บทที่ 2

สถิติทดสอบและผลงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้จะทำการศึกษาเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของสถิติทดสอบ 3 วิธี คือ สถิติทดสอบแบบ Brown and Forsythe สถิติทดสอบแบบ Marascuilo และสถิติทดสอบแบบ ANOVA F-TEST ภายใต้ลักษณะความแปรปรวนของประชากรที่กำหนด และปัจจัยต่าง ๆ ที่คาดว่าจะมีผลต่อการศึกษาดังได้กล่าวไว้ในขอบเขตของการวิจัยในบทที่ 1 แล้ว สำหรับสมมติฐานของการทดสอบความเท่ากันของค่าเฉลี่ยประชากร ในกรณีที่มีหลายประชากรมีรูปแบบดังนี้

H_o : $\mu_1 = \mu_2 = \ldots = \mu_c$

H_1 : มีค่าเฉลี่ยอย่างน้อย 1 คู่ที่แตกต่างกัน

โดยที่ μ_i แทนค่าเฉลี่ยของประชากรที่ i, $i = 1, 2, 3, \ldots c$

สำหรับการวิจัยครั้งนี้ กำหนดจำนวนประชากร (c) เท่ากับ 3 และ 6 ซึ่งในบทนี้จะกล่าวถึงรายละเอียดของสถิติทดสอบแต่ละวิธี ลักษณะการแจกแจงของประชากร รวมทั้งนำเสนอผลงานวิจัยที่เกี่ยวข้อง ดังรายละเอียดต่อไปนี้

2.1 สถิติทดสอบที่ใช้ในการวิจัยครั้งนี้

2.1.1 การทดสอบแบบ Brown and Forsythe (Brown and Forsythe Test) เสนอโดย Morton B. Brown and Alan B. Forsythe ในปี ค.ศ. 1974 เพื่อใช้เป็นวิธีการทดสอบความเท่ากันของค่าเฉลี่ยประชากรที่มากกว่า 2 กลุ่ม ในกรณีที่ลักษณะความแปรปรวนของประชากรไม่เท่ากัน ซึ่งมีรูปแบบดังนี้

$$BF = \frac{\sum_{i=1}^{c} n_i(\bar{X}_i - \bar{X})^2}{\sum_{i=1}^{c} (1 - n_i/N)\, S_i^2}$$

โดยที่ X_{ij} แทนค่าสังเกตที่ j ในกลุ่มตัวอย่างที่ i

$(j = 1,2,\ldots,n_i)$, $(i = 1,2,\ldots,c)$

$$N = \sum_{i=1}^{c} n_i$$

$$\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$$

$$\bar{X} = \sum_{i=1}^{c} n_i\bar{X}_i/N$$

$$S_i^2 = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2/(n_i-1)$$

ภายใต้สมมติฐานหลัก (H_o) ที่เป็นจริง กล่าวคือ $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_c$ แล้ว สถิติทดสอบแบบ Brown and Forsythe จะมีลักษณะการแจกแจงแบบเอฟ โดยประมาณ (Approximate F Distribution) ที่องศาความเป็นอิสระ $(c-1, f^*)$ โดยที่

$$f^* = \left[\sum_{i=1}^{c} v_i^2 /(n_i-1)\right]^{-1}$$

$$v_i = (1-n_i/N)\, S_i^2 \Big/ \left[\sum_{i=1}^{c} (1-n_i/N)\, S_i^2\right]$$

เกณฑ์การตัดสินใจจะปฏิเสธ H_o เมื่อ $BF > F_{\alpha,(c-1,\ f^*)}$ โดยที่ $F_{\alpha,(c-1,\ f^*)}$ คือค่าวิกฤตที่เปิดจากตาราง F ที่ระดับนัยสำคัญ α และองศาความเป็นอิสระ $(c-1,\ f^*)$ ตามลำดับ

2.1.2 การทดสอบแบบ Marascuilo (Marascuilo Test) เสนอโดย L.A. Marascuilo ในปี ค.ศ. 1987 ซึ่งได้ดัดแปลงจากการทดสอบของ Welch ทำให้มีรูปแบบการคำนวณที่ง่ายขึ้น ซึ่งมีรูปแบบดังนี้

$$M = \sum_{i=1}^{c} w_i (\bar{X}_i - \tilde{X})^2 / (c-1)$$

โดยที่ $w_i = n_i/S_i^2$

$$W = \sum_{i=1}^{c} w_i$$

$$\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$$

$$\tilde{X} = \sum_{i=1}^{c} w_i \bar{X}_i / W$$

$$S_i^2 = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2/(n_i-1)$$

ภายใต้สมมติฐานหลักที่เป็นจริง กล่าวคือ $H_o : \mu_1 = \mu_2 = \ldots = \mu_c$ แล้ว M จะมีลักษณะการแจกแจงแบบเอฟโดยประมาณ (Approximate F Distribution) ที่องศาความเป็นอิสระ (c-1, f) โดยที่

$$f = \left[(3/(c^2-1) \sum_{i=1}^{c} (1-w_i/W)^2/(n_i-1)\right]^{-1}$$

เกณฑ์การตัดสินใจ จะปฏิเสธ H_o เมื่อ $M > F_{\alpha}$, (c-1, f) โดยที่ F_{α} , (c-1 f) คือ ค่าวิกฤตที่เปิดตาราง F ที่ระดับนัยสำคัญ α และองศาความเป็นอิสระ (c-1, f) ตามลำดับ

2.1.3 การวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA F-TEST) เป็นวิธีการจำแนกความแปรปรวนทั้งหมดของกลุ่มข้อมูลที่ได้มาจากการวางแผนการทดลองไม่ว่าจะเป็นแผนการทดสอบแบบ Completely Randomized Design (CRD), Randomized Completely Block Design (RBD) หรือการทดลองแบบแฟคทอเรียล (Factorial Experiment) ก็ตาม โดยวิธีการของ ANOVA จะทำการจำแนกความแปรปรวนออกเป็นส่วนต่าง ๆ ซึ่งจะทำให้เราสามารถบอกถึงขนาดของความแปรปรวนจากแต่ละแหล่งว่า เป็นเท่าไรต่อจำนวนความแปรปรวนทั้งหมด

การวิเคราะห์ความแปรปรวน (ANOVA) ได้ถูกนำไปใช้กับงานทดลองแขนงต่าง ๆ มากมาย ซึ่งในแต่ละงานทดลองที่ผู้วิจัยต้องการศึกษานั้นผู้วิจัยจะต้องเข้าใจถึงวัตถุประสงค์ของการวิจัย เพื่อที่จะได้เลือกแผนการทดลองที่เหมาะสม ซึ่งจะนำไปสู่การตอบปัญหาที่สงสัยได้ ดังนั้นก่อนที่ผู้วิจัยจะลงมือทำการทดลอง ผู้วิจัยจะต้องระบุแหล่งที่มาของความแปรปรวน ทั้งนี้เพื่อที่จะได้เลือกแผนการทดลองที่สามารถทำให้ผู้วิจัยวัดขนาดของความแปรปรวนของแต่ละแหล่งต่อจำนวนความแปรปรวนทั้งหมดได้ยกตัวอย่างเช่น ในการทดสอบเพื่อดูว่า คุณภาพของสูตรอาหาร

4 ชนิดต่าง ๆ กัน จะมีผลทำให้น้ำหนักการเจริญเติบโตของสุกรแตกต่างกันหรือไม่นั้น ถ้าสมมติว่าหน่วยทดลองของเรามีความคล้ายคลึงกัน ผู้วิจัยอาจเลือกแผนการทดลองแบบ CRD ได้ (รายละเอียดเกี่ยวกับการวางแผนการทดลอง ศึกษาได้จากหนังสือ "Experimental Design" โดย Cochran and Cox (1957) และหนังสือ "Planning of Experiment" โดย Cox (1958) จากนั้นใช้การวิเคราะห์ความแปรปรวนวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากรการใช้ ANOVA เพื่อวิเคราะห์ข้อมูลมีวัตถุประสงค์สำคัญสองประการ คือ

1. เพื่อคาดประมาณและทดสอบสมมติฐานเกี่ยวกับค่าความแปรปรวนของประชากร
2. เพื่อคาดประมาณและทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร

สำหรับงานวิจัยครั้งนี้ จะพิจารณาถึงการใช้ ANOVA วิเคราะห์ข้อมูลเพื่อตอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร แต่อย่างไรก็ตามข้อสรุปเกี่ยวกับค่าเฉลี่ยประชากรก็จะต้องพิจารณาควบคู่ไปกับค่าความแปรปรวนของประชากรด้วย การวิเคราะห์ข้อมูลด้วย ANOVA จะต้องมีข้อสมมติบางอย่าง ซึ่งโดยปกติในการทดลองทั่ว ๆ ไปข้อมูลที่ได้มักจะมีคุณสมบัติไม่ครบตามต้องการ ด้วยเหตุนี้เองผู้ใช้เทคนิคของ ANOVA ควรจะตระหนักถึงข้อสมมติที่มีอยู่ เพราะผลลัพธ์ที่ได้อาจผิดพลาด หากข้อมูลไม่เป็นไปตามข้อสมมติ อย่างไรก็ตาม Cochran ได้ตั้งข้อสังเกตไว้ว่า ผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูลด้วย ANOVA น่าจะเป็นผลลัพธ์โดยประมาณมากกว่าที่จะเป็นผลลัพธ์ที่ถูกต้องแน่นอน

ในการวิจัยครั้งนี้ จะพิจารณาถึง การวิเคราะห์ความแปรปรวนสำหรับแผนการทดลองแบบสุ่มสมบูรณ์ (CRD) ซึ่งหน่วยทดลองทั้งหมดจะต้องมีความคล้ายคลึงกัน (Homogeneous) จากนั้นแบ่งออกเป็นกลุ่มต่าง ๆ แล้วให้วิธีการ (Treatment) แต่ละวิธีที่ต้องการทดสอบกับหน่วยทดลองภายในแต่ละกลุ่ม ดังนั้นเราสามารถแบ่งความผันแปรออกเป็น 2 แหล่งด้วยกัน คือ ความผันแปรภายในกลุ่ม อันเนื่องมาจากหน่วยทดลองที่มีอยู่อาจไม่สม่ำเสมอกันจริง และความผันแปรระหว่างกลุ่มอันเนื่องมาจากอิทธิพลของวิธีการ (Treatment) ที่ให้ ซึ่งสามารถสรุปความผันแปรทั้งหมดได้ดังตารางที่ 2.1 ดังนี้

ตารางที่ 2.1 การวิเคราะห์ความแปรปรวน สำหรับแผนการทดลองแบบสุ่มสมบูรณ์

สาเหตุของความแปรปรวน	ผลบวกกำลังสอง	ระดับความเป็นเสรี	ผลบวกกำลังสองเฉลี่ย	สถิติทดสอบ
ระหว่างกลุ่ม	$SS_B = \sum_{i=1}^{c} n_i(\bar{x}_i - \bar{x})^2$	c-1	$MS_B = SS_B/(c-1)$	$AF = \frac{MS_B}{MS_w}$
ภายในกลุ่ม	$SS_w = \sum_{i=1}^{c} (n_i-1)\, s_i^2$	N-c	$MS_w = SS_w/(N-c)$	
ผลรวม	$SS_T = \sum_{i=1}^{c} \sum_{j=1}^{n_i} (x_{ij}-\bar{x})^2$	N-1		

สัญลักษณ์ที่ใช้ในตาราง

X_{ij} แทนค่าสังเกตที่ j ในกลุ่มตัวอย่างที่ i
$(j = 1,2,..,n_i)$, $(i = 1,2,...,c)$

$$N = \sum_{i=1}^{c} n_i$$

$$\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$$

$$\bar{X} = \sum_{i=1}^{c} n_i\bar{X}_i/N$$

$$S_i^2 = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2/(n_i-1)$$

ดังนั้นสถิติทดสอบแบบ ANOVA F-TEST ที่ได้ คือ

$$AF = \frac{\sum_{i=1}^{c} n_i(\bar{X}_i-\bar{X})^2/(c-1)}{\sum_{i=1}^{c} (n_i-1)\ S_i^2/(N-c)}$$

ภายใต้สมมติฐานหลัก (H_o) ที่เป็นจริง กล่าวคือ $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_c$ แล้ว สถิติทดสอบแบบ ANOVA F-TEST จะมีลักษณะการแจกแจงแบบเอฟ (F-Distribution) ที่องศาความเป็นอิสระ (c-1, N-c)

เกณฑ์การตัดสินใจจะปฏิเสธ H_o เมื่อ $AF > F_{\alpha, (c-1, N-c)}$

$F_{\alpha,(c-1, N-c)}$ คือ ค่าวิกฤตที่เปิดได้จากตาราง F ที่ระดับนัยสำคัญ α และองศาความเป็นอิสระ (c-1, N-c)ตามลำดับ

2.2 <u>คุณสมบัติและลักษณะการแจกแจงของประชากรที่ศึกษา</u>

การแจกแจงปกติ (Normal Distribution) เป็นการแจกแจงที่ใช้อธิบายข้อมูลที่เกิดขึ้นโดยทั่วไปทางธรรมชาติ เช่น คะแนนสอบ หรือยอดขายสินค้าชนิดหนึ่งในรอบ 6 เดือนที่ผ่านมา เป็นต้น การแจกแจงแบบปกติค้นพบ โดย เดอมูวร์ (De Movire : 1667-1754)

นักคณิตศาสตร์ชาวฝรั่งเศส ต่อมาลาปลาซ (Laplace : 1749-1827) นักคณิตศาสตร์ชาวอังกฤษได้นำมาประยุกต์ใช้ทางด้านสังคมศาสตร์และวิทยาศาสตร์ ต่อมานักคณิตศาสตร์เยอรมัน เกาส์ (Gauss : 177-1855) ได้นำทฤษฎีนี้ไปศึกษาหาความคลาดเคลื่อน (Error) โดยการวัดซ้ำ ๆ ในกลุ่มที่มีขนาดคงเดิม และพบว่าการแจกแจงที่ได้จะมีลักษณะเป็นโค้งปกติ ดังนั้นการแจกแจงแบบปกติอาจเรียกอีกอย่างหนึ่งว่า Gaussian Distribution โดยมีฟังก์ชั่นความหนาแน่น (pdf) ดังนี้

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-(x-\mu)^2 / 2\sigma^2\right] \quad ; \quad -\infty < x < \infty$$

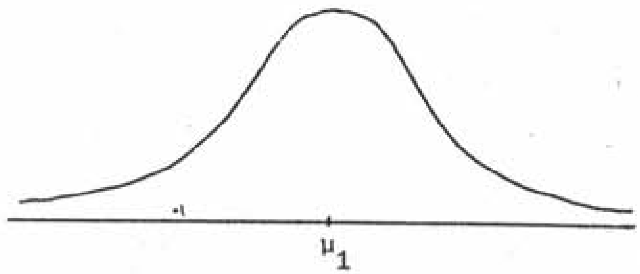
โดยที่ x เป็นค่าของข้อมูลจากกลุ่มตัวอย่าง

μ , σ^2 เป็นค่าเฉลี่ยและความแปรปรวนของประชากรตามลำดับ

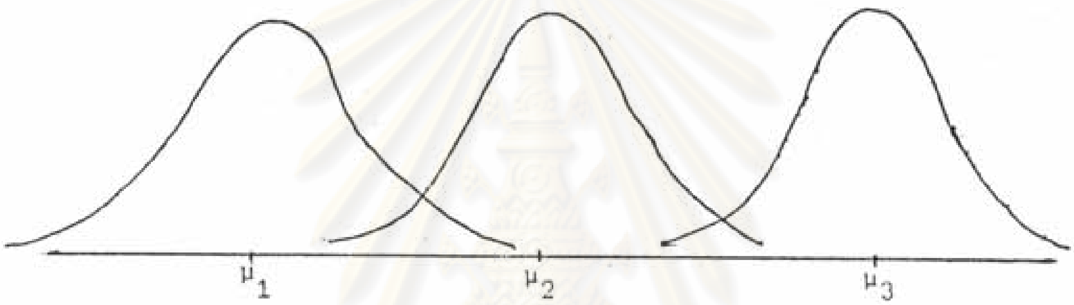
การแจกแจงแบบปกติมีลักษณะและคุณสมบัติดังนี้

1. ลักษณะของโค้งเป็นรูประฆังคว่ำ (Bell Shaped)
2. เส้นแบ่งครึ่งโค้งอยู่ที่จุดซึ่งเป็นค่าเฉลี่ยของข้อมูล และเส้นนี้แบ่งพื้นที่ทั้ง 2 ข้างสมมาตรกัน (Symmetry)
3. มีค่าเฉลี่ย มัธยฐาน และฐานนิยมเป็นจุดเดียวกัน หรือมีค่าเท่ากันนั่นเอง
4. มีความโด่ง (Kurtosis) ของเส้นโค้งเท่ากับ 3 ซึ่งเรียกว่า เมโซเคอร์ติค (Mesokurtic) และจุดเปลี่ยนโค้งทั้ง 2 ทั้งจะอยู่ ณ ตรง 1 ส่วนเบี่ยงเบนมาตรฐาน
5. ค่าความเบ้ (Skewness) เท่ากับศูนย์
6. ปลายทั้งสองข้างของเส้นโค้งจะค่อย ๆ ลดต่ำลงแต่ไม่จรดกับฐานของโค้งหรือแกนนอน
7. ถ้าลากเส้นตั้งฉากจากแกน x (ซึ่งเป็นแกนนอน) ไปยังเส้นโค้ง โดยที่เส้นดังกล่าวห่างจากค่าเฉลี่ย (μ) ทั้งด้านซ้ายและด้านขวาของระยะหนึ่งเท่า สองเท่าและสามเท่าของส่วนเบี่ยงเบนมาตรฐาน (σ) แล้ว พื้นที่ที่ปิดกั้นเส้นตั้งฉากกับเส้นโค้งจะเท่ากับ 68.2%, 95.4% และ 99.7% ของพื้นที่ทั้งหมด ตามลำดับ

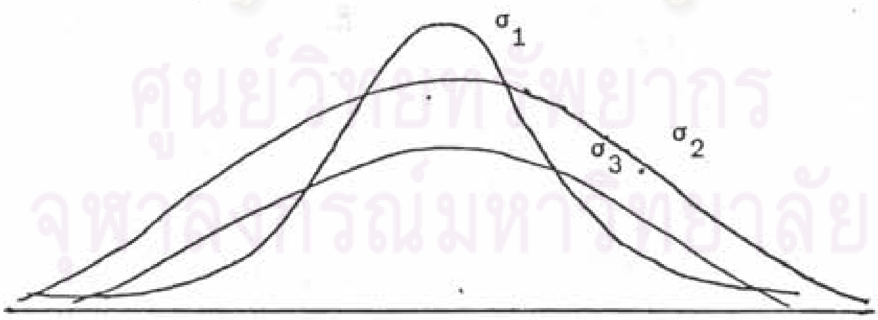
(ดูรูปที่ 2.1, 2.2, 2.3 ประกอบ)



รูปที่ 2.1 แสดงเส้นโค้งของการแจกแจงแบบปกติ



รูปที่ 2.2 แสดงการแจกแจงแบบปกติ 3 รูป ที่มีค่าเฉลี่ยต่าง ๆ กัน แต่มีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากัน



$\sigma_1 < \sigma_2 < \sigma_3$

รูปที่ 2.3 แสดงการแจกแจงแบบปกติ 3 รูปที่มีค่าส่วนเบี่ยงเบนมาตรฐานต่างกัน แต่มีค่าเฉลี่ยเท่ากัน

2.3 ผลงานวิจัยที่เกี่ยวข้อง

สำหรับผลงานวิจัยที่เกี่ยวข้องในการศึกษาเกี่ยวกับการทดสอบความเท่ากันของค่าเฉลี่ยประชากร ในกรณีที่มีหลายประชากรนั้น มีแนวความคิดพื้นฐานมาจากการทดสอบความแตกต่างระหว่างค่าเฉลี่ยที่มาจาก 2 ประชากร ประมาณปี ค.ศ. 1937 โดย Welch, B.L. ได้ค้นหาวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของ 2 ประชากร ในกรณีที่ค่าความแปรปรวนแตกต่างกัน โดยเสนอให้ใช้สถิติทดสอบที (t-test) ซึ่งยังให้ผลไม่เป็นที่น่าพอใจนัก ในกรณีที่การแจกแจงของประชากรไม่เป็นแบบปกติ จากแนวความคิดพื้นฐานเกี่ยวกับการทดสอบความเท่ากันของค่าเฉลี่ย 2 ประชากรนี้เอง จึงได้ทำให้นักสถิติหลานท่านพยายามขยายการทดสอบออกเป็นหลายประชากร เพื่อให้เหมาะสมกับสถานการณ์จริงที่เกิดขึ้น ซึ่งในปี ค.ศ. 1951 Welch และ James ได้เสนอวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ย k ประชากร เมื่อความแปรปรวนของประชากรไม่เท่ากัน ซึ่งก็ได้เป็นการขยายแนวความคิดมาจากวิธีการทดสอบของ Welch (1937) ที่ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของ 2 ประชากร ในกรณีที่ค่าความแปรปรวนต่างกัน และเรียกว่า สถิติทดสอบของ Welch

โดยในปี ค.ศ. 1974 Kohr และ Games ได้ศึกษาเปรียบเทียบวิธีการทดสอบของ Welch กับการทดสอบ F ซึ่งจากการศึกษาพบว่า สถิติทดสอบแบบ Welch มีความแกร่งมากกว่าสถิติทดสอบแบบ F เมื่อความแปรปรวนของประชากรแตกต่างกัน โดยเฉพาะอย่างยิ่งเมื่อขนาดตัวอย่างไม่เท่ากัน แต่เมื่อความแปรปรวนของประชากรเท่ากัน สถิติทดสอบแบบ Welch จะมีอำนาจของการทดสอบน้อยกว่าสถิติทดสอบแบบ F เพียงเล็กน้อย ซึ่งสอดคล้องกับผลจากศึกษาของ Morton B. Brown และ Alan B. Forsythe(1974) ที่ได้ศึกษาเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบที่ใช้ทดสอบความเท่ากันของค่าเฉลี่ยประชากร ในกรณีที่มีมากกว่า 2 ประชากร ในกรณีขนาดตัวอย่างเล็ก โดยศึกษากรณีที่ขนาดตัวอย่างเท่ากันและไม่เท่ากัน นอกจากนั้นได้ทดลองเปลี่ยนค่าเฉลี่ยและค่าความแปรปรวนที่ระดับต่าง ๆ กันภายใต้การแจกแจงแบบปกติ ซึ่งสถิติทดสอบที่เขานำมาศึกษา คือ สถิติทดสอบแบบ ANOVA F-TEST สถิติทดสอบแบบ Modified F สถิติทดสอบของ Welch and James ซึ่งจากการศึกษาพบว่า ในกรณีที่ประชากรที่นำมาทดสอบมีค่าความแปรปรวนแตกต่างกัน สถิติทดสอบแบบ ANOVA F-TEST ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งในกรณีที่ขนาดตัวอย่าง

เท่ากันและไม่เท่ากัน ที่ระดับนัยสำคัญ .05 และ .01 ในขณะที่สถิติทดสอบแบบ Modified F และสถิติทดสอบของ Welch and James สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งในกรณีที่ขนาดตัวอย่างเท่ากันและไม่เท่ากันที่ทำการศึกษา จากนั้นก็ได้มีนักสถิติหลายท่านทำการเสนอวิธีการต่าง ๆ เกี่ยวกับการทดสอบความเท่ากันของค่าเฉลี่ยประชากรที่มากกว่า 2 กลุ่ม เมื่อความแปรปรวนของประชากรไม่เท่ากัน ซึ่งต่อมา Marascuilo (1987) ได้เสนอวิธีการทดสอบที่ดัดแปลงจากสถิติทดสอบแบบ Welch เพียงเล็กน้อย ทำให้มีวิธีการคำนวณที่ง่ายด้วยเหตุดังกล่าวจึงเป็นที่น่าสนใจศึกษาเปรียบเทียบว่า สถิติทดสอบแบบ Brown and Forsythe (ดัดแปลงมาจากการวิเคราะห์ความแปรปรวน) และสถิติทดสอบแบบ Maraschilo (ดัดแปลงมาจากสถิติทดสอบแบบ Welch) สถิติทดสอบตัวใดจะมีความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 หรือค่าอำนาจการทดสอบดีกว่าสถิติทดสอบแบบ ANOVA F-TEST ภายใต้สถานการณ์ที่กำหนดไว้ในขอบเขตของการวิจัยต่อไป