การวิเคราะห์ความแปรปรวนทางเดียวและการเปรียบเทียบพหุคูณในงานวิจัยทางสาธารณสุข กรณีตัวอย่างการป้องกันโรคไข้เลือดออก

One-Way ANOVA and Multiple Comparison in Public Health Research: A Case Study of Hemorrhagic Fever Protection

จารุวรรณ วิโรจน์[1]
Jaruwan Viroj[1]
Received: 3 August 2014 ; Accepted: 12 November 2014

บทคัดย่อ

การวิเคราะห์ความแปรปรวนแบบทางเดียว เป็นสถิติอนุมานที่ใช้ในการวิเคราะห์ความแตกต่างของค่าเฉลี่ยของข้อมูลตั้งแต่ 3 กลุ่มขึ้นไป ใช้ในการวิเคราะห์ที่มีตัวแปรตามหนึ่งตัวและตัวแปรต้นหนึ่งตัว โดยตัวแปรตามที่ใช้ในการวิเคราะห์จะต้องมีลักษณะข้อมูลเป็นข้อมูลต่อเนื่อง ซึ่งการวิเคราะห์ดังกล่าวสามารถควบคุมความผิดพลาดประเภทที่ 1 ซึ่งมีข้อตกลงจำนวน 3 ข้อ ได้แก่ ข้อมูลที่นำมาวิเคราะห์ต้องอิสระต่อกัน ค่าความแปรปรวนของประชากรแต่ละกลุ่มต้องเท่ากัน และข้อมูลที่นำวิเคราะห์ต้องมาจากประชากรที่มีการแจกแจงแบบปกติ หากการวิเคราะห์พบว่ามีค่าเฉลี่ยอย่างน้อย 1 คู่แตกต่างกัน การที่จะหาคำตอบว่าค่าเฉลี่ยคู่ใดแตกต่างกันบ้าง จำเป็นต้องทำการเปรียบเทียบที่เรียกว่าการเปรียบเทียบพหุคูณ โดยสถิติแต่ละตัวมีความสามารถในการควบคุมความผิดพลาดในการทดสอบได้แตกต่างกัน ดังนั้นผู้วิจัยจึงควรเลือกใช้สถิติให้เหมาะสมเพื่อจะสามารถนำไปสู่การสรุปการวิเคราะห์ข้อมูลที่ถูกต้องต่อไป

คำสำคัญ: การวิเคราะห์ความแปรปรวนแบบทางเดียว การเปรียบเทียบพหุคูณ พฤติกรรมสุขภาพ

Abstract

One-way ANOVA is an inferential statistic for analyzing the mean difference between more than two groups. It is conducted with one dependent variable and one independent variable. The dependent variable is a continuous response variable. This statistic can control Type I errors. Three assumptions are made in ANOVA models: observations independent of one another, homogeneity of variance and normality. In those cases having a difference between the means, an additional exploration of the difference among means multiple comparisons test, is needed. There are several procedures for multiple comparisons, each having different control errors. As such, the researcher must choose an appropriate statistic that leads to a precise conclusion.

Keywords: One-way ANOVA, Multiple comparison, Health behavior

บทนำ

การวิเคราะห์ข้อมูลพฤติกรรมสาธารณสุขเป็นสิ่งที่สำคัญยิ่งที่จะนำไปสู่การแก้ไขปัญหาสาธารณสุขที่เหมาะสมและมีประสิทธิภาพ บ่อยครั้งที่นักวิจัยอาจได้พยายามคิดค้นและนำวิธีการที่หลากหลายมาใช้ในการแก้ไขพฤติกรรมสุขภาพ การหาคำตอบว่าวิธีการใดจะสามารถนำมาให้แล้วได้ประสิทธิผลมากที่สุดได้นั้น One-way ANOVA นับได้ว่าเป็นสถิติที่สามารถให้ผลการวิเคราะห์ที่น่าเชื่อถือ และสามารถนำประยุกต์ใช้ในการพิสูจน์สมมติฐานความแตกต่างข้อมูลซึ่งมีจำนวนข้อมูลตั้งแต่ 3 กลุ่มขึ้นไปบทความฉบับนี้ผู้เขียนได้การนำเสนอหลักการออกแบบงานวิจัย แนวคิด หลักการ รวมถึงหลักการพิจารณาการเลือกการวิเคราะห์ด้วย One-way ANOVA และการเปรียบเทียบพหุคูณ ซึ่งสามารถช่วยให้นักวิจัยสามารถตอบคำถามและสรุปผลการศึกษาได้อย่างถูกต้อง

[1] *อาจารย์, คณะสาธารณสุขศาสตร์ มหาวิทยาลัยมหาสารคาม เบอร์โทรศัพท์ 043 - 754353*

[1] *Lecture, Public Health Faculty, Mahasarakham University, Telephone number 043 - 754353*

การออกแบบการวิจัยเพื่อการวิเคราะห์ด้วย One-way ANOVA

ด้วยการเปลี่ยนแปลงทางสังคมในปัจจุบัน ส่งผลพฤติกรรมสุขภาพของคนเปลี่ยนไปจากอดีต และส่งผลให้ปัญหาด้านสุขภาพมีความซับซ้อนมากขึ้น โดยเฉพาะอย่างยิ่งปัญหาโรคไข้เลือดออก แม้ว่าในปัจจุบันหน่วยต่างๆ ได้ให้แนวทางในการจัดการปัญหาดังกล่าว แต่ก็ยังพบว่าอุบัติการณ์เกิดโรค และมีผู้เสียชีวิตสูงขึ้นในทุกๆ ปี ในปี 2556 พบผู้ป่วยด้วยโรคไข้เลือดออกมากกว่าปี 2555 ถึง 2.4 เท่า[1] แต่โรคไข้เลือดออกเป็นโรคที่สามารถป้องกันได้ หากประชาชนมีพฤติกรรมสุขภาพในการจัดการสภาพแวดล้อม และการป้องกันตนเองที่เหมาะสม[2] การนำกระบวนการวิเคราะห์ที่มีประสิทธิภาพมาใช้ในการแก้ไขปัญหาการแพร่ระบาดของโรคไข้เลือดออกนับว่าเป็นแนวทางที่สำคัญ ที่จะช่วยให้สามารถหาคำตอบถึงประสิทธิภาพในการดำเนินการแก้ไขปัญหาได้อย่างน่าเชื่อถือ และสามารถนำผลการศึกษาไปใช้ในการแก้ปัญหาจริงในพื้นที่ การวิจัยแบบกึ่งทดลอง(Quasi Experiment) เป็นรูปแบบการวิจัยที่มีการนำมาใช้ในงานวิจัยที่มีการทดลองที่มีหน่วยการทดลองเป็นประชาชนในชุมชน เนื่องจากนักวิจัยไม่สามารถควบคุมปัจจัยในการทดลองต่างๆ ได้ทุกปัจจัย[3] เช่น การจำกัดขอบเขตการรับข่าวสารในชุมชน การจำกัดด้านสิ่งแวดล้อม เป็นต้น ตัวอย่างเช่น นักวิจัยต้องการเปรียบเทียบคะแนนพฤติกรรมการป้องกันตนเองจากโรคไข้เลือดออกในประชาชนที่ได้รับความรู้ของเจ้าหน้าที่สาธารณสุขและประชาชนที่ได้รับความรู้จากกระบวนการมีส่วนร่วมในการป้องกันโรคไข้เลือดออก นักวิจัยอาจทำการออกแบบให้มีกลุ่มวิจัยจำนวน 3 กลุ่ม ได้แก่ กลุ่มควบคุม (control group) กลุ่มที่ได้รับความรู้จากเจ้าหน้าที่สาธารณสุข (group 1) กลุ่มที่ได้รับความรู้โดยการสร้างการมีส่วนร่วม (group 2) ในการวิจัยนักวิจัยควรออกแบบให้มีกลุ่มควบคุมทั้งนี้เพื่อให้สามารถยืนยันผลการวิจัยได้ชัดเจนถึงพฤติกรรมของประชาชนว่ามีการเปลี่ยนแปลงเนื่องมาจากการทดลองหรือไม่ หากนักวิจัยออกแบบการวิจัยโดยเก็บข้อมูลในกลุ่มตัวอย่าง (กรณีตัวอย่าง)ด้วยวิธีการสุ่มอย่างง่ายกลุ่มละ 10 ตัวอย่าง ซึ่งข้อมูลที่นักวิจัยทำการศึกษาได้ผลดัง Table 1

Table 1 Protective performance scores of Hemorrhagic fever

Group	Protective performance scores										mean (SD)
control group	18	15	14	16	15	13	12	13	16	17	14.9 (1.9)
group 1	19	17	17	15	16	18	19	17	18	17	17.3 (1.3)
group 2	19	20	18	18	15	17	18	17	18	16	17.6 (1.4)

จาก Table 1 แม้ว่านักวิจัยจะสามารถคำนวณค่าเฉลี่ยในแต่ละกลุ่มออกมาได้ แต่การที่จะหาคำตอบว่าความแตกต่างของค่าเฉลี่ยดังกล่าวมีความแตกต่างจริงในประชากรหรือไม่ จำเป็นต้องที่ต้องมีการคำนวณค่าตามวิธีทางสถิติมาอธิบายคำตอบนั้น จากข้อมูลดังกล่าวมีตัวแปรตามจำนวน 1 ตัวแปร คือ ตัวแปรคะแนนพฤติกรรมการป้องกันตนเองจากโรคไข้เลือดออก ซึ่งมีการวัดเป็นแบบข้อมูลเป็นข้อมูลต่อเนื่องและมีตัวแปรต้นจำนวน 1 ตัวแปร คือตัวแปรรูปแบบการให้ความรู้ มีลักษณะการวัดเป็นแบบข้อมูลแจงนับ (discrete variable) ในการเปรียบเทียบข้อมูลจำนวน 3 กลุ่ม หากนักวิจัยทำการเปรียบเทียบค่าเฉลี่ยทีละคู่โดยใช้ two sample t-test นักวิจัยทำการทดสอบความแตกต่างของข้อมูลรายคู่ของข้อมูลทั้ง 3 กลุ่ม โดยต้องทำการทดสอบ 3 ครั้ง คือ ค่าเฉลี่ยกลุ่มควบคุม = ค่าเฉลี่ยกลุ่มที่ 1, ค่าเฉลี่ยกลุ่มควบคุม = ค่าเฉลี่ยกลุ่มที่ 2 และ ค่าเฉลี่ยกลุ่มที่ 1 = ค่าเฉลี่ยกลุ่มที่ 2 ซึ่งวิธีการวิเคราะห์รายคู่นี้จะทำให้เกิดผิดพลาดประเภทที่ 1 (Type I error) เพิ่มขึ้น ซึ่งทำให้ความผิดพลาดที่เกิดจากการตัดสินในการไม่ยอมรับสมมติฐานหลักที่เป็นจริงเพิ่มขึ้น[4, 5] ดังนั้นการเลือกใช้สถิติ One-way ANOVA นับได้ว่าเป็นสถิติที่สามารถให้ผลการวิเคราะห์ที่น่าเชื่อถือกว่าการทดสอบรายคู่ เนื่องจากในการวิเคราะห์ความแตกต่างของค่าเฉลี่ยด้วย ANOVA จะทำการวิเคราะห์ค่าความแปรปรวนที่เกิดขึ้นจากข้อมูล และนำข้อมูลที่ได้จากกลุ่มตัวอย่าง ไปทดสอบด้วยสถิติ F ซึ่งคิดค้นมาจาก Sir Ronald Aylmer Fisher[6, 7] โดยมีแนวคิดว่าหากข้อมูลที่นำมาเปรียบเทียบในแต่ละกลุ่มมีค่าเฉลี่ยเท่ากัน ค่าความแปรปรวนในแต่ละกลุ่มจะเท่ากับศูนย์ แต่ถ้าในแต่ละกลุ่มมีค่าเฉลี่ยไม่เท่ากันดังนั้นค่าความแปรปรวนในแต่ละกลุ่มจะมากกว่าศูนย์ ซึ่งการวิเคราะห์ดังกล่าวสามารถควบคุมความผิดพลาดประเภทที่ 1 จากการทดสอบสมมติฐานได้ โดยความแปรปรวนของข้อมูลเกิดจากผลรวมของค่าความ

แปรปรวนระหว่างกลุ่ม (between –group portion of the variance) และค่าความแปรปรวนภายในกลุ่ม (within-group portion of the variance)[8] ดังสมการ

$$SS_{total} = SS_{betw} + SS_{with}$$

ซึ่งสามารถเขียนสมมติฐานในการทดสอบสมมติฐานได้ดังนี้

H_0 : ค่าเฉลี่ยกลุ่มควบคุม = ค่าเฉลี่ยกลุ่มที่ 1 = ค่าเฉลี่ยกลุ่มที่ 2

H_A : มีค่าเฉลี่ยอย่างน้อย 1 คู่ที่แตกต่างกัน

ข้อตกลงเบื้องต้นในการวิเคราะห์ข้อมูลด้วย One-way ANOVA

ข้อตกลงเบื้องต้นในการวิเคราะห์ข้อมูลด้วย One-way ANOVA ประกอบด้วยข้อตกลงจำนวน 3 ข้อ ได้แก่

1. ข้อมูลที่นำมาวิเคราะห์ต้องอิสระต่อกัน (Independent) ซึ่งนักวิจัยสามารถทดสอบข้อตกลงดังกล่าวได้จากการนำค่า residuals มาพล็อตเป็น Scatter plot จำแนกต่างกลุ่มข้อมูล ซึ่งหากพบว่าข้อมูลไม่อิสระต่อกันจะมีผลทำให้เกิดความคลาดเคลื่อนในการทดสอบ(Standard errors) โดยทำให้ค่าของ Type I error และ Type II error มีค่าเพิ่มขึ้น และจะนำไปสู่ความผิดพลาดในการสรุปผลการทดสอบสมมติฐาน[9]

2. ค่าความแปรปรวนของประชากรแต่ละกลุ่มต้องเท่ากัน (Homogeneity of Variance) ซึ่งสามารถทำการทดสอบได้จากการนำค่า residuals มาพล็อตด้วยแกน X หรือทำการทดสอบด้วย Levene's test หากไม่ผ่านข้อตกลงดังกล่าวจะทำให้เกิดอคติภายในกลุ่มวิเคราะห์ และเพิ่ม Type I error และ Type II error[9] และความผิดพลาดในการทดสอบสมมติฐานจะเพิ่มมากขึ้นถ้าหากข้อมูลที่นำมาทดสอบแต่ละกลุ่มมีจำนวนตัวอย่างไม่เท่ากัน แต่ถ้าหากขนาดตัวอย่างมากขึ้นอคติดังกล่าวจะเกิดลดลง[6]

3. ข้อมูลที่นำวิเคราะห์ต้องมาจากประชากรที่มีการแจกแจงแบบปกติ (Normality) ซึ่งสามารถทดสอบได้หลายวิธี เช่น การนำข้อมูลของตัวแปรตามแต่กลุ่มมาพล็อตด้วยกราฟฮิสโตแกรม แผนภูมิต้นและใบ นอกจากนี้ยังสามารถพิจารณาจากค่าทางสถิติได้ เช่น พิจารณาจากค่า Skewness และ Kutosis หรืออาจทำการทดสอบด้วยสถิติ Komogorov-Sminov และ Shapiro-Wilk หากพบว่าข้อมูลไม่แจกแจงปกติ จะมีผลทำให้การทดสอบเกิดคลาดเคลื่อนเล็กน้อย ซึ่งความคลาดเคลื่อนดังกล่าวจะลดลงหากข้อมูลที่นำมาทดสอบในแต่ละกลุ่มมีจำนวนเท่ากัน หรือมีจำนวนข้อมูลมาก[9]

ในการวิเคราะห์ข้อมูลด้วย One-way ANOVA ข้อมูลที่นำมาวิเคราะห์ต้องเป็นไปตามข้อตกลงเบื้องต้น 3 ข้อดังกล่าว หากข้อมูลไม่เป็นไปตามข้อตกลง ผู้วิจัยอาจเปลี่ยนมาใช้สถิติประเภทนอนพาราเมตริกซ์แทน โดยสถิติดังกล่าวคือการวิเคราะห์ด้วย Kruskal-Wallis test[10]

ประเภทการเปรียบเทียบพหุคูณ

เมื่อนักวิจัยทำการทดสอบด้วย One-way ANOVA และพบว่าผลการวิเคราะห์มีค่าเฉลี่ยของข้อมูลแตกต่างกันอย่างน้อย 1 คู่ การที่จะอธิบายต่อไปว่าข้อมูลชุดใดบ้างที่มีค่าเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญทางสถิติ จำเป็นต้องทำการเปรียบเทียบที่เรียกว่า การเปรียบเทียบพหุคูณ โดยสถิติในกลุ่มที่ใช้การเปรียบเทียบพหุคูณเป็นสถิติที่สามารถควบคุมความผิดพลาดประเภทต่อหนึ่งกลุ่มการทดสอบ (family-wise) ซึ่งทำให้ลดการเกิดอัตราความคลาดเคลื่อนชนิดที่ 1[11] โดยการเปรียบเทียบพหุคุณแบ่งได้เป็น 2 ประเภท คือ การเปรียบเทียบข้อมูลที่มีการวางแผนไว้ล่วงหน้า (Planned comparisons) และ การเปรียบเทียบข้อมูลภายหลัง (Post Hoc comparisons)

การเปรียบเทียบข้อมูลที่มีการวางแผนไว้ล่วงหน้า คือ การเปรียบเทียบที่นักวิจัยได้กำหนดไว้ล่วงหน้าไว้ว่าจะทำการเปรียบเทียบค่าเฉลี่ยคู่ใดบ้างก่อนการวิเคราะห์ความแปรปรวน โดยผู้วิจัยจะทำการเปรียบเทียบค่าเฉลี่ยเฉพาะกลุ่มตัวอย่างที่สนใจ โดยทั่วไปจะใช้ก่อนดำเนินการศึกษา และสามารถใช้วิเคราะห์ในกรณีที่ขนาดตัวอย่างเพียงเล็กน้อยได้[12] โดยนักวิจัยจะเป็นผู้กำหนดว่าต้องการทดสอบความแตกต่างของกลุ่มตัวแปรใดบ้าง เช่น นักวิจัยต้องการทดสอบความแตกต่างของค่าเฉลี่ยจำนวน 5 กลุ่ม นักวิจัยอาจทำการเปรียบเทียบเฉพาะกลุ่มที่ 1 กับกลุ่มที่ 5 และ กลุ่มที่ 4 กับกลุ่มที่ 5 เป็นต้น เทคนิคที่นิยมใช้มี 3 ประเภท ได้แก่

การเปรียบเทียบแบบ Trend

ใช้สำหรับการวิเคราะห์กรณีที่ตัวแปรต้นมีมาตรวัดแบบ ข้อมูลช่วง (Interval scale) หรือ ข้อมูลอัตราส่วน (Ratio scale) ทำการเปรียบเทียบความแตกต่างโดยใช้ standard t test ในการวิเคราะห์จะพิจารณาตามรูปแบบของกราฟเส้นว่าตัวแปรต้นส่งผลทำให้เกิดกราฟเส้นอย่างไร[6, 7, 9]

การเปรียบเทียบแบบ Planned orthogonal contrasts (POC)

โดยนักวิจัยจะกำหนดกลุ่มที่ต้องการเปรียบเทียบเอง โดยจะมีการกำหนดคู่การเปรียบเทียบไม่กี่คู่ ข้อมูลที่นำมาเปรียบเทียบต้องเป็นอิสระต่อกันและมีข้อมูลในแต่ละกลุ่มเท่ากัน ทำการเปรียบเทียบความแตกต่างโดยใช้ standard t-test แต่วิธีการเปรียบเทียบนี้ไม่ได้มีการควบคุมความผิดพลาดต่อหนึ่งกลุ่มทดสอบ (family-wise)[9]

การเปรียบเทียบแบบ Dunnett

ใช้เปรียบเทียบค่าเฉลี่ยของข้อมูลในกลุ่มที่ทดลองกับกลุ่มควบคุม[13] โดยทั่วไปใช้ทดสอบว่าสิ่งทดลองจะให้ผลดีกว่ากลุ่มควบคุมหรือไม่ ซึ่งทำการเปรียบเทียบความแตกต่างโดยใช้ standard t test โดยวิธีนี้มีประสิทธิในการทดสอบดีกว่าวิธีของ Bonferroni[9] โดยทั่วไปแล้วในการศึกษาจะมีการกำหนดให้จำนวนข้อมูลในกลุ่มควบคุมมากกว่ากลุ่มทดสอบตามกฎของ the rule of thumb กำหนดว่ากลุ่มควบคุมควรมีจำนวนกลุ่มตัวอย่างเท่ากับ $n\sqrt{k-1}$ (n คือ จำนวนกลุ่มตัวอย่างในแต่ละกลุ่มทดลอง และ k คือ จำนวนกลุ่มทดลอง)[14]

การเปรียบเทียบข้อมูลภายหลัง (Post Hoc Comparison or Posteriori Comparison)

เป็นการเปรียบเทียบโดยการนำข้อมูลของทุกกลุ่มที่ศึกษามาเปรียบเทียบเพื่อศึกษาความแตกต่างของค่าเฉลี่ย ทั้งนี้เพราะนักวิจัยไม่ทราบมาก่อนว่าค่าเฉลี่ยใดบ้างที่แตกต่างกัน เช่น มีจำนวนข้อมูลอยู่สามกลุ่ม ก็ทำการเปรียบเทียบทั้งสามกลุ่ม โดยวิธีการเปรียบเทียบแบบภายหลังสามารถทำได้หลายวิธี ซึ่งวิธีการทดสอบที่สำคัญอยู่ 8 วิธีการ ได้แก่ Fisher PLSD, Tukey HSD, Tukey-Kramer, Scheffe, Bonferroni, Duncan's Multiple Range, Student-Newman Keuls และ Game/Howell โดยแต่ละวิธีจะมีรายละเอียดดังนี้

Fisher LSD

วิธีนี้คิดค้นโดย Fisher โดยได้ทำการทดสอบที่เรียกว่า least significant different (LSD) ซึ่งใช้อัตราส่วน ที พหุคูณในการทดสอบเปรียบเทียบค่าเฉลี่ยรายคู่ เพื่อกำหนดค่าความต่างน้อยที่สุด เป็นเกณฑ์ตัดสินใจในการทดสอบสมมติฐาน ใช้ในการเปรียบเทียบข้อมูลที่มีจำนวนตัวอย่างในแต่ละกลุ่มเท่ากัน โดยคำนวณจาก Standard t test[9] วิธีนี้เหมาะกับการเปรียบเทียบข้อมูลที่มีจำนวนตัวอย่างและจำนวนกลุ่มที่นำมาเปรียบเทียบไม่มากนัก เนื่องจากไม่สามารถควบคุมความผิดพลาดภายในกลุ่มได้(familywise error) [12, 14, 15]

Tukey HSD

วิธีนี้คิดค้นโดย Tukey โดยได้เสนอวิธีการเปรียบเทียบพหุคูณ โดยใช้อัตราส่วน q (Studentized range statistic) และการกระจายตัวของ q ซึ่งได้เรียกวิธีดังกล่าวว่า Honestly Significant Different โดยใช้สัญลักษณ์ HSD ใช้ในการเทียบค่าเฉลี่ยกรณีกลุ่มตัวอย่างมีขนาดเท่ากัน ในการคำนวณใช้ Studentized range test วิธีนี้สามารถควบคุมความผิดพลาดประเภท family-wise ได้[9] โดยค่า HSD ที่ได้เป็นค่าต่ำสุดของผลต่างของค่าเฉลี่ยที่ใช้เปรียบเทียบ หากผลต่างของค่าเฉลี่ยคู่ใดเท่ากับหรือมากกว่าค่า HSD แสดงว่ากลุ่มตัวอย่างคู่นั้นแตกต่างกัน นอกจากนี้วิธีนี้สามารถควบคุมความผิดพลาดต่อหนึ่งการทดลอง (Experimentwise Error Rate) ซึ่งเป็นวิธีที่ให้ผลการทดสอบที่น่าเชื่อถือกว่าวิธี Bonferroni และ Scheffe's[9] วิธีนี้จะทำการเปรียบเทียบทีละคู่ ซึ่งจะเริ่มจากคู่ที่มีค่ามากที่สุดและน้อยที่สุด หากพบว่ามีความแตกต่างกันจะทำการเปรียบเทียบต่อ แต่ถ้าหากพบว่าคู่ดังกล่าวไม่แตกต่างกันจะหยุดการวิเคราะห์

Tukey-Kramer

เป็นวิธีการเปรียบเทียบซึ่งมีหลักการเปรียบเทียบเหมือนวิธีของ Tukey HSD แต่ใช้เปรียบเทียบกรณีกลุ่มตัวอย่างแต่กลุ่มที่นำมาเปรียบเทียบมีจำนวนตัวอย่างไม่เท่ากัน[9] เป็นวิธีที่เหมาะสมสำหรับงานวิจัยที่ต้องการเปรียบเทียบความแตกต่างของข้อมูลทุกกลุ่ม เนื่องจากให้ค่าความเชื่อมั่นในการทดสอบที่น่าเชื่อถือ[16, 17]

Scheffe

วิธีนี้คิดค้นโดย Scheffe โดยได้เสนอวิธีการเปรียบเทียบพหุคูณโดยใช้อัตราส่วน และการกระจายของตัวอย่าง F ซึ่งสามารถใช้เปรียบเทียบพหุคูณแบบคู่ (pair wise) และสามารถใช้ได้ทั้งกรณีที่ข้อมูลที่เปรียบเทียบมีกลุ่มตัวอย่างเท่าหรือไม่เท่ากันได้ วิธีนี้สามารถใช้ในการเปรียบเทียบ และสามารถควบคุมความผิดพลาดประเภท family-wise ได้ดี[15] และสามารถควบคุมให้อัตราความคลาดเคลื่อนชนิดที่ 1 ของการทดสอบทั้งชุดทดลองเท่ากับ ในการวิเคราะห์ใช้ Standard t test[9] ผลการวิเคราะห์ด้วยวิธีของ Scheffe จะมีความแม่นยำมากกว่าวิธีของ Tukey-Kramer[14]

Bonferroni

วิธีนี้คิดค้นโดย Dunn ซึ่งสร้างสูตรการคำนวณที่เรียกว่า Bonferroni สามารถใช้วิเคราะห์ได้ทั้งในกรณีที่ข้อมูลมีจำนวนกลุ่มตัวอย่างเท่ากัน หรือ ไม่เท่ากัน คำนวณโดยใช้ standard t – test[9] โดยจะทำการเปรียบเทียบความแตกต่างของค่าเฉลี่ยทีละคู่ และในการวิเคราะห์สามารถควบคุมค่า ในการทดสอบได้[14, 18] วิธีนี้เหมาะสมในการเปรียบเทียบข้อมูลที่มีจำนวนกลุ่มที่ต้องการเปรียบเทียบไม่เกิน 5 กลุ่ม นอกจากนี้หากมีจำนวนตัวอย่างที่นำมาวิเคราะห์ในแต่ละกลุ่มมากจะทำให้ผลการวิเคราะห์มีความเชื่อน่าถือต่ำ[10]

Duncan's Multiple Range

วิธีนี้คิดค้นโดย Duncan โดยการใช้อัตราส่วน q และการกระจายของอัตราส่วนดังกล่าว ประกอบกับคำนวณค่าสถิติ ซึ่งใช้ในการเปรียบเทียบข้อมูลที่มีจำนวนตัวอย่างในแต่ละกลุ่มเท่ากัน วิธีเหมาะสมจำนวนกลุ่มที่นำมาเปรียบเทียบไม่มากนักเนื่องจากไม่สามารถควบคุมความผิดพลาดชนิดที่ 1 ได้[12, 15]

Student-Newman Keuls (SNK)

ใช้ในการเปรียบเทียบข้อมูลที่มีจำนวนตัวอย่างในแต่ละกลุ่มเท่ากัน ในการเปรียบเทียบจะมีการนำค่า r (ช่วงห่างของค่าเฉลี่ยมาพิจารณาเปรียบเทียบด้วย) โดยการเปรียบเทียบตั้งแต่คู่ที่ต่างกันจากมากที่สุดก่อนไปเรื่อยๆ[4] และสามารถป้องกันความผิดพลาดประเภทที่ 2 ได้ดี[19] แต่ไม่สามารถควบคุมความผิดพลาดต่อหนึ่งการทดสอบ (Comparisonwise error rate) หรือ ความผิดพลาดต่อหนึ่งการทดลอง (Experimentwise error rate)[12, 14]

Games-Howell

ใช้วิเคราะห์กรณีที่มีค่าความแปรปรวนแต่ละกลุ่มไม่เท่ากัน ทำการเปรียบเทียบความแตกต่างโดยใช้ Studentized range test ข้อมูลที่นำมาวิเคราะห์ต้องมีขนาดตัวอย่างตั้ง 50 ตัวอย่างขึ้นไป[9]

แนวทางการเลือกวิธีการเปรียบเทียบพหุคูณ

ผู้เขียนได้เสนอแนะแนวทางในการเลือกวิธีการเปรียบเทียบโดยสามารถพิจารณาการเลือกใช้สถิติในการเปรียบเทียบพหุได้ตาม Figure 1

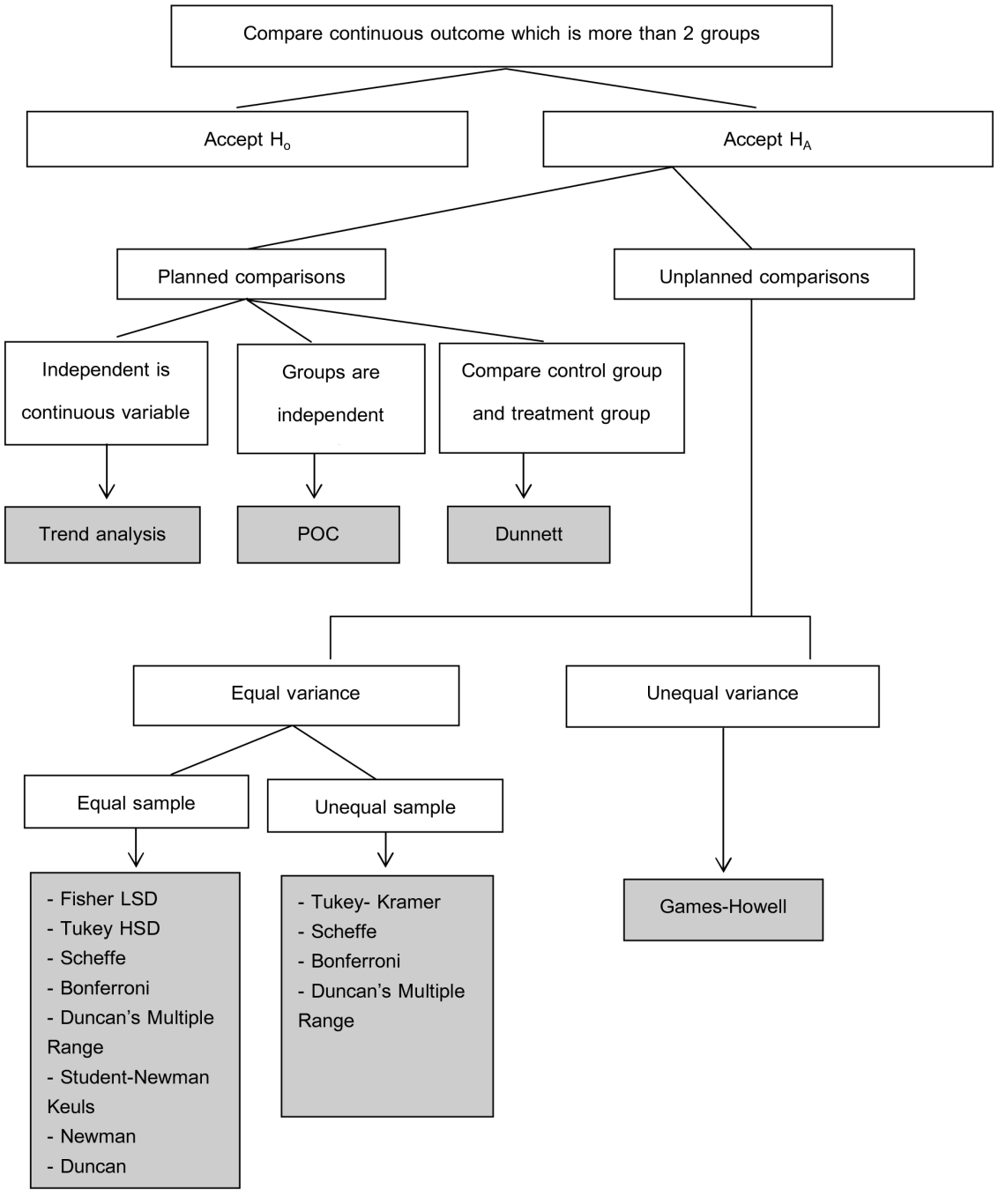

Figure 1 The procedure choosing multiple comparison which apply from Richard G.Lomax and Debbie L Hahs – Vaughm[9]

จากข้อมูลใน Table 1 ผู้วิจัยได้ทำการทดสอบข้อตกลงเบื้องต้นพบว่า ข้อมูลดังกล่าวอิสระต่อกัน มีค่าความแปรปรวนของประชากรแต่ละกลุ่มต้องเท่ากัน และตัวแปรตามแจกแจงแบบปกติ และนักวิจัยทำการทดสอบความแปรปรวนแล้วพบว่ายอมรับ H_A ซึ่งแสดงว่ามีค่าเฉลี่ยอย่างน้อย 1 กลุ่มแตกต่างกัน ดัง Table 2

Table 2 Compare protective performance scores of Hemorrhagic fever

Factor	Control Mean (SD)	Group 1 Mean (SD)	Group 2 Mean (SD)	F-test	P-value
Performance scores	14.9 (0.6)	17.3 (0.4)	17.6 (0.5)	9.04	0.001

ดังนั้นหากนักวิจัยต้องการทราบต่อไปว่าค่าเฉลี่ยคู่ใดบ้างที่แตกต่างกัน นักวิจัยต้องทำการวิเคราะห์ด้วยการเปรียบเทียบพหุคูณ โดยนักวิจัยสามารถพิจารณาเลือกวิธีเปรียบเทียบพหุคูณโดยพิจารณาจาก Figure 1 ประกอบได้ดังนี้

1. งานวิจัยนี้นักวิจัยไม่ทราบมาก่อนว่าค่าเฉลี่ยคู่ใดบ้างแตกต่างกัน ดังนั้นจึงจัดเป็นการเปรียบเทียบพหุคูณประเภทที่ไม่ได้มีการวางแผนการทดสอบ

2. พิจาณาต่อไปว่าความแปรปรวนเท่ากันหรือไม่

2.1) หากความแปรปรวนแต่ละกลุ่มไม่เท่ากัน ให้เลือกใช้วิธี Games-Howell

2.2) หากความแปรปรวนแต่ละกลุ่มเท่ากัน ให้พิจารณาต่อขั้นตอนที่ 3

3. ข้อมูลในแต่ละกลุ่มมีจำนวนกลุ่มตัวอย่างเท่ากันหรือไม่ จาก Table 1 พบว่าข้อมูลทั้ง 3 กลุ่มมีจำนวนตัวอย่างเท่ากัน เนื่องจากแต่ละกลุ่มมีจำนวนตัวอย่างกลุ่มละ 10 ตัวอย่าง ดังนั้นงานวิจัยนี้อาจเลือกใช้วิธี Fisher LSD, Tukey HSD, Scheffe, Bonferroni, Duncan's Multiple Range หรือ Student-Newman Keuls จากการวิเคราะห์ให้ผลการวิเคราะห์พหุคูณดัง Table 3

Table 3 Compare the result of multiple comparison (n = 30)

Post Hoc Comparison	p_value (95% CI)			Power full
	Control & group 1	Control & group 2	group 1& group 2	
Fisher LSD	0.002 (-0.972 : -3.828)	0.001 (-1.272 : -4.128)	0.670 (-1.728 : 1.128)	★★★★★
Tukey HSD	0.005 (-0.674 : -4.126)	0.002 (-0.974 : -4.426)	0.903 (1.426 : -2.026)	★★
Scheffe	0.007 (-0.597 : -4.203)	0.003 (-0.897 : -4.503)	0.912 (1.503 : -2.103)	★
Bonferroni	0.006 (-0.623 : -4.177)	0.002 (-0.923 : -4.477)	1.000 (1.477 : -2.077)	★★
Duncan's Multiple Range	0.002 (-3.828 : -0.972)	0.001 (-1.200 : -4.200)	0.670 (1.128 : -1.728)	★★★★★
Student-Newman Keuls	0.002 (-0.971 : -3.829)	0.002 (- 0.973 : -4.427)	0.670 (-1.729 : 1.129)	★★★

หมายเหตุ ★หมายถึงความสามารถในปฏิเสธสมมติฐานว่าง โดยพิจารณาการให้ ★จากความกว้างของช่วงเชื่อมั่น 95%

จากผลการเปรียบเทียบพหุคูณของข้อมูลจำนวน 30 ข้อมูล โดยใช้สถิติจำนวน 6 ตัวในการเปรียบเทียบค่าเฉลี่ยแต่ละกลุ่ม พบว่าวิธีดังกล่าวข้างต้นให้ผลการวิเคราะห์ให้ความน่าจะเป็นแตกต่างกันเล็กน้อย แต่ก็ยังนำไปสู่การสรุปการวิเคราะห์ที่เหมือนกัน คือ พบว่าคะแนนพฤติกรรมกลุ่มควบคุมและการได้รับความรู้จากเจ้าหน้าที่สาธารณสุข และ คะแนนพฤติกรรมกลุ่มควบคุมและการได้รับความรู้โดยการสร้างการมีส่วนร่วม มีคะแนนพฤติกรรมแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แต่ไม่พบความแตกต่างของคะแนนพฤติกรรมกลุ่มที่ได้รับความรู้จากเจ้าหน้าที่สาธารณสุขและการได้รับความรู้โดยการสร้างการมีส่วนร่วม จากการวิเคราะห์จะเห็นได้ว่าวิธีของ Fisher LSD และ Duncan's Multiple Range ให้ค่าความน่าจะเป็นในการทดสอบน้อยที่สุด และมีช่วงค่าความเชื่อมั่นแคบที่สุด ซึ่งจะมีโอกาสที่จะเกิดความคลาดเคลื่อนในการทดสอบสมมุติฐานน้อยกว่าวิธีอื่นๆ ช่วยให้นักวิจัยสามารถปฏิเสธสมมติฐานว่าง(null hypothesis) ได้อย่างมั่นใจมากขึ้น หากลำดับความสามารถในการปฏิเสธสมมติฐานว่างจะพบว่าวิธีของ (Fisher LSD = Duncan's Multiple Range) > Student-Newman Keuls > (Tukey HSD = Bonferroni) > Scheffe แต่การที่นักวิจัยจะเลือกสถิติตัวใดมาเปรียบเทียบพหุคูณ อาจพิจารณาจากวัตถุประสงค์ของนักวิจัยว่าต้องการป้องกันโอกาสที่จะเกิดความผิดพลาดในการทดสอบมากน้อยเพียงใด แต่ทั้งนี้ความถูกต้องของผลการวิจัยไม่ได้มากจากสถิติเพียงอย่างเดียว นักวิจัยต้องคำนึงถึงรูปแบบของการออกแบบงานวิจัยดังกล่าวร่วมด้วย

สรุป

One-way ANOVA เป็นสถิติอนุมานที่ใช้ในการวิเคราะห์ความแตกต่างของค่าเฉลี่ยของข้อมูลตั้งแต่ 3 กลุ่มขึ้นไป โดยตัวแปรตามที่ใช้ในการวิเคราะห์จะต้องมีลักษณะข้อมูลเป็นข้อมูลต่อเนื่อง หากการวิเคราะห์พบว่าข้อมูลมีค่าเฉลี่ยแตกต่างกันอย่างน้อย 1 คู่ และในการทดสอบนักวิจัยต้องการทราบว่าค่าเฉลี่ยคู่ใดบ้างมีความแตกต่างกัน นักวิจัยควรใช้วิธีวิเคราะห์พหุคูณ เนื่องจากสามารถควบคุมความผิดพลาดประเภทต่อหนึ่งกลุ่มการทดสอบ ซึ่งทำให้ลดการเกิดอัตราความคลาดเคลื่อนชนิดที่ 1 ซึ่งนำไปสู่ความถูกต้องของการสรุปผลการทดสอบ วิธีการเปรียบเทียบพหุด้วย Fisher LSD และ Duncan's Multiple Range ให้ประสิทธิภาพในการปฏิเสธสมมติฐานว่างได้มีประสิทธิภาพได้ดีกว่าวิธีอื่นๆ ทั้งนี้ข้อมูลต้องอิสระต่อกัน มีค่าความแปรปรวนของประชากรแต่ละกลุ่มต้องเท่ากัน และตัวแปรตามมีการแจกแจงแบบปกติ

บรรณานุกรม

1. สำนักโรคติดต่อนำโดยแมลง กรมควบคุมโรค. สถานการณ์โรคไข้เลือดออก. 12 พฤศจิกายน 2556; 2556 [cited 20 ตุลาคม 2557]; Available from: http://www.m-society.go.th/article_attach/ 11021/15445.pdf.
2. Bhopal RS. Concepts of Epidemiology: An integrated introduction to the ideas, theories, principles and methods of epidemiology. New York: Oxford University Press 2002.
3. Bernard HR. Social research method : qualitative and quantitative approaches Los Angeles: SAGE Publications; 2013.
4. Jerrold H. Zar. Biostatistical analysis. 4, editor. New Jersey: Prentice-Hall; 1999.
5. David G. Kleinbaum LLK, Keith E. Muller. Applied regression analysis and other multivariable methods / 2, editor. Boston: PWS-Kent Publishing Company; 1988.
6. Glenn Gamst LSMaAJG. Analysis of Variance Design. United Stated of America: Cambridge University; 2008.
7. David J. Weiss. Analysis of variance and functional measurement : a practical guide New York: Oxford University; 2006.
8. Weiss DJ. Analysis of variance and functional measurement : a practical guide New York: Oxford University Press; 2006.
9. Richard G.Lomax and Debbie L Hahs - Vaughm. Statistical Concept A Second Course. 4 ed. United States of America: Taylo & Francis Group; 2012.
10. Douglas G. Altman. Practical statistics for medical research. London: Chapman and Hall; 1991.
11. Jason C. Hsu. Multiple comparisons : theory and methods /. London: Chapman & Hall; 1996.
12. Gerry P. Quinn and Micheal J. Keough. Experimental Design and Data Analysis for Biologists. New York: Cambridge University; 2002.
13. Jason C. Hsu. Multiple Comparioson Theory and methods. London: Chaoman & Hall; 1996.
14. John A. Rafter MLAaJPB. Multiple Comparison Methods for Means. Society for Industrial and Applied Mathematics. 2002;44(2):259 - 78.

15. Guven Ozkaya and Ilker Ercan. Examining Multiple Comparison Procedures According to Error Rate, Power Type and False Discovery Rate. Journal of Modern Applied Statistical Methods. 2012;11(2): 348-60.
16. Michael R. Stoline. The Status of Multiple Comparisons: Simultaneous Estimation of All Pairwise Comparisons in One-Way ANOVA Designs. The American Statistician. 1981;3(35):134-41.
17. R. W. Day and G. P. Quinn. Comparisons of treatments after an analysis of variance in ecology. Ecological monographs. 1989;59(4):433-63.
18. Barbara G. Tabachnick and Linda S. Fidell. Experimental designs using ANOVA. Australia: Thomson Brooks; 2007.
19. R. A. Armstrong SVSaFE. An introduction to analysis of variance (ANOVA) with special reference to data from clinical experiments in optometry. Ophthal Physiol Opt. 2000;12:235-41.