

อำนาจการทดสอบของการใช้สถิติการทดสอบความแตกต่างของค่าเฉลี่ย

ปริญญานิพนธ์

ของ

นิภาพร ขำสอาด

เสนอต่อบัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา

หลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา

พฤษภาคม 2552

อำนาจการทดสอบของการใช้สถิติการทดสอบความแตกต่างของค่าเฉลี่ย

บทคัดย่อ

ของ

นิภาพร ขำสอาด

เสนอต่อบัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา

หลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา

พฤษภาคม 2552

นิภาพร ขำสอาด. (2552). อำนาจการทดสอบของการใช้สถิติการทดสอบความแตกต่างของค่าเฉลี่ย.

ปริญญาานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย

มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม: อาจารย์ชวลิต รวยอาจิณ,

รองศาสตราจารย์นิภา ศรีไพโรจน์.

การศึกษาค้นคว้าครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD จากกลุ่มตัวอย่าง 4 ขนาด คือ ขนาด 30 คน 60 คน 90 คน และ 120 คน

กลุ่มตัวอย่างที่ทำการศึกษาค้นคว้าเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ,2 และ 3 ของนักเรียนในสังกัดกรุงเทพมหานคร จำนวน 564 คน ที่ได้จากการสุ่มแบบหลายขั้นตอน เพื่อกำหนดเป็นประชากรเทียม เครื่องมือที่ใช้ในการศึกษาค้นคว้าเป็น แบบทดสอบวัดความถนัดด้านภาษา ประกอบด้วย ด้านศัพท์สัมพันธ์ ด้านจัดประเภท ด้านความเข้าใจภาษาและด้านความเข้าใจภาพ มีค่าความเชื่อมั่น 0.86 ,0.82 , 0.85 และ 0.91 ตามลำดับ

ผลการศึกษาค้นคว้า

1) เมื่อกลุ่มตัวอย่างเท่ากัน อำนาจการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ย ทั้ง 5 วิธี วิธี Fisher's LSDมีค่าอำนาจการทดสอบสูงสุดรองลงมาคือ วิธี Duncan ,วิธี Scheffe ' ,วิธี Newman Keuls และให้ค่าอำนาจการทดสอบน้อยที่สุด คือ วิธี Tukey-Kramer

2) อำนาจการทดสอบระหว่างกลุ่มตัวอย่าง 4 ขนาดพบว่า เมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นส่งผลให้ค่าอำนาจการทดสอบสูงขึ้น ซึ่งในการศึกษาค้นคว้าครั้งนี้ ขนาดกลุ่มตัวอย่างที่มีค่าอำนาจการทดสอบมากที่สุดคือ กลุ่มตัวอย่างขนาด 120 คนและ 90 คน รองลงมาคือ กลุ่มตัวอย่างขนาด 60 คน และมีค่าอำนาจการทดสอบน้อยที่สุด คือ กลุ่มตัวอย่างขนาด 30 คน

POWER OF TEST ON MULTIPLE COMPARISON

AN ABSTRACT

BY

NIPAPORN KHUMSAARD

Presented in Partial Fulfillment of the Requirements for the
Master of Education Degree in Educational Measurement
at Srinakharinwirot University

May 2009

Nipaporn Khumsaard. (2009). *Power of Test on Multiple Comparison*. Master Thesis. M.Ed. (Educational Measurement). Bangkok: Graduate School, Srinakharinwirot University. Advisor Committee: Mr. Chawalit Ruayajin, Assoc.Prof. Nipa Sripairot.

The purpose of this study to compare the power of test on multiple comparison procedures among five different methods : Tukey-Kramer test, Newman Keuls test, Duncan test, Scheffe ' test and Fisher's LSD test by using 4 sample sizes of 30, 60, 90, and 120 students to saw if there were any difference among the tests and which method yields to the most suitable result and at which sample sizes.

The samples consisted of 564 Mathayomsuksa I, II and III students of Bangkok Metropolitan administration by using the Multi – Stage Random Sampling. The instrument for the study included the verbal aptitude test with 40 items , 4 patterns and reliability were 0.86, 0.82, 0.85 and 0.91, respectively.

The results revealed that :

1) When equal group the power of test on multiple comparison procedures among five different methods presented ; the highest that of Fisher's LSD test and next that of Duncan test, Scheffe ' test , Newman Keuls test respectively. The lowest one included that of Tukey-Kramer test.

2) The power of test between sample size showed ; by multiple comparison procedures among five different methods when the sample size was increased, power of test was increased. This study was found that most power of test included that of sample sizes of 120 students, 90 students and 60 students respectively. The lowest one included that of sample sizes of 30 students.

อำนาจการทดสอบของการใช้สถิติการทดสอบความแตกต่างของค่าเฉลี่ย

ปริญญานิพนธ์
ของ
นิภาพร ขำสอาด

เสนอต่อบัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา

หลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา

พฤษภาคม 2552

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

ปริญญานิพนธ์
เรื่อง

อำนาจการทดสอบของการใช้สถิติการทดสอบความแตกต่างของค่าเฉลี่ย

ของ
นิภาพร ขำสอาด

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย

(รองศาสตราจารย์ ดร.สมชาย สันติวัฒนกุล)

วันที่ เดือนพฤษภาคม พ.ศ. 2552

คณะกรรมการควบคุมปริญญานิพนธ์

คณะกรรมการสอบปากเปล่า

.....ประธาน

.....ประธาน

(อาจารย์ชวลิต รวยอาจิณ)

(อาจารย์ ดร.สุวพร เข้มแข็ง)

.....กรรมการ

.....กรรมการ

(รองศาสตราจารย์นิภา ศรีไพโรจน์)

(อาจารย์ชวลิต รวยอาจิณ)

.....กรรมการ

(รองศาสตราจารย์นิภา ศรีไพโรจน์)

.....กรรมการ

(ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลุล่วงได้ด้วยความช่วยเหลือและความกรุณาอย่างสูงจาก อาจารย์ชวลิต รวยอาจิณ ประธานควบคุมปริญญานิพนธ์ และรองศาสตราจารย์นิภา ศรีไพโรจน์ ที่ได้กรุณาให้คำปรึกษา แนะนำ ชี้แนวทาง ตรวจ แก้ไขข้อบกพร่อง จนงานวิจัยนี้สำเร็จลุล่วงไปได้ ด้วยดี ผู้วิจัยขอกราบขอบพระคุณอย่างสูง

ขอกราบขอบพระคุณคุณอาจารย์ดร.สุพร เข้มแข็ง และผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช รวมทั้งอาจารย์ในภาควิชาการวัดผลและวิจัยการศึกษาทุกท่าน ที่ได้แนะแนวทางในการทำปริญญานิพนธ์ฉบับนี้

ขอกราบขอบพระคุณผู้อำนวยการ คณะครู อาจารย์ทุกท่านของโรงเรียนที่เป็นกลุ่มตัวอย่าง รวมทั้งนักเรียนทุกคนที่ได้ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี และขอขอบพระคุณ อาจารย์มิ่ง เทพครเมือง ที่ให้คำแนะนำ ให้คำปรึกษาตลอดมา ตลอดจน นายวุฒิไกร ขำสอาด เพื่อน พี่ น้อง สาขาวิชาการวัดผลการศึกษาและเพื่อนร่วมงานทุกคน ที่ให้คำแนะนำและเป็นกำลังใจ สำหรับผู้วิจัย

คุณค่าและประโยชน์อันพึงมีจากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาพระคุณ บิดา มารดา และครูอาจารย์ ตลอดจนผู้มีพระคุณที่ให้การอบรมสั่งสอนตลอดมา

นิภาพร ขำสอาด

สารบัญ

บทที่	หน้า
1 บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายของการศึกษา	4
ความสำคัญของการศึกษา	4
ขอบเขตของการศึกษา	4
ประชากรและกลุ่มตัวอย่าง	4
ตัวแปรที่ศึกษา	5
นิยามศัพท์เฉพาะ	5
สมมติฐานการศึกษา	6
กรอบแนวคิดในการวิจัย	7
2 เอกสารงานวิจัยที่เกี่ยวข้อง	8
ความหมายของการทดสอบความแตกต่างของค่าเฉลี่ย	8
การเปรียบเทียบระหว่างค่าเฉลี่ย	9
ประเภทของสถิติ	10
การทดสอบความเท่ากันของความแปรปรวน	15
วิธีของ Bartlett	15
วิธีของ Levene	15
อำนาจการทดสอบ	16
การทดสอบสมมติฐาน	16
ความหมายของอำนาจการทดสอบ	17
อิทธิพลที่ส่งผลต่ออำนาจการทดสอบ	18
ขนาดอิทธิพล	21
แนวทางการปรับปรุงอำนาจการทดสอบ	23
การคำนวณเกี่ยวกับอำนาจการทดสอบ	23
แนวคิดหรือกรอบสำหรับการวิเคราะห์ค่าอำนาจการทดสอบ	25

สารบัญ(ต่อ)

บทที่	หน้า
2 (ต่อ)	
สถิติทดสอบความแตกต่างของค่าเฉลี่ย.....	26
วิธีการทดสอบของ Tukey-Kramer	26
วิธีการทดสอบของ Newman Keuls	27
วิธีการทดสอบของ Duncan	28
วิธีการทดสอบของ Scheffe '	29
วิธีการทดสอบของ Fisher's LSD	30
ความถนัดและแบบทดสอบวัดความถนัด	34
งานวิจัยที่เกี่ยวข้อง	35
งานวิจัยต่างประเทศ	35
งานวิจัยภายในประเทศ	37
3 วิธีดำเนินการศึกษาค้นคว้า	44
ประชากรและกลุ่มตัวอย่างที่ใช้ในการศึกษา	44
เครื่องมือที่ใช้ในการศึกษา	45
วิธีการเก็บรวบรวมข้อมูล	48
การวิเคราะห์ข้อมูล	48
สถิติที่ใช้ในการศึกษา	49
4 ผลการวิเคราะห์ข้อมูล	55
การนำเสนอผลการวิเคราะห์ข้อมูล	55
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล	55
ผลการวิเคราะห์ข้อมูล	56

สารบัญ(ต่อ)

บทที่	หน้า
5 สรุป อภิปรายผล และข้อเสนอแนะ	75
สังเขปความมุ่งหมาย สมมติฐาน และวิธีดำเนินการศึกษาค้นคว้า	75
สรุปผลการศึกษาค้นคว้า	76
อภิปรายผล	77
ข้อเสนอแนะ	79
 บรรณานุกรม	 81
 ภาคผนวก	 88
 ประวัติย่อผู้วิจัย	 111

บัญชีตาราง

ตาราง		หน้า
1	แสดงการเปรียบเทียบขนาดของกลุ่มตัวอย่างและค่าวิกฤตของสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี	31
2	แสดงการเปรียบเทียบวิธีการเปรียบเทียบและขั้นตอนการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี	32
3	แสดงจำนวนโรงเรียนและห้องเรียนที่ทำการสุ่มได้ของชั้นมัธยมศึกษาปีที่ 1, 2, และ 3	45
4	ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างในการศึกษา	56
5	ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างขนาด 30 คน ชั้นละ 10 กลุ่ม	57
6	ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างขนาด 60 คน ชั้นละ 10 กลุ่ม	58
7	ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างขนาด 90 คน ชั้นละ 10 กลุ่ม	59
8	ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างขนาด 120 คน ชั้นละ 10 กลุ่ม	60
9	การเปรียบเทียบจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญจากการทดสอบความแตกต่างของค่าเฉลี่ยด้วยสถิติทดสอบ 5 วิธีจากกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน	62
10	การเปรียบเทียบจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญด้วยวิธี Tukey-Kramer, วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และ วิธี Fisher's LSD จากกลุ่มตัวอย่าง 4 ขนาด	63
11	การเปรียบเทียบค่าอำนาจการทดสอบของสถิติความแตกต่างของค่าเฉลี่ยที่ทำการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ย 5 วิธี จากกลุ่มตัวอย่าง 30 คน	64

บัญชีตาราง(ต่อ)

ตาราง	หน้า
12 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ย 5 วิธี จาก กลุ่มตัวอย่าง 60 คน	65
13 ผลการเปรียบเทียบค่าอำนาจการทดสอบของสถิติความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ย 5 วิธี จาก กลุ่มตัวอย่าง 90 คน	66
14 ผลการเปรียบเทียบค่าอำนาจการทดสอบของสถิติความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ย 5 วิธี จาก กลุ่มตัวอย่าง 120 คน	67
15 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของ ค่าเฉลี่ยวิธี Tukey-Kramer ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน	69
16 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของ ค่าเฉลี่ยวิธี Newman Keuls ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน	70
17 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของ ค่าเฉลี่ยวิธี Duncan ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน	71
18 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของ ค่าเฉลี่ยวิธี Scheffe ' ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน	72
19 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของ ค่าเฉลี่ยวิธี Fisher's LSD ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน	73
20 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วย วิธี Tukey-Kramer จากกลุ่มตัวอย่าง 4 ขนาด	90

บัญชีตาราง(ต่อ)

ตาราง	หน้า
21 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับขั้นที่คำนวณด้วย วิธี Newman Keuls จากกลุ่มตัวอย่าง 4 ขนาด	91
22 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับขั้นที่คำนวณด้วย วิธี Duncan จากกลุ่มตัวอย่าง 4 ขนาด	92
23 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับขั้นที่คำนวณด้วย วิธี Scheffe ' จากกลุ่มตัวอย่าง 4 ขนาด	93
24 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับขั้นที่คำนวณด้วย วิธี Fisher's LSD จากกลุ่มตัวอย่าง 4 ขนาด	94
25 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วย วิธี Tukey-Kramer จากกลุ่มตัวอย่าง 4 ขนาด	95
26 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วย วิธี Newman Keuls จากกลุ่มตัวอย่าง 4 ขนาด	96
27 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วย วิธี Duncan จากกลุ่มตัวอย่าง 4 ขนาด	97
28 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วย วิธี Scheffe ' จากกลุ่มตัวอย่าง 4 ขนาด	98
29 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วย วิธี Fisher's LSD จากกลุ่มตัวอย่าง 4 ขนาด	99

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 กรอบแนวคิดการวิจัย	7
2 ผลลัพธ์ที่เกิดจากการทดสอบสมมติฐานทางสถิติ	16
3 อิทธิพลของความแตกต่างระหว่าง μ_0 และ μ_x ที่ส่งผลต่ออำนาจการทดสอบ	18
4 อิทธิพลจากขนาดกลุ่มตัวอย่าง และความแปรปรวนที่ส่งผลต่ออำนาจการทดสอบ	20
5 อิทธิพลจากระดับนัยสำคัญทางสถิติที่ส่งผลต่ออำนาจการทดสอบ	21

บทที่ 1

บทนำ

ภูมิหลัง

ในงานวิจัยด้านต่าง ๆ เมื่อมีการทดสอบสมมติฐาน ผู้วิจัยจำเป็นที่จะต้องมีความรู้ในเรื่องที่จะศึกษาเป็นอย่างดี เมื่อได้ออกแบบการวิจัยและกำหนดแบบแผนการทดลอง หรือแม้กระทั่งกลุ่มตัวอย่างที่เหมาะสมแล้วนั้น ผู้วิจัยจะต้องเลือกสถิติที่เหมาะสมกับข้อมูลและขนาดของกลุ่มตัวอย่างด้วย ซึ่งแต่ละสถิติต่างมีข้อตกลงเบื้องต้น (Basic Assumptions) เกี่ยวกับข้อมูลที่จะนำมาวิเคราะห์แตกต่างกัน ดังนั้นการเลือกใช้สถิติที่นำมาทดสอบต้องมีความสอดคล้องกับลักษณะของข้อมูลที่กำลังศึกษา เพื่อให้ผลการสรุปของการวิจัยมีความถูกต้องและเชื่อถือได้มากขึ้น การทดสอบความแตกต่างของค่าเฉลี่ยเป็นวิธีการหนึ่งที่จะตอบปัญหาเกี่ยวข้องกับค่าพารามิเตอร์ของประชากร

ในการทำการวิจัยทุกครั้งมักจะมีการทดสอบความมีนัยสำคัญในค่าเฉลี่ยของประชากร หรือหาข้อสรุปว่า การจัดกระทำที่ให้ในแต่ละกลุ่มตัวอย่าง มีการจัดกระทำใดบ้างที่แตกต่างกัน ในการทดสอบสมมติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ยของประชากรที่มีมากกว่า 2 กลุ่มตัวอย่างขึ้นไปโดยทั่วไปแล้วจะการใช้การวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA) หากการทดสอบนั้นปฏิเสธสมมติฐานหลัก ($H_0 = \mu_1 = \mu_2 = \dots = \mu_k$) แสดงว่าค่าเฉลี่ยอย่างน้อยหนึ่งกลุ่มไม่เท่ากับกลุ่มอื่น ซึ่งในขั้นต่อไปผู้วิจัยจะต้องทำการเปรียบเทียบพหุคูณหรือการเปรียบเทียบภายหลัง (Posterior comparison) เป็นการทดสอบความต่างของค่าเฉลี่ยภายหลังการวิเคราะห์ความแปรปรวนเมื่อพบความแตกต่าง ซึ่งเป็นการเปรียบเทียบค่าเฉลี่ยแต่ละคู่ว่าค่าใดที่มีค่าต่างกัน โดยพิจารณาผลต่างของค่าเฉลี่ยแต่ละคู่กับค่าวิกฤตที่คำนวณได้ สถิติที่ใช้ในการเปรียบเทียบพหุคูณมีหลายรูปแบบ พอสรุปได้ 3 รูปแบบการทดสอบดังนี้ การเปรียบเทียบพหุคูณโดยการทดสอบแบบที (Multiple comparisons using the T-test) การเปรียบเทียบพหุคูณโดยใช้ค่าพิสัยสตีวเดนท์ (Multiple comparisons using the studentized range) การเปรียบเทียบพหุคูณโดยใช้การทดสอบแบบเอฟ (Multiple comparisons using the F – test) ยังมีนักสถิติหลายท่านได้พัฒนาและคิดค้นหาวิธีการเปรียบเทียบพหุคูณไว้มากหลายวิธี ซึ่ง เคิร์ก ได้รวบรวมไว้ในหนังสือที่ชื่อว่า Experimental design : porcedures for the behavioral science ได้ 22 วิธี (Kirk. 1995) และแบ่งได้เป็น 2 กลุ่มใหญ่ คือ กลุ่มแรกเป็นวิธีการเปรียบเทียบพหุคูณที่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติที่ว่า ความแปรปรวนต้องไม่แตกต่างกัน (Homogeneous Variances) ($\sigma^2_1 = \sigma^2_2 = \dots = \sigma^2_k$) วิธีการเปรียบเทียบพหุคูณในกลุ่มนี้มีหลายวิธี แบ่งได้เป็น การเปรียบเทียบแบบวางแผนล่วงหน้า (Planned contrast) เช่น Dunnett's test , Holm's test และการเปรียบเทียบ

ภายหลัง (Post hoc contrast) เช่น Tukey's test , Fisher-Hayter test, Scheffe' test, Duncan test , Newman Keuls's test กลุ่มที่สอง คือกลุ่มวิธีการเปรียบเทียบพหุคูณที่ไม่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติ คือความแปรปรวนต่างกัน (Heterogeneous Variances) วิธีการเปรียบเทียบพหุคูณในกลุ่มนี้มีหลายวิธี แบ่งได้เป็น การเปรียบเทียบแบบวางแผน เช่น Dunnett's test (Modified), Holm's test with Welch และการเปรียบเทียบภายหลัง เช่น Dunnett's T3 test, Dunnett's C test, Games – Howell test และ Brown – Forsythe test (Kirk. 1995)

วิธีการเปรียบเทียบภายหลัง (Post hoc contrast) มีอำนาจในการทดสอบสูงเนื่องจากการเปรียบเทียบพหุคูณแบบวางแผนล่วงหน้า มักมีข้อจำกัดในจำนวนการเปรียบเทียบ กล่าวคือ ถ้าจำนวนการเปรียบเทียบมีมาก และการเปรียบเทียบไม่เป็นอิสระต่อกัน จะทำให้ช่วงความเชื่อมั่นกว้างและอำนาจการทดสอบต่ำ (Hay. 1963 : 488) ซึ่งทำให้ยากต่อการตรวจสอบพบผลของการจัดกระทำของการทดลอง นักวิจัยควรออกแบบงานวิจัยให้มีอำนาจการทดสอบที่เหมาะสมและสอดคล้องกับสภาพการณ์ของการทำวิจัย แทนที่นักวิจัยจะสนใจแต่เพียงการออกแบบงานวิจัยเพื่อให้มีนัยสำคัญทางสถิติ

อำนาจการทดสอบคือ ความน่าจะเป็นในการปฏิเสธสมมติฐานกลางเมื่อสมมติฐานกลางเป็นเท็จ เขียนแทนด้วยสัญลักษณ์ $(1 - \beta)$ ซึ่งอำนาจการทดสอบมีความเกี่ยวข้องกับความคลาดเคลื่อนประเภทที่ 1 (Type II error) หรือระดับนัยสำคัญทางสถิติ เขียนแทนด้วยสัญลักษณ์ (α) คือความน่าจะเป็นในการปฏิเสธสมมติฐานกลางเมื่อสมมติฐานกลางเป็นจริง และความคลาดเคลื่อนประเภทที่ 2 (Type II error) เขียนแทนด้วยสัญลักษณ์ (β) คือความน่าจะเป็นในการยอมรับสมมติฐานกลางเมื่อสมมติฐานกลางเป็นเท็จ อำนาจการทดสอบใช้ประโยชน์สำหรับการวางแผนการวิจัย เช่น เป็นเกณฑ์ในการตัดสินใจเลือกขนาดกลุ่มตัวอย่าง และกำหนดนัยสำคัญทางสถิติโดยอาศัยการประมาณค่าอำนาจการทดสอบ การประมาณค่าอำนาจการทดสอบ (Power estimate) มี 2 ชนิด คือการประมาณค่าก่อนทำการทดลอง และการประมาณค่าหลังจากทำการทดลองแล้ว การประมาณค่าการทดสอบก่อนทำการทดลอง สามารถกระทำได้โดยการศึกษาค้นคว้าและงานวิจัยที่เกี่ยวข้อง ซึ่งชี้ให้เห็นขนาดของอิทธิพล (Effect size) หมายถึง ดัชนีที่ใช้วัดอิทธิพลของการจัดกระทำทางการทดลองซึ่งส่งผลต่อตัวแปรตามและนำไปเป็นแนวทางในการตัดสินใจเลือกขนาดของกลุ่มตัวอย่างที่มีพื้นฐานมาจากการประมาณค่าขนาดอิทธิพล การประมาณค่าอำนาจการทดสอบหลังจากทำการทดลองสามารถกระทำได้โดยการใช้โปรแกรมสำเร็จรูป ซึ่งมีด้วยกันหลายโปรแกรมดังนี้ SYSTAT , SAS , SPSS , GPOWER และ UCLA (Myer & Well , 2003 : 212 – 215) โดยสรุปอำนาจการทดสอบขึ้นอยู่กับองค์ประกอบ 3 ประการ คือ ระดับนัยสำคัญทางสถิติ (α - Level) ขนาดกลุ่มตัวอย่างและขนาดกลุ่มอิทธิพล

การออกแบบการวิจัยโดยทั่วไปมีจุดประสงค์เพื่อตอบคำถามการวิจัย และควบคุมความแปรปรวน (Kerlinger. 1986) การออกแบบการวิจัยครอบคลุมไปถึงการออกแบบการวัด การเลือกรูปแบบหรือวิธีการวัดผล การเลือกใช้สถิติให้เหมาะสมกับข้อตกลงเบื้องต้นและเป้าหมายของการวิจัยนั้นๆ การเลือกกลุ่มตัวอย่างหรือขนาดของกลุ่มตัวอย่าง และลักษณะของตัวแปรที่ต้องการศึกษา (ศิริชัย กาญจนวาสีและคณะ. 2537) จากที่กล่าวมาข้างต้น การออกแบบการสุ่มตัวอย่างเป็นการออกแบบที่มีความสำคัญต่อการวิจัย ในกรณีที่นักวิจัยไม่สามารถศึกษาข้อเท็จจริงจากประชากรทั้งหมดได้ จึงต้องใช้วิธีการสุ่มกลุ่มตัวอย่างจากกลุ่มของประชากรที่ต้องการศึกษา และใช้หลักฐานข้อเท็จจริงจากกลุ่มตัวอย่างนั้นสรุปผลเชิงสถิติจากกลุ่มประชากร เพื่อตอบคำถามการวิจัย

คุณลักษณะที่เป็นเป้าหมายกลุ่มสำคัญของการวิจัยที่ต้องการศึกษาจากกลุ่มตัวอย่าง คือ มุ่งให้การวิจัยมีความตรงโดยเฉพาะอย่างยิ่งความเที่ยงตรงภายนอก (External Validity) หรือความสามารถในการสรุปผลอ้างอิงผลการวิจัยไปสู่ประชากรทั้งหมดที่ศึกษา (Generalizability) การกำหนดกลุ่มตัวอย่างนั้นเป็นขั้นตอนหนึ่งในการออกแบบกลุ่มตัวอย่างที่นักวิจัยต้องดำเนินการก่อนเริ่มทำการวิจัย

กิลฟอร์ด และ ฟรชเตอร์ (Guilford and Fruchter. 1978) กล่าวว่า คำถามที่เกิดขึ้นเสมอเมื่อเริ่มวางแผนการวิจัย คือ กลุ่มตัวอย่างที่ใช้ในการวิจัยควรมีขนาดเท่าใด ฮิงเคิล และ โอลิเวอร์ (Hinkle and Oliver. 1983) กล่าวว่า การตอบคำถามเกี่ยวกับขนาดของกลุ่มตัวอย่างมีความยุ่งยากเนื่องจากความสัมพันธ์ระหว่างขนาดของกลุ่มตัวอย่างกับความมีนัยสำคัญทางสถิติของสถิติอนุมานที่ผู้วิจัยใช้ จุดวิกฤตของสถิติอนุมาน คืองานวิจัยที่ใช้สถิติอนุมาน ให้ความสำคัญสูงกับความมีนัยสำคัญกับผลการวิเคราะห์ เฮ (Hay. 1963) กล่าวว่างานวิจัยบางเรื่องสามารถแสดงผลที่มีนัยสำคัญได้ ถ้าใช้กลุ่มตัวอย่างขนาดใหญ่โดยที่ไม่มีการพิจารณาถึงความสมเหตุสมผลของสิ่งที่ต้องการศึกษา จุดสำคัญของวิธีการทางสถิติอนุมาน คือการแปลความหมายได้อย่างมั่นใจว่าความแตกต่างระหว่างพารามิเตอร์กับค่าทางสถิตินั้นเป็นความแตกต่างที่แท้จริง ไม่ใช่เป็นความแตกต่างที่เกิดขึ้นโดยบังเอิญอันเนื่องมาจากการสุ่ม การกำหนดขนาดของกลุ่มตัวอย่างมักถูกจำกัดไม่ให้มีขนาดใหญ่เกินไป ด้วยเหตุผลอันเนื่องมาจากค่าใช้จ่ายที่ต้องเสีย ระยะเวลาในการเก็บรวบรวมข้อมูลและความเป็นไปได้ในการเก็บข้อมูลให้ได้คุณภาพตรงกับข้อตกลงของการวิจัยหรือเป้าหมายของการวิจัยนั้น

ในการวิจัยเชิงสำรวจนักวิจัยที่มีเงินทุนและช่วงเวลาในการทำวิจัยที่จำกัดจำเป็นต้องเลือกขนาดของกลุ่มตัวอย่างที่มีขนาดเล็ก ส่วนในการวิจัยเชิงทดลอง ด้วยเหตุผลเดียวกันจำเป็นต้องใช้กลุ่มตัวอย่างที่มีขนาดเล็ก มีความเหมาะสมและเพียงพอ เพื่อให้ผลการวิจัยมีความเชื่อถือได้ ดังนั้นผู้วิจัยต้องควบคุมความคลาดเคลื่อนในการวัดที่จะเกิดขึ้นในระหว่างการทดลอง

ด้วยเหตุผลดังกล่าวข้างต้น ทำให้ผู้วิจัยตระหนักว่าการออกแบบการวิจัยที่ดีต้องคำนึงถึงปัจจัยที่เกี่ยวข้องหลายปัจจัย เช่น ระดับนัยสำคัญทางสถิติ ขนาดของกลุ่มตัวอย่าง สถิติที่ใช้ในการ

ทดสอบ เป็นต้น ผู้วิจัยจึงสนใจที่จะศึกษา อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนทางเดียวของสถิติการทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี คือ Tukey - Kramer วิธี Newman Keuls วิธี Duncan ซึ่งอยู่ในกลุ่ม t - statistic วิธี Scheffe ' ซึ่งอยู่ในกลุ่ม s - statistic และวิธี Fisher' s LSD ซึ่งอยู่ในกลุ่ม f - statistic ภายใต้ข้อตกลงที่เหมือนกันคือมีความแปรปรวนที่เท่ากัน ขนาดกลุ่มตัวอย่างต่างกัน ว่าให้ผลการทดสอบต่างกันหรือไม่

ความมุ่งหมายของการศึกษา

การศึกษาค้นคว้าครั้งนี้มีความมุ่งหมายสำคัญเพื่อศึกษาอำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว โดยกำหนดจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อศึกษาอำนาจการทดสอบของสถิติการทดสอบความแตกต่างของค่าเฉลี่ยต่างวิธีกัน โดยที่ขนาดของกลุ่มตัวอย่างเท่ากัน
2. เพื่อศึกษาอำนาจการทดสอบเมื่อขนาดของกลุ่มตัวอย่างต่างกัน โดยที่สถิติการทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน

ความสำคัญของการศึกษา

การศึกษาค้นคว้าครั้งนี้เป็นการศึกษาอำนาจการทดสอบของสถิติการทดสอบความแตกต่างของค่าเฉลี่ย ภายใต้ความแปรปรวนที่เท่ากัน และขนาดของกลุ่มตัวอย่างต่างกัน ผลการศึกษาค้นคว้าครั้งนี้เป็นการศึกษาถึงวิธีการทดสอบความแตกต่างของค่าเฉลี่ยวิธีใดใช้กับกลุ่มตัวอย่างเท่าใดจึงจะเหมาะสมเพื่อประโยชน์ในเชิงปฏิบัติ เป็นแนวทางให้ผู้ทำการวิจัยหรือนักคิดที่เกี่ยวข้องเลือกใช้สถิติได้อย่างเหมาะสม อีกทั้งยังเป็นหลักฐานสนับสนุนหรืออ้างอิงในโอกาสต่อไป

ขอบเขตของการศึกษา

ประชากรที่ใช้ในการศึกษา

ประชากรที่ใช้ในการศึกษาค้นคว้านี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 , 2 และ 3 ภาคเรียนที่ 2 ปีการศึกษา 2550 สังกัด กรุงเทพมหานคร จำนวน 62 โรงเรียน จำนวนนักเรียน 16,260 คน

กลุ่มตัวอย่างที่ใช้ในการศึกษา

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้านี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 , 2 และ 3 ภาคเรียนที่ 2 ปีการศึกษา 2550 สังกัดกรุงเทพมหานคร จำนวน 564 คน สุ่มจากประชากรโดยวิธี Multi - Stage Random Sampling

ตัวแปรที่ศึกษา

1. ตัวแปรอิสระ

1.1 วิธีการทดสอบความแตกต่างของค่าเฉลี่ย 5 วิธี คือ

1.1.1 วิธี Tukey - Kramer

1.1.2 วิธี Newman Keuls

1.1.3 วิธี Duncan

1.1.4 วิธี Scheffe '

1.1.5 วิธี Fisher' s LSD

1.2 ขนาดของกลุ่มตัวอย่าง 4 ขนาด คือ

1.2.1 30 คน

1.2.2 60 คน

1.2.3 90 คน

1.2.4 120 คน

2. ตัวแปรตาม ได้แก่ อำนาจการทดสอบของสถิติ

3. สมมติฐานของการทดสอบ

3.1 อำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยต่างวิธี
กัน ในกลุ่มตัวอย่างขนาดเท่ากันจะแตกต่างกัน

3.2 อำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี
เดียวกันในกลุ่มตัวอย่างขนาดต่างกันจะแตกต่างกัน

นิยามศัพท์เฉพาะ

1. การทดสอบความแตกต่างของค่าเฉลี่ย หมายถึง การเปรียบเทียบค่าเฉลี่ยทีละคู่
หรือมากกว่าหนึ่งคู่ ว่าคู่ใดที่มีค่าต่างกัน โดยพิจารณาค่าเฉลี่ยแต่ละคู่ว่ามีความแตกต่างกันอย่างมี
นัยสำคัญหรือไม่

2. วิธีการทดสอบความแตกต่างของค่าเฉลี่ย หมายถึง กระบวนการเปรียบเทียบการ
ทดสอบความแตกต่างของค่าเฉลี่ย ที่มีเงื่อนไขว่าความแปรปรวนเท่ากันแต่ขนาดของกลุ่มตัวอย่าง
ต่างกัน ซึ่งในการศึกษาได้ศึกษาวิธีดังนี้ วิธี Tukey - Kramer วิธี Newman Keuls วิธี Duncan
วิธี Scheffe ' และวิธี Fisher' s LSD

2.1 วิธี Tukey - Kramer ใช้กับกลุ่มตัวอย่างที่มีขนาดเท่ากันหรือใกล้เคียงกัน
เปรียบเทียบทีละคู่ ภายใต้ความแปรปรวนที่เท่ากัน

2.2 วิธี Newman Keuls ใช้กับกลุ่มตัวอย่างที่มีขนาดเท่ากันหรือใกล้เคียงกัน เปรียบเทียบหลายคู่พร้อมกัน ภายใต้ความแปรปรวนที่เท่ากัน

2.3 วิธี Duncan ใช้กับกลุ่มตัวอย่างที่มีขนาดเท่ากัน เปรียบเทียบหลายคู่พร้อมกัน ภายใต้ความแปรปรวนที่เท่ากัน

2.4 วิธี Scheffe' ใช้กับกลุ่มตัวอย่างที่มีขนาดเท่ากันหรือใกล้เคียงกัน เปรียบเทียบหลายคู่พร้อมกัน ภายใต้ความแปรปรวนที่เท่ากัน

2.5 วิธี Fisher's LSD ใช้กับกลุ่มตัวอย่างที่มีขนาดเท่ากันหรือใกล้เคียงกัน เปรียบเทียบทีละคู่ ภายใต้ความแปรปรวนที่เท่ากัน

3. ผลการทดสอบความแตกต่างของค่าเฉลี่ย หมายถึง จำนวนความแตกต่างของค่าเฉลี่ยรายคู่ ในการยอมรับหรือไม่ยอมรับสมมติฐานว่าเหมือนหรือต่างกันหรือไม่ระหว่างวิธี Tukey - Kramer วิธี Newman Keuls วิธี Duncan วิธี Scheffe' และวิธี Fisher's LSD

4. ขนาดของกลุ่มตัวอย่าง หมายถึง จำนวนนักเรียนที่ผู้วิจัยสุ่มมาจากกลุ่มประชากร โดยแบ่งเป็น 4 ขนาดคือ ขนาด 30 คน 60 คน 90 คน และ 120 คน

5. อำนาจการทดสอบ (Power of Test) ความหมายทางคณิตศาสตร์ หมายถึง ความน่าจะเป็นหรือโอกาสที่จะปฏิเสธสมมติฐานศูนย์ (H_0) เมื่อ H_0 นั้น ผิดหรือมีค่าเท่ากับ $1 - \beta$ ส่วนความหมายทางการวิจัยที่สามารถตรวจสอบพบอิทธิพลจริง ๆ ของการจัดกระทำทางการทดลอง ในการวิจัยครั้งนี้ ผู้วิจัยได้หาอำนาจการทดสอบจากการเปิดตารางของ แท็งค์ (Tang. Power Function for Analysis of Variance) (Kirk. 1995 : 816 – 823)

6. จำนวนครั้งของการทดสอบ หมายถึง การวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยด้วยวิธีต่างกัน จากขนาดกลุ่มตัวอย่างเท่ากัน หรือการวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกันแต่ขนาดของกลุ่มตัวอย่างต่างกัน ซึ่งในการวิจัยครั้งนี้ผู้วิจัยใช้จำนวนครั้งในการทดสอบ 10 ครั้ง

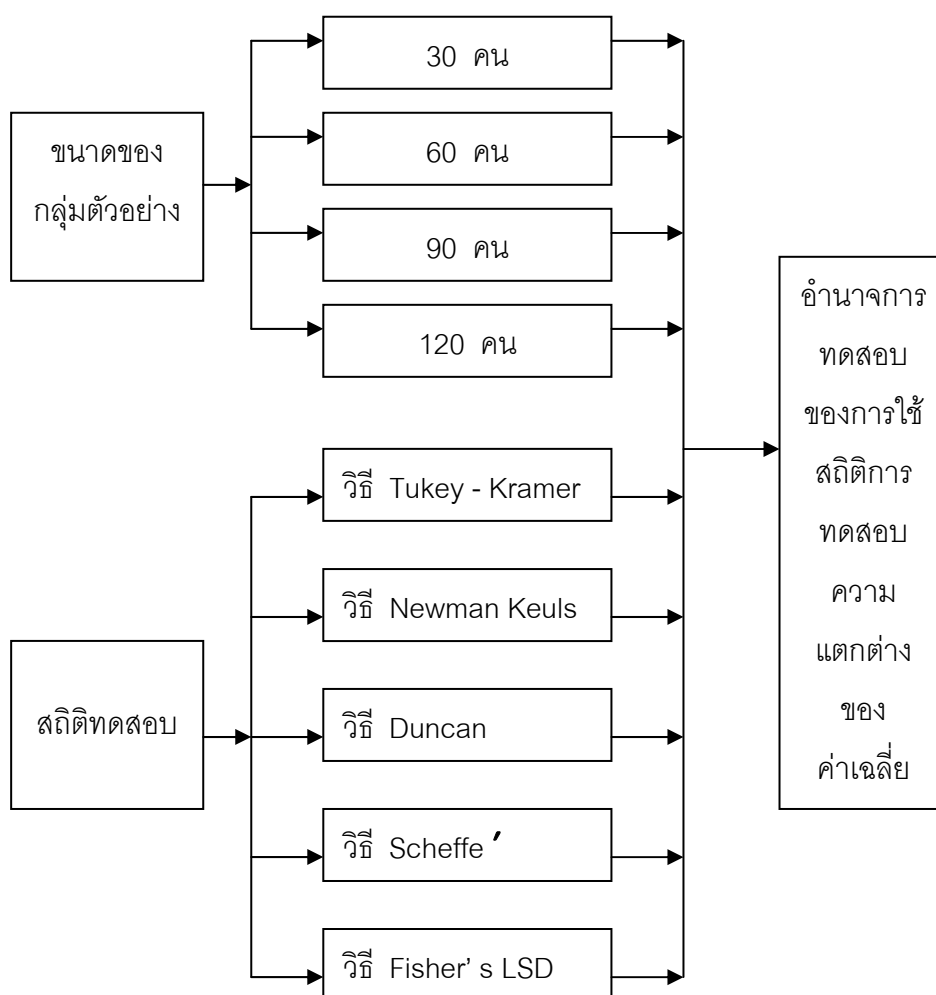
สมมติฐานการศึกษา

1. อำนาจการทดสอบ ที่ศึกษาโดยใช้ข้อมูลจริงวิเคราะห์ด้วยสถิติการทดสอบความแตกต่างของค่าเฉลี่ยด้วยต่างวิธีกัน จากกลุ่มตัวอย่างเท่ากัน อำนาจการทดสอบจะแตกต่างกัน

2. อำนาจการทดสอบ ที่ศึกษาโดยใช้ข้อมูลจริงวิเคราะห์ด้วยกลุ่มตัวอย่างขนาดต่างกันด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน อำนาจการทดสอบจะแตกต่างกัน

กรอบแนวคิดในการวิจัย

ในการศึกษาค้นคว้างานวิจัยเรื่อง ผลการเปรียบเทียบพหุคูณรายคู่บางวิธีภายใต้ความแปรปรวนที่แตกต่างกัน จากกลุ่มตัวอย่างขนาดต่างกัน (สิรินุช เอี่ยมเขียว.2544) และ งานวิจัยเรื่องการเปรียบเทียบอำนาจการทดสอบ (Power of Test) ของผลการวิเคราะห์ข้อมูลด้วยแบบแผนการวิเคราะห์แบบกลุ่มสุ่ม (RBD) กับแบบแผนการวิเคราะห์ความแปรปรวนร่วม (ANCOVA)(รุ่งโรจน์ ศรีจันทร์แก้ว.2547) ทำให้ผู้วิจัยสนใจที่จะศึกษา “อำนาจการทดสอบของการใช้สถิติการทดสอบความแตกต่างของค่าเฉลี่ย” โดยต้องการศึกษาความสัมพันธ์ของอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยต่างวิธีกัน ในกลุ่มตัวอย่างขนาดเท่ากัน และอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกันในกลุ่มตัวอย่างขนาดต่างกัน โดยแสดงความสัมพันธ์ของตัวแปรที่ต้องการศึกษา ดังภาพประกอบ 1



ภาพประกอบ 1 กรอบแนวคิดการวิจัย

บทที่ 2

เอกสารงานวิจัยที่เกี่ยวข้อง

ในการศึกษาครั้งนี้ผู้วิจัยได้ศึกษาเนื้อหาต่างๆ ดังนี้

1. ความหมายของการทดสอบความแตกต่างของค่าเฉลี่ย
2. การเปรียบเทียบระหว่างค่าเฉลี่ย
3. ประเภทของสถิติ
4. การทดสอบความเท่ากันของความแปรปรวน
5. อำนาจการทดสอบ (Power of Test)
6. สถิติทดสอบความแตกต่างของค่าเฉลี่ย

วิธีการทดสอบของ Tukey - Kramer

วิธีการทดสอบของ Newman Keuls

วิธีการทดสอบของ Duncan

วิธีการทดสอบของ Scheffe '

วิธีการทดสอบของ Fisher's LSD

7. ความถนัดและแบบทดสอบวัดความถนัด
8. งานวิจัยที่เกี่ยวข้อง

1. ความหมายของการทดสอบความแตกต่างของค่าเฉลี่ย

นักการศึกษาและนักจิตวิทยาได้นิยามหรือให้ความหมายเกี่ยวกับการทดสอบความแตกต่างของค่าเฉลี่ยไว้ ต่างๆ กันดังนี้

บุญเรียง ขจรศิลป์ (2533 : 51) ได้ให้ความหมายไว้ว่า การเปรียบเทียบระหว่างค่าเฉลี่ยของกลุ่มประชากรที่มากกว่า 2 กลุ่มคือ ความสัมพันธ์เชิงเส้นตรง (Linear Comparison) ของค่าเฉลี่ยของกลุ่มประชากรที่มากกว่า 2 กลุ่มนั้น

กัลยา วาณิชย์บัญชา (2539 : 160) ได้ให้ความหมายไว้ว่า เป็นการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยหลายๆ คู่ในเวลาเดียวกัน ก็ต่อเมื่อเกิดการปฏิเสธสมมติฐานหลัก

ระวีวรรณ พันธุ์พานิช (2540 : 215) ได้ให้ความหมายไว้ว่า ใช้เมื่อพบว่า F จากการวิเคราะห์ความแปรปรวนมีนัยสำคัญทางสถิติ ซึ่งตอบสมมติฐานการวิจัย เพียงว่า มีประชากรบางกลุ่มที่มีค่าเฉลี่ยแตกต่างกันจากกลุ่มอื่น แต่ไม่สามารถระบุได้ว่า กลุ่มใดบ้างที่มีค่าเฉลี่ยแตกต่างกัน

เนฟและวอร์ติงตัน (Neave; & Worthington. 1988 : 256) ได้ให้ความหมายไว้ว่า เป็นการศึกษาเปรียบเทียบเกี่ยวกับความแตกต่างระหว่างค่าเฉลี่ยของกลุ่มประชากร K ภายใต้สมมติฐาน

ทูทแฮกเกอร์ (Toothaker. 1993 : 2) ได้ให้ความหมายไว้ว่า เป็นการเปรียบเทียบค่าเฉลี่ยของกลุ่มประชากร J เช่น ความแตกต่างระหว่างประชากร J สองกลุ่ม หรือความแตกต่างระหว่างค่าเฉลี่ยกลุ่มที่หนึ่งและค่าเฉลี่ยกลุ่มที่สอง

โยเซฟและแทมเฮน (Yosef ; & Tamhane .1987 : 2) ได้ให้ความหมายไว้ว่า เป็นวิธีการทางสถิติที่ออกแบบมาเพื่อใช้ประกอบการตัดสินใจและการควบคุมผลกระทบในหลายๆ ด้านที่เกี่ยวกับการเปรียบเทียบในการวัดความคลาดเคลื่อน

วีส์ (Weiss. 1995 : 887) ได้ให้ความหมายไว้ว่า ในการตัดสินใจ เราต้องทราบว่าค่าเฉลี่ยแต่ละตัวแตกต่างกัน หรือค่าเฉลี่ยตัวใดมีค่ามากที่สุด และต้องทราบถึงความสัมพันธ์ระหว่างค่าเฉลี่ยทั้งหมด

2. การเปรียบเทียบระหว่างค่าเฉลี่ย

การเปรียบเทียบ (Comparison or Contrast) ระหว่างค่าเฉลี่ย เป็นความแตกต่างระหว่างค่าเฉลี่ย(วิรัช วรรณรัตน์ : 2535) ใช้สัญลักษณ์ทางพีชคณิต คือ ψ_i และ $\hat{\psi}_i$ โดย ψ_i แทนความแตกต่างระหว่างค่าเฉลี่ยของประชากรชุดที่ i และ $\hat{\psi}_i$ แทนความแตกต่างของค่าประมาณของกลุ่มตัวอย่างชุดที่ i ตัวอย่างเช่น $\psi_i = \mu_1 - \mu_2$ เป็นความแตกต่างของค่าเฉลี่ยประชากร μ_1 และ μ_2 $\hat{\psi}_i = \bar{Y}_1 - \bar{Y}_2$ เป็นความแตกต่างของค่าประมาณกลุ่มตัวอย่างถ้ามีการทดลอง 3 ระดับ จะได้ค่าเฉลี่ย 3 ค่า และได้ความแตกต่างดังนี้

$$\begin{array}{ll} \hat{\psi}_1 = \bar{Y}_1 - \bar{Y}_2 & \hat{\psi}_4 = \frac{\bar{Y}_1 + \bar{Y}_2}{2} - \bar{Y}_3 \\ \hat{\psi}_2 = \bar{Y}_1 - \bar{Y}_3 & \hat{\psi}_5 = \frac{\bar{Y}_1 + \bar{Y}_3}{2} - \bar{Y}_2 \\ \hat{\psi}_3 = \bar{Y}_2 - \bar{Y}_3 & \hat{\psi}_6 = \frac{\bar{Y}_2 + \bar{Y}_3}{2} - \bar{Y}_1 \end{array}$$

ความแตกต่างที่ 1 – 3 เป็นความแตกต่างของสองค่าเฉลี่ย ความแตกต่างที่ 4 – 6 เป็นความแตกต่างระหว่างจำนวนค่าเฉลี่ยของสองค่าเฉลี่ยกับอีกหนึ่งค่าเฉลี่ย ส่วนมากความแตกต่างระหว่างค่าเฉลี่ยจะอยู่ในรูปผลรวมเชิงเส้น มีค่าน้ำหนัก หรือค่าสัมประสิทธิ์แทนโดย C เมื่อค่าสัมประสิทธิ์อย่างน้อยหนึ่งตัวต้องไม่เท่ากับศูนย์

$$\hat{\psi}_i = C_1 \bar{Y}_1 + C_2 \bar{Y}_2 + \dots + C_k \bar{Y}_k$$

ได้ $\sum_{j=1}^k C_j = 0$ เมื่อ C_j อย่างน้อยหนึ่งตัวไม่เท่ากับศูนย์ เช่น

$$\hat{\psi}_1 = 1\bar{Y}_1 + (-1)\bar{Y}_2 + 0\bar{Y}_3 = \bar{Y}_1 - \bar{Y}_2$$

$$\hat{\psi}_2 = 1\bar{Y}_1 + 0\bar{Y}_2 + (-1)\bar{Y}_3 = \bar{Y}_1 - \bar{Y}_3$$

$$\hat{\psi}_3 = 0\bar{Y}_1 + 1\bar{Y}_2 + (-1)\bar{Y}_3 = \bar{Y}_2 - \bar{Y}_3$$

$$\hat{\psi}_4 = \frac{1}{2}\bar{Y}_1 + \frac{1}{2}\bar{Y}_2 + (-1)\bar{Y}_3 = \frac{\bar{Y}_1 + \bar{Y}_2}{2} - \bar{Y}_3$$

$$\hat{\psi}_5 = \frac{1}{2}\bar{Y}_1 + (-1)\bar{Y}_2 + \frac{1}{2}\bar{Y}_3 = \frac{\bar{Y}_1 + \bar{Y}_3}{2} - \bar{Y}_2$$

$$\hat{\psi}_6 = (-1)\bar{Y}_1 + \frac{1}{2}\bar{Y}_2 + \frac{1}{2}\bar{Y}_3 = \frac{\bar{Y}_2 + \bar{Y}_3}{2} - \bar{Y}_1$$

แต่การเปรียบเทียบมี C อย่างน้อยหนึ่งตัวเท่ากับศูนย์และผลรวมของ C_j เท่ากับศูนย์ หรือถ้าจะให้สะดวกในการเลือกเปรียบเทียบ คือ ผลรวมของค่าสัมบูรณ์ของ C_j เท่ากับสอง

$$\sum_{j=1}^k |C_j| = 2$$

จากตัวอย่าง ผลรวมของค่าสัมบูรณ์ของค่าสัมประสิทธิ์ $\hat{\psi}_1$ และ $\hat{\psi}_4$ คือ

$$|1| + |-1| + |0| = 1 + 1 + 0 = 2$$

$$\left|\frac{1}{2}\right| + \left|\frac{1}{2}\right| + |-1| = \frac{1}{2} + \frac{1}{2} + 1 = 2$$

3. ประเภทของสถิติ

คล็อกการและแซก (Klockars; & Sax. 1986 : 19 –22) ได้จำแนกรูปแบบทางสถิติออกเป็น 3 รูปแบบ คือ

สถิติแบบ t

ผู้ที่คิดค้นคือ Gosset ในปี 1908 การทดสอบแบบ t เป็นความคลาดเคลื่อนมาตรฐานของความแตกต่างระหว่างค่าเฉลี่ย ความคลาดเคลื่อนมาตรฐาน คือ หน้าทีของตัวแปรของข้อสังเกตและขนาดของกลุ่มตัวอย่าง เป็นการประมาณค่าตัวแปรของการสุ่มในการสังเกตที่ดีที่สุดโดยมาก คือ

ค่าเฉลี่ยที่ใช้เป็นเศษส่วนของการทดสอบ เพื่อการศึกษาผลกระทบของการจัดกระทำทั้งหมด ในการออกแบบกลุ่มสุ่มนี้ คือ กลุ่มที่รวมอยู่ในค่าเฉลี่ย

สถิติแบบ Q

โดยมากจะใช้สถิติแบบ Q ในการเปรียบเทียบรายคู่ เหมือนในการทดสอบแบบ t คือการวัดระยะห่างที่แน่นอนในหน่วยของความคลาดเคลื่อนมาตรฐาน Q จะแตกต่างจาก t เมื่อถูกใช้ในความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย เช่นเดียวกับเศษส่วนมากกว่าความคลาดเคลื่อนมาตรฐานของความแตกต่างของค่าเฉลี่ย ความแตกต่างในเศษส่วนเป็นการควบคุมโดยมุมมองที่แตกต่างในสิ่งที่สนใจ เมื่อการทดสอบ t ถูกใช้เป็นตัวแปรสุ่ม คือ ความแตกต่างระหว่างค่าเฉลี่ยสองค่า เมื่อ H_0 เป็นจริง ตัวแปรมีค่าเฉลี่ยเป็นศูนย์ และความคลาดเคลื่อนมาตรฐานจะเป็นตัวบอกถึงความไม่แน่นอนของความแตกต่างของค่าเฉลี่ยแต่อย่างไรก็ดีสถิติแบบ Q จะตั้งอยู่บนตัวแปรของค่าเฉลี่ยของกลุ่มตัวอย่างในการทดสอบหาค่าเฉลี่ยที่ใหญ่ที่สุดและเล็กที่สุดในกลุ่มของ k ค่าเฉลี่ยจะถูกกำหนดขึ้นจากประชากรที่เหมือนกัน และยังใช้ในการอธิบายจำนวนของความคลาดเคลื่อนมาตรฐานที่ค่าเฉลี่ยสองตัวสามารถแยกจากกันได้ และยังสามารถใช้ในการตัดสินใจจากการกระจายกลุ่มตัวอย่างที่เหมือนกัน หรืออาจใช้ในการตัดสินใจระดับของความมีนัยสำคัญ จำนวนกลุ่มของการจัดการกระทำจะถูกเปรียบเทียบ และจำนวนของระดับชั้นความเป็นอิสระของความคลาดเคลื่อนมาตรฐานของ Q ใช้ในการวัดว่าค่าเฉลี่ยสองตัวไหนที่มีนัยสำคัญที่แตกต่างกัน บางทีอาจจะใช้ในการทดสอบความเหมือนกันของค่าเฉลี่ยระดับกลาง

สถิติแบบ F

ผู้ที่คิดค้นคือ ฟิชเชอร์ (Fisher) การเปรียบเทียบเกี่ยวกับอัตราส่วน F เพื่อประกอบกรทดสอบสถิติตัวที่สาม ใช้ในการเปรียบเทียบพหุคูณ ข้อมูลเกี่ยวกับอัตราส่วน F และในเศษส่วนของระดับชั้นของความเป็นอิสระที่ 1 คือ ความเท่ากันของการทดสอบ t แบบสองทาง ค่าของจำนวนของอัตราส่วน F จะเท่ากับ t^2 ในการเปรียบเทียบนั้น แม้ว่าจะมีข้อมูลเท่ากัน การคำนวณสถิติสองตัวที่แตกต่างกันและยังต้องการความรอบคอบอีกด้วย

สถิติแบบ F คือ อัตราส่วนระหว่างค่าเฉลี่ยสองตัว เศษส่วนของอัตราส่วนนี้ คือ ค่าเฉลี่ยที่เกี่ยวข้องกับการเปรียบเทียบความน่าสนใจระหว่างค่าเฉลี่ยและเศษส่วนเป็นการวัดความแปรปรวนของการสุ่ม ถ้า H_0 เป็นจริงในการเปรียบเทียบแบบเจาะจง เศษส่วนจะประมาณค่าความแปรปรวนของประชากรถ้าเท็จเลขเศษส่วนจะมีความแปรปรวนมากกว่าเลขส่วนของลือก ที่ง่ายต่อการคำนวณหาค่าเฉลี่ยรวม แต่ละการจัดการกระทำของค่าเฉลี่ยแต่ละกลุ่มจะคูณด้วยน้ำหนักที่เป็นผลรวมของการ

ทดสอบจะมีค่าเป็นศูนย์ ถ้า H_0 เป็นจริงและมีค่าความแตกต่างตั้งแต่ศูนย์ ถ้า H_0 เป็นเท็จ เมื่อโครงร่างของน้ำหนักส่งผลต่อการเปรียบเทียบจะถูกพิสูจน์

ค่าน้ำหนักที่ซับซ้อนจะถูกใช้ในการเปรียบเทียบที่ซับซ้อน เช่น การเปรียบเทียบสิ่งที่น่าสนใจ อาจ会有ความแตกต่างระหว่างค่าเฉลี่ยตัวที่ 1 ในสองกลุ่มกับค่าเฉลี่ยที่คงที่ในกลุ่มที่ 3 ค่าน้ำหนักที่ใช้ในคำถามนี้ คือ 3, 3, -2, -2, -2 และค่าน้ำหนักส่วนมากจะมีผลรวมเป็นศูนย์ ดังนั้นถ้า H_0 เป็นจริง และค่าเฉลี่ยของประชากรเท่ากัน ผลรวมของการทดสอบจะไม่เท่ากับศูนย์ โดยจำนวนของค่าเฉลี่ยจะแตกต่างจากการสุ่มในการประมาณค่าเฉลี่ย

ผลรวมของการทดสอบค่าน้ำหนักและค่าเฉลี่ยจะเกี่ยวข้องถึง D และจะแสดงผลตามสูตร (1)

$$D = (a\bar{X}_1 + a\bar{X}_2 + \dots + a\bar{X}_k) \quad (1)$$

ค่าเฉลี่ยหรือความแปรปรวนสำหรับเศษส่วนในการเปรียบเทียบสามารถหาค่าได้ตามสูตร (2)

$$MSD = nD^2 / \sum a^2 \quad (2)$$

เมื่อ n แทนจำนวนของค่าเฉลี่ยในแต่ละกลุ่ม (โดยไม่ต้องสมมติจำนวนกลุ่มตัวอย่างให้เท่ากัน) และ $\sum a^2$ เป็นผลรวมของค่าน้ำหนัก (จากสูตรนี้สามารถปรับเปลี่ยนขนาดของกลุ่มตัวอย่างไม่ให้เท่ากันได้) ค่าเฉลี่ยของการเปรียบเทียบจะพบได้ โดยจะแบ่งจากค่าเฉลี่ยเพื่อสร้างสถิติ F อัตราส่วนจะมีนัยสำคัญต่อเมื่อ อัตราส่วนนั้นมีค่าเกินกว่าค่าวิกฤต F อัตราส่วนนี้จะวัดโดยการตัดสินใจเป็นที่ยอมรับกันทั่วไป

เคิร์ก (Kirk. 1995 : 124 – 125) ได้สรุปกระบวนการเปรียบเทียบพหุคูณโดยการให้รูปแบบของการเปรียบเทียบสถิติแบบใดแบบหนึ่ง คือ

$$\text{สถิติแบบ t : } \frac{\hat{\psi}}{\hat{\sigma}_{\psi}} = \frac{\sum_{j=1}^p C_j \bar{Y}_j}{\sqrt{MS_{error} \sum_{j=1}^p \frac{C_j^2}{n_j}}}$$

$$\text{สถิติแบบ Q : } \frac{\hat{\psi}}{\hat{\sigma}_{\bar{Y}}} = \frac{\sum_{j=1}^p C_j \bar{Y}_j}{\sqrt{\frac{MS_{error}}{n}}}$$

$$\text{สถิติแบบ F : } \frac{MS_{\text{set of means}}}{MS_{\text{error}}} = \frac{\sum_{j=1}^s n_j \bar{Y}_j^2 - \left(\sum_{j=1}^s n_j \bar{Y}_j \right)^2 / \sum_{j=1}^s n_j}{(s-1)MS_{\text{error}}}$$

เมื่อ p คือ จำนวนของค่าเฉลี่ย และ s คือ จำนวนของค่าเฉลี่ยในเซตของค่าเฉลี่ยตัวอักษร t , Q และ F เป็นการง่ายที่จะระบุถึงค่าสถิติและการกระจายของกลุ่มตัวอย่าง ได้แก่ การกระจายแบบ Student's t การกระจายแบบ Studentized range และการกระจายแบบ F ค่าวิกฤตสำหรับการกระจายเหล่านี้จะได้มาจากตาราง E เศษส่วนของสถิติแบบ t และสถิติแบบ Q จะถูกเปรียบเทียบว่าเป็นพิสัยประเภทไหน เศษส่วนคือการเปรียบเทียบความคลาดเคลื่อนมาตรฐาน ($\hat{\sigma}_\mu$) หรือความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย ($\hat{\sigma}_{\bar{Y}}$) เศษส่วนในสถิติแบบ F จะคำนวณมาจากค่าเฉลี่ยที่อยู่ในเซตของค่าเฉลี่ย ทั้งเศษและส่วนของสถิติแบบ F จะมีความแปรปรวน

สำหรับการเปรียบเทียบรายคู่ที่มีขนาดกลุ่มตัวอย่างเท่ากัน จะเกี่ยวข้องกับรูปแบบของการทดสอบทางสถิติ ดังนี้

$$t = \frac{Q}{\sqrt{2}} = \sqrt{F}$$

$$\frac{\hat{\psi}}{\hat{\sigma}} = \frac{\hat{\psi}}{\hat{\sigma}_{\bar{Y}} \sqrt{2}} = \sqrt{\frac{MS_{\text{set of means}}}{MS_{\text{error}}}}$$

โดยทั่วไปสถิติแบบ t จะมีแนวโน้มที่มีอำนาจมากกว่าสถิติแบบ F แต่เราเห็นว่าสถิติแบบ F จะต้องใช้การคำนวณหลายครั้งกว่าสถิติแบบ Q

การเปรียบเทียบความแตกต่างของค่าเฉลี่ยระหว่างกลุ่ม (Multiple Comparison Test) อาจเรียกอีกอย่างหนึ่งว่า การเปรียบเทียบพหุคูณ ใช้เมื่อพบว่า F จากการวิเคราะห์ความแปรปรวน มีนัยสำคัญทางสถิติ ซึ่งต้องตอบสมมติฐานการวิจัยเพียงว่า มีประชากรบางกลุ่มที่มีค่าเฉลี่ยแตกต่างจากกลุ่มอื่น แต่ไม่ทราบว่ากลุ่มใดบ้างที่มีค่าเฉลี่ยต่างกัน จึงต้องใช้การเปรียบเทียบพหุคูณ ซึ่งทำได้ 2 วิธี (ระวีวรรณ พันธุ์พานิช. 2540 : 51)

1. การเปรียบเทียบแบบวางแผน (Planned or Priori contrast) เป็นเทคนิควิธีการใช้เปรียบเทียบค่าเฉลี่ยของกลุ่มประชากรที่กำหนดไว้ล่วงหน้าก่อนการวิเคราะห์ ความแปรปรวน โดยการเปรียบเทียบเฉพาะค่าเฉลี่ยของบางกลุ่มที่ผู้วิจัยสนใจเท่านั้น

2. การเปรียบเทียบภายหลัง (Post – hoc or posteriori contrast) เป็นเทคนิคที่ใช้เปรียบเทียบค่าเฉลี่ยของกลุ่มประชากรหลังจากการปฏิเสธ H_0 การเปรียบเทียบความแตกต่างของค่าเฉลี่ยทุกกลุ่มที่เป็นไปได้

กรณีข้อมูลมีลักษณะบางประการไม่เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนจะมีผลต่อความเปลี่ยนแปลงของระดับนัยสำคัญและความไวของตัวทดสอบซึ่งมีผลต่ออำนาจการทดสอบดังนี้

1. ถ้าข้อมูลมีการแจกแจงไม่เป็นโค้งปกติ

ถ้าประชากรที่ศึกษา มีการแจกแจงคะแนนมีความเบ้ (Skewness) หรือมีความโด่ง (Kurtosis) ไม่มากนัก หรือแต่ละกลุ่มประชากรมีการแจกแจงคล้ายคลึงกัน การแจกแจง F จะไม่ค่อยได้รับความกระทบกระเทือน

เพื่อแก้ปัญหาการแจกแจงประชากรที่ไม่เป็นปกติ และเพื่อให้การทดสอบมีความคงทน (test is robust) ซึ่งหมายถึงกรณีที่ข้อมูลลักษณะไม่เป็นไปตามข้อตกลงเบื้องต้น แต่ค่าระดับนัยสำคัญของการทดสอบยังคงใกล้เคียงกับค่าที่กำหนดไว้ ควรกำหนดตัวอย่างให้ใหญ่เพียงพอหรือใช้สถิตินอนพาราเมตริก

2. ถ้าความแปรปรวนของประชากรไม่เท่ากันทุกกลุ่มหรือไม่มีความเป็นเอกพันธ์

เมื่อทดสอบความเป็นเอกพันธ์ของความแปรปรวนของกลุ่มประชากรที่มีข้อมูลในกลุ่มตัวอย่างต่างกันแล้วพบว่า ความแปรปรวนแตกต่างกันจะแก้ไขโดยเพิ่มขนาดของกลุ่มตัวอย่างให้ทุกกลุ่มมีขนาดเท่ากัน แต่หากจำเป็นต้องใช้จำนวนกลุ่มตัวอย่างไม่เท่ากัน การฝ่าฝืนข้อตกลงเบื้องต้นข้อนี้จะมีผลอย่างมากต่อการทดสอบนัยสำคัญ (ระวีวรรณ พันธุ์พานิช. 2540 : 40) การทดสอบความสำคัญของความแตกต่างระหว่างความแปรปรวนของสองกลุ่ม ในกรณีที่ผลวิจัยต้องการทดสอบความแปรปรวนของกลุ่มประชากรกลุ่มหนึ่ง มากกว่าความแปรปรวนของประชากรอีกกลุ่มหนึ่งหรือไม่ โดยใช้ข้อมูลจากกลุ่มตัวอย่างที่สุ่มมาจากประชากรและกรณีที่ต้องการทดสอบความเป็นเอกพันธ์ของความแปรปรวน (Homogeneous of variance) เพื่อทราบว่าที่ได้ assume ว่าความแปรปรวนของประชากรเท่ากัน หรือไม่เท่ากันนั้น มีความถูกต้องเที่ยงตรงหรือไม่ มีวิธีการทดสอบหลายวิธี เช่น วิธีใช้ F -test วิธีของ Hartley วิธีของ Cochran เป็นต้น (บุญชม ศรีสะอาด. 2538 : 233)

โดยมีข้อตกลงเบื้องต้นว่า

1. ประชากรที่สุ่มตัวอย่างมา มีการแจกแจงของค่าตัวแปรที่ศึกษาเป็นแบบปกติ
2. กลุ่มตัวอย่างเป็นอิสระต่อกัน
3. กลุ่มตัวอย่างแต่ละกลุ่มต้องได้มาจากประชากรที่มีความแปรปรวนเท่ากัน
4. ข้อมูลที่นำมาวิเคราะห์ต้องอยู่ในมาตราอันดับหรืออัตราส่วน

4. การทดสอบความเท่ากันของความแปรปรวน

เดดตัน (Dayton.1970 : 32 - 37) กล่าวว่า ในแบบแผนการวิจัยแบบกลุ่มสมบูรณ์นั้น ความแปรปรวนในแต่ละกลุ่มการจัดการกระทำจะประมาณได้จากความแปรปรวนของคะแนนประชากร โดยทำการทดสอบความเท่ากันของความแปรปรวนคือ $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_p^2 = \sigma^2$ เมื่อ σ_j^2 แทน ความแปรปรวนของคะแนนประชากรกลุ่มที่ j แต่เมื่อทำการทดสอบกับกลุ่มตัวอย่างจะได้ความแปรปรวนของกลุ่มตัวอย่าง คือ $S_1^2, \dots, S_p^2$ เป็นการประมาณค่าความแปรปรวนของคะแนนประชากรวิธีการทดสอบความเท่ากันของความแปรปรวนที่ใช้กันอย่างแพร่หลาย มีอยู่ 3 วิธี คือ

1. การทดสอบ Bartlett สามารถใช้กับกลุ่มตัวอย่างที่เท่ากัน หรือขนาดกลุ่มตัวอย่างที่แตกต่างกัน
2. การทดสอบ Hartley จะมีวิธีการคำนวณที่ง่ายกว่าการทดสอบของ Bartlett แต่ถูกออกแบบมาเพื่อใช้กับขนาดกลุ่มตัวอย่างที่เท่ากันเพียงอย่างเดียว
3. การทดสอบของ Levene มีความเหมาะสมที่จะใช้มากแต่ไม่สามารถคำนวณได้โดยตรงจากความแปรปรวนของคะแนนกลุ่มตัวอย่าง

วิธีของ Bartlett

จะเป็นการทดสอบที่มีประสิทธิภาพ ถ้าแต่ละประชากรมีการแจกแจงแบบปกติ แต่ถ้าการแจกแจงของประชากรเป็นแบบหางยาว (Long – tailed distribution) หรือเป็นการแจกแจงที่มีความเบ้เป็นบวก (Positive skewness) ผลการทดสอบด้วยวิธีของ Bartlett จะปฏิเสธ H_0 เป็นส่วนใหญ่ ดังนั้นจึงควรใช้วิธีของ Bartlett ในการทดสอบความเท่ากันของค่าความแปรปรวนหลายประชากร เมื่อแต่ละประชากรมีการแจกแจงแบบปกติ

วิธีของ Levene

Levene ได้นำเสนอวิธีการทดสอบความเท่ากันของความแปรปรวนของประชากร ซึ่งมีวิธีการคำนวณคล้ายกับการวิเคราะห์ความแปรปรวนแบบทางเดียว โดยการแทนค่าสังเกตแต่ละค่าด้วยค่าสัมบูรณ์ของส่วนเบี่ยงเบนระหว่างค่าสังเกตแต่ละค่ากับค่าเฉลี่ยของกลุ่มตัวอย่างนั้นๆ เป็นตัวทดสอบที่มีความแข็งแกร่งต่อลักษณะการแจกแจงของประชากรที่ไม่เป็นปกติ (เยาวภา ไชยศรี. 2536 :25 อ้างอิงจาก Levene . 1960 : 278 – 292)

ตัวสถิติทดสอบคำนวณได้จาก

$$L = \frac{\sum n_i (\bar{Z}_i - \bar{Z}_{...})^2 / (k-1)}{\sum \sum (Z_{ij} - \bar{Z}_i)^2 / \sum (n_i - 1)} \quad ; \text{df} = k - 1, \sum (n_i - 1)$$

เมื่อ

$$Z_{ij} = |X_{ij} - \bar{X}_i|$$

$$\bar{Z}_i = \sum Z_{ij} / n_i$$

$$\bar{Z}_{...} = \sum \sum Z_{ij} / \sum n_i$$

ถ้า L คำนวณ $> F$ จากตารางการแจกแจง F จะไม่ยอมรับ H_0

5. อำนาจการทดสอบ (Power of Test)

5.1 การทดสอบสมมติฐาน (Hypothesis Testing)

ในการทดสอบสมมติฐานจะเกิดผลลัพธ์ 4 รูปแบบที่สามารถเกิดขึ้นเมื่อนักวิจัยนำผลจากการทดลองไปใช้ประโยชน์ เช่นการค้นหว่าการจัดการกระทำทางการทดลองใดมีผลมากกว่าการจัดการทดลองอื่น และผลที่ได้รับจากกลุ่มตัวอย่างจะอ้างอิงไปสู่ประชากร ซึ่งผลทั้ง 4 แบบแสดงในภาพประกอบ 2

	H_0 จริง	H_0 เท็จ
การตัดสินใจ ปฏิเสธ H_0	Type I error $P = \alpha$	ตัดสินใจถูกต้อง $P = 1 - \beta$ (Power)
ยอมรับ H_0	ตัดสินใจถูกต้อง $P = 1 - \alpha$	Type II error $P = \beta$

ภาพประกอบ 2 ผลลัพธ์ที่เกิดจากการทดสอบสมมติฐานทางสถิติ

ในภาพประกอบ 2 แสดงให้เห็นแนวทางที่จะเกิดความคลาดเคลื่อนเมื่อทำการทดสอบสมมติฐาน แนวทางแรก คือ ความน่าจะเป็นที่ผลของการจัดการกระทำทางการทดลองไม่มีอิทธิพลจริงในประชากร แต่ผลลัพธ์จากการทดสอบสมมติฐานของกลุ่มตัวอย่างจะนำไปสู่ความเชื่อมั่นว่าการจัดการกระทำทางการทดลองมีอิทธิพลบ้าง ถ้าผลจากการศึกษาครั้งนี้ถูกนำไปใช้ประโยชน์ ผลการทดลองจะสรุปได้ว่าวิธีการใหม่ดีกว่าวิธีการเดิม แต่ในความเป็นจริงแล้วทั้งสองวิธีการไม่แตกต่างกัน

นั่นคือเกิดความคลาดเคลื่อนประเภทที่ 1 (Type I error) ความคลาดเคลื่อนประเภทนี้มักทำให้นักวิจัยเสียเวลาและทรัพยากรเมื่อนำไปปฏิบัติจริง ความน่าจะเป็นในการเกิดความคลาดเคลื่อนประเภทที่ 1 เป็นส่วนหนึ่งในองค์ประกอบของมาตรฐานหรือหลักการตัดสินใจที่ใช้ในการทดสอบสมมติฐาน คือการตั้งระดับนัยสำคัญทางสถิติ หากตั้งในระดับที่ต่ำมากจะนำไปสู่ความถี่ในการเกิดความคลาดเคลื่อนประเภทที่ 1

แนวทางที่สองคือ การเกิดความคลาดเคลื่อนประเภทที่ 2 (Type II Error) เกิดขึ้นเมื่อผู้วิจัยเข้าข้างสมมติฐานกลาง เมื่อในความเป็นจริงแล้วสมมติฐานทางเลือกถูกต้อง การวิเคราะห์อำนาจการทดสอบทางสถิตินั้นมีความเกี่ยวข้องกับการเกิดความคลาดเคลื่อนประเภทที่ 2 เนื่องจากอำนาจการทดสอบถูกกำหนดให้เป็นหนึ่งลบด้วยความน่าจะเป็นในการเกิดความคลาดเคลื่อนประเภทที่ 2 คือ $\text{power} = 1 - \beta$ (Murphy; & Mayors. 2004 : 4 – 5) และการวางแผนการทดลองจะประสบความสำเร็จได้ต้องมีความสมดุลระหว่างความคลาดเคลื่อนทั้งสองชนิด (Stevens. 2002 : 5) ซึ่งสามารถจัดการด้วยมโนทัศน์ของอำนาจการทดสอบ (Runyon ; & Haber. 1991 : 452)

5.2 ความหมายของอำนาจการทดสอบ (Power of the test)

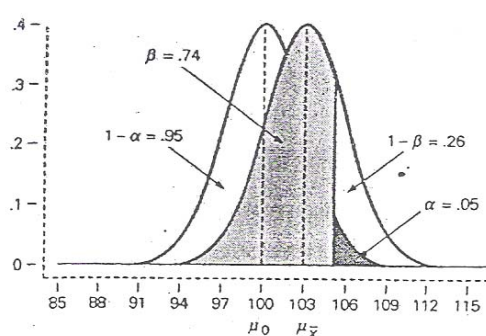
อำนาจการทดสอบ (Power of the test) คือ ความน่าจะเป็นที่มีค่าแปรผันตั้งแต่ 0.00 – 1.00 ค่าสูงสุดของอำนาจการทดสอบคือ 0.80 เป็นค่าที่นักวิจัยปรารถนามากแต่จะพบในการวิจัยด้านพฤติกรรมศาสตร์เท่านั้น ค่าอำนาจการทดสอบที่มีค่าช่วง 0.40 – 0.60 จะเป็นค่าปกติ (Pagano. 2004 : 244) เราสามารถที่จะคำนวณค่าอำนาจการทดสอบเมื่อสมมติฐานกลางเป็นเท็จเท่านั้น หรืออาจกล่าวได้ว่าเมื่อสมมติฐานกลางเป็นจริง เราไม่สามารถจะสร้างความคลาดเคลื่อนประเภทที่ 2 ได้ (Runyon ; & Haber. 1991 : 452) ด้วยเหตุนี้ จุดมุ่งหมายหลักของการวิเคราะห์อำนาจการทดสอบ ประกอบด้วย (1) การออกแบบการวิจัย (2) การประเมินความสำเร็จของงานวิจัย ในการออกแบบงานวิจัย การวิเคราะห์อำนาจการทดสอบถูกใช้เพื่อการตัดสินใจขนาดกลุ่มตัวอย่างที่ต้องการในการวิจัยที่ระดับนัยสำคัญทางสถิติสำหรับทุกขนาดอิทธิพลนั้นๆ โดยการใช้การวิเคราะห์อำนาจการทดสอบเพื่อช่วยตัดสินใจ นอกจากนั้นแล้วอำนาจการทดสอบยังเป็นสิ่งที่ชี้ให้เห็นถึงความล้มเหลวในการตรวจพบอิทธิพลซึ่งได้รับการกำหนดค่านัยสำคัญทางสถิติ ในการใช้กลุ่มตัวอย่างขนาดเล็ก การพิจารณาระดับของอำนาจการทดสอบขึ้นอยู่กับ (1) สถิติทดสอบที่ใช้ (2) ระดับนัยสำคัญทางสถิติที่เลือก (3) ขนาดกลุ่มตัวอย่างที่ใช้ (4) ขนาดอิทธิพลที่ทำการศึกษา (Rosenthal ; & Rosnow. 1991 : 440)

5.3 อิทธิพลที่ส่งผลต่ออำนาจการทดสอบ

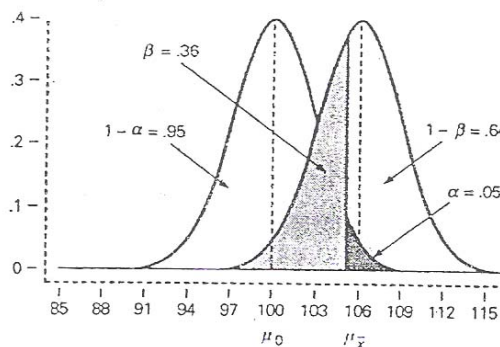
รันยอนและคณะ (Runyon ; et al. 1996 : 417) กล่าวว่า องค์ประกอบที่มีอิทธิพลที่ทำให้เกิดความเป็นไปได้ในการปฏิเสธสมมติฐานกลางที่เป็นเท็จนั้น บางองค์ประกอบสามารถควบคุมได้ บางองค์ประกอบไม่สามารถควบคุมได้

1. ความแตกต่างระหว่าง μ_0 และ $\mu_{\bar{x}}$

รันยอนและคณะ (Runyon ; et al. 1996 : 417- 418) ได้อธิบายอิทธิพลของความแตกต่างระหว่าง μ_0 และ $\mu_{\bar{x}}$ ที่ส่งผลต่ออำนาจการทดสอบว่า ถ้าสมมติฐานกลางเป็นเท็จและค่าเฉลี่ยของประชากรจริงแตกต่างจากค่าเฉลี่ยของสมมติฐาน จะมีการแจกแจงของกลุ่มตัวอย่าง 2 การแจกแจงซึ่งแสดงในภาพประกอบ 3 การแจกแจงทั้งสองสามารถอธิบายโดยทฤษฎีลิมิตสู่ส่วนกลาง การแจกแจงของกลุ่มตัวอย่างในสมมติฐานกลางเป็นการแจกแจงของค่าเฉลี่ยของกลุ่มตัวอย่างด้วย $\mu_0 = 100$ ส่วนการแจกแจงลำดับที่สองคือการแจกแจงค่าเฉลี่ยของกลุ่มตัวอย่างที่ควรได้รับ $\mu_{\bar{x}} = 103$



3(ก)



3(ข)

ภาพประกอบ 3 อิทธิพลของความแตกต่างระหว่าง μ_0 และ $\mu_{\bar{x}}$ ที่ส่งผลต่ออำนาจการทดสอบ

ที่มา : Runyon ; et al.(1996). Fundamental of behavioral Statistice.p. 418.

จากภาพประกอบ 3 จะเห็นว่าการแจกแจงทั้งสองซ้อนทับกัน และพื้นที่สีดำนั้นแสดง α หรือความน่าจะเป็นในการพบความคลาดเคลื่อนประเภทที่ 1 และส่วนที่เป็นสีเทาจึงแสดง β หรือความน่าจะเป็นในการพบความคลาดเคลื่อนประเภทที่ 2 การแจกแจงกลุ่มตัวอย่างทางด้านซ้ายแสดง μ_0 ส่วนที่ไม่ได้แรเงาของโค้งแสดง $(1 - \alpha)$ คือความน่าจะเป็นในการยอมรับสมมติฐานกลางที่เป็นจริง ในการแจกแจงทางด้านขวาคือ การแจกแจงที่แสดง $\mu_{\bar{x}}$ ส่วนที่ไม่ได้มีการแรเงาของโค้งแสดง $(1 - \beta)$ คือความน่าจะเป็นในการปฏิเสธสมมติฐานกลางที่เป็นเท็จ

สำหรับการแจกแจงในภาพประกอบ 3 (ก) ความแตกต่างระหว่างค่าเฉลี่ยของ 2 ประชากรที่มีความสัมพันธ์กัน $\mu_{\bar{x}} - \mu_0 = 103 - 100 = 3$ เปรียบเทียบความแตกต่างนี้กับความแตกต่างในภาพประกอบ 3 (ข) ซึ่งความแตกต่างระหว่างค่าเฉลี่ยที่มีขนาดใหญ่กว่า $\mu_{\bar{x}} - \mu_0 = 106 - 100 = 6$

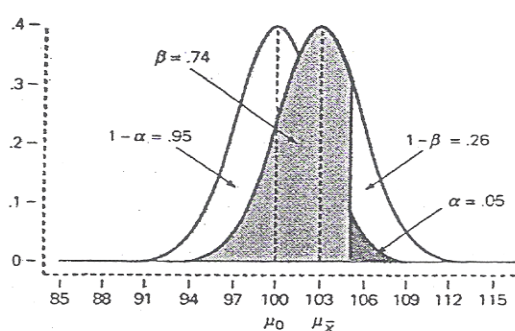
สรุปได้ว่าภายใต้เงื่อนไขของการวิจัยเมื่อ กลุ่มตัวอย่าง ความแปรปรวนของประชากร และระดับนัยสำคัญทางสถิติคงที่ ความน่าจะเป็นในการปฏิเสธสมมติฐานกลางจะมีค่าน้อยเมื่อ $(\mu_0 - \mu_{\bar{x}})$ มีค่าน้อย และจะเพิ่มขึ้นเมื่อ $(\mu_0 - \mu_{\bar{x}})$ มีค่ามาก

2. ขนาดกลุ่มตัวอย่าง

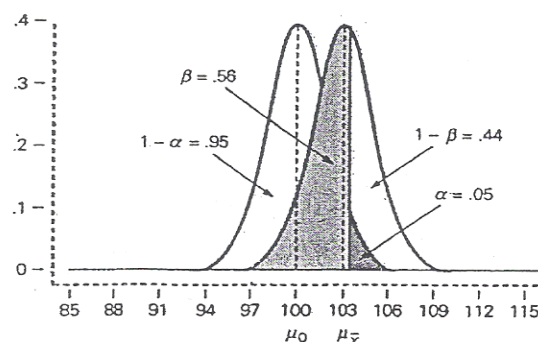
รันยอนและคณะ (Runyon ; et al. 1996 : 417 - 418) ได้อธิบายไว้ว่า กลุ่มตัวอย่างมีความสำคัญในการประมาณค่าพารามิเตอร์ของประชากร เมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นความแม่นยำในการประมาณค่าจะเพิ่มขึ้นด้วย เมื่อการกระจายของกลุ่มตัวอย่างลดลง ภาพประกอบ 4 (ก) แสดงให้เห็นการกระจายที่ควรได้รับถ้าทำการเลือกตัวอย่างจากประชากรโดยใช้กลุ่มตัวอย่างขนาดเล็ก ภาพประกอบ 4(ข) แสดงให้เห็นการกระจายทั้งสองถ้าได้รับจำนวนตัวอย่างเพิ่มขึ้น ปรากฏการณ์นี้อธิบายด้วยทฤษฎีลิมิตสู่ส่วนกลางความแปรปรวนของการกระจายของกลุ่มตัวอย่าง

$$\text{คือ } \sigma_{\bar{x}} = \sigma / \sqrt{N}$$

เมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้น ระดับของการกระจายของกลุ่มตัวอย่างจะลดลง และส่งผลให้อำนาจการทดสอบเพิ่มขึ้น ซึ่งสอดคล้องกับ ไมเนียม คิง และแบร์ (Minium ; King & Bear. 1993 : 359) ได้อธิบายอิทธิพลของขนาดกลุ่มตัวอย่างที่ส่งผลต่ออำนาจการทดสอบโดยใช้ทฤษฎีลิมิตสู่ส่วนกลาง กล่าวว่า ความน่าจะเป็นในการปฏิเสธสมมติฐานกลางที่เป็นเท็จจะมีความสัมพันธ์กับขนาดกลุ่มตัวอย่าง สมมติว่าในการทดสอบวิธีการเรียนรู้แบบใหม่ใช้กลุ่มตัวอย่างจำนวน 25 คน จะได้ค่าประมาณความเบี่ยงเบนมาตรฐานของค่าเฉลี่ยคือ $S_{\bar{x}} = S_x / \sqrt{25} = .02S_x$ แต่ถ้าใช้กลุ่มตัวอย่างจำนวน 1,000 คน $S_{\bar{x}} = S_x / \sqrt{1000} = 0.01S_x$ โดยทั่วไปกลุ่มตัวอย่างขนาดใหญ่จะมีค่าเบี่ยงเบนมาตรฐานของค่าเฉลี่ยของกลุ่มตัวอย่างขนาดเล็ก $(S_{\bar{x}})$ สรุปได้ว่ากลุ่มตัวอย่างขนาดใหญ่จะมีค่าคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยขนาดเล็กและจะมีค่าอำนาจการทดสอบขนาดใหญ่



4(ก)



4(ข)

ภาพประกอบ 4 อิทธิพลจากขนาดกลุ่มตัวอย่าง และความแปรปรวนที่ส่งผลต่ออำนาจการทดสอบ

ที่มา :Runyon ; et al. (1996). Fundamental of behavioral Statistic.p. 420.

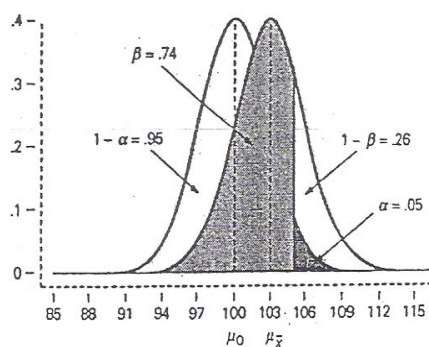
3. ความแปรปรวนในประชากร

ถ้าพิจารณาที่ความแปรปรวน สามารถพิจารณาจากภาพประกอบ 4 อีกครั้ง จากภาพประกอบ 4(ก) แสดงให้เห็นว่าเมื่อความแปรปรวนของประชากรมีค่ามาก เมื่อพิจารณาที่การซ้อนทับของการกระจาย 2 กลุ่มตัวอย่างพบว่าจะมีค่าขนาดเล็ก ซึ่งส่งผลให้อำนาจการทดสอบลดลง แตกต่างจากภาพประกอบ 4(ข) ความแปรปรวนของกลุ่มตัวอย่างมีค่าน้อย ทำให้การซ้อนทับของการกระจาย 2 กลุ่มตัวอย่างมีขนาดใหญ่ ซึ่งจะส่งผลให้อำนาจการทดสอบเพิ่มขึ้น (Runyon ; et al. 1996 : 419) หรืออธิบายด้วยความเบี่ยงเบนมาตรฐาน ซึ่งเป็นตัวสะท้อนให้เห็นความแปรปรวนอันสืบเนื่องมาจากองค์ประกอบที่ไม่เกี่ยวข้องกันสิ่งที่ต้องการศึกษา ความแปรปรวนที่อยู่นอกประเด็น ซึ่งมาจากแหล่งที่ไม่เกี่ยวข้อง จะมีแนวโน้มให้ค่าเบี่ยงเบนมาตรฐานเพิ่มขึ้น ดังนั้นอิทธิพลจากการพยายามกำจัดแหล่งความแปรปรวน ทำให้ความแปรปรวนมีแนวโน้มลดลงและเพิ่มค่าอำนาจการทดสอบ (Minium; King & Bear. 1993 : 360)

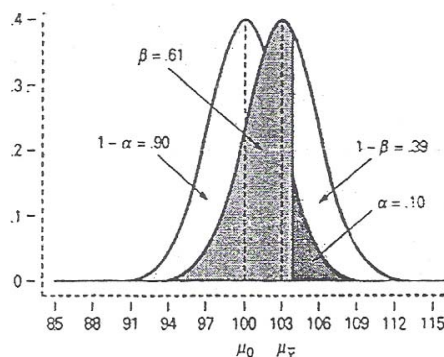
4. ระดับนัยสำคัญทางสถิติ

ระดับนัยสำคัญทางสถิติ (α) มีผลต่อความน่าจะเป็นในการปฏิเสธสมมติฐานกลางที่เป็นจริง ซึ่งแสดงในภาพประกอบ 5 ถ้าระดับนัยสำคัญทางสถิติ มีระดับนัยสำคัญทางสถิติ มีระดับต่ำ ทำให้ β (ความคลาดเคลื่อนชนิดที่ 2) มีค่าสูงขึ้น จะทำให้อำนาจการทดสอบลดลง แต่ถ้าเพิ่มระดับนัยสำคัญทางสถิติจะส่งผลให้ β มีค่าน้อยลงและอำนาจการทดสอบจะมีค่าเพิ่มขึ้น (Runyon; et al. 1996 : 420) ซึ่งสอดคล้องกับ เวลโควิทซ์ อีเวน และโคเฮน (Welkowitz; Ewen; & Cohen. 2000 : 206) ที่กล่าวว่า ระดับนัยสำคัญทางสถิติมีความสัมพันธ์ต่อการปฏิเสธสมมติฐานกลางแน่นอน ระดับนัยสำคัญทางสถิติเท่ากับ ความน่าจะเป็นในการเกิดความคลาดเคลื่อนประเภทที่ 1 ที่

ใช้บ่อยๆ คือ 0.05 ซึ่งการเข้มงวดกับเกณฑ์นี้จะยากต่อการปฏิเสธสมมติฐานกลาง และอำนาจการทดสอบจะต่ำ



5(ก)



5(ข)

ภาพประกอบ 5 อิทธิพลจากระดับนัยสำคัญทางสถิติที่ส่งผลต่ออำนาจการทดสอบ

ที่มา : Runyon; et al. (1996). Fundamental of Behavioral Statistics. P 421.

5. การทดสอบแบบมีทิศทาง กับ การทดสอบแบบไม่มีทิศทาง

รันยอน และคณะ (Runyon ; et al. 1996 : 421) กล่าวว่า การเลือกการทดสอบแบบมีทิศทางและไม่มีทิศทางมีอิทธิพลต่ออำนาจการทดสอบ การทดสอบแบบมีทิศทางจะให้อำนาจการทดสอบมากกว่าการทดสอบแบบไม่มีทิศทาง เหตุผลคือ การทดสอบแบบไม่มีทิศทางเขตวิกฤตจะถูกแบ่งครึ่ง แต่ละเขตจะอยู่ที่ปลายสุดของการแจกแจง ผลที่ตามมาคือเมื่อใช้การทดสอบแบบไม่มีทิศทาง จึงมีความน่าจะเป็นในการปฏิเสธสมมติฐานกลางมากกว่าการใช้การทดสอบแบบมีทิศทาง ซึ่งจะเพิ่มความสามารถในการปฏิเสธสมมติฐานกลาง

5.4 ขนาดอิทธิพล

โคเฮน (Cohen. 1988 : 9) ได้กล่าวถึง ปรากฏการณ์ภายใต้การทดสอบสมมติฐานทางสถิติของประชากรนั้นต้องทำการพิจารณาทั้งปรากฏการณ์ที่ไม่ปรากฏให้เห็น (สมมติฐานกลางเป็นจริง) และปรากฏการณ์ที่ปรากฏให้เห็น (สมมติฐานกลางเป็นเท็จ) ปรากฏการณ์ที่ไม่ปรากฏให้เห็นจะแสดงนัยของพารามิเตอร์ของแต่ละประชากร เช่น ในการศึกษาเพื่อตัดสินว่าเพศต่างกันส่งผลต่ออัตราการเกิดโรคจิตแตกต่างกันหรือไม่ ในการทดสอบอาจนำกลุ่มตัวอย่างที่เป็นคนไข้ซึ่งมีอาการนำมาวินิจฉัยและตัดสินด้วยคำตัดสิน ส่วนจากการทดสอบสมมติฐานพบว่าสัดส่วนของประชากรเพศชายเป็น 0.50 ของกลุ่มตัวอย่างทั้งหมด อาจกล่าวได้ว่าอิทธิพลจากเพศที่ปรากฏจากการวินิจฉัยมี

ขนาดเป็นศูนย์ ดังนั้นปรากฏการณ์ที่ไม่ปรากฏให้เห็นภายใต้การศึกษาจะแสดงสมมติฐานกลาง ซึ่งแสดงด้วยค่าพารามิเตอร์ของประชากร และแทนด้วยคำว่า “ขนาดอิทธิพล (effect size)” มีความหมายว่า ระดับของแต่ละปรากฏการณ์ที่ปรากฏในประชากร ขนาดอิทธิพลเกี่ยวข้องกับ 2 ส่วนที่สำคัญ คือ ส่วนที่ 1 คือขนาดอิทธิพลมีความสำคัญอย่างยิ่งในการตัดสินใจเกี่ยวกับอำนาจการทดสอบ และการกำหนดขนาดกลุ่มตัวอย่าง ส่วนที่ 2 ขนาดอิทธิพลเป็นความสัมพันธ์อย่างหนึ่งของมโนทัศน์แวดล้อมของสถิติอ้างอิง ที่ปฏิบัติกันอยู่ในงานวิจัยด้านพฤติกรรมศาสตร์ ความสัมพันธ์ระหว่างขนาดอิทธิพลและอำนาจการทดสอบที่ชัดเจนคือ เมื่อระดับนัยสำคัญทางสถิติและอำนาจการทดสอบคงที่ กลุ่มตัวอย่างขนาดเล็กจำเป็นต้องการพบขนาดอิทธิพล ต่อมาโฮเวลล์ (Howell, 1992 : 207 – 208) ได้ให้แนวคิดเพิ่มเติมเกี่ยวกับอำนาจการทดสอบว่าขึ้นอยู่กับระดับของการซ้อนทับระหว่างการกระจายของกลุ่มตัวอย่าง 2 กลุ่มภายใต้สมมติฐานกลาง (H_0) และสมมติฐานทางเลือก (H_1) และความคลาดเคลื่อนมาตรฐาน (Standard error) รูปแบบการวัดระดับการซ้อนทับในแต่ละสมมติฐานกลางเมื่อสมมติฐานกลางเป็นเท็จมักใช้ระยะทางระหว่าง μ_0 ถึง μ_1 เพื่อแสดงความคลาดเคลื่อนมาตรฐานการวัดระยะทางหรือเรียกว่าขนาดอิทธิพล แทนด้วย d

$$d = \frac{\mu_1 - \mu_0}{\sigma}$$

ดังนั้น d คือค่าที่ได้จากการวัดขนาดใหญ่ในแต่ละ μ_0 และ μ_1 ที่ต่างกันในเทอมของความเบี่ยงเบนมาตรฐานของประชากร

นอกจากนี้ รันยอน และคณะ (Runyon ; et al. 1996 : 453) ได้อธิบายแนวทางในการแปลความหมายของดัชนีอิทธิพล d ตามแนวคิดของโคเฮน โดยสรุปไว้ว่าโคเฮนได้สร้างกรอบในการแปลผลค่าดัชนี d เป็น ขนาดอิทธิพลขนาดเล็ก กลาง และใหญ่

ขนาดอิทธิพลขนาดเล็ก $d = 0.2$ โดยทั่วไปแล้วความแตกต่างนี้มีขนาดเล็กมากและยากที่จะพบนอกจากกลุ่มตัวอย่างขนาดใหญ่ นักวิจัยบางคนเชื่อว่าเมื่อ $d \leq 0.1$ ตัวแปรทั้งสองนั้นไม่มีความสัมพันธ์กัน

ขนาดอิทธิพลขนาดกลาง $d = 0.5$ อิทธิพลขนาดกลางมีขนาดใหญ่เพียงพอที่จะมองเห็นความแตกต่าง ซึ่งสามารถมองเห็นด้วยสายตา หรือถูกรวบรวมโดยประสบการณ์เช่นความแตกต่างระหว่างความสูงของผู้หญิงอายุ 14 ปี และ 18 ปี

ขนาดอิทธิพลขนาดใหญ่ $d = 0.8$ ความแตกต่างนี้สามารถทราบได้ เช่น ความแตกต่างของค่าเฉลี่ยเซวาร์ปัญญาาระหว่างนักศึกษามหาวิทยาลัยปี 1 และผู้ที่จบการศึกษาระดับปริญญาเอก ความแตกต่างของค่าเฉลี่ยเซวาร์ปัญญาาระหว่างนักศึกษาปริญญาโท และนักเรียนระดับมัธยมศึกษา

ในบางบริบทนักวิจัยสามารถควบคุมขนาดอิทธิพลโดยการระมัดระวังในการออกแบบงานวิจัย เช่น นักวิจัยสามารถเลือกกลุ่มหรือเงื่อนไขของสิ่งทดลองที่จะให้ค่า $\mu_1 - \mu_2$ ขนาดใหญ่

5.5 แนวทางการปรับปรุงอำนาจการทดสอบ

สตีเวน (Stevens. 2002 : 195 – 196) ได้เสนอแนวทางในการปรับปรุงอำนาจการทดสอบไว้ว่า การที่อำนาจการทดสอบต่ำโดยทั่วไปแล้วจะมีหน่วยการทดลองน้อยกว่า 20 หน่วยทดลองต่อกลุ่ม ซึ่งมี 4 วิธีการในการปรับปรุงอำนาจการทดสอบดังนี้

1. ผ่อนปรนการตั้งระดับนัยสำคัญทางสถิติ อาจจะเป็น $\alpha = 0.10$ และ $\alpha = 0.15$
2. ในการทดสอบแบบทางเดียว เมื่อมีเอกสารและงานวิจัยสนับสนุนให้ใช้การทดสอบสมมติฐานแบบทางเดียว วิธีการนี้ไม่สามารถใช้กับการทดลองพหุระดับเพราะว่าจะมีการทดสอบสองทางอยู่ถาวรแล้ว
3. ลดความแปรปรวนภายในกลุ่ม วิธีนี้จะทำให้แบบแผนมีความไวต่อการตรวจสอบพบอิทธิพลจากการกระทำทางการทดลอง คือ ใช้วิธีการสุ่มจากตัวอย่างแบบเจาะจง ซึ่งจะทำให้หน่วยทดลองมีความเป็นเอกพันธ์มาก ใช้แบบแผนแฟคทอเรียล (Factorial Design) ใช้การวิเคราะห์ความแปรปรวนร่วม (ANCOVA) ซึ่งจะช่วยลดความแปรปรวนในกลุ่ม และใช้แบบแผนแบบวัดซ้ำ (Repeated – measure Design)
4. นักวิจัยต้องทำการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องจนมั่นใจได้ว่าการจัดกระทำทางการทดลองที่นำมาศึกษานั้นมีอิทธิพลต่อตัวแปรตาม

5.6 การคำนวณเกี่ยวกับอำนาจของการทดสอบ

การคำนวณเกี่ยวกับอำนาจของการทดสอบนั้นจัดได้ 2 ลักษณะดังนี้คือ

1. คำนวณค่าอำนาจการทดสอบเมื่อทราบขนาดของกลุ่มตัวอย่าง
2. คำนวณหาขนาดของกลุ่มตัวอย่างเมื่อกำหนดอำนาจของการทดสอบ

การหาค่าอำนาจของการทดสอบนั้นหาได้โดยตรงจากการใช้ตารางอำนาจการทดสอบของ Tang (Power Function for Analysis of Variance) (Kirk. 1995 : 182 – 183) ซึ่งจะต้องทราบค่าดังต่อไปนี้

1. ดัชนีความแตกต่างระหว่างค่าเฉลี่ยของประชากร ซึ่งคำนวณได้จากสูตร(Kirk.1982 :142)

$$\phi = \sqrt{\frac{\sum_{j=1}^p \alpha_j^2 / p}{\sigma_\epsilon^2 / n}}$$

เมื่อ	ϕ	แทน	ดัชนีความแตกต่างระหว่างค่าเฉลี่ยของประชากร
	α	แทน	treatment effect
	$\sum_j^p \alpha_j^2$	แทน	ผลรวมกำลังสองของ treatment effect
	p	แทน	ระดับตัวแปรทดลอง
	n	แทน	ขนาดของกลุ่มตัวอย่าง
	σ_ϵ^2	แทน	ความคลาดเคลื่อนของความแปรปรวน (Error Variance)

และกรณีที่แต่ละกลุ่มทดลองมีขนาดใหญ่มากเกินไปคำนวณได้จากสูตร

$$\phi = \sqrt{\frac{\sum_{j=1}^p (\mu_j - \mu)^2 / p}{\sigma_\epsilon^2}}$$

2. หาค่าขั้นของความอิสระ (degrees of freedom) ระหว่างกลุ่ม $\nu_1 = p - 1$ และภายในกลุ่ม $\nu_2 = p(n - 1)$
3. ระดับความมีนัยสำคัญทางสถิติหรือโอกาสที่จะเปิดให้เกิดความคลาดเคลื่อนประเภทที่ 1 (α)

การอ่านค่าอำนาจการทดสอบจากตาราง แท็งค์ (Tang. Power Function for Analysis of Variance) (Kirk. 1995 : 816 – 823)

- เลือกตารางให้ตรงกับค่า ν_1 ซึ่งหาได้จาก $\nu_1 = p - 1$ (ระดับตัวแปรจัดกระทำ -1) ค่า ν_1 จะปรากฏอยู่มุมซ้ายด้านบนของตาราง
- เลือก α ให้เท่ากับระดับความมีนัยสำคัญของการวิจัย ดูที่ส่วนกลางของตาราง
- เลือก $\nu_2 = p(n - 1)$ ดูที่ส่วนบนของตาราง
- คำนวณค่า ϕ จากสูตรและเลือก ϕ จากแกนนอน (X) อยู่ด้านล่างของตาราง

5. อ่านค่าที่ตัดกันของ ϕ และ v_2 ในตาราง ลากเส้นขนานกับแกน Y อ่านค่าอำนาจการทดสอบ $(1 - \beta)$

5.7 แนวคิดหรือกรอบสำหรับการวิเคราะห์ค่าอำนาจการทดสอบ

โคเฮน (Cohen. 1973 : 225 – 236) ได้อธิบายส่วนประกอบสำคัญของการวิเคราะห์ค่าอำนาจการทดสอบ ดังนี้

1. การตัดสินใจ หรือระดับนัยสำคัญ (α)
2. ประสิทธิภาพความสัมพันธ์ (effect size)
3. อำนาจการทดสอบ $(1 - \beta)$
4. ขนาดกลุ่มตัวอย่าง (sample size)

การประมาณค่าอำนาจการทดสอบจากตารางของ Cohen

การประมาณค่าอำนาจการทดสอบของการทดลองสามารถแสดงได้โดยการคำนวณจากตารางของ โคเฮน (Cohen. 1988 : 289 – 354) โดยต้องคำนวณหาค่า f แทนค่า ϕ_A ซึ่งสามารถคำนวณ f^2 ได้จากสูตร

$$f^2 = \frac{\hat{\omega}_A^2}{1 - \hat{\omega}_A^2}$$

โดยหาค่า $\hat{\omega}_A^2$ ได้จากการคำนวณค่า F ในการวิเคราะห์ความแปรปรวนของระดับตัวแปรทดลอง A

$$\hat{\omega}_A^2 = \frac{(a-1)(F-1)}{(a-1)(F-1) + (a)(n)}$$

เมื่อ F คือ ค่าเอฟจากการคำนวณในการวิเคราะห์ความแปรปรวน

a คือ ระดับตัวแปรทดลอง

n คือ จำนวนกลุ่มตัวอย่าง

เมื่อได้ค่า f นำไปเปิดตารางของ Cohen ซึ่งต้องกำหนดระดับนัยสำคัญ และค่า

$$df_{num} = a - 1$$

การประมาณค่าอำนาจการทดสอบจาก Pearson – Hartley Chart

เราสามารถประมาณค่าอย่างง่ายของค่าอำนาจการทดสอบของการทดลองได้จาก Pearson – Hartley Chart โดยการประมาณค่า Omega Squared แล้วนำค่า Omega Squared ไปคำนวณหาค่า ϕ แล้วเปิดตาราง Pearson – Hartley Chart โดยหาได้จากสูตร (Keppel. 1991 : 87 -88)

$$\hat{\omega}_A^2 = \frac{(a-1)(F-1)}{(a-1)(F-1) + (a)(n)}$$

การใช้ Pearson – Hartley Chart จำเป็นต้องคำนวณค่า ϕ โดยคำนวณได้จากสูตร

$$\hat{\phi}_A^2 = n \frac{\hat{\omega}_A^2}{1 - \hat{\omega}_A^2}$$

แล้วนำค่า $\hat{\phi}_A^2$ ไปเปิดตาราง Pearson – Hartley Chart (Pearson and Hartley . 1951 : 115 – 122) โดยมี $df_{num} = a - 1 = \nu_1$ $df_{denom} = (a)(n - 1) = \nu_2$

6. สถิติทดสอบความแตกต่างของค่าเฉลี่ย

6.1 วิธีการทดสอบของ Tukey – Kramer (Kirk. 1995 : 144 – 155)

วิธีการของ Tukey – Kramer มีความแตกต่างจาก Tukey (1953) และ Kramer (1956) สำหรับกรณีที่กลุ่มตัวอย่างไม่เท่ากันและข้อตกลงพื้นฐานที่เป็นปกติ ความแปรปรวนเท่ากัน วิธีการทดสอบทางสถิตินี้แยกเป็น 2 กรณี

6.1.1 เมื่อขนาดของกลุ่มตัวอย่างแต่ละชุดเท่ากัน ($n_1 = n_2 = \dots = n_k = n_c$) โดยที่ค่าวิกฤตจะมีค่าดังนี้

$$T = q_\alpha(k, n - k) \sqrt{\frac{MSE}{n_c}}$$

เมื่อ	T	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey - Kramer
	q_α	แทน	ค่าวิกฤตของ Studentized Rang
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน
	n	แทน	ขนาดของกลุ่มตัวอย่าง
	n_c	แทน	จำนวนสมาชิกในกลุ่มตัวอย่างชุดที่ C
	k	แทน	จำนวนประชากร

6.1.2 เมื่อขนาดของกลุ่มตัวอย่างไม่เท่ากัน กรณีที่ $n_i \neq n_j$ อย่างน้อย 1 คู่

$$T = q_\alpha(k, n-k) \sqrt{\frac{MSE}{2} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

เมื่อ	T	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey - Kramer
	q_α	แทน	ค่าวิกฤตของ Studentized Rang
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน
	n	แทน	ขนาดของกลุ่มตัวอย่าง
	n_c	แทน	จำนวนสมาชิกในกลุ่มตัวอย่างชุดที่ C
	k	แทน	จำนวนประชากร

ถ้าการทดสอบสมมติฐาน $H_0 : \mu_i = \mu_j$ แบบสองทางถูกปฏิเสธ เปรียบเทียบค่า $|\bar{x}_i - \bar{x}_j|$ กับค่า T ที่ได้ถ้า $|\bar{x}_i - \bar{x}_j|$ มากกว่าหรือเท่ากับค่าวิกฤต T จะสรุปว่า $\mu_i \neq \mu_j$ ที่ระดับนัยสำคัญ α

6.2 วิธีการทดสอบของ Newman – Keuls (Kirk. 1995 : 144 – 155)

วิธีการทดสอบแบบนี้ นิวแมน (Newman) และคูลส์ (Keuls) เป็นผู้นำเสนอ มีชื่อเรียกว่าวิธีทดสอบนัยสำคัญโดยลำดับขั้น (Stairstep approach) การทดสอบนี้ใช้ค่า q เช่นเดียวกับแบบ HSD แต่แตกต่างกันที่มีการจัดเรียงลำดับขั้นของค่าเฉลี่ยที่นำมาเปรียบเทียบ ดังนั้นค่าวิกฤตที่จะเป็นเกณฑ์ในการเปรียบเทียบจะมีหลายค่า กล่าวคือค่าวิกฤตของการทดสอบจะเปลี่ยนไปตามจำนวนค่าเฉลี่ยแต่ละชุดที่เป็นคู่เปรียบเทียบตามลำดับขั้นของความมากขึ้นน้อยที่เรียงกัน หลักเกณฑ์เกี่ยวกับการทดสอบของ Newman – Keuls แบ่งออกเป็น 2 กรณี คือ

6.2.1 เมื่อขนาดของกลุ่มตัวอย่างแต่ละชุดเท่ากัน ($n_1 = n_2 = \dots = n_k = n_c$) โดยที่ค่าวิกฤตจะมีค่าดังนี้

$$NK = q_\alpha(r, n-k) \sqrt{\frac{MSE}{n_c}}$$

เมื่อ	NK	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Newman Keuls
	q	แทน	ค่าวิกฤตในตารางค่าวิกฤตของ Studentized rang
	n_c	แทน	จำนวนสมาชิกในกลุ่มตัวอย่างชุดที่ C
	k	แทน	จำนวนกลุ่มตัวอย่าง
	r	แทน	ระยะห่างของค่าเฉลี่ยที่ต้องการเปรียบเทียบ

6.2.2 เมื่อขนาดของกลุ่มตัวอย่างไม่เท่ากัน กรณีที่ $n_i \neq n_j$ อย่างน้อย 1 คู่ ; $i \neq j$ วิธีของ NK จะปฏิเสธสมมติฐาน $H_0 : \mu_i = \mu_j$ หรือสรุปว่า $\mu_i \neq \mu_j$ เมื่อ $|\bar{x}_i - \bar{x}_j| > NK$ โดยที่

$$NK = q_\alpha(r, n-k) \sqrt{\frac{MSE}{2} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

เมื่อ	NK	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Newman Keuls
	q	แทน	ค่าวิกฤตในตารางค่าวิกฤตของ Studentized rang
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน
	n	แทน	จำนวนสมาชิกในกลุ่มตัวอย่าง
	k	แทน	จำนวนกลุ่มตัวอย่าง
	r	แทน	ระยะห่างของค่าเฉลี่ยที่ต้องการเปรียบเทียบ

6.3 วิธีการทดสอบของ Duncan's New Multiple Range Test (Kirk. 1995 : 144 – 155)

วิธีของ Duncan จะคล้ายกับวิธี NK มีข้อแตกต่างกันตรงที่ q ของวิธีการแบบ Duncan's New Multiple Range Test นี้จะไม่เปิดตารางจาก Studentized Range แต่จะเปิดตารางจาก Duncan's New Multiple Range Test แทน ส่วนวิธีการเปิดตาราง และวิธีการคำนวณก็กระทำเหมือนการหาค่า q ของวิธี NK สูตรที่ใช้คำนวณ

$$W = P_\alpha(r, n-k) \sqrt{\frac{MSE}{n_c}}$$

เมื่อ	W	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Duncan
	n_c	แทน	จำนวนสมาชิกในกลุ่มตัวอย่างชุดที่ C
	k	แทน	จำนวนกลุ่มตัวอย่าง
	r	แทน	ระยะห่างของค่าเฉลี่ยที่ต้องการเปรียบเทียบ
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน
	n	แทน	จำนวนสมาชิกในกลุ่มตัวอย่าง
	P_α	แทน	ค่าที่ได้จากการเปิดตาราง

6.4 วิธีการทดสอบของ Scheffe'

การวิจัยสถานการณ์ที่เกี่ยวข้องกับการเปรียบเทียบการทดสอบที่ไม่ระบุข้อมูล เมื่อไม่สามารถระบุหรือเจาะจงการเปรียบเทียบรายคู่ได้ กระบวนการที่จะสามารถใช้ได้เมื่อพบสถานการณ์แบบนี้คือการทดสอบวิธีของ Scheffe' (1953) การทดสอบการควบคุมความคลาดเคลื่อนประเภทที่ 1 ที่น้อยกว่า α สำหรับการเปรียบเทียบจำนวนที่ไม่จำกัดระหว่างค่าเฉลี่ย $\rho \geq 3$ วิธีการของ Scheffe' ก็จะเป็นวิธีการที่มีอำนาจการทดสอบมากกว่าวิธีการของ Tukey's HSD และในกรณีที่ได้แนะนำไปว่า เมื่อเป็นการเปรียบเทียบที่ไม่ใช่รายคู่ กระบวนการของ Scheffe' จะใช้การกระจายกลุ่มตัวอย่างแบบ F ที่เหมือนกับการวิเคราะห์ความแปรปรวน และมีความแกร่งที่ไม่ปกติ และกระบวนการดังกล่าวนี้ สามารถนำไปใช้กับกลุ่มตัวอย่างที่ไม่เท่ากันได้ วิธีการทดสอบของ Scheffe' มีสูตรดังนี้

$$F_{\alpha(\bar{X}_1, \bar{X}_2)} = \frac{(\bar{X}_1 - \bar{X}_2)^2}{MS_w \left(\frac{1}{n_1} + \frac{1}{n_2} \right) (k-1)} \quad (\text{ในกรณีที่เปรียบเทียบ } \bar{X}_1 \text{ กับ } \bar{X}_2)$$

เมื่อ	df = k-1, n-k	
F	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Scheffe'
k	แทน	จำนวนกลุ่มตัวอย่าง
n	แทน	จำนวนตัวอย่างทั้งหมด
n_1, n_2	แทน	จำนวนกลุ่มตัวอย่างในกลุ่มที่ 1 และกลุ่มที่ 2

กระบวนการของ Scheffe' ส่วนมากจะใช้กับการวิเคราะห์ความแปรปรวนแบบ F ถ้าการทดสอบแบบ F มีนัยสำคัญทางสถิติ ก็จะมีการเปรียบเทียบระหว่างค่าเฉลี่ยอย่างน้อยหนึ่งคู่ที่มีนัยสำคัญทางสถิติ ตามวิธีการของ Scheffe'

กระบวนการของ Scheffe' จะเป็นกระบวนการที่มีความซับซ้อนของข้อมูลมากที่สุด ความซับซ้อนที่ว่่านี ก็คือ ข้อมูลที่มีอำนาจการทดสอบน้อย กระบวนการนี้จะใช้สำหรับสมมติฐานที่เกี่ยวข้องกับการเปรียบเทียบที่ไม่ใช่รายคู่เท่านั้น

6.5 วิธีการทดสอบของ Fisher's Least Significant Difference (LSD)

Fisher ได้พัฒนาวิธีการเปรียบเทียบผลต่างระหว่างค่าเฉลี่ยประชากรครั้งละคู่ โดยต้องสุ่มตัวอย่างแต่ละชุดเป็นอิสระกัน ซึ่งเรียกว่า Fisher's LSD (Kirk. 1995 : 144 – 155) มีสูตรดังนี้

$$LSD = t_{1-\frac{\alpha}{2}; n-k} \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \quad \text{โดยที่ } n = \sum_{i=1}^k n_i$$

เมื่อ	LSD	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Fisher's LSD
	k	แทน	จำนวนกลุ่มตัวอย่าง
	n	แทน	จำนวนตัวอย่างทั้งหมด
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน

ผู้วิจัยได้เลือกใช้สถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี สรุปได้ดังนี้

ตาราง 1 แสดงการเปรียบเทียบขนาดของกลุ่มตัวอย่างและค่าวิกฤตของสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี

สถิติทดสอบ	ขนาด ของกลุ่มตัวอย่าง	ค่าวิกฤต
Tukey - Kramer	กลุ่มตัวอย่างเท่ากัน หรือใกล้เคียงกัน	กรณีขนาดกลุ่มตัวอย่างเท่ากัน $T = q_\alpha(k, n - k) \sqrt{\frac{MSE}{n_c}}$ กรณีขนาดกลุ่มตัวอย่างไม่เท่ากัน $T = q_\alpha(k, n - k) \sqrt{\frac{MSE}{2} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$
Newman Keuls	กลุ่มตัวอย่างเท่ากัน หรือใกล้เคียงกัน	กรณีขนาดกลุ่มตัวอย่างเท่ากัน $NK = q_\alpha(r, n - k) \sqrt{\frac{MSE}{n_c}}$ กรณีขนาดกลุ่มตัวอย่างไม่เท่ากัน $NK = q_\alpha(r, n - k) \sqrt{\frac{MSE}{2} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$
Duncan	กลุ่มตัวอย่างเท่ากัน	$W = P_\alpha(r, n - k) \sqrt{\frac{MSE}{n_c}}$
Scheffe '	กลุ่มตัวอย่างเท่ากัน หรือใกล้เคียงกัน	$F_{\alpha(\bar{X}_1, \bar{X}_2)} = \frac{(\bar{X}_1 - \bar{X}_2)^2}{MS_w \left(\frac{1}{n_1} + \frac{1}{n_2} \right) (k - 1)}$
Fisher's LSD	กลุ่มตัวอย่างเท่ากัน หรือใกล้เคียงกัน	$LSD = t_{1-\frac{\alpha}{2}; n-k} \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$

ตาราง 2 แสดงการเปรียบเทียบวิธีการเปรียบเทียบและขั้นตอนการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี

สถิติทดสอบ	วิธีการเปรียบเทียบ	ขั้นตอนการทดสอบ
Tukey - Kramer	ทีละคู่	ขั้นที่ 1 ทำการทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ ถ้าปฏิเสธ H_0 ทำขั้นที่ 2 ต่อ ขั้นที่ 2 เรียงลำดับค่าเฉลี่ยจากน้อยไปหามาก ขั้นที่ 3 หาค่าวิกฤต T ขั้นที่ 4 หาค่า $\bar{X}_i - \bar{X}_j$ ทุกค่า $i \neq j; i, j = 1, 2, \dots, k$ ขั้นที่ 5 ถ้า $\bar{X}_i - \bar{X}_j \geq T$ หมายถึง $\mu_i \text{ แตกต่างจาก } \mu_j$
Newman Keuls	หลายคู่พร้อมกัน	ขั้นที่ 1 ทำการทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ ถ้าปฏิเสธ H_0 ทำขั้นที่ 2 ต่อ ขั้นที่ 2 เรียงลำดับค่าเฉลี่ยจากน้อยไปหามาก ขั้นที่ 3 หาค่าวิกฤต ขั้นที่ 4 หาค่า $\bar{X}_i - \bar{X}_j$ ทุกค่า $i \neq j; i, j = 1, 2, \dots, k$ ขั้นที่ 5 ถ้า $\bar{X}_i - \bar{X}_j \geq NK$ หมายถึง $\mu_i \text{ แตกต่างจาก } \mu_j$
Duncan	หลายคู่พร้อมกัน	ขั้นที่ 1 ทำการทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ ถ้าปฏิเสธ H_0 ทำขั้นที่ 2 ต่อ ขั้นที่ 2 หาค่าวิกฤต W ขั้นที่ 3 หาค่า $\bar{X}_i - \bar{X}_j$ ทุกค่า $i \neq j; i, j = 1, 2, \dots, k$ ขั้นที่ 4 ถ้า $\bar{X}_i - \bar{X}_j \geq W$ หมายถึง $\mu_i \text{ แตกต่างจาก } \mu_j$

ตาราง 2(ต่อ)

สถิติทดสอบ	วิธีการเปรียบเทียบ	ขั้นตอนการทดสอบ
Scheffe'	- หลายคู่พร้อมกัน - ทดสอบ Linear Contrasts	ขั้นที่ 1 ทำการทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ ถ้าปฏิเสธ H_0 ทำขั้นที่ 2 ต่อ ขั้นที่ 2 สร้าง 1-WAY ANOVA ได้ค่า $S^2 = \text{MSE}$ ขั้นที่ 3 หาค่าวิกฤต F ขั้นที่ 4 หาค่า $\bar{X}_i - \bar{X}_j$ ทุกค่า $i \neq j; i, j = 1, 2, \dots, k$ ขั้นที่ 5 ถ้า $\bar{X}_i - \bar{X}_j \geq F$ หมายถึง $\mu_i \text{ แตกต่างจาก } \mu_j$
Fisher's LSD	ทีละคู่	ขั้นที่ 1 ทำการทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ ถ้าปฏิเสธ H_0 ทำขั้นที่ 2 ต่อ ขั้นที่ 2 หาค่าวิกฤต LSD ขั้นที่ 3 หาค่า $\bar{X}_i - \bar{X}_j$ ทุกค่า $i \neq j; i, j = 1, 2, \dots, k$ ขั้นที่ 4 ถ้า $\bar{X}_i - \bar{X}_j \geq \text{LSD}$ หมายถึง $\mu_i \text{ แตกต่างจาก } \mu_j$

7. ความถนัดและแบบทดสอบวัดความถนัด

คำว่า “ความถนัด” ซึ่งในภาษาอังกฤษใช้คำว่า Aptitude นี้ ในวงการวัดผลปัจจุบัน หมายถึง สมรรถวิสัยและทิศทางแห่งความงอกงามของสมอง หรือจะกล่าวให้ง่ายขึ้นก็หมายถึง ขีดระดับความสามารถขั้นสูงสุดของบุคคลที่เขาอาจมีได้ต่อการเรียนรู้และการฝึกฝนในวิทยาการและทักษะต่างๆ ถ้าหากเขาได้รับการฝึกสอนและประสบการณ์ที่เหมาะสม ถ้าจะขยายความนี้ให้กว้างตามความเข้าใจของบุคคลทั่วไปแล้ว ความถนัดก็มีส่วนคล้ายกับความฉลาดของแต่ละบุคคล (บทความพิเศษของ ขวาล แพรัตกุล : 2517) ซึ่งนักการศึกษาได้ให้ความหมายแตกต่างกันดังนี้

เพ็ญศรี เตมียกุล (2525 : 123) ได้ให้ความหมายไว้ว่า ความถนัดหมายถึง ความสามารถของบุคคลเมื่อได้ใช้เวลาในการฝึกฝนเรียนรู้ ทำให้พัฒนาความรู้และทักษะบางอย่างหรือการจัดระบบการตอบสนอง ซึ่งแต่ละคนจะมีขีดระดับความสามารถสูงสุดแตกต่างกันไป

ล้วน สายยศและและอังคณา สายยศ (2541 : 15) ได้ให้ความหมายไว้ว่า ความถนัด (aptitude) เป็นความต่อเนื่องจากเชาวน์ปัญญา (intelligence) มองความสามารถที่ได้จากการฝึกฝน ซึ่งสมมาตั้งแต่เกิดจนถึงปัจจุบัน บางทีเรียกว่า เป็นประสบการณ์ในการเรียนรู้หรือความสามารถเฉพาะด้าน

จิตติรัตน์ ฐานวัฒน์ (2542 : 7) ได้ให้ความหมายไว้ว่า ความถนัดหมายถึง ความสามารถทางสมองในด้านต่างๆ ของแต่ละบุคคล มีลักษณะเป็นกลุ่มของความสามารถเฉพาะด้านที่เกิดจากการเรียนรู้การฝึกฝนและสะสมไว้เป็นประสบการณ์ของตน ซึ่งทำให้แต่ละบุคคลสามารถแสดงความสามารถเฉพาะด้านของตนออกมาได้อย่างมีประสิทธิภาพ ซึ่งลักษณะดังกล่าวสามารถเป็นตัวทำนายความสำเร็จทางการเรียนและการงานในอนาคตของบุคคลผู้นั้นได้

ในการวิจัยครั้งนี้ผู้วิจัยเลือกใช้แบบทดสอบความถนัดทางภาษา ซึ่งล้วน สายยศและอังคณา สายยศ (2538 : 177) ได้ให้ความหมายของแบบทดสอบความถนัดไว้ว่า เป็นแบบทดสอบที่วัดสมรรถภาพของมนุษย์ว่ามีขีดความสามารถในการเรียนรู้มากเท่าใด หรือวัดกำลังภายในสมองของมนุษย์ และได้กล่าวถึงการวัดความถนัดทางภาษา ประกอบด้วย 7 รูปแบบ ดังนี้ (2541 : 79 – 92)

1. แบบทดสอบคำตรงข้าม
2. แบบทดสอบคำที่มีความหมายใกล้เคียง
3. แบบทดสอบศัพท์สัมพันธ์
4. แบบทดสอบแบบผิดความ
5. แบบทดสอบความเข้าใจภาษา
6. แบบทดสอบความเข้าใจภาพ
7. แบบทดสอบการสังเคราะห์ข้อความ

ผู้วิจัยได้เลือกมาประกอบการวิจัยด้วยกัน 4 รูปแบบ ดังนี้คือ

1. แบบทดสอบแบบศัพท์สัมพันธ์
2. แบบทดสอบแบบผิดความ
3. แบบทดสอบความเข้าใจภาษา
4. แบบทดสอบความเข้าใจภาพ

ประโยชน์ของแบบทดสอบความถนัด (ล้วน สายยศและอังคณา สายยศ .2541: 22 –

27) สรุปได้ดังนี้

1. เพื่อการสอบคัดเลือก
2. เพื่อการแยกประเภทนักเรียน
3. เพื่อวินิจฉัยความสามารถ
4. เพื่อพยากรณ์ความสำเร็จ
5. เพื่อวัดพัฒนาการ
6. เหมาะสำหรับเปรียบเทียบสติกับปัญญา
7. เพื่อประเมินผลการศึกษา
8. เพื่อการวิจัย

สรุปได้ว่า ความถนัด หมายถึง ความสามารถของแต่ละบุคคลที่ได้รับการฝึกฝนเรียนรู้ในเรื่องนั้น ทำให้เกิดพัฒนาการทักษะการเรียนรู้เฉพาะด้านนั้นๆ ในการนำไปปฏิบัติหรือเรียกอีกอย่างว่าความสามารถเฉพาะด้าน เช่น ความถนัดทางภาษา หมายถึงความสามารถในการสื่อความหมาย เข้าใจความหมาย สามารถแปลความตีความ ขยายความและวิเคราะห์ความหมายจากคำ วลี ประโยค ข้อความประเภทต่างๆ หรือตามที่กำหนดให้ได้อย่างคล่องแคล่วและสมเหตุสมผล

8. งานวิจัยที่เกี่ยวข้อง

งานวิจัยต่างประเทศ

แรมเซย์ (Ramsey. 1978 : 479 – 485) ได้ศึกษาเปรียบเทียบอำนาจการทดสอบระหว่างการเปรียบเทียบพหุคูณแบบรายคู่ศึกษาถึง Experimentwise Type I Error Rate โดยวิธีเทคนิคอนติคาร์โล ผลการศึกษาพบว่า วิธี Peritz's F เป็นวิธีที่ดีที่สุดวิธีที่นำมาศึกษา ซึ่งมีอำนาจการทดสอบมากกว่าวิธีของ Tukey

ดักลาส และสตีเฟน (Douglas & Stephen.1978 : 265 –282) ได้ศึกษาเรื่องการเปรียบเทียบภายหลังสำหรับการทดสอบนอนพาราเมตริก การเปรียบเทียบพหุคูณแบบภายหลังใช้ในการประมาณความแตกต่างของประชากรหลายกลุ่มได้กำหนดสูตรสำหรับการทดสอบสถิติแบบนอนพาราเมตริกไว้ 3 ทางเลือก การทดสอบทั้ง 3 ของสถิตินอนพาราเมตริกอ้างอิงมาจาก Puri K-sample ซึ่งขยายมาจากการทดสอบของ Siegel – Tukey, Mood และ Klotz ทฤษฎีแวดล้อมในการพัฒนาวิธีการเปรียบเทียบภายหลังทั้ง 3 เป็นแนวทางเลือกพื้นฐานของไค-สแควร์ตามทฤษฎีของ Scheffe' การนำเสนอเป็นการแสดงเพื่อให้เห็นการประยุกต์ ถึงวิธีการเปรียบเทียบพหุคูณแบบเปรียบเทียบภายหลัง ผู้วิจัยเลือกใช้สำหรับสถิตินอนพาราเมตริก

เบลล์ และเจมส์ (Blair & James. 1980 : 309 – 335) ได้ศึกษาเรื่องการเปรียบเทียบอำนาจการทดสอบของผลรวมสถิติของ Wilcoxon ที่วิธี Student's statistic ภายใต้การแจกแจงที่ไม่ปกติหลายรูปแบบ ในการศึกษาค้นคว้านี้ได้ใช้เทคนิคอนติคาร์โลในการสมมติข้อมูลเพื่อการเปรียบเทียบค่าเฉลี่ยที่เป็นอิสระสองกลุ่มโดยเป็นการสร้างรูปแบบการแจกแจงข้อมูล 6 รูปแบบดังนี้ (1) uniform (2) laplace (3) half-normal (4) exponential (5) mix-normal และ (6) mix-uniform ขนาดของกลุ่มตัวอย่างที่นำมาศึกษาคือ $(n_1, n_2) = (3,9), (6,6), (9,27), (18,18), (27,18)$ และ $(54,54)$ ผลการศึกษาพบว่า วิธีการ Wilcoxon นั้นมีอำนาจการทดสอบมากกว่าวิธี Student's t statistic และผลที่ได้รับจากกลุ่มตัวอย่างขนาดเล็กแตกต่างจากกลุ่มตัวอย่างที่มีขนาดใหญ่กว่า

ลี่ ยู ฟาง (Li Yu – Fang. 1998 : 6489) ได้ศึกษาพัฒนาอำนาจการประมาณและขนาดกลุ่มตัวอย่างของวิธีการเปรียบเทียบพหุคูณ 7 วิธี ผลการศึกษาพบว่า Fisher' LSD และวิธี Duncan multiple range test เป็นวิธีทดสอบที่ดีที่สุดใน 7 วิธีที่นำมาศึกษา ขณะที่ Bonferoni และ Sidak เป็นสถิติทดสอบที่มีอำนาจน้อยที่สุด

วิลคอกซ์ (Wilcox. 1989 : 269 – 278) ได้ศึกษาเรื่องการประมาณค่าสำหรับความแปรปรวนที่ไม่เท่ากันเมื่อมีการเปรียบเทียบค่าเฉลี่ยในการวิเคราะห์ความแปรปรวนแบบทางเดียวและแบบสองทาง มีเอกสารมากมายที่แสดงถึงการไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบค่า F คือค่าความแปรปรวนไม่เท่ากัน ทำให้ไม่มีความแกร่งในการวิเคราะห์ความแปรปรวน ในการศึกษาครั้งนี้ Wilcox ได้อธิบายและเปรียบเทียบ 2 วิธีการเปรียบเทียบในกรณีค่าความแปรปรวนแตกต่างกันของการวิเคราะห์ความแปรปรวนแบบสองทาง วิธีการหนึ่งมีพื้นฐานมาจากการปรับปรุงวิธีของ Wilcox (1988) ที่สำหรับการวิเคราะห์ความแปรปรวนแบบทางเดียวซึ่งเป็นรูปแบบพื้นฐานสำหรับการพิจารณาวิธีนี้ เพื่อการวิเคราะห์ความแปรปรวนแบบสองทาง อีกวิธีหนึ่งเป็นการขยายผลการใช้วิธีของ James ผลการศึกษาพบว่า วิธีที่มีรากฐานมาจากวิธีของ Wilcox นั้นสามารถนำไปใช้ได้ง่ายกว่าและมีประสิทธิภาพมากกว่า วิธีที่นำมาจากวิธีของ James

เจฟเฟอเรีย และมิเชลลา (Jefferey & Michela. 1995 : 343 – 362) ได้ศึกษาอำนาจการทดสอบและอัตราความคลาดเคลื่อนประเภทที่ 1 กับวิธีการเปรียบเทียบพหุคูณแบบรายคู่ ภายใต้อำนาจการทดสอบที่เท่ากัน โดยศึกษาจากวิธีการเปรียบเทียบพหุคูณ 9 วิธี ผลการศึกษาพบว่า วิธี Dun และวิธีที่พัฒนามาจากวิธีของ Bonferroni มีความแข็งแกร่งดีกว่าวิธีของ Tukey วิธีของ Peritz วิธีของ Ryan และวิธีของ Fisher – Hayter มีอำนาจการทดสอบมากที่สุด

งานวิจัยภายในประเทศ

วิไลลักษณ์ องค์จิระวุฒิ (2522 : บทคัดย่อ) ได้ศึกษาเรื่องการเปรียบเทียบวิธีการต่างๆ ที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากร โดยพิจารณาจากความผิด 3 ชนิด ในการวิเคราะห์เพื่อเปรียบเทียบค่าเฉลี่ยของประชากรหลายๆ ประชากรว่ามีความแตกต่างกันอย่างมีนัยสำคัญหรือไม่นั้น มีระเบียบวิธีในการทดสอบหลายวิธีด้วยกันซึ่งแต่ละวิธีก็มีความยากง่ายและความสะดวกในการนำไปใช้ต่างกัน ในการศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อประเมินว่า วิธีใดที่จะให้ผลสรุปแม่นยำและถูกต้องในการทดสอบมากกว่าวิธีอื่น โดยพิจารณาจากความคลาดเคลื่อนทั้ง 3 ประเภท ที่เกิดขึ้นในการทดสอบแต่ละวิธีว่าวิธีใดเกิดความคลาดเคลื่อนน้อยที่สุด ความคลาดเคลื่อนทั้ง 3 ประเภทได้แก่ ความคลาดเคลื่อนประเภทที่ 1 (Type I Error) ความคลาดเคลื่อนประเภทที่ 2 (Type II Error) และความคลาดเคลื่อนประเภทที่ 3 (Type III Error) ซึ่งเป็นความคลาดเคลื่อนจากการสรุปผลความแตกต่างระหว่างค่าเฉลี่ยของประชากรตรงข้ามกับความเป็นจริง อันเป็นผลเนื่องมาจากตัวอย่างที่ใช้ ไม่เป็นตัวแทนที่ดีของประชากร ในการศึกษาครั้งนี้ได้นำข้อมูลมาวิเคราะห์ในลักษณะเป็นข้อมูลแจกแจงทางเดียวและข้อมูลแจกแจงสองทาง และทำการทดสอบความมีนัยสำคัญของความแตกต่างระหว่างค่าเฉลี่ยของประชากรโดยใช้ระเบียบวิธีการต่างๆ แล้วนำผลที่ได้จากการทดสอบทั้งหมดทุกวิธีมาเปรียบเทียบกัน ซึ่งผลที่ได้จากการสรุปได้ว่า ในการทดสอบความมีนัยสำคัญของความแตกต่างระหว่างค่าเฉลี่ยประชากร 2 ประชากร วิธีของ Murphy gap LSD สามารถให้ผลสรุปหรือผลการทดสอบได้ดีกว่าวิธีอื่น ทั้งยังเป็นวิธีที่มีความสะดวกในการนำไปใช้พอสมควรอีกด้วย ในกรณีที่ผู้วิจัยมีเวลาจำกัด ในการทดสอบนัยสำคัญของความแตกต่างระหว่างค่าเฉลี่ยของประชากรคู่ใดๆ วิธีทดสอบโดยใช้ Least Significant Difference (LSD), Tukey's W-Procedure และ Scheffe's method เป็นวิธีที่สะดวกและง่ายในการนำไปใช้ เพราะเป็นการคำนวณค่าสถิติเพียงค่าเดียวในการใช้เป็นตัวตัดสินความมีนัยสำคัญของความแตกต่างระหว่างค่าเฉลี่ยของประชากร 2 ประชากรใดๆ แต่ผลที่ได้จากการศึกษาครั้งนี้สามารถสรุปได้ว่า โดยทั่วไปวิธีของ Least Significant Difference test ให้ผลสรุปหรือผลการทดสอบที่น่าเชื่อถือได้ดีกว่าอีก 2 วิธีที่กล่าวมาแล้ว

นันทา วงษ์วิโรจน์ (2532 : บทคัดย่อ) ได้ศึกษาเรื่องสถิติทดสอบที่มีความแข็งแกร่งสำหรับทดสอบความเท่ากันของค่าเฉลี่ยประชากร การศึกษาครั้งนี้มีวัตถุประสงค์ที่จะศึกษาเปรียบเทียบสถิติทดสอบที่มีความแข็งแกร่งสำหรับทดสอบความเท่ากันของค่าเฉลี่ยประชากรโดยใช้สถิติทดสอบ 3 วิธี คือ สถิติทดสอบแบบ ANOVA F-TEST และสถิติทดสอบแบบ Trimmed W และสถิติทดสอบแบบ Trimmed F โดยจะศึกษาถึงความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบทั้ง 3 วิธี ในที่นี้ศึกษากรณีที่ประชากรทั้ง 3 ชุดมีการแจกแจงแบบเดียวกัน คือการแจกแจงแบบปกติ และการแจกแจงแบบปกติปลอมปน ขนาดตัวอย่างที่ใช้แบ่งเป็น 2 กรณี คือขนาดตัวอย่างเท่ากันและไม่เท่ากัน สำหรับข้อมูลที่ใช้ในการทดลองครั้งนี้จำลองด้วยโปรแกรมคอมพิวเตอร์ โดยใช้เทคนิคมอนติคาร์โล ซึ่งกระทำซ้ำกันหนึ่งพันครั้งในแต่ละกรณี ผลการศึกษาสรุปได้ว่า ความแข็งแกร่งของการทดสอบ โดยพิจารณาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 โดยใช้เกณฑ์ในการพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 (Cochran; & Cox.1957) คือ ถ้าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 อยู่ในช่วง (0.007, 0.015) ที่ระดับนัยสำคัญ 0.01 และถ้าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 อยู่ในช่วง (0.04, 0.06) ที่ระดับนัยสำคัญ 0.05 จะถือว่าการทดสอบนี้สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ พบว่าการแจกแจงประชากร ขนาดตัวอย่าง อัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน เปอร์เซ็นต์การตัดข้อมูลที่ปลายหางของการแจกแจง และระดับนัยสำคัญมีผลต่อความแข็งแกร่งของการทดสอบสูงสุดในทุกขนาดตัวอย่างและระดับนัยสำคัญที่ศึกษา แต่เมื่อประชากรมีการแจกแจงแบบปกติปลอมปนพบว่า สถิติทดสอบแบบ Trimmed W และ Trimmed F จะให้ค่าอำนาจการทดสอบสูงกว่าสถิติทดสอบแบบ ANOVA F-TEST ซึ่งการเลือกใช้สถิติแบบ Trimmed W และ Trimmed F นั้นพบว่าความสัมพันธ์ระหว่างอัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน อัตราส่วนของค่าเฉลี่ย และเปอร์เซ็นต์การตัดข้อมูลที่ปลายหางการแจกแจง จะมีผลต่ออำนาจการทดสอบทางสถิติทดสอบทั้ง 2 นอกจากนี้พบว่า ปัจจัยที่มีผลต่อค่าอำนาจการทดสอบของค่าสถิติทดสอบทั้ง 3 มากที่สุดเมื่อกำหนดให้ปัจจัยอื่นคงที่คือ อัตราส่วนของส่วนเบี่ยงเบนมาตรฐานของประชากร ส่วนปัจจัยที่มีผลกระทบรองลงมาคือ ลักษณะการแจกแจงของประชากร ขนาดตัวอย่าง เปอร์เซ็นต์การตัดข้อมูลที่ปลายหางของการแจกแจง ที่ระดับนัยสำคัญตามลำดับ

นุปผา จงเลิศตระกูล (2532 : บทคัดย่อ) การเปรียบเทียบค่าเฉลี่ยของประชากรนั้นโดยทั่วไปจะใช้วิธีการวิเคราะห์ความแปรปรวน เมื่อทำการปฏิเสธสมมติฐานกลางแล้ว อาจจะมีการเปรียบเทียบเชิงพหุหรือหาช่วงความเชื่อมั่นร่วม โดยงานวิจัยนี้ได้เสนอวิธีการเลือกประชากรที่ดี เมื่อประชากรมีการแจกแจงแบบเอกซ์โพเนนเชียล เมื่อทำการเปรียบเทียบพารามิเตอร์ของ K ประชากร

กับประชากรควบคุมหรือประชากรมาตรฐานตามค่าพารามิเตอร์แสดงตำแหน่งที่ไม่ทราบค่าได้แบ่งการพิจารณา เป็นดังนี้

การเลือกประชากรที่มีค่าพารามิเตอร์แสดงตำแหน่งมากกว่าหรือน้อยกว่าพารามิเตอร์แสดงตำแหน่งของประชากรควบคุม เป็นประชากรที่ดีเมื่อส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากันในทุกๆ ประชากรแต่ไม่ทราบค่าและขนาดตัวอย่างของแต่ละประชากรเท่ากันเป็นค่าที่กำหนดให้ แบ่งเป็น 2 กรณีคือ กรณีข้อมูลทั่วไปและกรณีข้อมูล เช่น เซอร์ประเภทที่ 2 ค่า lower และ upper percentage points ที่คำนวณได้จากทั้ง 2 กรณี แสดงในตารางจำนวนประชากร (k) ขนาดตัวอย่าง (n) ของตำแหน่งความเป็นอิสระ (v) และสัมประสิทธิ์เชื่อมั่นร่วม (p^*) เมื่อกำหนดให้ $k = 2, 3, \dots, a$ $v = 3, 4, \dots, a$ $p^* = 0.90, 0.99$ ซึ่งจะนำผลที่คำนวณได้ไปใช้ในการหาช่วงความเชื่อมั่นร่วมขอบเขตล่าง และขอบเขตบนแต่ละกรณีเพื่อนำมาใช้เป็นเกณฑ์การตัดสินใจเลือกประชากรที่ดี การเลือกประชากรที่มีค่าพารามิเตอร์แสดงตำแหน่งมากกว่าพารามิเตอร์แสดงตำแหน่งของประชากรมาตรฐาน เป็นประชากรที่ดีเมื่อส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากันในทุกๆ ประชากร คือมีค่าเท่ากับ 1 นั้น เพื่อศึกษาประสิทธิภาพของวิธีอนุบรรพโดยการเปรียบเทียบวิธีอนุบรรพกับวิธีทั่วไป ซึ่งพิจารณาจากขนาดตัวอย่างโดยเฉลี่ยของวิธีอนุบรรพและขนาดตัวอย่างที่กำหนดไว้คงที่โดยวิธีทั่วไป ผลจากการเปรียบเทียบขนาดตัวอย่างของวิธีการทั้ง 2 นี้ขนาดตัวอย่างที่ใช้เพื่อการตัดสินใจเลือกประชากรที่ดี โดยวิธีอนุบรรพจะมีขนาดเล็กกว่าขนาดตัวอย่างที่กำหนดไว้คงที่โดยวิธีทั่วไปสำหรับค่า a, B และสามเหลี่ยม แต่ละชุดที่กำหนดให้ คือ 0.01, 0.05, .010 และ 0.20, 0.50, 1.00 ตามลำดับ

เลิศลักษณ์ กลิ่นหอม (2532 : 316 – 318) ได้ทำการศึกษาความคลาดเคลื่อนชนิดที่ 1 และอำนาจการทดสอบ ในแบบแผนการทดลองแบบสุ่มภายในบล็อกที่มีปฏิสัมพันธ์ระหว่างสิ่งทดลองกับบล็อก พบว่า เมื่อปฏิสัมพันธ์ระหว่างสิ่งทดลองกับบล็อกมีขนาดใหญ่ขึ้นอำนาจการทดสอบจะลดลง และเมื่อจำนวนสิ่งทดลอง จำนวนบล็อก และค่าสหสัมพันธ์ระหว่างตัวแปรบล็อกกับตัวแปรตามเพิ่มขึ้นอำนาจการทดสอบจะเพิ่มขึ้นในการแจกแจงประชากรทั้ง 3 ลักษณะ ยกเว้นเมื่อสิ่งทดลองเพิ่มขึ้นในการแจกแจงแบบปกติ อำนาจการทดสอบจะลดลง ในกรณี RBD มีปฏิสัมพันธ์ระหว่างสิ่งทดลองและบล็อก การแปลงข้อมูลเป็นสแควร์รุธ และลอกการิทึมนั้น ค่าความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 และอำนาจของการทดสอบจะแตกต่างไปจากเกณฑ์เกือบทั้งหมด

พหล ศักดิ์คะทศน์ (2534 : บทคัดย่อ) ได้ศึกษาเรื่อง การเปรียบเทียบประสิทธิภาพของวิธีการทดสอบความแตกต่างของค่าเฉลี่ยในการวิเคราะห์ความแปรปรวน ในการศึกษารังนี้มีจุดประสงค์เพื่อเปรียบเทียบประสิทธิภาพการทดสอบของวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ย

3 วิธี คือ วิธี Unrestricted LSD (U-LSD) วิธี Bonferoni (Dunn) T-test (Bon) และวิธี Murphys Gap LSD (MG-LSD) โดยข้อมูลมีการแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 ทำการทดลองด้วยเทคนิคอนติคาร์โล โดยจำลองการทดสอบด้วยเครื่องคอมพิวเตอร์จำนวน 100 รอบ สำหรับแต่ละสถานการณ์ในการคำนวณหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 โดยใช้เกณฑ์ของ Braley กล่าวคือความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ต้องอยู่ในช่วง (0.05, 0.015) ที่ระดับนัยสำคัญที่ 0.01 และช่วง (0.025, 0.075) ที่ระดับนัยสำคัญที่ 0.05 จึงถือว่าสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบได้ โดยมีจำนวนสิ่งทดลองเท่ากับ 2 ถึง 10 ทั้งแผนการทดลองสุ่มสมบูรณ์ และแผนการทดลองแบบบล็อกสุ่มสมบูรณ์ และจำนวนบล็อกมี 3 ระดับ คือ 5, 7 และ 10 ในแผนการทดลองแบบบล็อกสุ่มสมบูรณ์ ผลการศึกษาพบว่า Murphys Gap LSD เหมาะสำหรับแผนการทดลองแบบสุ่มสมบูรณ์ เมื่อจำนวนสิ่งทดลองเท่ากับ 6 – 10 แต่เมื่อแผนการทดลองเป็นแบบบล็อกสุ่มสมบูรณ์กรณีที่สิ่งทดลองเท่ากับ 9 – 10 ปรากฏว่าทั้ง 3 วิธีดังกล่าวไม่มีวิธีใดที่เหมาะสมที่จะใช้ในการทดสอบ

วิชชุดา ศรีโสภา (2539 :บทคัดย่อ) ได้ศึกษาเรื่องการศึกษาเปรียบเทียบหุ การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยในแผนการทดลองแบบสุ่มสมบูรณ์ 4 วิธี คือ วิธี Unrestricted LSD , Murphys Gap LSD , Tukey's (H) และ Murphys Gap Unrestricted LSD และ 30% มีจำนวนสิ่งทดลองเท่ากับ 3,4,5,6 และ 10 จำนวนซ้ำในแต่ละสิ่งทดลองมีจำนวนไม่เท่ากัน 3 ระดับคือ เพิ่มทีละ 3 เพิ่มทีละ 5 เพิ่มทีละ 10 โดยมีจำนวนซ้ำเริ่มต้นเป็น 5 ทำการทดลองด้วยเทคนิคอนติคาร์โล โดยจำลองการทดลองด้วยเครื่องคอมพิวเตอร์จำนวน 1000 รอบ สำหรับแต่ละสถานการณ์ที่กำหนดในการคำนวณหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ซึ่งใช้เกณฑ์ในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของ Bradley และจะทำการหาอำนาจการทดสอบเมื่อวิธีการนับผ่านเกณฑ์ควบคุมความคลาดเคลื่อนประเภทที่ 1 ผลการศึกษาพบว่าวิธีการ Unrestricted LSD จะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีที่สุดและได้ค่าอำนาจการทดสอบสูงที่สุดในทุกสถานการณ์ที่ทำการทดลอง

สุกัญญา หนูกล้า (2542: 106 – 107) ทำการศึกษาอำนาจการทดสอบของตัวสถิติทดสอบค่าเฉลี่ยเมื่อประชากรมีการแจกแจงแบบเบ้ขวา พบว่าถ้าระดับนัยสำคัญทางสถิติ และระดับความโด่งเพิ่มขึ้นอำนาจการทดสอบของตัวสถิติทดสอบค่าเฉลี่ย คือสถิติทดสอบที สถิติทดสอบทีของจอห์นสัน สถิติทดสอบทีดัดแปลงของจอห์นสัน มีค่าเพิ่มขึ้น แต่เมื่อความเบ้เพิ่มขึ้นอำนาจการทดสอบของสถิติทดสอบทั้ง 3 ชนิดดังกล่าวมีค่าลดลง

ศิริพร กิตติสุภาพ (2544 : บทคัดย่อ) ได้ศึกษาเรื่องผลกระทบของการแจกแจงข้อมูลที่มีต่อการทดสอบความแตกต่างของค่าเฉลี่ยโดยใช้วิธี Tukey' s HSD , วิธี Scheffe ' และวิธี Brown – Forsythe จากการแจกแจงข้อมูล 5 ลักษณะ คือ การแจกแจงข้อมูลแบบโค้งปกติ การแจกแจงของข้อมูลเป็นโค้งเบ้ทางขวา การแจกแจงข้อมูลเป็นโค้งเบ้ทางซ้าย การแจกแจงของข้อมูลเป็นแบบสูงแหลม และการแจกแจงของข้อมูลเป็นแบบแบนราบ ผลการศึกษาพบว่า 1) ผลการเปรียบเทียบพหุคูณสามวิธีคือวิธี Tukey' s HSD , วิธี Scheffe ' และวิธี Brown – Forsythe ในการแจกแจงข้อมูลแบบเดียวกัน แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 2) ผลการเปรียบเทียบพหุคูณสามวิธีคือวิธี Tukey' s HSD , วิธี Scheffe ' และวิธี Brown – Forsythe ที่ทดสอบด้วยพหุคูณเดียวกัน ในการแจกแจงของข้อมูลต่างกัน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

สิรินุช เอี่ยมเขียว (2544 : บทคัดย่อ) ได้ศึกษาเรื่องผลการเปรียบเทียบพหุคูณรายคู่ ภายใต้ความแปรปรวนที่แตกต่างกัน ด้วยวิธีการทดสอบพหุคูณรายคู่ 3 วิธี คือ วิธี Dunnett's T3 วิธี Games – Howell และวิธี Brown – Forsythe จากกลุ่มตัวอย่าง 4 ขนาดคือ ขนาด 20 คน 40 คน 80 คน และ 120 คน ว่าจะให้ผลการเปรียบเทียบพหุคูณแตกต่างกันหรือไม่ และการเปรียบเทียบพหุคูณวิธีใด ใช้กลุ่มตัวอย่างเท่าใดจึงจะเหมาะสมที่สุด กลุ่มตัวอย่างที่ทำการศึกษาคือ นักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ของนักเรียนในสังกัดกรมสามัญศึกษา กรุงเทพมหานคร จำนวน 583 คน ที่ได้จากการสุ่มแบบ 2 ขั้นตอน แล้วนำมากำหนดเป็นประชากรเทียม เพื่อสุ่มเป็นกลุ่มตัวอย่างขนาด 20 คน 40 คน 80 คน และ 120 คน ขนาดละ 10 ครั้ง ด้วยการสุ่มแบบใส่คืน เครื่องมือที่ใช้ในการศึกษาคือ แบบทดสอบวัดความถนัดด้านเหตุผล แบบสรุปความ ตามแนวของเธอร์สตัน จำนวน 30 ข้อ มีค่าความเชื่อมั่น 0.634 ผลการศึกษาพบว่า การเปรียบเทียบพหุคูณ จากการทดสอบพหุคูณรายคู่ 3 วิธี ในกลุ่มตัวอย่างขนาด 120 คน มีจำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนขนาด 20 คน 40 คน และ 80 คน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนผลการเปรียบเทียบพหุคูณระหว่างกลุ่มตัวอย่าง 4 ขนาด ที่ทดสอบพหุคูณรายคู่ 3 วิธีมีจำนวนผลการยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

มนทกานติ หรรษวรพงษ์ (2545 : 145) ทำการเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบความแตกต่างระหว่างค่าเฉลี่ยของ 2 ประชากร พบว่าเมื่อประชากรมีการแจกแจงแบบปกติ พบว่าตัวสถิติที่มีอำนาจการทดสอบสูงสุด สำหรับทุกสถานการณ์ที่ศึกษา และเมื่อประชากรมีการแจกแจงไม่เป็นแบบปกติ เมื่อสัมประสิทธิ์ความเบ้และสัมประสิทธิ์ความโด่งของการแจกแจงแบบปกติ พบว่าตัวสถิติดัดแปลงของโอเกอร์แมนมีอำนาจการทดสอบสูงสุด ยกเว้นที่ระดับนัยสำคัญทางสถิติ

.01 และขนาดตัวอย่างเท่ากับ 10 และ 15 ตัวสถิติแมน-วิทนี มีอำนาจการทดสอบสูง การเพิ่มระดับนัยสำคัญทางสถิติ ขนาดตัวอย่าง และเปอร์เซ็นต์ความแตกต่างระหว่างค่าเฉลี่ยมีผลให้ค่าอำนาจการทดสอบของตัวสถิติทั้ง 4 ประเภทเพิ่มขึ้นและมีค่าใกล้เคียงกัน

รุ่งโรจน์ ศรีจันทร์แก้ว (2547 : 83 – 84) ศึกษาการเปรียบเทียบอำนาจการทดสอบของผลการวิเคราะห์ข้อมูลด้วยแบบแผนวิเคราะห์แบบกลุ่มสุ่ม กับแบบแผนการวิเคราะห์ความแปรปรวนร่วม พบว่าผลของค่าอำนาจการทดสอบในการวิเคราะห์ข้อมูลด้วยแบบแผนการวิเคราะห์แบบกลุ่มสุ่มกับแบบแผนการวิเคราะห์ความแปรปรวนร่วม ในกลุ่มตัวอย่างขนาด 90 คน ให้ผลแตกต่างกัน คือแบบแผนทีวิเคราะห์ด้วยแบบแผนกลุ่มสุ่มให้ค่าอำนาจการทดสอบน้อยกว่า ค่าอำนาจของแบบแผนการวิเคราะห์ความแปรปรวนร่วมทุกค่า ส่วนการวิเคราะห์ข้อมูลด้วยแบบแผนการวิเคราะห์แบบกลุ่มสุ่มและการวิเคราะห์ความแปรปรวนร่วม ในกลุ่มตัวอย่างขนาด 180 คน และ 270 คน ให้ผลค่าอำนาจการทดสอบเท่ากัน

สมกิต วงศ์มาก (2549 : บทคัดย่อ) ศึกษาอำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียวภายหลังการแปลงข้อมูลจากกลุ่มตัวอย่างขนาดเล็ก โดยใช้เทคนิคการจำลองแบบมอนติคาร์โลและการใช้ข้อมูลจริง พบว่า อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ที่ศึกษาโดยใช้เทคนิคการจำลองแบบมอนติคาร์โล เมื่อกลุ่มตัวอย่างมีขนาดเพิ่มขึ้น โดยที่วิธีการแปลงข้อมูลระดับนัยสำคัญทางสถิติ และค่าความเบี่ยงที่ อำนาจการทดสอบมีค่าสูงขึ้น อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ที่ศึกษาโดยใช้เทคนิคการจำลองแบบมอนติคาร์โล เมื่อวิธีการแปลงข้อมูลต่างกัน โดยที่ขนาดกลุ่มตัวอย่างระดับนัยสำคัญทางสถิติ และค่าความเบี่ยงที่ อำนาจการทดสอบมีค่าไม่แตกต่างกัน อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ที่ศึกษาโดยใช้เทคนิคการจำลองแบบมอนติคาร์โล เมื่อระดับนัยสำคัญทางสถิติเพิ่มขึ้น โดยที่ขนาดกลุ่มตัวอย่าง วิธีการแปลงข้อมูล และค่าความเบี่ยงที่ อำนาจการทดสอบมีค่าสูงขึ้น อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ที่ศึกษาโดยใช้เทคนิคการจำลองแบบมอนติคาร์โล เมื่อค่าความเบี่ยงเพิ่มขึ้น โดยที่ขนาดกลุ่มตัวอย่าง วิธีการแปลงข้อมูลและระดับนัยสำคัญทางสถิติคงที่ อำนาจการทดสอบมีค่าไม่แตกต่างกัน อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ที่ศึกษาโดยใช้ข้อมูลจริง เมื่อกลุ่มตัวอย่างมีขนาดเพิ่มขึ้น โดยที่วิธีการแปลงข้อมูล และระดับนัยสำคัญทางสถิติคงที่ อำนาจการทดสอบมีค่าเพิ่มขึ้น อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ที่ศึกษาโดยใช้ข้อมูลจริง เมื่อมีการแปลงข้อมูลแตกต่างกัน โดยที่ขนาดกลุ่มตัวอย่าง และระดับนัยสำคัญทางสถิติคงที่ อำนาจการทดสอบมีค่าไม่แตกต่างกัน อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ที่ศึกษาโดยใช้ข้อมูลจริง เมื่อระดับนัยสำคัญทางสถิติเพิ่มขึ้น โดยที่ขนาดกลุ่มตัวอย่าง และวิธีการ

แปลงข้อมูลคงที่ อำนาจการทดสอบมีค่าสูงขึ้น อำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ที่ศึกษาโดยใช้เทคนิคการจำลองแบบมอนติคาร์โลและข้อมูลจริง มีความสอดคล้องกันในทุกสถานการณ์

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องในข้างต้น จะเห็นว่างานวิจัยที่ต้องการเปรียบเทียบค่าเฉลี่ยนั้น จำเป็นต้องเลือกใช้วิธีการเปรียบเทียบค่าเฉลี่ยที่เหมาะสมกับข้อมูลที่ผู้วิจัยต้องการจะศึกษา ขนาดของกลุ่มตัวอย่างเป็นกรณีศึกษากรณีหนึ่งที่ถือว่าเป็นตัวแปรสำคัญในการศึกษาเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ย ซึ่งมีการศึกษาทั้งข้อมูลจริงและข้อมูลจำลอง ดังนั้นผู้วิจัยจึงสนใจจะศึกษา การทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี คือ Tukey - Kramer วิธี Newman Keuls วิธี Duncan ซึ่งอยู่ในกลุ่ม t -statistic วิธี Scheffe ' ซึ่งอยู่ในกลุ่ม s -statistic และวิธี Fisher' s LSD ซึ่งอยู่ในกลุ่ม f -statistic ภายใต้ข้อตกลงที่เหมือนกันคือมีความแปรปรวนที่เท่ากัน มีการแจกแจงเป็นแบบปกติ กลุ่มตัวอย่างเป็นอิสระต่อกัน กลุ่มตัวอย่างมีจำนวนเท่ากันหรือไม่ก็ได้ เป็นการเปรียบเทียบรายคู่ ว่าให้ค่าอำนาจการทดสอบต่างกันหรือไม่ เมื่อสถิติทดสอบความแตกต่างของค่าเฉลี่ยต่างวิธีกัน โดยที่ขนาดของกลุ่มตัวอย่างเท่ากัน และเมื่อขนาดของกลุ่มตัวอย่างมีขนาดต่างกัน โดยที่สถิติการทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ภาคเรียนที่ 2 ปีการศึกษา 2550 สังกัด กรุงเทพมหานคร จำนวน 62 โรงเรียน จำนวนนักเรียน 16,260 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ภาคเรียนที่ 2 ปีการศึกษา 2550 สังกัดกรุงเทพมหานคร จำนวน 564 คน โดยการสุ่มแบบหลายขั้นตอน

(Multi - Stage Random Sampling)

ขั้นแรก สุ่มจากโรงเรียนมัธยมศึกษาตอนต้นในสังกัดของกรุงเทพมหานคร จำนวนทั้งหมด 6 กลุ่มโรงเรียนโดยวิธีการสุ่มอย่างง่าย (Simple Random Sampling) ได้กลุ่มโรงเรียนศรีนครินทร์ ซึ่งมีจำนวนโรงเรียนมัธยมศึกษาขยายโอกาสจำนวน 7 โรงเรียน ซึ่งมีจำนวนนักเรียน 2,853 คน ขั้นที่สอง สุ่มจากกลุ่มโรงเรียนศรีนครินทร์ 7 โรงเรียนโดยวิธีการสุ่มอย่างง่าย (Simple Random Sampling) ได้ 5 โรงเรียน จำนวนนักเรียนทั้งหมด 564 คน ขั้นที่สาม สุ่มแบบแบ่งชั้น จากโรงเรียนมัธยมศึกษาขยายโอกาส 5 โรงเรียน ได้นักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 195 คน นักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 182 คน และนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 187 คน โดยให้ระดับชั้น (มัธยมศึกษาปีที่ 1, 2 และ 3) เป็นชั้นของการสุ่ม (Strata) และห้องเรียนเป็นหน่วยการสุ่ม (Sampling unit) ดังแสดงในตาราง 3

ตาราง 3 แสดงจำนวนโรงเรียนและห้องเรียนที่ทำการสุ่มได้ของชั้นมัธยมศึกษาปีที่ 1 , 2 และ 3

โรงเรียน	ม.1			ม.2			ม.3		
	จำนวน			จำนวน			จำนวน		
	ห้อง	ห้อง ที่สุ่ม ได้	นักเรียน	ห้อง	ห้อง ที่สุ่ม ได้	นักเรียน	ห้อง	ห้อง ที่สุ่ม ได้	นักเรียน
สุเหร่าทรายกองดิน	3	1	45	3	1	33	2	1	33
คลองกุ่ม	2	1	32	2	1	32	2	1	32
บ้านบางกะปิ	4	2	45	4	2	45	4	2	45
วัดคูบอน	2	1	33	2	1	32	2	1	32
ไทยนิยมสงเคราะห์	5	2	40	5	2	40	4	2	45
รวม	16	7	195	16	7	182	14	7	187

เครื่องมือที่ใช้ในการศึกษา

เครื่องมือที่ใช้ในการศึกษาครั้งนี้เป็นแบบทดสอบวัดความถนัดด้านภาษา ซึ่งผู้วิจัยนำมาจากเครื่องมือของ โสภา ศรีสุข (2547 : 87- 96) ที่ทำการศึกษารื่อง การศึกษาความสัมพันธ์ระหว่างความถนัดทางภาษาและความคล่องแคล่วในการใช้คำกับความสามารถในการเขียนเรียงความ ของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยใช้แบบทดสอบวัดความถนัดด้านภาษา เป็นแบบทดสอบปรนัยแบบเลือกตอบ มีทั้งหมด 4 รูปแบบ มีคำตอบที่ถูกต้องเพียงข้อเดียว จำนวน 40 ข้อ ซึ่งมีความเชื่อมั่นทั้ง 4 รูปแบบดังนี้ รูปแบบที่ 1 แบบทดสอบศัพท์สัมพันธ์ มีความเชื่อมั่น 0.81 รูปแบบที่ 2 แบบทดสอบจัดประเภท มีความเชื่อมั่น 0.81 รูปแบบที่ 3 แบบทดสอบความเข้าใจภาษา มีความเชื่อมั่น 0.80 รูปแบบที่ 4 แบบทดสอบความเข้าใจภาพ มีความเชื่อมั่น 0.83 โดยผู้วิจัยนำมาหาคุณภาพของแบบทดสอบอีกครั้ง มีวิธีการหาคุณภาพของเครื่องมือดังนี้

1. นำแบบทดสอบวัดความถนัดด้านภาษาไปทดสอบครั้งที่ 1 กับนักเรียนชั้นมัธยมศึกษาปีที่ 1 , 2 และ 3 ที่ไม่ใช่กลุ่มตัวอย่างในการศึกษา จำนวน 350 คน และนำไปวิเคราะห์หาค่าความเชื่อมั่นทั้งฉบับ ด้วยวิธี KR – 20 ของคูเดอร์ – ริชาร์ดสัน ซึ่งได้ค่าความเชื่อมั่นทั้ง 4 รูปแบบดังนี้ รูปแบบที่ 1 แบบทดสอบศัพท์สัมพันธ์ มีความเชื่อมั่น 0.86 รูปแบบที่ 2 แบบทดสอบ

จัดประเภท มีค่าความเชื่อมั่น 0.82 รูปแบบที่ 3 แบบทดสอบความเข้าใจภาษา มีค่าความเชื่อมั่น 0.85 รูปแบบที่ 4 แบบทดสอบความเข้าใจภาพ มีค่าความเชื่อมั่น 0.91

2. จัดทำรูปเล่มของแบบทดสอบวัดความถนัดทางด้านภาษา พร้อมคำชี้แจงในการทำแบบทดสอบ เพื่อนำไปสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างในการศึกษาครั้งนี้

ตัวอย่างแบบทดสอบวัดความถนัดด้านภาษา

คำชี้แจง 1. แบบทดสอบฉบับนี้เป็นแบบทดสอบวัดความถนัดทางด้านภาษา เป็นแบบทดสอบปรนัยแบบเลือกตอบ มีทั้งหมด 4 รูปแบบ มีคำตอบที่ถูกต้องเพียงข้อเดียว จำนวน 40 ข้อ ใช้เวลาในการทำข้อสอบ 60 นาที

2. ให้นักเรียนเลือกพิจารณาคำตอบที่ถูกต้องที่สุดเพียงข้อเดียว จากตัวเลือก 1, 2, 3, 4 หรือ 5 ที่กำหนดให้

3. ในการตอบแบบตัวเลือกเมื่อเลือกคำตอบใดให้ทำเครื่องหมาย X ให้เต็มช่องสี่เหลี่ยมใต้ตัวเลขนั้นในกระดาษคำตอบ

4. ถ้าต้องการเปลี่ยนคำตอบให้ทำเครื่องหมาย = ทับเครื่องหมาย X ในคำตอบเดิมก่อน

ก. **ศัพท์สัมพันธ์** เป็นแบบให้หาความหมายของคำต่างๆ แล้วนำมาพิจารณาเปรียบเทียบว่าเกี่ยวข้องกันอย่างไร ดังตัวอย่าง

(0) **ดินสอ** เกี่ยวข้องกับคำใด ?

1. สี
2. ปากกา
3. ยางลบ
4. กระเป๋
5. กระดาษ

เฉลย 5

ข. แบบจัดประเภท เป็นแบบทดสอบที่กำหนดประโยคหรือวลีไว้แล้ว ให้หาว่าประโยคหรือวลีใดมีความหมายผิดแผกไปจากข้ออื่น ๆ

(0) ข้อความใดไม่เข้าพวก ?

1. เขาช่วยเหลือเธอสินะ
2. เขาไปเถาะฉันมีแล้ว
3. วันนี้เขาซื้อขนมมาให้คุณแม่ด้วย
4. วันนี้คุณครูให้รางวัลนักเรียนนะจ๊ะ
5. ฉันทวงเงินเธอทุกวันแล้วนะยังไม่ให้อีก

เฉลย 5

ค. ความเข้าใจภาษา เป็นแบบทดสอบที่มีโจทย์กำหนดข้อความ ซึ่งอาจเป็นร้อยแก้วหรือร้อยกรองมาให้แล้ว ให้ใช้ข้อความที่ให้มาเป็นหลัก

คำชี้แจง ให้ใช้ข้อความนี้ตอบคำถามข้อ (0)

หนึ่งคนด้านความคิด
หนึ่งคนด้านความสำคัญ

หนึ่งชีวิตด้านความฝัน
แต่หนึ่งเธอนั้นสำคัญกว่าล้านคน

(0) ผู้เขียนข้อความนี้ต้องการเน้นในเรื่องใด ?

1. ความคิด
2. ความฝัน
3. คนสำคัญ
4. ความสำคัญ
5. จำนวนคนที่ต่างกัน

เฉลย 3

ง. แบบเข้าใจภาพ เป็นแบบทดสอบที่กำหนดภาพมาให้พิจารณาภาพจากนั้นตอบคำถาม



(0) อาการของคนในภาพเป็นอย่างไร ?

1. เศร้า
2. สงสัย
3. เสียใจ
4. ตื่นเต้น
5. ประหลาดใจ

เฉลย 4

เกณฑ์การให้คะแนน ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน

วิธีเก็บรวบรวมข้อมูล

1. ขอลหนังสือจากบัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒเพื่อขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูลในภาคเรียนที่ 2 ปีการศึกษา 2550 ช่วงเดือน พฤศจิกายน พ.ศ. 2550
2. ติดต่อโรงเรียนที่เป็นกลุ่มตัวอย่างในการศึกษาขออนุญาตผู้บริหารโรงเรียนเพื่อกำหนด วัน เวลา ที่ทำการทดสอบ และวางแผนในการดำเนินการสอบโดยผู้วิจัยเป็นผู้ดำเนินการด้วยตนเอง จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่ทำการทดสอบในแต่ละครั้ง
3. นำแบบทดสอบไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างทั้งหมด ตามวัน เวลา ที่นัดหมาย จากนั้นชี้แจงให้นักเรียนที่เป็นกลุ่มตัวอย่างเข้าใจและทราบถึงวัตถุประสงค์ในการทำแบบทดสอบครั้งนี้ พร้อมทั้งอธิบายวิธีการทำแบบทดสอบวัดความสามารถด้านภาษา ซึ่งมีทั้งหมด 40 ข้อ เป็นข้อสอบปรนัยแบบเลือกตอบ ให้นักเรียนอ่านคำชี้แจงในแต่ละรูปแบบให้เข้าใจ แล้วพิจารณาเลือกข้อที่ถูกต้องเพียงข้อเดียวจากตัวเลือก 1, 2, 3, 4 และ 5
4. หลังจากได้ดำเนินการสอบกลุ่มตัวอย่างครบทุกโรงเรียนที่ทำการสุ่มมาแล้วก็นำกระดาษคำตอบมาตรวจให้คะแนน ดังนี้ ถ้าตอบถูกต้องให้ 1 คะแนน ถ้าตอบผิดหรือไม่ตอบหรือตอบมากกว่า 1 ข้อให้ 0 คะแนน แล้วนำผลมาบันทึกข้อมูลและวิเคราะห์ทางสถิติ

การวิเคราะห์ข้อมูล

1. นำข้อมูลที่ได้จากการทดสอบของกลุ่มตัวอย่างขนาด 564 มาบันทึกเป็นประชากรเทียม
2. สุ่มข้อมูลจากประชากรเทียม ให้มีขนาด 30 คน 60 คน 90 คน และ 120 คน ซึ่งแต่ละขนาดจะสุ่มมา 10 กลุ่ม โดยทำการสุ่มข้อมูลแต่ละกลุ่มแต่ละขนาดด้วยการสุ่มแบบใส่คืน

3. นำข้อมูลทุกกลุ่มทุกขนาดมาทดสอบความเป็นเอกพันธ์ของความแปรปรวนของกลุ่มตัวอย่างด้วยสถิติทดสอบเอฟ ถ้ามีกลุ่มตัวอย่างใดที่ไม่เป็นเอกพันธ์ของความแปรปรวนให้ทำการสุ่มแบบใส่คืนใหม่จนครบตามจำนวนกลุ่มที่ต้องการ

4. นำข้อมูลทุกกลุ่มทุกขนาดที่ไม่เป็นเอกพันธ์ของความแปรปรวนมาวิเคราะห์ความแปรปรวน แล้วจึงทดสอบความแตกต่างของค่าเฉลี่ยด้วยสถิติทดสอบทั้ง 5 วิธี คือ วิธี Tukey – Kramer, วิธี Newman Keuls , วิธี Duncan , วิธี Scheffe ' และวิธี Fisher's LSD ด้วยระบบคอมพิวเตอร์กับทุกกลุ่ม

5. ทดสอบจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญทางสถิติ จากการวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยที่ต่างวิธีกัน ในกลุ่มตัวอย่างที่มีขนาดเดียวกัน โดยวิธี χ^2 - test

6. ทดสอบจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญทางสถิติ จากการวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกันในกลุ่มตัวอย่างที่มีขนาดต่างกัน โดยวิธี χ^2 - test

7. หาค่าดัชนีความแตกต่างระหว่างค่าเฉลี่ยของประชากร (ϕ) แล้วทำการเปิดตารางของแท็งค์ (Tang. Power Function for Analysis of Variance) (Kirk. 1995 : 816 – 823) ซึ่งจะได้ค่า $1 - \beta$

8. เปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธีในแต่ละกลุ่มแต่ละขนาดของกลุ่มตัวอย่างทั้ง 4 ขนาด โดยการนำเสนอในรูปแบบตาราง

สถิติที่ใช้ในการศึกษา

1. สถิติที่ใช้ในการหาคุณภาพของเครื่องมือ
ในการทำแบบทดสอบครั้งที่ 1 เพื่อหาคุณภาพของเครื่องมือ และหาค่าความเชื่อมั่นของแบบทดสอบ จากสูตร KR-20 ของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson Procedure) (ล้วน สายยศและ อังคณา สายยศ. 2539 : 215)

$$r_{tt} = \frac{n}{n-1} \left\{ 1 - \frac{\sum pq}{s_t^2} \right\}$$

เมื่อ	r_{tt}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
	n	แทน	จำนวนข้อของเครื่องมือวัด
	p	แทน	สัดส่วนของผู้ทำถูกในข้อที่วิเคราะห์
	q	แทน	สัดส่วนของผู้ทำผิดในข้อที่วิเคราะห์ หรือ $1-p$
	s_t^2	แทน	คะแนนความแปรปรวนของเครื่องมือ

2. สถิติทดสอบสมมติฐาน

2.1 ค่าสถิติพื้นฐาน คือ ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐานและสัมประสิทธิ์การกระจาย

2.2 สถิติทดสอบความแตกต่างของค่าเฉลี่ย ซึ่งในวิจัยครั้งนี้ มี 5 วิธี คือ

2.2.1 วิธี Tukey - Kramer (Kirk. 1995 : 144 – 155)

เมื่อขนาดของกลุ่มตัวอย่างแต่ละชุดเท่ากัน ($n_1 = n_2 = \dots n_k = n_c$) โดยที่ค่าวิกฤตจะมีค่า

ดังนี้

$$T = q_\alpha(k, n-k) \sqrt{\frac{MSE}{n_c}}$$

เมื่อ	T	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey - Kramer
	q_α	แทน	ค่าวิกฤตของ Studentized Rang
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน
	n	แทน	ขนาดของกลุ่มตัวอย่าง
	n_c	แทน	จำนวนสมาชิกในกลุ่มตัวอย่างชุดที่ C
	k	แทน	จำนวนประชากร

เมื่อขนาดของกลุ่มตัวอย่างไม่เท่ากัน

$$T = q_\alpha(k, n-k) \sqrt{\frac{MSE}{2} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

เมื่อ	T	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey - Kramer
	q_α	แทน	ค่าวิกฤตของ Studentized Rang
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน
	n	แทน	ขนาดของกลุ่มตัวอย่าง
	n_c	แทน	จำนวนสมาชิกในกลุ่มตัวอย่างชุดที่ C
	k	แทน	จำนวนประชากร

2.2.2 วิธี Newman Keuls (Kirk. 1995 : 144 – 155)

เมื่อขนาดของกลุ่มตัวอย่างแต่ละชุดเท่ากัน ($n_1 = n_2 = \dots n_k = n_c$) โดยที่ค่าวิกฤตจะมีค่า
ดังนี้

$$NK = q_\alpha(r, n-k) \sqrt{\frac{MSE}{n_c}}$$

เมื่อ	NK	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Newman Keuls
	q	แทน	ค่าวิกฤตในตารางค่าวิกฤตของ Studentized rang
	n_c	แทน	จำนวนสมาชิกในกลุ่มตัวอย่างชุดที่ C
	k	แทน	จำนวนกลุ่มตัวอย่าง
	r	แทน	ระยะห่างของค่าเฉลี่ยที่ต้องการเปรียบเทียบ

เมื่อขนาดของกลุ่มตัวอย่างไม่เท่ากัน

กรณีที่ $n_i \neq n_j$ อย่างน้อย 1 คู่ ; $i \neq j$ วิธีของ NK จะปฏิเสธสมมติฐาน $H_0 : \mu_i = \mu_j$
หรือสรุปว่า $\mu_i \neq \mu_j$ เมื่อ $|\bar{X}_i - \bar{X}_j| > NK$ โดยที่

$$NK = q_\alpha(r, n-k) \sqrt{\frac{MSE}{2} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

เมื่อ	NK	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Newman Keuls
	q	แทน	ค่าวิกฤตในตารางค่าวิกฤตของ Studentized rang
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน
	n	แทน	จำนวนสมาชิกในกลุ่มตัวอย่าง
	k	แทน	จำนวนกลุ่มตัวอย่าง
	r	แทน	ระยะห่างของค่าเฉลี่ยที่ต้องการเปรียบเทียบ

2.2.3 วิธี Duncan (Kirk. 1995 : 144 – 155)

$$W = P_\alpha(r, n - k) \sqrt{\frac{MSE}{n_c}}$$

เมื่อ	W	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Duncan
	n_c	แทน	จำนวนสมาชิกในกลุ่มตัวอย่างชุดที่ C
	k	แทน	จำนวนกลุ่มตัวอย่าง
	r	แทน	ระยะห่างของค่าเฉลี่ยที่ต้องการเปรียบเทียบ
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน
	n	แทน	จำนวนสมาชิกในกลุ่มตัวอย่าง
	P_α	แทน	ค่าที่ได้จากการเปิดตาราง

2.2.4 วิธี Scheffe' (Scheffe' . 1959)

$$F_{\alpha(\bar{X}_1, \bar{X}_2)} = \frac{(\bar{X}_1 - \bar{X}_2)^2}{MS_w \left(\frac{1}{n_1} + \frac{1}{n_2} \right) (k - 1)} \quad (\text{ในกรณีที่เปรียบเทียบ } \bar{X}_1 \text{ กับ } \bar{X}_2)$$

เมื่อ	df	=	$k - 1, n - k$
	F	แทน	ค่าวิกฤตในตารางค่าวิกฤตของ Studentized rang
	k	แทน	จำนวนกลุ่มตัวอย่าง
	n	แทน	จำนวนตัวอย่างทั้งหมด
	n_1, n_2	แทน	จำนวนกลุ่มตัวอย่างในกลุ่มที่ 1 และกลุ่มที่ 2

2.2.5 วิธี Fisher' LSD (Kirk. 1995 : 144 – 155)

$$LSD = t_{1-\frac{\alpha}{2}; n-k} \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \quad \text{โดยที่ } n = \sum_{i=1}^k n_i$$

เมื่อ	LSD	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Fisher's LSD
	k	แทน	จำนวนกลุ่มตัวอย่าง
	n	แทน	จำนวนตัวอย่างทั้งหมด
	MSE	แทน	ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน

2.3 สถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวน (ระวีวรรณ พันธุ์พานิช.

2540 :36)

$$F = \frac{MS_B}{MS_W} \quad ; \quad df = j-1, N-1$$

เมื่อ F	แทน	การวิเคราะห์ความแปรปรวน
MS _B	แทน	ค่าเฉลี่ยของค่าเบี่ยงเบนกำลังสองระหว่างกลุ่ม
MS _W	แทน	ค่าเฉลี่ยของค่าเบี่ยงเบนกำลังสองภายในกลุ่ม
ถ้า F คำนวณ > F ตาราง จะไม่ยอมรับ H ₀		

2.4 สถิติทดสอบนัยสำคัญของความแตกต่างจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญที่ได้มาจากการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยต่างวิธีกัน ในกลุ่มตัวอย่างที่มีขนาดเท่ากัน และการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน ในกลุ่มตัวอย่างที่มีขนาดต่างกัน โดยใช้ χ^2 -test (ล้วน สายยศ และอังคณา สายยศ. 2540 : 261)

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

เมื่อ	χ^2	แทน	การทดสอบไค-สแควร์
	O	แทน	ค่าความถี่ที่ได้มาจริง (จำนวนกลุ่มที่ปฏิเสธ H_0)
	E	แทน	ค่าความถี่ที่ได้มาตามทฤษฎีหรือที่คาดหวัง (จำนวนกลุ่มที่คาดหวังว่าจะปฏิเสธ H_0)

นำค่า χ^2 ที่คำนวณได้ไปเปรียบเทียบกับ χ^2 วิกฤต ซึ่งเปิดตารางไค-สแควร์ที่ $df = n - 1$ ถ้าค่า χ^2 ที่คำนวณมากกว่า χ^2 วิกฤต หมายความว่า ขนาดกลุ่มตัวอย่างที่ต่างกัน ให้ผลต่างกันในการใช้

2.5 ดัชนีความแตกต่างระหว่างค่าเฉลี่ยของประชากร ซึ่งคำนวณได้จากสูตร (Kirk. 1982 : 142)

$$\phi = \sqrt{\frac{\sum_{j=1}^p \alpha_j^2 / p}{\sigma_\epsilon^2 / n}}$$

เมื่อ	ϕ	แทน	ดัชนีความแตกต่างระหว่างค่าเฉลี่ย
	α	แทน	treatment effect
	$\sum_j^p \alpha_j^2$	แทน	ผลรวมกำลังสองของ treatment effect
	p	แทน	ระดับตัวแปรทดลอง
	n	แทน	ขนาดของกลุ่มตัวอย่าง
	σ_ϵ^2	แทน	ความคลาดเคลื่อนของความแปรปรวน (Error Variance)

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การนำเสนอผลการวิเคราะห์ข้อมูล

การศึกษาครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับดังนี้

1. ค่าสถิติพื้นฐานของคะแนนที่ได้จากการทำแบบทดสอบวัดความถนัดด้านภาษาจากกลุ่มตัวอย่างขนาด 564 คน ที่เป็นประชากรเทียม
2. การทดสอบลักษณะของข้อมูลเบื้องต้น คือ ความแตกต่างของค่าเฉลี่ย ซึ่งใช้ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐาน ที่คำนวณได้จากคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1,2 และ 3 ที่เป็นประชากรเทียม โดยการสุ่มแบบใส่คืน ทั้งหมด 4 ขนาด คือ ขนาด 30 คน 60 คน 90 คน และ 120 คน
3. การทดสอบความแตกต่างของค่าเฉลี่ย ด้วยสถิติทดสอบ 5 วิธี จากกลุ่มตัวอย่างขนาดเดียวกัน คือ ขนาด 30 คน 60 คน 90 คน และ 120 คน
4. การทดสอบความแตกต่างของค่าเฉลี่ย จากกลุ่มตัวอย่างขนาดต่างกัน 4 ขนาด ในแต่ละสถิติทดสอบ 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD
5. ค่าสถิติทดสอบเอฟด้วยการวิเคราะห์ความแปรปรวนแบบทางเดียว
6. การเปรียบเทียบค่าอำนาจการทดสอบที่วิเคราะห์ด้วยสถิติทดสอบ 5 วิธี จากกลุ่มตัวอย่างขนาดเดียวกัน คือ ขนาด 30 คน 60 คน 90 คน และ 120 และการเปรียบเทียบค่าอำนาจการทดสอบที่วิเคราะห์ด้วยกลุ่มตัวอย่างขนาดต่างกัน 4 ขนาด ในแต่ละสถิติทดสอบ 5 วิธี คือ Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD จำนวน 10 ครั้ง

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลและการแปลความหมายของผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

$\bar{X}$	แทน	คะแนนเฉลี่ย
s	แทน	ส่วนเบี่ยงเบนมาตรฐาน
C.V.	แทน	สัมประสิทธิ์การกระจาย
k	แทน	คะแนนเต็มข้อสอบ

T	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยของ Tukey-Kramer
NK	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยของ Newman Keuls
W	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยของ Duncan
S	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยของ Scheffe'
LSD	แทน	สถิติทดสอบความแตกต่างของค่าเฉลี่ยของ Fisher's LSD
F	แทน	ค่าการวิเคราะห์ความแปรปรวน
χ^2	แทน	สถิติทดสอบไค-สแควร์
ϕ	แทน	ดัชนีความแตกต่างระหว่างค่าเฉลี่ยของประชากร
$1 - \beta$	แทน	ค่าอำนาจการทดสอบ

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบวัดความถนัดด้านภาษา ซึ่งทดสอบจากกลุ่มตัวอย่าง ขนาด 564 คน ที่เป็นประชากรเทียม

การวิเคราะห์ครั้งนี้ ได้จากการนำคะแนนผลการทดสอบด้วยแบบทดสอบวัดความถนัดด้านภาษา กับกลุ่มตัวอย่าง ขนาด 564 คน ที่เป็นประชากรเทียมมาหาค่าสถิติพื้นฐาน คือ ค่าเฉลี่ย และค่าส่วนเบี่ยงเบนมาตรฐาน ดังปรากฏในตาราง 4

ตาราง 4 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างในการศึกษา

ชั้น	k	N	$\bar{X}$	s	C.V.
ม.1	40	195	21.28	6.21	29.18
ม.2	40	182	28.10	7.13	25.37
ม.3	40	187	32.40	4.90	15.12
รวม	40	564	25.34	7.48	29.51

ผลการวิเคราะห์ตามตาราง 4 ปรากฏว่าจากกลุ่มตัวอย่างซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 จำนวนทั้งหมด 564 คน ทำการสอบด้วยแบบทดสอบความถนัดด้านภาษาจำนวน 40 ข้อ 40 คะแนน เมื่อจำแนกตามระดับชั้น พบว่า นักเรียนชั้น ม. 3 มีคะแนนเฉลี่ยสูงสุด คือ 32.40

รองลงมาคือ ม.2 และ ม.1 มีคะแนนเฉลี่ย 28.10 และ 21.28 ตามลำดับ มีค่าส่วนเบี่ยงเบนมาตรฐาน 7.48 และค่าสัมประสิทธิ์การกระจาย 29.51

2. การทดสอบลักษณะข้อมูลเบื้องต้น

การวิเคราะห์ในตอนนี้ได้นำคะแนนมาทดสอบความแตกต่างของค่าเฉลี่ย โดยใช้ค่าเฉลี่ยและความแปรปรวนที่คำนวณได้จากคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1 , 2 และ 3 ซึ่งสุ่มมาจากประชากรเทียมเป็นการสุ่มแบบใส่คืน ทั้งหมด 4 ขนาด ตามลำดับดังนี้

2.1 การทดสอบความแตกต่างของค่าเฉลี่ย ของกลุ่มตัวอย่างที่สุ่มมาจากประชากรเทียมขนาด 30 คน ซึ่งสุ่มมา ชั้นละ 10 กลุ่ม ปรากฏดังตาราง 5

ตาราง 5 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างขนาด 30 คน ชั้นละ 10 กลุ่ม

กลุ่ม ตัวอย่างที่	ม.1		ม.2		ม.3		F	p
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s		
1	25.73	7.51	25.30	6.42	34.53	5.22	50.10	.000
2	25.00	7.29	27.86	7.07	34.56	5.23	48.02	.000
3	27.63	8.06	28.83	7.31	34.53	5.22	62.09	.000
4	23.76	6.93	30.43	7.72	33.53	5.07	70.10	.000
5	22.23	6.49	31.30	7.94	34.53	5.22	106.84	.000
6	19.30	5.63	30.53	7.74	31.93	4.83	112.89	.000
7	19.26	5.62	32.23	8.18	29.90	4.52	40.91	.000
8	19.26	5.62	32.23	8.18	29.90	4.52	46.70	.000
9	23.96	6.99	28.30	7.18	33.73	5.10	37.42	.000
10	23.56	6.87	30.33	7.69	29.23	4.42	23.74	.000

$P^{**} < .01$

ผลการวิเคราะห์ตามตาราง 5 ปรากฏว่า ค่าเฉลี่ยของคะแนนความถนัดด้านภาษาของกลุ่มตัวอย่างขนาด 30 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 มีค่าตั้งแต่ 19.26 ถึง 27.63 ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 25.30 ถึง 32.23 และของ นักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าตั้งแต่ 29.23 ถึง 34.56 ค่าส่วนเบี่ยงเบนมาตรฐานของนักเรียนชั้นมัธยมศึกษาปีที่ 1 มีค่าตั้งแต่ 5.62 ถึง 8.06 ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 6.42 ถึง 8.18 และของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าตั้งแต่ 4.42 ถึง 5.23 จากการวิเคราะห์ความแปรปรวน พบว่า กลุ่มตัวอย่างทุกกลุ่มมีค่าเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.2 การทดสอบความแตกต่างของค่าเฉลี่ย ของกลุ่มตัวอย่างที่สุ่มมาจากประชากรเทียม ขนาด 60 คน ซึ่งสุ่มมา ชั้นละ 10 กลุ่ม ปรากฏดังตาราง 6

ตาราง 6 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างขนาด 60 คน ชั้นละ 10 กลุ่ม

กลุ่ม ตัวอย่างที่	ม.1		ม.2		ม.3		F	p
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s		
1	25.68	7.49	26.31	6.67	34.96	5.28	26.03	.000
2	23.45	6.84	29.58	7.50	33.38	5.04	26.46	.000
3	23.38	6.82	29.68	7.53	34.45	5.21	14.48	.000
4	21.20	6.18	30.71	7.79	32.73	4.95	18.11	.000
5	21.43	6.25	31.76	8.06	34.05	5.15	47.30	.000
6	19.71	5.75	31.51	7.99	31.90	4.82	72.28	.000
7	22.30	6.51	29.90	7.58	31.81	4.81	46.52	.000
8	23.08	6.73	28.68	7.27	32.80	4.96	46.52	.000
9	21.76	6.35	28.95	7.34	29.33	4.43	24.06	.000
10	24.28	7.08	27.91	7.08	31.41	4.75	15.73	.000

$P^{**} < .01$

ผลการวิเคราะห์ตามตาราง 6 ปรากฏว่า ค่าเฉลี่ยของคะแนนความถนัดด้านภาษาของกลุ่มตัวอย่างขนาด 60 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 มีค่าตั้งแต่ 19.71 ถึง 25.68 ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 26.31 ถึง 31.76 และของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าตั้งแต่ 29.33 ถึง 34.96 ค่าส่วนเบี่ยงเบนมาตรฐานของนักเรียนชั้นมัธยมศึกษาปีที่ 1 มีค่าตั้งแต่ 5.75 ถึง 7.49 ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 6.67 ถึง 8.06 และของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าตั้งแต่ 4.43 ถึง 5.28 จากการวิเคราะห์ความแปรปรวน พบว่า กลุ่มตัวอย่างทุกกลุ่มมีค่าเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.3 การทดสอบความแตกต่างของค่าเฉลี่ย ของกลุ่มตัวอย่างที่สุ่มมาจากประชากร เทียมขนาด 90 คน ซึ่งสุ่มมา ชั้นละ 10 กลุ่ม ปรากฏดังตาราง 7

ตาราง 7 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างขนาด 90 คน ชั้นละ 10 กลุ่ม

กลุ่ม ตัวอย่างที่	ม.1		ม.2		ม.3		F	p
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s		
1	23.85	6.96	28.70	7.28	33.72	5.10	64.58	.000
2	22.75	6.64	26.15	6.63	34.47	5.21	89.56	.000
3	24.39	7.12	26.23	6.65	34.12	5.16	65.02	.000
4	21.78	6.35	26.14	6.63	34.40	5.20	98.14	.000
5	20.07	5.85	29.48	7.48	34.20	5.17	137.59	.000
6	20.25	5.91	30.13	6.64	34.23	5.18	162.60	.000
7	22.42	6.54	30.93	7.85	33.40	5.05	119.70	.000
8	21.93	6.40	31.57	8.01	31.98	4.83	113.99	.000
9	20.96	6.11	30.15	7.65	31.18	4.71	76.45	.000
10	21.58	6.30	30.28	7.68	31.18	4.71	81.42	.000

P**< .01

ผลการวิเคราะห์ตามตาราง 7 ปรากฏว่า ค่าเฉลี่ยของคะแนนความถนัดด้านภาษาของกลุ่มตัวอย่างขนาด 90 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 มีค่าตั้งแต่ 20.07 ถึง 24.39 ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 26.14 ถึง 31.57 และของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าตั้งแต่ 31.18 ถึง 34.47 ค่าส่วนเบี่ยงเบนมาตรฐานของนักเรียนชั้นมัธยมศึกษาปีที่ 1 มีค่าตั้งแต่ 5.85 ถึง 7.12 ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 6.63 ถึง 8.01 และของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าตั้งแต่ 4.71 ถึง 5.21 จากการวิเคราะห์ความแปรปรวน พบว่า กลุ่มตัวอย่างทุกกลุ่มมีค่าเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.4 การทดสอบความแตกต่างของค่าเฉลี่ย ของกลุ่มตัวอย่างที่สุ่มมาจากประชากรเทียมขนาด 120 คน ซึ่งสุ่มมา ชั้นละ 10 กลุ่ม ปรากฏดังตาราง 8

ตาราง 8 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านภาษาของนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ที่เป็นกลุ่มตัวอย่างขนาด 120 คน ชั้นละ 10 กลุ่ม

กลุ่ม ตัวอย่างที่	ม.1		ม.2		ม.3		F	p
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s		
1	26.06	7.60	26.78	6.79	34.70	5.24	84.05	.000
2	25.85	7.54	26.17	6.64	35.61	5.38	120.63	.000
3	23.95	6.99	30.17	7.65	33.41	5.05	93.70	.000
4	22.60	6.59	29.30	7.43	34.33	5.19	146.76	.000
5	20.63	6.02	30.82	7.82	33.61	5.08	196.32	.000
6	22.05	6.43	31.75	8.06	34.15	5.16	227.94	.000
7	19.09	5.57	31.72	8.05	32.80	4.96	313.66	.000
8	22.77	6.64	30.11	7.64	31.85	4.81	72.07	.000
9	23.13	6.75	28.95	7.34	33.40	5.05	126.41	.000
10	20.88	6.09	28.29	7.18	30.01	4.53	91.34	.000

P**< .01

ผลการวิเคราะห์ตามตาราง 8 ปรากฏว่า ค่าเฉลี่ยของคะแนนความถนัดด้านภาษาของกลุ่มตัวอย่างขนาด 120 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 มีค่าตั้งแต่ 19.09 ถึง 26.06 ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 26.17 ถึง 31.75 และของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าตั้งแต่ 30.01 ถึง 35.61 ค่าส่วนเบี่ยงเบนมาตรฐานของนักเรียนชั้นมัธยมศึกษาปีที่ 1 มีค่าตั้งแต่ 5.57 ถึง 7.60 ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 6.64 ถึง 8.06 และของนักเรียนชั้นมัธยมศึกษาปีที่ 3 มีค่าตั้งแต่ 4.53 ถึง 5.38 จากการวิเคราะห์ความแปรปรวน พบว่า กลุ่มตัวอย่างทุกกลุ่มมีค่าเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. การทดสอบความแตกต่างของผลการเปรียบเทียบสถิติทดสอบความแตกต่างของค่าเฉลี่ยจากกลุ่มตัวอย่างขนาดเดียวกัน

การวิเคราะห์ครั้งนี้ได้นำจำนวนผลการเปรียบเทียบสถิติทดสอบความแตกต่างของค่าเฉลี่ยจากการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ย 5 วิธี คือ Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD ในแต่ละกลุ่มตัวอย่าง 4 ขนาด คือ ขนาด 30 คน 60 คน 90 คน และ 120 มาทดสอบความแตกต่างจำนวนผลการเปรียบเทียบสถิติทดสอบความแตกต่างของค่าเฉลี่ย ซึ่งวิเคราะห์เปรียบเทียบระหว่างกลุ่มตัวอย่างเดียวกันโดยใช้สถิติ χ^2 -test ปรากฏผลดังตาราง 9

ตาราง 9 การเปรียบเทียบจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญ จากการทดสอบความแตกต่างของค่าเฉลี่ยด้วยสถิติทดสอบ 5 วิธี จากกลุ่มตัวอย่าง 4 ขนาด

ขนาดของกลุ่มตัวอย่าง	ผลการทดสอบ	T	NK	W	S	LSD	χ^2
30 คน	ไม่ปฏิเสธ H_0	8	8	8	9	7	0.250
	ปฏิเสธ H_0	22	22	22	21	23	0.090
60 คน	ไม่ปฏิเสธ H_0	5	5	5	5	5	0.000
	ปฏิเสธ H_0	25	25	25	25	25	0.000
90 คน	ไม่ปฏิเสธ H_0	4	4	4	4	3	0.210
	ปฏิเสธ H_0	26	26	26	26	27	0.030
120 คน	ไม่ปฏิเสธ H_0	5	5	5	5	3	0.695
	ปฏิเสธ H_0	25	25	25	25	27	0.125

$P^{**} < .01$

ผลการวิเคราะห์ตามตาราง 9 ปรากฏว่าการเปรียบเทียบสถิติทดสอบความแตกต่างของค่าเฉลี่ย 5 วิธี คือ Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD จากกลุ่มตัวอย่าง 4 ขนาด คือ ขนาด 30 คน 60 คน 90 คน และ 120 คน ให้ผลการทดสอบไม่แตกต่างกัน

4. การทดสอบความแตกต่างของผลการเปรียบเทียบสถิติทดสอบความแตกต่างของค่าเฉลี่ยจากกลุ่มตัวอย่างขนาดต่างกัน การวิเคราะห์ครั้งนี้ได้นำผลการเปรียบเทียบสถิติทดสอบความแตกต่างของค่าเฉลี่ยจากกลุ่มตัวอย่าง 4 ขนาด คือ 30 คน 60 คน 90 คน และ 120 คน ในแต่ละวิธีการทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD มาทดสอบความแตกต่างของจำนวนการเปรียบเทียบความแตกต่างของค่าเฉลี่ย ซึ่งวิเคราะห์เปรียบเทียบระหว่างสถิติทดสอบวิธีเดียวกัน โดยใช้ χ^2 -test ตามลำดับดังนี้

ตาราง 10 การเปรียบเทียบจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญ จากกลุ่มตัวอย่าง 4 ขนาด ในแต่ละวิธีการทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี

สถิติทดสอบ	ผลการทดสอบ	ขนาดของกลุ่มตัวอย่าง (คน)				χ^2
		30	60	90	120	
วิธี Tukey-Kramer	ไม่ปฏิเสธ H_0	8	5	4	5	1.636
	ปฏิเสธ H_0	22	25	26	25	0.367
วิธี Newman Keuls	ไม่ปฏิเสธ H_0	8	5	4	5	1.636
	ปฏิเสธ H_0	22	25	26	25	0.367
วิธี Duncan	ไม่ปฏิเสธ H_0	8	5	4	5	1.636
	ปฏิเสธ H_0	22	25	26	25	0.367
วิธี Scheffe'	ไม่ปฏิเสธ H_0	9	5	4	5	2.565
	ปฏิเสธ H_0	21	25	26	25	0.608
วิธี Fisher's LSD	ไม่ปฏิเสธ H_0	7	5	3	3	2.444
	ปฏิเสธ H_0	23	25	27	27	0.712

$$P^{**} < .01$$

ผลการวิเคราะห์ตามตาราง 10 ปรากฏว่าการเปรียบเทียบสถิติทดสอบความแตกต่างของค่าเฉลี่ย จากกลุ่มตัวอย่าง 4 ขนาด คือ 30 คน 60 คน 90 คน และ 120 คน ในแต่ละวิธีการทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD ด้วยวิธี Tukey-Kramer ให้ผลการทดสอบไม่แตกต่างกัน

5. ผลการเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี จำแนกตามขนาดของกลุ่มตัวอย่างทั้ง 4 ขนาด

การวิเคราะห์ครั้งนี้ ได้นำค่าจากการวิเคราะห์ข้อมูลด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD ในแต่ละกลุ่มซึ่งมีทั้งหมด 10 กลุ่ม มาคำนวณหาค่า ϕ (ดัชนีความแตกต่างระหว่างค่าเฉลี่ย) แล้วนำไปเปิดตารางการอ่านค่าอำนาจการทดสอบของ แท็งค์ (Tang. Power Function for Analysis of Variance) (Kirk. 1995 : 816 – 823) ซึ่งจะได้ค่า $1 - \beta$ ตามลำดับดังนี้

ตาราง 11 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี จากกลุ่มตัวอย่างขนาด 30 คน

รอบที่	T		NK		W		S		LSD	
	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
1	4.237	0.79	4.237	0.79	4.237	0.79	4.237	0.79	4.237	0.79
2	4.237	0.79	4.237	0.79	4.237	0.79	4.237	0.79	4.237	0.79
3	3.161	0.68	3.161	0.68	3.161	0.68	3.161	0.68	3.161	0.68
4	3.535	0.71	3.535	0.71	3.535	0.71	3.535	0.71	3.535	0.71
5	5.712	0.98	5.712	0.98	5.712	0.98	5.712	0.98	5.712	0.98
6	7.061	0.99	7.061	0.99	7.061	0.99	7.061	0.99	7.061	0.99
7	5.665	0.98	5.665	0.98	5.665	0.98	5.665	0.98	5.665	0.98
8	5.665	0.98	5.665	0.98	5.665	0.98	5.665	0.98	5.665	0.98
9	4.074	0.76	4.074	0.76	4.074	0.76	4.074	0.76	4.074	0.76
10	3.294	0.69	3.294	0.69	3.294	0.69	3.294	0.69	3.294	0.69
1 - β มีค่า ระหว่าง	0.68 – 0.99		0.68 – 0.99		0.68 – 0.99		0.68 – 0.99		0.68 – 0.99	

ผลการวิเคราะห์ตามตาราง 11 ปรากฏว่า ในการวิเคราะห์ในกลุ่มตัวอย่างขนาด 30 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า สถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี ให้ค่าอำนาจการทดสอบไม่แตกต่างกัน มีค่าระหว่าง 0.68 – 0.99

ตาราง 12 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี จากกลุ่มตัวอย่างขนาด 60 คน

รอบที่	T		NK		W		S		LSD	
	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
1	5.828	0.98	5.828	0.98	5.828	0.98	5.828	0.98	5.828	0.98
2	5.706	0.98	5.706	0.98	5.706	0.98	5.706	0.98	5.706	0.98
3	6.488	0.99	6.488	0.99	6.488	0.99	6.488	0.99	6.488	0.99
4	6.894	0.99	6.894	0.99	6.894	0.99	6.894	0.99	6.894	0.99
5	8.511	0.99	8.511	0.99	8.511	0.99	8.511	0.99	8.511	0.99
6	8.749	0.99	8.749	0.99	8.749	0.99	8.749	0.99	8.749	0.99
7	5.267	0.95	5.267	0.95	5.267	0.95	5.267	0.95	5.267	0.95
8	5.627	0.98	5.627	0.98	5.627	0.98	5.627	0.98	5.627	0.98
9	5.037	0.93	5.037	0.93	5.037	0.93	5.037	0.93	5.037	0.93
10	4.012	0.76	4.012	0.76	4.012	0.76	4.012	0.76	4.012	0.76
1 - β มีค่า ระหว่าง	0.76 – 0.99		0.76 – 0.99		0.76 – 0.99		0.76 – 0.99		0.76 – 0.99	

ผลการวิเคราะห์ตามตาราง 12 ปรากฏว่า ในการวิเคราะห์ในกลุ่มตัวอย่างขนาด 60 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า สถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี ให้ค่าอำนาจการทดสอบไม่แตกต่างกัน มีค่าระหว่าง 0.76 – 0.99

ตาราง 13 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี จากกลุ่มตัวอย่างขนาด 90 คน

รอบที่	T		NK		W		S		LSD	
	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
1	6.599	0.99	6.599	0.99	6.599	0.99	6.599	0.99	6.599	0.99
2	7.77	0.99	7.77	0.99	7.77	0.99	7.77	0.99	7.77	0.99
3	6.621	0.99	6.621	0.99	6.621	0.99	6.621	0.99	6.621	0.99
4	8.134	0.99	8.134	0.99	8.134	0.99	8.134	0.99	8.134	0.99
5	9.631	0.99	9.631	0.99	9.631	0.99	9.631	0.99	9.631	0.99
6	10.47	0.99	10.47	0.99	10.47	0.99	10.47	0.99	10.47	0.99
7	8.983	0.99	8.983	0.99	8.983	0.99	8.983	0.99	8.983	0.99
8	8.766	0.99	8.766	0.99	8.766	0.99	8.766	0.99	8.766	0.99
9	7.179	0.99	7.179	0.99	7.179	0.99	7.179	0.99	7.179	0.99
10	7.409	0.99	7.409	0.99	7.409	0.99	7.409	0.99	7.409	0.99
1 - β มีค่า	0.99		0.99		0.99		0.99		0.99	

ผลการวิเคราะห์ตามตาราง 13 ปรากฏว่า ในการวิเคราะห์ในกลุ่มตัวอย่างขนาด 90 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า สถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี ให้ค่าอำนาจการทดสอบไม่แตกต่างกัน คือมีค่า 0.99

ตาราง 14 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำ
การทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธีจากกลุ่มตัวอย่างขนาด 120 คน

รอบที่	T		NK		W		S		LSD	
	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
1	7.517	0.99	7.517	0.99	7.517	0.99	7.517	0.99	7.517	0.99
2	9.005	0.99	9.005	0.99	9.005	0.99	9.005	0.99	9.005	0.99
3	7.937	0.99	7.937	0.99	7.937	0.99	7.937	0.99	7.937	0.99
4	9.933	0.99	9.933	0.99	9.933	0.99	9.933	0.99	9.933	0.99
5	11.49	0.99	11.49	0.99	11.49	0.99	11.49	0.99	11.49	0.99
6	12.38	0.99	12.38	0.99	12.38	0.99	12.38	0.99	12.38	0.99
7	14.52	0.99	14.52	0.99	14.52	0.99	14.52	0.99	14.52	0.99
8	6.961	0.99	6.961	0.99	6.961	0.99	6.961	0.99	6.961	0.99
9	9.219	0.99	9.219	0.99	9.219	0.99	9.219	0.99	9.219	0.99
10	7.836	0.99	7.836	0.99	7.836	0.99	7.836	0.99	7.836	0.99
1 - β มีค่า	0.99		0.99		0.99		0.99		0.99	

ผลการวิเคราะห์ตามตาราง 14 ปรากฏว่า ในการวิเคราะห์ในกลุ่มตัวอย่างขนาด 120 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า สถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี ให้ค่าอำนาจการทดสอบไม่แตกต่างกัน คือมีค่า 0.99

จากการเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยด้วยวิธีต่างกันทั้ง 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD เมื่อขนาดของกลุ่มตัวอย่างคงที่ทั้ง 4 ขนาดคือ 30 คน 60 คน 90 คน และ 120 คน พบว่า

1. ในกลุ่มตัวอย่างขนาดเดียวกันทั้ง 4 ขนาดคือ 30 คน 60 คน 90 คน และ 120 คน เมื่อทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยต่างกันทั้ง 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD พบว่าให้ค่าอำนาจการทดสอบไม่แตกต่างกันในทุกกลุ่ม

2. ในกลุ่มตัวอย่างขนาด 90 คน และ 120 คน สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD มีค่าอำนาจการทดสอบเท่ากันในทุกกลุ่ม และสูงกว่าค่าอำนาจการทดสอบในกลุ่มตัวอย่างขนาด 30 คน และ 60 คน

3. ในการเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยด้วยวิธีต่างกันทั้ง 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD เมื่อขนาดของกลุ่มตัวอย่างคงที่ทั้ง 4 ขนาดคือ 30 คน 60 คน 90 คน และ 120 คน พบว่า ขนาดกลุ่มตัวอย่างที่มีค่าอำนาจการทดสอบมากที่สุดคือ กลุ่มตัวอย่างขนาด 120 คนและ 90 คน รองลงมาคือ กลุ่มตัวอย่างขนาด 60 คน และมีค่าอำนาจการทดสอบน้อยที่สุด คือ กลุ่มตัวอย่างขนาด 30 คน

6. ผลการเปรียบเทียบค่าอำนาจการทดสอบเมื่อขนาดของกลุ่มตัวอย่างต่างกันทั้ง 4 ขนาด ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี

การวิเคราะห์ครั้งนี้ ได้นำค่าจากการวิเคราะห์ข้อมูลด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD ในแต่ละกลุ่มซึ่งมีทั้งหมด 10 กลุ่ม มาคำนวณหาค่า ϕ (ดัชนีความแตกต่างระหว่างค่าเฉลี่ย) แล้วนำไปเปิดตารางการอ่านค่าอำนาจการทดสอบของ แท็งค์ (Tang. Power Function for Analysis of Variance) (Kirk. 1995 : 816 – 823) ซึ่งจะได้ค่า $1 - \beta$ จากนั้นนำเสนอโดยการจำแนกตามสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธีเมื่อวิเคราะห์ในกลุ่มตัวอย่างต่างกันทั้ง 4 ขนาดคือ 30 คน 60 คน 90 คน และ 120 คน ตามลำดับดังนี้

ตาราง 15 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย

วิธี Tukey-Kramer ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คนและ 120 คน

สถิติ ทดสอบ	กลุ่ม ที่	ขนาดของกลุ่มตัวอย่าง							
		30 คน		60 คน		90 คน		120 คน	
		ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
Tukey- Kramer	1	3.021	0.67	4.235	0.79	4.231	0.79	4.350	0.80
	2	2.053	0.57	4.235	0.79	4.235	0.79	4.236	0.79
	3	2.090	0.57	3.256	0.69	4.026	0.76	4.687	0.86
	4	2.119	0.58	3.045	0.67	4.566	0.84	4.598	0.84
	5	2.714	0.64	4.026	0.76	4.026	0.76	5.013	0.93
	6	1.801	0.55	3.025	0.67	5.130	0.94	5.012	0.93
	7	2.900	0.66	4.566	0.84	4.236	0.79	5.023	0.93
	8	2.900	0.66	3.210	0.69	3.980	0.75	4.690	0.86
	9	1.628	0.53	3.210	0.69	4.013	0.76	5.698	0.98
	10	1.498	0.51	3.025	0.67	4.445	0.81	5.558	0.97
1 - β มีค่าระหว่าง		0.51 – 0.67		0.67 – 0.84		0.76 – 0.94		0.79 – 0.98	

ผลการวิเคราะห์ตามตาราง 15 ปรากฏว่า ในการวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer ในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า ในกลุ่มตัวอย่างทั้ง 4 ขนาด ให้ค่าอำนาจการทดสอบ มีค่าระหว่าง 0.51 - 0.98 และค่าอำนาจการทดสอบมีแนวโน้มสูงขึ้นเมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้น

ตาราง 16 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย

วิธี Newman Keuls ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คนและ 120 คน

สถิติ ทดสอบ	กลุ่ม ที่	ขนาดของกลุ่มตัวอย่าง							
		30 คน		60 คน		90 คน		120 คน	
		ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
Newman Keuls	1	4.348	0.80	3.265	0.69	4.562	0.84	5.160	0.94
	2	4.348	0.80	3.569	0.71	4.563	0.84	4.255	0.79
	3	2.712	0.64	4.356	0.80	4.055	0.76	4.567	0.84
	4	2.023	0.57	4.026	0.76	4.026	0.76	5.463	0.97
	5	2.572	0.62	4.135	0.78	4.032	0.76	5.169	0.94
	6	2.712	0.64	4.056	0.76	5.558	0.97	5.988	0.98
	7	1.660	0.53	4.023	0.76	4.568	0.84	5.690	0.98
	8	2.572	0.62	2.059	0.57	4.356	0.80	6.012	0.99
	9	2.252	0.59	5.034	0.93	4.235	0.79	6.022	0.99
	10	2.572	0.62	4.036	0.76	4.044	0.76	6.032	0.99
1 - β มีค่าระหว่าง		0.53 – 0.80		0.57 – 0.93		0.76 – 0.97		0.79 – 0.99	

ผลการวิเคราะห์ตามตาราง 16 ปรากฏว่า ในการวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Newman Keuls ในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า ในกลุ่มตัวอย่างทั้ง 4 ขนาด ให้ค่าอำนาจการทดสอบ มีค่าระหว่าง 0.53 - 0.99 และค่าอำนาจการทดสอบมีแนวโน้มสูงขึ้นเมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้น

ตาราง 17 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย

วิธี Duncan ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คนและ 120 คน

สถิติ ทดสอบ	กลุ่ม ที่	ขนาดของกลุ่มตัวอย่าง							
		30 คน		60 คน		90 คน		120 คน	
		ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
Duncan	1	4.467	0.81	4.569	0.84	5.268	0.95	5.789	0.98
	2	4.129	0.78	4.569	0.84	6.012	0.99	6.589	0.99
	3	4.160	0.78	4.236	0.79	5.167	0.94	6.235	0.99
	4	4.129	0.78	4.025	0.76	6.014	0.99	6.589	0.99
	5	4.181	0.78	4.358	0.80	6.235	0.99	6.235	0.99
	6	4.318	0.80	4.560	0.84	6.255	0.99	5.369	0.96
	7	4.129	0.78	5.260	0.95	6.435	0.99	6.235	0.99
	8	4.253	0.79	5.480	0.97	6.665	0.99	6.014	0.99
	9	4.832	0.89	6.132	0.99	4.578	0.84	6.158	0.99
	10	4.129	0.78	5.046	0.93	6.158	0.99	6.013	0.99
1 - β มีค่าระหว่าง		0.78 – 0.89		0.76 – 0.99		0.84 – 0.99		0.98 – 0.99	

ผลการวิเคราะห์ตามตาราง 17 ปรากฏว่า ในการวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Duncan ในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า ในกลุ่มตัวอย่างทั้ง 4 ขนาด ให้ค่าอำนาจการทดสอบ มีค่า ระหว่าง 0.78 - 0.99 และค่าอำนาจการทดสอบมีแนวโน้มสูงขึ้นเมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้น

ตาราง 18 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย

วิธี Scheffe' ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คนและ 120 คน

สถิติ ทดสอบ	กลุ่ม ที่	ขนาดของกลุ่มตัวอย่าง							
		30 คน		60 คน		90 คน		120 คน	
		ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
Scheffe'	1	3.054	0.67	4.568	0.87	5.712	0.98	5.125	0.94
	2	3.054	0.67	4.630	0.86	5.665	0.98	5.134	0.94
	3	3.639	0.72	4.465	0.81	5.828	0.98	5.258	0.95
	4	3.054	0.67	4.181	0.78	5.706	0.98	5.025	0.93
	5	3.639	0.72	4.318	0.80	5.276	0.95	6.668	0.99
	6	3.054	0.67	4.257	0.79	5.627	0.98	6.758	0.99
	7	3.639	0.72	5.236	0.95	5.037	0.93	5.123	0.94
	8	3.054	0.67	5.023	0.93	6.488	0.99	5.254	0.95
	9	3.639	0.72	4.568	0.84	6.894	0.99	5.347	0.96
	10	3.054	0.67	5.456	0.97	6.621	0.99	5.235	0.95
1 - β มีค่าระหว่าง		0.67 – 0.72		0.78 – 0.97		0.93 – 0.99		0.93 – 0.99	

ผลการวิเคราะห์ตามตาราง 18 ปรากฏว่า ในการวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Scheffe' ในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า ในกลุ่มตัวอย่างทั้ง 4 ขนาด ให้ค่าอำนาจการทดสอบมีค่า ระหว่าง 0.67 - 0.99 และค่าอำนาจการทดสอบมีแนวโน้มสูงขึ้นเมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้น

ตาราง 19 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ย

วิธี Fisher's LSD ที่ทำการทดสอบในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คนและ 120 คน

สถิติ ทดสอบ	กลุ่ม ที่	ขนาดของกลุ่มตัวอย่าง							
		30 คน		60 คน		90 คน		120 คน	
		ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$	ϕ	$1 - \beta$
Fisher's LSD	1	6.598	0.99	6.599	0.99	7.548	0.99	7.517	0.99
	2	7.458	0.99	7.778	0.99	9.002	0.99	9.026	0.99
	3	6.265	0.99	6.659	0.99	7.568	0.99	10.235	0.99
	4	7.181	0.99	8.153	0.99	8.023	0.99	12.023	0.99
	5	8.125	0.99	9.326	0.99	9.126	0.99	10.122	0.99
	6	6.325	0.99	9.120	0.99	10.254	0.99	14.568	0.99
	7	6.259	0.99	6.355	0.99	11.235	0.99	10.235	0.99
	8	6.023	0.99	7.165	0.99	12.410	0.99	7.026	0.99
	9	6.030	0.99	8.265	0.99	6.941	0.99	8.015	0.99
	10	6.013	0.99	6.255	0.99	9.219	0.99	9.989	0.99
1 - β มีค่า		0.99		0.99		0.99		0.99	

ผลการวิเคราะห์ตามตาราง 19 ปรากฏว่า ในการวิเคราะห์ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Fisher's LSD ในกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน กลุ่มที่ 1 – 10 ทั้ง 10 กลุ่ม พบว่า ในกลุ่มตัวอย่างทั้ง 4 ขนาด ให้ค่าอำนาจการทดสอบไม่แตกต่างกัน คือมีค่า 0.99 ทุกกลุ่ม

จากการเปรียบเทียบค่าอำนาจการทดสอบเมื่อขนาดของกลุ่มตัวอย่างต่างกัน ทั้ง 4 ขนาด คือ 30 คน 60 คน 90 คน และ 120 คน ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน ทั้ง 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe ' และวิธี Fisher's LSD พบว่า

1. ในสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกันทั้ง 5 วิธี ค่าอำนาจการทดสอบมีแนวโน้มสูงขึ้นเมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้น
2. สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan และ วิธี Scheffe ' ให้ค่าอำนาจการทดสอบแตกต่างกันในทุกขนาดของกลุ่มตัวอย่าง ส่วนวิธี Fisher's LSD ให้ค่าอำนาจการทดสอบไม่แตกต่างกันในทุกขนาดของกลุ่มตัวอย่าง
3. สถิติทดสอบความแตกต่างของค่าเฉลี่ยที่ให้ค่าอำนาจการทดสอบสูงสุดคือ วิธี Fisher's LSD รองลงมาคือ วิธี Duncan ,วิธี Scheffe ' , วิธี Newman Keuls และสถิติทดสอบความแตกต่างของค่าเฉลี่ยที่ให้ค่าอำนาจการทดสอบน้อยที่สุด คือ วิธี Tukey-Kramer

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

สังเขปความมุ่งหมาย สมมติฐาน และวิธีดำเนินการศึกษาค้นคว้า

การวิจัยครั้งนี้เป็นการศึกษาเชิงทฤษฎี มีวัตถุประสงค์เพื่อศึกษาอำนาจการทดสอบของสถิติ การทดสอบความแตกต่างของค่าเฉลี่ยต่างวิธีกัน โดยที่ขนาดของกลุ่มตัวอย่างเท่ากันและเพื่อศึกษาอำนาจการทดสอบเมื่อขนาดของกลุ่มตัวอย่างต่างกัน โดยที่สถิติการทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน โดยมีสมมติฐานดังนี้คือ อำนาจการทดสอบ ที่ศึกษาโดยใช้ข้อมูลจริงวิเคราะห์ ด้วยสถิติการทดสอบความแตกต่างของค่าเฉลี่ยต่างวิธีกัน จากกลุ่มตัวอย่างเท่ากัน อำนาจการทดสอบจะแตกต่างกัน และอำนาจการทดสอบ ที่ศึกษาโดยใช้ข้อมูลจริงวิเคราะห์ด้วยกลุ่มตัวอย่างขนาดต่างกันด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน อำนาจการทดสอบจะแตกต่างกัน

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้านี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1, 2 และ 3 ภาคเรียนที่ 2 ปีการศึกษา 2550 สังกัดกรุงเทพมหานคร จำนวน 564 คน สุ่มจากประชากรโดยวิธี การสุ่มแบบหลายขั้นตอน (Multi - Stage Random Sampling) เครื่องมือที่ใช้ในการศึกษาค้นคว้านี้เป็นแบบทดสอบวัดความถนัดด้านภาษา ซึ่งผู้วิจัยนำมาจากเครื่องมือของ ไสภา ศรีสุข (2547 : 87 – 96) ทำการศึกษาเรื่อง การศึกษาความสัมพันธ์ระหว่างความถนัดทางภาษาและความคล่องแคล่วในการใช้คำกับความสามารถในการเขียนเรียงความ โดยใช้แบบทดสอบวัดความถนัดด้านภาษา เป็นแบบทดสอบปรนัย มีทั้งหมด 4 รูปแบบ มีคำตอบที่ถูกต้องเพียงข้อเดียว จำนวน 40 ข้อ ซึ่งผู้วิจัยได้นำมาหาคุณภาพของแบบทดสอบวัดความถนัดด้านภาษาอีกครั้ง มีค่าความเชื่อมั่นทั้ง 4 รูปแบบ ดังนี้ รูปแบบที่ 1 แบบทดสอบศัพท์สัมพันธ์ มีค่าความเชื่อมั่น 0.86 รูปแบบที่ 2 แบบทดสอบจัดประเภท มีค่าความเชื่อมั่น 0.82 รูปแบบที่ 3 แบบทดสอบความเข้าใจภาษา มีค่าความเชื่อมั่น 0.85 รูปแบบที่ 4 แบบทดสอบความเข้าใจภาพ มีค่าความเชื่อมั่น 0.91 จัดทำแบบทดสอบให้เพียงพอแก่นักเรียนที่เป็นกลุ่มตัวอย่าง จากนั้นทำการเก็บรวบรวมข้อมูลจากโรงเรียนที่เป็นกลุ่มตัวอย่างทั้งหมด 5 โรงเรียนในภาคเรียนที่ 2 ปีการศึกษา 2550

หลังจากได้ดำเนินการสอบกลุ่มตัวอย่างครบทุกโรงเรียนที่ทำการสุ่มมาแล้วก็นำกระดาษคำตอบมาตรวจให้คะแนน ดังนี้ ถ้าถูกต้องให้ 1 คะแนน ถ้าตอบผิดหรือไม่ตอบหรือตอบมากกว่า 1 ข้อให้ 0 คะแนน แล้วนำผลมาบันทึกข้อมูลที่ได้จากการทดสอบของกลุ่มตัวอย่างขนาด 564 มาบันทึกเป็นประชากรเทียมสุ่มข้อมูลจากประชากรเทียม ให้มี 4 ขนาด คือ 30 คน 60 คน 90 คน

และ 120 คน ซึ่งแต่ละขนาดจะสุ่มมา 10 กลุ่ม โดยทำการสุ่มข้อมูลแต่ละกลุ่มแต่ละขนาดด้วยการสุ่มแบบใส่คืน นำข้อมูลทุกกลุ่มทุกขนาดมาทดสอบความเป็นเอกพันธ์ของความแปรปรวนของกลุ่มตัวอย่างด้วยสถิติทดสอบเอฟ ถ้ามีกลุ่มตัวอย่างใดที่ไม่เป็นเอกพันธ์ของความแปรปรวนให้ทำการสุ่มแบบใส่คืนใหม่จนครบตามจำนวนกลุ่มที่ต้องการ นำข้อมูลทุกกลุ่มทุกขนาดที่ไม่เป็นเอกพันธ์ของความแปรปรวนมาวิเคราะห์ความแปรปรวน แล้วจึงทดสอบความแตกต่างของค่าเฉลี่ยด้วยสถิติทดสอบทั้ง 5 วิธี คือ วิธี Tukey – Kramer, วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD ด้วยระบบคอมพิวเตอร์กับทุกกลุ่ม จากนั้นทดสอบความแตกต่างของจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญ ที่ได้มาจากการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยที่ต่างกัน 5 วิธี จากกลุ่มตัวอย่างที่มีขนาดเดียวกัน 4 ขนาด โดยวิธี χ^2 -test และทดสอบความแตกต่างของจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญ ในกลุ่มตัวอย่างที่มีขนาดต่างกัน 4 ขนาด ที่ได้จากการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน 5 วิธี โดยวิธี χ^2 -test

หลังจากที่ข้อมูลได้รับการทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี และ 4 ขนาด พร้อมทั้งทดสอบความเป็นเอกพันธ์ของความแปรปรวนแล้ว จึงคำนวณค่าดัชนีความแตกต่างระหว่างค่าเฉลี่ยของประชากร (ϕ) เมื่อได้ค่า ϕ แล้วนำไปหาค่าอำนาจการทดสอบโดยการเปิดตารางของแท็งค์ (Tang. Power Function for Analysis of Variance) (Kirk. 1995 : 816 – 823) ซึ่งจะได้ค่า $1 - \beta$ นำค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธีในแต่ละกลุ่มแต่ละขนาดของกลุ่มตัวอย่างทั้ง 4 ขนาด มาเปรียบเทียบกัน โดยการนำเสนอในรูปแบบตาราง

สรุปผลการศึกษาค้นคว้า

1. จากการเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยด้วยวิธีต่างกันทั้ง 5 วิธี คือ วิธี Tukey-Kramer, วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD เมื่อขนาดของกลุ่มตัวอย่างคงที่ทั้ง 4 ขนาดคือ 30 คน 60 คน 90 คน และ 120 คน พบว่า

1.1 ในกลุ่มตัวอย่างขนาดเดียวกันทั้ง 4 ขนาดคือ 30 คน 60 คน 90 คน และ 120 คน เมื่อทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยต่างกันทั้ง 5 วิธี คือ วิธี Tukey-Kramer, วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD พบว่าให้ค่าอำนาจการทดสอบไม่แตกต่างกันในทุกกลุ่ม

1.2 ในกลุ่มตัวอย่างขนาด 90 คน และ 120 คน สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD มีค่าอำนาจการทดสอบเท่ากันในทุกกลุ่ม และสูงกว่าค่าอำนาจการทดสอบในกลุ่มตัวอย่างขนาด 30 คน และ 60 คน

1.3 ในการเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบความแตกต่างของค่าเฉลี่ยด้วยวิธีต่างกันทั้ง 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD เมื่อขนาดของกลุ่มตัวอย่างคงที่ทั้ง 4 ขนาดคือ 30 คน 60 คน 90 คน และ 120 คน พบว่า ขนาดกลุ่มตัวอย่างที่มีค่าอำนาจการทดสอบมากที่สุดคือ กลุ่มตัวอย่างขนาด 120 คนและ 90 คน รองลงมาคือ กลุ่มตัวอย่างขนาด 60 คน และมีค่าอำนาจการทดสอบน้อยที่สุด คือ กลุ่มตัวอย่างขนาด 30 คน

2. จากการเปรียบเทียบค่าอำนาจการทดสอบเมื่อขนาดของกลุ่มตัวอย่างต่างกัน ทั้ง 4 ขนาด คือ 30 คน 60 คน 90 คน และ 120 คน ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกันทั้ง 5 วิธี คือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD พบว่า

2.1 ในสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกันทั้ง 5 วิธี ค่าอำนาจการทดสอบมีแนวโน้มสูงขึ้นเมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้น

2.2 สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan และ วิธี Scheffe' ให้ค่าอำนาจการทดสอบแตกต่างกันในทุกขนาดของกลุ่มตัวอย่าง ส่วนวิธี Fisher's LSD ให้ค่าอำนาจการทดสอบไม่แตกต่างกันในทุกขนาดของกลุ่มตัวอย่าง

2.3 สถิติทดสอบความแตกต่างของค่าเฉลี่ยที่ให้ค่าอำนาจการทดสอบสูงสุดคือ วิธี Fisher's LSD รองลงมาคือ วิธี Duncan ,วิธี Scheffe', วิธี Newman Keuls และสถิติทดสอบความแตกต่างของค่าเฉลี่ยที่ให้ค่าอำนาจการทดสอบน้อยที่สุด คือ วิธี Tukey-Kramer

อภิปรายผล

อำนาจการทดสอบสามารถใช้ประโยชน์ในการวางแผน ออกแบบการวิจัย ใช้เป็นเกณฑ์ในการตัดสินใจเพื่อเลือกขนาดของกลุ่มตัวอย่าง และกำหนดนัยสำคัญทางสถิติโดยอาศัยการประมาณค่าอำนาจการทดสอบ อำนาจการทดสอบเปรียบเสมือนเครื่องมือที่ใช้วิเคราะห์เพื่อช่วยในการตัดสินใจในเรื่องที่ต้องการศึกษานั้นๆ โดยเฉพาะขนาดของกลุ่มตัวอย่างถือว่าเป็นปัจจัยอย่างหนึ่งที่ส่งผลต่อค่าอำนาจการทดสอบ เมอร์ฟี และเอ็มเมอร์ส (Murphy & Myers. 2004 : 18) ได้กล่าวไว้ว่า อำนาจการทดสอบที่เพียงพอจะต้องพบสองสิ่งคือ อำนาจการทดสอบควรมีค่ามากกว่า 0.50 หากอำนาจ

การทดสอบต่ำกว่า 0.50 การศึกษานี้คล้ายกับว่าประสบความสำเร็จมากกว่าประสบความสำเร็จ และเมื่ออำนาจการทดสอบมีค่ามากกว่า 0.80 ถือว่ามีค่าเพียงพอต่อการศึกษา ซึ่งในการศึกษาค้างนี้ ผู้วิจัยได้แยกการอภิปรายผล เป็น 2 หัวข้อดังนี้

1. อำนาจการทดสอบของสถิติการทดสอบความแตกต่างของค่าเฉลี่ยต่างวิธีกัน ทั้ง 5 วิธีคือ วิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD โดยที่ขนาดของกลุ่มตัวอย่างเท่ากัน ผลจากการศึกษาในครั้งนี้ชี้ให้เห็นว่า อำนาจการทดสอบของ สถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธี มีแนวโน้มสูงขึ้น เมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้น ผลจากการศึกษาค้างนี้สอดคล้องกับการศึกษาของ สมกิต วงศ์มาก (2459 : 148) ศึกษาอำนาจการ ทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียว ภายหลังการแปลงข้อมูลจากกลุ่มตัวอย่าง ขนาดเล็ก โดยใช้เทคนิคการจำลองแบบมอนติคาร์โลและการใช้ข้อมูลจริง จากการศึกษาพบว่าเมื่อ ขนาดกลุ่มตัวอย่างเพิ่มขึ้นมีผลให้ค่าอำนาจการทดสอบเพิ่มขึ้น นอกจากนี้ มนทกานติ หรรษวรพงศ์ (2545 : 145) ศึกษาการเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบความแตกต่างของ ค่าเฉลี่ยของ 2 ประชากร จากการศึกษาพบว่า เมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้นมีผลให้ค่าอำนาจ การทดสอบของสถิติทดสอบเพิ่มขึ้น ผู้วิจัยมีความเห็นว่าเนื่องจากการศึกษาค้างนี้กำหนดขนาดของ กลุ่มตัวอย่างเป็น 30 คน 60 คน 90 คน และ 120 คน ยิ่งกลุ่มตัวอย่างมีขนาดเพิ่มขึ้น การแจกแจงของข้อมูลก็จะเป็นการแจกแจงแบบปกติมากขึ้น และมีแนวโน้มให้ค่าอำนาจการทดสอบสูงขึ้น

2. อำนาจการทดสอบเมื่อขนาดของกลุ่มตัวอย่างต่างกันทั้ง 4 ขนาด คือ 30 คน 60 คน 90 คน และ 120 คน เมื่อทดสอบด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีเดียวกัน ผล การศึกษาค้างนี้ชี้ให้เห็นว่า สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer , วิธี Newman Keuls, วิธี Duncan และ วิธี Scheffe ' ให้ค่าอำนาจการทดสอบแตกต่างกันในทุก ขนาดของกลุ่มตัวอย่าง ส่วน วิธี Fisher's LSD ให้ค่าอำนาจการทดสอบไม่แตกต่างกันในทุกขนาด ของกลุ่มตัวอย่าง สถิติทดสอบความแตกต่างของค่าเฉลี่ยที่ให้ค่าอำนาจการทดสอบสูงสุดคือ วิธี Fisher's LSD รองลงมาคือ วิธี Duncan , วิธี Scheffe ' , วิธี Newman Keuls และสถิติ ทดสอบความแตกต่างของค่าเฉลี่ยที่ให้ค่าอำนาจการทดสอบน้อยที่สุด คือ วิธี Tukey-Kramer จาก การศึกษาค้างนี้สอดคล้องกับการศึกษาของ ลี ยู ฟาง (Li Yu – Fang. 1998 : 6489) ได้ศึกษา พัฒนาอำนาจการประมาณและขนาดกลุ่มตัวอย่างของวิธีการเปรียบเทียบพหุคูณ 7 วิธี ผล การศึกษาพบว่า วิธี Fisher's LSD และวิธี Duncan เป็นวิธีทดสอบที่ดีที่สุดใน 7 วิธีที่นำมาศึกษา ขณะที่ วิธี Bonferoni และ วิธี Sidak เป็นสถิติทดสอบที่มีอำนาจน้อยที่สุด นอกจากนี้ วิไลลักษณ์ องค์จิระวุฒิ (2522 : บทคัดย่อ) ได้ศึกษาเรื่องการเปรียบเทียบวิธีการต่างๆ ที่ใช้ทดสอบความ แตกต่างระหว่างค่าเฉลี่ยของประชากร โดยพิจารณาจากความผิด 3 ชนิด ในการวิเคราะห์เพื่อ

เปรียบเทียบค่าเฉลี่ยของประชากรหลายๆ ประชากรว่ามีความแตกต่างกันอย่างมีนัยสำคัญหรือไม่ ผลการศึกษาพบว่าวิธี Fisher's LSD ให้ผลสรุปหรือผลการทดสอบที่น่าเชื่อถือได้ดีกว่าวิธี Murphy gap LSD และวิธี Scheffe' ผู้วิจัยมีความเห็นว่า สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Tukey-Kramer, วิธี Newman Keuls และ วิธี Scheffe' มีขั้นตอนการทดสอบและข้อตกลงเบื้องต้นเหมือนกัน ส่งผลให้ผลที่ได้มีค่าใกล้เคียงกัน ส่วนวิธี Fisher's LSD เป็นวิธีที่นักวิจัยนิยมกันมาก เพราะ สามารถใช้ทดสอบกับกลุ่มตัวอย่างที่มีขนาดเท่ากันหรือไม่เท่ากันได้ และการเปรียบเทียบจะทำการเปรียบเทียบทีละคู่ ข้อมูลเป็นอิสระจากกันและมีการแจกแจงเป็นแบบปกติ ส่งผลให้อำนาจการทดสอบด้วยวิธี Fisher's LSD มีค่าสูงขึ้น

ข้อเสนอแนะ

ผลการวิจัยครั้งนี้สามารถสรุปเป็นข้อเสนอแนะในการนำไปใช้ในการวิจัยครั้งต่อไปดังนี้

1. ข้อเสนอแนะเพื่อการเลือกใช้สถิติทดสอบความแตกต่างของค่าเฉลี่ยและขนาดของกลุ่มตัวอย่าง

1.1 จากผลการวิจัยพบว่า ในการวิเคราะห์หาความแตกต่างของค่าเฉลี่ยจากคะแนนความถนัดทางภาษาของกลุ่มตัวอย่างขนาด 30 คน 60 คน 90 คน และ 120 คน ด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ย ทั้ง 5 วิธี คือ วิธี Tukey-Kramer, วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD ถ้าต้องการผลการทดสอบว่ามีความแตกต่างระหว่างค่าเฉลี่ยอย่างมั่นใจ มีอำนาจการทดสอบสูง ควรเลือกใช้วิธี Fisher's LSD

1.2 ถ้าผู้วิจัยเลือกใช้สถิติทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 5 วิธีคือ วิธี Tukey-Kramer, วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD มีขนาดของกลุ่มตัวอย่างให้เลือกศึกษา 4 ขนาด คือ 30 คน 60 คน 90 คน และ 120 คน ผลการศึกษาพบว่าขนาดของกลุ่มตัวอย่างเพิ่มขึ้นส่งผลให้อำนาจการทดสอบสูงขึ้น ดังนั้นควรจะเลือกใช้กลุ่มตัวอย่างที่มีขนาด 120 คน เพราะจะให้ค่าอำนาจการทดสอบสูงกว่าขนาดอื่น แต่เนื่องด้วยปัจจัยต่างๆ ในการทำการวิจัย ไม่ว่าจะเป็นระยะเวลา ค่าใช้จ่าย อาจเลือกใช้กลุ่มตัวอย่างขนาด 90 คน ได้ เพราะให้ค่าอำนาจการทดสอบได้ใกล้เคียงกัน ทั้งยังเป็นการประหยัดเวลาและลดค่าใช้จ่ายได้บ้าง อีกกรณีหนึ่ง คือ เมื่อกลุ่มตัวอย่างมีขนาด 30 คน และ 60 คน ถึงแม้กลุ่มตัวอย่างขนาด 60 คน จะมีอำนาจการทดสอบสูงกว่าแต่ให้ค่าอำนาจการทดสอบไม่แตกต่างกันมากนักกับกลุ่มตัวอย่างขนาด 30 คน ดังนั้นผู้วิจัยเห็นว่าควรจะเลือกใช้กลุ่มตัวอย่างขนาด 30 คน เพราะเป็นขนาดของกลุ่มตัวอย่างใน 1 ห้องเรียนที่พบได้โดยทั่วไป สะดวกในการเก็บข้อมูลประหยัดเวลาและลดค่าใช้จ่ายได้

2. ข้อเสนอแนะในการวิจัย

2.1 นอกจากศึกษาขนาดของกลุ่มตัวอย่างที่ต่างกันแล้ว ควรศึกษาภายใต้เงื่อนไขอื่น เช่น การแจกแจงที่ต่างกัน ณ ระดับนัยสำคัญทางสถิติที่ต่างกัน

2.2 ควรศึกษาด้วยสถิติทดสอบความแตกต่างของค่าเฉลี่ยที่ต่างกันภายใต้ความแปรปรวนที่ต่างกัน

บรรณานุกรม

บรรณานุกรม

- กัลยา วานิชย์บัญชา. (2539). *การวิเคราะห์สถิติเพื่อการตัดสินใจ*. กรุงเทพฯ : ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย.
- ชวาล แพรัตกุล. (2517). *การทดสอบเพื่อค้นหาและพัฒนาสมรรถภาพ*. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา วิทยาลัยการศึกษา ประสานมิตร.
- จิตติรัตน์ ฐานวัฒน์. (2542). *การศึกษาความสัมพันธ์ระหว่างความสามารถทางสมองกับผลสัมฤทธิ์ทางการเรียนด้านการบัญชี นักศึกษาสาขาการบัญชี. ปริญญาโท กศ.ม. (การวัดผลการศึกษา)*. กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
- นันทา วงษ์วิโรจน์. (2532). *สถิติทดสอบที่มีความแกร่งสำหรับทดสอบความเท่ากันของค่าเฉลี่ยประชากร*. วิทยานิพนธ์ พ.ม. (สถิติ) . กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- บุญชม ศรีสะอาด. (2538). *วิธีการทางสถิติสำหรับวิธีการวิจัย 2*. กรุงเทพฯ : สุวีริยาสาส์น.
- (2541). *วิธีการทางสถิติสำหรับการวิจัย. พิมพ์ครั้งที่ 2*. กรุงเทพฯ : สุวีริยาสาส์น
- บุญเชิด ภิญโญนนตพงษ์. (2527). *การทดสอบอิงเกณฑ์แนวคิดและวิธีการ*. พิมพ์ครั้งที่ 1. กรุงเทพฯ : สำนักพิมพ์โอเดียนสโตร์.
- บุญเรียง ขจรศิลป์. (2533). *สถิติวิจัย 2. ฟิสิกส์เซนเตอร์ การพิมพ์. พิมพ์ครั้งที่ 1*.
- บุปผา จงเลิศตระกูล. (2532). *การเปรียบเทียบเชิงพหุระหว่างประชากรใดๆ กับประชากร ควบคุมภายใต้การแจกแจงแบบเอกซ์โพเนนเชียล*. วิทยานิพนธ์ พ.ม. (สถิติ) . กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- พหล ศักดิ์กะทัศน์. (2534). *การเปรียบเทียบประสิทธิภาพของวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยในการวิเคราะห์ความแปรปรวน*. วิทยานิพนธ์ พ.ม. (สถิติ) . กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- เพ็ญศรี เตมีกุล. (2525). *การสร้างแบบทดสอบวัดความถนัดทางภาษาไทย สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในจังหวัดสิงห์บุรี*. ปริญญาโท กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
- มนทกานติ หรรษวรพงศ์. (2545). *การเปรียบเทียบอำนาจของตัวสถิติทดสอบความแตกต่างระหว่างค่าเฉลี่ยของ 2 ประชากร*. ปริญญาโท พ.ม.(สถิติ) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.

- ระวีวรรณ พันธุ์พานิชย์. (2540). แบบแผนเชิงสถิติของการทดสอบ. กรุงเทพฯ : ภาควิชาวัดผลและ
วิจัยทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- . (2541). สถิติเพื่อการวิจัย. ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ.
- รุ่งโรจน์ ศรีจันทร์แก้ว. (2547). การเปรียบเทียบอำนาจการทดสอบ (Power of Test) ของผลการ
วิเคราะห์ข้อมูลด้วยแบบแผนการวิเคราะห์แบบกลุ่มสุ่ม (RBD) กับแบบแผนการวิเคราะห์
ความแปรปรวนร่วม (ANCOVA). ปรินูญานิพนธ์ กศ.ม. (การวิจัยและสถิติทางการ
ศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
- เลิศลักษณ์ กลิ่นหอม. (2532). การศึกษาความคลาดเคลื่อนชนิดที่ 1 และอำนาจการทดสอบใน
แบบแผนการทดลองแบบสุ่มภายในบล็อกที่มีปฏิสัมพันธ์ระหว่างสิ่งทดลองกับบล็อก.
ปรินูญานิพนธ์ กศ.ด. กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
ถ่ายเอกสาร.
- เยาวภา ไชยศรี. (2536). การเปรียบเทียบประสิทธิภาพของตัวทดสอบสถิติบาร์เลด เลอวิน และ
บราว-ฟอร์ลิตี สำหรับทดสอบความเท่ากันของความแปรปรวนของประชากร. วิทยานิพนธ์
วท.ม.(สถิติ). กรุงเทพฯ : มหาวิทยาลัยเกษตรศาสตร์. ถ่ายเอกสาร.
- ล้วน สายยศ และอังคณา สายยศ. (2538). เทคนิคการวิจัยทางการศึกษา. กรุงเทพฯ :
สุวีริยาสาส์น.
- . (2539). เทคนิคการวัดผลการเรียนรู้. กรุงเทพฯ : สุวีริยาสาส์น.
- . (2540). สถิติวิทยาทางการวิจัย. พิมพ์ครั้งที่ 3. กรุงเทพฯ : สุวีริยาสาส์น.
- . (2541) .เทคนิคการสร้างและสอบข้อสอบความถนัดทางการเรียน. กรุงเทพฯ:
สุวีริยาสาส์น.
- วิชชุดา ศรีโสภา. (2539). การเปรียบเทียบพหุ. วิทยานิพนธ์ พ.ม. (สถิติ) . กรุงเทพฯ :
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- วิรัช วรรณรัตน์. (2535). วิธีการทางสถิติสำหรับการวิจัย. กรุงเทพฯ : สำนักทดสอบทางการศึกษา
และจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ
- วิไลลักษณ์ องค์จิระวุฒิ. (2522). การเปรียบเทียบวิธีการต่างๆ ที่ใช้ในการทดสอบความแตกต่าง
ระหว่างค่าเฉลี่ยของประชากร โดยพิจารณาจากความผิด 3 ชนิด. วิทยานิพนธ์ พ.ม.
(สถิติ) . กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.

- ศิริพร กิตติสุภาพ.(2544). ผลกระทบของการแจกแจงข้อมูลที่มีต่อการทดสอบความแตกต่างของค่าเฉลี่ยโดยใช้วิธี *Tukey's HSD* , วิธี *Scheffe'* และวิธี *Brown – Forsythe*.
ปริญญาานิพนธ์ กศ.ม. (การวิจัยและสถิติทางการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ศิริชัย กาญจนวาสี และคณะ. (2537). การเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัย. กรุงเทพฯ : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- สมกิต วงศ์มาก.(2549). การศึกษาอำนาจการทดสอบของการวิเคราะห์ความแปรปรวนแบบทางเดียวภายหลังการแปลงข้อมูลจากกลุ่มตัวอย่างขนาดเล็กโดยใช้เทคนิคอนติคาร์โลและใช้ข้อมูลจริง. วิทยานิพนธ์ กศ.ม. (การวิจัยและสถิติทางการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
- สินุช เอี่ยมเขียว.(2544). ผลการเปรียบเทียบพหุคูณรายคู่บางวิธีภายใต้ความแปรปรวนที่แตกต่างกัน จากกลุ่มตัวอย่างขนาดต่างกัน. ปริญญาานิพนธ์ กศ.ม. (การวิจัยและสถิติทางการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ
- สุกัญญา หนูกล้า. (2542). อำนาจการทดสอบของตัวสถิติทดสอบค่าเฉลี่ยเมื่อประชากรมีการแจกแจงแบบเบ้ขวา. ปริญญาานิพนธ์ พ.ม. (สถิติ) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- โสภา ศรีสุข. (2547). การศึกษาความสัมพันธ์ระหว่างความถนัดทางภาษาและความคล่องแคล่วในการใช้คำกับความสามารถในการเขียนเรียงความ ของนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนวัดอุดมรังสี. สารนิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
- Brown, M.B. ; & Forsythe, A.B. (1974b). *The ANOVA and multiple comparisons for data with heterogeneous variances*. *Biometrics*, 30. 719 -724.
- Blair, Clifford R.; & James, Hinggin J.(1980). “ *A comparison of The Power of Willcox's Rank-Sum Statistic to That of Student's t Statistic Under Various nonnormal Distributions*”. *Journal of Educational Statistics*. 5: 309 – 335.
- Cochran. W.G.;& Cox, G.M. (1957). *Experimental Designs*. New York : Wiley.

- Cohen, Jacob. (1973). *Statistical power analysis and Research Results*. American Educational Research Journal. 10, 225 – 230.
- . (1977). *Statistical power analysis for the behavioral sciences*. (rev.ed) New York : Academic Press.
- . (1988). *Statistical power analysis (2nd ed)*. Hillsdale. NJ : Erlbaum
- Cronbach, L J.;N. Rajaratnam. ;& G.C. Gleser.(1970). *Essential of Psychological Testing*. New York : Harper ;& Row Publishers.
- Dayton, C. Mitchell. (1970). *The Design of Educational Experiment*. Mc. Grae-Hill University of Marryland. 32 -37.
- Douglas A. Penfield ;& Stephen L. Koffler. (1978). “ *Post Hoc Procedures For Some K-sample Nonparametric Tests for Scale*”. Journal of Education Statistic. 3: 265 – 282.
- Games; & Howell. (1976). “*Pairwise multiple comparison procedures with unequal N's and/or variance*”. Journal of Education Statistics. 1: 113 – 125.
- Guildford, J.P. ;& Fruchteer, B. (1978). *Fundamental Statistics in Psychology and Education*. Singapore : McGraw-Hill .
- Hay , William L.(1963). *Statistic for Psychologists*. New York : Holt.
- Hinkle,D.E. ; & Oliver J.D. (1983).*How Large Should the Sample Be? A Question With Nosimple Answer ? Educational and Psychological Measurement*. 43 : 1051 1060.
- Howell, David C. (1992). *Statistic Method for Psychology*. 3 rd ed. Belmont, CA: Duxbury Pr.
- Jefferey , D. Kromerey ;& Michela , A. La Rocca. (1995). “*Power and Type I error Rate of New multiple Comparison Procedures Under Heterogeneous Variances*”. Journal of the American Statistical Association. 63 : 343 – 362.
- Keppel, Geoffrcy. (1991). *Design and Analysis : A researcher's Handbook*. Englewood Cliffs. NJ : Prentice Hall.
- Kerlinger , F . N. (1986) .*Fundamentals of Behavioral Research*. Florida : Holt, Rinehart and Winston.

- Kirk, Roger E. (1982). *Experimental design : procedures for the behavioral sciences*. 2nd ed. California : Brooks/Cole Publishing Company.
- (1995). *Experimental design : procedures for the behavioral sciences*. 3rd ed. California : Brooks/Cole Publishing Company.
- Klockars , Alan J.;& Sax. (1986). *Multiple Comparisons*. Sage Publications Inc.
- Levene, H. (1960). *Robust test for equality of variances*. Contributions to probability and Statistics . Essaya in hour of Horold Hotelling. Stanford, CA : Stanford University. 278 – 292 , Press.
- Li Yu – Fang. (1998). “ *The Development of power Estimates and Sample Size Requirements for Seven Multiple Comparisons Procedures*”. Dissertation Abstracts.
- Minium, Edward W. ; King, Bruce M.; & Bear, Gondon. (1993). *Statistic Reasoning in Psychology and Education*. 3rd ed. John Wiley & Sons.
- Murphy, Kevin R.; & Myers, Brett. (2004). *Statistic Power Analysis*. 2nd ed. New Jersey : Lawrence Erlbaum.
- Myer, Jerome L. ; & Well, Arnold D. (2003). *Research Design and Statistic Analysis*. 2nd ed. New Jersey : Lawrence Erlbaum.
- Pagano, Robert R. (2004). *Understanding Statistics in The Behavioral Science*. 7th ed. Ontario : Thomson Learning.
- Pearson, E.S.; & Hartley, H.O. (1951). *Chart of the power function for analysis of variance test*. Derived from the noncentral F- Distribution. Biometrika. 38, 112 – 130 [816].
- Ramsey ,Philip H. (1978). “*Power Differences Between Pairwise multiple Comparison*”. Journal of the American Statistic Association. 73 : 479 – 487.
- Richard C. Sprinthall. (1994). *Basic Statistical Analysis*. A Division of Paramount Publishing.
- Rosenthal, Robert.; & Rosnow, Ralph L. (1991). *Essentials of Behavioral Research : Method and Data Analysis*. 2nd ed. New York : McGraw-Hill.
- Runyon, Richard P.; et al. (1996). *Fundamentals of Behavioral Statistic*. 7th ed. McGraw-Hill.

- Runyon, Richard P.; & Haber, Audrey. (1991). *Fundamentals of Behavioral Statistics*. 7th ed. McGraw-Hill.
- Scheffe, H. (1953). "A method for judging all contrasts in the analysis of variance". *Biometrika*, 40. 87-104.
- (1959). *The analysis of variance*. New York : Wiley. [101 , 173 , 210, 240, 283, 330, 335, 399, 408 , 411].
- (1970). "Practical solutions of the Behrens-Fisher problem". *Journal of the American Statistical Association*, 65.1501 – 1508.
- Stevens, James P. (2002). *Applied Multivariate Statistics for The Social Science*. 4 th ed. New Jersey : Lawrence Erlbaum.
- Toothaker, Larry E. L.E. (1993). *Multiple comparison Procedures*. Sage Publication, Inc. University of Oklahoma.
- Tukey, J.W. (1949). "One degree of freedom for nonadditivity". *Biometrics*, 5. 232 – 242. [104, 269].
- (1953). *The problem of multiple comparisons*. Unpublished manuscript. Princeton University, Princeton, NJ. [120, 144, 146, 173, 399]
- (1955). *Queries. Biometrics*. 11. [111] – 113. [330].
- (1977). *Exploratory data analysis*. Reading, M.A. : Addison – Wesley. [166]
- Weiss. Neil A. (1995). *Introduction Statistics*. Addison – Wesley Publishing Company, Inc.
- Welkowitz, Joan.; Ewen, Robert B.; & Cohen, Jacob. (2000). *Introductory Statistics for The Behavioral Science*. 5 th ed. Orlando : Harcourt Brace.
- Wilcox, Rand R. (1989). "Adjusting for Unequal Variance When Comparing Means in One-Way and Two-Way Fixed Effects ANOVA Models". *Journal of Educational Statistics*. 14:270 – 278.
- Yosef Hochberg.; & Tamhane, Ajit C. (1987). *Multiple comparison procedures*. New York : John Wiley & Sons.

ภาคผนวก

ภาคผนวก ก.

- ค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Tukey – Kramer จากกลุ่มตัวอย่าง 4 ขนาด
- ค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Newman Keuls จากกลุ่มตัวอย่าง 4 ขนาด
- ค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Duncan จากกลุ่มตัวอย่าง 4 ขนาด
- ค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Scheffe ' จากกลุ่มตัวอย่าง 4 ขนาด
- ค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Fisher's LSD จากกลุ่มตัวอย่าง 4 ขนาด
- จำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี Tukey – Kramer จากกลุ่มตัวอย่าง 4 ขนาด
- จำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี Newman Keuls จากกลุ่มตัวอย่าง 4 ขนาด
- จำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี Duncan จากกลุ่มตัวอย่าง 4 ขนาด
- จำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี Scheffe ' จากกลุ่มตัวอย่าง 4 ขนาด
- จำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี Fisher's LSD จากกลุ่มตัวอย่าง 4 ขนาด

ตาราง 20 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Tukey – Kramer จาก
กลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ระหว่างชั้น	ขนาดกลุ่มตัวอย่าง (คน)			
		30	60	90	120
1	1,2	-4.33	-.63	-4.84*	-.71
2	1,3	-8.80*	-9.28*	-9.86*	-8.63*
3	2,3	-9.23*	-8.65*	-5.02*	-7.91*
4	1,2	-2.86	-6.13*	-3.40*	-.31
5	1,3	-9.56*	-9.93*	-11.72*	-9.75*
6	2,3	-6.70	-3.80*	-8.32*	-9.44*
7	1,2	-1.20	-6.30*	-1.84	-6.22*
8	1,3	-6.9*	-11.06*	-9.73*	-9.46*
9	2,3	-5.70*	-4.76*	-7.88*	-3.24*
10	1,2	-6.66*	-9.51*	-4.35*	-6.70*
11	1,3	-9.76*	-11.53*	-12.61*	-11.73*
12	2,3	-3.10	-2.01	-8.25*	-5.02*
13	1,2	-9.06*	-10.33*	-9.41*	-10.19*
14	1,3	-12.30*	-12.61*	-14.12*	-12.98*
15	2,3	-3.23*	-2.28	-4.71*	-2.79*
16	1,2	-11.23	-11.80*	-9.87*	-9.69*
17	1,3	-12.63*	-12.18*	-13.97*	-12.09*
18	2,3	-1.4	-.38	-4.10*	-2.40*
19	1,2	-12.96*	-7.60*	-8.51*	-12.63*
20	1,3	-10.63*	-9.51*	-10.97*	-13.71*
21	2,3	-.233	-1.91	-2.46*	-1.08
22	1,2	-12.96*	-5.60*	-9.64*	-7.34*
23	1,3	-10.63*	-9.71*	-10.05*	-9.08*
24	2,3	-2.33	-4.11*	-.41	-1.74
25	1,2	-4.33*	-7.18*	-9.18*	-5.82*
26	1,3	-9.76*	-7.56*	-10.22	-10.27*
27	2,3	-5.43*	-.38	-1.03	-4.45*
28	1,2	-6.76*	-3.63*	-8.70*	-7.40
29	1,3	-5.66*	-7.13*	-9.60*	-9.13*
30	2,3	-1.10	-3.50	-.90	-1.72*

ตาราง 21 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Newman Keuls
จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ระหว่างชั้น	ขนาดกลุ่มตัวอย่าง (คน)			
		30	60	90	120
1	1,2	-.433	-.63	-4.84*	-.71
2	1,3	-8.80*	-9.28*	-9.86*	-8.63*
3	2,3	-9.23*	-8.65*	-5.02*	-7.91*
4	1,2	-2.86	-6.13*	-3.40*	-.31
5	1,3	-9.56*	-9.93*	-11.72*	-9.75*
6	2,3	-6.70	-3.80*	-8.32*	-9.44*
7	1,2	-1.20	-6.30*	-1.84	-6.22*
8	1,3	-6.9*	-11.06*	-9.73*	-9.46*
9	2,3	-5.70*	-4.76*	-7.88*	-3.24*
10	1,2	-6.66*	-9.51*	-4.35*	-6.70*
11	1,3	-9.76*	-11.53*	-12.61*	-11.73*
12	2,3	-3.10	-2.01	-8.25*	-5.02*
13	1,2	-9.06*	-10.33*	-9.41*	-10.19*
14	1,3	-12.30*	-12.61*	-14.12*	-12.98*
15	2,3	-3.23*	-2.28	-4.71*	-2.79*
16	1,2	-11.23	-11.80*	-9.87*	-9.69*
17	1,3	-12.63*	-12.18*	-13.97*	-12.09*
18	2,3	-1.4	-.38	-4.10*	-2.40*
19	1,2	-12.96*	-7.60*	-8.51*	-12.63*
20	1,3	-10.63*	-9.51*	-10.97*	-13.71*
21	2,3	-.233	-1.91	-2.46*	-1.08
22	1,2	-12.96*	-5.60*	-9.64*	-7.34*
23	1,3	-10.63*	-9.71*	-10.05*	-9.08*
24	2,3	-2.33	-4.11*	-.41	-1.74
25	1,2	-4.33*	-7.18*	-9.18*	-5.82*
26	1,3	-9.76*	-7.56*	-10.22	-10.27*
27	2,3	-5.43*	-.38	-1.03	-4.45*
28	1,2	-6.76*	-3.63*	-8.70*	-7.40
29	1,3	-5.66*	-7.13*	-9.60*	-9.13*
30	2,3	-1.10	-3.50	-.90	-1.72*

ตาราง 22 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Duncan จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ระหว่างชั้น	ขนาดกลุ่มตัวอย่าง (คน)			
		30	60	90	120
1	1,2	-.433	-.63	-4.84*	-.71
2	1,3	-8.80*	-9.28*	-9.86*	-8.63*
3	2,3	-9.23*	-8.65*	-5.02*	-7.91*
4	1,2	-2.86	-6.13*	-3.40*	-.31
5	1,3	-9.56*	-9.93*	-11.72*	-9.75*
6	2,3	-6.70	-3.80*	-8.32*	-9.44*
7	1,2	-1.20	-6.30*	-1.84	-6.22*
8	1,3	-6.9*	-11.06*	-9.73*	-9.46*
9	2,3	-5.70*	-4.76*	-7.88*	-3.24*
10	1,2	-6.66*	-9.51*	-4.35*	-6.70*
11	1,3	-9.76*	-11.53*	-12.61*	-11.73*
12	2,3	-3.10	-2.01	-8.25*	-5.02*
13	1,2	-9.06*	-10.33*	-9.41*	-10.19*
14	1,3	-12.30*	-12.61*	-14.12*	-12.98*
15	2,3	-3.23*	-2.28	-4.71*	-2.79*
16	1,2	-11.23	-11.80*	-9.87*	-9.69*
17	1,3	-12.63*	-12.18*	-13.97*	-12.09*
18	2,3	-1.4	-.38	-4.10*	-2.40*
19	1,2	-12.96*	-7.60*	-8.51*	-12.63*
20	1,3	-10.63*	-9.51*	-10.97*	-13.71*
21	2,3	-.233	-1.91	-2.46*	-1.08
22	1,2	-12.96*	-5.60*	-9.64*	-7.34*
23	1,3	-10.63*	-9.71*	-10.05*	-9.08*
24	2,3	-2.33	-4.11*	-.41	-1.74
25	1,2	-4.33*	-7.18*	-9.18*	-5.82*
26	1,3	-9.76*	-7.56*	-10.22	-10.27*
27	2,3	-5.43*	-.38	-1.03	-4.45*
28	1,2	-6.76*	-3.63*	-8.70*	-7.40
29	1,3	-5.66*	-7.13*	-9.60*	-9.13*
30	2,3	-1.10	-3.50	-.90	-1.72*

ตาราง 23 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Scheffe'
จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ระหว่างชั้น	ขนาดกลุ่มตัวอย่าง (คน)			
		30	60	90	120
1	1,2	-.43	-.63	-4.84*	-.71
2	1,3	-8.80*	-9.28*	-9.86*	-8.63*
3	2,3	-9.23*	-8.65*	-5.02*	-7.91*
4	1,2	-2.86	-6.13*	-3.40*	-.31
5	1,3	-9.56*	-9.93*	-11.72*	-9.75*
6	2,3	-6.70*	-3.80*	-8.32*	-9.44*
7	1,2	-1.20	-6.30*	-1.84	-6.22*
8	1,3	-6.90*	-11.06*	-9.73*	-9.46*
9	2,3	-5.70*	-4.76*	-7.88*	-3.24*
10	1,2	-6.66*	-9.51*	-4.35*	-6.70*
11	1,3	-9.76*	-11.53*	-12.61*	-11.73*
12	2,3	-3.10*	-2.01	-8.25*	-5.02*
13	1,2	-9.06*	-10.33*	-9.41*	-10.19*
14	1,3	-12.30*	-12.61*	-14.12*	-12.98*
15	2,3	-3.23	-2.28*	-4.71*	-2.79*
16	1,2	-11.23*	-11.80*	-9.87*	-9.69*
17	1,3	-12.63*	-12.18*	-13.97*	-12.09*
18	2,3	-1.40*	-.38	-4.10*	-2.40*
19	1,2	-12.96*	-7.60*	-8.51*	-12.63*
20	1,3	-10.63*	-9.51*	-10.97*	-13.71*
21	2,3	-2.33	-1.91	-2.46*	-1.08
22	1,2	-12.96*	-5.60*	-9.64*	-7.34*
23	1,3	-10.63*	-9.71*	-10.05*	-9.08*
24	2,3	-2.33	-4.11*	-.41	-1.74
25	1,2	-4.33*	-7.18*	-9.18*	-5.82*
26	1,3	-9.76*	-7.56*	-10.22*	-10.27*
27	2,3	-5.43*	-.38	-1.03	-4.45*
28	1,2	-6.76*	-3.63*	-8.70*	-7.40*
29	1,3	-5.66*	-7.13*	-9.60*	-9.13*
30	2,3	-1.10	-3.50*	-.90	-1.72

ตาราง 24 แสดงค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี Fisher's LSD
จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ระหว่างชั้น	ขนาดกลุ่มตัวอย่าง (คน)			
		30	60	90	120
1	1,2	-.43	-.63	-4.84*	-.71
2	1,3	-8.80*	-9.28*	-9.86*	-8.63*
3	2,3	-9.23*	-8.65*	-5.02*	-7.91*
4	1,2	-2.86*	-6.13*	-3.40*	-.31
5	1,3	-9.56*	-9.93*	-11.72*	-9.75*
6	2,3	-6.70*	-3.80*	-8.32*	-9.44*
7	1,2	-1.20	-6.30*	-1.84*	-6.22*
8	1,3	-6.90*	-11.06*	-9.73*	-9.46*
9	2,3	-5.70*	-4.76*	-7.88*	-3.24*
10	1,2	-6.66*	-9.51*	-4.35*	-6.70*
11	1,3	-9.76*	-11.53*	-12.61*	-11.73*
12	2,3	-3.10	2.01	-8.25*	-5.02*
13	1,2	-9.06*	-10.33*	-9.41*	-10.19*
14	1,3	-12.30*	-12.61*	-14.12*	-12.98*
15	2,3	-3.23*	-2.28*	-4.71*	-2.79*
16	1,2	-11.23*	-11.80*	-9.87*	-9.69*
17	1,3	-12.63*	-12.18*	-13.97*	-12.09*
18	2,3	-1.40	-.38	-4.10*	-2.40*
19	1,2	-12.96*	-7.60*	-8.51*	-12.63*
20	1,3	-10.63*	-9.51*	-10.97*	-13.71*
21	2,3	2.33	-1.91	-2.46*	-1.08
22	1,2	-12.96*	-5.60*	-9.64*	-7.34*
23	1,3	-10.63*	-9.71*	-10.05*	-9.08*
24	2,3	-2.33	-4.11*	-.41	-1.74*
25	1,2	-4.33*	-7.18*	-9.18*	-5.82*
26	1,3	-9.76*	-7.56*	-10.22*	-10.27*
27	2,3	-5.43*	-.38	-1.03	-4.45*
28	1,2	-6.76*	-3.63*	-8.70*	-7.40*
29	1,3	-5.66*	-7.13*	-9.60*	-9.13*
30	2,3	-1.10	-3.50*	-.90	-1.72*

ตาราง 25 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี

Tukey – Kramer จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ขนาดกลุ่มตัวอย่าง (คน)			
	30	60	90	120
1	0	0	1	0
2	1	1	1	1
3	1	1	1	1
4	0	1	1	0
5	1	1	1	1
6	1	1	1	1
7	0	1	0	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	0	0	1	0
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	0	0	1	1
19	1	1	1	1
20	1	1	1	1
21	0	0	1	0
22	1	1	1	1
23	1	1	1	1
24	0	1	0	0
25	1	1	1	1
26	1	1	1	1
27	1	0	0	1
28	1	1	1	1
29	1	1	1	1
30	0	1	0	1

ตาราง 26 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี
Newman Keuls จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ขนาดกลุ่มตัวอย่าง (คน)			
	30	60	90	120
1	0	0	1	0
2	1	1	1	1
3	1	1	1	1
4	0	1	1	0
5	1	1	1	1
6	1	1	1	1
7	0	1	0	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	0	0	1	0
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	0	0	1	1
19	1	1	1	1
20	1	1	1	1
21	0	0	1	0
22	1	1	1	1
23	1	1	1	1
24	0	1	0	0
25	1	1	1	1
26	1	1	1	1
27	1	0	0	1
28	1	1	1	1
29	1	1	1	1
30	0	1	0	1

ตาราง 27 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี Duncan จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ขนาดกลุ่มตัวอย่าง (คน)			
	30	60	90	120
1	0	0	1	0
2	1	1	1	1
3	1	1	1	1
4	0	1	1	0
5	1	1	1	1
6	1	1	1	1
7	0	1	0	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	0	0	1	0
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	0	0	1	1
19	1	1	1	1
20	1	1	1	1
21	0	0	1	0
22	1	1	1	1
23	1	1	1	1
24	0	1	0	0
25	1	1	1	1
26	1	1	1	1
27	1	0	0	1
28	1	1	1	1
29	1	1	1	1
30	0	1	0	1

ตาราง 28 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี
Scheffe' จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ขนาดกลุ่มตัวอย่าง (คน)			
	30	60	90	120
1	0	0	1	0
2	1	1	1	1
3	1	1	1	1
4	0	1	1	0
5	1	1	1	1
6	1	1	1	1
7	0	1	0	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	0	0	1	1
13	1	1	1	1
14	1	1	1	1
15	0	1	1	1
16	1	1	1	1
17	1	1	1	1
18	0	0	1	1
19	1	1	1	1
20	1	1	1	1
21	0	0	1	0
22	1	1	1	1
23	1	1	1	1
24	0	1	0	0
25	1	1	1	1
26	1	1	1	1
27	1	0	0	1
28	1	1	1	1
29	1	1	1	1
30	0	1	0	0

ตาราง 29 แสดงจำนวนผลการทดสอบความแตกต่างของค่าเฉลี่ย ที่ทำการทดสอบด้วยวิธี
Fisher's LSD จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ขนาดกลุ่มตัวอย่าง (คน)			
	30	60	90	120
1	0	0	0	0
2	1	1	1	1
3	1	1	1	1
4	1	1	1	0
5	1	1	1	1
6	1	1	1	1
7	0	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	0	0	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	0	0	1	1
19	1	1	1	1
20	1	1	1	1
21	0	0	1	0
22	1	1	1	1
23	1	1	1	1
24	0	1	0	1
25	1	1	1	1
26	1	1	1	1
27	1	0	0	1
28	1	1	1	1
29	1	1	1	1
30	0	1	0	1

ภาคผนวก ข
แบบทดสอบที่ใช้ในการวิจัย

แบบทดสอบวัดความถนัดทางภาษา แบ่งเป็น 4 รูปแบบ

รูปแบบที่ 1 แบบศัพท์สัมพันธ์ หมายถึง ชุดของข้อคำถามที่ใช้ในการวัดความสามารถในการหาความสัมพันธ์ของคำต่างๆที่กำหนดให้ว่ามีความสัมพันธ์กับตัวเลือกใดที่กำหนดให้มากที่สุด

คำชี้แจง จากข้อ 1 – 10 ให้นักเรียนเลือกคำที่เกี่ยวข้องกับคำที่กำหนดให้มากที่สุด

1. บุปผา

1. นก
2. ผึ้ง
3. ปลวก
4. ค้างคาว
5. แมลงทับ

2. กังหัน

1. ครู
2. ชาวนา
3. วิศวกร
4. พระสงฆ์
5. พระเจ้าอยู่หัว

3. โยกย้าย

1. คน
2. ที่อยู่
3. ฐานะ
4. ตำแหน่ง
5. เงินเดือน

4. จนมุม

1. คดี
2. ศาล
3. โจทก์
4. จำเลย
5. ตำรวจ

5. ร่ม

1. พายุ
2. แสงไฟ
3. แสงแดด
4. แสงดาว
5. แสงเดือน

6. นกเอี้ยง

1. ม้า
2. วัว
3. แพะ
4. ช้าง
5. ควาย

7. กรรไกร

1. ไม้
2. หิน
3. ท่อน้ำ
4. กระดาษ
5. กระเบื้อง

8. พริ้ว

1. ยิง
2. ต่อย
3. ตัด
4. เป่า
5. เย็บ

9. กลัวย

1. วัว
2. ม้า
3. ลิง
4. สิงโต
5. ควาย

10. นักเรียน

1. มีด
2. จอบ
3. เสียม
4. เลื่อย
5. ดินสอ

รูปแบบที่ 2 แบบทดสอบแบบจัดประเภท หมายถึง ชุดของข้อคำถามที่ใช้วัดความสามารถด้านความเข้าใจ ตลอดจนการวิเคราะห์พิจารณาประเมินดูว่าประโยคหรือวลีใดที่ให้ไว้มีความหมายผิดแผกไปจากข้ออื่นๆ

คำชี้แจง ให้นักเรียนเลือกประโยคหรือวลีที่มีความหมายผิดแผกไปจากข้ออื่นจากประโยคหรือวลีที่กำหนดให้ต่อไปนี้

- | | |
|---|--|
| <p>1. 1. ปลายทางชีวิต</p> <p>2. ปลายทางแห่งรัก</p> <p>3. ปลายทางแห่งฝัน</p> <p>4. ส่งให้ถึงปลายทาง</p> <p>5. จุดหมายปลายทาง</p> <p>2. 1. ฟุ้งเต็มบริเวณ</p> <p>2. ปลิวฟุ้งไปทั่ว</p> <p>3. ฟุ้งเพื่อเห่อเหิม</p> <p>4. ลอยฟุ้งจนจับไม่ได้</p> <p>5. ฟุ้งกระจายเต็มไปหมด</p> <p>3. 1. ชั่วชีวิตที่มี</p> <p>2. ชั่วกับชั่วกัลป์</p> <p>3. ชั่วกาลปาวสาน</p> <p>4. ชั่วฟ้าดินสลาย</p> <p>5. ชั่วเจ็ดทีดีเจ็ดหน</p> <p>4. 1. ฉันรักเธอนะ</p> <p>2. เธอรักฉันนะ</p> <p>3. นะเธอ นะรักฉัน</p> <p>4. รักฉันเถอะนะเธอ</p> <p>5. ฉันอยากให้เธอรักฉัน</p> <p>5. 1. พ่อขานหนูหยิบกระป๋าให้ค่ะ</p> <p>2. หนูหยิบสมุดงานให้พี่หน่อย</p> <p>3. หยิบขนมให้แม่หน่อยค่ะลูก</p> <p>4. หนูหยิบปากกามานะค่ะแม่</p> <p>5. เพื่อนหนูทำงานหยิบหยิ่งจริง</p> | <p>6. 1. พ่อส่งมาเลยคะ</p> <p>2. เดี่ยวเราส่งให้เองจ๊ะ</p> <p>3. เธอพูดจาส่งเดชไปได้</p> <p>4. ฉันไปส่งเธอที่บ้านนะจ๊ะ</p> <p>5. นำส่งทางไปรษณีย์นะจ๊ะ</p> <p>7. 1. เธอนี้ทำตัวเป็นกาฝากจริง</p> <p>2. ฉันฝากของไว้ประเดี๋ยวนะ</p> <p>3. พ่อนำเงินไปฝากที่ธนาคาร</p> <p>4. แม่ฉันฝากขนมมาให้เธอด้วย</p> <p>5. ฝากฟ้าไปบอกเขาว่าคิดถึงจัง</p> <p>8. 1. จำใจมาละสิ</p> <p>2. ช่างจดจำจริงนะ</p> <p>3. เธอจำฉันได้จริงๆ</p> <p>4. เธอยังจำได้อีกหรือ</p> <p>5. เขาทำเป็นจำฉันไม่ได้นะ</p> <p>9. 1. แม่สอนให้เราทำดี</p> <p>2. พ่อสอนว่าเราต้องขยัน</p> <p>3. เขาถูกฉันสอนมวยเข้าให้</p> <p>4. ลุงสอนไม่ให้พูดหยาบคาย</p> <p>5. พี่สอนว่าอย่าหยิบของของผู้อื่นมา</p> <p>10. 1. ฉันคิดถึงเขามาก</p> <p>2. เขาคิดถึงฉันมากเลย</p> <p>3. เขารู้ใหม่ว่าฉันคิดถึงเขา</p> <p>4. เขาเป็นคนที่ฉันคิดถึงมาก</p> <p>5. คนเดียวที่ฉันคิดถึงก็คือเขา</p> |
|---|--|

รูปแบบที่ 3 ความเข้าใจภาษา

คำชี้แจง ให้ใช้บทกลอนต่อไปนี้ตอบคำถามข้อ 1-3

คุณค่าของหนังสือ	
หนังสือคือกุญแจแก้ความเหงา	หนังสือเฝาคความทุกข์ให้สุขได้
หนังสือเปรียบพลังธรรมอันอำไพ	หนังสือใช้เป็นครูผู้ฝึกคน
หนังสือคือหมอยารักษาใจ	หนังสือคือสายโซ่คล้องภาษา
หนังสือคือชุมทรัพย์สรรพปัญญา	หนังสือพามนุษย์ชาติฉลาดเอย

1. ผู้เขียนรู้สึกละอย่างไร

1. ตกใจ
2. ภูมิใจ
3. ชื่นชม
4. ชื่นชอบ
5. สรรเสริญ

2. กลอนบทนี้เหมาะสมกับใคร

1. ครู
2. ใครก็ได้
3. เด็กนักเรียน
4. เจ้าของหนังสือ
5. บุคคลทั่วไปทุกคน

3. จากข้อความ “หนังสือใช้เป็นครูผู้ฝึกคน” หมายความว่าอย่างไร

1. ครูก็คือหนังสือ
2. คนมีหนังสือเป็นครู
3. ผู้ฝึกฝนคนคือหนังสือ
4. หนังสือเปรียบเสมือนครู
5. หนังสือใช้ให้ครูไปสอนคน

คำชี้แจง ให้ใช้ข้อความต่อไปนี้ตอบคำถามข้อ 4

“เรียนรู้ตลอดชีวิต คิดอย่างสร้างสรรค์ ก้าวทันเทคโนโลยี”

4. บุคคลใดต่อไปนี้เป็นไปตามคำกล่าวข้างต้น

1. ตูฝึกเล่นเกมจากคอมพิวเตอร์เมื่อว่าง
2. ตั้มเรียนทุกวิชาแต่ก็เรียนไม่จบสักวิชา
3. ต้นเรียนหนังสือตั้งแต่เด็กจนแก่แล้วยังไม่จบ
4. ต้นเป็นครูจนเกษียณ ทุกวันนี้ก็ยังศึกษาเรื่องคอมพิวเตอร์
5. ต้อยเรียนจบปริญญาเอกสาขาคอมพิวเตอร์ตามที่แม่ต้องการ

คำชี้แจง ให้ใช้ข้อความต่อไปนี้ตอบคำถามข้อ 5

“เรียนรู้คู่รอยยิ้ม”

5. ผู้กล่าวต้องการสื่อถึงเรื่องใด

1. ให้เรียนไปยิ้มไป
2. ให้เรียนอย่างมีความสุข
3. เข้าห้องเรียนทุกครั้งต้องยิ้ม
4. ให้ยิ้มตลอดเวลาขณะที่เรียนอยู่
5. ต้องยิ้มตลอดเวลาถึงจะเรียนรู้ได้ดี

คำชี้แจง ให้ใช้บทกลอนต่อไปนี้ตอบคำถามข้อ 6-7

พ่อแม่ไม่มีเงินทองมากองให้ หาวิชาความรู้เป็นคู่มือ พอกับแม่มีแต่จะแก่เฒ่า ใช้วิชาช่วยตนไปจนตาย	จงตั้งใจพากเพียรเรียนหนังสือ ให้ยืดอกเอาไว้มั่นกาย จะเลี้ยงเราเรื่อยไปนั้นอย่าหมาย สุขสบายแม่พอก็พอใจ
---	--

6. ผู้กล่าวต้องการสื่อในลักษณะใด

1. สัจสอน
2. ชักชวน
3. ตักเตือน
4. โฆษณา
5. บอกกล่าว

7. อะไรเป็นสิ่งที่จะบอกได้ว่าพ่อแม่พอใจ

1. ลูกมีเงินทอง
2. ลูกเลี้ยงดูพ่อแม่ได้
3. ลูกเรียนหนังสือเก่ง
4. ลูกมีหนังสือมากมาย
5. ลูกมีความรู้ติดตัวช่วยเหลือตัวเองได้

คำชี้แจง

ให้ใช้บทกลอนต่อไปนี้ตอบคำถามข้อ 8 - 10

คนจะงาม	งามที่ใจ	ไซ้ไบหน้า
คนจะสวย	สวยจรรยา	ไซ้ตาหวาน
คนจะแก่	แก่ความรู้	ไซ้ยืนนาน
คนจะรวย	รวยศีลทาน	ไซ้บ้านใด

8. คนจะงามต้องงามแบบใด

1. จิตใจดี
2. รูปร่างดี
3. ไบหน้าดี
4. แต่งตัวดี
5. ผิวพรรณดี

9. คนจะสวยต้องสวยที่ใด

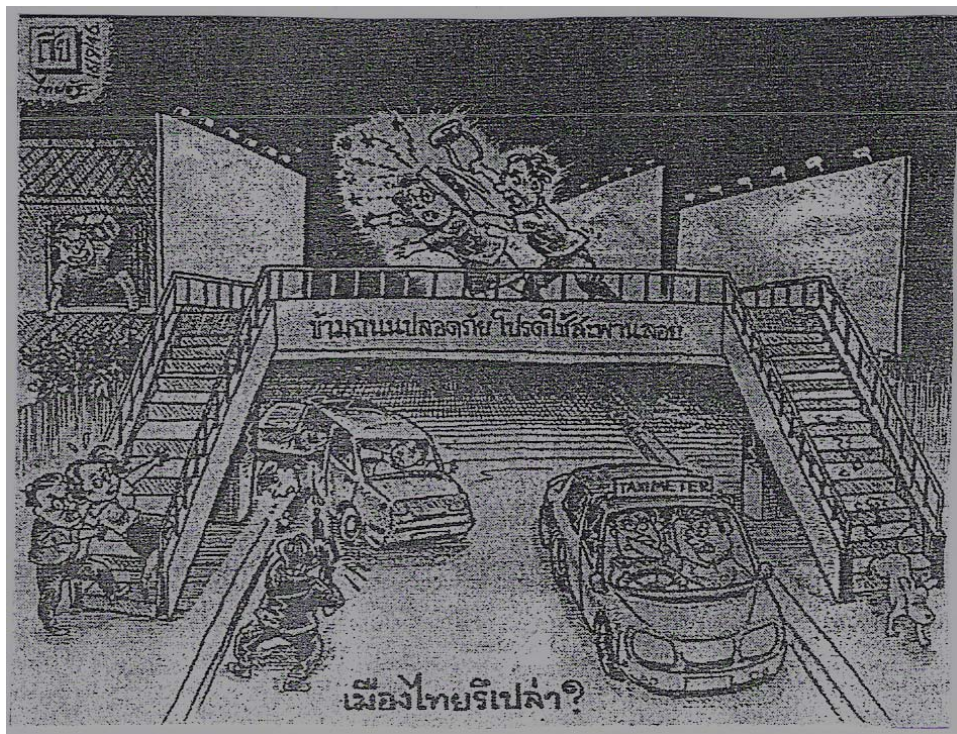
1. ตาหวาน
2. หน้าหวาน
3. มารยาทดี
4. หน้าสมส่วน
5. จรรยาบรรณ

10. คนจะแก่ต้องแก่แบบใดถึงจะเหมาะสม

1. แก่ตัว
2. แก่วัย
3. แก่อายุ
4. แก่ที่ใจ
5. แก่ความรู้

รูปแบบที่ 4 ความเข้าใจภาพ

คำชี้แจง ให้ใช้ภาพต่อไปนี้ตอบคำถามข้อ 1 – 2



1. ภาพนี้สื่อถึงเรื่องใด

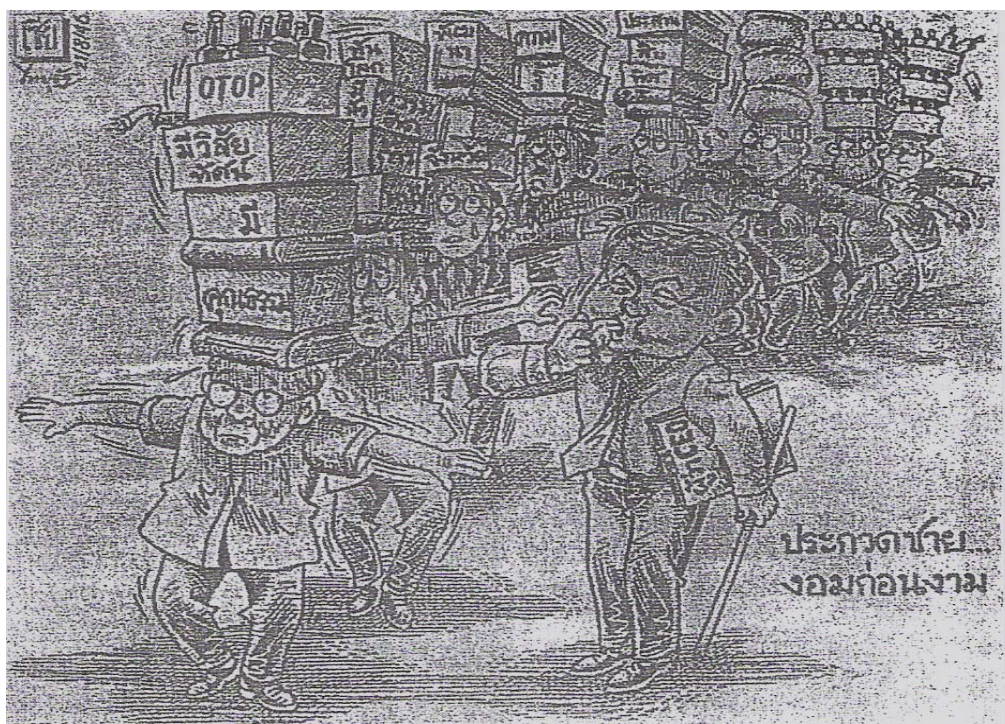
1. ความโลภ
2. ภัยอันตราย
3. ความขาดสติ
4. ความโหดร้าย
5. ความโกรธชัง

2. จากภาพเกิดขึ้นที่ใด

1. บนถนน
2. ทางเปลี่ยว
3. ใต้ทางด่วน
4. บนทางด่วน
5. บนและใต้สะพานลอย

คำชี้แจง

ให้ใช้ภาพต่อไปนี้ตอบคำถามข้อ 3 – 4



3. ภาพนี้สื่อเกี่ยวกับเรื่องใด

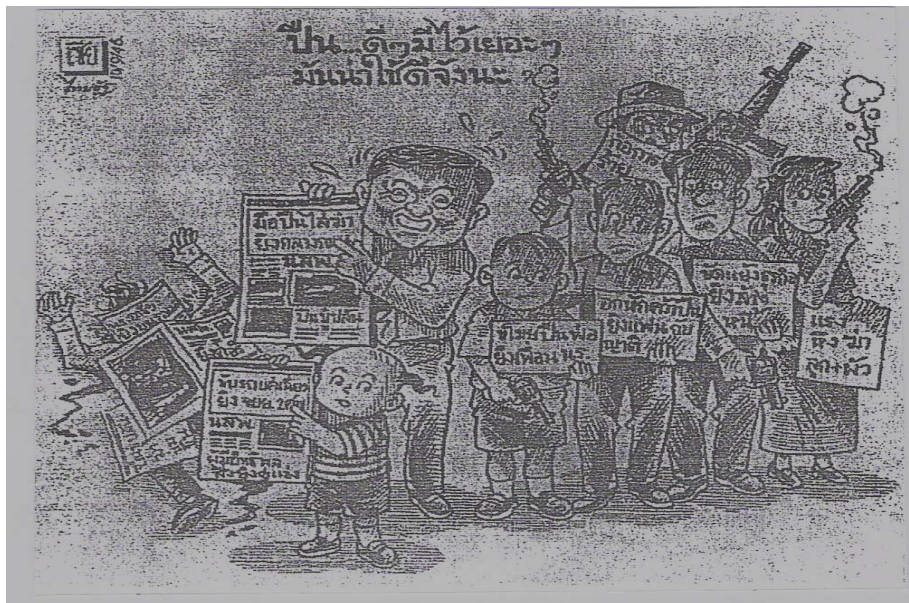
1. ผู้ชนะสิบทิศ
2. การฝึกโยคะ
3. การทำงานหนัก
4. การเดินมาราธอน
5. การประกวดชายงาม

4. ภาพนี้สอนใครได้ดี

1. ครู
2. นักเรียน
3. ผู้ปกครอง
4. ข้าราชการ
5. ประชาชนทั่วไป

คำชี้แจง

ให้ใช้ภาพต่อไปนี้ตอบคำถามข้อ 5 – 7



5. จากภาพเกิดขึ้นเพราะเหตุใด

1. ค่านิยมพื้นฐาน
2. การสั่งสอนและสั่งสม
3. การขาดสติในการยั้งคิด
4. การเลียนแบบในสิ่งที่เห็น
5. สืบเนื่องต่อกันมาทางสายเลือด

6. ข้อคิดของภาพนี้คือข้อใด

1. ปิ่นที่ดีจะนำโชค
2. ปิ่นดี ๆ ควรมีไว้หลายๆ
3. ควรจะมีปิ่นเก็บไว้ที่บ้านทุกคน
4. ควรรักษาปิ่นไว้ให้ดีเผื่อจะได้ใช้
5. ควรคิดแก้ปัญหาเรื่องปิ่นกันได้แล้ว

7. ใครควรเป็นผู้แก้ปัญหาเรื่องนี้

1. ครู
2. เด็ก
3. ผู้ปกครอง
4. นายกรัฐมนตรี
5. ประชาชนทุกคน

คำชี้แจง

ให้ใช้ภาพต่อไปนี้ตอบคำถามข้อ 8 - 10



8. ถ้าเชือกขาดจะเกิดอะไรขึ้น

1. จบเกม
2. ตำรวจชนะ
3. นักการเมืองชนะ
4. ผู้ดูเข้าไปซ้ำเติม
5. ก้าวไม่พร้อมกันอาจไปคนละทาง

9. จะเกิด “ครอบครัวเข้มแข็ง ชีวิตมั่นคง” ได้เมื่อใด

1. นักการเมืองเดินนำทหาร
2. ทหารกอดคอกับนักการเมือง
3. ทหารติดตามนักการเมืองตลอด
4. ทหารต้องอยู่ห่างจากนักการเมือง
5. ทหารกับนักการเมืองร่วมมือร่วมใจ

10. จะเกิดประโยชน์กับใครมากที่สุด ถ้าความจริงเป็นไปตามภาพ

1. ตำรวจ
2. นักธุรกิจ
3. นักการเมือง
4. เด็กนักเรียน
5. ประชาชนทุกคน

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ ชื่อสกุล	นางนิภาพร ขำสอาด
วันเดือนปีเกิด	12 กุมภาพันธ์ 2521
สถานที่เกิด	79 หมู่ 12 ตำบลช่วงเปา อำเภอจอมทอง จังหวัดเชียงใหม่
สถานที่อยู่ปัจจุบัน	26/58 – 59 แขวงคลองถนน เขตสายไหม กรุงเทพมหานคร
ตำแหน่งหน้าที่การงานปัจจุบัน	ครู คศ.1 โรงเรียนวัดเกาะสุวรรณาราม
สถานที่ทำงานปัจจุบัน	โรงเรียนวัดเกาะสุวรรณาราม สำนักงานเขตสายไหม สังกัดกรุงเทพมหานคร
ประวัติการศึกษา	
พ.ศ. 2538	มัธยมศึกษาตอนปลาย โรงเรียนจอมทอง จังหวัดเชียงใหม่
พ.ศ. 2543	ค.บ. (คณิตศาสตร์) มหาวิทยาลัยราชภัฏเชียงใหม่
พ.ศ. 2552	กศ.ม. (การวัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ

