

ชยุตม์ ภิรมย์สมบัติ: คุณสมบัติของตัวประมาณค่าความเข้มของอิทธิพล: การเปรียบเทียบระหว่างทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ (EFFECT MAGNITUDE ESTIMATORS' PROPERTIES: A COMPARISON BETWEEN CLASSICAL TEST THEORY AND ITEM RESPONSE THEORY) อ.ที่ปรึกษา: ผศ.ดร.สุชาติ บวรกิติวงศ์, 129 หน้า. ISBN 974-53-2253-9

การวิจัยครั้งนี้มีวัตถุประสงค์ 4 ข้อ ได้แก่ (1) เพื่อเปรียบเทียบคุณสมบัติของตัวประมาณค่าในด้านความลำเอียง ความคงเส้นคงวา และประสิทธิภาพสัมพัทธ์ระหว่างขนาดอิทธิพลที่ได้จากทฤษฎีการทดสอบแบบดั้งเดิม (d_{CTT}) ขนาดอิทธิพลที่ได้จากทฤษฎีการตอบสนองข้อสอบที่โมเดลประมาณค่าสอดคล้องกับข้อมูล (d_{IRT1}) ขนาดอิทธิพลที่ได้จากทฤษฎีการตอบสนองข้อสอบที่โมเดลประมาณค่าไม่สอดคล้องกับข้อมูล (d_{IRT2}) สัมประสิทธิ์สหสัมพันธ์ที่ได้จากทฤษฎีการทดสอบแบบดั้งเดิม (r_{CTT}) สัมประสิทธิ์สหสัมพันธ์ที่ได้จากทฤษฎีการตอบสนองข้อสอบที่โมเดลประมาณค่าสอดคล้องกับข้อมูล (r_{IRT1}) และสัมประสิทธิ์สหสัมพันธ์ที่ได้จากทฤษฎีการตอบสนองข้อสอบที่โมเดลประมาณค่าไม่สอดคล้องกับข้อมูล (r_{IRT2}) (2) เพื่อเปรียบเทียบค่าเฉลี่ยระหว่าง d_{CTT} , d_{IRT1} และ d_{IRT2} และระหว่าง r_{CTT} , r_{IRT1} และ r_{IRT2} (3) เพื่อศึกษาความสัมพันธ์ระหว่าง d_{CTT} และ d_{IRT1} และสร้างสมการถดถอยของ d_{IRT1} บน d_{CTT} และ (4) เพื่อศึกษาความสัมพันธ์ระหว่าง r_{CTT} และ r_{IRT1} และสร้างสมการถดถอยของ r_{IRT1} บน r_{CTT} ภายใต้สถานการณ์การสอบ 540 สถานการณ์ ตามเงื่อนไขของค่าความเข้มของอิทธิพลที่แท้จริง (.2, .5, .8, 1.2, 2.6) ขนาดกลุ่มตัวอย่าง (20, 50, 500, 2,000) ความยาวแบบสอบ (10, 50, 90) โมเดลฐาน (โมเดลโลจิสติกแบบหนึ่ง, สอง และสามพารามิเตอร์) และโมเดลประมาณค่า (โมเดลการทดสอบแบบดั้งเดิม, โมเดลการตอบสนองข้อสอบที่สอดคล้องกับข้อมูล และโมเดลการตอบสนองข้อสอบที่ไม่สอดคล้องกับข้อมูล)

ผลการวิจัยโดยสรุปพบว่า (1) ในภาพรวมตัวประมาณค่าที่มีความลำเอียงต่ำที่สุดคือ r_{IRT1} ตัวประมาณค่าที่มีความคงเส้นคงวาสูงสุดคือ r_{CTT} และตัวประมาณค่าที่มีประสิทธิภาพสัมพัทธ์สูงสุดคือ r_{IRT1} นอกจากนี้ r_{IRT1} ยังเป็นตัวประมาณค่าที่มีคุณสมบัติทุกด้านเป็นที่น่าพอใจที่สุด (2) ค่าเฉลี่ยของ d_{CTT} , d_{IRT1} และ d_{IRT2} มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดย d_{CTT} มีค่าเฉลี่ยสูงสุด เช่นเดียวกับค่าเฉลี่ยของ r_{CTT} , r_{IRT1} และ r_{IRT2} ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดย r_{CTT} มีค่าเฉลี่ยสูงสุด (3) สัมประสิทธิ์สหสัมพันธ์ระหว่าง d_{CTT} และ d_{IRT1} มีค่า .626 และมีนัยสำคัญทางสถิติที่ระดับ .05 สมการถดถอยของ d_{IRT1} บน d_{CTT} ในรูปคะแนนดิบคือ $d_{IRT1} = .004 + .065d_{CTT}$ สมการในรูปคะแนนมาตรฐานคือ $Z_{d_{IRT1}} = .626Z_{d_{CTT}}$ (4) สัมประสิทธิ์สหสัมพันธ์ระหว่าง r_{CTT} และ r_{IRT1} มีค่า .570 และมีนัยสำคัญทางสถิติที่ระดับ .05 สมการถดถอยของ r_{IRT1} บน r_{CTT} ในรูปคะแนนดิบคือ $r_{IRT1} = .003 + .079r_{CTT}$ และสมการในรูปคะแนนมาตรฐานคือ $Z_{r_{IRT1}} = .570Z_{r_{CTT}}$

170047

4684218027: MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION

KEYWORD: EFFECT MAGNITUDE / EFFECT SIZE / CLASSICAL TEST THEORY /

ITEM RESPONSE THEORY / SIMULATION

CHAYUT PIROMSOMBAT: EFFECT MAGNITUDE ESTIMATORS' PROPERTIES: A
COMPARISON BETWEEN CLASSICAL TEST THEORY AND ITEM RESPONSE THEORY.THESIS ADVISOR: ASST. PROF. SUCHADA BOWARNKITIWONG, Ph.D., 129 pp. ISBN 974-
53-2253-9

There were 4 objectives for this research: (1) to compare the estimators' properties in biasedness, consistency, and relative efficiency aspects among the effect size derived from Classical Test Theory (d_{CTT}), effect size derived from Item Response Theory which estimation models fit for the data (d_{IRT1}), effect size derived from Item Response Theory which estimation models unfit for the data (d_{IRT2}), correlation coefficient derived from Classical Test Theory (r_{CTT}), correlation coefficient derived from Item Response Theory which estimation models fit for the data (r_{IRT1}), and correlation coefficient derived from Item Response Theory which estimation models unfit for the data (r_{IRT2}); (2) to compare the means among d_{CTT} , d_{IRT1} , and d_{IRT2} as well as among r_{CTT} , r_{IRT1} , and r_{IRT2} ; (3) to study the relationship between d_{CTT} and d_{IRT1} and express the regression equation of d_{IRT1} on d_{CTT} ; (4) to study the relationship between r_{CTT} and r_{IRT1} and express the regression equation of r_{IRT1} on r_{CTT} . The 540 examination situations were built up from the conditions of the true effect magnitudes (.2, .5, .8, 1.2, 2.6), sample sizes (20, 50, 500, 2,000), test lengths (10, 50, 90), based models (one-, two-, and three-parameter logistic model), and estimation models (classical test model, item response models which fit for the data, item response models which unfit for the data).

The summarized findings were: (1) in the overview, the lowest biased estimator was r_{IRT1} , the highest consistency estimator was r_{CTT} , and the highest relative efficiency estimator was r_{IRT1} , in addition, r_{IRT1} was the most appropriate estimator for all properties; (2) the means of d_{CTT} , d_{IRT1} , and d_{IRT2} were different at the .05 significance level, in fact, the mean of d_{CTT} was the highest, in the same way, the means of r_{CTT} , r_{IRT1} , and r_{IRT2} were different at the .05 significance level and the mean of r_{CTT} was the highest; (3) the correlation coefficient between d_{CTT} and d_{IRT1} was .626 and significant at the .05 level, the regression of d_{IRT1} on d_{CTT} could be expressed by $d_{IRT1} = .004 + .065d_{CTT}$ for the raw scores and $Z_{d_{IRT1}} = .626Z_{d_{CTT}}$ for the standardized scores; (4) the correlation coefficient between r_{CTT} and r_{IRT1} was .570 and significant at the .05 level, the regression of r_{IRT1} on r_{CTT} could be expressed by $r_{IRT1} = .003 + .079r_{CTT}$ for the raw scores and $Z_{r_{IRT1}} = .570Z_{r_{CTT}}$ for the standardized scores.