การประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และ
โมเดลโลจีสติก 3 พารามิเตอร์

ปริญญานิพนธ์
ของ
กิตติศักดิ์ นิวรัตน์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาดุษฎีบัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา
พฤษภาคม 2554

การประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และ
โมเดลโลจีสติก 3 พารามิเตอร์

ปริญญานิพนธ์
ของ
กิตติศักดิ์ นิวรัตน์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาดุษฎีบัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา
พฤษภาคม 2554

งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัย

จาก

มหาวิทยาลัยศรีนครินทรวิโรฒ

ประเภททุนงบประมาณแผ่นดิน ประจำปีงบประมาณ 2551

การประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และ
โมเดลโลจีสติก 3 พารามิเตอร์

บทคัดย่อ
ของ
กิตติศักดิ์ นิวรัตน์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาดุษฎีบัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา
พฤษภาคม 2554

กิตติศักดิ์ นิวรัตน์. (2554). *การประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และ โมเดลโลจีสติก 3 พารามิเตอร์*. ปริญญานิพนธ์ กศ.ด.(การทดสอบและวัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม:รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ , รองศาสตราจารย์ ดร.องอาจ นัยพัฒน์ , ผู้ช่วยศาสตราจารย์ ดร.สุรชัย มีชาญ.

ในการวิจัยครั้งนี้ มีความมุ่งหมายหลักเพื่อศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ ซึ่งประกอบด้วยสัมประสิทธิ์ r_B กับสัมประสิทธิ์โครงสร้าง และโมเดลโลจีสติก 3 พารามิเตอร์ ได้แก่ สัมประสิทธิ์ ρ_{IRT} และใช้สัมประสิทธิ์ KR-20 เป็นเกณฑ์เทียบ

ประชากรเทียมที่ใช้เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย จำนวน 4 เขตพื้นที่ ปีการศึกษา 2550 โดยใช้วิธีการสุ่มแบบแบ่งชั้น(Stratified Random Sampling) จำนวน 6,000 คน และใช้วิธีการสุ่มแบบใส่คืนจากประชากรเทียม แบ่งเป็นกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ 300 500 1,000 และ 1,500 คน กลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ 30 60 120 และ 240 คน แต่ละขนาดจะมีจำนวน 30 กลุ่ม ใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 ที่ผู้วิจัยสร้างขึ้น วิเคราะห์ข้อมูลโดยหาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน การประมาณค่าความเชื่อมั่น ความคลาดเคลื่อนมาตรฐาน ความลำเอียง และการวิเคราะห์เส้นภาพ ผลการวิจัยสรุปได้ดังนี้

1. การประมาณค่าความเชื่อมั่นของแบบทดสอบ จากประชากรเทียมและจากกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ ขนาด 300 500 1,000 และ 1,500 คน พบว่า ค่าความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ ρ_{IRT} ได้ค่าความเชื่อมั่นสูงสุด รองลงมาคือ สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิโครงสร้าง ตามลำดับ

2.การประมาณค่าความเชื่อมั่นของแบบทดสอบ จากกลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ ขนาด 30 60 120 และ 240 คน พบว่า สัมประสิทธิ์ r_B ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุด รองลงมาคือ สัมประสิทธิ์ KR-20 และสัมประสิทธิโครงสร้าง ตามลำดับ

3. ที่กลุ่มตัวอย่างขนาด 300 500 1,000 1,500 คน ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B และสัมประสิทธิโครงสร้าง มีค่าอยู่ในช่วงความเชื่อมั่น 95%

4.ความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสูตรประมาณค่าความเชื่อมั่น มีความเหมาะสมและมีคุณภาพในการประมาณค่า

5.การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง มีเส้นภาพความเชื่อมั่นที่ไม่ขนานกัน

THE ESTIMATION OF THE RELIABILITY OF A TEST UNDER THE CONGENERIC MODEL AND THREE - PARAMETER LOGISTIC MODEL

AN ABSTRACT

BY

KITTISAK NEWRAT

Presented in Partial Fulfillment of the Requirements for the

Doctor of Education degree in Testing and Measurement

Srinakarinwirot University

May 2011

Kittisak Newrat. (2011). *The Estimation of The Reliability of A Test Under The Congeneric Model and Three - Parameter Logistic Model*. Dissertation, Ed.D. (Testing and Measurement). Bangkok: Graduate School, Srinakarinwirot University. Advisor Committee: Assoc. Prof. Dr.Boonchird Pinyoanuntapong, Assoc. Prof. Dr.Ong-art Naiyapatana, Asst. Prof. Dr. Surachai Meechan.

The main purpose of this dissertation was to investigate the reliability of the test estimated with the congeneric model (r_B and $r_{construct}$ co-efficiencies) and three - parameter logistic model (and KR-20 co-efficiencies).

For data gathering, 6,000 Prathomsuksa 5 students out of the 4 Chiangrai Education Services Area Offices divided into 4 population sizes: 300, 500, 1,000, and 1,500 were selected by the stratified random sampling technique. Otherwise, 30 sampling groups out of each the small-sized sample divided into 4 sizes: 30, 60,120, and 240 were all conducted with an achievement test on mathematical learning achievement designed for Prathomsuksa 5 students. The data were systematically analyzed through mean, standard deviation, the estimation of reliability, the standard error of the estimate, partiality, and Profile analysis.

The findings of the study were as follows:

1.In terms of the estimation of the test's reliability compared the psudo population with the 4 sample groups, including 300,500,1,000, and 1,500, it was stated that the reliability of the test with ρ_{IRT} , r_B , KR-20, and $r_{construct}$ co-efficiencies were respectively estimated;

2. In terms of the estimation of the test's reliability of a test compared with the 4 small-sized sample groups of 30, 60, 120, and 240, it was stated that the r_B , KR-20, and $r_{construct}$ co-efficiencies were respectively estimated;

3. In terms of the reliability of the test with the sample groups of 300, 500, 1,000, and 1,500 estimated with the KR-20 co-efficiency, it was stated that the r_B co-efficiency and $r_{construct}$ co-efficiency was averaged at its reliability of 95%;

4. In terms of the estimation of the test's reliability found in the standard error and partiality of the estimate with its reliability of $0 \geq 1$, it was stated that the appropriateness and quality of the test were found in the estimation of the reliability;

5. As compared the reliability of the test that was estimated with the KR-20, $\mathbf{r_B}$,

$r_{construct}$, and ρ_{IRT} co-efficiencies with the medium-sized sample group, it was also stated that the unparallel.

ปริญญานิพนธ์
เรื่อง
การประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และ
โมเดลโลจีสติก 3 พารามิเตอร์
ของ
กิตติศักดิ์ นิวรัตน์

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษาดุษฎีบัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

....คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร.สมชาย สันติวัฒนกุล)
วันที่..........เดือน พฤษภาคม พ.ศ. 2554

คณะกรรมการควบคุมปริญญานิพนธ์	คณะกรรมการสอบปากเปล่า
....ประธาน (รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์)	....ประธาน (อาจารย์ ดร.ปรมินทร์ อริเดช)
....กรรมการ (รองศาสตราจารย์ ดร.องอาจ นัยพัฒน์)	....กรรมการ (รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์)
....กรรมการ (ผู้ช่วยศาสตราจารย์ ดร.สุรชัย มีชาญ)	....กรรมการ (รองศาสตราจารย์ ดร.องอาจ นัยพัฒน์)
	....กรรมการ (ผู้ช่วยศาสตราจารย์ ดร.สุรชัย มีชาญ)
	....กรรมการ (อาจารย์ ดร.สุวพร เซ็มเฮง)

ประกาศคุณูปการ

ปริญญานิพนธ์นี้สำเร็จได้ด้วยดีเป็นเพราะผู้วิจัยได้รับความกรุณาอย่างยิ่ง ในการให้คำปรึกษา ให้ข้อเสนอแนะทั้งทางด้านทฤษฎี ด้านกระบวนการและคำแนะนำแนวทางในการวิจัย จาก รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ ประธานกรรมการควบคุมปริญญานิพนธ์ ผู้วิจัยขอกราบขอบพระคุณในความเสียสละและความเมตตาที่ท่านมอบให้

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.องอาจ นัยพัฒน์ และผู้ช่วยศาสตราจารย์ ดร.สุรชัย มีชาญ กรรมการควบคุมปริญญานิพนธ์ ผู้วิจัยรู้สึกซาบซึ้งในความเสียสละและความเมตตาที่ทุกท่านมอบให้ด้วยดีเสมอมา และได้รับคำแนะนำเพิ่มเติมจากอาจารย์ ดร.ปรมินทร์ อริเดช ประธานกรรมการสอบปากเปล่าปริญญานิพนธ์ และอาจารย์ ดร.สุวพร เซ็มเฮง กรรมการสอบปากเปล่าปริญญานิพนธ์ ทำให้ปริญญานิพนธ์นี้สมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณอาจารย์ทุกท่านที่อบรมสั่งสอน ให้ความรู้ และคุณธรรมของการเป็นนักวิชาการแก่ผู้วิจัยในการศึกษาตามหลักสูตรการทดสอบและวัดผลการศึกษา ทำให้ผู้วิจัยได้ความรู้ได้รู้คุณค่าของการศึกษาในระดับปริญญาเอก

ขอขอบคุณผู้เชี่ยวชาญด้านคณิตศาสตร์และวัดผลการศึกษาทุกท่านที่เสียสละเวลาในการตรวจสอบคุณภาพเครื่องมือที่ใช้ในการทำปริญญานิพนธ์นี้

ขอขอบคุณศึกษานิเทศก์ บุคลากรสำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ครูและนักเรียนที่เป็นกลุ่มประชากรเทียม และเพื่อนทุกคนที่ให้ความช่วยเหลือในการเก็บข้อมูลการวิจัย

ขอขอบคุณเพื่อนๆการทดสอบและวัดผลการศึกษา รหัส 48 ที่ให้คำปรึกษาและเป็นกำลังใจในการทำปริญญานิพนธ์ตลอดมา

สุดท้ายนี้ขอขอบพระคุณ พ่ออุ่นเรือน แม่ศรีจันทร์ พี่เกียรติชัย ที่คอยห่วงใย สนับสนุนในทุกๆด้านและเป็นกำลังใจที่สำคัญแก่ผู้วิจัยตลอดมา

กิตติศักดิ์ นิวรัตน์

สารบัญ

สารบัญ(ต่อ)

บทที่ หน้า

บัญชีตาราง

ตาราง **หน้า**

บัญชีตาราง (ต่อ)

บัญชีภาพประกอบ

ภาพประกอบ **หน้า**

บทที่ 1
บทนำ

ภูมิหลัง

ความเชื่อมั่น(Reliability) ของเครื่องมือวัดเป็นค่าที่แสดงถึงการมีคุณภาพของเครื่องมือวัดใช้เป็นหลักฐานแสดงว่า การวัดมีความเป็นอิสระหรือปลอดจากความคลาดเคลื่อน ได้ผลการวัดที่สอดคลองกัน(บุญเชิด ภิญโญอนันตพงษ์. 2547: 198) หรือคำนวณได้จากอัตราส่วนของความแปรปรวนของคะแนนจริงกับความแปรปรวนที่ได้จากการทดสอบ โดยคะแนนจริงจากการทดสอบนั้นเป็นคุณลักษณะภายในหรือคุณลักษณะแฝง(Latent Traits) ส่วนคะแนนที่ได้จากการทดสอบเป็นคะแนนที่สังเกตได้โดยตรง ในการประมาณค่าความเชื่อมั่นของเครื่องมือวัดนั้นสามารถประมาณค่าความเชื่อมั่นได้หลายวิธี มีวิธีประมาณค่าความเชื่อมั่นที่น่าสนใจ 2 วิธี คือ ทฤษฎีทดสอบคะแนนจริงสัมพันธ์(Congeneric Test Theory) และทฤษฎีการตอบข้อสอบ(Item Response Theory : IRT)

ทฤษฎีทดสอบคะแนนจริงสัมพันธ์ เป็นทฤษฎีย่อยของทฤษฎีการทดสอบมาตรฐานเดิม (Classical Test Theory: CTT) ซึ่งเป็นทฤษฎีแรกที่ได้พัฒนาขึ้นมาและมีการพัฒนาอย่างต่อเนื่องจนถึงปัจจุบัน ทฤษฎีการทดสอบมาตรฐานเดิมเชื่อว่าการวัดใด ๆจะมีผลกระทบจากความคลาดเคลื่อนเสมอ ดังนั้นคะแนนสอบ X (Observed Score) จึงประกอบด้วยคะแนนจริง T (True Score) และคะแนนคลาดเคลื่อน E (Error Score) สามารถเขียนเป็นสมการ ได้ $X_i = T_i + E_i$ เมื่อ คะแนนจริงและคะแนนคลาดเคลื่อนจากแบบทดสอบฉบับเดียวกันไม่มีความสัมพันธ์กัน ($Cov(T_i,E_i)=0$) คะแนนคลาดเคลื่อนจากแบบทดสอบต่างฉบับกันไม่มีความสัมพันธ์กัน ($Cov(E_i,E_j)=0 \quad ; \quad i \neq j$) และค่าเฉลี่ยของคะแนนคลาดเคลื่อนมีค่าเป็นศูนย์($\varepsilon(E_i)=0$) เมื่อแบบทดสอบแบ่งออกเป็นส่วนย่อย คะแนนรวม(Composite Score) X จะประกอบด้วย $X=X_1+X_2+\ldots+X_k= \sum X_i$ และคะแนนจริงรวม(True Score Composite) T จะประกอบด้วย $T = T_1+T_2+ \ldots +T_k= \sum T_i$ กรอบแนวคิดของโมเดลคะแนนจริงสัมพันธ์(*Jöreskog* .1971 :107-112) กำหนดคะแนนสอบมีส่วนประกอบ คือ $X_i = a_iT_i + b_i + E_i$ เมื่อ a_i และ b_i เป็นค่าคงที่ โดย a_i เป็นค่าอำนาจจำแนกของข้อสอบ b_i เป็นค่าความยากของข้อสอบ คะแนนจริง(T_i และ T_j) ของคะแนนสอบสองค่าใด ๆ(X_i และ X_j) จะมีความสัมพันธ์เป็นเส้นตรงเมื่อหาค่า a_{ij} และ b_{ij} ที่เหมาะสมได้ ทำให้ $T_i = a_{ij}T_j + b_{ij}$ ($1 \leq I , j \leq k$)

ถ้าผ่อนปรนให้คะแนนจริงของแต่ละส่วนไม่จำเป็นต้องเท่ากันพอดี แต่ยอมให้ต่างกันได้เท่ากับความยากที่ต่างกันในแต่ละส่วนและไม่จำเป็นต้องเท่ากับศูนย์เสมอไป ($b_i \neq b_j \neq 0$) ค่าอำนาจจำแนกของข้อสอบมีค่าเท่ากับ 1 ($a_i = a_j = 1$) และแต่ละส่วนมีค่าความแปรปรวนของคะแนนสอบต่างกันได้เล็กน้อย($Var(E_i) \approx Var(E_j)$) แต่ความแตกต่างเหล่านี้ต้องไม่โตมากนักเพื่อให้แต่ละส่วนยังคงความคู่ขนานไว้เพราะว่าแต่ละส่วนของแบบทดสอบมีขนาดความยาวเท่ากันหรือมีจำนวนข้อเท่ากัน ก็จะได้โมเดลคะแนนจริงสมมูล(Tau-Equivalent Model) คือ $T_i = T_j + b_{ij}$ แต่ในทาง

ปฏิบัติแบบทดสอบมีการแบ่งส่วนแต่ละส่วนตามลักษณะของแบบทดสอบทำให้มีขนาดความยาวหรือจำนวนข้อไม่เท่ากัน ทำให้มีผลกระทบต่อความแปรปรวนร่วมของคะแนนสอบ นอกจากนั้นถึงแม้ว่าจะสร้างแบบทดสอบที่แต่ละส่วนมีจำนวนข้อเท่ากัน แต่เมื่อไปทดสอบกับกลุ่มตัวอย่างแล้วพบว่า แต่ละส่วนมีการกระจายของคะแนนมากน้อยต่างกันมากๆ แสดงว่าความยาวที่แท้จริงของมัน (Functional Lengths) หรือความยาวที่เป็นผลมาจากการสอบ(Functionally) ของแต่ละส่วนมีขนาดไม่เท่ากัน นั่นก็คือแม้จำนวนข้อจะเท่ากันแต่ยังไม่เพียงพอที่จะบ่งชี้ความยาวที่แท้จริง ซึ่งโดยปกติจะไม่ทราบค่าจนกว่าจะนำแบบทดสอบไปทดสอบเสียก่อน นอกจากนี้ยังมีการวัดบางอย่างที่คะแนนเต็มของแต่ละส่วนไม่เท่ากัน หรือใช้ผู้ตรวจหลายคน ทำให้การกระจายของคะแนนของผู้ตรวจไม่เท่ากัน ส่งผลให้ไม่สอดคล้องกับโมเดลคะแนนจริงสมมูล โมเดลคะแนนจริงสมมูลสามารถประมาณค่าความเชื่อมั่นเมื่อแบ่งแบบทดสอบออกเป็นสองส่วนได้จากสูตรของรูลอน เรียกว่าสัมประสิทธิ์ r_{RU} สูตรของฟลานาแกน เรียกว่าสัมประสิทธิ์ r_{FL} และสูตรของกัตต์แมน เรียกว่าสัมประสิทธิ์ r_{Gutt} หากแบ่งแบบทดสอบออกเป็นหลายส่วนและมีการให้คะแนนแบบ 2 ค่า(Dichotomous) สามารถประมาณค่าความเชื่อมั่นได้จากสูตรของคูเดอร์และริชาร์ดสัน เรียกว่าสัมประสิทธิ์ KR-20 ใช้เมื่อข้อสอบแต่ละข้อมีความยากไม่เท่ากัน และ KR-21 ใช้เมื่อข้อสอบแต่ละข้อมีความยากเท่ากัน และสูตรของกัตแมน เรียกว่าสัมประสิทธิ์ r_{Gutt} แต่ถ้าหากแบ่งแบบทดสอบออกเป็นหลายส่วนและมีการให้คะแนนแบบหลายค่า(Polytomous) สามารถประมาณค่าความเชื่อมั่นได้จากสูตรของครอนบัค เรียกว่าสัมประสิทธิ์ r_{α} และสูตรของฮอยท์ เรียกว่าสัมประสิทธิ์ r_{Hoyt}

ถ้าผ่อนปรนให้คะแนนจริงของแต่ละส่วนไม่จำเป็นต้องเท่ากันพอดี แต่ยอมให้ต่างกันได้เท่ากับความยากที่ต่างกันในแต่ละส่วนและไม่จำเป็นต้องเท่ากับศูนย์เสมอไป ($b_i \neq b_j \neq 0$) ค่าอำนาจจำแนกของข้อสอบมีค่าไม่เท่ากัน ($a_i \neq a_j$) แต่ละส่วนมีคะแนนจริงต่างกันเท่ากับค่าคงที่ที่มาสัมพันธ์สมบูรณ์แบบ (Perfectly Correlation) หรือมีความสัมพันธ์เป็นเส้นตรง ค่าคงที่ a_{ij} และ b_{ij} จะแปรเปลี่ยนไปตามความสัมพันธ์ของแต่ละส่วนของแบบทดสอบ แต่ไม่ขึ้นกับการแปรเปลี่ยนของแต่ละบุคคล แต่ละส่วนมีค่าความแปรปรวนของคะแนนสอบไม่เท่ากัน ($Var(E_i) \approx Var(E_j)$) และแต่ละส่วนของแบบทดสอบมีขนาดความยาวไม่เท่ากันหรือมีจำนวนข้อไม่เท่ากัน แต่มีเนื้อหาเป็นเอกพันธ์กันหรือวัดคุณลักษณะเดียวกัน ซึ่งเป็นข้อตกลงที่สำคัญของการวัดทั้งในทฤษฎีการทดสอบมาตรฐานเดิม(CTT) และทฤษฎีการตอบสนองข้อสอบ(IRT) ก็จะได้โมเดลคะแนนจริงสัมพันธ์ (Congeneric Model) คือ $T_i = a_{ij}T_j + b_{ij}$ สามารถกล่าวได้ว่าโมเดลคะแนนจริงสมมูลเป็นกรณีเฉพาะของโมเดลคะแนนจริงสัมพันธ์ และเมื่อพิจารณาในทางปฏิบัติแล้ว โมเดลคะแนนจริงสมมูลมีเงื่อนไขมากกว่าโมเดลคะแนนจริงสัมพันธ์ทำให้มีโอกาสที่จะละเมิดเงื่อนไขมากกว่า ส่งผลให้การประมาณค่าความเชื่อมั่นขาดความถูกต้องแม่นยำ ดังนั้นการประมาณค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์จึงมีข้อดีกว่าแบบคะแนนจริงสมมูล

โมเดลคะแนนจริงสัมพันธ์สามารถประมาณค่าความเชื่อมั่นได้จากการคำนวณค่าความแปรปรวน-ความแปรปรวนร่วม และจากโมเดลการวัด(Measurement Model) โดยการประมาณค่า

ความเชื่อมั่นจากการคำนวณค่าความแปรปรวน-ความแปรปรวนร่วมสามารถคำนวณตามสูตรต่างๆ ตามเงื่อนไขของการแบ่งส่วนของแบบทดสอบออกเป็น 2 ส่วน 3 ส่วน และหลายส่วน ดังนี้ คริสทอฟ(Kristof. 1974: 491-499) เป็นผู้นำเสนอสูตรการประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งออกเป็น 3 ส่วนที่มีความยาวแต่ละส่วนไม่เท่ากัน เรียกสัมประสิทธิ์ r_K ต่อมา เฟลดต์(Feldt. 1975: 557-561) ได้ปรับปรุงวิธีการของคริสทอฟ ให้สามารถประมาณค่าความเชื่อมั่นของแบบทดสอบเมื่อแบ่งแบบทดสอบออกเป็น 2 ส่วนที่มีความยาวไม่เท่ากัน เรียกสัมประสิทธิ์ r_{FS} หลังจากนั้นก็มีการพัฒนาการประมาณค่าความเชื่อมมั่นเมื่อมีการแบ่งแบบทดสอบออกเป็นหลายส่วนที่มีความยาวของแต่ละส่วนไม่เท่ากัน โดยเริ่มจาก ราชู(Raju. 1977: 549-565) ที่พัฒนาจากสัมประสิทธิ์แอลฟ่า ของครอนบัค เรียกสัมประสิทธิ์ r_R ต่อมากิลเมอร์และเฟลดต์(Gilmer and Feldt. 1983: 99-111) ได้พัฒนาสูตรของคริสทอฟให้สามารถประมาณค่าความเชื่อมั่นของแบบทดสอบอัตนัยหรือการประเมินบุคคลด้วยกรรมการ ได้สัมประสิทธิ์ r_{F1} และ r_{F2} ต่อมาในปี 1989 เลียว(Liou. 1989: 153-163) ได้พัฒนาสูตรจากคริสทอฟเป็นสัมประสิทธิ์ r_{L1} และพัฒนาสูตรของเฟลดต์เป็นสัมประสิทธิ์ r_{L2} ให้สามารถประมาณค่าความเชื่อมั่นจากแบบทดสอบที่แบ่งเป็นหลายส่วน ซึ่งผลการทดลองสามารถคำนวณได้ง่ายกว่าและให้ผลตรงกับสูตร r_{F1} และ r_{F2} ของกิลเมอร์และเฟลดต์ และในปีเดียวกัน เฟลดต์และเบรนนอน (Feldt and Brennan. In Linn. 1989 อ้างอิงจาก บุญเชิด ภิญโญอนันตพงษ์. 2537: 5) ได้พัฒนาสูตรการประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งออกเป็นสองส่วน(r_{FS}) มาเป็นแบบแบ่งออกเป็นหลายส่วน เรียกสัมประสิทธิ์ r_{FK} ต่อมาในปี 2538 บุญเชิด ภิญโญอนันตพงษ์ ได้พัฒนาสัมประสิทธิ์ r_{L2} ของเลียว(Liou. 1989) เพื่อประมาณค่าความเชื่อมั่นของแบบทดสอบที่มีการให้คะแนนแบบ 2 ค่าและแต่ละข้อมีค่าความยากไม่เท่ากัน เรียกสัมประสิทธิ์ r_B โดยสัมประสิทธิ์ r_B จะคล้ายกับสัมประสิทธิ์ KR-20 แต่ใช้เงื่อนไขข้อตกลงของโมเดลคะแนนจริงสัมพันธ์ สัมประสิทธิ์ r_B จึงเป็นสัมประสิทธิ์ปรับขยายสัมประสิทธิ์ KR-20

การประมาณค่าความเชื่อมั่นจากโมเดลการวัด(Measurement Model) โจเรสคอก (*Jöreskog* .1971: 107-112) เป็นบุคคลแรกที่ใช้คำว่าการทดสอบคะแนนจริงสัมพันธ์(Congeneric Tests) ในทฤษฎีการทดสอบมาตรฐานเดิม (CTT) โดยได้เสนอโมเดลการทดสอบคะแนนจริงสัมพันธ์ที่วัดองค์ประกอบเดี่ยว (Model for One Set of Congeneric Test Score) ซึ่งใช้วิธีการประมาณค่าด้วยวิธีความน่าจะเป็นสูงสุด (maximum likelihood method) กำหนดให้ τ เป็นตัวแปรสุ่ม คะแนนจริงรวม(True Score Composite) T จะประกอบด้วย $T=T_1+T_2+\ldots+T_k=\sum T_i$ และแต่ละค่ามีความสัมพันธ์เชิงเส้นตรงกับ τ ก็จะได้โมเดลการทดสอบคะแนนจริงสัมพันธ์ที่วัดองค์ประกอบเดี่ยวคือ $T_i=\lambda_i\tau+\mu_i$ โดย λ_i และ μ_i เป็นค่าคงที่ซึ่งมีข้อตกลงว่า $\varepsilon(\tau)=0$ และ $Var(\tau)=1$ แต่วิธีประมาณค่าความเชื่อมั่นด้วยวิธีความน่าจะเป็นสูงสุด (ML) เป็นวิธีการคำนวณที่มีความสลับซับซ้อนมากในขณะนั้น (เมื่อปี1971) ต่อมา เรยูเตอร์เบิร์กและกูสตาฟสัน(Reuterberg and Gustafsson.1992: 795-811) ได้ทำการศึกษาการวิเคราะห์องค์ประกอบเชิงยืนยันและความ

เชื่อมั่น พร้อมทั้งนำเสนอการประมาณค่าความเชื่อมั่นของแต่ละองค์ประกอบและการประมาณค่าความเชื่อมั่นรวมด้วยโมเดลคะแนนจริงมาตรฐานเดิม โมเดลคะแนนจริงสมมูลและโมเดลคะแนนจริงสัมพันธ์ ต่อมาเรย์คอฟ (Raykov. 1997: 174-178) ได้นำวิธีการของ โจเรสคอก (*Jöreskog* .1971) มาใช้ในการประมาณค่าความเชื่อมั่นรวม(composite reliability) สำหรับการวัดคะแนนจริงสัมพันธ์ซึ่งเป็นการประมาณค่าโดยใช้โมเดลสมการโครงสร้าง (Structural Equation Model) จากโปรแกรม LISREL ที่ใช้วิธีประมาณค่าด้วยวิธีความน่าจะเป็นสูงสุด(ML) ที่พัฒนาโดย โจเรสคอก มาใช้ในการวัดคะแนนจริงสัมพันธ์ ที่เรียกว่า โมเดลการวัดคะแนนจริงสัมพันธ์ (Congeneric Measurement Model) โดยการประมาณค่าความเชื่อมั่นจะมีค่าเท่ากับผลรวมของน้ำหนักองค์ประกอบยกกำลังสองหารด้วยผลรวมของน้ำหนักองค์ประกอบยกกำลังสองบวกกับผลรวมของความแปรปรวนของความคลาดเคลื่อน เรียกว่า สัมประสิทธิ์โครงสร้าง ที่ได้จากการประมาณค่าความเชื่อมั่นทางโครงสร้าง (Construct Reliability)

ทฤษฎีการตอบข้อสอบ (Item Response Theory: IRT) มีฐานความคิดที่สำคัญ 2 ประการ คือ 1. ผลการตอบของผู้สอบสามารถอธิบายได้ด้วยความสามารถที่อยู่ภายในของผู้สอบ 2.ความสัมพันธ์ระหว่างผลการสอบกับความสามารถที่มีอยู่ภายใน สามารถอธิบายได้ด้วยฟังก์ชั่นลักษณะข้อสอบหรือโค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) อันมีลักษณะเป็นฟังก์ชันทางคณิตศาสตร์ เรียกว่าฟังก์ชันโลจีสติก (Logistic function) หรือฟังก์ชันปกติสะสม (Normal ogive function) โดยนิยมใช้ฟังก์ชันโลจีสติก มากกว่าฟังก์ชันปกติสะสม เนื่องจากฟังก์ชันโลจีสติก คำนวณง่ายกว่า และมีค่าเข้าใกล้แอซิมโถด (Asymptote) ช้ากว่าฟังก์ชันปกติสะสม (ผจงจิต อินทสุวรรณ.2545: 14 ; ศิริชัย กาญจนวาสี. 2545: 45-46) สำหรับโมเดลการตอบข้อสอบเมื่อมีการให้คะแนนแบบ 2 ค่า(Binary or Dichotomous IRT) ได้แก่ โมเดล 1 พารามิเตอร์ ที่มีเพียงพารามิเตอร์ของค่าความยาก(b) โมเดล 2 พารามิเตอร์ ที่มีพารามิเตอร์ของค่าความยาก(b) และค่าอำนาจจำแนก(a) และโมเดล 3 พารามิเตอร์ที่มีพารามิเตอร์ของค่าความยาก(b) ค่าอำนาจจำแนก(a) และ ค่าการเดา(c) ในปี 2003 ดิมิทรอฟ(Dimitrov. 2003: 440-458) ได้นำเสนอการแปลงค่าพารามิเตอร์ที่ได้จากทฤษฎีการตอบข้อสอบ(IRT) มาเป็นค่าความเชื่อมั่น จากแบบทดสอบที่ให้คะแนนแบบ 2 ค่า โดยทำการแปลงค่าพารามิเตอร์อำนาจจำแนก(a) ความยาก(b) การเดา(c) มาเป็นค่าความแปรปรวนของคะแนนจริงและความแปรปรวนของความคลาดเคลื่อนแล้วนำมาประมาณค่าความเชื่อมั่นด้วยการหาสัดส่วนของความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนจริงบวกความแปรปรวนของความคลาดเคลื่อน เรียกว่า ρ_{IRT}

เมื่อนำสูตรต่าง ๆของการประมาณค่าความเชื่อมั่นมาพิจารณาร่วมกับบริบทของประเทศไทยที่มีการวัดประเมินในระดับชั้นเรียน(Classroom Assessment) เพื่อสนับสนุนส่งเสริมการเรียนรู้ของผู้เรียน กับการวัดประเมินในระดับใหญ่(Large Scale)ที่เป็นการวัดประเมินมาตรฐานความรู้ในระดับกลุ่มโรงเรียน เขตพื้นที่การศึกษาและระดับประเทศ เป็นการวัดประเมินเพื่อสนับสนุนคุณภาพโดยรวมของโรงเรียน ในการวัดประเมินแต่ละระดับมีจำนวนนักเรียนต่างกัน เครื่องมือที่ใช้ในการวัด

ประเมินนิยมใช้แบบทดสอบวัดผลสัมฤทธิ์ชนิดเลือกตอบที่ให้คะแนนแบบ 2 ค่า(Dichotomous) คือ 0 กับ 1 และมีความยากของข้อสอบต่างกัน มีความเป็นเอกพันธ์หรือวัดเนื้อหาเดียวกัน มีจำนวนข้อสอบต่างกัน ซึ่งสามารถใช้ทฤษฎีการทดสอบมาตรฐานเดิมโมเดลคะแนนจริงสมมูลด้วยสัมประสิทธิ์ KR-20 ซึ่งเป็นกรณีเฉพาะของโมเดลคะแนนจริงสัมพันธ์ นักวิจัยส่วนใหญ่รู้จักและนิยมใช้โดยทั่วไป โมเดลคะแนนจริงสัมพันธ์ ด้วยสัมประสิทธิ์ r_B โมเดลการวัดคะแนนจริงสัมพันธ์ ด้วยสัมประสิทธิโครงสร้าง และทฤษฎีการตอบข้อสอบโมเดลโลจีสติก 3 พารามิเตอร์ ด้วยสัมประสิทธิ์ ρ_{IRT} มาใช้ในการประมาณค่าความเชื่อมั่นได้เหมาะสมกับบางระดับเท่านั้น ผู้วิจัยจึงมีความสนใจที่จะศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ด้วยสัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง และโมเดลโลจีสติก 3 พารามิเตอร์ ด้วยสัมประสิทธิ์ ρ_{IRT} ว่าสามารถประมาณค่าความเชื่อมั่นได้เมื่อกลุ่มตัวอย่างมีขนาดใด โดยใช้โมเดลคะแนนจริงสมมูลด้วยสัมประสิทธิ์ KR-20 เป็นเกณฑ์เทียบ แต่ละวิธีให้ค่าความคลาดเคลื่อนมาตรฐานและความลำเอียงมากน้อยเพียงใดและเมื่อกลุ่มตัวอย่างมีขนาดต่างกัน จะประมาณค่าความเชื่อมั่นของแบบทดสอบจาก 4 สูตร แตกต่างกันหรือไม่

ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีความมุ่งหมายหลักเพื่อศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ ซึ่งประกอบด้วยสัมประสิทธิ์ r_B กับสัมประสิทธิโครงสร้าง และโมเดลโลจีสติก 3 พารามิเตอร์ ได้แก่ สัมประสิทธิ์ ρ_{IRT} และใช้สัมประสิทธิ์ KR-20 เป็นเกณฑ์เทียบ โดยมีความมุ่งหมายเฉพาะดังนี้

1. เพื่อประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากประชากรเทียม

2. เพื่อประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ ขนาด 300 500 1,000 และ 1,500 คน

3. เพื่อประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง จากกลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ ขนาด 30 60 120 และ 240 คน

4. เพื่อศึกษาความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง และสัมประสิทธิ์ ρ_{IRT}

5. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง

ความสำคัญของการวิจัย

การวิจัยครั้งนี้ เป็นการศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และโมเดลโลจีสติก 3 พารามิเตอร์ โดยมีความสำคัญในทางทฤษฎีและทางปฏิบัติดังนี้

1.เนื่องจากในยุคของการพัฒนาการประมาณค่าความเชื่อมั่นได้มีการคิดค้นและพัฒนาสูตรต่าง ๆอย่างมากมาย มีสูตรที่น่าสนใจในกรณีที่แบบทดสอบมีการให้คะแนน 2 ค่า คือ 0 กับ 1 และมีความยากของข้อสอบต่างกัน วัดจากแบบทดสอบที่มีความเป็นเอกพันธ์หรือวัดเนื้อหาเดียวกัน ได้แก่ สัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} โดยสัมประสิทธิ์ KR-20 เป็นสูตรที่มีชื่อเสียงและใช้กันอย่างแพร่หลาย เป็นสูตรหนึ่งในโมเดลคะแนนจริงสมมูล ส่วนสัมประสิทธิ์ r_B และสัมประสิทธิ์โครงสร้าง เป็นสูตรในโมเดลคะแนนจริงสัมพันธ์ ที่ประมาณค่าจากความแปรปรวน-ความแปรปรวนร่วม กับ โมเดลการวัด ตามลำดับ สัมประสิทธิ์ r_B มีการพัฒนาจากสัมประสิทธิ์ r_{L2} ส่วนสัมประสิทธิ์โครงสร้าง ประมาณค่าจากโมเดลการวัดมีการคิดค้นขึ้นมาเป็นวิธีการแรกใช้วิธีประมาณค่าด้วยวิธีความน่าจะเป็นสูงสุด(Maximum Likelihood method :ML) มีความซับซ้อนในเรื่องของการคำนวณในขณะนั้น ปัจจุบันสามารถประมาณค่าด้วยวิธีความน่าจะเป็นสูงสุดได้จากโปรแกรม LISREL แต่นักวิจัยส่วนใหญ่นิยมนำไปใช้ในการวิเคราะห์องค์ประกอบ การวิเคราะห์อิทธิพล การวิเคราะห์กลุ่มพหุ และโมเดลโค้งพัฒนาการ มีผู้สนใจที่จะใช้ในการประมาณค่าความเชื่อมั่นน้อย สำหรับสัมประสิทธิ์ ρ_{IRT} เป็นการประมาณค่าความเชื่อมั่นจากทฤษฎีการตอบข้อสอบ(IRT) ด้วยการแปลงค่าพารามิเตอร์ของโมเดลโลจีสติก 3 พารามิเตอร์มาเป็นค่าความเชื่อมั่น โดยใช้โปรแกรม BILOG-MG3 ในการหาค่าพารามิเตอร์แล้วใช้โปรแกรม SPSS ในการแปลงค่าพารามิเตอร์เป็นค่าความเชื่อมั่น ดังนั้นการประมาณค่าความเชื่อมั่นของแบบทดสอบทั้ง 4 สูตร จึงเป็นการขยายขอบเขตของการศึกษาค้นคว้าเกี่ยวกับการประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่มีการใช้ทฤษฎีทดสอบคะแนนจริงสัมพันธ์และทฤษฎีการตอบข้อสอบและใช้โปรแกรมคอมพิวเตอร์เข้ามาช่วยในการวิเคราะห์

2.การจะนำสูตรการประมาณค่าความเชื่อมั่นทั้ง 4 สูตรไปใช้ในบริบทของประเทศไทยก็ต้องมาพิจารณาที่ระดับของการวัดประเมิน ที่มีอยู่ 2 ระดับ คือระดับชั้นเรียนและระดับใหญ่ ในแต่ละระดับก็จะมีขนาดกลุ่มนักเรียนแตกต่างกัน เครื่องมือที่นิยมใช้ในทั้งสองระดับก็เป็นแบบทดสอบที่มีการให้คะแนนแบบ 2 ค่า คือ 0 กับ 1 หากจะนำสูตรทั้งสี่มาใช้ในการประมาณค่าความเชื่อมั่นของแบบทดสอบให้ตรงกับระดับการวัดประเมินและเครื่องมือได้อย่างเหมาะสม ควรที่จะมีงานวิจัยที่ทำการศึกษาการประมาณค่าความเชื่อมั่นที่บอกได้ว่าสูตรใดเหมาะสมกับกลุ่มนักเรียนขนาดใด สูตรไหนคงที่ คงเส้นคงวา มีความถูกต้อง ซึ่งผลการศึกษาครั้งนี้จะเป็นแนวทางในการตัดสินใจเลือกใช้สูตรการประมาณค่าความเชื่อมั่นได้อย่างเหมาะสมกับบริบทของโรงเรียนและผู้ใช้ ซึ่งเป็นอีกทางหนึ่งที่ใช้ในการพิจารณาคุณภาพของแบบทดสอบ ให้ผู้ใช้พิจารณาเลือกใช้ได้อย่างเหมาะสม

ขอบเขตของการวิจัย

การวิจัยครั้งนี้จำกัดอยู่ในขอบเขตของการศึกษาดังต่อไปนี้

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ปีการศึกษา 2550 จำนวน 15,395 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างขนาดใหญ่ที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย จำนวน 4 เขตพื้นที่ ปีการศึกษา 2550 โดยใช้วิธีการสุ่มแบบแบ่งชั้น(Stratified Random Sampling) จำนวน 6,000 คน

ซึ่งกลุ่มตัวอย่างขนาดใหญ่นี้ ผู้วิจัยกำหนดให้เป็นประชากรเทียม(Pseudo Population)

กลุ่มตัวอย่างขนาดกลางและขนาดเล็กที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย จำนวน 4 เขตพื้นที่ ปีการศึกษา 2550 โดยใช้วิธีการสุ่มแบบใส่คืนจากประชากรเทียม แบ่งเป็นกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ 300 500 1,000 และ 1,500 คน กลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ 30 60 120 และ 240 คน แต่ละขนาดจะมีจำนวน 30 กลุ่ม

ตัวแปรที่ศึกษา

ตัวแปรอิสระ ได้แก่

1. สูตรการประมาณค่าความเชื่อมั่นของแบบทดสอบ 4 สูตร คือ

1.1 สัมประสิทธิ์ KR-20

1.2 สัมประสิทธิ์ r_B

1.3 สัมประสิทธิ์โครงสร้าง ที่ได้จากการประมาณค่าความเชื่อมั่นจากโมเดลการวัด

1.4 สัมประสิทธิ์ ρ_{IRT} ที่ได้จากโมเดลโลจีสติก 3 พารามิเตอร์ของทฤษฎีการตอบข้อสอบ

2. ขนาดกลุ่มตัวอย่าง ได้แก่

2.1 กลุ่มตัวอย่างขนาดกลาง ได้แก่ กลุ่มตัวอย่างขนาด 300 500 1,000 และ 1,500 คน โดยแต่ละขนาดจะมีจำนวน 30 กลุ่ม

2.2 กลุ่มตัวอย่างขนาดเล็ก ได้แก่ กลุ่มตัวอย่างขนาด 30 60 120 และ 240 คน โดยแต่ละขนาดจะมีจำนวน 30 กลุ่ม

ตัวแปรตาม คือ ค่าความเชื่อมั่นของแบบทดสอบ

ข้อตกลงของการประมาณค่าความเชื่อมั่น (อยู่ก่อนนิยามศัพท์)

การวิจัยครั้งนี้เป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากประชากรเทียม กลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ 300 500 1,000 1,500 คน กลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ 30 60 120 240 คน โดยเป็นการประมาณค่าของแต่ละสูตรเทียบกับการประมาณค่าความเชื่อมั่นที่แท้จริงที่ได้จากการประมาณค่าจากประชากรเทียม ไม่ได้ประมาณค่าความเชื่อมั่นที่แท้จริงกลางจากทุกสูตรรวมกัน

นิยามศัพท์เฉพาะ

1. ความเชื่อมั่นของแบบทดสอบ หมายถึง สัดส่วนของความแปรปรวนของคะแนนจริงต่อความแปรปรวนของคะแนนที่สอบได้

2. โมเดลคะแนนสมมูล(Tau-Equivalent Model) หมายถึง โมเดลที่มีลักษณะของคะแนนจริง คือ $T_i = T_j + b_{ij}$ เมื่อ T_i และ T_j แทนคะแนนจริงของแต่ละข้อ b_{ij} แทนความยาก ประมาณค่าความเชื่อมั่นด้วยวิธีการคำนวณของคูเดอร์-ริชาร์ดสัน ที่เรียกว่า สัมประสิทธิ์ KR-20

3. โมเดลคะแนนจริงสัมพันธ์(Congeneric Model) หมายถึง โมเดลที่มีลักษณะของคะแนนจริง คือ $T_i = a_{ij}T_j + b_{ij}$ เมื่อ T_i และ T_j แทนคะแนนจริงของแต่ละข้อ b_{ij} แทนความยาก a_{ii} แทนอำนาจจำแนกของข้อสอบ สามารถประมาณค่าได้ 2 วิธีคือ คำนวณจากค่าความแปรปรวน-ความแปรปรวนร่วม ด้วยวิธีการคำนวณของบุญเชิด (2540) ที่เรียกว่า สัมประสิทธิ์ r_B และคำนวณจากโมเดลการวัดคะแนนจริงสัมพันธ์

4. โมเดลการวัดคะแนนจริงสัมพันธ์(Congeneric Measurement Model) หมายถึง โมเดลการวัด(Measurement Model)ย่อยหนึ่งของโมเดลสมการโครงสร้าง(Structural Equation Model) และโมเดลสมการโครงสร้างก็เป็นโมเดลย่อยหนึ่งของโมเดลลิสเรล (LISREL Model) ที่ใช้ในการวัดคะแนนจริงสัมพันธ์ ประกอบไปด้วยตัวแปรภายนอกแฝงและตัวแปรภายนอกสังเกตได้ ไม่มีตัวแปรภายใน โดยตัวแปรภายนอกแฝงก็คือผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ตัวแปรภายนอกสังเกตได้คือข้อสอบวิชาคณิตศาสตร์ จำนวน 50 ข้อ ตามโมเดลการวัดองค์ประกอบเดี่ยวคะแนนจริงสัมพันธ์

5. สัมประสิทธิ์โครงสร้าง หมายถึง ค่าที่ได้จากการประมาณค่าความเชื่อมั่นทางโครงสร้าง (Construct Reliability) ที่ได้จากการประมาณค่าความเชื่อมั่นจากโมเดลการวัดคะแนนจริงสัมพันธ์จากโปรแกรม LISREL มีค่าเท่ากับผลรวมของน้ำหนักองค์ประกอบยกกำลังสองหารด้วยผลรวมของน้ำหนักองค์ประกอบยกกำลังสองบวกกับผลรวมของความแปรปรวนของความคลาดเคลื่อน

6. โมเดลโลจีสติก 3 พารามิเตอร์ หมายถึง โมเดลตามทฤษฎีการตอบข้อสอบ (Item Response Theory: IRT) เมื่อมีการให้คะแนนแบบ 2 ค่า โดยใช้ฟังก์ชันโลจีสติก (Logistic function)

ซึ่งเป็นฟังก์ชันทางคณิตศาสตร์ ในการอธิบายความสัมพันธ์ระหว่างผลการสอบกับความสามารถที่มีอยู่ภายใน มีพารามิเตอร์ของค่าอำนาจจำแนก(a) ค่าความยาก(b)และ ค่าการเดา (c)

7. สัมประสิทธิ์ ρ_{IRT} หมายถึง ค่าที่ได้จากการประมาณค่าความเชื่อมั่นจากโมเดลโลจีสติก 3 พารามิเตอร์ของทฤษฎีการตอบข้อสอบ(IRT)มาเป็นค่าความเชื่อมั่นด้วยการหาสัดส่วนของความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนจริงบวกความแปรปรวนของความคลาดเคลื่อน

8. แบบทดสอบ หมายถึง แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 จำนวน 50 ข้อ

9. กลุ่มตัวอย่างขนาดใหญ่ หมายถึง นักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ปีการศึกษา 2550 ที่ได้จากการสุ่มแบบแบ่งชั้นจากประชากร จำนวน 6,000 คน ซึ่งกลุ่มตัวอย่างขนาดใหญ่นี้ผู้วิจัยกำหนดให้เป็นประชากรเทียม

10. กลุ่มตัวอย่างขนาดกลาง หมายถึง นักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ปีการศึกษา 2550 ที่ได้จากการสุ่มแบบใส่คืนจากประชากรเทียม โดยมีขนาดกลุ่มละ 300 500 1,000 และ 1,500 คน แต่ละขนาดมีจำนวนจำนวน 30 กลุ่ม

11. กลุ่มตัวอย่างขนาดเล็ก หมายถึง นักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ปีการศึกษา 2550 ที่ได้จากการสุ่มแบบใส่คืนจากประชากรเทียม โดยมีขนาดกลุ่มละ 30 60 120 และ 240 คน แต่ละขนาดมีจำนวนจำนวน 30 กลุ่ม

12. ความเชื่อมั่นที่แท้จริง(True Reliability) หมายถึง ค่าความเชื่อมั่นที่ประมาณด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากประชากรเทียม

กรอบแนวคิดในการวิจัย

ในการศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามทฤษฎีการทดสอบมาตรฐานเดิมโมเดลคะแนนจริงสมมูลและโมเดลคะแนนจริงสัมพันธ์นั้น บุญเชิด ภิญโญอนันตพงษ์ (2542: บทคัดย่อ) ได้ทำการศึกษาสัมประสิทธิ์ r_B : การประมาณค่าความเชื่อมั่นสำหรับแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยความยากรายข้อต่างกัน เพื่อพัฒนาและศึกษาผลการใช้สูตร r_B ที่ผู้วิจัยพัฒนาขึ้นมาสำหรับประมาณค่าความเชื่อมั่นของแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยข้อสอบที่ให้คะแนนระบบ 0, 1 และมีความยากรายข้อต่างกัน โดยศึกษาจากแบบทดสอบที่วัดเนื้อหาเดียวกัน จำนวนข้อต่างกัน จำนวน 6 ฉบับ คือ ฉบับ 25, 30, 35, 40, 45, และ 50 ข้อ ทั้งจากกลุ่มตัวอย่างขนาดใหญ่ จำนวน 6 กลุ่มคือ 412, 400, 410, 403, 405, 406 คน รวม 2,436 คน และกลุ่มตัวอย่างขนาดเล็ก คือ 30, 50 และ 80 คนว่าเป็นอย่างไรเมื่อเทียบกับสูตร KR-20 ผลการศึกษาพบว่า สูตร r_B คำนวณค่าความเชื่อมั่นของแบบทดสอบจำนวนข้อต่างกันจากกลุ่มตัวอย่างขนาดใหญ่ได้สูงสอดคล้องสัมพันธ์กับสูตร KR-20 โดยสูตร r_B คำนวณค่าความ

เชื่อมั่นได้สูงกว่าสูตร KR-20 ทุกฉบับ เมื่อศึกษาจากกลุ่มตัวอย่างขนาดเล็ก พบว่า สูตร r_B คำนวณจากแบบทดสอบฉบับเดียวกันและกลุ่มตัวอย่างขนาดเดียวกันได้สอดคล้องสัมพันธ์กับสูตร KR-20 จากแบบทดสอบทั้ง 6 ฉบับและกลุ่มตัวอย่างทุกขนาด สูตร r_B คำนวณค่าความเชื่อมั่นสูงกว่าสูตร KR-20 ส่วนจำนวนข้อสอบและขนาดกลุ่มตัวอย่างมีผลต่อค่าความเชื่อมั่นที่คำนวณได้จากทั้งสองสูตรในลักษณะคล้ายคลึงกัน และทั้งสองสูตรประมาณค่าความเชื่อมั่นจากกลุ่มตัวอย่างขนาดเล็กเทียบกับกลุ่มตัวอย่างขนาดใหญ่มีค่าความลำเอียงในทิศทางเดียวกัน แสดงให้เห็นว่าสัมประสิทธิ์ KR-20 และ r_B สามารถนำไปใช้ในการคำนวณค่าความเชื่อมั่นจากกลุ่มตัวอย่างขนาดเล็ก คือ 30 50 และ 80 คน และกลุ่มตัวอย่างขนาดใหญ่ ได้

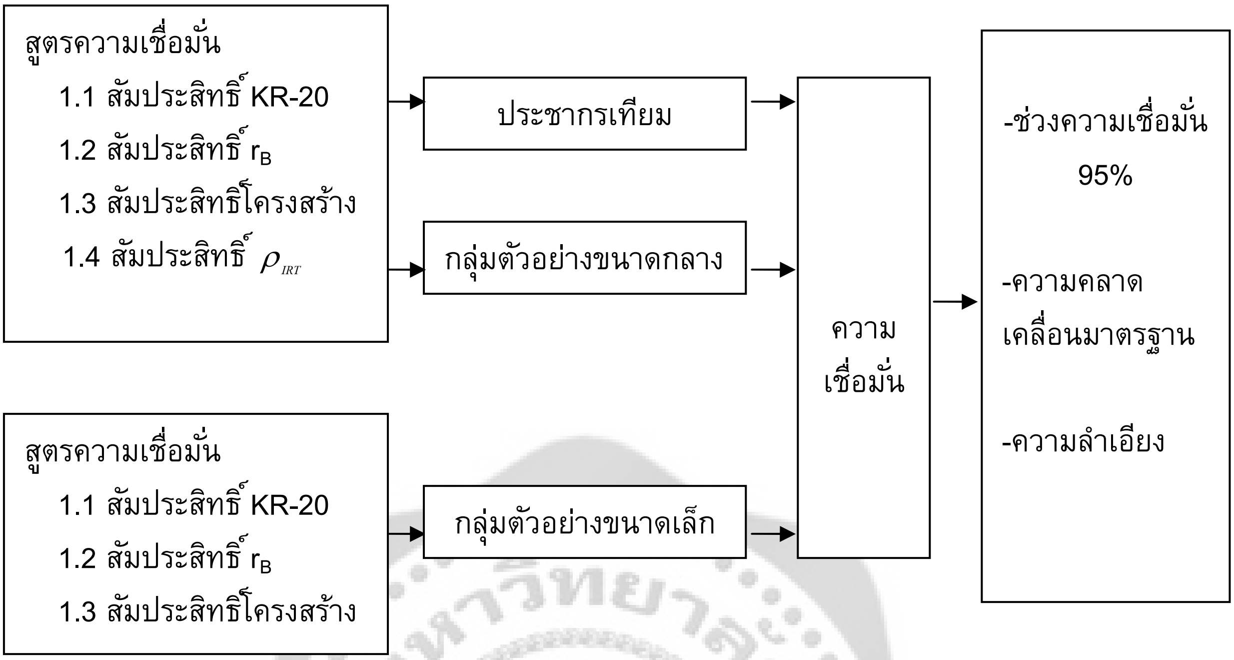
โมเดลการวัดคะแนนจริงสัมพันธ์ ซึ่งเป็นโมเดลย่อยหนึ่งของโมเดลสมการโครงสร้างและโมเดลสมการโครงสร้างก็เป็นโมเดลย่อยหนึ่งของโมเดลลิสเรล โดยชูมัคเกอร์และโลแมค (Schumacker and Lomax. 1996 อ้างอิงจาก นงลักษณ์ วิรัชชัย.2542: 311) ได้สรุปงานวิจัยที่มีการศึกษาเรื่องขนาดของกลุ่มตัวอย่างที่เหมาะสมสำหรับใช้โมเดลลิสเรลว่า จากงานวิจัยของดิง เวลิเซอร์และฮาร์โล(Ding, Velicer and Harlow) ในปี 1995 พบว่างานวิจัยที่ใช้โมเดลลิสเรลส่วนใหญ่ใช้ขนาดกลุ่มตัวอย่าง 100-150 คนและให้ผลการวิจัยที่น่าพอใจ บูมสมา(Boomsma) เสนอไว้ในบทความเมื่อปี 1983 ว่าขนาดกลุ่มตัวอย่างที่เหมาะสมควรเป็น 400 คน นอกจากนี้ ชูมัคเกอร์และโลแมค(Schumacker and Lomax. 1996) แฮร์และคณะ(Hait and other. 1998) เสนอให้ใช้กฎแห่งความชัดเจน(rule of thump) ที่นักสถิติตัวแปรพหุใช้กันมากคือ ใช้ขนาดกลุ่มตัวอย่าง 10-20 คนต่อตัวแปรในการวิจัยหนึ่งตัวแปร ทั้งนี้ต้องพิจารณาความซับซ้อนของโมเดลที่ใช้ด้วย ถ้าโมเดลมีความซับซ้อนมากและการแจกแจงของตัวแปรไม่เป็นโค้งปกติ ต้องเพิ่มขนาดกลุ่มตัวอย่างให้รับกัน เมื่อนำโมเดลการวัดคะแนนจริงสัมพันธ์มาใช้ในการประมาณค่าความเชื่อมั่นตามที่เรย์คอฟ (Raykov.1997: 174-178) ได้เสนอไว้ โดยการประมาณค่าความเชื่อมั่นจะมีค่าเท่ากับผลรวมของน้ำหนักองค์ประกอบยกกำลังสองหารด้วยผลรวมของน้ำหนักองค์ประกอบยกกำลังสองบวกกับผลรวมของความแปรปรวนของความคลาดเคลื่อน ก็จะได้ค่าความเชื่อมั่นที่ประมาณค่าได้จากโมเดลการวัดคะแนนจริงสัมพันธ์

โมเดลโลจีสติก 3 พารามิเตอร์ ซึ่งเป็นโมเดลหนึ่งของทฤษฎีการตอบข้อสอบ(IRT) เมื่อมีการให้คะแนนแบบ 2 ค่า โดยฮูลิน ลิสแสคและดราสโกว(Hulin, Lissak and Drasgow. 1982 249-260) ได้กล่าวถึงจำนวนข้อสอบที่มีมากนั้นไม่มีความจำเป็นเท่ากับจำนวนผู้เข้าสอบ โดยได้เสนอว่าโมเดลโลจิสติก 3 พารามิเตอร์ ที่มีความยาวข้อสอบจำนวน 50 ข้อ ขนาดกลุ่มตัวอย่างที่เพียงพอก็คือ 1,00 คน ในปีต่อมาฮูลิน ดราสโกวและพาร์สัน(Hulin, Drasgow and Parson. 1983: 99-105) ได้กล่าวถึงปัจจัยที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ ได้แก่ ความยาวแบบทดสอบ ขนาดกลุ่มตัวอย่าง และจำนวนพารามิเตอร์ของโมเดลโลจีสติก จากผลการวิจัยพบว่า ความยาวของแบบทดสอบและขนาดกลุ่มตัวอย่างเป็นปัจจัยสำคัญที่ส่งผลต่อการประมาณค่าพารามิเตอร์ โดยได้เสนอแนะว่าโมเดลโลจีสติก 3 พารามิเตอร์เพื่อประมาณค่าพารามิเตอร์ด้วยวิธีความน่าจะเป็นสูงสุด

(ML)ให้ได้ผลถูกต้องนั้นควรใช้แบบทดสอบที่มีจำนวน 60 ข้อกับกลุ่มตัวอย่างจำนวนไม่น้อยกว่า 200 คน หรือถ้าใช้แบบทดสอบที่มีจำนวนข้อน้อยกว่า 30 ข้อ จะต้องใช้กับกลุ่มตัวอย่างไม่น้อยกว่า 1,000 คน ในปีเดียวกันสวามินาธานและกิฟฟอร์ด (Swaninathan and Gifford. 1983) ได้ศึกษาการประมาณค่าพารามิเตอร์ในโมเดล 3 พารามิเตอร์ โดยใช้กลุ่มตัวอย่างขนาด 50, 200, และ 1,000 คน พบว่าขนาดกลุ่มตัวอย่างที่เพิ่มขึ้นส่งผลน้อยมากต่อความแม่ยำในการประมาณค่าพารามิเตอร์ความยาก(b)และพารามิเตอร์การเดา(c) แต่การเพิ่มขึ้นทั้งขนาดกลุ่มตัวอย่างและความยาวของแบบทดสอบ สามารถทำให้มีความแม่นยำในการประมาณค่าพารามิเตอร์ เช่น แบบทดสอบจำนวน 20 ข้อ ทดสอบกับผู้สอบจำนวน 1,000 คน สามารถประมาณค่าพารามิเตอร์ความยากและการเดาได้เป็นอย่างดี และสำหรับแบบทดสอบจำนวน 80 ข้อ ที่ทดสอบกับผู้สอบจำนวน 1,000 คน สามารถประมาณค่าพารามิเตอร์ได้ดีทุกค่า ในขณะที่ลอร์ด (Lord.1968 อ้างอิงจาก Hulin, Drasgow and Parson. 1983: 100-139) ได้กล่าวว่าการประมาณค่าพารามิเตอร์ด้วยวิธีความน่าจะเป็นสูงสุด (ML) ในโมเดลโลจีสติก 3 พารามิเตอร์ ให้มีความถูกต้องนั้น ควรใช้ข้อสอบประมาณ 50 ข้อกับกลุ่มตัวอย่างอย่างน้อย 1,000 คน เมื่อได้ค่าพารามิเตอร์จากการประมาณค่าแล้วก็จะนำมาแปลงเป็นค่าความเชื่อมั่น โดยในปี 2003 ดิมิทรอฟ(Dimitrov. 2003: 440) ได้นำเสนอการแปลงค่าพารามิเตอร์ที่ได้จากทฤษฎีการตอบข้อสอบ(IRT)มาเป็นค่าความเชื่อมั่น จากแบบทดสอบที่ให้คะแนนแบบ 2 ค่า โดยทำการแปลงค่าพารามิเตอร์อำนาจจำแนก(a) ความยาก(b) การเดา(c) มาเป็นค่าความแปรปรวนของคะแนนจริงและความแปรปรวนของความคลาดเคลื่อนแล้วนำมาประมาณค่าความเชื่อมั่นด้วยการหาสัดส่วนของความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนจริงบวกความแปรปรวนของความคลาดเคลื่อน

เมื่อพิจารณาในบริบทของโรงเรียนที่ประกอบด้วยโรงเรียนขนาดเล็ก ที่มีนักเรียนในแต่ละห้องประมาณ 30 คนเพียงห้องเดียว และโรงเรียนขนาดใหญ่ที่มีนักเรียนในแต่ละห้องประมาณ 60 คนและมีจำนวนหลายห้องเรียน การประมาณค่าความเชื่อมั่นด้วยสูตรทั้ง 4 สูตร เมื่อใช้ประชากรเทียมจะสามารถประมาณค่าได้เท่าไหร่ เมื่อใช้กลุ่มตัวอย่างขนาดกลางและเล็กทั้ง 4 สูตรจะสามารถประมาณค่าได้จริงในทางปฏิบัติที่ขนาดกลุ่มตัวอย่างขนาดใด และประมาณค่าได้ถูกต้อง ที่ขนาดกลุ่มตัวอย่างขนาดใด

จากที่กล่าวมาข้างต้นสามารถสรุปเป็นกรอบแนวคิดในการวิจัยได้ดังนี้



สมมติฐานในการวิจัย

1. สูตรการประมาณค่าความเชื่อมั่น 4 สูตร สามารถประมาณค่าความเชื่อมั่นอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริง

2. สมมติฐานของการวิเคราะห์เส้นภาพ(Profile Analysis)

2.1 กลุ่มตัวอย่างแต่ละขนาดมีเส้นภาพความเชื่อมั่น ขนานกัน

2.2 ค่าความเชื่อมั่นที่คำนวณได้แต่ละสูตรเป็นระนาบกัน

2.3 กลุ่มตัวอย่างแต่ละขนาดมีเส้นภาพความเชื่อมั่นอยู่ในระดับเดียวกัน

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้เป็นการศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และโมเดลโลจีสติก 3 พารามิเตอร์ ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง และได้นำเสนอตามหัวข้อต่อไปนี้

1. ทฤษฎีการทดสอบมาตรฐานเดิม
2. การประมาณค่าความเชื่อมั่นโมเดลคะแนนจริงสมมูล
3. การประมาณค่าความเชื่อมั่นโมเดลคะแนนจริงสัมพันธ์
   3.1 ความแปรปรวน-ความแปรปรวนร่วม
   3.2 โมเดลการวัดคะแนนจริงสัมพันธ์
4. ทฤษฎีการตอบสนองข้อสอบ
   4.1 โมเดลโลจีสติก 3 พารามิเตอร์
   4.2 การประมาณค่าความเชื่อมั่นจากโมเดลโลจีสติก 3 พารามิเตอร์
5. คำอธิบายรายวิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5
6. งานวิจัยที่เกี่ยวข้อง

ทฤษฎีการทดสอบมาตรฐานเดิม(Classical Test Theory=CTT)

ทฤษฎีการทดสอบมาตรฐานเดิมมีข้อตกลงว่า คะแนนที่สอบได้(Observed Score) แทนด้วย X จะประกอบด้วยคะแนนสองส่วนคือ คะแนนจริง(True Score) แทนด้วย T และคะแนนคลาดเคลื่อน(Error Score) แทนด้วย E สามารถเขียนเป็นสมการได้ดังนี้

X	=	T	+	E
คะแนนสอบ	=	คะแนนจริง	+	ความคลาดเคลื่อน

คะแนนจริงเป็นคะแนนที่แทนความสามารถจริงของบุคคล เป็นค่าวัดทางทฤษฎีไม่สามารถหาได้โดยตรงว่ามีค่าเท่าใด แต่ถือว่าเป็นคะแนนเฉพาะบุคคลที่ปลอดจากสภาวะภายนอกและภายใน ที่จะทำให้คะแนนเพิ่มหรือลดจากความเป็นจริง เรียกได้ว่าสภาพการทดสอบอยู่ในสภาพสมบูรณ์(Perfect)และใช้เครื่องมือวัดที่สมบูรณ์ทุกประการ ในทางทฤษฎีการกะประมาณคะแนนจริงของคนที่ j นิยามว่าเป็นค่าเฉลี่ยของคะแนนที่ได้จากการทดสอบซ้ำจำนวนมากด้วยแบบทดสอบคู่ขนาน ดังนั้นก็จะได้คะแนนจริงมีค่าเท่ากับ $T_j = \frac{\sum_{j=1}^{N} X}{N}$ ค่าเฉลี่ยของคะแนนที่สังเกตได้จึงเป็น

ตัวกะประมาณของคะแนนจริง และคะแนนจริงของแต่ละคนมีค่าคงที่ การที่คะแนนสอบมีค่าต่างกันเป็นเพราะคะแนนความคลาดเคลื่อน

คะแนนความคลาดเคลื่อนจากการวัดมี 2 ชนิด คือ ความคลาดเคลื่อนที่เป็นระบบ กับความคลาดเคลื่อนสุ่ม

ระดับของความคู่ขนาน (Degree of Parallelism)

ความเชื่อมั่นของเครื่องมือวัดต้องนิยามจากความคู่ขนานซึ่งตามทฤษฎีมาตรฐานเดิมได้กล่าวถึงความคู่ขนาน 3 ระดับ(บุญเชิด ภิญโญอนันตพงษ์. 2547) คือ

1. ความคู่ขนานแบบมาตรฐานเดิม (Classical Parallel)
2. ความคู่ขนานแบบคะแนนจริงสมมูล (Tau-Equivalent)
3. ความคู่ขนานแบบคะแนนจริงสัมพันธ์ (Congeneric)

ข้อตกลง	**ระดับความคู่ขนาน**		
	Classical Parallel	Tau-Equivalent	Congeneric
1) มีความเอกพันธ์ในเนื้อหา หรือวัดคุณลักษณะเดียวกัน (Homogeneous)	ต้องมี	ต้องมี	ต้องมี
2) มีคะแนนจริงเท่ากัน ($T_{i1} = T_{i2} = T_{i3} = \ldots$) และมีความแปรปรวนความคลาดเคลื่อนเท่ากัน ($\sigma_{E1}^2 = \sigma_{E2}^2 = \sigma_{E3}^2 = \ldots$)	ต้องมี	ผ่อนปรน $T_i = T_j + b_{ij}$	ผ่อนปรน $T_i = a_{ij}T_j + b_{ij}$
3) มีคะแนนสอบ(X)เฉลี่ยเท่ากัน ($\mu_1 = \mu_2 = \mu_3 = \ldots$)	ต้องมี	ผ่อนปรน	ผ่อนปรน
4) ความแปรปรวนของคะแนนสอบ (X)เท่ากัน ($\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \ldots$)	ต้องมี	ผ่อนปรน	ผ่อนปรน
5) ความแปรปรวนร่วมของคะแนนสอบ(X) เท่ากัน ($\sigma_{12} = \sigma_{13} = \sigma_{23} = \ldots$)	ต้องมี	ต้องมี	ผ่อนปรน
6) ความแปรปรวนร่วมของคะแนนสอบ(X)กับคะแนนเกณฑ์ภายนอกเท่ากัน ($\sigma_{1Y} = \sigma_{2Y} = \sigma_{3Y} = \ldots$)	ต้องมี	ต้องมี	ผ่อนปรน

ความคู่ขนานแบบมาตรฐานเดิม(Classical Parallel) ต้องมีข้อตกลงครบทุกข้อ ก็จะได้โมเดลคะแนนจริงมาตรฐานเดิมคือ $T_i = T_j$ ในส่วนของ**ความคู่ขนานแบบคะแนนจริงสมมูล**(Tau-Equivalent) มีการผ่อนปรนข้อตกลงของความคู่ขนานของแต่ละส่วนของแบบทดสอบลงในข้อที่ 2. ให้คะแนนจริงของแต่ละส่วนไม่จำเป็นต้องเท่ากันพอดี แต่ยอมให้ต่างกันได้เท่ากับความยากที่ต่างกันในแต่ละส่วน ได้โมเดลคะแนนจริงสมมูลคือ $T_i = T_j + b_{ij}$ เมื่อ $i \neq j=1,2,\ldots,K$ และ b_{ij} ไม่จำเป็นต้องเท่ากับศูนย์เสมอไป ข้อที่ 3.ผ่อนปรนให้แต่ละส่วนมีคะแนนเฉลี่ยต่างกัน ข้อที่ 4.ผ่อนปรนให้แต่ละส่วนมีค่าความแปรปรวนของคะแนนสอบต่างกันได้เล็กน้อย แต่ความแตกต่างเหล่านี้ต้องไม่โตมากนักเพื่อให้แต่ละส่วนยังคงความคู่ขนานไว้เพราะว่าแต่ละส่วนของแบบทดสอบมีขนาดความยาวเท่ากันหรือมีจำนวนข้อเท่ากัน แม้จะมีการผ่อนปรนเงื่อนไขข้อ 2 , 3 , 4 แล้ว แต่ในทางปฏิบัติอาจมีการแบ่งส่วนตามลักษณะของแบบทดสอบ ทำให้แต่ละส่วนมีความยาวหรือจำนวนข้อไม่เท่ากัน มีผลกระทบต่อเงื่อนไขข้อ 5 และ 6

นอกจากนั้น แม้จะมีจำนวนข้อที่เท่ากัน แต่เมื่อทดสอบกับกลุ่มตัวอย่าง ปรากฏว่า แต่ละส่วนมีการกระจายมากน้อยต่างกัน แสดงว่า ความยาวที่เป็นจริงของมัน (Functional Lengths) หรือ ความยาวที่เป็นผลมาจากการสอบของแต่ละส่วนมีขนาดไม่เท่ากัน นั้นคือ จำนวนข้อสอบที่เท่ากันยังไม่เพียงพอในการบ่งชี้ความยาวที่แท้จริง ซึ่งโดยปกติจะไม่ทราบจนกว่าจะนำไปทดสอบเสียก่อน

และอีกประการหนึ่ง การวัดบางอย่างอาจทำให้การกระจายของคะแนนไม่เท่ากัน เช่น คะแนนข้อสอบอัตนัยที่มีคะแนนเต็มแต่ละข้อไม่เท่ากัน, คณะกรรมการประเมินผลงานหรือประเมินพฤติกรรม ที่แต่ละคนให้ไม่เท่ากัน ทั้งหมดเหล่านี้ทำให้ไม่สอดคล้องกับเงื่อนไขของแบบคะแนนจริงสมมูล

ความคู่ขนานแบบคะแนนจริงสัมพันธ์(Congeneric) มีการผ่อนปรนข้อตกลงของความคู่ขนานลงมาตามลำดับเหลือเงื่อนไขเพียงข้อ 1.แบบทดสอบต้องมีเนื้อหาเป็นเอกพันธ์กันหรือวัดคุณลักษณะเดียวกัน เพียงข้อเดียว และในการผ่อนปรนเงื่อนไขข้อ 2 ให้แต่ละส่วนมีคะแนนจริงต่างกันเท่ากับค่าคงที่ที่มาสัมพันธ์สมบูรณ์แบบ หรือมีความสัมพันธ์เป็นเส้นตรง ได้โมเดลคะแนนจริงสัมพันธ์คือ $T_i = a_{ij}T_j + b_{ij}$ ค่าคงที่ a_{ij} และ b_{ij} จะแปรเปลี่ยนไปตามความสัมพันธ์ของแต่ละส่วนของแบบทดสอบ แต่ไม่ขึ้นกับการแปรเปลี่ยนของแต่ละบุคคล

การหาค่าความเชื่อมั่นของเครื่องมือวัดจากการแบ่งส่วนภายในฉบับมีสูตรต่าง ๆ (บุญเชิด ภิญโญอนันตพงษ์. 2547) สรุปได้ดังนี้

ระดับความคู่ขนาน	แบ่งเป็นสองส่วน(Split-Half)	แบ่งหลายส่วน (Multiparts)
มาตรฐานเดิม	สูตรสเปียร์แมน-บราวน์ เฉพาะสองส่วน($r_{(S-B)2}$)	สูตรสเปียร์แมน-บราวน์ ทั่วไป ($r_{(S-B)k}$)
คะแนนจริงสมมูล	สูตรรูลอน(r_{RU}) สูตรฟลานาแกน(r_{FL}) สูตรกัตต์แมน(r_{Gutt})	ให้คะแนนแบบ 2 ค่า (Dichotomous) สูตรคูเดอร์และริชาร์สัน(KR-20 และ KR- 21) สูตรกัตต์แมน(r_{Gutt})
		ให้คะแนนแบบหลายค่า (Polytomous) สูตรของครอนบัค(r_{α}) สูตรฮอยท์(r_{Hoyt})
คะแนนจริงสัมพันธ์	สูตรคริสทอฟ(r_K)เมื่อแบ่ง3ส่วน สูตรเฟลดต์(r_{FS})เมื่อแบ่ง 2 ส่วน	สูตรราชู(r_R) สูตรกิลเมอร์-เฟลดต์ (r_{F1} , r_{F2}) สูตรเลียว(r_{L1} , r_{L2}) สูตรเฟลดต์(r_{FK}) สูตรบุญเชิด(r_B)

การประมาณค่าความเชื่อมั่นโมเดลมาตรฐานเดิม (Classical Parallel Model)

การประมาณค่าความเชื่อมั่นโมเดลมาตรฐานเดิมสามารถประมาณค่าความเชื่อมั่นได้จากสูตรของสเปียร์แมน-บราวด์ หากแบ่งแบบทดสอบออกเป็นสองส่วนก็ใช้สูตร $r_{(S-B)2}$ หากแบ่งแบบทดสอบออกเป็นหลายส่วนก็ใช้สูตร $r_{(S-B)K}$

การประมาณค่าความเชื่อมั่นโมเดลคะแนนจริงสมมูล(Tau-Equivalent Model)

การประมาณค่าความเชื่อมั่นโมเดลคะแนนจริงสมมูลสามารถประมาณค่าความเชื่อมั่นเมื่อแบ่งแบบทดสอบออกเป็นสองส่วนได้จากสูตรของรูลอน เรียกว่าสัมประสิทธิ์ r_{RU} สูตรของฟลานาแกน เรียกว่าสัมประสิทธิ์ r_{FL} และสูตรของกัตต์แมน เรียกว่าสัมประสิทธิ์ r_{Gutt} หากแบ่งแบบทดสอบออกเป็นหลายส่วนและมีการให้คะแนนแบบ 2 ค่า สามารถประมาณค่าความเชื่อมั่นได้จากสูตรของคูเดอร์และริชาร์ดสัน เรียกว่าสัมประสิทธิ์ KR-20 ใช้เมื่อข้อสอบแต่ละข้อมีความยากไม่เท่ากัน และ KR-21 ใช้เมื่อข้อสอบแต่ละข้อมีความยากเท่ากัน และสูตรของกัตแมน เรียกว่าสัมประสิทธิ์ r_{Gutt} แต่ถ้าหากแบ่งแบบทดสอบออกเป็นหลายส่วนและมีการให้คะแนนแบบหลายค่า

(Polytomous) สามารถประมาณค่าความเชื่อมั่นได้จากสูตรของครอนบัค เรียกว่าสัมประสิทธิ์ r_{α} และสูตรของฮอยท์ เรียกว่าสัมประสิทธิ์ r_{Hoyt}

แต่เมื่อพิจารณาตามบริบทของโรงเรียนที่ยังมีการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่มีการให้คะแนนแบบ 2 ค่า คือ 0 กับ 1 และแต่ละข้อมีความยากไม่เท่ากัน มีความเป็นเอกพันธ์หรือวัดเนื้อหาเดียวกัน วิธีประมาณค่าความเชื่อมั่นที่นักวิจัยส่วนใหญ่นิยมใช้และเป็นที่รู้จักกันอย่างแพร่หลาย มีความเหมาะสมที่จะนำไปใช้ ก็คือสัมประสิทธิ์ KR-20 ของคูเดอร์-ริชาร์ดสัน โดยมีสูตรในการคำนวณคือ

$$r_{tt} = \frac{k}{k-1}\left[1 - \frac{\sum pq}{S_x^2}\right]$$

เมื่อ k แทน จำนวนข้อของเครื่องมือวัด

p แทน สัดส่วนของผู้ตอบถูกหรือความยากของแต่ละข้อ

q แทน สัดส่วนของผู้ตอบผิด ซึ่งเท่ากับ 1-p

S_X^2 แทน ความแปรปรวนของคะแนนรวมทั้งฉบับของเครื่องมือวัด

การประมาณค่าความเชื่อมั่นโมเดลคะแนนจริงสัมพันธ์(Congeneric Model)

การประมาณค่าความเชื่อมั่นโมเดลคะแนนจริงสัมพันธ์สามารถประมาณค่าความเชื่อมั่นได้จากการคำนวณค่าความแปรปรวน-ความแปรปรวนร่วม และจากโมเดลการวัด

1.การประมาณค่าความเชื่อมั่นโมเดลคะแนนจริงสัมพันธ์จากค่าความแปรปรวน-ความแปรปรวนร่วม

สูตรการประมาณค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์จากค่าความแปรปรวน-ความแปรปรวนร่วม สามารถจัดประเภทตามจำนวนการแบ่งแบบทดสอบออกเป็นส่วนๆ คือ

1.เมื่อมีการแบ่งแบบทดสอบออกเป็น 2 ส่วน ได้แก่ วิธีการของเฟลดต์(Feldt. 1975: 558-560) สัมประสิทธิ์ r_{FS}

2.เมื่อมีการแบ่งแบบทดสอบออกเป็น 3 ส่วน ได้แก่ วิธีการของคริสทอฟ (Kristof. 1974: 492 - 496) สัมประสิทธิ์ r_k

3.เมื่อมีการแบ่งแบบทดสอบออกเป็นหลายส่วน ได้แก่ วิธีการของราชู (Raju 1977: 550-561) สัมประสิทธิ์ β_k หรือ r_R วิธีการของกิลเมอร์และเฟลดต์ (Gilmer and Feldt. 1983: 99-111) สัมประสิทธิ์ r_{F1} และ r_{F2} วิธีการของเฟลดต์และเบรนนอน (Feldt and Brennan. 1989: 115) สัมประสิทธิ์ r_{FK} วิธีการของเลียว (Liou. 1989: 154-158) สัมประสิทธิ์ r_{L1} และ r_{L2} วิธีการของบุญเชิด (บุญเชิด ภิญโญอนันตพงษ์ (2540) สัมประสิทธิ์ r_B

โดยได้อธิบายสูตรการคำนวณในแต่ละประเภทได้ดังนี้

1.1.กรณีแบ่งแบบทดสอบออกเป็น 2 ส่วน

วิธีการของเฟลดต์(Feldt. 1975 : 558-560) สัมประสิทธิ์ r_{FS}

เฟลดต์ได้ทำการศึกษาความเชื่อมั่นของแบบทดสอบที่แบ่งสามส่วนที่ยาวไม่เท่ากันของคริสทอฟ (Kristof. 1974) ให้สามารถวิเคราะห์ด้วยการแบ่งส่วนไม่เท่ากัน โดยกำหนดข้อตกลงเกี่ยวกับความแปรปรวนคลาดเคลื่อนตามทฤษฎีมาตรฐานเดิม เพิ่มขึ้นอีกหนึ่งข้อ คือ ให้ $\sigma_{E_g}^2 = \lambda_g \sigma_E^2$ ซึ่งข้อตกลงที่เพิ่มขึ้นนี้ มิได้ทำให้เกิดข้อจำกัดที่สำคัญประการใดตามมา และจากการเพิ่มข้อตกลงดังกล่าวยังทำให้สัมประสิทธิ์นี้จัดอยู่ในประเภทคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม (Classical Congeneric) (บุญเชิด ภิญโญอนันตพงษ์. 2537:13)

การแสดงที่มาของสัมประสิทธิ์ r_{FS} ของ เฟลดต์ (Feldt. 1975 : 558-560) มีดังนี้

สมมติให้แบบทดสอบประกอบด้วย คะแนนรวม และคะแนนส่วนย่อยสองส่วน โดยคะแนนที่สอบได้ แทนด้วย X, X_1 และ X_2 ตามลำดับ คะแนนจริง แทนด้วย T,T_1และT_2 ตามลำดับ และคะแนนคลาดเคลื่อนแทนด้วย E, E_1 และ E_2 ตามลำดับ ให้ λ_1 และ λ_2 แทนค่าสัดส่วนคะแนนจริงของ T_1 และ T_2 กับ T ตามลำดับและไม่ทราบค่า แสดงเป็นสูตรต่าง ๆ ได้ดังต่อไปนี้

$$X = X_1 + X_2 \quad = T + E \tag{1}$$
$$T = T_1 + T_2 \quad = \lambda_1 T + \lambda_2 T \tag{2}$$
$$E = E_1 + E_2 \tag{3}$$
$$X_1 = \lambda_1 T + C + E_1 \quad ; \quad C = \text{ค่าคงที่} \tag{4}$$
$$X_2 = \lambda_2 T + C + E_2 \tag{5}$$

และ

$$1 = \lambda_1 + \lambda_2 \tag{6}$$

ความแปรปรวนของคะแนนสอบ และความแปรปรวนร่วมของคะแนนสอบสองส่วนจะได้

$$\sigma_{X_1}^2 = \lambda_1^2 \sigma_T^2 + \sigma_{E_1}^2 \tag{7}$$
$$\sigma_{X_2}^2 = \lambda_2^2 \sigma_T^2 + \sigma_{E_2}^2 \tag{8}$$

และ

$$\sigma_{X_1 X_2} = \lambda_1 \lambda_2 \sigma_T^2 \tag{9}$$

ในการประมาณค่าความเชื่อมั่นของแบบทดสอบทั้งสองฉบับ ต้องคำนวณค่า $\rho = \sigma_T^2 / \sigma_X^2$ อย่างไรก็ตามสมการดังกล่าวข้างต้น กับข้อกำหนดที่ว่า $\lambda_1 + \lambda_2 = 1$ จะมีความสัมพันธ์เพียงสี่สมการ แต่มีตัวไม่ทราบค่าห้าตัว ทำให้ไม่สามารถหาค่า σ_T^2 และ ρ ให้ได้เพียงค่าเดียว (Unique

Solution) การที่จะทำให้มีเพียงค่าเดียวได้นั้น ข้อมูลที่ได้ต้องแบ่งเป็นสามส่วนตามเทคนิคของคริสทอฟ (Kristof. 1974) อย่างไรก็ตามถ้าต้องการหาค่าให้ได้เพียงค่าเดียวแล้วต้องสมมติต่อไปอีกว่า สองส่วนย่อยที่มีความยาวสั้นกว่าทั้งฉบับ และมีขนาดไม่เท่ากันนั้น มีข้อตกลงดังนี้

$$\sigma_{E_1}^2 = \lambda_1\sigma_E^2 \qquad (10)$$

$$\sigma_{E_2}^2 = \lambda_2\sigma_E^2 \qquad (11)$$

ความสัมพันธ์ระหว่าง ความแปรปรวนของคะแนนคลาดเคลื่อน ของแบบทดสอบส่วนที่ขยายยาวออกกับส่วนที่ยุบสั้นลง เป็นไปตามทฤษฎีการวัดแบบมาตรฐานเดิม (Gulliksen. 1950:72) จากข้อตกลงดังกล่าวทำให้สามารถลดเซตของตัวไม่ทราบค่าเหลือเพียงสี่ตัว $(\lambda_1, \lambda_2, \sigma_T^2$ และ $\sigma_E^2)$ สมการดังกล่าวข้างต้นจึงลดลงได้โดย

นำสมการ (10) แทนค่าในสมการ (7) ได้

$$\sigma_{X_1}^2 = \lambda_1^2\sigma_T^2 + \lambda_1\sigma_E^2 \qquad (12)$$

นำสมการ (11) แทนค่าในสมการ (8) ได้

$$\sigma_{X_2}^2 = \lambda_2^2\sigma_T^2 + \lambda_2\sigma_E^2 \qquad (13)$$

และจากสมการ (9) ที่ว่า $\sigma_{X_1X_2} = \lambda_1\lambda_2\sigma_T^2$

สมการ (6) ที่ว่า $1 = \lambda_1 + \lambda_2$

เพื่อที่จะแก้สมการให้ σ_T^2 เป็นสมการของพารามิเตอร์ของคะแนนสอบได้ ให้นำสมการ (12) ลบด้วย (13) และอาศัยข้อเท็จจริงของสมการ (6) ที่ว่า $1 = \lambda_1 + \lambda_2$ แล้วจะได้ดังนี้

(12)-(13) $\sigma_{X_1}^2 - \sigma_{X_2}^2 = \lambda_1^2\sigma_T^2 + \lambda_1\sigma_E^2 - \lambda_2^2\sigma_T^2 - \lambda_2\sigma_E^2$

ดึงตัวร่วม ได้

$$\begin{aligned}\sigma_{X_1}^2 - \sigma_{X_2}^2 &= (\lambda_1^2 - \lambda_2^2)\sigma_T^2 + (\lambda_1 - \lambda_2)\sigma_E^2 \\ &= (\lambda_1 + \lambda_2)(\lambda_1 - \lambda_2)\sigma_T^2 + (\lambda_1 - \lambda_2)\sigma_E^2 \\ &= (\lambda_1 - \lambda_2)(\sigma_T^2 + \sigma_E^2) \\ &= (\lambda_1 - \lambda_2)\sigma_X^2\end{aligned}$$

หารทั้งสองข้างด้วย σ_X^2 ทำให้ได้

$$\frac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X^2} = (\lambda_1 - \lambda_2) \qquad (14)$$

จากความสัมพันธ์ของ สมการ (14) และ $\lambda_1 + \lambda_2 = 1$ จาก สมการ (6) สามารถแก้สมการพร้อมกันได้

จาก (6) $\lambda_2 = 1 - \lambda_1$

แทนค่า λ_2 ลงในสมการ (14) ได้

$$\frac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X^2} = \left(\lambda_1 (1 - \lambda_1\right)$$

$$\frac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X^2} = -1 + 2\lambda_1$$

$$1 + \frac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X^2} = 2\lambda_1$$

$$\frac{\sigma_X^2 + \left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)}{\sigma_X^2} = 2\lambda_1$$

$$\frac{\sigma_X^2 + \left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)}{2\sigma_X^2} = \lambda_1 \qquad (15)$$

ทำนองเดียวกันก็จะได้

$$\frac{\sigma_X^2 - \left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)}{2\sigma_X^2} = \lambda_2 \qquad (16)$$

ขั้นสุดท้ายนำค่าที่ได้จากสมการ (15) กับ (16) แทนลงสมการ (9) ที่ว่า $\sigma_{X_{-1}X_2} = \lambda_1\lambda_2\sigma_T^2$

ทำให้คำนวณหา $\lambda_1\lambda_2$ ได้ดังนี้

$$\lambda_1\lambda_2 = \left[\frac{\sigma_X^2 + \left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)}{2\sigma_X^2}\right]\left[\frac{\sigma_X^2 - \left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)}{2\sigma_X^2}\right]$$

$$= \frac{\sigma_X^4 + \sigma_X^2\left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right) - \sigma_X^2\left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right) - \left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)^2}{4\sigma_X^4}$$

$$= \frac{\sigma_X^4 - \left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)^2}{4\sigma_X^4}$$

จากสมการ (9) คือ $\sigma_{X_{-1}X_2} = \lambda_1\lambda_2\sigma_T^2$ หา σ_T^2 ได้ดังนี้

$$\sigma_T^2 = \frac{\sigma_{X_1}\sigma_{X_2}}{\lambda_1\lambda_2}$$

แทนค่า $\lambda_1\lambda_2$ เข้าไป

$$\sigma_T^2 = \frac{4\sigma_{X_1}\sigma_{X_2}\sigma_X^4}{\sigma_X^4 - \left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)^2}$$

$$= \frac{4\sigma_{X_1}\sigma_{X_2}}{1 - \frac{\left(\sigma_{X_1}^2 - \sigma_{X_2}^2\right)^2}{\sigma_X^4}}$$

$$= \frac{4\sigma_{X_1}\sigma_{X_2}}{1 - \left(\frac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X^2}\right)^2} \qquad (17)$$

จาก $\rho = \frac{\sigma_T^2}{\sigma_X^2}$ นำสมการ (17) แทนค่าเข้าไป ได้ดังนี้

$$\rho = \frac{4\sigma_{X_1}\sigma_{X_2}}{\sigma_X^2\left[1 - \left(\frac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X^2}\right)^2\right]}$$

$$\boxed{\rho = \frac{4\sigma_{X_1}\sigma_{X_2}}{\sigma_X^2 - \left(\frac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X}\right)^2} = r_{FS}} \qquad (18)$$

ถ้าไม่คำนึงถึงขนาดความยาวที่ไม่เท่ากันของส่วนย่อย และใช้สัมประสิทธิ์ แอลฟ่า กับข้อมูลส่วนย่อย กับทั้งฉบับ สัมประสิทธิ์ดังกล่าวจะเป็น $4\sigma_{X_1X_2} / \sigma_X^2$ จะเห็นได้ชัดเจนว่า สมการ (18) จะให้ค่าที่สูงกว่าแอลฟ่า เมื่อความแปรปรวนของส่วนย่อยไม่เท่ากัน

1.2.กรณีแบ่งแบบทดสอบออกเป็น 3 ส่วน

วิธีการของคริสทอฟ (Kristof. 1974: 492 - 496) สัมประสิทธิ์ r_k

คริสทอฟเป็นบุคคลแรกที่ได้คิดเทคนิคการประมาณค่าความเชื่อมั่นของแบบทดสอบที่วัดเนื้อหาที่เป็นเอกพันธ์กัน โดยแบ่งเป็นสามส่วนที่มีความยาวไม่เท่ากัน และคะแนนจริงของแต่ละส่วนต้องมีสหสัมพันธ์เป็นเส้นตรง โดยไม่ต้องใช้จำนวนข้อของแต่ละส่วนมาคำนวณ
(บุญเชิด ภิญโญอนันตพงษ์. 2537:12)

การแสดงที่มาของสัมประสิทธิ์ r_k ของคริสทอฟ (Kristof. 1974: 492 - 496) มีดังนี้

สมมติให้แบบทดสอบฉบับหนึ่งแบ่งเป็นสามส่วนโดยมีเนื้อหาเป็นเอกพันธ์กัน มีคะแนนรวมทั้งฉบับและคะแนนส่วนย่อยสามส่วนดังนี้

คะแนนที่สอบได้ แทนด้วย X, X_1, X_2 และ X_3

คะแนนจริง แทนด้วย T, T_1, T_2 และ T_3

คะแนนคลาดเคลื่อน แทนด้วย E, E_1, E_2 และ E_3 ตามลำดับ

$$X = X_1 + X_2 + X_3 \tag{19}$$

$$T = T_1 + T_2 + T_3 \tag{20}$$

$$E = E_1 + E_2 + E_3 \tag{21}$$

$$X_j = T_j + E_j \quad ; \quad j = 1,2,3, \tag{22}$$

ให้ σ_{jk} แทน ความแปรปรวนร่วมระหว่างตัวแปร X_j และ X_k

σ_X^2 แทน ความแปรปรวนของคะแนนสอบทั้งฉบับ

σ_T^2 แทน ความแปรปรวนของคะแนนจริงทั้งฉบับ

σ_E^2 แทน ความแปรปรวนคะแนนคลาดเคลื่อนทั้งฉบับ

ρ แทน ความเชื่อมั่นของแบบทดสอบ

λ_j และ μ_j เป็นจำนวนที่ไม่เป็นลบ ตามแบบจำลองทฤษฎีทดสอบมาตรฐานเดิม ความแปรปรวนของคะแนนจริง และความแปรปรวนของคะแนนคลาดเคลื่อนของส่วนย่อย j สามารถแทนด้วย $\lambda_j^2\sigma_T^2$ และ $\mu_j\sigma_E^2$ ตามลำดับ สามารถนำมาสร้างเป็นสมการต่าง ๆ ได้ดังนี้

$$\sigma_{jj} = \lambda_j^2\sigma_j^2 + \mu_j\sigma_E^2 \quad ; \quad j = 1,2,3, \tag{23}$$

$$\sigma_{jk} = \lambda_j\lambda_k\sigma_T^2 \quad ; \quad j,k = 1,2,3; \quad j \neq k \tag{24}$$

$$1 = \lambda_1 + \lambda_2 + \lambda_3 \quad ; \quad \lambda_1, \lambda_2, \lambda_3 \geq 0 \tag{25}$$

$$1 = \mu_1 + \mu_2 + \mu_3 \quad ; \quad \mu_1, \mu_2, \mu_3 \geq 0 \tag{26}$$

จากข้างต้นมีสมการ 8 สมการ(สมการที่ (19) ถึง สมการที่ (26) และมีตัวไม่ทราบค่า 8 ตัว และทราบค่าของ σ_{jk} เมื่อ j , k = 1, 2, 3 ในกรณีนี้เมื่อกลุ่มตัวอย่างมีขนาดโตพอแล้วสามารถใช้ค่าสถิติ แทนค่าพารามิเตอร์ซึ่งทำให้การแก้สมการต่อไปนี้ได้ค่าคำตอบค่าเดี่ยว (Unique Solution)

$$\sigma_T^2 = \frac{\sigma_{12}\sigma_{13}}{\sigma_{23}} + \frac{\sigma_{12}\sigma_{23}}{\sigma_{12}} + \frac{\sigma_{13}\sigma_{23}}{\sigma_{12}} + 2(\sigma_{12} + \sigma_{13} + \sigma_{23}) \tag{27}$$

$$\sigma_E^2 = \sigma_X^2 - \sigma_T^2 \tag{28}$$

$$\lambda_j = \frac{\sqrt{\sigma_{jk}\sigma_{jl}}}{\sigma_T\sqrt{\sigma_{kl}}} \quad ; \quad j,k,l = 1,2,3 \text{ และแตกต่างกันทั้งหมด} \tag{29}$$

$$\mu_j = \frac{\sigma_{jj} - \lambda_j^2 \sigma_T^2}{\sigma_E^2} \qquad ; \qquad j = 1,2,3 \qquad (30)$$

ขนาดของ λ_j และ μ_j จะแสดงคุณลักษณะของส่วนย่อย j ซึ่งไม่เปลี่ยนแปลงไปตามเงื่อนไขใดนอกจากผลรวมของมันต้องเท่ากับหนึ่ง ดังสมการ (25) และ (26) ความแปรปรวนของคะแนนจริง σ_T^2 ในสมการ (27) จะไม่ขึ้นอยู่กับการแบ่งแบบทดสอบออกเป็นสามส่วนที่มีเนื้อหาเป็นเอกพันธ์กัน (ส่วนย่อยเหล่านี้ถ้าคะแนนจริงมีความสัมพันธ์กันเป็นเส้นตรงแล้ว เรียกว่า มีเนื้อหาเป็นเอกพันธ์กัน หรือเรียกว่าคะแนนจริงสัมพันธ์ โดยไม่ได้บังคับว่าข้อสอบภายในส่วนนั้นต้องมีความเป็นเอกพันธ์ทางเนื้อหาหรือเนื้อหาเดียวกัน) ดังนั้นไม่ว่าจะแบ่งทดสอบเป็นสามส่วนใด ๆ ก็ตาม ก็ถือว่าใช้ได้ทั้งนั้น และจะให้ค่า σ_T^2 เท่ากันเสมอ และจำนวนดังกล่าวนี้จะไม่ใช้ค่าขีดจำกัดล่าง และถ้าแบ่งแบบทดสอบเป็นสามส่วนให้มีเนื้อหาเป็นเอกพันธ์กันแล้ว ค่าด้านขวาของสมการ (27) จะไม่แปรเปลี่ยน (Invariant) ไปตามการแบ่งแบบทดสอบเป็นสามส่วนแบบใด ๆ ซึ่งถือว่าเป็นผลสำคัญที่ได้มาของทฤษฎี การแสดงที่มาดังกล่าวนี้คล้ายคลึงกับการพัฒนาการประมาณค่าความเชื่อมั่นจากแบบทดสอบที่มีฟอร์มไม่คู่ขนานกันของ ลอร์ด และ โนวิก (Lord and Novick. 1968 : 216)

จะเห็นได้ชัดเจนว่า หากประมาณค่าความแปรปรวนของคะแนนคลาดเคลื่อน σ_E^2 ตามสมการ (10) กับความเชื่อมั่น ρ หรือ r_k ที่คำนวณจากสูตร

$$\boxed{\rho = \frac{\sigma_T^2}{\sigma_X^2} = r_k = \frac{(\sigma_{12}\sigma_{13} + \sigma_{12}\sigma_{23} + \sigma_{13}\sigma_{23})^2}{(\sigma_{12}\sigma_{13}\sigma_{23})\sigma_x^2}} \qquad (31)$$

มีค่าที่ไม่ขึ้นอยู่กับการแบ่งแบบทดสอบเป็นสามส่วนแบบใด ๆ สำหรับค่าของ σ_T^2 ในสมากร (31) หาได้จากสมการ (27)

ถ้ากลุ่มตัวอย่างมีขนาดไม่โตพอ และแทนค่าพารามิเตอร์ด้วยสถิติตัวอย่างธรรมดาแล้ว ค่า σ_T^2, σ_E^2 และ ρ ในสมการ (27), (28) และ (31) จะให้ค่าพารามิเตอร์ที่สอดคล้องกัน และในขณะเดียวกัน ค่าเหล่านี้จะเป็นพารามิเตอร์แบบความน่าจะเป็นสูงสุด (Maximum Likelihood) เมื่อค่าโมเม้นต์ที่สองของคะแนนเป็นไปตามการแจกแจงของ วิสฮาร์ด (Wishart Distribution) และมีข้อสังเกตว่า การทดสอบทางสถิติเกี่ยวกับแบบจำลอง ตามสมการ (23) ถึง (26) จะเป็นไปไม่ได้ที่โมเม้นต์ที่สองของคะแนนตามสมการดังกล่าวจะให้ค่าคำตอบที่แน่นอนเสมอไป

ถ้ากลุ่มตัวอย่างมีขนาดโตมาก จนกระทั่งสามารถใช้ค่าพารามิเตอร์ได้แล้ว การศึกษาว่า σ_T^2 และ ρ ที่คำนวณจากสมการ (27) และ (31) จะมีค่ามากกว่าพารามิเตอร์แท้จริง ถ้าไม่สามารถทำให้ส่วนย่อยของแบบทดสอบที่แบ่งมีเนื้อหาเป็นเอกพันธ์กัน และมีความเพียงพอที่จะพิจารณาว่า σ_T^2 ของ ρ เป็นฟังก์ชันเส้นตรงอย่างง่าย ถ้าใช้แบบจำลองทฤษฎีการทดสอบแบบมาตรฐานเดิมจากสมการที่ (20) สามารถเขียนสมการใหม่ในรูปแบบของความแปรปรวนและความแปรปรวนร่วม ดังนั้น

$$\sigma_T^2 = \sigma_{T_1}^2 + \sigma_{T_2}^2 + \sigma_{T_3}^2 + 2(\sigma_{12} + \sigma_{13} + \sigma_{23}) \quad (32)$$

นำด้านขวาของสมการ (32) ลบด้วย ด้านขวาของสมการ (27) จะได้ผลต่าง Δ ดังนี้

$$\Delta = \sigma_{T_1}^2 + \sigma_{T_2}^2 + \sigma_{T_3}^2 - \frac{\sigma_{12}\sigma_{13}}{\sigma_{23}} - \frac{\sigma_{12}\sigma_{23}}{\sigma_{13}} - \frac{\sigma_{12}\sigma_{23}}{\sigma_{12}}$$

$$= \sigma_{T_1}^2\left(1 - \frac{\rho_{12}\rho_{13}}{\rho_{23}}\right) + \sigma_{T_2}^2\left(1 - \frac{\rho_{12}\rho_{23}}{\rho_{13}}\right) + \sigma_{T_3}^2\left(1 - \frac{\rho_{13}\rho_{23}}{\rho_{12}}\right) \quad (33)$$

เนื่องด้วย ρ_{jk} เป็นตัวบ่งชี้สหสัมพันธ์ระหว่างคะแนนจริง T_j กับ T_k จึงมีเหตุผลเพียงพอที่จะสมมุติว่า สหสัมพันธ์ ρ_{jk} ทั้งสามมีค่าเป็นบวก แต่ส่วนมากแล้วหนึ่งสหสัมพันธ์บางส่วน(Partial Correlations) ระหว่างคะแนนส่วนย่อยสองส่วนโดยควบคุมคะแนนส่วนที่สามารถมีค่าเป็นลบได้ ดังนั้นผลต่างในวงเล็บของสมการที่ (33) อย่างน้อยหนึ่งค่าสามารถมีค่าเป็นลบ และในกรณีเช่นนี้องค์ประกอบ σ_T^2 ควรมีค่าโตเพียงพอ เพื่อที่จะมีผลให้ $\Delta < 0$ ถ้าสัมประสิทธิ์ ρ_{jk} มีค่าไม่แตกต่างกันมากนักแล้ว จะเป็นไปไม่ได้ที่ค่า $\Delta < 0$ ดังนั้นโดยธรรมชาติแล้วจึงควรกำหนดข้อตกลงว่า $\Delta \geq 0$ หรือกล่าวอีกอย่างหนึ่งได้ว่าสูตรสมการ (27) และ (31) จะให้ค่าประมาณที่ต่ำกว่าพารามิเตอร์ σ_T^2 และ ρ เมื่อส่วนย่อยทั้งสามของแบบทดสอบไม่มีเนื้อหาเป็นเอกพันธ์กัน อย่างไรก็ตามการประมาณค่าที่เกินกว่าพารามิเตอร์นั้น แทบจะไม่เกิดขึ้นในทางปฏิบัติ ดังนั้นจึงไม่กำหนดให้เป็นข้อตกลง

ในการเปรียบเทียบ σ_T^2 ตามสมการ (27) กับ σ_T^2 ที่ประมาณค่าจากสัมประสิทธิ์แอลฟา (coefficient α) คือ $\tilde{\sigma}_T^2 = \alpha\sigma_X^2$ เมื่อใช้ข้อมูลเดียวกัน

$$\tilde{\sigma}_T^2 = 3(\sigma_{12} + \sigma_{13} + \sigma_{23}) \quad (34)$$

และนำสมการ (27) ลบด้วย สมการ (34) ได้ดังนี้

$$\sigma_T^2 - \tilde{\sigma}_T^2 = \frac{\sigma_{12}\sigma_{13}}{\sigma_{23}} + \frac{\sigma_{12}\sigma_{23}}{\sigma_{13}} + \frac{\sigma_{13}\sigma_{23}}{\sigma_{12}} - \sigma_{12} - \sigma_{13} - \sigma_{23}$$

$$= \frac{1}{6}\left[\left(\sqrt{\frac{\sigma_{12}\sigma_{13}}{\sigma_{23}}} + \sqrt{\frac{\sigma_{12}\sigma_{23}}{\sigma_{13}}} - 2\sqrt{\frac{\sigma_{13}\sigma_{23}}{\sigma_{12}}}\right)^2 + \left(\sqrt{\frac{\sigma_{12}\sigma_{13}}{\sigma_{23}}} - 2\sqrt{\frac{\sigma_{12}\sigma_{23}}{\sigma_{13}}} + \sqrt{\frac{\sigma_{13}\sigma_{23}}{\sigma_{12}}}\right)^2 + \left(-2\sqrt{\frac{\sigma_{12}\sigma_{13}}{\sigma_{23}}} + \sqrt{\frac{\sigma_{12}\sigma_{23}}{\sigma_{13}}} + \sqrt{\frac{\sigma_{13}\sigma_{23}}{\sigma_{12}}}\right)^2\right] \geq 0 \quad (35)$$

ซึ่ง $\sigma_T^2 - \tilde{\sigma}_T^2 = 0$ ก็ต่อเมื่อผลรวมของนิพจณ์ ในแต่ละวงเล็บมีค่าเป็นศูนย์ อันเนื่องมาจาก $\sigma_{12} = \sigma_{13} = \sigma_{23}$ สำหรับกรณีอื่น ๆ แล้ว สูตรสมการ(27) จะให้ค่าประมาณที่ดีกว่า

$\tilde{\sigma}_T^2$ และจากการเทียบกันแล้วจะนิยมใช้ สมการ (31) มากกว่าใช้ สัมประสิทธิ์แอลฟ่า(coefficient α) เมื่อ σ_T^2 ในสมการ (31) คำนวณจากสมการ(27)

เมื่อเปรียบเทียบ σ_T^2 ที่ได้จาก สมการ (27) กับที่ได้จาก ค่าขีดจำกัดล่างของกัตต์แมน (Guttman. 1945) r_G หรือ λ_3 โดย $\hat{\sigma}_T^2 = \lambda_3 \sigma_X^2$ จะได้

$$\hat{\sigma}_T^2 = 2(\sigma_{12} + \sigma_{13} + \sigma_{23}) + \sqrt{3(\sigma_{12}^2 + \sigma_{13}^2 + \sigma_{23}^2)} \qquad (36)$$

จากข้อเท็จจริงใน สมการ (35) ที่ว่าจำนวน $\sigma_{12} = \sigma_{13} = \sigma_{23}$ ใด ๆ มีค่าเป็นบวก และกำลังสองของมันก็เป็นบวกด้วย ดังนั้น

$$\begin{aligned} &\frac{\sigma_{12}^2\sigma_{13}^2}{\sigma_{23}^2} + \frac{\sigma_{12}^2\sigma_{23}^2}{\sigma_{13}^2} + \frac{\sigma_{13}^2\sigma_{23}^2}{\sigma_{12}^2} - \sigma_{12}^2 - \sigma_{13}^2 - \sigma_{23}^2 \\ &= \left(\frac{\sigma_{12}\sigma_{13}}{\sigma_{23}} + \frac{\sigma_{12}\sigma_{23}}{\sigma_{13}} + \frac{\sigma_{13}\sigma_{23}}{\sigma_{12}} \right)^2 - 3(\sigma_{12}^2 + \sigma_{13}^2 + \sigma_{23}^2) \geq 0 \end{aligned} \qquad (37)$$

หมายความว่า

$$\sigma_T^2 - \hat{\sigma}_T^2 = \frac{\sigma_{12}\sigma_{13}}{\sigma_{23}} + \frac{\sigma_{12}\sigma_{23}}{\sigma_{13}} + \frac{\sigma_{13}\sigma_{23}}{\sigma_{12}} - \sqrt{3(\sigma_{12}^2 + \sigma_{13}^2 + \sigma_{23}^2)} \geq 0 \qquad (38)$$

ซึ่ง $\sigma_T^2 - \hat{\sigma}_T^2 = 0$ ก็ต่อเมื่อ $\sigma_{12} = \sigma_{13} = \sigma_{23}$ ดังอธิบายใต้สมการ (35) ในกรณีอื่น ๆ สูตรสมการ (27) จะให้ค่าประมาณที่ดีกว่า $\hat{\sigma}_T^2$ จากการเปรียบเทียบมีผู้นิยม สมการ (31) มากกว่า λ_3 เมื่อ σ_T^2 ในสมการ (31) ได้มาจากสมการ (27)

ดังนั้นสามารถลงสรุปได้ว่า สมการ (31) ที่ได้จาก สมการ (27) จะประมาณค่าพารามิเตอร์ความเชื่อมั่นได้ดีกว่า ทั้งสัมประสิทธิ์แอลฟ่าของครอนบัด (Cronbach. 1951) และสัมประสิทธิ์ของกัตต์แมน (Guttman. 1945) เสมอ

อย่างไรก็ตาม แม้ว่าจะยังไม่มีการสร้างทฤษฎีตัวอย่างสุ่ม ที่จะนำตัวสถิติใช้แทน ตัวพารามิเตอร์ แต่ในช่วงเวลานั้น ทั้งสัมประสิทธิ์ แลมด้า ของกัตต์แมน ก็ไม่มีเช่นเดียวกัน ส่วนสัมประสิทธิ์ แอลฟ่า ก็ไม่มีทฤษฎีการแจกแจงตัวอย่างสุ่มสำหรับตัวอย่างขนาดเล็ก ยกเว้นแต่การแบ่งส่วนย่อยของแบบทดสอบให้มีความคู่ขนานแบบคะแนนจริงสมมูล และการวัดมีการแจกแจงแบบอิสระเท่านั้น

1.3.กรณีแบ่งแบบทดสอบออกเป็น k ส่วน

วิธีการของราชู (Raju 1977:550-561) สัมประสิทธิ์ β_k หรือ r_R

ราชูได้พิจารณาเห็นว่า สัมประสิทธิ์แอลฟ่า ของครอนบัค แม้จะสามารถใช้วิเคราะห์ความเชื่อมั่นของแบบทดสอบที่แบ่งออกเป็นหลาย ๆส่วนได้ก็ตาม แต่ถ้าส่วนย่อย ๆเหล่านั้นมีขนาดความยาวไม่เท่ากันแล้ว สัมประสิทธิ์แอลฟ่าจะมีปัญหาทำให้ประมาณค่าต่ำความเป็นจริง ดังนั้นราชู จึงปรับแก้สัมประสิทธิ์แอลฟ่า ให้เป็นสัมประสิทธิ์ทั่วไปที่ใช้กับแบบทดสอบที่แบ่งส่วนย่อยหลายส่วนที่ยาวไม่เท่ากัน เรียก สัมประสิทธิ์เบต้า (Coefficient β) ถ้าส่วนย่อยที่มีขนาดความยาวหรือจำนวนข้อเท่ากัน สัมประสิทธิเบต้าจะให้ค่าเท่ากันกับสัมประสิทธิ์แอลฟ่า เทคนิคนี้สามารถคำนวณจากแบบทดสอบที่แบ่งเป็นกี่ส่วนก็ได้ แต่ที่สำคัญต้องทราบจำนวนข้อของแต่ละส่วนย่อย
(บุญเชิด ภิญโญอนันตพงษ์. 2537:13)

การแสดงที่มาของสัมประสิทธิ์ r_R ของราชู (Raju 1977:550-561) ได้ดังนี้

สมมติให้ S เป็นเซตของข้อสอบ n ข้อ และถูกแบ่งออกเป็นส่วน ๆ อิสระจากกัน k ส่วน แต่ละส่วนแทนด้วน S_1, S_2, ….., S_k และแต่ละส่วนมีข้อสอบจำนวน n_1, n_2, ….., n_k ข้อ ตามลำดับ ให้ $p_i = n_i / n$ แทนสัดส่วนจำนวนข้อ ใน S_i และให้ X_1, X_2, ….., X_k แทนคะแนนรวมจากข้อสอบจำนวน n_1, n_2, ….., n_k ข้อตามลำดับ จากคำนิยามข้างต้นสิ่งที่ตามมาก็คือ $\sum_1^k S_i = S$, $\sum_1^k n_i = n$, $\sum_1^k X_i = X$ และ $\sum_1^k p_i = 1$ และสามารถนิยามสัมประสิทธิ์เบต้า ของแบบทดสอบที่แบ่งเป็น K ส่วนได้ดังนี้

$$\beta_k = \frac{\sum\sum_{i \neq j} \sigma_{x_1 x_2}}{\sigma_x^2 \sum\sum_{i \neq j} p_i p_j} \qquad (39)$$

เมื่อ $\sigma_{x_i x_j}$ เป็นความแปรปรวนร่วมระหว่างคะแนนส่วนย่อย X_i กับ X_j และ σ_x^2 เป็นความแปรปรวนของ X และสูตรสมการ (39) เขียนใหม่ได้ดังนี้

$$\beta_k = \frac{\sigma_x^2 - \sum_1^k \sigma_{x_1}^2}{\sigma_x^2 \left(1 - \sum_1^k p_i^2\right)} \qquad (40)$$

เมื่อ $\sigma_{x_i}^2$ เป็นความแปรปรวนของแบบทดสอบย่อย X_i

จากความรู้เดิมเป็นที่รู้จักกันทั่วไปตามที่ โนวิก และลีวิส (Novick and Lewis. 1968) ได้แสดงให้เห็นว่า สัมประสิทธิ์แอลฟ่า r_α ของครอนบัค (Cronbach. 1951) เป็นค่าขีดจำกัดล่างของความเชื่อมั่นของแบบทดสอบ และเป็นค่าเดียวกันกับความเชื่อมั่นได้พอดี เมื่อคะแนนจริงจาก

ส่วนย่อยของแบบทดสอบเป็นคะแนนจริงสมมูล ($\tau - equivalent$) ทำนองเดียวกัน β_k หรือ r_R มีค่าเดียวกันกับความเชื่อมั่น เมื่อเป็นไปตามเงื่อนไขดังกล่าว

ในการศึกษาสัมประสิทธิ์ β_k เริ่มต้นจาก กำหนดให้ u และ v เป็นข้อสอบสองข้อที่มีข้อตกลงแบบคะแนนจริงสมมูล เมื่อ $\tau_u = \tau_v + a_{uv}$ เมื่อ τ_u และ τ_v เป็นคะแนนจริงจากคะแนนสอบ u และ v และ a_{uv} เป็นค่าคงที่ขึ้นอยู่กับ u และ v เท่านั้น ซึ่งหมายความว่า คะแนนจริงของ u และ v ต่างกันเท่ากับค่าคงที่ และค่าคงที่นี้ขึ้นอยู่เฉพาะ u และ v เท่านั้น ให้ X = T + E และ $X_i = T_i + E_i$ เมื่อ T และ T_i เป็นคะแนนจริง E และ E_i เป็น คะแนนคลาดเคลื่อน จาก X และ X_i ตามลำดับ และจากทฤษฎีการวัดทางจิตวิทยาแบบมาตรฐานเดิมความเชื่อมั่นของ X นิยามไว้ดังนี้

$$r_{xx} = \frac{\sigma_T^2}{\sigma_x^2}$$

เมื่อ σ_T^2 เป็นความแปรปรวนของคะแนนจริงจาก X

หลักการ

ให้ X เป็นคะแนนของแบบทดสอบฉบับหนึ่ง ประกอบด้วยข้อสอบที่มีคะแนนจริงสมมูลจำนวน n ข้อ และแบ่งคะแนน X เป็น K ส่วนย่อยอิสระจากกัน เป็น $X_1, X_2, \ldots, X_k$ และมี n_i แทนจำนวนข้อของ X_i แล้ว ความเชื่อมั่นของ X จะเท่ากับ β_k นั่นคือ

$$\beta_k = \frac{\sigma_T^2}{\sigma_x^2}$$

พิสูจน์ได้ เพราะว่า $X = X_1 + X_2 + \ldots + X_k$; และ $X_i = T_i + E_i$

$$\sigma_T^2 = \sum_1^k \sigma_{T_i}^2 + \sum_{i \neq j}\sum \sigma_{T_i T_j}$$

ตามข้อตกลงที่ว่า คะแนนคลาดเคลื่อนเป็นอิสระกัน จะได้ว่า $\sigma_{T_i T_j} = \sigma_{x_i x_j}$ ดังนั้น

$$\sigma_T^2 = \sum_1^k \sigma_{T_i}^2 + \sum_{i \neq j}\sum \sigma_{x_i x_j} \qquad (41)$$

เนื่องจากข้อสอบเป็นแบบคะแนนจริงสมมูล ดังนั้นข้อสอบแต่ละข้อ จะมีความแปรปรวนของคะแนนจริงเท่ากัน และความแปรปรวนร่วมระหว่างข้อสอบสองข้อใด ๆ เท่ากับ ความแปรปรวนร่วม

ระหว่างข้อสอบสองข้ออื่นๆ นอกจากนั้นยังสามารถแสดงให้เห็นว่าความแปรปรวนของคะแนนจริงของข้อสอบสมมติ (Arbitary Item) เท่ากับ ความแปรปรวนร่วมของข้อสอบสมมติสองข้อ ซึ่งแสดงเป็นค่าพีชคณิตได้ดังนี้ $\sigma^2_{T_u} = \sigma_{uv} = \sigma^2_{T_v}$ เมื่อ τ_u และ τ_v เป็นคะแนนจริงจากข้อสอบ u และ v และให้ค่าที่ร่วมกันแทนด้วย g ดังนั้นความแปรปรวนของคะแนนจริง T_i เขียนได้ดังนี้

$$\begin{aligned}\sigma^2_{T_i} &= \sum_1^{n_i} \sigma^2_{T_u} + \sum_{u \neq v; u,v \in S_i}\sum \sigma_{uv} \\ &= n_i g + n_i(n_i - 1)g \\ &= n_i^2 g\end{aligned}$$

จากผลดังกล่าวนี้ สมการ (41) สามารถเขียนใหม่ได้ดังนี้

$$\sigma^2_T = g\sum_1^k n_i^2 + \sum_{i \neq j}\sum \sigma_{x_i x_j} \qquad (42)$$

เพื่อที่จะประเมินค่าคงที่ g ในเทอมของความแปรปรวนร่วมระหว่างแบบทดสอบย่อยที่ทราบค่า จึงดำเนินการต่อไปดังนี้ ความแปรปรวนร่วมแต่ละเทอม $\sigma_{x_i x_j}$ ประกอบด้วยความแปรปรวนร่วมระหว่างข้อสอบ $n_i n_j$ เนื่องจากข้อสอบแต่ละคู่เป็นแบบคะแนนจริงสมมูล ความแปรปรวนร่วมระหว่างข้อสอบ $n_i n_j$ ต้องมีค่าเท่ากับ g และดังนั้น

$$\begin{aligned}\sigma_{x_i x_j} &= g n_i n_j \\ \sum_{i \neq j}\sum \sigma_{x_i x_j} &= g\sum_{i \neq j}\sum n_i n_j \\ g &= \frac{\sum_{i \neq j}\sum \sigma_{x_i x_j}}{\sum_{i \neq j}\sum \sigma_{n_i n_j}}\end{aligned}$$

แทนค่า g ลงในสมการ (42) จะได้

$$\sigma^2_T = \frac{\left(\sum_1^k n_i^2 + \sum_{i \neq j}\sum n_i n_j\right)\sum_{i \neq j}\sum \sigma_{x_i x_j}}{\sum_{i \neq j}\sum n_i n_j}$$

$$= \frac{n^2 \sum\sum_{i \neq j} \sigma_{x_i x_j}}{\sum\sum_{i \neq j} n_i n_j}$$

เนื่องจาก $\frac{n_i n_j}{n^2} = p_i p_j$

สุดท้ายจะได้

$$\sigma_T^2 = \frac{\sum\sum_{i \neq j} \sigma_{x_i x_j}}{\sum\sum_{i \neq j} p_i p_j}$$

$$r_{xx} = \frac{\sigma_T^2}{\sigma_x^2} = \frac{\sum\sum_{i \neq j} \sigma_{x_i x_j}}{\sigma_x^2 \sum\sum_{i \neq j} p_i p_j} = \beta_k$$

เนื่องจาก $\sum\sum_{i \neq j} p_i p_j = 1 - \sum_1^k p_i^2$ และ $\sum\sum_{i \neq j} \sigma_{x_i x_j} = \sigma_x^2 - \sum_1^k \sigma_{x_i}^2$

ดังนั้นสามารถเขียนสัมประสิทธิ์เบต้าได้ดังนี้

$$\boxed{\beta_k = \frac{\sigma_x^2 - \sum_1^k \sigma_{x_1}^2}{\sigma_x^2 \left(1 - \sum_1^k p_i^2\right)}}$$

วิธีการของกิลเมอร์และเฟลดต์ (Gilmer & Feldt. 1983 : 101-105) สัมประสิทธิ์ r_{F1} และ r_{F2}

กิลเมอร์ และเฟลดต์ (Gilmer & Feldt. 1983 : 101-105) ได้เสนอสัมประสิทธิ์ความเชื่อมั่นสองตัวคือ สัมประสิทธิ์ r_{F1} และ r_{F2}พร้อมทั้งแสดงที่มาดังนี้

แบบจำลองคะแนนจริงสัมพันธ์ สำหรับแบบทดสอบที่ประกอบด้วย K ส่วน สามารถแสดงเป็นสมการต่าง ๆ ได้ดังนี้

$$X = X_1 + X_2 + X_3 + \cdots + X_K$$

$$T = T_1 + T_2 + T_3 + \cdots + T_K$$

$$E = E_1 + E_2 + E_3 + \cdots + E_K$$

$$X = T + E$$
$$X_j = T_j + E_j = \lambda_j T + b_j + E_j \qquad (j = 1,...,K)$$
$$\sum \lambda_j = 1.0 \qquad \lambda_j > 0$$
$$\sum b_j = 0$$

เมื่อ X, T และ E แทนคะแนนสอบ คะแนนจริง และคะแนนคลาดเคลื่อนของแบบทดสอบทั้งฉบับตามลำดับ และให้ X_j, T_j และ E_j (j = 1, …., K) แทนคะแนนสอบ คะแนนจริงและ คะแนนความคลาดเคลื่อนของส่วนย่อยของแบบทดสอบ (Part - Test) ตามลำดับ ส่วน λ_j เป็นค่าคงที่แทนสัดส่วนของคะแนนจริงทั้งฉบับที่เป็นผลมาจากส่วนย่อย หรือเป็น “ค่าสัดส่วนความยาวที่เป็นผลมาจากการสอบ” ของแบบทดสอบที่แบ่งเป็นส่วนย่อย และ b_j เป็นค่าคงที่ ส่วนคะแนนความคลาดเคลื่อน E_j สมมติให้ค่าคาดหวังมีค่าเป็นศูนย์ มีความอิสระจากกันเด็ดขาดและไม่ขึ้นอยู่กับคะแนนจริง และคะแนนจริงของแต่ละส่วนมีความสัมพันธ์เป็นเส้นตรง หมายความว่า คะแนนจริงของแต่ละส่วนมีสหสัมพันธ์สมบูรณ์แบบกับคะแนนจริงรวมทั้งฉบับ และนิยามส่วนต่อไปได้ดังนี้

$\sigma_{ij} = Cov(X_i, X_j)$ = ความแปรปรวนร่วมของคะแนนสอบระหว่างแต่ละส่วน

$\sigma_j^2 = Var(X_j)$ = ความแปรปรวนของคะแนนสอบแต่ละส่วน

σ_T^2 = ความแปรปรวนของคะแนนจริงทั้งฉบับ

σ_X^2 = ความแปรปรวนของคะแนนสอบทั้งฉบับ

กำหนดให้ E_i, E_j และT_j มีความอิสระซึ่งกันและกัน ดังนั้นความแปรปรวนและความแปรปรวนร่วมของคะแนนสอบเขียนได้ดังนี้

$$\sigma_j^2 = \lambda_j^2 \sigma_T^2 + \sigma_{E_j}^2 \qquad (j = 1,...,K)$$
$$\sigma_{ij} = \lambda_i \lambda_j \sigma_T^2 \qquad (i \neq j)$$

สัมประสิทธิ์ r_{F1}

เริ่มต้นจากการพิจารณาผลรวมของความแปรปรวนร่วมในแต่ละแถวของเมตริกซ์ความแปรปรวนของส่วนย่อยของแบบทดสอบ สำหรับแถว j ผลรวมของความแปรปรวนร่วมแทนได้ดังนี้

$$\sum_{i \neq j} \sigma_{ij} = \lambda_j \sum_{i \neq j} \lambda_i \sigma_T^2 = \lambda_j (1 - \lambda_j) \sigma_T^2 \qquad (43)$$

$$\sigma_T^2 = \frac{\sum_{i \neq j} \sigma_{ij}}{\lambda_j (1 - \lambda_j)} \qquad (44)$$

จากสมการนี้บ่งชี้ว่า ถ้าสามารถประมาณค่า λ ตัวใดตัวหนึ่ง ซึ่งเรียกว่า λ_j แล้วก็จะสามารถประมาณค่า σ_T^2 ได้ โดยหารผลรวมของแถวที่ j ด้วย $\lambda_j(1-\lambda_j)$ ให้ $\hat{\sigma}^2$ แทนค่าประมาณของตัวอย่าง ความเชื่อมั่นของแบบทดสอบทั้งฉบับ สามารถได้จาก คำนิยามอัตราส่วนความแปรปรวน

$$rel = \frac{\hat{\sigma}_T^2}{\hat{\sigma}_X^2} \tag{45}$$

แทนค่าสมการ (2) ลงใน (3) ได้

$$\boxed{rel = r_{F1} = \frac{\sum_{i \neq j} \hat{\sigma}_{ij}}{\lambda_j (1-\lambda_j) \hat{\sigma}_x^2}} \tag{46}$$

จะเห็นได้ชัดเจนจากสมการ(43) ว่า แถวที่มีผลรวมของความแปรปรวนร่วมที่มีค่ามากที่สุด (เรียกว่า แถว l) จะต้องมีค่าของ $\lambda_j(1-\lambda_j)$ มากที่สุด และ $\lambda_j(1-\lambda_j)$ ที่มีค่ามากที่สุด จะมีความสัมพันธ์กับ λ_j ที่มีค่ามากที่สุด หรือ λ_l ข้อเท็จจริงของข้อความหลังนี้สามารถจะเข้าใจได้ง่าย เมื่อทราบความจริงที่ว่า $\lambda_j(1-\lambda_j)$ เพิ่มขึ้น เมื่อ λ_j เข้าใกล้ .50 (ไม่ว่าจากค่าที่มากกว่าหรือน้อยกว่า) แต่จะไม่มีค่า λ_j ใดเข้าใกล้ .50 ได้เท่ากับ λ_l ดังนั้นจึงไม่มีค่าของ $\lambda_j(1-\lambda_j)$ ใด ๆ ที่มีค่ามากกว่า $\lambda_l(1-\lambda_l)$

ให้หารผลรวมของความแปรปรวนร่วมในแต่ละแถวด้วย ผลรวมของความแปรปรวนร่วมที่มีค่ามากที่สุด

$$C_1 = \frac{\sum_{i \neq 1} \sigma_{1i}}{\sum_{i \neq l} \sigma_{li}} = \frac{\lambda_1(1-\lambda_1)}{\lambda_l(1-\lambda_l)}$$

$$\dots$$

$$C_K = \frac{\sum_{i \neq K} \sigma_{Ki}}{\sum_{i \neq l} \sigma_{li}} = \frac{\lambda_K(1-\lambda_K)}{\lambda_l(1-\lambda_l)} \tag{47}$$

สมการข้างต้นอาจนำมาเขียนใหม่ได้ดังนี้

$$\lambda_j^2 - \lambda_j + C_j(\lambda_l)(1-\lambda_l) = 0$$

และแก้สมการหาค่า λ_j ได้ดังนี้

$$\lambda_j = .5 - \left[.25 - C_j(\lambda_l)(1-\lambda_l)\right]^{1/2} \qquad j \neq l$$

เครื่องหมายหน้ารากที่สองจะต้องเป็นลบ เพราะว่ามีแต่เฉพาะ λ_l เท่านั้นที่อาจมีค่ามากกว่า .50 ดังนั้นปริมาณที่อยู่ภายใต้รากที่สองต้องไม่เป็นลบ เพราะว่า $C_j \leq 1$ และ $\lambda_l(1-\lambda_l) \leq .25$

จากการแก้สมการ โดยทำให้ λ_j อยู่ในเทอมของ λ_l และนิยาม $f(\lambda_l)$ ว่าเป็นผลรวมของ λ_j ลบด้วย 1.0 ดังนี้

$$f(\lambda_l) = \lambda_l + (.5)(K-1) - \sum_{j \neq k}\left[.25 - C_j(\lambda_l)(1-\lambda_l)\right]^{1/2} - 1 \qquad (48)$$

เมื่อ $0 \leq \lambda_l \leq 1$ ขั้นต่อไป คือการหารากที่สองของ $f(\lambda_l) = 0$

ค่าอนุพันธ์ที่หนึ่งของ $f(\lambda_l)$ ที่ $\lambda_l = 0$ มีค่าเท่ากับ $1 + \sum_{j \neq 1} C_j$ ดังนั้น ความลาดชัน (Slope) ของ $f(\lambda_l)$ ที่ $\lambda_l = 0$ มีค่าเป็นบวก ข้อสังเกต $f(1) = -1$ และ $f(1) = 0$ ค่าอนุพันธ์ที่สองของ $f(\lambda_l)$ จะเป็นลบตลอดพิสัย(range)ของ λ_l แสดงว่า $f(\lambda_l)$ จะเว้าต่ำลง และจากข้อเท็จจริงเหล่านี้ แสดงว่าฟังก์ชันมีรากเพียงเล็กน้อยที่ $\lambda_l = 1$ และอย่างมากที่สุดที่มีค่ามากเพียงค่าเดียวอยู่ระหว่าง 0 และ 1 ค่าอนุพันธ์ที่หนึ่งของ $f(\lambda_l)$ ที่ $\lambda_l = 1$ มีค่าเท่ากับ $1 - \sum_{j \neq 1} C_j$ ถ้า $\sum_{j \neq 1} C_j \leq 1$ สมการของ $f(\lambda_l) = 0$ จะมีค่ารากเพียงเล็กน้อยและเมตริกซ์ความแปรปรวนร่วมของตัวอย่างจะไม่ได้คำตอบแบบคะแนนจริงสัมพันธ์ ในเหตุการณ์ เช่นนี้ r_{F1} มีค่าเท่ากับ r_α

ค่าประมาณค่าแรกของ λ_l คือ $\lambda_l = 1/\sum_{j \neq 1} C_j$ ค่าสมการที่แก้ได้นี้ ได้มาโดยการใช้รากที่สองแต่ละค่าในสมการที่ (48) แทนค่าสองเทอมแรกของการกระจายแบบทวินาม(binomial) ของ $\left[.25 - C_j(\lambda_l)(1-\lambda_l)\right]^{1/2}$ และแก้สมการจากสมการกำลังสองของ λ_l ค่าประมาณนี้จะมากกว่า รากที่แท้จริง (true root) และ $f\left(1/\sum C_j\right) > 0$ ปริมาณ 1/K เป็นค่าขอบเขตล่างของรากที่แท้จริง และ $f(1/K) < 0$ วิธีการที่ทำซ้ำ ๆ ซึ่งครอบคลุมการคาดคะเนภายใน (Interpolation) แบบเส้นตรงที่ติดต่อกัน โดยเริ่มที่ขอบเขตเหล่านี้ของ λ_l อาจนำมาใช้การประมาณรากของ $f(\lambda_l) = 0$ ด้วยระดับความแม่นยำใด ๆ ที่ต้องการ

ถ้าใช้ผลรวมของความแปรปรวนร่วมของแถวใดแถวหนึ่ง แทนแถวที่มีผลรวมของความแปรปรวนร่วมมากที่สุด เป็นตัวหารของผลรวมของแถวอื่นๆ สมการที่ได้จะมีค่าแตกต่างได้เล็กน้อย แต่ผลที่ได้จะมีค่าเท่ากัน การประมาณค่าความเชื่อมั่นที่ได้จะมีค่าเพียงค่าเดียวโดยที่มันจะไม่ขึ้นอยู่กับการเลือกว่าจะใช้แถวใดเป็นตัวหาร

ตัวอย่าง

จาก เมตริกซ์ความแปรปรวน-ความแปรปรวนร่วม ที่ $\sigma_x^2 = 115.2845$

$$\begin{bmatrix} 25.3540 & 11.1009 & 8.9309 & 4.4168 \\ 11.1009 & 11.0477 & 5.6798 & 2.5414 \\ 8.9309 & 5.6793 & 6.5121 & 1.8746 \\ 4.4168 & 2.5414 & 1.8746 & 3.2829 \end{bmatrix} \quad \begin{matrix} \sum \hat{\sigma}_{1i} = 24.4486 \\ \sum \hat{\sigma}_{2i} = 19.3216 \\ \sum \hat{\sigma}_{3i} = 16.4848 \\ \sum \hat{\sigma}_{4i} = 8.8328 \end{matrix}$$

แถวที่ 1 มีผลรวมของความแปรวนสูงสุด จึงกำหนดให้เป็นแถวที่ I เราจะได้ C_2 = .7903, C_3 = .6743, และ C_4 = .3613 ขอบเขตของ λ_l อยู่ที่ .25 ถึง .5477

แทนค่า λ_1 ลงใน λ_l ของสมการ (48) ได้

$$f(\lambda_1) = \lambda_1 + .5 - [.25 - (.7903)(\lambda_1)(1-\lambda_1)]^{1/2} - [.25 - (.6743)(\lambda_1)(1-\lambda_1)]^{1/2}$$
$$- [.25 - (.3613)(\lambda_1)(1-\lambda_1)]^{1/2} = 0$$

เมื่อแก้สมการแล้ว จะได้ $\lambda_1 = .4301$ เมื่อได้ λ_1 แล้วสามารถประมาณค่าความเชื่อมั่นจากสมการ (43) กับ (45) ได้

$$r_{F1} = \frac{\sum \hat{\sigma}_{li}}{\lambda_l (1-\lambda_l)(\hat{\sigma}_x^2)}$$

$$= \frac{24.4486}{(.4301)(.5699)(115.2845)} = .8652$$

สัมประสิทธิ์ r_{F2}

การแสดงที่มาของสัมประสิทธิ์ตัวที่สองก็คล้ายคลึงกับการแสดงที่มาของสัมประสิทธิ์ตัวแรก โดยเริ่มตั้งแต่จากการพัฒนาการประมาณค่าความแปรปรวนของคะแนนจริง

ผลรวมของความแปรปรวนร่วมทั้งหมดเท่ากับ ความแปรปรวนคะแนนสอบรวม ลบด้วย ผลรวมของความแปรปรวนส่วนย่อยของแบบทดสอบ นั่นคือ

$$\sum_{i \neq j}\sum \lambda_i \lambda_j \sigma_T^2 = \sigma_X^2 - \sum \sigma_i^2$$

แก้สมการหาค่า σ_T^2 จะได้

$$\sigma_T^2 = \frac{1}{\sum\sum_{i \neq j} \lambda_i \lambda_j}\left(\sigma_X^2 - \sum \sigma_i^2\right) \tag{49}$$

นำค่าของ $\sum \lambda_i = 1$ มา ยกกำลังสองทั้งสองข้าง ได้ผลดังนี้

$$\sum \lambda_i^2 + \sum\sum_{i \neq j} \lambda_i \lambda_j = 1$$

และด้วยเหตุนี้ จะได้

$$\sum\sum_{i \neq j} \lambda_i \lambda_j = 1 - \sum \lambda_i^2$$

นำไปแทนค่าด้านขวาของสมการ (49)

$$\sigma_T^2 = \frac{1}{1 - \sum \lambda_i^2}\left(\sigma_X^2 - \sum \sigma_i^2\right) \tag{50}$$

จากสมการ (45) และ(50) และใช้ค่าของตัวอย่าง จะได้

$$rel = \frac{\hat{\sigma}_T^2}{\hat{\sigma}_X^2} = \frac{1}{1 - \sum \hat{\lambda}_i^2} \cdot \frac{\hat{\sigma}_X^2 - \sum \hat{\sigma}_i^2}{\hat{\sigma}_X^2} \tag{51}$$

การแสดงที่มาของ r_{F2} เริ่มต้นจาก การพิจารณาผลรวมของความแปรปรวนร่วมของแถวใดๆ ลบด้วย ความแปรปรวนร่วมหนึ่งตัวในแถวนั้น ตัวอย่าง เช่น พิจารณารวมในแถวที่ 1 ลบด้วย σ_{12}

$$\begin{aligned} Cov(X_1, X_2) + Cov(X_1, X_3) + \cdots + Cov(X_1, X_K) - Cov(X_1, X_2) \\ = \lambda_1 \lambda_2 \sigma_T^2 + \cdots \lambda_1 \lambda_K \sigma_T^2 - \lambda_1 \lambda_2 \sigma_T^2 \\ = \lambda_1 (1 - \lambda_1 - \lambda_2) \sigma_T^2 \end{aligned}$$

จากนั้น ให้ระบุแถวที่มีผลรวมของความแปรปรวนร่วมมากที่สุด และให้แถวนั้นเป็นแถวที่ I เหมือนที่ได้นิยามไว้แต่แรก และนิยาม D_i ดังนี้

$$D_1 = \frac{\sum_{j\neq 1}\sigma_{1j} - \sigma_{1l}}{\sum_{j\neq l}\sigma_{1l} - \sigma_{1l}} = \frac{\lambda_1\left(1-\lambda_1-\lambda_l\right)\sigma_T^2}{\lambda_l\left(1-\lambda_l-\lambda_1\right)\sigma_T^2} = \frac{\lambda_1}{\lambda_l}$$

$$\ldots$$

$$D_K = \frac{\sum_{j\neq K}\sigma_{Kj} - \sigma_{Kl}}{\sum_{j\neq l}\sigma_{1l} - \sigma_{Ke}} = \frac{\lambda_K\left(1-\lambda_K-\lambda_l\right)\sigma_T^2}{\lambda_l\left(1-\lambda_l-\lambda_K\right)\sigma_T^2} = \frac{\lambda_K}{\lambda_l} \tag{52}$$

ข้อสังเกต $D_l = 1.0$ และทำสมการข้างต้นต่อไปจะได้

$$\lambda_l D_1 = \lambda_1$$
$$\ldots$$
$$\lambda_l D_K = \lambda_K \tag{53}$$

รวมผลด้านซ้ายและด้านขวาทั้งหมดได้

$$\lambda_l \sum D_j = \sum \lambda_j = 1.0$$
$$\lambda_l = \frac{1}{\sum D_j} \quad , \qquad D_l = 1.0 \tag{54}$$

นำสมการ (53) ยกกำลังสองทั้งสองข้างแล้วรวมผลทั้งหมดได้

$$\lambda_l^2 = \sum D_j^2 = \sum \lambda_j^2 \tag{55}$$

จากสมการ (54) และ(55) จะได้

$$\sum \lambda_j^2 = \frac{\sum D_j^2}{\left(\sum D_j\right)^2} \tag{56}$$

ขั้นสุดท้ายแทนค่าสมการ (56) ในสมการ (51) ได้

$$\boxed{r_{F2} = \frac{\left(\sum D_j\right)^2}{\left(\sum D_j\right)^2 - \sum D_j^2} \cdot \frac{\hat{\sigma}_X^2 - \sum \hat{\sigma}_j^2}{\hat{\sigma}_X^2}} \tag{57}$$

วิธีการของเลียว (Liou. 1989 : 154-158) สัมประสิทธิ์ r_{L1} และ r_{L2}

เลียว (Liou. 1989 : 154-158) ได้ทำวิจัยและเสนอสัมประสิทธิ์ใหม่ขึ้นมา สองสัมประสิทธิ์ คือ r_{L1} และ r_{L2} ซึ่งมีข้อดีสามารถคำนวณด้วยเครื่องคำนวณธรรมดา และยังสามารถพิสูจน์ได้ทั้งเชิงทฤษฎี และเลขจำนวน ว่ามีค่าเท่ากับสัมประสิทธิ์ r_k และr_{FS} ของ คริสทอฟ (Kristof. 1974) และของเฟลดด์(Feldt. 1975) เมื่อ K เท่ากับ สาม และสอง ตามลำดับ โดยสัมประสิทธิ์ r_{L1} ปรับมาจากสัมประสิทธิ์ r_k ของ คริสทอฟ (Kristof. 1974) สัมประสิทธิ์ r_{L2} ปรับมาจาก สัมประสิทธิ์ r_{FS} ของเฟลดด์(Feldt. 1975) สัมประสิทธิ์ทั้งสองตัวสามารถแสดงที่มาดังนี้

การแสดงที่มาของสัมประสิทธิ์

ให้ σ_X^2 แทน ความแปรปรวนของคะแนนสอบรวมทั้งฉบับ และ σ_E^2 แทน ความแปรปรวนของคะแนนความคลาดเคลื่อนรวมทั้งฉบับ ตามทฤษฎีคะแนนจริงแบบมาตรฐานเดิมความเชื่อมั่นของแบบทดสอบรวมทั้งฉบับ คือ อัตราส่วน σ_T^2/σ_X^2 หรือเท่ากับ $1-\left(\sigma_E^2/\sigma_X^2\right)$ ผลรวมของความแปรปรวนส่วนย่อย หารด้วย ความแปรปรวนรวมทั้งฉบับได้ดังนี้

$$\frac{\sum \sigma_i^2}{\sigma_X^2} = \frac{\sum \lambda_i^2 \sigma_T^2}{\sigma_X^2} + (1-\rho)$$

เมื่อ ρ แทนความเชื่อมั่นของแบบทดสอบรวมทั้งฉบับ ดังนั้น

$$\boxed{\rho = \frac{\sum \lambda_i^2 \sigma_T^2}{\sigma_X^2} + \left(1-\frac{\sum \sigma_i^2}{\sigma_X^2}\right)} \qquad (58)$$

ถ้าสามารถหาค่า $\sum \lambda_i^2$ ได้แล้วความเชื่อมั่นสามารถเขียนเป็นสมการใหม่ได้ดังนี้

$$\boxed{\rho = \frac{1}{1-\sum \lambda_i^2}\left(1-\frac{\sum \sigma_i^2}{\sigma_X^2}\right)} \qquad (59)$$

บุคคลแรกที่ได้แสดงที่มาของ สมการ (59) คือราชู (Raju. 1977) และได้ตรวจสอบว่า $\sum \lambda_i^2 \geq 1/K$ และถ้ามีค่าเท่ากันแล้ว สมการ (59) จะกลายเป็นสัมประสิทธิ์ แอลฟ่า ในการประมาณค่าของ ρ ต้องคำนวณค่าของ $\sum \lambda_i^2 \sigma_T^2$ และ $\sum \lambda_i^2$ ก่อนแล้วนำไปแทนค่าลงในสมการ (58) และ (59)

สัมประสิทธิ์ r_{L1}

สัมประสิทธิ์ r_{L1} ตัวนี้ คริสทอฟ (Kristof. 1974) เป็นผู้คิดขึ้นครั้งแรก โดยบังคับว่าต้องแบ่งส่วนย่อยของแบบทดสอบเป็นสามเท่านั้น เพื่อให้การแก้สมการได้ค่าออกมาเพียงค่าเดียว ถ้าแทนค่า $\lambda_i^2\sigma_T^2$ (i = 1,2,3) แต่ละค่า ด้วย Q_1, Q_2 และQ_3 ตามลำดับ ซึ่งค่าเหล่านี้คำนวณได้จากค่าความแปรปรวนร่วมของคะแนนสอบของแต่ละส่วนย่อย ดังนั้น

$$\sum \lambda_i^2\sigma_T^2 \equiv Q_1 + Q_2 + Q_3$$
$$= \frac{\sigma_{12}\sigma_{13}}{\sigma_{23}} + \frac{\sigma_{12}\sigma_{23}}{\sigma_{13}} + \frac{\sigma_{13}\sigma_{23}}{\sigma_{12}}$$

เมื่อแทนค่าลงในสมการ (58) แล้ว สัมประสิทธิ์ r_k ของคริสทอฟ (Kristof. 1974) เป็นดังนี้

$$r_k = \frac{Q_1 + Q_2 + Q_3}{\sigma_X^2} + \left[1 - \frac{\left(\sigma_1^2 + \sigma_2^2 + \sigma_3^2\right)}{\sigma_X^2}\right]$$

เมื่อ K>3 จะมีพารามิเตอร์ที่ไม่ทราบค่ามากเกินกว่าที่จะหาได้จากข้อมูล ถ้าสมมติว่าคะแนนส่วนย่อยของแบบทดสอบที่สอบได้ มีการแจกแจงของตัวแปรพหุแบบปกติแล้ว (multivariate normal distribution) การประมาณค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับต้องใช้วิธีการของโจเรสคอก (Joreskog's approach) ซึ่งต้องใช้คอมพิวเตอร์ในการคำนวณแบบ maximum likelihood มีวิธีการคำนวณที่ยุ่งยาก ไม่เหมาะที่จะใช้ในทางปฏิบัติ ดังนั้นจึงได้พัฒนาวิธีการคิดคำนวณแบบง่ายๆ ขึ้นมาโดยหาค่าประมาณของ $\sum \lambda_i^2\sigma_T^2$ จากเมตริกซ์ความแปรปรวนร่วมระหว่างส่วนย่อยของแบบทดสอบ ผลรวมของความแปรปรวนร่วมของหลัก j (column j) สามารถเขียนแทนได้ดังนี้

$$\sum_{i \neq j} \sigma_{ij} = \sum_{i \neq j} \lambda_i \lambda_j \sigma_T^2 = \lambda_j \sum_{i \neq j} \lambda_i \sigma_T^2$$

และผลรวมของความแปรปรวนร่วมกำลังสองคือ

$$\sum_{i \neq j} \sigma_{ij}^2 = \sum_{i \neq j} \lambda_i^2 \lambda_j^2 \sigma_T^4 = \lambda_j^2 \sum_{i \neq j} \lambda_i^2 \sigma_T^4$$

เมื่อ K > 3 แล้ว Q_i (i = 1,….., K) สามารถที่จะหาค่าได้ดังนี้

$$\left.\begin{aligned} Q_1 &= \frac{\left\lfloor \left(\sum \sigma_{i1}\right)^2 - \sum \sigma_{i1}^2 \right\rfloor}{\sum\sum_{i \neq j} \sigma_{ij}} \\ &= \frac{\left\lfloor \lambda_1^2 \sigma_T^4 \left(\sum \lambda_i\right)^2 - \lambda_1^2 \sigma_T^4 \left(\sum \lambda_i^2\right) \right\rfloor}{\sum\sum_{i \neq j} \lambda_i \lambda_j \sigma_T^2} \\ &= \frac{\left[\lambda_1^2 \sigma_T^4 \left(\sum\sum_{i \neq j} \lambda_i \lambda_j \right) \right]}{\sigma_T^2 \left(\sum\sum_{i \neq j} \lambda_i \lambda_j \right)} \\ &= \lambda_1^2 \sigma_T^2 \end{aligned}\right\} \qquad (i.j \neq 1)$$

$$\ldots$$

$$\left.\begin{aligned} Q_K &= \frac{\left\lfloor \left(\sum \sigma_{ik}\right)^2 - \sum \sigma_{ik}^2 \right\rfloor}{} \\ &= \lambda_K^2 \sigma_T^2 \end{aligned}\right\} \qquad (i, j \neq K)$$

รวมด้านซ้ายทั้งหมด และรวมด้านขวาทั้งหมดจะได้ $\sum Q_i = \sum \lambda_i^2 \sigma_T^2$ การสร้างตัวประมาณค่าความเชื่อมั่น r_{L1} ทำได้โดยแทนค่า $\sum Q_i$ ลงในสมการ (58) และสามารถแสดงให้เห็นว่าสัมประสิทธิ์ของคริสทอฟ (Kristof. 1974) เป็นกรณีเฉพาะของ r_{L1} เมื่อ K = 3 โดยใช้การกระจายทางพีชคณิตอย่างง่าย

ถ้ากระจายสัมประสิทธิ์แอลฟ่า (r_α) ในรูปแบบของสมการ (58) ตัวดัชนีดังกล่าวจะได้ดังนี้

$$r_\alpha = \frac{\sum\sum_{i \neq j} \sigma_{ij}}{\left[(K-1)\sigma_X^2\right]} + \left(1 - \frac{\sum \sigma_i^2}{\sigma_X^2}\right)$$

ความแตกต่างระหว่างขนาดของ r_{L1} และ r_α เท่ากับ Δ

$$\Delta = (K-1)\sum Q_i - \sum\sum_{i \neq j} \sigma_{ij}$$

ถ้าแต่ละ $\lambda_i^2 \sigma_T^2$ แก้สมการหาค่าได้ค่าเดียว และ Q_i มีค่าเป็นบวกทั้งหมดแล้ว จะได้

$$\Delta = (K-1)\sum \left(Q_i^{1/2}\right)^2 - \sum_{i \neq j}\sum \left(Q_i^{1/2}\right)\left(Q_j^{1/2}\right)$$
$$= \sum_{i<J}\sum \left(Q_i^{1/2} - Q_j^{1/2}\right)^2 \qquad \geq 0$$

เมื่อ K = 3 แล้ว r_{L1} จะให้ค่าประมาณที่ดีกว่า r_α เสมอ ยกเว้นเฉพาะที่ Q_i มีค่าเท่ากันหมดเท่านั้น (Kristof . 1974:495) ถ้ากำหนดให้ ข้อมูลเป็นแบบจำลองคะแนนจริงสัมพันธ์แล้ว ผลต่างโดยเฉลี่ย(Δ)มีค่าประมาณเท่ากับ $\left(K\sum \lambda_i^2 - 1\right)\sigma_T^2$ ซึ่งมีค่ามากกว่าหรือเท่ากับ ศูนย์ กล่าวได้ว่า สัมประสิทธิ์ r_{L1} จะประมาณค่าความเชื่อมั่น ได้ดีกว่า r_α เมื่อแบบทดสอบเป็นแบบคะแนนจริงสัมพันธ์

สัมประสิทธิ์ r_{L2}

ถ้าส่วนย่อยของแบบทดสอบเป็นแบบคะแนนจริงสัมพันธ์ที่แทนขนาดความยาวที่สั้นกว่าแบบทดสอบฉบับรวมแล้ว ความแปรปรวน $\sigma_{E_i}^2$ จะมีความสัมพันธ์ กับ σ_E^2 ถ้าอาศัยทฤษฎีคะแนนจริงแบบมาตรฐานเดิมแล้ว $\sigma_{E_i}^2 = \lambda_i \sigma_E^2$ จากข้อตกลงดังกล่าวนี้สามารถที่จะประมาณค่าของ ρ ได้จากการแบ่งแบบทดสอบเป็นสองส่วน ตามที่ เฟลดต์ (Feldt. 1975) ได้พิสูจน์ให้เห็นข้างต้นในเชิงสถิติแล้ว ข้อตกลงของ $\sigma_{E_i}^2$ ทำให้มีข้อมูลมากเกินจำเป็น เมื่อแบบทดสอบประกอบด้วยส่วนย่อยหลายส่วน คริสทอฟ (Kristof. 1971) ใช้วิธีความน่าจะเป็นสูงสุด (ML) พัฒนาการประมาณค่าพารามิเตอร์ และทดสอบสมมติฐานภายใต้ แบบจำลองของ r_{FS} การแก้สมการโดยหาค่าประมาณนี้ ทำให้การคำนวณตามวิธีการของคริสทอฟง่ายลง ซึ่ง r_{L2} ก็ใช้หลักการเดียวกัน

การประมาณค่าของ λ_i (i = 1, ….., K) หาได้โดยคำนวณจาก $\sum \sigma_{ij}$ (และที่ i = j แล้ว $\sigma_{ij} = \sigma_i^2$)ซึ่งได้มาดังนี้

$$\sum_j \sigma_{1j} = \lambda_1\left(\sum_j \lambda_j\right)\sigma_T^2 + \lambda_1 \sigma_E^2$$
$$= \lambda_1\left(\sigma_T^2 + \sigma_E^2\right)$$
$$= \lambda_1 \sigma_X^2$$
$$\ldots$$
$$\lambda_K = \frac{\sum \sigma_{Kj}}{\sigma_X^2}$$

แทนค่า λ_i ในสมการ (59) จะได้การประมาณค่าของ ρ ซึ่งแทนด้วย r_{L2} และตามที่ระบุไว้ตอนต้นว่า $\sum \lambda_i^2$ จะมีค่ามากกว่า หรือ เท่ากับ 1/K ยกเว้นแต่ค่าที่ได้จะเป็นลบ ดังนั้นตัวประมาณค่าที่

ได้มาต้องไม่น้อยกว่า สัมประสิทธิ์ แอลฟ่า ในทุกกรณี และยังสามารถแสดงให้เห็นว่า เมื่อแบ่งแบบทดสอบเป็นสองส่วน สูตรของเฟลดต์ (Feldt. 1975) เป็นกรณีเฉพาะของ r_{L2}นั่นคือ

$$\lambda_1^2 + \lambda_2^2 = \frac{\left[\sigma_X^2 + \left(\sigma_1^2 - \sigma_2^2\right)\right]^2}{4\sigma_X^4} + \frac{\left[\sigma_X^2 - \left(\sigma_1^2 - \sigma_2^2\right)\right]^2}{4\sigma_X^4}$$

$$= \frac{\sigma_X^4 + \left(\sigma_1^2 - \sigma_2^2\right)^2}{2\sigma_X^4}$$

$$= \frac{\left(\sigma_1^2 + \sigma_{12}\right)^2 + \left(\sigma_2^2 + \sigma_{21}\right)^2}{\sigma_x^4}$$

และ $r_F = r_{L2}$

วิธีการของบุญเชิด (บุญเชิด ภิญโญอนันตพงษ์ (2540) สัมประสิทธิ์ r_B

บุญเชิด ภิญโญอนันตพงษ์ (2540) เสนอสูตรความเชื่อมั่นสำหรับแบบทดสอบที่ให้คะแนนแบบ 0,1 และมีความยากรายข้อไม่เท่ากัน และคำนวณโดยไม่ใช้จำนวนข้อมาคิดคำนวณ มีการแสดงที่มาของสูตรดังนี้

กำหนดให้แบบทดสอบประกอบด้วยข้อสอบ K ข้อ คะแนนสอบ คะแนนจริงและคะแนนคลาดเคลื่อนของแต่ละข้อ และของทั้งฉบับแทนด้วยสัญลักษณ์ดังนี้

คะแนนสอบแต่ละข้อและทั้งฉบับ แทนด้วย X_1, X_2, X_3 , … ,X_K และX

คะแนนจริงแต่ละข้อและทั้งฉบับ แทนด้วย T_1, T_2,T_3 , ... ,T_K และ T

คะแนนคลาดเคลื่อนแต่ละข้อและทั้งฉบับ แทนด้วย E_1, E_2 ,E_3 , ... E_K และ E ตามลำดับ

ดังนั้น

$$X = X_1 + X_2 + X_3 + \cdots + X_K \qquad (60)$$

$$T = T_1 + T_2 + T_3 + \cdots + T_K \qquad (61)$$

$$E = E_1 + E_2 + E_3 + \cdots + E_K \qquad (62)$$

จากแบบจำลองคะแนนของทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) ได้คะแนนแต่ละข้อดังนี้

$$X_i = T_i + E_i \qquad (63)$$

และกำหนดข้อตกลงว่า แบบทดสอบประกอบด้วยข้อสอบที่วัดคุณลักษณะเดียวกัน และเป็นไปตามข้อตกลงแบบคะแนนจริงสัมพันธ์ กล่าวคือ คะแนนจริงของแต่ละข้อมีความสัมพันธ์เป็นเส้นตรงกับคะแนนจริงทั้งฉบับ

$$T_i = \lambda_i T + C_i \qquad (i = 1,..., K) \qquad (64)$$

เมื่อ λ_i เป็นลักษณะประจำข้อที่ i ซึ่งกำหนดให้เป็นจำนวนจริงบวก

C_i เป็นค่าคงที่ของข้อสอบ i

โดยกำหนดเงื่อนไขว่า $\sum \lambda_i = 1$ และ $\sum C_i = 0$ นำสมการ (64) แทนค่าลงในสมการ (63) ได้

$$X_i = \lambda_i T + C_i + E_i \qquad (i = 1,...,K) \tag{65}$$

ถ้านำแบบทดสอบไปสอบกับกลุ่มตัวอย่างจำนวนหนึ่ง และหาค่าความแปรปรวนของข้อสอบแต่ละข้อ (σ_i^2) และหาค่าความแปรปรวนของข้อสอบแต่ละคู่ (σ_{ij}) ได้ดังนี้

$$\sigma_i^2 = \lambda_i^2 \sigma_T^2 + \sigma_{E_i}^2 \tag{66}$$

$$\sigma_{ji} = \lambda_j \lambda_i \sigma_T^2 \qquad (j \neq i) \tag{67}$$

ผลรวมของความแปรปรวนของข้อสอบทุกข้อคือ

$$\sum \sigma_i^2 = \sum \lambda_i^2 \sigma_T^2 + \sum \sigma_{E_i}^2 \tag{68}$$

นำ σ_X^2 ไปหารสมการ (68) และนิยามความเชื่อมั่นของแบบทดสอบ $\rho_{XX'} = \sigma_T^2 / \sigma_X^2$ จะได้สูตรความเชื่อมั่นซึ่งราชู (Raju.1977) เป็นผู้เสนอไว้เป็นคนแรกดังนี้

$$\rho_{XX'} = \frac{1}{1 - \sum \lambda_i^2}\left[1 - \frac{\sum \sigma_i^2}{\sum \sigma_X^2}\right] \tag{69}$$

เลียว (Liou.1989) ได้อาศัยวิธีการของเฟลดต์ (Feldt. 1975) ที่จะประมาณค่า λ_i โดยพิจารณาว่าข้อสอบแต่ละข้อก็คือแบบทดสอบที่ลดจำนวนข้อลงจนเหลือข้อสอบเพียงข้อเดียว ดังนั้นข้อสอบก็คือแบบทดสอบฟอร์มหนึ่งนั่นเอง ซึ่งตามทฤษฎีการทดสอบแบบมาตรฐานเดิม ความแปรปรวนของคะแนนคลาดเคลื่อนแต่ละข้อและทั้งฉบับจะมีความสัมพันธ์กันดังนี้ $\sigma_{E_i}^2 = \lambda_i \sigma_E^2$ และถ้านำแบบทดสอบแต่ละข้อมาสร้างเป็นเมตริกซ์ความแปรปรวน-ความแปรปรวนร่วมแล้ว จะสามารถหาค่า λ_i ของแต่ละข้อจากแต่ละแถว (Row) ของเมตริกซ์ดังกล่าวได้ โดยนำค่า σ_X^2 ไปหารผลรวมของความแปรปรวน-ความแปรปรวนร่วมในแต่ละแถว ดังนี้

$$\begin{aligned} \lambda_1 &= \sum \sigma_{1i} / \sigma_X^2 \\ \lambda_2 &= \sum \sigma_{2i} / \sigma_X^2 \\ &\ \vert \qquad\quad \vert \\ \lambda_K &= \sum \sigma_{Ki} / \sigma_X^2 \end{aligned} \tag{70}$$

จากสมการ (70) นำมาเขียนใหม่โดยแยกความแปรปรวนร่วมของคะแนนตัวมันเองออกมา ซึ่งก็คือ ความแปรปรวนของข้อนั้น เช่น

$$\lambda_1 = \frac{\sigma_1^2 + \sum \sigma_{1i}}{\sigma_X^2} \tag{71}$$

ถ้าสามารถหาค่าความแปรปรวนของคะแนนข้อ 1 กับข้ออื่นๆ ทั้งหมดได้ ก็สามารถหาค่า λ_i ได้ การหาค่าความแปรปรวนร่วมของข้อสอบแต่ละคู่จากสูตรโดยตรง มีความยุ่งยากและใช้เวลานานมาก วิธีที่ง่ายและรวดเร็วกว่าคือ อาศัยคำนิยามของสัมประสิทธิ์สหสัมพันธ์ $\rho_{XY} = \sigma_{XY} / \sigma_X \sigma_Y$ ดังนั้น ความแปรปรวนร่วมของข้อสอบข้อ 1 และข้อ 2 ได้จากการนำค่าส่วนเบี่ยงเบนมาตรฐานคูณค่าสัมประสิทธิ์สหสัมพันธ์คือ $\sigma_{12} = \sigma_1 \sigma_2 \rho_{12}$ แต่เนื่องจากข้อสอบแต่ละข้อให้คะแนนในระบบ (0,1) ดังนั้น

$$\sigma_1 = \sqrt{P_1(1-P_1)} \text{ และ } \sigma_2 = \sqrt{P_2(1-P_2)} \tag{72}$$

เมื่อ P เป็นสัดส่วนของผู้ตอบถูก หรือความยากง่ายของข้อสอบ

สำหรับสัมประสิทธิ์สหสัมพันธ์จะต้องใช้สูตรกรณีเฉพาะของสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson) คือ สัมประสิทธิ์ ϕ ซึ่งมีสูตรดังนี้

$$\phi = \frac{BC - AD}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} \tag{73}$$

เมื่อ A, B, C และ D เป็นสัดส่วนจำนวนผู้ตอบสองข้อถูกหรือผิด ดังแสดงในตาราง

		ข้อ 2 0	1	
	1	A	B	A+B
ข้อ 1	0	C	D	C+D
		A+C	B+D	1.0

จากตารางจะได้ว่า

$A+B = P_1$

$B+D = P_2$

$C+D = 1-P_1$

$A+C = 1-P_2$

และจัดค่าต่างๆ อยู่ในรูปของ P_1 , P_2 และ B จะได้

$A = P_1-B$, $D = P_2-B$, $C = 1-P_1-P_2+B$

นำค่าดังกล่าวแทนในสมการ (73) ได้

$$\begin{aligned}\phi_{12} &= \frac{B_{12}\left(1-P_1-P_2+B_{12}\right)-\left(P_1-B_{12}\right)\left(P_2-B_{12}\right)}{\sqrt{P_1\left(1-P_1\right)\left(1-P_2\right)P_2}} \\ &= \frac{B_{12}-P_1P_2}{\sqrt{P_1\left(1-P_1\right)}\sqrt{P_2\left(1-P_2\right)}}\end{aligned} \tag{74}$$

ดังนั้นจากสมการ (72) และ (74) ได้ความแปรปรวนร่วมดังนี้

$$\sigma_{12} = \sqrt{P_1\left(1-P_1\right)}\sqrt{P_2\left(1-P_2\right)}\frac{B_{12}-P_1P_2}{\sqrt{P_1\left(1-P_1\right)}\sqrt{P_2\left(1-P_2\right)}}$$

$$\sigma_{12} = B_{12}-P_1P_2$$

ทำนองเดียวกัน

$$\sigma_{13} = B_{13}-P_1P_3$$
$$\sigma_{14} = B_{14}-P_1P_4$$

หรือเขียนเป็นรูปทั่วไปได้ดังนี้

$$\begin{aligned}\sigma_{1i} &= B_{1i}-P_1P_i \\ \sigma_{2i} &= B_{2i}-P_2P_i \\ &\vdots \\ \sigma_{Ki} &= B_{Ki}-P_KP_i\end{aligned}$$

ดังนั้นจากสมการ (70) และ (71) นำมาเขียนใหม่ได้

$$\begin{aligned}
\lambda_1 &= \frac{P_1(1-P_1)+\sum(B_{1i}-P_1P_i)}{\sigma_X^2} && i \neq 1 \\
&= \frac{P_1(1-P_1)+\sum B_{1i}-P_1\sum P_i}{\sigma_X^2} \\
&= \frac{P_1\left(1-P_1+\sum P_i\right)+\sum B_{1i}}{\sigma_X^2} \\
&= \frac{P_1(1-\overline{X})+\sum B_{1i}}{\sigma_X^2} \\
\lambda_1 &= \frac{\sum B_{1i} - P_1(\overline{X}-1)}{\sigma_X^2} && i \neq 1
\end{aligned}$$

ทำนองเดียวกัน

$$\begin{aligned}
\lambda_2 &= \frac{\sum B_{2i} - P_2(\overline{X}-1)}{\sigma_X^2} && i \neq 2 \\
\vdots \quad & \quad \vdots \\
\lambda_K &= \frac{\sum B_{Ki} - P_K(\overline{X}-1)}{\sigma_X^2} && i \neq K
\end{aligned}$$

ดังนั้นการประมาณค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม (Classical-Congeneric) จากแบบทดสอบที่ให้คะแนนระบบ (0,1) จำนวน K ข้อ โดยใช้สัมประสิทธิ์ r_B ที่สร้างขึ้นใหม่ มีสูตรดังนี้

$$r_B = \frac{1}{1-\sum \lambda_i^2}\left[1 - \frac{\sum P_i(1-P_i)}{S_X^2}\right] \qquad (75)$$

เมื่อ P_i = ค่าความยากง่ายของข้อสอบแต่ละข้อ

S_X^2 = ความแปรปรวนของคะแนนทั้งฉบับ

B_{ji} = สัดส่วนจำนวนผู้ตอบสองข้อใด ๆ ตรงกัน

$$\lambda_1 = \left[\sum B_{1i} - P_1\left(\overline{X} - 1\right)\right] / S_x^2$$
$$\lambda_2 = \left[\sum B_{2i} - P_2\left(\overline{X} - 1\right)\right] / S_x^2$$
$$| \qquad |$$
$$\lambda_K = \left[\sum B_{Ki} - P_K\left(\overline{X} - 1\right)\right] / S_x^2$$
$$\overline{X} = \sum P_i$$

จากการแสดงที่มาของสูตรสามารถเขียนเป็นขั้นตอนการหาค่าสัมประสิทธิ์ r_B ได้ดังต่อไปนี้

ขั้นตอนการหาค่าสัมประสิทธิ์ r_B

การหาค่าสัมประสิทธิ์ r_B จากแบบทดสอบ k ข้อ ที่ให้คะแนนในระบบ 0, 1 ดำเนินการเป็นขั้นตอนดังนี้

ขั้นที่ 1 สร้างเมตริกซ์คะแนนสอบของแต่ละข้อและคะแนนรวม โดยมีข้อสอบอยู่ในแต่ละหลัก (column) และนักเรียนแต่ละคนอยู่ในแต่ละแถว (row)

ขั้นที่ 2 หาค่า P_i ของแต่ละข้อ ซึ่งเป็นสัดส่วนจำนวนผู้ตอบถูกในแต่ละข้อ และ $\sum P_i\left(1 - P_i\right)$

ขั้นที่ 3 หาค่า $\overline{X} = \sum P_i$ และ $S_X^2 = \dfrac{N\sum X^2 - \left(\sum X\right)^2}{N(N-1)}$

ขั้นที่ 4 สร้างเมตริกซ์ของสัดส่วนจำนวนผู้ตอบถูกทั้งสองข้อของข้อสอบแต่ละคู่ทุกคู่ (B_{ji})

ขั้นที่ 5 คำนวณค่า λ_k

ขั้นที่ 6 แทนค่า P_i , λ_k และ S_X^2 ในสมการ (75)

เมื่อพิจารณาตามบริบทของโรงเรียนที่ยังมีการใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่มีการให้คะแนนแบบ 2 ค่า คือ 0 กับ 1 และแต่ละข้อมีความยากไม่เท่ากัน มีความเป็นเอกพันธ์หรือวัดเนื้อหาเดียวกัน วิธีการของบุญเชิด (บุญเชิด ภิญโญอนันตพงษ์ (2540) ซึ่งเรียกว่าสัมประสิทธิ์ r_B จึงมีความเหมาะสมที่จะนำมาใช้ในการวิจัยครั้งนี้

2.การประมาณค่าความเชื่อมั่นโมเดลคะแนนจริงสัมพันธ์จากโมเดลการวัดคะแนนจริงสัมพันธ์

โมเดลการวัดคะแนนจริงสัมพันธ์ (Congeneric Measurement Model)

โมเดลการวัดคะแนนจริงสัมพันธ์เป็นโมเดลการวัดย่อยหนึ่งของโมเดลสมการโครงสร้างและโมเดลสมการโครงสร้างก็เป็นโมเดลย่อยหนึ่งของโมเดลลิสเรล

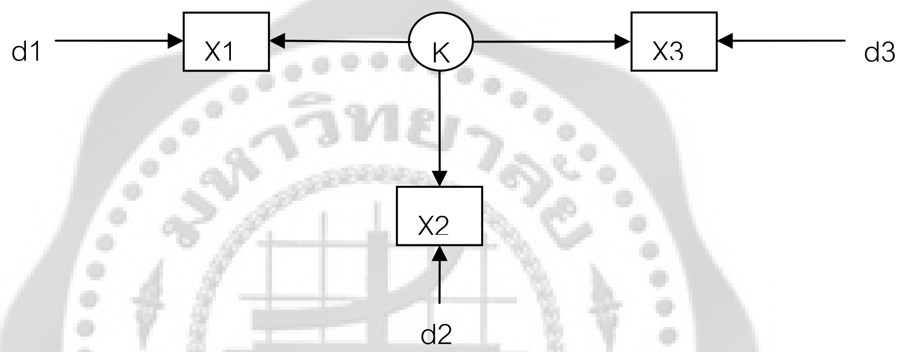
ตามทฤษฎีการทดสอบมาตรฐานเดิม เมื่อมีตัวแปรสังเกตได้ X1, X2, X3 มีค่าคะแนนจริงเป็น T1, T2, T3 ตามลำดับ จะแบ่งลักษณะของตัวแปรสังเกตได้ เป็น 3 แบบ ตามคุณสมบัติของคะแนนจริง(นงลักษณ์ วิรัชชัย.2542 : 37) กล่าวคือ ถ้าสหสัมพันธ์ระหว่างคะแนนจริงแต่ละคู่เป็นค่าสหสัมพันธ์สมบูรณ์ คือมีค่าเท่ากับ 1 เรียกว่า ตัวแปรคะแนนจริงสัมพันธ์ คือ ตัวแปรมีคะแนนจริงร่วมกัน

ถ้ามีความแปรปรวนของคะแนนจริงทุกตัวเท่ากัน และมีความแปรปรวนของความคลาดเคลื่อนเท่ากันด้วย เรียกว่าเป็น ตัวแปรคู่ขนาน

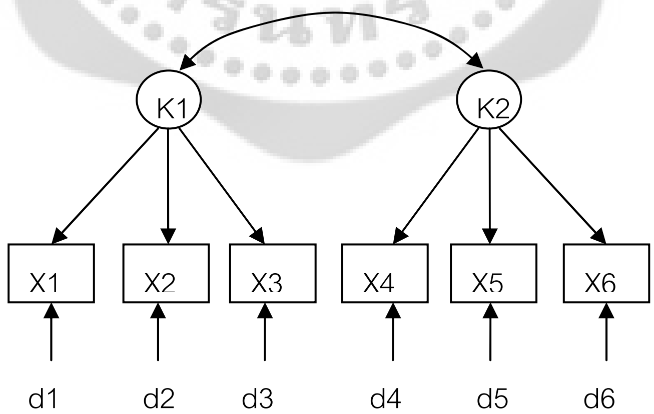
ถ้าความแปรปรวนของคะแนนจริงทุกตัวเท่ากัน แต่ความแปรปรวนของความคลาดเคลื่อนไม่เท่ากันด้วย เรียกว่าเป็น ตัวแปรสมมูล

จากนิยามของตัวแปรคะแนนจริงสัมพันธ์สามารถเขียนโมเดลการวัดคะแนนจริงสัมพันธ์ ได้ 2 โมเดล คือ

1. โมเดลการวัดองค์ประกอบเดี่ยวคะแนนจริงสัมพันธ์ (one factor congeneric measurement model)



2. โมเดลการวัดพหุองค์ประกอบคะแนนจริงสัมพันธ์ (multifactor congeneric measurement model)



การประมาณค่าความเชื่อมั่นจากโมเดลการวัดคะแนนจริงสัมพันธ์

โจเรสคอก (*Jöreskog* .1971 :107-112) เป็นบุคคลแรกที่ใช้คำว่า Congeneric Tests ในทฤษฎีทดสอบคะแนนจริงมาตรฐานเดิม (CTT) โดยได้เสนอโมเดลการทดสอบคะแนนจริงสัมพันธ์ที่วัดองค์ประกอบเดียว (Model for One Set of Congeneric Test Score) โดยกำหนดให้

τ เป็นตัวแปรสุ่ม

μ_i เป็นค่าเฉลี่ยของ X_i

ถ้ากำหนดให้ τ เป็นตัวแปรสุ่มและ $T_i(T_1,T_2,\ldots,T_K)$ แต่ละค่ามีความสัมพันธ์เชิงเส้นตรงกับ τ นั่นคือ

$T_i = \beta_i\tau + \mu_i$ หรือ $T_i = \lambda_i\tau + \mu_i$

โดย β_i และ λ_i เป็นค่าน้ำหนักองค์ประกอบ และ μ_i เป็นค่าคงที่ซึ่งมีข้อตกลงว่า

$\varepsilon(\tau)=0$ และ $Var(\tau)=1$

ก็จะได้สมการ

$X_i = \beta_i\tau + \mu_i + e$ หรือ $X_i = \lambda_i\tau + \mu_i + e$

ถ้าให้ X, β , μ และ e เป็นเวกเตอร์หลัก(Column Vector) สามารถเขียนสมการในรูปของเมทริกซ์ได้ดังนี้

$$x = \mu + \beta\tau + e$$

ความเชื่อมั่นของการทดสอบที่ i ก็จะได้

$$\rho_i = \frac{\beta_i^2}{\beta_i^2 + \theta_i^2} \quad \text{หรือ} \quad \rho_i = \frac{\lambda_i^2}{\lambda_i^2 + \theta_i^2} \qquad (76)$$

เมื่อ θ_i^2 เป็นความแปรปรวนของ e_i

แล้วนำเสนอการประมาณค่าด้วยวิธีความน่าจะเป็นสูงสุด (maximum likelihood method) ได้ค่าความเชื่อมั่นที่ประมาณค่าด้วยวิธีความน่าจะเป็นสูงสุด (ML) จาก

$$\hat{\rho}_i = \frac{\hat{\beta}_i^2}{\hat{\beta}_i^2 + \hat{\theta}_i^2} \qquad (77)$$

กำหนด $\alpha' = (\alpha_1, \alpha_2, \ldots, \alpha_K)$ เป็นน้ำหนักที่สัมพันธ์กัน และ

$$y = \alpha' X$$

โดย $x = \mu + \beta\tau + e$ ก็จะได้

$$y = \alpha'\mu + (\alpha'\beta)\tau + \alpha'\tau$$

ค่าความเชื่อมั่น ก็จะได้

$$\rho = \frac{(\alpha'\beta)^2}{(\alpha'\beta)^2 + \alpha'\Theta^2\alpha} \quad (78)$$

เมื่อ β เป็นค่าประมาณด้วยวิธีความน่าจะเป็นสูงสุด ของ β

μ เป็นค่าประมาณด้วยวิธีความน่าจะเป็นสูงสุด ของ μ

Θ เป็นค่าประมาณด้วยวิธีความน่าจะเป็นสูงสุด ของ θ

วิธีประมาณค่าความเชื่อมั่นด้วยวิธีความน่าจะเป็นสูงสุด (ML) เป็นวิธีการคำนวณที่ยุ่งยากในขณะนั้น (เมื่อปี 1971) ต่อมา เรย์คอฟ (Raykov . 1997: 174-178) ได้นำวิธีการของ โจเรสคอก (*Jöreskog* .1971) มาใช้ในการหาความเชื่อมั่นรวม (composite reliability) ซึ่งเป็นการประมาณค่าโดยใช้โมเดลสมการโครงสร้าง โดยได้เสนอสูตรสัมประสิทธิ์ความเชื่อมั่นรวม ดังนี้

$$\begin{aligned}\rho_{xx} &= \frac{Var(T_1 + ... + T_k)}{Var(X_1 + ... + X_k)} \\ &= \frac{Var\left(\sum a_i + T_1 \sum b_i\right)}{Var\left(\sum a_i + T_1 \sum b_i + \sum E_i\right)} \\ &= \frac{\left(\sum b_i\right)^2 Var(T_1)}{\left(\sum b_i\right)^2 Var(T_1) + \sum Var(E_i)}\end{aligned} \quad (79)$$

โดย $X_{(w)} = w_1 X_1 + w_2 X_2 + ... + w_k X_k$

เมื่อ $w, w_1, w_2, ..., w_k$ เป็นน้ำหนักแต่ละองค์ประกอบ

จากนั้นในปี 2003 เรย์คอฟ(Raykov. 2003: 143-159) ได้นำเสนอการประมาณค่าความเชื่อมั่นรวมอีกครั้งโดยมีฐานคิดจากเดิม เพียงแต่เปลี่ยนสัญลักษณ์ โดยเริ่มจากฐานคิดของทฤษฎีการทดสอบมาตรฐานเดิม(CTT) คือ

$$T_i = a_{ij} + b_{ij}T_{j'} \quad (80)$$

สามารถเขียนสมการให้มีความเหมาะสม ได้สมการ

$$T_i = a_i + b_i\eta \quad (81)$$

เมื่อ η ถูกสร้างขึ้นมาภายใต้เงื่อนไขของโครงสร้างคะแนนจริงสัมพันธ์(Congeneric Construct) ตามโมเดลการทดสอบคะแนนจริงสัมพันธ์ หลังจากเอาส่วนประกอบออกจากคะแนนสังเกตได้ เมื่อนำมาเขียนในรูปของคะแนนสังเกตได้ก็จะได้สมการ

$$Y_i = a_i + b_i\eta + E_i \qquad (82)$$

เมื่อ E_i คือคะแนนความคลาดเคลื่อน(i= 1, 2,…,k) เพื่อให้สามารถใช้อ้างอิงทั่วไปได้จะกำหนดให้ Var(η)=1 เมื่อ Var(.) แสดงถึงความแปรปรวนในประชากร ค่าสัมประสิทธิ์ความเชื่อมั่นของมาตรวัด Y จะมีค่าเท่ากับ $Y = Y_1 + Y_2 + \ldots + Y_k$ ผลก็คือ

$$\rho_{Y,K} = \frac{\left(\sum_{i=1}^{k} b_i\right)^2 Var(\eta)}{\left(\sum_{i=1}^{k} b_i\right)^2 Var(\eta) + \sum_{i=1}^{k} \theta_i} = \frac{\left(\sum_{i=1}^{k} b_i\right)^2}{\left(\sum_{i=1}^{k} b_i\right)^2 + \sum_{i=1}^{k} \theta_i} \qquad (83)$$

เมื่อ $\theta_i = \mathrm{Var}(E_i)$ แทนความแปรปรวนของความคลาดเคลื่อน(i=1, 2,…,k) และ k คือจำนวนข้อสอบ

โดยในการศึกษาครั้งนี้จะใช้สมการที่ (83) มาใช้ในการประมาณค่าความเชื่อมั่นของโมเดลการวัดคะแนนจริงสัมพันธ์ เรียกว่า สัมประสิทธิ์โครงสร้าง

ทฤษฎีการตอบข้อสอบ (Item Response Theory : IRT)

ข้อตกลงเบื้องต้น

ทฤษฎีการตอบข้อสอบ มีข้อตกลงเบื้องต้น ดังนี้

1. ความเป็นมิติเดียว (Unidimensionality)

แบบทดสอบวัดคุณลักษณะเด่นเพียงลักษณะเดียว ถ้าความสามารถของคนมีอยู่ทั้งหมด k ด้าน ซึ่งความสามารถแต่ละด้านต่างก็ส่งผลต่อการตอบข้อสอบข้อต่าง ๆที่รวมกันเป็นแบบทดสอบ ถ้าผลที่ได้จากการตอบแบบทดสอบ (Test Performance) หรือคะแนนของผู้ตอบนั้นสามารถอธิบายได้เพียงความสามารถเดียว ก็จะเรียกว่ามีความเป็นมิติเดียว

2. ความเป็นอิสระในการตอบข้อสอบ (Local Independent)

หมายถึงความน่าจะเป็น หรือโอกาส (Probability) ในการตอบข้อสอบแต่ละข้อได้ถูกต้องเป็นอิสระจากกัน นั่นคือ การตอบข้อสอบข้อใดข้อหนึ่งได้ถูกหรือผิด จะไม่มีผลกระทบต่อการตอบข้ออื่นๆด้วย หรือจะกล่าวในเชิงคณิตศาสตร์ได้ว่า ความเป็นอิสระในการตอบข้อสอบ หมายถึง ความน่าจะเป็นหรือโอกาสที่จะตอบข้อสอบถูกทั้งหมดมีค่าเท่ากับผลคูณของความน่าจะเป็นหรือโอกาสที่จะตอบข้อสอบถูกเป็นรายข้อ

อย่างไรก็ตาม แฮมเบิลตันและสวามินาธาน (Hambleton and Swaminathan. 1985) กล่าวว่า ถ้าแบบทดสอบมีความเป็นมิติเดียวกันแล้ว ความเป็นอิสระในการตอบข้อสอบก็จะเกิดขึ้นตามไปด้วย

3. โค้งคุณลักษณะข้อสอบ (Item Characteristic Curve)

เป็นฟังก์ชั่นคณิตศาสตร์ สามารถใช้อธิบายความสัมพันธ์ระหว่างความน่าจะเป็นหรือโอกาสที่ผู้ตอบข้อสอบได้ถูก กับระดับความสามารถที่วัดได้โดยใช้ชุดของข้อสอบหรือแบบสอบทั้งฉบับนั้น ทั้งนี้ความน่าจะเป็นหรือโอกาสในการตอบข้อสอบถูกจะขึ้นอยู่กับโค้งคุณลักษณะข้อสอบในแต่ละโมเดลที่เลือกใช้ โดยมีรูปร่าง (Shape) ของโค้งคุณลักษณะข้อสอบในแต่ละข้อมีคุณสมบัติไม่แปรเปลี่ยน (Invariant) ไปตามกลุ่มตัวอย่างที่ใช้ ดังนั้นจึงทำให้ความน่าจะเป็นหรือโอกาสในการตอบข้อสอบถูกในแต่ละข้อไม่แปรเปลี่ยนไปด้วย คุณสมบัตินี้ถือเป็นลักษณะเด่นของโมเดลต่าง ๆ ในทฤษฎีการตอบข้อสอบ โค้งคุณลักษณะข้อสอบมีหลายรูปแบบขึ้นอยู่กับว่าเลือกใช้พารามิเตอร์ของข้อสอบกี่ตัว

4. ข้อสอบต้องไม่เป็นข้อสอบประเภทความเร็ว

ฟังก์ชันสารสนเทศ (Information Function)

ฟังก์ชันสารสนเทศ จะเกี่ยวข้องกับการคัดเลือกข้อสอบ การพัฒนาแบบทดสอบ และการประเมินความแม่นยำของการประมาณค่าความสามารถของผู้สอบ ได้แก่

1. ฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function: IIF)

ฟังก์ชันสารสนเทศของข้อสอบ เป็นความสัมพันธ์ของอัตราส่วนระหว่างกำลังสองค่าอนุพันธ์ของโอกาสในการตอบถูกของผู้สอบที่ระดับความสามารถนั้นๆ กับผลคูณของโอกาสในการตอบถูกและผิดของผู้สอบในระดับความสามารถนั้นๆ เขียนเป็นสมการได้ดังนี้

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

เมื่อ $I_i(\theta)$ หมายถึง สารสนเทศของข้อสอบข้อที่ I ที่ระดับความสามารถ θ

$P_i^{'}(\theta)$ หมายถึงค่าอนุพันธ์(Derivative)ของโอกาสในการตอบข้อสอบข้อที่ i ถูก

$P_i(\theta)$ หมายถึง โอกาสในการตอบข้อสอบข้อที่ i ถูกที่ระดับความสามารถ θ

$Q_i(\theta)$ หมายถึง โอกาสในการตอบข้อสอบข้อที่ i ผิดที่ระดับความสามารถ θ

2. ฟังก์ชันสารสนเทศของแบบสอบ (Test Information Function: TIF)

ฟังก์ชันสารสนเทศของแบบสอบ เป็นผลรวมของค่าฟังก์ชันสารสนเทศของข้อสอบ

$$I(\theta) = \sum_{j=1}^{n} I_i(\theta)$$

โมเดลโลจีสติก 3 พารามิเตอร์

ในปี คศ. 1968 เบิร์นบอม (Birnbaum, 1968) ได้พัฒนาโมเดลโลจีสติก 2 พารามิเตอร์ ซึ่งเป็นฟังก์ชั่นที่สามารถคำนวณได้ง่ายกว่าฟังก์ชั่นนอร์มัลโอไจฟ์ แล้วในปี คศ.1974 ลอร์ด (Lord, 1974) ได้เสนอโมเดลโลจีสติก 3 พารามิเตอร์ โดยเพิ่มพารามิเตอร์โอกาสในการเดาข้อสอบ (c_i) ในการประมาณค่าพารามิเตอร์ของผู้สอบด้วย ดังนี้

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

เมื่อ $P_i(\theta)$ คือ โอกาสของผู้สอบที่มีความสามารถระดับ θ ทำข้อสอบข้อที่ i ได้ถูกต้อง

c_i คือ ค่าความน่าจะเป็นในการเดาถูก

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i คือ ค่าความยากของข้อสอบข้อที่ i

e คือ ค่าคงที่ของลอการิทึมธรรมชาติ ซึ่งมีค่าประมาณ 2.71828

D คือ ค่าองค์ประกอบของการปรับสเกล ซึ่งมีค่าเท่ากับ 1.70

การประมาณค่าความเชื่อมั่นจากโมเดลโลจีสติก 3 พารามิเตอร์

ดิมิทรอฟ(Dimitrov. 2003: 440-458) ได้เสนอวิธีการประมาณค่าความเชื่อมั่นจากค่าพารามิเตอร์ของแบบทดสอบที่มีการให้คะแนนแบบ 2 ค่า คือ 0 กับ 1 โดยได้เสนอในรูปของ 2 พารามิเตอร์ และ 3 พารามิเตอร์ โดยสามารถหาค่าคาดหวังของคะแนนข้อสอบ(Expected Item Score) ของ 2 พารามิเตอร์ได้จาก

$$\pi_i = \frac{1-erf(X_i)}{2} \qquad (84)$$

เมื่อ

$$X_i = \frac{a_i b_i}{\sqrt{2(1+a_i^2)}}$$

erf แทนฟังก์ชั่นทางคณิตศาสตร์ เรียกว่าฟังก์ชั่นความคลาดเคลื่อน โดยหาก $X_i > 0$ สามารถคำนวณ erf ได้จาก

$$\text{erf}(X) = 1 - (1+m_1X + m_2X^2 + m_3X^3 + m_4X^4)^{-4} \qquad (85)$$

เมื่อ $m_1 = .278393$

$m_2 = .230389$

$m_3 = .000972$

m_4 = .078108

แต่ถ้า $X_i < 0$ สามารถที่จะใช้ erf(-X) = -erf(X)

กรณีของ 3 พารามิเตอร์สามารถคำนวณจาก

$$\pi_i^* = c_i + (1 - c_i)\pi_i \quad (86)$$

โดยที่ π_i มาจากสมการ (25) จึงสามารถเขียนสมการที่(27)ใหม่ได้ดังนี้

$$\pi_i^* = c_i + (1 - c_i)\left[\frac{1 - erf(X_i)}{2}\right] \quad (87)$$

สามารถคำนวณค่าคาดหวังของความแปรปรวนคลาดเคลื่อนของข้อสอบ(Expectrd Item Error Variance) ของ 2 พารามิเตอร์ได้จาก

$$\sigma^2(e_i) = m_i \exp\left[-0.5\left(\frac{b_i}{d_i}\right)^2\right] \quad (88)$$

เมื่อ b_i แทนความยาก และ m_i กับ d_i เป็นค่าจำแนกข้อสอบที่ได้จาก

$m_i = 0.2646 – 0.118a_i + 0.0187\, a_i^2$

$d_i = 0.7427 + 0.7081/a_i + 0.0074/\, a_i^2$

กรณีของ 3 พารามิเตอร์สามารถคำนวณจาก 2 พารามิเตอร์เป็น 3 พารามิเตอร์ได้ดังนี้

$$\sigma^2(e_i^*) = c_i(1 - c_i)(1 - \pi_i) + (1 - c_i)^2\sigma^2(e_i) \quad (89)$$

หรือสามารถแทนค่าจากสมการ(84)และสมการ(89) มาเป็น 3 พารามิเตอร์ก็จะได้

$$\sigma^2(e_i^*) = c_i(1 - c_i)\left(1 - \left[\frac{1 - erf(X_i)}{2}\right]\right) + (1 - c_i)^2\left(m_i \exp\left[-0.5\left(\frac{b_i}{d_i}\right)^2\right]\right) \quad (90)$$

โดยสามารถคำนวณค่าคาดหวังของความแปรปรวนของคะแนนจริงของข้อสอบ(Expectrd Item True Variance) ของ 2 พารามิเตอร์ได้จาก

$$\sigma^2(\tau_i) = \pi_i(1-\pi_i) - \sigma^2(e_i) \qquad (91)$$

โดยที่ π_i สามารถหาได้จากสมการ (84)
ส่วนกรณีของ 3 พารามิเตอร์สามารถคำนวณได้จาก

$$\sigma^2(\tau_i^*) = \pi_i^*(1-\pi_i^*) - \sigma^2(e_i^*) \qquad (92)$$

ค่าความเชื่อมั่นรายข้อของ 2 พารามิเตอร์หาได้จากสมการ(91)และสมการ(88)ดังนี้

$$\rho_{ii} = \frac{\sigma^2(\tau_i)}{\sigma^2(\tau_i) + \sigma^2(e_i)} \qquad (93)$$

ค่าความเชื่อมั่นรายข้อของ 3 พารามิเตอร์หาได้จากสมการ(92)และสมการ(89)ดังนี้

$$\rho_{ii}^* = \frac{\sigma^2(\tau_i^*)}{\sigma^2(\tau_i^*) + \sigma^2(e_i^*)} \qquad (94)$$

กำหนดให้ค่าคาดหวังของคะแนนข้อสอบ(Expected Item Score: π_i จากสมการ(84)ในกรณี 2 พารามิเตอร์ และ π_i^* จากสมการ(86),(88)ในกรณี 3 พารามิเตอร์)ของข้อสอบแต่ละข้อ ในแบบทดสอบที่ให้คะแนนแบบ 2 ค่า จำนวน n ข้อ ค่าคาดหวังของคะแนนข้อที่ตอบถูก(Expected number-right Score) ของแบบทดสอบ ในกรณี 2 พารามิเตอร์ คือ

$$\mu = \sum_{i=1}^{n} \pi_i \qquad (95)$$

ในกรณี 3 พารามิเตอร์ คือ

$$\mu = \sum_{i=1}^{n} \pi_i^* \qquad (37)$$

กำหนดให้ค่าคาดหวังของความแปรปรวนของความคลาดเคลื่อน(Expected Error Variance: $\sigma^2(e_i)$ จากสมการ(88)ในกรณี 2 พารามิเตอร์ และ $\sigma^2(e_i^2)$ จากสมการ(89),(90)ในกรณี 3 พารามิเตอร์)ของข้อสอบแต่ละข้อ ในแบบทดสอบที่ให้คะแนนแบบ 2 ค่า จำนวน n ข้อ ค่าคาดหวังของความแปรปรวนของความคลาดเคลื่อนสำหรับคะแนนข้อที่ตอบถูกของแบบทดสอบ ในกรณี 2 พารามิเตอร์คือ

$$\sigma_e^2 = \sum_{i=1}^{n} \sigma^2(e_i) \tag{97}$$

ในกรณี 3 พารามิเตอร์ คือ

$$\sigma_e^2 = \sum_{i=1}^{n} \sigma^2(e_i^*) \tag{98}$$

ค่าคาดหวังของความแปรปรวนของคะแนนจริงสำหรับคะแนนข้อที่ตอบถูกของแบบทดสอบ แบบทดสอบจะมีความเป็นเอกพันธ์หรือวัดเนื้อหาเดียวกัน มีความสัมพันธ์สมบูรณ์แบบ(Perfect Correlation)ระหว่างคะแนนจริงสัมพันธ์(τ_i และ τ_j)ของทุกๆ 2 ข้อ i และ j เพราะมีความสัมพันธ์เป็นเชิงเส้นตามโมเดลคะแนนจริงสัมพันธ์คือ $\tau_i = a_{ij}\tau_j + b_{ij}$ (*Jöreskog* .1971 :107-112) เมื่อ b_{ij} เป็น 0,1 ดังนั้นความแปรปรวนร่วมของ τ_i และ τ_j คือ $\sigma(\tau_i, \tau_j) = \sigma(\tau_i)\sigma(\tau_j)$ ความแปรปรวนของคะแนนจริงของข้อที่ตอบถูกของแบบทดสอบที่ให้คะแนนแบบ 2 ค่า จำนวน n ข้อ $\tau = \sum \tau_i \ (i = 1, \ldots, n)$ สามารถหาได้จาก

$$\sigma_\tau^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma(\tau_i, \tau_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma(\tau_i)\sigma(\tau_j) \tag{99}$$

โดยแทนที่ $\sigma(\tau_i)$และ $\sigma(\tau_j)$ในด้านขวาของสมการ(99)ทีเดียวจากสมการ(91) ซึ่งจะทำการแทนค่าจำนวน n ครั้งตามจำนวนข้อสอบ ก็จะได้สมการในกรณีของ 2 พารามิเตอร์ดังนี้

$$\sigma_\tau^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \sqrt{\left[\pi_i(1-\pi_i) - \sigma^2(e_i)\right]\left[\pi_j(1-\pi_j) - \sigma^2(e_j)\right]} \tag{100}$$

เมื่อ π_i และ $\sigma^2(e_i)$ (หรือ π_j และ $\sigma^2(e_j)$) ได้มาจากสมการ(84) และสมการ(88) ตามลำดับ

กรณี 3 พารามิเตอร์ก็จะได้สมการ

$$\sigma_\tau^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \sqrt{\left[\pi_i^*(1-\pi_i^*) - \sigma^2(e_i^*)\right]\left[\pi_j^*(1-\pi_j^*) - \sigma^2(e_j^*)\right]} \tag{101}$$

เมื่อ π_i และ $\sigma^2(e_i)$ (หรือ π_j และ $\sigma^2(e_j)$) ได้มาจากสมการ(84) และสมการ(29)ตามลำดับ

ภายใต้โมเดลคะแนนจริงของลอร์ดและโนวิค(Lord & Novick, 1968) ค่าความเชื่อมั่นของแบบทดสอบก็จะได้เท่ากับ

$$\rho_{IRT} = \frac{\rho_{\tau}^{2}}{\rho_{\tau}^{2} + \rho_{e}^{2}} \qquad (102)$$

เมื่อ กรณีของ 2 พารามิเตอร์ ρ_{τ}^{2} ได้จากสมการ(100) ρ_{e}^{2} ได้จากสมการ(97)
กรณีของ 3 พารามิเตอร์ ρ_{τ}^{2} ได้จากสมการ(101) ρ_{e}^{2} ได้จากสมการ(98)

การวิจัยครั้งนี้จะใช้สมการ(102) กรณี 3 พารามิเตอร์ มาใช้ในการประมาณค่าความเชื่อมั่น

คำอธิบายรายวิชา คณิตศาสตร์ชั้นประถมศึกษาปีที่ 5

ศึกษา ฝึกทักษะการคิดคำนวณ และฝึกการแก้ปัญหาในสาระต่อไปนี้

จำนวนนับ การอ่านและการเขียนตัวหนังสือ ตัวเลขฮินดูอารบิก ตัวเลขไทยแทนจำนวน ชื่อหลัก ค่าของตัวเลขในแต่ละหลัก การเขียนในรูปกระจาย การเรียงลำดับจำนวน การประมาณค่าใกล้เคียงเป็นจำนวนเต็มสิบ เต็มร้อย เต็มพัน สมบัติการสลับที่และสมบัติการเปลี่ยนหมู่ของการบวก สมบัติการสลับที่และสมบัติการเปลี่ยนหมู่ของการคูณ สมบัติการแจกแจง

การบวก การลบ การคูณ การหารจำนวนนับ และโจทย์ปัญหา การบวก การลบ การคูณ และการหารจำนวนนับ การบวก ลบ คูณ หารระคน โจทย์ปัญหา

เศษส่วน เศษส่วน เศษเกิน จำนวนคละ เศษส่วนของจำนวนนับ เศษส่วนที่เท่ากัน เศษส่วนอย่างต่ำ การเปรียบเทียบเศษส่วนที่มีตัวส่วนเป็นพหุคูณของกันและกัน การเรียงลำดับเศษส่วน

การบวก การลบ การคูณ การหารเศษส่วน และโจทย์ปัญหา การบวกและการลบเศษส่วนที่มีตัวส่วนเป็นพหุคูณของกันและกัน การคูณและการหารเศษส่วน การบวก ลบ คูณ เศษส่วนระคน โจทย์ปัญหา

ทศนิยม การอ่านและการเขียนทศนิยมไม่เกินสองตำแหน่ง หลักและค่าประจำหลัก การเขียนในรูปกระจาย การเปรียบเทียบและเรียงลำดับทศนิยม การเขียนทศนิยมไม่เกินสองตำแหน่งให้อยู่ในรูปเศษส่วนและเขียนเศษส่วนที่มีตัวส่วนเป็น 10 หรือ 100 ให้อยู่ในรูปทศนิยม

การบวก การลบ การคูณทศนิยม และโจทย์ปัญหา การบวกและการลบทศนิยมไม่เกินสองตำแหน่ง การคูณทศนิยมที่มีผลคูณเป็นทศนิยมไม่เกินสองตำแหน่งการบวก ลบ คูณทศนิยมระคนที่ผลลัพธ์เป็นทศนิยมไม่เกินสองตำแหน่ง โจทย์ปัญหา

ร้อยละ และโจทย์ปัญหา การเขียนเศษส่วนที่ตัวส่วนเป็นตัวประกอบของ 100 ให้อยู่ในรูปร้อยละ การเขียนร้อยละให้อยู่ในรูปเศษส่วนและทศนิยม การเปรียบเทียบเศษส่วน ทศนิยม และร้อยละ ร้อยละของจำนวนนับ โจทย์ปัญหาร้อยละที่มีผลลัพธ์เป็นจำนวนนับ

การประมาณค่าจำนวนนับ การประมาณค่าใกล้เคียงเป็นจำนวนเต็มสิบ เต็มร้อย และเต็มพัน

การหาความยาว ความยาวของเส้นรอบรูปของรูปสามเหลี่ยมและรูปสี่เหลี่ยม โจทย์ปัญหาและสถานการณ์

การหาพื้นที่ การหาพันที่ของรูปสามเหลี่ยมและรูปสี่เหลี่ยมมุมฉาก การคาดคะเนพื้นที่เป็นตารางเมตร ตารางเซนติเมตร และตารางวา โจทย์ปัญหาและสถานการณ์

การหาปริมาตร การหาปริมาตรและ / หรือความจุของทรงสี่เหลี่ยมมุมฉาก

รูปเรขาคณิตและสมบัติบางประการของรูปเรขาคณิต

- **มุม** จุดยอดมุม แขนของมุม การเรียกชื่อมุม การเขียนสัญลักษณ์แทนมุม ชนิดของมุม การวัดขนาดของมุมเป็นองศา การสร้างมุมโดยใช้ไม้โพรแทรกเตอร์(ครึ่งวงกลม
- **รูปสี่เหลี่ยม** รูปสี่เหลี่ยมจัตุรัส รูปสี่เหลี่ยมผืนผ้า รูปสี่เหลี่ยมด้านขนาน รูปสี่เหลี่ยมขนมเปียกปูน รูปสี่เหลี่ยมคางหมู รูปสี่เหลี่ยมรูปว่าว การสร้างรูปสี่เหลี่ยมมุมฉาก
- **รูปสามเหลี่ยม** รูปสามเหลี่ยมด้านเท่า รูปสามเหลี่ยมหน้าจั่ว รูปสามเหลี่ยมด้านไม่เท่า รูปสามเหลี่ยมมุมฉาก รูปสามเหลี่ยมมุมแหลม รูปสามเหลี่ยมมุมป้าน ส่วนประกอบของรูปสามเหลี่ยม ขนาดของมุมภายใน การสร้างรูปสามเหลี่ยม
- **รูปวงกลม** ส่วนประกอบของรูปวงกลม การสร้างรูปวงกลม
- **การประดิษฐ์ลวดลายโดยใช้รูปเรขาคณิต**
- **เส้นขนาน** เส้นขนานและการใช้สัญลักษณ์ // แสดงการขนาน การสร้างเส้นขนาน
- **ทรงสี่เหลี่ยมมุมฉาก ทรงกลม ทรงกระบอก กรวย ปริซึม และพีระมิด**

แบบรูปและความสัมพันธ์ แบบรูปของจำนวน การเขียนประโยคสัญลักษณ์แสดงความสัมพันธ์ของสถานการณ์หรือปัญหา

สถิติและความน่าจะเป็นเบื้องต้น การอ่านแผนภูมิแท่งและแผนภูมิแท่งเปรียบเทียบ การเก็บรวบรวมข้อมูลและการเขียนแผนภูมิแท่ง ความหมายและการนำไปใช้ในชีวิตประจำวันของเหตุการณ์ที่เกิดขึ้นแน่นอน อาจจะเกิดขึ้นหรือไม่เกิดขึ้น และไม่เกิดขึ้นอย่างแน่นอน

การจัดประสบการณ์หรือสร้างสถานการณ์ที่ใกล้ตัวให้ผู้เรียนได้ศึกษาค้นคว้าโดยปฏิบัติจริง ทดลอง สรุป รายงาน เพื่อพัฒนาทักษะ / กระบวนการในการคิดคำนวณ การแก้ปัญหา การให้เหตุผล การสื่อความหมายทางคณิตศาสตร์ และนำประสบการณ์ด้านความรู้ ความคิด ทักษะกระบวนการที่ได้ไปใช้ในการเรียนรู้สิ่งต่างๆ และใช้ในชีวิตประจำวันอย่างสร้างสรรค์ รวมทั้งเห็นคุณค่าและมีเจตคติที่ดีต่อคณิตศาสตร์ สามารถทำงานอย่างเป็นระบบระเบียบ รอบคอบ มีความรับผิดชอบ มีวิจารณญาณและเชื่อมั่นในตนเอง

การวัดและประเมินผล ใช้วิธีการหลากหลายตามสภาพความเป็นจริงของเนื้อหา และทักษะที่ต้องการวัด

งานวิจัยที่เกี่ยวข้อง

สำหรับงานวิจัยที่มีการศึกษาเกี่ยวกับการประมาณค่าความเชื่อมั่นของแบบทดสอบและมีส่วนหนึ่งเกี่ยวของกับสูตรทั้ง 4 สูตร ผู้วิจัยได้ศึกษางานวิจัยที่เกี่ยวข้อง แยกตามสูตรการคำนวณคือ สัมประสิทธิ์ KR-20 ทฤษฎีคะแนนจริงสัมพันธ์ สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และทฤษฎีการทดสอบมาตรฐานเดิมกับทฤษฎีการตอบข้อสอบ สัมประสิทธิ์ ρ_{IRT}

งานวิจัยที่เกี่ยวข้องกับสัมประสิทธิ์ KR-20 จากส่วนหนึ่งของงานวิจัย ซึ่งมีหลายคนศึกษา อาทิ

ฉัตรศิริ ปิยะพิมลสิทธิ์ (2540: 42-43) ได้ทำการศึกษาการเปรียบเทียบความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบเลือกตอบความสามารถในการอ่านภาษาไทยที่มีรูปแบบต่างกัน กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนสังกัดกรมสามัญศึกษา กรุงเทพมหานคร กลุ่ม 5 จำนวน 10 โรงเรียน เป็นนักเรียนทั้งสิ้น 2,366 คน เครื่องมือที่ใช้เป็นแบบทดสอบเลือกตอบความสามารถในการอ่านภาษาไทยชนิด 5 ตัวเลือก คือ แบบตัวเลือกซ้อน แบบตัวเลือกถูกผิดที่ตรวจให้คะแนนด้วยวิธี 0-1 แบบตัวเลือกธรรมดา หาค่าความเชื่อมั่นโดยใช้สูตร KR-20 และแบบตัวเลือกถูกผิดที่ตรวจให้คะแนนทุกตัวเลือก หาค่าความเชื่อมั่นโดยใช้สูตรสหสัมพันธ์แบบเพียร์สัน ผลการวิจัยพบว่า ค่าความเชื่อมั่นของแบบทดสอบเลือกตอบแบบตัวเลือกธรรมดา มีค่าเท่ากับ 0.759 แบบตัวเลือกถูกผิดที่ตรวจให้คะแนนด้วยวิธี 0-1 มีค่าเท่ากับ 0.780 แบบตัวเลือกถูกผิดที่ตรวจให้คะแนนทุกตัวเลือก มีค่าเท่ากับ 0.743 และแบบตัวเลือกซ้อน มีค่าเท่ากับ 0.813 โดยแบบตัวเลือกซ้อนมีค่าความเชื่อมั่นสูงสุด รองลงมาคือ แบบตัวเลือกถูกผิดที่ตรวจให้คะแนนด้วยวิธี 0-1 แบบตัวเลือกธรรมดา และแบบตัวเลือกถูกผิดที่ตรวจให้คะแนนทุกตัวเลือกมีค่าความเชื่อมั่นต่ำสุด

ชุติมา สุขสว่าง (2445: 50-51) ได้ทำการศึกษาการเปรียบเทียบความเชื่อมั่นของแบบทดสอบความสามารถด้านผลการคิดเอกนัยที่มีเนื้อหาและจำนวนตัวเลือกต่างกัน กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นประถมศึกษาปีที่ 4 ของปีการศึกษา 2544 ในโรงเรียนสังกัดสำนักงานคณะกรรมการการประถมศึกษา จังหวัด ชลบุรี จำนวน 1,218 คน เครื่องมือที่ใช้เป็นแบบทดสอบเลือกตอบวัดความสามารถด้านผลการคิดเอกนัยที่มีเนื้อหาเป็นภาษาและเป็นรูปภาพชนิดเลือกตอบ 5 4 และ 3 ตัวเลือก ผลการวิจัยพบว่า ความเชื่อมั่นของแบบทดสอบวัดความสามารถด้านการคิดเอกนัยทางภาษาที่มีผลการคิดแบบหน่วย แบบกลุ่มและแบบความสัมพันธ์ ชนิดเลือกตอบ 5 4 และ3 ตัวเลือก มีค่าตั้งแต่ 0.573 ถึง 0.829 โดยแบบทดสอบวัดความสามารถด้านการคิดเอกนัยทางภาษาที่มีผลการคิดแบบกลุ่ม ชนิดเลือกตอบ 5 ตัวเลือก มีความเชื่อมั่นสูงสุด และ

แบบทดสอบวัดความสามารถด้านการคิดเอกนัยทางภาษาที่มีผลการคิดแบบหน่วย ชนิดเลือกตอบ 3 ตัวเลือก มีค่าความเชื่อมั่นต่ำสุด และเมื่อพิจารณาความเชื่อมั่นของแบบทดสอบวัดความสามารถด้านการคิดเอกนัยทางรูปภาพที่มีผลการคิดแบบหน่วย แบบกลุ่ม และแบบความสัมพันธ์ชนิดเลือกตอบ 5 4 และ3 ตัวเลือก มีค่าตั้งแต่ 0.564 ถึง 0.860 โดยแบบทดสอบวัดความสามารถด้านการคิดเอกนัยทางรูปภาพที่มีผลการคิดแบบกลุ่ม ชนิดเลือกตอบ 4 ตัวเลือก มีความเชื่อมั่นสูงสุด และแบบทดสอบวัดความสามารถด้านการคิดเอกนัยทางรูปภาพที่มีผลการคิดแบบหน่วย ชนิดเลือกตอบ 3 ตัวเลือก มีค่าความเชื่อมั่นต่ำสุด

บังอร กมลวัฒนา (2542: 54) ได้ทำการศึกษาการเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบและการจัดเรียงปัญหาต่างกัน กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2541 ของโรงเรียนในสังกัดคณะกรรมการการศึกษาเอกชน กรุงเทพมหานคร จำนวน 1,600 คน ได้มาโดยวิธีการแบ่งแบบแบ่งชั้น เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเป็นแบบทดสอบวัดความสามารถในการแก้ปัญหาตามกระบวนการแก้ปัญหาของเวียร์ ซึ่งผู้วิจัยสร้างขึ้นจำนวน 4 ฉบับ หาค่าความเชื่อมั่นโดยใช้สูตร KR-20 ผลการวิจัยพบว่า แบบทดสอบวัดความสามารถในการแก้ปัญหาแบบเลือกตอบที่มีการจัดเรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก และโดยวิธีการสุ่มค่ามีค่าความเชื่อมั่นเท่ากับ 0.788 และ 0.793 ตามลำดับ แบบทดสอบวัดความสามารถในการแก้ปัญหาแบบตอบสั้นที่มีการจัดเรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก และโดยวิธีการสุ่มมีค่าความเชื่อมั่นเท่ากับ 0.804 และ 0.816 ตามลำดับ

บุญเชิด ภิญโญอนันตพงษ์ (2542: บทคัดย่อ) ได้ทำการศึกษาสัมประสิทธิ์ r_B : การประมาณค่าความเชื่อมั่นสำหรับแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยความยากรายข้อต่างกัน เพื่อพัฒนาและศึกษาผลการใช้สูตร r_B ที่ผู้วิจัยพัฒนาขึ้นมาสำหรับประมาณค่าความเชื่อมั่นของแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยข้อสอบที่ให้คะแนนระบบ 0, 1 และมีความยาดรายข้อต่างกัน โดยศึกษาจากแบบทดสอบที่วัดเนื้อหาเดียวกัน จำนวนข้อต่างกัน จำนวน 6 ฉบับ คือ ฉบับ 25, 30, 35, 40, 45, และ 50 ข้อ ทั้งจากกลุ่มตัวอย่างขนาดใหญ่และกลุ่มตัวอย่างขนาดเล็ก คือ 30, 50 และ80 คนว่าเป็นอย่างไรเมื่อเทียบกับสูตร KR-20 ผลการศึกษาพบว่า สูตร r_B คำนวณค่าความเชื่อมั่นของแบบทดสอบจำนวนข้อต่างกันจากกลุ่มตัวอย่างขนาดใหญ่ได้สูงสอดคล้องสัมพันธ์กับสูตร KR-20 โดยสูตร r_B คำนวณค่าความเชื่อมั่นได้สูงกว่าสูตร KR-20 ทุกฉบับ เมื่อศึกษาจากกลุ่มตัวอย่างขนาดเล็ก พบว่า สูตร r_B คำนวณจากแบบทดสอบฉบับเดียวกันและกลุ่มตัวอย่างขนาดเดียวกันได้สอดคล้องสัมพันธ์กับสูตร KR-20 จากแบบทดสอบทั้ง 6 ฉบับและกลุ่มตัวอย่างทุกขนาด สูตร r_B คำนวณค่าความเชื่อมั่นสูงกว่าสูตร KR-20 ส่วนจำนวนข้อสอบและขนาดกลุ่มตัวอย่างมีผลต่อค่าความเชื่อมั่นที่คำนวณได้จากทั้งสองสูตรในลักษณะคล้ายคลึงกัน และทั้งสองสูตรประมาณค่าความเชื่อมั่นจากกลุ่มตัวอย่างขนาดเล็กเทียบกับกลุ่มตัวอย่างขนาดใหญ่มีค่าความ

ลำเอียงในทิศทางเดียวกัน จึงสรุปได้ว่า สูตร r_B ที่พัฒนาขึ้นครั้งนี้ สามารถนำไปใช้ในการประมาณค่าความเชื่อมั่นของแบบทดสอบได้ดีและมีค่าความเชื่อมั่นสูงกว่าสูตร KR-20

โคเลน แฮนสัน และเบรนนัน (Kolen, Hanson and Brennan. 1992: 285-307) ได้ศึกษาผลของการแปลงคะแนน จำนวนค่าของคะแนนแปลง และการเทียบคะแนนที่มีต่อการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ใช้โมเดลของคะแนนจริงแบบเข้มงวด และหาค่าความเชื่อมั่นโดยใช้สูตร KR-20 โดยใช้การแปลงคะแนน 4 แบบ ได้แก่ Normalization, ACT Scale, Stannine และแบบ Log-Odds ข้อมูลที่นำมาวิเคราะห์ได้จากแบบทดสอบ ACT Assessment ของปี ค.ศ.1988 ที่เป็นการสอบคัดเลือกเข้าวิทยาลัย วิชาภาษาอังกฤษ คณิตศาสตร์ การอ่าน และเหตุผลทางวิทยาศาสตร์ ซึ่งเป็นแบบทดสอบชนิดเลือกตอบ ผลการวิจัยพบว่า ความเชื่อมั่นของคะแนนแปลงรูปแบบ Normalization มีค่าน้อยกว่าของคะแนนที่แปลงแบบ ACT Scale โดยเฉพาะวิชาคณิตศาสตร์และเหตุผลทางวิทยาศาสตร์ และความเชื่อมั่นของคะแนนที่แปลงแบบ Stannine มีค่าความเชื่อมั่นต่ำกว่าของคะแนนแปลงรูปแบบ Normalization

สุทธิวรรณ พีรศักดิ์โสภณ (2546: 62-68) ได้ทำการศึกษาการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปเมื่อใช้วิธีการประมาณค่าโดยทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด-ควอลส์ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาตอนปลาย ปีการศึกษา 2544 จำนวน 2,000 คน และสุ่มมาเป็นตัวอย่างจำลองสามขนาดคือ 300 500 800 คน แต่ละขนาดได้มาโดยการสุ่มซ้ำ 30 ครั้ง เครื่องมือที่ใช้เป็นแบบทดสอบวัดความถนัดทางการเรียนสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลายจำนวน 3 ฉบับคือ ฉบับด้านตัวเลข ด้านภาษา และด้านเหตุผล ผลการวิจัยพบว่า แบบทดสอบวัดความถนัดทางการเรียนทั้งสามด้านคือด้านตัวเลข ด้านภาษา ด้านเหตุผล จากกลุ่มตัวอย่างขนาดใหญ่ สูตร r_B ให้ค่าความเชื่อมั่นสูงกว่าสูตร KR-20 และ KR-21 จากกลุ่มตัวอย่างจำลองขนาด 300 500 800 คน ค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR-20 และ KR-21

วิชุดา กิจธรธรรม (2544: 37-38) ได้ทำการศึกษาการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนของแบบวัดความถนัดทางวิชาชีพพยาบาลตามแนวคิดของฟลานาแกนด้วยวิธีการวิเคราะห์ความแปรปรวน วิธีการตามทฤษฎีการตอบข้อสอบ และวิธีการของคีตกลุ่มตัวอย่างที่ใช้เป็นนักศึกษาพยาบาลชั้นปีที่ 4 ปีการศึกษา 2543 ในสถาบันอุดมศึกษาเอกชนเขตกรุงเทพมหานคร จำนวน 500 คน ได้จากการสุ่มแบบสองขั้นตอน เครื่องมือที่ใช้เป็นแบบทดสอบวัดความถนัดทางวิชาชีพพยาบาลของฟลานาแกน จำนวน 5 ฉบับ ผลการวิจัยพบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดทางวิชาชีพพยาบาลของฟลานาแกน ทั้ง 5 ฉบับ ที่วิเคราะห์โดยใช้สูตร r_B ให้ค่าความเชื่อมั่นสูงกว่าสูตร KR-20 และ KR-21

เรวดี อินทสะระ (2530: 43) ได้ทำการศึกษาการเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตราระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์

องค์ประกอบ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2529 โรงเรียนสังกัดกรมสามัญศึกษา จังหวัดพัทลุง จำนวน 2,823 คน เครื่องมือที่ใช้เป็นแบบทดสอบวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 เรื่องอัตราส่วน ร้อยละ ที่ผู้วิจัยสร้างขึ้นจากการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบข้อสอบ โดยใช้โปรแกรมโลจีส 5 แล้วจัดข้อสอบออกเป็น 2 ฉบับๆ ละ 30 ข้อ ผลการวิจัยพบว่า ค่าความเชื่อมั่นที่คำนวณจากสูตร KR-20 ของแบบทดสอบฉบับที่ 1 มีค่าเท่ากับ 0.8782 ฉบับที่ 2 มีค่าเท่ากับ 0.8980

ออสเตอร์ฮอพ และพาเมลา (Oosterhof and Pamela. 1984: 287-293) ได้ทำการศึกษาการเปรียบเทียบคุณภาพของแบบทดสอบเลือกตอบที่มีตัวเลือกแตกต่างกัน 3 ชุด กับแบบทดสอบเติมคำ(Completion) 1 ชุด ในวิชาคณิตศาสตร์ที่เป็นโจทย์ปัญหา แบบทดสอบเลือกตอบที่มีรูปแบบของตัวเลือกตอบที่มีรูปแบบของตัวเลือกตอบแตกต่างกัน 3 รูปแบบ คือ แบบตัวเลือกเดี่ยว หรือแบบตัวเลือกธรรมดา แบบตัวเลือกปลายเปิด และแบบตัวเลือกเป็นช่วง กลุ่มตัวอย่างเป็นนักศึกษาธุรกิจศึกษา จำนวน 232 คน ผลการวิจัยพบว่า ค่าความยากของแบบทดสอบเลือกตอบทั้ง 3 ชุด แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนแบบทดสอบเติมคำยากกว่าแบบทดสอบเลือกตอบทั้ง 3 ชุด อย่างมีนัยสำคัญทางสถิติที่ระดับ .02 ค่าความเชื่อมั่นของแบบทดสอบทั้ง 3 ชุด ที่คำนวณโดยใช้สูตร KR-20 พบว่า แบบทดสอบแบบเติมคำมีค่าความเชื่อมั่นสูง รองลงมาคือชนิดตัวเลือกปลายเปิด และชนิดตัวเลือกเป็นช่วง

ราชู (Raju. 1977: 549-565) ได้พัฒนาสัมประสิทธิ์แอลฟ่า r_c ของครอนบัค(Cronbach. 1951) ให้สามารถประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนด้วยความยาวไม่เท่ากัน ราชู เรียกว่า สัมประสิทธิ์เบต้า(βk Coefficient) ใช้สัญลักษณ์ r_R และได้ทำการทดลองเปรียบเทียบสัมประสิทธิ์ r_R กับสัมประสิทธิ์ r_c ของครอนบัค สัมประสิทธิ์ r_{FS} ของเฟลดต์(Feldt. 1975) สัมประสิทธิ์ r_H ของฮอร์ส(Horse. 1951) สัมประสิทธิ์ r_k ของคริสทอฟ(Kritstof. 1974) โดยศึกษาจากนักเรียนเกรด 6 จำนวน 300 คน ที่สอบแบบทดสอบการคิดคำนวณจำนวน 40 ข้อ จากชุดอนุกรมผลสัมฤทธิ์ เอส อาร์ เอ ระดับต้น ในปี 1975 ผลการวิจัยพบว่า ค่าความเชื่อมั่นของสัมประสิทธิ์ KR-20 มีค่าเท่ากับ 0.854 และจากการแบ่งแบบทดสอบออกเป็นหลายๆ ส่วนเป็น 4 แบบต่างกันคือ แบ่งสองส่วน 2 แบบ แบ่งสามส่วน 1 แบบ และแบ่งสี่ส่วน 1 แบบ ในการแบ่งแบบทดสอบแต่ละแบบจะกำหนดแบบทดสอบใส่ในแต่ละส่วนโดยการสุ่ม ผลพบว่า สัมประสิทธิ์ r_R ประมาณค่าได้ดีเมื่อเทียบกับ KR-20 ยกเว้นเฉพาะการแบ่งสองส่วนที่มีขนาด(35, 5) ส่วนสัมประสิทธิ์ r_{FS} r_H r_k ก็ได้ผลดีเช่นเดียวกัน แต่สัมประสิทธิ์ r_c ให้ค่าความเชื่อมั่นต่ำที่สุดในทุกกรณี

งานวิจัยที่เกี่ยวข้องกับทฤษฎีคะแนนจริงสัมพันธ์ จากส่วนหนึ่งของงานวิจัย ซึ่งมีหลายคนศึกษา อาทิ

บุญเชิด ภิญโญอนันตพงษ์ (2537: บทคัดย่อ) ได้ทำการศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์ เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ ที่มีความยาวส่วนย่อยไม่เท่ากัน ที่แบ่งให้มีส่วนย่อยจำนวนไม่เท่ากัน ที่คำนวณได้กับค่าพารามิเตอร์ความเชื่อมั่นที่แท้จริงของแบบทดสอบ ที่คำนวณจากสูตรแต่ละสูตร และที่คำนวณจากกลุ่มตัวอย่างขนาดต่างกัน กลุ่มตัวอย่างเป็นนักเรียนจำนวน 3000 คน สุ่มมาจากนักเรียนที่สอบคัดเลือกเข้าเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2536 ของโรงเรียนต่าง ๆที่ใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนฟอร์มเดียวกัน จำนวน 4 วิชา ของสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ผลการศึกษาพบว่า เมื่อแบ่งเป็น 2 ส่วน คำนวณจากกลุ่มตัวอย่างขนาดใหญ่และเล็ก สูตร r_{L2} และสูตร r_{FK} ประมาณค่าความเชื่อมั่นได้คงที่เป็นส่วนใหญ่มากกว่าสูตร r_R เมื่อแบ่งเป็น 3 ส่วน คำนวณจากกลุ่มตัวอย่างขนาดใหญ่ สูตร r_{L2} และ r_{FK} ประมาณค่าความเชื่อมั่นได้คงที่ไม่แปรเปลี่ยนไปตามความยาวของส่วนย่อย ส่วนสูตร r_{L1} และ r_{F2} โดยเฉลี่ยประมาณค่าความเชื่อมั่นได้คงที่เป็นส่วนใหญ่เกือบทั้งหมด แต่เมื่อคำนวณจากกลุ่มตัวอย่างขนาดเล็ก สูตร r_{L1} และ r_{F2} โดยเฉลี่ยประมาณค่าความเชื่อมั่นได้คงที่ไม่แปรเปลี่ยนไปตามความยาวของส่วนย่อย และเมื่อแบ่งเป็น 4 ส่วน คำนวณจากกลุ่มตัวอย่างขนาดใหญ่ สูตร r_{L2} , r_{FK} ,r_{L1} และ r_{F2} ประมาณค่าความเชื่อมั่นได้คงที่ไม่แปรเปลี่ยนไปตามความยาวของส่วนย่อย ส่วนสูตร r_R ประมาณค่าความเชื่อมั่นได้คงที่เป็นส่วนใหญ่ แต่เมื่อคำนวณจากกลุ่มตัวอย่างขนาดเล็ก สูตร r_{L2} และ r_{FK} โดยเฉลี่ยประมาณค่าความเชื่อมั่นได้คงที่ไม่แปรเปลี่ยนไปตามความยาวของส่วนย่อย ส่วนสูตร r_{L1} และ r_{F2} โดยเฉลี่ยประมาณค่าความเชื่อมั่นได้คงที่เป็นส่วนใหญ่เกือบทั้งหมด

สนอง เที่ยงตรง (2540: 61-63) ได้ศึกษาการแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างและความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ของการทดสอบวัดการประเมินค่าทางสัญลักษณ์และแบบทดสอบวัดการประเมินค่าทางภาษา แบบการแปลงรูปตามโครงสร้างทางสติปัญญาของกิลฟอร์ด กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ของโรงเรียนสังกัดกรมสามัญศึกษาในจังหวัดนนทบุรี จำนวน 610 คนวิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบเชิงสำรวจและวิธีวิเคราะห์องค์ประกอบเชิงยืนยันด้วยวิธีความน่าจะเป็นสูงสุด และคำนวณค่าความเชื่อมั่นด้วยสูตรเฟลดต์-ราชู ผลการศึกษาความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบ จากการวิเคราะห์องค์ประกอบเชิงสำรวจพบว่า แบบทดสอบทั้งสองวัดองค์ประกอบร่วมกันคือ องค์ประกอบการประเมินค่าทางสัญลักษณ์แบบการแปลงรูปและองค์ประกอบการประเมินค่าทางภาษาแบบการแปลงรูป สำหรับการวิเคราะห์เชิงยืนยันด้วยวิธีความน่าจะเป็นสูงสุดนั้นพบว่า แบบทดสอบฉบับถอดรหัส ฉบับการสลับที่ตัวอักษร และฉบับการตัดสินความถูกต้องของนิพจน์ วัดองค์ประกอบการประเมินค่าทางสัญลักษณ์แบบการแปลงรูปอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และแบบทดสอบตัดสินการปรับใช้ประโยชน์ของสิ่งของ และฉบับการเปรียบเทียบประโยคที่ใช้บรรยายภาพ วัดองค์ประกอบ

การประเมินค่าทางภาษาแบบการแปลงรูปอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แบบทดสอบการประเมินค่าทางสัญลักษณ์แบบการแปลงรูปจำนวน 3 ฉบับ มีความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ คือ .740 .749 และ .763 ตามลำดับ แบบทดสอบการประเมินค่าทางภาษาแบบการแปลงรูป จำนวน 2 ฉบับ มีความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ คือ .717 และ .742 ตามลำดับ

นภดล กองศิลป์ (2540: บทคัดย่อ) ได้ทำการศึกษา การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างและความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ของแบบทดสอบวัดความรู้ความเข้าใจทางภาษาแบบประยุกต์ตามโครงสร้างทางสติปัญญาของกิลฟอร์ด กลุ่มตัวอย่าง เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 900 คน วิเคราะห์ข้อมูลโดยใช้วิธีวิเคราะห์องค์ประกอบเชิงสำรวจและวิธีวิเคราะห์องค์ประกอบเชิงยืนยันด้วยวิธีความน่าจะเป็นสูงสุด และคำนวณค่าความเชื่อมั่นด้วยสูตรเฟลดต์-ราชู ผลการศึกษาความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบวัดความรู้ความเข้าใจทางภาษาแบบประยุกต์ทั้ง 5 ฉบับ คือการให้ทางเลือก การทดสอบเครื่องมือเครื่องใช้ การตั้งคำถามให้ตรงกับเนื้อเรื่อง การมองเห็นความบกพร่องและการมองเห็นปัญหา พบว่า แบบทดสอบแต่ละฉบับวัดองค์ประกอบร่วมกัน 1 องค์ประกอบ โดยมีน้ำหนักร่วมกันอยู่ระหว่าง .548 ถึง .713 ความเที่ยงตรงเชิงโครงสร้างที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยวิธีความน่าจะเป็นสูงสุด พบว่า แบบทดสอบทั้ง 5 ฉบับ มีความเที่ยงตรงในการวัดองค์ประกอบความรู้ความเข้าใจทางภาษาแบบประยุกต์ โดยมีน้ำหนักองค์ประกอบที่มีนัยสำคัญทางสถิติที่ระดับ .01 และมีค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์อยู่ระหว่าง .784 ถึง .887

วิศิษฏ์ พหลยุทธ (2539: บทคัดย่อ) ได้ทำการศึกษาเรื่อง การประมาณค่าความเชื่อมั่นของแบบทดสอบที่มีคำถามหลายรูปแบบตามข้อตกลงแบบคู่ขนาน แบบคะแนนจริงสมมูลและแบบคะแนนจริงสัมพันธ์ โดยมีจุดมุ่งหมายเพื่อ 1.หาความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่น ของ สเปียร์แมน ($r_{S\text{-}B}$) ของครอนบัค (r_C) ของครอนบัค โซแมนน์และแม็กกี (α r_{stat}) ของเฟลดต์-ราชู ($r_{F\text{-}R}$) และของเลียว (r_{L1}) 2.เพื่อหาความลำเอียงทางสถิติของการประมาณค่าความเชื่อมั่นของ สเปียร์แมน ($r_{S\text{-}B}$) ของครอนบัค (r_C) ของครอนบัค โซแมนน์และแม็กกี (α r_{stat}) ของเฟลดต์-ราชู ($r_{F\text{-}R}$) และของเลียว (r_{L1}) 3.เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่คำนวณด้วยสูตรของสเปียร์แมน ($r_{S\text{-}B}$) ของครอนบัค (r_C) ของครอนบัค โซแมนน์และแม็กกี (α r_{stat}) ของเฟลดต์-ราชู ($r_{F\text{-}R}$) และของเลียว (r_{L1}) กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 จำนวน 600 คน ผลการศึกษาพบว่า ค่าความคลาดเคลื่อนที่คำนวณด้วยสูตร $r_{F\text{-}R}$ มีค่ามากที่สุด และ $r_{S\text{-}B}$ มีค่าน้อยที่สุด ค่าความลำเอียงทางสถิติที่คำนวณด้วยสูตร r_{L1} มีค่ามากที่สุด และ $r_{S\text{-}B}$ มีค่าน้อยที่สุด ค่าความเชื่อมั่นที่คำนวณด้วยสูตร $r_{S\text{-}B}$ มีค่ามากที่สุด และ r_C มีค่าน้อยที่สุด

สุมนา โสตถิผลอนันต์ (2539: บทคัดย่อ) ได้ศึกษาการเปรียบเทียบวิธีการเทียบคะแนนเส้นตรงตามแบบจำลองคะแนนจริงสัมพันธ์ที่ได้จากคะแนนสอบมาคำนวณด้วยสูตร 4 วิธี (วิธีของแมคแคน 2 วิธี ของวู๊ดดรัฟ 2 วิธี) โดยมีข้อตกลงว่า แบบทดสอบที่นำมาเปรียบเทียบนั้นต้องเป็นไปตามแบบจำลองคะแนนจริงสัมพันธ์และมีความเชื่อมั่นเท่าเทียมกัน กลุ่มตัวอย่างเป็น

นักศึกษาพยาบาลชั้นปีที่ 1 และ 2 หลักสูตรพยาบาลศาสตร์ และพยาบาลศาสตร์ระดับต้น ของวิทยาลัยพยาบาล สังกัดสำนักงานพัฒนากำลังคนด้านสาธารณสุข สำนักงานปลัดกระทรวงสาธารณสุข กระทรวงสาธารณสุข ในเขตภาคกลาง จำนวน 864 คน ผลการศึกษาพบว่า คะแนนแปลงที่ได้จาก 4 วิธีให้ผลแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อเปรียบเทียบผลการใช้แบบทดสอบแบบเชื่อมโยงกับข้อสอบร่วม พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 รวมทั้งผลของขนาดความยาวของแบบทดสอบหลักก็พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในทำนองเดียวกัน ผลของการเปรียบเทียบความคลาดเคลื่อนมาตรฐานที่เกิดจากการใช้ 4 วิธีก็พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เช่นกัน

แมคแคน (Robert G. McCann .1989: 236-276)ได้ศึกษาการเปรียบเทียบวิธีเทียบคะแนนที่สังเกตได้ 2 ค่า ที่สมมติว่ามีความเชื่อมั่นเท่าเทียมกัน และใช้แบบทดสอบที่เป็นคะแนนจริงสัมพันธ์ โดยวิธีการเปรียบเทียบความคลาดเคลื่อนมาตรฐาน 4 วิธี คือ วิธีของแมคแคน 2 วิธี ของทัคเกอร์ 1 วิธีและของลีไวน์ 1 วิธี พบว่า วิธีที่ 1 ของแมคแคนให้ผลใกล้เคียงกับวิธีของทัคเกอร์ ส่วนวิธีที่ 2 ให้ผลใกล้เคียงกับวิธีของลีไวน์ และทั้ง 4 วิธี ให้ผลความคลาดเคลื่อนมาตรฐานในระดับที่น่าพอใจเป็นส่วนมาก ข้อดีของวิธีที่ 1 และ 2 นั้นไม่จำเป็นต้องใช้ข้อตกลงที่ว่าการกระจายของคะแนนจากแบบทดสอบ 2 ฉบับที่นำมาเปรียบเทียบกันต้อมีการแจกแจงเป็นโค้งปกติ

ราชู (Raju. 1977: 549-565) ได้พัฒนาสัมประสิทธิ์แอลฟ่า r_c ของครอนบัค(Cronbach. 1951) ให้สามารถประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนด้วยความยาวไม่เท่ากัน ราชู เรียกว่า สัมประสิทธิ์เบต้า(βk Coefficient)หรือใช้สัญลักษณ์ r_R และได้ทำการทดลองเปรียบเทียบสัมประสิทธิ์ r_R กับสัมประสิทธิ์ r_c ของครอนบัค สัมประสิทธิ์ r_{FS} ของเฟลดต์(Feldt. 1975) สัมประสิทธิ์ r_H ของฮอร์ส(Horse. 1951) สัมประสิทธิ์ r_k ของคริสทอฟ(Kritstof. 1974) โดยศึกษาจากนักเรียนเกรด 6 จำนวน 300 คน ที่สอบแบบทดสอบการคิดคำนวณจำนวน 40 ข้อ จากชุดอนุกรมผลสัมฤทธิ์ เอส อาร์ เอ ระดับต้น ในปี 1975 ผลการวิจัยพบว่า ค่าความเชื่อมั่นของสัมประสิทธิ์ KR-20 มีค่าเท่ากับ 0.854 และจากการแบ่งแบบทดสอบออกเป็นหลายๆ ส่วนเป็น 4 แบบต่างกันคือ แบ่งสองส่วน 2 แบบ แบ่งสามส่วน 1 แบบ และแบ่งสี่ส่วน 1 แบบ ในการแบ่งแบบทดสอบแต่ละแบบจะกำหนดแบบทดสอบใส่ในแต่ละส่วนโดยการสุ่ม ผลพบว่า สัมประสิทธิ์ r_R ประมาณค่าได้ดีเมื่อเทียบกับ KR-20 ยกเว้นเฉพาะการแบ่งสองส่วนที่มีขนาด(35, 5) ส่วนสัมประสิทธิ์ r_{FS} r_H r_k ก็ได้ผลดีเช่นเดียวกัน แต่สัมประสิทธิ์ r_c ให้ค่าความเชื่อมั่นต่ำที่สุดในทุกกรณี

คริสทอฟ (Kristof. 1974: 491-499) ได้ทำการศึกษาการประมาณค่าความเชื่อมั่นและความแปรปรวนของคะแนนจริงจากการแบ่งแบบทดสอบออกเป็น 3 ส่วน กลุ่มตัวอย่างเป็นนักเรียนจำนวน 2,000 คน โดยใช้แบบทดสอบมาตรฐานวัดความถนัดชุดคำศัพท์ จำนวน 144 ข้อ ซึ่งแบ่งแบบทดสอบออกเป็นส่วนย่อย 3 ส่วนที่แตกต่างกัน 7 แบบ คือ แบบ A ส่วนที่ 1 ได้แก่ข้อ 1-

48 จำนวน 48 ข้อ ส่วนที่ 2 ได้แก่ข้อ 49-96 จำนวน 48 ข้อ ส่วนที่ 3 ได้แก่ ข้อ 97-144 จำนวน 48 ข้อ แบบ B ส่วนที่ 1 ได้แก่ข้อ 1, 3, 5, 7,… จำนวน 72 ข้อ ส่วนที่ 2 ได้แก่ ข้อ 2, 6, 10, 14,… จำนวน 36 ข้อ ส่วนที่ 3 ได้แก่ ข้อ 4, 8, 12, 16,… จำนวน 36 ข้อ แบบ C_1 แบ่งด้วยการสุ่มด้วยความน่าจะเป็น 1/3 ทั้งสามส่วน ได้ส่วนที่ 1 จำนวน 47 ข้อ ส่วนที่ 2 จำนวน 49 ข้อ ส่วนที่ 3 จำนวน 48 ข้อ แบบ C_2 ทำเช่นเดียวกับแบบ C_1 ได้ส่วนที่ 1 จำนวน 45 ข้อ ส่วนที่ 2 จำนวน 59 ข้อ ส่วนที่ 3 จำนวน 40 ข้อ แบบ D_1 และแบบ D_2 แบ่งการสุ่มด้วยความน่าจะเป็น 1/6, 1/3 และ ½ ตามลำดับ แบบ D_1 ได้จำนวนข้อสอบเท่ากับ 27, 47 และ 70 ข้อ และแบบ D_2 ได้จำนวนข้อสอบเท่ากับ 19, 61 และ 64 ข้อ ตามลำดับ แบบ E ส่วนที่ 1 ได้แก่ ข้อ 1, 4, 7, 10,… จำนวน 48 ข้อ ส่วนที่ 2 ได้แก่ ข้อ 2, 5, 8, 11,… จำนวน 48 ข้อ ส่วนที่ 3 ได้แก่ ข้อ 3, 6, 9, 12,… จำนวน 48 ข้อ จากนั้นคำนวณค่าสัมประสิทธิ์ r_k ของเขา สัมประสิทธิ์ r_g ของกัตต์แมน(Guttman. 1945) และสัมประสิทธิ์ r_c ของครอนบัค(Cronbach. 1951) จากการคำนวณค่าสัมประสิทธิ์โดยใช้สูตรทั้งสามและนำค่าสัมประสิทธิ์มาเปรียบเทียบกัน พบว่า ในการแบ่งแบบทดสอบแต่ละฉบับดังกล่าวข้างต้น ได้ค่าสัมประสิทธิ์ความเชื่อมั่น $r_k > r_g > r_c$

งานวิจัยที่เกี่ยวข้องกับสัมประสิทธ์ r_B จากส่วนหนึ่งของงานวิจัย ซึ่งมีหลายคนศึกษา อาทิ

บุญเชิด ภิญโญอนันตพงษ์ (2542: บทคัดย่อ) ได้ทำการศึกษาสัมประสิทธิ์ r_B : การประมาณค่าความเชื่อมั่นสำหรับแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยความยากรายข้อต่างกัน เพื่อพัฒนาและศึกษาผลการใช้สูตร r_B ที่ผู้วิจัยพัฒนาขึ้นมาสำหรับประมาณค่าความเชื่อมั่นของแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยข้อสอบที่ให้คะแนนระบบ 0, 1 และมีความยาดรายข้อต่างกัน โดยศึกษาจากแบบทดสอบที่วัดเนื้อหาเดียวกัน จำนวนข้อต่างกัน จำนวน 6 ฉบับ คือ ฉบับ 25, 30, 35, 40, 45, และ 50 ข้อ ทั้งจากกลุ่มตัวอย่างขนาดใหญ่และกลุ่มตัวอย่างขนาดเล็ก คือ 30, 50 และ80 คนว่าเป็นอย่างไรเมื่อเทียบกับสูตร KR-20 ผลการศึกษาพบว่า สูตร r_B คำนวณค่าความเชื่อมั่นของแบบทดสอบจำนวนข้อต่างกันจากกลุ่มตัวอย่างขนาดใหญ่ได้สูงสอดคล้องสัมพันธ์กับสูตร KR-20 โดยสูตร r_B คำนวณค่าความเชื่อมั่นได้สูงกว่าสูตร KR-20 ทุกฉบับ เมื่อศึกษาจากกลุ่มตัวอย่างขนาดเล็ก พบว่า สูตร r_B คำนวณจากแบบทดสอบฉบับเดียวกันและกลุ่มตัวอย่างขนาดเดียวกันได้สอดคล้องสัมพันธ์กับสูตร KR-20 จากแบบทดสอบทั้ง 6 ฉบับและกลุ่มตัวอย่างทุกขนาด สูตร r_B คำนวณค่าความเชื่อมั่นสูงกว่าสูตร KR-20 ส่วนจำนวนข้อสอบและขนาดกลุ่มตัวอย่างมีผลต่อค่าความเชื่อมั่นที่คำนวณได้จากทั้งสองสูตรในลักษณะคล้ายคลึงกัน และทั้งสองสูตรประมาณค่าความเชื่อมั่นจากกลุ่มตัวอย่างขนาดเล็กเทียบกับกลุ่มตัวอย่างขนาดใหญ่มีค่าความลำเอียงในทิศทางเดียวกัน จึงสรุปได้ว่า สูตร r_B ที่พัฒนาขึ้นครั้งนี้ สามารถนำไปใช้ในการประมาณค่าความเชื่อมั่นของแบบทดสอบได้ดีและมีค่าความเชื่อมั่นสูงกว่าสูตร KR-20

วิชุดา กิจธรธรรม (2544: 37-38) ได้ทำการศึกษาการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนของแบบวัดความถนัดทางวิชาชีพพยาบาลตามแนวคิดของฟลานาแกนด้วยวิธีการวิเคราะห์ความแปรปรวน วิธีการตามทฤษฎีการตอบข้อสอบ และวิธีการของคีต กลุ่มตัวอย่างที่ใช้เป็นนักศึกษาพยาบาลชั้นปีที่ 4 ปีการศึกษา 2543 ในสถาบันอุดมศึกษาเอกชนเขตกรุงเทพมหานคร จำนวน 500 คน ได้จากการสุ่มแบบสองขั้นตอน เครื่องมือที่ใช้เป็นแบบทดสอบวัดความถนัดทางวิชาชีพพยาบาลของฟลานาแกน จำนวน 5 ฉบับ ผลการวิจัยพบว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดทางวิชาชีพพยาบาลของฟลานาแกน ทั้ง 5 ฉบับ ที่วิเคราะห์โดยใช้สูตร r_B ให้ค่าความเชื่อมั่นสูงกว่าสูตร KR-20 และ KR-21

สุทธิวรรณ พีรศักดิ์โสภณ (2546: 62-68) ได้ทำการศึกษาการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปเมื่อใช้วิธีการประมาณค่าโดยทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลดขควอลส์ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาตอนปลาย ปีการศึกษา 2544 จำนวน 2,000 คน และสุ่มมาเป็นตัวอย่างจำลองสามขนาดคือ 300 500 800 คน แต่ละขนาดได้มาโดยการสุ่มซ้ำ 30 ครั้ง เครื่องมือที่ใช้เป็นแบบทดสอบวัดความถนัดทางการเรียนสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย จำนวน 3 ฉบับคือ ฉบับด้านตัวเลข ด้านภาษา และด้านเหตุผล ผลการวิจัยพบว่า แบบทดสอบวัดความถนัดทางการเรียนทั้งสามด้านคือด้านตัวเลข ด้านภาษา ด้านเหตุผล จากกลุ่มตัวอย่างขนาดใหญ่ สูตร r_B ให้ค่าความเชื่อมั่นสูงกว่าสูตร KR-20 และ KR-21 จากกลุ่มตัวอย่างจำลองขนาด 300 500 800 คน ค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR-20 และ KR-21

อ่อนนุช เต็มเปี่ยม (2545: บทคัดย่อ) ได้ศึกษา หลักฐานความเที่ยงตรงเชิงโครงสร้างและความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ของแบบทดสอบความสามารถทางสมอง เพื่อศึกษาหลักฐานความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบความความสามารถทางสมองโดยวิธีวิเคราะห์แบบหลายลักษณะหลายวิธี วิธีวิเคราะห์องค์ประกอบเชิงสำรวจ และเพื่อศึกษาหลักฐานความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ของแบบทดสอบความสามารถทางสมอง ที่คำนวณด้วยสูตรสัมประสิทธิ์ r_B กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ของโรงเรียนในสังกัดสำนักงานการศึกษาเอกชนอำเภอหาดใหญ่ จังหวัดสงขลา ผลการศึกษาพบว่า วิธีหลายลักษณะหลายวิธีมีค่าสัมประสิทธิ์ความเที่ยงตรงร่วมอยู่ระหว่าง .5516 ถึง .7097 มีสัมประสิทธิ์ความเที่ยงตรงแยกอยู่ระหว่าง .4983 ถึง .7299 โดยค่าความเที่ยงตรงร่วมของคุณลักษณะที่วัดจะมีค่าสูงกว่าความเที่ยงตรงแยกเกือบทุกค่า วิธีวิเคราะห์องค์ประกอบเชิงสำรวจ พบว่า ข้อคำถามส่วนใหญ่มีความเที่ยงตรงในการวัดองค์ประกอบได้สอดคล้องกัน ซึ่งประกอบด้วยความสามารถด้านตัวเลข ด้านเหตุผล ด้านภาษา เมื่อคำนวณด้วยสูตรสัมประสิทธิ์ r_B ของแบบเติมคำ มีค่า .9288 แบบเลือกตอบ มีค่า .8903

งานวิจัยที่เกี่ยวข้องกับสัมประสิทธิ์โครงสร้าง จากส่วนหนึ่งของงานวิจัย ซึ่งมีหลายคนศึกษา อาทิ

เรยูเตอร์เบิร์กและกูสตาฟสัน(Reuterberg and Gustafsson. 1992: 795-811) ได้ทำการศึกษาการวิเคราะห์องค์ประกอบเชิงยืนยันและความเชื่อมั่น : การทดสอบข้อตกลงเบื้องต้นของโมเดลการวัด เครื่องมือที่ใช้เป็นแบบทดสอบทางภาษา(คำตรงข้าม) จำนวน 40 ข้อ 5 องค์ประกอบ องค์ประกอบละ 8 ข้อ วิเคราะห์ข้อมูลโดยใช้โปรแกรม LISREL ผลการวิจัยพบว่า เมื่อใช้โมเดลการวัดมาตรฐานเดิม(Parallel Measurement Model) ได้ค่า χ^2 =316.07 , p-value< .01, AGFI = .884 และ PGFI = .779 แสดงว่าโมเดลไม่สอดคล้องกับข้อมูลเชิงประจักษ์ มีค่าน้ำหนักองค์ประกอบในแต่ละองค์ประกอบเท่ากับ 1.02, 1.02, 1.02, 1.02, 1.02 ตามลำดับ มีค่าความแปรปรวนคลาดเคลื่อนเท่ากับ 1.38, 1.38, 1.38, 1.38, 1.38 ตามลำดับ มีค่าความเชื่อมั่นเท่ากับ .790 เมื่อใช้โมเดลคะแนนจริงสัมพันธ์(Congereric Model) ได้ค่า χ^2 =7.98 , p-value> .05, AGFI = .992 และPGFI = .332 แสดงว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์ มีค่าน้ำหนักองค์ประกอบในแต่ละองค์ประกอบเท่ากับ .81, 1.30, 1.37, 0.77, 0.88 ตามลำดับ มีค่าความแปรปรวนคลาดเคลื่อนเท่ากับ 1.26, 1.52, 1.29, 1.03, 1.41 ตามลำดับ มีค่าความเชื่อมั่นเท่ากับ .803

เรคอฟ (Raykov, T. 2003: 143-159). ได้ศึกษาการพัฒนามาตรวัดโดยการทดสอบการเปลี่ยนแปลงความเชื่อมั่นรวม โดยใช้โมเดลคะแนนจริงสัมพันธ์ ของโมเดลสมการโครงสร้าง กำหนดกลุ่มตัวอย่าง 2 วิธีคือ วิธีการจำลองข้อมูล(Simulated Data) กับเก็บข้อมูลจริง วิธีการจำลองข้อมูลมีขนาด 500 คน ศึกษากับ 5 ตัวแปร ข้อมูลจริงมีกลุ่มตัวอย่างเป็นนักเรียนชายจำนวน 300 คน ที่ทำแบบทดสอบเหตุผล(SAT1)ในปี 1998 จำนวน 35 ข้อ โดยให้ข้อ 1- 10 เป็นตัวแปรที่ 1 ข้อที่ 11-20 เป็นตัวแปรที่ 2 ข้อที่ 21-30 เป็นตัวแปรที่ 3 และข้อที่ 31-35 เป็นตัวแปรที่ 4 ทำการวิเคราะห์ข้อมูลด้วยโปรแกรม LISREL ผลการวิจัยพบว่า เมื่อใช้โมเดลคะแนนจริงสัมพันธ์กับวิธีการจำลองข้อมูล 5 ตัวแปร ได้ค่า χ^2 =4.06 , df=5 , p-value= .54 ได้ค่าความเชื่อมั่นได้เท่ากับ .90 ตัวแปรตัวที่ 5 มีค่านำหนักองค์ประกอบน้อยสุดคือ .11 ความแปรปรวนของความคลาดเคลื่อนเท่ากับ 1.48 แสดงว่า ตัวแปรที่ 5 อาจไม่ได้วัดคุณลักษณะเดียวกัน การตัดตัวแปรที่ 5 ออกจึงเป็นการปรับปรุงมาตรวัด แล้วทำการวิเคราะห์จาก 4 ตัวแปรได้ค่าความเชื่อมั่นเท่ากับ .93 สูงกว่าอย่างมีนัยสำคัญและมีค่า χ^2 =.68 , df=2 , p-value= .71 เมื่อใช้โมเดลการวัดคะแนนจริงสัมพันธ์กับข้อมูลจริง 4 ตัวแปร ได้ค่า χ^2 =5.87 , df=2 , p-value= .053 ได้ค่าความเชื่อมั่นได้เท่ากับ .83 หากนำข้อสอบ 5 ข้อสุดท้ายในตัวแปรที่ 4 ออก ก็จะได้ค่า χ^2 =15.42 , df=3 , p-value= .0 ค่าความเชื่อมั่นได้เท่ากับ .81 แสดงให้เห็นว่าความเชื่อมั่นของข้อสอบ 35 ข้อไม่เท่ากับความเชื่อมั่นของข้อสอบ 30 ข้อ ดังนั้นความยาวของข้อสอบที่ลดลง 5 ข้อ นำไปสู่การสูญเสียความเชื่อมั่น

งานวิจัยที่เกี่ยวข้องกับทฤษฎีการทดสอบมาตรฐานเดิมกับทฤษฎีการตอบข้อสอบสัมประสิทธ์ ρ_{IRT} จากส่วนหนึ่งของงานวิจัย ซึ่งมีหลายคนศึกษา อาทิ

นภดล ยิ่งยงสกุล (2540: บทคัดย่อ) ได้ศึกษาความเที่ยง(ความเชื่อมั่น) ความตรง (ความเที่ยงตรง)ของแบบทดสอบและความสัมพันธ์ของคะแนนสอบระหว่างการให้คะแนนตามทฤษฎีการทดสอบมาตรฐานเดิมกับการให้คะแนนตามทฤษฎีการตอบข้อสอบ กลุ่มตัวอย่างคือนักเรียนชั้นประถมศึกษาปีที่ 5 จังหวัดยะลา จำนวน 231 คน เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตสาสตร์ จำนวน 3 ฉบับ พบว่า ความเชื่อมั่นของแบบทดสอบจากวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามความสามารถของผู้สอบ สูงกว่าวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อสอบ และวิธีการให้น้ำหนักคะแนนรายข้อเท่ากันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อสอบ มีค่าสูงกว่าวิธีการให้น้ำหนักคะแนนรายข้อเท่ากัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เบญจพร ยนตจักรวิถี (2540: บทคัดย่อ) ได้ทำการเปรียบเทียบผลการวิเคราะห์ข้อสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ระหว่างทฤษฎีการทดสอบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ กับนักเรียนในสังกัดกรมสามัญศึกษา จังหวัดสงขลา จำนวน 12,545 คน ผลการวิจัยพบว่า 1.การวิเคราะห์ข้อสอบด้วยทฤษฎีมาตรฐานเดิม ได้ข้อสอบที่มีคุณภาพดีตามเกณฑ์ 32 ข้อ ส่วนโมเดลของราสค์ได้ข้อสอบที่มีคุณภาพดีตามเกณฑ์ 18 ข้อ จำนวนข้อสอบที่ได้คัดเลือกไว้ 2 วิธี มีสัดส่วนแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และไม่มีความสอดคล้องกัน 2.สัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความยากของข้อสอบที่ได้จากการวิเคราะห์ด้วยทฤษฎีมาตรฐานเดิมและโมเดลของราสค์มีค่าเท่ากับ 0.99 ซึ่งมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 3.สัมประสิทธิ์สหสัมพันธ์ระหว่างค่าคะแนนความสามารถของผู้สอบที่ได้จากการวิเคราะห์ด้วยทฤษฎีมาตรฐานเดิมและโมเดลของราสค์มีค่าเท่ากับ 0.93 ซึ่งมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

วัฒนา ขัดสี (2533: บทคัดย่อ) ได้ทำการศึกษาความสัมพันธ์ระหว่างค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ข้อสอบและคะแนนความสามารถในการสอบโดยทฤษฎีมาตรฐานเดิมกับทฤษฎีการตอบข้อสอบของแบบทดสอบเลือกตอบ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2532 จังหวัดสกลนคร จำนวน 1,008 คน เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบเลือกตอบวิชาคณิตศาสตร์ ชนิด 5 ตัวเลือก จำนวน 36 ข้อ ผลการวิจัยพบว่า 1.ค่าความยากของข้อสอบที่ได้จากการวิเคราะห์ข้อสอบโดยทฤษฎีมาตรฐานเดิมและทฤษฎีการตอบข้อสอบมีค่าอยู่ระหว่าง 10.6 ถึง 16.6 และ -1.146 ถึง 4.478 ตามลำดับ มีค่าความยากเฉลี่ยเท่ากับ 13.14 และ 0.647 ตามลำดับ ซึ่งแสดงให้เห็นว่าเป็นแบบทดสอบที่มีค่าความยากง่ายเบี่ยงเบนไปทางยาก และค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความยากของข้อสอบที่ได้จากการวิเคราะห์โดยทฤษฎีมาตรฐานเดิมกับทฤษฎีการตอบข้อสอบมีค่าเท่ากับ 0.7969 ซึ่งมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 2.ค่าอำนาจจำแนกของข้อสอบที่ได้จากทฤษฎีมาตรฐานเดิมกับทฤษฎีการตอบข้อสอบมีค่าอยู่ระหว่าง 0.19 ถึง 0.65 และ 0.358 ถึง 2.00 ตามลำดับ มีค่า

อำนาจจำแนกเฉลี่ยเท่ากับ 0.485 และ 0.840 ตามลำดับ ค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวิเคราะห์โดยทฤษฎีทั้งสองมีค่าเท่ากับ 0.0702 ซึ่งมีความสัมพันธ์กันอย่างไม่มีนัยสำคัญทางสถิติ 3.คะแนนความสามารถในการสอบวิชาคณิตศาสตร์เฉลี่ยที่ได้จากการวิเคราะห์โดยทฤษฎีมาตรฐานเดิมกับทฤษฎีการตอบข้อสอบมีค่าเม่ากับ 16.25 และ -0.165 สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนความสามารถที่ได้จากทฤษฎีทั้งสองมีค่าเท่ากับ 0.7722 ซึ่งมีนัยสำคัญทางสถิติที่ระดับ .01

วิรัตน์ ธรรมาภรณ์และคณะ (2545: บทคัดย่อ) ได้ทำการศึกษาการวิเคราะห์ข้อสอบความถนัดทางการเรียนโดยใช้ทฤษฎีการทดสอบดั้งเดิมและทฤษฎีการตอบข้อสอบ มีวัตถุประสงค์เพื่อวิเคราะห์คุณภาพข้อสอบของแบบทดสอบวัดความถนัดทางการเรียนที่ใช้ในการสอบคัดเลือกผู้เข้าศึกษาต่อระดับปริญญาโท สังกัดคณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ สำหรับนักศึกษาปีการศึกษา 2539 โดยใช้การวิเคราะห์ข้อสอบตามทฤษฎีการทดสอบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ เพื่อเปรียบเทียบวิธีทั้งสอง และเพื่อเปรียบเทียบคุณภาพของแบบทดสอบย่อยซึ่งประกอบด้วยแบบทดสอบวัดความถนัดเชิงจำนวน และความถนัดเชิงเหตุผล กลุ่มตัวอย่างที่ใช้ในการวิจัยคือ ผู้สอบคัดเลือกเข้าศึกษาระดับปริญญาโทของคณะศึกษาศาสตร์ ปีการศึกษา 2539 จำนวน 450 คน ข้อมูลที่ใช้ในการวิเคราะห์ได้จากกระดาษคำตอบของผู้สอบและใช้โปรแกรม BILOG Version 3.04 ในการวิเคราะห์หาค่าความยากและอำนาจจำแนกตามทฤษฎีการทดสอบดั้งเดิมและวิเคราะห์ค่าอำนาจจำแนก ค่าความยากและค่าการเดาตามทฤษฎีการตอบสนองข้อสอบใช้การทดสอบไค-สแควร์ ของแมคนีมาร์ และการทดสอบค่า Z เพื่อเปรียบเทียบผลการวิเคราะห์ระหว่าง 2 วิธี และเพื่อเปรียบเทียบสัดส่วนของข้อสอบที่ผ่านการคัดเลือกระหว่างข้อสอบทั้ง 2 ฉบับ ผลการวิจัยพบว่า 1.ข้อสอบวัดความถนัดเชิงจำนวนผ่านการคัดเลือกด้วยวิธีการวิเคราะห์ตามทฤษฎีการทดสอบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบร้อยละ 18.3 และ 26.7 ตามลำดับ 2.ข้อสอบวัดความถนัดเชิงเหตุผลผ่านการคัดเลือกด้วยวิธีวิเคราะห์ตามทฤษฎีการทดสอบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบร้อยละ 37.8 และ 48.6 ตามลำดับ 3.วิธีวิเคราะห์ตามทฤษฎีการทดสอบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบได้ผลการคัดเลือกข้อสอบที่มีคุณภาพไม่แตกต่างกัน ($p>.05$) ทั้งสองฉบับ 4.ข้อสอบวัดความถนัดเชิงเหตุผลมีข้อสอบที่ผ่านการคัดเลือกมากกว่าข้อสอบวัดความถนัดเชิงจำนวน ($p<.001$) ทั้งการวิเคราะห์ตามทฤษฎีการทดสอบดั้งเดิม ตามทฤษฎีการตอบสนองข้อสอบและจากการผ่านคัดเลือกทั้งสองวิธีรวมกัน

สยมภู รักษ์สังข์ (2547: บทคัดย่อ) ได้ศึกษาความสำพันธ์ของผลการวิเคราะห์ข้อสอบแบบทฤษฎีมาตรฐานเดิม(CTT)และทฤษฎีการตอบข้อสอบ(IRT) โดยมีจุดมุ่งหมายเพื่อศึกษาความสำพันธ์ของค่าความยากและค่าอำนาจจำแนกของการวิเคราะห์ข้อสอบแบบทฤษฎีมาตรฐานเดิม (CTT)และทฤษฎีการตอบข้อสอบ(IRT)ที่คำนวณจากกลุ่มตัวอย่างขนาดใหญ่และขนาดเล็กลง กลุ่มตัวอย่างที่ใช้ในการวิจัยคือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานในเขตกรุงเทพมหานคร จำนวน 2,000 คน และกำหนดให้เป็นประชากรเทียม จากนั้นทำการสุ่มแบบใส่คืน ขนาด 1,600 800 400 200 100 และ 50 คน ขนาดละ 10 รอบ เครื่องมือที่

ใช้คือ แบบทดสอบวัดความสามารถในการอ่าน วิชาภาษาไทย ชนิดเลือกตอบ จำนวน 40 ข้อ ผลการวิจัยพบว่า การวิเคราะห์ข้อสอบระหว่างทฤษฎีมาตรฐานเดิมกับทฤษฎีการตอบข้อสอบ ส่วนใหญ่มีค่าความยาก(P กับ b)ที่มีความสัมพันธ์กันทางลบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้นกลุ่มตัวอย่างขนาด 2,000 และ 100 คน ที่ไม่พบความสัมพันธ์กัน ส่วนค่าอำนาจจำแนก(D กับ a) และ (r กับ a) ทุกขนาดกลุ่มตัวอย่างมีความสัมพันธ์ทางบวก อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ดิมิทรอฟ(Dimitrov. 2003: 440-458) ได้ศึกษาการวัดขอบเขตคะแนนจริงและความเชื่อมั่นของแบบทดสอบที่ให้คะแนนแบบ 2 ค่าด้วยฟังก์ชั่นพารามิเตอร์ของทฤษฎีการตอบข้อสอบ ศึกษาจากกลุ่มตัวอย่าง 2 กลุ่ม คือ กลุ่มที่หนึ่งเป็นการจำลองข้อมูลจำนวน 8,000 คน แบบทดสอบมีจำนวน 20 ข้อ ทำการศึกษาโดยใช้โมเดลโลจีสติก 2 พารามิเตอร์ กลุ่มที่สองเป็นนักเรียนชั้นประถมศึกษาปีที่ 5 จำนวน 4,854 คน ทำแบบทดสอบเลือกตอบของ Ohio Off-Grade Proficiency Test(OOPT) จำนวน 24 ข้อ ทำการศึกษาโดยใช้โมเดลโลจีสติก 3 พารามิเตอร์ วิเคราะห์ข้อมูลด้วยโปรแกรม XCALIBRE และ SPSS ผลการวิจัยพบว่า โมเดลโลจีสติก 2 พารามิเตอร์ ได้ค่าคาดหวังของความแปรปรวนของคะแนนจริงเท่ากับ 6.315 ค่าคาดหวังของความแปรปรวนของคะแนนคลาดเคลื่อนเท่ากับ 3.719 ค่าความเชื่อมั่นเท่ากับ .63 โมเดลโลจีสติก 3 พารามิเตอร์ ได้ค่าคาดหวังของความแปรปรวนของคะแนนจริงเท่ากับ 16.520 ค่าคาดหวังของความแปรปรวนของคะแนนคลาดเคลื่อนเท่ากับ 4.471 ค่าความเชื่อมั่นเท่ากับ .789

จากเอกสารและงานวิจัยที่เกี่ยวข้องที่กล่าวมาข้างต้นจะเห็นได้ว่า ในงานวิจัยโดยทั่วไป ได้ใช้การประมาณค่าความเชื่อมั่นด้วยสัมประสิทธิ์ KR-20 อย่างมากมาย นักการศึกษาและนักวิจัยส่วนใหญ่รู้จักและใช้ในการประมาณค่าความเชื่อมั่นของแบบทดสอบ สำหรับการประมาณค่าความเชื่อมั่นด้วยสัมประสิทธิ์ r_B ที่ได้มีการพัฒนามาจากสัมประสิทธิ์ r_{L2} ก็ได้มีการศึกษาและใช้มาตั้งแต่ปี 42 เพื่อประมาณค่าความเชื่อมั่นภายใต้โมเดลคะแนนจริงสัมพันธ์ ส่วนสัมประสิทธิ์โครงสร้างงานวิจัยโดยส่วนใหญ่นิยมใช้ในการวิเคราะห์องค์ประกอบ การวิเคราะห์อิทธิพล การวิเคราะห์กลุ่มพหุ และโมเดลโค้งพัฒนาการ มีส่วนน้อยที่จะใช้ในการประมาณค่าความเชื่อมั่น เช่นเดียวกันกับสัมประสิทธ์ ρ_{IRT} ก็มีส่วนน้อยที่ใช้ในการประมาณค่าความเชื่อมั่น ส่วนใหญ่นิยมใช้ในการประมาณค่าพารามิเตอร์และฟังก์ชั่นสารสนเทศของข้อสอบ เมื่อทำการศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบส่วนใหญ่ได้ทำการศึกษาภายใต้สูตรแต่ละสูตร ไม่ได้มีการศึกษารวมกันหลายสูตร ผู้วิจัยจึงมีความสนใจที่ศึกษาทั้ง 4 สูตร อีกทั้งการประมาณค่าความเชื่อมั่นของแบบทดสอบ มีทฤษฎีที่สามารถใช้ในการวัดประเมิน(Assessing) ที่เหมาะสมกับระดับการประเมินในระดับชั้นเรียน(Classroom Assessment) กับระดับใหญ่(Large Assessment) และเมื่อนำมาใช้ในการประมาณค่าของแบบทดสอบที่มีลักษณะเป็นแบบเลือกตอบ มีการตรวจให้คะแนนเป็นแบบ 2 ค่า คือ ผิดให้ 0 คะแนน ,ถูกให้ 1 คะแนน มีค่าความยากของข้อสอบต่างกัน คือ ทฤษฎีการทดสอบ

มาตรฐานเดิมที่มีความคู่ขนานแบบ คะแนนจริงสมมูล ได้แก่ KR-20 และแบบคะแนนจริงสัมพันธ์ ได้แก่ r_B โมเดลการทดสอบคะแนนจริงสัมพันธ์ที่วัดองค์ประกอบเดี่ยว ได้แก่ สัมประสิทธิ์โครงสร้าง และทฤษฎีการตอบข้อสอบ ได้แก่ ρ_{IRT} ผู้วิจัยจึงมีความสนใจที่จะศึกษาสูตรที่มีการใช้จากทฤษฎีเหล่านี้สามารถใช้ได้ในการวัดประเมินในลักษณะใด

บทที่ 3
วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้เป็นการศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และโมเดลโลจีสติก 3 พารามิเตอร์ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. การกำหนดประชากรและกลุ่มตัวอย่าง
2. การสร้างเครื่องมือที่ใช้ในการวิจัย
3. การเก็บรวบรวมข้อมูล
4. การจัดกระทำและการวิเคราะห์ข้อมูล

การกำหนดประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ปีการศึกษา 2550 จำนวน 15,395 คน ประกอบด้วยเขตพื้นที่การศึกษาจังหวัดเชียงราย เขต 1 จำนวน 3,535 คน เขต 2 จำนวน 4,076 คน เขต 3 จำนวน 4,607 คน และเขต 4 จำนวน 3,177 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างขนาดใหญ่

กลุ่มตัวอย่างขนาดใหญ่ที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ปีการศึกษา 2550 จำนวน 6,000 คน ทำการสุ่มแบบชั้นภูมิจากเขตพื้นที่การศึกษาจังหวัดเชียงราย ที่มีอยู่ 4 เขต ได้กลุ่มตัวอย่างขนาดใหญ่ของแต่ละเขตพื้นที่การศึกษาดังนี้ เขต 1 จำนวน 1,378 คน เขต 2 จำนวน 1,588 คน เขต 3 จำนวน 1,796 คน และเขต 4 จำนวน 1,238 คน

ซึ่งกลุ่มตัวอย่างขนาดใหญ่นี้ ผู้วิจัยกำหนดให้เป็นประชากรเทียม(Pseudo Population)

กลุ่มตัวอย่างขนาดกลางและขนาดเล็ก

กลุ่มตัวอย่างขนาดกลางและขนาดเล็กที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ปีการศึกษา 2550 โดยใช้วิธีการสุ่มแบบใส่คืนจากประชากรเทียม แบ่งเป็นกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ 300 500 1,000 และ 1,500 คน กลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ 30 60 120 และ 240 คน แต่ละขนาดมีจำนวน 30 กลุ่ม

การสร้างเครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 ที่ผู้วิจัยสร้างขึ้น มีลักษณะเป็นแบบทดสอบเลือกตอบ 4 ตัวเลือก โดยมีขั้นตอนของการสร้างและพัฒนาแบบทดสอบดังนี้

1. ศึกษาหลักสูตรและเอกสารงานวิจัยที่เกี่ยวข้องเพื่อเป็นแนวทางในการสร้างแบบทดสอบ วิเคราะห์มาตรฐานการเรียนรู้ กรอบสาระการเรียนรู้ รายละเอียดสาระการเรียนรู้และผลการเรียนที่คาดหวัง เพื่อกำหนดมวลความรู้หลักและมวลความรู้ย่อย โดยดำเนินการดังนี้

1.1 วิเคราะห์เนื้อหาวิชาคณิตศาสตร์ ระดับประถมศึกษาปีที่ 5 เพื่อกำหนดเป็นขอบเขตเนื้อหาใหญ่และวิเคราะห์เนื้อหาย่อยในแต่ละเนื้อหาใหญ่

1.2 วิเคราะห์จุดประสงค์เฉพาะวิชาของเนื้อหาวิชาคณิตศาสตร์ เพื่อหาพฤติกรรมหลัก (Conceptual Objective) ซึ่งเป็นพฤติกรรมใหญ่ที่ต้องการวัดโดยใช้แนวทางการแบ่งพฤติกรรมด้านพุทธิพิสัย(Cognitive) ตามแนวคิดของบลูม

1.3 นำผลที่ได้จากข้อ 1.1 และ 1.2 มาวิเคราะห์เพื่อกำหนดขอบเขตของมวลความรู้หลักและมวลความรู้ย่อย กำหนดตารางวิเคราะห์หลักสูตรและจำนวนเวลาที่ใช้ในการจัดการเรียนรู้

2. ตรวจสอบความเหมาะสมของมวลความรู้ย่อยในแต่ละมวลความรู้ภายในขอบเขตเนื้อหาที่กำหนด โดยให้ผู้เชี่ยวชาญการสอนคณิตศาสตร์ไม่น้อยกว่า 5 ปี จำนวน 5 ท่าน เป็นผู้พิจารณาตัดสินความครอบคลุมของมวลความรู้ย่อยในมวลความรู้หลักที่กำหนดขึ้น โดยวิธีการตัดสินความสอดคล้องด้วยการหาค่าดัชนีความสอดคล้อง ถ้าค่าเฉลี่ยของคะแนนการตัดสินของผู้เชี่ยวชาญในแต่ละมวลความรู้ย่อยมีค่าสูงกว่า 0.5 แสดงว่ามวลความรู้ย่อยที่กำหนดขึ้นมีความครอบคลุมมวลความรู้ในขอบเขตเนื้อหานั้น โดยมีการหาค่าความเที่ยงตรงเชิงเนื้อหาดังนี้

ค่าความเที่ยงตรงเชิงเนื้อหา ใช้ดัชนีความสอดคล้อง(Item Index of Objective Congruent) โดยมีค่าเฉลี่ยของคะแนนตัดสินของผู้เชี่ยวชาญในแต่ละมวลความรู้มีค่าตั้งแต่ 0.5 ขึ้นไปแสดงว่ามวลความรู้ย่อยที่กำหนดขึ้นมีความครอบคลุมมวลความรู้ในขอบเขตเนื้อหาและข้อสอบข้อนั้นวัดตามคุณลักษณะเฉพาะที่ข้อสอบจากมวลความรู้ย่อยที่กำหนดไว้ โดยมีสูตรดังนี้
(บุญเชิด ภิญโญอนันตพงษ์. 2547: 179)

$$IOC = \frac{\sum R}{N}$$

เมื่อ R แทนความคิดเห็นของผู้เชี่ยวชาญ
ถ้าแน่ใจว่ามีความสอดคล้อง ให้ค่า +1
ถ้าไม่แน่ใจว่ามีความสอดคล้อง ให้ค่า 0
ถ้าแน่ใจว่าไม่สอดคล้อง ให้ค่า -1
N แทนจำนวนผู้เชี่ยวชาญ

3. พิจารณาน้ำหนักความสำคัญของมวลความรู้ย่อย เพื่อใช้เป็นเกณฑ์ในการกำหนดจำนวนข้อสอบที่เหมาะสมของแต่ละมวลความรู้ย่อย โดยคิดจากมวลความรู้หลักเป็นร้อยละ 100

4. สร้างลักษณะเฉพาะข้อสอบ(Item specification) จากมวลความรู้ย่อยที่กำหนดไว้ เพื่อกำหนดโครงสร้างของข้อสอบรายข้อ คือ กำหนดกรอบลักษณะคำถาม ที่มาของตัวลวงแต่ละตัว ทำการตรวจสอบความเหมาะสมของลักษณะเฉพาะข้อสอบโดยผู้เชี่ยวชาญด้านการสอนคณิตศาสตร์และด้านวัดผล 10 ท่าน แล้วทำการปรับปรุงพัฒนาตามข้อเสนอแนะ

5. สร้างข้อสอบตามโครงสร้างของลักษณะเฉพาะข้อสอบ จำนวน 100 ข้อ ซึ่งคิดเป็น 2 เท่าของจำนวนที่ใช้ในการศึกษา แล้วนำไปให้ผู้เชี่ยวชาญด้านคณิตศาสตร์และด้านวัดผล จำนวน 10 ท่าน ตรวจสอบความเที่ยงตรงเชิงเนื้อหา ความเป็นปรนัยของข้อคำถาม โดยวิธีการตัดสินความสอดคล้องด้วยการหาค่าดัชนีความสอดคล้อง ถ้าค่าเฉลี่ยของคะแนนการตัดสินของผู้เชี่ยวชาญในแต่ข้อมีค่าสูงกว่า 0.5 แสดงว่าข้อสอบข้อนั้นวัดตามคุณลักษณะเฉพาะที่ข้อสอบจากมวลความรู้ย่อยที่กำหนดไว้ และนำมาปรับปรุง พัฒนา

6. จัดข้อสอบออกเป็น 2 ฉบับ ๆละ 50 ข้อ เพื่อให้แบบทดสอบมีจำนวนข้อสอบที่เหมาะสมกับเวลาและความตั้งใจของผู้สอบ โดยใช้วิธีการสุ่มข้อสอบที่สร้างจากลักษณะเฉพาะข้อสอบตามมวลความรู้ย่อยเดียวกันออกเป็น 2 กลุ่มให้ได้จำนวนข้อตามที่กำหนดน้ำหนักความสำคัญไว้ในข้อ 3

7. ทดลองใช้แบบทดสอบครั้งที่ 1 โดยนำแบบทดสอบทั้ง 2 ฉบับที่ได้รับการปรับปรุงพัฒนาแล้วไปทดลองใช้กับนักเรียนชั้นประถมศึกษาปีที่ 5 ที่ไม่ใช่กลุ่มตัวอย่าง จำนวนฉบับละ 200 คน โดยทำการแจกแบบทดสอบสลับกันในแต่ห้องสอบ เพื่อให้แบบทดสอบแต่ละฉบับได้รับการทดสอบจากผู้สอบที่มีความสามารถใกล้เคียงกัน

8. วิเคราะห์คุณภาพข้อสอบตามทฤษฎีการทดสอบแบบมาตรฐานเดิม เพื่อคัดเลือกและปรับปรุงพัฒนาข้อสอบให้ได้ข้อสอบที่มีคุณภาพตามเกณฑ์ จำนวน 60 ข้อ โดยพิจารณาจากค่าความยากง่ายอยู่ในช่วง 0.2 ถึง 0.8 ค่าอำนาจจำแนกมีค่าตั้งแต่ +0.2 ขึ้นไป รวมทั้งพิจารณาเวลาที่เหมาะสมสำหรับการทำแบบทดสอบแต่ละฉบับโดยคิดจากเวลาที่นักเรียนส่วนใหญ่ประมาณร้อยละ 90 ทำข้อสอบเสร็จ โดยมีสูตรการคำนวณดังนี้

8.1 ค่าความยากง่าย ใช้วิธีการวิเคราะห์แบบกลุ่ม โดยทำการทดสอบเพียงครั้งเดียว แล้วใช้สูตรคำนวณค่าความยากง่ายดังนี้

$$p = \frac{R}{N}$$

เมื่อ p แทน ค่าความยากง่ายของข้อสอบ

R แทน จำนวนคนที่ตอบข้อนั้นถูก

N แทน จำนวนคนทั้งหมด

8.2 ค่าอำนาจจำแนก ใช้วิธีการวิเคราะห์แบบพอยท์ไบซีเรียล แล้วใช้สูตรคำนวณค่าความยากง่ายดังนี้(กนกทิพย์ พัฒนาพัวพันธ์. 2541: 14)

$$r_{pbis} = \left[\frac{\bar{X}_p - \bar{X}_q}{SD_X} \right] \sqrt{pq}$$

เมื่อ r_{pbis} แทน ค่าอำนาจจำแนกของข้อสอบ

$\bar{X}_p$ แทน ค่าเฉลี่ยของคนในกลุ่มตอบถูก

$\bar{X}_q$ แทน ค่าเฉลี่ยของคนในกลุ่มตอบผิด

SD_X แทน ส่วนเบี่ยงเบนมาตรฐานของคะแนนทั้งหมด

p แทน สัดส่วนของคนตอบถูก

q แทน สัดส่วนของคนตอบผิด

9. ทดลองใช้แบบทดสอบครั้งที่ 2 โดยนำแบบทดสอบที่ปรับปรุงแก้ไขแล้วจากการทดลองใช้ครั้งที่ 1 จำนวน 60 ข้อ ไปสอบกับนักเรียนชั้นประถมศึกษาปีที่ 5 ที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 1,000 คน

10. ตรวจสอบความเป็นมิติเดียว(Unidimensional) ด้วยการวิเคราะห์องค์ประกอบเชิงสำรวจ(Exploratory Factor Analysis) โดยสกัดองค์ประกอบด้วยวิธีวิเคราะห์องค์ประกอบหลัก (Principle Component) ด้วยโปรแกรม SPSS พิจารณาค่าความแปรปรวนสูงสุดจากองค์ประกอบหลักตัวแรกมีสัดส่วนมากกว่าองค์ประกอบอื่นๆมาก และวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) ที่มีตัวแปรภายนอกแฝงเพียงตัวเดียว เพื่อทดสอบความสอดคล้องระหว่างโมเดลกับข้อมูลเชิงประจักษ์ ด้วยโปรแกรม LISREL

11. วิเคราะห์ค่าความยาก อำนาจจำแนก การเดาและฟังก์ชั่นสารสนเทศของข้อสอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าและค่าฟังก์ชั่นสารสนเทศของแบบทดสอบตามทฤษฏีการตอบข้อสอบ(IRT) โดยพิจารณาค่าความยาก(b)อยู่ระหว่าง -2.5 ถึง +2.5 ค่าอำนาจจำแนก(a)มากกว่า 0.8 และค่าการเดา(c)น้อยกว่า 0.3 ด้วยโมเดลโลจีสติก 3 พารามิเตอร์จากโปรแกรม BOLOG-MG3 เพื่อคัดเลือกข้อสอบที่มีคุณภาพจำนวน 50 ข้อ

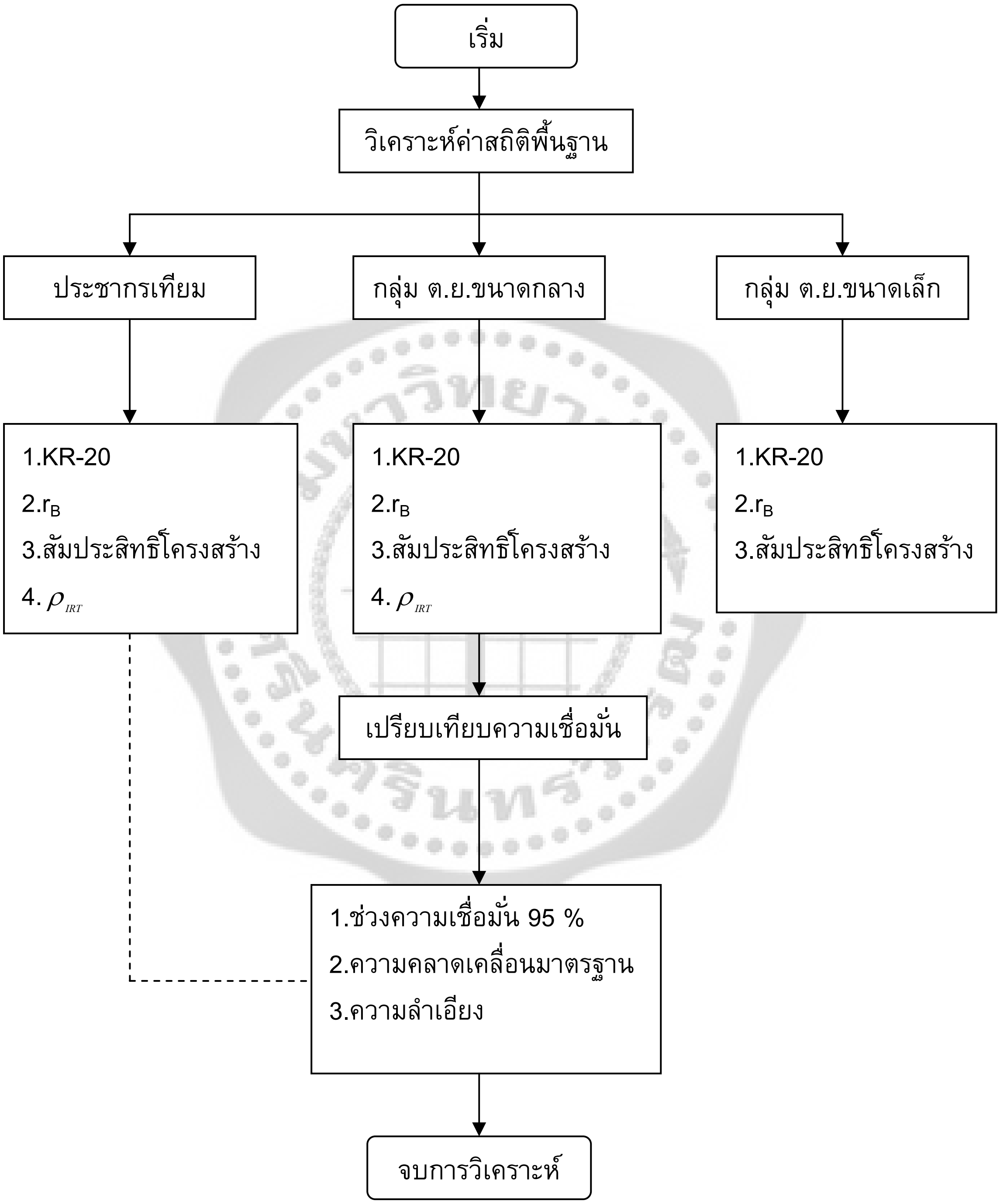
12. ได้แบบทดสอบฉบับสมบูรณ์จำนวน 50 ข้อ เพื่อนำไปทดสอบกับประชากรเทียม

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลสำหรับการศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และโมเดลโลจีสติก 3 พารามิเตอร์ในครั้งนี้ ได้กำหนดการเก็บรวบรวมข้อมูลด้วยการนำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 ไปทดสอบกับประชากรเทียม ปีการศึกษา 2550 จำนวน 6,000 คน ในช่วงเดือนธันวาคม 2550 โดยทำการเก็บเพียงครั้งเดียว

การจัดกระทำข้อมูลและการวิเคราะห์ข้อมูล

การจัดกระทำและการวิเคราะห์ข้อมูลได้ดำเนินการดังนี้



1. วิเคราะห์ค่าสถิติบรรยายของคะแนนสอบ ได้แก่ ค่าเฉลี่ย มัธยฐาน ฐานนิยม ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ความโด่ง คะแนนต่ำสุด และคะแนนสูงสุด โดยใช้โปรแกรม SPSS

2. ประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากประชากรเทียม

3. ประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง

4. ประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง จากกลุ่มตัวอย่างขนาดเล็ก

5. ศึกษาความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT}

6. เปรียบเทียบความเชื่อมั่นจากการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 ด้วยสัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} ที่กลุ่มตัวอย่างขนาดกลาง

7. สถิติที่ใช้ในการวิเคราะห์ข้อมูล

7.1 ค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสมมูล ด้วยสัมประสิทธิ์ KR-20 วิเคราะห์ข้อมูลโดยใช้โปรแกรม Formula r_B จากสูตร

$$r_{tt} = \frac{k}{k-1}\left[1 - \frac{\sum pq}{S_x^2}\right]$$

เมื่อ k แทน จำนวนข้อของเครื่องมือวัด

p แทน สัดส่วนของผู้ตอบถูกหรือความยากของแต่ละข้อ

q แทน สัดส่วนของผู้ตอบผิด ซึ่งเท่ากับ 1-p

S_X^2 แทน ความแปรปรวนของคะแนนรวมทั้งฉบับของเครื่องมือวัด

7.2 ค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ด้วยสัมประสิทธิ์ r_B วิเคราะห์ข้อมูลโดยใช้โปรแกรม Formula r_B จากสูตร

$$r_B = \frac{1}{1 - \sum \lambda_i^2}\left[1 - \frac{\sum P_i(1-P_i)}{S_X^2}\right]$$

เมื่อ P_i แทน ค่าความยากง่ายของข้อสอบแต่ละข้อ

S_X^2 แทน ความแปรปรวนของคะแนนทั้งฉบับ

B_{ji} แทน สัดส่วนจำนวนผู้ตอบสองข้อใด ๆ ตรงกัน

$\lambda_1 = \left[\sum B_{1i} - P_1\left(\overline{X} - 1\right)\right] / S_x^2$

$\lambda_2 = \left[\sum B_{2i} - P_2\left(\overline{X} - 1\right)\right] / S_x^2$

| |

$\lambda_K = \left[\sum B_{Ki} - P_K\left(\overline{X} - 1\right)\right] / S_x^2$

$\overline{X} = \sum P_i$

7.3 ค่าความเชื่อมั่นของสัมประสิทธิ์โครงสร้าง จากโมเดลการวัดคะแนนจริงสัมพันธ์ วิเคราะห์ข้อมูลโดยใช้โปรแกรม LISREL ในการวิเคราะห์องค์ประกอบเชิงยืนยันโมเดลการทดสอบคะแนนจริงสัมพันธ์ที่วัดองค์ประกอบเดี่ยว แล้วนำค่าน้ำหนักองค์ประกอบมาประมาณค่าความเชื่อมั่นตามวิธีการของเรย์คอฟ(Raykov. 2003: 143-159) จากสูตร

$$\rho_{Y,K} = \frac{\left(\sum_{i=1}^{k} b_i\right)^2}{\left(\sum_{i=1}^{k} b_i\right)^2 + \sum_{i=1}^{k} \theta_i}$$

เมื่อ b_i แทน น้ำหนักองค์ประกอบ(LX) ที่ได้จากการประมาณค่าด้วยวิธีความน่าจะเป็นสูงสุด(Maximum Likelihood)

θ_i แทน น้ำหนักองค์ประกอบความแปรปรวนของความคลาดเคลื่อน

7.4 ค่าความเชื่อมั่นของสัมประสิทธิ์ ρ_{IRT} จากโมเดลโลจีสติก 3 พารามิเตอร์ วิเคราะห์ข้อมูลโดยใช้โปรแกรม BILOG-MG3 เพื่อประมาณค่าพารามิเตอร์ค่าอำนาจจำแนก(a) ค่าความยาก (b) และค่าการเดา(c) แล้วนำค่าพารามิเตอร์ทั้ง 3 ตัวมาวิเคราะห์ดดยใช้โปรแกรม SPSS ในการหาค่าความแปรปรวนของคะแนนจริง และความแปรปรวนของความคลาดเคลื่อนแล้วนำมาประมาณค่าความเชื่อมั่น โดยทำการประมาณค่าความเชื่อมั่นตามวิธีการของดิมิทรอฟ(Dimitrov. 2003: 440) จากสูตร

$$\rho_{IRT} = \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma_e^2}$$

เมื่อ σ_τ^2 แทน ความแปรปรวนของคะแนนจริง

σ_e^2 แทน ความแปรปรวนของความคลาดเคลื่อน

โดย

$$\sigma_{\tau}^{2}=\sum_{i=1}^{n}\sum_{j=1}^{n}\sqrt{\left[\pi_{i}^{*}\left(1-\pi_{i}^{*}\right)-\sigma^{2}(e_{i}^{*})\right]\left[\pi_{j}^{*}\left(1-\pi_{j}^{*}\right)-\sigma^{2}(e_{j}^{*})\right]}$$

$$\pi=\sum_{i=1}^{n}\pi_{i}$$

$$\sigma_{e}^{2}=\sum_{i=1}^{n}\sigma^{2}(e_{i}^{*})$$

7.5 ช่วงความเชื่อมั่น 95% เป็นการคำนวณช่วงของความถูกต้องของการประมาณค่า เพื่อดูว่าค่าความเชื่อมั่นที่ประมาณค่าได้จาก 4 สูตร ที่กลุ่มตัวอย่างขนาดกลาง 30 กลุ่ม มีค่าอยู่ในช่วงของความเชื่อมั่น 95% หรือไม่ หากค่าที่ประมาณได้มีค่าอยู่ในช่วงนี้ก็แสดงว่าสูตรนั้นสามารถประมาณค่าได้ถูกต้อง มีสูตรดังนี้

$$CI=\rho_{T(r_i)}\pm 1.96SD(r_i)$$

เมื่อ CI แทน ช่วงความเชื่อมั่น 95%

$\rho_{T(r_i)}$ แทน ค่าความเชื่อมั่นที่แท้จริงของสูตร i โดยคำนวณจากประชากรเทียม

1.96 แทน ค่าคงที่ของการแจกแจงแบบที (t- distribution) ที่ .05

$SD(r_i)$ แทน ส่วนเบี่ยงเบนมาตรฐานของค่าเฉลี่ยความเชื่อมั่น ของสูตร i จากกลุ่มตัวอย่าง 30 กลุ่ม

7.6 ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่น(Standard error of estimate: SE) พิจารณาค่า SE ที่มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสูตรประมาณค่าความเชื่อมั่น มีความเหมาะสมและมีคุณภาพในการประมาณค่า โดยปรับจากสูตรของพาร์แชล เฮาทตัน ครอมเรย์ (Parshall, Houghton, Kromrey. 1995: 37-54) ดังนี้

$$SE(r_i)=\sqrt{\frac{\sum\left(r_{ij}-\bar{r}_i\right)^2}{N-1}}$$

เมื่อ $SE(r_i)$ แทน ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นของสูตร i

r_{ij} แทน ค่าความเชื่อมั่นที่คำนวณได้จากสูตรที่ i จากกลุ่มตัวอย่างที่ j $(j=1,\ldots,30)$

$\bar{r}_i$ แทน ค่าความเชื่อมั่นเฉลี่ยจากสูตร i ที่คำนวณจากกลุ่มตัวอย่างขนาดกลาง 30 กลุ่ม

N แทน จำนวนกลุ่มตัวอย่าง 30 กลุ่ม

7.7 ค่าความลำเอียง(Bias) เป็นค่าความลำเอียงเฉลี่ยที่คิดเครื่องหมาย ซึ่งเป็นการบอกทิศทางของความลำเอียง ค่าความลำเอียงมีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าวิธีประมาณค่าความเชื่อมั่น 4 วิธี มีความเหมาะสมและมีคุณภาพในการประมาณค่า โดยปรับจากสูตรของพาร์แชล เฮาทตัน ครอมเรย์(Parshall, Houghton, Kromrey. 1995: 37-54) ดังนี้

$$Bias = \bar{r}_i - \rho_{T(r_i)}$$

เมื่อ $\bar{r}_i$ แทน ค่าความเชื่อมั่นเฉลี่ยจากสูตร i ที่คำนวณคำนวณจากกลุ่มตัวอย่างขนาดกลาง 30 กลุ่ม

$\rho_{T(r_i)}$ แทนค่าความเชื่อมั่นที่แท้จริงของสูตร i โดยคำนวณจากประชากรเทียม

7.8 สถิติทดสอบนัยสำคัญของความแตกต่างของค่าความเชื่อมั่นเฉลี่ยหลายค่า ซึ่งประมาณจากกลุ่มตัวอย่างขนาดกลางหลายกลุ่ม ใช้วิธีการวิเคราะห์เส้นภาพ(Profile Analysis) ทำการทดสอบดังนี้

7.8.1 การทดสอบความคู่ขนาน(Test of Parallelism) เป็นการทดสอบว่ากลุ่มตัวอย่างแต่ละขนาดมีเส้นภาพความเชื่อมั่นคู่ขนานกันหรือไม่ หรือความแตกต่างของค่าความเชื่อมั่นที่คำนวณจากแต่ละสูตรของกลุ่มตัวอย่างแต่ละขนาดจะเท่ากันหรือไม่ โดยใช้ค่าแลมบ์ดาของวิลค์(Wills' Lambda : Λ) ทดสอบ ด้วยสูตรดังนี้

$$Approximate\ F(df_1, df_2) = \left[\frac{1-y}{y}\right]\left[\frac{df_1}{df_2}\right]$$

$$\Lambda = \frac{|S_{wg}|}{|S_{wg} + S_{bg}|}$$

เมื่อ Λ แทน ค่าแลมบ์ดาของวิลค์

S_{wg} แทน ผลบวกกำลังสองภายในขนาดกลุ่มตัวอย่าง

S_{bg} แทน ผลบวกกำลังสองระหว่างขนาดกลุ่มตัวอย่าง

p แทน จำนวนตัวแปรตาม

โดย

$y = \Lambda^{1/2}$

$df_1 = p(df_{bg})$

$$df_2 = S\left[\left(df_{wg}\right) - \frac{p - df_{bg} + 1}{2}\right] - \left[\frac{p\left(df_{bg}\right) - 2}{2}\right]$$

$$S = \sqrt{\frac{P^2\left(df_{bg}\right)^2 - 4}{p^2\left(df_{bg}\right)^2 - 5}}$$

7.8.2 การทดสอบระดับ(Level Test) เป็นการทดสอบกลุ่มตัวอย่างแต่ละขนาดจะให้เส้นภาพ(Profile) ของความเชื่อมั่นอยู่ในระดับเดียวกันหรือไม่ หรือทดสอบว่าค่าความเชื่อมั่นเฉลี่ยที่คำนวณจากกลุ่มตัวอย่างขนาดหนึ่งจะแตกต่างไปจากกลุ่มตัวอย่างขนาดอื่นๆหรือไม่ ใช้สูตรดังนี้

$$F = \frac{MS_{bg}}{MS_{wg}}, \qquad df = (k-1),(N-k)$$

เมื่อ MS_{bg} แทน ค่าเฉลี่ยกำลังสองระหว่างขนาดกลุ่มตัวอย่าง

MS_{wg} แทน ค่าเฉลี่ยกำลังสองภายในขนาดกลุ่มตัวอย่าง

k แทน จำนวนขนาดกลุ่มตัวอย่าง

N แทน จำนวนกลุ่มย่อยทั้งหมด

7.8.3 การทดสอบความเป็นระนาบ(Test of Flatness) เป็นการทดสอบว่าค่าความเชื่อมั่นที่คำนวณได้แต่ละสูตรโดยเฉลี่ยมีค่าเท่ากันหรือไม่ การทดสอบความเป็นระนาบนี้จะทดสอบหลังจากที่พบว่าเส้นภาพมีความคู่ขนานกันก่อนเท่านั้น โดยใช้สูตรดังนี้

$$F = \frac{N - k - p + 2}{p - 1}\left[\frac{1 - \Lambda}{\Lambda}\right]; \qquad df = (p-1),(N-k-p+2)$$

เมื่อ k แทน จำนวนขนาดกลุ่มตัวอย่าง

p แทน จำนวนตัวแปรตาม

Λ แทน ค่าแลมป์ดาของวิลค์

N แทน จำนวนกลุ่มย่อยทั้งหมด

7.8.4 การทดสอบนัยสำคัญของความแตกต่างของความเชื่อมั่นเฉลี่ยรายคู่ ซึ่งประมาณค่าจากกลุ่มตัวอย่างขนาดกลาง 30 กลุ่ม ตามวิธีการของเชฟเฟ่(Scheffe') แบบวัดซ้ำ โดยมีสูตรดังนี้(Hays. 1988 : 401-531)

$$F = \frac{SS_{(\tilde{\psi}_g)}}{MS_{resg}} \qquad df = (j-1),(n-1)$$

เมื่อ $SS_{(\tilde{\psi}_g)}$ แทน ผลบวกกำลังสองของการเปรียบเทียบ

MS_{resg} แทน ค่าเฉลี่ยกำลังสองของส่วนที่เหลือ(Residual)

บทที่ 4
ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีความมุ่งหมายหลักเพื่อศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ ซึ่งประกอบด้วยสัมประสิทธิ์ r_B กับสัมประสิทธิ์โครงสร้าง และโมเดลโลจีสติก 3 พารามิเตอร์ ได้แก่ สัมประสิทธิ์ ρ_{IRT} และใช้สัมประสิทธิ์ KR-20 เป็นเกณฑ์เทียบ ในครั้งนี้ ได้กำหนดสัญลักษณ์และอักษรย่อต่าง ๆดังต่อไปนี้

สัญลักษณ์และอักษรย่อ

สัญลักษณ์และอักษรย่อที่ใช้แทนความหมายในการวิเคราะห์ข้อมูล มีดังนี้

r_{20}	แทน	ค่าความเชื่อมั่นที่คำนวณจากสูตรสัมประสิทธิ์ KR-20
r_B	แทน	ค่าความเชื่อมั่นที่คำนวณจากสูตรสัมประสิทธิ์ r_B
$r_{construct}$	แทน	ค่าความเชื่อมั่นที่คำนวณจากสูตรสัมประสิทธิ์โครงสร้าง
ρ_{IRT}	แทน	ค่าความเชื่อมั่นที่คำนวณจากสูตรสัมประสิทธิ์ ρ_{IRT}
$\bar{x}$	แทน	ค่าเฉลี่ย
$S.E._{mean}$	แทน	ค่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย
SD	แทน	ค่าส่วนเบี่ยงเบนมาตรฐาน
Sk	แทน	ค่าความเบ้
$S.E._{Sk}$	แทน	ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้
Ku	แทน	ค่าความโด่ง
$S.E._{Ku}$	แทน	ค่าความคลาดเคลื่อนมาตรฐานของค่าโด่ง
Sig	แทน	นัยสำคัญทางสถิติ

ผลการวิเคราะห์ข้อมูล

ผู้วิจัยได้นำเสนอผลการวิเคราะห์ข้อมูลตามลำดับความมุ่งหมายของการวิจัยแบ่งออกเป็น 4 ตอนดังนี้

ตอนที่ 1 ค่าสถิติบรรยายของแบบทดสอบที่ได้จากกลุ่มประชากรเทียมและกลุ่มตัวอย่างแต่ละขนาดจำนวน 8 ขนาด

ตอนที่ 2 การประมาณค่าความเชื่อมั่นของแบบทดสอบที่คำนวณจากสูตร 4 สูตร

ตอนที่ 2.1 การคำนวณจากประชากรเทียม

ตอนที่ 2.2 การคำนวณจากกลุ่มตัวอย่างขนาดกลาง 4 ขนาด

ตอนที่ 2.3 การคำนวณจากกลุ่มตัวอย่างขนาดเล็ก 4 ขนาด

ตอนที่ 3 ความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณจากสูตร 4 สูตร

ตอนที่ 4 การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่ประมาณค่าที่คำนวณจากสูตร 4 สูตร จากกลุ่มตัวอย่างขนาดกลาง

โดยมีรายละเอียดดังต่อไปนี้

ตอนที่ 1 ค่าสถิติบรรยายของแบบทดสอบที่ได้จากกลุ่มประชากรเทียมและกลุ่มตัวอย่างแต่ละขนาด จำนวน 8 ขนาด

ผู้วิจัยนำคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 ที่เก็บรวบรวมมาได้จากกลุ่มตัวอย่างขนาดใหญ่ที่ผู้วิจัยกำหนดให้เป็นประชากรเทียม จำนวน 6,000 คน จากนั้นทำการสุ่มแบบใส่คืนเพื่อแบ่งเป็นกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ ขนาด 300 500 1,000 และ 1,500 คน แต่ละขนาดมีจำนวน 30 กลุ่ม และกลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ ขนาด 30 60 120 และ 240 คน แต่ละขนาดมีจำนวน 30 กลุ่ม มาหาค่าเฉลี่ย ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้ ความโด่ง ค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง โดยผู้วิจัยได้นำเสนอในตารางที่ 1 ถึง 3 ดังนี้

ตาราง 1 ค่าเฉลี่ย ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้ ความโด่ง ค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง ของแบบทดสอบ จากประชากรเทียม

สถิติบรรยาย	ค่าที่ได้
ค่าเฉลี่ย ($\bar{x}$)	28.519
ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย($S.E._{mean}$)	.134
ส่วนเบี่ยงเบนมาตรฐาน (SD)	10.341
ความเบ้ (Sk)	-.049
ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้($S.E._{Sk}$)	.032
ความโด่ง (Ku)	-1.023
ค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง($S.E._{Ku}$)	.063

ผลการวิเคราะห์ตามตาราง 1 พบว่าค่าสถิติบรรยายของแบบทดสอบที่ได้จากประชากรเทียม มีค่าค่าเฉลี่ย เท่ากับ 28.519 ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย เท่ากับ .134 ส่วน

เบี่ยงเบนมาตรฐาน เท่ากับ 10.341 ความเบ้เท่ากับ -.049 ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้ เท่ากับ .032ความโด่ง เท่ากับ -1.023 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่งเท่ากับ .063

ตาราง 2 ค่าเฉลี่ย ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้ ความโด่ง ค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่งของแบบทดสอบ จากกลุ่มตัวอย่างขนาดกลาง

สถิติบรรยาย	1,500	1,000	500	300
$\bar{x}$	28.230 - 29.064	27.997 - 29.122	27.688 - 29.594	27.347 - 29.410
$S.E._{mean}$	.262 - .274	.319 - .333	.441 - .491	.557 - .630
SD	10.107 - 10.603	10.071 - 10.530	9.860 - 10.983	9.643 - 10.914
Sk	-.139 - .017	-.133 - .018	-.191 - .190	-.211 - .910
$S.E._{Sk}$	.063	.077	.109	.141
Ku	-1.098 – (-.962)	-1.113 – (-.974)	-1.101 – (-.898)	-1.174 – (-.804)
$S.E._{Ku}$	.126	.155	.218	.281

ผลการวิเคราะห์ตามตาราง 2 พบว่าค่าสถิติบรรยายของแบบทดสอบที่ได้จากกลุ่มตัวอย่างขนาด 1,500 คน มีค่าค่าเฉลี่ยระหว่าง 28.230 ถึง 29.064 ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย มีค่าระหว่าง .262 ถึง .274 ส่วนเบี่ยงเบนมาตรฐาน มีค่าระหว่าง 10.107 ถึง 10.603 ความเบ้ มีค่าระหว่าง -.139 ถึง .017 ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้ มีค่าเท่ากับ .063 ความโด่ง มีค่าระหว่าง -1.098 ถึง -.962 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง มีค่าเท่ากับ .126

กลุ่มตัวอย่างขนาด 1,000 คน มีค่าค่าเฉลี่ยระหว่าง 27.997 ถึง 29.122 ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย มีค่าระหว่าง .319 ถึง .333 ส่วนเบี่ยงเบนมาตรฐาน มีค่าระหว่าง 10.071 ถึง 10.530 ความเบ้ มีค่าระหว่าง -.133 ถึง .018 ความคลาดเคลื่อนมาตรฐานของค่าความเบ้ มีค่าเท่ากับ .077 ความโด่ง มีค่าระหว่าง -1.113 ถึง -.974 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง มีค่าเท่ากับ .155

กลุ่มตัวอย่างขนาด 500 คน มีค่าค่าเฉลี่ย ระหว่าง 27.688 ถึง 29.594 ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย มีค่าระหว่าง .441 ถึง .491 ส่วนเบี่ยงเบนมาตรฐาน มีค่าระหว่าง 9.860 ถึง 10.983 ความเบ้ มีค่าระหว่าง -.191 ถึง .190 ความคลาดเคลื่อนมาตรฐานของค่าความเบ้มีค่า เท่ากับ .109 ความโด่ง มีค่าระหว่าง -1.101 ถึง -.898 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง มีค่าเท่ากับ .218

กลุ่มตัวอย่างขนาด 300 คน มีค่าค่าเฉลี่ย ระหว่าง 27.347 ถึง 29.410 ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย มีค่าระหว่าง .557 ถึง .630 ส่วนเบี่ยงเบนมาตรฐาน มีค่าระหว่าง 9.643 ถึง 10.914 ความเบ้ มีค่าระหว่าง -.211 ถึง .091 ความคลาดเคลื่อนมาตรฐานของค่าความเบ้มีค่า เท่ากับ .141 ความโด่ง มีค่าระหว่าง -1.174 ถึง -.804 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง มีค่าเท่ากับ .281

ตาราง 3 ค่าเฉลี่ย ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ความคลาดเคลื่อนมาตรฐานของค่าความเบ้ ความโด่ง ค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่งของแบบทดสอบ จากกลุ่มตัวอย่างขนาดเล็ก

สถิติบรรยาย	240	120	60	30
$\bar{x}$	27.279 - 30.604	25.192 – 29.292	25.467 – 29.500	23.833 – 30.800
$S.E._{mean}$	.640 - .702	.840 – 1.004	1.153 – 1.421	1.574 – 2.183
SD	9.919 – 10.871	9.198 – 10.994	8.934 – 11.008	8.619 – 11.954
Sk	-.265 - .134	-.313 - .141	-.661 - .119	-.418 – 436
$S.E._{Sk}$	.157	.221	.309	.427
Ku	-1.191 – (-.895)	-1.137 – (-.713)	-1.341 – (-.240)	-1.488 – (-.122)
$S.E._{Ku}$	.313	.438	.608	.833

ผลการวิเคราะห์ตามตาราง 3 พบว่าค่าสถิติบรรยายของแบบทดสอบที่ได้จากกลุ่มตัวอย่างขนาด 240 คน มีค่าค่าเฉลี่ย ระหว่าง 27.279 ถึง 30.ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย มีค่าระหว่าง .640 ถึง .702 ส่วนเบี่ยงเบนมาตรฐาน มีค่าระหว่าง 9.919 ถึง 10.871 ความเบ้ มีค่าระหว่าง -.265 ถึง .134 ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้ มีค่าเท่ากับ .157 ความโด่ง มีค่าระหว่าง -1.191 ถึง -.895 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง มีค่าเท่ากับ .313

กลุ่มตัวอย่างขนาด 120 คน มีค่าค่าเฉลี่ย ระหว่าง 25.192 ถึง 29.292 ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย มีค่าระหว่าง .840 ถึง 1.004 ส่วนเบี่ยงเบนมาตรฐาน มีค่าระหว่าง 9.198 ถึง 10.994 ความเบ้ มีค่าระหว่าง -.313 ถึง .141 ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้ มีค่าเท่ากับ .221 ความโด่ง มีค่าระหว่าง -1.137 ถึง -.713 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง มีค่าเท่ากับ .438

กลุ่มตัวอย่างขนาด 60 คน มีค่าค่าเฉลี่ย ระหว่าง 25.467 ถึง 29.500 ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย มีค่าระหว่าง 1.153 ถึง 1.421 ส่วนเบี่ยงเบนมาตรฐาน มีค่าระหว่าง 8.934 ถึง 11.008 ความเบ้ มีค่าระหว่าง -.661 ถึง .119 ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้

มีค่าเท่ากับ .309 ความโด่ง มีค่าระหว่าง -1.341 ถึง -.240 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง มีค่าเท่ากับ .608

กลุ่มตัวอย่างขนาด 30 คน มีค่าค่าเฉลี่ย ระหว่าง 23.833 ถึง 30.800 ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย มีค่าระหว่าง 1.574 ถึง 2.183 ส่วนเบี่ยงเบนมาตรฐาน มีค่าระหว่าง 8.619 ถึง 11.954 ความเบ้ มีค่าระหว่าง -.418 ถึง .436 ค่าความคลาดเคลื่อนมาตรฐานของค่าความเบ้ มีค่าเท่ากับ .427 ความโด่ง มีค่าระหว่าง -1.488 ถึง -.122 และค่าความคลาดเคลื่อนมาตรฐานของค่าความโด่ง มีค่าเท่ากับ .833

ตอนที่ 2 การประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณจากสูตร 4 สูตร

ตอนที่ 2.1 การคำนวณจากประชากรเทียม

ผู้วิจัยนำคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 ที่เก็บรวบรวมมาได้จากกลุ่มตัวอย่างขนาดใหญ่ที่ผู้วิจัยกำหนดให้เป็นประชากรเทียม จำนวน 6,000 คน มาประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B จากโปรแกรม Formula r_B สำหรับสัมประสิทธิ์โครงสร้างนั้นได้ทำการวิเคราะห์ จากการวิเคราะห์องค์ประกอบเชิงยืนยันโมเดลการทดสอบคะแนนจริงสัมพันธ์ที่วัดองค์ประกอบเดี่ยว แล้วนำค่าน้ำหนักองค์ประกอบมาประมาณค่าความเชื่อมั่นตามวิธีการของเรย์คอฟ(Raykov. 2003: 143-159) ส่วนสัมประสิทธิ์ ρ_{IRT} จากโมเดลโลจีสติก 3 พารามิเตอร์ ผู้วิจัยวิเคราะห์ข้อมูลโดยใช้โปรแกรม BILOG-MG3 เพื่อประมาณค่าพารามิเตอร์ได้แก่ค่าอำนาจจำแนก(a) ค่าความยาก(b) และค่าการเดา(c) แล้วนำค่าพารามิเตอร์ทั้ง 3 ตัวมาวิเคราะห์โดยใช้โปรแกรม SPSS ในการหาค่าความแปรปรวนของคะแนนจริง และความแปรปรวนของความคลาดเคลื่อนแล้วนำค่าที่ได้มาประมาณค่าความเชื่อมั่น โดยทำการประมาณค่าความเชื่อมั่นตามวิธีการของดิมิทรอฟ(Dimitrov. 2003: 440) ผู้วิจัยได้นำเสนอในตาราง 4 ดังนี้

ตาราง 4 ค่าความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากประชากรเทียม

การประมาณค่าความเชื่อมั่นของแบบทดสอบ	ค่าความเชื่อมั่น
สัมประสิทธิ์ r_{20}	.910
สัมประสิทธิ์ r_B	.911
สัมประสิทธิ์ $r_{construct}$	.908
สัมประสิทธิ์ ρ_{IRT}	.974

ผลการวิเคราะห์ตามตาราง 4 พบว่า ค่าความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ ρ_{IRT} ได้ค่าความเชื่อมั่นสูงสุด รองลงมาคือ สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ตอนที่ 2.2 การคำนวณจากกลุ่มตัวอย่างขนาดกลาง 4 ขนาด

ผู้วิจัยนำผลการทดสอบจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 5 จากกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ ขนาด 300 500 1,000 และ 1,500 คน แต่ละขนาดมีจำนวน 30 กลุ่ม มาประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B จากโปรแกรม Formula r_B สำหรับสัมประสิทธิ์โครงสร้างนั้นได้ทำการวิเคราะห์ จากการวิเคราะห์องค์ประกอบเชิงยืนยันโมเดลการทดสอบคะแนนจริงสัมพันธ์ที่วัดองค์ประกอบเดี่ยว แล้วนำค่าน้ำหนักองค์ประกอบที่ได้มาประมาณค่าความเชื่อมั่นตามวิธีการของเรย์คอฟ(Raykov. 2003: 143-159) ส่วนสัมประสิทธิ์ ρ_{IRT} จากโมเดลโลจีสติก 3 พารามิเตอร์ ผู้วิจัยทำการวิเคราะห์ข้อมูลโดยใช้โปรแกรม BILOG-MG3 เพื่อประมาณค่าพารามิเตอร์ ได้แก่ ค่าอำนาจจำแนก(a) ค่าความยาก(b) และค่าการเดา(c) แล้วนำค่าพารามิเตอร์ทั้ง 3 ตัวมาวิเคราะห์โดยใช้โปรแกรม SPSS ในการหาค่าความแปรปรวนของคะแนนจริง และความแปรปรวนของความคลาดเคลื่อนแล้วนำค่าที่ได้มาประมาณค่าความเชื่อมั่น โดยทำการประมาณค่าความเชื่อมั่นตามวิธีการของดิมิทรอฟ(Dimitrov. 2003: 440) และผู้วิจัยได้คำนวณช่วงความเชื่อมั่น 95% ซึ่งเป็นการคำนวณช่วงของความถูกต้องของการประมาณค่าเพื่อดูว่าค่าความเชื่อมั่นที่ประมาณค่าได้จาก 4 สูตร ที่กลุ่มตัวอย่างขนาดกลาง 4 ขนาด แต่ละขนาดมีจำนวน 30 กลุ่ม มีค่าอยู่ในช่วงของความเชื่อมั่น 95% หรือไม่ หากค่าที่ประมาณได้มีค่าอยู่ในช่วงนี้ก็แสดงว่าสูตรนั้นสามารถประมาณค่าได้ถูกต้อง ผู้วิจัยได้นำเสนอในตาราง 5 ถึงตาราง 8 ดังนี้

ตาราง 5 ค่าความเชื่อมั่นและช่วงความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาด 300 คน

กลุ่มที่	r_{20}	r_B	$r_{construct}$	ρ_{IRT}
1	.895	.896	.894	.959
2	.921	.922	.921	.969
3	.921	.922	.921	.969
4	.916	.917	.917	.968
5	.913	.914	.789	.966
6	.911	.912	.909	.963

ตาราง 5 (ต่อ)

กลุ่มที่	r_{20}	r_B	$r_{construct}$	ρ_{IRT}
7	.916	.917	.916	.966
8	.900	.901	.900	.961
9	.910	.911	.894	.964
10	.911	.912	.911	.964
11	.918	.919	.917	.967
12	.909	.910	.910	.964
13	.918	.919	.918	.968
14	.908	.909	.907	.964
15	.909	.909	.908	.965
16	.916	.917	.916	.966
17	.917	.918	.917	.968
18	.906	.907	.906	.964
19	.900	.901	.900	.962
20	.918	.918	.918	.968
21	.919	.919	.918	.968
22	.912	.913	.913	.965
23	.913	.913	.912	.966
24	.907	.908	.907	.964
25	.908	.908	.908	.965
26	.916	.917	.915	.967
27	.906	.907	.905	.964
28	.912	.913	.913	.965
29	.910	.910	.909	.965
30	.918	.919	.918	.966
ช่วงความเชื่อมั่น 95%	.898 - .923	.899 - .922	.862 - .954	.969 - .979

ผลการวิเคราะห์ตามตาราง 5 พบว่า ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 300 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .895 ถึง .921 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .896 ถึง .922 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .789 ถึง .921 และสัมประสิทธิ์ ρ_{IRT} มีค่าอยู่ระหว่าง .959 ถึง .969

สัมประสิทธิ์ ρ_{IRT} ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 1 ค่า ที่ประมาณค่าได้ต่ำกว่าช่วงความเชื่อมั่น 95% และสัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยค่าที่ประมาณได้มีค่าต่ำกว่าช่วงความเชื่อมั่น 95%

ตาราง 6 ค่าความเชื่อมั่นและช่วงความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาด 500 คน

กลุ่มที่	r_{20}	r_B	$r_{construct}$	ρ_{IRT}
1	.915	.915	.914	.969
2	.914	.915	.914	.970
3	.916	.916	.915	.970
4	.907	.908	.906	.966
5	.906	.907	.906	.967
6	.910	.910	.910	.967
7	.915	.915	.915	.970
8	.911	.911	.910	.968
9	.900	.901	.899	.965
10	.911	.911	.910	.968
11	.917	.917	.916	.970
12	.902	.903	.903	.964
13	.915	.916	.914	.970
14	.904	.905	.902	.965
15	.919	.919	.919	.972
16	.915	.916	.915	.970
17	.908	.908	.907	.969
18	.911	.912	.911	.969
19	.912	.913	.911	.969

ตาราง 6 (ต่อ)

กลุ่มที่	$\mathbf{r_{20}}$	$\mathbf{r_B}$	$\mathbf{r_{construct}}$	ρ_{IRT}
20	.910	.911	.909	.968
21	.909	.910	.908	.967
22	.917	.917	.917	.970
23	.908	.909	.907	.968
24	.913	.914	.913	.969
25	.923	.923	.922	.973
26	.910	.911	.909	.968
27	.910	.910	.908	.967
28	.912	.912	.911	.968
29	.905	.905	.904	.966
30	.908	.909	.907	.967
ช่วงความเชื่อมั่น 95%	.900 - .920	.901 - .921	.898 - .918	.970 - .978

ผลการวิเคราะห์ตามตาราง 6 พบว่า ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 500 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .900 ถึง .923 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .901 ถึง .923 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .899 ถึง .922 และสัมประสิทธิ์ ρ_{IRT} มีค่าอยู่ระหว่าง .964 ถึง .973

สัมประสิทธิ์ ρ_{IRT} ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 1 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% สัมประสิทธิ์โครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 2 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% และสัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยค่าที่ประมาณได้มีค่าต่ำกว่าช่วงความเชื่อมั่น 95%

ตาราง 7 ค่าความเชื่อมั่นและช่วงความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาด 1,000 คน

กลุ่มที่	r_{20}	r_B	$r_{construct}$	ρ_{IRT}
1	.912	.913	.912	.972
2	.911	.911	.910	.972
3	.914	.915	.913	.972
4	.906	.907	.9005	.970
5	.910	.911	.909	.971
6	.906	.907	.905	.969
7	.913	.913	.912	.972
8	.909	.910	.908	.971
9	.907	.907	.905	.970
10	.905	.906	.904	.969
11	.910	.911	.909	.971
12	.909	.910	.909	.971
13	.911	.911	.909	.971
14	.913	.913	.912	.972
15	.914	.914	.913	.972
16	.911	.912	.909	.971
17	.909	.910	.909	.970
18	.905	.905	.903	.969
19	.913	.914	.911	.972
20	.905	.906	.904	.969
21	.909	.910	.907	.970
22	.909	.910	.908	.970
23	.911	.912	.910	.972
24	.906	.907	.904	.970
25	.912	.913	.911	.972
26	.913	.913	.912	.973
27	.913	.914	.913	.971
28	.913	.913	.912	.972
29	.912	.912	.910	.972

ตาราง 7 (ต่อ)

กลุ่มที่	r20	rB	rconstruct	ρ_{IRT}
30	.908	.909	.907	.970
ช่วงความเชื่อมั่น 95%	.904 - .916	.905 - .916	.901 - .915	.972 - .976

ผลการวิเคราะห์ตามตาราง 7 พบว่า ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 1,000 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .905 ถึง .914 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .905 ถึง .915 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .903 ถึง .913 และสัมประสิทธิ์ ρ_{IRT} มีค่าอยู่ระหว่าง .969 ถึง .973

สัมประสิทธิ์ ρ_{IRT} ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง ค่าที่ประมาณได้ทุกค่าอยู่ในช่วงความเชื่อมั่น 95% ส่วนสัมประสิทธิ์ ρ_{IRT} ค่าที่ประมาณได้มีค่าที่ไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยมีบางค่าที่ได้ต่ำกว่าช่วงความเชื่อมั่น 95%

ตาราง 8 ค่าความเชื่อมั่นและช่วงความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาด 1,500 คน

กลุ่มที่	r_{20}	r_B	$r_{construct}$	ρ_{IRT}
1	.906	.907	.904	.971
2	.909	.910	.908	.972
3	.913	.914	.912	.974
4	.908	.908	.906	.972
5	.906	.907	.905	.971
6	.910	.910	.909	.972
7	.911	.911	.908	.972
8	.908	.909	.907	.971
9	.910	.911	.909	.972
10	.911	.912	.910	.972
11	.908	.909	.905	.972
12	.908	.908	.906	.971
13	.910	.910	.908	.972
14	.916	.916	.915	.973
15	.910	.911	.908	.972
16	.915	.915	.913	.973
17	.909	.909	.906	.972
18	.905	.906	.903	.971
19	.912	.912	.911	.973
20	.916	.916	.915	.974
21	.914	.915	.912	.973
22	.911	.911	.910	.972
23	.912	.913	.911	.973
24	.907	.907	.905	.971
25	.913	.913	.911	.973
26	.910	.911	.907	.972
27	.911	.912	.908	.973
28	.910	.910	.908	.972
29	.911	.911	.908	.972

ตาราง 8 (ต่อ)

กลุ่มที่	r20	rB	rconstruct	ρ_{IRT}
30	.913	.913	.912	.973
ช่วงความเชื่อมั่น 95%	.904 - .916	.906 - .916	.902 - .914	.972 - .976

ผลการวิเคราะห์ตามตาราง 8 พบว่า ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 1,500 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .905 ถึง .916 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .906 ถึง .916 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .903 ถึง .915 และสัมประสิทธิ์ ρ_{IRT} มีค่าอยู่ระหว่าง .971 ถึง .974

สัมประสิทธิ์ ρ_{IRT} ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B ค่าที่ประมาณได้ทุกค่าอยู่ในช่วงความเชื่อมั่น 95% สัมประสิทธิ์โครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 2 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% สัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มี 6 ค่า ที่ประมาณค่าได้ต่ำกว่าช่วงความเชื่อมั่น 95%

เพื่อให้สามารถมองเห็นภาพรวมของค่าความเชื่อมั่นที่ประมาณได้จากสูตร 4 สูตร มีค่าอยู่ในช่วงของความเชื่อมั่น 95% หรือไม่ จากกลุ่มตัวอย่างขนาดกลาง 4 ขนาด แต่ละขนาดมีจำนวน 30 กลุ่ม หากค่าที่ประมาณได้มีค่าอยู่ในช่วงนี้ก็แสดงว่าสูตรนั้นสามารถประมาณค่าได้ถูกต้อง ผู้วิจัยได้สรุปข้อมูลดังตาราง 9

ตาราง 9 ช่วงความเชื่อมั่น 95% ของการประมาณค่าความเชื่อมั่นที่ประมาณค่าได้จาก 4 สูตร ที่กลุ่มตัวอย่างขนาดกลาง 4 ขนาด จำนวน 30 กลุ่ม

ขนาด	**สูตร**	**ค่าที่ประมาณได้**	**ค่าเฉลี่ย**	**SD**	**ช่วงความเชื่อมั่น95%**	**จำนวนค่าที่อยู่นอกกรอบช่วงความเชื่อมั่น 95%**
300	r_{20}	.895 - .921	.91180	.006376	.898 - .923	1
	r_B	.896 - .922	.91260	.006355	.899 - .922	1
	$r_{construct}$	.789 -.921	.90690	.023414	.862 - .954	1
	ρ_{IRT}	.959 - .969	.96533	.002339	.969 - .979	28

ตาราง 9 (ต่อ)

ขนาด	สูตร	ค่าที่ประมาณได้	ค่าเฉลี่ย	SD	ช่วงความเชื่อมั่น95%	จำนวนค่าที่อยู่นอกกรอบช่วงความเชื่อมั่น 95%
500	r_{20}	.900 - .923	.91110	.005074	.900 - .920	1
	r_B	.901 - .923	.91163	.004867	.901 - .921	1
	$r_{construct}$	.899 - .922	.91040	.005184	.898 - .918	2
	ρ_{IRT}	.964 - .973	.96830	.002020	.970 - .978	21
1,000	r_{20}	.905 - .914	.90997	.002895	.904 - .916	0
	r_B	.905 - .915	.91063	.002785	.905 - .916	0
	$r_{construct}$	.903 - .913	.90868	.003354	.901 - .915	0
	ρ_{IRT}	.969 - .973	.97093	.001143	.972 - .976	18
1,500	r_{20}	.905 – .916	.91043	.002825	.904 - .916	0
	r_B	.906 - .916	.91090	.002708	.906 - .916	0
	$r_{construct}$	.903 - .915	.90867	.003089	.902 - .914	2
	ρ_{IRT}	.971 - .974	.97220	.000847	.972 - .976	6

ผลการวิเคราะห์ตามตาราง 9 พบว่า ช่วงความเชื่อมั่น 95% ของการประมาณค่าความเชื่อมั่นที่กลุ่มตัวอย่างขนาดกลาง 300 คน ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 1 ค่า ที่ประมาณค่าได้ต่ำกว่าช่วงความเชื่อมั่น 95% และสัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยค่าที่ประมาณได้มีค่าต่ำกว่าช่วงความเชื่อมั่น 95%

ที่กลุ่มตัวอย่างขนาดกลาง 500 คน ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 1 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% สัมประสิทธิโครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 2 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% และสัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยค่าที่ประมาณได้มีค่าต่ำกว่าช่วงความเชื่อมั่น 95%

ที่กลุ่มตัวอย่างขนาดกลาง 1,000 คน ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง ค่าที่ประมาณได้มีค่าอยู่ในช่วงความเชื่อมั่น 95% ส่วนสัมประสิทธิ์ ρ_{IRT} ค่าที่ประมาณได้มีค่าที่ไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยมีบางค่าที่ได้ต่ำกว่าช่วงความเชื่อมั่น 95%

ที่กลุ่มตัวอย่างขนาดกลาง 1,500 คน ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B ค่าที่ประมาณได้มีค่าอยู่ในช่วงความเชื่อมั่น 95% สัมประสิทธิโครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 2 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% สัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มี 6 ค่า ที่ประมาณค่าได้ต่ำกว่าช่วงความเชื่อมั่น 95%

ตอนที่ 2.3 การคำนวณจากกลุ่มตัวอย่างขนาดเล็ก 4 ขนาด

ผู้วิจัยนำคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 จากกลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ ขนาด 30 60 120 และ 240 คน แต่ละขนาดมีจำนวน 30 กลุ่ม มาประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B จากโปรแกรม Formula r_B ส่วนสัมประสิทธิโครงสร้าง ผู้วิจัยทำการวิเคราะห์จากการวิเคราะห์องค์ประกอบเชิงยืนยันโมเดลการทดสอบคะแนนจริงสัมพันธ์ที่วัดองค์ประกอบเดี่ยว แล้วนำค่าน้ำหนักองค์ประกอบที่ได้มาประมาณค่าความเชื่อมั่นตามวิธีการของเรย์คอฟ(Raykov. 2003: 143-159) ผู้วิจัยได้นำเสนอในตาราง 10 ถึงตาราง 13 ดังนี้

ตาราง 10 ค่าความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง จากกลุ่มตัวอย่างขนาด 30 คน

กลุ่มที่	r_{20}	r_B	$r_{construct}$
1	.907	.909	.903
2	.905	.906	.901
3	.911	.912	.900
4	.870	.872	.866
5	.920	.921	.916
6	.923	.924	.920
7	.913	.915	.866
8	.873	.875	.866
9	.918	.919	.914
10	.899	.901	.896
11	.926	.926	.924
12	.892	.895	.891
13	.929	.930	.927

ตาราง 10 (ต่อ)

กลุ่มที่	r_{20}	r_B	$r_{construct}$
14	.899	.901	.896
15	.895	.895	.892
16	.938	.939	.936
17	.926	.931	.926
18	.929	.930	.926
19	.912	.914	.910
20	.875	.877	.871
21	.900	.901	.892
22	.895	.897	.893
23	.920	.921	.919
24	.910	.912	.907
25	.891	.893	.890
26	.920	.922	.918
27	.904	.906	.900
28	.910	.913	.912
29	.928	.929	.927
30	.906	.908	.900

ผลการวิเคราะห์ตามตาราง 10 พบว่า ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 30 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .870 ถึง .938 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .872 ถึง .939 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .866 ถึง .936

สัมประสิทธิ์ r_B ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ตาราง 11 ค่าความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง จากกลุ่มตัวอย่างขนาด 60 คน

กลุ่มที่	r_{20}	r_B	$r_{construct}$
1	.899	.901	.899
2	.915	.915	.914
3	.912	.914	.909
4	.875	.877	.875
5	.888	.889	.887
6	.899	.901	.894
7	.905	.907	.909
8	.917	.918	.918
9	.923	.924	.923
10	.898	.900	.897
11	.881	.882	.878
12	.912	.913	.911
13	.902	.904	.902
14	.913	.914	.913
15	.898	.899	.897
16	.892	.892	.889
17	.902	.903	.903
18	.912	.913	.912
19	.912	.913	.911
20	.903	.905	.900
21	.923	.924	.923
22	.907	.908	.907
23	.915	.916	.914
24	.903	.904	.903
25	.900	.901	.899
26	.909	.911	.908
27	.881	.883	.874
28	.879	.881	.879
29	.916	.918	.917
30	.904	.906	.904

ผลการวิเคราะห์ตามตาราง 11 พบว่า ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 60 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .875 ถึง .923 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .877 ถึง .924 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .874 ถึง .923

สัมประสิทธิ์ r_B ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ตาราง 12 ค่าความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง จากกลุ่มตัวอย่างขนาด 120 คน

กลุ่มที่	r_{20}	r_B	$r_{construct}$
1	.915	.916	.916
2	.891	.892	.900
3	.899	.901	.898
4	.907	.907	.906
5	.913	.913	.912
6	.884	.885	.883
7	.901	.902	.901
8	.882	.882	.880
9	.896	.897	.895
10	.907	.908	.906
11	.897	.898	.896
12	.899	.900	.899
13	.886	.887	.885
14	.917	.917	.916
15	.895	.895	.896
16	.903	.905	.903
17	.899	.899	.898
18	.892	.893	.893
19	.894	.895	.895
20	.895	.896	.893
21	.907	.907	.906
22	.892	.893	.890
23	.906	.908	.907

ตาราง 12 (ต่อ)

กลุ่มที่	r_{20}	r_B	$r_{construct}$
24	.909	.910	.907
25	.900	.901	.900
26	.912	.913	.913
27	.898	.899	.898
28	.903	.904	.902
29	.908	.909	.909
30	.922	.922	.906

ผลการวิเคราะห์ตามตาราง 12 พบว่า ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 120 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .882 ถึง .922 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .882 ถึง .922 สัมประสิทธิโครงสร้าง มีค่าอยู่ระหว่าง .880 ถึง .916

สัมประสิทธิ์ r_B ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ KR-20 และสัมประสิทธิโครงสร้าง ตามลำดับ

ตาราง 13 ค่าความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิโครงสร้าง จากกลุ่มตัวอย่างขนาด 240 คน

กลุ่มที่	r_{20}	r_B	$r_{construct}$
1	.913	.914	.914
2	.913	.913	.912
3	.907	.908	.906
4	.919	.920	.919
5	.913	.913	.912
6	.912	.912	.912
7	.912	.912	.912
8	.917	.918	.918
9	.911	.912	.911
10	.914	.915	.915
11	.907	.908	.905
12	.907	.908	.907

ตาราง 13 (ต่อ)

กลุ่มที่	r_{20}	r_B	$r_{construct}$
13	.910	.910	.910
14	.920	.921	.921
15	.913	.914	.913
16	.906	.907	.906
17	.918	.918	.918
18	.904	.905	.904
19	.914	.915	.915
20	.909	.910	.908
21	.912	.913	.912
22	.907	.908	.908
23	.913	.914	.914
24	.916	.917	.915
25	.912	.913	.913
26	.903	.903	.902
27	.914	.914	.914
28	.909	.910	.909
29	.917	.918	.917
30	.915	.916	.914

ผลการวิเคราะห์ตามตาราง 13 พบว่า ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 240 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .903 ถึง .920 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .903 ถึง .921 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .902 ถึง .921

สัมประสิทธิ์ r_B ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ KR-20 ตามลำดับ

ตอนที่ 3 ความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณจากสูตร 4 สูตร

ผู้วิจัยได้นำค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตร 4 สูตรที่กลุ่มตัวอย่างขนาดกลาง 4 ขนาด ขนาดละ 30 กลุ่ม มาหาความคลาดเคลื่อนมาตรฐาน และความลำเอียงของการประมาณค่า เพื่อดูความเหมาะสมและคุณภาพในการประมาณค่า ผู้วิจัยได้นำเสนอในตาราง 14 ถึง 15 ดังนี้

ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณจากสูตร 4 สูตร

ความคลาดเคลื่อนมาตรฐาน(Standard error of estimate: SE) ของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} พิจารณาค่า SE ที่มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสูตรประมาณค่าความเชื่อมั่น มีความเหมาะสมและมีคุณภาพในการประมาณค่า ผู้วิจัยได้นำเสนอในตาราง 14 ดังนี้

ตาราง 14 ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณจากสูตร 4 สูตร กลุ่มตัวอย่างขนาดกลาง 4 ขนาด จำนวน 30 กลุ่ม

ขนาด	สูตร	ความคลาดเคลื่อนมาตรฐาน(SE)
300	r_{20}	.006372
	r_B	.006348
	$r_{construct}$	.023452
	ρ_{IRT}	.002323
500	r_{20}	.005099
	r_B	.004867
	$r_{construct}$	.005183
	ρ_{IRT}	.002019
1,000	r_{20}	.002894
	r_B	.002785
	$r_{construct}$	.003354
	ρ_{IRT}	.001142

ตาราง 14 (ต่อ)

ขนาด	**สูตร**	**ความคลาดเคลื่อนมาตรฐาน(SE)**
1,500	r_{20}	.002824
	r_B	.002708
	$r_{construct}$	.003082
	ρ_{IRT}	.000846

ผลการวิเคราะห์ตามตาราง 14 พบว่า ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} ที่กลุ่มตัวอย่างขนาดกลาง 300, 500, 1,000, 1,500 คน มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสูตรประมาณค่าความเชื่อมั่น มีความเหมาะสมและมีคุณภาพในการประมาณค่า

ความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณจากสูตร 4 สูตร

ความลำเอียง(Bias) ของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} เป็นค่าความลำเอียงเฉลี่ยที่คิดเครื่องหมาย ซึ่งเป็นการบอกทิศทางของความลำเอียง ค่าความลำเอียงมีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าวิธีประมาณค่าความเชื่อมั่น 4 วิธี มีความเหมาะสมและมีคุณภาพในการประมาณค่า ผู้วิจัยได้นำเสนอในตาราง 15 ดังนี้

ตาราง 15 ความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณจากสูตร 4 สูตร กลุ่มตัวอย่างขนาดกลาง 4 ขนาด จำนวน 30 กลุ่ม

ขนาด	**สูตร**	**ความลำเอียง(Bias)**
300	r_{20}	.0018
	r_B	.0016
	$r_{construct}$	-.0011
	ρ_{IRT}	-.0086
500	r_{20}	.0011
	r_B	.0006
	$r_{construct}$	.0024
	ρ_{IRT}	-.0057

ตาราง 15 (ต่อ)

ขนาด	**สูตร**	**ความลำเอียง(Bias)**
1,000	r_{20}	-.00003
	r_B	-.0003
	$r_{construct}$	.0006
	ρ_{IRT}	-.0030
1,500	r_{20}	.0004
	r_B	-.0001
	$r_{construct}$	.0006
	ρ_{IRT}	-.0018

ผลการวิเคราะห์ตามตาราง 15 พบว่า ความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} ที่กลุ่มตัวอย่างขนาดกลาง 300, 500, 1,000, 1,500 คน มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสูตรประมาณค่าความเชื่อมั่น มีความเหมาะสมและมีคุณภาพในการประมาณค่า

ตอนที่ 4 การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่คำนวณจากสูตร 4 สูตร จากกลุ่มตัวอย่างขนาดกลาง

ผู้วิจัยได้นำค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตร 4 สูตร ได้แก่ สัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} ที่กลุ่มตัวอย่างขนาดกลาง 4 ขนาด ขนาดละ 30 กลุ่ม มาหาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ผู้วิจัยได้นำเสนอในตาราง 16 ดังนี้

ตาราง 16 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ของความเชื่อมั่นของแบบทดสอบที่คำนวณจากสูตร 4 สูตร กลุ่มตัวอย่างขนาดกลาง 4 ขนาด

ขนาด	r_{20}		r_B		$r_{construct}$		ρ_{IRT}	
	$\overline{r_{20}}$	SD	$\overline{r_B}$	SD	$\overline{r_{construct}}$	SD	$\overline{\rho_{IRT}}$	SD
300	.912	.0064	.913	.0051	.907	.0029	.965	.0028
500	.911	.0064	.912	.0049	.910	.0028	.968	.0027
1,000	.910	.0234	.911	.0052	.909	.0034	.971	.0031
1,500	.910	.0023	.911	.0020	.909	.0011	.972	.0008

ผลการวิเคราะห์ตามตาราง 16 พบว่า ค่าเฉลี่ยความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาดกลาง 300 , 500 , 1,000, 1,500 คน ที่ประมาณค่าด้วยสัมประสิทธิ์ ρ_{IRT} มีค่าสูงที่สุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าส่วนเบี่ยงเบนมาตรฐาน ของความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาดกลาง 300 , 500 , 1,000, 1,500 คน ที่ประมาณค่าด้วยสัมประสิทธิ์ ρ_{IRT} มีค่าต่ำที่สุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

จากนั้นผู้วิจัยได้นำค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตร 4 สูตร ได้แก่ สัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} ที่กลุ่มตัวอย่างขนาดกลาง 4 ขนาด ขนาดละ 30 กลุ่ม มาทำการเปรียบเทียบความเชื่อมั่นของแบบทดสอบด้วยการวิเคราะห์เส้นภาพ ซึ่งมีการทดสอบอยู่ 3 รายการตามสมมติฐานการวิจัยคือ ทำการทดสอบความคู่ขนานกันของเส้นภาพก่อน หากพบว่ามีความคู่ขนานกันก็จะทำการทดสอบความเรียบของเส้นภาพและระดับของเส้นภาพ ผู้วิจัยได้นำเสนอในตาราง 17 - 20 ดังนี้

ความคู่ขนานกันของเส้นภาพ (Parallelism of Profiles)

การทดสอบความคู่ขนาน(Test of Parallelism) เป็นการทดสอบว่ากลุ่มตัวอย่างแต่ละขนาดมีเส้นภาพความเชื่อมั่นคู่ขนานกันหรือไม่ หรือความแตกต่างของค่าความเชื่อมั่นที่คำนวณจากแต่ละสูตรของกลุ่มตัวอย่างแต่ละขนาดจะเท่ากันหรือไม่ ผู้วิจัยได้นำเสนอในตาราง 17 ดังนี้

ตาราง 17 การทดสอบความคู่ขนาน(Test of Parallelism)

Test Name	**Approx. F**	**df**	**Sig.**
Wilks	13.585	9	.000

ผลการวิเคราะห์ตามตาราง 17 พบว่า ค่า **Wilks F = 13.585** มีนัยสำคัญที่ระดับ .01 แสดงว่ากลุ่มตัวอย่างแต่ละขนาดมีเส้นภาพความเชื่อมั่นที่ไม่ขนานกัน

ความเรียบของเส้นภาพ (Flatness of Profiles (Coincident))

การทดสอบความเรียบ(Test of Flatness) เป็นการทดสอบว่าค่าความเชื่อมั่นที่คำนวณได้แต่ละสูตรโดยเฉลี่ยมีค่าเท่ากัน(ระนาบกัน)หรือไม่ ผู้วิจัยได้นำเสนอในตารางที่ 18 ดังนี้

ตารางที่ 18 การทดสอบความเรียบ(Test of Flatness)

Test Name	Approx. F	df	Sig.
Wilks	15226.414	3	.000

ผลการวิเคราะห์ตามตาราง 18 พบว่า **Wilks F = 15226.414** มีนัยสำคัญที่ระดับ .01 แสดงว่าค่าความเชื่อมั่นที่คำนวณได้แต่ละสูตรมีเส้นภาพที่ไม่ระนาบกัน

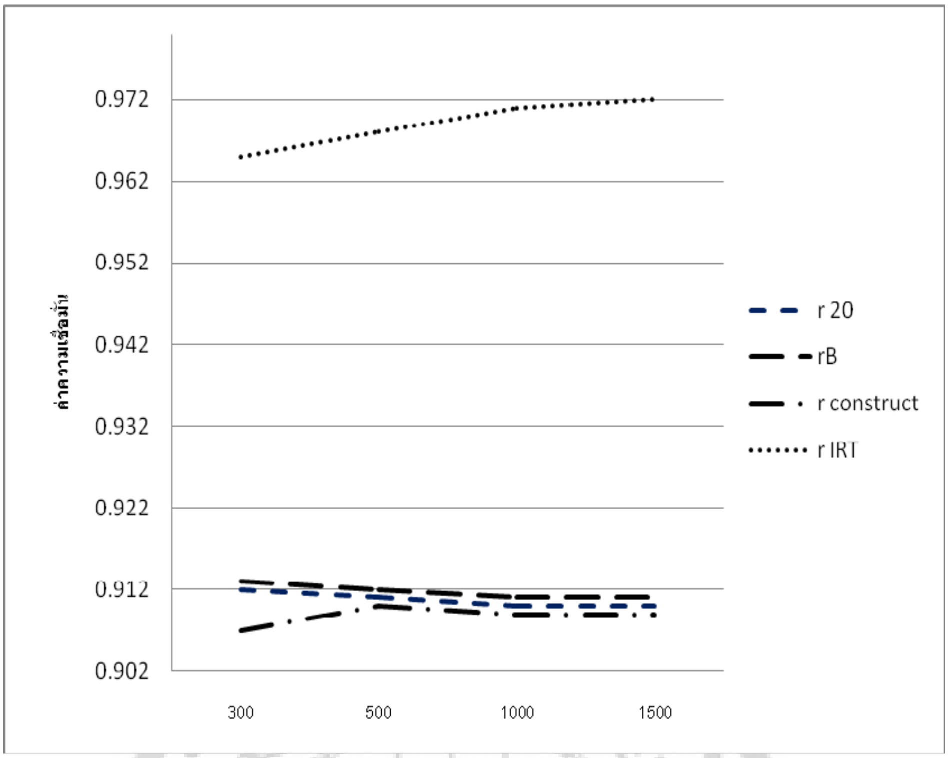
การทดสอบระดับ(Level Test)

การทดสอบระดับ(Level Test) เป็นการทดสอบกลุ่มตัวอย่างแต่ละขนาดจะให้เส้นภาพ (Profile) ของความเชื่อมั่นอยู่ในระดับเดียวกันหรือไม่ หรือทดสอบว่าค่าความเชื่อมั่นเฉลี่ยที่คำนวณจากกลุ่มตัวอย่างขนาดหนึ่งจะแตกต่างไปจากกลุ่มตัวอย่างขนาดอื่นๆหรือไม่ ผู้วิจัยได้นำเสนอในตารางที่ 19 ดังนี้

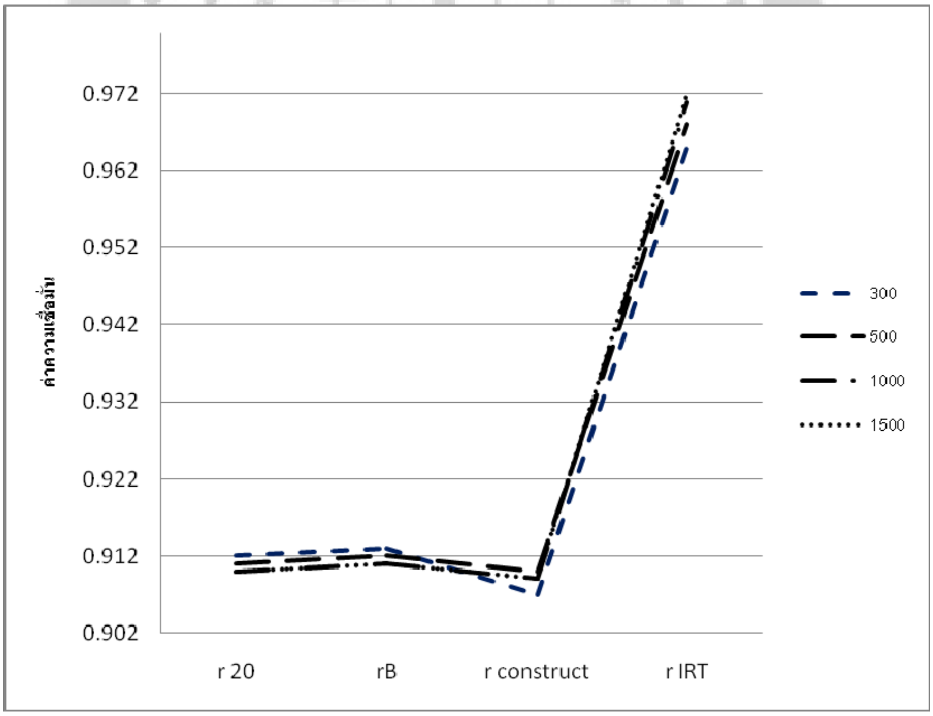
ตารางที่ 19 การทดสอบระดับ(Level Test)

Source of Variance	SS	df	MS	F	Sig
Within+Residual	.01	116	.00		
Between (Plan)	.00	3	.00	.47	.702

ผลการวิเคราะห์ตามตาราง 19 พบว่า กลุ่มตัวอย่างแต่ละขนาดจะให้เส้นภาพ(Profile) ของความเชื่อมั่นรวมในทุกขนาดแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าอยู่ระดับเดียวกัน



ภาพประกอบ 1 ค่าความเชื่อมั่นของแบบทดสอบของการทดสอบความคู่ขนาน



ภาพประกอบ 2 ค่าความเชื่อมั่นของแบบทดสอบที่สูตรการประมาณค่าความเชื่อมั่นเดียวกัน ขนาดกลุ่มตัวอย่างต่างกัน

เนื่องจากเส้นภาพความเชื่อมั่นที่ไม่ขนานกัน แสดงว่ามี Interaction ผู้วิจัยจึงทำการทดสอบ Simple effect ได้ผลดังตาราง 20 - 21 ดังนี้

ตาราง 20 การทดสอบ Simple effect ของการประมาณค่าความเชื่อมั่นของแบบทดสอบ

แหล่ง	SS	df	MS	F	Sig.
Sample size	3.416E-05	3	1.139E-05	.514	.673
Formula	.313	1.121	.279	3012.427	.000

ผลการวิเคราะห์ตามตาราง 20 พบว่า ที่สูตรการประมาณค่าความเชื่อมั่นเดียวกัน ขนาดกลุ่มตัวอย่างต่างกัน มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ที่ขนาดกลุ่มตัวอย่างเดียวกัน สูตรการประมาณค่าความเชื่อมั่นต่างกัน มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ผู้วิจัยจึงทำการเปรียบเทียบรายคู่ ได้ผลดังตาราง 21 ดังนี้

ตาราง 21 การเปรียบเทียบรายคู่สูตรการประมาณค่าความเชื่อมั่น

สูตร (ค่าเฉลี่ย)		Mean Difference	Std. Error	Sig.	Lower Bound	Upper Bound
r_{20} (0.910825)	r_B (0.9114417)	-.0006167**	.000	.000	-.0007336	-.0004998
	$r_{construct}$ (0.9086625)	.002162	.001	.236	-.0006228	.004948
	ρ_{IRT} (0.9691917)	-.05837**	.000	.000	-.05911	-.05763
r_B (0.9114417)	$r_{construct}$ (0.9086625)	.002779	.001	.052	-.0000128	.005571
	ρ_{IRT} (0.9691917)	-.05775**	.000	.000	-.05847	-.05703
$r_{construct}$ (0.9086625)	ρ_{IRT} (0.9691917)	-.06053**	.001	.000	-.06341	-.05765

** นัยสำคัญทางสถิติที่ระดับ .01

ผลการวิเคราะห์ตามตาราง 21 พบว่า สัมประสิทธิ์ KR-20 กับ สัมประสิทธิ์ r_B , สัมประสิทธิ์ ρ_{IRT} กับ สัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

บทที่ 5

สรุป อภิปรายผลและข้อเสนอแนะ

การวิจัยการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ และโมเดลโลจีสติก 3 พารามิเตอร์ มีความมุ่งหมายหลักเพื่อศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ ซึ่งประกอบด้วยสัมประสิทธิ์ r_B กับสัมประสิทธิ์โครงสร้าง และโมเดลโลจีสติก 3 พารามิเตอร์ ได้แก่ สัมประสิทธิ์ ρ_{IRT} และใช้สัมประสิทธิ์ KR-20 เป็นเกณฑ์เทียบ โดยมีความมุ่งหมายเฉพาะดังนี้

1. เพื่อประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากประชากรเทียม
2. เพื่อประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ ขนาด 300 500 1,000 และ 1,500 คน
3. เพื่อประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง จากกลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ ขนาด 30 60 120 และ 240 คน
4. เพื่อศึกษาความคลาดเคลื่อนมาตรฐาน และ ความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT}
5. เพื่อเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย ปีการศึกษา 2550 จำนวน 15,395 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างขนาดใหญ่ที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย จำนวน 4 เขตพื้นที่ ปีการศึกษา 2550 โดยใช้วิธีการสุ่มแบบแบ่งชั้น(Stratified Random Sampling) จำนวน 6,000 คน

ซึ่งกลุ่มตัวอย่างขนาดใหญ่นี้ ผู้วิจัยกำหนดให้เป็นประชากรเทียม

กลุ่มตัวอย่างขนาดกลางและขนาดเล็กที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 สำนักงานเขตพื้นที่การศึกษาจังหวัดเชียงราย จำนวน 4 เขตพื้นที่ ปีการศึกษา 2550 โดยใช้วิธีการสุ่มแบบใส่คืนจากประชากรเทียม แบ่งเป็นกลุ่มตัวอย่างขนาดกลาง 4 ขนาด

ได้แก่ 300 500 1,000 และ 1,500 คน กลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ 30 60 120 และ 240 คน แต่ละขนาดจะมีจำนวน 30 กลุ่ม

ตัวแปรที่ศึกษา

ตัวแปรอิสระ ได้แก่

1. สูตรการประมาณค่าความเชื่อมั่นของแบบทดสอบ 4 สูตร คือ

1.1 สัมประสิทธิ์ KR-20

1.2 สัมประสิทธิ์ r_B

1.3 สัมประสิทธิ์โครงสร้าง ที่ได้จากการประมาณค่าความเชื่อมั่นจากโมเดลการวัด

1.4 สัมประสิทธิ์ ρ_{IRT} ที่ได้จากโมเดลโลจีสติก 3 พารามิเตอร์ของทฤษฎีการตอบข้อสอบ

2. ขนาดกลุ่มตัวอย่าง ได้แก่

2.1 กลุ่มตัวอย่างขนาดกลาง ได้แก่ กลุ่มตัวอย่างขนาด 300 500 1,000 และ 1,500 คน โดยแต่ละขนาดจะมีจำนวน 30 กลุ่ม

2.2 กลุ่มตัวอย่างขนาดเล็ก ได้แก่ กลุ่มตัวอย่างขนาด 30 60 120 และ 240 คน โดยแต่ละขนาดจะมีจำนวน 30 กลุ่ม

ตัวแปรตาม คือ ค่าความเชื่อมั่นของแบบทดสอบ

สมมติฐานในการวิจัย

1. สูตรการประมาณค่าความเชื่อมั่น 4 สูตร สามารถประมาณค่าความเชื่อมั่นอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริง

2. สมมติฐานของการวิเคราะห์เส้นภาพ(Profile Analysis)

2.1 กลุ่มตัวอย่างแต่ละขนาดมีเส้นภาพความเชื่อมั่น ขนานกัน

2.2 ค่าความเชื่อมั่นที่คำนวณได้แต่ละสูตรเป็นระนาบกัน

2.3 กลุ่มตัวอย่างแต่ละขนาดมีเส้นภาพความเชื่อมั่นอยู่ในระดับเดียวกัน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 ที่ผู้วิจัยสร้างขึ้น มีลักษณะเป็นแบบทดสอบเลือกตอบ 4 ตัวเลือก จำนวน 50 ข้อ

การจัดกระทำข้อมูลและการวิเคราะห์ข้อมูล

การจัดกระทำและการวิเคราะห์ข้อมูลได้ดำเนินการดังนี้

1. วิเคราะห์ค่าสถิติพื้นฐานของคะแนนสอบ ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน คะแนนสูงสุด คะแนนต่ำสุด มัธยฐาน ฐานนิยม ความเบ้ ความโด่ง และความคลาดเคลื่อนมาตรฐาน โดยใช้โปรแกรม SPSS

2. ประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากประชากรเทียม

3. ประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง

4. ประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง จากกลุ่มตัวอย่างขนาดเล็ก

5. ศึกษาความคลาดเคลื่อนมาตรฐาน และ ความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และ สัมประสิทธิ์ ρ_{IRT}

6. เปรียบเทียบความเชื่อมั่น จากการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 ด้วยสัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} ที่กลุ่มตัวอย่างขนาดกลาง โดยใช้วิธีการวิเคราะห์เส้นภาพ(Profile Analysis)

สรุปผลการวิจัย

1. การประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากประชากรเทียม พบว่า ค่าความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ ρ_{IRT} ได้ค่าความเชื่อมั่นสูงสุด รองลงมาคือ สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

2.การประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง 4 ขนาด ได้แก่ ขนาด 300 500 1,000 และ 1,500 คน แต่ละขนาดมีจำนวน 30 กลุ่ม พบว่า

2.1. ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 300 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .895 ถึง .921 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .896 ถึง .922 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .789 ถึง .921 และสัมประสิทธิ์ ρ_{IRT} มีค่าอยู่ระหว่าง .959 ถึง .969 โดย สัมประสิทธิ์ ρ_{IRT} ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B

สัมประสิทธิ์โครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 1 ค่า ที่ประมาณค่าได้ต่ำกว่าช่วงความเชื่อมั่น 95% และสัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยค่าที่ประมาณได้มีค่าต่ำกว่าช่วงความเชื่อมั่น 95%

2.2.ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 500 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .900 ถึง .923 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .901 ถึง .923 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .899 ถึง .922 และสัมประสิทธิ์ ρ_{IRT} มีค่าอยู่ระหว่าง .964 ถึง .973 โดยสัมประสิทธิ์ ρ_{IRT} ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 1 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% สัมประสิทธิ์โครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 2 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% และสัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยค่าที่ประมาณได้มีค่าต่ำกว่าช่วงความเชื่อมั่น 95%

2.3.ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 1,000 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .905 ถึง .914 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .905 ถึง .915 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .903 ถึง .913 และสัมประสิทธิ์ ρ_{IRT} มีค่าอยู่ระหว่าง .969 ถึง .973 โดย สัมประสิทธิ์ ρ_{IRT} ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง ค่าที่ประมาณได้ทุกค่าอยู่ในช่วงความเชื่อมั่น 95% ส่วนสัมประสิทธิ์ ρ_{IRT} ค่าที่ประมาณได้มีค่าที่ไม่ได้อยู่ในช่วงความเชื่อมั่น 95% โดยมีบางค่าที่ได้ต่ำกว่าช่วงความเชื่อมั่น 95%

2.4.ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 1,500 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .905 ถึง .916 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .906 ถึง .916 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .903 ถึง .915 และสัมประสิทธิ์ ρ_{IRT} มีค่าอยู่ระหว่าง .971 ถึง .974 โดย สัมประสิทธิ์ ρ_{IRT} ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าความเชื่อมั่นที่ประมาณค่าได้จากสูตรสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B ค่าที่ประมาณได้ทุกค่าอยู่ในช่วงความเชื่อมั่น 95% สัมประสิทธิ์โครงสร้าง ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มีเพียง 2 ค่า ที่ประมาณค่าได้สูงกว่าช่วงความเชื่อมั่น 95% สัมประสิทธิ์ ρ_{IRT} ส่วนใหญ่ประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% มี 6 ค่า ที่ประมาณค่าได้ต่ำกว่าช่วงความเชื่อมั่น 95%

3.การประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง จากกลุ่มตัวอย่างขนาดเล็ก 4 ขนาด ได้แก่ ขนาด 30 60 120 และ 240 คน แต่ละขนาดมีจำนวน 30 กลุ่ม พบว่า

3.1.ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 30 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .870 ถึง .938 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .872 ถึง .939 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .866 ถึง .936 โดย สัมประสิทธิ์ r_B ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

3.2.ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 60 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .875 ถึง .923 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .877 ถึง .924 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .874 ถึง .923 โดยสัมประสิทธิ์ r_B ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

3.3.ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 120 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .882 ถึง .922 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .882 ถึง .922 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .880 ถึง .916 โดยสัมประสิทธิ์ r_B และสัมประสิทธิ์ KR-20 ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์โครงสร้าง ตามลำดับ

3.4.ค่าความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาด 240 คนที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 มีค่าอยู่ระหว่าง .903 ถึง .920 สัมประสิทธิ์ r_B มีค่าอยู่ระหว่าง .903 ถึง .921 สัมประสิทธิ์โครงสร้าง มีค่าอยู่ระหว่าง .902 ถึง .921 โดยสัมประสิทธิ์ r_B และสัมประสิทธิ์โครงสร้าง ประมาณค่าความเชื่อมั่นของแบบทดสอบได้ค่าสูงสุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ KR-20 ตามลำดับ

4. ความคลาดเคลื่อนมาตรฐาน และ ความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT}

4.1.ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} พบว่า ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นของแบบทดสอบที่คำนวณจาก 4 สูตร ที่กลุ่มตัวอย่างขนาดกลาง 300, 500, 1,000, 1,500 คน มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสูตรประมาณค่าความเชื่อมั่น มีความเหมาะสมและมีคุณภาพในการประมาณค่า

4.2.ความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ ด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} พบว่า ค่าความลำเอียงของการ

ประมาณค่าความเชื่อมั่นของแบบทดสอบที่คำนวณจาก 4 สูตร ที่กลุ่มตัวอย่างขนาดกลาง 300, 500, 1,000, 1,500 คน มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสูตรประมาณค่าความเชื่อมั่น มีความเหมาะสมและมีคุณภาพในการประมาณค่า

5.การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง

5.1.ค่าเฉลี่ยความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาดกลาง 300 , 500 , 1,000, 1,500 คน ที่ประมาณค่าด้วยสัมประสิทธิ์ ρ_{IRT} มีค่าสูงที่สุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

ค่าส่วนเบี่ยงเบนมาตรฐาน ของความเชื่อมั่นของแบบทดสอบจากกลุ่มตัวอย่างขนาดกลาง 300 , 500 , 1,000, 1,500 คน ที่ประมาณค่าด้วยสัมประสิทธิ์ ρ_{IRT} มีค่าต่ำที่สุดในทุกกลุ่ม รองลงมาคือ สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง ตามลำดับ

5.2.ความคู่ขนานกันของเส้นภาพ (Parallelism of Profiles) พบว่า ค่า **Wilks F = 13.585** มีนัยสำคัญที่ระดับ .01 แสดงว่ากลุ่มตัวอย่างแต่ละขนาดมีเส้นภาพความเชื่อมั่นที่ไม่ขนานกัน

5.3.ความเรียบของเส้นภาพ (Test of Flatness) พบว่า **Wilks F = 15226.414** มีนัยสำคัญที่ระดับ .01 แสดงว่าค่าความเชื่อมั่นที่คำนวณได้แต่ละสูตรมีเส้นภาพที่ไม่ระนาบกัน

5.4.ระดับของเส้นภาพ(Level Test) พบว่า กลุ่มตัวอย่างแต่ละขนาดจะให้เส้นภาพ (Profile) ของความเชื่อมั่นรวมในทุกขนาดแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าอยู่ระดับเดียวกัน

เนื่องจากเส้นภาพความเชื่อมั่นที่ไม่ขนานกัน แสดงว่ามี Interaction ผู้วิจัยจึงทำการทดสอบ Simple effect ได้ผลดังนี้

5.5.เมื่อสูตรการประมาณค่าความเชื่อมั่นเป็นสูตรเดียวกัน ขนาดกลุ่มตัวอย่างต่างกันมีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

5.6.เมื่อขนาดกลุ่มตัวอย่างเป็นขนาดเดียวกัน สูตรการประมาณค่าความเชื่อมั่นต่างกันมีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ผู้วิจัยจึงทำการเปรียบเทียบรายคู่ พบว่า สัมประสิทธิ์ KR-20 กับ สัมประสิทธิ์ r_B , สัมประสิทธิ์ ρ_{IRT} กับ สัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง มีค่าความเชื่อมั่นของแบบทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

อภิปรายผล

การวิจัยครั้งนี้มีความมุ่งหมายหลักเพื่อศึกษาการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ ซึ่งประกอบด้วยสัมประสิทธิ์ r_B กับสัมประสิทธิ์โครงสร้าง และโมเดลโลจีสติก 3 พารามิเตอร์ ได้แก่ สัมประสิทธิ์ ρ_{IRT} และใช้สัมประสิทธิ์ KR-20 เป็นเกณฑ์เทียบว่าทั้ง 4 สูตรสามารถประมาณค่าความเชื่อมั่นได้เมื่อกลุ่มตัวอย่างมีขนาดใด แต่ละวิธีให้ค่าความคลาดเคลื่อนมาตรฐานและความลำเอียงมากน้อยเพียงใดและเมื่อกลุ่มตัวอย่างมีขนาดต่างกัน จะประมาณค่าความเชื่อมั่นของแบบทดสอบจาก 4 สูตร แตกต่างกันหรือไม่ สามารถอภิปรายผลตามสมมติฐานการวิจัยได้ดังนี้

1.การประมาณค่าโดยสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 สามารถประมาณค่าได้กับกลุ่มตัวอย่างขนาดเล็กและขนาดกลาง คือ 30 60 120 240 คน และ 300 500 1,000 1,500 คน โดยสัมประสิทธิ์ r_B สามารถประมาณค่าความเชื่อมั่นได้สูงกว่าสัมประสิทธิ์ KR-20 ทั้ง 2 สูตรสามารถประมาณค่าความเชื่อมั่นอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริง ความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 มีความเหมาะสมในการประมาณค่าความเชื่อมั่น และมีคุณภาพในการประมาณค่า ผลการวิจัยส่วนนี้อาจเนื่องมาจาก ทฤษฎีการทดสอบมาตรฐานเดิม (CTT)ได้แบ่งระดับความคู่ขนานไว้ 3 ระดับคือ 1.ความคู่ขนานแบบมาตรฐานเดิม (Classical Parallel) 2.ความคู่ขนานแบบคะแนนจริงสมมูล (Tau-Equivalent) 3.ความคู่ขนานแบบคะแนนจริงสัมพันธ์ (Congeneric) แต่ละระดับก็มีข้อตกลงเบื้องต้นและมีการผ่อนปรนข้อตกลงเบื้องต้นบางข้อในระดับที่ 2 และ 3 **สัมประสิทธิ์ r_B** เป็นสูตรที่อยู่ในระดับความคู่ขนานแบบคะแนนจริงสัมพันธ์ (Congeneric) ที่มีการผ่อนปรนข้อตกลงของความคู่ขนานลงมาตามลำดับเหลือเงื่อนไขเพียง ข้อ 1.แบบทดสอบต้องมีเนื้อหาเป็นเอกพันธ์กันหรือวัดคุณลักษณะเดียวกัน เพียงข้อเดียว และในการผ่อนปรนเงื่อนไขข้อ 2 ให้แต่ละส่วนมีคะแนนจริงต่างกันเท่ากับค่าคงที่ที่มาสัมพันธ์สมบูรณ์แบบหรือมีความสัมพันธ์เป็นเส้นตรง ได้โมเดลคะแนนจริงสัมพันธ์คือ $T_i= a_{ij}T_j + b_{ij}$ ส่วน**สัมประสิทธิ์ KR-20** เป็นสูตรที่อยู่ในระดับความคู่ขนานแบบคะแนนจริงสมมูล (Tau-Equivalent)ที่มีการผ่อนปรนข้อตกลงของความคู่ขนานลง แต่ก็มีมากกว่าความคู่ขนานแบบคะแนนจริงสัมพันธ์ คือต้องมีความแปรปรวนร่วมของคะแนนสอบ(X) เท่ากัน ($\sigma_{12} = \sigma_{13} = \sigma_{23} = \ldots$) และความแปรปรวนร่วมของคะแนนสอบ(X)กับคะแนนเกณฑ์ภายนอกเท่ากัน ($\sigma_{1Y} = \sigma_{2Y} = \sigma_{3Y} = \ldots$) จึงทำให้ค่าความเชื่อมั่นที่ประมาณค่าได้จากสัมประสิทธิ์ r_B มีค่าสูงกว่าสัมประสิทธิ์ KR-20 สอดคล้องกับ บุญเชิด ภิญโญอนันตพงษ์ (2542: บทคัดย่อ) ได้ทำการศึกษาสัมประสิทธิ์ r_B : การประมาณค่าความเชื่อมั่นสำหรับแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยความยากรายข้อต่างกัน เพื่อพัฒนาและศึกษาผลการใช้สูตร r_B ที่ผู้วิจัยพัฒนาขึ้นมาสำหรับประมาณค่าความเชื่อมั่นของแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยข้อสอบที่ให้คะแนนระบบ 0, 1 และมีความยากรายข้อต่างกัน โดยศึกษาจากแบบทดสอบที่วัดเนื้อหาเดียวกัน จำนวนข้อต่างกัน จำนวน 6 ฉบับ คือ

ฉบับ 25, 30, 35, 40, 45, และ 50 ข้อ ทั้งจากกลุ่มตัวอย่างขนาดใหญ่ จำนวน 6 กลุ่มคือ 412, 400, 410, 403, 405, 406 คน รวม 2,436 คน และกลุ่มตัวอย่างขนาดเล็ก คือ 30, 50 และ 80 คนว่าเป็นอย่างไรเมื่อเทียบกับสูตร KR-20 ผลการศึกษาพบว่า สูตร r_B คำนวณค่าความเชื่อมั่นของแบบทดสอบจำนวนข้อต่างกันจากกลุ่มตัวอย่างขนาดใหญ่ได้สูงสอดคล้องสัมพันธ์กับสูตร KR-20 โดยสูตร r_B คำนวณค่าความเชื่อมั่นได้สูงกว่าสูตร KR-20 ทุกฉบับ เมื่อศึกษาจากกลุ่มตัวอย่างขนาดเล็ก พบว่า สูตร r_B คำนวณจากแบบทดสอบฉบับเดียวกันและกลุ่มตัวอย่างขนาดเดียวกันได้สอดคล้องสัมพันธ์กับสูตร KR-20 จากแบบทดสอบทั้ง 6 ฉบับและกลุ่มตัวอย่างทุกขนาด สูตร r_B คำนวณค่าความเชื่อมั่นสูงกว่าสูตร KR-20 ส่วนจำนวนข้อสอบและขนาดกลุ่มตัวอย่างมีผลต่อค่าความเชื่อมั่นที่คำนวณได้จากทั้งสองสูตรในลักษณะคล้ายคลึงกัน และทั้งสองสูตรประมาณค่าความเชื่อมั่นจากกลุ่มตัวอย่างขนาดเล็กเทียบกับกลุ่มตัวอย่างขนาดใหญ่มีค่าความลำเอียงในทิศทางเดียวกัน แสดงให้เห็นว่าสัมประสิทธิ์ KR-20 และ r_B สามารถนำไปใช้ในการคำนวณค่าความเชื่อมั่นจากกลุ่มตัวอย่างขนาดเล็ก คือ 30 50 และ 80 คน และกลุ่มตัวอย่างขนาดใหญ่ ได้

2.การประมาณค่าโดยสัมประสิทธิโครงสร้าง สามารถประมาณค่าได้กับกลุ่มตัวอย่างขนาดเล็กและขนาดกลาง คือ 30 60 120 240 คน และ 300 500 1,000 1,500 คน แต่เมื่อพิจารณาความสามารถประมาณค่าความเชื่อมั่นอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริง พบว่าสามารถประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริง เมื่อกลุ่มตัวอย่างมีขนาด 300 500 1,000 1,500 คน ความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสัมประสิทธิโครงสร้าง มีความเหมาะสมในการประมาณค่าความเชื่อมั่น และมีคุณภาพในการประมาณค่า ผลการวิจัยส่วนนี้อาจเนื่องมาจาก การประมาณค่าโดยสัมประสิทธิโครงสร้าง ได้ทำการวิเคราะห์โดยใช้โมเดลลิสเรล และใช้โปรแกรม ลิสเรลในการวิเคราะห์ข้อมูล ซึ่งมีขนาดของกลุ่มตัวอย่างเข้ามาเกี่ยวข้องในการใช้โปรแกรม โดยชูมัคเกอร์และโลแมค(Schumacker and Lomax. 1996 อ้างอิงจาก นงลักษณ์ วิรัชชัย.2542: 311) ได้สรุปงานวิจัยที่มีการศึกษาเรื่องขนาดของกลุ่มตัวอย่างที่เหมาะสมสำหรับใช้โมเดลลิสเรลว่า จากงานวิจัยของดิง เวลิเซอร์และฮาร์โล(Ding, Velicer and Harlow) ในปี 1995 พบว่างานวิจัยที่ใช้โมเดลลิสเรลส่วนใหญ่ใช้ขนาดกลุ่มตัวอย่าง 100-150 คนและให้ผลการวิจัยที่น่าพอใจ บูมสมา(Boomsma) เสนอไว้ในบทความเมื่อปี 1983 ว่าขนาดกลุ่มตัวอย่างที่เหมาะสมควรเป็น 400 คน นอกจากนี้ ชูมัคเกอร์และโลแมค(Schumacker and Lomax. 1996) แฮร์และคณะ(Hait and other. 1998) เสนอให้ใช้กฎแห่งความชัดเจน(rule of thump) ที่นักสถิติตัวแปรพหุใช้กันมากคือใช้ขนาดกลุ่มตัวอย่าง 10-20 คนต่อตัวแปรในการวิจัยหนึ่งตัวแปร ซึ่งจากงานวิจัยครั้งนี้ ได้ทำการศึกษาจากข้อสอบจำนวน 50 ข้อ เมื่อทำการวิเคราะห์โดยโมเดลลิสเรล จึงทำให้ค่าความเชื่อมั่นที่ประมาณค่าได้จะมีความเหมาะสมเมื่อกลุ่มตัวอย่างมีขนาด 300 500 1,000 1,500 คน

3.การประมาณค่าโดยสัมประสิทธิ์ ρ_{IRT} สามารถประมาณค่าได้กับกลุ่มตัวอย่างขนาดกลาง คือ 300 500 1,000 1,500 คน โดยสัมประสิทธิ์ ρ_{IRT} สามารถประมาณค่าความเชื่อมั่นได้สูงกว่าสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง สามารถประมาณค่าความเชื่อมั่นอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริงที่กลุ่มตัวอย่างขนาด 1,000 1,500 คน ความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าสัมประสิทธิ์ ρ_{IRT} มีความเหมาะสมในการประมาณค่าความเชื่อมั่น และมีคุณภาพในการประมาณค่า ผลการวิจัยส่วนนี้อาจเนื่องมาจากการประมาณค่าสัมประสิทธิ์ ρ_{IRT} ได้ทำการวิเคราะห์จากโมเดลโลจิสติก 3 พารามิเตอร์ ซึ่งเป็นโมเดลหนึ่งของทฤษฎีการตอบข้อสอบ (IRT) เมื่อมีการให้คะแนนแบบ 2 ค่า เมื่อได้ค่าพารามิเตอร์จากการประมาณค่าแล้วก็จะนำมาแปลงเป็นค่าความเชื่อมั่น โดยสูตรของ ดิมิทรอฟ(Dimitrov. 2003 : 440) ขนาดกลุ่มตัวอย่างจึงส่งผลต่อค่าพารามิเตอร์ทำให้ส่งผลต่อการประมาณค่าความเชื่อมั่นด้วย สอดคล้องกับ ฮูลิน ดราสโกวและพาร์สัน(Hulin, Drasgow and Parson. 1983: 99-105) ที่ได้กล่าวถึงปัจจัยที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ ได้แก่ ความยาวแบบทดสอบ ขนาดกลุ่มตัวอย่าง และจำนวนพารามิเตอร์ของโมเดลโลจีสติก จากผลการวิจัยพบว่า ความยาวของแบบทดสอบและขนาดกลุ่มตัวอย่างเป็นปัจจัยสำคัญที่ส่งผลต่อการประมาณค่าพารามิเตอร์โดยได้เสนอแนะว่าโมเดลโลจีสติก 3 พารามิเตอร์เพื่อประมาณค่าพารามิเตอร์ด้วยวิธีความน่าจะเป็นสูงสุด(ML)ให้ได้ผลถูกต้องนั้นควรใช้แบบทดสอบที่มีจำนวน 60 ข้อกับกลุ่มตัวอย่างจำนวนไม่น้อยกว่า 200 คน หรือถ้าใช้แบบทดสอบที่มีจำนวนข้อน้อยกว่า 30 ข้อ จะต้องใช้กับกลุ่มตัวอย่างไม่น้อยกว่า 1,000 คน ในขณะที่ลอร์ด (Lord.1968 อ้างอิงจาก Hulin, Drasgow and Parson. 1983: 100-139)ได้กล่าวว่าการประมาณค่าพารามิเตอร์ด้วยวิธีความน่าจะเป็นสูงสุด(ML)ในโมเดลโลจีสติก 3 พารามิเตอร์ ให้มีความถูกต้องนั้น ควรใช้ข้อสอบประมาณ 50 ข้อกับกลุ่มตัวอย่างอย่างน้อย 1,000 คน ดังนั้นผลการวิจัยครั้งนี้ที่ใช้ข้อสอบจำนวน 50 ข้อ จึงได้กลุ่มตัวอย่างที่เหมาะสมในการประมาณค่าความเชื่อมั่นคือ 1,000 1,500 คน

4.การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากกลุ่มตัวอย่างขนาดกลาง มีเส้นภาพความเชื่อมั่นที่ไม่ขนานกัน ไม่ความสอดคล้องกันและอยู่ในระดับเดียวกัน ผลการวิจัยส่วนนี้อาจเนื่องมาจาก การประมาณค่าความเชื่อมั่นของทั้ง 4 วิธีมีข้อตกลงเบื้องต้นที่ต่างกัน ขนาดกลุ่มตัวอย่างมีผลต่อการวิเคราะห์ข้อมูล เมื่อนำมาเปรียบเทียบกันจึงมีเส้นภาพที่ไม่ขนานกัน ไม่สอดคล้องกัน แต่อยู่ในระดับเดียวกัน ซึ่งผลการวิจัยครั้งนี้ทำให้ทราบว่า **สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20** สามารถประมาณค่าได้กับกลุ่มตัวอย่างขนาดเล็กและขนาดกลาง คือ 30 60 120 240 คน และ 300 500 1,000 1,500 คน โดยสัมประสิทธิ์ r_B สามารถประมาณค่าความเชื่อมั่นได้สูงกว่าสัมประสิทธิ์ KR-20 ส่วน **สัมประสิทธิ์โครงสร้าง** สามารถประมาณค่าได้กับกลุ่มตัวอย่างขนาดเล็กและขนาดกลาง คือ 30 60 120 240 คน และ 300 500 1,000 1,500 คน แต่

เมื่อพิจารณาความสามารถประมาณค่าความเชื่อมั่นอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริง พบว่าสามารถประมาณค่าอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริง เมื่อกลุ่มตัวอย่างมีขนาด 500 1,000 1,500 คน **สัมประสิทธิ์** ρ_{IRT} สามารถประมาณค่าได้กับกลุ่มตัวอย่างขนาดกลาง คือ 300 500 1,000 1,500 คน โดยสัมประสิทธิ์ ρ_{IRT} สามารถประมาณค่าความเชื่อมั่นได้สูงกว่าสัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20 และสัมประสิทธิ์โครงสร้าง สามารถประมาณค่าความเชื่อมั่นอยู่ในช่วงความเชื่อมั่น 95% ของความเชื่อมั่นที่แท้จริงที่กลุ่มตัวอย่างขนาด 1,000 1,500 คน โดยทั้ง 4 สูตรมีความคลาดเคลื่อนมาตรฐานและความลำเอียงของการประมาณค่าความเชื่อมั่นของแบบทดสอบ มีค่าต่ำ เข้าใกล้ศูนย์ แสดงว่าทั้ง 4 สูตรประมาณค่าความเชื่อมั่น มีความเหมาะสมและมีคุณภาพในการประมาณค่า เมื่อนำมาพิจารณาร่วมกับบริบทของการประเมินในระดับชั้นเรียน(Classroom Assessment) เพื่อสนับสนุน ส่งเสริมการเรียนรู้ของผู้เรียน การใช้ **สัมประสิทธิ์ r_B สัมประสิทธิ์ KR-20** จึงมีความเหมาะสมที่จะนำไปใช้ในระดับนี้ ซึ่งวิธีการทั้ง 2 วิธีนี้ มีขั้นตอนและวิธีคำนวณที่ง่ายต่อการนำไปใช้ ไม่ต้องอาศัยทักษะทางคอมพิวเตอร์มากนัก อีกทั้งครูผู้สอนโดยส่วนใหญ่มีความคุ้นเคยกับการใช้ KR-20 ในการประมาณค่าความเชื่อมั่น หากทำการวิเคราะห์ค่า KR-20 โดยใช้โปรแกรม SPSS ให้พึงระลึกไว้ว่า ค่าที่ได้จะเป็นค่าความเชื่อมั่นที่เป็นขอบเขตขั้นต่ำของการประมาณค่า ส่วนการวัดประเมินในระดับใหญ่(Large Scale)ที่เป็นการวัดประเมินมาตรฐานความรู้ในระดับกลุ่มโรงเรียน เขตพื้นที่การศึกษาและระดับประเทศ เป็นการวัดประเมินเพื่อสนับสนุนคุณภาพโดยรวมของโรงเรียน การใช้**สัมประสิทธิ์โครงสร้าง** และ **สัมประสิทธิ์** ρ_{IRT} จึงมีความเหมาะสมที่จะนำไปใช้ในระดับนี้ ซึ่งวิธีการทั้ง 2 วิธีสามารถประมาณค่าได้จากนักเรียนจำนวนมาก ได้ค่าพารามิเตอร์ที่บ่งบอกถึงประชากรและความไม่แปรเปลี่ยนตามกลุ่มตัวอย่างซึ่งเป็นข้อตกลงของทฤษฎีการตอบข้อสอบ(IRT) แต่ต้องใช้ความรู้และทักษะทางคอมพิวเตอร์ในการศึกษาทฤษฎีและการวิเคราะห์ข้อมูลจากโมเดลลิสเรลและทฤษฎีการตอบข้อสอบ (IRT)

ข้อเสนอแนะ

จากผลการวิจัยครั้งนี้ สามารถสรุปเป็นข้อเสนอแนะในการนำไปใช้และในการวิจัยครั้งต่อไปดังนี้

ข้อเสนอแนะในการนำไปใช้

เมื่อใช้แบบทดสอบที่มีลักษณะเป็นแบบเลือกตอบและมีการให้คะแนนแบบ 0 , 1

1.สัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B และสัมประสิทธิ์โครงสร้าง เหมาะสำหรับนำไปใช้เมื่อกลุ่มตัวอย่างมีขนาดเล็กและขนาดกลาง คือ 30 60 120 240 คน และ 300 500 1,000 1,500 คน

2. สัมประสิทธิ์ ρ_{IRT} เหมาะสำหรับนำไปใช้เมื่อกลุ่มตัวอย่างมีขนาดกลาง คือ 1,000 1,500 คน

ข้อเสนอในการวิจัยครั้งต่อ

1.ควรทำการศึกษาจำนวนข้อสอบที่เหมาะสมในการประมาณค่าความเชื่อมั่นของแบบทดสอบตามโมเดลคะแนนจริงสัมพันธ์ ซึ่งประกอบด้วยสัมประสิทธิ์ r_B กับสัมประสิทธิ์โครงสร้าง และโมเดลโลจีสติก 3 พารามิเตอร์ ได้แก่ สัมประสิทธิ์ ρ_{IRT} และสัมประสิทธิ์ KR-20

2.ควรทำการเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่ประมาณค่าด้วยสัมประสิทธิ์ KR-20 สัมประสิทธิ์ r_B สัมประสิทธิ์โครงสร้าง และสัมประสิทธิ์ ρ_{IRT} จากค่าความเชื่อมั่นที่แท้จริงที่เป็นค่ากลางของทั้ง 4 วิธี

บรรณานุกรม

บรรณานุกรม

กนกทิพย์ พัฒนาพัวพันธ์. (2541). *สถิติอ้างอิงเพื่อการวิจัยทางการศึกษา.* เชียงใหม่: ภาควิชาประเมินผลและวิจัยการศึกษา มหาวิทยาลัยเชียงใหม่.

เกียรติศักดิ์ ส่องแสง. (2547). *ฟังก์ชั่นสารสนเทศของแบบทดสอบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์.* ปริญญานิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.

จรัญ จันทลักขณา. (2549). *สถิติการวิเคราะห์และการวางแผนงานวิจัย.* กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.

ฉัตรศิริ ปิยะพิมลสิทธิ์. (2540). *การเปรียบเทียบความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบเลือกตอบความสามารถในการอ่านภาษาไทยที่มีรูปแบบต่างกัน.* ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

ชุติมา สุขสว่าง. (2445). *การเปรียบเทียบความเชื่อมั่นของแบบทดสอบความสามารถด้านผลการคิดเอกนัยที่มีเนื้อหาและจำนวนตัวเลือกต่างกัน.* ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

ชูศรี วงค์รัตนะ. (2541). *เทคนิคการใช้สถิติเพื่อการวิจัย.* พิมพ์ครั้งที่ 7. กรุงเทพฯ: เทพเนรมิตการพิมพ์.

นงลักษณ์ วิรัชชัย. (2542). *โมเดลลิสเรล : สถิติวิเคราะห์สำหรับการวิจัย.* กรุงเทพฯ: ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

นพดล กองศิลป์ . (2540). *การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างและความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ของแบบทดสอบวัดความรู้ความเข้าใจทางภาษาแบบประยุกต์ตามโครงสร้างทางสติปัญญาของกิลฟอร์ด.* ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

นภดล ยิ่งยงสกุล. (2540). ความเที่ยง ความตรงของแบบทดสอบและความสัมพันธ์ของคะแนนสอบระหว่างการให้คะแนนตามทฤษฎีการทดสอบมาตรฐานเดิมกับการให้คะแนนตามทฤษฎีการตอบข้อสอบ. วิทยานิพนธ์ ศศ.ม. (การวัดผลและวิจัยการศึกษา). สงขลา: บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์.

ผจงจิต อินทสุวรรณ. (2545). *ทฤษฎีการตอบข้อคำถาม.* สถาบันวิจัยพฤติกรรมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.

บังอร กมลวัฒนา. (2542). *การเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบและการจัดเรียงปัญหาต่างกัน.* ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

บุญเชิด ภิญโญอนันต์พงษ์. (2533). Congeneric Part Reliability. *วารสารวัดผลการศึกษา* 12(34): 28 – 32 ; พฤษภาคม – สิงหาคม.

_______. (2536). เทคนิคการประมาณค่าความเชื่อถือได้ของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric Model). *วารสารวัดผลการศึกษา* 14(42) : 8 – 20 ; มกราคม – เมษายน.

_______. (2537). *การประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบคะแนนจริงสัมพันธ์.* ปริญญานิพนธ์ดุษฎีบัณฑิต สาขาการทดสอบและวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

_______. (2538). *แบบทดสอบคะแนนจริงสัมพันธ์ : การวิเคราะห์ทางสถิติ.* ภาควิชาการวัดผลและการวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.

_______. (2542). *รายงานการวิจัย สัมประสิทธิ์ r_b : การประมาณค่าความเชื่อมั่นสำหรับแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยความยากรายข้อต่างกัน.* ภาควิชาการวัดผลและการวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.

_______. (2549). *ทฤษฎีทดสอบคะแนนจริงสัมพันธ์ (Congeneric Test Theory).* เอกสารประกอบการประชุมสัมมนาทางวิชาการ การวิจัย การวัดและประเมินทางการศึกษาแห่งประเทศไทย ครั้งที่ 14, ภาควิชาการวัดผลและการวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.

เบญจพร ยนตจักรวิถี. (2540). *การเปรียบเทียบผลการวิเคราะห์ข้อสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ระหว่างทฤษฎีการทดสอบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ.* วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). สงขลา: มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี.

ปรมินทร์ อริเดช. (2547). *การเปรียบเทียบฟังก์ชั่นสารสนเทศในการใช้โมเดลโลจีสติก จีอาร์เอ็ม และ จีพีซีเอ็ม ของการวัดเจตคติแบบลิเคิร์ตและมาตรวัดแบบตัวเลือกบังคับตอบที่มีวิธีการให้คะแนนแบบสองค่า และแบบหลายค่า.* ปริญญานิพนธ์ดุษฎีบัณฑิต สาขาการทดสอบและวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

รังสรรค์ มณีเล็ก. (2540). *ผลของตัวแปรบางตัวต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์.* ปริญญานิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

เรวดี อินทสะระ. (2530). *การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตราระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ.* ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

วัฒนา ขัดสี. (2533). *การศึกษาความสัมพันธ์ระหว่างค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ข้อสอบและคะแนนความสามารถในการสอบโดยทฤษฎีมาตรฐานเดิมกับทฤษฎีการตอบข้อสอบของแบบทดสอบเลือกตอบ*. ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

วิชุดา กิจธรธรรม. (2544). *การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนของแบบวัดความถนัดทางวิชาชีพพยาบาลตามแนวคิดของฟลานาแกนด้วยวิธีการวิเคราะห์ความแปรปรวน วิธีการตามทฤษฎีการตอบข้อสอบ และวิธีการของคีต*. ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

วิรัตน์ ธรรมาภรณ์ และคณะ. (2545). *การวิเคราะห์ข้อสอบความถนัดทางการเรียนโดยใช้ทฤษฎีการทดสอบดั้งเดิมและทฤษฎีการตอบข้อสอบ*. สงขลา: คณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์

วิศิษฏ์ พหลยุทธ. (2539). *การประมาณค่าความเชื่อมั่นของแบบทดสอบที่มีคำถามหลายรูปแบบตามข้อตกลงแบบคู่ขนาน แบบคะแนนจริงสมมูลและแบบคะแนนจริงสัมพันธ์*. ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

ศิริชัย กาญจนวาสี. (2545). *ทฤษฎีการทดสอบแนวใหม่*. ปรับปรุงครั้งที่ 2. กรุงเทพฯ: ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

________. (2548). *ทฤษฎีการทดสอบแบบดั้งเดิม*. ปรับปรุงครั้งที่ 5. กรุงเทพฯ: ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

สนอง เที่ยงตรง. (2540). *การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างและความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ของการทดสอบวัดการประเมินค่าทางสัญลักษณ์และแบบทดสอบวัดการประเมินค่าทางภาษา แบบการแปลงรูปตามโครงสร้างทางสติปัญญาของกิลฟอร์ดสัมพันธ์*. ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

สยมภู รักษ์สังข์. (2547). *ความสำพันธ์ของผลการวิเคราะห์ข้อสอบแบบทฤษฎีมาตรฐานเดิม (CTT)และทฤษฎีการตอบข้อสอบ(IRT)*. ปริญญานิพนธ์การศึกษามหาบัณฑิต สาขาการวิจัยและสถิติทางการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

สุทธิวรรณ พีรศักดิ์โสภณ. สมกิจ กิจพูนวงศ์. (2545). *การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อประมาณค่าความเชื่อมั่นของเครื่องมือวัดตามข้อตกลงแบบคะแนนจริงสัมพันธ์.* คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม.

สุทธิวรรณ พีรศักดิ์โสภณ. (2546). *การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปเมื่อใช้วิธีการประมาณค่าโดยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์-ควอลส์.* ปริญญานิพนธ์ดุษฎีบัณฑิต สาขาการทดสอบและวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

สุมนา โสตถิผลอนันต์. (2539). *การศึกษาวิธีการเทียบคะแนนแบบเส้นตรงตามแบบจำลองคะแนนจริงสัมพันธ์สัมพันธ์*. ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.

อ่อนนุช เต็มเปี่ยม. (2545). *การศึกษาหลักฐานความเที่ยงตรงเชิงโครงสร้างและความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ของแบบทดสอบความสามารถทางสมองสัมพันธ์*. ปริญญานิพนธ์มหาบัณฑิต สาขาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

Allen, M.J. and Yen, W.M. (1979). *Introduction to Measurement Theory.* California. Brook/Cole Publishing Company.

Armor, D. J. (1974). *Theta reliability and factor scaling*. In Costner, H. L. (ed.), Sociological Methodology. Jossey-Bass, San Francisco. 17-50.

Baker, Frank B. (1992). *Item Response Theory: Parameter Estimation Techniques.* New York: Marcel Dekker, Inc.

Byrne, Barbara M. (1998). *Structural Equation Modeling.* Lawrence Erlbaum Associates, Inc., Publishers. New Jersey.

Dimitrov, Dimiter M. (2003). Marginal True-Score Measures and Reliability for Birary Item as a function of Their IRT Parameters. *Applied Psychological Measurement.* 27, 440-458.

Emberson, S.E. & Reise. S.T. (2000). *Item Response Theory for Psychologists*. New Jersey:Lawrence Erlbaum Associates.

Feldt, L.S. (1975). Estimation of the Reliability of a Test Devided into Two Parts of Unequal Lengths. *Psychometrika,* 40, 557 – 561.

Gilmer, J.S.& Feldt, L.S. (1983). Reliability Estimation for a Test with Parts of Unknown Lengths. *Psychometrika.* 48, 99 – 111.

Gulliksen, Harold. (1950). *Theory of Mental Tests*. New York: John Wiley & Sons, Inc.

Hair, Joseph F. And Other. (1995). *Multivariate Data Analysis With Readings.* Fourth Edition.Prentice-Hall,Inc.New Jersey.

Hambleton, Ronald K. and Swaminathan, Hariharan. (1985). *Item Response Theory.* U.S.A.: Kluwer -Nijhoff Publishing.

Hambleton, R. K., and Cook, L. L. (1977). Latent Trait Models and Their Use in the Analysis of Educational Test Data. *Journal of Educational Measurement*. 14.

Hays, W.L. (1988). *Statistics*. New York: Holt, Rinehart and Winston, Inc.

Heise, D. R., and Bohrnstedt, G. W. (1970). Validity, invalidity and reliability. In Borgatta, E. F. and Bohrnstedt, G. W. (eds.), *Sociological Methodology.* Jossey-Bass, San Francisco. 104-129.

Hulin, C. L., Lissak, R.I. and Drasgow, F. (1982). Recovery of two and Three-parameter logistic item characteristic curve: A monte carlo study. *Applied Psychological Measurement.* 6, 249 – 260.

Jöreskog, K.G. (1971). Statistical Analysis of Sets of Congeneric Test. *Psychometrika,* 36, 109 – 133.

Kolen, Michael J., Hanson, Bradley A. & Brennan, Robert L. (1992, Winter). Conditional Standard Error of Measurement for Scale Score. *Journal of Educational Measurement.* 29(4): 285-307.

Kristof, W. (1974). Estimation of Reliability and True Score Variance from a Split of a Test into Three Arbitary Parts. *Psychometrika*, 39, 491 – 499.

Linn, Robert L. (1989). *Educational Measurement*. New York: Macmillan Publishing Company.

Liou, M. (1989). A note on reliability estimation for a test with components of unknown functional lengths. *Psychometrika*, 54, 153-163.

Lord, F. M. (1980). *Applicational of Item Response Theory to Practical Testing Problems*. Hillsdale, New York: Lawrence Erlbaum Associates.

Lord, Frederic M. and Novick, Melvin R. (1968). *Statistical Theories of Mental Test Scores*. Canada: Addison-Wesley Publishing Company, Inc.

Lucke, Joseph F. (2005). The α and the ω of Congeneric Test Theory : An Extension of Reliability and Internal Consistency to Heterogeneous Test. *Applied Psychological Measurement*, 29 , 65-81.

_______. (2005). Rassling the Hog" : The Influence of Correlated Item Error on Internal Consistency, Classical Reliability, and Congeneric Reliability. *Applied Psychological Measurement*, 29 , 106-125.

MacCann,R.A. (1989). A Comparison of Two Observed-scores Equating Methods that Assume Equally Reliable Congeneric Test, *Applied Psychological Measurement.* 13(3) : 236 – 276.

Oosterhof,Albert C. and Pamela K. (1984). Comparison of Difficulties and Reliabilities of Qualitative Word Problam in Completion and Multiple-Choice Item Forates, *Applied Psychological Measurement*. 8: 287-293.

Parshall, Cynthia G.Houghton,Pansy Du B.Kromrey, Jeffrey D. (1995). Equating Error and Statistical Bias in Small Sample Linear Equating. *Journal of Educational Measurement*. 32(1): 37-54.

Raju, N. S. (1977). A generalization of coefficient alpha. *Psychometrika*, 42, 549-565.

_______. (1979). Note on two generalizations of coefficient alpha. *Psychometrika*, 44, 347-349.

Raykov, T. (1997). Estimation of Composite Reliability for Congeneric Measures. *Applied Psychological Measurement*, 21 , 173-184.

______. (1998). A Method for Obtaining Standard Error and Confidence Interval of Composite Reliability for Congeneric Items. *Applied Psychological Measurement*, 22 , 369-374.

_______. (1998). Coefficient Alpha and Composite Reliability With Interrelated Nonhomogeneous Item. *Applied Psychological Measurement*, 22 , 375-385.

_______. (2001). Bias of Coefficient α for Fixed Congeneric Measures With Correlated Error. *Applied Psychological Measurement*, 25 , 69-76.

Raykov, T. and Grayson,David. (2003). A Test for Chang of Composite Reliability in Scale Development. *Multivariate Behavioral Research*, 38 , 146-159.

Reuterberg, Sven-Eric. Gustafsson, Jan-Eric. (1992). Confirmatory Factor Analysis and Reliability : Testing Measurement Model Assumptions. *Education and Psychological Measurement.* 52, 795-811.

Swaminathan, H. and Gifford, J. A. (1983). Estimation of parameters in the three-parameter latent trait model. In D.J. Weiss (ED.) *New Horizons in Testing: Latent Traits Test Theory and Computerized Adaptive Testing*. New York: Academic Press.

Werts, C. E., Rock, R. D., Linn, R. L., and Joreskog, K. G. (1973). A Congeneric method for Platonic True Scores. *Educational and Psychological Measurement*, 33, 331-318.

ภาคผนวก

ภาคผนวก ก
การสร้างและพัฒนาเครื่องมือการวิจัย

- รายนามผู้เชี่ยวชาญ
- การวิเคราะห์องค์ประกอบเชิงสำรวจ
- การวิเคราะห์องค์ประกอบเชิงยืนยัน
- การวิเคราะห์ค่าความยาก อำนาจจำแนก การเดาและฟังก์ชั่นสารสนเทศของข้อสอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าและค่าฟังก์ชั่นสารสนเทศของแบบทดสอบตามทฤษฎีการตอบข้อสอบ(IRT) ด้วยโมเดลโลจีสติก 3 พารามิเตอร์จากโปรแกรม BOLOG-MG3

รายนามผู้เชี่ยวชาญประเมินความสอดคล้อง

ผู้เชี่ยวชาญการสอนคณิตศาสตร์

1.อาจารย์ลัดดา มิตกิตติ ครู คศ.3 โรงเรียนบ้านแม่ขาน จังหวัดเชียงใหม่

2.อาจารย์นิยม ไชยวงศ์ ครู คศ.3 โรงเรียนวัดทุ่งหลุก จังหวัดเชียงใหม่

3.อาจารย์อำไพ ใจวัง ครู คศ.3 โรงเรียนวัดสารภีวิทยาคาร จังหวัดเชียงใหม่

4.อาจารย์ปัทมา อินเทพ ครู คศ.3 โรงเรียนดอนชัยวิทยาคาร จังหวัดเชียงใหม่

5.อาจารย์ปรีชา ต๊ะมา ครู คศ.3 โรงเรียนวัดท่ากาน จังหวัดเชียงใหม่

ผู้เชี่ยวชาญด้านวัดผล

1.รองศาสตราจารย์ นิโลบล นิ่มกิ่งรัตน์ ภาควิชาประเมินผลและวิจัยการศึกษา
มหาวิทยาลัยเชียงใหม่

2.รองศาสตราจารย์ ดร.เกียรติสุดา ศรีสุข ภาควิชาประเมินผลและวิจัยการศึกษา
มหาวิทยาลัยเชียงใหม่

3.ผู้ช่วยศาสตราจารย์ ดร.สุชาติ ลี้ตระกูล โปรแกรมวิจัยและประเมินผลการศึกษา
มหาวิทยาลัยราชภัฏเชียงราย

4.อาจารย์ ดร.ปรมินทร์ อริเดช โปรแกรมวิจัยและประเมินผลการศึกษา
มหาวิทยาลัยราชภัฏเชียงราย

5.อาจารย์ ดร.นิคม นาคอ้าย คณะครุศาสตร์
มหาวิทยาลัยราชภัฏพิบูลสงคราม

ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ

(Exploratiry Factor Analysis)

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
IT1	.83300	.373162	1000
IT2	.75700	.429110	1000
IT3	.77900	.415128	1000
IT4	.55800	.496873	1000
IT5	.76400	.424835	1000
IT6	.70700	.455366	1000
IT7	.46700	.499159	1000
IT8	.38800	.487538	1000
IT9	.77100	.420399	1000
IT10	.52400	.499674	1000
IT11	.69000	.462725	1000
IT12	.74900	.433805	1000
IT13	.59700	.490746	1000
IT14	.62600	.484106	1000
IT15	.63700	.481105	1000
IT16	.63100	.482775	1000
IT17	.53100	.499288	1000
IT18	.39400	.488879	1000
IT19	.39700	.489521	1000
IT20	.49300	.500201	1000
IT21	.51300	.500081	1000
IT22	.61500	.486839	1000
IT23	.43100	.495464	1000
IT24	.55200	.497537	1000
IT25	.49600	.500234	1000
IT26	.64500	.478753	1000
IT27	.58600	.492795	1000
IT28	.53400	.499092	1000
IT29	.38900	.487767	1000
IT30	.60400	.489309	1000
IT31	.68200	.465932	1000
IT32	.70500	.456271	1000
IT33	.74300	.437198	1000
IT34	.76900	.421683	1000
IT35	.49100	.500169	1000
IT36	.49300	.500201	1000
IT37	.54700	.498035	1000
IT38	.69100	.462312	1000
IT39	.65100	.476892	1000
IT40	.65700	.474949	1000
IT41	.62500	.484365	1000
IT42	.66800	.471167	1000
IT43	.37900	.485381	1000
IT44	.66800	.471167	1000
IT45	.58300	.493310	1000
IT46	.54800	.497940	1000
IT47	.48600	.500054	1000
IT48	.38100	.485876	1000
IT49	.51600	.499994	1000
IT50	.43400	.495873	1000
IT51	.60100	.489938	1000
IT52	.64800	.477833	1000
IT53	.29700	.457165	1000
IT54	.55600	.497103	1000
IT55	.27600	.447240	1000
IT56	.32400	.468234	1000
IT57	.57000	.495323	1000
IT58	.46700	.499159	1000
IT59	.51100	.500129	1000
IT60	.43200	.495602	1000

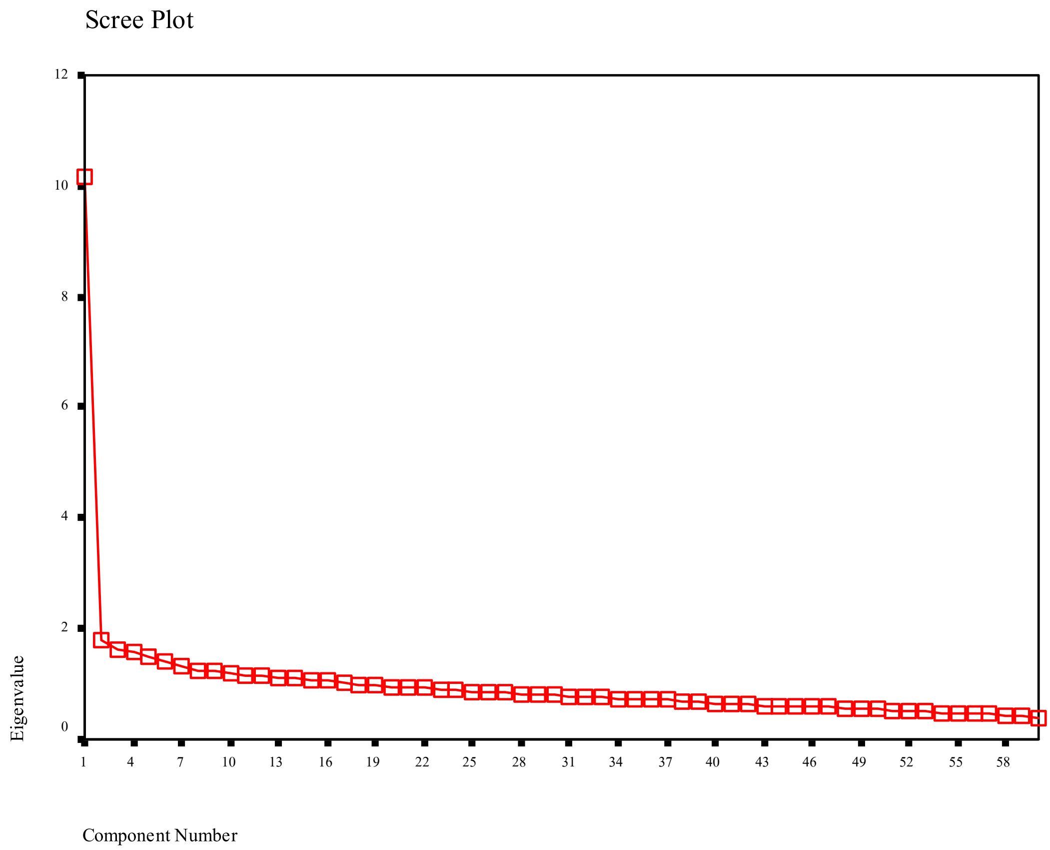
KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.920
Bartlett's Test of Sphericity	Approx. Chi-Square	11468.84
	df	1770
	Sig.	.000

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	10.145	16.909	16.909	10.145	16.909	16.909	3.270	5.450	5.450
2	1.779	2.966	19.874	1.779	2.966	19.874	3.142	5.237	10.687
3	1.622	2.704	22.578	1.622	2.704	22.578	2.502	4.170	14.857
4	1.593	2.655	25.234	1.593	2.655	25.234	2.459	4.098	18.955
5	1.480	2.467	27.701	1.480	2.467	27.701	1.919	3.199	22.154
6	1.411	2.352	30.052	1.411	2.352	30.052	1.898	3.163	25.317
7	1.325	2.209	32.261	1.325	2.209	32.261	1.651	2.752	28.068
8	1.260	2.099	34.360	1.260	2.099	34.360	1.641	2.734	30.803
9	1.221	2.034	36.395	1.221	2.034	36.395	1.639	2.732	33.535
10	1.189	1.981	38.376	1.189	1.981	38.376	1.521	2.536	36.070
11	1.144	1.906	40.282	1.144	1.906	40.282	1.507	2.512	38.583
12	1.138	1.896	42.178	1.138	1.896	42.178	1.391	2.318	40.901
13	1.108	1.847	44.025	1.108	1.847	44.025	1.378	2.297	43.197
14	1.096	1.826	45.851	1.096	1.826	45.851	1.234	2.057	45.254
15	1.053	1.755	47.606	1.053	1.755	47.606	1.224	2.040	47.294
16	1.050	1.750	49.356	1.050	1.750	49.356	1.143	1.905	49.199
17	1.023	1.705	51.061	1.023	1.705	51.061	1.117	1.862	51.061
18	.987	1.644	52.705						
19	.973	1.622	54.327						
20	.952	1.587	55.914						
21	.947	1.579	57.493						
22	.930	1.550	59.043						
23	.902	1.504	60.547						
24	.884	1.474	62.021						
25	.863	1.438	63.460						
26	.853	1.422	64.882						
27	.844	1.407	66.289						
28	.825	1.375	67.663						
29	.821	1.368	69.031						
30	.800	1.333	70.365						
31	.785	1.308	71.673						
32	.771	1.286	72.959						
33	.752	1.254	74.213						
34	.741	1.235	75.447						
35	.733	1.222	76.670						
36	.717	1.194	77.864						
37	.710	1.183	79.047						
38	.694	1.157	80.204						
39	.670	1.117	81.322						
40	.656	1.093	82.414						
41	.644	1.074	83.488						
42	.630	1.050	84.538						
43	.607	1.011	85.549						
44	.602	1.004	86.553						
45	.596	.993	87.546						
46	.583	.972	88.519						
47	.577	.962	89.481						
48	.563	.939	90.420						
49	.540	.899	91.319						
50	.538	.896	92.215						
51	.533	.888	93.104						
52	.512	.854	93.958						
53	.503	.839	94.796						
54	.487	.812	95.608						
55	.474	.789	96.398						
56	.465	.775	97.172						
57	.459	.765	97.938						
58	.431	.719	98.657						
59	.411	.685	99.342						
60	.395	.658	100.000						

Extraction Method: Principal Component Analysis.



จากผลการตรวจสอบความเป็นมิติเดียว(Unidimensional) ด้วยการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) โดยสกัดองค์ประกอบด้วยวิธีวิเคราะห์องค์ประกอบหลัก(Principle Component) ด้วยโปรแกรม SPSS พิจารณาค่าความแปรปรวนสูงสุดจากองค์ประกอบหลักตัวแรกมีสัดส่วนมากกว่าองค์ประกอบอื่นๆมาก โดยมีค่าไอเกนขององค์ประกอบแรกเท่ากับ 10.145 องค์ประกอบที่สองเท่ากับ 1.779 แสดงว่าแบบทดสอบมีความเป็นมิติเดียว

ผลการวิเคราะห์องค์ประกอบเชิงยืนยัน

(Confirmatory Factor Analysis)

```
                              Goodness of Fit Statistics

                               Degrees of Freedom = 1522
                 Minimum Fit Function Chi-Square = 1619.05 (P = 0.041)
      Normal Theory Weighted Least Squares Chi-Square = 1612.75 (P = 0.052)
                 Estimated Non-centrality Parameter (NCP) = 90.75
            90 Percent Confidence Interval for NCP = (0.0 ; 190.99)

                          Minimum Fit Function Value = 1.62
                 Population Discrepancy Function Value (F0) = 0.091
                90 Percent Confidence Interval for F0 = (0.0 ; 0.19)
            Root Mean Square Error of Approximation (RMSEA) = 0.0077
            90 Percent Confidence Interval for RMSEA = (0.0 ; 0.011)
              P-Value for Test of Close Fit (RMSEA < 0.05) = 1.00

                    Expected Cross-Validation Index (ECVI) = 2.23
            90 Percent Confidence Interval for ECVI = (2.14 ; 2.33)
                          ECVI for Saturated Model = 3.66
                        ECVI for Independence Model = 45.75

Chi-Square for Independence Model with 1770 Degrees of Freedom =
45585.42
                              Independence AIC = 45705.42
                                  Model AIC = 2228.75
                                Saturated AIC = 3660.00
                              Independence CAIC = 46059.88
                                  Model CAIC = 4048.34
                               Saturated CAIC = 14471.19

                             Normed Fit Index (NFI) = 0.96
                           Non-Normed Fit Index (NNFI) = 1.00
                        Parsimony Normed Fit Index (PNFI) = 0.83
                           Comparative Fit Index (CFI) = 1.00
                           Incremental Fit Index (IFI) = 1.00
                            Relative Fit Index (RFI) = 0.96

                               Critical N (CN) = 1021.12


                        Root Mean Square Residual (RMR) = 0.027
                              Standardized RMR = 0.027
                           Goodness of Fit Index (GFI) = 0.95
                     Adjusted Goodness of Fit Index (AGFI) = 0.94
                    Parsimony Goodness of Fit Index (PGFI) = 0.79
```

จากวิเคราะห์องค์ประกอบเชิงยืนยัน(Confirmatory Factor Analysis) ที่มีตัวแปรภายนอกแฝงเพียงตัวเดียว เพื่อทดสอบความสอดคล้องระหว่างโมเดลกับข้อมูลเชิงประจักษ์ ด้วยโปรแกรม LISREL พบว่า ได้ค่า $\chi^2 = 1612.75$ p=.052 GFI= .95 AGFI= .94 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

การวิเคราะห์วิเคราะห์ค่าความยาก อำนาจจำแนก การเดาและฟังก์ชั่นสารสนเทศของข้อสอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าและค่าฟังก์ชั่นสารสนเทศของแบบทดสอบตามทฤษฎีการตอบข้อสอบ(IRT) ด้วยโมเดลโลจีสติก 3 พารามิเตอร์จากโปรแกรม BOLOG-MG3

```
1
                                BILOG-MG V3.0
                             REV 19990329.1300

       BILOG-MG ITEM MAINTENANCE PROGRAM: LOGISTIC ITEM RESPONSE MODEL


                    ***  BILOG-MG ITEM MAINTENANCE PROGRAM  ***


                               *** PHASE 2  ***


 recode 0 to 2

 >CALIB ACCel = 1.0000,

        TPRior;

 CALIBRATION PARAMETERS
 ======================

 MAXIMUM NUMBER OF EM CYCLES:                     20

 MAXIMUM NUMBER OF NEWTON CYCLES:                  2
 CONVERGENCE CRITERION:                           0.0100
 ACCELERATION CONSTANT:                           1.0000

 LATENT DISTRIBUTION:                           NORMAL PRIOR FOR EACH GROUP
 PLOT EMPIRICAL VS. FITTED ICC'S:               NO
 DATA HANDLING:                                 DATA ON SCRATCH FILE
 CONSTRAINT DISTRIBUTION ON ASYMPTOTES:         YES
 CONSTRAINT DISTRIBUTION ON SLOPES:             YES
 CONSTRAINT DISTRIBUTION ON THRESHOLDS:         YES
 SOURCE OF ITEM CONSTRAINT DISTIBUTION
       MEANS AND STANDARD DEVIATIONS:           PROGRAM DEFAULTS
1
 -----------------------------------------------------------------------


                       ******************************

                          CALIBRATION OF MAINTEST

                                  TEST0001

                       ******************************
```

METHOD OF SOLUTION:

EM CYCLES (MAXIMUM OF 20)
FOLLOWED BY NEWTON-RAPHSON STEPS (MAXIMUM OF 2)

QUADRATURE POINTS AND PRIOR WEIGHTS:

	1	2	3	4	5
POINT	-0.4000E+01	-0.3429E+01	-0.2857E+01	-0.2286E+01	-0.1714E+01
WEIGHT	0.7648E-04	0.6387E-03	0.3848E-02	0.1673E-01	0.5245E-01

	6	7	8	9	10
POINT	-0.1143E+01	-0.5714E+00	-0.8882E-15	0.5714E+00	0.1143E+01
WEIGHT	0.1186E+00	0.1936E+00	0.2280E+00	0.1936E+00	0.1186E+00

	11	12	13	14	15
POINT	0.1714E+01	0.2286E+01	0.2857E+01	0.3429E+01	0.4000E+01
WEIGHT	0.5245E-01	0.1673E-01	0.3848E-02	0.6387E-03	0.7648E-04

CONSTRAINT DISTRIBUTIONS ON ITEM PARAMETERS
(THRESHOLDS, NORMAL; SLOPES, LOG-NORMAL; GUESSING, BETA)

	THRESHOLDS		SLOPES		ASYMPTOTES	
ITEM	MU	SIGMA	MU	SIGMA	ALPHA	BETA
ITEM0001	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0002	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0003	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0004	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0005	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0006	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0007	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0008	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0009	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0010	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0011	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0012	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0013	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0014	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0015	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0016	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0017	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0018	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0019	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0020	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0021	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0022	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0023	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0024	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0025	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0026	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0027	0.000	2.000	1.000	1.649	11.00	11.00

ITEM0028	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0029	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0030	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0031	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0032	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0033	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0034	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0035	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0036	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0037	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0038	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0039	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0040	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0041	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0042	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0043	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0044	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0045	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0046	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0047	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0048	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0049	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0050	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0051	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0052	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0053	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0054	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0055	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0056	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0057	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0058	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0059	0.000	2.000	1.000	1.649	11.00	11.00
ITEM0060	0.000	2.000	1.000	1.649	11.00	11.00

[E-M CYCLES]

-2 LOG LIKELIHOOD =   75181.567

CYCLE   1;   LARGEST CHANGE=   6.02061

-2 LOG LIKELIHOOD =   70451.314

CYCLE   2;   LARGEST CHANGE=   0.88930

-2 LOG LIKELIHOOD =   70083.471

CYCLE   3;   LARGEST CHANGE=   0.28815

-2 LOG LIKELIHOOD =   70040.254

CYCLE   4;   LARGEST CHANGE=   0.16607

```
-2 LOG LIKELIHOOD =      70034.866

CYCLE      5;   LARGEST CHANGE=   0.09052

-2 LOG LIKELIHOOD =      70035.736

CYCLE      6;   LARGEST CHANGE=   0.13572

-2 LOG LIKELIHOOD =      70036.396

CYCLE      7;   LARGEST CHANGE=   0.06145

-2 LOG LIKELIHOOD =      70038.138

CYCLE      8;   LARGEST CHANGE=   0.02152

-2 LOG LIKELIHOOD =      70038.407

CYCLE      9;   LARGEST CHANGE=   0.03050

-2 LOG LIKELIHOOD =      70038.650

CYCLE     10;   LARGEST CHANGE=   0.01204

-2 LOG LIKELIHOOD =      70038.564

CYCLE     11;   LARGEST CHANGE=   0.00282


[NEWTON CYCLES]


-2 LOG LIKELIHOOD:        70038.6003

CYCLE    12;   LARGEST CHANGE=   0.00173

INTERVAL COUNTS FOR COMPUTATION OF ITEM CHI-SQUARES
 -----------------------------------------------------------------------
        86.    72.    71.    80.   151.   146.   192.   113.    89.
 -----------------------------------------------------------------------

INTERVAL AVERAGE THETAS
 -----------------------------------------------------------------------
     -1.877 -1.438 -1.019 -0.584 -0.144  0.209  0.596  1.054  1.592
 -----------------------------------------------------------------------
```

1

SUBTEST TEST0001; ITEM PARAMETERS AFTER CYCLE 12

ITEM	INTERCEPT S.E.	SLOPE S.E.	THRESHOLD S.E.	LOADING S.E.	ASYMPTOTE S.E.	CHISQ (PROB)	DF
ITEM0001	1.355	1.445	-0.937	0.822	0.392	10.4	8.0
	0.256*	0.197*	0.258*	0.112*	0.087*	(0.2355)	
ITEM0002	0.855	1.497	-0.571	0.832	0.313	9.3	9.0
	0.219*	0.204*	0.198*	0.113*	0.071*	(0.4090)	
ITEM0003	0.716	1.065	-0.672	0.729	0.387	24.1	9.0
	0.275*	0.183*	0.340*	0.126*	0.089*	(0.0042)	
ITEM0004	-0.294	1.072	0.274	0.731	0.211	22.8	9.0
	0.216*	0.159*	0.175*	0.108*	0.053*	(0.0067)	
ITEM0005	0.951	1.546	-0.615	0.840	0.310	6.7	8.0
	0.217*	0.212*	0.196*	0.115*	0.072*	(0.5738)	
ITEM0006	0.209	1.087	-0.192	0.736	0.343	10.0	9.0
	0.273*	0.183*	0.274*	0.124*	0.075*	(0.3467)	
ITEM0007	-1.534	1.360	1.128	0.806	0.302	8.4	9.0
	0.426*	0.317*	0.148*	0.188*	0.041*	(0.4903)	
ITEM0008	-2.317	1.555	1.490	0.841	0.280	3.7	9.0
	0.552*	0.388*	0.150*	0.210*	0.031*	(0.9323)	
ITEM0009	0.772	1.139	-0.677	0.751	0.356	23.6	9.0
	0.251*	0.169*	0.288*	0.111*	0.083*	(0.0049)	
ITEM0010	-0.804	1.478	0.544	0.828	0.255	6.6	9.0
	0.283*	0.247*	0.127*	0.139*	0.044*	(0.6751)	
ITEM0011	0.407	1.638	-0.249	0.854	0.273	9.9	9.0
	0.211*	0.227*	0.152*	0.118*	0.058*	(0.3578)	
ITEM0012	1.047	2.019	-0.519	0.896	0.269	11.9	8.0
	0.189*	0.274*	0.137*	0.122*	0.058*	(0.1543)	
ITEM0013	-0.264	1.643	0.160	0.854	0.253	5.0	9.0
	0.229*	0.227*	0.123*	0.118*	0.046*	(0.8341)	
ITEM0014	0.089	1.736	-0.051	0.866	0.220	12.4	9.0
	0.197*	0.224*	0.117*	0.112*	0.047*	(0.1929)	
ITEM0015	-0.416	2.287	0.182	0.916	0.334	15.0	8.0
	0.284*	0.398*	0.100*	0.159*	0.041*	(0.0590)	
ITEM0016	-0.414	2.286	0.181	0.916	0.323	3.7	8.0
	0.269*	0.361*	0.096*	0.145*	0.040*	(0.8799)	
ITEM0017	-1.278	1.892	0.676	0.884	0.317	6.5	9.0
	0.351*	0.323*	0.103*	0.151*	0.035*	(0.6844)	

ITEM0018	-1.532	2.126	0.721	0.905	0.143	10.2	9.0
	0.266*	0.274*	0.066*	0.117*	0.025*	(0.3342)	
ITEM0019	-2.137	1.487	1.438	0.830	0.275	15.7	9.0
	0.516*	0.371*	0.149*	0.207*	0.033*	(0.0735)	
ITEM0020	-1.042	1.738	0.600	0.867	0.238	5.2	9.0
	0.290*	0.279*	0.099*	0.139*	0.037*	(0.8210)	
ITEM0021	-2.431	1.910	1.273	0.886	0.410	9.9	9.0
	0.680*	0.519*	0.134*	0.241*	0.030*	(0.3602)	
ITEM0022	-0.291	1.012	0.287	0.711	0.313	16.7	9.0
	0.297*	0.196*	0.252*	0.138*	0.067*	(0.0540)	
ITEM0023	-1.625	1.143	1.422	0.753	0.283	8.5	9.0
	0.441*	0.288*	0.188*	0.189*	0.042*	(0.4860)	
ITEM0024	-0.461	1.146	0.402	0.753	0.241	16.0	9.0
	0.255*	0.194*	0.174*	0.127*	0.054*	(0.0674)	
ITEM0025	-1.512	1.323	1.143	0.798	0.340	10.4	9.0
	0.444*	0.315*	0.161*	0.190*	0.042*	(0.3225)	
ITEM0026	0.086	1.167	-0.074	0.759	0.264	9.7	9.0
	0.230*	0.175*	0.204*	0.114*	0.063*	(0.3744)	
ITEM0027	-0.427	0.711	0.601	0.579	0.310	16.2	9.0
	0.311*	0.148*	0.356*	0.121*	0.072*	(0.0624)	
ITEM0028	-0.719	1.408	0.511	0.815	0.257	18.1	9.0
	0.284*	0.249*	0.137*	0.144*	0.048*	(0.0337)	
ITEM0029	-2.395	2.025	1.183	0.897	0.253	10.2	9.0
	0.499*	0.403*	0.095*	0.178*	0.027*	(0.3319)	
ITEM0030	-0.422	1.515	0.278	0.835	0.300	4.2	9.0
	0.272*	0.251*	0.145*	0.138*	0.050*	(0.9002)	
ITEM0031	0.539	1.558	-0.346	0.842	0.215	11.7	9.0
	0.176*	0.176*	0.137*	0.095*	0.053*	(0.2302)	
ITEM0032	0.282	1.831	-0.154	0.878	0.339	8.8	9.0
	0.236*	0.258*	0.143*	0.124*	0.053*	(0.4576)	
ITEM0033	0.918	1.837	-0.500	0.878	0.278	9.7	8.0
	0.198*	0.245*	0.150*	0.117*	0.061*	(0.2890)	
ITEM0034	0.888	1.589	-0.559	0.846	0.336	8.5	8.0
	0.228*	0.228*	0.197*	0.121*	0.071*	(0.3818)	
ITEM0035	-0.692	1.567	0.442	0.843	0.177	6.7	9.0
	0.221*	0.216*	0.100*	0.116*	0.038*	(0.6696)	
ITEM0036	-1.078	1.350	0.798	0.804	0.266	11.0	9.0
	0.321*	0.254*	0.136*	0.151*	0.043*	(0.2748)	
ITEM0037	-1.009	2.208	0.457	0.911	0.284	4.8	9.0

	0.304*	0.342*	0.086*	0.141*	0.034*	(0.8512)	
ITEM0038	0.586	1.976	-0.297	0.892	0.238	8.0	8.0
	0.185*	0.250*	0.117*	0.113*	0.050*	(0.4380)	
ITEM0039	-0.075	1.618	0.047	0.851	0.310	3.7	9.0
	0.250*	0.248*	0.149*	0.131*	0.052*	(0.9284)	
ITEM0040	-0.225	1.752	0.128	0.868	0.353	9.4	9.0
	0.283*	0.301*	0.144*	0.149*	0.050*	(0.3992)	
ITEM0041	0.148	1.183	-0.125	0.764	0.204	22.1	9.0
	0.185*	0.144*	0.166*	0.093*	0.054*	(0.0085)	
ITEM0042	0.075	2.090	-0.036	0.902	0.304	4.1	8.0
	0.230*	0.308*	0.113*	0.133*	0.046*	(0.8485)	
ITEM0043	-2.090	1.410	1.483	0.816	0.254	4.0	9.0
	0.481*	0.336*	0.155*	0.195*	0.033*	(0.9108)	
ITEM0044	0.194	1.558	-0.124	0.842	0.282	9.9	9.0
	0.223*	0.206*	0.153*	0.111*	0.055*	(0.3625)	
ITEM0045	-0.390	2.119	0.184	0.904	0.239	6.5	9.0
	0.219*	0.254*	0.088*	0.109*	0.036*	(0.6845)	
ITEM0046	-0.676	2.607	0.259	0.934	0.210	9.8	9.0
	0.237*	0.354*	0.067*	0.127*	0.032*	(0.3704)	
ITEM0047	-1.532	2.540	0.603	0.930	0.248	7.7	9.0
	0.324*	0.354*	0.069*	0.130*	0.028*	(0.5607)	
ITEM0048	-1.608	1.533	1.049	0.838	0.184	12.2	9.0
	0.330*	0.271*	0.105*	0.148*	0.032*	(0.2035)	
ITEM0049	-1.623	1.818	0.893	0.876	0.348	4.3	9.0
	0.441*	0.383*	0.111*	0.185*	0.035*	(0.8908)	
ITEM0050	-1.421	1.541	0.922	0.839	0.227	7.5	9.0
	0.326*	0.265*	0.109*	0.144*	0.035*	(0.5847)	
ITEM0051	-0.174	2.115	0.082	0.904	0.227	10.7	9.0
	0.208*	0.258*	0.091*	0.110*	0.038*	(0.3003)	
ITEM0052	-0.071	1.487	0.048	0.830	0.308	6.9	9.0
	0.259*	0.247*	0.167*	0.138*	0.057*	(0.6426)	
ITEM0053	-2.044	1.127	1.814	0.748	0.172	13.0	9.0
	0.412*	0.266*	0.209*	0.177*	0.032*	(0.1633)	
ITEM0054	-0.647	1.171	0.553	0.760	0.288	10.2	9.0
	0.302*	0.218*	0.183*	0.141*	0.054*	(0.3332)	
ITEM0055	-2.824	1.733	1.629	0.866	0.180	4.4	9.0
	0.558*	0.395*	0.139*	0.197*	0.024*	(0.8795)	
ITEM0056	-3.757	2.454	1.531	0.926	0.253	9.2	9.0
	0.945*	0.683*	0.118*	0.258*	0.022*	(0.4148)	

```
ITEM0057 |-0.995   | 2.077   | 0.479   | 0.901   |    0.323   |     4.7   8.0
         | 0.337*  | 0.383*  | 0.097*  | 0.166*  |    0.038*  | (0.7880)
         |         |         |         |         |            |
ITEM0058 |-1.186   | 2.434   | 0.487   | 0.925   |    0.177   |    11.8   9.0
         | 0.271*  | 0.335*  | 0.065*  | 0.127*  |    0.029*  | (0.2230)
         |         |         |         |         |            |
ITEM0059 |-1.468   | 1.468   | 1.000   | 0.826   |    0.343   |     5.8   9.0
         | 0.419*  | 0.318*  | 0.141*  | 0.179*  |    0.040*  | (0.7644)
         |         |         |         |         |            |
ITEM0060 |-1.295   | 0.911   | 1.421   | 0.674   |    0.250   |    11.9   9.0
         | 0.368*  | 0.218*  | 0.223*  | 0.161*  |    0.049*  | (0.2196)
-----------------------------------------------------------------------
                                                      * STANDARD ERROR

    LARGEST CHANGE =    0.001733                          600.1 530.0
                                                         (0.0185)
-----------------------------------------------------------------------


PARAMETER       MEAN  STN DEV
-----------------------------------
ASYMPTOTE      0.276    0.059
SLOPE          1.635    0.431
LOG(SLOPE)     0.456    0.274
THRESHOLD      0.402    0.679


QUADRATURE POINTS, POSTERIOR WEIGHTS, MEAN AND S.D.:


                  1            2            3            4            5
POINT        -0.3851E+01 -0.3302E+01 -0.2753E+01 -0.2205E+01 -0.1656E+01
POSTERIOR     0.1728E-03  0.1438E-02  0.7929E-02  0.2872E-01  0.6627E-01


                  6            7            8            9           10
POINT        -0.1107E+01 -0.5577E+00 -0.8757E-02  0.5402E+00  0.1089E+01
POSTERIOR     0.1019E+00  0.1439E+00  0.2185E+00  0.2243E+00  0.1355E+00


                 11           12           13           14           15
POINT         0.1638E+01  0.2187E+01  0.2736E+01  0.3285E+01  0.3834E+01
POSTERIOR     0.5625E-01  0.1292E-01  0.1986E-02  0.1873E-03  0.9743E-05

MEAN          0.00000
S.D.          1.00000


    94472 BYTES OF NUMERICAL WORKSPACE USED OF   8192000 AVAILABLE IN
PHASE-2

     5868 BYTES OF CHARACTER WORKSPACE USED OF   2048000 AVAILABLE IN
PHASE-2
          02/06/2008  23:24:21
```

การวิเคราะห์ค่าความยาก อำนาจจำแนก การเดาและฟังก์ชั่นสารสนเทศของข้อสอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าและค่าฟังก์ชั่นสารสนเทศของแบบทดสอบตามทฤษฎีการตอบข้อสอบ(IRT) โดยพิจารณาค่าความยาก(b)อยู่ระหว่าง -2.5 ถึง +2.5 ค่าอำนาจจำแนก(a)มากกว่า 0.8 และค่าการเดา(c)น้อยกว่า 0.3 ด้วยโมเดลโลจีสติก 3 พารามิเตอร์จากโปรแกรม BOLOG-MG3 สามารถคัดเลือกข้อสอบที่มีคุณภาพได้จำนวน 50 ข้อ

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา คณิตศาสตร์ชั้นประถมศึกษาปีที่ 5

คำชี้แจง

1. แบบทดสอบฉบับนี้เป็นแบบทดสอบสำหรับงานวิจัย ขอให้นักเรียนตั้งใจทำตามความสามารถของนักเรียน ผลการทำแบบทดสอบไม่มีผลกระทบใดๆต่อนักเรียนและโรงเรียน ผู้วิจัยจะเก็บข้อมูลไว้เป็นความลับ

2. แบบทดสอบมีทั้งหมด 50 ข้อ แต่ละข้อมี 4 ตัวเลือก ให้นักเรียนทำเครื่องหมายกากบาท (X)ลงในกระดาษคำตอบให้ตรงกับตัวเลือกที่นักเรียนคิดว่าถูกเพียงตัวเดียว

เช่น ข้อที่ 0 นักเรียนคิดว่าตัวเลือก ข เป็นคำตอบที่ถูกก็ทำเครื่องหมาย (X) ดังนี้

ข้อ	ก	ข	ค	ง
0		X		

หากต้องการเปลี่ยนคำตอบให้นักเรียนลบเครื่องหมายเดิมออกก่อน แล้วค่อยเลือกคำตอบใหม่

3. การทดสอบครั้งนี้ใช้เวลาในการทดสอบ 1 ชั่วโมง ห้ามลงมือทำก่อนจนกว่าอาจารย์ผู้ควบคุมการสอบจะอนุญาต

4. ไม่อนุญาตให้ใช้เครื่องคิดเลข ในการทำข้อสอบ และห้ามขีดเขียนหรือทำเครื่องหมายใดๆลงในแบบทดสอบ

ขอให้นักเรียนโชคดีและขอขอบคุณที่ตั้งใจทำแบบทดสอบ

1. สี่สิบหกล้านสองหมื่นเก้าพันสามร้อยเจ็ด เขียนเป็นตัวเลขฮินดูอารบิกได้ตรงกับข้อใด

ก. 406,029,307 ข. 46,029,307 ค. 46,129,137 ง. 4,629,307

2. จำนวน 2,000,000 + 900,000 + 8,000 + 2 เป็นรูปกระจายของข้อใด

ก. 29,080,002 ข. 20,908,002 ค.2,980,002 ง. 2,908,002

3. ข้อใดเรียงลำดับจาก**มาก**ไปหา**น้อย**ได้ถูกต้อง

ก. 10,530,985 1,053,098 2,011,584

ข. 657,214 6,842,457 9,875,001

ค. 268,999 6,842,457 105,309

ง. 10,530,501 969,954 268,999

4. พิจารณาแบบรูปต่อไปนี้ 5 10 20 ... 55 80 จำนวนที่หายไปคือค่าใด

ก. 25 ข. 30 ค. 35 ง. 40

5. พิจารณาแบบรูปต่อไปนี้ 69 62 56 51 ... 44 จำนวนที่หายไปคือค่าใด

ก. 46 ข. 47 ค. 48 ง. 49

6. (2 x 3) x 4 มีค่า**ไม่เท่ากับ**จำนวนในข้อใด

ก. 2 x (3 x 4) ข. 3 x (4 x 2) ค. (4 x 2) x 3 ง. (4 x 1) x 2

7. ผลลบของ 1,548,922 - 1,396,447 มีค่าเท่ากับข้อใด

ก. 2,945,369 ข. 252,525 ค. 152,485 ง. 152,475

8. ผลลบของ 325,412 - 127,481 มีค่าเท่ากับข้อใด

ก. 452,893 ข. 442,893 ค. 208,031 ง. 197,931

9. ข้อใดต่อไปนี้**ไม่จริง**

ก. 22 + 56 = 78 ข. 951 – 741 = 210 ค. 9 x 80 = 72 ง. 1250 ÷ 5 = 250

10. 45 x 11 มีค่าเท่ากับข้อใด

ก. 34 ข. 56 ค. 90 ง. 495

11. 7,200 ÷ 16 มีค่าเท่ากับข้อใด

ก. 450 ข. 7,184 ค. 7,216 ง. 115,200

12. (32 + 8) x 100 ÷ 5 มีค่าเท่ากับข้อใด

ก. 800 ข. 192 ค. 145 ง. 135

13. (1,080 ÷ 15) + (265 – 257) มีค่าเท่ากับข้อใด

ก. 72 ข. 80 ค. 1,103 ง. 1,073

14. (5,820 $\div$ 20) - (45 x 2) มีค่าเท่ากับข้อใด

ก. 5,793 ข. 381 ค. 244 ง. 201

15. "ฝรั่งกิโลกรัมละ 12 บาท วันแรกขายได้ 30 กิโลกรัม วันที่สองขายได้ 45 กิโลกรัม วันสามขายได้ 60 กิโลกรัม รวมสามวันขายได้กี่บาท" โจทย์ปัญหานี้ตรงกับประโยคสัญลักษณ์ข้อใด

ก. (12 + 30) + (12 + 45) + (12 + 60) = $\square$

ข. 12 x (30 + 45 + 60) = $\square$

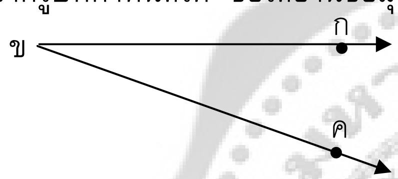
ค. (12 + 30) x (12 + 45) x (12 + 60) = $\square$

ง. (12 x 30) + 45 +60 = $\square$

16. แดนเลี้ยงไก่ไว้ 2,250 ตัว เป็นไก่ตัวผู้ 1,445 ตัว ที่เหลือเป็นไก่ตัวเมีย แดนขายไก่ตัวเมีย ราคาตัวละ 80 บาท แดนจะได้เงินจากการขายไก่เป็นเงินเท่าไร

ก. 295,600 บาท ข. 115,600 บาท ค. 113, 350 บาท ง. 64,400 บาท

17. จากรูปที่กำหนดให้ ข้อใดอ่านชื่อมุมได้ถูกต้อง



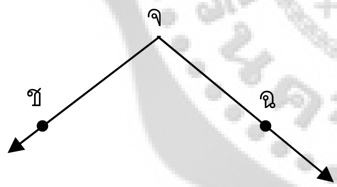
ก. มุม คขก หรือมุม กขค

ข. มุม ขกค หรือมุม คขก

ค. มุม ขคก

ง. มุม คกข

18. จากรูปที่กำหนดให้ ข้อใดกล่าวได้ถูกต้อง



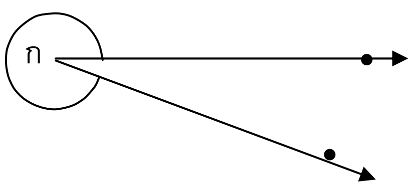
ก. $\overrightarrow{\text{ชจ}}$ และ $\overrightarrow{\text{จฉ}}$ เป็นแขนของมุม

ข. $\overrightarrow{\text{ชจ}}$ และ $\overrightarrow{\text{ฉจ}}$ เป็นแขนของมุม

ค. $\overrightarrow{\text{จช}}$ และ $\overrightarrow{\text{จฉ}}$ เป็นแขนของมุม

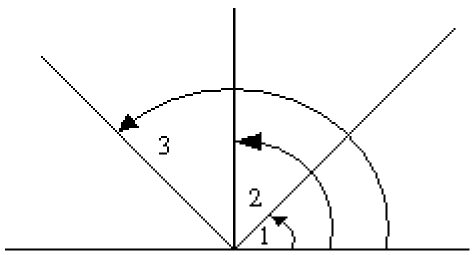
ง. $\overrightarrow{\text{จช}}$ และ $\overrightarrow{\text{ฉจ}}$ เป็นแขนของมุม

19. จากรูปที่กำหนดให้ เรียกว่ามุมอะไร



ก. มุมแหลม ข. มุมฉาก ค. มุมป้าน ง. มุมกลับ

จากรูปต่อไปนี้ ใช้ตอบคำถามข้อ 20 - 21



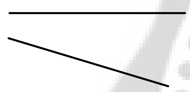
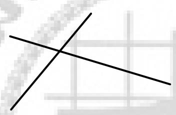
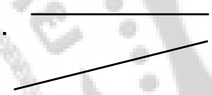
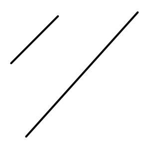
20. มุมที่ 1 มีขนาดเท่าไร

ก. 45° ข. 90° ค. 135° ง. 180°

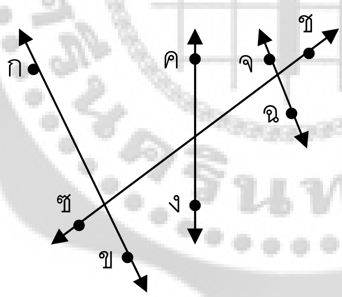
21. มุมที่ 2 มีขนาดเท่าไร

ก. 45° ข. 90° ค. 135° ง. 180°

22. ส่วนของเส้นตรงใดที่ขนานกัน

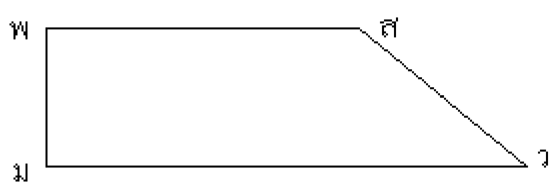
ก.  ข.  ค.  ง. 

23. จากรูป ข้อใดเขียนได้ถูกต้อง



ก. $\overleftrightarrow{\text{กข}}$ // $\overleftrightarrow{\text{คง}}$ ข. $\overleftrightarrow{\text{กข}}$ // $\overleftrightarrow{\text{จฉ}}$ ค. $\overleftrightarrow{\text{คง}}$ // $\overleftrightarrow{\text{จฉ}}$ ง. $\overleftrightarrow{\text{จฉ}}$ // $\overleftrightarrow{\text{ชซ}}$

24. จากรูป ส่วนของเส้นตรงใดขนานกัน



ก. $\overline{\text{พส}}$ // $\overline{\text{พม}}$ ข. $\overline{\text{สว}}$ // $\overline{\text{มว}}$ ค. $\overline{\text{พส}}$ // $\overline{\text{มว}}$ ง. $\overline{\text{สว}}$ // $\overline{\text{พส}}$

25. ในการเลือกหัวหน้าห้องชั้น ป.5 มีผู้สมัคร 4 คน นักเรียนทั้งห้องมีจำนวน 20 คน เลือกผู้สมัครดังหมายเลขต่อไปนี้

4 3 4 1 2 2 3 4 3 4 3 2 1 1 2 2 3 4 1 4

หมายเลขที่ได้คะแนนเป็นอันดับ 1 คือผู้สมัครหมายเลขใด

ก. หมายเลข 1 ข. หมายเลข 2 ค. หมายเลข 3 ง. หมายเลข 4

26. จากการสำรวจวันเกิดของนักเรียนห้องหนึ่ง จำนวน 30 คน พบว่ามีนักเรียนเกิดในวันต่าง ๆดังนี้

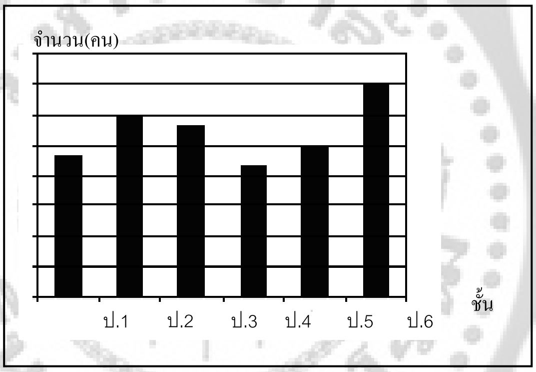
จ อ อ จ ศ พ พฤ ส ศ ส อา จ อา พ ศ ศ อ พ

พฤ ศ อา จ ศ จ พฤ ศ อ จ ส ศ

วันใดที่นักเรียนเกิดมากที่สุด

ก. จันทร์ (จ) ข. อังคาร (อ) ค. ศุกร์ (ศ) ง. เสาร์ (ส)

จากแผนภูมิแท่งแสดงจำนวนนักเรียนชั้นประถมศึกษาปีที่ 1– 6 ของโรงเรียนแห่งหนึ่ง **ใช้ตอบคำถามข้อ 27 – 30**



27. ชั้นที่มีจำนวนนักเรียนมากที่สุด มีจำนวนเท่าไร

ก. 24 คน ข. 21 คน ค. 18 คน ง. 17 คน

28. นักเรียนชั้น ป.1 มีจำนวนน้อยกว่าชั้น ป.2 กี่คน

ก. 1 คน ข. 3 คน ค. 4 คน ง. 5 คน

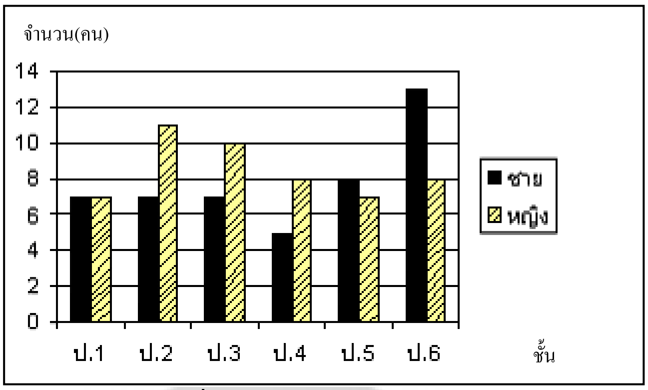
29. นักเรียนชั้น ป.3 มีจำนวนมากกว่ากว่าชั้น ป.4 กี่คน

ก. 1 คน ข. 3 คน ค. 4 คน ง. 5 คน

30. นักเรียนโรงเรียนนี้มีจำนวนนักเรียนทั้งหมด กี่คน

ก. 108 คน ข. 98 คน ค. 84 คน ง. 83 คน

จากแผนภูมิแท่งเปรียบเทียบแสดงจำนวนนักเรียนชาย-หญิงชั้นประถมศึกษาปีที่ 1– 6 ของโรงเรียนแห่งหนึ่ง **ใช้ตอบคำถามข้อ 31 - 39**



31. นักเรียนชายมีจำนวนสูงสุด มีจำนวนกี่คน

ก. 14 คน ข. 13 คน ค. 11 คน ง. 10 คน

32. นักเรียนหญิงมีจำนวนสูงสุด มีจำนวนกี่คน

ก. 13 คน ข. 11 คน ค. 10 คน ง. 8 คน

33. ชั้นที่มีนักเรียนชาย **เท่ากับ** นักเรียนหญิง คือชั้นใด

ก. ป.6 กับ ป.5 ข. ป.6 กับ ป.4 ค. ป.5 กับ ป.3 ง. ป.1

34. ชั้นใดที่มีนักเรียนชายและนักเรียนหญิง **ต่างกัน**มากที่สุด

ก. ป.6 ข. ป.4 ค. ป.3 ง. ป.2

35. ชั้นใดที่มีนักเรียนชายและนักเรียนหญิง **ต่างกัน** อยู่ 4 คน

ก. ป.4 ข. ป.3 ค. ป.2 ง. ป.1

36. ชั้น ป.3 มีนักเรียนชายและหญิง รวมกี่คน

ก. 18 คน ข. 17 คน ค. 15 คน ง. 14 คน

37. ชั้นที่มีนักเรียนชายและหญิงรวม 13 คน คือชั้นใด

ก. ป.1 ข. ป.3 ค. ป.4 ง. ป.5

38. โรงเรียนนี้มีนักเรียนชายกี่คน

ก. 98 คน ข. 84 คน ค. 51 คน ง. 47 คน

39. โรงเรียนนี้มีนักเรียนชายและหญิงรวมกี่คน

ก. 108 คน ข. 98 คน ค. 84 คน ง. 83 คน

40. นักเรียน 4 คน โยนลูกเต๋า(6 หน้า) คนละ 5 ครั้ง แล้วนำแต้มแต่ละครั้งมารวมกัน คำตอบของใคร **ไม่เกิดขึ้น**อย่างแน่นอน

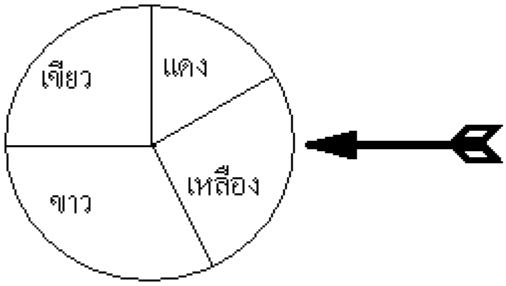
ก. คนที่ 1 รวมได้ 16 แต้ม

ข. คนที่ 2 รวมได้ 21 แต้ม

ค. คนที่ 3 รวมได้ 29 แต้ม

ง. คนที่ 4 รวมได้ 32 แต้ม

41. หมุนวงล้อดังรูป สีอะไรที่มี**โอกาสมากที่สุด**ที่จะหยุด ตรงกับลูกศร



ก. แดง ข. เหลือง ค. ขาว ง. เขียว

42. กล่องใบหนึ่งมีลูกกวาดสีเหลือง 1 เม็ด สีเขียว 2 เม็ด สีม่วง 3 เม็ด สีน้ำเงิน 4 เม็ด ทำการหยิบโดยไม่มองขึ้นมา 1 เม็ด ลูกกวาดสีใดที่มี**โอกาสมากที่สุด**ที่จะถูกหยิบขึ้นมา

ก. สีเหลือง ข. สีเขียว ค. สีม่วง ง. สีน้ำเงิน

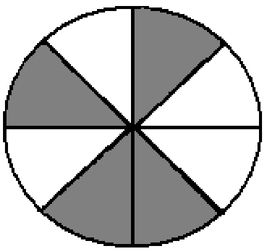
43. ข้อใดคือจำนวนคละ

ก. $\frac{5}{8}$ ข. $5\frac{1}{4}$ ค. $\frac{9}{3}$ ง. $\frac{4}{4}$

44. ข้อใดคือเศษเกิน

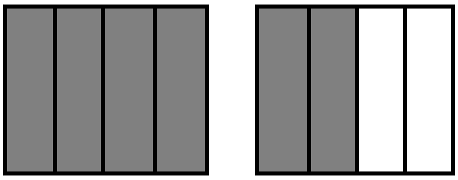
ก. $\frac{5}{8}$ ข. $5\frac{1}{4}$ ค. $\frac{1}{3}$ ง. $\frac{4}{4}$

45. จากรูป ส่วนที่แรงเงาสามารถเขียนเป็นเศษส่วนได้เท่ากับข้อใด



ก. $\frac{4}{4}$ ข. $\frac{4}{6}$ ค. $\frac{4}{8}$ ง. $\frac{6}{8}$

46. จากรูป ส่วนที่แรงเงาสามารถเขียนเป็นเศษส่วนได้เท่ากับข้อใด



ก. $4\frac{2}{4}$ ข. $\frac{4}{6}$ ค. $4\frac{2}{8}$ ง. $1\frac{2}{4}$

47. $\frac{2}{3}$ เป็นเศษส่วนอย่างต่ำของเศษส่วนในข้อใด

ก. $\frac{7}{14}$ ข. $\frac{24}{36}$ ค. $\frac{9}{12}$ ง. $\frac{24}{30}$

48. $\frac{7}{18}+\frac{3}{18}+\frac{4}{18}$ มีค่าเท่ากับเท่าไร

ก. $\frac{14}{18}$ ข. $\frac{14}{54}$ ค. $\frac{8}{54}$ ง. $\frac{8}{18}$

49. $\frac{15}{18}-\frac{3}{18}-\frac{4}{18}$ มีค่าเท่ากับเท่าไร

ก. $\frac{14}{18}$ ข. $\frac{14}{54}$ ค. $\frac{8}{54}$ ง. $\frac{8}{18}$

50. ถังใบหนึ่งมีน้ำอยู่ 5 ลิตร วันแรกใช้ไป $\frac{1}{5}$ ลิตร วันที่สองใช้ไป $\frac{2}{3}$ ลิตร เหลือน้ำในถังกี่ลิตร

ก. $\frac{24}{3}$ ลิตร ข. $4\frac{2}{15}$ ลิตร ค. $\frac{72}{15}$ ลิตร ง. $\frac{62}{75}$ ลิตร

ภาคผนวก ข

คำสั่งการแปลงค่าพารามิเตอร์เป็นค่าความแปรปรวน

คำสั่งการแปลงค่าพารามิเตอร์จากโมเดลโลจีสติก 3 พารามิเตอร์ เป็นค่าความแปรปรวน

```
COMPUTE p = .2646 - .118*a + .0187*(a**2).
COMPUTE s = .7427+.7081/a+.0074/(a**2).
COMPUTE ve = p*exp(-.5*((b/s)**2)).
COMPUTE X = (a*b)/sqrt(2*(1+a**2)).
COMPUTE erf =
(1+.278393*abs(X)+.230389*X**2+.000972*(abs(X))**3+.078108*X**4)**4.
COMPUTE erf = 1-1/erf.
IF(X<0)erf = -erf.
COMPUTE pi = (1-erf)/2.
COMPUTE vt = pi*(1-pi)-ve.
IF(vt<0)vt = 0.
COMPUTE ve = c*(1-c)*(1-pi)+ve*((1-c)**2).
COMPUTE pi = c+(1-c)*pi.
COMPUTE vt = pi*(1-pi)-ve.
IF(vt<0)vt = 0.
SET FORMAT = F8.3 ERROR = NONE RESULTS OFF HEATHER NO.
FLIP
         VARIABLES a b c pi ve vt.
VECTOR V = VAR001 TO VAR050.
COMPUTE Y = 0.
LOOP #I = 1 TO 50.
LOOP #J = 1 TO 50.
        COMPUTE Y = Y+SQRT(V(#I)*V(#J)).
END LOOP.
END LOOP.
FLIP VAR001 TO VAR024 Y.
COMPUTE roi = vt/(vt+ve).
SET RESULTS ON.
REPORT FORMAT = AUTOMATIC
   /VARIABLE = pi ' ' ve ' ' vt ' '
   /BREAK = (TOTAL)
   /SUMMARY = MAX(vt) 'True score variance:'
   /SUMMARY = SUBTRACT(SUM(ve) MAX(ve)) (vt(COMMA) (3)) 'Error
variance:'
   /SUMMARY = SUBTRACT(SUM(pi) MAX(pi)) (vt(COMMA) (3)) 'Marginal NR
score:' .
SELECT IF(CASE_LBL ~='Y').
RENAME VARIABLES (CASE_LBL = ITEM) (ve=var_err)(vt=var_tau).
VARIABLE LABEL pi 'item score'.
DESCRIPTIVES
     VARIABLE = pi
     /STATISTICS = VAR.
```

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ - ชื่อสกุล	นายกิตติศักดิ์ นิวรัตน์
วันเดือนปีเกิด	28 มกราคม 2520
สถานที่เกิด	เชียงใหม่
ที่อยู่ปัจจุบัน	80/2 หมู่ 3 ต.หนองหาร อ.สันทราย จ.เชียงใหม่ 50290
ตำแหน่งหน้าที่การงาน	อาจารย์
สถานที่ทำงานปัจจุบัน	มหาวิทยาลัยราชภัฏเชียงราย เชียงราย
ประวัติการศึกษา	
พ.ศ. 2542	ปทส. (ไฟฟ้าสื่อสาร) จากวิทยาลัยเทคนิคเชียงใหม่
พ.ศ. 2546	ศษ.ม. (วิจัยและสถิติการศึกษา) จากมหาวิทยาลัยเชียงใหม่
พ.ศ. 2554	กศ.ด. (การทดสอบและวัดผลการศึกษา) จากมหาวิทยาลัยศรีนครินทรวิโรฒ