

ผลการเปรียบเทียบพหุคูณรายคู่บางวิธีภายใต้ความแปรปรวนที่แตกต่างกัน
จากกลุ่มตัวอย่างขนาดต่างกัน

ปริญญาณิพนธ์
ของ
สิรินุช เอี่ยมเขียว

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวิจัยและสถิติทางการศึกษา

พฤษภาคม 2544

ลิขสิทธิ์เป็นของ มหาวิทยาลัยศรีนครินทรวิโรฒ

514.538

ค ๕ ๕1๗

๕ ๓

ผลการเปรียบเทียบพหุคูณรายคู่บางวิธีภายใต้ความแปรปรวนที่ต่างกัน
จากกลุ่มตัวอย่างขนาดต่างกัน

บทคัดย่อ

ของ

สิรินุช เอี่ยมเขียว

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวิจัยและสถิติทางการศึกษา

พฤษภาคม 2544

153222

สิรินุช เอี่ยมเขียว. (2544). ผลการเปรียบเทียบพหุคูณรายคู่บางวิธีภายใต้ความแปรปรวนที่แตกต่างกัน จากกลุ่มตัวอย่างขนาดต่างกัน. ปรินทิพนิพนธ์ กศ.ม. (การวิจัยและสถิติทางการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม: รองศาสตราจารย์ ดร. บุญเชิด ภิญโญอนันตพงษ์, อาจารย์ ขวาลิต รวยอาจิณ.

การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาผลการเปรียบเทียบพหุคูณรายคู่ ภายใต้ความแปรปรวนที่แตกต่างกัน ด้วยวิธีการทดสอบพหุคูณรายคู่ 3 วิธี คือ วิธี Dunnett's T3 วิธี Games-Howell และ วิธี Brown-Forsythe จากกลุ่มตัวอย่าง 4 ขนาด คือ ขนาด 20 คน 40 คน 80 คน และ 120 คน ว่าจะให้ผลการเปรียบเทียบพหุคูณแตกต่างกันหรือไม่ และการเปรียบเทียบพหุคูณวิธีใด ใช้กลุ่มตัวอย่างเท่าใดจึงจะเหมาะสมที่สุด

กลุ่มตัวอย่างที่ทำการศึกษานักเรียนชั้นมัธยมศึกษาปีที่ 2 , 4 และ 6 ของนักเรียนในสังกัดกรมสามัญศึกษา ในกรุงเทพมหานคร จำนวน 583 คน ที่ได้จากการสุ่มแบบ 2 ขั้นตอน แล้วนำมากำหนดเป็นประชากรเทียม เพื่อสุ่มเป็นกลุ่มตัวอย่างขนาด 20 คน 40 คน 80 คน และ 120 คน ขนาดละ 10 ครั้ง ด้วยการสุ่มแบบใส่คืน เครื่องมือที่ใช้ในการศึกษา คือ แบบทดสอบวัดความถนัดด้านเหตุผล แบบสรุปความ ตามแนวของเทอร์สโตน จำนวน 30 ข้อ มีค่าความเชื่อมั่น 0.634

ผลการศึกษาพบว่า

1) ผลการเปรียบเทียบพหุคูณ จากการทดสอบพหุคูณรายคู่ 3 วิธี ในกลุ่มตัวอย่างขนาด 120 คน มีจำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ.01 ส่วนขนาด 20 คน 40 คน และ 80 คน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2) ผลการเปรียบเทียบพหุคูณระหว่างกลุ่มตัวอย่าง 4 ขนาด ที่ทดสอบพหุคูณรายคู่ 3 วิธี มีจำนวนผลการยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

RESULTS OF PAIRWISE COMPARISON PROCEDURES UNDER
HETEROGENEOUS VARIANCES WITH DIFFERENT SAMPLE SIZES

AN ABSTRACT
BY
SIRINUCH EAIMKEOW

Presented in partial fulfillment of requirement for
the Master of Education degree in Educational Research and Statistics
At Srinakarinwirot University
May 2001

Sirinuch Eaimkeow. (2001). *Results of Pairwise Comparison Procedures Under Heterogeneous Variances With Different Sample Sizes*. Master Thesis. M.Ed. (Educational Research and Statistics). Bangkok : Graduate School, Srinakharinwirot University. Advisor Committee: Assoc.Prof.Dr. Boonchird Pinyoanunatapong, Mr. Chawalit Ruayajin.

The purpose of the research is to study the results of the pairwise comparison procedures among three different methods : Dunnett's T3 test , Games-Howell test and Brown-Forsythe test by using 4 sample sizes of 20, 40, 80 and 120 students to see if there are any difference among the tests and which method yields to the most suitable result and at which sample sizes.

The sample group consisted of 583 MathayomSuksa II , IV and VI students in academic year 2000 in Bangkok by using the two - stage random sampling as Pseudopopulation to be 20, 40, 80 and 120 students and each 8 group had 10 times by repeated random sampling.

The research tool was the aptitude verbal reasoning test , based on 30 items of Thurstone conception having the reliability 0.634.

The findings were as follows :

- 1) The results of pairwise comparison among three methods in the sample sizes of 120 students were significantly different at the level .01 and the samples of 20 , 40 and 80 students were not significantly different.
- 2) The results of pairwise comparison among the the four group sizes among three methods were significantly different at the level .01.

ปริญญานิพนธ์

เรื่อง

ผลการเปรียบเทียบพหุคูณรายคู่บางวิธีภายใต้ความแปรปรวนที่ต่างกัน
จากกลุ่มตัวอย่างขนาดต่างกัน

ของ

นางสาวสิรินุช เอี่ยมเขียว

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวิจัยและสถิติทางการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ



คณบดีบัณฑิตวิทยาลัย

(ศาสตราจารย์ ดร. เสริมศักดิ์ วิชาลาภรณ์)

วันที่ ๘ เดือน พฤษภาคม พ.ศ. 2544

คณะกรรมการสอบปริญญานิพนธ์



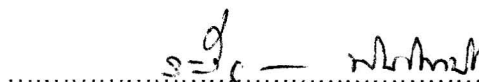
ประธาน

(รองศาสตราจารย์ ดร. บุญเชิด ภิญโญอนันตพงษ์)



กรรมการ

(อาจารย์ชวลิต รวยอาจิณ)



กรรมการที่แต่งตั้งเพิ่มเติม

(ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช)



กรรมการที่แต่งตั้งเพิ่มเติม

(ผู้ช่วยศาสตราจารย์ ดร. นิคม ตั้งกะพิภพ)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลุล่วงได้ด้วยความอนุเคราะห์และความกรุณาอย่างสูงจาก รองศาสตราจารย์ ดร. บุญเชิด ภิญโญอนันตพงษ์ และอาจารย์ ขวลิต รวยอาจิณ ที่ได้กรุณาให้ คำปรึกษา แนะนำ ชี้แนะแนวทาง ตรวจ แก้ไขข้อบกพร่อง จนงานวิจัยนี้สำเร็จลุล่วงไปได้ด้วยดี ผู้วิจัยขอกราบขอบพระคุณอย่างสูง

ขอกราบขอบพระคุณ กรรมการควบคุมการสอบปริญญานิพนธ์ ผู้ช่วยศาสตราจารย์ ระวีวรรณ พันธุ์พานิช และผู้ช่วยศาสตราจารย์ ดร.นิคม ดังคะพิภพ รวมทั้งอาจารย์ในภาควิชา การวัดผลและวิจัยทางการศึกษา ทุกท่านที่ได้แนะแนวทางในการทำปริญญานิพนธ์ฉบับนี้

ขอกราบขอบพระคุณผู้อำนวยการ คณะครู อาจารย์ทุกท่านของโรงเรียนที่เป็น กลุ่มตัวอย่าง รวมทั้งนักเรียนทุกคนที่ได้ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

ขอกราบขอบพระคุณคุณพ่อ นาวาอากาศโทสมคิด เอี่ยมเขียว และคุณแม่ ขนิษฐา เอี่ยมเขียว ที่ได้ให้ความอนุเคราะห์และกำลังใจอย่างมากในการทำปริญญานิพนธ์ฉบับนี้

ขอขอบคุณอาจารย์อุษา ธนาบุญฤทธิ์ ที่ให้ความช่วยเหลือในการเก็บข้อมูล และ ขอขอบคุณเพื่อนวิชาเอกวิจัยและสถิติทางการศึกษา ภาควิชา รุณที่1 ที่ให้คำแนะนำและเป็น กำลังใจ

คุณค่าและประโยชน์อันพึงมีจากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชา พระคุณบิดา มารดา และครูอาจารย์ ตลอดจนผู้มีพระคุณที่ได้ให้การอบรมสั่งสอนตลอดมา

สิรินุช เอี่ยมเขียว

สารบัญ

บทที่

หน้า

1 บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายของการศึกษา	4
ความสำคัญของการศึกษา	4
ขอบเขตของการศึกษา	5
ประชากรและกลุ่มตัวอย่าง	5
ตัวแปรที่ศึกษา	5
นิยามศัพท์เฉพาะ	5
สมมติฐานการศึกษา	7
2 เอกสารงานวิจัยที่เกี่ยวข้อง	8
การเปรียบเทียบพหุคูณ	8
ความหมายของการเปรียบเทียบพหุคูณ	9
การเปรียบเทียบระหว่างค่าเฉลี่ย	10
ชนิดของการเปรียบเทียบ	12
ประเภทของสถิติ	14
ประเภทของการเปรียบเทียบ	16
การทดสอบความเท่ากันของแปรปรวน	20
วิธีของ Levene	20
การแสดงสูตรคำนวณ	21
การทดสอบด้วย Dunnett's T3 test	21
การทดสอบด้วย Games-Howell test	22
การทดสอบด้วย Brown-Forsythe test	23

สารบัญ(ต่อ)

บทที่	หน้า
2 (ต่อ)ความถนัดทางด้านเหตุผล	24
ความหมายของความถนัดด้านเหตุผล	24
ทฤษฎีเชาว์ปัญญาและความถนัด	25
ทฤษฎีพัฒนาการทางสติปัญญาของเพียร์เจต์กับการคิดหาเหตุผล	27
การวัดความสามารถด้านเหตุผล	28
การใช้แบบทดสอบความถนัด	31
งานวิจัยที่เกี่ยวข้อง	32
งานวิจัยต่างประเทศ	32
งานวิจัยภายในประเทศ	33
3 วิธีดำเนินการศึกษา	40
ประชากรและกลุ่มตัวอย่างที่ใช้ในการศึกษา	40
เครื่องมือที่ใช้ในการศึกษา	41
วิธีการเก็บรวบรวมข้อมูล	43
วิธีการวิเคราะห์ข้อมูล	44
สถิติที่ใช้ในการศึกษา	45
4 ผลการวิเคราะห์ข้อมูล	50
การนำเสนอผลการวิเคราะห์ข้อมูล	50
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล	50
ผลการวิเคราะห์ข้อมูล	51
5 สรุป อภิปรายผล และข้อเสนอแนะ	66
ความมุ่งหมายของการศึกษา	66
สมมติฐานการศึกษา	66
กลุ่มตัวอย่างที่ใช้ในการศึกษา	66
เครื่องมือที่ใช้ในการศึกษา	66

สารบัญ(ต่อ)

บทที่	หน้า
5(ต่อ)วิธีการเก็บข้อมูล	67
การวิเคราะห์ข้อมูล	67
การเสนอผลการวิเคราะห์ข้อมูล	68
สรุปผลการวิจัย	68
อภิปรายผล	69
ข้อเสนอแนะ	71
 บรรณานุกรม	 72
 ภาคผนวก	 78
 ประวัติย่อผู้วิจัย	 97

บัญชีตาราง

ตาราง

หน้า

1	วิธีการเปรียบเทียบพหุคูณที่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติ	17
2	วิธีการเปรียบเทียบพหุคูณที่ไม่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติ	18
3	จำนวนโรงเรียนและห้องเรียนที่ทำการสุ่มได้ของนักเรียนชั้นมัธยมศึกษา ปีที่ 2, 4 และ 6	41
4	ค่าสถิติพื้นฐานของคะแนนจากแบบทดสอบของนักเรียนชั้นมัธยมศึกษา ปีที่ 2, 4 และ 6 ที่เป็นกลุ่มตัวอย่างขนาด 583 คน	51
5	ค่าสถิติพื้นฐานของคะแนนจากแบบทดสอบของนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ที่เป็นกลุ่มตัวอย่างขนาด 20 คน จำนวน 30 กลุ่ม	52
6	ค่าสถิติพื้นฐานของคะแนนจากแบบทดสอบของนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ที่เป็นกลุ่มตัวอย่างขนาด 40 คน จำนวน 30 กลุ่ม	53
7	ค่าสถิติพื้นฐานของคะแนนจากแบบทดสอบของนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ที่เป็นกลุ่มตัวอย่างขนาด 80 คน จำนวน 30 กลุ่ม	54
8	ค่าสถิติพื้นฐานของคะแนนจากแบบทดสอบของนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ที่เป็นกลุ่มตัวอย่างขนาด 120 คน จำนวน 30 กลุ่ม	55
9	เปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบ พหุคูณต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาด 20 คน	56
10	เปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบ พหุคูณต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาด 40 คน	57
11	เปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบ พหุคูณต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาด 80 คน	57
12	เปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบ พหุคูณต่างกัน 3 วิธี จากกลุ่มตัวอย่างขนาด 120 คน	58
13	เปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบ พหุคูณต่างกัน 3 วิธี ครั้งละ 2 วิธีจากกลุ่มตัวอย่าง 120 คน	59
14	เปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบด้วย วิธี Dunnett's T3 จากกลุ่มตัวอย่าง 4 ขนาด	60

บัญชีตาราง (ต่อ)

ตาราง

หน้า

15	เปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบด้วย วิธี Dunnett's T3 จากกลุ่มตัวอย่างครั้งละ 2 ขนาด	61
16	เปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบด้วย วิธี Games-Howell จากกลุ่มตัวอย่าง 4 ขนาด.....	61
17	เปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบด้วย วิธี Games- Howell จากกลุ่มตัวอย่างครั้งละ 2 ขนาด	62
18	เปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบด้วย วิธี Brown-Forsythe จากกลุ่มตัวอย่าง 4 ขนาด	63
19	เปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐานจากการทดสอบด้วย วิธี Brown-Forsythe จากกลุ่มตัวอย่างครั้งละ 2 ขนาด	64
20	จำนวนผลการเปรียบเทียบพหุคูณที่ไม่ยอมรับสมมติฐาน ที่ทำการ ทดสอบพหุคูณรายคู่ 3 วิธี จากกลุ่มตัวอย่าง 4 ขนาด	64
21	ค่าความแตกต่างระหว่างค่าเฉลี่ยของ 2 ระดับชั้นที่ทดสอบ ด้วยวิธี Dunnett's T3 จากกลุ่มตัวอย่าง 4 ขนาด	80
22	ค่าความแตกต่างระหว่างค่าเฉลี่ยของ 2 ระดับชั้นที่ทดสอบ ด้วยวิธี Games- Howell จากกลุ่มตัวอย่าง 4 ขนาด	81
23	ค่าความแตกต่างระหว่างค่าเฉลี่ยของ 2 ระดับชั้นที่ทดสอบ ด้วยวิธี Brown-Forsythe จากกลุ่มตัวอย่าง 4 ขนาด	82
24	ค่าความแตกต่างระหว่างค่าความแปรปรวนของ 3 ระดับชั้นที่ทดสอบ ด้วยวิธี Levene จากกลุ่มตัวอย่างขนาด 20 คน	83
25	ค่าความแตกต่างระหว่างค่าความแปรปรวนของ 3 ระดับชั้นที่ทดสอบ ด้วยวิธี Levene จากกลุ่มตัวอย่างขนาด 40 คน	84
26	ค่าความแตกต่างระหว่างค่าความแปรปรวนของ 3 ระดับชั้นที่ทดสอบ ด้วยวิธี Levene จากกลุ่มตัวอย่างขนาด 80 คน	85
27	ค่าความแตกต่างระหว่างค่าความแปรปรวนของ 3 ระดับชั้นที่ทดสอบ ด้วยวิธี Levene จากกลุ่มตัวอย่างขนาด 120 คน	86

บัญชีตาราง (ต่อ)

ตาราง

หน้า

28	จำนวนผลการเปรียบเทียบพหุคูณ ที่ทดสอบด้วยพหุคูณรายคู่ ด้วยวิธี Dunnett's T3 จากกลุ่มตัวอย่าง 4 ขนาด	87
29	จำนวนผลการเปรียบเทียบพหุคูณ ที่ทดสอบด้วยพหุคูณรายคู่ ด้วยวิธี Games-Howell จากกลุ่มตัวอย่าง 4 ขนาด	88
30	จำนวนผลการเปรียบเทียบพหุคูณ ที่ทดสอบด้วยพหุคูณรายคู่ ด้วยวิธี Brown-Forsythe จากกลุ่มตัวอย่าง 4 ขนาด	89

บทที่ 1

บทนำ

ภูมิหลัง

ในการทดสอบสมมติฐานสำหรับงานวิจัยด้านต่าง ๆ นั้น นอกจากผู้วิจัยจะต้องมีความรู้ความเข้าใจในเรื่องที่จะศึกษาเป็นอย่างดี เพื่อกำหนดแบบแผนการทดลอง หรือเทคนิคการสุ่มตัวอย่างที่เหมาะสมแล้ว ผู้วิจัยจะต้องเลือกสถิติทดสอบที่เหมาะสมกับลักษณะข้อมูล ซึ่งแต่ละสถิติต่างก็มีข้อตกลงเบื้องต้น (Assumptions) เกี่ยวกับลักษณะของข้อมูลที่จะนำมาวิเคราะห์แตกต่างกัน ดังนั้น การเลือกใช้สถิติทดสอบที่มีข้อตกลงเบื้องต้นสอดคล้องกับลักษณะข้อมูลจะส่งผลทำให้การสรุปผลของการวิจัยมีความถูกต้องและเชื่อถือได้มากขึ้น การทดสอบสมมติฐานทางสถิติ เป็นวิธีการหนึ่งที่จะตอบปัญหาเกี่ยวข้องกับค่าพารามิเตอร์ของประชากร ไม่ว่าจะเป็นพารามิเตอร์ที่สนใจจะเป็นค่าความแปรปรวน ค่าเฉลี่ย หรือค่าสัดส่วนของสิ่งที่สนใจในประชากรใดก็ตามต่างก็มีวิธีการทดสอบเพื่อตอบข้อสงสัยของผู้วิจัย

งานวิจัยส่วนใหญ่ทั้งการวิจัยเชิงทดลองและการวิจัยเชิงสำรวจ มักจะทำการทดสอบความมีนัยสำคัญในค่าเฉลี่ยของประชากร หรือหาข้อสรุปว่า การจัดการกระทำ (treatment) ที่ให้ในแต่ละกลุ่มตัวอย่าง มีการจัดการกระทำใดบ้างที่แตกต่างกัน ในการทดสอบสมมติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ยของประชากรที่มีมากกว่า 2 กลุ่มขึ้นไป ($k \geq 3$) การทดสอบดังกล่าวที่ใช้กันโดยทั่วไป จะใช้การวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA) ถ้าในการทดสอบปฏิเสธสมมติฐานหลัก (H_0) ที่ว่า ค่าเฉลี่ยของประชากรทุกกลุ่มเท่ากัน หรือ $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ แสดงว่า ค่าเฉลี่ยอย่างน้อยหนึ่งกลุ่มจะไม่เท่ากับกลุ่มอื่น ๆ ซึ่งผู้วิจัยยังสรุปรายละเอียดไม่ได้ว่าผลจากการจัดการกระทำใดหรือค่าเฉลี่ยของประชากรกลุ่มใดบ้างที่แตกต่างกัน ผู้วิจัยต้องทำการทดสอบเพื่อตัดสินใจต่อไปว่า ค่าเฉลี่ยใดไม่เท่ากันบ้างโดยใช้วิธีการเปรียบเทียบพหุคูณ การเปรียบเทียบพหุคูณ เป็นการเปรียบเทียบค่าเฉลี่ย k จำนวน ในรูปของผลรวมเชิงเส้นของค่าเฉลี่ย ซึ่งเป็นความแตกต่างระหว่างค่าเฉลี่ย 2 ค่าของค่าเฉลี่ย k จำนวน หรือเป็นความแตกต่างระหว่างค่าเฉลี่ยหนึ่งค่า กับจำนวนเฉลี่ยของค่าเฉลี่ยอื่น (Toothaker, 1991:2) ได้มีนักสถิติหลายท่านพัฒนาคิดค้นหาวิธีการเปรียบเทียบพหุคูณไว้หลากหลายวิธี ซึ่ง เคิร์ก ได้รวบรวมไว้ในหนังสือที่ชื่อว่า Experimental design : procedures for the behavioral science ได้ 22 วิธี (Kirk, 1995) และแบ่งได้เป็น 2 กลุ่มใหญ่ กลุ่มแรก คือ กลุ่มวิธีการเปรียบเทียบพหุคูณที่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติที่ว่า ความแปรปรวนต้องไม่แตกต่างกัน (Homogeneous Variances) ($\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$) วิธีการเปรียบเทียบพหุคูณในกลุ่มนี้มีหลายวิธี แบ่งได้เป็น การเปรียบเทียบแบบวางแผนล่วงหน้า

(Planned contrast) เช่น Dunnett's test, Holm's test และการเปรียบเทียบภายหลัง

(Post hoc contrast) เช่น Tukey's test, Fisher-Hayter test, Scheffe's test กลุ่มที่สองคือกลุ่มวิธีการเปรียบเทียบพหุคูณที่ไม่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติ คือความแปรปรวนแตกต่างกัน (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)

ผู้วิจัยมักจะเลือกวิธีการเปรียบเทียบภายหลัง ซึ่งให้อำนาจการทดสอบสูงเนื่องจากการเปรียบเทียบพหุคูณแบบวางแผนล่วงหน้า มักมีข้อจำกัดในจำนวนการเปรียบเทียบ คือถ้าจำนวนการเปรียบเทียบมีมาก และการเปรียบเทียบไม่เป็นอิสระต่อกัน จะทำให้ช่วงความเชื่อมั่นกว้าง และอำนาจการทดสอบต่ำ (Hay.1963:488) การใช้เทคนิคการเปรียบเทียบภายหลังต้องขึ้นกับการทดสอบเอฟ คือใช้การทดสอบเอฟเพื่อทดสอบความมีนัยสำคัญในค่าเฉลี่ยของประชากร เมื่อการทดสอบเอฟมีนัยสำคัญจึงใช้การเปรียบเทียบภายหลัง ถ้าการละเมิดข้อตกลงเบื้องต้นมีผลกระทบถึงการวิเคราะห์ความแปรปรวนด้วยสถิติทดสอบเอฟ โดยเฉพาะกรณีกลุ่มตัวอย่างมีขนาดไม่เท่ากัน และกลุ่มที่มีขนาดเล็กมีความแปรปรวนสูง กับกลุ่มที่มีขนาดใหญ่มีความแปรปรวนต่ำ ค่าระดับนัยสำคัญเอฟจะผิดจากความเป็นจริงมาก ทำให้อัตราความคลาดเคลื่อนประเภทที่ 1 ของการเปรียบเทียบพหุคูณภายหลังสูงขึ้นไปด้วย และถ้าข้อมูลเหมาะสมกับการทดสอบเอฟก็จะเหมาะสมกับการเปรียบเทียบพหุคูณด้วย (สมคิด ไวยวุฒินันท์. 2530:4) ในกรณีกลุ่มหลังนี้ที่ว่าไม่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติ คือความแปรปรวนแตกต่างกัน ผู้วิจัยควรทำการทดสอบความเท่ากันของความแปรปรวนก่อนการวิเคราะห์ความแปรปรวน เพื่อมิให้เกิดผลสรุปที่ผิดพลาดขึ้น ซึ่งเป็นกลุ่มที่น่าสนใจทำการศึกษา ในการปฏิบัติจริงส่วนใหญ่มักจะเผชิญปัญหาในการเลือกใช้สถิติเพื่อทำการวิเคราะห์ กล่าวคือ ประสบบัญญาว่าข้อมูลที่ผู้วิจัยมีอยู่ไม่สอดคล้องกับข้อตกลงเบื้องต้นของสถิตินั้นๆ จึงทำให้เกิดทางเลือกกับผู้วิจัยว่า ยังคงใช้สถิติตัวนั้นทดสอบโดยละเมิดข้อตกลงเบื้องต้น หรือจะเลือกใช้วิธีการทดสอบแบบอื่นทดแทน เช่น การทดสอบนอนพาราเมตริก ซึ่งสามารถนำไปใช้ทดสอบโดยไม่ต้องมีข้อตกลงเบื้องต้น แต่เนื่องจากการทดสอบดังกล่าวเป็นการทดสอบที่มีอำนาจการทดสอบต่ำ และการใช้สถิติทดสอบแบบดั้งเดิมมีพื้นฐานความเชื่อว่า การละเมิดข้อตกลงเบื้องต้นจะมีผลกระทบเพียงเล็กน้อยต่อการแจกแจงแบบสุ่มของค่าสถิติในการทดสอบ (ไพบรมา พจนพิมล. 2526:1) จึงมีผลทำให้นักวิจัยพยายามอ้างเหตุผลในการเลือกใช้สถิติให้สอดคล้องกับข้อตกลงเบื้องต้นของสถิติตัวนั้น หรือใช้สถิติตัวนั้นในการวิเคราะห์ข้อมูลโดยไม่คำนึงหรือเพิกเฉยในความผิดพลาดของข้อตกลงเบื้องต้นนั้นเสีย ข้อตกลงเบื้องต้นที่จะมี

อิทธิพลต่อการทดสอบมากคือ ข้อตกลงเบื้องต้นในเรื่องลักษณะการแจกแจงของประชากรเป็นปกติ และความเท่ากันของความแปรปรวนประชากร เมื่อผู้วิจัยไม่สามารถที่จะควบคุมให้ข้อมูลมีความแปรปรวนไม่แตกต่างกันได้ ข้อมูลที่ได้จากการวิจัยนั้นอาจไม่ตรงตามข้อตกลงเบื้องต้นที่กำหนดไว้ ดังนั้นค่าต่าง ๆ ที่นำมาวิเคราะห์นั้นจึงเป็นเพียงค่าประมาณมากกว่าค่าที่แท้จริง และการที่ข้อตกลงเบื้องต้นไม่เป็นจริงมีผลต่อระดับนัยสำคัญของการทดลอง (Cochran & Cox, 1971:91)

การออกแบบการวิจัย มีจุดมุ่งหมายเพื่อตอบคำถามการวิจัย และควบคุมความแปรปรวน (Kerlinger, 1986) การออกแบบการวิจัยครอบคลุมการออกแบบที่สำคัญได้แก่ การออกแบบการวัด ซึ่งเป็นการกำหนดรูปแบบและวิธีการวัดตัวแปรที่ศึกษา ออกแบบการสุ่มตัวอย่าง ซึ่งเป็นการกำหนดรูปแบบและวิธีการให้มาซึ่งประชากร และกลุ่มตัวอย่างที่ศึกษา และการออกแบบสถิติ ซึ่งเป็นการเลือกใช้สถิติที่เหมาะสมกับเป้าหมายของการวิจัย ลักษณะของประชากรและกลุ่มตัวอย่าง และลักษณะของตัวแปรที่ศึกษา (ศิริชัย กาญจนวาสี, 2537) การออกแบบการวิจัยทั้ง 3 ส่วนนั้น การออกแบบการสุ่มตัวอย่างเป็นการออกแบบที่มีความสำคัญต่อการวิจัย ในกรณีที่นักวิจัยไม่สามารถศึกษาข้อเท็จจริงจากประชากรทั้งหมดได้ จึงต้องใช้วิธีสุ่มตัวอย่างจากกลุ่มประชากรที่ศึกษา และใช้หลักฐานข้อเท็จจริงจากกลุ่มตัวอย่างนั้นสรุปผลเชิงสถิติจากกลุ่มประชากรเพื่อตอบปัญหาการวิจัย คุณลักษณะที่เป็นเป้าหมายกลุ่มสำคัญของการวิจัยที่ศึกษาจากกลุ่มตัวอย่าง คือ มุ่งให้การวิจัยมีความตรงโดยเฉพาะอย่างยิ่งความตรงภายนอก (external validity) หรือความสามารถในการสรุปผลอ้างอิงผลการวิจัยไปสู่ประชากรทั้งหมดที่ศึกษา (generalizability) การกำหนดตัวอย่างนั้นเป็นขั้นตอนหนึ่งในการออกแบบการสุ่มตัวอย่างที่นักวิจัยต้องดำเนินการก่อนเริ่มทำการวิจัย กิลฟอร์ด และ ฟร็ชเตอร์ (Guilford and Fruchter, 1978) กล่าวว่า คำถามที่เกิดขึ้นเสมอเมื่อเริ่มวางแผนการวิจัย คือ กลุ่มตัวอย่างที่ใช้ในการวิจัยควรมีขนาดเท่าใด ฮิงเคล และ โอลิเวอร์ (Hinkle and Oliver, 1983) กล่าวว่า การตอบคำถามเกี่ยวกับขนาดของกลุ่มตัวอย่างมีความยุ่งยากเนื่องจากความสัมพันธ์ระหว่างขนาดของกลุ่มตัวอย่างกับความมีนัยสำคัญทางสถิติของสถิติอนุมานที่นักวิจัยใช้ จุดวิกฤตของสถิติอนุมาน คือ งานวิจัยที่ใช้สถิติอนุมาน ให้ความสำคัญสูงกับความมีนัยสำคัญกับผลการวิเคราะห์ เฮ (Hay, 1973) กล่าวว่า งานวิจัยบางเรื่องสามารถแสดงผลที่มีนัยสำคัญได้ ถ้าใช้กลุ่มตัวอย่างขนาดใหญ่โดยที่ไม่มีการพิจารณาถึงความสมเหตุสมผลของสิ่งที่ศึกษา จุดสำคัญของวิธีการทางสถิติอนุมาน คือการแปลความหมายได้อย่างมั่นใจว่าความแตกต่างระหว่างพารามิเตอร์กับค่าทางสถิตินั้นเป็นความแตกต่างที่แท้จริง ไม่ใช่เป็นความแตกต่างที่เกิดขึ้นโดยบังเอิญอันเนื่องจากการสุ่ม การกำหนดขนาดของกลุ่มตัวอย่างมักถูกจำกัดไม่ให้มีขนาดใหญ่มากด้วยเหตุผลเกี่ยวกับค่าใช้จ่ายและเวลาในการเก็บข้อมูลและ

ความเป็นไปได้ในการจัดเก็บข้อมูลให้ได้คุณภาพ นักวิจัยที่มีเงินทุนและเวลาในการวิจัยที่จำกัด จึงมีความจำเป็นต้องใช้กลุ่มตัวอย่างที่มีขนาดเล็ก ส่วนในการวิจัยเชิงทดลอง นักวิจัยอาจต้องการควบคุมความคลาดเคลื่อนในการวัดที่จะเกิดในระหว่างการทดลอง จึงกำหนดตัวอย่างให้มีขนาดเล็ก แต่มีความเหมาะสมและเพียงพอ เพื่อให้ผลการวิจัยมีความเชื่อถือได้

ด้วยเหตุดังกล่าว ผู้วิจัยจึงสนใจที่จะศึกษาผลการเปรียบเทียบพหุคูณทั้ง 3 วิธีคือ Dunnett's T3 test ซึ่งอยู่ในกลุ่ม t - statistic Games-Howell test ซึ่งอยู่ในกลุ่ม q - statistic และ Brown-Forsythe test ซึ่งอยู่ในกลุ่ม F - statistic เนื่องจากมีข้อตกลงในการใช้เหมือนกันคือ กลุ่มตัวอย่างมีจำนวนเท่ากันหรือไม่ก็ได้ เป็นการเปรียบเทียบรายคู่ และ ความแปรปรวนแตกต่างกันว่าจะให้ผลการทดสอบต่างกันหรือไม่ เมื่อกลุ่มตัวอย่างมีขนาดต่างกัน โดยการศึกษาครั้งนี้ผู้วิจัยสนใจที่จะศึกษาเกี่ยวกับความถนัดด้านเหตุผล เพราะการใช้เหตุผลเป็นสิ่งที่จำเป็นในชีวิตประจำวัน เป็นความสามารถที่สำคัญที่สุดในการเรียนทุกชนิดทุกระดับชั้น

ความมุ่งหมายของการศึกษา

การศึกษานี้มีความมุ่งหมายสำคัญเพื่อศึกษาผลการเปรียบเทียบพหุคูณรายคู่ ด้วยวิธี Dunnett's T3 , วิธี Games-Howell และ วิธี Brown-Forsythe จากกลุ่มตัวอย่างขนาดต่างกัน ว่าให้ผลสอดคล้องกันมากน้อยเพียงใด โดยกำหนดเป็นความมุ่งหมายเฉพาะดังนี้

1. เพื่อเปรียบเทียบผลการทดสอบพหุคูณรายคู่ด้วย วิธี Dunnett's T3 วิธี Games-Howell และ วิธี Brown-Forsythe ในแต่ละกลุ่มตัวอย่างที่มีขนาดเดียวกัน 4 ขนาด คือ 20, 40, 80 และ 120
2. เพื่อเปรียบเทียบผลการทดสอบพหุคูณรายคู่ จากกลุ่มตัวอย่างขนาดต่างกันด้วยวิธีทดสอบเดียวกัน คือ วิธี Dunnett's T3 , วิธี Games-Howell และ วิธี Brown-Forsythe

ความสำคัญของการศึกษา

การศึกษานี้เป็นการศึกษาการเปรียบเทียบรายคู่ ภายใต้ความแปรปรวนที่แตกต่างกัน และขนาดกลุ่มตัวอย่างขนาดต่างกัน ผลการศึกษานี้เป็นวิธีการเปรียบเทียบรายคู่ วิธีใดจะคำนวณกับกลุ่มตัวอย่างขนาดเท่าใดจึงจะเหมาะสม เพื่อประโยชน์ในเชิงปฏิบัติ เพราะเป็นแนวทางให้ผู้มีเกณฑ์หรือหลักฐานสนับสนุน ทำให้ได้คำตอบที่ต้องใกล้เคียงความเป็นจริงมากที่สุดต่อการทำการศึกษาและการวิจัย

ขอบเขตของการศึกษา

1. ประชากรที่ใช้ในการศึกษาค้างนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ กรุงเทพมหานคร จำนวน 114 โรงเรียน จำนวนนักเรียน 122,828 คน

2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้างนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ กรุงเทพมหานคร จำนวน 583 คน สุ่มจากประชากรโดยวิธี Two-Stage Random Sampling

3. ตัวแปรที่ศึกษา

3.1 ตัวแปรอิสระ

3.1.1 วิธีการเปรียบเทียบพหุคูณแบบรายคู่ 3 วิธี คือ

- Dunnett's T3 test
- Games-Howell test
- Brown-Forsythe test

3.1.2 ขนาดของกลุ่มตัวอย่าง 4 ขนาด คือ

- 20 คน
- 40 คน
- 80 คน
- 120 คน

3.2 ตัวแปรตาม ได้แก่ ผลการเปรียบเทียบพหุคูณรายคู่

นิยามศัพท์เฉพาะ

1. การเปรียบเทียบพหุคูณ (Multiple Comparison) หมายถึงการเปรียบเทียบความแตกต่างของค่าเฉลี่ยทีละคู่มากกว่าหนึ่งคู่ ว่ามีความแตกต่างกันอย่างมีนัยสำคัญหรือไม่

2. วิธีการเปรียบเทียบพหุคูณรายคู่ภายใต้ความแปรปรวนที่ต่างกัน หมายถึงกระบวนการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยรายคู่ ที่มีเงื่อนไขว่าความแปรปรวนของประชากรของกลุ่มตัวอย่างต่างกัน ซึ่งในการศึกษาได้ศึกษาวิธีดังนี้

2.1 วิธี Dunnett's T3 คือวิธีทดสอบการเปรียบเทียบพหุคูณโดยมี

ข้อดกลงว่าความแปรปรวนของแต่ละกลุ่มต่างกัน มีรูปแบบสูตรดังนี้

$$T3 = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}}}$$

2.2. วิธี Games-Howell คือวิธีทดสอบการเปรียบเทียบพหุคูณโดยมี

ข้อดกลงว่าความแปรปรวนของแต่ละกลุ่มต่างกัน มีรูปแบบสูตรดังนี้

$$GH = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}\right) / 2}}$$

2.3. วิธี Brown-Forsythe test คือวิธีทดสอบการเปรียบเทียบพหุคูณโดยมีข้อ

ดกลงว่าความแปรปรวนของแต่ละกลุ่มต่างกัน มีรูปแบบสูตรดังนี้

$$BF = \frac{(\overline{X}_i - \overline{X}_j)^2}{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}}$$

3. ผลการเปรียบเทียบพหุคูณ หมายถึง จำนวนความแตกต่างระหว่างค่าเฉลี่ยรายคู่ ในการยอมรับหรือไม่ยอมรับสมมติฐานว่าเหมือนหรือต่างกันหรือไม่ ระหว่างวิธี Dunnett's T3 วิธี Games-Howell และ วิธี Brown-Forsythe.

4. ขนาดของกลุ่มตัวอย่าง หมายถึง จำนวนนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ปีการศึกษา 2543 ภาคเรียนที่ 2 ที่ได้ทำแบบทดสอบวัดความถนัดด้านเหตุผล แบบสรุปความ ซึ่งได้มาจากการสุ่มแบบใส่คืน จากกลุ่มตัวอย่าง 583 คน ที่กำหนดให้เป็นประชากรเทียม (Pseudopopulation) ในการศึกษาครั้งนี้ใช้ขนาดกลุ่มตัวอย่าง 4 ขนาด คือ ขนาด 20 คน 40 คน 80 คน และ 120 คน

5. ความถนัดทางการเรียน หมายถึง ชีตความสามารถของบุคคลอันเป็นผลมาจากการที่ได้ศึกษาและประสบการณ์ทั้งปวงที่ได้รับมา โดยอาศัยแบบทดสอบความถนัดทางการเรียนเป็นเครื่องมือในการวัด

6. ความถนัดด้านเหตุผล หมายถึง ความสามารถในการพิจารณาเหตุที่กำหนดให้ แล้วประเมินและลงสรุปเรื่องราวต่างๆ ได้อย่างสมเหตุสมผล ซึ่งในการศึกษาครั้งนี้วัดโดยใช้แบบทดสอบสรุปความ

สมมติฐานการศึกษา

1. การทดสอบพหุคูณรายคู่ด้วยวิธีต่างกันจากกลุ่มตัวอย่างเดียวกันมีสมมติฐานดังนี้

1.1 ผลการเปรียบเทียบพหุคูณด้วย วิธี Dunnett's T3 , วิธี Games-Howell และวิธี Brown-Forsythe ในกลุ่มตัวอย่างที่มีขนาด 20 คน ให้จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกัน

1.2 ผลการเปรียบเทียบพหุคูณด้วย วิธี Dunnett's T3 , วิธี Games-Howell และวิธี Brown-Forsythe ในกลุ่มตัวอย่างที่มีขนาด 40 คน ให้จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกัน

1.3 ผลการเปรียบเทียบพหุคูณด้วย วิธี Dunnett's T3 , วิธี Games-Howell และวิธี Brown-Forsythe ในกลุ่มตัวอย่างที่มีขนาด 80 คน ให้จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกัน

1.4 ผลการเปรียบเทียบพหุคูณด้วย วิธี Dunnett's T3, วิธี Games – Howell และ วิธี Brown – Forsythe . ในกลุ่มตัวอย่างที่มีขนาด 120 คน ให้จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกัน

2. การทดสอบพหุคูณรายคู่จากกลุ่มตัวอย่างขนาดต่างกันด้วยวิธีทดสอบเดียวกันมีสมมติฐานดังนี้

2.1 ผลการเปรียบเทียบพหุคูณวิธี Dunnett's T3 จากกลุ่มตัวอย่างขนาด 20, 40, 80 และ 120 คน ให้จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกัน

2.2 ผลการเปรียบเทียบพหุคูณวิธี Games-Howell จากกลุ่มตัวอย่างขนาด 20, 40, 80 และ 120 คน ให้จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกัน

2.3 ผลการเปรียบเทียบพหุคูณวิธี Brown-Forsythe จากกลุ่มตัวอย่างขนาด 20, 40, 80 และ 120 คน ให้จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกัน

บทที่ 2

เอกสารงานวิจัยที่เกี่ยวข้อง

ในการศึกษาครั้งนี้ผู้วิจัยได้ศึกษาเนื้อหาต่างๆ ดังนี้

การเปรียบเทียบพหุคูณ

1. ความหมายของการเปรียบเทียบพหุคูณ
2. การเปรียบเทียบระหว่างค่าเฉลี่ย
3. ชนิดของการเปรียบเทียบ
 - การเปรียบเทียบรายคู่
 - การเปรียบเทียบที่ไม่ใช่รายคู่
 - การเปรียบเทียบอิสระ
4. ประเภทของสถิติ
 - t - Statistic
 - q - Statistic
 - F - Statistic
5. การทดสอบความเท่ากันของความแปรปรวน
6. วิธีการของ Levene
7. การแสดงสูตรคำนวณ
 - การทดสอบ Dunnett's T3 test
 - การทดสอบ Games-Howell test
 - การทดสอบ Brown-Forsythe test

ความถนัดด้านเหตุผล

1. ความหมายของความถนัด
2. ทฤษฎีของเชอว์ปัญญา และความถนัด
3. ทฤษฎีพัฒนาทางสติปัญญาของเพียร์เจต์กับการคิดหาเหตุผล
4. การวัดความสามารถด้านเหตุผล
5. การใช้แบบทดสอบความถนัด

งานวิจัยที่เกี่ยวข้อง

1. งานวิจัยต่างประเทศ
2. งานวิจัยภายในประเทศ

ความหมายของการเปรียบเทียบพหุคูณ

มีนักสถิติและนักการศึกษาหลายท่านได้ให้ความหมายของการเปรียบเทียบพหุคูณไว้ต่าง ๆ กันดังนี้

เกลส และ สแตนเลย์ (Glass & Staley.1970:381) กล่าวว่า การเปรียบเทียบพหุคูณเป็นวิธีการเปรียบเทียบระหว่างค่าเฉลี่ยเพื่อปฏิเสธสมมติฐานหลัก H_0

เจมส์ (James. 1982) กล่าวว่า การคำนวณค่า F มีนัยสำคัญสามารถบอกได้ว่าการกระทำที่นำมาพิจารณานั้นแตกต่างกัน ในการศึกษาพฤติกรรมนั้นการวัดตัวแปรอิสระ ซึ่งเป็นการเปรียบเทียบแต่ละการกระทำที่นำมาพิจารณาเป็นการคำนวณทางสถิติ การทดสอบนี้เกี่ยวกับการตัดสินใจความแตกต่างระหว่างค่าเฉลี่ย หรือผลรวมของความน่าจะเป็นการรวมของ k กลุ่ม วิธีการทางสถิติที่ใช้เรียกว่า การทดสอบการเปรียบเทียบพหุคูณ การเปรียบเทียบพหุคูณส่วนใหญ่จะหมายถึงการเปรียบเทียบภายหลัง การเปรียบเทียบภายหลังเป็นการกระทำเมื่อพฤติกรรมที่เราศึกษาไม่มีการตัดสินใจและการเปรียบเทียบไว้ล่วงหน้า และต้องการที่จะค้นหาผลรวมตัวแปรของการกระทำที่นำมาศึกษาให้พบนัยสำคัญที่แตกต่าง

ทูแทคเกอร์ (Toothaker. 1991:2) กล่าวว่า การเปรียบเทียบพหุคูณ เป็นการเปรียบเทียบค่าเฉลี่ย k จำนวน ในรูปการผสมเชิงเส้นของค่าเฉลี่ย ซึ่งมีความแตกต่างระหว่างค่าเฉลี่ย 2 ค่า ของค่าเฉลี่ย k จำนวน หรือเป็นความแตกต่างระหว่างค่าเฉลี่ยหนึ่งค่ากับจำนวนเฉลี่ยของสองค่าเฉลี่ยอื่น

ริชาร์ด (Richard.1994) กล่าวว่า ถ้าการคำนวณค่า F พบว่ามีนัยสำคัญ คือมีความแม่นยำในการค้นพบว่า มีกลุ่มตัวอย่างที่แตกต่างกัน การเปรียบเทียบพหุคูณเป็นขั้นตอนหนึ่งที่ทำให้ได้ผลของกลุ่มที่แตกต่าง

เนล (Neil.1995) กล่าวว่า เมื่อวิเคราะห์ความแปรปรวนแบบทางเดียวแล้ว และผลที่ได้ปฏิเสธสมมติฐานหลัก สามารถที่จะพิจารณาได้ว่าค่าเฉลี่ยของประชากรที่นำมาพิจารณานั้นมีค่าไม่เท่ากันทั้งหมด ที่สุดก็ทำการพิจารณาครั้งสุดท้าย ซึ่งเป็นการหาค่าเฉลี่ยที่แตกต่างกัน หรือหาค่าเฉลี่ยที่มากที่สุด หรือหาค่าเฉลี่ยที่มากกว่าโดยปกติ เป็นความสัมพันธ์ของค่าเฉลี่ยทั้งหมด วิธีการที่จะนำมาใช้ในการค้นหานี้เรียกว่า การเปรียบเทียบพหุคูณ

สมจิต วัฒนายากุล (สมจิต วัฒนายากุล.2529:13) กล่าวว่า การเปรียบเทียบเชิงซ้อน หมายถึง การเปรียบเทียบทรีเมนต์ต่าง ๆ เพื่อหาว่าทรีเมนต์ใดบ้างที่ยังผลให้สมมติฐานหลักถูกปฏิเสธ

วิรัช วรรณรัตน์ (วิรัช วรรณรัตน์.2535:193) กล่าวว่า การเปรียบเทียบพหุคูณหรือการเปรียบเทียบภายหลังเป็นการทดสอบความแตกต่างของค่าเฉลี่ยภายหลังการวิเคราะห์ความแปรปรวนเพื่อหาความแตกต่าง

ดาร์วาร์ตัน พิกุลทอง (ดาร์วาร์ตัน พิกุลทอง.2537:351) กล่าวว่า การเปรียบเทียบพหุคูณเป็นวิธีการทดสอบเพื่อหาความแตกต่างระหว่างค่าเฉลี่ยประชากร

กัลยา วาณิชย์บัญชา (กัลยา วาณิชย์บัญชา.2539:160) กล่าวว่า การเปรียบเทียบพหุคูณเป็นการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยหลายๆคู่ในเวลาเดียวกันก็ต่อเมื่อปฏิเสธสมมติฐานหลัก

ระวีวรรณ พันธุ์พานิช (ระวีวรรณ พันธุ์พานิช.2540:51) กล่าวว่า การเปรียบเทียบพหุคูณอาจเรียกอีกอย่างหนึ่งว่า การเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่ม

จากความหมายต่างๆ ที่นักสถิติและนักการศึกษาทั้งหลายได้ให้ไว้ อาจจะสรุปได้ว่าการเปรียบเทียบพหุคูณ หมายถึง วิธีการทางสถิติที่มีขึ้นเพื่อเป็นการหาค่าเฉลี่ยทางสถิติที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ซึ่งมีทั้งการเปรียบเทียบแบบวางแผน และการเปรียบเทียบภายหลัง

การเปรียบเทียบระหว่างค่าเฉลี่ย

การเปรียบเทียบ (Comparison or Contrast)ระหว่างค่าเฉลี่ย เป็นความแตกต่างระหว่างค่าเฉลี่ย 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

$\hat{\psi}_i = \bar{Y}_1 - \bar{Y}_2$ เป็นความแตกต่างของค่าประมาณกลุ่มตัวอย่าง ถ้ามีการทดลอง 3 ระดับ จะได้ค่าเฉลี่ย 3 ค่า และได้ความแตกต่างดังนี้

$$\hat{\psi}_1 = \bar{Y}_1 - \bar{Y}_2$$

$$\hat{\psi}_2 = \bar{Y}_1 - \bar{Y}_3$$

$$\hat{\psi}_3 = \bar{Y}_2 - \bar{Y}_3$$

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

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

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

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

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

ได้ $\sum_{j=1}^k C_j = 0$ เมื่อ C_j อย่างน้อยหนึ่งตัวไม่เท่ากับ 0 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 หรือถ้าจะให้สะดวกในการเลือกการเปรียบเทียบ คือ ผลรวมของค่าสัมบูรณ์ของ C_j เท่ากับสอง

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

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

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

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

ชนิดของการเปรียบเทียบ

1. การเปรียบเทียบเป็นรายคู่ (Pairwise comparison)

เมื่อสัมประสิทธิ์ทุกตัวของการเปรียบเทียบครั้งหนึ่งยกเว้น 2 สัมประสิทธิ์เท่ากับศูนย์ การเปรียบเทียบนั้นเรียกว่า การเปรียบเทียบรายคู่ (Pairwise comparison) นอกเหนือจากนี้ไม่เป็นการเปรียบเทียบรายคู่ เมื่อมีค่าเฉลี่ย k จำนวน จะเท่ากับ $k(k-1)/2$ จากตัวอย่างข้างต้น มีจำนวนค่าเฉลี่ย 3 ค่า จึงมีการเปรียบเทียบรายคู่ $3(3-1)/2 = 3$ ดังนี้ (Kirk. 1995:115)

$$\psi_1 = 1\bar{Y}_1 + (-1)\bar{Y}_2 + 0\bar{Y}_3 = \bar{Y}_1 - \bar{Y}_2$$

$$\psi_2 = 1\bar{Y}_1 + 0\bar{Y}_2 + (-1)\bar{Y}_3 = \bar{Y}_1 - \bar{Y}_3$$

$$\psi_3 = 0\bar{Y}_1 + 1\bar{Y}_2 + (-1)\bar{Y}_3 = \bar{Y}_2 - \bar{Y}_3$$

ซึ่งเปรียบเทียบแบบรายคู่นี้ ค่าสัมประสิทธิ์ในสมการจะมีเพียง 2 ค่า

2. การเปรียบเทียบที่ไม่ใช่รายคู่หรือการเปรียบเทียบเชิงซ้อน (Nonpairwise or Complex comparison)

เป็นการเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่างมากกว่า 2 กลุ่มและจำนวนการเปรียบเทียบได้นั้นมีจำนวนมาก พิจารณาตัวอย่างในกรณี จำนวนเฉลี่ยสองค่าเฉลี่ยเปรียบเทียบกับค่าเฉลี่ยที่สาม(ระวีวรรณ พันธุ์พานิช.2540:52)

$$\psi_4 = \frac{1}{2}\bar{Y}_1 + \frac{1}{2}\bar{Y}_2 + (-1)\bar{Y}_3 = \frac{\bar{Y}_1 + \bar{Y}_2}{2} - \bar{Y}_3$$

$$\psi_7 = \frac{1}{3}\bar{Y}_1 + \frac{2}{3}\bar{Y}_2 + (-1)\bar{Y}_3 = \frac{(1)\bar{Y}_1 + (2)\bar{Y}_2}{3} - \bar{Y}_3$$

$$\psi_8 = \frac{1}{4}\bar{Y}_1 + \frac{3}{4}\bar{Y}_2 + (-1)\bar{Y}_3 = \frac{(1)\bar{Y}_1 + (3)\bar{Y}_2}{4} - \bar{Y}_3$$

$$\psi_9 = \frac{1}{5}\bar{Y}_1 + \frac{4}{5}\bar{Y}_2 + (-1)\bar{Y}_3 = \frac{(1)\bar{Y}_1 + (4)\bar{Y}_2}{5} - \bar{Y}_3$$

3.การเปรียบเทียบอิสระ (Orthogonality of Multiple Comparisons)

การเปรียบเทียบอิสระเป็นการพิจารณาสำหรับการเปรียบเทียบแบบคู่ หรือการเปรียบเทียบสองค่าในหนึ่งครั้ง ถ้ามี 3 การเปรียบเทียบในหนึ่งกลุ่มของการเปรียบเทียบ การเปรียบเทียบอิสระจะเป็นการตัดสินใจสำหรับการเปรียบเทียบค่าที่หนึ่งกับค่าที่สอง การเปรียบเทียบค่าที่หนึ่งกับค่าที่สาม และการเปรียบเทียบค่าที่สองกับค่าที่สาม ถ้าขนาดของกลุ่มตัวอย่างเท่ากันแล้ว การเปรียบเทียบอิสระของสองค่าเป็นการตัดสินใจโดยน้ำหนักของค่าเฉลี่ย ถ้า

$$\sum_{j=1}^J c_{1j}c_{2j} = 0$$

ซึ่งการเปรียบเทียบค่าหนึ่งกับสองเป็นอิสระ ในที่นี้ c_{1j} และ c_{2j} เป็นค่าน้ำหนักสำหรับการเปรียบเทียบค่าหนึ่งกับค่าสองที่ค่าเฉลี่ย j ถ้าผลรวมของค่าน้ำหนักไม่เท่ากับศูนย์แล้วการเปรียบเทียบนั้นไม่เป็นอิสระ (nonorthogonal) การเปรียบเทียบสองค่าในหนึ่งครั้งและการพิจารณาแบบคู่เป็นอิสระ ถ้าขนาดกลุ่มตัวอย่างเท่ากัน: (Toothaker.1991:17)

$$\begin{aligned}\hat{\psi}_1 &= \bar{Y}_1 - \bar{Y}_2 \\ \hat{\psi}_2 &= \bar{Y}_1 - \bar{Y}_3 \\ \hat{\psi}_3 &= \bar{Y}_3 - \bar{Y}_4 \\ \hat{\psi}_4 &= \bar{Y}_1 - \frac{1}{2}\bar{Y}_3 - \frac{1}{2}\bar{Y}_4\end{aligned}$$

สามารถเขียน C สำหรับค่าเฉลี่ย J ค่าในทุกการเปรียบเทียบลงในตาราง และคำนวณผลรวมของค่าน้ำหนักการเปรียบเทียบแต่ละคู่:

	ค่าเฉลี่ยที่ J			
	1	2	3	4
การเปรียบเทียบที่ 1	1	-1	0	0
การเปรียบเทียบที่ 2	1	0	-1	0
การเปรียบเทียบที่ 3	0	0	1	-1
การเปรียบเทียบที่ 4	1	0	-1/2	-1/2

$$1 \text{ vs } 2: \sum_{j=1}^J c_{1j} c_{2j} = (1)(1) + (-1)(0) + (0)(-1) + (0)(0) = 1$$

$$1 \text{ vs } 3: \sum_{j=1}^J c_{1j} c_{3j} = (1)(0) + (-1)(0) + (0)(1) + (0)(-1) = 0$$

$$1 \text{ vs } 4: \sum_{j=1}^J c_{1j} c_{4j} = (1)(1) + (-1)(0) + (0)(-1/2) + (0)(-1/2) = 1$$

$$2 \text{ vs } 3: \sum_{j=1}^J c_{2j} c_{3j} = (1)(0) + (0)(0) + (1)(-1) + (0)(-1) = -1$$

$$2 \text{ vs } 4: \sum_{j=1}^J c_{2j} c_{4j} = (1)(1) + (0)(0) + (-1)(-1/2) + (0)(-1/2) = 1.5$$

$$3 \text{ vs } 4: \sum_{j=1}^J c_{3j} c_{4j} = (0)(1) + (0)(0) + (1)(-1/2) + (-1)(-1/2) = 0$$

ดังนั้นการเปรียบเทียบค่าหนึ่งกับค่าสี่เป็นการเปรียบเทียบอิสระ แต่การเปรียบเทียบอื่นไม่เป็น การเปรียบเทียบอิสระ ภายในกลุ่มการเปรียบเทียบแบบคู่ที่เป็นไปได้ส่วนมากการเปรียบเทียบ ไม่เป็นอิสระ เช่นการเปรียบเทียบค่าที่หนึ่งกับค่าสอง สำหรับหนึ่งกลุ่มค่าเฉลี่ย J ค่าจำนวนที่ มากที่สุดของการเปรียบเทียบเป็นการเปรียบเทียบอิสระซึ่งแต่ละอันเป็น $J-1$

ประเภทของสถิติ

ส่วนหนึ่งของลักษณะเชิงซ้อนของหัวข้อการเปรียบเทียบพหุคูณนั้นมีความ แตกต่างโดยทั่วไปของสถิติ (Toothaker, 1991:12)

1. t - Statistic

สำหรับการเปรียบเทียบจะต้องทราบข้อมูล 3 อย่างที่เกี่ยวกับกลุ่มตัวอย่างคือ ค่าเฉลี่ย ค่าความแปรปรวน และการแจกแจงของกลุ่มตัวอย่างเป็นแบบปกติ เมื่อทราบข้อมูล ทั้ง 2 แล้วก็สามารถหาค่า t statistic ได้ดังนี้

$$t_{\hat{\psi}} = \frac{\hat{\psi} - \psi}{\sqrt{MS_w \sum_{j=1}^J \frac{c_j^2}{n}}}$$

สมมติฐานที่ทดสอบโดย t คือ

$$H_0 : \psi = 0$$

ที่

$$\hat{\psi} = \overline{Y_j} - \overline{Y_{j'}} \quad \text{สำหรับ } j \neq j'$$

ถ้ามีการเปรียบเทียบหนึ่งที่น่าสนใจ ซึ่งไปตามข้อตกลงเบื้องต้นของ ANOVA ความแปรปรวนเท่ากันและเป็นอิสระ t statistic สามารถที่จะทำการทดสอบสมมติฐานได้

2. q - Statistic

วิธีหนึ่งในสถิติพื้นฐานของการเปรียบเทียบพหุคูณ คือ range statistic หรือ Studentized range มีรูปแบบดังนี้

$$q = \frac{\bar{Y}_j - \bar{Y}_{j'}}{\sqrt{\frac{MS_W}{n}}} \quad \text{สำหรับ } j \neq j'$$

Studentized range มี สองค่าพารามิเตอร์ คือ J จำนวนของค่าเฉลี่ย และ df จากตารางมีให้ เลือกสำหรับ $\alpha = .05$ และ $\alpha = .01$ สำหรับการเปรียบเทียบรายคู่มีความสัมพันธ์ระหว่าง t กับ q statistic:

$$t = \frac{\bar{Y}_j - \bar{Y}_{j'}}{\sqrt{\frac{MS_W}{n} \sum_{j=1}^J c_j^2}} = \frac{q}{\sqrt{2}} = \frac{\bar{Y}_j - \bar{Y}_{j'}}{\sqrt{\frac{MS_W}{n}}} \frac{1}{\sqrt{2}}$$

ความแตกต่างของสองรูปแบบคือ t มีตัวส่วนเป็นค่าประมาณค่าเบี่ยงเบนมาตรฐานของการเปรียบเทียบซึ่งเกี่ยวกับ $\sum c_j^2 = 2$ สำหรับการเปรียบเทียบแบบรายคู่ ขณะที่ q มีตัวส่วนเป็นค่าประมาณความคลาดเคลื่อนมาตรฐานของหนึ่งค่าเฉลี่ยในการเปรียบเทียบ สำหรับ การเปรียบเทียบรายคู่ $\sum c_j^2$ เท่ากับ 2 เพราะ C_j คือ 1 และ -1 สำหรับสองค่าเฉลี่ยที่เริ่มการเปรียบเทียบ และเท่ากับศูนย์ที่ค่าเฉลี่ยตัวอื่น ดังนั้นตัวส่วนของ t เป็นค่าประมาณความคลาดเคลื่อนมาตรฐานของความแตกต่างระหว่างค่าเฉลี่ยของสองกลุ่มตัวอย่าง

3. F - Statistic

บางการเปรียบเทียบพหุคูณใช้ F statistic สำหรับการเปรียบเทียบ F นี้ มีความสัมพันธ์กับ t statistic คือ $t^2 = F$ ดังนั้นจึงมีค่า degree of freedom เพียงค่าเดียวในตัวเอง เพราะว่าทุกค่า F statistic เป็นสัดส่วนของสอง mean square คือ หนึ่งค่า F - statistic สำหรับการเปรียบเทียบเป็นสัดส่วนของ mean square สำหรับการเปรียบเทียบหารด้วย MS_w ผลรวมของ square และ mean square สำหรับการเปรียบเทียบสามารถพบได้เริ่มกับความสัมพันธ์

$$t^2 = F$$

$$t^2 = \frac{(\hat{\psi} - \psi)^2}{\frac{MS_w}{n} \sum_{j=1}^J c_j^2} = \frac{(\hat{\psi} - \psi)^2 / \left[\sum_{j=1}^J c_j^2 \right]}{MS_w} = F$$

จากสมการ เพราะ F เป็นสัดส่วนของสอง mean squares ได้

$$SS_{\hat{\psi}} = MS_{\hat{\psi}} = (\hat{\psi} - \psi)^2 / \left[\sum_{j=1}^J c_j^2 / n \right]$$

ที่ MS และ SS เท่ากันเพราะ มี degree of freedom เพียงค่าเดียวสำหรับการเปรียบเทียบ

การเปรียบเทียบพหุคูณ (Multiple Comparison Test) อาจจะเรียกอีกอย่างหนึ่งว่า การเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่ม ใช้เมื่อพบว่า F จากการวิเคราะห์ความแปรปรวนมีนัยสำคัญทางสถิติ ซึ่งต้องตอบสมมติฐานการวิจัยเพียงว่า มีประชากรบางกลุ่มที่มีค่าเฉลี่ยแตกต่างจากกลุ่มอื่น แต่ไม่ทราบว่ากลุ่มใดบ้างที่มีค่าเฉลี่ยต่างกัน จึงต้องใช้การเปรียบเทียบพหุคูณ ซึ่งทำได้ 2 วิธี

1. การเปรียบเทียบแบบวางแผน (Planned or Priori contrast) เป็นเทคนิควิธีการใช้เปรียบเทียบค่าเฉลี่ยของกลุ่มประชากรที่กำหนดไว้ล่วงหน้าก่อนการวิเคราะห์ความแปรปรวน โดยการเปรียบเทียบเฉพาะค่าเฉลี่ยของบางกลุ่มที่ผู้วิจัยสนใจเท่านั้น

2. การเปรียบเทียบภายหลัง (Post-hoc or Posteriori contrast) เป็นเทคนิคที่ใช้เปรียบเทียบค่าเฉลี่ยของกลุ่มประชากรหลังจากการปฏิเสธ H_0 การเปรียบเทียบความแตกต่างของค่าเฉลี่ยทุกกลุ่มที่เป็นไปได้ (ระวีวรรณ พันธุ์พานิช, 2540:51)

วิธีการเปรียบเทียบพหุคูณนั้นเคิร์ก (Kirk.1995:127-128) ได้นำเสนอวิธีการเปรียบเทียบพหุคูณแบบต่าง ๆ ดังแสดงในตาราง 1-2

ตาราง 1 วิธีการเปรียบเทียบพหุคูณที่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติ

	Orthogonal Contrast	Nonorthogonal Contrast
การทดสอบแบบวางแผน (A Priori Contrast)	1.ทดสอบ p-1 contrast สถิติทดสอบ t-test	2.ทดสอบ p-1 contrast แบบ มีกลุ่มร่วม สถิติทดสอบ Dunnett's Test 3.ทดสอบ C contrast สถิติทดสอบ Dunn-Sidak Test Holm's Test
การทดสอบภายหลัง (A Posteriori Contrast)		4.ทดสอบค่าเฉลี่ยรายคู่ทุกคู่ สถิติทดสอบ Turkey's Test Fisher-Hayter Test REGW F FQ and Q Test 5.ทดสอบทุก contrast สถิติทดสอบ Scheff's Test

ตาราง 2 วิธีการเปรียบเทียบพหุคูณที่ไม่เป็นไปตามข้อตกลงเบื้องต้นของค่าสถิติ

	Orthogonal Contrast	Nonorthogonal Contrast
การทดสอบแบบวางแผน (A Priori Contrast)	1.ทดสอบ p-1 contrast ที่ความแปรปรวนของประชากรแตกต่างกัน สถิติทดสอบ t-test ที่ชั้นของความเป็นอิสระ คำนวณด้วยวิธีของ Welch	2.ทดสอบ p-1 contrast แบบมีกลุ่มร่วม และกลุ่มตัวอย่างไม่เท่ากัน หรือ ความแปรปรวนของประชากรแตกต่างกัน สถิติทดสอบ Dunnett's Test (modified) 3.ทดสอบ C contrast ที่ความแปรปรวนของประชากรแตกต่างกัน สถิติทดสอบ Dunn-Sidak Test Holm's Test (ที่ชั้นของความเป็นอิสระของ Welch)
การทดสอบภายหลัง (A Posteriori Contrast)		4.ทดสอบค่าเฉลี่ยรายคู่ทั้งหมด ถ้ากลุ่มตัวอย่างไม่เท่ากัน สถิติทดสอบ Turkey-Kramer Test Fisher-Hayter Test ถ้าประชากรมีความแปรปรวนแตกต่างกัน สถิติทดสอบ Dunnett's T3 Test Dunnett's C Test Games-Howell Test 5.ทดสอบทุก contrast ถ้าประชากรมีความแปรปรวนแตกต่างกัน สถิติทดสอบ Brown-Forsythe Test

กรณีข้อมูลมีลักษณะบางประการไม่เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวน จะมีผลต่อความเปลี่ยนแปลงของระดับนัยสำคัญและความไวของตัวทดสอบซึ่งมีผลต่ออำนาจการทดสอบดังนี้

1. ถ้าข้อมูลมีการแจกแจงไม่เป็นโค้งปกติ

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

เพื่อแก้ปัญหาการแจกแจงประชากรที่ไม่เป็นโค้งปกติ และเพื่อให้การทดสอบมีความคงทน (test is robust) ซึ่งหมายถึงกรณีที่ข้อมูลมีลักษณะไม่เป็นไปตามข้อตกลงเบื้องต้น แต่ค่าระดับนัยสำคัญของการทดสอบยังคงใกล้เคียงกับค่าที่กำหนดไว้ควรกำหนดตัวอย่างให้ใหญ่เพียงพอหรือใช้สถิตินอนพาราเมตริก

2. ถ้าความแปรปรวนของประชากรไม่เท่ากันทุกกลุ่มหรือไม่มีความเป็นเอกพันธ์

เมื่อทดสอบความเป็นเอกพันธ์ของความแปรปรวนของกลุ่มประชากรที่มีข้อมูลในกลุ่มตัวอย่างต่างกันแล้วพบว่า ความแปรปรวนแตกต่างกันจะแก้ไขโดยเพิ่มขนาดของกลุ่มตัวอย่างให้ทุกกลุ่มมีขนาดเท่ากัน แต่หากจำเป็นต้องใช้จำนวนกลุ่มตัวอย่างไม่เท่ากัน การฝ่าฝืนข้อตกลงเบื้องต้นข้อนี้ จะมีผลอย่างมากต่อการทดสอบนัยสำคัญ (ระวีวรรณ พันธุ์พานิช.

2540:40) การทดสอบนัยสำคัญของความแตกต่างระหว่างความแปรปรวนของสองกลุ่ม ในกรณีที่ผู้วิจัยต้องการทดสอบความแปรปรวนของประชากรกลุ่มหนึ่ง มากกว่าความแปรปรวนของประชากรอีกกลุ่มหนึ่งหรือไม่ โดยใช้ข้อมูลจากกลุ่มตัวอย่างที่สุ่มมาจากประชากรและกรณีที่ต้องการทดสอบความเป็นเอกพันธ์ของความแปรปรวน (Homogeneous of variance) เพื่อทราบว่าได้ assume ว่าความแปรปรวนของประชากรเท่ากัน หรือไม่เท่ากันนั้น มีความถูกต้องเที่ยงตรงหรือไม่ มีวิธีการทดสอบหลายวิธี เช่น วิธีใช้ F – test วิธีของ Hartley วิธีของ Cochran เป็นต้น (บุญชม ศรีสะอาด. 2541:233)

โดยมีข้อตกลงเบื้องต้นว่า

1. ประชากรที่สุ่มตัวอย่างมา มีการแจกแจงของค่าตัวแปรที่ศึกษาเป็นแบบปกติ
2. กลุ่มตัวอย่างเป็นอิสระแก่กัน

การทดสอบความแปรปรวนที่เท่ากัน

เดตัน (Dayton.1970:32) กล่าวว่า ในแบบแผนการวิจัยแบบสุ่มสมบูรณ์นั้น ความแปรปรวนในแต่ละกลุ่มการจัดกระทำจะประมาณได้จากความแปรปรวนของคะแนนประชากร โดยทำการทดสอบความเท่ากันของความแปรปรวน คือ $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_p^2 = \sigma^2$ เมื่อ σ_j^2 แทน ความแปรปรวนของคะแนนประชากรกลุ่มที่ j แต่เมื่อทำการทดสอบกับกลุ่มตัวอย่างจะได้ ความแปรปรวนของคะแนนกลุ่มตัวอย่าง คือ $s_1^2, \dots, s_p^2$ เป็นการประมาณค่าความแปรปรวนของคะแนนประชากรวิธีการทดสอบความเท่ากันของความแปรปรวนที่ใช้กันอย่างแพร่หลายมีอยู่สองสามวิธี เช่น

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

วิธีของ Levene

Levene ได้นำเสนอวิธีการทดสอบความเท่ากันของความแปรปรวนของประชากร ซึ่งมีวิธีการคำนวณคล้ายกับการวิเคราะห์ความแปรปรวนแบบทางเดียว โดยการแทนค่าสังเกตแต่ละค่าด้วยค่าสัมบูรณ์ของส่วนเบี่ยงเบนระหว่างค่าสังเกตแต่ละค่ากับค่าเฉลี่ยของกลุ่มตัวอย่างนั้นๆ เป็นตัวทดสอบที่มีความแรงต่อลักษณะการแจกแจงของประชากรที่ไม่ปกติ (Levene, H. 1960:278-292 อ้างอิงจาก เขียวภา ไชยศรี. 2536:25)

ตัวสถิติทดสอบคำนวณได้จาก

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

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

การทดสอบด้วยวิธี Dunnett's T3 test

เสนอโดย Charles W. Dunnett ในปี ค.ศ. 1980 เพื่อเป็นวิธีการทดสอบความแตกต่างของค่าเฉลี่ยประชากรสองกลุ่ม ในกรณีที่มีความแปรปรวนของประชากรไม่เท่ากัน ซึ่งมีรูปแบบดังนี้ (Dunnett, 1980)

$$T3 = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}}}$$

เมื่อ	T3	แทน	การเปรียบเทียบพหุคูณของ Dunnett
	$\bar{X}_i$	แทน	ค่าเฉลี่ยของกลุ่ม i
	$\bar{X}_j$	แทน	ค่าเฉลี่ยของกลุ่ม j
	$\hat{\sigma}_i^2$	แทน	ค่าความแปรปรวนของกลุ่ม i
	$\hat{\sigma}_j^2$	แทน	ค่าความแปรปรวนของกลุ่ม j
	n_i	แทน	จำนวนตัวอย่างของกลุ่ม i
	n_j	แทน	จำนวนตัวอย่างของกลุ่ม j

ภายใต้สมมติฐานไร้ัยสำคัญ H_0 เป็นจริง กล่าวคือ $\mu_1 = \mu_2$ แล้วสถิติทดสอบ Dunnett's T3 test จะมีลักษณะการแจกแจงแบบ Studentized maximum modulus distribution of k ที่องศาความเป็นอิสระ (degree of freedom) (c,df) โดยที่ องศาความเป็นอิสระ(degree of freedom)นั้นเป็นค่าที่คำนวณโดยวิธีของ Welch's modified

$$df = \frac{\left[\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j} \right]^2}{\frac{\hat{\sigma}_i^4}{n_i^2(n_i - 1)} + \frac{\hat{\sigma}_j^4}{n_j^2(n_j - 1)}}$$

เกณฑ์การตัดสินใจจะปฏิเสธ H_0 เมื่อ $|T3| \geq m_{\alpha;c,df}$ โดยที่ $m_{\alpha;c,df}$ คือค่าวิกฤตที่เปิดตาราง Studentized maximum modulus distribution ที่มีระดับ ัยสำคัญ และองศาความเป็นอิสระ(degree of freedom) (c,df) ตามลำดับ

การทดสอบด้วยวิธี Games-Howell test

เสนอโดย Paul A. Games และ Johe f. Howell ในปี ค.ศ. 1976 เพื่อเป็นวิธีการทดสอบความแตกต่างของค่าเฉลี่ยประชากรสองกลุ่ม ในกรณีที่มีความแปรปรวนของประชากรไม่เท่ากัน ซึ่งมีรูปแบบดังนี้ (Games & Howell.1976)

$$GH = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j} \right) / 2}}$$

เมื่อ	GH	แทน	การเปรียบเทียบพหุคูณของ Games และ Howell
	$\overline{X}_i$	แทน	ค่าเฉลี่ยของกลุ่มi
	$\overline{X}_j$	แทน	ค่าเฉลี่ยของกลุ่มj
	$\hat{\sigma}_i^2$	แทน	ค่าความแปรปรวนของกลุ่มi
	$\hat{\sigma}_j^2$	แทน	ค่าความแปรปรวนของกลุ่มj
	n_i	แทน	จำนวนตัวอย่างของกลุ่มi
	n_j	แทน	จำนวนตัวอย่างของกลุ่มj

ภายใต้สมมติฐานไร้นัยสำคัญ H_0 เป็นจริง กล่าวคือ $\mu_1 = \mu_2$ แล้วสถิติทดสอบ Games-Howell test มีองศาความเป็นอิสระ(degree of freedom) (p,df) โดยที่องศาความเป็นอิสระ(degree of freedom)นั้นเป็นค่าที่คำนวณโดยวิธีของ Welch's modified

$$df = \frac{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}\right)^2}{\frac{\hat{\sigma}_i^4}{n_i^2(n_i-1)} + \frac{\hat{\sigma}_j^4}{n_j^2(n_j-1)}}$$

เกณฑ์การตัดสินใจจะปฏิเสธ H_0 เมื่อ $|GH| \geq q_{\alpha,p,df}$ โดยที่ $q_{\alpha,p,df}$ คือค่าวิกฤตที่เปิดตาราง Studentized rang distribution ที่มีระดับนัยสำคัญและองศาความเป็นอิสระ(degree of freedom) (p,df) ตามลำดับ

การทดสอบด้วยวิธี Brown-Forsythe test

เสนอโดย Morton b. Brown และ Alan B. Forsythe ในปี ค.ศ. 1974 เพื่อเป็นวิธีการทดสอบความแตกต่างของค่าเฉลี่ยประชากรสองกลุ่ม ในกรณีที่มีความแปรปรวนของประชากรไม่เท่ากัน ซึ่งมีรูปแบบดังนี้ (Brown & Forsythe. 1974)

$$BF = \frac{(\overline{X}_i - \overline{X}_j)^2}{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}}$$

เมื่อ	BF	แทน	การเปรียบเทียบพหุคูณของ Brown และ Fosite
	$\overline{X}_i$	แทน	ค่าเฉลี่ยของกลุ่มi
	$\overline{X}_j$	แทน	ค่าเฉลี่ยของกลุ่มj
	$\hat{\sigma}_i^2$	แทน	ค่าความแปรปรวนของกลุ่มi
	$\hat{\sigma}_j^2$	แทน	ค่าความแปรปรวนของกลุ่มj
	n_i	แทน	จำนวนตัวอย่างของกลุ่มi
	n_j	แทน	จำนวนตัวอย่างของกลุ่มj

ภายใต้สมมติฐานไร้นัยสำคัญ H_0 เป็นจริง กล่าวคือ $\mu_1 = \mu_2$ แล้วสถิติทดสอบ Brown-Forsythe test จะมีลักษณะการแจกแจงแบบ F ที่องศาความเป็นอิสระ (degree of freedom) โดยที่ df_2 องศาความเป็นอิสระ (degree of freedom) นั้นเป็นค่าที่คำนวณโดยวิธีของ Welch's modified

$$df_1 = n_i - 1$$

$$\text{และ} \quad df_2 = \frac{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j} \right)^2}{\frac{\hat{\sigma}_i^4}{n_i^2(n_i - 1)} + \frac{\hat{\sigma}_j^4}{n_j^2(n_j - 1)}}$$

เกณฑ์การตัดสินใจจะปฏิเสธ H_0 เมื่อ $BF \geq F_{\alpha; df_1, df_2}$ โดยที่ $F_{\alpha; df_1, df_2}$ คือค่าวิกฤตที่เปิดตาราง F ที่มีระดับนัยสำคัญและองศาความเป็นอิสระ (degree of freedom) ตามลำดับ

ความถนัดด้านเหตุผล

ความหมายของความถนัด

นักจิตวิทยาและนักการศึกษาได้ให้ความหมายของความถนัด แตกต่างกันไป ดังนี้

ฟรีแมน (Freeman.1966:431) กล่าวว่า ความถนัดเป็นผลรวมของคุณลักษณะต่างๆ ที่จะชี้ให้เห็นนิสัยของแต่ละคนในการที่จะได้มาซึ่งความรู้และทักษะหรือการตอบสนอง ดังนั้นแบบทดสอบที่วัดความถนัดจึงออกแบบเพื่อวัดศักยภาพของความสามารถในกิจกรรมเฉพาะภายในพิสัยที่จำกัด

ครอนบาค (Cronbach.1970:38) กล่าวว่า ความถนัดทางการเรียน หมายถึง ความสามารถทางสมองที่ร่วมกันทำงานเพื่อเพิ่มพูนความสำเร็จในกิจกรรมทางปัญญา

บิงแฮม (Bingham.1987:17) กล่าวว่า ความถนัดเป็นสภาวะที่แสดงถึงความเหมาะสมของบุคคลที่สำคัญประการแรกคือ ความพร้อมของบุคคลในการเพิ่มพูนความชำนาญให้แก่ตนเองหรือเป็นศักยภาพของบุคคลนั้น หรืออีกนัยหนึ่งคือความพร้อมที่จะสนใจในความสามารถของบุคคลนั้นๆ

ซวาล แพร์ตุนกุล (2517:50) กล่าวว่า ความถนัดเป็นความสามารถของบุคคลที่จะทำกิจกรรมใดกิจกรรมหนึ่ง จึงให้ความหมายของความถนัดว่า ความพยายามที่จะคาดคะเนหรือพยากรณ์ภายหน้าอันเป็นเรื่องราวของอนาคตกาล โดยอาศัยของเท็จจริงในปัจจุบันเป็นรากฐานว่าเด็กสามารถไปได้ไกลเพียงใด จะเรียนรู้สิ่งนั้น สิ่งนี้ได้เท่าไร ถ้าเขาได้รับประสบการณ์การฝึกฝนที่เหมาะสม

ทองหล่อ วิภาวรินทร์ (2523:16-17) ได้ให้ความหมายของความถนัดว่า

1. ความถนัด หมายถึง ปัญญาของบุคคล หรือเรียกว่าง่าย ๆ ว่าไหวพริบ หรือความฉลาด หรือความสามารถในการแก้ปัญหา โดยการนำความรู้หรือหลักวิชาไปแก้ปัญหาเฉพาะหน้า ซึ่งเป็นลักษณะของพลังงานความคิดที่สำคัญในการแก้ไขอุปสรรคและตัดสินใจเหตุการณ์เฉพาะหน้าได้ถูกต้อง

2. ความถนัดเป็นสมรรถภาพทางสมอง ที่ประกอบด้วยความสามารถเฉพาะหลายด้าน ซึ่งความสามารถแต่ละด้านมีคุณภาพแตกต่างกันออกไป ดังนั้นแต่ละคนจึงแตกต่างกันในเรื่องความสามารถของสมองอีกด้วย ฉะนั้นการวัดจึงต้องวัดความสามารถหลายด้าน ถ้าบุคคลใดมีความสามารถทางสมองด้านใดมากก็จะมี ความถนัดทางด้านนั้นด้วย

3. ความถนัดไม่ได้หมายถึงกรรมพันธุ์และพรหมลิขิต ไม่ใช่เป็นสมรรถภาพทางสมองที่ติดตัวมาแต่กำเนิด หากเป็นผลที่เกิดจากการที่ได้มีโอกาสฝึกฝน และคิดปัญหาหรือมีประสบการณ์ในด้านนั้นมากพอ

ลัวัน สายยศ และ อังคณา สายยศ (2541:27) กล่าวว่า ความถนัดหมายถึง ความสามารถที่บุคคลได้รับประสบการณ์ฝึกฝนตนเองและมีการสั่งสมไว้มากจนเกิดทักษะพิเศษที่เด่นชัดด้านใดด้านหนึ่งพร้อมที่จะปฏิบัติด้านนั้นได้เป็นอย่างดี

จากความหมายต่างๆ สามารถสรุปได้ว่า ความถนัดคือความสามารถของบุคคลที่สร้างมาจากการเรียนรู้และประสบการณ์ ทำให้บุคคลนั้นสามารถเรียนและประกอบกรงานต่างๆ ตามความสามารถนั้นได้ประสบความสำเร็จ

ทฤษฎีของเชาว์ปัญญาและความถนัด

ทฤษฎีเชาว์ปัญญา และความถนัดมีหลายทฤษฎี แต่ที่สำคัญดังนี้ (ลัวัน สายยศ และ อังคณา สายยศ.2541:42-52)

1. ทฤษฎีองค์ประกอบเดียว (Uni-Factor Theory) บางทีทฤษฎีนี้เรียกว่า Global Theory ผู้คิดทฤษฎีนี้คือ บิเนท์ และซิมอน (Binet & Simon) ทฤษฎีนี้เสนอโครงสร้างทางเชาว์ปัญญาเป็นลักษณะเป็นอันหนึ่งอันเดียวไม่แบ่งออกเป็นส่วนย่อย คล้ายกับความสามารถทั่วไป (General Ability)นั่นเอง ในปี ค.ศ.1905หรือปี พ.ศ.2448 บิเนท์ และซิมอน ได้สร้างข้อสอบวัดตามแนวความคิดของเขาเป็นครั้งแรก ข้อสอบนี้สร้างวัดเชาว์ปัญญาเป็นแบบ Global Measure คือ วัดออกมาเป็นคะแนนเดียว แล้วแปลความหมายมีปัญญาระดับใด

2. ทฤษฎีสององค์ประกอบ (Bi-Factor Theory) ทฤษฎีนี้เสนอโดยนักจิตวิทยาชาวอังกฤษ ชื่อ สเปียร์แมน (Charles Spearman) ในปี ค.ศ. 1927 เป็นทฤษฎีที่อาศัยการวิเคราะห์ทางสถิติ พบว่า กิจกรรมทางสมองทั้งหลาย เมื่อวิเคราะห์ดูแล้วมีองค์ประกอบร่วมกันอันหนึ่ง เรียกชื่อองค์ประกอบนี้ว่า องค์ประกอบทั่วไป (General Factor) เรียกว่า G – Factor และอีก

องค์ประกอบหนึ่งเป็นองค์ประกอบเฉพาะ (Specific Factor) เรียกว่า S – Factor และ
องค์ประกอบเฉพาะเหล่านี้มีกิจกรรมเฉพาะของมันเอง

3. ทฤษฎีหลายองค์ประกอบ (Multiple-Factor Theory) ทฤษฎีนี้เป็นที่ยอมรับกัน
อย่างกว้างขวางจากนักจิตวิทยาชาวอเมริกัน ผู้นำทฤษฎีนี้คือ เฮอร์สโตน (L.L. Thurstone) เสนอ
ทฤษฎีนี้เมื่อปี ค.ศ.1933 โดยการทำวิจัยโครงสร้างทางสมองอย่างกว้างขวาง และใช้หลักการ
วิเคราะห์องค์ประกอบ (Factor Analysis) ทำให้สามารถแยกแยะความสามารถทางสมองออกได้
หลายอย่าง แต่ที่เห็นชัดและสำคัญมี 7 ประการ คือ

- 3.1 องค์ประกอบด้านภาษา (Verbal Factor ใช้ตัวย่อว่า V)
- 3.2 องค์ประกอบด้านความคล่องในการใช้ถ้อยคำ (Word Fluency Factor
ใช้ตัวย่อว่า W)
- 3.3 องค์ประกอบด้านจำนวน (Number Factor ใช้ตัวย่อว่า N)
- 3.4 องค์ประกอบด้านมิติสัมพันธ์ (Space Factor ใช้ตัวย่อว่า S)
- 3.5 องค์ประกอบด้านความจำ (Memory Factor ใช้ตัวย่อว่า M)
- 3.6 องค์ประกอบด้านสังเกต (Perceptual Speed Factor ใช้ตัวย่อว่า P)
- 3.7 องค์ประกอบด้านเหตุผล (Reasoning Factor ใช้ตัวย่อว่า R)

4. ทฤษฎีไฮราซิคอล (Hierarchical Theory) ผู้นำทฤษฎีนี้คือ เวอร์นอน (Vernon)
เบิร์ท (Burt) ชาวอังกฤษ และฮัมฟรีย์ (Humphreys) ชาวอเมริกัน โดยเฉพาะ
เวอร์นอน ได้เสนอโครงสร้างทางสมองตามทฤษฎีสององค์ประกอบของสเปียร์แมนเสนอไว้ใน
ส่วนขององค์ประกอบทั่วไป (General Factor : G – Factor) แบ่งออกเป็นสององค์ประกอบใหญ่ๆ
คือ ความถนัดทางภาษา (Verbal Education ; V : ed) ความถนัดทางช่าง (Practical
Mechanical ; K : m) ซึ่งองค์ประกอบใหญ่ทั้งสองเรียกว่า Major Group องค์ประกอบใหญ่ทั้ง
สองยังแยกเป็นองค์ประกอบย่อยได้อีก เรียกว่า Minor Group Factors เช่น ด้าน V : ed ยังแบ่ง
ย่อยเป็นองค์ประกอบด้านภาษา (Verbal) องค์ประกอบด้านตัวเลข (Number) และอื่นๆ อีก ส่วน
K : m ได้แบ่งย่อยออกเป็นความรู้เชิงกล (Mechanic Information) มิติสัมพันธ์ (Spatial) ความ
สามารถในการใช้กลไกของกล้ามเนื้อ (Psychomotor Abilities) และอื่นๆ อีก ซึ่งแต่ละ
องค์ประกอบย่อย ยังแบ่งเป็นองค์ประกอบเฉพาะ (Specific Factor)

5. ทฤษฎีโครงสร้างทางสติปัญญา (The Structure of Intellect Theory) ผู้ตั้งทฤษฎีนี้
คือ กิลฟอร์ด (Guilford) นักจิตวิทยาชาวอเมริกัน ได้ศึกษาพัฒนามาจากทฤษฎีหลายองค์
ประกอบของเฮอร์สโตน ด้วยวิธีการวิเคราะห์องค์ประกอบแบบทดสอบวัดสติปัญญา แล้วเสนอ

โครงสร้างทางสมองในปี ค.ศ.1967 อธิบายโครงสร้างทางสมองในรูปแบบจำลองสามมิติ (Three – Dimensional Model) ดังนี้ (Guilford and Hoepfner.1971:20-21)

มิติที่ 1 วิธีการคิด (Operation) เป็นการปฏิบัติงานทางสมอง หรือขบวนการแบบต่าง ๆ ซึ่งจะเกิดขึ้นตามลำดับจากง่ายไปยาก เริ่มจากการรู้และเข้าใจ การจำ การคิดแบบบอกนกนัย การคิดแบบเอกนัย และการประมาณค่า

มิติที่ 2 เนื้อหา (Content) เป็นสิ่งเร้าต่าง ๆ ที่ได้รับจากตัวระบบประสาทสัมผัสทั้งหลายของมนุษย์ แบ่งเป็น 4 อย่างคือ ภาพ สัญลักษณ์ ภาษา และพฤติกรรม

มิติที่ 3 ผลการคิด (Products) แบ่งออกเป็น 6 ลักษณะ คือ หน่วย จำนวน ความสัมพันธ์ ระบบ การแปลงรูป และการประยุกต์

6. ทฤษฎีความสามารถทางสมองสองระดับ (Two – Level Theory of Mental Ability) ทฤษฎีนี้เสนอโดย เจนเซน (Jensen) เจนเซนเสนอทฤษฎีว่า ความสามารถทางสมองมีอยู่สองระดับ คือ ระดับ 1 (Level I) เป็นความสามารถด้านการเรียนรู้และการจำอย่างนกนัย นั่นคือเป็นความสามารถที่จะสะสมหรือเก็บข้อมูลไว้ได้และพร้อมที่จะระลึกนึกออกได้ ระดับ II (Level II) เป็นระดับการจัดกระทำทางสมองเป็นขั้นสร้างมโนภาพ เหตุผล และการแก้ปัญหา ระดับ II นี้เหมือนกับองค์ประกอบทั่วไป (G – Factor) นั่นเอง

ทฤษฎีพัฒนาการทางสติปัญญาของเพียร์เจต์กับการคิดหาเหตุผล

เพียร์เจต์ (De Cecco.1967:264-265) แบ่งลำดับขั้นของการพัฒนาทางสติปัญญาของเด็ก ซึ่งเกี่ยวข้องกับการคิดหาเหตุผลไว้ ดังนี้

ขั้นที่ 1 ขั้นการคิดก่อนสร้างสัจกับ พบในเด็กอายุ 2 – 4 ปี เด็กในขั้นนี้จะเรียนรู้โดยการกระทำซ้ำ ๆ กัน โดยใช้ประสาทสัมผัส แต่ก็ยังไม่สามารถสร้างสัจกับได้ จึงไม่รู้จักใช้สัญลักษณ์ เด็กจะสร้างแผนการรู้ขึ้นเอง เช่นเด็กอายุ 3 ขวบ ยังไม่มีความเข้าใจในเรื่องการเท่ากัน การจัดลำดับสิ่งของที่มีค่ากลาง ๆ การคิดหาเหตุผลจะมีลักษณะจากส่วนย่อยสู่ส่วนย่อย

ขั้นที่ 2 ขั้นการคิดโดยใช้ญาณ พบในเด็กอายุระหว่าง 4 – 7 ปี เด็กเริ่มมีการคิดหาเหตุผล แต่มีอย่างจำกัดและยังไม่อิสระ เด็กเริ่มเข้าใจสัญลักษณ์ และสามารถสร้างสัจกับได้บ้างแล้ว

ขั้นที่ 3 ขั้นการคิดด้วยรูปธรรม พบในเด็กอายุ 7 – 11 ปี ในขั้นนี้น้อยอย่างน้อยที่สุดจะมีการคิดหาเหตุผลที่สอดคล้องตามหลักตรรกศาสตร์ที่เพียร์เจต์เรียกว่าการใช้ตัวปฏิบัติการณ์เด็กจะรวบรวมสัจกับแล้วนำมาใช้สามารถคิดหาเหตุผลถูกต้องตามหลักตรรกศาสตร์เชิงคณิตศาสตร์ได้

ขั้นที่ 4 ขั้นคิดตามหลักตรรกศาสตร์พบในเด็กอายุ 11 – 15 ปี เด็กจะนำความคิดที่ถูกต้องตามหลักตรรกศาสตร์ไปใช้แก้ปัญหาในสถานการณ์ที่เป็นรูปธรรม มีการคิดที่อิสระขึ้น และที่สำคัญก็คือ จะมีการคิดย้อนกลับอย่างสมบูรณ์ และยังสามารถคิดในสิ่งที่เป็นนามธรรมได้ สามารถสรุปและอนุมานจากสมมติฐานได้ และสามารถหาความรู้ในลักษณะที่เป็นวิทยาศาสตร์ โดยสอดคล้องตามหลักตรรกศาสตร์ได้ จึงนับว่าเด็กในขั้นนี้มีการคิดหาเหตุผลสูงสุด

ดี เซคโค (De Cecco.1967 : 265-266) กล่าวถึงพัฒนาการทางสติปัญญาตามทฤษฎีของเพียเจท์ว่า ในขั้นการคิดเป็นรูปธรรมเด็กมีการคิดหาเหตุผลแล้ว และสิ่งที่ปรากฏมาขั้นนี้ ได้แก่

1. การจำแนกประเภท เป็นความสามารถที่จะจำแนกวัตถุสิ่งของออกเป็นหมวดหมู่
2. การหาความสัมพันธ์ เป็นการคิดหาความสัมพันธ์ไม่สมมาตรได้
3. การเข้าใจเกี่ยวกับจำนวน เป็นผลร่วมของความสามารถในการจำแนกประเภทและการหาความสัมพันธ์

การวัดความสามารถด้านเหตุผล

ความเหตุผล (Reasoning) หมายถึง การคิดอย่างมีวิจารณ์ญาณ แล้ววินิจฉัยลงสรุปอย่างถูกต้อง ในแบบทดสอบมาตรฐานนั้นการวัดเหตุผลมักจะอยู่ในกลุ่มความสามารถใหญ่ ๆ 3 ด้าน คือ ด้านความสามารถในการจำแนก ด้านความสามารถในการหาความสัมพันธ์หรืออุปมาอุปไมย และความสามารถในการสรุปอ้างอิงโดยหลักตรรกะ แต่จากการจัดจำแนกทักษะการคิดแบบเหตุผลของลิปแมน, จาคอบส์ และโคลแมน จำแนกความสามารถในการคิดเป็น 3 กลุ่มคือ Mental acts, Reasoning skill และ Inquiry skill ในส่วนของทักษะทางเหตุผลจะ (ล้วน สายยศ และอังคณา สายยศ. 2541:106)

1. การจำแนกประเภท (Classification)

บางทีก็เรียกว่าการจัดประเภท ผู้ที่มีความถนัดทางด้านนี้จะต้องเป็นผู้ที่จัดจัดจำแนกแยกประเภทของสิ่งต่างๆ ที่กำหนดไว้ให้อย่างชาญฉลาด ผู้ตอบต้องใช้ความสามารถด้านการวิเคราะห์ด้วย เพราะจะจำแนกสิ่งหนึ่งสิ่งใดโดยอาศัยโครงสร้าง มโนภาพ หน้าที่ ฯ เปรียบเทียบกับกลุ่มนั้นจะต้องแยกให้ดีจึงจะมองเห็นความแตกต่างกันได้ การจำแนกประเภทนี้จะแยกได้ทั้งด้านภาษาและรูปภาพ

1.1 การจำแนกประเภทแบบภาษา ส่วนใหญ่ใช้คำศัพท์ที่เป็นพื้นๆ แต่มีสิ่งที่ยากที่จำแนกได้ยาก ซึ่งอาจเป็นตัวร่วมบางประการในโครงสร้างและคุณสมบัติของคำนั้นก็ได้ การเขียนข้อสอบจะต้องให้ผู้ตอบคิดหลายแง่มุมแต่ก็ให้มีทางที่ถูกที่สุดเพียงทางเดียว ถ้าแยกแยะแล้วมีส่วนถูกได้หลายแง่หลายมุมไม่ควรนำมาออกข้อสอบ ควรนำกลุ่มคำชุดใหม่ที่เหมาะสมกว่ามาแทน

1.2 การจำแนกประเภทแบบภาพ วิธีการออกคล้ายกับการจำแนกประเภทแบบภาษา ที่ต่างกันเพียงแปลภาษาให้เป็นภาพเท่านั้นเองการออกข้อสอบการจำแนกประเภทแบบภาพนี้สามารถแบ่งได้ 2 อย่าง คือ ภาพจริงหรือภาพเสมือนจริง เช่น รูปนก รูปกา รูปคน รูปต้นไม้ เป็นต้น กับภาพทรงเรขาคณิต ภาพแบบนี้จะขีดเขียนอะไรก็ได้แต่ก็ต้องมีแนวความคิดเป็นเครื่องพิจารณาเพื่อจะได้ทราบว่าภาพใดไม่อยู่ในพวกเดียวกับเขา

2. แบบทดสอบอุปมาอุปไมย (Analogy)

เป็นแบบทดสอบที่วัดความถนัดทางการวิเคราะห์ โดยเฉพาะด้านการวิเคราะห์ความสัมพันธ์นั้นคือ ผู้สอบหรือผู้ที่ทำข้อสอบจะต้องมีความถนัดทางด้านนี้อย่างดีด้วยถึงจะทำได้ หลักการใหญ่ๆ ของอุปมาอุปไมยอยู่ที่ผู้ออกข้อสอบจะต้องวิเคราะห์ความสัมพันธ์ของสิ่งต่างๆ ให้ได้ก่อน แล้วเอาความสัมพันธ์อาจเป็นมโนภาพ โครงสร้าง หน้าที่ หรือลักษณะต่างๆ มาสัมพันธ์กันให้เหมาะสมกันมากที่สุด รูปแบบของการอุปมาอุปไมยจะเป็นการหาสิ่งที่คู่กันมาคู่หนึ่ง แล้วนำไปเปรียบเทียบกับคู่อื่นๆ ที่มีลักษณะการสัมพันธ์ในแนวเดียวกัน ในการเขียนข้อสอบอุปมาอุปไมยทั้งด้านภาษาและภาพ

2.1 อุปมาอุปไมยด้านภาษา อาศัยภาษาที่เป็นศัพท์ทั่วไป มีมโนภาพ โครงสร้าง และหน้าที่ร่วมกันดังกล่าวนมาแล้วข้างต้น แล้วนำมาให้ผู้สอบวิเคราะห์หาความสัมพันธ์ที่มีโอกาสใกล้เคียงความเป็นจริงให้มากที่สุด แบบทดสอบชนิดนี้ต้องการวัดความถนัดในการหาความสัมพันธ์ ระหว่างสิ่งของและเรื่องราวต่างๆ และขยายสถานการณ์นั้นออกไปสู่สิ่งอื่นหรือสถานการณ์อื่นที่มีความสัมพันธ์ในทำนองเดียวกันหรือลักษณะเดียวกันกับของเดิม

2.2 อุปมาอุปไมยด้านภาพ แบ่งออกเป็น 2 ประเภท คือภาพประเภทเสมือนกับภาพประเภททรงเรขาคณิต

ก. อุปมาอุปไมยภาพเสมือน วิธีการสร้างใช้หลักการเกี่ยวกับการอุปมาอุปไมยทางด้านภาษาดังที่กล่าวมาแล้ว ถ้านึกถึงภาพยากก็ใช้วิธีเขียนเป็นภาพที่สามารถเขียนเป็นภาพได้ แล้วให้ผู้ชำนาญการวาดภาพ เขียนหรือวาดเป็นภาพอีกทีหนึ่ง การวาดอย่างให้เป็นแบบที่อูๆ มากนัก ควรเป็นภาพที่มีอาการประกอบด้วยก็จะดี อุปมาอุปไมยแบบภาพเสมือนใช้ได้ดีกับเด็กอายุน้อยๆ หรือระดับประถมลงไปมีความเหมาะสมกว่าเด็กระดับโต

ข. อุปมาอุปไมยภาพทรงเรขาคณิต หลักการสร้างอาศัยความสัมพันธ์ในโครงสร้าง มโนภาพ เช่นเดียวกับการเขียนข้อสอบอุปมาอุปไมยที่กล่าวไปแล้ว สิ่งที่ต้องเปลี่ยนแปลงสำคัญก็คือ การกำหนดภาพทรงเรขาคณิตในรูปแบบต่างๆ ที่จะต้องมีตัวร่วมบางอย่างเกี่ยวพันกันสองภาพแรกก่อน แล้วนำความคิดจากสองภาพแรกนั้นไปสร้างภาพที่เกี่ยวข้องอีกคู่หนึ่ง เริ่มแรกนั้นต้องมองหาทางออกให้ดีกว่าภาพสองคู่ที่จะต้องเกี่ยวพันกันหรือมีความสำคัญได้ในทาง

หนึ่งทางใดแน่ๆ เมื่อเกิดความแน่ใจแล้วจึงเขียนภาพที่เป็นตัวลงชั้นนอกนอกเลียนแบบตัวที่ถูกโดยพยายามให้ผิดเพี้ยนจากตัวถูก แต่ก็ต้องสร้างภาพให้มีโอกาสทำท่าจะเป็นไปได้ด้วยจึงจะดี

3. แบบอนุกรมมิติภาพหรืออนุกรมมิติ

ข้อสอบลักษณะนี้เป็นอนุกรมมิติแบบภาพ ซึ่งคล้าย ๆ กับอนุกรมคณิตศาสตร์แบบสัมพันธ์กันดังที่กล่าวมาแล้ว อนุกรมแบบภาพธรรมดาจะเป็นไปในทิศทางเดียว แต่ก็ถือว่าเป็นการวัดเหตุผลเหมือนกัน หลักการสร้างพยายามหาแนวโน้มของความเป็นไปของภาพไว้ในชุดแรกก่อนแล้วถามว่าต่อไปน่าจะเป็นภาพใด

4. แบบทดสอบสรุปความ

การออกข้อสอบแบบนี้เป็น การออกที่มุ่งวัดเหตุผลเป็นประการสำคัญ และเป็นแบบทดสอบเหตุผลชนิดหนึ่งที่ต้องใช้ภาษาค่อนข้างมาก ถึงแม้ว่าข้อสอบแบบนี้จะใช้ภาษาเป็นส่วนใหญ่ แต่ก็เป็นการใช้ภาษาเพื่อใส่เรียงหาเหตุผลเรียกตามภาษาทดสอบว่า Verbal Reasoning โครงสร้างของคำถามเป็นแบบคณิตศาสตร์แนวหนึ่งที่เรียกว่า ตรรกวิทยา นั่นคือการเขียนข้อสอบแบบนี้ประกอบด้วยเหตุใหญ่และเหตุย่อย เมื่อมีเหตุมาเป็นเครื่องพิจารณาแล้วก็สามารถประเมินลงสรุปได้ว่าเป็นอย่างไร สิ่งที่ต้องระวังคือแต่ละเหตุที่ยกมาจะต้องถือเป็นจริงในสาระของมัน ถ้าเหตุใดเหตุหนึ่งไม่เป็นจริงการลงสรุปจะทำได้ยาก ขนาดบางทีทุก ๆ เหตุเป็นจริง การลงสรุปไม่สามารถเป็นจริงได้ก็มี ในการเขียนข้อสอบประเภทนี้มักจะมีตัวเลือกหนึ่งที่เขียนว่า “ไม่สามารถลงสรุปได้” หรือจะใช้คำอื่นใดก็ได้ในทำนองเดียวกัน กล่าวอย่างย่อข้อสอบประเภทนี้วัดความถนัดในการรวบรวมสาเหตุหรือข้อเท็จจริงต่าง ๆ ลงเป็นข้อยุติที่มีความหมายได้อย่างสมเหตุสมผลตามเงื่อนไขที่กำหนดไว้นั่นเอง

5. ตัวร่วมหรือตัวต่าง

ข้อสอบประเภทนี้จะยกสิ่งต่าง ๆ มาให้พิจารณา ส่วนใหญ่จะเป็นคำหรือใช้ภาพแทนก็ได้ เมื่อยกมาแล้วให้ผู้สอบพิจารณาดูตัวร่วมของมันว่าน่าจะเป็นอะไร ก่อนจะสามารถหาตัวร่วมหรือมโนภาพนี้จำเป็นจะต้องวิเคราะห์ทุก ๆ คำให้ดี อาจจะต้องใช้จินตนาการโยงความสัมพันธ์แต่ละสิ่งอย่างดีแล้วนำมาผสมกลมกลืนกันเป็นสิ่งที่ใหม่ที่สามารถเอ่ยขึ้นแล้วรับรู้และเข้าใจได้ตรงกัน วิธีการเขียนจึงมักแยกออกเป็น 2 วิธี คืออาจหาตัวร่วมโดยตรง กับหาตัวที่ทำให้สิ่งนั้นแตกต่างกันได้

5.1 หากตัวร่วม พยายามแยกแยะดูว่าคำหรือศัพท์ที่กำหนดมาให้มีอะไรเป็นตัวร่วมกันผู้ออกข้อสอบจะต้องมีความสามารถสูงทั้งด้านภาษาและด้านการวิเคราะห์แต่เนื่องจากวิธีการคิดใช้สมองด้านเหตุผลมากกว่าด้านภาษาจึงให้มาอยู่ในองค์ประกอบด้านเหตุผล

5.2 หากตัวต่าง ข้อสอบลักษณะนี้หนักไปทางเหตุผลด้านภาษาเหมือนกัน เพราะมุ่งให้ผู้สอบพยายามวิเคราะห์ความหมายหรือสิ่งที่เกี่ยวข้องของคำเหล่านั้นว่าที่แตกต่างกันเพราะอะไร

6. ความสามารถด้านวิเคราะห์

ความจริงการวิเคราะห์ก็คือเหตุผลนั่นเอง แต่การออกข้อสอบที่ให้ข้อเฉพาะลงไปอย่างนี้มีจุดประสงค์จะเป็นข้อสอบให้ผู้ตอบคิดหาความสัมพันธ์เกี่ยวข้องของตัวแปร จากสถานการณ์ที่สมมติขึ้นมา แม้กระทั่งชื่อ สถานที่ เหตุการณ์ก็สมมติทั้งนั้น การสร้างสถานการณ์จึงต้องเขียนให้มีความเกี่ยวพันกันซับซ้อนมากหน่อย มิฉะนั้นผู้อ่านก็ไม่ได้ใช้ความสามารถด้านวิเคราะห์ความยากในการออกข้อสอบอยู่ที่จุดนี้

การใช้แบบทดสอบความถนัด

การประเมินผลสิ่งใดๆ จะให้ได้ผลดีมีความถูกต้องที่เป็นจริงจำเป็นจะต้องวัดสิ่งนั้นให้ละเอียดทุกแง่มุมและจะต้องได้ข้อมูลที่มีความเที่ยงตรงแน่นอนจากสิ่งที่วัดนั้นด้วยเขาวีปัญญาหรือความถนัดเป็นคุณลักษณะชนิดหนึ่งที่มีอยู่ในแต่ละบุคคลเมื่อเราเชื่อว่าคุณลักษณะนี้มีอยู่ก็ควรจะสามารสอบวัดได้แต่จะวัดออกมาโดยวิธีการใดเป็นเรื่องของเทคนิคการวัด

ข้อมูลจากการวัดความถนัดของมนุษย์จึงเป็นข้อมูลที่มีประโยชน์ในการประเมินความสามารถของแต่ละบุคคลอย่างดี เป็นข้อมูลที่เพิ่มความสมบูรณ์จากการประเมินโดยข้อมูลด้านผลสัมฤทธิ์ ความรู้สึกนึกคิดและการประพฤติปฏิบัติจริงๆ การใช้แบบทดสอบความถนัดชุดใดสำหรับสอบวัดเพื่อประเมินนักเรียนกลุ่มหนึ่งนั้น จะต้องขึ้นอยู่กับความมุ่งหมายของผู้ใช้แบบทดสอบ อย่างไรก็ตามการใช้แบบทดสอบความถนัดมาใช้เพื่อประเมินผลนั้น มักใช้เพื่อพยากรณ์ผลภายหน้าเป็นประการสำคัญ ผลภายหน้าดังกล่าวอาจเป็นผลสัมฤทธิ์อย่างใดอย่างหนึ่งก็ได้ การใช้แบบทดสอบในกรณีนี้มักเป็นชุด ชุดหนึ่งๆ จะประกอบด้วยฉบับย่อยๆ ที่วัดองค์ประกอบความสามารถหรือความถนัดด้านต่างๆ สาเหตุที่ใช้แบบทดสอบความถนัดเป็นชุดหรือหลายองค์ประกอบ ก็เพื่อความแม่นยำในการพยากรณ์ เพราะผลสัมฤทธิ์ในกิจกรรมหนึ่งๆ นั้นมิใช่มาจากองค์ประกอบเดียว แต่เป็นผลมาจากหลายๆ องค์ประกอบ ตัวพยากรณ์ผลสัมฤทธิ์ตัวหนึ่งๆ จึงไม่จำเป็นต้องเหมือนกันเสมอไป (ล้วน สายยศ และอังคณา สายยศ. 2541:241)

งานวิจัยที่เกี่ยวข้อง

งานวิจัยต่างประเทศ

แรมเซย์ (Ramsey.1978:479-485) ได้ศึกษาเปรียบเทียบอำนาจการทดสอบระหว่างการเปรียบเทียบพหุคูณแบบรายคู่ ศึกษาถึง Experimentwise Type I Error Rate โดยวิธีเทคนิคมอนติคาร์โล ผลการศึกษาพบว่า วิธี Peritz's F เป็นวิธีที่สูงสุดในวิธีที่นำมาศึกษา 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 ทฤษฎีแวกล้อมในการพัฒนาวิธีการเปรียบเทียบภายหลังทั้ง 3 เป็นแนวทางเลือกพื้นฐานของไค-สแควร์ ตามทฤษฎีของ Scheffe การนำเสนอนี้เป็นการแสดงให้เห็นการประยุกต์วิธีการเปรียบเทียบพหุคูณแบบเปรียบเทียบภายหลังสำหรับสถิตินอนพาราเมตริก

แบลร์ และเจมส์ (Blair & James.1980:309-335) ได้ศึกษาเรื่องการเปรียบเทียบอำนาจการทดสอบของผลรวมสถิติของ Wilcoxon ที่วิธี Student's t statistic ภายใต้การแจกแจงที่ไม่ปกติหลายรูปแบบ ในการศึกษาครั้งนี้ได้ใช้เทคนิคมอนติคาร์โลในการสมมติข้อมูลเพื่อการเปรียบเทียบค่าเฉลี่ยที่เป็นอิสระสองกลุ่มโดยเป็นการสร้างรูปแบบการแจกแจงข้อมูล 6 รูปแบบ ดังนี้ (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

ลี ยู-ฟาง (Li YU-Fang.1988:6489) ได้ศึกษาพัฒนาอำนาจการประมาณและขนาดกลุ่มตัวอย่างของวิธีการเปรียบเทียบพหุคูณ 7 วิธี ผลการศึกษาพบว่า Fisher LSD และวิธี Duncan multiple rang test เป็นวิธีทดสอบที่ดีที่สุด ใน 7 วิธีที่นำมาศึกษา ขณะที่ Bonferoni และ Sidak เป็นสถิติทดสอบที่มีอำนาจน้อยที่สุด

วิลคอกซ์ (Wilcox.1989:269-278) ได้ศึกษาเรื่องการประมาณค่าสำหรับความแปรปรวนที่ไม่เท่ากันเมื่อมีการเปรียบเทียบค่าเฉลี่ยในการวิเคราะห์ความแปรปรวนแบบทางเดียวและ

แบบสองทาง มีเอกสารมากมายที่แสดงถึงการไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบค่า F คือค่าความแปรปรวนไม่เท่ากัน ทำให้ไม่มีความแกร่งในการวิเคราะห์ความแปรปรวน ในการศึกษาครั้งนี้ Wilcoxon ได้อธิบายและเปรียบเทียบ 2 วิธีการเปรียบเทียบในกรณีค่าความแปรปรวนแตกต่างกันของการวิเคราะห์ความแปรปรวนแบบสองทางวิธีการหนึ่งมีพื้นฐานมาจากการปรับปรุงวิธีของ Wilcoxon (1988) ที่สำหรับการวิเคราะห์ความแปรปรวนแบบทางเดียว ซึ่งเป็นรูปแบบพื้นฐานสำหรับการพิจารณาวิธีนี้เพื่อการวิเคราะห์ความแปรปรวนแบบสองทาง อีกวิธีหนึ่งเป็นการขยายผลการใช้ของวิธีของ James ผลการศึกษาพบว่า วิธีที่มีรากฐานมาจากวิธีของ Wilcoxon นั้นสามารถนำไปใช้ได้ง่ายกว่าและมีประสิทธิภาพมากกว่า วิธีที่นำมาจากวิธีของ James

เจฟเฟอเรีย และมิเชลลา (Jefferey & Michela, 1995:343-362) ได้ศึกษาอำนาจการทดสอบและอัตราความคลาดเคลื่อนประเภทที่ 1 กับวิธีการเปรียบเทียบพหุคูณแบบรายคู่ ภายใต้ความแปรปรวนไม่เท่ากัน โดยศึกษาจากวิธีการเปรียบเทียบพหุคูณ 9 วิธี ผลการศึกษาพบว่า วิธี Dunn และวิธีที่พัฒนามาจากวิธีของ Bonferroni มีความแกร่งดีกว่าวิธีของ 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 อันเป็นผลเนื่องมาจากตัวอย่างที่ใช้ไม่เป็นตัวแทนที่ดีของประชากร ในการศึกษาครั้งนี้ได้นำข้อมูลมาวิเคราะห์ในลักษณะเป็นข้อมูลแจกแจงทางเดียวและข้อมูลแจกแจงสองทาง และทำการทดสอบความมีนัยสำคัญของความแตกต่างระหว่างค่าเฉลี่ยของประชากรโดยใช้ระเบียบวิธีการต่างๆ แล้วนำผลที่ได้จากการทดสอบทั้งหมดทุกวิธีมาเปรียบเทียบกัน ซึ่งผลที่ได้จากการวิจัยสรุปได้ว่า ในการทดสอบความมีนัยสำคัญของความ

แตกต่างระหว่างค่าเฉลี่ยประชากร 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 โดยทั่วๆ ไปวิธีของ Least Significant Difference test ให้ผลสรุปหรือผลการทดสอบที่น่าเชื่อถือได้ดีกว่าอีก 2 วิธีที่กล่าวมาแล้ว

บุปผา จงเลิศตระกูล (2532:บทคัดย่อ) การเปรียบเทียบค่าเฉลี่ยของประชากรนั้น โดยทั่วไปจะใช้วิธีการวิเคราะห์ความแปรปรวน เมื่อทำการปฏิเสธสมมติฐานว่างแล้ว อาจจะทำให้การเปรียบเทียบเชิงพหุหรือหาช่วงความเชื่อมั่นร่วม โดยงานวิจัยนี้ได้เสนอวิธีการเลือกประชากรที่ดี เมื่อประชากรมีการแจกแจงแบบเอกซ์โพเนนเชียล เมื่อทำการเปรียบเทียบพารามิเตอร์ของ K ประชากรกับประชากรควบคุมหรือประชากรมาตรฐานตามค่าพารามิเตอร์แสดงตำแหน่งที่ไม่ทราบค่าได้แบ่งการพิจารณา เป็นดังนี้

การเลือกประชากรที่มีค่าพารามิเตอร์แสดงตำแหน่งมากกว่าหรือน้อยกว่าพารามิเตอร์แสดงตำแหน่งของประชากรควบคุม เป็นประชากรที่ดีเมื่อส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากันในทุก ๆ ประชากรแต่ไม่ทราบค่าและขนาดตัวอย่างของแต่ละประชากรเท่ากันเป็นค่าที่กำหนดให้ แบ่งเป็น 2 กรณี คือ กรณีข้อมูลทั่วไปและกรณีข้อมูล เช่น เซอร์ประเภทที่ 2 ค่า lower และ upper percentage points ที่คำนวณได้จากทั้ง 2 กรณี แสดงในตารางจำนวนประชากร (k) ขนาดตัวอย่าง (n) องค์ตำแหน่งความเป็นอิสระ (v) และสัมประสิทธิ์เชื่อมั่นร่วม (p^*) เมื่อกำหนดให้ $k = 2, 3, \dots, 10$ $n = 2, 3, \dots, a$ $v = 3, 4, \dots, a$ $p^* = 0.90, 0.95, 0.99$ ซึ่งจะนำผลที่คำนวณได้ไปใช้ในการหาช่วงความเชื่อมั่นร่วมขอบเขตล่าง และขอบเขตบนแต่ละกรณีเพื่อนำมาใช้เป็นเกณฑ์การตัดสินใจเลือกประชากรที่ดีการเลือกประชากรที่มีค่าพารามิเตอร์แสดงตำแหน่งมากกว่าพารามิเตอร์แสดงตำแหน่งของประชากรมาตรฐาน เป็นประชากรที่ดีเมื่อส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากันในทุก ๆ ประชากร คือมีค่าเท่ากับ 1 นั้น เพื่อศึกษาประสิทธิภาพของวิธีอนุบรรพโดยการเปรียบเทียบวิธีอนุบรรพกับวิธีทั่วไป ซึ่งพิจารณาจากขนาดตัวอย่างโดยเฉลี่ยของวิธีอนุบรรพและขนาดตัวอย่างที่กำหนดไว้คงที่โดยวิธีทั่วไป ผลจากการเปรียบเทียบขนาดตัวอย่างของวิธีการทั้ง 2 นี้ขนาดตัวอย่างที่ใช้เพื่อการตัดสินใจเลือกประชากรที่ดี โดยวิธีอนุบรรพจะมีขนาดเล็กกว่าขนาดตัวอย่างที่กำหนดไว้คงที่โดยวิธีทั่วไป สำหรับค่า a, B และสามเหลี่ยม แต่ละชุดที่กำหนดให้ คือ 0.01, 0.05, 0.10 และ 0.2, 0.5, 1.0 ตามลำดับ

นันทา วงษ์วิโรจน์ (2532:บทคัดย่อ) ได้ศึกษาเรื่องสถิติทดสอบที่มีความแกร่งสำหรับทดสอบความเท่ากันของค่าเฉลี่ยประชากร การศึกษาครั้งนี้มีวัตถุประสงค์ที่จะศึกษาเปรียบเทียบสถิติทดสอบที่มีความแกร่งสำหรับทดสอบความเท่ากันของค่าเฉลี่ยประชากรโดยใช้สถิติทดสอบ 3 วิธีคือ สถิติทดสอบแบบ ANOVA F-TEST และสถิติทดสอบแบบ Trimmed W และสถิติทดสอบแบบ Trimmed F โดยจะศึกษาถึงความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบทั้ง 3 วิธี ในที่นี้ศึกษากรณีที่ประชากรทั้ง 3 ชุด มีการแจกแจงแบบเดียวกัน คือการแจกแจงแบบปกติ และการแจกแจงแบบปกติปลอมปน ขนาดตัวอย่างที่ใช้แบ่งเป็น 2 กรณีคือขนาดตัวอย่างเท่ากันและไม่เท่ากัน สำหรับข้อมูลที่ใช้ในการทดลองครั้งนี้จำลองด้วยโปรแกรมคอมพิวเตอร์ โดยใช้เทคนิคมอนติคาร์โล ซึ่งกระทำซ้ำกันหนึ่งพันครั้งในแต่ละกรณี ผลการศึกษาสรุปได้ว่า ความแกร่งของการทดสอบ โดยพิจารณาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 โดยใช้เกณฑ์ในการพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของ Cochran (อ้างจาก Ramsey 1980:337-349) คือ ถ้าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 อยู่ในช่วง (0.007,0.015) ที่ระดับนัยสำคัญ .01 และถ้าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 อยู่ในช่วง (0.04,0.06) ที่ระดับนัยสำคัญ .05 จะถือว่าการทดสอบนี้สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ พบว่า การแจกแจงประชากร ขนาดตัวอย่าง อัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน เปอร์เซ็นต์การตัดข้อมูลที่ปลายหางของการแจกแจง และระดับนัยสำคัญมีผลต่อความแกร่งของการทดสอบอำนาจการทดสอบ เมื่อประชากรมีการแจกแจงแบบปกติที่มีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานเท่ากันทุกประชากรแล้ว สถิติทดสอบแบบ ANOVA F-TEST จะมีอำนาจการทดสอบสูงสุดในทุกขนาดตัวอย่างและระดับนัยสำคัญที่ศึกษา แต่เมื่อประชากรมีการแจกแจงแบบปกติปลอมปนพบว่า สถิติทดสอบแบบ Trimmed W และ Trimmed F จะให้ค่าอำนาจการทดสอบสูงกว่าสถิติทดสอบแบบ ANOVA F-TEST ซึ่งการเลือกใช้สถิติทดสอบแบบ Trimmed W หรือ Trimmed F นั้นพบว่า ความสัมพันธ์ระหว่างอัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน อัตราส่วนของค่าเฉลี่ยและเปอร์เซ็นต์การตัดข้อมูลที่ปลายหางของการแจกแจง จะมีผลต่ออำนาจการทดสอบของสถิติทดสอบทั้ง 2 นอกจากนี้พบว่า ปัจจัยที่มีผลกระทบต่อค่าอำนาจการทดสอบของค่าสถิติทดสอบทั้ง 3 มากที่สุดเมื่อกำหนดให้ปัจจัยอื่นคงที่คือ อัตราส่วนของส่วนเบี่ยงเบนมาตรฐานของประชากร ส่วนปัจจัยที่มีผลกระทบรองลงมาก็คือ ลักษณะการแจกแจงของประชากร ขนาดตัวอย่าง เปอร์เซ็นต์การตัดข้อมูลที่ปลายหางของการแจกแจง ที่ระดับนัยสำคัญตามลำดับ

นันทวัน บำรุงสวัสดิ์ (2533:บทคัดย่อ) การวิจัยครั้งนี้มีวัตถุประสงค์ที่จะศึกษาเปรียบเทียบวิธีทดสอบความเท่ากันของค่าเฉลี่ยประชากรเมื่อความแปรปรวนของประชากรไม่เท่ากัน โดยสถิติทดสอบ 3 วิธี คือ สถิติทดสอบแบบ Brown and Forsythe สถิติทดสอบแบบ Marascuilo และสถิติทดสอบแบบ ANOVA F-Test โดยจะศึกษาถึงความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบทั้ง 3 วิธี ในวิทยานิพนธ์นี้ ศึกษาจากกรณีที่ประชากรเป็น 3 และ 6 กลุ่ม ที่มีการแจกแจงแบบเดียวกัน คือ การแจกแจงแบบปกติ ที่มีอัตราส่วนของความแปรปรวนเท่ากันและไม่เท่ากัน สำหรับข้อมูลที่ใช้ในการทดลองครั้งนี้จำลองด้วยโปรแกรมคอมพิวเตอร์ โดยใช้เทคนิคมอนติคาร์โล ซึ่งกระทำซ้ำกัน 1,000 ครั้ง ในแต่ละกรณี ผลการศึกษาสรุปผลที่สำคัญได้ดังนี้ 1. เมื่อประชากรมีอัตราส่วนของความแปรปรวนเท่ากัน สถิติทดสอบแบบ Brown and Forsythe สถิติทดสอบแบบ Marascuilo และสถิติทดสอบแบบ ANOVA F-Test สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ แต่ประชากรมีอัตราส่วนของความแปรปรวนไม่เท่ากัน สถิติทดสอบแบบ Brown and Forsythe และสถิติแบบ Marascuilo สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ดีกว่าสถิติทดสอบแบบ ANOVA F-Test

2. อำนาจของการทดสอบ เมื่อขนาดตัวอย่างเล็กและประชากรมีการแจกแจงแบบปกติที่มีค่าอัตราส่วนของความแปรปรวนเท่ากันทุกประชากร สถิติทดสอบแบบ ANOVA F-Test จะมีอำนาจการทดสอบสูงสุด แต่ที่ขนาดตัวอย่างใหญ่ สถิติทดสอบทั้ง 3 วิธี มีอำนาจการทดสอบสูงสุดเท่ากัน ในทุกระดับนัยสำคัญที่ศึกษา แต่เมื่อประชากรมีอัตราส่วนของความแปรปรวนไม่เท่ากันแล้วพบว่า สถิติทดสอบแบบ Brown and Forsythe และสถิติทดสอบแบบ Marascuilo ให้ค่าอำนาจการทดสอบสูงกว่าสถิติทดสอบแบบ ANOVA F-Test ซึ่งการเลือกใช้สถิติทดสอบแบบ Brown and Forsythe หรือสถิติทดสอบแบบ Marascuilo นั้น พบว่าความสัมพันธ์ระหว่างอัตราส่วนของความแปรปรวน และอัตราส่วนค่าเฉลี่ย มีผลต่ออำนาจการทดสอบของสถิติทดสอบทั้ง 2 วิธีดังกล่าว

พหล ศักดิ์กะทัศน์ (2534:บทคัดย่อ) ได้ศึกษาเรื่อง การเปรียบเทียบประสิทธิภาพของวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยในการวิเคราะห์ความแปรปรวน ในการศึกษาครั้งนี้มีจุดประสงค์เพื่อเปรียบเทียบประสิทธิภาพการทดสอบของวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ย 3 วิธี คือวิธี Unrestricted LSD (U-LSD) วิธี Bonferroni (Dunn) T-test(Bon) และวิธี Murphys Gap LSD(MG-LSD) โดยข้อมูลมีการแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 ทำการทดลองด้วยเทคนิคมอนติคาร์โล โดยจำลองการทดสอบด้วยเครื่องคอมพิวเตอร์จำนวน 1000 รอบ สำหรับแต่ละสถานการณ์ในการคำนวณหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 โดยใช้เกณฑ์ของ Braley กล่าวคือความน่าจะเป็นของ

ความคลาดเคลื่อนประเภทที่ 1 ต้องอยู่ในช่วง (0.05,0.15) ที่ระดับนัยสำคัญที่.01 และช่วง (0.025,0.075) ที่ระดับนัยสำคัญที่ .05 จึงถือว่าสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และอำนาจการทดสอบ โดยมีจำนวนสิ่งทดลองเท่ากับ 2 ถึง 10 ทั้งแผนการทดลองแบบสุ่มสมบูรณ์ และแผนการทดลองแบบบล็อกสุ่มสมบูรณ์ และจำนวนซ้ำมี 4 ระดับ คือ 5, 10, 15 และ 20 ในการทดลองแบบสุ่มสมบูรณ์ และจำนวนบล็อกมี 3 ระดับคือ 5, 7 และ 10 ในแผนการทดลองแบบบล็อกสุ่มสมบูรณ์ ผลการศึกษาพบว่า วิธี Murphys Gap LSD เหมาะสำหรับแผนการทดลองแบบสุ่มสมบูรณ์ เมื่อจำนวนสิ่งทดลองเท่ากับ 2—5 และเหมาะสมสำหรับแบบแผนการทดลองแบบบล็อกสุ่มสมบูรณ์ เมื่อจำนวนสิ่งทดลองเท่ากับ 2-8 ส่วนวิธี Unrestricted LSD เหมาะสำหรับแผนการทดลองแบบสุ่มสมบูรณ์ เมื่อจำนวนสิ่งทดลองเท่ากับ 6-10 แต่เมื่อแผนการทดลองเป็นแบบบล็อกสุ่มสมบูรณ์กรณีที่สิ่งทดลองเท่ากับ 9-10 ปรากฏว่าทั้ง 3 วิธีดังกล่าวไม่มีวิธีใดที่เหมาะสมที่จะใช้ในการทดสอบ

กึ่งทอง ยงยุทธมีชัย (2538:บทคัดย่อ) ได้ศึกษาเรื่องการเปรียบเทียบสถิติทดสอบค่าเฉลี่ยของประชากรที่มีความแปรปรวนไม่เท่ากัน : กรณีศึกษาสำหรับแผนการทดลองแบบสุ่มสมบูรณ์ การศึกษาครั้งนี้จุดประสงค์เพื่อศึกษาเปรียบเทียบตัวสถิติทดสอบค่าเฉลี่ยประชากรที่มีความแปรปรวนไม่เท่ากัน โดยใช้ตัวสถิติทดสอบแบบ ANOVA F TEST การแปลงข้อมูลเป็นค่าลอการิทึม ตัวสถิติทดสอบแบบ Trimmed F ตัวสถิติทดสอบแบบ Brown & Forsythe และตัวสถิติทดสอบแบบเอฟที่ใช้ค่าเฉลี่ยของ Graybill & Deal โดยใช้หลักการพิจารณาถึงความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ซึ่งใช้เกณฑ์การพิจารณาของ Bradley และอำนาจการทดสอบ ประชากรที่ใช้ในการศึกษามีการแจกแจงแบบปกติที่มีอัตราส่วนความแปรปรวนเท่ากันและไม่เท่ากัน ขนาดตัวอย่างที่ใช้ในการศึกษาแบ่งเป็น 2 กรณีคือขนาดตัวอย่างเท่ากันและไม่เท่ากันระดับนัยสำคัญของการทดลองคือ .01 และ .05 ข้อมูลที่ใช้ในการทดลองครั้งนี้จำลองด้วยโปรแกรมคอมพิวเตอร์โดยใช้เทคนิคมอนติคาร์โล ทำซ้ำ 600 ครั้งในแต่ละกรณี ผลการศึกษพบว่า ตัวสถิติทดสอบทั้ง 5 วิธี สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ในกรณีที่ประชากรมีการแจกแจงแบบปกติที่มีอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานเท่ากัน แต่กรณีที่อัตราส่วนของส่วนเบี่ยงเบนมาตรฐานแตกต่างกัน ตัวสถิติทดสอบแบบเอฟที่ใช้ค่าเฉลี่ยของ Graybill & Deal และตัวสถิติทดสอบแบบ Brown & Forsythe สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ตัวสถิติทดสอบ ANOVA F TEST และตัวสถิติทดสอบแบบเอฟที่ใช้ค่าเฉลี่ยของ Graybill & Deal มีค่าอำนาจการทดสอบสูงในกรณีที่ประชากรมีการแจกแจงแบบปกติ ที่มีอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานเท่ากัน แต่ในกรณีที่อัตราส่วนของส่วน

เบี่ยงเบนมาตรฐานแตกต่างกันเป็น (1:1.1:1.2), (1:1.3:1.4) และ (1:1.8:2) ทุกขนาดตัวอย่าง ตัวสถิติทดสอบแบบเอฟที่ใช้ค่าเฉลี่ยของ Graybill & Deal มีค่าอำนาจในการทดสอบสูงแต่ในกรณีที่อัตราส่วนของส่วนเบี่ยงเบนมาตรฐานแตกต่างกันเป็น (1:2:3) และ (1:3:5) สามารถใช้ตัวสถิติทดสอบแบบ Brown & Forsythe ทุกขนาดตัวอย่าง หรือตัวสถิติทดสอบแบบเอฟที่ใช้ค่าเฉลี่ยของ Graybill & Deal แต่มีข้อจำกัดว่าขนาดตัวอย่างแต่ละกลุ่มต้องแตกต่างกันมาก ค่าอำนาจการทดสอบของตัวสถิติทดสอบจะมีค่าลดลงเมื่ออัตราส่วนของส่วนเบี่ยงเบนมาตรฐานแตกต่างกันมากขึ้น และค่าอำนาจการทดสอบของตัวสถิติทดสอบจะมีค่าเพิ่มมากขึ้น เมื่อเพิ่มขนาดตัวอย่างหรือระดับนัยสำคัญเพิ่มขึ้น

วิชชุดา ศรีโสภา (2539:บทคัดย่อ) ได้ศึกษาเรื่องการศึกษาการเปรียบเทียบพหุ การศึกษาครั้งนี้มีจุดประสงค์เพื่อเปรียบเทียบวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยในแผนการทดลองแบบสุ่มสมบูรณ์ 4 วิธี คือ วิธี Unrestricted LSD, Murphys Gap LSD, Tukey's(H) และ Murphys Gap Unrestricted LSD โดยข้อมูลมีการแจกแจงแบบปกติมีค่าสัมประสิทธิ์ความผันแปร 5%, 10%, 15%, 20% และ 30% มีจำนวนสิ่งทดลองเท่ากับ 3, 4, 5, 6 และ 10 จำนวนซ้ำในแต่ละสิ่งทดลองมีจำนวนไม่เท่ากัน 3 ระดับคือ เพิ่มทีละ 3 เพิ่มทีละ 5 และเพิ่มทีละ 10 โดยมีจำนวนซ้ำเริ่มต้นเป็น 5 ทำการทดลองด้วยเทคนิคมอนติคาร์โล โดยจำลองการทดลองด้วยเครื่องคอมพิวเตอร์จำนวน 1000 รอบ สำหรับแต่ละสถานการณ์ที่กำหนดในการคำนวณหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ซึ่งใช้เกณฑ์ในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของ Bradley และจะทำการหาอำนาจการทดสอบเมื่อวิธีการนั้นผ่านเกณฑ์ควบคุมความคลาดเคลื่อนประเภทที่ 1 ผลการศึกษาพบว่าวิธีการ Unrestricted LSD จะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีที่สุดและได้ค่าอำนาจการทดสอบสูงที่สุดในทุกสถานการณ์ที่ทำการทดลอง

กลมทิพย์ ปรัชญชรินทร์ (2541:บทคัดย่อ) การวิจัยครั้งนี้มีวัตถุประสงค์ที่จะศึกษาเปรียบเทียบอำนาจของการทดสอบ (power of the test) ของสถิติทดสอบความแตกต่างของค่าเฉลี่ยของประชากร 3 และ 5 กลุ่ม ได้แก่ตัวสถิติทดสอบเอฟ (F) ตัวสถิติทดสอบครัสคัล-วัลลิส (KW) ตัวสถิติทดสอบนอร์มอลสกอว์ (NS) และตัวสถิติทดสอบแบบดัดแปลงแบบอย่างต่อเนื่อง (CA) เมื่อสุ่มอย่างมาจากประชากรที่มีการแจกแจงเหมือนกัน ได้แก่การแจกแจงแลมดาของคูร์กการแจกแจงปกติ การแจกแจงแกมมา และการแจกแจงลอการิทึมขนาดตัวอย่างที่กำหนดเท่ากัน เท่ากับ 5, 10, 20, 30, 40 และ 50 ระดับในสำคัญเท่ากับ 0.01, 0.05 และ 0.10 ในการวิจัยครั้งนี้ ใช้เทคนิคการจำลองแบบมอนติคาร์โลกระทำซ้ำ 1,000 ครั้ง ในแต่ละกรณี ผลการวิจัยสรุปได้ดังนี้ อำนาจการทดสอบพิจารณา 2

กรณี คือ กรณีแรก ประชากรมีการแจกแจงปกติ พบว่าระดับนัยสำคัญ 0.01 และ 0.05 ตัวสถิติทดสอบเอฟมีอำนาจการทดสอบสูงสุด และที่ระดับนัยสำคัญ 0.10 ตัวสถิติทดสอบนอร์มอลสกอว์มีอำนาจการทดสอบสูงสุด กรณีที่สอง การแจกแจงของประชากรไม่ใช่แบบปกติ โดยทั่วไป ตัวสถิติทดสอบครัสคัล-วัลลิสมีอำนาจการทดสอบสูงสุด ยกเว้น กรณีความเบ้อยู่ระหว่าง 0.0-0.5 และความโด่งอยู่ในระดับต่ำปานกลางพบว่า ตัวสถิติทดสอบเอฟมีอำนาจการทดสอบสูงกว่า เมื่อ $n \leq 10$ และตัวสถิติทดสอบนอร์มอลสกอว์มีอำนาจการทดสอบสูงกว่า เมื่อ $n \geq 10$ ค่าอำนาจการทดสอบจะแปรผันตามความเบ้ ความโด่ง จำนวนกลุ่มตัวอย่าง ขนาดตัวอย่างและระดับนัยสำคัญ โดยเรียงลำดับจากมากไปหาน้อย

งานวิจัยส่วนใหญ่ที่ศึกษาเกี่ยวกับวิธีทดสอบความแตกต่างของค่าเฉลี่ยนั้น จะเป็นการศึกษาถึงอำนาจการทดสอบ และค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่มีการทดสอบทางสถิติที่ต่างกัน และศึกษาถึงรูปแบบของลักษณะข้อมูลที่น่ามาใช้กับสถิติทดสอบนั้นๆ ว่าให้ผลการทดสอบเป็นอย่างไร

จากการศึกษาเอกสารและผลงานวิจัยที่เกี่ยวข้องข้างต้น จะเห็นว่างานวิจัยที่ต้องการเปรียบเทียบค่าเฉลี่ยนั้น นอกจากเลือกใช้เทคนิคการวิเคราะห์ความแปรปรวนที่เหมาะสมกับข้อมูลที่ศึกษาแล้ว ยังต้องพบปัญหาในการเลือกวิธีการเปรียบเทียบพหุคูณอีกด้วย ซึ่งมีหลายวิธีให้เลือก เช่น วิธี Dunnett's T3 วิธี Games-Howell และวิธี Brown-Forsythe ที่มีวิธีการคำนวณที่คล้ายกันแต่จะแตกต่างกันที่การแจกแจงของแต่ละวิธีนั้นแตกต่างกัน โดยวิธี Dunnett's T3 และวิธี Games-Howell นั้นที่ผ่านมายังไม่ค่อยมีผู้ใดทำการศึกษา และยังพบว่า มีตัวแปรบางตัวที่ส่งผลต่อการเปรียบเทียบพหุคูณวิธีต่าง ๆ เช่น ขนาดกลุ่มตัวอย่าง ลักษณะการแจกแจงข้อมูล เป็นต้น ซึ่งผลการวิจัยที่ผ่านมายังไม่มีผู้ศึกษาเปรียบเทียบผลการเปรียบเทียบพหุคูณรายคู่ด้วยวิธี Dunnett's T3 วิธี Games-Howell และวิธี Brown-Forsythe ว่าแต่ละวิธี แต่ละขนาดกลุ่มตัวอย่างเท่าไรถึงจะให้ผลการเปรียบเทียบแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ ใน กรุงเทพมหานคร จำนวน 114 โรงเรียน จำนวนนักเรียน 122,828 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ กรุงเทพมหานคร จำนวน 583 คน เชื่อมั่นได้ 95% (ล้วน สายยศ และอังคณา สายยศ. 2540:400 อ้างอิงจาก Yamane.1967) สุ่มจากประชากรโดยวิธี Two-Stage Random Sampling

ขั้นแรก สุ่มโรงเรียนจากโรงเรียนมัธยมศึกษาในกรุงเทพมหานคร โดยวิธีการสุ่มอย่างง่าย (Simple Random Sampling) ได้ 5 โรงเรียน

ขั้นที่สอง สุ่มห้องเรียนโดยวิธีการสุ่มแบบแบ่งชั้น(Stratified Random Sampling) โดยให้ระดับชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 เป็นชั้นของการสุ่ม(Strata) และห้องเรียนเป็นหน่วยการสุ่ม(Sampling unit) ได้ 16 ห้อง ในการวิจัยครั้งนี้กลุ่มตัวอย่าง 583 คน ดังแสดงในตาราง 3

ตาราง 3 จำนวนโรงเรียนและห้องเรียนที่ทำการสุ่มได้ของชั้นมัธยมศึกษาปีที่ 2, 4 และ 6

โรงเรียน	ม.2			ม.4			ม.6		
	จำนวน			จำนวน			จำนวน		
	ห้อง	ห้อง ที่สุ่ม ได้	นัก เรียน	ห้อง	ห้อง ที่สุ่ม ได้	นัก เรียน	ห้อง	ห้อง ที่สุ่ม ได้	นัก เรียน
ดอนเมืองทหารอากาศบำรุง	14	2	59	14	2	79	11	-	-
เศรษฐบุตธำเพ็ญ	15	1	32	6	-	-	6	1	37
พุทธจักรวิทยา	12	-	-	16	2	88	16	2	77
บางกะปิสุขุมวิท	14	1	33	12	1	38	12	2	84
สีกัน(วัฒนารักษ์อุปถัมภ์)	14	2	56	7	-	-	6	-	-
รวม	69	6	180	55	5	205	51	5	198

เครื่องมือที่ใช้ในการศึกษา

เครื่องมือที่ใช้ในการศึกษาค้างนี้เป็นแบบทดสอบวัดความถนัดด้านเหตุผล ซึ่งผู้วิจัยนำมาจากเครื่องมือของ พิกุล เกตุประดิษฐ์ ที่ทำการศึกษารื่อง การวิเคราะห์องค์ประกอบความถนัดที่สัมพันธ์กับผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนมัธยมศึกษาตอนปลาย โดยใช้แบบทดสอบสรุปความ ที่สร้างขึ้นเป็นแบบทดสอบปรนัย แบบเลือกตอบมีคำตอบที่ถูกเพียงข้อเดียวจำนวนฉบับละ 15 ข้อ ซึ่งมีค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ .668 และของวิชา ภาควิชา ที่ศึกษาเรื่อง สมรรถภาพสมองบางประการที่สัมพันธ์กับผลสัมฤทธิ์ทางการเรียนวิชาฟิสิกส์ของนักเรียนชั้นมัธยมศึกษาตอนปลาย โดยใช้แบบทดสอบสรุปความ ที่สร้างขึ้นเป็นแบบทดสอบปรนัย แบบเลือกตอบมีคำตอบที่ถูกเพียงข้อเดียวจำนวนฉบับละ 15 ข้อ ซึ่งมีค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ .726 โดยผู้วิจัยนำมาหาคูณภาพเครื่องมืออีกครั้ง มีวิธีการหาคูณภาพเครื่องมือดังนี้

1. นำแบบทดสอบวัดความถนัดด้านเหตุผลไปทดลองสอบครั้งที่ 1 กับนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ที่ไม่ใช่กลุ่มตัวอย่างในการศึกษาจำนวนประมาณ 200 คน แล้วนำไปวิเคราะห์หาความเชื่อมั่นทั้งฉบับ ด้วยวิธี KR – 20 ของ คูเดอร์-ริชาร์ดสัน ซึ่งได้ค่าความเชื่อมั่น 0.634

2. จัดทำรูปเล่มของแบบทดสอบวัดความถนัดด้านเหตุผล พร้อมคำชี้แจงในการทำแบบทดสอบ เพื่อนำไปสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างในการศึกษา

ตัวอย่างแบบทดสอบวัดความถนัดด้านเหตุผล

คำชี้แจง 1. แบบทดสอบฉบับนี้เป็นแบบทดสอบปรนัยชนิดเลือก มีจำนวน 30 ข้อให้เวลาในการทำ 20 นาที

2. การตอบให้นักเรียนพิจารณาข้อความที่กำหนดให้ในการหาข้อสรุปอย่างสมเหตุผล แล้วเลือกข้อสรุปว่าควรเป็นข้อใดจาก ก. ถึง จ.

(0) เปิดบางตัวเป็นนก ไก่ทุกตัวเป็นนก นกทุกตัวกินหนอน สรุปว่าอย่างไร

- ก. เปิดทุกตัวเป็นไก่
- ข. ไก่ทุกตัวกินหนอน
- ค. เปิดทุกตัวกินหนอน
- ง. ไก่บางตัวเป็นเปิด
- จ. ยังสรุปไม่ได้แน่นอน

เฉลย ข. ยังสรุปไม่ได้แน่นอน

(00) หมอทุกคนเป็นเศรษฐี เศรษฐีบางคนตระหนี่ นายสมศักดิ์เป็นหมอ ดังนั้น

- ก. คนตระหนี่เป็นหมอ
- ข. นายสมศักดิ์เป็นเศรษฐี
- ค. คนตระหนี่เป็นเศรษฐี
- ง. นายสมศักดิ์เป็นตระหนี่
- จ. ยังสรุปไม่ได้แน่นอน

เฉลย ข. นายสมศักดิ์เป็นเศรษฐี

(00) นางงามทุกคนสวย คนสวยส่วนมากผิวขาว คนผิวขาวบางคนกลัวแดด ดังนั้น

- ก. นางงามทุกคนผิวขาว
- ข. คนสวยทุกคนผิวขาว
- ค. คนสวยส่วนมากกลัวแดด
- ง. นางงามบางคนกลัวแดด
- จ. ยังสรุปไม่ได้แน่นอน

เฉลย ง. นางงามบางคนกลัวแดด

เกณฑ์การให้คะแนน

ตอบถูกได้ 1 คะแนน

ตอบผิดได้ 0 คะแนน

วิธีเก็บรวบรวมข้อมูล

1. ขอนหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูล
2. ติดต่อโรงเรียนที่เป็นกลุ่มตัวอย่างในการศึกษา ขออนุญาตผู้บริหารโรงเรียน เพื่อกำหนด วัน เวลา ที่ทำการทดสอบ
3. วางแผนในการดำเนินการสอบโดยผู้วิจัยเป็นผู้ดำเนินการสอบด้วยตนเอง จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่ทำการทดสอบในแต่ละครั้ง
4. นำแบบทดสอบไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างทั้งหมด ตามวัน เวลาที่นัดหมายไว้
5. ชี้แจงให้นักเรียนที่เป็นกลุ่มตัวอย่างเข้าใจ และทราบถึงวัตถุประสงค์ในการทำแบบทดสอบครั้งนี้พร้อมทั้งขอความร่วมมือในการทำแบบทดสอบด้วย
6. อธิบายวิธีการทำแบบทดสอบวัดความถนัดด้านเหตุผล ซึ่งมีทั้งหมด 30 ข้อ เป็นข้อสอบปรนัยแบบเลือกตอบ ให้นักเรียนพิจารณาข้อความที่กำหนดให้ในการหาข้อสรุปอย่างสมเหตุสมผลแล้วเลือกข้อสรุปว่าควรเลือกจากข้อใดจาก ก ถึง จ
7. หลังจากได้ดำเนินการสอบกลุ่มตัวอย่างครบทุกโรงเรียนที่ทำการสุ่มมาแล้วนั้น นำกระดาษคำตอบมาตรวจให้คะแนน ดังนี้ ถ้าตอบถูกให้ 1 คะแนน ถ้าตอบผิดหรือไม่ตอบหรือตอบมากกว่าหนึ่งข้อให้ 0 คะแนน แล้วนำผลมาบันทึกข้อมูลและวิเคราะห์ทางสถิติ

การวิเคราะห์ข้อมูล

1. นำข้อมูลที่ได้มาจากการทดสอบของกลุ่มตัวอย่างขนาด 583 คน มาบันทึกเป็นประชากรเทียม
2. สุ่มข้อมูลจากกลุ่มตัวอย่างให้มีขนาด 20 คน 40 คน 80 คน และ 120 คน ซึ่งแต่ละขนาดจะสุ่มมา 30 กลุ่ม โดยทำการสุ่มข้อมูลแต่ละกลุ่มแต่ละขนาดด้วยการสุ่มแบบใส่คืน
3. นำข้อมูลทุกกลุ่มทุกขนาดมาทดสอบความเป็นเอกพันธ์ของความแปรปรวนของกลุ่มตัวอย่างด้วยวิธีของ Levene ถ้ามีกลุ่มตัวอย่างใดที่ไม่เป็นเอกพันธ์ของความแปรปรวน ให้ทำการสุ่มแบบใส่คืนใหม่จนครบตามจำนวนกลุ่มที่ต้องการ
4. นำข้อมูลทุกกลุ่มทุกขนาดที่ไม่เป็นเอกพันธ์ของความแปรปรวนมาวิเคราะห์ความแปรปรวน แล้วจึงทดสอบการเปรียบเทียบพหุคูณทั้ง 3 วิธี คือ วิธี Dunnett's T3 วิธี Games-Howell และ วิธี Brown-Forsythe
5. ทดสอบความแตกต่างของจำนวนกลุ่มที่เมื่อทดสอบการเปรียบเทียบพหุคูณแล้วแตกต่างอย่างมีนัยสำคัญทางสถิติ จากการวิเคราะห์ด้วยวิธีการเปรียบเทียบพหุคูณที่ต่างกัน ในกลุ่มตัวอย่างที่มีขนาดเดียวกัน โดยวิธี Q - test
6. ทดสอบความแตกต่างของจำนวนกลุ่มที่เมื่อทดสอบการเปรียบเทียบพหุคูณแล้วแตกต่างอย่างมีนัยสำคัญทางสถิติ จากการวิเคราะห์ด้วยวิธีการเปรียบเทียบพหุคูณเดียวกันในกลุ่มตัวอย่างที่มีขนาดต่างกัน โดยวิธี χ^2 - test

สถิติที่ใช้ในการศึกษา

1. สถิติที่ใช้ในการหาคุณภาพเครื่องมือ

นำแบบทดสอบไปทดลองครั้งที่1 แล้วนำมาหาความเชื่อมั่นของแบบทดสอบ จากสูตร KR-20 ของ คูเดอร์-ริชาร์ดสัน (Kuder-Richardson Procedure) (ล้วน สายยศและอังคณา 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	สัดส่วนของผู้ทำถูกในข้อหนึ่งๆ
	q	แทน	สัดส่วนของผู้ทำผิดในข้อหนึ่งๆหรือคือ1-p
	s_t^2	แทน	คะแนนความแปรปรวนของเครื่องมือฉบับนั้น

2. สถิติทดสอบสมมติฐาน

2.1 ค่าสถิติพื้นฐาน คือ ค่าเฉลี่ย และค่าความแปรปรวน

2.2. สถิติการทดสอบการเปรียบเทียบพหุคุณรายคู่ในการวิจัยครั้งนี้มี 3 วิธีคือ

2.2.1 Dunnett's T3 test (Kirk. 1995:146)

$$T3 = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

เมื่อ	T3	แทน	การเปรียบเทียบพหุคุณของ Dunnett
	$\overline{X}_1$	แทน	ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1
	$\overline{X}_2$	แทน	ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 2
	s_1^2	แทน	ค่าความแปรปรวนของกลุ่มตัวอย่างที่ 1
	s_2^2	แทน	ค่าความแปรปรวนของกลุ่มตัวอย่างที่ 2
	n_1	แทน	จำนวนตัวอย่างของกลุ่มตัวอย่างที่ 1
	n_2	แทน	จำนวนตัวอย่างของกลุ่มตัวอย่างที่ 2

โดย
$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^2}{n_1^2(n_1 - 1)} + \frac{s_2^2}{n_2^2(n_2 - 1)}}$$

เกณฑ์การตัดสินใจจะปฏิเสธ H_0 เมื่อ $|T3| \geq m_{\alpha;c,df}$ โดยที่ $m_{\alpha;c,df}$ คือค่าวิกฤตที่เปิดตาราง Studentized maximum modulus distribution ที่มีระดับนัยสำคัญและองศาความเป็นอิสระ(c,df) ตามลำดับ

2.2.2. Games-Howell test (Kirk.1995:147)

$$GH = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right) / 2}}$$

เมื่อ	GH	แทน	การเปรียบเทียบพหุคูณของ Games และ Howell
	$\overline{X}_1$	แทน	ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1
	$\overline{X}_2$	แทน	ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 2
	s_1^2	แทน	ค่าความแปรปรวนของกลุ่มตัวอย่างที่ 1
	s_2^2	แทน	ค่าความแปรปรวนของกลุ่มตัวอย่างที่ 2
	n_1	แทน	จำนวนตัวอย่างของกลุ่มตัวอย่างที่ 1
	n_2	แทน	จำนวนตัวอย่างของกลุ่มตัวอย่างที่ 2

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^2}{n_1^2(n_1 - 1)} + \frac{s_2^2}{n_2^2(n_2 - 1)}}$$

เกณฑ์การตัดสินใจจะปฏิเสธ H_0 เมื่อ $|GH| \geq q_{\alpha,p,df}$ โดยที่ $q_{\alpha,p,df}$ คือค่าวิกฤตที่เปิดตาราง Studentized rang distribution ที่มีระดับนัยสำคัญและองศาความเป็นอิสระ(p,df) ตามลำดับ

2.2.3. Brown-Forsythe test (Kirk.1995:155)

$$BF = \frac{(\bar{X}_1 - \bar{X}_2)^2}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

เมื่อ	BF	แทน	การเปรียบเทียบพหุคูณของ Brown และ Forsythe
	$\bar{X}_1$	แทน	ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1
	$\bar{X}_2$	แทน	ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 2
	s_1^2	แทน	ค่าความแปรปรวนของกลุ่มตัวอย่างที่ 1
	s_2^2	แทน	ค่าความแปรปรวนของกลุ่มตัวอย่างที่ 2
	n_1	แทน	จำนวนตัวอย่างของกลุ่มตัวอย่างที่ 1
	n_2	แทน	จำนวนตัวอย่างของกลุ่มตัวอย่างที่ 2

$$df_1 = n_1 - 1$$

และ

$$df_2 = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{s_1^2}{n_1^2(n_1 - 1)} + \frac{s_2^2}{n_2^2(n_2 - 1)}}$$

เกณฑ์การตัดสินใจจะปฏิเสธ H_0 เมื่อ $BF \geq F_{\alpha;df_1,df_2}$ โดยที่ $F_{\alpha;df_1,df_2}$ คือค่าวิกฤตที่เปิดตาราง F ที่มีระดับนัยสำคัญและองศาความเป็นอิสระ ตามลำดับ

2.3. วิธีทดสอบความเป็นเอกพันธ์ของความแปรปรวนของ Levene (Levene, H. 1960:278-292. อ้างอิงจาก เยาวภา ไชยศรี. 2536:25)

ตัวสถิติทดสอบคำนวณได้จาก

$$L = \frac{\sum n_i (\bar{Z}_{i.} - \bar{Z}_{..})^2 / (k-1)}{\sum \sum (Z_{ij} - \bar{Z}_{i.})^2 / \sum (n_i - 1)}$$

เมื่อ

$$Z_{ij} = |X_{ij} - \bar{X}_{i.}|$$

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

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

การใช้วิธี Levene's Test จะใช้สูตรคล้ายคลึงกับสูตร sum of squares เพียงแต่เปลี่ยนค่า X_{ij} เป็น Z_{ij} ; $df = k-1, \sum (n_i - 1)$

ถ้า L คำนวณ > F จากตารางการแจกแจง F จะไม่ยอมรับ H_0

2.4 สถิติวิเคราะห์ความแปรปรวนเพื่อเปรียบเทียบค่าเฉลี่ย

(ระวีวรรณ พันธุ์พานิช. 2540:36)

โดยใช้ F - test

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

เมื่อ

F แทน การวิเคราะห์ความแปรปรวน

MS_B แทน ค่าเฉลี่ยของค่าเบี่ยงเบนกำลังสองระหว่างกลุ่ม

MS_W แทน ค่าเฉลี่ยของค่าเบี่ยงเบนกำลังสองภายในกลุ่ม

ถ้า F คำนวณ > F ตาราง จะไม่ยอมรับ H_0

2.5. สถิติทดสอบนัยสำคัญของความแตกต่างจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญทางสถิติที่ได้มาจากการเปรียบเทียบพหุคูณที่แตกต่างกัน ในกลุ่มตัวอย่างที่มีขนาดเดียวกัน โดยใช้ Q – test (บุญเชิด ภิญโญนนีตพงษ์. 2527:77)

$$Q = \frac{(k-1)(k\sum p^2 - T^2)}{kT - \sum X^2}$$

เมื่อ	Q	แทน	การทดสอบแบบ Q
	k	แทน	จำนวนการเปรียบเทียบพหุคูณ
	p	แทน	จำนวนกลุ่มที่ปฏิเสธ H_0 ของแต่ละสูตร
	T	แทน	ผลรวมของจำนวนกลุ่มที่ปฏิเสธ H_0 หรือ $T = \sum p$
	X	แทน	จำนวนสูตรที่ปฏิเสธ H_0 ในหนึ่งกลุ่มตัวอย่าง

นำค่า Q ที่คำนวณได้มาเปรียบเทียบกับค่า Qวิกฤติ ซึ่งเปิดจากตารางไค-สแควร์ ที่ $df = k-1$ ถ้าค่า Q ที่คำนวณมากกว่า Q วิกฤติ หมายความว่า วิธีการเปรียบเทียบพหุคูณที่ต่างกันให้ผลต่างกันในการใช้

2.6. สถิติทดสอบนัยสำคัญของความแตกต่างจำนวนกลุ่มที่มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญที่ได้มาจากการเปรียบเทียบพหุคูณวิธีเดียวกัน 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 ที่คำนวณได้ไปเทียบกับ χ^2 วิกฤติ ซึ่งเปิดจากตาราง ไค-สแควร์ ที่ $df = n-1$ ถ้าค่า χ^2 ที่คำนวณได้มากกว่า χ^2 วิกฤติ หมายความว่า ขนาดกลุ่มตัวอย่างที่ต่างกันให้ผลต่างกันในการใช้

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การนำเสนอผลการวิเคราะห์ข้อมูล

การศึกษาครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับดังนี้

1. ค่าสถิติพื้นฐานของคะแนนที่ได้จากการทำแบบทดสอบความถนัดด้านเหตุผล จากกลุ่มตัวอย่างขนาด 583 คน ที่เป็นประชากรเทียม

2. การทดสอบลักษณะข้อมูลเบื้องต้น คือ ความแตกต่างระหว่างค่าเฉลี่ย ซึ่งใช้ค่าเฉลี่ย และค่าความแปรปรวนที่คำนวณได้จากคะแนนความถนัด ด้านเหตุผลของนักเรียน ชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ที่เป็นประชากรเทียม โดยการสุ่มแบบใส่คืน ทั้งหมด 4 ขนาด คือ ขนาด 20 คน 40 คน 80 คน และ 120 คน

3. การทดสอบความแตกต่างของผลการเปรียบเทียบ พหุคูณจากการทดสอบพหุคูณ รายคู่ 3 วิธี จากกลุ่มตัวอย่างขนาดเดียวกัน คือ ขนาด 20 คน 40 คน 80 คน และ 120 คน

4. การทดสอบความแตกต่างของผลการเปรียบเทียบพหุคูณจากกลุ่มตัวอย่างขนาดต่างกัน 4 ขนาด ในแต่ละวิธีการทดสอบพหุคูณรายคู่ 3 วิธี คือ วิธี Dunett's T3 วิธี Games – Howell และ วิธี Brown - Forsythe

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลและการแปลความหมายของผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

X	แทน	คะแนนเฉลี่ย
S^2	แทน	ค่าความแปรปรวน
T3	แทน	วิธีการเปรียบเทียบพหุคูณของ Dunett
GH	แทน	วิธีการเปรียบเทียบพหุคูณของ Games และ Howell
BF	แทน	วิธีการเปรียบเทียบพหุคูณของ Brown และ Forsythe
Q	แทน	สถิติทดสอบแบบ Q
χ^2	แทน	สถิติทดสอบแบบไค – สแควร์
F	แทน	ค่าการวิเคราะห์ความแปรปรวน
ม.2	แทน	ชั้นมัธยมศึกษาปีที่ 2
ม.4	แทน	ชั้นมัธยมศึกษาปีที่ 4
ม.6	แทน	ชั้นมัธยมศึกษาปีที่ 6

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของคะแนนที่ได้จากแบบทดสอบวัดความถนัดด้านเหตุผล ซึ่งทดสอบจากกลุ่มตัวอย่าง ขนาด 583 คน ที่เป็นประชากรเทียม

การวิเคราะห์ครั้งนี้ ได้จากการนำคะแนนผลการทดสอบด้วยแบบทดสอบวัดความถนัดด้านเหตุผล กับกลุ่มตัวอย่าง ขนาด 583 คน ที่เป็นประชากรเทียมมาหาค่าสถิติพื้นฐาน คือ ค่าเฉลี่ย และค่าความแปรปรวน ดังปรากฏในตารางที่ 4

ตาราง 4 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านเหตุผลของนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ที่ เป็น กลุ่มตัวอย่างในการศึกษา

ชั้น	N	$\bar{X}$	S^2
ม.2	180	9.53	25.50
ม.4	205	14.64	34.10
ม.6	198	17.69	47.75
รวม	583	13.89	46.78

ผลการวิเคราะห์ตาราง 4 ปรากฏว่าค่าเฉลี่ยของคะแนนความถนัดด้านเหตุผลของกลุ่มตัวอย่างจำนวน 583 คน มีค่า 13.89 ค่าความแปรปรวน 46.78 โดยนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าเฉลี่ย 9.53 ค่าความแปรปรวน 25.50 นักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าเฉลี่ย 14.64 ค่าความแปรปรวน 34.10 และนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีค่าเฉลี่ย 17.69 ค่าความแปรปรวน 47.75

2. การทดสอบลักษณะข้อมูลเบื้องต้น

การวิเคราะห์ในตอนนี้ได้้นำคะแนนมาทดสอบความแตกต่างของค่าเฉลี่ย โดยใช้ค่าเฉลี่ย และค่าความแปรปรวนที่คำนวณได้จากคะแนนความถนัดด้านเหตุผลของนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ซึ่งสุ่มมาจากประชากรเทียมเป็นการสุ่มแบบใส่คืน ทั้งหมด 4 ขนาด ตามลำดับดังนี้

2.1 การทดสอบความแตกต่างของค่าเฉลี่ย ของกลุ่มตัวอย่างที่สุ่มมาจากประชากร
เทียม โดยมีขนาด 20 คน ซึ่งสุ่มมา ชั้นละ 10 กลุ่ม ปรากฏดังตาราง 5

ตาราง 5 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านเหตุผลของนักเรียนชั้นมัธยมศึกษาปีที่
2, 4 และ 6 ค่าการทดสอบความแตกต่างของค่าเฉลี่ย ที่เป็นกลุ่มตัวอย่างขนาด 20 คน
ชั้นละ 10 กลุ่ม

กลุ่ม ตัวอย่าง ที่	$\bar{X}$			S^2			F
	ม.2	ม.4	ม.6	ม.2	ม.4	ม.6	
1	9.35	15.15	17.80	3.50	48.97	45.74	11.40**
2	9.95	14.10	17.45	8.26	27.88	36.99	11.58**
3	9.50	15.30	17.50	14.47	29.48	43.63	11.70**
4	10.55	14.20	17.20	15.73	41.22	6.58	10.47**
5	10.45	15.95	17.00	28.26	48.89	11.68	8.34**
6	11.00	15.85	17.00	24.00	47.81	12.21	7.24**
7	11.15	14.65	17.95	23.81	11.50	31.94	10.31**
8	11.10	14.70	17.90	25.56	7.69	41.57	9.28**
9	9.80	14.50	17.40	10.58	11.94	32.35	16.07**
10	11.45	13.50	17.10	29.41	7.52	44.93	5.99**

P**<.01

ผลการวิเคราะห์ตามตาราง 5 ปรากฏว่า ค่าเฉลี่ยของคะแนนความถนัดด้านเหตุผล
ของกลุ่มตัวอย่างขนาด 20 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 9.35 ถึง 11.45
ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าตั้งแต่ 13.50 ถึง 15.95 และของนักเรียนชั้นมัธยมศึกษา
ปีที่ 6 มีค่าตั้งแต่ 17.00 ถึง 17.95 ค่าความแปรปรวนของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มี
ค่าตั้งแต่ 3.50 ถึง 29.41 ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าตั้งแต่ 7.52 ถึง 48.99
และของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีค่าตั้งแต่ 11.68 ถึง 45.74 จากการวิเคราะห์ความ
แปรปรวน พบว่า กลุ่มตัวอย่างทุกกลุ่มมีค่าเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่
ระดับ .01

2.2 การทดสอบความแตกต่างของค่าเฉลี่ย ของกลุ่มตัวอย่างที่สุ่มมาจากประชากร
เทียม โดยมีขนาด 40 คน ซึ่งสุ่มมา ชั้นละ 10 กลุ่ม ปรากฏดังตาราง 6

ตาราง 6 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านเหตุผลของนักเรียนชั้นมัธยมศึกษาปีที่
2, 4 และ 6 ค่าการทดสอบความแตกต่างของค่าเฉลี่ย ที่เป็นกลุ่มตัวอย่างขนาด 40 คน
ชั้นละ 10 กลุ่ม

กลุ่ม ตัวอย่าง ที่	$\bar{X}$			S^2			F
	ม.2	ม.4	ม.6	ม.2	ม.4	ม.6	
1	10.75	14.45	16.85	11.89	39.84	57.92	12.51**
2	10.00	15.07	17.50	5.94	28.17	45.53	22.06**
3	10.75	14.15	18.22	6.40	30.64	32.69	27.87**
4	11.25	15.50	17.52	29.08	52.35	6.71	14.65**
5	10.85	14.92	17.40	18.28	36.63	10.86	19.95**
6	10.62	14.45	16.77	16.95	44.45	9.51	17.04**
7	10.57	13.95	17.55	25.68	8.40	33.53	21.58**
8	10.65	13.97	17.67	28.90	3.76	36.53	21.41**
9	11.42	14.15	17.47	37.43	6.07	34.87	14.05**
10	10.37	14.00	17.67	26.18	6.35	32.22	24.68**

P**<.01

ผลการวิเคราะห์ตามตาราง 6 ปรากฏว่า ค่าเฉลี่ยของคะแนนความถนัดด้านเหตุผล
ของกลุ่มตัวอย่างขนาด 40 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 10.00 ถึง
11.42 ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าตั้งแต่ 13.95 ถึง 15.50 และของนักเรียน
ชั้นมัธยมศึกษาปีที่ 6 มีค่า ตั้งแต่ 16.77 ถึง 18.22 ค่าความแปรปรวนของนักเรียนชั้นมัธยม
ศึกษาปีที่ 2 มีค่าตั้งแต่ 5.94 ถึง 37.43 ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าตั้งแต่ 3.76
ถึง 52.35 และของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีค่าตั้งแต่ 6.71 ถึง 57.92 จากการ
วิเคราะห์ความแปรปรวน พบว่า กลุ่มตัวอย่างทุกกลุ่มมีค่าเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญ
ทางสถิติที่ระดับ .01

2.3 การทดสอบความแตกต่างของค่าเฉลี่ย ของกลุ่มตัวอย่างที่สุ่มมาจากประชากร
เทียม โดยมีขนาด 80 คน ซึ่งสุ่มมา ชั้นละ 10 กลุ่ม ปรากฏดังตาราง 7

ตาราง 7 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านเหตุผลของนักเรียนชั้นมัธยมศึกษาปีที่
2, 4 และ 6 ค่าการทดสอบความแตกต่างของค่าเฉลี่ย ที่เป็นกลุ่มตัวอย่างขนาด 80 คน
ชั้นละ 10 กลุ่ม

กลุ่ม ตัวอย่าง ที่	$\bar{X}$			S^2			F
	ม.2	ม.4	ม.6	ม.2	ม.4	ม.6	
1	9.83	14.41	17.81	5.25	30.49	38.15	52.00**
2	10.12	14.76	18.61	6.97	28.33	39.32	52.34**
3	9.78	14.25	17.71	5.13	28.79	40.10	51.16**
4	11.05	14.62	17.66	25.08	38.38	6.63	37.50**
5	10.51	14.81	17.01	24.68	38.33	9.96	35.94**
6	11.42	14.48	17.58	29.41	39.06	8.57	29.57**
7	10.90	13.97	17.81	31.71	9.89	36.43	36.88**
8	10.83	13.90	17.85	28.23	7.63	29.95	45.06**
9	10.96	14.33	17.76	28.89	11.23	32.48	38.20**
10	11.06	13.92	17.73	29.57	6.98	29.53	40.71**

P**<.01

ผลการวิเคราะห์ตามตาราง 7 ปรากฏว่า ค่าเฉลี่ยของคะแนนความถนัดด้านเหตุผล
ของกลุ่มตัวอย่างขนาด 80 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 9.78 ถึง 11.42
ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าตั้งแต่ 13.90 ถึง 14.81 และของนักเรียนชั้นมัธยมศึกษา
ปีที่ 6 มีค่าตั้งแต่ 17.01 ถึง 18.16 ค่าความแปรปรวนของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มี
ค่าตั้งแต่ 5.13 ถึง 37.71 ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าตั้งแต่ 6.98 ถึง 39.06
และของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีค่าตั้งแต่ 6.63 ถึง 40.10 จากการวิเคราะห์ความ
แปรปรวน พบว่า กลุ่มตัวอย่างทุกกลุ่มมีค่าเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่
ระดับ .01

2.4 การทดสอบความแตกต่างของค่าเฉลี่ย และการทดสอบความแตกต่างของความแปรปรวนของกลุ่มตัวอย่างที่สุ่มมาจากประชากรเทียม โดยมีขนาด 120 คน ซึ่งสุ่มมาชั้นละ 10 กลุ่ม ปรากฏดังตาราง 8

ตาราง 8 ค่าสถิติพื้นฐานของคะแนนความถนัดด้านเหตุผลของนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ค่าการทดสอบความแตกต่างของค่าเฉลี่ย และค่าการทดสอบความแตกต่างของความแปรปรวน ที่เป็น กลุ่มตัวอย่างขนาด 120 คน ชั้นละ 10 กลุ่ม

กลุ่ม ตัวอย่าง ที่	$\bar{X}$			S^2			F
	ม.2	ม.4	ม.6	ม.2	ม.4	ม.6	
1	9.77	14.70	17.54	9.65	27.15	38.62	80.07**
2	9.69	14.93	17.81	3.74	31.57	41.00	80.02**
3	9.95	14.05	17.86	4.78	35.37	39.91	70.48**
4	10.39	14.48	17.34	19.33	37.15	9.42	66.60**
5	10.33	14.55	17.34	28.82	30.87	8.93	55.91**
6	10.96	14.60	17.50	26.15	39.19	6.62	53.71**
7	11.14	14.36	17.45	33.93	10.30	34.55	45.58**
8	11.03	14.35	17.64	35.12	8.24	31.15	52.73**
9	10.78	14.28	17.44	26.82	7.93	36.29	56.19**
10	11.05	14.45	17.28	29.93	11.79	44.59	40.50**

P**<.01

ผลการวิเคราะห์ตามตาราง 8 ปรากฏว่า ค่าเฉลี่ยของคะแนนความถนัดด้านเหตุผลของกลุ่มตัวอย่างขนาด 120 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 9.69 ถึง 11.14 ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าตั้งแต่ 14.08 ถึง 14.93 และของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีค่าตั้งแต่ 17.28 ถึง 17.86 ค่าความแปรปรวนของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีค่าตั้งแต่ 3.74 ถึง 35.12 ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 มีค่าตั้งแต่ 7.93 ถึง 39.19 และของนักเรียนชั้นมัธยมศึกษาปีที่ 6 มีค่าตั้งแต่ 6.62 ถึง 44.59 จากการวิเคราะห์ความแปรปรวน พบว่า กลุ่มตัวอย่างทุกกลุ่มมีค่าเฉลี่ยที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. การทดสอบความแตกต่างของผลการเปรียบเทียบพหุคูณ จากกลุ่มตัวอย่างขนาดเดียวกัน

การวิเคราะห์ครั้งนี้ได้นำจำนวนผลการเปรียบเทียบพหุคูณ จากการทดสอบพหุคูณรายคู่ 3 วิธี คือ วิธี Dunnett's T3 วิธี Games – Howell และวิธี Brown – Forsythe ในแต่ละกลุ่มตัวอย่าง 4 ขนาด คือ ขนาด 20 คน 40 คน 80 คน และ 120 คน มาทดสอบความแตกต่างจำนวนผลการเปรียบเทียบพหุคูณ ซึ่งวิเคราะห์เปรียบเทียบระหว่างกลุ่มตัวอย่างเดียวกันโดยใช้ Q – Test ตามลำดับดังนี้

3.1 การทดสอบกับกลุ่มตัวอย่างขนาด 20 คน ปรากฏผลดังตาราง 9

ตาราง 9 การเปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบพหุคูณรายคู่ต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 20 คน

จำนวนการยอมรับและ ไม่ยอมรับสมมติฐาน	T3	GH	BF	Q
ยอมรับสมมติฐาน	18	18	24	6
ไม่ยอมรับสมมติฐาน	12	12	6	

ผลการวิเคราะห์ตาราง 9 ปรากฏว่าการเปรียบเทียบพหุคูณจากกลุ่มตัวอย่างขนาด 20 คน ที่ทดสอบด้วยวิธี Dunnett's T3 ยอมรับสมมติฐาน 18 กลุ่ม ไม่ยอมรับสมมติฐาน 12 กลุ่ม การเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Games – Howell ยอมรับสมมติฐาน 18 กลุ่ม ไม่ยอมรับสมมติฐาน 12 กลุ่ม และการเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Brown – Forsythe ยอมรับสมมติฐาน 24 กลุ่ม ไม่ยอมรับสมมติฐาน 6 กลุ่ม เมื่อทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐาน ด้วยสูตร Q – test พบว่ามีจำนวนแตกต่างอย่างไม่มีนัยสำคัญทางสถิติที่แสดงว่า ผลการเปรียบเทียบพหุคูณที่ได้จากการทดสอบพหุคูณ 3 วิธี ให้ผลการยอมรับและไม่ยอมรับสมมติฐานไม่แตกต่างกัน

3.2 การทดสอบกับกลุ่มตัวอย่างขนาด 40 คน ปรากฏผลดังตาราง 10

ตาราง 10 การเปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบ
พหุคูณรายคู่ต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 40 คน

จำนวนการยอมรับและ ไม่ยอมรับสมมติฐาน	T3	GH	BF	Q
ยอมรับสมมติฐาน	6	6	12	5.538
ไม่ยอมรับสมมติฐาน	24	24	18	

ผลการวิเคราะห์ตาราง 10 ปรากฏว่าการเปรียบเทียบพหุคูณจากกลุ่มตัวอย่างขนาด 40 คน ที่ทดสอบด้วยวิธี Dunnett's T3 ยอมรับสมมติฐาน 6 กลุ่ม ไม่ยอมรับสมมติฐาน 24 กลุ่ม การเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Games – Howell ยอมรับสมมติฐาน 6 กลุ่ม ไม่ยอมรับสมมติฐาน 24 กลุ่ม และการเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Brown – Forsythe ยอมรับสมมติฐาน 12 กลุ่ม ไม่ยอมรับสมมติฐาน 18 กลุ่ม เมื่อทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐาน ด้วยสูตร Q – test พบว่ามีจำนวนแตกต่างอย่างไม่มีนัยสำคัญทางสถิติแสดงว่า ผลการเปรียบเทียบพหุคูณที่ได้จากการทดสอบพหุคูณ 3 วิธี ให้ผลการยอมรับและไม่ยอมรับสมมติฐานไม่แตกต่างกัน

3.3 การทดสอบกับกลุ่มตัวอย่างขนาด 80 คน ปรากฏผลดังตาราง 11

ตาราง 11 การเปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบ
พหุคูณรายคู่ต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 80 คน

จำนวนการยอมรับและ ไม่ยอมรับสมมติฐาน	T3	GH	BF	Q
ยอมรับสมมติฐาน	1	1	7	9.00
ไม่ยอมรับสมมติฐาน	29	29	23	

ผลการวิเคราะห์ตาราง 11 ปรากฏว่าการเปรียบเทียบพหุคูณจากกลุ่มตัวอย่างขนาด 80 คน ที่ทดสอบด้วยวิธี Dunnett's T3 ยอมรับสมมติฐาน 1 กลุ่ม ไม่ยอมรับสมมติฐาน 29 กลุ่ม การเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Games – Howell ยอมรับสมมติฐาน 1 กลุ่ม ไม่ยอมรับสมมติฐาน 29 กลุ่ม และการเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Brown – Forsythe ยอมรับสมมติฐาน 7 กลุ่ม ไม่ยอมรับสมมติฐาน 23 กลุ่ม เมื่อทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐาน ด้วยสูตร Q – test พบว่ามีจำนวนแตกต่างอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า ผลการเปรียบเทียบพหุคูณที่ได้จากการทดสอบพหุคูณ 3 วิธี ให้ผลการยอมรับและไม่ยอมรับสมมติฐานไม่แตกต่างกัน

3.4 การทดสอบกับกลุ่มตัวอย่างขนาด 120 คน ปรากฏผลดังตาราง 12

ตาราง 12 การเปรียบเทียบจำนวนผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบพหุคูณรายคู่ต่างกัน 3 วิธี จากกลุ่มตัวอย่าง 120 คน

จำนวนการยอมรับและ ไม่ยอมรับสมมติฐาน	T3	GH	BF	Q
ยอมรับสมมติฐาน	0	0	6	12.00**
ไม่ยอมรับสมมติฐาน	30	30	24	

P** < .01

ผลการวิเคราะห์ตาราง 12 ปรากฏว่าการเปรียบเทียบพหุคูณจากกลุ่มตัวอย่างขนาด 120 คนที่ทดสอบด้วยวิธี Dunnett's T3 ยอมรับสมมติฐาน 0 กลุ่ม ไม่ยอมรับสมมติฐาน 30 กลุ่ม การเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Games – Howell ยอมรับสมมติฐาน 0 กลุ่ม ไม่ยอมรับสมมติฐาน 30 กลุ่ม และการเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Brown – Forsythe ยอมรับสมมติฐาน 6 กลุ่ม ไม่ยอมรับสมมติฐาน 24 กลุ่ม เมื่อทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐาน ด้วยสูตร Q – test พบว่ามีจำนวนแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการเปรียบเทียบพหุคูณที่ได้จากการทดสอบพหุคูณ 3 วิธี ให้ผลการยอมรับและไม่ยอมรับสมมติฐานแตกต่างกัน เพื่อให้ทราบผลการเปรียบเทียบพหุคูณคู่ใดต่างกันบ้าง จึงนำมาเปรียบเทียบด้วยสูตร Q – test เป็นรายคู่ และเพื่อเป็นการควบคุมอัตราความคลาดเคลื่อนให้เท่ากับระดับนัยสำคัญที่ .01 ผู้วิจัยกำหนดระดับนัยสำคัญทางสถิติที่ระดับ .003 ดังแสดงในตาราง 13

ตาราง 13 การเปรียบเทียบผลการเปรียบเทียบพหุคูณ จากการทดสอบพหุคูณด้วยวิธี 3 วิธี
ที่ละคู่วิธี จากกลุ่มตัวอย่างขนาด 120 คน

วิธีการเปรียบเทียบพหุคูณ	T3	GH	BF
T3	-	0.00	6.00*
GH		-	6.00*
BF			-

P* < .003

ผลการวิเคราะห์ตาราง 13 ปรากฏว่า การเปรียบเทียบพหุคูณที่ทดสอบความแตกต่างของจำนวนผลการเปรียบเทียบพหุคูณด้วยสูตร Q – test นั้น ในขนาดกลุ่มตัวอย่างขนาด 120 คน วิธี Dunnett's T3 กับวิธี Games- Howell แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ วิธี Dunnett's T3 กับวิธี Brown-Forsythe แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .003 และวิธี Games- Howell กับวิธี Brown-Forsythe แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .003

4. การทดสอบความแตกต่างของผลการเปรียบเทียบพหุคูณจากกลุ่มตัวอย่างขนาดต่างกันการวิเคราะห์ครั้งนี้ได้นำจำนวนผลการเปรียบเทียบพหุคูณจากกลุ่มตัวอย่าง 4 ขนาด คือ 20 คน 40 คน 80 คน และ 120 คน ในแต่ละวิธีการทดสอบพหุคูณ 3 วิธี คือ วิธี Dunnett's T3 วิธี Brown – Forsythe และวิธี Games – Howell มาทดสอบความแตกต่างของจำนวนผลการเปรียบเทียบพหุคูณ ซึ่งวิเคราะห์เปรียบเทียบระหว่างกลุ่ม โดยใช้ χ^2 – test ตามลำดับดังนี้

4.1 การทดสอบพหุคูณด้วยวิธี Dunnett's T3 ปรากฏผลดังแสดงในตาราง 14

ตาราง 14 การเปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบพหุคูณด้วยวิธี Dunnett's T3 จากกลุ่มตัวอย่าง 4 ขนาด

จำนวนการยอมรับและ ไม่ยอมรับสมมติฐาน	20	40	80	120	χ^2
ยอมรับสมมติฐาน	18	6	1	0	28.67**
ไม่ยอมรับสมมติฐาน	12	24	29	30	

$P^{**} < .01$

ผลการวิเคราะห์ตาราง 14 ปรากฏว่า การเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Dunnett's T3 จากกลุ่มตัวอย่างขนาด 20 คน ยอมรับสมมติฐาน 18 กลุ่ม ไม่ยอมรับสมมติฐาน 12 กลุ่ม จากกลุ่มตัวอย่างขนาด 40 คน ยอมรับสมมติฐาน 6 กลุ่ม ไม่ยอมรับสมมติฐาน 24 กลุ่ม จากกลุ่มตัวอย่างขนาด 80 คน ยอมรับสมมติฐาน 1 กลุ่ม ไม่ยอมรับสมมติฐาน 29 กลุ่ม และจากกลุ่มตัวอย่างขนาด 120 คน ยอมรับสมมติฐาน 0 กลุ่ม ไม่ยอมรับสมมติฐาน 30 กลุ่ม เมื่อทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานด้วย χ^2 - test พบว่า มีจำนวนแตกต่างอย่างมีนัยสำคัญ ทางสถิติที่ระดับ .01 แสดงว่าจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานที่ได้จากการทดสอบพหุคูณรายคู่จากกลุ่มตัวอย่าง 4 ขนาดด้วยวิธี Dunnett's T3 ให้ผลการยอมรับและไม่ยอมรับสมมติฐานแตกต่างกัน เพื่อให้ทราบว่าการเปรียบเทียบพหุคูณของขนาดกลุ่มตัวอย่างคู่ใดบ้างที่ต่างกัน จึงนำมาเปรียบเทียบด้วย χ^2 - test เป็นรายคู่เพื่อเป็นการควบคุมอัตราความคลาดเคลื่อนให้เท่ากับระดับนัยสำคัญทางสถิติที่ระดับ .01 ผู้วิจัยจึงกำหนดระดับนัยสำคัญเท่ากับ .0025 ดังแสดงในตาราง 15

ตาราง 15 การเปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบด้วยวิธี Dunnett's T3 จากกลุ่มตัวอย่างที่ละคู่ขนาด ทั้งหมด 4 ขนาด

ขนาดกลุ่มตัวอย่าง	20	40	80	120
20	-	6.00	13.76*	15.60*
40		-	18.46*	20.40*
80			-	28.06*
120				-

$P^* < .0025$

ผลการวิเคราะห์ตาราง 15 ปรากฏว่าการเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Dunnett's T3 ซึ่งทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานด้วย χ^2 - test นั้น กลุ่มตัวอย่างขนาด 20 คน กับ 40 คน แตกต่างอย่างไม่มีนัยสำคัญทางสถิติ กลุ่มตัวอย่างขนาด 20 คน กับ 80 คน ขนาด 20 คน กับ 120 คน ขนาด 40 คน กับ 80 คน ขนาด 40 คน กับ 120 คน และขนาด 80 คน กับ 120 คน แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025

4.2 การทดสอบพหุคูณด้วยวิธี Games – Howell ปรากฏผลดังแสดงในตาราง 16

ตาราง 16 การเปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบพหุคูณด้วย วิธี Games – Howell จากกลุ่มตัวอย่าง 4 ขนาด

จำนวนการยอมรับและ ไม่ยอมรับสมมติฐาน	20	40	80	120	χ^2
ยอมรับสมมติฐาน	18	6	1	0	28.67**
ไม่ยอมรับสมมติฐาน	12	24	29	30	

$P^{**} < .01$

ผลการวิเคราะห์ตาราง 16 ปรากฏว่า การเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Games – Howell จากกลุ่มตัวอย่างขนาด 20 คน ยอมรับสมมติฐาน 18 กลุ่ม ไม่ยอมรับสมมติฐาน 12 กลุ่ม จากกลุ่มตัวอย่างขนาด 40 คน ยอมรับสมมติฐาน 6 กลุ่ม ไม่ยอมรับสมมติฐาน 24 กลุ่ม จากกลุ่มตัวอย่างขนาด 80 คน ยอมรับสมมติฐาน 1 กลุ่ม ไม่ยอมรับสมมติฐาน 29 กลุ่ม และจากกลุ่มตัวอย่างขนาด 120 คน ยอมรับสมมติฐาน 0 กลุ่ม ไม่ยอมรับสมมติฐาน 30 กลุ่ม เมื่อทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานด้วย χ^2 - test พบว่า มีจำนวนแตกต่างอย่างมีนัยสำคัญ ทางสถิติที่ระดับ .01 แสดงว่าจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานที่ได้จากการทดสอบพหุคูณรายคู่จาก

กลุ่มตัวอย่าง 4 ขนาดด้วยวิธี Games – Howell ให้ผลการยอมรับและไม่ยอมรับสมมติฐานแตกต่างกัน เพื่อให้ทราบว่าการเปรียบเทียบพหุคูณของขนาดกลุ่มตัวอย่างคู่ใดบ้างที่ต่างกัน จึงนำมาเปรียบเทียบด้วย χ^2 - test เป็นรายคู่เพื่อเป็นการควบคุมอัตราความคลาดเคลื่อนให้เท่ากับระดับนัยสำคัญทางสถิติที่ χ^2 ระดับ .01 ผู้วิจัยจึงกำหนดระดับนัยสำคัญเท่ากับ .0025 ดังแสดงในตาราง 17

ตาราง 17 การเปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบด้วยวิธี Games-Howell จากกลุ่มตัวอย่างที่ละคู่ขนาด ทั้งหมด 4 ขนาด

ขนาดกลุ่มตัวอย่าง	20	40	80	120
20	-	6.00	13.76*	15.60*
40		-	18.46*	20.40*
80			-	28.06*
120				-

P* < .0025

ผลการวิเคราะห์ตาราง 17 ปรากฏว่าการเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Games-Howell ซึ่งทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานด้วย χ^2 - test นั้น กลุ่มตัวอย่างขนาด 20 คน กับ 40 คน แตกต่างอย่างไม่มีนัยสำคัญทางสถิติ กลุ่มตัวอย่างขนาด 20 คน กับ 80 คน ขนาด 20 คน กับ 120 คน ขนาด 40 คน กับ 80 คน ขนาด 40 คน กับ 120 คน และขนาด 80 คน กับ 120 คน แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025

4.3 การทดสอบพหุคูณด้วยวิธี Brown-Forsythe ปรากฏผลดังแสดงในตาราง 18

ตาราง 18 การเปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบพหุคูณด้วยวิธี Brown-Forsythe จากกลุ่มตัวอย่าง 4 ขนาด

จำนวนการยอมรับและ ไม่ยอมรับสมมติฐาน	20	40	80	120	χ^2
ยอมรับสมมติฐาน	24	12	7	6	15.66**
ไม่ยอมรับสมมติฐาน	6	18	23	24	

$P^{**} < .01$

ผลการวิเคราะห์ตาราง 18 ปรากฏว่า การเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Brown-Forsythe จากกลุ่มตัวอย่างขนาด 20 คน ยอมรับสมมติฐาน 24 กลุ่ม ไม่ยอมรับสมมติฐาน 6 กลุ่ม จากกลุ่มตัวอย่างขนาด 40 คน ยอมรับสมมติฐาน 12 กลุ่ม ไม่ยอมรับสมมติฐาน 18 กลุ่ม จากกลุ่มตัวอย่างขนาด 80 คน ยอมรับสมมติฐาน 7 กลุ่ม ไม่ยอมรับสมมติฐาน 23 กลุ่ม และจากกลุ่มตัวอย่างขนาด 120 คน ยอมรับสมมติฐาน 6 กลุ่ม ไม่ยอมรับสมมติฐาน 24 กลุ่ม เมื่อทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานด้วย χ^2 - test พบว่า มีจำนวนแตกต่างอย่างมีนัยสำคัญ ทางสถิติที่ระดับ .01 แสดงว่าจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานที่ได้จากการทดสอบพหุคูณรายคู่จากกลุ่มตัวอย่าง 4 ขนาดด้วยวิธี Brown-Forsythe ให้ผลการยอมรับและไม่ยอมรับสมมติฐานแตกต่างกัน เพื่อให้ทราบว่าการเปรียบเทียบพหุคูณของขนาดกลุ่มตัวอย่างคู่ใดบ้างที่ต่างกัน จึงนำมาเปรียบเทียบด้วย χ^2 - test เป็นรายคู่เพื่อเป็นการควบคุมอัตราความคลาดเคลื่อนให้เท่ากับระดับนัยสำคัญทางสถิติที่ระดับ .01 ผู้วิจัยจึงกำหนดระดับนัยสำคัญเท่ากับ .0025 ดังแสดงในตาราง 19

ตาราง 19 การเปรียบเทียบผลการยอมรับและไม่ยอมรับสมมติฐาน จากการทดสอบพหุคูณด้วยวิธี Brown-Forsythe จากกลุ่มตัวอย่างที่ละคู่ขนาด ทั้งหมด 4 ขนาด

ขนาดกลุ่มตัวอย่าง	20	40	80	120
20	-	6.00	9.16*	10.80*
40		-	4.86	6.00
80			-	9.66*
320				-

P* < .0025

ผลการวิเคราะห์ตาราง 19 ปรากฏว่าการเปรียบเทียบพหุคูณที่ทดสอบด้วยวิธี Brown-Forsythe ซึ่งทดสอบความแตกต่างของจำนวนผลการยอมรับและไม่ยอมรับสมมติฐานด้วย χ^2 - test นั้น กลุ่มตัวอย่างขนาด 20 คน กับ 40 คน ขนาด 40 คน กับ 80 คน และ ขนาด 40 คน กับ 120 คน แตกต่างอย่างไม่มีนัยสำคัญทางสถิติที่ และ กลุ่มตัวอย่างขนาด 20 คน กับ 80 คน ขนาด 20 คน กับ 120 คน และ ขนาด 80 คน กับ 120 คน แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025

ตาราง 20 ผลการเปรียบเทียบพหุคูณที่ไม่ยอมรับสมมติฐาน ที่ทำการทดสอบพหุคูณรายคู่ด้วย 3 วิธี จากกลุ่มตัวอย่าง 4 ขนาด

ขนาดกลุ่มตัวอย่าง (คน)	ผลการเปรียบเทียบพหุคูณที่ไม่ยอมรับสมมติฐาน(กลุ่ม)		
	T3	GH	BF
20	12	12	6
40	24	24	18
80	29	29	23
120	30	30	24

ผลการวิเคราะห์ตาราง 20 ปรากฏว่า ที่ขนาดกลุ่มตัวอย่าง 20 คน วิธี Dunnett's T3 และ วิธี Games-Howell มีผลการไม่ยอมรับสมมติฐาน 12 กลุ่ม วิธี Brown-Forsythe มีผลการไม่ยอมรับสมมติฐาน 6 กลุ่ม รวม 30 กลุ่ม จากการเปรียบเทียบพหุคูณทั้ง 3 วิธี ด้วยขนาดกลุ่มตัวอย่าง 20 คน ทั้งหมด 90 กลุ่ม ที่ขนาดกลุ่มตัวอย่าง 40 คน วิธี Dunnett's T3 และวิธี Games-Howell มีผลการไม่ยอมรับสมมติฐาน 24 กลุ่ม วิธี Brown-Forsythe มีผลการไม่ยอมรับสมมติฐาน 18 กลุ่ม รวม 30 กลุ่ม จากการเปรียบเทียบพหุคูณทั้ง 3 วิธี ด้วยขนาดกลุ่มตัวอย่าง 40 คน ทั้งหมด 90 กลุ่ม ที่ขนาดกลุ่มตัวอย่าง 80 คน วิธี Dunnett's T3 และวิธี Games-Howell มีผลการไม่ยอมรับสมมติฐาน 29 กลุ่ม วิธี Brown-Forsythe มีผลการไม่ยอมรับสมมติฐาน 23 กลุ่ม รวม 30 ครั้ง จากการเปรียบเทียบพหุคูณทั้ง 3 วิธี ด้วยขนาดกลุ่มตัวอย่าง 80 คน ทั้งหมด 90 กลุ่ม ที่ขนาดกลุ่มตัวอย่าง 120 คน วิธี Dunnett's T3 และวิธี Games-Howell มีผลการไม่ยอมรับสมมติฐาน 30 กลุ่ม วิธี Brown-Forsythe มีผลการไม่ยอมรับสมมติฐาน 24 กลุ่ม รวม 30 กลุ่ม จากการเปรียบเทียบพหุคูณทั้ง 3 วิธี ด้วยขนาดกลุ่มตัวอย่าง 120 คน ทั้งหมด 90 กลุ่ม ส่วนวิธี Dunnett's T3 และวิธี Games-Howell ให้ผลการไม่ยอมรับสมมติฐาน 95 กลุ่ม จากทุกขนาดกลุ่มตัวอย่าง และวิธี Brown-Forsythe ให้ผลการไม่ยอมรับสมมติฐาน 71 กลุ่มจากทุกขนาดกลุ่มตัวอย่าง เมื่อเปรียบเทียบแต่ละวิธีในทุกขนาดกลุ่มตัวอย่างมีจำนวน 120 กลุ่ม โดย วิธี Dunnett's T3 และวิธี Games-Howell ให้ผลการไม่ยอมรับสมมติฐานเท่ากัน และวิธี Brown-Forsythe ให้ผลการไม่ยอมรับสมมติฐานน้อยที่สุด

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษา

1. เพื่อเปรียบเทียบจำนวนผลการเปรียบเทียบพหุคูณ ที่ได้จากการทดสอบการเปรียบเทียบพหุคูณรายคู่ ด้วยวิธีต่างกัน 3 วิธี คือ 1) วิธี Dunnett's T3 2) วิธี Games – Howell และ 3) วิธี Brown – Forsythe โดยศึกษาแต่ละขนาดกลุ่มตัวอย่าง คือ ขนาด 20 คน 40 คน 80 คน และ 120 คน

2. เพื่อเปรียบเทียบจำนวนผลการเปรียบเทียบพหุคูณ ที่ได้จากการทดสอบกลุ่มตัวอย่าง 4 ขนาด คือ ขนาด 20 คน 40 คน 80 และ 120 คน ที่ทดสอบพหุคูณรายคู่แต่ละวิธี คือ 1) Dunnett's T3 2) วิธี Games – Howell และ 3) วิธี Brown – Forsythe

สมมติฐานของการศึกษา

1. การทดสอบพหุคูณรายคู่ด้วยวิธีที่แตกต่างกันจากกลุ่มตัวอย่างเดียวกันให้ผลการยอมรับและไม่ยอมรับสมมติฐานแตกต่างกัน
2. การทดสอบพหุคูณรายคู่จากกลุ่มตัวอย่างขนาดต่างกัน ด้วยวิธีการทดสอบเดียวกันให้ผลการยอมรับและไม่ยอมรับสมมติฐานแตกต่างกัน

กลุ่มตัวอย่างที่ใช้ในการศึกษา

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ภาคเรียนที่ 2 ปีการศึกษา 2543 จำนวน 583 คน ซึ่งได้มาโดยวิธีการสุ่มแบบสองขั้นตอน (Two – Stage Random Sampling)

เครื่องมือที่ใช้ในการศึกษา

เครื่องมือที่ใช้ในการศึกษาค้นคว้าเป็นแบบทดสอบความถนัดด้านเหตุผลแบบสรุปความตามแนวของเธอร์สโตน รวม 30 ข้อ เป็นแบบทดสอบปรนัยชนิดเลือกตอบ 5 ตัวเลือก ซึ่งผู้วิจัยนำมาปรับปรุงจากแบบทดสอบความถนัดด้านเหตุผล แบบสรุปความ ของ พิกุล เกตุประดิษฐ์ และ วิภา ภัทรมัย มีความเชื่อมั่น 0.634

วิธีเก็บรวบรวมข้อมูล

1. ขอหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูล
2. ติดต่อโรงเรียนที่เป็นกลุ่มตัวอย่างในการศึกษา ขออนุญาตผู้บริหารโรงเรียน เพื่อกำหนด วัน เวลา ที่ทำการทดสอบ
3. วางแผนในการดำเนินการสอบโดยผู้วิจัยเป็นผู้ดำเนินการสอบด้วยตนเอง จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่ทำการทดสอบในแต่ละครั้ง
4. นำแบบทดสอบไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างทั้งหมด ตามวัน เวลาที่นัดหมายไว้
5. ชี้แจงให้นักเรียนที่เป็นกลุ่มตัวอย่างเข้าใจ และทราบถึงวัตถุประสงค์ในการทำแบบทดสอบครั้งนี้พร้อมทั้งขอความร่วมมือในการทำแบบทดสอบด้วย
6. อธิบายวิธีการทำแบบทดสอบวัดความถนัดด้านเหตุผล ซึ่งมีทั้งหมด 30 ข้อ เป็นข้อสอบปรนัยแบบเลือกตอบ ให้นักเรียนพิจารณาข้อความที่กำหนดให้ในการหาข้อสรุปอย่างสมเหตุสมผลแล้วเลือกข้อสรุปว่าควรเลือกจากข้อใดจาก ก. ถึง จ.
7. หลังจากได้ดำเนินการสอบกลุ่มตัวอย่างครบทุกโรงเรียนที่ทำการสุ่มมาแล้วนั้น นำกระดาษคำตอบมาตรวจให้คะแนน ดังนี้ ถ้าตอบถูกให้ 1 คะแนน ถ้าตอบผิดหรือไม่ตอบหรือตอบมากกว่าหนึ่งข้อให้ 0 คะแนน แล้วนำผลมาบันทึกข้อมูลและวิเคราะห์ทางสถิติ

การวิเคราะห์ข้อมูล

1. นำข้อมูลที่ได้มาจากการทดสอบของกลุ่มตัวอย่างขนาด 583 คน มาบันทึกเป็นข้อมูลของประชากรเทียม
2. รวบรวมข้อมูลจากกลุ่มตัวอย่างให้มีขนาด 20 คน 40 คน 80 คน และ 120 คน ซึ่งแต่ละขนาดจะสุ่มมา 30 กลุ่ม โดยทำการสุ่มข้อมูลแต่ละกลุ่มแต่ละขนาดด้วยการสุ่มแบบใส่คืน
3. นำข้อมูลทุกกลุ่มทุกขนาดมาทดสอบความเป็นเอกพันธ์ของความแปรปรวนของกลุ่มตัวอย่างด้วยวิธีของ Levene ถ้ามีกลุ่มตัวอย่างใดที่เป็นเอกพันธ์ของความแปรปรวน ให้ทำการสุ่มแบบใส่คืนใหม่จนครบตามจำนวนกลุ่มที่ต้องการ

4. นำข้อมูลทุกกลุ่มทุกขนาดที่ไม่เป็นเอกพันธ์ของความแปรปรวนมาวิเคราะห์ความแปรปรวน แล้วจึงทดสอบการเปรียบเทียบพหุคูณทั้ง 3 วิธี คือ วิธี Dunnett's T3 วิธี Games-Howell และ วิธี Brown-Forsythe

5. ทดสอบความแตกต่างของจำนวนกลุ่มที่เมื่อทดสอบการเปรียบเทียบพหุคูณแล้วแตกต่างอย่างมีนัยสำคัญทางสถิติ จากการวิเคราะห์ด้วยวิธีการเปรียบเทียบพหุคูณที่ต่างกัน ในกลุ่มตัวอย่างที่มีขนาดเดียวกัน โดยวิธี Q - test

6. ทดสอบความแตกต่างของจำนวนกลุ่มที่เมื่อทดสอบการเปรียบเทียบพหุคูณแล้วแตกต่างอย่างมีนัยสำคัญทางสถิติ จากการวิเคราะห์ด้วยวิธีการเปรียบเทียบพหุคูณเดียวกันในกลุ่มตัวอย่างที่มีขนาดต่างกัน โดยวิธี χ^2 - test

การนำเสนอผลการวิเคราะห์ข้อมูล

การศึกษาครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับดังนี้

1. คำสถิติพื้นฐานของคะแนนที่ได้จากการทำแบบทดสอบความถนัดด้านเหตุผล จากกลุ่มตัวอย่างขนาด 583 คน ที่เป็นประชากรเทียม

2. การทดสอบลักษณะข้อมูลเบื้องต้น คือ ความแตกต่างระหว่างค่าเฉลี่ย ซึ่งสุ่มมาจากคะแนนความถนัดด้านเหตุผลของนักเรียนชั้นมัธยมศึกษาปีที่ 2, 4 และ 6 ที่เป็นประชากรเทียม โดยสุ่มแบบใส่คืน ทั้งหมด 4 ขนาด คือ ขนาด 20 คน 40 คน 80 คน และ 120 คน

3. การทดสอบความแตกต่างของผลการเปรียบเทียบ พหุคูณจากการทดสอบพหุคูณรายคู่ 3 วิธี จากกลุ่มตัวอย่างขนาดเดียวกัน คือ ขนาด 20 คน 40 คน 80 คน และ 120 คน

4. การทดสอบความแตกต่างของผลการเปรียบเทียบพหุคูณจากกลุ่มตัวอย่างขนาดต่างกัน 4 ขนาด ในแต่ละวิธีการทดสอบพหุคูณรายคู่ 3 วิธี คือ วิธี Dunnett's T3 วิธี Games-Howell และ วิธี Brown - Forsythe

สรุปผลการวิจัย

1. ผลการเปรียบเทียบพหุคูณด้วย วิธี Dunnett's T3 วิธี Games - Howell และ วิธี Brown - Forsythe ในกลุ่มตัวอย่างขนาด 20 คน ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และในกลุ่มตัวอย่างขนาด 40 คน และ ขนาด 80 คนให้ผลการไม่ยอมรับ

สมมติฐานแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติเช่นเดียวกัน ส่วนในกลุ่มตัวอย่างขนาด 120 คน ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยวิธี Dunnett's T3 กับ วิธี Brown - Forsythe มีผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .003 และวิธี Games - Howell กับ วิธี Brown - Forsythe มีผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .003

2. ผลการเปรียบเทียบพหุคูณด้วยวิธี Dunnett's T3 จากกลุ่มตัวอย่างขนาด 20 คน 40 คน 80 คน และ 120 คน ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยกลุ่มตัวอย่างขนาด 20 คน กับ 80 คน ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025 และขนาด 20 คน กับ 120 คน ขนาด 40 คน กับ 80 คน ขนาด 40 คน กับ 120 คน และขนาด 80 คน กับ 120 คน ก็ให้ผลการยอมรับและไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025 เช่นเดียวกัน ผลการเปรียบเทียบพหุคูณด้วยวิธี Games - Howell จากกลุ่มตัวอย่างขนาด 20 คน 40 คน 80 คน และ 120 คน ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยกลุ่มตัวอย่างขนาด 20 คน กับ 80 คน ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025 และขนาด 20 คน กับ 120 คน ขนาด 40 คน กับ 80 คน ขนาด 40 คน กับ 120 คน และขนาด 80 คน กับ 120 คน ก็ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025 เช่นเดียวกัน และผลการเปรียบเทียบพหุคูณด้วยวิธี Brown - Forsythe จากกลุ่มตัวอย่างขนาด 20 คน 40 คน 80 คน และ 120 คน ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยกลุ่มตัวอย่างขนาด 20 คน กับ 80 คน ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025 และขนาด 20 คน กับ 120 คน และขนาด 80 คน กับ 120 คน ก็ให้ผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .0025 เช่นเดียวกัน

อภิปรายผล

การวิจัยครั้งนี้ มีจุดมุ่งหมายสำคัญเพื่อการศึกษาผลการเปรียบเทียบพหุคูณ จากการทดสอบพหุคูณรายคู่ 3 วิธี และกลุ่มตัวอย่างขนาดต่างกัน 4 ขนาดว่าจะได้จำนวนผลการเปรียบเทียบพหุคูณที่ไม่ยอมรับสมมติฐานต่างกันหรือไม่ เมื่อลักษณะของข้อมูลที่น่ามาวิเคราะห์ไม่เป็นไปตามข้อตกลงเบื้องต้นทางสถิติ คือข้อมูลของกลุ่มตัวอย่างมีความแปรปรวนแตกต่างกัน ผลการวิเคราะห์สามารถนำมาอภิปรายได้ดังนี้

1. การศึกษาผลการเปรียบเทียบพหุคูณที่ไม่ยอมรับสมมติฐาน จากการทดสอบพหุคูณรายคู่ 3 วิธี ในแต่ละขนาดกลุ่มตัวอย่าง พบว่า ขนาด 120 คน จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญที่ .01 โดยวิธี Brown-Forsythe มีจำนวนผลการไม่ยอมรับสมมติฐานน้อยที่สุด ส่วนวิธี Dunnett's T3 และวิธี Games-Howell มีจำนวนผลการไม่ยอมรับสมมติฐานเท่ากัน ซึ่งสอดคล้องกับสมมติฐานของการศึกษาข้อ 1.4 ที่ตั้งไว้ และสอดคล้องกับการศึกษาของ พหล ศักดิ์กะทัศน์ (2534:บทคัดย่อ) ได้ศึกษาการเปรียบเทียบวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยในการวิเคราะห์ความแปรปรวน 3 วิธีในกลุ่มตัวอย่างเดียวกัน ซึ่งพบว่า การวิเคราะห์ทั้ง 3 วิธีที่ศึกษานั้นแตกต่างกัน ส่วนที่กลุ่มตัวอย่างขนาด 20 คน 40 คน และ 80 คน มีจำนวนผลการเปรียบเทียบพหุคูณที่ไม่ยอมรับสมมติฐานแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ โดยวิธี Brown-Forsythe มีจำนวนผลการไม่ยอมรับสมมติฐานน้อยที่สุด ส่วนวิธี Dunnett's T3 และวิธี Games-Howell มีจำนวนผลการไม่ยอมรับสมมติฐานเท่ากัน ซึ่งไม่สอดคล้องกับสมมติฐานของการศึกษา ข้อ 1.1 ข้อ 1.2 และข้อ 1.3 ที่ตั้งไว้ เมื่อข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นทางสถิติ คือความแปรปรวนแตกต่างกัน จึงต้องทำการทดสอบก่อนที่จะนำมาวิเคราะห์หาความแตกต่างของค่าเฉลี่ย อาจจะต้องมีการเปลี่ยนแปลงข้อมูล เพราะจะมีผลกระทบต่อความมีนัยสำคัญทางสถิติของกลุ่มจัดกระทำ (Dayton.1970:33) หรือใช้สถิติที่เอื้อประโยชน์กับข้อมูลที่ไม่เป็นไปตามข้อตกลงเบื้องต้นทางสถิติ จากการวิจัย พบว่า จำนวนผลการไม่ยอมรับสมมติฐานของทั้ง 3 วิธีนั้นแตกต่างกัน เมื่อรวมทุกขนาดกลุ่มตัวอย่างที่ทำการศึกษาแล้ว จำนวนผลการไม่ยอมรับสมมติฐานของวิธี Brown-Forsythe มีจำนวนน้อยที่สุด เมื่อเทียบกับอีกสองวิธี วิธี Brown-Forsythe จึงเป็นวิธีที่อยู่ในลักษณะอนุรักษนิยม(Conservative) (Kirk.1995:155)

2. การศึกษาจำนวนผลการไม่ยอมรับสมมติฐานในระหว่างขนาดกลุ่มตัวอย่าง 4 ขนาด จากการทดลองพหุคูณรายคู่ 3 วิธี พบว่า จำนวนผลการไม่ยอมรับสมมติฐานแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 ซึ่งสอดคล้องกับสมมติฐานของการศึกษาข้อ 2 ที่ตั้งไว้ จากการศึกษาจะเห็นว่า ขนาดของกลุ่มตัวอย่างมีผลต่อจำนวนผลการไม่ยอมรับสมมติฐาน ซึ่งสอดคล้องกับการศึกษาของ สมคิด ไวยวุฒินันท์ (2530:บทคัดย่อ) ได้ศึกษาการเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของวิธีการเปรียบเทียบพหุคูณ 3 วิธี โดยศึกษากับกลุ่มตัวอย่าง 4 ขนาด พบว่า ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 จะแตกต่างกันไปตามขนาดของกลุ่มตัวอย่าง ซิงเคิล และ โอลิเวอร์ (Hinkle & Oliver.1983:1051-1060) กล่าวว่า ขนาดกลุ่มตัวอย่างมีความสัมพันธ์กับระดับนัยสำคัญทางสถิติ จากการศึกษาพบอีกว่า

จำนวนผลการไม่ยอมรับสมมติฐานนั้น จะแปรผันตามขนาดกลุ่มตัวอย่าง เช่น ขนาดกลุ่มตัวอย่าง 20 คน มีจำนวนการไม่ยอมรับสมมติฐาน 30 กลุ่ม จากทั้งหมด 90 กลุ่ม จากทุกวิธีการเปรียบเทียบพหุคูณ ส่วนขนาดกลุ่มตัวอย่าง 120 คน มีจำนวนการไม่ยอมรับสมมติฐาน 84 กลุ่ม จากทั้งหมด 90 กลุ่ม จากทุกวิธีการเปรียบเทียบพหุคูณ เนื่องจากเมื่อใช้กลุ่มตัวอย่างขนาดใหญ่ การแจกแจงของค่าสถิติของกลุ่มตัวอย่างจะมีลักษณะใกล้เคียงกับลักษณะการแจกแจงของประชากร ทำให้ความคลาดเคลื่อนมาตรฐานมีค่าน้อย ค่าความคลาดเคลื่อนประเภทที่ 1 และประเภทที่ 2 มีค่าน้อยกว่าการวิจัยที่ใช้กลุ่มตัวอย่างขนาดเล็ก และการทดสอบสมมติฐานจะปฏิเสธสมมติฐานหลักอย่างมีนัยสำคัญทางสถิติ อันจะส่งผลให้การสรุปเกี่ยวกับประชากรแม่นยำ (นงลักษณ์ วิรัชชัย.2537:22)

ข้อเสนอแนะ

ผลการวิจัยครั้งนี้สามารถสรุปเป็นข้อเสนอแนะในการนำไปใช้และการวิจัยครั้งต่อไป ดังนี้

1. ข้อเสนอแนะเพื่อการเลือกใช้วิธีการเปรียบเทียบพหุคูณสำหรับข้อมูลที่มีความแปรปรวนแตกต่างกัน

1.1 ในการวิเคราะห์หาความแตกต่างของค่าเฉลี่ยจากคะแนนความถนัดด้านเหตุผลแบบสรุปความของกลุ่มตัวอย่าง สามารถเลือกใช้วิธีการเปรียบเทียบพหุคูณวิธี Dunnett's T3 วิธี Games-Howell และวิธี Brown-Forsythe วิธีใดก็ได้ แต่ถ้าต้องการผลการทดสอบ ว่ามีความแตกต่างระหว่างค่าเฉลี่ยอย่างมั่นใจ ควรเลือกใช้วิธี Brown-Forsythe

1.2 ในการเลือกขนาดกลุ่มตัวอย่างเพื่อใช้ในการวิเคราะห์คะแนนความถนัดด้านเหตุผล แบบสรุปความ ควรเลือกกลุ่มตัวอย่างที่มีขนาดใหญ่ตั้งแต่ 80 คนขึ้นไป

2. ข้อเสนอแนะในการวิจัย

2.1 ควรศึกษาผลการเปรียบเทียบพหุคูณที่ได้จากการทดสอบพหุคูณวิธีอื่นอีก เช่น วิธี Dunnett วิธี Holm ซึ่งเป็นวิธีการเปรียบเทียบพหุคูณแบบวางแผน

2.2 ควรศึกษาผลการเปรียบเทียบพหุคูณที่ทำการศึกษาด้วยวิธีการจำลองข้อมูล (Simulation) ด้วยคอมพิวเตอร์

บรรณานุกรม

บรรณานุกรม

- กมลทิพย์ ปรัชญชรินทร์. (2541). การเปรียบเทียบตัวสถิติทดสอบความแตกต่างของค่าเฉลี่ย
ประชากรหลายกลุ่ม. วิทยานิพนธ์ พ.ม. (สถิติ). กรุงเทพฯ: บัณฑิตวิทยาลัย
จุฬาลงกรณ์มหาวิทยาลัย, ถ่ายเอกสาร.
- กัลยา วาณิชยปัญญา. (2539). การวิเคราะห์สถิติ:สถิติเพื่อการตัดสินใจ. กรุงเทพฯ: ภาควิชา
สถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย.
- กึ่งทอง ยงยุทธมีชัย. (2538). การเปรียบเทียบสถิติทดสอบค่าเฉลี่ยของประชากรที่มีความ
แปรปรวนไม่เท่ากัน. วิทยานิพนธ์ พ.ม.(สถิติ) . กรุงเทพฯ : บัณฑิตวิทยาลัย
จุฬาลงกรณ์มหาวิทยาลัย, ถ่ายเอกสาร.
- ชวาล แพรัตกุล. (2517). การทดสอบเพื่อค้นหาและพัฒนาสมรรถภาพ. กรุงเทพฯ :
สำนักทดสอบทางการศึกษาและจิตวิทยา วิทยาลัยการศึกษา ประสานมิตร.
- ดาร์รัตน์ พิกุลทอง. (2537). หลักสถิติ. กรุงเทพฯ: ภาควิชาคณิตศาสตร์และสถิติ
สถาบันราชภัฏพระนคร.
- ทวีศิลป์ กลอนภาดล.(2538).การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อกำหนดขนาดตัวอย่างและ
อำนาจทางสถิติสำหรับการทดสอบความแตกต่างของค่ามัธยฐานเลขคณิต.
วิทยานิพนธ์ ค.ม.(วิจัยการศึกษา).กรุงเทพฯ:จุฬาลงกรณ์มหาวิทยาลัย,
ถ่ายเอกสาร.
- ทองหล่อ วิภาวิณ. (2523). การวัดความถนัด. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและ
จิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ.
- นันทา วงษ์วิโรจน์. (2532). สถิติทดสอบที่มีความแข็งแกร่งสำหรับทดสอบความเท่ากันของค่าเฉลี่ย
ประชากร. วิทยานิพนธ์ พ.ม.(สถิติ) . กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์
มหาวิทยาลัย, ถ่ายเอกสาร.
- นันทวัน บำรุงสวัสดิ์. (2533). การเปรียบเทียบวิธีทดสอบความเท่ากันของค่าเฉลี่ย เมื่อความ
แปรปรวนของประชากรไม่เท่ากัน. วิทยานิพนธ์ พ.ม.(สถิติ) . กรุงเทพฯ :
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, ถ่ายเอกสาร.
- บุญชม ศรีสะอาด. (2538). วิธีการทางสถิติสำหรับวิธีการวิจัย 2. กรุงเทพฯ: สุวีริยาสาส์น
- บุญเชิด ภิญโญอนันตพงษ์.(2527).การทดสอบอิงเกณฑ์: แนวคิดและวิธีการ. พิมพ์ครั้งที่ 1
สำนักพิมพ์โอเดียนสโตร์

- บุปผา จงเลิศตระกูล. (2532). การเปรียบเทียบเชิงพหุระหว่างประชากรใดๆ กับประชากร ควบคุมภายใต้การแจกแจงแบบเอกซ์โพเนนเชียล. วิทยานิพนธ์ พ.ม.(สถิติ) . กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, ถ่ายเอกสาร.
- ไพบรมา พจนพิมล. (2526). การศึกษาแบบมอนติคาโล: การเปรียบเทียบอำนาจการทดสอบของ ที-เทส วิลค็อกซอน เทส เกอร์-โฮฟฟ์ดิง นอร์มอล-สกออร์ เทส และ แวน เดอ แวร์เดน นอร์มอล-สกออร์ เทส ภายใต้ลักษณะการแจกแจงของประชากร 3 แบบ. วิทยานิพนธ์ ค.ม.(วิจัยการศึกษา). กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย, ถ่ายเอกสาร.
- พหล ศักดิ์คะทนต์. (2534). การเปรียบเทียบประสิทธิภาพของวิธีทดสอบความแตกต่างระหว่างค่าเฉลี่ยในการวิเคราะห์ความแปรปรวน. วิทยานิพนธ์ พ.ม.(สถิติ) . กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, ถ่ายเอกสาร.
- ระวีวรรณ พันธุ์พานิชย์. (2540). แบบแผนเชิงสถิติของการทดลอง. กรุงเทพฯ: ภาควิชาวัดผลและวิจัยทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- เยาวภา ไชยศรี. (2536). การเปรียบเทียบประสิทธิภาพของตัวทดสอบสถิติบาร์เลต เลอวิน และ บราว-ฟอร์ลิตส์ สำหรับการทดสอบความเท่ากันของความแปรปรวนของประชากร. วิทยานิพนธ์ วท.ม.(สถิติ). กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์, ถ่ายเอกสาร.
- ล้วน สายยศ และอังคณา สายยศ. (2538). เทคนิคการวิจัยทางการศึกษา. กรุงเทพฯ: สุวีริยาสาส์น.
- (2538). เทคนิคการวิจัยทางการศึกษา. กรุงเทพฯ: สุวีริยาสาส์น.
- (2539). เทคนิคการวัดผลการเรียนรู้. กรุงเทพฯ: สุวีริยาสาส์น.
- (2540). สถิติวิทยาทางการวิจัย. กรุงเทพฯ: สุวีริยาสาส์น.
- (2541) เทคนิคการสร้างและสอบข้อสอบความถนัดทางการเรียนและความสามารถทั่วไป. กรุงเทพฯ: สุวีริยาสาส์น.
- วิชชุดา ศรีโสภณ. (2539). การเปรียบเทียบพหุ. วิทยานิพนธ์ พ.ม.(สถิติ) . กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, ถ่ายเอกสาร.
- วิรัช วรรณรัตน์. (2535). วิธีการทาสถิติสำหรับการวิจัย. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ.
- วิไลลักษณ์ องค์จิระวุฒิ. (2522). การเปรียบเทียบวิธีการต่างๆ ที่ใช้ในการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากร โดยพิจารณาจากความผิด 3 ชนิด. วิทยานิพนธ์ พ.ม. (สถิติ) . กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย; ถ่ายเอกสาร.

- ศิริชัย กาญจนวสี และคณะ. (2537). *การเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัย*. กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- สมคิด ไวยวุฒินันท์. (2530). *การเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของการเปรียบเทียบพหุคูณระหว่าง บอนเฟอโรนี ไคสแควร์ของมาร์ซูโล และวิธีของทัมฮานน์*. วิทยานิพนธ์ ค.ม.(สถิติการศึกษา) . กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, ถ่ายเอกสาร.
- สมจิต วัฒนชายงกุล. (2529). *สถิติวิเคราะห์เบื้องต้นสำหรับการวิจัย*. กรุงเทพฯ: สำนักพิมพ์ประกายพริก.
- สายทิพย์ แดด้สี. (2536). *การสร้างแบบทดสอบความถนัดทางด้านเหตุผล*. ปรินญานิพนธ์ กศ.ม.(วัดผลทางการศึกษา). พิษณุโลก: บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร, ถ่ายเอกสาร.
- อุทุมพร ทองอุไทย. (2523). *แผนวิเคราะห์ข้อมูลพฤติกรรมศาสตร์*. กรุงเทพฯ: โรงพิมพ์เจริญผล
- Bingham,Walt van Duke. (1987). *Aptitude and Aptitude testing*. New York:Harper and Borthers.
- Bradley,J.V. (1978). Robustness?. *British Journal of Mathematic and Statistical Psychology*, 31, 144-152.
- Brown,M.B. & Forsythe,A.B. (1974). The ANOVO and mutiple comparison for data with heterogeneous variance," *Biometrics*. 30: 791-724.
- Cochran,William.G and Cox,Geetrude.M. (1975). *Experimental design*. New York : John Wiley & Son.
- Cronbach,L.J.,N.Rajaratnam and G.C.Glaser. (1970). *Essentials of Psychological Testing*. New York : Harper & Row Publish.
- Dayton,C.Mitchell.(1970).*The Design of Educational Experiment*. Mc. Grae-Hill University of Marryland.
- De Cecco, John Paul. (1967). *Thepsychology of language, thought and instruction : read*. New York :Holt,Rinehart and Winston.
- Douglas A. Penfield & Stephen L. Koffler. (1978). "Post Hoc Procedures For Some K-Sample Nonparametric Tests For Scale," *Journal of Educational Statistic*. 3 : 265-282

- Dunnett, C.W. (1980). "Pairwise multiple comparisons in the unequal variance case." *Journal of the American Statistical Association*. 75 : 796-800.
- Freeman, Frank S. (1966). *Theory and Practice of Psychological testing*. 3rd ed. New York : Rinehart and Winston .
- Games, P.A. & Howell, J.F. (1976). "Pairwise multiple comparison procedures with unequal N's and/or variance," *Journal of Educational Statistics*. 1: 113-125.
- Gene, V. Glass. & Julian, C. Stanley. (1970). *Statistical Methods In Education and Psychology*. Prentice Hall, Inc, New Jersey.
- Guildford, J.P. and Fruchter, B. (1978). *Fundamental Statistics in Psychology and Education*. Singapore: McGraw-Hill Co., .
- 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 No simple Answer ? *Educational and Psychological Measurement*. 43:1051-1060.
- 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.
- James V. Couch. (1982). *Fundamentals of Statistics For the Behavioral Science*. New York: St. Martin's Press.
- Kerlinger, F.N. (1986). *Fundamentals of Behavioral Research*. Florida : Holt, Rinehart and Winston.
- Kirk, R.E. (1995). *Experimental design: procedures for the behavioral science*.
- Levene, H. (1960). Robust test for equality of variances. *Contributions to probability and Statistics : Essays in honor of Harold 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*.

- Neil A. Weiss. (1995). *Introduction Statistics*. Addison-Wesley Publishing Company, Inc.
- Philip H. Ramsey. (1978). "Power Differences Between Pairwise multiple Comparisons," *Journal of the American Statistical Association*. 73:479-487.
- R. Clifford Blair & James J. Hinggin. (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.
- Rand R. Wilcox. (1989). "Adjusting for Unequal Variances When Comparing Means in One-Way and Two-Way Fixed Effects ANOVA Models," *Journal of Educational Statistics*. 14:270-278.
- Richard C. Sprinthall. (1994). *Basic Statistical Analysis*. A Division of Paramount Publishing.
- Toothaker, L.E. (1991). *Multiple comparisons*. Newbury Park, CA : Sage.

ภาคผนวก

ภาคผนวก ก.

- ค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี
Dunnett's T3 จากกลุ่มตัวอย่าง 4 ขนาด
- ค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี
Games-Howell จากกลุ่มตัวอย่าง 4 ขนาด
- ค่าความแตกต่างของค่าเฉลี่ย 2 ระดับชั้นที่คำนวณด้วยวิธี
Brown-Forsythte จากกลุ่มตัวอย่าง 4 ขนาด
- ค่าความแตกต่างของค่าความแปรปรวน 3 ระดับชั้นที่คำนวณด้วยวิธี
Levene จากกลุ่มตัวอย่าง ขนาด 20 คน
- ค่าความแตกต่างของค่าความแปรปรวน 3 ระดับชั้นที่คำนวณด้วยวิธี
Levene จากกลุ่มตัวอย่าง ขนาด 40 คน
- ค่าความแตกต่างของค่าความแปรปรวน 3 ระดับชั้นที่คำนวณด้วยวิธี
Levene จากกลุ่มตัวอย่าง ขนาด 80 คน
- ค่าความแตกต่างของค่าความแปรปรวน 3 ระดับชั้นที่คำนวณด้วยวิธี
Levene จากกลุ่มตัวอย่าง ขนาด 120 คน
- จำนวนผลการเปรียบเทียบพหุคูณ ที่ทำการทดสอบพหุคูณรายคู่ ด้วยวิธี
Dunnett's T3 จากกลุ่มตัวอย่าง 4 ขนาด
- จำนวนผลการเปรียบเทียบพหุคูณ ที่ทำการทดสอบพหุคูณรายคู่ ด้วยวิธี
Games-Howell จากกลุ่มตัวอย่าง 4 ขนาด
- จำนวนผลการเปรียบเทียบพหุคูณ ที่ทำการทดสอบพหุคูณรายคู่ ด้วยวิธี
Brown-Forsythte จากกลุ่มตัวอย่าง 4 ขนาด

ตาราง 21 แสดงค่าความแตกต่างระหว่างค่าเฉลี่ยของ 2 ระดับชั้น ที่ทำการทดสอบด้วยพหุคูณ
รายคู่ด้วยวิธี Dunnett's T3 จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ระหว่างชั้น	ขนาดกลุ่มตัวอย่าง (คน)			
		20	40	80	120
1	2 , 4	-5.80*	-4.27*	-4.57*	-4.92*
2	2 , 6	-8.45*	-6.67*	-7.97*	-7.76*
3	4 , 6	-2.65	-2.40	-3.40*	-2.84*
4	2 , 4	-4.15	-5.07*	-4.63*	-5.24*
5	2 , 6	-7.50*	-7.50*	-8.03*	-8.12*
6	4 , 6	-3.35	-2.42	-3.40*	-2.88*
7	2 , 4	-5.80*	-3.97*	-4.46*	-4.13*
8	2 , 6	-8.00*	-8.05*	-7.92*	-7.91*
9	4 , 6	-2.20	-4.07*	-3.46*	-3.78*
10	2 , 4	-3.65	-4.37*	-3.57*	-4.09*
11	2 , 6	-6.65*	-6.40*	-6.61*	-6.95*
12	4 , 6	-3.00	-2.02	-3.03*	-2.85*
13	2 , 4	-5.50	-4.07*	-4.30*	-3.71*
14	2 , 6	-6.55*	-6.55*	-6.50*	-6.50*
15	4 , 6	-1.05	-2.47	-2.20	-2.79*
16	2 , 4	-4.85	-3.82*	-3.06*	-3.64*
17	2 , 6	-6.00*	-6.15*	-6.16*	-6.54*
18	4 , 6	-1.15	-2.32	-3.10*	-2.90*
19	2 , 4	-3.50	-3.37*	-3.07*	-3.22*
20	2 , 6	-6.80*	-6.97*	-6.91*	-6.31*
21	4 , 6	-3.30	-3.60*	-3.83*	-3.09*
22	2 , 4	-3.60	-3.32*	-3.06*	-3.32*
23	2 , 6	-6.80*	-7.02*	-7.01*	-6.60*
24	4 , 6	-3.20	-3.70*	-3.95*	-3.28*
25	2 , 4	-4.70*	-2.72	-3.375*	-3.50*
26	2 , 6	-7.60*	-6.05*	-6.80*	-6.65*
27	4 , 6	-2.90	-3.32*	-3.42*	-3.15*
28	2 , 4	-2.05	-3.62*	-2.86*	-3.39*
29	2 , 6	-5.60	-7.30*	-6.67*	-6.25*
30	4 , 6	-3.60	-3.67*	-3.81*	-2.83*

ตาราง 22 แสดงค่าความแตกต่างระหว่างค่าเฉลี่ยของ 2 ระดับชั้น ที่ทำการทดสอบด้วยพหุคูณ
รายคู่ด้วยวิธี Games-Howell จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่าง ที่	ระหว่างชั้น	ขนาดกลุ่มตัวอย่าง (คน)			
		20	40	80	120
1	2 , 4	-5.80*	-4.27*	-4.57*	-4.92*
2	2 , 6	-8.45*	-6.67*	-7.97*	-7.76*
3	4 , 6	-2.65	-2.40	-3.40*	-2.84*
4	2 , 4	-4.15	-5.07*	-4.63*	-5.24*
5	2 , 6	-7.50*	-7.50*	-8.03*	-8.12*
6	4 , 6	-3.35	-2.42	-3.40*	-2.88*
7	2 , 4	-5.80*	-3.97*	-4.46*	-4.13*
8	2 , 6	-8.00*	-8.05*	-7.92*	-7.91*
9	4 , 6	-2.20	-4.07*	-3.46*	-3.78*
10	2 , 4	-3.65	-4.37*	-3.57*	-4.09*
11	2 , 6	-6.65*	-6.40*	-6.61*	-6.95*
12	4 , 6	-3.00	-2.02	-3.03*	-2.85*
13	2 , 4	-5.50	-4.07*	-4.30*	-3.71*
14	2 , 6	-6.55*	-6.55*	-6.50*	-6.50*
15	4 , 6	-1.05	-2.47	-2.20	-2.79*
16	2 , 4	-4.85	-3.82*	-3.06*	-3.64*
17	2 , 6	-6.00*	-6.15*	-6.16*	-6.54*
18	4 , 6	-1.15	-2.32	-3.10*	-2.90*
19	2 , 4	-3.50	-3.37*	-3.07*	-3.22*
20	2 , 6	-6.80*	-6.97*	-6.91*	-6.31*
21	4 , 6	-3.30	-3.60*	-3.83*	-3.09*
22	2 , 4	-3.60	-3.32*	-3.06*	-3.32*
23	2 , 6	-6.80*	-7.02*	-7.01*	-6.60*
24	4 , 6	-3.20	-3.70*	-3.95*	-3.28*
25	2 , 4	-4.70*	-2.72	-3.375*	-3.50*
26	2 , 6	-7.60*	-6.05*	-6.80*	-6.65*
27	4 , 6	-2.90	-3.32*	-3.42*	-3.15*
28	2 , 4	-2.05	-3.62*	-2.86*	-3.39*
29	2 , 6	-5.60	-7.30*	-6.67*	-6.25*
30	4 , 6	-3.60	-3.67*	-3.81*	-2.83*

ตาราง 23 แสดงค่าความแตกต่างระหว่างค่าเฉลี่ยของ 2 ระดับชั้น ที่ทำการทดสอบด้วยพหุคูณ
รายคู่ด้วยวิธี Brown-Forsythe จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ระหว่างชั้น	ขนาดกลุ่มตัวอย่าง (คน)			
		20	40	80	120
1	2, 4	-5.80	-4.27*	-4.57*	-4.92*
2	2, 6	-8.45*	-6.67*	-7.97*	-7.76*
3	4, 6	-2.65	-2.40	-3.40	-2.84*
4	2, 4	-4.15*	-5.07*	-4.63*	-5.24*
5	2, 6	-7.50	-7.50*	-8.03*	-8.12*
6	4, 6	-3.35	-2.42	-3.40*	-2.88*
7	2, 4	-5.80	-3.97*	-4.46*	-4.13*
8	2, 6	-8.00	-8.05*	-7.92*	-7.91*
9	4, 6	-2.20	-4.07	-3.46*	-3.78*
10	2, 4	-3.65	-4.37	-3.57	-4.09
11	2, 6	-6.65*	-6.40*	-6.61*	-6.95*
12	4, 6	-3.00	-2.02*	-3.03*	-2.85*
13	2, 4	-5.50	-4.07	-4.30	-3.71
14	2, 6	-6.55	-6.55*	-6.50*	-6.50*
15	4, 6	-1.05	-2.47	-2.20*	-2.79*
16	2, 4	-4.85	-3.82	-3.06	-3.64
17	2, 6	-6.00*	-6.15*	-6.16*	-6.54*
18	4, 6	-1.15	-2.32*	-3.10*	-2.90*
19	2, 4	-3.50	-3.37	-3.07	-3.22
20	2, 6	-6.80	-6.97*	-6.91*	-6.31*
21	4, 6	-3.30	-3.60*	-3.83*	-3.09*
22	2, 4	-3.60	-3.32	-3.063	-3.32
23	2, 6	-6.80	-7.02*	-7.013*	-6.60*
24	4, 6	-3.20	-3.70*	-3.95*	-3.28*
25	2, 4	-4.70	-2.72	-3.375*	-3.50
26	2, 6	-7.60	-6.05*	-6.80*	-6.65*
27	4, 6	-2.90*	-3.32*	-3.42*	-3.15*
28	2, 4	-2.05	-3.62	-2.86	-3.39*
29	2, 6	-5.60	-7.30	-6.67*	-6.25*
30	4, 6	-3.60*	-3.67*	-3.81*	-2.83*

ตาราง 24 แสดงค่าความแตกต่างระหว่างความแปรปรวนของ 3 ระดับชั้น ที่ทำการทดสอบด้วยวิธีของ Levene เมื่อกลุ่มตัวอย่างมีขนาด 20 คน

กลุ่ม ตัวอย่าง ที่	$\bar{X}$			S^2			L
	ม.2	ม.4	ม.6	ม.2	ม.4	ม.6	
1	9.35	15.15	17.80	3.50	48.97	45.74	9.87**
2	9.95	14.10	17.45	8.26	27.88	36.99	4.24**
3	9.50	15.30	17.50	14.47	29.48	43.63	3.95**
4	10.55	14.20	17.20	15.73	41.22	6.58	6.10**
5	10.45	15.95	17.00	28.26	48.89	11.68	4.60*
6	11.00	15.85	17.00	24.00	47.81	12.21	5.70**
7	11.15	14.65	17.95	23.81	11.50	31.94	1.09*
8	11.10	14.70	17.90	25.56	7.69	41.57	3.74*
9	9.80	14.50	17.40	10.58	11.94	32.35	3.44*
10	11.45	13.50	17.10	29.41	7.52	44.93	3.79*

ตาราง 25 แสดงค่าความแตกต่างระหว่างความแปรปรวนของ 3 ระดับชั้น ที่ทำการทดสอบด้วยวิธีของ Levene เมื่อกลุ่มตัวอย่างมีขนาด 40 คน

กลุ่ม ตัวอย่าง ที่	$\bar{X}$			S^2			L
	ม.2	ม.4	ม.6	ม.2	ม.4	ม.6	
1	10.75	14.45	16.85	11.89	39.84	57.92	11.84**
2	10.00	15.07	17.50	5.94	28.17	45.53	14.93**
3	10.75	14.15	18.22	6.40	30.64	32.69	12.46**
4	11.25	15.50	17.52	29.08	52.35	6.71	14.96**
5	10.85	14.92	17.40	18.28	36.63	10.86	7.95**
6	10.62	14.45	16.77	16.95	44.45	9.51	10.60**
7	10.57	13.95	17.55	25.68	8.40	33.53	5.31**
8	10.65	13.97	17.67	28.90	3.76	36.53	8.83**
9	11.42	14.15	17.47	37.43	6.07	34.87	7.92**
10	10.37	14.00	17.67	26.18	6.35	32.22	6.43**

ตาราง 26 แสดงค่าความแตกต่างระหว่างความแปรปรวนของ 3 ระดับชั้น ที่ทำการทดสอบด้วยวิธีของ Levene เมื่อกลุ่มตัวอย่างมีขนาด 80 คน

กลุ่ม ตัวอย่าง ที่	$\bar{X}$			S^2			L
	ม.2	ม.4	ม.6	ม.2	ม.4	ม.6	
1	9.83	14.41	17.81	5.25	30.49	38.15	26.98**
2	10.12	14.76	18.61	6.97	28.33	39.32	26.11**
3	9.78	14.25	17.71	5.13	28.79	40.10	34.49**
4	11.05	14.62	17.66	25.08	38.38	6.63	23.85**
5	10.51	14.81	17.01	24.68	38.33	9.96	19.98**
6	11.42	14.48	17.58	29.41	39.06	8.57	20.49**
7	10.90	13.97	17.81	31.71	9.89	36.43	9.89**
8	10.83	13.90	17.85	28.23	7.63	29.95	11.87**
9	10.96	14.33	17.76	28.89	11.23	32.48	7.20**
10	11.06	13.92	17.73	29.57	6.98	29.53	11.40**

ตาราง 27 แสดงค่าความแตกต่างระหว่างความแปรปรวนของ 3 ระดับชั้น ที่ทำการทดสอบด้วยวิธีของ Levene เมื่อกลุ่มตัวอย่างมีขนาด 120 คน

กลุ่ม ตัวอย่าง ที่	$\bar{X}$			S^2			L
	ม.2	ม.4	ม.6	ม.2	ม.4	ม.6	
1	9.77	14.70	17.54	9.65	27.15	38.62	50.36**
2	9.69	14.93	17.81	3.74	31.57	41.00	43.79**
3	9.95	14.05	17.86	4.78	35.37	39.91	46.68**
4	10.39	14.48	17.34	19.33	37.15	9.42	30.70**
5	10.33	14.55	17.34	28.82	30.87	8.93	22.13**
6	10.96	14.60	17.50	26.15	39.19	6.62	33.65**
7	11.14	14.36	17.45	33.93	10.30	34.55	14.97**
8	11.03	14.35	17.64	35.12	8.24	31.15	17.62**
9	10.78	14.28	17.44	26.82	7.93	36.29	16.63**
10	11.05	14.45	17.28	29.93	11.79	44.59	15.22**

ตาราง 28 แสดงจำนวนผลการเปรียบเทียบพหุคูณ ที่ทำการทดสอบด้วยพหุคูณรายคู่ด้วยวิธี Dunnett's T3 จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ขนาดกลุ่มตัวอย่าง (คน)			
	20	40	80	120
1	1	1	1	1
2	1	1	1	1
3	0	0	1	1
4	0	1	1	1
5	1	1	1	1
6	0	0	1	1
7	1	1	1	1
8	1	1	1	1
9	0	1	1	1
10	0	1	1	1
11	1	1	1	1
12	0	0	1	1
13	0	1	1	1
14	1	1	1	1
15	0	0	0	1
16	0	1	1	1
17	1	1	1	1
18	0	0	1	1
19	0	1	1	1
20	1	1	1	1
21	0	1	1	1
22	0	1	1	1
23	1	1	1	1
24	0	1	1	1
25	1	0	1	1
26	1	1	1	1
27	0	1	1	1
28	0	1	1	1
29	0	1	1	1
30	0	1	1	1

ตาราง 29 แสดงจำนวนผลการเปรียบเทียบพหุคูณ ที่ทำการทดสอบด้วยพหุคูณรายคู่ด้วยวิธี Games-Howell จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ขนาดกลุ่มตัวอย่าง (คน)			
	20	40	80	120
1	1	1	1	1
2	1	1	1	1
3	0	0	1	1
4	0	1	1	1
5	1	1	1	1
6	0	0	1	1
7	1	1	1	1
8	1	1	1	1
9	0	1	1	1
10	0	1	1	1
11	1	1	1	1
12	0	0	1	1
13	0	1	1	1
14	1	1	1	1
15	0	0	0	1
16	0	1	1	1
17	1	1	1	1
18	0	0	1	1
19	0	1	1	1
20	1	1	1	1
21	0	1	1	1
22	0	1	1	1
23	1	1	1	1
24	0	1	1	1
25	1	0	1	1
26	1	1	1	1
27	0	1	1	1
28	0	1	1	1
29	0	1	1	1
30	0	1	1	1

ตาราง 30 แสดงจำนวนผลการเปรียบเทียบพหุคูณ ที่ทำการทดสอบด้วยพหุคูณรายคู่ด้วยวิธี
Brown-Forsythe จากกลุ่มตัวอย่าง 4 ขนาด

กลุ่มตัวอย่างที่	ขนาดกลุ่มตัวอย่าง (คน)			
	20	40	80	120
1	0	1	1	1
2	1	1	1	1
3	0	0	0	1
4	0	1	1	1
5	1	1	1	1
6	0	0	1	1
7	0	1	1	1
8	0	1	1	1
9	0	0	1	1
10	0	0	0	0
11	1	1	1	1
12	0	1	1	1
13	0	0	0	0
14	0	1	1	1
15	0	0	1	1
16	0	0	0	0
17	1	1	1	1
18	0	1	1	1
19	0	0	0	0
20	0	1	1	1
21	0	1	1	1
22	0	0	0	0
23	0	1	1	1
24	0	1	1	1
25	0	0	1	0
26	0	1	1	1
27	1	1	1	1
28	0	0	0	1
29	0	0	1	1
30	1	1	1	1

หมายเหตุ 0 คือ ไม่ยอมรับสมมติฐาน 1 คือ ยอมรับสมมติฐาน

ภาคผนวก ข.
แบบทดสอบที่ใช้ในการวิจัย

แบบทดสอบสรุปความ

คำชี้แจง

- 1. แบบทดสอบฉบับนี้เป็นแบบทดสอบแบบปรนัยชนิดเลือกตอบ มีจำนวน 30 ข้อ ให้เวลาทำ 20 นาที ฉะนั้นนักเรียนควรรีบตอบโดยเร็วให้ครบทุกข้อจึงจะได้คะแนนดี
- 2. ให้นักเรียนใช้เหตุผลหาข้อสรุปจากข้อความที่กำหนดให้ว่าสรุปอย่างสมเหตุสมผลแล้วจะเป็นข้อใด ดังตัวอย่าง

ตัวอย่าง

- (0) เปิดบางตัวเป็นนก ไก่ทุกตัวเป็นนก นกทุกตัวกินหนอน ดังนั้น
- | | | |
|----------------------|--|----------------------------|
| ก. เปิดทุกตัวเป็นไก่ | | คำตอบที่ถูกต้องคือ ข้อ ข. |
| ข. ไก่ทุกตัวกินหนอน | | เพราะเป็นการสรุปอย่างสม |
| ค. เปิดทุกตัวกินหนอน | | เหตุผลผลดังนั้นให้นักเรียน |
| ง. ไก่บางตัวเป็นเปิด | | กากบาทลงในช่องของ ข้อ ข. |
| จ. สรุปไม่ได้แน่นอน | | ในกระดาษคำตอบดังนี้ |

ข้อ	ก	ข	ค	ง	จ
0		X			

- 3. ถ้านักเรียนพบข้อยากอย่าท้อใจ จงข้ามไปทำข้ออื่นก่อน มีเวลาเหลือจึงย้อนกลับมาทำข้อนั้นใหม่
- 4. อย่าขีดเขียนหรือทำเครื่องหมายใดๆ ลงในข้อสอบฉบับนี้
- 5. ถ้านักเรียนสงสัยข้อใดให้ยกมือถามเสียก่อน เมื่อบอกให้ลงมือทำข้อสอบแล้วนักเรียนจะถามอะไรอีกไม่ได้

1.ทุกคนที่ดื่มเหล้าจะเป็นโรคตับแข็ง นายเดชเป็นโรคตับแข็ง ดังนั้น

- ก. นายเดชต้องตาย
- ข. นายเดชไม่ดื่มเหล้า
- ค. นายเดชต้องไปหาหมอ
- ง. นายเดชต้องดื่มเหล้าทุกเย็น
- จ. ยังสรุปไม่ได้แน่นอน

2.นกเป็นสัตว์ชนิดเดียวที่บินได้ สัตว์ที่บินได้เป็นสัตว์ที่ออกลูกเป็นไข่ ไม่มีเปิดตัวใดเป็นนก ดังนั้น

- ก. เปิดออกลูกเป็นไข่
- ข. เปิดบางตัวบินไม่ได้
- ค. สัตว์บินได้บางตัวออกลูกเป็นไข่
- ง. สัตว์ออกลูกเป็นไข่บางตัวเป็นเปิด
- จ. ยังสรุปไม่ได้แน่นอน

3.หมอบุคคลเป็นเศรษฐกิจ เศรษฐีบางคนตระหนี่ นายสมศักดิ์เป็นหมอบ ดังนั้น

- ก. คนตระหนี่เป็นหมอบ
- ข. นายสมศักดิ์เป็นเศรษฐกิจ
- ค. คนตระหนี่เป็นเศรษฐกิจ
- ง. นายสมศักดิ์เป็นคนตระหนี่
- จ. ยังสรุปไม่ได้แน่นอน

4.ถ้านายออดดื่มน้ำส้มและน้ำกระเจี๊ยบแล้ว นายอู๊ดจะดื่มโค้ก ถ้านายออดดื่มน้ำกระเจี๊ยบดังนั้น

- ก. นายอู๊ดดื่มโค้ก
- ข. นายออดดื่มน้ำส้ม
- ค. นายออดไม่ดื่มโค้ก
- ง. นายอู๊ดไม่ดื่มน้ำส้ม
- จ. ยังสรุปไม่ได้แน่นอน

5.นายชินมีลูกสาวคนเดียวชื่อนางสาย นายแสงเป็นลูกของนางสาย นางจันเป็นลูกนายแสง นายกรุงเป็นลูกเขยนายชิน ดังนั้น

- ก. นายกรุงเป็นปู่นางจัน
- ข. นายชินเป็นปู่นายแสง
- ค. นางสายเป็นยายนางจัน
- ง. นายกรุงเป็นตานายแสง
- จ. ยังสรุปไม่ได้แน่นอน

6.ถ้าเรือแล่นไปตามคลองหลอดเรือจะเร็ว ถ้าเรือแล่นออกแม่น้ำเจ้าพระยาเรือจะช้า แต่เรือต้องแล่นไปตามคลองหลอดหรือแม่น้ำเจ้าพระยา ดังนั้น

- ก. เรือเร็ว
- ข. เรือช้า
- ค. เรือเร็วและจม
- ง. เรือต้องเร็วและช้า
- จ. เรืออาจจะเร็วหรือช้า

7.ทุกคนที่เรียนจบปริญญาตรี ตั้งใจที่จะเรียนต่อปริญญาโท นายทวยจบปริญญาตรีและสอบเรียนต่อปริญญาโทไม่ได้ ดังนั้น

- ก. นายทวยผิดหวัง
- ข. นายทวยเรียนหนังสือไม่เก่ง
- ค. นายทวยดูหนังสือไม่ตรงกับที่ข้อสอบออก
- ง. นายทวยไม่มีสิทธิ์สอบเรียนต่อปริญญาโท
- จ. ยังสรุปไม่ได้แน่นอน

8.นายดำเข้าแถวอยู่ระหว่างนายแดงกับนายเขียว นายแดงยืนอยู่ระหว่างนายดำกับนายขาว นายขาวยืนระหว่างนายแดงกับนายเหลือง อยากทราบว่าใครยืนเป็นอันดับรองสุดท้าย

- ก. นายดำ
- ข. นายขาว
- ค. นายแดง
- ง. นายเขียว
- จ. นายเหลือง

9. คนไม่ดูหนังสือจะสอบไม่ได้ แต่งดูหนังสือ ดังนั้น

- ก. แต่งสอบได้
- ข. แต่งสอบไม่ได้
- ค. แต่งอาจจะสอบไม่ได้
- ง. แต่งสอบได้คะแนนมากกว่าคนไม่ดูหนังสือ
- จ. ยังสรุปไม่ได้แน่นอน

10. ปูวิ่งช้ากว่าป้อม แต่เร็วกว่าป่อง ปีกวิ่งช้ากว่าปอง ใครวิ่งเร็วที่สุด

- ก. ปู
- ข. ป้อม
- ค. ปีก
- ง. ป่อง
- จ. ยังสรุปไม่ได้แน่นอน

11. คนที่เรียนดีแต่ยากจนจะได้ทุน นักกีฬาทุกคนเรียนดี นายชาติเป็นคนเรียนดี ดังนั้น

- ก. นายชาติได้ทุน
- ข. นายชาติยากจน
- ค. นักกีฬาได้ทุน
- ง. นายชาติเป็นนักกีฬา
- จ. ยังสรุปไม่ได้แน่นอน

12. น้ำลตตอผุด น้ำท่วมมีปลามาก ถ้าตอไม่ผุดสรุปว่า

- ก. ปลามีมาก
- ข. น้ำไม่ลต
- ค. ไม่มีตอ
- ง. น้ำไม่ลตและมีปลามาก
- จ. ยังสรุปไม่ได้แน่นอน

13. ถ้าคนไม่ลืมนอะไรเลย คนก็ต้องเป็นบ้า นายสันติจำได้ว่า $\sqrt{4321} = 65.73431$ ดังนั้น

- ก. นายสันติเป็นคนบ้า
- ข. นายสันติมีความจำเป็นยอดเยี่ยม
- ค. นายสันติขี้เกียจถอด $\sqrt{4321}$ เอง

ง. นายสันติไม่เคยลืมนค่า $\sqrt{4321}$

จ. ยังสรุปไม่ได้แน่นอน

14. ค้างคาวเป็นสัตว์ที่มีใบหูและบินได้อย่างนก แต่นกไม่มีใบหู ดังนั้น

- ก. ค้างคาวไม่ใช่นก
- ข. ค้างคาวคือนกที่มีใบหู
- ค. สัตว์ที่บินได้ทุกชนิดเป็นนก
- ง. สัตว์มีหูทุกชนิดเป็นค้างคาว
- จ. ยังสรุปไม่ได้แน่นอน

15. ด้อยถึงโรงเรียนเป็นคนแรก อ้อยถึงโรงเรียนทีหลังด้อย แต่ถึงก่อนตีม อู๊ดถึงโรงเรียนทีหลังด้อย ดันไปโรงเรียนไม่ทัน ดังนั้น

- ก. อู๊ดและอ้อยไปถึงโรงเรียนพร้อมกัน
- ข. อู๊ดไปถึงโรงเรียนก่อนตีม
- ค. ตีมไปถึงโรงเรียนก่อนโรงเรียนเข้า
- ง. ดันไปถึงโรงเรียนเป็นคนสุดท้าย
- จ. ยังสรุปไม่ได้แน่นอน

16. นาย ก. จะมาก็ตอเมื่อนาย ข. มา นาย ข. จะมาก็ตอเมื่อนาย ค. ไม่มา นาย ค. จะมาก็ตอเมื่อนาย ง. มา ดังนั้น

- ก. นาย ก. มาคนเดียว
- ข. นาย ก. และนาย ค. มา
- ค. นาย ข. และนาย ง. มา
- ง. นาย ก. นาย ค. และนาย ง. มา
- จ. ยังสรุปไม่ได้แน่นอน

17. ผลไม้เปรี้ยวทุกชนิดเราเรียกว่าส้ม ส้มทุกชนิดมีเปลือก ส้มบางชนิดกลม มะขวิดเป็นผลไม้รสเปรี้ยวแต่ไม่กลม ดังนั้น

- ก. มะขวิดเป็นส้ม
- ข. มะขวิดมีเปลือก
- ค. มะขวิดบางลูกกลม
- ง. มะขวิดบางลูกไม่เปรี้ยว
- จ. ยังสรุปได้ไม่แน่นอน

18.ปูเงินเดือนมากกว่าปลา ปลาเงินเดือนมากกว่า กุ้ง กุ้งเงินเดือนเท่ากับปู ทราয়เงินเดือนมากกว่า กุ้ง กุ้งจะเงินเดือนขึ้นก็ต่อเมื่อกุ้งเงินเดือนขึ้น แต่ปีนี้กุ้งเงินเดือนขึ้น ดังนั้น

- ก. กุ้งได้เงินเดือนมากที่สุด
- ข. ทราয়ได้เงินเดือนมากที่สุด
- ค. กุ้งได้เงินเดือนมากกว่า ปู
- ง. กุ้งได้เงินเดือนมากกว่าปู
- จ. ปูได้เงินเดือนมากกว่าทราয়

19.ถ้าณราไปเที่ยว นรินทร์จะไปเที่ยว ถ้าดาวไปเที่ยว เดือนจะไปเที่ยว ทุกครั้งที่ดาวไปเที่ยวกับ ณรา แสดงว่า

- ก. ดาว เดือน ณรา นรินทร์ ไปเที่ยวด้วยกันทุกครั้ง
- ข. ถ้าดาวไปเที่ยว นรินทร์ก็ไปเที่ยว
- ค. ดาวและเดือนไปเที่ยวจำนวนครั้งเท่ากัน
- ง. เดือนและนรินทร์ไปเที่ยวด้วยกัน
- จ. ยังสรุปไม่ได้แน่นอน

20.คนทุกคนเมื่อหิวต้องกิน และเมื่อเหนื่อยต้องหยุดพัก นายแสงเดินทางตัดทุ่งมาเพื่อไปตลาด เมื่อเดินทางไปได้ $\frac{1}{3}$ ของระยะทาง เขารู้สึกร้อนเหนื่อย และหิว ดังนั้น

- ก. นายแสงเร่งเดินให้ถึงตลาดโดยเร็ว
- ข. นายแสงหยุดพักที่ได้ต้นไม้ให้หาย
- ค. เหนื่อยเสียก่อนจึงเดินต่อ
- ง. นายแสงกลับบ้าน
- จ. นายแสงจะไม่เลือกเดินทางนี้อีก
- ฉ. ยังสรุปไม่ได้แน่นอน

21.ถ้าฝนตก กบจะร้อง ถ้าฝนตกและแดดออก

นกกระจอกจะเข้ารัง ถ้าฝนตกฟ้าร้องและฟ้าผ่า คนจะรีบเข้าบ้าน เมื่อวานนี้ฝนตก แดดออก และฟ้าร้อง ดังนั้น

- ก. กบร้อง
- ข. คนรีบเข้าบ้าน
- ค. นกกระจอกเข้ารัง
- ง. กบร้องและนกกระจอกเข้ารัง
- จ. กบร้อง นกกระจอกเข้ารังและคนรีบเข้าบ้าน

22.พ่อบอกลูกว่า ถ้าลูกสอบได้ที่ 1 ถึง 5 พ่อจะซื้อรถจักรยานยนต์ให้ ถ้าลูกสอบได้ที่ 6 ถึง 10 พ่อจะซื้อรถจักรยานให้ ถ้าลูกสอบได้ที่ 11 ถึง 15 พ่อจะซื้อไม้ปิงปองให้ ถ้าลูกสอบตกจะโดนตี 15 ที ปรากฏว่าลูกสอบได้ที่ 14 แต่ได้ 47% ดังนั้นพ่อควรจะทำอย่างไร

- ก. ซื้อไม้ปิงปองให้
- ข. ตี 15 ที
- ค. ซื้อไม้ปิงปองให้และตี 15 ที
- ง. พ่อถามลูกว่าควรได้รับรางวัลหรือตี
- จ. ยังสรุปไม่ได้แน่นอน

23.หยิบลูกหินมา 8 ลูก เป็นสีแดง 3 ลูก สีเขียว 3 ลูก สีขาว 2 ลูก ในการหยิบเมื่อหยิบได้ลูกหินสีแดง ครั้งต่อไปจะได้ลูกหินสีขาวเสมอ หยิบครั้งที่ 3 ได้สีขาว ครั้งที่ 4 และครั้งที่ 5 ได้สีเขียว ครั้งที่ 6 ได้สีแดง ดังนั้น

- ก. ครั้งแรกหยิบได้สีขาว
- ข. ครั้งแรกหยิบได้สีแดง
- ค. ครั้งสุดท้ายหยิบได้สีเขียว
- ง. ครั้งสุดท้ายหยิบได้สีขาว
- จ. ครั้งสุดท้ายหยิบได้สีแดง

24.ถ้าไก่เป็นคนดีแล้วเปิดจะเป็นคนนินทา คนนินทาเป็นคนไม่มีศีล กบนินทาเปิดดังนั้น

- ก. ไก่เป็นคนดี
- ข. เปิดเป็นคนดี
- ค. เปิดเป็นคนนินทา

ง. กบเป็นคนที่ไม่มีศีล

จ. ยังสรุปไม่ได้แน่นอน

25.บ้านเรณูอยู่ทางทิศตะวันตกของบ้านปัญญา

บ้านปรีชาอยู่ทางทิศเหนือของบ้านปัญญา บ้านสุดาอยู่ทางทิศตะวันออกเฉียงเหนือของตลาด และอยู่ทางทิศตะวันตกเฉียงใต้ของบ้านปัญญา ตลาดอยู่ทางทิศตะวันตกเฉียงใต้ของบ้านเรณู ดังนั้น

ก. บ้านสุดาอยู่ทางทิศตะวันออกเฉียงใต้ของบ้านเรณู

ข. บ้านสุดาอยู่ทางทิศใต้ของบ้านปรีชา

ค. บ้านเรณูอยู่ทางทิศตะวันตกเฉียงใต้ของบ้านสุดา

ง. บ้านปัญญาอยู่ทางทิศตะวันตกเฉียงใต้ของบ้านสุดา

จ. บ้านปัญญาอยู่ทางทิศตะวันตกเฉียงเหนือของตลาด

26.ถ้าข้าราชการได้รับการเพิ่มเงินเดือนแล้วสินค้าจะราคาแพงขึ้นทุกครั้ง ถ้าข้าราชการได้รับเงินเดือนเพิ่มขึ้นแล้วประชาชนจะเดือดร้อน แต่สินค้าแพงขึ้นเพราะราคาน้ำมันขึ้น ดังนั้น

ก. น้ำมันขึ้นราคาส่งข้าราชการจะได้รับเงินเดือนขึ้น

ข. ถ้าน้ำมันขึ้นราคาแล้วประชาชนเดือดร้อน

ค. ถ้าสินค้านำเข้าแพงขึ้นประชาชนเดือดร้อน

ง. ถ้าประชาชนไม่เดือดร้อนแสดงว่าข้าราชการไม่ได้รับเงินเดือนขึ้น

จ. ถ้าประชาชนไม่เดือดร้อนแสดงว่าสินค้าไม่เพิ่มราคา

27.นาย ก. กินส้มตำ ไก่ย่าง ปวดท้อง

นาย ข. กินลาบ ไก่ย่าง ไม่ปวดท้อง

นาย ค. กินส้มตำ เนื้อย่าง ปวดท้อง

นาย ง. กินลาบ เนื้อย่าง ไม่ปวดท้อง

ดังนั้นสิ่งไหนทำให้ปวดท้อง

ก. ลาบ

ข. ส้มตำ

ค. เนื้อย่าง

ง. ไก่ย่าง

จ. ยังสรุปไม่ได้แน่นอน

28.เด็กชอบอ่านหนังสือการ์ตูน ผู้ใหญ่ชอบอ่านหนังสือนิยาย สดากำลังอ่านหนังสือนิยาย ดังนั้น

ก. สดากเป็นผู้ใหญ่

ข. สดากชอบอ่านหนังสือนิยาย

ค. สดากไม่ชอบอ่านการ์ตูน

ง. สดากชอบทั้งการ์ตูนและนิยาย

จ. ยังสรุปไม่ได้แน่นอน

29.เด่นบอกแม่ว่า อยากกินข้าวต้มปลา แต่ไม่ให้ใส่ข้าวต้ม ดังนั้น

ก. เต้นหิวข้าว

ข. เต้นอยากกินปลา

ค. แม่ของเด่นขายข้าวต้ม

ง. เต้นอยากกินข้าวต้มปลา

จ. ยังสรุปไม่ได้แน่นอน

30.นางงามทุกคนสวย คนสวยส่วนมากผิวขาว คนผิวขาวบางคนกลัวแดด ดังนั้น

ก. นางงามทุกคนผิวขาว

ข. คนสวยทุกคนผิวขาว

ค. คนสวยส่วนมากกลัวแดด

ง. นางงามบางคนกลัวแดด

จ. ยังสรุปไม่ได้แน่นอน

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ	นางสาวสิรินุช เอี่ยมเขียว
วันเดือนปีเกิด	21 ตุลาคม 2520
สถานที่เกิด	อำเภอเมือง จังหวัดอุดรธานี
สถานที่อยู่ปัจจุบัน	553/121 ถ. เดชะตุงคะ แขวงสีกัน เขตดอนเมือง กรุงเทพมหานคร 10210
สถานที่ทำงาน	-
ประวัติการศึกษา	
พ.ศ.2537	ชั้นมัธยมศึกษาปีที่ 6 โรงเรียนอุดรพิทยานุกูล
พ.ศ.2541	คบ. (วิชาเอกคณิตศาสตร์) สถาบันราชภัฏพระนคร
พ.ศ.2544	กศ.ม. (การวิจัยและสถิติทางการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร