

การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบด้วยวิธีของราจูและ
วิธีของเลียว ที่มีความยากและจำนวนข้อแตกต่างกัน

**A Comparison of Raju's and Liou's Test Reliability with
Different Difficulty and Number of Items**

สุรศักดิ์ เมตตาวิมล

Surasak Mettavimon

วิทยานิพนธ์ศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวัดผลและวิจัยการศึกษา
มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี

Master of Education Thesis in Educational Measurement and Research

Prince of Songkla University

2542

ชื่อวิทยานิพนธ์ การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบด้วยวิธีของราฐ และวิธีของเลียว ที่มีความยากและจำนวนข้อแตกต่างกัน
ผู้เขียน นายสุรศักดิ์ เมตตาวิมล
สาขาวิชา การวัดผลและวิจัยการศึกษา
ปีการศึกษา 2541

บทคัดย่อ

การวิจัยครั้งนี้เพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบปรนัยชนิดเลือกตอบ วิชาคณิตศาสตร์ที่มีช่วงความยากต่างกัน คือ .20 - .80 , .21 - .40 , .41 - .60 และ .61 - .80 ด้วยวิธีของราฐ เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีจำนวนข้อต่างกัน คือ 40 ข้อ , 30 ข้อ และ 20 ข้อ เมื่อคำนวณค่าความเชื่อมั่นด้วยวิธีของราฐ และวิธีของเลียว ตามลำดับ และเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่คำนวณด้วยสูตร F_{BK} ของราฐ และ สูตร F_{L2} ของเลียวเมื่อแบบทดสอบมีช่วงความยากและจำนวนข้อเท่ากัน โดยทดสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2541 สังกัดกรมสามัญศึกษา จังหวัดปัตตานี จำนวน 960 คน ซึ่งเลือกมาโดยวิธีการสุ่มแบบแบ่งชั้น

ผลการวิจัยพบว่า เมื่อคำนวณค่าความเชื่อมั่นด้วยสูตร F_{BK} ของราฐ แบบทดสอบที่มีจำนวนข้อเท่ากันแต่มีช่วงความยากต่างกันคือ .20 - .80 , .21 - .40 , .41 - .60 และ .61 - .80 มีค่าความเชื่อมั่นไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แต่มีแนวโน้มว่าแบบทดสอบที่มีช่วงความยาก .41 - .60 มีค่าความเชื่อมั่นสูงสุด ขณะที่แบบทดสอบที่มีช่วงความยาก .21 - .40 มีค่าความเชื่อมั่นต่ำสุด เมื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีช่วงความยากเท่ากัน แต่มีจำนวนข้อต่างกันคือ 40 ข้อ 30 ข้อ และ 20 ข้อ พบว่ามีค่าความเชื่อมั่นไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในกรณีที่คำนวณค่าความเชื่อมั่นด้วยสูตร F_{L2} ของเลียว พบว่า แบบทดสอบที่มีจำนวนข้อเท่ากัน แต่มีช่วงความยากต่างกันคือ .20 - .80 , .21 - .40 , .41 - .60 และ .61 - .80 มีค่าความเชื่อมั่นไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีช่วงความยากเท่ากัน

แต่มีจำนวนข้อต่างกันคือ 40 ข้อ 30 ข้อ และ 20 ข้อ มีค่าความเชื่อมั่นไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่คำนวณโดยใช้สูตร r_{BK} ของราชู กับ สูตร r_{L2} ของเลียว เมื่อแบบทดสอบมีช่วงความยากและจำนวนข้อเท่ากัน พบว่ามีค่าความเชื่อมั่นไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

Abstract

This research was intended to compare the reliability estimated by Raju's formula and Lien's formula for multiple-choice mathematics tests which had different difficulty levels (20 - 30, 31 - 40, 41 - 50, and 51 - 60), and those of the tests of which the number of items was either 40, 30, or 20 using Raju's and Lien's methods respectively. The aim was to compare the reliability estimated by Raju's r_{BK} formula and Lien's r_{L2} formula of the tests with the same difficulty levels and number of items. The tests were administered to 1000 non-Sukse 1 students in schools under the Department of General Education, Chachoengsao Patana, during the academic year 1998. They were selected by a stratified random sampling for the purpose of the study.

The findings were that, with Raju's r_{BK} formula, the reliability estimates of the tests with different difficulty ranges (20 - 30, 31 - 40, 41 - 50, and 51 - 60) were not significantly different at .05. There was a tendency that the reliability of the tests of which the difficulty was 41 - 50 was the highest and that of the 21 - 40 range was the lowest. When the reliability estimates of the tests with equal difficulty but different number of items were compared, they were not significantly different at .05. When Lien's r_{L2} formula was used, the reliability estimates of the tests with equal number of items but different difficulty levels, either 20 - 30, 31 - 40, 41 - 50, or 51 - 60, were not significantly different at .05. Comparing the tests with equal

Thesis Title A Comparison of Raju's and Liou's Tests Reliability with Different Difficulty and Number of Items
Author Mr. Surasak Mettavimon
Major Program Educational Measurement and Research
Academic Year 1998

Abstract

This research was intended to compare the reliability estimated by Raju's method of multiple-choice mathematics tests which had different difficulty levels : .20 - .80, .21 - .40, .41 - .60, and .61 - .80, and those of the tests of which the number of items was either 40, 30, or 20 using Raju's and Liou's methods consecutively, then to compare the reliability estimated by Raju's $r_{\beta\kappa}$ formula and Liou's r_{L2} formula of the tests with the same difficulty levels and number of items. The tests were administered to Matayom Suksa 1 students in schools under the Department of General Education, Changwat Pattani, during the academic year 1998. They were selected by a stratified random sampling for 960 samples for the study.

The findings were that, with Raju's $r_{\beta\kappa}$ formula, the reliability estimates of the tests with different difficulty ranges: .20 - .80, .21 - .40, .41 - .60, and .61 - .80, were not significantly different at .05. There was a tendency that the reliability of the tests of which the difficulty was .41 - .60 was the highest whereas that of the .21 - .40 range was the lowest. When the reliability estimates of the tests with equal difficulty but different number of items were compared, they were not significantly different at .05. When Liou's r_{L2} formula was used, the reliability estimates of the tests with equal number of items but different difficulty levels, either .20 - .80, .21 - .40, .41 - .60, or .61 - .80, were not significantly different at .05. Comparing the tests with equal

difficulty but different number of items, either 40, 30, or 20, the reliability estimates were not significantly different at .05. When the reliability levels calculated by Raju's $r_{\beta k}$ formula and Liou's r_{L2} formula of the tests with the same ranges of difficulty and number of items were compared, it was found that they were not different at the level of significance of .05.