บทที่ 2

สถิติทดสอบและผลงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้จะทำการศึกษาเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของสถิติทดสอบ 3 วิธีคือ สถิติทดสอบแบบ ANOVA F-TEST สถิติทดสอบแบบ Trimmed W และสถิติทดสอบแบบ Trimmed F ภายใต้ลักษณะการแจกแจงของประชากรที่กำหนดและปัจจัยต่าง ๆ ที่คาดว่าจะมีผลต่อการศึกษาดังได้กล่าวไว้ในขอบเขตของการวิจัยในบทที่ 1 แล้ว สำหรับสมมติฐานของการทดสอบความเท่ากันของค่าเฉลี่ยประชากรในกรณีที่มีหลายประชากรมีรูปแบบดังนี้

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 ซึ่งในบทนี้จะกล่าวถึงรายละเอียดของสถิติทดสอบแต่ละวิธี ลักษณะการแจกแจงของประชากร รวมทั้งนำเสนอผลงานวิจัยที่เกี่ยวข้อง ดังรายละเอียดต่อไปนี้

2.1 สถิติทดสอบที่ใช้ในการวิจัยครั้งนี้

2.1.1 การวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA F-TEST) เป็นวิธีการจำแนกความแปรปรวนทั้งหมดของกลุ่มข้อมูลที่ได้มาจากการวางแผนการทดลองไม่ว่าจะเป็นแผนการทดลองแบบ Completely Randomized Design (CRD), Randomized Complete Block Design (RBD) หรือการทดลองแบบแฟคทอเรียล (Factorial Experiment) ก็ตาม โดยวิธีการของ ANOVA จะทำการจำแนกความแปรปรวนออกเป็นส่วนต่าง ๆ ซึ่งจะทำให้เราสามารถบอกถึงขนาดของความแปรปรวนจากแต่ละแหล่งว่าเป็นเท่าไรต่อจำนวนความแปรปรวนทั้งหมด

การวิเคราะห์ความแปรปรวน (ANOVA) ได้ถูกนำไปใช้กับงานทดลองแขนงต่าง ๆ มากมาย ซึ่งในแต่ละงานทดลองที่ผู้วิจัยต้องการศึกษานั้นผู้วิจัยจะต้องเข้าใจถึงวัตถุประสงค์ของการวิจัย เพื่อที่จะได้เลือกแผนการทดลองที่เหมาะสม ซึ่งจะนำไปสู่การตอบปัญหาที่สงสัยได้ ดังนั้นก่อนที่ผู้วิจัยจะลงมือทำการทดลอง ผู้วิจัยจะต้องระบุแหล่งที่มาของความแปรปรวนที่จะเกิดขึ้นได้ ทั้งนี้เพื่อที่จะได้เลือกแผนการทดลองที่สามารถทำให้ผู้วิจัยวัดขนาดของความแปรปรวนของแต่ละแหล่ง ต่อจำนวนความแปรปรวนทั้งหมดได้ยกตัวอย่างเช่น ในการทดสอบเพื่อดูว่าคุณภาพของสูตรอาหาร 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)) จากนั้นใช้การวิเคราะห์ความแปรปรวนวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร การใช้ ANOVA เพื่อวิเคราะห์ข้อมูลมีวัตถุประสงค์สำคัญสองประการคือ

1. เพื่อคาดประมาณและทดสอบสมมติฐานเกี่ยวกับค่าความแปรปรวนของประชากร

2. เพื่อคาดประมาณและทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร

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

ในการวิจัยครั้งนี้ จะพิจารณาถึง การวิเคราะห์ความแปรปรวนสำหรับแผนการทดลองแบบสุ่มสมบูรณ์ (CRD) ซึ่งหน่วยทดลองทั้งหมดจะต้องมีความคล้ายคลึงกัน (Homogeneous) จากนั้นแบ่งออกเป็นกลุ่มต่าง ๆ แล้วให้วิธีการ (Treatment) แต่ละวิธีที่ต้องการทดสอบกับหน่วยทดลองภายในแต่ละกลุ่ม ดังนั้นเราสามารถแบ่งความผันแปรออกเป็น 2 แหล่ง ด้วยกันคือ 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

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

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

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

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

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

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 สถิติทดสอบแบบANOVA F-Test จะมีลักษณะการแจกแจงแบบเอฟ (F-Distribution) ที่องศาความเป็นอิสระ (c-1,N-c)

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

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

สถิติทดสอบที่จะกล่าวต่อไปนี้ เป็นสถิติทดสอบที่ได้มีการกำหนดร้อยละของการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจง จากนั้นใช้ค่าเฉลี่ยของข้อมูลส่วนที่เหลือ (g-times trimmed mean) เป็นตัวประมาณค่าเฉลี่ยของประชากร นอกจากนั้นตัวประมาณค่าความแปรปรวนของประชากร จะคำนวณมาจากผลรวมของส่วนเบี่ยงเบนกำลังสองแบบวินเซอร์ไรซ์ที่ระดับการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจงด้านละ g%รายละเอียดของวิธีการคำนวณมีดังต่อไปนี้

การคำนวณค่าเฉลี่ยแบบทริมด์และผลรวมของส่วนเบี่ยงเบนกำลังสองแบบวินเซอร์ไรซ์ (Trimmed Mean and Winsorized Sum of Squared Deviations)

ถ้ากำหนดให้ n คือขนาดตัวอย่าง และ $X_1, X_2, \ldots, X_n$ เป็นค่าสังเกตที่เรียงลำดับจากน้อยไปมาก จะได้ว่า

- ค่าเฉลี่ยแบบทริมด์ (The g-trimmed mean) จะคำนวณจาก

$$\bar{X}_{tg} = \left[\sum_{j=k+1}^{n-k} X_j + (1-\varepsilon)(X_k + X_{n-k+1})\right] / (n-2gn) \quad \ldots\ldots\ldots\ldots (2.1)$$

- ค่าเฉลี่ยแบบวินเซอร์ไรซ์(The g-Winsorized mean)

จะคำนวณจาก

$$\bar{X}_{wg} = \left[\sum_{j=k+1}^{n-k} X_j + k\{(1-\varepsilon)(X_k + X_{n-k+1}) + \varepsilon(X_{k+1}+X_{n-k})\}\right] / n \quad .(2.2)$$

โดยที่ g แทนร้อยละของการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจงด้านละ g%

$k = [gn] + 1$ เมื่อ $[gn]$ เป็นจำนวนเต็มมากที่สุด $\leq gn$

$\varepsilon = gn - [gn]$

- ผลรวมของส่วนเบี่ยงเบนกำลังสองแบบวินเซอร์ไรซ์ (The g-Winsorized Sum of Squared Deviation) คำนวณจาก

$$SSD_{wg} \quad \sum_{j=k+1}^{n-k} (X - \bar{X}_{wg})^2 + k\left[\{(1-\varepsilon)X_k + \varepsilon X_{k+1} - \bar{X}_{wg}\}^2 + \{(1-\varepsilon)\, X_{n-k+1} + \varepsilon X_{n-k} - \bar{X}_{wg}\}^2\right] \quad \ldots(2.3)$$

2.1.2 สถิติทดสอบแบบ Trimmed F เสนอโดย H. LEE and K.Y.FUNG ใน ปี ค.ศ. 1983 โดยพัฒนามาจากสถิติทดสอบแบบ Modified F Test ทั้งนี้เนื่องจากสถิติทดสอบแบบ Modified F Test สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ในกรณีที่ประชากรมีความแปรปรวนแตกต่างกัน ภายใต้การแจกแจงแบบปกติ แต่ถ้าลักษณะการแจกแจงของประชากรเป็นแบบสมมาตรหางยาว (Long-Tailed Distribution) สถิติทดสอบแบบ Modified F จะมีอำนาจการทดสอบลดลงด้วยเหตุผลดังกล่าว ทำให้ H. LEE and K.Y. FUNG ได้นำแนวความคิด (Concept) เกี่ยวกับการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจงมาประยุกต์ใช้กับสถิติทดสอบแบบ Modified F Test ซึ่งมีรูปแบบดังนี้

$$F_{t(g)} = \frac{\sum_{i=1}^{c} h_i(\bar{X}_{tg_i} - \bar{X}_{tg})^2}{\sum_{i=1}^{c} (1-hi/H)\, S^2_{wg_i}} \quad \ldots\ldots\ldots\ldots\ldots(2.4)$$

โดยที่ $h_i = (1-2g)\, n_i$

$$H = \sum_{i=1}^{c} hi = (1-2g)N$$

$$\bar{X}_{tg_i} = \left[\sum_{j=k+1}^{n_i-k} X_{ij} + (1-\varepsilon)(X_{ik} + X_{i,n_i-k+1})\right] / (n_i - 2gn_i) \ldots(2.5)$$

สำหรับสูตรที่ใช้ในการคำนวณ $\bar{X}_{tg}$ (Overall Trimmed Mean) เริ่มแรกจะทำการเรียงลำดับค่า X_{ij} ทั้งหมด แล้วคำนวณค่า $\bar{X}_{tg}$ จากสูตรต่อไปนี้

$$\bar{X}_{tg} = \left[\sum_{j=k+1}^{N-k} X_j + (1-\varepsilon)(X_k + X_{N-k+1})\right] / (N-2gn) \quad \ldots\ldots\ldots\ldots(2.6)$$

โดยที่ $N = \sum_{i=1}^{c} n_i$

$k = [gn] + 1$ เมื่อ $[gn]$ เป็นจำนวนเต็มมากที่สุดที่ $\leq gn$

$\varepsilon = gn - [gn]$

สำหรับสูตรที่ใช้ในการคำนวณ $S^2_{wg_i}$ ได้จากสมการที่ (2.7) ดังนี้

$$S^2_{wg_i} = SSD_{wg_i} / (h_i-1) \quad \text{....................(2.7)}$$

$$SSD_{wg_i} = \sum_{j=k+1}^{n_i-k} (X_{ij}-\bar{X}_{wg_i})^2 + k\,[\{(1-\varepsilon)X_{ik} + \varepsilon X_{i,k+1} - \bar{X}_{wg_i}\}^2 + \{(1-\varepsilon)\,X_{i,n_i-k+1} + X_{i,n_i-k} - \bar{X}_{wg_i}\}^2] \quad \text{.............(2.8)}$$

และ
$$\bar{X}_{wg_i} = [\sum_{j=k+1}^{n_i-k} X_{ij} + k\,\{(1-\varepsilon)(X_{ik}+X_{i,n_i-k+1}) + \varepsilon(X_{i,k+1} + X_{i,n_i-k})\}]\,/n_i \quad \text{............(2.9)}$$

ภายใต้สมมติฐานหลักที่เป็นจริง กล่าวคือ $H_o : \mu_1 = \mu_2 = \ldots = \mu_c$ $F_{t(g)}$ จะมีลักษณะการแจกแจงแบบเอฟโดยประมาณ (Approximate F Distribution) ที่องศาความเป็นอิสระ $(c-1,\ f_t)$ โดยที่

$$f_t = [\sum_{i=1}^{c} v_i^2 / (h_i-1)]^{-1}$$

$$v_i = (1-h_i/H)\, S^2_{wg_i} / [\sum_{i=1}^{c} (1-h_i/H)\, S^2_{wg_i}]$$

เกณฑ์การตัดสินใจ จะปฏิเสธ H_o เมื่อ $F_{t(g)} > F_{\alpha,(c-1,f_t)}$ โดยที่ $F_{\alpha,(c-1,f_t)}$ คือค่าวิกฤตที่เปิดจากตาราง F ที่ระดับนัยสำคัญ α และองศาความเป็นอิสระ $(c-1, f_t)$ ตามลำดับ

2.1.3 สถิติทดสอบแบบ Trimmed W เสนอโดย H. LEE and K.Y. FUNG ในปี ค.ศ. 1983 ซึ่งพัฒนามาจากสถิติทดสอบ Modified Welch โดยนำแนวความคิดเกี่ยวกับการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจงมาประยุกต์ใช้กับสถิติทดสอบดังกล่าว ซึ่งมีรูปแบบดังนี้

$$W_{t(g)} = \frac{\sum_{i=1}^{c} w_i (\bar{X}_{tg_i} - \tilde{X}_{tg})^2 / (c-1)}{1+\left[2(c-2)/(c^2-1)\right] \sum_{i=1}^{c} (1-w_i/W)^2/(n_i-1)}$$

$$w_i = h_i/S^2_{wg_i}$$

$$S^2_{wg_i} = SSD_{wg_i}/(h_i-1)$$

$$h_i = (1-2g)n_i$$

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

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

ภายใต้สมมติฐานหลักที่เป็นจริง กล่าวคือ $H_o : \mu_1 = \mu_2 = \ldots = \mu_c$ แล้ว $W_{t(g)}$ จะมีลักษณะการแจกแจงแบบเอฟโดยประมาณ (Approximate F Distribtuion) ที่องศาความเป็นอิสระ $(c-1, f_t)$ โดยที่

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

เกณฑ์การตัดสินใจ จะปฏิเสธ H_o เมื่อ $W_t(g) > F_{\alpha,(c-1,f_t)}$ โดยที่ $F_{\alpha,(c-1,f_t)}$ คือค่าวิกฤตที่เปิดจากตาราง F ที่ระดับนัยสำคัญ α และองศาความเป็นอิสระ $(c-1\ ,\ f_t)$ ตามลำดับ

เนื่องจากในการวิจัยครั้งนี้สนใจศึกษาความแกร่งของสถิติทดสอบทั้ง 3 ตัว ในกรณีที่ข้อมูลที่นำมาทดสอบมีลักษณะการแจกแจงเป็นแบบปกติ (Normal Distribution) การแจกแจงแบบปกติปลอมปน (Contaminate Normal Distribution) ซึ่งรายละเอียดและคุณสมบัติต่าง ๆ เกี่ยวกับการแจกแจงของประชากรจะนำเสนอดังต่อไปนี้

2.2 คุณสมบัติและลักษณะการแจกแจงของประชากรที่ศึกษา

2.2.1 การแจกแจงปกติ (Normal Distribution) เป็นการแจกแจงที่ใช้อธิบายข้อมูลที่เกิดขึ้นโดยทั่วไปทางธรรมชาติ เช่น คะแนนสอบ หรือยอดขายสินค้าชนิดหนึ่งในรอบ 6 เดือนที่ผ่านมา เป็นต้น การแจกแจงแบบปกติค้นพบโดยเดอมูวร์ (De Movire : 1667-1754) นักคณิตศาสตร์ชาวฝรั่งเศส ต่อมาลาปลาซ (Laplace : 1749-1827) นักคณิตศาสตร์ชาวอังกฤษได้นำมาประยุกต์ใช้ทางด้านสังคมศาสตร์และวิทยาศาสตร์ ต่อมานักคณิตศาสตร์เยอรมันเกาส์ (Gauss: 177-1855) ได้นำทฤษฎีนี้ไปศึกษาหาความคลาดเคลื่อน (Error) โดยการวัดซ้ำ ๆ ในกลุ่มที่มีขนาดคงเดิม และพบว่าการแจกแจงที่ได้จะมีลักษณะเป็นโค้งปกติ ดังนั้นการแจกแจงแบบปกติอาจเรียกอีกอย่างหนึ่งว่า Gaussian Distribution โดยมีฟังก์ชันความหนาแน่น (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

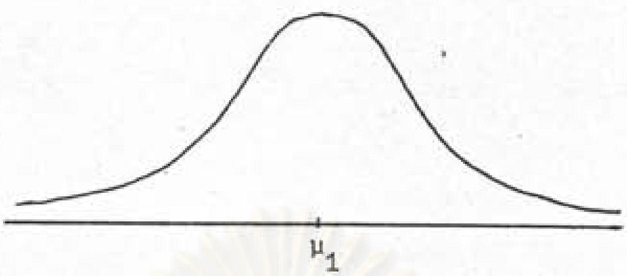
4. มีความโด่ง (Kurtosis) ของเส้นโค้งเท่ากับ 3 ซึ่งเรียกว่าเมโซเคอร์-ติค (Mesokurtic)และจุดเปลี่ยนโค้งทั้ง 2 ข้างจะอยู่ ณ ตรง 1 ส่วนเบี่ยงเบนมาตรฐาน

5. ค่าความเบ้ (Skewness) เท่ากับศูนย์

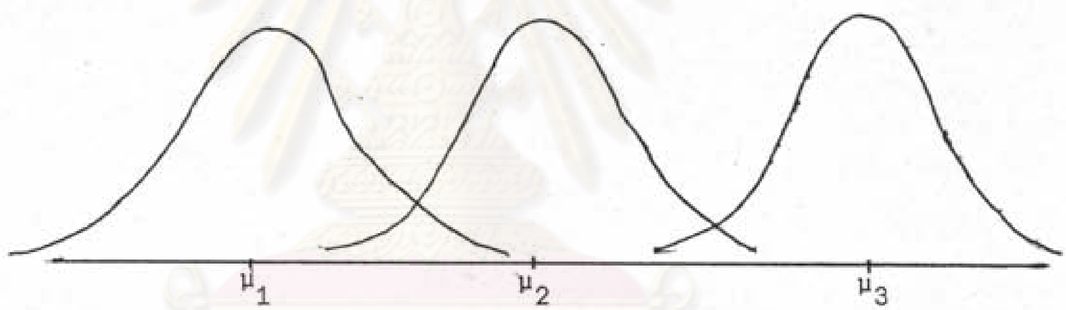
6. ปลายทั้งสองข้างของเส้นโค้งจะค่อย ๆ ลดต่ำลงแต่ไม่จรดกับฐานของโค้งหรือแกนนอน

7. ถ้าลากเส้นตั้งฉากจากแกน X (ซึ่งเป็นแกนนอน) ไปยังเส้นโค้ง โดยที่เส้นดังกล่าวห่างจากค่าเฉลี่ย (μ) ทั้งด้านซ้ายและด้านขวาของระยะหนึ่งเท่า สองเท่าและสามเท่าของส่วนเบี่ยงเบนมาตรฐาน (σ) แล้ว พื้นที่ที่ปิดกั้นเส้นตั้งฉากกับเส้นโค้งจะเท่ากับ 68% 95% และ 99.7% ของพื้นที่ทั้งหมด ตามลำดับ

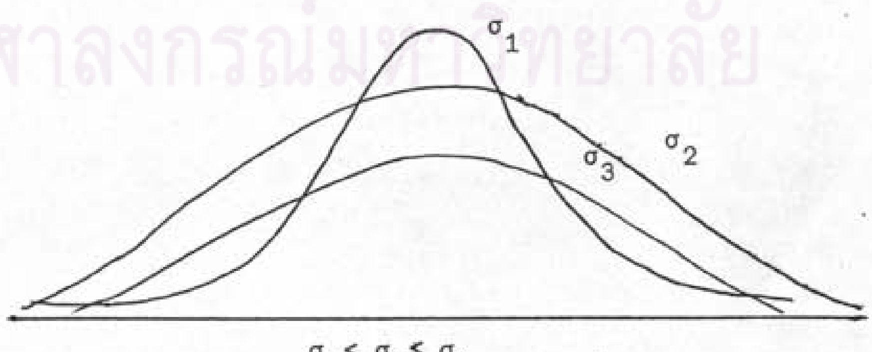
(ดูรูปที่ 2.1, 2.2,2.3 ประกอบ)



รูปที่ 2.1 แสดงเส้นโค้งของการแจกแจงแบบปกติ



รูปที่ 2.2 แสดงการแจกแจงแบบปกติ 3 รูป ที่มีค่าเฉลี่ยต่าง ๆ กัน แต่มีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากัน



รูปที่ 2.3 แสดงการแจกแจงปบบปกติ 3 รูปที่มีค่าส่วนเบี่ยงเบนมาตรฐานต่างกัน แต่มีค่าเฉลี่ยเท่ากัน

2.2.2 การแจกแจงแบบปกติปลอมปน (Contaminate Normal Distribution) เป็นการแจกแจงที่สร้างขึ้นเพื่อให้มีลักษณะเป็นการแจกแจงชนิดสมมาตรหางยาว โดยส่วนหนึ่งของประชากรมาจากการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ ความแปรปรวน σ^2 ด้วยความน่าจะเป็น $1-p$ และอีกส่วนหนึ่งของประชากรมาจากการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ เช่นกัน แต่มีความแปรปรวนเป็น $r^2\sigma^2$ ด้วยความน่าจะเป็น p ซึ่ง p เป็นเปอร์เซ็นต์การปลอมปน และ r เป็นค่าสเกลแฟคเตอร์ กำหนดสัญลักษณ์รูปแบบของการแจกแจงดังนี้

$$f(x) = (1-p)\, N(\mu, \sigma^2) + pN(\mu, r^2\sigma^2)$$

โดยที่ $N(\mu, \sigma^2)$ แทนการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ และความแปรปรวนเท่ากับ σ^2

สำหรับการวิจัยครั้งนี้ ศึกษาการแจกแจงแบบปกติปลอมปน 4 รูปแบบ โดยกำหนดเปอร์เซ็นต์การปลอมปน 2 ระดับ คือ p = 10,20% และสเกลแฟคเตอร์ 2 ระดับ คือ r = 5,10 เพื่อให้มีรูปแบบของการแจกแจงต่าง ๆ กันดังนี้

2.2.2.1 เปอร์เซ็นต์การปลอมปน 10 และสเกลแฟคเตอร์ 5

2.2.2.2 เปอร์เซ็นต์การปลอมปน 10 และสเกลแฟคเตอร์ 10

2.2.2.3 เปอร์เซ็นต์การปลอมปน 20 และสเกลแฟคเตอร์ 5

2.2.2.4 เปอร์เซ็นต์การปลอมปน 20 และสเกลแฟคเตอร์ 10

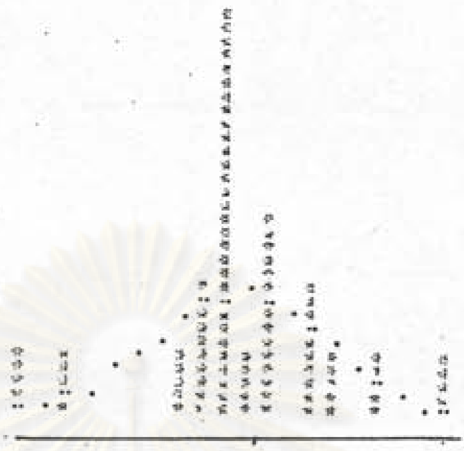
(ดูรูปที่ 2.4- 2.7 ประกอบ)

E(X)

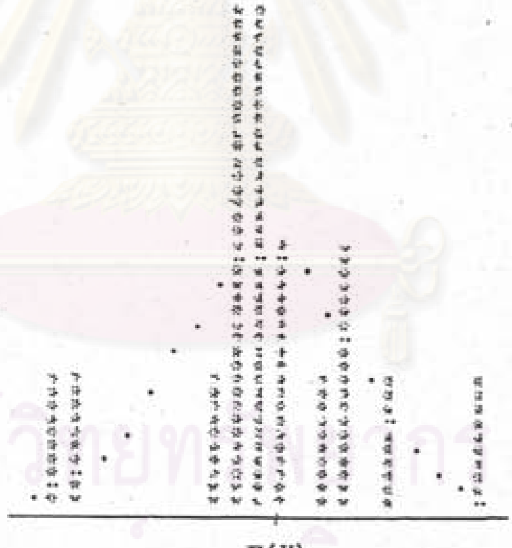
รูปที่ 2.4 เส้นโค้งของการแจกแจงแบบปกติเทียบกับเส้นโค้งของการแจกแจงแบบปกติปลอมปนที่มี p = 10% , r = 5

E(X)

รูปที่ 2.5 เส้นโค้งของการแจกแจงแบบปกติเทียบกับเส้นโค้งของการแจกแจงแบบปกติปลอมปนที่มี p = 10% , r = 10



รูปที่ 2.6 เส้นโค้งของการแจกแจงแบบปกติเทียบกับเส้นโค้งของการแจกแจงแบบปกติปลอมปนที่มี p = 20% , r = 5



รูปที่ 2.7 เส้นโค้งของการแจกแจงแบบปกติเทียบกับเส้นโค้งของการแจกแจงแบบปกติปลอมปนที่มี p = 20% , r - 10

<u>หมายเหตุ</u> (•) แทนการแจกแจงแบบปกติ

(*) แทนการแจกแจงแบบปกติปลอมปน

2.3 <u>ผลงานวิจัยที่เกี่ยวข้อง</u>

สำหรับผลงานวิจัยที่เกี่ยวข้องในการศึกษาเกี่ยวกับการทดสอบการเท่ากันของค่าเฉลี่ยประชากรในกรณีที่มีหลายประชากรนั้น มีแนวความคิดพื้นฐานมาจากการทดสอบความแตกต่างระหว่างค่าเฉลี่ยที่มาจาก 2 ประชากร ประมาณ ปี ค.ศ. 1937 โดย Welch, B.L ได้ค้นหาวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของ 2 ประชากรในกรณีที่ค่าความแปรปรวนแตกต่างกัน โดยเสนอให้ใช้สถิติทดสอบที (t-test) ซึ่งยังให้ผลไม่เป็นที่น่าพอใจนักในกรณีที่การแจกแจงของประชากรไม่เป็นแบบปกติ ต่อมาได้นักสถิติหลายท่านได้พยายามหาตัวสถิติทดสอบที่มีประสิทธิภาพสูง และมีความแกร่ง (Robust) สำหรับการแจกแจงที่ไม่เป็นแบบปกติ เช่น สถิติทดสอบทริมด์ที (Trimmed t-test)สถิติทดสอบวินเซอร์ไรซ์ ที (Winsorized t-test) ซึ่งต่างเป็นสถิติทดสอบที่สร้างขึ้นเพื่อใช้สำหรับการแจกแจงสมมาตรหางยาว (Long-Tailed Distribution) โดย Yuen และ Dixon ได้เสนอสถิติทดสอบทริมด์ ที สำหรับทดสอบความแตกต่างของค่าเฉลี่ยประชากรที่เป็นอิสระต่อกัน 2 ชุด โดยอาศัยข้อมูลจากกลุ่มนักสถิติ Princeton ที่ศึกษาเกี่ยวกับความแกร่งของตัวประมาณตำแหน่งของประชากร (Robust Estimates of Population Location) พบว่า ค่าเฉลี่ยทริมด์ (Trimmed Mean) มีประสิทธิภาพสูงสำหรับการแจกแจงในหลายรูปแบบ และที่การแจกแจงแบบปกติประสิทธิภาพของตัวประมาณแบบค่าเฉลี่ยทริมด์ เมื่อเทียบกับค่าเฉลี่ยเลขคณิต มีค่าค่อนข้างสูง คือมีเศษส่วนการสูญเสีย (fractional loss) เพียงประมาณ $\frac{2}{3}$ g/n เท่านั้น ในขณะที่ตัวประมาณแบบอื่น เช่น มัธยฐานมีเศษส่วนของการสูญเสียถึงหนึ่งในสาม ต่อมาในปี ค.ศ. 1974 Hogg ได้เสนอวิธีการเลือกระดับการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจง ซึ่งจะขึ้นอยู่กับแต่ละสถานการณ์ กล่าวคือ ขึ้นกับลักษณะความมากหรือน้อยของข้อมูลที่ผิดปกติ (Outlier) ตัวอย่างเช่น ภายใต้การแจกแจงที่เป็นสมมาตรแบบหางยาว ระหว่างการแจกแจงแบบปกติและการแจกแจงแบบคอชี (Cauchy Distribution) ใช้ระดับการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจง ประมาณ $\frac{1}{4}$ หรือ $\frac{1}{3}$ ของขนาดตัวอย่างและภายใต้การแจกแจงแบบปกติและการแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียล จะใช้ระดับการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจงประมาณ $\frac{1}{5}$ ของขนาดตัวอย่าง ซึ่งการศึกษาทั้งสองกรณีนี้พบว่า ค่าเฉลี่ยทริมด์ (Trimmed Mean) เป็นตัวประมาณที่มีความแกร่ง Hogg ได้กล่าวต่อไปว่า ระดับการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจงจะถูกกำหนดคงที่โดยพิจารณาล่วงหน้าก่อนที่จะทำการสุ่มตัวอย่าง ซึ่งมีนักสถิติหลายท่านเห็นว่าไม่เหมาะสม เนื่องจากในทางปฏิบัติ เราไม่ทราบลักษณะการแจกแจงของข้อมูลที่ได้แน่ชัด

และได้แนะนำให้เลือกระดับการตัดค่าสังเกต ภายหลังจากได้ทำการสุ่มตัวอย่างแล้ว โดยใช้ข้อมูลตัวอย่างประกอบการพิจารณา ดังเช่น Tukey และ Mc Laughlin รวมทั้ง Jaeckel ได้ใช้วิธีเลือกระดับการตัดค่าสังเกตที่จะทำให้ค่าประมาณความแปรปรวนของค่าเฉลี่ยทริมด์ มีค่าน้อยที่สุด โดยการทำการจำลองที่ขนาดตัวอย่างเท่ากับ 20 ภายใต้การแจกแจงแบบปกติ แบบโลจิสติค และแบบคอชี เมื่อกำหนดระดับการตัดค่าสังเกต $g = [n\alpha]$ โดยที่ $\hat{g}$ และ $\hat{\alpha}$ เป็นค่าที่จะทำให้ค่าประมาณความแปรปรวนของค่าเฉลี่ยทริมด์มีค่าน้อยที่สุด (Minimized $S_m^2(\hat{\alpha})$ ซึ่งจากการศึกษาเขาพบว่า เมื่อช่วงของ α อยู่ระหว่าง 0 ถึง $\left[\frac{1}{4}n\right]$ จะทำให้ค่าความแปรปรวนของค่าเฉลี่ยทริมด์ มีค่าน้อยที่สุด และมีค่าใกล้เคียงกัน ภายใต้การแจกแจงที่ศึกษา

จากแนวความคิดพื้นฐานเกี่ยวกับการทดสอบความเท่ากันของค่าเฉลี่ย 2 ประชากรนี้เอง จึงได้ทำให้นักสถิติหลายท่านพยายามขยายการทดสอบออกเป็นหลายประชากร เพื่อให้เหมาะสมกับสถานการณ์จริงที่เกิดขึ้น ซึ่งในปี ค.ศ. 1974 Morton B. Brown และ Alan B. Forsythe ได้ศึกษาเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบที่ใช้ทดสอบความเท่ากันของค่าเฉลี่ยประชากร ในกรณีที่มีมากกว่า 2 ประชากร ในกรณีขนาดตัวอย่างเล็ก โดยศึกษาทั้งกรณีที่ขนาดตัวอย่างเท่ากันและไม่เท่ากัน นอกจากนั้นได้ทดลองเปลี่ยนค่าเฉลี่ยและค่าความแปรปรวนที่ระดับต่าง ๆ กัน ภายใต้การแจกแจงแบบปกติ ซึ่งสถิติทดสอบที่เขานำมาศึกษา คือ สถิติทดสอบแบบ ANOVA F-Test สถิติทดสอบแบบ Modified F สถิติทดสอบของ Welch and James ซึ่งจากการศึกษาเขาพบว่า ในกรณีที่ประชากรที่นำมาทดสอบมีค่าความแปรปรวนแตกต่างกัน สถิติทดสอบแบบ ANOVA F-Test ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ทั้งในกรณีที่ขนาดตัวอย่างเท่ากันและไม่เท่ากัน ที่ระดับนัยสำคัญ .05 และ .01 ในขณะที่สถิติทดสอบแบบ Modified F และสถิติทดสอบของ Welch and James สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ทั้งในกรณีที่ขนาดตัวอย่างเท่ากันและไม่เท่ากันที่ทำการศึกษา ต่อมาในปี ค.ศ. 1983 Hyunshik Lee และ Karen Yuen Fung ได้นำค่าเฉลี่ยแบบทริมด์ (Trimmed Mean)และตัวประมาณแบบ Sine-Wave M มาประยุกต์ใช้กับสถิติทดสอบ Modified Welch และ Modified F โดยได้ทำการศึกษาค่าอำนาจของการทดสอบในกรณีที่ข้อมูลมีลักษณะการแจกแจงแบบปกติและการแจกแจงที่มีลักษณะสมมาตรหางยาว (Long-Tail Distribution) ในกรณีขนาดตัวอย่างเล็ก จากการศึกษาสรุปได้ว่า สถิติ

ทดลองที่มีการพิจารณาระดับการตัดค่าสังเกตที่ปลายทั้ง 2 ด้านของการแจกแจง (Trimmed Mean) จะมีค่าอำนาจของการทดลองสูงกว่า สถิติทดลองที่ใช้ตัวประมาณแบบ Sine-Wave M ในกรณีที่ประชากรมีการแจกแจงแบบปกติ แต่ในกรณีที่ประชากรมีการแจกแจงแบบสมมาตรหางยาว สถิติทดลองที่มีการใช้ค่าเฉลี่ยแบบทริมด์ และสถิติทดลองที่ใช้ตัวประมาณแบบ Sine-Wave M จะให้ค่าอำนาจการทดลองดีพอกัน ในกรณีตัวอย่างขนาดเล็ก(10,10,10) นอกจากนี้สถิติทดลองที่มีการใช้ค่าเฉลี่ยแบบทริมด์ ยังมีวิธีการคำนวณที่ง่าย ด้วยเหตุดังกล่าวจึงเป็นที่น่าสนใจศึกษาเปรียบเทียบว่าสถิติทดลองที่มีการใช้ค่าเฉลี่ยแบบทริมด์ ซึ่งในที่นี้คือสถิติทดลองแบบ Trimmed W (พัฒนามาจากสถิติทดลอง Modified Welch) และสถิติทดลองแบบ Trimmed F (พัฒนามาจากสถิติทดลอง Modified F) สถิติทดลองตัวใดจะมีความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 หรือมีค่าอำนาจการทดลองดีกว่าสถิติทดลองแบบ ANOVA F-TEST ภายใต้สถานการณ์ที่กำหนดไว้ในขอบเขตของการวิจัยต่อไป