บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในการทดลองทุกประเภท มีหลักสำคัญข้อหนึ่งคือ ผู้ทดลองจะต้องพยายามใช้วิธีการต่าง ๆเพื่อควบคุมหรือลดความแปรปรวนของหน่วยทดลอง(Experimental Unit) ให้มีน้อยที่สุด ซึ่งการควมคุมหรือลดความแปรปรวนของหน่วยทดลองนั้น หากพิจารณาอย่างกว้าง ๆ อาจแยกได้เป็น 2 ประเภทคือ การควมคุมโดยตรงและการควมคุมโดยทางอ้อม โดยใช้หลักสถิติเข้าช่วย

การควมคุมความแปรปรวนโดยตรงนี้ สามารถทำได้โดยอาศัยเทคนิคต่าง ๆ ประกอบการทดลอง เช่น ใช้หน่วยทดลองที่สม่ำเสมอ เลือกสภาพแวดล้อมที่เหมาะสม เป็นต้น แต่ในบางครั้ง การควมคุมความแปรปรวนของแผนการทดลองโดยตรงนั้น ไม่สามารถทำได้อย่างมีประสิทธิภาพเพียงพอ อันเนื่องมาจากสาเหตุหลายประการ ในกรณีเช่นนี้ ผู้ทดลองหรือนักวิจัยควรใช้หลักสถิติเข้าช่วยในการวิเคราะห์และตีความหมายของข้อมูล เพื่อให้ผลการทดลองมีความถูกต้องยิ่งขึ้น โดยเฉพาะอย่างยิ่ง การวิเคราะห์ข้อมูลเพื่อทดสอบความแตกต่างระหว่างอิทธิพลของทรีทเมนต์ (Treatment Effect) เมื่อมีการเก็บข้อมูลของหน่วยทดลอง ก่อนที่จะได้รับอิทธิพลของทรีทเมนต์ (ตัวแปรร่วม) และเก็บข้อมูลของหน่วยทดลอง หลังจากที่ได้รับ - อิทธิพลของทรีทเมนต์แล้ว (ตัวแปรตาม) ในกรณีเช่นนี้วิธีการวิเคราะห์ทางสถิติที่เหมาะสม ที่นิยมใช้กัน ก็คือ **วิธีวิเคราะห์ความแปรปรวนร่วม (Analysis of Covariance : ANCOVA)**

การวิเคราะห์ความแปรปรวนร่วม เป็นเทคนิคการควมคุมทางสถิติวิธีหนึ่งที่นำมาใช้ลดความแปรปรวนที่เกิดขึ้นในการทดลอง โดยรวมการวิเคราะห์ความแปรปรวนกับการวิเคราะห์การถดถอยเข้าด้วยกัน วิธีนี้ใช้การวัดตัวแปรเพิ่มขึ้นซึ่งเรียกว่า ตัวแปรร่วม (Covariate หรือ Concomitant Variables) เพิ่มเติมจากการวัดตัวแปรตามที่สนใจศึกษา ตัวแปรร่วม (X) นี้เป็นตัวแปรที่มีผลต่อตัวแปรตาม (Y) ดังนั้นเพื่อให้การวิเคราะห์ผลการวัดตัวแปรตามถูกต้องมากยิ่งขึ้น จึงต้องอาศัยตัวแปรร่วมมาเป็นตัวปรับในการวิเคราะห์ข้อมูล ด้วยเหตุนี้ จึงอาจกล่าวได้อีกนัยหนึ่งว่า การวิเคราะห์ความแปรปรวนร่วม เป็นการวิเคราะห์ความแปรปรวนของสิ่งที่ต้องการศึกษา หลังจากที่ได้ขจัดค่าของตัวแปรที่เกี่ยวข้อง หรือตัวแปรร่วมออกไปก่อน (สมจิต วัฒนาชยากุล , 2527 หน้า 220)

จุดมุ่งหมายหลัก ในการวิเคราะห์ความแปรปรวนร่วมก็คือ ต้องการหาข้อสรุปว่าอิทธิพลของทรีทเมนต์แตกต่างกันหรือไม่ ซึ่งผลของการวิเคราะห์จะถูกต้องได้ผลตรงตามเป้าหมาย ก็ต่อเมื่อข้อมูลที่จะนำมาวิเคราะห์ มีคุณสมบัติเป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนร่วมเท่านั้น ซึ่งข้อตกลงเบื้องต้นข้อหนึ่งก็คือ ลักษณะการแจกแจงของความคลาดเคลื่อนจะต้องมีลักษณะการแจกแจงแบบปกติ แต่ในทางปฏิบัติปัญหาที่นักวิจัยมักพบอยู่เสมอก็คือ ลักษณะการแจกแจงของความคลาดเคลื่อน อาจจะไม่เป็นแบบปกติ เช่น อาจจะมีการแจกแจงแบบหางยาว โดยที่ปลายหางของการแจกแจงมีลักษณะยาวกว่าปลายหางของการแจกแจงแบบปกติ ซึ่งเราจะเรียกการแจกแจงชนิดนี้ว่าการแจกแจงชนิดลอง-เทลด์(Long-tailed Distribution) เมื่อข้อมูลไม่มีคุณสมบัติเป็นไปตามข้อตกลงเบื้องต้นเช่นนี้ หากนักวิจัยยังคงใช้วิธีการวิเคราะห์ความแปรปรวนร่วมอยู่ ผลสรุปที่ได้อาจจะมีความเชื่อถือได้น้อย จากปัญหาที่เกิดนี้เอง ผู้วิจัยจึงต้องการหาตัวสถิติทดสอบแบบไม่ใช้พารามิเตอร์ (Nonparametric Test) มาช่วยแก้ปัญหาในกรณีที่ข้อมูลไม่สอดคล้องกับข้อตกลงเบื้องต้น ว่าสถิติทดสอบแบบไม่ใช้พารามิเตอร์ตัวใด จะให้ผลการทดสอบที่มีความถูกต้องน่าเชื่อถือได้มากกว่ากัน เนื่องจากสถิติทดสอบแบบไม่ใช้พารามิเตอร์เป็นสถิติทดสอบที่ไม่มีข้อจำกัดเกี่ยวกับลักษณะการแจกแจงความคลาดเคลื่อน อีกทั้งยังสามารถคำนวณได้รวดเร็ว เข้าใจง่ายและสะดวกในการนำไปใช้อีกด้วย

1.2 วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบประสิทธิภาพของการทดสอบแบบไม่ใช้พารามิเตอร์ 2 วิธี ในการวิเคราะห์ความแปรปรวนร่วม คือ วิธี Quade's Rank และวิธี Rank Transformation เมื่อความคลาดเคลื่อนมีการแจกแจงแบบโลจิสติก ดับเบิ้ลเอ็กซ์โปเนนเชียล และปกติปลอมปน โดยพิจารณาความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 (Type I Error) และอำนาจของการทดสอบ (Power of the Test)

1.3 สมมุติฐานของการวิจัย

1.3.1 ภายใต้ลักษณะการแจกแจงของความคลาดเคลื่อนต่าง ๆ กัน วิธีวิเคราะห์แบบ Rank Transformation จะมีค่าอำนาจการทดสอบสูงกว่า วิธีวิเคราะห์แบบ Quade' s Rank

1.3.2 ขนาดตัวอย่าง สัมประสิทธิ์สหสัมพันธ์ และสัมประสิทธิ์ความแปรปรวนจะมีผลทำให้ค่าอำนาจการทดสอบแตกต่างกัน

1.4 ขอบเขตของการวิจัย

1.4.1 แบบหุ่น (Model) ที่ใช้ศึกษาในการวิจัยครั้งนี้เป็นแบบอิทธิพลกำหนด (Fixed Effect Model) ในแผนการทดลองแบบสุ่มสมบูรณ์ (Completely Randomized Design : CRD) ดังนี้

$$Y_{ij} = \mu + T_i + \beta(X_{ij} - \overline{X}..) + \epsilon_{ij} \quad ; \quad i = 1,2,\ldots,t \quad j = 1,2,\ldots,n_i$$

โดยที่ Y_{ij} หมายถึง ค่าสังเกตของตัวแปรตามจากหน่วยทดลองที่ j ในทรีทเมนต์ที่ i

μ หมายถึง ค่าเฉลี่ยของประชากร

T_i หมายถึง อิทธิพลของทรีทเมนต์ที่ i

β หมายถึง สัมประสิทธิ์การถดถอยรวมของ Y บน X

$X_{ij} - \overline{X}..$ หมายถึง ค่าเบี่ยงเบนของตัวแปรร่วมที่ (i,j) จากค่าเฉลี่ยทั้งหมดของตัวแปรร่วม

ϵ_{ij} หมายถึง ค่าความคลาดเคลื่อนของการทดลองจากหน่วยทดลองที่ (i,j) โดยกำหนดให้มีตัวแปรร่วม 1 ตัวแปร เท่านั้น

1.4.2 ลักษณะการแจกแจงความคลาดเคลื่อนที่นำมาทดสอบ จะมีค่าเฉลี่ย (μ) เป็น 0 และความแปรปรวนเป็น σ^2 ทุกรูปแบบการศึกษา โดยมีลักษณะการแจกแจงความคลาดเคลื่อนที่สนใจศึกษา ดังนี้

1.4.2.1 การแจกแจงแบบโลจิสติก (Logistic Distribution)

1.4.2.2 การแจกแจงแบบดับเบิ้ลเอ็กซ์โปเนนเชียล (Double Exponential Distribution)

1.4.2.3 การแจกแจงแบบปกติปลอมปน (Scale Contaminated Normal Distribution) โดยจะศึกษาเมื่อเปอร์เซนต์การปลอมปน (PC) เท่ากับ 10% และ 25% สำหรับสเกลแฟคเตอร์ (Scale factor : C) 2 ระดับ คือ 3 และ 10

1.4.3 กำหนดจำนวนกลุ่มตัวอย่างเท่ากับ 3 กลุ่ม

1.4.4 ขนาดตัวอย่างที่ใช้ในการวิจัย เป็นดังนี้

1.4.4.1 กรณีขนาดตัวอย่างทั้ง 3 กลุ่มเท่ากัน จะศึกษาเมื่อขนาดตัวอย่างแต่ละกลุ่มตัวอย่างเป็น 4,5,6,8 และ 10 คือ $n = 12,15,18,24$ และ 30

1.4.4.2 กรณีขนาดตัวอย่างทั้ง 3 กลุ่มไม่เท่ากัน จะศึกษาเมื่อขนาดตัวอย่างแต่ละกลุ่มตัวอย่างเป็น(3,4,5 : n=12),(4,5,6 : n=15),(4,6,8:n=18),(6,8,10 : n=24) และ (8,10,12 : n=30)

1.4.5 กำหนดให้ค่าเฉลี่ยของประชากร (μ) เท่ากับ 100 และมีความแปรปรวนของประชากร (σ^2) เท่ากันทุกกลุ่มเป็น 100

1.4.6 กำหนดสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรร่วม (ρ_{xy})ดังนี้ 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 และ 1.0

1.4.7 กำหนดสัมประสิทธิ์ความแปรปรวน (Coefficient of variation:C.V.) 3 ขนาด คือ 10% 15% และ 20%

1.4.8 กำหนดให้ค่าสัมประสิทธิ์การถดถอยรวม (β) เท่ากับ 1.2

1.4.9 กำหนดอิทธิพลของสิ่งทดลอง (T_i) โดยให้ $\sum_i T_i = 0$

1.4.10 กำหนดระดับนัยสำคัญของการทดสอบ 2 ระดับ คือ 0.01 และ 0.05

หมายเหตุ

1. สเกลแฟคเตอร์ เป็นค่าที่ปรับให้ข้อมูลมีการกระจายมากขึ้น โดยสเกลแฟคเตอร์ที่มีค่าสูง อาจทำให้เกิดค่าผิดปกติขึ้นในข้อมูล และจากการศึกษาที่ผ่านมาสเกลแฟคเตอร์ที่มีค่าน้อยกว่า 3 จะมีโอกาสของการเกิดค่าผิดปกติน้อย ส่วนสเกลแฟคเตอร์ที่มีค่ามากกว่า 10 จะมีโอกาสของการเกิดค่าผิดปกติมาก จึงทำการศึกษาเพียง 2 ระดับดังกล่าว (จีรพร วีระพันธุ์,2530)

2. กำหนดให้มีขนาดตัวอย่างเล็กเป็นส่วนมาก เพราะว่าการวิจัยนี้เป็นการเปรียบเทียบประสิทธิภาพของสถิติทดสอบแบบไม่ใช้พารามิเตอร์ ถ้ากำหนดขนาดตัวอย่างที่ใหญ่แล้ว การใช้สถิติทดสอบแบบไม่ใช้พารามิเตอร์ จะให้ผลใกล้เคียงกับสถิติทดสอบแบบใช้พารามิเตอร์

3. ค่าความแปรปรวนเท่ากับ 100 นี้ไม่มีผลต่อค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 (เลิศสรรพ์ เมฆสุด,2530)

4. ในกรณีที่ใช้จำนวนกลุ่มตัวอย่างเท่ากับ 3 กลุ่ม ค่า $\beta = 1.2$ จะให้ค่าอำนาจการทดสอบสูงสุด (D.R.COX : 139-140)

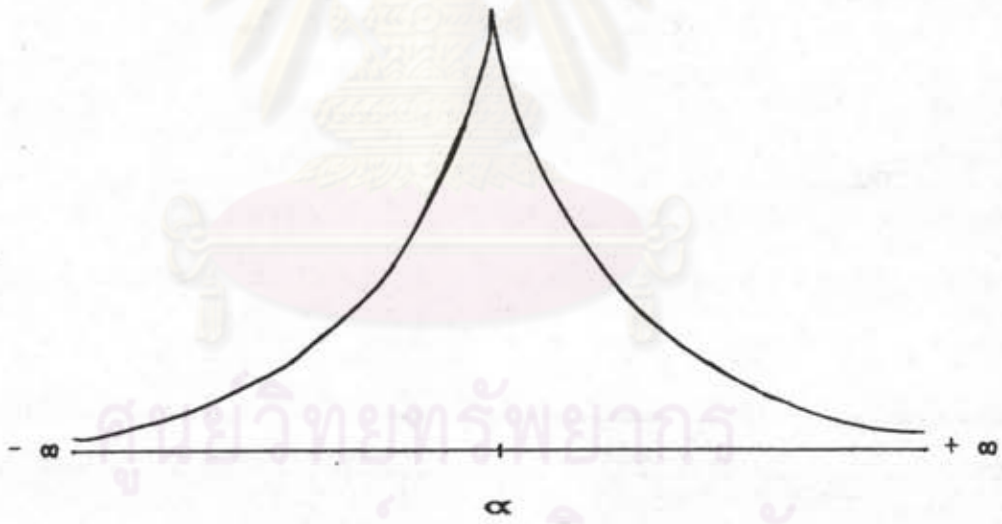
1.4.11 สร้างแบบจำลองข้อมูลให้มีสถานการณ์ตามที่ต้องการศึกษา ด้วยวิธีมอนติคาร์โลซิมูเลชั่น (Monte Carlo Simulation Technique) จากเครื่องคอมพิวเตอร์

IBM 370/3031 โดยทำการทดลองซ้ำ ๆ กัน 1,000 ครั้ง ในแต่ละสถานการณ์ ซึ่งศึกษาเมื่อความคลาดเคลื่อนมีการแจกแจงแบบโลจิสติก แบบดับเบิ้ลเอ็กซ์โปเนนเชียล และ แบบปกติปลอมปน โดยมีค่าฟังก์ชันความน่าจะเป็น ค่าคาดหวัง และค่าความแปรปรวนของการแจกแจงแต่ละรูปแบบเป็นดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1. การแจกแจงแบบโลจิสติก
จะมีฟังก์ชันความน่าจะเป็น ดังนี้คือ

$$f(x) = \frac{1}{\beta}\,\frac{e^{-\frac{(x-\alpha)}{\beta}}}{\left[1+e^{-\frac{(x-\alpha)}{\beta}}\right]^2}, \quad -\infty < x < +\infty, \quad -\infty < \alpha < +\infty, \quad \beta > 0$$



ค่าคาดหวัง $E(X) = \alpha$

ค่าความแปรปรวน $V(X) = (1/3)\,\pi^2\,\beta^2$

ความเบ้ (skewness) = 0

ความโด่ง (kurtosis) = 4.2

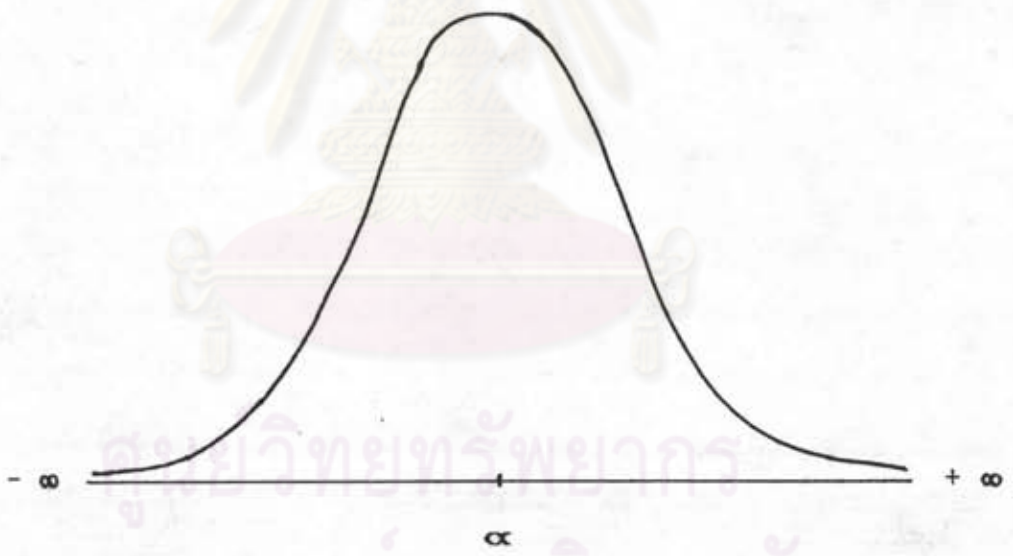
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2. การแจกแจงแบบดับเบิ้ลเอ็กซ์โปเนนเชียล (Double Exponential Distribution) จะมีฟังก์ชันความน่าจะเป็น ดังนี้คือ

$$f(x) = \frac{1}{2\beta} e^{\frac{-|x-\alpha|}{\beta}}, \quad -\infty < x < +\infty$$

$$-\infty < \alpha < +\infty$$

$$\beta > 0$$



ค่าคาดหวัง $E(X) = \alpha$

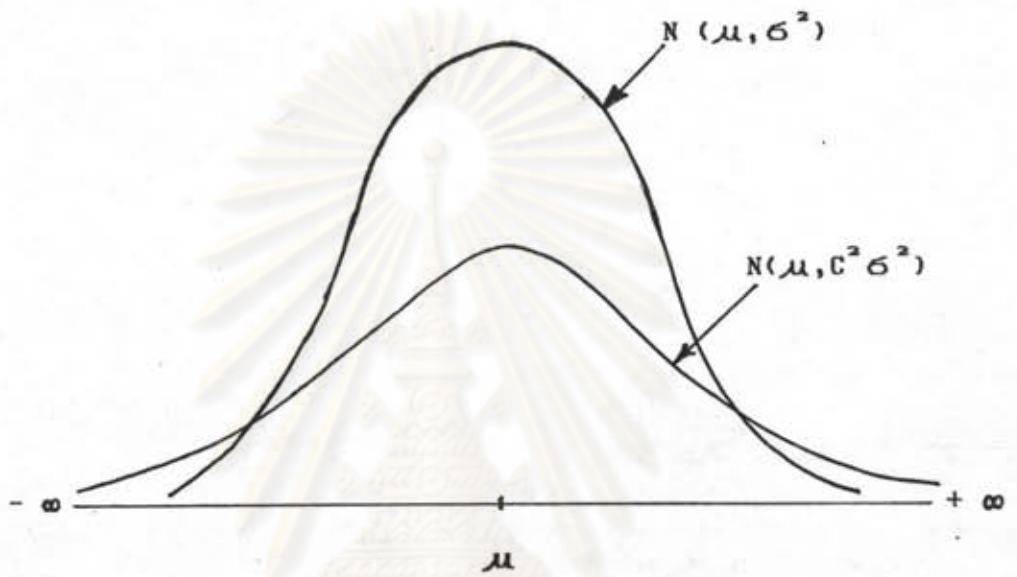
ค่าความแปรปรวน $V(X) = 2\beta^2$

ความเบ้ $= 0$

ความโด่ง $= 6.0$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3. การแจกแจงแบบปกติปลอมปน



ลักษณะการแจกแจงแบบปกติปลอมปนที่พิจารณาในการวิจัยนี้ เป็นการแจกแจงที่แปลงมาจากการแจกแจงแบบปกติ ซึ่งมีฟังก์ชันการแปลงเป็น ดังนี้

$$F = (1-PC)N(\mu,\sigma^2) + (PC)N(\mu,C^2\sigma^2), \qquad C>0$$

หมายความว่า ค่า x จะมาจากการแจกแจง $N(\mu,\sigma^2)$ ด้วยความน่าจะเป็น 1-PC และการแจกแจง $N(\mu,C^2\sigma^2)$ ด้วยความน่าจะเป็น PC

μ และ σ^2 เป็นพารามิเตอร์ที่กำหนดค่าเฉลี่ย และค่าความแปรปรวนของความคลาดเคลื่อน

PC และ C เป็นพารามิเตอร์ที่กำหนดเปอร์เซนต์ของการปลอมปนและสเกลแฟคเตอร์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 เพื่อเป็นแนวทางให้ผู้วิจัย สามารถเลือกใช้สถิติทดสอบแบบไม่ใช้พารามิเตอร์ของแผนงานทดลองแบบทดลองแบบสุ่มสมบูรณ์ในการวิเคราะห์ความแปรปรวนร่วม ได้อย่างเหมาะสม เมื่อข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงของความคลาดเคลื่อน

1.5.2 เพื่อช่วยให้นักวิจัยสามารถทำการวิเคราะห์ข้อมูลได้สะดวก รวดเร็ว โดยไม่ต้องมาคำนึงถึงลักษณะของข้อมูลในการทดลองมากนัก อีกทั้งให้ผลถูกต้องตรงกับการวิเคราะห์โดยใช้ F-test ในการวิเคราะห์ความแปรปรวนร่วม

1.6 คำจำกัดความของคำต่าง ๆ ที่ใช้ในการวิจัย

1.6.1 ตัวแปรตาม หมายถึง ตัวแปรที่มีค่าเป็นค่าของหน่วยทดลองที่วัดได้หลังจากที่หน่วยทดลองได้รับอิทธิพลของทรีทเมนต์แล้ว

1.6.2 ตัวแปรร่วม หมายถึง ตัวแปรที่มีค่าเป็นค่าของหน่วยทดลองที่วัดได้ก่อนที่หน่วยทดลองจะได้รับอิทธิพลของทรีทเมนต์ ซึ่งเป็นค่าของตัวแปรที่แฝงมากับหน่วยทดลอง

1.6.3 ความผิดพลาดประเภทที่ 1 (Type I Error) คือความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐานว่าง (Null Hypothesis:H_0) เมื่อสมมติฐานว่างนั้นเป็นจริง

1.6.4 ความผิดพลาดประเภทที่ 2 (Type II Error) คือความผิดพลาดที่เกิดจากการยอมรับสมมติฐานว่าง เมื่อสมมติฐานว่างนั้นเป็นเท็จ

1.6.5 อำนาจการทดสอบ คือ ความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างนั้นเป็นเท็จ

1.6.6 สถิติที่ใช้พารามิเตอร์ คือสถิติอนุมานซึ่งมีลักษณะดังนี้ ข้อมูลที่ใช้เป็นแบบ interval scale หรือ ratio scale ซึ่งจะเกี่ยวข้องกับค่าพารามิเตอร์ และมีข้อตกลงต่างๆ เช่น ตัวอย่างต้องสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ ค่าสังเกตเป็นอิสระต่อกัน และสำหรับกรณีการอนุมานที่มีสองหรือมากกว่าสองตัวอย่างแล้ว จะมีข้อตกลงเพิ่มเติมว่า จะต้องเป็นตัวอย่างที่สุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ ซึ่งมีความแปรปรวนเท่ากัน

1.6.7 สถิติที่ไม่ใช้พารามิเตอร์ คือสถิติอนุมานที่ใช้เมื่อ ข้อตกลงเบื้องต้นของการอนุมานแบบพาราเมตริกไม่เป็นไปตามต้องการ และเป็นวิธีที่ไม่ขึ้นอยู่กับลักษณะการแจกแจงของประชากรซึ่งไม่เกี่ยวข้องกับค่าพารามิเตอร์ของประชากร