000
	70	-3.389 ถึง 3.138	0.000	1.000
	80	-3.492 ถึง 3.273	0.000	1.000
	90	-3.566 ถึง 3.376	0.000	1.000
	100	-3.750 ถึง 3.562	0.000	1.000
2,000	50	-3.022 ถึง 2.701	0.000	1.000
	60	-3.304 ถึง 3.031	0.000	1.000
	70	-3.427 ถึง 3.184	0.000	1.000
	80	-3.542 ถึง 3.335	0.000	1.000
	90	-3.626 ถึง 3.439	0.000	1.000
	100	-3.724 ถึง 3.540	0.000	1.000

ตาราง 15 (ต่อ)

จำนวนคน	จำนวนข้อ	e	$\bar{e}$	e_{SD}
2,500	50	-2.994 ถึง 2.639	0.000	1.000
	60	-3.226 ถึง 2.939	0.000	1.000
	70	-3.325 ถึง 3.073	0.000	1.000
	80	-3.452 ถึง 3.227	0.000	1.000
	90	-3.560 ถึง 3.363	0.000	1.000
	100	-3.666 ถึง 3.273	0.000	1.000
3,000	50	-3.052 ถึง 2.693	0.000	1.000
	60	-3.292 ถึง 3.008	0.000	1.000
	70	-3.403 ถึง 3.155	0.000	1.000
	80	-3.525 ถึง 3.311	0.000	1.000
	90	-3.601 ถึง 3.408	0.000	1.000
	100	-3.745 ถึง 3.556	0.000	1.000

จากตาราง 15 พบว่า ค่าความสามารถของผู้สอบมีค่าต่ำสุดเท่ากับ -4.042 เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาด 100 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาด 100 ข้อ ค่าความสามารถของผู้สอบจะมีค่าสูงสุดเท่ากับ 3.673 สำหรับค่าความสามารถเฉลี่ยของผู้สอบมีค่าเท่ากับ 0.000 ในทุกขนาดกลุ่มตัวอย่าง ส่วนความเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบมีค่าต่ำสุดเท่ากับ 1.000 และมีค่าสูงสุดเท่ากับ 1.003

2.11 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอิวาโรสติด
เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 16

ตาราง 16 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอิวาโรสติด เมื่อ
ขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	o	$\bar{o}$	σ_{SD}
200	50	-0.699 ถึง 1.894	0.750	0.414
	60	-0.938 ถึง 2.324	1.027	0.500
	70	-0.916 ถึง 2.167	0.932	0.451
	80	-0.912 ถึง 2.161	0.926	0.453
	90	-0.755 ถึง 2.082	0.901	0.455
	100	-0.776 ถึง 2.076	0.924	0.432
300	50	-1.171 ถึง 2.290	1.082	0.600
	60	-1.009 ถึง 2.431	1.090	0.565
	70	-0.891 ถึง 2.236	0.955	0.507
	80	-0.805 ถึง 2.214	0.951	0.486
	90	-0.643 ถึง 2.123	0.917	0.474
	100	-0.619 ถึง 2.133	0.948	0.467
400	50	-1.232 ถึง 1.911	0.704	0.458
	60	-1.603 ถึง 2.385	0.960	0.572
	70	-1.400 ถึง 2.249	0.866	0.539
	80	-1.073 ถึง 2.017	0.893	0.516
	90	-0.980 ถึง 1.865	0.879	0.491
	100	-0.859 ถึง 1.949	0.898	0.454

ตาราง 16 (ต่อ)

จำนวนคน	จำนวนข้อ	e	$\bar{e}$	e_{SD}
500	50	-0.934 ถึง 2.357	1.079	0.497
	60	-0.859 ถึง 2.371	1.073	0.510
	70	-0.846 ถึง 2.202	0.965	0.459
	80	-0.918 ถึง 2.201	0.965	0.454
	90	-0.889 ถึง 2.100	0.941	0.425
	100	-0.832 ถึง 2.120	0.978	0.420
600	50	-1.014 ถึง 2.433	1.125	0.557
	60	-0.928 ถึง 2.200	1.100	0.486
	70	-0.674 ถึง 2.066	0.973	0.450
	80	-0.743 ถึง 2.058	0.970	0.430
	90	-0.704 ถึง 1.978	0.940	0.417
	100	-0.707 ถึง 1.904	0.959	0.400
700	50	-1.532 ถึง 2.419	1.036	0.571
	60	-1.544 ถึง 2.572	1.025	0.567
	70	-1.353 ถึง 2.382	0.919	0.532
	80	-1.035 ถึง 2.356	0.931	0.516
	90	-0.945 ถึง 2.126	0.901	0.482
	100	-0.860 ถึง 2.096	0.906	0.455
800	50	-1.489 ถึง 2.268	0.973	0.507
	60	-1.453 ถึง 2.258	0.992	0.483
	70	-1.314 ถึง 2.032	0.899	0.486
	80	-0.996 ถึง 1.935	0.904	0.463
	90	-0.996 ถึง 2.126	0.892	0.460
	100	-1.037 ถึง 1.889	0.883	0.937

ตาราง 16 (ต่อ)

จำนวนคน	จำนวนข้อ	e	$\bar{e}$	e_{SD}
900	50	-0.965 ถึง 2.519	1.048	0.480
	60	-0.818 ถึง 2.268	1.020	0.442
	70	-0.793 ถึง 2.112	0.920	0.401
	80	-0.818 ถึง 2.238	0.929	0.400
	90	-0.858 ถึง 2.034	0.911	0.382
	100	-0.818 ถึง 2.036	0.912	0.379
1,000	50	-1.276 ถึง 2.324	0.904	0.594
	60	-1.099 ถึง 2.429	0.884	0.535
	70	-1.076 ถึง 2.283	0.822	0.511
	80	-1.005 ถึง 2.281	0.865	0.495
	90	-0.921 ถึง 2.061	0.836	0.467
	100	-0.850 ถึง 2.043	0.853	0.445
1,500	50	-1.443 ถึง 2.385	1.004	0.574
	60	-1.185 ถึง 2.319	0.989	0.514
	70	-1.364 ถึง 2.148	0.860	0.481
	80	-1.318 ถึง 2.155	0.899	0.461
	90	-1.276 ถึง 2.066	0.864	0.450
	100	-1.327 ถึง 2.079	0.911	0.436
2,000	50	-1.612 ถึง 2.573	0.970	0.582
	60	-1.571 ถึง 2.490	0.960	0.536
	70	-1.345 ถึง 2.330	0.888	0.501
	80	-1.070 ถึง 2.280	0.905	0.463
	90	-0.992 ถึง 2.171	0.868	0.451
	100	-0.890 ถึง 2.054	0.866	0.439

ตาราง 16 (ต่อ)

จำนวนคน	จำนวนข้อ	๑	๑	๑ SD
2,500	50	-1.370 ถึง 2.335	0.983	0.527
	60	-1.073 ถึง 2.294	0.959	0.498
	70	-1.003 ถึง 2.176	0.864	0.486
	80	-1.019 ถึง 1.983	0.907	0.472
	90	-0.957 ถึง 1.912	0.867	0.456
	100	-0.979 ถึง 1.913	0.880	0.431
3,000	50	-1.469 ถึง 2.479	0.981	0.491
	60	-1.458 ถึง 2.436	0.974	0.478
	70	-1.318 ถึง 2.107	0.849	0.454
	80	-1.295 ถึง 2.238	0.875	0.444
	90	-1.034 ถึง 2.162	0.854	0.458
	100	-1.330 ถึง 2.226	0.873	0.423

จากตาราง 16 พบว่า ค่าความสามารถของผู้สอบมีค่าต่ำสุดเท่ากับ -4.612 เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด 50 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาด 50 ข้อ ค่าความสามารถของผู้สอบจะมีค่าสูงสุดเท่ากับ 2.573 สำหรับค่าความสามารถเฉลี่ยของผู้สอบมีค่าต่ำสุดเท่ากับ 0.704 เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาด 50 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 600 คน และข้อสอบมีขนาด 50 ข้อ ค่าความสามารถเฉลี่ยของผู้สอบจะมีค่าสูงสุดเท่ากับ 1.125 ส่วนความเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบมีค่าต่ำสุดเท่ากับ 0.379 เมื่อกลุ่มตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาด 100 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาด 100 ข้อ ค่าความเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบจะมีค่าสูงสุดเท่ากับ 0.937

2.12 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 17

ตาราง 17 ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	e	$\bar{e}$	e_{SD}
200	50	-3.266 ถึง 3.246	0.001	1.007
	60	-3.069 ถึง 3.228	0.001	1.008
	70	-3.219 ถึง 3.263	0.000	1.000
	80	-3.196 ถึง 3.183	-0.001	1.004
	90	-3.242 ถึง 3.175	0.000	1.007
	100	-3.225 ถึง 3.162	0.000	1.010
300	50	-3.129 ถึง 3.232	-0.001	1.010
	60	-3.219 ถึง 3.228	0.000	1.004
	70	-3.198 ถึง 3.245	0.001	1.006
	80	-3.204 ถึง 3.288	0.001	1.006
	90	-3.196 ถึง 3.241	0.000	1.010
	100	-3.204 ถึง 3.271	0.000	1.005
400	50	-3.205 ถึง 3.204	0.001	1.009
	60	-3.184 ถึง 3.264	0.000	1.010
	70	-3.296 ถึง 3.232	0.001	1.003
	80	-3.207 ถึง 3.328	-0.001	1.005
	90	-3.251 ถึง 3.237	-0.002	1.002
	100	-3.156 ถึง 3.267	0.001	1.006

ตาราง 17 (ต่อ)

จำนวนคน	จำนวนข้อ	σ	$\bar{\sigma}$	σ_{SD}
500	50	-3.213 ถึง 3.209	0.000	1.006
	60	-3.283 ถึง 3.283	0.000	1.006
	70	-3.166 ถึง 3.243	0.001	1.008
	80	-3.602 ถึง 3.172	-0.001	1.000
	90	-3.199 ถึง 3.141	0.000	1.009
	100	-3.201 ถึง 3.232	0.000	1.004
600	50	-3.280 ถึง 3.278	0.000	1.007
	60	-3.211 ถึง 3.273	0.001	1.003
	70	-3.214 ถึง 3.231	0.000	1.009
	80	-3.224 ถึง 3.236	0.000	1.000
	90	-3.294 ถึง 3.232	0.000	1.004
	100	-3.298 ถึง 3.271	0.000	1.012
700	50	-3.223 ถึง 3.254	0.000	1.006
	60	-3.235 ถึง 3.218	0.000	1.003
	70	-3.290 ถึง 3.122	0.000	1.004
	80	-3.326 ถึง 3.220	0.001	1.010
	90	-3.252 ถึง 3.377	0.001	1.006
	100	-3.217 ถึง 3.243	0.000	1.001
800	50	-3.244 ถึง 3.227	-0.001	1.009
	60	-3.293 ถึง 3.298	-0.001	1.008
	70	-3.253 ถึง 3.242	-0.001	1.001
	80	-3.286 ถึง 3.286	0.000	1.010
	90	-3.171 ถึง 3.295	0.000	1.003
	100	-3.248 ถึง 3.188	0.000	1.009

ตาราง 17 (ต่อ)

จำนวนคน	จำนวนข้อ	σ	$\bar{\sigma}$	σ_{SD}
900	50	-3.201 ถึง 3.283	-0.001	1.001
	60	-3.292 ถึง 3.312	0.000	1.006
	70	-3.157 ถึง 3.298	0.000	1.011
	80	-3.161 ถึง 3.283	0.000	1.007
	90	-3.779 ถึง 3.238	0.001	1.011
	100	-3.219 ถึง 3.282	-0.001	1.011
1,000	50	-3.287 ถึง 3.264	-0.001	1.005
	60	-3.271 ถึง 3.278	0.000	1.004
	70	-3.203 ถึง 3.288	0.000	1.011
	80	-3.201 ถึง 3.254	0.000	1.002
	90	-3.209 ถึง 3.215	0.000	1.002
	100	-3.298 ถึง 3.297	0.000	1.003
1,500	50	-3.209 ถึง 3.233	0.000	1.012
	60	-3.208 ถึง 3.285	0.000	1.007
	70	-3.199 ถึง 3.241	0.000	1.003
	80	-3.230 ถึง 3.353	0.000	1.003
	90	-3.280 ถึง 3.252	0.001	1.003
	100	-3.277 ถึง 3.325	0.000	1.008
2,000	50	-3.191 ถึง 3.294	0.000	1.013
	60	-3.247 ถึง 3.267	0.000	1.000
	70	-3.248 ถึง 3.278	0.000	1.013
	80	-3.203 ถึง 3.228	0.000	1.007
	90	-3.213 ถึง 3.282	0.000	1.003
	100	-3.198 ถึง 3.287	0.000	1.011

ตาราง 17 (ต่อ)

จำนวนคน	จำนวนข้อ	e	$\bar{e}$	e_{SD}
2,500	50	-3.286 ถึง 3.206	0.000	1.001
	60	-3.212 ถึง 3.261	0.000	1.007
	70	-3.276 ถึง 3.281	0.000	1.004
	80	-3.259 ถึง 3.259	0.000	1.000
	90	-3.354 ถึง 3.356	0.000	1.003
	100	-3.250 ถึง 3.260	0.000	1.012
3,000	50	-3.321 ถึง 3.300	0.000	1.005
	60	-3.305 ถึง 3.367	0.000	1.001
	70	-3.302 ถึง 3.305	0.000	1.007
	80	-3.283 ถึง 3.261	0.000	1.004
	90	-3.297 ถึง 3.255	0.000	1.001
	100	-3.299 ถึง 3.297	0.000	1.011

จากตาราง 17 พบว่า ค่าความสามารถของผู้สอบมีค่าต่ำสุดเท่ากับ -3.779 เมื่อกลุ่มตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาด 90 ข้อ แต่เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน ข้อสอบมีขนาด 60 ข้อ ค่าความสามารถของผู้สอบจะมีค่าสูงสุดเท่ากับ 3.367 สำหรับค่าความถ่วงเฉลี่ยของผู้สอบมีค่าต่ำสุดเท่ากับ -0.002 และมีค่าสูงสุดเท่ากับ 0.001 ส่วนความเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบมีค่าต่ำสุดเท่ากับ 1.000 และมีค่าสูงสุดเท่ากับ 1.003

3. ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าพารามิเตอร์ของข้อสอบ
และความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีเอสดี และวิธีของ
เบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

3.1 ผลการเปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่
ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง
18

ตาราง 18 เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วย
วิธีแมกซิมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _B	MS _w	F-test
50	12	1.261	0.105	1.273	0.083
60	12	1.635	0.136	1.264	0.108
70	12	1.221	0.102	1.212	0.084
80	12	1.218	0.102	1.055	0.096
90	12	0.958	0.080	1.096	0.073
100	12	0.807	0.067	1.142	0.059

จากตาราง 18 พบว่า ค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วย
วิธีแมกซิมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญ
ทางสถิติ

3.2 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีอิวิวิสถิติ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 19 - 20

๔

ตาราง 19 เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีอิวิวิสถิติ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _b	MS _w	F-test
50	12	24.603	2.050	1.205	1.701
60	12	26.203	2.184	1.247	1.751
70	12	11.837	0.986	1.075	0.918
80	12	74.278	6.190	2.000	3.095*
90	12	29.361	2.447	1.143	2.142
100	12	62.532	5.211	4.367	1.193

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 19 พบว่า ค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีอิวิวิสถิติ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังนั้น จึงทำการทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังแสดงในตาราง 20

ตาราง 20 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ที่ประมาณค่า
ด้วยวิธีทวิวิสัย เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	อ	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.538	0.262	0.479	0.497	0.370	0.326	0.472	0.575	0.524	0.616	0.555	0.483	1.438
200	0.538	-	0.276	0.059	0.041	0.168	0.212	0.066	0.037	0.014	0.078	0.017	0.055	0.900
300	0.262		-	0.217	0.235	0.108	0.064	0.210	0.313	0.262	0.354	0.293	0.221	1.176*
400	0.479			-	0.018	0.109	0.153	0.007	0.096	0.045	0.137	0.076	0.004	0.959
500	0.497				-	0.127	0.171	0.025	0.078	0.027	0.119	0.058	0.014	0.941
600	0.370					-	0.044	0.102	0.205	0.154	0.246	0.185	0.113	1.068*
700	0.326						-	0.146	0.249	0.198	0.290	0.229	0.157	1.112*
800	0.472							-	0.103	0.052	0.144	0.083	0.011	1.011
900	0.575								-	0.051	0.041	0.020	0.092	0.908
1,000	0.524									-	0.092	0.031	0.041	0.914
1,500	0.616										-	0.061	0.133	0.822
2,000	0.555											-	0.072	0.888
2,500	0.483												-	0.955
3,000	1.438													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 20 พบว่า ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 80 ข้อ
ของกลุ่มตัวอย่าง 300 คน 600 คน และ 700 คน มีความแตกต่างจากกลุ่มตัวอย่าง
3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.3 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าอำนาจจำแนกของ
ข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง
21 - 26

ตาราง 21 เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธี
ของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _B	MS _w	F-test
50	12	24.904	2.075	0.686	3.025
60	12	21.551	1.796	0.194	9.267*
70	12	31.008	2.584	0.222	11.657*
80	12	24.818	2.068	0.399	5.188*
90	12	25.516	2.126	0.410	5.187*
100	12	30.165	2.514	0.289	8.701*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 21 พบว่า ค่าอำนาจจำแนกของข้อสอบ 60 ข้อ 70 ข้อ
80 ข้อ 90 ข้อ และ 100 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่าง
กัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังนั้น จึงทำการทดสอบความ
แตกต่างของค่าอำนาจจำแนกของข้อสอบดังกล่าวด้วยการเปรียบเทียบพหุคูณ ได้ผลดังแสดงใน
ตาราง 22 - 26

ตาราง 22 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 60 ข้อ ที่ประมาณ
ค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{a}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.842	0.779	0.707	0.718	0.680	0.724	0.669	0.624	0.592	0.549	0.238	0.362	0.436
200	0.842	-	0.063	0.135	0.124	0.162	0.118	0.173	0.218	0.250	0.293	0.604*	0.480*	0.406*
300	0.779		-	0.072	0.155	0.117	0.055	0.110	0.155	0.187	0.230	0.541*	0.417*	0.343
400	0.707			-	0.011	0.027	0.017	0.038	0.083	0.115	0.158	0.469*	0.345	0.271
500	0.718				-	0.038	0.006	0.049	0.094	0.126	0.169	0.480*	0.356	0.282
600	0.680					-	0.044	0.011	0.056	0.088	0.131	0.442*	0.318	0.244
700	0.724						-	0.055	0.100	0.132	0.175	0.486*	0.362	0.288
800	0.669							-	0.045	0.077	0.120	0.431*	0.307	0.233
900	0.624								-	0.032	0.075	0.386*	0.262	0.188
1,000	0.592									-	0.043	0.354	0.230	0.156
1,500	0.549										-	0.311	0.187	0.113
2,000	0.238											-	0.124	0.198
2,500	0.362												-	0.074
3,000	0.436													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 22 พบว่า ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 60 ข้อ ของ
กลุ่มตัวอย่าง 200 คน 300 คน 600 คน 700 คน 800 คน และ 900 คน มีความแตกต่าง
จากกลุ่มตัวอย่าง 2,000 คน ค่าอำนาจจำแนกเฉลี่ยของกลุ่มตัวอย่าง 200 คน และ 300
คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน และค่าอำนาจจำแนกเฉลี่ยของกลุ่มตัวอย่าง
200 คน มีความแตกต่างจากกลุ่มตัวอย่าง 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 23 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 70 ข้อ ที่ประมาณ
ค่าตัววัดของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	ล	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.992	0.849	0.631	0.630	0.570	0.671	0.651	0.622	0.789	0.516	0.229	0.490	0.428
200	0.992	-	0.143	0.361	0.362	0.422*	0.321	0.341	0.370*	0.203	0.476*	0.763*	0.502*	0.564*
300	0.849		-	0.218	0.219	0.279	0.178	0.198	0.227	0.060	0.333	0.620*	0.359	0.421*
400	0.631			-	0.001	0.061	0.040	0.020	0.009	0.158	0.115	0.402*	0.141	0.203
500	0.630				-	0.060	0.041	0.021	0.008	0.159	0.114	0.401*	0.140	0.202
600	0.570					-	0.101	0.081	0.052	0.219	0.054	0.341	0.080	0.142
700	0.671						-	0.020	0.049	0.118	0.155	0.442*	0.181	0.243
800	0.651							-	0.029	0.138	0.135	0.422*	0.161	0.223
900	0.622								-	0.167	0.106	0.393*	0.132	0.194
1,000	0.789									-	0.273	0.560*	0.299	0.361
1,500	0.516										-	0.287	0.026	0.088
2,000	0.229											-	0.261	0.199
2,500	0.490												-	0.062
3,000	0.428													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 23 พบว่า ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 70 ข้อ ของ
กลุ่มตัวอย่าง 200 คน มีความแตกต่างจากกลุ่มตัวอย่าง 600 คน 900 คน 1,500 คน และ
2,500 คน ค่าอำนาจจำแนกเฉลี่ยของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 500 คน
700 คน 800 คน 900 คน และ 1,000 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน
ค่าอำนาจจำแนกเฉลี่ยของกลุ่มตัวอย่าง 200 คน และ 300 คน มีความแตกต่างจากกลุ่ม
ตัวอย่าง 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 24 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบกับพหุคูณ

จำนวน คน	$\bar{a}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.837	0.820	0.562	0.697	0.643	0.683	0.629	0.634	0.531	0.515	0.228	0.580	0.411
200	0.837	-	0.017	0.275	0.140	0.194	0.154	0.208	0.203	0.306	0.322	0.609*	0.257	0.426
300	0.820		-	0.258	0.123	0.177	0.137	0.191	0.186	0.289	0.305	0.592*	0.240	0.409
400	0.562			-	0.135	0.081	0.121	0.067	0.072	0.031	0.047	0.334	0.018	0.151
500	0.697				-	0.054	0.014	0.068	0.063	0.166	0.182	0.469*	0.117	0.286
600	0.643					-	0.040	0.014	0.009	0.112	0.128	0.415	0.063	0.232
700	0.683						-	0.054	0.049	0.152	0.168	0.455	0.103	0.272
800	0.629							-	0.005	0.098	0.114	0.401	0.049	0.218
900	0.634								-	0.103	0.119	0.406	0.054	0.223
1,000	0.531									-	0.016	0.303	0.049	0.120
1,500	0.515										-	0.287	0.065	0.104
2,000	0.228											-	0.352	0.183
2,500	0.580												-	0.169
3,000	0.411													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 24 พบว่า ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 80 ข้อของกลุ่มตัวอย่าง 200 คน 300 คน และ 500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 25 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 90 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{a}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		1.043	0.897	0.813	0.839	0.627	0.706	0.655	0.634	0.689	0.474	0.686	0.657	0.522
200	1.043	-	0.146	0.230	0.204	0.416	0.337	0.388	0.409	0.354	0.569*	0.357	0.386	0.521
300	0.897		-	0.084	0.058	0.270	0.191	0.242	0.263	0.208	0.423	0.211	0.240	0.375
400	0.813			-	0.026	0.186	0.107	0.158	0.179	0.124	0.339	0.127	0.156	0.291
500	0.839				-	0.212	0.133	0.184	0.205	0.150	0.365	0.153	0.182	0.317
600	0.627					-	0.079	0.028	0.007	0.062	0.153	0.059	0.030	0.105
700	0.706						-	0.051	0.072	0.017	0.232	0.020	0.049	0.184
800	0.655							-	0.021	0.034	0.181	0.031	0.002	0.133
900	0.634								-	0.055	0.160	0.052	0.023	0.112
1,000	0.689									-	0.215	0.003	0.032	0.167
1,500	0.474										-	0.212	0.183	0.048
2,000	0.686											-	0.029	0.164
2,500	0.657												-	0.135
3,000	0.522													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 25 พบว่า ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 90 ข้อของกลุ่มตัวอย่าง 200 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน และ 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 26 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ 100 ข้อ ที่ประมาณ
ค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{a}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.824	0.778	0.841	0.718	0.607	0.716	0.625	0.639	0.722	0.421	0.708	0.388	0.371
200	0.824	-	0.046	0.017	0.106	0.217	0.108	0.199	0.185	0.102	0.403*	0.116	0.436*	0.453*
300	0.778		-	0.063	0.060	0.171	0.062	0.153	0.139	0.056	0.357*	0.070	0.390*	0.407*
400	0.841			-	0.123	0.234	0.125	0.216	0.202	0.119	0.420*	0.133	0.453*	0.470*
500	0.718				-	0.111	0.002	0.093	0.079	0.004	0.297	0.010	0.330	0.347
600	0.607					-	0.109	0.127	0.032	0.115	0.186	0.101	0.219	0.236
700	0.716						-	0.091	0.077	0.006	0.295	0.008	0.328	0.345
800	0.625							-	0.014	0.097	0.204	0.083	0.237	0.254
900	0.639								-	0.083	0.218	0.069	0.251	0.268
1,000	0.722									-	0.301	0.014	0.334	0.351*
1,500	0.421										-	0.287	0.033	0.050
2,000	0.708											-	0.320	0.337
2,500	0.388												-	0.017
3,000	0.371													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 26. พบว่า ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 100 ข้อ
ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน
2,500 คน และ 3,000 คน ค่าอำนาจจำแนกเฉลี่ยของกลุ่มตัวอย่าง 1,000 คน มีความ
แตกต่างจากกลุ่มตัวอย่าง 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.4 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าอำนาจจำแนกของ
ข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของ
กลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 27 - 28

ตาราง 27 เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วย
วิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
50	200	2	14.677	7.339	1.349	5.442*
	300	2	18.553	9.277	1.450	6.397*
	400	2	0.193	0.097	2.306	0.042
	500	2	1.915	0.957	0.729	1.313
	600	2	9.250	4.625	0.886	5.220*
	700	2	4.482	2.241	0.806	2.782
	800	2	3.458	1.729	1.264	1.369
	900	2	2.632	1.316	0.715	1.840
	1,000	2	0.662	0.331	0.801	0.413
	1,500	2	2.681	1.341	0.804	1.668
	2,000	2	3.345	1.673	0.686	2.440
	2,500	2	2.584	1.292	0.936	1.380
	3,000	2	0.779	0.390	0.982	0.397
60	200	2	14.814	7.407	1.683	4.402*
	300	2	21.691	10.846	1.301	8.335*
	400	2	7.058	3.529	1.225	2.880
	500	2	6.042	3.021	0.611	4.948*
	600	2	6.710	3.355	0.819	4.097*
	700	2	7.477	3.739	0.835	4.480*

ตาราง 27 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _b	MS _w	F-test
60	800	2	3.697	1.849	0.662	2.794
	900	2	4.144	2.072	0.807	2.566
	1,000	2	1.307	0.654	0.810	0.807
	1,500	2	2.980	1.490	0.746	1.999
	2,000	2	2.522	1.261	0.777	1.623
	2,500	2	0.703	0.352	0.593	0.593
	3,000	2	0.880	0.440	0.853	0.516
70	200	2	17.847	8.924	1.734	5.146*
	300	2	21.372	10.686	1.171	9.124*
	400	2	5.329	2.664	0.509	5.234*
	500	2	4.319	2.159	0.781	2.764
	600	2	3.459	1.730	0.756	2.287
	700	2	7.043	3.521	1.030	3.419*
	800	2	4.781	2.390	0.584	4.090*
	900	2	5.067	2.534	0.715	3.543*
	1,000	2	8.030	4.015	0.746	5.379*
	1,500	2	1.189	0.594	0.764	0.777
	2,000	2	4.061	2.031	0.637	3.188*
	2,500	2	1.536	0.768	0.647	1.187
	3,000	2	1.254	0.627	0.792	0.792
80	200	2	9.077	4.538	2.208	2.056
	300	2	15.068	7.534	1.332	5.657*
	400	2	2.320	1.190	1.052	1.131
	500	2	7.263	3.632	0.975	3.724*

ตาราง 27 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
80	600	2	4.705	2.352	0.954	2.466
	700	2	7.753	3.877	0.663	5.849*
	800	2	3.206	1.603	0.738	2.174
	900	2	6.383	3.191	0.711	4.486*
	1,000	2	2.679	1.339	0.782	1.713
	1,500	2	2.824	1.412	0.635	2.225
	2,000	2	5.546	2.773	0.601	4.612*
	2,500	2	2.586	1.293	0.704	1.837
	3,000	2	1.651	0.825	0.664	1.244
90	200	2	44.584	22.292	2.256	9.882*
	300	2	29.771	14.885	1.250	11.908*
	400	2	12.269	6.134	0.680	9.021*
	500	2	15.795	7.898	0.856	9.227*
	600	2	4.208	2.104	0.784	2.683
	700	2	9.511	4.755	0.649	7.323*
	800	2	8.381	4.191	0.673	6.223*
	900	2	6.987	3.493	0.732	4.776*
	1,000	2	6.169	3.084	0.646	4.776*
	1,500	2	3.739	1.869	0.658	2.840
	2,000	2	7.867	3.934	0.771	5.104*
	2,500	2	5.016	2.508	0.750	3.347*
	3,000	2	3.011	1.506	0.770	1.956

ตาราง 27 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
100	200	2	30.141	15.071	2.456	6.136*
	300	2	44.033	22.017	2.092	10.527*
	400	2	14.903	7.451	2.829	2.634
	500	2	22.672	11.336	1.391	8.151*
	600	2	3.318	1.659	1.461	1.136
	700	2	26.674	13.337	2.831	4.711*
	800	2	4.263	2.132	1.772	1.203
	900	2	5.674	2.837	1.935	1.466
	1,000	2	11.166	5.583	1.261	4.429*
	1,500	2	4.595	2.298	2.382	0.965
	2,000	2	14.037	7.019	1.900	3.694*
	2,500	2	0.456	0.228	1.407	0.162
	3,000	2	3.590	1.795	1.409	1.274

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 27 พบว่า ค่าอำนาจจำแนกของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่างขนาด 200 คน 300 คน และ 600 คน ค่าอำนาจจำแนกของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 500 คน 600 คน และ 700 คน ค่าอำนาจจำแนกของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 700 คน 800 คน 900 คน 1,000 คน และ 2,000 คน ค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 300 คน 500 คน 700 คน 900 คน และ 2,000 คน ค่าอำนาจจำแนกของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 500 คน 700 คน 800 คน 900 คน 1,000 คน 2,000 คน และ 2,500 คน และค่าอำนาจจำแนกของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง

200 คน 300 คน 500 คน 700 คน 1,000 คน และ 2,000 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จึงได้ทำการทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบดังกล่าวด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 28

ตาราง 28 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส์	อีวริสติก-เบส์
50	200	0.005*	0.661*	0.666*
	300	0.359	0.498	0.857*
	400	-	-	-
	500	-	-	-
	600	0.241*	0.363	0.604*
	700	-	-	-
	800	-	-	-
	900	-	-	-
	1,000	-	-	-
	1,500	-	-	-
	2,000	-	-	-
	2,500	-	-	-
	3,000	-	-	-
60	200	0.187	0.493	0.680*
	300	0.411	0.439	0.850*
	400	-	-	-
	500	0.085	0.424*	0.339

ตาราง 28 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไคส์ลิต-อีวริสติก	แมกซิมัม ไคส์ลิต-เบส	อีวริสติก-เบส
60	600	0.062	0.375	0.437*
	700	0.077	0.466*	0.389
	800	-	-	-
	900	-	-	-
	1,000	-	-	-
	1,500	-	-	-
	2,000	-	-	-
	2,500	-	-	-
	3,000	-	-	-
70	200	0.064	0.648*	0.584*
	300	0.209	0.547*	0.756*
	400	0.260	0.382*	0.122
	500	-	-	-
	600	-	-	-
	700	0.113	0.433*	0.320
	800	0.234	0.364*	0.130
	900	0.283	0.362*	0.079
	1,000	0.063	0.443*	0.380*
	1,500	-	-	-
	2,000	0.287	0.015	0.302*
	2,500	-	-	-
	3,000	-	-	-

ตาราง 28 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไคส์ลิสต์-อีวริสติก	แมกซิมัม ไคส์ลิสต์-เบส	อีวริสติก-เบส
80	200	-	-	-
	300	0.057	0.501*	0.558*
	400	-	-	-
	500	0.226	0.426*	0.200
	600	-	-	-
	700	0.046	0.403*	0.357*
	800	-	-	-
	900	0.314	0.370*	0.056
	1,000	-	-	-
	1,500	-	-	-
	2,000	0.281	0.071	0.352*
	2,500	-	-	-
	3,000	-	-	-
90	200	0.263	0.700*	0.963*
	300	0.160	0.611*	0.771*
	400	0.109	0.497*	0.388*
	500	0.102	0.556*	0.454*
	600	-	-	-
	700	0.160	0.453*	0.293
	800	0.352*	0.393*	0.041
	900	0.282	0.380*	0.098
	1,000	0.215	0.369*	0.154
	1,500	-	-	-
	2,000	0.258	0.414*	0.156

ตาราง 28 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไคส์ลียูด-อีวริสติก	แมกซิมัม ไคส์ลียูด-เบส	อีวริสติก-เบส
90	2,500	0.189	0.333*	0.144
	3,000	-	-	-
100	200	0.321	0.452	0.773*
	300	0.479	0.459	0.938*
	400	-	-	-
	500	0.268	0.401	0.669*
	600	-	-	-
	700	0.294	0.432	0.726*
	800	-	-	-
	900	-	-	-
	1,000	0.055	0.379	0.434*
	1,500	-	-	-
	2,000	0.091	0.407	0.498*
	2,500	-	-	-
	3,000	-	-	-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 28 พบว่า ค่าอำนาจจำแนกของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 600 คน และค่าอำนาจจำแนกของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 800 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไคส์ลียูด และวิธีอีวริสติก มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าอำนาจจำแนกของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน ค่าอำนาจจำแนกของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 500 คน และ 700 คน ค่าอำนาจ

จำแนกของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 700 คน 800 คน 900 คน และ 1,000 คน ค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 300 คน 500 คน 700 คน และ 900 คน และค่าอำนาจจำแนกของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 500 คน 700 คน 800 คน 900 คน 1,000 คน 2,000 คน และ 2,500 คน ที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิฮูด และวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าอำนาจจำแนกของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน และ 600 คน ค่าอำนาจจำแนกของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 600 คน ค่าอำนาจจำแนกของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 1,000 คน และ 2,000 คน ค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 300 คน 700 คน และ 2,000 คน และค่าอำนาจจำแนกของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 500 คน 700 คน 1,000 คน และ 2,000 คน ที่ประมาณค่าด้วยวิธีอีวริสติก และวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.5 ผลการเปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 29

ตาราง 29 เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _b	MS _w	F-test
50	12	20.705	1.725	6.211	0.278
60	12	11.242	0.937	6.403	0.146
70	12	12.133	1.011	6.385	0.158
80	12	6.321	0.527	5.601	0.094
90	12	5.378	0.448	5.765	0.078
100	12	7.136	0.595	5.747	0.104

จากตาราง 29 พบว่า ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.6 ผลการเปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 30

ตาราง 30 เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _B	MS _w	F-test
50	12	307.255	25.605	19.612	1.306
60	12	4,952.636	412.720	397.344	1.039
70	12	546.430	45.536	30.578	1.489
80	12	159.431	13.286	12.139	1.095
90	12	824.003	68.667	61.271	1.121
100	12	113.683	9.474	13.595	0.697

จากตาราง 30 พบว่า ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.7 ผลการเปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 31

ตาราง 31 เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _b	MS _w	F-test
50	12	38.004	3.167	3.333	0.950
60	12	43.534	3.628	2.767	1.311
70	12	26.098	2.175	2.692	0.808
80	12	41.291	3.441	2.814	1.223
90	12	48.287	4.024	3.514	1.145
100	12	36.301	3.025	2.694	1.123

จากตาราง 31 พบว่า ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.8 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม โคล์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 32 - 33

ตาราง 32 เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธี
แมกซิมัม ไลค์ลิสต์ วิธีทิวาสติค และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
50	200	2	300.716	150.358	29.966	5.018*
	300	2	73.163	36.582	4.883	7.492*
	400	2	58.224	29.112	4.656	6.253*
	500	2	65.253	32.627	5.436	6.002*
	600	2	70.775	35.388	2.931	12.074*
	700	2	33.936	16.968	3.191	5.318*
	800	2	28.275	14.137	5.777	2.447
	900	2	110.217	55.109	3.620	15.224*
	1,000	2	12.771	6.385	31.891	0.200
	1,500	2	32.368	16.184	6.178	2.620
	2,000	2	59.365	29.682	10.510	2.824
	2,500	2	120.720	60.360	10.961	5.507*
	3,000	2	39.707	19.853	6.341	3.131
60	200	2	4,452.181	2,226.090	1,673.513	1.330
	300	2	109.013	54.506	3.915	13.923*
	400	2	85.761	42.880	10.329	4.151
	500	2	83.470	41.735	5.839	7.148*
	600	2	92.700	46.350	3.873	11.966*
	700	2	51.108	25.554	3.953	6.465*
	800	2	101.013	50.507	3.430	14.725*
	900	2	115.503	57.752	4.146	13.930*
	1,000	2	159.243	79.621	6.860	11.607*
	1,500	2	55.473	27.736	4.762	5.825*

ตาราง 32 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
60	2,000	2	65.247	32.623	4.362	7.479*
	2,500	2	16.170	8.085	32.172	0.251
	3,000	2	61.889	30.944	4.410	7.017*
70	200	2	87.214	43.607	101.445	0.430
	300	2	177.844	88.922	6.871	12.942*
	400	2	19.860	9.930	19.415	0.512
	500	2	90.965	45.482	4.170	10.907*
	600	2	139.008	69.504	5.258	13.219*
	700	2	132.661	66.331	3.572	18.568*
	800	2	108.376	54.188	2.902	18.675*
	900	2	139.595	69.797	3.914	17.832*
	1,000	2	124.333	62.167	3.567	17.431*
	1,500	2	117.183	58.592	4.885	11.995*
	2,000	2	74.950	37.475	5.833	6.424*
	2,500	2	58.634	29.317	5.668	5.173*
	3,000	2	86.527	43.264	4.337	9.976*
80	200	2	81.526	40.763	18.990	2.147
	300	2	122.060	61.030	4.131	14.152*
	400	2	165.815	82.908	6.019	13.774*
	500	2	115.335	57.667	5.433	10.615*
	600	2	242.657	121.329	8.111	14.960*
	700	2	58.375	29.188	10.585	2.757
	800	2	144.082	72.041	2.194	32.839*
	900	2	226.759	113.380	8.780	12.914*

ตาราง 32 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
80	1,000	2	150.805	75.402	4.259	17.705*
	1,500	2	89.179	44.589	4.275	10.431*
	2,000	2	115.327	57.664	4.083	14.122*
	2,500	2	108.060	54.030	4.475	12.075*
	3,000	2	142.156	71.078	5.010	14.190*
90	200	2	29.761	14.881	124.363	0.120
	300	2	151.807	75.904	4.707	15.913*
	400	2	269.566	134.783	15.118	8.915*
	500	2	162.039	81.020	4.660	17.386*
	600	2	144.603	72.301	4.768	15.164*
	700	2	102.475	51.237	3.710	13.812*
	800	2	144.204	72.102	3.080	23.407*
	900	2	676.224	338.112	87.208	3.877*
	1,000	2	174.204	87.102	4.217	20.656*
	1,500	2	92.636	46.318	3.821	12.124*
	2,000	2	142.302	71.151	6.961	10.221*
	2,500	2	138.326	69.163	5.866	11.791*
	3,000	2	315.797	157.898	37.176	4.247*
100	200	2	119.704	59.852	12.175	4.916*
	300	2	167.993	83.997	4.421	19.001*
	400	2	194.104	97.052	7.485	12.967*
	500	2	233.906	116.953	4.825	24.240*
	600	2	208.308	104.154	4.563	22.825*
	700	2	135.675	67.838	3.507	19.345*

ตาราง 32 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
100	800	2	164.875	82.438	2.360	34.933*
	900	2	148.121	74.061	13.133	5.639*
	1,000	2	361.646	180.823	20.317	8.900*
	1,500	2	117.428	58.714	3.764	15.599*
	2,000	2	178.579	89.290	5.798	15.401*
	2,500	2	165.745	82.872	4.338	19.104*
	3,000	2	73.492	36.746	8.802	4.175*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 32 พบว่า ค่าความยากของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่างขนาด 800 คน 1,000 คน 1,500 คน 2,000 คน และ 3,000 คน ค่าความยากของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 200 คน 400 คน และ 2,500 คน ค่าความยากของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 400 คน ค่าความยากของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 700 คน ค่าความยากของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 200 คน ที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จึงได้ทำการทดสอบความแตกต่างของค่าความยากของข้อสอบดังกล่าว ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 33

ตาราง 33 การทดสอบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธี
แมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน
ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส์	อีวริสติก-เบส์
50	200	3.229*	0.519	2.710*
	300	1.679*	0.556	1.123*
	400	1.526*	0.750	0.776
	500	1.567*	0.443	1.124
	600	1.650*	0.532	1.112*
	700	1.146*	0.390	0.756
	800	-	-	-
	900	2.096*	0.941	1.155*
	1,000	-	-	-
	1,500	-	-	-
	2,000	-	-	-
	2,500	2.081*	0.430	1.651*
	3,000	-	-	-
60	200	-	-	-
	300	1.904	0.863	1.041*
	400	-	-	-
	500	1.592*	0.364	1.228*
	600	1.720*	0.543	1.177*
	700	1.305*	0.622	0.683
	800	1.802*	0.601	1.201*
	900	1.924*	0.627	1.297*
	1,000	2.153*	0.365	1.788*

ตาราง 33 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไคส์ลียูต-อีวริสติก	แมกซิมัม ไคส์ลียูต-เบส	อีวริสติก-เบส
60	1,500	1.352*	0.547	0.805
	2,000	1.475*	0.719	0.756
	2,500	-	-	-
	3,000	1.428*	0.587	0.841
70	200	-	-	-
	300	2.200*	0.675	1.525*
	400	-	-	-
	500	1.555*	0.409	1.146*
	600	1.963*	0.686	1.277*
	700	1.860*	0.432	1.428*
	800	1.703*	0.466	1.159*
	900	1.944*	0.575	1.369*
	1,000	1.881*	0.840*	1.041*
	1,500	1.799*	0.606	1.165*
	2,000	1.452*	0.881	0.571
	2,500	1.261*	0.380	0.881
	3,000	1.517*	0.401	1.116*
80	200	-	-	-
	300	1.743*	0.770	0.973*
	400	1.983*	0.592	1.391*
	500	1.614*	0.351	1.263*
	600	2.345*	0.519	1.826*
	700	-	-	-
	800	1.880*	0.712*	1.168*

ตาราง 33 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
80	900	2.292*	0.587	1.705*
	1,000	1.811*	0.297	1.514*
	1,500	1.493*	0.792	0.701
	2,000	1.694*	0.948*	0.746
	2,500	1.630*	0.628	1.002*
	3,000	1.750*	0.269	0.683*
90	200	-	-	-
	300	0.802*	0.212	0.590*
	400	2.251*	0.294	1.957*
	500	1.845*	0.536	1.309*
	600	1.751*	0.552	1.119*
	700	1.497*	0.584	0.913*
	800	1.738*	0.496	1.242*
	900	3.578*	0.498	3.080
	1,000	1.942*	0.699	1.243*
	1,500	1.362*	0.292	1.070*
	2,000	1.753*	1.135*	0.618
	2,500	1.695*	0.460	1.235*
	3,000	2.568*	0.722	1.846
100	200	1.482*	1.126	0.356
	300	1.833*	0.881*	0.952*
	400	1.949*	0.726	1.223*
	500	2.100*	0.600	1.500*
	600	2.033*	0.859	1.174*

ตาราง 33 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
100	700	1.626*	0.582	1.044*
	800	1.806*	0.737*	1.069*
	900	1.711*	0.697	1.014
	1,000	2.461*	0.292	2.169*
	1,500	1.530*	0.701*	0.829*
	2,000	1.867*	0.680	1.187*
	2,500	1.786*	0.584	1.202*
	3,000	1.212*	0.607	0.501

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 33 พบว่า ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ และวิธีอีวริสติก มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้น ค่าความยากของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 800 คน 1,000 คน 1,500 คน 2,000 คน และ 3,000 คน ค่าความยากของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน และ 2,500 คน ค่าความยากของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 400 คน ค่าความยากของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 700 คน และค่าความยากของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 200 คน ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ค่าความยากของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 1,000 คน และ 2,000 คน ค่าความยากของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 800 คน ค่าความยากของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 2,000 คน และค่าความยากของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 300 คน 800 คน และ 1,500 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ และวิธีของเบส มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าความยากของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 600 คน 900 คน และ 2,500 คน ค่าความยากของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 300 คน 500 คน 600 คน 800 คน 900 คน และ 1,000 คน ค่าความยากของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 300 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน 1,500 คน และ 3,000 คน ค่าความยากของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 500 คน 700 คน 800 คน 900 คน 2,500 คน และ 3,000 คน ค่าความยากของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 1,000 คน 1,500 คน และ 2,500 คน ค่าความยากของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 1,000 คน 1,500 คน 2,000 คน และ 2,500 คน ที่ประมาณค่าด้วยวิธีวิธีวิสถิติ และวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.9 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าการเดาของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง

ตาราง 34 เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธี
แมกซีมัม ไลค์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _b	MS _w	F-test
50	12	0.185	0.015	0.003	5.856*
60	12	0.240	0.020	0.003	7.965*
70	12	0.193	0.016	0.003	5.874*
80	12	0.191	0.016	0.003	5.926*
90	12	0.242	0.020	0.003	7.608*
100	12	0.337	0.028	0.003	9.960*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 34 พบว่า ค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธี
แมกซีมัม ไลค์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญ
ทางสถิติที่ระดับ .05 จึงทำการทดสอบความแตกต่างของค่าการเคาะของข้อสอบด้วยการ
เปรียบเทียบพหุคูณ ได้ผลดังตาราง 35 - 40

ตาราง 35 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 50 ข้อ ที่ประมาณค่าด้วย
วิธีแมกซีมัม ไลค์ลียูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.247	0.226	0.243	0.251	0.253	0.230	0.220	0.241	0.216	0.221	0.209	0.202	0.206
200	0.247	-	0.021	0.004	0.004	0.006	0.017	0.027	0.006	0.031	0.026	0.038	0.045	0.041
300	0.226		-	0.017	0.025	0.027	0.004	0.006	0.015	0.010	0.005	0.017	0.024	0.020
400	0.243			-	0.008	0.010	0.013	0.023	0.002	0.027	0.022	0.034	0.041	0.037
500	0.251				-	0.002	0.021	0.031	0.010	0.035	0.030	0.042	0.049*	0.045
600	0.253					-	0.023	0.033	0.012	0.037	0.032	0.044	0.051*	0.047
700	0.230						-	0.010	0.011	0.014	0.009	0.021	0.028	0.024
800	0.220							-	0.021	0.004	0.001	0.011	0.018	0.014
900	0.241								-	0.025	0.020	0.032	0.039	0.035
1,000	0.216									-	0.005	0.007	0.014	0.010
1,500	0.221										-	0.012	0.019	0.015
2,000	0.209											-	0.007	0.003
2,500	0.202												-	0.004
3,000	0.206													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 35 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 50 ข้อ ของกลุ่ม
ตัวอย่าง 500 คน และ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน อย่างมี
นัยสำคัญทางสถิติที่ระดับ .05

ตาราง 36 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 60 ข้อ ที่ประมาณค่าด้วย
วิธีแมกซิมัม ไลค์ลิชชูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.231	0.219	0.233	0.243	0.241	0.214	0.210	0.230	0.202	0.209	0.198	0.189	0.191
200	0.231	-	0.012	0.002	0.012	0.010	0.017	0.021	0.001	0.029	0.022	0.033	0.042	0.040
300	0.219		-	0.014	0.024	0.022	0.005	0.009	0.011	0.017	0.010	0.021	0.030	0.028
400	0.233			-	0.010	0.008	0.019	0.023	0.003	0.031	0.024	0.035	0.044*	0.042
500	0.243				-	0.002	0.029	0.033	0.013	0.041	0.034	0.045*	0.054*	0.052
600	0.241					-	0.027	0.031	0.011	0.039	0.032	0.043*	0.052*	0.050
700	0.214						-	0.004	0.016	0.012	0.005	0.016	0.025	0.023
800	0.210							-	0.020	0.008	0.001	0.012	0.021	0.019
900	0.230								-	0.028	0.021	0.032	0.041	0.039
1,000	0.202									-	0.007	0.004	0.013	0.011
1,500	0.209										-	0.011	0.020	0.018
2,000	0.198											-	0.009	0.007
2,500	0.189												-	0.002
3,000	0.191													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 36 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 60 ข้อ ของกลุ่ม
ตัวอย่าง 400 คน 500 คน และ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน
และ 3,000 คน ค่าการเคาะเฉลี่ยของกลุ่มตัวอย่าง 500 คน และ 600 คน มีความแตกต่าง
จากกลุ่มตัวอย่าง 2,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 37 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 70 ข้อ ที่ประมาณค่าด้วย
วิธีแมกซีมัม ไลค์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.224	0.212	0.232	0.231	0.233	0.210	0.210	0.225	0.203	0.209	0.197	0.192	0.189
200	0.224	-	0.012	0.008	0.007	0.009	0.014	0.014	0.001	0.021	0.015	0.027	0.032	0.035
300	0.212		-	0.020	0.019	0.021	0.002	0.002	0.013	0.009	0.003	0.015	0.020	0.023
400	0.232			-	0.001	0.001	0.022	0.022	0.007	0.029	0.023	0.035	0.040	0.043
500	0.231				-	0.002	0.021	0.021	0.006	0.028	0.022	0.034	0.039	0.042
600	0.233					-	0.023	0.023	0.008	0.030	0.024	0.036	0.041	0.044
700	0.210						-	0.000	0.015	0.007	0.001	0.013	0.018	0.021
800	0.210							-	0.015	0.007	0.001	0.013	0.018	0.021
900	0.225								-	0.022	0.016	0.028	0.033	0.036
1,000	0.203									-	0.006	0.006	0.011	0.014
1,500	0.209										-	0.012	0.017	0.020
2,000	0.197											-	0.005	0.008
2,500	0.192												-	0.003
3,000	0.189													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 37 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 70 ข้อ ของกลุ่ม
ตัวอย่าง 400 คน 500 คน และ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 3,000 คน
อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 38 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิชต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	ค	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.219	0.215	0.220	0.235	0.226	0.208	0.204	0.217	0.201	0.200	0.193	0.189	0.192
200	0.219	-	0.004	0.001	0.016	0.007	0.011	0.015	0.002	0.018	0.019	0.026	0.030	0.027
300	0.215		-	0.005	0.020	0.011	0.007	0.011	0.002	0.014	0.015	0.022	0.026	0.023
400	0.220			-	0.015	0.006	0.012	0.016	0.003	0.019	0.020	0.027	0.031	0.028
500	0.235				-	0.009	0.027	0.031	0.018	0.034	0.035	0.042*	0.046*	0.043*
600	0.226					-	0.018	0.022	0.009	0.025	0.026	0.033	0.037	0.034
700	0.208						-	0.004	0.009	0.007	0.008	0.015	0.019	0.016
800	0.204							-	0.013	0.003	0.004	0.011	0.015	0.012
900	0.217								-	0.016	0.017	0.024	0.028	0.025
1,000	0.201									-	0.001	0.008	0.012	0.009
1,500	0.200										-	0.007	0.011	0.008
2,000	0.193											-	0.004	0.001
2,500	0.189												-	0.003
3,000	0.192													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 38 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน 2,500 คน และ 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 39 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 90 ข้อ ที่ประมาณค่าด้วย
วิธีแฟกทิม ไลค์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.230	0.223	0.234	0.241	0.224	0.218	0.207	0.228	0.213	0.213	0.201	0.197	0.191
200	0.230	-	0.007	0.004	0.011	0.006	0.012	0.023	0.002	0.017	0.017	0.029	0.033	0.039
300	0.223		-	0.011	0.018	0.001	0.005	0.016	0.005	0.010	0.010	0.022	0.026	0.032
400	0.234			-	0.007	0.010	0.016	0.027	0.006	0.021	0.021	0.033	0.037*	0.043
500	0.241				-	0.017	0.023	0.034	0.013	0.028	0.028	0.040*	0.044*	0.050
600	0.224					-	0.006	0.017	0.004	0.011	0.011	0.023	0.027	0.033
700	0.218						-	0.011	0.010	0.005	0.005	0.017	0.021	0.027
800	0.207							-	0.021	0.006	0.006	0.006	0.010	0.016
900	0.228								-	0.015	0.015	0.027	0.031	0.037
1,000	0.213									-	0.000	0.012	0.016	0.022
1,500	0.213										-	0.012	0.016	0.022
2,000	0.201											-	0.004	0.010
2,500	0.197												-	0.006
3,000	0.191													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 39 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 90 ข้อ ของกลุ่ม
ตัวอย่าง 200 คน มีความแตกต่างจากกลุ่มตัวอย่าง 3,000 คน ค่าการเคาะเฉลี่ยของกลุ่ม
ตัวอย่าง 400 คน และ 500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน และค่าการ
เคาะเฉลี่ยของกลุ่มตัวอย่าง 500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน อย่างมี
นัยสำคัญทางสถิติที่ระดับ .05

ตาราง 40 การทดสอบความแตกต่างของค่าการเดาของข้อสอบ 100 ข้อ ที่ประมาณค่า
ด้วยวิธีแมกซีมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.253	0.227	0.239	0.239	0.223	0.222	0.207	0.225	0.216	0.215	0.206	0.201	0.193
200	0.253	-	0.026	0.014	0.014	0.030	0.031	0.046*	0.028	0.037*	0.038*	0.047*	0.052*	0.060
300	0.227		-	0.012	0.012	0.004	0.005	0.020	0.002	0.011	0.012	0.021	0.026	0.034
400	0.239			-	0.000	0.016	0.017	0.032	0.014	0.023	0.024	0.033	0.038*	0.046
500	0.239				-	0.016	0.017	0.032	0.014	0.023	0.024	0.033	0.038*	0.046
600	0.223					-	0.001	0.016	0.002	0.007	0.008	0.017	0.022	0.030
700	0.222						-	0.015	0.003	0.006	0.007	0.016	0.021	0.029
800	0.207							-	0.018	0.009	0.008	0.001	0.006	0.014
900	0.225								-	0.009	0.010	0.019	0.024	0.032
1,000	0.216									-	0.001	0.010	0.015	0.023
1,500	0.215										-	0.009	0.014	0.022
2,000	0.206											-	0.005	0.013
2,500	0.201												-	0.008
3,000	0.193													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 40 พบว่า ค่าการเดาเฉลี่ยของข้อสอบ 100 ข้อ ของกลุ่ม
ตัวอย่าง 200 คน 400 คน และ 500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน
และ 3,000 คน และค่าการเดาเฉลี่ยของกลุ่มตัวอย่าง 200 คน มีความแตกต่างจากกลุ่ม
ตัวอย่าง 800 คน 1,000 คน 1,500 คน และ 2,000 คน อย่างมีนัยสำคัญทางสถิติที่
ระดับ .05

3.10 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 41 - 47

ตาราง 41 เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _B	MS _w	F-test
50	12	0.285	0.024	0.006	4.013*
60	12	0.660	0.055	0.009	6.017*
70	12	0.556	0.046	0.008	5.928*
80	12	5.123	4.594	0.754	6.096*
90	12	0.555	0.046	0.006	7.206*
100	12	1.684	0.140	0.009	15.348*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 41 พบว่า ค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จึงทำการทดสอบความแตกต่างของค่าการเคาะของข้อสอบด้วยการเปรียบเทียบ พบว่า ได้ผลดังตาราง 42 - 47

ตาราง 42 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 50 ข้อ ที่ประมาณค่าด้วย
วิธีวิธีสถิติ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.126	0.093	0.108	0.107	0.105	0.085	0.149	0.142	0.159	0.131	0.128	0.111	0.127
200	0.126	-	0.033	0.018	0.019	0.021	0.041	0.023	0.016	0.033	0.005	0.002	0.015	0.001
300	0.093		-	0.015	0.014	0.012	0.008	0.056	0.049	0.066*	0.038	0.035	0.018	0.034
400	0.108			-	0.001	0.003	0.023	0.041	0.034	0.051	0.023	0.020	0.003	0.019
500	0.107				-	0.002	0.022	0.042	0.035	0.052*	0.024	0.021	0.004	0.020
600	0.105					-	0.020	0.044	0.037	0.054*	0.026	0.023	0.006	0.022
700	0.085						-	0.064	0.057	0.074*	0.046	0.043	0.026	0.042
800	0.149							-	0.007	0.010	0.018	0.021	0.038	0.022
900	0.142								-	0.017	0.011	0.014	0.031	0.015
1,000	0.159									-	0.028	0.031	0.048	0.032
1,500	0.131										-	0.003	0.020	0.004
2,000	0.128											-	0.017	0.001
2,500	0.111												-	0.016
3,000	0.127													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 42 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 50 ข้อ ของกลุ่ม
ตัวอย่าง 700 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,000 คน อย่างมีนัยสำคัญทางสถิติที่
ระดับ .05

ตาราง 43 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 60 ข้อ ที่ประมาณค่าด้วยวิธีธีวรีสถิติ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.124	0.086	0.126	0.099	0.103	0.090	0.132	0.156	0.194	0.133	0.162	0.128	0.127
200	0.124	-	0.038	0.002	0.025	0.021	0.034	0.008	0.032	0.070	0.009	0.038	0.004	0.003
300	0.086		-	0.040	0.013	0.017	0.004	0.046	0.070	0.108	0.047	0.076	0.042	0.041
400	0.126			-	0.027	0.023	0.036	0.006	0.030	0.068	0.007	0.036	0.002	0.001
500	0.099				-	0.004	0.009	0.033	0.057	0.095	0.034	0.063	0.029	0.028
600	0.103					-	0.013	0.029	0.053	0.091	0.030	0.059	0.025	0.024
700	0.090						-	0.042	0.066	0.104	0.043	0.072	0.038	0.037
800	0.132							-	0.024	0.062	0.001	0.030	0.004	0.005
900	0.156								-	0.038	0.023	0.006	0.028	0.029
1,000	0.194									-	0.061	0.032	0.066	0.067
1,500	0.133										-	0.029	0.005	0.006
2,000	0.162											-	0.034	0.035
2,500	0.128												-	0.001
3,000	0.127													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 43 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 300 คน 500 คน 600 คน และ 700 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 44 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 70 ข้อ ที่ประมาณค่าด้วย
วิธีอีวริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.126	0.099	0.118	0.100	0.112	0.098	0.145	0.128	0.140	0.180	0.110	0.146	0.165
200	0.126	-	0.027	0.008	0.026	0.014	0.028	0.019	0.002	0.014	0.054	0.016	0.020	0.039
300	0.099		-	0.019	0.001	0.013	0.001	0.046	0.029	0.041	0.081*	0.011	0.047	0.066
400	0.118			-	0.018	0.006	0.020	0.027	0.010	0.022	0.062	0.008	0.028	0.047
500	0.100				-	0.012	0.002	0.045	0.028	0.040	0.080*	0.010	0.046	0.065
600	0.112					-	0.014	0.033	0.016	0.028	0.068	0.002	0.034	0.053
700	0.098						-	0.047	0.030	0.042	0.082*	0.012	0.048	0.067
800	0.145							-	0.017	0.005	0.035	0.035	0.001	0.020
900	0.128								-	0.012	0.052	0.018	0.018	0.037
1,000	0.140									-	0.040	0.030	0.006	0.025
1,500	0.180										-	0.070*	0.034	0.015
2,000	0.110											-	0.036	0.055
2,500	0.146												-	0.019
3,000	0.165													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 44 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 70 ข้อ ของกลุ่ม
ตัวอย่าง 300 คน 500 คน และ 700 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน
และค่าการเคาะเฉลี่ยของกลุ่มตัวอย่าง 1,500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000
คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 45 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีอิวริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.130	0.103	0.107	0.096	0.113	0.086	0.139	0.132	0.137	0.135	0.114	0.115	1.000
200	0.130	-	0.027	0.023	0.034	0.017	0.044	0.009	0.002	0.007	0.005	0.016	0.015	0.870
300	0.103		-	0.004	0.007	0.010	0.017	0.036	0.029	0.034	0.032	0.011	0.012	0.897
400	0.107			-	0.011	0.006	0.021	0.032	0.025	0.030	0.028	0.007	0.008	0.893
500	0.096				-	0.017	0.010	0.043	0.036	0.041	0.039	0.018	0.019	0.904
600	0.113					-	0.027	0.026	0.019	0.024	0.022	0.001	0.002	0.887
700	0.086						-	0.053	0.046	0.051	0.049	0.028	0.029	0.914
800	0.139							-	0.007	0.002	0.004	0.025	0.024	0.861
900	0.132								-	0.005	0.003	0.018	0.017	0.868
1,000	0.137									-	0.002	0.023	0.022	0.863
1,500	0.135										-	0.021	0.020	0.865
2,000	0.114											-	0.001	0.886
2,500	0.115												-	0.885
3,000	1.000													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 45 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 200 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน และ 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 46 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 90 ข้อ ที่ประมาณค่าด้วย
วิธีธีวริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.152	0.100	0.103	0.084	0.117	0.093	0.127	0.130	0.134	0.151	0.121	0.131	0.154
200	0.152	-	0.052	0.049	0.068*	0.035	0.059*	0.025	0.022	0.018	0.001	0.031	0.021	0.002
300	0.100		-	0.003	0.016	0.017	0.007	0.027	0.030	0.034	0.051	0.021	0.031	0.054
400	0.103			-	0.019	0.014	0.010	0.024	0.027	0.031	0.048	0.018	0.028	0.051
500	0.084				-	0.033	0.009	0.043	0.046	0.050	0.067*	0.037	0.047	0.070*
600	0.117					-	0.024	0.010	0.013	0.017	0.034	0.004	0.014	0.037
700	0.093						-	0.034	0.037	0.041	0.058*	0.028	0.038	0.061*
800	0.127							-	0.003	0.007	0.024	0.006	0.004	0.027
900	0.130								-	0.004	0.021	0.009	0.001	0.024
1,000	0.134									-	0.017	0.013	0.003	0.020
1,500	0.151										-	0.030	0.020	0.003
2,000	0.121											-	0.010	0.033
2,500	0.131												-	0.023
3,000	0.154													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 46 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 90 ข้อ ของกลุ่ม
ตัวอย่าง 200 คน มีความแตกต่างจากกลุ่มตัวอย่าง 500 คน และ 700 คน และค่าการเคาะ
เฉลี่ยของกลุ่มตัวอย่าง 500 คน และ 700 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน
และ 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 47 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 100 ข้อ ที่ประมาณค่าด้วย
วิธีนิวาสติค เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.175	0.120	0.090	0.085	0.120	0.110	0.167	0.156	0.210	0.123	0.183	0.138	0.158
200	0.175	-	0.055	0.085*0.090	0.055	0.065*0.008	0.019	0.035	0.052	0.008	0.037	0.017		
300	0.120		-	0.030	0.035	0.000	0.010	0.047	0.036	0.090*0.003	0.063*0.018	0.038		
400	0.090			-	0.005	0.030	0.020	0.077*0.066	0.120*0.033	0.093*0.048	0.068			
500	0.085				-	0.035	0.025	0.082*0.071	0.125*0.038	0.098*0.053	0.073			
600	0.120					-	0.010	0.047	0.036	0.090*0.003	0.063*0.018	0.038		
700	0.110						-	0.057	0.046	0.100*0.013	0.073*0.028	0.048		
800	0.167							-	0.011	0.043	0.044	0.016	0.029	0.009
900	0.156								-	0.054	0.033	0.027	0.018	0.002
1,000	0.210									-	0.087*0.027	0.072*0.052		
1,500	0.123										-	0.060	0.015	0.035
2,000	0.183											-	0.045	0.025
2,500	0.138												-	0.020
3,000	0.158													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 47 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 100 ข้อ ของ
กลุ่มตัวอย่าง 200 คน มีความแตกต่างจากกลุ่มตัวอย่าง 400 คน 500 คน และ 700 คน
ค่าการเคาะเฉลี่ยของกลุ่มตัวอย่าง 300 คน 400 คน 500 คน 600 คน และ 700 คน มี
ความแตกต่างจากกลุ่มตัวอย่าง 1,000 คน และ 2,000 คน และค่าการเคาะเฉลี่ยของกลุ่ม
ตัวอย่าง 400 คน 500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 800 คน 900 คน และ 3000

คน และค่าการเดาเฉลี่ยของกลุ่มตัวอย่าง 1,000 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน และ 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.11 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าการเดาของข้อสอบ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 48 - 53

ตาราง 48 เปรียบเทียบความแตกต่างของค่าการเดาของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _B	MS _w	F-test
50	12	0.714	0.060	0.008	7.560*
60	12	0.612	0.051	0.006	7.954*
70	12	0.823	0.069	0.007	9.383*
80	12	0.686	0.057	0.006	9.854*
90	12	1.118	0.093	0.010	9.013*
100	12	0.190	0.016	0.006	2.606

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 48 พบว่า ค่าการเดาของข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ และ 90 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังนั้น จึงได้ทำการทดสอบความแตกต่างของค่าการเดาของข้อสอบดังกล่าว ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 49 - 53

ตาราง 49 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 50 ข้อ ที่ประมาณค่าด้วย
วิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.187	0.232	0.226	0.217	0.249	0.239	0.239	0.234	0.243	0.252	0.149	0.296	0.237
200	0.187	-	0.045	0.039	0.030	0.062	0.052	0.052	0.047	0.056	0.065	0.038	0.109*	0.050
300	0.232		-	0.006	0.015	0.017	0.007	0.007	0.002	0.011	0.020	0.083*	0.064	0.005
400	0.226			-	0.009	0.023	0.013	0.013	0.008	0.017	0.026	0.077	0.070	0.011
500	0.217				-	0.032	0.022	0.022	0.017	0.026	0.035	0.068	0.079	0.020
600	0.249					-	0.010	0.010	0.015	0.006	0.003	0.100*	0.047	0.012
700	0.239						-	0.000	0.005	0.004	0.013	0.090*	0.057	0.002
800	0.239							-	0.005	0.004	0.013	0.090*	0.057	0.002
900	0.234								-	0.009	0.018	0.085*	0.062	0.003
1,000	0.243									-	0.009	0.094*	0.053	0.006
1,500	0.252										-	0.103*	0.044	0.015
2,000	0.149											-	0.147*	0.088*
2,500	0.296												-	0.059
3,000	0.237													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 49 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 50 ข้อ ของ
กลุ่มตัวอย่าง 200 คน และ 2,000 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน และ
ค่าการเคาะเฉลี่ยของกลุ่มตัวอย่าง 300 คน 600 คน 700 คน 800 คน 900 คน 1,000
คน และ 1,500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน อย่างมีนัยสำคัญทางสถิติ
ที่ระดับ .05

ตาราง 50 การทดสอบความแตกต่างของค่าการเดาของข้อสอบ 60 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.192	0.237	0.210	0.225	0.221	0.246	0.234	0.243	0.232	0.246	0.147	0.253	0.247
200	0.192	-	0.045	0.018	0.033	0.029	0.054	0.042	0.051	0.040	0.054	0.045	0.061	0.055
300	0.237		-	0.027	0.012	0.016	0.009	0.003	0.006	0.005	0.009	0.090*	0.016	0.010
400	0.210			-	0.015	0.011	0.036	0.024	0.033	0.022	0.036	0.063	0.043	0.037
500	0.225				-	0.004	0.021	0.009	0.018	0.007	0.021	0.078*	0.028	0.022
600	0.221					-	0.025	0.013	0.022	0.011	0.025	0.074*	0.032	0.026
700	0.246						-	0.012	0.003	0.014	0.000	0.099*	0.007	0.001
800	0.234							-	0.009	0.002	0.012	0.087*	0.019	0.013
900	0.243								-	0.011	0.003	0.096*	0.010	0.004
1,000	0.232									-	0.014	0.085*	0.021	0.015
1,500	0.246										-	0.099*	0.007	0.001
2,000	0.147											-	0.106*	0.100*
2,500	0.253												-	0.006
3,000	0.247													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 50 พบว่า ค่าการเดาเฉลี่ยของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 300 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน และ 1,500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน และค่าการเดาเฉลี่ยของกลุ่มตัวอย่าง 2,000 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน และ 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 51 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 70 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.200	0.221	0.214	0.227	0.233	0.235	0.244	0.242	0.243	0.242	0.133	0.251	0.245
200	0.200	-	0.021	0.014	0.027	0.033	0.035	0.044	0.042	0.043	0.042	0.067*	0.051	0.045
300	0.221		-	0.007	0.006	0.012	0.014	0.023	0.021	0.022	0.021	0.088*	0.030	0.024
400	0.214			-	0.013	0.019	0.021	0.030	0.028	0.029	0.028	0.081*	0.037	0.031
500	0.227				-	0.006	0.008	0.017	0.015	0.016	0.015	0.094*	0.024	0.018
600	0.233					-	0.002	0.011	0.009	0.010	0.009	0.100*	0.018	0.012
700	0.235						-	0.009	0.007	0.008	0.007	0.102*	0.016	0.010
800	0.244							-	0.002	0.001	0.002	0.111*	0.007	0.001
900	0.242								-	0.001	0.000	0.109*	0.009	0.003
1,000	0.243									-	0.001	0.110*	0.008	0.002
1,500	0.242										-	0.109*	0.009	0.003
2,000	0.133											-	0.118*	0.112*
2,500	0.251												-	0.006
3,000	0.245													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 51 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน และ 1,500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน และค่าการเคาะเฉลี่ยของกลุ่มตัวอย่าง 2,000 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน และ 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 52 การทดสอบความแตกต่างของค่าการเคาของข้อสอบ 80 ข้อ ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.198	0.227	0.200	0.228	0.242	0.245	0.242	0.246	0.229	0.257	0.160	0.243	0.239
200	0.198	-	0.029	0.002	0.030	0.044	0.047	0.044	0.048	0.031	0.059*	0.038	0.045	0.041
300	0.227		-	0.027	0.001	0.015	0.018	0.015	0.019	0.002	0.030	0.067*	0.016	0.012
400	0.200			-	0.028	0.042	0.045	0.042	0.046	0.029	0.057*	0.040	0.043	0.039
500	0.228				-	0.014	0.017	0.014	0.018	0.001	0.029	0.068*	0.015	0.011
600	0.242					-	0.003	0.000	0.004	0.013	0.015	0.082*	0.001	0.003
700	0.245						-	0.003	0.001	0.016	0.012	0.085*	0.002	0.006
800	0.242							-	0.004	0.013	0.015	0.082*	0.001	0.003
900	0.246								-	0.017	0.011	0.086*	0.003	0.007
1,000	0.229									-	0.028	0.069*	0.014	0.010
1,500	0.257										-	0.097*	0.014	0.018
2,000	0.160											-	0.083*	0.079
2,500	0.243												-	0.004
3,000	0.239													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 52 พบว่า ค่าการเคาเฉลี่ยของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 400 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน ค่าการเคาเฉลี่ยของกลุ่มตัวอย่าง 300 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน และ 1,500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน และค่าการเคาเฉลี่ยของกลุ่มตัวอย่าง 2,000 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน และ 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 53 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบ 90 ข้อ ที่ประมาณค่าด้วย
วิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.206	0.230	0.222	0.216	0.249	0.239	0.245	0.246	0.227	0.244	0.170	0.158	0.168
200	0.206	-	0.024	0.016	0.010	0.043	0.033	0.039	0.040	0.021	0.038	0.036	0.048	0.038
300	0.230		-	0.008	0.014	0.019	0.009	0.015	0.016	0.003	0.014	0.060	0.072*	0.062
400	0.222			-	0.006	0.027	0.017	0.023	0.024	0.005	0.022	0.052	0.064	0.054
500	0.216				-	0.033	0.023	0.029	0.030	0.011	0.028	0.046	0.058	0.048
600	0.249					-	0.010	0.004	0.003	0.022	0.005	0.079*	0.091*	0.081*
700	0.239						-	0.006	0.007	0.012	0.005	0.069	0.081*	0.071*
800	0.245							-	0.001	0.018	0.001	0.075*	0.087*	0.077*
900	0.246								-	0.019	0.002	0.076*	0.088*	0.078*
1,000	0.227									-	0.017	0.057	0.069	0.059
1,500	0.244										-	0.074*	0.086*	0.076*
2,000	0.170											-	0.012	0.002
2,500	0.158												-	0.010
3,000	0.168													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 53 พบว่า ค่าการเคาะเฉลี่ยของข้อสอบ 90 ข้อ ของ
กลุ่มตัวอย่าง 300 คน 600 คน 700 คน 800 คน 900 คน และ 1,500 คน มีความ
แตกต่างจากกลุ่มตัวอย่าง 2,500 คน ค่าการเคาะเฉลี่ยของกลุ่มตัวอย่าง 600 คน 800 คน
900 คน และ 1,500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน และค่าการเคาะ
เฉลี่ยของกลุ่มตัวอย่าง 600 คน 700 คน 800 คน 900 คน และ 1,500 คน มีความ
แตกต่างจากกลุ่มตัวอย่าง 3,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.12 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าการเคาะของข้อสอบ
ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวีริสต์ และวิธีของเบส์ เมื่อขนาดของกลุ่ม
ตัวอย่างต่างกัน ดังแสดงในตาราง 54 - 55

ตาราง 54 เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธี
แมกซิมัม ไลค์ลิสต์ วิธีอีวีริสต์ และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	จำนวนคน	df	SS	MS _w	MS _b	F-test
50	200	2	0.363	0.182	0.006	30.146*
	300	2	0.616	0.308	0.004	77.462*
	400	2	0.541	0.271	0.007	40.304*
	500	2	0.560	0.280	0.006	49.929*
	600	2	0.708	0.354	0.004	79.614*
	700	2	0.746	0.373	0.002	188.664*
	800	2	0.221	0.111	0.006	19.549*
	900	2	0.308	0.154	0.005	28.897*
	1,000	2	0.186	0.093	0.005	17.704*
	1,500	2	0.389	0.195	0.006	34.000*
	2,000	2	0.175	0.088	0.006	13.833*
	2,500	2	0.864	0.432	0.006	72.145*
	3,000	2	0.319	0.160	0.008	19.588*
60	200	2	0.354	0.177	0.006	29.052*
	300	2	0.810	0.405	0.004	97.366*
	400	2	0.377	0.188	0.007	27.688*
	500	2	0.737	0.369	0.007	49.580*
	600	2	0.660	0.330	0.004	85.219*
	700	2	0.821	0.410	0.003	157.696*

ตาราง 54 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _w	MS _b	F-test
60	800	2	0.344	0.172	0.003	60.372*
	900	2	0.266	0.133	0.134	9.832*
	1,000	2	0.049	0.024	0.005	4.571*
	1,500	2	0.394	0.197	0.006	32.431*
	2,000	2	0.082	0.041	0.012	3.401*
	2,500	2	0.467	0.234	0.004	63.898*
	3,000	2	0.434	0.217	0.004	57.182*
70	200	2	0.363	0.182	0.007	24.854*
	300	2	0.648	0.324	0.006	56.135*
	400	2	0.524	0.262	0.004	60.983*
	500	2	0.772	0.386	0.007	59.122*
	600	2	0.682	0.341	0.004	76.924*
	700	2	0.744	0.372	0.003	115.350*
	800	2	0.360	0.180	0.004	43.276*
	900	2	0.532	0.266	0.004	64.117*
	1,000	2	0.372	0.186	0.008	24.685*
	1,500	2	0.138	0.069	0.011	6.406*
	2,000	2	0.280	0.140	0.005	30.260*
	2,500	2	0.386	0.193	0.006	33.590*
	3,000	2	0.238	0.119	0.009	13.411*
80	200	2	0.342	0.171	0.007	26.006*
	300	2	0.757	0.379	0.005	79.220*
	400	2	0.578	0.289	0.006	47.719*
	500	2	0.980	0.490	0.006	89.069*

ตาราง 54 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _w	MS _B	F-test
80	600	2	0.796	0.398	0.004	90.512*
	700	2	1.113	0.557	0.002	233.554*
	800	2	0.430	0.215	0.004	58.477*
	900	2	0.551	0.276	0.005	55.914*
	1,000	2	0.383	0.192	0.006	34.053*
	1,500	2	0.566	0.283	0.008	34.706*
	2,000	2	0.234	0.117	0.005	24.436*
	2,500	2	0.712	0.356	0.003	104.483*
	3,000	2	0.339	0.169	0.006	30.370*
90	200	2	0.286	0.143	0.009	15.942*
	300	2	0.964	0.482	0.005	95.485*
	400	2	0.938	0.469	0.006	74.562*
	500	2	1.276	0.638	0.006	100.423*
	600	2	0.889	0.444	0.003	132.088*
	700	2	1.116	0.558	0.003	168.833*
	800	2	0.656	0.328	0.003	110.545*
	900	2	0.706	0.353	0.005	65.285*
	1,000	2	0.454	0.227	0.007	32.799*
	1,500	2	0.400	0.200	0.007	27.157*
	2,000	2	0.298	0.149	0.008	18.063*
	2,500	2	0.198	0.099	0.010	10.247*
	3,000	2	0.064	0.032	0.010	3.116*

ตาราง 54 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _w	MS _b	F-test
100	200	2	0.301	0.151	0.012	12.741*
	300	2	0.769	0.384	0.005	73.984*
	400	2	1.339	0.670	0.006	106.149*
	500	2	1.413	0.707	0.006	124.582*
	600	2	0.918	0.459	0.004	123.467*
	700	2	1.040	0.520	0.003	172.241*
	800	2	0.311	0.156	0.005	30.324*
	900	2	0.381	0.191	0.007	26.949*
	1,000	2	0.035	0.018	0.009	2.020
	1,500	2	0.859	0.430	0.005	84.841*
	2,000	2	0.187	0.093	0.007	13.527*
	2,500	2	0.597	0.299	0.005	61.737*
	3,000	2	0.407	0.204	0.005	42.819*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 54 พบว่า ค่าการเดาของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้น ค่าการเดาของข้อสอบ 100 ข้อของกลุ่มตัวอย่าง 1,000 คน ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ผู้วิจัยจึงได้ทำการทดสอบความแตกต่างของค่าการเดาของข้อสอบที่มีความแตกต่างกัน ด้วยการเปรียบเทียบ พบว่า ได้ผลดังตาราง 55

ตาราง 55 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม
ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส์	อีวริสติก-เบส์
50	200	0.121*	0.060*	0.061*
	300	0.132*	0.007	0.139*
	400	0.135*	0.017	0.118*
	500	0.144*	0.034	0.110*
	600	0.148*	0.004	0.144*
	700	0.144*	0.010	0.154*
	800	0.071*	0.019	0.090*
	900	0.100*	0.008	0.092*
	1,000	0.159*	0.027	0.084*
	1,500	0.090*	0.031	0.121*
	2,000	0.081*	0.060*	0.021
	2,500	0.091*	0.094*	0.185*
	3,000	0.079*	0.031	0.110*
60	200	0.107*	0.039*	0.068*
	300	0.133*	0.018	0.151*
	400	0.107*	0.023	0.084*
	500	0.144*	0.018	0.126*
	600	0.138*	0.020	0.118*
	700	0.124*	0.032*	0.156*
	800	0.078*	0.024	0.102*
	900	0.074*	0.013	0.087*
	1,000	0.008	0.030	0.038*

ตาราง 55 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลียูด-อีวริสติก	แมกซิมัม ไลค์ลียูด-เบส	อีวริสติก-เบส
60	1,500	0.076*	0.037*	0.113*
	2,000	0.036	0.051*	0.015
	2,500	0.061*	0.064*	0.125*
	3,000	0.064*	0.056*	0.120*
70	200	0.098*	0.024	0.074*
	300	0.113*	0.009	0.122*
	400	0.114*	0.018	0.096*
	500	0.131*	0.004	0.127*
	600	0.121*	0.000	0.121*
	700	0.112*	0.025*	0.137*
	800	0.065*	0.034*	0.099*
	900	0.097*	0.017	0.114*
	1,000	0.063*	0.040*	0.103*
	1,500	0.029	0.033	0.062*
	2,000	0.087*	0.064*	0.023
	2,500	0.046*	0.059*	0.105*
	3,000	0.024	0.056*	0.080*
80	200	0.089*	0.021	0.068*
	300	0.112*	0.012	0.124*
	400	0.113*	0.020	0.093*
	500	0.139*	0.007	0.132*
	600	0.113*	0.016	0.129*
	700	0.122*	0.037*	0.159*
	800	0.065*	0.038*	0.103*

ตาราง 55 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไคส์ลิสต์-ฮิวริสติก	แมกซิมัม ไคส์ลิสต์-เบส	ฮิวริสติก-เบส
80	900	0.089*	0.022	0.111*
	1,000	0.065*	0.030*	0.095*
	1,500	0.067*	0.052*	0.119*
	2,000	0.076*	0.031*	0.045*
	2,500	0.081*	0.051*	0.132*
	3,000	0.036*	0.056*	0.092*
90	200	0.078*	0.024	0.054*
	300	0.123*	0.007	0.130*
	400	0.131*	0.012	0.119*
	500	0.157*	0.025	0.132*
	600	0.107*	0.025*	0.132*
	700	0.125*	0.021*	0.146*
	800	0.080*	0.038*	0.118*
	900	0.098*	0.018	0.116*
	1,000	0.079*	0.014	0.093*
	1,500	0.062*	0.031	0.093*
	2,000	0.080*	0.031	0.049*
	2,500	0.066*	0.039*	0.027
	3,000	0.037*	0.023	0.014
100	200	0.078*	0.042*	0.036
	300	0.104*	0.001	0.105*
	400	0.149*	0.016	0.133*
	500	0.154*	0.019	0.135*
	600	0.103*	0.025*	0.128*

ตาราง 55 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซีมัม ไลค์ลิสต์-อีวริสติก	แมกซีมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
100	700	0.112*	0.023*	0.135*
	800	0.040*	0.039*	0.079*
	900	0.069*	0.012	0.081*
	1,000	-	-	-
	1,500	0.092*	0.034*	0.126*
	2,000	0.023	0.038*	0.061*
	2,500	0.063*	0.045*	0.108*
	3,000	0.035*	0.054*	0.089*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 55 พบว่า ค่าการเดาของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ และวิธีอีวริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกันมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้น ค่าการเดาของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 1,000 คน และ 2,000 คน ค่าการเดาของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 1,500 คน และ 3,000 คน และค่าการเดาของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 1,000 คน และ 2,000 คน ที่มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ค่าการเดาของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน 2,000 คน และ 2,500 คน ค่าการเดาของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 200 คน 700 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน ค่าการเดาของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 700 คน 800 คน 1,000 คน 2,000 คน 2,500 คน และ 3,000 คน ค่าการเดาของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 700 คน 800 คน 1,000 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน ค่าการเดาของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 600 คน 700 คน 800 คน และ 2500 คน และค่าการเดาของข้อสอบ 100 ข้อ ของกลุ่ม

ตัวอย่าง 200 คน 600 คน 700 คน 800 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด และวิธีของเบส์ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าการเคาของข้อสอบที่ประมาณค่าด้วยวิธีอีวีเอสติค และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้น ค่าการเคาของข้อสอบ 50 ข้อ 60 ข้อ และ 70 ข้อ ของกลุ่มตัวอย่าง 2,000 คน ค่าการเคาของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 2,500 คน และ 3,000 คน และค่าการเคาของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 1,000 คน ที่มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.13 ผลการเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 56

ตาราง 56 เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _B	MS _W	F-test
50	12	0.000	0.000	1.001	0.000
60	12	0.000	0.000	1.001	0.000
70	12	0.000	0.000	1.001	0.000
80	12	0.000	0.000	1.001	0.000
90	12	0.000	0.000	1.001	0.000
100	12	0.000	0.000	1.001	0.000

จากตาราง 56 พบว่า ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญ

ทางสถิติ

3.14 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอีวีรีสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 57 - 62

ตาราง 57 เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอีวีรีสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _B	MS _w	F-test
50	12	74.618	6.218	0.286	21.723*
60	12	31.594	2.633	0.257	10.249*
70	12	21.477	1.790	0.230	7.785*
80	12	10.551	0.879	0.204	4.310*
90	12	10.458	0.872	0.214	4.082
100	12	12.460	1.038	0.185	5.610*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 57 พบว่า ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอีวีรีสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ของข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ และ 100 ข้อ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติทางสถิติที่ระดับ .05 ผู้วิจัยจึงได้ทำการทดสอบความแตกต่างของค่าความสามารถของผู้สอบด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 58 - 62

ตาราง 58 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
อิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน และข้อสอบมีจำนวน 50 ข้อ ด้วยการ
เปรียบเทียบพหุคูณ

จำนวน คน	อ	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.750	1.081	0.704	1.079	1.125	1.036	0.973	1.048	0.903	1.004	0.970	0.983	0.981
200	0.750	-	0.331*0.046	0.329*0.375*0.286*0.223*0.298*0.153	0.254*0.220*0.233*0.231*									
300	1.081		-	0.377*0.002	0.044	0.045	0.108	0.033	0.178*0.077	0.111	0.098	0.100		
400	0.704			-	0.375*0.421*0.332*0.269*0.344*0.199*0.300*0.268*0.279*0.277*									
500	1.079				-	0.046	0.043	0.106	0.031	0.176*0.075	0.109	0.096	0.098	
600	1.125					-	0.089	0.152*0.077	0.222*0.121*0.155*0.142*0.144*					
700	1.036						-	0.063	0.012	0.133*0.032	0.066	0.053	0.055	
800	0.973							-	0.075	0.070	0.031	0.003	0.010	0.008
900	1.048								-	0.145*0.044	0.078	0.065	0.067	
1,000	0.903									-	0.101*0.067	0.080	0.078	
1,500	1.004										-	0.034	0.021	0.023
2,000	0.970											-	0.013	0.011
2,500	0.983												-	0.002
3,000	0.981													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 58 พบว่า เมื่อข้อสอบมีขนาด 50 ข้อ ค่าความสามารถเฉลี่ยของผู้สอบ 200 คน มีความแตกต่างจากกลุ่มตัวอย่าง 300 คน ค่าความสามารถเฉลี่ยของผู้สอบ 300 คน มีความแตกต่างจากกลุ่มตัวอย่าง 400 คน ค่าความสามารถเฉลี่ยของผู้สอบ 200 คน และ 400 คน มีความแตกต่างจากกลุ่มตัวอย่าง 500 คน 600 คน 700 คน

และ 900 คน ค่าความสามารถเฉลี่ยของผู้สอบ 200 คน 400 คน และ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 800 คน 2,000 คน 2,500 คน และ 3,000 คน ค่าความสามารถเฉลี่ยของผู้สอบ 300 คน 400 คน 600 คน 700 คน และ 900 คน มีความแตกต่างจากกลุ่มตัวอย่าง 600 คน และ 1,000 คน ค่าความสามารถเฉลี่ยของผู้สอบ 200 คน 400 คน 600 คน และ 1,000 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 59 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีวิวิธวิธี
เมื่อขนาดของกลุ่มตัวอย่างต่างกัน และข้อสอบมีจำนวน 60 ข้อ ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	อ	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		1.027	1.090	0.960	1.073	1.100	1.025	0.992	1.020	0.884	0.989	0.960	0.959	0.974
200	1.027	-	0.063	0.067	0.046	0.073	0.002	0.035	0.007	0.143	0.038	0.067	0.068	0.053
300	1.090		-	0.130	0.017	0.010	0.065	0.098	0.070	0.206*	0.101	0.130	0.131	0.116
400	0.960			-	0.113	0.140	0.065	0.032	0.060	0.076	0.029	0.000	0.001	0.014
500	1.073				-	0.027	0.048	0.081	0.053	0.189*	0.084	0.113	0.114*	0.099
600	1.100					-	0.075	0.108	0.080	0.216*	0.111	0.140*	0.141*	0.126*
700	1.025						-	0.033	0.005	0.141*	0.036	0.065	0.066	0.051
800	0.992							-	0.028	0.108	0.003	0.032	0.033	0.018
900	1.020								-	0.136*	0.031	0.060	0.061	0.046
1,000	0.884									-	0.105*	0.076	0.075	0.090*
1,500	0.989										-	0.029	0.030	0.015
2,000	0.960											-	0.001	0.014
2,500	0.959												-	0.015
3,000	0.974													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 59 พบว่า เมื่อข้อสอบมีขนาด 60 ข้อ ค่าความสามารถเฉลี่ยของผู้สอบ 300 คน 500 คน 600 คน 700 คน และ 900 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,000 คน ค่าความสามารถเฉลี่ยของผู้สอบ 500 คน และ 800 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,500 คน ค่าความสามารถเฉลี่ยของผู้สอบ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 2,000 คน ค่าความสามารถเฉลี่ยของผู้สอบ 600 คน และ 1,000 คน มีความแตกต่างจากกลุ่มตัวอย่าง 3,000 คน และค่าความสามารถเฉลี่ยของผู้สอบ 1,000

คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 60 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอิวิริสติด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน และข้อสอบมีจำนวน 70 ข้อ ด้วยการเปรียบเทียบพหุคูณ

จำนวน คน	$\bar{c}$	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.932	0.955	0.866	0.965	0.973	0.919	0.899	0.920	0.822	0.860	0.888	0.864	0.849
200	0.932	-	0.023	0.066	0.033	0.041	0.013	0.033	0.012	0.110	0.072	0.044	0.068	0.083
300	0.955		-	0.089	0.010	0.018	0.036	0.056	0.035	0.133	0.095	0.067	0.091	0.106
400	0.866			-	0.099	0.107	0.053	0.033	0.054	0.044	0.006	0.022	0.002	0.017
500	0.965				-	0.008	0.046	0.066	0.045	0.143*	0.105	0.077	0.101	0.116*
600	0.973					-	0.054	0.074	0.053	0.151*	0.113*	0.085	0.109*	0.124*
700	0.919						-	0.020	0.001	0.097	0.059	0.031	0.055	0.070
800	0.899							-	0.021	0.077	0.039	0.011	0.035	0.050
900	0.920								-	0.098	0.060	0.032	0.056	0.071
1,000	0.822									-	0.038	0.066	0.042	0.027
1,500	0.860										-	0.028	0.004	0.011
2,000	0.888											-	0.024	0.039
2,500	0.864												-	0.015
3,000	0.849													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 60 พบว่า เมื่อข้อสอบมีขนาด 70 ข้อ ค่าความสามารถเฉลี่ยของผู้สอบ 500 คน และ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,000 คน และ 3,000 คน และค่าความสามารถเฉลี่ยของผู้สอบ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,500 คน และ 2,500 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 61 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
อิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกัน และข้อสอบมีจำนวน 80 ข้อ ด้วยการ
เปรียบเทียบพหุคูณ

จำนวน คน	อ	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.926	0.951	0.893	0.965	0.970	0.931	0.904	0.929	0.865	0.899	0.905	0.907	0.875
200	0.926	-	0.025	0.033	0.039	0.044	0.005	0.022	0.003	0.061	0.027	0.021	0.019	0.051
300	0.951		-	0.058	0.014	0.019	0.020	0.047	0.022	0.086	0.052	0.046	0.044	0.076
400	0.893			-	0.072	0.077	0.038	0.011	0.036	0.028	0.006	0.012	0.014	0.018
500	0.965				-	0.005	0.034	0.061	0.036	0.100	0.066	0.060	0.058	0.090
600	0.970					-	0.039	0.066	0.041	0.105	0.071	0.065	0.063	0.095*
700	0.931						-	0.027	0.002	0.066	0.032	0.026	0.024	0.056
800	0.904							-	0.025	0.039	0.005	0.001	0.003	0.029
900	0.929								-	0.064	0.030	0.024	0.022	0.054
1,000	0.865									-	0.034	0.040	0.042	0.010
1,500	0.899										-	0.006	0.008	0.024
2,000	0.905											-	0.002	0.030
2,500	0.907												-	0.032
3,000	0.875													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 61 พบว่า เมื่อข้อสอบมีขนาด 80 ข้อ ค่าความสามารถ
เฉลี่ยของผู้สอบ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 3,000 คน อย่างมีนัยสำคัญทาง
สถิติที่ระดับ .05

ตาราง 62 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
 อีวริสติก เมื่อขนาดของกลุ่มตัวอย่างต่างกันและข้อสอบมีจำนวน 100 ข้อ ด้วยการ
 เปรียบเทียบบุคคล

จำนวน คน	อ	200	300	400	500	600	700	800	900	1,000	1,500	2,000	2,500	3,000
		0.924	0.948	0.898	0.978	0.959	0.906	0.883	0.912	0.853	0.911	0.867	0.889	0.873
200	0.924	-	0.024	0.026	0.054	0.035	0.018	0.041	0.012	0.071	0.013	0.057	0.035	0.051
300	0.948		-	0.050	0.030	0.011	0.042	0.065	0.036	0.095	0.037	0.081	0.059	0.075
400	0.898			-	0.080	0.061	0.008	0.015	0.014	0.045	0.013	0.031	0.009	0.025
500	0.978				-	0.019	0.072	0.095	0.066	0.125*	0.067	0.111*	0.089	0.105*
600	0.959					-	0.053	0.076	0.047	0.106*	0.048	0.092*	0.070	0.086
700	0.906						-	0.023	0.006	0.053	0.005	0.039	0.017	0.033
800	0.883							-	0.029	0.030	0.028	0.016	0.006	0.010
900	0.912								-	0.059	0.001	0.045	0.023	0.039
1,000	0.853									-	0.058	0.014	0.036	0.020
1,500	0.911										-	0.044	0.022	0.038
2,000	0.867											-	0.022	0.006
2,500	0.889												-	0.016
3,000	0.873													-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 62 พบว่า เมื่อข้อสอบมีขนาด 100 ข้อ ค่าความสามารถ
 เฉลี่ยของผู้สอบ 500 คน มีความแตกต่างจากกลุ่มตัวอย่าง 3,000 คน และค่าความสามารถ
 เฉลี่ยของผู้สอบ 500 คน และ 600 คน มีความแตกต่างจากกลุ่มตัวอย่าง 1,000 คน และ
 2,000 คน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.15 ผลการเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่
ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 63

ตาราง 63 เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
ของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	df	SS	MS _B	MS _w	F-test
50	12	0.002	0.000	1.000	0.000
60	12	0.002	0.000	1.001	0.000
70	12	0.002	0.000	1.000	0.000
80	12	0.002	0.000	1.001	0.000
90	12	0.002	0.000	1.001	0.000
100	12	0.002	0.000	1.001	0.000

จากตาราง 63 พบว่า ค่าความสามารถของผู้สอบที่ประมาณค่าด้วย
วิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทาง
สถิติ

3.16 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าความสามารถของ
ผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของ
กลุ่มตัวอย่างต่างกัน ดังแสดงในตาราง 64 - 65

ตาราง 64 เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
แมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน

จำนวนข้อ	จำนวนคน	df	SS	MS _b	MS _w	F-test
50	200	2	74.939	37.470	1.226	30.571*
	300	2	234.127	117.063	1.240	94.443*
	400	2	132.058	66.029	1.301	50.748*
	500	2	388.332	194.166	1.198	162.073*
	600	2	505.748	252.874	1.185	213.330*
	700	2	500.097	250.049	1.232	202.953*
	800	2	505.558	252.779	1.093	231.383*
	900	2	658.492	329.246	0.950	346.521*
	1,000	2	544.358	272.179	1.057	257.528*
	1,500	2	1,007.269	503.634	1.114	452.036*
	2,000	2	1,254.167	627.084	1.003	625.462*
	2,500	2	1,310.000	805.000	1.038	775.902*
	3,000	2	1,923.104	961.552	0.944	1,018.462*
60	200	2	140.509	70.254	1.221	57.541*
	300	2	237.699	118.850	1.388	85.653*
	400	2	245.968	122.984	1.268	96.957*
	500	2	384.068	192.034	1.195	160.680*
	600	2	483.942	241.971	1.192	203.079*
	700	2	490.131	245.066	1.102	222.476*

ตาราง 64 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
60	800	2	525.070	262.535	1.103	238.074*
	900	2	623.473	311.736	1.017	306.488*
	1,000	2	520.412	260.206	1.054	246.943*
	1,500	2	978.095	489.048	1.064	459.631*
	2,000	2	1,227.925	613.963	0.874	702.734*
	2,500	2	1,531.521	765.761	1.028	745.049*
	3,000	2	1,895.957	947.979	0.932	1,017.447*
70	200	2	115.844	57.922	1.261	45.917*
	300	2	182.068	91.033	1.220	74.647*
	400	2	199.495	99.748	1.270	78.570*
	500	2	310.075	155.037	1.205	128.689*
	600	2	378.637	189.319	1.166	162.433*
	700	2	394.144	197.072	1.083	185.376*
	800	2	431.127	215.564	0.966	223.201*
	900	2	507.112	253.556	0.931	272.388*
	1,000	2	450.670	225.335	1.101	204.723*
	1,500	2	739.308	369.654	0.958	385.820*
	2,000	2	1,051.559	525.780	0.903	582.521*
	2,500	2	1,244.855	622.427	1.040	598.614*
	3,000	2	1,440.796	720.398	0.919	783.854*
80	200	2	114.485	57.243	1.258	45.508*
	300	2	182.299	91.149	1.253	72.756*
	400	2	212.932	106.466	1.287	82.734*
	500	2	310.612	155.306	1.232	126.091*

ตาราง 64 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
80	600	2	378.487	188.244	1.221	154.139*
	700	2	404.241	202.120	1.014	199.331*
	800	2	435.621	217.810	1.007	218.265*
	900	2	517.194	258.597	0.965	267.865*
	1,000	2	499.119	249.560	0.907	275.060*
	1,500	2	808.053	404.027	1.072	376.811*
	2,000	2	1,091.974	545.987	0.827	659.875*
	2,500	2	1,370.415	685.208	1.073	638.376*
	3,000	2	1,531.149	765.575	0.958	799.430*
90	200	2	108.109	54.055	1.198	45.119*
	300	2	168.265	84.133	1.280	65.754*
	400	2	206.391	103.196	1.295	79.695*
	500	2	295.191	147.596	1.171	126.032*
	600	2	352.997	176.499	1.160	152.213*
	700	2	378.588	189.294	0.957	197.892*
	800	2	424.154	212.077	0.940	225.705*
	900	2	497.147	248.570	0.930	267.290*
	1,000	2	465.072	232.536	0.907	256.381*
	1,500	2	744.909	372.455	1.029	362.075*
	2,000	2	1,005.610	502.805	0.915	549.337*
	2,500	2	1,252.951	626.476	1.025	611.425*
	3,000	2	1,458.113	729.057	0.924	788.302*

ตาราง 64 (ต่อ)

จำนวนข้อ	จำนวนคน	df	SS	MS _B	MS _w	F-test
100	200	2	113.856	56.928	1.157	49.204*
	300	2	179.669	89.834	1.148	78.577*
	400	2	214.761	107.380	1.202	89.351*
	500	2	318.438	159.219	1.215	131.060*
	600	2	337.856	183.928	1.141	161.268*
	700	2	382.426	191.213	1.032	185.288*
	800	2	415.198	207.599	0.968	214.385*
	900	2	499.208	249.604	0.966	258.464*
	1,000	2	485.148	242.574	0.971	249.897*
	1,500	2	829.672	414.836	1.132	366.547*
	2,000	2	1,001.548	500.774	1.105	458.385*
	2,500	2	1,317.144	685.572	0.967	681.151*
	3,000	2	1,524.251	762.126	1.033	737.592*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 64 พบว่า ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผู้วิจัยจึงได้ทำการทดสอบความแตกต่างของค่าความสามารถของผู้สอบด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 65

ตาราง 65 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
แมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน
ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส์	อีวริสติก-เบส์
50	200	0.750*	0.001	0.749*
	300	0.958*	0.001	1.083*
	400	0.704*	0.001	0.708*
	500	1.079*	0.000	1.079*
	600	1.125*	0.000	1.125*
	700	1.036*	0.000	1.036*
	800	0.973*	0.001	0.974*
	900	1.048*	0.001	1.049*
	1,000	0.903*	0.001	0.904*
	1,500	1.004*	0.000	1.004*
	2,000	0.970*	0.000	0.970*
	2,500	0.983*	0.000	0.983*
	3,000	0.981*	0.000	0.981*
60	200	1.207*	0.001	1.206*
	300	1.090*	0.000	1.090*
	400	0.960*	0.000	0.960*
	500	1.073*	0.000	1.073*
	600	1.100*	0.001	1.099*
	700	1.205*	0.001	1.204*
	800	0.992*	0.001	0.993*
	900	1.020*	0.000	1.020*
	1,000	0.884*	0.001	0.883*

ตาราง 65 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
60	1,500	0.989*	0.000	0.989*
	2,000	0.960*	0.000	0.960*
	2,500	0.959*	0.000	0.959*
	3,000	0.974*	0.000	0.974*
70	200	0.932*	0.000	0.932*
	300	0.955*	0.001	0.954*
	400	0.867*	0.001	0.866*
	500	0.965*	0.001	0.964*
	600	0.973*	0.000	0.973*
	700	0.919*	0.000	0.919*
	800	0.899*	0.001	0.900*
	900	0.920*	0.000	0.920*
	1,000	0.822*	0.000	0.822*
	1,500	0.880*	0.000	0.860*
	2,000	0.888*	0.000	0.888*
	2,500	0.864*	0.000	0.864*
80	3,000	0.849*	0.000	0.849*
	200	0.926*	0.001	0.927*
	300	0.958*	0.001	0.957*
	400	0.893*	0.001	0.894*
	500	0.965*	0.001	0.966*
	600	0.970*	0.001	0.971*
	700	0.931*	0.000	0.931*
	800	0.904*	0.000	0.904*

ตาราง 65 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไคส์ลิสต์-ฮิวริสติก	แมกซิมัม ไคส์ลิสต์-เบส	ฮิวริสติก-เบส
80	900	0.929*	0.000	0.929*
	1,000	0.865*	0.000	0.865*
	1,500	0.899*	0.000	0.899*
	2,000	0.905*	0.000	0.905*
	2,500	0.907*	0.000	0.907*
	3,000	0.875*	0.000	0.875*
90	200	0.901*	0.000	0.901*
	300	0.917*	0.000	0.917*
	400	0.879*	0.002	0.881*
	500	0.941*	0.000	0.941*
	600	0.940*	0.000	0.940*
	700	0.901*	0.001	0.901*
	800	0.892*	0.000	0.892*
	900	0.911*	0.001	0.911*
	1,000	0.836*	0.000	0.836*
	1,500	0.863*	0.000	0.863*
	2,000	0.869*	0.000	0.869*
	2,500	0.867*	0.000	0.867*
	3,000	0.854*	0.000	0.854*
100	200	0.924*	0.001	0.925*
	300	0.948*	0.000	0.948*
	400	0.898*	0.001	0.897*
	500	0.978*	0.000	0.978*
	600	0.959*	0.000	0.959*

ตาราง 65 (ต่อ)

จำนวนข้อ	จำนวนคน	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
100	700	0.906*	0.000	0.905*
	800	0.883*	0.001	0.882*
	900	0.912*	0.000	0.913*
	1,000	0.853*	0.000	0.853*
	1,500	0.911*	0.000	0.911*
	2,000	0.867*	0.000	0.867*
	2,500	0.889*	0.000	0.889*
	3,000	0.873*	0.000	0.873*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 65 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ กับวิธีอีวริสติก และวิธีอีวริสติก กับวิธีของเบส เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับค่าความสามารถเฉลี่ยของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ กับวิธีของเบส เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีความนัยสำคัญทางสถิติ

4. ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าพารามิเตอร์ของข้อสอบ
และความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของ
เบส์ เมื่อจำนวนข้อสอบต่างกัน

4.1 ผลการเปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่
ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 66

ตาราง 66 เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกันใน
ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด

จำนวนคน	df	SS	MS _B	MS _w	F-test
200	5	0.706	0.021	1.046	0.020
300	5	0.125	0.025	1.097	0.023
400	5	0.774	0.155	1.030	0.150
500	5	0.179	0.036	1.127	0.032
600	5	0.514	0.103	1.158	0.089
700	5	0.205	0.041	1.202	0.034
800	5	0.430	0.086	1.137	0.076
900	5	1.120	0.224	1.296	0.173
1,000	5	0.380	0.076	1.034	0.073
1,500	5	0.225	0.045	1.227	0.037
2,000	5	0.234	0.047	1.256	0.037
2,500	5	0.376	0.075	1.145	0.066
3,000	5	0.066	0.013	1.309	0.010

จากตาราง 66 พบว่า ค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกันใน
ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.2 ผลการเปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่
ประมาณค่าด้วยวิธีอิวิวิสถิติ เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 67

ตาราง 67 เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกันที่
ประมาณค่าด้วยวิธีอิวิวิสถิติ

จำนวนคน	df	SS	MS _B	MS _w	F-test
200	5	15.665	3.133	4.220	0.742
300	5	9.906	1.981	3.122	0.635
400	5	2.721	0.544	2.342	0.232
500	5	12.342	2.469	1.001	2.466
600	5	5.121	1.024	1.450	0.706
700	5	10.655	2.131	2.482	0.866
800	5	2.003	0.401	1.747	0.229
900	5	11.987	2.398	1.887	1.271
1,000	5	4.088	0.818	1.263	0.647
1,500	5	16.752	3.350	1.712	1.958
2,000	5	8.763	1.753	1.397	1.254
2,500	5	1.413	0.283	1.134	0.249
3,000	5	0.858	0.172	1.276	0.135

จากตาราง 67 พบว่า ค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกัน
ที่ประมาณค่าด้วยวิธีอิวิวิสถิติ มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.3 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 68 - 69

ตาราง 68 เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกันที่ประมาณค่าด้วยวิธีของเบส์

จำนวนคน	df	SS	MS _B	MS _w	F-test
200	5	3.772	0.754	0.838	0.900
300	5	0.857	0.171	0.187	0.917
400	5	10.211	2.042	0.980	2.084
500	5	2.523	0.505	0.686	0.735
600	5	0.713	0.143	0.322	0.443
700	5	0.521	0.104	0.054	1.930
800	5	0.116	0.023	0.054	0.432
900	5	0.121	0.024	0.097	0.250
1,000	5	3.915	0.783	0.288	2.717
1,500	5	0.930	0.186	0.310	0.600
2,000	5	23.706	4.741	0.236	20.062*
2,500	5	5.448	1.090	0.337	3.234
3,000	5	2.076	0.415	0.209	1.985

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 68 พบว่า ค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกันที่ประมาณค่าด้วยวิธีของเบส์ มีความแตกต่างกันอย่างไรไม่มีนัยสำคัญทางสถิติ ยกเว้น ค่าอำนาจจำแนกของข้อสอบของกลุ่มตัวอย่าง 2,500 คน ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังนั้น ผู้วิจัยจึงได้ทำการทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ

ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 69

ตาราง 69 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกันของ
กลุ่มตัวอย่างขนาด 2,000 คน ที่วิเคราะห์ด้วยวิธีของเบส์ ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{g}$	50	60	70	80	90	100
		0.240	0.238	0.229	0.228	0.686	0.708
50	0.240	-	0.002	0.011	0.012	0.446*	0.468*
60	0.238		-	0.001	0.010	0.448*	0.470*
70	0.229			-	0.001	0.457*	0.479*
80	0.228				-	0.458*	0.480*
90	0.686					-	0.022
100	0.708						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 69 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ และ 80 ข้อ มีความแตกต่างกับค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 90 ข้อ และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.4 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าอำนาจจำแนกของ
ข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวน
ข้อสอบต่างกัน ดังแสดงในตาราง 70 - 71

ตาราง 70 เปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกันที่
ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
200	50	2	14.110	7.055	1.374	5.133*
	60	2	14.579	7.290	1.709	4.265*
	70	2	16.752	8.376	1.774	4.722*
	80	2	7.874	3.937	2.262	1.741
	90	2	43.876	21.938	2.305	9.519*
	100	2	29.908	14.954	2.528	5.915*
300	50	2	18.470	9.235	1.478	6.248*
	60	2	21.684	10.842	1.322	8.204*
	70	2	21.057	10.529	1.192	8.830*
	80	2	14.609	7.305	1.353	5.400*
	90	2	29.341	14.671	1.268	11.572*
	100	2	44.063	22.031	2.119	10.399*
400	50	2	0.244	0.122	2.383	0.051
	60	2	6.269	3.135	1.252	2.503
	70	2	4.412	2.206	0.522	4.224*
	80	2	1.686	0.843	1.081	0.780
	90	2	10.934	5.467	0.698	7.837*
	100	2	13.510	6.755	2.913	2.319

ตาราง 70 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
500	50	2	1.468	0.734	0.746	0.984
	60	2	5.658	2.829	0.620	4.567*
	70	2	3.830	1.915	0.795	2.409
	80	2	6.605	3.302	0.994	3.323*
	90	2	14.761	7.381	0.874	8.449*
	100	2	22.514	11.257	1.417	7.943*
600	50	2	9.158	4.579	0.908	5.046*
	60	2	6.423	3.212	0.836	3.843*
	70	2	3.098	1.549	0.770	2.013
	80	2	4.256	2.128	0.972	2.189
	90	2	3.630	1.815	0.799	2.271
	100	2	2.905	1.453	1.488	0.976
700	50	2	4.312	2.156	0.816	2.643
	60	2	7.226	3.613	0.843	4.284*
	70	2	6.626	3.313	1.049	3.160
	80	2	7.333	3.667	0.673	5.451*
	90	2	8.929	4.464	0.658	6.783*
	100	2	26.568	13.284	2.869	4.631*
800	50	2	3.118	1.559	1.280	1.218
	60	2	3.450	1.725	0.668	2.583
	70	2	4.165	2.082	0.594	3.505*
	80	2	2.552	1.276	0.753	1.694
	90	2	7.805	3.902	0.680	5.737*
	100	2	3.845	1.922	1.794	1.071

ตาราง 70 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
900	50	2	2.169	1.084	0.735	1.476
	60	2	3.657	1.828	0.820	2.230
	70	2	4.435	2.218	0.728	3.046
	80	2	5.899	2.950	0.695	4.246*
	90	2	6.343	3.171	0.742	4.276*
	100	2	4.988	2.494	1.966	1.268
1,000	50	2	0.309	0.155	0.824	0.188
	60	2	0.992	0.496	0.828	0.598
	70	2	7.179	3.590	0.765	4.694*
	80	2	2.116	1.058	0.797	1.328
	90	2	5.203	2.601	0.658	3.952*
	100	2	10.592	5.296	1.288	4.112*
1,500	50	2	2.384	1.192	0.822	1.450
	60	2	2.681	1.340	0.759	1.766
	70	2	0.747	0.373	0.786	0.475
	80	2	2.125	1.063	0.648	1.641
	90	2	3.170	1.585	0.669	2.369
	100	2	4.736	2.368	2.437	0.972
2,000	50	2	3.224	1.162	0.694	2.322
	60	2	2.447	1.224	0.785	1.558
	70	2	3.906	1.953	0.645	3.026
	80	2	5.368	2.684	0.610	4.400*
	90	2	7.178	3.589	0.781	4.594*
	100	2	13.745	6.872	1.925	3.570*

ตาราง 70 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
2,500	50	2	2.126	1.063	0.954	1.115
	60	2	0.639	0.320	0.602	0.530
	70	2	1.219	0.610	0.658	0.926
	80	2	2.014	1.007	0.717	1.406
	90	2	4.310	2.155	0.762	2.829
	100	2	0.278	0.139	1.434	0.097
3,000	50	2	0.707	0.354	1.001	0.354
	60	2	0.680	0.340	0.866	0.393
	70	2	0.901	0.451	0.809	0.557
	80	2	1.306	0.653	0.676	0.967
	90	2	2.443	1.222	0.783	1.561
	100	2	3.243	1.622	1.431	1.133

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 70 พบว่า ค่าอำนาจจำแนกของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน และ 600 คน ค่าอำนาจจำแนกของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 600 คน และ 700 คน ค่าอำนาจจำแนกของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 800 คน และ 1,000 คน ค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 300 คน 500 คน 700 คน และ 900 คน ค่าอำนาจจำแนกของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 500 คน 700 คน 800 คน 900 คน 1,000 คน และ 2,000 คน และค่าอำนาจจำแนกของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 500 คน 700 คน 1,000 คน และ 2,000 คน ที่ประมาณค่าด้วยวิธีแมกซีมัม โคลลิชูด วิธีอีวริสติก และวิธีของเบส์ มีความแตกต่าง

กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผู้วิจัยจึงได้ทำการทดสอบความแตกต่างของค่าอำนาจ
จำแนกของข้อสอบ ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังแสดงในตาราง 71

ตาราง 71 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่มีจำนวนต่างกันที่
ประมาณค่าด้วยวิธีแมกซีмум โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ ด้วยการเปรียบเทียบ
พหุคูณ

จำนวนคน	จำนวนข้อ	แมกซีмум โลคัลลิสต์-อีวริสติก	แมกซีмум โลคัลลิสต์-เบส์	อีวริสติก-เบส์
200	50	0.026	0.640*	0.666*
	60	0.206	0.474	0.680*
	70	0.038	0.622*	0.584*
	80	-	-	-
	90	0.287	0.676*	0.963*
	100	0.358	0.415	0.773*
300	50	0.379	0.478	0.857*
	60	0.429	0.421	0.850*
	70	0.227	0.529*	0.756*
	80	0.074	0.484*	0.558*
	90	0.174	0.597*	0.771*
	100	0.493	0.445	0.938*
400	50	-	-	-
	60	-	-	-
	70	0.236	0.358*	0.122
	80	-	-	-
	90	0.078	0.466*	0.388*
	100	-	-	-

ตาราง 71 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไคส์ลิสต์-ฮิวริสติก	แมกซิมัม ไคส์ลิสต์-เบส	ฮิวริสติก-เบส
500	50	-	-	-
	60	0.069	0.408*	0.339
	70	-	-	-
	80	0.217	0.411*	0.194
	90	0.082	0.536*	0.454*
	100	0.288	0.381	0.669*
600	50	0.270	0.334	0.604*
	60	0.084	0.353	0.437*
	70	-	-	-
	80	-	-	-
	90	-	-	-
	100	-	-	-
700	50	-	-	-
	60	0.068	0.457*	0.389
	70	-	-	-
	80	0.031	0.388*	0.357*
	90	0.148	0.441*	0.293*
	100	0.306	0.420	0.726*
800	50	-	-	-
	60	-	-	-
	70	0.217	0.347*	0.130
	80	-	-	-
	90	0.343*	0.384*	0.041
	100	-	-	-

ตาราง 71 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไลค์ลิสต์-ฮิวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	ฮิวริสติก-เบส
900	50	-	-	-
	60	-	-	-
	70	-	-	-
	80	0.301	0.359*	0.058
	90	0.270	0.388*	0.098
	100	-	-	-
1,000	50	-	-	-
	60	-	-	-
	70	0.031	0.411*	0.380*
	80	-	-	-
	90	0.192	0.346*	0.154
	100	0.080	0.354	0.434*
1,500	50	-	-	-
	60	-	-	-
	70	-	-	-
	80	-	-	-
	90	-	-	-
	100	-	-	-
2,000	50	-	-	-
	60	-	-	-
	70	-	-	-
	80	0.265	0.087	0.352*
	90	0.245	0.401*	0.156
	100	0.103	0.395	0.498*

ตาราง 71 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
2,500	50	-	-	-
	60	-	-	-
	70	-	-	-
	80	-	-	-
	90	-	-	-
3,000	100	-	-	-
	50	-	-	-
	60	-	-	-
	70	-	-	-
	80	-	-	-
	90	-	-	-
	100	-	-	-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 71 พบว่า ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 800 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ กับวิธีอีวริสติก มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 500 คน และ 700 คน ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 800 คน และ 1,000 คน ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 300 คน 500 คน 700 คน และ 900 คน และค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 700 คน 800 คน 900 คน 1,000 คน และ 2,000

คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ กับวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าอำนาจจำแนกเฉลี่ยของข้อสอบที่มีจำนวนต่างกันที่ประมาณค่าด้วยวิธีอิวิริสติก กับวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้นค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 500 คน และ 700 คน ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 400 คน และ 800 คน ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 500 คน และ 900 คน และค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 800 คน 900 คน และ 2,000 คน ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.5 ผลการเปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 72

ตาราง 72 เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _B	MS _w	F-test
200	5	1.932	0.387	4.990	0.077
300	5	0.885	0.177	4.941	0.036
400	5	5.401	1.080	5.068	0.213
500	5	1.575	0.315	5.729	0.055
600	5	7.143	1.429	6.256	0.228
700	5	7.143	1.429	6.256	0.228
800	5	4.504	0.901	5.790	0.156
900	5	3.451	0.690	7.528	0.092
1,000	5	0.667	0.133	4.716	0.028
1,500	5	4.804	0.961	5.982	0.161
2,000	5	8.300	1.660	6.928	0.240
2,500	5	4.886	0.977	7.499	0.130
3,000	5	4.830	0.966	6.782	0.142

จากตาราง 72 พบว่า ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.6 ผลการเปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 73

ตาราง 73 เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _B	MS _w	F-test
200	5	6,145.821	1,229.164	007.988	1.521
300	5	11.596	2.319	7.581	0.306
400	5	88.806	17.761	23.378	0.760
500	5	11.476	2.295	4.866	0.472
600	5	21.098	4.220	6.640	0.636
700	5	21.098	4.220	6.640	0.636
800	5	13.785	2.757	3.126	0.882
900	5	198.537	39.707	61.738	0.643
1,000	5	195.555	39.111	24.172	1.618
1,500	5	7.287	1.457	3.654	0.399
2,000	5	5.595	1.119	5.447	0.206
2,500	5	107.652	21.530	18.185	1.184
3,000	5	98.248	19.650	26.757	0.734

จากตาราง 73 พบว่า ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.7 ผลการเปรียบเทียบความแตกต่างของค่าความฮากของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 74

ตาราง 74 เปรียบเทียบความแตกต่างของค่าความฮากของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _b	MS _w	F-test
200	5	25.463	5.093	4.447	1.145
300	5	4.372	0.874	1.986	0.440
400	5	25.866	5.173	3.576	1.447
500	5	1.222	0.244	4.406	0.056
600	5	7.934	1.587	2.319	0.684
700	5	7.934	1.587	2.319	0.684
800	5	1.130	0.226	0.320	0.707
900	5	2.030	0.406	0.915	0.444
1,000	5	21.181	4.236	4.467	0.948
1,500	5	18.443	3.689	3.682	1.002
2,000	5	15.609	3.122	5.807	0.538
2,500	5	2.734	0.547	2.230	0.245
3,000	5	30.082	6.016	3.257	1.847

จากตาราง 74 พบว่า ค่าความฮากของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.8 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าความยากของ
ข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวน
ข้อสอบต่างกัน ดังแสดงในตาราง 75 - 76

ตาราง 75 เปรียบเทียบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธี
แมกซีมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	df	SS	MS _b	MS _w	F-test
200	50	2	300.581	150.291	30.582	4.914*
	60	2	4,418.855	2,209.427	1,702.357	1.298
	70	2	87.483	43.742	103.940	0.421
	80	2	86.000	43.000	19.464	2.209
	90	2	30.553	15.277	127.209	0.120
	100	2	125.763	62.882	12.534	5.017*
300	50	2	74.206	37.103	4.977	7.455*
	60	2	110.476	55.238	3.973	13.903*
	70	2	179.313	89.657	6.997	12.814*
	80	2	124.227	62.114	4.377	14.191*
	90	2	153.287	76.643	4.836	15.847*
	100	2	169.566	84.783	4.475	18.947*
400	50	2	58.973	29.487	4.812	6.128*
	60	2	85.235	42.618	10.565	4.034*
	70	2	20.897	10.448	19.989	0.523
	80	2	167.889	83.944	6.191	13.559*
	90	2	271.823	135.912	15.574	8.727*
	100	2	197.742	98.871	7.704	12.834*

ตาราง 75 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
500	50	2	66.820	33.410	5.577	5.991*
	60	2	84.488	42.244	5.934	7.119*
	70	2	92.369	46.185	4.245	10.880*
	80	2	116.640	58.320	5.520	10.565*
	90	2	164.179	82.090	4.758	17.253*
	100	2	237.048	118.524	4.918	24.125*
600	50	2	71.863	35.931	3.004	11.960*
	60	2	93.413	46.707	3.935	11.869*
	70	2	139.287	69.644	5.357	13.001*
	80	2	244.919	122.459	8.274	14.800*
	90	2	148.350	74.175	4.863	15.253*
	100	2	213.209	106.605	4.640	22.975*
700	50	2	33.886	16.943	3.234	5.240*
	60	2	51.529	25.764	3.995	6.449*
	70	2	133.176	66.588	3.621	18.888*
	80	2	60.074	30.037	10.759	2.792
	90	2	103.744	51.872	3.761	13.792*
	100	2	137.143	68.572	3.549	19.320*
800	50	2	28.557	14.279	5.855	2.439
	60	2	101.425	50.712	3.466	14.632*
	70	2	109.265	54.632	2.954	18.497*
	80	2	146.574	73.287	2.230	32.868*
	90	2	144.692	72.346	8.113	23.242*
	100	2	166.318	83.159	2.387	34.845*

ตาราง 75 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
900	50	2	113.389	56.695	3.723	15.228*
	60	2	116.452	58.226	4.211	13.828*
	70	2	140.824	70.412	3.984	17.673*
	80	2	233.437	116.718	8.739	13.356*
	90	2	675.006	337.503	88.529	3.812*
	100	2	150.233	75.116	13.305	5.646*
1,000	50	2	14.312	7.156	32.996	0.217
	60	2	162.451	81.226	7.040	11.537*
	70	2	127.516	63.758	3.656	17.439*
	80	2	153.018	76.509	4.341	17.625*
	90	2	176.842	88.421	4.303	20.551*
	100	2	364.240	182.120	1.417	8.787*
1,500	50	2	32.898	16.449	6.334	2.593
	60	2	56.811	28.406	4.864	5.840*
	70	2	120.873	60.436	5.036	12.001*
	80	2	92.022	46.022	4.373	10.524*
	90	2	94.110	47.055	3.888	12.103*
	100	2	119.924	59.962	3.832	15.646*
2,000	50	2	58.705	29.352	10.654	2.755
	60	2	65.400	32.700	4.410	7.416*
	70	2	75.085	37.543	5.917	6.345*
	80	2	116.664	58.332	4.146	14.068*
	90	2	143.012	71.506	7.062	10.125*
	100	2	179.753	89.877	5.872	15.307*

ตาราง 75 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
2,500	50	2	120.079	60.039	11.188	5.367*
	60	2	17.499	8.749	32.907	0.266
	70	2	59.212	29.606	5.776	5.126*
	80	2	109.876	54.938	4.563	12.041*
	90	2	139.822	69.911	5.972	11.707*
	100	2	167.602	83.801	4.420	18.959*
3,000	50	2	39.614	19.807	6.471	3.061
	60	2	62.804	31.402	4.480	7.009*
	70	2	88.251	44.126	4.435	9.949*
	80	2	144.539	72.269	5.107	14.152*
	90	2	316.860	158.430	37.877	4.183*
	100	2	75.764	37.882	8.945	4.235*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 75 พบว่า ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม โคลลิฮูด วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้น ค่าความยากของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 800 คน 1,000 คน 1,500 คน 2,000 คน และ 3,000 คน ค่าความยากของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 2,500 คน ค่าความยากของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 400 คน ค่าความยากของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 700 คน และค่าความยากของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 200 คน ที่มีความแตกต่างกันอย่างไม่เป็นนัยสำคัญทางสถิติ ซึ่งผู้วิจัยได้ทำการทดสอบความแตกต่างของค่าความยากของข้อสอบ ด้วยการเปรียบเทียบพหุคูณ ผลการทดสอบดังตาราง 76

ตาราง 76 การทดสอบความแตกต่างของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม
โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน ด้วยการเปรียบเทียบพหุคูณ

จำนวนคน	จำนวนข้อ	แมกซีมัม โลคัลลิสต์-อีวริสติก	แมกซีมัม โลคัลลิสต์-เบส	อีวริสติก-เบส
200	50	3.267*	0.557	2.710
	60	-	-	-
	70	-	-	-
	80	-	-	-
	90	-	-	-
	100	1.560*	1.204	0.356
300	50	1.717*	0.594	1.123*
	60	1.941*	0.900	1.041
	70	2.240*	0.715	1.525*
	80	1.781*	0.808	0.973*
	90	1.831*	0.619	1.212*
	100	1.861*	0.909*	0.952*
400	50	1.578*	0.802	0.776
	60	1.751*	0.850	0.865
	70	-	-	-
	80	2.039*	0.648	1.391*
	90	2.307	0.350	1.957*
	00	2.013*	0.789	1.224*
500	50	1.619*	0.495	1.124
	60	1.622*	0.394	1.228*
	70	1.590*	0.444	1.146*
	80	1.644*	0.381	1.263*
	90	1.888*	0.579	1.309*

ตาราง 76 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไลค์ลิสต์-อีวาสิติค	แมกซิมัม ไลค์ลิสต์-เบส	อีวาสิติค-เบส
500	100	2.146*	0.646	1.500*
600	50	1.696*	0.584	1.112*
	60	1.747*	0.570	1.177*
	70	1.992*	0.715	1.277*
	80	2.392*	0.566	1.828*
	90	1.807*	0.594	1.213*
	100	2.089*	0.903*	1.186*
700	50	1.156*	0.400	0.756
	60	1.322*	0.639	0.683
	70	1.882*	0.454*	1.428*
	80	-	-	-
	90	1.523*	0.610	0.913*
	100	1.651*	0.607	0.994*
800	50	0.736*	0.619	1.201*
	60	2.932*	0.497	1.237*
	70	1.927*	0.739*	1.188*
	80	1.754*	0.513	1.242*
	90	1.833*	0.764*	1.069*
	100	2.184*	1.029*	1.155*
900	50	1.955*	0.658	1.297*
	60	1.980*	0.611	1.369*
	70	2.323*	0.555	1.634*
	80	3.605*	0.525	2.946
	90	1.741*	0.727	0.974

ตาราง 76 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
900	100	-	-	-
1,000	50	-	-	-
	60	0.543*	0.435	0.978*
	70	0.374*	0.907*	1.041*
	80	1.854*	0.340	0.814*
	90	1.990*	0.747*	1.243*
	100	2.504*	0.335	2.169*
1,500	50	1.390*	0.570	0.343
	60	1.874*	0.664	1.210*
	70	1.547*	0.846*	0.701
	80	1.393*	0.323	0.598*
	90	1.571*	0.732*	0.839*
	100	1.489*	0.733	0.756
2,000	50	1.471*	0.900	0.571
	60	1.727*	0.981*	0.746*
	70	1.779*	1.161*	0.618
	80	1.892*	0.705	1.187*
	90	2.100*	0.449	1.652
	100	1.892*	0.705	1.187*
2,500	50	2.100*	0.449	1.652
	60	-	-	-
	70	1.285*	0.404	0.881
	80	1.670*	0.668*	1.002*
	90	1.728*	0.493	1.235*

ตาราง 76 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส์	อีวริสติก-เบส์
2,500	100	1.823*	0.621	1.202*
3,000	50	-	-	-
	60	1.457*	0.616	0.841
	70	1.561*	0.445*	1.116*
	80	1.795*	0.314*	0.683*
	90	2.606*	0.760	1.846
	100	1.247*	0.642	0.605*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 76 พบว่า ค่าความยากเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ กับวิธีอีวริสติก เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้น ค่าความยากเฉลี่ยของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 400 คน ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ค่าความยากเฉลี่ยของข้อสอบ 60 ข้อ ของกลุ่มตัวอย่าง 2,000 คน ค่าความยากเฉลี่ยของข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 700 คน 2,000 คน และ 3,000 คน ค่าความยากเฉลี่ยของข้อสอบ 70 ข้อ และ 90 ข้อ ของกลุ่มตัวอย่าง 800 คน 1,000 คน และ 1,500 คน ค่าความยากเฉลี่ยของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 2,500 คน และ 3,000 คน และค่าความยากเฉลี่ยของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 300 คน 600 คน และ 800 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ กับวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ .05

ค่าความยากเฉลี่ยของข้อสอบของกลุ่มตัวอย่างที่มีจำนวนต่างกันที่ประมาณค่าด้วยวิธีอีวริสติก กับวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้น ค่าความยากเฉลี่ยของข้อสอบ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน 400 คน 500 คน

700 คน 1,500 คน 2,000 คน และ 2,500 คน ค่าความชากเฉลี่ยของข้อสอบ 60 ข้อ
 ของกลุ่มตัวอย่าง 300 คน 400 คน 700 คน และ 3,000 คน ค่าความชากเฉลี่ยของ
 ข้อสอบ 70 ข้อ ของกลุ่มตัวอย่าง 1,500 คน 2,000 คน และ 2,500 คน ค่าความชาก
 เฉลี่ยของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่าง 900 คน ค่าความชากเฉลี่ยของข้อสอบ 90 ข้อ
 ของกลุ่มตัวอย่าง 900 คน 2,000 คน และ 3,000 คน และค่าความชากเฉลี่ยของข้อสอบ
 100 ข้อ ของกลุ่มตัวอย่าง 200 คน และ 1,500 คน ที่มีความแตกต่างกันอย่างมีนัยสำคัญ
 ทางสถิติ

4.9 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าการเดาของข้อสอบ
ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 77 -
79

ตาราง 77 เปรียบเทียบความแตกต่างของค่าการเดาของข้อสอบที่ประมาณค่าด้วยวิธี
แมกซิมัม ไลค์ลิสต์ เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _b	MS _w	F-test
200	5	0.070	0.014	0.001	11.860*
300	5	0.013	0.003	0.001	1.879
400	5	0.023	0.005	0.003	1.716
500	5	0.014	0.003	0.003	1.066
600	5	0.042	0.008	0.002	3.622*
700	5	0.021	0.004	0.003	1.554
800	5	0.009	0.002	0.002	0.976
900	5	0.016	0.003	0.002	1.789
1,000	5	0.022	0.004	0.003	1.739
1,500	5	0.010	0.002	0.004	0.571
2,000	5	0.018	0.003	0.004	0.838
2,500	5	0.009	0.002	0.004	0.464
3,000	5	0.017	0.003	0.005	0.714

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 77 พบว่า ค่าการเดาของข้อสอบของกลุ่มตัวอย่าง 200
คน และ 600 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ เมื่อจำนวนข้อสอบต่างกัน มีความ
แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังนั้น จึงได้ทำการทดสอบความแตกต่าง

ของค่าการเคาของข้อสอบ ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 78 - 79

ตาราง 78 การทดสอบความแตกต่างของค่าการเคาของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าด้วยวิธีแมกซีมัม ไลค์ลิตูด เมื่อกลุ่มตัวอย่างมีจำนวน 200 คน ด้วยการเปรียบเทียบ
พหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.247	0.231	0.224	0.219	0.230	0.252
50	0.247	-	0.016	0.023*	0.028*	0.017	0.005
60	0.231		-	0.007	0.012	0.001	0.021*
70	0.224			-	0.005	0.006	0.028*
80	0.219				-	0.011	0.033*
90	0.230					-	0.022*
100	0.252						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 78 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 200 คน ค่าการเคาเฉลี่ยของข้อสอบ 50 ข้อ มีความแตกต่างจากค่าการเคาเฉลี่ยของข้อสอบ 70 ข้อ และ 80 ข้อ และค่าการเคาเฉลี่ยของข้อสอบ 60 ข้อ 70 ข้อ 80 ข้อ และ 90 ข้อ มีความแตกต่างจากค่าการเคาเฉลี่ยของข้อสอบ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 79 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าตัวสถิติแมกซีมัม ไลค์ลิสต์ เมื่อกลุ่มตัวอย่างมีจำนวน 600 คน ด้วยการเปรียบเทียบ
พหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.253	0.241	0.233	0.226	0.224	0.223
50	0.253	-	0.012	0.020	0.027	0.029	0.030*
60	0.241		-	0.008	0.015	0.017	0.018
70	0.233			-	0.007	0.009	0.010
80	0.226				-	0.002	0.003
90	0.224					-	0.001
100	0.223						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 79 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 600 คน ค่าการ
เคาะเฉลี่ยของข้อสอบ 50 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 100 ข้อ
อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.10 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าการเคาะของข้อสอบ
ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 80 - 84

ตาราง 80 เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธี
อิวิริสติก เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _b	MS _w	F-test
200	5	0.177	0.035	0.009	3.843
300	5	0.054	0.011	0.005	2.248
400	5	0.062	0.012	0.005	2.283
500	5	0.030	0.006	0.004	1.566
600	5	0.015	0.003	0.006	0.498
700	5	0.038	0.008	0.004	1.962
800	5	0.091	0.018	0.009	2.088
900	5	8.173	1.635	0.374	4.370*
1,000	5	0.459	0.092	0.007	12.707*
1,500	5	0.156	0.031	0.010	3.255*
2,000	5	0.370	0.074	0.010	7.518*
2,500	5	0.070	0.014	0.007	2.027
3,000	5	0.081	0.016	0.009	1.776

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 80 พบว่า ค่าการเคาะของข้อสอบของกลุ่มตัวอย่าง 900
คน 1,000 คน 1,500 คน และ 2,000 คน ที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวน
ข้อสอบต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังนั้น จึงได้ทำการ
ทดสอบความแตกต่างของค่าการเคาะของข้อสอบ ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง

81 - 84

ตาราง 81 การทดสอบความแตกต่างของค่าการเคาของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าด้วยวิธีวิวิธวิธี เมื่อกลุ่มตัวอย่างมีจำนวน 900 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.142	0.156	0.128	0.135	0.129	0.466
50	0.142	-	0.014	0.014	0.007	0.013	0.324
60	0.156		-	0.028	0.021	0.027	0.310
70	0.128			-	0.007	0.001	0.338*
80	0.135				-	0.006	0.331*
90	0.129					-	0.337*
100	0.466						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 81 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 900 คน ค่าการเคาเฉลี่ยของข้อสอบ 70 ข้อ 80 ข้อ และ 90 ข้อ มีความแตกต่างจากค่าการเคาเฉลี่ยของข้อสอบ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 82 การทดสอบความแตกต่างของค่าการเคาของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าด้วยวิธีวิวิริสติก เมื่อกลุ่มตัวอย่างมีจำนวน 1,000 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.159	0.194	0.140	0.134	0.134	0.210
50	0.159	-	0.035	0.019	0.025	0.025	0.051
60	0.194		-	0.054*	0.060*	0.060*	0.016
70	0.140			-	0.006	0.006	0.070*
80	0.134				-	0.000	0.076*
90	0.134					-	0.076*
100	0.210						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 82 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน ค่าการเคาเฉลี่ยของข้อสอบ 60 ข้อ มีความแตกต่างจากค่าการเคาเฉลี่ยของข้อสอบ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 83 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าตัวสถิติวิวิธคติ เมื่อกลุ่มตัวอย่างมีจำนวน 1,500 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.131	0.133	0.180	0.138	0.151	0.123
50	0.131	-	0.002	0.049	0.007	0.020	0.008
60	0.133		-	0.047	0.005	0.018	0.010
70	0.180			-	0.042	0.029	0.057*
80	0.138				-	0.013	0.015
90	0.151					-	0.028
100	0.123						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 83 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน ค่า
การเคาะเฉลี่ยของข้อสอบ 70 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 100 ข้อ
อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 84 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าตัวสถิติวิวิธคดี เมื่อกลุ่มตัวอย่างมีจำนวน 2,000 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.128	0.162	0.110	0.115	0.121	0.183
50	0.128	-	0.034	0.018	0.013	0.007	0.055
60	0.162		-	0.052	0.047	0.041	0.021
70	0.110			-	0.005	0.011	0.073*
80	0.115				-	0.006	0.068*
90	0.121					-	0.062*
100	0.183						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 84 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน ค่าการเคาะเฉลี่ยของข้อสอบ 70 ข้อ 80 ข้อ และ 90 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.11 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าการเคาะของข้อสอบ
ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 85 - 88

ตาราง 85 เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีของ
เบส์ เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _b	MS _w	F-test
200	5	0.028	0.006	0.014	0.400
300	5	0.009	0.002	0.009	0.214
400	5	0.035	0.007	0.010	0.697
500	5	0.011	0.002	0.012	0.177
600	5	0.044	0.009	0.004	2.519
700	5	0.007	0.001	0.002	0.777
800	5	0.007	0.002	0.002	0.983
900	5	0.009	0.002	0.004	0.487
1,000	5	0.016	0.003	0.011	0.291
1,500	5	0.011	0.002	0.008	0.262
2,000	5	0.700	0.140	0.008	18.338*
2,500	5	0.775	0.155	0.006	25.120*
3,000	5	0.421	0.084	0.007	12.269*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 85 พบว่า ค่าการเคาะของข้อสอบของกลุ่มตัวอย่าง
2,000 คน 2,500 คน และ 3,000 คน ที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบ
ต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังนั้น จึงได้ทำการทดสอบ
ความแตกต่างของค่าการเคาะของข้อสอบ ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 86 -

ตาราง 86 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าตัวสถิติของเบส์ เมื่อกลุ่มตัวอย่างมีจำนวน 2,000 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.149	0.147	0.133	0.160	0.170	0.244
50	0.149	-	0.002	0.016	0.011	0.021	0.095*
60	0.147		-	0.014	0.013	0.023	0.097*
70	0.133			-	0.027	0.037	0.111*
80	0.160				-	0.010	0.084*
90	0.170					-	0.074*
100	0.244						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 86 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน ค่าการเคาะเฉลี่ยของข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ และ 90 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 87 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าตัววิธีของเบส์ เมื่อกลุ่มตัวอย่างมีจำนวน 2,500 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.296	0.253	0.251	0.243	0.158	0.246
50	0.296	-	0.043	0.045	0.053*	0.138*	0.050*
60	0.253		-	0.002	0.010	0.095*	0.007
70	0.251			-	0.008	0.093*	0.005
80	0.243				-	0.085*	0.003
90	0.158					-	0.088*
100	0.246						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 87 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน ค่าการเคาะเฉลี่ยของข้อสอบ 50 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 80 ข้อ ค่าการเคาะเฉลี่ยของข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ และ 80 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 90 ข้อ และค่าการเคาะเฉลี่ยของข้อสอบ 50 ข้อ และ 90 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 88 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่มีจำนวนต่างกันที่ประมาณ
ค่าด้วยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีจำนวน 3,000 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{c}$	50	60	70	80	90	100
		0.237	0.247	0.245	0.240	0.168	0.247
50	0.237	-	0.010	0.008	0.003	0.069*	0.010
60	0.247		-	0.002	0.007	0.079*	0.000
70	0.245			-	0.005	0.077*	0.002
80	0.240				-	0.072*	0.007
90	0.168					-	0.079*
100	0.247						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 88 พบว่า เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน ค่าการเคาะเฉลี่ยของข้อสอบ 50 ข้อ 60 ข้อ และ 70 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 90 ข้อ และค่าการเคาะเฉลี่ยของข้อสอบ 90 ข้อ มีความแตกต่างจากค่าการเคาะเฉลี่ยของข้อสอบ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.12 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าการเคาะของข้อสอบ
ที่ประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวีรอสติค และวิธีของเบส์ เมื่อจำนวนข้อสอบ
ต่างกัน ดังแสดงในตาราง 89 - 90

ตาราง 89 เปรียบเทียบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธี
แมกซิมัม โลคัลลิสต์ วิธีอีวีรอสติค และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
200	50	2	0.363	0.182	0.006	30.146*
	60	2	0.354	0.177	0.006	29.052*
	70	2	0.363	0.182	0.007	24.854*
	80	2	0.342	0.171	0.007	26.006*
	90	2	0.290	0.145	0.009	16.415*
	100	2	0.301	0.151	0.012	12.741*
300	50	2	0.616	0.308	0.004	77.462*
	60	2	0.810	0.405	0.004	97.366*
	70	2	0.648	0.324	0.006	56.135*
	80	2	0.757	0.379	0.005	79.220*
	90	2	0.964	0.482	0.005	95.485*
	100	2	0.769	0.384	0.005	73.983*
400	50	2	0.541	0.271	0.007	40.304*
	60	2	0.374	0.188	0.007	27.668*
	70	2	0.535	0.268	0.004	65.387*
	80	2	0.578	0.289	0.006	47.719*
	90	2	0.938	0.469	0.006	74.562*
	100	2	1.339	0.670	0.006	106.149*

ตาราง 89 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
500	50	2	0.560	0.280	0.006	49.929*
	60	2	0.737	0.369	0.007	49.580*
	70	2	0.772	0.386	0.007	59.122*
	80	2	0.980	0.490	0.006	98.069*
	90	2	1.276	0.638	0.006	100.423*
	100	2	1.413	0.707	0.006	124.582*
600	50	2	0.708	0.354	0.004	79.614*
	60	2	0.680	0.330	0.004	85.219*
	70	2	0.682	0.341	0.004	76.924*
	80	2	0.796	0.398	0.004	90.512*
	90	2	0.889	0.444	0.003	132.088*
	100	2	0.918	0.459	0.004	123.467*
700	50	2	0.746	0.373	0.002	188.664*
	60	2	0.821	0.410	0.003	157.696*
	70	2	0.744	0.372	0.003	115.350*
	80	2	1.113	0.557	0.002	233.554*
	90	2	1.116	0.558	0.003	168.833*
	100	2	1.040	0.520	0.003	172.241*
800	50	2	0.221	0.111	0.006	19.549*
	60	2	0.344	0.172	0.003	60.372*
	70	2	0.360	0.180	0.004	43.276*
	80	2	0.480	0.215	0.004	58.477*
	90	2	0.656	0.328	0.003	110.545*
	100	2	0.311	0.156	0.005	30.324*

ตาราง 89 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _b	MS _w	F-test
900	50	2	0.308	0.154	0.005	28.897*
	60	2	0.266	0.133	0.014	9.832*
	70	2	0.532	0.266	0.004	64.117*
	80	2	0.578	0.289	0.005	59.101*
	90	2	0.706	0.353	0.005	65.285*
	100	2	0.396	0.198	0.007	28.626*
1,000	50	2	0.186	0.093	0.005	17.704*
	60	2	0.049	0.024	0.005	4.571*
	70	2	0.372	0.186	0.008	24.685*
	80	2	0.383	0.192	0.006	34.053*
	90	2	0.454	0.227	0.007	32.799*
	100	2	0.035	0.018	0.009	2.020
1,500	50	2	0.384	0.195	0.006	34.000*
	60	2	0.394	0.197	0.006	32.431*
	70	2	0.138	0.069	0.011	6.406*
	80	2	0.566	0.283	0.008	34.706*
	90	2	0.400	0.200	0.007	27.157*
	100	2	0.859	0.430	0.005	84.841*
2,000	50	2	0.175	0.088	0.006	13.833*
	60	2	0.082	0.041	0.012	3.401*
	70	2	0.280	0.140	0.005	30.260*
	80	2	0.234	0.117	0.005	24.436*
	90	2	0.298	0.149	0.008	18.063*
	100	2	0.187	0.093	0.007	13.527*

ตาราง 90 การทดสอบความแตกต่างของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม
โลคัลลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ด้วยการเปรียบเทียบ
พหุคูณ

จำนวนคน	จำนวนข้อ	แมกซิมัม โลคัลลิสต์-อีวริสติก	แมกซิมัม โลคัลลิสต์-เบส์	อีวริสติก-เบส์
200	50	0.121*	0.060*	0.061*
	60	0.107*	0.039*	0.068*
	70	0.098*	0.024	0.074*
	80	0.089*	0.021	0.068*
	90	0.078*	0.022	0.056*
	100	0.078*	0.042*	0.036
300	50	0.132*	0.007	0.139*
	60	1.333*	0.018	0.151*
	70	0.113*	0.009	0.122*
	80	0.112*	0.012	0.124*
	90	0.123*	0.007	0.130*
	100	0.107*	0.001	0.108*
400	50	0.135*	0.017	0.118*
	60	0.107*	0.023	0.084*
	70	0.114*	0.015	0.099*
	80	0.113*	0.020	0.093*
	90	0.131*	0.012	0.119*
	100	0.149*	0.016	0.133*
500	50	0.144*	0.034	0.110*
	60	0.144*	0.018	0.126*
	70	0.131*	0.004	0.127*
	80	0.139*	0.007	0.132*

ตาราง 90 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
900	90	0.098*	0.018	0.116*
	100	0.069*	0.014	0.083*
1,000	50	0.057*	0.027	0.084*
	60	0.008	0.030	0.038*
	70	0.063*	0.040*	0.103*
	80	0.065*	0.030*	0.095*
	90	0.079*	0.014	0.093*
	100	-	-	-
1,500	50	0.090*	0.031	0.121*
	60	0.076*	0.037*	0.113*
	70	0.029	0.033	0.062*
	80	0.067*	0.052*	0.119*
	90	0.062*	0.031	0.093*
	100	0.092*	0.034*	0.126*
2,000	50	0.081*	0.060*	0.021
	60	0.036	0.051*	0.015
	70	0.087*	0.064*	0.023
	80	0.076*	0.031*	0.045*
	90	0.080*	0.031	0.049*
	100	0.023	0.038*	0.061*
2,500	50	0.091*	0.094*	0.185*
	60	0.061*	0.064*	0.125*
	70	0.046*	0.059*	0.105*
	80	0.081*	0.051*	0.132*

ข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 600 คน 700 คน 800 คน และ 2,500 คน และค่าการเดาเฉลี่ยของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 200 คน 600 คน 700 คน 800 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด กับวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ค่าการเดาเฉลี่ยของข้อสอบของกลุ่มตัวอย่างที่มีจำนวนต่างกันที่ประมาณค่าด้วยวิธีอิวิริสติก กับวิธีของเบส์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ยกเว้น ค่าการเดาเฉลี่ยของข้อสอบ 50 ข้อ 60 ข้อ และ 70 ข้อ ของกลุ่มตัวอย่าง 2,000 คน ค่าการเดาเฉลี่ยของข้อสอบ 90 ข้อ ของกลุ่มตัวอย่าง 2,500 คน และ 3,000 คน ค่าการเดาเฉลี่ยของข้อสอบ 100 ข้อ ของกลุ่มตัวอย่าง 200 คน ที่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.13 ผลการเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่
ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 91

ตาราง 91 เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
แมกซิมัม ไลค์ลิตูด เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _B	MS _w	F-test
200	5	0.000	0.000	1.005	0.000
300	5	0.001	0.000	1.003	0.000
400	5	0.000	0.000	1.003	0.000
500	5	0.000	0.000	1.002	0.000
600	5	0.000	0.000	1.002	0.000
700	5	0.000	0.000	1.001	0.000
800	5	0.000	0.000	1.001	0.000
900	5	0.000	0.000	1.001	0.000
1,000	5	0.000	0.000	1.001	0.000
1,500	5	0.000	0.000	1.001	0.000
2,000	5	0.000	0.000	1.001	0.000
2,500	5	0.000	0.000	1.000	0.000
3,000	5	0.000	0.000	1.000	0.000

จากตาราง 91 พบว่า ค่าความสามารถของผู้สอบที่ประมาณค่าด้วย
วิธีแมกซิมัม ไลค์ลิตูด เมื่อจำนวนข้อสอบต่างกัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.14 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 92 - 105

ตาราง 92 เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _b	MS _w	F-test
200	5	8.059	1.612	0.202	7.963*
300	5	8.483	1.697	0.269	6.300*
400	5	14.815	2.963	0.257	11.549*
500	5	9.030	1.806	0.213	8.462*
600	5	19.071	3.814	0.211	18.070*
700	5	12.966	2.599	0.273	9.536*
800	5	8.666	1.733	0.224	7.743*
900	5	16.630	3.326	0.173	19.260*
1,000	5	4.535	0.970	0.260	3.485*
1,500	5	28.662	5.732	0.238	24.051*
2,000	5	20.288	4.058	0.248	16.366*
2,500	5	30.063	6.013	0.230	26.171*
3,000	5	54.162	10.832	0.210	51.503*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 92 พบว่า ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังนั้น ผู้วิจัยจึงได้ทำการทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ด้วยการเปรียบเทียบพหุคูณ ผลการทดสอบดังแสดงในตาราง 93 - 105

ตาราง 93 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
อิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 200 คน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		0.750	1.027	0.932	0.926	0.901	0.924
50	0.750	-	0.277*	0.182*	0.176*	0.151*	0.174*
60	1.027		-	0.095	0.101	0.126	0.103
70	0.932			-	0.006	0.031	0.008
80	0.926				-	0.025	0.002
90	0.901					-	0.003
100	0.924						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 93 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 200 คน
เมื่อข้อสอบมีขนาด 50 ข้อ มีความแตกต่างจากข้อสอบ 60 ข้อ 70 ข้อ 90 ข้อ และ 100
ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 94 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
 อีวริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 300 คน ด้วยการ
 เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		1.082	1.090	0.955	0.951	0.917	0.948
50	1.082	-	0.008	0.127	0.131	0.165*	0.134
60	1.090		-	0.135	0.139	0.173*	0.142*
70	0.955			-	0.004	0.038	0.007
80	0.951				-	0.034	0.003
90	0.917					-	0.031
100	0.948						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 94 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 300 คน
 เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 90 ข้อ และค่าความ
 สามารถเฉลี่ยของผู้สอบ เมื่อข้อสอบมีขนาด 60 ข้อ มีความแตกต่างจากข้อสอบ 100 ข้อ
 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 95 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
ธีวรีสถิติ เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 400 คน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{a}$	50	60	70	80	90	100
		0.704	0.960	0.866	0.893	0.879	0.898
50	0.704	-	0.256*	0.162*	0.189*	0.175*	0.194*
60	0.960		-	0.094	0.067	0.081	0.062
70	0.866			-	0.027	0.013	0.032
80	0.893				-	0.014	0.005
90	0.879					-	0.019
100	0.898						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 95 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 400 คน
เมื่อข้อสอบมีขนาด 50 ข้อ มีความแตกต่างจากข้อสอบ 60 ข้อ 70 ข้อ 80 ข้อ 90 ข้อ
และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 96 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
 สิวรีสถิต เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 500 คน ด้วยการใช้
 เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		1.079	1.073	0.965	0.965	0.941	0.978
50	1.079	-	0.006	0.114*	0.114*	0.138*	0.101*
60	1.073		-	0.108*	0.108*	0.132*	0.095
70	0.965			-	0.000	0.024	0.013
80	0.965				-	0.024	0.013
90	0.941					-	0.037
100	0.978						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 96 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 500 คน
 เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ 80 ข้อ และ
 90 ข้อ และค่าความสามารถเฉลี่ยของผู้สอบ เมื่อข้อสอบมีขนาด 50 ข้อ มีความแตกต่าง
 จากข้อสอบ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 97 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
ฮิวริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 600 คน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		1.125	1.100	0.973	0.970	0.940	0.959
50	1.125	-	0.025	0.152*	0.155*	0.185*	0.166*
60	1.100		-	0.127*	0.130*	0.160*	0.141*
70	0.973			-	0.003	0.033	0.014
80	0.970				-	0.030	0.011
90	0.940					-	0.019
100	0.959						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 97 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 600 คน
เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ 80 ข้อ 90 ข้อ
และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 98 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
อิวิรส์ติค เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 700 คน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		1.036	1.025	0.919	0.931	0.901	0.906
50	1.036	-	0.011	0.117*	0.105*	0.135*	0.130*
60	1.025		-	0.106*	0.094*	0.124*	0.119*
70	0.919			-	0.012	0.018	0.013
80	0.931				-	0.030	0.025
90	0.901					-	0.005
100	0.906						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 98 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 700 คน
เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ 80 ข้อ 90 ข้อ
และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 99 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
อีวริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 800 คน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		0.973	0.992	0.899	0.904	0.892	0.883
50	0.973	-	0.019	0.074	0.069	0.081*	0.090*
60	0.992		-	0.093*	0.088*	0.100	0.109*
70	0.899			-	0.005	0.007	0.016
80	0.904				-	0.012	0.021
90	0.892					-	0.009
100	0.883						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 99 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 800 คน
เมื่อข้อสอบมีขนาด 50 ข้อ มีความแตกต่างจากข้อสอบ 90 ข้อ และ 100 ข้อ และค่าความ
สามารถเฉลี่ยของผู้สอบ เมื่อข้อสอบมีขนาด 60 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ
80 ข้อ และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 100 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
ชีววิธี เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 900 คน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		1.048	1.020	0.920	0.929	0.911	0.912
50	1.048	-	0.028	0.128*	0.119*	0.137*	0.136*
60	1.020		-	0.100*	0.091*	0.109*	0.108*
70	0.920			-	0.009	0.009	0.008
80	0.929				-	0.018	0.017
90	0.911					-	0.001
100	0.912						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 100 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 900 คน
เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ 80 ข้อ 90
ข้อ และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 101 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วย
วิธีสัณฐานวิธี เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 1,000 คน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	$\bar{a}$	50	60	70	80	90	100
		0.903	0.884	0.822	0.865	0.836	0.853
50	0.903	-	0.019	0.081*	0.038	0.067	0.050
60	0.884		-	0.062	0.019	0.048	0.031
70	0.822			-	0.043	0.014	0.031
80	0.865				-	0.029	0.012
90	0.836					-	0.017
100	0.853						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 101 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 1,000 คน
เมื่อข้อสอบมีขนาด 50 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ อย่างมีนัยสำคัญทางสถิติที่
ระดับ .05

ตาราง 102 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีฮิวริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 1,500 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		1.004	0.989	0.860	0.899	0.863	0.911
50	1.004	-	0.015	0.144*	0.105*	0.141*	0.093*
60	0.989		-	0.129*	0.090*	0.126*	0.078*
70	0.860			-	0.039	0.003	0.051
80	0.899				-	0.036	0.012
90	0.863					-	0.048
100	0.911						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 102 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 1,500 คน เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ 80 ข้อ 90 ข้อ และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 103 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีวิธีสถิติ เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 2,000 คน ด้วยการเปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		0.970	0.960	0.888	0.905	0.869	0.867
50	0.970	-	0.010	0.082*	0.065*	0.101*	0.103*
60	0.960		-	0.072*	0.055	0.091*	0.093*
70	0.888			-	0.017	0.019	0.021
80	0.905				-	0.036	0.038
90	0.869					-	0.002
100	0.867						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 103 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 2,000 คน เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 90 ข้อ และ 100 ข้อ และค่าความสามารถเฉลี่ยของผู้สอบ เมื่อข้อสอบมีขนาด 50 ข้อ มีความแตกต่างจากข้อสอบ 80 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 104 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วย
วิธีชีววิธีคิด เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 2,500 คน ด้วยการ
เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		0.983	0.959	0.864	0.907	0.867	0.889
50	0.983	-	0.024	0.119*	0.076*	0.116*	0.094*
60	0.959		-	0.095*	0.052*	0.092*	0.070*
70	0.864			-	0.043	0.003	0.025
80	0.907				-	0.040	0.018
90	0.867					-	0.022
100	0.889						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 104 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 2,500 คน
เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ 80 ข้อ 90 ข้อ
และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 105 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วย
วิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ของกลุ่มตัวอย่างจำนวน 3,000 คน ด้วยการใช้
เปรียบเทียบพหุคูณ

จำนวนข้อ	อ	50	60	70	80	90	100
		0.981	0.974	0.849	0.875	0.854	0.873
50	0.981	-	0.007	0.132*	0.106*	0.127*	0.108*
60	0.974		-	0.125*	0.099*	0.120*	0.101*
70	0.849			-	0.026	0.005	0.024
80	0.875				-	0.021	0.002
90	0.854					-	0.019
100	0.873						-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 105 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบ 3,000 คน
เมื่อข้อสอบมีขนาด 50 ข้อ และ 60 ข้อ มีความแตกต่างจากข้อสอบ 70 ข้อ 80 ข้อ 90 ข้อ
และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.15 ผลการเปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ดังแสดงในตาราง 106

ตาราง 106 เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	df	SS	MS _b	MS _w	F-test
200	5	0.000	0.000	1.006	0.000
300	5	0.001	0.000	1.005	0.000
400	5	0.001	0.000	1.001	0.000
500	5	0.001	0.000	1.001	0.000
600	5	0.001	0.000	1.001	0.000
700	5	0.000	0.000	1.001	0.000
800	5	0.000	0.000	1.001	0.000
900	5	0.001	0.000	1.002	0.000
1,000	5	0.000	0.000	1.001	0.000
1,500	5	0.001	0.000	1.001	0.000
2,000	5	0.000	0.000	1.001	0.000
2,500	5	0.000	0.000	1.001	0.000
3,000	5	0.000	0.000	1.001	0.000

จากตาราง 106 พบว่า ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.16 ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าความสามารถของ
ผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวน
ข้อสอบต่างกัน ดังแสดงในตาราง 107 - 108

ตาราง 107 เปรียบเทียบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธี
แมกซิมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
200	50	2	74.939	37.470	1.226	30.571*
	60	2	140.509	70.254	1.221	57.541*
	70	2	115.844	57.922	1.261	45.917*
	80	2	114.485	57.243	1.258	45.508*
	90	2	108.109	54.055	1.198	45.119*
	100	2	113.856	56.928	1.157	49.204*
300	50	2	234.127	117.063	1.240	94.443*
	60	2	237.699	118.850	1.388	85.653*
	70	2	182.066	91.033	1.220	74.647*
	80	2	182.299	91.149	1.253	72.756*
	90	2	168.265	84.133	1.280	65.754*
	100	2	179.669	89.834	1.143	78.577*
400	50	2	132.058	66.029	1.301	50.748*
	60	2	245.968	122.984	1.268	96.957*
	70	2	199.495	99.748	1.270	78.570*
	80	2	212.932	106.466	1.287	82.734*
	90	2	206.391	103.196	1.295	79.695*
	100	2	214.761	107.380	1.202	89.351*

ตาราง 107 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
500	50	2	388.332	194.166	1.198	162.073*
	60	2	384.068	192.034	1.195	160.680*
	70	2	310.075	155.037	1.205	128.689*
	80	2	310.612	155.306	1.232	126.091*
	90	2	295.191	147.596	1.171	126.032*
	100	2	318.438	159.219	1.215	131.060*
600	50	2	505.748	252.874	1.185	213.330*
	60	2	483.942	241.971	1.192	203.079*
	70	2	378.637	189.319	1.166	162.433*
	80	2	376.487	188.244	1.221	154.139*
	90	2	352.997	176.499	1.160	152.213*
	100	2	337.856	183.928	1.141	161.268*
700	50	2	500.097	250.049	1.232	202.953*
	60	2	490.131	245.066	1.102	222.476*
	70	2	394.144	197.072	1.063	185.376*
	80	2	404.241	202.120	1.014	199.331*
	90	2	378.588	189.294	0.957	197.892*
	100	2	382.426	191.213	1.032	185.288*
800	50	2	505.558	252.779	1.093	231.383*
	60	2	525.070	262.535	1.103	238.074*
	70	2	431.127	215.564	0.966	223.201*
	80	2	435.621	217.810	1.007	216.265*
	90	2	424.154	212.077	0.940	225.705*
	100	2	415.198	207.599	0.968	214.385*

ตาราง 107 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
900	50	2	658.492	329.246	0.950	346.521*
	60	2	623.473	311.736	1.017	306.488*
	70	2	507.112	253.556	0.931	272.388*
	80	2	517.194	258.597	0.965	267.865*
	90	2	497.147	248.570	0.930	267.290*
1,000	100	2	499.208	249.604	0.966	258.464*
	50	2	544.358	272.179	1.057	257.526*
	60	2	520.412	260.206	1.054	246.943*
	70	2	450.670	225.335	1.101	204.723*
	80	2	499.119	249.560	0.907	275.060*
1,500	90	2	465.072	232.536	0.907	256.381*
	100	2	485.148	242.574	0.971	249.897*
	50	2	1,007.269	503.634	1.114	452.036*
	60	2	978.095	489.048	1.064	459.631*
	70	2	739.308	369.654	0.958	385.820*
2,000	80	2	808.053	404.027	1.072	376.811*
	90	2	744.909	372.455	1.029	362.075*
	100	2	829.672	414.836	1.132	366.547*
	50	2	1,254.167	627.084	1.003	625.462*
	60	2	1,227.925	613.963	0.874	702.734*
	70	2	1,051.559	525.780	0.903	582.521*
	80	2	1,091.974	545.987	0.827	659.875*
	90	2	1,005.610	502.805	0.915	549.337*
	100	2	1,001.548	500.774	1.105	453.385*

ตาราง 107 (ต่อ)

จำนวนคน	จำนวนข้อ	df	SS	MS _B	MS _w	F-test
2,500	50	2	1,310.000	805.000	1.038	775.902*
	60	2	1,531.521	765.761	1.028	745.049*
	70	2	1,244.855	622.427	1.040	598.614*
	80	2	1,370.415	685.208	1.073	638.376*
	90	2	1,252.951	626.476	1.025	611.425*
	100	2	1,317.144	685.572	0.967	681.151*
3,000	50	2	1,923.104	961.552	0.944	1,018.462*
	60	2	1,895.957	947.979	0.932	1,017.447*
	70	2	1,440.796	720.398	0.919	783.854*
	80	2	1,531.149	765.575	0.958	799.430*
	90	2	1,458.113	729.057	0.924	788.302*
	100	2	1,524.251	762.126	1.033	737.592*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 107 พบว่า ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซีมัม โคลลิฮูด วิธีอีวีรติค และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผู้วิจัยจึงได้ทำการทดสอบความแตกต่างของค่าความสามารถของผู้สอบ ด้วยการเปรียบเทียบพหุคูณ ได้ผลดังตาราง 108

ตาราง 108 การทดสอบความแตกต่างของค่าความสามารถของผู้สอบที่ประมาณค่าด้วย
วิธีแมกซีมัม ไลค์ลิสต์ วิธีอีวาวิสตติค และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ด้วย
การเปรียบเทียบพหุคูณ

จำนวนคน	จำนวนข้อ	แมกซีมัม ไลค์ลิสต์-อีวาวิสตติค	แมกซีมัม ไลค์ลิสต์-เบส์	อีวาวิสตติค-เบส์
200	50	0.750*	0.001	0.749*
	60	1.207*	0.001	1.206*
	70	0.932*	0.000	0.932*
	80	0.926*	0.001	0.927*
	90	0.901*	0.000	0.901*
	100	0.924*	0.001	0.925*
300	50	0.958*	0.001	1.083*
	60	1.090*	0.000	1.090*
	70	0.955*	0.001	0.954*
	80	0.958*	0.001	0.957*
	90	0.917*	0.000	0.917*
	100	0.948*	0.000	0.948*
400	50	0.704*	0.001	0.703*
	60	0.960*	0.000	0.960*
	70	0.867*	0.001	0.866*
	80	0.893*	0.001	0.894*
	90	0.879*	0.002	0.881*
	100	0.898*	0.001	0.897*
500	50	1.079*	0.000	1.079*
	60	1.073*	0.000	1.073*
	70	0.965*	0.001	0.964*
	80	0.965*	0.001	0.966*

ตาราง 108 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไลค์ลิสต์-อีวริสติก	แมกซิมัม ไลค์ลิสต์-เบส	อีวริสติก-เบส
500	90	0.941*	0.000	0.941*
	100	0.978*	0.000	0.978*
600	50	1.125*	0.000	1.125*
	60	1.100*	0.001	1.099*
	70	0.973*	0.000	0.973*
	80	0.970*	0.001	0.971*
	90	0.940*	0.000	0.940*
	100	0.959*	0.000	0.959*
700	50	1.038*	0.000	1.038*
	60	1.205*	0.001	1.204*
	70	0.919*	0.000	0.919*
	80	0.931*	0.000	0.931*
	90	0.901*	0.001	0.901*
	100	0.906*	0.000	0.905*
800	50	0.973*	0.001	0.974*
	60	0.992*	0.001	0.993*
	70	0.899*	0.001	0.900*
	80	0.904*	0.000	0.904*
	90	0.892*	0.000	0.892*
	100	0.883*	0.001	0.882*
900	50	1.048*	0.001	1.049*
	60	1.020*	0.000	1.020*
	70	0.920*	0.000	0.920*
	80	0.929*	0.000	0.929*

ตาราง 108 (ต่อ)

จำนวนคน	จำนวนข้อ	แมกซิมัม ไคส์ลิสต์-ฮิวริสติก	แมกซิมัม ไคส์ลิสต์-เบส	ฮิวริสติก-เบส
900	90	0.911*	0.001	0.911*
	100	0.912*	0.000	0.913*
1,000	50	0.903*	0.001	0.904*
	60	0.884*	0.001	0.883*
	70	0.822*	0.000	0.822*
	80	0.865*	0.000	0.865*
	90	0.836*	0.000	0.836*
	100	0.853*	0.000	0.853*
1,500	50	1.004*	0.000	1.004*
	60	0.989*	0.000	0.989*
	70	0.860*	0.000	0.860*
	80	0.899*	0.000	0.899*
	90	0.863*	0.000	0.863*
	100	0.911*	0.000	0.911*
2,000	50	0.970*	0.000	0.970*
	60	0.960*	0.000	0.960*
	70	0.888*	0.000	0.888*
	80	0.905*	0.000	0.905*
	90	0.869*	0.000	0.869*
	100	0.867*	0.000	0.867*
2,500	50	0.983*	0.000	0.983*
	60	0.959*	0.000	0.959*
	70	0.864*	0.000	0.864*
	80	0.907*	0.000	0.907*

ตาราง 108 (ต่อ)

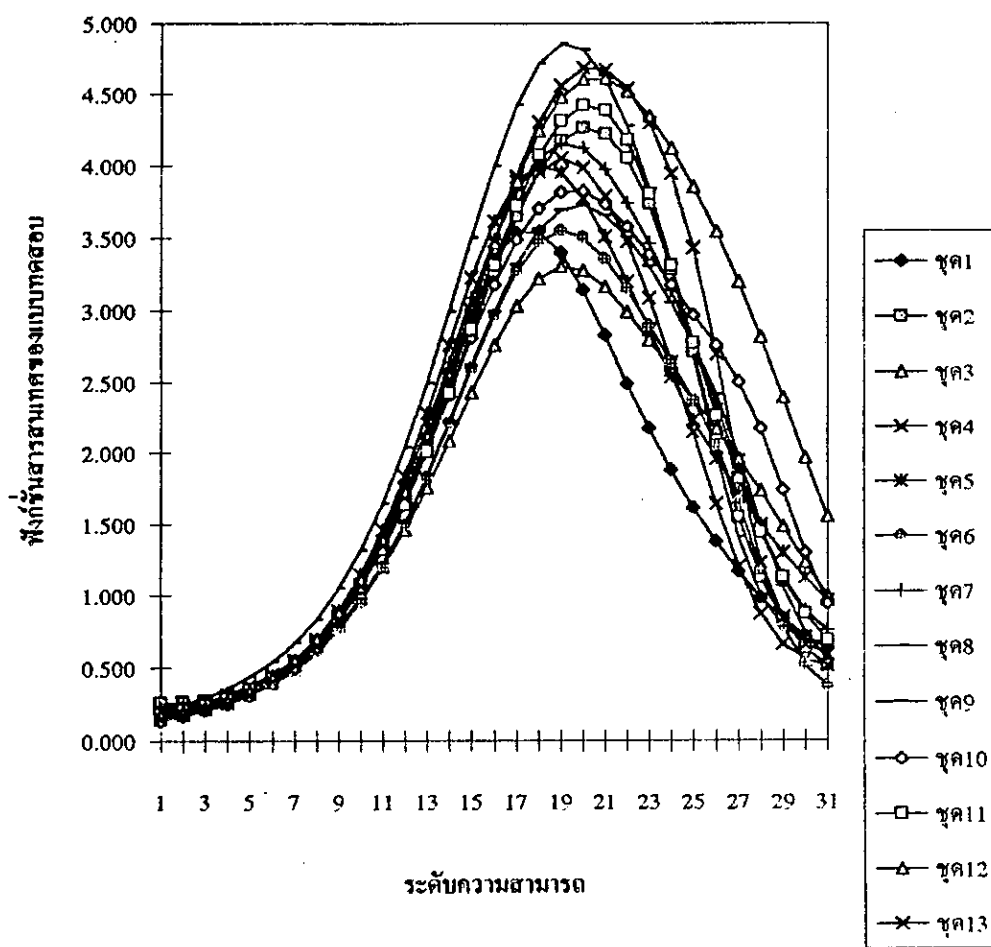
จำนวนคน	จำนวนข้อ	แมกซิมัม ไคส์ลิสต์-ฮิวริสติก	แมกซิมัม ไคส์ลิสต์-เบส	ฮิวริสติก-เบส
2,500	90	0.867*	0.000	0.867*
	100	0.889*	0.000	0.889*
3,000	50	0.981*	0.000	0.981*
	60	0.974*	0.000	0.974*
	70	0.849*	0.000	0.849*
	80	0.875*	0.000	0.875*
	90	0.854*	0.000	0.854*
	100	0.873*	0.000	0.873*

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 108 พบว่า ค่าความสามารถเฉลี่ยของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไคส์ลิสต์ กับวิธีฮิวริสติก และวิธีฮิวริสติก กับวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และค่าความสามารถเฉลี่ยของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไคส์ลิสต์ กับวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

5. ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไคส์ลิสต์ วิธีฮิวริสติก และวิธีของเบส เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความน่าเชื่อถือต่าง ๆ

5.1 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไคส์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความน่าเชื่อถือต่าง ๆ ดังภาพประกอบ 7 - 12

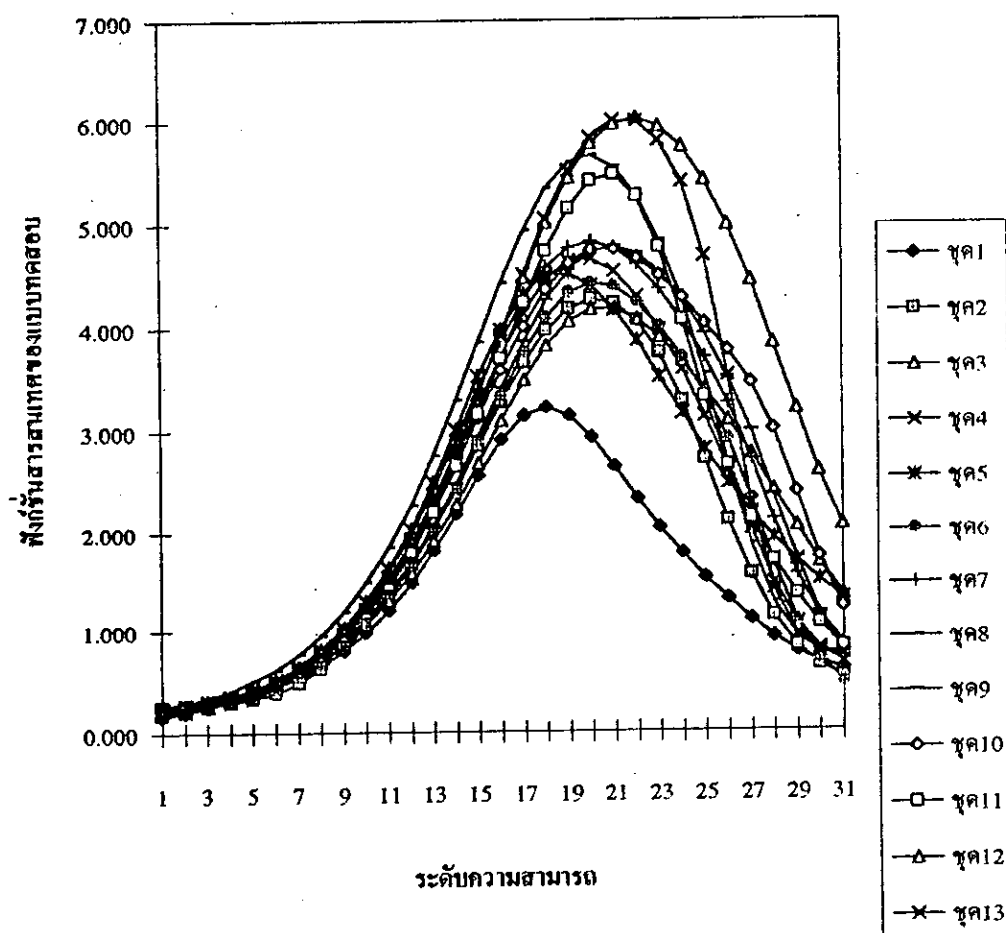


ภาพประกอบ 7 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

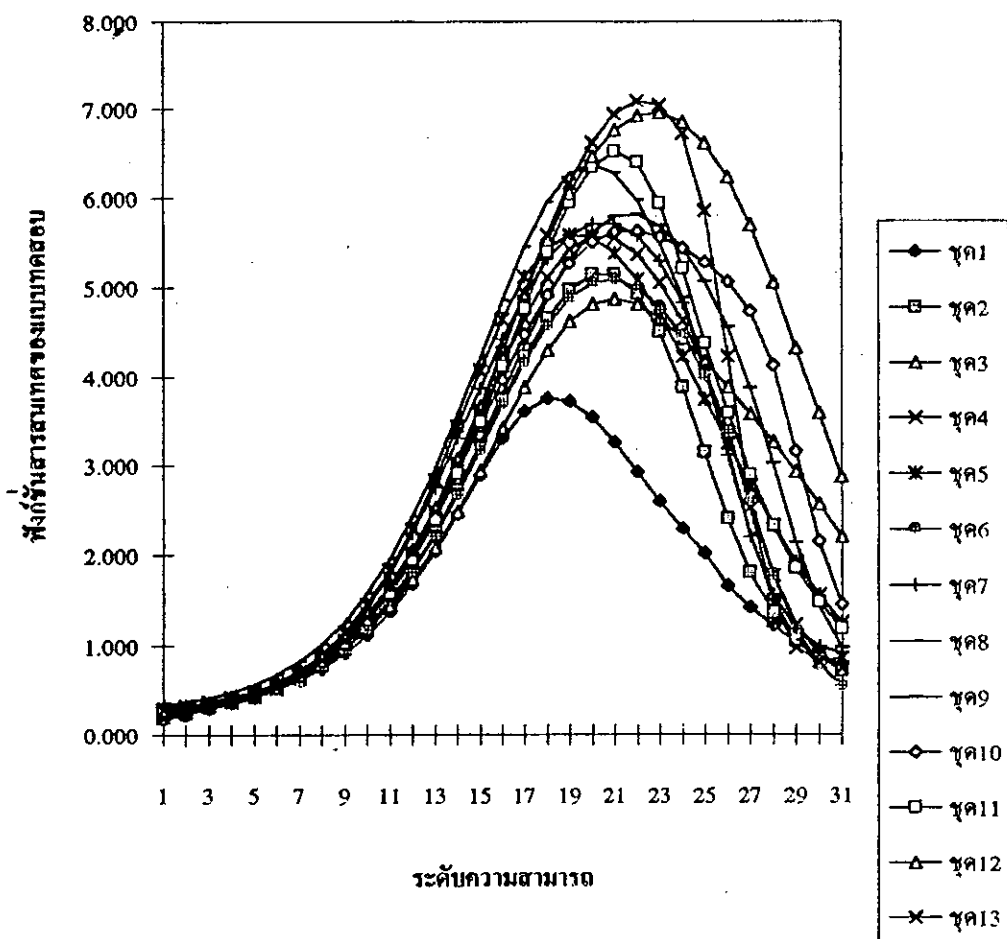


ภาพประกอบ 8 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซ์ิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

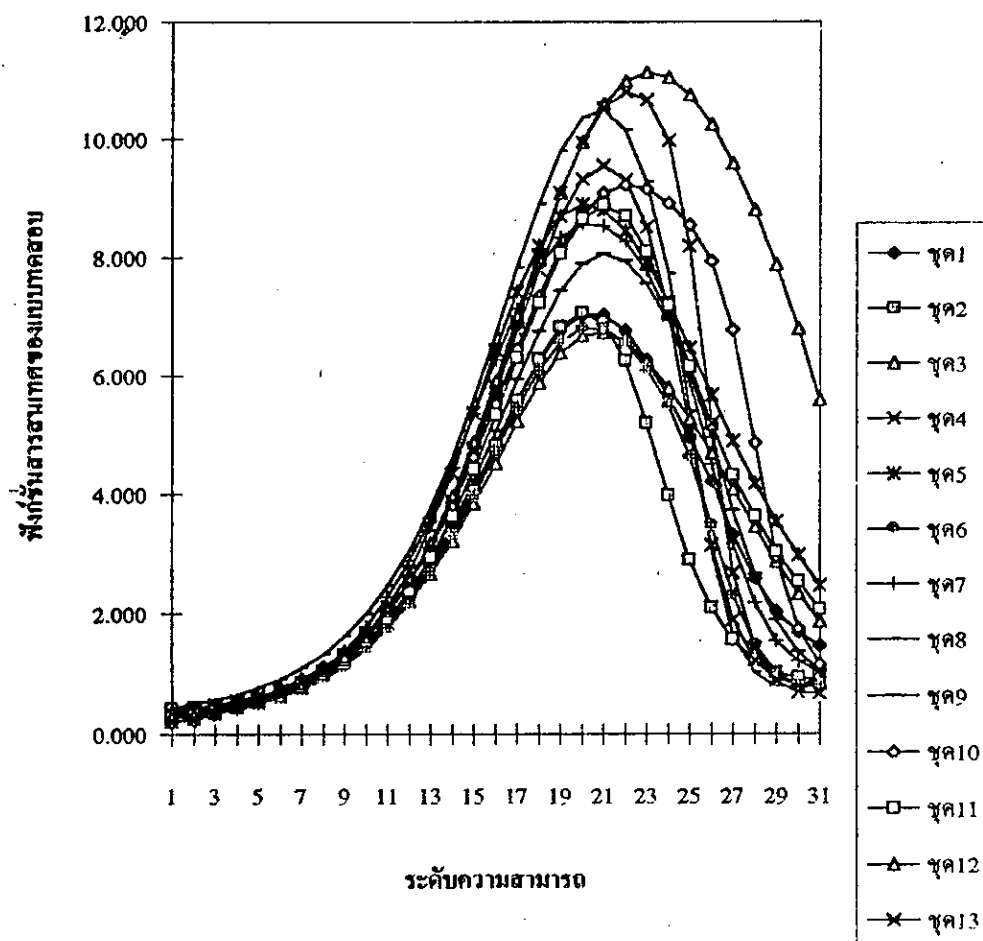


ภาพประกอบ 9 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิสต์ เมื่อข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

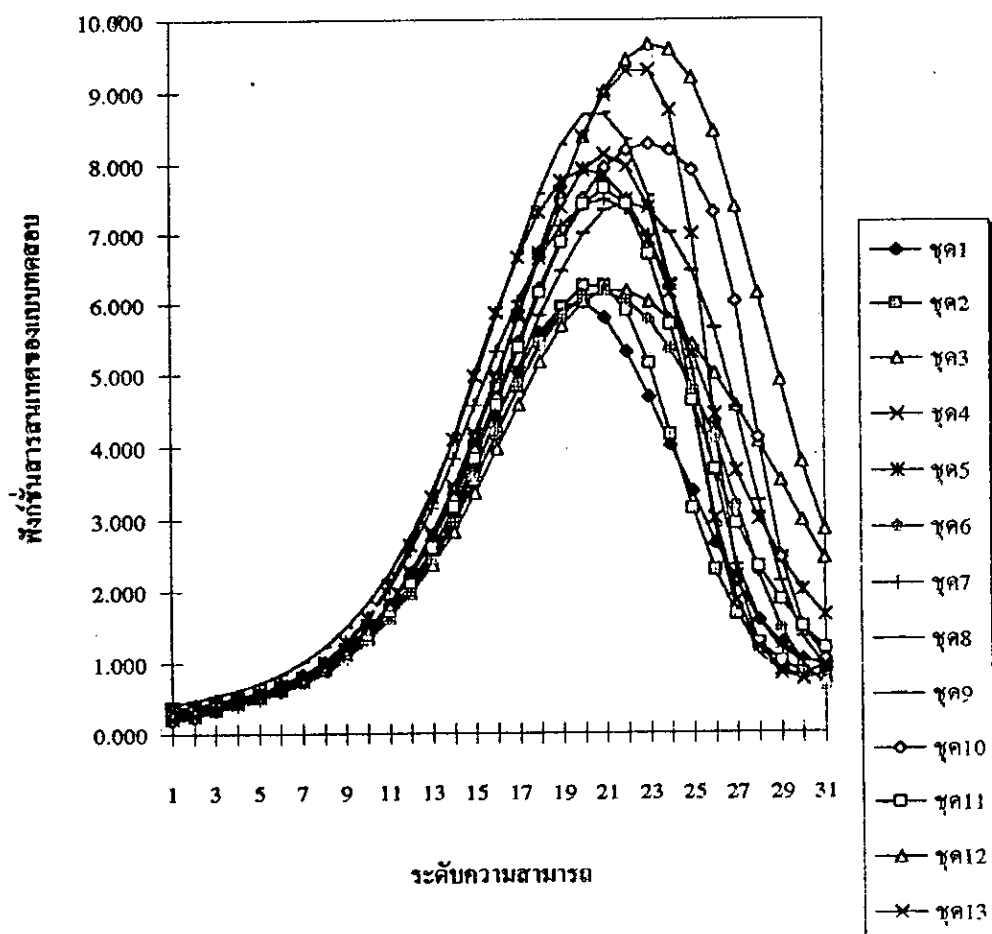


ภาพประกอบ 10 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซ์ิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

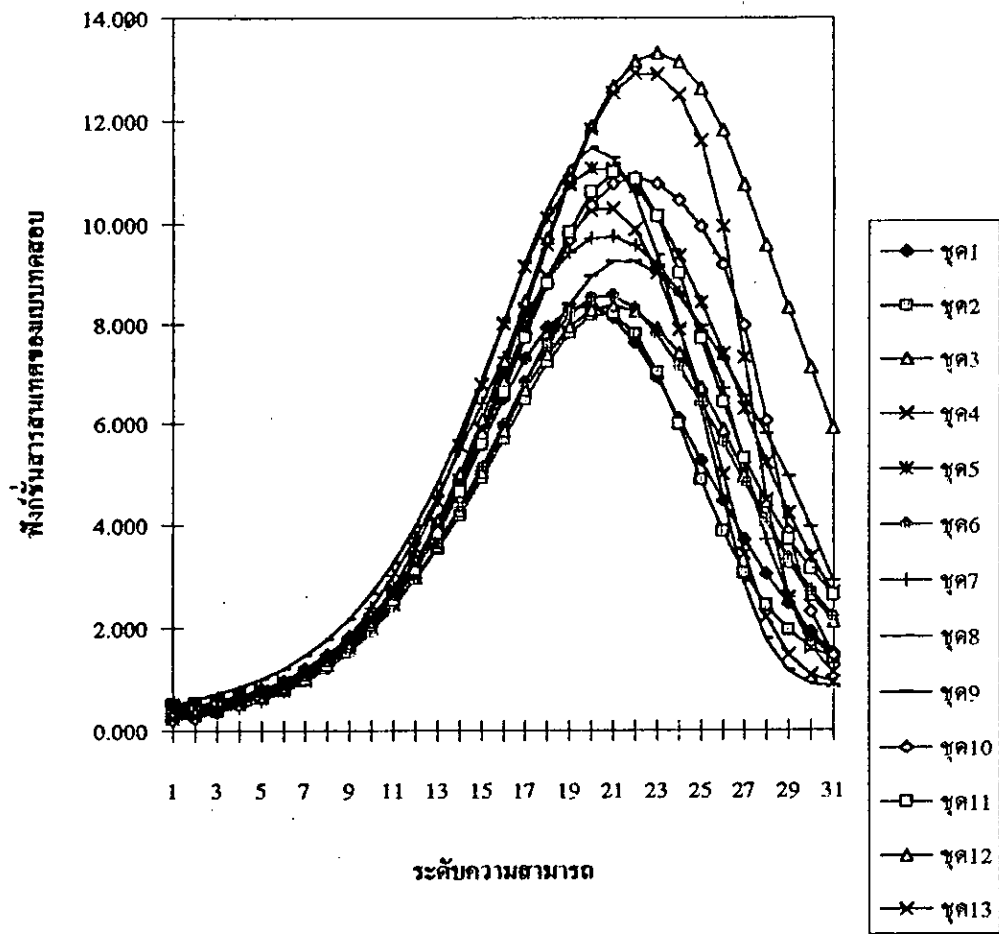


ภาพประกอบ 11 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน



ภาพประกอบ 12 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่า พารามิเตอร์โดยวิธีแมกซ์ิมัม โลคัลลิสต์ เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมี ขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

จากภาพประกอบ 7 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 50 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.296; e = 0.600\}$ เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 900 คน $\{I(e) = 4.853; e = 0.600\}$

จากภาพประกอบ 8 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 60 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.219; e = 0.400\}$ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน $\{I(e) = 6.019; e = 1.200\}$

จากภาพประกอบ 9 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 70 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.741; e = 0.400\}$ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน $\{I(e) = 7.084; e = 1.200\}$

จากภาพประกอบ 10 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 80 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 6.008; e = 0.800\}$ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน $\{I(e) = 9.642; e = 1.400\}$

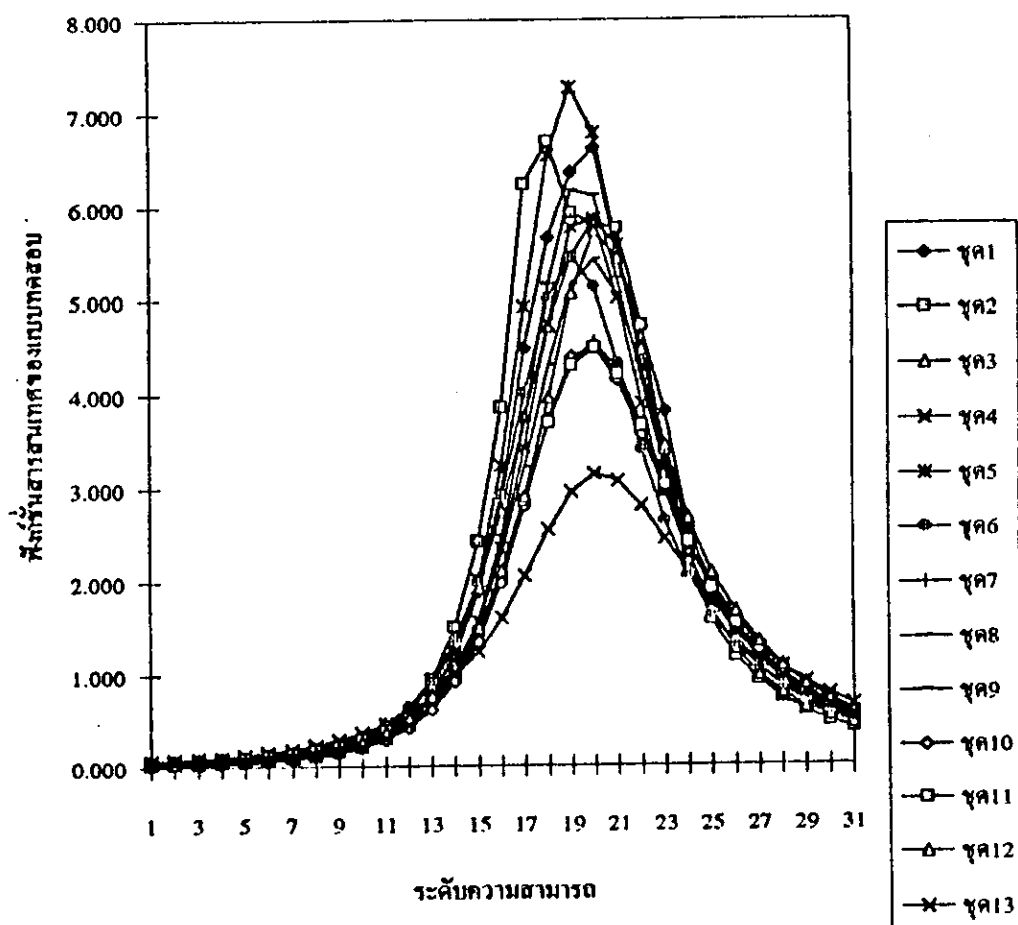
จากภาพประกอบ 11 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 90 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 6.713; e = 1.000\}$ เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน $\{I(e) = 11.141; e = 1.400\}$

จากภาพประกอบ 12 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 100 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 8.189; e = 1.000\}$ เมื่อกลุ่มตัวอย่างมีขนาด 300 คน และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน $\{I(e) = 13.304; e = 1.400\}$

$$\theta = 1.400\}$$

ดังนั้น จากภาพประกอบ 7 - 12 แสดงว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 3.219 ; \theta = 0.400\}$ เมื่อข้อสอบมีขนาด 60 ข้อ ขนาดของกลุ่มตัวอย่าง 200 คน และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อข้อสอบมีขนาด 100 ข้อ ขนาดของกลุ่มตัวอย่าง 2,500 คน $\{I(\theta) = 13.304 ; \theta = 1.400\}$

5.2 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีเอสดีค เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 13 - 18

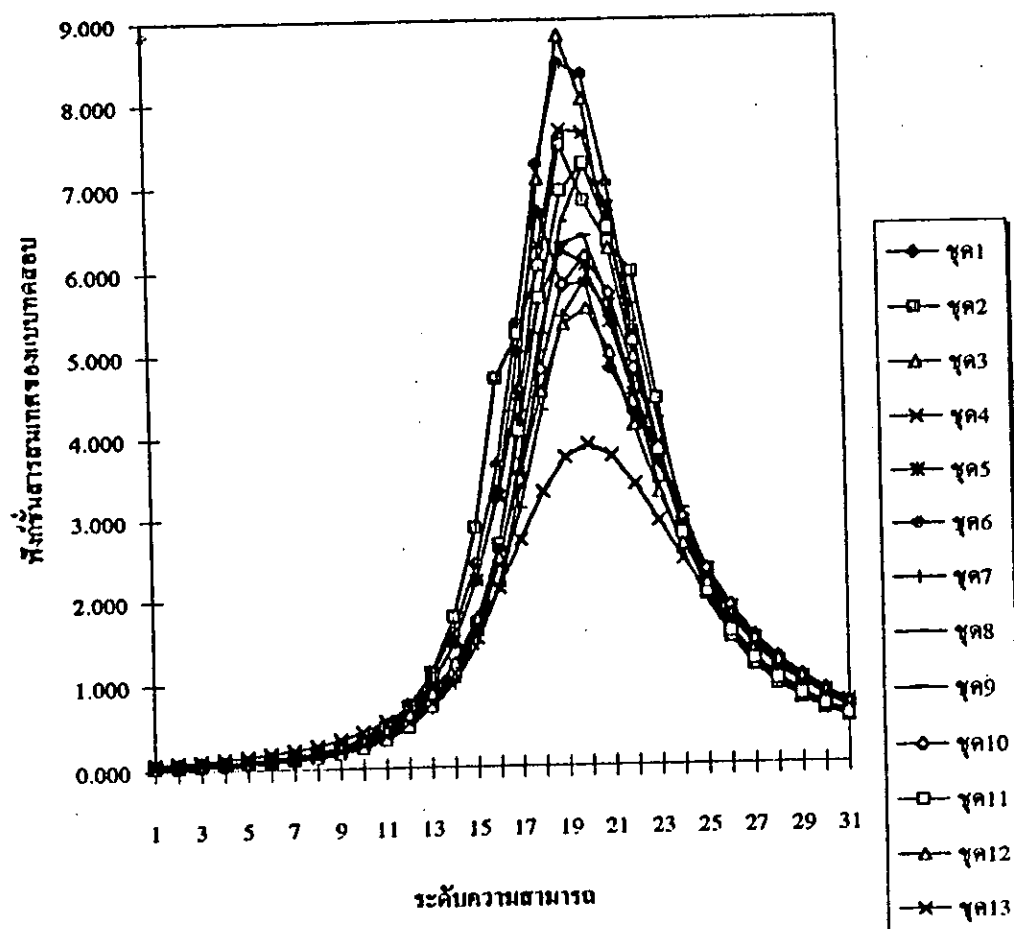


ภาพประกอบ 13 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอวิวิธิต เมื่อข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

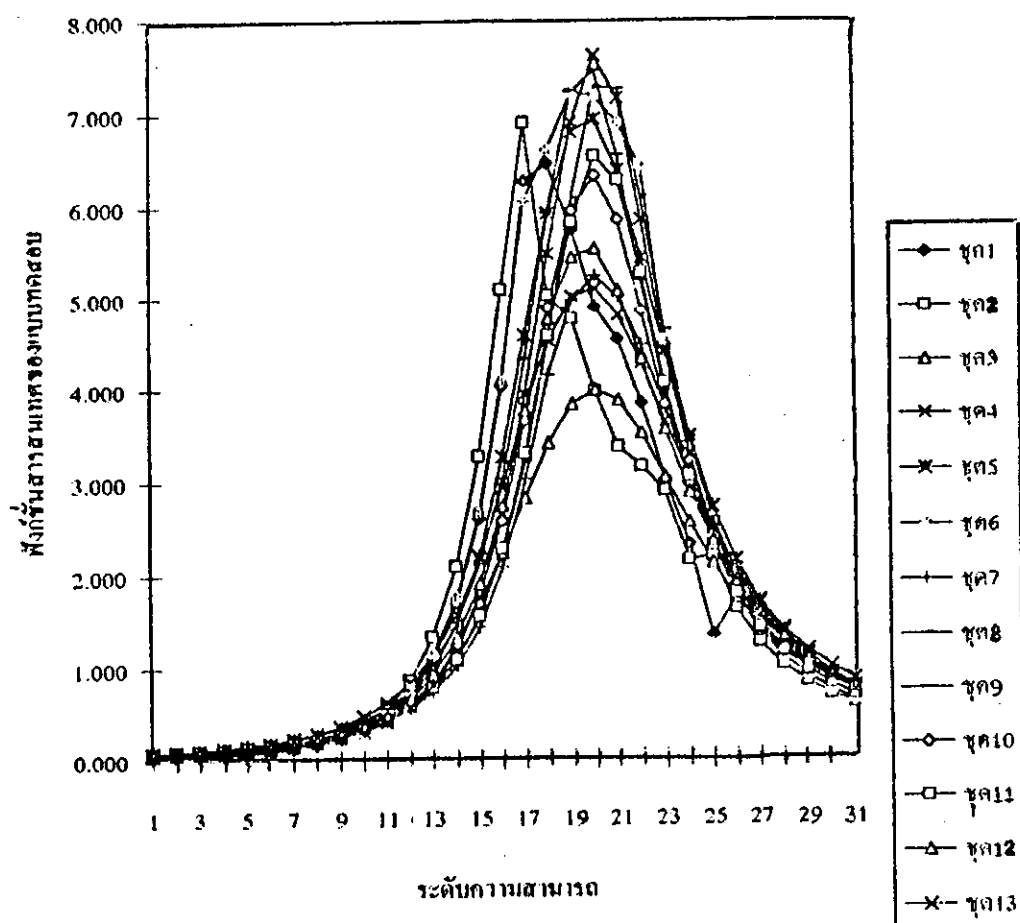


ภาพประกอบ 14 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่า พารามิเตอร์โดยวิธีอีวารีสดิค เมื่อข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

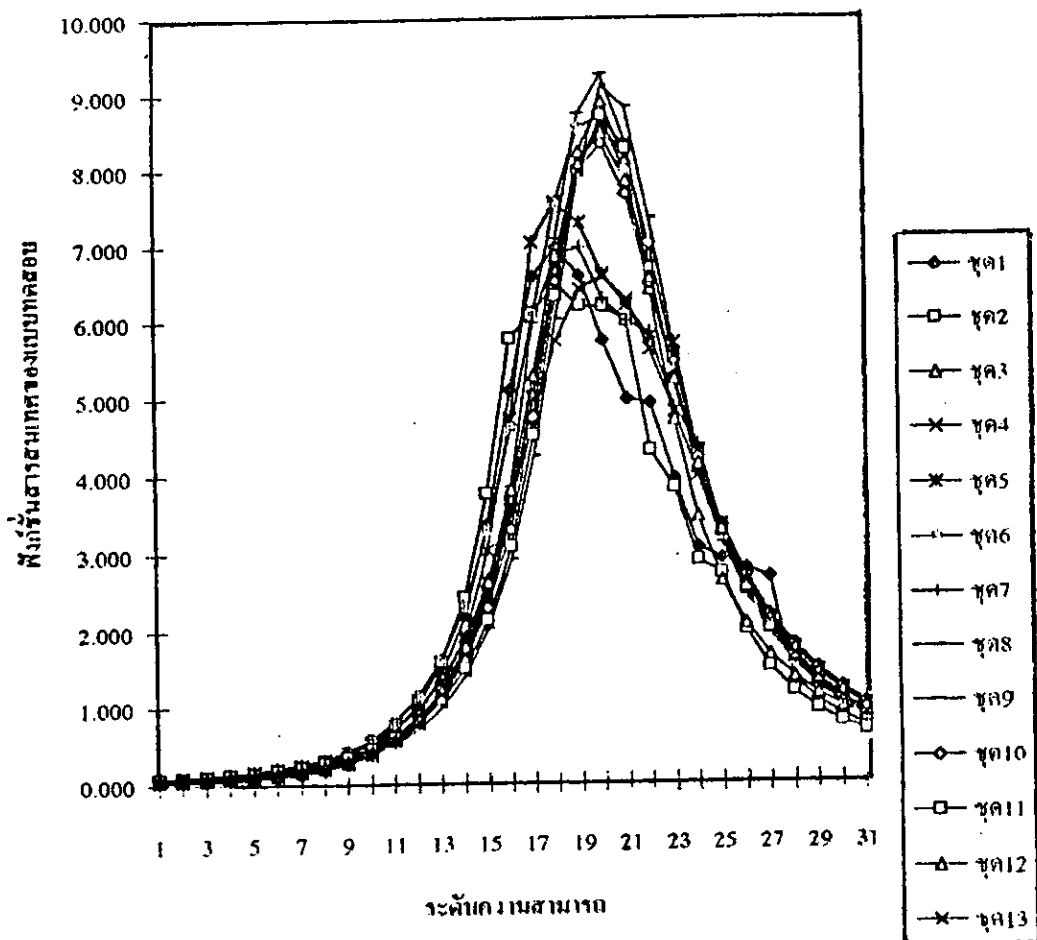


ภาพประกอบ 15 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน-ราสัน เมื่อข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

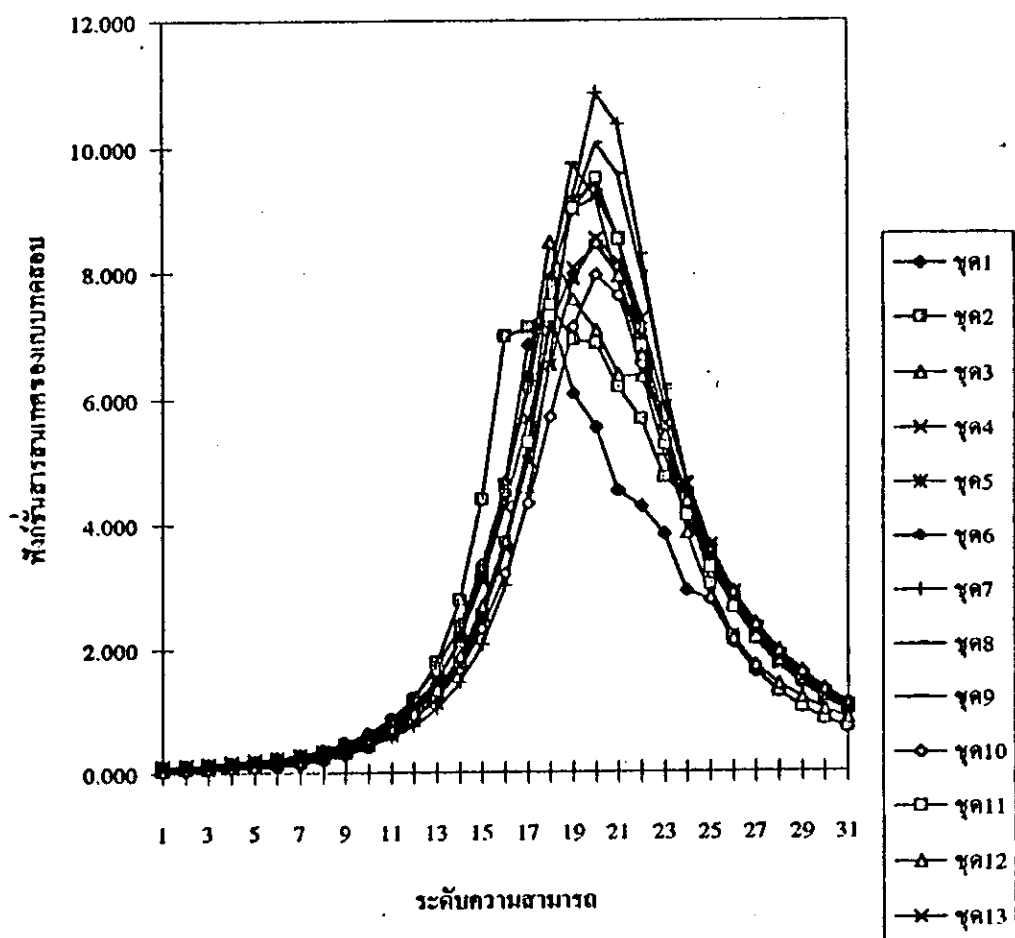


ภาพประกอบ 16 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

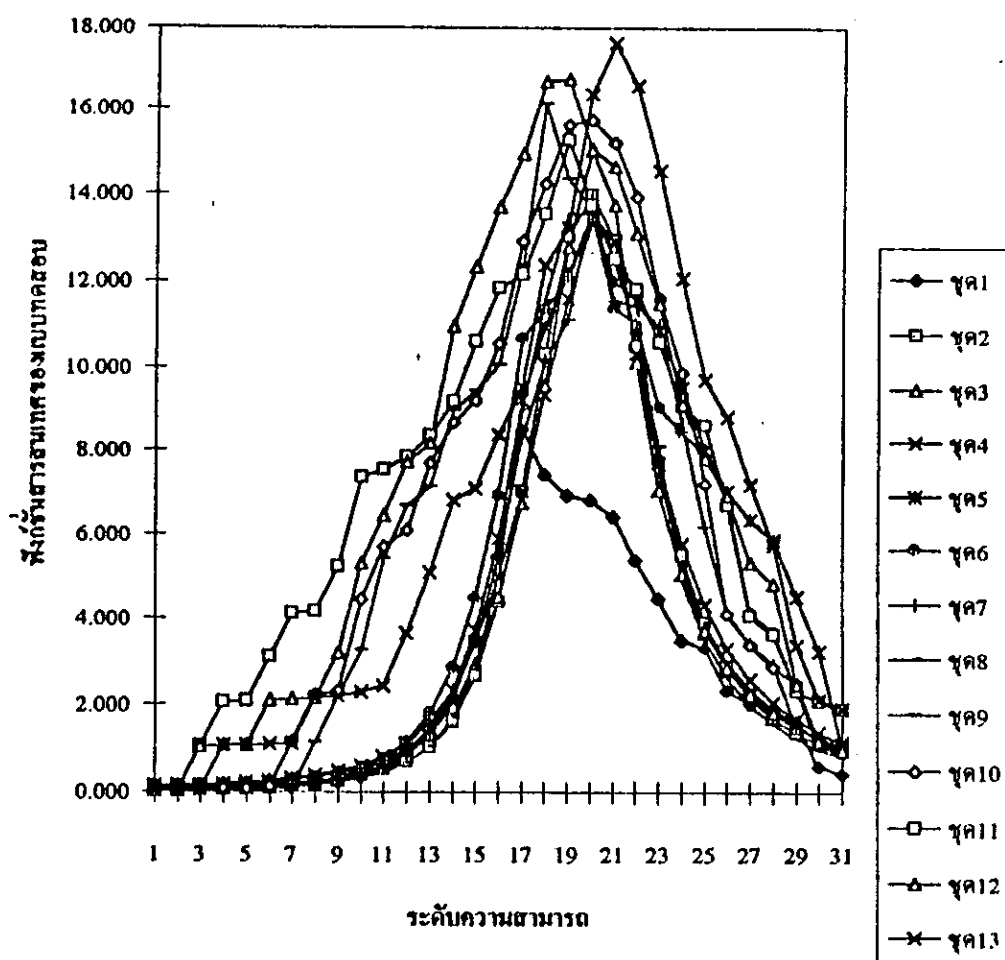


ภาพประกอบ 17 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่า พารามิเตอร์โดยวิธีนิวตัน-ราฟสัน เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน



ภาพประกอบ 18 เปรียบเทียบโค้งการกระจายของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

จากภาพประกอบ 13 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 50 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.147 ; e = 0.800\}$ เมื่อกลุ่มตัวอย่างมีขนาด 500 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 600 คน $\{I(e) = 7.272 ; e = 0.600\}$

จากภาพประกอบ 14 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 60 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.878 ; e = 0.800\}$ เมื่อกลุ่มตัวอย่างมีขนาด 500 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 400 คน $\{I(e) = 8.800 ; e = 0.600\}$

จากภาพประกอบ 15 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 70 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.985 ; e = 0.800\}$ เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน $\{I(e) = 7.619 ; e = 0.800\}$

จากภาพประกอบ 16 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 80 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 6.478 ; e = 0.400\}$ เมื่อกลุ่มตัวอย่างมีขนาด 300 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน $\{I(e) = 9.256 ; e = 0.800\}$

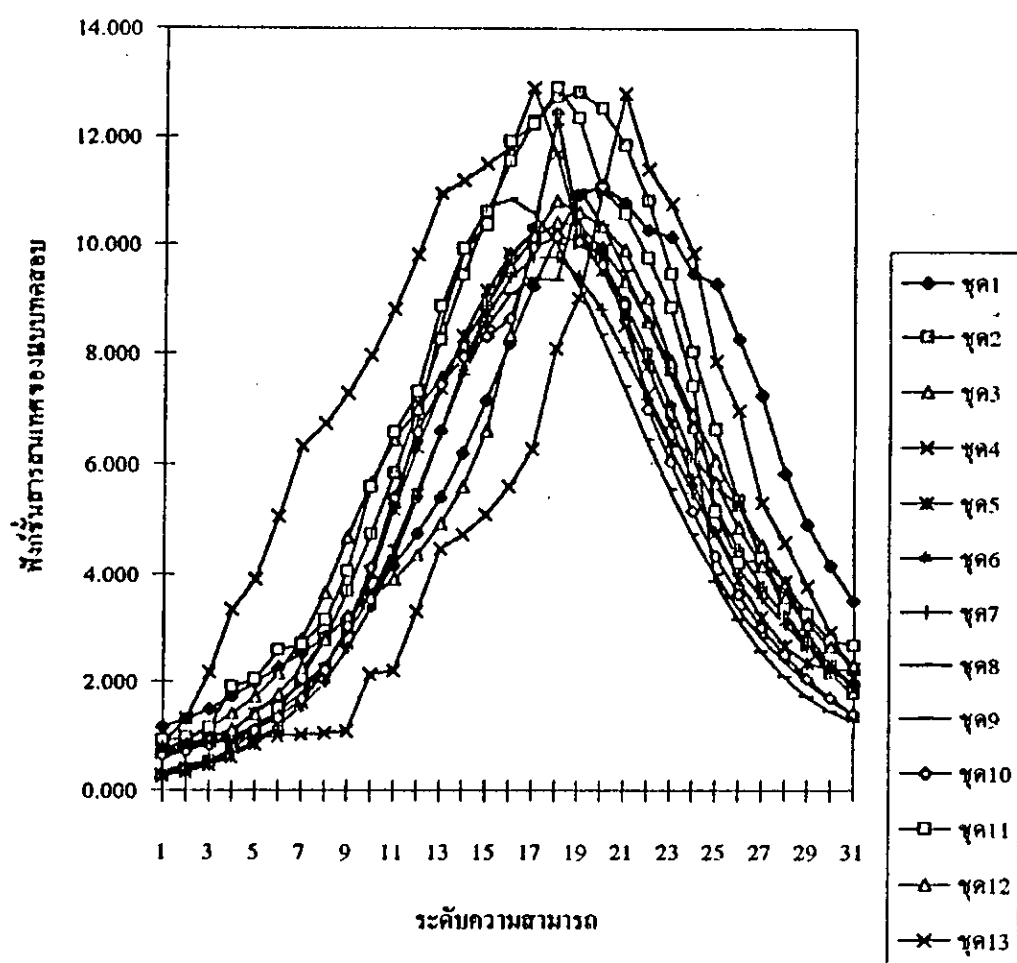
จากภาพประกอบ 17 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 90 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 7.401 ; e = 0.400\}$ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 800 คน $\{I(e) = 10.865 ; e = 0.800\}$

จากภาพประกอบ 18 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 100 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 8.514 ; e = 0.200\}$ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน $\{I(e) = 17.630 ;$

$$e = 1.000)$$

ดังนั้น จากภาพประกอบ 13 - 18 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอิมพิวติค มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.147; e = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ ขนาดของกลุ่มตัวอย่าง 500 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ ขนาดของกลุ่มตัวอย่าง 3,000 คน $\{I(e) = 17.630; e = 1.000\}$

5.3 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 19 - 24

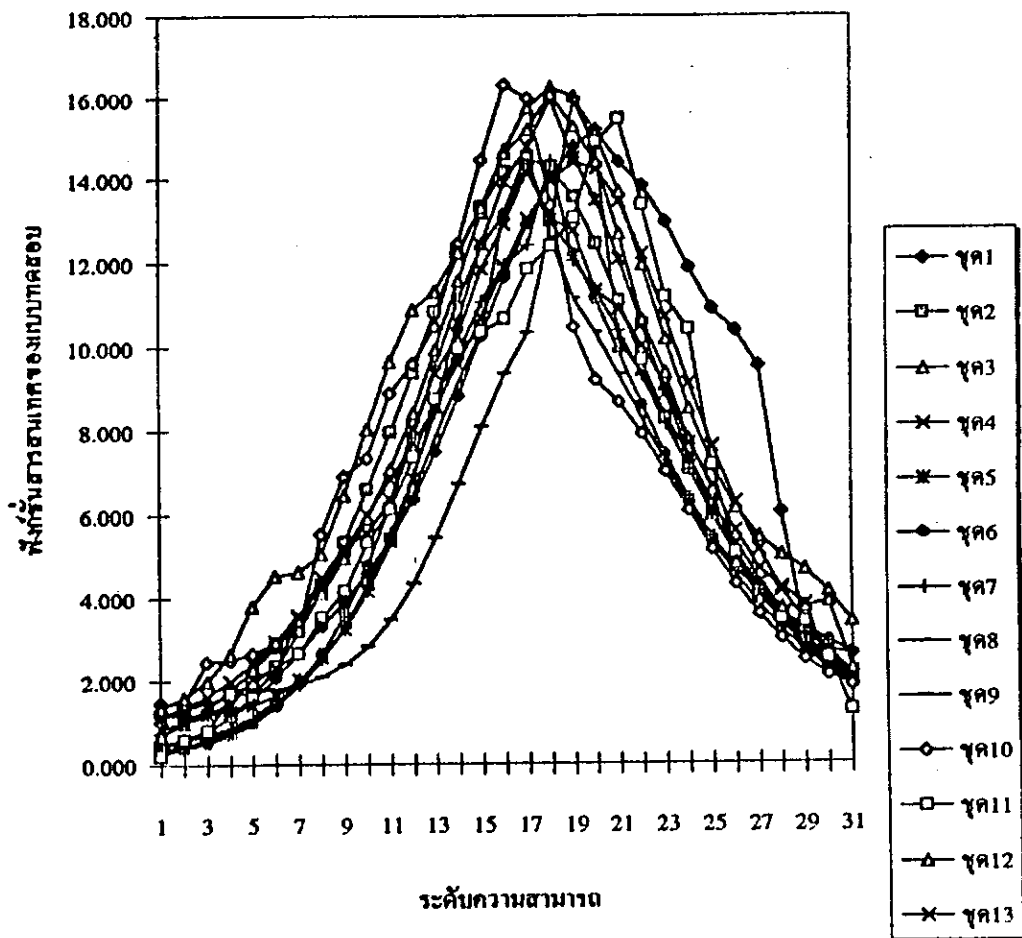


ภาพประกอบ 19 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โคซวีซีของเบส์ เมื่อข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

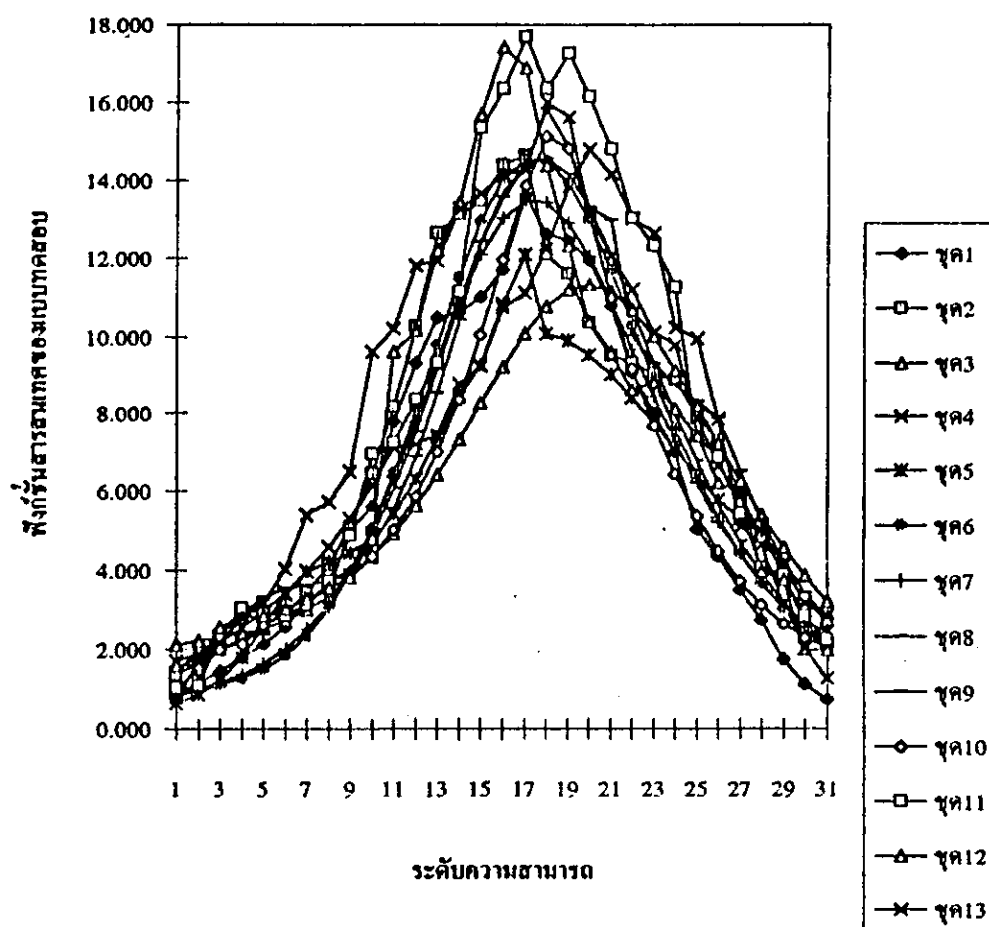


ภาพประกอบ 20 เปรียบเทียบโค้งสำรวจสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบย์ เมื่อข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

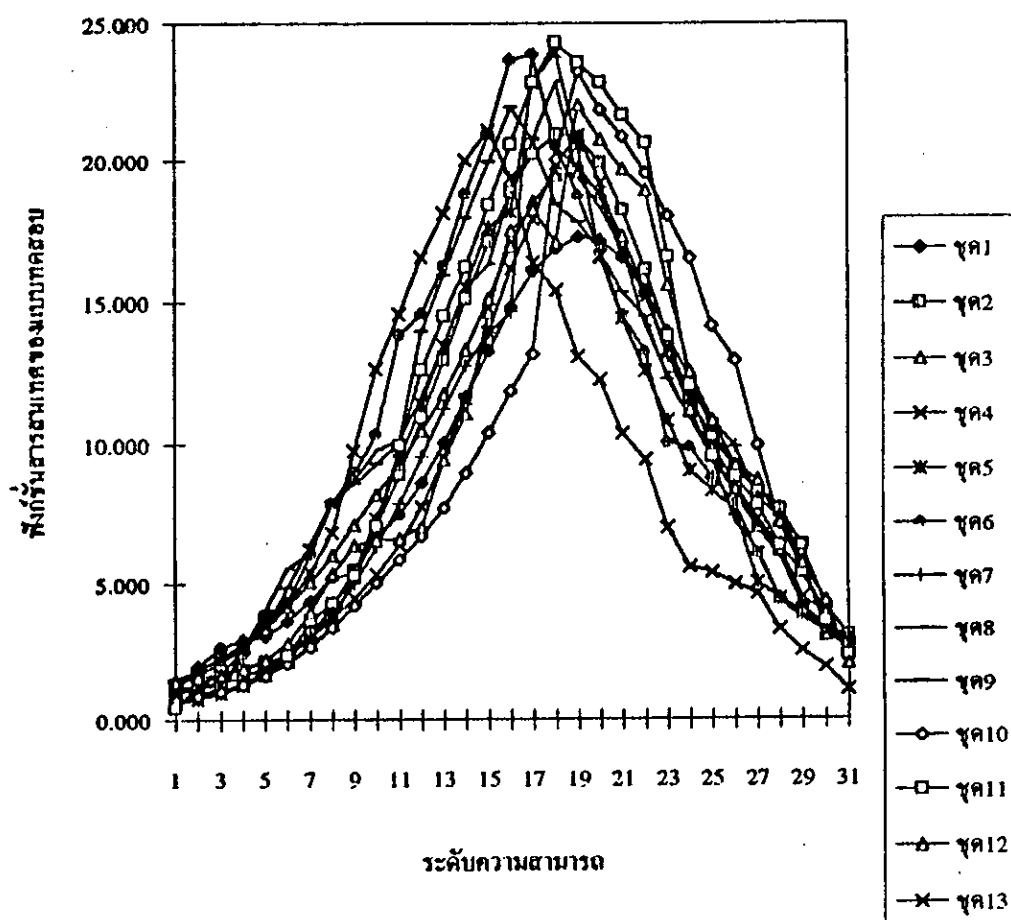


ภาพประกอบ 21 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

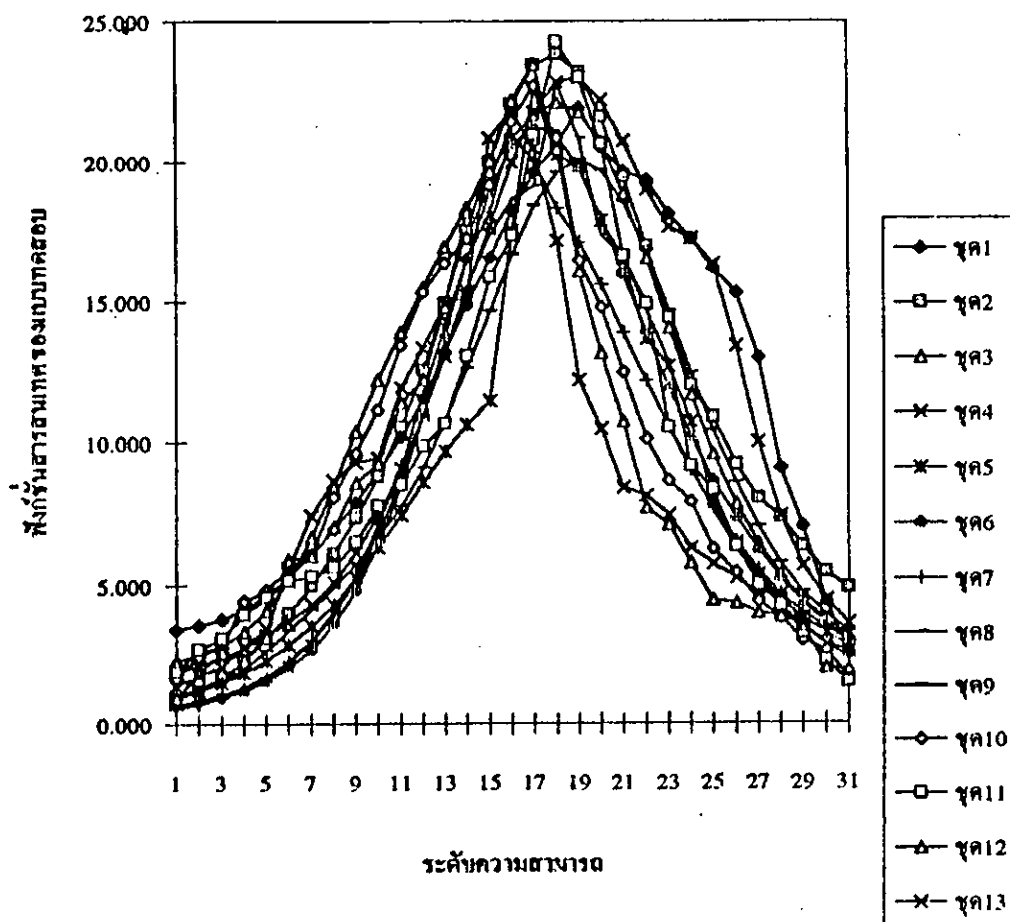


ภาพประกอบ 22 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่า พหาวมิเตอร์โดวี่ของเบส์ เมื่อข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

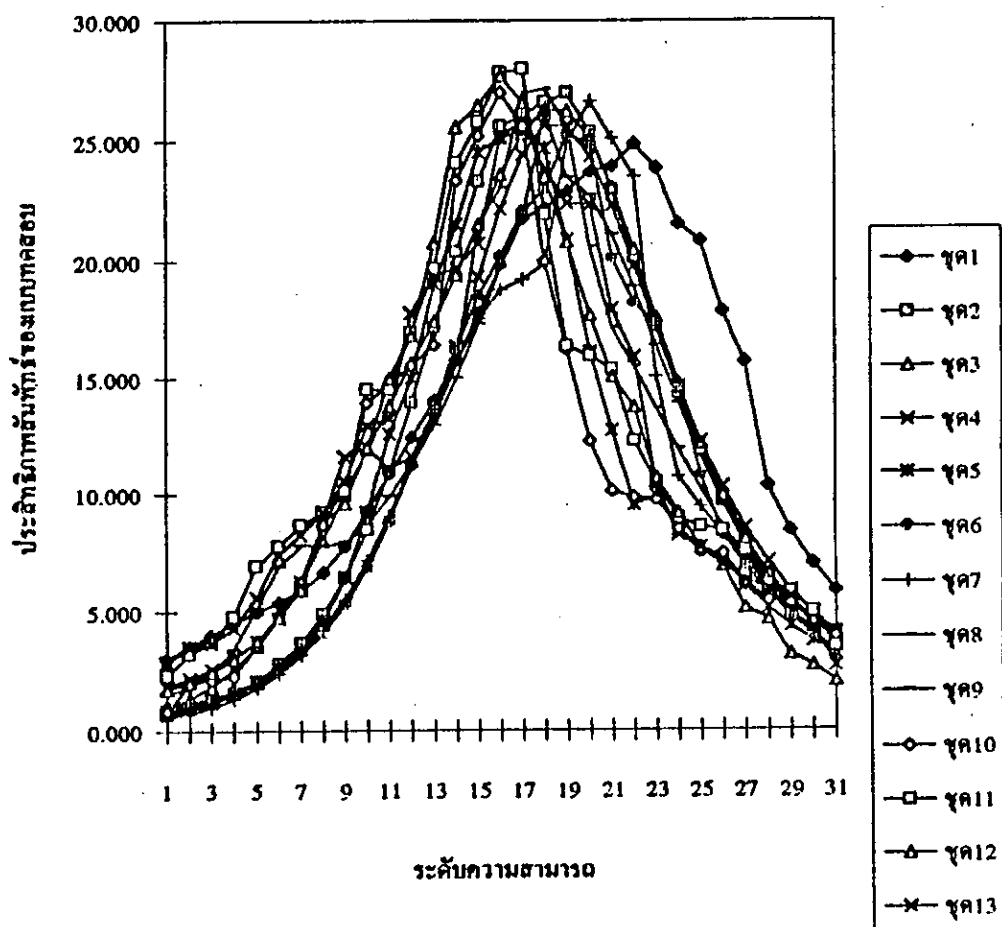


ภาพประกอบ 23 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
 ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
 ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
 ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
 ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
 ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
 ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
 ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
 ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
 ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
 ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
 ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน



ภาพประกอบ 24 เปรียบเทียบโค้งสำรสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่า
พารามิเตอร์โคชีวิชีของเบส์ เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน
ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวนกลุ่มตัวอย่าง 200 คน
ชุด 2 จำนวนกลุ่มตัวอย่าง 300 คน
ชุด 3 จำนวนกลุ่มตัวอย่าง 400 คน
ชุด 4 จำนวนกลุ่มตัวอย่าง 500 คน
ชุด 5 จำนวนกลุ่มตัวอย่าง 600 คน
ชุด 6 จำนวนกลุ่มตัวอย่าง 700 คน
ชุด 7 จำนวนกลุ่มตัวอย่าง 800 คน

ชุด 8 จำนวนกลุ่มตัวอย่าง 900 คน
ชุด 9 จำนวนกลุ่มตัวอย่าง 1,000 คน
ชุด 10 จำนวนกลุ่มตัวอย่าง 1,500 คน
ชุด 11 จำนวนกลุ่มตัวอย่าง 2,000 คน
ชุด 12 จำนวนกลุ่มตัวอย่าง 2,500 คน
ชุด 13 จำนวนกลุ่มตัวอย่าง 3,000 คน

จากภาพประกอบ 19 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 50 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 9.766 ; e = 0.200\}$ เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน $\{I(e) = 12.917 ; e = 0.400\}$

จากภาพประกอบ 20 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 60 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 14.164 ; e = 0.200\}$ เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน $\{I(e) = 16.310 ; e = 0.000\}$

จากภาพประกอบ 21 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 70 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 11.318 ; e = 0.800\}$ เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน $\{I(e) = 17.658 ; e = 0.200\}$

จากภาพประกอบ 22 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 80 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 17.266 ; e = 0.600\}$ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน $\{I(e) = 24.281 ; e = 0.400\}$

จากภาพประกอบ 23 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 90 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 20.091 ; e = 0.600\}$ เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และมีค่าสูงสุดที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน $\{I(e) = 24.251 ; e = 0.400\}$

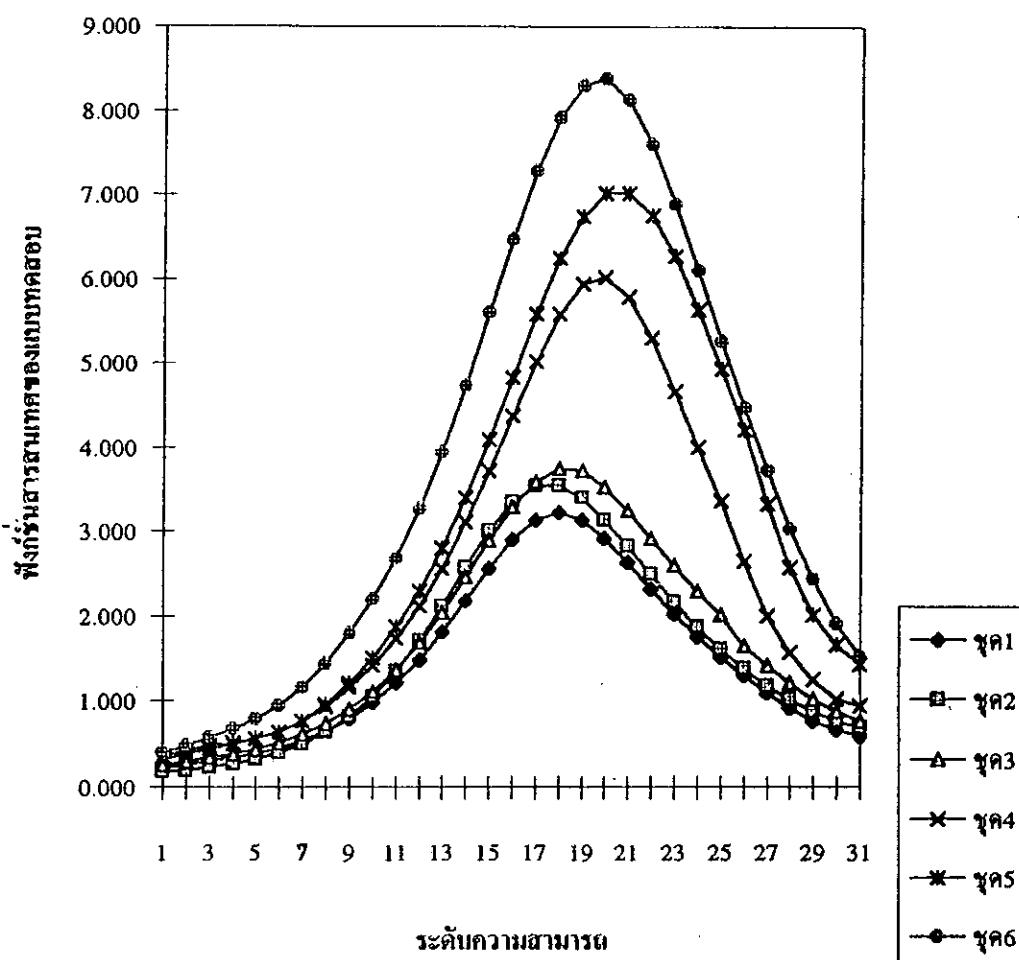
จากภาพประกอบ 24 พบว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบ 100 ข้อ มีค่าต่ำสุด ที่ระดับความสามารถสูง $\{I(e) = 24.166 ; e = 1.200\}$ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน $\{I(e) = 27.954 ;$

$$\alpha = 0.200\}$$

ดังนั้น จากภาพประกอบ 19 - 24 แสดงว่า เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ มีค่าต่ำสุด ที่ระดับความสามารถต่ำ $\{I(\alpha) = 9.766 ; \alpha = 0.200\}$ เมื่อข้อสอบมีขนาด 50 ข้อ ขนาดของกลุ่มตัวอย่าง 800 คน และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ ขนาดของกลุ่มตัวอย่าง 2,000 คน $\{I(\alpha) = 27.954 ; \alpha = 0.200\}$

6. ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีอาร์เอสทีค และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ

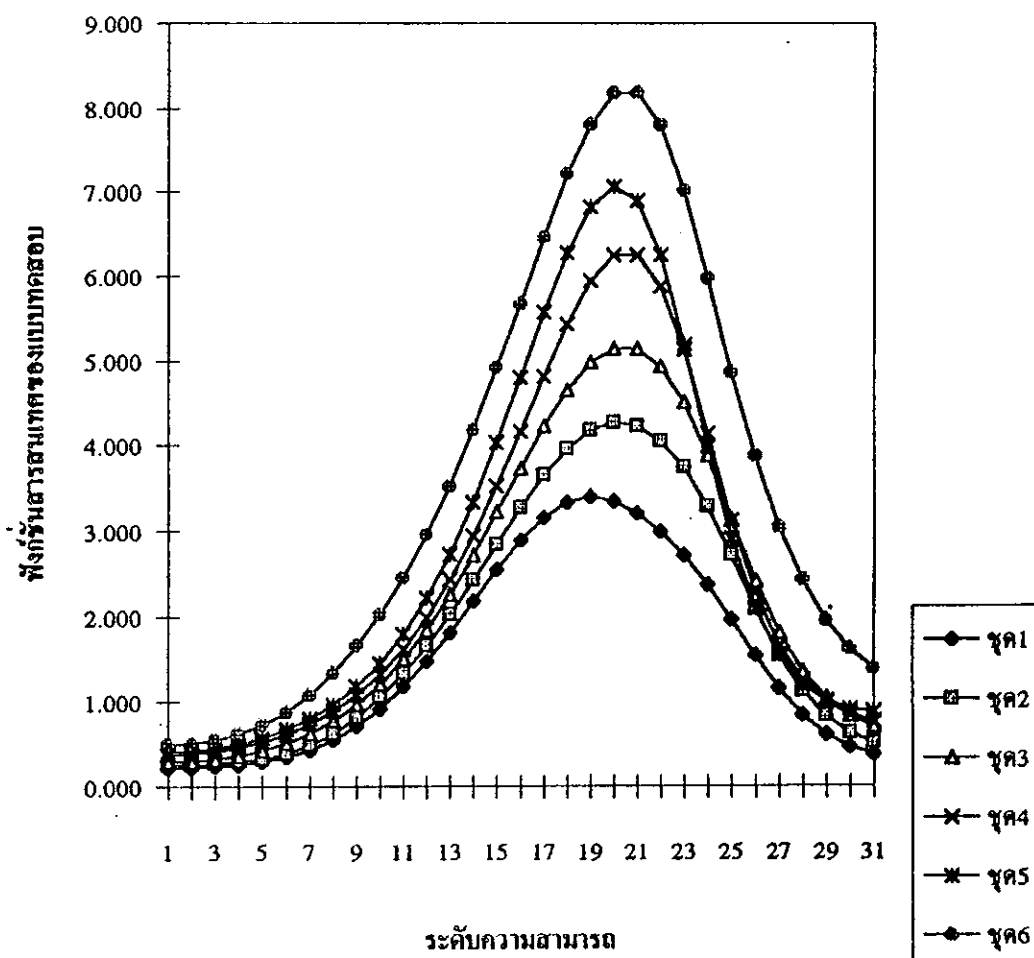
6.1 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 25 - 37



ภาพประกอบ 25 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

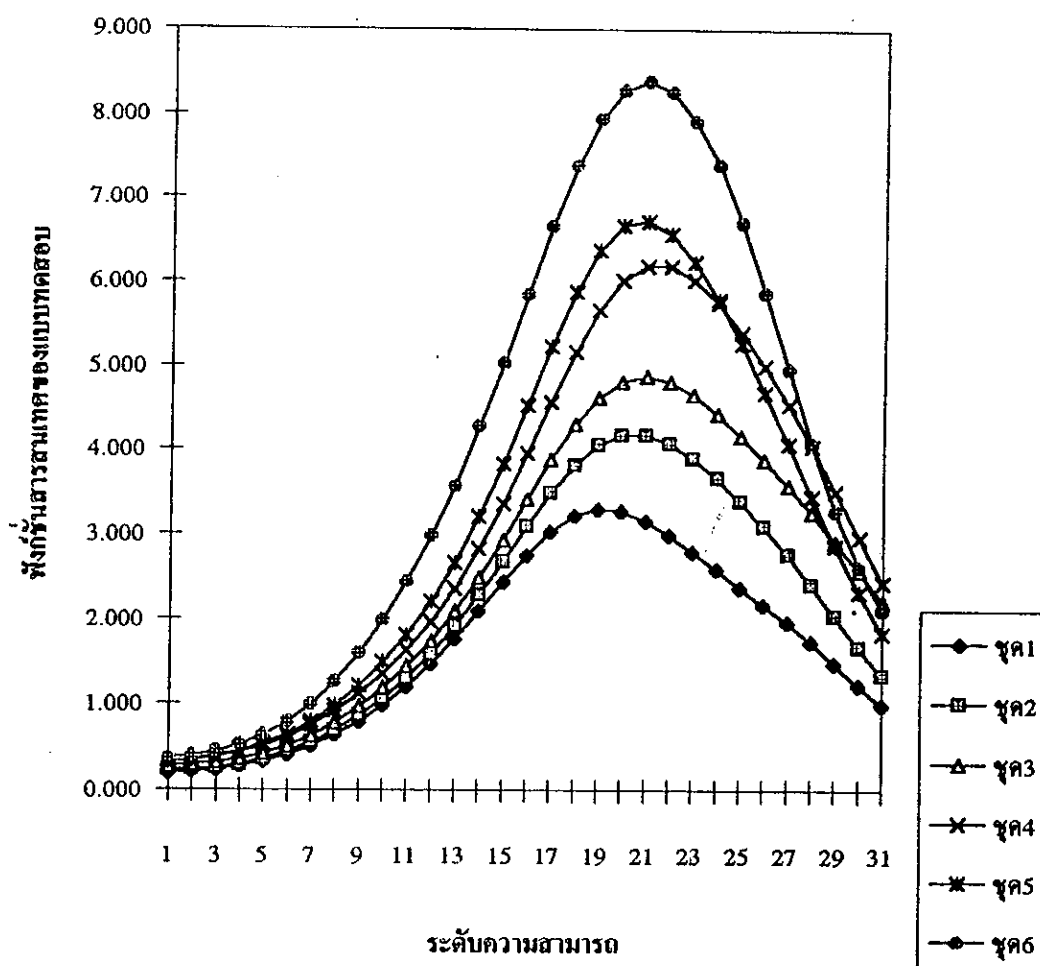
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 26 เปรียบเทียบโค้งการตอบสนองของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 300 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

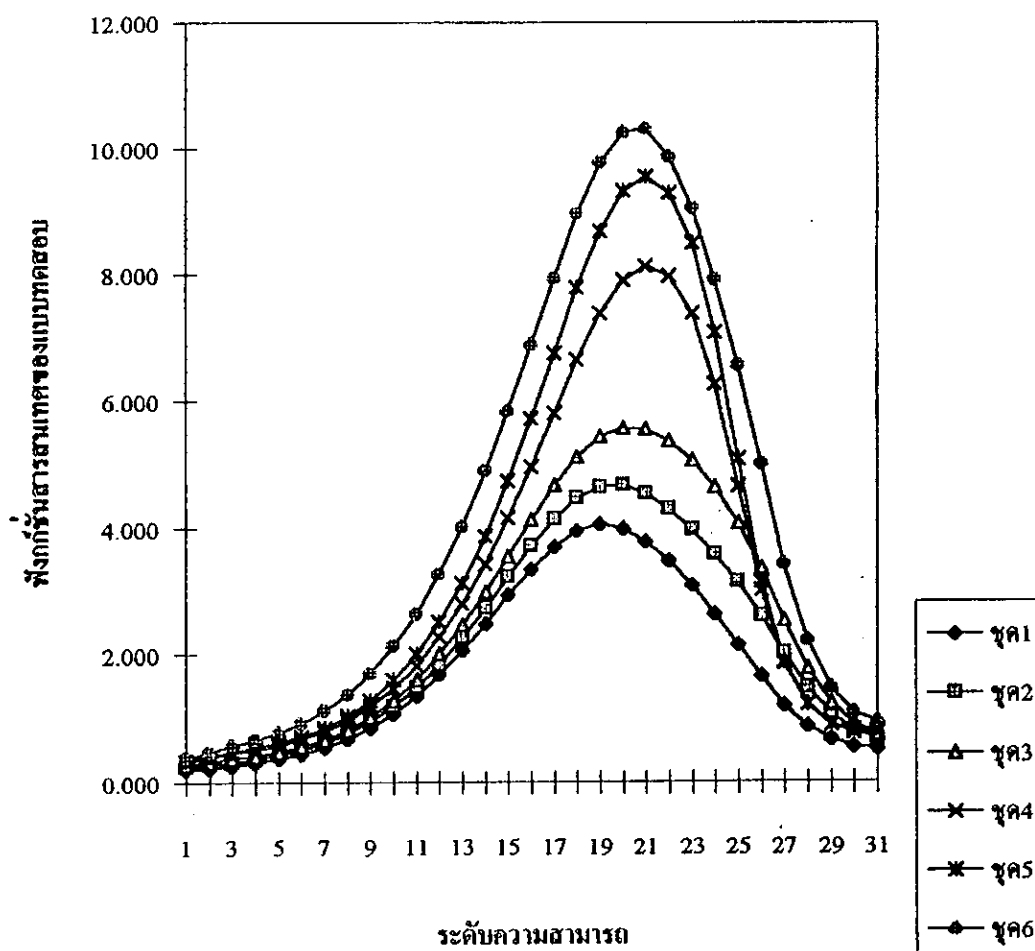
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
ชุด 2 จำนวนข้อสอบ 60 ข้อ
ชุด 3 จำนวนข้อสอบ 70 ข้อ
ชุด 4 จำนวนข้อสอบ 80 ข้อ
ชุด 5 จำนวนข้อสอบ 90 ข้อ
ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 27 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

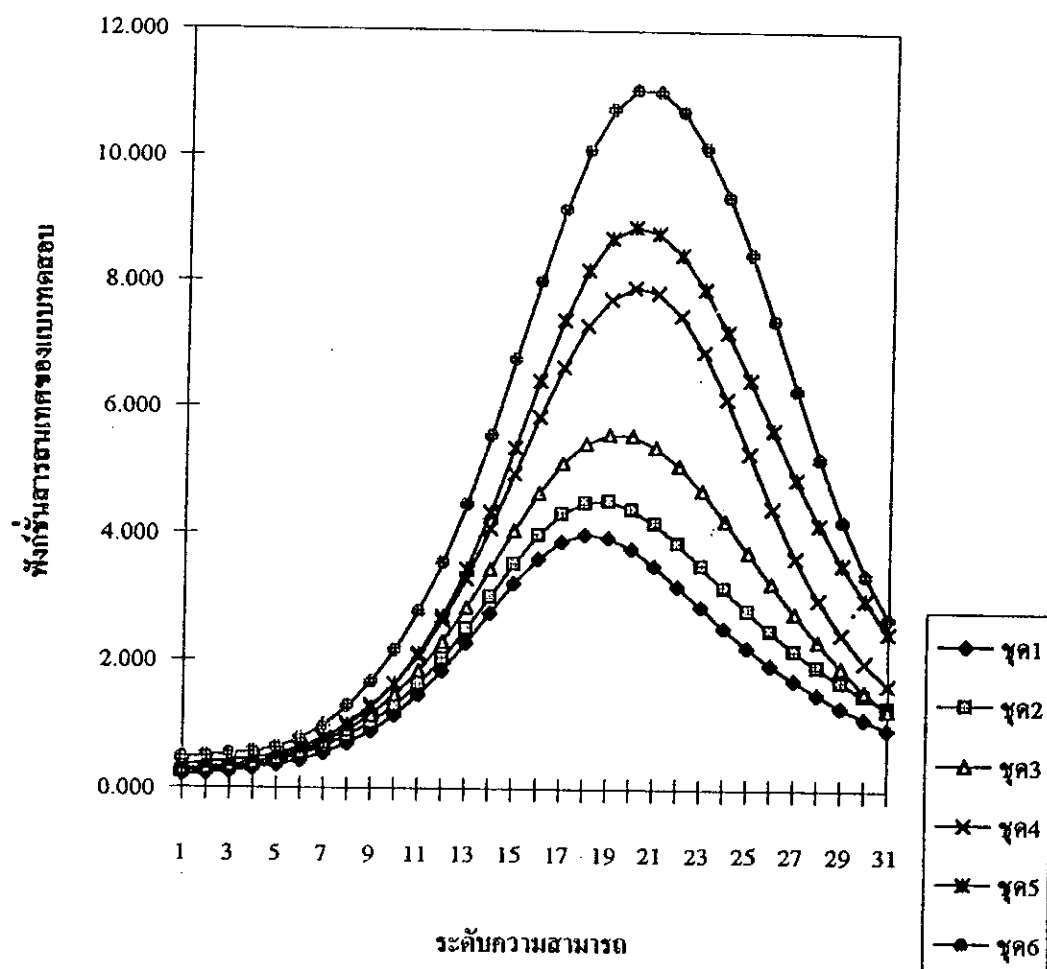
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
 ชุด 2 จำนวนข้อสอบ 60 ข้อ
 ชุด 3 จำนวนข้อสอบ 70 ข้อ
 ชุด 4 จำนวนข้อสอบ 80 ข้อ
 ชุด 5 จำนวนข้อสอบ 90 ข้อ
 ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 28 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 500 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

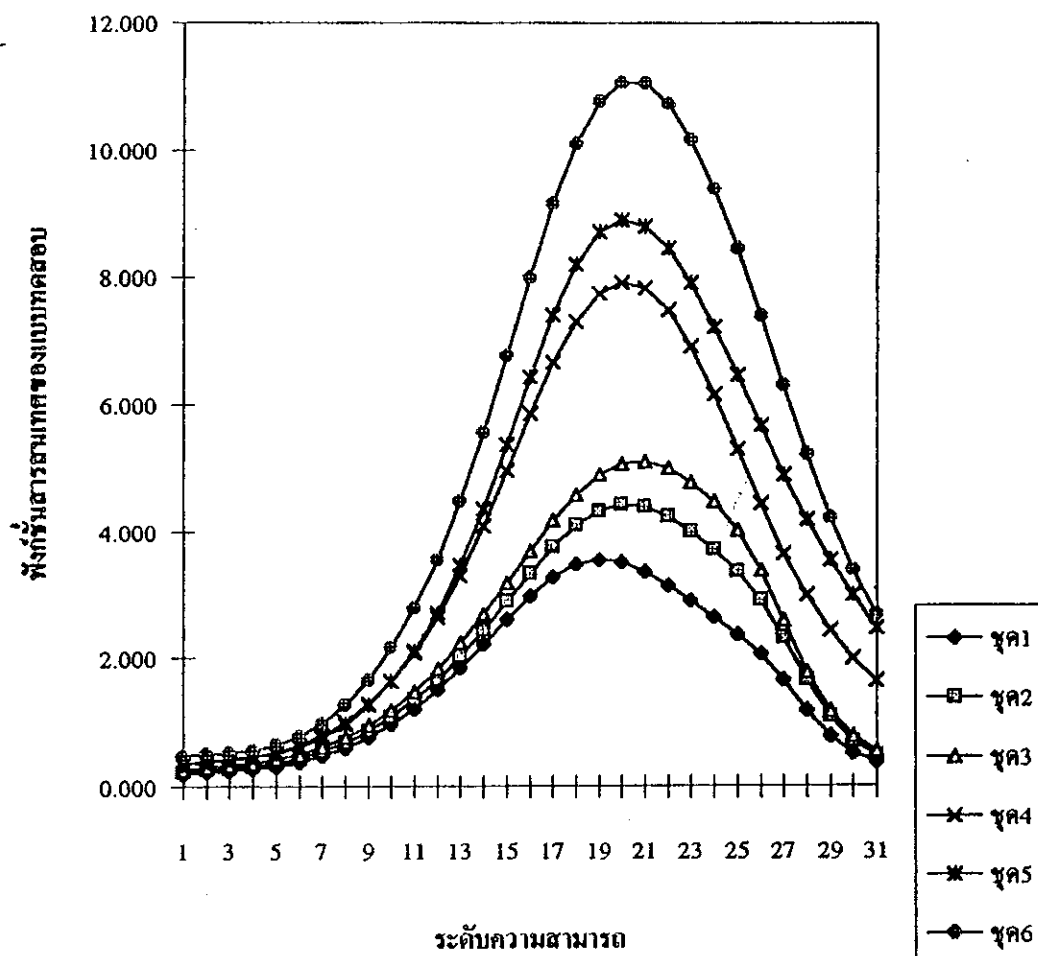
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 29 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแบบถ่วงน้ำหนัก โดยเฉลี่ย เมื่อกลุ่มตัวอย่างมีขนาด 600 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

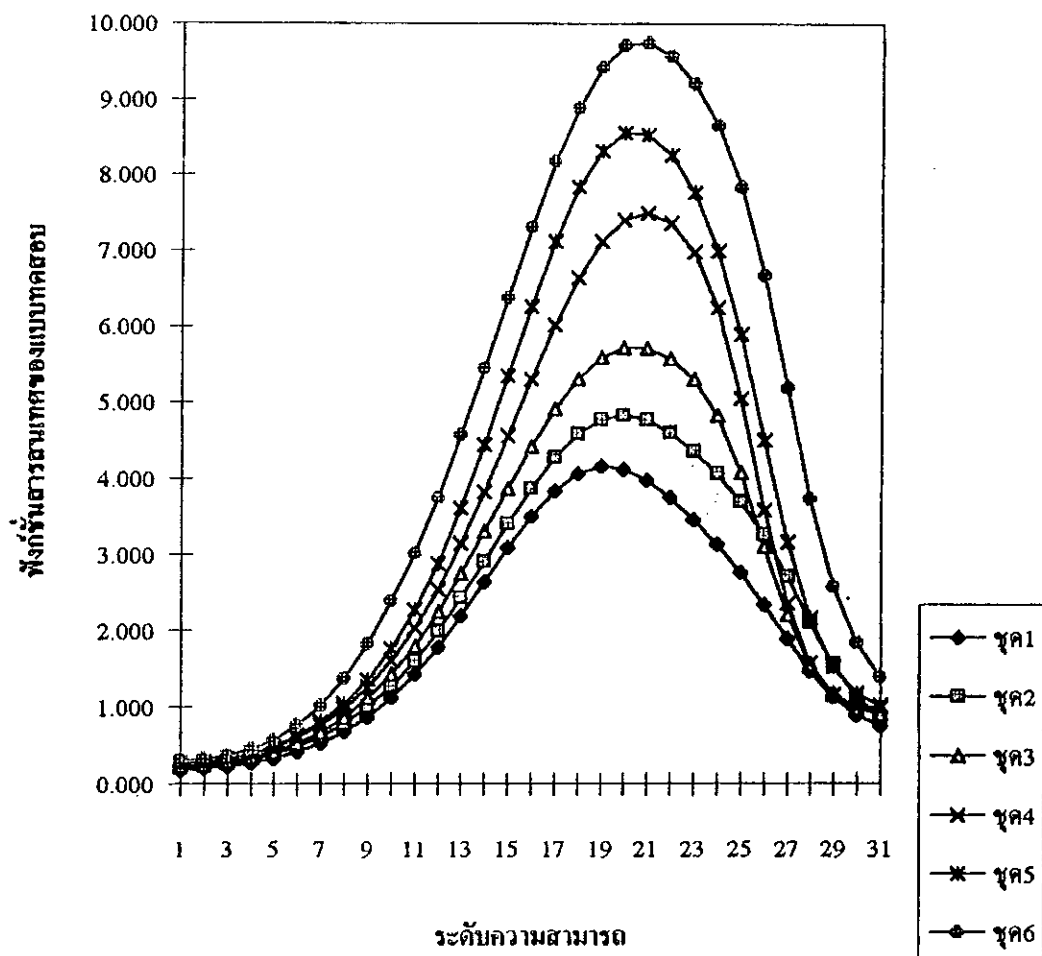
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 30 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิตูด เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

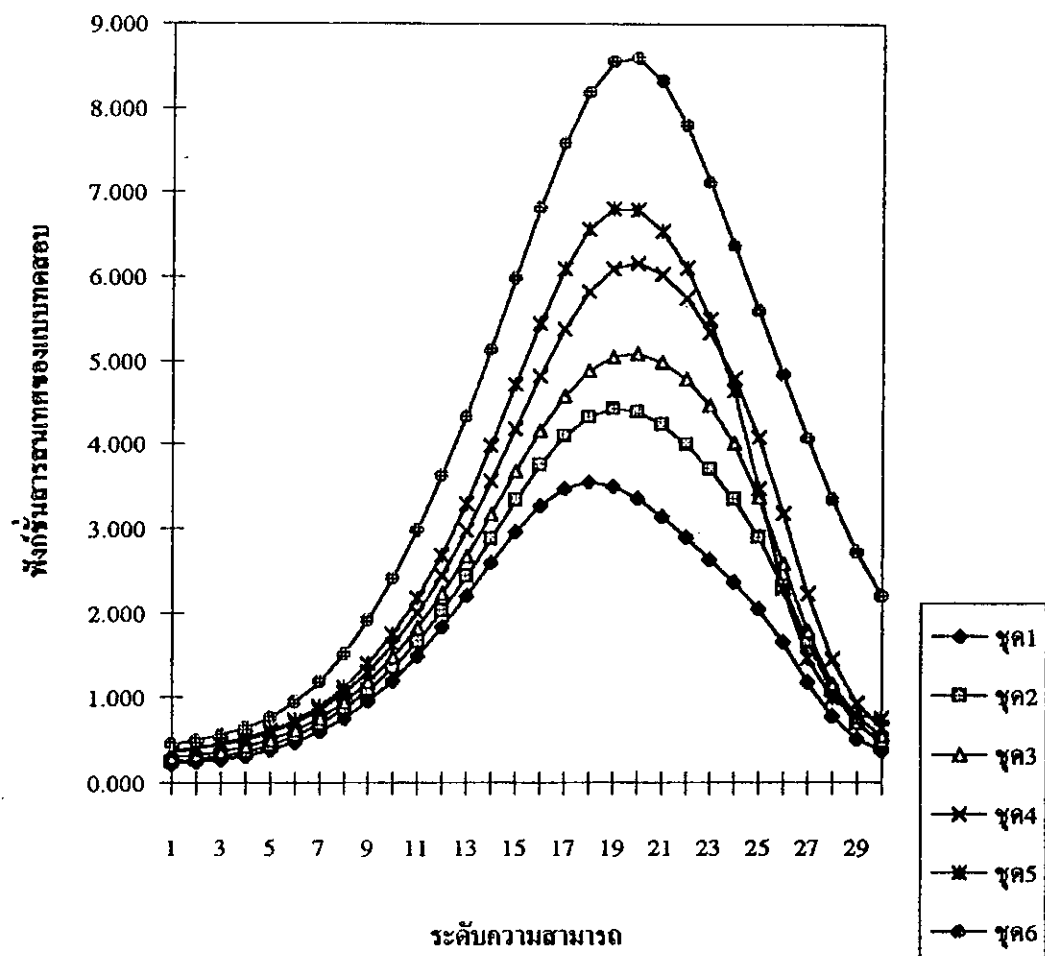
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 31 เปรียบเทียบโค้งการตอบสนองของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซ์ิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

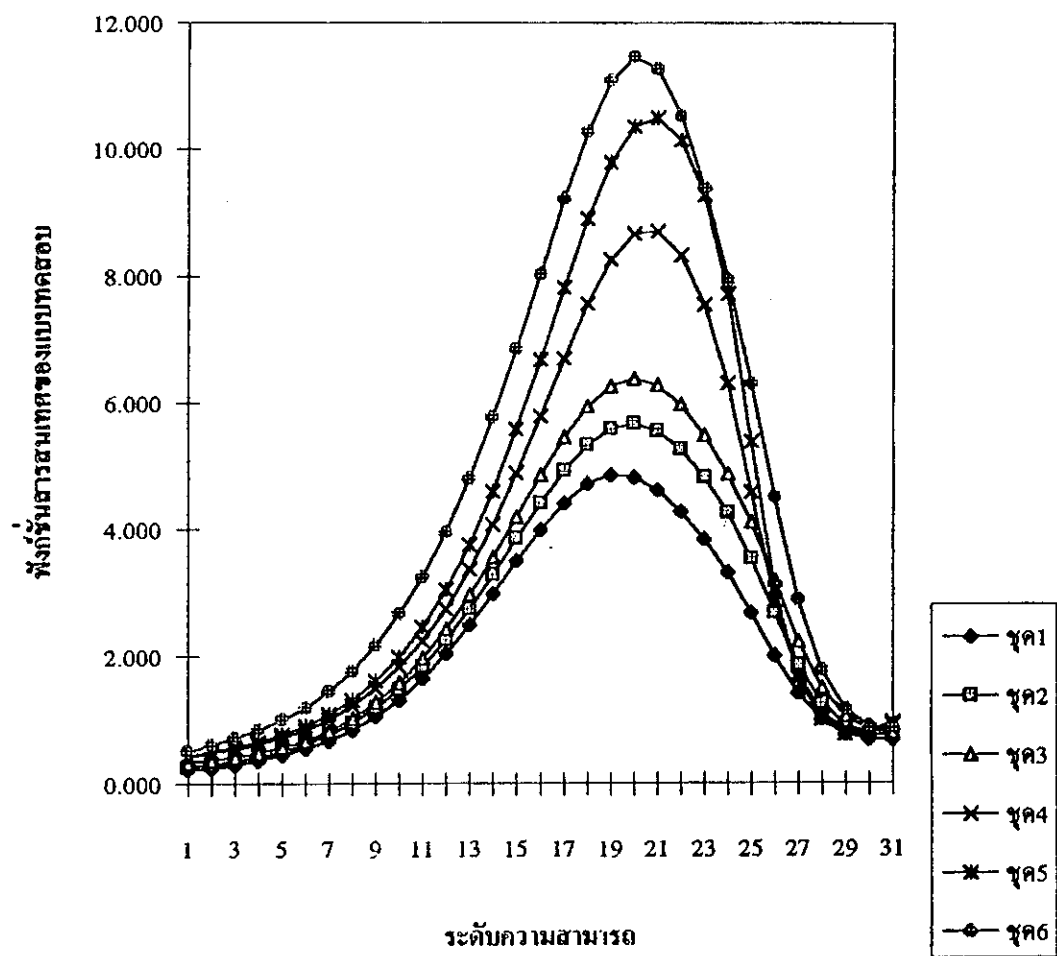
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 32 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

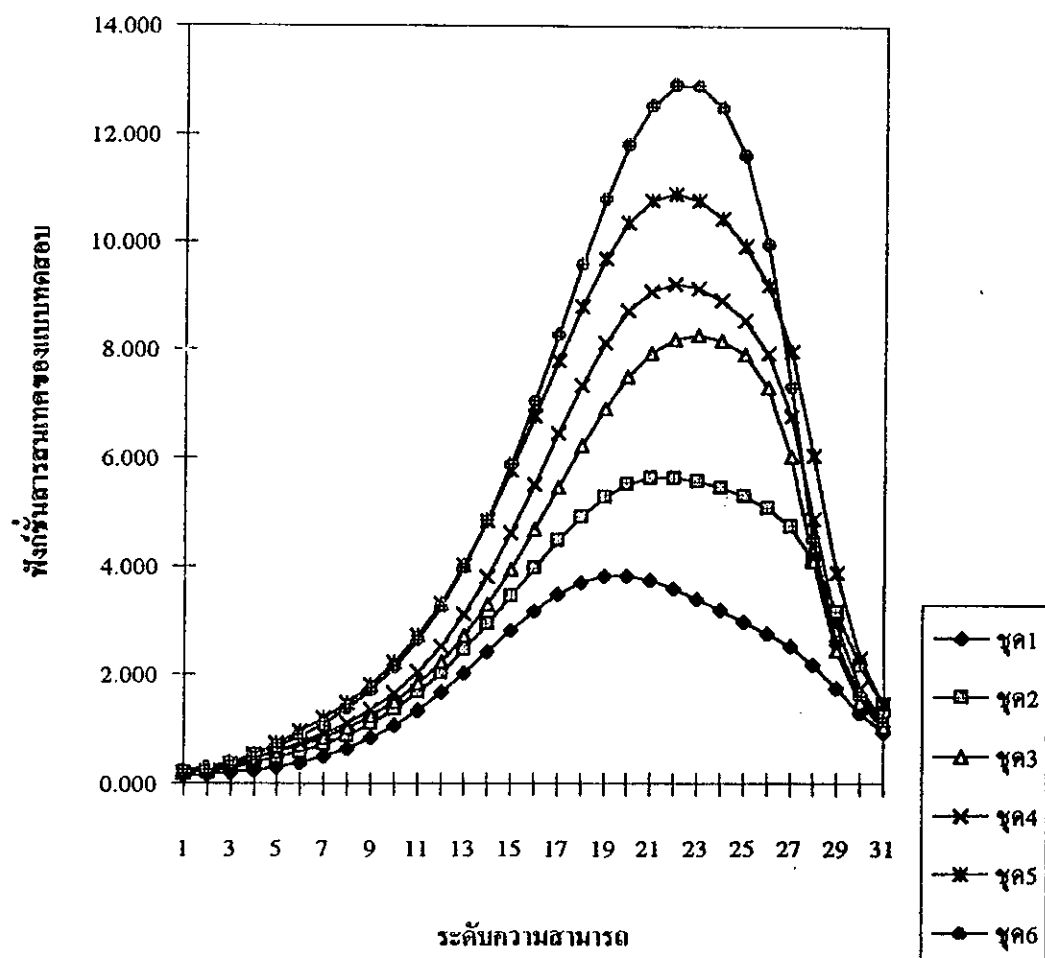
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 33 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โคลด์ลิสต์ เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาดต่างกัน ๘ ระดับความสามารถต่าง ๆ

หมายเหตุ

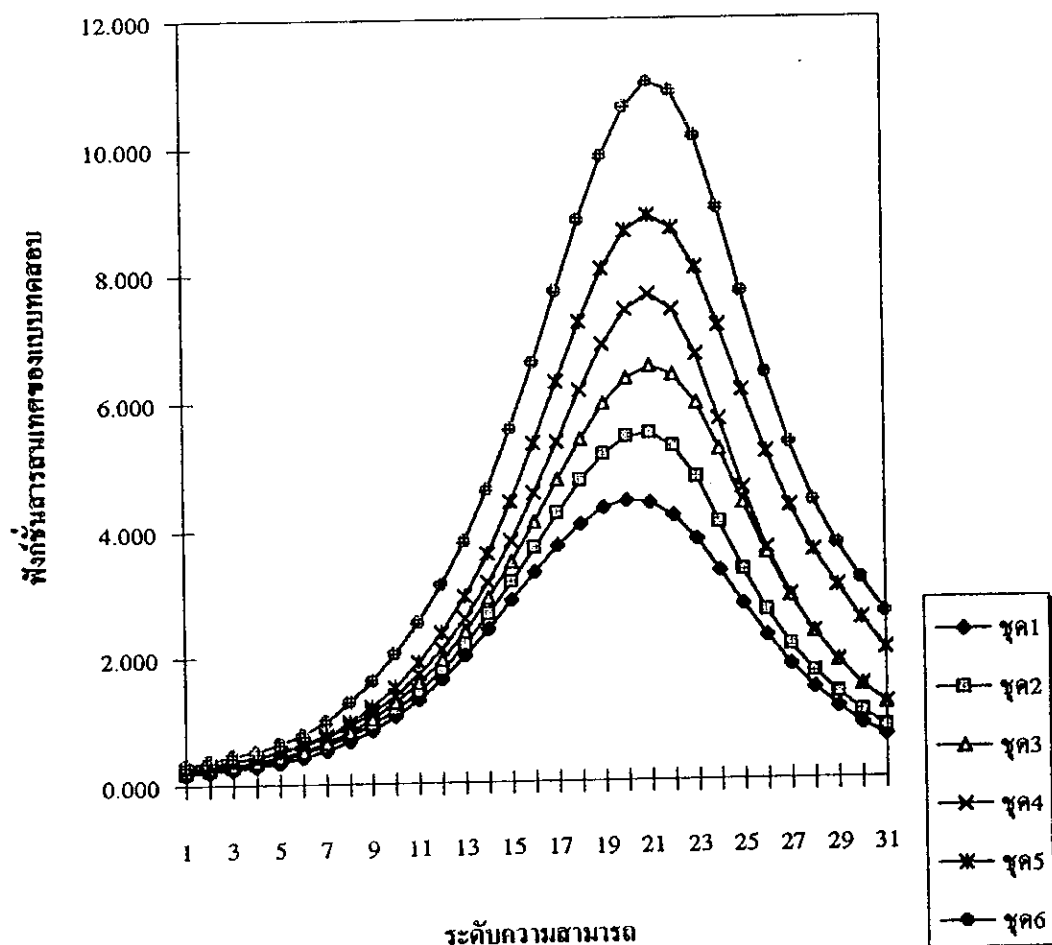
ชุด 1	จำนวนข้อสอบ	50 ข้อ
ชุด 2	จำนวนข้อสอบ	60 ข้อ
ชุด 3	จำนวนข้อสอบ	70 ข้อ
ชุด 4	จำนวนข้อสอบ	80 ข้อ
ชุด 5	จำนวนข้อสอบ	90 ข้อ
ชุด 6	จำนวนข้อสอบ	100 ข้อ



ภาพประกอบ 34 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน และข้อสอบมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

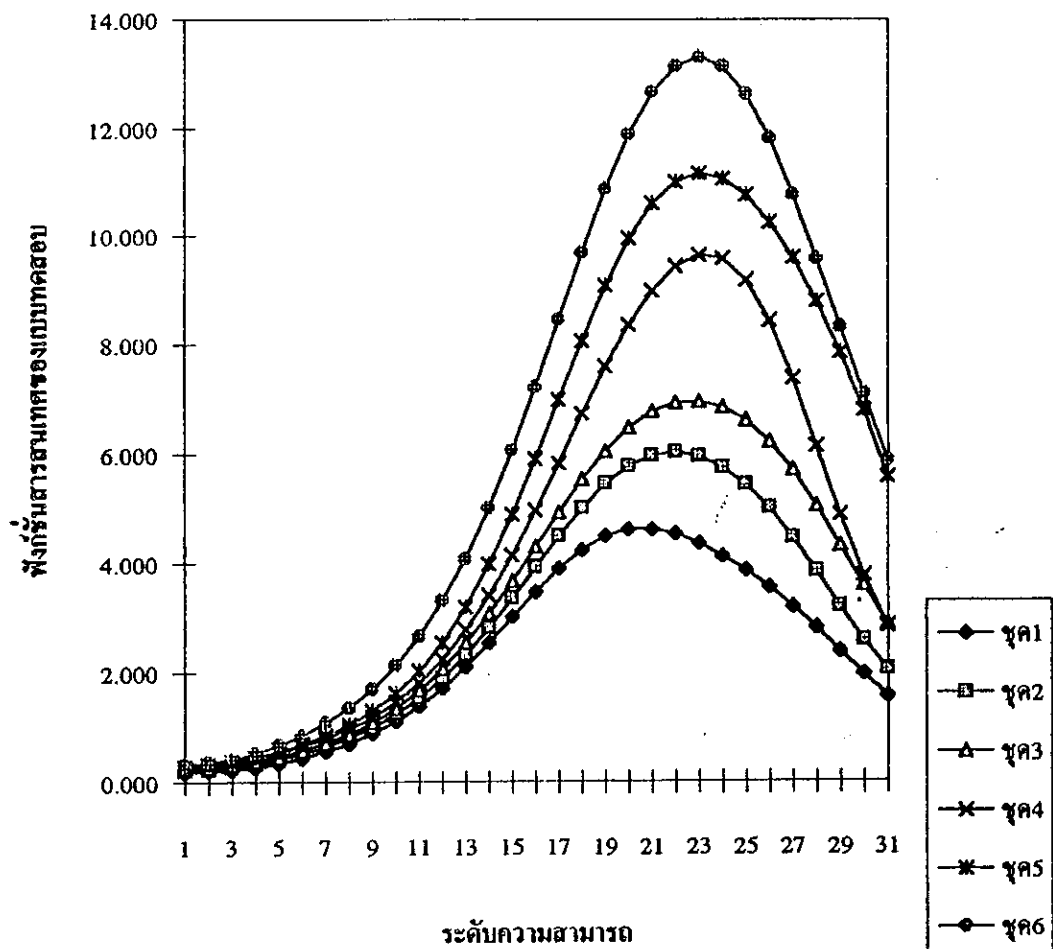
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 35 เปรียบเทียบโค้งการตอบสนองของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลไลซ์ เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

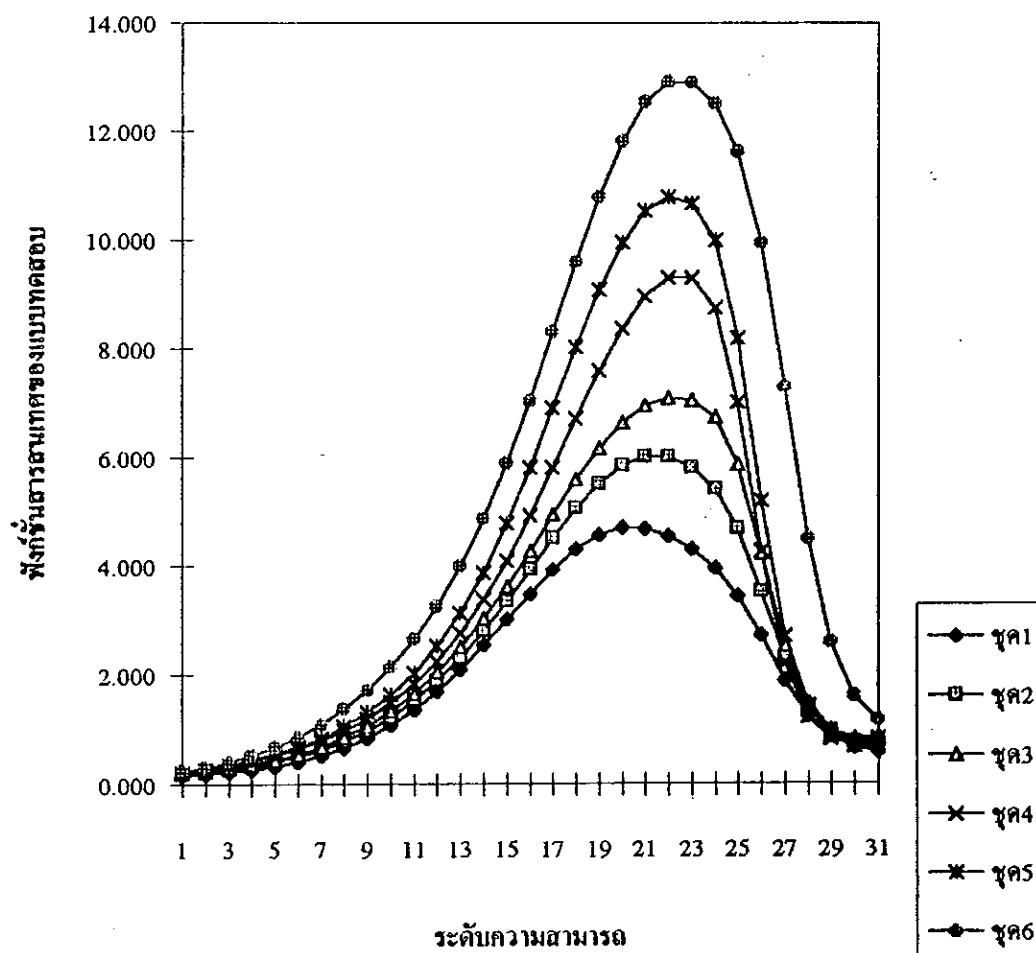
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 36 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน และข้อสอบมีขนาดต่างกัน ๕ ระดับความสามารถต่าง ๆ

หมายเหตุ

- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 37 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซ์ิมัม โลคัลลิสต์ เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ

จากภาพประกอบ 25 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ไลกelihood ของกลุ่มตัวอย่างขนาด 200 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.136 ; e = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 8.370 ; e = 0.800\}$

จากภาพประกอบ 26 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ลิเกอิต ของกลุ่มตัวอย่างขนาด 300 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 3.393 ; \theta = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 8.189 ; \theta = 1.000\}$

จากภาพประกอบ 27 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ไลค์ลียูด ของกลุ่มตัวอย่างขนาด 400 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 3.296 ; \theta = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 8.383 ; \theta = 1.000\}$

จากภาพประกอบ 28 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ลิคูด ของกลุ่มตัวอย่างขนาด 500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 4.045 ; \theta = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 10.306 ; \theta = 1.000\}$

จากภาพประกอบ 29 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ไลกelihood ของกลุ่มตัวอย่างขนาด 600 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 3.997 ; \theta = 0.400\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 11.064 ; \theta = 0.800\}$

จากภาพประกอบ 30 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ไลค์ลิต ของกลุ่มตัวอย่างขนาด 700 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.547 ; e = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับ

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 11.064 ; e = 0.800\}$

จากภาพประกอบ 31 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ลิคิลิต์ของกลุ่มตัวอย่างขนาด 800 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 4.149 ; \theta = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 9.743 ; \theta = 1.000\}$

จากภาพประกอบ 32 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ลิคelihood ของกลุ่มตัวอย่างขนาด 900 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 3.547 ; \theta = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 8.581 ; \theta = 1.000\}$

จากภาพประกอบ 33 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ลิเกอิต ของกลุ่มตัวอย่างขนาด 1,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 4.853 ; \theta = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 11.457 ; \theta = 0.800\}$

จากภาพประกอบ 34 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซ์ิมไลค์ลิสต์ ของกลุ่มตัวอย่างขนาด 1,500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 3.822 ; \theta = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 12.911 ; \theta = 1.200\}$

จากภาพประกอบ 35 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมไลซ์ลิคelihood ของกลุ่มตัวอย่างขนาด 2,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 4.422 ; \theta = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 10.990 ; \theta = 1.000\}$

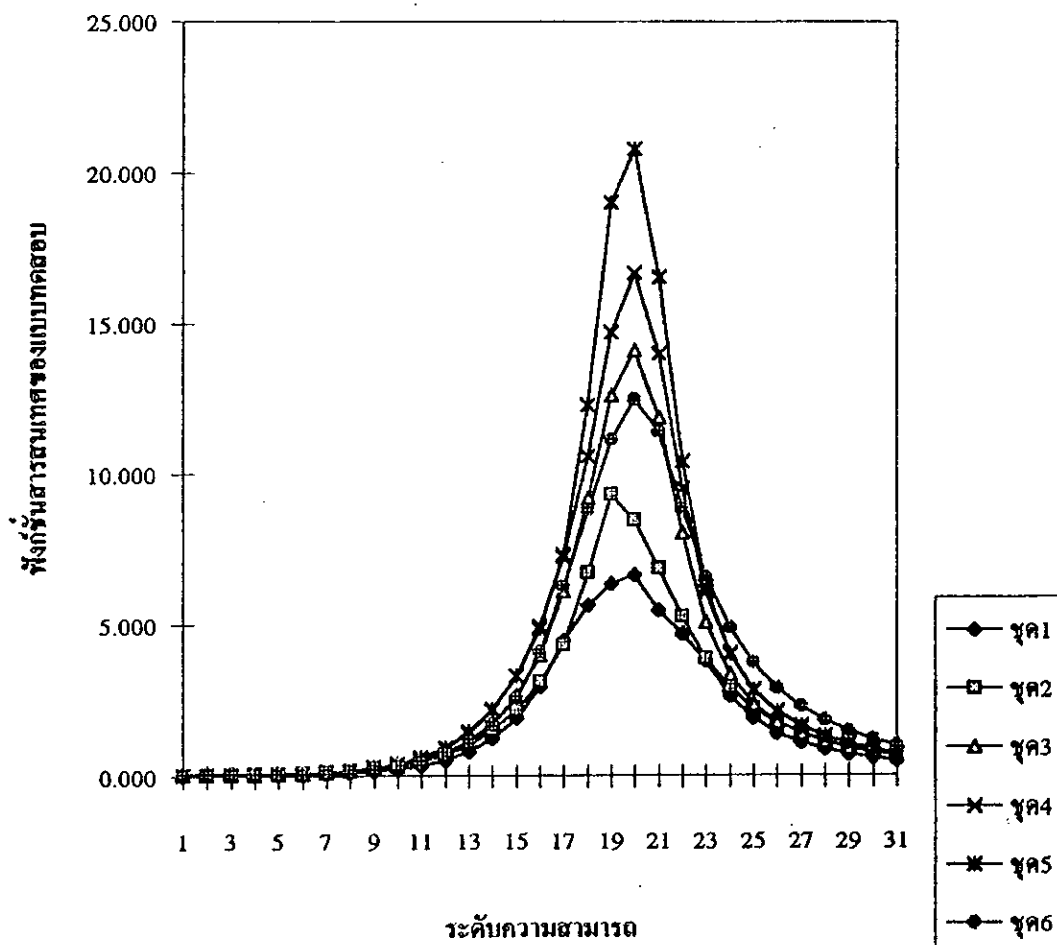
จากภาพประกอบ 36 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซ์ิมไลค์ลิจิต ของกลุ่มตัวอย่างขนาด 2,500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

$\{I(e) = 4.604 ; e = 1.000\}$ เมื่อทดสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อทดสอบมีขนาด 100 ข้อ $\{I(e) = 13.304 ; e = 1.400\}$

จากภาพประกอบ 37 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัมไลค์ลิฮูด ของกลุ่มตัวอย่างขนาด 3,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 4.684 ; e = 0.800\}$ เมื่อทดสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อทดสอบมีขนาด 100 ข้อ $\{I(e) = 12.911 ; e = 1.200\}$

ดังนั้น จากภาพประกอบ 25 - 37 แสดงว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัมไลค์ลิฮูด มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.136 ; e = 0.600\}$ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน ข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อทดสอบมีขนาด 100 ข้อ ขนาดของกลุ่มตัวอย่าง 2,500 คน $\{I(e) = 13.304 ; e = 1.400\}$

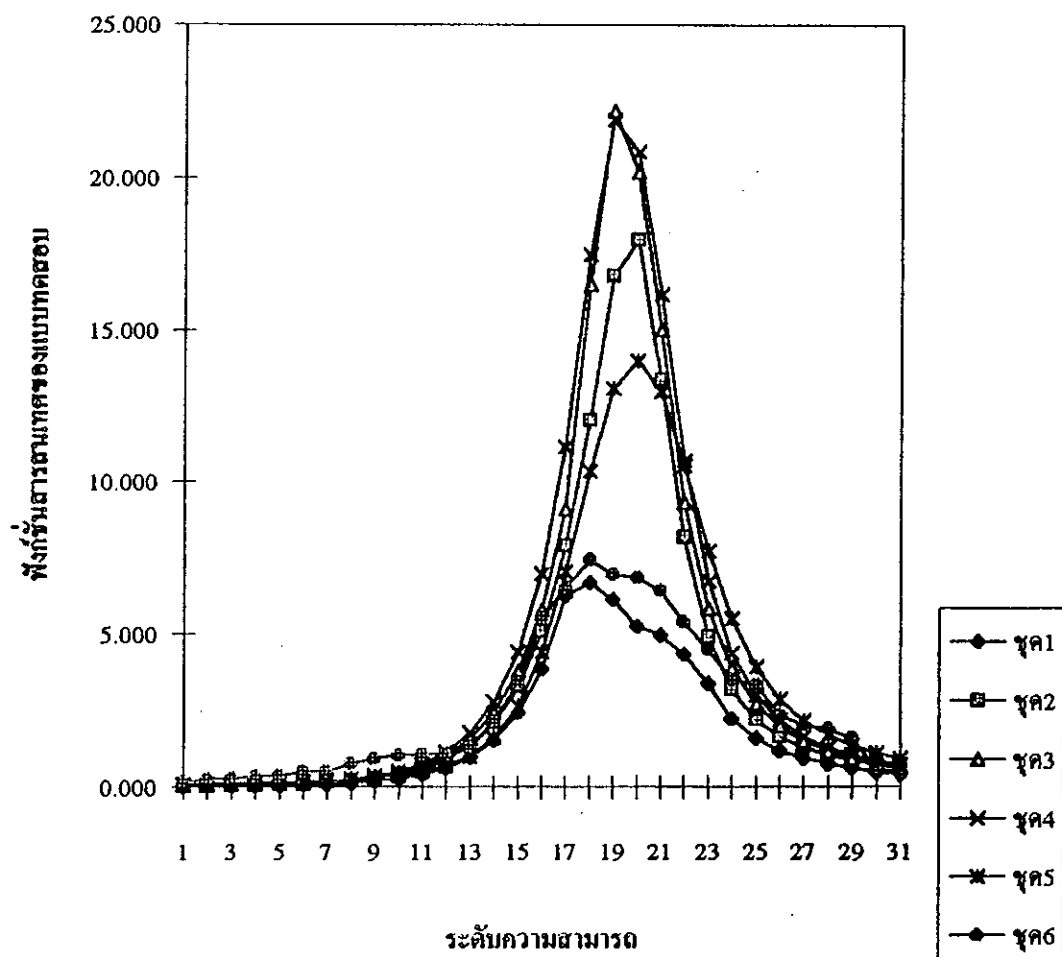
6.2 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอิมพิริสติก เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 38 - 50



ภาพประกอบ 38 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอิมพิวริสติก เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

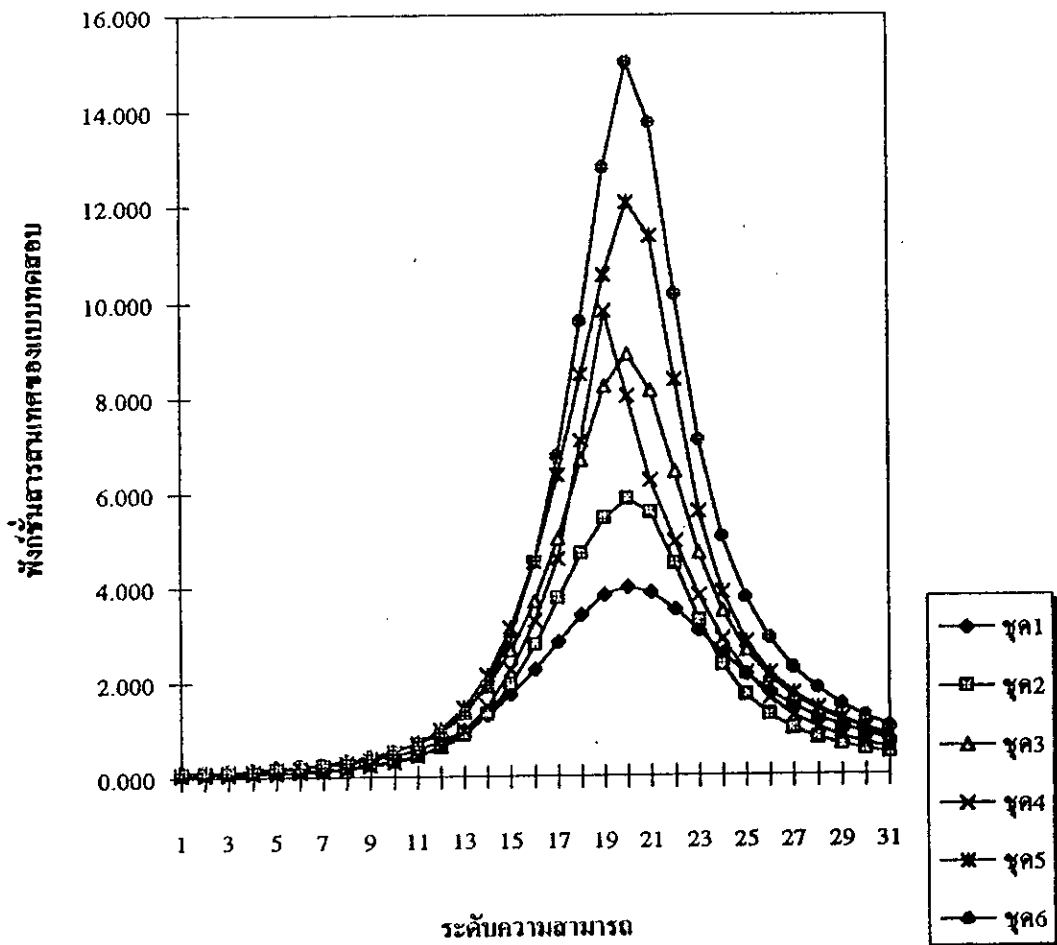
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 39 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อกลุ่มตัวอย่างมีขนาด 300 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

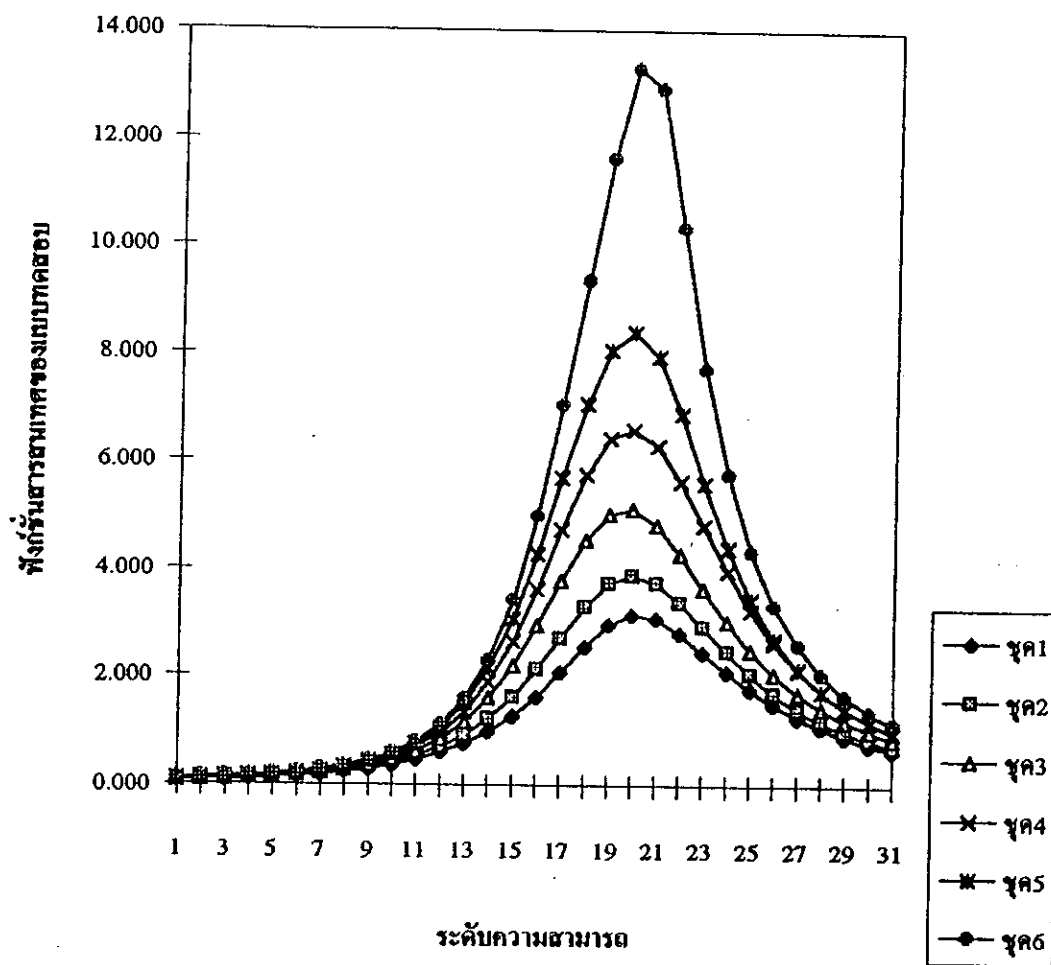
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 40 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอิมพิริสติก เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

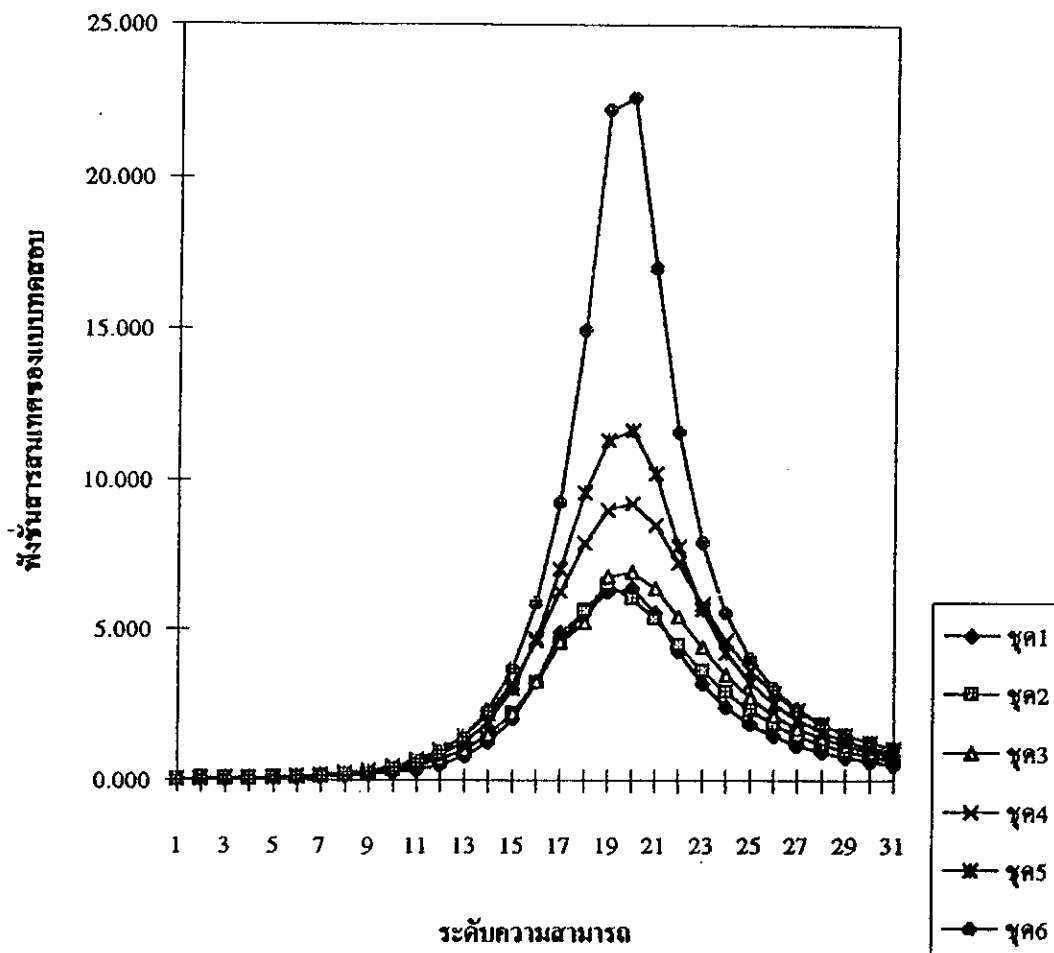
- ชุค 1 จำนวนข้อสอบ 50 ข้อ
 ชุค 2 จำนวนข้อสอบ 60 ข้อ
 ชุค 3 จำนวนข้อสอบ 70 ข้อ
 ชุค 4 จำนวนข้อสอบ 80 ข้อ
 ชุค 5 จำนวนข้อสอบ 90 ข้อ
 ชุค 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 41 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน-ราสัน เมื่อกลุ่มตัวอย่างมีขนาด 500 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

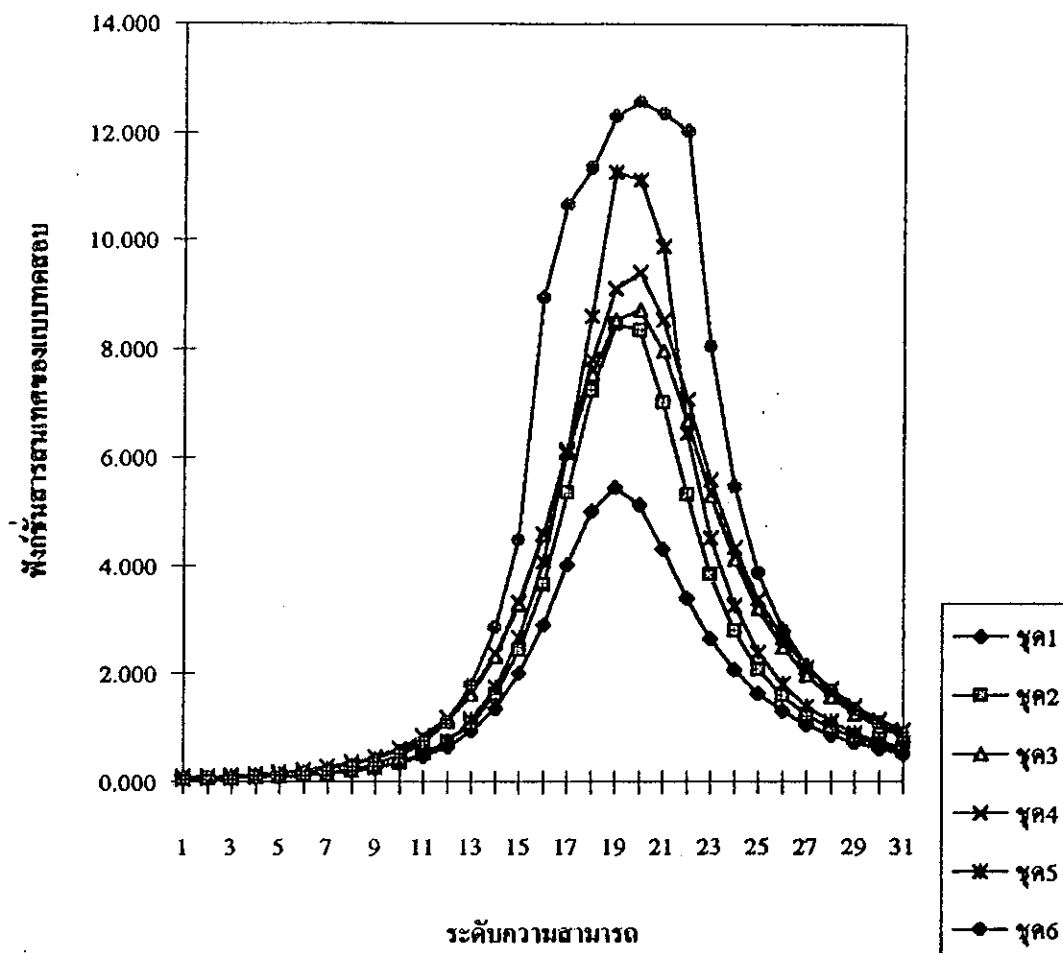
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 42 เปรียบเทียบโค้งความสัมพันธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีเอสติด เมื่อกลุ่มตัวอย่างมีขนาด 600 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

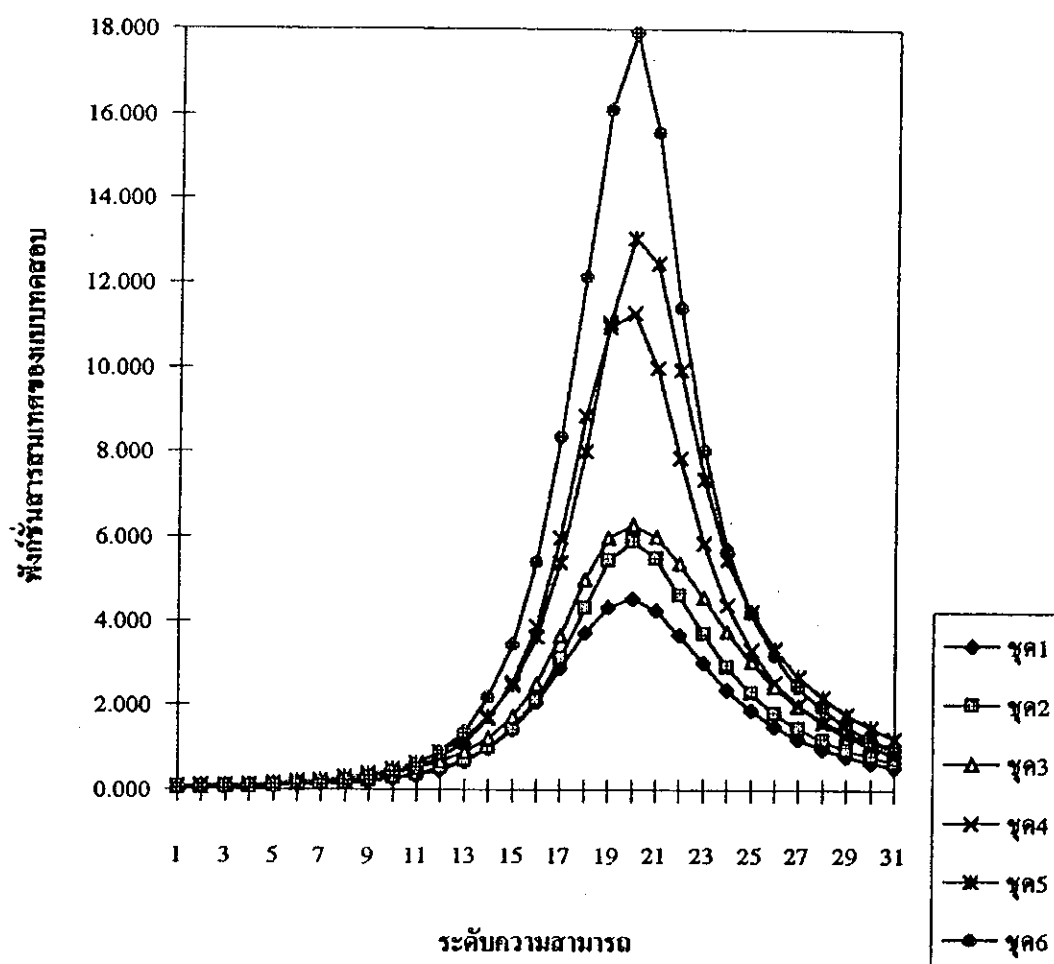
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 43 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน-ราสัน เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

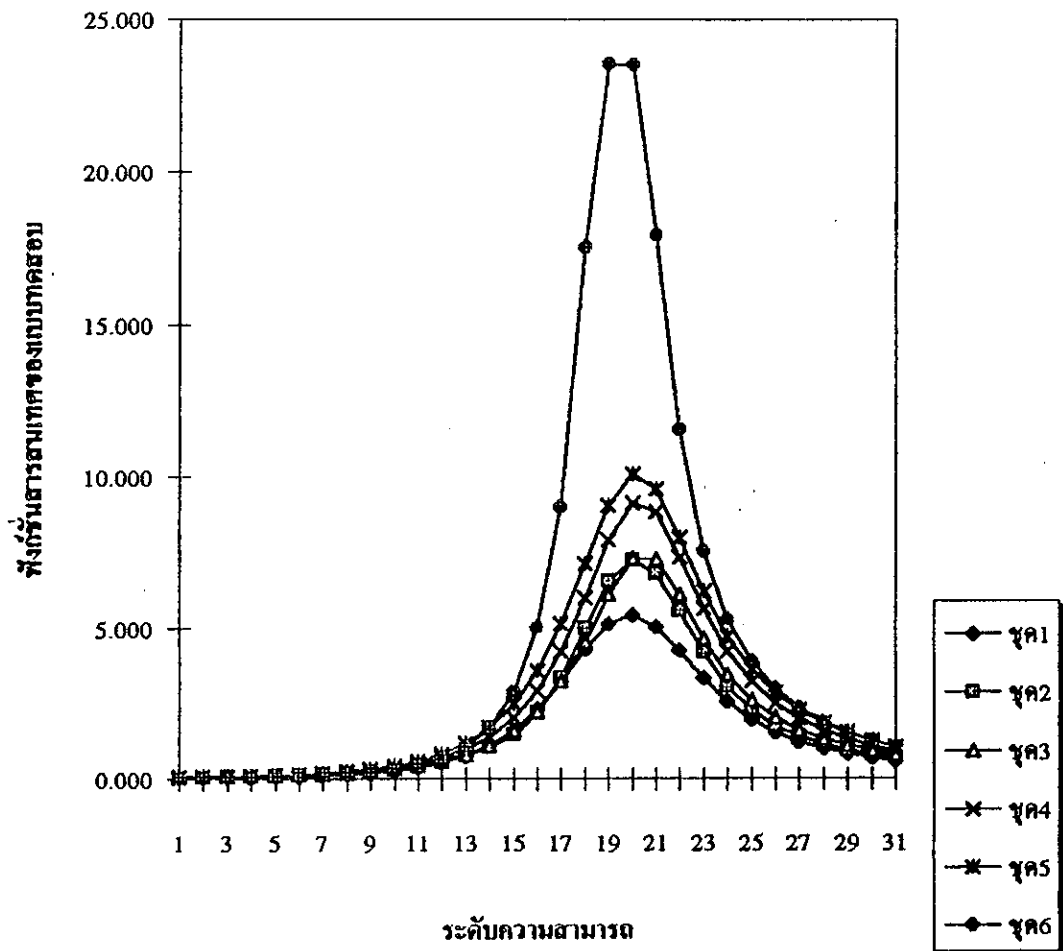
- ชค 1 จำนวนข้อสอบ 50 ข้อ
- ชค 2 จำนวนข้อสอบ 60 ข้อ
- ชค 3 จำนวนข้อสอบ 70 ข้อ
- ชค 4 จำนวนข้อสอบ 80 ข้อ
- ชค 5 จำนวนข้อสอบ 90 ข้อ
- ชค 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 44 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน-ราฟสัน เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

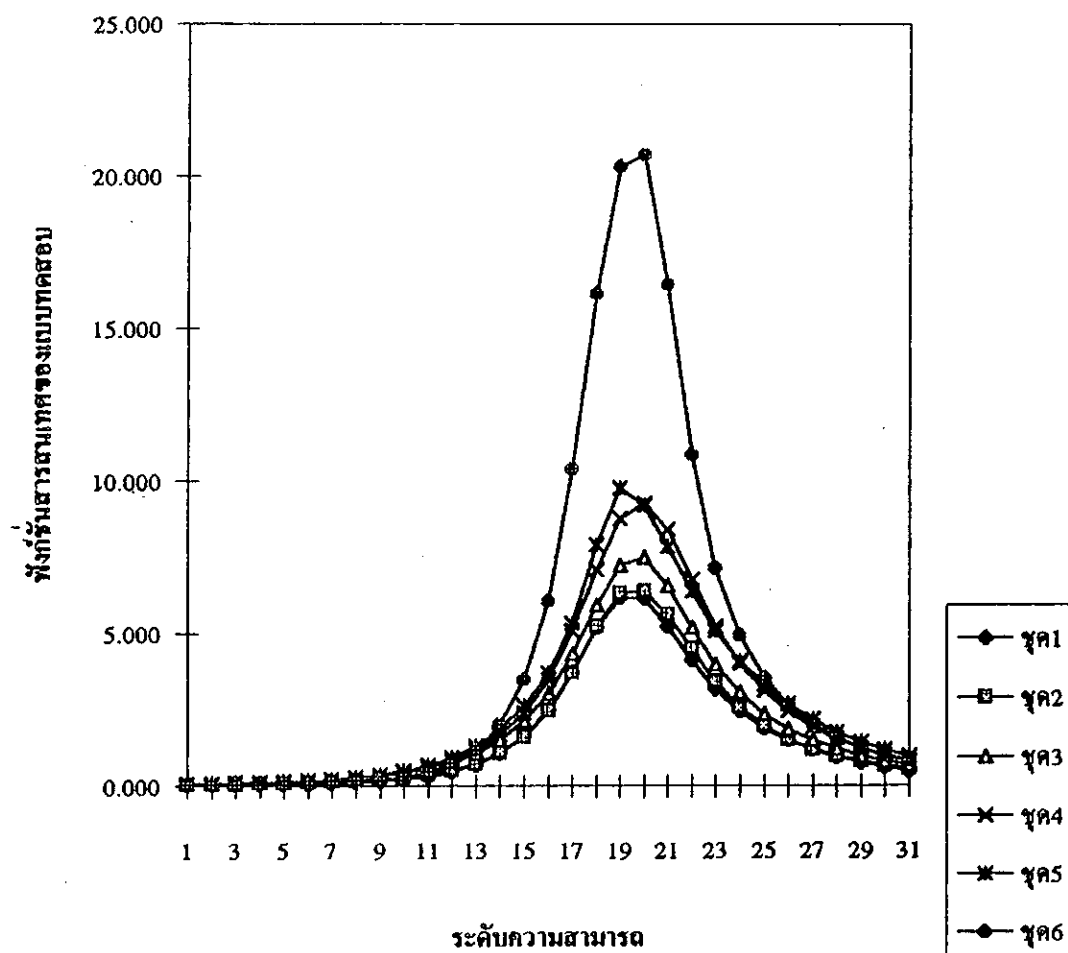
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 45 เปรียบเทียบโด่งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน-ราสัน เมื่อกลุ่มตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

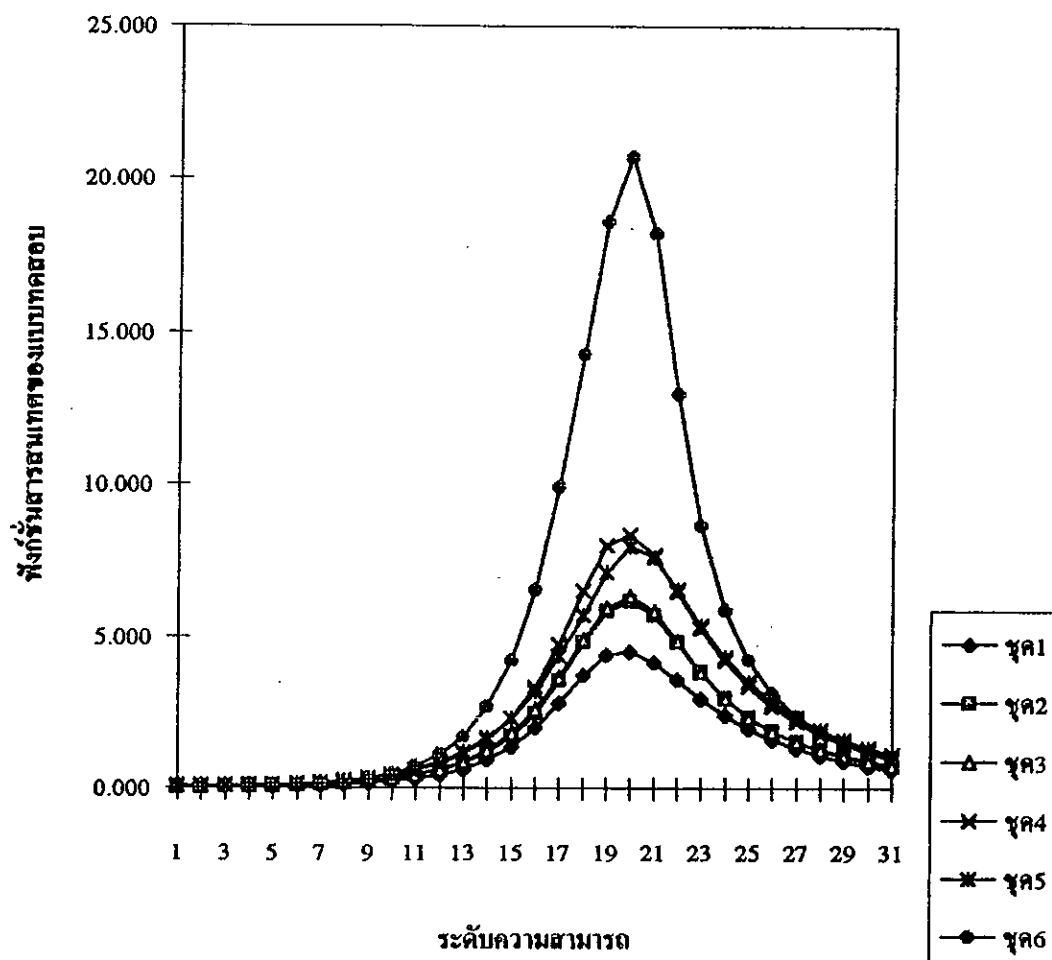
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 46 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

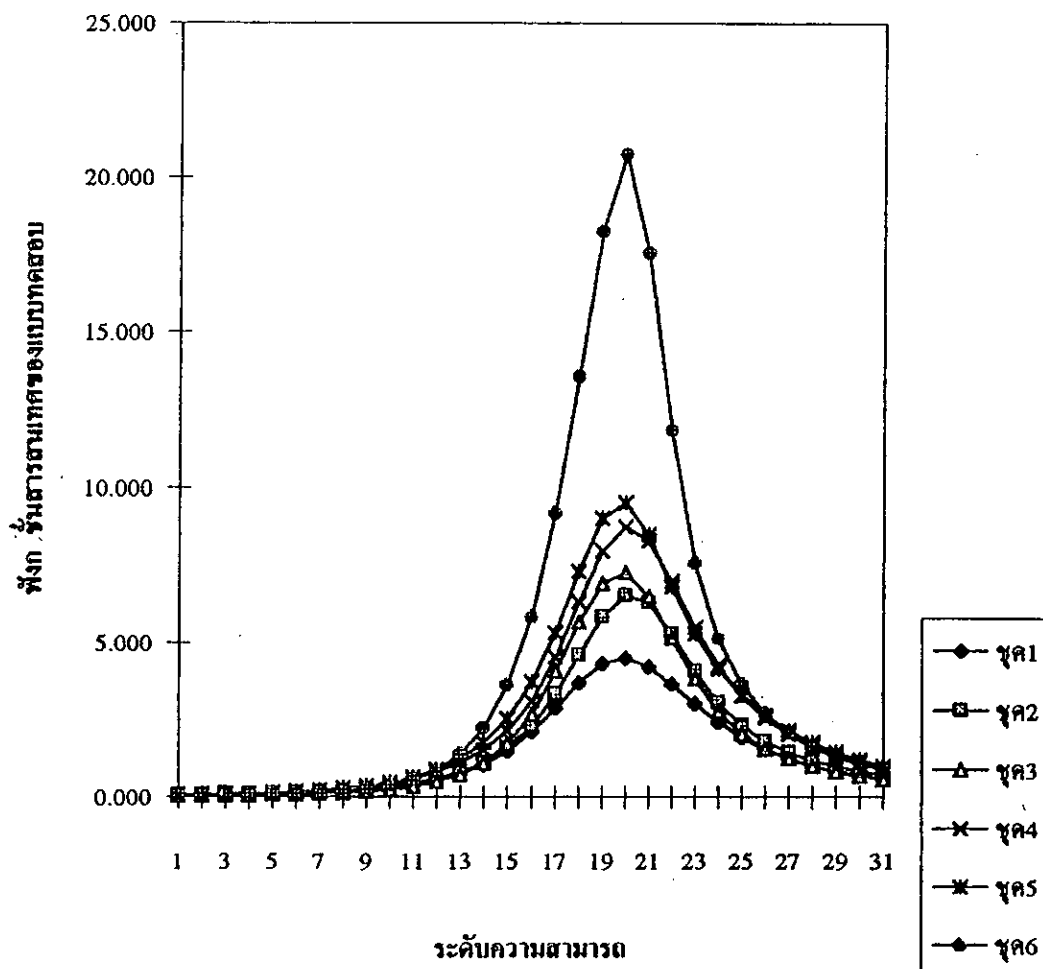
- ชุค 1 จำนวนข้อสอบ 50 ข้อ
- ชุค 2 จำนวนข้อสอบ 60 ข้อ
- ชุค 3 จำนวนข้อสอบ 70 ข้อ
- ชุค 4 จำนวนข้อสอบ 80 ข้อ
- ชุค 5 จำนวนข้อสอบ 90 ข้อ
- ชุค 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 47 เปรียบเทียบโด่งสาร์สนะตงแบบทคสอบที่เป็นผลมาจากการประมาธ
ค่าพารามิเตอร์โดยวิธีวิธีสถิติ เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน และข้อสอบมีขนาด
ต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

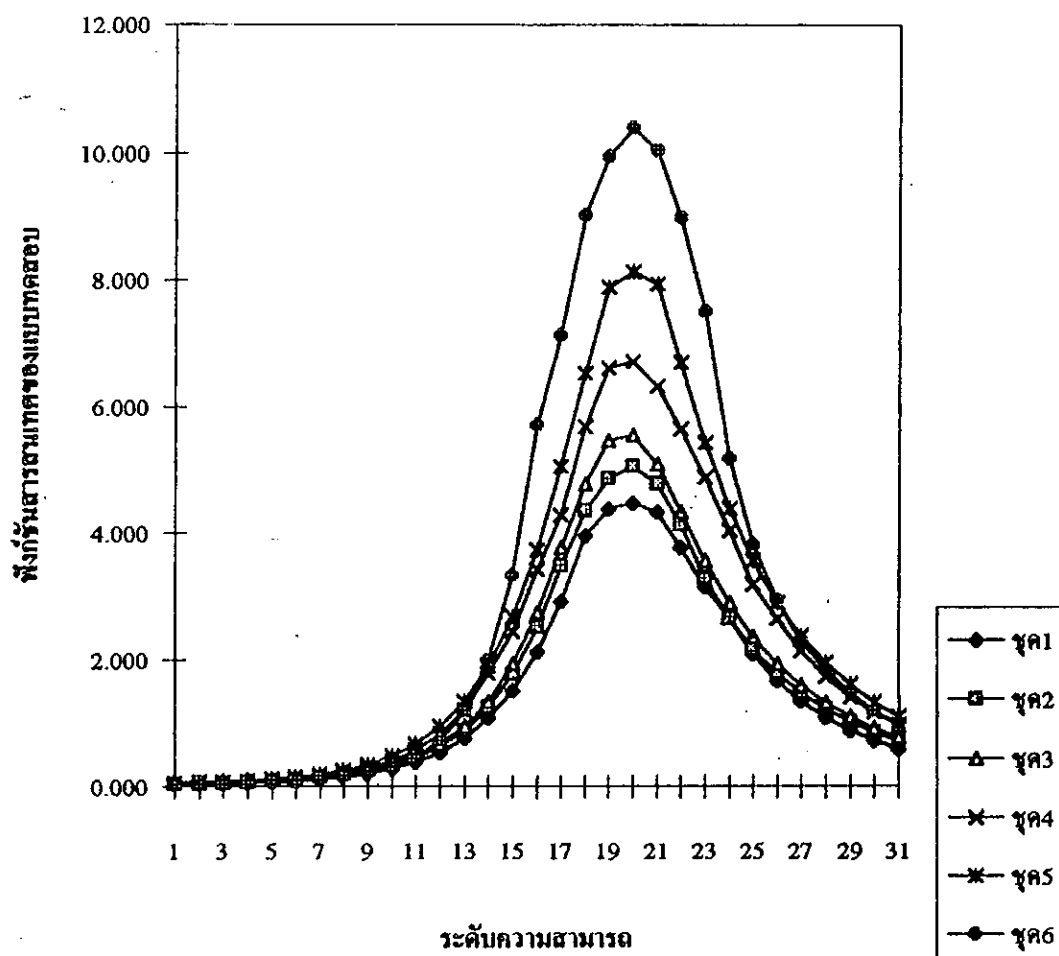
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 48 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน-ราฟสัน เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาดต่างกัน ๕ ระดับความสามารถต่าง ๆ

หมายเหตุ

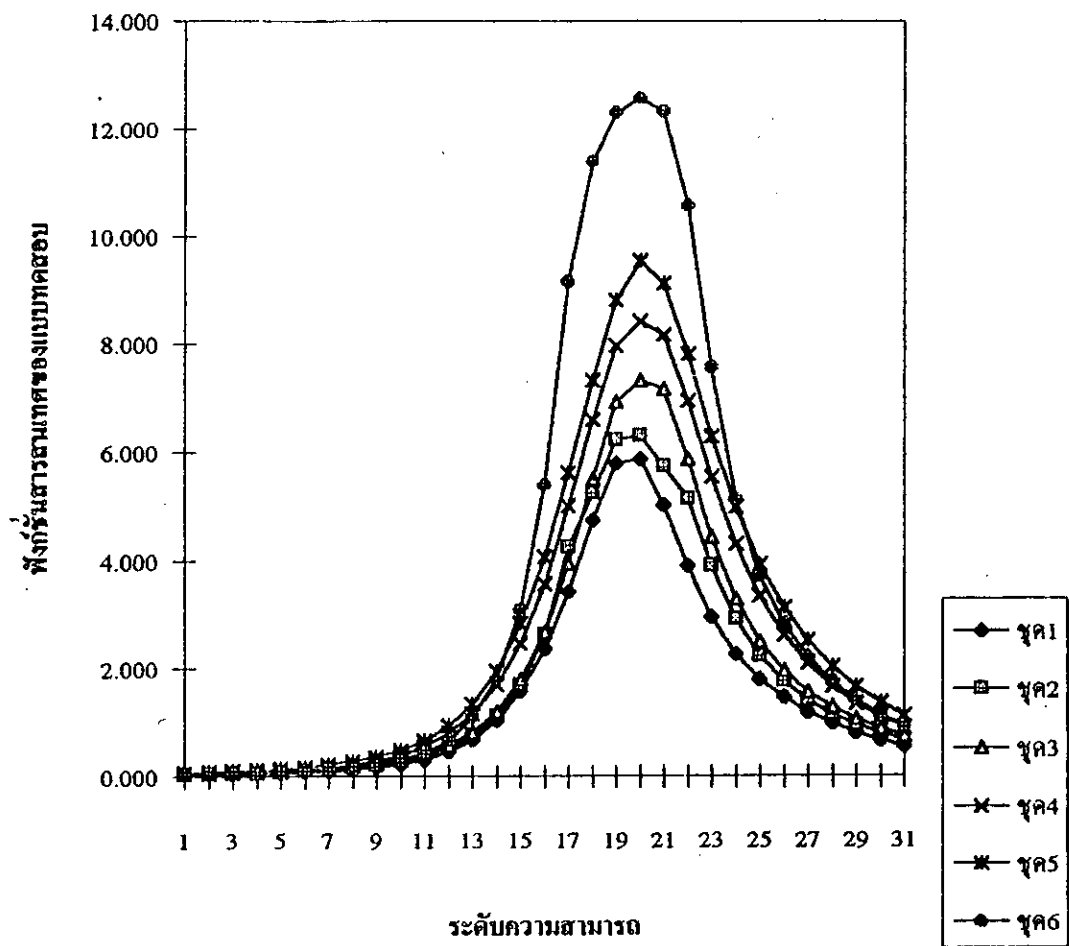
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 49 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน-ราสัน. เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

- ชุค 1 จำนวนข้อสอบ 50 ข้อ
- ชุค 2 จำนวนข้อสอบ 60 ข้อ
- ชุค 3 จำนวนข้อสอบ 70 ข้อ
- ชุค 4 จำนวนข้อสอบ 80 ข้อ
- ชุค 5 จำนวนข้อสอบ 90 ข้อ
- ชุค 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 50 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ

จากภาพประกอบ 38 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีวิวิธิตของกลุ่มตัวอย่างขนาด 200 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 6.626 ; \theta = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 12.439 ; \theta = 0.800\}$

จากภาพประกอบ 39 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีวีริสติกของกลุ่มตัวอย่างขนาด 300 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 6.693 ; e = 0.400\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 90 ข้อ $\{I(e) = 22.195 ; e = 0.600\}$

จากภาพประกอบ 40 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีวิธีสถิติของกลุ่มตัวอย่างขนาด 400 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 3.985 ; e = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 14.998 ; e = 0.800\}$

จากภาพประกอบ 41 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอิมพิวติคของกลุ่มตัวอย่างขนาด 500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 3.147 ; \theta = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 13.308 ; \theta = 0.800\}$

จากภาพประกอบ 42 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีวิธีคิดของกลุ่มตัวอย่างขนาด 600 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 6.387 ; e = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 22.586 ; e = 0.800\}$

จากภาพประกอบ 43 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีวิวิธิตของกลุ่ตัวอย่างขนาด 700 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง ($I(0) = 5.451$; $\sigma = 0.600$) เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับ

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 12.578 ; \theta = 0.800\}$

จากภาพประกอบ 44 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีวีริสติกของกลุ่มตัวอย่างขนาด 800 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อทดสอบมีขนาด 100 ข้อ $\{I(e) = 17.890 ; e = 0.800\}$

จากภาพประกอบ 45 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีเอสติคของกลุ่มตัวอย่างขนาด 900 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 23.548 ; e = 0.600\}$

จากภาพประกอบ 46 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีอาร์เอสติคของกลุ่มตัวอย่างขนาด 1,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 20.670 ; \theta = 0.800\}$

จากภาพประกอบ 47 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวีเอสติกของกลุ่มตัวอย่างขนาด 1,500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อทดสอบมีขนาด 100 ขั้ว $\{I(\phi) = 20.726 ; \phi = 0.800\}$

จากภาพประกอบ 48 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวีเอสติคของกลุ่มตัวอย่างขนาด 2,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 20.746 ; \theta = 0.800\}$

จากภาพประกอบ 49 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีอาร์สติกของกลุ่มตัวอย่างขนาด 2,500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

$\{I(\alpha) = 4.467 ; \alpha = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\alpha) = 10.387 ; \alpha = 0.800\}$

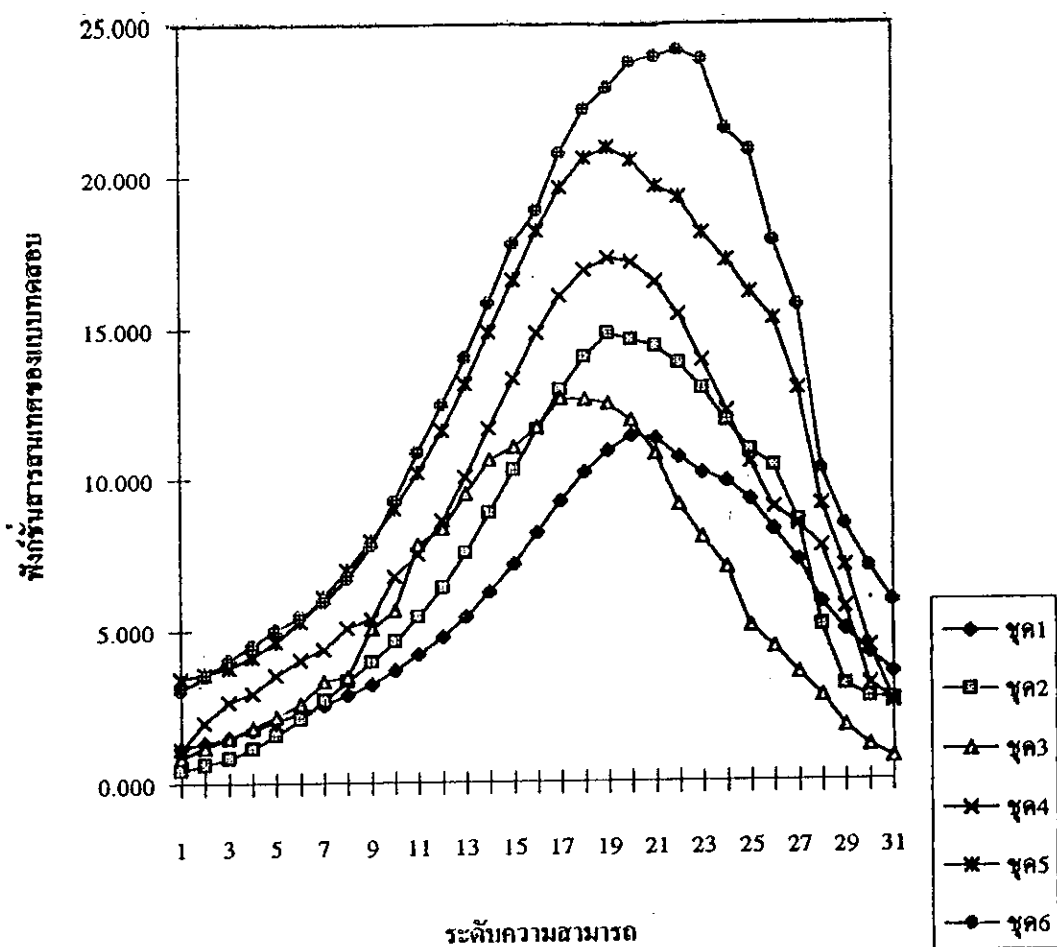
จากภาพประกอบ 50 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีเอสติคของกลุ่มตัวอย่างขนาด 3,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

$\{I(\alpha) = 5.858 ; \alpha = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\alpha) = 12.571 ; \alpha = 0.800\}$

ดังนั้น จากภาพประกอบ 38 - 50 แสดงว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีเอสติค มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\alpha) = 3.147 ;$

$\alpha = 0.800\}$ เมื่อกลุ่มตัวอย่างมีขนาด 500 คน ข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ ขนาดของกลุ่มตัวอย่าง 900 คน $\{I(\alpha) = 23.548 ; \alpha = 0.800\}$

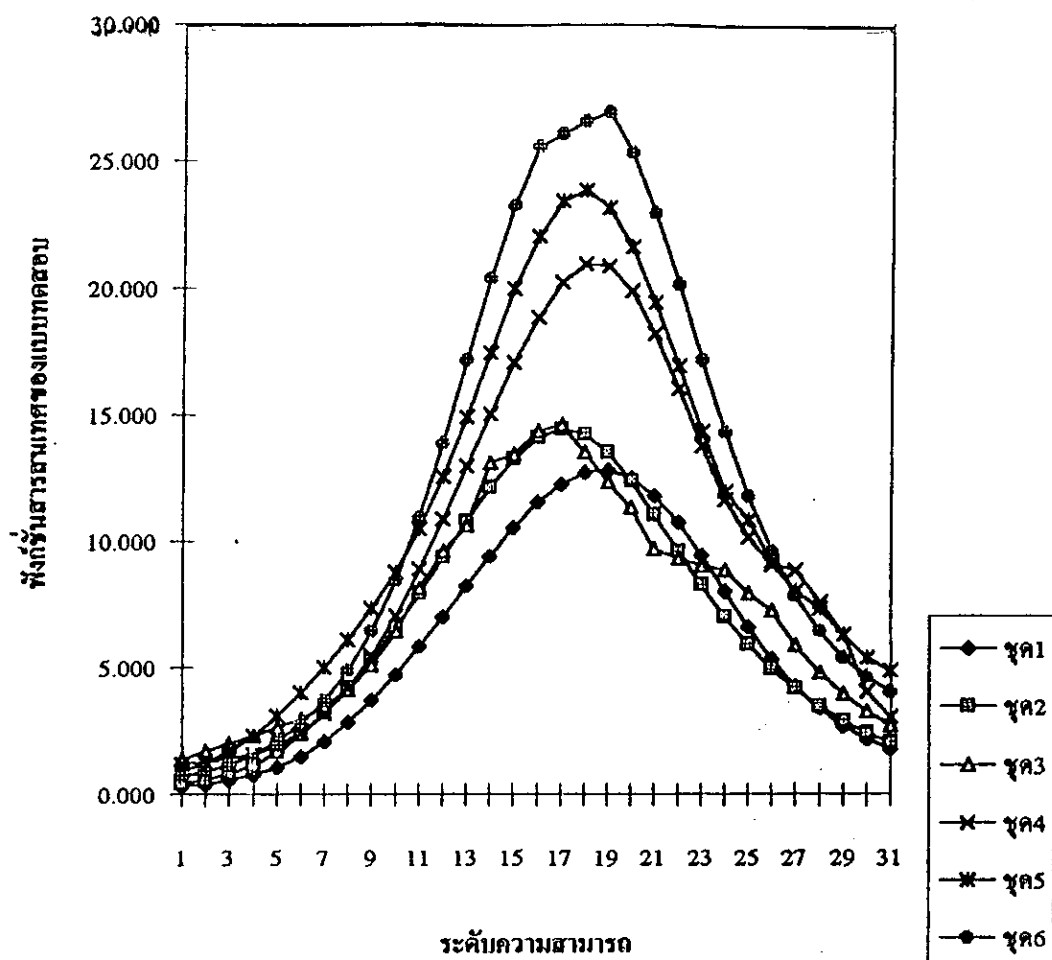
6.3 ผลการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 51 - 63



ภาพประกอบ 51 เปรียบเทียบโค้งสาร์สน์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

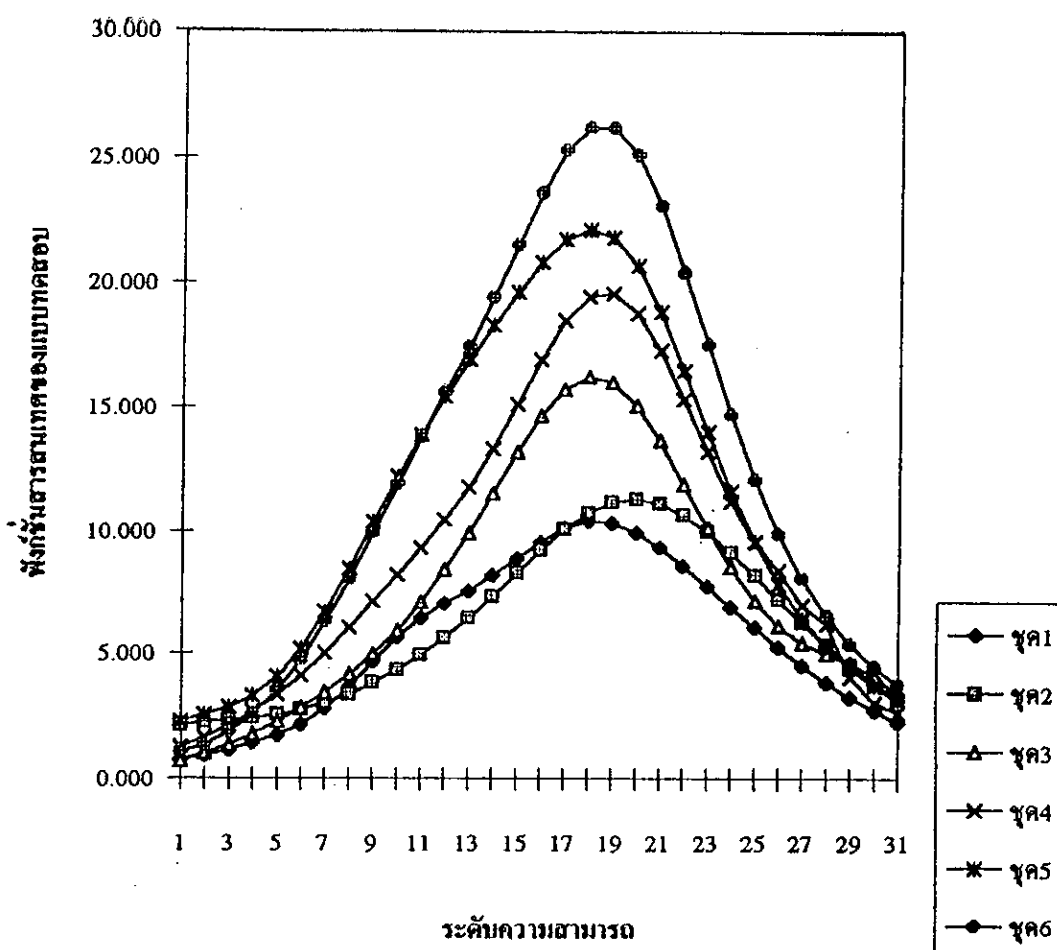
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 52 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมวลค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 300 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

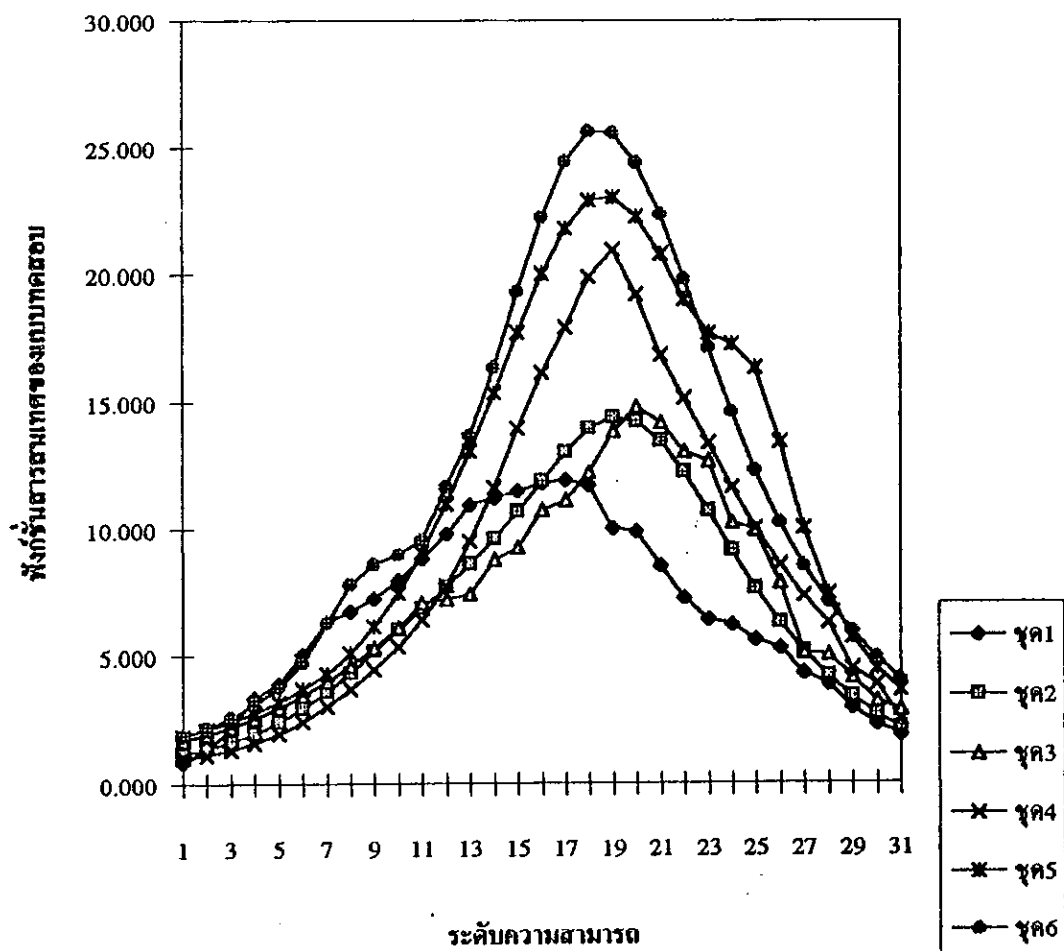
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 53 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 400 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

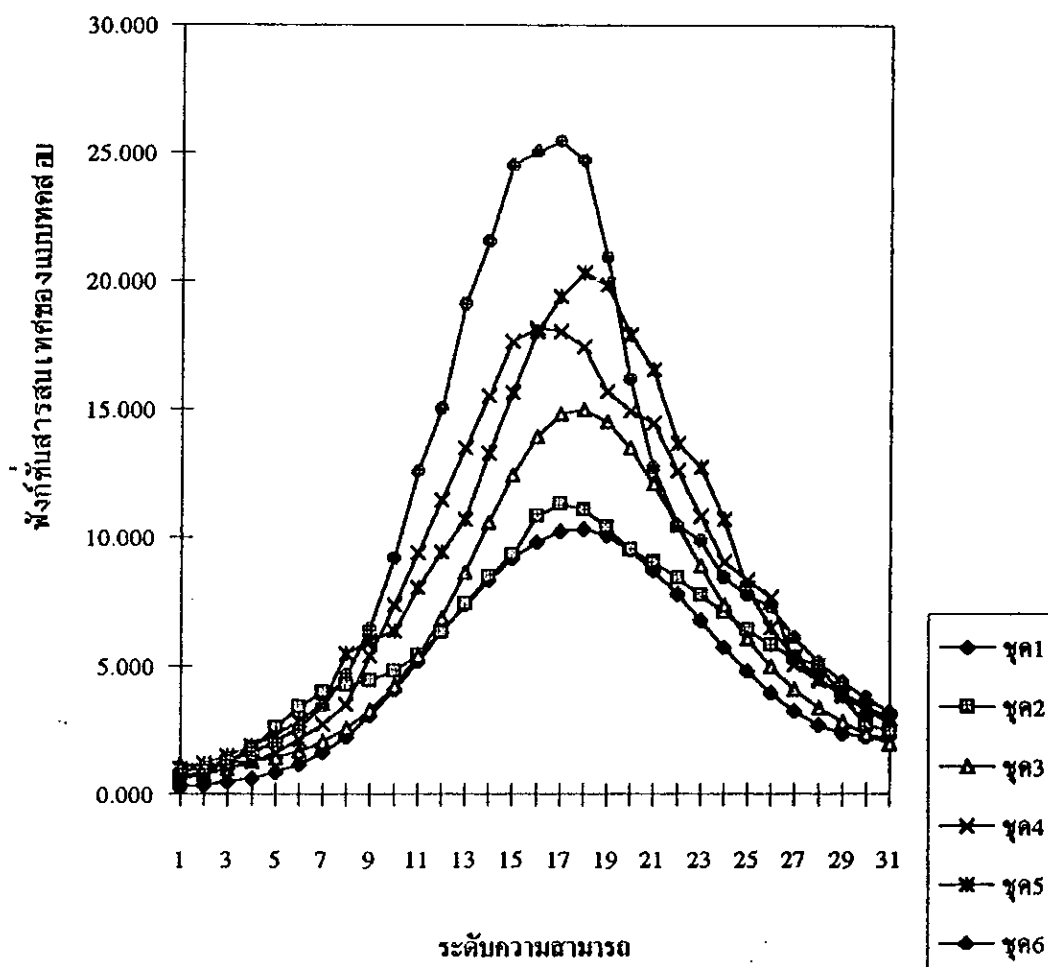
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
ชุด 2 จำนวนข้อสอบ 60 ข้อ
ชุด 3 จำนวนข้อสอบ 70 ข้อ
ชุด 4 จำนวนข้อสอบ 80 ข้อ
ชุด 5 จำนวนข้อสอบ 90 ข้อ
ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 54 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 500 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

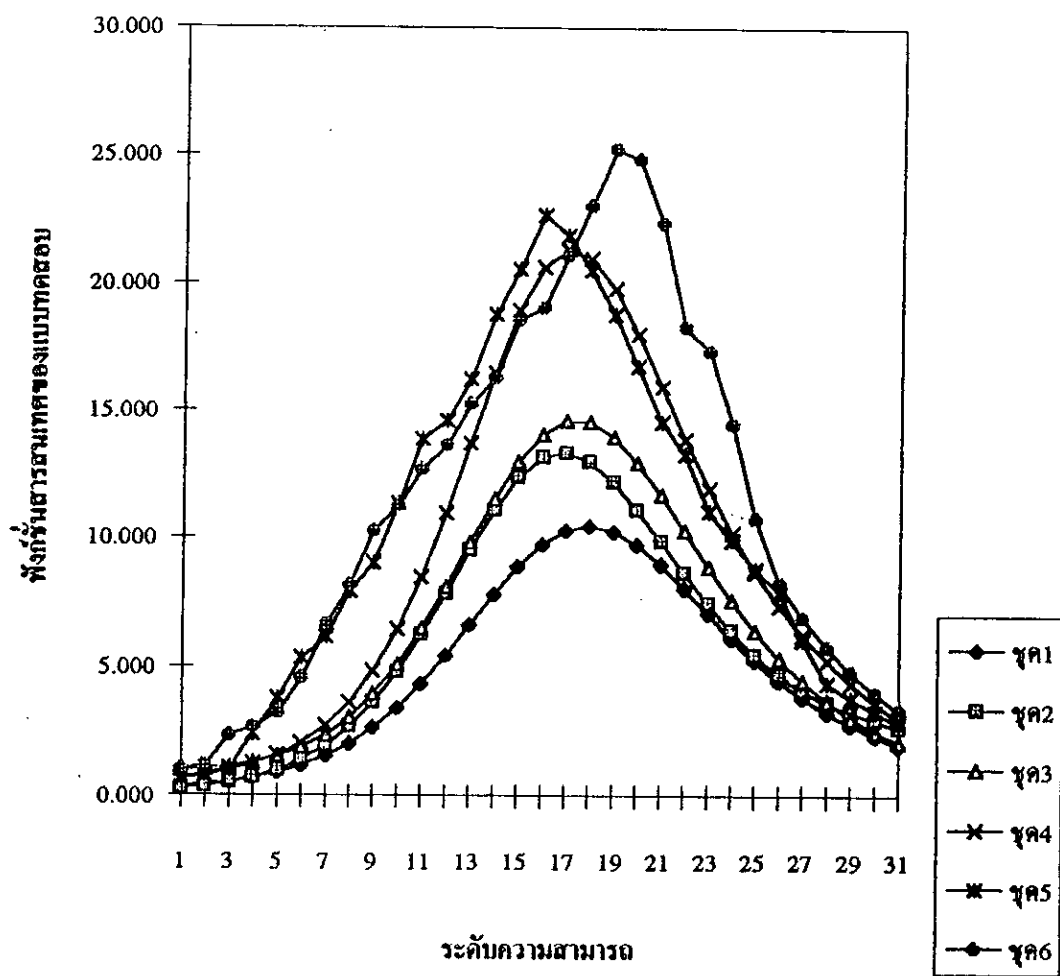
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
ชุด 2 จำนวนข้อสอบ 60 ข้อ
ชุด 3 จำนวนข้อสอบ 70 ข้อ
ชุด 4 จำนวนข้อสอบ 80 ข้อ
ชุด 5 จำนวนข้อสอบ 90 ข้อ
ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 55 เปรียบเทียบโค้งสำหรับแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 600 คน และข้อสอบมีขนาดต่างกัน ๕ ระดับความสามารถต่าง ๆ

หมายเหตุ

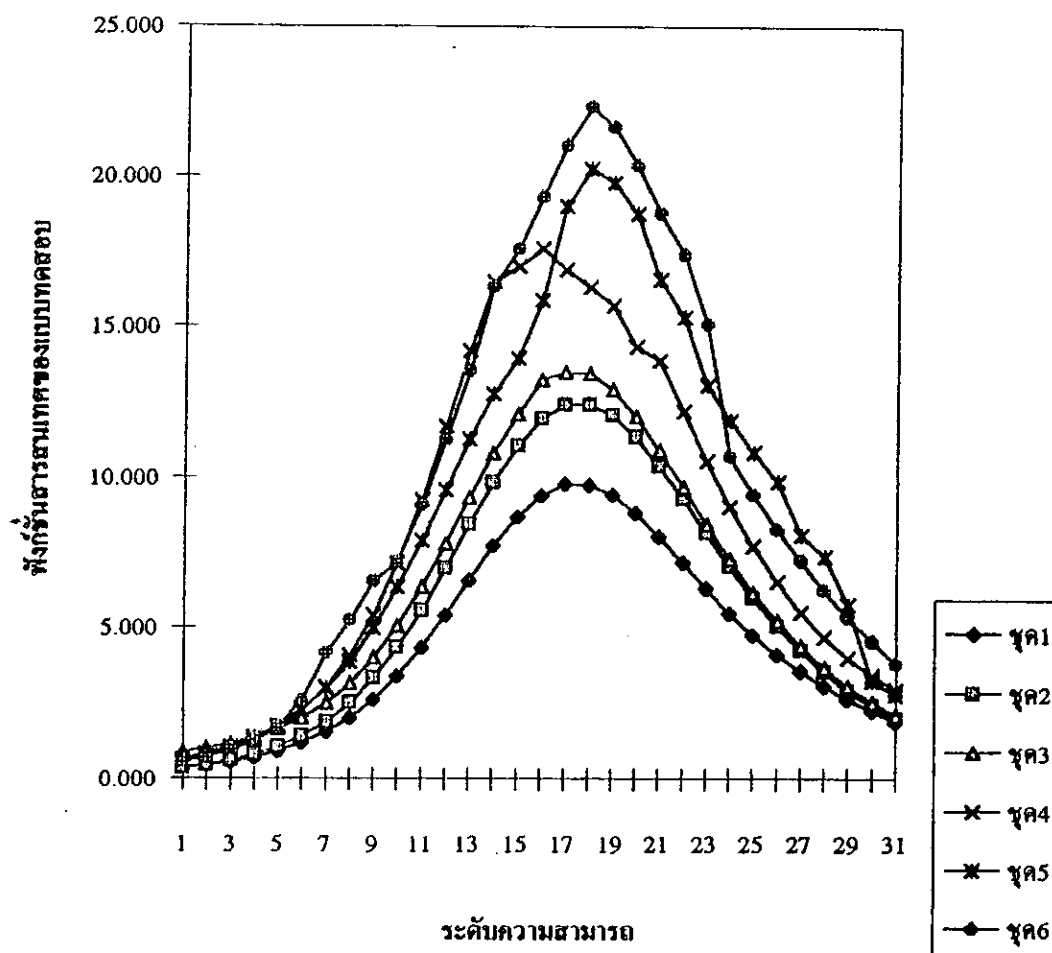
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 56 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 700 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

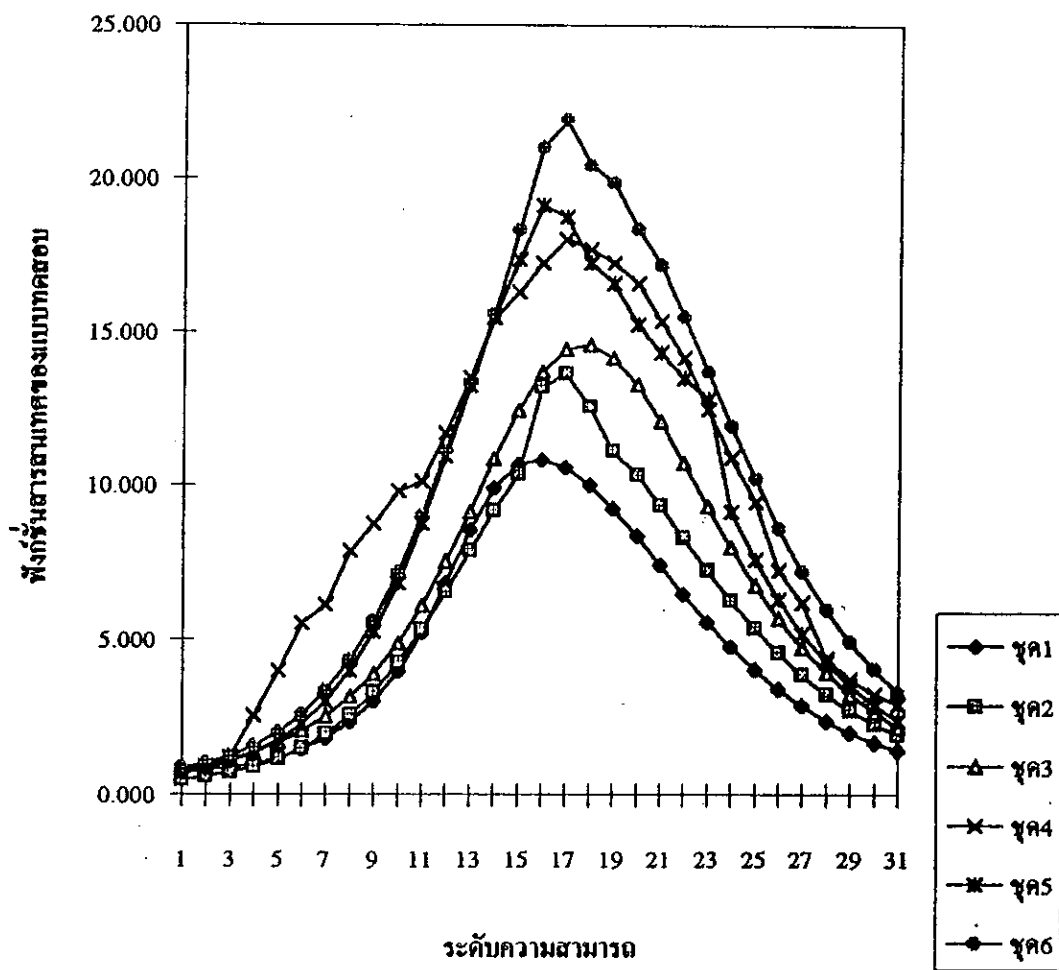
- ชุค 1 จำนวนข้อสอบ 50 ข้อ
- ชุค 2 จำนวนข้อสอบ 60 ข้อ
- ชุค 3 จำนวนข้อสอบ 70 ข้อ
- ชุค 4 จำนวนข้อสอบ 80 ข้อ
- ชุค 5 จำนวนข้อสอบ 90 ข้อ
- ชุค 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 57 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 800 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

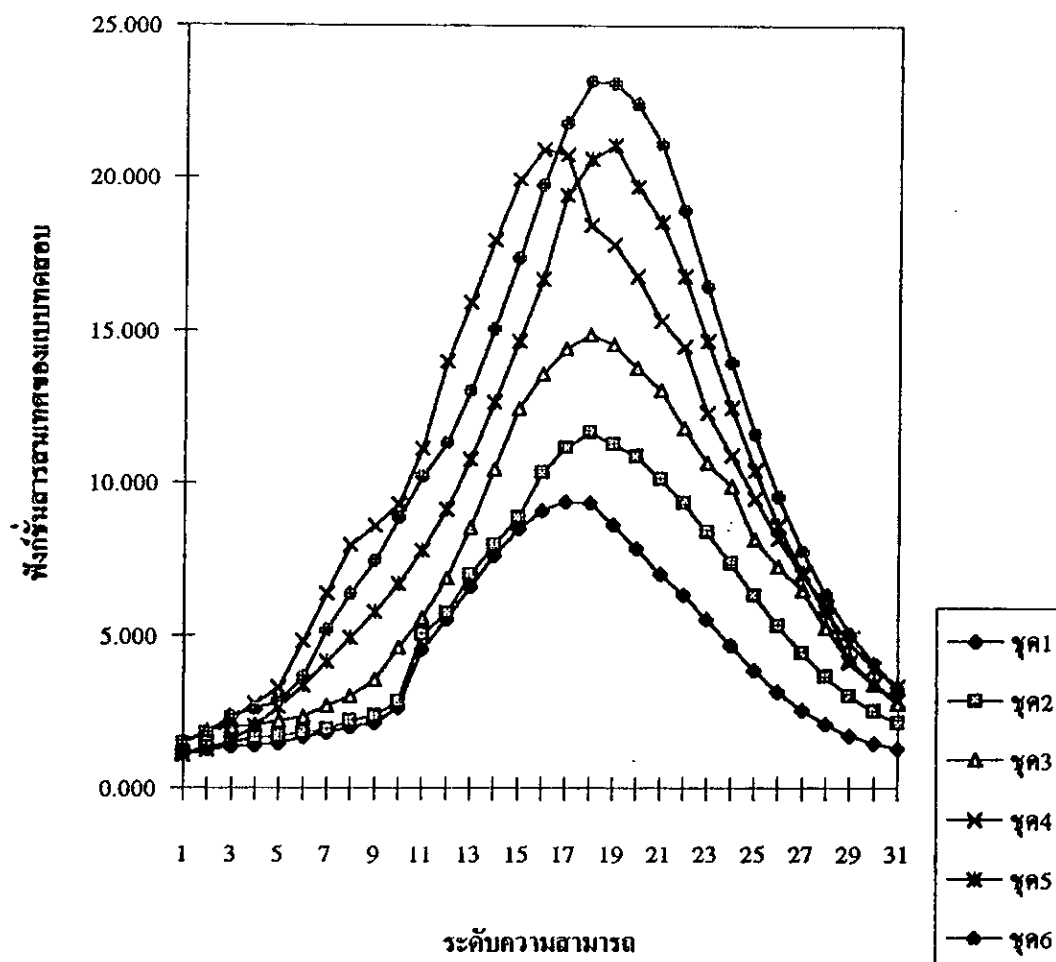
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
 ชุด 2 จำนวนข้อสอบ 60 ข้อ
 ชุด 3 จำนวนข้อสอบ 70 ข้อ
 ชุด 4 จำนวนข้อสอบ 80 ข้อ
 ชุด 5 จำนวนข้อสอบ 90 ข้อ
 ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 58 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 900 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

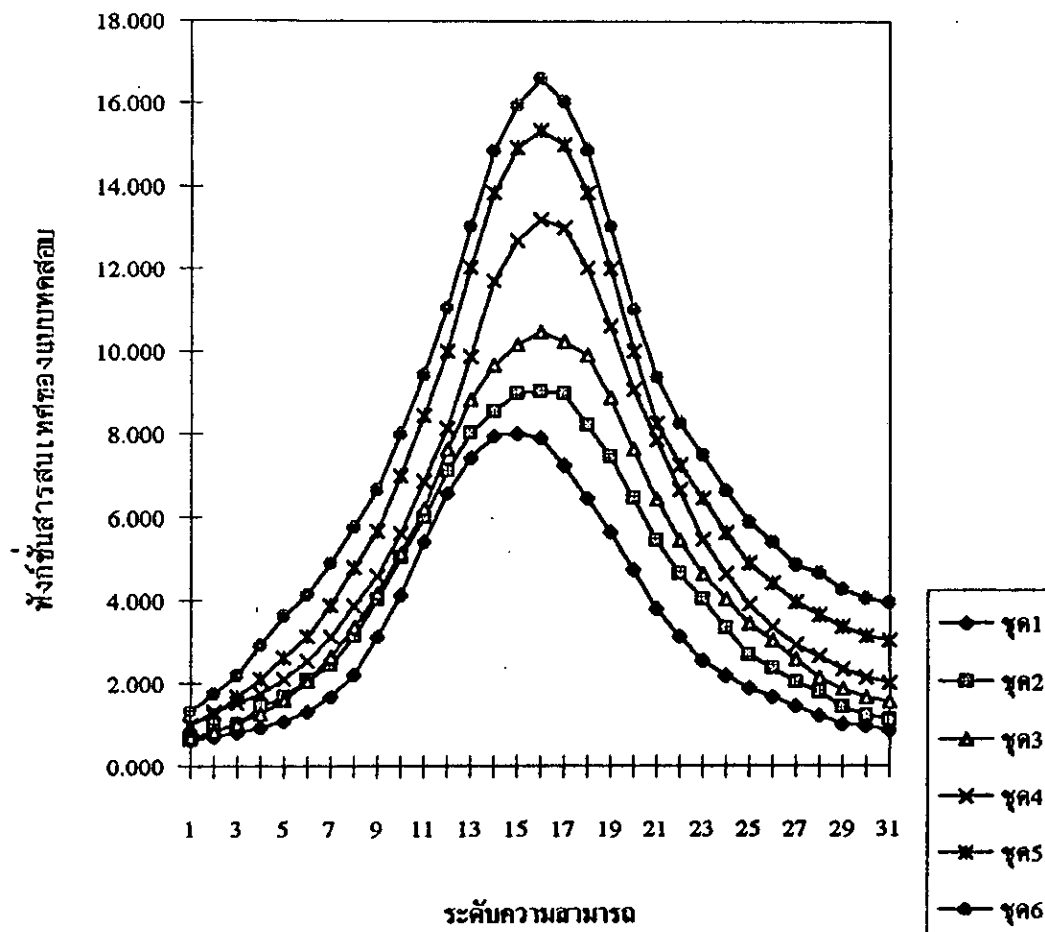
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 59 เปรียบเทียบโค้งการหาขนาดของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

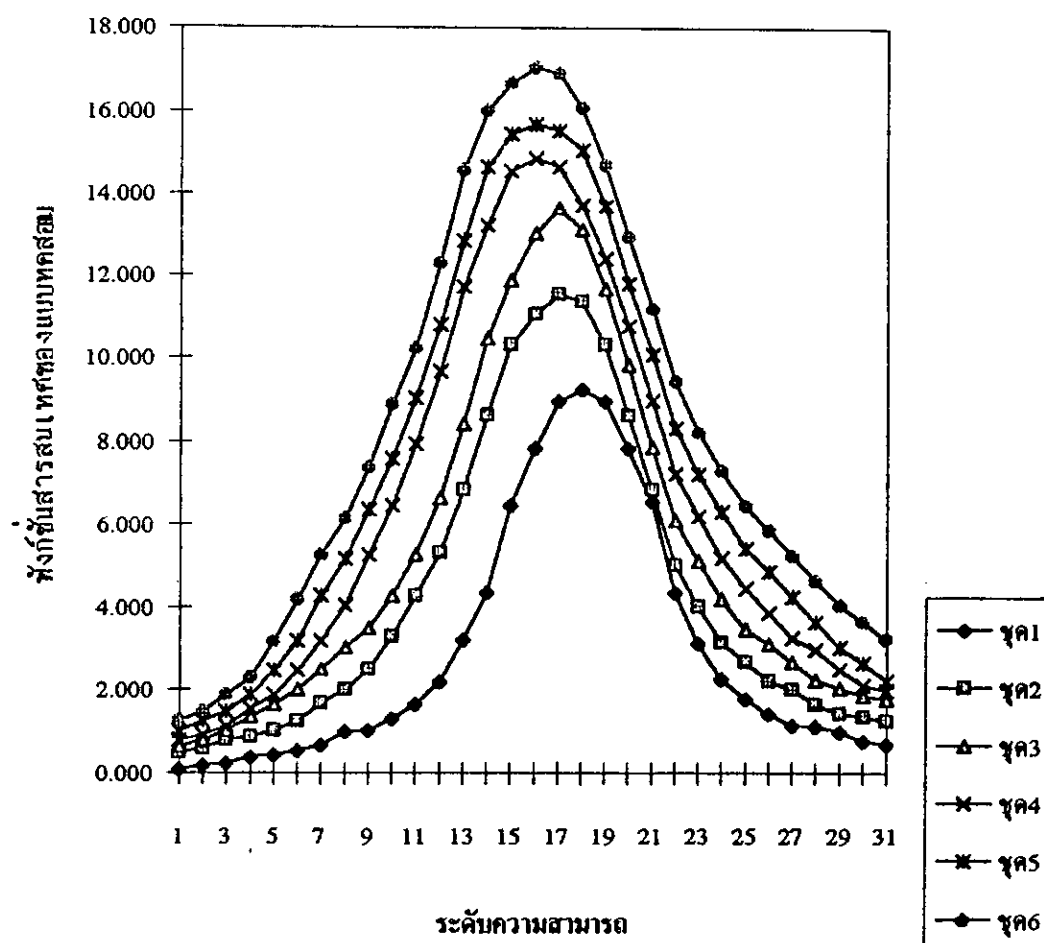
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 60 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

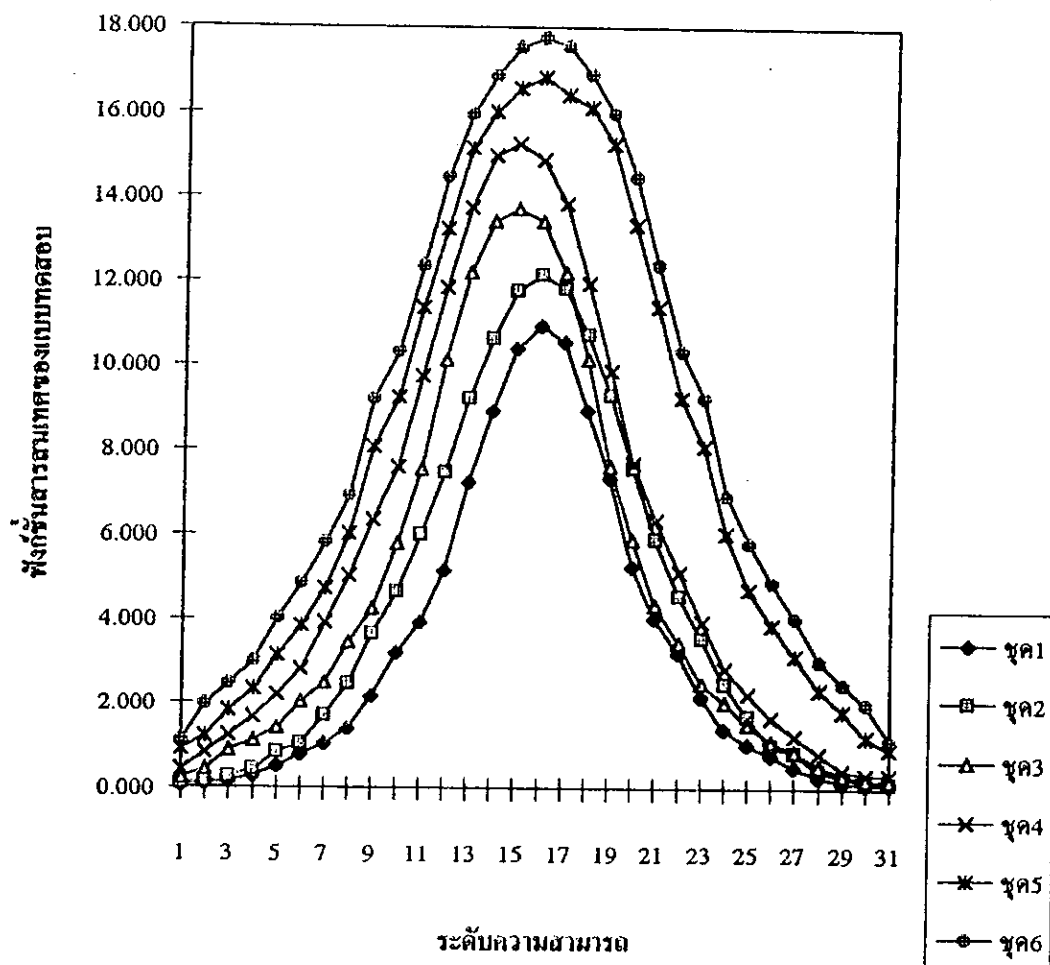
- จุด 1 จำนวนข้อสอบ 50 ข้อ
- จุด 2 จำนวนข้อสอบ 60 ข้อ
- จุด 3 จำนวนข้อสอบ 70 ข้อ
- จุด 4 จำนวนข้อสอบ 80 ข้อ
- จุด 5 จำนวนข้อสอบ 90 ข้อ
- จุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 61 เปรียบเทียบโค้งการสืบแทนของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

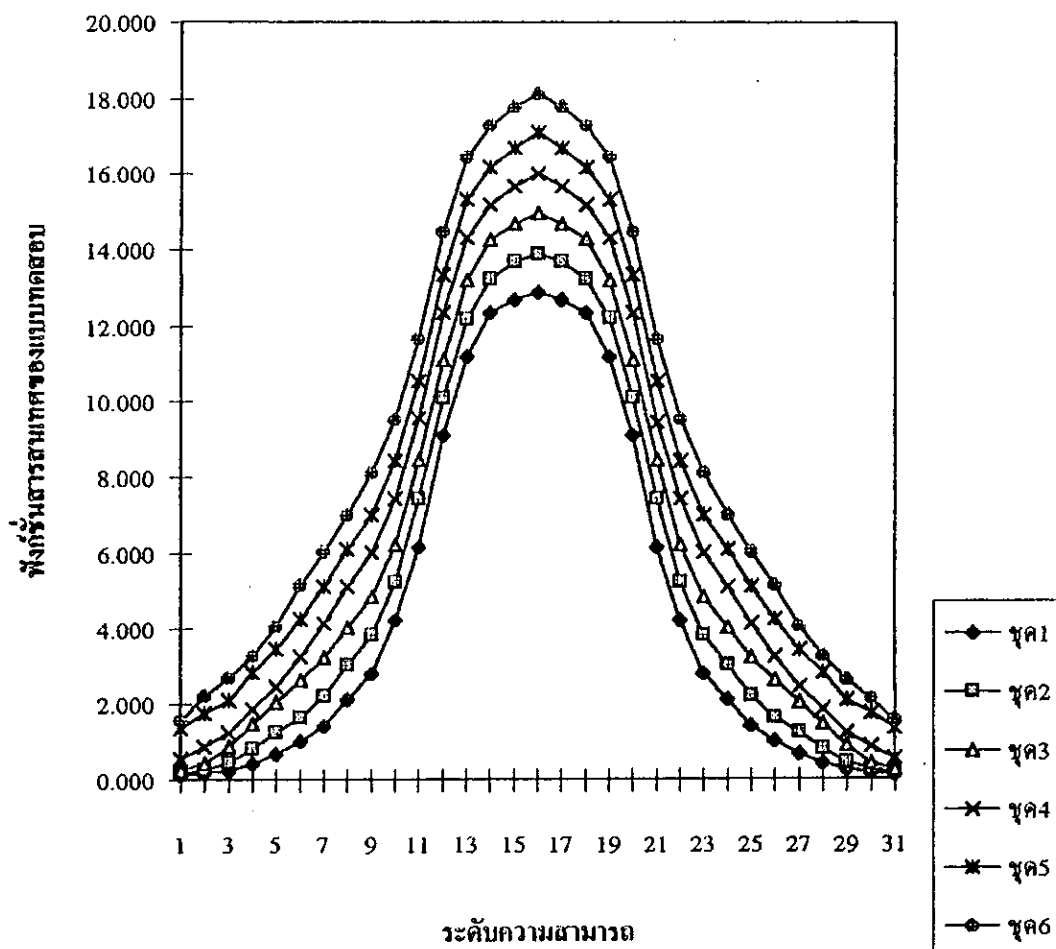
- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 62 เปรียบเทียบโค้งสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน และข้อสอบมีขนาดต่างกัน ๖ ระดับความสามารถต่าง ๆ

หมายเหตุ

- ชุด 1 จำนวนข้อสอบ 50 ข้อ
ชุด 2 จำนวนข้อสอบ 60 ข้อ
ชุด 3 จำนวนข้อสอบ 70 ข้อ
ชุด 4 จำนวนข้อสอบ 80 ข้อ
ชุด 5 จำนวนข้อสอบ 90 ข้อ
ชุด 6 จำนวนข้อสอบ 100 ข้อ



ภาพประกอบ 63 เปรียบเทียบโค้งกระจายของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน และข้อสอบมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

- ชุด 1 จำนวนข้อสอบ 50 ข้อ
- ชุด 2 จำนวนข้อสอบ 60 ข้อ
- ชุด 3 จำนวนข้อสอบ 70 ข้อ
- ชุด 4 จำนวนข้อสอบ 80 ข้อ
- ชุด 5 จำนวนข้อสอบ 90 ข้อ
- ชุด 6 จำนวนข้อสอบ 100 ข้อ

จากภาพประกอบ 51 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 200 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 11.388 ; e = 0.800\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถสูง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 24.166 ; e = 1.200\}$

จากภาพประกอบ 52 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์ของกลุ่มตัวอย่างขนาด 300 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 12.823 ; e = 0.600\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 26.938 ; e = 0.600\}$

จากภาพประกอบ 53 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 400 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 10.396 ; \theta = 0.400\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 26.230 ; \theta = 0.400\}$

จากภาพประกอบ 54 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 11.912 ; e = 0.200\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 25.637 ; e = 0.400\}$

จากภาพประกอบ 55 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 600 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(e) = 10.326 ; e = 0.400\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 25.432 ; e = 0.200\}$

จากภาพประกอบ 56 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 700 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\theta) = 10.443 ; \theta = 0.400\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับ

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 25.224 ; \theta = 0.600\}$

จากภาพประกอบ 57 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 800 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 22.319 ; e = 0.400\}$

จากภาพประกอบ 58 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์ของกลุ่มตัวอย่างขนาด 900 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับ
 {I(๑) = 10.823 ; ๑ = 0.000} เมื่อข้อสอบมีขนาด 100 ข้อ {I(๑) = 21.914 ; ๑ = 0.200}

จากภาพประกอบ 59 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 1,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\alpha) = 23.171 ; \alpha = 0.400\}$

จากภาพประกอบ 60 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 1,500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\theta) = 16.568 ; \theta = 0.000\}$

จากภาพประกอบ 61 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 2,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

ความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(e) = 17.027 ; e = 0.000\}$

จากภาพประกอบ 62 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 2,500 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง

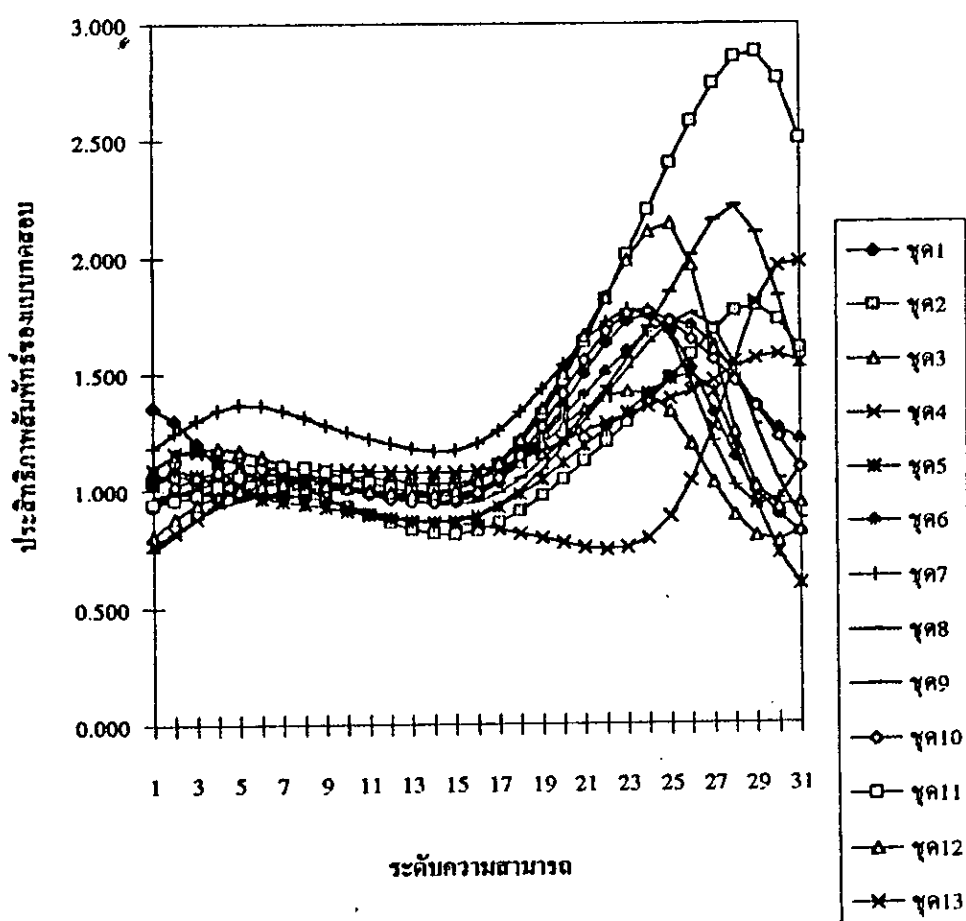
$\{I(\alpha) = 10.957 ; \alpha = 0.000\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\alpha) = 17.739 ; \alpha = 0.000\}$

จากภาพประกอบ 63 พบว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ของกลุ่มตัวอย่างขนาด 3,000 คน มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\alpha) = 12.887 ; \alpha = 0.000\}$ เมื่อข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ $\{I(\alpha) = 18.140 ; \alpha = 0.000\}$

ดังนั้น จากภาพประกอบ 51 - 63 แสดงว่า เมื่อจำนวนข้อสอบมีขนาดต่างกัน ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ มีค่าต่ำสุด ที่ระดับความสามารถปานกลาง $\{I(\alpha) = 8.016 ; \alpha = -0.200\}$ เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน ข้อสอบมีขนาด 50 ข้อ และมีค่าสูงสุด ที่ระดับความสามารถปานกลาง เมื่อข้อสอบมีขนาด 100 ข้อ ขนาดของกลุ่มตัวอย่าง 400 คน $\{I(\alpha) = 26.230 ; \alpha = 0.400\}$

7. ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีรติค และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ

7.1 ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 64 - 69

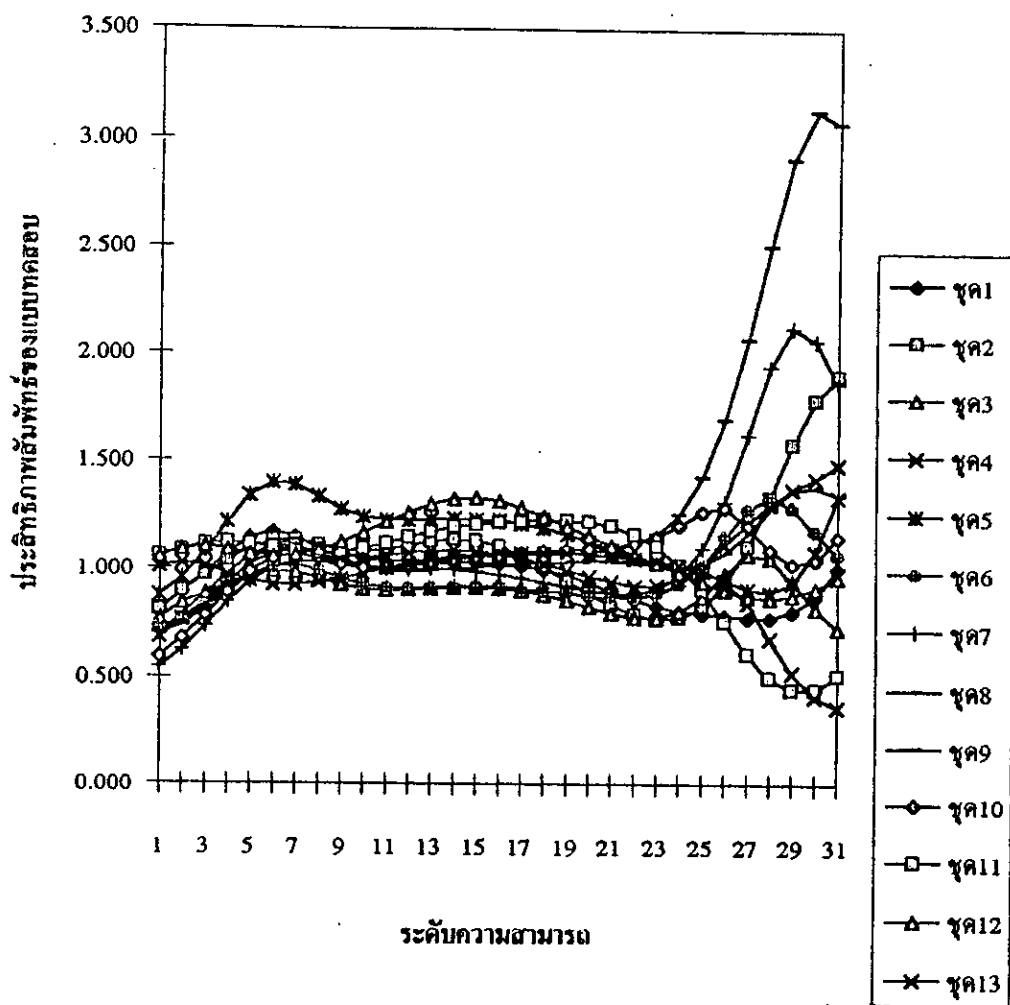


ภาพประกอบ 64 เปรียบเทียบได้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแบบกลุ่ม ใดก็ได้ เมื่อข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

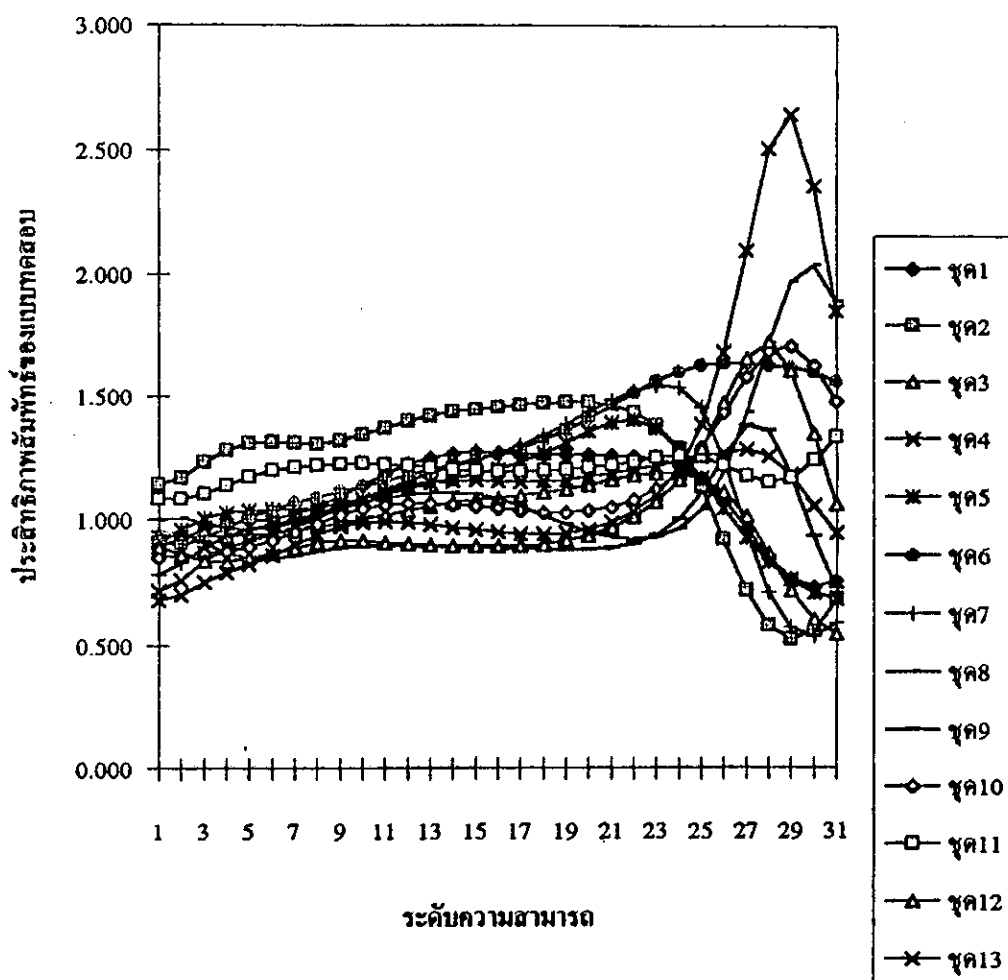


ภาพประกอบ 64 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

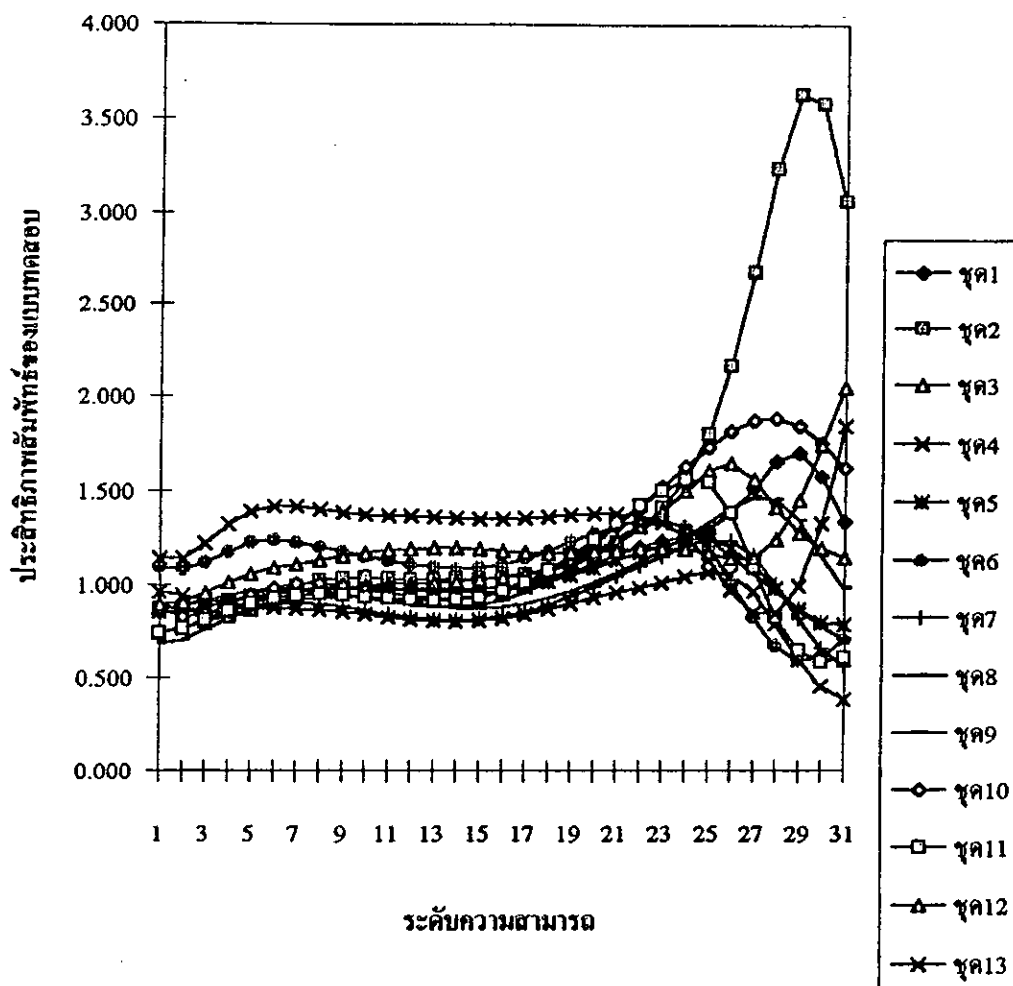
ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน



ภาพประกอบ 64 (ต่อ)

หมายเหตุ

ชด 1	จำนวน	800/400 คน	ชด 8	จำนวน	600/500 คน
ชด 2	จำนวน	900/400 คน	ชด 9	จำนวน	700/500 คน
ชด 3	จำนวน	1,000/400 คน	ชด 10	จำนวน	800/500 คน
ชด 4	จำนวน	1,500/400 คน	ชด 11	จำนวน	900/500 คน
ชด 5	จำนวน	2,000/400 คน	ชด 12	จำนวน	1,000/500 คน
ชด 6	จำนวน	2,500/400 คน	ชด 13	จำนวน	1,500/500 คน
ชด 7	จำนวน	3,000/400 คน			

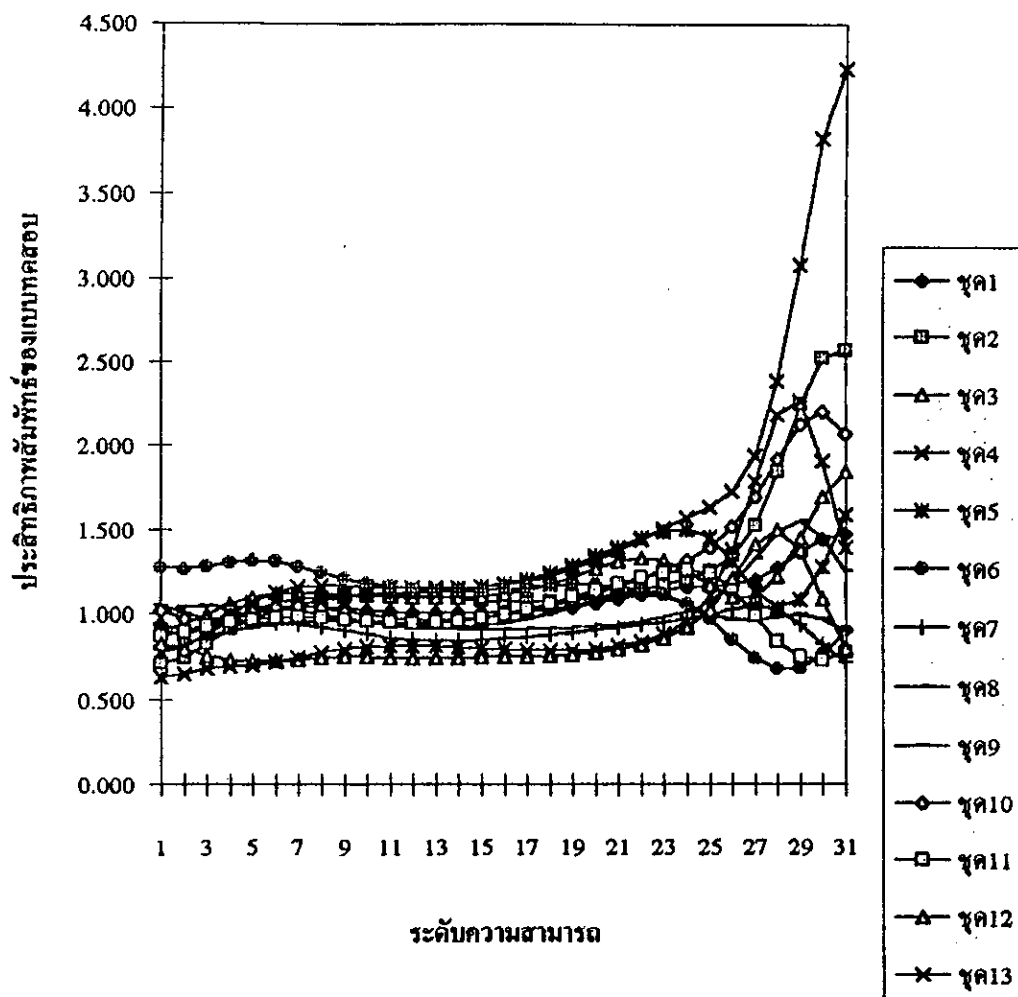


ภาพประกอบ 64 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 2,000/500 คน
 ชด 2 จำนวน 2,500/500 คน
 ชด 3 จำนวน 3,000/500 คน
 ชด 4 จำนวน 700/600 คน
 ชด 5 จำนวน 800/600 คน
 ชด 6 จำนวน 900/600 คน
 ชด 7 จำนวน 1,000/600 คน

ชด 8 จำนวน 1,500/600 คน
 ชด 9 จำนวน 2,000/600 คน
 ชด 10 จำนวน 2,500/600 คน
 ชด 11 จำนวน 3,000/600 คน
 ชด 12 จำนวน 800/700 คน
 ชด 13 จำนวน 900/700 คน

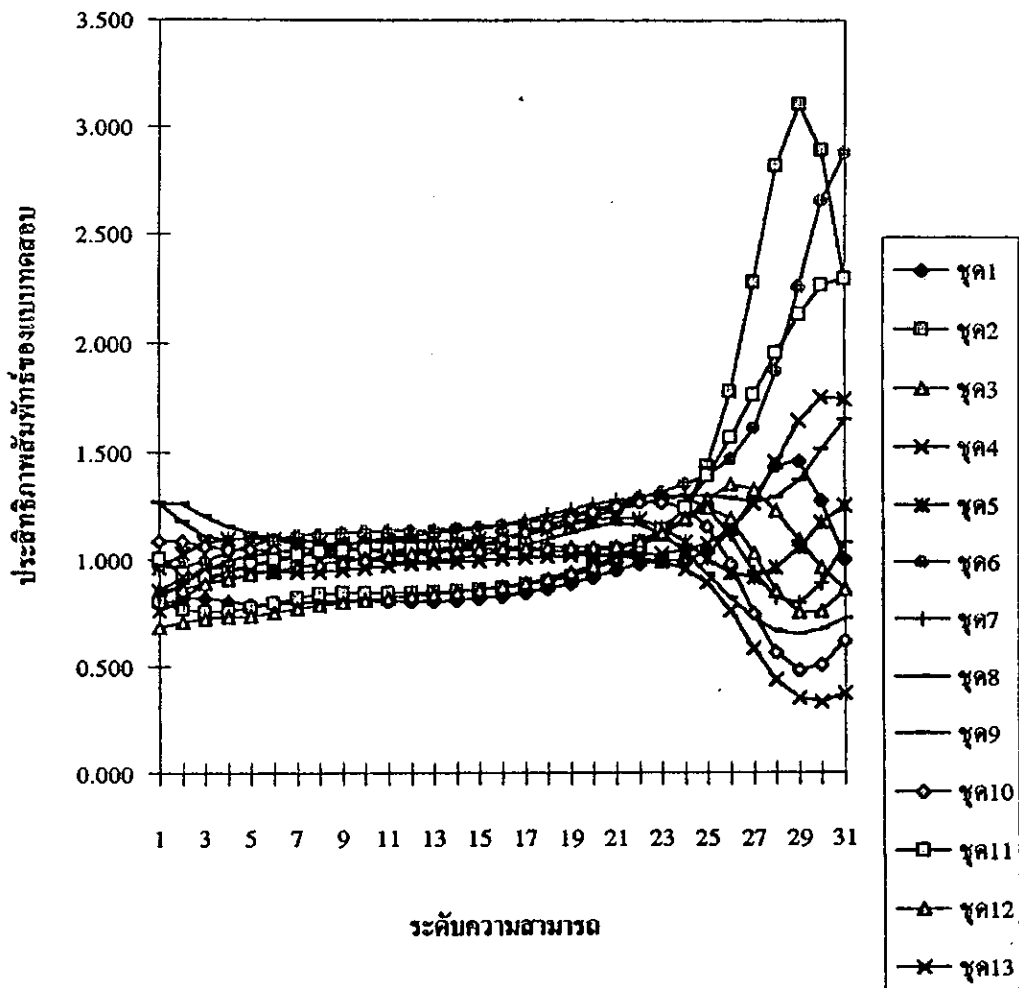


ภาพประกอบ 64 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

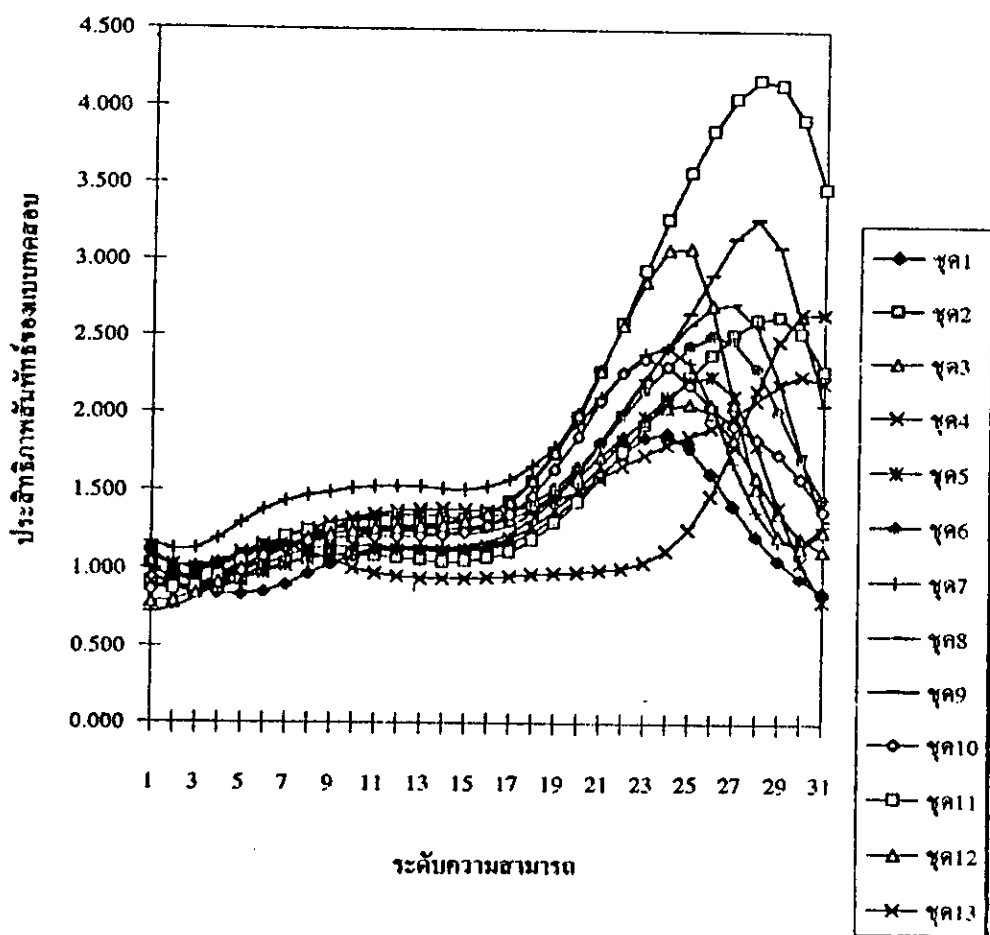


ภาพประกอบ 64 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 2,000/900 คน
 ชด 2 จำนวน 2,500/900 คน
 ชด 3 จำนวน 3,000/900 คน
 ชด 4 จำนวน 1,500/1,000 คน
 ชด 5 จำนวน 2,000/1,000 คน
 ชด 6 จำนวน 2,500/1,000 คน
 ชด 7 จำนวน 3,000/1,000 คน

ชด 8 จำนวน 2,000/1,500 คน
 ชด 9 จำนวน 2,500/1,500 คน
 ชด 10 จำนวน 3,000/1,500 คน
 ชด 11 จำนวน 2,500/2,000 คน
 ชด 12 จำนวน 3,000/2,000 คน
 ชด 13 จำนวน 3,000/2,500 คน

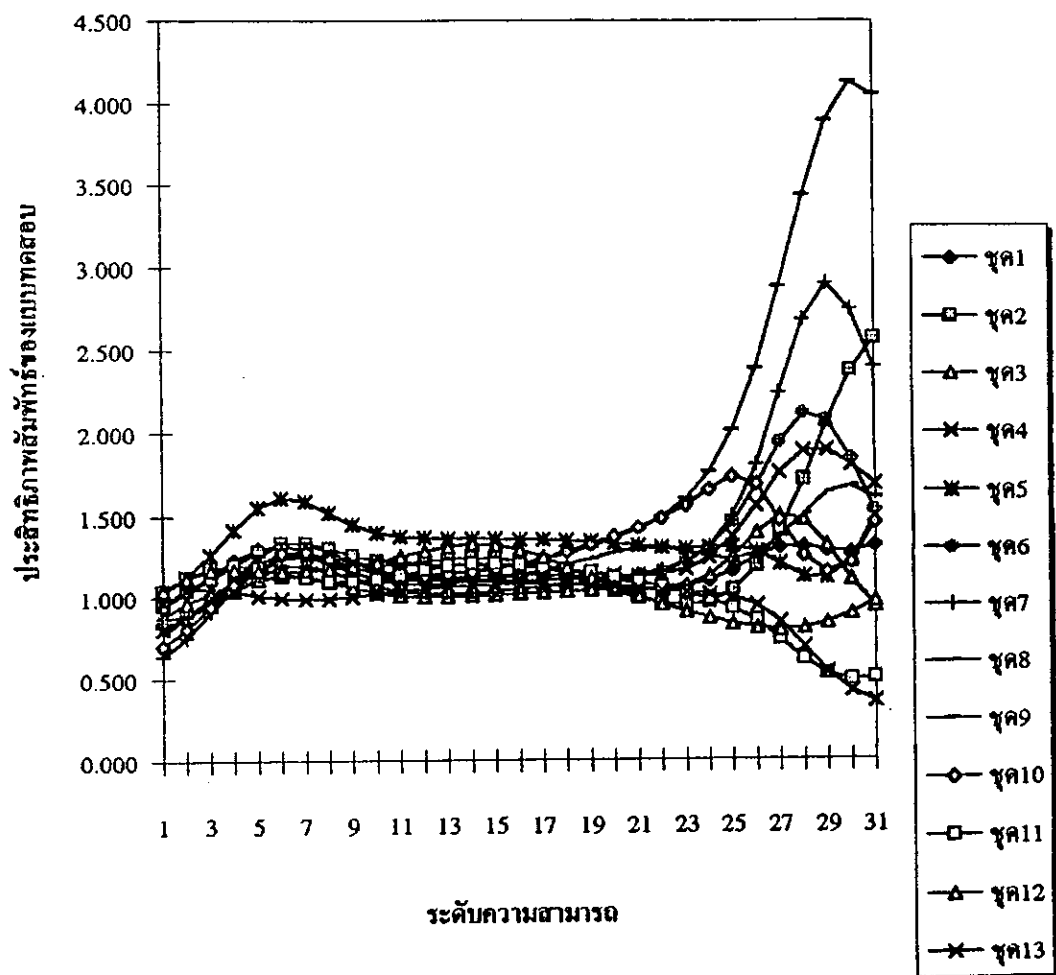


ภาพประกอบ 65 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

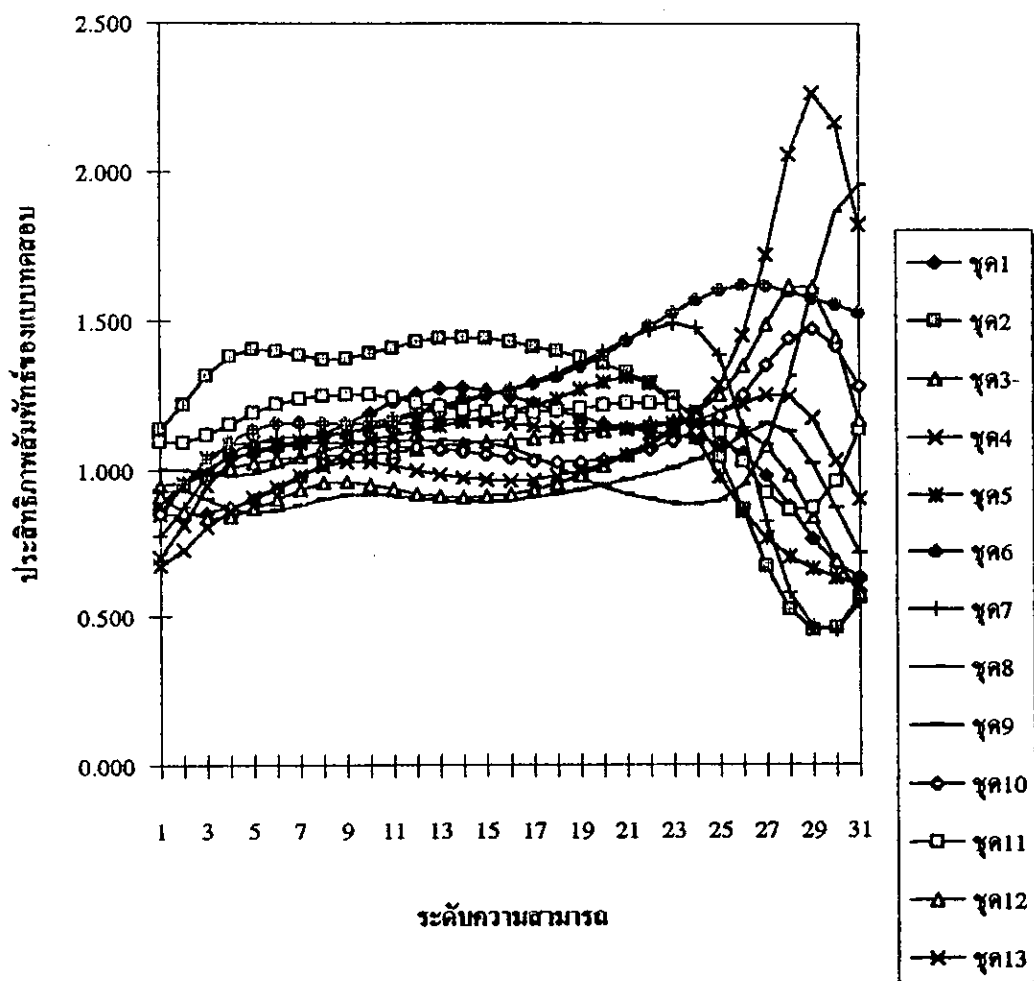


ภาพประกอบ 65 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 500/300 คน
 ชด 2 จำนวน 600/300 คน
 ชด 3 จำนวน 700/300 คน
 ชด 4 จำนวน 800/300 คน
 ชด 5 จำนวน 900/300 คน
 ชด 6 จำนวน 1,000/300 คน
 ชด 7 จำนวน 1,500/300 คน

ชด 8 จำนวน 2,000/300 คน
 ชด 9 จำนวน 2,500/300 คน
 ชด 10 จำนวน 3,000/300 คน
 ชด 11 จำนวน 500/400 คน
 ชด 12 จำนวน 600/400 คน
 ชด 13 จำนวน 700/400 คน

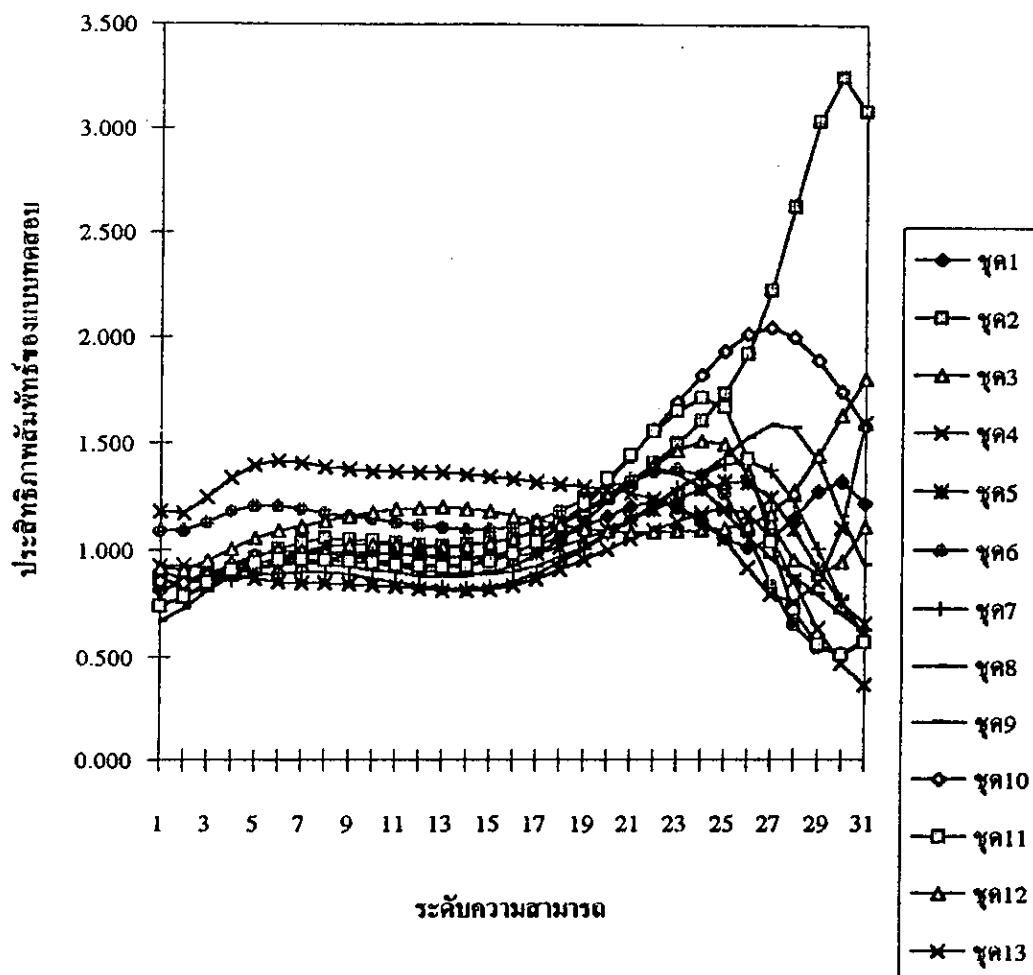


ภาพประกอบ 65 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

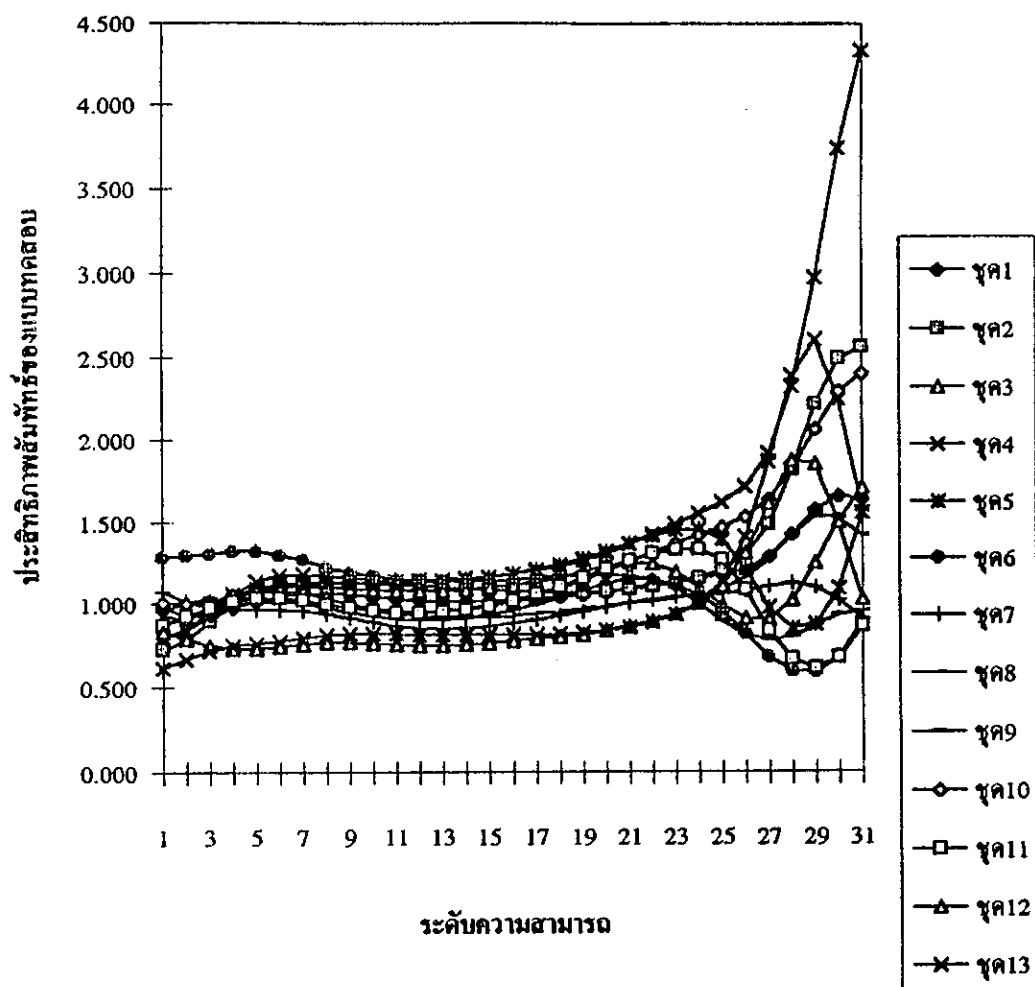


ภาพประกอบ 65 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

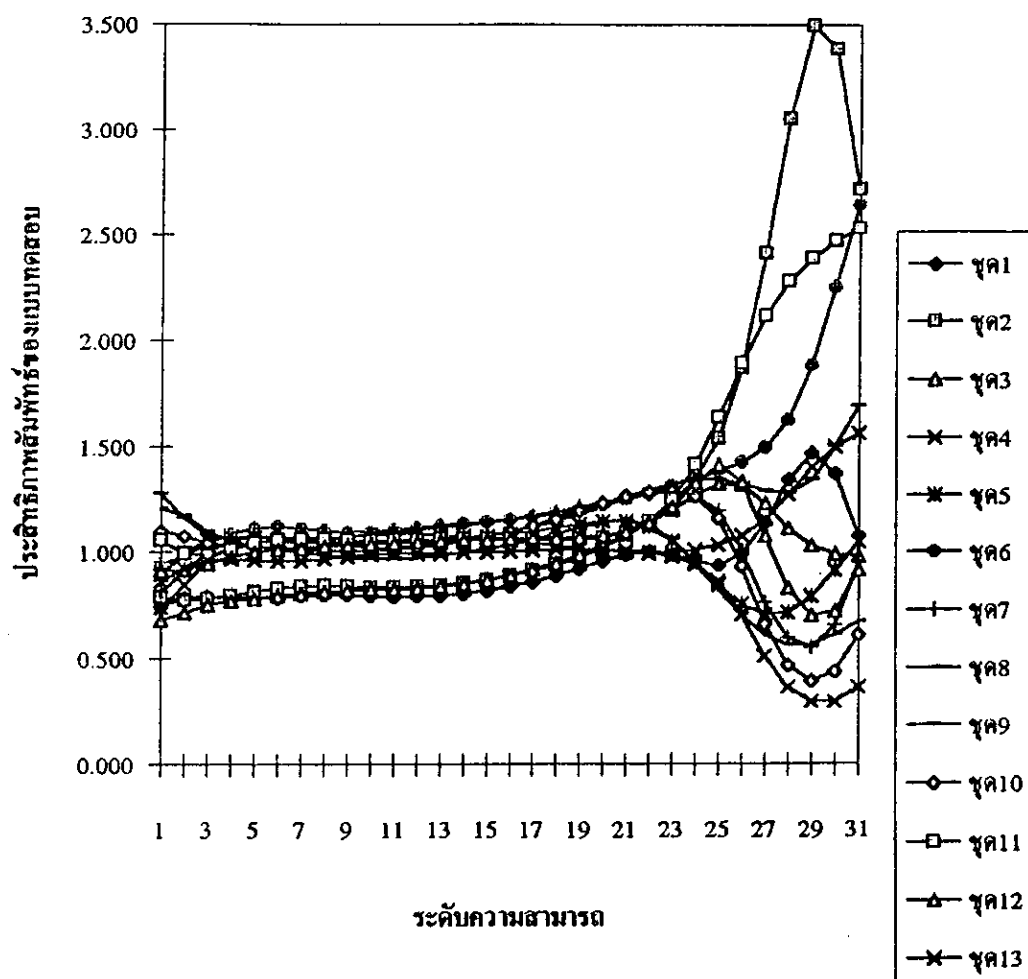


ภาพประกอบ 65 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 1,000/700 คน
ชด 2 จำนวน 1,500/700 คน
ชด 3 จำนวน 2,000/700 คน
ชด 4 จำนวน 2,500/700 คน
ชด 5 จำนวน 3,000/700 คน
ชด 6 จำนวน 900/800 คน
ชด 7 จำนวน 1,000/800 คน

ชด 8 จำนวน 1,500/800 คน
ชด 9 จำนวน 2,000/800 คน
ชด 10 จำนวน 2,500/800 คน
ชด 11 จำนวน 3,000/800 คน
ชด 12 จำนวน 1,000/900 คน
ชด 13 จำนวน 1,500/900 คน

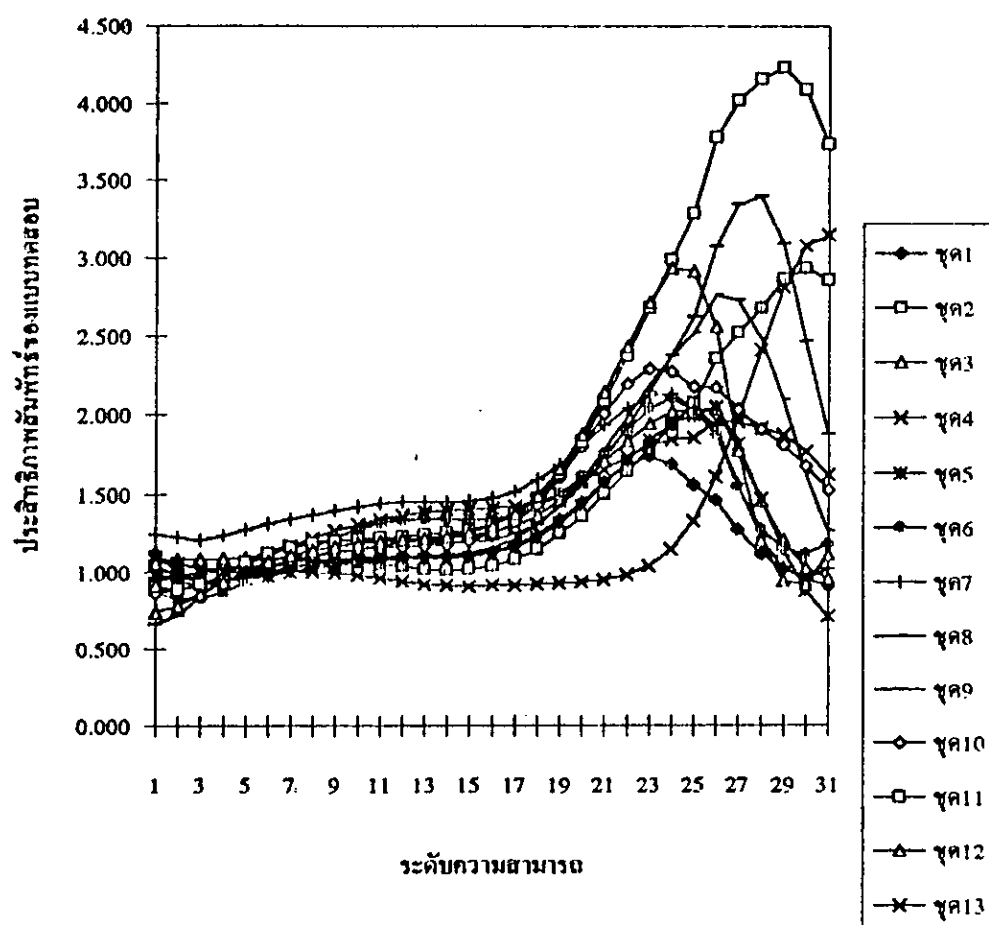


ภาพประกอบ 65 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 2,000/900 คน
 ชด 2 จำนวน 2,500/900 คน
 ชด 3 จำนวน 3,000/900 คน
 ชด 4 จำนวน 1,500/1,000 คน
 ชด 5 จำนวน 2,000/1,000 คน
 ชด 6 จำนวน 2,500/1,000 คน
 ชด 7 จำนวน 3,000/1,000 คน

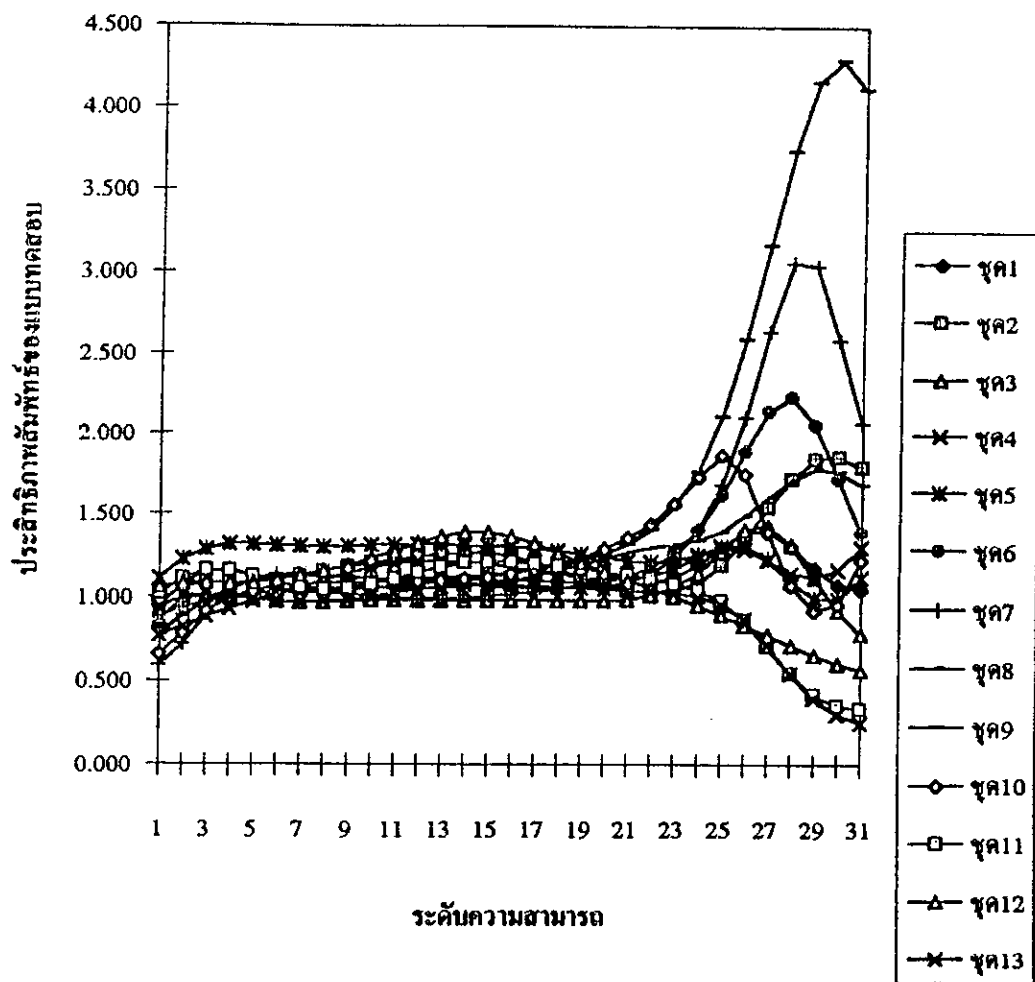
ชด 8 จำนวน 2,000/1,500 คน
 ชด 9 จำนวน 2,500/1,500 คน
 ชด 10 จำนวน 3,000/1,500 คน
 ชด 11 จำนวน 2,500/2,000 คน
 ชด 12 จำนวน 3,000/2,000 คน
 ชด 13 จำนวน 3,000/2,500 คน



ภาพประกอบ 66 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิสต์ เมื่อข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน	ชุด 8 จำนวน 1,000/200 คน
ชุด 2 จำนวน 400/200 คน	ชุด 9 จำนวน 1,500/200 คน
ชุด 3 จำนวน 500/200 คน	ชุด 10 จำนวน 2,000/200 คน
ชุด 4 จำนวน 600/200 คน	ชุด 11 จำนวน 2,500/200 คน
ชุด 5 จำนวน 700/200 คน	ชุด 12 จำนวน 3,000/200 คน
ชุด 6 จำนวน 800/200 คน	ชุด 13 จำนวน 400/300 คน
ชุด 7 จำนวน 900/200 คน	

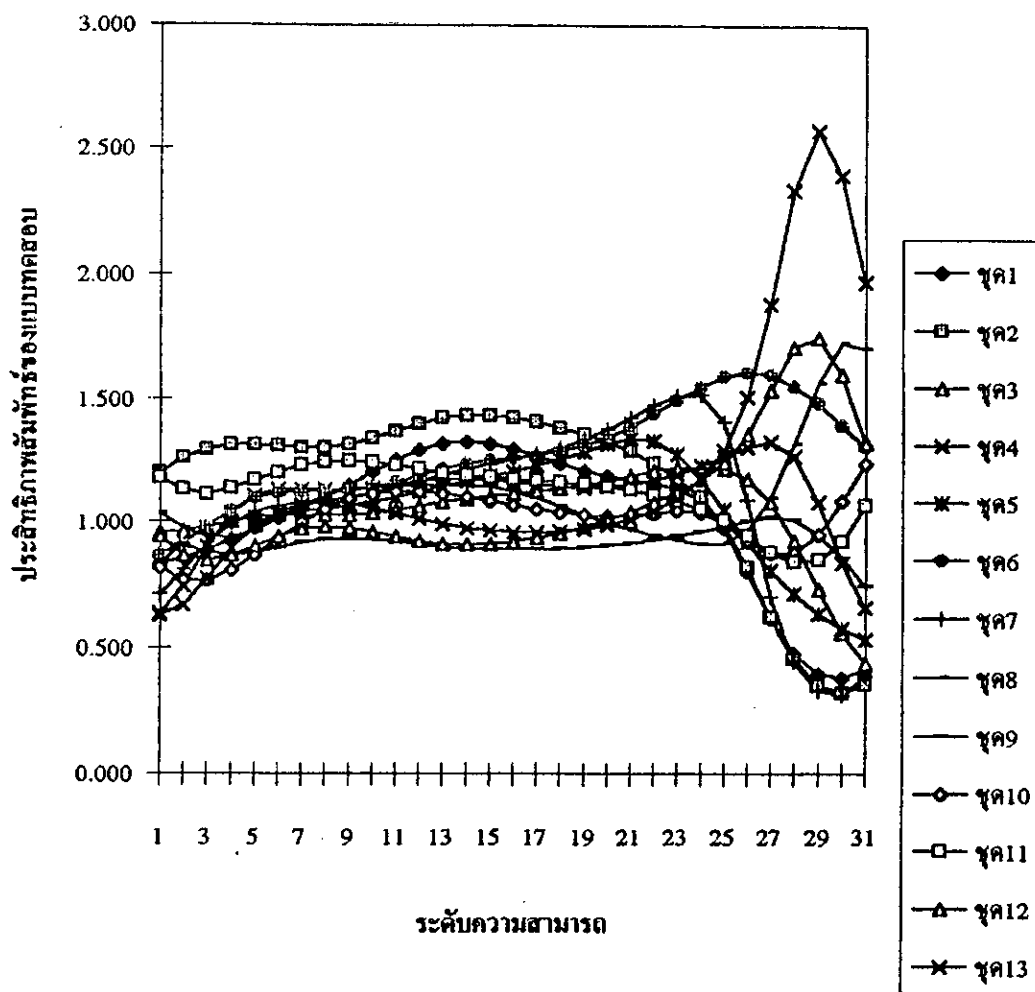


ภาพประกอบ 66 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

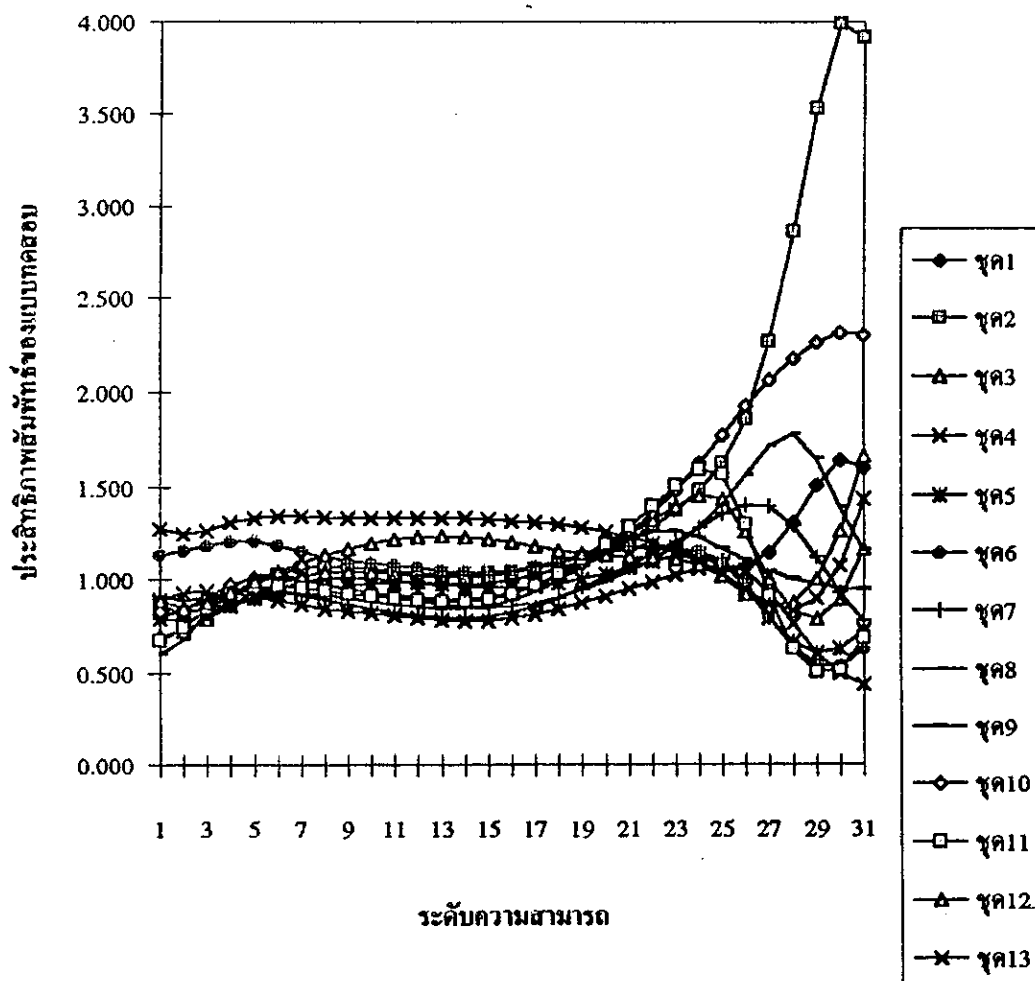


ภาพประกอบ 66 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

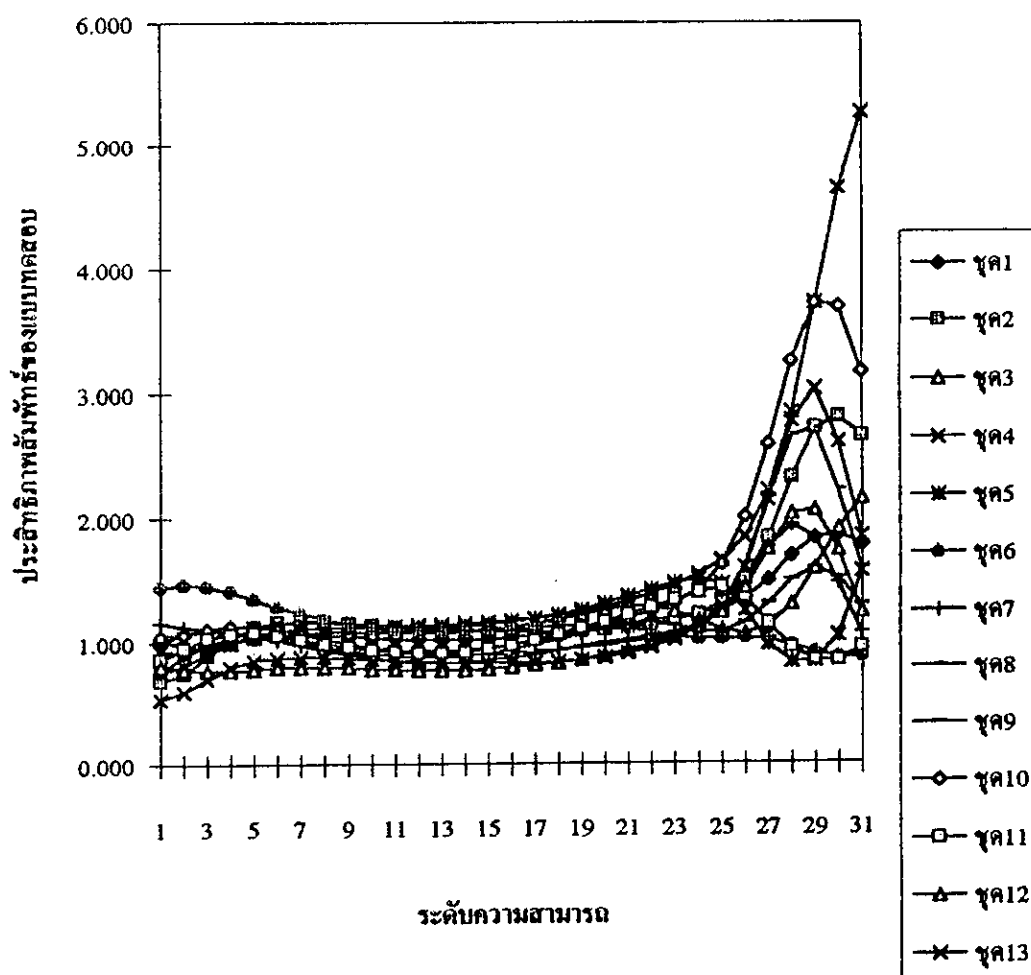


ภาพประกอบ 66 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

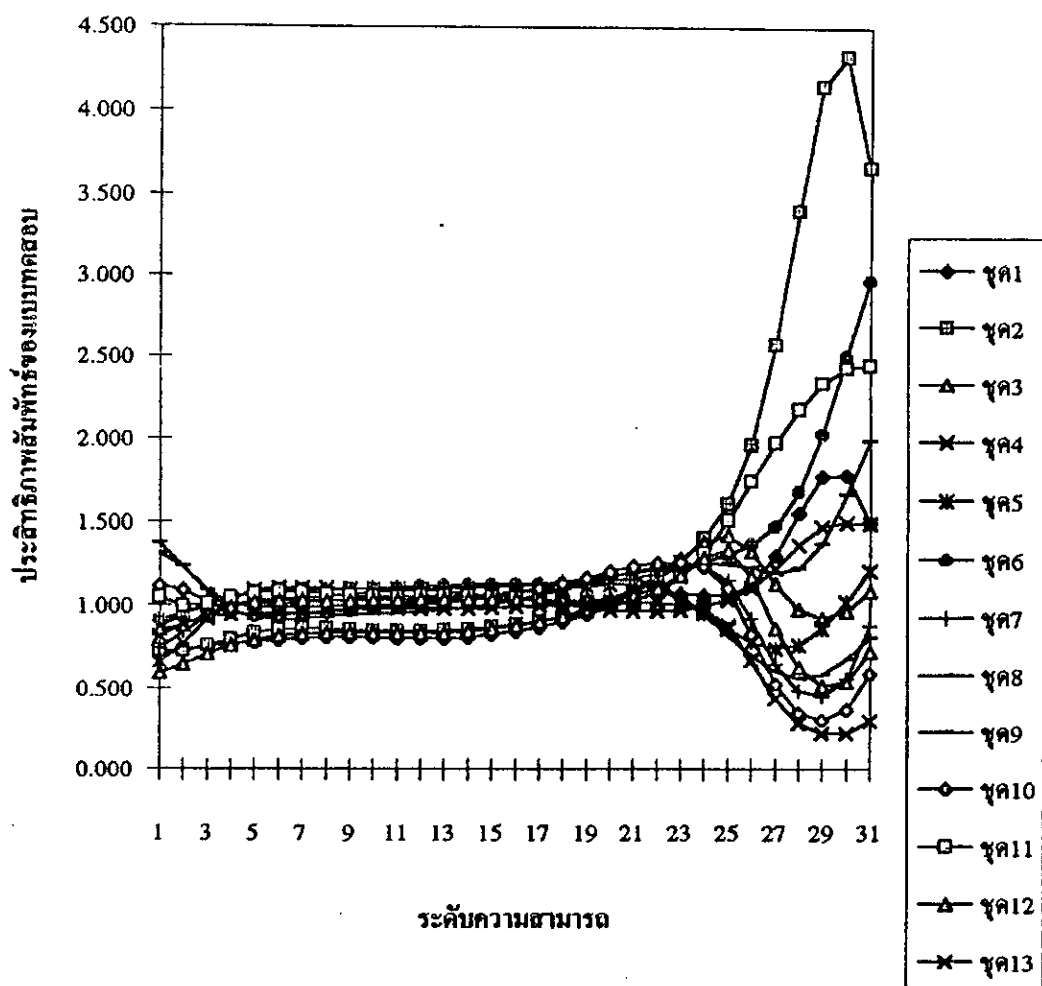


ภาพประกอบ 66 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

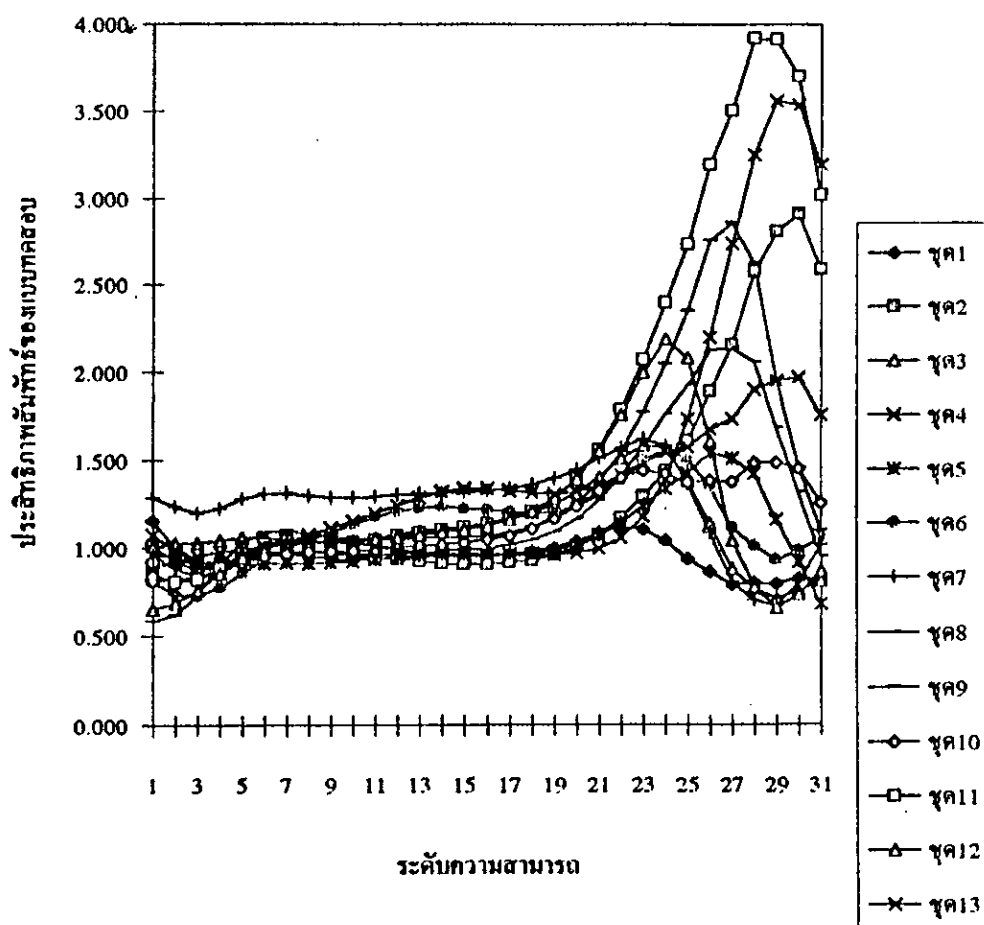


ภาพประกอบ 66 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

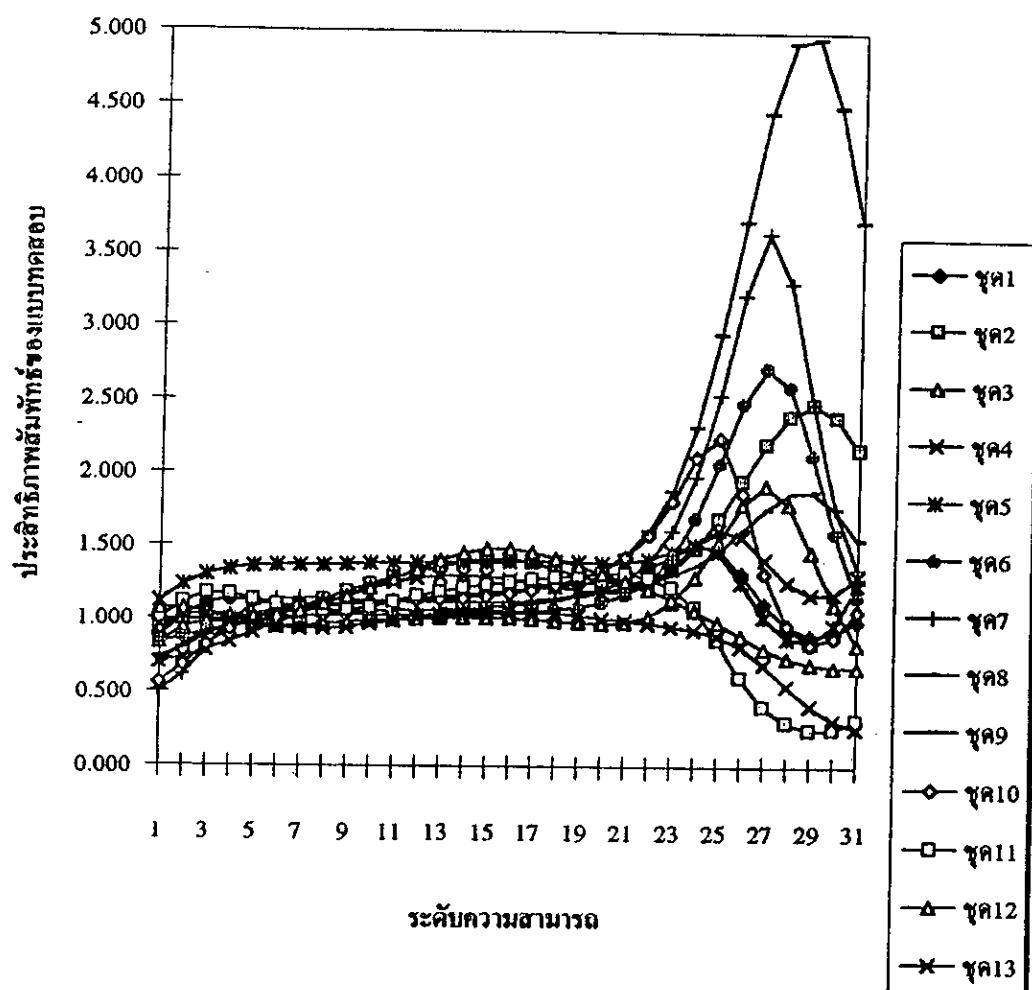


ภาพประกอบ 67 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีแมกซ์ิม ไลค์ลิสต์ เมื่อข้อสอบมีขนาด 80 ข้อ และกลุ่ม
ตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 300/200 คน
ชด 2 จำนวน 400/200 คน
ชด 3 จำนวน 500/200 คน
ชด 4 จำนวน 600/200 คน
ชด 5 จำนวน 700/200 คน
ชด 6 จำนวน 800/200 คน
ชด 7 จำนวน 900/200 คน

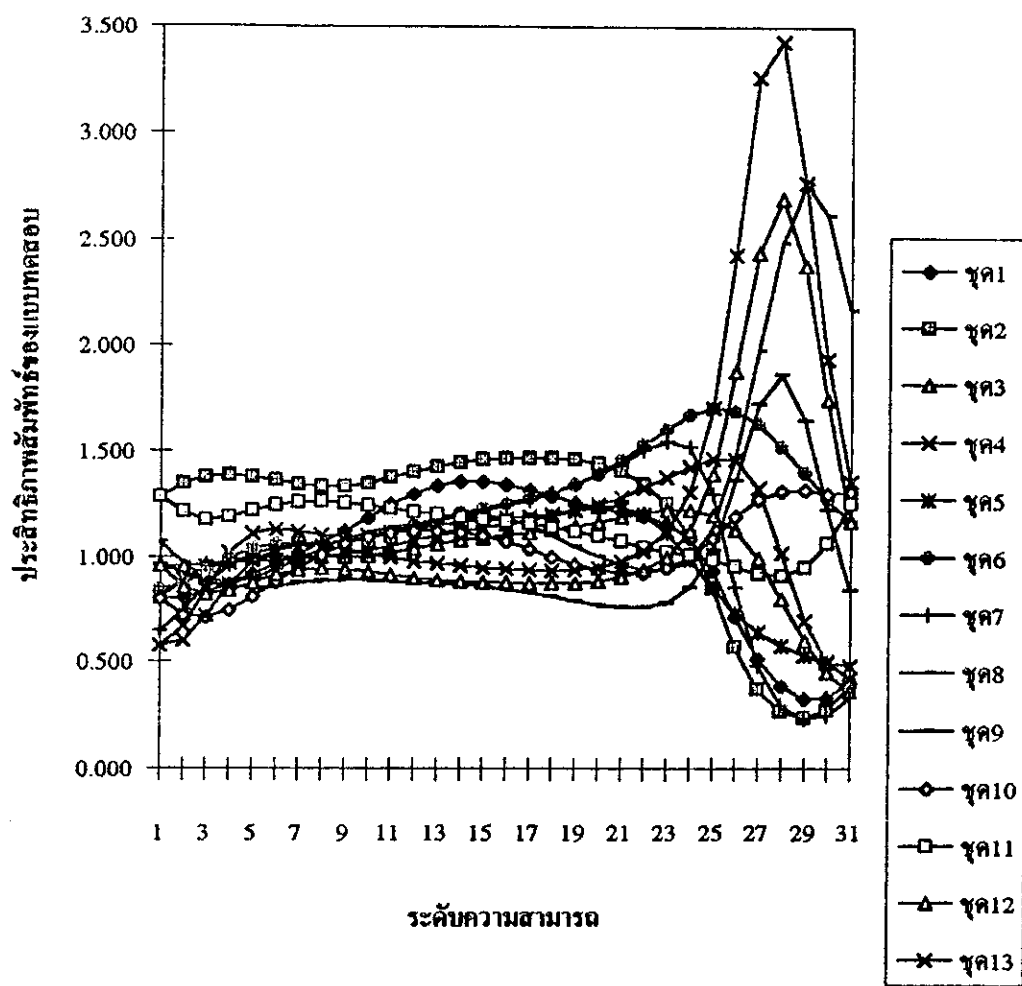
ชด 8 จำนวน 1,000/200 คน
ชด 9 จำนวน 1,500/200 คน
ชด 10 จำนวน 2,000/200 คน
ชด 11 จำนวน 2,500/200 คน
ชด 12 จำนวน 3,000/200 คน
ชด 13 จำนวน 400/300 คน



ภาพประกอบ 67 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน	ชุด 8 จำนวน 2,000/300 คน
ชุด 2 จำนวน 600/300 คน	ชุด 9 จำนวน 2,500/300 คน
ชุด 3 จำนวน 700/300 คน	ชุด 10 จำนวน 3,000/300 คน
ชุด 4 จำนวน 800/300 คน	ชุด 11 จำนวน 500/400 คน
ชุด 5 จำนวน 900/300 คน	ชุด 12 จำนวน 600/400 คน
ชุด 6 จำนวน 1,000/300 คน	ชุด 13 จำนวน 700/400 คน
ชุด 7 จำนวน 1,500/300 คน	

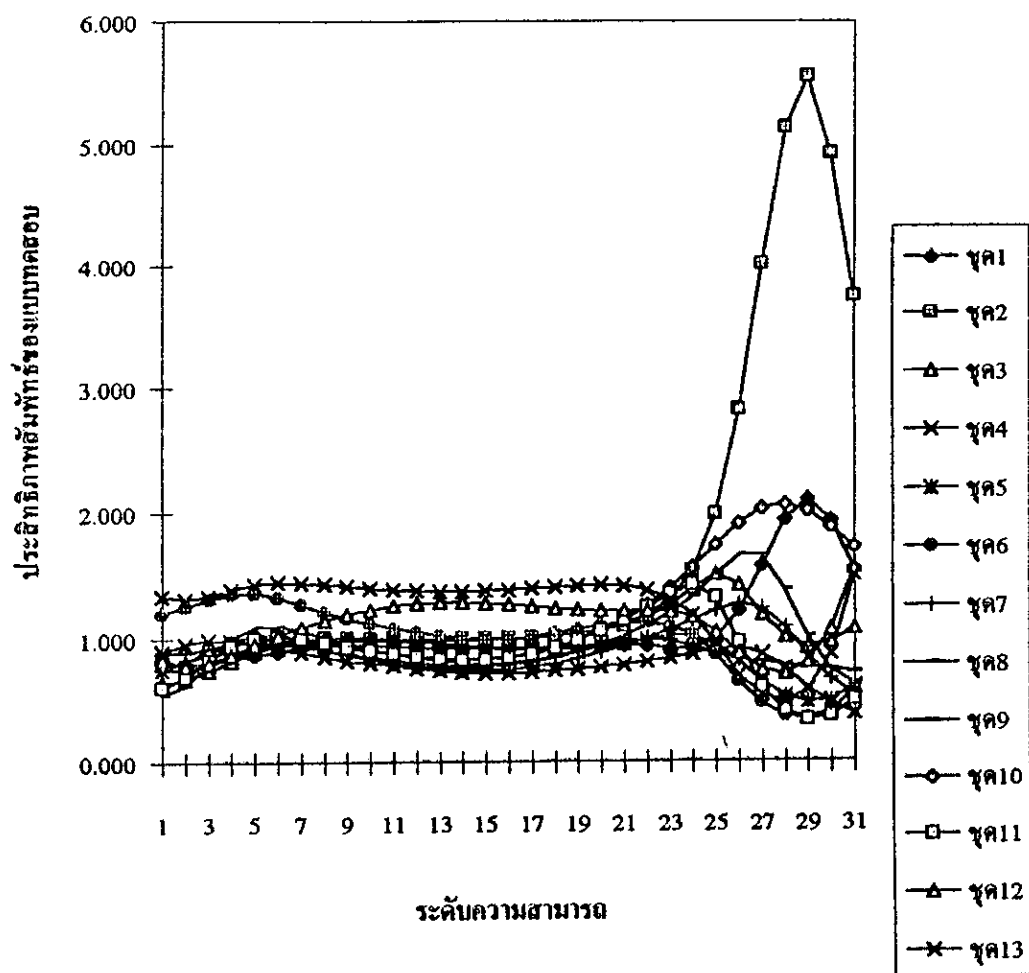


ภาพประกอบ 67 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

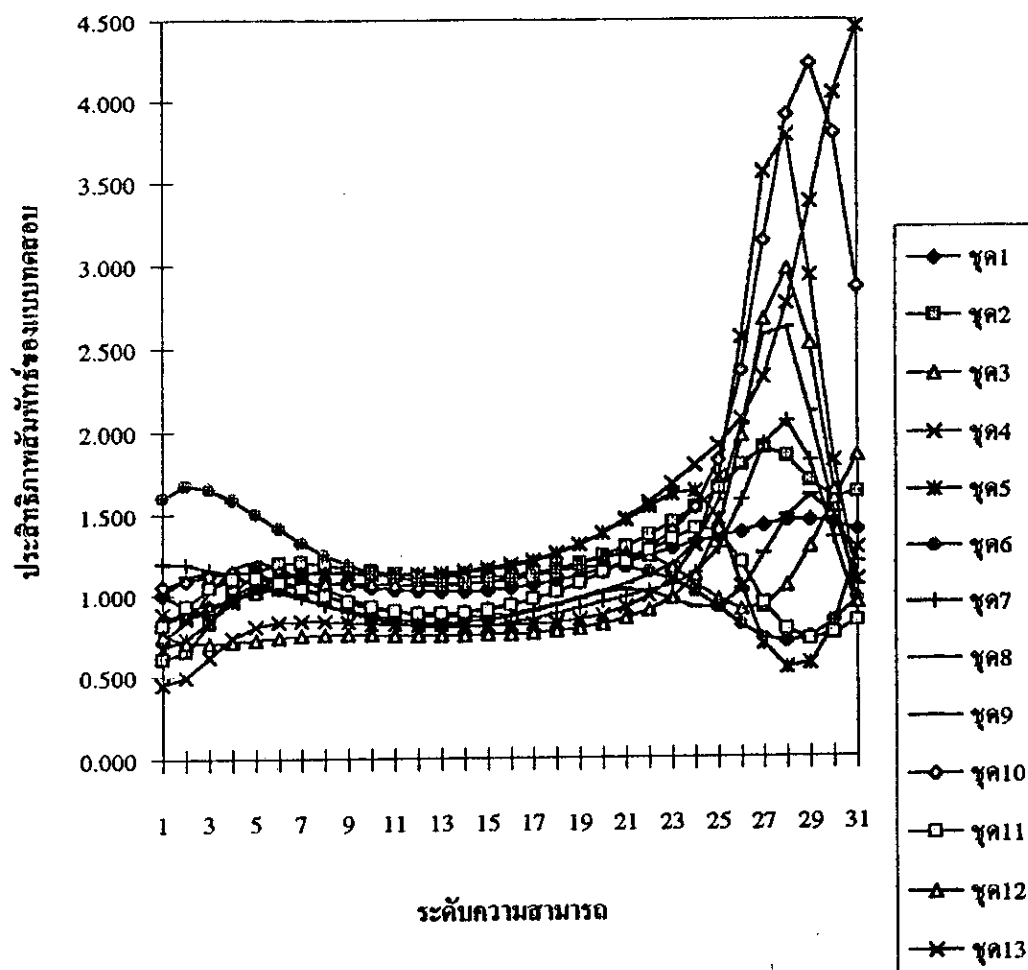


ภาพประกอบ 67 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

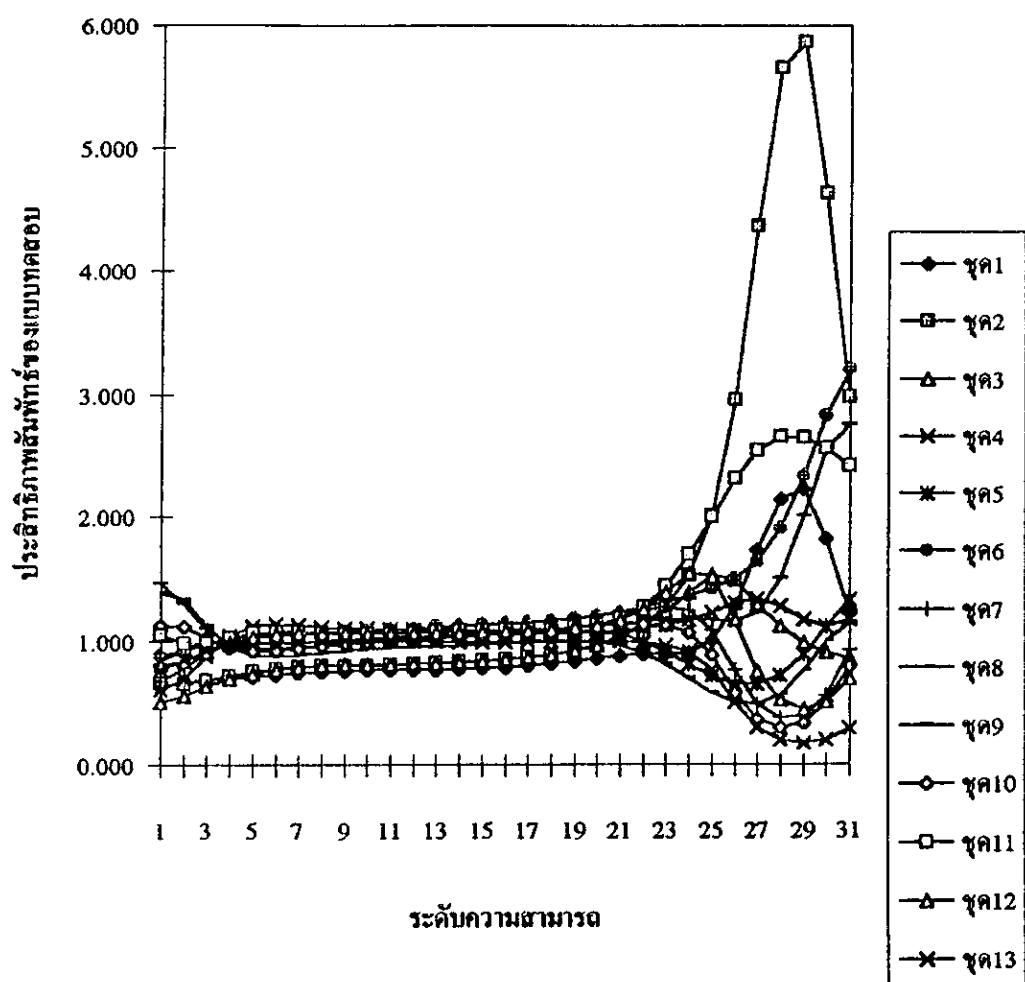


ภาพประกอบ 67 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 1,000/700 คน
 ชด 2 จำนวน 1,500/700 คน
 ชด 3 จำนวน 2,000/700 คน
 ชด 4 จำนวน 2,500/700 คน
 ชด 5 จำนวน 3,000/700 คน
 ชด 6 จำนวน 900/800 คน
 ชด 7 จำนวน 1,000/800 คน

ชด 8 จำนวน 1,500/800 คน
 ชด 9 จำนวน 2,000/800 คน
 ชด 10 จำนวน 2,500/800 คน
 ชด 11 จำนวน 3,000/800 คน
 ชด 12 จำนวน 1,000/900 คน
 ชด 13 จำนวน 1,500/900 คน

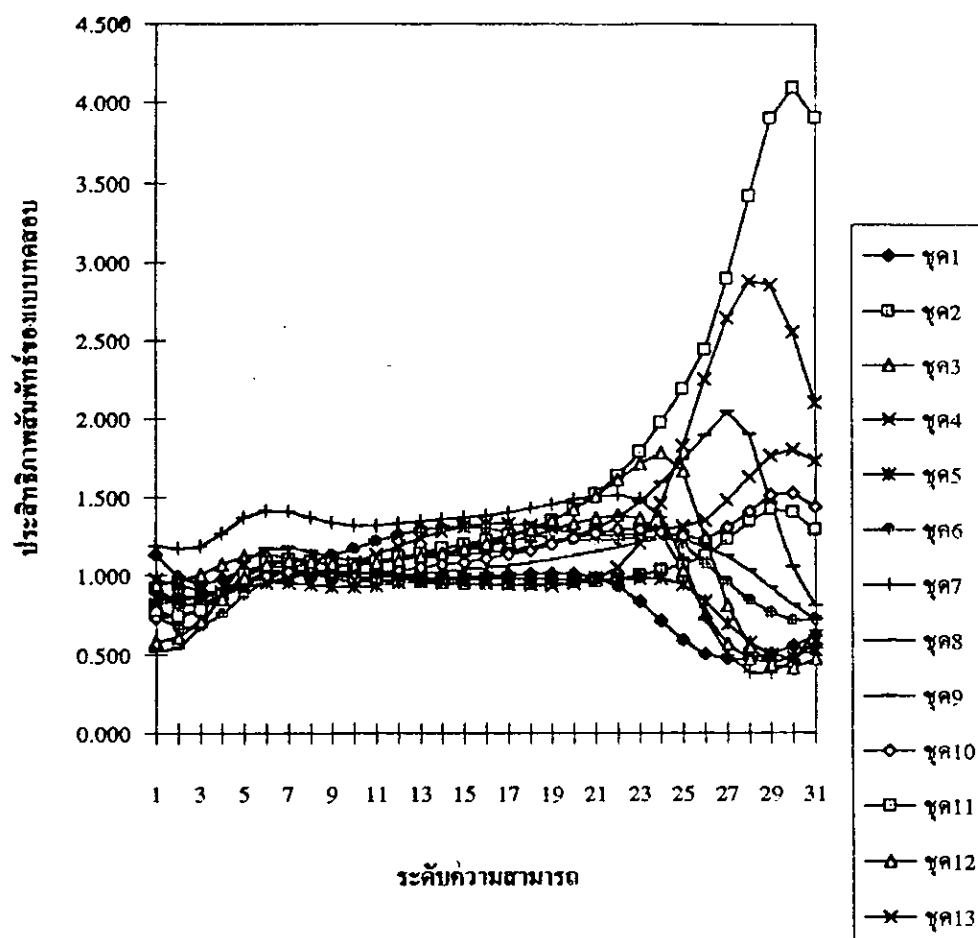


ภาพประกอบ 67 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

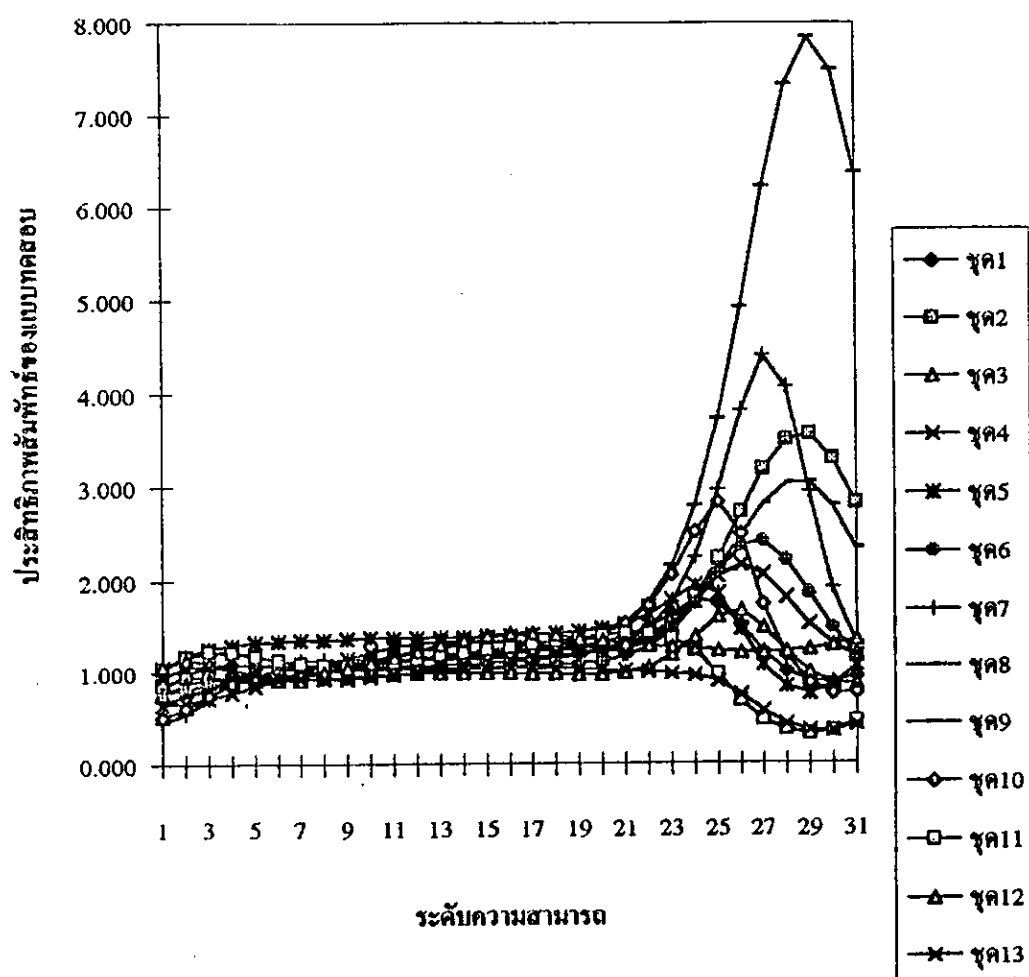


ภาพประกอบ 68 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซ์ิมัม โลคัลลิสต์ เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

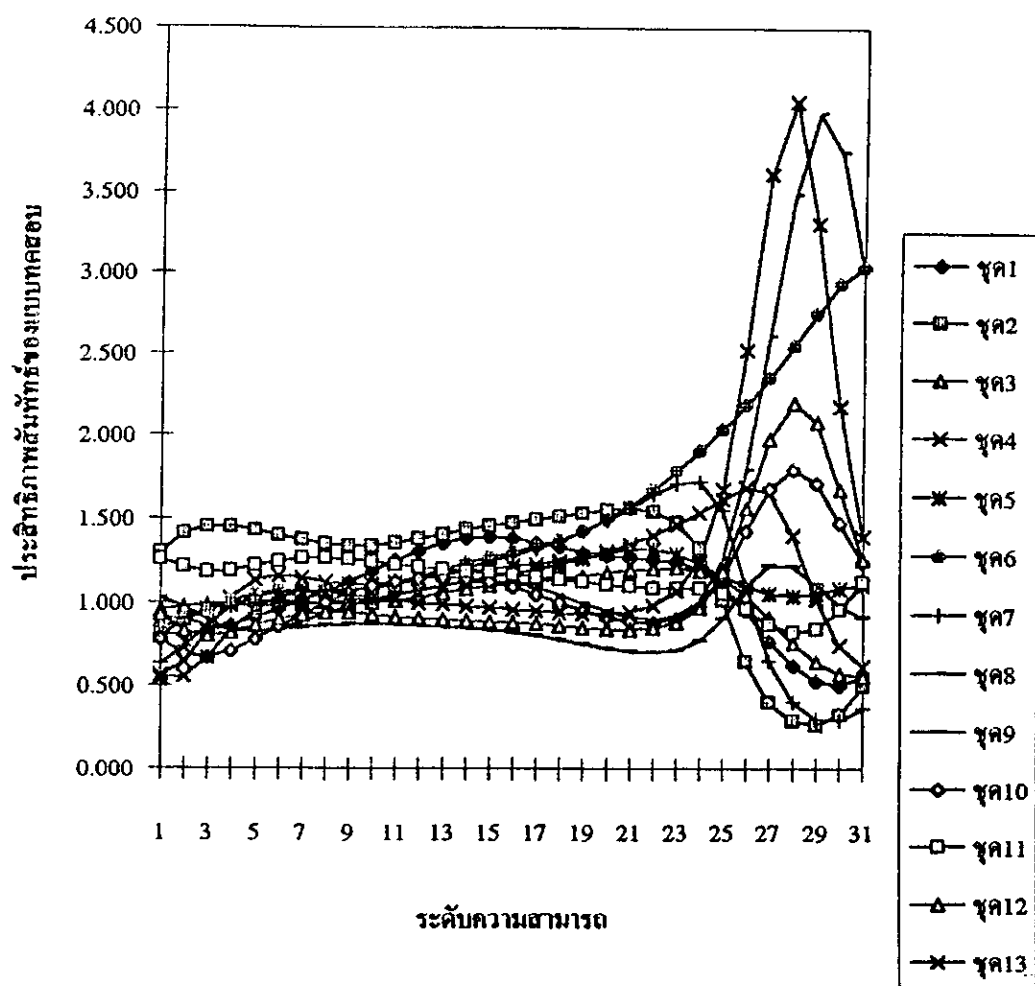


ภาพประกอบ 68 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
ชุด 2 จำนวน 600/300 คน
ชุด 3 จำนวน 700/300 คน
ชุด 4 จำนวน 800/300 คน
ชุด 5 จำนวน 900/300 คน
ชุด 6 จำนวน 1,000/300 คน
ชุด 7 จำนวน 1,500/300 คน

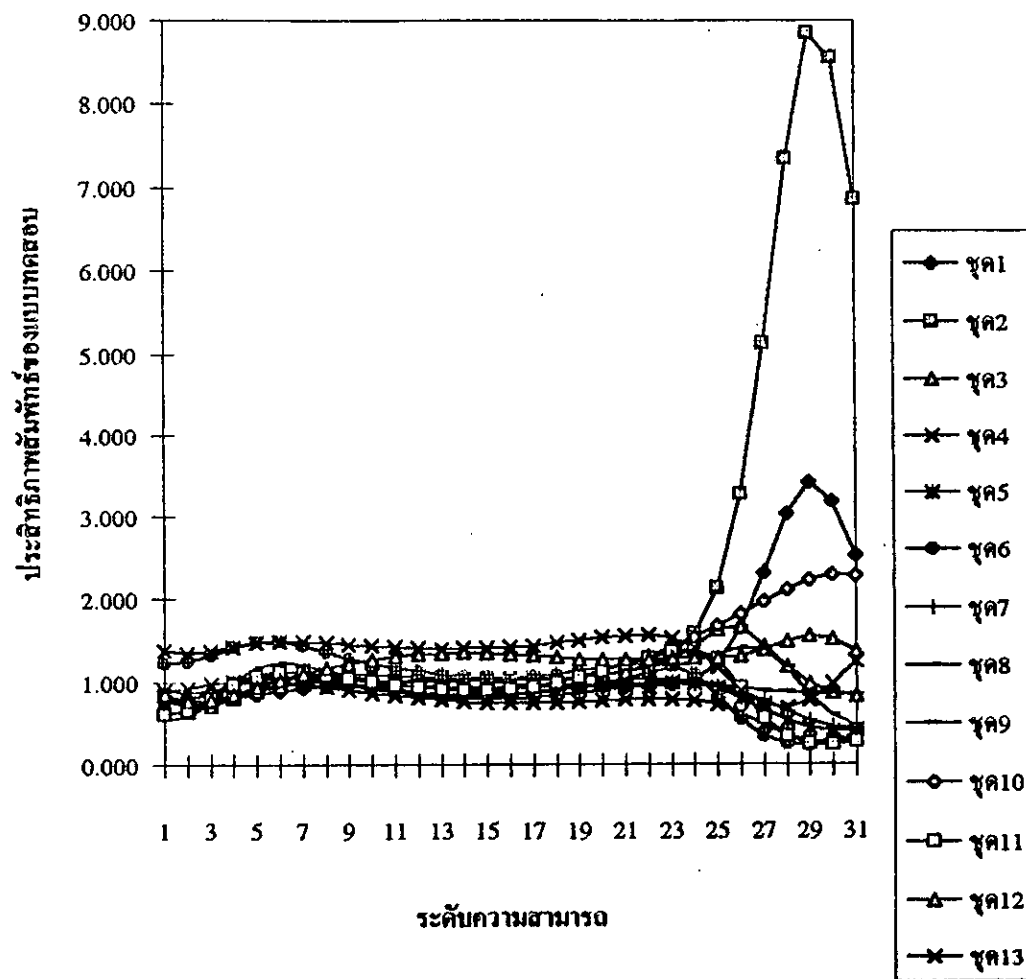
ชุด 8 จำนวน 2,000/300 คน
ชุด 9 จำนวน 2,500/300 คน
ชุด 10 จำนวน 3,000/300 คน
ชุด 11 จำนวน 500/400 คน
ชุด 12 จำนวน 600/400 คน
ชุด 13 จำนวน 700/400 คน



ภาพประกอบ 68 (ต่อ)

หมายเหตุ

ชุด 1	จำนวน	800/400 คน	ชุด 8	จำนวน	600/500 คน
ชุด 2	จำนวน	900/400 คน	ชุด 9	จำนวน	700/500 คน
ชุด 3	จำนวน	1,000/400 คน	ชุด 10	จำนวน	800/500 คน
ชุด 4	จำนวน	1,500/400 คน	ชุด 11	จำนวน	900/500 คน
ชุด 5	จำนวน	2,000/400 คน	ชุด 12	จำนวน	1,000/500 คน
ชุด 6	จำนวน	2,500/400 คน	ชุด 13	จำนวน	1,500/500 คน
ชุด 7	จำนวน	3,000/400 คน			

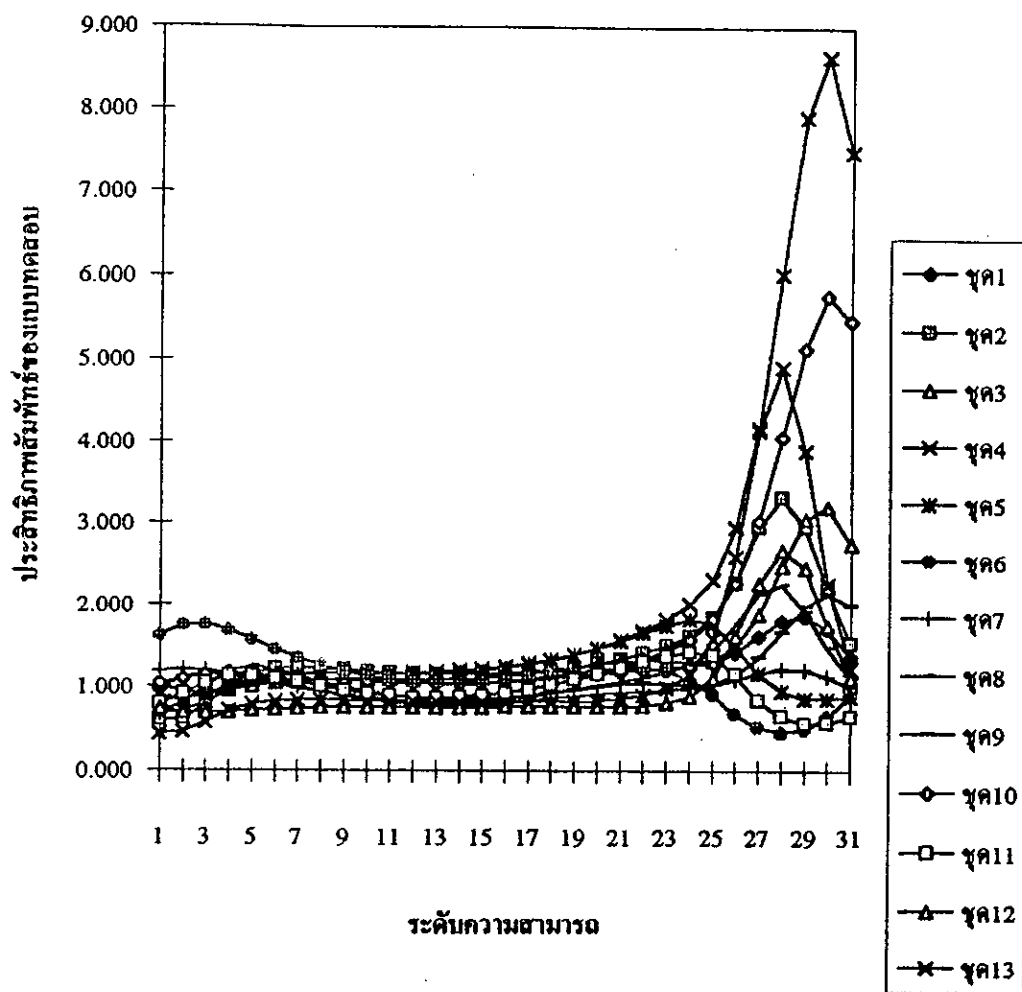


ภาพประกอบ 68 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

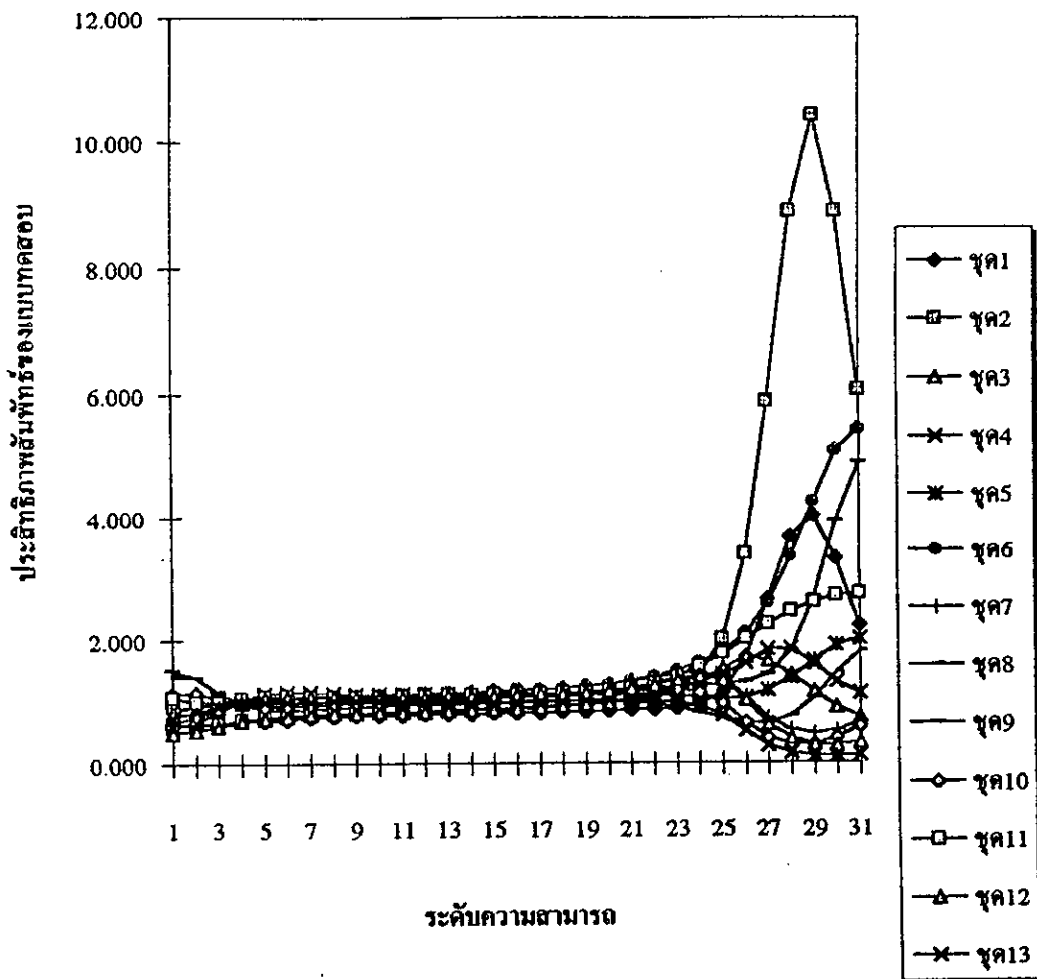


ภาพประกอบ 68 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

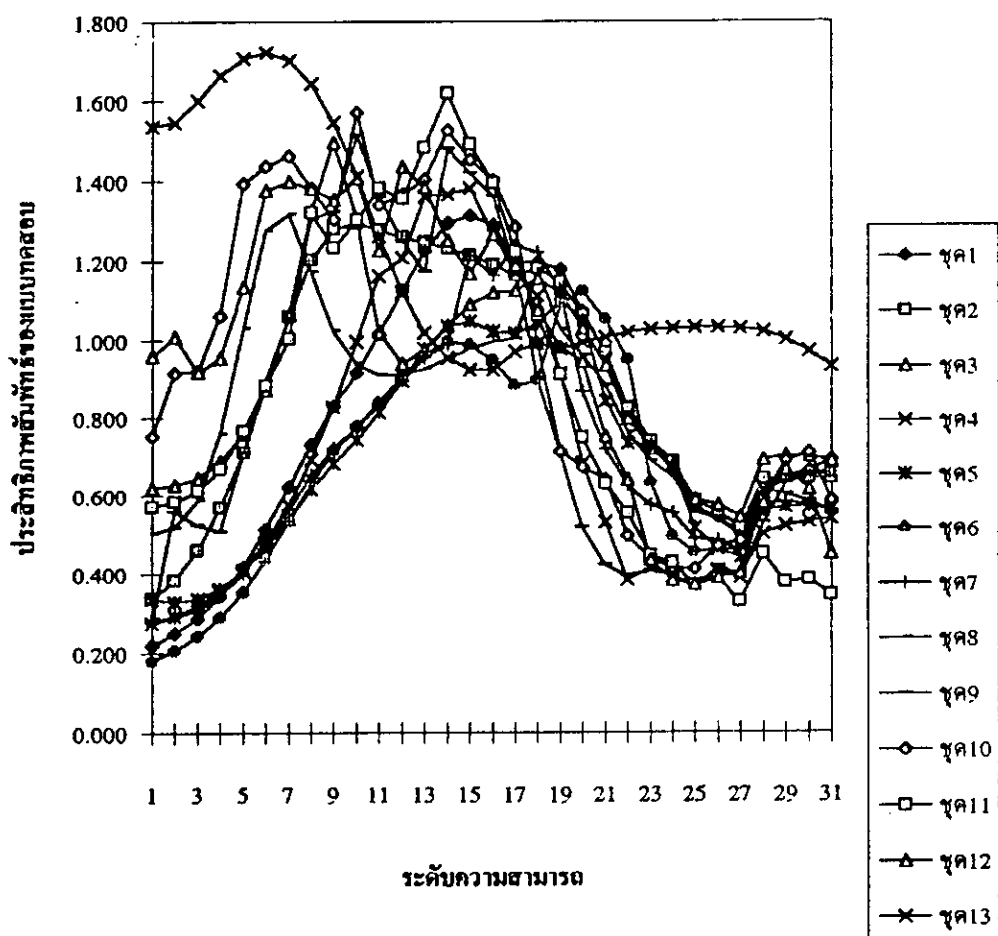


ภาพประกอบ 68 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

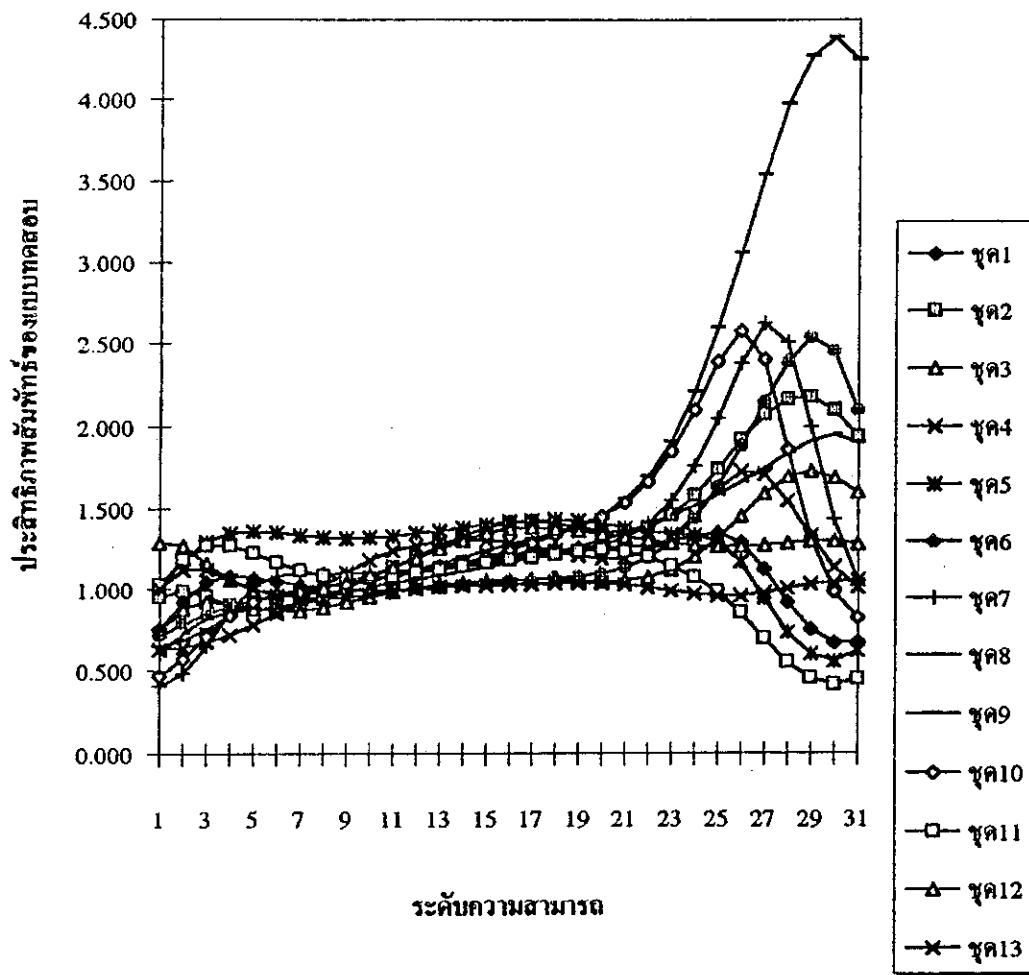
ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน



ภาพประกอบ 69 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่ม
ตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

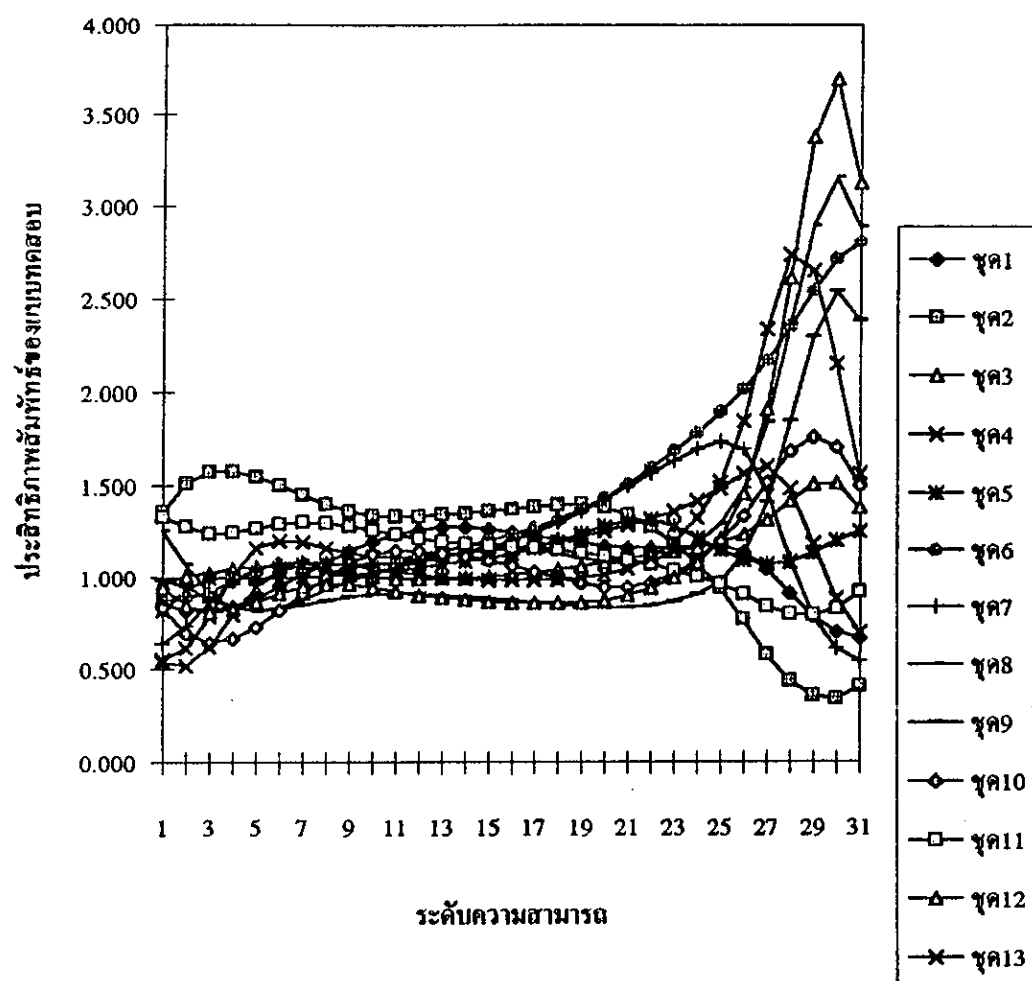
ชุด 1 จำนวน 300/200 คน	ชุด 8 จำนวน 1,000/200 คน
ชุด 2 จำนวน 400/200 คน	ชุด 9 จำนวน 1,500/200 คน
ชุด 3 จำนวน 500/200 คน	ชุด 10 จำนวน 2,000/200 คน
ชุด 4 จำนวน 600/200 คน	ชุด 11 จำนวน 2,500/200 คน
ชุด 5 จำนวน 700/200 คน	ชุด 12 จำนวน 3,000/200 คน
ชุด 6 จำนวน 800/200 คน	ชุด 13 จำนวน 400/300 คน
ชุด 7 จำนวน 900/200 คน	



ภาพประกอบ 69 (ต่อ)

หมายเหตุ

ชด 1	จำนวน	500/300 คน	ชด 8	จำนวน	2,000/300 คน
ชด 2	จำนวน	600/300 คน	ชด 9	จำนวน	2,500/300 คน
ชด 3	จำนวน	700/300 คน	ชด 10	จำนวน	3,000/300 คน
ชด 4	จำนวน	800/300 คน	ชด 11	จำนวน	500/400 คน
ชด 5	จำนวน	900/300 คน	ชด 12	จำนวน	600/400 คน
ชด 6	จำนวน	1,000/300 คน	ชด 13	จำนวน	700/400 คน
ชด 7	จำนวน	1,500/300 คน			

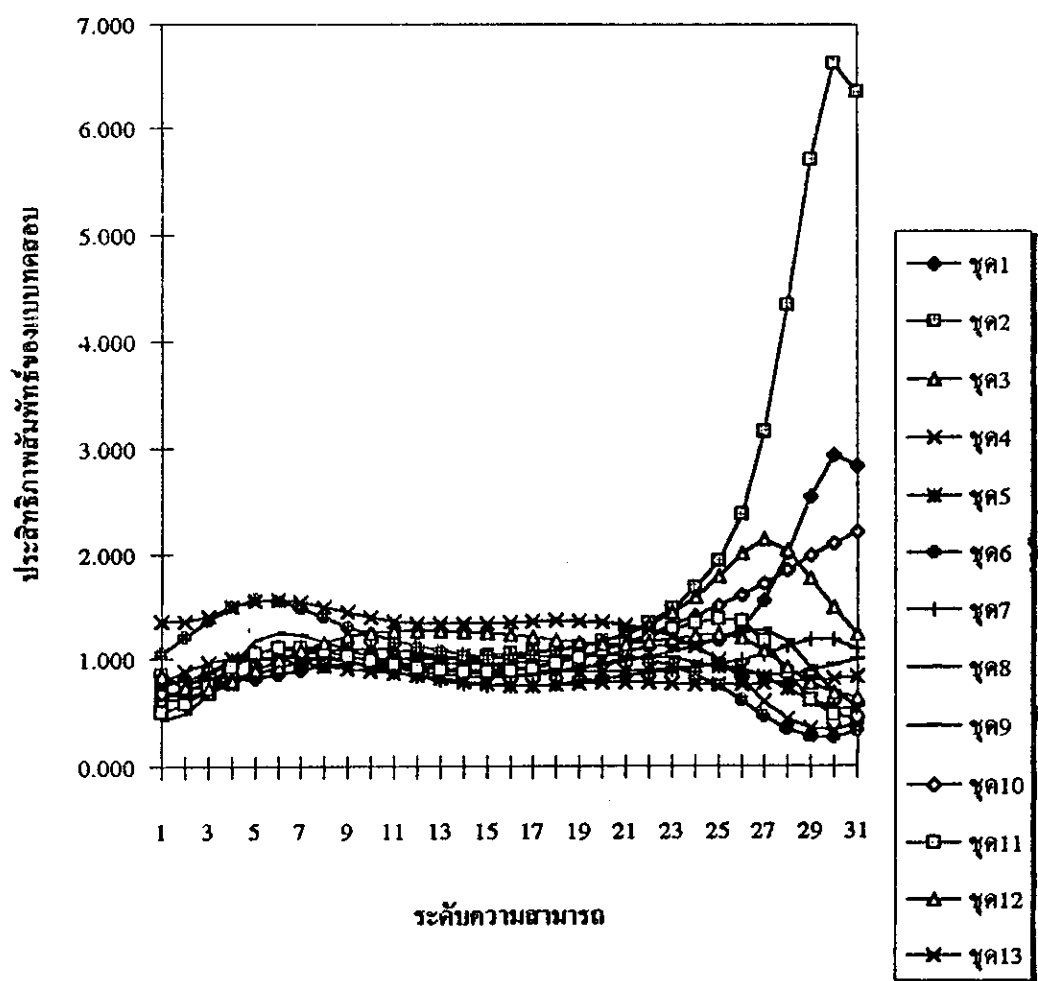


ภาพประกอบ 69 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 800/400 คน
 ชด 2 จำนวน 900/400 คน
 ชด 3 จำนวน 1,000/400 คน
 ชด 4 จำนวน 1,500/400 คน
 ชด 5 จำนวน 2,000/400 คน
 ชด 6 จำนวน 2,500/400 คน
 ชด 7 จำนวน 3,000/400 คน

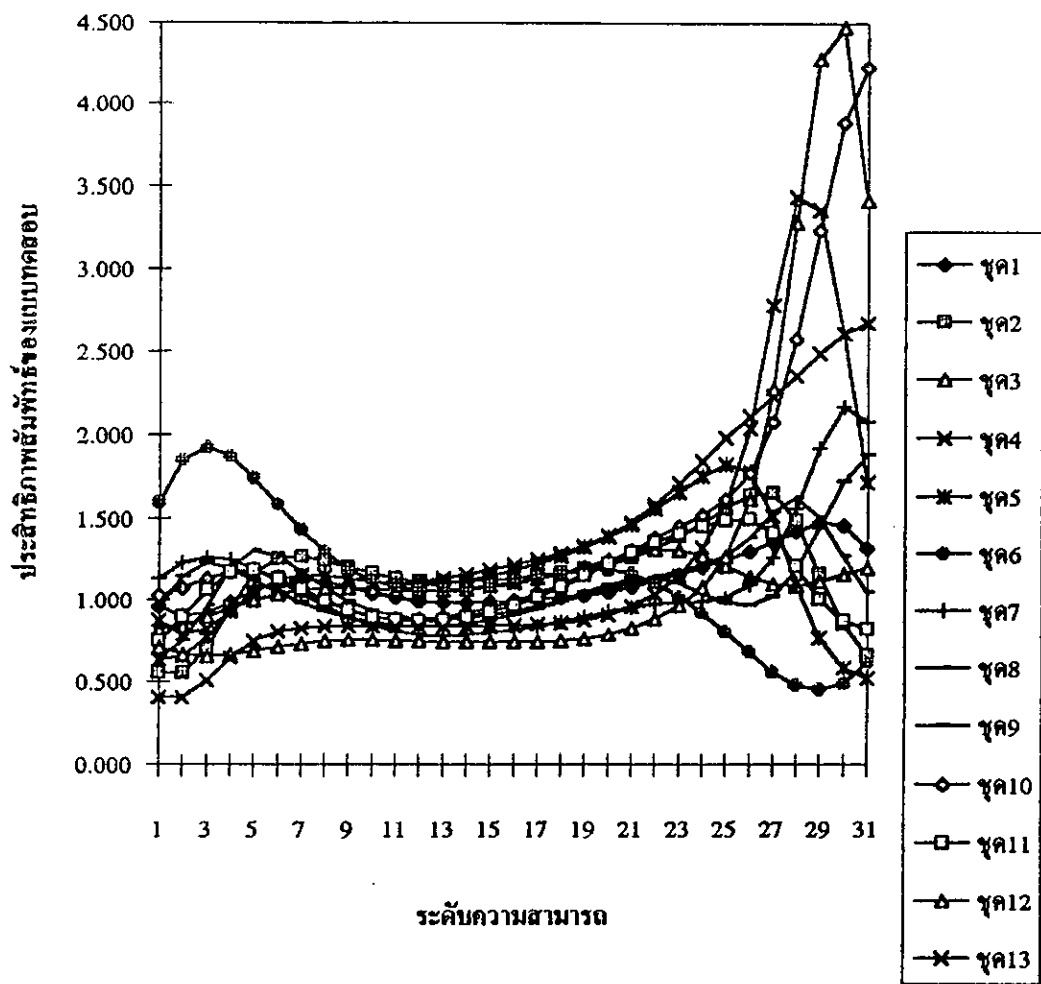
ชด 8 จำนวน 600/500 คน
 ชด 9 จำนวน 700/500 คน
 ชด 10 จำนวน 800/500 คน
 ชด 11 จำนวน 900/500 คน
 ชด 12 จำนวน 1,000/500 คน
 ชด 13 จำนวน 1,500/500 คน



ภาพประกอบ 69 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน	ชุด 8 จำนวน 1,500/600 คน
ชุด 2 จำนวน 2,500/500 คน	ชุด 9 จำนวน 2,000/600 คน
ชุด 3 จำนวน 3,000/500 คน	ชุด 10 จำนวน 2,500/600 คน
ชุด 4 จำนวน 700/600 คน	ชุด 11 จำนวน 3,000/600 คน
ชุด 5 จำนวน 800/600 คน	ชุด 12 จำนวน 800/700 คน
ชุด 6 จำนวน 900/600 คน	ชุด 13 จำนวน 900/700 คน
ชุด 7 จำนวน 1,000/600 คน	

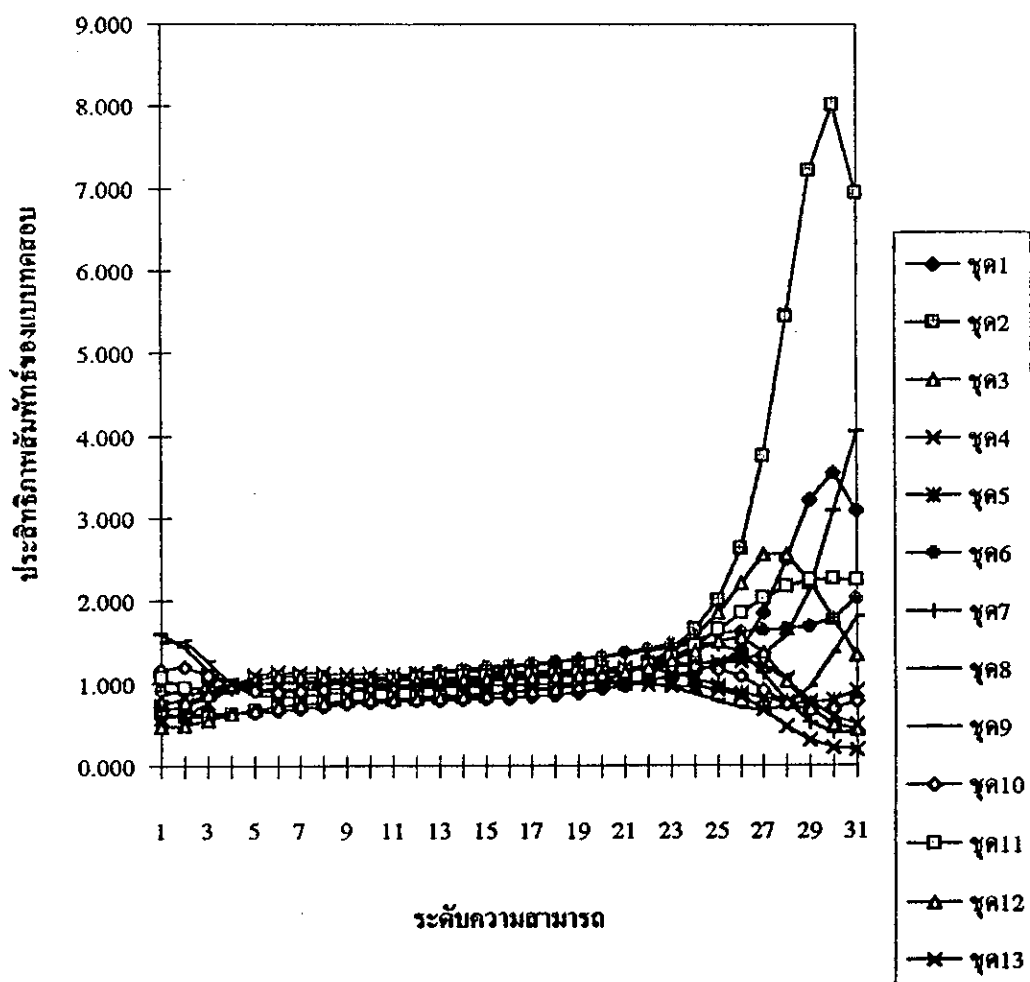


ภาพประกอบ 69 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 1,000/700 คน
 ชด 2 จำนวน 1,500/700 คน
 ชด 3 จำนวน 2,000/700 คน
 ชด 4 จำนวน 2,500/700 คน
 ชด 5 จำนวน 3,000/700 คน
 ชด 6 จำนวน 900/800 คน
 ชด 7 จำนวน 1,000/800 คน

ชด 8 จำนวน 1,500/800 คน
 ชด 9 จำนวน 2,000/800 คน
 ชด 10 จำนวน 2,500/800 คน
 ชด 11 จำนวน 3,000/800 คน
 ชด 12 จำนวน 1,000/900 คน
 ชด 13 จำนวน 1,500/900 คน



ภาพประกอบ 69 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 2,000/900 คน
 ชด 2 จำนวน 2,500/900 คน
 ชด 3 จำนวน 3,000/900 คน
 ชด 4 จำนวน 1,500/1,000 คน
 ชด 5 จำนวน 2,000/1,000 คน
 ชด 6 จำนวน 2,500/1,000 คน
 ชด 7 จำนวน 3,000/1,000 คน

ชด 8 จำนวน 2,000/1,500 คน
 ชด 9 จำนวน 2,500/1,500 คน
 ชด 10 จำนวน 3,000/1,500 คน
 ชด 11 จำนวน 2,500/2,000 คน
 ชด 12 จำนวน 3,000/2,000 คน
 ชด 13 จำนวน 3,000/2,500 คน

จากภาพประกอบ 64 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 50 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิฮูด ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 700 คน มีค่าสูงสุด ที่ระดับความสามารถสูง
 $\{RE(2,500, 700) = 4.235 ; \sigma = 3.000\}$

จากภาพประกอบ 65 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 60 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิฮูด ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 700 คน มีค่าสูงสุด ที่ระดับความสามารถสูง
 $\{RE(2,500, 700) = 4.333 ; \sigma = 3.000\}$

จากภาพประกอบ 66 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 70 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิฮูด ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 700 คน มีค่าสูงสุด ที่ระดับความสามารถสูง
 $\{RE(2500, 700) = 5.261 ; \sigma = 3.000\}$

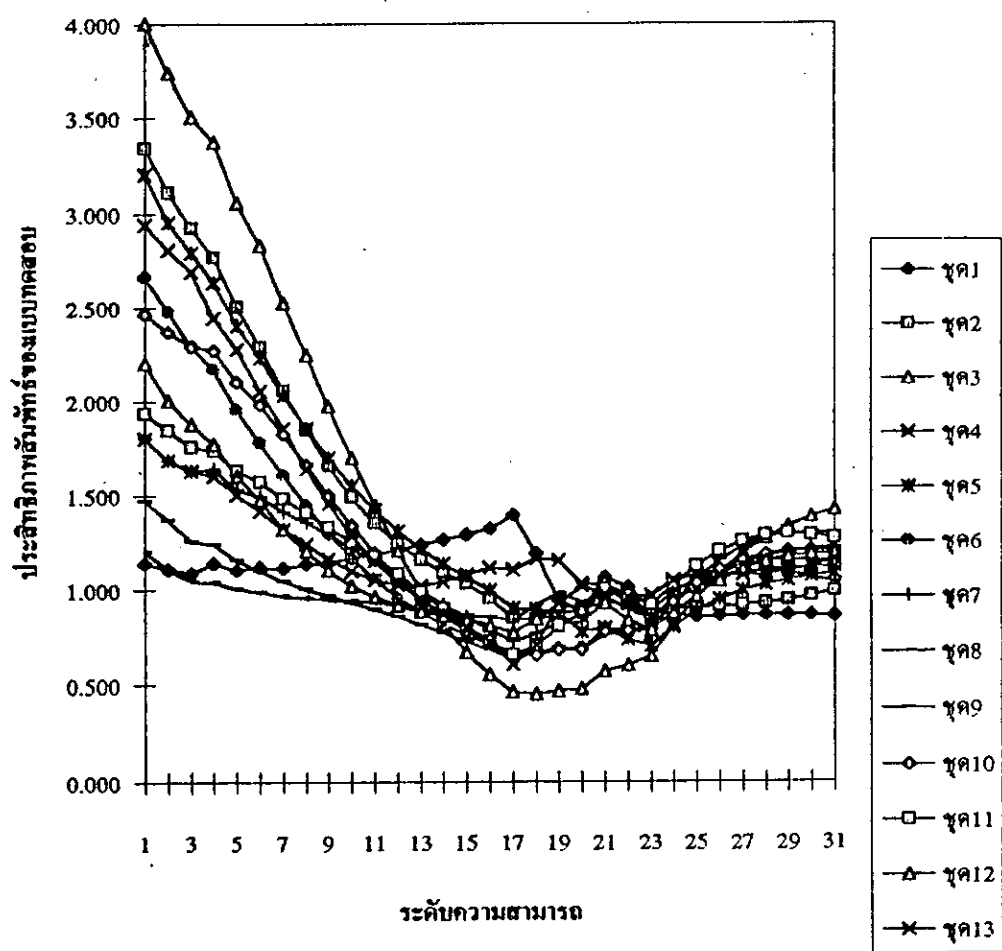
จากภาพประกอบ 67 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 80 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิฮูด ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 900 คน มีค่าสูงสุด ที่ระดับความสามารถสูง
 $\{RE(2,500, 900) = 5.864 ; \sigma = 2.600\}$

จากภาพประกอบ 68 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 90 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิฮูด ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 900 คน มีค่าสูงสุด ที่ระดับความสามารถสูง
 $\{RE(2,500, 900) = 10.412 ; \sigma = 2.600\}$

จากภาพประกอบ 69 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 100 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิฮูด ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 900 คน มีค่าสูงสุด ที่ระดับความสามารถสูง
 $\{RE(2,500, 900) = 8.019 ; \sigma = 2.800\}$

ดังนั้น จากภาพประกอบ 64 - 69 แสดงว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 50 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิฮูด ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 900 คน มีค่าสูงสุด ที่ระดับความสามารถสูง
 $\{RE(2,500, 900) = 10.412 ; \sigma = 2.600\}$

7.2 ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีเอสดีค เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 70 - 75

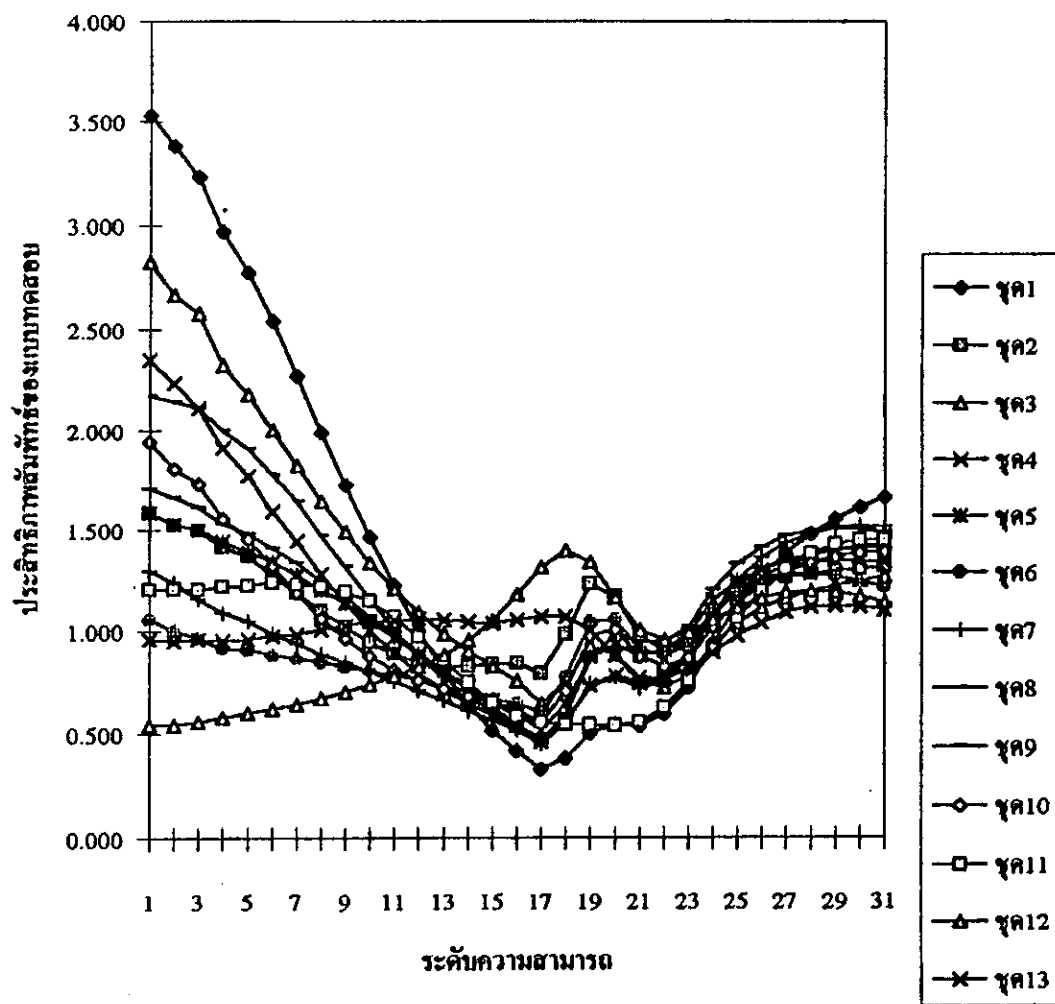


ภาพประกอบ 70 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

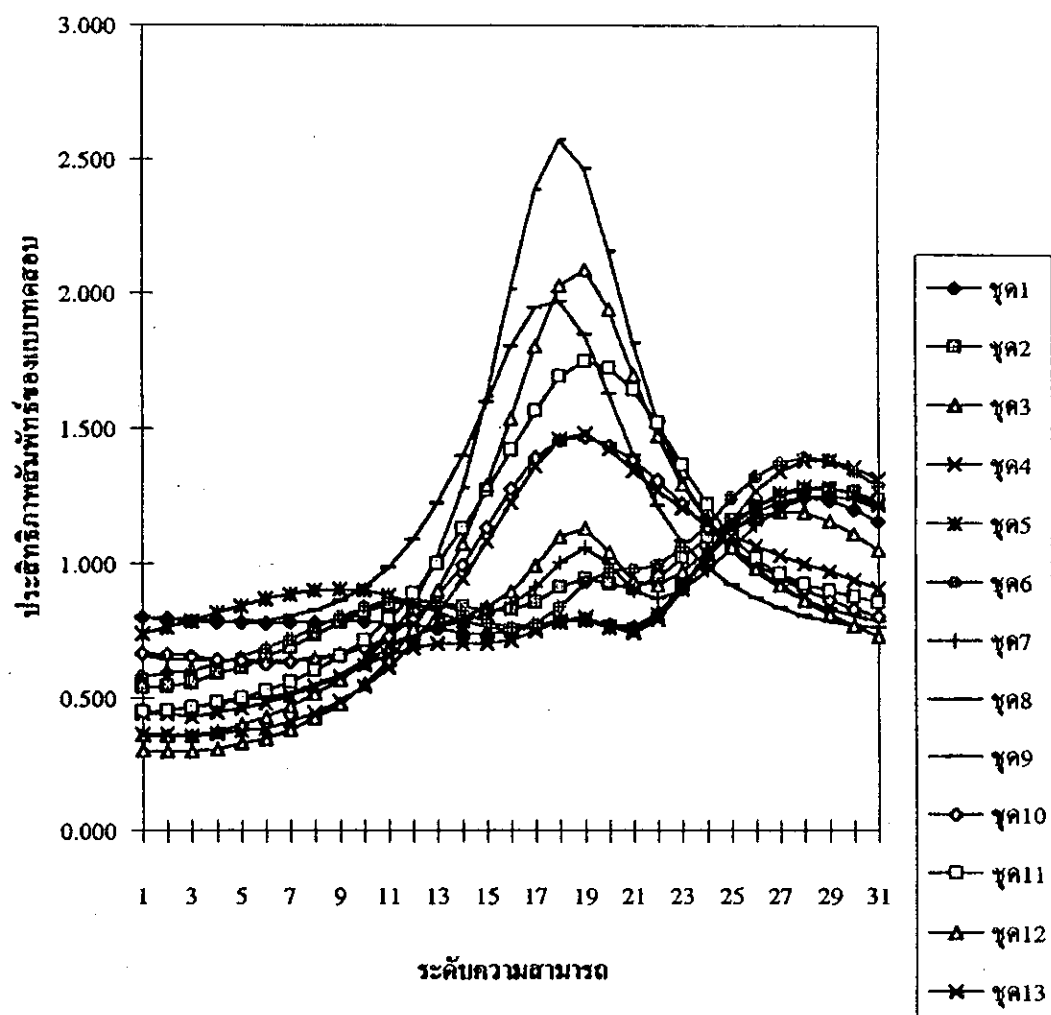


ภาพประกอบ 70 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

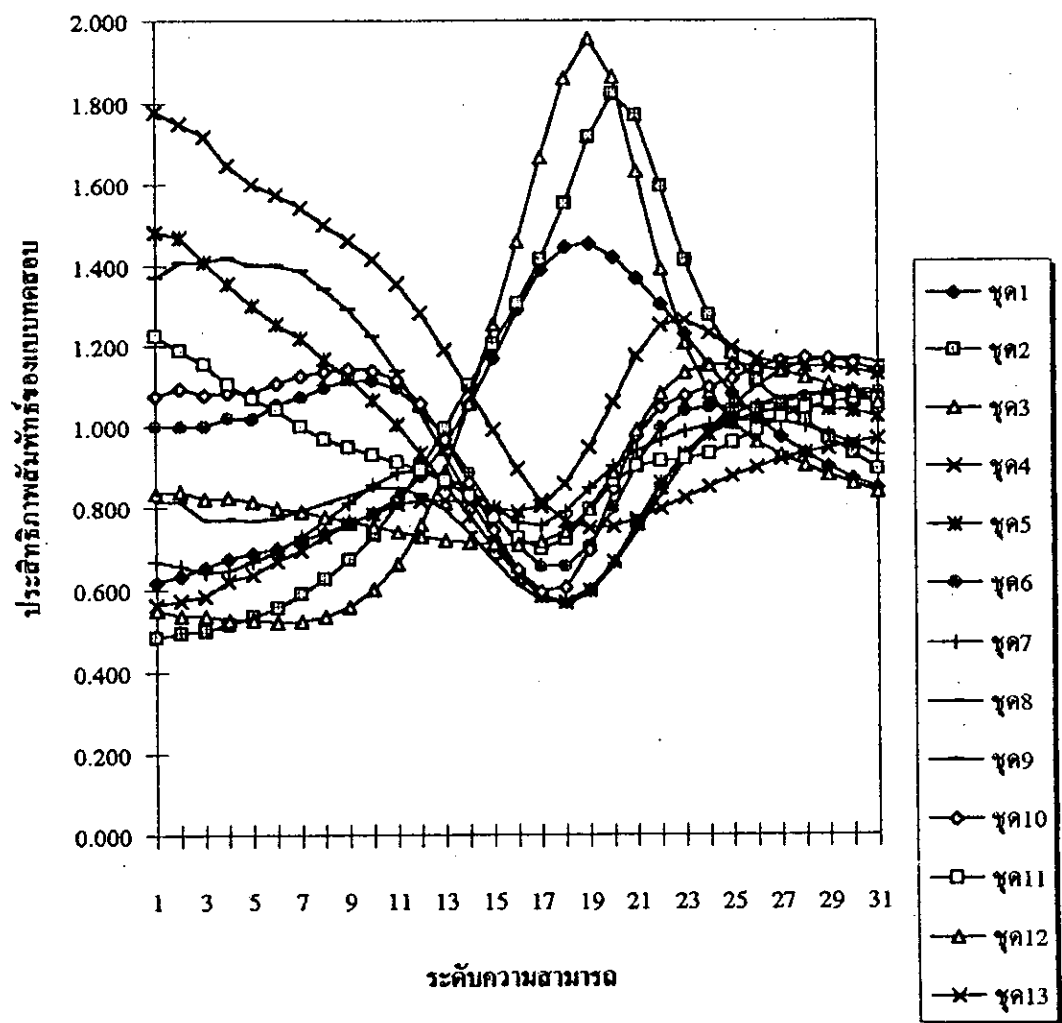


ภาพประกอบ 70 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 800/400 คน
 ชด 2 จำนวน 900/400 คน
 ชด 3 จำนวน 1,000/400 คน
 ชด 4 จำนวน 1,500/400 คน
 ชด 5 จำนวน 2,000/400 คน
 ชด 6 จำนวน 2,500/400 คน
 ชด 7 จำนวน 3,000/400 คน

ชด 8 จำนวน 600/500 คน
 ชด 9 จำนวน 700/500 คน
 ชด 10 จำนวน 800/500 คน
 ชด 11 จำนวน 900/500 คน
 ชด 12 จำนวน 1,000/500 คน
 ชด 13 จำนวน 1,500/500 คน

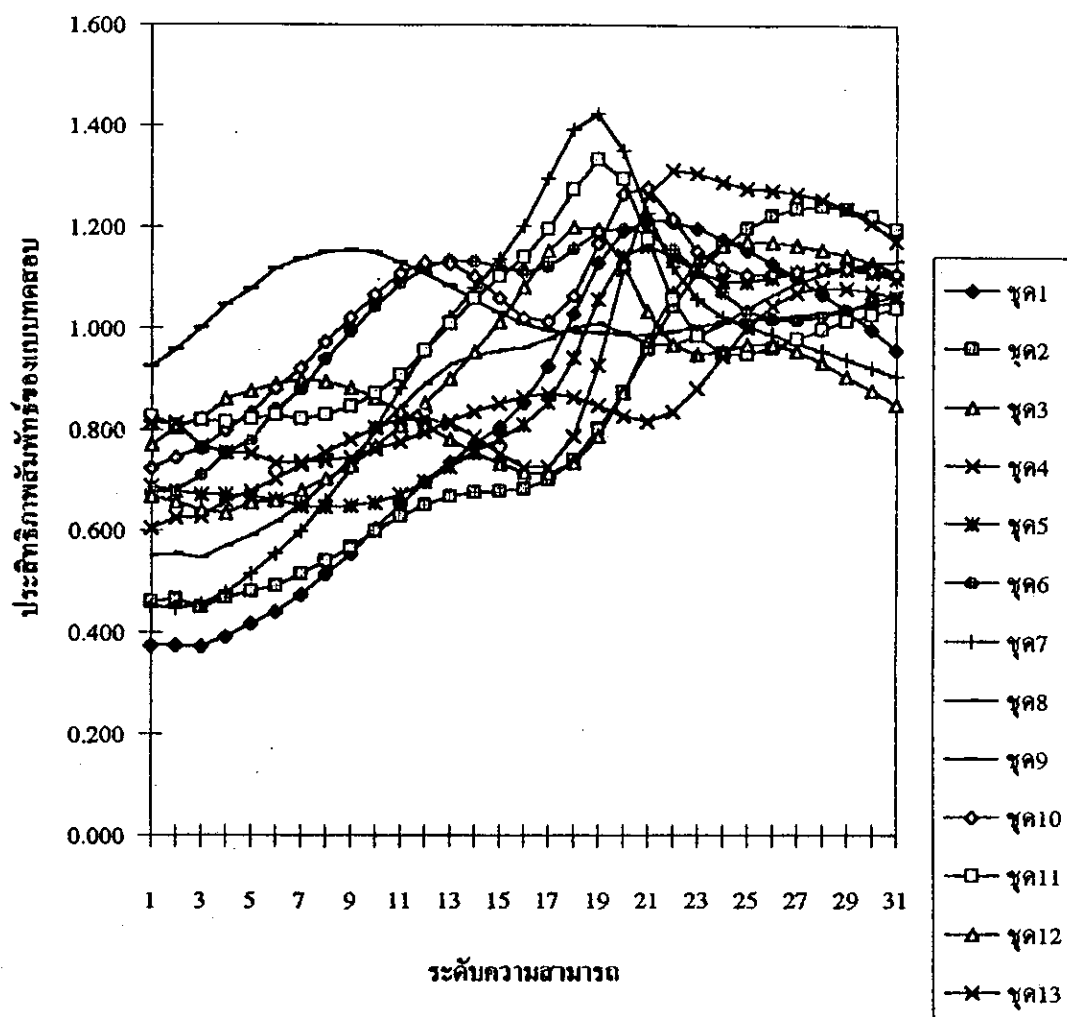


ภาพประกอบ 70 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
ชุด 2 จำนวน 2,500/500 คน
ชุด 3 จำนวน 3,000/500 คน
ชุด 4 จำนวน 700/600 คน
ชุด 5 จำนวน 800/600 คน
ชุด 6 จำนวน 900/600 คน
ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
ชุด 9 จำนวน 2,000/600 คน
ชุด 10 จำนวน 2,500/600 คน
ชุด 11 จำนวน 3,000/600 คน
ชุด 12 จำนวน 800/700 คน
ชุด 13 จำนวน 900/700 คน

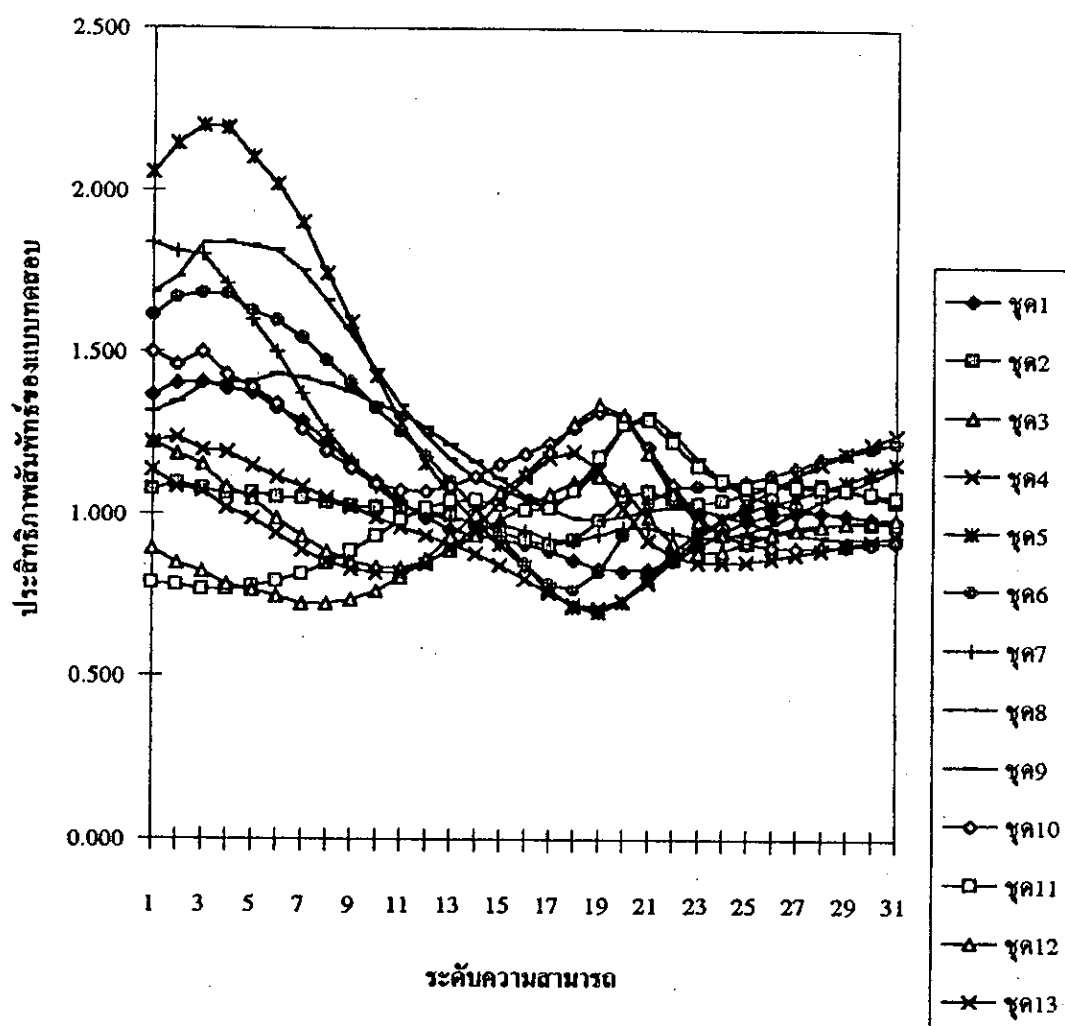


ภาพประกอบ 70 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
ชุด 2 จำนวน 1,500/700 คน
ชุด 3 จำนวน 2,000/700 คน
ชุด 4 จำนวน 2,500/700 คน
ชุด 5 จำนวน 3,000/700 คน
ชุด 6 จำนวน 900/800 คน
ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
ชุด 9 จำนวน 2,000/800 คน
ชุด 10 จำนวน 2,500/800 คน
ชุด 11 จำนวน 3,000/800 คน
ชุด 12 จำนวน 1,000/900 คน
ชุด 13 จำนวน 1,500/900 คน

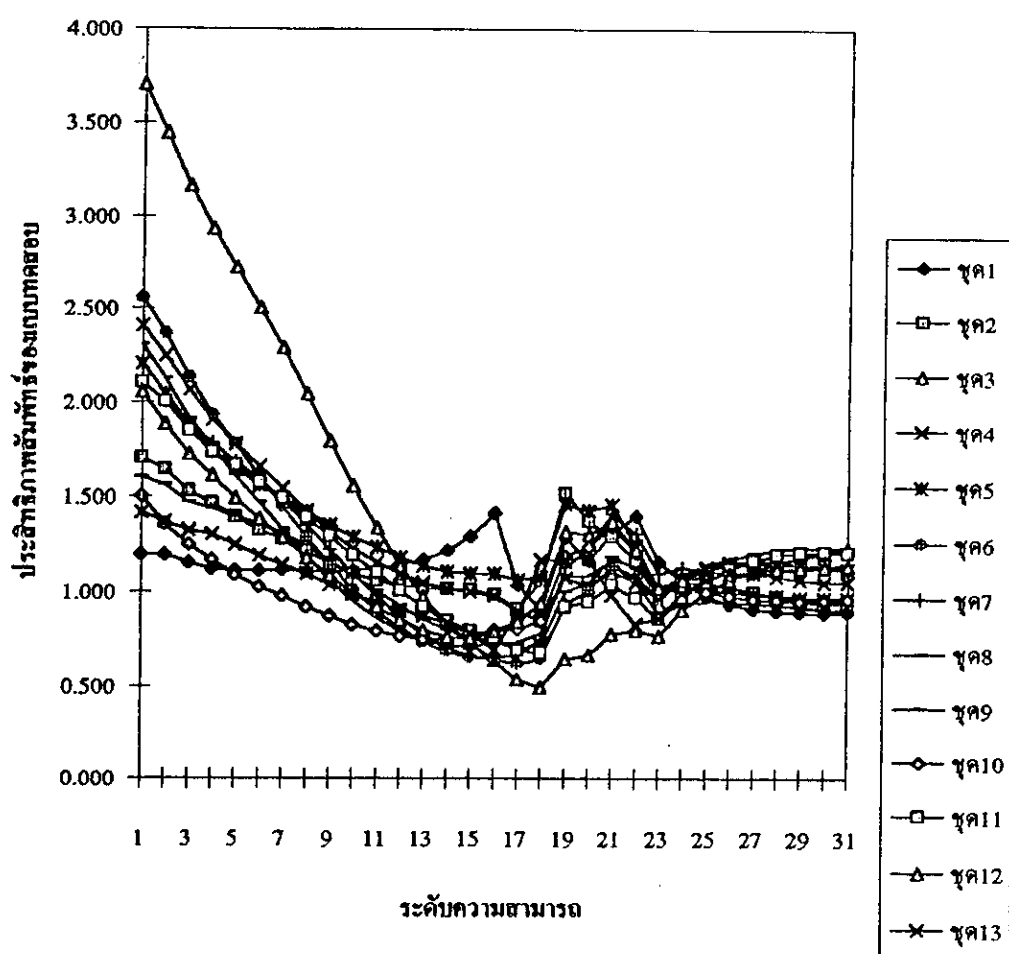


ภาพประกอบ 70 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

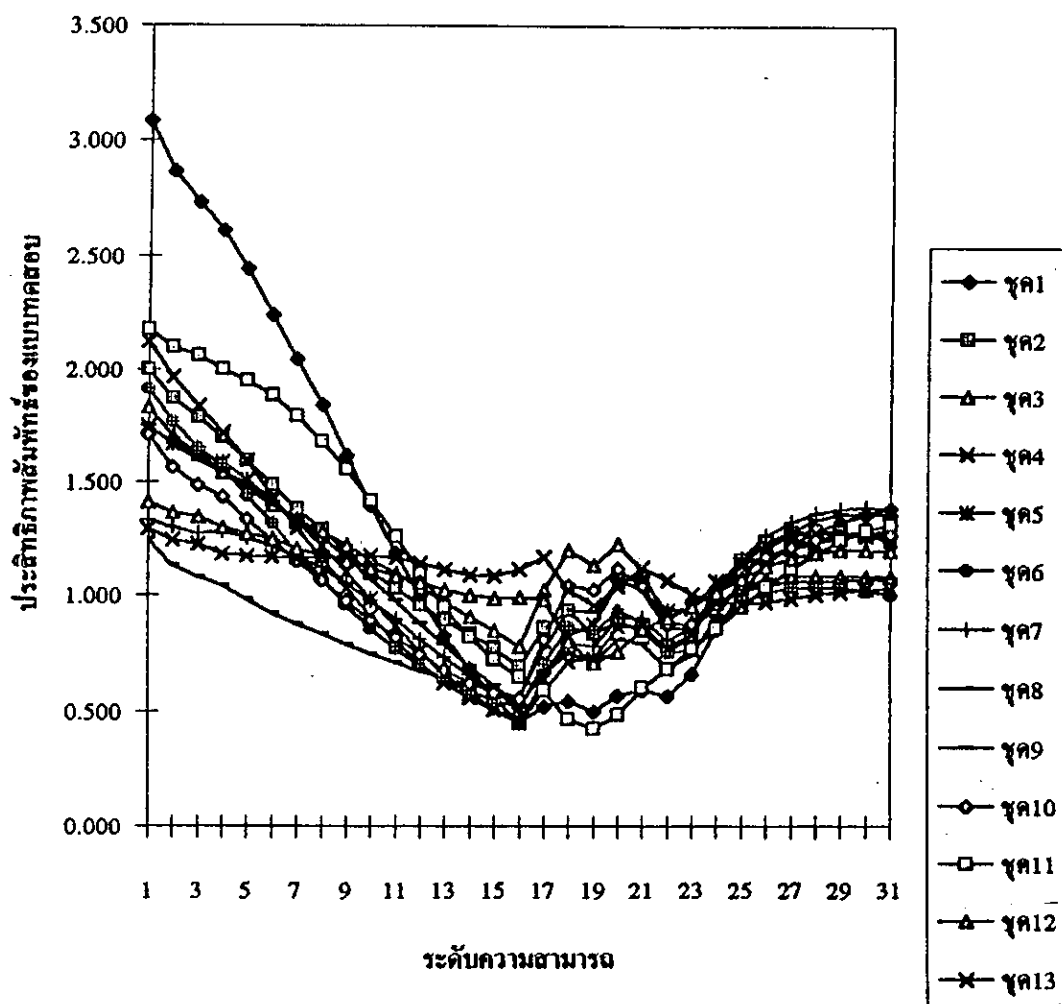
ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน



ภาพประกอบ 71 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน	ชุด 8 จำนวน 1,000/200 คน
ชุด 2 จำนวน 400/200 คน	ชุด 9 จำนวน 1,500/200 คน
ชุด 3 จำนวน 500/200 คน	ชุด 10 จำนวน 2,000/200 คน
ชุด 4 จำนวน 600/200 คน	ชุด 11 จำนวน 2,500/200 คน
ชุด 5 จำนวน 700/200 คน	ชุด 12 จำนวน 3,000/200 คน
ชุด 6 จำนวน 800/200 คน	ชุด 13 จำนวน 400/300 คน
ชุด 7 จำนวน 900/200 คน	

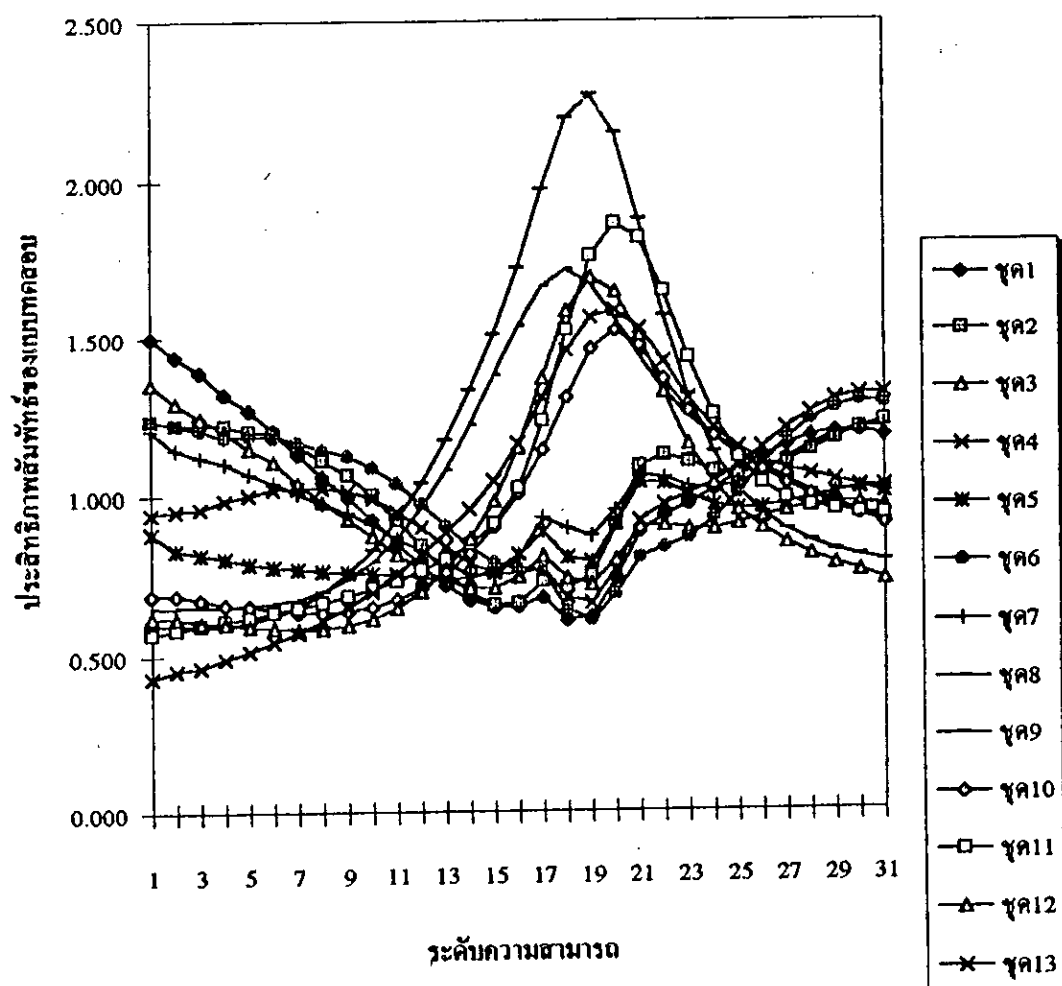


ภาพประกอบ 71 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

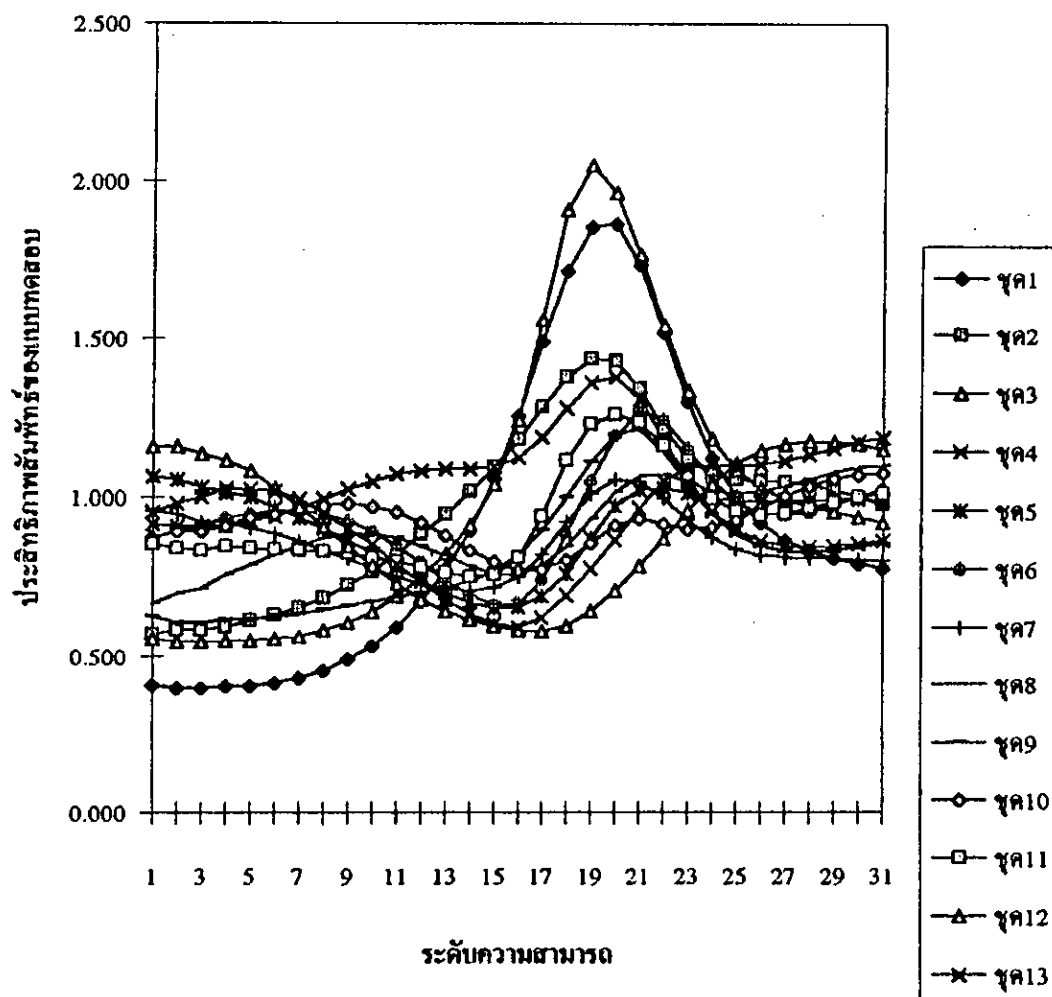


ภาพประกอบ 71 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 800/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

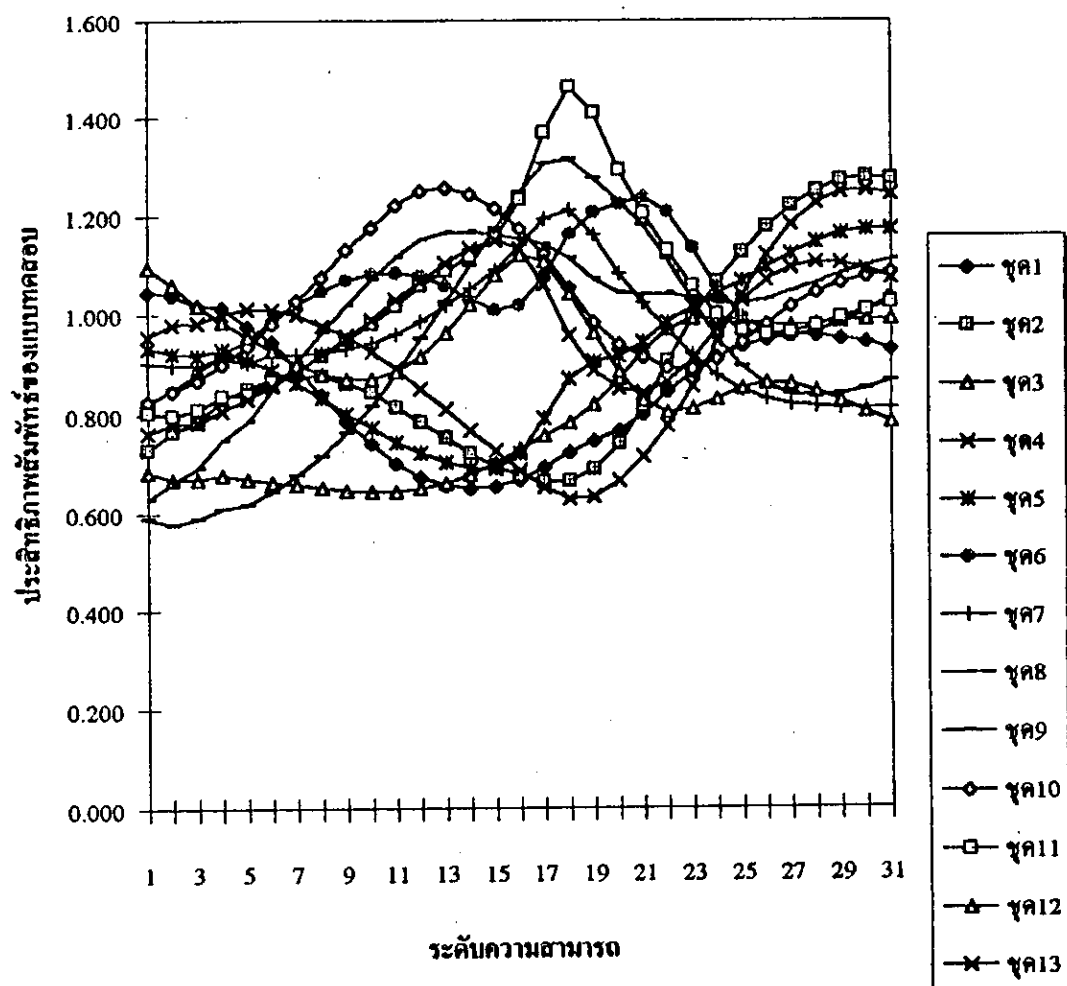


ภาพประกอบ 71 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 2,000/500 คน
 ชด 2 จำนวน 2,500/500 คน
 ชด 3 จำนวน 3,000/500 คน
 ชด 4 จำนวน 700/600 คน
 ชด 5 จำนวน 800/600 คน
 ชด 6 จำนวน 900/600 คน
 ชด 7 จำนวน 1,000/600 คน

ชด 8 จำนวน 1,500/600 คน
 ชด 9 จำนวน 2,000/600 คน
 ชด 10 จำนวน 2,500/600 คน
 ชด 11 จำนวน 3,000/600 คน
 ชด 12 จำนวน 800/700 คน
 ชด 13 จำนวน 900/700 คน

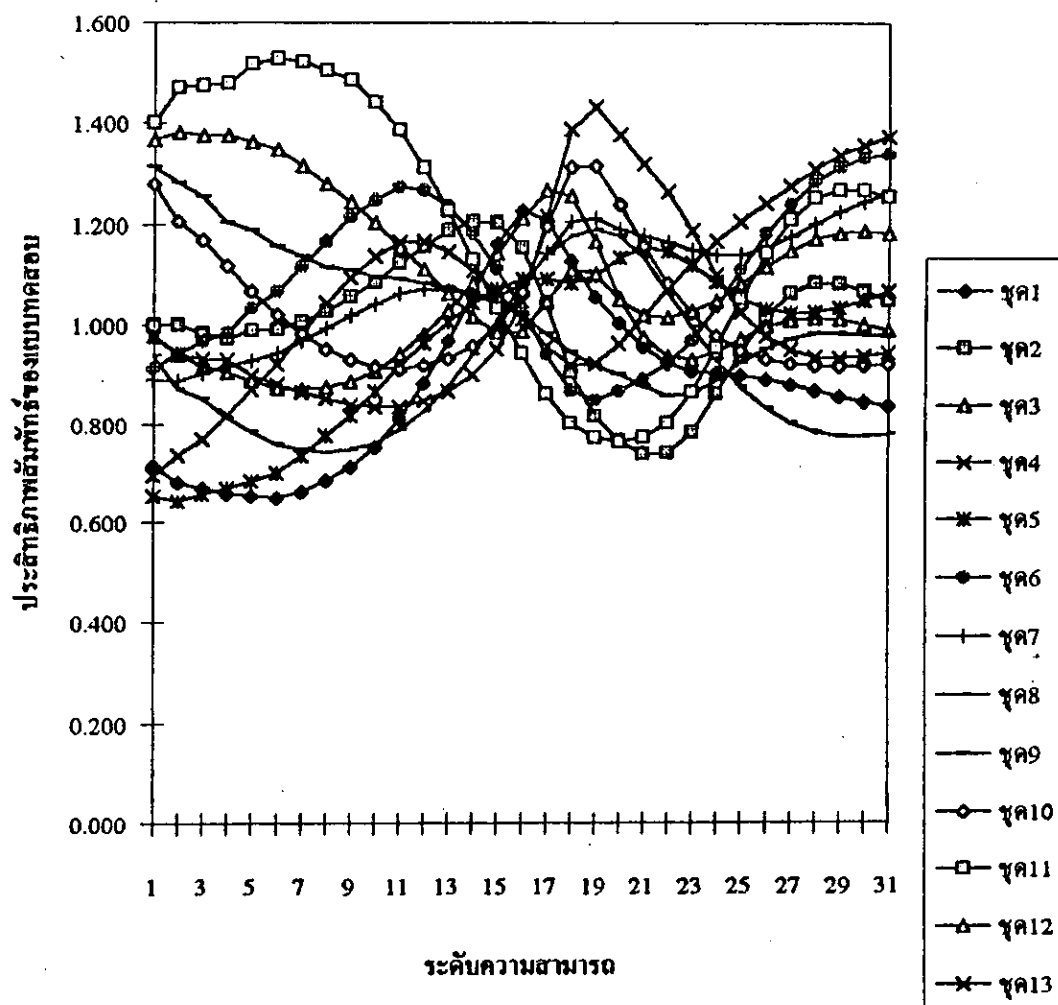


ภาพประกอบ 71 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

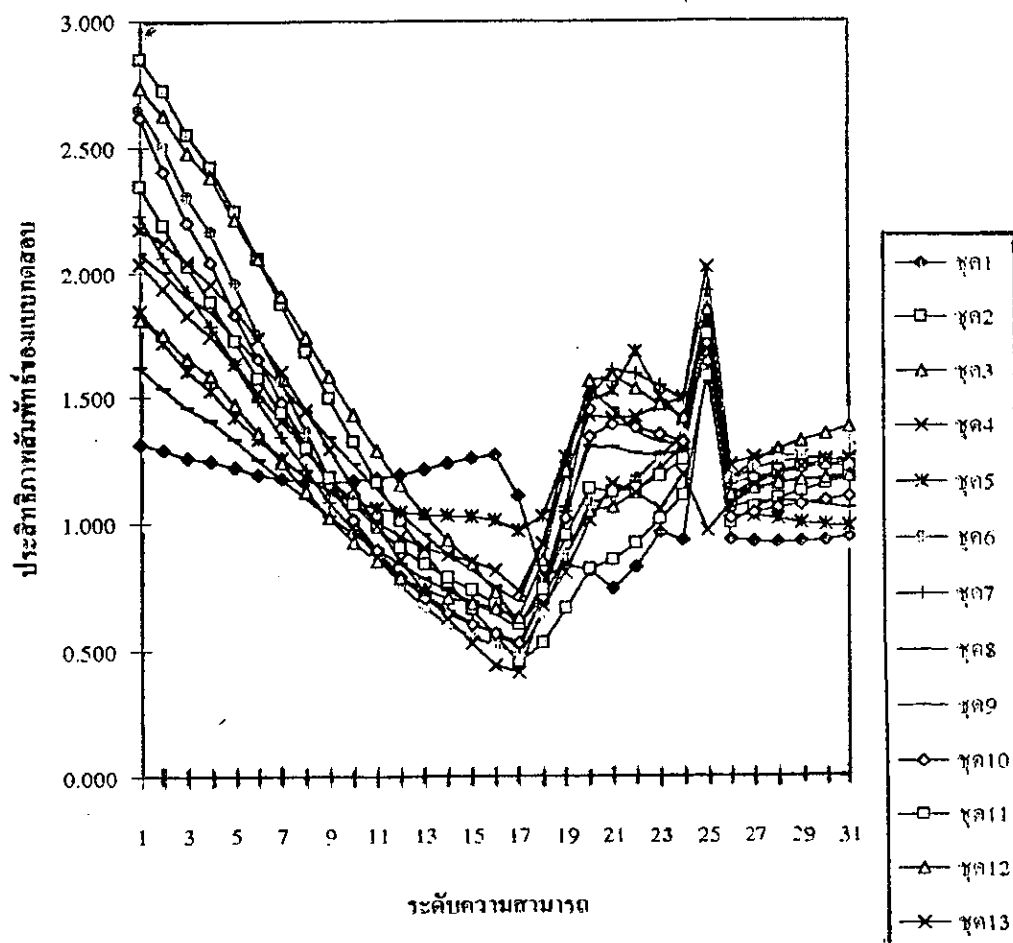


ภาพประกอบ 71 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

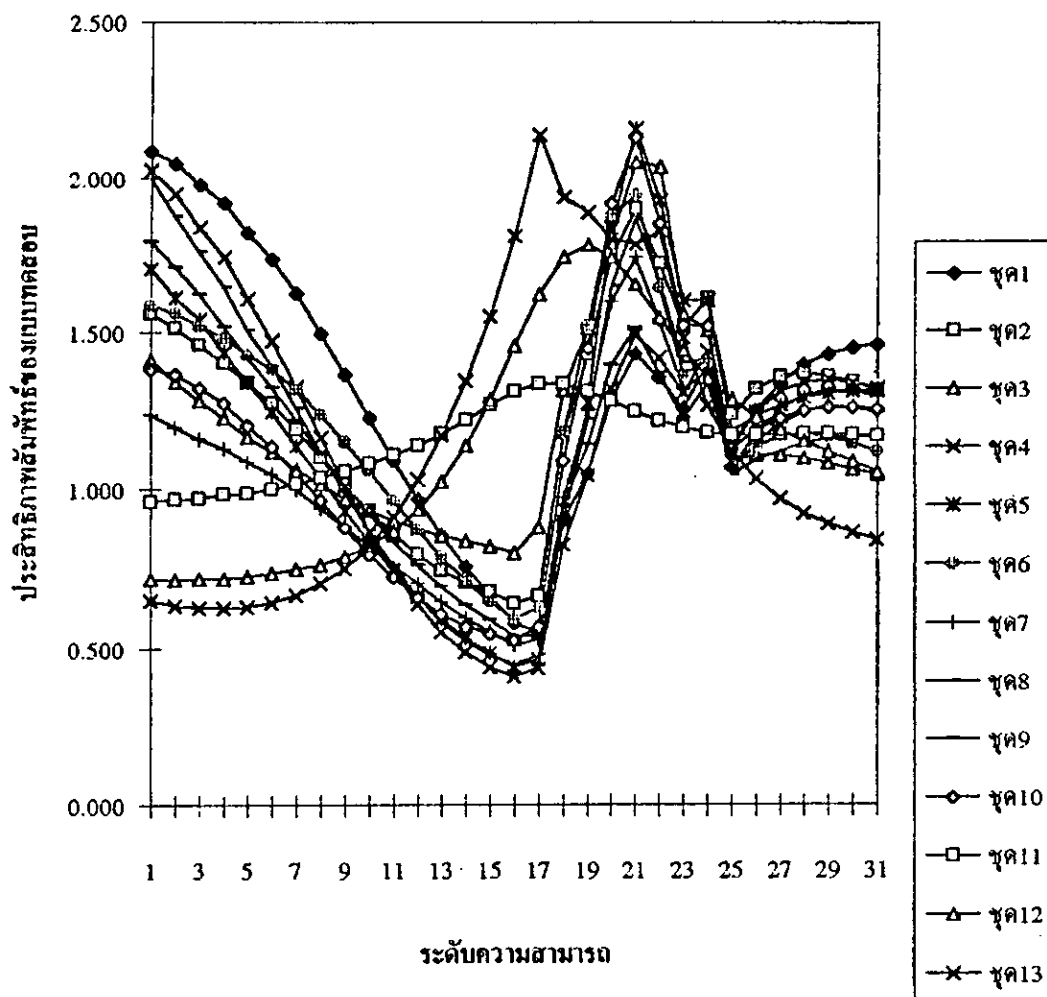


ภาพประกอบ 72 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวริสติก เมื่อข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

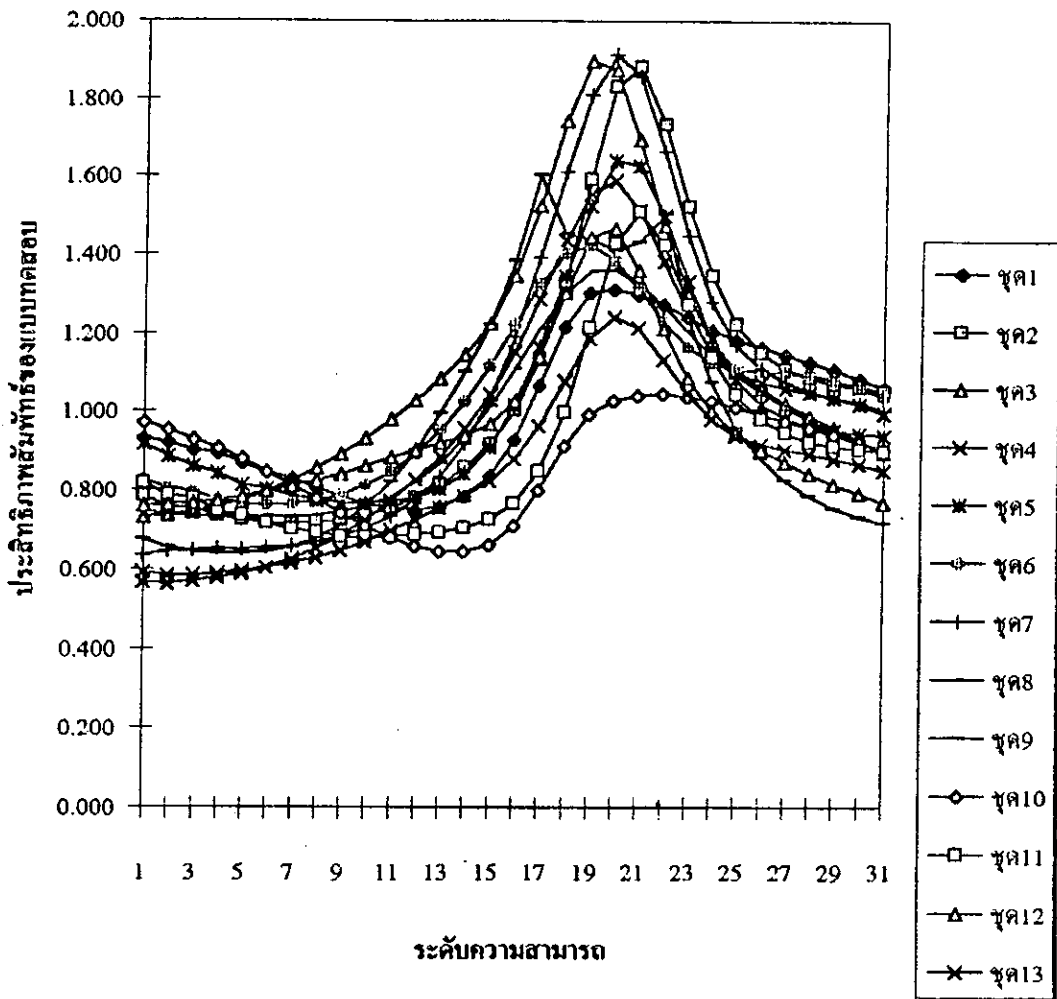


ภาพประกอบ 72 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 500/300 คน
 ชด 2 จำนวน 600/300 คน
 ชด 3 จำนวน 700/300 คน
 ชด 4 จำนวน 800/300 คน
 ชด 5 จำนวน 900/300 คน
 ชด 6 จำนวน 1,000/300 คน
 ชด 7 จำนวน 1,500/300 คน

ชด 8 จำนวน 2,000/300 คน
 ชด 9 จำนวน 2,500/300 คน
 ชด 10 จำนวน 3,000/300 คน
 ชด 11 จำนวน 500/400 คน
 ชด 12 จำนวน 600/400 คน
 ชด 13 จำนวน 700/400 คน

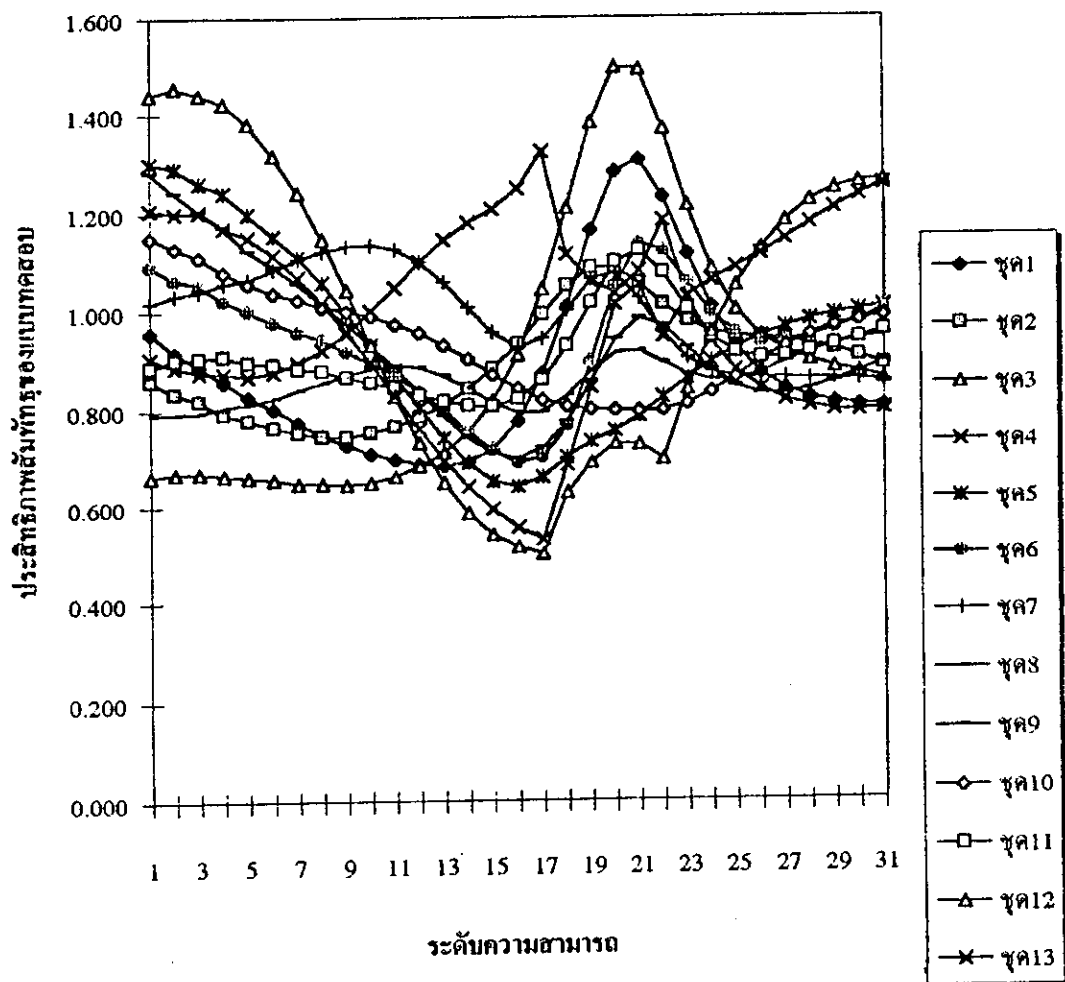


ภาพประกอบ 72 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 800/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

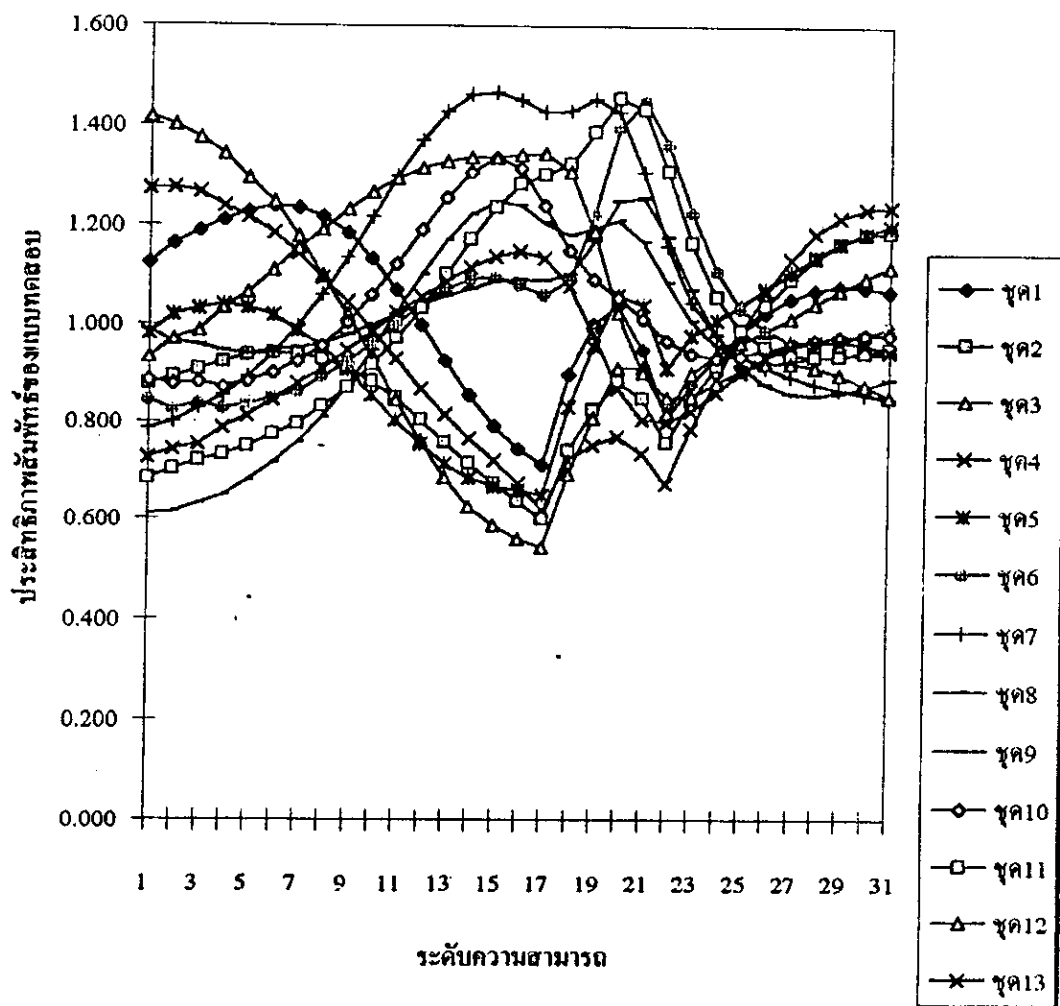


ภาพประกอบ 72 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
ชุด 2 จำนวน 2,500/500 คน
ชุด 3 จำนวน 3,000/500 คน
ชุด 4 จำนวน 700/600 คน
ชุด 5 จำนวน 800/600 คน
ชุด 6 จำนวน 900/600 คน
ชุด 7 จำนวน 1,000/600 คน

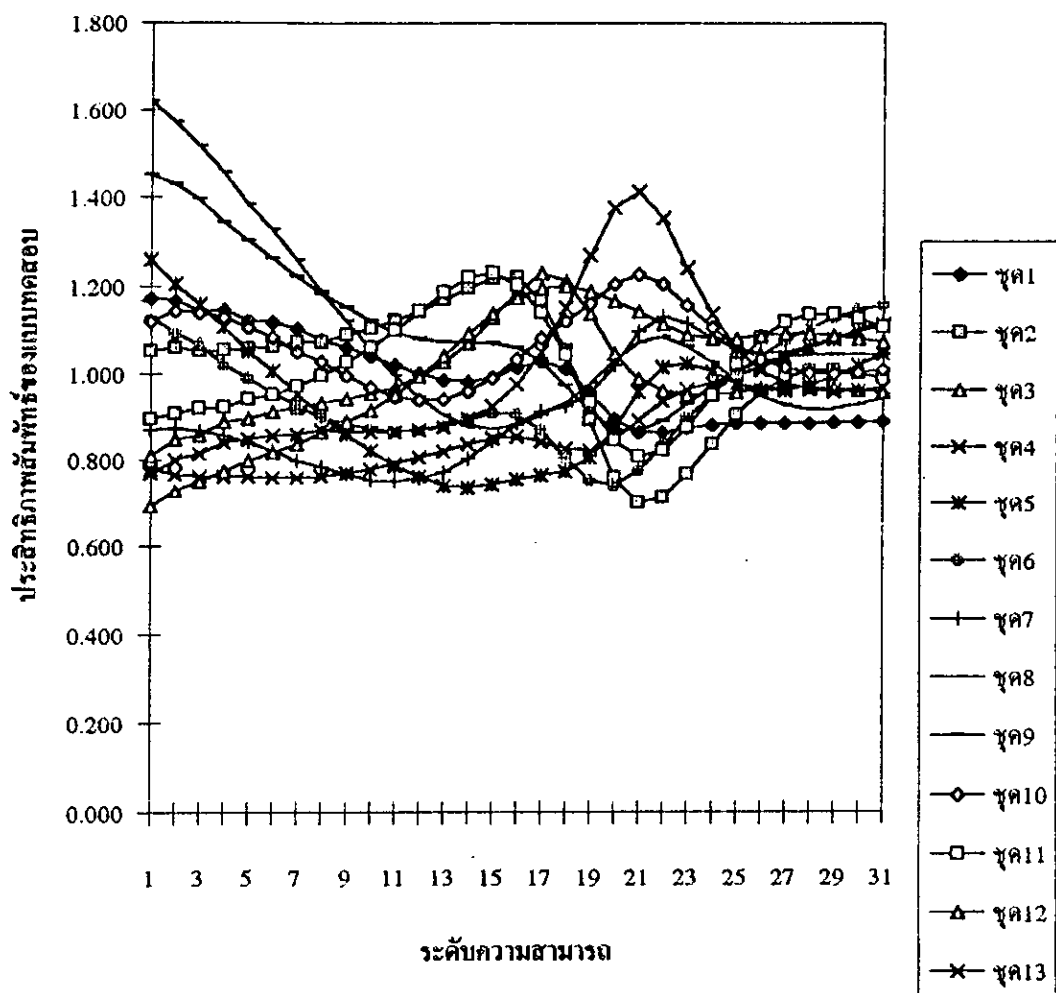
ชุด 8 จำนวน 1,500/600 คน
ชุด 9 จำนวน 2,000/600 คน
ชุด 10 จำนวน 2,500/600 คน
ชุด 11 จำนวน 3,000/600 คน
ชุด 12 จำนวน 800/700 คน
ชุด 13 จำนวน 900/700 คน



ภาพประกอบ 72 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน	ชุด 8 จำนวน 1,500/800 คน
ชุด 2 จำนวน 1,500/700 คน	ชุด 9 จำนวน 2,000/800 คน
ชุด 3 จำนวน 2,000/700 คน	ชุด 10 จำนวน 2,500/800 คน
ชุด 4 จำนวน 2,500/700 คน	ชุด 11 จำนวน 3,000/800 คน
ชุด 5 จำนวน 3,000/700 คน	ชุด 12 จำนวน 1,000/900 คน
ชุด 6 จำนวน 900/800 คน	ชุด 13 จำนวน 1,500/900 คน
ชุด 7 จำนวน 1,000/800 คน	

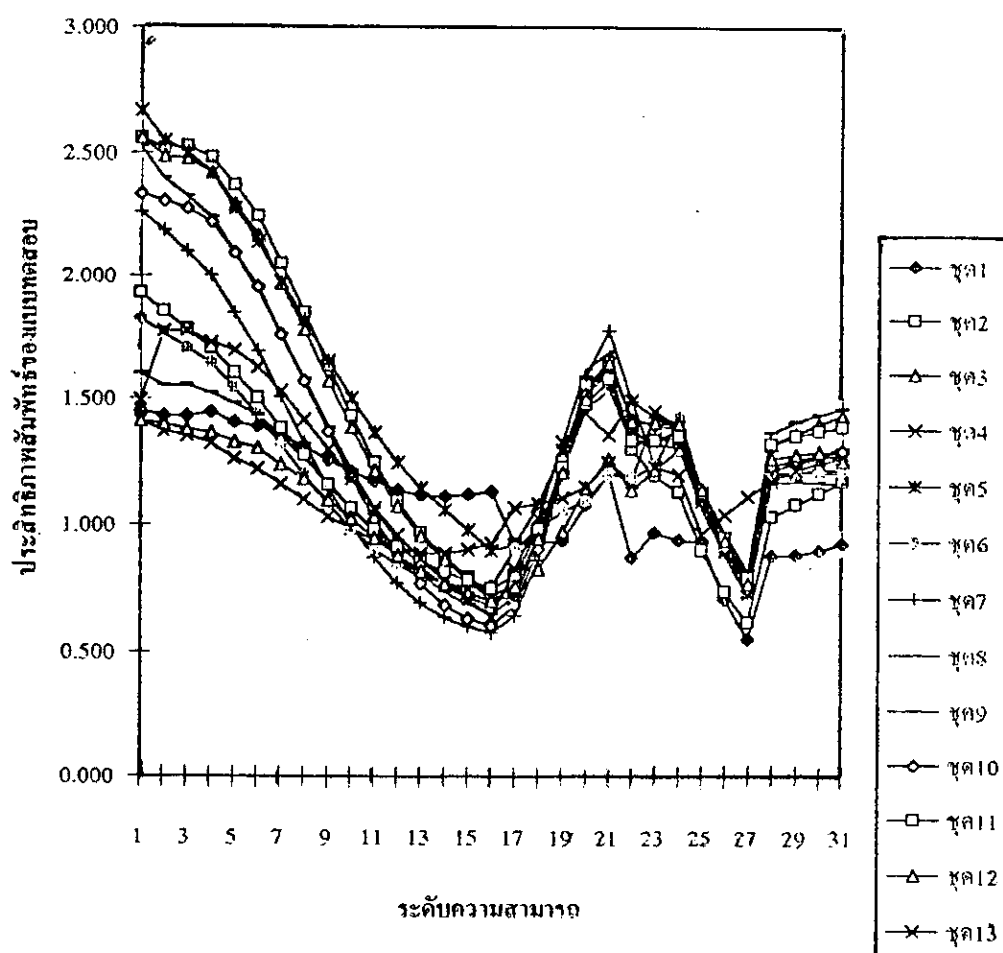


ภาพประกอบ 72 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
ชุด 2 จำนวน 2,500/900 คน
ชุด 3 จำนวน 3,000/900 คน
ชุด 4 จำนวน 1,500/1,000 คน
ชุด 5 จำนวน 2,000/1,000 คน
ชุด 6 จำนวน 2,500/1,000 คน
ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
ชุด 9 จำนวน 2,500/1,500 คน
ชุด 10 จำนวน 3,000/1,500 คน
ชุด 11 จำนวน 2,500/2,000 คน
ชุด 12 จำนวน 3,000/2,000 คน
ชุด 13 จำนวน 3,000/2,500 คน



ภาพประกอบ 73 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีวิสติด เมื่อข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน

ชุด 2 จำนวน 400/200 คน

ชุด 3 จำนวน 500/200 คน

ชุด 4 จำนวน 600/200 คน

ชุด 5 จำนวน 700/200 คน

ชุด 6 จำนวน 800/200 คน

ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน

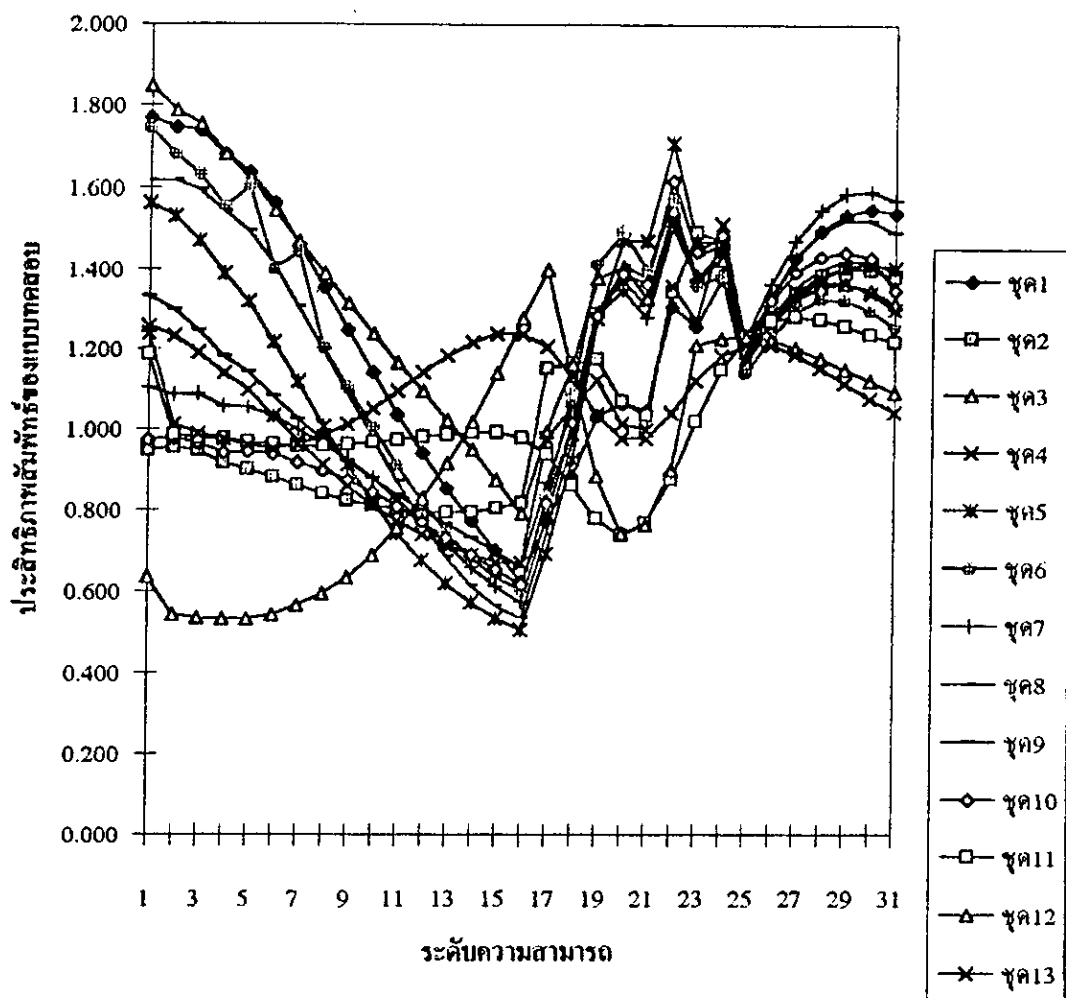
ชุด 9 จำนวน 1,500/200 คน

ชุด 10 จำนวน 2,000/200 คน

ชุด 11 จำนวน 2,500/200 คน

ชุด 12 จำนวน 3,000/200 คน

ชุด 13 จำนวน 400/300 คน

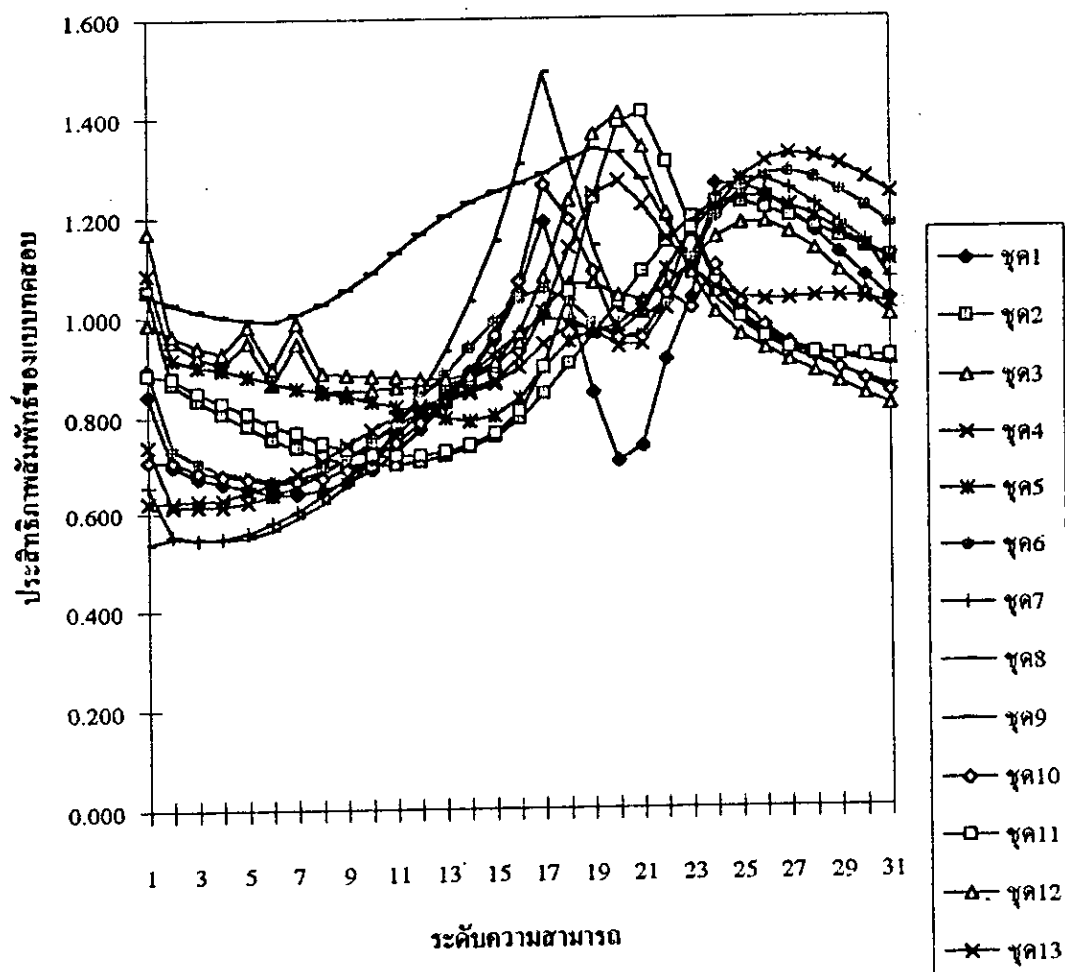


ภาพประกอบ 73 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/800 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

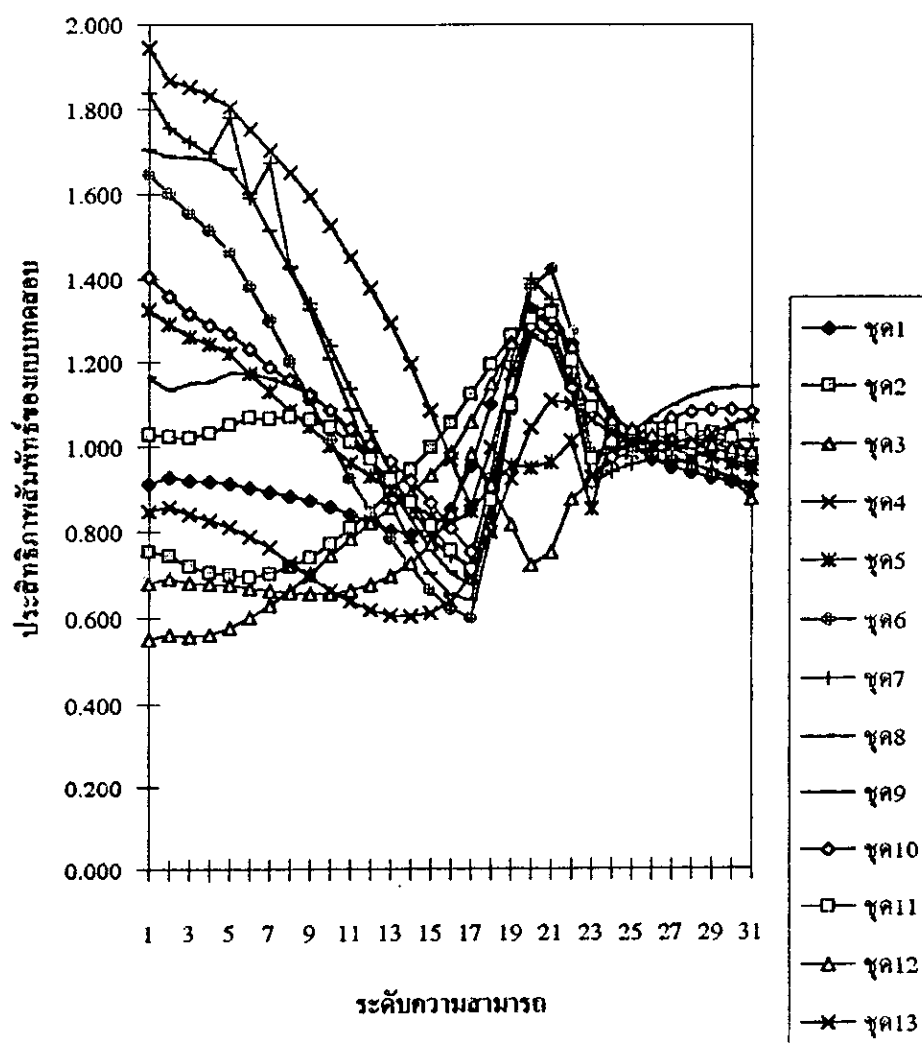


ภาพประกอบ 73 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

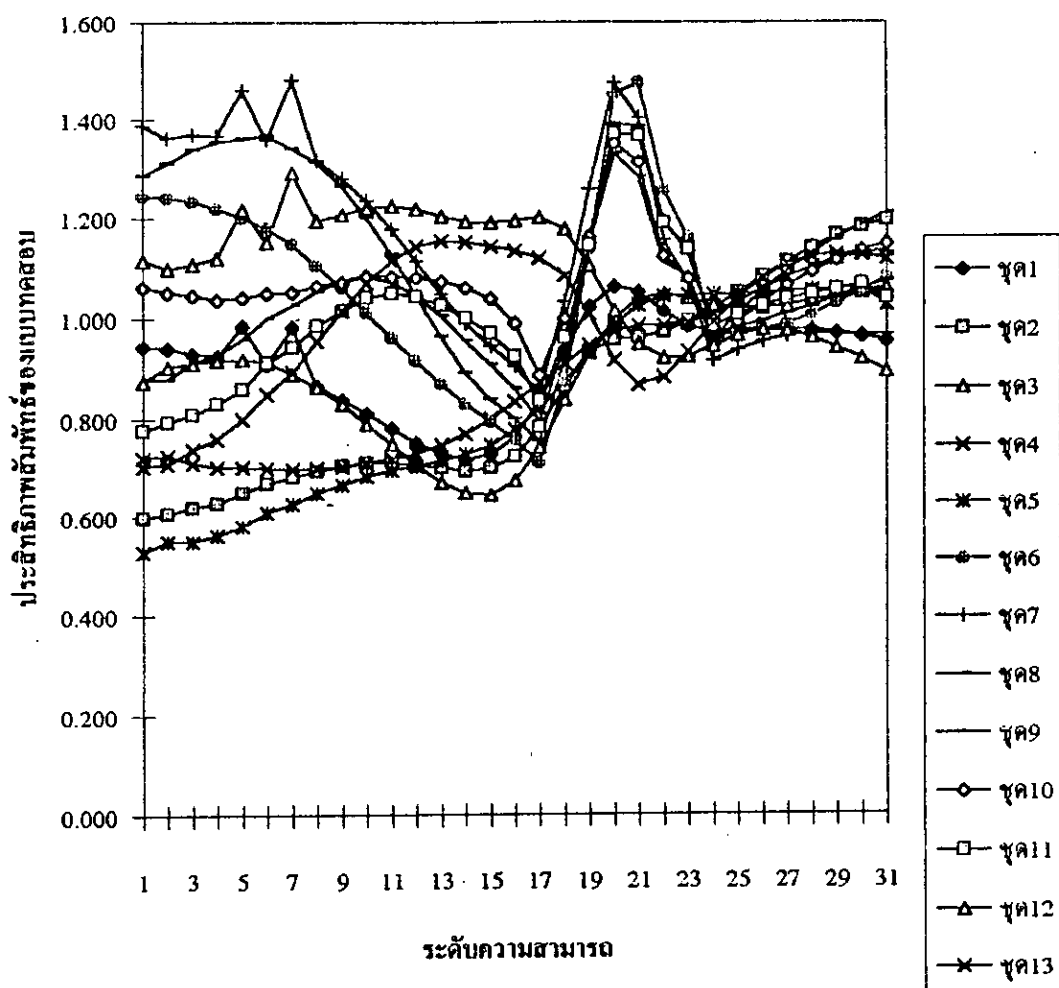


ภาพประกอบ 73 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/800 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

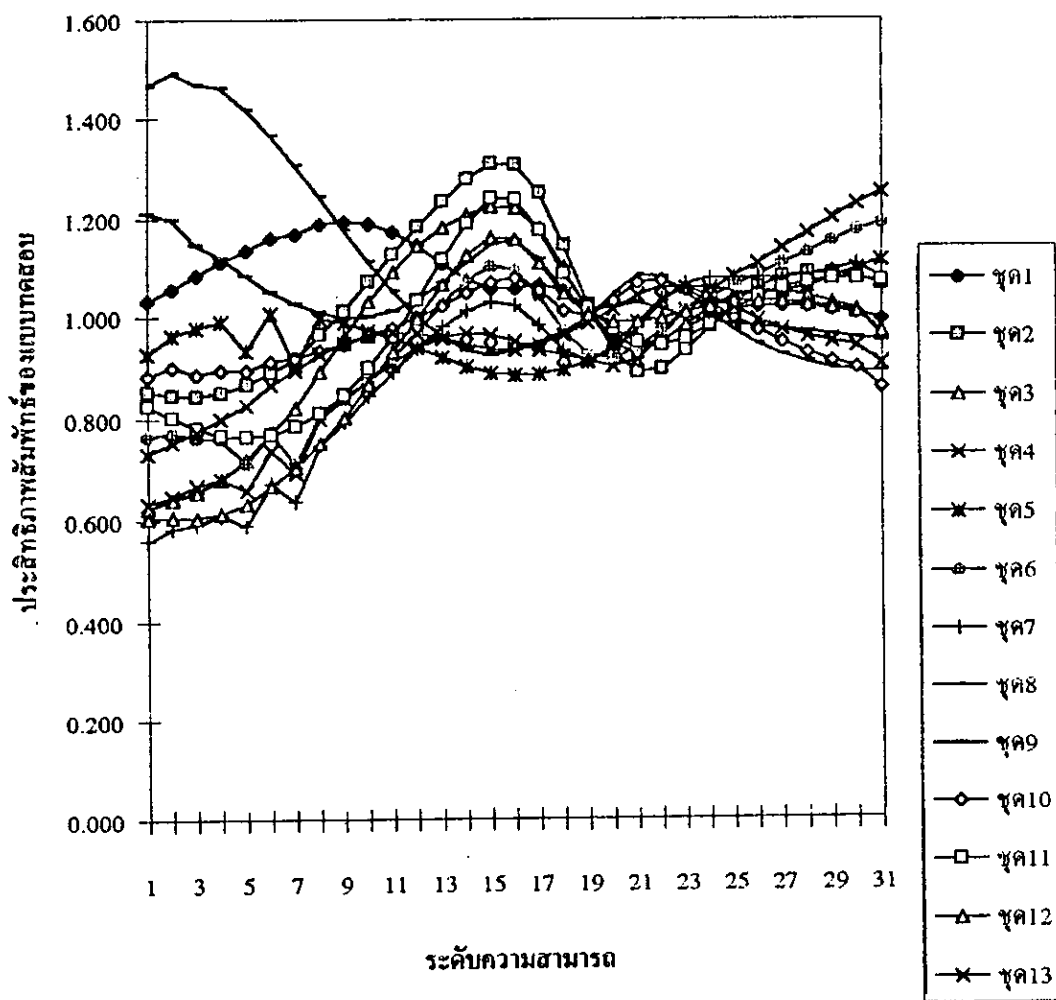


ภาพประกอบ 73 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

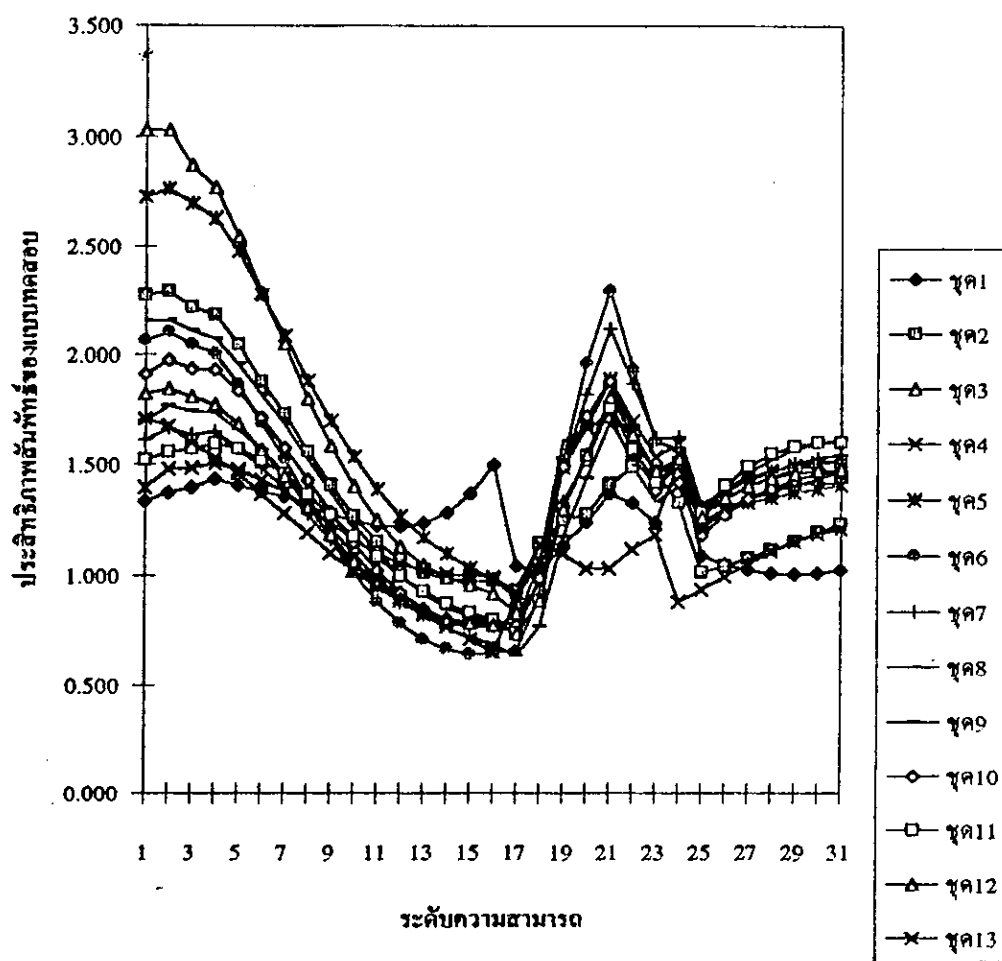


ภาพประกอบ 73 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

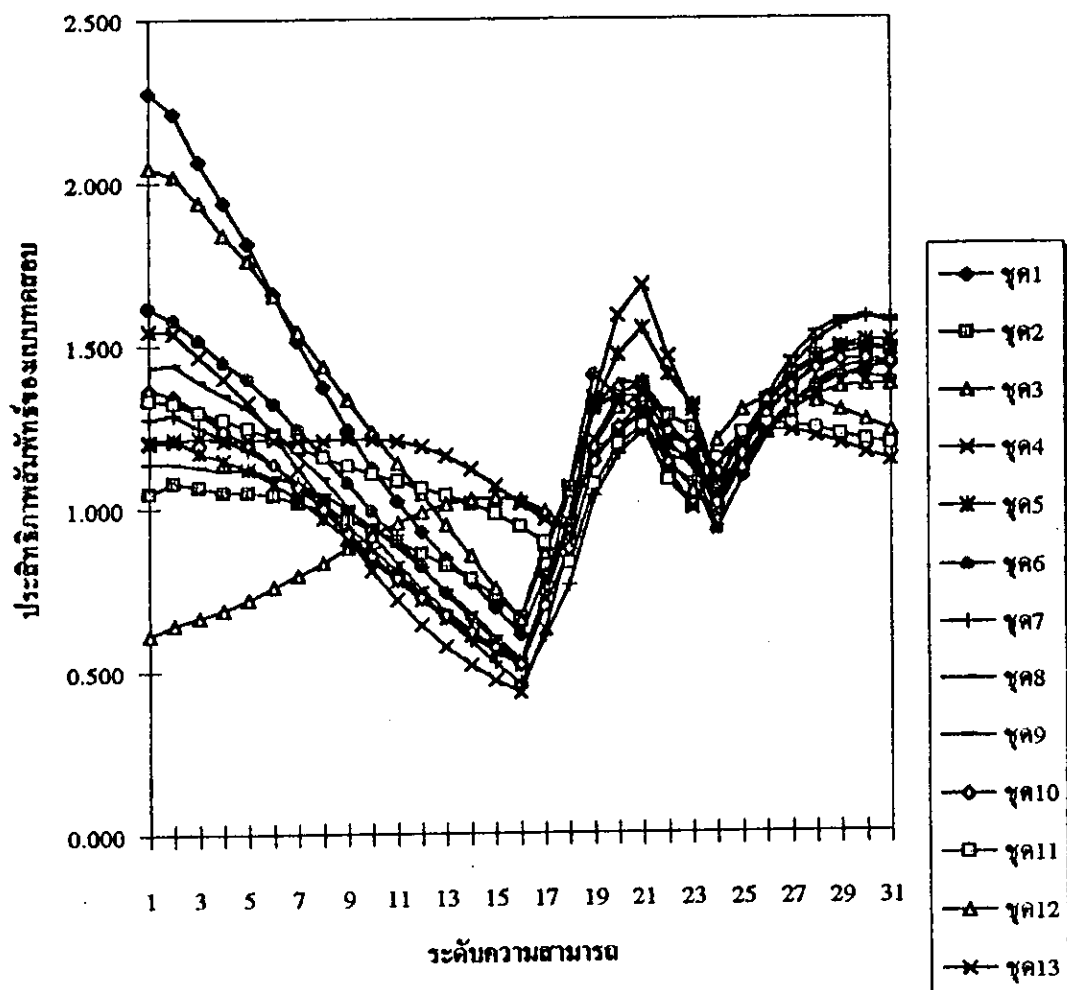


ภาพประกอบ 74 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีเอสดี เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

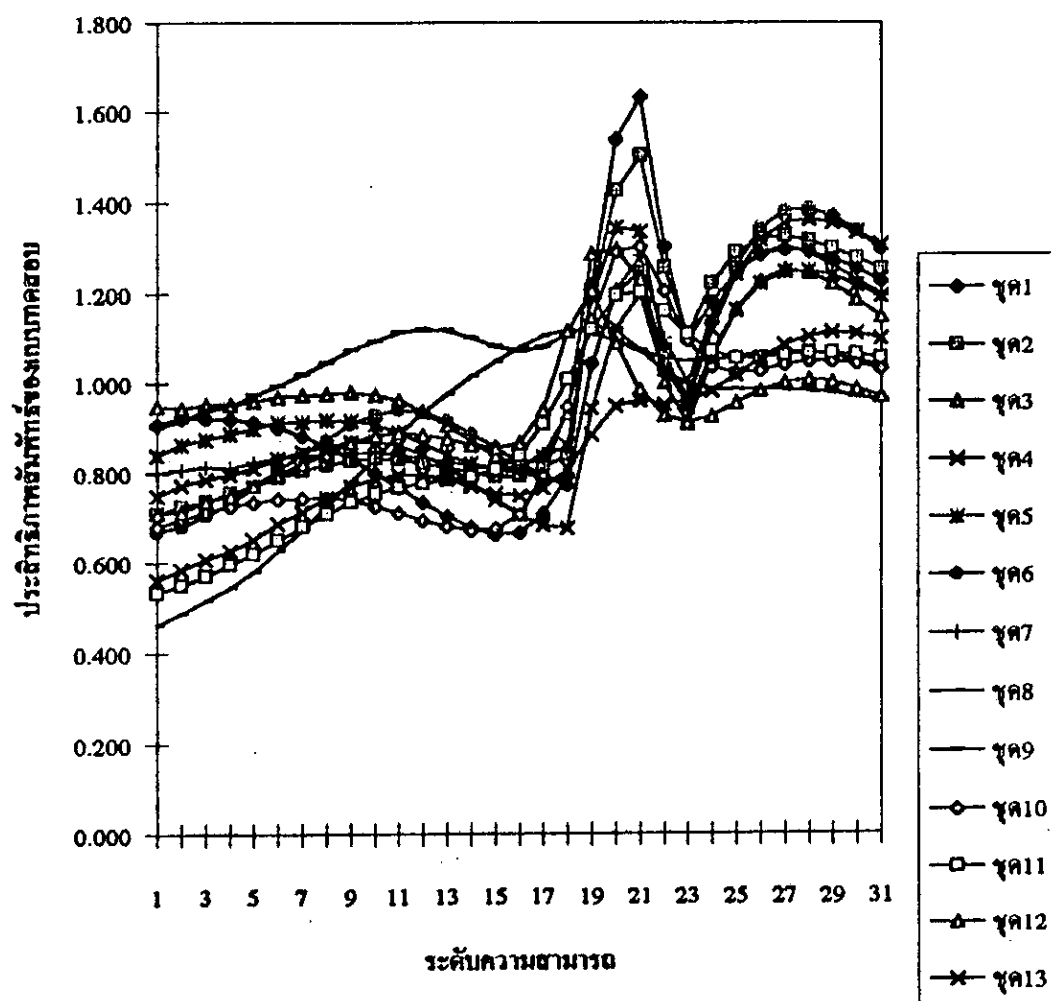


ภาพประกอบ 74 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 18 จำนวน 700/400 คน

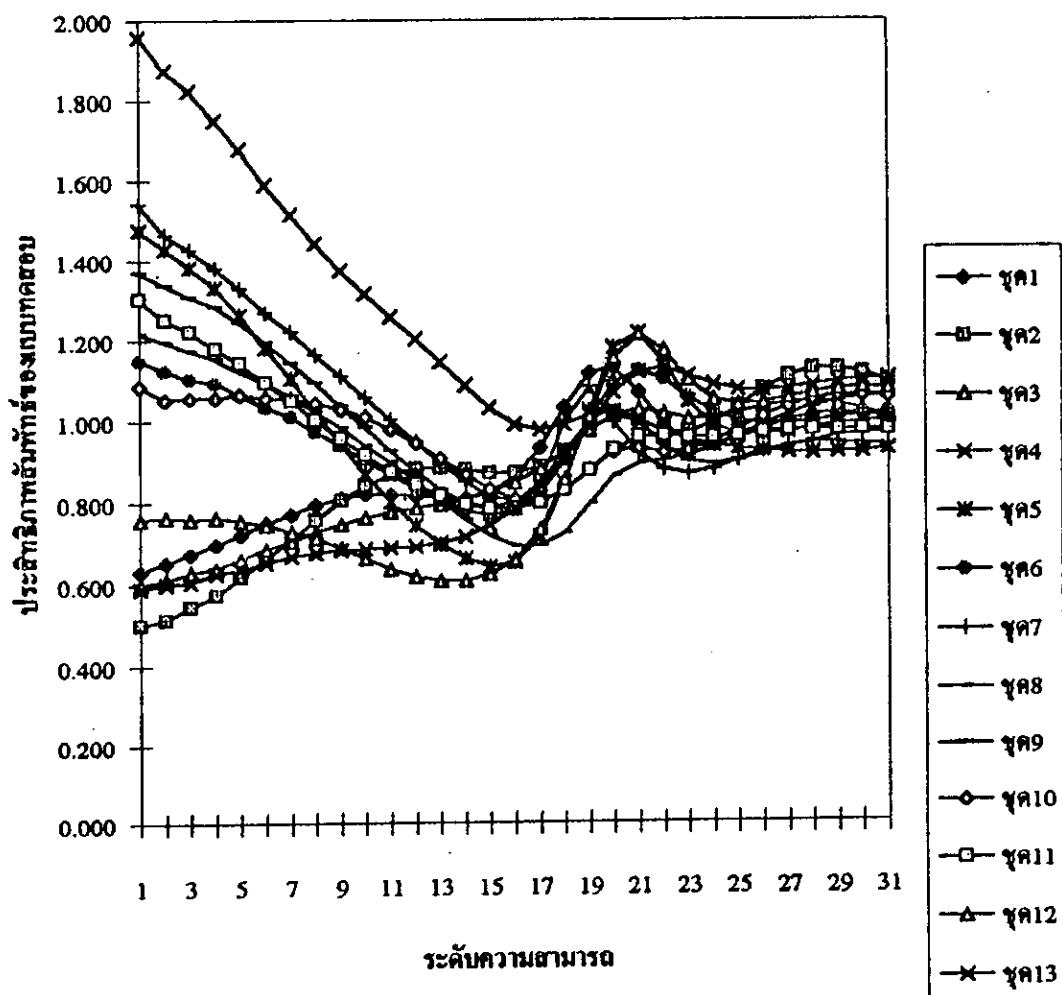


ภาพประกอบ 74 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

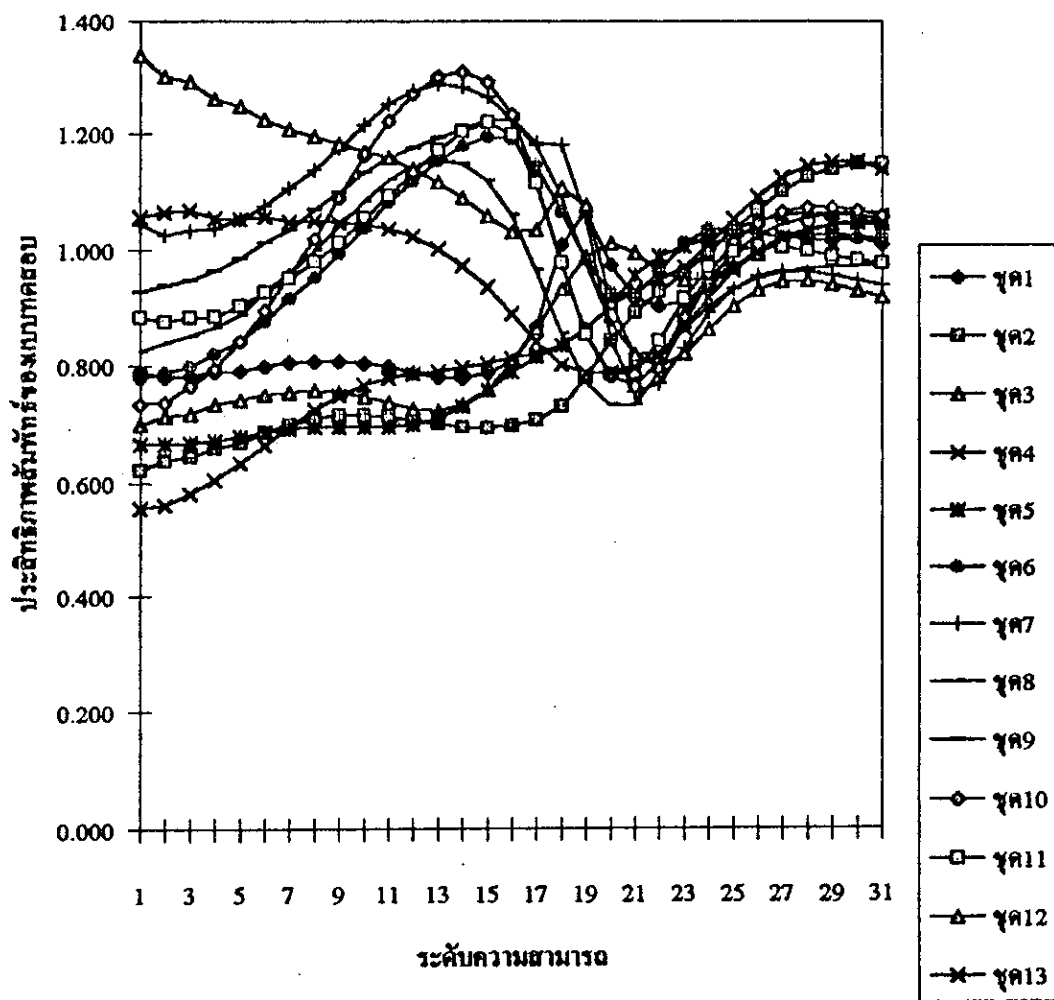


ภาพประกอบ 74 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

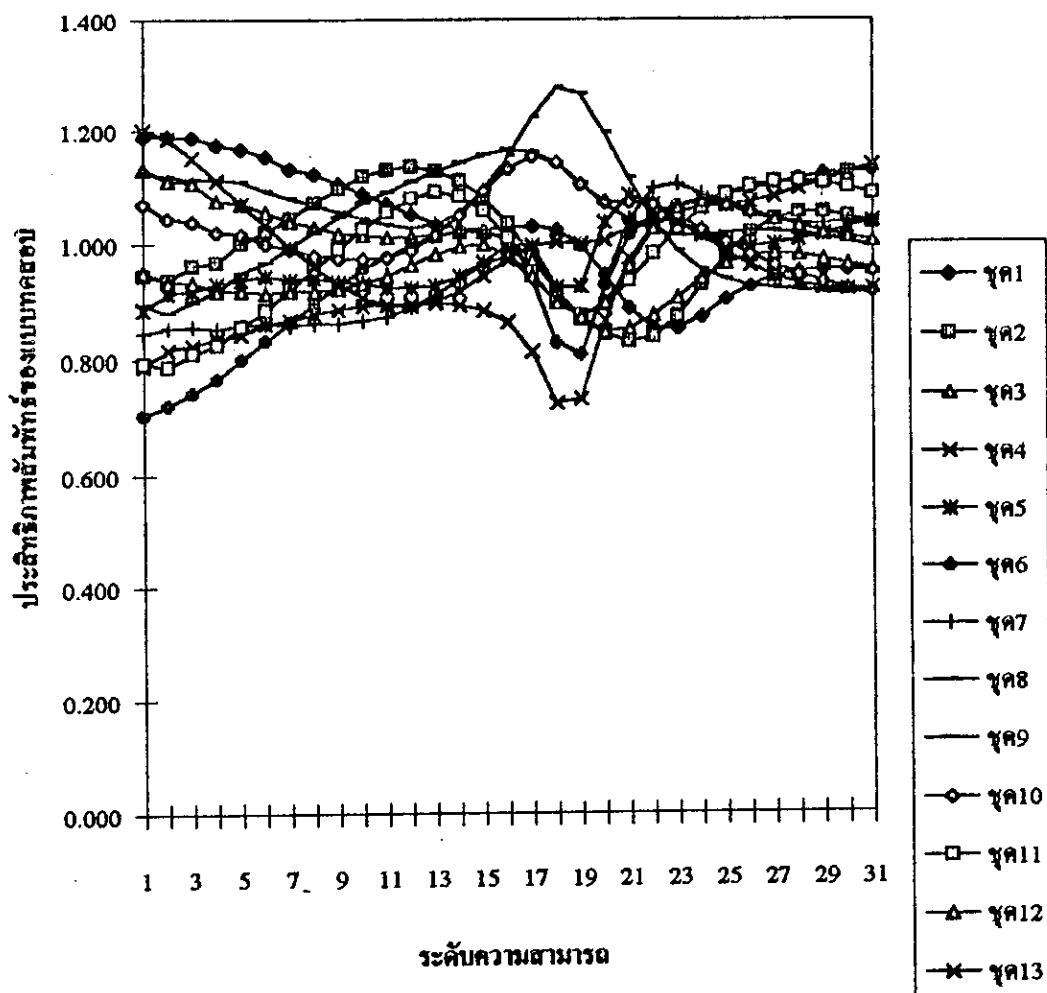


ภาพประกอบ 74 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

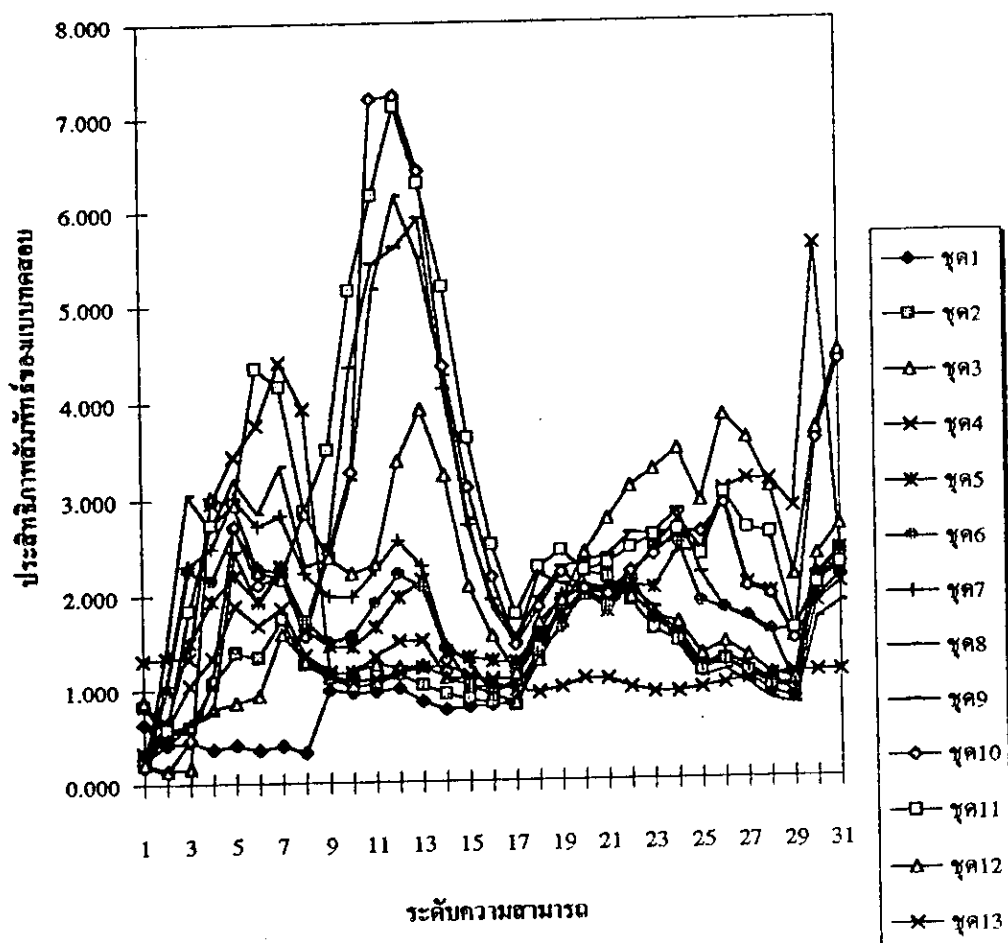


ภาพประกอบ 74 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

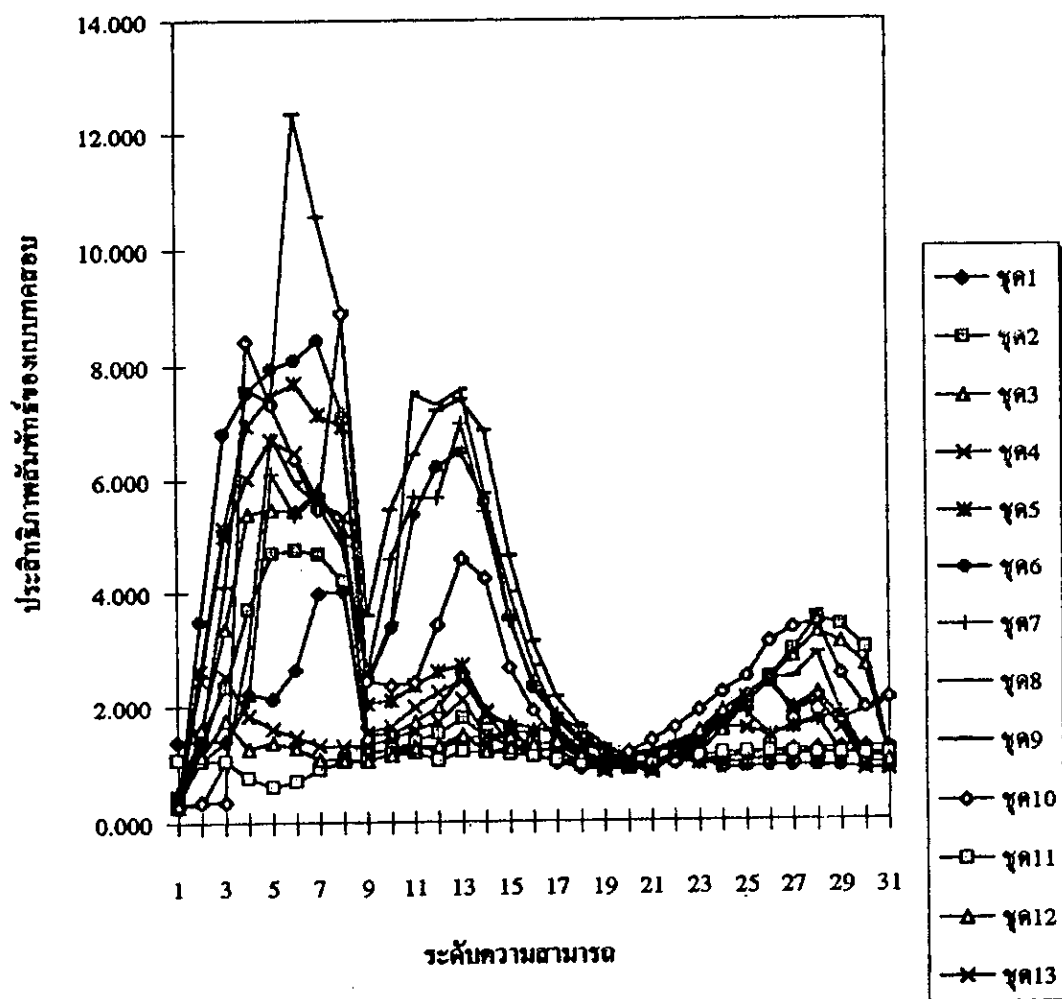


ภาพประกอบ 75 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีวิสต์ค เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

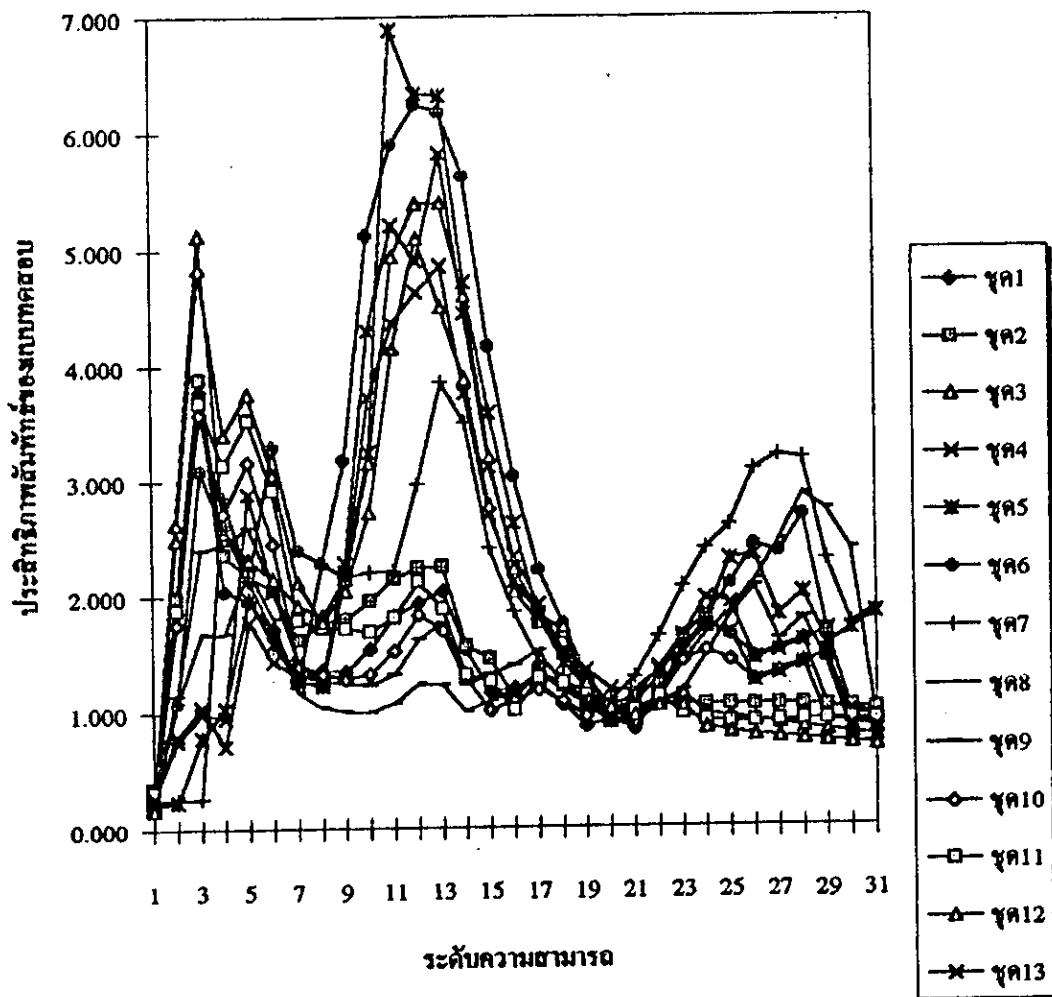


ภาพประกอบ 75 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

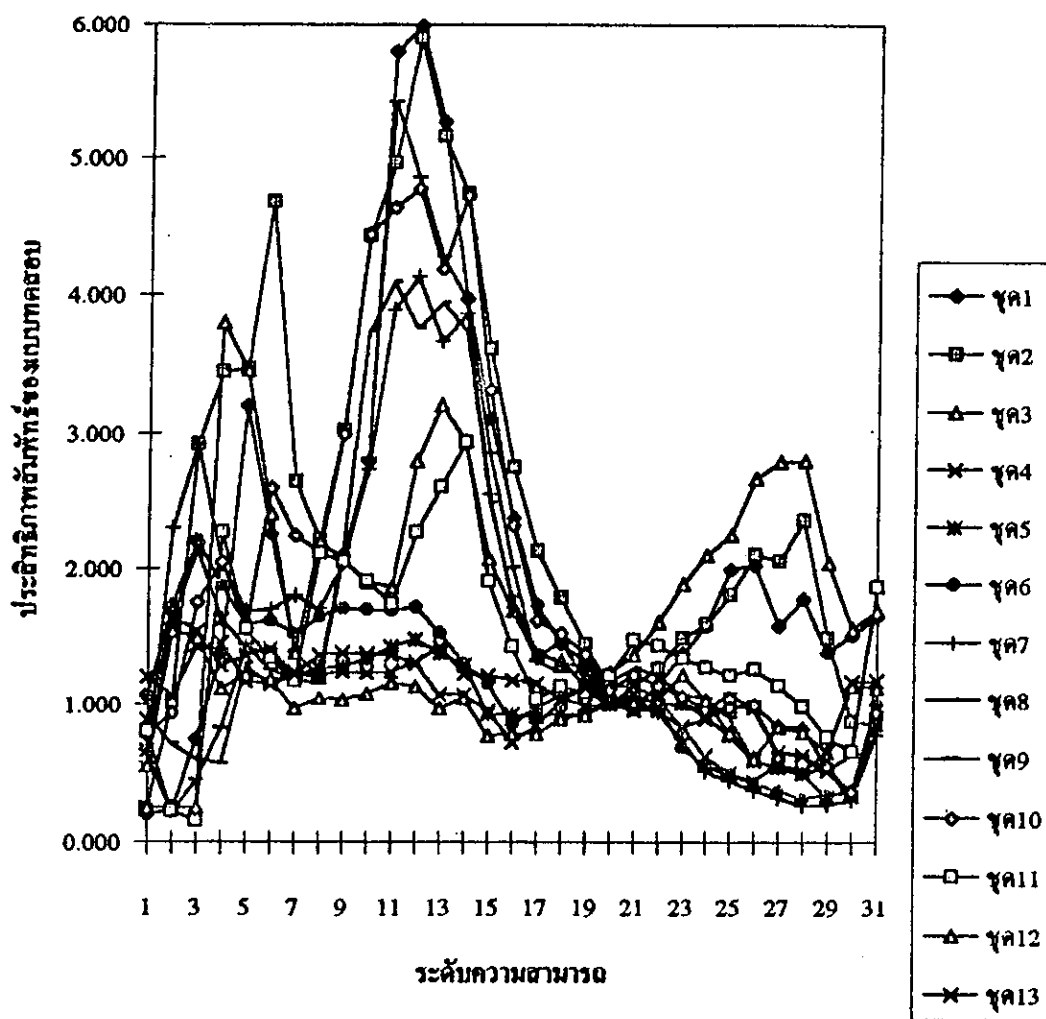


ภาพประกอบ 75 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

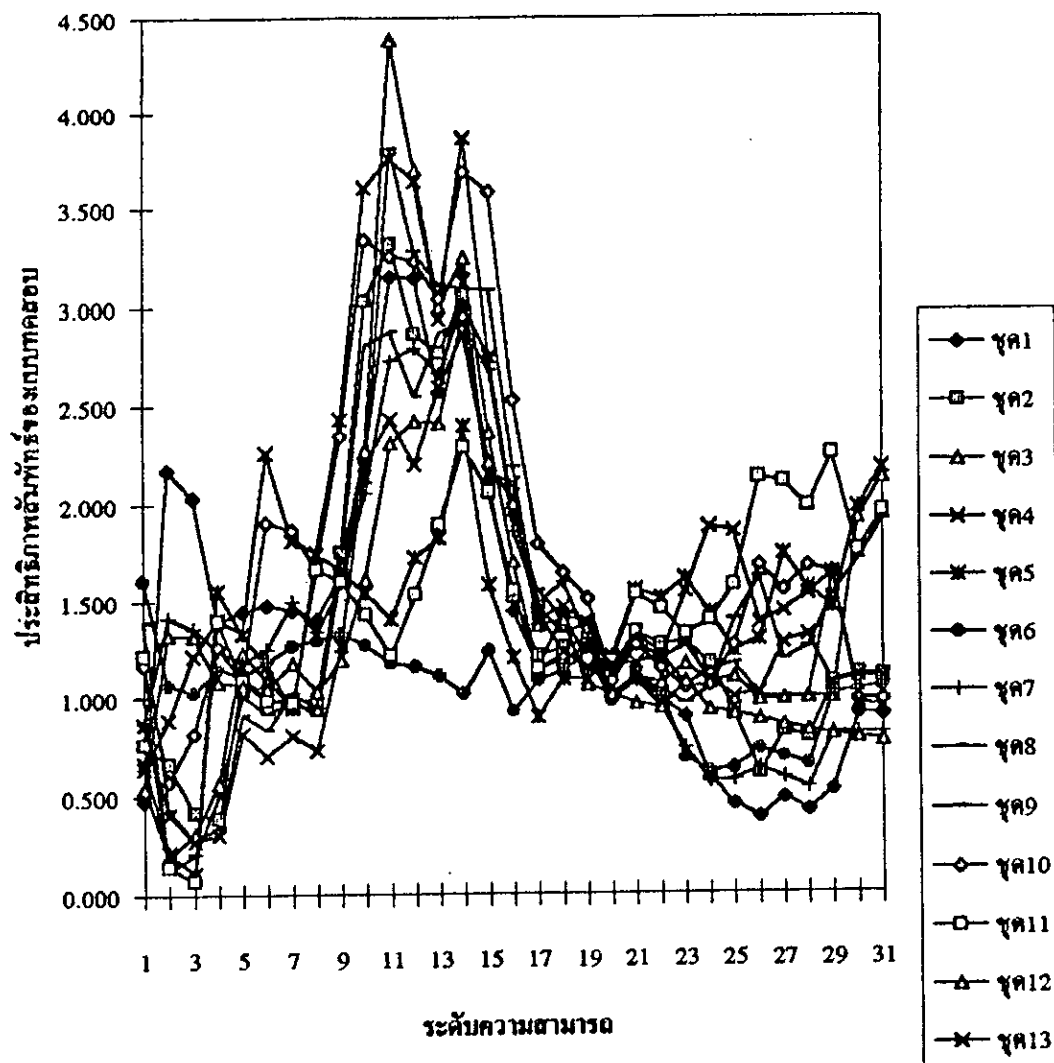


ภาพประกอบ 75 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

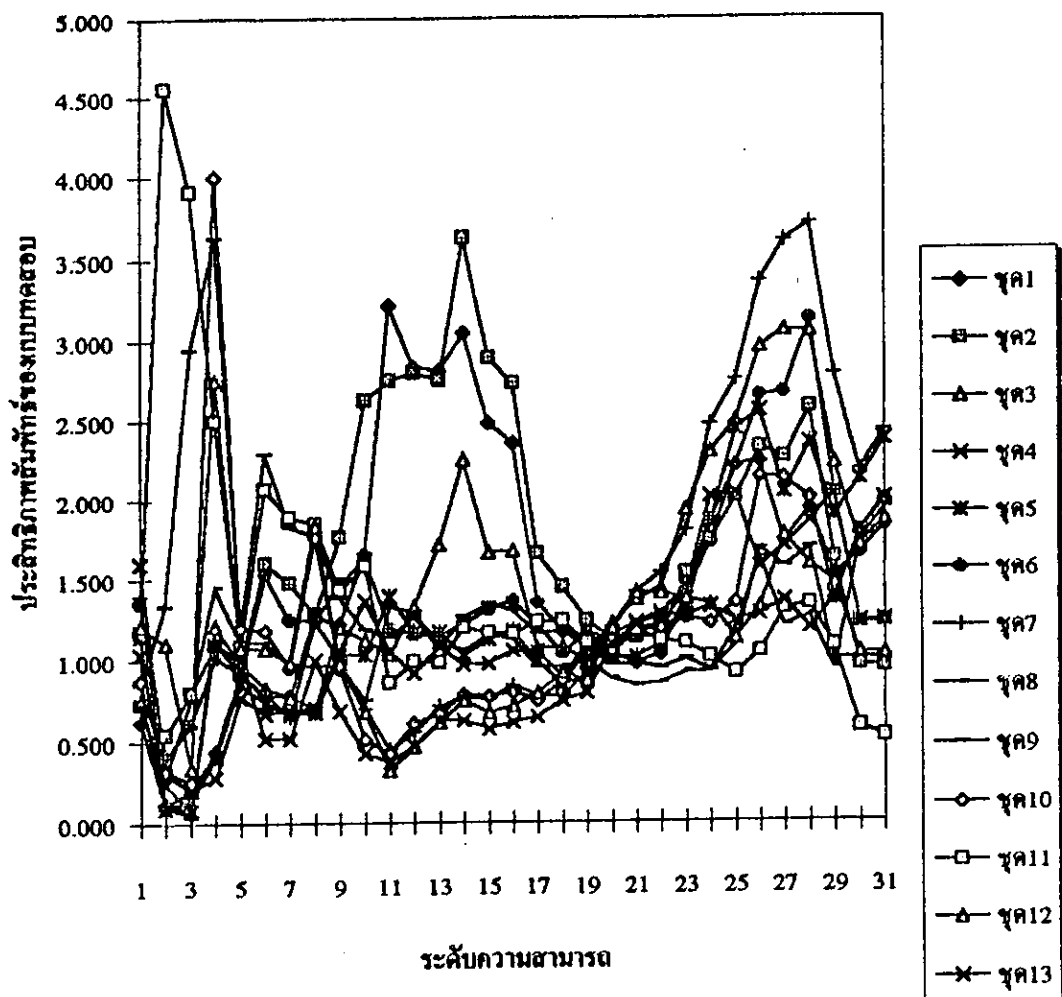


ภาพประกอบ 75 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน



ภาพประกอบ 75 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

จากภาพประกอบ 70 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 50 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวิธีสถิติ ระหว่างกลุ่มตัวอย่าง 500 คน กับ 200 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(500, 200) = 4.000 ; e = -3.000\}$

จากภาพประกอบ 71 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 60 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวิธีสถิติ ระหว่างกลุ่มตัวอย่าง 500 คน กับ 200 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(500, 200) = 3.700 ; e = -3.000\}$

จากภาพประกอบ 72 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 70 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวิธีสถิติ ระหว่างกลุ่มตัวอย่าง 400 คน กับ 200 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(400, 200) = 2.864 ; e = -3.000\}$

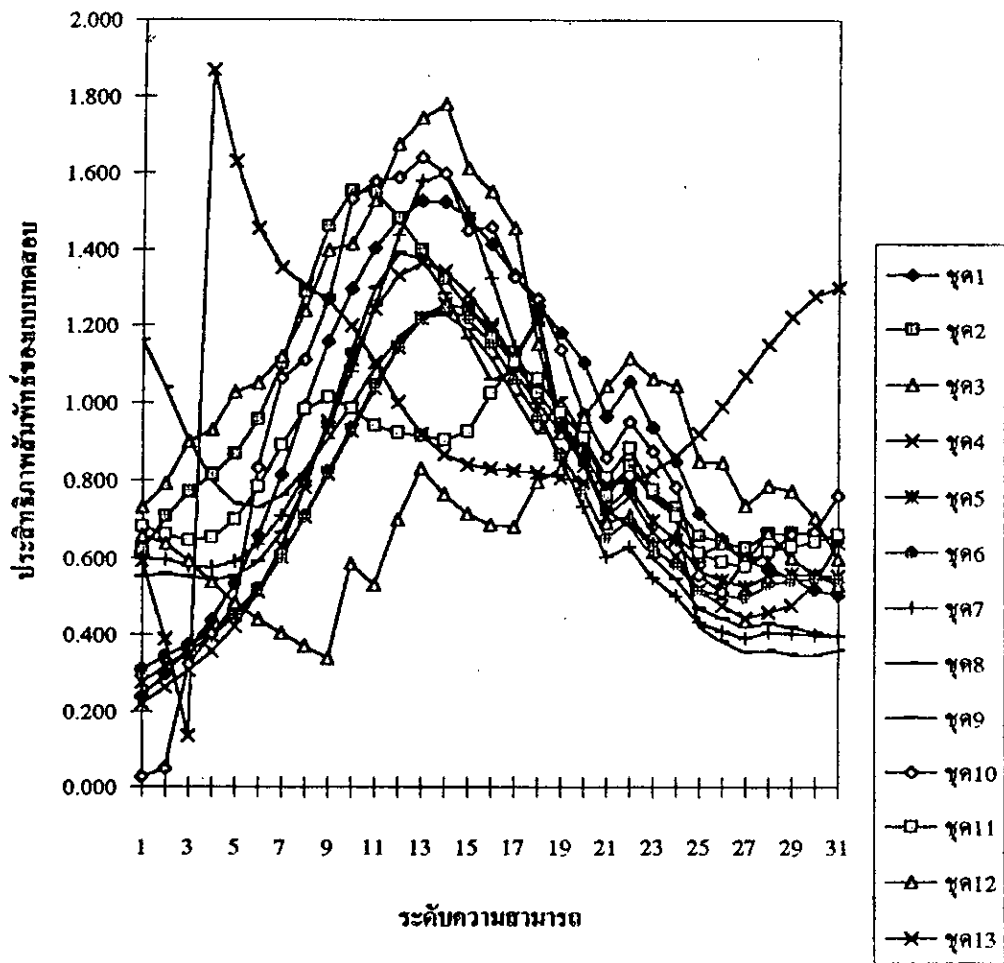
จากภาพประกอบ 73 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 80 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวิธีสถิติ ระหว่างกลุ่มตัวอย่าง 700 คน กับ 200 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(700, 200) = 2.667 ; e = -3.000\}$

จากภาพประกอบ 74 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 90 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวิธีสถิติ ระหว่างกลุ่มตัวอย่าง 500 คน กับ 200 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(500, 200) = 3.030 ; e = -3.000\}$

จากภาพประกอบ 75 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 100 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวิธีสถิติ ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 300 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(2,500, 300) = 12.351 ; e = -2.000\}$

ดังนั้น จากภาพประกอบ 70 - 75 แสดงว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 100 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวิธีสถิติระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 300 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(2,500, 300) = 12.351 ; e = -2.000\}$

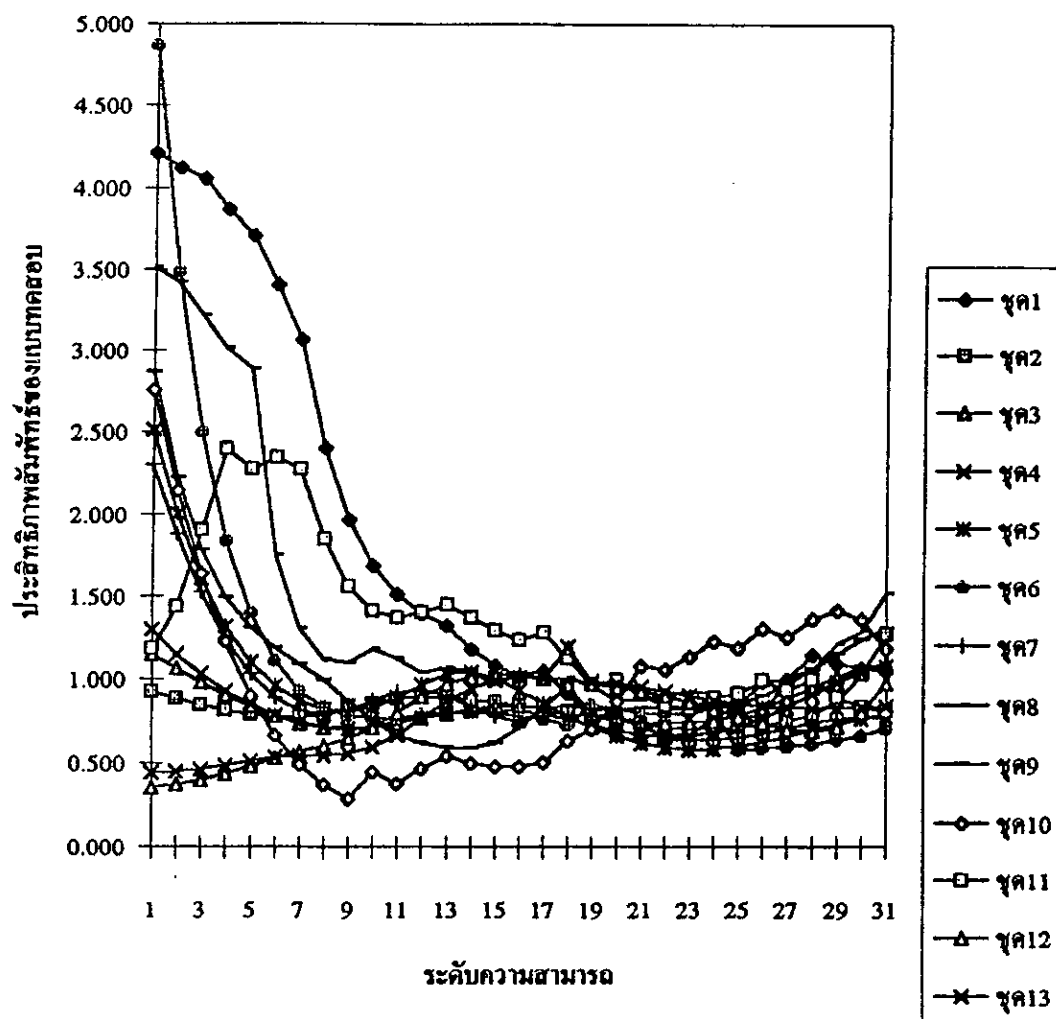
7.3 ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 76 - 81



ภาพประกอบ 76 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดฮวี่ของเบส์ เมื่อข้อสอบมีขนาด 50 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความหนาแน่นต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน	ชุด 8 จำนวน 1,000/200 คน
ชุด 2 จำนวน 400/200 คน	ชุด 9 จำนวน 1,500/200 คน
ชุด 3 จำนวน 500/200 คน	ชุด 10 จำนวน 2,000/200 คน
ชุด 4 จำนวน 600/200 คน	ชุด 11 จำนวน 2,500/200 คน
ชุด 5 จำนวน 700/200 คน	ชุด 12 จำนวน 3,000/200 คน
ชุด 6 จำนวน 800/200 คน	ชุด 13 จำนวน 400/300 คน
ชุด 7 จำนวน 900/200 คน	

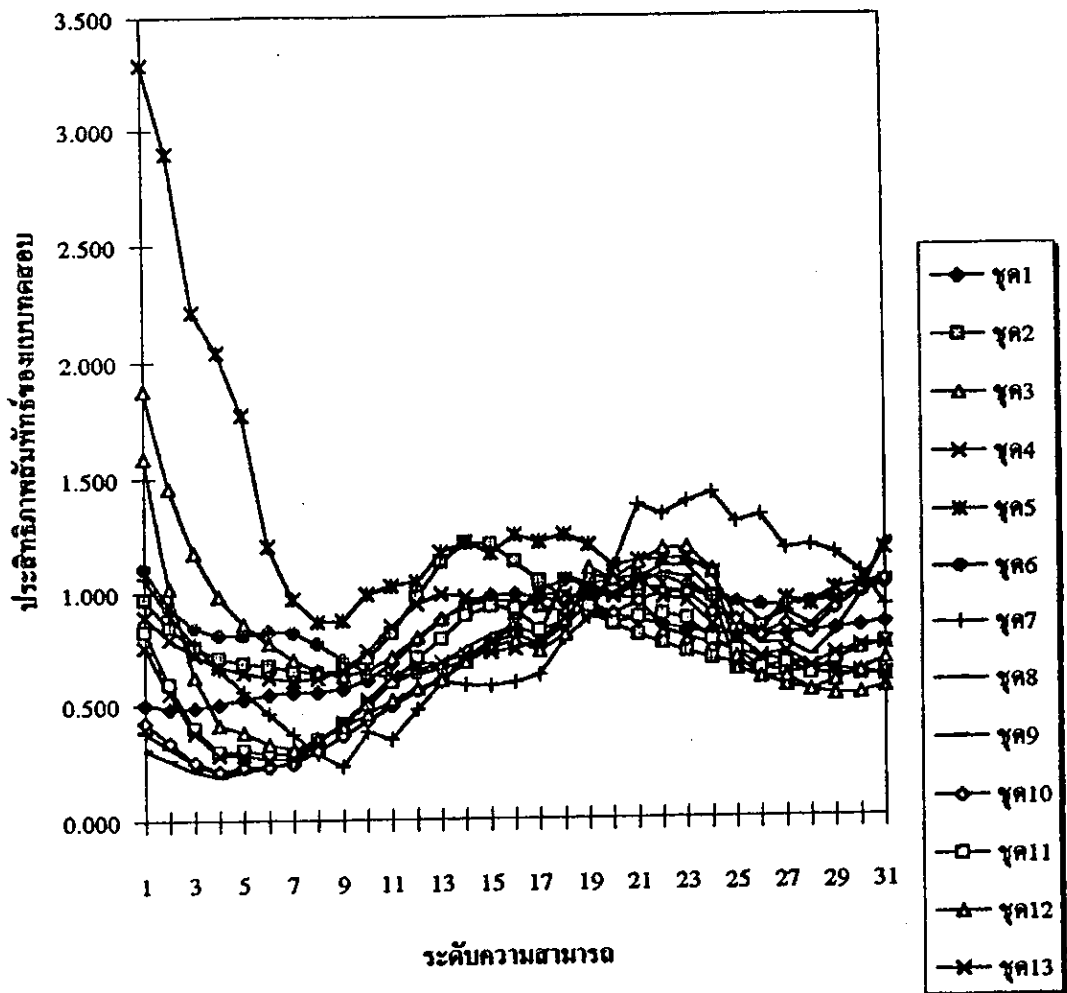


ภาพประกอบ 76 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

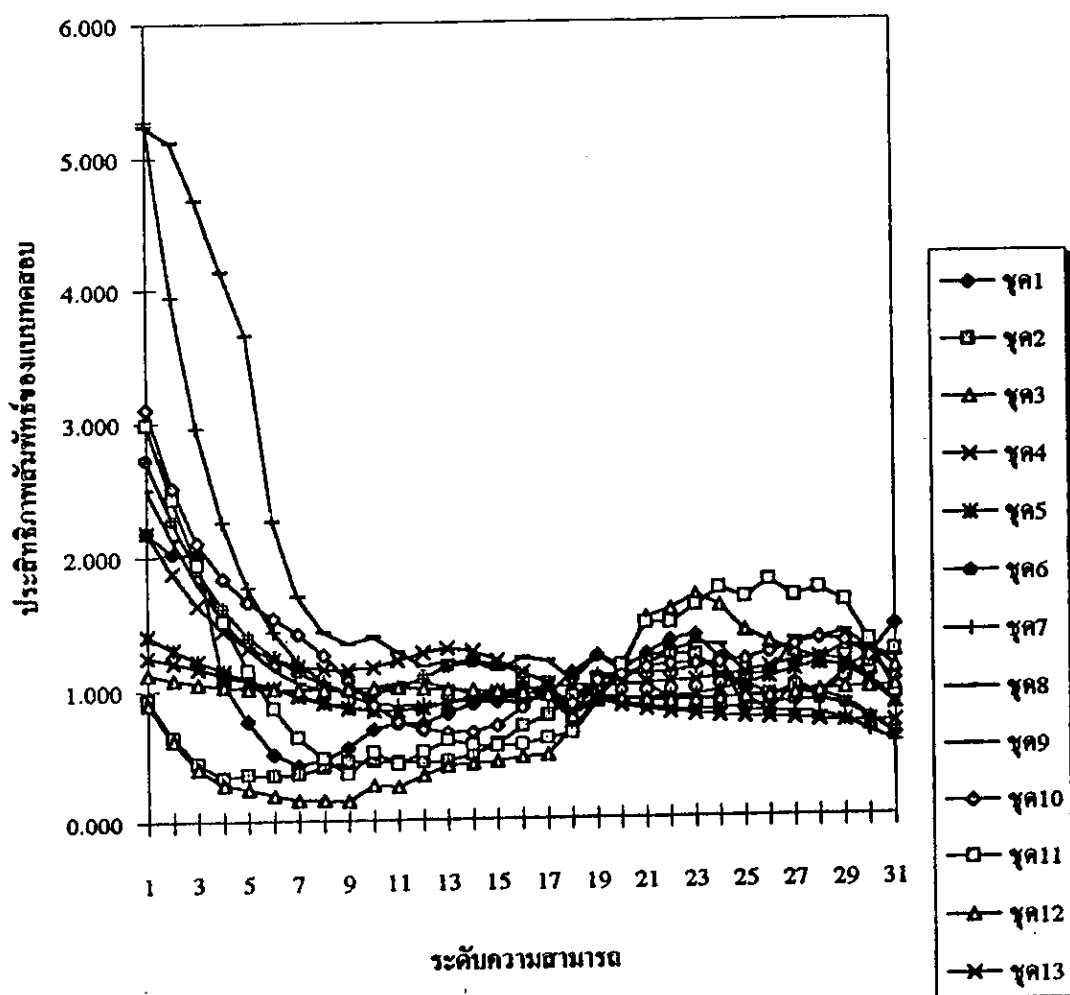


ภาพประกอบ 76 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
ชุด 2 จำนวน 900/400 คน
ชุด 3 จำนวน 1,000/400 คน
ชุด 4 จำนวน 1,500/400 คน
ชุด 5 จำนวน 2,000/400 คน
ชุด 6 จำนวน 2,500/400 คน
ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
ชุด 9 จำนวน 700/500 คน
ชุด 10 จำนวน 800/500 คน
ชุด 11 จำนวน 900/500 คน
ชุด 12 จำนวน 1,000/500 คน
ชุด 13 จำนวน 1,500/500 คน

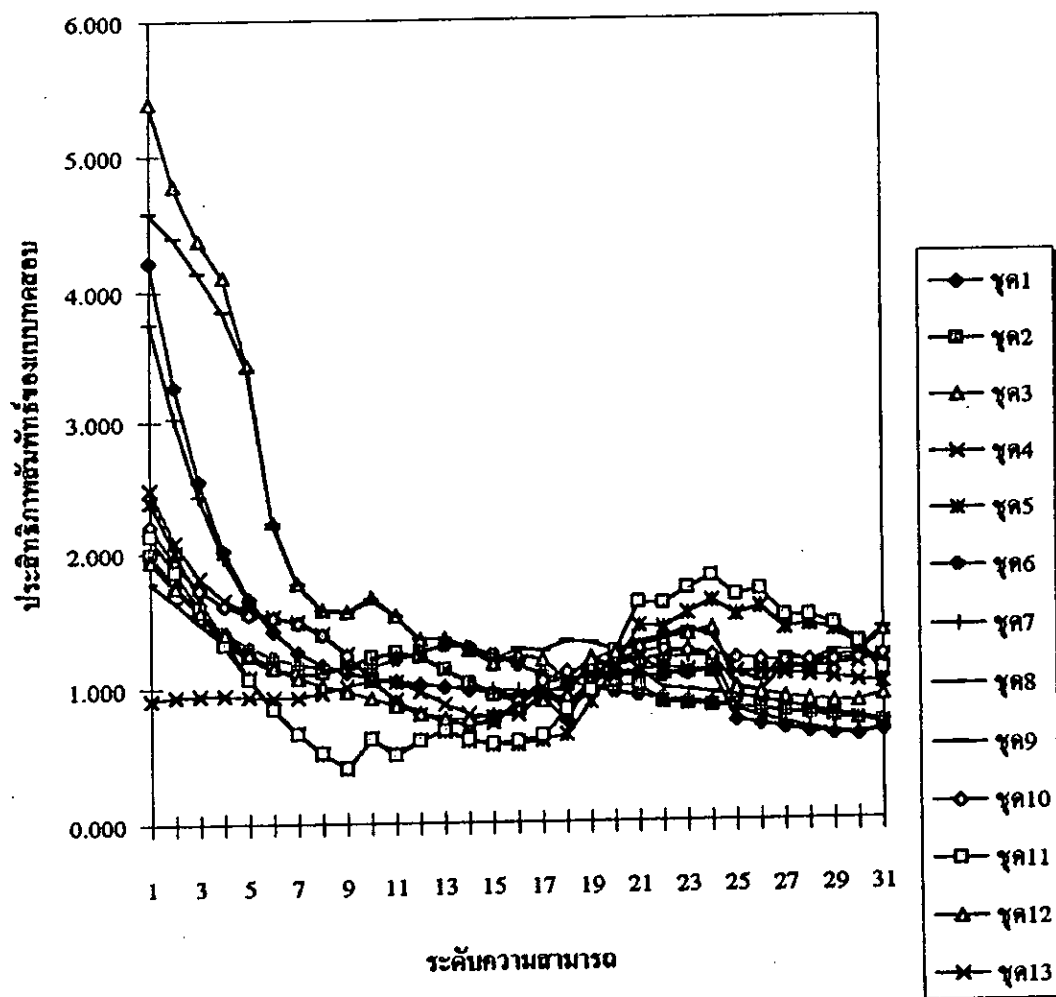


ภาพประกอบ 76 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/800 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

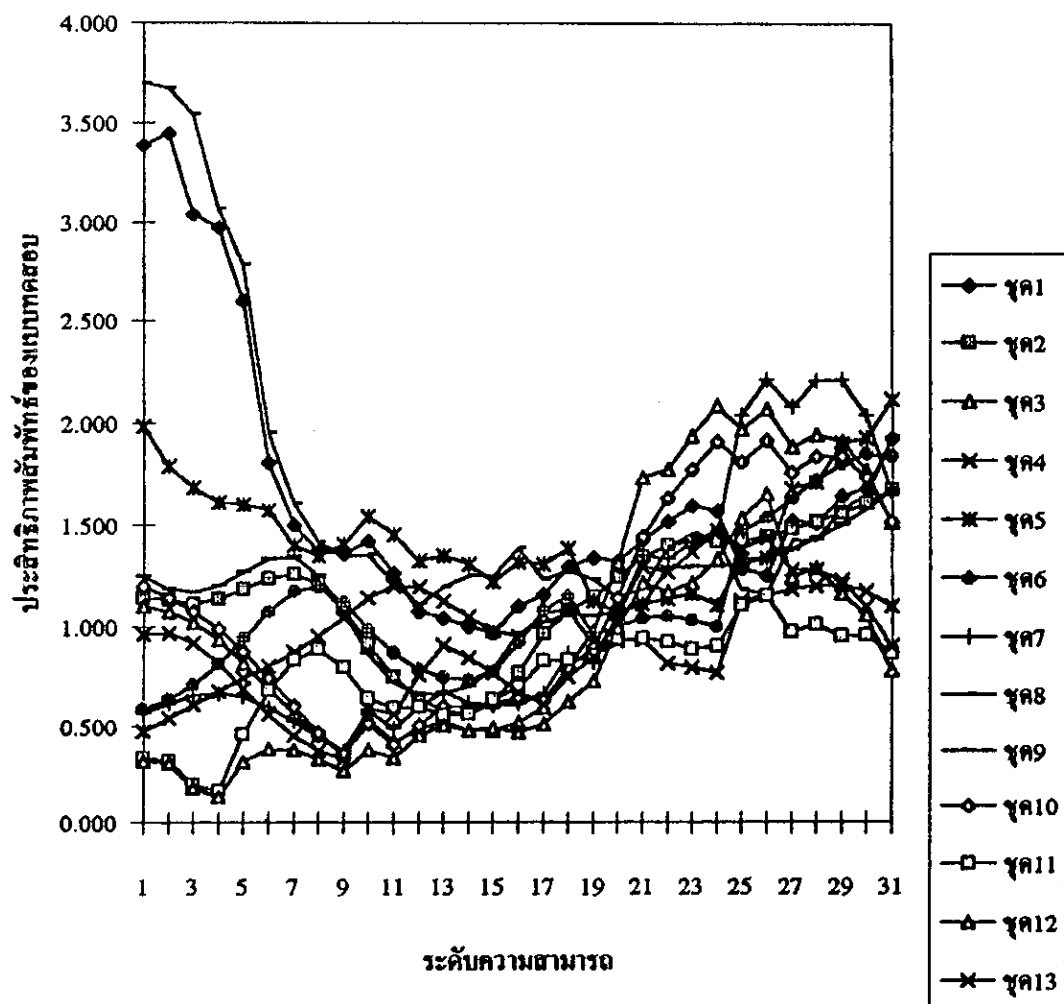


ภาพประกอบ 76 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
ชุด 2 จำนวน 1,500/700 คน
ชุด 3 จำนวน 2,000/700 คน
ชุด 4 จำนวน 2,500/700 คน
ชุด 5 จำนวน 3,000/700 คน
ชุด 6 จำนวน 900/800 คน
ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
ชุด 9 จำนวน 2,000/800 คน
ชุด 10 จำนวน 2,500/800 คน
ชุด 11 จำนวน 3,000/800 คน
ชุด 12 จำนวน 1,000/900 คน
ชุด 13 จำนวน 1,500/900 คน

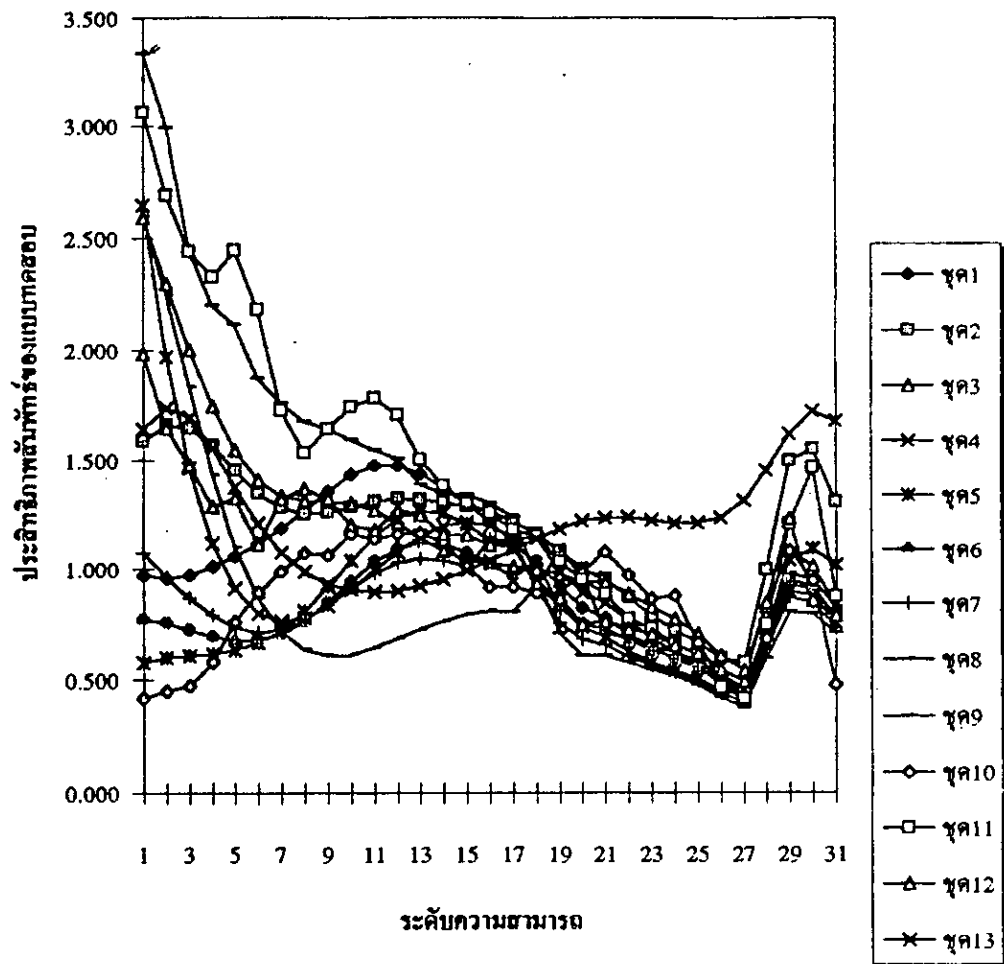


ภาพประกอบ 76 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

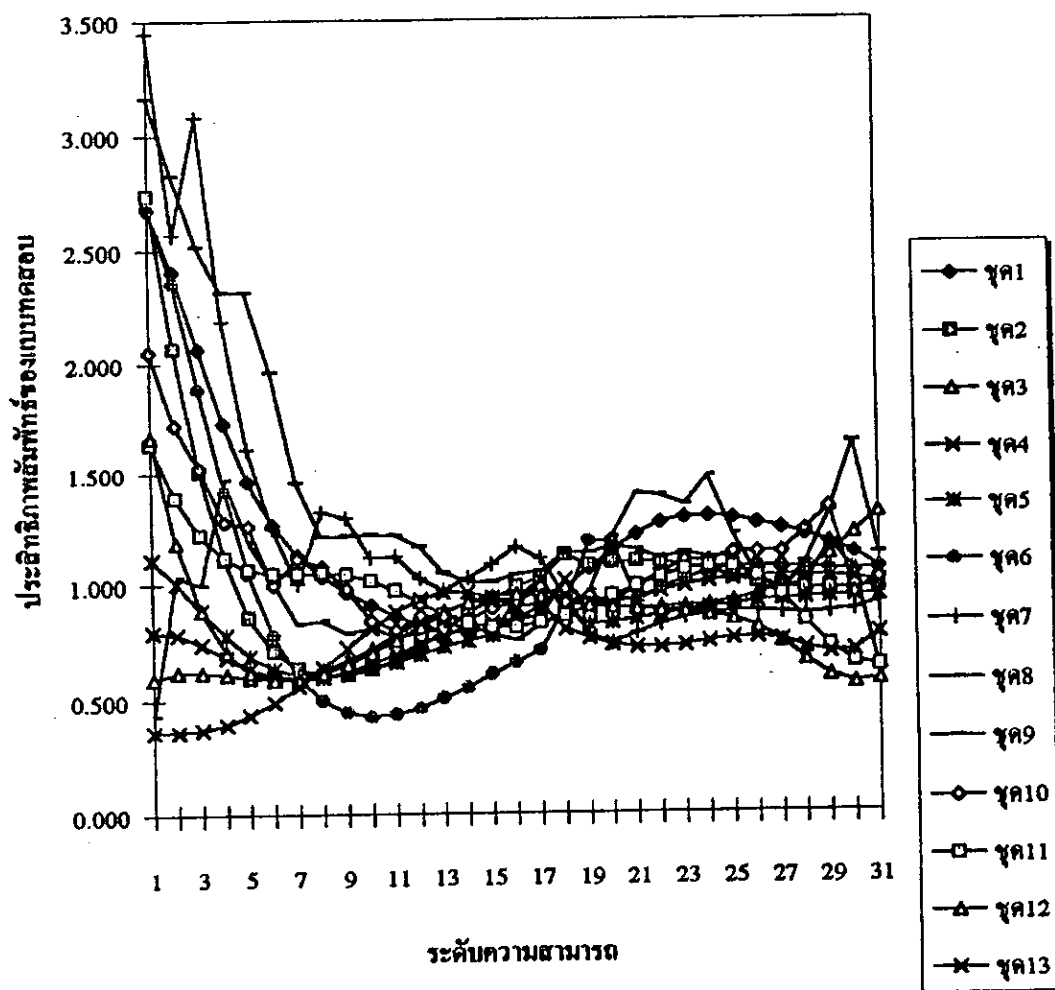


ภาพประกอบ 77 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบมีขนาด 60 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

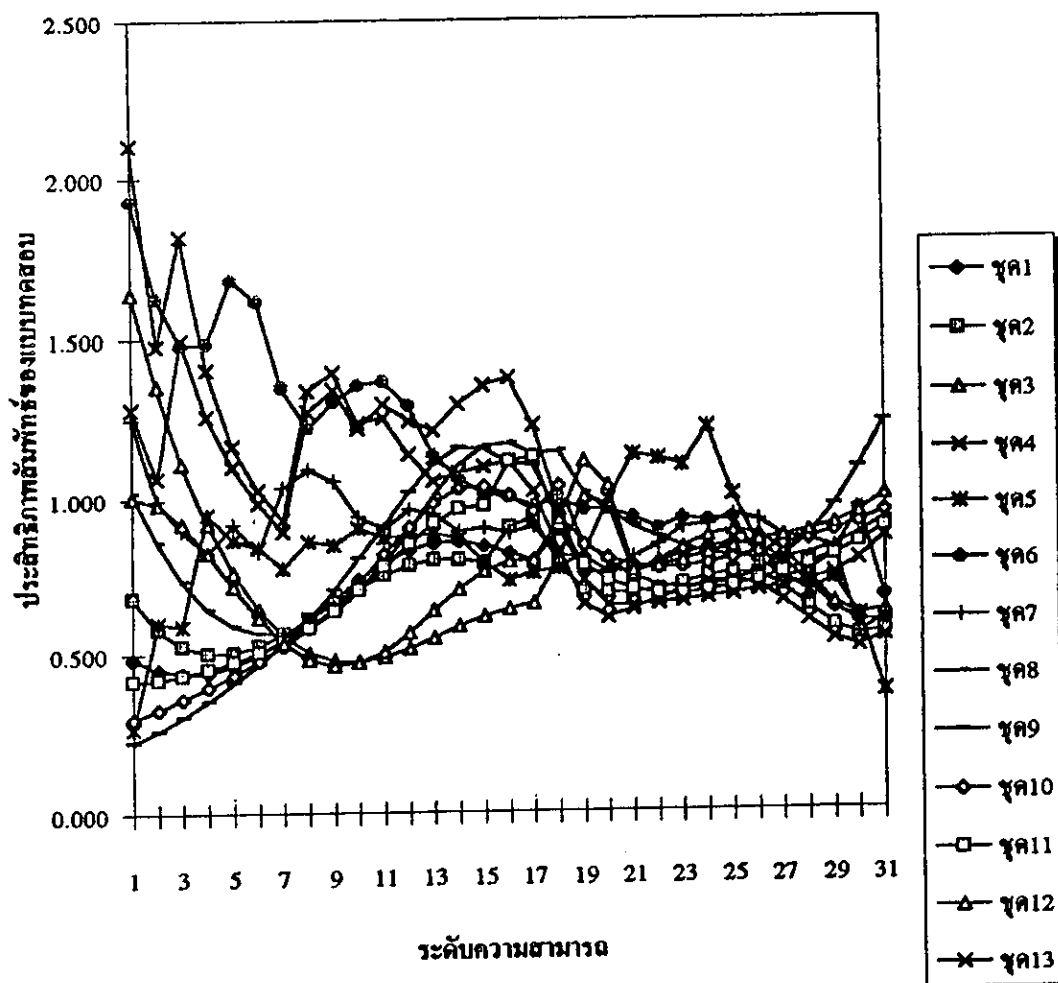


ภาพประกอบ 77 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

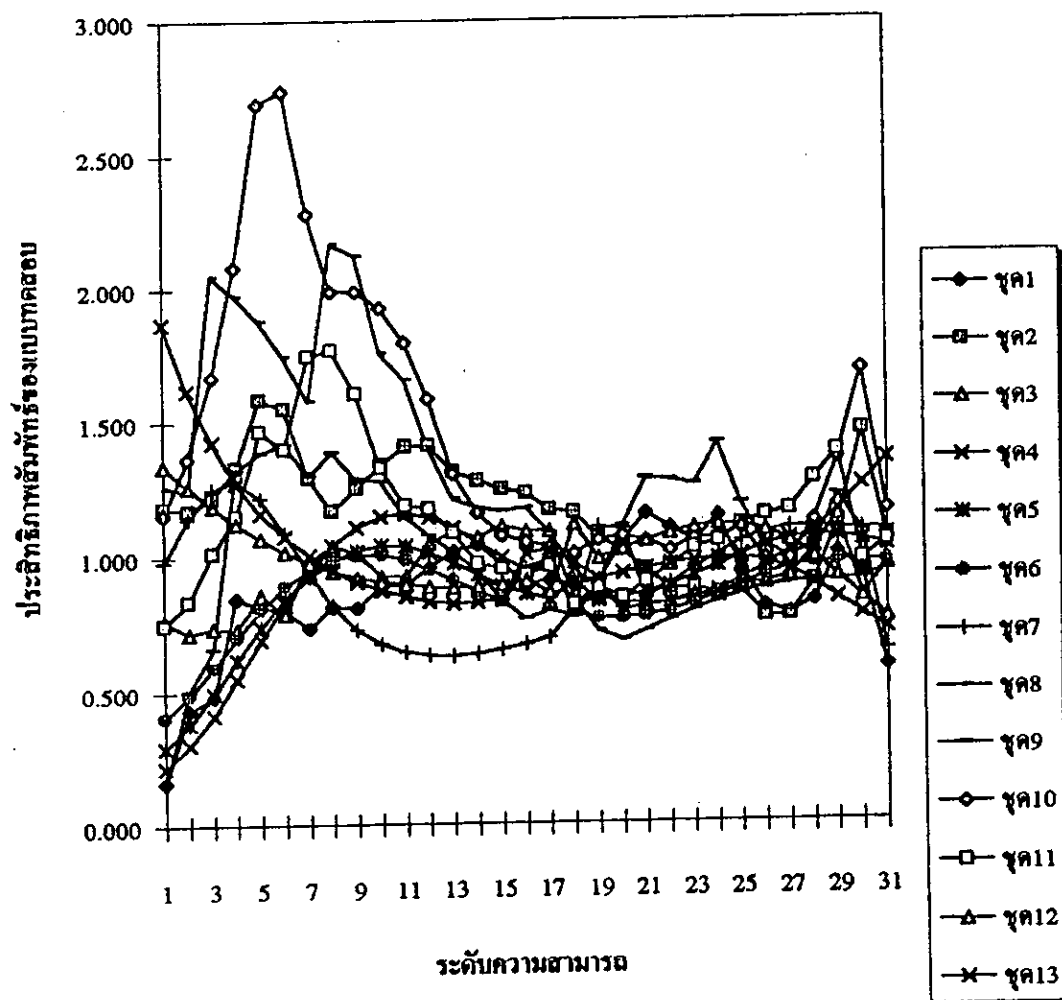


ภาพประกอบ 77 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

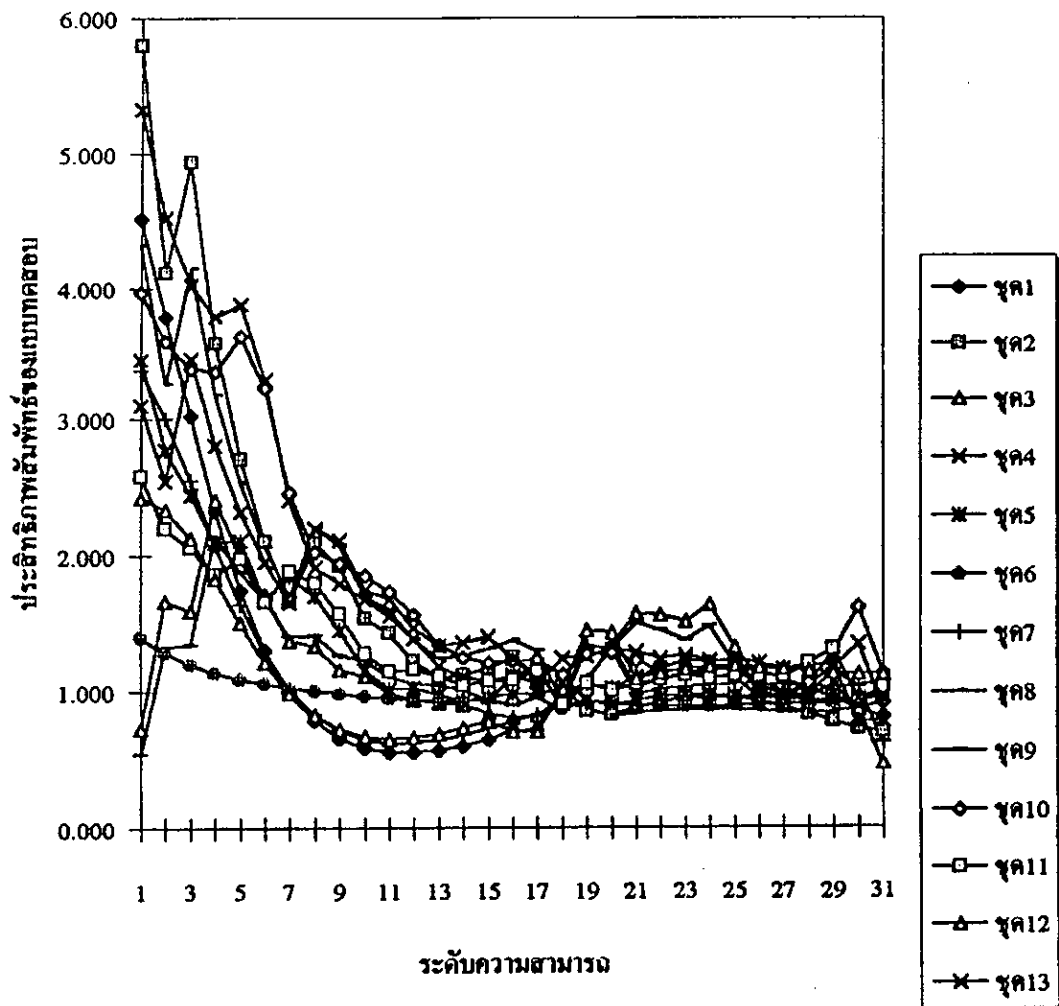


ภาพประกอบ 77 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 2,000/500 คน
 ชด 2 จำนวน 2,500/500 คน
 ชด 3 จำนวน 3,000/500 คน
 ชด 4 จำนวน 700/600 คน
 ชด 5 จำนวน 800/600 คน
 ชด 6 จำนวน 900/600 คน
 ชด 7 จำนวน 1,000/600 คน

ชด 8 จำนวน 1,500/600 คน
 ชด 9 จำนวน 2,000/600 คน
 ชด 10 จำนวน 2,500/600 คน
 ชด 11 จำนวน 3,000/600 คน
 ชด 12 จำนวน 800/700 คน
 ชด 13 จำนวน 900/700 คน

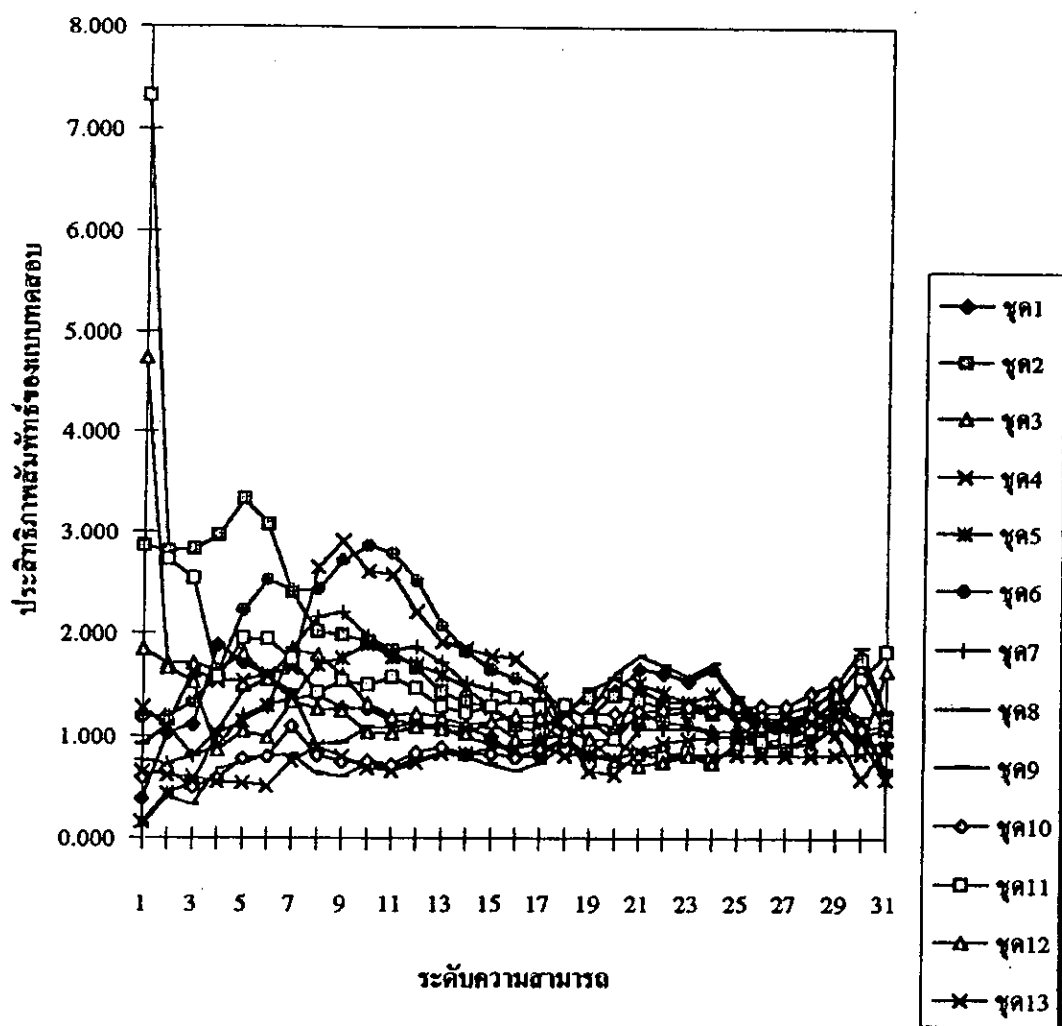


ภาพประกอบ 77 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

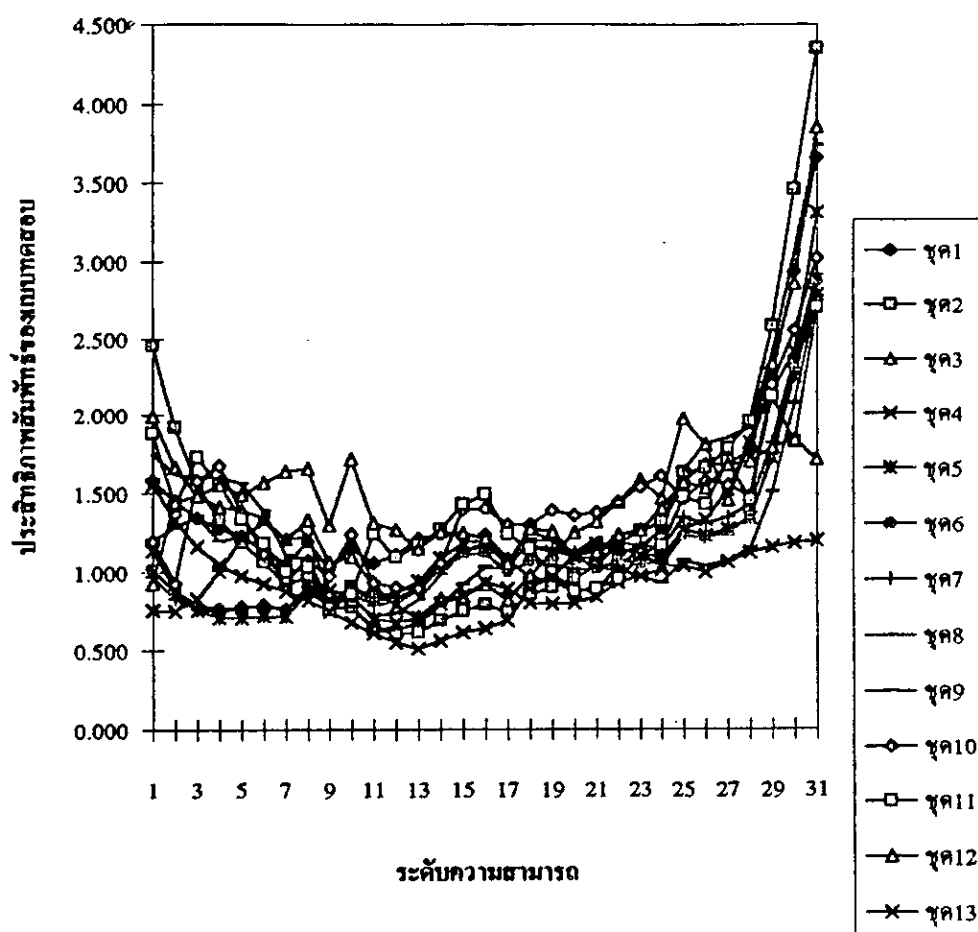


ภาพประกอบ 77 (ต่อ)

หมายเหตุ

ชค 1 จำนวน 2,000/900 คน
 ชค 2 จำนวน 2,500/900 คน
 ชค 3 จำนวน 3,000/900 คน
 ชค 4 จำนวน 1,500/1,000 คน
 ชค 5 จำนวน 2,000/1,000 คน
 ชค 6 จำนวน 2,500/1,000 คน
 ชค 7 จำนวน 3,000/1,000 คน

ชค 8 จำนวน 2,000/1,500 คน
 ชค 9 จำนวน 2,500/1,500 คน
 ชค 10 จำนวน 3,000/1,500 คน
 ชค 11 จำนวน 2,500/2,000 คน
 ชค 12 จำนวน 3,000/2,000 คน
 ชค 13 จำนวน 3,000/2,500 คน

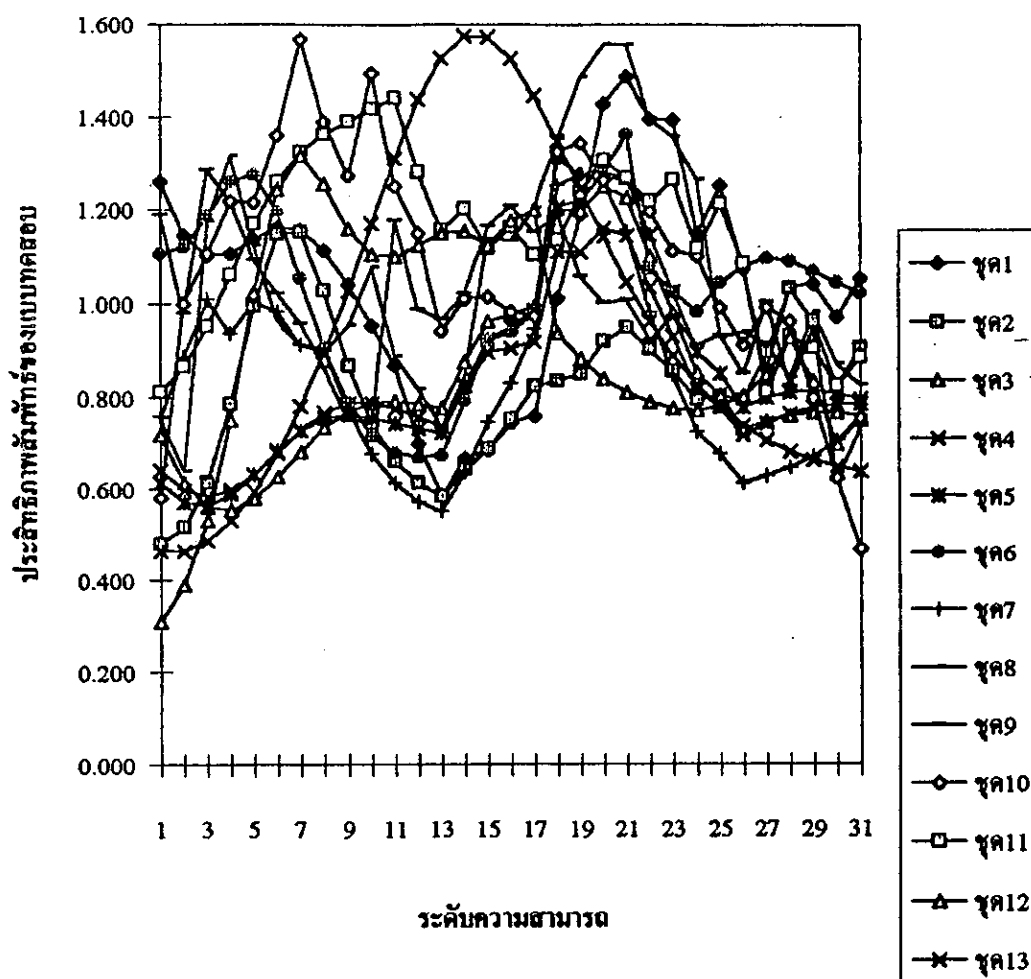


ภาพประกอบ 78 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โคชีวีซีของเบส์ เมื่อข้อสอบมีขนาด 70 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 300/200 คน
 ชด 2 จำนวน 400/200 คน
 ชด 3 จำนวน 500/200 คน
 ชด 4 จำนวน 600/200 คน
 ชด 5 จำนวน 700/200 คน
 ชด 6 จำนวน 800/200 คน
 ชด 7 จำนวน 900/200 คน

ชด 8 จำนวน 1,000/200 คน
 ชด 9 จำนวน 1,500/200 คน
 ชด 10 จำนวน 2,000/200 คน
 ชด 11 จำนวน 2,500/200 คน
 ชด 12 จำนวน 3,000/200 คน
 ชด 13 จำนวน 400/300 คน

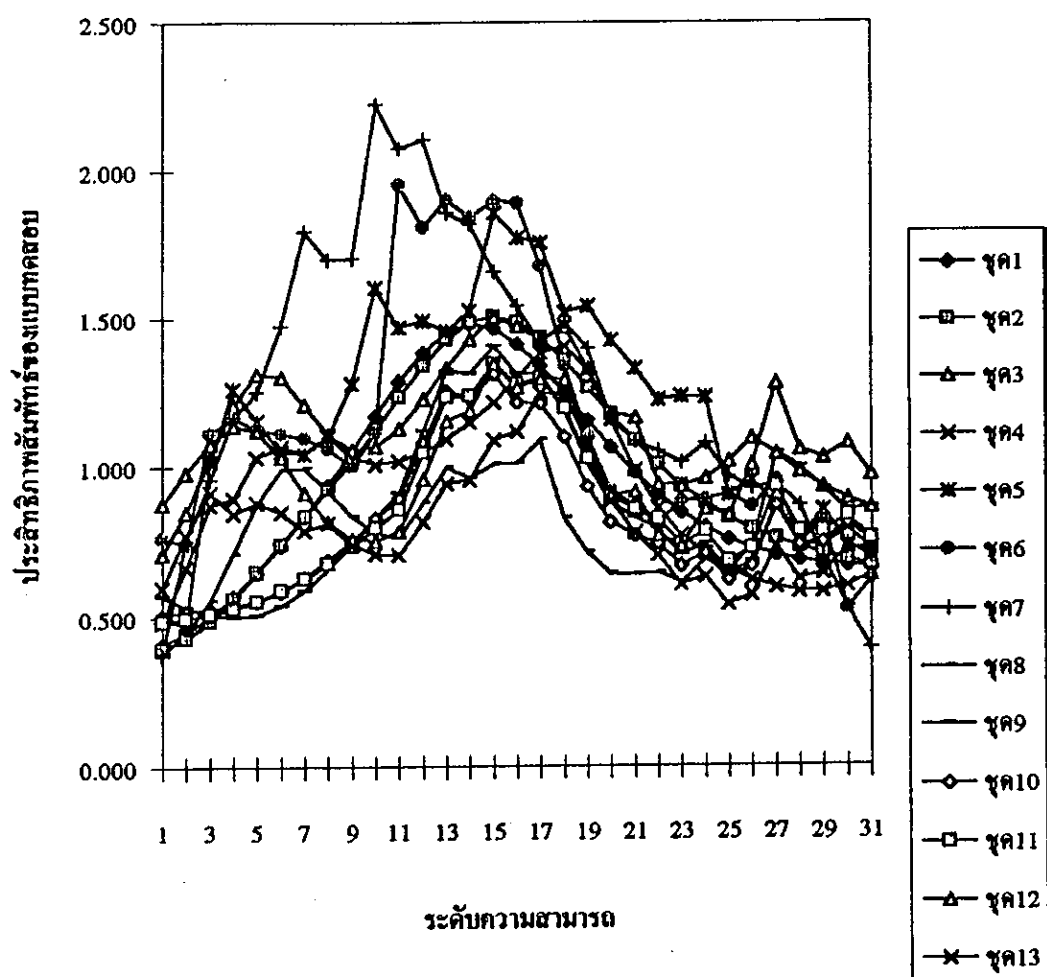


ภาพประกอบ 78 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

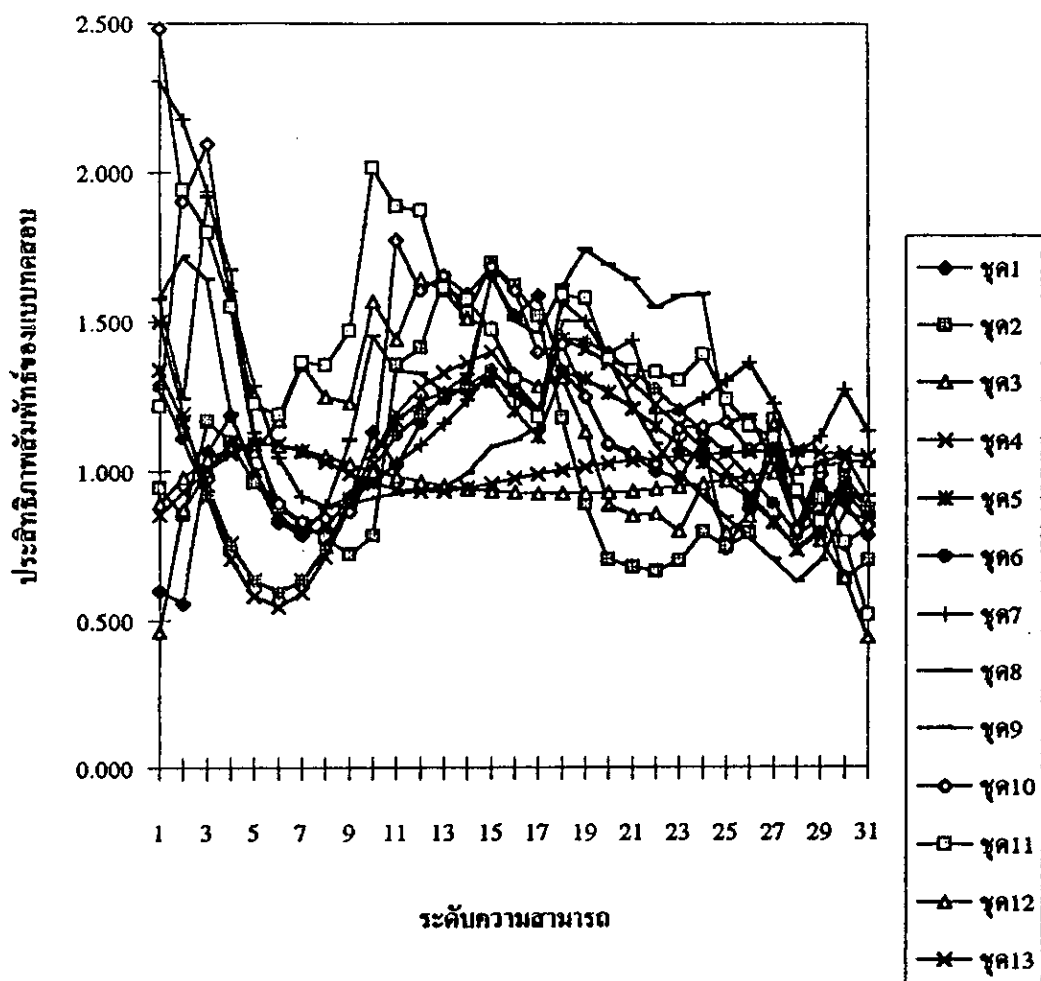


ภาพประกอบ 78 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

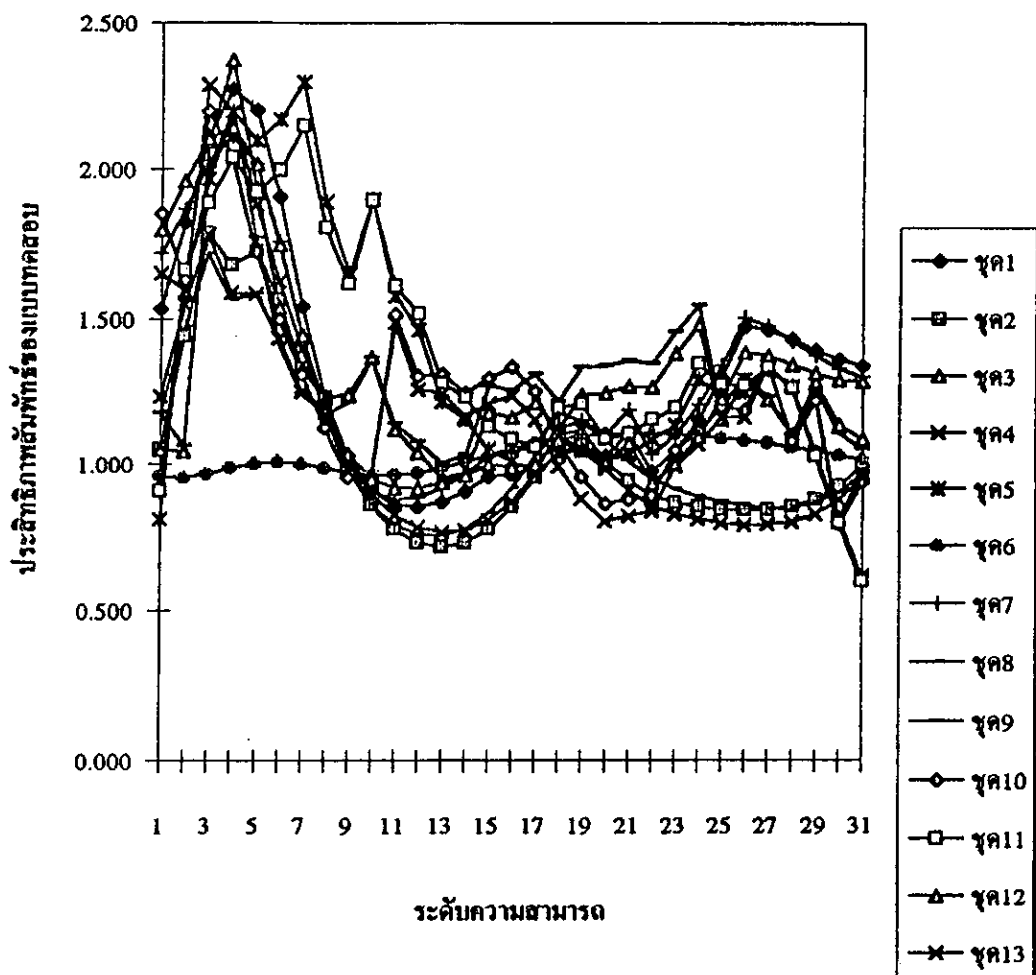


ภาพประกอบ 78 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

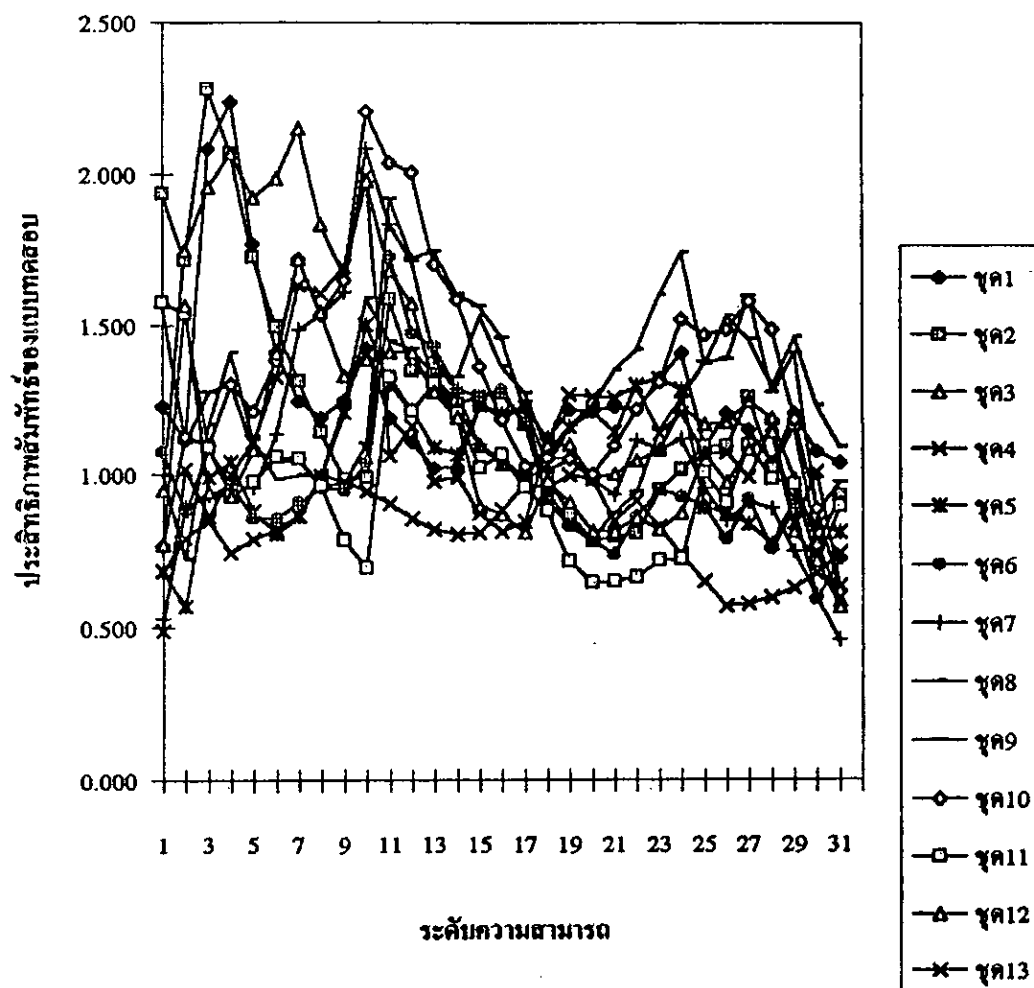


ภาพประกอบ 78 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

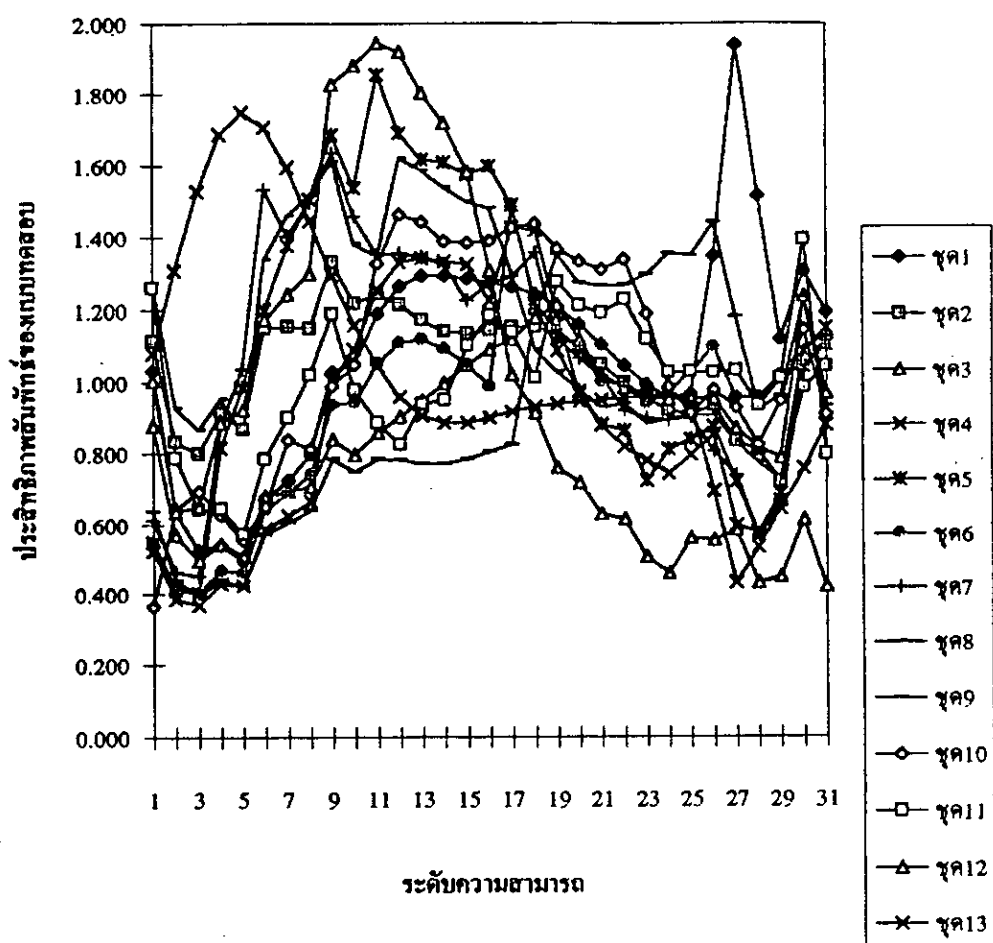


ภาพประกอบ 78 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

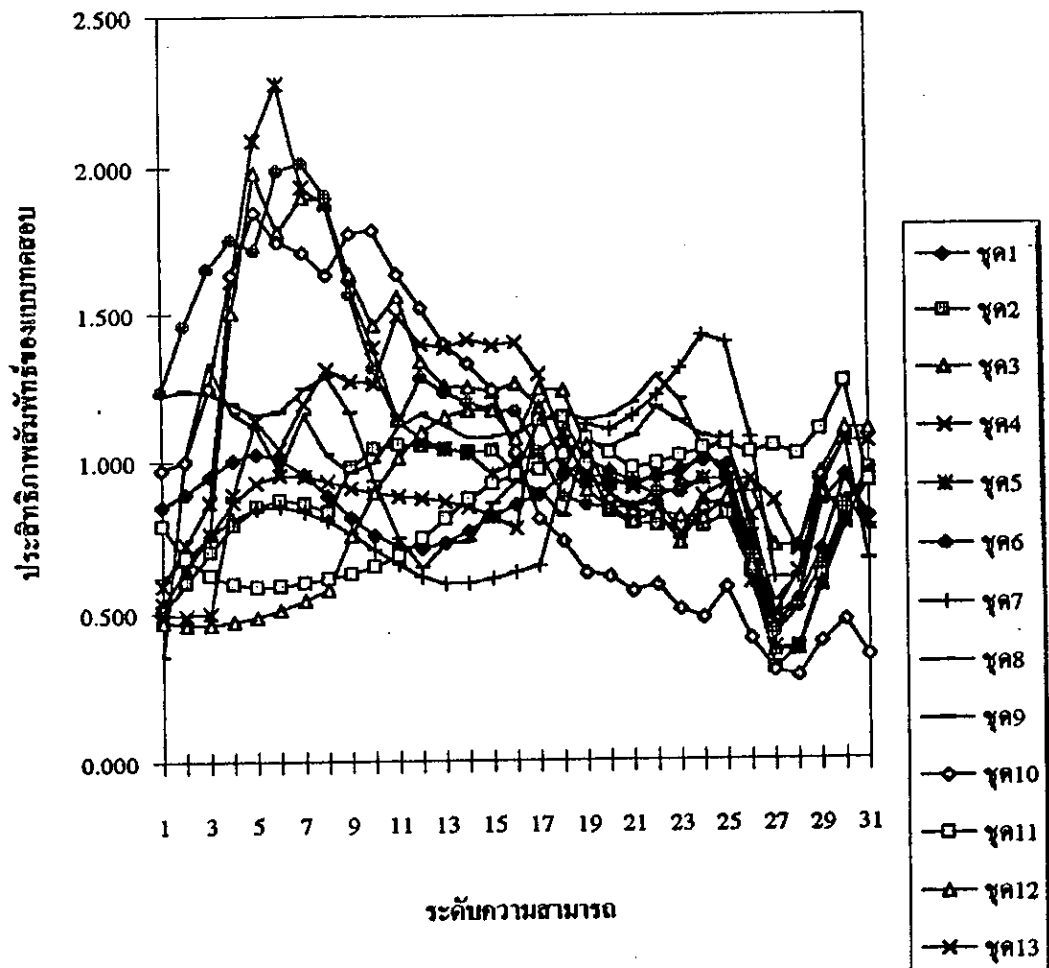


ภาพประกอบ 79 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดฮิวอี้ของเบส์ เมื่อข้อสอบมีขนาด 80 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

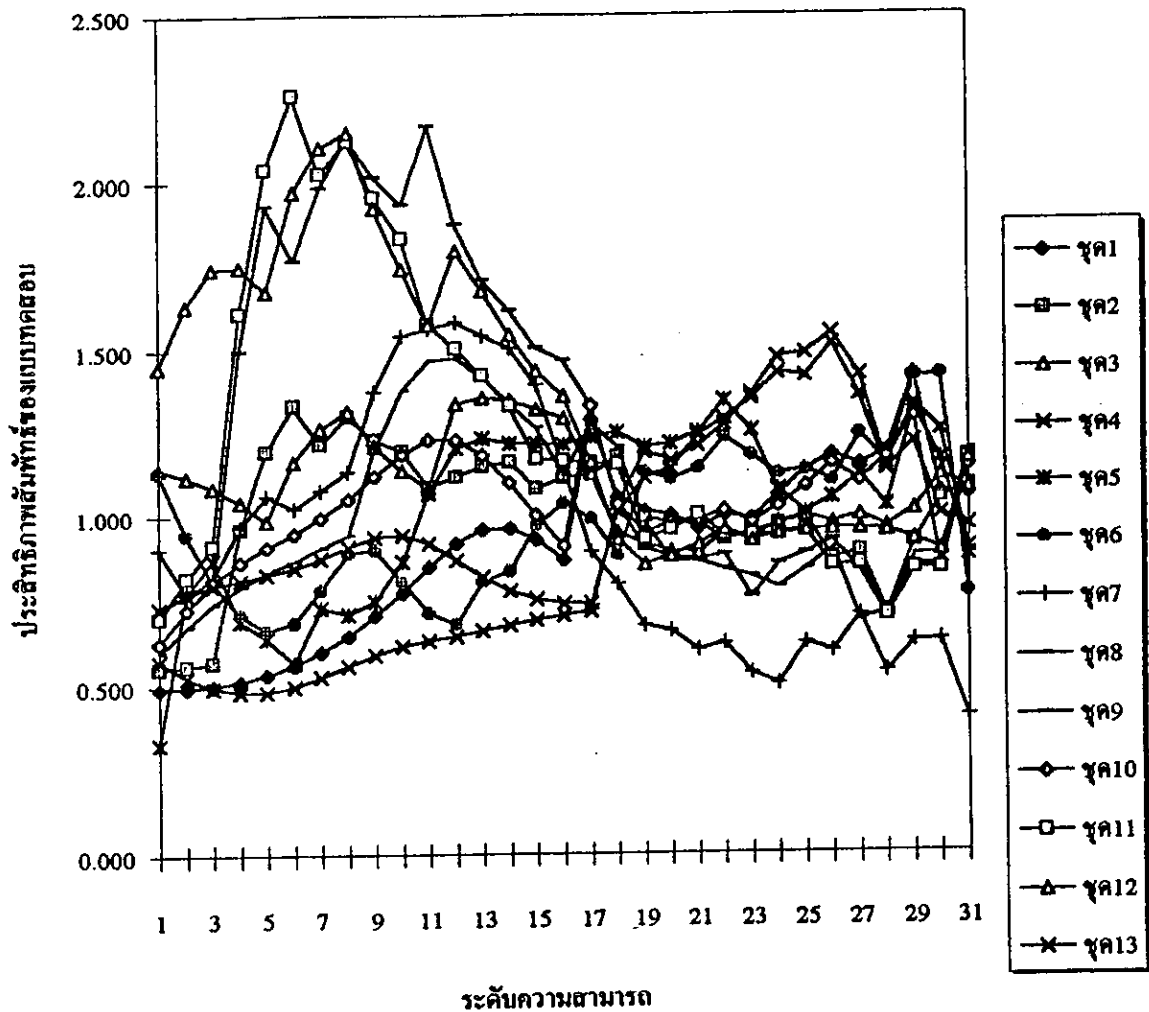


ภาพประกอบ 79 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 500/300 คน
 ชุด 2 จำนวน 600/300 คน
 ชุด 3 จำนวน 700/300 คน
 ชุด 4 จำนวน 800/300 คน
 ชุด 5 จำนวน 900/300 คน
 ชุด 6 จำนวน 1,000/300 คน
 ชุด 7 จำนวน 1,500/300 คน

ชุด 8 จำนวน 2,000/300 คน
 ชุด 9 จำนวน 2,500/300 คน
 ชุด 10 จำนวน 3,000/300 คน
 ชุด 11 จำนวน 500/400 คน
 ชุด 12 จำนวน 600/400 คน
 ชุด 13 จำนวน 700/400 คน

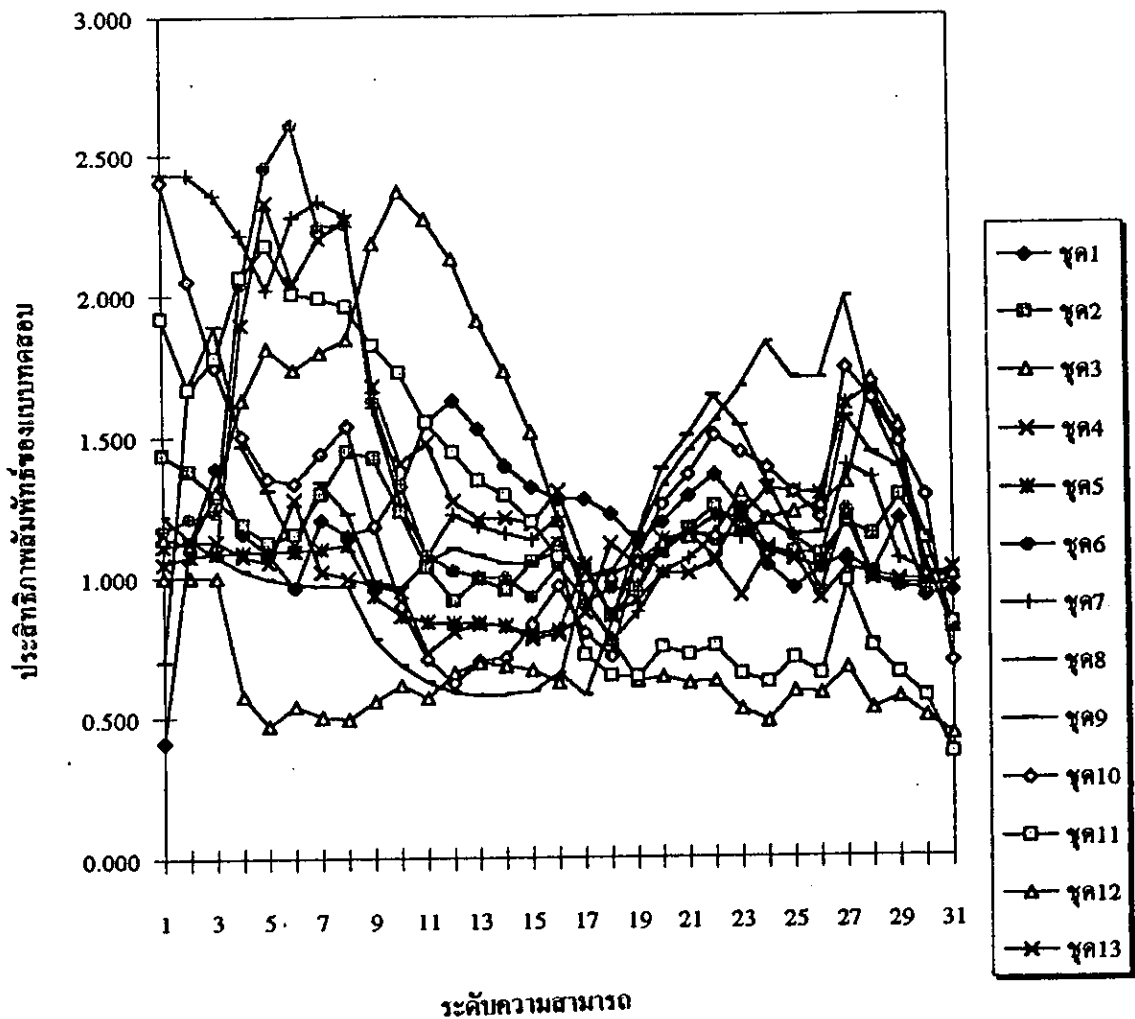


ภาพประกอบ 79 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

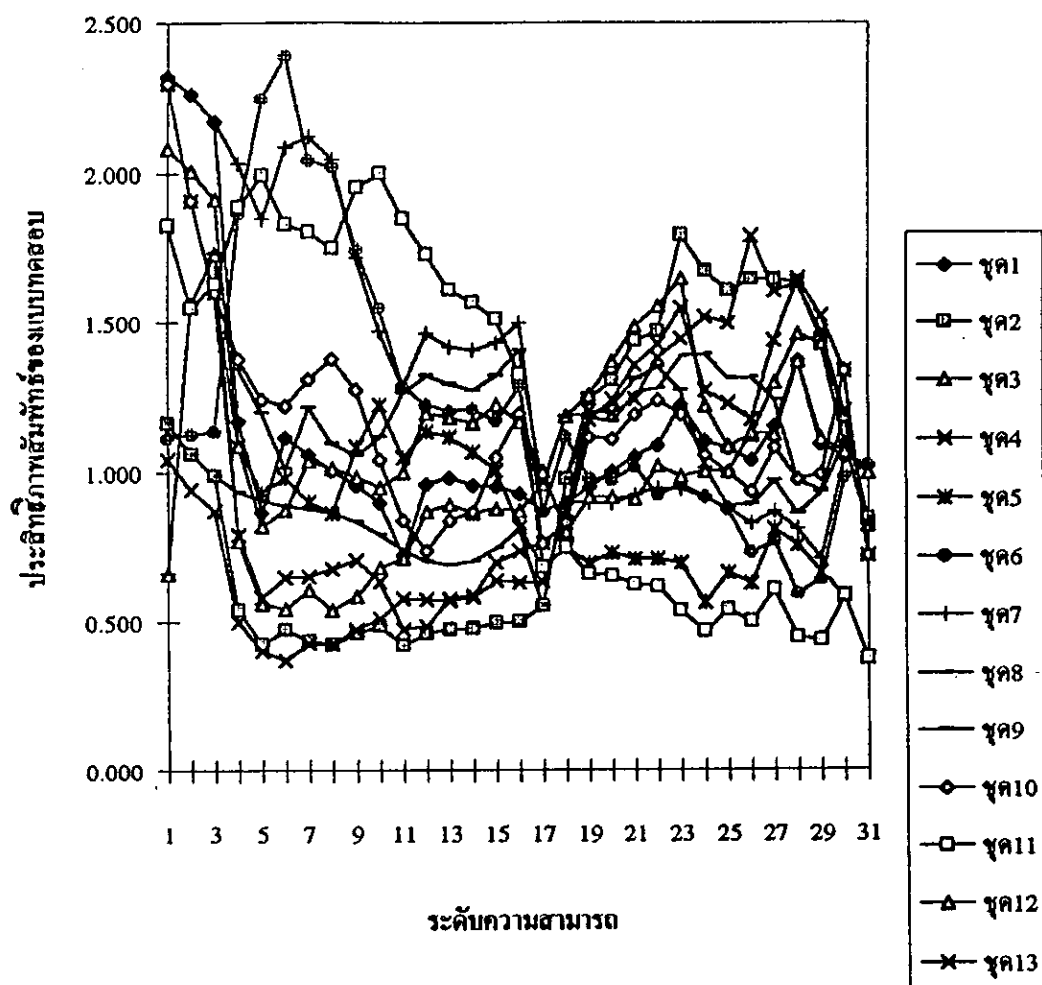


ภาพประกอบ 79 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

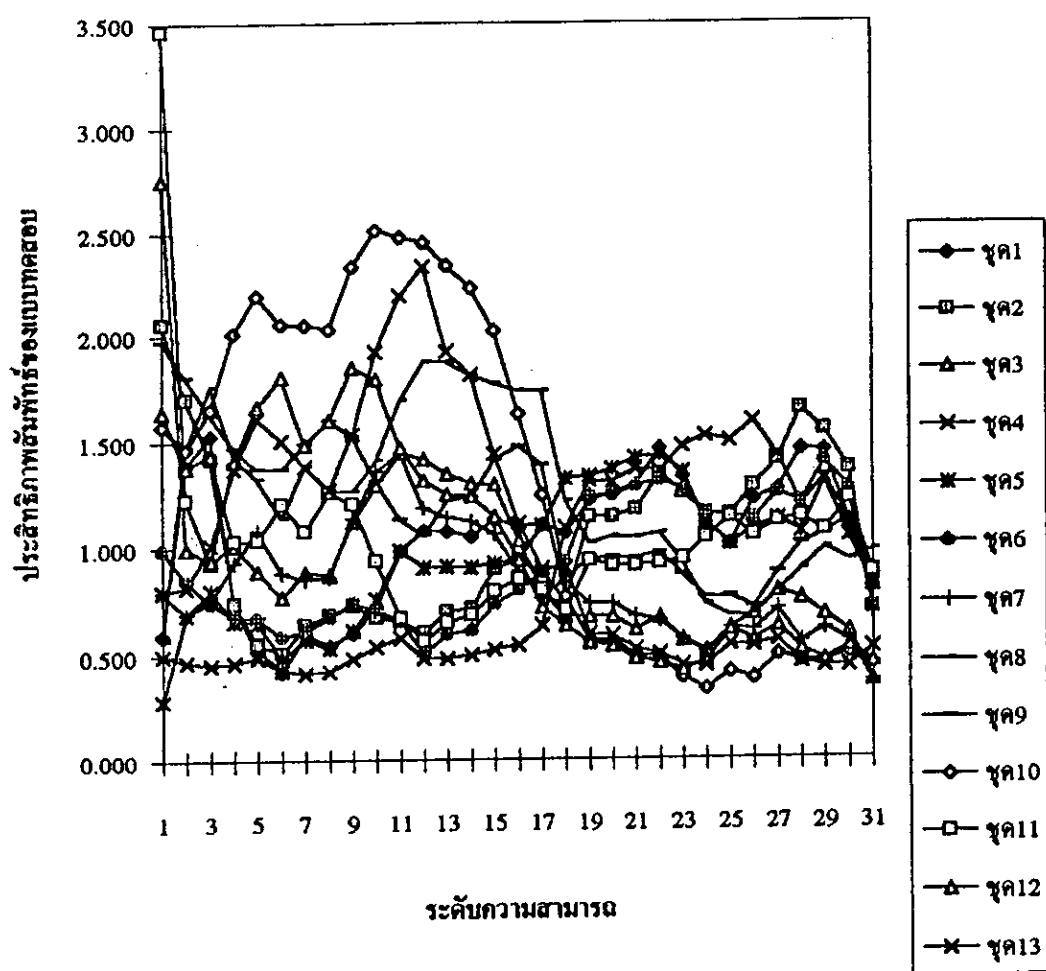


ภาพประกอบ 79 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน

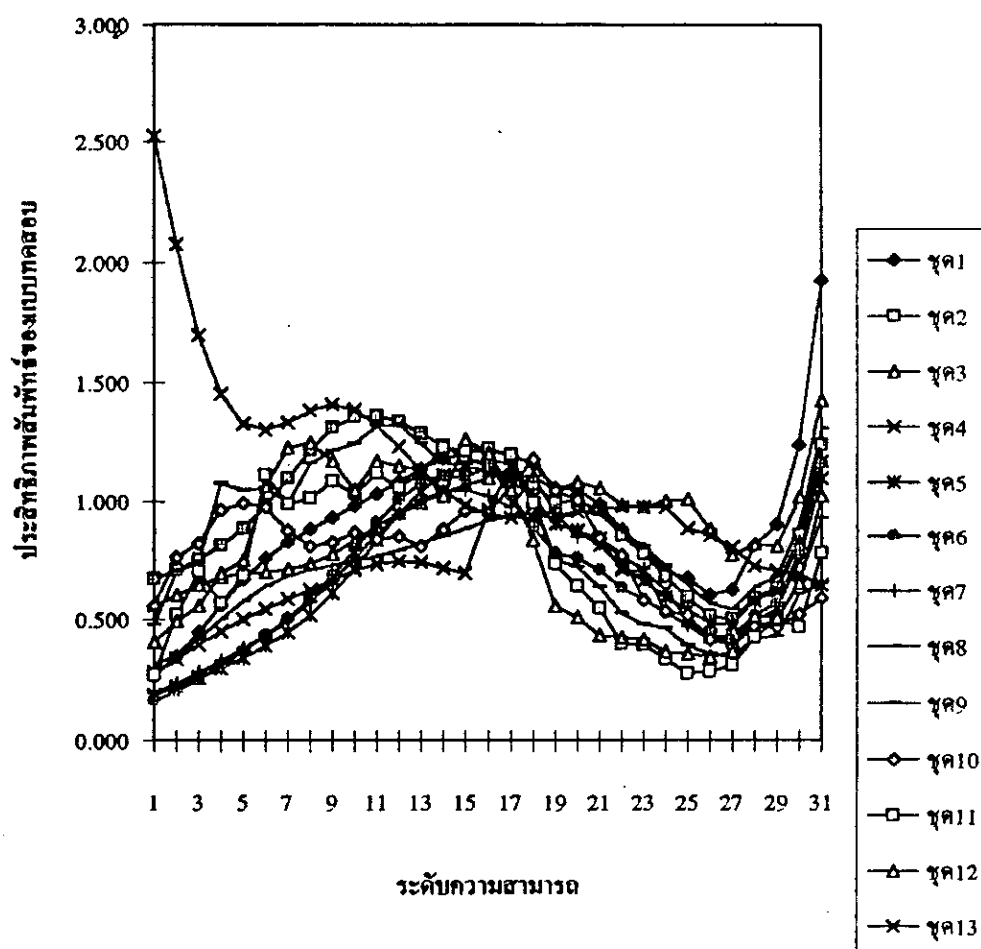


ภาพประกอบ 79 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

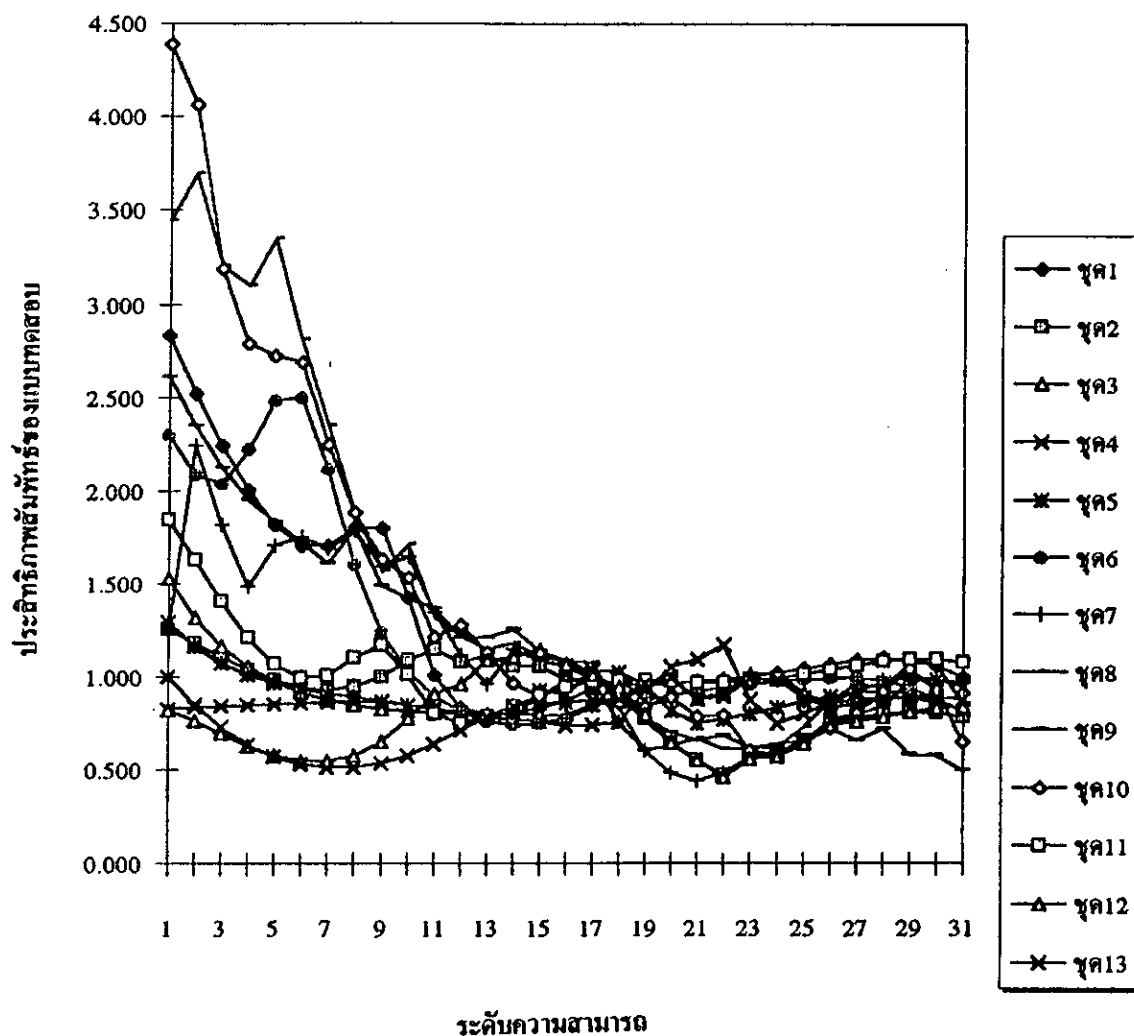


ภาพประกอบ 80 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อข้อสอบมีขนาด 90 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ๗ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 300/200 คน
ชด 2 จำนวน 400/200 คน
ชด 3 จำนวน 500/200 คน
ชด 4 จำนวน 600/200 คน
ชด 5 จำนวน 700/200 คน
ชด 6 จำนวน 800/200 คน
ชด 7 จำนวน 900/200 คน

ชด 8 จำนวน 1,000/200 คน
ชด 9 จำนวน 1,500/200 คน
ชด 10 จำนวน 2,000/200 คน
ชด 11 จำนวน 2,500/200 คน
ชด 12 จำนวน 3,000/200 คน
ชด 13 จำนวน 400/300 คน

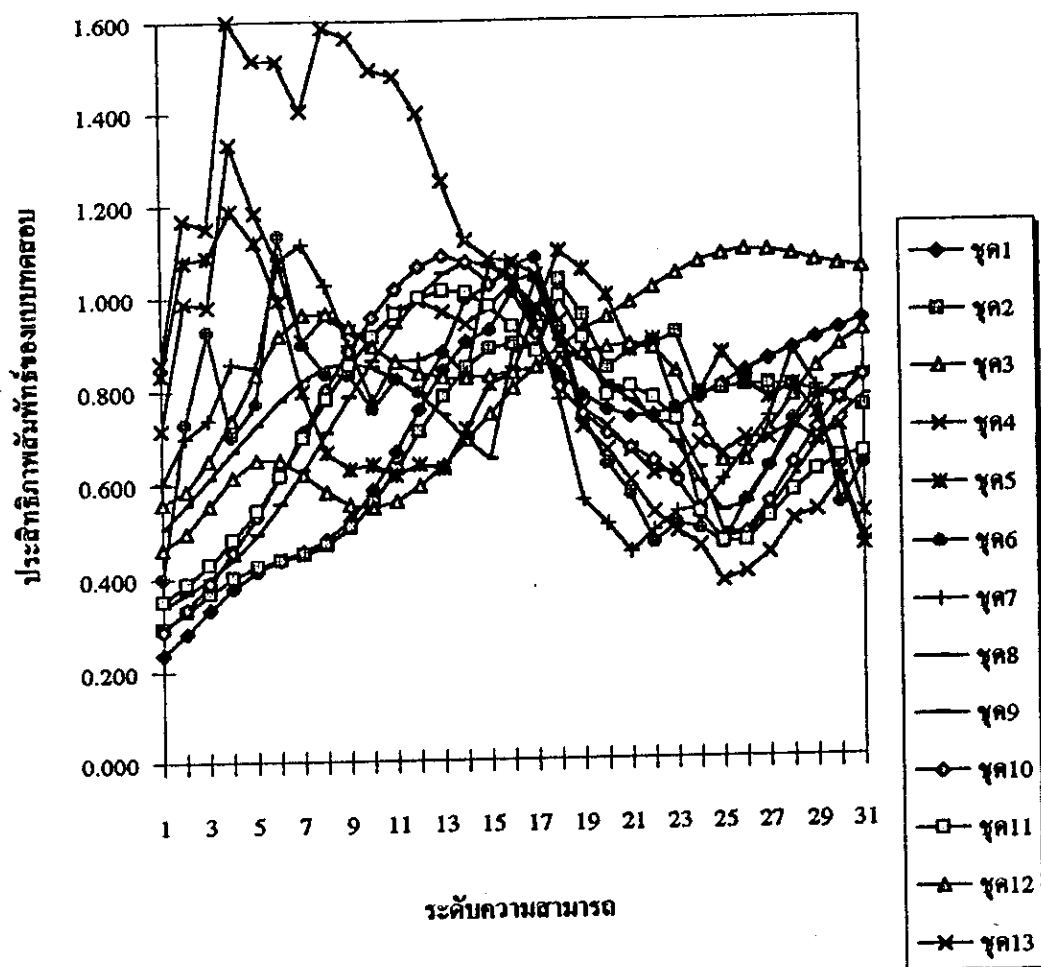


ภาพประกอบ 80 (ต่อ)

หมายเหตุ

ชุค 1 จำนวน 500/300 คน
 ชุค 2 จำนวน 600/300 คน
 ชุค 3 จำนวน 700/300 คน
 ชุค 4 จำนวน 800/300 คน
 ชุค 5 จำนวน 900/300 คน
 ชุค 6 จำนวน 1,000/300 คน
 ชุค 7 จำนวน 1,500/300 คน

ชุค 8 จำนวน 2,000/300 คน
 ชุค 9 จำนวน 2,500/300 คน
 ชุค 10 จำนวน 3,000/300 คน
 ชุค 11 จำนวน 500/400 คน
 ชุค 12 จำนวน 800/400 คน
 ชุค 13 จำนวน 700/400 คน

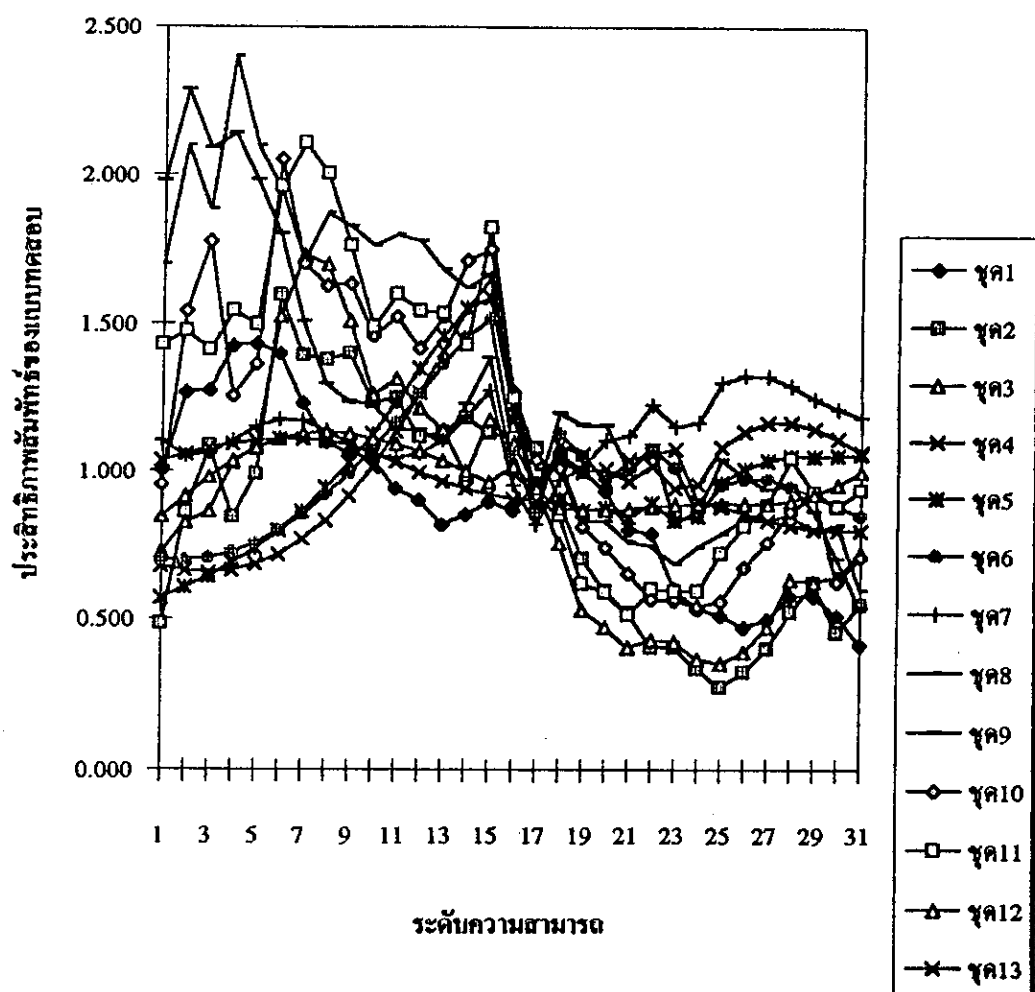


ภาพประกอบ 80 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
ชุด 2 จำนวน 900/400 คน
ชุด 3 จำนวน 1,000/400 คน
ชุด 4 จำนวน 1,500/400 คน
ชุด 5 จำนวน 2,000/400 คน
ชุด 6 จำนวน 2,500/400 คน
ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
ชุด 9 จำนวน 700/500 คน
ชุด 10 จำนวน 800/500 คน
ชุด 11 จำนวน 900/500 คน
ชุด 12 จำนวน 1,000/500 คน
ชุด 13 จำนวน 1,500/500 คน

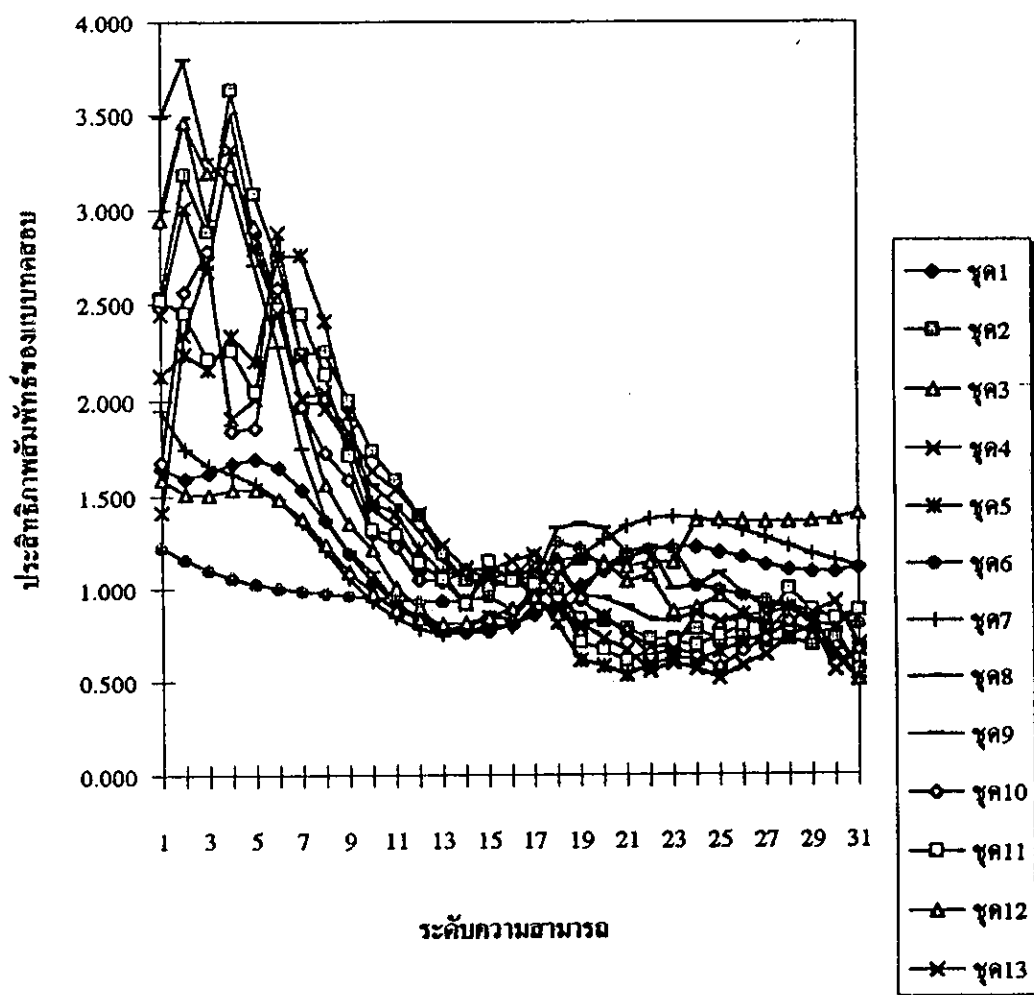


ภาพประกอบ 80 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

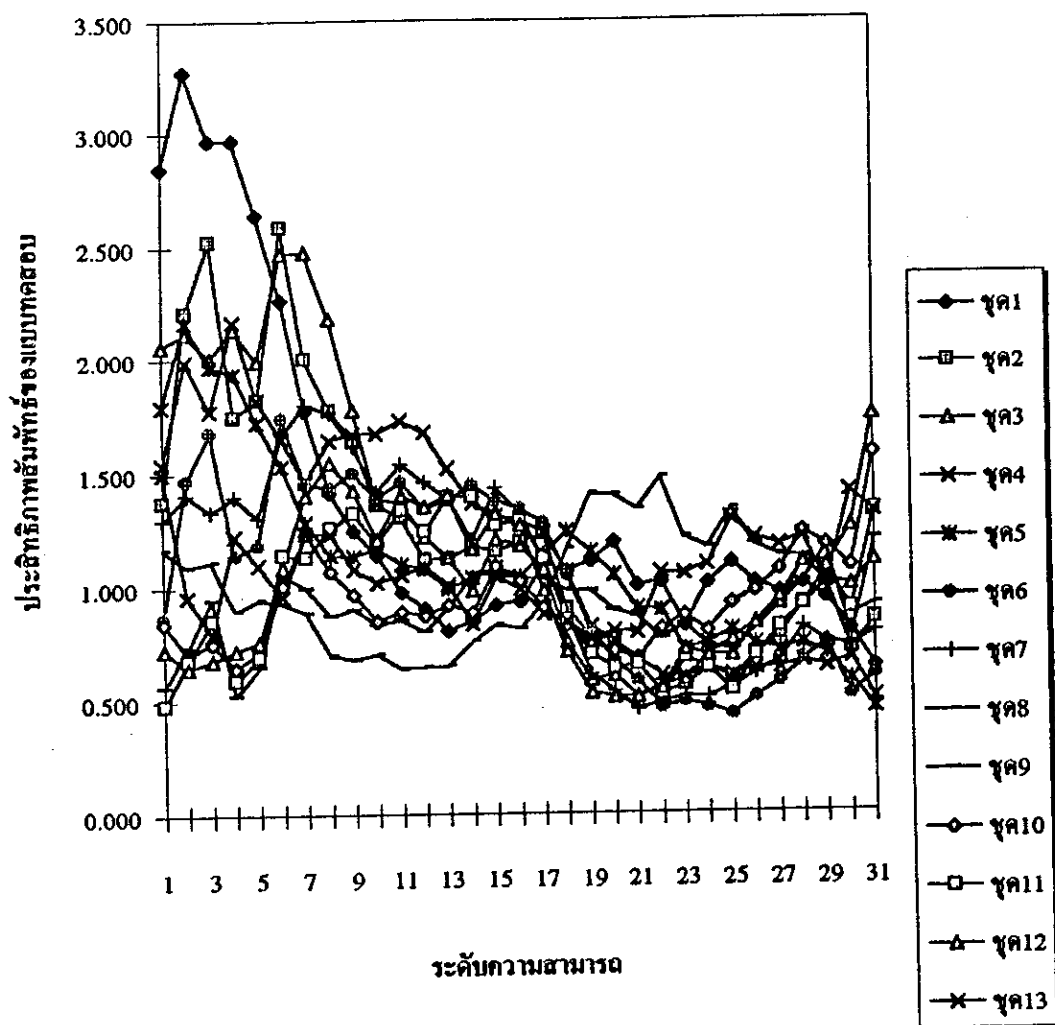


ภาพประกอบ 80 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 1,000/700 คน
 ชด 2 จำนวน 1,500/700 คน
 ชด 3 จำนวน 2,000/700 คน
 ชด 4 จำนวน 2,500/700 คน
 ชด 5 จำนวน 3,000/700 คน
 ชด 6 จำนวน 900/800 คน
 ชด 7 จำนวน 1,000/800 คน

ชด 8 จำนวน 1,500/800 คน
 ชด 9 จำนวน 2,000/800 คน
 ชด 10 จำนวน 2,500/800 คน
 ชด 11 จำนวน 3,000/800 คน
 ชด 12 จำนวน 1,000/900 คน
 ชด 13 จำนวน 1,500/900 คน

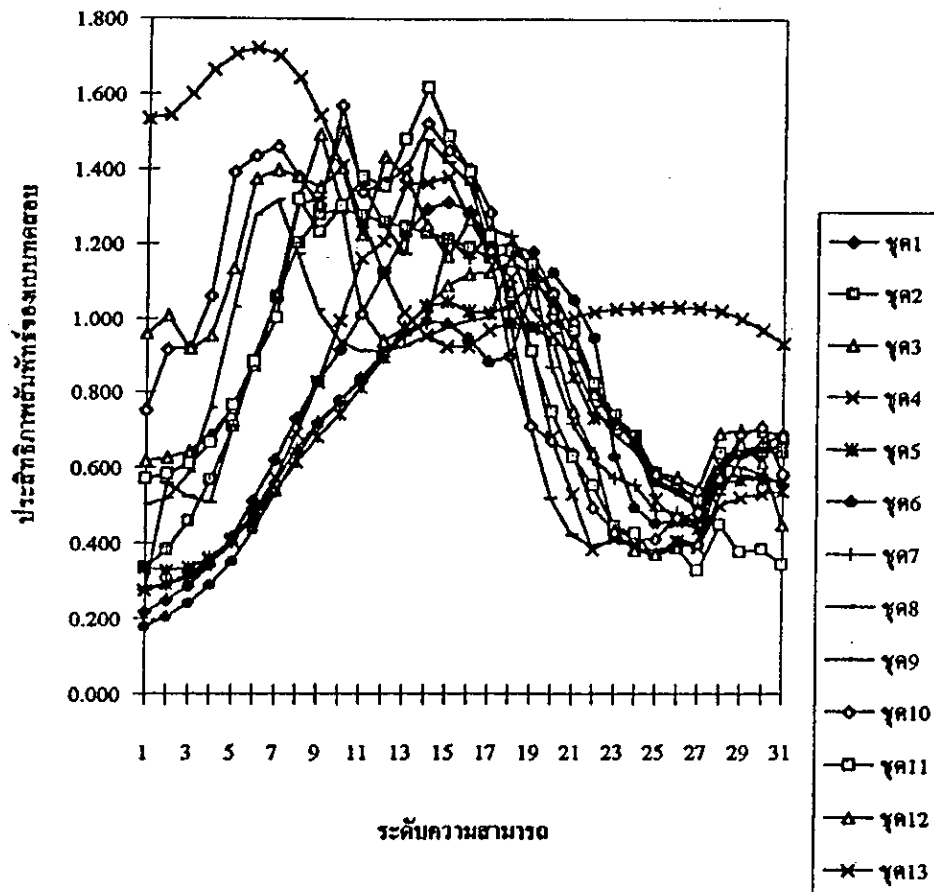


ภาพประกอบ 80 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

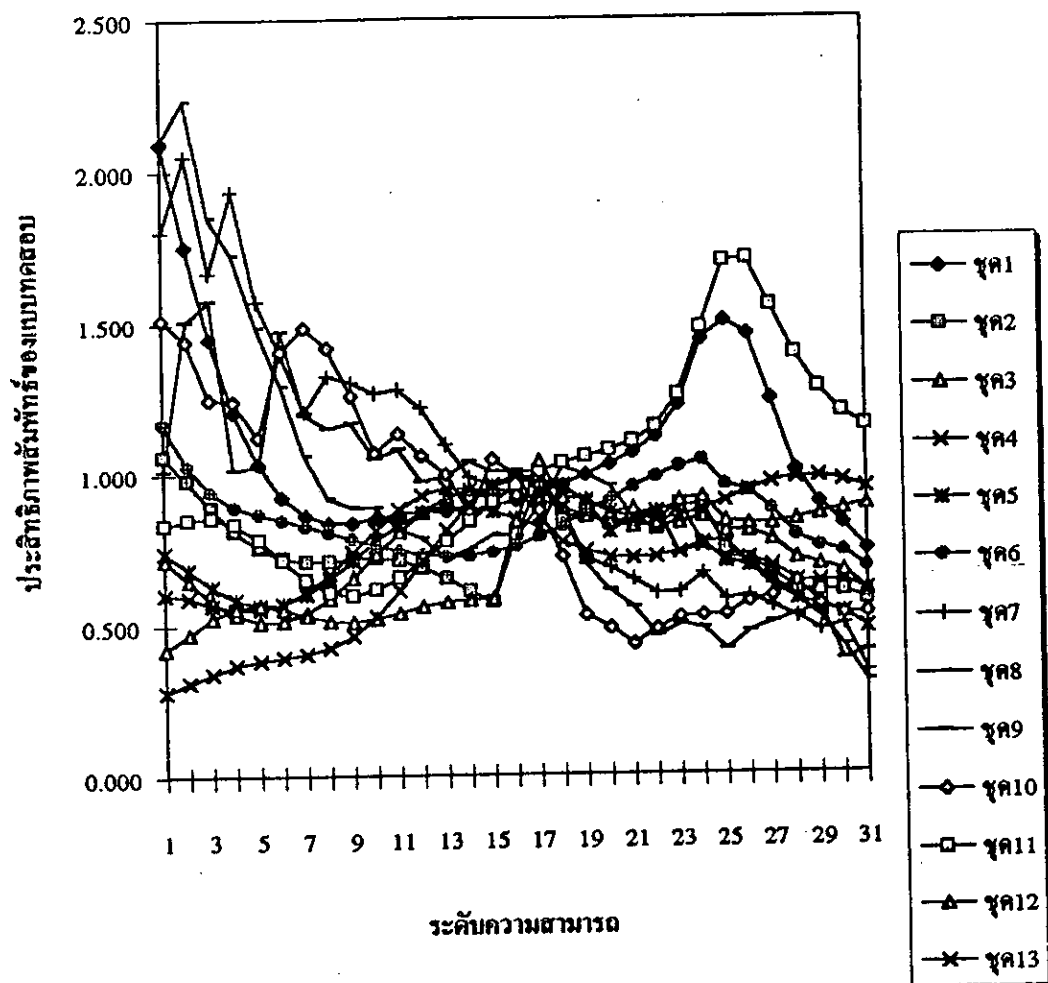


ภาพประกอบ 81 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โคสไวซีของเบส์ เมื่อข้อสอบมีขนาด 100 ข้อ และกลุ่มตัวอย่างมีขนาดต่างกัน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 300/200 คน
ชุด 2 จำนวน 400/200 คน
ชุด 3 จำนวน 500/200 คน
ชุด 4 จำนวน 600/200 คน
ชุด 5 จำนวน 700/200 คน
ชุด 6 จำนวน 800/200 คน
ชุด 7 จำนวน 900/200 คน

ชุด 8 จำนวน 1,000/200 คน
ชุด 9 จำนวน 1,500/200 คน
ชุด 10 จำนวน 2,000/200 คน
ชุด 11 จำนวน 2,500/200 คน
ชุด 12 จำนวน 3,000/200 คน
ชุด 13 จำนวน 400/300 คน

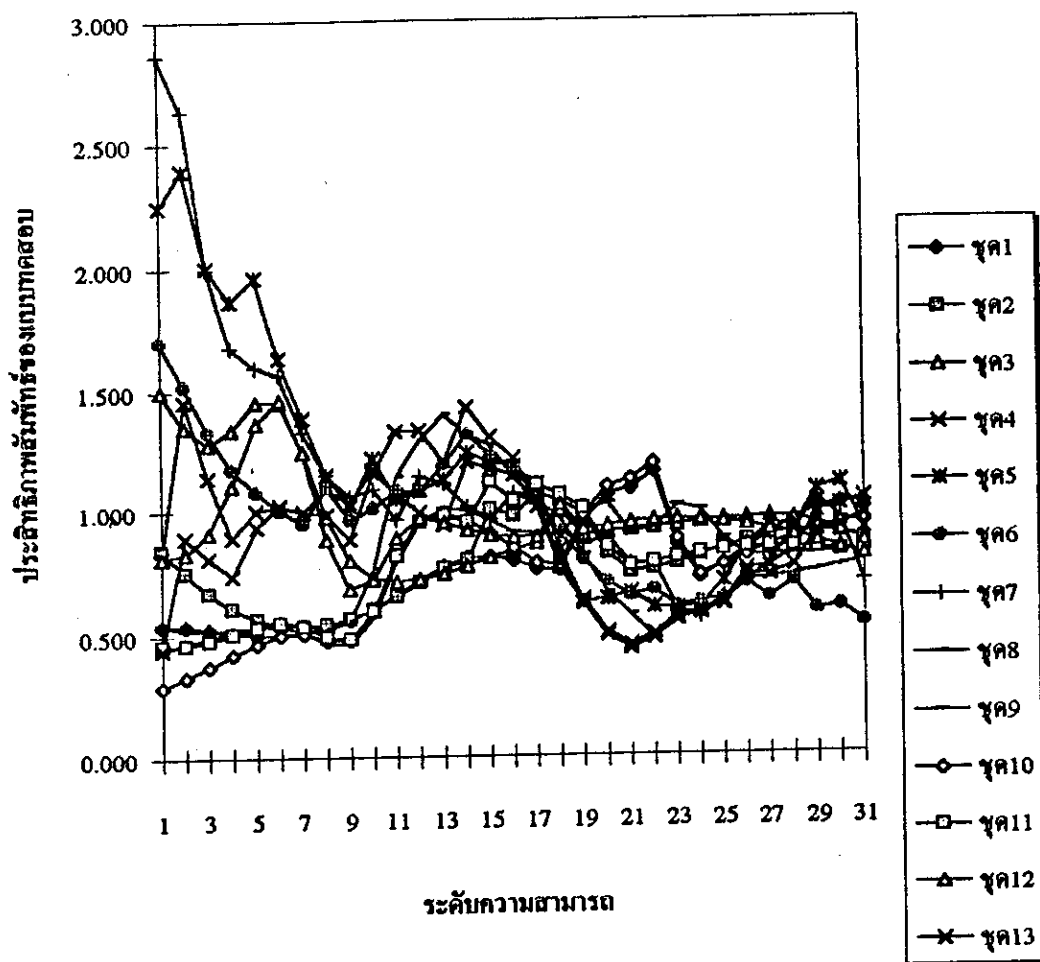


ภาพประกอบ 81 (ต่อ)

หมายเหตุ

ชด 1 จำนวน 500/300 คน
 ชด 2 จำนวน 600/300 คน
 ชด 3 จำนวน 700/300 คน
 ชด 4 จำนวน 800/300 คน
 ชด 5 จำนวน 900/300 คน
 ชด 6 จำนวน 1,000/300 คน
 ชด 7 จำนวน 1,500/300 คน

ชด 8 จำนวน 2,000/300 คน
 ชด 9 จำนวน 2,500/300 คน
 ชด 10 จำนวน 3,000/300 คน
 ชด 11 จำนวน 500/400 คน
 ชด 12 จำนวน 600/400 คน
 ชด 13 จำนวน 700/400 คน

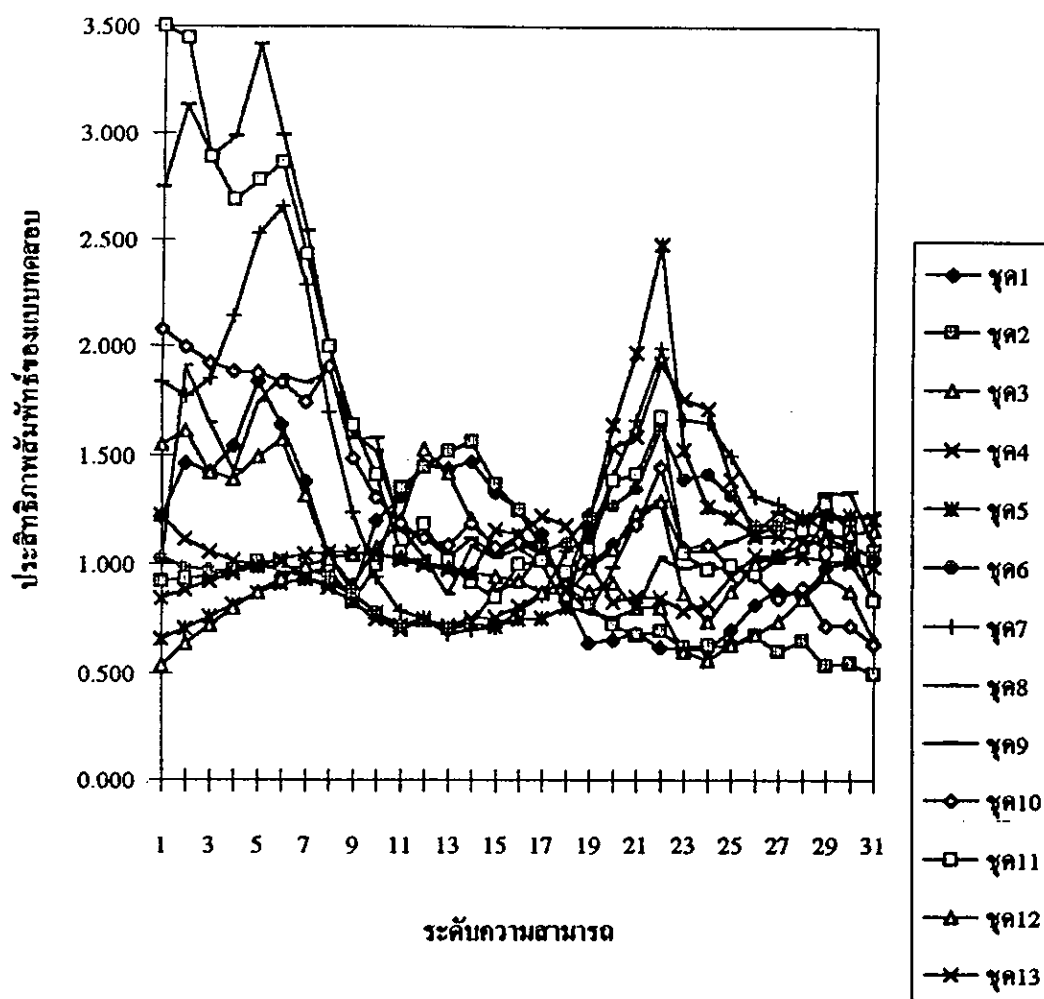


ภาพประกอบ 81 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 800/400 คน
 ชุด 2 จำนวน 900/400 คน
 ชุด 3 จำนวน 1,000/400 คน
 ชุด 4 จำนวน 1,500/400 คน
 ชุด 5 จำนวน 2,000/400 คน
 ชุด 6 จำนวน 2,500/400 คน
 ชุด 7 จำนวน 3,000/400 คน

ชุด 8 จำนวน 600/500 คน
 ชุด 9 จำนวน 700/500 คน
 ชุด 10 จำนวน 800/500 คน
 ชุด 11 จำนวน 900/500 คน
 ชุด 12 จำนวน 1,000/500 คน
 ชุด 13 จำนวน 1,500/500 คน

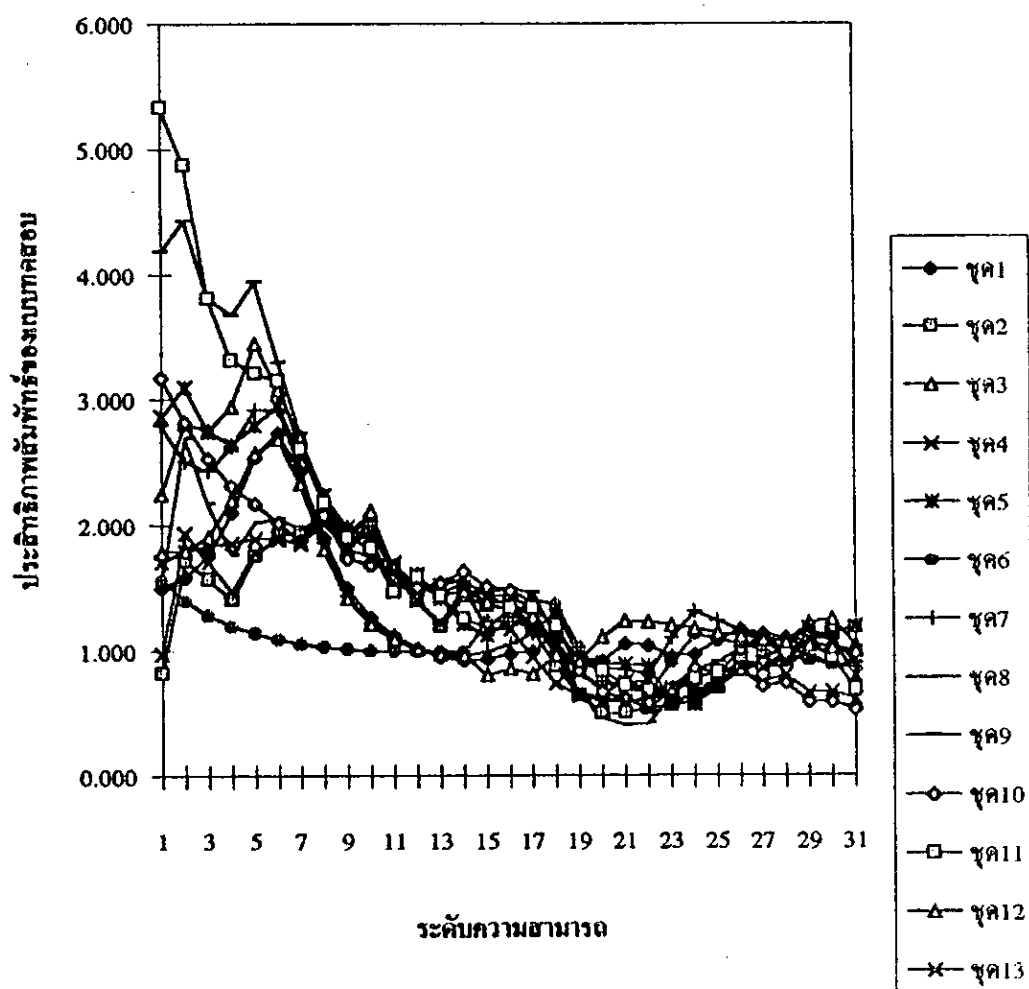


ภาพประกอบ 81 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/500 คน
 ชุด 2 จำนวน 2,500/500 คน
 ชุด 3 จำนวน 3,000/500 คน
 ชุด 4 จำนวน 700/600 คน
 ชุด 5 จำนวน 800/600 คน
 ชุด 6 จำนวน 900/600 คน
 ชุด 7 จำนวน 1,000/600 คน

ชุด 8 จำนวน 1,500/600 คน
 ชุด 9 จำนวน 2,000/600 คน
 ชุด 10 จำนวน 2,500/600 คน
 ชุด 11 จำนวน 3,000/600 คน
 ชุด 12 จำนวน 800/700 คน
 ชุด 13 จำนวน 900/700 คน

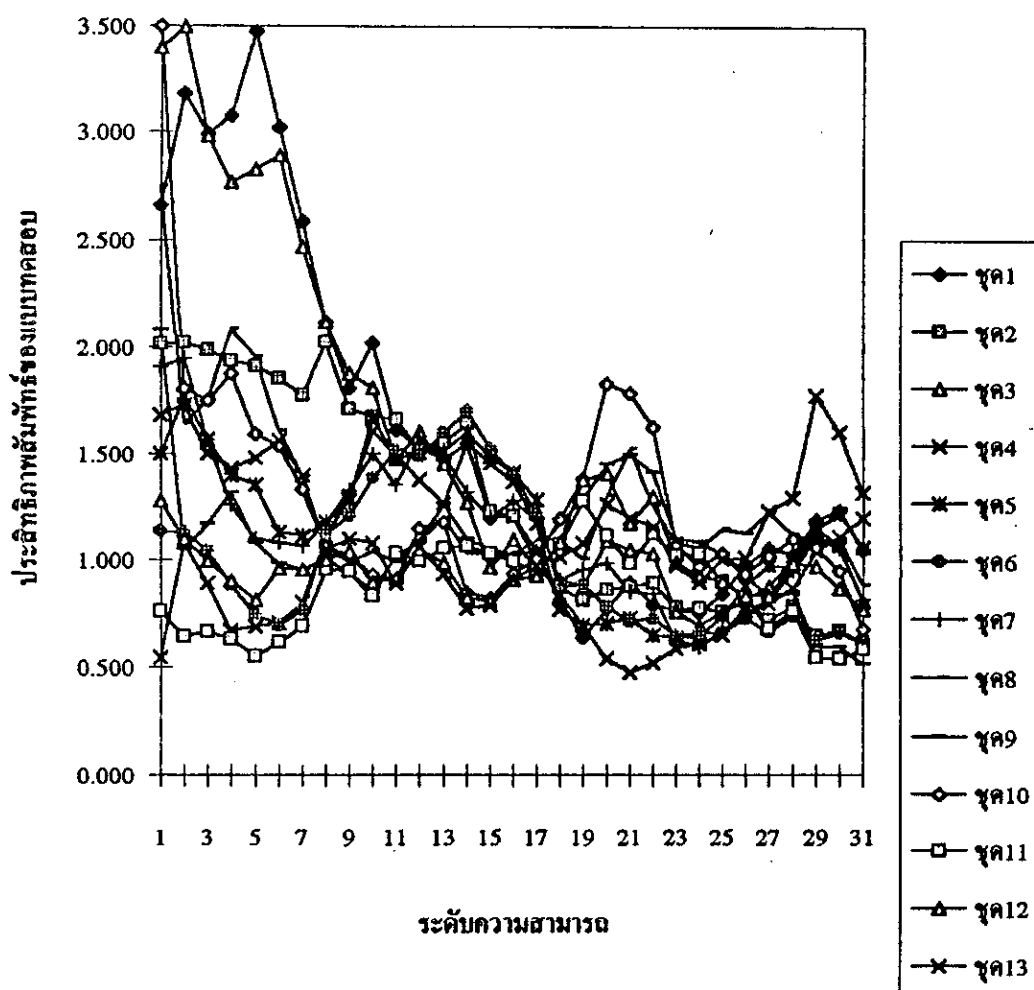


ภาพประกอบ 81 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 1,000/700 คน
 ชุด 2 จำนวน 1,500/700 คน
 ชุด 3 จำนวน 2,000/700 คน
 ชุด 4 จำนวน 2,500/700 คน
 ชุด 5 จำนวน 3,000/700 คน
 ชุด 6 จำนวน 900/800 คน
 ชุด 7 จำนวน 1,000/800 คน

ชุด 8 จำนวน 1,500/800 คน
 ชุด 9 จำนวน 2,000/800 คน
 ชุด 10 จำนวน 2,500/800 คน
 ชุด 11 จำนวน 3,000/800 คน
 ชุด 12 จำนวน 1,000/900 คน
 ชุด 13 จำนวน 1,500/900 คน



ภาพประกอบ 81 (ต่อ)

หมายเหตุ

ชุด 1 จำนวน 2,000/900 คน
 ชุด 2 จำนวน 2,500/900 คน
 ชุด 3 จำนวน 3,000/900 คน
 ชุด 4 จำนวน 1,500/1,000 คน
 ชุด 5 จำนวน 2,000/1,000 คน
 ชุด 6 จำนวน 2,500/1,000 คน
 ชุด 7 จำนวน 3,000/1,000 คน

ชุด 8 จำนวน 2,000/1,500 คน
 ชุด 9 จำนวน 2,500/1,500 คน
 ชุด 10 จำนวน 3,000/1,500 คน
 ชุด 11 จำนวน 2,500/2,000 คน
 ชุด 12 จำนวน 3,000/2,000 คน
 ชุด 13 จำนวน 3,000/2,500 คน

จากภาพประกอบ 76 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 50 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ ระหว่างกลุ่มตัวอย่าง 400 คน กับ 300 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(400, 300) = 5.386 ; \phi = -3.000\}$

จากภาพประกอบ 77 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 60 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 2,000 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(2,500, 2,000) = 7.313 ; \phi = -3.000\}$

จากภาพประกอบ 78 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 70 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ ระหว่างกลุ่มตัวอย่าง 400 คน กับ 200 คน มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(400, 200) = 4.348 ; \phi = 3.000\}$

จากภาพประกอบ 79 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 80 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 2,000 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(2,500, 2,000) = 3.460 ; \phi = -3.000\}$

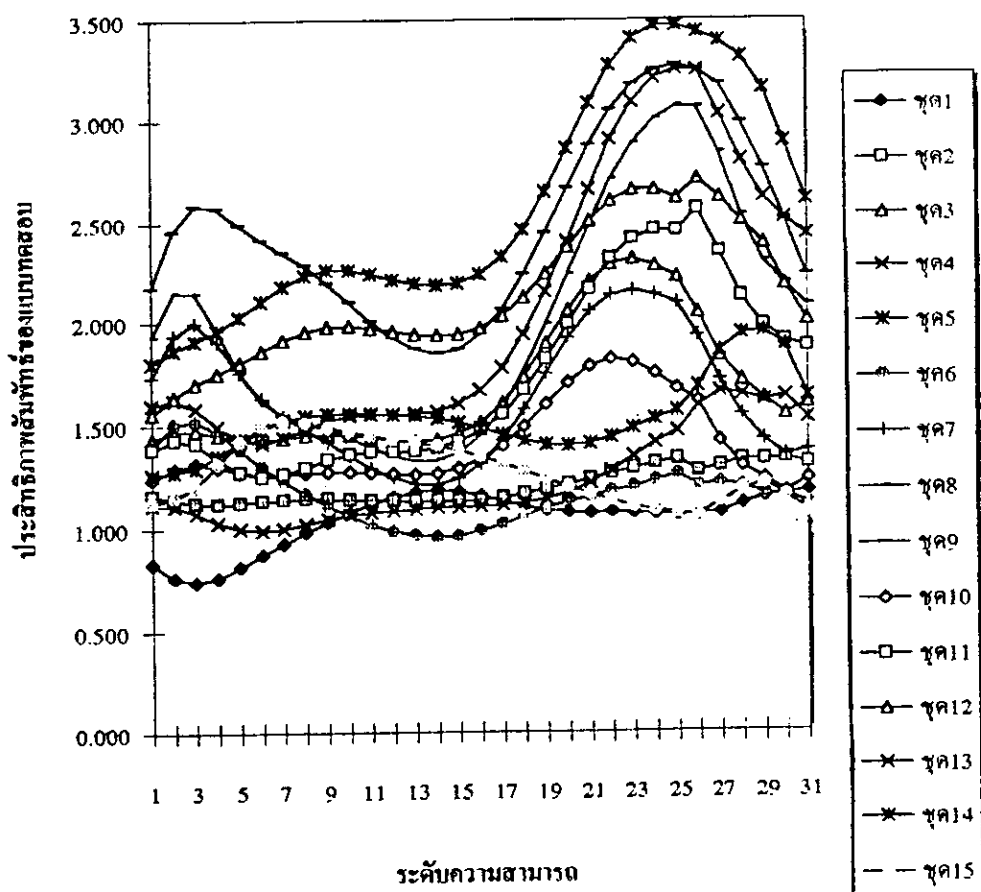
จากภาพประกอบ 80 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 90 ข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ ระหว่างกลุ่มตัวอย่าง 2,000 คน กับ 800 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(2,000, 300) = 3.792 ; \phi = -2.800\}$

จากภาพประกอบ 81 พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 100 ข้อ ที่เป็นผลจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ ระหว่างกลุ่มตัวอย่าง 3,000 คน กับ 800 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(3,000, 800) = 5.329 ; \phi = -3.000\}$

ดังนั้น จากภาพประกอบ 76 - 81 แสดงว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ 60 ข้อ ที่เป็นมาผลจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ ระหว่างกลุ่มตัวอย่าง 2,500 คน กับ 2,000 คน มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(2,500, 2,000) = 7.313 ; \phi = -3.000\}$

8. ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ

8.1 ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 82 - 94

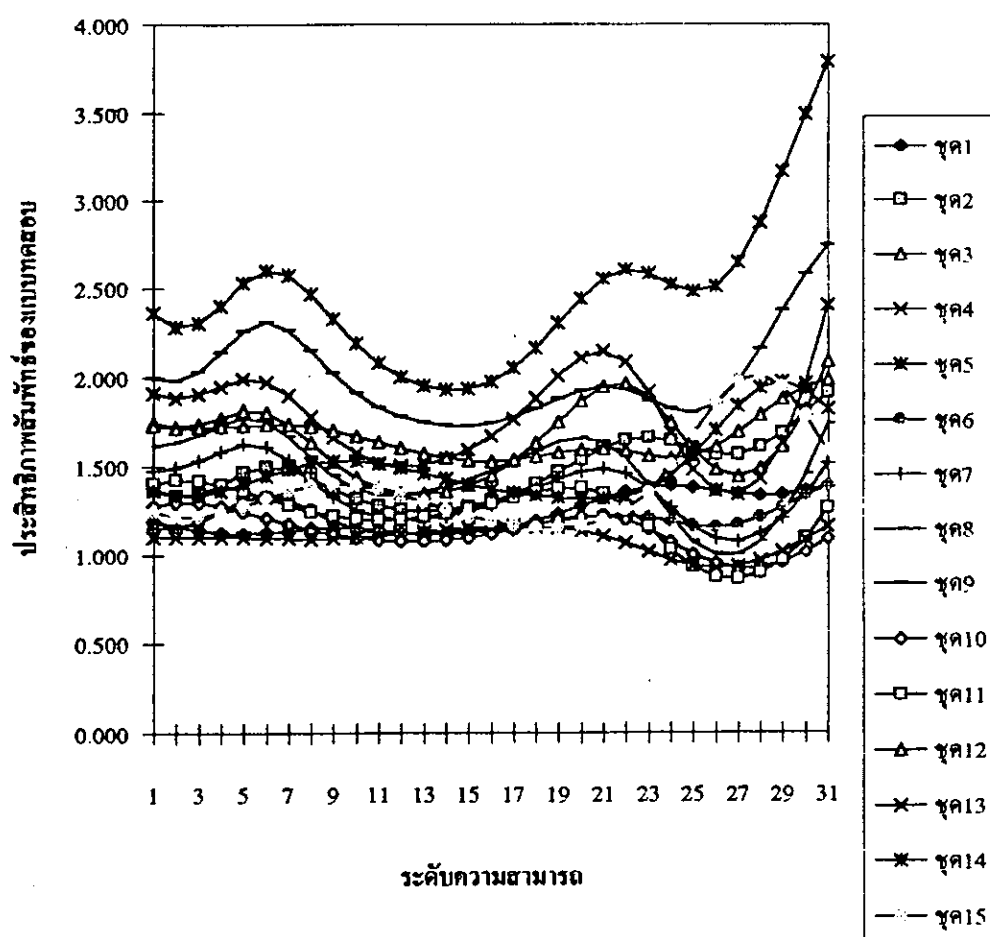


ภาพประกอบ 82 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 200 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
ชด 2 จำนวน 70/50 ข้อ
ชด 3 จำนวน 80/50 ข้อ
ชด 4 จำนวน 90/50 ข้อ
ชด 5 จำนวน 100/50 ข้อ
ชด 6 จำนวน 70/60 ข้อ
ชด 7 จำนวน 80/60 ข้อ
ชด 8 จำนวน 90/60 ข้อ

ชด 9 จำนวน 100/60 ข้อ
ชด 10 จำนวน 80/70 ข้อ
ชด 11 จำนวน 90/70 ข้อ
ชด 12 จำนวน 100/70 ข้อ
ชด 13 จำนวน 90/80 ข้อ
ชด 14 จำนวน 100/80 ข้อ
ชด 15 จำนวน 100/90 ข้อ

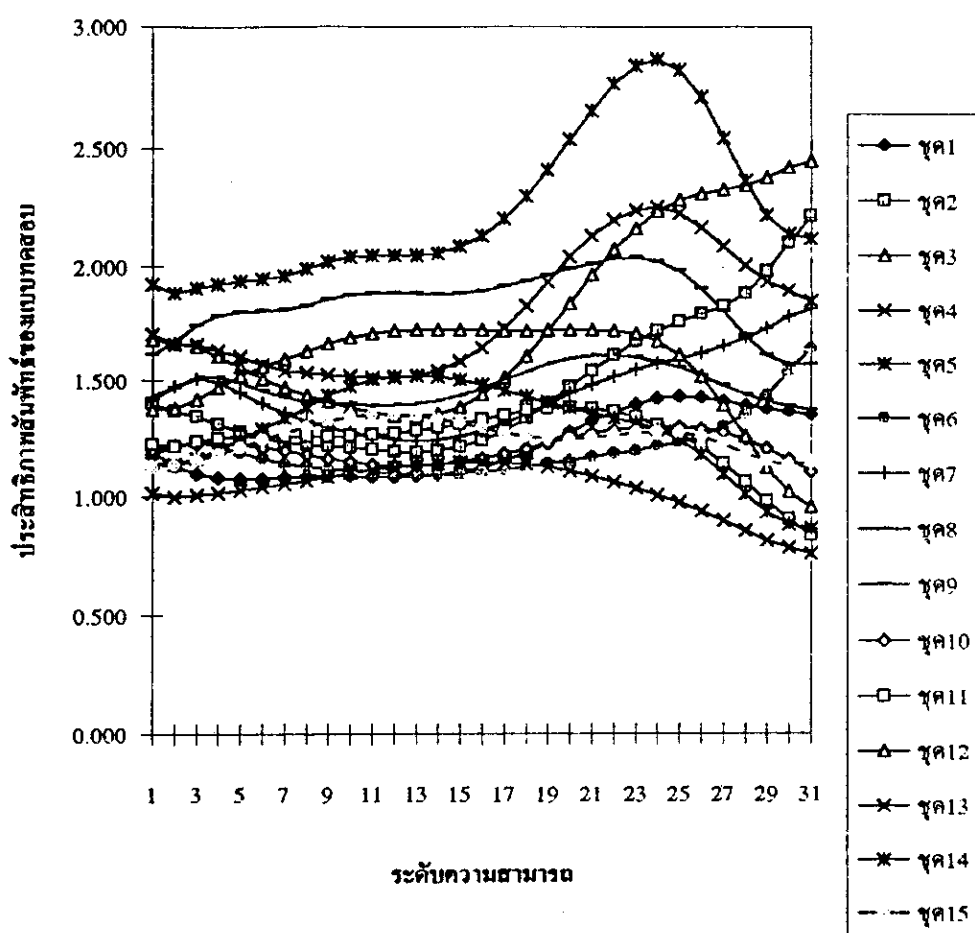


ภาพประกอบ 83 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 300 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
ชด 2 จำนวน 70/50 ข้อ
ชด 3 จำนวน 80/50 ข้อ
ชด 4 จำนวน 90/50 ข้อ
ชด 5 จำนวน 100/50 ข้อ
ชด 6 จำนวน 70/60 ข้อ
ชด 7 จำนวน 80/60 ข้อ
ชด 8 จำนวน 90/60 ข้อ

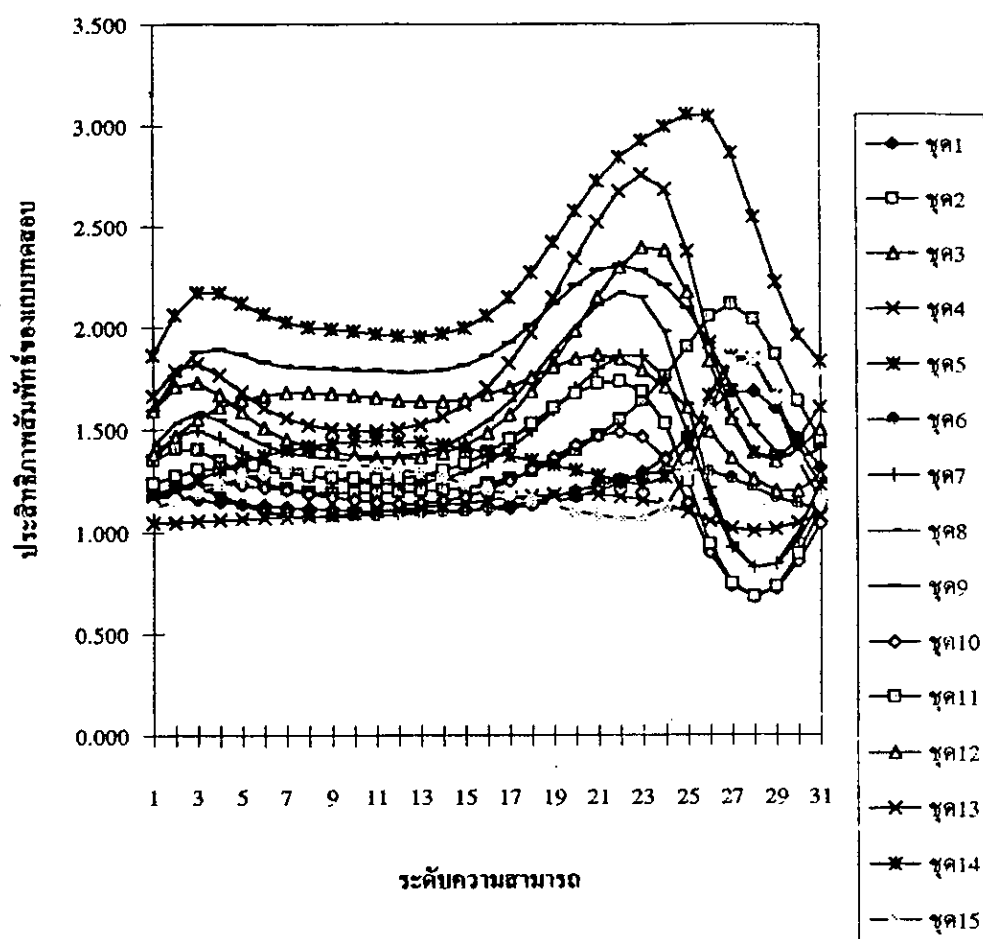
ชด 9 จำนวน 100/60 ข้อ
ชด 10 จำนวน 80/70 ข้อ
ชด 11 จำนวน 90/70 ข้อ
ชด 12 จำนวน 100/70 ข้อ
ชด 13 จำนวน 90/80 ข้อ
ชด 14 จำนวน 100/80 ข้อ
ชด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 84 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 400 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ	ชุด 9 จำนวน 100/60 ข้อ
ชุด 2 จำนวน 70/50 ข้อ	ชุด 10 จำนวน 80/70 ข้อ
ชุด 3 จำนวน 80/50 ข้อ	ชุด 11 จำนวน 90/70 ข้อ
ชุด 4 จำนวน 90/50 ข้อ	ชุด 12 จำนวน 100/70 ข้อ
ชุด 5 จำนวน 100/50 ข้อ	ชุด 13 จำนวน 90/80 ข้อ
ชุด 6 จำนวน 70/60 ข้อ	ชุด 14 จำนวน 100/80 ข้อ
ชุด 7 จำนวน 80/60 ข้อ	ชุด 15 จำนวน 100/90 ข้อ
ชุด 8 จำนวน 90/60 ข้อ	

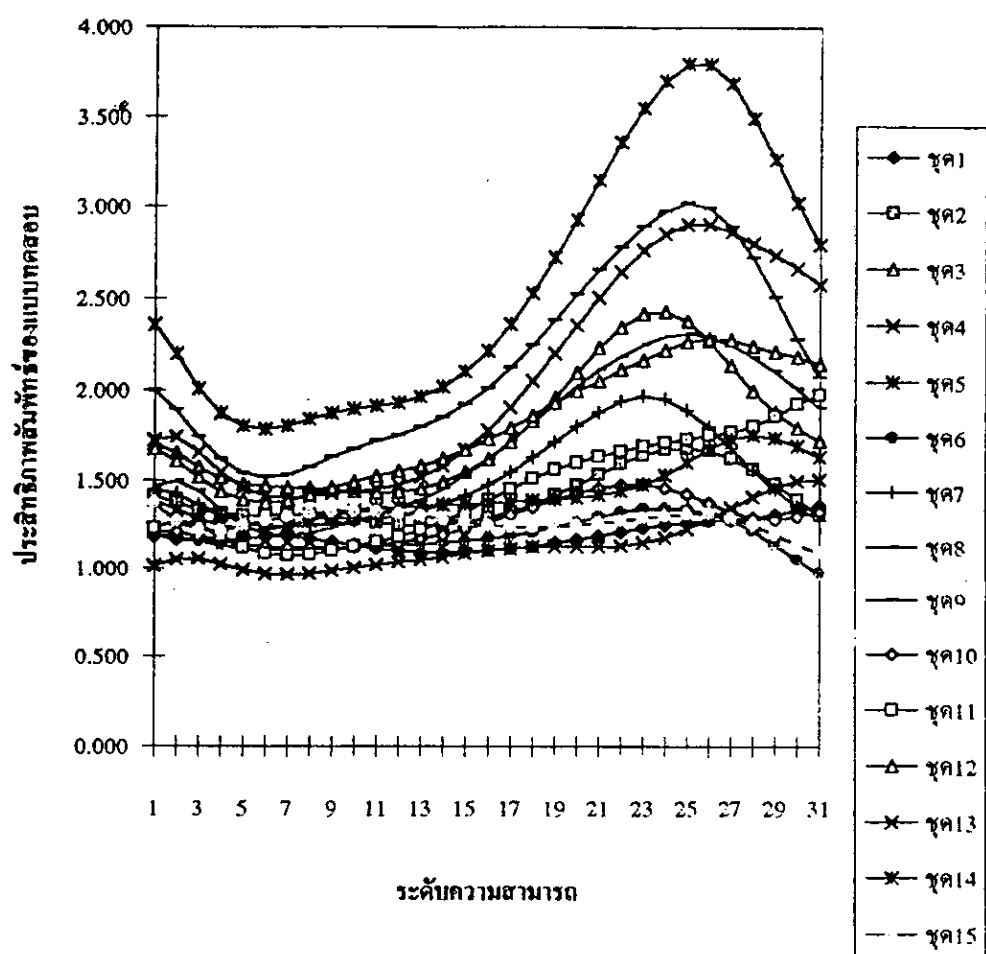


ภาพประกอบ 85 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 500 คน ณ
ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

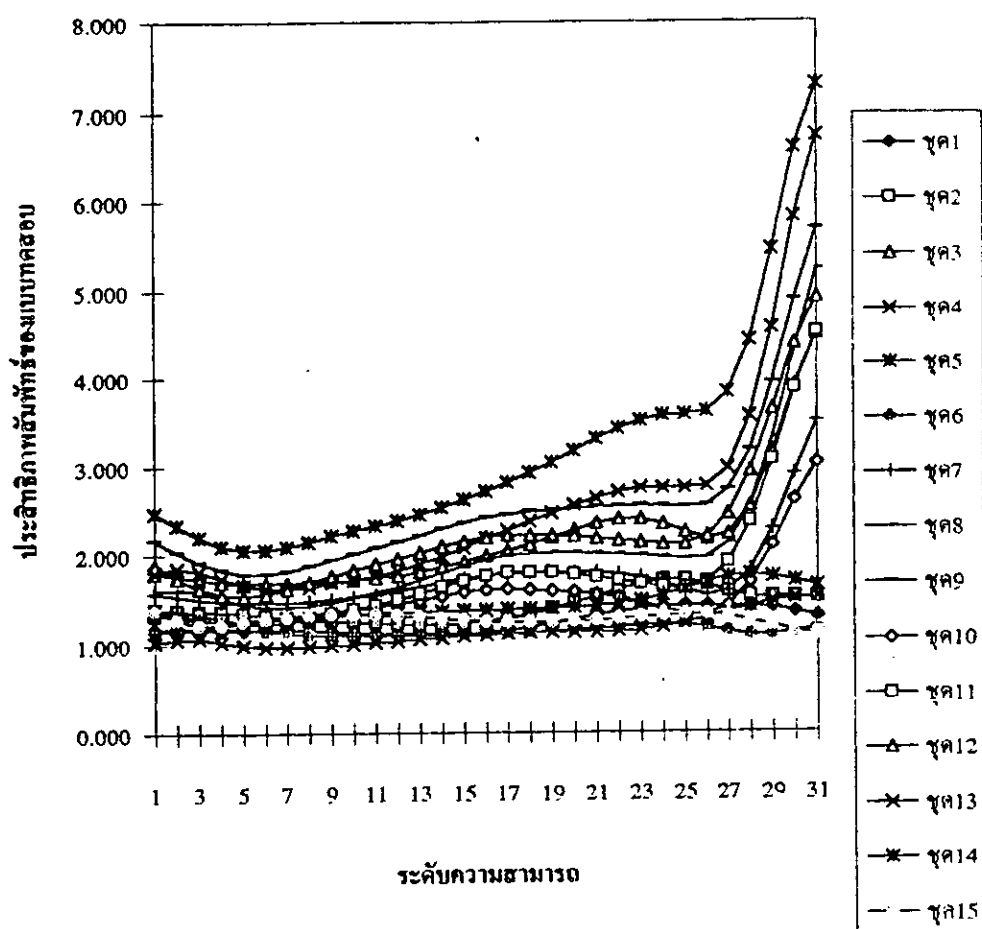
ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 86 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 600 คน ณ
ระดับความสามารถต่าง ๆ

หมายเหตุ

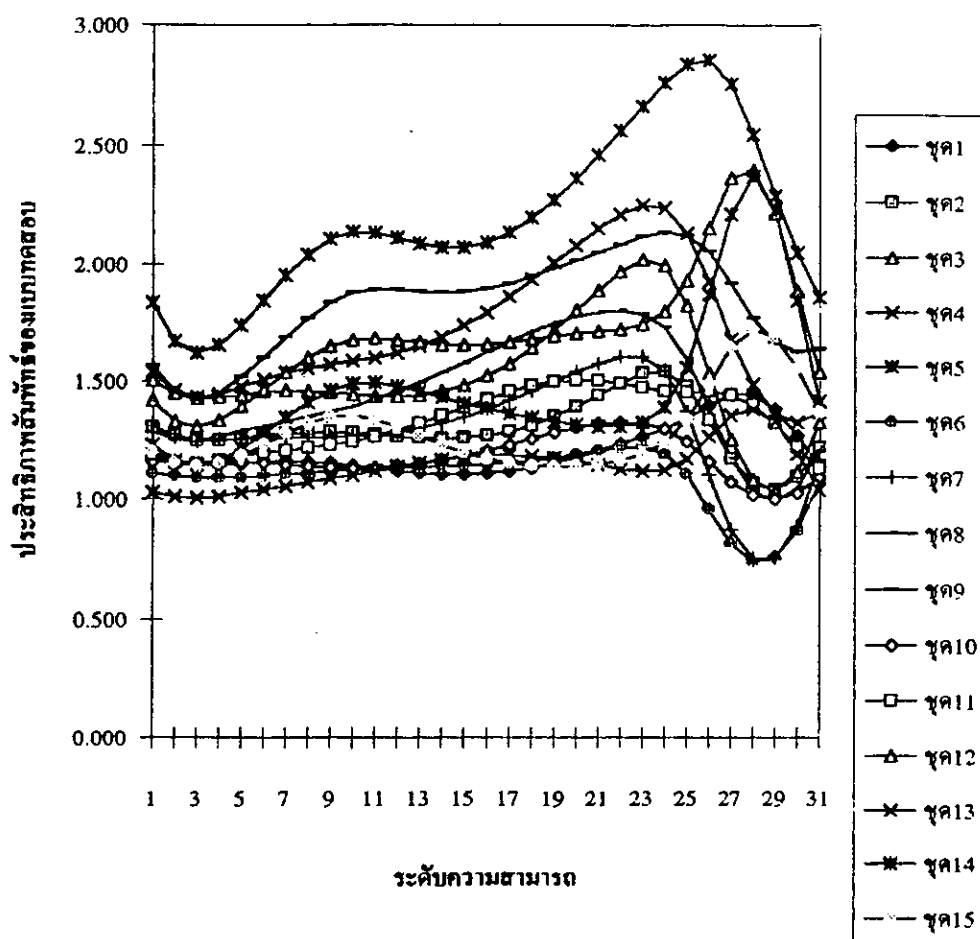
ชุด 1 จำนวน 60/50 ข้อ	ชุด 9 จำนวน 100/60 ข้อ
ชุด 2 จำนวน 70/50 ข้อ	ชุด 10 จำนวน 80/70 ข้อ
ชุด 3 จำนวน 80/50 ข้อ	ชุด 11 จำนวน 90/70 ข้อ
ชุด 4 จำนวน 90/50 ข้อ	ชุด 12 จำนวน 100/70 ข้อ
ชุด 5 จำนวน 100/50 ข้อ	ชุด 13 จำนวน 90/80 ข้อ
ชุด 6 จำนวน 70/60 ข้อ	ชุด 14 จำนวน 100/80 ข้อ
ชุด 7 จำนวน 80/60 ข้อ	ชุด 15 จำนวน 100/90 ข้อ
ชุด 8 จำนวน 90/60 ข้อ	



ภาพประกอบ 87 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 700 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

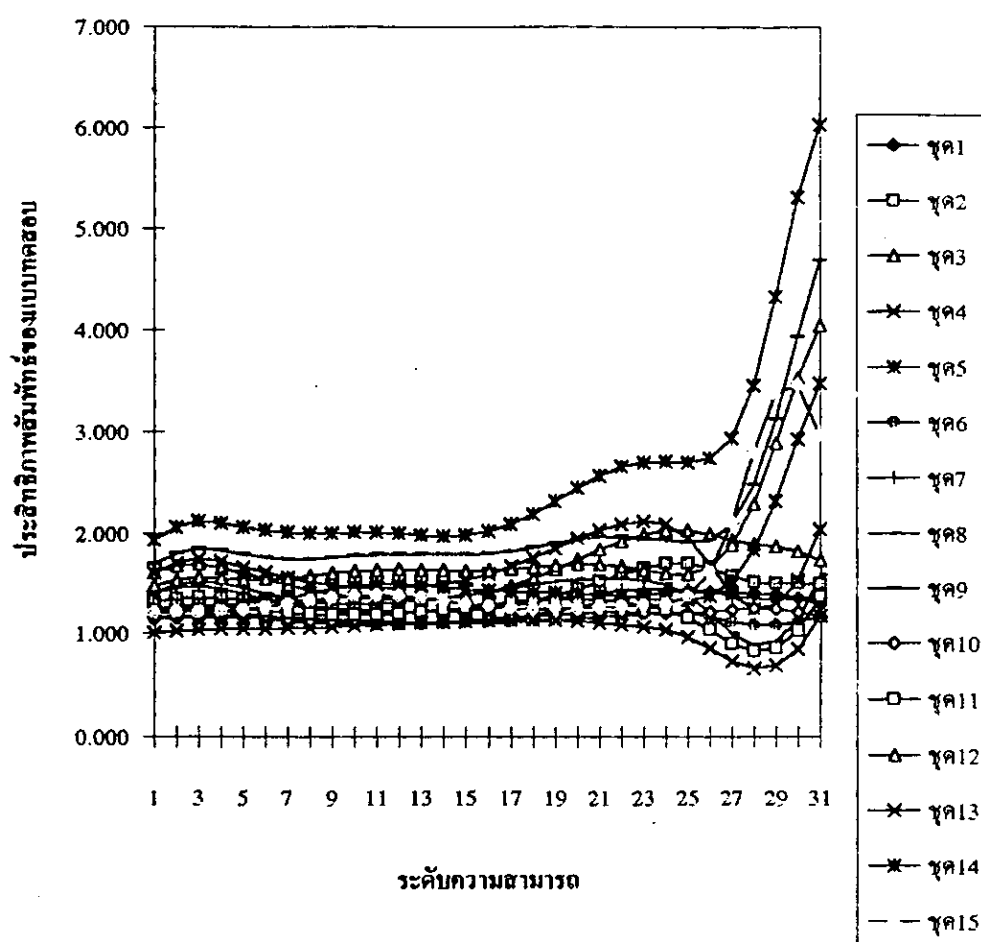
ชด 1	จำนวน 60/50 ข้อ	ชด 9	จำนวน 100/60 ข้อ
ชด 2	จำนวน 70/50 ข้อ	ชด 10	จำนวน 80/70 ข้อ
ชด 3	จำนวน 80/50 ข้อ	ชด 11	จำนวน 90/70 ข้อ
ชด 4	จำนวน 90/50 ข้อ	ชด 12	จำนวน 100/70 ข้อ
ชด 5	จำนวน 100/50 ข้อ	ชด 13	จำนวน 90/80 ข้อ
ชด 6	จำนวน 70/60 ข้อ	ชด 14	จำนวน 100/80 ข้อ
ชด 7	จำนวน 80/60 ข้อ	ชด 15	จำนวน 100/90 ข้อ
ชด 8	จำนวน 90/60 ข้อ		



ภาพประกอบ 88 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 800 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ	ชุด 9 จำนวน 100/60 ข้อ
ชุด 2 จำนวน 70/50 ข้อ	ชุด 10 จำนวน 80/70 ข้อ
ชุด 3 จำนวน 80/50 ข้อ	ชุด 11 จำนวน 90/70 ข้อ
ชุด 4 จำนวน 90/50 ข้อ	ชุด 12 จำนวน 100/70 ข้อ
ชุด 5 จำนวน 100/50 ข้อ	ชุด 13 จำนวน 90/80 ข้อ
ชุด 6 จำนวน 70/60 ข้อ	ชุด 14 จำนวน 100/80 ข้อ
ชุด 7 จำนวน 80/60 ข้อ	ชุด 15 จำนวน 100/90 ข้อ
ชุด 8 จำนวน 90/60 ข้อ	

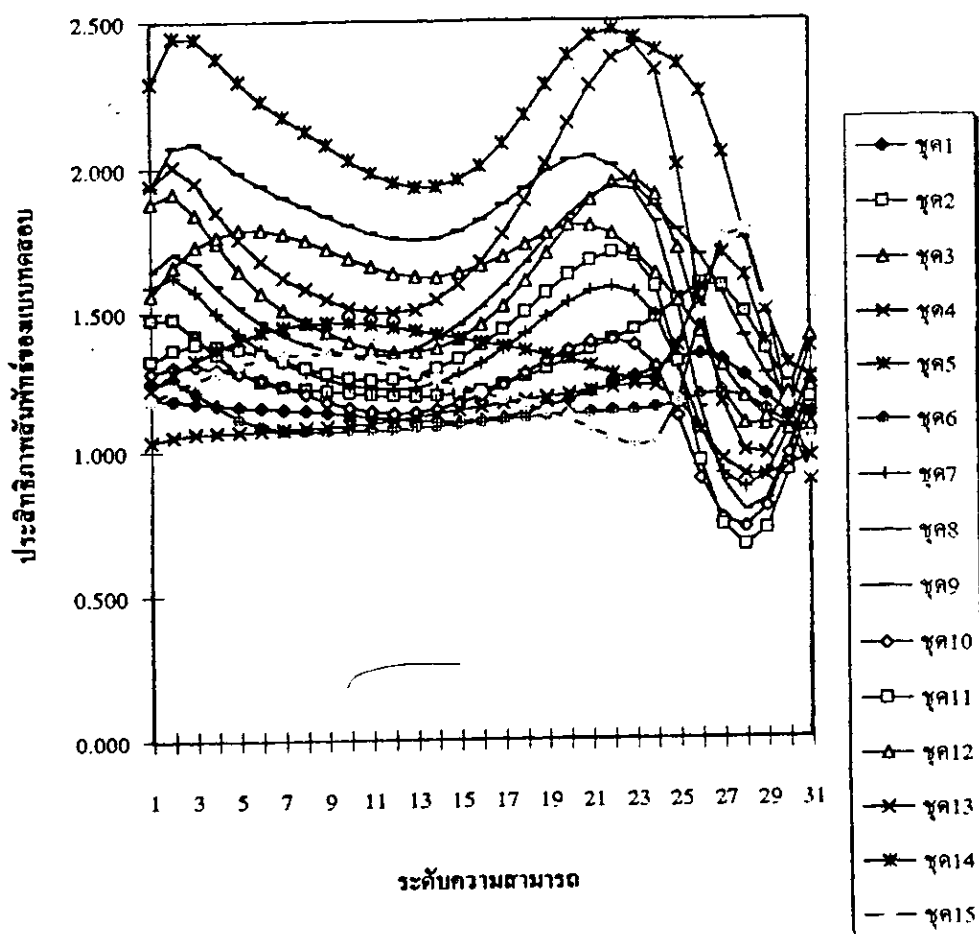


ภาพประกอบ 89 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 900 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
ชด 2 จำนวน 70/50 ข้อ
ชด 3 จำนวน 80/50 ข้อ
ชด 4 จำนวน 90/50 ข้อ
ชด 5 จำนวน 100/50 ข้อ
ชด 6 จำนวน 70/60 ข้อ
ชด 7 จำนวน 80/60 ข้อ
ชด 8 จำนวน 90/60 ข้อ

ชด 9 จำนวน 100/60 ข้อ
ชด 10 จำนวน 80/70 ข้อ
ชด 11 จำนวน 90/70 ข้อ
ชด 12 จำนวน 100/70 ข้อ
ชด 13 จำนวน 90/80 ข้อ
ชด 14 จำนวน 100/80 ข้อ
ชด 15 จำนวน 100/90 ข้อ

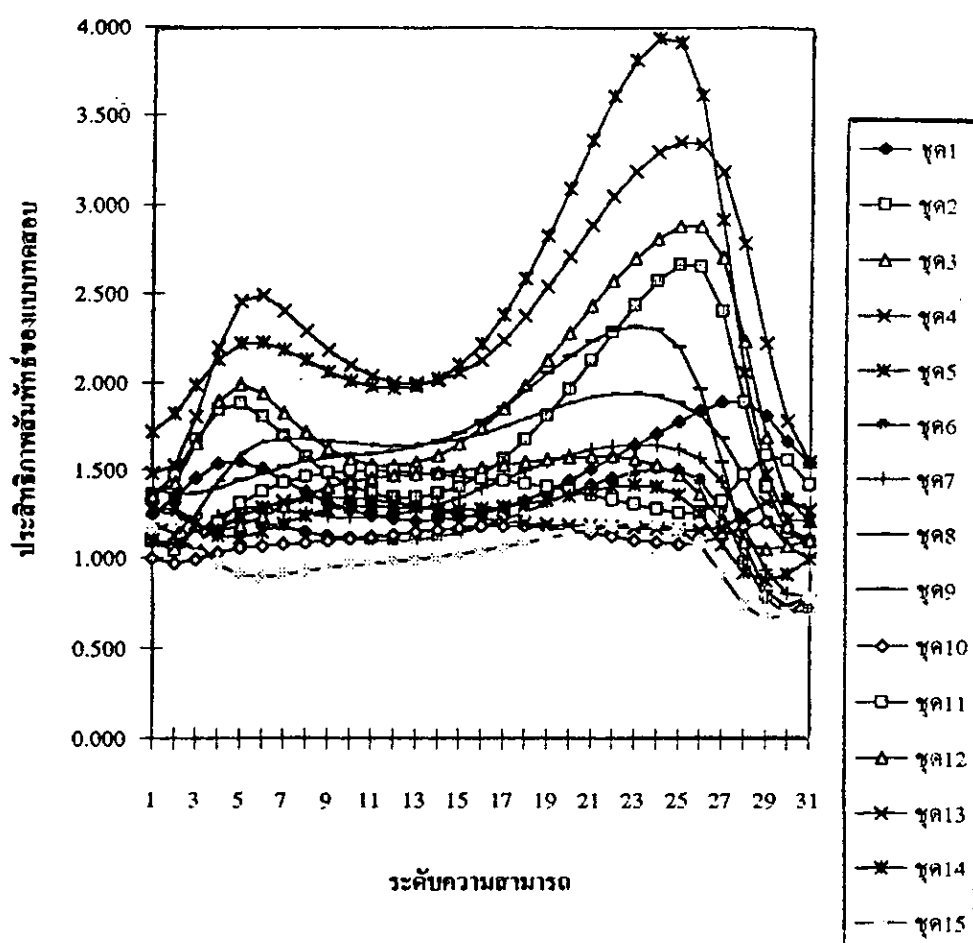


ภาพประกอบ 90 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

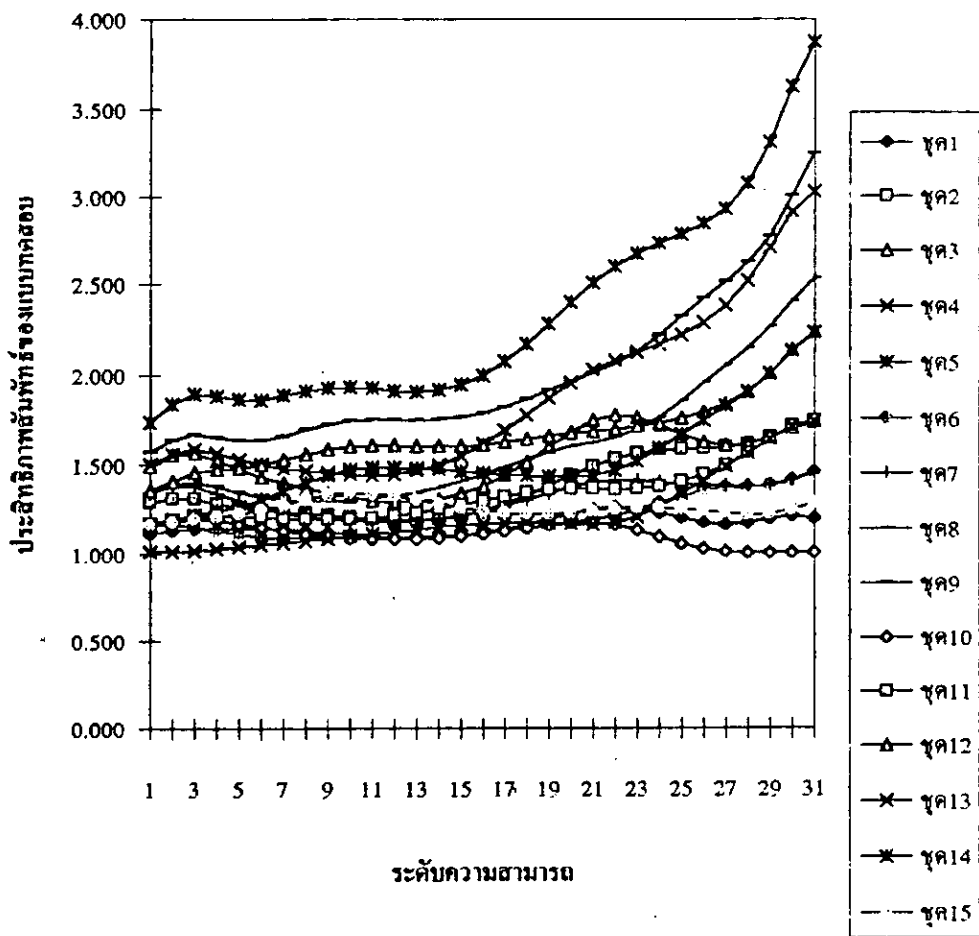


ภาพประกอบ 91 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิสต์ เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน
ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

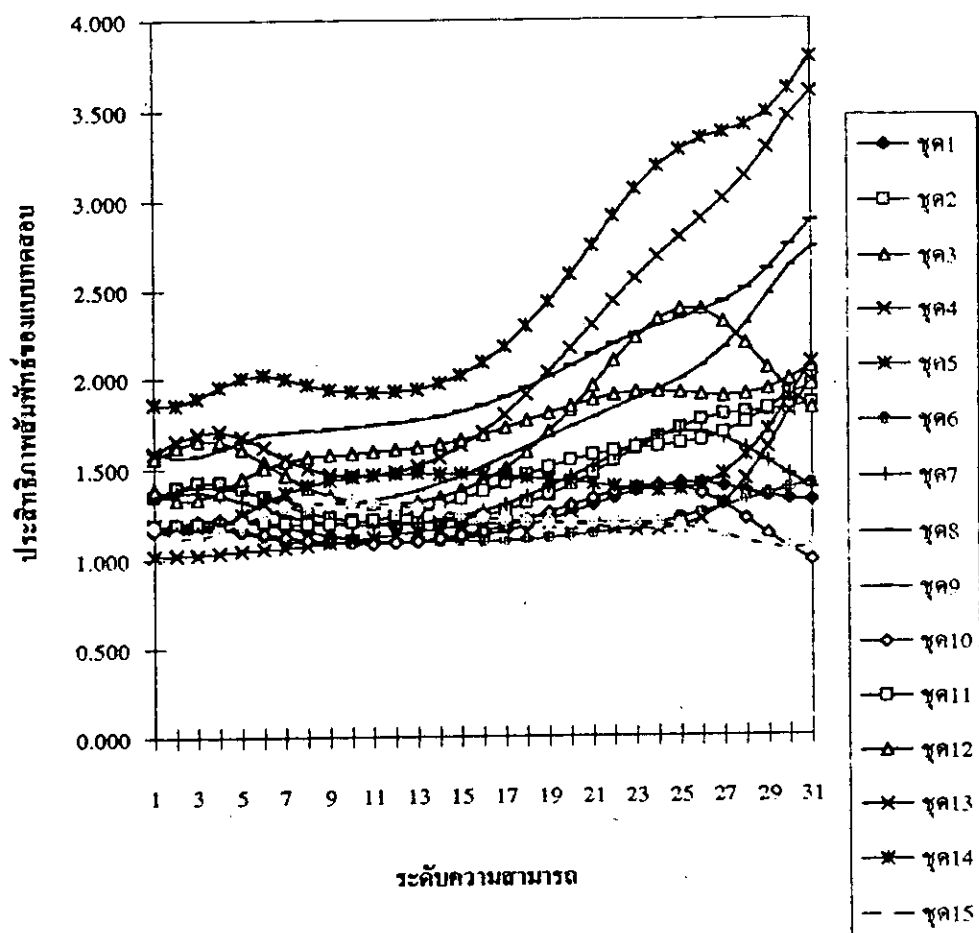


ภาพประกอบ 92 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน
ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
ชด 2 จำนวน 70/50 ข้อ
ชด 3 จำนวน 80/50 ข้อ
ชด 4 จำนวน 90/50 ข้อ
ชด 5 จำนวน 100/50 ข้อ
ชด 6 จำนวน 70/60 ข้อ
ชด 7 จำนวน 80/60 ข้อ
ชด 8 จำนวน 90/60 ข้อ

ชด 9 จำนวน 100/60 ข้อ
ชด 10 จำนวน 80/70 ข้อ
ชด 11 จำนวน 90/70 ข้อ
ชด 12 จำนวน 100/70 ข้อ
ชด 13 จำนวน 90/80 ข้อ
ชด 14 จำนวน 100/80 ข้อ
ชด 15 จำนวน 100/90 ข้อ

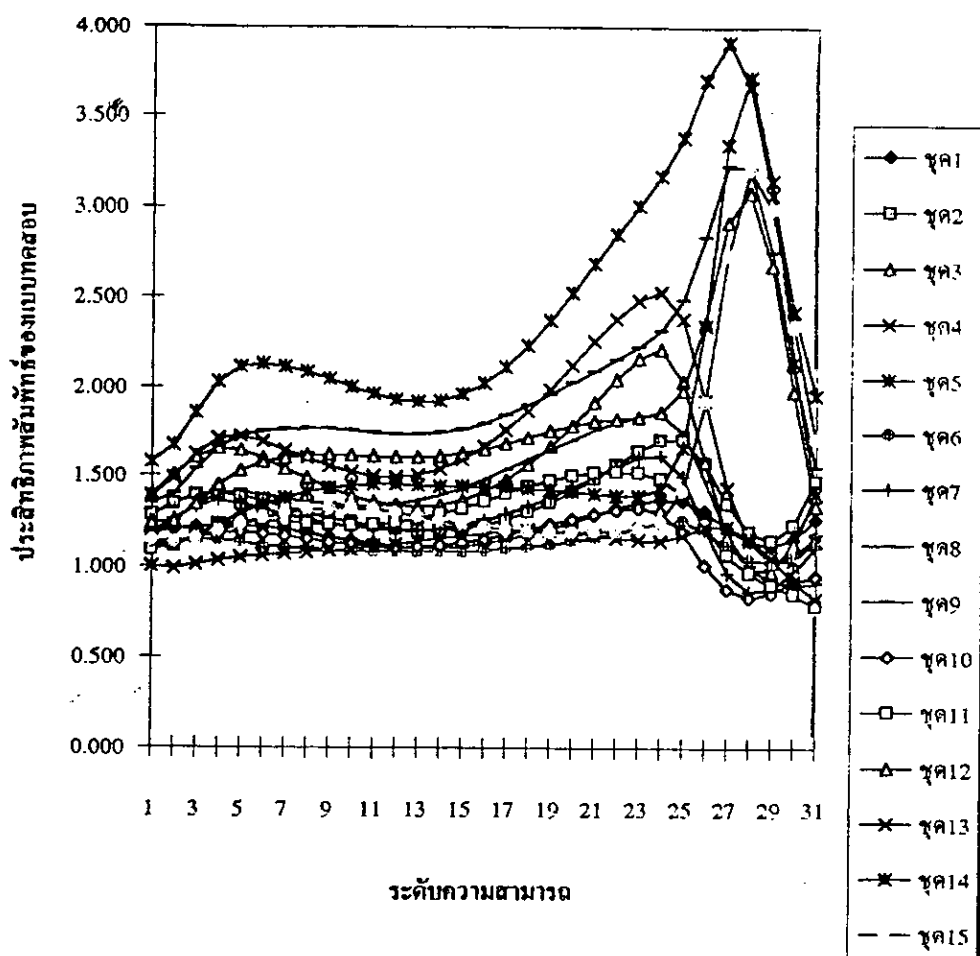


ภาพประกอบ 93 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม ไลค์ลิฮูด เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 94 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีแมกซิมัม โลคัลลิสต์ เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน
ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

จากภาพประกอบ 91 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 1,500 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(100, 50) = 3.936 ; \theta = 1.600\}$

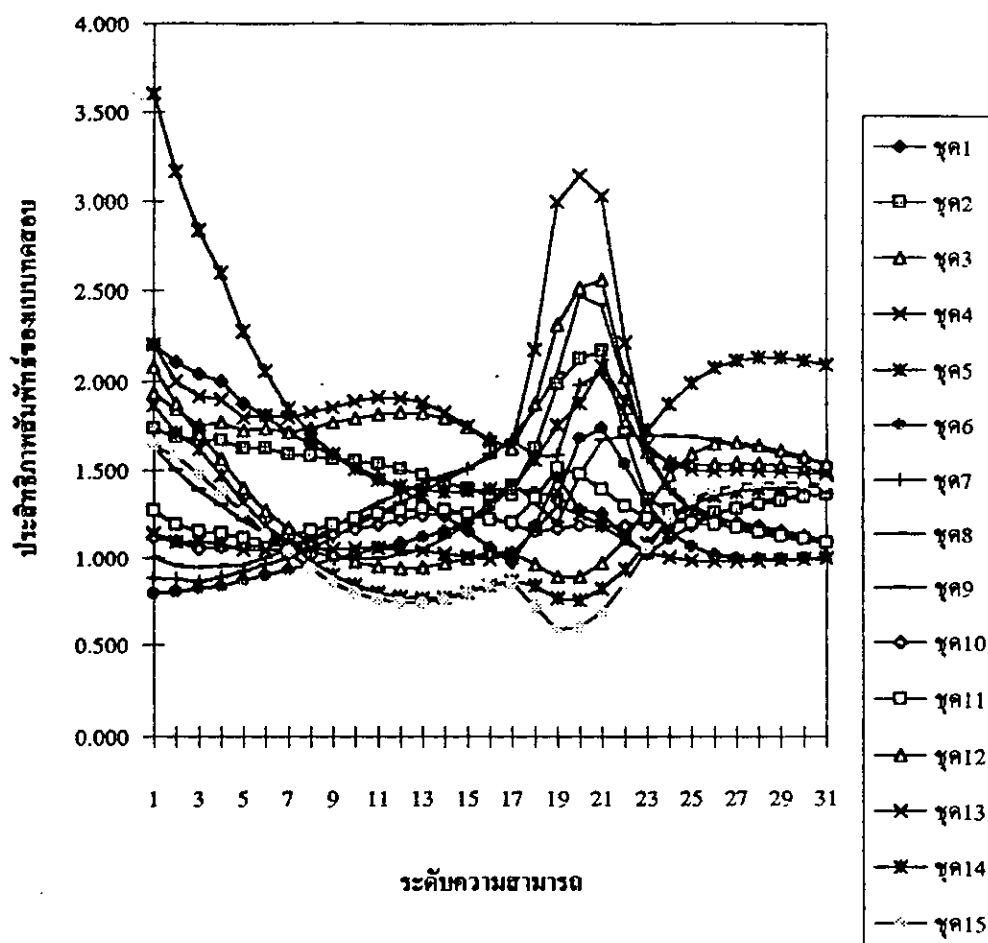
จากภาพประกอบ 92 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 2,000 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(100, 50) = 3.387 ; \theta = 3.000\}$

จากภาพประกอบ 93 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 2,500 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(100, 50) = 3.595 ; \theta = 3.000\}$

จากภาพประกอบ 94 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 3,000 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(100, 50) = 3.918 ; \theta = 2.200\}$

ดังนั้น จากภาพประกอบ 82 - 94 แสดงว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 700 คน ที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิตูด มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(100, 50) = 7.295 ; \theta = 3.000\}$

8.2 ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีริสติก เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 95 - 107

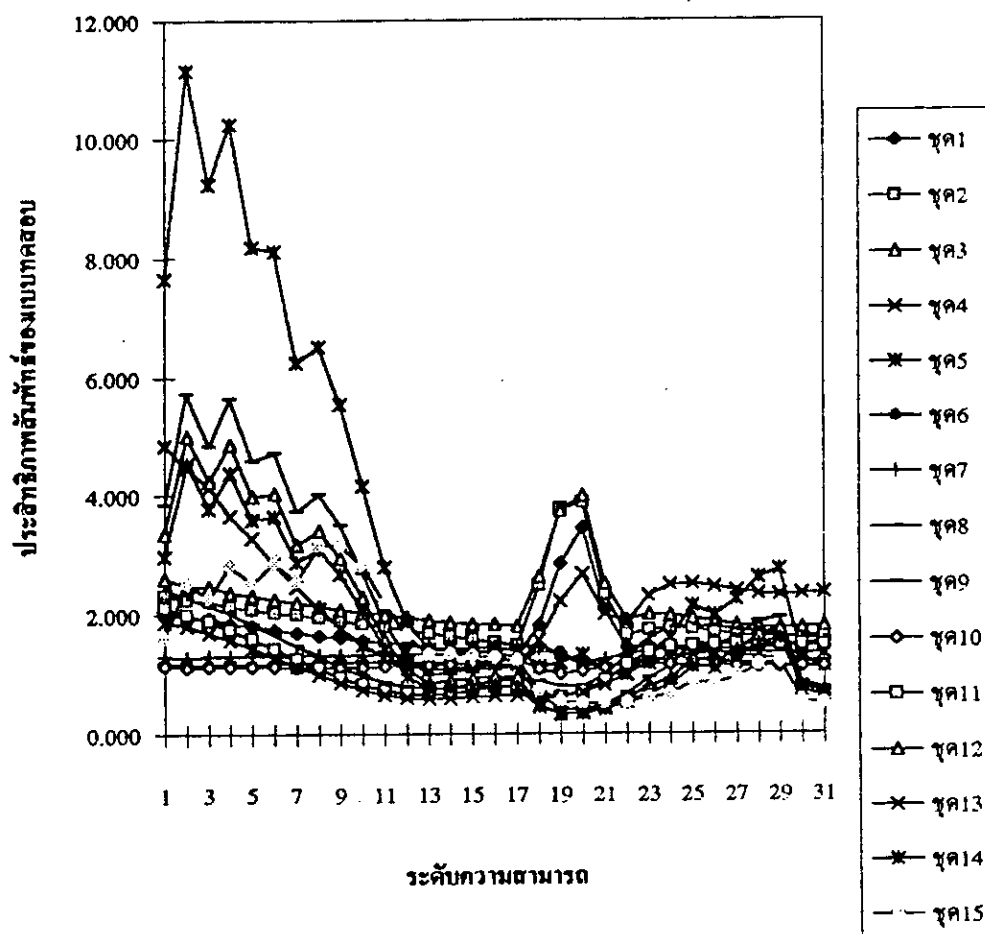


ภาพประกอบ 95 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีริสติก เมื่อกลุ่มตัวอย่างมีขนาด 200 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

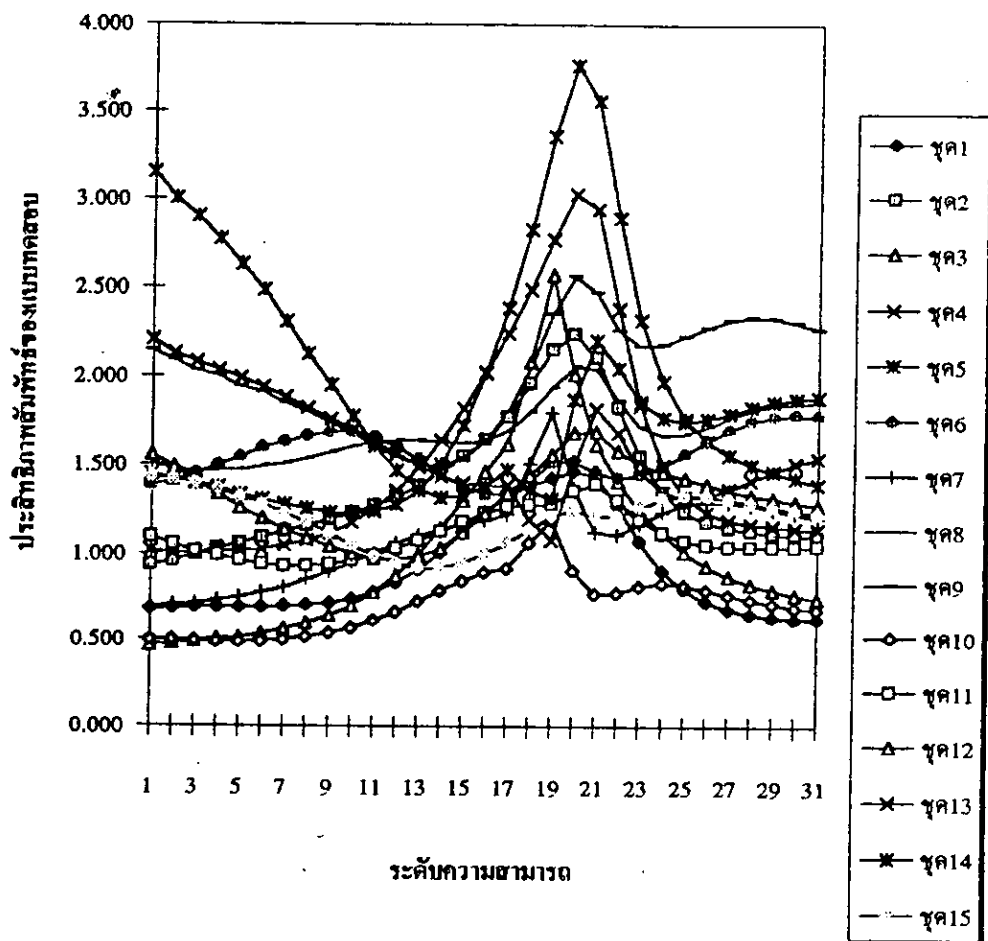
ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 96 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อกลุ่มตัวอย่างมีขนาด 300 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

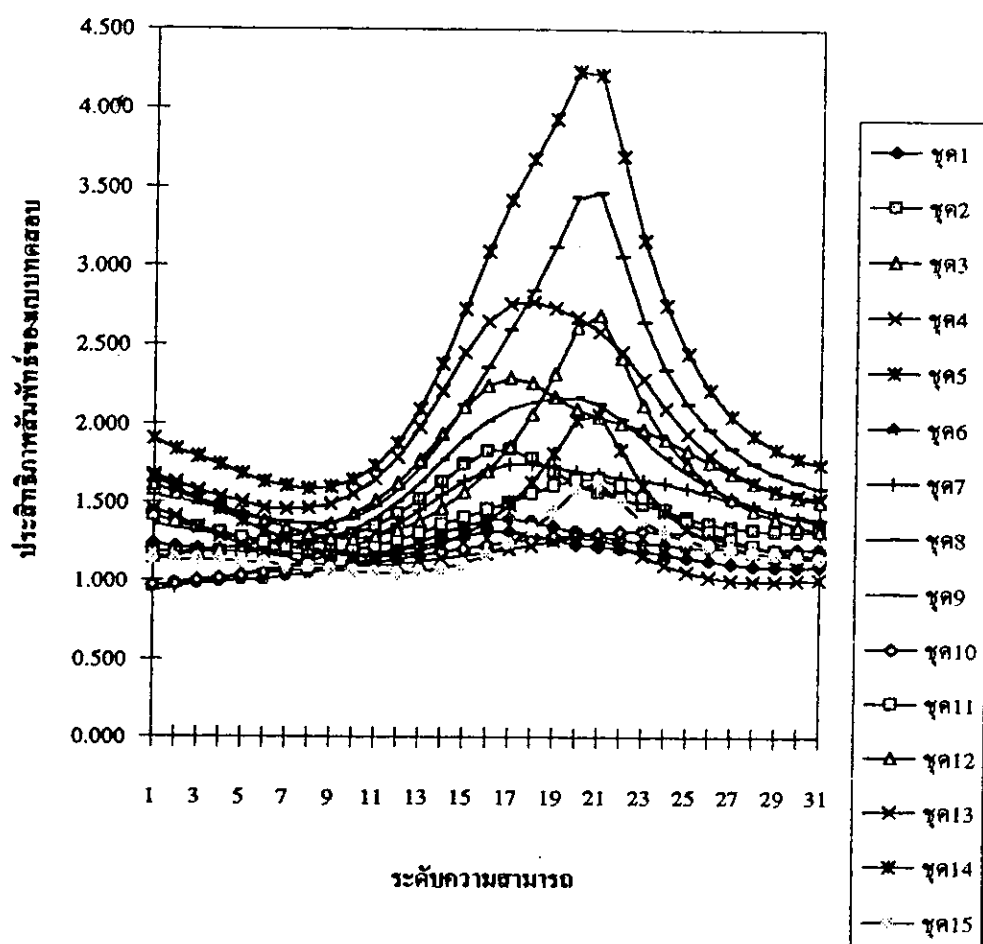
ชด 1 จำนวน 60/50 ข้อ	ชด 9 จำนวน 100/60 ข้อ
ชด 2 จำนวน 70/50 ข้อ	ชด 10 จำนวน 80/70 ข้อ
ชด 3 จำนวน 80/50 ข้อ	ชด 11 จำนวน 90/70 ข้อ
ชด 4 จำนวน 90/50 ข้อ	ชด 12 จำนวน 100/70 ข้อ
ชด 5 จำนวน 100/50 ข้อ	ชด 13 จำนวน 90/80 ข้อ
ชด 6 จำนวน 70/60 ข้อ	ชด 14 จำนวน 100/80 ข้อ
ชด 7 จำนวน 80/60 ข้อ	ชด 15 จำนวน 100/90 ข้อ
ชด 8 จำนวน 90/60 ข้อ	



ภาพประกอบ 97 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีริสติก เมื่อกลุ่มตัวอย่างมีขนาด 400 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

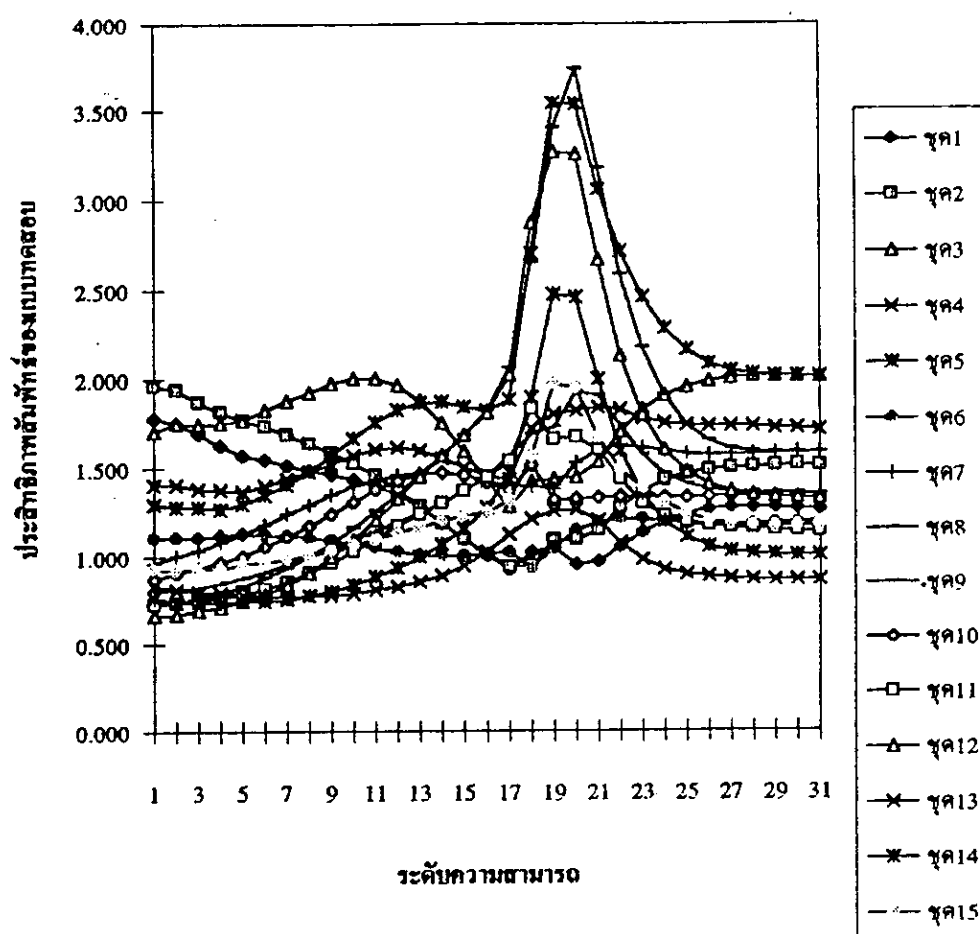
ชุด 1 จำนวน 60/50 ข้อ	ชุด 9 จำนวน 100/60 ข้อ
ชุด 2 จำนวน 70/50 ข้อ	ชุด 10 จำนวน 80/70 ข้อ
ชุด 3 จำนวน 80/50 ข้อ	ชุด 11 จำนวน 90/70 ข้อ
ชุด 4 จำนวน 90/50 ข้อ	ชุด 12 จำนวน 100/70 ข้อ
ชุด 5 จำนวน 100/50 ข้อ	ชุด 13 จำนวน 90/80 ข้อ
ชุด 6 จำนวน 70/60 ข้อ	ชุด 14 จำนวน 100/80 ข้อ
ชุด 7 จำนวน 80/60 ข้อ	ชุด 15 จำนวน 100/90 ข้อ
ชุด 8 จำนวน 90/60 ข้อ	



ภาพประกอบ 98 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีวิธีสถิติ เมื่อกลุ่มตัวอย่างมีขนาด 500 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ	ชด 9 จำนวน 100/60 ข้อ
ชด 2 จำนวน 70/50 ข้อ	ชด 10 จำนวน 80/70 ข้อ
ชด 3 จำนวน 80/50 ข้อ	ชด 11 จำนวน 90/70 ข้อ
ชด 4 จำนวน 90/50 ข้อ	ชด 12 จำนวน 100/70 ข้อ
ชด 5 จำนวน 100/50 ข้อ	ชด 13 จำนวน 90/80 ข้อ
ชด 6 จำนวน 70/60 ข้อ	ชด 14 จำนวน 100/80 ข้อ
ชด 7 จำนวน 80/60 ข้อ	ชด 15 จำนวน 100/90 ข้อ
ชด 8 จำนวน 90/60 ข้อ	

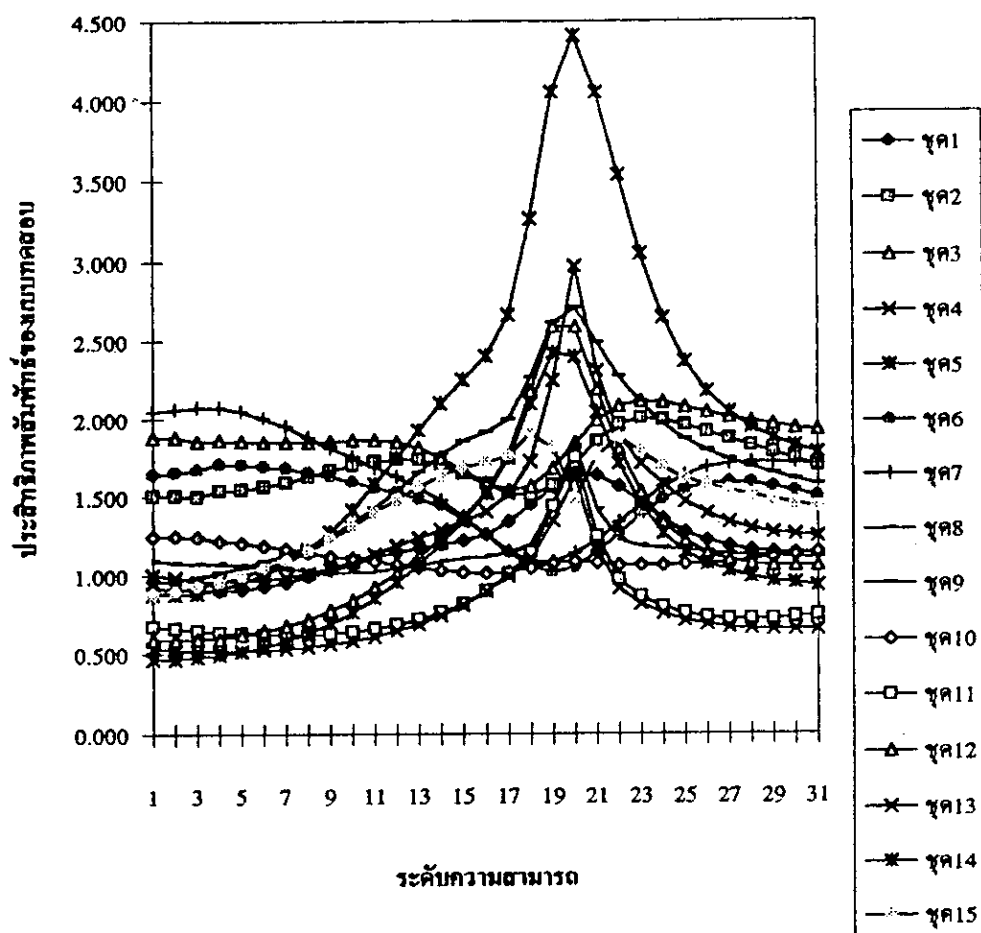


ภาพประกอบ 99 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อกลุ่มตัวอย่างมีขนาด 600 คน ณ ระดับ
ความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

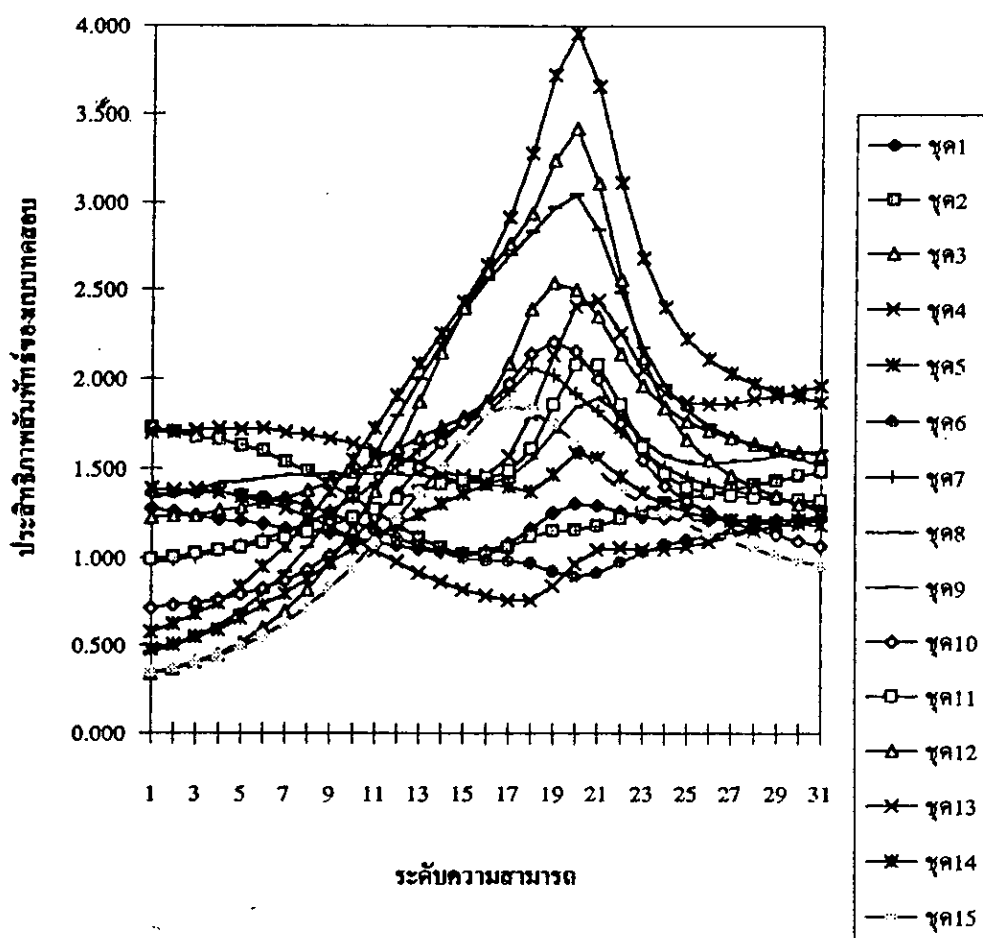


ภาพประกอบ 100 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน เมื่อกลุ่มตัวอย่างมีขนาด 700 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
 ชุด 2 จำนวน 70/50 ข้อ
 ชุด 3 จำนวน 80/50 ข้อ
 ชุด 4 จำนวน 90/50 ข้อ
 ชุด 5 จำนวน 100/50 ข้อ
 ชุด 6 จำนวน 70/60 ข้อ
 ชุด 7 จำนวน 80/60 ข้อ
 ชุด 8 จำนวน 90/60 ข้อ

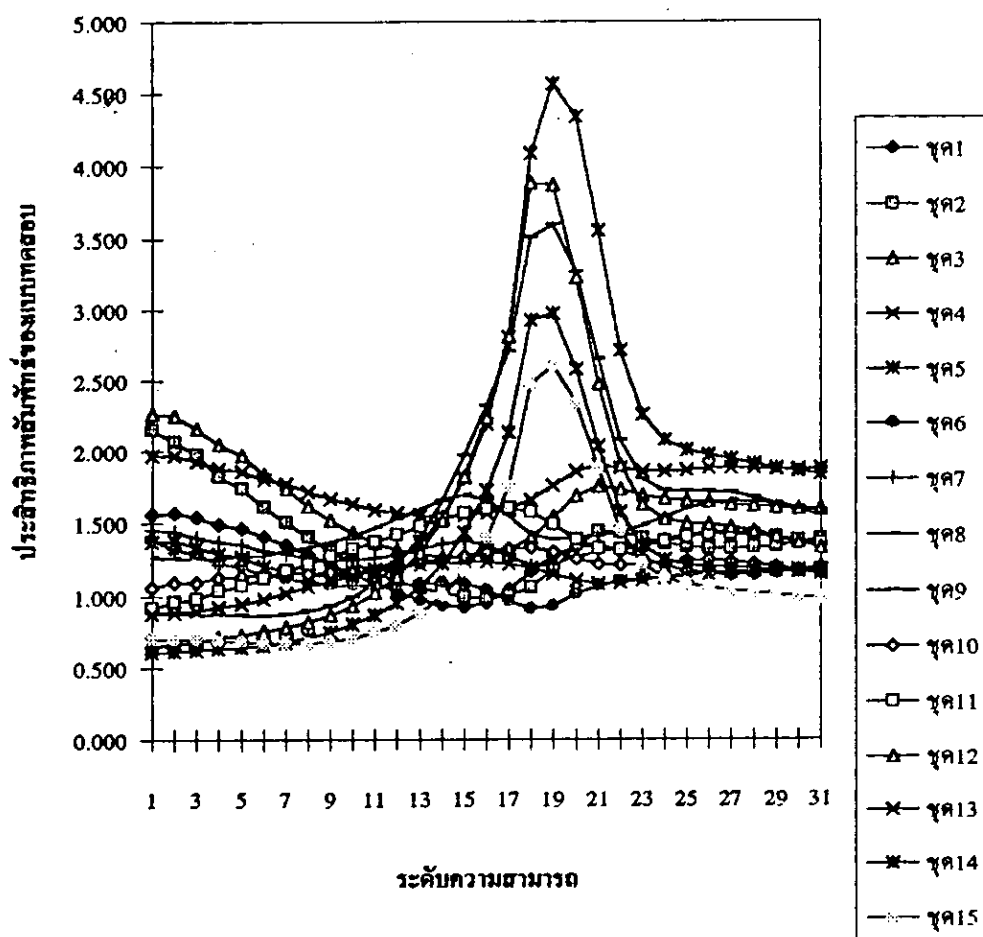
ชุด 9 จำนวน 100/60 ข้อ
 ชุด 10 จำนวน 80/70 ข้อ
 ชุด 11 จำนวน 90/70 ข้อ
 ชุด 12 จำนวน 100/70 ข้อ
 ชุด 13 จำนวน 90/80 ข้อ
 ชุด 14 จำนวน 100/80 ข้อ
 ชุด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 101 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีอีวีวิสต์ค เมื่อกลุ่มตัวอย่างมีขนาด 800 คน ณ ระดับ
ความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ	ชด 9 จำนวน 100/60 ข้อ
ชด 2 จำนวน 70/50 ข้อ	ชด 10 จำนวน 80/70 ข้อ
ชด 3 จำนวน 80/50 ข้อ	ชด 11 จำนวน 90/70 ข้อ
ชด 4 จำนวน 90/50 ข้อ	ชด 12 จำนวน 100/70 ข้อ
ชด 5 จำนวน 100/50 ข้อ	ชด 13 จำนวน 90/80 ข้อ
ชด 6 จำนวน 70/60 ข้อ	ชด 14 จำนวน 100/80 ข้อ
ชด 7 จำนวน 80/60 ข้อ	ชด 15 จำนวน 100/90 ข้อ
ชด 8 จำนวน 90/60 ข้อ	

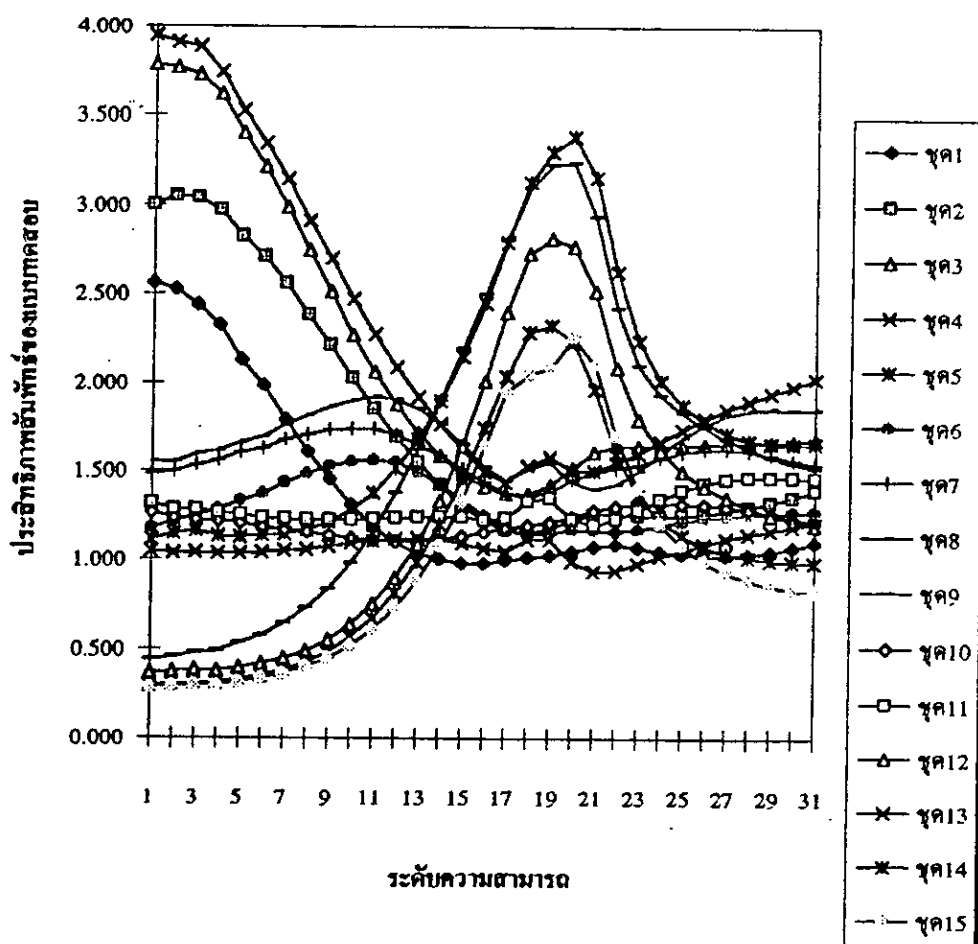


ภาพประกอบ 102 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีวิวิธวิธี เมื่อกลุ่มตัวอย่างมีขนาด 900 คน ณ ระดับ
ความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

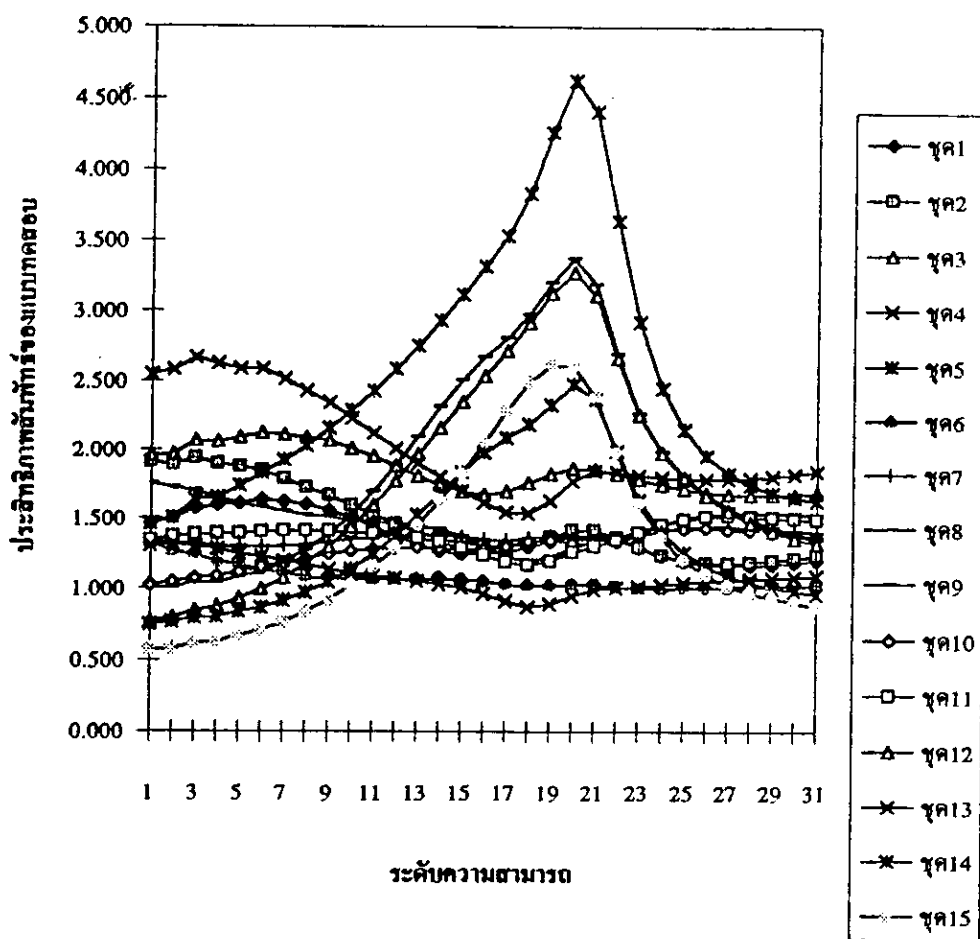


ภาพประกอบ 103 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีริสติก เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

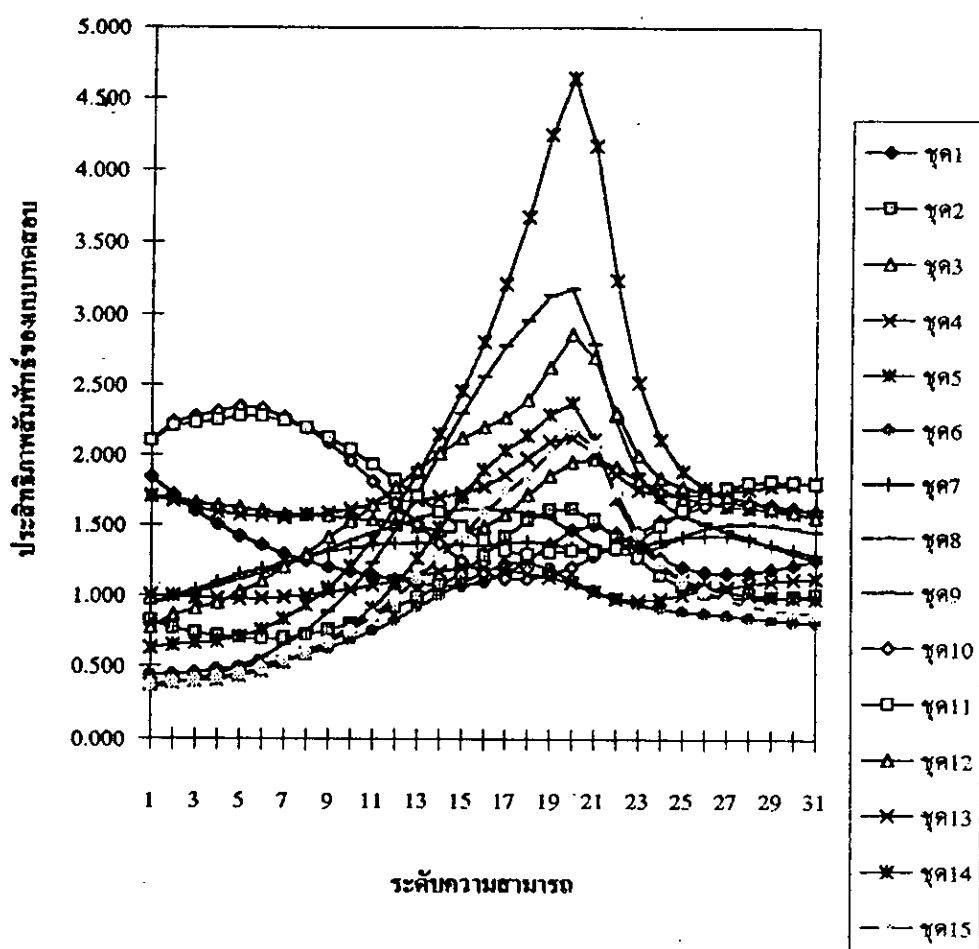


ภาพประกอบ 104 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีวิธีสถิติ เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน ณ ระดับ
ความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

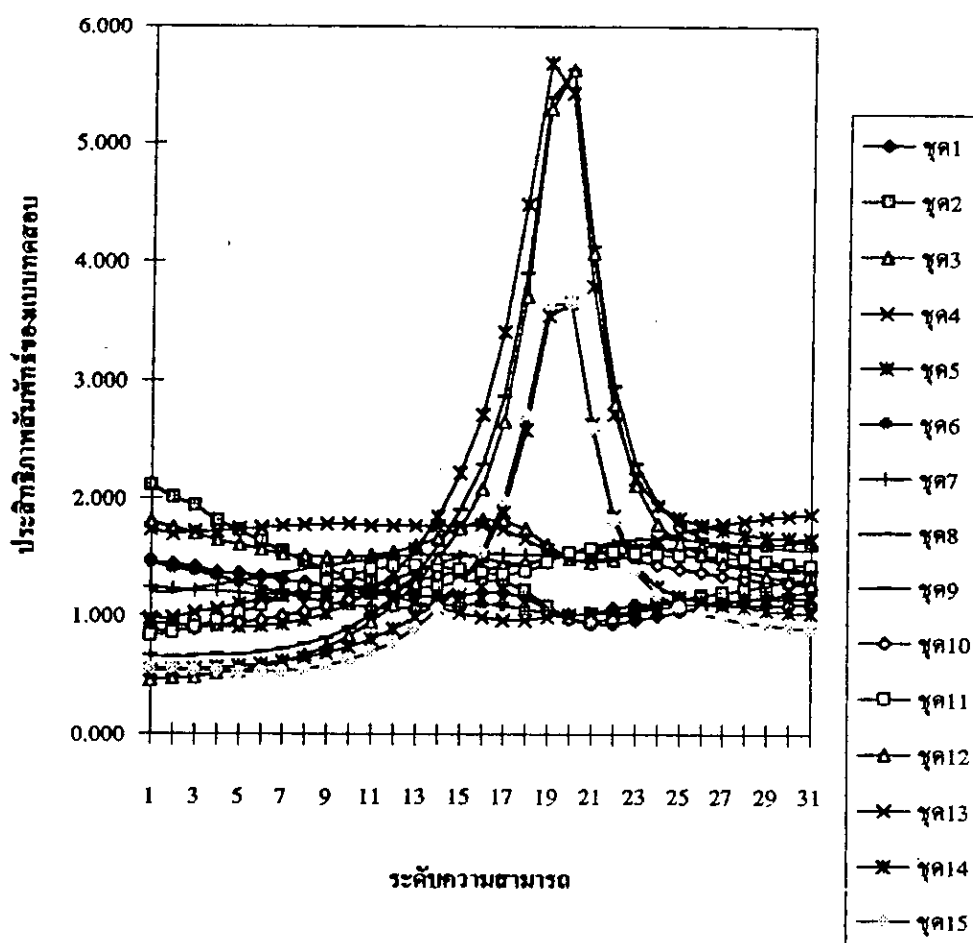


ภาพประกอบ 105 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีวิธีสถิติ เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน ณ ระดับ
ความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 หัสน
ชด 2 จำนวน 70/50 หัสน
ชด 3 จำนวน 80/50 หัสน
ชด 4 จำนวน 90/50 หัสน
ชด 5 จำนวน 100/50 หัสน
ชด 6 จำนวน 70/60 หัสน
ชด 7 จำนวน 80/60 หัสน
ชด 8 จำนวน 90/60 หัสน

ชด 9 จำนวน 100/60 หัสน
ชด 10 จำนวน 80/70 หัสน
ชด 11 จำนวน 90/70 หัสน
ชด 12 จำนวน 100/70 หัสน
ชด 13 จำนวน 90/80 หัสน
ชด 14 จำนวน 100/80 หัสน
ชด 15 จำนวน 100/90 หัสน

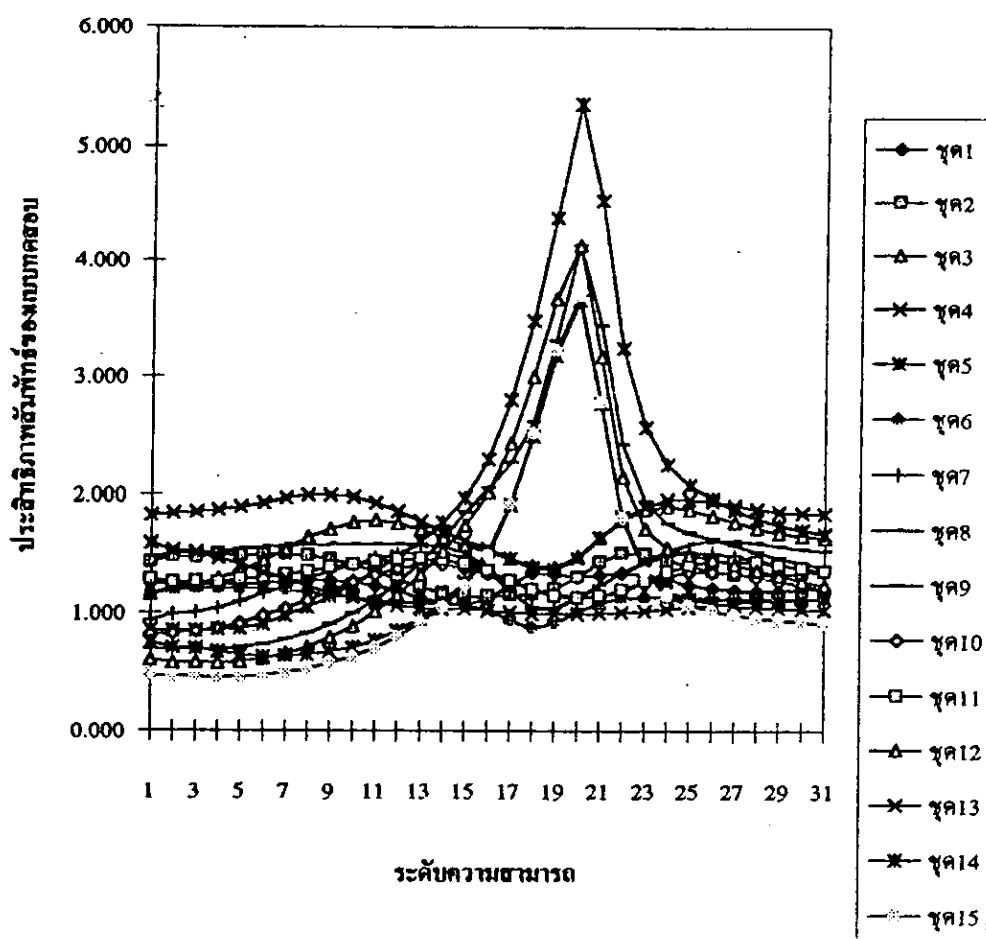


ภาพประกอบ 106 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีอีวีริสติค เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 107 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีนิวตัน-ราฟสัน เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
 ชด 2 จำนวน 70/50 ข้อ
 ชด 3 จำนวน 80/50 ข้อ
 ชด 4 จำนวน 90/50 ข้อ
 ชด 5 จำนวน 100/50 ข้อ
 ชด 6 จำนวน 70/60 ข้อ
 ชด 7 จำนวน 80/60 ข้อ
 ชด 8 จำนวน 90/60 ข้อ

ชด 9 จำนวน 100/60 ข้อ
 ชด 10 จำนวน 80/70 ข้อ
 ชด 11 จำนวน 90/70 ข้อ
 ชด 12 จำนวน 100/70 ข้อ
 ชด 13 จำนวน 90/80 ข้อ
 ชด 14 จำนวน 100/80 ข้อ
 ชด 15 จำนวน 100/90 ข้อ

จากภาพประกอบ 95 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 200 คน ที่ประมาณค่าด้วยวิธีวิวิธวิธี มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(100, 50) = 3.600 ; \sigma = -3.000\}$

จากภาพประกอบ 96 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 300 คน ที่ประมาณค่าด้วยวิธีวิธีวิวัตติค มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(100, 50) = 11.143 ; \sigma = -2.800\}$

จากภาพประกอบ 97 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 400 คน ที่ประมาณค่าด้วยวิธีอิวิริสติก มีค่าสูงสุด ที่ระดับความสามารถปานกลาง ($RE(100, 50) = 3.764$; $e = 0.800$)

จากภาพประกอบ 98 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 500 คน ที่ประมาณค่าด้วยวิธีวิธีวิเศษิต มีค่าส่งผล ที่ระดับความสามารถปานกลาง $\{RE(100, 50) = 4.229 ; \sigma = 0.800\}$

จากภาพประกอบ 99 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 60 ข้อ ของกลุ่มตัวอย่าง 600 คน ที่ประมาณค่าด้วยวิธีอวิริสติก มีค่าสูงสุด ที่ระดับความสามารถปานกลาง ($RE(100, 60) = 3.733$; $\sigma = 0.800$)

จากภาพประกอบ 100 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 700 คน ที่ประมาณค่าด้วยวิธีวิวิธิต มีค่าสูงสุด ที่ระดับความสามารถปานกลาง $\{RE(100, 50) = 4.402 ; \sigma = 0.800\}$

จากภาพประกอบ 101 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 800 คน ที่ประมาณค่าด้วยวิธีอวิวิธิต มีค่าสูงสุด ที่ระดับความสามารถปานกลาง $\{RE(100, 50) = 3.954 ; \sigma = 0.800\}$

จากภาพประกอบ 102 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 900 คน ที่ประมาณค่าด้วยวิธีอิมพิริสติก มีค่าสูงสุด ที่ระดับความสามารถปานกลาง $\{RE(100, 50) = 4.566 ; \sigma = 0.600\}$

จากภาพประกอบ 103 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 90 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 1,000 คน ที่ประมาณค่าด้วยวิธีอิมพิริสติก มีค่าส่งผล ที่ระดับความสามารถต่ำ $\{RE(90, 50) = 3.944 ; \sigma = -3.000\}$

จากภาพประกอบ 104 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 1,500 คน ที่ประมาณค่าด้วยวิธีอิวิริสติก มีค่าสูงสุด ที่ระดับความสามารถปานกลาง $\{RE(100, 50) = 4.620 ; e = 0.800\}$

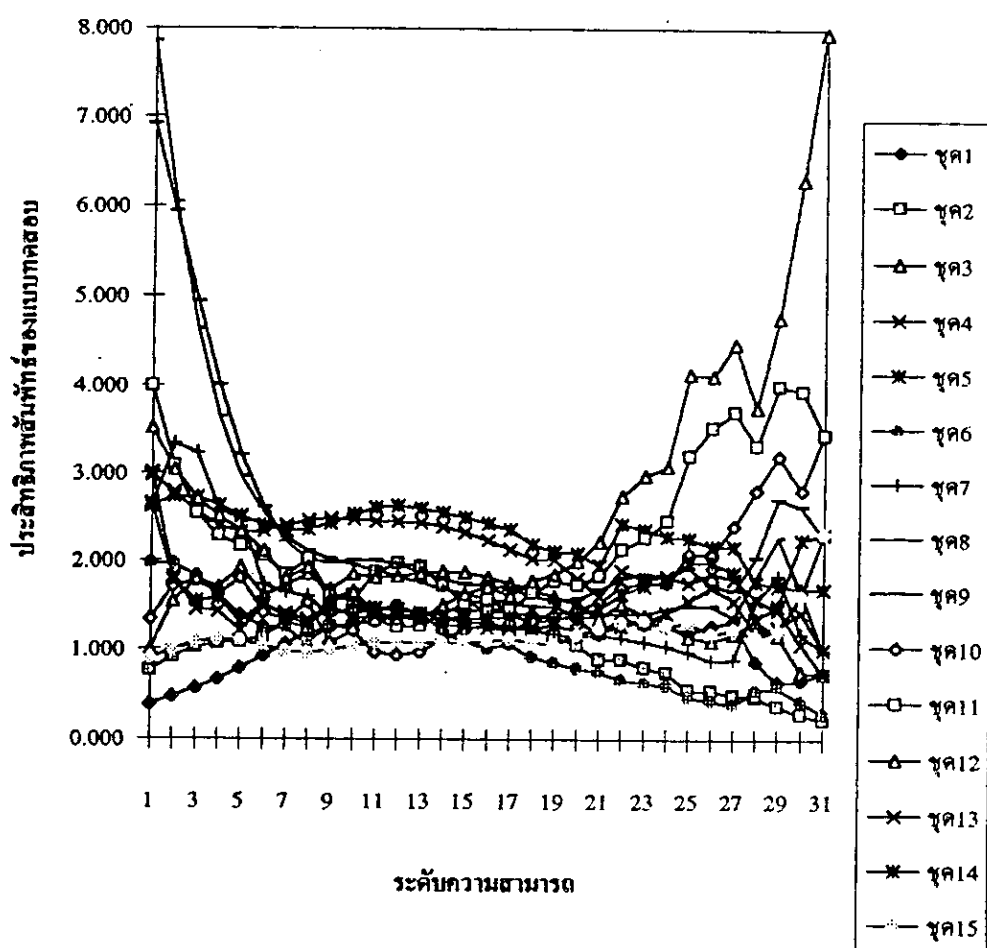
จากภาพประกอบ 105 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 2,000 คน ที่ประมาณค่าด้วยวิธีอิวิริสติก มีค่าสูงสุด ที่ระดับความสามารถปานกลาง $\{RE(100, 50) = 4.640 ; e = 0.800\}$

จากภาพประกอบ 106 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 2,500 คน ที่ประมาณค่าด้วยวิธีอิวิริสติก มีค่าสูงสุด ที่ระดับความสามารถปานกลาง $\{RE(100, 50) = 5.674 ; e = 0.600\}$

จากภาพประกอบ 107 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 3,000 คน ที่ประมาณค่าด้วยวิธีอิวิริสติก มีค่าสูงสุด ที่ระดับความสามารถปานกลาง $\{RE(100, 50) = 5.354 ; e = 0.800\}$

ดังนั้น จากภาพประกอบ 95 - 107 แสดงว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 300 คน ที่ประมาณค่าด้วยวิธีอิวิริสติก มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(100, 50) = 11.143 ; e = -2.800\}$

8.3 ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ ดังภาพประกอบ 108 - 120

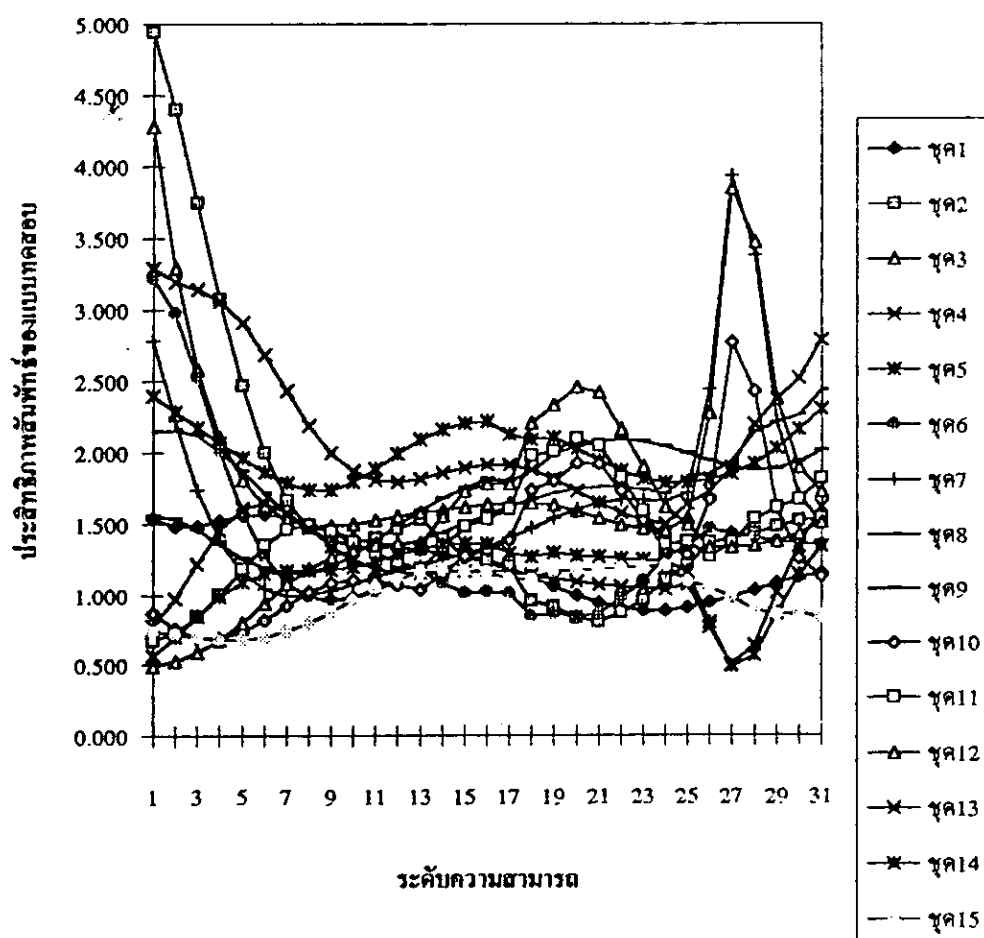


ภาพประกอบ 108 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 200 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
 ชุด 2 จำนวน 70/50 ข้อ
 ชุด 3 จำนวน 80/50 ข้อ
 ชุด 4 จำนวน 90/50 ข้อ
 ชุด 5 จำนวน 100/50 ข้อ
 ชุด 6 จำนวน 70/60 ข้อ
 ชุด 7 จำนวน 80/60 ข้อ
 ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
 ชุด 10 จำนวน 80/70 ข้อ
 ชุด 11 จำนวน 90/70 ข้อ
 ชุด 12 จำนวน 100/70 ข้อ
 ชุด 13 จำนวน 90/80 ข้อ
 ชุด 14 จำนวน 100/80 ข้อ
 ชุด 15 จำนวน 100/90 ข้อ

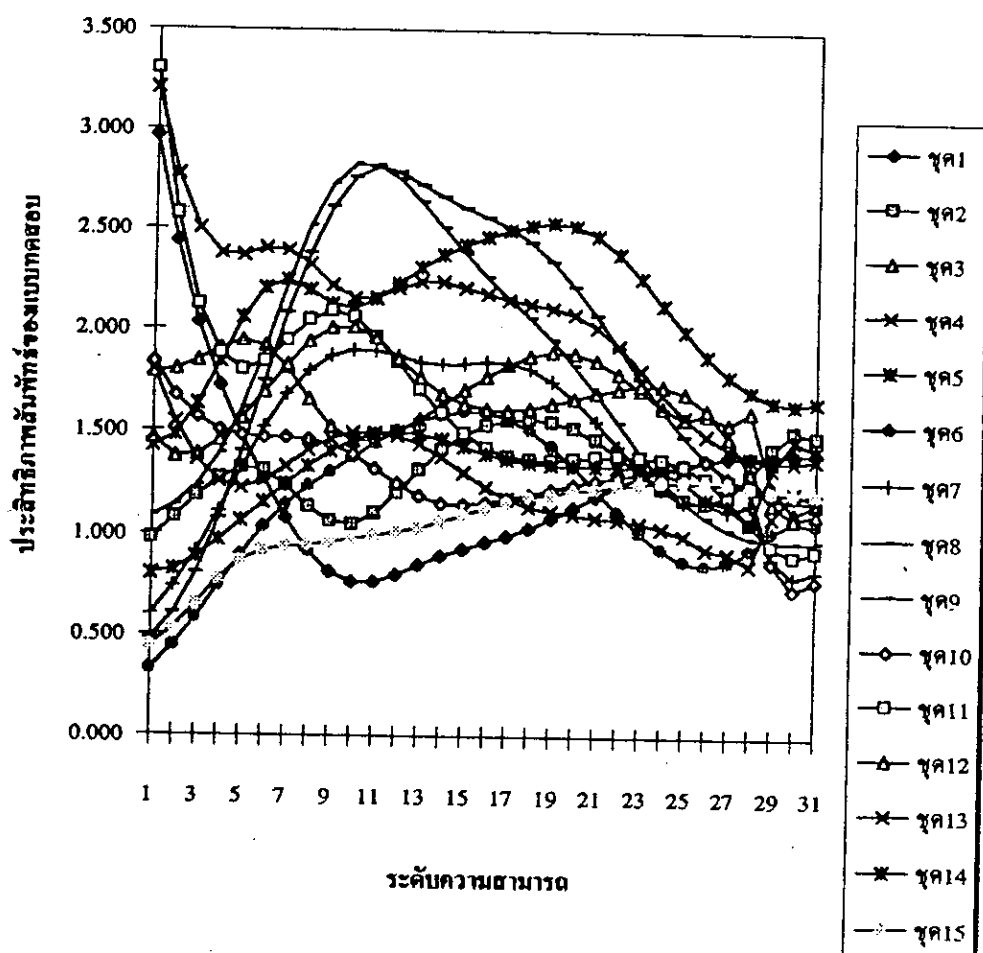


ภาพประกอบ 109 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 300 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

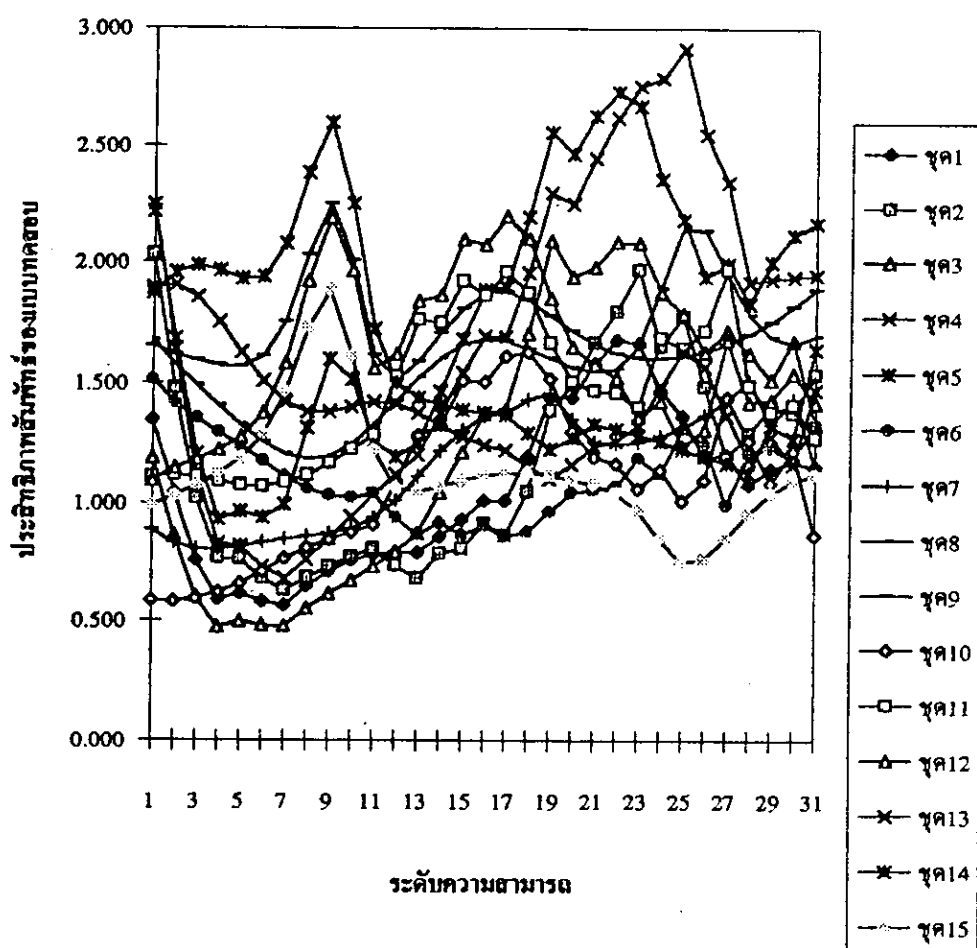


ภาพประกอบ 110 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 400 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

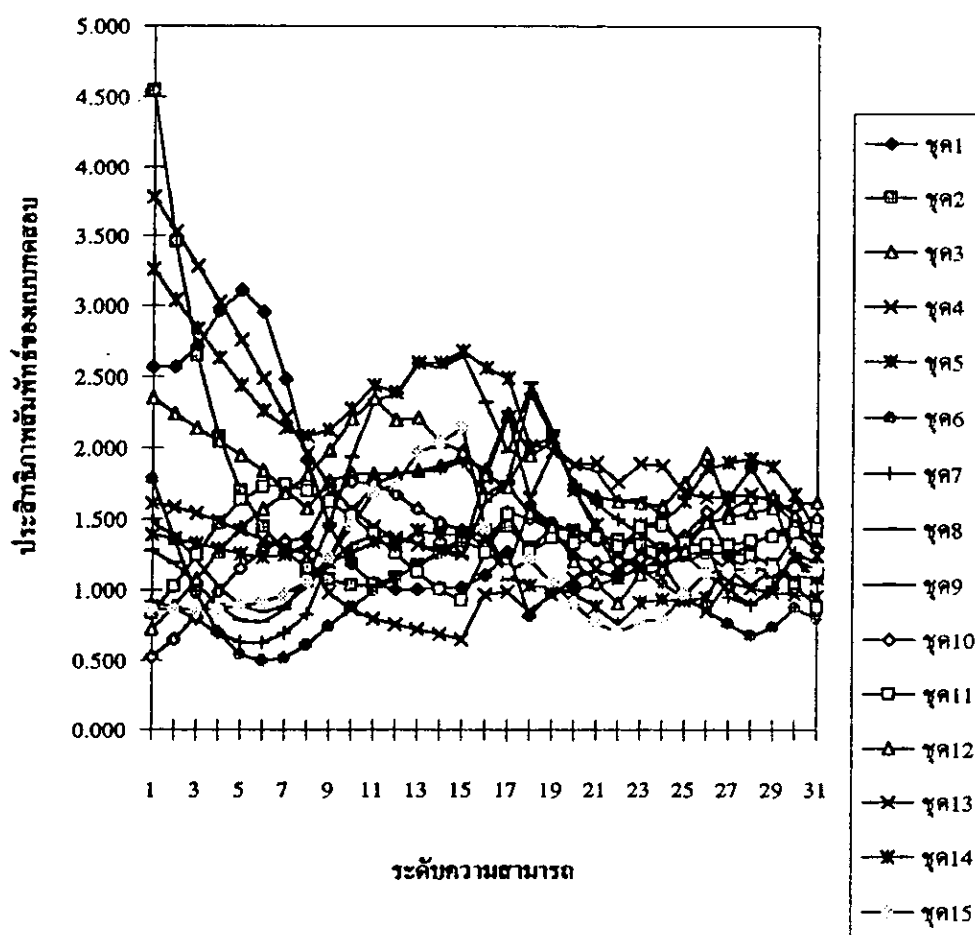


ภาพประกอบ 111 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โคยวี่ของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 500 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

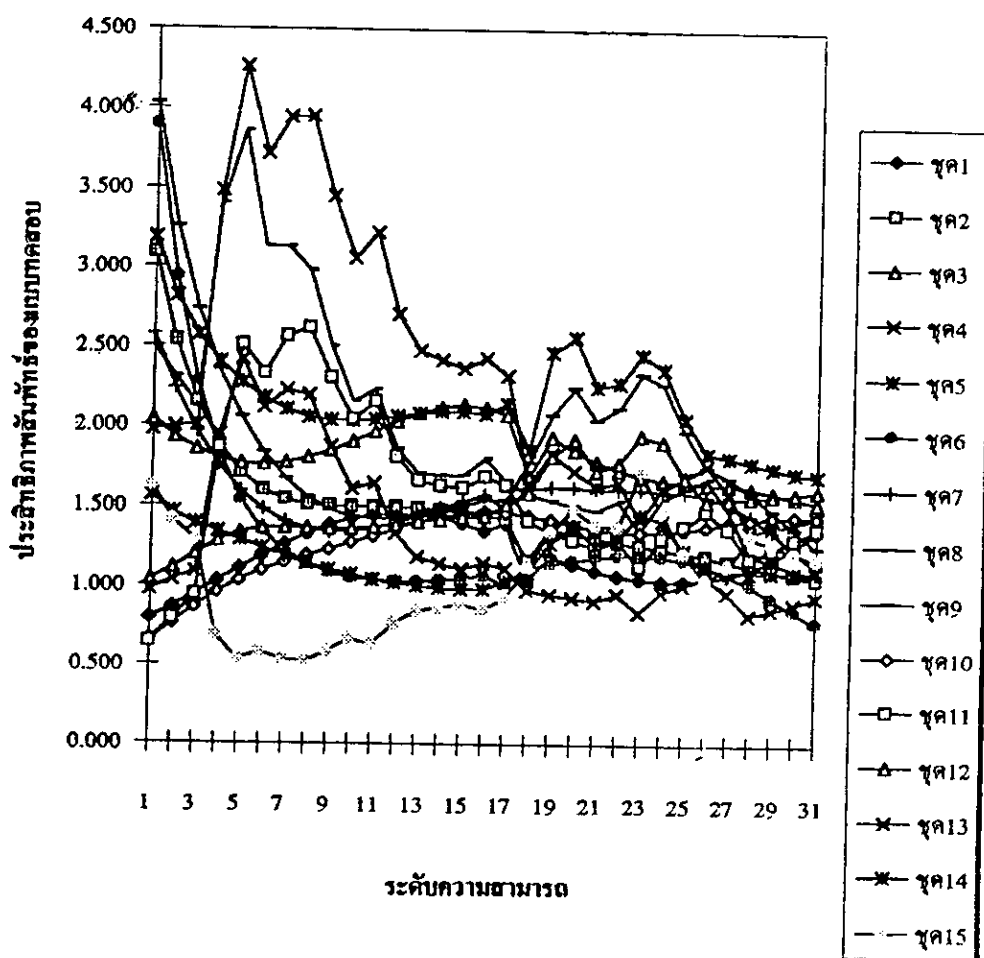


ภาพประกอบ 112 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการ
ประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 600 คน ณ ระดับ
ความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

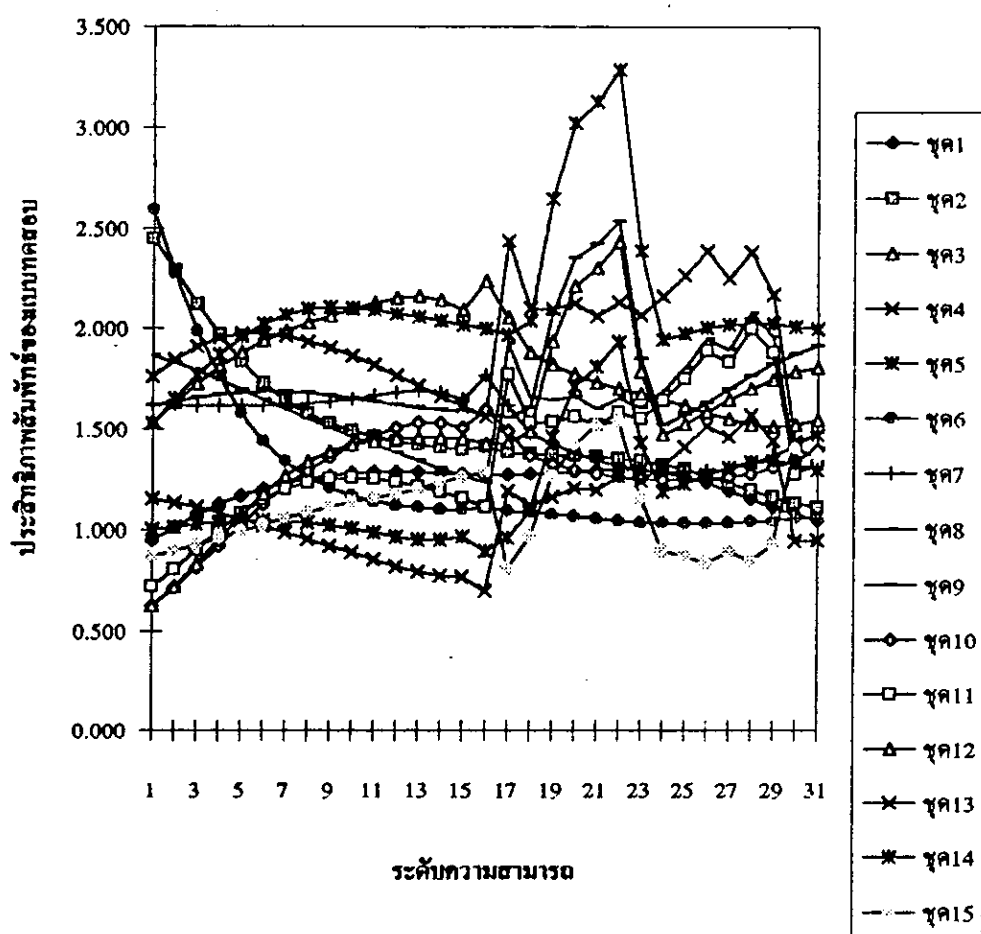
ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 113 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดฮิวีซีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 700 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ	ชุด 9 จำนวน 100/60 ข้อ
ชุด 2 จำนวน 70/50 ข้อ	ชุด 10 จำนวน 80/70 ข้อ
ชุด 3 จำนวน 80/50 ข้อ	ชุด 11 จำนวน 90/70 ข้อ
ชุด 4 จำนวน 90/50 ข้อ	ชุด 12 จำนวน 100/70 ข้อ
ชุด 5 จำนวน 100/50 ข้อ	ชุด 13 จำนวน 90/80 ข้อ
ชุด 6 จำนวน 70/60 ข้อ	ชุด 14 จำนวน 100/80 ข้อ
ชุด 7 จำนวน 80/60 ข้อ	ชุด 15 จำนวน 100/90 ข้อ
ชุด 8 จำนวน 90/60 ข้อ	

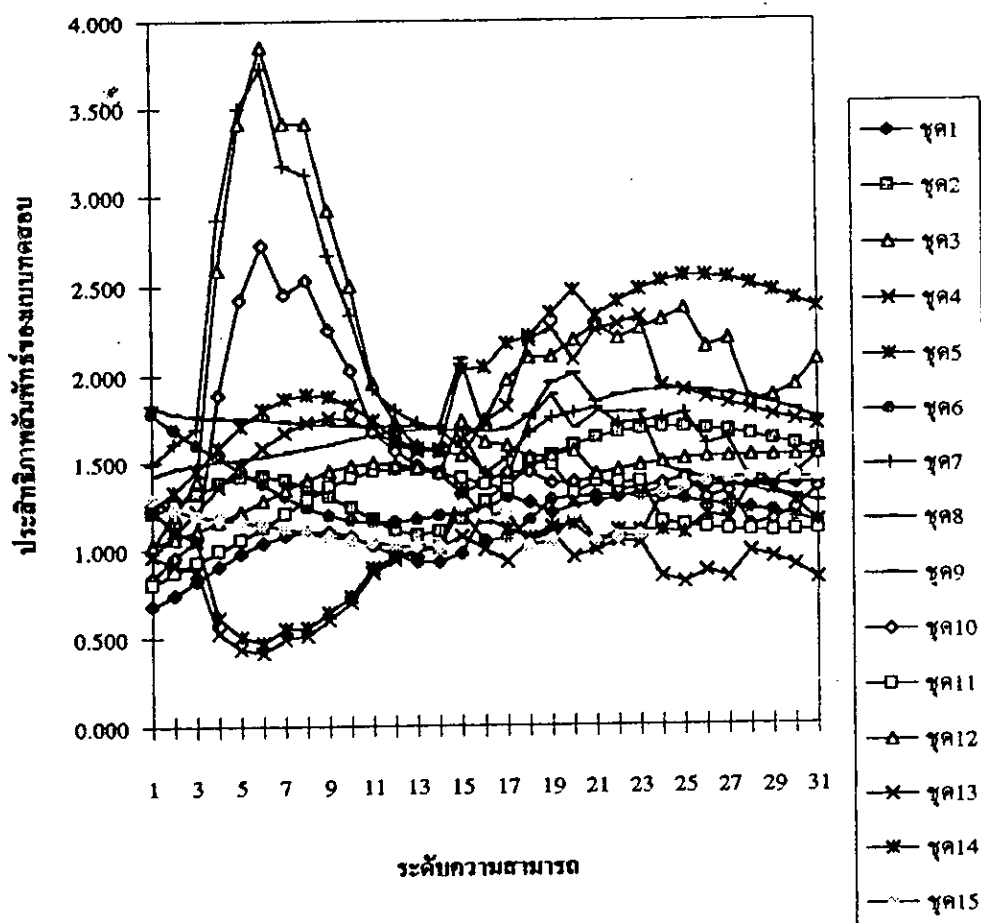


ภาพประกอบ 114 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 800 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ

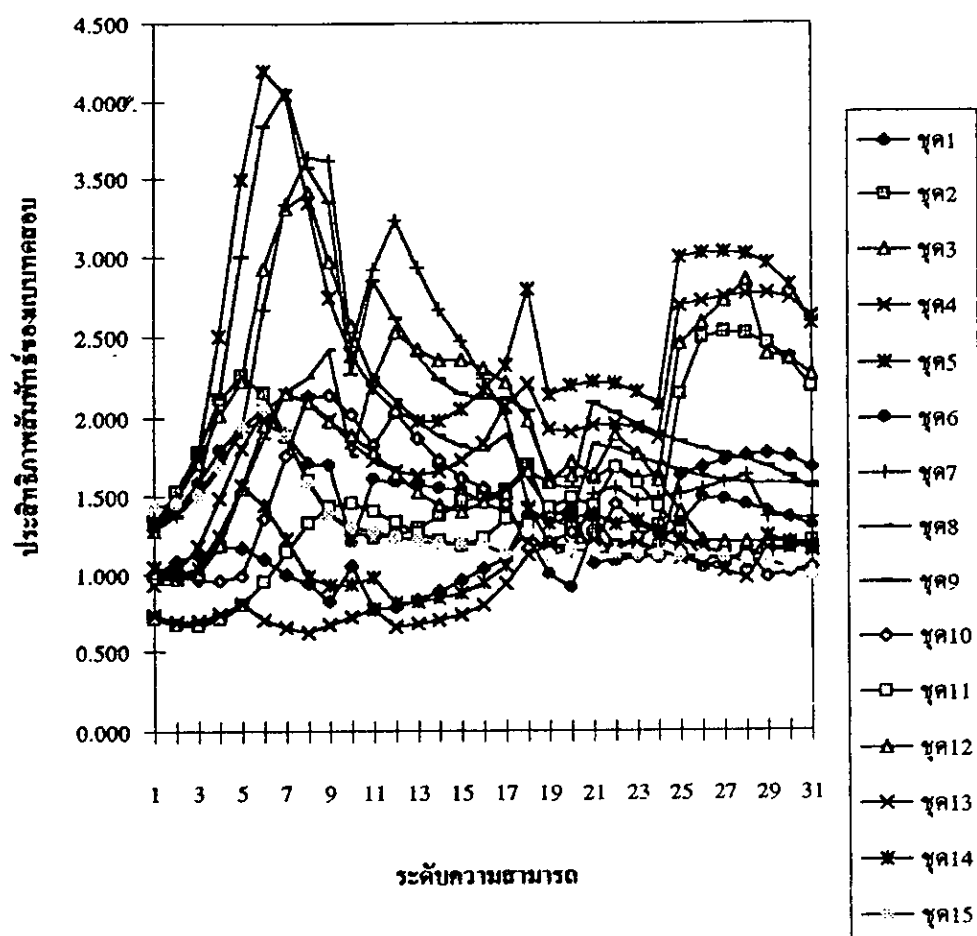


ภาพประกอบ 115 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดฮิวอี้ของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 900 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
ชด 2 จำนวน 70/50 ข้อ
ชด 3 จำนวน 80/50 ข้อ
ชด 4 จำนวน 90/50 ข้อ
ชด 5 จำนวน 100/50 ข้อ
ชด 6 จำนวน 70/60 ข้อ
ชด 7 จำนวน 80/60 ข้อ
ชด 8 จำนวน 90/60 ข้อ

ชด 9 จำนวน 100/60 ข้อ
ชด 10 จำนวน 80/70 ข้อ
ชด 11 จำนวน 90/70 ข้อ
ชด 12 จำนวน 100/70 ข้อ
ชด 13 จำนวน 90/80 ข้อ
ชด 14 จำนวน 100/80 ข้อ
ชด 15 จำนวน 100/90 ข้อ

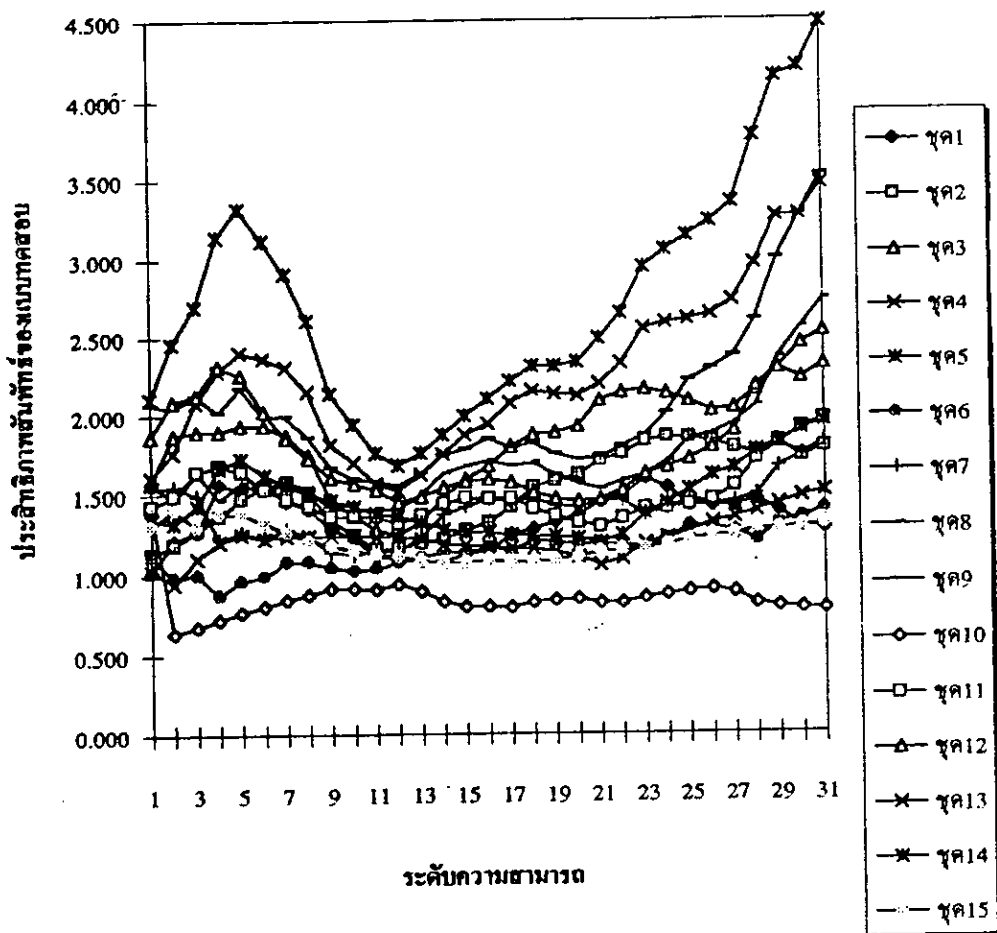


ภาพประกอบ 116 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 1,000 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
ชด 2 จำนวน 70/50 ข้อ
ชด 3 จำนวน 80/50 ข้อ
ชด 4 จำนวน 90/50 ข้อ
ชด 5 จำนวน 100/50 ข้อ
ชด 6 จำนวน 70/60 ข้อ
ชด 7 จำนวน 80/60 ข้อ
ชด 8 จำนวน 90/60 ข้อ

ชด 9 จำนวน 100/60 ข้อ
ชด 10 จำนวน 80/70 ข้อ
ชด 11 จำนวน 90/70 ข้อ
ชด 12 จำนวน 100/70 ข้อ
ชด 13 จำนวน 90/80 ข้อ
ชด 14 จำนวน 100/80 ข้อ
ชด 15 จำนวน 100/90 ข้อ

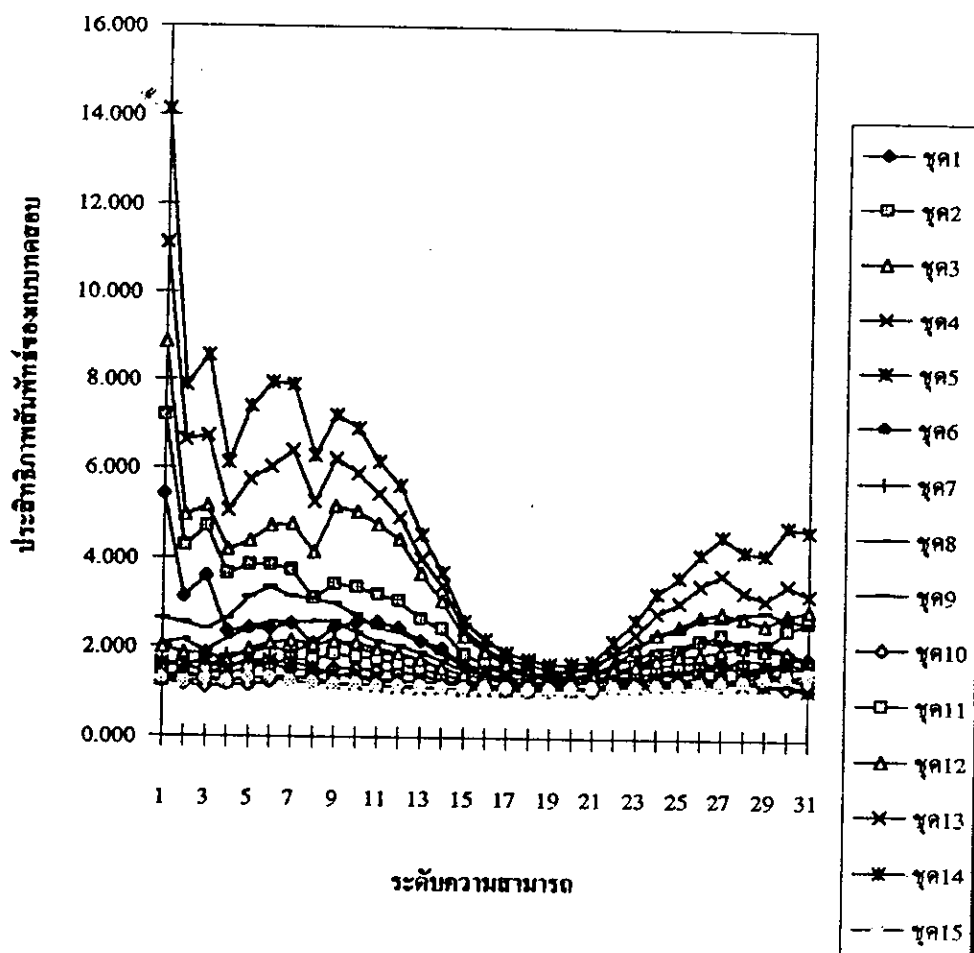


ภาพประกอบ 117 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 1,500 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ
ชุด 2 จำนวน 70/50 ข้อ
ชุด 3 จำนวน 80/50 ข้อ
ชุด 4 จำนวน 90/50 ข้อ
ชุด 5 จำนวน 100/50 ข้อ
ชุด 6 จำนวน 70/60 ข้อ
ชุด 7 จำนวน 80/60 ข้อ
ชุด 8 จำนวน 90/60 ข้อ

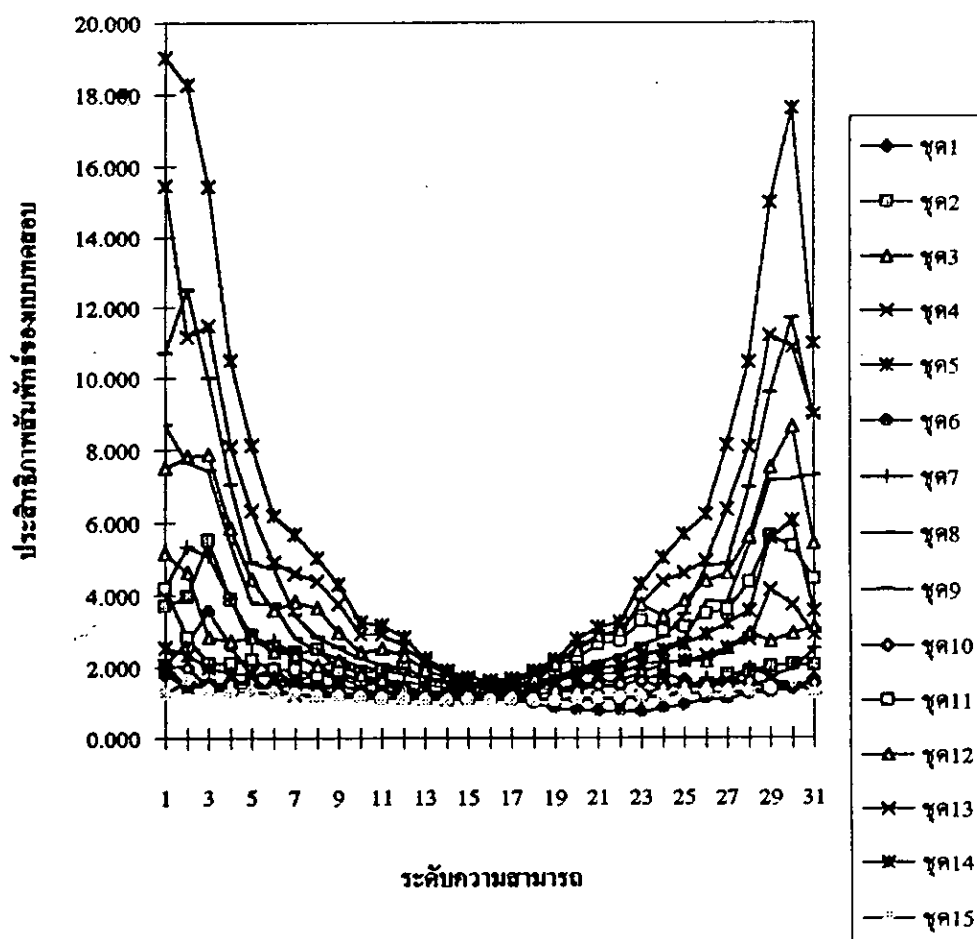
ชุด 9 จำนวน 100/60 ข้อ
ชุด 10 จำนวน 80/70 ข้อ
ชุด 11 จำนวน 90/70 ข้อ
ชุด 12 จำนวน 100/70 ข้อ
ชุด 13 จำนวน 90/80 ข้อ
ชุด 14 จำนวน 100/80 ข้อ
ชุด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 118 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 2,000 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชุด 1 จำนวน 60/50 ข้อ	ชุด 9 จำนวน 100/60 ข้อ
ชุด 2 จำนวน 70/50 ข้อ	ชุด 10 จำนวน 80/70 ข้อ
ชุด 3 จำนวน 80/50 ข้อ	ชุด 11 จำนวน 90/70 ข้อ
ชุด 4 จำนวน 90/50 ข้อ	ชุด 12 จำนวน 100/70 ข้อ
ชุด 5 จำนวน 100/50 ข้อ	ชุด 13 จำนวน 90/80 ข้อ
ชุด 6 จำนวน 70/60 ข้อ	ชุด 14 จำนวน 100/80 ข้อ
ชุด 7 จำนวน 80/60 ข้อ	ชุด 15 จำนวน 100/90 ข้อ
ชุด 8 จำนวน 90/60 ข้อ	

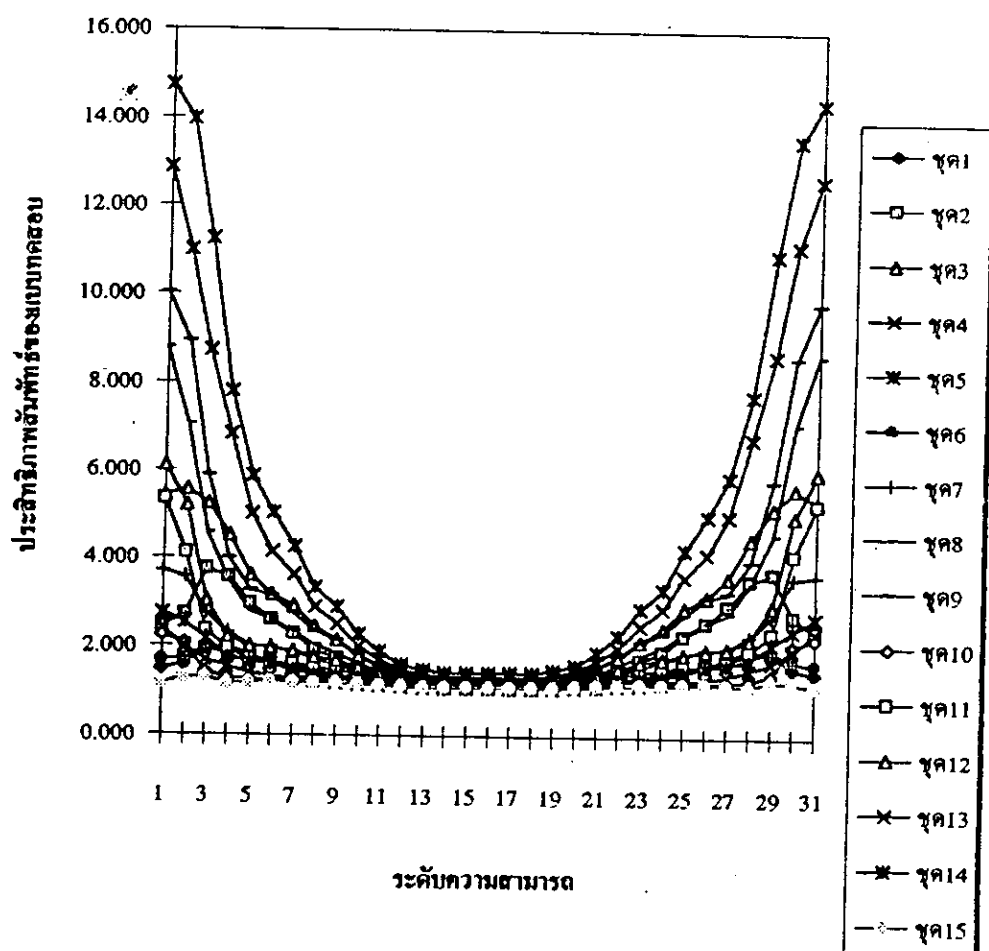


ภาพประกอบ 119 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โคฮิวซีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
 ชด 2 จำนวน 70/50 ข้อ
 ชด 3 จำนวน 80/50 ข้อ
 ชด 4 จำนวน 90/50 ข้อ
 ชด 5 จำนวน 100/50 ข้อ
 ชด 6 จำนวน 70/60 ข้อ
 ชด 7 จำนวน 80/60 ข้อ
 ชด 8 จำนวน 90/60 ข้อ

ชด 9 จำนวน 100/60 ข้อ
 ชด 10 จำนวน 80/70 ข้อ
 ชด 11 จำนวน 90/70 ข้อ
 ชด 12 จำนวน 100/70 ข้อ
 ชด 13 จำนวน 90/80 ข้อ
 ชด 14 จำนวน 100/80 ข้อ
 ชด 15 จำนวน 100/90 ข้อ



ภาพประกอบ 120 เปรียบเทียบโค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์โดยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาด 3,000 คน ณ ระดับความสามารถต่าง ๆ

หมายเหตุ

ชด 1 จำนวน 60/50 ข้อ
ชด 2 จำนวน 70/50 ข้อ
ชด 3 จำนวน 80/50 ข้อ
ชด 4 จำนวน 90/50 ข้อ
ชด 5 จำนวน 100/50 ข้อ
ชด 6 จำนวน 70/60 ข้อ
ชด 7 จำนวน 80/60 ข้อ
ชด 8 จำนวน 90/60 ข้อ

ชด 9 จำนวน 100/60 ข้อ
ชด 10 จำนวน 80/70 ข้อ
ชด 11 จำนวน 90/70 ข้อ
ชด 12 จำนวน 100/70 ข้อ
ชด 13 จำนวน 90/80 ข้อ
ชด 14 จำนวน 100/80 ข้อ
ชด 15 จำนวน 100/90 ข้อ

จากภาพประกอบ 108 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 100 ข้อ กับ 70 ข้อ ของกลุ่มตัวอย่าง 200 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(100, 70) = 7.945 ; e = 3.000\}$

จากภาพประกอบ 109 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 70 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 300 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(70, 50) = 4.938 ; e = -3.000\}$

จากภาพประกอบ 110 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 90 ข้อ กับ 70 ข้อ ของกลุ่มตัวอย่าง 400 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(90, 70) = 3.294 ; e = -3.000\}$

จากภาพประกอบ 111 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 90 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 500 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(90, 50) = 2.913 ; e = 1.800\}$

จากภาพประกอบ 112 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 70 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 600 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(70, 50) = 4.545 ; e = -3.000\}$

จากภาพประกอบ 113 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 90 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 700 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(90, 50) = 4.260 ; e = -2.200\}$

จากภาพประกอบ 114 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 800 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(100, 50) = 3.284 ; e = 1.200\}$

จากภาพประกอบ 115 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 80 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 900 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(80, 50) = 3.845 ; e = -2.000\}$

จากภาพประกอบ 116 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบ-ทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 1,000 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(100, 50) = 4.194 ; e = -2.000\}$

จากภาพประกอบ 117 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 1,500 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถสูง $\{RE(100, 50) = 4.477 ; \sigma = 3.000\}$

จากภาพประกอบ 118 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 2,000 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(100, 50) = 14.124 ; \sigma = -3.000\}$

จากภาพประกอบ 119 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 2,500 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(100, 50) = 19.017 ; \sigma = -3.000\}$

จากภาพประกอบ 120 พบว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 3,000 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(100, 50) = 14.726 ; \sigma = -3.000\}$

ดังนั้น จากภาพประกอบ 108 - 120 แสดงว่า ค่าประสิทธิภาพสัมพัทธ์ระหว่างแบบทดสอบ 100 ข้อ กับ 50 ข้อ ของกลุ่มตัวอย่าง 2,500 คน ที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ $\{RE(100, 50) = 19.017 ; \sigma = -3.000\}$

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

วัตถุประสงค์ในการวิจัย

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซ์ิมัม โลคัลลิสต์ วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยเทคนิควิธีมอนติ คาร์โล โดยมีวัตถุประสงค์เฉพาะของการวิจัย ดังนี้

1. เพื่อศึกษาค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซ์ิมัม โลคัลลิสต์ วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน
2. เพื่อศึกษาค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซ์ิมัม โลคัลลิสต์ วิธีอีวีริสติก และวิธีของเบส์ เมื่อจำนวนข้อสอบต่างกัน
3. เพื่อเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซ์ิมัม โลคัลลิสต์ วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นคะแนนของผู้เข้าสอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี พ.ศ. 2537 วิชาภาษาไทย ของ อ.ก.ค. กรมสามัญศึกษา ซึ่งสุ่มออกเป็น 13 ขนาด คือ ขนาดกลุ่ม 200 คน 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน โดยกลุ่มตัวอย่างแต่ละกลุ่มจะตอบข้อสอบที่มีขนาดต่างกัน 6 ขนาด คือ ขนาดข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ 90 ข้อ และ 100 ข้อ ด้วยการทดลองซ้ำกลุ่มละ 40 ครั้ง ด้วยเทคนิควิธีมอนติ คาร์โล

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบทดสอบวิชาภาษาไทย แบบปรนัยชนิด 4 ตัวเลือก จำนวน 100 ข้อ ซึ่ง อ.ก.ค.กรมสามัญศึกษา เป็นผู้สร้างตามหลักสูตรของ ก.ค. ที่ใช้ในการสอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู

การวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ ผู้วิจัยทำการวิเคราะห์ข้อมูลด้วยเครื่องไมโครคอมพิวเตอร์ โดยใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0 และโปรแกรม BILOG 3 รวมทั้งได้พัฒนาโปรแกรมคอมพิวเตอร์ ภาษา TURBO PASCAL Version 7.0 เพื่อให้สำหรับเทคนิควิธีมอนติ คาร์โล ในการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีเอสดีค และวิธีของเบส์ รวมทั้งใช้ในการคำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบ และค่าสถิติต่าง ๆ ที่ใช้ในการทดสอบสมมติฐาน โดยมีขั้นตอนดังต่อไปนี้

1. การตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ ตามข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม

1.1 การตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว ผู้วิจัยได้ตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว โดยการนำคะแนนผลการสอบของแบบทดสอบวิชาภาษาไทยมาทำการวิเคราะห์องค์ประกอบ โดยใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0 ด้วยวิธีการสกัดองค์ประกอบแบบวิธีองค์ประกอบหลัก (Principal Component) สำหรับการหมุนแกนใช้การหมุนแกนแบบหมุนจากด้วยวิธีนารีแมกซ์ (Varimax)

1.2 การตรวจสอบคุณสมบัติความเป็นอิสระในการการตอบข้อสอบ ให้วิธีการพิจารณาว่า ถ้าแบบทดสอบมีคุณสมบัติการวัดเพียงมิติเดียวแล้ว ก็จะมีคุณสมบัติความเป็นอิสระในการตอบข้อสอบด้วย

2. การสุ่มตัวอย่างแบบใส่คืน (Random Sampling With Replacement) เพื่อสุ่มตัวอย่างผลการสอบของผู้สอบออกเป็น 13 ขนาดกลุ่ม คือ ขนาดกลุ่มตัวอย่าง 200 คน 300 คน 400 คน 500 คน 600 คน 700 คน 800 คน 900 คน 1,000 คน 1,500 คน 2,000 คน 2,500 คน และ 3,000 คน โดยกลุ่มตัวอย่างแต่ละกลุ่มจะต้องตอบข้อสอบที่มีขนาดต่างกัน 6 ขนาด คือ ขนาดข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ 80 ข้อ 90 ข้อ และ 100 ข้อ ซึ่งได้มาด้วยการสุ่มจากตารางวิเคราะห์เนื้อหาที่แบ่งออกเป็น 3 เนื้อหาใหญ่ ๆ โดยการทดลองซ้ำกลุ่มละ 40 ครั้ง ด้วยเทคนิควิธีมอนติ คาร์โล

3. การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบด้วยวิธีอีวีเอสดีค และวิธีของเบส์ โดยผู้วิจัยได้พัฒนาโปรแกรมคอมพิวเตอร์ขึ้นเพื่อใช้ในการวิจัยครั้งนี้ 2 โปรแกรม คือ

3.1 โปรแกรม HEU-IRT เป็นโปรแกรมที่ใช้สำหรับการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบด้วยวิธีอีวีเอสดีค

3.2 โปรแกรม BAY-IRT เป็นโปรแกรมที่ใช้สำหรับการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบด้วยวิธีของเบย์

โดยการพัฒนาโปรแกรมคอมพิวเตอร์ทั้งสองโปรแกรมนี้ มีขั้นตอนในการพัฒนาดังต่อไปนี้

ขั้นที่ 1 เลือกโปรแกรมแปลภาษาของ TURBO PASCAL Version 7.0 ซึ่งเป็นแบบ Compiler

ขั้นที่ 2 การวิเคราะห์ระบบงานที่จะนำมาใช้กับโปรแกรมคอมพิวเตอร์ ซึ่งเป็นขั้นตอนเกี่ยวกับการวิเคราะห์ข้อมูลทางสถิติ

ขั้นที่ 3 การเขียนโปรแกรมเพื่อสร้าง Source Code แล้วนำไป Compiler โดยผู้วิจัยเลือกใช้ Editor ของโปรแกรม TURBO PASCAL Version 7.0

ขั้นที่ 4 นำโปรแกรมที่เขียนไปใช้กับข้อมูลทดสอบ เพื่อตรวจสอบความถูกต้องของการประมวลผล หากพบข้อผิดพลาดจะต้องดำเนินการแก้ไขให้ถูกต้องก่อนนำไปใช้จริง

ขั้นที่ 5 นำโปรแกรมที่พัฒนาแล้วไปใช้กับข้อมูลจริง

4. การประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบจากผลการสอบที่มีจำนวนผู้สอบและจำนวนข้อสอบแตกต่างกัน 3 วิธี คือ วิธีแมกซีมัม ไลค์ลิสต์ โดยใช้โปรแกรมสำเร็จรูป BILOG 3 ของมิสเลวี และบ็อค (Mislevy and Bock. 1990 : 6 - 8) วิธีอีวริสติก และวิธีของเบย์ โดยใช้โปรแกรมคอมพิวเตอร์ภาษา TURBO PASCAL Version 7.0 ที่ผู้วิจัยเป็นผู้พัฒนา

5. เปรียบเทียบความแตกต่างของค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบ ซึ่งประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบย์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน โดยใช้สถิติทดสอบแบบ Repeated Measure (AxS) Design ถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติแล้ว ทดสอบความแตกต่างรายคู่ด้วยวิธีการของเชฟเฟ (Scheffe's Test) โดยใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version 6.0

6. เปรียบเทียบความแตกต่างของค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบ ซึ่งประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบย์ เมื่อจำนวนข้อสอบต่างกัน โดยใช้สถิติทดสอบแบบ Repeated Measure (AxS) Design ถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติแล้ว ทดสอบความแตกต่างรายคู่ด้วยวิธีการของเชฟเฟ (Scheffe's Test) โดยใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS Version

6.0

7. คำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบที่ได้จากวิธีการประมาณค่าพารามิเตอร์แต่ละวิธี ณ ระดับความสามารถต่าง ๆ 31 ระดับ คือ ตั้งแต่ระดับความสามารถ -3.0, -2.8, -2.6, ..., 2.8, 3.0 .

8. คำนวณค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่ได้จากวิธีการประมาณค่าพารามิเตอร์แต่ละวิธี เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ (θ_j) โดยนำเอาค่าฟังก์ชันสารสนเทศของแบบทดสอบที่ได้จากการประมาณค่าแต่ละวิธี เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ณ ระดับความสามารถต่าง ๆ มาคำนวณ โดยใช้สูตรในการคำนวณ ดังนี้

$$RE(y, x) = \frac{I(\theta, y)}{I(\theta, x)}$$

9. ทดสอบค่าประสิทธิภาพสัมพัทธ์ของวิธีประมาณค่าพารามิเตอร์ ดังนี้

เมื่อ $RE(y, x) = 1$ แสดงว่า แบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีคุณภาพเท่ากัน

$RE(y, x) > 1$ แสดงว่า แบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีคุณภาพไม่เท่ากัน

$RE(y, x) < 1$ แสดงว่า แบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิฮูด วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีคุณภาพไม่เท่ากัน

สรุปผลการวิเคราะห์ข้อมูล

1. ผลการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ พบว่า เมื่อนำแบบทดสอบจำนวน 100 ข้อ มาวิเคราะห์องค์ประกอบ จะมีองค์ประกอบที่มีค่าไอเกนเกิน 1.00 อยู่ 32 องค์ประกอบ โดยองค์ประกอบที่ 1 มีค่าไอเกนสูงสุด เท่ากับ 5.786 องค์ประกอบที่ 2 มีค่าไอเกน เท่ากับ 1.887 และองค์ประกอบถัดไปมีค่าไอเกนลดลงเรื่อย ๆ แสดงว่า แบบทดสอบมีการวัดเพียงมิติเดียว โดยองค์ประกอบที่มีค่าไอเกนสูงสุด เป็นองค์ประกอบที่แทนความสามารถ หรือคุณลักษณะที่เด่นหรือสำคัญ

(Dominant Factor) เพื่อสังเกตเฉพาะตัวที่เป็นตัวกำหนดพฤติกรรมการตอบข้อสอบ และเมื่อแบบทดสอบมีคุณสมบัติการวัดเพียงมิติเดียวแล้ว แสดงว่า แบบทดสอบมีคุณสมบัติความเป็นอิสระในการตอบข้อสอบด้วย

2. ค่าสถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ วิธีอีวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน พบว่า

2.1 ค่าอำนาจจำแนกเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงกว่าค่าอำนาจจำแนกเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ และวิธีอีวริสติก เมื่อกลุ่มตัวอย่างมีขนาดเล็กและข้อสอบมีจำนวนน้อยข้อ ในขณะที่ค่าความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ มีค่าต่ำกว่าค่าความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ และวิธีอีวริสติก เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ และข้อสอบมีจำนวนมากข้อ

นอกจากนี้ ในการประมาณค่าอำนาจจำแนกของข้อสอบด้วยวิธีอีวริสติก และวิธีของเบส์ ไม่พบว่า มีข้อสอบที่ไม่เหมาะสมกับโมเดลจากการประมาณค่าทั้ง 2 วิธี แต่ถ้าประมาณค่าอำนาจจำแนกของข้อสอบด้วยวิธีแมกซีมัม ไลค์ลิสต์ แล้ว พบว่า มีข้อสอบที่ไม่เหมาะสมกับโมเดล ดังนั้น เมื่อต้องการเปรียบเทียบความแตกต่างของค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ แล้ว จึงจำเป็นต้องตัดข้อสอบเหล่านี้ออกเมื่อต้องการเปรียบเทียบค่าอำนาจจำแนกของข้อสอบ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

2.2 ค่าความยากเฉลี่ยของข้อสอบ และค่าความเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอีวริสติก มีค่าสูงกว่าค่าความยากเฉลี่ยของข้อสอบ และค่าความเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบ ที่ประมาณค่าด้วยวิธีแมกซีมัม ไลค์ลิสต์ และวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาดเล็ก และข้อสอบมีจำนวนน้อยข้อ

ในการประมาณค่าความยากของข้อสอบด้วยวิธีของเบส์ นี้ ไม่พบว่ามีข้อสอบที่ไม่เหมาะสมกับโมเดล แต่ในการประมาณค่าความยากของข้อสอบด้วยวิธีอีวริสติก และวิธีแมกซีมัม ไลค์ลิสต์ นี้ พบว่า มีข้อสอบที่ไม่เหมาะสมกับโมเดล จึงจำเป็นต้องตัดข้อสอบเหล่านี้ออก เมื่อต้องการเปรียบเทียบความแตกต่างของค่าความยากของข้อสอบ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน

2.3 ค่าการเคาะเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีวิธีวิริสติก เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ และข้อสอบมีจำนวนมากข้อ มีค่าต่ำกว่าค่าการเคาะเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ และวิธีแมกซิมัม โคล์ลิสต์ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ แต่ข้อสอบมีจำนวนน้อยข้อ ในขณะที่ค่าความเบี่ยงเบนมาตรฐานของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีวิริสติก เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ และข้อสอบมีจำนวนมากข้อ มีค่าสูงกว่าค่าความเบี่ยงเบนมาตรฐานของค่าการเคาะของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ และวิธีแมกซิมัม โคล์ลิสต์ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ และข้อสอบมีจำนวนน้อยข้อ

2.4 ค่าความสามารถเฉลี่ยของผู้สอบ และค่าความเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีวิริสติก มีค่าสูงกว่าค่าความสามารถเฉลี่ยของผู้สอบ และค่าความเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ และวิธีของเบส์ ในทุกขนาดกลุ่มตัวอย่าง

3. ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าพารามิเตอร์ของข้อสอบ และความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ วิธีวิริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน พบว่า

3.1 ค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.2 ค่าอำนาจจำแนกของข้อสอบ 80 ข้อ ของกลุ่มตัวอย่างที่มีขนาดเล็ก ที่ประมาณค่าด้วยวิธีวิริสติก มีความแตกต่างจากกลุ่มตัวอย่างขนาดใหญ่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.3 ค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาดเล็ก มีความแตกต่างจากกลุ่มตัวอย่างขนาดใหญ่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.4 ค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ กับวิธีวิริสติก เมื่อกลุ่มตัวอย่างมีขนาดเล็ก มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในขณะที่ค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ กับวิธีของเบส์ วิธีวิริสติก กับวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.5 ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีวิริสติก และค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธี

ของเบส เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไรไม่มีนัยสำคัญทางสถิติ

3.6 ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม โคลด์ลิสต์ กับวิธีอิวิริสติก เมื่อข้อสอบมีจำนวนน้อยข้อ และกลุ่มตัวอย่างมีขนาดใหญ่ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซีมัม โคลด์ลิสต์ กับวิธีของเบส์ เมื่อข้อสอบมีจำนวนมากข้อ และกลุ่มตัวอย่างมีขนาดใหญ่ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก กับวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.7 ค่าการเผาของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิฮูด ค่าการเผาของข้อสอบที่ประมาณค่าด้วยวิธีวิธีวิธีสถิติ และค่าการเผาของข้อสอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาดเล็ก มีความแตกต่างจากค่าการเผาของข้อสอบ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.8 ค่าการเดาของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ กับวิธี
อิวิริสติก และค่าการเดาของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก กับวิธีของเบส์ เมื่อกลุ่ม
ตัวอย่างมีขนาดเล็ก มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับค่าการ
เดาของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ กับวิธีของเบส์ เมื่อกลุ่มตัวอย่างมี
ขนาดปานกลางและขนาดใหญ่ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.9 ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ และค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไรมีนัยสำคัญทางสถิติ แต่ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอาร์ลิสต์ เมื่อกลุ่มตัวอย่างมีขนาดเล็ก มีความแตกต่างจากกลุ่มตัวอย่างที่มีขนาดใหญ่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3.10 ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ กับวิธีฮาร์วาร์ด และวิธีฮาร์วาร์ด กับวิธีของเบส เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แต่ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ กับวิธีของเบส เมื่อขนาดของกลุ่มตัวอย่างต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. ผลการเปรียบเทียบและทดสอบความแตกต่างของค่าพารามิเตอร์ของข้อสอบ
และความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวริสติก และวิธีของ

เบส เมื่อจำนวนข้อสอบต่างกัน พบว่า

4.1 ค่าอำนาจจำแนกเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ ค่าอำนาจจำแนกเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก และค่าอำนาจจำแนกเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ยกเว้น เมื่อกลุ่มตัวอย่างมีขนาด 2,500 คน ค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 50 ข้อ 60 ข้อ 70 ข้อ และ 80 ข้อ ที่ประมาณค่าด้วยวิธีของเบส จะมีความแตกต่างจากค่าอำนาจจำแนกเฉลี่ยของข้อสอบ 90 ข้อ และ 100 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.2 ค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ กับวิธีอิวิริสติก เมื่อจำนวนข้อสอบต่างกัน ส่วนใหญ่มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แต่ค่าอำนาจจำแนกของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ กับวิธีของเบส และวิธีอิวิริสติก กับวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.3 ค่าความยากเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ ค่าความยากเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก และค่าความยากเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4.4 ค่าความยากของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ กับวิธีอิวิริสติก วิธีแมกซิมัม โคล์ลิสต์ กับวิธีของเบส และวิธีอิวิริสติก กับวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน ส่วนใหญ่แล้วมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.5 ค่าการเดาเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ ค่าการเดาเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีอิวิริสติก และค่าการเดาเฉลี่ยของข้อสอบที่ประมาณค่าด้วยวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน ส่วนใหญ่แล้วมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.6 ค่าการเดาของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ กับวิธีอิวิริสติก วิธีแมกซิมัม โคล์ลิสต์ กับวิธีของเบส และวิธีอิวิริสติก กับวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.7 ค่าความสามารถเฉลี่ยของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม โคล์ลิสต์ และค่าความสามารถเฉลี่ยของผู้สอบที่ประมาณค่าด้วยวิธีของเบส เมื่อจำนวนข้อสอบต่างกัน

7. เมื่อขนาดของกลุ่มตัวอย่างกัน ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีวีส์ติค มีค่าสูงสุด ที่ระดับความสามารถต่ำ รองลงมา คือ ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลียูด ที่ระดับความสามารถสูง และค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลจากการประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ ที่ระดับความสามารถต่ำ ตามลำดับ

8. เมื่อจำนวนข้อสอบต่างกัน พบว่า ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่ประมาณค่าด้วยวิธีของเบส์ มีค่าสูงสุด ที่ระดับความสามารถต่ำ รองลงมาคือ ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่ประมาณค่าด้วยวิธีอีวีวีส์ติค ที่ระดับความสามารถต่ำ และค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลียูด ที่ระดับความสามารถสูง ตามลำดับ

อภิปรายผล

1. จากผลการศึกษา ที่พบว่า ค่าพารามิเตอร์ของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลียูด วิธีอีวีวีส์ติค และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีความแตกต่างกันนั้น เนื่องจากวิธีประมาณค่าทั้งสามวิธีนี้มีแนวคิดพื้นฐานทางสถิติที่แตกต่างกัน โดยวิธีแมกซิมัม ไลค์ลียูด และวิธีอีวีวีส์ติค ใช้สถิตินอนเบส์เชียน (Non Bayesian Statistics) ซึ่งแนวคิดนี้จะพิจารณาค่าพารามิเตอร์ว่าเป็นค่าคงที่ ที่ไม่ทราบค่า และต้องการประมาณ ในขณะที่วิธีของเบส์ใช้สถิติแบบเบส์เชียน (Bayesian Statistics) โดยมีการกำหนดการแจกแจงความน่าจะเป็นของค่าพารามิเตอร์ที่เรียกว่า การแจกแจงเริ่มแรก (Prior Distribution) ของค่าอำนาจจำแนก และค่าการเดา (สุชาดา กิระนันท์. 2525 : 26 - 27) จึงทำให้วิธีการประมาณค่าพารามิเตอร์ด้วยแมกซิมัม ไลค์ลียูด และวิธีอีวีวีส์ติค ไม่สามารถประมาณค่าอำนาจจำแนก และค่าการเดาได้ดีพอ เพราะไม่มีการกำหนดขอบเขตจำกัดของช่วงค่าอำนาจจำแนก และค่าการเดา โดยเฉพาะวิธีอีวีวีส์ติคที่ประมาณค่าพารามิเตอร์ของข้อสอบจากค่าสถิติความยากและค่าอำนาจจำแนก ที่ได้จากการวิเคราะห์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม และการนำวิธีอีวีวีส์ติคไปใช้อย่างมีประสิทธิภาพนั้น แบบทดสอบควรประกอบด้วยจำนวนข้อสอบอย่างน้อย 80 ข้อ และมีความเชื่อมั่น (KR-20) ประมาณ 0.90 (Schmidt. 1977 : 613 - 620) แต่ความเชื่อมั่นของแบบทดสอบที่ใช้ในการวิจัยครั้งนี้มีค่าไม่ถึง 0.90 จึงทำให้การประมาณค่าพารามิเตอร์อาจมีความคลาดเคลื่อนเกิดขึ้นได้ ซึ่งทำให้การประมาณค่าความสามารถของ

ผู้สอบด้วยวิธีแมกซิมัม ไลค์ลิสต์ และวิธีอิวิริสติก ต้องตัดข้อที่มีค่าอำนาจจำแนกและค่าความยากสูงหรือต่ำเกินไป ที่อยู่ภายนอกขอบเขตที่ยอมรับได้ออก เมื่อต้องการประมาณค่าความสามารถของผู้สอบ จึงอาจทำให้เกิดปัญหาเกี่ยวกับความเที่ยงตรงของค่าที่ประมาณได้ (Hambleton and Swaminathan. 1985 : 132) ในขณะที่การประมาณค่าด้วยวิธีของเบส์นั้นไม่พบปัญหาดังกล่าว (Swaminathan and Gifford. 1986 : 589 - 601)

2. จากผลการศึกษาที่พบว่า ค่าความสามารถเฉลี่ยของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม ไลค์ลิสต์ และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ไม่มีความแตกต่างกันนั้น เนื่องจากในการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิสต์ ขึ้นแต่ละครั้งนั้น ค่าความสามารถจะถูกปรับให้มีค่าเฉลี่ยเท่ากับ 0 และค่าความเบี่ยงเบนมาตรฐานเป็น 1 ซึ่งจะทำให้ค่าอำนาจจำแนก ค่าความยาก และค่าการเดา ต้องถูกปรับตามไปด้วย (Hambleton and Swaminathan. 1985 : 127 - 138) และวิธีของเบส์ก็ได้กำหนดไว้ว่า ค่าความสามารถเป็นตัวแปรสุ่มที่มีการแจกแจงเป็นปกติ (Swaminathan and Gifford. 1985 : 35) ในขณะที่ค่าความสามารถเฉลี่ยของผู้สอบที่ประมาณค่าด้วยวิธีอิวิริสติก เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีความแตกต่างกันนั้น แม้ว่าจะใช้วิธีแมกซิมัม ไลค์ลิสต์ แบบมีเงื่อนไข ก็ตาม (Hambleton and Swaminathan. 1985 : 613 - 620) เพราะค่าพารามิเตอร์ความสามารถของผู้สอบสามารถคำนวณมาจากค่าพารามิเตอร์ของข้อสอบที่ได้จากทฤษฎีการทดสอบแบบดั้งเดิม จึงอาจทำให้ค่าพารามิเตอร์ที่ได้มีความแตกต่างจากค่าพารามิเตอร์ที่ได้จากวิธีแมกซิมัม ไลค์ลิสต์ และวิธีของเบส์ ที่ทำการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ด้วยกระบวนการหาค่าซ้ำ ๆ จนกว่าจะได้ค่าคงที่

ดังนั้น เมื่อค่าพารามิเตอร์ที่ได้จากการประมาณค่าทั้ง 3 วิธี มีความแตกต่างกันย่อมส่งผลต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบ และค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ณ ระดับความสามารถต่าง ๆ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ให้มีความแตกต่างกันด้วย

2. ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม ไลค์ลิสต์ มีคุณภาพสูงกว่าวิธีอิวิริสติก และวิธีของเบส์ ที่ระดับความสามารถสูง เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ในขณะที่ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอิวิริสติก และวิธีของเบส์ มีคุณภาพสูงกว่าวิธีแมกซิมัม ไลค์ลิสต์ ที่ระดับความสามารถต่ำ เมื่อขนาดของกลุ่ม-

ตัวอย่างและจำนวนข้อสอบต่างกัน น่าจะมีสาเหตุมาจากการคำนวณหาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบนั้น ขึ้นอยู่กับค่าฟังก์ชันสารสนเทศของแบบทดสอบที่นำมาเปรียบเทียบกัน ซึ่งค่าฟังก์ชันสารสนเทศของแบบทดสอบนี้ เป็นผลรวมของค่าฟังก์ชันสารสนเทศของข้อสอบที่เป็นดัชนีกำหนดคุณภาพของข้อสอบที่ได้จากค่าพารามิเตอร์ทุกตัวในแบบจำลอง ดังนั้น ค่าฟังก์ชันสารสนเทศของแบบทดสอบที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ข้อสอบแต่ละวิธี ย่อมมีความแตกต่างกันด้วย ทั้งนี้ขึ้นอยู่กับแนวคิด กระบวนการ และผลการประมาณค่าพารามิเตอร์แต่ละวิธี ซึ่งสวามินาธาน และกิฟฟอร์ด (Swaminathan and Gifford, 1986 : 589 - 601) กล่าวว่า เมื่อทำการเปรียบเทียบผลการประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัมไลค์ลิฮูด และวิธีของเบส์ นั้น วิธีแมกซิมัมไลค์ลิฮูด ไม่สามารถประมาณค่าพารามิเตอร์ให้คงที่ได้ หรือประมาณค่าได้คงที่แต่ไม่ตกอยู่ในขอบเขตของค่าพารามิเตอร์ที่จะยอมรับได้ ซึ่งในสถานการณ์เช่นนี้ การคำนวณโดยผู้ใช้โปรแกรม BILOG 3 จึงจำเป็นต้องตัดค่าที่อยู่นอกขอบเขตออก ทำให้ค่าพารามิเตอร์เหล่านี้ส่งผลกระทบต่อแบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิฮูด เพราะทำให้ฟังก์ชันสารสนเทศของแบบทดสอบมีค่าสูงสุด เมื่อกลุ่มตัวอย่างมีความสามารถสูง ในขณะที่แบบทดสอบที่ประมาณค่าด้วยวิธีของเบส์ และวิธีอีวีริสติก จะมีค่าฟังก์ชันสารสนเทศสูงสุด ที่ระดับความสามารถปานกลาง (วิชุดา บัวคง, 2533 : 155)

นอกจากนี้ การที่พบว่าแบบทดสอบที่ทำการประมาณค่าด้วยวิธีอีวีริสติก และวิธีของเบส์ มีคุณภาพสูงกว่าแบบทดสอบที่ประมาณค่าด้วยวิธีแมกซิมัมไลค์ลิฮูด เฉพาะในกลุ่มตัวอย่างที่มีระดับความสามารถต่ำนั้น อาจเนื่องมาจากการประมาณค่าพารามิเตอร์ด้วยวิธีอีวีริสติก และวิธีของเบส์ นั้น ข้อสอบมีค่าพิสัยของค่าอำนาจจำแนกและค่าความยากสูงเกินจริง และไม่ได้ถูกตัดออก ดังนั้น เมื่อนำมาหาค่าฟังก์ชันสารสนเทศของแบบทดสอบวาร์ม (Warm, 1978 : 70) จึงกล่าวว่า ถ้าข้อสอบมีค่าอำนาจจำแนกสูงหรือต่ำ ก็จะมีผลทำให้ค่าฟังก์ชันสารสนเทศสูงหรือต่ำเช่นกัน ประกอบกับ แฮมเบิลตัน และสวามินาธาน (Hambleton and Swaminathan, 1985 : 106 - 111) ได้กล่าวว่า ถ้าข้อสอบมีความยากอยู่ในระดับใดแล้ว จะมีค่าฟังก์ชันสารสนเทศสูงสุดในระดับความสามารถนั้น ๆ ด้วย ดังนั้น ข้อสอบที่ค่อนข้างยากถึงยากมาก จึงมีค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงมากในกลุ่มตัวอย่างที่มีระดับความสามารถสูง ในขณะที่ข้อสอบที่ค่อนข้างง่ายถึงง่ายมาก น่าจะมีค่าฟังก์ชันสารสนเทศของแบบทดสอบสูง ในกลุ่มตัวอย่างที่มีความสามารถต่ำด้วย แต่ก็ไม่เป็นเช่นนั้น เพราะอิทธิพลของค่าอำนาจจำแนกของข้อสอบที่ต่ำเกินจริง นั้นเอง ดังนั้น เมื่อนำค่าฟังก์ชันสารสนเทศของแบบทดสอบทั้งฉบับ ที่เกิดจากผลรวมของค่าฟังก์ชันสารสนเทศของ

ข้อสอบทุกข้อ ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ของข้อสอบด้วยวิธีแมกซิมัม ไลค์ลิสต์ มาเปรียบเทียบกับวิธีอีวีริสต์ติค และวิธีของเบส์ ที่มีค่าฟังก์ชันสารสนเทศของแบบทดสอบสูง ในกลุ่มตัวอย่างที่มีระดับความสามารถปานกลางถึงต่ำแล้ว จึงพบว่า ในกลุ่มตัวอย่างที่มีระดับ ความสามารถต่ำ แบบทดสอบที่ประมาณค่าด้วยวิธีอีวีริสต์ติค และวิธีของเบส์ จะมีคุณภาพสูง กว่าวิธีแมกซิมัม ไลค์ลิสต์

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ในการหาคุณภาพของแบบทดสอบนั้น ผู้วิจัยเสนอแนะให้ผู้ต้องการหาคุณภาพ ของแบบทดสอบ ดำเนินการหาคุณภาพของแบบทดสอบเป็นรายข้อตามทฤษฎีการทดสอบแบบ ดั้งเดิมก่อน หากพบข้อบกพร่องให้ดำเนินการปรับปรุงข้อคำถามให้เรียบร้อยก่อน แล้วจึงนำ แบบทดสอบที่ผ่านการวิเคราะห์แล้ว มาหาคุณภาพตามทฤษฎีการตอบข้อคำถาม เพื่อประโยชน์ ในการใช้ข้อต่อไป เป็นต้นว่า งานคลังข้อสอบ

2. สิ่งที่ต้องคำนึงถึงในการเลือกวิธีประมาณค่าพารามิเตอร์ของข้อสอบ และ ความสามารถของผู้สอบ ก็คือ ผู้ที่ต้องการประมาณค่าพารามิเตอร์ ต้องรู้จักเลือกใช้วิธีประมาณค่า แต่ละวิธีได้อย่างถูกต้อง โดยคำนึงถึงข้อตกลงเบื้องต้นของวิธีการประมาณค่าพารามิเตอร์แต่ละวิธีเป็นสำคัญ เพื่อให้ผลการวิเคราะห์ที่ได้มีความถูกต้อง

3. ในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบนั้น สามารถเลือกใช้ วิธีประมาณค่าพารามิเตอร์ได้ทั้ง 3 วิธี ทั้งนี้ขึ้นอยู่กับหน้าที่ของการสอบแต่ละครั้ง เช่น ใน การสอบแข่งขัน ควรเลือกใช้วิธีแมกซิมัม ไลค์ลิสต์ เพราะสามารถประมาณค่าความสามารถ ของผู้สอบที่มีระดับความสามารถสูง

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. ควรมีการเปรียบเทียบความแปรปรวนของวิธีประมาณค่าพารามิเตอร์แต่ละวิธี พร้อมทั้งมีการตรวจสอบความแกร่งของกระบวนการประมาณค่าพารามิเตอร์ เมื่อนำไปใช้ใน สถานการณ์ต่าง ๆ ทั้งที่เป็นไปและไม่เป็นไปตามข้อตกลงเบื้องต้นของวิธีการประมาณค่า พารามิเตอร์แต่ละวิธี

2. ความรู้การศึกษาความคลาดเคลื่อนมาตรฐานของแบบทดสอบ ที่มีขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยวิธีการประมาณค่าพารามิเตอร์ระหว่างวิธีแมกซิมัมไลค์ลิฮูด กับวิธีอีวีริสติก วิธีแมกซิมัมไลค์ลิฮูด กับวิธีของเบส์ และวิธีอีวีริสติก กับวิธีของเบส์ หรืออาจใช้วิธีประมาณค่าพารามิเตอร์วิธีเดียวกัน โดยใช้เทคนิควิชันอนติ คาร์โล

3. ความรู้การศึกษาวิจัยในทำนองนี้อีก แต่เลือกเฉพาะกลุ่มตัวอย่างที่มีระดับความสามารถเหมือนกัน เพื่อให้ได้ข้อมูลเพิ่มเติมในเรื่องนี้ อันจะเป็นแนวทางในการพัฒนาวิธีการประมาณค่าพารามิเตอร์ในทฤษฎีการตอบข้อคำถามให้มีประสิทธิภาพยิ่งขึ้น

ข้อเสนอแนะเกี่ยวกับการใช้เทคนิคมอนติ คาร์โล

ในการประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ เค็มที่ผู้วิจัยกำหนดจำนวนการทดลองซ้ำไว้ 40 ครั้ง เป็นจำนวนขั้นสูงสุดของการทำซ้ำ ที่กำหนดขึ้นนั้น อาจไม่จำเป็นเพราะถ้าผลต่างของการประมาณค่าพารามิเตอร์ ครั้งที่ ๓ และ ๓+1 มีค่าน้อยกว่าหรือเท่ากับ .001 แล้ว ก็สามารถยุติการทำซ้ำได้ เพื่อเป็นการประหยัดเวลาและพื้นที่ของเครื่องคำนวณ เพราะค่าที่ประมาณได้นั้นคงที่แล้ว

מדורגות

บรรณานุกรม

- คำย เชื้องฉ. การศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของ
ผู้สอบ จากการทดสอบเทเลอ์รูปิรามิดที่มีรูปแบบ จำนวนขึ้น และวิธีการให้
คะแนนที่แตกต่างกัน โดยใช่วิธีมอนติ คาร์โล. ปรินญานิพนธ์ กศ.ค. กรุงเทพฯ :
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2534. อัดสำเนา.
- บัญชา พนเจริญศักดิ์. "ข้อมูลมอนติคาร์โลสำหรับการกระจายปกติ," วิทยาศาสตร์.
11(1) : 700 - 704; พฤศจิกายน 2527.
- วิชา บัวคง. การเปรียบเทียบประสิทธิผลของวิธีประมาณค่าพารามิเตอร์ของแบบจำลอง
โลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซ์ิมไลค์ลิสต์ วิธีอีวีรอสติค และวิธีของ
เบส์ ในแบบสอบวัดผลสัมฤทธิ์และแบบสอบความถนัด. วิทยานิพนธ์ ค.ม.
กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2533. อัดสำเนา.
- วินัย วงศ์ฤทธิวัฒนา. การเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของโมเดลโลจิสติก
แบบสองพารามิเตอร์ ระหว่างวิธีของเบส์กับวิธีแมกซ์ิมไลค์ลิสต์. วิทยานิพนธ์
ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2532. อัดสำเนา.
- ล้วน สายยศ และอังคณา สายยศ. หลักการวิจัยทางการศึกษา. กรุงเทพฯ :
ศึกษาพร, 2536.
- เลิศลักษณ์ กลิ่นหอม. การศึกษาความคลาดเคลื่อนชนิดที่ 1 และอำนาจของการทดสอบใน
แบบแผนการทดลองแบบสุ่มภายในบล็อกที่มีปฏิสัมพันธ์ระหว่างสิ่งทดลองกับบล็อก.
ปรินญานิพนธ์ กศ.ค. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร,
2532. อัดสำเนา.
- ศิริจันทร์ ทองประเสริฐ. การจำลองแบบปัญหา. ภาควิชาวิศวกรรมอุตสาหการจุฬาลงกรณ์
มหาวิทยาลัย, 2537.
- สงบ ลักษณะ. "ความสามารถเกี่ยวกับการตอบข้อสอบ," วารสารการวัดผลการศึกษา.
4(1) : 47 - 54, 88 - 94 : พฤษภาคม-สิงหาคม, 2525.
- สนชา สีนาน. การศึกษาโดยวิธีมอนติ คาร์โล เปรียบเทียบอำนาจของการทดสอบการ
เท่ากันของความแปรปรวนระหว่างประชากรสองกลุ่ม. ปรินญานิพนธ์ สด.ม.
กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2528. อัดสำเนา.

- สุชาดา กิระนันท์. การอนุมานเชิงสถิติ : ทฤษฎีขั้นต้น. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2525.
- สุชาติ ประสิทธิ์รัฐสินธุ์. เทคนิคทางสถิติขั้นสูงสำหรับการวิเคราะห์ข้อมูลด้วยไมโครคอมพิวเตอร์และโปรแกรมสำเร็จรูป SPSS PC+ เล่ม 1. กรุงเทพฯ : สถาบันบัณฑิตพัฒนบริหารศาสตร์, 2533.
- เสรี เปรมสวัสดิโรจน์. การศึกษาความเที่ยงตรงในการจำแนกความรอบรู้และการประมาณค่าความสามารถในวิชาคณิตศาสตร์ของแบบทดสอบเฟล็กซีเบิล และแบบทดสอบรูปปริamid ที่ใช้ขนาดขึ้นแปรผัน. ปรินซ์ตันพังก์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2531. อัดสำเนา.
- Baker, F.B. "Advance in Item Analysis," Review of Education Research. 47 : 151 - 178 ; Winter, 1977.
- Baton, Richard F. A Primer on Simulation and Gaming. Englewood Cliffs : Princetice-Hall, 1970.
- Birnbaum, Allan. "Some Latent Trait Models and Their Use in Inferring in Examinee's Ability," In F.M. Lord and M.R. Novick. Statistical Theories of Mental Test Scores. Reading, Massachusetts : Addison-Wesley, 1968.
- Hambleton, R.K. and Cook, L.L. "Latent Triat Model and Their Use in The Analysis of Education Test Data," Journal of Educational Measurement. 14(2) : 75-96; Summer, 1977.
- ✕ Hambleton, R.K. and Swaminathan, Hariharan. Item Response Theory. Boston : Kluwer-Nijhoff, 1985.
- Hammersley, J.M. and Handscomb, D.C. Monte Carlo Methods. London : Nethuen and Co.Ltd., 1964.
- Hulin, Drasgow and Parsons. Item Response Theory Application to Psychological Measurement. Illinois : Dow Jones-Irwing Homewood, 1983.

- Jonosky, Janine Eileen. "A Study of Microcomputer Based Pseudo-Random Number Generators for the Uniform and Normal Probability Distribution," Dissertation Abstracts International. 47(12) : 4367 - A; June, 1987.
- Keppel, Geoffrey. Design and Analysis : a Researcher's Handbook. 3 rd ed. Englewood Cliffs, N.J. : Prentice-Hall , 1973.
- Kerlinger, Fred N. Foundation of Behavioral Research. Hong Kong : Holt, Rinehart and Winston. Inc., 1988.
- Krobe, Ronald W. and Forge, L.L. The Manager's Guide to Statistics and Quantitative Methods. New York : McGraw-Hill, 1980.
- Lord, F.M. "An Analysis of the Verbal Scholastic Aptitude Test Using Birnbaum's Three-Parameter Logistic Model," Educational and Psychological Measurement. 28 : 989 - 1020, 1968.
- Lord, F.M. and Novick, M.R. Statistic Theories of Mental Test Score. Reading, Massachusetts : Addison-Wesly, 1968.
- V Lord, F.M. Application of Item Response Theory to Practical Testing Problem. Hillsdale, New Jersey : Lawrence Erlbaum Associates, 1980.
- Lord, F.M. "Maximum Likelihood and Bayesian Parameter Estimation in Item Response Theory," Journal of Educational Measurement. 23(2) : 157 - 162; Summer, 1986.
- Meyer, Herbert A. Symposium on Monte Carlo Methods. New York : John Willy and Son, 1956.
- Mihram, G. Arthur. Simulation : Statistical Foundation and Methodology. New York : Academic Press, 1972.
- Mislevy, R.J. and Bock, R.D. BILOG : Maximum Likelihood Item Analysis and Test Scoring with Logistic Models. Mooresville, IN : Scientific Software, 1990.

ประวัติย่อของผู้วิจัย

ชื่อ นางรัตนา ศรีเหรียญ

เกิด วันที่ 8 ธันวาคม พ.ศ. 2497

สถานที่เกิด กรุงเทพมหานคร

สถานที่อยู่ปัจจุบัน บ้านเลขที่ 481/514 หมู่บ้านจรัญวิลล่า ซอยจรัญสนิทวงศ์ 37

ถนนจรัญสนิทวงศ์ เขตบางกอกน้อย กรุงเทพฯ 10700

ตำแหน่งหน้าที่ปัจจุบัน

ประวัติการศึกษา

- พ.ศ. 2508 จบชั้นประถมศึกษาปีที่ 4 จากโรงเรียนสตรีวัฒนา กรุงเทพฯ
- พ.ศ. 2511 จบชั้นประถมศึกษาปีที่ 7 จากโรงเรียนสตรีวัฒนา กรุงเทพฯ
- พ.ศ. 2514 จบชั้นมัธยมศึกษาปีที่ 3 จากโรงเรียนสตรีวัฒนา กรุงเทพฯ
- พ.ศ. 2516 จบชั้นมัธยมศึกษาปีที่ 5 จากโรงเรียนสาธิตน้ำผึ้ง กรุงเทพฯ
- พ.ศ. 2520 จบ กศ.บ. (เอกชีววิทยา โทเคมี) จากมหาวิทยาลัยศรีนครินทรวิโรฒ
ปทุมวัน กรุงเทพฯ
- พ.ศ. 2529 จบ กศ.ม. (เอกวัดผลการศึกษา) จากมหาวิทยาลัยศรีนครินทรวิโรฒ
ประสานมิตร กรุงเทพฯ
- พ.ศ. 2539 จบ กศ.ด. (เอกการทดสอบและวัดผลการศึกษา) จากมหาวิทยาลัย
ศรีนครินทรวิโรฒ ประสานมิตร กรุงเทพฯ

การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลอง
โลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซ์ิมัม โลค์ลิสต์ วิธีฮิวริสติก
และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน
ด้วยเทคนิคมอนติ คาร์โล

บทคัดย่อ
ของ
รัตนา ศรีเทริญ

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาบัณฑิต สาขาการทดสอบและวัดผลการศึกษา
มิถุนายน 2539

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อศึกษาค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ รวมทั้งเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ ระหว่างวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวีริสติก และวิธีของเบส์ ของแบบจำลองโลจิสติก 3 พารามิเตอร์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน โดยการสุ่มผลการตอบข้อสอบของผู้สอบแข่งขันเพื่อบรรจุบุคคลเข้ารับราชการเป็นข้าราชการครู ครั้งที่ 1 ปี 2537 วิชาภาษาไทย ของ อ.ก.ค.กรมสามัญศึกษา ออกเป็น 13 ขนาดกลุ่ม คือ ขนาดกลุ่ม 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,500, 2,000, 2,500 และ 3,000 คน จากจำนวนผู้สอบแข่งขัน 15,258 คน โดยแบบทดสอบที่ใช้สอบจริงประกอบด้วยข้อสอบจำนวน 100 ข้อ แต่กลุ่มตัวอย่างแต่ละคนในแต่ละกลุ่มต้องตอบข้อสอบที่มีขนาดต่างกัน 6 ขนาด คือ ขนาดข้อสอบ 50, 60, 70, 80, 90 และ 100 ข้อ ด้วยการทดลองซ้ำ 40 ครั้ง โดยใช้เทคนิคมอนติ คาร์โล แล้วนำผลการตอบดังกล่าว มาศึกษาค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ที่ประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ โดยใช้โปรแกรม BILOG 3 ของมิลเลวี และบ็อค วิธีอีวีริสติก โดยใช้โปรแกรม HUE-IRT และวิธีของเบส์ โดยใช้โปรแกรม BAY-IRT ที่ผู้วิจัยเป็นผู้พัฒนา จากนั้นนำค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบที่วิเคราะห์ได้ของแต่ละวิธี มาเปรียบเทียบความแตกต่างของค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ ด้วยสถิติทดสอบแบบ Repeated Measure (A x S) Design และทดสอบความแตกต่างรายคู่ด้วยวิธีการของเชฟเฟ (Scheffe's Test) ผลการวิจัย พบว่า

1. ค่าพารามิเตอร์ของข้อสอบที่ประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ วิธีอีวีริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2. ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีแมกซิมัม โลคัลลิสต์ และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ไม่มีความแตกต่างกัน ในขณะที่ค่าความสามารถของผู้สอบที่ประมาณค่าด้วยวิธีอีวีริสติก เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน มีความแตกต่างกันที่ระดับ .05

3. เมื่อขนาดของกลุ่มตัวอย่างต่างกัน การประมาณค่าพารามิเตอร์ด้วยวิธีอีวีริสติก สำหรับผู้สอบที่มีระดับความสามารถต่ำ มีคุณภาพสูงสุด รองลงมา คือ การประมาณค่าพารามิเตอร์ด้วยวิธีแมกซิมัม โลคัลลิสต์ สำหรับผู้สอบที่มีระดับความสามารถสูง และวิธีของเบส์ สำหรับผู้สอบที่มีระดับความสามารถต่ำ ตามลำดับ

4. เมื่อจำนวนข้อสอบมีขนาดต่างกัน การประมาณค่าพารามิเตอร์ด้วยวิธีของเบส์ สำหรับผู้สอบที่มีระดับความสามารถต่ำ มีคุณภาพสูงสุด รองลงมา คือ การประมาณค่าพารามิเตอร์ด้วยวิธีอีอาร์สติค สำหรับผู้สอบที่มีระดับความสามารถต่ำ และวิธีแมกซิมัม ไลค์ลียูด สำหรับผู้สอบที่มีระดับความสามารถสูง ตามลำดับ

จากการวิจัยดังกล่าว ข้อเสนอแนะที่ได้จากการวิจัยครั้งนี้ คือ

1. ควรเลือกใช้วิธีแมกซิมัม ไลค์ลียูด หรือวิธีของเบส์ ในการประมาณค่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ และใช้วิธีของเบส์ เมื่อกลุ่มตัวอย่างมีขนาดเล็ก
2. ควรเลือกใช้วิธีของเบส์ ในการประมาณค่าเมื่อผู้สอบมีระดับความสามารถต่ำ และใช้วิธีแมกซิมัม ไลค์ลียูด เมื่อผู้สอบมีระดับความสามารถสูง

A COMPARISON OF EFFICIENCY OF PARAMETER ESTIMATION OF THE THREE
PARAMETER LOGISTIC MODEL AMONG MAXIMUM LIKELIHOOD, HUERISTIC AND
BAYESIAN METHODS ON DIFFERENT SAMPLE SIZES AND NUMBER OF ITEMS
USING MONTE CARLO TECHNIQUE

ABSTRACT

BY

RATANA SRIHARUN

Presented in partial fulfillment of the requirements for the Doctor
of Education Degree in Testing and Education Measurement
at Srinakharinwirot University

June 1996

The purposes of the research were to estimate parameters of examinee's ability and the test items parameters using three different methods : Maximum Likelihood , Hueristic and Bayesian estimates for three parameter of logistic model; and to compare the efficiency of the estimations, focusing on different sample sizes and number of items. The sample was drawn from the first competitive examination to recruit and appoint personal government teachers in 1994 on Thai Language test preparing by the General Education department sub-comittees. Thirteen sample sizes : 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,500, 2,000, 2,500 and 3,000 subjects were drawn from 15,258 candidates. The original test consisted of 100 items. It was designed to divided into 50, 60, 70, 80, 90 and 100 items, subtest lenghts using Monte Carlo technique with 40 iteration. Computer programs used were three programs : BILOG 3, HUE-IRT and BAY-IRT program, which the last two programs developed by the researcher. In estimating basic statistical item parameter and examinees' ability parameter through the three different methods, compare item and examinee's ability parameter by reapeated method (AxS) measure design. Post hoc multiple comparisons were done by Scheffe's Test. The relative efficiency of tests on differrent sample sizes and number of items were also computed. The findings were as follows :-

1. Item parameter estimations among Maximum Likelihood, Hueristic and Bayesian estimates on different sample sizes and number of items were different at the .05 significance level.

2. Examinees' ability estimation between Maximum Likelihood method were equivalent to Bayesian estimated but differed from Hueristic method on different sample sizes and number of items at the .05 significance level.

3. On the various sample sizes, Hueristic method was the most efficient for estimating the low ability group, next by the Maximum Likelihood estimated for the high ability group and Bayesian estimated for the low ability group, respectively.

4. On the different number of items, Bayesian method was the most efficient for estimating the low ability group, next by the Hueristic estimated for the low ability group and Maximum Likelihood estimated for the high ability group, respectively.

The major suggestions of this study were :-

1. Using Maximum Likelihood or Bayesian method for estimating the large sample and Bayesian estimate for the small sample group.

2. Using Bayesian method for estimating the low ability group and Maximum Likelihood estimated for the high ability group.