

# หลักการวิเคราะห์ความแปรปรวนหลายตัวแปร

## Multivariate Analysis of Variance: MANOVA

รองศาสตราจารย์ ดร.ไพฑูริย์ สุขศรีงาม

คณะสิ่งแวดล้อมและทรัพยากรศาสตร์ มหาวิทยาลัยมหาสารคาม

### บทนำ

การวิเคราะห์ความแปรปรวนหลายตัวแปร (MANOVA) เป็นเทคนิคทางสถิติแบบหนึ่งที่ใช้เปรียบเทียบความแตกต่างค่าเฉลี่ยของตัวแปรตามจำนวนหลายตัวของกลุ่มตัวอย่าง หรือประชากรตั้งแต่ 2 กลุ่มขึ้นไป โดยการนำตัวแปรตามทั้งหมดมาจัดรวมกัน (combination) ทำให้เกิดการสร้างตัวแปรตามชุดใหม่ขึ้นมาจากชุดตัวแปรตามที่มีอยู่ก่อนแล้ว ก่อนการพัฒนาวิธีการของการวิเคราะห์แบบ MANOVA นักวิจัยนิยมใช้การเปรียบเทียบความแตกต่างค่าเฉลี่ยของตัวแปรตามทีละตัวแปร โดยใช้การวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA) จึงมีการทดสอบทางสถิติหลายครั้ง เท่ากับจำนวนตัวแปรตาม อย่างไรก็ตามการใช้ ANOVA ดังกล่าว ในการทดสอบตัวแปรตามหลายๆ ตัวแปร ทำให้เกิดปัญหาในการเพิ่มค่าความคลาดเคลื่อนในการตัดสินใจประเภทที่ 1 (type I error) หรือค่าแอลฟาจริง (actual alpha) (ในทางสถิติเรียกค่านี้ว่า overall type I error) ตามหลักการ Bonferroni Inequality [1],[2] มีค่ามากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (คือค่า  $\alpha$ )

การวิเคราะห์ความแปรปรวนหลายตัวแปร (MANOVA) มีข้อกำหนดหรือข้อจำกัดในการใช้ที่สำคัญหลายประการ ซึ่งนักวิจัยจำเป็นต้องตรวจสอบข้อมูลให้สอดคล้องกับข้อกำหนด ถ้าข้อมูลไม่สอดคล้องจะมีผลทำให้อำนาจการทดสอบ (power of the test) มีค่าลดลงหรือมีความน่าเชื่อถือได้น้อยกว่าการวิเคราะห์ความแปรปรวน (ANOVA) นอกจากนี้ถ้าผลการทดสอบทางสถิติมีนัยสำคัญ นักวิจัยต้องทำการทดสอบความแตกต่างค่าเฉลี่ยของตัวแปรตามทีละตัว เพื่อตรวจสอบว่ามีตัวแปรตามใดบ้างที่มีความแตกต่างจริงเกิดขึ้น อันเป็นผลมาจากตัวแปรอิสระ โดยจะต้องปรับค่านัยสำคัญทางสถิติ (ค่า  $\alpha$ ) เดิมที่ใช้ให้มีค่าเหมาะสมตามหลักการ Bonferroni Correction เสียก่อน

### ข้อจำกัดในการวิเคราะห์ความแปรปรวนหลายตัวแปร

การใช้การวิเคราะห์แบบหลายตัวแปร มีข้อจำกัดทั้งทางทฤษฎีและทางปฏิบัติจริง ดังนี้

#### 1. ข้อจำกัดเชิงทฤษฎี

ในการวิจัย นักวิจัยจะต้องเลือกตัวแปรตามที่มีความเหมาะสม สามารถวัดค่าเชิงปริมาณได้ อย่างน่าเชื่อถือ ตัวแปรตาม que เลือกมาจะต้องไม่มีความสัมพันธ์ซึ่งกันและกัน [1],[3] เนื่องจากค่าที่วัดได้ของตัวแปรตามแต่ละตัว เป็นค่าที่เกิดจากการวัดลักษณะเฉพาะที่เกิดจากอิทธิพลของตัวแปรอิสระ ถ้าตัวแปรตามหลายตัวแปรมีความสัมพันธ์กัน ค่าสังเกต (observation) ที่วัดได้จะเป็นค่าของลักษณะเดียวกัน หรือลักษณะที่คล้ายคลึงกัน แต่ได้มาจากการใช้วิธีการวัดที่แตกต่างกัน

การทดสอบโดยใช้ MANOVA ถ้าผลการทดสอบมีนัยสำคัญทางสถิติ ผู้วิจัยจะต้องทำการทดสอบตัวแปรตามทีละตัว เพื่อประเมินว่าตัวแปรตามใดที่มีส่วนร่วมทำให้ผลทดสอบมีความแตกต่าง ถ้าตัวแปรตามแต่ละตัวไม่มีความสัมพันธ์ซึ่งกันและกัน หรือมีความเป็นอิสระต่อกันก็จะทำให้ตัวแปรตามแต่ละตัวจะไม่มีส่วนร่วมซ้อนกัน (overlap) ในการรวมเชิงเส้นของแต่ละตัวแปรทั้งหมด การทดสอบตัวแปรตามแต่ละตัว (univariate test) จึงแสดงให้เห็นการมีส่วนร่วมอย่างบริสุทธิ์ของตัวแปรตาม que ส่งผลต่อการเกิดความแตกต่างของการรวมเชิงเส้น ดังกล่าว [3]

## 2. ข้อจำกัดในการปฏิบัติ

2.1 จำนวนสมาชิกในกลุ่มตัวอย่างและอำนาจการทดสอบในการวิเคราะห์ความแปรปรวนหลายตัวแปรจำนวนสมาชิกของกลุ่มตัวอย่างในแต่ละกลุ่มย่อย จะต้องมีความมากกว่าจำนวนตัวแปรตาม ถ้ามีจำนวนสมาชิก

ดังกล่าวน้อยกว่าจำนวนตัวแปรตาม จะไม่สามารถทดสอบข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ของวาเรียนซ์-โควาเรียนซ์เมทริกซ์ได้ (homogeneity of variance-covariance matrices) และถ้ามีจำนวนสมาชิกมากกว่าตัวแปรตาม จำนวน 1 หรือ 2 ตัว จะทำให้ผลการทดสอบข้อตกลงเบื้องต้นดังกล่าวเกิดการปฏิเสธ [1] เมื่อพิจารณาในเรื่องอำนาจการทดสอบ ซึ่งเป็นโอกาสของการตัดสินใจถูกต้องที่ปฏิเสธสมมุติฐานว่างที่เป็นเท็จ (false null hypothesis) ค่านี้จะมีค่าเพิ่มมากขึ้นเมื่อจำนวนสมาชิกในกลุ่มย่อย มากกว่าจำนวนตัวแปรตาม เนื่องจากมีการลดค่าระดับของความเสรีของความคลาดเคลื่อนนั่นเอง (degrees of freedom for error: df error)

2.2 ค่าสังเกตมีความเป็นอิสระ หรือไม่สัมพันธ์กัน (independent observations) การวิจัยทางสังคมศาสตร์ ศึกษาศาสตร์มักพบว่า ค่าสังเกตที่วัดได้จากสมาชิกแต่ละคนในกลุ่มตัวอย่างย่อย มีความสัมพันธ์กัน เช่น คะแนนผลสัมฤทธิ์ทางการเรียนแต่ละคนที่เรียนแบบกลุ่ม หรือเรียนเป็นคู่ จะได้รับอิทธิพลจากสมาชิกคนอื่น ๆ ในกลุ่มเสมอ เนื่องจากมีปฏิสัมพันธ์ระหว่างสมาชิกของกลุ่ม การมีความสัมพันธ์ของค่าสังเกตดังกล่าวมีผลทำให้ค่าแอลฟาจริง (actual alpha) ซึ่งหมายถึง รั้อยละของโอกาสปฏิเสธสมมุติฐานว่าง (null hypothesis) อย่างผิดพลาด มีค่ามากกว่าค่า แอลฟาปกติ (normal alpha) ซึ่งเป็นค่า  $\alpha$  ที่ผู้วิจัยกำหนดไว้ในการวิจัย หลายเท่า

ในการเลือกกลุ่มตัวอย่าง ถึงแม้ว่านักวิจัยใช้เทคนิคการชักตัวอย่างแบบสุ่ม (random sampling techniques) และเลือกสมาชิกแต่ละตัวเข้าสู่กลุ่มทดลอง หรือควบคุม (random assignment) ก็ไม่เป็นหลักประกันว่าจะไม่มีความสัมพันธ์ระหว่างค่าสังเกตของสมาชิก การตรวจสอบความสัมพันธ์ระหว่างค่าสังเกตดังกล่าว สามารถทำได้หลายวิธี เช่น การหาความสัมพันธ์ระหว่างกลุ่ม (interclass correlation) สำหรับการวิเคราะห์ความแปรปรวน ซึ่งมีสูตร ดังนี้ [2]

$$R = \frac{MS_b - MS_w}{MS_b + (n-1)(MS_w)}$$

เมื่อ  $MS_b$  หมายถึง ค่ากำลังสองเฉลี่ยระหว่างกลุ่ม (between group mean square)

$MS_w$  หมายถึง ค่ากำลังสองเฉลี่ยภายในกลุ่ม (within group mean square)

n หมายถึง ขนาดกลุ่มตัวอย่าง

นอกจากนี้ การทดสอบความเป็นอิสระของค่าสังเกต สามารถใช้การทดสอบ run test for randomness ก็ได้ [4]

2.3 การแจกแจงปกติ (normality) การแจกแจงของตัวแปรตาม ในแต่ละกลุ่มตัวอย่างย่อย ต้องมาจากค่าสังเกตของประชากรที่มีการแจกแจงเป็นเส้นโค้งปกติ ถ้าการแจกแจงดังกล่าวมีลักษณะแบบเบ้ (skewness) โด่ง (kurtosis) และแบน (platykurtosis) จะมีผลทำให้ค่า แอลฟาจริง (actual alpha) และอำนาจการทดสอบของการวิเคราะห์ความแปรปรวนมีค่าเปลี่ยนแปลงไปแต่ไม่มากนัก เนื่องจากสถิติทดสอบ F มีความแกร่ง (robustness) ถ้าใช้กลุ่มตัวอย่างขนาดเท่ากับ 20 [1] อย่างไรก็ตามใช้การวิเคราะห์ความแปรปรวนหลายตัวแปร จะเน้นการแจกแจงตัวแปรตามทั้งหมดรวมกัน เป็นเส้นโค้งปกติ (multivariate normality) ซึ่งโดยทั่วไปแล้ว หมายถึง การแจกแจงค่าเฉลี่ยของตัวแปรตามแต่ละตัวในกลุ่มย่อย และของการรวมเชิงเส้น (linear combination) ของตัวแปรตามทุกตัวรวมกันเป็นเส้นโค้งปกติ [1] การทดสอบแจกแจงปกติทีละตัวแปรตาม มิได้เป็นหลักประกันว่า การแจกแจงตัวแปรตามรวมกันจะเป็นเส้นโค้งปกติ เนื่องจากการเกิดการแจกแจงของตัวแปรตามรวมกันเป็นเส้นโค้งปกติมีคุณสมบัติ 3 ประการ คือ [2]

- 1) การแจกแจงของแต่ละตัวแปรตามต้องเป็นโค้งแบบปกติ
- 2) ผลรวมเชิงเส้น (linear combination) ของตัวแปรตาม มีการแจกแจงปกติ
- 3) ชุดย่อยทั้งหมดของชุดตัวแปรตามทั้งชุด มีการแจกแจงปกติ (multivariate normality)

การทดสอบการแจกแจงปกติของตัวแปรตามแต่ละตัวของกลุ่มตัวอย่างย่อย สามารถทำได้หลายวิธี เช่น ใช้การสร้างแผนภาพการกระจาย (scatter plot) และการคำนวณจากสูตร Kolmogorov – Smirnov test สำหรับกรณีหนึ่งตัวอย่าง [5] และ Shapiro-Wilk test [2] ส่วนการหา multivariate normality ในทางปฏิบัติยังไม่มียุทธวิธีการคำนวณที่เหมาะสม และยอมรับกันโดยทั่วไปจึงแนะนำให้ทดสอบการแจกแจงปกติสำหรับตัวแปรตามทีละตัว ในกรณีที่ข้อมูลมีการแจกแจงของตัวแปรตามพหุนาม ไม่เป็นเส้นโค้งปกติ อาจมีผลกระทบต่อค่าเพิ่มค่าแอลฟาจริง (actual alpha) แต่ไม่มากนัก เช่นเพิ่มขึ้นประมาณ .02 จากค่าระดับนัยสำคัญที่กำหนดไว้ (เช่น .01) อย่างไรก็ตาม กรณีที่แต่ละกลุ่มย่อยมีการแจกแจงตัวแปรตามเป็นแบบแบน (platykurtosis) หรือแบบโด่ง (kurtosis) จะมีผลในการลดอำนาจการทดสอบอย่างมาก แต่ถ้าวการแจกแจงของตัวแปรตามเป็นแบบเบ้ก็ยังมีผลต่อการลดอำนาจการทดสอบได้เหมือนกัน แต่ไม่มากนัก [6]

นักสถิติมีเชื่อว่าสถิติทดสอบ F (F-test) มีความแกร่งต่อการแจกแจงที่ไม่เป็นโค้งปกติของตัวแปรตามแต่ละตัวดังกล่าวมาแล้ว จึงอนุมานว่าสถิติทดสอบ F ก็มีความแกร่งต่อการแจกแจงตัวแปรหลายตัวแปรที่ไม่เป็นโค้งปกติ หรือเป็นแบบเบ้ด้วย ทั้งในเรื่องการเพิ่มแอลฟาจริง และค่าอำนาจการทดสอบ

2.4 ค่าสุดโต่ง (outliers) การวิเคราะห์ความแปรปรวนหลายตัวแปร มีหลักการเหมือนกับการวิเคราะห์ความแปรปรวนในเรื่องค่าสังเกตของตัวแปรตามจะต้องไม่มีค่าที่มากหรือน้อยผิดปกติจากค่าสังเกตอื่น ๆ เรียกค่าสังเกตนี้ว่า ค่าสุดโต่ง (outliers) ค่าสุดโต่งอาจเกิดจากการบันทึกค่าสังเกตผิดพลาด การทดลองที่บกพร่องหรือไม่ทราบสาเหตุชัดเจน [7] ถ้าข้อมูลของตัวแปรใดมีค่าสุดโต่ง จะส่งผลต่อการเพิ่มค่าของแอลฟาจริงและค่าอำนาจการทดสอบลดลง การทดสอบค่าสุดโต่งใช้ run test สำหรับกรณีตัวแปรตามทีละตัว [4] และ run test สำหรับค่าสุดโต่งของตัวแปรตามรวม (multivariate outliers) ของแต่ละกลุ่มย่อย [1] ถ้าข้อมูลมีค่าสังเกตตัวใดเป็นค่าสุดโต่งควรตัดทิ้งไป หรือใช้วิธีการแปลงข้อมูล (transformation of data) [7]

2.5 ความเป็นเอกพันธ์ของความแปรปรวนประชากร (homogeneity of variance) ในการวิเคราะห์ความแปรปรวนของตัวแปรตามแต่ละตัว ค่าสังเกตของกลุ่มตัวอย่างต้องมาจากประชากรที่มีความแปรปรวนเท่ากัน หรือเรียกว่า ความแปรปรวนเป็นเอกพันธ์ ซึ่งสถิติทดสอบ F (F-test) มีความแกร่งต่อการมีความแปรปรวนประชากรไม่เท่ากัน (เรียก ความแปรปรวนเป็น วิวิธพันธ์ : heterogeneity of variance) เมื่อแต่ละกลุ่มตัวอย่างหรือกลุ่มย่อยมีขนาดกลุ่มตัวอย่างเท่ากันหรือมีขนาดใกล้เคียงกัน โดยมีอัตราส่วนของกลุ่มตัวอย่างขนาดใหญ่ ต่อขนาดเล็กน้อยกว่า 1.5 [2] ถ้ากลุ่มตัวอย่างหรือกลุ่มตัวอย่างย่อย มีขนาดไม่เท่ากันมาก ๆ หรืออัตราส่วนดังกล่าวมากกว่า 1.5 และถ้ากลุ่มตัวอย่างขนาดเล็ก มีความแปรปรวนขนาดใหญ่จะมีผลทำให้โอกาสกระทำผิดพลาดในการตัดสินใจประเภทที่ 1 ( $\alpha$ -error) มากยิ่งขึ้นหรือแอลฟาจริงมีค่ามากกว่าค่าระดับนัยสำคัญทางสถิติ ( $\alpha$ ) มากขึ้น แต่ถ้ากลุ่มตัวอย่างขนาดใหญ่มีความแปรปรวนขนาดใหญ่ F-test จะมีความแกร่งที่เหมาะสมหรือมีความอนุรักษ์ (conservation) คือ มีค่าแอลฟาจริง น้อยกว่าค่าระดับนัยสำคัญทางสถิติ [2]

การทดสอบความเป็นเอกพันธ์ของความแปรปรวนประชากรมีสถิติทดสอบหลายตัว เช่น Bartlett's test [7] Cochran's test และ Harley test [9] แต่สถิติทดสอบเหล่านี้มีความไวต่อการแจกแจงไม่เป็นโค้งปกติ และ Levene's test [2] ซึ่งมีความแกร่งต่อการแจกแจงไม่เป็นโค้งปกติ

2.6 ความเป็นเอกพันธ์ของ วาเรียนซ์-โควาเรียนซ์ เมทริกซ์ (homogeneity of variance-covariance matrices) การวิเคราะห์ความแปรปรวนหลายตัวแปร ให้ความสำคัญต่อข้อตกลงของความเป็นเอกพันธ์ ของ วาเรียนซ์-โควาเรียนซ์ เมทริกซ์ มากโดยกำหนดว่าวาเรียนซ์-โควาเรียนซ์ เมทริกซ์ ภายในแต่ละ

กลุ่มย่อยถูกสุ่มมาจากวาเรียนซ์-โควาเรียนซ์ เมทริกซ์ ของประชากรเดียวกัน อย่างไรก็ตามสถิติทดสอบ F (F-test) มีความแข็งแกร่งต่อการมีข้อมูลไม่สอดคล้องกับหลักการดังกล่าว ถ้ากลุ่มตัวอย่างแต่ละกลุ่มมีขนาดเท่ากัน แต่ถ้ากลุ่มตัวอย่าง มีขนาดแตกต่างกันมาก และมีตัวแปรตามจำนวนมาก จะมีผลทำให้แอลฟาจริง มีค่าเพิ่มมากขึ้น และมากกว่า ค่านัยสำคัญของสถิติ ( $\alpha$ ) นอกจากนี้ถ้ากลุ่มตัวอย่างขนาดเล็กมี วาเรียนซ์-โควาเรียนซ์ ขนาดใหญ่ จะทำให้ค่าแอลฟาจริง เพิ่มมากยิ่งขึ้น จึงมีโอกาสปฏิเสธสมมติฐานว่าง ( $H_0$ ) อย่างผิดพลาดเพิ่มมากขึ้น ถ้ามีเหตุการณ์เช่นนี้ ควรใช้สถิติทดสอบ Pillai's trace แทนการใช้ Wilk's lambda [1]

ในการทดสอบข้อตกลงเบื้องต้นดังกล่าว นิยมใช้วิธีการ Box test หรือ (Box's M method) โดย Box ได้เขียนสูตรในการคำนวณไว้ทั้งในรูป  $\chi^2$  approximation และ F approximation อย่างไรก็ตามวิธีการของ Box test มีความไวต่อการแจกแจงที่ไม่เป็นโค้งปกติ จึงควรทดสอบการแจกแจงปกติของแต่ละตัวแปรตามในแต่ละกลุ่ม ก่อนการใช้ Box test ถ้าข้อมูลไม่มีการแจกแจงแบบโค้งปกติ ก็แปลงข้อมูลเสียก่อนแล้วจึงใช้ Box test ต่อไป

ในการเลือกใช้ Box test ในรูปของสถิติทดสอบ  $\chi^2$  หรือ F มีผู้เสนอว่าถ้ากลุ่มตัวอย่างมีขนาดเท่ากับ 20 และมีตัวแปรตามจำนวน 6 ตัว ควรใช้  $\chi^2$ -test แต่ถ้าไม่เป็นตามเงื่อนไขดังกล่าว ควรใช้ F-test [2] อย่างไรก็ตามโปรแกรมวิเคราะห์ทางสถิติส่วนใหญ่ใช้ F-test [10]

2.7 ความสัมพันธ์เชิงเส้น (linearity) ทั้งการวิเคราะห์ความแปรปรวนหลายตัวแปร (MANOVA) และ การวิเคราะห์ความแปรปรวนร่วมหลายตัวแปร (MANCOVA) เน้นความสำคัญของข้อตกลงนี้ โดยกำหนดว่าจะต้องมีความสัมพันธ์เชิงเส้น (linear relationship) ระหว่างทุกคู่ของความแปรปรวน ทุกคู่ของความแปรปรวนร่วม (covariates) และทุกคู่ของตัวแปรตาม ตัวแปรร่วม ในแต่ละกลุ่มย่อย ถ้าข้อมูลไม่สอดคล้องกับข้อตกลงนี้ จะมีผลทำให้อำนาจการทดสอบมีค่าลดลง เนื่องจาก (1) ผลรวมเชิงเส้น (linear combination) ของตัวแปรตามไม่สามารถแบ่งแยกกลุ่มสำหรับตัวแปรอิสระได้มากที่สุด และ (2) ตัวแปรร่วมไม่สามารถช่วยปรับค่าความคลาดเคลื่อนได้มากที่สุด [1] ถ้าพบว่าตัวแปรร่วมมีความสัมพันธ์ไม่เป็นเชิงเส้นควรตัดตัวแปรร่วมทิ้งไป ถ้าพบว่าตัวแปรตามมีความสัมพันธ์ไม่เป็นเชิงเส้น ควรใช้วิธีการแปลงข้อมูลเพื่อให้ค่าอำนาจการทดสอบมีค่าเพิ่มขึ้น การทดสอบข้อตกลงนี้ อาจใช้วิธีสร้างแผนภาพการกระจาย (scatter plot) หรือใช้คำนวณจากโปรแกรมคำนวณสถิติ [10]

2.8 ความเป็นเอกพันธ์ของความชันการถดถอย (homogeneity of regression slope) เฉพาะการวิเคราะห์ความแปรปรวนร่วมหลายตัวแปร (MANOVA) ให้มีความสำคัญกับข้อตกลงเบื้องต้นข้อนี้ เหมือนกับการวิเคราะห์ความแปรปรวนร่วม (ANCOVA) โดยกำหนดว่า เส้นการถดถอยระหว่างตัวแปรร่วมกับตัวแปรตาม ของแต่ละกลุ่มมีลักษณะเหมือนกัน หรือเส้นการถดถอยของแต่ละกลุ่มเรียงตัวขนานกัน หรือจะต้องไม่มีอิทธิพลร่วมหรือปฏิสัมพันธ์ (interaction) ระหว่างตัวแปรร่วมกับตัวแปรอิสระในแต่ละกลุ่มถ้ามีปฏิสัมพันธ์ระหว่างตัวแปรดังกล่าวเกิดขึ้น จะทำให้อำนาจการทดสอบมีค่าลดลงและการแปลความหมายของการวิจัยไม่สมเหตุสมผล

2.9 การไม่มีภาวะร่วมเส้นตรงพหุ (multicollinearity) และภาวะเอกฐาน (singularity) การวิเคราะห์ความแปรปรวนหลายตัวแปร และการวิเคราะห์ความแปรปรวนร่วมหลายตัวแปร กำหนดว่าจะต้องมีความสัมพันธ์ระหว่างตัวแปรตามอยู่ในระดับที่ไม่สูงนัก เช่น ไม่เกิน .80 [2] หรือไม่เกิน .90 [1] ถ้าตัวแปรมีความสัมพันธ์กันระดับสูงมาก ๆ เรียกว่า มีภาวะร่วมเส้นตรงพหุ (multicollinearity) หรือตัวแปรตามหนึ่งตัวแปรมีการรวมเกือบเป็นเชิงเส้น (near-linear combination) กับตัวแปรตามอื่น ๆ เรียกว่ามีภาวะเอกฐาน (singularity) ซึ่งทั้ง 2 กรณีดังกล่าว จะทำให้ตัวแปรตามตัวใดตัวหนึ่งมีลักษณะซ้ำซ้อนกับลักษณะของตัวแปรอื่น ๆ

นอกจากนี้ จะต้องมีความสัมพันธ์ระหว่างตัวแปรตามอยู่ในระดับไม่สูงนัก ถ้ามีความสัมพันธ์ดังกล่าวระดับสูง ควรตัดตัวแปรร่วมออกไป เนื่องจากไม่สามารถปรับค่าตัวแปรตามได้ดีกว่าตัวแปรร่วมอื่น ๆ ซึ่งการหา

ความสัมพันธ์ระหว่างตัวแปรร่วม ใช้วิธีการหา squared multiple correlation (SMC :  $R^2$ ) ถ้ามีค่าเกิน 0.50 แสดงว่ามีภาวะเอกฐาน ให้ตัดตัวแปรร่วมออกไป [2]

2.10 ความเชื่อถือได้ของตัวแปรร่วม (reliability of covariates) ในการวิจัยจำเป็นต้องมีเหตุผลที่น่าเชื่อถือได้ในการคัดเลือกตัวแปรร่วม เช่น ตัวแปรร่วมต้องมีความสัมพันธ์กับตัวแปรตาม ตัวแปรร่วมต้องมีความสัมพันธ์ต่อกันอยู่ในระดับต่ำ และสามารถวัดค่าเชิงปริมาณของตัวแปรร่วมอย่างแม่นยำเชื่อถือได้ ถ้าเลือกตัวแปรร่วมที่ไม่สมเหตุผล หรือไม่น่าเชื่อถือ จะส่งผลทำให้มีการเพิ่มโอกาสในการตัดสินใจผิดพลาด ประเภทที่ 1 ( $\alpha$ -error) และประเภทที่ 2 ( $\beta$ -error) เพิ่มมากขึ้น สำหรับการวิจัยที่ไม่ใช่แบบทดลอง (non experimental research) ความสัมพันธ์ระหว่างตัวแปรร่วมกับตัวแปรตาม ควรมีค่ามากกว่า 0.80 ขึ้นไป [1]

### การกำหนดค่านัยสำคัญสำหรับการทดสอบความแตกต่างของตัวแปรตามทีละตัว (Univariate Test)

หลังจากทดสอบความแตกต่างค่าเฉลี่ยของตัวแปรตามหลายตัวแปร โดยใช้การวิเคราะห์ความแปรปรวนหลายตัวแปร และพบว่าผลการทดสอบมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญทางสถิติค่าหนึ่งกำหนดไว้ (เช่น  $\alpha = .05$ ) นักวิจัยจะต้องทำการทดสอบความแตกต่างของตัวแปรตามทีละตัว (เรียก univariate test) เพื่อหาว่า ตัวแปรใด มีส่วนทำให้ผลการทดสอบดังกล่าวมีนัยสำคัญทางสถิติ โดยมากใช้สถิติทดสอบ (independent t-test) และสถิติทดสอบ F (F-test) สำหรับการวิเคราะห์ความแปรปรวน (ANOVA) [2] ซึ่งต้องทำการทดสอบหลายครั้งเท่ากับจำนวนตัวแปรตาม แต่เนื่องจากการทดสอบดังกล่าว แต่ละครั้งมีความเป็นอิสระต่อกัน ทำให้ค่า แอลฟารวม (overall alpha rate) มีค่าเพิ่มขึ้น ตามหลักการของ Bonferroni Inequality [1],[2] ซึ่งมีสูตรดังนี้

$$\text{Overall } \alpha \text{ rate} = 1 - (1 - \alpha)^k$$

โดย k หมายถึง จำนวนครั้งของการทดสอบ

$\alpha$  หมายถึง ค่าระดับนัยสำคัญทางสถิติ

ทั้งนี้ Bonferroni Inequality กล่าวไว้ว่า ระดับ แอลฟารวม (overall alpha level) ของชุดการทดสอบ 1 ชุด ที่ประกอบด้วย การทดสอบตัวแปรตามแต่ละตัว ที่เป็นอิสระต่อกัน มีค่าน้อยกว่า หรือเท่ากับ ผลรวมของระดับ แอลฟาของการทดสอบแต่ละครั้ง [2]

ดังนั้น ในการทดสอบตัวแปรตามทีละตัว (univariate test) จึงต้องนำจำนวนครั้งของการทดสอบไปหารค่าระดับแอลฟา หรือระดับนัยสำคัญทางสถิติที่ใช้ในการทดสอบ MANOVA ซึ่งจะได้  $\frac{\alpha}{k}$  และเป็นไปตาม Bonferroni Correction [1],[2],[11] เช่น ผู้วิจัยกำหนดค่าระดับนัยสำคัญทางสถิติของการทดสอบ MANOVA ไว้เท่ากับ .05 และมีตัวแปรตาม 2 ตัว ดังนั้นค่า แอลฟา สำหรับการทดสอบตัวแปรตามแต่ละตัวมีค่า  $\frac{.05}{2} = .025$  ซึ่งจะใช้ในการกำหนดนัยสำคัญทางสถิติต่อไป

### การเลือกสถิติทดสอบสำหรับการวิเคราะห์ความแปรปรวนหลายตัวแปร

การทดสอบ MANOVA มีสถิติทดสอบที่ใช้อยู่ 4 วิธี คือ Pillai's trace, Hotelling's trace, Wilk's lambda และ Roy's largest root ในการคำนวณโดยใช้โปรแกรมคำนวณสถิติ เช่น SPSS จะแสดงผลการคำนวณโดยใช้ทั้ง 4 วิธีทุกครั้ง โดยนำเสนอในรูปสถิติทดสอบ F (F-test) ถ้างานวิจัยมีปัจจัยหรือตัวแปรอิสระที่แบ่งออกเป็น 2 ระดับ (factor level) เช่น เพศ แบ่งเป็นเพศชาย กับ เพศหญิง ค่า F ที่คำนวณได้ของทุกวิธี จะมีค่าเท่ากัน เนื่องจากมีระดับปัจจัย 2 ระดับ ทำให้ระดับขั้นความเสรีของสิ่งทดลองหรือสถานการณ์มีค่าเท่ากับ 1 ซึ่งมีเพียงวิธีเดียวที่ใช้ในการสร้างผลรวมเชิงเส้น (linear combination) ของตัวแปรตาม เพื่อแบ่งระดับย่อยหรือกลุ่ม แต่ถ้า

งานวิจัยที่ประกอบด้วย 3 ระดับขึ้นไป จะมีวิธีการสร้างผลรวมเชิงเส้นของตัวแปรตามได้หลายวิธีเพื่อแบ่งกลุ่ม ค่า F ที่คำนวณได้ จะมีค่าแตกต่างกัน โดยค่า F จากวิธีหนึ่งอาจทำให้เกิดนัยสำคัญทางสถิติ แต่ค่า F จากอีกวิธีหนึ่งไม่ทำให้นัยสำคัญทางสถิติ ปัญหาจึงเกิดขึ้นกับผู้วิจัยว่าจะเลือกรายงานค่า F จากวิธีการของใคร

การที่ MANOVA มีสถิติทดสอบหลายวิธีการนั้น ก็เนื่องจากไม่มีวิธีการใดที่มีความเหมาะสมที่สุดในทุกสถานการณ์นั่นเอง ซึ่งต่างจากการวิเคราะห์ ANOVA ที่มีตัวแปรตามเพียงตัวเดียว ถ้าข้อมูลสอดคล้องกับข้อตกลงเบื้องต้น F-test ของ ANOVA เป็นวิธีการทางสถิติที่เหมาะสมที่สุด เนื่องจากมีอำนาจการทดสอบมากที่สุด และวิธีการทดสอบไม่มีความลำเอียง เมื่อนำวิธีการทั้ง 4 วิธีของ MANOVA ไปหาอำนาจการทดสอบในสถานการณ์ต่างๆ กัน พบว่า ไม่สามารถสรุปได้อย่างชัดเจนและเชื่อถือได้ว่าวิธีการของใครดีที่สุด

ในการเลือกใช้สถิติทดสอบใดมีความเหมาะสมมากกว่ากัน จำเป็นต้องมีความรู้ความเข้าใจเกี่ยวกับการวิเคราะห์เชิงจำแนก (discrimination analysis) และความแกร่ง (robustness) ของแต่ละวิธีการเมื่อข้อมูลไม่สอดคล้องกับข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนพหุนาม ดังนั้นในบทความนี้จะนำเสนอแนวทางไว้ดังต่อไปนี้ [1],[2]

#### 1. กรณีความแกร่ง (robustness)

1.1 กลุ่มตัวอย่างมีขนาดเล็กมีจำนวนไม่เท่ากัน และข้อมูลไม่สอดคล้องกับข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ของวาเรียนซ์-โควาเรียนซ์เมทริกซ์ (homogeneity of variance-covariance matrices) วิธีการของ Pillai's trace มีความแกร่งมากกว่าวิธีอื่นๆ โดยมีค่าของแอลฟาจริง (actual alpha) น้อยกว่าของวิธีอื่น

1.2 กลุ่มย่อยมีความแปรปรวนแตกต่างกัน และกลุ่มตัวอย่างมีขนาดเท่ากันหรือใกล้เคียงกัน (อัตราส่วนระหว่างกลุ่มตัวอย่างขนาดใหญ่ที่สุดต่อขนาดเล็กที่สุดเท่ากับหรือน้อยกว่า 1.5) วิธีการทั้ง 3 วิธีคือ Pillai's trace Hotelling's trace และ Wilk's lambda มีความแกร่งเท่าๆกัน

1.3 ข้อมูลมีการแจกแจงไม่เป็นเส้นโค้งปกติ (multivariate normality) วิธีการทั้ง 4 วิธีมีความแกร่งเท่า ๆ กัน

1.4 ข้อมูลไม่สอดคล้องกับข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ของวาเรียนซ์-โควาเรียนซ์เมทริกซ์ วิธีการของ Roy's largest root มีความแกร่งน้อยกว่าทั้ง 3 วิธี

1.5 กลุ่มย่อยมีขนาดไม่เท่ากัน ข้อมูลสอดคล้องกับข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ของวาเรียนซ์-โควาเรียนซ์เมทริกซ์ และมีความสอดคล้องกับการแจกแจงแบบโค้งปกติ วิธีการของ Pillai's trace มีความแม่นยำหรือมีความแกร่งมากที่สุด

#### 2. กรณีอำนาจการทดสอบ (power of the test)

2.1 ข้อมูลสอดคล้องกับข้อตกลงเบื้องต้น วิธีในทั้ง 4 วิธีมี ค่าว่าการทดสอบแตกต่างกันเพียงเล็กน้อย (<0.06) (Olson.1974)

2.2 มีความแตกต่างระหว่างกลุ่ม เนื่องจากการใช้ discriminant function ที่ 1 หรือใช้ตัวแปรที่ 1 ข้อมูลมีความเป็นเอกพันธ์ของวาเรียนซ์-โควาเรียนซ์ เมทริกซ์ วิธีการของ Roy's largest root มีค่าอำนาจการทดสอบมากที่สุด รองลงมาได้แก่ Hotelling trace Wilk's lambda และ Pillai's trace ตามลำดับ

2.3 มีความแตกต่างระหว่างกลุ่ม เนื่องการใช้ discriminant function 2 ค่า หรือมากกว่าหรือกลุ่มมีความแตกต่างกับความตัวแปรหลายตัว วิธีการของ Pillai's trace มีค่าอำนาจการทดสอบมากที่สุด

นอกจากนี้ Todman และ Dugard [12] ได้เสนอแนะว่าถ้าค่าไอเกน (eigen value) ของ Hotelling trace และ Pillai's trace มีค่าเท่ากันหรือเกือบเท่ากัน ผลการทดสอบสมมุติฐานจะไม่มีความสำคัญทางสถิติ และถ้าค่าของ Hotelling's trace และ Roy's largest root มีค่าเท่ากันหรือเกือบเท่ากัน ผลการทดสอบสมมุติฐานจะไม่มีความสำคัญทางสถิติ เนื่องจากมีตัวแปรตามหนึ่งตัวที่มีอิทธิพลมากที่สุด หรือมีความสัมพันธ์ระหว่างตัวแปรตามค่อนข้างสูงมาก

อย่างไรก็ตามหนังสือสถิติศาสตร์และเอกสารที่เกี่ยวข้องกับวิธีการใช้โปรแกรมการคำนวณสถิติส่วนมากแนะนำให้ใช้วิธีการของ Wilk's lambda

### บทสรุป

การวิเคราะห์ความแปรปรวนพหุนาม (MANOVA) เป็นเทคนิคทางสถิติขั้นสูง ที่ใช้การทดสอบความแตกต่างค่าเฉลี่ย ของตัวแปรตามหลายตัวแปร อันเนื่องมาจากอิทธิพลของตัวแปรอิสระ 1 ตัวแปรหรือหลายตัวแปร วิธีการดังกล่าว มีความเชื่อถือได้สูงมาก ถ้านักวิจัยได้ปฏิบัติตามเงื่อนไขหรือข้อตกลงเบื้องต้นของการใช้

### เอกสารอ้างอิง

- [1] Tabachnick,B.G. and Fidell, L.S. (2007). **Using Multivariate Statistics.5<sup>th</sup>** edition New York: Pearson Education,Inc.
- [2] Stevens,J.R. (2009). **Applied Multivariate Statistics for Social Sciences.** New York: Routledge Talyer Francis Company.
- [3] Dancy,C.P. and Reidy, J. (1999). **Statistics without Maths for Psychology:Using SPSS for Windows.** New York: Prentice-Hall.
- [4] Andersen, R.L. (1997). **Practical Statistics for Analytical Chemistry,** New York: Van Nostrand Reinhold Company.
- [5] Freund,R.J.,Wilson, W.J. and Mohr,D.L.(2010). **Statistical Methods.3<sup>rd</sup>** editon, London: Academic Press.
- [6] Olson,C.L. (1974). Comparative Robustness of Six Tests in Multivariate Analysis of Variance. **Journal of the American Statistical Association,** 69, 894-908.
- [7] Milton,J.s.(1992).**Statistical Methods in the Biological and Health Sciencs.2<sup>nd</sup>** edition New York: Hill Book company.
- [8] Dixon,W.J and Messey,F.J. (1993). **Introduction to Statistical Analysis.** London: McGraw-Hill Book company.
- [9] Ott,R.L.and Longnecker, M. (2001). **An Introduction to Statistical Methods and Data Analysis. 5<sup>th</sup>** edition, Singapore: Duxbury Thompson Learning.
- [10] Brace,N., Kemp, R. and Snelgar R. (2006). **SPSS for Psychologists: A Guide to Data Analysis Using SPSS for Windows. 3<sup>rd</sup>** editon, London: Laerence Erlbrum Associates Puvlishers.
- [11] Johnson,D.E. (1998). **Applied Multivariate Methods for Data Analysis,** New York: Duxbury Press.
- [12] Todman, J. and Dugard, P. (2007). **Approaching Multivariate Aratysis:An Introduction for Psychotugy.** New York: Paychotugy Press.