

## การทดสอบภายหลังการวิเคราะห์ความแปรปรวน

## Post hoc analysis of variance test

อุทัยวรรณ สายพัฒนา (Uthaiwan Saipattana) \*

Received: June 5, 2024

Revised: June 18, 2024

Accepted: July 20, 2024

\* Corresponding author, Email: uthaiwan.sa@gmail.com

## บทคัดย่อ

บทความวิชาการนี้มีวัตถุประสงค์ เพื่อนำเสนอสาระเกี่ยวกับวิธีเปรียบเทียบรายคู่ หรือการเปรียบเทียบพหุคูณ (Multiple comparison) สำหรับงานวิจัยที่มีการศึกษาประชากรมากกว่า 2 กลุ่ม บางครั้งก็เรียกว่า Post hoc comparisons เพราะเป็นการตรวจสอบหลังจากการวิเคราะห์ความแปรปรวน (ANOVA: Analysis of variance) ด้วยสถิติทดสอบ F หรือ F-test แล้วพบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างโดยรวม (Omnibus test) มีความแตกต่างกัน จึงต้องทำการตรวจสอบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างคู่ใดบ้างที่แตกต่างกัน ซึ่งวิธีการตรวจสอบโดยการเปรียบเทียบเป็นรายคู่มีหลายวิธี แต่ละวิธีมีข้อดีข้อด้อย มีวัตถุประสงค์และคุณลักษณะเฉพาะในการเปรียบเทียบแตกต่างกันออกไป ทั้งนี้ผู้วิจัยควรพิจารณาเลือกใช้วิธีการเปรียบเทียบให้สอดคล้องและเหมาะสมกับลักษณะของข้อมูลและลักษณะของกลุ่มตัวอย่าง เพื่อให้ได้ผลของการตรวจสอบสมมติฐานที่มีความคลาดเคลื่อนน้อยที่สุด และตอบปัญหาการวิจัยได้ชัดเจนมากที่สุด ในที่นี้ผู้เขียนนำเสนอวิธีการเปรียบเทียบพหุคูณที่นักวิจัยส่วนใหญ่ นิยมนำมาใช้ในการตรวจสอบความแตกต่าง 3 วิธี คือ วิธี LSD วิธี Tukey-Kramer และวิธี Scheffe' และยกตัวอย่างขั้นตอนวิธีการวิเคราะห์จากโปรแกรม SPSS ด้วยวิธี Scheffe' และการนำเสนอการแปลผลการวิจัย

**คำสำคัญ:** การเปรียบเทียบเป็นรายคู่, วิธี LSD, วิธี Tukey-Kramer, วิธี Scheffe'

## Abstract

This academic article aimed to present information about how to compare pairs or multiple comparisons (Multiple Comparison), sometimes called Post Hoc Comparison because it is checked after the analysis of variance (ANOVA) with the F test statistic (F-test) when it was found that the average of the overall sample (Omnibus test) was different. Therefore, it must be checked that which pairs of sample means are different? There are many ways to check by comparing pairs. Each method has its own terms. There are different objectives and specific characteristics of comparison. The researcher should consider selecting the comparison method accordingly and appropriate to the nature of the data and the characteristics of the sample group, in order to get the results of the inspection with the least deviation and answer research questions as clearly as possible. Here, the author would like to mention 3 methods of multiple comparison that most researchers like to use to check the differences: LSD method, Tukey-Kramer method, and Scheffe' method. Also, give an example of the steps, analysis methods from the SPSS program using the Scheffe' method, and present the interpretation of the research results.

**Keywords:** Multiple comparison, LSD method, Tukey-Kramer method, Scheffe' method

\* อาจารย์ประจำ คณะศึกษาศาสตร์ มหาวิทยาลัยนอร์ทกรุงเทพ  
Lecturer, Faculty of Education, North Bangkok University

## บทนำ

ในการวิจัยเชิงเปรียบเทียบที่ต้องการทราบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างกลุ่มใดมีความแตกต่างจากกลุ่มอื่นหรือไม่ กรณีที่มีการเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่าง 2 กลุ่ม ก็นิยมใช้สถิติ Z-test หรือ t-test ซึ่งการเลือกใช้สถิติทดสอบแบบใดขึ้นอยู่กับความแปรปรวนของข้อมูลในประชากร และขนาดของกลุ่มตัวอย่างนั้น สำหรับกรณีที่ต้องการเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่างที่มีมากกว่า 2 กลุ่ม หรือตั้งแต่ 3 กลุ่มขึ้นไป และต้องการทดสอบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างแต่ละกลุ่มมีความแตกต่างกันหรือไม่ ถ้าใช้การทดสอบทีละคู่จะทำให้ไม่สะดวก เสียเวลาในการทดสอบ และที่สำคัญทำให้เกิดความคลาดเคลื่อนประเภทที่ 1 (Type I error) หรือเรียกว่า Pairwise error: PE) สูงเกินไป เช่น ในการทดสอบค่าเฉลี่ยของประชากร 4 กลุ่ม ถ้าใช้ t-test ทดสอบค่าเฉลี่ยทีละคู่ จำนวน 6 ครั้ง คือ

ครั้งที่ 1 ทำการเปรียบเทียบค่าเฉลี่ยของกลุ่มที่ 1 กับค่าเฉลี่ยของกลุ่มที่ 2

ครั้งที่ 2 ทำการเปรียบเทียบค่าเฉลี่ยของกลุ่มที่ 1 กับค่าเฉลี่ยของกลุ่มที่ 3

ครั้งที่ 3 ทำการเปรียบเทียบค่าเฉลี่ยของกลุ่มที่ 1 กับค่าเฉลี่ยของกลุ่มที่ 4

ครั้งที่ 4 ทำการเปรียบเทียบค่าเฉลี่ยของกลุ่มที่ 2 กับค่าเฉลี่ยของกลุ่มที่ 3

ครั้งที่ 5 ทำการเปรียบเทียบค่าเฉลี่ยของกลุ่มที่ 2 กับค่าเฉลี่ยของกลุ่มที่ 4

และ ครั้งที่ 6 ทำการเปรียบเทียบค่าเฉลี่ยของกลุ่มที่ 3 กับค่าเฉลี่ยของกลุ่มที่ 4

การทดสอบแต่ละครั้ง เราจะต้องกำหนดให้ระดับนัยสำคัญเป็น .05 ( $\alpha = .05$  คือมีความเชื่อมั่น 95%) เมื่อต้องการทดสอบทีละคู่ จำนวน 6 ครั้ง ระดับนัยสำคัญจะมากกว่า .05 ที่กำหนดไว้ หรือมีความเชื่อมั่นลดลง นั่นคือ การทดสอบค่าเฉลี่ยมากกว่า 1 คู่ จะมีค่าความคลาดเคลื่อนที่เกิดขึ้นทั้งหมด เรียกว่า Familywise Error: FE

$$FE = 1 - (1 - \alpha)^c$$

เมื่อ c เป็นจำนวนครั้งของการทดสอบโดยใช้ t-test และ  $\alpha$  เป็นระดับนัยสำคัญของการทดสอบ t-test แต่ละครั้ง

ถ้ามีกลุ่มตัวอย่าง 4 กลุ่ม เมื่อใช้ t-test ทดสอบทีละคู่ จำนวน 6 ครั้ง และกำหนดระดับนัยสำคัญไว้ที่ .05 นั่นคือ จะเกิดความคลาดเคลื่อนประเภทที่ 1 เท่ากับ  $1 - (1 - .05)^6 = .19$  ซึ่งค่ามากกว่า .05 ดังนั้นจึงมีการนำเทคนิค การวิเคราะห์ความแปรปรวน (Analysis of variance: ANOVA) มาใช้ในการทดสอบเกี่ยวกับความแตกต่างของค่าเฉลี่ยของกลุ่มตัวอย่างที่มากกว่า 2 กลุ่ม หรือตั้งแต่ 3 กลุ่มขึ้นไป (k กลุ่ม โดยที่  $k \geq 3$ ) โดยทำการทดสอบเพียงครั้งเดียวด้วยสถิติการทดสอบรวม (Omnibus test) ที่นิยมเรียกกันทั่วไปว่า F-test เป็นการทดสอบว่าค่าเฉลี่ยของกลุ่มตัวอย่างทุกกลุ่มในการวิจัยแตกต่างกันอย่างมีนัยสำคัญหรือไม่ ซึ่งการวิเคราะห์ความแปรปรวนเป็นเทคนิควิธีที่ดี เพราะไม่ต้องเสียเวลามาทดสอบหลายครั้ง และมีความคลาดเคลื่อนประเภทที่ 1 เท่าเดิมตามที่กำหนดไว้

## การวิเคราะห์ความแปรปรวน

การวิเคราะห์ความแปรปรวน (Analysis of variance: ANOVA) โดยใช้ F-test ที่เรียกว่า A posteriori multiple comparisons เป็นการวิเคราะห์เพื่อศึกษาอิทธิพลของตัวแปรอิสระ (Independent variable) หรือตัวแปรจัดกระทำ (Treatment variable) ที่มีต่อตัวแปรตาม (Dependent variable) ถ้าการวิเคราะห์ความแปรปรวนที่มีตัวแปรอิสระ 1 ตัวและตัวแปรตาม 1 ตัว เรียกว่า การวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA) ถ้ามีตัวแปรอิสระ 2 ตัวและตัวแปรตาม 1 ตัว เรียกว่า การวิเคราะห์ความแปรปรวนสองทาง (Two-way ANOVA) ในบทความนี้จะกล่าวเฉพาะการวิเคราะห์ความแปรปรวนทางเดียว

## การวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA)

การวิเคราะห์ความแปรปรวนทางเดียวเป็นการศึกษาอิทธิพลของตัวแปรอิสระ 1 ตัว ว่ามีผลต่อตัวแปรตามอย่างไร โดยตัวแปรอิสระมีหลายระดับ (k ระดับ) หรือเป็นการทดสอบว่า ค่าเฉลี่ยของกลุ่มตัวอย่าง k กลุ่มแตกต่างกันหรือไม่ ยกตัวอย่างเช่น การเปรียบเทียบความคิดเห็นของครูที่มีต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียน โดยสนใจศึกษาตัวแปรอิสระ คือ ระดับการศึกษา ที่ต่างกัน 3 ระดับ คือ ครูที่มีการศึกษาระดับปริญญาตรี ปริญญาโท และปริญญาเอก เป็นต้น

ซึ่งก่อนที่จะทำการวิเคราะห์ความแปรปรวน จะต้องพิจารณาข้อตกลงเบื้องต้น (Assumption) ของการวิเคราะห์ก่อนว่า ข้อมูลที่นำมาวิเคราะห์นั้นมีลักษณะเป็นไปตามข้อตกลงเบื้องต้นหรือไม่

ข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวน ทางเดี่ยว (Clinch & Keselman, 1982, pp.42-43; Cochran & Cox, 1976, pp.104-106) มีดังนี้

1. กลุ่มตัวอย่างได้มาจากการสุ่มอย่างเป็นอิสระ
2. ข้อมูลของกลุ่มตัวอย่างมีการแจกแจงเป็นโค้งปกติ (Sample normal distribution)
3. คุณลักษณะของตัวแปรตาม (Dependent Variable characteristic) จะต้องมีความเป็นตัวแปรต่อเนื่องหรืออยู่ในมาตราวัดระดับอันตรภาคชั้น (Interval scale) ขึ้นไป
4. กลุ่มตัวอย่างแต่ละกลุ่มที่ใช้ในการทดสอบต้องเป็นอิสระจากกัน
5. ความแปรปรวนระหว่างกลุ่มตัวอย่าง (Heterogeneity between sample groups) แต่ละกลุ่มเท่ากัน

เทคนิคการวิเคราะห์ความแปรปรวน เป็นการทดสอบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างจำนวน  $k$  กลุ่มนั้น มีความแตกต่างกันหรือไม่ ถ้าการทดสอบพบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (significant) ก็เป็นการบอกเพียงว่ามีค่าเฉลี่ยของกลุ่มตัวอย่างอย่างน้อย 1 คู่ที่มีความแตกต่างกัน แต่จะไม่แสดงให้เห็นว่าคู่ใดที่แตกต่างกัน ดังนั้น ผู้วิจัยจะต้องทำการทดสอบภายหลัง (Post hoc tests หรือ Post hoc comparisons) การวิเคราะห์อีก เพื่อจะทดสอบว่าค่าเฉลี่ยของกลุ่มใดหรือคู่ใดบ้างที่มีความแตกต่างกันอย่างมีนัยสำคัญ เป็นการตรวจสอบในรายละเอียด ซึ่งวิธีการทดสอบมีหลายวิธี แต่ละวิธีมีวัตถุประสงค์และคุณสมบัติเฉพาะในการเปรียบเทียบแตกต่างกันออกไป ผู้วิจัยจึงต้องเลือกใช้ให้เหมาะสม เพื่อให้ผลการทดสอบมีความคลาดเคลื่อนน้อยที่สุด และตอบปัญหาการวิจัยได้ชัดเจนที่สุด

วิธีการเปรียบเทียบความแตกต่างของค่าเฉลี่ยภายหลังการทดสอบรวม ที่พบว่ามีนัยสำคัญทางสถิติได้แก่

- วิธี Least -significant different test: LSD
- วิธี Fisher least-significant different test: FLSD

- วิธี Duncan's new multiple range test: DC หรือ DMRT

- วิธี Student Newman-Keuls test or Newman-Keuls multiple range test: S-N-K

- วิธี Tukey's honestly significance difference test: Tukey A หรือ TA หรือ HSD

- วิธี Tukey's alternate test: Tukey B หรือ TB

- วิธี Scheffe' test: Scheffe' หรือ S

- วิธี Dunnett's test: DN

- วิธี Stepwise multiple comparison test: SMC

- วิธี Bayes significance difference test: BSD หรือ Bayes LSD

- วิธี Bayes exact test: BET

- วิธี Bonferroni t statistic และ

- วิธี Williams' method เป็นต้น

ซึ่งวิธีการแต่ละวิธีนั้น จะต้องมีความเหมาะสมกับลักษณะของข้อมูลและลักษณะของกลุ่มตัวอย่าง ดังนั้น ผู้วิจัยจะต้องพิจารณาปัจจัยที่สำคัญประกอบการเลือกใช้วิธีการเปรียบเทียบต่าง ๆ ให้สอดคล้อง และเหมาะสมด้วย ซึ่งปัจจัยที่สำคัญนั้น ได้แก่

### 1. ลักษณะความคลาดเคลื่อนทางสถิติ

ความคลาดเคลื่อนทางสถิติในการวิเคราะห์ข้อมูลมี 3 ประเภท คือ

1.1 ความคลาดเคลื่อนประเภทที่ 1 (Type I error หรือ  $\alpha$  error) เป็นความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมติฐานที่เป็นกลาง (null hypothesis:  $H_0$ ) ซึ่งความเป็นจริงนั้น null hypothesis มีความถูกต้อง (Probability of rejecting a true null hypothesis)

1.2 ความคลาดเคลื่อนประเภทที่ 2 (Type II error หรือ  $\beta$  error) เป็นความคลาดเคลื่อนที่เกิดจากการยอมรับสมมติฐานที่เป็นกลาง (null hypothesis:  $H_0$ ) ทั้ง ๆ ที่ความเป็นจริงนั้น null hypothesis ไม่ถูกต้อง (Probability of failing to rejecting a false null hypothesis)

1.3 ความคลาดเคลื่อนประเภทที่ 3 (Type III error หรือ  $\gamma$  error) เป็นความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมติฐานที่เป็นกลาง (null hypothesis:  $H_0$ ) และผลของการยอมรับสมมติฐานทางเลือก (alternative hypothesis) นั้นขัดแย้งกันเอง ในลักษณะตรงข้ามกับความเป็นจริง (Probability of declaring one treatment superior to another when the reverse is actually true)

ในทางปฏิบัติของการวิเคราะห์ข้อมูลแล้ว โอกาสที่จะเกิดความคลาดเคลื่อนประเภทที่ 3 มีน้อยมาก แต่มักจะเกิดความคลาดเคลื่อนประเภทที่ 1 และ 2 ซึ่งนักสถิติมักจะกำหนดระดับนัยสำคัญเพื่อป้องกันการเกิดความคลาดเคลื่อนประเภทที่ 1 มากกว่าป้องกันการเกิดความคลาดเคลื่อนประเภทที่ 2

## 2. ลักษณะการเปรียบเทียบของแต่ละวิธี

การเปรียบเทียบแต่ละวิธีมีลักษณะเฉพาะ ผู้วิจัยจะต้องรู้และทำความเข้าใจลักษณะเฉพาะเหล่านั้น เพื่อจะได้เลือกใช้ได้อย่างถูกต้อง และเหมาะสมกับเรื่องที่ตนศึกษา ซึ่งแบ่งเป็น 2 กลุ่ม ดังนี้

2.1 กลุ่มที่ใช้ค่าพิสัยค่าเดียว (The fixed range test) เป็นกลุ่มที่ใช้ค่าพิสัยค่าเดียวในการเปรียบเทียบทั้งหมด สถิติเปรียบเทียบในกลุ่มนี้ ได้แก่ วิธี LSD วิธี FLSD วิธี Tukey (HSD) วิธี Scheffe' และวิธี Dunnett

2.2 กลุ่มที่ใช้ค่าพิสัยหลายค่า (The multiple range test) เป็นกลุ่มที่ใช้ค่าพิสัยหลายค่าตามจำนวนครั้งของการเปรียบเทียบ หรือจำนวนชั้นของระยะห่างระหว่างค่าเฉลี่ยของกลุ่มที่เปรียบเทียบกัน สถิติเปรียบเทียบในกลุ่มนี้ ได้แก่ วิธี S-N-K, วิธี DMRT และวิธี Tukey's alternate test: Tukey B

## 3. ขนาดของกลุ่มตัวอย่างในแต่ละกลุ่มที่เปรียบเทียบ

วิธีการเปรียบเทียบเกือบทุกวิธี ใช้สำหรับเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่างที่มีขนาดเท่ากัน ยกเว้นวิธี Scheffe' ซึ่งในทางปฏิบัติ ผู้วิจัยส่วนมากมักใช้ขนาดของกลุ่มตัวอย่างไม่เท่ากัน ด้วยเหตุผลที่ว่า ในระหว่างการทดลองนั้น จำนวนของกลุ่มตัวอย่างอาจมีจำนวนลดลง หรือเรื่องวิธีการสุ่มกลุ่มตัวอย่าง หรือเหตุผลใดก็ตาม

จากวิธีการเปรียบเทียบหรือวิธีการทดสอบความแตกต่างของค่าเฉลี่ยของกลุ่มตัวอย่างที่กล่าวมาข้างต้นนั้น มีวิธีการเพียงบางวิธีเท่านั้น ที่นักวิจัยส่วนใหญ่นิยมนำมาใช้ในการทดสอบความแตกต่าง ในที่นี้จะขอกล่าวถึงเพียง 3 วิธีเท่านั้น คือ วิธี LSD วิธี Tukey-Kramer และวิธี Scheffe' ซึ่งมีรายละเอียดของแต่ละวิธีดังนี้

## 1. วิธี Least-significant different (LSD)

วิธีการเปรียบเทียบพหุคูณแบบ LSD หรือ Fisher's least – significant different เป็นเทคนิคที่ R.A. Fisher (1959) เป็นผู้พัฒนาขึ้น เป็นการเปรียบเทียบค่าเฉลี่ยประชากรครั้งละหลายคู่ โดยต้องสุ่มตัวอย่างแต่ละชุดเป็นอิสระกัน และควรเป็นกรณีที่มีการวางแผนไว้ก่อนที่จะเก็บข้อมูล วิธีนี้บางครั้งเรียกว่า Fisher's LSD โดยใช้สูตร

$$LSD = t_{\alpha/2; n-k} \sqrt{MSE} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

$$\text{โดยที่ } n = \sum_{i=1}^k n_i$$

เมื่อ LSD แทน สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธี Fisher's LSD

MSE แทน ค่าประมาณของความแปรปรวนของความคลาดเคลื่อน

k แทน จำนวนกลุ่มตัวอย่าง

n แทน จำนวนตัวอย่างทั้งหมด

$n_i$  แทน จำนวนข้อมูลกลุ่มที่ i

$n_j$  แทน จำนวนข้อมูลกลุ่มที่ j

การทดสอบของวิธีการ Fisher's LSD มีขั้นตอนดังนี้

ขั้นที่ 1 ทำการทดสอบสมมติฐาน

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

ถ้าปฏิเสธ  $H_0$  ทำขั้นที่ 2 ต่อ

ขั้นที่ 2 หาค่าวิกฤต LSD

ขั้นที่ 3 หาค่า

$$|\bar{X}_i - \bar{X}_j| \text{ ทุกค่า } i \neq j; i, j = 1, 2, \dots, k$$

ขั้นที่ 4 การตัดสินใจ

ถ้า  $|\bar{X}_i - \bar{X}_j| \geq LSD$  แสดงว่า  $\mu_i \neq \mu_j$  หรือค่าเฉลี่ยของกลุ่มตัวอย่างมีความแตกต่างกัน

## 2. วิธี Turkey – Kramer

การทดสอบความแตกต่างของค่าเฉลี่ยตามวิธี Turkey-Kramer (Kirk. 1995, pp.144-155) นี้ มีความแตกต่างจากวิธีการของ Tukey (1953) และ Kramer (1956) สำหรับกรณีที่กลุ่มตัวอย่างมีขนาดไม่เท่ากัน และมีข้อตกลงเบื้องต้นที่เป็นปกติ มีความแปรปรวน

เท่ากัน วิธีการทดสอบทางสถิติของวิธีนี้ แบ่งเป็น 2 กรณี ดังนี้

2.1 กรณีที่กลุ่มตัวอย่างมีขนาดเท่ากัน (Equal sample size) กรณีที่  $n_1 = n_2 = \dots = n_k = n_c$  โดยที่ค่าวิกฤตหาได้จากสูตร

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

เมื่อ  $T$  แทน สถิติทดสอบความแตกต่างของ

ค่าเฉลี่ยตามวิธี Tukey-Kramer

$q_\alpha$  แทน ค่าวิกฤตของ Studentized Range

MSE แทน ค่าประมาณของความแปรปรวนของ

ความคลาดเคลื่อน

$k$  แทน จำนวนกลุ่มตัวอย่าง

$n$  แทน ขนาดของกลุ่มตัวอย่าง

$n_c$  แทน จำนวนสมาชิกในกลุ่มตัวอย่างของ

กลุ่มที่  $c$

2.2 กรณีที่กลุ่มตัวอย่างมีขนาดไม่เท่ากัน (Unequal sample size) กรณีที่  $n_1 \neq n_2$  อย่างน้อย 1 คู่ โดยที่ค่าวิกฤตหาได้จากสูตร

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

เมื่อ  $T$  แทน สถิติทดสอบความแตกต่างของ

ค่าเฉลี่ยตามวิธี Tukey-Kramer

$q_\alpha$  แทน ค่าวิกฤตของ Studentized Range

MSE แทน ค่าประมาณของความแปรปรวนของ

ความคลาดเคลื่อน

$k$  แทน จำนวนกลุ่มตัวอย่าง

$n$  แทน ขนาดของกลุ่มตัวอย่าง

$n_c$  แทน จำนวนสมาชิกในกลุ่มตัวอย่างของ

กลุ่มที่  $c$

ถ้าการทดสอบสมมติฐาน  $H_0: \mu_i = \mu_j$  ถูกปฏิเสธเปรียบเทียบค่า  $|\bar{X}_i - \bar{X}_j|$  กับค่า  $T$  ที่ได้

ถ้า  $|\bar{X}_i - \bar{X}_j|$  มากกว่าหรือเท่ากับค่าวิกฤต  $T$  จะสรุปได้ว่า  $\mu_i \neq \mu_j$  ที่ระดับนัยสำคัญ  $\alpha$

ขั้นตอนการทดสอบของวิธี Tukey-Kramer มีดังนี้

ขั้นที่ 1 ทำการทดสอบสมมติฐาน

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

ถ้าปฏิเสธ  $H_0$  ทำขั้นที่ 2 ต่อ

ขั้นที่ 2 เรียงลำดับค่าเฉลี่ยจากน้อยไปมาก

ขั้นที่ 3 หาค่าวิกฤต  $T$

ขั้นที่ 4 หาค่า  $|\bar{X}_i - \bar{X}_j|$  ทุกค่า  $i \neq j$ ;

$$i, j = 1, 2, \dots, k$$

ขั้นที่ 5 ถ้า  $|\bar{X}_i - \bar{X}_j| \geq T$  แสดงว่า  $\mu_i \neq \mu_j$  หรือค่าเฉลี่ยของกลุ่มตัวอย่างมีความแตกต่างกัน

### 3. วิธี The Scheffe's post hoc comparison (Scheffe')

วิธีการเปรียบเทียบพหุคูณแบบ Scheffe' เป็นเทคนิคที่พัฒนาขึ้นโดย Scheffe (1953) การทดสอบการควบคุมความคลาดเคลื่อนประเภทที่ 1 ที่น้อยกว่า  $\alpha$  สำหรับการเปรียบเทียบจำนวนที่ไม่จำกัดระหว่างค่าเฉลี่ย  $p \geq 3$  วิธีของ Scheffe' จะเป็นวิธีการที่ให้ผลการทดสอบไม่ต่างกับวิธี LSD แต่วิธี LSD มีอำนาจการทดสอบสูงกว่าวิธีการของ Scheffe' เมื่อกลุ่มตัวอย่างมีขนาดเท่ากัน (นิภาพร ขำสะอาด, 2552) และเมื่อเป็นการเปรียบเทียบที่ไม่ใช่รายคู่ วิธีการของ Scheffe' จะใช้การกระจายกลุ่มตัวอย่างแบบ F ที่เหมือนกับการวิเคราะห์ความแปรปรวน และมีความแกร่ง และเป็นวิธีการเปรียบเทียบค่าเฉลี่ยที่สามารถใช้ได้กับกลุ่มตัวอย่างที่มีขนาดเท่ากันหรือไม่เท่ากันก็ได้ วิธีการทดสอบของ Scheffe' มีสูตร ดังนี้

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

ในกรณีที่เปรียบเทียบ  $\bar{X}_1$  กับ  $\bar{X}_2$

เมื่อ  $F$  แทน สถิติทดสอบความแตกต่างของค่าเฉลี่ยวิธีของ Scheffe' โดยที่  $df_1 = k-1$ ,  $df_2 = n-k$

$k$  แทน จำนวนกลุ่มตัวอย่าง

$n$  แทน จำนวนตัวอย่างทั้งหมด

$n_1, n_2$  แทน จำนวนกลุ่มตัวอย่างของกลุ่มที่ 1 และกลุ่มที่ 2

วิธีการเปรียบเทียบของ Scheffe' มีขั้นตอนการทดสอบ ดังนี้

ขั้นที่ 1 ทำการทดสอบสมมติฐาน

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

ถ้าปฏิเสธ  $H_0$  ทำขั้นที่ 2 ต่อ

ขั้นที่ 2 สร้าง 1-way ANOVA ได้ค่า  $S_2 = MSE$

ขั้นที่ 3 หาค่าวิกฤต  $F$

ขั้นที่ 4 หาค่า  $|\bar{X}_i - \bar{X}_j|$  ทุกค่า  $i \neq j$ ;  
 $i, j = 1, 2, \dots, k$

ขั้นที่ 5 ถ้า  $|\bar{X}_i - \bar{X}_j| \geq F$  แสดงว่า  $\mu_i \neq \mu_j$   
 หรือค่าเฉลี่ยของกลุ่มตัวอย่างมีความแตกต่างกัน

อย่างไรก็ตาม การเปรียบเทียบพหุคูณแต่ละวิธีที่กล่าวมานั้น ปัจจุบันมีโปรแกรมที่เอื้อให้มีความสะดวกในการวิเคราะห์ เช่น โปรแกรม SPSS SAS MINITAB หรือ STATISTICA เป็นต้น ในบทความนี้จะขอยกตัวอย่างแนวทางการวิเคราะห์โดยใช้โปรแกรม SPSS ในการตรวจสอบความแตกต่างของค่าเฉลี่ยด้วยวิธี Scheffe' ดังนี้

ตัวอย่างงานวิจัยที่มีการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของข้อมูลที่มีมากกว่า 2 กลุ่ม

**วัตถุประสงค์ของการวิจัย:** เพื่อเปรียบเทียบความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียนของครูที่มีระดับการศึกษาแตกต่างกัน คือ ระดับปริญญาตรี ปริญญาโท และปริญญาเอก เป็นต้น

**สมมติฐานการวิจัย:** ครูที่มีระดับการศึกษาแตกต่างกันมีความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาล แตกต่างกันได้ เริ่มใช้โปรแกรม SPSS เพื่อการคำนวณโดยเปิดโปรแกรม SPSS แล้วบันทึกข้อมูลที่เก็บรวบรวมมา ดังภาพประกอบ

ตัวอย่าง: การนำเสนอผลการวิเคราะห์ข้อมูลตามวัตถุประสงค์ของการวิจัย

	ระดับการศึกษา	a1	a2	a3	a4	a5	a6	a7	a8	A_mean
1	2	5	4	5	4	5	2	5	3	4.13
2	2	4	4	4	3	4	3	5	3	3.75
3	1	4	3	4	5	4	4	4	5	4.13
4	1	5	4	5	4	5	4	4	5	4.50
5	3	5	4	5	4	5	4	4	4	4.38
6	3	4	4	4	4	4	4	4	4	4.00
7	1	5	4	5	4	5	4	4	5	4.50
8	2	4	4	3	4	4	4	4	4	3.88
9	1	4	5	4	5	5	5	4	5	4.63
10	3	3	4	5	4	5	4	5	2	4.00
11	3	5	4	5	4	5	4	4	2	4.13
12	2	4	4	4	4	4	4	4	3	3.88

ตัวแปรต้น คือ ระดับการศึกษา  
 มี 3 กลุ่ม คือ ปริญญาตรี  
 ปริญญาโท และปริญญาเอก

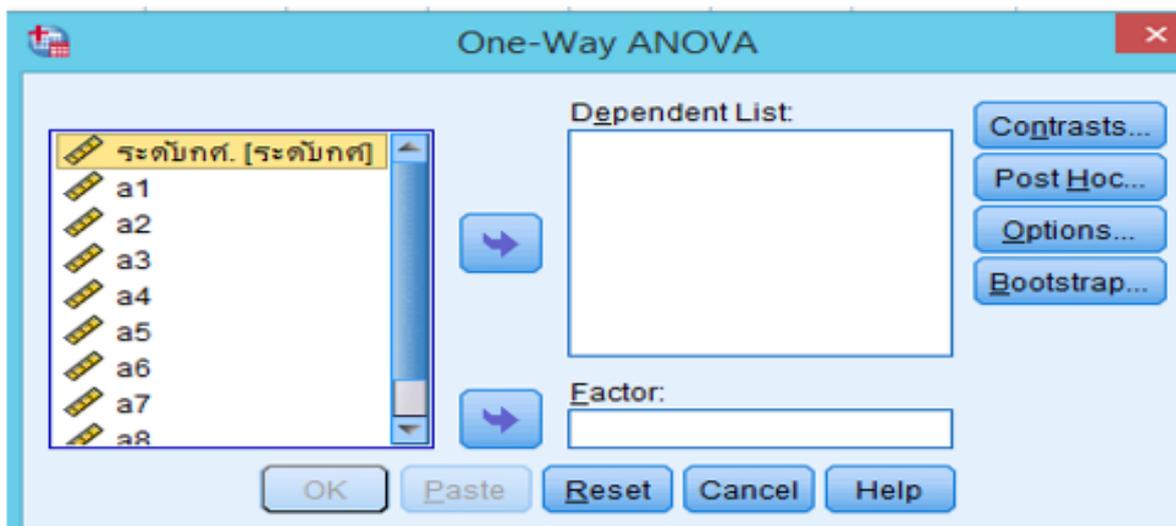
ตัวแปรตาม คือ ค่าเฉลี่ยของระดับความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาล

ภาพประกอบ 1 ไฟล์ข้อมูลที่ต้องเตรียม

เริ่มวิเคราะห์โดยคลิกเมนู Analyze เลือกคำสั่ง Compare means แล้วเลือก One-way ANOVA ดังภาพประกอบ 2

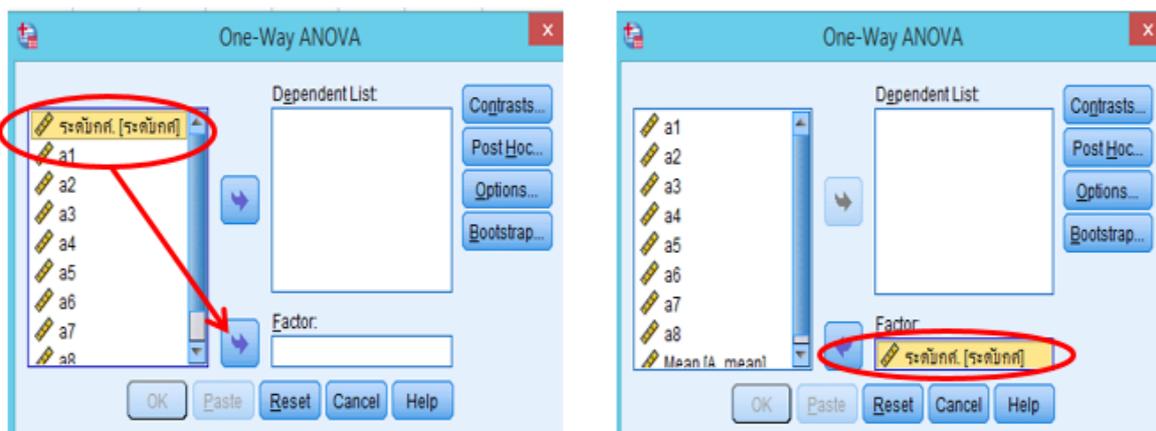
ภาพประกอบ 2 การวิเคราะห์ One-way ANOVA

จะได้หน้าต่าง One-way ANOVA ดังภาพประกอบ 3



ภาพประกอบ 3 การนำข้อมูลเข้า

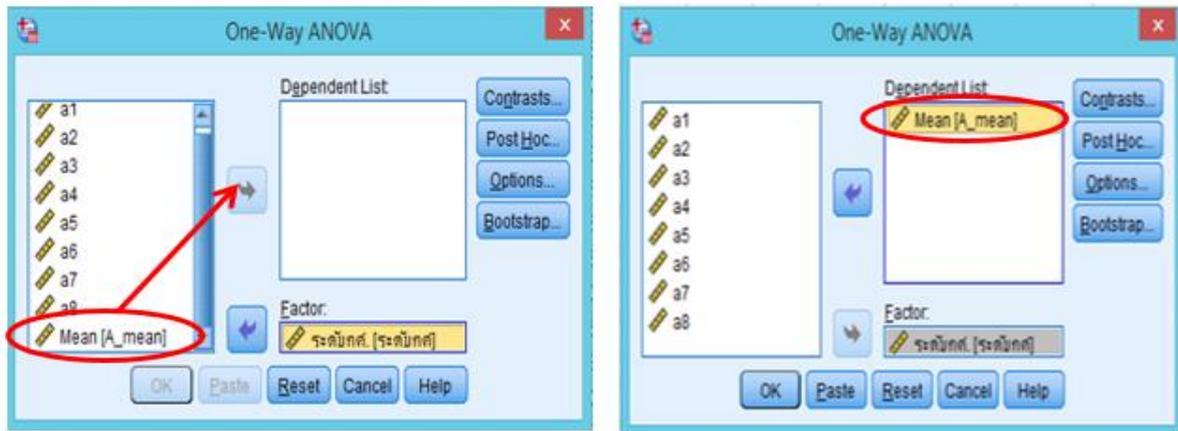
คลิกเลือกตัวแปรต้น คือ ระดับการศึกษา จากช่องซ้าย เพื่อนำตัวแปรนี้ไปใส่ในช่อง Factor ด้านขวามือ ดังภาพประกอบ 4



ภาพประกอบ 4 กำหนดค่าตัวแปร

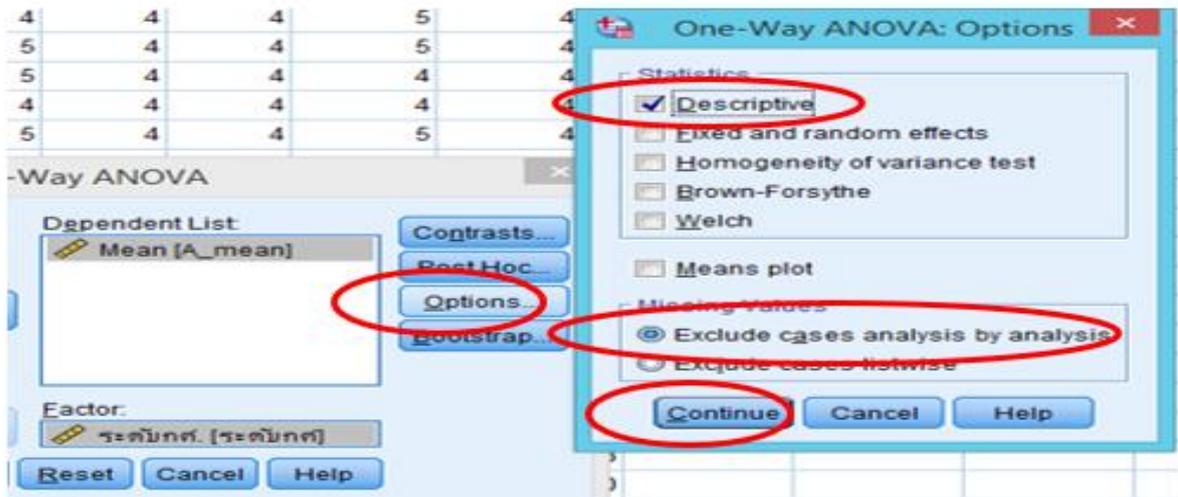
หมายเหตุ: ในกรณีที่ข้อมูลที่มียังไม่ได้มีการแบ่งกลุ่ม ผู้วิจัยจะต้องทำการแบ่งกลุ่มของข้อมูลให้เรียบร้อยก่อนที่จะนำมาวิเคราะห์โดยวิธีนี้

หลังจากนั้นให้คลิกเลือกตัวแปรตาม ได้แก่ ค่าเฉลี่ยของระดับความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาล (Mean คือ ค่าเฉลี่ยของข้อคำถามข้อ 1 ถึงข้อ 8) จากช่องซ้าย เพื่อนำตัวแปรนี้ไปด้านขวามือ (Dependent list) ดังภาพประกอบ 5



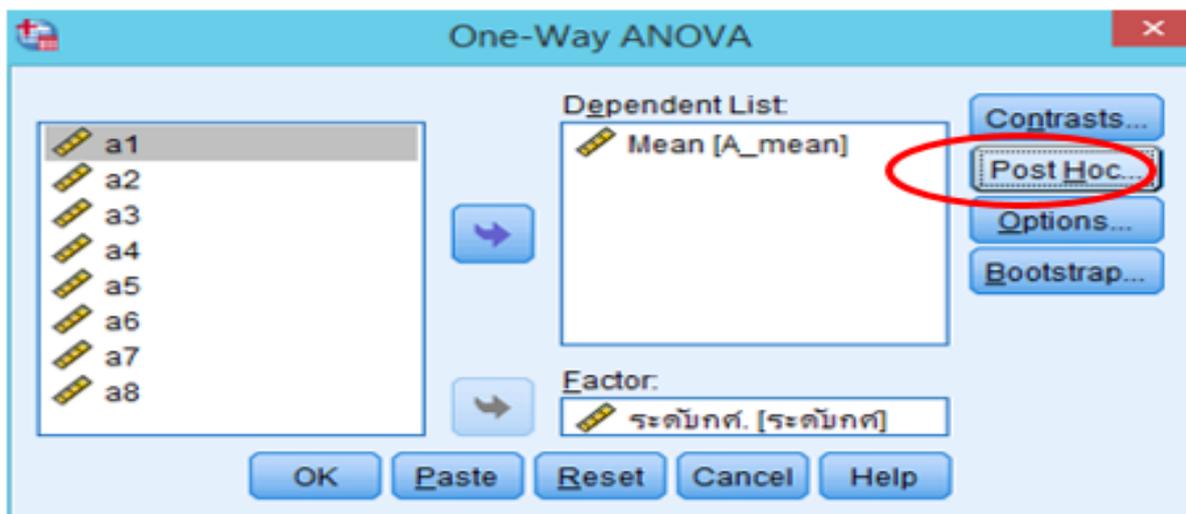
ภาพประกอบ 5 กำหนดค่าตัวแปร

หลังจากนั้นที่หน้าต่าง One-way ANOVA ให้คลิกที่ปุ่ม Options จะขึ้นหน้าต่าง Options ให้คลิกเลือก Descriptive (ในกรอบ Statistics) และ เลือก Exclude cases analysis by analysis (ในกรอบ Missing values) แล้วคลิกปุ่ม Continue จะได้ภาพที่เป็นหน้าต่าง One-way ANOVA อีกครั้ง ดังภาพประกอบ 6



ภาพประกอบ 6 กำหนดค่า Descriptive

ให้คลิก Post hoc เพื่อเลือกสถิติในการวิเคราะห์ค่าความแตกต่างเป็นรายคู่ ดังภาพประกอบ 7



ภาพประกอบ 7 วิเคราะห์ค่าความแตกต่างเป็นรายคู่

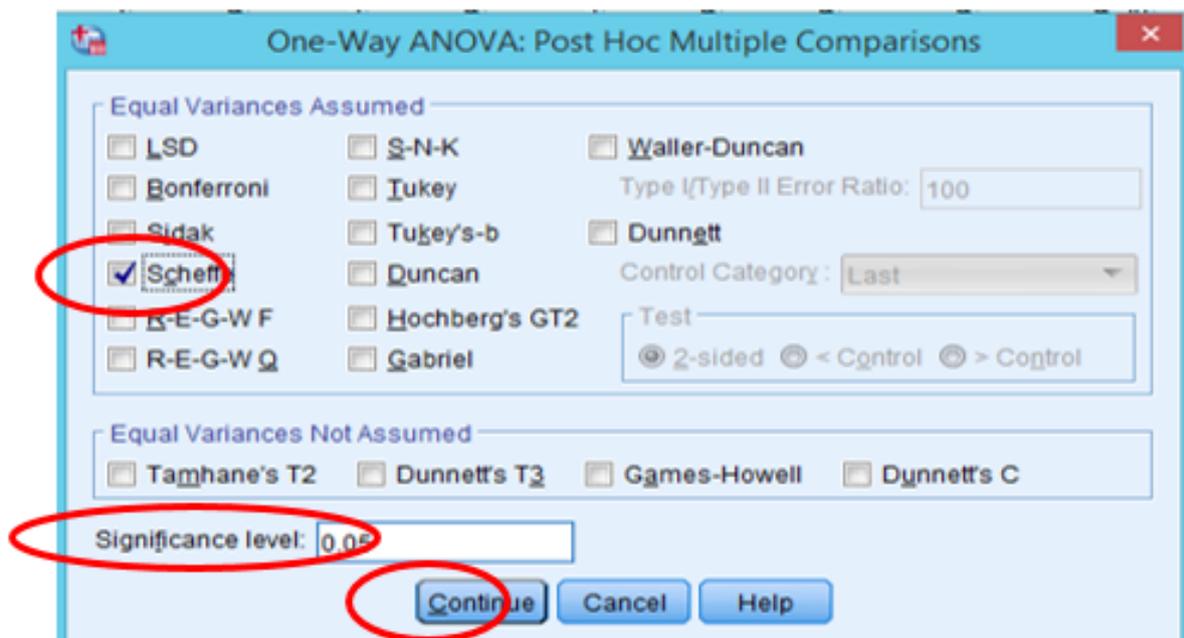
เมื่อคลิก Post hoc จะปรากฏหน้าต่าง One-way ANOVA: Post hoc multiple comparisons ดังภาพประกอบ 8



ภาพประกอบ 8 Multiple comparisons

ในส่วนหน้าต่าง One-way ANOVA: Post hoc multiple comparisons จะปรากฏวิธีการเปรียบเทียบรายคู่วิธีต่าง ๆ ให้เลือกวิธีใดวิธีหนึ่งตามความเหมาะสมของลักษณะของข้อมูล ลักษณะของกลุ่มตัวอย่าง และข้อตกลงของแต่ละวิธี ในบทความนี้เลือกวิธี Scheffe' แล้วใส่ค่า Significance level: (.05 หรือ .01) ตามที่ผู้วิจัยกำหนดขึ้น

แล้วคลิก Continue เพื่อกลับไปยังหน้าต่าง One-way ANOVA ดังภาพประกอบ 9



ภาพประกอบ 9 One-Way ANOVA

แล้วคลิกปุ่ม OK จะปรากฏผลการคำนวณที่ Output Window ในส่วนของ Descriptive ดังภาพประกอบ 10

→ Oneway

	จำนวนข้อมูล N	ค่าเฉลี่ยของข้อมูลแต่ละกลุ่ม Mean	ส่วนเบี่ยงเบนมาตรฐาน Std. Deviation	ความคลาดเคลื่อนมาตรฐาน Std. Error	ขอบเขตของค่าความเชื่อมั่น 95% Confidence Interval for Mean		ค่าต่ำสุดของกลุ่ม Minimum	ค่าสูงสุด ของกลุ่ม Maximum
					Lower Bound	Upper Bound		
บ ดร์	11	4.5000	.18540	.05590	4.3754	4.6246	4.13	4.75
บ ๊ว	11	4.0114	.21253	.06408	3.8686	4.1541	3.75	4.38
บ เอก	8	4.1250	.16366	.05786	3.9882	4.2618	3.88	4.38
Total	30	4.2208	.28746	.05248	4.1135	4.3282	3.75	4.75

ภาพประกอบ 10 ส่วนของ Descriptive

และผลการวิเคราะห์ความแปรปรวน (ANOVA) หรือ F-test ดังภาพประกอบ 11

ANOVA

Mean	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.413	2	.707	19.412	.000
Within Groups	.983	27	.036		
Total	2.396	29			

มีค่าน้อยกว่า .05 ถือว่า Sig

ภาพประกอบ 11 ผลการวิเคราะห์ความแปรปรวน

ผลการวิเคราะห์ความแปรปรวน พบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างทุกกลุ่มในการวิจัยแตกต่างกันอย่างมีนัยสำคัญ โดยดูค่า Sig (หรือ p-Value) ถ้า p มีค่าน้อยกว่าหรือเท่ากับค่าแอลฟา จะถือว่ามีนัยสำคัญ (Sig) แต่ถ้าค่า p มีค่ามากกว่าค่าแอลฟา ถือว่าไม่มีนัยสำคัญ (ไม่ Sig) จากผลการวิเคราะห์นี้ กำหนดค่าแอลฟา เท่ากับ .05 ( $\alpha = .05$ ) ดังนั้น ถือว่าแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 โปรแกรมจึงทำการเปรียบเทียบเป็นรายคู่ตามวิธี Scheffe' ตามที่กำหนดไว้(ผู้เขียนได้ลองวิเคราะห์ด้วยวิธี LSD เพื่อนำมาเปรียบเทียบผลการทดสอบ) ผลปรากฏดังภาพประกอบ 12-13

**Multiple Comparisons**

Dependent Variable: Mean

**Post Hoc Tests**

**Scheffe**

(I) ระดับกศ.	(J) ระดับกศ.	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
ม.ตรี	ม.โท	.48864*	.08136	.000	.2779	.6994
	ม.เอก	.37500*	.08866	.001	.1454	.6046
ม.โท	ม.ตรี	-.48864*	.08136	.000	-.6994	-.2779
	ม.เอก	-.11364	.08866	.450	-.3433	.1160
ม.เอก	ม.ตรี	-.37500*	.08866	.001	-.6046	-.1454
	ม.โท	.11364	.08866	.450	-.1160	.3433

\*. The mean difference is significant at the 0.05 level.

ภาพประกอบ 12 วิเคราะห์ด้วยวิธี Scheffe'

**Multiple Comparisons**

Dependent Variable: A\_mean

**LSD**

(I) ระดับกศ.	(J) ระดับกศ.	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
ม.ตรี	ม.โท	.48864*	.08136	.000	.3217	.6556
	ม.เอก	.37500*	.08866	.000	.1931	.5569
ม.โท	ม.ตรี	-.48864*	.08136	.000	-.6556	-.3217
	ม.เอก	-.11364	.08866	.211	-.2955	.0683
ม.เอก	ม.ตรี	-.37500*	.08866	.000	-.5569	-.1931
	ม.โท	.11364	.08866	.211	-.0683	.2955

\*. The mean difference is significant at the 0.05 level.

ภาพประกอบ 13 วิเคราะห์ด้วยวิธี LSD

การดูว่า ค่าเฉลี่ยของกลุ่มใดมีความแตกต่างอย่างมีนัยสำคัญหรือไม่ พิจารณาจากผลต่างระหว่างค่าเฉลี่ยของแต่ละคู่ที่มีดอกจัน (\*) ถ้าคู่ใดมีดอกจัน แสดงว่า คู่ นั้นมีความแตกต่างอย่างมีนัยสำคัญ (หรือ Sig) หรือดูที่ช่อง Sig ก็ได้ คู่ใดที่ค่า Sig น้อยกว่าหรือเท่ากับ .05 แสดงว่า คู่ นั้นแตกต่างอย่างมีนัยสำคัญที่ระดับ .05 และแต่ละคู่ที่เปรียบเทียบกันนั้น ผลการเปรียบเทียบจะเป็นค่าบวก 1 ตัว และค่าลบ 1 ตัว จากตัวอย่างผลการวิเคราะห์เพื่อเปรียบเทียบความแตกต่างของค่าเฉลี่ยเป็นรายคู่ด้วยวิธี Scheffe' (วิธี LSD เพื่อเปรียบเทียบ) พบว่าทั้ง 2 วิธีตรวจสอบความแตกต่างได้ผลที่ตรงกัน

ตาราง 1 ผลการวิเคราะห์ความแปรปรวนของระดับความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียน ของครูที่มีระดับการศึกษาต่างกัน

แหล่งความแปรปรวน	df	SS	MS	F	p-value
ระหว่างกลุ่ม	2	1.413	.707	19.412	.00
ภายในกลุ่ม	27	.983	0.36		
รวม	29	2.396			

จากตาราง 1 พบว่า ความแปรปรวนของระดับความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียน ของครูที่มีระดับการศึกษาต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ดังนั้น ค่าเฉลี่ยความคิดเห็นของครูที่มีระดับการศึกษาต่างกัน มีความแตกต่างกันอย่างน้อย 1 คู่ จึงนำค่าเฉลี่ยไปทดสอบเป็นรายคู่โดยวิธีของเชฟเฟ (Scheffe' method) ผลปรากฏดังตาราง 2

ตาราง 2 ผลการเปรียบเทียบเป็นรายคู่ของค่าเฉลี่ยความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียน ของครูที่มีระดับการศึกษาต่างกัน

ระดับการศึกษา	ปริญญาตรี (ค่าเฉลี่ย 4.5000)	ปริญญาโท (ค่าเฉลี่ย 4.0114)	ปริญญาเอก (ค่าเฉลี่ย 4.1250)
ปริญญาตรี (ค่าเฉลี่ย 4.5000)	1.000	.48864*	.37500*
ปริญญาโท (ค่าเฉลี่ย 4.0114)		1.000	.11364
ปริญญาเอก (ค่าเฉลี่ย 4.1250)			1.000

จากตาราง 2 พบว่า ค่าเฉลี่ยความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียนของครูที่มีการศึกษาระดับปริญญาตรี กับค่าเฉลี่ยความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียน ของครูที่มีการศึกษาระดับปริญญาโท และปริญญาเอก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยค่าเฉลี่ยความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียนของครูที่มีการศึกษาระดับปริญญาตรีสูงกว่าค่าเฉลี่ยความคิดเห็นต่อการบริหารตามหลักธรรมาภิบาลของผู้บริหารโรงเรียน ของครูที่มีการศึกษาระดับปริญญาโทและปริญญาเอก นอกนั้นไม่แตกต่างกัน

### บทสรุป

วิธีการทดสอบภายหลังที่การวิเคราะห์ความแปรปรวน พบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างโดยรวม (Omnibus test) มีความแตกต่างกัน จึงต้องทำการตรวจสอบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างคู่ใดบ้างที่แตกต่างกัน วิธีการทดสอบเพื่อเปรียบเทียบรายคู่

มีหลายวิธี แต่ละวิธีมีข้อดีข้อด้อย มีวัตถุประสงค์และคุณลักษณะเฉพาะในการเปรียบเทียบแตกต่างกัน ในงานวิจัยทางการศึกษาที่เป็นเชิงเปรียบเทียบส่วนใหญ่จะนิยมใช้วิธี Scheffe' เนื่องจากเป็นเทคนิคที่ดีในการตรวจสอบนัยสำคัญและมีความละเอียดอ่อนกว่าวิธีการทดสอบอื่น ๆ ในการเปรียบเทียบแบบหลายคู่พร้อมกัน แต่จะมีอำนาจการทดสอบ (Power of test) ต่ำ (Conservative) กว่าทดสอบที่ออกแบบมาเพื่อการเปรียบเทียบที่วางแผนไว้ล่วงหน้า นั่นคือ ถ้าผลการทดสอบพบว่า มีนัยสำคัญ แสดงว่าค่าเฉลี่ยของกลุ่มตัวอย่างมีความแตกต่างกันจริง ๆ อีกทั้ง สามารถใช้ได้ทั้งในกรณีที่กลุ่มที่เปรียบเทียบมีขนาดเท่ากันหรือไม่เท่ากันก็ได้ ส่วนวิธี LSD มีความไว (Sensitivity) ในการทดสอบสูงมาก จึงอาจเกิด Type I error มากกว่าวิธี Scheffe' และในกรณีที่แต่ละกลุ่มมีขนาดเท่ากัน และมีค่าความแปรปรวนเท่ากันแล้ว พบว่า วิธี LSD มีอำนาจการทดสอบสูงสุด ส่วนวิธี Turkey-Kramer มีอำนาจการทดสอบต่ำสุด ปัจจุบันมีความเจริญก้าวหน้าทางเทคโนโลยีรวดเร็ว

มาก ทำให้สามารถใช้วิธีการทดสอบภายหลัง (Post hoc comparison) ด้วยโปรแกรมวิเคราะห์ SPSS เวอร์ชันต่าง ๆ มีความสะดวก ง่าย และเอื้อต่อการคำนวณมาก เพียงแต่ผู้ใช้ควรพิจารณาเลือกใช้ให้สอดคล้องและเหมาะสมกับลักษณะของข้อมูลและกลุ่มตัวอย่าง เพื่อให้ได้ผลของการตรวจสอบที่มีความคลาดเคลื่อนน้อยที่สุด และตอบปัญหาการวิจัยได้ชัดเจนมากที่สุด

#### เอกสารอ้างอิง

- นิภาพร ขำสะอาด. (2552). *อำนาจการทดสอบของการใช้สถิติการทดสอบความแตกต่างของค่าเฉลี่ย*. [ปริญญาณิพนธ์ปริญาการศึกษา มหาวิทยาลัย สาขาวิชาการวัดผลการศึกษา]. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- สุพัฒน์ สุกมลสันต์. (2560). การเปรียบเทียบก่อนและหลังการทดสอบรวมเพื่อการวิจัย. *วารสารวิชาการ มหาวิทยาลัยราชภัฏบุรีรัมย์*, 9(2), 51-70.
- Clinch, J. J., & Keselman, H. J. (1982). Parametric alternatives to the analysis of variance. *Journal of Educational Statistics*, 7(3), 42-43.
- Cochran, W.G. & Cox, G.M. (1976). *Experimental design*. John Wiley and Sons.
- Fisher, R. A. (1959). *Statistical methods and scientific inference*. Oliver and Boyd.
- Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences*. (3<sup>rd</sup> ed.). Thomson Brooks/Cole Publishing Co.
- Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications, *Biometrics*, 12, 309-310.
- Scheffe, Henry. (1953). A method for judging all contrasts in the analysis of variance. *Biometrics*, 40(1-2), 87-110.
- Tukey, J. W. (1953). Comparing individual means in the analysis of variance. *Biometrics*, 5(2), 99-114.