A Comparison of Statistical Properties of the Indices for Detecting Answer Copying

Kridsada Thirasophon

Sirichai Kanjanawasee

ABSTRACT

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. Final examination test responses of 250 students of Faculty of Education, Chulalongkorn University in ED MEAS EVA course were used in answer copying simulation. The S-plus computer program was employed in analysis processes. 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. The detection rate of increased with test length but not for K_2 index, S_1 index and S_2 index. 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.

Journal of Research Methodology, Volume 21, Number 1 (January-April 2008)

วารสารวิธีวิทยาการวิจัย ปีที่ 21 ฉบับที่ 1 (มกราคม-เมษายน 2551)

การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ

กฤษฎา ถิระโสภณ
ศิริชัย กาญจนวาสี

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์และเปรียบเทียบความคลาดเคลื่อนประเภทที่ 1 และอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบ ได้แก่ ดัชนี K_2 ดัชนี S_1 ดัชนี S_2 และ ดัชนี ω ภายใต้สถานการณ์ซึ่งแตกต่างในเงื่อนไขด้านความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก โดยใช้วิธีวิจัยเชิงทดลอง ซึ่งนำข้อมูลผลการตอบข้อสอบปลายภาคประจำรายวิชา การวัดและการประเมินผลทางการศึกษาของนิสิตคณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยจำนวน 250 คน มาจัดกระทำด้วยวิธีการจำลองสถานการณ์และวิเคราะห์ข้อมูลด้วยโปรแกรม S-Plus ผลการวิจัยพบว่า ดัชนี K_2 และดัชนี S_1 สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในทุกสถานการณ์ ดัชนี S_2 สามารถควบคุมค่าความคลาดเคลื่อน ประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ความยาวของแบบสอบเป็น 65 ข้อและดัชนี ω สามารถควบคุมค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ความยาวของแบบสอบเป็น 35 ข้อ ตัวแปรความยาวของแบบสอบที่เพิ่มขึ้นมีอิทธิพลทำให้ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีเพิ่มขึ้น แต่ไม่มีอิทธิพลสำหรับดัชนี K_2, ดัชนี S_1 และ ดัชนี S_2 ตัวแปรวิธีการลอกแบบลอกเฉพาะข้อยากและตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกที่เพิ่มขึ้นมีอิทธิพลทำให้ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เพิ่มขึ้น

ความเป็นมาและความสำคัญของปัญหา

ในยุคปัจจุบันโลกได้กลายเป็นสังคมฐานความรู้ (Knowledge-based society) มนุษย์ไม่ว่าชนชาติใด ศาสนาใดล้วนแล้วแต่ที่จะต้องมีความรู้ และใช้ความรู้ที่มีไปแก้ไขปัญหาและพัฒนาประเทศของตน ซึ่งการที่จะทำให้ประชาชนในชาตินั้นมีความรู้ได้ วิธีการที่สำคัญที่สุดคือ การจัดการเรียนการสอนแก่ประชากรให้มีความรู้ความสามารถ แต่เพียงแค่การจัดการเรียนการสอนแก่ผู้เรียนนั้นไม่เพียงพอที่จะทำให้เราทราบได้ว่าผู้เรียนมีความรู้ความสามารถในเรื่องที่เรียนแล้ว ดังนั้น จึงต้องมีการประเมินผลเข้ามาเป็นส่วนหนึ่งในการจัดการเรียนการสอน โดยกระบวนการการประเมินผลจะช่วยให้การสอนมีประสิทธิภาพมากขึ้นและทำให้ผู้เรียนเกิดการเรียนรู้มากขึ้นอีกด้วย (Norman, 2003)

ในการประเมินผลการเรียนรู้นั้นจำเป็นต้องทำการวัดผลการเรียนของผู้เรียน ซึ่งวิธีการวัดผลการเรียนที่เป็นที่นิยมมากในปัจจุบัน คือ การสอบโดยใช้แบบสอบเลือกตอบ (Multiple-Choice Tests) โดยการจัดสอบเพื่อที่วัดผลสัมฤทธิ์ซึ่งเกิดจากการเรียนรู้ของผู้เรียนให้ได้ถูกต้องตามสภาพความเป็นจริงนั้น ผู้สอนต้องหาวิธีการที่จะป้องกันไม่ให้เกิดการโกงข้อสอบขึ้น (Cheating) เนื่องจากหากเกิดการโกงข้อสอบขึ้นไม่ว่าในรูปแบบใดก็ตามจะส่งผลให้ผลการสอบที่ได้นั้นมีความคลาดเคลื่อนไปจากระดับความสามารถที่แท้จริงของผู้สอบซึ่งจะทำให้ผลการประเมินระดับความสามารถของผู้เรียนคลาดเคลื่อนไปจากความเป็นจริงด้วย

การลอกข้อสอบ (Answer Copying) เป็นประเภทหนึ่งของการโกงข้อสอบ โดยการลอกข้อสอบนั้นจะทำให้ผลการสอบของผู้สอบไม่ตรงตามความเป็นจริง ซึ่งผลกระทบจากการลอกข้อสอบไม่เพียงแต่จะมีต่อการวัดและการประเมินผลการเรียนระดับรายวิชาหรือระดับชั้นเรียนเท่านั้น แต่ยังส่งผลกระทบต่อการสอบระดับที่สูงขึ้น เช่น การสอบเข้าศึกษาต่อ การสอบคัดเลือกบุคคลเข้าทำงาน การสอบเพื่อรับใบอนุญาตประกอบวิชาชีพ เป็นต้น โดยผลกระทบต่อการสอบระดับที่สูงขึ้นนี้อาจจะก่อให้เกิดผลเสียในระดับที่รุนแรงต่อ สถาบัน องค์กร สังคม รวมถึงประเทศด้วย ดังนั้น การป้องกันไม่ให้เกิดการลอกข้อสอบจึงมีความสำคัญและจำเป็นอย่างยิ่งต่อการจัดสอบทุกระดับ วิธีการป้องกันไม่ให้เกิดการลอกข้อสอบที่ใช้กันโดยทั่วไป ได้แก่ การจัดการคุมสอบ แต่บางครั้งแม้ผู้คุมสอบจะสังเกตพฤติกรรมบางอย่างซึ่งส่อว่าผู้สอบอาจทำการลอกข้อสอบแต่ผู้คุมสอบไม่สามารถหาหลักฐานยืนยันได้ว่าผู้สอบทำการลอกข้อสอบกันจริง ทำให้ผู้ที่ทำการลอกข้อสอบกันไม่ได้รับการลงโทษใด ๆ ซึ่งการที่นักเรียนผู้ที่ทำการลอกข้อสอบไม่ได้รับการลงโทษใด ๆ เนื่องจากไม่มีหลักฐานยืนยันการลอกข้อสอบที่เพียงพอนั้นไม่เพียงแต่จะส่งผลต่อความถูกต้องของการวัดและประเมินผลการเรียนเท่านั้น แต่ยังส่งผลต่อคุณลักษณะนิสัยของผู้เรียนที่จะไม่เกรงกลัวและละอายต่อการทุจริต

ในการสอบอีกด้วย ในปัจจุบันได้มีการพัฒนาวิธีการที่จะสามารถตรวจสอบได้ว่าผู้สอบที่ถูกสงสัยว่าจะทำการลอกข้อสอบกันนั้นมีแนวโน้มที่จะลอกข้อสอบกันหรือไม่ วิธีการดังกล่าว คือ วิธีการทางสถิติ ซึ่งวิธีการทางสถิตินี้จะทำให้ได้หลักฐานเพิ่งเติมเกี่ยวกับการลอกข้อสอบที่หนักแน่นและเพียงพอ (Sotaridona and Meijer, 2002)

หนึ่งในวิธีการทางสถิติดังที่กล่าวข้างต้นนั้นคือ ค่าสถิติตรวจจับการลอกข้อสอบ หรือ ดัชนีตรวจจับการลอกข้อสอบ (Answer-Copying statistic) โดยดัชนีตรวจจับการลอกข้อสอบนั้นถูกคิดค้นและนำเสนอมาเป็นเวลามากกว่า 25 ปีแล้ว 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) โดยดัชนีตรวจจับการลอกข้อสอบแต่ละตัวคำนวณอยู่บนพื้นฐานทางทฤษฎีที่แตกต่างกัน และ ดัชนีตรวจจับการลอกข้อสอบแต่ละตัวจะมีประสิทธิภาพที่แตกต่างกันไปตามสถานการณ์การสอบที่ต่างกัน

ด้วยเหตุผลดังที่กล่าวข้างต้น ผู้วิจัยจึงสนใจที่จะทำการวิเคราะห์คุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 (type I error) และ อำนาจการตรวจจับการลอกข้อสอบ (detection rates) ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω และทำการเปรียบเทียบคุณสมบัติทางสถิติระหว่างดัชนีตรวจจับการลอกข้อสอบชนิดต่าง ๆ ดังกล่าวภายใต้เงื่อนไขตัวแปรต้นที่แตกต่างกัน เพื่อให้ได้สารสนเทศเกี่ยวกับสภาพเงื่อนไขหรือสถานการณ์ที่มีประสิทธิภาพที่สุดสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละตัว ซึ่งจะเป็นประโยชน์ต่อการเลือกใช้ค่าดัชนีตรวจจับการลอกข้อสอบที่เหมาะสมของผู้ที่จะนำเอาดัชนีตรวจจับการลอกข้อสอบไปใช้ในสภาพการสอบแบบต่าง ๆ

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

ขอบเขตการวิจัย

1. ตัวแปรที่ศึกษา

1.1 ตัวแปรต้น ประกอบด้วย

1) ความยาวของแบบสอบ แบ่งออกเป็นแบบสอบ 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) และอำนาจของการตรวจจับการลอกข้อสอบ (detection rate)

2. ประชากร คือ นิสิตระดับปริญญาบัณฑิตที่ลงทะเบียนเรียนในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษา ต้น ปีการศึกษา 2548 จำนวน 250 คน

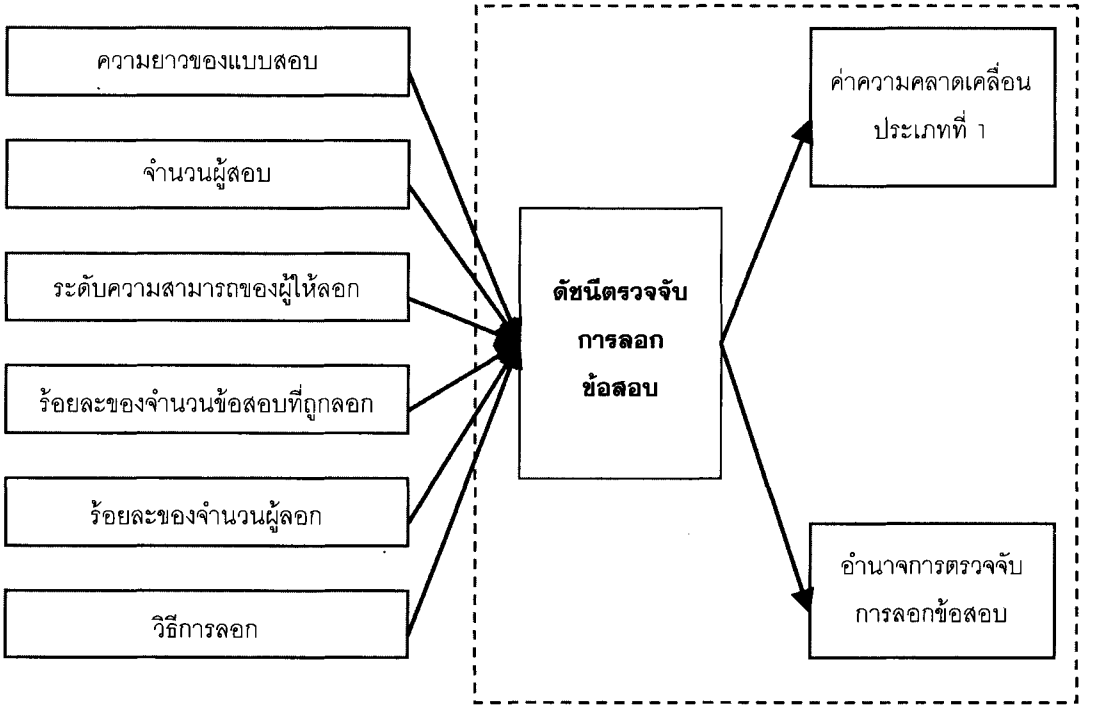
ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ได้ข้อมูลเกี่ยวกับคุณสมบัติทางสถิติ ซึ่งได้แก่ ความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบ ของดัชนีตรวจจับการลอกข้อสอบ อันประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω ภายใต้สถานการณ์ที่ต่างกันในเงื่อนไขตัวแปรต้นด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก ซึ่งจะทำให้ผู้ประสงค์ที่จะนำดัชนีตรวจจับการลอกข้อสอบไปใช้สามารถเลือกใช้ค่าดัชนีตรวจจับการลอกข้อสอบให้เหมาะสมกับสถานการณ์การสอบของตน

2. ข้อค้นพบที่ได้จากการวิเคราะห์และเปรียบเทียบค่าดัชนีตรวจจับการลอกข้อสอบสามารถนำไปใช้ในการจัดแผนผังที่นั่งสอบ หรือ สภาพแวดล้อมในการสอบเพื่อหลีกเลี่ยงการเกิดการลอกข้อสอบได้อย่างมีประสิทธิภาพ

กรอบแนวคิดในการวิจัย

กรอบแนวคิดในการวิจัยสามารถเขียนได้ ดังนี้



หมายเหตุ กรอบเส้นประสี่เหลี่ยมแสดงถึงการศึกษาและการเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบตรวจจับการลอกข้อสอบ

ภาพที่ 1 กรอบความคิดของการวิจัย

คำจำกัดความของดัชนีตรวจจับการลอกข้อสอบ

ดัชนี K_2 หมายถึง ดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบทวินาม (Binomial distribution) โดยหากค่าดัชนี K_2 ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้วผู้สอบในคู่ผู้สอบซึ่งถูกสงสัยว่าจะเป็นผู้ลอกจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก

ดัชนี S_1 หมายถึง ดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบปัวซง(Poisson distribution) โดยหากค่าดัชนี S_1 ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้วผู้สอบในคู่ผู้สอบซึ่งถูกสงสัยว่าจะเป็นผู้ลอกจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก

ดัชนี S_2 หมายถึง ดัชนีตรวจจับการลอกข้อสอบซึ่งจะถูกคำนวณโดยมีพื้นฐานอยู่บนการแจกแจงของจำนวนคำตอบผิดและคำตอบถูกที่ตรงกันระหว่างผู้ให้ลอกและผู้ลอกเป็นการแจกแจงแบบปัวซง (Poisson distribution) โดยหากค่าดัชนี S_2 ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าระดับนัยสำคัญที่กำหนดแล้วผู้สอบในคู่ผู้สอบซึ่งถูกสงสัยว่าจะเป็นผู้ลอกจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอก

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

วิธีดำเนินการวิจัย

การศึกษาเรื่อง การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ ใช้ระเบียบวิธีวิจัยเชิงทดลอง (Experimental Research) โดยผู้วิจัยใช้ข้อมูลผลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษาต้น ปีการศึกษา 2548 ส่วนข้อสอบปรนัย ของนิสิตระดับปริญญาบัณฑิตที่ลงทะเบียนเรียนและเข้าสอบในรายวิชาดังกล่าว จำนวนทั้งสิ้น 250 คน ซึ่งเป็นข้อมูลเชิงประจักษ์ (empirical data) มาจัดกระทำด้วยวิธีการจำลองสถานการณ์ (Simulation) และ วิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์ S-Plus 2000 (S-Plus 2000, MathSoft, Inc.) ซึ่งผู้วิจัยเป็นผู้เขียนคำสั่งที่ใช้ในการวิเคราะห์ประมวลผลสำหรับดัชนีตรวจจับการลอกข้อสอบแต่ละดัชนีภายใต้ความแตกต่างของเงื่อนไขตัวแปรต้นในด้าน ความยาวของแบบสอบ จำนวนผู้สอบ ระดับความสามารถของผู้ให้ลอก ร้อยละของจำนวนข้อสอบที่ถูกลอก ร้อยละของจำนวนผู้ลอก และ วิธีการลอก

ประชากร กลุ่มตัวอย่าง และ การสุ่มกลุ่มตัวอย่าง

ประชากร ได้แก่ นิสิตระดับปริญญาบัณฑิต ที่ลงทะเบียนและเข้าสอบในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา

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

เนื่องจากขนาดของกลุ่มผู้สอบเป็นตัวแปรต้นหนึ่งที่ผู้วิจัยศึกษาซึ่งแบ่งออกเป็นสองขนาด คือ 100 คน และ 250 คน ดังนั้น สำหรับกลุ่มผู้สอบขนาดใหญ่ คือ 250 คน ผู้วิจัยจะไม่ทำการสุ่ม แต่ผู้วิจัยจะใช้ผลการสอบของนิสิตที่ลงทะเบียนเรียนและเข้าสอบในรายวิชา 2702303 ภาคการศึกษาต้นปีการศึกษา 2548 ทั้งหมด ส่วนกลุ่มผู้สอบขนาดเล็ก คือ 100 คน ผู้วิจัยจะใช้วิธีการสุ่มอย่างง่ายสุ่มผู้สอบจำนวน 100 คนจากกลุ่มผู้สอบจำนวน 250 คนที่เป็นนิสิตที่ลงทะเบียนเรียนและเข้าสอบในรายวิชา 2702303 ภาคการศึกษาต้น ปีการศึกษา 2548

การเก็บรวบรวมข้อมูล และ เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบปลายภาคประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษาต้น ปีการศึกษา 2548 ซึ่งถือเป็นแบบสอบมาตรฐาน โดยมีค่าเฉลี่ยของค่าความยากง่ายของข้อสอบเป็น 0.682 ค่าเฉลี่ยของอำนาจจำแนกของข้อสอบเป็น 0.233 และ ค่าความเที่ยงแบบสอดคล้องภายในวิธีแบ่งครึ่งข้อสอบเป็นครึ่งแรก-ครึ่งหลังเป็น 0.79

การเก็บรวบรวมข้อมูล งานวิจัยชิ้นนี้ผู้วิจัยไม่ได้ดำเนินการเก็บข้อมูลจริง แต่เป็นการนำเอาผลการตอบข้อสอบส่วนข้อสอบปรนัย ของนิสิตที่ลงทะเบียนและเข้าสอบ ในรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษาต้น ปีการศึกษา 2548 จำนวน 250 คน ซึ่งเป็นข้อมูลเชิงประจักษ์ (empirical data) มาใช้ในการทดลอง และ วิเคราะห์ข้อมูล

ขั้นตอนการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (Simulation)

การวิจัยนี้ศึกษาตัวแปรต้น 6 ตัวแปร ได้แก่ 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

เนื่องจากในงานวิจัยนี้ผู้วิจัยทำการศึกษาตัวแปรต้นหลายตัวและตัวแปรต้นแต่ละตัวมีระดับย่อยของตัวแปรต้นต่างกัน ดังนั้น ผู้วิจัยจึงใช้สัญลักษณ์ย่อภาษาอังกฤษแทนตัวแปรต้นและระดับย่อยของตัวแปรต้นที่ผู้วิจัยศึกษา ซึ่งสามารถศึกษาของสัญลักษณ์ย่อภาษาอังกฤษได้จากตารางที่ 1

ตารางที่ 1 สัญลักษณ์ย่อภาษาอังกฤษแทนตัวแปรต้นที่ศึกษาในงานวิจัย

ตัวแปรต้น	สัญลักษณ์ย่อ	ระดับย่อยของตัวแปรต้น	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)

โดยการจัดกระทำข้อมูลในแต่ละสถานการณ์มีรายละเอียดดังนี้

ปัจจัยด้าน ความยาวของแบบสอบ และ จำนวนผู้สอบ

นำผลการตอบข้อสอบปลายภาค ประจำรายวิชา 2702303 การวัดและการประเมินผลทางการศึกษา ภาคการศึกษาต้น ปีการศึกษา 2548 ในส่วนข้อสอบหลายตัวเลือก (4 ตัวเลือก) จำนวน 65 ข้อ ของผู้สอบจำนวน 250 คนมาทำการสุ่มผู้สอบจำนวน 100 คน และ สุ่มข้อสอบ 35 ข้อจากแบบสอบความยาว 65 ข้อ ด้วยวิธีการสุ่มอย่างง่าย เพื่อสร้างชุดข้อมูล 4 ชุด ตามปัจจัยด้านความยาวของแบบสอบ และ จำนวนผู้สอบ ได้แก่ ชุดข้อมูลแบบสอบขนาด 35 ข้อ จำนวนผู้สอบ 100 คน ชุดข้อมูลแบบสอบขนาด 35 ข้อ จำนวนผู้สอบ 250 คน ชุดข้อมูลแบบสอบขนาด 65 ข้อ จำนวนผู้สอบ 100 คน และ ชุดข้อมูลแบบสอบขนาด 65 ข้อ จำนวนผู้สอบ 250 คน จากนั้นนำข้อมูลทั้ง 4 ชุดไปวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ (a, b, c) พารามิเตอร์ระดับความสามารถของผู้สอบ (θ) ตามทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories) และ วิเคราะห์หาค่าพารามิเตอร์ความชันของแต่ละรายการคำตอบ (Slope of the trace lines) และค่าพารามิเตอร์จุดตัดของแต่ละรายการคำตอบ (Intercept parameter) ตามทฤษฎีการทดสอบแนวใหม่แบบโมเดล Nominal

Response Model ด้วยโปรแกรม MULTILOG เพื่อนำผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบไปใช้ในการคัดเลือกข้อสอบที่จะอยู่ในกลุ่มข้อสอบระดับยาก ทำการเรียงลำดับผู้สอบในแต่ละชุดข้อมูลตามค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) จากมากไปน้อย และ นำค่าพารามิเตอร์ความชันและค่าพารามิเตอร์จุดตัดของแต่ละรายการคำตอบที่ได้ไปวิเคราะห์หาดัชนี ω

ปัจจัยด้านระดับความสามารถของผู้ให้ลอก และ ร้อยละของจำนวนผู้ลอก

ในแต่ละชุดข้อมูลทั้ง 4 ชุดข้างต้นให้เลือกผู้สอบที่อยู่ ณ ตำแหน่งเปอร์เซ็นไทล์ที่ 90 และ 60 มาเพื่อให้เป็น ผู้ให้ลอก (source) จากนั้นทำการสุ่มผู้ลอก (copiers) ด้วยวิธีการสุ่มอย่างง่าย ซึ่งการสุ่มผู้ลอกมีข้อตกลงเบื้องต้นว่า ให้สุ่มผู้สอบที่มีตำแหน่งเปอร์เซ็นไทล์ ต่ำกว่าตำแหน่งเปอร์เซ็นต์ไทล์ของผู้ให้ลอกเท่านั้น โดยสุ่มผู้ลอกมา 2 ขนาดคือ ขนาดร้อยละ 5 และ ร้อยละ 10 ของผู้สอบทั้งหมด โดยในแต่ละกลุ่มเงื่อนไขนั้นผู้ให้ลอกจะมีเพียง 1 คนเท่านั้น (ผู้สอบที่มีตำแหน่งเปอร์เซ็นไทล์ที่ 60 หรือ 90) ส่วนผู้ลอกจะมีจำนวนร้อยละ 5 หรือ 10 ของผู้สอบทั้งหมดขึ้นอยู่กับเงื่อนไขตัวแปรต้นด้านร้อยละของจำนวนผู้ลอกที่ต่างกัน

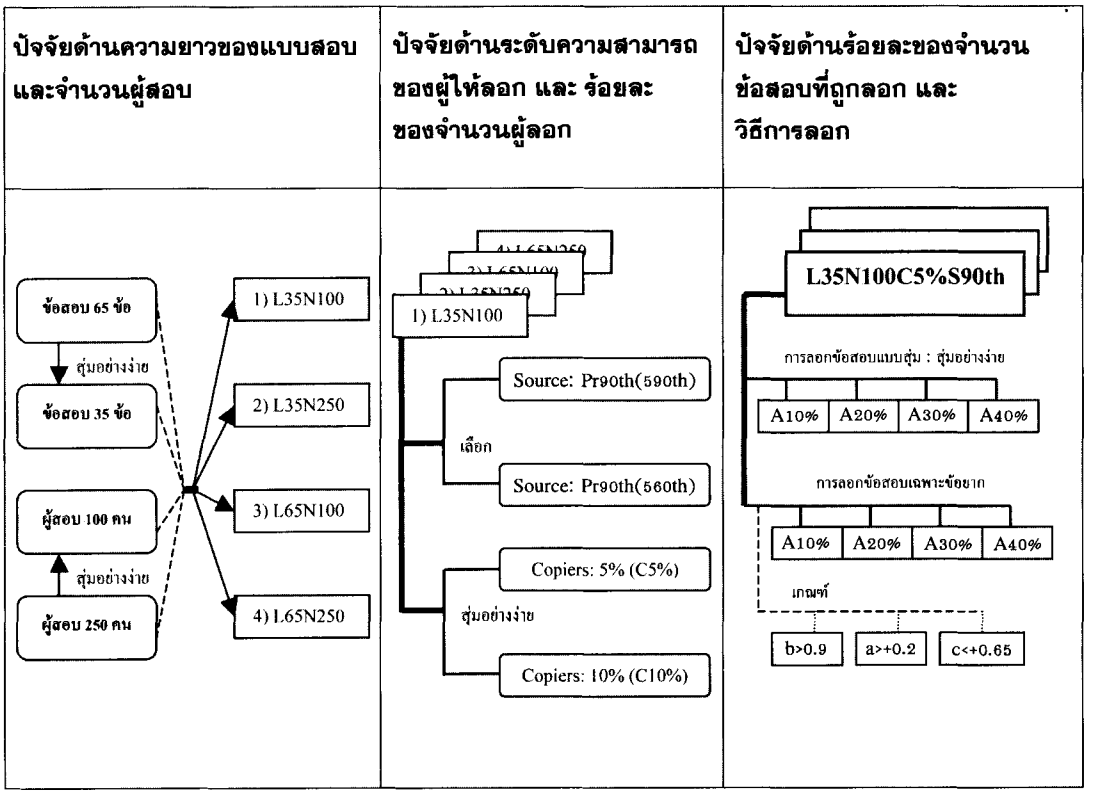
ปัจจัยด้านร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอก

จากนั้นในแต่ละชุดข้อมูลที่ถูกแยกย่อยออกมา ก็จะต้องทำการศึกษาเพิ่มเติมในเงื่อนไขเกี่ยวกับ ร้อยละของจำนวนข้อสอบที่ถูกลอก ซึ่งแบ่งเป็น 4 ระดับคือ ร้อยละ 10, 20, 30 และ 40 ของข้อสอบทั้งหมด และ วิธีการลอกที่แตกต่างกัน ซึ่งแบ่งออกเป็น 2 วิธี คือ การลอกข้อสอบแบบสุ่ม และการลอกข้อสอบเฉพาะข้อยาก ในส่วนเงื่อนไขการลอกข้อสอบแบบสุ่มผู้วิจัยใช้วิธีการสุ่มอย่างง่ายเพื่อให้ได้ข้อสอบตามจำนวนร้อยละของจำนวนข้อสอบที่ถูกลอกทั้ง 4 ระดับ และในส่วนของเงื่อนไขวิธีการลอกข้อสอบเฉพาะข้อยาก 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

จากนั้นใช้วิธีการสุ่มอย่างง่ายในการสุ่มข้อสอบในกลุ่มข้อสอบระดับยากให้ได้จำนวนเท่ากับจำนวนร้อยละของจำนวนข้อสอบที่ถูกลอกทั้ง 4 ระดับ

โดยขั้นตอนการจัดกระทำข้อมูลในแต่ละสถานการณ์ตามปัจจัยทั้ง 6 ตัวแปรที่ผู้วิจัยศึกษาสามารถแสดงเป็นภาพที่ 2



ภาพที่ 2 ขั้นตอนการจัดกระทำข้อมูลในแต่ละสถานการณ์ตามปัจจัยที่ผู้วิจัยศึกษา

วิธีการจัดกระทำข้อมูลด้วยวิธีการจำลองสถานการณ์ (Simulation)

ผู้วิจัยจัดกระทำกับข้อมูลด้วยวิธีการจำลองสถานการณ์ (Simulation) ซึ่งเป็นการเปลี่ยนคำตอบของผู้ลอกให้ตรงกับคำตอบของผู้ให้ลอกโดยเปลี่ยนคำตอบให้สอดคล้องกับเงื่อนไขที่ศึกษาคือ ร้อยละของจำนวนข้อสอบที่ถูกลอก และ วิธีการลอกข้อสอบ

การวิเคราะห์ซ้ำ (Replication)

การวิเคราะห์หาค่าดัชนีตรวจจับการลอกข้อสอบ ดัชนี K_2, ดัชนี S_1 ดัชนี S_2 และดัชนี ω นั้น ผู้วิจัยจะทำการวิเคราะห์ข้อมูลของแต่ละสถานการณ์จาก 128 สถานการณ์ซ้ำจำนวน 100 รอบ เพื่อเป็นการตรวจสอบความเที่ยงและถูกต้องของผลการวิเคราะห์ ดังนั้น ในแต่ละสถานการณ์ที่ถูกศึกษา ผู้วิจัยจำเป็นจะต้องสร้างชุดข้อมูลซึ่งมีลักษณะการจัดกระทำข้อมูลตรงตามลักษณะสถานการณ์นั้น ๆ จำนวน 100 ชุด เพื่อใช้ในการวิเคราะห์ซ้ำ 100 รอบ โดยชุดข้อมูล 100 ชุด ของสถานการณ์ที่ศึกษาหนึ่ง ๆ นั้น ผู้วิจัยกำหนดให้แต่ละชุดข้อมูลมีความแตกต่างกันในด้าน ผู้ลอก ซึ่งข้อมูลแต่ละ

ชุดมีผู้สอบที่ถูกกำหนดให้เป็นผู้ลอกแตกต่างกันไป และข้อสอบซึ่งข้อมูลแต่ละชุดข้อสอบที่ถูกกำหนดให้เป็นข้อสอบที่ถูกลอกจะแตกต่างกันไป

เนื่องจากงานวิจัยนี้มีสถานการณ์ที่ต้องจัดกระทำข้อมูลรวมทั้งสิ้น 128 สถานการณ์ ซึ่งในแต่ละสถานการณ์ต้องสร้างชุดข้อมูล 100 ชุดเพื่อวิเคราะห์ซ้ำ ดังนั้น ผู้วิจัยจึงสร้างชุดข้อมูลที่ผ่านการจัดกระทำด้วยการจำลองสถานการณ์ (simulation) แล้วทั้งสิ้น 12,800 ชุด เพื่อนำไปวิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์ S-Plus 2000

การระบุการลอกของดัชนีตรวจจับการลอกข้อสอบ

สำหรับดัชนี K_2 ดัชนี S_1 และ ดัชนี S_2 นั้น ผู้สอบจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอกก็ต่อเมื่อค่าดัชนีที่คำนวณได้มีค่า *น้อยกว่าหรือเท่ากับ* ค่าระดับนัยสำคัญที่ผู้วิจัยกำหนดเป็นเกณฑ์ในการเทียบค่าดัชนีตรวจจับการลอกข้อสอบเพื่อระบุการลอก สำหรับดัชนี ω ผู้สอบจะถูกระบุหรือถูกบ่งชี้ว่าเป็นผู้ลอกก็ต่อเมื่อดัชนี ω มีค่า *มากกว่า* ค่าวิกฤติของการทดสอบสมมติฐานทางเดียว (one-tailed critical value) ซึ่งขอบเขตวิกฤติมีพื้นที่อยู่ทางด้านขวาของโค้งปกติ ณ ระดับนัยสำคัญที่ผู้วิจัยกำหนดเป็นเกณฑ์ โดยผู้วิจัยใช้ระดับนัยสำคัญจำนวน 7 ค่า ดังนี้ .0001, .0005, .001, .0025, .005, .01 และ .05

คุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ

คุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ตัวนั้นประกอบด้วยความคลาดเคลื่อนประเภทที่ 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)

2. อัตราตรวจจับการลอกข้อสอบ (Detection Rates)

หลักการคำนวณค่าอำนาจการตรวจจับการลอกข้อสอบ คือ การคำนวณสัดส่วนของผู้ลอกจริง (true copier) จะถูกระบุว่าเป็นผู้ลอกหรือถูกตรวจจับได้โดยค่าดัชนีตรวจจับการลอกข้อสอบ

การศึกษาคุณสมบัติทางสถิติของค่าดัชนีตรวจจับการลอกข้อสอบ

ผู้วิจัยนำค่าความคลาดเคลื่อนประเภทที่ 1 และอำนาจการตรวจจับการลอกข้อสอบซึ่งได้จากการวิเคราะห์ด้วยโปรแกรม S-Plus 2000 มาวาดเป็นกราฟเส้นด้วยโปรแกรม Microsoft Excel

เพื่อให้เกิดความง่ายต่อการศึกษาทำความเข้าใจ และง่ายต่อการเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ สำหรับค่าความคลาดเคลื่อนประเภทที่ ๑ กราฟเส้นที่วาดนั้นเป็นกราฟเส้นแสดงค่าความคลาดเคลื่อนประเภทที่ ๑ ของดัชนีตรวจจับการลอกข้อสอบแต่ละตัว ณ ระดับนัยสำคัญต่าง ๆ ที่ผู้วิจัยศึกษา โดยผู้วิจัยวาดเส้นกราฟขอบเขต (**Boundary line**) ซึ่งเป็นเส้นกราฟที่ค่าความคลาดเคลื่อนประเภทที่ ๑ ณ ทุกระดับนัยสำคัญที่ผู้วิจัยศึกษามีค่าเท่ากับดัชนีนัยสำคัญต่าง ๆ เพิ่มอีกหนึ่งเส้น ซึ่งเส้นกราฟขอบเขตจะแสดงขอบเขตค่าความคลาดเคลื่อนที่ยอมรับได้ และสำหรับค่าอำนาจการตรวจจับการลอกข้อสอบ กราฟเส้นที่วาดนั้นเป็นกราฟเส้นแสดงค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบแต่ละตัว ณ ระดับนัยสำคัญต่าง ๆ ที่ผู้วิจัยศึกษา

สำหรับกราฟเส้นความคลาดเคลื่อนประเภทที่ ๑ หากกราฟเส้นค่าความคลาดเคลื่อนประเภทที่ ๑ อยู่เหนือเส้นกราฟขอบเขต เรียกว่า **liberal** แต่หากเส้นกราฟค่าความคลาดเคลื่อนประเภทที่ ๑ อยู่ต่ำกว่าเส้นกราฟขอบเขต เรียกว่า **conservative** สำหรับกราฟเส้นอำนาจการตรวจจับการลอกข้อสอบแต่ละเส้นแสดงถึงแนวโน้มของค่าอำนาจการตรวจจับการลอกข้อสอบที่เปลี่ยนไปตามระดับนัยสำคัญของแต่ละค่าดัชนีตรวจจับการลอกข้อสอบ

โดยดัชนีตรวจจับการลอกข้อสอบที่ดีควรมีค่าความคลาดเคลื่อนประเภทที่ ๑ ซึ่งต่ำและมีค่าอำนาจการตรวจจับการลอกข้อสอบที่สูง

การเปรียบเทียบคุณสมบัติทางสถิติของดัชนีตรวจจับการลอกข้อสอบ

การเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ ๑ ระหว่างดัชนีตรวจจับการลอกข้อสอบพิจารณาเปรียบเทียบ โดยมีหลักการคือ ค่าความคลาดเคลื่อนประเภทที่ ๑ ของดัชนีตรวจจับการลอกข้อสอบที่ดีต้องมีค่าสูงหรือต่ำกว่าเส้นกราฟขอบเขตเพียงเล็กน้อยหากกราฟค่าความคลาดเคลื่อนประเภทที่ ๑ ของดัชนีตรวจจับการลอกข้อสอบใดที่สามารถควบคุมระดับให้มีค่าสูงหรือต่ำกว่าเส้นกราฟขอบเขตเพียงเล็กน้อยได้สำหรับทุกระดับนัยสำคัญที่กำหนด และเส้นกราฟอยู่ใกล้เส้นกราฟขอบเขตที่สุด ดัชนีตรวจจับการลอกข้อสอบนั้นเป็นดัชนีตรวจจับการลอกข้อสอบที่สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ ๑ ดีที่สุดสำหรับสถานการณ์นั้น ๆ ส่วนเส้นกราฟของดัชนีตรวจจับการลอกข้อสอบที่อยู่ใต้เส้นกราฟขอบเขตถัดลงมาจากเส้นกราฟของดัชนีตรวจจับการลอกข้อสอบที่สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ ๑ ดีที่สุดนั้นจะสามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ ๑ ดีรองลงมาตามลำดับ

การเปรียบเทียบอำนาจการตรวจจับการลอกข้อสอบระหว่างดัชนีตรวจจับการลอกข้อสอบพิจารณาเปรียบเทียบโดยมีหลักการว่า หากกราฟเส้นของดัชนีตรวจจับการลอกข้อสอบใดมีค่าสูงสุดหรือ อยู่เหนือกราฟเส้นของดัชนีตรวจจับการลอกข้อสอบอื่น ๆ ดัชนีตรวจจับการลอกข้อสอบนั้น ๆ

เป็นดัชนีตรวจจับการลอกข้อสอบที่มีอำนาจการตรวจจับการลอกข้อสอบที่ดีที่สุดสำหรับสถานการณ์นั้น ๆ ส่วนเส้นกราฟของดัชนีตรวจจับการลอกข้อสอบที่อยู่ถัดลงมาจากเส้นที่อยู่สูงที่สุดนั้น ก็จะมีระดับอำนาจการตรวจจับที่ดีรองลงมาตามลำดับ

ผลการวิจัย

ผู้วิจัยแบ่งการนำเสนอผลการวิจัยออกเป็น 2 ส่วน คือ ผลการศึกษาและเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 และผลการศึกษาและเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบ ของค่าดัชนีตรวจจับการลอกข้อสอบ รายละเอียดดังนี้

ผลการศึกษาและเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 ของค่าดัชนีตรวจจับการลอกข้อสอบ

1. ดัชนี S_2 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ และมีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ

2. ดัชนี ω สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ และมีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ

3. ดัชนี K_2 และ ดัชนี S_1 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในทุกสถานการณ์

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

ตารางที่ 4 ค่าสัมประสิทธิสหสัมพันธ์แบบเพียร์สัน ระหว่างตัวแปรสังเกตได้ของกลุ่มตัวอย่างทั้งหมด (n=430)

ตัวแปร	DEC	PROF	STA	SELF	AUTO	IMPA	LEAD	MOTI	TEAM	COOP	ATMO
DEC	1.000										
PROF	.493**	1.000									
STA	.567**	.437**	1.000								
SELF	.380**	.428**	.479**	1.000							
AUTO	.496**	.476**	.526**	.533**	1.000						
IMPA	.265**	.126**	.287**	.121*	.315**	1.000					
LEAD	.506**	.396**	.369**	.356**	.455**	.105*	1.000				
MOTI	.559**	.417**	.308**	.310**	.387**	.128**	.670**	1.000			
TEAM	.539**	.442**	.370**	.349**	.453**	.147**	.601**	.718**	1.000		
COOP	.581**	.406**	.314**	.289**	.391**	.135**	.584**	.754**	.745**	1.000	
ATMO	.558**	.479**	.359**	.315**	.458**	.200**	.567**	.725**	.743**	.797**	1.000
$\bar{X}$	3.508	3.933	3.401	4.019	3.719	2.928	3.860	3.723	3.789	3.665	3.752
S.D.	.619	.643	.663	.519	.654	.937	.572	.612	.623	.690	.676

(a) (b)

Bartlett's Test of Sphericity = 2623.853 df = 55 P = .000 Kaiser-Meyer-Olkin Measure of Sampling Adequacy = .910

(c) (d)

ภาพที่ 3 การเปรียบค่าความคลาดเคลื่อนประเภทที่ ๑ เชิงประจักษ์ (empirical type I error) ณ ระดับนัยสำคัญ 7 ระดับ ของดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และ ดัชนี ω (Omega) โดยแยกตามตัวแปรขนาดจำนวนผู้สอบ และความยาวของแบบสอบ

ผลการศึกษาและเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบของค่าดัชนีตรวจจับการลอกข้อสอบ

๑. ตัวแปรความยาวของแบบสอบมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω แต่จะไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2, ดัชนี S_1 และดัชนี S_2

2. ตัวแปรจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่

3. ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่

4. ตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เล็กน้อยแต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้

5. ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับ

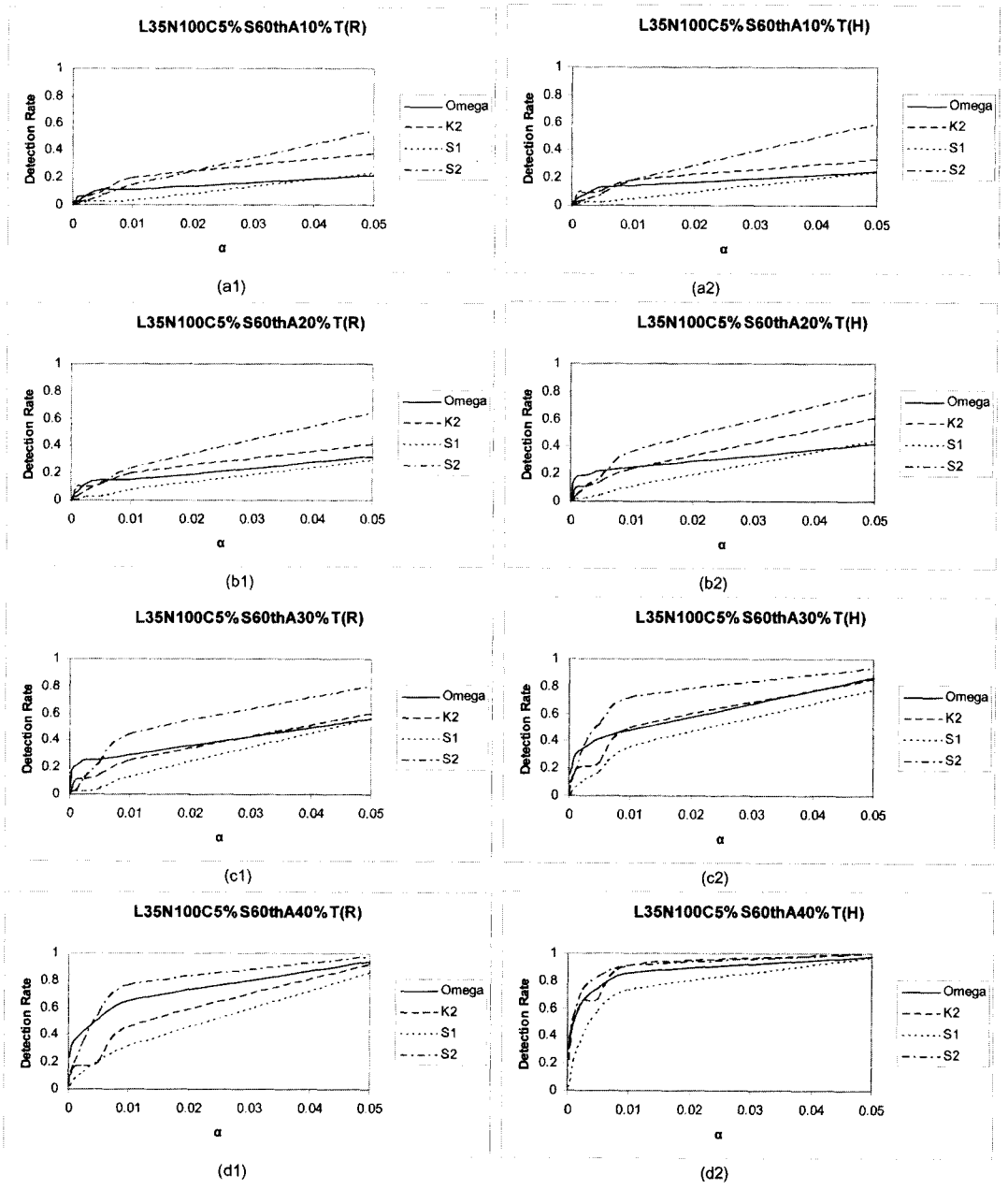
การลอกข้อสอบทั้งสี่ กรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อยากจะสูงกว่ากรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบแบบสุ่มเล็กน้อย

6. ตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่มีแนวโน้มที่จะมีค่าสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น

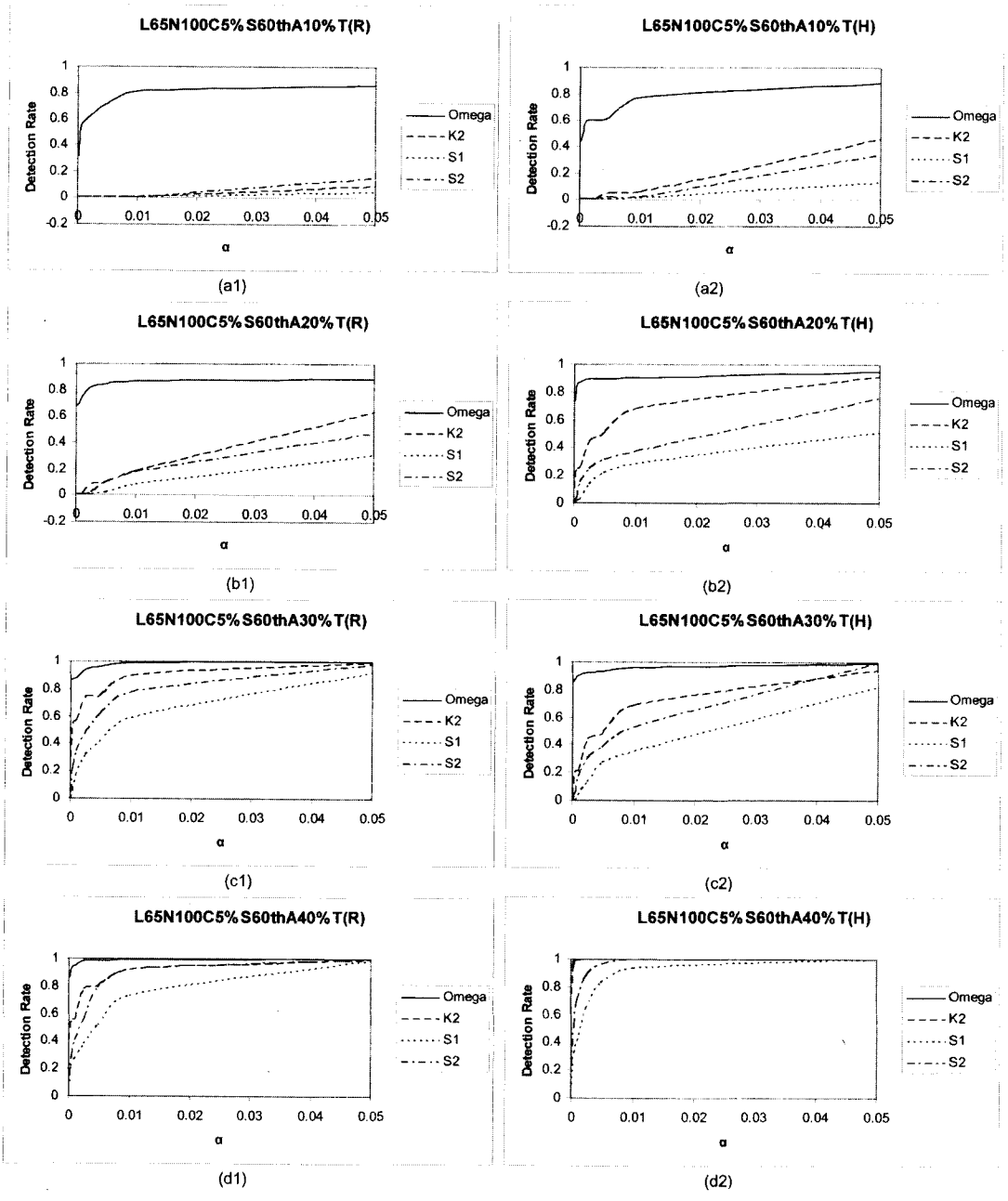
7. ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ดัชนี S_2 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด

8. ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ดัชนี ω เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด

9. ในเกือบทุกสถานการณ์ที่ถูกจำลองขึ้น ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบต่ำที่สุด



ภาพที่ 4 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



ภาพที่ 5 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

อภิปรายผล

ผลการศึกษาและเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 (type I error) ของค่าดัชนีตรวจจับการลอกข้อสอบ

จากภาพที่ 3 แสดงให้เห็นว่า ดัชนี S_2 สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ และมีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวแบบสอบเป็น 35 ข้อ โดยดัชนี ω สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ และมีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ ซึ่งข้อค้นพบนี้แตกต่างจากผลการวิจัยของ Wollack (1997) ที่พบว่า ดัชนี ω สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ใต้เส้นขอบเขตในทุกสถานการณ์ ส่วนดัชนี K_2 และ ดัชนี S_1 นั้นสามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในทุกสถานการณ์ แต่ระดับค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนี S_1 ต่ำกว่าเส้นขอบเขตมากกว่าของดัชนี K_2 กล่าวได้ว่า ดัชนี S_1 เป็นดัชนีที่มีระดับค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำที่สุดในทุกสถานการณ์ ซึ่งจะส่งผลต่ออำนาจการตรวจจับการลอกข้อสอบของดัชนี S_1 ที่จะมีค่าน้อย ดังจะเห็นได้จากสรุปผลการวิจัยส่วนผลการศึกษาและเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบของค่าดัชนีตรวจจับการลอกข้อสอบ ที่พบว่า ดัชนี S_1 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบที่ต่ำที่สุดในเกือบทุกสถานการณ์ ซึ่งสอดคล้องกับหลักการที่ว่า ค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบไม่ควรที่จะมีค่าต่ำกว่าเส้นขอบเขตมาก เนื่องจากอำนาจการตรวจจับผู้สอบซึ่งเป็นผู้ลอกจริงจะต่ำมาก (Seaman, Levin และ Serlin, 1991 อ้างถึงใน Wollack, 1997)

ผลการวิจัยพบว่าระดับค่าความคลาดเคลื่อนประเภทที่ 1 ของดัชนีตรวจจับการลอกข้อสอบทั้งสี่มีแนวโน้มที่จะมีค่าสูงกว่าเส้นกราฟขอบเขตทั้งในระดับที่ยอมรับได้และไม่สามารถยอมรับได้ สาเหตุอาจเนื่องมาจากปัจจัยด้านจำนวนผู้สอบซึ่งงานวิจัยชิ้นนี้ศึกษาตัวแปรจำนวนผู้สอบ 2 ระดับ คือ 100 และ 250 คน ซึ่งเป็นตัวแทนของจำนวนผู้สอบกลุ่มเล็กและกลุ่มใหญ่ตามลำดับ แต่ในงานวิจัยของ Sotaridona และ Meijer (2003) จะศึกษาตัวแปรจำนวนผู้สอบ 2 ระดับ คือ 100 และ 500 คน ซึ่งเป็นตัวแทนของจำนวนผู้สอบกลุ่มเล็กและกลุ่มใหญ่ตามลำดับ โดยจำนวนผู้สอบซึ่งเป็นตัวแทนของผู้สอบกลุ่มใหญ่ในงานวิจัยชิ้นนี้ขนาดเป็นครึ่งหนึ่งของจำนวนผู้สอบที่เป็นตัวแทนของผู้สอบกลุ่มใหญ่ในงานวิจัยของ Sotaridona และ Meijer (2003) ซึ่งตัวแปรจำนวนผู้สอบที่ต่างกันนี้

อาจจะส่งผลต่อการคำนวณค่าดัชนีตรวจจับการลอกข้อสอบทั้งสี่ โดยเฉพาะ ดัชนี ω ที่มีพื้นฐานการคำนวณอยู่บนค่าพารามิเตอร์ที่ถูกประมาณค่าจากโมเดล Nominal Response Model ซึ่งค่าพารามิเตอร์ ที่ประมาณค่าได้จากโมเดลนี้จะมีความถูกต้องมากขึ้นหากมีจำนวนผู้สอบมากขึ้น

นอกจากนี้ข้อค้นพบที่ว่าดัชนี ω สามารถควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ และมีแนวโน้มที่จะมีระดับค่าความคลาดเคลื่อนประเภทที่ 1 สูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อนั้นแตกต่างจากผลการวิจัยของ Wollack (1997) ที่พบว่า ดัชนี ω สามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ใต้เส้นขอบเขตในทุกสถานการณ์ อาจเนื่องมาจากรูปแบบการทดลองหรือจัดกระทำข้อมูลที่ต่างกัน โดยในงานวิจัยของ Wollack นั้นมีการใช้แผนผังที่นั่งสอบ (seating chart) ในการจัดกระทำให้เกิดการลอกขึ้นแต่ในงานวิจัยชิ้นนี้ไม่ได้ใช้โดยแผนผังที่นั่งสอบถูกนำมาช่วยในการจับคู่ผู้ลอกและผู้ให้ลอกให้เป็นคู่ผู้สอบซึ่งนั่งอยู่ในรัศมีที่การลอกสามารถเกิดขึ้นได้เท่านั้น ซึ่งจากงานวิจัยของ Wollack (1997) พบว่าแผนผังที่นั่งสอบสามารถ ช่วยให้ค่าความคลาดเคลื่อนประเภทที่ 1 ลดลงได้

ผลการศึกษาและเปรียบเทียบค่าอำนาจการตรวจจับการลอกข้อสอบของค่าดัชนีตรวจจับการลอกข้อสอบ

ผลการวิจัยพบว่า ตัวแปรความยาวของแบบสอบมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω (ภาพที่ 4 และ 5) แต่ตัวแปรจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ซึ่งสอดคล้องกับผลการวิจัยของ Wollack (1997) และ Sotaridona และ Meijer (2002 และ 2003) แต่ตัวแปรความยาวของแบบสอบและจำนวนผู้สอบไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี K_2 ดัชนี S_1 และ ดัชนี S_2 (ภาพที่ 4 และ 5) 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

ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ แต่ผู้วิจัยคาดว่าตัวแปรร้อยละของจำนวนผู้ลอกน่าจะมีอิทธิพล

ต่อดัชนี ω หากนำไปใช้ในสถานการณ์จริง เนื่องจากค่าสถิติ ω เป็นค่าสถิติที่คำนวณโดยใช้ข้อมูลพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) ซึ่งคำนวณอยู่บนพื้นฐานของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) และ พารามิเตอร์จุดตัดและความชันซึ่งคำนวณโดยมีพื้นฐานอยู่บนโมเดล Nominal Response Model (Bock, 1972) ซึ่งพารามิเตอร์ที่กล่าวมาทั้งหมดนั้นจะมีค่าเปลี่ยนแปลงไปจากเดิมหรือบิดเบือนไปจากเดิมหากเกิดการลอกขึ้น โดยถ้าร้อยละของจำนวนผู้ลอกไม่เท่ากันแล้วจะส่งผลให้ดัชนี ω ที่คำนวณได้แตกต่างกันด้วย ซึ่งเมื่อดัชนี ω มีความแตกต่างกันอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω ย่อมต่างกันด้วย แต่สาเหตุที่ผลการวิเคราะห์ข้อมูลในงานวิจัยชิ้นนี้ พบว่า ตัวแปรร้อยละของจำนวนผู้ลอกไม่มีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนี ω อาจเนื่องมาจากผู้วิจัยทำการคำนวณค่าสถิติ ω โดยใช้ข้อมูลค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และ พารามิเตอร์จุดตัดและความชันซึ่งคำนวณโดยมีพื้นฐานอยู่บนโมเดล Nominal Response Model ของชุดข้อมูลที่ยังไม่จัดกระทำ (simulation) ให้เกิดการลอก ซึ่งดัชนี ω ที่คำนวณ โดยทราบค่าค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และ พารามิเตอร์จุดตัดและความชัน ซึ่งคำนวณโดยมีพื้นฐานอยู่บนโมเดล 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 ดังนั้น ดัชนี ω ที่คำนวณได้จะแตกต่างกันซึ่งจะส่งผลให้อำนาจการตรวจจับการลอกข้อสอบของดัชนี ω แตกต่างกันด้วย

ส่วนตัวแปรระดับความสามารถของผู้ให้ลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เพียงเล็กน้อยแต่ไม่สามารถระบุแนวโน้มของอิทธิพลที่ชัดเจนได้ และ ตัวแปรวิธีการลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่เล็กน้อย กล่าวคือ ค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ กรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบเฉพาะข้อยากจะสูงกว่ากรณีสถานการณ์ที่ตัวแปรวิธีการลอกเป็นการลอกข้อสอบแบบสุ่มเล็กน้อย (ภาพที่ 4 และ 5) ซึ่งข้อค้นพบนี้สำหรับดัชนี ω นั้นสอดคล้องกับงานวิจัยของ Wollack (1997)

ตัวแปรร้อยละของจำนวนข้อสอบที่ถูกลอกมีอิทธิพลต่อค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่ โดยค่าอำนาจการตรวจจับการลอกข้อสอบของดัชนีตรวจจับการลอกข้อสอบทั้งสี่มีแนวโน้มที่จะมีค่าสูงขึ้นหากร้อยละของจำนวนข้อสอบที่ถูกลอกมีค่ามากขึ้น โดยอาจกล่าวได้ว่า ผู้สอบที่ลอกข้อสอบจำนวนน้อยมีแนวโน้มต่ำที่จะถูกตรวจจับว่าเป็นผู้ลอก แต่ผู้สอบที่ลอกข้อสอบจำนวนมากมีแนวโน้มสูงที่จะถูกตรวจจับว่าเป็นผู้ลอก (ภาพที่ 4 และ 5) ซึ่งข้อค้นพบนี้สอดคล้องกับงานวิจัยของ Wollack (1997) และ Sotaridona และ Meijer (2002 และ 2003)

นอกจากนี้ผลการวิเคราะห์ข้อมูลยังพบว่าในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ดัชนี S_2 เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด (ภาพที่ 4) ส่วนในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ดัชนี ω เป็นดัชนีตรวจจับการลอกข้อสอบที่มีค่าอำนาจการตรวจจับการลอกข้อสอบสูงที่สุด (ภาพที่ 5) แต่เมื่อเราพิจารณาในแง่ของการควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 จะพบว่า ดัชนี S_2 มีแนวโน้มที่ระดับค่าความคลาดเคลื่อนประเภทที่ 1 จะสูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 ข้อ (ภาพที่ 3 (a) และ (b)) ข้อส่วนดัชนี ω มีแนวโน้มที่ระดับค่าความคลาดเคลื่อนประเภทที่ 1 จะสูงกว่าระดับที่ยอมรับได้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 ข้อ (ภาพที่ 4 (c) และ (d)) กล่าวได้ว่า ดัชนี S_2 เป็นดัชนีตรวจจับการลอกข้อสอบที่ไม่เหมาะที่จะใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 35 และ ดัชนี ω เป็นดัชนีตรวจจับการลอกข้อสอบที่ไม่เหมาะที่จะใช้ในสถานการณ์ที่ตัวแปรความยาวของแบบสอบเป็น 65 เนื่องจากค่าดัชนีตรวจจับการลอกข้อสอบที่ดีควรมีอำนาจการตรวจจับการลอกข้อสอบที่สูงและสามารถควบคุมระดับความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่ยอมรับได้

จากข้อค้นพบทั้งในด้านความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการตรวจจับการลอกข้อสอบทั้งหมด ผู้วิจัยสามารถสรุปได้ว่า ไม่มีดัชนีตรวจจับการลอกข้อสอบตัวใดที่สามารถใช้ได้ดีในทุกๆ สถานการณ์ ดัชนีตรวจจับการลอกข้อสอบแต่ละตัวสามารถใช้ได้อย่างมีประสิทธิภาพในสถานการณ์ที่แตกต่างกัน

ข้อเสนอแนะในการนำไปใช้

1. เนื่องจากความเหมาะสมในการเลือกใช้ดัชนีตรวจจับการลอกข้อสอบในแต่ละสถานการณ์นั้นต้องพิจารณาใน 2 ส่วน คือ ความสามารถในการควบคุมระดับค่าความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ในระดับที่สามารถยอมรับได้ กล่าวคือ ค่าความคลาดเคลื่อนประเภทที่ 1 ณ ระดับนัยสำคัญต่างๆ มีค่าสูงหรือต่ำกว่าเส้นขอบเขตเพียงเล็กน้อย เนื่องจากผลที่ตามมาของความคลาดเคลื่อนประเภทที่ 1 หรือ การระบุผู้ลอกผิดพลาดเป็นสิ่งที่รุนแรง และ การมีอำนาจการตรวจ

จับการลอกข้อสอบที่สูงเนื่องจากอำนาจการตรวจจับการลอกข้อสอบที่สูงแสดงให้เห็นถึงความสามารถในการตรวจจับผู้ลอกข้อสอบที่แท้จริงที่มีประสิทธิภาพ ซึ่งจากข้อค้นพบทั้งหมดในงานวิจัยนี้ผู้วิจัยสามารถสรุปแนวทางในการเลือกใช้ดัชนีตรวจจับการลอกข้อสอบ ได้ดังตารางที่ 2

ตารางที่ 2 แนวทางในการเลือกใช้ดัชนีตรวจจับการลอกข้อสอบภายใต้ปัจจัยด้านความยาวของแบบสอบ และจำนวนผู้สอบ

ปัจจัยการทดสอบ		ความยาวของแบบสอบ (L)	
		L35 (แบบสอบข้อคำถามน้อย)	L65 (แบบสอบข้อคำถามมาก)
จำนวนผู้สอบ (N)	N100 (กลุ่มผู้สอบขนาดเล็ก)	ω	K_2
	N250 (กลุ่มผู้สอบขนาดใหญ่)	S_1 (A10% และ A20%) ω (A30% และ A40%)	S_2

2. หากผู้นำดัชนีตรวจจับการลอกข้อสอบไปใช้ต้องการความแน่นอนหรือความเชื่อมั่นในผลการตรวจจับการลอกข้อสอบ ผู้วิจัยแนะนำให้ตรวจสอบการลอกข้อสอบด้วยดัชนีตรวจจับการลอกข้อสอบทั้ง 4 ดัชนี จากนั้นพิจารณาแนวโน้มของผลการตรวจจับการลอกข้อสอบจากดัชนีตรวจจับการลอกข้อสอบทั้ง 4 ดัชนี เพื่อระบุหรือบ่งชี้ว่าเกิดการลอกข้อสอบหรือไม่

3. ค่าดัชนีตรวจจับการลอกข้อสอบทั้งสี่ ซึ่งประกอบด้วย ดัชนี K_2, ดัชนี S_1, ดัชนี S_2 และดัชนี ω นั้นเพียงแต่ระบุแนวโน้มของการเกิดการลอกข้อสอบเท่านั้น แต่ไม่ได้ระบุว่าในคู่ผู้สอบนั้นผู้สอบคนใดเป็นผู้ลอก (copier) และ ผู้สอบคนใดเป็นผู้ให้ลอก (source) ถ้าหากอยากทราบว่าในคู่ ผู้สอบหนึ่งๆ ที่ถูกตรวจจับว่าน่าจะเกิดการลอกข้อสอบกัน ผู้สอบคนใดเป็นผู้ลอก และ ผู้สอบคนใดเป็นผู้ลอกอาจจะใช้วิธีการสัมภาษณ์เหตุผลในการตอบข้อคำถามต่างๆ ในแบบสอบนั้น

4. ดัชนีตรวจจับการลอกข้อสอบควรนำมาใช้หลังจากที่ผู้คุมสอบสังเกตเห็นพฤติกรรมที่ผิดปกติของคู่ผู้สอบหนึ่งๆ หรือ มีหลักฐานพยานบุคคลแล้ว เพื่อใช้ผลการวิเคราะห์ค่าดัชนีตรวจจับการลอกข้อสอบเป็นหลักฐานทางสถิติ (statistical evidence) เพื่อยืนยันการเกิดการลอกข้อสอบโดยคู่ผู้สอบที่ถูกระบุว่าลอกข้อสอบกันควรจะถูกคุมสอบอย่างเข้มงวด หรือ นั่งสอบแยกกันในระยะที่ไม่สามารถลอกข้อสอบกันได้ในการสอบครั้งต่อไป

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรเพิ่มจำนวนผู้สอบของระดับปัจจัยจำนวนผู้สอบขนาดใหญ่ให้มากกว่า 250 คนเพื่อให้การประมาณค่าพารามิเตอร์ตามโมเดล Nominal Response Model และการวิเคราะห์ค่าดัชนีตรวจจับการลอกข้อสอบมีความถูกต้องมากขึ้น

2. ควรทำการศึกษาค่าดัชนี ω ประเภทที่คำนวณโดยไม่ทราบค่าพารามิเตอร์ระดับความสามารถของผู้สอบ (θ) และ พารามิเตอร์ข้อสอบอยู่ก่อน (ω computed using unknown item and trait parameters)

3. เนื่องจากงานวิจัยชิ้นนี้ใช้ข้อมูลจริงซึ่งเป็นข้อมูลทุติยภูมิ (secondary data) มาใช้ในการจัดกระทำข้อมูลด้วยการจำลองสถานการณ์ (Simulation) ซึ่งจากการทำวิจัยพบว่าการใช้ข้อมูลจริงทำให้มีข้อจำกัดหลายประการ เช่น ข้อจำกัดด้านจำนวนผู้สอบซึ่งส่งผลกระทบต่อการประมาณค่าพารามิเตอร์ตามทฤษฎีการทดสอบแนวใหม่ เป็นต้น ดังนั้นในการวิจัยครั้งต่อไปควรใช้ข้อมูลที่ไม่ใช่ข้อมูลจริงมาจัดกระทำ โดยอาจใช้วิธีการจำลองชุดข้อมูลผลการตอบข้อสอบที่มีลักษณะตรงตามเงื่อนไขที่ศึกษาจากโปรแกรมคอมพิวเตอร์เพื่อนำชุดข้อมูลที่จำลองขึ้นไปทำการจัดกระทำข้อมูลต่อไป

4. เนื่องจากเกณฑ์ในการพิจารณาเพื่อระบุข้อสอบระดับยากที่ใช้ในงานวิจัยชิ้นนี้ยังไม่สามารถระบุข้อสอบที่อยู่ในระดับยากในความเป็นจริงได้ทั้งนี้เพราะข้อจำกัดของการใช้ข้อมูลทุติยภูมิ ดังนั้นในการวิจัยครั้งต่อไปหากมีการศึกษาปัจจัยด้านร้อยละของจำนวนข้อสอบที่ถูกลอกและวิธีการลอกโดยมีการลอกข้อสอบเฉพาะข้อยากเป็นระดับย่อยของปัจจัยวิธีการลอก ควรที่จะใช้แบบสอบที่มีจำนวนข้อสอบที่ถูกระบุว่าเป็นข้อสอบยากมากเพียงพอที่จะศึกษาในปัจจัยด้านร้อยละของจำนวนข้อสอบที่ถูกลอก โดยเกณฑ์ที่ใช้ในการระบุข้อสอบที่อยู่ในระดับยากต้องสามารถระบุข้อสอบที่เป็นข้อสอบยากในความเป็นจริงได้

เอกสารอ้างอิง

ศิริชัย กาญจนวาสี. (2550). *ทฤษฎีการทดสอบแนวใหม่.* (พิมพ์ครั้งที่ 3). กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สุชาดา บวรกิติวงศ์. (2548). *สถิติประยุกต์ทางพฤติกรรมศาสตร์.* (พิมพ์ครั้งที่ 1). กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

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

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

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

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

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