

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การทดสอบเป็นกระบวนการหนึ่งของการเรียนการสอน ซึ่งเป็นกระบวนการวัดผลสัมฤทธิ์การเรียนรู้ของผู้เรียนหลังจากเรียนจบหลักสูตร โดยทั่วไปจะใช้แบบทดสอบชนิดใช้กระดาษคำตอบ (Paper Pencil Test) เป็นเครื่องมือวัดผลสัมฤทธิ์แบบทดสอบที่นำมาใช้ในการวัดผลสัมฤทธิ์การเรียนรู้มีอยู่หลายรูปแบบ เช่น แบบทดสอบแบบหลายตัวเลือก แบบทดสอบแบบจับคู่ แบบทดสอบแบบเขียนตอบ เป็นต้น ผู้สอบจะได้ทำแบบทดสอบชุดเดียวกันและทดสอบในเวลาเดียวกัน ซึ่งเป็นแบบทดสอบที่สร้างขึ้นมาให้ครอบคลุมองค์ประกอบของคุณลักษณะหรือเนื้อหาที่ต้องการวัด เป็นแบบทดสอบที่มีค่าความยากของข้อสอบที่หลากหลาย ไม่มีการกำหนดสัดส่วนของค่าความยากที่แน่นอน นิยมใช้ข้อสอบที่มีค่าความยากของข้อสอบที่ระดับ 0.20 ถึง 0.80 เรียกอีกอย่างหนึ่งว่าการทดสอบแบบประเพณีนิยม (Classical Testing)

การที่ผู้สอบได้ทำแบบทดสอบชุดเดียวกันและทดสอบในเวลาเดียวกัน ความแม่นยำของคะแนนที่ได้จะไม่เหมือนกัน กล่าวคือ ผู้สอบบางคนอาจจะได้คะแนนที่มีความคลาดเคลื่อนน้อย แต่คะแนนของบางคนอาจจะได้คะแนนที่มีความคลาดเคลื่อนมาก ดังนั้นการจัดการทดสอบควรเป็นการจัดการทดสอบเป็นรายบุคคล โดยการคัดเลือกข้อสอบที่มีความเหมาะสมกับผู้สอบ เพื่อประมาณค่าความสามารถของผู้สอบได้อย่างแม่นยำและมีความคลาดเคลื่อนต่ำจึงยุติการทดสอบ สามารถเรียกอีกอย่างหนึ่งว่าการทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้สอบ (Adaptive Testing)

การทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้สอบ (Adaptive Testing) เป็นการทดสอบที่ใช้แบบทดสอบแตกต่างกับสำหรับผู้สอบแต่ละคน โดยมีการคัดเลือกข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบ (Weiss, 1974) ทำให้จำนวนข้อสอบที่ใช้ในการทดสอบน้อยลง ประหยัดเวลาที่ใช้ในการทดสอบ นอกจากนั้นยังช่วยผ่อนคลายความเครียดของผู้สอบได้ เช่น ผู้สอบที่มีความสามารถต่ำไม่เครียดกับการทำข้อสอบที่ยากเกินความสามารถของตนเอง ขณะที่ผู้สอบที่มีความสามารถสูงไม่จำเป็นต้องทำข้อสอบที่ง่ายเกินไป ซึ่งอาจจะส่งผลให้ผู้สอบขาดความระมัดระวังในการทำการทดสอบ เป็นต้น

การคัดเลือกข้อสอบที่จะนำมาใช้ในการทดสอบแบบปรับเหมาะนั้น เป็นขั้นตอนที่มีความสำคัญขั้นตอนหนึ่ง กล่าวคือ หากพิจารณาคัดเลือกข้อสอบที่ไม่มีความเหมาะสมกับระดับความสามารถของผู้สอบไม่จำเป็นที่จะเป็นการคัดเลือกข้อสอบข้อเริ่มต้นหรือการคัดเลือกข้อสอบข้อถัดไป จะส่งผลต่อการประมาณค่าความสามารถของผู้สอบและทำให้มีความคลาดเคลื่อนของความสามารถของผู้สอบสูงตามไปด้วย ดังนั้นการคัดเลือกข้อสอบที่จะนำมาใช้จะต้องพิจารณาจากปัจจัยต่าง ๆ ของผู้สอบร่วมด้วย เช่น คะแนนดิบหรือเกรดเฉลี่ยของผู้สอบในการคัดเลือกข้อสอบข้อเริ่มต้น หากเป็นการพิจารณาคัดเลือกข้อสอบข้อถัดไปอาจจะพิจารณาจากค่าความสามารถของผู้สอบที่ได้จากการตอบข้อสอบข้อสอบข้อที่ผ่านมา ซึ่งสามารถนำทฤษฎีการตัดสินใจมาประยุกต์ใช้ได้

การทดสอบชนิดใช้กระดาษมีจุดด้อย คือ ต้องใช้เวลาในการตรวจให้คะแนนผู้สอบมาก เกิดความผิดพลาดในการตรวจให้คะแนนผู้สอบได้ง่าย ข้อสอบสูญหายได้ง่าย ต่อมามีการนำคอมพิวเตอร์มาใช้ในการทดสอบ เพื่อแก้จุดด้อยของการทดสอบ โดยใช้กระดาษคำตอบ เรียกว่า การทดสอบด้วยคอมพิวเตอร์ (Computer Based Testing: CBT) ซึ่งมีจุดเด่น คือ สามารถตรวจให้คะแนนผู้สอบได้เร็วมากขึ้น สามารถนำข้อสอบมาจัดทำเป็นคลังข้อสอบ เพื่อนำข้อสอบไปใช้ในการทดสอบในครั้งถัดไปได้

การทดสอบด้วยคอมพิวเตอร์ (Computer Based Testing: CBT) เป็นการนำคอมพิวเตอร์มาใช้ในการบริหารการทดสอบ แบ่งออกเป็น 4 ยุค คือ 1) การทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) เป็นยุคนำคอมพิวเตอร์มาทดสอบแทนการทดสอบโดยใช้กระดาษคำตอบ โดยที่ผู้สอบทำแบบทดสอบฉบับเดียวกัน ใช้เวลาสอบเท่ากัน และสอบในสถานการณ์เดียวกัน 2) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นยุคนำคอมพิวเตอร์มาทดสอบ โดยให้ผู้สอบทำแบบทดสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ ซึ่งผู้สอบแต่ละคนอาจจะทำข้อสอบที่ไม่เหมือนกันและจำนวนข้อที่ไม่เท่ากัน 3) การวัดอย่างต่อเนื่อง (Continuous Measurement: CM) เป็นยุคนำการสอบและการทดสอบมาใช้ในระบบคอมพิวเตอร์เดียวกัน ซึ่งจะมีการเรียนกับการทดสอบสลับกันไปมาตลอดบทเรียน กล่าวคือ เมื่อเรียนเรื่องใดจบ จะมีการทดสอบในเรื่องนั้น 4) การวัดอย่างชาญฉลาด (Intelligent Measurement: IM) เป็นยุคนำคอมพิวเตอร์มาใช้ในการบริหารการทดสอบครบทุกขั้นตอน กล่าวคือ เป็นการนำคอมพิวเตอร์มาตรวจให้คะแนนแบบทดสอบประเภทเขียนตอบที่มีคำตอบที่แน่นอนตายตัวและสามารถวิเคราะห์ผลงาน จุดแข็งและจุดอ่อน รวมทั้งให้ข้อเสนอแนะกับผู้เรียนหรือผู้สอบได้ (Bunderson, & Others, 1989, quoted in Linn, 1990)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นการทดสอบที่จัดข้อสอบให้เหมาะสมกับความสามารถของผู้สอบ ซึ่งอยู่บนพื้นฐานของการตอบข้อสอบข้อที่ผ่านมาของผู้สอบ กล่าวคือ เมื่อผู้สอบทำข้อสอบข้อเริ่มต้นหรือชุดเริ่มต้น (ขึ้นอยู่กับ การออกแบบ) แล้วนำผลการตอบข้อสอบมาวิเคราะห์หรือประเมินระดับความสามารถของผู้สอบ เพื่อที่จะคัดเลือกข้อสอบข้อถัดไปที่เหมาะสมโดยอาศัยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เป็นพื้นฐานแบ่งออกเป็น 2 วิธี คือ การทดสอบแบบสองขั้นตอน (Two-Stage Strategies) และการทดสอบแบบหลายขั้นตอน (Multi-Stage Strategies) โดยการทดสอบแบบหลายขั้นตอน ยังแบ่งออกเป็น 2 รูปแบบ คือ รูปแบบการแยกทางแบบคงที่ (Fixed Branching Model) และรูปแบบการแยกทางแบบแปรผัน (Variable Branching Model) สำหรับการเลือกรูปแบบการทดสอบนั้น ผู้ใช้ควรเลือกรูปแบบการทดสอบให้เหมาะสมกับวัตถุประสงค์ของการทดสอบ (Weiss, 1974)

สำหรับการดำเนินการในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มี 5 ขั้นตอน คือ ขั้นตอนที่ 1 การสร้างคลังข้อสอบ (Create Item Bank) เป็นขั้นตอนที่ต้องพิจารณาขนาดของคลังข้อสอบและเงื่อนไขต่าง ๆ ที่สอดคล้องกับหลักการของทฤษฎีการตอบสนองข้อสอบ เพื่อคัดเลือกข้อสอบที่ตรงตามเงื่อนไขที่กำหนดมาจัดเก็บในคลังข้อสอบ ขั้นตอนที่ 2 การคัดเลือกข้อสอบข้อเริ่มต้น (First Item Selection) เป็นขั้นตอนที่ต้องพิจารณาคัดเลือกข้อสอบข้อเริ่มต้นที่มีความเหมาะสมกับ

ผู้สอบ ซึ่งควรเป็นข้อสอบที่มีค่าความยากของข้อสอบระดับปานกลาง กล่าวคือ การเลือกข้อสอบที่มีค่าความยากของข้อสอบใกล้เคียงกับค่าเฉลี่ยของค่าความยากของข้อสอบทั้งหมด การคัดเลือกข้อสอบข้อเริ่มต้นจากการสุ่มในช่วงระดับความยากของข้อสอบที่กำหนดหรือการคัดเลือกข้อสอบข้อเริ่มต้นตามเกณฑ์ที่กำหนด ขั้นตอนที่ 3 การคัดเลือกข้อสอบข้อถัดไป (Next Item Selection) เป็นขั้นตอนที่ต้องพิจารณาคัดเลือกข้อสอบตามเงื่อนไขที่กำหนด โดยอาศัยผลการตอบข้อสอบข้อก่อนหน้ามาร่วมพิจารณาด้วย ขั้นตอนที่ 4 การประมาณค่าความสามารถของผู้สอบ (Calculate Possible Ability Levels) เป็นการประมาณค่าความสามารถของผู้สอบหลังจากตอบข้อสอบแล้ว ซึ่งวิธีการประมาณค่าความสามารถของผู้สอบที่นิยมใช้กัน คือ 1) การประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ (Bayesian) 2) การประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ปรับใหม่ (Bayesian Updating) 3) การประมาณค่าความสามารถของผู้สอบด้วยวิธีความน่าจะเป็นสูงสุด (Maximum Likelihood) ซึ่งการประมาณค่าความสามารถของผู้สอบด้วยวิธีความน่าจะเป็นสูงสุดยังมีข้อจำกัดอยู่ คือ ถ้าผู้สอบตอบข้อสอบถูกทุกข้อหรือผิดทุกข้อจะไม่สามารถประมาณค่าความสามารถของผู้สอบได้ ขั้นตอนที่ 5 เกณฑ์การยุติการทดสอบ (Termination Criterion) คือ การสิ้นสุดการทดสอบ เมื่อผู้สอบทดสอบครบตามเงื่อนไขหรือเกณฑ์ที่กำหนดการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะขาดขั้นตอนใดขั้นตอนหนึ่งไม่ได้ เนื่องจากแต่ละขั้นตอนมีการดำเนินการที่มีความสัมพันธ์กัน (Thompson, & Weiss, 2011)

ในการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะคัดเลือกข้อสอบที่สอดคล้องกับวิธีการประมาณค่าความสามารถของผู้สอบ (Hulin, Drasgow, & Parson, 1983) หากคัดเลือกข้อสอบที่ไม่มีความเหมาะสมกับระดับความสามารถของผู้สอบมาใช้ทำให้ใช้เวลาในการทดสอบและจำนวนข้อสอบที่มากขึ้น ซึ่งวิธีการคัดเลือกข้อสอบข้อถัดไป มีนักวิชาการนำเสนอไว้หลายวิธี เช่น การคัดเลือกข้อสอบข้อถัดไป โดยใช้สารสนเทศคุลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (Kullback-Leibler Information with a Posterior Distribution: KLP) เสนอโดย Veldkamp, and Van der Linden (2002) เป็นฟังก์ชันที่แก้ไขฟังก์ชันสารสนเทศฟิชเชอร์ในค่าสารสนเทศคุลเบค-ไลเบลอร์ $KL_j(\theta|\hat{\theta}_n)$ โดยคัดเลือกข้อสอบที่มีค่า KLP สูงสุดเป็นข้อสอบข้อถัดไป การคัดเลือกข้อสอบข้อถัดไป โดยใช้ประสิทธิภาพของสารสนเทศ (Efficiency Balanced Information: EBI) นำเสนอโดย Han (2010) เป็นวิธีเลือกข้อสอบข้อถัดไป โดยพิจารณาจากการประเมินช่วงของค่าความสามารถของผู้สอบ (θ) ซึ่งประเมินจากประสิทธิภาพของค่าสารสนเทศ ซึ่งหาได้จากค่าความคลาดเคลื่อนมาตรฐานและการกำหนดค่าความคลาดเคลื่อนมาตรฐานที่คาดว่าจะเกิดขึ้นจากค่าความสามารถของผู้สอบหลังจากผู้สอบตอบข้อสอบข้อที่ j แล้ว โดยเลือกข้อสอบที่มีค่า EBI สูงสุดเป็นข้อสอบข้อถัดไป

วิธีการคัดเลือกข้อสอบข้อถัดไปที่นิยมใช้กันมากคือ 1) การคัดเลือกข้อสอบข้อถัดไปที่ค่าความยากของข้อสอบ (b-Parameter) ใกล้เคียงกับความสามารถของผู้สอบ 2) การคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ (Maximum Information Criterion: MI)

วิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ (MI) เป็นการคัดเลือกข้อสอบข้อถัดไปที่นำปัจจัยของการเดาข้อสอบมาพิจารณาร่วมกับค่าความสามารถของผู้สอบ (θ) ขณะนั้น ซึ่งจะคัดเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบมาเป็นข้อสอบข้อถัดไป (Birnbbaum, 1968)

ในการคัดเลือกข้อสอบข้อถัดไป หากคัดเลือกข้อสอบข้อใดข้อหนึ่งซ้ำกันหลายครั้ง อาจจะทำให้ผลการประมาณค่าความสามารถของผู้สอบไม่ตรงกับความเป็นจริง Cheng, Ankenmann, and Spray (2003) ได้เสนอวิธีการคำนวณดัชนีการคาบเกี่ยวกันโดยเฉลี่ย ซึ่งเป็นฟังก์ชันของดัชนีการใช้ข้อสอบซ้ำในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีจำนวนข้อสอบที่แน่นอน ชี้ให้เห็นว่าการควบคุมการใช้ข้อสอบซ้ำ โดยระบุค่าควบคุมอัตราการใช้ข้อสอบซ้ำ (Maximum Item Exposure Rate: r_{max}) ให้ได้รับการควบคุมโดยตรงมากที่สุดทั้งระดับข้อสอบและระดับแบบทดสอบ ต่อมา Chen and Lei (2005) ได้เสนอให้นำวิธีการของ Sympon and Hetter (1985) ซึ่งสามารถควบคุมการเลือกข้อสอบซ้ำในระดับของข้อสอบแต่ละข้อและระดับของแบบทดสอบได้ในเวลาเดียวกัน มาใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผลการวิจัยแสดงว่า การควบคุมอัตราการใช้ข้อสอบซ้ำ (Item Exposure Rate) สามารถทำให้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีประสิทธิภาพและมีความแม่นยำในการประมาณค่าความสามารถของผู้สอบมากยิ่งขึ้น เมื่อดัชนีการคาบเกี่ยวกันระหว่างข้อสอบสูงสุดที่กำหนดไว้ล่วงหน้านั้น สามารถคัดเลือกข้อสอบมาใช้ซ้ำกันในอัตราส่วนที่กำหนด

ทฤษฎีการตัดสินใจ (Decision Theory) เป็นทฤษฎีที่อธิบายถึงการนำหลักเกณฑ์หรือเครื่องมือต่าง ๆ เข้ามาช่วยในการตัดสินใจ เพื่อให้ผู้ตัดสินใจมีโอกาสผิดพลาดน้อยลงภายใต้สถานการณ์ต่าง ๆ ที่สามารถประเมินได้ โดยใช้กฎเกณฑ์หรือเครื่องมือดังกล่าวมาพิจารณาทางเลือกที่เหมาะสม ถ้ามีทางเลือกเพียงทางเลือกเดียวการตัดสินใจก็ไม่เกิดขึ้น เนื่องจากต้องเลือกตามวิธีทางเลือกที่มีอยู่ กล่าวคือ จะไม่มีการเปรียบเทียบว่าผลลัพธ์หรือความเหมาะสมมากที่สุดหรือไม่ ถ้ามีวิธีที่มีความเหมาะสมมากกว่าหนึ่งทางเลือกแล้ว จะต้องมีการตัดสินใจเลือกทางหรือวิธีที่มีความเหมาะสมมากที่สุด (Hansson, 2005)

หลักการของทฤษฎีการตัดสินใจ (Decision Theory) เป็นการนำเงื่อนไข หลักเกณฑ์ของแต่ละกรณีมาใช้ในการพิจารณา เพื่อหาทางเลือกที่มีความเหมาะสมมากที่สุด ซึ่งสามารถแบ่งวิธีการตัดสินใจออกเป็น 3 ประเภท คือ 1) การตัดสินใจในสภาพแวดล้อมภายใต้ความแน่นอน เป็นวิธีที่ผู้ตัดสินใจทราบถึงความแน่นอนของสิ่งที่จะตามมาของทุกทางเลือก โดยผู้ตัดสินใจจะเลือกทางเลือกที่มีความเหมาะสมมากที่สุด 2) การตัดสินใจในสภาพแวดล้อมภายใต้ความไม่แน่นอน เป็นวิธีที่ผู้ตัดสินใจไม่ทราบถึงความน่าจะเป็นของการเกิดเหตุการณ์ต่าง ๆ ดังนั้น การตัดสินใจจึงอาศัยสารสนเทศที่มีอยู่ เจตคติ (Attitude) ส่วนตัวของผู้ตัดสินใจและความเหมาะสมตามเงื่อนไขที่มีอยู่ 3) การตัดสินใจภายใต้สภาพแวดล้อมที่มีความเสี่ยง เป็นวิธีที่ผู้ตัดสินใจทราบถึงความน่าจะเป็นของการเกิดเหตุการณ์ต่าง ๆ โดยพิจารณาจากค่าคาดหวังของผลที่ได้จากผลคูณของค่าความน่าจะเป็นของเหตุการณ์กับผลตอบแทนที่ได้รับของเหตุการณ์นั้น ซึ่งมีเกณฑ์การคัดเลือกค่าคาดหวัง คือ เกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) และเกณฑ์ทางเลือกที่มีค่าความคาดหวังของค่าเสียโอกาสน้อยที่สุด (Minimum Expected Value of Opportunity Loss: EOL)

ในกรณีที่ใช้เกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) ต้องพิจารณาภายใต้เงื่อนไข 3 กรณี ดังนี้ 1) กรณีที่ค่าความคาดหวัง (EMV) เท่ากัน จะพิจารณาเลือกทางเลือกที่ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD) ของข้อมูลที่เหมาะสม (ขึ้นอยู่กับ การนำไปใช้ในด้านใด) 2) กรณีที่ค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูลเท่ากัน จะเลือกทางเลือกที่มีค่าความคาดหวังสูงสุด 3) กรณีที่ค่าความคาดหวังในแต่ละทางเลือกไม่เท่ากันและมีค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูลแต่ละทางเลือกไม่เท่ากันจะต้องพิจารณาทางเลือกใหม่ โดยใช้ค่าสัมประสิทธิ์การกระจาย (Coefficient of Variation: CV) ซึ่งจะพิจารณาเลือกทางเลือกที่ค่าสัมประสิทธิ์การกระจายที่เหมาะสม (ขึ้นอยู่กับ การนำไปใช้ในด้านใด)

จากหลักการของทฤษฎีการตัดสินใจ (Decision Theory) ข้างต้น สามารถประยุกต์กับการจัดการทดสอบได้ โดยในการเลือกแบบทดสอบ จะพิจารณาทั้งกลุ่มที่ได้คะแนนสูงและกลุ่มที่ได้คะแนนต่ำ หากเป็นการเลือกข้อสอบ จะพิจารณาทั้งกรณีที่ตอบข้อสอบถูกและกรณีที่ตอบข้อสอบผิด ซึ่งตามหลักการคัดเลือกแบบทดสอบหรือข้อสอบที่จะนำมาใช้ในการทดสอบนั้น จะต้องมีการกระจายของข้อมูลสูง กล่าวคือ แบบทดสอบหรือข้อสอบมีค่าการกระจายของข้อมูลสูง แสดงว่าแบบทดสอบหรือข้อสอบนั้น มีความเหมาะสมกับผู้สอบในหลายระดับความสามารถ นอกจากนั้นยังสามารถใช้ค่าสัมประสิทธิ์การกระจายเป็นค่าในการตัดสินใจเลือกแบบทดสอบหรือข้อสอบได้ โดยจะเลือกแบบทดสอบหรือข้อสอบที่มีค่าสัมประสิทธิ์การกระจายสูง ซึ่งเป็นการแสดงถึงประสิทธิภาพในการจำแนกความสามารถของผู้สอบได้ดี

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เป็นการทดสอบที่มีการคัดเลือกข้อสอบข้อถัดไปให้มีความเหมาะสมกับระดับความสามารถของผู้สอบ เป็นขั้นตอนที่ 3 จาก 5 ขั้นตอนของการดำเนินการในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งเป็นขั้นตอนที่ต้องพิจารณาคัดเลือกข้อสอบข้อถัดไปตามเงื่อนไขที่กำหนด โดยอาศัยผลการตอบข้อสอบข้อก่อนหน้านี้นำมาพิจารณาด้วยสำหรับการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจต้องพิจารณาเงื่อนไขหรือหลักเกณฑ์ต่าง ๆ ให้สอดคล้องกับหลักการของทฤษฎีการตัดสินใจที่กล่าวถึงการพิจารณาเงื่อนไขที่มีอยู่ในแต่ละทางเลือก โดยพิจารณาผลตอบแทนที่ได้ตามเงื่อนไขในแต่ละทางเลือก ซึ่งในที่นี้คือ ค่าประมาณความสามารถของผู้สอบก่อนที่จะตอบข้อสอบที่มีอยู่ในคลังข้อสอบ โดยไม่นำข้อสอบที่ใช้แล้วมาพิจารณา กรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นถูกและกรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นผิด กล่าวคือ ค่าความคาดหวังสูงสุดได้มาจากการคำนวณผลรวมของผลคูณค่าความสามารถของผู้สอบกับความน่าจะเป็นในกรณีที่ผู้สอบตอบข้อสอบที่มีอยู่ในคลังข้อสอบได้ถูก และกรณีที่ผู้สอบตอบข้อสอบที่มีอยู่ในคลังข้อสอบผิด โดยไม่นำข้อสอบที่ใช้แล้วมาพิจารณา ซึ่งจะพิจารณาคัดเลือกข้อสอบที่มีค่าสัมประสิทธิ์การกระจายสูงสุดเป็นข้อสอบข้อถัดไป

Han (2009) ได้ศึกษาวิธีอัตราส่วนของสารสนเทศสูงสุด (Maximum Information Ratio: MIR) สำหรับการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เพื่อศึกษาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้วยการเปรียบเทียบวิธีการคัดเลือกข้อสอบข้อถัดไปด้วยการสุ่มกับวิธีการใช้สารสนเทศสูงสุดของฟิชเชอร์ (Maximum Fisher Information: MFI) โดยศึกษากับคลังข้อสอบขนาด 500 ข้อ ผลการศึกษาปรากฏว่าวิธีอัตราส่วนของสารสนเทศสูงสุดสำหรับการคัดเลือกข้อสอบข้อถัดไป มีประสิทธิภาพในการบริหาร

จัดการคลังข้อสอบดีกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปด้วยการสุ่ม และวิธีการใช้สารสนเทศสูงสุดของฟิชเชอร์ (MFI) ซึ่งสามารถสร้างความสมดุลของคลังข้อสอบ รวมทั้งช่วยลดจำนวนข้อสอบและความคลาดเคลื่อนในการยุติการทดสอบได้

การตัดสินใจเลือกทางเลือกที่มีความเหมาะสมที่จะนำไปใช้ในการแก้ไขปัญหาที่เกิดขึ้น หากตัดสินใจเลือกทางเลือกที่ไม่มีความเหมาะสมกับปัญหาจะทำให้ผลที่ของการแก้ไขปัญหาไม่ได้ผลดีเท่าที่ควรหรืออาจจะทำให้เกิดปัญหามากยิ่งขึ้นได้ ซึ่งสามารถนำมาประยุกต์กับขั้นตอนการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้ เนื่องจากการคัดเลือกข้อถัดไปมีการตัดสินใจเลือกข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบ หากตัดสินใจเลือกข้อสอบที่ไม่มีความเหมาะสมกับความสามารถของผู้สอบทำให้ประมาณค่าความสามารถของผู้สอบคลาดเคลื่อนได้

จากเหตุผลข้างต้นประกอบกับการศึกษาเอกสารและงานวิจัยทั้งในต่างประเทศและต่างประเทศ เกี่ยวกับวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า ยังไม่มีการนำค่าประมาณความสามารถของผู้สอบก่อนที่จะตอบข้อสอบที่ยังไม่ได้ใช้ในคลังข้อสอบ กรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นถูกและกรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นผิดมาพิจารณาคัดเลือกข้อสอบข้อถัดไปร่วมกัน ซึ่งการนำผลที่ได้จากการคาดการณ์ทั้งทางบวกและทางลบ ย่อมมีผลต่อการตัดสินใจเลือกข้อสอบข้อถัดไปที่เหมาะสมมากยิ่งขึ้นผู้วิจัยจึงสนใจที่จะพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจและพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เพื่อนำเสนอวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีประสิทธิภาพมากยิ่งขึ้น ด้านจำนวนข้อสอบที่ใช้ในการทดสอบและเวลาที่ใช้ในการทดสอบและโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่สามารถนำไปใช้งานได้จริง รวมทั้งเป็นแนวทางการศึกษาวิธีการพัฒนาเงื่อนไขต่าง ๆ ที่นำมาใช้กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้อีกด้วย

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจ (Decision Theory)
2. เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
3. เพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไประหว่างวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ (Maximum Information Criterion) กับวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจ ด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบ

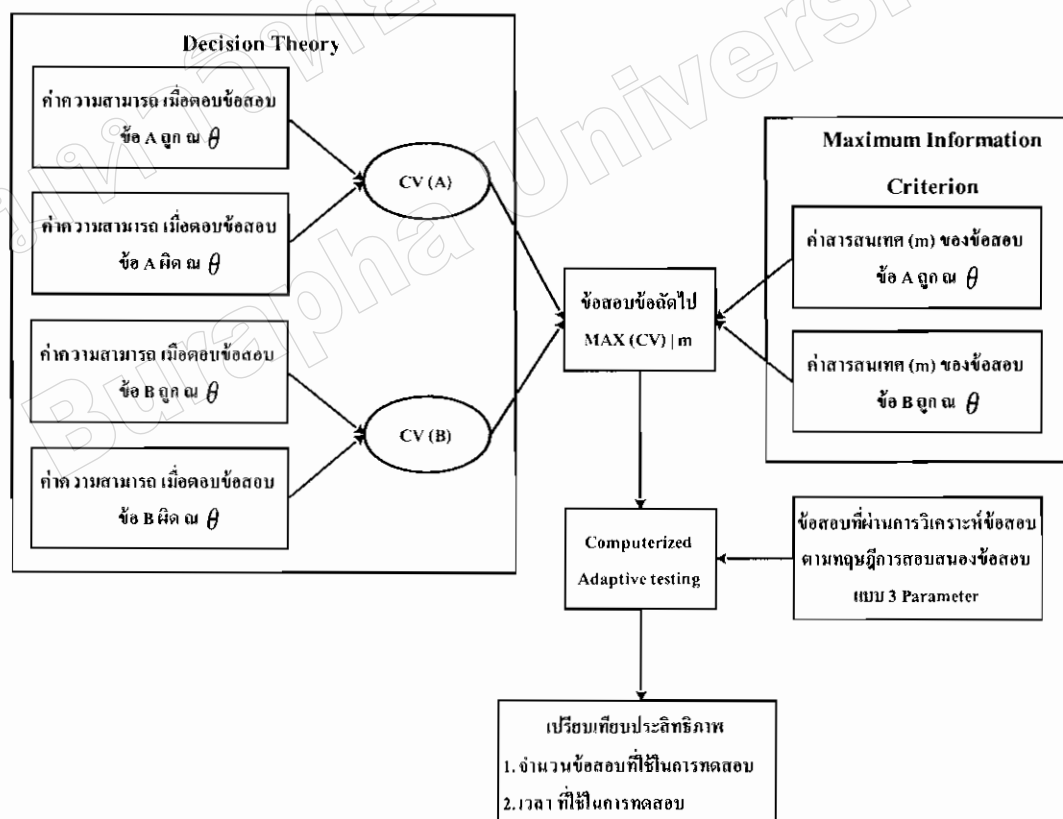
กรอบแนวคิดของการวิจัย

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจครั้งนี้ เป็นการศึกษาวิธีการตัดสินใจภายใต้ความเสี่ยง (Decision Making Under Risk) ตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) โดยใช้ค่าสัมประสิทธิ์การกระจายมาประยุกต์ในการคัดเลือกข้อสอบข้อถัดไปใน

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยการนำค่าประมาณความสามารถของผู้สอบก่อนที่จะตอบข้อสอบที่มีอยู่ในคลังข้อสอบ กรณีที่คาดว่าผู้สอบจะตอบข้อสอบข้อนั้นถูกและกรณีที่คาดว่าจะตอบข้อสอบข้อนั้นผิด มาพิจารณาคัดเลือกข้อสอบข้อถัดไปร่วมกัน โดยไม่นำข้อสอบที่ใช้ไปแล้วมาพิจารณา โดยเปรียบเทียบประสิทธิภาพกับวิธีการคัดเลือกข้อสอบข้อถัดไป ด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบระหว่างวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ

ข้อสอบที่นำมาใช้ในการวิจัย คือข้อสอบวิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้ผ่านการวิเคราะห์หาค่าความยาก (b-Parameter) ค่าอำนาจจำแนก (a-Parameter) และค่าการเดาของข้อสอบ (c-Parameter) ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) แบบ 3 พารามิเตอร์ โดยคัดเลือกข้อสอบที่มีค่าความยากของข้อสอบ (b-Parameter) ในช่วง -2.50 ถึง 2.50 ค่าอำนาจจำแนกของข้อสอบ (a-Parameter) ในช่วง 0.50 ถึง 2.50 และค่าการเดาของข้อสอบ (c-Parameter) น้อยกว่า 0.30 มาเก็บไว้ในคลังข้อสอบของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้ข้อสอบ

สามารถเขียนเป็นกรอบแนวคิดของการวิจัยได้ตามภาพที่ 1



ภาพที่ 1 กรอบแนวคิดการวิจัยของการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจ

จากภาพที่ 1 แสดงกรอบแนวคิดการวิจัยของการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจ สัญลักษณ์ประกอบในภาพที่ 1 มีความหมาย ดังนี้

1. θ คือ ความสามารถปัจจุบันของผู้สอบ
2. $MAX(CV)$ คือ ค่าสัมประสิทธิ์การกระจายสูงสุด ซึ่งเป็นค่าที่ใช้ในการเลือกข้อสอบข้อถัดไป
3. m คือ ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ

สมมติฐานการวิจัย

จากวัตถุประสงค์ของการวิจัย เพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไประหว่างวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ (Maximum Information Criterion) กับวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจ ด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบ ผู้วิจัยจึงตั้งสมมติฐานการวิจัย ดังนี้

1. จำนวนข้อสอบที่ใช้ในการทดสอบของวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ
2. เวลาที่ใช้ในการทดสอบของวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ

ประโยชน์ที่คาดว่าจะได้รับการวิจัย

1. ได้วิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ทำให้สามารถลดจำนวนข้อสอบที่ใช้ทดสอบและลดเวลาที่ใช้ในการทดสอบ ส่งผลให้เกิดประสิทธิภาพสูงสุด ประหยัดเวลาและงบประมาณตลอดจนแรงงานคนที่ใช้ในการจัดการทดสอบ
2. ได้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ที่มีวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจมาใช้งานจริงกับการทดสอบวัดความรู้ทางการอ่านภาษาอังกฤษของหน่วยงานอื่น ส่งผลให้ผู้สอบทราบระดับความรู้ความสามารถทางการอ่านภาษาอังกฤษ
3. เป็นแนวทางศึกษาวิธีการพัฒนาเงื่อนไขต่าง ๆ ที่นำมาใช้กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ต่อไป เช่น การสร้างคลังข้อสอบ การคัดเลือกข้อสอบข้อเริ่มต้น การประมาณค่าความสามารถของผู้สอบ และเกณฑ์การยุติการทดสอบ เป็นต้น

ขอบเขตของการวิจัย

การวิจัยครั้งนี้มุ่งเน้นที่จะพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้ทฤษฎีการตัดสินใจ (Decision Theory) พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือก

ข้อสอบข้อถัดไป ด้านจำนวนข้อสอบที่ใช้ในการทดสอบและด้านเวลาที่ใช้ในการทดสอบระหว่างวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ (Maximum Information Criterion) กับวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจ กับซึ่งมีขอบเขตของการวิจัย ดังนี้

1. ประชากรที่ใช้ในการวิจัยเป็นนักศึกษาระดับปริญญาตรี ที่ลงทะเบียนเรียนวิชาภาษาอังกฤษเพื่อทักษะการเรียนในภาคเรียนที่ 1 ปีการศึกษา 2555 มหาวิทยาลัยราชภัฏสวนดุสิต กรุงเทพมหานคร จำนวน 1,200 คน

2. การทดสอบวัดความรู้ในทักษะการอ่านภาษาอังกฤษในการวิจัยครั้งนี้ เป็นความรู้ในทักษะการอ่านภาษาอังกฤษของนักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏสวนดุสิต กรุงเทพมหานคร

3. ตัวแปรที่ใช้ในการวิจัย มีดังนี้

3.1 ตัวแปรต้น คือ วิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 2 วิธี ดังนี้

3.1.1 วิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ (Maximum Information Criterion)

3.1.2 วิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจ (Decision Theory)

3.2 ตัวแปรตาม คือ ประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป พิจารณาจาก

3.2.1 ด้านจำนวนข้อสอบที่ใช้ในการทดสอบ

3.2.2 ด้านเวลาที่ใช้ในการทดสอบ

นิยามศัพท์เฉพาะ

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing Programs) หมายถึง โปรแกรมที่ผู้วิจัยพัฒนาขึ้นตามหลักการของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) และหลักการของทฤษฎีการตัดสินใจ (Decision Theory) แบบวิธีการตัดสินใจภายใต้ความเสี่ยง (Decision Making Under Risk)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing) หมายถึง การทดสอบด้วยคอมพิวเตอร์ที่มีการคัดเลือกข้อสอบให้มีความเหมาะสมกับระดับความสามารถของผู้สอบ

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) หมายถึง ความสัมพันธ์ระหว่างคุณลักษณะของข้อสอบกับความสามารถของผู้สอบมีลักษณะเป็นโค้งแบบ Logistic

วิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจ (Next Item Selection Procedure Using Decision Theory) หมายถึง การคัดเลือกข้อสอบข้อถัดไป โดยใช้วิธีการตัดสินใจภายใต้ความเสี่ยง (Decision Making Under Risk) ตามเกณฑ์ค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) โดยใช้ค่าสัมประสิทธิ์การกระจาย (Coefficient of Variation: CV)

วิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้สอบ (Maximum Information Criterion) หมายถึง การคัดเลือกข้อสอบข้อถัดไปที่ค่าสารสนเทศสูงสุดใกล้เคียงกับค่าความสามารถของผู้สอบ

ค่าอำนาจจำแนกของข้อสอบ (a-Parameter) หมายถึง ค่าความชันโค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve: ICC) ที่ตำแหน่งค่าความยากของข้อสอบ (b-Parameter) ซึ่งมีค่าตั้งแต่ -2.50 ถึง +2.50 วิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้

ค่าความยากของข้อสอบ (b-Parameter) หมายถึง ตำแหน่งโค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve: ICC) บนระดับความสามารถที่ทำให้มีโอกาสตอบข้อสอบได้ถูกต้อง ซึ่งมีค่าตั้งแต่ 0.00 ถึง +2.50 วิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้

ค่าการเดาของข้อสอบ (c-Parameter) หมายถึง ค่าความน่าจะเป็นที่ผู้สอบจะเดาข้อสอบข้อนั้นได้ถูกต้อง ซึ่งเป็นค่ากำกับต่ำสุดของโค้งคุณลักษณะข้อสอบ (Item Characteristic Curve: ICC) ซึ่งมีค่าไม่เกิน 0.30 วิชาภาษาอังกฤษเพื่อทักษะการเรียนรู้

ทฤษฎีการตัดสินใจ (Decision Theory) หมายถึง ขั้นตอน หลักการ หรือวิธีการที่นำมาช่วยในการตัดสินใจ เพื่อทำให้ตัดสินใจผิดพลาดน้อยลงภายใต้สถานการณ์ต่าง ๆ ที่สามารถประเมินได้

ค่าความสามารถของผู้สอบ (Ability: θ) หมายถึง ระดับความสามารถของผู้สอบที่ประเมินได้จากการตอบข้อสอบที่ผ่านมาในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ค่าความน่าจะเป็นที่จะตอบข้อสอบถูก (Probability of a Correct Response on Item) หมายถึง ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบที่มีอยู่ในคลังข้อสอบถูกต้องที่ระดับความสามารถปัจจุบัน โดยไม่นำข้อสอบที่ใช้ไปแล้วมาพิจารณา

ค่าความน่าจะเป็นที่จะตอบข้อสอบผิด (Probability of an Incorrect Response on Item) หมายถึง ความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบที่มีอยู่ในคลังข้อสอบผิดที่ระดับความสามารถปัจจุบัน โดยไม่นำข้อสอบที่ใช้ไปแล้วมาพิจารณา

ค่าความสามารถของผู้สอบที่คาดว่าจะตอบข้อสอบถูก (Ability of a Correct Response on Item) หมายถึง ค่าประมาณความสามารถของผู้สอบก่อนที่ผู้สอบจะตอบข้อสอบที่มีอยู่ในคลังข้อสอบ โดยไม่นำข้อสอบที่ใช้ไปแล้วมาพิจารณา กรณีที่คาดว่าจะตอบข้อสอบข้อนั้นถูก

ค่าความสามารถของผู้สอบที่คาดว่าจะตอบข้อสอบผิด (Ability of an Incorrect Response on Item) หมายถึง ค่าประมาณความสามารถของผู้สอบก่อนที่ผู้สอบจะตอบข้อสอบที่มีอยู่ในคลังข้อสอบ โดยไม่นำข้อสอบที่ใช้ไปแล้วมาพิจารณา กรณีที่คาดว่าจะตอบข้อสอบข้อนั้นผิด

ค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) หมายถึง อัตราส่วนของผลตอบแทนที่คาดว่าจะได้จากเหตุการณ์ต่าง ๆ หาได้จากผลรวมของผลคูณระหว่างความน่าจะเป็นกับผลตอบแทนทุกเหตุการณ์ที่พิจารณา ซึ่งมี 2 เหตุการณ์ คือ กรณีที่คาดว่าจะตอบข้อสอบถูก และกรณีที่คาดว่าจะตอบข้อสอบผิด ผลตอบแทนของแต่ละเหตุการณ์ คือ ค่าประมาณความสามารถของผู้สอบ กรณีที่คาดว่าจะตอบข้อสอบถูก และค่าประมาณความสามารถของผู้สอบ กรณีที่คาดว่าจะตอบข้อสอบผิด ความน่าจะเป็นของเหตุการณ์ คือ ความน่าจะเป็นกรณีที่คาดว่าจะตอบข้อสอบถูก และความน่าจะเป็นกรณีที่คาดว่าจะตอบข้อสอบผิด

ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD) หมายถึง การวัดการกระจายของข้อมูลที่เบี่ยงเบนออกจากค่าเฉลี่ยของข้อมูลชุดนั้น

ค่าสัมประสิทธิ์การกระจาย (Coefficient of Variation: CV) หมายถึง ค่าอัตราส่วนระหว่างส่วนเบี่ยงเบนมาตรฐานกับค่าเฉลี่ยของข้อมูลชุดเดียวกัน

มหาวิทยาลัยบูรพา
Burapha University