บทที่ 2

สถิติทดสอบและผลงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้จะทำการศึกษาเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของสถิติทดสอบ 5 วิธี คือตัวสถิติทดสอบแบบ ANOVA F TEST การแปลงข้อมูลเป็นค่าลอคการิทึม ตัวสถิติทดสอบแบบ TRIMMED F ตัวสถิติทดสอบแบบ Brown & Forsythe และสถิติทดสอบแบบเอฟที่ใช้ค่าเฉลี่ยของ Graybill & Deal ภายใต้ลักษณะการแจกแจงแบบปกติและลักษณะความแปรปรวนของประชากรที่กำหนด รวมทั้งปัจจัยต่างๆ ที่คาดว่าจะมีผลต่อการศึกษา ดังได้กล่าวไว้ในขอบเขตของการวิจัยในบทที่ 1 แล้ว สำหรับสมมติฐานของการทดสอบความเท่ากันของค่าเฉลี่ยประชากร ในกรณีที่มีหลายประชากรมีรูปแบบดังนี้

H_0 : $\mu_1 = \mu_2 = \mu_3 = \ldots\ldots = \mu_i$

H_1 : มีค่าเฉลี่ยอย่างน้อย 1 คู่ ที่แตกต่างกัน

โดยที่ μ_i แทนค่าเฉลี่ยของประชากรที่ i , i = 1,2,...,k

สำหรับการวิจัยครั้งนี้ กำหนดจำนวนประชากร (k) เท่ากับ 3 ซึ่งในบทนี้จะกล่าวถึงรายละเอียดของสถิติทดสอบแต่ละวิธี ลักษณะการแจกแจงของประชากรรวมทั้งนำเสนอผลงานวิจัยที่เกี่ยวข้อง ดังรายละเอียดต่อไปนี้

สถิติทดสอบที่ใช้ในการวิจัย

1. การทดสอบเอฟ (F test) ในการวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว (One Way Classification)

การวิเคราะห์ความแปรปรวน (analysis of variance) เป็นวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากร ตั้งแต่สามประชากรขึ้นไป หลักการที่สำคัญที่ใช้ในการทดสอบสมมติฐาน คือการแยกความแปรปรวนหรือความแตกต่างของข้อมูลที่เกิดขึ้นทั้งหมดออกจาก

สาเหตุต่างๆ แล้วพิจารณาสัดส่วนของความแปรปรวนหรือความแตกต่างระหว่างประชากรและความแปรปรวนหรือความแตกต่างภายในประชากรเดียวกันว่ามีค่ามากน้อยเพียงใด ถ้าอัตราส่วนดังกล่าวมีมากแสดงว่าความแปรปรวนหรือความแตกต่างระหว่างประชากรมีมาก เมื่อเทียบกับความแปรปรวนหรือความแตกต่างภายในประชากรเดียวกัน สามารถสรุปได้ว่าจำนวนประชากรทั้งหมดที่นำมาทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากรอย่างน้อยหนึ่งประชากรที่แตกต่างจากประชากรอื่นๆ ที่นำมาทดสอบ ซึ่งวิธีวิเคราะห์ความแปรปรวนจะแตกต่างกันไปตามลักษณะแผนการทดลอง แผนการทดลองที่ใช้ในการวิจัยนี้ใช้แผนการทดลองแบบสุ่มโดยสมบูรณ์ (completely randomized design) หรือข้อมูลแบบแจกแจงทางเดียว ซึ่งเกี่ยวกับตัวแปรหรือปัจจัยที่ต้องการศึกษาซึ่งมีเพียงปัจจัยเดียว โดยแยกเป็นหลายระดับหรือสิ่งทดลองหรือทรีทเมนท์ การวิเคราะห์ความแปรปรวนจะแยกสาเหตุของความแปรปรวนของข้อมูลทั้งหมดว่า เนื่องมาจากอิทธิพลของทรีทเมนต์ (treatment effect) แต่เพียงอย่างเดียว ดังนั้น เพื่อให้แผนการทดลองนี้มีประสิทธิภาพมากที่สุด หน่วยทดลอง (experimental unit) ที่นำมาใช้ควรมีลักษณะเหมือนกันหรือคล้ายคลึงกันมากที่สุด (homogeneous) หรือให้มีความแปรปรวนระหว่างหน่วยทดลองน้อยที่สุด

ในการวิเคราะห์ทางเดียว เราจำแนกข้อมูลด้วยปัจจัยเพียงปัจจัยเดียว คือแยกตามสิ่งทดลองที่ใช้ และในการทดลองทุกครั้งจะต้องมีความคลาดเคลื่อน (error) เกิดขึ้นเสมอ ดังนั้นในการทดสอบแบบทางเดียว เราอาจแยกความแปรผันทั้งหมดที่เกิดขึ้นออกเป็น ความแปรผันที่เกิดขึ้นเนื่องจากการใช้สิ่งทดลองที่ต่างกันในแต่ละกลุ่ม ซึ่งจะเรียกว่าความแปรผันระหว่างกลุ่มหรือความแปรผันระหว่างสิ่งทดลองกับความคลาดเคลื่อนของการทดลอง อาจเรียกได้ว่าเป็นความแปรผันภายในกลุ่ม ซึ่งถ้าไม่มีผลอันเนื่องมาจากสิ่งทดลองที่ใช้แตกต่างกันแล้ว ความแปรผันทั้งคู่ก็ไม่ควรจะแตกต่างกัน นั่นคือ ความแปรผันทั้งคู่จะเป็นตัวประมาณค่าของความแปรปรวน (σ^2) ของข้อมูล เพื่อความสะดวกในการวิเคราะห์ อาจจัดส่วนประกอบของความแปรผันที่เกิดขึ้นในการทดลองลงในตารางการวิเคราะห์ความแปรปรวน (ANOVA) ได้ดังตาราง

ตารางการวิเคราะห์ความแปรปรวนของแผนการทดลองแบบสุ่มตลอด

ที่มาของความแปรผัน	d.f.	ความแปรผัน	M.S.
ระหว่างสิ่งทดลองหรือทรีทเมนต์	k - 1	$SS_{Tr} = \sum_{i}^{k} n_i(\bar{x}_{i} - \bar{x}_{..})^2$	$MS_{Tr} = \dfrac{SS}{k-1}$
ภายในสิ่งทดลองเดียวกัน	N - k	$SS_E = \sum_{i}^{k}\sum_{j}^{n_i}(x_{ij} - \bar{x}_{i.})^2$	$MS_E = \dfrac{SS_E}{N-k}$
รวม	N - 1	$SS_T = \sum_{i}^{k}\sum_{j}^{n_i}(x_{ij} - \bar{x}_{..})^2$	

ตัวสถิติที่ใช้ในการทดสอบ คือ

$$F_{k-1, N-k} = \frac{MS_{Tr}}{MS_E}$$

ซึ่งสมมติฐานที่กำหนดจะถูกปฏิเสธไป เมื่อ F ที่คำนวณได้จากตัวอย่างมีค่ามากกว่าค่า F ที่เปิดจากตารางที่กำหนดที่ระดับนัยสำคัญที่กำหนด

2. การทดสอบเอฟที่ใช้ค่าเฉลี่ยของ Graybill & Deal

จากตัวสถิติ F TEST ที่กล่าวมาแล้วข้างต้น จะเห็นได้ว่าการสร้างตัวแปรขึ้นมาใหม่โดยใช้ชื่อว่า $\bar{x}$ นั้นใช้ค่าเฉลี่ยซึ่งเป็นค่าวัดแนวโน้มสู่ส่วนกลาง (central location) เพื่อดูความแตกต่างของค่าสังเกตแต่ละตัวในกลุ่มกับค่าเฉลี่ย (mean) ซึ่งเป็นตัวแทนของกลุ่มดังกล่าว ทั้งนี้หากว่าการกระจายหรือการแจกแจงของข้อมูลเป็นแบบสมมาตร (symmetry) แล้วจะถือว่าค่าเฉลี่ยเป็นค่าที่เหมาะสมสำหรับนำมาใช้เป็นตัวแทนของกลุ่ม แต่ถ้าข้อมูลดังกล่าวมีการกระจายแตกต่างไปจากกลุ่มอื่น ค่าวัดแนวโน้มสู่ส่วนกลางคือค่าเฉลี่ยคงไม่เหมาะสมสำหรับจะเป็นตัวแทนของกลุ่ม เพื่อให้ตัวสถิติใหม่ที่จะได้มานั้นมีความเหมาะสมกับการที่จะนำไปทดสอบข้อมูลดังกล่าวได้ดียิ่งขึ้น ตัวสถิติที่สร้างขึ้นมาใหม่โดยใช้ชื่อว่า การทดสอบเอฟที่ใช้ค่าเฉลี่ยของ Graybill & Deal

$$\bar{x}_{i.} = \frac{\sum_{j}^{n_i} x_{ij} \Big/ s_i^2}{\sum_{j}^{n_i} 1 \Big/ s_i^2}$$

$$= \frac{\sum_{j}^{n_i} x_{ij}}{n_i}$$

$$\bar{x}_{..} = \frac{\sum_{i}^{k} \bar{x}_{i.}\, n_i \Big/ s_i^2}{\sum_{i}^{k} n_i \Big/ s_i^2}$$

นำ $\bar{x}_{i.}$, $\bar{x}_{..}$ มาแทนในตารางวิเคราะห์ความแปรปรวนของแผนการทดลองแบบสุ่มตลอด

ตารางการวิเคราะห์ความแปรปรวนของแผนการทดลองแบบสุ่มตลอด

ที่มาของความแปรผัน	d.f.	ความแปรผัน	M.S.
ระหว่างสิ่งทดลองหรือทรีทเมนต์	k - 1	$SS_{Tr} = \sum_{i}^{k} n_i(\bar{x}_{i.} - \bar{x}_{..})^2$	$MS_{Tr} = \frac{SS}{k-1}$
ภายในสิ่งทดลองเดียวกัน	N - k	$SS_E = \sum_{i}^{k}\sum_{j}^{n_i}(x_{ij} - \bar{x}_{i.})^2$	$MS_E = \frac{SS_E}{N-k}$
รวม	N - 1	$SS_T = \sum_{i}^{k}\sum_{j}^{n_i}(x_{ij} - \bar{x}_{..})^2$	

ตัวสถิติที่ใช้ในการทดสอบ คือ

$$F_{k-1,\, N-k} = \frac{MS_{Tr}}{MS_E}$$

ซึ่งมีการแจกแจงแบบเอฟ และมีองศาความเป็นอิสระเท่ากับ k - 1 และ N - k สมมติฐานที่กำหนดจะถูกปฏิเสธไป เมื่อ F ที่คำนวณได้จากตัวอย่างมีค่ามากกว่าค่า F ที่เปิดจากตารางที่กำหนดที่ระดับนัยสำคัญที่กำหนด

3. การทดสอบเอฟที่เสนอโดย Brown & Forsythe

เสนอโดย Morton B. and Alan B. Forsythe ในปี ค.ศ.1974 เพื่อใช้เป็นวิธีการทดสอบความเท่ากันของค่าเฉลี่ยประชากรที่มากกว่าสองกลุ่ม ในกรณีที่มีความแปรปรวนของประชากรไม่เท่ากันซึ่งมีรูปแบบดังนี้

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

โดยที่ x_{ij} แทนค่าสังเกตที่ j ในกลุ่มตัวอย่างที่ i

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

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

$$\bar{x}_i = \frac{\sum_{j}^{n_i} x_{ij}}{n_i}$$

$$\bar{x} = \frac{\sum_{j}^{c} n_i \bar{x}_i}{N}$$

$$s_i^2 = \frac{\sum_{j}^{n_i}(x_{ij} - \bar{x}_i)^2}{(n_i - 1)}$$

ภายใต้สมมติฐานหลัก (H_0) ที่เป็นจริง กล่าวคือ $\mu_1 = \mu_2 = \ldots = \mu_c$ แล้วสถิติทดสอบแบบ Brown&Forsythe จะมีลักษณะการแจกแจงแบบเอฟโดยประมาณ (Approximate F Distribution) ที่องศาความเป็นอิสระ (c-1 , f^*) โดยที่

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

$$v_i = \frac{\left(1 - n_i / N\right) s_i^2}{\sum_{i}^{c}\left(1 - n_i / N\right) s_i^2}$$

เกณฑ์การตัดสินใจจะปฏิเสธ H_0 เมื่อ BF มากกว่า $F_{\alpha, (c-1, f^*)}$ โดยที่ $F_{\alpha, (c-1, f^*)}$ คือค่าวิกฤต ที่เปิดจากตาราง F ที่ระดับนัยสำคัญ α และองศาความเป็นอิสระ (c-1, f^*) ตามลำดับ

4. การแปลงข้อมูลเป็นค่าลอคการิทึม (The Logarithmic Transformation)

ในทางทฤษฎีแล้ว ผู้วิจัยควรจะทดสอบข้อมูลที่ได้มาเสียก่อนว่ามีคุณสมบัติเป็นไปตามข้อสมมติหรือไม่ หากข้อมูลไม่เป็นไปตามข้อสมมติก็จำเป็นที่จะต้องแก้ไขด้วยการแปลงข้อมูล(transformation of data) เสียก่อน แล้วจึงนำข้อมูลที่แปลงแล้วมาวิเคราะห์ จึงจะได้ผลถูกต้องตามความเป็นจริง การทดสอบแบบนี้เรียกว่าการทดสอบแบบใช้พารามิเตอร์ (parameter test)

วิธีการแปลงข้อมูลที่เหมาะสม คือการแปลงเป็นค่าลอคการิทึม (the logarithmic transformation) เป็นวิธีการแปลงข้อมูลที่ขจัดปัญหาข้อมูลไม่เป็นไปตามข้อสมมติได้ดีกว่าวิธีอื่นๆ ข้อมูลที่ได้จากการแปลงจะมีลักษณะการแจกแจงแบบปกติโดยประมาณ (has an approximate normal distribution)

วิธีการแปลงข้อมูล

กำหนดให้ X เป็นค่าสังเกตที่ไม่เป็นไปตามข้อสมมติ (original data)

Y เป็นค่าสังเกตที่แปลงแล้ว (transformed data)

การแปลงเป็นค่าลอคการิทึม (the logarithmic transformation) มี 2 แบบ

1. Y = log (X)
2. Y = log (X+1)

การแปลงเป็นค่าลอคการิทึมแบบที่ 2 ให้ผลใกล้เคียงกับแบบที่ 1 แต่ในกรณีที่ค่าสังเกตมีค่าเป็นศูนย์ (0) การแปลงเป็นค่าลอคการิทึมแบบที่ 2 จะให้ผลดีกว่าแบบที่ 1 การแปลงเป็นค่าลอคการิทึม เหมาะสำหรับข้อมูลที่เป็นเลขจำนวนเต็มบวก และมีช่วงกว้างใช้ในกรณีที่ความแปรปรวนมีค่าเป็นสัดส่วนกับกำลังสองของค่าเฉลี่ยของทรีทเมนท์ หรือส่วนเบี่ยงเบนมาตรฐานเป็นสัดส่วนกับค่าเฉลี่ย การแปลงเป็นค่าลอคการิทึม จะมีผลทำให้ข้อมูลในแต่ละทรีทเมนท์มีความแปรปรวนเท่ากัน นอกจากนั้นการแปลงโดยวิธีนี้จะทำให้อิทธิพลต่างๆ ที่มีลักษณะเป็นแบบผลคูณในข้อมูลเดิมเปลี่ยนเป็นแบบบวกในมาตราลอคการิทึม ภายหลังจากการแปลงข้อมูลแล้ว จึงวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานที่ต้องการโดยใช้ F TEST

5. Trimmed F

สถิติทดสอบที่จะกล่าวต่อไปนี้ เป็นสถิติทดสอบที่ได้มีการกำหนดร้อยละของการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจง จากนั้นใช้ค่าเฉลี่ยของข้อมูลส่วนที่เหลือ (g-times trimmed mean) เป็นตัวประมาณค่าเฉลี่ยของประชากร และตัวประมาณค่าความแปรปรวนของประชากร จะคำนวณมาจากผลรวมของส่วนเบี่ยงเบนกำลังสองแบบวินเซอร์ไรซ์ ที่ระดับการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจงด้านละ g % รายละเอียดของการคำนวณมีดังต่อไปนี้

การคำนวณค่าเฉลี่ยแบบทริมด์และผลรวมของส่วนเบี่ยงเบนกำลังสองแบบวินเซอร์ไรซ์
(Trimmed Mean and Winsorized Sum of Squared Deviations)

ถ้ากำหนดให้ n คือขนาดตัวอย่าง และ $X_1, X_2, ..., X_n$ เป็นค่าสังเกต ที่เรียงลำดับจากน้อยไปมาก จะได้ว่า

- ค่าเฉลี่ยแบบทริมด์ (the g-trimmed mean) คำนวณจาก

$$\bar{x}_{tg} = \left[\sum_{j=k+1}^{n-k} x_j + (1-\varepsilon)\left(x_k + x_{n-k+1}\right) \right] \Big/ (n - 2gn)$$

- ค่าเฉลี่ยแบบวินเซอร์ไรซ์ (the g-Winsorized mean) คำนวณจาก

$$\bar{x}_{wg} = \left[\sum_{j=k+1}^{n-k} x_j + k\left\{(1-\varepsilon)\left(x_k + x_{n-k+1}\right) + \varepsilon\left(x_{k+1} + x_{n-k}\right)\right\} \right] \Big/ n$$

โดยที่ g แทนร้อยละของการตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจงด้านละ g%

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

$\varepsilon = gn-[gn]$

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

$$SSD_{wg} = \sum_{j=k+1}^{n-k} \left(x_j - \bar{x}_{wg}\right)^2 + k\left[\left\{(1-\varepsilon)x_k + \varepsilon\, x_{k+1} - \bar{x}_{wg}\right\}^2 + \left\{(1-\varepsilon)x_{n-k+1} + \varepsilon\, x_{n-k} - \bar{x}_{wg}\right\}^2\right]$$

สถิติทดสอบแบบ Trimmed F เสนอโดย H. LEE และ K.Y. FUNG ในปีค.ศ.1983 โดยพัฒนามาจากสถิติทดสอบแบบ Modified F Test ทั้งนี้เนื่องจากสถิติทดสอบแบบ Modified F Test สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ในกรณีที่ประชากรมีความแปรปรวนต่างกันภายใต้การแจกแจงแบบปกติ แต่ถ้าลักษณะการแจกแจงของประชากรเป็น แบบสมมาตรหางยาว (Long-Tailed Distribution) สถิติทดสอบแบบ Modified F Test จะมีอำนาจการทดสอบลดลง ด้วยเหตุผลดังกล่าวทำให้ H. LEE และ K.Y. FUNG ได้นำแนวความคิดเกี่ยวกับ การตัดค่าสังเกตที่ปลายทั้งสองด้านของการแจกแจง มาประยุกต์ใช้กับ
สถิติทดสอบแบบ Modified F Test ซึ่งมีรูปแบบดังนี้

$$F_{t(g)} = \frac{\sum_{i=1}^{c} h_i\,(\bar{x}_{tg_i} - \bar{x}_{tg})^2}{\sum_{i=1}^{c} (1-\frac{h_i}{H})\, s^2_{wg_i}}$$

โดยที่

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

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

$$\bar{x}_{tg_i} = [\sum_{j=k+1}^{n_i-k} x_{ij} + (1-\varepsilon)(x_{ik} + x_{i,n-k+1})] / (n_i - 2gn_i)$$

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

$$\bar{x}_{tg} = \frac{\sum_{j=k+1}^{N-k} x_j + (1-\varepsilon)(x_k + x_{N-k+1})}{N - 2gn}$$

โดยที่

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

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

ε = gn - [gn]

และ

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

$$SSD_{wg_i} = \sum_{j=k+1}^{n_i-k} (x_{ij} - \bar{x}_{wg_i})^2 + k\Big[\big\{(1-\varepsilon)x_{ik} + \varepsilon\, x_{i,k+1} - \bar{x}_{wg_i}\big\}^2 +$$

$$\big\{(1-\varepsilon)x_{i,n_i-k+1} + \varepsilon(i)\, x_{i,n_i-k} - \bar{x}_{wg_i}\big\}^2\Big]$$

$$\bar{x}_{wg_i} = \left[\sum_{j=k+1}^{n_i-k} x_{ij} + k\big\{(1-\varepsilon)(x_{ik} + x_{i,n_i-k+1}) + \varepsilon(x_{i,k+1} + x_{i,n_i-k})\big\}\right] \Big/ n_i$$

ภายใต้สมมติฐานหลักที่เป็นจริง กล่าวคือ H_0 : $\mu_1 = \mu_2 = ... = \mu_c$

$F_{t(g)}$ จะมีลักษณะการแจกแจงแบบเอฟโดยประมาณ (Approximate F Distribution) ที่องศาความเป็นอิสระ (c-1, f_t) โดยที่

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

$$v_i = (1 - h_i / H)s^2_{wg_i} \Big/ \sum_{i=1}^{c}(1 - h_i / H)s^2_{wg_i}$$

เกณฑ์การตัดสินใจ จะปฏิเสธ H_0 เมื่อ $F_{t(g)} > F_{\alpha,(c-1,f_t)}$ โดยที่ $F_{\alpha,(c-1,f_t)}$ คือค่าวิกฤตที่เปิดจากตารางเอฟ ที่ระดับนัยสำคัญ α และองศาความเป็นอิสระ (c-1 , f_t) ตามลำดับ

คุณสมบัติและลักษณะการแจกแจงของประชากรที่ศึกษา

การแจกแจงแบบปกติ (Normal Distribution)

การแจกแจงแบบปกติ เป็นการแจกแจงที่ใช้อธิบายข้อมูลที่เกิดขึ้นโดยทั่วไปทางธรรมชาติ เช่น น้ำหนัก ความสูง ผู้ค้นพบคือ เดอมูลร์ (De Movire : 1667-1754) ชาวฝรั่งเศส ในปี 1733 เดอมูลร์ ได้แสดงว่าการแจกแจงแบบไบโนเมียล (Binomial distribution) สามารถประมาณ ด้วยการแจกแจงแบบปกติ เมื่อ n มีขนาดใหญ่ ต่อมา ลาปลาซ (Lapace : 1749 - 1827) ชาวอังกฤษ ได้นำไปประยุกต์ใช้ในทางสังคมและวิทยาศาสตร์

ฟังก์ชันความหนาแน่น (p.d.f.) ของการแจกแจงแบบปกติคือ

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

$$\sigma^2 > 0$$

โดยที่ f(x) = แทนความสูงของโค้งที่วัดจากแกนนอน ณ จุดใดๆ

σ = ความเบี่ยงเบนมาตรฐานของประชากร

σ^2 = ความแปรปรวนของประชากร

μ = ค่าเฉลี่ยของประชากร

π = 3.14159

e = 2.71828

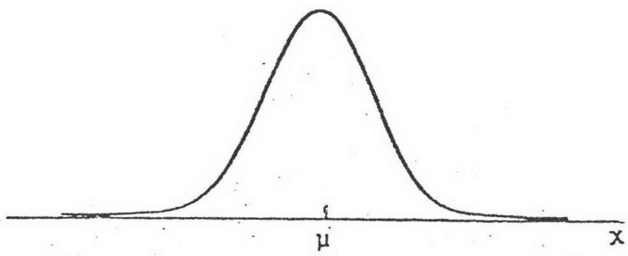
x = ค่าของข้อมูลจากกลุ่มตัวอย่าง

และค่า μ , σ เป็นพารามิเตอร์ที่บอกลักษณะของประชากรว่าประชากรนั้นมีตำแหน่งอยู่ที่ใดและมีการกระจายมากน้อยเพียงใด

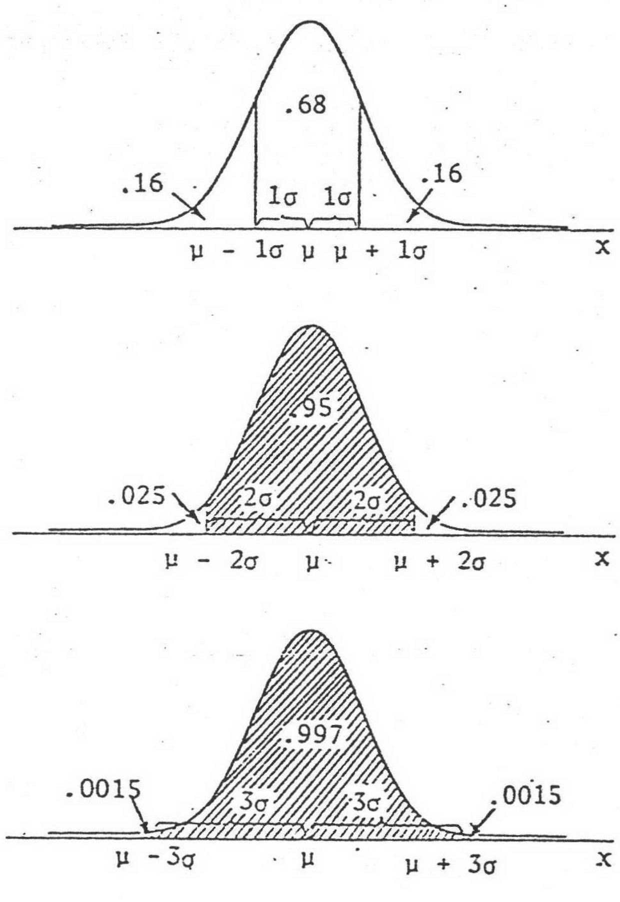
การแจกแจงแบบปกติมีคุณสมบัติดังนี้

1. ลักษณะของโค้งเป็นรูประฆังคว่ำ (Bell Shaped)
2. เส้นแบ่งครึ่งโค้งอยู่ที่จุดที่เป็นค่าเฉลี่ยของข้อมูล และเส้นนี้ทำให้เส้นโค้งที่อยู่สองข้างมีลักษณะสมมาตร (Symmetry)
3. จุดที่เป็นค่าเฉลี่ย มัธยฐาน และฐานนิยม เป็นจุดเดียวกันหรือมีค่าเท่ากัน
4. มีความโด่ง (kurtosis) ของเส้นโค้งเท่ากับ 3 ซึ่งเรียกว่าเมโซเคอร์ติค (Mesokurtic) และจุดเปลี่ยนโค้งทั้งสองข้างอยู่ ณ ตำแหน่ง 1 ส่วนเบี่ยงเบนมาตรฐาน
5. ค่าความเบ้ (Skewness) เท่ากับ 0
6. ปลายทั้งสองข้างของเส้นโค้งจะค่อยๆ ลดต่ำลง และไม่จรดกับฐานของโค้งหรือแกนนอน (x)

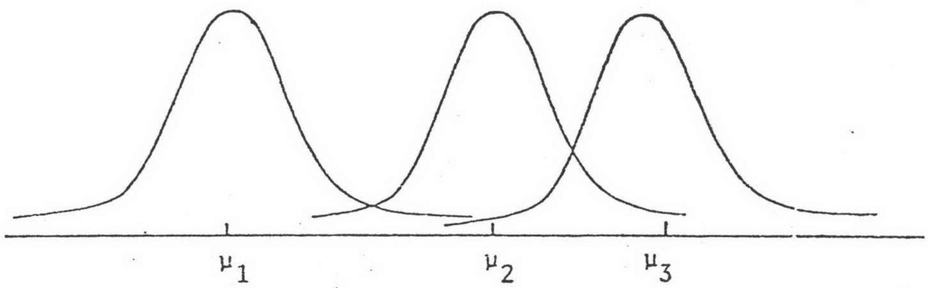
7. ถ้าลากเส้นตั้งฉากจากแกน x ไปยังเส้นโค้ง โดยที่เส้นดังกล่าวห่างจากจุดเฉลี่ย (μ) ทั้งด้านซ้ายและขวาด้วยระยะ 1เท่า 2เท่า และ 3เท่าของค่าเบี่ยงเบนมาตรฐาน ($1\sigma, 2\sigma, 3\sigma$) พื้นที่ที่ปิดกั้นเส้นตั้งฉากกับเส้นโค้งจะเท่ากับ 68% 95% และ 99.7% ของพื้นที่ทั้งหมดตามลำดับ ดูรูปที่ 2.1, 2.2, 2.3 และ 2.4 ประกอบ



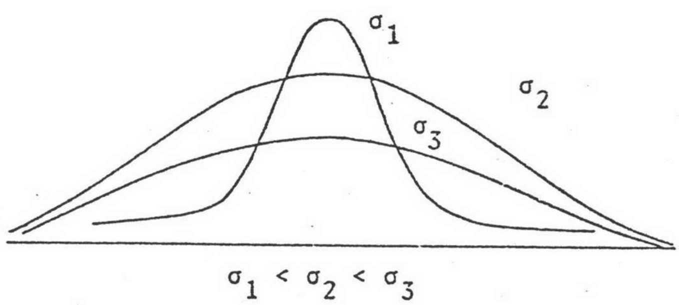
รูปที่ 2.1 แสดงเส้นโค้งการแจกแจงแบบปกติ



รูปที่ 2.2 แสดงพื้นที่ 68% 95% และ99.7% ของเส้นโค้งปกติ



รูปที่ 2.3 แสดงการแจกแจงแบบปกติ 3 รูป ซึ่งมีค่าเฉลี่ยต่างๆ กัน แต่มีค่าเบี่ยงเบนมาตรฐานเท่ากัน



รูปที่ 2.4 แสดงการแจกแจงแบบปกติ 3 รูป ซึ่งมีค่าเบี่ยงเบนมาตรฐานต่างกัน แต่มีค่าเฉลี่ยเท่ากัน

ผลงานวิจัยที่เกี่ยวข้อง

สำหรับผลงานวิจัยที่เกี่ยวข้องในการศึกษาเกี่ยวกับการทดสอบความเท่ากันของค่าเฉลี่ยประชากรในกรณีที่มีหลายประชากรนั้นมีแนวความคิดพื้นฐานมาจากการทดสอบความแตกต่างระหว่างค่าเฉลี่ยที่มาจากสองประชากร ประมาณปี ค.ศ.1937 โดย Welch,B.L. ได้ค้นหาวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของสองประชากร กรณีที่ค่าความแปรปรวนแตกต่างกัน โดยเสนอให้ใช้สถิติทดสอบที (t test) ซึ่งให้ผลไม่เป็นที่พอใจนัก ในกรณีที่การแจกแจงของประชากรไม่เป็นแบบปกติ ปี ค.ศ.1938 Welch,B.L. ได้สรุปว่า ถ้าเราทราบสัดส่วนของความแปรปรวนทั้ง 2 ประชากร ก็สามารถใช้สถิติ U ได้ แต่ถ้าไม่ทราบความแปรปรวนประชากรนั้น R.A. Fisher ได้ใช้ความคิดของการแจกแจง fiducial (fiducial distributions) โดยใช้สถิติ V ทดสอบความแตกต่างของค่าเฉลี่ย 2 กลุ่ม เมื่อความแปรปรวนประชากรไม่เท่ากัน โดยประมาณว่า ทั้ง U และ V มีการแจกแจงแบบที จากแนวความคิดพื้นฐานเกี่ยวกับการทดสอบความเท่ากันของค่าเฉลี่ยสองประชากรนี้เอง จึงทำให้นักสถิติหลายท่านพยายามขยายการทดสอบออกเป็นหลายประชากร เพื่อให้เหมาะสมกับสถานการณ์จริงที่เกิดขึ้น ซึ่งในปี ค.ศ. 1951 Welch และ James ได้เสนอวิธีการทดสอบความแตกต่างระหว่างค่าเฉลี่ย k ประชากร เมื่อความแปรปรวนของประชากรไม่เท่ากัน ซึ่งก็ได้เป็นการขยายแนวความคิดมาจากวิธีการทดสอบของ Welch (1937) ที่ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของสองประชากรในกรณีที่ความแปรปรวนของประชากรต่างกัน และเรียกว่าสถิติทดสอบของ Welch การ์บิลและดีล (Graybill and Deal (1959)) ได้แก้ปัญหาการรวมค่าเฉลี่ยจากตัวอย่างที่สุ่มจากประชากร 2 กลุ่ม เมื่อทราบว่ามีความแปรปรวนไม่เท่ากัน โดยใช้การหาค่าเฉลี่ยถ่วงน้ำหนักโดยให้น้ำหนักนี้เป็นน้ำหนักแบบสุ่ม (random weight) และทำให้ตัวประมาณค่าถ่วงน้ำหนักที่ได้เป็นตัวประมาณค่าที่ไม่เอนเอียงที่ดีกว่า สม่ำเสมอ (uniformly) (มีความแปรปรวนน้อยกว่าตัวประมาณค่าที่ไม่เอนเอียงของประชากรแต่ละกลุ่ม) โดยมีเงื่อนไขว่าขนาดของตัวอย่างสุ่มทั้งสองประชากรต้องมากกว่า 9 นั่นคือ n_1 และ $n_2 > 9$ ปี ค.ศ.1974 Kohr และ Games ได้ศึกษาเปรียบเทียบวิธีการทดสอบของ Welch กับการทดสอบ F ซึ่งจากการศึกษาพบว่า สถิติทดสอบแบบ Welch มีความแกร่งมากกว่าสถิติทดสอบแบบ F เมื่อความแปรปรวนของประชากรแตกต่างกัน โดยเฉพาะอย่างยิ่งเมื่อขนาดตัวอย่างไม่เท่ากัน แต่เมื่อความแปรปรวนของประชากรเท่ากัน สถิติทดสอบแบบ Welch จะมีอำนาจการทดสอบน้อยกว่าสถิติทดสอบแบบ F เพียงเล็กน้อย ซึ่งสอดคล้องกับผลการศึกษาของ Morton B. Brown และ Alan B. Forsythe (1974) ที่ได้ศึกษาเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบที่ใช้ทดสอบความเท่ากันของ

ค่าเฉลี่ยประชากรในกรณีที่มีมากกว่าสองประชากร ในกรณีขนาดตัวอย่างเล็ก โดยศึกษาในกรณีที่ขนาดตัวอย่างเท่ากันและไม่เท่ากัน นอกจากนั้นได้ทดลองเปลี่ยนค่าเฉลี่ยและ
ค่าความแปรปรวนที่ระดับต่างๆกัน ภายใต้การแจกแจงแบบปกติ ซึ่งสถิติทดสอบที่เขานำมาศึกษา คือสถิติทดสอบแบบ ANOVA F TEST สถิติทดสอบแบบ Modified F สถิติทดสอบของ Welch และ James ซึ่งจากการศึกษาพบว่า ในกรณีที่ประชากรที่นำมาทดสอบมีค่าความแปรปรวนแตกต่างกัน สถิติทดสอบแบบ ANOVA F TEST ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีที่ขนาดตัวอย่างเท่ากันและไม่เท่ากันที่ระดับนัยสำคัญ 0.01 และ 0.05 ในขณะที่สถิติทดสอบแบบ Modified F และสถิติทดสอบของ Welch และ James สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีที่ขนาดตัวอย่างเท่ากันและไม่เท่ากัน ที่ระดับนัยสำคัญ0.01 และ 0.05 จากนั้นก็ได้มีนักสถิติหลายท่านทำการเสนอวิธีการต่างๆ เกี่ยวกับการทดสอบความเท่ากันของค่าเฉลี่ยประชากรที่มากกว่าสองกลุ่ม เมื่อความแปรปรวนของประชากรไม่เท่ากัน ในทางปฏิบัติสมมติฐานของความเท่ากันของความแปรปรวนของความคลาดเคลื่อน (equality of error variances) มักไม่ทราบ และจนถึงปัจจุบันนี้ยัง ไม่มีวิธีการพัฒนาที่ถูกต้องที่กล่าวว่าสมมติฐานข้อนี้ต้องตัดทิ้งได้ ซาร์การ์ (Sarkar (1991) ได้ใช้วิธีพิทมันโคลสเนส
(Pitman closeness) ในการเปรียบเทียบตัวประมาณค่าค่าเฉลี่ยร่วมของการ์บิลและคีลกับค่าเฉลี่ยตัวอย่างแต่ละตัว พบว่า ตัวประมาณค่าเฉลี่ยร่วมของการ์บิลและคีลเป็นตัวประมาณค่าที่ดีกว่าค่าเฉลี่ยตัวอย่างแต่ละตัว นอกจากนี้ ยังได้หาเงื่อนไขที่จำเป็นและพอเพียงคือ ขนาดของแต่ละกลุ่มตัวอย่างต้องมากกว่า 5 เมื่อมีประชากร 2 กลุ่ม และขนาดตัวอย่างต้องมากกว่า 8 เมื่อมีประชากรมากกว่า 2 กลุ่ม

จากวิทยานิพนธ์ของนางสาวนันทวัน บำรุงสวัสดิ์ (1991) ได้เสนอวิธีการเปรียบเทียบสถิติทดสอบแบบ Brown และ Forsythe (ดัดแปลงมาจากการวิเคราะห์ความแปรปรวน) และสถิติทดสอบของMarascuilo (ดัดแปลงมาจากสถิติทดสอบของ Welch) ได้ผลสรุปดังนี้

1. เมื่อประชากรมีอัตราส่วนของความแปรปรวนเท่ากันสถิติทดสอบของ Brown และ Forsythe สถิติทดสอบของ Marascuilo และสถิติทดสอบแบบ ANOVA F TEST สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ แต่ถ้าประชากรมีอัตราส่วนของ ความแปรปรวนไม่เท่ากัน สถิติทดสอบของ Brown และ Forsythe และสถิติทดสอบของ Marascuilo สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ดีกว่าสถิติทดสอบแบบ ANOVA F TEST

2. อำนาจของการทดสอบ เมื่อขนาดตัวอย่างเล็กและประชากรมีการแจกแจงแบบปกติ ที่มีค่าอัตราส่วนของความแปรปรวนเท่ากันทุกประชากร สถิติทดสอบแบบ ANOVA F TEST จะมีอำนาจการทดสอบสูงสุด แต่เมื่อขนาดตัวอย่างใหญ่ สถิติทดสอบทั้ง 3 วิธี มีอำนาจการทดสอบสูงสุดเท่ากัน ในทุกระดับนัยสำคัญที่ศึกษา ($\alpha = 0.01$ และ 0.05) แต่เมื่อประชากรมีอัตราส่วนของความแปรปรวนไม่เท่ากัน พบว่าสถิติทดสอบของ Brown และ Forsythe และสถิติทดสอบของ Marascuilo ให้ค่าอำนาจของการทดสอบสูงกว่าสถิติทดสอบแบบ ANOVA F TEST ซึ่งการเลือกใช้สถิติทดสอบของBrown และ Forsythe หรือสถิติทดสอบของ Marascuilo นั้น พบว่าความสัมพันธ์ระหว่างอัตราส่วนของความแปรปรวนและอัตราส่วนของค่าเฉลี่ย มีผลต่ออำนาจของการทดสอบของสถิติทดสอบทั้ง 2 วิธีดังกล่าว