บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

เป้าหมายสำคัญยิ่งของการวิจัยคือ การหาผลสรุปที่ถูกต้องตรงความเป็นจริง ฉะนั้น การดำเนินการวิจัยจึงต้องมีการวางแผนและตรวจสอบข้อมูลทุกขั้นตอน เพื่อป้องกันความคลาดเคลื่อนในการวิจัย และขั้นตอนที่สำคัญขั้นตอนหนึ่งคือ การวิเคราะห์ข้อมูล ซึ่งต้องใช้ระเบียบวิธีทางสถิติในการสรุปผลและตัดสินใจ โดยเฉพาะการอ้างอิงด้านการทดสอบสมมติฐาน เพื่อสรุปผลเกี่ยวกับค่าพารามิเตอร์ของประชากร ทั้งนี้เพราะงานวิจัยส่วนมากไม่สามารถที่จะศึกษาจากประชากรทั้งหมดได้โดยตรง เนื่องจากขนาดของประชากรที่ศึกษามักมีขนาดใหญ่มากในแต่ละโครงการวิจัย ในทางปฏิบัติจึงมักสุ่มตัวอย่างจากประชากรมาเพียงบางส่วน แล้วศึกษาข้อมูลจากตัวอย่างเพื่ออนุมานคุณสมบัติของประชากรจากคุณสมบัติของตัวอย่างที่ได้ ความถูกต้องและความเชื่อถือได้ในการคาดคะเนเกี่ยวกับค่าพารามิเตอร์ของประชากรจึงขึ้นอยู่กับการใช้สถิติทดสอบ สถิติทดสอบแต่ละวิธีมีข้อตกลงเบื้องต้นเฉพาะของวิธีนั้น ๆ ผู้วิจัยจะต้องเลือกใช้วิธีทดสอบทางสถิติที่เหมาะสม จึงจะทำให้ได้ผลสรุปที่มีความแม่นตรงสูง มีโอกาสผิดพลาดเพียงเท่ากับโอกาสของความคลาดเคลื่อนที่ผู้วิจัยยอมให้มีได้ในการทดสอบสมมติฐานซึ่งก็คือระดับนัยสำคัญที่กำหนดไว้

งานวิจัยที่ต้องการทดสอบความมีนัยสำคัญในค่าเฉลี่ยของประชากร หรือหาข้อสรุปว่าทรีทเมนต์ (treatment) ที่ให้ในแต่ละกลุ่มตัวอย่าง มีทรีทเมนต์ใดบ้างที่ให้ผลแตกต่างกัน สำหรับกรณีกลุ่มตัวอย่างมากกว่าสองกลุ่ม เทคนิคการทดสอบสมมติฐานทางสถิติที่นำมาใช้กันมาก คือ การทดสอบเอฟ (F-test) หรือนิยมเรียกกันว่าการวิเคราะห์ความแปรปรวน (ANOVA) ซึ่งมีข้อตกลงเบื้องต้นคือ กลุ่มตัวอย่างถูกสุ่มมาจากประชากรที่แจกแจงแบบปกติ ความแปรปรวนของประชากรแต่ละกลุ่มเท่ากัน และกลุ่มตัวอย่างแต่ละกลุ่มเป็นอิสระต่อกัน (Glass and Stanley 1970: 340) การทดสอบสมมติฐานด้วยสถิติทดสอบเอฟ เป็นการทดสอบรวม (Over all test) เมื่อการทดสอบเอฟมีนัยสำคัญ หมายถึงค่าเฉลี่ยอย่างน้อยหนึ่งกลุ่มจะไม่เท่ากับ

กลุ่มอื่น ๆ ซึ่งผู้วิจัยยังสรุปรายละเอียดไม่ได้ว่าผลจากทรีทเมนต์ใดหรือค่าเฉลี่ยของประชากรกลุ่มใดบ้างที่แตกต่างกัน ผู้วิจัยต้องตรวจสอบวิเคราะห์เพิ่มเติมด้วยการเปรียบเทียบพหุคูณ (Multiple Comparisons) ซึ่งเรียกว่าการเปรียบเทียบภายหลังหรือการวิเคราะห์ภายหลัง (Posteriori Comparisons or Post hoc Analysis)

โดยทั่วไปก่อนการวิเคราะห์ความแปรปรวนผู้วิจัยอาจตรวจสอบความแตกต่างของค่าเฉลี่ยแต่ละกลุ่มได้จำนวนหนึ่งตามที่ผู้วิจัยต้องการ การเปรียบเทียบพหุคูณในลักษณะนี้เรียกว่าการเปรียบเทียบภายแรกหรือการเปรียบเทียบตามแผน (Priori Comparisons or Planed Comparisons) อย่างไรก็ตามการเปรียบเทียบภายแรกมักมีข้อจำกัดในจำนวนการเปรียบเทียบกล่าวคือ ถ้าจำนวนการเปรียบเทียบมีมาก และการเปรียบเทียบไม่เป็นอิสระต่อกัน จะทำให้ช่วงความเชื่อมั่นกว้าง และอำนาจการทดสอบต่ำในกรณีเช่นนี้ผู้วิจัยมักจะเลือกใช้การเปรียบเทียบภายหลัง ซึ่งจะให้อำนาจการทดสอบสูงกว่า (Hays 1963: 488, Davis 1969: 441)

การเปรียบเทียบภายหลังนั้น นักสถิติได้เสนอและพัฒนาขึ้นมาหลายวิธี ซึ่งทุกวิธีมีข้อตกลงเบื้องต้นเช่นเดียวกันคือ

1. กลุ่มตัวอย่างถูกสุ่มมาจากประชากรที่แจกแจงปกติหรือกลุ่มตัวอย่างมีขนาดใหญ่
2. ความแปรปรวนของประชากรเท่ากัน
3. กลุ่มตัวอย่างแต่ละกลุ่มเป็นอิสระต่อกัน
4. การแจกแจงค่าสถิติทดสอบอยู่บนพื้นฐานของสมมติฐานสูญที่เป็นจริง (True Null Hypothesis)

Glass และ Stanley (1970: 383) กล่าวว่าการเปรียบเทียบภายหลังหลาย ๆ วิธีนั้น มี 2 วิธีที่เป็นประโยชน์มากที่สุดคือวิธีทีของทูกี้ (Tukey's T-Method) และวิธีเอสของเชเฟ (Scheffé's S-Method) ซึ่งสอดคล้องกับ Kirk (1982: 116, 121) ที่กล่าวว่าวิธีของทูกี้ใช้กันอย่างกว้างขวางมากในการทดสอบรายคู่ เมื่อกลุ่มตัวอย่างทุกกลุ่มมีขนาดเท่ากัน และวิธีของเชเฟ (Scheffé) มีความยืดหยุ่นสูง ใช้ได้กับการทดสอบที่เป็นรายคู่ (Pairwise) และไม่เป็นรายคู่ (Nonpairwise) ทั้งขนาดของกลุ่มตัวอย่างที่เท่ากันและไม่เท่ากัน และเป็นวิธีที่มีลักษณะอนุรักษ์นิยม (Conservative)

Miller (1966: 43-44) กล่าวถึงวิธีที่ควรเลือกใช้ในการเปรียบเทียบภายหลัง คือวิธีของทูกี้และวิธีของเชเฟ โดยให้เหตุผลว่า ทั้งสองวิธีมีความไว (Sensitivity) กว่าวิธีอื่น ๆ และช่วงความเชื่อมั่นแคบกว่าวิธีอื่น ๆ

สุญาณี จิตตะยโศธร (2525: 40) ได้ศึกษาอัตราความคลาดเคลื่อนประเภทที่ 1 จากผลการทดลอง (Actual Type I Error Rate) ของการเปรียบเทียบพหุคูณ 5 วิธี คือ วิธีของทูกี้ วิธีของดันน์ วิธีของเชเฟ วิธีของดันเนตต์ และวิธีทดสอบของนิวแมนคูลส์ โดยใช้กลุ่มตัวอย่างขนาด 5, 10 และ 15 เมื่อข้อมูลเป็นไปตามข้อตกลงเบื้องต้นพบว่า วิธีการเปรียบเทียบภายหลังที่สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากับอัตราความคลาดเคลื่อนที่ระบุ (α) คือวิธีของทูกี้และวิธีของเชเฟ ส่วนวิธีอื่นไม่สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากับอัตราความคลาดเคลื่อนที่ระบุ

ในการวิเคราะห์ความแปรปรวนด้วยสถิติทดสอบเอฟและการเปรียบเทียบภายหลังด้วยวิธีของทูกี้และวิธีของเชเฟ มักจะพบว่าข้อมูลที่นำมาวิเคราะห์ไม่เป็นไปตามข้อตกลงเบื้องต้นข้อใดข้อหนึ่ง ดังที่ Cochran และ Cox (1957: 91) กล่าวว่าในทางปฏิบัติเรามักไม่แน่ใจว่าข้อมูลที่ได้จะเป็นไปตามข้อตกลงเบื้องต้นทั้งหมดและบ่อยครั้งที่ข้อมูลไม่เป็นไปตามข้อตกลงบางข้อ สำหรับกรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงด้านความเป็นเอกพันธ์ของความแปรปรวน มีผู้ศึกษาไว้ดังนี้

Petrinovich และ Hardyck (1969: 53) กล่าวว่าเมื่อความแปรปรวนของประชากรไม่เท่ากันและขนาดของกลุ่มตัวอย่างแต่ละกลุ่มน้อยกว่า 10 เมื่อทำการเปรียบเทียบพหุคูณด้วยวิธีใด ๆ ก็ตามอำนาจการทดสอบจะต่ำมาก

สุญาณี จิตตะยโศธร (2525: 40-41) ได้ศึกษาเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 จากผลการทดลองของการเปรียบเทียบพหุคูณกรณีการเปรียบเทียบรายคู่ด้วยกลุ่มตัวอย่าง 3 กลุ่ม และขนาดกลุ่มตัวอย่างแต่ละกลุ่มเท่ากันคือขนาด 5, 10, 15 เมื่อเปรียบเทียบระหว่างวิธีของทูกี้และวิธีของเชเฟ กรณีฝ่าฝืนข้อตกลงเบื้องต้นด้านความเป็นเอกพันธ์ของความแปรปรวนโดยอัตราส่วนความแปรปรวนของประชากรเป็น 0.9 : 1.0 : 1.1 และ 0.8:1.0:1.2 พบว่า ที่ระดับอัตราความคลาดเคลื่อนที่ระบุเท่ากับ .05 วิธีของทูกี้สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เฉพาะกรณีกลุ่มตัวอย่างขนาด 10 ในขณะที่วิธีของเชเฟสามารถควบคุมได้ทุกกรณีการทดลอง ซึ่งสอดคล้องกับ Scheffé (1959: 77) ที่กล่าวว่าวิธีของเชเฟแกร่ง (robust) ต่อการฝ่าฝืนข้อตกลงเบื้องต้น

การใช้เทคนิคการเปรียบเทียบภายหลังต้องขึ้นกับการทดสอบเอฟ คือใช้การทดสอบเอฟเพื่อทดสอบความมีนัยสำคัญในค่าเฉลี่ยของประชากร เมื่อการทดสอบเอฟมีนัยสำคัญ จึงใช้วิธีการเปรียบเทียบภายหลังด้วยวิธีของทูกี้หรือวิธีของเชเฟ และการเปรียบเทียบภายหลังด้วยวิธีของเชเฟ ใช้การแจกแจงเอฟ (F-distribution) ในการทดสอบนัยสำคัญ ดังนั้นการทดสอบเอฟอาจมีผลต่อการเปรียบเทียบภายหลัง ดังที่ Petrinovich และ Hardyck (1969: 53) ได้สรุปผลจากการทดลองไว้ว่า การฝ่าฝืนข้อตกลงเบื้องต้นมีผลกระทบทำให้เกิดความแตกต่างระหว่างวิธีของทูกี้ และวิธีของเชเฟไม่มากนัก ยกเว้นการฝ่าฝืนข้อตกลงเบื้องต้นนั้นมีผลกระทบถึงการวิเคราะห์ความแปรปรวนด้วยสถิติทดสอบเอฟ โดยเฉพาะกรณีกลุ่มตัวอย่างมีขนาดไม่เท่ากัน และกลุ่มที่มีขนาดเล็กมีความแปรปรวนสูง กับกลุ่มขนาดใหญ่มีความแปรปรวนต่ำ ค่าระดับนัยสำคัญเอฟ (Significant F value) จะผิดจากความเป็นจริงมาก ทำให้อัตราความคลาดเคลื่อนประเภทที่ 1 ของการเปรียบเทียบภายหลังสูงขึ้นไปด้วย และถ้าข้อมูลเหมาะสมสำหรับการทดสอบเอฟ ก็จะเหมาะสมกับการเปรียบเทียบภายหลังด้วย

Tomarken และ Serlin (1986: 90-99) ได้ทำการศึกษาผลที่มีต่อการฝ่าฝืนข้อตกลงเบื้องต้นในด้านความเป็นเอกพันธ์ของความแปรปรวนของสถิติทดสอบเอฟ โดยศึกษากับกลุ่มตัวอย่างขนาดเท่ากัน คือ ขนาด 12 และ 20 เมื่อกำหนดจำนวนกลุ่มตัวอย่างเป็น 3 กลุ่ม และ 4 กลุ่ม กำหนดความแตกต่างของความแปรปรวนเป็นอัตราส่วนคือ 12 : 4 : 1 สำหรับกลุ่มตัวอย่าง 3 กลุ่ม และ 12 : 6 : 4 : 1 สำหรับกลุ่มตัวอย่าง 4 กลุ่ม พบว่า การทดสอบเอฟไม่สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ตามที่กำหนด เมื่อกลุ่มตัวอย่าง 3 กลุ่ม ทั้งระดับอัตราความคลาดเคลื่อนที่ระบุเท่ากับ .05 และ .01 และเมื่อกลุ่มตัวอย่างมี 4 กลุ่ม ที่ระดับอัตราความคลาดเคลื่อนที่ระบุเท่ากับ .01 โดยทุกกรณีที่ไม่สามารถควบคุมได้นี้อัตราความคลาดเคลื่อนประเภทที่ 1 มากกว่าที่กำหนด ซึ่งสอดคล้องกับผลสรุปจากการศึกษาของ Scheffé (1959: 354) ที่พบว่า กรณีกลุ่มตัวอย่างเท่ากัน เมื่อความแปรปรวนของประชากรแตกต่างกันอัตราความคลาดเคลื่อนประเภทที่ 1 จะมากกว่าอัตราความคลาดเคลื่อนที่ระบุ

Glass, Peckham และ Senders (1972: 245) ได้สรุปผลจากการศึกษาเมื่อความแปรปรวนแตกต่างกัน และกลุ่มตัวอย่างไม่เท่ากันของสถิติทดสอบเอฟ จากการศึกษาของ Hsu (1938), Scheffé (1959), Pratt (1964), Box (1954), Boneau (1960) van der Vaart (1961) ดังนี้

1. เมื่อกลุ่มตัวอย่างที่มีจำนวนน้อยถูกสุ่มมาจากประชากรที่มีความแปรปรวนมาก ระดับนัยสำคัญที่เป็นจริง จะมากกว่าระดับนัยสำคัญที่กำหนด (α)

2. เมื่อกลุ่มตัวอย่างที่มีจำนวนน้อยถูกสุ่มมาจากประชากรที่มีความแปรปรวนน้อยระดับนัยสำคัญที่เป็นจริงจะน้อยกว่าระดับนัยสำคัญที่กำหนด (α)

Kirk (1982: 77-78) กล่าวว่าการแจกแจง เอฟจะ เปลี่ยนแปลงได้ เมื่อฝ่าฝืนข้อตกลง เบื้องต้นด้านความ เป็น เอกพันธ์ของความแปรปรวน (Homogeneity of variance) สำหรับกลุ่มตัวอย่าง เท่ากัน เมื่อความแปรปรวนของประชากรแตกต่างกัน Rogan และ Keselman (1977) พบว่าระดับนัยสำคัญที่ เป็นจริงใหญ่กว่าระดับนัยสำคัญที่กำหนด สำหรับกรณีขนาดกลุ่มตัวอย่างไม่ เท่ากัน การฝ่าฝืนข้อตกลง เบื้องต้นข้อนี้มีผลอย่างมากต่อการกำหนดนัยสำคัญ

เมื่อความแปรปรวนของประชากรไม่ เท่ากัน มีผลทำให้การทดสอบ เอฟ เปลี่ยนแปลงไป และมีผลสืบ เนื่องถึงการ เปรียบ เทียบภายหลังดังที่ Federer (1955: 140) ได้กล่าวว่า แม้ว่าการทดสอบ เอฟ เพื่อตรวจสอบความมีนัยสำคัญ จำ เป็นต้องทำก่อนที่จะ เปรียบ เทียบภายหลัง แต่การทดสอบ เอฟก็อาจจะ เปลี่ยนแปลงคุณสมบัติบางประการของการ เปรียบ เทียบภายหลังทำให้ไม่ได้รับการตรวจสอบ ผลการตัดสินใจจึงผิดพลาด ซึ่งสอดคล้องกับที่ Games, Winkler และ Probert (1972: 887) กล่าวว่า ถ้าผู้วิจัยต้องการทำการ เปรียบ เทียบพหุคูณที่มีข้อตกลง เบื้องต้น เกี่ยวกับความ เท่ากันของความแปรปรวน ควรจะทำการทดสอบความ เท่ากันของความแปรปรวนก่อนการวิ เคราะห์ความแปรปรวน เพื่อมิให้ เกิดผลสรุปที่ผิดพลาดขึ้น ดังนั้น เมื่อผู้วิจัยต้องการทำการ เปรียบ เทียบพหุคูณในกรณีที่ความแปรปรวนของประชากรไม่ เท่ากัน ผู้วิจัยควร เลือกใช้ เทคนิคการวิ เคราะห์ความแปรปรวนวิธีอื่น ๆ ที่ไม่ฝ่าฝืนข้อตกลง เบื้องต้น ดังที่ Ferguson (1981: 245) กล่าวว่า เมื่อใช้สถิติทดสอบ เอฟ การฝ่าฝืนข้อตกลง เบื้องต้นด้านความ เป็น เอกพันธ์ของความแปรปรวนอาจมีผลอย่างมากต่อความแปรปรวนของทรีท เมนต์ที่ผู้วิจัยศึกษาในแต่ละกลุ่มตัวอย่าง และภายใต้ เงื่อนไขบางอย่างผู้วิจัยควร เลือกใช้สถิติทดสอบอื่น

Marascuilo (1966: 286) ได้ เสนอให้ใช้สถิติทดสอบยู (U-test) ในการทดสอบความมีนัยสำคัญของค่า เฉลี่ย เมื่อประชากรมากกว่า 2 กลุ่ม และข้อมูลไม่ เป็นไปตามข้อตกลง เบื้องต้น เกี่ยวกับความ เป็น เอกพันธ์ในความแปรปรวนของประชากร โดยที่การแจกแจงของสถิติทดสอบยูจะประมาณได้ด้วยการแจกแจงไคสแควร์ (Chi Square) ที่ชั้นความ เป็นอิสระ เท่ากับ K-1 เมื่อ K หมายถึงจำนวนกลุ่มตัวอย่าง และข้อตกลง เบื้องต้นของสถิติทดสอบยู (U-test) คือ

1. ประชากรมีการแจกแจงแบบปกติ หรือกลุ่มตัวอย่างมีขนาดใหญ่

2. กลุ่มตัวอย่างได้มาโดยการสุ่มจากประชากร

3. กลุ่มตัวอย่างแต่ละกลุ่ม เป็นอิสระต่อกัน

การเปรียบเทียบภายหลังการทดสอบด้วยสถิติทดสอบยูนั้น Marascuilo (1966: 286) เสนอให้ใช้ Marascuilo's χ^2 ซึ่งพัฒนามาในแนวคิดเดียวกับทฤษฎีของเชเฟ แต่ใช้การแจกแจงไคสแควร์ ที่ชั้นความเป็นอิสระ เท่ากับ K-1 แทนการแจกแจงเอฟในวิธีของเชเฟ ซึ่งการเปรียบเทียบภายหลังด้วย Marascuilo's χ^2 นั้น มีความยืดหยุ่น (Flexibility) ได้เช่นเดียวกับวิธีของเชเฟ คือใช้ได้กับการเปรียบเทียบที่เป็นรายคู่ และไม่เป็นรายคู่ ทั้งขนาดกลุ่มตัวอย่างที่เท่ากันและไม่เท่ากัน

Marascuilo (1966: 280) กล่าวว่าการเปรียบเทียบพหุคูณด้วย Marascuilo's χ^2 นั้นได้ทำการศึกษากับข้อมูลที่กลุ่มตัวอย่างมีขนาดใหญ่ สำหรับกลุ่มตัวอย่างขนาดเล็กน่าสนใจเป็นอย่างยิ่ง ที่จะทำการศึกษากลุ่มตัวอย่างด้วยเทคนิคมอนติคาร์โล

Tamhane (1977: Cited by Kirk 1982: 120-121) ได้เสนอวิธีการเปรียบเทียบภายหลังอีกวิธีหนึ่งที่ใช้กับการแจกแจงที (t-distribution) วิธีนี้ใช้เทคนิคของ Welch (1938) และใช้อสมการผลคูณของไซด์ดัคค์ (Šidák's multiplicative Inequality) การเปรียบเทียบภายหลังวิธีนี้เป็นไปตามข้อตกลงของการทดสอบที (t-test) ซึ่งเหมือนกับการทดสอบเอฟ ยกเว้นข้อตกลงเรื่องความเป็นเอกพันธ์ของความแปรปรวน

Tamhane (1979: 473) กล่าวว่าวิธีนี้ใช้เทคนิคของ Welch (1938) ทำให้ใช้ได้กับกรณีกลุ่มตัวอย่างมีขนาดเล็ก เมื่อความแปรปรวนแตกต่างกัน และการใช้อสมการผลคูณของไซค์ดัคค์ทำให้วิธีนี้เป็นวิธีอนุรักษ์นิยม (Conservative) และ Kirk (1982: 121) กล่าวว่าเมื่อความแปรปรวนของประชากรแตกต่างกันและสนใจเฉพาะการเปรียบเทียบรายคู่ วิธีนี้เป็นอีกวิธีหนึ่งที่ควรเลือกใช้ในการทดสอบนัยสำคัญ

นอกจากนี้มีวิธีการเปรียบเทียบภายแรกวิธีหนึ่งที่นิยมใช้กันมากโดยทั่ว ๆ ไป คือวิธีของบอนเฟอโรนี (Bonferroni Method) ซึ่งวิธีนี้เมื่อใช้กับการแจกแจงที Miller (1966: 67) เรียกว่าสถิติบอนเฟอโรนีที (Bonferroni t Statistics) และมีข้อตกลงเช่นเดียวกับการทดสอบที

Marascuilo และ Levin (1983: 47) กล่าวว่า สถิติบอนเฟอโรนีที สามารถใช้ทดสอบกับข้อมูลที่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบยู (U-test) นั่นคือ สถิติบอนเฟอโรนีทีแกร่ง (robust) เมื่อฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ของความแปรปรวน ซึ่งสอดคล้องกับที่ Myers (1973: 315) กล่าวว่า สถิติบอนเฟอโรนีที มีประโยชน์มากในสถานการณ์ที่ฝ่าฝืนข้อตกลงเบื้องต้น ซึ่งวิธีอื่น ๆ ไม่อาจใช้ประโยชน์ได้

สถิติบอนเฟอโรนีที เป็นเทคนิคที่มาจากอสมการบอนเฟอโรนี (Bonferroni Inequality) ที่กำหนดว่า โอกาสการเกิดอัตราความคลาดเคลื่อนประเภทที่ 1 อย่างน้อยหนึ่งการเปรียบเทียบในเซ็ตการเปรียบเทียบใด ๆ ที่เป็นอิสระหรือไม่เป็นอิสระต่อกัน (Orthogonal Contrasts or Nonorthogonal Contrasts) จะน้อยกว่าหรือเท่ากับ ผลรวมของอัตราความคลาดเคลื่อนที่กำหนดในแต่ละการเปรียบเทียบ (Games 1971: 550) ซึ่งอสมการบอนเฟอโรนีแสดงในรูปสัญลักษณ์คือ $\alpha_{EW} \leq \alpha_1 + \alpha_2 + \ldots + \alpha_m$ เมื่อ α_{EW} หมายถึงโอกาสการเกิดอัตราความคลาดเคลื่อนประเภทที่ 1 อย่างน้อยหนึ่งการเปรียบเทียบในเซ็ตการทดสอบทั้งหมด และ $\alpha_1, \alpha_2, \ldots, \alpha_m$ หมายถึงโอกาสการเกิดอัตราความคลาดเคลื่อนประเภทที่ 1 ในการเปรียบเทียบ (Contrast) ที่ 1, 2,...., m ตามลำดับ โดยที่ $\alpha_1 + \alpha_2 + \ldots + \alpha_m = \alpha$ เมื่อ α หมายถึงอัตราความคลาดเคลื่อนที่ระบุสำหรับการทดสอบการเปรียบเทียบทั้งเซ็ต หรือกล่าวอีกนัยหนึ่ง สถิติบอนเฟอโรนีทีเป็นวิธีเปรียบเทียบพหุคูณที่แบ่งอัตราความคลาดเคลื่อน (α) ให้กับการเปรียบเทียบแต่ละครั้ง และเมื่อรวมทุกครั้งแล้ว อัตราความคลาดเคลื่อนก็เท่ากับ α พอดี (Marascuilo and Levin 1983: 45)

ในการกำหนดอัตราความคลาดเคลื่อนประเภทที่ 1 เพื่อควบคุมแต่ละเซ็ตของการเปรียบเทียบนั้นโดยทั่วไปมักจะกำหนดให้เท่ากันทุกการเปรียบเทียบในเซ็ตนั้น ๆ (Morrison 1976: 33, Marascuilo and McSweeney 1977: 28-29) แต่บางกรณีผู้วิจัยอาจกำหนดให้แตกต่างกันได้ ถ้าผู้วิจัยมีเหตุผลหรือทฤษฎีสนับสนุนที่ทำให้เห็นว่า แต่ละการเปรียบเทียบนั้นมีความสำคัญแตกต่างกัน โดยที่การเปรียบเทียบใดมีความสำคัญมากผู้วิจัยก็กำหนดอัตราความคลาดเคลื่อนให้มาก เพื่อให้การทดสอบนั้นมีอำนาจการทดสอบสูง (Myers 1979: 299, Rosenthal and Rubin 1984: 1028-1033)

การกำหนดให้ทุกการเปรียบเทียบมีอัตราความคลาดเคลื่อนเท่ากันหรือทดสอบที่ระดับนัยสำคัญระดับเดียวกันนั้น ถ้าจำนวนการเปรียบเทียบมีมากระดับนัยสำคัญที่ทดสอบแต่ละการเปรียบเทียบจะน้อย ซึ่งจะทำให้อำนาจการทดสอบต่ำลงไปด้วย ดังนั้นการเปรียบเทียบพหุคูณด้วยสถิติบอนเฟอโรนีที จึงขึ้นอยู่กับจำนวนการเปรียบเทียบ ซึ่ง Perlmutter และ Myers (1973: 182) กล่าวว่า ช่วงความเชื่อมั่นของสถิติบอนเฟอโรนีที ขึ้นอยู่กับจำนวนการเปรียบเทียบ ถ้าจำนวนการเปรียบเทียบมีจำนวนน้อย ช่วงความเชื่อมั่นจะแคบถ้ามีจำนวนการเปรียบเทียบมากขึ้น ช่วงความเชื่อมั่นจะกว้างขึ้น ซึ่งสอดคล้องกับ Mudholkar และ Subbiah (1976: 429) ที่กล่าวว่า เมื่อจำนวนการเปรียบเทียบที่สนใจมีจำนวนน้อย ควรทำการเปรียบเทียบภายแรกด้วยสถิติบอนเฟอโรนีที ซึ่งเป็นสถิติที่มีความไวมากที่สุด และอนุรักษ์นิยม

Marascuilo และ McSweeney (1977: 28) กล่าวว่า ถ้าผู้วิจัยสนใจการเปรียบเทียบจำนวนหนึ่งที่ไม่มากกว่าจำนวน $\binom{k}{2}$ เมื่อ k คือจำนวนกลุ่มตัวอย่างสถิติบอนเฟอโรนีที ก็ยังคงมีความเหมาะสมที่จะใช้ทดสอบ

ในงานวิจัยที่ต้องการเปรียบเทียบค่าเฉลี่ยหรือต้องการสรุปผลเกี่ยวกับทรีทเมนต์นั้น ส่วนมากมักจะต้องการเปรียบเทียบความแตกต่างเป็นรายคู่เท่านั้น (Keppel 1982: 153, Ury 1976: 89, Perlmutter and Myers 1973: 183) โดยเฉพาะงานวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์ตัวแปรอิสระส่วนมากจะมีลักษณะเป็นตัวแปรแบบแยกประเภท (Categories) มากกว่าตัวแปรประเภทต่อเนื่อง เช่น เพศ แยกเป็นเพศชาย และเพศหญิง ทัศนคติ แยกเป็นเห็นด้วยมาก เห็นด้วยปานกลาง และไม่เห็นด้วย เป็นต้น (สุชาติ ประสิทธิรัฐสินธ์ และคนอื่น ๆ 2523: 1) ซึ่งจะแปลความหมายได้เมื่อเปรียบเทียบระหว่างกลุ่มหรือการเปรียบเทียบรายคู่เท่านั้น (Ferguson 1981: 295)

จากที่กล่าวมาข้างต้น จะเห็นว่าในงานวิจัยที่ต้องการเปรียบเทียบค่าเฉลี่ยนั้น นอกจากจะต้องเลือกใช้เทคนิคการวิเคราะห์ความแปรปรวนที่เหมาะสมกับข้อมูลที่ศึกษาแล้ว ยังจะต้องพบปัญหาการเลือกวิธีเปรียบเทียบพหุคูณอีกด้วยคือ การเปรียบเทียบพหุคูณด้วยสถิติบอนเฟอโรนีที ไคสแควร์ของมาร์ชูโล และวิธีของทัมฮาเนน์ วิธีใดสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ตามที่กำหนด เมื่อกำหนดให้ความแปรปรวนของประชากรแตกต่างกัน และเปรียบเทียบเฉพาะรายคู่ที่กำหนดขนาดกลุ่มตัวอย่างเท่ากันทุกกรณี คือทั้งกลุ่มตัวอย่างขนาดเล็ก ขนาดกลาง และขนาดใหญ่ และกำหนดจำนวนกลุ่มตัวอย่างเป็น 3, 4 และ 5 กลุ่ม ซึ่งสามารถศึกษาได้โดยวิธี

ชิมูเลท (Simulate) ซึ่งจะให้ผลสรุปที่เด่นชัดภายใต้สถานการณ์คล้ายการทดลองคือสามารถระบุ หรือจำกัดขนาดของกลุ่มตัวอย่างและลักษณะการแจกแจงของประชากรได้ด้วย เครื่องคอมพิวเตอร์ โดยเทคนิคที่เรียกว่า มอนติคาร์โล (Monte Carlo)

<u>วัตถุประสงค์ของการวิจัย</u>

การวิจัยครั้งนี้ต้องการ เปรียบ เทียบและหาผลสรุปลักษณะของอัตราความคลาด เคลื่อนประเภทที่ 1 จากผลการทดลอง (Actual Type I Error Rate) ของวิธีเปรียบ เทียบพหุคูณ 3 วิธีคือ

1. สถิติบอน เฟอโรนีที (Bonferroni t Statistic)
2. ไคสแควร์ของมาร์ชูไล (Marascuilo's χ^2)
3. วิธีของทัมฮานน์ (Tamhane Procedure)

ภายใต้เงื่อนไขดังต่อไปนี้

1. ลักษณะการแจกแจงของประชากร เป็นแบบปกติ (Normal Distribution)
2. ความแปรปรวนของประชากร (Population Variance) ทั้งที่กำหนดให้เท่ากันและแตกต่างกัน
3. จำนวนกลุ่มตัวอย่าง 3, 4 และ 5 กลุ่ม
4. ขนาดกลุ่มตัวอย่างทุกกลุ่มมีขนาด เท่ากัน

<u>สมมติฐานการวิจัย</u>

วิธี เปรียบ เทียบพหุคูณ สถิติบอน เฟอโรนีที ไคสแควร์ของมาร์ชูไล และวิธีของทัมฮานน์สามารถควบคุมอัตราความคลาด เคลื่อนประเภทที่ 1 ได้เท่ากับอัตราความคลาด เคลื่อนที่ระบุ

<u>ข้อตกลงเบื้องต้น</u>

1. การวิจัยครั้งนี้ถือว่า อัตราความคลาด เคลื่อนประเภทที่ 1 จากผลการทดลอง เป็นดัชนีที่จะใช้เป็น เกณฑ์ในการ เลือกใช้วิธีการ เปรียบ เทียบพหุคูณ
2. การวิจัยครั้งนี้ เลือกใช้โปรแกรมย่อยสัปรูทีน (Subroutine Subprogram) ที่มีหลักฐานและการศึกษาในรูปการแปลงข้อมูลไปยังประชากรที่ต้องการวิจัย

ขอบเขตการวิจัย

1. การวิจัยครั้งนี้ศึกษาตัวแปรดังต่อไปนี้

1.1 ตัวแปรอิสระ (Independent variable) คือขนาดกลุ่มตัวอย่าง ความแปรปรวนของประชากร จำนวนกลุ่มตัวอย่าง และวิธีเปรียบเทียบพหุคูณ 3 วิธี

1.2 ตัวแปรตาม (Dependent variable) คือ อัตราความคลาดเคลื่อนประเภทที่ 1 จากผลการทดลอง ของวิธีเปรียบเทียบพหุคูณ 3 วิธี คือ สถิติบอนเฟอโรนีที ไคสแควร์ของมาร์ชูโลและวิธีของทัมฮานน์ โดยคำนวณเปรียบเทียบกับอัตราความคลาดเคลื่อนที่ระบุ (Nominated α)

2. ศึกษาความสามารถในการควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 จากผลการทดลองของสถิติบอนเฟอโรนีที ไคสแควร์ของมาร์ชูโล และวิธีของทัมฮานน์ เพื่อเปรียบเทียบกับอัตราความคลาดเคลื่อนที่ระบุ เฉพาะการแจกแจงของประชากรแบบปกติ

3. ศึกษากรณีจำนวนกลุ่มตัวอย่างเป็น 3, 4 และ 5 กลุ่ม ขนาดกลุ่มตัวอย่างเท่ากันทุกกลุ่ม คือขนาด 5, 10, 15, 20, 25 และ 30

4. กำหนดค่าพารามิเตอร์ $\mu = 500$, $\sigma^2 = 100$ สำหรับค่าความแปรปรวนที่น้อยที่สุด และจำกัดขนาดของความแปรปรวนดังนี้

4.1 กรณีกลุ่มตัวอย่าง 3 กลุ่มมี 3 รูปแบบ คือ

4.1.1 $\sigma_1^2 = 100$, $\sigma_2^2 = 100$, $\sigma_3^2 = 100$ (ตามแผนการทดลองลักษณะ 1 : 1 : 1)

4.1.2 $\sigma_1^2 = 100$, $\sigma_2^2 = 200$, $\sigma_3^2 = 300$ (ตามแผนการทดลองลักษณะ 1 : 2 : 3)

4.1.3 $\sigma_1^2 = 100$, $\sigma_2^2 = 400$, $\sigma_3^2 = 700$ (ตามแผนการทดลองลักษณะ 1 : 4 : 7)

4.2 กรณีกลุ่มตัวอย่าง 4 กลุ่มมี 3 รูปแบบคือ

4.2.1 $\sigma_1^2 = 100$, $\sigma_2^2 = 100$, $\sigma_3^2 = 100$, $\sigma_4^2 = 100$
(ตามแผนการทดลองลักษณะ 1 : 1 : 1 : 1)

4.2.2 $\sigma_1^2 = 100$, $\sigma_2^2 = 200$, $\sigma_3^2 = 300$, $\sigma_4^2 = 400$
(ตามแผนการทดลองลักษณะ 1 : 2 : 3 : 4)

4.2.3 $\sigma_1^2 = 100, \sigma_2^2 = 400, \sigma_3^2 = 700, \sigma_4^2 = 1000$

(ตามแผนการทดลองลักษณะ 1 : 4 : 7 : 10)

4.3 กรณีกลุ่มตัวอย่าง 5 กลุ่ม มี 3 รูปแบบ คือ

4.3.1 $\sigma_1^2 = 100, \sigma_2^2 = 100, \sigma_3^2 = 100, \sigma_4^2 = 100, \sigma_5^2 = 100$

(ตามแผนการทดลองลักษณะ 1 : 1 : 1 : 1 : 1)

4.3.2 $\sigma_1^2 = 100, \sigma_2^2 = 100, \sigma_3^2 = 200, \sigma_4^2 = 200, \sigma_5^2 = 300$

(ตามแผนการทดลองลักษณะ 1 : 1 : 2 : 2 : 3)

4.3.3 $\sigma_1^2 = 100, \sigma_2^2 = 100, \sigma_3^2 = 400, \sigma_4^2 = 400, \sigma_5^2 = 700$

(ตามแผนการทดลองลักษณะ 1 : 1 : 4 : 4 : 7)

5. ศึกษาเฉพาะการเปรียบเทียบรายคู่ (pairwise)

6. ศึกษาเฉพาะกรณีสถิติบอนเฟอโรนีที เฉลี่ยระดับนัยสำคัญเท่า ๆ กันทุกการเปรียบเทียบรายคู่

7. เปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ของสถิติบอนเฟอโรนีที ไคสแควร์ของมาร์ชูโล และวิธีของทัมฮานน์ เมื่อ α เป็น .05 และ .01

8. การทดลองจะกระทำซ้ำ 4,000 ครั้ง ในแต่ละสถานการณ์การทดลอง

9. สมมติฐานที่ทดสอบเป็นสมมติฐานที่ไม่กำหนดทิศทาง

<u>คำจำกัดความ</u>

ความแกร่งหมายถึง คุณสมบัติของการทดสอบทางสถิติที่ไม่แสดงถึงความไว (Sensitive) ต่อการเบี่ยงเบนหรือ การผ่าฝืนข้อตกลงเบื้องต้นของการทดสอบที่มีผลต่ออัตราการเกิดความคลาดเคลื่อนประเภทที่ 1 และความคลาดเคลื่อนประเภทที่ 2

ความคลาดเคลื่อนประเภทที่ 1 หมายถึง ความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมติฐานสูญ เมื่อสมมติฐานสูญนั้นถูกต้อง โอกาสที่จะทำให้เกิดความคลาดเคลื่อนประเภทที่ 1 กำหนดด้วย α

อัตราความคลาดเคลื่อนประเภทที่ 1 จากผลการทดลอง หมายถึง อัตราความคลาดเคลื่อนประเภทที่ 1 ที่เกิดขึ้นจริงในการทดสอบสมมติฐานของการทดสอบตามแผนการทดลองครั้งนี้

อัตราความคลาดเคลื่อนที่ระบุ หมายถึง อัตราความคลาดเคลื่อนประเภทที่ 1 ที่ควบคุมด้วยระดับนัยสำคัญ α ซึ่งผู้วิจัยเป็นผู้กำหนดขึ้น

ไคสแควร์ของมาร์ซูโล หมายถึง วิธีเปรียบเทียบพหุคูณ สำหรับการทดสอบความมีนัยสำคัญของความแตกต่างของค่าเฉลี่ยของประชากร ค่าวิกฤตของวิธีเปรียบเทียบพหุคูณนี้ได้จากรากที่สองของค่าไคสแควร์ที่อ่านจากตาราง

<u>ประโยชน์ของการวิจัย</u>

เพื่อช่วยให้ผู้ใช้สถิติเลือกสถิติทดสอบหรือวิธีเปรียบเทียบพหุคูณที่เหมาะสมกับข้อมูลที่นำมาศึกษา ซึ่งจะทำให้ได้ผลสรุปที่มีความแม่นตรงสูง มีโอกาสคลาดเคลื่อนเท่ากับโอกาสที่ผู้วิจัยกำหนดในการทดสอบสมมติฐาน