การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ

นายกฤษฎา ถิระโสภณ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2550
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A COMPARISON OF STATISTICAL PROPERTIES OF THE INDICES FOR DETECTING ANSWER COPYING

Mr. Kridsada Thirasophon

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education Program
in Educational Measurement and Evaluation
Department of Educational Research and Psychology
Faculty of Education
Chulalongkorn University
Academic Year 2007

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ
โดย	นายกฤษฎา ถิระโสภณ
สาขาวิชา	การวัดและประเมินผลการศึกษา
อาจารย์ที่ปรึกษา	ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีคณะครุศาสตร์

(รองศาสตราจารย์ ดร. พฤทธิ์ ศิริบรรณพิทักษ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ

(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

..... อาจารย์ที่ปรึกษา

(ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร. ณัฏฐภรณ์ หลาวทอง)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กฤษฎา ถิระโสภณ : การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ. (A COMPARISON OF STATISTICAL PROPERTIES OF THE INDICES FOR DETECTING ANSWER COPYING) อ. ที่ปรึกษา : ศ.ดร.ศิริชัย กาญจนวาสี, 167 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์และเปรียบเทียบคุณสมบัติทางสถิติซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้สถานการณ์จำลองซึ่งแตกต่างในเงื่อนไขของตัวแปรต้นด้านความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก รวม 128 สถานการณ์ ในการดำเนินการวิจัยใช้วิธีวิจัยเชิงทดลอง โดยนำข้อมูลผลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 ของนิสิตที่เข้าสอบจำนวนทั้งสิ้น 250 คน มาจัดกระทำด้วยวิธี การจำลองสถานการณ์ และ วิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์ TAP, MULTILOG และ S-Plus ซึ่งผู้วิจัยเป็นผู้เขียนคำสั่งที่ใช้ในการวิเคราะห์ในโปรแกรม S-Plus เพื่อวิเคราะห์หาค่าความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ

ผลการวิจัยพบว่า

ดัชนี K_2 และ ดัชนี S_1 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในทุกสถานการณ์ โดยดัชนี S_1 เป็นดัชนีที่มีระดับค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุดในเกือบทุกสถานการณ์ ดัชนี S_2 สามารถควบคุมค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ และ ดัชนี ω สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ

ตัวแปรความยาวของแบบสอบที่เพิ่มขึ้นมีอิทธิพลทำให้ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω เพิ่มขึ้น แต่ไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2, ดัชนี S_1 และ ดัชนี S_2 ตัวแปรจำนวนผู้สอบและตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจของการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ ตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่แต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้ ตัวแปรวิธีการลอกแบบลอกข้อสอบเฉพาะข้อยากและตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกที่เพิ่มขึ้นมีอิทธิพลทำให้ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เพิ่มขึ้น และ โดยในเกือบทุกสถานการณ์ ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบต่ำที่สุด

ภาควิชา วิจัยและจิตวิทยาการศึกษา
สาขาวิชา การวัดและประเมินผลการศึกษา
ปีการศึกษา 2550

ลายมือชื่อนิสิต.....
ลายมือชื่ออาจารย์ที่ปรึกษา .....
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

4883654027 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEY WORDS : ANSWER COPYING INDICES / DETECTION RATE / TYPE I ERROR

KRIDSADA THIRASOPHON : A COMPARISON OF STATISTICAL PROPERTIES OF THE INDICES FOR DETECTING ANSWER COPYING. THESIS ADVISOR : PROF. SIRICHAI KANJANAWASEE, Ph.D., 167 pp.

The purposes of this study were to investigate and to compare the type I error rate and the detection rate of copying indices which were K_2 index, S_1 index, S_2 index and ω statistic for the conditions of 35-and 65-item tests, 100 and 250 sample sizes, 90^{th} and 60^{th} percentile rank-sources, 5% and 10 % of copiers, random- and difficulty weighted-copying and 10%, 20%, 30% and 40% answer copying. Experimental Research was employed in answer copying simulation from 2702303 final examination test responses of 250 examinees. The TAP, MULTILOG and S-plus computer programs were employed in analysis processes. The programs in S-plus which were written and developed by researcher were used to analyze the type I error rate and the detection rate of copying indices.

Results showed that K_2 index and S_1 index were able to control the type I error rates for all simulated situations. The S_2 index was able to control the type I error rates for 65-item test. The ω statistic was able to control the type I error rates for 35-item test. S_1 index was the most conservative copying index. Results also showed that the detection rate of ω increased with test length but not for K_2 index, S_1 index and S_2 index. The sample size and the percentage of copiers didn't affect the detection rates of the four copying indices. The ability levels of source slightly affect the detection rates of the indices with no clear tendency. The detection rate of the indices slightly increased with difficulty weighted-copying. The detection rate of the indices increased if the percentage of copied items was higher. It was also found that the S_1 index had the lowest detection rate.

Department Educational Research and Psychology — Student's signature.....
Field of study Educational Measurement and Evaluation — Advisor's signature.....
Academic year 2007 — Co-advisor's signature.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความกรุณาของศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งกรุณาให้คำแนะนำ แก้ไขข้อบกพร่องต่างๆ และให้กำลังใจแก่ผู้วิจัยเป็นอย่างดีโดยตลอดระยะเวลาการทำวิจัย รวมทั้ง รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ และ ผู้ช่วยศาสตราจารย์ ดร. ณัฏฐภรณ์ หลาวทอง ซึ่งได้ให้ปรึกษา แนะนำ และ อนุญาตให้สามารถนำผลการตอบข้อสอบของนิสิตในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 มาใช้ในการทำวิจัยชิ้นนี้

ขอกราบขอบพระคุณ James A. Wollack, Leonardo S. Sotaridona และ Rob R. Meijer ซึ่งกรุณาให้คำแนะนำเกี่ยวกับกระบวนการทำวิจัย แก้ไขข้อบกพร่องของคำสั่งซึ่งใช้ในการประมวลผลหาค่าดัชนีตรวจจับการลอกข้อสอบทั้งสี่จนสมบูรณ์

ขอกราบขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ที่ให้คำแนะนำ และ ชี้แนะแนวทางในการแก้ไขข้อบกพร่องต่างๆ ของงานวิจัย ทำให้งานวิจัยสมบูรณ์มากขึ้น

ขอกราบขอบพระคุณคณาจารย์ภาควิชาวิจัยและจิตวิทยาการศึกษาทุกท่านที่ประสิทธิ์ประสาทวิชาให้ผู้วิจัยมีความรู้ ความเข้าใจ และ ปฏิบัติงานด้านวิจัยอย่างมีประสิทธิภาพมากยิ่งขึ้น

ขอขอบพระคุณ คุณ วราภรณ์ ยิ่งยวด สำหรับการจัดหาโปรแกรมคอมพิวเตอร์ S-PLUS 2000 ที่ใช้ในงานวิจัยชิ้นนี้

ขอขอบพระคุณสำหรับความรัก ความห่วงใย ความเข้าใจ และ กำลังใจเสมอมาของครอบครัวถิระโสภณ และ ญาติพี่น้องทุกคน รวมทั้ง คุณพจนา ทรรพนันทน์ ตลอดจนน้ำใจและความปรารถนาดีของพี่น้องและเพื่อนภาควิชาวิจัยและจิตวิทยาการศึกษาทุกคนที่ให้ความช่วยเหลือ แนะนำ และ เป็นกำลังใจตั้งแต่เริ่มทำการศึกษาจนงานวิจัยสำเร็จได้ด้วยดี ผู้วิจัยขอขอบคุณไว้ ณ ที่นี้

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... ง

บทคัดย่อภาษาอังกฤษ..... จ

กิตติกรรมประกาศ..... ฉ

สารบัญ..... ช

สารบัญตาราง..... ญ

สารบัญภาพ..... ฑ

บทที่

1 บทนำ..... 1

ความเป็นมาและความสำคัญของปัญหา..... 1

คำถามของการวิจัย..... 5

วัตถุประสงค์ของการวิจัย..... 5

สมมติฐานของการวิจัย..... 5

ขอบเขตของการวิจัย..... 7

ข้อตกลงเบื้องต้น..... 8

คำจำกัดความที่ใช้ในการวิจัย..... 8

ประโยชน์ที่ได้รับจากการวิจัย..... 10

2 เอกสารและงานวิจัยที่เกี่ยวข้อง..... 11

ตอนที่ 1 องค์ความรู้เชิงทฤษฎี..... 11

2.1 การทุจริตทางการศึกษา (Academic Dishonesty) หรือ การโกง (Cheating)..... 11

2.2 สาเหตุที่ทำให้ผู้เรียนทำการโกง..... 13

2.3 สาเหตุของการแพร่หลายของการโกง..... 14

2.4 วิธีการในการตรวจจับการลอกข้อสอบ..... 14

2.5 ประวัติความเป็นมาของดัชนีตรวจจับการลอกข้อสอบ..... 15

2.6 แนวทางการศึกษาคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ. 22

2.7 เกณฑ์สำหรับเปรียบเทียบคุณภาพของดัชนีตรวจจับการลอกข้อสอบ..... 23

2.8 ทฤษฎีของดัชนีตรวจจับการลอกข้อสอบที่ใช้ในการวิจัย..... 25

บทที่		หน้า
	2.8.1 ดัชนี K	25
	2.8.2 ดัชนี ω	31
	2.8.3 ดัชนี S_1.....	34
	2.8.4 ดัชนี S_2.....	37
	2.9 ทฤษฎีที่เป็นพื้นฐานของดัชนีการตรวจจับการลอกข้อสอบ.....	42
	ตอนที่ 2 งานวิจัยที่เกี่ยวข้องกับดัชนีตรวจจับการลอกข้อสอบ.....	50
3	วิธีดำเนินการวิจัย.....	61
	ขั้นตอนการดำเนินการวิจัย.....	63
	ประชากร กลุ่มตัวอย่าง และ การสุ่มกลุ่มตัวอย่าง.....	63
	ตัวแปรที่เกี่ยวข้อง.....	64
	การเก็บรวบรวมข้อมูล และ เครื่องมือที่ใช้ในการวิจัย.....	64
	ขั้นตอนการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (Simulation)	64
	วิธีการจัดกระทำข้อมูลด้วยวิธีการจำลองสถานการณ์ (Simulation).....	71
	การวิเคราะห์ซ้ำ (Replication).....	71
	ระดับนัยสำคัญ (Alpha Level: α).....	72
	การเขียนคำสั่งเพื่อทำการวิเคราะห์ค่าสถิติตรวจจับการลอกข้อสอบ.....	72
	การวิเคราะห์ข้อมูล.....	74
	โปรแกรมที่ใช้ในการวิเคราะห์ข้อมูล.....	74
	การระบุการลอกของดัชนีตรวจจับการลอกข้อสอบ.....	75
	คุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ.....	76
	การศึกษาคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ.....	79
	การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ.....	80
4	ผลการวิเคราะห์ข้อมูล.....	84
	ตอนที่ 1 ผลการวิเคราะห์ข้อมูลในขั้นตอนก่อนและระหว่างการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์(Simulation).....	85

บทที่ หน้า

1. ผลการวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT)..... 85
2. ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) 89
3. ผลการวิเคราะห์ค่าพารามิเตอร์ของผู้สอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)..... 94

ตอนที่ 2 ผลการวิเคราะห์คุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ..... 99

1. ค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error Rate)..... 99
2. อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate)..... 103

5 สรุปผลการวิจัย อภิปรายผล และ ข้อเสนอแนะ..... 136

สรุปผลการวิจัย..... 136

อภิปรายผลการวิจัย..... 138

ข้อเสนอแนะในการนำไปใช้..... 143

ข้อเสนอแนะในการวิจัยครั้งต่อไป..... 145

รายการอ้างอิง..... 147

ภาคผนวก..... 149

ประวัติผู้เขียนวิทยานิพนธ์..... 167

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

ตารางที่		หน้า
1	คุณภาพของการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ.....	24
2	วิเคราะห์งานวิจัยเรื่อง คุณสมบัติทางสถิติของดัชนี K สำหรับการตรวจจับการลอกข้อสอบ (Sataridona และ Meijer, 2002) และ เรื่องสถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Sataridona และ Meijer, 2003) และข้อสรุปเกี่ยวกับแนวทางในการวิจัยของผู้วิจัย.....	55
3	ความถี่ของผู้ตอบตัวเลือกต่างๆ ของนิสิตคณะพาณิชยศาสตร์และการบัญชี คณะครุศาสตร์ และ คณะวิทยาศาสตร์.....	59
4	สัญลักษณ์ย่อภาษาอังกฤษแทนตัวแปรต้นที่ศึกษาในงานวิจัย.....	69
5	ค่าวิกฤติของการทดสอบสมมติฐานทางเดียว ณ ระดับนัยสำคัญที่ศึกษา.....	76
6	จำนวนผู้สอบที่ถูกใช้คำนวณหาค่าความคลาดเคลื่อนประเภทที่ 1 แยกตามตัวแปรจำนวนผู้สอบ และ ร้อยละของผู้ลอก.....	77
7	จำนวนผู้สอบที่ถูกใช้คำนวณหาค่าอำนาจตรวจจับการลอกข้อสอบ แยกตามตัวแปรจำนวนผู้สอบ และ ร้อยละของผู้ลอก.....	78
8	คะแนนสูงสุด คะแนนต่ำสุด ค่ามัธยฐาน ค่าเฉลี่ยเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน ความเบ้ ความโด่ง ของกลุ่มข้อมูลที่ถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบจำนวน 4 กลุ่ม.....	86
9	ค่าเฉลี่ยของค่าความยากง่ายของข้อสอบ ค่าเฉลี่ยของอำนาจจำแนกของข้อสอบ และ ค่าความเที่ยงแบบความสอดคล้องภายในของกลุ่มข้อมูลที่ถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบจำนวน 4 กลุ่ม.....	87
10	ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยากของข้อสอบแต่ละข้อ ของกลุ่มข้อมูล ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน.....	90
11	ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยากของข้อสอบแต่ละข้อ ของกลุ่มข้อมูล ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน.....	91

ตารางที่		หน้า
12	ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยากของข้อสอบแต่ละข้อ ของกลุ่มข้อมูล ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน.....	92
13	ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยากของข้อสอบแต่ละข้อ ของกลุ่มข้อมูล ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน.....	93
14	ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม MULTILOG ของผู้สอบแต่ละคน จากกลุ่มความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100) เรียงตามค่าเปอร์เซ็นไทล์.....	95
15	ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม MULTILOG ของผู้สอบแต่ละคน จากกลุ่มความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250) เรียงตามค่าเปอร์เซ็นไทล์.....	96
16	ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม Multilog ของผู้สอบแต่ละคน จากกลุ่มความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100) เรียงตามค่าเปอร์เซ็นไทล์.....	97
17	ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม Multilog ของผู้สอบแต่ละคน จากกลุ่มความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250) เรียงตามค่าเปอร์เซ็นไทล์.....	98
18	ความคลาดเคลื่อนประเภทที่ 1 (Type I error) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ณ ระดับนัยสำคัญ 7 ระดับ แยกตามตัวแปร จำนวนผู้สอบ (N) และ ความยาวของแบบสอบ (L).....	100
19	สัญลักษณ์ย่อภาษาอังกฤษแทนตัวแปรต้นที่ศึกษาในงานวิจัย.....	104
20	อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2 และ ดัชนี ω ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 1-สถานการณ์ที่ 32).....	105

ตารางที่ หน้า

21 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของ ดัชนี S_1 และ ดัชนี S_2 ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 1-สถานการณ์ที่ 32)..... 106

22 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2 และ ดัชนี ω ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 33-สถานการณ์ที่ 64)..... 107

23 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของ ดัชนี S_1 และ ดัชนี S_2 ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 33-สถานการณ์ที่ 64)..... 108

24 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2 และ ดัชนี ω ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 65-สถานการณ์ที่96)..... 109

25 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของ ดัชนี S_1 และ ดัชนี S_2 ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 65-สถานการณ์ที่ 96)..... 110

26 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2 และ ดัชนี ω ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 97-สถานการณ์ที่ 128)..... 111

27 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของ ดัชนี S_1 และ ดัชนี S_2 ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 97-สถานการณ์ที่ 128)..... 112

28 แนวโน้มค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error) และ อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนีตรวจจับการลอกข้อสอบทั้ง 4 ดัชนี ได้แก่ ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้ตัวแปรต้นด้านความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนผู้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก..... 134

ตารางที่		หน้า
29	ความหมายของสัญลักษณ์ที่ปรากฏในตาราง 26 แยกตาม คุณสมบัติทางสถิติด้านความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ…..	135
30	แนวทางในการเลือกใช้ดัชนีตรวจจับการลอกข้อสอบภายใต้ปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ……..	144

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

ภาพที่		หน้า
1	δ_{i^*jr} ซึ่งเป็นฟังก์ชันของ P_{i^*jr} โดย p_g= 0.25 และ p_g=0.20.....	40
2	โค้งรายการคำตอบสำหรับข้อคำถามตัวอย่าง โดยใช้ Nominal Response Model.....	44
3	กรอบความคิดของการวิจัย.....	60
4	ขั้นตอนในการดำเนินการวิจัย.....	62
5	สถานการณ์ที่ได้จากการจำลองเหตุการณ์ (simulation) 128 สถานการณ์.....	70
6	กราฟค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω เมื่อผู้สอบมีจำนวน 100 คน และ แบบสอบยาว 40 ข้อ.....	81
7	กราฟค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω เมื่อแบบสอบยาว 40 ข้อ และ ร้อยละของจำนวนข้อสอบที่ถูกลอกเป็น 30.....	83
8	การเปรียบค่าความคลาดเคลื่อนประเภทที่ 1 เชิงประจักษ์ (empirical type I error) ณ ระดับนัยสำคัญ 7 ระดับ ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) โดยแยกตามตัวแปรขนาดจำนวนผู้สอบ (simulee size) และ ความยาวของแบบสอบ (test length).....	101
9	อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก.....	113
10	อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก.....	114
11	อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก.....	115

ภาพที่ หน้า

12 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และวิธีการลอก..... 116

13 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก.... 117

14 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก.... 118

15 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และวิธีการลอก..... 119

16 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และวิธีการลอก..... 120

17 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก.... 121

ภาพที่ หน้า

18 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก..... 122

19 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และวิธีการลอก..... 123

20 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และวิธีการลอก..... 124

21 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก..... 125

22 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก.... 126

23 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และวิธีการลอก..... 127

ภาพที่ หน้า

24 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และวิธีการลอก..... 128

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในยุคปัจจุบันโลกได้กลายเป็นสังคมฐานความรู้ (Knowledge - based society) มนุษย์ไม่ว่าชนชาติใด ศาสนาใดล้วนแล้วแต่ที่จะต้องมีความรู้ และใช้ความรู้ที่มีในไปแก้ไขปัญหาที่เกิดขึ้น หากประเทศใดมีประชากรที่มีความรู้อยู่เป็นส่วนมากของประเทศ ประเทศนั้นก็จะมีการพัฒนาในด้านต่างๆ อย่างรวดเร็วและมีประสิทธิภาพ ซึ่งการที่จะทำให้ประชาชนในชาตินั้นมีความรู้ได้ วิธีการที่สำคัญที่สุดคือ การจัดการเรียนการสอนแก่ประชากรให้มีความรู้ความสามารถ แต่เพียงแค่การจัดการเรียนการสอนแก่ผู้เรียนนั้นไม่เพียงพอที่จะทำให้เราทราบได้ว่าผู้เรียนมีความรู้ความสามารถในเรื่องที่เรียนแล้ว ดังนั้นจึงต้องมีการประเมินผลเข้ามาเป็นส่วนหนึ่งในการจัดการเรียนการสอน โดยกระบวนการการประเมินผลจะช่วยให้การสอนมีประสิทธิภาพมากขึ้นและทำให้ผู้เรียนมีการเรียนรู้มากขึ้นอีกด้วย (Norman, 2003)

ในการประเมินผลการเรียนรู้นั้นจำเป็นต้องทำการวัดผลการเรียนของผู้เรียน ซึ่งวิธีการวัดผลการเรียนที่เป็นที่นิยมมากในปัจจุบัน คือ การสอบโดยใช้แบบสอบแบบหลายตัวเลือก (Multiple-Choice Tests) ซึ่งการจัดสอบเพื่อที่วัดผลสัมฤทธิ์ซึ่งเกิดจากการเรียนรู้ของผู้เรียนให้ได้ถูกต้องตามสภาพความเป็นจริงนั้น ผู้สอนต้องทำการจัดการคุมสอบให้รัดกุมเพื่อไม่ให้เกิดการโกงข้อสอบขึ้น (Cheating) เนื่องจากหากเกิดการโกงข้อสอบขึ้นไม่ว่าในรูปแบบใดก็ตามจะส่งผลให้ผลการสอบที่ได้นั้นมีความคลาดเคลื่อนไปจากระดับความสามารถที่แท้จริงของผู้สอบซึ่งจะทำให้ผลการประเมินระดับความสามารถของผู้เรียนคลาดเคลื่อนไปจากความเป็นจริง

การโกงข้อสอบนั้นเกิดขึ้นมานานแล้วโดยจากหนังสือ Cheating on tests: How to do It, Detect It, and Prevent It ของ Cizek (1999) ในบทที่ 3 ซึ่งมีชื่อบทว่า “How to Cheat: A Compendium of Methods” ได้กล่าวถึงวิธีการลอกข้อสอบหลากหลายวิธี เช่น การใช้วัสดุต่างๆ ที่ซ่อนไว้ การใช้ลูกอมช็อคโกแล็ตในการบอกใบ้คำตอบ การใช้เทคโนโลยีสมัยใหม่ ตัวอย่างเช่น การใช้ เครื่องส่งสัญญาณวิทยุ เวปไซต์ ในการโกงข้อสอบ เป็นต้น ซึ่งในปัจจุบันได้มีการศึกษาเกี่ยวกับการโกงข้อสอบมากมาย ซึ่งหนึ่งในวิธีการโกงนั้นคือ การลอกข้อสอบ (Answer Copying) ซึ่งโดยส่วนใหญ่แล้วการลอกข้อสอบเป็นวิธีการที่ผู้สอบลอกคำตอบจากผู้สอบที่นั่งสอบในบริเวณใกล้เคียง แต่ในสถานการณ์จริงการลอกข้อสอบนั้นไม่จำเป็นที่ผู้ลอกต้องลอกคำตอบจากผู้สอบที่นั่งในบริเวณใกล้เคียงกันเท่านั้น การลอกข้อสอบสามารถทำโดยการส่งสัญญาณเป็นรหัสจากผู้ให้ลอกซึ่งนั่งสอบในระยะที่ไกลได้ด้วย เช่น การเคาะเท้า กดปากกา เป็นต้น เนื่องด้วยการลอก

ข้อสอบนั้นจะทำให้ผลการสอบของผู้สอบไม่ตรงตามความเป็นจริง (Meijer, 1998) ดังนั้นเราจึงต้องทำการป้องกันการลอกข้อสอบโดยใช้การคุมสอบที่ดีและการจัดที่นั่งสอบที่มีประสิทธิภาพ

ปัญหาซึ่งเกิดจากการลอกข้อสอบไม่เพียงแต่มีความสำคัญต่อการวัดและประเมินผลการเรียนระดับรายวิชา หรือ ระดับชั้นเรียนเท่านั้น แต่ในการสอบระดับที่สูงขึ้น เช่น การสอบเพื่อเข้าศึกษาต่อในระดับมหาวิทยาลัย การสอบเพื่อเข้าทำงาน หรือ การสอบเพื่อรับใบอนุญาตประกอบวิชาชีพต่างๆ นั้นหากเกิดการลอกข้อสอบขึ้นจะก่อให้เกิดปัญหาระดับที่ใหญ่มากขึ้น เนื่องจากการสอบในระดับที่สูงขึ้นนั้นมักเป็นการสอบเพื่อคัดเลือกผู้ที่มีความสามารถเข้าศึกษา ทำงาน หรือได้รับใบอนุญาตประกอบวิชาชีพต่างๆ หากเกิดการลอกขึ้นจะทำให้เกิดความไม่ยุติธรรม และความผิดพลาดในการคัดเลือกบุคคลเข้าศึกษาต่อ ทำงาน หรือ ได้รับใบอนุญาตประกอบวิชาชีพ ซึ่งก่อให้เกิดผลเสียต่อ สถาบัน องค์กร บริษัท สังคม รวมถึงประเทศด้วย ดังนั้นการป้องกันไม่ให้เกิดการลอกข้อสอบจึงมีความสำคัญและจำเป็นอย่างยิ่งต่อการจัดสอบทุกระดับ

แต่อย่างไรก็ตามในบางครั้งผู้คุมสอบสังเกตพฤติกรรมผิดปกติบางอย่างซึ่งส่อว่าผู้สอบอาจทำการลอกข้อสอบ แต่ผู้คุมสอบไม่สามารถที่จะหาหลักฐานยืนยันได้ว่าผู้สอบทำการลอกข้อสอบกันจริง ทำให้ผู้ที่ทำการลอกข้อสอบไม่ได้รับการลงโทษใดๆ จากการลอกข้อสอบ ซึ่งการที่นักเรียนผู้ที่ทำการลอกข้อสอบไม่ได้รับการลงโทษใดๆ เนื่องจากไม่มีหลักฐานยืนยันการลอกข้อสอบที่เพียงพอนั้นไม่เพียงแต่จะส่งผลต่อความถูกต้องแม่นยำของการวัดและประเมินผลการเรียนเท่านั้น แต่ยังส่งผลต่อคุณลักษณะนิสัยของผู้เรียนที่จะไม่เกรงกลัวและไม่ละอายต่อการทุจริตในการสอบอีกด้วย โดยในปัจจุบันได้มีการพัฒนาวิธีการที่จะสามารถช่วยให้ผู้คุมสอบสามารถตรวจสอบได้ว่าผู้สอบที่ถูกสงสัยว่าจะทำการลอกข้อสอบกันนั้นมีแนวโน้มที่จะลอกข้อสอบกันหรือไม่ วิธีการดังกล่าวคือ วิธีการทางสถิติ โดยวิธีการทางสถิตินั้นจะทำให้ได้หลักฐานเพิ่มเติมที่หนักแน่นและเพียงพอเกี่ยวกับการลอกข้อสอบ (Sotaridona and Meijer, 2002)

หนึ่งในวิธีการทางสถิติดังที่กล่าวข้างต้นนั้นคือ ค่าสถิติเกี่ยวกับการลอกข้อสอบ (Answer-Copying statistic) โดยค่าสถิติเกี่ยวกับการลอกข้อสอบนั้นถูกคิดค้น และ นำเสนอมาเป็นเวลามากกว่า 25 ปีแล้ว (Assessment Systems Corporation, 1993; Bay, 1995; Bellezza & Bellezza, 1989; Cody, 1985; Frary, Tideman & Watts, 1977; Hanson, Harris & Brennan, 1987; Holland, 1996, Kadane, 1999; Sotaridona & Meijer, 2002,2003; Wollack, 1997) ซึ่งดัชนีตรวจจับการลอกข้อสอบที่ถูกเสนอและถูกศึกษานั้นมีอยู่หลายค่าดัชนี เช่น ดัชนี g_2, ค่าสถิติ ω, ดัชนี K^*, ดัชนี K_1, ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 เป็นต้น

โดยแนวทางในการศึกษาเกี่ยวกับคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบนั้นอยู่ 3 แนวทาง คือ แนวทางแรก การคำนวณค่าดัชนีโดยใช้ข้อมูลจริงซึ่งได้จากการสอบโดยเชื่อว่าจะมีข้อมูลที่เกิดจากการลอกข้อสอบปนอยู่ แล้วนำมาเปรียบเทียบกับค่าดัชนีซึ่งคำนวณจากชุด

ของข้อมูลที่เชื่อว่าไม่มีการลอกข้อสอบเกิดขึ้น (Angoff, 1974; Bellezza & Bellezza, 1989; Cody, 1985; Frary et al., 1977; Holland, 1996; Kadane, 1999) แนวทางการศึกษาคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบแนวทางที่สอง คือ การจำลองสถานการณ์ (simulation) (Sotaridona & Meijer, 2002, 2003; Wollack, 1997) และ แนวทางการศึกษาคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบแนวทางที่สาม คือ การใช้ข้อมูลจริงซึ่งถูกรวบรวมด้วยวิธีการซึ่งชุดข้อมูลจะไม่มีผู้สอบที่ลอกข้อสอบจากผู้สอบคนอื่นในชุดข้อมูลนั้นได้เลย (Bay, 1995; Hanson et al., 1987; Iwamoto, Nungester, Watson, & Luecht, 1997)

ปัจจุบันแนวทางในการศึกษาคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบที่ผู้วิจัยส่วนใหญ่นิยมใช้ คือ แนวทางที่สอง การจำลองสถานการณ์ (simulation) เนื่องจากในการศึกษาคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ ซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 (type I error) และ อำนาจการตรวจจับการลอกข้อสอบ (detection rates) นั้นผู้วิจัยจะทำการศึกษาในหลากหลายสถานการณ์ ซึ่งเป็นการยากที่จะดำเนินการเก็บข้อมูลจริง เนื่องจากในสถานการณ์จริงผู้วิจัยไม่สามารถกำหนดให้เกิดการลอกข้อสอบขึ้นตามทุกสถานการณ์ที่ผู้วิจัยศึกษาได้

ในปัจจุบันดัชนีตรวจจับการลอกข้อสอบที่ถูกเสนอและถูกศึกษานั้น คือ ดัชนี K_2 (K_2 index) (Sotaridona and Meijer, 2002), ดัชนี ω (Wollack, 1997), ดัชนี S_1 และ ดัชนี S_2 (Sotaridona and Meijer, 2003) โดยดัชนีตรวจจับการลอกข้อสอบแต่ละตัวคำนวณอยู่บนพื้นฐานของทฤษฎีที่แตกต่างกัน และ ดัชนีตรวจจับการลอกข้อสอบแต่ละตัวจะมีประสิทธิภาพที่แตกต่างกันไปตามเงื่อนไขที่ต่างกัน

ในด้านพื้นฐานทฤษฎีของดัชนีตรวจจับการลอกข้อสอบนั้น ดัชนี ω เป็นค่าสถิติตรวจสอบการลอกข้อสอบที่มีพื้นฐานอยู่บนโมเดลของทฤษฎีการตอบสนองข้อสอบ (item response theory) แบบ nominal response model (NRM) (Bock, 1972) ส่วนค่าดัชนี K_2 นั้นถูกประมาณค่าอยู่บนพื้นฐานของการแจกแจงทวินาม (binomial distribution) โดยนำเอาเฉพาะจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ลอกและผู้ให้ลอกมาวิเคราะห์หาค่าดัชนี ส่วนดัชนี S_1 มีการวิเคราะห์คล้ายคลึงกับดัชนี K_2 คือนำเอาเฉพาะจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ลอกและผู้ให้ลอกมาวิเคราะห์หาค่าดัชนีแต่ดัชนี S_1 จะมีพื้นฐานอยู่บนการแจกแจงแบบปัวซง (Poisson distribution) ส่วนดัชนี S_2 นั้นถูกประมาณค่าอยู่บนพื้นฐานการแจกแจงแบบปัวซงเช่นเดียวกับดัชนี S_1 แต่ ดัชนี S_2 จะนำเอาทั้งคำตอบผิดและคำตอบถูกที่ตรงกันระหว่างผู้ลอกและผู้ให้ลอกมาใช้ในการวิเคราะห์หาค่าดัชนี

ในด้านประสิทธิภาพของดัชนีตรวจจับการลอกข้อสอบนั้น ดัชนี K_2 (K_2 index), ดัชนี ω, ดัชนี S_1 และ ดัชนี S_2 จะมีประสิทธิภาพแตกต่างกันในสภาพเงื่อนไขที่แตกต่างกัน ดังจะเห็นได้จาก

งานวิจัยของ Leonardo S. Sotaridona และ Rob R. Meijer เรื่อง คุณสมบัติทางสถิติของดัชนี K สำหรับการตรวจจับการลอกข้อสอบ (Statistical Properties of the K-Index for detecting answer copying) ซึ่งทำการศึกษาเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบระหว่าง ดัชนี K ,ดัชนี K_1, ดัชนี K_2 และ ดัชนี ω ในสภาวะเงื่อนไขตัวแปรด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก และ ร้อยละของจำนวนข้อสอบที่ถูกลอกที่แตกต่างกัน ซึ่งใช้วิธีการจำลองสถานการณ์ (simulation) ในการศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ โดยผลการศึกษาและเปรียบเทียบพบว่าดัชนีตรวจจับการลอกข้อสอบแต่ละดัชนีจะมีประสิทธิภาพต่างกันไปตามเงื่อนไขตัวแปรที่แตกต่างกัน

นอกจากนี้ยังมีงานวิจัยของ Leonardo S. Sotaridona และ Rob R. Meijer เรื่อง สถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Two new statistics to detect answer copying) ซึ่งทำการศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบระหว่าง ดัชนี K2, ดัชนี ω และ ค่าดัชนีตรวจจับการลอกข้อสอบแบบใหม่ 2 ดัชนีที่ Leonardo S. Sotaridona และ Rob R. Meijer พัฒนาขึ้น คือ ดัชนี S_1 และ ดัชนี S_2 ในสถานการณ์ซึ่งตัวแปรด้าน ความยาวของแบบสอบ จำนวนผู้สอบ และ ร้อยละของจำนวนข้อสอบที่ถูกลอกที่แตกต่างกัน ซึ่งใช้วิธีการจำลองสถานการณ์ (simulation) ในการศึกษาและเปรียบเทียบ โดยผลการศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบทั้งสี่พบว่า ดัชนีตรวจจับการลอกข้อสอบทั้งสี่มีประสิทธิภาพต่างกันไปตามเงื่อนไขตัวแปรที่แตกต่างกัน หรืออาจกล่าวได้ว่า หากผู้ประสงค์จะนำค่าดัชนีตรวจจับการลอกข้อสอบไปใช้เลือกใช้ดัชนีตรวจจับการลอกข้อสอบที่เหมาะกับสภาพเงื่อนไขการสอบที่ถูกจัดขึ้นของตน ค่าดัชนีตรวจจับการลอกข้อสอบที่ได้จากการวิเคราะห์ก็จะมีความถูกต้องและเหมาะสมมากขึ้น ซึ่งจะส่งผลให้ความผิดพลาดในการตรวจจับการลอกข้อสอบน้อยลง

ด้วยเหตุผลดังที่กล่าวข้างต้น ผู้วิจัยจึงสนใจที่จะทำการวิเคราะห์คุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ ซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 (type I error) และ อำนาจการตรวจจับการลอกข้อสอบ (detection rates) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω และทำการเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบระหว่างดัชนีตรวจจับการลอกข้อสอบชนิดต่างๆ ดังกล่าวภายใต้เงื่อนไขตัวแปรต้นที่แตกต่างกัน เพื่อให้ได้สารสนเทศเกี่ยวกับสภาพเงื่อนไขหรือสถานการณ์ที่มีประสิทธิภาพที่สุดสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัว ซึ่งจะเป็นประโยชน์ต่อการเลือกใช้ค่าดัชนีตรวจจับการลอกข้อสอบที่เหมาะสมของผู้ที่จะนำเอาดัชนีตรวจจับการลอกข้อสอบไปใช้ในสภาพการสอบแบบต่างๆ

คำถามของการวิจัย

1. เมื่อเงื่อนไขตัวแปรต้นต่างในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก แล้วคุณสมบัติทางสถิติ ซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω จะมีลักษณะเป็นอย่างไร
2. ค่าความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบระหว่างดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้เงื่อนไขตัวแปรต้นซึ่งต่างในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก จะมีความเหมือนหรือแตกต่างกันอย่างไร

วัตถุประสงค์ของการวิจัย

1. เพื่อวิเคราะห์คุณสมบัติทางสถิติซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้เงื่อนไขตัวแปรต้นที่แตกต่างกันในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก
2. เพื่อเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ ระหว่างค่าดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้เงื่อนไขตัวแปรต้นที่แตกต่างกันในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก

สมมติฐานของการวิจัย

จากการศึกษางานวิจัยเรื่อง คุณสมบัติทางสถิติของดัชนี K สำหรับการตรวจจับการลอกข้อสอบ (Statistical Properties of the K-Index for detecting answer copying) ของ Leonardo S. Sotaridona และ Rob R. Meijer (2002) ซึ่งได้ทำการเปรียบเทียบคุณสมบัติทางสถิติ ซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ ระหว่างดัชนี K (Holland, 1996), ดัชนี K_1 (Sotaridona and Meijer, 2002), ดัชนี K_2 (Sotaridona and Meijer, 2002) และ ดัชนี ω (Wollack, 1997) ภายใต้เงื่อนไขตัวแปรต้นที่แตกต่างกันในด้าน ความยาว

ของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก และ ร้อยละของจำนวนข้อสอบที่ถูกลอก ผลการศึกษาพบว่า

1. ค่าดัชนี K ทุกแบบสามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 (type I Error) ให้ต่ำกว่าเส้นกราฟขอบเขต (boundary line) ในทุกสถานการณ์ที่ถูกจำลองขึ้น
2. ดัชนี ω สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 (type I Error) ให้อยู่ต่ำกว่าเส้นกราฟของเขต (boundary line) ในเกือบทุกสถานการณ์ที่ถูกจำลองขึ้น ยกเว้นสถานการณ์ที่แบบสอบความยาว 80 และจำนวนผู้สอบเป็น 500 และ 2000 คนที่ระดับความคลาดเคลื่อนประเภทที่ 1 (type I Error) มีค่ามากกว่า เส้นกราฟของเขต (boundary line)
3. ดัชนี ω จะมีอำนาจการตรวจจับการลอกข้อสอบ (Detection rate) สูงกว่าค่าดัชนี K ทุกแบบในเกือบทุกสถานการณ์ที่ถูกจำลองขึ้น
4. จำนวนของผู้สอบที่เพิ่มขึ้นจะไม่ส่งผลกระทบอย่างเห็นได้ชัดเจนต่ออำนาจตรวจจับการลอกข้อสอบ (detection rate) ของดัชนี ω แต่ในทางตรงกันข้ามการเพิ่มขึ้นของจำนวนผู้สอบจะส่งผลกระทบทำให้อำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของค่าดัชนี K ทุกแบบเพิ่มขึ้น

นอกจากนี้ผู้วิจัยยังศึกษางานวิจัย เรื่อง สถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Two new statistics to detect answer copying) ของ Leonardo S. Sotaridona และ Rob R. Meijer (2003) โดยสถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบนั้น Leonardo S. Sotaridona และ Rob R. Meijer เป็นผู้คิดค้นขึ้น ซึ่งได้แก่ ดัชนี S_1 และ ดัชนี S_2 โดยวัตถุประสงค์ของงานวิจัยชิ้นนี้คือ เพื่อเปรียบเทียบคุณสมบัติทางสถิติ ซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ ระหว่างดัชนี S_1, ดัชนี S_2, ดัชนี K_2 (Sotaridona and Meijer, 2002) และ ดัชนี ω (Wollack, 1997) ภายใต้เงื่อนไขตัวแปรต้นที่แตกต่างกันในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ และ ร้อยละของจำนวนข้อสอบที่ถูกลอก ผลการศึกษาพบว่า

1. ค่าดัชนี S_2 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 (type I Error) ให้อยู่ต่ำกว่าเส้นกราฟของเขต (boundary line) สำหรับสถานการณ์ที่จำนวนผู้สอบเป็น 100 คน และมีแนวโน้มที่ระดับค่าความคลาดเคลื่อนประเภทที่ 1 จะสูงกว่าเส้นกราฟขอบเขตจนไม่สามารถยอมรับได้สำหรับสถานการณ์ที่จำนวนผู้สอบเป็น 500 คน
2. ระดับความคลาดเคลื่อนประเภทที่ 1 (type I Error) ของดัชนี ω มีแนวโน้มที่จะสูงกว่าเส้นกราฟขอบเขตจนไม่สามารถยอมรับได้ในสถานการณ์ที่จำนวนผู้สอบเป็น 100 คน และ มีแนวโน้มที่จะมีต่ำกว่าเส้นกราฟขอบเขตในสถานการณ์ที่จำนวนผู้สอบเป็น 500 คน

3. ดัชนี S_1 และดัชนี K_2 จะสามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 (type I Error) ให้ต่ำกว่าเส้นกราฟของเขต (boundary line) ในเกือบทุกสถานการณ์โดยจะมีค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนี S_2 และดัชนี ω
4. อำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของค่าดัชนีทุกค่าจะเพิ่มขึ้นเมื่อร้อยละของจำนวนข้อสอบที่ถูกลอกเพิ่มขึ้น
5. อำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของดัชนี ω จะเพิ่มขึ้นเมื่อความยาวของแบบสอบเพิ่มขึ้น แต่การเปลี่ยนแปลงของจำนวนผู้สอบจะไม่ส่งผลต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ซึ่งข้อค้นพบนี้สอดคล้องกับข้อค้นพบจากงานวิจัยของ Wollack (1997) และ Sotaridoan และ Meijer (2002) แต่จำนวนผู้สอบ และ ความยาวของแบบสอบ ที่เพิ่มขึ้นจะส่งผลทำให้อำนาจการตรวจจับการลอกข้อสอบ ของค่าดัชนี K_2, S_1 และ S_2 มีค่าเพิ่มขึ้น
6. โดยทั่วไปแล้วค่าดัชนี K_2 จะมีอำนาจการตรวจจับการลอกข้อสอบ (detection rate) ต่ำที่สุดเมื่อเทียบกับค่าดัชนีตัวอื่นๆ
7. ไม่มีดัชนีตรวจจับการลอกข้อสอบใดที่จะใช้ได้ดีที่สุดในทุกสถานการณ์ที่ถูกจำลองขึ้น

จากผลการศึกษางานวิจัยทั้งสองชิ้นข้างต้นทำให้ผู้วิจัยตั้งสมมติฐานของงานวิจัยชิ้นนี้ได้ดังนี้

1. ค่าดัชนี K_2 จะสามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 (type I error) ให้ต่ำกว่าเส้นกราฟของเขต (boundary line) ในทุกสถานการณ์ที่จำลองขึ้น แต่ดัชนี ω จะไม่สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 (type I error) ให้ต่ำกว่าเส้นกราฟของเขต (boundary line) ได้ในทุกสถานการณ์ที่จำลองขึ้น
2. ดัชนีตรวจจับการลอกข้อสอบแต่ละตัวจะสามารถใช้ได้อย่างมีประสิทธิภาพที่สุดในสถานการณ์ที่แตกต่างกัน
3. อำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของดัชนีตรวจจับการลอกข้อสอบทุกตัวจะเพิ่มขึ้นเมื่อร้อยละของจำนวนข้อสอบที่ถูกลอกเพิ่มขึ้น

ขอบเขตการวิจัย

1. ตัวแปรที่ศึกษา

1.1 ตัวแปรต้น ประกอบด้วย

1. ความยาวของแบบสอบ แบ่งออกเป็นแบบสอบขนาดยาว 35 ข้อ และ 65 ข้อ
2. จำนวนผู้สอบ แบ่งออกเป็น จำนวนผู้สอบขนาด 100 และ 250 คน

3. ระดับความสามารถของผู้ให้ลอก แบ่งออกเป็น ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 และ 60
4. ร้อยละของจำนวนผู้ลอก แบ่งออกเป็น ร้อยละ 5 และ ร้อยละ 10 ของผู้สอบทั้งหมด
5. ร้อยละของจำนวนข้อสอบที่ถูกลอก แบ่งออกเป็น ร้อยละ 10, 20, 30 และ 40 ของข้อสอบทั้งหมด
6. วิธีการลอก แบ่งออกเป็น 2 วิธี คือ
    6.1) การลอกข้อสอบแบบสุ่ม
    6.2) การลอกข้อสอบเฉพาะข้อยาก

1.2 ตัวแปรตาม คือ ความคลาดเคลื่อนประเภทที่ 1 (type I error) และ อำนาจของการตรวจจับการลอกข้อสอบ (detection rate)

2. ประชากร คือ นิสิตระดับปริญญาบัณฑิตทั้งหมดที่ลงทะเบียนเรียนในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548

ข้อตกลงเบื้องต้น

งานวิจัยชิ้นนี้ผู้วิจัยใช้ข้อมูลผลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ส่วนข้อสอบปรนัย ซึ่งถือเป็นแบบสอบมาตรฐาน ของนิสิตระดับปริญญาบัณฑิตที่ลงทะเบียนและเข้าสอบในรายวิชาดังกล่าว จำนวนทั้งสิ้น 250 คน ซึ่งเป็นข้อมูลเชิงประจักษ์ (empirical data) มาจัดกระทำด้วยวิธีการจำลองสถานการณ์ (Simulation) โดยข้อมูลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ดังกล่าวถือเป็นข้อมูลผลการตอบข้อสอบที่ได้จากการสอบที่มีการคุมสอบซึ่งมีมาตรฐานไม่มีการลอกข้อสอบเกิดขึ้น

คำจำกัดความที่ใช้ในการวิจัย

การลอกข้อสอบ หมายถึง การทุจริตทางการศึกษาประเภทหนึ่ง ซึ่งเกิดจากการที่ผู้สอบซึ่งมีระดับความสามารถต่ำกว่ารับทราบหรือรับรู้คำตอบของผู้สอบอีกคนหนึ่งที่มีระดับความสามารถสูงกว่า จากนั้นตอบข้อสอบด้วยการเลือกตัวเลือกตามคำตอบของผู้สอบที่มีระดับความสามารถสูงกว่า

ดัชนีตรวจจับการลอกข้อสอบ หมายถึง ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ซึ่งใช้ในการตรวจสอบว่าผู้สอบที่ถูกสงสัยว่าจะลอกข้อสอบจากผู้ให้ลอกนั้นมีแนวโน้มที่จะลอกข้อสอบจริงหรือไม่

ดัชนี K_2 หมายถึง ดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบทวินาม (Binomial distribution) สัญลักษณ์ที่ใช้ คือ K_2 โดยหากค่าดัชนี K_2 ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้วผู้สอบในคู่ผู้สอบซึ่งถูกสงสัยว่าจะเป็นผู้ลอกจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก

ดัชนี S_1 หมายถึง ดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบปัวซง (Poisson distribution) สัญลักษณ์ที่ใช้ คือ S_1 โดยหากค่าดัชนี S_1ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้วผู้สอบในคู่ผู้สอบซึ่งถูกสงสัยว่าจะเป็นผู้ลอกจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก

ดัชนี S_2 หมายถึง ดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดและคำตอบถูกที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบปัวซง (Poisson distribution) สัญลักษณ์ที่ใช้ คือ S_2 โดยหากค่าดัชนี S_2 ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้วผู้สอบในคู่ผู้สอบซึ่งถูกสงสัยว่าจะเป็นผู้ลอกจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก

ดัชนี ω หรือ ค่าสถิติ ω หมายถึง ดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนโมเดลของทฤษฎีการตอบสนองข้อสอบ (item response theory) แบบ nominal response model (NRM) สัญลักษณ์ที่ใช้ คือ ω โดยหากค่าดัชนี ω มีค่ามากกว่าค่าวิกฤติของการทดสอบสมมติฐานทางเดียว (one-tailed critical value) โดยขอบเขตวิกฤติมีพื้นที่อยู่ทางด้านขวาของโค้งปกติแล้วผู้สอบในคู่ผู้สอบซึ่งถูกสงสัยว่าจะเป็นผู้ลอกจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก

คุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ หมายถึง ค่าความคลาดเคลื่อนประเภทที่ 1 (type I error) และ อำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของดัชนีตรวจจับการลอกข้อสอบ ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรมคอมพิวเตอร์

ความคลาดเคลื่อนประเภทที่ 1 หมายถึง สัดส่วนของผู้ที่ไม่ใช่ผู้ลอกอย่างแท้จริง (true noncopier) แต่ถูกระบุผิดว่าเป็นผู้ลอก (copier)

อำนาจการตรวจจับการลอกข้อสอบ หรือ อำนาจการตรวจจับการลอกข้อสอบ หมายถึง สัดส่วนของผู้ลอกจริง (true copier) ซึ่งถูกระบุว่าเป็นผู้ลอกหรือถูกตรวจจับได้โดยดัชนีตรวจจับการลอกข้อสอบ

ประโยชน์ที่ได้รับจากการวิจัย

1. ได้ข้อมูลเกี่ยวกับคุณสมบัติทางสถิติ ซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ ของดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้สถานการณ์ที่ต่างกันในเงื่อนไขตัวแปรต้นด้านความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก ซึ่งจะทำให้ผู้ประสงค์ที่จะนำดัชนีตรวจจับการลอกข้อสอบไปใช้สามารถเลือกใช้ค่าดัชนีตรวจจับการลอกข้อสอบให้เหมาะสมกับสถานการณ์การสอบของตน
2. ผลที่ได้จากการวิเคราะห์ค่าดัชนีตรวจจับการลอกข้อสอบสามารถนำไปใช้ในการจัดแผนผังที่นั่งสอบ หรือ สภาพแวดล้อมในการสอบเพื่อหลีกเลี่ยงการเกิดการลอกข้อสอบได้อย่างมีประสิทธิภาพ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องในการวิจัยครั้งนี้ แบ่งออกเป็น 2 ตอนดังนี้

ตอนที่ 1 องค์ความรู้เชิงทฤษฎี โดยแบ่งออกเป็นหัวข้อ ดังนี้

2.1 การทุจริตทางการศึกษา (Academic Dishonesty) หรือ การโกง (Cheating)

2.2 สาเหตุที่ทำให้ผู้เรียนทำการโกง

2.3 สาเหตุของการแพร่หลายของการโกง

2.4 วิธีการในการตรวจจับการลอกข้อสอบ

2.5 ประวัติความเป็นมาของดัชนีตรวจจับการลอกข้อสอบ

2.6 แนวทางการศึกษาคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ

2.7 เกณฑ์สำหรับเปรียบเทียบคุณภาพของดัชนีตรวจจับการลอกข้อสอบ

2.8 ทฤษฎีของดัชนีตรวจจับการลอกข้อสอบที่ใช้ในการวิจัย

2.9 ทฤษฎีที่เป็นพื้นฐานของดัชนีการตรวจจับการลอกข้อสอบ

ตอนที่ 2 งานวิจัยที่เกี่ยวข้องกับดัชนีตรวจจับการลอกข้อสอบ

ตอนที่ 1 องค์ความรู้เชิงทฤษฎีเกี่ยวกับดัชนีตรวจจับการลอกข้อสอบ

2.1 การทุจริตทางการศึกษา (Academic Dishonesty) หรือ การโกง (Cheating)

Maramark, Sheilah, Maline และ Mindi Barth (1993) กล่าวว่า การทุจริตทางการศึกษา (Academic Dishonesty) หรือ การโกง (Cheating) นั้นมีหลากหลายรูปแบบตั้งแต่รูปแบบธรรมดา คือ การลอกข้อสอบหรือลอกรายงานของนักเรียนคนอื่น จนถึง การขโมยแบบสอบ และรูปแบบอื่นๆ อีกมากมาย แต่นักวิจัยผู้ศึกษาด้านการทุจริตทางการศึกษาส่วนใหญ่นั้นมักให้ความสนใจกับการทุจริตที่เกิดขึ้นในขณะที่มีการสอบหรือการคัดลอกรายงานประจำภาค (plagiarism of team papers) นักวิจัยผู้ศึกษาด้านการทุจริตทางการศึกษาส่วนใหญ่นั้นไม่ให้ความสนใจกับรูปแบบอื่นๆ ของการทุจริต และด้วยสาเหตุนี้จึงทำให้ไม่ค่อยมีการศึกษาเกี่ยวกับการทุจริตลักษณะอื่นๆ

จากการศึกษางานวิจัยของ Maramark, Sheilah, Maline และ Mindi Barth (1993) เรื่อง การทุจริตทางการศึกษาของนักศึกษามหาวิทยาลัย ประเด็นทางการศึกษา (Acdemic

Dishonesty Among College Students. Issues in Education) ผู้วิจัยสามารถสรุปประเภทของการทุจริตทางการศึกษา (Academic Dishonesty) ได้ดังนี้ คือ

1. การทุจริตเชิงการลอก (copying cheating) เช่น
   การลอกคำตอบจากผู้สอบคนอื่น (Copied from another student's exam)
   ให้ผู้อื่นลอกข้อสอบระหว่างการสอบ (Gave another student answer during an exam)
   แบ่งปันคำตอบระหว่างการสอบโดยใช้การส่งสัญญาณระหว่างกัน (Shared answers during an exam by using a system of signals)
2. การทุจริตที่กระทำร่วมกันหรือช่วยเหลือกัน (collaborated cheating) เช่น
   ทำข้อสอบแทนผู้อื่น (Took an exam for someone else)
   จ้างผู้อื่นทำรายงานประจำภาค (Purchased term papers and turned in as own work)
   ทำหรือเขียนรายงานประจำภาคให้ผู้อื่น (Wrote term paper for another student)
   จ้างให้ผู้อื่นสอบหรือทำรายงานแทน (Hired a ghostwriter)
3. การทุจริตด้วยการกระทำที่ไม่สมควร (misconduct cheating) เช่น
   การลอกเนื้อหาโดยปราศจากการอ้างอิง (Copied materials without footnoting)
   การส่งรายงานฉบับเดียวกันกับหลายวิชาโดยไม่ได้รับอนุญาตจากผู้สอน (Submitted same term paper to another class without permission)
   การเปิดหนังสือหรือเอกสารระหว่างการสอบโดยไม่ได้รับอนุญาต (Used notes or books during exam when prohibited)
   ส่งรายงานผลการทดลองทั้งๆ ที่ไม่ได้ทำการทดลองจริง (Turned in a dry lab report without doing the experiment)
   บอกข้อคำถามในแบบสอบแก่ผู้เรียนห้องอื่นที่ยังไม่ได้ทำการสอบ (Gave test questions to students in another class)
4. การทุจริตโดยการเตรียมตัวก่อน (pre-perceiving cheating) เช่น
   การศึกษาสำเนาแบบสอบเพื่อศึกษาเนื้อหาเพิ่มเติม (Studied copy of exam prior to taking make-up)
   การศึกษาสำเนาแบบสอบฉบับก่อนๆ (Reviewed previous copies of an instructor's test)
   การศึกษาสำเนาแบบสอบที่ถูกขโมยมา (Reviewed a stolen copy of an exam)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การศึกษาแบบสอบหรือรายงานของรุ่นพี่หรือบุคคลที่ตนรู้จัก (Studied tests or used term papers from fraternity or sorority files)

5. การทุจริตแบบอื่นๆ เช่น

การแกล้งป่วยเพื่อหลบเลี่ยงการสอบ (Feigned illness to avoid a test)

การทำลายงานของผู้อื่น (Sabotaged someone else's work)

การสร้างความสัมพันธ์อันดีกับผู้สอนเพื่อให้ได้มาซึ่งข้อมูลเกี่ยวกับแบบสอบ (Developed a relationship with a instructor to get test information)

การพยายามต่อต้านผลการตัดเกรดของผู้สอนหลังการสอบ (Attempted to bias instructor's grading after an exam)

จะเห็นได้ว่า การลอกข้อสอบ (answer copying) เป็นประเภทหนึ่งของการโกง หรือ การทุจริตทางการศึกษา แต่ไม่ว่าจะเป็นการทุจริตแบบใดก็ตาม การทุจริตทางการศึกษาจะส่งผลต่อการวัดและประเมินผลผู้เรียน ทำให้ผลการวัดและการประเมินมีความคลาดเคลื่อนไม่ตรงตามความเป็นจริง

2.2 **สาเหตุที่ทำให้ผู้เรียนทำการทุจริตทางการศึกษา**

Maramark, Sheilah, Maline และ Mindi Barth (1993) กล่าวไว้ว่า สาเหตุที่ทำให้ผู้เรียนทุจริต นั้นมีความซับซ้อน แต่สาเหตุหลักซึ่งทำให้ผู้เรียนทุจริตมีอยู่สองสาเหตุ ก็คือ ความเครียด (Stress) และ การแข่งขัน (Competition) เช่น การแข่งขันเพื่อเข้าศึกษาต่อในระดับมหาวิทยาลัย การแข่งขันเพื่อการเข้าทำงานหลังเรียนจบ บรรยากาศทางสังคมหรือครอบครัวที่ทำให้เกิดความเครียด เป็นต้น

McCabe, Trevino และ Butterfield (2001) ได้กล่าวไว้ว่า ภาวะการแข่งขันที่เพิ่มขึ้นเพื่อให้ได้เข้าทำงานในตำแหน่งงานที่ใฝ่ฝันซึ่งมีตำแหน่งงานว่างไม่มากนัก หรือ การแข่งขันเพื่อให้ได้เข้าศึกษาต่อในมหาวิทยาลัยด้านธุรกิจ กฎหมาย และ แพทย์ชั้นนำระดับชาติ เป็นตัวแปรสำคัญที่ก่อให้เกิดแรงผลักดันให้ผู้เรียนต้องทำให้ได้ดี โดยมีผลการวิจัยพบว่าแรงผลักดันเหล่านี้ส่วนใหญ่ก่อให้เกิดการทุจริตทางการศึกษาขึ้น ดังจะเห็นได้จากงานวิจัยของ McCabe (1993), Nuss (1984), Singhal (1982)

จากการศึกษาเอกสารและงานวิจัยข้างต้นผู้วิจัยสามารถสรุปได้ว่า สาเหตุหลักของการทำให้ผู้เรียนโกงหรือกระทำการทุจริตทางการศึกษานั้นประกอบด้วยสองส่วน คือ การแข่งขัน (competition) และ ความเครียด (stress) ซึ่งการแข่งขันที่มีมากขึ้นในปัจจุบันทำให้ผู้เรียนเกิดแรงผลักดันมากจนเกิดเป็นความเครียดซึ่งความเครียดนี้เป็นสาเหตุหนึ่งที่ทำให้ผู้เรียนตัวสินใจโกงหรือกระทำการทุจริตทางการศึกษา

2.3 สาเหตุของการแพร่หลายของการทุจริตทางการศึกษา

สาเหตุของการแพร่หลายของการทุจริตทางการศึกษาหรือปัจจัยที่ส่งผลต่อการก่อให้เกิดการทุจริตทางการศึกษามากขึ้นนั้นสามารถแบ่งออกได้เป็น 2 ปัจจัย คือ

1. ปัจจัยส่วนบุคคล (individual factor) (Bower, 1964) เช่น เพศ (gender) ระดับคะแนนเฉลี่ย (Grade Point Average: GPA) ศีลธรรมในการทำงาน (work ethic) ความมุ่งมั่นในการแข่งขันเพื่อประสบความสำเร็จ (competitive achievement striving) การนับถือตนเอง (self-esteem) เป็นต้น
2. ปัจจัยบริบท (contextual factor) (Canning, 1956; Jendrek, 1989; Michaels และ Miethe, 1989; Tittle และ Rowe, 1973) เช่น การตอบโต้ของครูต่อการโกง (faculty response to cheating) บทลงโทษ (sanction threats) การศึกษาทางสังคม (social learning) กฎระเบียบ (honor code) เป็นต้น

McCabe และ Trevino (1997) ได้ทำการศึกษากับนักศึกษามหาวิทยาลัยขนาดกลางและใหญ่จำนวน 9 มหาวิทยาลัย จำนวนเกือบ 1,800 คน ในปีการศึกษา 1993-1994 โดย McCabe และ Trevino ทำการศึกษาเปรียบเทียบผลกระทบจากปัจจัยบริบทและปัจจัยส่วนบุคคลที่จะส่งผลต่อพฤติกรรมการทุจริตทางการศึกษา ผลการศึกษาพบว่า ปัจจัยบริบท (เช่น พฤติกรรมการโกงของเพื่อน การรับรู้ถึงความรุนแรงของบทลงโทษหากถูกจับได้ว่าโกง เป็นต้น) มีผลกระทบอย่างมีนัยสำคัญมากกว่าปัจจัยส่วนบุคคล (เช่น เพศ ระดับคะแนนเฉลี่ย การมีส่วนร่วมในกิจกรรมเสริมหลักสูตร เป็นต้น) โดยปัจจัยที่เกี่ยวข้องกับเพื่อน (Peer-related factors) จะมีความสัมพันธ์กับพฤติกรรมการลอกสูงที่สุด

2.4 วิธีการในการตรวจจับการลอกข้อสอบ

การลอกข้อสอบ (answer copying) นั้นเป็นประเภทหนึ่งของการทุจริตทางการศึกษา (academic dishonesty) ซึ่งการลอกข้อสอบนั้นจะทำให้ผลการวัดและการประเมินคลาดเคลื่อนไม่ตรงกับความเป็นจริง ดังนั้นในการจัดการทดสอบแต่ละครั้งจึงต้องมีการตรวจจับว่าได้มีการเกิดการลอกข้อสอบเกิดขึ้นหรือไม่ โดย Cizek (2000) กล่าวไว้ว่ามีวิธีการสองวิธีการในการตรวจจับการลอกข้อสอบ คือ

1. โดยการสังเกต (observational method)
2. โดยใช้กระบวนการทางสถิติ (statistical method)

โดย Cizek กล่าวไว้ในหนังสือ Detecting Cheating on Tests ว่าวิธีการที่ใช้กระบวนการทางสถิติจะมีความถูกต้อง ความน่าเชื่อถือ มากกว่า และ เป็นเครื่องมือที่มีประสิทธิภาพในการตรวจจับการโกงซึ่งเกิดขึ้นในการสอบ

ซึ่งวิธีการทางสถิติ หรือ ผู้วิจัยจะใช้คำว่า ดัชนีตรวจจับการลอกข้อสอบ นั้นได้มีการคิดค้นพัฒนา มาเป็นเวลายาวนานมาก ซึ่งสามารถศึกษาพัฒนาการในการคิดค้น และ พัฒนาดัชนีตรวจจับการลอกข้อสอบ ได้ในหัวข้อถัดไป

2.5 ประวัติความเป็นมาของดัชนีตรวจจับการลอกข้อสอบ

ในยุคแรกของการใช้แบบสอบหลายตัวเลือก (multiple-choice test) นั้นมักจะเกิดการลอกข้อสอบ (answer copying) ซึ่งเป็นประเภทหนึ่งของการทุจริตทางการศึกษา (academic dishonesty) และการตรวจจับการทุจริตชนิดนี้ทำได้ยากหรืออาจกล่าวได้ว่าเป็นไปไม่ได้ ไม่มีวิธีการใดๆ ที่สามารถใช้ในการตรวจจับการทุจริตประเภทนี้

แต่ต่อมาได้มีผู้คิดค้นวิธีการตรวจจับการลอกข้อสอบขึ้น โดยค่าสถิติเกี่ยวกับการลอกข้อสอบนั้นถูกคิดค้น และ นำเสนอมาเป็นเวลามากกว่า 25 ปีแล้ว (Assessment Systems Corporation, 1993; Bay, 1995; Bellezza & Bellezza, 1989; Cody, 1985; Frary, Tideman & Watts, 1977; Hanson, Harris & Brennan, 1987; Holland, 1996, Kadane, 1999; Sotaridona & Meijer, 2002,2003; Wollack, 1997)

โดยในหัวข้อนี้ผู้วิจัยจะทำการเสนอข้อมูลของประวัติความเป็นมาของดัชนีตรวจจับการลอกข้อสอบ โดยแบ่งออกเป็น 2 ยุค คือ

1. การศึกษาในยุคแรก (Early Efforts) (ค.ศ. 1927-1960)
2. การศึกษาในยุคต่อมา (ค.ศ. 1974-ปัจจุบัน)

1. การศึกษาในยุคแรก (Early Efforts)

Bird (1927,1929) เสนอวิธีการเชิงประจักษ์ (empirical approaches) 3 วิธี ในการตรวจจับ โดยมีพื้นฐานอยู่บนการตรวจจับการแจกแจงที่สังเกตได้ (observed distributions) ของคำตอบผิดที่ตรงกันของคู่ผู้สอบ

Crowford (1930) เสนอวิธีการที่คล้ายคลึงกัน โดยมีพื้นฐานอยู่บนร้อยละของคำตอบผิดที่ตรงกันของคู่ผู้สอบ

Dickenson (1945) เสนออัตราส่วนที่มีชื่อว่า “ร้อยละของความคลาดเคลื่อนที่ตรงกันที่เป็นไปได้ (probable percentage of identical errors)” ซึ่งมีพื้นฐานอยู่บนจำนวนของตัวเลือกต่อข้อและไม่ใช้คุณลักษณะการกระจายของร้อยละความคลาดเคลื่อนที่ตรงกันซึ่งสามารถสังเกตได้

Anikeef (1954) เปรียบเทียบจำนวนของคำตอบผิดที่ตรงกัน กับ Np เมื่อ N คือ จำนวนของคำตอบผิดทั้งหมดของผู้สอบที่ถูกสงสัย และ p คือส่วนกลับของจำนวนตัวเลือกในข้อคำถาม โดยสถิติตัวนี้จะใช้การแจกแจงทวินาม (binomial distribution) โดยมีส่วนเบี่ยงเบนมาตรฐานเป็น $\sqrt{Np(1-p)}$ ถ้าเป็นไปตามข้อสันนิษฐานที่ว่า ผู้สอบจะเดาอย่างสุ่มเมื่อไม่รู้คำตอบ โดย Anikeef ยอมรับในข้อจำกัด (inadequacy) ของข้อสันนิษฐานนี้ และ Anikeef ได้อ้างว่าวิธีการของเขานั้นจะมีประสิทธิภาพสำหรับผู้โกงข้อสอบที่ถูกระบุตัวแล้ว (identifying cheater)

โดยจากวิธีการศึกษาดัชนีตรวจจับการลอกข้อสอบที่กล่าวมาข้างต้นทั้งหมดมีข้อจำกัดในด้านการขาดวิธีการคำนวณ (lack of computational resources) รายงานฉบับแรกของการใช้คอมพิวเตอร์ในการตรวจจับการโกงได้ถูกเสนอโดย Saupe (1960) ซึ่งได้ทำการวิเคราะห์คำตอบโดยใช้อุปกรณ์อ่านเครื่องหมายด้วยแสง (Optical Mark Reader: OMR) Saupe ได้เสนอว่าจำนวนของคำตอบถูกหรือผิดที่ตรงกันของคู่ผู้สอบใดๆ นั้นจะขึ้นอยู่กับจำนวนคำตอบถูกหรือผิดทั้งหมดของคู่ผู้สอบแต่ละคู่ Saupe ใช้การวิเคราะห์ถดถอยเชิงเส้นตรง (linear regression) ในการทำนายจำนวนคำตอบถูกหรือผิดที่ตรงกันของคู่ผู้สอบแต่ละคู่จากผลคูณของจำนวนของคำตอบถูกและจำนวนข้อคำถามที่ตอบผิดเหมือนกัน จากพื้นฐานของค่าคลาดเคลื่อนมาตรฐานของการประมาณ (standard error of prediction) จะสามารถประมาณค่าความเป็นไปได้ 2 ค่า คือ ความเป็นไปได้ซึ่งความแตกต่างระหว่างจำนวนที่สังเกตได้กับจำนวนที่ได้จากการประมาณของคำตอบถูกและผิดที่ตรงกันโดยบังเอิญ (due to chance) ตามทฤษฎีแล้วสำหรับทุกๆ คู่ผู้สอบค่าประมาณทั้งสองค่าควรที่จะเป็นอิสระต่อกันทางสถิติ ซึ่งผลการศึกษาของ Saupe ได้เป็นไปตามทฤษฎีนี้ วิธีการนี้มีประสิทธิภาพในการระบุคู่ผู้สอบซึ่งลอกข้อสอบกันแต่จะยุ่งยากในการนำไปใช้เนื่องจากความจำเป็นที่จะต้องสร้างสมการถดถอย (regression equations) ทุกๆ การวิเคราะห์

2. การศึกษาในยุคต่อมา

Angoff (1974) รายงานผลการศึกษาซึ่งทำการเปรียบเทียบตัวบ่งชี้การลอกข้อสอบ (indicators of answer copying) ซึ่งคล้ายคลึงกับรายงานของ Saupe (1960) ตรงที่ Angoff ได้ใช้จำนวนของคำตอบถูกและผิดที่ตรงกัน แต่ส่วนที่ต่างคือ Angoff ได้ศึกษาจำนวนของการละเว้นที่จะตอบที่ตรงกัน (common omission) จำนวนของคำตอบผิดที่ตรงกันและการละเว้นที่จะตอบที่ตรงกันและ จำนวนของข้อคำถามที่มีคำตอบผิดที่ตรงกันและการละเว้นที่จะตอบที่ตรงกัน Angoff ได้นำเอา ตัวแปรอิสระ (independent variables) มาใช้ร่วมกับตัวบ่งชี้ที่ได้กล่าวไปแล้วข้างต้นด้วย โดยมีตัวแปรอิสระสองตัว คือ ตัวแปรทำนายสองตัวของ Saupe (Saupe's predictor variables) ส่วนตัวแปรอิสระอื่นๆ สำหรับคู่ผู้สอบ คือ ผลลัพธ์ของจำนวนคำตอบผิด ผลลัพธ์ของ

จำนวนการละเว้นที่จะตอบที่ตรงกัน และสุดท้ายคือผลรวมของจำนวนการละเว้นที่จะตอบกับคำตอบผิดของคู่ผู้สอบที่มีจำนวนคำตอบผิดที่น้อย ตัวแปรอิสระเหล่านี้จะถูกจับคู่กับตัวบ่งชี้อย่างเหมาะสมเพื่อให้ได้สิ่งที่ Angoff เรียกว่า ดัชนี (indices) จำนวน 8 ตัว อย่างไรก็ตามแทนที่จะใช้การวิเคราะห์ถดถอยเพื่อทำนายค่าดัชนีจากตัวแปรอิสระเช่นเดียวกับ Saupe ทำ Angoff ได้แบ่งข้อมูลของเขาออกเป็นชั้นๆ ตามตัวแปรอิสระและใช้การกระจายของตัวบ่งชี้ภายในชั้นข้อมูลเพื่อประมาณค่าตัวบ่งชี้ของคู่ผู้สอบใดๆ ในการทำให้ดัชนีทั้ง 8 ตัวมีความถูกต้อง Angoff ใช้ตัวอย่างการตอบ 3 กลุ่มของแบบสอบย่อยในส่วนการพูด (verbal) และ คณิตศาสตร์ (mathematical) ของการสอบ Scholastic Aptitude Tests (SAT) กลุ่มที่หนึ่งประกอบด้วยคู่ผู้สอบผู้ซึ่งไม่ได้ลอกข้อสอบกันเลยเนื่องจากอยู่คนละสถานที่ กลุ่มที่สองประกอบด้วยคู่ผู้สอบผู้ซึ่งถูกจัดให้สอบในสถานที่เดียวกันแต่ไม่มีรายงานว่าผู้สอบลอกข้อสอบกัน กลุ่มที่สามจะเหมือนกับกลุ่มที่หนึ่งยกเว้นแต่รูปแบบของแบบสอบ SAT ที่จะต่างกัน ผลปรากฏว่าได้ค่าดัชนีทุกตัวที่เหมือนกันสำหรับตัวอย่างทั้งสามกลุ่ม จากนั้นคำตอบ 50 คู่ของผู้ลอกที่ถูกรู้ (known copiers) จะถูกบอกและดัชนีต่างๆ จะถูกประเมินความมีประสิทธิภาพในการตรวจจับกรณีนี้ นอกจากนี้ยังมีการพิจารณาขนาดของความสัมพันธ์ระหว่างกันของค่าดัชนีต่างๆ ด้วย ผลการพิจารณาพบว่ามีดัชนีสองค่าที่ถูกจัดว่าดีที่สุด ดัชนีที่หนึ่งจะรวบรวมเอาคำตอบผิดที่ตรงกัน ผลลัพธ์ของจำนวนคำตอบผิด จำนวนอื่นๆ(ที่ระบุข้างต้น)และ จำนวนของคำตอบผิดและการละเว้นที่จะตอบของผู้สอบที่มีจำนวนคำตอบผิดที่น้อย น่าประหลาดที่ค่าดัชนีที่มีพื้นฐานบนจำนวนคำตอบถูกที่ตรงกันและผลลัพธ์ของจำนวนคำตอบถูก สามารถตรวจจับได้ดีเทียบเท่ากับค่าดัชนีที่มีพื้นฐานอยู่บนคำตอบผิดที่ตรงกันและมีข้อได้เปรียบในด้านความเป็นอิสระทางสถิติอีกด้วย (ตัวแปรสองตัวที่ถูกเสนอโดย Saupe, 1960) อย่างไรก็ตามค่าดัชนีไม่ถูกยอมรับโดย Angoff เนื่องจากว่า Angoff เกรงว่าผู้ที่ไม่ใช่นักสถิติอาจจะอ้างได้ว่าคำตอบถูกที่สอดคล้องกันนั้นเนื่องมาจากการมีความรู้เหมือนกัน

Frary, Tideman และ Watts (1977) พัฒนาและศึกษาดัชนีตรวจจับการลอกข้อสอบสองตัวซึ่งมีพื้นฐานที่แตกต่างจากค่าดัชนีที่ถูกเสนอข้างต้นที่กล่าวมา วิธีการคือการประมาณค่าความน่าจะเป็นซึ่งผู้สอบแต่ละคนจะเลือกคำตอบจากตัวเลือกที่มีอยู่ในข้อสอบ รวมทั้งการละเว้นที่จะเลือกคำตอบด้วย ค่าความน่าจะเป็นเหล่านี้จะถูกประมาณจากคะแนนสอบของผู้สอบและสัดส่วนของผู้สอบที่เลือกตัวเลือกแต่ละตัว ค่าดัชนีทั้งสองตัวจะมีพื้นฐานอยู่บนจำนวนคำตอบที่ตรงกัน (ถูก ผิด และการละเว้นที่ตอบ) ของคู่ผู้สอบ (examinee pair) แต่อย่างไรก็ตามดัชนีทั้งสองตัวแตกต่างกันในด้านสมมติฐานที่ถูกประเมิน ดัชนีตัวแรกประเมินสมมติฐานซึ่งจำนวนของคำตอบที่เหมือนกันนั้นเกิดขึ้นโดยปราศจากการลอกเพียงอย่างเดียว ดัชนีตัวที่สองพิจารณาผู้สอบหนึ่งคนที่อาจจะเป็นผู้ลอกและผู้สอบอีกหนึ่งคนที่อาจจะเป็นผู้ให้ลอก ซึ่งนำไปสู่การพิจารณาสมมติฐาน 2 สมมติฐาน คือ ผู้สอบคนแรกจะตอบคำตอบที่ตรงกันกับคำตอบของผู้สอบคนที่สองโดยปราศจาก

การลอก และในทางกลับกันด้วย ค่าดัชนีของสมมติฐานทั้งสองสมมติฐานจะมีค่าไม่เท่ากันถ้าผู้สอบทั้งสองคนมีคำตอบที่ตรงกันไม่ทั้งหมด ค่าที่คาดหวัง (expected values) และ ความแปรปรวน (variances) ของจำนวนคำตอบที่ตรงกันซึ่งเกิดขึ้นโดยบังเอิญได้มาจากสมมติฐานทั้งสองแบบ โดยมีความเป็นไปได้ที่จะทำให้เกิดค่าสถิติมาตรฐาน (standardized statistics) ที่การกระจายเป็นปกติ ดังนั้นเมื่อกำหนดให้คู่ผู้สอบใดๆ (สำหรับดัชนีตัวที่ 1) หรือ ผู้สอบหนึ่งคนที่อาจจะลอกข้อสอบจากบุคคลอื่น (สำหรับดัชนีตัวที่ 2) จะมีความเป็นไปได้ที่จะประมาณความเป็นไปได้ที่คำตอบที่ตรงกันจะเกิดขึ้นโดยปราศจากการลอก

Frary และคณะ พัฒนาค่าดัชนีของพวกเขาโดยการใช้คำตอบจากแบบสอบสองแบบที่มีข้อสอบเหมือนกันแต่เรียงลำดับแตกต่างกันซึ่งจัดสอบคนละห้องสอบ การกระจายของดัชนีถูกศึกษาเป็นคู่ คือ คู่ที่แบบสอบต่างแบบกันและอยู่คนละห้องสอบ กับ คู่ที่แบบสอบเดียวกันและอยู่ในห้องสอบเดียวกัน การกระจายของดัชนีที่แบบสอบเดียวกันและอยู่ในห้องสอบเดียวกันจะมีค่าดัชนีทั้งสองตัวที่สูงมากกว่าปกติ ในขณะที่การกระจายของดัชนีที่คู่แบบสอบต่างแบบกันและอยู่คนละห้องสอบ จะมีค่าของดัชนีสูงมากแค่ไม่กี่กรณี ค่าดัชนีแรกจะไม่สามารถบ่งชี้การลอกที่เป็นไปได้ได้หลายกรณีมากเท่ากับดัชนีตัวที่สองและ ไม่ถูกแนะนำให้ใช้ต่อไป

Schumacher (1980) รายงานวิธีการซึ่งต้องการความรู้ในเรื่องตำแหน่งที่นั่ง (seating locations) ของผู้สอบสองคนที่ถูกสงสัยว่าจะโกงข้อสอบ นอกจากนั้นผู้สอบทั้งสองคนนั้นยังต้องถูกจัดให้นั่งด้วยกันในส่วนหนึ่งของการสอบและนั่งแยกจากกันในการสอบส่วนที่เหลือ จากนั้นคำนวณ ค่าสถิติ chi-square ขนาด 2x2 (2x2 chi-square statistic) เพื่อประเมินความเป็นอิสระของจำนวนคำตอบที่เหมือนกันและไม่เหมือนกันระหว่างสองตำแหน่ง วิธีการนี้จะเหมาะเฉพาะสำหรับการสอบที่มีการจัดที่นั่งที่หลากหลาย เช่น ใช้กับการทดสอบเพื่อขอใบอนุญาตประกอบวิชาชีพ เป็นต้น Stegman และ Barnhill (1981, 1982) ได้รายงานถึงการใช้วิธีการของ Schumacher ที่กว้างขวางซึ่งเป็นข้อมูลจากคณะกรรมการของผู้ตรวจจับทางการแพทย์ระดับชาติ (the National Board of Medical Examiners) โดยพวกเขาใช้เฉพาะคำตอบผิดที่ตรงกันและไม่ตรงกัน ซึ่งสอดคล้องกับความกังวลของ Angoff (1974) เกี่ยวกับข้อโต้แย้งของผู้ที่ไม่ได้เป็นนักสถิติที่ว่าคำตอบถูกที่ตรงกันซึ่งมีปริมาณมากผิดปกติก็สามารถระบุผู้ลอกข้อสอบได้ การศึกษาส่วนใหญ่จะมุ่งเน้นไปเพื่อตรวจจับการแจกแจงของ chi-square ภายใต้สถานการณ์ที่ได้อธิบายข้างต้น โดยจากผลการศึกษาพบว่า จำนวนของคำตอบผิดที่เหมือนกันจะมากขึ้นเล็กน้อยหากผู้สอบมีพื้นทางการศึกษา (academic backgrounds) ที่เหมือนกัน

Cody (1985) เขียนบทความโดยไม่ได้มีการอ้างอิงเอกสารใดๆ เพียงแต่นำเสนอแนวคิดซึ่งเป็นการปรับอย่างหยาบๆ ในด้านพื้นฐาน ของดัชนีตัวที่สองของ Frary และคณะ เท่านั้น สำหรับการประมาณค่าความเป็นไปได้ซึ่งผู้สอบจะเลือกตัวเลือกหนึ่งๆ Cody จะใช้เฉพาะสัดส่วนของ

ผู้สอบที่เลือกตัวเลือกนั้น ในขณะที่ Frary และคณะ พิจารณาคะแนนของผู้สอบด้วย นอกจากนี้ Cody จะใช้เฉพาะคำตอบผิด ในขณะที่ Frary และคณะ จะใช้คำตอบทั้งหมดของผู้ลอกข้อสอบที่ถูกสงสัย และด้วยเหตุข้างกล่าว วิธีการของ Cody จึงมันถูกคาดหวังว่าจะมีความไวต่อการตรวจจับการลอกข้อสอบน้อยกว่าดัชนีตัวที่สองของ Frary และคณะ

Hanson, Harris และ Brennan (1987) เสนอดัชนีใหม่จำนวน 2 ตัว โดยมีพื้นฐานอยู่บนการวิเคราะห์ข้อมูลที่ครอบคลุม พวกเขาได้ให้ข้อสังเกตว่า สำหรับคู่ตัวบ่งชี้การลอกข้อสอบ (pairs of indicator of copying) คู่ตัวบ่งชี้ทั้งสองจะมีค่าสูงพร้อมๆ กับการที่ตัวบ่งชี้เพียงตัวเดียวซึ่งถูกปรับสำหรับระดับคะแนนของผู้สอบจะมีค่าสูง คู่ของตัวบ่งชี้คู่ที่หนึ่งได้แก่ จำนวนของคำตอบผิดที่ตรงกัน และ ความยาวของช่วงที่ยาวที่สุดของคำตอบที่ตรงกันหรือการละเว้นไม่ตอบ คู่ของตัวบ่งชี้คู่ที่สอง ได้แก่ จำนวนของคำตอบผิดในช่วงที่ยาวที่สุดของคำตอบที่ตรงกัน และ จำนวนของคำตอบผิดที่ตรงกันที่แสดงอยู่ในรูปร้อยละของจำนวนของคำตอบผิดที่เป็นไปได้สูงสุด

Hanson และคณะ เสนอและศึกษาพัฒนาค่าดัชนีของ Cody พวกเขาได้ประมาณความเป็นไปได้ซึ่งผู้ลอกข้อสอบซึ่งถูกสงสัยจะเลือกตัวเลือกหนึ่งๆ จากสัดส่วนของผู้สอบที่เลือกตัวเลือกนั้นๆ ภายในชั้นคะแนน (score stratum) ที่ผู้ถูกสงสัยอยู่ อย่างไรก็ตามพวกเขาไม่ได้ใช้คำตอบถูกและการละเว้นที่จะตอบที่ตรงกันในการคำนวณค่าดัชนีที่ถูกปรับปรุง (modified index) ถ้าพวกเขาใช้ข้อมูลดังกล่าวผลที่ได้น่าจะเหมือนกับค่าดัชนีตัวที่สองของ Frary และคณะ

Roberts (1987) ประเมินวิธีการในการตรวจจับการโกงข้อสอบซึ่งตามรายงานแล้วถูกใช้จากผู้สอนในมหาวิทยาลัยบางแห่งที่มีชุดของแบบสอบหลากหลายในห้องสอบเดียวกันโดยผู้สอบไม่รู้ จากนั้นการสอบข้ามชุดข้อสอบ (crossform copying) สามารถตรวจวัดโดยการตรวจให้คะแนนคำตอบของผู้สอบโดยใช้เฉลยจากทุกชุดข้อสอบ และหาผลต่างซึ่งมีค่าลบระหว่างคะแนนที่ได้จากการตรวจโดยใช้เฉลยที่เหมาะสม (applicable key) กับคะแนนที่ได้จากการตรวจโดยหนึ่งในเฉลยอื่นๆ Roberts ให้ข้อสรุปว่าวิธีการนี้สามารถก่อให้เกิดผลลัพธ์เชิงบวกที่ผิด (false positive outcomes)

Bellezza และ Bellezza (1989) รายงานว่าวิธีการตรวจจับที่มีพื้นฐานอยู่บนการแจกแจงทวินาม (binomial distribution) ซึ่งโดยคร่าวๆ จะเหมือนกับวิธีแรกของ Frary และคณะ (บทความนี้ของเขาไม่ได้ถูกอ้างอิง) ไม่เพียงแค่ดัชนีตัวแรกของ Frary และคณะเท่านั้นที่มีประสิทธิภาพต่ำกว่าดัชนีตัวที่สองของพวกเขา ดัชนีของ Bellezza และ Bellezza ก็มีประสิทธิภาพต่ำกว่าเช่นกันโดยไม่สามารถนำข้อมูล คำตอบถูกที่ตรงกัน หรือระดับคะแนนของผู้สอบ และ ความแพร่หลายของตัวเลือกผิดที่ตรงกัน เข้าไปใช้ในการคำนวณค่าสถิติของพวกเขา

Hanson และคณะ (1987) ศึกษาเปรียบเทียบวิธีการตรวจจับการลอกข้อสอบ โดยเปรียบเทียบวิธีการที่เขาได้พัฒนาขึ้นกับวิธีการของ Angoff (1974), Frary และคนอื่นๆ (1977)

และ Cody (1985) เพื่อเปรียบเทียบค่า false positive rates ของแต่ละค่าดัชนี โดยใช้ข้อมูลจากแบบสอบที่มีข้อสอบ 100 ข้อ แต่ละข้อมี 4 ตัวเลือก โดยมีจำนวนผู้สอบมากกว่า 19,000 คน ผู้สอบแต่ละคนจะถูกจับคู่กับอีกคนหนึ่งที่ทำการสอบในอีกสถานที่หนึ่ง ได้คู่ผู้สอบทั้งสิ้น 9,143 คู่ โดยคู่ผู้สอบจำนวน 8,643 คู่จะถูกใช้เป็นข้อมูลเกณฑ์ (benchmark data) เพื่อให้แสดงค่า false positive rates ของดัชนีแต่ละตัวที่ทำการศึกษา ส่วนคู่ผู้สอบที่เหลือ 500 คู่นั้นผู้ศึกษาจะให้ผู้สอบคนที่สองเป็นผู้ลอกข้อสอบ (copier) โดยขั้นตอนนี้จะลดความสมจริงของการศึกษาเนื่องจากคะแนนสอบของผู้สอบในคู่ผู้สอบไม่ได้ถูกนำมาพิจารณา ซึ่งอาจจะก่อให้เกิดการลอกที่ผู้ที่มีคะแนนมากกว่าลอกผู้ที่มีคะแนนน้อยกว่า โดยศึกษาระดับการลอก 5 ระดับ (ร้อยละ 10 ถึง 50 ของคำตอบของผู้ให้ลอก) และศึกษาวิธีการลอก 5 วิธีซึ่งจะถูกจำลองสถานการณ์ขึ้น 5 วิธีการนั้นได้แก่ 1) การลอกโดยการสุ่ม 2) การลอกข้อยาก 3) การลอกข้อสอบในช่วงแรกของชุดข้อสอบ 4) การลอกข้อสอบในช่วงท้ายของชุดข้อสอบ และ 5) การลอกข้อสอบเป็นชุด ชุดละ 5 ข้อ โดยแต่ละชุดได้จากการสุ่ม ผลการศึกษาของ Hanson และคนอื่นๆ ค่อนข้างที่จะครอบคลุมและซับซ้อน แต่กล่าวโดยสรุปได้ว่า อัตรา false positive rate จะมีค่าแตกต่างกันไปตามลำดับการลอกข้อสอบ และ วิธีการลอก ไม่มีดัชนีตัวใดที่จะใช้ได้ดีสำหรับระดับการลอกข้อสอบร้อยละ 10 ของจำนวนข้อสอบทั้งหมด สำหรับระดับร้อยละของจำนวนข้อสอบที่ถูกลอกที่มากขึ้นค่าดัชนีจะใช้ได้ดีแตกต่างกันไปตามวิธีการลอก ข้อค้นพบเพิ่มเติมของ Hanson และคนอื่นๆ คือ ระดับนัยสำคัญเชิงทฤษฎี (theoretical significance levels) ของค่าดัชนีของ Angoff ค่าดัชนีของ Frary และคณะ และ ค่าดัชนีของ Cody ไม่เข้ากันหรือไม่สอดคล้องกับอัตรา false positive rates ของข้อมูลมาตรฐาน (benchmark data) เนื่องจากระดับนัยสำคัญเชิงทฤษฎีจะทำให้ค่าระดับ false positive rates มีค่ามากเกินกว่าที่ควรจะเป็น ดังนั้น ควรที่จะใช้ระดับนัยสำคัญเชิงประจักษ์แทน (empirical significance levels) ซึ่งข้อค้นพบนี้สอดคล้องกับการศึกษาของ Stegman และ Barnhill (1982) ข้อจำกัดของการศึกษาของ Hanson และคณะ คือจุดมุ่งหมายของการศึกษาของ Hanson และ คณะ เป็นการหาว่าค่าดัชนีใดที่ควรใช้ แต่คุณลักษณะสำคัญที่เราใช้ในการประเมินอย่างหนึ่งคือ การที่ผู้ซึ่งไม่ใช่นักสถิติจะสามารถเข้าใจดัชนีได้ นอกจากนี้การศึกษาของเขายังไม่พิจารณาความง่ายของการคำนวณอีกด้วย ซึ่งจะส่งผลต่อการนำไปใช้ในสถานการณ์ที่ข้อสอบมีความยาวแตกต่างกันและใช้กับกลุ่มผู้สอบกลุ่มเล็ก

โดยผลการศึกษาข้างต้นที่กล่าวมาสามารถสรุปได้ว่า ดัชนีที่ 2 ของ Frary และคณะ (the second index of Frary et al.) เป็นค่าดัชนีที่สามารถใช้งานได้จริงในห้องเรียน หรือ ในสถานการณ์ที่คล้ายกันนี้ โดยต่อมาดัชนีที่ 2 ของ Frary และคณะ ได้ถูกพัฒนา และ เป็นที่รู้จักกันในชื่อของ g_2 ซึ่งค่าดัชนี g_2 ได้ถูกใช้ในสถาบันโพลีคลินิกและมหาวิทยาลัยของรัฐ Virginia (Virginia Polytechic Institute and State University) เป็นเวลามากกว่า 15 ปี โดยจุดประสงค์

ของการคำนวณค่าดัชนีนี้ส่วนใหญ่จะใช้ในการตรวจจับแนวโน้มของการลอกข้อสอบเพื่อตรวจดูความมีประสิทธิภาพของการคุมสอบมากกว่าที่จะคำนวณเพื่อที่จะให้ได้หลักฐานสำหรับใช้ในการพิจารณาตัดสินการลอกข้อสอบ

ได้มีการศึกษาเกี่ยวกับดัชนี g_2 ในสถานการณ์ที่หลากหลาย เช่น ผู้สอนใช้แบบสอบชุดเดียวในห้องสอบจะทำให้ได้ค่าดัชนี g_2 ที่สูงมาก ต่อมาผู้สอนจัดการสอบขึ้นอีกโดยใช้แบบสอบที่หลากหลาย (สลับลำดับของข้อสอบในแบบสอบ) ค่าดัชนี g_2 จะสูงเฉพาะคู่ผู้สอบที่ได้แบบสอบเหมือนกันและนั่งใกล้กันมากพอที่ผู้สอบคนหนึ่งจะสามารถมองเห็นคำตอบของผู้สอบอีกคนได้ โดยผลการทดลองมีความคงที่มากแม้ขนาดกลุ่มผู้สอบต่างกัน (ขนาดต่ำสุด 30 คน) และความยาวของแบบสอบที่ต่างกัน (จำนวนข้อสอบน้อยสุด 20 ข้อ) และ เนื้อหาในแบบสอบที่แตกต่างกัน จากนั้นผู้สอนต้องทำการสังเกตพฤติกรรมของคู่นักเรียนที่มีค่าดัชนี g_2 สูง เพื่อให้ได้หลักฐานยืนยันการลอกข้อสอบ (visual confirmation of copying)

ในปี 1990 แผ่นพับของสำนักงาน การบริการด้านการวัดผลและการวิจัย (office of Measurement and Research services) ได้สรุปการใช้ดัชนี g_2 โดยแผ่นพับได้ชี้แจงไว้อย่างชัดเจนว่าหลักฐานทางสถิติเพียงอย่างเดียวนั้นไม่เพียงพอที่จะตัดสินความผิดของการลอกข้อสอบ เพราะ ค่า false positive rates สามารถเกิดขึ้นได้ในทุกๆ ค่าระดับนัยสำคัญ อย่างไรก็ตามถ้าคู่ผู้สอบที่มีค่าดัชนี g_2 สูง นั้นถูกจัดให้นั่งอยู่ในบริเวณที่มีการลอกข้อสอบเกิดขึ้น โอกาสหรือความเป็นไปได้ที่ผู้สอบสองคนนี้จะลอกข้อสอบกันก็จะมีสูงมาก

Frary (1978) ทำการศึกษาการสอบในห้องเรียนของนักศึกษาจำนวน 100 คน ซึ่งมีแนวโน้มสูงมากที่จะลอกข้อสอบกัน (ค่าดัชนี g_2 มีค่ามาก) จากนั้นทำการสุ่มกลุ่มตัวอย่างจากนักศึกษา 100 คนนี้เพื่อส่งแบบสอบถามให้นักศึกษาประเมิน ซึ่งแบบสอบถามจะเกี่ยวกับบทลงโทษที่ควรจะถูกกำหนดขึ้นสำหรับการทุจริตทางการศึกษาหลากหลายรูปแบบ (รวมการลอกข้อสอบไว้ด้วย) มีอัตราการตอบกลับมากกว่าร้อยละ 90 โดยกลุ่มตัวอย่างที่มีแนวโน้มสูงที่จะลอกกันจะตอบข้อคำถามด้วยคำตอบในลักษณะผ่อนผันให้กับบทลงโทษที่ควรกำหนดขึ้น

Frary และ Olson (1985) รายงานผลการใช้ค่าดัชนี g_2 ในการตรวจจับการลอกข้อสอบ เช่น ใช้ในการตรวจจับการลอกข้อสอบของการจัดการสอบมาตรฐานสำหรับโรงเรียนประถมทั่วรัฐ โดยมีจุดประสงค์เพื่อตรวจจับระดับความหย่อนยานของการจัดการสอบซึ่งอาจก่อให้เกิดการลอกข้อสอบได้ โดยผลปรากฏว่ามีเพียงไม่กี่ห้องเรียนที่มีค่าดัชนี g_2 สูง

วิธีการหรือค่าดัชนีตรวจจับการลอกข้อสอบส่วนใหญ่จะใช้ข้อมูลของจำนวนคำตอบผิดที่ตรงกัน ซึ่งอาจจะทำให้ผู้ที่ไม่ใช่นักสถิติเกิดข้อกังขาได้ว่าคำตอบถูกที่ตรงกันอาจจะมีปริมาณมากกว่าที่คาดไว้ซึ่งอาจจะก่อให้เกิดความคลาดเคลื่อนของค่าดัชนี Buss และ Novick ได้กล่าวถึงผลกระทบที่อาจจะเกิดขึ้นได้เมื่อไม่รวมคำตอบถูกที่ตรงกันในการคำนวณค่าดัชนีต่างๆ หรือ การ

พิจารณาไว้ว่าเป็นการไม่ยุติธรรม และ อาจก่อให้การนำไปใช้เกิดความผิดพลาดทางด้านวิธีการทางสถิติ

การรวมเอาคำตอบถูกที่ตรงกัน (identical right answer) เข้าไปในการคำนวณจะทำให้ได้หลักฐานทางสถิติของการลอกข้อสอบเพิ่มเติม โดยจากค่าดัชนีต่างๆ ซึ่งได้อธิบายไว้ข้างต้นมีเพียงค่าดัชนี g_2 ของ Frary และคณะ (Frary et al.) เท่านั้นซึ่งใช้ทั้งคำตอบถูกและผิดที่ตรงกัน แต่ค่าดัชนีอื่นๆ ก็สามารถประยุกต์นำเอาคำตอบถูกที่ตรงกันมาใช้คำนวณได้เช่นกัน แต่การใช้ข้อมูลทั้งคำตอบถูกและคำตอบผิดที่ตรงกันจะก่อให้เกิดความยุ่งยากในการอธิบายให้เกิดความเข้าใจ

โดยในยุคนี้ค่าสถิติตรวจจับการลอกข้อสอบที่ถูกคิดค้นขึ้นมีอยู่มากมาย แต่ก็จะถูกใช้กับการสอบที่มีปริมาณผู้สอบมากๆ เช่น การสอบคัดเลือกบุคคลเพื่อศึกษาต่อ เป็นต้น จะไม่ค่อยใช้ค่าดัชนีการตรวจจับการลอกข้อสอบกับการทดสอบมาตรฐานหรือการทดสอบในห้องเรียน เนื่องจากสาเหตุต่างๆ เช่น ไม่มีเครื่องมือหรืออุปกรณ์คำนวณ ไม่มีผู้เชี่ยวชาญหรือผู้มีความรู้ทางด้านดัชนีการตรวจจับการลอกข้อสอบ การไม่เห็นความจำเป็นของผลการตรวจจับการลอกข้อสอบ เป็นต้น

ในยุคปัจจุบันได้มีผู้คิดค้นค่าดัชนีตรวจจับการลอกข้อสอบขึ้นเพิ่มเติมมากมาย เช่น ดัชนี K (Holland, 1996), ค่าสถิติหรือค่าดัชนี ω (Wollack, 1997), ดัชนี K_2 (Sotaridona และ Meijer, 2002), ดัชนี S_1 และ ดัชนี S_2 (Sotaridona และ Meijer, 2003) เป็นต้น ซึ่งเป็นดัชนีตรวจจับการลอกข้อสอบที่ถูกคิดค้นและถูกศึกษาอยู่ในปัจจุบัน โดยรายละเอียดของดัชนีตรวจจับการลอกข้อสอบเหล่านี้ผู้วิจัยจะนำเสนอในหัวข้อถัดไป

2.6 แนวทางการศึกษาคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ

การศึกษาเกี่ยวกับคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบนั้นอยู่ 3 แนวทาง (Wollack, 2003) คือ

แนวทางแรก การคำนวณค่าดัชนีโดยใช้ข้อมูลจริงซึ่งได้จากการสอบซึ่งเชื่อว่าจะมีข้อมูลการลอกข้อสอบปนอยู่ แล้วนำมาเปรียบเทียบกับค่าดัชนีซึ่งคำนวณจากชุดของข้อมูลที่เชื่อว่าไม่มีการลอกข้อสอบ (Angoff, 1974; Bellezza และ Bellezza, 1989; Cody, 1985; Frary และคณะ, 1977; Holland, 1996; Kadane, 1999) นอกจากนี้ยังสามารถคำนวณโดยการเปรียบเทียบการกระจายของค่าดัชนีตรวจจับการลอกข้อสอบ 2 และเปรียบเทียบค่าเฉลี่ยและค่าดัชนีสูงสุด (extreme index) เพื่อแสดงให้เห็นว่าค่าเฉลี่ยของดัชนี และ มีค่าดัชนีสูงสุดของชุดข้อมูลที่เชื่อว่ามีการลอกข้อสอบอยู่นั้นสูงกว่า หรือ มีค่ามากกว่า ชุดข้อมูลที่เชื่อว่าไม่มีการลอกข้อสอบอยู่ (Bellezza และ Bellezza; Frary และคณะ) โดยแนวทางในการศึกษารูปแบบนี้มีจุดแข็งคือ การใช้ข้อมูลจริง (real data) ไม่มีข้อตกลงเบื้องต้นเกี่ยวกับ รูปแบบการตอบข้อคำถามของนักเรียน ธรรมชาติของผู้ลอก (copiers) และ ผู้ให้ลอก (sources) จำนวนผู้ลอก จำนวนข้อคำถามที่ถูกลอก

ข้อคำถามใดที่ถูกลอก ส่วนจุดด้อยของแนวทางการศึกษานี้ คือ เนื่องจากการศึกษาแนวทางนี้นั้นไม่ได้ควบคุมธรรมชาติของการลอกข้อสอบจึงทำให้ไม่สามารถศึกษาได้ว่าค่าดัชนีจะสามารถระบุตัวผู้ลอกได้ดีเพียงใด ภายใต้เงื่อนไขที่แตกต่างกัน เช่น ความยาวของแบบสอบ ขนาดของกลุ่มตัวอย่าง จำนวนผู้สอบ ร้อยละของจำนวนข้อคำถามที่ถูกลอก และ วิธีการลอก ปัญหาที่สองคือแนวทางการศึกษานี้ไม่สามารถตรวจวัดได้ว่าผู้สอบไม่ได้ลอกข้อสอบได้อย่างสมบูรณ์แน่นอน ซึ่งการที่ผู้คุมสอบไม่เห็นว่าผู้สอบลอกกันไม่เพียงพอที่จะสันนิษฐานว่าผู้สอบไม่ได้ลอกกัน สุดท้ายแนวทางการศึกษาแบบนี้เป็นการวิเคราะห์ข้อมูลที่ค่อนข้างน้อย ทำให้ได้ข้อมูลหรือคำตอบที่ค่อนข้างน้อยด้วย

แนวทางที่สอง คือ การจำลองสถานการณ์ (simulation) (Sotaridona และ Meijer, 2002, 2003; Wollack, 1997) การศึกษาแนวทางนี้ทำการจำลองข้อมูลการสอบ และ การลอกข้อสอบตามโมเดลการสอบที่ต้องกำหนดขึ้นก่อน (pre-specified test model) ภายใต้เงื่อนไขที่แตกต่างกัน และ พิจารณาค่าความคลาดเคลื่อนประเภทที่ 1 (Type I error) และกำลังของดัชนีชนิดต่างๆ ภายใต้เงื่อนไขตัวแปรต่างๆ การศึกษาโดยการจำลองสถานการณ์มีข้อดี คือ ผู้วิจัยสามารถควบคุมเงื่อนไขต่างๆ ที่สนใจได้อย่างเต็มที่ ข้อเสีย คือ ผลที่ได้มีเงื่อนไขอยู่บนการเป็นไปตามข้อตกลงเบื้องต้นของการจำลองสถานการณ์ และที่สำคัญ แนวทางการศึกษานี้ข้อมูลจะถูกทำให้สอดคล้องกับโมเดลการวัด (measurement model) ที่ใช้ในการจำลองคำตอบของข้อคำถามซึ่งในความเป็นจริง ข้อมูลอาจจะไม่สอดคล้องกับโมเดลก็ได้

แนวทางที่สาม คือ การใช้ข้อมูลจริงซึ่งถูกรวบรวมด้วยวิธีการที่ชุดข้อมูลจะไม่มีผู้สอบที่ลอกข้อสอบจากผู้สอบคนอื่นในชุดข้อมูลนั้นได้ (Bay, 1995; Hanson และคณะ, 1987; Iwamoto, Nungester, Watson และ Luecht, 1997) ชุดข้อมูลที่ใช้ในการศึกษามักจะได้มาจากการรวมกันของชุดข้อมูลที่ได้จากการสอบในสถานที่และเวลาที่ต่างกัน การศึกษาแนวทางนี้มีข้อดีทั้งหมดที่เคยกล่าวมาข้างต้นเนื่องจากใช้ข้อมูลจริง โดยเราจะทราบคุณสมบัติทางสถิติบางอย่างล่วงหน้าได้ นั่นคือ ชุดข้อมูลซึ่งไม่มีข้อมูลจากการลอก การลอกจะถูกศึกษาโดยการจับคู่ผู้สอบอย่างสุ่ม จากนั้นเปลี่ยนคำตอบของข้อคำถามที่ได้จากการสุ่มของผู้สอบคนหนึ่งให้ตรงกับคำตอบของผู้สอบอีกคนหนึ่ง แต่การศึกษาแนวทางนี้มักที่จะขาดความซับซ้อนของการศึกษาแบบการจำลองสถานการณ์สัมบูรณ์ (pure simulated studies)

2.7 เกณฑ์สำหรับเปรียบเทียบคุณภาพของดัชนีตรวจจับการลอกข้อสอบ

คุณภาพของดัชนีตรวจจับการลอกข้อสอบนิยมพิจารณาจากดัชนีบ่งชี้สำคัญ 2 ตัว ได้แก่ อำนาจการทดสอบ หรือ อัตราความถูกต้องของการตรวจพบและระบุผู้ลอกจริงว่าเป็นผู้ลอกข้อสอบ (Power rate หรือ Detection rate) และ ความคลาดเคลื่อนของการตรวจจับ หรือ อัตรา

ความคลาดเคลื่อนประเภทที่ 1 (Type I error rate) ซึ่งเป็นโอกาสของการเกิดความคลาดเคลื่อนในขณะที่ตรวจจับซึ่งพบว่าผู้ที่ไม่ใช่ผู้ลอกอย่างแท้จริง แต่ถูกระบุโดยดัชนีตรวจจับการลอกข้อสอบว่าเป็นผู้ลอก

ในการคำนวณค่าดัชนีตรวจจับการลอกข้อสอบ มีจุดมุ่งหมายเพื่อทดสอบนัยสำคัญของผลการตรวจจับ โดยมีสมมติฐานศูนย์ของการทดสอบ คือ ผู้สอบที่ถูกสงสัยไม่ได้เป็นผู้ลอกข้อสอบ โดยผลการทดสอบสมมติฐานนำไปสู่การตัดสินใจว่า ยอมรับสมมติฐานศูนย์ (Accept H_0) หรือปฏิเสธสมมติฐานศูนย์ (Reject H_0) ผลของการตัดสินใจมีโอกาสเกิดเหตุการณ์ 4 ลักษณะ ดังตาราง 1

ตาราง 1 คุณภาพของการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ

การตัดสินใจตามผลการตรวจจับ	H_0: ผู้สอบที่ถูกสงสัยไม่ได้เป็นผู้ลอกข้อสอบ	
	ความเป็นจริง	
	H_0 ถูก	H_0 ผิด
ยอมรับ H_0	• ตัดสินถูก (True negative) • ระดับความเชื่อมั่น (1-α)	• ตัดสินผิด (Type II Error, β) • False negative
ปฏิเสธ H_0	• ตัดสินผิด (Type I Error, α) • False positive	• ตัดสินถูก (True positive) • อำนาจการทดสอบ (1-β)

ค่าดัชนีตรวจจับการลอกข้อสอบที่คำนวณได้จะนำไปสู่การตัดสินใจสรุปผลการตรวจจับดังนี้

1) ตัดสินใจถูก มีโอกาสเกิดขึ้น 2 ลักษณะ คือ การสรุปถูกว่า
   - ผู้สอบที่ถูกสงสัยไม่ได้เป็นผู้ลอกข้อสอบตามความเป็นจริง (True negative)
   - ผู้สอบที่ถูกสงสัยเป็นผู้ลอกข้อสอบตามความเป็นจริง (True positive)
2) ตัดสินผิด มีโอกาสเกิดขึ้น 2 ลักษณะ คือ การสรุปผิดว่า
   - ผู้สอบที่ถูกสงสัยเป็นผู้ลอกข้อสอบ ทั้งๆ ที่ความจริงไม่ได้เป็นผู้ลอกข้อสอบ (False positive)
   - ผู้สอบที่ถูกสงสัยไม่ได้เป็นผู้ลอกข้อสอบ ทั้งๆ ที่ความจริงเป็นผู้ลอกข้อสอบ (False negative)

เนื่องจากอำนาจการทดสอบ (1-β) กับ β เป็นค่าดัชนีซึ่งมีสเกลที่ผกผัน และ ความคลาดเคลื่อนประเภทที่ 1(α) กับ 1-α ต่างก็เป็นดัชนีสเกลที่ผกผันเช่นเดียวกัน ดังนั้น การพิจารณาดัชนีบ่งชี้คุณภาพ 2 ตัว คือ อำนาจการทดสอบ หรือ อัตราความถูกต้องของการตรวจพบและระบุผู้ลอกจริงว่าเป็นผู้ลอกข้อสอบ (Power rate หรือ Detection rate) และ อัตราความคลาดเคลื่อนประเภทที่ 1 (Type I Error rate) ก็เพียงพอที่จะให้สารสนเทศครบทั้ง 4 เหตุการณ์

2.8 ทฤษฎีของดัชนีตรวจจับการลอกข้อสอบที่ใช้ในการวิจัย

ในส่วนนี้จะนำเสนอทฤษฎีของดัชนีตรวจจับการลอกข้อสอบที่ใช้ในการวิจัยครั้งนี้อย่างละเอียด โดยจะนำเสนอเป็นลำดับดังนี้

2.8.1 ดัชนี K (The K-index)

2.8.2 ดัชนี ω

2.8.3 ดัชนี S_1

2.8.4 ดัชนี S_2

2.8.1 ดัชนี K (The K-index)

ดัชนี K คือ สถิติที่สามารถใช้ในการประเมินระดับความสอดคล้องซึ่งผิดปกติของคำตอบผิดของการสอบข้อสอบปรนัยระหว่างผู้สอบ 2 คน คือ ผู้ให้ลอก (Source: s) และ ผู้ลอก (Copier: c) ผู้ลอกนั้นถูกสงสัยว่าเป็นผู้ลอกคำตอบจากผู้ให้ลอก โดยดัชนี K จะนำเฉพาะคำตอบผิดของผู้สอบมาคำนวณเท่านั้น (Holland,1996)

โดยในที่นี้ผู้วิจัยจะนำเสนอรายละเอียดเกี่ยวกับดัชนี K เป็นลำดับหัวข้อต่อไปนี้

2.8.1.1 ดัชนี K ที่มีพื้นฐานจากการกระจายเชิงประจักษ์ (Empirical Distribution)

2.8.1.2 ดัชนี K ที่มีพื้นฐานจากการประมาณทางทฤษฎี (Theoretical Approximation)

2.8.1.3 ดัชนี K ที่ใช้ในยุคปัจจุบัน

คำชี้แจง

สัญลักษณ์ที่จะถูกใช้ในทฤษฎีพื้นฐานของดัชนีตรวจจับการลอกข้อสอบ มีดังนี้ คือ

j (j = 1,…. , J) คือ ผู้สอบ

i (i = 1, …, I) คือ ข้อสอบ

v (v = 1, … , V) คือ ประเภทของคำตอบของข้อสอบ หรือ ตัวเลือกของข้อสอบ

s คือ ผู้สอบที่ถูกกำหนดให้เป็น ผู้ให้ลอก (Source)

c คือ ผู้สอบที่ถูกสงสัยว่าจะลอกข้อสอบจาก s (Copier)

w_j คือ จำนวนของคำตอบผิด ของผู้สอบ j

M หรือ m คือ จำนวนของคำตอบผิดที่เหมือนกันระหว่างผู้สอบ j และ s

r = 1, ... , c' , ... R คือ กลุ่มย่อยของผู้สอบที่แต่ละกลุ่มย่อยจะมีจำนวนของคำตอบผิดซึ่งแตกต่างกัน และ c' คือกลุ่มซึ่งผู้สอบ c อยู่

$j' = 1, \ldots, n_r$ คือ ผู้สอบในกลุ่มย่อย r ซึ่งแต่ละกลุ่มย่อยมีผู้สอบอย่างน้อย 1 คน และ

$$\sum_{r=1}^{R} n_r = J - 1$$

$M_r = (M_{r1}, \ldots, M_{rj}, \ldots, M_{m_r})$ คือ เวกเตอร์ของจำนวนของคำตอบผิดที่เหมือนกัน (matching wrong answer) ในกลุ่มย่อยเฉพาะ r

$M_{c'} = (M_{c'1}, \ldots, M_{c'n_{c'}})$ คือ เวกเตอร์ของจำนวนคำตอบผิดที่เหมือนกัน (matching wrong answer) ของผู้สอบ $n_{c'}$ คนในกลุ่มย่อย c' เมื่อกลุ่มย่อย c' ประกอบด้วยผู้สอบที่มีจำนวนคะแนนผิด (incorrect score) เท่ากันกับผู้ลอก

$Q_r = \frac{w_r}{I}$ คือ สัดส่วนของคำตอบผิดของกลุ่มย่อย r เมื่อ I คือ จำนวนข้อสอบทั้งหมดของแบบสอบ

2.8.1.1 **ดัชนี K ที่มีพื้นฐานจากการกระจายเชิงประจักษ์**
(K-Index Based on the Empirical Distribution)

ดัชนี K สามารถคำนวณโดยใช้ข้อมูลเชิงประจักษ์ (empirical data) ของผู้สอบ J คน ซึ่งตอบข้อสอบจำนวน I ข้อ เพื่อคำนวณดัชนี K ซึ่งมีพื้นฐานจากข้อมูลเชิงประจักษ์ ขั้นตอนแรกเราต้องพิจารณากลุ่มของผู้สอบที่มีจำนวนคะแนนผิดเหมือนกันกับผู้ลอก (กลุ่มย่อย c') จากนั้นสำหรับผู้สอบแต่ละคนในกลุ่มย่อย c' เราจะพิจารณาจำนวนของข้อสอบที่คำตอบผิดของผู้สอบในกลุ่มตรงกับคำตอบผิดของผู้ให้ลอก ซึ่งก็คือ เวกเตอร์ $M_{c'}$ และการกระจายของ $M_{c'}$ ซึ่งประกอบด้วยการกระจายแบบสอดคล้องเชิงประจักษ์ (empirical agreement distribution) สำหรับผู้สอบ c เราจะกำหนด $m_{cc'}$ ให้เป็นจำนวนของคำตอบผิดที่ตรงกันระหว่าง c กับ s และตัวแปรสุ่ม M_{rj} จะถูกเขียนอย่างง่ายๆ เป็น M ถ้าไม่มีความจำเป็นต้องระบุกลุ่มที่ผู้สอบ j เป็นสมาชิกอยู่

ดัชนี K ถูกจำกัดความไว้ว่า คือ สัดส่วนของผู้สอบที่มีจำนวนคะแนนผิดเหมือนกับผู้ลอกเมื่อจำนวนของคะแนนผิดที่เหมือนกันกับผู้ให้ลอกจะต้องมีค่ามากกว่าหรือเท่ากับ $m_{cc'}$ เป็นอย่างน้อย

สำหรับ $J' = 1, \ldots, n_{c'}$ และ ให้ $I_{c'j'}$ เป็นตัวแปรชี้นำ (indicator variable) โดยกำหนดให้เป็น 1 สำหรับ $m_{c'j'} \geq m_{c'c'}$ และ เป็น 0 ในกรณีอื่นๆ เขียนได้เป็น

$$K = \frac{\sum_{j=1}^{n_{c'}} I_{c'j'}}{n_{c'}} \qquad (1)$$

เมื่อค่า K มีค่าน้อย นั่นหมายความว่า จากผลการคำนวณชี้ให้เห็นว่ามีหลักฐานทางสถิติที่ผู้สอบ c จะลอกข้อสอบจากผู้สอบ s

โดยทั่วไปจำนวนของคะแนนผิดที่เหมือนกันจะขึ้นอยู่กับระดับความสามารถของ s และ c จำนวนของคำตอบผิดที่ตรงกันจะน้อยเมื่อไม่ s ก็ c หรือ ทั้งคู่มีคะแนนถูกมาก (มีระดับความสามารถสูง) ในขณะที่จะมีค่ามาก เมื่อผู้สอบทั้งคู่มีคะแนนผิดมาก (ระดับความสามารถต่ำ)

ดัชนี K คำนวณอย่างมีเงื่อนไขอยู่บนจำนวนคะแนนผิดที่ตรงกันของผู้ซึ่งถูกสงสัยว่าจะเป็นผู้ลอก โดยจำนวนของคะแนนผิดที่ตรงกัน (M) ต้องขึ้นอยู่กับหรือมีความสัมพันธ์กับระดับความสามารถของประชากรผู้สอบดังที่กล่าวไว้ข้างต้น ดังนั้นสิ่งที่ตามมาก็คือจำนวนของผู้สอบที่ถูกใช้ในการคำนวณจริงของดัชนี K (กลุ่มย่อย c') อาจจะน้อยมาก ซึ่งจำนวนของผู้สอบในกลุ่มย่อย c' จะมีผลต่อความถูกต้องของค่าดัชนี K ดังนั้นเมื่อขนาดกลุ่มตัวอย่างเล็ก (J=100) แทนที่เราจะคำนวณค่าดัชนี K โดยมีพื้นฐานอยู่บนการกระจายแบบสอดคล้องเชิงประจักษ์ (empirical agreement distribution) แล้วทางเลือกหนึ่งที่จะสามารถใช้ในการคำนวณค่าดัชนี K ได้ดีกว่าคือการใช้การประมาณทางทฤษฎี (Theoretical Approximations)

2.8.1.2 ดัชนี K ที่มีพื้นฐานจากการประมาณทางทฤษฎี (Theoretical Approximations)

การใช้ดัชนี K สิ่งที่ต้องระบุเป็นอย่างแรก คือ ความคลาดเคลื่อนประเภทที่ 1 (Type 1 error: α) ซึ่งถูกนิยามเป็น ความเป็นไปได้ในการระบุผิดว่าผู้สอบเป็นผู้ลอก ตามทฤษฎีแล้ว เราต้องการที่จะได้ค่าสถิติที่ ค่าความคลาดเคลื่อนประเภทที่ 1 ที่ใช้เป็นเกณฑ์ (nominal type I error) และ ค่าความคลาดเคลื่อนประเภทที่ 1 เชิงประจักษ์ (empirical type I error) ที่มีค่าคล้ายกัน หรืออาจกล่าวได้ว่าการนำไปใช้นั้นเราไม่ต้องการสถิติที่ค่าความคลาดเคลื่อนประเภทที่ 1 เชิงประจักษ์มีค่ามากกว่าค่าความคลาดเคลื่อนประเภทที่ 1 ที่ใช้เป็นเกณฑ์จนเกินยอมรับได้(liberal) สาเหตุเพราะว่าผลที่ตามมาเนื่องจากการระบุผิดพลาดว่าผู้ที่เป็นผู้สอบสุจริตเป็นผู้ลอกข้อสอบมีมากและรุนแรง

ดัชนีตรวจจับการลอกข้อสอบที่ไม่สามารถควบคุมระดับ ความคลาดเคลื่อนประเภทที่ 1 เชิงประจักษ์ (empirical type I error) ให้อยู่ต่ำกว่าความคลาดเคลื่อนประเภทที่ 1 ที่ใช้เป็นเกณฑ์ (nominal type I error) ได้นั้นเป็นดัชนีที่ไม่ควรนำมาใช้ ในขณะเดียวกันดัชนีตรวจจับการลอกข้อสอบก็ไม่ควรที่จะมีค่าความคลาดเคลื่อนประเภทที่ 1 เชิงประจักษ์ (empirical type I error) ต่ำกว่าความคลาดเคลื่อนประเภทที่ 1 ที่ใช้เป็นเกณฑ์ (nominal type I error) มากๆ (conservative) เพราะ อำนาจในการตรวจจับผู้สอบซึ่งเป็นผู้ลอกข้อสอบจริงๆ จะต่ำมาก (Seaman, Levin และ Serlin, 1991 อ้างถึงใน Wollack, 1997)

โดยทั่วไปข้อด้อยของการใช้การกระจายเชิงประจักษ์แบบไม่ต่อเนื่อง (discrete empirical distribution) สำหรับกลุ่มตัวอย่างขนาดเล็ก คือ ตัวแปรสุ่ม M จะมีจำนวนไม่มากพอซึ่งจะส่งผลให้ไม่สามารถคำนวณค่าความคลาดเคลื่อนประเภทที่ 1 ในระดับนัยสำคัญที่ถูกระบุไว้ก่อน เช่น ขนาด 0.01 ได้ (Agresti, 1996: 43)

Holland (1996) กล่าวว่าการกระจายของตัวแปร M สามารถที่จะประมาณได้โดยการแจกแจงทวินาม (binomial distribution) นั่นคือ $M \overset{approx}{\sim} B(w_s, p)$ เมื่อ w_s คือจำนวนคำตอบผิดของผู้ให้ลอกที่เราทราบค่า และ p คือค่าที่เราไม่ทราบค่า Holland (1996) แนะนำ 2 วิธีการในการประมาณค่า p

<u>วิธีการแรก</u> ตัวแปร p จะถูกคำนวณโดยมีข้อตกลงเบื้องต้นว่า การแจกแจงทวินาม และการกระจายแบบ empirical ของตัวแปร M มีค่าเฉลี่ยเท่ากัน ให้ $\overline{m}_{c'}$ คือ ค่าเฉลี่ยของการกระจายแบบสอดคล้องเชิงประจักษ์ (empirical agreement distribution) ซึ่งมีค่าเท่ากับ

$$\overline{m}_{c'} = \frac{\sum_{j=1}^{n_{c'}} m_{c'j}}{n_{c'}} \qquad (2)$$

ดังนั้นค่าประมาณของ p ซึ่งเขียนได้เป็น $p^*_{c'}$ มีค่าเป็น

$$p^*_{c'} = \frac{\overline{m}_{c'}}{w_s} \qquad (3)$$

ให้ K^* คือ ดัชนี K ซึ่งคำนวณจากการใช้ $p^*_{c'}$ ดังนั้น K^* เขียนได้เป็น

$$K^* = P(M \geq m_{c'c}) = \sum_{g=m_{c'c}}^{w_s} \binom{w_s}{g} (p^*_{c'})^g (1-p^*_{c'})^{w_s-g} \qquad (4)$$

Holland (1996) กล่าวว่าการใช้ข้อมูลเชิงประจักษ์ขนาดใหญ่ซึ่งมีการแจกแจงแบบทวินามและมีการใช้การประมาณค่า p^*_c ทำให้ได้การประมาณสำหรับการกระจายแบบสอดคล้องเชิงประจักษ์ที่ conservative หรือกล่าวได้ว่า ดัชนี K ที่มีพื้นฐานจากการประมาณค่าแบบ binomial มักที่จะมีค่าสูงกว่าดัชนี K ที่มีพื้นฐานจากการกระจายเชิงประจักษ์ (the empirical distribution) (Agresti, 1990: 9)

<u>วิธีการที่สอง</u> เนื่องจากการคำนวณ p^*_c เราจะต้องมีข้อมูลรูปแบบการตอบข้อสอบของผู้สอบในกลุ่มย่อย c' โดยค่าของ p^*_c จะได้รับผลกระทบจากขนาดของกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่างที่เล็กจะทำให้ได้ค่าประมาณของ p^*_c ที่มีความน่าเชื่อถือน้อย Holland (1996) แนะนำวิธีการ การประมาณค่า p^*_c โดยการใช้ การวิเคราะห์ถดถอยเชิงเส้น (linear regression) ซึ่งเป็นวิธีการที่ 2 โดยการใช้สัดส่วนของคำตอบผิด (Q_r) ของผู้สอบแต่ละคนในแต่ละกลุ่มย่อยที่มีคะแนนผิดต่างกัน r =1, …, R เป็นตัวทำนาย (regressors) จากการใช้ข้อมูลขนาดใหญ่จากศูนย์ ETS นั้น Holland (1996) แสดงให้เห็นว่า p^*_r ที่ถูกนิยามให้คล้ายคลึงกับสมการ 3 และมีความสัมพันธ์เชิงเส้นกับ Q_r ให้ $\hat{p}_r$ คือค่าประมาณของความเป็นไปได้แบบ binomial ของ p^*_r โดยใช้ Q_r เป็นตัวทำนาย จะสามารถเขียน $\hat{p}_r$ เป็นฟังก์ชันเชิงเส้นแยกเป็น 2 ส่วนโดย a และ b คือ พารามิเตอร์ จุดตัดและ ความชัน ตามลำดับ

$$\hat{p}_r = \begin{cases} a + bQ_r & \text{ถ้า } 0 < Q_r \le 0.3 \\ [a + .3b] + .4b[Q_r - .3] & 0.3 < Q_r \le 1 \end{cases} \qquad (5)$$

ค่า a และ b ต้องถูกระบุในการประมาณค่า $\hat{p}_r$ ในสมการที่ 5 โดย Holland (1996) ใช้ค่า a = 0.085 และใช้ค่า b ที่แตกต่างกันขึ้นอยู่กับชุดข้อสอบที่ใช้ อย่างไรก็ตามการศึกษาของเขายังไม่เป็นที่แน่ชัดว่าค่าต่างๆ เหล่านี้ได้มาได้อย่างไร นอกจากนี้ค่าพวกนี้ยังแปรเปลี่ยนไปตามชุดข้อสอบอีกด้วย

2.8.1.3 ดัชนี K ที่ใช้ในยุคปัจจุบัน

การศึกษาในปัจจุบัน Sotaridona และ Meijer (2002) เสนอ $\hat{p}^*_1$ และ $\hat{p}^*_2$ ($\hat{p}_r$ คือค่าประมาณของความเป็นไปได้แบบ binomial ของ p^*_r โดยใช้ Q_r เป็นตัวทำนาย) ซึ่งเป็นค่าประมาณของ p^*_r ที่มีพื้นฐานจากการวิเคราะห์ถดถอยเชิงเส้นและกำลังสอง (linear and quadratic regression) จากการประมาณค่า p^* สองแบบทำให้ได้ค่าดัชนี K สองแบบคือ $\overline{K}_1$ และ $\overline{K}_2$ (ในงานวิจัยชิ้นนี้เขียนย่อเป็น K_2) ซึ่งเขียนได้ดังนี้

$$\overline{K}_1 = P(M \geq m_{c'c}) = \sum_{g=m_{c'c}}^{w_s} \binom{w_s}{g} (\hat{p}_1^*)^g (1-\hat{p}_1^*)^{w_s-g} \quad (6)$$

และ

$$\overline{K}_2 = P(M \geq m_{c'c}) = \sum_{g=m_{c'c}}^{w_s} \binom{w_s}{g} (\hat{p}_2^*)^g (1-\hat{p}_2^*)^{w_s-g} \quad (7)$$

ซึ่งการประมาณค่า p แบบ $p_{c'}^*$ จะใช้ข้อมูลที่เกี่ยวข้องกับผู้สอบที่อยู่ในกลุ่มย่อย c' เท่านั้น ในขณะที่ $\hat{p}_1^*$ และ $\hat{p}_2^*$ จะใช้ข้อมูลที่เกี่ยวข้องจากกลุ่มย่อย R กลุ่ม ดังนั้น $\hat{p}_1^*$ และ $\hat{p}_2^*$ น่าจะให้การประมาณค่า p ที่ดีกว่า $p_{c'}^*$

โดยการประมาณค่า $\hat{p}_1^*$ และ $\hat{p}_2^*$ สามารถประมาณได้จาก

$$\hat{p}_1^* = \beta_0 + \beta_1 Q_r + \varepsilon_r$$

$$\hat{p}_2^* = \beta_0 + \beta_1 Q_r + \beta_2 Q_r^2 + \varepsilon_r$$

โดย β_0 คือ intercept parameter β_1 คือ slope parameter

β_2 คือ regression parameter ε_r คือ error term

Q_r คือ สัดส่วนของจำนวนข้อผิดของผู้สอบในกลุ่มย่อยที่มีจำนวนข้อผิด r

โดยขั้นตอนในการหาค่าพารามิเตอร์ต่างๆ ในสมการที่ $\hat{p}_2^* = \beta_0 + \beta_1 Q_r + \beta_2 Q_r^2 + \varepsilon_r$ คือ เมื่อทำการตรวจให้คะแนนผู้สอบทุกคนแล้ว จัดผู้สอบที่มีจำนวนคะแนนผิดเท่ากันให้อยู่ในกลุ่มเดียวกัน เราจะได้กลุ่มผู้สอบย่อยๆ จำนวน R กลุ่มซึ่งแต่ละกลุ่มมีคะแนนผิดที่แตกต่างกันระหว่างกลุ่ม แต่จะมีคะแนนผิดเท่ากันภายในกลุ่มย่อยเดียวกันซึ่งผู้วิจัยจะคำนวณค่าสัดส่วนของจำนวนข้อผิดของผู้สอบในแต่ละกลุ่มย่อยที่มีจำนวนข้อผิด r (Q_r) จากนั้นในแต่ละกลุ่มย่อยให้หาค่า เวกเตอร์ของจำนวนของคำตอบผิดที่เหมือนกัน (M_r) สำหรับผู้สอบทุกคนในกลุ่มย่อย จากนั้นนำค่า M_r ของแต่ละคนในกลุ่มย่อยนั้นมาหาค่าเฉลี่ย ($\overline{M}_r$) ซึ่งเขียนเป็นความสัมพันธ์ได้ว่า

$$\overline{M}_r = \frac{\sum_{j=1}^{n_r} M_{rj}}{n_r}$$

พอเราคำนวณค่า $\overline{M}_r$ ของทุกกลุ่มย่อยเสร็จแล้วเราก็นำค่า $\overline{M}_r$ ที่คำนวณได้ไปหารด้วยจำนวนคำตอบผิดของผู้ให้ลอก (w_s) ซึ่งเขียนเป็นความสัมพันธ์ได้ว่า

$$p_r^* = \frac{\overline{M}_r}{w_s}$$

จากนั้นเราจะนำเอาค่า p_r^* และ Q_r ซึ่งคำนวณได้ไปวิเคราะห์ถดถอยกำลังสอง (quadratic regression) โดยใช้ Q_r เป็นตัวทำนาย เพื่อให้ได้ค่าพารามิเตอร์ β_0, β_1, β_2, ε_r แล้วนำค่าพารามิเตอร์ต่างๆ ไปใส่ในสมการ $\hat{p}_2^* = E(\beta_0 + \beta_1 Q_r + \beta_2 Q_r^2 + \varepsilon_r)$ เพื่อคำนวณหาค่า $\hat{p}_2^*$ เพื่อนำไปใช้ในการคำนวณดัชนี K_2

หากค่าดัชนี K ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้ว ผู้สอบจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก

2.8.2 **ดัชนี** ω

ผู้วิจัยจะนำเสนอรายละเอียดเกี่ยวกับดัชนี ω เป็นลำดับหัวข้อต่อไปนี้

2.8.2.1 การคำนวณดัชนี ω

2.8.2.2 ความแตกต่างระหว่างดัชนี ω และ ดัชนี g_2

2.8.2.3 ความแตกต่างระหว่างดัชนี ω และ ดัชนี K

2.8.2.1 **การคำนวณดัชนี** ω

Wollack (1997) เสนอดัชนีการตรวจจับการลอกข้อสอบ ω ซึ่งมีฐานแนวคิดอยู่บน nominal response model (NRM) (Bock, 1972)

ดัชนี ω เปรียบเทียบจำนวนของข้อสอบที่ตอบตรงกันระหว่างผู้ลอกกับผู้ให้ลอก กับจำนวนของข้อสอบที่ตอบตรงกันระหว่างผู้ลอกกับผู้ให้ลอกที่คาดว่าน่าจะเกิดขึ้นเมื่อผู้สอบทั้งสองคนตอบข้อคำถามอย่างเป็นอิสระต่อกัน (ผู้สอบทั้งสองไม่ได้ลอกกัน)

ดัชนี ω จะมีพื้นฐานอยู่บนการประมาณค่าความเป็นไปได้ที่ผู้สอบจะเลือกตัวเลือกต่างๆ คือ ตัวเลือกที่ถูก และ ตัวเลือกที่ผิด ซึ่งในการคำนวณดัชนี ω โมเดล NRM จะถูกใช้ในการประมาณความเป็นไปได้ที่ผู้สอบจะตอบหนึ่งในตัวเลือกหรือประเภทคำตอบของข้อคำถาม (the item response categories) $v\left[=1,...,h,...,V\right]$ ภายใต้โมเดล NRM ความเป็นไปได้ของผู้สอบ j ที่มีระดับความสามารถ θ_j ที่เลือกตัวเลือก h ของข้อคำถาม i ซึ่งพารามิเตอร์จุดตัดและความชันเป็น ζ_{ih} และ λ_{ih} สามารถเขียนได้ดังนี้

$$P_{ih}(\theta_j) = \frac{\exp(\zeta_{ih} + \lambda_{ih}\theta_j)}{\sum_{v=1}^{V} \exp(\zeta_{iv} + \lambda_{iv}\theta_j)} \qquad (8)$$

ความคิดรวบยอด (concept) ของดัชนี ω สามารถเขียนได้ดังนี้

$$\text{Index} = \frac{\left[\left(\text{Number of matches}\right) - \left(\text{Expected number of matches}\right)\right]}{\text{Standard error}}$$

ถ้าให้ h_{cs} เป็นจำนวนของข้อคำถามที่คำตอบระหว่าง s และ c มีลักษณะเหมือนกันทุกประการ ให้ $E\left(h_{cs}|\theta_c, U_s, \xi\right)$ เป็นค่าที่คาดหวังของ h_{cs} ซึ่งมีเงื่อนไขเกี่ยวกับระดับความสามารถของผู้ลอก (θ_c) เวกเตอร์คำตอบของข้อคำถามของผู้ให้ลอก (U_s) และ พารามิเตอร์ของข้อสอบ (ξ) และให้ $\sigma_{h_{cs}}$ เป็นส่วนเบี่ยงเบนมาตรฐานของ h_{cs} แล้วเราสามารถเขียนแสดงความสัมพันธ์ของดัชนี ω ได้ดังนี้

$$\omega = \frac{h_{cs} - E\left(h_{cs}|\theta_c, U_s, \xi\right)}{\sigma_{h_{cs}}} \tag{9}$$

เมื่อ

$$E\left(h_{cs}|\theta_c, U_s, \xi\right) = \sum_{i=1}^{I} P\left(u_{ic} = u_{is}|\theta_c, U_s, \xi\right)$$

และ

$$\sigma_{h_{cs}} = \sqrt{\sum_{i=1}^{I}\left[P\left(u_{ic} = u_{is}|\theta_c, U_s, \xi\right)\right]\left[1 - P\left(u_{ic} = u_{is}|\theta_c, U_s, \xi\right)\right]}$$

ในการคำนวณดัชนี ω นั้นจะใช้โมเดล NRM คำนวณความเป็นไปได้ของ c ในการเลือกคำตอบเดียวกับ s ในทุกๆ ข้อคำถาม

สำหรับคู่ของผู้สอบ s และ c ใดๆ ที่มีการตอบข้อคำถามอย่างเป็นอิสระต่อกันการกระจายของ ω จะมีลักษณะเป็นปกติมาตรฐาน (standard normal distribution) ซึ่งมีค่าเฉลี่ยเป็นศูนย์และส่วนเบี่ยงเบนมาตรฐานเป็นหนึ่ง (Wollack, 1997) ดังนั้น ดัชนี ω ที่คำนวณได้จะสามารถประเมินหรือเปรียบเทียบที่ระดับนัยสำคัญทางสถิติโดยการใช้การกระจายแบบปกติมาตรฐานได้

เมื่อดัชนี ω ที่คำนวณได้มีค่ามากกว่าค่าวิกฤติของการทดสอบสมมติฐานทางเดียว (one-tailed critical value) โดยขอบเขตวิกฤติมีพื้นที่อยู่ทางด้านขวาของโค้งปกติแล้วผู้สอบจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก ค่าสถิติ ω ที่เป็นค่าบวกมากๆ จะสื่อถึงความเป็นไปได้ที่น้อยว่าจำนวนของข้อคำถามที่ผู้ลอกและผู้ให้ลอกตอบตรงกันจะเกิดขึ้นโดยบังเอิญ หรือ อาจจะกล่าวได้ว่า ดัชนี

ω ที่มากบ่งบอกถึงความแข็งแกร่งที่มากขึ้นของหลักฐานทางสถิติที่ระบุว่าคู่ผู้สอบนั้นๆ ลอกข้อสอบกัน

2.8.2.2 **ความแตกต่างระหว่างดัชนี ω และ ดัชนี g_2**

ดัชนี ω มีความคล้ายคลึงกับ ดัชนี g_2 ที่ถูกเสนอโดย Frary, Tideman และ Watts (1977) ข้อแตกต่างระหว่างดัชนีทั้งสองค่า คือ วิธีการในการคำนวณ ค่าคาดหวัง h_{cs} สำหรับ ω จะใช้ โมเดล NRM โดยมีเงื่อนไขเกี่ยวกับค่า θ_c, U_s และ ξ ในขณะที่ ดัชนี g_2 จะใช้ตัวลวงของข้อสอบ และ ค่าความยากจากทฤษฎีการตอบสนองข้อสอบแบบดั้งเดิม (Classical test theory) และ อัตราส่วนของจำนวนคะแนนถูกของผู้ลอกกับจำนวนคะแนนถูกเฉลี่ยของผู้สอบทุกคน

Wollack (1997) ได้ทำการศึกษาเปรียบเทียบ ความคลาดเคลื่อนประเภทที่ 1 แบบ empirical และ อำนาจในการตรวจจับการลอกข้อสอบระหว่างดัชนี ω และ ดัชนี g_2 ผลลัพธ์แสดงให้เห็นว่า ดัชนี ω มีประสิทธิภาพดีกว่าดัชนี g_2 ภายใต้สถานการณ์ที่ถูกสร้างขึ้น โดยเฉพาะ ดัชนี g_2 ไม่สามารถควบคุมระดับของ ความคลาดเคลื่อนประเภทที่ 1 ไว้ได้ซึ่งพบว่าการควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 มีลักษณะ liberal ในทุกสถานการณ์ที่ศึกษา

2.8.2.3 **ความแตกต่างระหว่างดัชนี ω และ ดัชนี K**

แม้ว่าทั้งดัชนี K และดัชนี ω จะใช้ความคล้ายคลึงกันของการตอบข้อคำถาม แต่ดัชนี ω เปรียบเทียบคำตอบของผู้ลอกกับเวกเตอร์การตอบทั้งหมดของผู้ให้ลอก ในขณะที่ ดัชนี K คำตอบผิดของผู้ลอกจะถูกเปรียบเทียบกับคำตอบผิดของผู้ให้ลอก โดย Wollack (1996) ชี้ให้เห็นว่ากำลังของค่าสถิติซึ่งไม่ได้นำข้อมูลจากข้อคำถามที่ตอบถูกไปคิดด้วยนั้นมีแนวโน้มที่จะลดลงเนื่องจากการลดลงของจำนวนข้อคำถามที่นำไปใช้

ดัชนี ω มีพื้นฐานจากทฤษฎี IRT นั่นคือ โมเดล NRM ซึ่งมีสิ่งที่ควรพิจารณา ดังนี้ ประการแรก ข้อสันนิษฐานที่ว่าข้อมูลต้องสอดคล้อง (fit) กับโมเดล เป็นสิ่งสำคัญสำหรับ ดัชนี ω ประการที่สอง ถ้าผู้สอบที่ถูกสงสัยว่าจะลอกข้อสอบได้ทำการลอกข้อสอบจำนวนมากจากผู้ให้ลอก ระดับความสามารถของผู้ลอกจะสูงกว่าความเป็นจริงซึ่งจะส่งผลต่อค่า ดัชนี ω สุดท้ายการประมาณค่าพารามิเตอร์ที่ใช้ในโมเดล NRM ต้องการจำนวนผู้สอบจำนวนมาก (Wollack, 1997) ซึ่งจะเป็นข้อด้อยของค่าสถิตินี้ในกรณีที่ข้อมูลที่มีอยู่มีจำนวนไม่มากพอ แม้ว่า Wollack และ Cohen (1998) จะกล่าวว่าการประมาณค่าพารามิเตอร์ของข้อสอบสำหรับกลุ่มตัวอย่างขนาดเล็กเท่ากับ 100 คน สำหรับ แบบสอบที่มีข้อสอบจำนวน 40 และ 80 ข้อไม่ได้ส่งผลให้ค่าความคลาดเคลื่อนประเภทที่ 1 เพิ่มขึ้น หรือ กำลังที่ลดลง

ในทางตรงกันข้าม ดัชนี K ไม่ได้สันนิษฐานอยู่บนโมเดล IRT ดังนั้นจึงง่ายที่จะนำไปใช้ในทางปฏิบัติ อย่างไรก็ตามข้อด้อยของดัชนีนี้ก็คือ จำนวนของผู้สอบในแต่ละกลุ่มที่มีจำนวนคะแนนผิดเท่ากัน ควรมีจำนวนมากพอที่จะได้รับการประมาณค่า binomial p ที่น่าเชื่อถือ

2.8.3 **ดัชนี** S_1

สำหรับดัชนี S_1 ซึ่ง Sotaridona และ Meijer เป็นผู้คิดค้นและนำเสนอนั้น ผู้วิจัยจะนำเสนอรายละเอียดโดยแบ่งเป็นลำดับหัวข้อ ดังนี้

2.8.3.1 ที่มาของดัชนี S_1

2.8.3.2 การคำนวณหาค่าดัชนี S_1

2.8.3.3 โมเดลของการประมาณค่า μ

2.8.3.4 การตรวจสอบความสอดคล้องของโมเดล loglinear

2.8.3.1 **ที่มาของดัชนี** S_1

นักวิจัยหลายท่านที่ศึกษาเกี่ยวกับดัชนีการตรวจจับการลอกข้อสอบได้เสนอการกระจายสำหรับตัวแปรสุ่ม M แบบต่างๆ Bay (1995) ได้ใช้การกระจายแบบ compound binomial ในการพัฒนาดัชนีตรวจจับการลอกข้อสอบ B_m โดยทุกข้อคำถามในข้อสอบจะถูกนำมาใช้พิจารณา ส่วนกรณีที่เฉพาะคำตอบผิดเท่านั้นที่ถูกนำมาใช้ ได้แก่ ดัชนีตรวจจับการลอกข้อสอบ ESA (ESA copying index) (Belleza และ Belleza, 1989) ดัชนี K (K index)(Holland, 1996) และ ดัชนี K_2 ($\overline{K}_2$ index) (Sotaridona & Meijer, 2002) ซึ่งใช้การแจกแจงทวินาม สำหรับตัวแปร M Wollack (1997: 309) ได้วิจารณ์ค่าดัชนี B_m และ ESA เกี่ยวกับการที่ค่าดัชนีทั้งสองไม่สามารถปรับความน่าจะเป็นของการตอบคำถามของผู้สอบให้อยู่ในรูปฟังก์ชันของคะแนนสอบ Wollack (1997) ยังค้นพบอีกว่า ดัชนี B_m และ ESA มีอำนาจการตรวจจับการลอกข้อสอบ (detection rate) ที่ต่ำกว่าหากเทียบกับดัชนีตัวอื่นๆ ที่มีพื้นฐานมาจากทฤษฎีการตอบสนองข้อสอบแบบดั้งเดิมเหมือนกัน เช่น ดัชนี g_2 (g_2 index) (Frary และคณะ, 1977)

เนื่องจากการที่คำตอบของผู้ให้ลอกจะถูกกำหนดให้คงที่ จากนั้นเราจะทำการนับจำนวนคำตอบที่ผิดของผู้ลอกที่ตรงกับของผู้ให้ลอก ซึ่งจะแทนด้วย M โดยการแจกแจงทวินาม นั้นจะให้ค่าอำนาจของการตรวจจับ (detection rate) ที่ไม่สูงสำหรับดัชนี K และ K_2 (Lewis และ Thayer, 1998; Sotaridona และ Meijer, 2002) Sotaridona และ Meijer (2003) จึงเสนอดัชนี s_1 (s_1 index) ซึ่งใช้การแจกแจงแบบปัวซง (Poisson Distribution) ในการประมาณค่าการกระจายของ M

2.8.3.2 **การคำนวณหาค่าดัชนี** S_1

ดัชนี S_1 มีความคล้ายคลึงกับดัชนี K_2 ซึ่งมีพื้นฐานอยู่บนตัวแปรสุ่ม M แต่ดัชนี S_1 แตกต่างจากดัชนี K_2 ในเรื่องต่อไปนี้ ประการแรก สำหรับดัชนี K_2 ตัวแปร M จะถูกประมาณค่าด้วยการแจกแจงทวินาม ในขณะที่ สำหรับดัชนี S_1 แล้วตัวแปร M จะถูกประมาณค่าด้วยการแจกแจงแบบ Poisson ประการที่สอง พารามิเตอร์ Poisson (Poisson parameter) ของดัชนี S_1 ซึ่งก็คือพารามิเตอร์ μ จะถูกประมาณโดยการใช้โมเดล loglinear ในขณะที่ดัชนี K_2 พารามิเตอร์ binomial ซึ่งก็คือพารามิเตอร์ p จะถูกประมาณค่าโดยการใช้ โมเดลการถดถอย หากค่าประมาณของ μ สำหรับจำนวนผิดของกลุ่ม c' เป็น $\hat{\mu}_c$ ดัชนี S_1 สามารถคำนวณได้จาก

$$S_1 = \sum_{w=m_{c'c}}^{w_s} \frac{e^{-\hat{\mu}_{c'}} \hat{\mu}_{c'}^{w}}{w!} \qquad (10)$$

ดัชนี S_1 ไม่ใช่ความเป็นไปได้แบบ upper tail เนื่องจากการแจกแจงแบบปัวซง ไม่มีขอบเขตบนสำหรับจำนวนของคำตอบผิดที่ตรงกันในขณะที่มีขอบเขตบนสำหรับตัวแปร M ซึ่งคือ $m_{c'c}$ ซึ่งความหมายของสมการ (10) คือ ความเป็นไปได้ที่ผู้สอบซึ่งถูกสงสัยว่าเป็นผู้ลอก ว่าจะมีเวกเตอร์คำตอบผิดที่ตรงกันกับผู้ให้ลอกมากกว่าค่า $m_{c'c}$

หากค่าดัชนี S_1 ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้ว ผู้สอบจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก ค่าดัชนี S_1 ยิ่งน้อย ยิ่งทำให้หลักฐานทางสถิติในการลอกข้อสอบระหว่างผู้ลอก c กับ ผู้ให้ลอก s มีความแข็งแกร่งมากขึ้น

2.8.3.3 **โมเดลของการประมาณค่า** μ

การคำนวณดัชนี s_1 จากสมการ 10 จะต้องทราบค่า w_s, $m_{c'c}$, และ μ โดยค่าของ w_s และ $m_{c'c}$ เราสามารถทราบได้ แต่ μ เราต้องทำการประมาณ ซึ่งค่าเฉลี่ยของ M จะแตกต่างตามระดับความสามารถ ค่าของ M มีค่าน้อยถ้าผู้สอบส่วนใหญ่มีระดับความสามารถสูง เพราะ จำนวนข้อคำถามที่ตอบผิดของผู้ลอกที่ตรงกับของผู้ให้ลอกมีจำนวนน้อย ในทางกลับกันถ้าผู้สอบส่วนใหญ่มีระดับความสามารถต่ำ จำนวนข้อคำถามที่ตอบผิดจะมาก และ จำนวนของข้อคำถามที่ตอบผิดเหมือนกันก็จะมีแนวโน้มที่จะมีจำนวนมาก ข้อมูลเหล่านี้จะถูกนำมาใช้ในการประมาณค่า μ โดยการแบ่งผู้สอบเป็นกลุ่มชั้น ตามจำนวนของคำตอบผิดของพวกเขา

เนื่องจากดัชนี S_1 ใช้การแจกแจงแบบปัวซง ในการประมาณค่าตัวแปร M ทำให้เราจะต้องใช้ โมเดล loglinear (loglinear model) ในการแสดงค่า log ของค่าเฉลี่ยของตัวแปร M (Agresti, 1996: 73) ซึ่งการใช้โมเดลนี้ อนุญาตให้ตัวแปร μ มีความสัมพันธ์แบบไม่ใช่เชิงเส้นตรงกับ ตัว

แปรที่ใช้ในการทำนายซึ่งในที่นี้คือจำนวนของคำตอบผิด การศึกษาของ Hanson (1994) พบว่าโมเดล loglinear ใช้ได้ดีสำหรับตัวแปร M เมื่อตัวแปร M ถูกสันนิษฐานโดยการกระจายแบบ compound binomial

ข้อมูลที่จำเป็นสำหรับการประมาณค่า μ คือ จำนวนคำตอบผิด และ ค่าเฉลี่ยของจำนวนคะแนนผิดที่ตรงกันของแต่ละกลุ่มย่อยที่มีจำนวนข้อผิดเป็น r ให้ μ_r เป็นค่าเฉลี่ย หรือ ค่าที่คาดหวัง (expected value) ของตัวแปร Poisson M_{rj} โมเดล loglinear สามารถเขียนได้ดังนี้

$$\log(\mu_r) = \beta_0 + \beta_1 w_r, \forall r, \tag{11}$$

เมื่อ β_0 คือ จุดตัดที่แสดงถึง logarithm ของค่าเฉลี่ยประชากรของกลุ่มที่มีจำนวนข้อผิดเท่ากัน จำนวน R กลุ่ม และ β_1 คือ พารามิเตอร์ความชัน

การจะทราบค่าดัชนี S_1 จำเป็นต้องทราบ ค่าเฉลี่ยที่เหมาะสม (fitted mean) สำหรับกลุ่มจำนวนข้อผิดที่ผู้ลอกเป็นสมาชิกอยู่ โดยค่าเฉลี่ยที่เหมาะสม (fitted mean) เป็นค่าเฉลี่ยที่จะเกิดขึ้นเมื่อโมเดลมีความสอดคล้องกับข้อมูลเท่านั้น ซึ่งเราสามารถที่จะทราบได้ว่าข้อมูลและโมเดล loglinear ที่ใช้มีความสอดคล้องกันหรือไม่โดยการใช้ค่าสถิติ likelihood-ratio goodness-of-fit (G^2) ตรวจสอบซึ่งจะกล่าวอย่างละเอียดในหัวข้อที่ 2.7.3.4 ซึ่งหากโมเดลและข้อมูลมีความสอดคล้องกันแล้ว ค่าเฉลี่ยที่เหมาะสม (fitted mean) สามารถคำนวณได้จาก

$$\hat{\mu}_{c'} = e^{(\hat{\beta}_0 + \hat{\beta}_1 w_{c'})}$$

2.8.3.4 การตรวจสอบความสอดคล้องโมเดล loglinear

ความสอดคล้องของโมเดล loglinear ในสมการที่ 11 ถูกตรวจจับโดยการใช้สถิติ likelihood-ratio goodness-of-fit, G^2 (Agresti, 1996: 89) สถิติ G^2 สามารถใช้ในการทดสอบสมมติฐานศูนย์ที่ว่าโมเดลมีความสอดคล้องกับข้อมูล ส่วนสมมติฐานทางเลือกเป็นโมเดลไม่สอดคล้องกับข้อมูล โดยให้ $\hat{\mu}_r$ เป็นจำนวนค่าเฉลี่ยที่เหมาะสม (fitted mean) ของคำตอบผิดที่ตรงกันของกลุ่มที่มีจำนวนคำตอบผิด r สถิติ G^2 สามารถคำนวณได้จาก

$$G^2 = 2\sum_{r=1}^{R} \mu_r \log\left(\frac{\mu_r}{\hat{\mu}_r}\right) \tag{12}$$

ถ้าโมเดลสอดคล้องกับข้อมูลอย่างสมบูรณ์แล้ว $\hat{\mu}_r = \mu_r$ ในกรณีนี้ $\log\left(\frac{\mu_r}{\hat{\mu}_r}\right) = 0$ ทำให้ $G^2 = 0$ การกระจายของ G^2 จะถูกประมาณด้วยสถิติ chi-squared ด้วยค่าองศาอิสระ (degrees of freedom) R ลบด้วยจำนวนของพารามิเตอร์ของโมเดล สำหรับโมเดล loglinear ในสมการที่ 11 จำนวนพารามิเตอร์ของโมเดลเป็น 2 ค่า p (p value) ที่ใช้ในการทดสอบสมมติฐานศูนย์เป็นความเป็นไปได้แบบ right-tail ค่า G^2 ที่มาก หรือ ค่า p ที่น้อย ตัวอย่างเช่นน้อยกว่า .01 จะทำให้ความสอดคล้องของโมเดลไม่ดีนัก (Agresti 1996: 89) ถ้าความสอดคล้องของโมเดลกับข้อมูลนั้นไม่ดี เราก็ไม่ควรที่จะใช้สถิติ S_1 ซึ่งสามารถคำนวณได้จากสมการที่ 10 ตรวจจับการลอกข้อสอบ

เนื่องจากจากการศึกษาของ Sotaridona และ Meijer (2003) พบว่า การแจกแจงแบบปัวซงมีความสอดคล้องกับการกระจายของตัวแปร M เป็นอย่างดี ดังนั้นเราสามารถหาค่าเฉลี่ยที่เหมาะสม (fitted mean) สำหรับกลุ่มย่อยที่มีผู้ลอกอยู่แล้วนำไปแทนค่าในสมการ (10) เพื่อหาค่าดัชนี S1 ได้เลยไม่มีความจำเป็นต้องทำการตรวจสอบโมเดลว่ามีความสอดคล้องกับข้อมูลอีก

2.8.4 **ดัชนี** S_2

ดัชนีตรวจจับการลอกข้อสอบที่มีพื้นฐานอยู่บนคำตอบผิดที่เหมือนกัน เช่น ดัชนี K และ K_2 จะไม่นำเอารายละเอียดเพิ่มเติมเกี่ยวกับการลอกซึ่งเป็นคำตอบถูกที่เหมือนกันมาใช้ในการคำนวณหรือวิเคราะห์ ในการไม่นำจำนวนของคำตอบถูกที่เหมือนกันมาวิเคราะห์การลอกข้อสอบนั้น เนื่องจากข้อสันนิษฐานที่ว่า ผู้ลอก c รู้คำตอบของข้อสอบ i อย่างแท้จริง เมื่อใดก็ตามที่ผู้ลอก c และ ผู้ให้ลอก s ตอบข้อสอบ i ได้ถูกต้อง อย่างไรก็ตามข้อสันนิษฐานนี้ไม่ได้เป็นจริงเสมอไป ผู้สอบอาจจะตอบคำถามได้ถูกต้องเพราะการลอก หรือ การเดาก็ได้

ดัชนีK และ K_2ไม่ไวต่อผู้ลอกที่ทำการลอกเฉพาะคำตอบถูกจากผู้ให้ลอก ซึ่งอาจจะเกิดขึ้นได้ในกรณี เช่น ผู้ให้ลอก s และ ผู้ลอก c เป็นเพื่อนกันและ ผู้ให้ลอก s แบ่งคำตอบของเธอหรือเขาให้กับผู้ลอก c ในข้อคำถามที่เธอหรือเขามั่นใจว่าทำถูก อีกตัวอย่างหนึ่ง คือ ในกรณีการสอบที่ผู้ลอก c ได้ทำการติดสินบนผู้ให้ลอก s ให้บอกคำตอบเฉพาะข้อที่คิดว่าทำได้ถูกแก่ผู้ลอก c

ดัชนีตรวจวัดการลอกข้อสอบ S_2 ถูกเสนอเพื่อเอาชนะจุดบกพร่องข้างต้นที่กล่าวมา เนื่องจากดัชนี S_2 จะรวมเอาข้อมูลการลอกข้อสอบทั้ง คำตอบถูกที่ตรงกัน และ คำตอบผิดที่ตรงกัน ไปใช้ในการวิเคราะห์

สำหรับ ดัชนี K และ K_2 นั้นหลักฐานของการลอกคำตอบจะมีค่าเป็น 1 ถ้าผู้ให้ลอก s และผู้ลอก c เลือกตัวเลือกผิดเดียวกัน และเป็น 0 ถ้าทั้งคู่ตอบถูกหรือคำตอบผิดของพวกเขาไม่ตรงกัน แต่สำหรับ ดัชนี S_2 หลักฐานจะมีค่าเป็น 1 ถ้าผู้ให้ลอก s และ ผู้ลอก c เลือกตัวเลือกผิดเดียวกัน

เป็น δ (จะอธิบายต่อไป) ถ้าผู้ให้ลอก s และ ผู้ลอก c ตอบถูก และเป็น 0 ในกรณีอื่น ตัวแปร δ แสดงถึงข้อมูลปริมาณการลอกข้อสอบถูกสำหรับคู่ผู้ลอกและผู้ให้ลอกหนึ่งๆ

ให้ i^* คือข้อสอบที่ผู้ให้ลอก s ตอบถูก และ U_{i^*rj} คือ คำตอบของผู้สอบ rj ที่ตอบข้อสอบ i^* ดังนั้น δ_{i^*rj} เป็นค่าประมาณของข้อมูลการลอกข้อสอบข้อ i^* โดยผู้สอบ rj โดยค่าของ δ_{i^*rj} จะอยู่ในช่วง

$$1 \geq \delta_{i^*rj} \geq 0$$

นั่นคือ $\delta_{i^*rj} = 0$ ถ้า rj รู้คำตอบที่ถูกของข้อสอบข้อ i^* และ $\delta_{i^*rj} = 1$ถ้า rj ไม่รู้คำตอบของข้อสอบข้อ i^* เลยอย่างสิ้นเชิง ปัญหาคือการบ่งบอกปริมาณของความรู้ที่ rj มีในการทำข้อสอบข้อ i^* ซึ่งในการบ่งบอกเราจะเป็นต้องทราบความเป็นไปได้ของผู้สอบ rj ในการตอบข้อสอบข้อ i^* ได้ถูกต้อง ความเป็นไปได้นี้สามารถที่จะประมาณค่าได้จากสัดส่วนของผู้สอบในกลุ่มที่ตอบผิดจำนวน r ที่สามารถตอบข้อสอบข้อ i^* ได้ถูกต้อง จุดด้อยของวิธีการนี้คือการประมาณค่านั้นจะขึ้นอยู่กับประชากรของผู้สอบที่ทำการสอบเป็นอย่างมาก

ให้ P_{i^*rj} คือ ความเป็นไปได้ของผู้สอบ rj ที่จะสามารถตอบข้อสอบข้อ i^* ได้ถูกต้อง และ A_{i^*rj} เป็นตัวแปรบ่งชี้ (indicator variable) ซึ่งจะมีค่าเท่ากับ 1 ถ้า $U_{i^*rj} = u_{i^*s}$ และมีค่าเป็น 0 ในกรณีอื่นๆ

P_{i^*rj} คือความเป็นไปได้ที่เป็นเงื่อนไข (conditional probability) ไม่ใช่ความเป็นไปได้ร่วม (joint probability) ซึ่ง s และ rj จะให้คำตอบเหมือนกันในการทำข้อสอบข้อ i^* กำหนดให้มี U_{i^*s} แล้วความเป็นไปได้คือ

$$P_{i^*rj} = \Pr\left(U_{i^*rj} = u_{i^*s} \middle| U_{i^*s}\right), \tag{13}$$

และการประมาณค่าแบบ maximum likelihood ของ P_{i^*rj} คือ

$$\hat{P}_{i^*rj} = \frac{\sum_{j=1}^{Jr} A_{i^*rj}}{J_r} \tag{14}$$

กำหนดให้มี ค่าประมาณของ P_{i^*rj} สิ่งที่ต้องทำคือการแปลงค่าประมาณนี้ให้เป็น δ_{i^*rj} โดยใช้ฟังก์ชันการแปลงที่เหมาะสมหรือ $f\left(P_{i^*rj}\right)$ ซึ่งจะต้องเป็นฟังก์ชันที่เป็นไปตามเงื่อนไขต่อไปนี้

1. $f\left(P_{i^*rj}\right)$ จะมีค่าเข้าใกล้ 0 หาก P_{i^*rj} มีค่าเข้าใกล้ 1 นั่นคือ หลักฐานทางสถิติของการลอกข้อสอบจะลดลงหาก P_{i^*rj} มีค่าเข้าใกล้ 1
2. $f\left(P_{i^*rj}\right)$ จะมีค่าเข้าใกล้ 1 หาก P_{i^*rj} มีค่าเข้าใกล้ 0 นั่นคือหลักฐานทางสถิติของการลอกข้อสอบจะมีค่าเข้าใกล้ 1 ถ้าผู้ลอกที่สงสัยตอบข้อสอบได้อย่างถูกต้อง แม้ว่าจะมีความเป็นไปได้น้อยมากที่จะตอบข้อสอบข้อนั้นถูก
3. แบบสอบที่มีจำนวนตัวเลือกต่างกันจะมีน้ำหนักฟังก์ชันที่ต่างกัน ให้ f และ f' คือฟังก์ชันสองฟังก์ชันที่มีน้ำหนักต่างกัน และ i^* และ $i^{*'}$ คือข้อสอบที่ถูกเลือกจากแบบสอบสองฉบับ โดยมีจำนวนตัวเลือกเป็น V และ V' โดย $V < V'$ จะทำให้ $f\left(P_{i^*rj}\right) > f'\left(P_{i^*rj}\right)$ เมื่อ $P_{i^*rj} = P_{i^{*'}rj}$

เงื่อนไขข้อ 1 และ 2 นั้นสามารถทำความเข้าใจได้จากการอภิปราย ส่วนเงื่อนไขข้อที่ 3 เกิดขึ้นจากแนวความคิดที่ว่าแบบสอบเลือกตอบที่มีจำนวนตัวเลือกต่างกันควรที่จะมีฟังก์ชันการแปลงค่าที่แตกต่างกันโดยความแตกต่างนี้เกิดจากปัจจัยฟังก์ชันของจำนวนตัวเลือก โดยควรมีการพิจารณาฟังก์ชันที่รวมเอาความเป็นไปได้ของการเดาคำตอบเป็น scaling factor

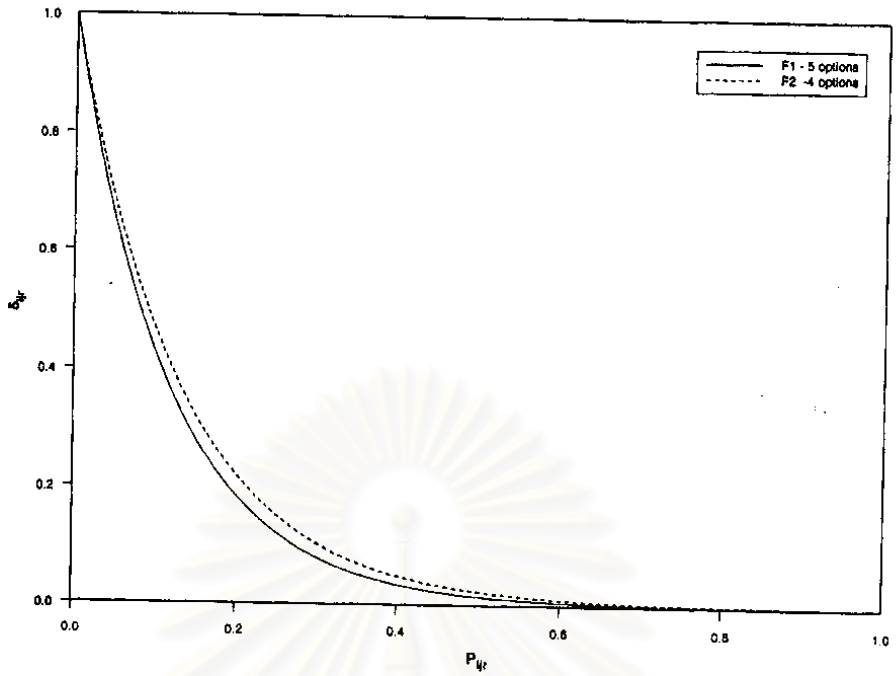
เพื่อความสะดวก กำหนดให้ g เป็นความเป็นไปได้ในการตอบคำถามข้อ i ได้ถูกจากการเดา โดยค่า g ที่ใช้บ่อย คือ 0.20 สำหรับแบบสอบที่มี 5 ตัวเลือก และ ค่า g มีค่าเป็น 0.25 สำหรับแบบสอบที่มี 4 ตัวเลือก ฟังก์ชันที่เหมาะสมและเป็นไปตามเงื่อนไขทั้ง 3 ข้อเขียนได้ดังนี้

$$\delta_{i^*rj} = f\left(P_{i^*rj}\right) = d_1 e^{d_2 P_{i^*rj}} \tag{15}$$

เมื่อ

$$d_2 = -\left(\frac{1+g}{g}\right) \quad \text{และ} \quad d_1 = \left(\frac{1+g}{1-g}\right)^{d_2 P_{i^*rj}}$$

สมการที่ 15 เป็นฟังก์ชันของ P_{i^*rj} ที่ลงลงอย่างคงที่ โดย g คือ ค่าคงที่แบบ scaling (scaling constant) ภาพ 1 แสดงกราฟของสมการที่ 15 ด้วยค่า g เป็น 0.2 (F1- 5 ตัวเลือก) และ g เป็น 0.25 (F2 – 4 ตัวเลือก) จากกราฟจะเห็นว่า ค่าของ δ_{i^*rj} ของกราฟทั้ง F1 และ F2 จะเข้าใกล้ 0 เมื่อ P_{i^*rj} มีค่าเข้าใกล้ 1 และ δ_{i^*rj} มีค่าเข้าใกล้ 1 เมื่อ P_{i^*rj} มีค่าเข้าใกล้ 0 (เงื่อนไขข้อ 1 และ 2) นอกจากนี้ $F1\left(P_{i^*rj}\right) < F2\left(P_{i^*rj}\right)$ สำหรับ $P_{i^*rj} \in (0,1)$ (เงื่อนไขข้อ 3)



ภาพ 1 δ_{i^*jr} ซึ่งเป็นฟังก์ชันของ P_{i^*jr} โดย p_g= 0.25 และ p_g=0.20
ที่มา: Sotaridona และ Meijer (2003)

ให้ M_{rj}^* คือ ผลรวมของจำนวนคำตอบผิดที่ตรงกัน และ คำตอบถูกที่ตรงกันแบบถ่วงน้ำหนัก ระหว่างผู้สอบ rj และ ผู้สอบ s โดยเขียนได้ดังนี้

$$M_{rj}^* = M_{rj} + \sum_{i^*} \delta_{i^*rj} \tag{16}$$

ในสมการที่ 16 การกระจายของค่า M_{rj}^* ของแต่ละข้อคำถามจะเป็น 0 ถ้าคำตอบของ rj ไม่ตรงกับ s และมีค่าเป็น 1 หากคำตอบที่ผิดของผู้สอบ rj ตรงกับคำตอบของผู้สอบ s และมีค่าเป็น δ_{i^*rj} ถ้าคำตอบถูกของผู้สอบ rj ตรงกับคำตอบถูกของผู้สอบ s ค่าของ M_{rj}^* ควรที่จะมีค่ามากถ้าคำตอบผิดของผู้สอบ rj ส่วนใหญ่ตรงกับคำตอบผิดของผู้สอบ s หรือถ้า P_{i^*rj} มีค่าน้อย และคำตอบถูกของผู้สอบ rj ส่วนใหญ่ตรงกับคำตอบถูกของผู้สอบ s ค่าของ M_{rj}^* ที่มีค่ามากเทียบกับจำนวนของข้อสอบแสดงถึงหลักฐานทางสถิติของการลอกข้อสอบที่แข็งแกร่ง

หากว่าไม่มีการตรงกันของคำตอบถูกระหว่างผู้สอบ s และ ผู้สอบ rj พจน์ที่สองของสมการ 16 จะมีค่าเป็นศูนย์ และ $M_{rj}^* = M_{rj}$ ดังนั้น M_{rj} จะกลายเป็นกรณีพิเศษของ M_{rj}^* ในทางตรงกันข้ามหากไม่มีการตรงกันของคำตอบผิด มีแต่การตรงกันของคำตอบถูก ดังนั้น $M_{rj} = 0$ และ

$M_{rj}^{*} = \sum_{i^{*}} \delta_{i^{*}rj}$ โดย $\mathbf{M}_{rj}$ จะมีความไวต่อการลอกข้อสอบผิดเพียงอย่างเดียว แต่ M_{rj}^{*} จะมีความไวต่อทั้งการลองข้อสอบผิดและถูก

ในความเป็นจริงตัวแปรสุ่ม $\mathbf{M}_{rj}^{*}$ เป็นตัวแปรสุ่มซึ่งเป็นค่าจริง และ ไม่ติดลบ เราจะทำให้ตัวแปรสุ่ม $\mathbf{M}_{rj}^{*}$ เป็นจำนวนเต็มโดยการปัดเศษ แม้ว่าการกระทำนี้จะก่อให้เกิดความคลาดเคลื่อน แต่ Leonardo S. Sotaridona และ Rob R. Meijer ผู้คิดค้นและเสนอดัชนีการตรวจวัดการลอกข้อสอบนี้คาดว่าความคลาดเคลื่อนที่เกิดขึ้นนี้จะกระทบต่อประสิทธิภาพของดัชนีเพียงเล็กน้อย โดย $\mathbf{M}_{rj}^{*}$ จะถูกสันนิษฐานว่ามีการแจกแจงแบบปัวซง เช่นเดียวกันกับ $\mathbf{M}_{rj}$ และจะใช้โมเดล loglinear ในการประมาณค่าเฉลี่ย $\hat{\mu}_r$ นอกจากนี้ผลจากการศึกษาของ Sotaridona และ Meijer (2003) พบว่า การแจกแจงแบบปัวซงมีความสอดคล้องกับการกระจายของตัวแปร $\mathbf{M}_{rj}^{*}$ เป็นอย่างดี ดังนั้นหากสามารถคำนวณค่าเฉลี่ยที่เหมาะสม $\hat{\mu}_r$ (fitted mean) ของกลุ่มย่อยที่มีผู้ลอกอยู่ได้ก็จะสามารถคำนวณหาดัชนี S_2 ได้ดังนี้

$$S_2 = \sum_{w=m_{c'c}^{*}}^{I} \frac{e^{-\hat{\mu}_{c'}} \hat{\mu}_{c'}^{w}}{w!} \tag{17}$$

เมื่อ $M_{c'c}^{*}$ เป็นตัวแทนของ $m_{c'c}^{*}$ คือผลรวมของจำนวนคำตอบผิดที่เหมือนกัน และ คำตอบถูกที่เหมือนกันแบบถ่วงน้ำหนักระหว่างผู้สอบ c และผู้สอบ s

หากค่าดัชนี S_2 ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้ว ผู้สอบจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก ค่า S_2 ที่มีค่าน้อยจะแสดงถึงแนวโน้มของการลอกข้อสอบที่มาก

จากรายละเอียดของค่าดัชนี และ ค่าสถิติ ข้างต้น ผู้วิจัยสามารถสรุปรายละเอียดของดัชนีตรวจจับการลอกข้อสอบแต่ละตัวได้ว่า ดัชนี ω เป็นค่าสถิติตรวจจับการลอกข้อสอบที่มีพื้นฐานอยู่บนโมเดลของทฤษฎีการตอบสนองข้อสอบ (item response theory) แบบ nominal response model (NRM) (Bock, 1972) ส่วนค่าดัชนี K_2 เป็นดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบทวินาม (Binomial distribution) ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบปัวซง (Poisson distribution) ส่วนดัชนี S_2 จะเป็นดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดและคำตอบถูกที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบปัวซง (Poisson distribution)

2.9 ทฤษฎีที่เป็นพื้นฐานของดัชนีการตรวจจับการลอกข้อสอบ

ดัชนีตรวจจับการลอกข้อสอบแต่ละตัวนั้นมีพื้นฐานและมีทฤษฎีที่เกี่ยวข้องไม่เหมือนกัน ดังที่กล่าวไว้ในหัวข้อที่แล้ว กล่าวคือ ดัชนี ω เป็นค่าสถิติตรวจจับการลอกข้อสอบที่มีพื้นฐานอยู่บน nominal response model (NRM) ค่าดัชนี K จะถูกประมาณค่าอยู่บนพื้นฐานการแจกแจงทวินาม (binomial distribution) ดัชนี S_1 จะมีการวิเคราะห์ทางคณิตศาสตร์คล้ายคลึงกับ ดัชนี K_2 แต่ดัชนี S_1 จะใช้การกระจายของจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบปัวซง (Poisson distribution) ส่วนดัชนี S_2 จะเป็นดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดและคำตอบถูกที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบปัวซง (Poisson distribution) ในหัวข้อนี้ผู้วิจัยจึงนำเสนอทฤษฎีที่เป็นพื้นฐานหรือทฤษฎีที่เกี่ยวข้องกับดัชนีตรวจจับการลอกข้อสอบดังที่กล่าวข้างต้น โดยจะแบ่งการนำเสนอออกเป็น 5 หัวข้อดังนี้ คือ

2.9.1 Nominal Response Model (NRM)

2.9.2 การแจกแจงทวินาม (Binomial Distribution)

2.9.3 การแจกแจงแบบปัวซง (Poisson Distribution)

2.9.4 Generalized Linear Models (GLM)

2.9.5 Generalized Linear Models for Count Data: Poisson Regression

2.9.1 Nominal Response Model (NRM)

Nominal Response Model (NRM) ถูกคิดค้นโดย บอค (Bock, 1972) โดยโมเดล NRM นี้ใช้สำหรับวิเคราะห์ข้อสอบหรือข้อคำถามที่มีรายการคำตอบไม่เรียงลำดับ เช่น ข้อสอบแบบหลายตัวเลือก (Multiple choices) ข้อคำถามวัดเจตคติ ข้อคำถามสำหรับประเมินบุคลิกลักษณะ เป็นต้น

Nominal Response Model มีจุดมุ่งหมายเพื่อประมาณค่าความสามารถของผู้สอบให้มีความถูกต้องสูงสุดโดยใช้สารสนเทศจากโค้งลักษณะการสอบแต่ละตัวเลือก โดยต่อมาธิสเซน (Thissen) ได้นำแนวคิดนี้ไปพัฒนาต่อโดยกำหนดน้ำหนักสำหรับตัวเลือกแต่ละตัวทำให้การประมาณค่าความสามารถผู้สอบที่ระดับความสามารถต่ำมีความถูกต้องมากขึ้น

Nominal Response Model (NRM) มีลักษณะเป็นโมเดลทั่วไปที่ใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละรายการคำตอบของข้อคำถามโดยตรงแบบขั้นตอนเดียว (Direct IRT Model)

โมเดล

ใน NRM ลักษณะข้อคำถามแต่ละข้อ (i) อธิบายได้ด้วยค่าพารามิเตอร์ความชันของแต่ละรายการคำตอบ (Slope of the trace lines; α_{ix}) และค่าพารามิเตอร์จุดตัดของแต่ละรายการคำตอบ (Intercept parameter, c_{ix}) โอกาสในการตอบรายการคำตอบ x (x=0,1, …, m_i) ตามโมเดล NRM สามารถเขียนเป็นฟังก์ชันได้ดังนี้

$$P_{ix} = \frac{\exp(\alpha_{ix}\theta + c_{ix})}{\sum_{x=0}^{m_i} \exp(\alpha_{ix}\theta + c_{ix})}$$

เมื่อ $P_{ix}(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบมีคุณลักษณะ θ จะเลือกรายการคำตอบที่ x ในเมื่อ

x=0,1, …,m_i

α_{ix} คือ ค่าพารามิเตอร์ความชันแต่ละรายการคำตอบ (slope parameter)

c_{ix} คือ ค่าพารามิเตอร์จุดตัดของแต่ละรายการคำตอบ (intercept parameter)

เพื่อให้สามารถประมาณค่าพารามิเตอร์ในสมการได้ จึงจำเป็นต้องกำหนดเงื่อนไขให้ $\sum \alpha_{ix} = \sum c_{ix} = 0$ และในบางกรณีนั้นจะมีการกำหนดให้ค่าพารามิเตอร์ของรายการคำตอบต่ำสุด $\alpha_{i1} = c_{i1} = 0$ ตามโมเดล NRM จะมีการประมาณค่าพารามิเตอร์ α_{ix} และ c_{ix} สำหรับแต่ละรายการคำตอบ ซึ่งมีทั้งหมดจำนวน m_i+1 รายการคำตอบในแต่ละข้อ

ธิสเซน (Thissen, 1993) ได้แสดงตัวอย่างข้อคำถาม ซึ่งวัดความเสี่ยงต่อการเป็นโรงเบื่ออาหาร และ ทำการประมาณค่าพารามิเตอร์ของแต่ละรายการคำตอบโดยใช้โมเดล NRM

ตัวอย่างข้อคำถาม : ท่านชอบบรรยากาศการรับประทานอาหารลักษณะใด

รายการคำตอบ : (1) ที่บ้านตามลำพัง

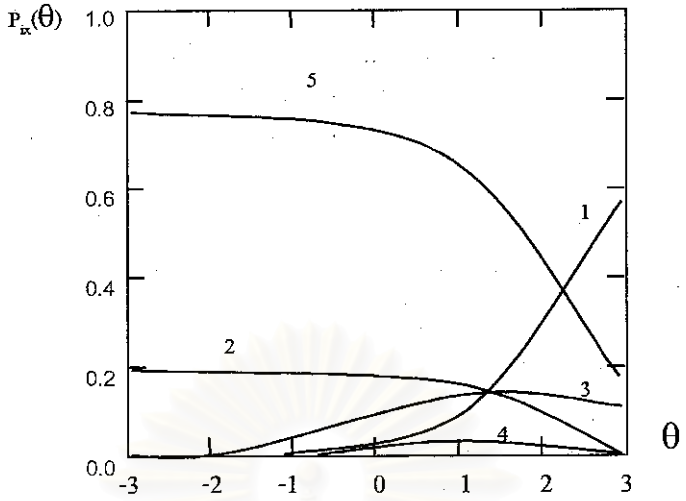
(2) ที่บ้านกับครอบครัว

(3) ที่ภัตตาคาร

(4) ที่บ้านเพื่อน

(5) แบบใดก็ได้

ธิสเซน ได้ประมาณค่าพารามิเตอร์โดยใช้โมเดล NRM พบว่าข้อคำถามส่วนใหญ่มีรายการคำตอบแบบเรียงลำดับ แต่บางข้อไม่มีการเรียงลำดับที่ชัดเจน สำหรับข้อคำถามตัวอย่างมีค่าความชันของแต่ละรายการคำตอบดังนี้ α_1= -0.39, α_2= -0.39, α_3= 0.24, α_4= -0.39 และ α_5= 0.93 ส่วนค่าจุดตัดของแต่ละรายการคำตอบมีค่าดังนี้ c_1= 2.02, c_2= -1.49, c_3= -0.26, c_4= 0.57 และ c_5= -0.83 ซึ่งมีได้รายการคำตอบดังแสดงในภาพ 2



ภาพ 2 โค้งรายการคำตอบสำหรับข้อคำถามตัวอย่าง โดยใช้ Nominal Response Model
ที่มา: ศิริชัย กาญจนวาสี (2545)

จากรูปจะเห็นได้ว่าผู้ที่มีความเสี่ยงต่ำในการเป็นโรงเบื่ออาหารสูง (θ ต่ำ) มักเลือกตัวเลือกหรือรายการคำตอบ (5) คือแบบใดก็ได้ ในขณะที่ผู้ที่มีความเสี่ยงสูงในการเป็น โรคเบื่ออาหารสูง (θ สูง) มักเลือกตัวเลือกหรือรายการคำตอบ (1) คือ ที่บ้านตามลำพัง

2.9.2 **การแจกแจงทวินาม** (Binomial Distribution)

การแจกแจงทวินามเป็นการแจกแจงของตัวแปรสุ่มแบบไม่ต่อเนื่องที่เกิดขึ้นในการทดลองที่ทำซ้ำๆ กันโดยผลการทดลองแต่ละครั้งมีได้เพียง 2 อย่าง คือ ความสำเร็จ (success) และ ความไม่สำเร็จ (failure) เช่น โยนเหรียญ 1 อัน 10 ครั้ง จะสังเกตว่าในการโยนแต่ละครั้งจะให้ผลที่เป็นอิสระต่อกัน นั่นคือ การเกิดผลในการโยนครั้งที่ 1 จะไม่ส่งผลหรือไม่มีผลต่อการเกิดผลในการโยนครั้งที่ 2 สมมติให้เหตุการณ์ที่สนใจจะศึกษาคือ การเกิดหัว ดังนั้นการได้หัวจากการโยนแต่ละครั้งจะเป็นความสำเร็จ ทำให้การเกิดก้อยแต่ละครั้งเป็นความไม่สำเร็จ เป็นที่ทราบกันดีว่าโอกาสหรือความน่าจะเป็นในการเกิดหัวจากการโยนเหรียญแต่ละครั้ง จะเท่ากับ 0.5 ซึ่งในที่นี้จะเรียกว่า p และความน่าจะเป็นในการเกิดก้อยจะเท่ากับ 0.5 จะเรียกว่า q

ลักษณะทั่วไปของการทดลองทวินาม มีดังนี้

1. เป็นการทดลองที่กระทำซ้ำๆ กัน n ครั้ง
2. ในการทดลองแต่ละครั้ง ผลที่เกิดมีเพียง 2 อย่าง คือ ความสำเร็จและความไม่สำเร็จ
3. ความน่าจะเป็นของความสำเร็จที่เกิดจากการทดลองแต่ละครั้ง มีค่าคงที่เท่ากับ p และ ความน่าจะเป็นของความไม่สำเร็จที่เกิดจากการทดลองแต่ละครั้งมีค่าเท่ากับ q โดยที่ q=1-p

4. การทดลองแต่ละครั้งเป็นอิสระต่อกัน

นิยาม การทดลองที่ให้ผลลัพธ์ปรากฏได้เพียง 2 อย่าง คือ ความสำเร็จ (success) กับความไม่สำเร็จ (failure) โดยที่ความน่าจะเป็นของการเกิดความสำเร็จจากการทดลองแต่ละครั้งเป็น p และความน่าจะเป็นของการเกิดความไม่สำเร็จเป็น 1-p ถ้ากำหนดให้ X เป็นตัวแปรสุ่ม X จะมีค่าได้เพียง 2 ค่าเท่านั้น คือ x=0,1 โดยที่

x=0 แทนเหตุการณ์ของการเกิดความไม่สำเร็จ (failure)

x=1 แทนเหตุการณ์ของการเกิดความสำเร็จ (success)

การแจกแจงความน่าจะเป็นของ X จะเขียนได้ดังนี้

$$Pr(X=x) = {}^{n}C_{x}p^{x}q^{n-x} \qquad ; x=0,1,2,\ldots,n$$

หรือ

$$P(X=x)=\binom{n}{x}p^{x}(1-p)^{n-x},\ x=0,\ 1,\ 2,...,n$$

เมื่อ n= จำนวนครั้งของการทดลอง

X= จำนวนครั้งของการเกิดความสำเร็จ

p= ความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้ง

q= ความน่าจะเป็นของความไม่สำเร็จในการทดลองแต่ละครั้ง = 1-p

ค่าเฉลี่ยเลขคณิต (μ) ละความแปรปรวน (σ^2)ของการแจกแจงทวินาม คือ

$$\mu = E(X) = np$$
$$\sigma^2 = Var(X) = npq = np(1-p)$$

2.9.3 **การแจกแจงแบบปัวซง** (Poisson Distribution)

การแจกแจงแบบปัวซง (Poisson Distribution) ซึ่งเป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มไม่ต่อเนื่องประเภทหนึ่งซึ่งเป็นการแจกแจงความน่าจะเป็นของค่าจำนวนนับโดยค่าจำนวนนับนั้นเป็นค่าที่เกิดขึ้นในช่วงเวลาใดเวลาหนึ่ง หรือเกิดขึ้นในพื้นที่ใดพื้นที่หนึ่งต่อเนื่องกัน โดยที่หน่วยของเวลาอาจจะเป็นวินาที นาที ชั่วโมง วัน สัปดาห์ เดือน หรือ ปี และ หน่วยพื้นที่อาจจะเป็นตารางมิลลิเมตร ตารางฟุต หรืออาจจะเป็นปริมาตรก็ได้ ตัวอย่างของการทดลองที่เป็นการทดลองปัวซง เช่น จำนวนหนูในนาข้าวต่อหนึ่งหน่วยพื้นที่ จำนวนผู้เข้ามาใช้บริการตู้ ATM ในช่วงเวลา 16.00-17.00 น. จำนวนคำที่พิมพ์ผิดต่อหนึ่งหน้ากระดาษ จำนวนอุบัติเหตุที่เกิดขึ้นต่อปีในเขตกรุงเทพฯ จำนวนครั้งที่โทรศัพท์เข้ามาในทุก 5 นาที ฯลฯ

อาจจะกล่าวได้ว่า คุณสมบัติของการแจกแจงแบบปัวซง คือ เป็นปรากฏการณ์หรือการทดลองที่เกิดขึ้นในช่วงเวลาหนึ่งหรือขอบเขตหนึ่ง

นิยาม ถ้า X เป็นจำนวนครั้งของความสำเร็จที่เกิดขึ้นในช่วงเวลาใดช่วงเวลาหนึ่ง หรือในพื้นที่ใดพื้นที่หนึ่ง จำนวนครั้งของความสำเร็จจะมีการแจกแจงปัวซง ที่มีพารามิเตอร์ λ โดยที่ λ คือค่าเฉลี่ยของการเกิดเหตุการณ์ของ X ต่อหน่วยเวลาหรือหน่วยพื้นที่

สมบัติของการแจกแจงปัวซง

1. ค่าเฉลี่ยของจำนวนครั้งของความสำเร็จ ในช่วงเวลาที่กำหนด หรือ ในพื้นที่ที่กำหนดจะต้องเป็นค่าที่ทราบแน่นอน
2. ความน่าจะเป็นของการเกิดความสำเร็จ 1 ครั้งในช่วงเวลาสั้นๆ หรือในพื้นที่เล็กๆ แปรตามขนาดของช่วงเวลาหรือขนาดของพื้นที่ และจะต้องไม่ขึ้นอยู่กับความสำเร็จที่เกิดภายนอกช่วงเวลาที่กำหนด หรือพื้นที่ที่กำหนด
3. ความน่าจะเป็นที่จะเกิดความสำเร็จมากกว่า 1 ครั้ง ในช่วงเวลาสั้นๆ หรือในพื้นที่เล็กๆ ถือว่ามีค่าน้อยมาก

ถ้ากำหนดให้ X เป็นจำนวนครั้งของความสำเร็จในช่วงเวลาที่กำหนด หรือในสิ่งที่ที่กำหนด ความน่าจะเป็นที่จะเกิดความสำเร็จ X ครั้งเท่ากับ

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!},\ x = 0,1,2,...$$

เมื่อ λ = ค่าเฉลี่ยของความสำเร็จในช่วงที่กำหนด

e =ค่าคงที่ = 2.71828

ค่าเฉลี่ยและความแปรปรวนของการแจกแจงปัวซง คือ

$$E(x) = \lambda$$
$$Var(x) = \lambda$$

2.9.4 Generalized Linear Models (GLM)

เราสามารถเขียนย่อ Generalized Linear Models สามารถเขียนย่อได้ว่า GLM โดย GLM เป็นวิธีการที่จะช่วยในการศึกษาอิทธิพล (effect) ของตัวแปรต้น (explanatory variables) ที่มีต่อตัวแปรตาม (response variables) ซึ่ง การวิเคราะห์ถดถอย (regression) และ การวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) เป็นวิธีการวิเคราะห์ส่วนหนึ่งของการวิเคราะห์แบบ GLM

ส่วนประกอบของ Generalized Linear Models (GLM)

ทุกๆ โมเดล GLM จะมีส่วนประกอบ 3 อย่าง คือ *random component* ซึ่งเป็นส่วนที่ระบุตัวแปรตาม และ สันนิษฐานรูปแบบการแจกแจงความน่าจะเป็นของมัน *systematic component* เป็นส่วนที่ระบุตัวแปรต้นที่จะใช้เป็นตัวทำนาย (predictors) ในโมเดล และ *link* เป็นส่วนที่จะอธิบายความสัมพันธ์ทางฟังก์ชันระหว่าง systematic component และ ค่าคาดหวังหรือค่าเฉลี่ยของ random component

1. Random Component เป็นส่วนที่ระบุตัวแปรตาม Y และ เลือกรูปแบบการแจกแจงความน่าจะเป็นสำหรับตัวแปรตามนั้น ตัวอย่างเช่น เราอาจจะกำหนดให้ตัวแปร Y เป็นผลลัพธ์ที่เป็นไปได้ของการทดลองหนึ่งซึ่งอาจจะเป็นจำนวนครั้งของความสำเร็จของการทดลอง ซึ่ง Y ไม่มีทางที่จะมีค่าติดลบได้ ดังนั้นเราจึงเลือก Poisson Distribution สำหรับตัวแปรตามนั้น

   โดยเราสามารถเขียนให้ค่าคาดหวังของ Y หรือ ค่าเฉลี่ยของการแจกแจงความน่าจะเป็น เป็น $\mu = E(Y)$ ซึ่งในโมเดล GLM ค่าของ μ จะแปรเปลี่ยนไปตามค่าของตัวแปรต้น
2. Systematic Component เป็นส่วนที่ระบุตัวแปรต้น ซึ่งจะถูกใส่เข้าไปเป็นตัวทำนายแบบเชิงเส้นทางด้านขวาของสมการทำนายหรืออาจกล่าวได้ว่า systematic component จะเป็นส่วนระบุตัวแปร x ในความสัมพันธ์

$$\alpha + \beta_1 x_1 + ... + \beta_k x_k$$

   ซึ่งผลรวมเชิงเส้นของตัวแปรต้นนี้ ถูกเรียกว่า ตัวทำนายเชิงเส้น (linear predictor)
3. Link เป็นส่วนที่ระบุว่า $\mu = E(Y)$ มีความสัมพันธ์กับตัวแปรต้นอย่างไร โดยเราจะให้ฟังก์ชัน $g(.)$ คือ link function

   โดย $g(\mu) = \mu$ เป็น link function ที่ธรรมดาที่สุด เราเรียก link function นี้ว่า identity link ซึ่งมันระบุถึงโมเดลเชิงเส้นของค่าเฉลี่ยของตัวแปรตาม เขียนได้ว่า

$$\mu = \alpha + \beta_1 x_{1.} + ... \beta_k x_k$$

   ซึ่งรูปแบบนี้เป็นรูปแบบทั่วไปของโมเดลวิเคราะห์ถดถอยสำหรับตัวแปรตามแบบต่อเนื่อง แต่ในบางครั้ง link function อาจจะไม่ได้ระบุความสัมพันธ์แบบเชิงเส้น แต่จะระบุความสัมพันธ์ระหว่าง ค่าเฉลี่ยของตัวแปรตามกับตัวทำนายไม่ใช่แบบเชิงเส้น (nonlinear) เช่น $g(\mu) = \log(\mu)$ ซึ่งฟังก์ชัน log จะใช้กับจำนวนบวก ดังนั้น

"log link" เหมาะสมในกรณีที่ μ ไม่สามารถมีค่าติดลบได้ โมเดล GLM ที่ใช้ log link ถูกเรียกว่า loglinear model ซึ่งสามารถเขียนได้ดังนี้

$$\log(\mu) = \alpha + \beta_1 x_1 + ... + \beta_k x_k$$

และในบางครั้งเราอาจจะพบ link function แบบ $g(\mu) = \log\left[\frac{\mu}{(1-\mu)}\right]$ ซึ่งถูกเรียกว่า logit link ซึ่งเป็น link function ที่เหมาะสมที่จะใช้เมื่อ μ มีค่าอยู่ระหว่าง 0 ถึง 1 เช่น ค่าความน่าจะเป็น เป็นต้น โมเดล GLM ที่ใช้ logit link จะถูกเรียกว่า logit model

แต่ละการแจกแจงที่เป็นไปได้ของตัวแปรตามจะมีฟังก์ชันพิเศษของค่าเฉลี่ยของตัวแปรตามนั้น เรียกว่า *natural parameter* โดย

สำหรับ การแจกแจงแบบปกติ (Normal Distribution) ฟังก์ชัน natural parameter คือ ค่า mean

สำหรับ การแจกแจงแบบปัวซง (Poisson Distribution) ฟังก์ชัน natural parameter คือ ค่า log ของ mean

สำหรับ การแจกแจงแบบทวินาม (Binomial Distribution) ฟังก์ชัน natural parameter คือ ค่า logit ของความเป็นไปได้ที่จะสำเร็จ

2.9.5 Generalized Linear Models for Count Data: Poisson Regression

ตัวแปรตามที่ไม่ต่อเนื่องมากมายที่เป็นจำนวนผลลัพธ์ที่เป็นไปได้ เช่น จำนวนเมืองต่างๆ ในโลก จำนวนรถที่ถูกขโมยในปี 1995 เป็นต้น โมเดล GLM ที่จะใช้กับข้อมูลประเภทนี้ หรือ ข้อมูลแบบ count data จะใช้การแจกแจงแบบปัวซง (Poisson Distribution) สำหรับตัวแปรตาม

การกระจายแบบปัวซง มีค่าเฉลี่ยเป็นบวกโดย log ของ mean จะเป็น natural parameter ของการกระจายแบบ Poisson

Poisson Loglinear Model

คือ โมเดล GLM ที่ใช้การแจกแจงแบบปัวซงสำหรับตัวแปรตาม Y และใช้ log link

ถ้าให้ μ คือ ค่าคาดหวังของตัวแปร Y และ x เป็นตัวแปรต้น แล้ว Poisson Loglinear Model สามารถเขียนได้ดังนี้

$$\log \mu = \alpha + \beta x$$

ดังนั้น ค่าเฉลี่ยของตัวแปร Y สามารถเขียนให้อยู่ในรูปความสัมพันธ์แบบ exponential ได้ดังนี้

$$\mu = \exp(\alpha + \beta x) = e^{\alpha}(e^{\beta})^{x}$$

ถ้า β >0 จะทำให้ e^{β} >1 แล้วค่าเฉลี่ยของ Y จะมีค่าเพิ่มขึ้นถ้าค่าของ x เพิ่ม และ ถ้า β <0 ค่าเฉลี่ยของ Y จะลดลงถ้าค่าของ x ลดลง

หากโมเดล Poisson Loglinear Model สอดคล้อง (fit) กับข้อมูลแล้ว ค่าเฉลี่ยของ Y หรือ μ จะถูกเรียกว่า fitted value

Poisson Model Checking

โมเดล GLM จะสามารถอธิบายหรือใช้งานได้อย่างถูกต้องถ้าโมเดลสอดคล้อง (fit) เป็นอย่างดีกับข้อมูล โดยค่าสถิติ likelihood-ratio goodness-of-fit (G^2) เป็นค่าที่สามารถบ่งบอกความเพียงพอของความสอดคล้องของโมเดล GLM ได้

สถิติ G^2 ใช้ทดสอบสมมติฐานศูนย์ที่ว่า โมเดลมีความสอดคล้องกับข้อมูล ส่วนสมมติฐานทางเลือก คือ โมเดลไม่สอดคล้องกับข้อมูล ในการทดลองที่ประกอบด้วยตัวแปรต้นครั้งที่ i จากทั้งหมด N ครั้ง จะได้ผลลัพธ์หรือตัวแปรตาม y_i และมีค่าเฉลี่ยที่เหมาะสมเป็น $\hat{\mu}_i$ สถิติ G^2 สามารถคำนวณจาก

$$G^2 = 2\sum y_i \log\left(\frac{y_i}{\hat{\mu}_i}\right)$$

เมื่อค่า fitted value ($\hat{\mu}_i$) ค่อนข้างสูง (เข้าใกล้ 5) และเมื่อจำนวนครั้งของการทดลองคงที่ ค่าสถิติ G^2 จะมีการแจกแจงแบบ chi-squared โดย df คือจำนวนของตัวแปรตามลบด้วยจำนวนของพารามิเตอร์ของโมเดล ค่า p (p-value) คือ ค่าความเป็นไปได้แบบ right-tail

ค่าสถิติ G^2 ที่มาก และ ค่า p ที่น้อย แสดงถึงว่าโมเดลไม่สอดคล้องกับข้อมูล

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 2 งานวิจัยที่เกี่ยวข้องกับดัชนีตรวจจับการลอกข้อสอบ

งานวิจัยต่างประเทศ

Francis S. Belleza และ Suzanne F. Belleza (1989) ทำวิจัยเรื่อง การตรวจจับการโกงข้อสอบหลายตัวเลือกโดยการวิเคราะห์ความผิดพลาดที่เหมือนกัน (Detection of Cheating on Multiple-Choice Tests by Using Error-Similarity Analysis) โดยมีรายละเอียดดังนี้ การโกงข้อสอบหลายตัวเลือกเป็นปัญหาสำคัญที่ไม่สามารถแก้ได้โดยการใช้เพียงแบบสอบที่หลากหลายแบบ จำนวนผู้คุมสอบที่มากขึ้น หรือ ใช้ห้องสอบขนาดใหญ่ขึ้น แต่เราสามารถใช้กระบวนการทางสถิติ (Statistical procedure) ซึ่งจะเปรียบเทียบคำตอบของคู่นักเรียน โดยใช้ข้อมูลจากข้อคำถามที่นักเรียนทั้งคู่ตอบผิด ถ้าจำนวนของคำตอบผิดที่ตรงกัน (identical wrong answer) มีค่ามากพอและมากกว่าจำนวนคำตอบผิดที่คาดไว้ และ นักเรียนทั้งสองนั่งสอบใกล้กันแล้ว มีแนวโน้มที่จะเกิดการโกงขึ้น (cheating) โดยผลการวิจัยของพวกเขาซึ่งใช้การวิเคราะห์ข้อมูลทางสถิติ จากข้อมูลผู้สอบ 90 คนทำให้ได้ข้อสรุปเกี่ยวกับการโกง และ ข้อจำกัดของขบวนการทางสถิติ

James A. Wollack (1997) ทำการวิจัย เรื่อง Nominal Response Model สำหรับการตรวจจับการลอกข้อสอบ (A Nominal Response Model Approach for Detecting Answer Copying) โดยมีเนื้อหาโดยสรุป ดังนี้ เมื่อผู้สอบลอกคำตอบของข้อคำถามจากผู้สอบคนอื่น จะส่งผลให้ความตรงของการทดสอบลดลง ขบวนการทางสถิติที่ใช้ในการตรวจจับการลอกข้อสอบส่วนใหญ่จะถูกพัฒนาขึ้นโดยมีพื้นฐานจาก ทฤษฎีการทดสอบแบบดั่งเดิม (Classical test theory) ดังนั้นจะทำให้ค่าสถิติเหล่านั้นมีจุดด้อยในด้านคะแนนซึ่งไม่เป็นอิสระจากกลุ่มตัวอย่าง ค่าสถิติของข้อสอบ และ การประมาณค่าจำนวนของคำตอบที่ตรงกันระหว่างคู่ผู้สอบที่ลำเอียง ขบวนการทางสถิติที่มีพื้นฐานจากทฤษฎีการตอบสนองข้อสอบ (Item response theory) จะช่วยลดปัญหาดังกล่าว โดยงานวิจัยชิ้นนี้จะทำการเปรียบเทียบค่าสถิติที่มีพื้นฐานจากทฤษฎีการตอบสนองข้อสอบตัวใหม่ คือ ดัชนี ω กับค่าดัชนีที่มีพื้นฐานจากทฤษฎีการทดสอบแบบดั่งเดิม คือ g_2 ภายใต้สถานการณ์การลอกที่แตกต่างกันตามตัวแปรต้นที่ศึกษา คือ จำนวนข้อสอบที่ถูกลอกซึ่งแตกต่าง ความยาวของแบบสอบที่แตกต่าง รูปแบบการลอกข้อสอบที่แตกต่าง และ ขนาดกลุ่มตัวอย่างที่แตกต่าง ผลการวิจัย พบว่าค่าสถิติ ω สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 (Type I error) ได้อย่างดี อำนาจการตรวจจับการลอกข้อสอบของดัชนี ω จะแตกต่างกันไปตามความยาวของแบบสอบและร้อยละของจำนวนข้อสอบที่ถูกลอกโดยอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω จะยิ่งมากขึ้นเมื่อมีจำนวนของข้อคำถามที่ถูกลอกมากขึ้น ซึ่งอาจ

กล่าวได้ว่าผู้สอบที่ลอกข้อสอบจำนวนมากมีโอกาสที่จะถูกตรวจจับว่าลอกข้อสอบมากกว่าผู้สอบที่ลอกข้อสอบจำนวนน้อย ดัชนี ω มีอำนาจการตรวจจับผู้ลอกข้อสอบดีสำหรับการลอกข้อสอบร้อยละ 20 ของแบบสอบจำนวน 80 ข้อ และ การลอกข้อสอบร้อยละ 30 สำหรับแบบสอบจำนวน 40 ข้อ นอกจากนี้ รูปแบบการลอกข้อสอบมีผลเพียงเล็กน้อยต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ข้อจำกัดของดัชนี ω คือ ดัชนี ω คำนวณอยู่บนพื้นฐานของ nominal response IRT model ซึ่งการที่จะทำให้ค่าประมาณพารามิเตอร์มีความถูกต้องและเสถียรจำเป็นที่จะต้องใช้จำนวนผู้สอบที่มากหรือกลุ่มตัวอย่างที่มีขนาดใหญ่ แต่อย่างไรก็ตามสามารถกล่าวได้ว่าดัชนี ω เป็นดัชนีตรวจจับการลอกข้อสอบที่ดีกว่าดัชนี g_2

James A. Wollack และ Allan S. Cohen (1998) ได้ทำวิจัยเรื่อง การตรวจจับการลอกข้อสอบเมื่อไม่ทราบค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ของผู้สอบ (Detection of Answer Copying With Unknown Item and Trait Parameter) โดยมีเนื้อหาโดยสรุป ดังนี้ จากงานวิจัยเรื่อง Nominal Response Model สำหรับการตรวจจับการลอกข้อสอบ (A Nominal Response Model Approach for Detecting Answer Copying) ของ James A. Wollack (1997) ได้ศึกษาการใช้ดัชนี ω ตรวจสอบการลอกข้อสอบโดยมีพื้นฐานอยู่บนข้อตกลงเบื้องต้นว่า ทราบค่าพารามิเตอร์ของข้อสอบซึ่งได้วิเคราะห์จาก nominal response model อยู่ก่อนแต่ข้อตกลงเบื้องต้นนี้ไปจำกัดความเป็นประโยชน์ของดัชนี ω โดยเฉพาะเมื่อนำเอาดัชนี ω ไปใช้ในสถานการณ์ห้องเรียนจริง เนื่องจากครูผู้สอนส่วนใหญ่ไม่มีชุดข้อสอบที่ได้ทำการวัดหาค่าพารามิเตอร์ของข้อสอบไว้ก่อน งานวิจัยชิ้นนี้จึงทำการศึกษาเกี่ยวกับ อัตราความคลาดเคลื่อนประเภทที่ 1 (Type I error) เชิงประจักษ์ และ อำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ซึ่งถูกคำนวณโดยไม่ทราบค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ของผู้สอบอยู่ก่อน แต่จะทำการประมาณค่าพารามิเตอร์จากชุดข้อมูลที่มีผู้สอบจำนวน 100 คน และ 500 คน ผลการวิจัย พบว่าการประมาณค่าพารามิเตอร์จากชุดข้อมูลดังกล่าวไม่กระทบต่อค่าอัตราความคลาดเคลื่อนประเภทที่ 1 (Type I error) แต่อำนาจการตรวจจับการลอกข้อสอบของดัชนี ω จะลดลงเล็กน้อยสำหรับชุดข้อมูลที่มีผู้สอบจำนวน 100 คน แต่จะมีค่าเท่าเดิมสำหรับชุดข้อมูลที่มีผู้สอบจำนวน 500 คน

Leonardo S. Sotaridona และ Rob R. Meijer (2002) ทำการวิจัยเรื่อง คุณสมบัติทางสถิติของดัชนี K สำหรับการตรวจจับการลอกข้อสอบ (Statistical Properties of the K-Index for detecting answer copying) โดยตรวจสอบคุณสมบัติทางสถิติของดัชนี K (Holland, 1996) ซึ่งสามารถใช้ในการตรวจจับพฤติกรรมการลอกข้อสอบ การศึกษาเป็นแบบการศึกษาจำลอง

สถานการณ์ (simulation study) ซึ่งศึกษาการนำดัชนี K ไปใช้กับชุดข้อมูลที่มีขนาดเล็ก กลาง และ ใหญ่ นอกจากนั้น อัตราความคลาดเคลื่อนประเภทที่ 1 (Type I error) และ อำนาจการตรวจจับการลอกข้อสอบ (detection rates) ของดัชนีนี้จะถูกเปรียบเทียบกับดัชนี ω (Wollack, 1997) โดยการประมาณค่าหลายแบบถูกใช้ในการคำนวณหาดัชนี K ผลการวิจัยพบว่า ทุกการประมาณค่าจะสามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ภายใต้ระดับ nominal (nominal level) ผลการวิจัยยังพบว่าการใช้ดัชนี ω จะให้ค่าอำนาจการตรวจจับการลอกข้อสอบที่สูงกว่าดัชนี K ทุกแบบ สำหรับสถานการณ์ที่กลุ่มตัวอย่างเป็นขนาดเล็กและปานกลาง (100 และ 500 คน) อำนาจการตรวจจับการลอกข้อสอบของทั้งดัชนี K และ ดัชนี ω จะมีค่าเพิ่มขึ้นเมื่อร้อยละของจำนวนข้อสอบที่ถูกลอกเพิ่มขึ้นหรืออาจกล่าวได้ว่า ผู้สอบที่ลอกข้อสอบจำนวนมากมีแนวโน้มที่จะถูกตรวจจับว่าเป็นผู้ลอกมากกว่าผู้สอบที่ลอกข้อสอบจำนวนน้อย การเพิ่มจำนวนของผู้สอบไม่ได้มีผลมากต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ซึ่งอาจจะเป็นเพราะ การคำนวณดัชนี ω จะขึ้นอยู่กับรูปแบบการตอบข้อคำถามของผู้ลอกและผู้ให้ลอกเท่านั้น ไม่ได้สนใจรูปแบบการตอบข้อคำถามของผู้สอบคนอื่นๆ ในทางกลับกันอำนาจการตรวจจับการลอกข้อสอบของดัชนี K จะมีค่าเพิ่มขึ้นเมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อคำถามมีค่าเพิ่มขึ้น นอกจากนี้ยังพบว่าระดับความสามารถของผู้ให้ลอกจะส่งผลให้อำนาจการตรวจจับการลอกข้อสอบของดัชนี K และ ดัชนี ω มีค่าเพิ่มขึ้นเล็กน้อยสำหรับสถานการณ์ที่มีร้อยละของจำนวนข้อสอบที่ถูกลอกเป็น ร้อยละ 40 และ ร้อยละ 30 แต่ระดับความสามารถของผู้ให้ลอกจะไม่ส่งผลทำให้อำนาจการตรวจจับการลอกข้อสอบของดัชนี K และ ดัชนี ω เปลี่ยนแปลงสำหรับสถานการณ์ที่มีร้อยละของจำนวนข้อสอบที่ถูกลอกเป็น ร้อยละ 20 และ ร้อยละ 10

Leonardo S. Sotaridona และ Rob R. Meijer (2003) ทำการวิจัยเรื่อง สถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Two new statistics to detect answer copying) ดัชนีตรวจจับการลอกข้อสอบใหม่ 2 ดัชนีที่ถูกนำเสนอ ได้แก่ ดัชนี S_1 และ S_2 ดัชนี S_1 จะคล้ายคลึงกับดัชนี K (Holland, 1996) และดัชนี K_2 (Sotaridona & Meijer, 2002) แต่การแจกแจงของจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกนั้นใช้การแจกแจงแบบปัวซง แทนที่จะใช้การแจกแจงทวินาม เพื่อพัฒนาอำนาจการตรวจจับการลอกข้อสอบของดัชนี K และ K_2 ดัชนี S_2 ถูกเสนอเพื่อกำจัดข้อจำกัดของดัชนี K และ K_2 นั่นคือความไม่ไวต่อการลอกข้อสอบถูกของดัชนี K และ K_2 ดัชนี S_2 จะรวมเอาคำตอบถูกที่ตรงกันเข้ากับคำตอบผิดที่ตรงกัน การศึกษาเป็นแบบการศึกษาเชิงจำลองสถานการณ์ (simulation study) จะทำการตรวจสอบความเป็นประโยชน์ของดัชนี S_1 และ S_2 สำหรับชุดแบบสอบที่มีจำนวนข้อสอบ 40 และ 80 ข้อ ขนาดกลุ่มตัวอย่างจำนวน 100 และ 500 คน และ ร้อยละการลอกข้อสอบจำนวน 10% 20% 30% และ 40% ความคลาดเคลื่อน

ประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 และ S_2 ถูกเปรียบเทียบกับความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 และ ดัชนี ω ผลการวิจัยด้านค่าความคลาดเคลื่อนประเภทที่ 1 ชี้ให้เห็นว่า ค่าดัชนีทั้ง 4 ดัชนีสามารถที่จะควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ได้โดยของดัชนี S_1 และ K_2 จะค่อนข้าง conservative หากเปรียบเทียบกับดัชนี S_2 และ ω กล่าวคือค่าดัชนี S_2 สามารถรักษาระดับค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error) ให้อยู่ต่ำกว่าเส้นกราฟขอบเขต (Boundary line) สำหรับกลุ่มผู้สอบจำนวน 100 คน และจะมีแนวโน้มที่จะ liberal สำหรับกลุ่มผู้สอบจำนวน 500 คน และ ค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error) ของดัชนี ω จะมีแนวโน้มที่ liberal สำหรับกลุ่มผู้สอบจำนวน 100 คน และ มีแนวโน้มที่จะ conservative สำหรับกลุ่มผู้สอบจำนวน 500 คน โดยดัชนีตรวจจับการลอกข้อสอบที่ค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error) มีแนวโน้มจะ conservative ที่สุดสำหรับกลุ่มผู้สอบจำนวน 100 คน คือ ค่าดัชนี S_1 และ สำหรับกลุ่มผู้สอบจำนวน 500 คน คือ ค่าดัชนี K_2 ส่วนข้อค้นพบเกี่ยวกับอำนาจการตรวจจับการลอกข้อสอบ มีดังนี้ อำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของดัชนีตรวจจับการลอกข้อสอบทุกค่าจะเพิ่มขึ้นเมื่อร้อยละของจำนวนข้อสอบที่ถูกลอกที่เพิ่มขึ้น โดยอำนาจการตรวจจับการลอกข้อสอบ (Detection rate) ของดัชนี ω จะเพิ่มขึ้นเมื่อความยาวของแบบสอบเพิ่มขึ้น แต่การเปลี่ยนแปลงของจำนวนผู้สอบจะไม่มีอิทธิพลต่ออำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของดัชนี ω ซึ่งข้อค้นพบนี้สอดคล้องกับข้อค้นพบจากงานวิจัยของ Wollack (1997) และ Sotaridoan และ Meijer (2002) แต่ จำนวนผู้สอบ และ ความยาวของแบบสอบ ที่เพิ่มขึ้นจะมีผลทำให้อำนาจการตรวจจับการลอกข้อสอบ (Detection rate) ของค่าดัชนี K_2, S_1 และ S_2 มีค่าเพิ่มขึ้น อาจกล่าวได้ว่าไม่มีดัชนีตรวจจับการลอกข้อสอบใดที่จะใช้ได้ดีที่สุดในทุกสถานการณ์ที่ถูกจัดกระทำขึ้น ดัชนีตรวจจับการลอกข้อสอบแต่ละตัวสามารถใช้ได้อย่างมีประสิทธิภาพแตกต่างกันในสถานการณ์ที่ต่างกัน แต่โดยรวมแล้วดัชนี S_1 มีอำนาจการตรวจจับการลอกข้อสอบสูงกว่าดัชนี K_2 และดัชนี S_2 มีการพัฒนาของอำนาจการตรวจจับการลอกข้อสอบมากขึ้นอย่างมีนัยสำคัญหากเปรียบเทียบกับดัชนี K และ K_2

ซึ่งจากการศึกษางานวิจัยสองงาน คือ งานวิจัยเรื่องคุณสมบัติทางสถิติของดัชนี K สำหรับการตรวจจับการลอกข้อสอบ (Statistical Properties of the K-Index for detecting answer copying) และ เรื่องสถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Two new statistics to detect answer copying) ของผู้วิจัยสองท่านคือ Leonardo S. Sotaridona และ Rob R. Meijer อย่างละเอียดในด้าน วัตถุประสงค์ของการวิจัย ระเบียบวิธีวิจัย ตัวแปรต้น ตัวแปรตาม จำนวนปัจจัยหรือตัวแปรต้นที่ศึกษา จำนวนสถานการณ์ที่ศึกษา แบบสอบที่ใช้ จำนวนตัวเลือกในแบบ

สอบที่ใช้ และโปรแกรมที่ใช้ในการวิเคราะห์ข้อมูล ผู้วิจัยสามารถสรุปเป็นตารางวิเคราะห์งานวิจัย และข้อสรุปเกี่ยวกับแนวทางในการวิจัยของผู้วิจัยได้ดังต่อไปนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 2 วิเคราะห์งานวิจัยเรื่อง คุณสมบัติทางสถิติของดัชนี K สำหรับการตรวจจับการลอกข้อสอบ (Sataridona และ Meijer, 2002) และ เรื่องสถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Sataridona และ Meijer, 2003) และข้อสรุปเกี่ยวกับแนวทางในการวิจัยของผู้วิจัย

หัวข้อ	Sataridona และ Meijer (2002)	Sataridona และ Meijer (2003)	**ข้อสรุป**
1. วัตถุประสงค์	เปรียบเทียบ type I error และ detection rate ระหว่างค่าดัชนี K*, K1,K2 และ ค่าสถิติ ω	เปรียบเทียบ type I error และ detection rate ระหว่างค่าดัชนี K2, S1,S2 และ ค่าสถิติ ω	เปรียบเทียบ type I error และ detection rate ระหว่างค่าดัชนี K2, S1,S2 และ ค่าสถิติ ω
2. ระเบียบวิธีวิจัย	เชิงทดลอง	เชิงทดลอง	เชิงทดลอง
3. ตัวแปรต้นที่ศึกษา			
3.1 ความยาวของแบบสอบ	**40 และ 80 ข้อ**	**40 และ 80 ข้อ**	**35 และ 65 ข้อ**
3.2 จำนวนผู้สอบ	**100, 500 และ 2,000 คน**	**100 และ 500 คน**	**100 และ 250 คน**
3.3 ระดับความสามารถของผู้ให้ลอก	**เปอร์เซ็นไทล์ที่ 90 และ 60**	ช่วงเปอร์เซ็นไทล์ 40-90 (สุ่มเลือก) (ไม่ได้ศึกษาปัจจัยนี้)	**เปอร์เซ็นไทล์ที่ 90 และ 60**
3.4 ร้อยละของจำนวนข้อสอบที่ถูกลอก	**ร้อยละ 10,20,30 และ 40 ของข้อสอบทั้งหมด**	**ร้อยละ 10,20,30 และ 40 ของข้อสอบทั้งหมด**	**ร้อยละ 10,20,30 และ 40 ของข้อสอบทั้งหมด**
3.5 ร้อยละของจำนวนผู้ลอก	ร้อยละ 5 (ไม่ได้ศึกษาปัจจัยนี้)	ร้อยละ 5 (ไม่ได้ศึกษาปัจจัยนี้)	**ร้อยละ 5 และ ร้อยละ 10**
3.6 วิธีการลอก	(ไม่ได้ศึกษาปัจจัยนี้)	(ไม่ได้ศึกษาปัจจัยนี้)	**2 วิธี คือ ลอกแบบสุ่ม และ ลอกเฉพาะข้อยาก**
4. จำนวนปัจจัยที่ศึกษา	4 ปัจจัย (4 factors)	3 ปัจจัย (3 factors)	6 ปัจจัย (6 factors)
5. จำนวนสถานการณ์ที่ศึกษา	2x3x4x2 = 48 สถานการณ์	2x2x4 = 16 สถานการณ์	2x2x2x4x2x2 = 128 สถานการณ์
6. ตัวแปรตาม	type I error และdetection rate	type I error และdetection rate	type I error และdetection rate
7. แบบสอบที่ใช้	Mathematic college placement test	English college placement test และ Mathematic college placement test	แบบสอบปลายภาคของรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา จำนวน 65 ข้อ
8. จำนวนตัวเลือกในแบบสอบที่ใช้	5 ตัวเลือก	5 ตัวเลือก	4 ตัวเลือก
9. โปรแกรมที่ใช้ในการวิเคราะห์ข้อมูล	S-plus MULTILOG	S-plus MULTILOG	S-plus MULTILOG

จากการศึกษางานวิจัยต่างประเทศซึ่งทำการศึกษาและเปรียบเทียบคุณสมบัติทางสถิติทางสถิติของดัชนีตรวจจับการลอกข้อสอบนั้นพบว่า ดัชนีตรวจจับการลอกข้อสอบซึ่งถูกพัฒนาขึ้นและถูกใช้การศึกษาในปัจจุบันมีอยู่ 4 ค่าคือ ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω โดยดัชนีตรวจจับการลอกข้อสอบทั้ง 4 ดัชนีมีประสิทธิภาพแตกต่างกันไปตามเงื่อนไขที่แตกต่างกัน ดังนั้นผู้วิจัยจึงเลือกที่จะศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ดัชนีดังกล่าว

วิธีดำเนินการวิจัย ผู้วิจัยใช้ระเบียบวิธีวิจัยเชิงทดลอง (Experimental Research) ตามงานวิจัยต่างประเทศของ Sataridona และ Meijer (2002 และ 2003) ซึ่งเป็นการนำเอาข้อมูลเชิงประจักษ์ (Empirical data) มาจัดกระทำด้วยวิธี การจำลองสถานการณ์ (Simulation)

ทางด้านตัวแปร จากการศึกษางานวิจัยต่างประเทศของ Sataridona และ Meijer (2002 และ 2003) ผู้วิจัยได้ข้อสรุปด้านตัวแปรดังนี้ ผู้วิจัยศึกษาและเปรียบเทียบ ตัวแปรตาม คือ ความคลาดเคลื่อนประเภทที่ 1 (Type I error) และ อำนาจการตรวจจับการลอกข้อสอบ (detection rate) ส่วนตัวแปรต้น ผู้วิจัยจะทำการศึกษาตัวแปรต้นตามงานวิจัยต่างประเทศของ Sataridona และ Meijer (2002 และ 2003) จำนวน 4 ตัวแปร คือ ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก และศึกษาตัวแปรต้นเพิ่มเติมจากข้อเสนอแนะจากวิจัย เรื่อง สถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Two new statistics to detect answer copying) ที่ทำโดย Leonardo S. Sotaridona และ Rob R. Meijer (2003) จำนวน 2 ตัวแปร คือ ร้อยละของจำนวนผู้ลอก และ วิธีการลอก

ตัวแปรต้น ด้าน ความยาวของแบบสอบ ผู้วิจัยศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบภายใต้เงื่อนไข ด้านความยาวของแบบสอบจำนวน 2 ระดับ คือ แบบสอบความยาว 65 ข้อ และ แบบสอบความยาว 35 ข้อ ซึ่งแบบสอบความยาว 65 ข้อเป็นตัวแทนของแบบสอบขนาดยาวที่มีข้อคำถามมาก และ แบบสอบความยาว 35 ข้อเป็นตัวแทนของแบบสอบขนาดสั้นที่มีข้อคำถามไม่มาก โดยสาเหตุที่ผู้วิจัยไม่ศึกษาและเปรียบเทียบแบบสอบความยาว 40 ข้อ และ 80 ข้อ ตามงานวิจัยของ Sataridona และ Meijer (2002 และ 2003) เนื่องจากผู้วิจัยใช้ฐานข้อมูลการสอบปลายภาคประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษาต้น ปีการศึกษา 2548 ซึ่งลักษณะของแบบสอบจะแบบออกเป็นสองส่วน คือ ส่วนข้อสอบปรนัย และ ข้อสอบอัตนัย ซึ่งผู้วิจัยนำผลการตอบข้อสอบเฉพาะส่วนข้อสอบปรนัยมาใช้เท่านั้น ซึ่งส่วนของข้อสอบปรนัยของแบบสอบชุดดังกล่าวมีจำนวนข้อสอบทั้งสิ้น 65 ข้อ ดังนั้นผู้วิจัยจึงศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบภายใต้เงื่อนไขด้าน ความยาวของแบบสอบจำนวนสองระดับ คือ แบบสอบความยาว 65 ข้อ และ แบบสอบความยาว 35 ข้อ ซึ่งเป็นแบบสอบที่มีความยาวประมาณครึ่งหนึ่งของแบบสอบ

ความยาว 65 ข้อ โดยแบบสอบความยาว 35 ข้อนั้นได้จากการสุ่มข้อสอบจากแบบสอบความยาว 65 ข้อด้วย วิธีการสุ่มอย่างง่าย

ตัวแปรต้น ด้าน จำนวนผู้สอบ ผู้วิจัยศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบภายใต้เงื่อนไข ด้านจำนวนผู้สอบ 2 ขนาด คือ ขนาด 250 คน และ 100 คน ซึ่งกลุ่มผู้สอบขนาด 100 คนเป็นตัวแทนของกลุ่มผู้สอบขนาดเล็ก และ กลุ่มจำนวนผู้สอบขนาด 250 คนเป็นตัวแทนของกลุ่มผู้สอบขนาดใหญ่ โดยสาเหตุที่ผู้วิจัยไม่ศึกษาและเปรียบเทียบจำนวนผู้สอบขนาด 100 คน และ 500 คน ตามงานวิจัยของ Sataridona และ Meijer (2003) เนื่องจากผู้วิจัยใช้ฐานข้อมูลการสอบปลายภาคประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษาต้น ปีการศึกษา 2548 ซึ่งมีผู้ที่เข้าสอบทั้งสิ้น 250 คน ดังนั้นผู้วิจัยจึงศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบภายใต้เงื่อนไขด้าน จำนวนผู้สอบ เพียง 2 ขนาด คือ ผู้สอบจำนวน 250 คน และ ผู้สอบจำนวน 100 คน โดยกลุ่มผู้สอบจำนวน 100 คนนั้นได้จากการสุ่มผู้สอบจากกลุ่มผู้สอบจำนวน 250 คนด้วยวิธีการสุ่มอย่างง่าย

ตัวแปรต้น ด้าน ระดับความสามารถของผู้ให้ลอก ผู้วิจัยศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบภายใต้เงื่อนไข ด้านระดับความสามารถของผู้ให้ลอก 2 ระดับ คือ ผู้ให้ลอกที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 90 และ 60 ตามงานวิจัยของ Sataridona และ Meijer (2002) โดยผู้ให้ลอกที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 60 เป็นตัวแทนของผู้ให้ลอกที่มีความรู้ความสามารถระดับปานกลาง และ ผู้ให้ลอกที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 90 เป็นตัวแทนของผู้ให้ลอกที่มีความรู้ความสามารถระดับสูง

ตัวแปรต้น ด้าน ร้อยละของจำนวนข้อสอบที่ถูกลอก ผู้วิจัยศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบภายใต้เงื่อนไข ด้านร้อยละของจำนวนข้อสอบที่ถูกลอกจำนวน 4 ระดับ คือ ร้อยละ 10, 20, 30 และ 40 ของข้อสอบทั้งหมด ตามงานวิจัยของ Sataridona และ Meijer (2002 และ 2003) โดยร้อยละของจำนวนข้อสอบที่ถูกลอกทั้ง 4 ระดับได้แก่ ร้อยละ 10, 20, 30 และ 40 เป็นตัวแทนของจำนวนข้อสอบที่ถูกลอกจากระดับที่น้อยจนไปถึงระดับที่มากตามลำดับ

จากงานวิจัยเรื่อง เรื่องสถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Two new statistics to detect answer copying) ของ Leonardo S. Sotaridona และ Rob R. Meijer (2003) นั้นนักวิจัยของงานวิจัยดังกล่าวทั้งสองท่านได้ให้ข้อเสนอแนะไว้ว่าควรมีการศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบอันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ในปัจจัยหรือตัวแปรต้นด้าน วิธีการลอกข้อสอบ (varying way of

answer copying) และ ด้านร้อยละของจำนวนผู้ลอก (different percentage of copiers) เพิ่มเติม

ตัวแปรต้น ด้านวิธีการลอกข้อสอบ (varying way of answer copying) นั้นจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องพบว่า Hanson et al (1987) ได้ทำการศึกษาทดลองเกี่ยวกับดัชนีตรวจจับการลอกข้อสอบ ซึ่งได้ทำการจำลองการลอกข้อสอบไว้ 5 วิธี คือ

1) การลอกข้อสอบแบบสุ่ม

2) การลอกข้อสอบเฉพาะข้อยาก

3) การลอกข้อสอบในช่วงแรกของชุดข้อสอบ

4) การลอกข้อสอบในช่วงท้ายของชุดข้อสอบ

5) การลอกข้อสอบเป็นช่วง ช่วงละ 5 ข้อ ซึ่งแต่ละช่วงข้อสอบได้รับการเลือกอย่างสุ่ม

ผู้วิจัยทำการสัมภาษณ์นิสิตระดับปริญญาบัณฑิตชั้นปีที่ 1 ถึง ชั้นปีที่ 3 ของคณะวิทยาศาสตร์ พาณิชยศาสตร์และการบัญชี และ ครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย รวมกัน 3 คณะ จำนวนทั้งสิ้น 80 คน แบ่งเป็นเพศชาย 40 คน และ หญิง 40 คน โดยทำการสัมภาษณ์ในวันที่ 9 สิงหาคม พ.ศ. 2549 และ วันที่ 10 สิงหาคม พ.ศ. 2549 โดยคำถามที่ใช้ในการสัมภาษณ์คือ

“หากท่านลอกข้อสอบหลายตัวเลือก โดยปกติท่านลอกด้วยวิธีการแบบใด” โดยตัวเลือกมีดังนี้ คือ

1) การลอกข้อสอบแบบสุ่ม กล่าวคือ ลอกคำตอบข้อใดก็ได้เมื่อมีโอกาส

2) การลอกข้อสอบเฉพาะข้อยาก

3) การลอกข้อสอบในช่วงแรกของชุดข้อสอบ

4) การลอกข้อสอบในช่วงท้ายของชุดข้อสอบ

5) การลอกข้อสอบเป็นช่วงๆ ช่วงละ 3 ข้อ, 4 ข้อ หรือ 5 ข้อ เป็นต้น ซึ่งแต่ละช่วงข้อสอบที่ลอกไม่สามารถกำหนดได้ว่าจะเป็นช่วงใด ขึ้นอยู่กับโอกาส

ผลการสัมภาษณ์ ปรากฏว่า มีเพียงสามตัวเลือกเท่านั้นที่นิสิตทั้งสามคณะตอบ คือ ตัวเลือกที่ 1) การลอกข้อสอบแบบสุ่ม 2) การลอกข้อสอบเฉพาะข้อยาก และ 5) การลอกข้อสอบเป็นช่วง ช่วงละ 5 ข้อ ซึ่งแต่ละช่วงข้อสอบได้รับการเลือกอย่างสุ่ม

โดยสามารถแสดงผลการตอบแยกตามคณะ และ เพศ ได้ดังตาราง 3

ตาราง 3 ความถี่ของผู้ตอบตัวเลือกต่างๆ ของนิสิตคณะพาณิชยศาสตร์และการบัญชี คณะครุศาสตร์ และ คณะวิทยาศาสตร์

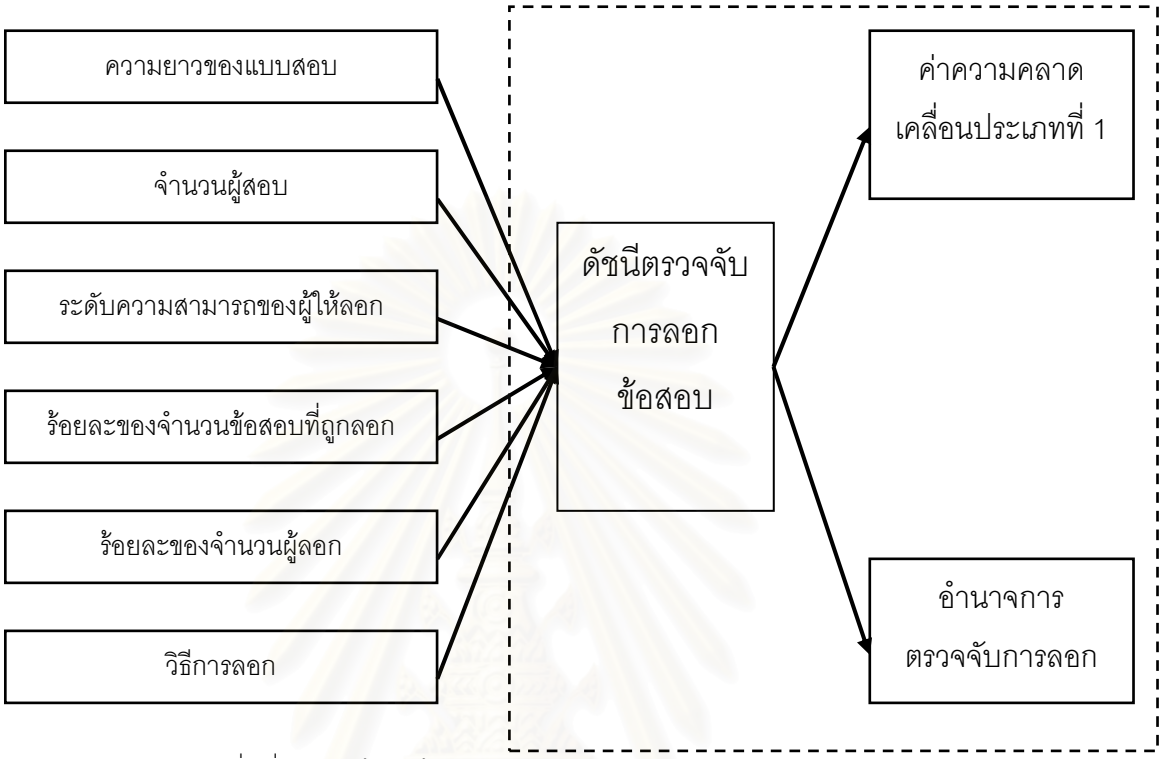
ตัวเลือก	คณะพาณิชยศาสตร์และการบัญชี (30 คน)		คณะครุศาสตร์ (20 คน)		คณะวิทยาสตร์ (30 คน)		รวม
	ชาย (15 คน)	หญิง (15 คน)	ชาย (11 คน)	หญิง (9 คน)	ชาย (15 คน)	หญิง (15 คน)	
1)การลอกข้อสอบแบบสุ่ม	9	7	5	3	6	6	36
2)การลอกข้อสอบเฉพาะข้อยาก	6	8	6	6	8	9	43
3)การลอกข้อสอบในช่วงแรกของชุดข้อสอบ	0	0	0	0	0	0	0
4)การลอกข้อสอบในช่วงท้ายของชุดข้อสอบ	0	0	0	0	0	0	0
5)การลอกข้อสอบเป็นช่วง ช่วงละ 5 ข้อ ซึ่งแต่ละช่วงข้อสอบได้รับการเลือกอย่างสุ่ม	0	0	0	0	1	0	1

จากผลการสัมภาษณ์ข้างต้นผู้วิจัยจึงเลือกวิธีการลอกมาใช้ในการวิจัยเพียง 2 วิธี คือ การลอกข้อสอบแบบสุ่ม และ การลอกข้อสอบเฉพาะข้อยาก

ด้านร้อยละของจำนวนผู้ลอก (different percentage of copiers) จากการวิเคราะห์งานวิจัยสองเรื่อง คือ งานวิจัยเรื่องคุณสมบัติทางสถิติของดัชนี K สำหรับการตรวจจับการลอกข้อสอบ (Statistical Properties of the K-Index for detecting answer copying) และ เรื่องสถิติตรวจจับการลอกข้อสอบแบบใหม่ 2 แบบ (Two new statistics to detect answer copying) ของผู้วิจัยสองท่านคือ Leonardo S. Sotaridona และ Rob R. Meijer พบว่างานวิจัยทั้งสองชิ้นใช้ร้อยละของจำนวนผู้ลอกเท่ากัน คือ ร้อยละ 5 ดังนั้นผู้วิจัยจึงกำหนดร้อยละของจำนวนผู้ลอกเพื่อศึกษาในงานวิจัยเพิ่มเติมคือ ร้อยละ 10 ดังนั้นในงานวิจัยชิ้นนี้ของผู้วิจัยจึงศึกษาปัจจัยด้านร้อยละของจำนวนผู้ลอก 2 ระดับ คือ ร้อยละ 5 และ ร้อยละ 10 โดยจำนวนผู้ลอกร้อยละ 5 เป็นตัวแทนของจำนวนผู้ลอกในการสอบที่น้อย และ จำนวนผู้ลอกร้อยละ 10 เป็นตัวแทนของจำนวนผู้ลอกในการสอบที่มาก

กรอบแนวคิดในการวิจัย

จากการศึกษาแนวคิด ทฤษฎี และ งานวิจัยที่เกี่ยวข้องกับดัชนีตรวจจับการลอกข้อสอบ ผู้วิจัยสามารถเขียนกรอบแนวคิดในการวิจัยได้ ดังนี้



หมายเหตุ กรอบเส้นประสี่เหลี่ยมแสดงถึงการศึกษาและการเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ
ตรวจจับการลอกข้อสอบ

ภาพ 3 กรอบความคิดของการวิจัย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีดำเนินการวิจัย

การศึกษาเรื่อง “การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ” ใช้ระเบียบวิธีวิจัยเชิงทดลอง (Experimental Research) ในการศึกษา โดยมีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบคุณสมบัติทางสถิติซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้เงื่อนไขตัวแปรต้นที่แตกต่างกันในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก

ผู้วิจัยใช้ข้อมูลผลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ส่วนข้อสอบปรนัย ของนิสิตระดับปริญญาบัณฑิตที่ลงทะเบียนเรียนและเข้าสอบในรายวิชาดังกล่าว จำนวนทั้งสิ้น 250 คน ซึ่งเป็นข้อมูลเชิงประจักษ์ (empirical data) มาจัดกระทำด้วยวิธี การจำลองสถานการณ์ (Simulation) และ วิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์ซึ่งผู้วิจัยเป็นผู้เขียนคำสั่งที่ใช้ในการวิเคราะห์ประมวลผลสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัวภายใต้ความแตกต่างของเงื่อนไขตัวแปรต้นในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก โดยมีขั้นตอนการดำเนินการวิจัยดังแผนภาพต่อไปนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาพ 4 ขั้นตอนในการดำเนินการวิจัย

ขั้นตอนการดำเนินการวิจัย

1. ศึกษาเอกสาร และงานวิจัยที่เกี่ยวข้องกับดัชนีตรวจจับการลอกข้อสอบ เพื่อให้ได้ประเด็นปัญหาที่สนใจ นิยามศัพท์ที่ใช้ในการวิจัยและกรอบมโนทัศน์ในการวิจัย

2. ออกแบบการวิจัย ซึ่งเป็นการวิจัยเชิงทดลอง ใช้ข้อมูลผลการตอบข้อสอบซึ่งเป็นข้อมูลเชิงประจักษ์ มาจัดกระทำด้วยวิธีการจำลองสถานการณ์ และ วิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์

3. ศึกษาและสังเคราะห์วิธีการดำเนินการวิจัยของงานวิจัยเพื่อคัดเลือกปัจจัยที่จะใช้ศึกษา

4. ติดต่อผู้เกี่ยวข้องเพื่อขออนุญาตใช้ผลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ของนิสิตระดับปริญญาบัณฑิตที่ลงทะเบียนและเข้าสอบรายวิชาดังกล่าว

5. ดำเนินการทดลองกับผลการตอบข้อสอบของกลุ่มตัวอย่างหรือประชากร ด้วยวิธีการจำลองสถานการณ์ (Simulation) และ วิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์ซึ่งผู้วิจัยเป็นผู้เขียนคำสั่งที่ใช้ในการวิเคราะห์ประมวลผลสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัวภายใต้ความแตกต่างของเงื่อนไขตัวแปรต้นในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และวิธีการลอก

6. สรุปและอภิปรายผลการวิจัย

ประชากร กลุ่มตัวอย่าง และ การสุ่มกลุ่มตัวอย่าง

ประชากร ได้แก่ นิสิตระดับปริญญาบัณฑิต ที่ลงทะเบียนและเข้าสอบในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา

กลุ่มตัวอย่าง ได้แก่ นิสิตระดับปริญญาบัณฑิต ที่ลงทะเบียนและเข้าสอบในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 จำนวน 250 คน และ 100 คน

เนื่องจากขนาดของกลุ่มผู้สอบเป็นตัวแปรต้นหนึ่งที่ผู้วิจัยศึกษา ซึ่งขนาดของกลุ่มผู้สอบจะมีสองขนาด คือ 100 คน และ 250 คน ดังนั้น สำหรับกลุ่มผู้สอบขนาดใหญ่ คือ 250 คน ผู้วิจัยจะไม่ทำการสุ่ม แต่ผู้วิจัยจะใช้ผลการสอบของนิสิตที่ลงทะเบียนเรียนและเข้าสอบในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ทั้งหมด ส่วนกลุ่มผู้สอบขนาดเล็ก คือ 100 คน ผู้วิจัยจะใช้วิธีการสุ่มอย่างง่ายสุ่มผู้สอบจำนวน 100 คน

จากกลุ่มผู้สอบจำนวน 250 คนที่เป็นนิสิตที่ลงทะเบียนเรียนและเข้าสอบในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548

ตัวแปรที่เกี่ยวข้อง

1.ตัวแปรต้น ประกอบด้วย

1. ความยาวของแบบสอบ แบ่งออกเป็นแบบสอบขนาดยาว 35 ข้อ และ 65 ข้อ
2. จำนวนผู้สอบ แบ่งออกเป็น จำนวนผู้สอบขนาด 100 และ 250 คน
3. ระดับความสามารถของผู้ให้ลอก แบ่งออกเป็น ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 และ 60
4. ร้อยละของจำนวนผู้ลอก แบ่งออกเป็น ร้อยละ 5 และ ร้อยละ 10 ของผู้สอบทั้งหมด
5. ร้อยละของจำนวนข้อสอบที่ถูกลอก แบ่งออกเป็น ร้อยละ 10, 20, 30 และ 40 ของข้อสอบทั้งหมด
6. วิธีการลอก แบ่งออกเป็น 2 วิธี คือ
    6.1) การลอกข้อสอบแบบสุ่ม
    6.2) การลอกข้อสอบเฉพาะข้อยาก

2.ตัวแปรตาม คือ ความคลาดเคลื่อนประเภทที่ 1 (Type I error) และ อำนาจของการตรวจจับการลอกข้อสอบ (detection rates)

การเก็บรวบรวมข้อมูล และ เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบปลายภาคประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ซึ่งถือเป็นแบบสอบมาตรฐาน

การเก็บรวบรวมข้อมูล งานวิจัยชิ้นนี้ผู้วิจัยไม่ได้ดำเนินการเก็บข้อมูลจริง แต่เป็นการนำเอาผลการตอบข้อสอบส่วนข้อสอบปรนัย ของนิสิตที่ลงทะเบียนและเข้าสอบ ในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 จำนวน 250 คน ซึ่งเป็นข้อมูลทุติยภูมิ (Secondary Data) และ เป็นข้อมูลเชิงประจักษ์ (empirical data) มาใช้ในการทดลอง และ วิเคราะห์ข้อมูล

ขั้นตอนการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (Simulation)

หลังจากที่ผู้วิจัยติดต่อผู้ที่เกี่ยวข้องเพื่อขอนำผลการทดสอบและแบบสอบปลายภาคประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปี

การศึกษา 2548 มาใช้ ซึ่งผู้วิจัยนำเพียงรูปแบบการตอบข้อสอบของกลุ่มตัวอย่างและประชากรมาใช้ในการจัดกระทำและวิเคราะห์แต่เพียงส่วนข้อสอบหลายตัวเลือก (4 ตัวเลือก) จำนวน 65 ข้อ

เนื่องด้วยงานวิจัยชิ้นนี้ผู้วิจัยได้ทำการศึกษาตัวแปรต้นหลายตัวแปร ดังนั้น เพื่อความสะดวกและความเหมาะสมในการนำเสนอข้อมูลต่างๆ ผู้วิจัยจึงกำหนดสัญลักษณ์ย่อภาษาอังกฤษสำหรับตัวแปรต้นที่ผู้วิจัยทำการศึกษา ดังนี้

1. ความยาวของแบบสอบ (test length) ย่อเป็น L
   แบ่งออกเป็น
   1.1 แบบสอบขนาดยาว 35 ข้อ ย่อเป็น L35
   1.2 แบบสอบขนาดยาว 65 ข้อ ย่อเป็น L65
2. จำนวนผู้สอบ (the number of simulees) ย่อเป็น N
   แบ่งออกเป็น
   2.1 จำนวนผู้สอบขนาด 100 คน ย่อเป็น N100
   2.2 จำนวนผู้สอบขนาด 250 คน ย่อเป็น N250
3. ระดับความสามารถของผู้ให้ลอก (the ability levels of source) ย่อเป็น S
   แบ่งออกเป็น
   3.1 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 ย่อเป็น S90th
   3.2 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 ย่อเป็น S60th
4. ร้อยละของจำนวนผู้ลอก (the percentage of copiers) ย่อเป็น C
   แบ่งออกเป็น
   4.1 ร้อยละ 5 ของจำนวนผู้สอบทั้งหมด ย่อเป็น C5%
   4.2 ร้อยละ 10 ของจำนวนผู้สอบทั้งหมด ย่อเป็น C10%
5. ร้อยละของจำนวนข้อสอบที่ถูกลอก (the percentage of answer copied) ย่อเป็น A
   แบ่งออกเป็น
   5.1 ร้อยละ 10 ของจำนวนข้อสอบทั้งหมด ย่อเป็น A10%
   5.2 ร้อยละ 20 ของจำนวนข้อสอบทั้งหมด ย่อเป็น A20%
   5.3 ร้อยละ 30 ของจำนวนข้อสอบทั้งหมด ย่อเป็น A30%
   5.4 ร้อยละ 40 ของจำนวนข้อสอบทั้งหมด ย่อเป็น A40%
6. วิธีการลอก (type of copying) ย่อเป็น T
   แบ่งออกเป็น
   6.1 การลอกข้อสอบแบบสุ่ม ย่อเป็น T(R)
   6.2 การลอกข้อสอบเฉพาะข้อยาก ย่อเป็น T(H)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากตัวแปรต้นทั้ง 6 ตัวแปรข้างต้น คือ ความยาวของแบบสอบ (แบ่งเป็น 2 ระดับย่อย), จำนวนผู้สอบ(แบ่งเป็น 2 ระดับย่อย), ระดับความสามารถของผู้ให้ลอก(แบ่งเป็น 2 ระดับย่อย), ร้อยละของจำนวนผู้ลอก(แบ่งเป็น 2 ระดับย่อย), ร้อยละของจำนวนข้อสอบที่ถูกลอก(แบ่งเป็น 4 ระดับย่อย) และ วิธีการลอก(แบ่งเป็น 2 ระดับย่อย) ทำให้เราได้จำนวนสถานการณ์ที่ต้องจัดกระทำข้อมูลรวมทั้งสิ้น 2 x 2 x 2 x 2 x 4 x 2 = 128 สถานการณ์ โดยในแต่ละสถานการณ์จะมีการวิเคราะห์ข้อมูลซ้ำ (replication) จำนวน 100 รอบ

รายละเอียดในการจัดกระทำข้อมูลแต่ละสถานการณ์

ในแต่ละสถานการณ์ จากจำนวนสถานการณ์ทั้งสิ้น 128 สถานการณ์ที่ผู้วิจัยต้องทำการจัดกระทำข้อมูลมีรายละเอียดดังต่อไปนี้

ปัจจัยด้าน ความยาวของแบบสอบ และ จำนวนผู้สอบ

1. เมื่อได้รับผลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ในส่วนข้อสอบหลายตัวเลือก (4 ตัวเลือก) จำนวน 65 ข้อ ของผู้สอบจำนวน 250 คนมาแล้ว ผู้วิจัยทำการสุ่มผู้สอบจำนวน 100 คน จากกลุ่มผู้สอบจำนวน 250 คน และ สุ่มข้อสอบ 35 ข้อจากแบบสอบความยาว 65 ข้อ ด้วยวิธีการสุ่มอย่างง่าย เพื่อสร้างชุดข้อมูล 4 ชุด ตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ ดังนี้

1.1 แบบสอบขนาด 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100)

1.2 แบบสอบขนาด 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250)

1.3 แบบสอบขนาด 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100)

1.4 แบบสอบขนาด 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250)

2. นำข้อมูลทั้ง 4 ชุดไปวิเคราะห์ตรวจให้คะแนนแก่ผู้สอบแต่ละคน รวมทั้งวิเคราะห์ข้อมูลเพื่อหาค่า คะแนนต่ำสุด คะแนนสูงสุด ค่าเฉลี่ย ค่ามัธยฐาน ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน ความเบ้ ความโด่ง ด้วยโปรแกรม TAP แล้วนำผลการวิเคราะห์มาศึกษาลักษณะโดยภาพรวมของกลุ่มผู้สอบในชุดข้อมูลแต่ละชุด

3. นำข้อมูลทั้ง 4 ชุดไปวิเคราะห์หาค่าความยากง่ายและอำนาจจำแนกของข้อสอบ รวมทั้ง ความเที่ยงแบบสอดคล้องภายใน ตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) ด้วยโปรแกรม TAP แล้วนำผลการวิเคราะห์มาศึกษาลักษณะโดยภาพรวมของชุดข้อสอบ

4. นำข้อมูลทั้ง 4 ชุดไปวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ และ พารามิเตอร์ระดับความสามารถของผู้สอบ (θ) ตามทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories) โดยใช้โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (Three-Parameter Model) ด้วยโปรแกรม

MULTILOG จากนั้นนำผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบไปใช้ในการคัดเลือกข้อสอบที่จะอยู่ในกลุ่มข้อสอบระดับยาก และ ทำการเรียงลำดับผู้สอบในแต่ละชุดข้อมูลตามค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) จากมากไปน้อย

5. นำข้อมูลทั้ง 4 ชุดไปวิเคราะห์หาค่าพารามิเตอร์ความชันของแต่ละรายการคำตอบ (Slop of the trace lines) และ ค่าพารามิเตอร์จุดตัดของแต่ละรายการคำตอบ (Intercept parameter) ตามทฤษฎีการทดสอบแนวใหม่แบบโมเดล Nominal Response Model ด้วยโปรแกรม MULTILOG เพื่อนำค่าพารามิเตอร์ที่ได้ไปวิเคราะห์หาดัชนี ω ในภายหลัง

ปัจจัยด้านระดับความสามารถของผู้ให้ลอก และ ร้อยละของจำนวนผู้ลอก

6. จากนั้นนำชุดข้อมูลทั้ง 4 ชุดมาศึกษาตามเงื่อนไขของตัวแปรต้นเกี่ยวกับระดับความสามารถของผู้ให้ลอก และ ร้อยละของจำนวนผู้ลอกของผู้สอบทั้งหมด กล่าวคือ ในแต่ละชุดข้อมูลให้เลือกผู้สอบที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 และ 60 มาเพื่อให้เป็น ผู้ให้ลอก (source)

จากนั้นทำการสุ่มผู้ลอก (copiers) ด้วยวิธีการสุ่มอย่างง่าย โดยการสุ่มมีข้อตกลงเบื้องต้นว่า ให้สุ่มผู้สอบที่มีตำแหน่งเปอร์เซ็นไทล์ ต่ำกว่าตำแหน่งเปอร์เซ็นต์ไทล์ของผู้ให้ลอกเท่านั้น โดยสุ่มผู้ลอกมา 2 ขนาดคือ ขนาดร้อยละ 5 และ ร้อยละ 10 ของผู้สอบทั้งหมด กล่าวคือ ถ้าชุดข้อมูลมีจำนวนผู้สอบ 100 คน ก็ทำการสุ่มผู้ลอกจำนวน 5 คน สำหรับขนาดผู้ลอกร้อยละ 5 และ สุ่มผู้ลอกจำนวน 10 คนสำหรับขนาดผู้ลอกร้อยละ 10 และ ถ้าชุดข้อมูลมีจำนวนผู้สอบ 250 คน ก็ทำการสุ่มผู้ลอกจำนวน 13 คน สำหรับขนาดผู้ลอกร้อยละ 5 และ สุ่มผู้ลอกจำนวน 25 คน สำหรับขนาดผู้ลอกร้อยละ 10

ดังนั้นในแต่ละชุดข้อมูลทั้ง 4 ชุดที่ผู้วิจัยทำการแบ่งไว้ก่อนหน้านี้ก็จะมีการศึกษาในเงื่อนไขเกี่ยวกับร้อยละของจำนวนผู้ลอกและผู้ให้ลอกดังนี้ คือ

6.1 ผู้ลอกร้อยละ 5 ของจำนวนผู้สอบทั้งหมด ลอกข้อสอบจากผู้ให้ลอกที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 90

6.2 ผู้ลอกร้อยละ 5 ของจำนวนผู้สอบทั้งหมด ลอกข้อสอบจากผู้ให้ลอกที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 60

6.3 ผู้ลอกร้อยละ 10 ของจำนวนผู้สอบทั้งหมด ลอกข้อสอบจากผู้ให้ลอกที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 90

6.4 ผู้ลอกร้อยละ 10 ของจำนวนผู้สอบทั้งหมด ลอกข้อสอบจากผู้ให้ลอกที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 60

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

โดยในแต่ละกลุ่มเงื่อนไขนั้นผู้ให้ลอกจะมีเพียง 1 คนเท่านั้น (ผู้สอบที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 60 หรือ 90) ส่วนผู้ลอกจะมีจำนวนร้อยละ 5 หรือ 10 ของผู้สอบทั้งหมดขึ้นอยู่กับเงื่อนไขตัวแปรต้นด้านร้อยละของจำนวนผู้ลอกที่ต่างกัน

ปัจจัยด้านร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก

7. จากนั้นในแต่ละชุดข้อมูลที่ถูกแยกย่อยออกมา ก็จะต้องทำการศึกษาเพิ่มเติมในเงื่อนไขเกี่ยวกับร้อยละของจำนวนข้อสอบที่ถูกลอก ซึ่งแบ่งเป็น 4 ระดับคือ ร้อยละ 10, 20, 30 และ 40 ของข้อสอบทั้งหมด และ วิธีการลอกที่แตกต่างกัน ซึ่งแบ่งออกเป็น 2 วิธี คือ

7.1) การลอกข้อสอบแบบสุ่ม

7.2) การลอกข้อสอบเฉพาะข้อยาก

ในส่วนเงื่อนไขการลอกข้อสอบแบบสุ่มผู้วิจัยใช้วิธีการสุ่มอย่างง่ายสุ่มข้อสอบให้ได้ตามจำนวนร้อยละของจำนวนข้อสอบที่ถูกลอกทั้ง 4 ระดับ กล่าวคือ

สำหรับข้อสอบจำนวน 35 ข้อ

ทำการสุ่มข้อสอบจำนวน 4 ข้อ สำหรับขนาดของข้อสอบที่ถูกลอกร้อยละ 10

ทำการสุ่มข้อสอบจำนวน 7 ข้อ สำหรับขนาดของข้อสอบที่ถูกลอกร้อยละ 20

ทำการสุ่มข้อสอบจำนวน 11 ข้อ สำหรับขนาดของข้อสอบที่ถูกลอกร้อยละ 30

ทำการสุ่มข้อสอบจำนวน 14 ข้อ สำหรับขนาดของข้อสอบที่ถูกลอกร้อยละ 40

สำหรับข้อสอบจำนวน 65 ข้อ

ทำการสุ่มข้อสอบจำนวน 7 ข้อ สำหรับขนาดของข้อสอบที่ถูกลอกร้อยละ 10

ทำการสุ่มข้อสอบจำนวน 13 ข้อ สำหรับขนาดของข้อสอบที่ถูกลอกร้อยละ 20

ทำการสุ่มข้อสอบจำนวน 20 ข้อ สำหรับขนาดของข้อสอบที่ถูกลอกร้อยละ 30

ทำการสุ่มข้อสอบจำนวน 26 ข้อ สำหรับขนาดของข้อสอบที่ถูกลอกร้อยละ 40

ในส่วนของเงื่อนไขวิธีการลอกข้อสอบเฉพาะข้อยาก ผู้วิจัยได้ใช้เกณฑ์ในการพิจารณาเพื่อระบุข้อสอบระดับยากโดยมีพื้นฐานอยู่บนค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories) โดยข้อสอบที่จะถูกระบุว่าเป็นข้อสอบระดับยาก ต้องมีค่าพารามิเตอร์ ตามเกณฑ์ต่อไปนี้

1. ค่าพารามิเตอร์ความยากของข้อสอบ (b) มีค่ามากกว่า -0.90
2. ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a) มีค่ามากกว่า +0.20
3. ค่าพารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก (c) ไม่เกิน +0.65

โดยเหตุผลที่ผู้วิจัยระบุเกณฑ์ในการพิจารณาเพื่อระบุข้อสอบระดับยากให้เป็นดังข้างต้นนั้นจะอยู่ในส่วนของบทที่ 4 ผลการวิเคราะห์ข้อมูล

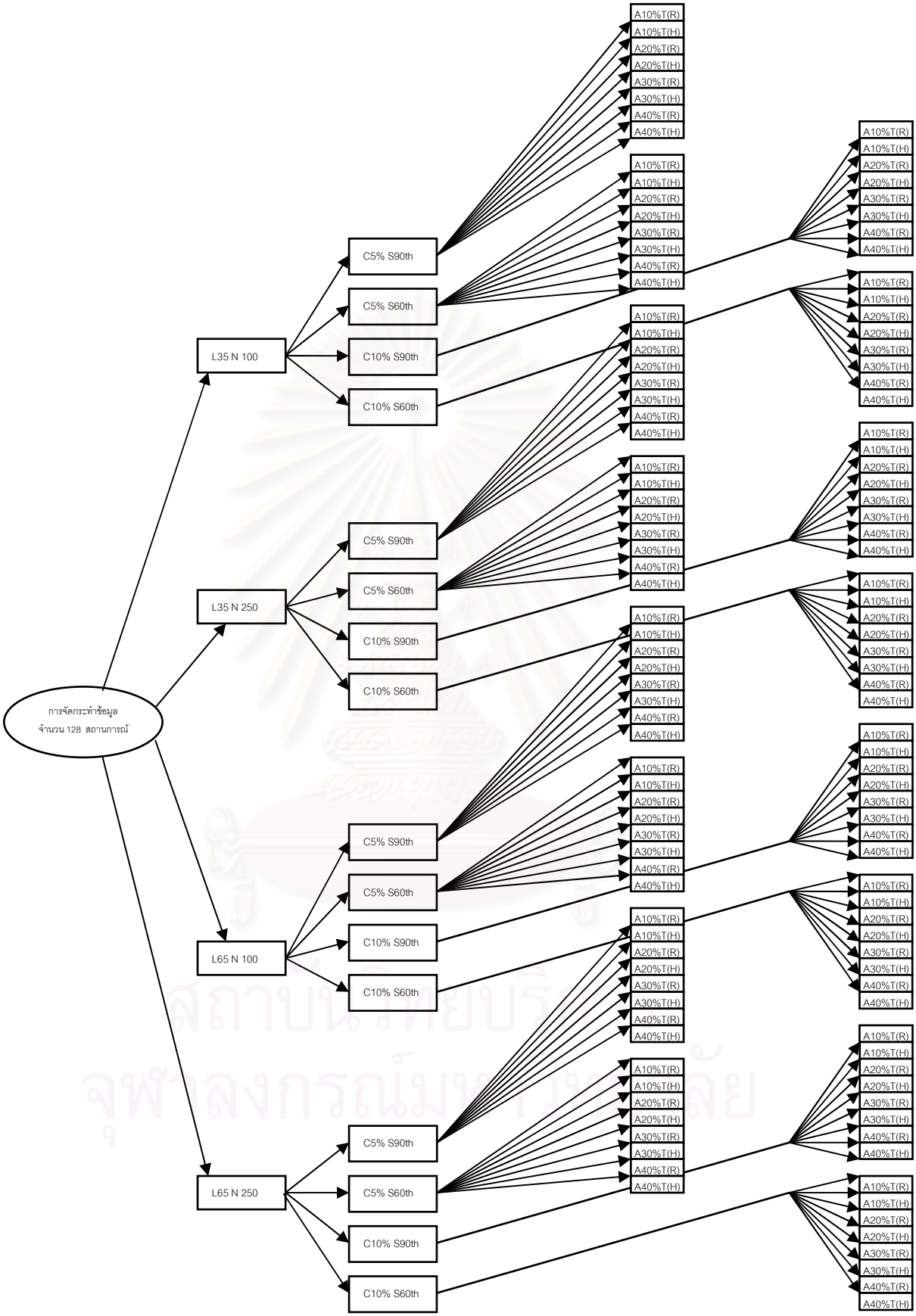
จากนั้นใช้วิธีการสุ่มอย่างง่ายในการสุ่มข้อสอบในกลุ่มข้อสอบระดับยากให้ได้จำนวนเท่ากับจำนวนร้อยละของจำนวนข้อสอบที่ถูกลอกทั้ง 4 ระดับ

สถานการณ์ที่ต้องจัดกระทำข้อมูลทั้ง 128 สถานการณ์ถูกแสดงเป็นภาพ 5 ในหน้า 70 โดยผู้วิจัยใช้สัญลักษณ์ย่อภาษาอังกฤษแทนตัวแปรต้นที่ผู้วิจัยศึกษาในการนำเสนอภาพ ซึ่งสามารถศึกษาของสัญลักษณ์ย่อภาษาอังกฤษได้จากตาราง 4

ตาราง 4 สัญลักษณ์ย่อภาษาอังกฤษแทนตัวแปรต้นที่ศึกษาในงานวิจัย

ตัวแปรต้น	สัญลักษณ์ย่อ	ระดับย่อของตัวแปรต้น	สัญลักษณ์ย่อ
1. ความยาวของแบบสอบ (test length)	L	• แบบสอบความยาว 35 ข้อ	L35
		• แบบสอบความยาว 65 ข้อ	L65
2. จำนวนผู้สอบ (the number of simulees)	N	• จำนวนผู้สอบขนาด 100 คน	N100
		• จำนวนผู้สอบขนาด 250 คน	N250
3. ระดับความสามารถของผู้ให้ลอก (the ability levels of source)	S	• ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90	S90th
		• ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60	S60th
4. ร้อยละของจำนวนผู้ลอก (the percentage of copiers)	C	• ร้อยละ 5 ของจำนวนผู้สอบทั้งหมด	C5%
		• ร้อยละ10 ของจำนวนผู้สอบทั้งหมด	C10%
5. ร้อยละของจำนวนข้อสอบที่ถูกลอก (the percentage of answer copied)	A	• ร้อยละ 10 ของจำนวนข้อสอบทั้งหมด	A10%
		• ร้อยละ 20 ของจำนวนข้อสอบทั้งหมด	A20%
		• ร้อยละ 30 ของจำนวนข้อสอบทั้งหมด	A30%
		• ร้อยละ 40 ของจำนวนข้อสอบทั้งหมด	A40%
6. วิธีการลอก (type of copying)	T	• การลอกข้อสอบแบบสุ่ม	T(R)
		• การลอกข้อสอบเฉพาะข้อยาก	T(H)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ภาพ 5 สถานการณ์ที่ได้จากการจำลองเหตุการณ์ (simulation) 128 สถานการณ์

วิธีการจัดกระทำข้อมูลด้วยวิธีการจำลองสถานการณ์ (Simulation)

ผู้วิจัยจะจัดกระทำกับข้อมูลด้วยวิธีการจำลองสถานการณ์ (Simulation) ซึ่งเป็นการเปลี่ยนคำตอบของผู้ลอกให้ตรงกับคำตอบของผู้ให้ลอก โดยเปลี่ยนคำตอบให้สอดคล้องกับเงื่อนไขที่ศึกษา คือ ร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอกข้อสอบ

การวิเคราะห์ซ้ำ (Replication)

การวิเคราะห์หาค่าดัชนีตรวจจับการลอกข้อสอบ ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω นั้นผู้วิจัยจะทำการวิเคราะห์ข้อมูลของแต่ละสถานการณ์จาก 128 สถานการณ์ซ้ำจำนวน 100 รอบ เพื่อเป็นการตรวจสอบความเที่ยงและถูกต้องของผลการวิเคราะห์ ซึ่งการวิเคราะห์ข้อมูลของแต่ละสถานการณ์ซ้ำนี้ยังส่งผลถึงความถูกต้องของตัวแปรตามที่ผู้วิจัยศึกษา นั่นคือ ความคลาดเคลื่อนประเภทที่ 1 (type I error) และ อำนาจการตรวจจับการลอกข้อสอบการลอกข้อสอบ (detection rate) อีกด้วย

ซึ่งการวิเคราะห์ข้อมูลในแต่ละสถานการณ์ซ้ำนั้นไม่สามารถใช้ข้อมูลชุดเดิมของสถานการณ์นั้นๆ ซ้ำในการวิเคราะห์ได้เนื่องจากหากใช้ข้อมูลชุดเดิมผลการวิเคราะห์จากโปรแกรมคอมพิวเตอร์ย่อมมีค่าเท่าเดิมทุกครั้ง ซึ่งผู้วิจัยไม่ต้องการทดสอบความเที่ยงตรงของการประมวลผลของโปรแกรมคอมพิวเตอร์แต่ต้องการทดสอบความเที่ยงตรงของผลที่ได้จากการวิเคราะห์ว่าเป็นผลการวิเคราะห์ของสถานการณ์นั้นๆ จริง ดังนั้น ในแต่ละสถานการณ์ที่ถูกศึกษาผู้วิจัยจำเป็นจะต้องสร้างชุดข้อมูลซึ่งมีลักษณะการจัดกระทำข้อมูลตรงตามลักษณะสถานการณ์นั้นๆ จำนวน 100 ชุด เพื่อใช้ในการวิเคราะห์ซ้ำ 100 รอบ

โดยชุดข้อมูล 100 ชุด ของสถานการณ์ที่ศึกษาหนึ่งๆ นั้นผู้วิจัยกำหนดให้แต่ละชุดข้อมูลมีความแตกต่างกันในด้านต่อไปนี้

1. ผู้ลอก ซึ่งข้อมูลแต่ละชุดมีผู้สอบที่ถูกกำหนดให้เป็นผู้ลอกแตกต่างกันไป
2. ข้อสอบ ซึ่งข้อมูลแต่ละชุด ข้อสอบที่ถูกกำหนดให้เป็นข้อสอบที่ถูกลอกจะแตกต่างกันไป

เนื่องจากงานวิจัยชิ้นนี้มีสถานการณ์ที่ต้องจัดกระทำข้อมูลรวมทั้งสิ้น 128 สถานการณ์ ซึ่งในแต่ละสถานการณ์ต้องสร้างชุดข้อมูล 100 ชุดเพื่อวิเคราะห์ซ้ำ ดังนั้น ผู้วิจัยต้องสร้างชุดข้อมูลที่ผ่านการจัดกระทำด้วยการจำลองสถานการณ์ (simulation) แล้วทั้งสิ้น 12,800 ชุด

จากนั้นนำข้อมูลซึ่งเป็นข้อมูลที่ผ่านการจัดกระทำแล้ว หรือ ผ่านการเปลี่ยนคำตอบของผู้ที่ถูกกำหนดให้เป็นผู้ลอกให้ตรงกับคำตอบของผู้ที่ถูกกำหนดให้เป็นผู้ให้ลอก ซึ่งจะมีสถานการณ์ทั้งสิ้น 128 สถานการณ์ และ มีจำนวนชุดข้อมูลทั้งสิ้น 12,800 ชุด ไปวิเคราะห์ด้วยโปรแกรม

คอมพิวเตอร์ S-Plus 2000 ซึ่งผู้วิจัยเป็นผู้เขียนคำสั่งที่ใช้ในการวิเคราะห์ประมวลผลสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัว

ระดับนัยสำคัญ (Alpha Level: α)

ระดับนัยสำคัญ (alpha level: α) ซึ่งใช้เป็นเกณฑ์ในการเทียบค่าดัชนีตรวจจับการลอกข้อสอบเพื่อระบุการลอก โดยผู้วิจัยใช้ระดับนัยสำคัญจำนวน 7 ค่า ดังนี้ .0001, .0005, .001, .0025, .005, .01 และ .05 โดยเพิ่มเติมระดับนัยสำคัญขึ้นสองค่าจากงานวิจัยของ Sotaridona, L.S. และ Meijer, R.R. (2003) ซึ่งใช้ระดับนัยสำคัญในการศึกษา 5 ค่า คือ .0001, .0005, .001, .005 และ .01

การเขียนคำสั่งเพื่อทำการวิเคราะห์ดัชนีตรวจจับการลอกข้อสอบ

โปรแกรมที่ใช้วิเคราะห์หาดัชนีตรวจจับการลอกข้อสอบ ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และดัชนี ω ในงานวิจัยชิ้นนี้ คือ โปรแกรม S-Plus 2000 ซึ่งบริษัท StatSci ซึ่งเป็นบริษัทในเครือของบริษัท MathSoft เป็นผู้ผลิตและจัดจำหน่าย โดยโปรแกรม S-Plus 2000 มีลักษณะคล้ายคลึงกับโปรแกรมคอมพิวเตอร์ที่ใช้ในการวิเคราะห์ผลการวิจัยอื่นๆ ที่จะมีฟังก์ชันเพื่อไว้ใช้ในการทดสอบสมมติฐานทางสถิติมากมาย เช่น SPSS, MATLAB เป็นต้น แต่ส่วนที่โปรแกรม S-Plus 2000 มีความแตกต่างจากโปรแกรมอื่นๆ คือ S-Plus 2000 เป็นโปรแกรมผู้ใช้สามารถที่จะสร้างฟังก์ชันทางสถิติซึ่งอาจจะไม่ได้ถูกบรรจุอยู่ในตัวโปรแกรมขึ้นมาเพื่อให้โปรแกรมทำการวิเคราะห์หรือคำนวณได้ ซึ่งการสร้างฟังก์ชันทางสถิติขึ้นนั้นผู้ใช้ต้องมีความรู้ด้าน “ภาษา S” (S language) ซึ่งเป็นภาษาที่ใช้ในการเขียนคำสั่งในโปรแกรม S-Plus 2000

ดัชนีตรวจจับการลอกข้อสอบ ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ที่ผู้วิจัยศึกษาในงานวิจัยชิ้นนี้นั้นก็ไม่ได้เป็นฟังก์ชันทางสถิติที่ถูกบรรจุไว้ในโปรแกรมคอมพิวเตอร์โปรแกรมต่างๆ รวมทั้งโปรแกรม S-Plus 2000 ด้วย ดังนั้นผู้วิจัยจึงต้องสร้างฟังก์ชันทางสถิติเพื่อใช้ในการวิเคราะห์ค่าดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ขึ้นมาด้วยการเขียนคำสั่งด้วยภาษา S สำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัว ซึ่งมีขั้นตอนกระบวนการต่อไปนี้

1. ศึกษาภาษา S เพื่อให้มีความรู้ความเข้าใจที่จะสามารถเขียนคำสั่งสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัวได้
2. ศึกษาเนื้อหาและทฤษฎีของค่าดัชนีตรวจจับการลอกข้อสอบแต่ละตัวอย่างละเอียด เพื่อที่จะนำความรู้และความเข้าใจในดัชนีแต่ละตัวไปเขียนเป็นคำสั่งภาษา S เพื่อใช้ในการวิเคราะห์ข้อมูล

3. ศึกษางานวิจัย เรื่อง การตรวจสอบการโกงโดยการใช้ดัชนีตรวจจับการลอกข้อสอบ S_2 (Cheating Detection Using the S_2 Copying Index)(Leonardo S. Sotaridona, 2003) ซึ่งมีส่วนของคำสั่งที่ใช้ในการวิเคราะห์หาค่าดัชนี S_2 อยู่ในงานวิจัย อย่างละเอียด แล้วนำคำสั่งที่อยู่ในงานวิจัยเล่มดังกล่าวไปปรับปรุงและใช้เป็นต้นแบบในการเขียนคำสั่งเพื่อทำการวิเคราะห์หาค่าดัชนีตรวจจับการลอกข้อสอบตัวอื่นๆ ที่เหลือ
4. ทำการติดต่อกับผู้คิดค้นค่าดัชนีตรวจจับการลอกข้อสอบแต่ละค่า คือ James A. Wollack ซึ่งเป็นผู้คิดค้นดัชนี ω และ Leonardo S. Sotaridona ซึ่งเป็นผู้คิดค้นค่าดัชนี K_2, ดัชนี S_1 และ ดัชนี S_2 ทางจดหมายอิเล็กทรอนิค (e-mail) เพื่อขอแนวคิดและคำแนะนำในการเขียนคำสั่ง
5. ปรับปรุงคำสั่งที่ใช้ในการวิเคราะห์หาค่าดัชนี S_2 ให้เหมาะสมกับปัจจัยที่ผู้วิจัยศึกษา และเขียนคำสั่งที่ใช้ในการวิเคราะห์เป็นภาษา S สำหรับดัชนี K_2, ดัชนี S_1 และ ดัชนี ω
6. ส่งคำสั่งที่ได้ปรับปรุงและเขียนขึ้นใหม่ดังที่กล่าวในข้อ 5 ไปให้ผู้คิดค้นดัชนีนั้นๆ ตรวจสอบความถูกต้อง กล่าวคือ ผู้วิจัยส่งคำสั่งที่ใช้ในการวิเคราะห์เป็นภาษา S ของดัชนี ω ให้แก่ James A. Wollack ทำการตรวจสอบ และ ผู้วิจัยส่งคำสั่งที่ใช้ในการวิเคราะห์เป็นภาษา S ของดัชนี K_2, ดัชนี S_1 และ ดัชนี S_2 ให้แก่ Leonardo S. Sotaridona ทำการตรวจสอบ
7. แก้ไขคำสั่งของค่าดัชนีตรวจจับการลอกข้อสอบแต่ละตัวตามคำแนะนำของผู้คิดค้นค่าดัชนีนั้นๆ จนมีความถูกต้องสมบูรณ์
8. ทดลองคำสั่งของค่าดัชนีตรวจจับการลอกข้อสอบแต่ละตัวที่ได้ปรับแก้แล้วกับข้อมูลที่ผู้วิจัยได้จัดกระทำขึ้นเพื่อตรวจสอบว่าคำสั่งทุกคำสั่งสามารถที่จะใช้วิเคราะห์หาค่าดัชนีตรวจจับการลอกข้อสอบได้จริง
9. นำคำสั่งที่ใช้ในการวิเคราะห์หาค่าดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ไปใช้กับข้อมูลที่ผ่านการจัดกระทำด้วยการจำลองสถานการณ์ (Simulation) ทั้ง 128 สถานการณ์ 12,800 ชุดข้อมูล โดยคำสั่งที่ผู้วิจัยได้เขียนขึ้นนั้นเป็นคำสั่งที่ให้โปรแกรมทำการวิเคราะห์จนได้ค่าตัวแปรตามที่ผู้วิจัยศึกษา คือ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ ที่ทุกระดับนัยสำคัญที่ผู้วิจัยกำหนดเป็นเกณฑ์เพื่อระบุการลอก

การวิเคราะห์ข้อมูล

โปรแกรมที่ใช้ในการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในงานวิจัยชิ้นนี้ใช้โปรแกรมคอมพิวเตอร์ ดังนี้ คือ

1. TAP (Test Analysis Program)

โดยโปรแกรม TAP (Brooks, G.P., 2003) ใช้ในการตรวจให้คะแนนแก่ผู้สอบแต่ละคน รวมทั้งวิเคราะห์ข้อมูลเพื่อหาค่า คะแนนต่ำสุด คะแนนสูงสุด ค่าเฉลี่ย ค่ามัธยฐาน ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน ความเบ้ ความโด่ง แล้วนำผลการวิเคราะห์มาศึกษาลักษณะโดยภาพรวมของกลุ่มผู้สอบแต่ละกลุ่มซึ่งเป็นกลุ่มที่ได้จากการจัดกระทำในปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ ดังนี้

1. กลุ่มแบบสอบขนาด 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100)
2. กลุ่มแบบสอบขนาด 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250)
3. กลุ่มแบบสอบขนาด 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100)
4. กลุ่มแบบสอบขนาด 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250)

นอกจากนี้โปรแกรม TAP จะใช้ในการวิเคราะห์หาค่าความยากง่ายและอำนาจจำแนกของข้อสอบ รวมทั้ง ความเที่ยงแบบสอดคล้องภายใน ของกลุ่มผู้สอบแต่ละกลุ่มซึ่งเป็นกลุ่มที่ได้จากการจัดกระทำในปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบดังที่กล่าวมาแล้วข้างต้น ตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) แล้วนำผลการวิเคราะห์มาศึกษาลักษณะโดยภาพรวมของชุดข้อสอบ

2. MULTILOG

โปรแกรม MULTILOG (Thissen, 1991) ถูกใช้ในการวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ และ พารามิเตอร์ระดับความสามารถของผู้สอบ (θ) ตามทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories) โดยใช้โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (Three-Parameter Model) เพื่อนำผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบไปใช้ในการคัดเลือกข้อสอบที่อยู่ในกลุ่มข้อสอบระดับยาก และ นำค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) มาใช้ในการเรียงลำดับผู้สอบตามระดับความสามารถจากมากไปน้อย ในขั้นตอนการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (Simulation)

นอกจากนี้โปรแกรม MULTILOG ยังใช้ในการวิเคราะห์หาค่าพารามิเตอร์ความชันของแต่ละรายการคำตอบ (Slop of the trace lines) และ ค่าพารามิเตอร์จุดตัดของแต่ละรายการ

(Intercept parameter) ตามทฤษฎีการทดสอบแนวใหม่แบบโมเดล Nominal Response Model เพื่อนำค่าพารามิเตอร์ที่ได้ไปวิเคราะห์หาดัชนี ω

3. S-Plus 2000

โปรแกรม S-Plus 2000 (S-Plus 2000, MathSoft, Inc.) ถูกใช้ในขั้นตอนวิเคราะห์หาค่าดัชนีตรวจจับการลอกข้อสอบ ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω และนำค่าดัชนีตรวจจับการลอกข้อสอบไปวิเคราะห์หาค่า ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ ที่ทุกระดับนัยสำคัญที่ผู้วิจัยกำหนดเป็นเกณฑ์เพื่อระบุการลอก

โดยผู้วิจัยต้องนำชุดข้อมูลที่ได้จากขั้นตอนการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (Simulation) ซึ่งเป็นข้อมูลของสถานการณ์ที่ศึกษา 128 สถานการณ์ มีชุดข้อมูลทั้งสิ้น 12,800 ชุด กล่าวคือ หนึ่งสถานการณ์จะมีชุดข้อมูล 100 ชุดเพื่อการวิเคราะห์ซ้ำ โดยผลการวิเคราะห์ด้วยโปรแกรม S-Plus 2000 นั้นผู้วิจัยผู้เขียนคำสั่งที่ใช้ในการวิเคราะห์ประมวลผลสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัวโดยให้มีผลการวิเคราะห์ออกมาเป็น ค่าของตัวแปรตามที่ผู้วิจัยศึกษา คือ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ ที่ทุกระดับนัยสำคัญที่ผู้วิจัยกำหนดเป็นเกณฑ์เพื่อระบุการลอก

การระบุการลอกของดัชนีตรวจจับการลอกข้อสอบ

สำหรับดัชนี K_2, ดัชนี S_1 และ ดัชนี S_2 นั้น ผู้สอบจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอกก็ต่อเมื่อค่าดัชนีที่คำนวณได้มีค่า *น้อยกว่าหรือเท่ากับ* ค่าระดับนัยสำคัญที่ผู้วิจัยกำหนดเป็นเกณฑ์ในการเทียบค่าดัชนีตรวจจับการลอกข้อสอบเพื่อระบุการลอก โดยผู้วิจัยใช้ระดับนัยสำคัญจำนวน 7 ค่า ดังนี้ .0001, .0005, .001, .0025, .005, .01 และ .05

สำหรับดัชนี ω ผู้สอบจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอกก็ต่อเมื่อดัชนี ω มีค่า *มากกว่า* ค่าวิกฤติของการทดสอบสมมติฐานทางเดียว (one-tailed critical value) ซึ่งขอบเขตวิกฤติมีพื้นที่อยู่ทางด้านขวาของโค้งปกติ โดยผู้วิจัยแสดงค่าวิกฤติของการทดสอบสมมติฐานทางเดียว ณ ระดับนัยสำคัญต่างๆ ที่ผู้วิจัยศึกษา ซึ่งใช้ในการเทียบดัชนี ω เพื่อระบุการลอก ในตาราง 5

ตาราง 5 ค่าวิกฤติของการทดสอบสมมติฐานทางเดียว ณ ระดับนัยสำคัญที่ศึกษา

ระดับนัยสำคัญ	ค่าวิกฤติของการทดสอบสมมติฐานทางเดียว
.0001	3.719
.0005	3.291
.001	3.09
.0025	2.807
.005	2.576
.01	2.326
.05	1.645

คุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ

คุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ตัวนั้นประกอบด้วย ความคลาดเคลื่อนประเภทที่ 1 (Type I error) และ อำนาจการตรวจจับการลอกข้อสอบ (Detection Rates) ซึ่งเป็นตัวแปรตามที่ผู้วิจัยศึกษา โดยคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบทั้งสองมีหลักการพื้นฐานดังนี้

1. ความคลาดเคลื่อนประเภทที่ 1 (Type I error)

หลักการของการคำนวณค่าความคลาดเคลื่อนประเภทที่ 1 นั้น คือ การคำนวณจำนวนครั้งหรือสัดส่วนของผู้ที่ไม่ใช่ผู้ลอกอย่างแท้จริง (true noncopier) ถูกระบุผิดว่าเป็นผู้ลอก (copier)

ค่าความคลาดเคลื่อนประเภทที่ 1 สนใจเพียงแต่ผู้สอบที่ไม่ได้เป็นผู้ลอกเท่านั้น ดังนั้น สถานการณ์ที่ศึกษาจำนวน 128 สถานการณ์ จะมีจำนวนของผู้สอบที่จะนำมาคำนวณหาค่าความคลาดเคลื่อนประเภทที่ 1 แตกต่างกันไป แสดงเป็นตารางได้ดังนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 6 จำนวนผู้สอบที่ถูกใช้คำนวณหาค่าความคลาดเคลื่อนประเภทที่ 1 แยกตามตัวแปรจำนวนผู้สอบ และ ร้อยละของจำนวนผู้ลอก

สถานการณ์		จำนวนผู้สอบที่จะนำมาคำนวณหาค่าความคลาดเคลื่อนประเภทที่ 1
จำนวนผู้สอบ	ร้อยละของจำนวนผู้ลอก	
100 คน	5%	(ผู้ไม่ได้ลอก 94 คนต่อ 1 ชุดข้อมูล x 100 รอบของการวิเคราะห์ซ้ำ) = 9,400 คน ต่อ 1 สถานการณ์
	10%	(ผู้ไม่ได้ลอก 89 คนต่อ 1 ชุดข้อมูล x 100 รอบของการวิเคราะห์ซ้ำ) = 8,900 คน ต่อ 1 สถานการณ์
250 คน	5%	(ผู้ไม่ได้ลอก 236 คนต่อ 1 ชุดข้อมูล x 100 รอบของการวิเคราะห์ซ้ำ) = 23,600 คน ต่อ 1 สถานการณ์
	10%	(ผู้ไม่ได้ลอก 224 คนต่อ 1 ชุดข้อมูล x 100 รอบของการวิเคราะห์ซ้ำ) = 22,400 คน ต่อ 1 สถานการณ์

ตัวอย่างของการนำไปใช้ เช่น หากสถานการณ์ที่ศึกษา คือ สถานการณ์ที่ ความยาวของแบบสอบเป็น 35 ข้อ ผู้สอบจำนวน 100 คน ผู้ลอกร้อยละ 5 ผู้ให้ลอกอยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 โดยผู้ลอกลอกข้อสอบร้อยละ 30 จากข้อสอบทั้งหมด โดยลอกข้อสอบแบบสุ่ม (L35N100C5%S90thA30%T(R)) โดยทำการวิเคราะห์ด้วยดัชนี K_2 จากสถานการณ์จะเห็นว่าในข้อมูลหนึ่งชุดจะมีผู้ซึ่งไม่ใช่ผู้ลอกจำนวน 94 คน (เนื่องจากไม่รวมผู้ลอกร้อยละ 5 ซึ่งคือ 5 คน และผู้ให้ลอก 1 คน) ซึ่งพอทำการวิเคราะห์ซ้ำ 100 รอบ จะได้ค่าดัชนี K_2 จำนวน ทั้งสิ้น 9,400 ค่า และสมมติหากเรานำค่าดัชนี K_2 ทั้งสิ้น 9,400 ค่า ไปเทียบกับระดับนัยสำคัญ 0.05 เพื่อระบุการลอกปรากฏว่ามีผู้ที่ถูกตรวจจับได้ว่าลอกข้อสอบทั้งสิ้น 8 คน (ซึ่งทั้ง 8 คนไม่ได้ทำการลอกข้อสอบ) ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญ 0.05 สามารถหาได้โดย นำ 8 ไปหารด้วย 9,400 ซึ่งได้ค่าประมาณ 0.00085

2. อัตราตรวจจับการลอกข้อสอบ (Detection Rates)

หลักการคำนวณค่าอำนาจการตรวจจับการลอกข้อสอบ คือ การคำนวณสัดส่วนของผู้ลอกจริง (true copier) จะถูกระบุว่าเป็นผู้ลอกหรือถูกตรวจจับได้โดยค่าดัชนีตรวจจับการลอกข้อสอบ ซึ่งประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω

อัตราตรวจจับการลอกข้อสอบสนใจเพียงแต่ผู้สอบที่เป็นผู้ลอกจริงเท่านั้น ดังนั้น สถานการณ์ที่ศึกษาจำนวน 128 สถานการณ์ จะมีจำนวนของผู้สอบที่จะนำมาคำนวณหาค่าอำนาจตรวจจับการลอกข้อสอบแตกต่างกันไป แสดงเป็นตารางได้ดังนี้

ตาราง 7 จำนวนผู้สอบที่ถูกใช้คำนวณหาค่าอำนาจตรวจจับการลอกข้อสอบ แยกตามตัวแปรจำนวนผู้สอบ และ ร้อยละของจำนวนผู้ลอก

สถานการณ์		จำนวนผู้สอบที่จะนำมาคำนวณหาค่าอำนาจตรวจจับการลอกข้อสอบ
จำนวนผู้สอบ	ร้อยละของจำนวนผู้ลอก	
100 คน	5%	(ผู้ลอกจริง 5 คนต่อ 1 ชุดข้อมูล x 100 รอบของการวิเคราะห์ซ้ำ) = 500 คน ต่อ 1 สถานการณ์
	10%	(ผู้ลอกจริง 10 คนต่อ 1 ชุดข้อมูล x 100 รอบของการวิเคราะห์ซ้ำ) = 1,000 คน ต่อ 1 สถานการณ์
250 คน	5%	(ผู้ลอกจริง 13 คนต่อ 1 ชุดข้อมูล x 100 รอบของการวิเคราะห์ซ้ำ) = 1,300 คน ต่อ 1 สถานการณ์
	10%	(ผู้ลอกจริง 25 คนต่อ 1 ชุดข้อมูล x 100 รอบของการวิเคราะห์ซ้ำ) = 2,500 คน ต่อ 1 สถานการณ์

ตัวอย่างของการนำไปใช้ เช่น หากสถานการณ์ที่ศึกษา คือ สถานการณ์ที่ ความยาวของแบบสอบเป็น 35 ข้อ ผู้สอบจำนวน 100 คน ผู้ลอกร้อยละ 5 ผู้ให้ลอกอยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 โดยผู้ลอกลอกข้อสอบร้อยละ 30 จากข้อสอบทั้งหมด โดยลอกข้อสอบแบบสุ่ม (L35N100C5%S90thA30%T(R)) โดยทำการวิเคราะห์ด้วยดัชนี K_2 จากสถานการณ์จะเห็นว่าในข้อมูลหนึ่งชุดจะมีผู้ซึ่งเป็นผู้ลอกจริงจำนวน 5 คน (เนื่องจากไม่นับผู้ที่ไม่ได้ทำการลอกจำนวน 94 คน และ ผู้ให้ลอก 1 คน) ซึ่งทำการวิเคราะห์ซ้ำ 100 รอบ จะได้ค่าดัชนี K_2 จำนวนทั้งสิ้น 500 ค่า และสมมติหากเรานำค่าดัชนี K_2 ทั้งสิ้น 500 ค่า ไปเทียบกับระดับนัยสำคัญ 0.05 เพื่อระบุการลอกปรากฏว่ามีผู้ที่ถูกตรวจจับได้ว่าลอกข้อสอบทั้งสิ้น 423 คน (ผู้สอบทั้ง 500 คนทำการลอกข้อสอบจริง) อำนาจตรวจจับการลอกข้อสอบ ณ ระดับนัยสำคัญ 0.05 สามารถหาได้โดย นำ 423 ไปหารด้วย 500 ซึ่งได้ค่าประมาณ 0.846

การศึกษาคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ

เพื่อให้การศึกษาคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบนั้นง่ายต่อการทำความเข้าใจมากขึ้น และ สามารถที่จะทำการเปรียบเทียบคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบได้อย่างชัดเจน ผู้วิจัยจึงนำคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ ซึ่งก็คือ ค่าความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ มาวาดเป็นกราฟเส้นก่อนที่จะทำการศึกษาค่าความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ โดยมีรายละเอียดดังนี้

1. **การศึกษาค่าความคลาดเคลื่อนประเภทที่ 1**

ค่าความคลาดเคลื่อนประเภทที่ 1 มีขั้นตอนในการวาดกราฟเส้นดังนี้

1.1 นำค่าความคลาดเคลื่อนประเภทที่ 1 ณ ทุกระดับนัยสำคัญที่ผู้วิจัยศึกษาของทุกค่าดัชนีตรวจจับการลอกข้อสอบ ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม S-Plus 2000 ไปทำการวาดกราฟเส้นด้วยโปรแกรม Microsoft Excel ซึ่งกราฟเส้นที่ได้จากการวาดนั้นแกนในแนวนอนจะแสดงถึงค่าระดับนัยสำคัญ ส่วนแกนในแนวแกนตั้งจะแสดงถึงค่าความคลาดเคลื่อนประเภทที่ 1

1.2 ข้อมูลที่ใส่เข้าไปในโปรแกรม Microsoft Excel เพื่อทำการวาดกราฟเส้นนั้น นอกจากจะใส่ข้อมูลของค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญที่ผู้วิจัยศึกษาของทุกค่าดัชนีตรวจจับการลอกข้อสอบแล้ว ผู้วิจัยจะใส่ข้อมูลขอบเขตซึ่งเป็นข้อมูลที่ค่าความคลาดเคลื่อนประเภทที่ 1 จะมีค่าเท่ากับระดับนัยสำคัญต่างๆ เพื่อทำการวาดกราฟที่มีชื่อว่า เส้นกราฟขอบเขต (Boundary line) หรือ เส้นกราฟเอกลักษณ์ (Identity line) ซึ่งเป็นเส้นกราฟแสดงขอบเขตค่าความคลาดเคลื่อนที่ยอมรับได้

1.3 กราฟเส้นที่ได้จากการวาดด้วยโปรแกรม Microsoft Excel นั้นจะมี 5 เส้น โดยกราฟ 4 เส้นนั้นจะเป็นกราฟค่าความคลาดเคลื่อนประเภทที่ 1 ของค่าดัชนีตรวจจับการลอกข้อสอบ ซึ่งประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ส่วนกราฟอีก 1 เส้นนั้นจะเป็นเส้นกราฟขอบเขต (Boundary line) หรือ เส้นกราฟเอกลักษณ์ (Identity line) ซึ่งเป็นเส้นกราฟแสดงขอบเขตค่าความคลาดเคลื่อนประเภทที่ 1 ที่ยอมรับได้

โดยหากค่าความคลาดเคลื่อนประเภทที่ 1 ซึ่งได้จากการวิเคราะห์มีค่าเหนือจากเส้นกราฟขอบเขต จะถือว่าความคลาดเคลื่อนประเภทที่ 1 ที่วิเคราะห์ได้นั้น liberal แต่หากค่าความคลาดเคลื่อนประเภทที่ 1 ซึ่งได้จากการวิเคราะห์มีค่าต่ำกว่าหรืออยู่ต่ำกว่าเส้นกราฟขอบเขต จะถือว่าความคลาดเคลื่อนประเภทที่ 1 ที่วิเคราะห์ได้นั้น conservative

ค่าดัชนีตรวจจับการลอกข้อสอบที่ดี จะต้องสามารถรักษาระดับความคลาดเคลื่อนประเภทที่ 1 ให้มีค่าอยู่ใต้หรือต่ำกว่าเส้นกราฟขอบเขตได้สำหรับทุกระดับนัยสำคัญที่ทำการทดสอบ แต่ไม่ควรจะมีค่าอยู่ใต้เส้นกราฟขอบเขตมาก เนื่องจาก หากค่าความคลาดเคลื่อนประเภทที่ 1 มีค่าต่ำกว่าเส้นกราฟขอบเขตมากๆ จะส่งผลทำให้อำนาจการตรวจจับการลอกข้อสอบมีค่าต่ำ

2. การศึกษาค่าอำนาจการตรวจจับการลอกข้อสอบ

อำนาจการตรวจจับการลอกข้อสอบ มีขั้นตอนในการวาดกราฟเส้นดังนี้

2.1 นำค่าอำนาจการตรวจจับการลอกข้อสอบ ณ ทุกระดับนัยสำคัญของทุกค่าดัชนีตรวจจับการลอกข้อสอบ ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม S-Plus 2000 ไปทำการวาดกราฟเส้นด้วยโปรแกรม Microsoft Excel ซึ่งกราฟเส้นที่ได้จากการวาดนั้นแกนในแนวนอนจะแสดงถึงค่าระดับนัยสำคัญ ส่วนแกนในแนวตั้งจะแสดงถึงค่าอำนาจการตรวจจับการลอกข้อสอบ

2.2 กราฟเส้นที่ได้จากการวาดด้วยโปรแกรม Microsoft Excel นั้นจะมี 4 เส้น โดยแต่ละเส้นจะแสดงถึงค่าอำนาจการตรวจจับการลอกข้อสอบของค่าดัชนีตรวจจับการลอกข้อสอบทั้งสี่ ซึ่งประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω

กราฟเส้นที่แสดงค่าอำนาจการตรวจจับการลอกข้อสอบแต่ละเส้นนั้นจะแสดงให้เห็นถึงแนวโน้มของค่าอำนาจการตรวจจับการลอกข้อสอบที่เปลี่ยนไปตามระดับระดับนัยสำคัญของแต่ละค่าดัชนีตรวจจับการลอกข้อสอบ ซึ่งประกอบด้วยดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω

โดยดัชนีตรวจจับการลอกข้อสอบที่ดีควรมีค่าความคลาดเคลื่อนประเภทที่ 1 (Type I error) ที่ต่ำ และ มีค่าอำนาจการตรวจจับการลอกข้อสอบ (Detection rate) ที่สูง

การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ

การเปรียบเทียบคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ สามารถแบ่งออกได้เป็น

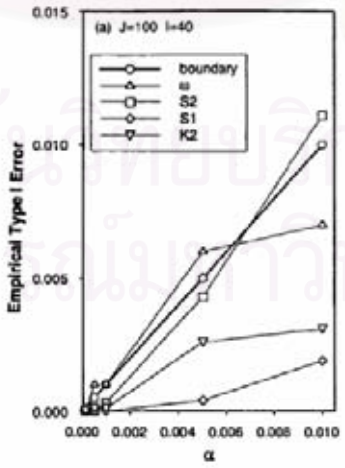
1. การเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1

เนื่องจากค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบที่ได้จากการวิเคราะห์ด้วยโปรแกรม S-Plus 2000 นั้นถูกนำไปวาดเป็นกราฟเส้น ด้วยโปรแกรม Microsoft Excel ดังที่กล่าวไว้แล้วในหัวข้อ การศึกษาคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ

การเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 ระหว่างดัชนีตรวจจับการลอกข้อสอบ มีหลักการ ดังนี้

1) ค่าดัชนีตรวจจับการลอกข้อสอบที่ดีต้องสามารถรักษาระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้มีค่าอยู่ใต้เส้นกราฟขอบเขตได้สำหรับทุกระดับนัยสำคัญที่กำหนด
2) ดัชนีตรวจจับการลอกข้อสอบที่ดีไม่ควรมีค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าเส้นกราฟขอบเขตมากเกินไป

ดังนั้น อาจกล่าวได้ว่าค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบที่ดีต้องมีค่าต่ำกว่าเส้นกราฟขอบเขตแต่มีค่าต่ำกว่าเพียงเล็กน้อย โดยการเปรียบเทียบกราฟค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบแต่ละเส้นนั้นใช้สายตาในการเปรียบเทียบ หากกราฟค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบใดที่สามารถควบคุมระดับให้มีค่าต่ำกว่าเส้นกราฟขอบเขตได้สำหรับทุกระดับนัยสำคัญที่กำหนด และเส้นกราฟอยู่ใกล้เส้นกราฟขอบเขตที่สุด ดัชนีตรวจจับการลอกข้อสอบนั้นเป็นดัชนีตรวจจับการลอกข้อสอบที่สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ดีที่สุดสำหรับสถานการณ์นั้นๆ ส่วนเส้นกราฟของดัชนีตรวจจับการลอกข้อสอบที่อยู่ห่างจากเส้นกราฟขอบเขตถัดลงมาจากเส้นกราฟของดัชนีตรวจจับการลอกข้อสอบที่สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ดีที่สุดนั้นจะสามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ดีรองลงมาตามลำดับ ศึกษาวิธีการเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 ได้จากภาพ 6



ภาพ 6 กราฟค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω เมื่อผู้สอบมีจำนวน 100 คน และ แบบสอบยาว 40 ข้อ

ที่มา: Sotaridona และ Meijer (2003)

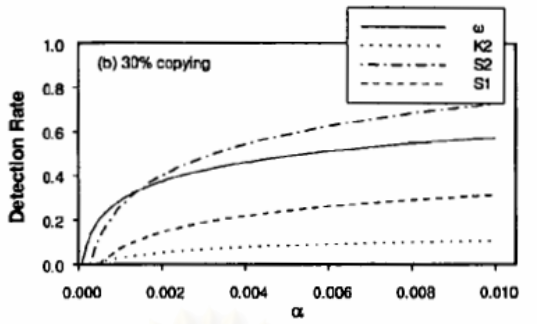
จากกราฟ เส้นกราฟหนาจะเป็นเส้นกราฟขอบเขต (Boundary line) และ เส้นกราฟอีก 4 เส้นที่เหลือจะแสดงถึงค่าความคลาดเคลื่อนประเภทที่ 1 ของค่าดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และดัชนี ω ที่ระดับนัยสำคัญต่างๆ

จากกราฟจะเห็นได้ว่าภายใต้เงื่อนไขที่มีผู้สอบจำนวน 100 คน และ แบบสอบยาว 40 ข้อ ดัชนี S_2 และ ดัชนี ω ไม่สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้มีค่าอยู่ใต้เส้นกราฟขอบเขตได้สำหรับทุกระดับนัยสำคัญที่กำหนด ส่วนดัชนี K_2 และ ดัชนี S_1 สามารถรักษาระดับความคลาดเคลื่อนประเภทที่ 1 ให้มีค่าอยู่ใต้เส้นกราฟขอบเขตได้สำหรับทุกระดับนัยสำคัญที่กำหนด โดยดัชนี K_2 นั้นเป็นค่าดัชนีที่สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ได้ดีกว่า ดัชนี S_1 เนื่องจากเส้นกราฟของดัชนี K_2 อยู่ใกล้เส้นกราฟขอบเขตมากกว่าเส้นกราฟของค่าดัชนี S_1

2. การเปรียบเทียบอำนาจการตรวจจับการลอกข้อสอบ

ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม S-Plus 2000 นั้นถูกนำไปวาดเป็นกราฟเส้น ด้วยโปรแกรม Microsoft Excel ดังที่กล่าวในหัวข้อ การศึกษาคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบการเปรียบเทียบนั้นสามารถทำได้โดยเปรียบเทียบกราฟเส้นของแต่ละค่าดัชนีตรวจจับการลอกข้อสอบโดยใช้สายตา ว่ากราฟเส้นของดัชนีตรวจจับการลอกข้อสอบใดที่มีค่าสูงที่สุด หากกราฟเส้นของดัชนีตรวจจับการลอกข้อสอบใดมีค่าสูงสุด หรือ อยู่เหนือกราฟเส้นของดัชนีตรวจจับการลอกข้อสอบอื่นๆ ดัชนีตรวจจับการลอกข้อสอบนั้นๆ เป็นดัชนีตรวจจับการลอกข้อสอบที่มีอำนาจการตรวจจับการลอกข้อสอบที่ดีที่สุดสำหรับสถานการณ์นั้นๆ ส่วนเส้นกราฟของดัชนีตรวจจับการลอกข้อสอบที่อยู่ถัดลงมาจากเส้นที่อยู่สูงที่สุดนั้นก็จะมีระดับอำนาจการตรวจจับที่ดีรองลงมาตามลำดับ ศึกษาวิธีการเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ ได้จากภาพ 7

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ภาพ 7 กราฟค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω เมื่อแบบสอบยาว 40 ข้อ และ ร้อยละของจำนวนข้อสอบที่ถูกลอกเป็น 30
ที่มา: Sotaridona และ Meijer (2003)

จากกราฟด้านบน แสดงให้เห็นว่า ภายใต้เงื่อนไข ที่แบบสอบยาว 40 ข้อ และ ร้อยละของจำนวนข้อสอบที่ถูกลอกเป็น 30 นั้นค่าดัชนี S_2 มีค่าอำนาจการตรวจจับการลอกข้อสอบที่ดีที่สุด รองลงมาคือ ดัชนี ω, ค่าดัชนี S_1 และ ค่าดัชนี K_2 ตามลำดับ เนื่องจากกราฟค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 อยู่เหนือกราฟอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบอื่นๆ ทุกตัว เส้นกราฟที่อยู่ถัดมาคือ เส้นกราฟค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω, ค่าดัชนี S_1 และ ค่าดัชนี K_2 ตามลำดับ ทำให้ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω, ค่าดัชนี S_1 และ ค่าดัชนี K_2 นั้นดีรองจากดัชนี S_2 ตามลำดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การศึกษาเรื่อง “การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ” นั้นผู้วิจัยเริ่มทำการวิเคราะห์ข้อมูลตั้งแต่ขั้นตอนก่อนการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (simulation) เพื่อศึกษาลักษณะโดยภาพรวมของกลุ่มผู้สอบ และ ชุดข้อสอบที่ถูกแบ่งออกเป็นกลุ่มตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ จากนั้นในขั้นตอนจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ ผู้วิจัยวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ และพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) ตามทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories) โดยใช้โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (Three-Parameter Model) เพื่อนำผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบไปใช้ในการคัดเลือกข้อสอบที่จัดอยู่ในกลุ่มข้อสอบระดับยาก และ นำข้อมูลค่าพารามิเตอร์ระดับความสามารถของผู้สอบไปใช้ในการเลือกผู้ลอกและผู้ให้ลอกในจัดกระทำข้อมูลแต่ละสถานการณ์ที่ผู้วิจัยศึกษา และ สุดท้ายทำการวิเคราะห์หาค่าความคลาดเคลื่อนประเภทที่ 1 และ อำนาจของการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ ณ ระดับนัยสำคัญต่างๆ ที่ผู้วิจัยศึกษา จากนั้นนำผลการวิเคราะห์ไปเขียนกราฟเส้นเพื่อศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ

โดยผู้วิจัยแบ่งการเสนอผลการวิเคราะห์ข้อมูลออกเป็น 2 ตอน คือ

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลในขั้นตอนก่อนและระหว่างการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (Simulation) ซึ่งประกอบด้วยผลการวิเคราะห์ข้อมูล ดังนี้

1. ผลการวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT)
2. ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)
3. ผลการวิเคราะห์ค่าพารามิเตอร์ของผู้สอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)

ตอนที่ 2 ผลการวิเคราะห์คุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ ซึ่งประกอบด้วย

1. ความคลาดเคลื่อนประเภทที่ 1 (Type I Error Rate)
2. อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate)

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลในขั้นตอนก่อนและระหว่างการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์(Simulation)

1. ผลการวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT)

ผู้วิจัยวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม โดยตรวจให้คะแนนแก่ผู้สอบแต่ละคนในแต่ละกลุ่มผู้สอบซึ่งถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ จากนั้นวิเคราะห์ข้อมูลเพื่อหาค่า คะแนนต่ำสุด คะแนนสูงสุด ค่าเฉลี่ย ค่ามัธยฐาน ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน ความเบ้ ความโด่ง เพื่อนำผลการวิเคราะห์ดังกล่าวมาศึกษาลักษณะโดยภาพรวมของกลุ่มผู้สอบแต่ละกลุ่มซึ่งถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ ว่ามีลักษณะเหมือนหรือแตกต่างกันอย่างไร

นอกจากนี้ผู้วิจัยวิเคราะห์หาค่าความยากง่ายและอำนาจจำแนกของข้อสอบ รวมทั้งความเที่ยงแบบสอดคล้องภายใน ตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) ของแต่ละกลุ่มผู้สอบซึ่งถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ เพื่อนำผลการวิเคราะห์ดังกล่าวมาศึกษาลักษณะของชุดข้อสอบแต่ละชุดของกลุ่มผู้สอบแต่ละกลุ่มที่ถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ ว่ามีลักษณะเหมือนหรือแตกต่างกันอย่างไร

ผู้วิจัยนำเสนอผลการวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิมของกลุ่มผู้สอบหรือกลุ่มข้อมูลซึ่งถูกแบ่งออกตามปัจจัย ความยาวของแบบสอบ และ จำนวนผู้สอบ ซึ่งได้แก่

1. กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100)
2. กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250)
3. กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100)
4. กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250)

ออกเป็น 2 ส่วน คือ

1.1 ผลการวิเคราะห์ คะแนนต่ำสุด คะแนนสูงสุด ค่าเฉลี่ย ค่ามัธยฐาน ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน ความเบ้ ความโด่ง

1.2 ผลการวิเคราะห์ ค่าความยากง่าย อำนาจจำแนกของข้อสอบ และ ความเที่ยงแบบสอดคล้องภายใน ตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory)

1.1 ผลการวิเคราะห์ คะแนนต่ำสุด คะแนนสูงสุด ค่าเฉลี่ย ค่ามัธยฐาน ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน ความเบ้ ความโด่ง

ตาราง 8 คะแนนสูงสุด คะแนนต่ำสุด ค่ามัธยฐาน ค่าเฉลี่ยเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน ความเบ้ ความโด่ง ของกลุ่มข้อมูลที่ถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบจำนวน 4 กลุ่ม

	L35N100	L35N250	L65N100	L65N250
คะแนนต่ำสุด	12	9	23	20
คะแนนสูงสุด	33	34	57	62
ค่าเฉลี่ย	24.04	24.384	43.610	44.356
ค่ามัธยฐาน	24	25	44.5	46
ส่วนเบี่ยงเบนมาตรฐาน (S)	4.728	4.96	8.862	8.999
ความแปรปรวน (S^2)	22.358	24.605	78.538	80.981
ความเบ้ (sk)	-0.488	-0.577	-0.488	-0.594
ความโด่ง (ku)	-0.374	-0.157	-0.712	-0.482

กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100) มีคะแนนต่ำสุด 12 คะแนน คะแนนสูงสุด 33 คะแนน มีค่าเฉลี่ยของคะแนนเป็น 24.04 คะแนน มีค่าส่วนเบี่ยงเบนมาตรฐานเป็น 4.728 ผู้สอบในกลุ่มนี้ส่วนใหญ่มีคะแนนรวมที่สูง (sk = -0.488) และข้อมูลในกลุ่มนี้การกระจายของข้อมูลมาก (ku = -0.374)

กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250) มีคะแนนต่ำสุด 9 คะแนน คะแนนสูงสุด 34 คะแนน มีค่าเฉลี่ยของคะแนนเป็น 24.384 คะแนน มีค่าส่วนเบี่ยงเบนมาตรฐานเป็น 4.96 ผู้สอบในกลุ่มนี้ส่วนใหญ่มีคะแนนรวมที่สูง (sk = -0.577) และข้อมูลในกลุ่มนี้การกระจายของข้อมูลมาก (ku = -0.157)

กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100) มีคะแนนต่ำสุด 23 คะแนน คะแนนสูงสุด 57 คะแนน มีค่าเฉลี่ยของคะแนนเป็น 43.610 คะแนน มีค่าส่วนเบี่ยงเบนมาตรฐานเป็น 8.862 ผู้สอบในกลุ่มนี้ส่วนใหญ่มีคะแนนรวมที่สูง (sk = -0.488) และข้อมูลในกลุ่มนี้การกระจายของข้อมูลมาก (ku = -0.712)

กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250) มีคะแนนต่ำสุด 20 คะแนน คะแนนสูงสุด 62 คะแนน มีค่าเฉลี่ยของคะแนนเป็น 44.356 คะแนน มีค่าส่วน

เบี่ยงเบนมาตรฐานเป็น 8.999 ผู้สอบในกลุ่มนี้ส่วนใหญ่มีคะแนนรวมที่สูง (sk = -0.594) และข้อมูลในกลุ่มนี้การกระจายของข้อมูลมาก (ku = -0.482)

1.2 ผลการวิเคราะห์ ค่าความยากง่าย อำนาจจำแนกของข้อสอบ และ ความเที่ยงแบบสอดคล้องภายใน ตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory)

ตาราง 9 ค่าเฉลี่ยของค่าความยากง่ายของข้อสอบ ค่าเฉลี่ยของอำนาจจำแนกของข้อสอบ และ ค่าความเที่ยงแบบความสอดคล้องภายในของกลุ่มข้อมูลที่ถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบจำนวน 4 กลุ่ม

	L35N100	L35N250	L65N100	L65N250
ค่าเฉลี่ยของค่าความยากง่ายของข้อสอบ (p_m)	0.687	0.697	0.671	0.682
ค่าเฉลี่ยของอำนาจจำแนกของข้อสอบ (r_m)	0.242	0.245	0.224	0.233
ค่าเฉลี่ยของอำนาจจำแนกของข้อสอบซึ่งคำนวณโดยวิธี Point-Biserial Correlation Coefficient (r_{pbism})	0.305	0.322	0.303	0.312
ค่าความเที่ยงแบบความสอดคล้องภายในวิธีแบ่งครึ่งข้อสอบเป็นครึ่งแรก-ครึ่งหลัง	0.636	0.677	0.778	0.79
ค่าความเที่ยงแบบความสอดคล้องภายในวิธีแบ่งครึ่งข้อสอบเป็นข้อคู่-ข้อคี่	0.759	0.759	0.898	0.898
ค่าความเที่ยงแบบความสอดคล้องภายในวิธีของคูเดอร์-ริชาร์ดสันแบบสูตร KR20	0.711	0.744	0.848	0.857
ค่าความเที่ยงแบบความสอดคล้องภายในวิธีของคูเดอร์-ริชาร์ดสันแบบสูตร KR21	0.683	0.72	0.83	0.839

เมื่อวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม โดยใช้ข้อมูลกลุ่มความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100) แล้วพบว่า ข้อสอบฉบับ 35 ข้อนี้นั้นค่อนข้างง่าย (p_m=0.687) ข้อสอบส่วนใหญ่มีอำนาจจำแนกพอใช้ได้ (r_m= 0.242 และ r_{pbism}= 0.305) และแบบสอบฉบับนี้มีความเที่ยงแบบความสอดคล้องภายใน เนื่องจากค่าความเที่ยงแบบสอดคล้องภายในที่คำนวณจากทั้ง 4 วิธีตามตารางข้างต้นมีค่าใกล้เคียงกันโดยมีค่าเป็นบวกและมีค่าที่ค่อนข้างสูง

เมื่อวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม โดยใช้ข้อมูลกลุ่มความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250) แล้วพบว่า ข้อสอบฉบับ 35 ข้อนี้นั้นค่อนข้างง่าย (p_m= 0.697) ข้อสอบส่วนใหญ่มีอำนาจจำแนกพอใช้ได้ (r_m= 0.245 และ r_{pbism}= 0.322) และแบบสอบฉบับนี้มีความเที่ยงแบบความสอดคล้องภายใน เนื่องจากค่าความเที่ยงแบบสอดคล้องภายในที่คำนวณจากทั้ง 4 วิธีตามตารางข้างต้นมีค่าใกล้เคียงกันโดยมีค่าเป็นบวกและมีค่าที่ค่อนข้างสูง

เมื่อวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม โดยใช้ข้อมูลกลุ่มความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100) แล้วพบว่า ข้อสอบฉบับ 35 ข้อนี้นั้นค่อนข้างง่าย (p_m= 0.671) ข้อสอบส่วนใหญ่มีอำนาจจำแนกพอใช้ได้ (r_m= 0.224 และ r_{pbism}= 0.303) และแบบสอบฉบับนี้มีความเที่ยงแบบความสอดคล้องภายใน เนื่องจากค่าความเที่ยงแบบสอดคล้องภายในที่คำนวณจากทั้ง 4 วิธีตามตารางข้างต้นมีค่าใกล้เคียงกันโดยมีค่าเป็นบวกและมีค่าที่ค่อนข้างสูง

เมื่อวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม โดยใช้ข้อมูลกลุ่มความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250) แล้วพบว่า ข้อสอบฉบับ 35 ข้อนี้นั้นค่อนข้างง่าย (p_m= 0.682) ข้อสอบส่วนใหญ่มีอำนาจจำแนกพอใช้ได้ (r_m= 0.233 และ r_{pbism}= 0.312) และแบบสอบฉบับนี้มีความเที่ยงแบบความสอดคล้องภายใน เนื่องจากค่าความเที่ยงแบบสอดคล้องภายในที่คำนวณจากทั้ง 4 วิธีตามตารางข้างต้นมีค่าใกล้เคียงกันโดยมีค่าเป็นบวกและมีค่าที่ค่อนข้างสูง

จากตาราง 8 และ 9 ข้างต้นสามารถสรุปได้ว่า ข้อมูลทั้ง 4 กลุ่มคล้ายคลึงกัน ถึงแม้ได้มีการสุ่มอย่างง่ายเพื่อแบ่งให้จำนวนผู้สอบ 250 คน เป็น 100 คน และ สุ่มอย่างง่ายเพื่อแบ่งขุดแบบสอบ 65 ข้อเป็น 35 ข้อก็ไม่ได้ทำให้คุณลักษณะของข้อมูลคะแนนผู้สอบและคุณลักษณะของข้อสอบและแบบสอบ แตกต่างกัน กล่าวคือ ข้อมูลทุกกลุ่มนั้นผู้สอบส่วนใหญ่จะมีคะแนนรวมที่สูงซึ่งพิจารณาจากค่าความเบ้ซึ่งมีค่าเป็นลบ นอกจากนี้ข้อมูลทุกกลุ่มมีการกระจายของข้อมูลมากซึ่งพิจารณาจากค่าความโด่งซึ่งมีค่าเป็นลบ

เมื่อพิจารณาถึงผลการวิเคราะห์ข้อสอบและแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิมโดยใช้ข้อมูลจากทั้งสี่กลุ่ม แล้วพบว่า ได้ผลการวิเคราะห์ที่คล้ายคลึงกัน กล่าวคือ โดยภาพรวมแล้วข้อสอบอยู่ในระดับค่อนข้างง่าย มีอำนาจจำแนกพอใช้ได้ และ แบบสอบมีความเที่ยงแบบความสอดคล้องภายใน

2. ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนอง ข้อสอบ (Item Response Theory: IRT)

ผู้วิจัยทำการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ โดยใช้โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (Three-Parameter Model) เพื่อนำผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบไปใช้ในการคัดเลือกข้อสอบที่จะอยู่ในกลุ่มข้อสอบระดับยาก

โดยจากผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ ของทั้ง 4 กลุ่มข้อมูล พบว่า ข้อสอบส่วนใหญ่จะเป็นข้อสอบที่ค่อนข้างง่าย และหากผู้วิจัยใช้เกณฑ์ในการพิจารณาเพื่อระบุว่าข้อสอบระดับยากเป็นข้อสอบข้อใดตามเกณฑ์ที่ระบุไว้ในหนังสือ ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories) ของ ศ.ดร. ศิริชัย กาญจนวาสี ซึ่งระบุไว้ว่าข้อสอบที่จัดอยู่ในระดับยากตามทฤษฎีการตอบสนองข้อสอบ ควรมีค่าพารามิเตอร์ความยากของข้อสอบ (b) อยู่ใกล้ +2.50 ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a) ควรอยู่ในช่วง +0.50 ถึง +2.50 และ ค่าพารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก (c) ไม่ควรเกิน 0.30 โดยหลังจากการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแล้วพบว่าข้อสอบที่เข้าเกณฑ์ดังกล่าวจะมีจำนวนข้อที่น้อยซึ่งไม่เพียงพอต่อการนำไปจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (simulation) ดังนั้นผู้วิจัยจึงได้ปรับเกณฑ์ในการพิจารณาเพื่อระบุข้อสอบระดับยากเป็น ข้อสอบที่จะถูกระบุว่าเป็นข้อสอบระดับยาก ต้องมีค่าพารามิเตอร์ ดังต่อไปนี้

1. ค่าพารามิเตอร์ความยากของข้อสอบ (b) มีค่ามากกว่า -0.90
2. ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a) มีค่ามากกว่า +0.20
3. ค่าพารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก (c) ไม่เกิน +0.65

ซึ่งผู้วิจัยแบ่งการนำเสนอผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ ออกเป็น 4 กลุ่มซึ่งแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และจำนวนผู้สอบ ดังนี้คือ

2.1 กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100)

2.2 กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250)

2.3 กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100)

2.4 กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250)

2.1 **กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน** (L35N100)

ตาราง 10 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยากของข้อสอบแต่ละข้อ ของกลุ่มข้อมูล ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน

ข้อสอบ	a*	b*	c*	ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก**
1	9.83	0.53	0.46	✓
2	0.56	0.84	0.08	✓
3	0.24	-3.37	0	
4	1.18	-1.08	0	
5	0	-116.89	0.2	
6	1.09	-0.7	0	✓
7	0.48	-0.41	0	✓
8	0.14	-1.51	0	
9	0.72	-1.11	0	
10	0.49	0.22	0.38	✓
11	2.31	0.9	0.58	✓
12	6.91	0.74	0.63	✓
13	0.22	-1.21	0	
14	0.11	-3.17	0	
15	1.16	2.12	0.26	✓
16	0.8	0.63	0.42	✓
17	1.3	-0.59	0.11	✓
18	5.19	0.46	0.78	
19	0.5	-1.31	0.26	
20	3.43	-0.18	0.72	
21	0.2	-1.44	0	
22	0.2	-4.54	0	
23	0.61	-2.38	0	
24	0.74	-0.55	0.22	✓
25	1.42	1.81	0.55	✓
26	0.19	-7.15	0	
27	0.11	-2.13	0	
28	0.6	0.54	0.38	✓
29	0.64	-0.32	0	✓
30	0.46	-1.19	0	
31	0.49	1.42	0.48	✓
32	-0.05	10.96	0	
33	2.51	0.81	0.62	✓
34	1.09	-0.43	0.75	
35	0.73	-0.2	0	✓

* a หมายถึง พารามิเตอร์อำนาจจำแนกของข้อสอบ b หมายถึง พารามิเตอร์ความยากของข้อสอบ c หมายถึง พารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก

** เกณฑ์ที่ใช้ในการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก คือ a > +0.20 , b > -0.90 และ c < 0.65

จากตารางข้างต้น ข้อสอบที่สามารถนำไปใช้ในการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ ในสถานการณ์การลอกข้อสอบที่ลอกเฉพาะข้อสอบยากได้แก่ ข้อ 1, 2, 6, 7, 10, 11, 12, 15, 16, 17, 24, 25, 28, 29, 31, 33 และ 35

2.2 **กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน** (L35N250)

ตาราง 11 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยากของข้อสอบแต่ละข้อ ของกลุ่มข้อมูล ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน

ข้อสอบ	a*	b*	c*	ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก**
1	0.58	-0.9	0	✓
2	0.75	0.82	0.2	✓
3	0.51	-2.35	0	
4	0.75	-0.76	0.4	✓
5	0.45	-1.3	0	
6	0.54	-1.08	0	
7	0.38	-0.56	0	✓
8	0.12	0	0	
9	0.97	-1.06	0	
10	1.3	0.2	0.53	✓
11	5.6	0.42	0.6	✓
12	7.52	0.38	0.54	✓
13	0.44	-0.57	0	✓
14	0.33	-1.4	0	
15	0.05	8.9	0	
16	0.42	-1.2	0	
17	1.21	-0.53	0.26	✓
18	0.21	-0.94	0.69	
19	0.29	-2.82	0	
20	0.74	-0.25	0.63	✓
21	0	-51.11	0.19	
22	0.39	-1.9	0	
23	0.65	-1.04	0.6	
24	0.34	-1.48	0	
25	1.03	1.28	0.52	✓
26	0.35	-3.9	0	
27	0.42	0.84	0.37	✓
28	0.47	-0.96	0	
29	0.74	-0.63	0	✓
30	0.4	-0.84	0.22	✓
31	0.58	1	0.41	✓
32	0.06	-9.18	0	
33	0.45	-1.33	0	
34	0.83	-2.09	0	
35	0.67	-0.21	0	✓

* a หมายถึง พารามิเตอร์อำนาจจำแนกของข้อสอบ b หมายถึง พารามิเตอร์ความยากของข้อสอบ c หมายถึง พารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก

** เกณฑ์ที่ใช้ในการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก คือ a > +0.20 , b > -0.90 และ c < 0.65

จากตารางข้างต้น ข้อสอบที่สามารถนำไปใช้ในการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ ในสถานการณ์การลอกข้อสอบที่ลอกเฉพาะข้อสอบยากได้แก่ ข้อ 1, 2, 4, 7, 10, 11, 12, 13, 17, 20, 25, 27, 29, 30, 31 และ 35

2.3 **กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน** (L65N100)

ตาราง 12 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยากของข้อสอบแต่ละข้อ ของกลุ่มข้อมูล ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน

ข้อสอบ	a*	b*	c*	ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก**	ข้อสอบ	a*	b*	c*	ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก**
1	1.82	0.7	0.47	✓	34	5.05	0.42	0.77	
2	0.51	0.15	0	✓	35	0.52	-0.06	0	✓
3	0.43	0.84	0.03	✓	36	0.7	-0.22	0.53	✓
4	0.62	-0.03	0.28	✓	37	-0.15	-4.47	0	
5	0.37	-0.31	0.53	✓	38	0.54	-2.56	0	
6	4.96	0.7	0.59	✓	39	0.09	-3.17	0	
7	5.57	-0.12	0.45	✓	40	0.24	-0.74	0	✓
8	1.26	-0.02	0.15	✓	41	0.18	-5.11	0	
9	-0.02	13.05	0.17		42	0.41	0.02	0	✓
10	5.56	0.06	0.39	✓	43	2.01	-0.2	0.7	
11	1.08	-0.63	0	✓	44	6.43	0.72	0.78	
12	1.49	0.26	0.3	✓	45	0.61	-0.75	0.14	✓
13	0.81	0.06	0.14	✓	46	5.11	1.07	0.79	
14	-3.31	-1.26	0.49		47	1.62	1.9	0.56	✓
15	0.19	-1.1	0		48	0.57	-1.54	0	
16	1.34	-0.19	0.24	✓	49	0.06	-24.57	0	
17	1.28	-0.15	0.39	✓	50	0.24	-2.74	0	
18	1	-0.27	0	✓	51	0.24	-0.94	0	
19	0.47	-0.93	0		52	0.33	-0.54	0	✓
20	0.77	-1.06	0		53	0.38	-0.82	0	✓
21	0.91	1.04	0.55	✓	54	1.52	-0.86	0	✓
22	1.39	-0.34	0.22	✓	55	0.64	-0.27	0	✓
23	1.33	0.47	0.54	✓	56	0.04	-10.89	0.42	
24	0.79	-0.41	0.11	✓	57	0.73	0.15	0.41	✓
25	0.27	-0.94	0		58	3.6	1.03	0.82	
26	4.79	0.82	0.57	✓	59	0.77	1.9	0.56	✓
27	0.16	-2.19	0		60	0.08	-7.72	0	
28	0.54	-0.32	0	✓	61	6.88	0.12	0.76	
29	0.05	9.68	0		62	5.57	0.86	0.63	✓
30	1.17	-0.56	0.34	✓	63	4.57	1.28	0.52	✓
31	1.76	0.6	0.44	✓	64	1.03	-0.55	0.71	
32	1.05	-0.73	0	✓	65	1.22	0.28	0.2	✓
33	-0.37	-3.49	0.04						

* a หมายถึง พารามิเตอร์อำนาจจำแนกของข้อสอบ b หมายถึง พารามิเตอร์ความยากของข้อสอบ c หมายถึง พารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก

** เกณฑ์ที่ใช้ในการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก คือ a > +0.20 , b > -0.90 และ c < 0.65

จากตารางข้างต้น ข้อสอบที่สามารถนำไปใช้ในการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ ในสถานการณ์การลอกข้อสอบที่ลอกเฉพาะข้อสอบยากได้แก่ ข้อ 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 16, 17, 18, 21, 22, 23, 24, 26, 28, 30, 31, 32, 35, 36, 40, 42, 45, 47, 52, 53, 54, 55, 57, 59, 62, 63 และ 65

2.4 **กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน** (L65N250)

ตาราง 13 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยากของข้อสอบแต่ละข้อ ของกลุ่มข้อมูล ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน

ข้อสอบ	a*	b*	c*	ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก**	ข้อสอบ	a*	b*	c*	ผลการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก**
1	0.58	-0.85	0	✓	34	0.35	-0.67	0.69	
2	0.56	0.47	0.12	✓	35	0.76	0.09	0.18	✓
3	0.81	0.8	0.19	✓	36	0.29	-2.83	0	
4	0.8	0.01	0.35	✓	37	-2.04	-2.57	0.21	
5	0.51	-1.58	0.36		38	1.18	0.31	0.72	
6	0.64	0.14	0.35	✓	39	-0.21	-3.54	0.56	
7	0.88	-0.44	0.47	✓	40	0.3	-0.89	0	✓
8	1.27	0.03	0.31	✓	41	0.39	-1.89	0	
9	0.46	-0.62	0.24	✓	42	0.26	-0.55	0	✓
10	1.74	-0.09	0.34	✓	43	1.49	-0.07	0.75	
11	0.73	-0.68	0.08	✓	44	0.72	0.14	0.7	
12	0.83	0.43	0.38	✓	45	0.36	-1.36	0	
13	0.87	0.44	0.3	✓	46	0.29	-3.21	0	
14	-0.62	-2.64	0.47		47	0.61	1.02	0.43	✓
15	0.1	0.09	0		48	0.69	-1.55	0	
16	1.53	0.12	0.47	✓	49	0.36	-1.98	0.59	
17	0.9	-1.04	0		50	0.51	-0.09	0.46	✓
18	0.97	-0.47	0	✓	51	0.67	0.85	0.41	✓
19	1.16	0.12	0.49	✓	52	0.37	-0.62	0	✓
20	1.62	-0.43	0.41	✓	53	0.52	-0.83	0	✓
21	1.18	0.11	0.49	✓	54	1.47	-0.48	0.35	✓
22	1.01	-0.77	0	✓	55	0.77	-0.57	0	✓
23	1.37	0.07	0.43	✓	56	2.05	1.21	0.78	
24	0.6	-0.94	0		57	0.71	0.17	0.44	✓
25	0.44	-0.5	0	✓	58	0.03	-22.03	0.19	
26	0.51	-1.06	0		59	0.68	1.21	0.45	✓
27	0.44	-0.29	0.27	✓	60	0.07	-8.94	0	
28	0.71	-0.57	0	✓	61	0.34	-3.04	0	
29	0.06	7.71	0		62	0.74	0.21	0.46	✓
30	1.85	-0.45	0.42	✓	63	0.43	1.12	0.35	✓
31	0.74	0.02	0.35	✓	64	0.81	-1.77	0.26	
32	1.21	-0.52	0.23	✓	65	0.67	-0.16	0	✓
33	-0.28	-4.23	0						

* a หมายถึง พารามิเตอร์อำนาจจำแนกของข้อสอบ b หมายถึง พารามิเตอร์ความยากของข้อสอบ c หมายถึง พารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก

** เกณฑ์ที่ใช้ในการพิจารณาเพื่อระบุความเป็นข้อสอบระดับยาก คือ a > +0.20 , b > -0.90 และ c < 0.65

จากตารางข้างต้น ข้อสอบที่สามารถนำไปใช้ในการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ ในสถานการณ์การลอกข้อสอบที่ลอกเฉพาะข้อสอบยากได้แก่ ข้อ 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 16, 18, 19, 20, 21, 22, 23, 25, 27, 28, 30, 31, 32, 35, 40, 42, 47, 50, 51, 52, 53, 54, 55, 57, 59, 62, 63 และ 65

3. ผลการวิเคราะห์ค่าพารามิเตอร์ของผู้สอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)

ผู้วิจัยทำการวิเคราะห์ค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) ตามทฤษฎีการตอบสนองข้อสอบ เพื่อนำค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) มาเรียงลำดับผู้สอบในแต่ละกลุ่มซึ่งได้ถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบตามระดับความสามารถจากมากไปน้อย จากนั้นกำหนดผู้ให้ลอกซึ่งเป็นผู้สอบที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 และ 90 ของแต่ละกลุ่มข้อมูล

ผู้วิจัยแบ่งการนำเสนอผลการวิเคราะห์ค่าพารามิเตอร์ของผู้สอบตามทฤษฎีการตอบสนองข้อสอบ ออกเป็น 4 กลุ่มซึ่งได้ถูกแบ่งออกตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ ดังนี้คือ

1. กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100)
2. กลุ่ม ความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250)
3. กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100)
4. กลุ่ม ความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 14 ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม MULTILOG ของผู้สอบแต่ละคน จากกลุ่มความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน (L35N100) เรียงตามค่าเปอร์เซ็นไทล์

No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P
11	-2.268	1	90	-0.428	26	88	0.27	51	10	0.821	76
22	-1.991	2	60	-0.425	27	87	0.287	52	71	0.821	76
63	-1.769	3	42	-0.412	28	94	0.306	53	35	0.84	78
36	-1.707	4	95	-0.407	28	4	0.32	54	43	0.851	79
39	-1.649	5	20	-0.396	30	34	0.355	55	69	0.879	80
21	-1.585	6	18	-0.393	31	96	0.395	56	85	0.882	81
28	-1.564	7	41	-0.289	32	56	0.427	56	52	0.892	82
57	-1.435	8	19	-0.282	33	78	0.43	57	64	0.912	83
8	-1.377	9	98	-0.252	34	72	0.534	59	23	0.984	84
32	-1.325	10	13	-0.247	35	58	0.555	60	46	1	85
25	-1.312	11	83	-0.241	36	82	0.556	61	30	1.008	86
24	-1.168	12	33	-0.167	37	77	0.593	62	5	1.044	87
79	-1.123	13	15	-0.152	38	50	0.602	63	93	1.052	88
14	-1.061	14	29	-0.093	39	100	0.605	64	73	1.075	89
70	-1	14	12	-0.016	40	47	0.609	65	92	1.131	90
31	-0.939	16	44	-0.006	41	16	0.612	66	62	1.193	91
59	-0.858	17	45	0.034	42	75	0.626	67	67	1.193	91
26	-0.749	18	86	0.077	43	61	0.644	68	3	1.198	93
38	-0.615	19	7	0.078	44	97	0.647	69	2	1.225	94
81	-0.594	20	84	0.086	45	9	0.693	70	40	1.289	95
65	-0.54	21	99	0.086	45	51	0.725	71	68	1.297	96
74	-0.532	22	54	0.098	47	76	0.765	72	53	1.311	97
27	-0.493	23	55	0.11	48	37	0.778	73	91	1.423	98
6	-0.489	24	80	0.147	49	17	0.796	74	49	1.47	99
48	-0.479	25	1	0.232	50	89	0.805	75	66	1.905	100

จากตารางข้างต้น ผู้วิจัยสามารถทำการระบุผู้สอบที่จะถูกกำหนดให้เป็นผู้ให้ลอก สำหรับข้อมูลกลุ่มความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน ได้ คือ ผู้ให้ลอกซึ่งอยู่ในตำแหน่งเปอร์เซ็นไทล์ที่ 60 คือ ผู้สอบเลขที่ 58 และผู้ให้ลอกซึ่งอยู่ในตำแหน่งเปอร์เซ็นไทล์ที่ 90 คือ ผู้สอบเลขที่ 92

ตาราง 15 ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม MULTILOG ของผู้สอบแต่ละคน จากกลุ่มความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน (L35N250) เรียงตามค่าเปอร์เซ็นไทล์

No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P
49	-3.033	0.2	99	-0.942	13	185	-0.573	25.8	246	-0.23	38.6	33	0.232	51.4	140	0.521	64.2	110	0.679	77	214	0.985	89.8
53	-2.766	0.6	58	-0.94	13.4	45	-0.558	26.2	146	-0.226	39	144	0.241	51.8	219	0.524	64.6	102	0.681	77.4	9	0.993	90.2
21	-2.082	1	130	-0.931	13.8	80	-0.538	26.6	212	-0.22	39.4	134	0.253	52.2	171	0.527	65	152	0.682	77.8	4	0.997	90.6
18	-1.927	1.4	250	-0.925	14.2	34	-0.534	27	237	-0.208	39.8	172	0.255	52.6	139	0.533	65.4	64	0.697	78.2	236	1.056	91
48	-1.856	1.8	60	-0.923	14.6	15	-0.518	27.4	187	-0.203	40.2	2	0.265	53	129	0.547	65.8	104	0.698	78.6	136	1.073	91.4
113	-1.759	2.2	54	-0.916	15	43	-0.485	27.8	20	-0.191	40.6	114	0.302	53.4	12	0.554	66.4	143	0.699	79	83	1.087	92
79	-1.736	2.6	23	-0.91	15.4	201	-0.47	28.2	203	-0.174	41	123	0.351	53.8	100	0.554	66.4	209	0.714	79.4	154	1.087	92
16	-1.697	3	57	-0.832	15.8	164	-0.425	28.6	178	-0.157	41.6	199	0.364	54.2	174	0.557	67	106	0.717	79.8	141	1.094	92.6
137	-1.693	3.4	70	-0.801	16.2	105	-0.424	29	242	-0.157	41.6	168	0.38	54.6	38	0.561	67.4	135	0.73	80.2	3	1.096	93
247	-1.635	3.8	124	-0.79	16.6	200	-0.405	29.4	177	-0.1	42.2	156	0.381	55	76	0.564	67.8	163	0.735	80.6	243	1.126	93.4
24	-1.629	4.2	26	-0.783	17	224	-0.403	29.8	145	0.02	42.6	208	0.39	55.4	94	0.573	68.2	82	0.744	81	78	1.138	93.8
61	-1.552	4.6	44	-0.779	17.4	91	-0.393	30.2	72	0.028	43	160	0.403	55.8	245	0.575	68.6	7	0.745	81.4	149	1.163	94.2
207	-1.49	5	31	-0.762	17.8	239	-0.39	30.6	121	0.032	43.4	37	0.405	56.2	153	0.589	69	69	0.76	81.8	109	1.176	94.6
47	-1.475	5.4	241	-0.76	18.2	180	-0.387	31.2	159	0.037	43.8	116	0.411	56.8	220	0.595	69.4	148	0.792	82.2	215	1.184	95
55	-1.416	5.8	56	-0.742	18.6	202	-0.387	31.2	229	0.041	44.2	218	0.411	56.8	226	0.597	69.8	150	0.804	82.6	221	1.198	95.4
232	-1.409	6.2	39	-0.708	19	132	-0.38	31.8	191	0.042	44.6	235	0.428	57.4	162	0.604	70.2	10	0.832	83	206	1.255	95.8
125	-1.399	6.6	126	-0.696	19.4	19	-0.372	32.4	8	0.086	45	196	0.434	57.8	204	0.606	70.6	165	0.856	83.4	228	1.285	96.2
32	-1.397	7	107	-0.691	19.8	40	-0.372	32.4	1	0.087	45.4	96	0.438	58.2	97	0.607	71	225	0.878	83.8	222	1.288	96.6
28	-1.351	7.4	188	-0.69	20.2	138	-0.344	33	233	0.088	45.8	197	0.44	58.6	151	0.621	71.4	103	0.88	84.2	216	1.332	97
50	-1.334	8	59	-0.675	20.6	101	-0.337	33.4	67	0.097	46.2	117	0.441	59.2	211	0.622	71.8	92	0.885	84.6	183	1.362	97.4
52	-1.334	8	81	-0.67	21	166	-0.334	33.8	238	0.118	46.6	223	0.441	59.2	147	0.63	72.4	169	0.887	85	190	1.614	97.8
71	-1.303	8.6	11	-0.668	21.6	29	-0.329	34.2	170	0.119	47	194	0.443	59.8	217	0.63	72.4	205	0.89	85.4	192	1.615	98.2
176	-1.285	9	13	-0.668	21.6	122	-0.309	34.6	182	0.144	47.4	120	0.448	60.2	5	0.637	73	87	0.896	85.8	65	1.797	98.6
158	-1.223	9.4	184	-0.661	22.2	41	-0.305	35	6	0.168	48	186	0.45	60.6	173	0.64	73.4	161	0.912	86.2	195	1.806	99
167	-1.202	9.8	17	-0.649	22.6	249	-0.295	35.4	234	0.168	48	27	0.464	61.2	66	0.642	73.8	157	0.915	86.6	89	1.838	99.4
30	-1.123	10.2	25	-0.639	23	35	-0.281	35.8	127	0.182	48.6	75	0.464	61.2	133	0.645	74.2	118	0.927	87	131	1.956	99.8
155	-1.12	10.6	73	-0.619	23.4	62	-0.28	36.2	175	0.192	49	181	0.466	62	193	0.646	74.6	231	0.943	87.4			
46	-1.07	11	198	-0.614	23.8	119	-0.271	36.6	36	0.194	49.4	213	0.466	62	95	0.652	75	128	0.945	87.8			
68	-1.036	11.4	84	-0.604	24.2	63	-0.252	37.2	240	0.213	49.8	85	0.476	62.6	112	0.654	75.4	22	0.954	88.2			
108	-1.035	11.8	179	-0.597	24.6	90	-0.252	37.2	14	0.221	50.2	189	0.505	63	244	0.661	75.8	227	0.957	88.6			
86	-0.967	12.2	142	-0.589	25	98	-0.247	37.8	248	0.222	50.6	51	0.506	63.4	77	0.662	76.2	210	0.959	89			
115	-0.962	12.6	42	-0.576	25.4	93	-0.24	38.2	111	0.225	51	74	0.509	63.8	230	0.672	76.6	88	0.976	89.4			

จากตารางข้างต้น ผู้วิจัยสามารถทำการระบุผู้สอบที่จะถูกกำหนดให้เป็นผู้ให้ลอก สำหรับข้อมูลกลุ่มความยาวของแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน ได้ คือ ผู้ให้ลอกซึ่งอยู่ในตำแหน่งเปอร์เซ็นไทล์ที่ 60 คือ ผู้สอบเลขที่ 120 และ ผู้ให้ลอกซึ่งอยู่ในตำแหน่งเปอร์เซ็นไทล์ที่ 90 คือ ผู้สอบเลขที่ 9

ตาราง 16 ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม Multilog ของผู้สอบแต่ละคน จากกลุ่มความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน (L65N100) เรียงตามค่าเปอร์เซ็นไทล์

No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P
32	-2.598	1	16	-0.448	26	83	0.231	51	68	0.848	76
8	-2.230	2	90	-0.429	27	55	0.232	52	92	0.854	77
24	-2.033	3	13	-0.418	28	75	0.238	53	23	0.862	78
36	-1.951	4	19	-0.295	28	99	0.239	54	10	0.873	79
31	-1.841	5	82	-0.213	30	87	0.314	55	50	0.885	80
28	-1.783	6	29	-0.203	31	96	0.340	56	78	0.905	81
11	-1.652	7	9	-0.146	32	80	0.344	56	30	0.910	82
25	-1.601	8	81	-0.145	33	4	0.346	57	89	0.920	83
70	-1.395	9	60	-0.125	34	58	0.370	59	51	0.940	84
21	-1.134	10	42	-0.065	35	86	0.380	60	52	0.994	85
26	-1.101	11	98	-0.062	36	6	0.435	61	56	1.016	86
79	-1.073	12	72	-0.054	37	84	0.436	62	35	1.090	87
22	-1.038	13	43	-0.016	38	94	0.439	63	69	1.111	88
38	-1.035	14	5	0.027	39	47	0.444	64	93	1.221	89
63	-1.031	14	12	0.034	40	85	0.444	64	61	1.247	90
57	-0.954	16	88	0.044	41	76	0.476	66	46	1.297	91
39	-0.921	17	48	0.060	42	100	0.497	67	71	1.378	92
65	-0.826	18	1	0.064	43	77	0.530	68	3	1.381	93
27	-0.777	19	44	0.067	44	97	0.565	69	73	1.399	94
74	-0.702	20	45	0.073	45	37	0.590	70	67	1.406	95
59	-0.633	21	17	0.082	46	64	0.614	71	40	1.412	96
14	-0.561	22	54	0.082	46	2	0.632	72	53	1.450	97
15	-0.487	23	7	0.142	48	34	0.782	73	66	1.582	98
41	-0.477	24	20	0.151	49	33	0.812	74	49	1.646	99
18	-0.456	25	95	0.174	50	62	0.812	74	91	2.041	100

จากตารางข้างต้น ผู้วิจัยสามารถทำการระบุผู้สอบที่จะถูกกำหนดให้เป็นผู้ให้ลอก สำหรับข้อมูลกลุ่มความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน ได้ คือ ผู้ให้ลอกซึ่งอยู่ในตำแหน่งเปอร์เซ็นไทล์ที่ 60 คือ ผู้สอบเลขที่ 86 และ ผู้ให้ลอกซึ่งอยู่ในตำแหน่งเปอร์เซ็นไทล์ที่ 90 คือ ผู้สอบเลขที่ 61

ตาราง 17 ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) ซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม Multilog ของผู้สอบแต่ละคน จากกลุ่มความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน (L65N250) เรียงตามค่าเปอร์เซ็นไทล์

No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P	No.	θ	P
53	-3.03	0.2	125	-1.038	13	241	-0.437	25.6	96	-0.08	38.6	127	0.175	51.4	233	0.42	64.2	123	0.745	77	154	1.087	89.8
48	-2.846	0.6	17	-1.024	13.6	42	-0.435	26.2	62	-0.075	39	6	0.183	51.8	75	0.421	64.6	148	0.759	77.4	118	1.093	90.2
50	-2.539	1.2	58	-1.024	13.6	43	-0.409	26.6	166	-0.058	39.4	186	0.184	52.2	10	0.424	65	129	0.772	78	4	1.143	90.6
52	-2.539	1.2	158	-0.994	14.2	122	-0.408	27	80	-0.051	39.8	178	0.186	52.6	171	0.428	65.4	227	0.772	78	102	1.149	91
49	-2.446	1.8	84	-0.989	14.6	93	-0.404	27.4	73	-0.048	40.2	105	0.194	53	114	0.457	65.8	244	0.774	78.6	9	1.199	91.4
79	-2.344	2.2	23	-0.945	15	101	-0.401	27.8	11	-0.047	40.6	174	0.21	53.4	111	0.458	66.2	82	0.779	79	215	1.209	91.8
21	-1.872	2.6	108	-0.933	15.4	98	-0.379	28.2	219	-0.043	41	38	0.221	53.8	210	0.461	66.6	134	0.785	79.6	225	1.216	92.2
18	-1.848	3	86	-0.88	15.8	35	-0.356	28.6	25	-0.039	41.4	72	0.232	54.4	172	0.486	67	173	0.785	79.6	83	1.234	92.6
16	-1.763	3.4	26	-0.859	16.4	15	-0.35	29	1	-0.032	41.8	193	0.232	54.4	150	0.487	67.4	149	0.815	80.2	161	1.239	93
113	-1.72	3.8	34	-0.859	16.4	202	-0.349	29.4	242	-0.029	42.2	14	0.245	55	106	0.495	67.8	222	0.827	80.6	141	1.26	93.4
61	-1.577	4.2	142	-0.824	17	20	-0.332	29.8	94	-0.016	42.6	240	0.247	55.4	133	0.513	68.2	5	0.836	81	183	1.267	93.8
55	-1.56	4.6	31	-0.807	17.4	200	-0.306	30.2	27	0.007	43.2	76	0.248	55.8	69	0.522	68.6	2	0.858	81.4	87	1.289	94.2
176	-1.526	5	59	-0.806	17.8	19	-0.295	30.6	197	0.007	43.2	223	0.261	56.2	140	0.544	69	136	0.868	81.8	128	1.332	94.6
68	-1.512	5.4	126	-0.803	18.2	138	-0.287	31	121	0.011	43.8	235	0.262	56.6	110	0.552	69.4	204	0.882	82.2	135	1.333	95
71	-1.482	5.8	250	-0.723	18.6	37	-0.283	31.4	224	0.013	44.2	175	0.268	57	162	0.553	69.8	153	0.885	82.8	228	1.362	95.4
167	-1.468	6.2	60	-0.694	19	212	-0.28	31.8	146	0.015	44.6	168	0.278	57.4	165	0.563	70.2	163	0.885	82.8	109	1.388	95.8
207	-1.455	6.6	39	-0.693	19.4	119	-0.275	32.2	185	0.048	45	189	0.281	57.8	209	0.578	70.6	230	0.895	83.4	103	1.469	96.2
24	-1.355	7	90	-0.692	19.8	107	-0.258	32.6	180	0.058	45.4	170	0.284	58.4	169	0.586	71	66	0.905	84	145	1.479	96.6
44	-1.348	7.4	56	-0.674	20.2	132	-0.241	33	239	0.059	45.8	248	0.284	58.4	238	0.592	71.4	143	0.905	84	221	1.548	97
247	-1.325	7.8	81	-0.645	20.6	117	-0.235	33.4	196	0.076	46.2	213	0.287	59	139	0.61	71.8	152	0.965	84.6	192	1.589	97.4
155	-1.32	8.2	188	-0.613	21	194	-0.205	33.8	181	0.083	46.6	13	0.288	59.4	116	0.618	72.2	22	0.978	85	195	1.797	97.8
28	-1.319	8.6	29	-0.609	21.4	63	-0.188	34.2	120	0.102	47	67	0.291	59.8	74	0.631	72.6	64	0.987	85.4	190	1.837	98.2
137	-1.308	9	40	-0.606	21.8	177	-0.186	34.6	8	0.115	47.4	218	0.302	60.2	157	0.633	73	88	0.992	85.8	206	1.872	98.6
232	-1.297	9.4	133	-0.522	22.2	45	-0.18	35	115	0.124	47.8	151	0.332	60.6	217	0.647	73.4	3	0.997	86.2	65	1.95	99
46	-1.29	9.8	99	-0.52	22.6	159	-0.167	35.4	187	0.138	48.2	231	0.333	61	199	0.657	73.8	77	1.003	86.6	89	2.004	99.4
47	-1.274	10.2	164	-0.513	23	246	-0.162	35.8	191	0.14	48.6	12	0.356	61.4	205	0.66	74.2	78	1.004	87	131	2.257	99.8
57	-1.196	10.6	36	-0.466	23.4	182	-0.14	36.2	144	0.148	49	51	0.367	61.8	95	0.667	74.6	216	1.007	87.4			
32	-1.193	11	124	-0.447	23.8	203	-0.139	36.6	160	0.155	49.4	245	0.369	62.2	220	0.671	75	156	1.024	87.8			
54	-1.15	11.4	201	-0.445	24.2	208	-0.135	37	85	0.16	50	226	0.388	62.6	243	0.686	75.4	7	1.041	88.2			
30	-1.147	11.8	198	-0.443	24.6	33	-0.117	37.4	97	0.16	50	104	0.397	63	112	0.692	75.8	214	1.047	88.6			
70	-1.101	12.2	184	-0.442	25	234	-0.11	37.8	249	0.168	50.6	91	0.415	63.4	211	0.693	76.2	147	1.069	89			
41	-1.087	12.6	179	-0.437	25.6	237	-0.087	38.2	229	0.174	51	100	0.419	63.8	92	0.717	76.6	236	1.083	89.4			

จากตารางข้างต้น ผู้วิจัยสามารถทำการระบุผู้สอบที่จะถูกกำหนดให้เป็นผู้ให้ลอก สำหรับข้อมูลกลุ่มความยาวของแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน ได้ คือ ผู้ให้ลอกซึ่งอยู่ในตำแหน่งเปอร์เซ็นไทล์ที่ 60 คือ ผู้สอบเลขที่ 218 และผู้ให้ลอกซึ่งอยู่ในตำแหน่งเปอร์เซ็นไทล์ที่ 90 คือ ผู้สอบเลขที่ 118

จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 2 ผลการวิเคราะห์คุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ

ผลการวิเคราะห์คุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ ซึ่งได้แก่ ค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error Rate) และ อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) มีรายละเอียดดังนี้

1. ค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error Rate)

ผลการวิเคราะห์ค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error Rate) ของดัชนีตรวจจับการลอกข้อสอบแสดงในรูปแบบตารางและภาพโดยแยกตามปัจจัย จำนวนผู้สอบ และ ความยาวของแบบสอบ โดยค่าความคลาดเคลื่อนประเภทที่ 1 ซึ่งแสดงในรูปแบบตารางและภาพโดยแยกตามปัจจัย จำนวนผู้สอบ และ ความยาวของแบบสอบนั้นเป็นค่าเฉลี่ยของค่าความคลาดเคลื่อนประเภทที่ 1 จากสถานการณ์ที่ถูกแยกย่อยต่อจากตัวแปรต้นด้านจำนวนผู้สอบ และ ความยาวของแบบสอบ ซึ่งคือตัวแปรต้น ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนผู้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก โดยสาเหตุที่ผู้วิจัยไม่นำเสนอค่าความคลาดเคลื่อนประเภทที่ 1 แยกย่อยเป็น 128 สถานการณ์เนื่องจาก ค่าความคลาดเคลื่อนประเภทที่ 1 จะถูกคำนวณโดยมีพื้นฐานอยู่บนค่าดัชนีตรวจจับการลอกข้อสอบของผู้ที่ไม่ใช่ผู้ลอก (non-copiers) เท่านั้น ซึ่งตัวแปรต้น ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนผู้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก จะไม่ได้ส่งผลต่อค่าดัชนีตรวจจับการลอกข้อสอบของผู้ที่ไม่ใช่ผู้ลอก ดังนั้น ผู้วิจัยจึงทำการเฉลี่ยค่าความคลาดเคลื่อนประเภทที่ 1 ของกรณีต่างๆ ที่ถูกแยกออกตามตัวแปรต้นระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนผู้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก และ นำเสนอผลการวิเคราะห์ค่าความคลาดเคลื่อนประเภทที่ 1 แยกตามปัจจัยจำนวนผู้สอบ และ ความยาวแบบสอบ เท่านั้น

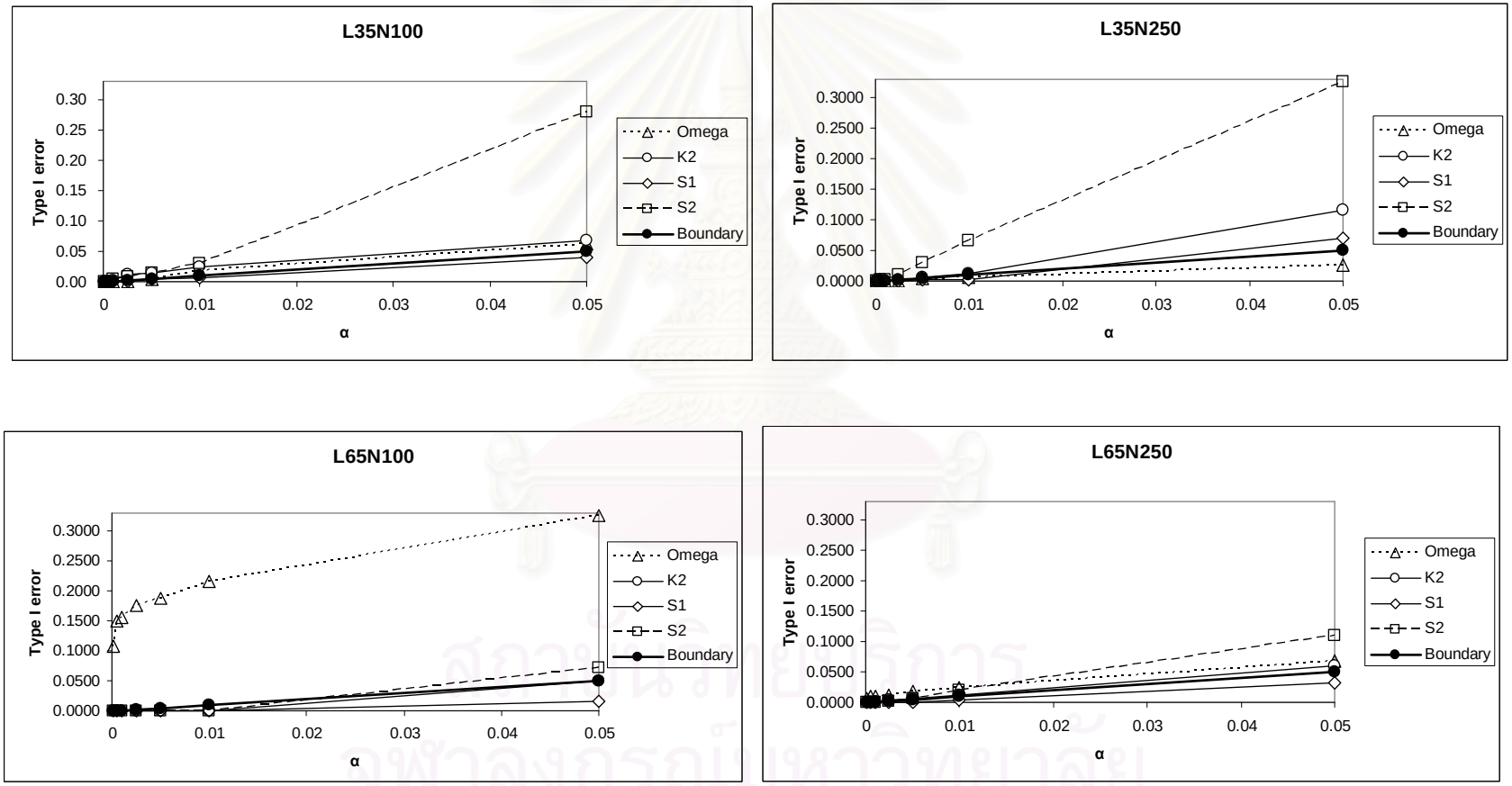
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 18 ความคลาดเคลื่อนประเภทที่ 1 (Type I error) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ณ ระดับนัยสำคัญ 7 ระดับ แยกตามตัวแปร จำนวนผู้สอบ (N) และ ความยาวของแบบสอบ (L)

สถานการณ์	α level	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
L35N100	ω	0.000003	0.000042	0.000042	0.000108	0.004767	0.018756	0.062533
	K2	0.000000	0.000107	0.002983	0.012002	0.014770	0.024907	0.068941
	S1	0.000000	0.000007	0.000018	0.000759	0.004177	0.006156	0.040589
	S2	0.000018	0.000406	0.003883	0.008216	0.014299	0.030718	0.280278
L35N250	ω	0.000236	0.000249	0.000256	0.000266	0.004420	0.005958	0.026048
	K2	0.000000	0.000003	0.000004	0.000360	0.006575	0.011503	0.115862
	S1	0.000000	0.000000	0.000003	0.000928	0.001429	0.002511	0.070155
	S2	0.000000	0.001298	0.001546	0.009574	0.029008	0.066286	0.326165
L65N100	ω	0.108133	0.150990	0.155906	0.176116	0.187067	0.216296	0.326672
	K2	0.000000	0.000004	0.000025	0.000042	0.000053	0.000912	0.049360
	S1	0.000000	0.000000	0.000000	0.000000	0.000004	0.000028	0.015405
	S2	0.000000	0.000000	0.000000	0.000004	0.000039	0.000183	0.072343
L65N250	ω	0.005082	0.010800	0.010889	0.012768	0.017566	0.024123	0.068561
	K2	0.000124	0.000286	0.000310	0.004288	0.006892	0.011968	0.059855
	S1	0.000000	0.000037	0.000085	0.000215	0.000761	0.003237	0.031899
	S2	0.000040	0.000148	0.000239	0.002331	0.006525	0.019006	0.109489

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาพ 8 การเปรียบค่าความคลาดเคลื่อนประเภทที่ 1 เชิงประจักษ์ (empirical type I error) ณ ระดับนัยสำคัญ 7 ระดับ ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) โดยแยกตามตัวแปรขนาดจำนวนผู้สอบ (simulee size) และ ความยาวของแบบสอบ (test length)



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ผลการศึกษาค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error) ของดัชนีตรวจจับการลอกข้อสอบ

จากตาราง 18 และ รูปภาพ 8 ผู้วิจัยสามารถสรุปผลการศึกษาค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบ ได้ดังนี้

1. **ดัชนี** ω

ดัชนี ω สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงหรือต่ำกว่าเส้นขอบเขต (boundary line) เพียงเล็กน้อย ในสถานการณ์ L35N100 และ L35N250 แต่มีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าเส้นขอบเขตจนไม่สามารถยอมรับได้ในสถานการณ์ L65N100 และ L65N250

2. **ดัชนี** K_2

ดัชนี K_2 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงหรือต่ำกว่าเส้นขอบเขต (boundary line) เพียงเล็กน้อย ในทุกสถานการณ์

3. **ดัชนี** S_1

ดัชนี S_1 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงหรือต่ำกว่าเส้นขอบเขต (boundary line) เพียงเล็กน้อย ในทุกสถานการณ์

4. **ดัชนี** S_2

ดัชนี S_2 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงหรือต่ำกว่าเส้นขอบเขต (boundary line) เพียงเล็กน้อย ในสถานการณ์ L65N100 และ L65N250 แต่มีระดับความคลาดเคลื่อนประเภทที่ 1 สูงกว่าเส้นขอบเขตจนไม่สามารถยอมรับได้ในสถานการณ์ L35N100 และ L35N250

การเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error) ของดัชนีตรวจจับการลอกข้อสอบ

จากตาราง 18 และ รูปภาพ 8 ผู้วิจัยสามารถเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ได้ว่า ดัชนี S_2 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) และ มีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) ดัชนี ω สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) และมีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) ส่วนดัชนี K_2 และ ดัชนี S_1 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในทุกสถานการณ์ แต่ค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนี S_1 ต่ำกว่าเส้นขอบเขตมากกว่าดัชนี K_2 ดังนั้น อาจกล่าวได้ว่า ดัชนี S_1 เป็นดัชนีที่มีระดับค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุด (the most conservative index) ในเกือบทุกสถานการณ์

2. **อำนาจการตรวจจับการลอกข้อสอบ** (Detection Rate)

ผลการวิเคราะห์ค่าอำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนีตรวจจับการลอกข้อสอบจะแสดงในรูปแบบตารางและภาพโดยแยกตามปัจจัยต้นที่ผู้วิจัยศึกษา คือ ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนผู้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก รวมทั้งสิ้น 128 สถานการณ์ โดยมีรายละเอียด ดังตาราง

โดยการนำเสนอผลการวิเคราะห์ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ทั้ง 128 สถานการณ์นั้นผู้วิจัยใช้สัญลักษณ์ย่อภาษาอังกฤษแทนตัวแปรต้นที่ผู้วิจัยศึกษาในการนำเสนอ ซึ่งสามารถศึกษาของสัญลักษณ์ย่อภาษาอังกฤษได้จากตาราง 19

ตาราง 19 สัญลักษณ์ย่อภาษาอังกฤษแทนตัวแปรต้นที่ศึกษาในงานวิจัย

ตัวแปรต้น	สัญลักษณ์ย่อ	ระดับย่อของตัวแปรต้น	สัญลักษณ์ย่อ
1. ความยาวของแบบสอบ (test length)	L	• แบบสอบความยาว 35 ข้อ • แบบสอบความยาว 65 ข้อ	L35 L65
2. จำนวนผู้สอบ (the number of simulees)	N	• จำนวนผู้สอบขนาด 100 คน • จำนวนผู้สอบขนาด 250 คน	N100 N250
3. ระดับความสามารถของผู้ให้ลอก (the ability levels of source)	S	• ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 • ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60	S90th S60th
4. ร้อยละของจำนวนผู้ลอก (the percentage of copiers)	C	• ร้อยละ 5 ของจำนวนผู้สอบทั้งหมด • ร้อยละ10 ของจำนวนผู้สอบทั้งหมด	C5% C10%
5. ร้อยละของจำนวนข้อสอบที่ถูกลอก (the percentage of answer copied)	A	• ร้อยละ 10 ของจำนวนข้อสอบทั้งหมด • ร้อยละ 20 ของจำนวนข้อสอบทั้งหมด • ร้อยละ 30 ของจำนวนข้อสอบทั้งหมด • ร้อยละ 40 ของจำนวนข้อสอบทั้งหมด	A10% A20% A30% A40%
6. วิธีการลอก (type of copying)	T	• การลอกข้อสอบแบบสุ่ม • การลอกข้อสอบเฉพาะข้อยาก	T(R) T(H)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 20 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2 และ ดัชนี ω ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 1-สถานการณ์ที่ 32)

Sit.	test length (L)	sample (N)	percentage of copier (C)	ability levels of source (S)	type of copying (T)	percentage of answer copied (A)	OMEGA INDEX DETECTION RATE							K2 INDEX DETECTION RATE						
							0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
001	L35	N100	C5%	S90th	T(R)	A10%	0.000	0.008	0.014	0.014	0.018	0.026	0.110	0.004	0.004	0.004	0.048	0.048	0.048	0.200
002						A20%	0.012	0.018	0.022	0.028	0.046	0.066	0.208	0.008	0.008	0.008	0.070	0.072	0.072	0.322
003						A30%	0.026	0.046	0.064	0.116	0.146	0.258	0.568	0.058	0.084	0.084	0.298	0.350	0.350	0.700
004						A40%	0.066	0.118	0.152	0.236	0.284	0.398	0.656	0.070	0.102	0.102	0.334	0.376	0.376	0.732
005					T(H)	A10%	0.002	0.016	0.020	0.020	0.026	0.042	0.120	0.006	0.006	0.006	0.040	0.040	0.040	0.220
006						A20%	0.020	0.036	0.042	0.052	0.064	0.136	0.364	0.032	0.040	0.040	0.198	0.218	0.218	0.494
007						A30%	0.050	0.084	0.130	0.234	0.276	0.472	0.804	0.092	0.154	0.154	0.584	0.684	0.684	0.992
008						A40%	0.100	0.156	0.232	0.362	0.450	0.640	0.906	0.034	0.056	0.056	0.462	0.462	0.462	0.898
009				S60th	T(R)	A10%	0.000	0.024	0.032	0.072	0.110	0.112	0.218	0.020	0.026	0.058	0.060	0.094	0.192	0.374
010						A20%	0.002	0.048	0.062	0.116	0.148	0.148	0.322	0.010	0.028	0.094	0.094	0.114	0.192	0.410
011						A30%	0.050	0.184	0.212	0.252	0.252	0.290	0.558	0.022	0.042	0.106	0.114	0.142	0.250	0.600
012						A40%	0.242	0.310	0.356	0.424	0.518	0.652	0.944	0.094	0.106	0.166	0.172	0.194	0.456	0.916
013					T(H)	A10%	0.002	0.040	0.052	0.090	0.132	0.140	0.244	0.024	0.026	0.090	0.090	0.108	0.178	0.332
014						A20%	0.024	0.144	0.176	0.188	0.222	0.246	0.420	0.020	0.044	0.096	0.106	0.142	0.234	0.604
015						A30%	0.158	0.218	0.298	0.350	0.416	0.480	0.868	0.086	0.106	0.188	0.212	0.242	0.492	0.852
016						A40%	0.296	0.424	0.504	0.654	0.746	0.858	0.980	0.296	0.326	0.494	0.650	0.666	0.910	1.000
017			C10%	S90th	T(R)	A10%	0.006	0.052	0.066	0.074	0.095	0.114	0.259	0.009	0.029	0.031	0.077	0.115	0.115	0.427
018						A20%	0.041	0.064	0.072	0.096	0.121	0.149	0.317	0.015	0.033	0.035	0.106	0.162	0.162	0.505
019						A30%	0.139	0.167	0.188	0.288	0.346	0.488	0.710	0.022	0.147	0.147	0.202	0.452	0.452	0.835
020						A40%	0.161	0.201	0.230	0.370	0.410	0.506	0.787	0.012	0.029	0.035	0.214	0.261	0.261	0.630
021					T(H)	A10%	0.002	0.037	0.065	0.088	0.107	0.125	0.214	0.020	0.020	0.025	0.058	0.058	0.058	0.201
022						A20%	0.109	0.126	0.129	0.142	0.161	0.208	0.499	0.021	0.028	0.029	0.105	0.127	0.127	0.505
023						A30%	0.160	0.189	0.216	0.341	0.396	0.564	0.839	0.032	0.165	0.165	0.308	0.769	0.769	0.990
024						A40%	0.212	0.308	0.362	0.562	0.629	0.820	0.963	0.050	0.189	0.189	0.354	0.809	0.809	0.990
025				S60th	T(R)	A10%	0.000	0.020	0.040	0.089	0.122	0.132	0.250	0.009	0.015	0.065	0.096	0.096	0.163	0.380
026						A20%	0.012	0.067	0.092	0.132	0.197	0.236	0.409	0.035	0.086	0.098	0.260	0.320	0.439	0.803
027						A30%	0.076	0.162	0.203	0.242	0.294	0.382	0.674	0.035	0.055	0.089	0.193	0.195	0.326	0.691
028						A40%	0.188	0.282	0.373	0.482	0.569	0.677	0.895	0.141	0.329	0.355	0.615	0.623	0.748	0.969
029					T(H)	A10%	0.008	0.035	0.049	0.094	0.124	0.135	0.236	0.009	0.015	0.065	0.096	0.096	0.163	0.380
030						A20%	0.047	0.096	0.124	0.171	0.223	0.258	0.431	0.016	0.022	0.047	0.087	0.093	0.236	0.586
031						A30%	0.140	0.198	0.257	0.293	0.408	0.485	0.771	0.049	0.058	0.095	0.136	0.136	0.357	0.723
032						A40%	0.270	0.408	0.550	0.663	0.767	0.867	0.988	0.120	0.306	0.334	0.711	0.747	0.868	0.995

ตาราง 21 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของ ดัชนี S_1 และ ดัชนี S_2 ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 1-สถานการณ์ที่ 32)

Sit.	test length (L)	sample (N)	percentage of copier (C)	ability levels of source (S)	type of copying (T)	percentage of answer copied (A)	S1 INDICES DETECTION RATE							S2 INDICES DETECTION RATE						
							0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
001	L35	N100	C5%	S90th	T(R)	A10%	0.000	0.000	0.002	0.008	0.018	0.054	0.276	0.000	0.010	0.026	0.096	0.162	0.278	0.606
002						A20%	0.000	0.002	0.002	0.022	0.042	0.088	0.384	0.002	0.028	0.060	0.154	0.258	0.376	0.770
003						A30%	0.000	0.012	0.046	0.130	0.218	0.352	0.736	0.016	0.162	0.244	0.434	0.560	0.708	0.952
004						A40%	0.004	0.026	0.050	0.154	0.270	0.394	0.788	0.032	0.194	0.300	0.506	0.646	0.772	0.988
005					T(H)	A10%	0.000	0.000	0.000	0.010	0.018	0.056	0.284	0.000	0.012	0.030	0.112	0.172	0.288	0.660
006						A20%	0.000	0.006	0.016	0.060	0.138	0.236	0.540	0.008	0.096	0.156	0.298	0.398	0.542	0.902
007						A30%	0.000	0.028	0.082	0.238	0.462	0.658	0.982	0.048	0.332	0.510	0.760	0.906	0.968	0.994
008						A40%	0.000	0.018	0.036	0.218	0.400	0.542	0.926	0.028	0.276	0.430	0.660	0.838	0.922	0.996
009				S60th	T(R)	A10%	0.000	0.004	0.008	0.020	0.022	0.032	0.228	0.004	0.020	0.020	0.032	0.066	0.142	0.538
010						A20%	0.000	0.000	0.002	0.020	0.024	0.072	0.288	0.000	0.020	0.020	0.068	0.122	0.230	0.640
011						A30%	0.010	0.020	0.020	0.022	0.038	0.126	0.556	0.020	0.022	0.024	0.122	0.228	0.440	0.802
012						A40%	0.022	0.030	0.064	0.128	0.200	0.310	0.858	0.034	0.130	0.194	0.346	0.550	0.762	0.980
013					T(H)	A10%	0.000	0.004	0.010	0.020	0.020	0.048	0.236	0.006	0.020	0.020	0.046	0.084	0.178	0.586
014						A20%	0.000	0.004	0.014	0.022	0.048	0.102	0.442	0.010	0.022	0.040	0.108	0.168	0.356	0.802
015						A30%	0.000	0.022	0.058	0.110	0.172	0.348	0.776	0.024	0.116	0.164	0.368	0.520	0.714	0.934
016						A40%	0.020	0.078	0.226	0.392	0.584	0.728	0.970	0.094	0.380	0.540	0.732	0.826	0.914	1.000
017			C10%	S90th	T(R)	A10%	0.000	0.000	0.000	0.007	0.036	0.069	0.411	0.000	0.017	0.039	0.118	0.201	0.388	0.766
018						A20%	0.000	0.000	0.003	0.017	0.043	0.113	0.480	0.000	0.022	0.043	0.168	0.257	0.447	0.820
019						A30%	0.000	0.003	0.036	0.110	0.203	0.323	0.777	0.017	0.123	0.210	0.390	0.609	0.751	0.973
020						A40%	0.000	0.000	0.003	0.050	0.138	0.248	0.625	0.000	0.076	0.154	0.294	0.470	0.611	0.970
021					T(H)	A10%	0.000	0.000	0.000	0.010	0.012	0.050	0.247	0.000	0.010	0.013	0.077	0.123	0.250	0.626
022						A20%	0.000	0.002	0.005	0.019	0.059	0.125	0.460	0.002	0.032	0.062	0.181	0.319	0.454	0.843
023						A30%	0.000	0.000	0.035	0.110	0.329	0.576	0.924	0.005	0.143	0.352	0.631	0.793	0.896	0.989
024						A40%	0.000	0.001	0.071	0.168	0.419	0.661	0.972	0.011	0.217	0.451	0.709	0.867	0.951	0.990
025				S60th	T(R)	A10%	0.000	0.000	0.010	0.020	0.020	0.031	0.194	0.000	0.020	0.020	0.031	0.047	0.130	0.485
026						A20%	0.000	0.001	0.006	0.023	0.051	0.120	0.515	0.000	0.022	0.042	0.113	0.228	0.381	0.752
027						A30%	0.001	0.010	0.012	0.026	0.068	0.165	0.594	0.010	0.019	0.060	0.165	0.307	0.467	0.842
028						A40%	0.010	0.030	0.071	0.177	0.349	0.522	0.904	0.029	0.143	0.291	0.504	0.661	0.786	0.993
029					T(H)	A10%	0.000	0.000	0.002	0.020	0.022	0.030	0.211	0.000	0.020	0.022	0.028	0.059	0.136	0.496
030						A20%	0.000	0.000	0.011	0.020	0.028	0.071	0.401	0.000	0.019	0.028	0.077	0.168	0.297	0.766
031						A30%	0.000	0.012	0.025	0.037	0.080	0.176	0.631	0.014	0.037	0.077	0.176	0.302	0.519	0.917
032						A40%	0.000	0.037	0.096	0.262	0.468	0.647	0.968	0.039	0.243	0.429	0.635	0.789	0.905	0.995

ตาราง 22 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2 และ ดัชนี ω ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 33-สถานการณ์ที่ 64)

Sit.	test length (L)	sample (N)	percentage of copier (C)	ability levels of source (S)	type of copying (T)	percentage of answer copied (A)	OMEGA INDICES DETECTION RATE							K2 INDICES DETECTION RATE						
							0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
033	L35	N250	C5%	S90th	T(R)	A10%	0.000	0.000	0.000	0.000	0.000	0.030	0.078	0.000	0.000	0.000	0.004	0.055	0.055	0.246
034						A20%	0.000	0.002	0.016	0.048	0.069	0.103	0.195	0.000	0.005	0.005	0.005	0.121	0.121	0.522
035						A30%	0.049	0.068	0.068	0.082	0.102	0.148	0.409	0.002	0.035	0.035	0.035	0.174	0.174	0.529
036						A40%	0.084	0.121	0.160	0.226	0.292	0.432	0.779	0.013	0.129	0.129	0.129	0.470	0.470	1.000
037					T(H)	A10%	0.000	0.000	0.000	0.002	0.002	0.034	0.093	0.000	0.005	0.005	0.005	0.072	0.072	0.262
038						A20%	0.000	0.005	0.012	0.038	0.052	0.095	0.254	0.000	0.005	0.005	0.005	0.077	0.077	0.303
039						A30%	0.015	0.049	0.066	0.094	0.145	0.273	0.663	0.000	0.017	0.017	0.017	0.195	0.195	0.597
040						A40%	0.100	0.186	0.256	0.390	0.532	0.708	0.945	0.028	0.205	0.205	0.205	0.726	0.726	0.948
041				S60th	T(R)	A10%	0.000	0.000	0.000	0.003	0.007	0.008	0.084	0.000	0.000	0.004	0.008	0.008	0.067	0.227
042						A20%	0.000	0.002	0.003	0.008	0.015	0.026	0.124	0.000	0.002	0.003	0.020	0.020	0.132	0.327
043						A30%	0.020	0.042	0.086	0.198	0.341	0.446	0.801	0.058	0.208	0.210	0.506	0.572	0.633	0.948
044						A40%	0.053	0.138	0.192	0.308	0.408	0.511	0.808	0.008	0.028	0.042	0.097	0.097	0.235	0.479
045					T(H)	A10%	0.000	0.002	0.005	0.011	0.031	0.042	0.137	0.000	0.000	0.002	0.042	0.042	0.153	0.409
046						A20%	0.011	0.028	0.042	0.073	0.095	0.132	0.388	0.000	0.027	0.027	0.160	0.163	0.306	0.768
047						A30%	0.051	0.113	0.145	0.268	0.466	0.591	0.873	0.025	0.168	0.176	0.538	0.561	0.723	0.978
048						A40%	0.183	0.372	0.492	0.635	0.780	0.849	0.981	0.028	0.155	0.178	0.528	0.528	0.753	0.985
049			C10%	S90th	T(R)	A10%	0.000	0.001	0.001	0.006	0.014	0.016	0.047	0.000	0.000	0.000	0.005	0.016	0.016	0.148
050						A20%	0.000	0.001	0.006	0.019	0.032	0.048	0.099	0.000	0.000	0.000	0.006	0.032	0.032	0.188
051						A30%	0.032	0.054	0.065	0.098	0.139	0.194	0.520	0.003	0.106	0.106	0.106	0.158	0.854	0.994
052						A40%	0.040	0.072	0.094	0.144	0.186	0.268	0.590	0.000	0.007	0.007	0.007	0.106	0.112	0.519
053					T(H)	A10%	0.000	0.000	0.000	0.006	0.012	0.028	0.091	0.000	0.010	0.010	0.016	0.074	0.085	0.316
054						A20%	0.000	0.002	0.002	0.010	0.023	0.055	0.218	0.000	0.014	0.014	0.017	0.112	0.112	0.388
055						A30%	0.016	0.043	0.064	0.108	0.174	0.299	0.662	0.002	0.045	0.045	0.045	0.275	0.321	0.901
056						A40%	0.062	0.140	0.190	0.277	0.380	0.553	0.862	0.017	0.137	0.137	0.137	0.266	0.554	0.870
057				S60th	T(R)	A10%	0.000	0.002	0.002	0.006	0.008	0.014	0.082	0.000	0.000	0.000	0.002	0.005	0.030	0.158
058						A20%	0.006	0.008	0.008	0.010	0.018	0.032	0.145	0.000	0.000	0.000	0.004	0.004	0.047	0.190
059						A30%	0.013	0.024	0.039	0.086	0.147	0.213	0.490	0.005	0.019	0.020	0.066	0.069	0.137	0.409
060						A40%	0.051	0.132	0.191	0.283	0.393	0.495	0.848	0.000	0.026	0.031	0.158	0.172	0.230	0.580
061					T(H)	A10%	0.000	0.002	0.002	0.007	0.020	0.034	0.146	0.000	0.003	0.003	0.047	0.064	0.089	0.310
062						A20%	0.004	0.014	0.030	0.060	0.096	0.135	0.335	0.003	0.016	0.017	0.087	0.114	0.142	0.481
063						A30%	0.042	0.102	0.136	0.219	0.346	0.455	0.763	0.003	0.058	0.062	0.196	0.280	0.344	0.742
064						A40%	0.215	0.417	0.527	0.655	0.786	0.835	0.987	0.206	0.460	0.601	0.744	0.847	0.854	0.998

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 23 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของ ดัชนี S_1 และ ดัชนี S_2 ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 33-สถานการณ์ที่ 64)

Sit.	test length (L)	sample (N)	percentage of copier (C)	ability levels of source (S)	type of copying (T)	percentage of answer copied (A)	S1 INDICES DETECTION RATE							S2 INDICES DETECTION RATE						
							0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
033	L35	N250	C5%	S90th	T(R)	A10%	0.000	0.000	0.000	0.023	0.023	0.029	0.185	0.000	0.023	0.028	0.050	0.128	0.195	0.554
034						A20%	0.000	0.002	0.003	0.028	0.042	0.076	0.409	0.002	0.039	0.066	0.162	0.288	0.412	0.932
035						A30%	0.002	0.023	0.026	0.032	0.083	0.136	0.486	0.026	0.066	0.107	0.217	0.338	0.505	0.955
036						A40%	0.012	0.034	0.072	0.132	0.297	0.425	0.957	0.062	0.238	0.368	0.511	0.779	0.964	1.000
037					T(H)	A10%	0.000	0.000	0.000	0.023	0.028	0.041	0.208	0.000	0.032	0.044	0.071	0.141	0.214	0.575
038						A20%	0.000	0.000	0.000	0.023	0.034	0.051	0.257	0.005	0.031	0.042	0.105	0.198	0.265	0.682
039						A30%	0.000	0.000	0.002	0.046	0.098	0.160	0.558	0.000	0.071	0.136	0.268	0.484	0.571	0.965
040						A40%	0.000	0.072	0.138	0.203	0.493	0.656	0.936	0.124	0.388	0.615	0.755	0.907	0.937	0.999
041				S60th	T(R)	A10%	0.000	0.000	0.000	0.000	0.008	0.015	0.161	0.000	0.000	0.004	0.027	0.058	0.104	0.376
042						A20%	0.000	0.000	0.000	0.002	0.016	0.038	0.223	0.000	0.002	0.009	0.045	0.088	0.176	0.477
043						A30%	0.000	0.028	0.056	0.155	0.314	0.482	0.847	0.027	0.148	0.277	0.481	0.662	0.796	0.952
044						A40%	0.000	0.005	0.022	0.070	0.113	0.211	0.519	0.005	0.069	0.107	0.223	0.345	0.457	0.786
045					T(H)	A10%	0.000	0.000	0.003	0.005	0.020	0.064	0.284	0.002	0.005	0.018	0.059	0.112	0.202	0.601
046						A20%	0.000	0.005	0.020	0.058	0.102	0.197	0.687	0.005	0.045	0.094	0.205	0.356	0.575	0.891
047						A30%	0.002	0.037	0.067	0.213	0.395	0.584	0.909	0.036	0.204	0.345	0.580	0.778	0.862	0.952
048						A40%	0.010	0.083	0.154	0.341	0.526	0.753	0.938	0.095	0.334	0.476	0.751	0.862	0.912	0.972
049			C10%	S90th	T(R)	A10%	0.000	0.000	0.001	0.006	0.006	0.010	0.120	0.001	0.007	0.010	0.035	0.069	0.128	0.446
050						A20%	0.000	0.000	0.000	0.012	0.013	0.025	0.156	0.000	0.012	0.015	0.046	0.113	0.164	0.502
051						A30%	0.000	0.004	0.021	0.054	0.112	0.462	0.949	0.017	0.068	0.249	0.647	0.847	0.940	0.996
052						A40%	0.000	0.002	0.002	0.010	0.026	0.090	0.475	0.002	0.018	0.053	0.109	0.270	0.478	0.975
053					T(H)	A10%	0.000	0.002	0.002	0.014	0.023	0.049	0.260	0.003	0.023	0.038	0.086	0.174	0.267	0.654
054						A20%	0.000	0.002	0.002	0.016	0.033	0.073	0.324	0.002	0.026	0.048	0.116	0.222	0.336	0.751
055						A30%	0.000	0.004	0.007	0.034	0.082	0.230	0.816	0.008	0.053	0.145	0.290	0.507	0.816	0.992
056						A40%	0.000	0.014	0.031	0.119	0.159	0.418	0.838	0.025	0.139	0.282	0.509	0.676	0.841	0.994
057				S60th	T(R)	A10%	0.000	0.000	0.000	0.000	0.000	0.005	0.118	0.000	0.000	0.000	0.014	0.043	0.098	0.318
058						A20%	0.000	0.000	0.000	0.000	0.004	0.014	0.148	0.000	0.000	0.004	0.022	0.069	0.123	0.378
059						A30%	0.000	0.000	0.001	0.020	0.055	0.096	0.392	0.000	0.017	0.048	0.100	0.192	0.321	0.716
060						A40%	0.000	0.001	0.006	0.051	0.125	0.244	0.642	0.001	0.046	0.103	0.248	0.384	0.542	0.885
061					T(H)	A10%	0.000	0.000	0.000	0.008	0.018	0.048	0.275	0.000	0.007	0.016	0.047	0.102	0.200	0.550
062						A20%	0.000	0.001	0.004	0.023	0.055	0.115	0.496	0.002	0.022	0.047	0.118	0.222	0.383	0.800
063						A30%	0.000	0.011	0.022	0.080	0.166	0.292	0.768	0.012	0.072	0.142	0.291	0.486	0.659	0.941
064						A40%	0.013	0.094	0.183	0.395	0.604	0.790	0.968	0.090	0.354	0.533	0.772	0.881	0.934	0.978

ตาราง 24 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2 และ ดัชนี ω ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 65-สถานการณ์ที่ 96)

Sit.	test length (L)	sample (N)	percentage of copier (C)	ability levels of source (S)	type of copying (T)	percentage of answer copied (A)	OMEGA INDICES DETECTION RATE							K2 INDICES DETECTION RATE						
							0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
065	L65	N100	C5%	S90th	T(R)	A10%	0.138	0.182	0.182	0.200	0.200	0.240	0.342	0.000	0.000	0.000	0.000	0.000	0.000	0.324
066						A20%	0.200	0.200	0.200	0.226	0.258	0.298	0.584	0.000	0.000	0.000	0.000	0.000	0.008	0.312
067						A30%	0.200	0.260	0.436	0.558	0.630	0.640	0.780	0.012	0.184	0.248	0.296	0.544	0.544	0.868
068						A40%	0.300	0.308	0.446	0.664	0.866	0.940	0.986	0.000	0.050	0.054	0.082	0.206	0.206	0.780
069					T(H)	A10%	0.156	0.194	0.194	0.210	0.240	0.280	0.500	0.000	0.000	0.000	0.014	0.020	0.032	0.324
070						A20%	0.226	0.284	0.302	0.332	0.422	0.444	0.838	0.012	0.078	0.090	0.202	0.300	0.300	0.630
071						A30%	0.354	0.442	0.476	0.578	0.756	0.820	1.000	0.162	0.272	0.328	0.408	0.594	0.596	0.900
072						A40%	0.508	0.674	0.838	0.898	0.984	1.000	1.000	0.416	0.556	0.696	0.708	0.928	0.928	1.000
073				S60th	T(R)	A10%	0.316	0.532	0.578	0.644	0.718	0.812	0.860	0.000	0.000	0.000	0.000	0.006	0.006	0.090
074						A20%	0.680	0.700	0.752	0.822	0.846	0.864	0.886	0.000	0.000	0.000	0.020	0.094	0.184	0.636
075						A30%	0.862	0.874	0.884	0.942	0.966	0.996	1.000	0.302	0.554	0.564	0.736	0.750	0.892	0.994
076						A40%	0.868	0.932	0.946	0.982	0.982	0.994	1.000	0.414	0.550	0.554	0.774	0.796	0.918	0.990
077					T(H)	A10%	0.440	0.496	0.598	0.600	0.624	0.776	0.880	0.000	0.000	0.000	0.004	0.052	0.056	0.456
078						A20%	0.742	0.856	0.874	0.892	0.892	0.900	0.946	0.102	0.248	0.254	0.440	0.502	0.676	0.910
079						A30%	0.858	0.880	0.900	0.924	0.930	0.964	0.994	0.040	0.208	0.208	0.438	0.476	0.684	0.938
080						A40%	0.944	0.994	0.994	1.000	1.000	1.000	1.000	0.746	0.948	0.998	1.000	1.000	1.000	1.000
081			C10%	S90th	T(R)	A10%	0.235	0.286	0.295	0.308	0.383	0.433	0.535	0.000	0.000	0.000	0.025	0.060	0.060	0.311
082						A20%	0.334	0.356	0.392	0.436	0.498	0.502	0.695	0.010	0.022	0.037	0.057	0.277	0.277	0.458
083						A30%	0.363	0.427	0.478	0.575	0.743	0.789	0.902	0.110	0.143	0.317	0.332	0.527	0.605	0.834
084						A40%	0.549	0.717	0.848	0.880	0.921	0.971	0.990	0.363	0.650	0.780	0.957	0.964	0.989	0.990
085					T(H)	A10%	0.268	0.314	0.365	0.399	0.453	0.479	0.711	0.000	0.022	0.070	0.079	0.255	0.255	0.418
086						A20%	0.339	0.410	0.482	0.534	0.581	0.592	0.745	0.060	0.087	0.163	0.165	0.378	0.394	0.630
087						A30%	0.497	0.535	0.566	0.617	0.714	0.765	0.894	0.122	0.156	0.389	0.411	0.576	0.619	0.816
088						A40%	0.613	0.706	0.752	0.839	0.917	0.957	0.981	0.396	0.601	0.697	0.813	0.836	0.973	0.983
089				S60th	T(R)	A10%	0.377	0.481	0.519	0.582	0.593	0.606	0.675	0.000	0.004	0.004	0.020	0.034	0.043	0.193
090						A20%	0.537	0.570	0.607	0.636	0.709	0.795	0.898	0.016	0.049	0.050	0.092	0.110	0.322	0.482
091						A30%	0.594	0.756	0.820	0.855	0.894	0.903	0.936	0.040	0.085	0.135	0.217	0.243	0.414	0.726
092						A40%	0.841	0.892	0.913	0.942	0.958	0.984	1.000	0.061	0.200	0.338	0.421	0.508	0.640	0.865
093					T(H)	A10%	0.385	0.497	0.500	0.521	0.531	0.569	0.719	0.000	0.000	0.000	0.008	0.033	0.048	0.179
094						A20%	0.502	0.544	0.606	0.723	0.791	0.834	0.892	0.004	0.038	0.049	0.084	0.109	0.216	0.440
095						A30%	0.752	0.808	0.857	0.896	0.924	0.943	0.960	0.067	0.149	0.230	0.336	0.428	0.578	0.806
096						A40%	0.915	0.946	0.958	0.960	0.960	0.976	1.000	0.405	0.741	0.806	0.889	0.952	0.982	1.000

ตาราง 25 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของ ดัชนี S_1 และ ดัชนี S_2 ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 65-สถานการณ์ที่ 96)

Sit.	test length (L)	sample (N)	percentage of copier (C)	ability levels of source (S)	type of copying (T)	percentage of answer copied (A)	S1 INDICES DETECTION RATE							S2 INDICES DETECTION RATE						
							0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
065	L65	N100	C5%	S90th	T(R)	A10%	0.000	0.000	0.000	0.000	0.000	0.000	0.060	0.000	0.000	0.000	0.000	0.000	0.000	0.324
066						A20%	0.000	0.000	0.000	0.000	0.000	0.000	0.030	0.000	0.000	0.000	0.000	0.000	0.000	0.312
067						A30%	0.000	0.000	0.000	0.012	0.144	0.248	0.784	0.000	0.004	0.012	0.236	0.374	0.544	0.946
068						A40%	0.000	0.000	0.000	0.000	0.022	0.054	0.588	0.000	0.000	0.000	0.054	0.106	0.202	0.854
069					T(H)	A10%	0.000	0.000	0.000	0.000	0.000	0.000	0.188	0.000	0.000	0.000	0.000	0.020	0.020	0.328
070						A20%	0.000	0.000	0.000	0.012	0.046	0.090	0.462	0.000	0.008	0.012	0.090	0.228	0.300	0.668
071						A30%	0.000	0.006	0.014	0.146	0.190	0.328	0.784	0.000	0.062	0.162	0.318	0.436	0.596	0.926
072						A40%	0.000	0.034	0.142	0.360	0.446	0.692	0.974	0.030	0.184	0.406	0.630	0.736	0.928	1.000
073				S60th	T(R)	A10%	0.000	0.000	0.000	0.000	0.000	0.002	0.040	0.000	0.000	0.000	0.000	0.002	0.006	0.148
074						A20%	0.000	0.000	0.000	0.000	0.020	0.082	0.304	0.000	0.000	0.000	0.078	0.098	0.176	0.468
075						A30%	0.004	0.066	0.174	0.316	0.412	0.588	0.920	0.048	0.228	0.344	0.480	0.612	0.774	0.980
076						A40%	0.018	0.220	0.266	0.378	0.512	0.730	0.982	0.140	0.292	0.406	0.562	0.790	0.920	0.996
077					T(H)	A10%	0.000	0.000	0.000	0.000	0.000	0.010	0.126	0.000	0.000	0.000	0.000	0.016	0.018	0.334
078						A20%	0.002	0.016	0.024	0.134	0.216	0.280	0.510	0.014	0.034	0.150	0.246	0.310	0.374	0.754
079						A30%	0.000	0.002	0.034	0.134	0.270	0.350	0.824	0.000	0.050	0.168	0.306	0.378	0.528	0.990
080						A40%	0.096	0.332	0.412	0.652	0.830	0.936	1.000	0.250	0.520	0.702	0.884	0.962	0.998	1.000
081			C10%	S90th	T(R)	A10%	0.000	0.000	0.000	0.000	0.000	0.000	0.200	0.000	0.000	0.000	0.000	0.050	0.060	0.324
082						A20%	0.000	0.000	0.000	0.008	0.019	0.037	0.381	0.000	0.000	0.010	0.033	0.100	0.277	0.625
083						A30%	0.000	0.005	0.017	0.063	0.138	0.321	0.742	0.002	0.020	0.116	0.209	0.332	0.592	0.910
084						A40%	0.006	0.080	0.142	0.320	0.620	0.812	0.989	0.040	0.288	0.390	0.668	0.937	0.979	0.990
085					T(H)	A10%	0.000	0.000	0.000	0.000	0.004	0.070	0.394	0.000	0.000	0.000	0.069	0.110	0.255	0.556
086						A20%	0.000	0.000	0.009	0.051	0.063	0.164	0.597	0.000	0.010	0.060	0.152	0.168	0.405	0.825
087						A30%	0.000	0.020	0.057	0.080	0.145	0.399	0.700	0.018	0.062	0.124	0.239	0.426	0.601	0.899
088						A40%	0.000	0.092	0.145	0.318	0.565	0.694	0.974	0.044	0.276	0.393	0.618	0.778	0.866	0.988
089				S60th	T(R)	A10%	0.000	0.000	0.000	0.000	0.000	0.001	0.095	0.000	0.000	0.000	0.000	0.002	0.023	0.177
090						A20%	0.000	0.000	0.000	0.000	0.009	0.049	0.340	0.000	0.000	0.000	0.011	0.088	0.187	0.486
091						A30%	0.000	0.004	0.009	0.022	0.059	0.123	0.613	0.000	0.018	0.028	0.075	0.170	0.354	0.873
092						A40%	0.000	0.012	0.054	0.200	0.309	0.456	0.903	0.003	0.087	0.228	0.350	0.517	0.703	0.979
093					T(H)	A10%	0.000	0.000	0.000	0.000	0.000	0.002	0.061	0.000	0.000	0.000	0.002	0.007	0.015	0.171
094						A20%	0.000	0.000	0.000	0.002	0.006	0.044	0.348	0.000	0.000	0.002	0.011	0.058	0.140	0.512
095						A30%	0.000	0.001	0.018	0.078	0.197	0.322	0.746	0.000	0.036	0.088	0.248	0.364	0.512	0.918
096						A40%	0.006	0.100	0.280	0.507	0.684	0.825	0.997	0.044	0.366	0.547	0.744	0.869	0.952	1.000

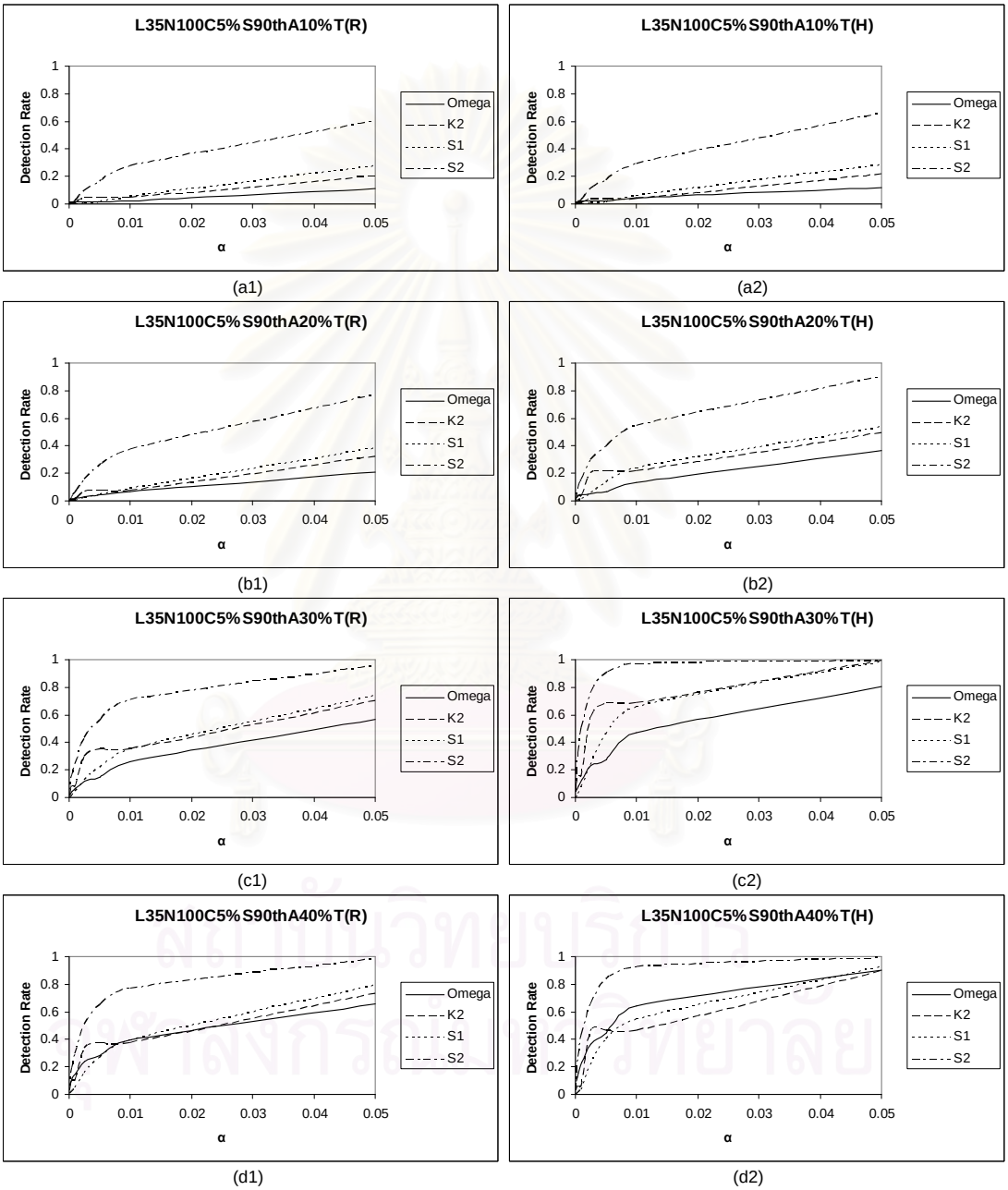
ตาราง 26 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2 และ ดัชนี ω ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 97-สถานการณ์ที่ 128)

Sit.	test length (L)	sample (N)	percentage of copier (C)	ability levels of source (S)	type of copying (T)	percentage of answer copied (A)	OMEGA INDICES							K2 INDICES						
							DETECTION RATE							DETECTION RATE						
							0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
097	L65	N250	C5%	S90th	T(R)	A10%	0.085	0.085	0.085	0.093	0.112	0.137	0.332	0.000	0.015	0.015	0.046	0.046	0.046	0.238
098						A20%	0.098	0.128	0.161	0.221	0.297	0.384	0.709	0.008	0.051	0.051	0.115	0.144	0.144	0.417
099						A30%	0.158	0.295	0.353	0.475	0.574	0.685	0.938	0.010	0.085	0.085	0.159	0.255	0.255	0.542
100						A40%	0.310	0.503	0.588	0.719	0.822	0.902	0.997	0.011	0.087	0.087	0.162	0.261	0.261	0.523
101					T(H)	A10%	0.082	0.090	0.094	0.102	0.116	0.136	0.305	0.000	0.002	0.002	0.029	0.029	0.029	0.155
102						A20%	0.102	0.182	0.202	0.238	0.294	0.369	0.727	0.005	0.053	0.053	0.112	0.139	0.139	0.434
103						A30%	0.220	0.284	0.335	0.428	0.527	0.703	0.922	0.029	0.059	0.059	0.178	0.220	0.220	0.541
104						A40%	0.335	0.460	0.528	0.719	0.823	0.900	0.976	0.022	0.062	0.062	0.174	0.182	0.182	0.526
105				S60th	T(R)	A10%	0.073	0.077	0.078	0.088	0.112	0.151	0.281	0.000	0.002	0.002	0.007	0.007	0.038	0.143
106						A20%	0.088	0.136	0.175	0.225	0.289	0.370	0.557	0.002	0.043	0.043	0.104	0.104	0.218	0.485
107						A30%	0.241	0.381	0.418	0.511	0.576	0.665	0.872	0.049	0.113	0.113	0.229	0.230	0.448	0.646
108						A40%	0.477	0.598	0.638	0.732	0.810	0.857	0.923	0.155	0.325	0.325	0.518	0.531	0.679	0.855
109					T(H)	A10%	0.065	0.077	0.082	0.087	0.112	0.161	0.312	0.002	0.006	0.006	0.036	0.037	0.112	0.277
110						A20%	0.111	0.187	0.235	0.283	0.318	0.409	0.709	0.019	0.081	0.088	0.225	0.231	0.459	0.712
111						A30%	0.277	0.362	0.404	0.567	0.693	0.798	0.912	0.095	0.240	0.245	0.452	0.487	0.690	0.873
112						A40%	0.640	0.784	0.827	0.880	0.901	0.912	0.923	0.505	0.706	0.816	0.864	0.906	0.909	0.931
113			C10%	S90th	T(R)	A10%	0.060	0.062	0.062	0.072	0.088	0.123	0.310	0.003	0.020	0.020	0.050	0.071	0.071	0.198
114						A20%	0.068	0.091	0.117	0.182	0.259	0.334	0.632	0.012	0.050	0.051	0.070	0.167	0.167	0.422
115						A30%	0.131	0.250	0.307	0.437	0.537	0.639	0.919	0.014	0.070	0.086	0.102	0.268	0.268	0.548
116						A40%	0.286	0.470	0.547	0.674	0.782	0.863	0.985	0.014	0.080	0.086	0.106	0.268	0.268	0.548
117					T(H)	A10%	0.060	0.063	0.066	0.081	0.101	0.132	0.330	0.000	0.011	0.011	0.039	0.048	0.053	0.183
118						A20%	0.085	0.120	0.136	0.196	0.235	0.287	0.534	0.000	0.035	0.035	0.062	0.088	0.088	0.246
119						A30%	0.202	0.271	0.341	0.440	0.558	0.671	0.898	0.036	0.090	0.094	0.108	0.306	0.306	0.641
120						A40%	0.318	0.455	0.503	0.666	0.778	0.864	0.968	0.026	0.071	0.071	0.084	0.218	0.218	0.588
121				S60th	T(R)	A10%	0.040	0.042	0.047	0.068	0.082	0.108	0.304	0.000	0.000	0.000	0.005	0.005	0.058	0.159
122						A20%	0.076	0.140	0.171	0.230	0.270	0.332	0.592	0.000	0.010	0.010	0.053	0.053	0.148	0.290
123						A30%	0.271	0.406	0.451	0.548	0.631	0.738	0.926	0.063	0.104	0.145	0.192	0.286	0.327	0.677
124						A40%	0.503	0.655	0.721	0.814	0.890	0.936	0.987	0.136	0.192	0.361	0.370	0.588	0.597	0.909
125					T(H)	A10%	0.040	0.040	0.052	0.076	0.104	0.166	0.383	0.000	0.002	0.002	0.021	0.022	0.145	0.270
126						A20%	0.062	0.150	0.204	0.291	0.348	0.441	0.748	0.017	0.038	0.065	0.114	0.184	0.257	0.611
127						A30%	0.296	0.456	0.508	0.652	0.764	0.838	0.952	0.070	0.118	0.176	0.216	0.361	0.394	0.849
128						A40%	0.702	0.856	0.898	0.945	0.960	0.974	0.986	0.242	0.537	0.804	0.846	0.963	0.969	0.988

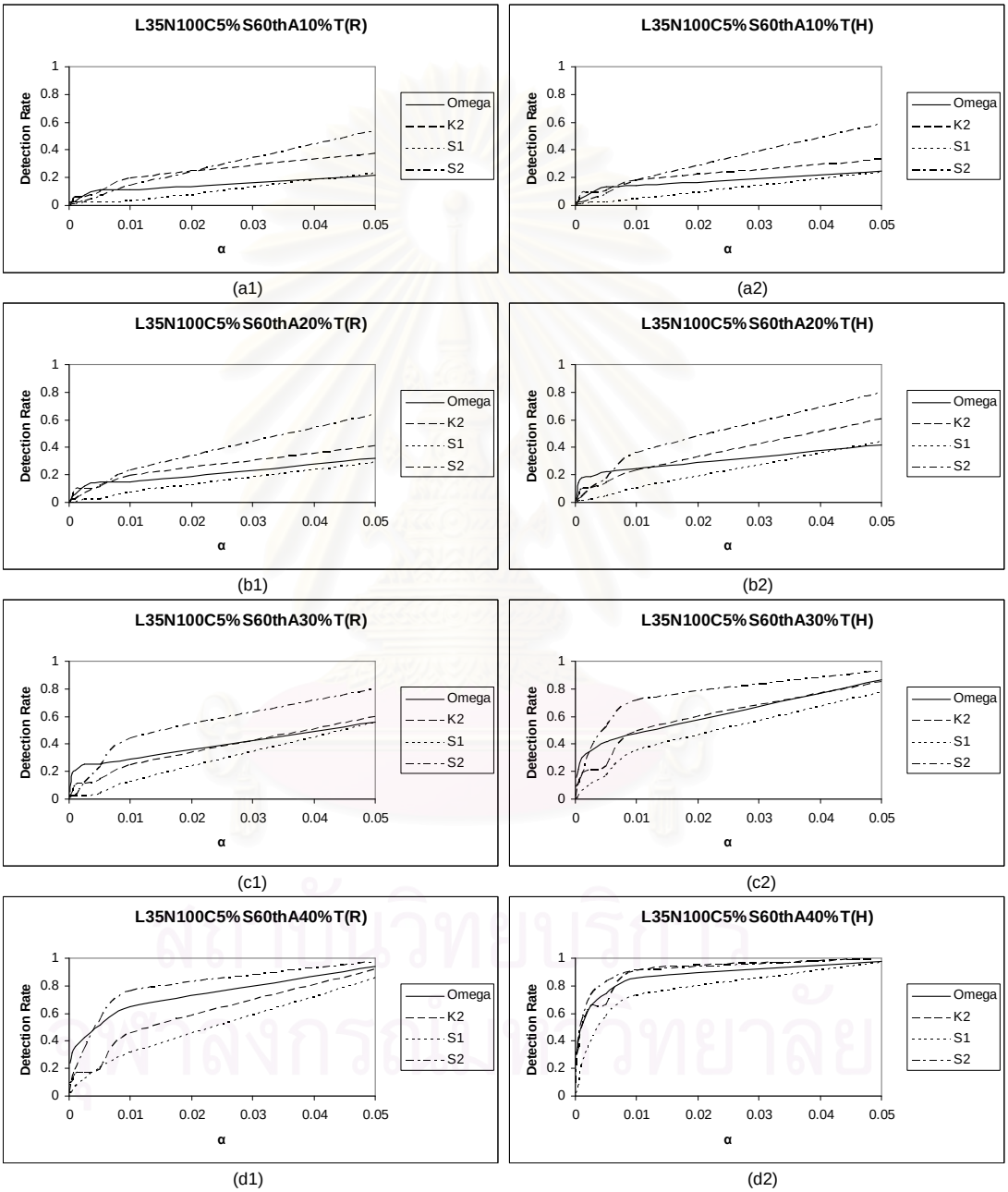
ตาราง 27 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของ ดัชนี S_1 และ ดัชนี S_2 ณ ระดับนัยสำคัญ 7 ระดับ แยกตามสถานการณ์ (สถานการณ์ที่ 97-สถานการณ์ที่ 128)

Sit.	test length (L)	sample (N)	percentage of copier (C)	ability levels of source (S)	type of copying (T)	percentage of answer copied (A)	S1 INDICES DETECTION RATE							S2 INDICES DETECTION RATE						
							0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05	0.0001	0.0005	0.001	0.0025	0.005	0.01	0.05
097	L65	N250	C5%	S90th	T(R)	A10%	0.000	0.000	0.000	0.000	0.008	0.035	0.192	0.000	0.000	0.000	0.031	0.054	0.114	0.513
098						A20%	0.000	0.002	0.002	0.013	0.037	0.091	0.491	0.002	0.014	0.028	0.079	0.134	0.296	0.782
099						A30%	0.000	0.001	0.005	0.030	0.065	0.188	0.613	0.001	0.028	0.041	0.163	0.251	0.430	0.855
100						A40%	0.000	0.002	0.008	0.035	0.075	0.205	0.637	0.002	0.022	0.046	0.178	0.285	0.459	0.876
101					T(H)	A10%	0.000	0.000	0.000	0.000	0.001	0.005	0.112	0.000	0.000	0.000	0.002	0.018	0.057	0.268
102						A20%	0.000	0.000	0.001	0.003	0.028	0.075	0.470	0.000	0.003	0.025	0.058	0.142	0.329	0.790
103						A30%	0.000	0.000	0.015	0.029	0.062	0.192	0.655	0.000	0.022	0.047	0.155	0.279	0.471	0.935
104						A40%	0.000	0.000	0.015	0.042	0.065	0.166	0.703	0.001	0.026	0.054	0.160	0.228	0.461	0.953
105				S60th	T(R)	A10%	0.000	0.000	0.000	0.000	0.000	0.003	0.066	0.000	0.000	0.000	0.002	0.008	0.015	0.254
106						A20%	0.000	0.000	0.000	0.002	0.022	0.055	0.338	0.000	0.004	0.012	0.048	0.109	0.193	0.614
107						A30%	0.000	0.008	0.022	0.058	0.139	0.235	0.587	0.005	0.050	0.083	0.190	0.315	0.456	0.824
108						A40%	0.002	0.074	0.115	0.245	0.363	0.525	0.833	0.052	0.186	0.305	0.458	0.603	0.732	0.915
109					T(H)	A10%	0.000	0.000	0.000	0.000	0.004	0.012	0.152	0.000	0.000	0.002	0.006	0.020	0.064	0.425
110						A20%	0.000	0.000	0.000	0.012	0.055	0.142	0.573	0.000	0.003	0.023	0.119	0.196	0.383	0.864
111						A30%	0.000	0.013	0.044	0.120	0.209	0.421	0.848	0.013	0.082	0.158	0.344	0.527	0.734	0.923
112						A40%	0.008	0.135	0.318	0.634	0.785	0.893	0.915	0.102	0.519	0.712	0.870	0.910	0.915	0.918
113			C10%	S90th	T(R)	A10%	0.000	0.000	0.000	0.001	0.008	0.027	0.180	0.000	0.001	0.003	0.032	0.050	0.114	0.440
114						A20%	0.000	0.000	0.002	0.009	0.026	0.085	0.412	0.000	0.006	0.021	0.068	0.114	0.264	0.742
115						A30%	0.000	0.001	0.002	0.018	0.045	0.150	0.543	0.000	0.010	0.035	0.106	0.215	0.394	0.824
116						A40%	0.000	0.002	0.004	0.018	0.057	0.158	0.546	0.001	0.011	0.037	0.113	0.238	0.399	0.832
117					T(H)	A10%	0.000	0.000	0.000	0.000	0.007	0.010	0.148	0.000	0.000	0.001	0.012	0.028	0.086	0.361
118						A20%	0.000	0.000	0.000	0.000	0.013	0.061	0.254	0.000	0.001	0.002	0.045	0.078	0.150	0.578
119						A30%	0.000	0.000	0.009	0.039	0.071	0.188	0.612	0.000	0.024	0.051	0.136	0.280	0.484	0.904
120						A40%	0.000	0.000	0.012	0.028	0.070	0.168	0.563	0.000	0.030	0.049	0.123	0.219	0.451	0.887
121				S60th	T(R)	A10%	0.000	0.000	0.000	0.000	0.000	0.002	0.072	0.000	0.000	0.000	0.002	0.004	0.031	0.262
122						A20%	0.000	0.000	0.000	0.000	0.004	0.026	0.232	0.000	0.000	0.001	0.017	0.047	0.132	0.408
123						A30%	0.000	0.010	0.024	0.059	0.128	0.248	0.602	0.008	0.046	0.082	0.178	0.308	0.469	0.825
124						A40%	0.000	0.046	0.086	0.193	0.345	0.508	0.867	0.037	0.138	0.235	0.418	0.584	0.736	0.960
125					T(H)	A10%	0.000	0.000	0.000	0.000	0.001	0.008	0.190	0.000	0.000	0.000	0.002	0.020	0.067	0.338
126						A20%	0.000	0.000	0.000	0.007	0.032	0.094	0.476	0.000	0.003	0.011	0.060	0.154	0.295	0.732
127						A30%	0.000	0.000	0.014	0.079	0.166	0.289	0.765	0.000	0.043	0.111	0.230	0.357	0.546	0.937
128						A40%	0.000	0.079	0.206	0.514	0.736	0.888	0.981	0.047	0.358	0.601	0.804	0.934	0.972	0.988

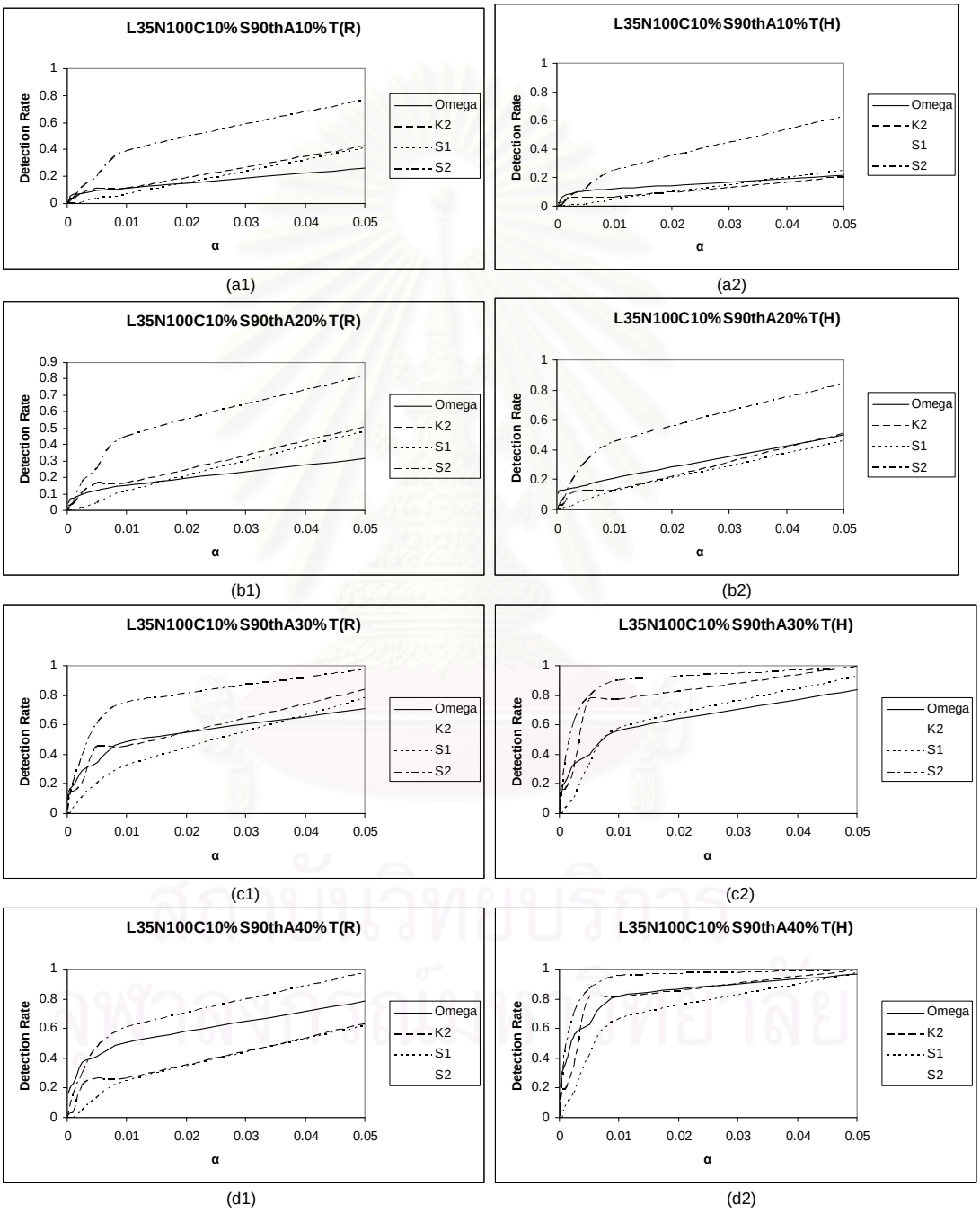
ภาพ 9 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



ภาพ 10 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



ภาพ 11 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



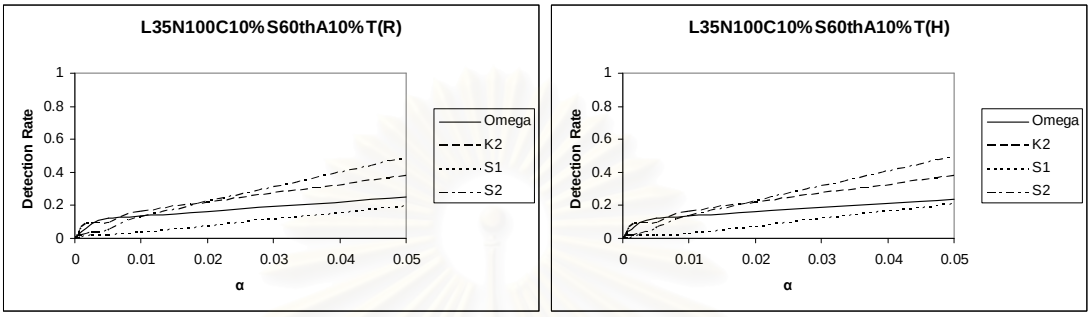
(a1) (a2)

(b1) (b2)

(c1) (c2)

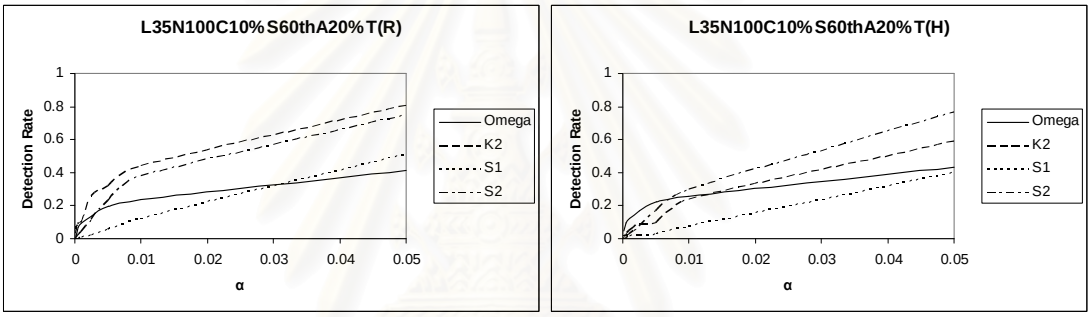
(d1) (d2)

ภาพ 12 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



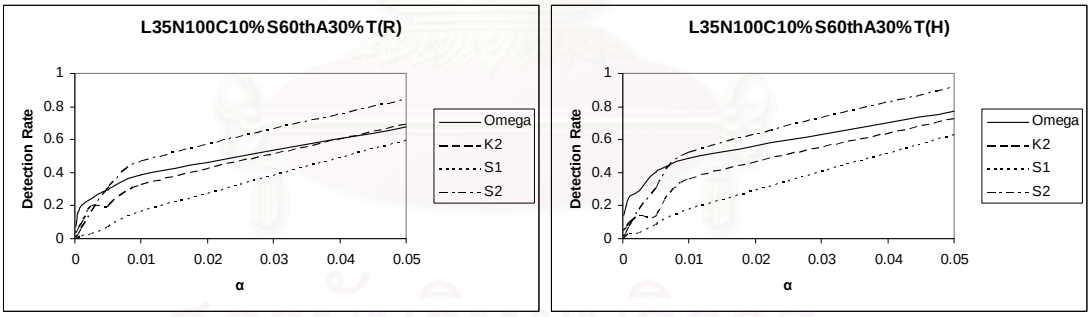
(a1)

(a2)



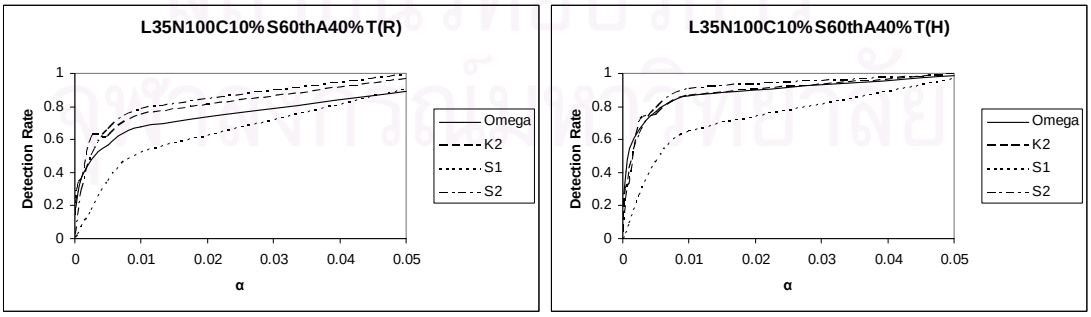
(b1)

(b2)



(c1)

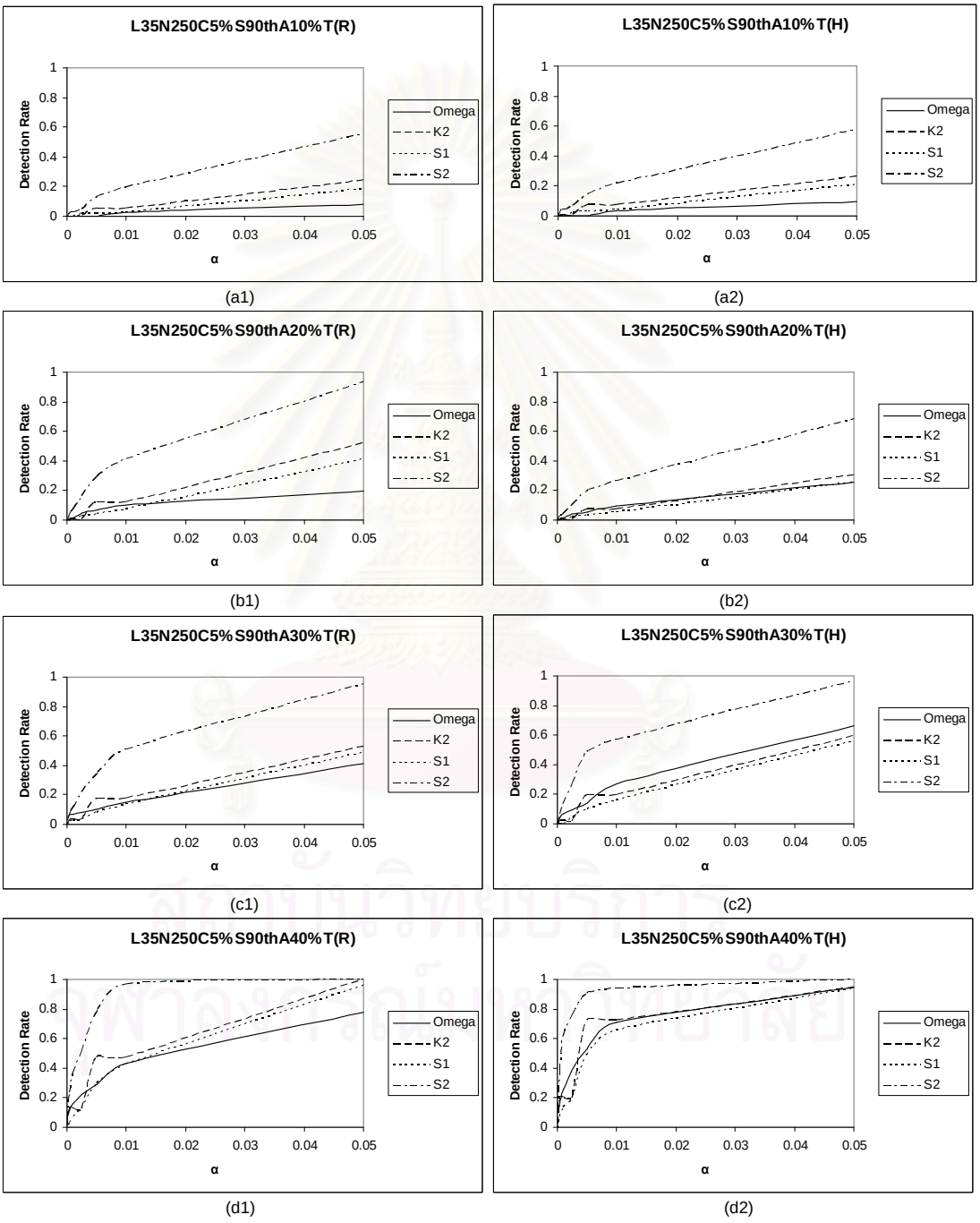
(c2)



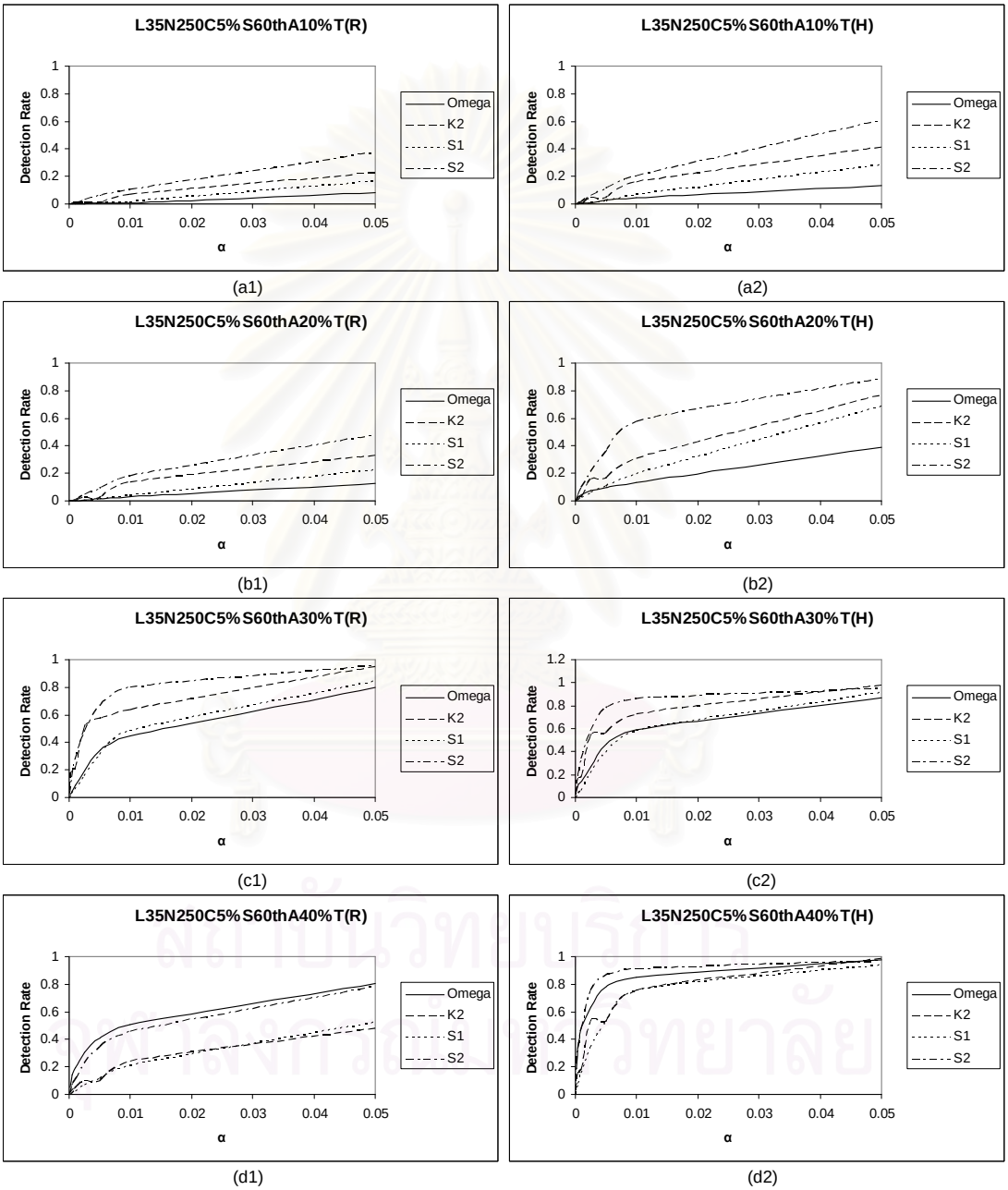
(d1)

(d2)

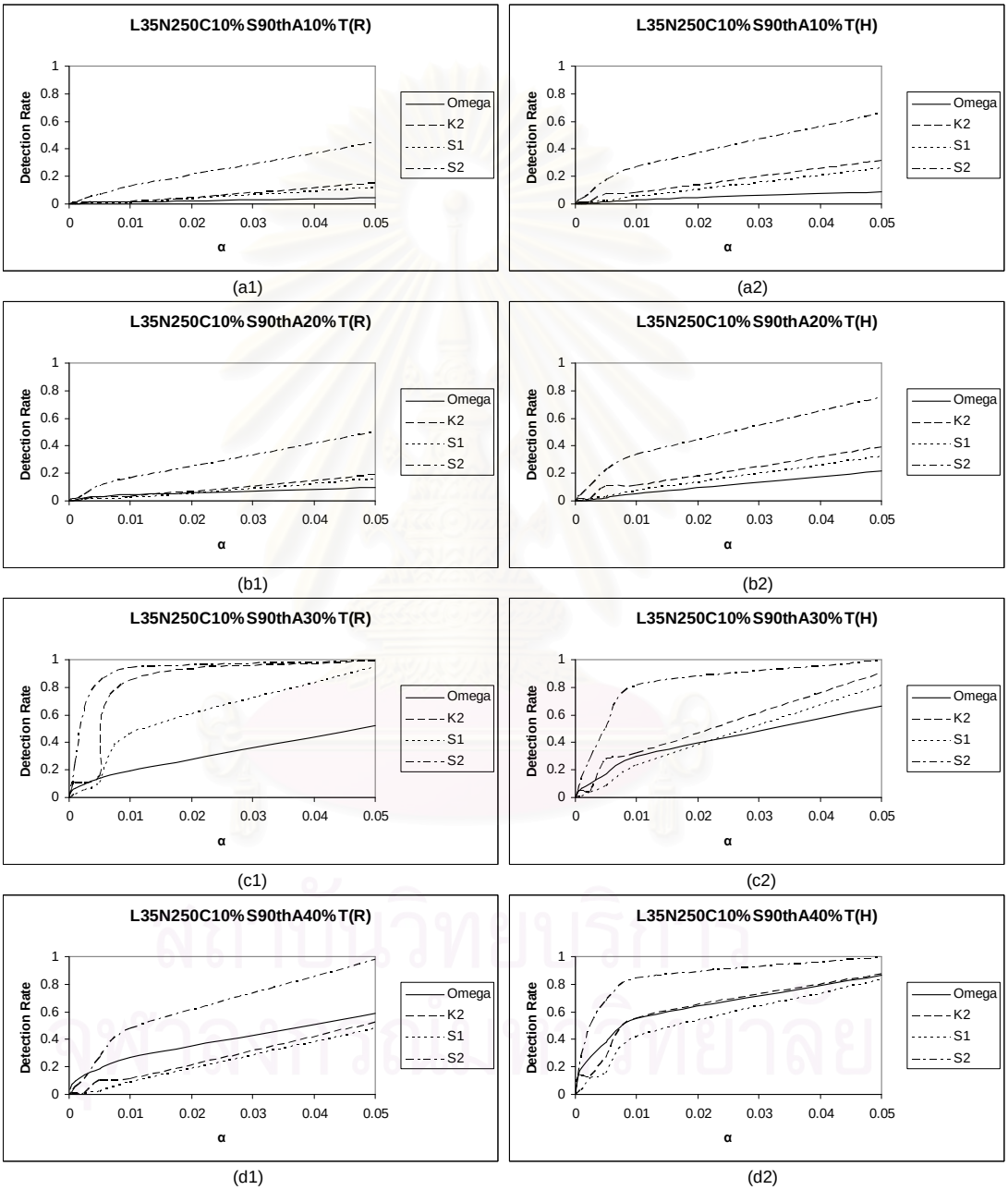
ภาพ 13 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



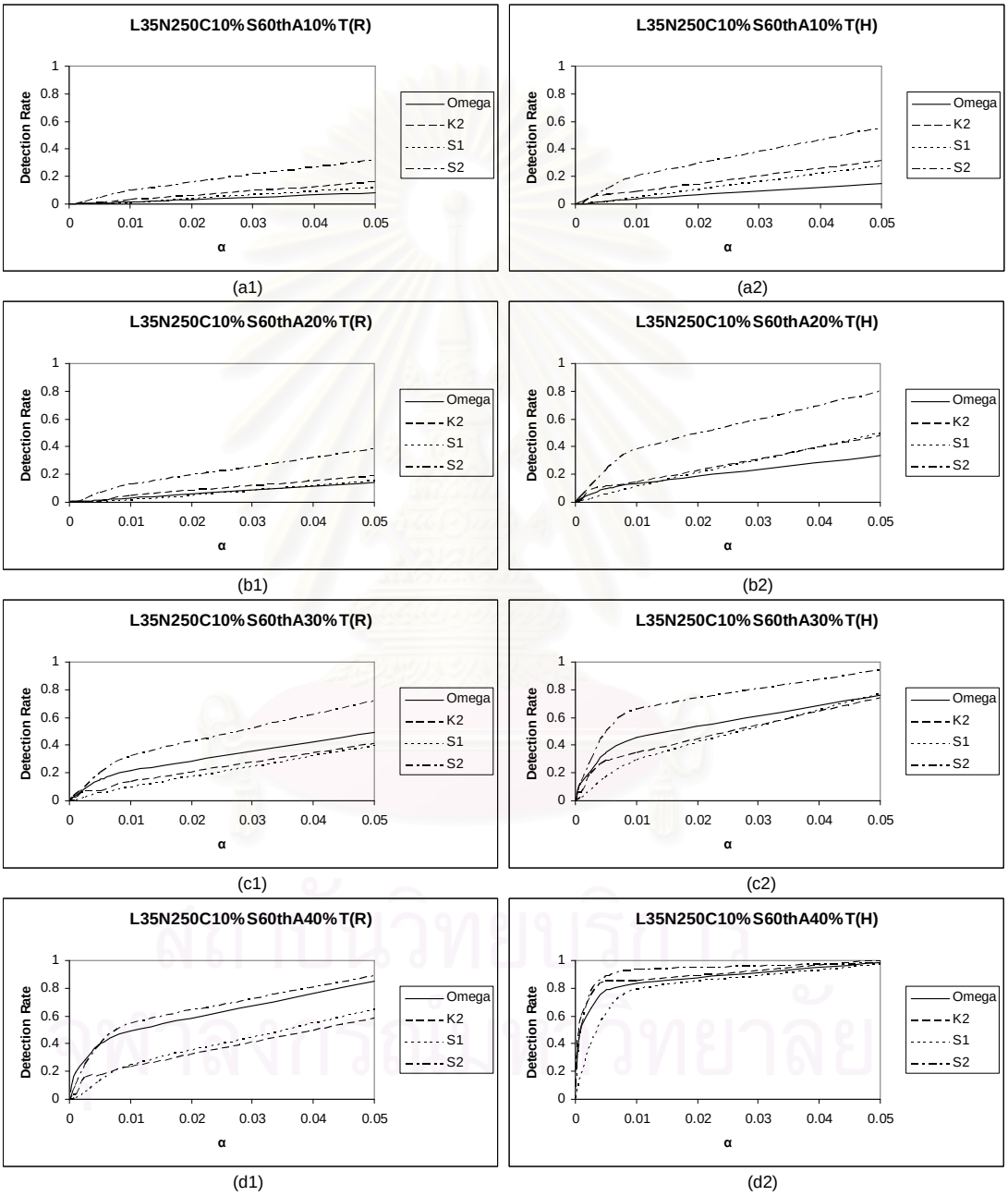
ภาพ 14 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



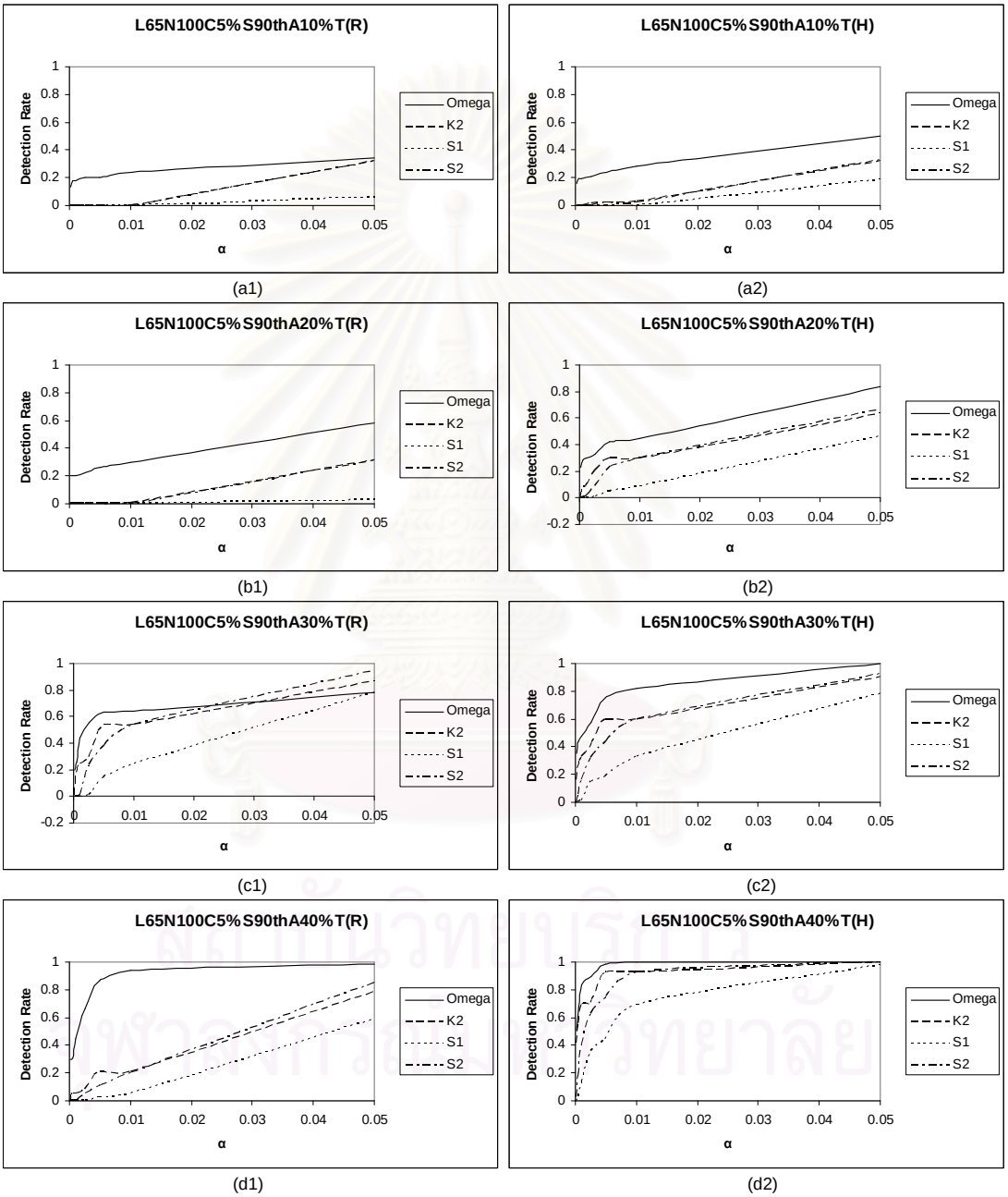
ภาพ 15 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



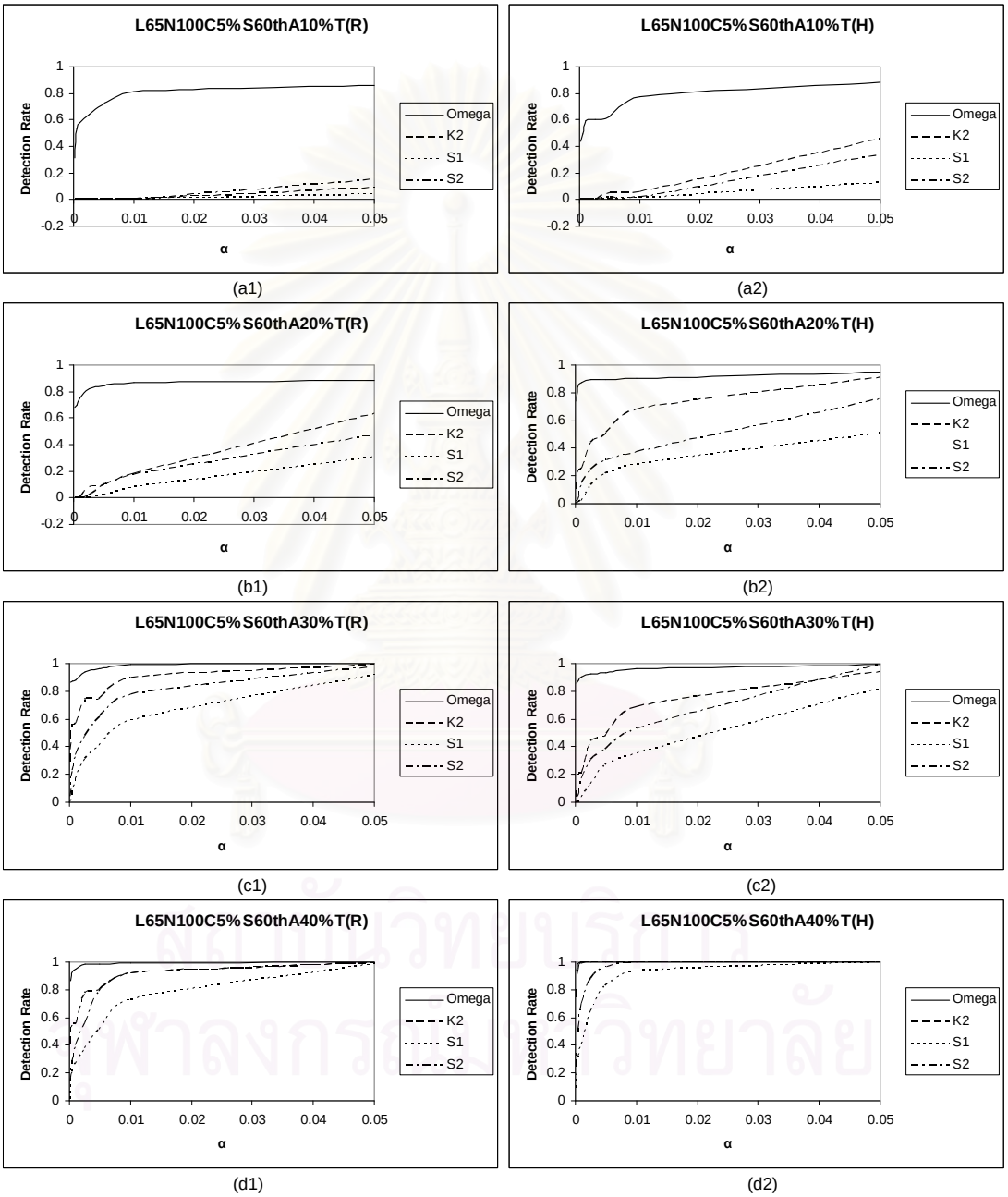
ภาพ 16 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 35 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



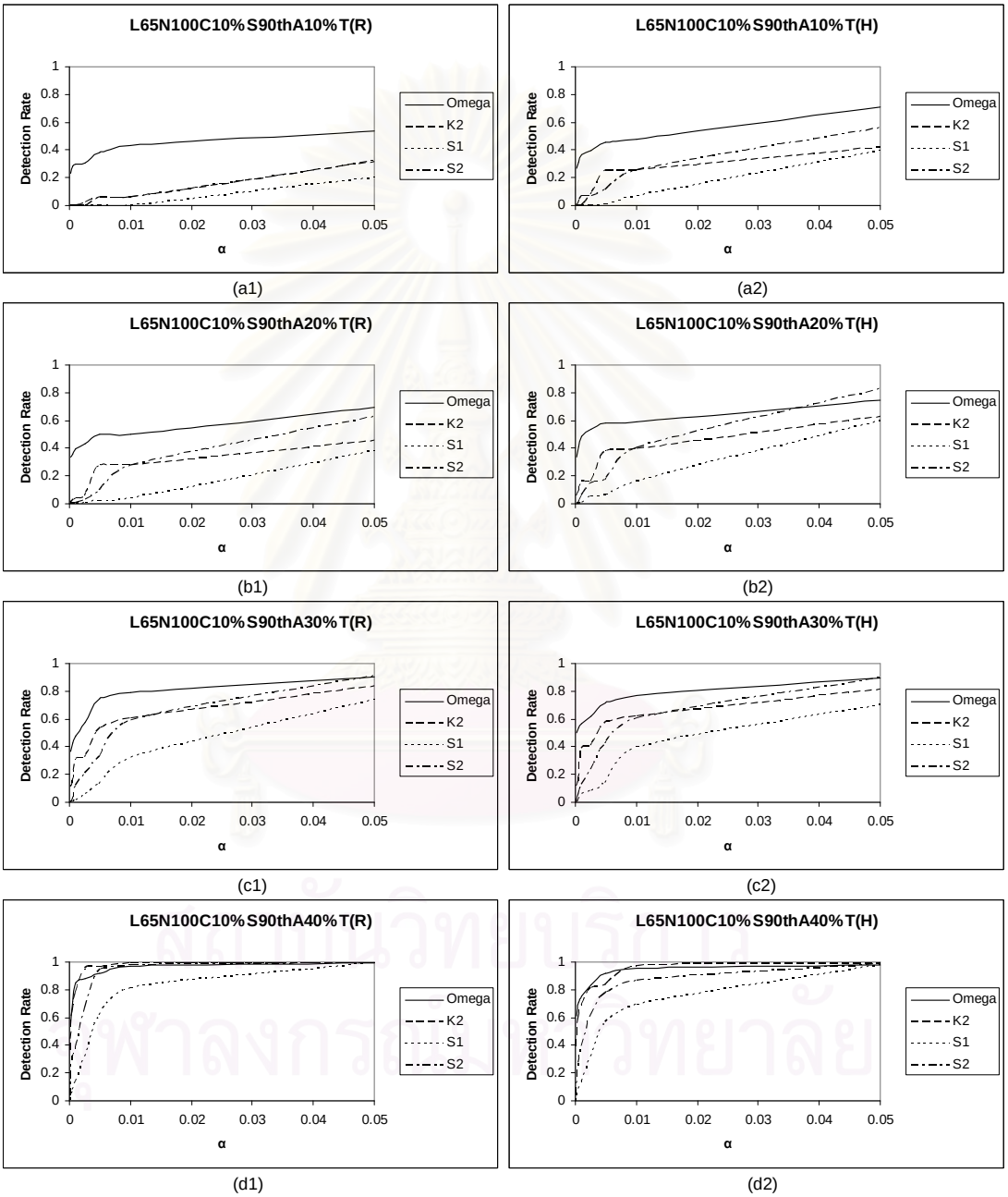
ภาพ 17 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



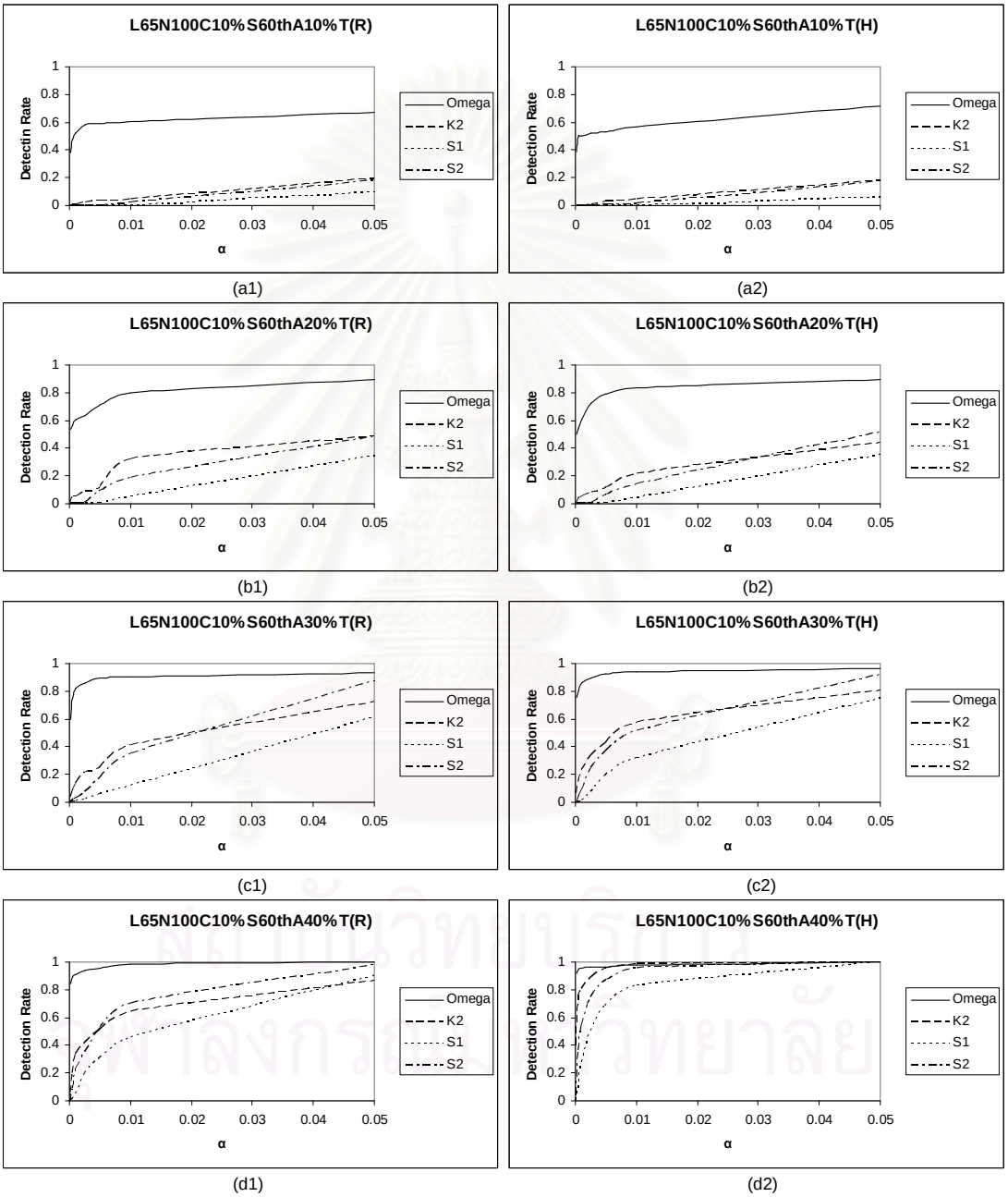
ภาพ 18 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



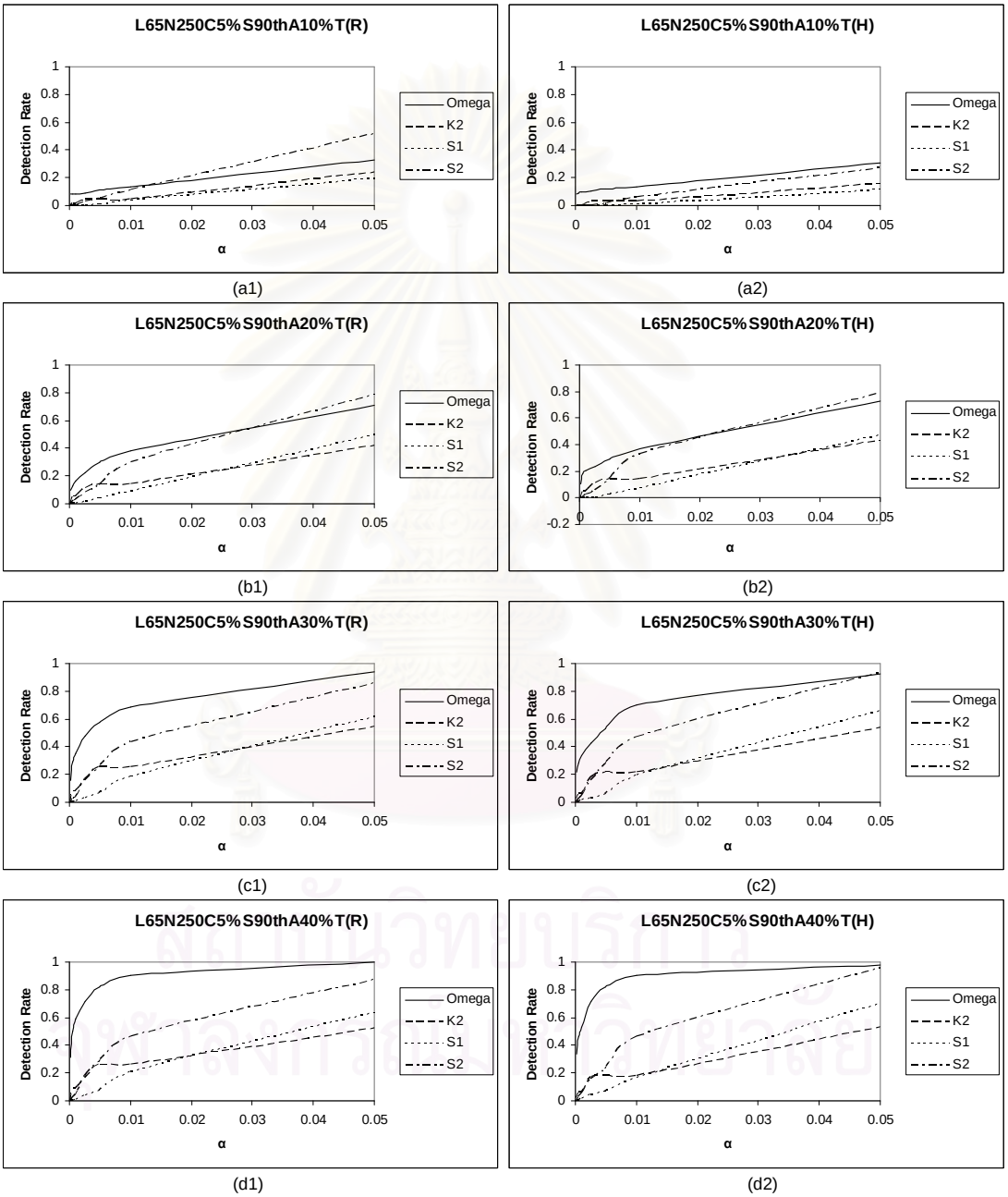
ภาพ 19 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



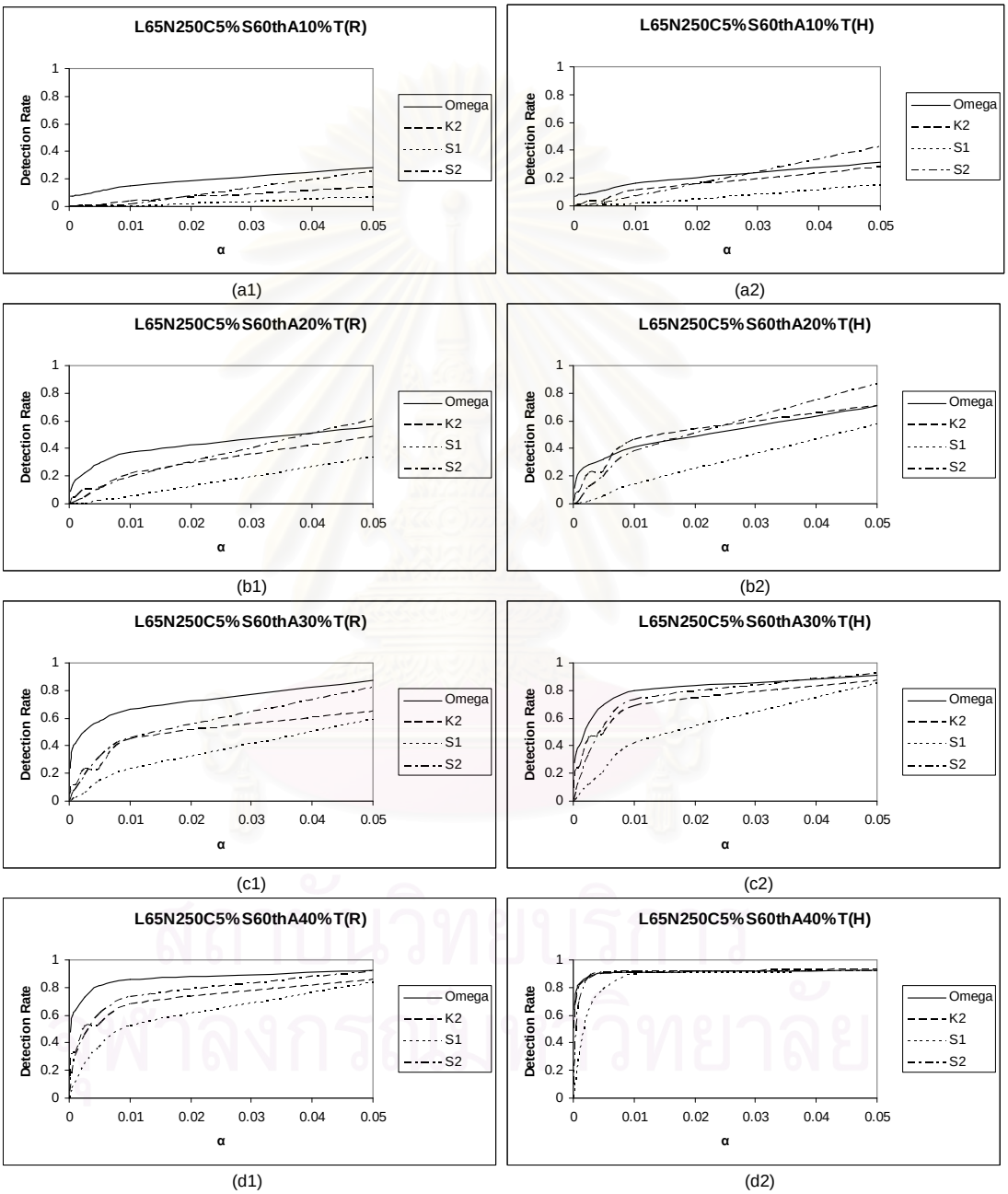
ภาพ 20 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 100 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



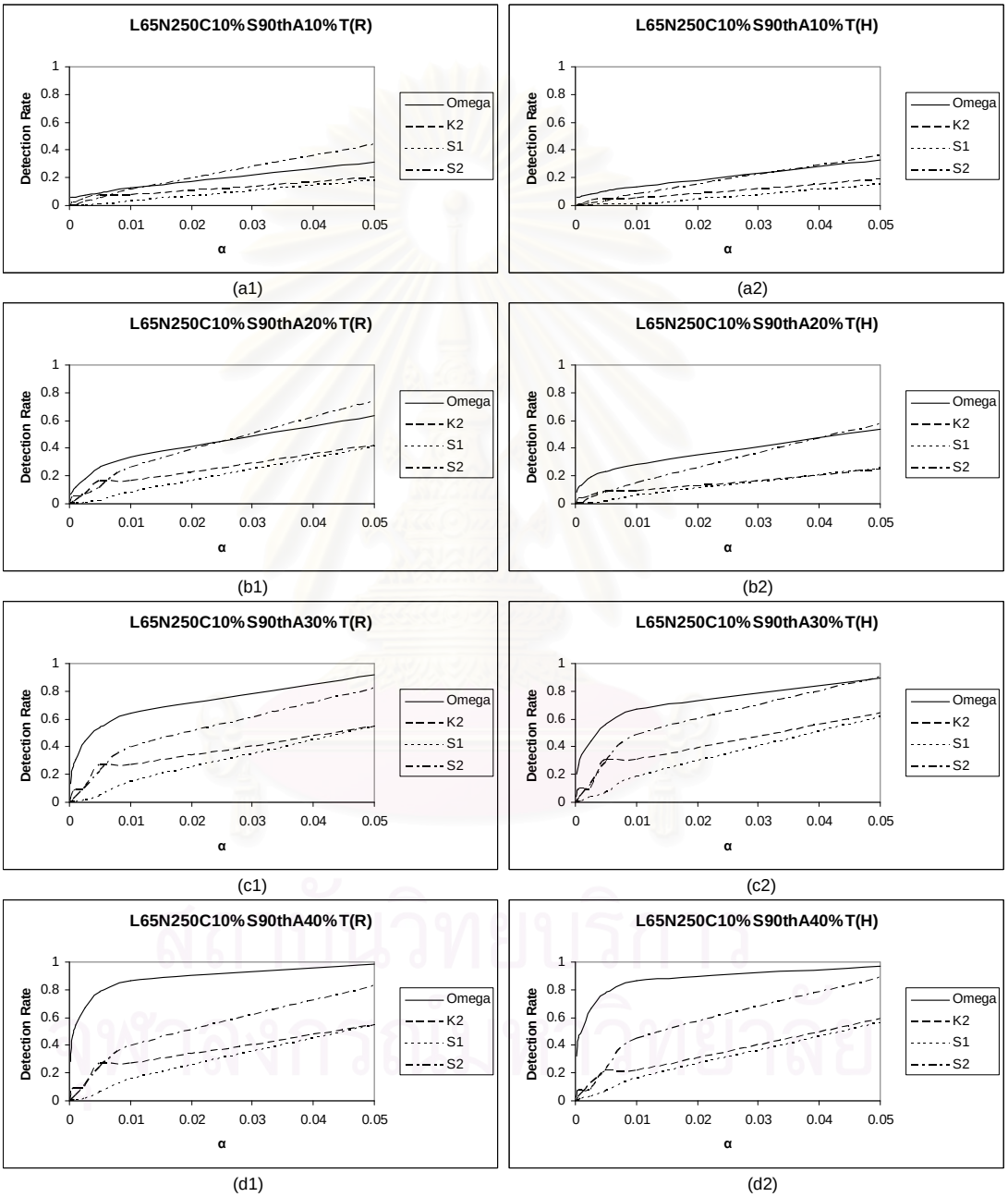
ภาพ 21 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



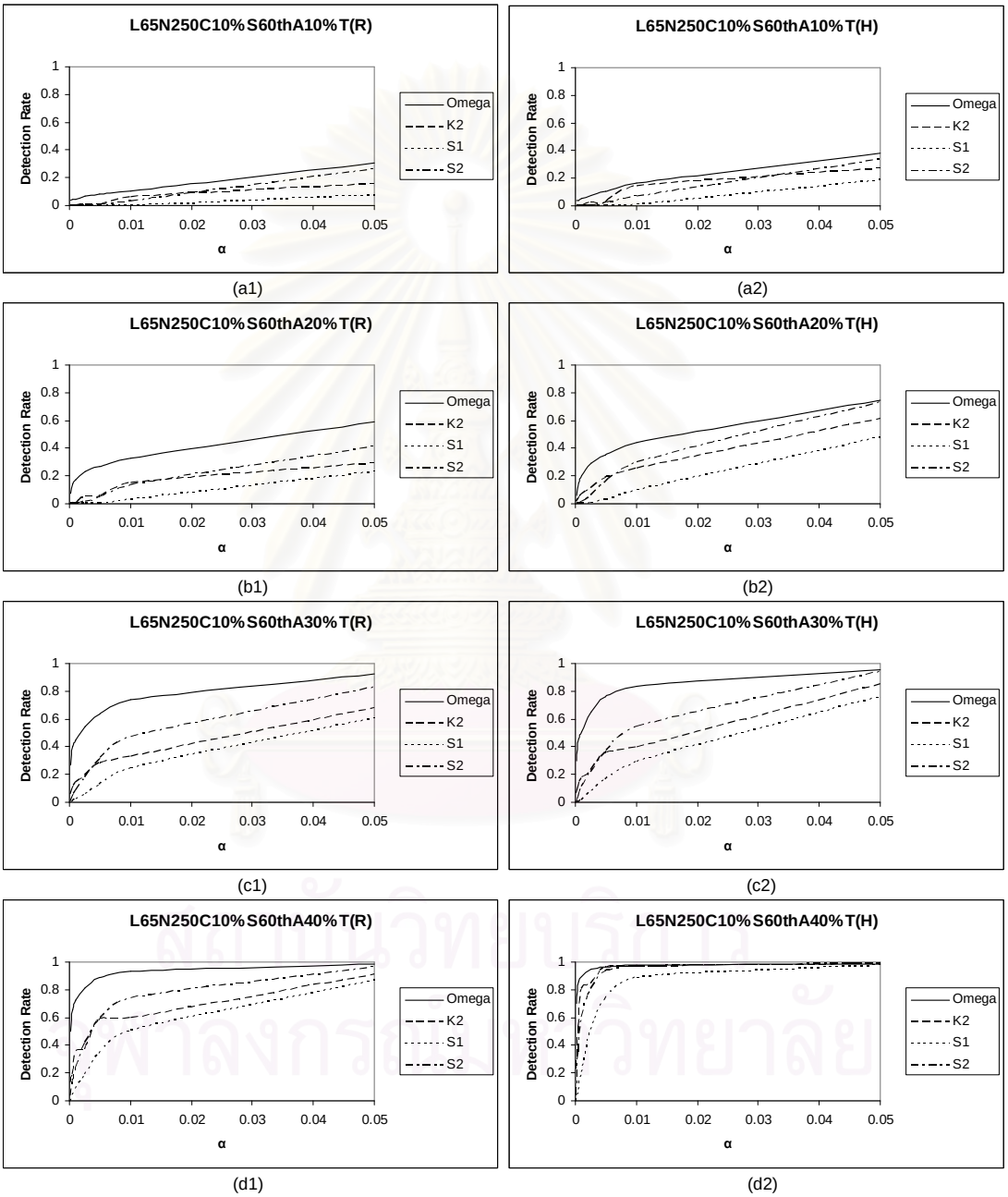
ภาพ 22 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 5 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



ภาพ 23 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



ภาพ 24 อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) ในสถานการณ์ ความยาวแบบสอบ 65 ข้อ จำนวนผู้สอบ 250 คน จำนวนผู้ลอกร้อยละ 10 ผู้ให้ลอกที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 60 แยกตามตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก



ผลการศึกษาค่าอำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) **ของดัชนีตรวจจับการลอกข้อสอบ**

จากตาราง 20-27 และ รูป 9-24 ผู้วิจัยสามารถสรุปผลการศึกษาค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบแต่ละตัวแยกตามตัวแปรที่ผู้วิจัยศึกษา ได้ดังนี้

1. **ดัชนี** ω

ความยาวของแบบสอบ ตัวแปรความยาวของแบบสอบมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω มีแนวโน้มที่จะสูงขึ้นหากแบบสอบมีจำนวนข้อเพิ่มขึ้น

จำนวนผู้สอบ ตัวแปรจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากจำนวนผู้สอบเปลี่ยนแปลง

ร้อยละของจำนวนผู้ลอก ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากร้อยละของจำนวนผู้ลอกเปลี่ยนแปลง

ระดับความสามารถของผู้ให้ลอก ตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี ω เล็กน้อย แต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้ กล่าวคือ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 90 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 60 เล็กน้อย และ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 60 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 90 เล็กน้อย

วิธีการลอก ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω มีแนวโน้มที่จะสูงขึ้นเล็กน้อยหากตัวแปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อยาก

ร้อยละของจำนวนข้อสอบที่ถูกลอก ตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω มีแนวโน้มที่จะสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น

2. **ดัชนี K_2**

<u>ความยาวของแบบสอบ</u> ตัวแปรความยาวของแบบสอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของค่าดัชนี K_2 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากความยาวของแบบสอบเปลี่ยนแปลง

<u>จำนวนผู้สอบ</u> ตัวแปรจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากจำนวนผู้สอบเปลี่ยนแปลง

<u>ร้อยละของจำนวนผู้ลอก</u> ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากร้อยละของจำนวนผู้ลอกเปลี่ยนแปลง

<u>ระดับความสามารถของผู้ให้ลอก</u> ตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 เล็กน้อย แต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้ กล่าวคือ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 ในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 90 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 60 เล็กน้อย และ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 ในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็นตำแหน่งเปอร์เซ็นไทล์ที่ 60 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็นตำแหน่งเปอร์เซ็นไทล์ที่ 90 เล็กน้อย

<u>วิธีการลอก</u> ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 มีแนวโน้มที่จะสูงขึ้นเล็กน้อยหากตัวแปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อยาก

<u>ร้อยละของจำนวนข้อสอบที่ถูกลอก</u> ตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 มีแนวโน้มที่จะสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น

3. **ดัชนี S_1**

<u>ความยาวของแบบสอบ</u> ตัวแปรความยาวของแบบสอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของค่าดัชนี S_1 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากความยาวของแบบสอบเปลี่ยนแปลง

จำนวนผู้สอบ ตัวแปรจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากจำนวนผู้สอบเปลี่ยนแปลง

ร้อยละของจำนวนผู้ลอก ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากร้อยละของจำนวนผู้ลอกเปลี่ยนแปลง

ระดับความสามารถของผู้ให้ลอก ตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 เล็กน้อย แต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้ กล่าวคือ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 ในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 90 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 60 เล็กน้อย และ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 ในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็นตำแหน่งเปอร์เซ็นไทล์ที่ 60 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็นตำแหน่งเปอร์เซ็นไทล์ที่ 90 เล็กน้อย

วิธีการลอก ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 มีแนวโน้มที่จะสูงขึ้นเล็กน้อยหากตัวแปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อยาก

ร้อยละของจำนวนข้อสอบที่ถูกลอก ตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 มีแนวโน้มที่จะสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น

4. **ดัชนี S_2**

ความยาวของแบบสอบ ตัวแปรความยาวของแบบสอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของค่าดัชนี S_2 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากความยาวของแบบสอบเปลี่ยนแปลง

จำนวนผู้สอบ ตัวแปรจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากจำนวนผู้สอบเปลี่ยนแปลง

ร้อยละของจำนวนผู้ลอก ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 ไม่มีแนวโน้มที่จะสูงขึ้นหรือลดลงอย่างชัดเจนหากร้อยละของจำนวนผู้ลอกเปลี่ยนแปลง

ระดับความสามารถของผู้ให้ลอก ตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 เล็กน้อย แต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้ กล่าวคือ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 ในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 90 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 60 เล็กน้อย และ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 ในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 60 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็น ตำแหน่งเปอร์เซ็นไทล์ที่ 90 เล็กน้อย

วิธีการลอก ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 มีแนวโน้มที่จะสูงขึ้นเล็กน้อยหากตัวแปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อยาก

ร้อยละของจำนวนข้อสอบที่ถูกลอก ตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี S_2 มีแนวโน้มที่จะสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น

การเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) **ของดัชนีตรวจจับการลอกข้อสอบ**

จากตาราง 20-27 และ รูป 9-24 ผู้วิจัยสามารถเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ได้ว่า ตัวแปรความยาวของแบบสอบมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω แต่จะไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2, ดัชนี S_1 และ ดัชนี S_2 ตัวแปรจำนวนผู้สอบและตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ ตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เล็กน้อยแต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้ ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ กรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อ

ยากจะสูงกว่ากรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบแบบสุ่มเล็กน้อย ส่วนตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่มีแนวโน้มที่จะมีค่าสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น

ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 และ ตัวแปรจำนวนผู้สอบเป็น 100 คน (L35N100) และ สถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 และ ตัวแปรจำนวนผู้สอบเป็น 250 คน (L35N250) ดัชนี S_2 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด

ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 และ ตัวแปรจำนวนผู้สอบเป็น 100 คน (L65N100) และ สถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 และ ตัวแปรจำนวนผู้สอบเป็น 250 คน (L65N250) ดัชนี ω เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด

ในเกือบทุกสถานการณ์ที่ถูกจำลองขึ้น ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบต่ำที่สุด

จากผลการศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ ได้แก่ ค่าความคลาดเคลื่อนประเภทที่ 1 (type I error) และ ค่าอำนาจการตรวจจับการลอกข้อสอบ (detection rate) ผู้วิจัยสามารถแสดงแนวโน้มของคุณสมบัติทางสถิติทั้งสองในรูปตารางได้ ดังตาราง 28

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 28 แนวโน้มค่าความคลาดเคลื่อนประเภทที่ 1 (Type I Error) และ อำนาจการตรวจจับการลอกข้อสอบ (Detection Rate) ของดัชนีตรวจจับการลอกข้อสอบทั้ง 4 ดัชนี ได้แก่ ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้ตัวแปรต้นด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนผู้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก และวิธีการลอก

ตัวแปรต้น	Type I Error						Detection Rate			
	ω		K_2	S_1	S_2		ω	K_2	S_1	S_2
1. ความยาวของแบบสอบ (test length: L)	L35	L65	✓	✓	L35	L65	+	0	0	0
2. จำนวนผู้สอบ (the number of simulees: N)	✓	×	✓	✓	×	✓	0	0	0	0
3. ระดับความสามารถของผู้ให้ลอก (the ability levels of source: S)	✓	×	✓	✓	×	✓	+/-	+/-	+/-	+/-
4. ร้อยละของจำนวนผู้ลอก (the percentage of copiers: C)	✓	×	✓	✓	×	✓	0	0	0	0
5. ร้อยละของจำนวนข้อสอบที่ถูกลอก (the percentage of answer copied: A)	✓	×	✓	✓	×	✓	+	+	+	+
6. วิธีการลอก (type of copying : T)	✓	×	✓	✓	×	✓	+	+	+	+

หมายเหตุ ความหมายของสัญลักษณ์ที่ปรากฏในตารางถูกอธิบายไว้ในตาราง 29

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 29 ความหมายของสัญลักษณ์ที่ปรากฏในตาราง 26 แยกตาม คุณสมบัติทางสถิติด้านความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ

1. Detection Rate							
สัญลักษณ์	**ความหมาย**	**ตัวแปรต้น**					
+	ตัวแปรต้นมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบในทิศทาง *บวก* กล่าวคือสำหรับ	ความยาวของแบบสอบ	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากแบบสอบมีจำนวนข้อ	มากขึ้น	
				ลดลง		น้อยลง	
		จำนวนผู้สอบ	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากจำนวนผู้สอบ	มากขึ้น	
				ลดลง		น้อยลง	
		ระดับความสามารถของผู้ให้ลอก	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากผู้ให้ลอกมีระดับความสามารถ	สูงขึ้น	
				ลดลง		น้อยลง	
		ร้อยละของจำนวนผู้ลอก	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากจำนวนผู้ลอก	มากขึ้น	
				ลดลง		น้อยลง	
		ร้อยละของจำนวนข้อสอบที่ถูกลอก	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากร้อยละของจำนวนข้อสอบที่ถูกลอก	มากขึ้น	
				ลดลง		น้อยลง	
		วิธีการลอก	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากผู้ลอกทำการลอกข้อสอบ	เฉพาะข้อยาก	
				ลดลง		แบบสุ่ม	
0	ตัวแปรต้น*ไม่มี*อิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบ						
-	ตัวแปรต้นมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบในทิศทาง *ลบ* กล่าวคือสำหรับ	ความยาวของแบบสอบ	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากแบบสอบมีจำนวนข้อ	น้อยลง	
				ลดลง		มากขึ้น	
		จำนวนผู้สอบ	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากจำนวนผู้สอบ	น้อยลง	
				ลดลง		มากขึ้น	
		ระดับความสามารถของผู้ให้ลอก	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากผู้ให้ลอกมีระดับความสามารถ	น้อยลง	
				ลดลง		สูงขึ้น	
		ร้อยละของจำนวนผู้ลอก	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากจำนวนผู้ลอก	น้อยลง	
				ลดลง		มากขึ้น	
		ร้อยละของจำนวนข้อสอบที่ถูกลอก	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากร้อยละของจำนวนข้อสอบที่ถูกลอก	น้อยลง	
				ลดลง		มากขึ้น	
		วิธีการลอก	ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบมีแนวโน้มที่จะ	สูงขึ้น	หากผู้ลอกทำการลอกข้อสอบ	แบบสุ่ม	
				ลดลง		เฉพาะข้อยาก	
+/-	ตัวแปรต้นมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบทั้งในทิศทาง *บวก* และ *ลบ* โดยไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้						

2. Type I Error	
สัญลักษณ์	**ความหมาย**
✓	ดัชนีตรวจจับการลอกข้อสอบ*สามารถควบคุมระดับ*ค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงหรือต่ำกว่าเส้นขอบเขต (boundary line) เพียงเล็กน้อย
×	ดัชนีตรวจจับการลอกข้อสอบ*ไม่สามารถควบคุมระดับ*ค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงกว่าเส้นขอบเขต (boundary line) จนไม่สามารถยอมรับได้

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และ ข้อเสนอแนะ

การศึกษาเรื่อง “การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ” มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบคุณสมบัติทางสถิติซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้เงื่อนไขตัวแปรต้นที่แตกต่างกันในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก

วิธีการศึกษาได้ใช้วิธีวิจัยเชิงทดลอง กล่าวคือ ผู้วิจัยนำข้อมูลผลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ของนิสิตระดับปริญญาบัณฑิตที่ลงทะเบียนเรียนและเข้าสอบในรายวิชาดังกล่าว จำนวนทั้งสิ้น 250 คน ซึ่งเป็นข้อมูลทุติยภูมิ (secondary data) และ ข้อมูลเชิงประจักษ์ (empirical data) มาจัดกระทำด้วยวิธี การจำลองสถานการณ์ (simulation) และ วิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์ S-Plus 2000 ซึ่งผู้วิจัยเป็นผู้เขียนคำสั่งที่ใช้ในการวิเคราะห์ประมวลผลสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัว ภายใต้ความแตกต่างของเงื่อนไขตัวแปรต้นในด้านความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก เพื่อวิเคราะห์หาค่าคุณสมบัติทางสถิติซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ แล้วนำผลการวิเคราะห์มาศึกษาและเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ ซึ่งผลการวิจัยสรุปได้ดังนี้

สรุปผลการวิจัย

1. ผลการศึกษาและเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 (type I error) ของค่าดัชนีตรวจจับการลอกข้อสอบ

1.1 ดัชนี S_2 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) และ มีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35)

1.2 ดัชนี ω สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) และมีแนวโน้มที่จะมี

ระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65)

1.3 ดัชนี K_2 และ ดัชนี S_1 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในทุกสถานการณ์

1.4 ดัชนี S_1 เป็นดัชนีที่มีระดับค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุด (the most conservative index) ในเกือบทุกสถานการณ์

2.ผลการศึกษาและเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของค่าดัชนีตรวจจับการลอกข้อสอบ

2.1 ตัวแปรความยาวของแบบสอบมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω แต่จะไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2, ดัชนี S_1 และ ดัชนี S_2

2.2 ตัวแปรจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่

2.3 ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่

2.4 ตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เล็กน้อยแต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้

2.5 ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ กรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อยากจะสูงกว่ากรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบแบบสุ่มเล็กน้อย

2.6 ตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่มีแนวโน้มที่จะมีค่าสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น

2.7 ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 (L35) ดัชนี S_2 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด

2.8 ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 (L65) ดัชนี ω เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด

2.9 ในเกือบทุกสถานการณ์ที่ถูกจำลองขึ้น ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบต่ำที่สุด

อภิปรายผลการวิจัย

1. ผลการศึกษาและเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 (type I error) ของค่าดัชนีตรวจจับการลอกข้อสอบ

มีการโต้เถียงกันว่าไม่ควรยอมรับอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ (power of indices) ที่ไม่สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ต่ำกว่าเส้นขอบเขต (boundary line) ได้ (Seaman, Levin และ Serlin, 1991 อ้างถึงใน Wollack, 1997) เนื่องจากผลที่ตามมาของการระบุผู้ลอกผิดพลาดเป็นสิ่งที่รุนแรง แต่ค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบก็ไม่ควรที่จะมีค่าต่ำกว่าเส้นขอบเขต (boundary line) มาก เนื่องจากอำนาจในการตรวจจับผู้สอบซึ่งเป็นผู้ลอกจริงจะต่ำมาก (Seaman, Levin และ Serlin, 1991 อ้างถึงใน Wollack, 1997) ดังนั้น ดัชนีตรวจจับการลอกข้อสอบที่ดีควรมีระดับความคลาดเคลื่อนประเภทที่หนึ่งต่ำกว่าเส้นขอบเขต (boundary line) เพียงเล็กน้อย

จากผลการวิเคราะห์ข้อมูลพบว่า ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) ดัชนี S_2 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงหรือต่ำกว่าเส้นขอบเขต (boundary line) เพียงเล็กน้อย และ มีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) โดยดัชนี ω นั้นสามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) และมีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) ซึ่งแตกต่างจากผลการวิจัยของ Wollack (1997) ที่พบว่า ดัชนี ω สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ใต้เส้นขอบเขต (boundary line) ในทุกสถานการณ์ ส่วนดัชนี K_2 และ ดัชนี S_1 นั้นสามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในทุกสถานการณ์ แต่ระดับค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนี S_1 ต่ำกว่าเส้นขอบเขตมากกว่าของดัชนี K_2 ดังนั้นดัชนี S_1 เป็นดัชนีที่มีระดับค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุดในทุกสถานการณ์ กล่าวได้ว่าหากพิจารณาในด้านของความคลาดเคลื่อน

ประเภทที่ 1 แล้ว ดัชนี S_2 เหมาะสมที่จะนำไปใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) ดัชนี ω เหมาะสมที่จะนำไปใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) ส่วน ดัชนี K_2 และ ดัชนี S_1 เหมาะสมที่จะนำไปใช้ในทุกสถานการณ์ แต่ดัชนี K_2 จะสามารถใช้ได้อย่างมีประสิทธิภาพมากกว่าดัชนี S_1 เนื่องจากระดับค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนี S_1 ต่ำกว่าเส้นขอบเขตมากกว่าของดัชนี K_2 ซึ่งจะส่งผลต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 ที่จะมีค่าน้อยกว่าของ ดัชนี K_2 ดังจะเห็นได้จากสรุปผลการวิจัย ส่วนผลการศึกษาและเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของค่าดัชนีตรวจจับการลอกข้อสอบซึ่งพบว่า ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบต่ำที่สุด ในเกือบทุกสถานการณ์

ซึ่งหากพิจารณาค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ ที่ผู้วิจัยศึกษาในตาราง 18 แล้วจะพบว่า ระดับความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ที่ยอมรับได้นั้นมีแนวโน้มที่จะมีค่าสูงกว่าเส้นขอบเขต (boundary line) เล็กน้อย ระดับความคลาดเคลื่อนประเภทที่ 1 ของดัชนี S_2 จะสูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) และ ระดับความคลาดเคลื่อนประเภทที่ 1 ของดัชนี ω จะสูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) สาเหตุที่ระดับความคลาดเคลื่อนประเภทที่ 1 มีแนวโน้มสูงกว่าเส้นขอบเขต (boundary line) ทั้งในระดับที่ยอมรับได้หรือไม่สามารถยอมรับได้ อาจเนื่องมาจากปัจจัยด้านจำนวนผู้สอบซึ่งงานวิจัยชิ้นนี้ศึกษาตัวแปรจำนวนผู้สอบ 2 ระดับ คือ 100 และ 250 คน ซึ่งเป็นตัวแทนของจำนวนผู้สอบกลุ่มเล็กและกลุ่มใหญ่ตามลำดับ แต่ในงานวิจัยของ Sotaridona และ Meijer (2003) จะศึกษาตัวแปรจำนวนผู้สอบ 2 ระดับ คือ 100 และ 500 คน ซึ่งเป็นตัวแทนของกลุ่มผู้สอบขนาดเล็กและขนาดใหญ่ตามลำดับ โดยจำนวนผู้สอบซึ่งเป็นตัวแทนของผู้สอบกลุ่มใหญ่ในงานวิจัยชิ้นนี้มีขนาดเป็นครึ่งหนึ่งของจำนวนผู้สอบซึ่งเป็นตัวแทนของผู้สอบกลุ่มใหญ่ในงานวิจัยของ Sotaridona และ Meijer (2003) ซึ่งตัวแปรจำนวนผู้สอบที่แตกต่างกันนี้อาจจะส่งผลต่อการคำนวณค่าดัชนีตรวจจับการลอกข้อสอบทั้งสี่ โดยเฉพาะ ดัชนี ω ที่มีพื้นฐานการคำนวณอยู่บนค่าพารามิเตอร์ที่ถูกประมาณค่าจากโมเดล Nominal Response Model (Bock, 1972) ซึ่งค่าพารามิเตอร์ที่ประมาณค่าได้จากโมเดลนี้จะมีความถูกต้องมากขึ้นหากมีจำนวนผู้สอบมากขึ้น นอกจากนี้รูปแบบการทดลองหรือจัดกระทำข้อมูล (simulation) ก็อาจเป็นสาเหตุหนึ่ง ดังจะเห็นได้จากข้อค้นพบในงานวิจัยชิ้นนี้ที่ดัชนี ω นั้นสามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) และมีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) ซึ่งแตกต่างจากผลการวิจัยของ Wollack (1997) ที่พบว่า ดัชนี ω

สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ใต้เส้นขอบเขต (boundary line) ในทุกกรณี ซึ่งงานวิจัยของ Wollack (1997) มีรูปแบบการจัดกระทำข้อมูลที่ต่างจากงานวิจัยชิ้นนี้ โดยในงานวิจัยของ Wollack นั้นมีการใช้ แผนผังที่นั่งสอบ (seating chart) ในการจัดกระทำให้เกิดการลอกขึ้นแต่ในงานวิจัยชิ้นนี้ไม่ได้ใช้ โดยแผนผังที่นั่งสอบถูกนำมาช่วยในการจับคู่ผู้ลอกและผู้ให้ลอกให้เป็นคู่ผู้สอบซึ่งนั่งอยู่ในรัศมีที่การลอกสามารถเกิดขึ้นได้เท่านั้น ซึ่งจากงานวิจัยของ Wollack (1997) พบว่าแผนผังที่นั่งสอบสามารถช่วยให้ค่าความคลาดเคลื่อนประเภทที่ 1 ลดลงได้

2. ผลการศึกษาและเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบ (detection rate) ของค่าดัชนีตรวจจับการลอกข้อสอบ

จากผลการวิเคราะห์ข้อมูลพบว่า ตัวแปรความยาวของแบบสอบมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω แต่ตัวแปรจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ซึ่งสอดคล้องกับผลการวิจัยของ Wollack (1997) และ Sotaridona และ Meijer (2002 และ 2003) แต่ตัวแปรความยาวของแบบสอบและจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2, ดัชนี S_1 และ ดัชนี S_2 ซึ่งแตกต่างจากผลการวิจัยของ Sotaridona และ Meijer (2003) ซึ่งพบว่าตัวแปรความยาวของแบบสอบและจำนวนผู้สอบมีอิทธิพลต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2, ดัชนี S_1 และดัชนี S_2 ซึ่งสาเหตุอาจเนื่องมาจากระดับของตัวแปรความยาวของแบบสอบและจำนวนผู้สอบที่ผู้วิจัยและ Sotaridona และ Meijer ศึกษามีความแตกต่างกัน กล่าวคือ ด้านตัวแปรความยาวของแบบสอบผู้วิจัยศึกษา 2 ระดับ คือ 35 และ 65 ข้อ แต่ Sotaridona และ Meijer ศึกษา 2 ระดับคือ 40 และ 80 ข้อ ส่วนในด้านตัวแปรจำนวนผู้สอบผู้วิจัยศึกษา 2 ระดับ คือ 100 และ 250 คน แต่ Sotaridona และ Meijer ศึกษา 2 ระดับคือ 100 และ 500 คน

ด้านตัวแปรร้อยละของจำนวนผู้ลอกพบว่า ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ แต่ผู้วิจัยคาดว่าตัวแปรร้อยละของจำนวนผู้ลอกน่าจะมีอิทธิพลต่อดัชนี ω หากนำไปใช้ในสถานการณ์จริง เนื่องจากค่าสถิติ ω เป็นค่าสถิติที่คำนวณโดยใช้ข้อมูลพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) ซึ่งคำนวณอยู่บนพื้นฐานของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) และพารามิเตอร์จุดตัดและความชันซึ่งคำนวณโดยมีพื้นฐานอยู่บนโมเดล Nominal Response Model (Bock, 1972) ซึ่งพารามิเตอร์ที่กล่าวมาทั้งหมดนั้นจะมีค่าเปลี่ยนแปลงไปจากเดิมหรือบิดเบือนไปจากเดิมหากเกิดการลอกขึ้น โดยถ้าร้อยละของจำนวนผู้ลอกไม่เท่ากัน ยกตัวอย่างเช่น ในงานงานวิจัยชิ้นนี้ผู้วิจัยศึกษาตัวแปรร้อยละของจำนวนผู้ลอก 2 ระดับ คือ ร้อยละ 5 และ ร้อยละ10 ค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และ พารามิเตอร์จุดตัดและความชันซึ่ง

คำนวณโดยมีพื้นฐานอยู่บนโมเดล Nominal Response Model ย่อมมีค่าแตกต่างกัน ซึ่งจะส่งผลให้ดัชนี ω ที่คำนวณได้แตกต่างกันด้วย ซึ่งเมื่อดัชนี ω มีความแตกต่างกันอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ย่อมต่างกันด้วย แต่สาเหตุที่ผลการวิเคราะห์ข้อมูลในงานวิจัยชิ้นนี้ชี้ให้เห็นว่า ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω อาจเนื่องมาจากผู้วิจัยทำการคำนวณค่าสถิติ ω โดยใช้ข้อมูลค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และ พารามิเตอร์จุดตัดและความชันซึ่งคำนวณโดยมีพื้นฐานอยู่บนโมเดล Nominal Response Model ของชุดข้อมูลที่ยังไม่จัดกระทำ (simulation) ให้เกิดการลอก ซึ่งดัชนี ω ที่คำนวณโดยทราบค่าค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และพารามิเตอร์จุดตัดและความชันซึ่งคำนวณโดยมีพื้นฐานอยู่บนโมเดล Nominal Response Model อยู่ก่อนแล้วนั้นเรียกว่า "ω computed using known item and trait parameters" แต่ในสถานการณ์จริงเมื่อนำดัชนี ω ไปใช้คงเป็นไปได้ยากที่ครูผู้สอนส่วนใหญ่จะมีชุดข้อสอบที่ได้ทำการวัดหาค่าพารามิเตอร์ของข้อสอบไว้ก่อนแล้ว ดังนั้น ในสถานการณ์จริงค่าดัชนี ω จะคำนวณโดยไม่ทราบค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และ พารามิเตอร์จุดตัดและความชันซึ่งคำนวณโดยมีพื้นฐานอยู่บนโมเดล Nominal Response Model อยู่ก่อนซึ่งดัชนี ω ประเภทนี้จะเรียกว่า "ω computed using unknown item and trait parameters" ซึ่งค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และ พารามิเตอร์จุดตัดและความชันซึ่งคำนวณโดยมีพื้นฐานอยู่บนโมเดล Nominal Response Model ในสถานการณ์จริงนั้นย่อมจะแตกต่างกันไปหากเกิดการลอกข้อสอบโดยร้อยละของจำนวนผู้ลอกแตกต่างกัน ดังนั้นดัชนี ω ที่คำนวณได้จะแตกต่างกันซึ่งจะส่งผลให้อำนาจการตรวจจับการลอกข้อสอบของดัชนี ω แตกต่างกันด้วย

ผลการวิเคราะห์ข้อมูลยังพบว่าตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เพียงเล็กน้อยแต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้ กล่าวคือ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็นตำแหน่งเปอร์เซ็นไทล์ที่ 90 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็นตำแหน่งเปอร์เซ็นไทล์ที่ 60 เล็กน้อย และ ในการเปรียบเทียบคู่สถานการณ์บางกรณีค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบในสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็นตำแหน่งเปอร์เซ็นไทล์ที่ 60 สูงกว่าสถานการณ์ที่ตัวแปรระดับความสามารถของผู้ให้ลอกเป็นตำแหน่งเปอร์เซ็นไทล์ที่ 90 เล็กน้อย

ผลการวิเคราะห์ข้อมูลด้านตัวแปรวิธีการลอกนั้นพบว่า ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ กรณีสถานการณ์ที่ตัว

แปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อยากจะสูงกว่ากรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบแบบสุ่มเล็กน้อย ซึ่งข้อค้นพบนี้สำหรับดัชนี ω นั้นสอดคล้องกับงานวิจัยของ Wollack (1997)

ตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ โดยค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่มีแนวโน้มที่จะมีค่าสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น โดยอาจกล่าวได้ว่า ผู้สอบที่ลอกข้อสอบจำนวนน้อยมีแนวโน้มต่ำที่จะถูกตรวจจับว่าเป็นผู้ลอก แต่ผู้สอบที่ลอกข้อสอบจำนวนมากมีแนวโน้มสูงที่จะถูกตรวจจับว่าเป็นผู้ลอก ซึ่งข้อค้นพบนี้สอดคล้องกับงานวิจัยของ Wollack (1997) และ Sotaridona และ Meijer (2002 และ 2003)

จากผลการวิเคราะห์ข้อมูลแสดงให้เห็นว่า ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบต่ำที่สุดในเกือบทุกสถานการณ์ที่ถูกจำลองขึ้น ซึ่งสอดคล้องกับคำกล่าวของ Seaman, Levin และ Serlin (1991) ที่ว่าค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบไม่ควรที่จะมีค่าต่ำกว่าเส้นขอบเขต (boundary line) มาก เนื่องจากอำนาจในการตรวจจับผู้สอบซึ่งเป็นผู้ลอกจริงจะต่ำมาก ซึ่งจากผลการวิเคราะห์ข้อมูลพบว่าดัชนี S_1 เป็นดัชนีที่มีระดับค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุดในเกือบทุกสถานการณ์ ดังนั้นจึงส่งผลให้ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบต่ำที่สุดในเกือบทุกสถานการณ์

นอกจากนี้ผลการวิเคราะห์ข้อมูลยังพบว่า ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) ดัชนี S_2 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด ส่วนในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) ดัชนี ω เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด แต่เมื่อเราพิจารณาในแง่ของการควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 จะพบว่า ดัชนี S_2 มีแนวโน้มที่ระดับค่าความคลาดเคลื่อนประเภทที่ 1 จะสูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) ส่วนดัชนี ω มีแนวโน้มที่ระดับค่าความคลาดเคลื่อนประเภทที่ 1 จะสูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) ดังนั้นกล่าวได้ว่า ดัชนี S_2 เป็นดัชนีตรวจจับการลอกข้อสอบที่ไม่เหมาะที่จะใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (L35) และ ดัชนี ω เป็นดัชนีตรวจจับการลอกข้อสอบที่ไม่เหมาะที่จะใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (L65) เนื่องจากค่าดัชนีตรวจจับการลอกข้อสอบที่ดีควรมีอำนาจการตรวจจับการลอกข้อสอบที่สูงและสามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้

จากข้อค้นพบทั้งในด้านความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบทั้งหมด ผู้วิจัยสามารถสรุปได้ว่า ไม่มีดัชนีตรวจจับการลอกข้อสอบตัวใดที่สามารถใช้ได้ดีในทุกๆ สถานการณ์ ดัชนีตรวจจับการลอกข้อสอบแต่ละตัวสามารถใช้ได้อย่างมีประสิทธิภาพในสถานการณ์ที่แตกต่างกัน

ข้อเสนอแนะในการนำไปใช้

1. เนื่องจากความเหมาะสมในการเลือกใช้ดัชนีตรวจจับการลอกข้อสอบในแต่ละสถานการณ์นั้น ต้องพิจารณาใน 2 ส่วน คือ ความสามารถในการควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่สามารถยอมรับได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงหรือต่ำกว่าเส้นขอบเขตเพียงเล็กน้อย เนื่องจากผลที่ตามมาของความคลาดเคลื่อนประเภทที่ 1 หรือ การระบุผู้ลอกผิดพลาดเป็นสิ่งที่รุนแรง และ การมีอำนาจการตรวจจับการลอกข้อสอบที่สูง เนื่องจากอำนาจการตรวจจับการลอกข้อสอบที่สูง แสดงให้เห็นถึงความสามารถในการตรวจจับผู้ลอกข้อสอบที่แท้จริงที่มีประสิทธิภาพ ซึ่งจากข้อค้นพบทั้งหมดในงานวิจัยนี้ผู้วิจัยสามารถสรุปแนวทางในการเลือกใช้ดัชนีตรวจจับการลอกข้อสอบได้ดังนี้
    1.1 ดัชนีตรวจจับการลอกข้อสอบที่เหมาะสมต่อการใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (แบบสอบที่มีจำนวนข้อคำถามน้อย) และ ตัวแปรจำนวนผู้สอบเป็น 100 คน (กลุ่มผู้สอบขนาดเล็ก) (L35N100) คือ ดัชนี ω เนื่องจาก มีอำนาจการตรวจจับการลอกข้อสอบที่ค่อนข้างสูง และ สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่สามารถยอมรับได้ในเกือบทุกสถานการณ์
    1.2 ดัชนีตรวจจับการลอกข้อสอบที่เหมาะสมต่อการใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (แบบสอบที่มีจำนวนข้อคำถามน้อย) และ ตัวแปรจำนวนผู้สอบเป็น 250 คน (กลุ่มผู้สอบขนาดใหญ่) (L35N250) กรณีตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกเป็น ร้อยละ 10 และร้อยละ 20 (A10% และ A20%)คือดัชนี S_1 ส่วนในกรณีตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกเป็น ร้อยละ 30 และร้อยละ 40 (A30% และ A40%) คือ ดัชนี ω เนื่องจาก มีอำนาจการตรวจจับการลอกข้อสอบที่ค่อนข้างสูง และ สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่สามารถยอมรับได้ในเกือบทุกสถานการณ์
    1.3 ดัชนีตรวจจับการลอกข้อสอบที่เหมาะสมต่อการใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (แบบสอบที่มีจำนวนข้อคำถามมาก) และ ตัวแปรจำนวนผู้สอบเป็น 100 คน (กลุ่มผู้สอบขนาดเล็ก) (L65N100) คือดัชนี K_2 เนื่องจาก มี

อำนาจการตรวจจับการลอกข้อสอบที่ค่อนข้างสูง และ สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่สามารถยอมรับได้ในเกือบทุกสถานการณ์

1.4 ดัชนีตรวจจับการลอกข้อสอบที่เหมาะสมต่อการใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (แบบสอบที่มีจำนวนข้อคำถามมาก) และ ตัวแปรจำนวนผู้สอบเป็น 250 คน (กลุ่มผู้สอบขนาดใหญ่) (L65N250) คือดัชนี S_2 เนื่องจาก มีอำนาจการตรวจจับการลอกข้อสอบที่ค่อนข้างสูง และ สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่สามารถยอมรับได้ในเกือบทุกสถานการณ์

โดยแนวทางในการเลือกใช้ดัชนีตรวจจับการลอกข้อสอบข้างต้น สามารถแสดงในรูปตารางได้ ดังนี้

ตาราง 30 แนวทางในการเลือกใช้ดัชนีตรวจจับการลอกข้อสอบภายใต้ปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ

ปัจจัยการทดสอบ		**ความยาวของแบบสอบ** (L)	
		L35 (แบบสอบข้อคำถามน้อย)	L65 (แบบสอบข้อคำถามมาก)
จำนวนผู้สอบ (N)	N100 (กลุ่มผู้สอบขนาดเล็ก)	ω	K2
	N250 (กลุ่มผู้สอบขนาดใหญ่)	S1 (A10% และ A20%) ω (A30% และ A40%)	S2

2. หากผู้นำดัชนีตรวจจับการลอกข้อสอบไปใช้ต้องการความแน่นอนหรือความเชื่อมั่นในผลการตรวจจับการลอกข้อสอบ ผู้วิจัยแนะนำให้ตรวจสอบการลอกข้อสอบด้วยดัชนีตรวจจับการลอกข้อสอบทั้ง 4 ดัชนี จากนั้นพิจารณาแนวโน้มของผลการตรวจจับการลอกข้อสอบจากดัชนีตรวจจับการลอกข้อสอบทั้ง 4 ดัชนี เพื่อระบุหรือบ่งชี้ว่าเกิดการลอกข้อสอบหรือไม่
3. ค่าดัชนีตรวจจับการลอกข้อสอบทั้งสี่ ซึ่งประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω นั้นเพียงแต่ระบุแนวโน้มของการเกิดการลอกข้อสอบเท่านั้น แต่ไม่ได้ระบุว่าในคู่ผู้สอบนั้นผู้สอบคนใดเป็นผู้ลอก (copier) และ ผู้สอบคนใดเป็นผู้ให้ลอก (source) ถ้าหากอยากทราบว่าในคู่ผู้สอบหนึ่งๆ ที่ถูกตรวจจับว่าน่าจะเกิดการลอกข้อสอบกัน ผู้สอบคนใดเป็นผู้ลอก และ

ผู้สอบคนใดเป็นผู้ลอกอาจจะใช้วิธีการสัมภาษณ์เหตุผลในการตอบข้อคำถามต่างๆ ในแบบสอบนั้น

4. ดัชนีตรวจจับการลอกข้อสอบควรนำมาใช้หลังจากที่ผู้คุมสอบสังเกตเห็นพฤติกรรมที่ผิดปกติของคู่ผู้สอบหนึ่งๆ หรือ มีหลักฐานพยานบุคคลแล้ว เพื่อใช้ผลการวิเคราะห์ค่าดัชนีตรวจจับการลอกข้อสอบเป็นหลักฐานทางสถิติ (statistical evidence) เพื่อยืนยันการเกิดการลอกข้อสอบ โดยคู่ผู้สอบที่ถูกระบุว่าลอกข้อสอบกันควรจะถูกคุมสอบอย่างเข้มงวด หรือ นั่งสอบแยกกันในระยะที่ไม่สามารถลอกข้อสอบกันได้ในการสอบครั้งต่อไป

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. เนื่องจากงานวิจัยชิ้นนี้ศึกษาด้านตัวแปรจำนวนผู้สอบเพียง 2 ระดับคือ 100 และ 250 คน ซึ่งเป็นตัวแทนของจำนวนผู้สอบกลุ่มเล็กและกลุ่มใหญ่ตามลำดับ ในการวิจัยครั้งต่อไปควรเพิ่มจำนวนผู้สอบกลุ่มใหญ่ให้มากกว่า 250 คนเพื่อให้การประมาณค่าพารามิเตอร์ตามโมเดล Nominal Response Model และการวิเคราะห์ค่าดัชนีตรวจจับการลอกข้อสอบมีความถูกต้องมากขึ้น
2. เนื่องจากงานวิจัยชิ้นนี้ศึกษาค่าดัชนี ω ประเภทที่คำนวณโดยทราบค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และ พารามิเตอร์ข้อสอบอยู่ก่อนแล้ว (ω computed using known item and trait parameters) ซึ่งในสถานการณ์จริงคงเป็นไปได้ยากที่จะทราบค่าพารามิเตอร์ของข้อสอบก่อน ดังนั้นในการวิจัยครั้งต่อไปควรทำการศึกษาค่าดัชนี ω ประเภทที่คำนวณโดยไม่ทราบค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และพารามิเตอร์ข้อสอบอยู่ก่อน (ω computed using unknown item and trait parameters)
3. เนื่องจากงานวิจัยชิ้นนี้ใช้ข้อมูลจริงซึ่งเป็นข้อมูลทุติยภูมิ (secondary data) ซึ่งได้แก่ ผลการตอบข้อสอบส่วนข้อสอบปรนัย ของนิสิตที่ลงทะเบียนและเข้าสอบ ในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 มาใช้ในการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (Simulation) ซึ่งจากการทำวิจัยพบว่าการใช้ข้อมูลจริงทำให้มีข้อจำกัดหลายประการ เช่น ข้อจำกัดด้านจำนวนผู้สอบซึ่งส่งผลกระทบต่อการประมาณค่าพารามิเตอร์ตามทฤษฎีการทดสอบแนวใหม่ เป็นต้น ดังนั้นในการวิจัยครั้งต่อไปควรใช้ข้อมูลที่ไม่ใช่ข้อมูลจริงมาจัดกระทำ โดยอาจใช้วิธีการจำลองชุดข้อมูลผลการตอบข้อสอบที่มีลักษณะตรงตามเงื่อนไขที่ศึกษาจากโปรแกรมคอมพิวเตอร์จากนั้นจึงนำชุดข้อมูลที่จำลองขึ้นไปทำการจัดกระทำข้อมูลต่อไป
4. เนื่องจากงานวิจัยชิ้นนี้เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 ซึ่งเมื่อนำผลการตอบ

ข้อสอบของนิสิตที่ลงทะเบียนและเข้าสอบในรายวิชาดังกล่าว จำนวน 250 คนไปจัดกระทำตามปัจจัยด้าน ความยาวของแบบสอบ และ จำนวนผู้สอบ จนได้ชุดข้อมูล 4 ชุด แล้วนำชุดข้อมูลทั้ง 4 ชุด ไปวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการทดสอบแนวใหม่เพื่อนำผลการวิเคราะห์มาคัดเลือกข้อสอบยากพบว่ามีจำนวนข้อคำถามที่จัดเป็นข้อสอบยากตามเกณฑ์ที่ระบุไว้ในหนังสือ ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories) ของ ศ.ดร. ศิริชัย กาญจนวาสี ไม่เพียงพอต่อการจัดกระทำข้อมูลในปัจจัยด้านร้อยละของจำนวนข้อสอบที่ถูกลอก ทำให้ผู้วิจัยต้องปรับเกณฑ์ในการพิจารณาเพื่อระบุข้อสอบระดับยากจนข้อสอบที่ถูกระบุว่าเป็นข้อสอบยากด้วยเกณฑ์ดังกล่าวส่วนหนึ่งมีลักษณะเป็นข้อสอบที่มีระดับความยากปานกลางในความเป็นจริง ดังนั้นในการวิจัยครั้งต่อไปหากมีการศึกษาปัจจัยด้านร้อยละของจำนวนข้อสอบที่ถูกลอกและวิธีการลอก โดยมีการลอกข้อสอบเฉพาะข้อยากเป็นระดับย่อยของปัจจัยวิธีการลอก ควรที่จะใช้แบบสอบที่มีจำนวนข้อสอบที่ถูกระบุว่าเป็นข้อสอบยากมากเพียงพอที่จะศึกษาในปัจจัยด้านร้อยละของจำนวนข้อสอบที่ถูกลอก โดยเกณฑ์ที่ใช้ในการระบุข้อสอบที่อยู่ในระดับยากต้องสามารถระบุข้อสอบที่เป็นข้อสอบยากในความเป็นจริงได้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

ภาษาไทย

ศิริชัย กาญจนวาสี. ทฤษฎีการทดสอบแนวใหม่. พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2550.

สุชาดา บวรกิติวงศ์. สถิติประยุกต์ทางพฤติกรรมศาสตร์. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2548.

ภาษาอังกฤษ

Agresti, A. An introduction to categorical data analysis. New York: Wiley, 1996.

Belleza, F. S. and Belleza S. F. Detection of cheating on multiple-choice tests by using error-similarity analysis. Teaching of Psychology 16(1989): 151-155.

Cizek, G. J. Cheating on tests: how to do it, detect it, and prevent it. Journal of College Student Development (2000).

Frary, R. B. Statistical detection of multiple-choice test answer copying: state of the art. Applied Measurement in Education. 6(1993): 152-165

Gronlund , N.E. Assessment of student achievement. 7 th ed.USA: Pearson Education, 2003.

Hanson, B. A., Harris, D. J. and Brennan, R. L. A comparison of several statistical methods for examining allegations of copying. ACT Research Report Series 87-15 (1987)

Maramark et al. Academic dishonesty among college students. Issues in education (1993): 1-14

MathSoft. S-Plus 2000 [Computer software and manual]. Seattle, WA, 2000.

McCabe, D. L. and Trevino, L. K. Individual and contextual influences on academic dishonesty: A multicampus investigation. Research in Higher Education 38(1997): 379-396

McCabe, D. L., Trevino, L. K. and Butterfield, K. D. Cheating in academic institutions: a decade of research. Ethics & Behavior 3(2001): 219-232

Rothwell, D. and Day, A. S-PLUS for S-PLUS 2000 for Windows. ,1997

Sotaridona, L.S. and Meijer, R.R. Statistical properties of the K-index for detecting answer copying. Journal of Educational Measurement 39(2002): 115-132

Sotaridona, L.S. and Meijer, R.R. Two new statistics to detect answer copying. Journal of Educational Measurement 40(2003): 53-69

Sotaridona, L.S. Cheating detection using the S_2 copying index. The Philippine Statistician 52(2003): 59-67

Thissen, D. MULTILOG version 6 user's guide. Chicago: Scientific Software, 1991.

Venables, B. and Smith, D. Notes on S-PLUS: a programming environment for data analysis and graphics. Australia, 1992.

Wollack, J.A. A nominal response model approach for detecting answer copying. Applied Psychological Measurement 21(1997): 307-320

Wollack, J.A. Detection of answer copying with unknown item and trait parameters. Applied Psychological Measurement 22(1998): 144-152

Wollack, J.A., Cohen, A.S. and Serlin, R.C. Defining error rates and power for detecting answer copying. Applied Psychological Measurement 25(2001): 385-404

Wollack, J.A. Comparison of answer copying indices with real data. Journal of Educational Measurement 40(2003): 189-205

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

THE S-PLUS FUNCTION OF K_2 INDEX

```
K2.index<-function(k,n,X,sou,sub,V)
{

# The “K2.index function” implement the K2 index proposed by Sotaridona and
# and Meijer (2002).

#Programmed by: Kridsada Thirasophon
#Date: April 2007

#Reference: Sotaridona,L. S., & Meijer, R. R. (2002). Statistical Properties of the K-index
# for Detecting Answer Copying

# NOTATION
# k- the number of items
# n- the number of examinees
# X- a matrix of response pattern (actual response, not the 0/1). The rows
#      of X for the items and the column for the examinees
# sou- the column number identifying the source’s location in X
# sub- the column number identifying the copier’s location in X
# V- a vector of answer keys, the length of V is the same is the number of
#          row of X

# (1) Number of wrong (nw) and number of matching wrong response (mw)
  nw<- matrix(0,1,n)
  mw<- matrix(0,1,n)
  for(i in 1:n)  {
          mw[,i]<- sum(X[,sou]==X[,i]&X[,sou]!=V)
          nw[,i]<- k-sum(X[,i]==V)
```

```
        }

  u<- matrix(sort(unique(nw)),1)          # unique number of wrong, needed in
conditioning
  r<- ncol(u)                             # r is the number of groups

# (2) The proportion of wrong answers for each subgroup r (Qr)
        Qr<- matrix(u/k,ncol=1)

# (3) Parameter Qr^2 (Qrs)
        Qrs<- matrix(Qr^2,ncol=1)

# (4) Parameter Pr for each subgroup
  # identify examinees with matching wrong answer (Mrj)
        Mrj<-matrix(0,r,n)
        for(i in 1:r)  {
                for(j in 1:n)  {
                if(u[1,i]==nw[j]) Mrj[i,j]<-mw[j]
                }
        }

  # compute cumulative value of matching wrong answer of each subgroups (Mrjc)
        Mrjc<- apply(Mrj,1,sum)

  # identify examinees with the same number of wrong answer in each subgroups (snw)
        snw<- matrix(0,r,n)
        for(i in 1:r)  {
                for(j in 1:n)  {
                if(u[1,i]==nw[j]) snw[i,j]<-1
                }
        }
```

```
  # total number of examinees with the same number of wrong answer in each
subgroups (Tsnw)
        Tsnw<-apply(snw,1,sum)

  # mean comulative value of matching wrong answer for each subgroup (Mrjm)
        Mrjm<-Mrjc/Tsnw

  # the matrix of ws (Mws)
        ws<-nw[,sou]
        Mws<- matrix(ws,r,1)

  # parameter Pr
        Pr<- Mrjm/Mws

#(5) regression process
  # build database (datab)
        datab<- data.frame(Pr,Qr,Qrs)

  # estimated regression coefficients (Pr.lm)
        Pr.lm<-lm(Pr~Qr+Qrs,data=datab)

  # matrix of estimated regression coefficients (MPr.Co)
        MPr.Co<-matrix(Pr.lm[[1]])

#(6) estimated regression equation
        Pr<-matrix(0,r,1)
              for(i in 1:r){
              Pr[i,]<-sum(MPr.Co[1,]+ MPr.Co[2,]*Qr[i,]+ MPr.Co[3,]*Qrs[i,])
              }
        sPr<-sum(Pr)            #sum of all element in matrix Pr
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
  #parameter P2(P2)
        P2<-sPr/r

#(7) estimate K2.index
        mwc<-mw[,sub]
        K2<-1-pbinom(mwc,ws,P2)

        list(K2=K2)
        }
#end
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

THE S-PLUS FUNCTION OF S_1 INDEX

```
S1.index<-function(k,n,X,sou,sub,v)
{
#The “S1.index function” implements the S1 index written by Sotaridona and Meijer
#(2003)

#Programmed by: Kridsada Thirasophon
#Date:  March 2007

#Reference:     Sotaridona, L. S., & Meijer, R.R.(2003). Two new statistics to detect
#answer copying. Journal of Educational Measument.

#NOTATION
# k -  the number of items
# n -  the number of examinees
# X - a matrix of response pattern (actual response, not the 0/1). The rows of X for the
#items and the columns for the examinees
# sou - the column number identifying the source’s location in X
# sub - the column number identifying the copier’s location in X
# v - a vector of answer keys, the length of v is the same is the number of row of X

#(1) Number of wrong (nw) and number of matching wrong responses (mw)
        nw<- matrix(0,1,n)
        mw<- matrix(0,1,n)
        for(i in 1:n) {
                mw[,i] <- sum(X[,sou]==X[,i]&X[,sou]!=v)
                nw[,i] <- k-sum(X[,i]==v)
        }

        u <- matrix(sort(unique(nw)),1)#unique number of wrong, needed in conditioning
```

```
        r <- ncol(u)                          # r is the number of subgroups

#(2) Estimation of the loglinear parameter
        ul <- matrix(0,1,r)
        z <- matrix(0,1,r)
        for(i in 1:r) {
           for(j in 1:n) {
                 if (u[,i]==nw[j]){
                       ul[,i] <- 1 + ul[,i]      # this count the number of examinees with
                                                 the same
                                                 # number –wrong

                       z[,i] <- z[,i] + mw[j]    # this computes the sum of match wrong
                                                 which is
                                                 # needed to obtained the mu
                       }
                 }
        }

        # mu is the mean number of match-wrong for each group r
                 mu <- z/ul

        #Estimation of the mean mu using loglinear model

        ##########################################################################
#########      ff is an array which stores the results of fitting the loglinear model
        ## ff[[1]][1] is the first element of the first array in ff
        ## pred is the “predicted value” assuming the model holds; also called the fitted
        ## value
        ##########################################################################
#########
```

```
        x <- u[1,]          # transform to become a vector
        y <- mu[1,]

        # Fitting the loglinear model using the "glim" function is S-Plus

                ff <- glim(x,y,error= "poisson", link= "log", resid= "p")

        #Estimate of the parameter of the loglinear model
                B0 <- ff[[1]][1]
                B1 <- ff[[1]][2]

        #Predicted value
                pred <- exp(B0+B1*u)

        #reflecting the predicted value to all examinees
                pred.i <- matrix(0,1,n)
                        for(i in 1:r) {
                                for(j in 1:n) {
                                        if (u[,i]==nw[j])
                                                pred.i[,j] <-pred[i]
                        }
                }

                S1<- 1-ppois(mw[sub],pred.i[sub])
        Out <- list(S1=S1)
}
#end
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

THE S-PLUS FUNCTION OF S_2 INDEX

```
S2.index<-function(k,n,X,sou,sub,V,m)
{
#The “S2.index function” implements the S2 index proposed by Sotaridona and
# Meijer(2003)

#Programmed by: Leonardo S. Sotaridona
#Developed by: Kridsada Thirasophon
#Date: April 2007

#Reference: Sotaridona, L.S., Meijer, R.R. (2003). Two new statistics to detect
# answer copying. Journal of Educational Measurement.

#NOTATION
# k-the number of items
# n-the number
# X-a matrix of response pattern (actual response, not the 0/1). The row of X for the
# items and columns for the examinees
# sou- the column number identifying the source’s location in X
# sub- the column number identifying the copier’s location in X
# V- a vector of answer keys, the length of v is the same is the number of row of X
# m- the number of options

#(1) Number or wrong(nw) and number of matching wrong responses(mw)
        nw<-matrix(0,1,n)
        mw<-matrix(0,1,n)
                for(i in 1:n){
                mw[,i]<-sum(X[,sou]==X[,i]&X[,sou]!=V)
                nw[,i]<-k-sum(X[,i]==V)
}
```

```
u<-matrix(sort(unique(nw)),1)  # unique number of wrong, needed in
                                # conditioning
r<-ncol(u)                      # r is the number of subgroups

#(2) Identify examinees with the same number of wrong

S<-matrix(NA,r,n)
        for(i in 1:r){
                for(j in 1:n){
                        if(u[1,i]==nw[j])S[i,j]<-j
                                }
                        }

        Usi<-X[,sou]            #Usi is the response pattern of the source

#(3) Computes Pij for each group r

        dummy1<-matrix(0,k,r)
        for(i in 1:r){
                ind<-S[i,][S[i,]!= "NA"]
                respp1<-X[,ind]

# Within a certain group r, determine the examinees with matching
# answer as the source in every item, note, the correct responses are included

                F1<-respp1-Usi
                F2<-F1*2

                F2[F2==0]<-1; F2[F2!=1]<-0
                dim(F2)<-c(k,length(ind))
```

```
        F3<-apply(F2,1,sum)

    #Prij
        Prij<-F3/length(ind)

    #weighted evidence is in q.i
        q.i<-round( ((((m+1)/(m-1))*2.718)^(-(m+1)*Prij)),4)
        dummy1[,i]<-q.i
    }

#(4) Reflecting the weighted of evidence

    weight<-matrix(0,k,n)
    for(i in 1:r){
        for(j in 1:n){
            if(u[1,i]==nw[j])
                weight[,j]<-dummy1[,i]
            }
        }

#(5) Computing M*rj (sum of weighted evidence for all the matching correct item)
    D2<-weight
    XX<-X                   #X is an input matrix of response patterns
    XX[,sou][XX[,sou]!=V]<-0    # source's incorrect answer is replaced with 0
    XX<-XX*XX[,sou]             # just a trict

    XX[XX==(V*V)]<-1            # source's correct response is replaced with 1
    XX[XX!=1]<-0

#the sum of weighted evidence is in Mrij=M*ij
```

```
        Mrij<-matrix(apply(XX*D2,2,sum),1,n)

#combining evidence: matching wrong plus weighted
#matching correct

        mwl<-mw+ceiling(Mrij)

#(6) Estimation of the loglinear parameters

        ul<-matrix(0,1,r)
        z<-matrix(0,1,r)
                for(i in 1:r){
                        for(j in 1:n){
                                if(u[,i]==nw[j]){

                                ul[,i]<-1+ul[,i]        #this counts the number of
                                                        #examinees with the same
                                                        #number-wrong
                                z[,i]<-z[,i]+mw[j]      #this computes the sum of match
                                                        #wrong which is needed to
                                                        #obtained the mu
                                        }
                                }
                        }

        #mu is the mean number of match-wrong for each group r
                mu<-z/ul

uul<-matrix(0,1,r)
zzz<-matrix(0,1,r)
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
for(i in 1:r){
        for(j in 1:n){
                if(u[,i]==nw[j]){
                uul[,i]<-1+uul[,i]                  #this counts the number of examinees
                                                    #with the same number-wrong
                zzz[,i]<-zzz[,i]+mwl[j]  #this computes the sum of match wrong
                                                    #(plus weighted match-correct) which
                                                    #is needed to obtained the mean mu
                                }
                        }
                }
        #Mean mul
                mul<-zzz/uul

# Estimation of the mean mu using loglinear model

#########################################################################
##ff is an array which stores the results of fitting the loglinear model
##ff[[1]][1] is the first element of the first array in ff
##pred is the “predicted value” assuming the model holds; also called the fitted value
#########################################################################

        x<-u[1,]        #transform to become a vector, note that u and mu1 are in
                        #matrix form
        y<-mu[1,]

#Fitting the loglinear model using the “glim” function in S-Plus
        ff<-glim(x, y, error=“poisson”, link=“log”, resid= “p”)

#Estimate of the parameter of the loglinear model
        B0<-ff[[1]][1]
```

```
        B1<-ff[[1]][2]


#Predicted value
        pred<-exp(B0+B1*u)


#reflecting the predicted value to all examinees
        pred.i<-matrix(0,1,n)
        for(i in 1:r){
                for(j in 1:n){
                        if(u[,i]==nw[j])
                                pred.i[,j]<-pred[i]
                        }
                }
        S2<-1-ppois(mwl[sub],pred.i[sub])
        list(S2=S2)
        }
#end
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

THE S-PLUS FUNCTION OF ω STATISTIC

```
Omega.index <- function(k,n,X,sou,sub,SL,IN,Z)
{

#The “Omega.index function” implement the omega index proposed by James A.
Wollack
#Programmed by Kridsada Thirasophon
#Date:  April 2007
#Reference: James A. Wollack(1997). A Nominal Response Model Approach for
#                Detecting Answer Copying

#NOTATION
# k- the number of item
# n- the number of examinees
# X- a matrix of response pattern (actual response, not the 0/1). The rows
#      of X for the items and the column for the examinees
# sou- the column number identifying the source’s location in X
# sub- the column number identifying the copier’s location in X
# m- the number of options
# Z- a matrix of ability level of all examinees. The column of the matrix
#         for the examinees
# SL- a matrix of slop parameter, the row of the matrix for the items and
#         the column for the options
# IN- a matrix of intercept parameter, the row of the matrix for the items
#         the column for the options

# (1) The number of identically answered items between source and other examinees
(his)
        his<-matrix(0,1,n)
                for(i in 1:n){
```

```
        his[,i]<-sum(X[,sou]==X[,i])
        }

  # the number of identically answered items between source and copier (hcs)
        hcs<-his[,sub]

  # m-the number of option (this number will be changed according to number
  #of options of test
        m<-4

# (2) Probability of copier selecting the answer provided by source for all items (Pc)
  # Part1: exp[(slope parameter)*(ability level of copier)+(intercept parameter)]
  #       for all items (Pa1)

  #matrix of unique option (MM)
        MM<-matrix(1:4,ncol=1)

  # matrix of slope parameter of copier who selected the answer provided by source
  #for all items (SLC) and  matrix of intercept parameter of copier who selected the
  # answer provided by source for all item (INC)

        SLC<- matrix(0,k,1)
        INC<- matrix(0,k,1)
        for(i in 1:k)  {
          for(j in 1:m)  {
                if(X[i,sou]==MM[j,]) {
                        SLC[i,]<-SL[i,j]
                        INC[i,]<-IN[i,j]
                        }
                }
        }
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
        Pa1<- matrix(0,k,1)

  # matrix of copier's ability level
        MZc<- matrix(Z[,sub],k,1)

        Pa1<-exp(SLC*MZc+INC)


  # Part2 : sum of exp[(slope parameter)*(ability level of copier)+
  # (intercept parameter)] for all options (Pa2)

        Pa2<-matrix(0,k,1)
  # the matrix of ability level of copier dim(k,m): (MZc2)
        MZc2<-matrix(Z[,sub],k,m)
        el3<-exp(SL*MZc2+IN)
        Pa2<-apply(el3,1,sum)

  # Probability of copier selecting the answer provided by source for all items: (Pc)
              Pc<-Pa1/Pa2

# (3) Expected number of matches: (EM)
              EM<-apply(Pc,2,sum)

# (4) Standard error (se)
  #element (Qc) : 1-Pc
              Qc<- 1-Pc

  #element (PQ) : Pc*Qc
              PQ<-Pc*Qc
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
  #variance of the number of answer matches (v)
                v<- apply(PQ,2,sum)

  #standard error (se)
                se<-sqrt(v)

# (5) omega index (w)
                w<-(hcs-EM)/se

                list(w=w)
        }
#end
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นาย กฤษฎา ถิระโสภณ เกิดเมื่อวันที่ 9 ธันวาคม 2526 สำเร็จการศึกษาครุศาสตรบัณฑิต (เกียรตินิยมอันดับ 1) จากภาควิชาหลักสูตร การสอนและเทคโนโลยีการศึกษา วิชาเอกฟิสิกส์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2547 และศึกษาต่อไปในระดับปริญญามหาบัณฑิต ภาควิชาวิจัยและจิตวิทยาการศึกษา สาขาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2548

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย