

### 1.1 ความเป็นมาและความสำคัญของปัญหา

หลักการวิเคราะห์โควาเรียนซ์หรือการวิเคราะห์ตัวแปรร่วม (Analysis of Covariance) ถูกนำมาใช้ในการวิเคราะห์เพื่อทดสอบสมมติฐานความแตกต่างระหว่างอิทธิพลของทรีตเมนต์ เมื่อค่าสังเกตที่ได้จากการศึกษาอยู่ที่อิทธิพลของทรีตเมนต์และปัจจัยอื่น ๆ นอกเหนือจาก Treatment ซึ่งเราเรียกว่า ตัวแปรร่วม (Concomitant Variable or Covariate) ถ้าเราไม่คำนึงถึงตัวแปรร่วม (Covariate) โดยการวิเคราะห์โควาเรียนซ์ การเปรียบเทียบทรีตเมนต์จะมีความคลาดเคลื่อนอันเนื่องมาจากตัวแปรร่วม การใช้วิธีวิเคราะห์โควาเรียนซ์จะลดความคลาดเคลื่อนนี้ได้ (จรัส สันทสักษณ์, 2523)

การทดสอบสมมติฐานเพื่อหาข้อสรุปในการวิจัย ผู้วิจัยจำเป็นต้องเลือกสถิติทดสอบที่เหมาะสมกับลักษณะของข้อมูล และต้องคำนึงถึงข้อตกลงเบื้องต้นของแต่ละวิธีนั้นด้วย ดังนั้น ถ้าเลือกใช้สถิติทดสอบที่เหมาะสมจะทำให้ผลสรุปของการวิจัยเป็นไปอย่างถูกต้อง สมเหตุสมผล และให้อำนาจการทดสอบสูง ในการวิเคราะห์ตัวแปรร่วมก็เช่นกัน การเลือกใช้วิธีเหมาะสมกับลักษณะของข้อมูล ถ้าลักษณะของข้อมูลเป็นไปตามข้อตกลงเบื้องต้น ผู้วิจัยส่วนมากเลือกใช้วิธี วิเคราะห์ความแปรปรวนร่วมแบบพาราเมตริกซ์ (Parametric ANCOVA) ในการทดสอบ ซึ่งเป็นสถิติพาราเมตริกซ์ (Parametric Statistics) ที่มีวิธีการ ยุ่งยากซับซ้อนและใช้เวลานานในการวิเคราะห์ห้มาก ตลอดจนความเข้มงวดต่อข้อสมมติเบื้องต้น ทำให้การวิจัยบางอย่างปฏิบัติไม่ได้ ภายใต้อำนาจการวิจัยบางอย่างเราไม่สามารถที่จะระบุลงไปอย่างชัดเจนว่าการแจกแจงของประชากรเป็นแบบใด และไม่อาจตั้งข้อสมมติฐานว่าการแจกแจงของประชากรเป็นไปตามเงื่อนไข การที่จะให้ข้อมูลที่เกิดจากผลการทดลองมีคุณสมบัติสอดคล้องกับข้อสมมติเบื้องต้นดังกล่าวจึงทำได้ยาก ในการวิจัยโดยทั่ว ๆ ไป ปัญหาที่ผู้วิจัยพบมากก็คือ ลักษณะการแจกแจงความคลาดเคลื่อน (error) ไม่เป็นไปตามข้อตกลงเบื้องต้น คือไม่เป็นการแจกแจงแบบปกติ เช่น อาจเป็นการแจกแจงที่มีหางยาว (long tails) ในกรณีเช่นนี้ ผู้วิจัยอาจจะเลือกใช้สถิติทดสอบนั้น โดยยอมฝ่าฝืนข้อตกลงเบื้องต้น หรือมีเช่นนั้นก็อาจจะเลือกใช้วิธีวิเคราะห์

ตัวแปรร่วม แบบนอนพาราเมตริกซ์ (Nonparametric ANCOVA) ซึ่งเป็นสถิตินอนพาราเมตริกซ์ (Nonparametric Statistics) และเป็นวิธีที่ไม่มีข้อจำกัดเกี่ยวกับลักษณะการแจกแจงความคลาดเคลื่อน (error) สามารถคำนวณได้รวดเร็ว เข้าใจง่าย และสะดวกในการนำไปใช้ นับเป็นอีกวิธีหนึ่งที่ผู้วิจัยสามารถเลือกใช้ได้

เมื่อลักษณะข้อมูลไม่สอดคล้องกับข้อตกลงเบื้องต้น ในกรณีที่มีการแจกแจงความคลาดเคลื่อนไม่เป็นแบบปกติ การฝ่าฝืนข้อตกลงจะมีผลกระทบทำให้ประสิทธิภาพในการวิเคราะห์ลดลง สิ่งมีข้อสงสัยในการนำวิธีวิเคราะห์ตัวแปรร่วมแบบพาราเมตริกซ์กับนอนพาราเมตริกซ์มาวิเคราะห์ข้อมูล ว่าการวิเคราะห์ความแปรปรวนร่วมแบบใดจะให้ผลสรุปที่ถูกต้อง และเหมาะสมกับลักษณะข้อมูลดังกล่าวแล้วได้ดีกว่ากัน

การพิจารณาความเหมาะสมของสถิติทดสอบในการทดสอบสมมติฐานว่าทรีตเมนต์ไม่มีความแตกต่างกันนั้น สิ่งที่เราควรพิจารณาก็คือ อำนาจการทดสอบ และความแกร่ง (Robustness) โดยพิจารณาในลักษณะที่ว่า สถิติทดสอบนั้นจะต้องมีความไว (Sensitive) ต่อการเปลี่ยนแปลงของปัจจัยที่ต้องการทดสอบ และจะต้องไม่มีความไว (insensitive) ต่อการเปลี่ยนแปลงของสิ่งอื่นที่ไม่ใช่ปัจจัยที่ต้องการทดสอบ ในการที่จะเลือกใช้สถิติทดสอบ เราต้องเริ่มพิจารณาถึงความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 ก่อน แล้วจึงจะพิจารณาถึงอำนาจการทดสอบเป็นลำดับต่อไป โดยมีขั้นตอนดังนี้คือ ให้ความน่าจะเป็นที่ยอมให้เกิดความผิดพลาดประเภทที่ 1 ไม่เกิน  $\alpha$  ที่กำหนดไว้ และเมื่อเป็นไปตามเงื่อนไขดังกล่าวแล้ว สิ่งที่จะใช้ในการพิจารณาเลือก สถิติทดสอบอีกก็คือ เลือกสถิติทดสอบที่มีโอกาสน้อยที่สุดที่จะยอมรับสมมติฐาน  $H_0$  เมื่อสมมติฐาน  $H_0$  นั้นผิด ซึ่งหมายความว่าให้อำนาจของการทดสอบสูงที่สุด

## 1.2 วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบอำนาจการทดสอบ (Power of the test) แบบพาราเมตริกซ์กับนอนพาราเมตริกซ์ เมื่อการแจกแจงความคลาดเคลื่อน (error) มีรูปแบบต่าง ๆ กันในการวิเคราะห์ตัวแปรร่วม

### 1.3 สมมติฐานของการวิจัย

1. เมื่อการแจกแจงความคลาดเคลื่อน (error) มีรูปแบบต่างกัน จะมีผลทำให้วิธีการทดสอบแบบพารามตริกซ์กับนอนพารา เมตริกซ์ จะมีอำนาจการทดสอบแตกต่างกัน
2. ขนาดตัวอย่าง และสัมประสิทธิ์ความแปรปรวน (Coefficient of Variation) มีผลทำให้ค่าอำนาจการทดสอบแตกต่างกัน

### 1.4 ข้อตกลงของการวิจัย

โดยพิจารณาจากโมเดล 
$$y_{ij} = \mu + \beta(x_{ij} - \bar{x}_{..}) + \tau_i + \epsilon_{ij}$$

;  $i = 1, 2, \dots, t$   
 $j = 1, 2, \dots, n$

ดังนี้คือ

1. ตัวแปรร่วมมีความสัมพันธ์เชิงเส้นตรงกับตัวแปรตาม
2.  $\tau_i$  เป็นอิทธิพลของทรีตเมนต์ (Treatment effect) ที่  $i$  โดยที่
 
$$\sum_{i=1}^t \tau_i = 0$$
3. ความคลาดเคลื่อนเป็นตัวแปรสุ่มต่อเนื่องมีการแจกแจงแบบเดียวกัน และเป็นอิสระซึ่งกันและกัน

### 1.5 ขอบเขตของการวิจัย

1. โมเดลเป็น Fixed effect
2. ลักษณะการแจกแจงความคลาดเคลื่อนที่นำมาทดสอบ โดยมีค่าเฉลี่ย ( $\mu$ ) เป็น 0 และความแปรปรวนเป็น  $\sigma^2$  ทุกรูปแบบการศึกษา โดยมีการแจกแจงความคลาดเคลื่อนดังนี้
  - 2.1 การแจกแจงแบบโลจิสติก (Logistic Distribution)
  - 2.2 การแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียล (Double Exponential Distribution)
  - 2.3 การแจกแจงแบบปกติ (Normal Distribution)

3. พิจารณาขนาดประชากรเป็น 3 ขนาด คือ 3 4 และ 5 ประชากร
4. ขนาดตัวอย่างแต่ละกลุ่มเท่ากัน 4 ขนาด คือ 5 15 30 และ 50
5. จำนวนตัวแปรร่วม (Covariate)  $X$  1 ตัวแปร แบ่งเป็น 2 ประเภท
  - 5.1 เมื่อ  $X$  มีลักษณะเป็นช่วงเท่ากัน
  - 5.2 เมื่อ  $X$  มีลักษณะเป็นช่วงไม่เท่ากัน เพื่อให้ได้ตัวแปรที่มีลักษณะคล้ายธรรม-

ชาติมากที่สุด จึงกำหนดให้ตัวแปรร่วม  $X \sim N(30, 100)$

6. พิจารณาระดับนัยสำคัญเป็น 3 ระดับ 0.05 0.01 และ 0.10
7. ประชากรที่ศึกษาล้างมาจาก โมเดล

$$Y_{ij} = \mu + \beta (X_{ij} - \bar{X}_{..}) + \tau_i + \epsilon_{ij}$$

$$i = 1, 2, \dots, t$$

$$j = 1, 2, \dots, n$$

7.1 กำหนดให้ค่าเฉลี่ยของประชากรเท่ากันทุกกลุ่ม โดยใช้  $\mu = 100$

7.2 กำหนดให้ค่าถดถอยรวมของตัวแปรร่วม ( $X$ ) และตัวแปรตาม ( $Y$ )

$$\text{คือ } \beta = 1$$

7.3 สร้างอิทธิพลของสิ่งทดลอง ( $\tau_i$ ) ให้แตกต่างกัน โดยพิจารณา

$$\sum_{i=1}^t \tau_i = 0$$

8. พิจารณาสัมประสิทธิ์ของความแปรปรวน 3 ขนาด คือ 5% 10% และ 15%
9. การจำลองกระทำซ้ำ ๆ กัน ไม่น้อยกว่า 300 ครั้ง ในแต่ละสถานการณ์

ของการทดลอง

หมายเหตุ 1. การสร้างค่าคงที่  $X$  สร้างจากการแจกแจงแบบปกติ เพื่อให้เกิดค่าที่เป็นธรรมชาติอย่างไรก็ตามที่กำหนดค่าเฉลี่ยเป็น 30 และค่าแปรปรวนเป็น 100 เพื่อให้ได้ค่าที่ไม่ติดลบ และมีค่าห่างกันพอสมควร

2. การกำหนดค่า  $\beta = 1$  เพราะจากเหตุผลทางทฤษฎีพิจารณาแล้วพบว่าจาก

$$\text{ตัวแบบ } y_{ij} = \mu + \beta (x_{ij} - \bar{x}_{..}) + \tau_i + \epsilon_{ij}$$

ในการปรับค่า  $y_{ij}$  ในการวิเคราะห์ห้ตัวแปรร่วมจะลบด้วย  $\beta (x_{ij} - \bar{x}_{..})$

$$y_{ij} (\text{adj}) = y_{ij} - \beta(x_{ij} - \bar{x}_{..}) = \mu + \tau_i + \epsilon_{ij}$$

ซึ่งไม่ขึ้นอยู่กับ  $\beta$  และอีกเหตุผลหนึ่งคือ จากการทดลองกระทำที่ขนาดตัวอย่าง 10 ทรีตเมนต์ 5 พบว่า ค่า  $\beta$  ไม่มีผลต่อค่าความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 และค่าอำนาจของการทดสอบ

#### 1.6 คำจำกัดความของค่าต่าง ๆ ที่ใช้ในการวิจัย

ตัวแปรตาม (Dependent Variable ; Y) หมายถึง ค่าของหน่วยทดลองที่วัดได้ หลังจากหน่วยทดลองได้รับอิทธิพลของทรีตเมนต์แล้ว

ตัวแปรร่วม (Concomitant Variable ; X) หมายถึง ค่าของหน่วยทดลองที่วัดได้ ก่อนที่หน่วยทดลองจะได้รับอิทธิพลของทรีตเมนต์ ซึ่งเป็นค่าของตัวแปรที่แฝงมากับหน่วยทดลอง

อำนาจการทดสอบ (Power of the test) คือค่าความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่าง ( $H_0$ ) เมื่อสมมติฐานว่างผิด

ความผิดพลาดประเภทที่ 1 (Type I error) คือความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐานว่าง ( $H_0$ ) เมื่อสมมติฐานว่างถูก

ความผิดพลาดประเภทที่ 2 (Type II error) คือความผิดพลาดที่เกิดจากการยอมรับสมมติฐานว่าง ( $H_0$ ) เมื่อสมมติฐานว่างผิด

สถิติพาราเมตริกซ์ (Parametric Statistics) เป็นสถิติอนุมานที่ว่าด้วยการทดสอบสมมติฐานและการประมาณค่า ซึ่งมีข้อตกลงเบื้องต้น (Assumption) บางอย่าง และเกี่ยวข้องกับค่าของพารามิเตอร์

สถิติอนพาราเมตริกซ์ (Nonparametric Statistics) เป็นสถิติอนุมานที่ไม่  
เกี่ยวข้องกับค่าพารามิเตอร์ของประชากร ไม่มีข้อตกลงเบื้องต้นเกี่ยวกับค่าของพารามิเตอร์

ความแกร่ง (Robustness) ของการทดสอบหมายถึง คุณสมบัติของการทดสอบที่  
ไม่ไวต่อการเปลี่ยนแปลงของปัจจัยอื่นที่ไม่ใช่ปัจจัยที่ต้องการทดสอบ เช่น การฝ่าฝืนข้อตกลง  
เบื้องต้นของการทดสอบสิ่งนั้น ซึ่งสิ่งที่ใช้พิจารณาความแกร่งคือ ความน่าจะเป็นของความผิด  
พลาดประเภทที่ 1

#### 1.7 ประโยชน์ของการวิจัย

ช่วยให้ผู้วิจัย เลือกใช้สถิติทดสอบในการวิเคราะห์ตัวแปรร่วมที่เหมาะสมและมีประสิทธิ-  
ภาพ เมื่อข้อมูลมีการแจกแจงความคลาดเคลื่อนไม่เป็นไปตามข้อตกลง

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย