การวิเคราะห์ความแปรปรวนแบบหลายองค์ประกอบ
(Factorial Analysis of Variance)

ฉัตรศิริ ปิยะพิมลสิทธิ์

จากที่ได้กล่าวไปแล้วเกี่ยวกับโมเดลการวิเคราะห์ความแปรปรวนแบบองค์ประกอบเดียว (One-way ANOVA) และวิธีการเปรียบเทียบพหุคูณ (MCPs) สำหรับโมเดล ในเอกสารนี้จะต่อเนื่องกัน คือโมเดลการวิเคราะห์ความแปรปรวนโดยขยายจากกรณีมีองค์ประกอบเดียวไปเป็นกรณีมี 2 และ 3 องค์ประกอบ นั่นคือในเอกสารนี้จะตอบคำถามที่ว่า "ถ้ามีหลายองค์ประกอบที่ต้องการเปรียบเทียบค่าเฉลี่ย ควรจะทำอย่างไร" ในอีกกรณีหนึ่ง ผู้วิจัยสนใจในอิทธิพลของตัวแปรอิสระที่มี 3 ตัวหรือมากกว่าที่มีต่อตัวแปรตาม ทำนองเดียวกับการวิเคราะห์การถดถอย เราจะมีการเปลี่ยนแปลงจากตัวแปรตัวเดียวในการวิเคราะห์ความแปรปรวนแบบหนึ่งองค์ประกอบไปเป็นกรณีมีตัวแปรอิสระหล่ายตัวในการวิเคราะห์ความแปรปรวนแบบหลายองค์ประกอบ ซึ่งในเนื้อหานี้จะกล่าวถึงโมเดลการวิเคราะห์ความแปรปรวนกรณีมี 2 องค์ประกอบ

ตัวอย่าง สมมติว่านักวิจัยสนใจในการเลือกหนังสือและช่วงเวลาเข้าเรียนที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนสถิติ ดังนั้นตัวแปรอิสระก็คือการเลือกหนังสือสำหรับเรียนในชั้นและตัวแปรอิสระตัวที่สองคือช่วงเวลาเรียนที่เข้าเรียน สมมติฐานของนักวิจัยศึกษาอิทธิพลที่มีต่อผลสัมฤทธิ์ทางการเรียน ตัวอย่างการศึกษาบอกได้ว่านักวิจัยสนใจในการเปรียบเทียบการเลือกหนังสือ 3 เล่ม (A, B และ C) และ 3 ช่วงเวลา (ช่วงเช้า, กลางวัน และเย็น) นักเรียนควรจะถูกสุ่มเข้าในแต่ละกลุ่มด้วยหลักทางสถิติเข้ากลุ่มของการเลือกหนังสือและช่วงเวลาเข้าเรียน กลุ่มหนึ่งนักเรียนอาจจะถูกให้เรียนในช่วงเย็นและใช้หนังสือ A

แนวคิดการวิเคราะห์จะทำนองเดียวกับการวิเคราะห์ความแปรปรวนแบบหนึ่งองค์ประกอบ นอกจากนี้มีแนวคิดใหม่ คือ อิทธิพลหลัก (main effects) ผลของปฏิสัมพันธ์ (interaction effects) วิธีการเปรียบเทียบพหุคูณ สำหรับอิทธิพลหลักและผลปฏิบัติสัมพันธ์ และการออกแบบที่ไม่เป็นอิสระ (nonorthogonal designs)

โมเดลการวิเคราะห์ความแปรปรวนสององค์ประกอบ (The Two-Factor ANOVA Model)

ในหัวข้อนี้จะอธิบายคุณลักษณะของโมเดล ANOVA 2 องค์ประกอบ โครงร่างของข้อมูล โมเดลเชิงเส้น อิทธิพลหลัก และปฏิสัมพันธ์ ข้อตกลงเบื้องต้นของโมเดลและการละเมิด การคำนวณผลรวมกำลังสอง ตารางสรุป ANOVA ค่าเฉลี่ยกำลังสองคาดหวัง และวิธีการเปรียบเทียบพหุคูณ

คุณลักษณะของโมเดล (Characteristics of the Model)

คุณลักษณะแรกของโมเดล ANOVA 2 องค์ประกอบควรมีความแจ่มชัด ความเข้าใจในการพิจารณาอิทธิพลของ 2 องค์ประกอบหรือตัวแปรอิสระ 2 ตัวบนตัวแปรตาม โดยแต่ละองค์ประกอบจะมีระดับตั้งแต่ 2 ระดับขึ้นไป ซึ่งจะเรียกว่า factorial design เพราะมีมากกว่า 1 องค์ประกอบในโมเดล เราจะเห็นว่า ANOVA 2 องค์ประกอบจะขยายเพิ่มจาก ANOVA 1 องค์ประกอบ ทำไมผู้วิจัยต้องคิดซับซ้อนขึ้นมาเป็นองค์ประกอบที่สอง มี 3 เหตุผลในขณะนี้ คือ เหตุผลแรก ผู้วิจัยอาจจะมีความสนใจคิดที่จะศึกษาองค์ประกอบที่สอง มากกว่าจะศึกษาแต่ละองค์ประกอบแยกเป็นอิสระจากกัน ผู้วิจัยรวมทั้งสององค์ประกอบอยู่ในการศึกษาเดียว ซึ่งในแบบทดสอบไม่ได้มีเฉพาะอิทธิพลของแต่ละองค์ประกอบเท่านั้นแต่มีอิทธิพลขององค์ประกอบทั้งสองที่เก็บรวบรวมไปพร้อมกัน นอกจากนี้องค์ประกอบที่สองจะส่งผลอย่างอิสระจากอีกองค์ประกอบหนึ่ง (ไม่มีปฏิสัมพันธ์กัน) หรือองค์ประกอบทั้งสองมีการส่งผลร่วมกันมในบางอิทธิพลที่เพิ่มขึ้นมา (มีปฏิสัมพันธ์) ถ้ามีการศึกษา 2 ชิ้นที่แยกกัน แต่ละชิ้นศึกษาแยกตัวแปรอิสระแต่ละตัว ไม่มีสารสนเทศที่จะบ่งบอกถึงผลของปฏิสัมพันธ์

เหตุผลที่สองสำหรับการรวมองค์ประกอบเข้าไปเป็นความพยายามในการลดความคลาดเคลื่อน (หรือภายในกลุ่ม) ซึ่งความแปรปรวนที่ไม่สามารถอธิบายได้ด้วยองค์ประกอบเดียว นั่นคือเหมือนกับการตัดสินใจใช้การวิเคราะห์การถดถอยพหุคูณที่รวมตัวแปรอิสระหลาย ๆ ตัวและวิเคราะห์ด้วยกัน (เป็นการลดความคลาดเคลื่อนของความแปรปรวนซึ่งจะช่วยเพิ่มค่า R^2) ในการใช้องค์ประกอบที่ 2 มีการกำหนดให้มีการประมาณค่าความคลาดเคลื่อนที่ถูกต้องแน่นอนสำหรับเหตุผลนี้ การออกแบบที่มี 2 องค์ประกอบ โดยทั่วไปจะมีอำนาจมากกว่าการออกแบบองค์ประกอบเดียว 2 ครั้ง ซึ่งองค์ประกอบที่สองจะช่วยควบคุมการเพิ่มขึ้นของความแปรปรวนที่เกิดจากภายนอก

เหตุผลที่สามสำหรับการพิจารณา 2 องค์ประกอบคือความสามารถในการสรุปอ้างอิงผลและมีประสิทธิภาพและมีความประหยัดในการใช้แหล่งทรัพยากร ดังนั้นผลที่ได้สามารถสรุปอ้างอิงไปยังกรณีอื่น ๆ แต่การศึกษาจะมีต้นทุนมากและในเทอมของเวลาและงบประมาณ

นอกจากนี้ ใน ANOVA 2 องค์ประกอบ ทุกระดับขององค์ประกอบแรก (หรือเรียกว่า factor A) จะจับคู่กับทุกระดับขององค์ประกอบสอง (หรือเรียกว่า factor B) ในอีกกรณีหนึ่ง ทุก ๆ กลุ่มขององค์ประกอบ A และ B จะรวมอยู่ในโมเดลการศึกษาที่ถูกอ้างอิงเรียกว่า fully crossed design ถ้ามีบางกลุ่มไม่ถูกรวมแล้ว การออกแบบจะ non fully crossed และอาจจะเรียกว่าเป็นแบบ nested design ในแต่ละค่าสังเกตจะถูกสุ่มมาให้ได้รับเพียงระดับเดียวของทั้งสององค์ประกอบ และแต่ละค่าสังเกตจะสอบวัดเพียงครั้งเดียว ถ้าแต่ละค่าสังเกตสอบวัดมากกว่า 1 ครั้งแล้วควรจะต้องเรียกว่า repeated measures design ในเอกสารนี้อธิบายภายใต้องค์ประกอบที่เป็นแบบอิทธิพลกำหนด (fixed-effects factors) ดังนั้นรูปแบบทั้งหมดจะเป็นโมเดลแบบอิทธิพลกำหนด (fixed-effects model) ถ้ามี 1 องค์ประกอบ หรือมากกว่าที่เป็น

อิทธิพลสุ่มแล้วการออกแบบจะไม่ใช่รูปแบบอิทธิพลกำหนด ซึ่งเราจะอธิบายในโอกาสต่อไป สำหรับในเอกสารนี้จะสอดคล้องกับตัวแปรตามที่ถูกวัดอย่างน้อยในระดับช่วง (interval level)

ในเอกสารนี้จะสมมติว่าจำนวนของค่าสังเกตในแต่ละกลุ่มจะเท่ากัน ซึ่งจะเป็นการออกแบบอิสระ (orthogonal design) ซึ่งอิทธิพลเนื่องจากองค์ประกอบอิสระ เราจะอธิบายในกรณีที่ n แต่ละกลุ่มไม่เท่ากัน นอกจากนี้จะสมมติว่ามีอย่างน้อย 2 ค่าสังเกตต่อกลุ่มและมีเทอมความคลาดเคลื่อนอันเนื่องมาจากความแปรปรวนภายในกลุ่มกับการทดสอบปฏิสัมพันธ์

ในการสรุป คุณลักษณะของโมเดลการวิเคราะห์องค์ประกอบแบบสององค์ประกอบแบบอิทธิพลกำหนด มีดังนี้ 1) มีตัวแปรอิสระ 2 ตัวที่แต่ละตัวจะวัดตั้งแต่ 2 ระดับขึ้นไป 2) ระดับของตัวแปรอิสระทั้งสองตัวจะถูกกำหนดโดยผู้วิจัย 3) กลุ่มตัวอย่างถูกสุ่มเข้าในแต่ละกลุ่มทดลองเพียงกลุ่มเดียว 4) ทั้งสององค์ประกอบจะแบบแบบ fully crossed และ 5) ตัวแปรตามจะถูกวัดอย่างน้อยในระดับช่วง ในเนื้อหานี้การออกแบบการทดลอง การวิเคราะห์ความแปรปรวนแบบสององค์ประกอบจะถูกอ้างอิงว่าเป็นรูปแบบสุ่มสมบูรณ์ (the conpletely randomized factorial design)

โครงร่างของข้อมูล (The Layout of the Data)

ก่อนจะกล่าวถึงทฤษฎีและลำดับของการวิเคราะห์ข้อมูล ต้องเข้าใจถึงโครงร่างของข้อมูล โดยจะให้ค่าสังเกตแต่ละค่าแทน Y_{ijk} เมื่อ j คือระดับขององค์ประกอบ A ส่วน k คือระดับขององค์ประกอบ B และ i คือลำดับที่ของค่าสังเกตที่อยู่ภายในระดับขององค์ประกอบ A และระดับขององค์ประกอบ B เช่น Y_{321} หมายถึงเป็นค่าสังเกตตัวที่ 3 ที่อยู่ระดับที่สองขององค์ประกอบ A และระดับที่หนึ่งขององค์ประกอบ B โดยตัวห้อยตัวแรกจะมีพิสัยจาก i = 1, …, n ตัวห้อยที่สองมีพิสัยจาก j = 1, …, J และตัวห้อยที่สามมีพิสัยจาก k = 1, …, K สังเกตว่าตัวอักษรที่เห็นตัวห้อยสองตัวจะบ่งบอกถึงเซลของค่าสังเกต จะได้ว่า Y_{321} เป็นค่าสังเกตค่าที่ 3 ในเซล 21 ดังนั้น J ระดับขององค์ประกอบ A, K ระดับขององค์ประกอบ B และกลุ่มตัวอย่าง n ในแต่ละเซล สำหรับกลุ่มตัวอย่างรวมคือ JKn = N เราจะสมมติว่ากลุ่มตัวอย่าง n ในแต่ละเซลซึ่งหมายถึงจำนวน n ในแต่ละกลุ่มเท่ากัน

โครงร่างของข้อมูลจากกลุ่มตัวอย่างแสดงในตาราง 1 เราจะเห็นว่าแต่ละแถวจะแสดงค่าสังเกตสำหรับแต่ละระดับขององค์ประกอบ A และแต่ละสดมภ์จะแสดงค่าสังเกตในแต่ละระดับขององค์ประกอบ B ในช่วงล่างของแต่ละสดมภ์จะเป็นค่าเฉลี่ยรวมสำหรับแต่ละสดมภ์ ($\bar{Y}_{..k}$) ด้านขวาของแต่ละแถวเป็นค่าเฉลี่ยรวมสำหรับแต่ละแถว ($\bar{Y}_{.j.}$) และด้านล่างขวามือตรงมุมจะเป็นค่าเฉลี่ยรวมทั้งหมด ($\bar{Y}_{...}$) ซึ่งจำเป็นต้องคำนวณค่าเฉลี่ยในแต่ละเซล ($\bar{Y}_{.jk}$) ซึ่งจะแสดงในด้านล่างของแต่ละเซลล์ โดยสรุปโครงร่างของข้อมูลจะมีรูปแบบที่ผู้วิจัยจะต้องแทนที่ข้อมูลลงไปในตารางก่อนเริ่มต้นวิเคราะห์

ตาราง 1 โครงร่างข้อมูลวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ

	ระดับขององค์ประกอบ B				Row Mean
	1	2	...	K	
1	Y_{111}	Y_{112}	...	Y_{11K}	$\bar{Y}_{.1.}$
	Y_{211}	Y_{212}	...	Y_{21K}	
	:	:	:	:	
	Y_{n11}	Y_{n12}	...	Y_{n1K}	
	$\bar{Y}_{.11}$	$\bar{Y}_{.12}$	...	$\bar{Y}_{.1K}$	
2	Y_{121}	Y_{122}	...	Y_{12K}	$\bar{Y}_{.2.}$
	Y_{221}	Y_{222}	...	Y_{22K}	
	:	:	:	:	
	Y_{n21}	Y_{n22}	...	Y_{n2K}	
	$\bar{Y}_{.21}$	$\bar{Y}_{.22}$	...	$\bar{Y}_{.2K}$	
ระดับขององค์ประกอบ A	:	:	:	:	:
J	Y_{1J1}	Y_{1J2}	...	Y_{1JK}	$\bar{Y}_{.J.}$
	Y_{2J1}	Y_{2J2}	...	Y_{2JK}	
	:	:	:	:	
	Y_{nJ1}	Y_{nJ2}	...	Y_{nJK}	
	$\bar{Y}_{.J1}$	$\bar{Y}_{.J2}$	...	$\bar{Y}_{.JK}$	
Column Mean	$\bar{Y}_{..1}$	$\bar{Y}_{..2}$	...	$\bar{Y}_{..K}$	$\bar{Y}_{...}$

โมเดลการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ

ในหัวข้อนี้จะแนะนำโมเดลการวิเคราะห์ความแปรปรวนเชิงเส้นตรง และการประมาณค่าพารามิเตอร์ในโมเดล โมเดลการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบเป็นโมเดลเชิงเส้นตรงคล้ายกับโมเดลการถดถอยอย่างง่ายและถดถอยพหุคูณ และโมเดลการวิเคราะห์องค์ประกอบแบบหนึ่งองค์ประกอบ โดยโมเดลนี้จะเป็นโมเดลอิทธิพลกำหนด (fixed-effects) ที่เขียนอยู่ในเทอมพารามิเตอร์ของประชากรได้ดังนี้

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk}$$

เมื่อ Y_{ijk} คือคะแนนสังเกตของตัวแปรตามสำหรับกลุ่มตัวอย่างที่ i ในระดับที่ j ขององค์ประกอบ A และระดับที่ k ขององค์ประกอบ B (หรือในเซลล์ jk), μ คือค่าเฉลี่ยรวมทั้งหมด

α_j คือ อิทธิพลของระดับ j ในองค์ประกอบ A , β_k คืออิทธิพลของระดับ k ในองค์ประกอบ B, $(\alpha\beta)_{jk}$ คือผลของปฏิสัมพันธ์สำหรับการรวมกันของระดับ j ในองค์ประกอบ A และระดับ k ในองค์ประกอบ B และ ε_{ijk} คือความคลาดเคลื่อนอย่างสุ่มของบุคคลที่ i ในเซล jk ความคลาดเคลื่อนได้มาจากความแตกต่างของแต่ละบุคคล ความคลาดเคลื่อนในการวัด และ/หรือองค์ประกอบอื่น ๆ หรืออิทธิพลอื่น ๆ ที่ไม่ได้อยู่ในการศึกษา อิทธิพลของประชากรและความคลาดเคลื่อนสามารถคำนวณได้ดังนี้

$$\alpha_j = \mu_{.j.} - \mu$$

$$\beta_k = \mu_{..k} - \mu$$

$$(\alpha\beta)_{jk} = \mu_{.jk} - (\mu_{.j.} + \mu_{..k} - \mu)$$

และ
$$\varepsilon_{ijk} = Y_{ijk} - \mu_{.jk}$$

นั่นคือ อิทธิพลแนวแถวจะเท่ากับความแตกต่างระหว่างค่าเฉลี่ยในระดับ j ขององค์ประกอบ A และค่าเฉลี่ยรวมทั้งหมด อิทธิพลแนวสดมภ์จะเท่ากับความแตกต่างระหว่างค่าเฉลี่ยในระดับ k ขององค์ประกอบ B และค่าเฉลี่ยรวมทั้งหมด ผลของปฏิสัมพันธ์เป็นอิทธิพลของการรวมกันในระดับขององค์ประกอบ A และ B และความคลาดเคลื่อนเท่ากับความแตกต่างระหว่างค่าสังเกตแต่ละค่ากับค่าเฉลี่ยของเซลล์ jk ความคลาดเคลื่อนในการวิเคราะห์ความแปรปรวนจะคล้ายกับความคลาดเคลื่อนในการวิเคราะห์การถดถอย ซึ่งทั้งสองกรณีจะนำเสนอความแปรปรวนของตัวแปรตามที่ไม่สามารถอธิบายได้ด้วยตัวแปรอิสระ

ทำไมดูเหมือนว่าผลของปฏิสัมพันธ์จะมีความแตกต่างน้อยกว่าอีก 2 อิทธิพล สมการของผลปฏิสัมพันธ์อาจจะเขียนใหม่ได้ดังนี้

$$(\alpha\beta)_{jk} = \mu_{.jk} - \alpha_j - \beta_k - \mu$$

ซึ่งการเขียนจะเขียนแสดงอิทธิพลแนวแถวและสดมภ์ ทีนี้เราจะเห็นว่าผลของปฏิสัมพันธ์จะเท่ากับ ค่าเฉลี่ยของเซลล์ ลบกับ 1) อิทธิพลแนวแถว 2) อิทธิพลแนวสดมภ์ และ 3) ค่าเฉลี่ยรวมทั้งหมด ในอีกกรณีหนึ่ง ปฏิสัมพันธ์จะเขียนเป็นฟังชั้นของค่าเฉลี่ยในเซลล์โดยปราศจากอิทธิพลของแถวและสดมภ์และค่าเฉลี่ยรวมทั้งหมด โดยการเปลี่ยนกลับเป็นอย่างเดิม

$$\begin{aligned}(\alpha\beta)_{jk} &= \mu_{.jk} - \alpha_j - \beta_k - \mu \\ &= \mu_{.jk} - (\mu_{.j.} - \mu) - (\mu_{..k} - \mu) - \mu \\ &= \mu_{.jk} - (\mu_{.j.} + \mu_{..k} - \mu)\end{aligned}$$

เงื่อนไขพิเศษของโมเดล สำหรับโมเดลที่มี n เท่ากันมีเงื่อนไขคือ

$$\sum_j \alpha_j = 0$$

$$\sum_k \beta_k = 0$$

$$\sum_j (\alpha\beta)_{jk} = 0$$

และ $\sum_{k} (\alpha\beta)_{jk} = 0$

ดังนั้น ผลรวมของอิทธิพลแนวแถวเท่ากับ 0 ผลรวมของอิทธิพลแนวสดมภ์เท่ากับ 0 และผลรวมของผลปฏิสัมพันธ์เท่ากับ 0 ทั้งคู่ของแนวแถวและสดมภ์ ในการนำไปใช้ ตัวอย่าง ถ้าอิทธิพลของแนวแถวไม่เป็น 0 แล้วอิทธิพลของแนวแถวจะสมดุลอยู่รอบ ๆ 0 ซึ่งมีบางค่าเป็นบวกและบางค่าเป็น 0

การประมาณค่าพารามิเตอร์ของโมเดล μ, α_j, β_k, $(\alpha\beta)_{jk}$ และ ε_{ijk} วิธีกำลังสองต่ำสุดจะถูกใช้ในการประมาณค่า วิธีกำลังสองต่ำสุดโดยทั่วไปจะเหมาะสมสำหรับโมเดลเชิงเส้นตรง (การถดถอย ANOVA และสถิติอื่น ๆ) การประมาณค่าของกลุ่มตัวอย่างจะแสดงด้วย $\bar{Y}_{...}$, a_j, b_k, $(ab)_{jk}$ และ e_{ijk} ตามลำดับ ทั้งสี่ค่าสามารถคำนวณได้ดังนี้

$$a_j = \bar{Y}_{.j.} - \bar{Y}$$

$$b_k = \bar{Y}_{..k} - \bar{Y}$$

$$(ab)_{jk} = \bar{Y}_{.jk} - (\bar{Y}_{.j.} + \bar{Y}_{..k} - \bar{Y})$$

และ

$$e_{ijk} = Y_{ijk} - \bar{Y}_{.jk}$$

สังเกตเกตว่า $\bar{Y}$ จะแสดงค่าเฉลี่ยรวมทั้งหมดของกลุ่มตัวอย่าง, $\bar{Y}_{.j.}$ แสดงค่าเฉลี่ยของกลุ่มตัวอย่างสำหรับระดับที่ j ขององค์ประกอบ A, $\bar{Y}_{..k}$ แสดงค่าเฉลี่ยของกลุ่มตัวอย่างสำหรับระดับที่ k ขององค์ประกอบ B และ $\bar{Y}_{.jk}$ แสดงค่าเฉลี่ยของกลุ่มตัวอย่างสำหรับเซลล์ jk เมื่อจุดที่ห้อยอยู่แสดงให้เห็นถึงค่าเฉลี่ยของทุกกลุ่มทั้ง i, j และ k

ต่อมาพิจารณาสมมติฐานสำหรับการถดถอยนัยสำคัญ ในกรณีของการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ จะมีสมมติฐานถึง 3 ชุดสำหรับในแต่ละอิทธิพลหลักและผลของปฏิสัมพันธ์ สมมติฐานศูนย์และสมมติฐานทางเลือก สำหรับการทดสอบอิทธิพลขององค์ประกอบ A คือ

$H_{01} : \mu_{.1.} = \mu_{.2.} = \ldots = \mu_{.J.}$

H_{11} : ไม่ใช่ทุก $\mu_{.J.}$ ที่จะเท่ากัน

สมมติฐานสำหรับทดสอบอิทธิพลขององค์ประกอบ B คือ

$H_{02} : \mu_{..1} = \mu_{..2} = \ldots = \mu_{..K}$

H_{12} : ไม่ใช่ทุก $\mu_{..K}$ ที่จะเท่ากัน

สุดท้าย สมมติฐานสำหรับการทดสอบผลปฏิสัมพันธ์คือ

$H_{03} : (\mu_{.jk} - \mu_{.j.} - \mu_{..k} + \mu) = 0$ สำหรับทุก j และ k

H_{13} : ไม่ใช่ทุก $(\mu_{.jk} - \mu_{.j.} - \mu_{..k} + \mu)$ ที่จะเท่ากัน

สมมติฐานศูนย์สามารถเขียนในเทอมของอิทธิพลของแนวแถว อิทธิพลของแนวสดมภ์ และผลปฏิสัมพันธ์ คือ

$H_{01} : \alpha_1 = \alpha_2 = \ldots = \alpha_J = 0$

$H_{02} : \beta_1 = \beta_2 = \ldots = \beta_K = 0$

$H_{03} : (\alpha\beta)_{jk} = 0$ สำหรับทุก j และ k

สมมติฐานศูนย์อาจจะมีการเขียนวิธีอื่นโดยเฉพาะการเขียนแบบไม่มีทิศทางที่เหมาะสม ถ้ามีสมมติฐานศูนย์ถูกปฏิเสธ แล้วผู้วิจัยอาจจะต้องการเปรียบเทียบพหุคูณเพื่อหาความแตกต่างของค่าเฉลี่ยที่ต่างกันอย่างมีนัยสำคัญ ซึ่งจะอธิบายต่อไป

อิทธิพลหลักและผลปฏิสัมพันธ์ (Main Effects and Interaction Effects)

ต่อมาจะอธิบายเกี่ยวกับอิทธิพลหลักและผลปฏิสัมพันธ์ อิทธิพลหลักขององค์ประกอบ A นิยามว่าเป็นอิทธิพลขององค์ประกอบ A บนตัวแปรตาม Y ซึ่งจะแสดงอิทธิพลขององค์ประกอบ A บน Y ในทางสถิติจะมีการควบคุมองค์ประกอบ B เอาไว้ด้วย ทำนองเดียวกันกับอิทธิพลหลักขององค์ประกอบ B

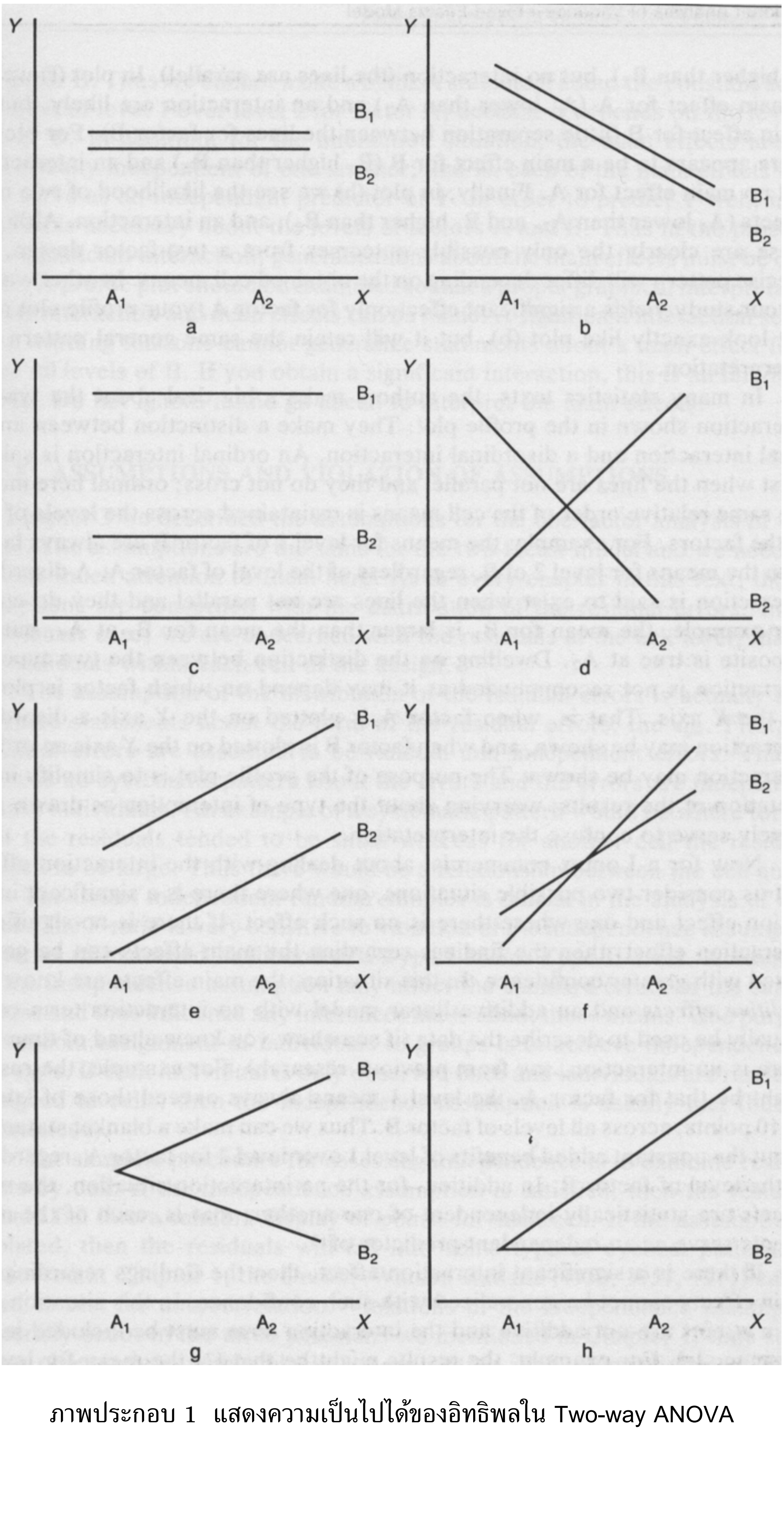
แนวคิดเกี่ยวกับปฏิสัมพันธ์ โดยมากจะซับซ้อน ปฏิสัมพันธ์สามารถนิยามในทางใดทางหนึ่งต่อไปนี้ 1) เป็นการรวมกันของ 2 องค์ประกอบหรือการมีอิทธิพลร่วมกันของทั้งสององค์ประกอบ 2) ความแตกต่างของค่าเฉลี่ยระหว่างระดับขององค์ประกอบ A ที่ไม่คงที่ในแต่ละระดับขององค์ประกอบ B 3) เป็นอิทธิพลร่วมขององค์ประกอบ A และ B บน Y 4) เป็นอิทธิพลที่ไม่สามารถถูกทำนายจากอิทธิพลหลักตัวเดียว ตัวอย่างของผลปฏิสัมพันธ์ ตัวอย่างแรก ปฏิสัมพันธ์ของความถนัดและการจัดกระทำ (Aptitude-Treatment Interaction : ATI) ความหมายนี้มีอิทธิพลของการให้การจัดกระทำขึ้นอยู่กับความถนัดของแต่ละบุคคล ในกรณีอื่น ๆ มีบางการจัดกระทำที่มีอิทธิพลเฉพาะสำหรับบุคคลกับคนที่มีความถนัดสูงและการจัดกระทำอื่น ๆ มีอิทธิพลโดยมากสำหรับความถนัดต่ำ ตัวอย่างที่สองปฏิสัมพันธ์ระหว่างการจัดกระทำและเพศ มีบางการจัดกระทำที่อาจจะมีอิทธิพลมากสำหรับเพศชายและอาจจะมีอิทธิพลมากสำหรับหญิงด้วย เป็นงานวิจัยที่พิจารณาความแตกต่างระหว่างเพศ

แผนภาพที่แสดงเป็นแผนภาพสำหรับอิทธิพลหลักและผลปฏิสัมพันธ์ มองดูแผนภาพในภาพประกอบ 1 ในแต่ละพล็อตจะแสดงชุดของค่าเฉลี่ยในเซล บางครั้งอ้างอิงว่า profile plot บนแกน X เป็นระดับขององค์ประกอบ A และแกน Y เป็นค่าเฉลี่ยในเซลล์บนตัวแปรตาม Y และเส้นภายในที่พล็อตลงไปจะแสดงระดับขององค์ประกอบ B Profile plots จะแสดงสารสนเทศเกี่ยวกับความเป็นไปได้ของอิทธิพลหลักสำหรับ A อิทธิพลหลักสำหรับ B และผลปฏิสัมพันธ์ อิทธิพลหลักขององค์ประกอบ A สามารถตรวจสอบโดยพิจารณาค่าเฉลี่ยในแต่ละระดับของ A และค่าเฉลี่ยของ A ในแต่ละระดับของ B ถ้าค่าเฉลี่ยรวมสำหรับระดับของ A มีค่าเหมือนกันหรือใกล้เคียงกัน จะบ่งชี้การไม่มีอิทธิพลหลักสำหรับองค์ประกอบ A อิทธิพลหลักสำหรับองค์ประกอบ B ควรจะถูกประเมินโดยการนำค่าเฉลี่ยสำหรับแต่ละระดับของ B และค่าเฉลี่ยของ B ในแต่ละระดับของ A ถ้าค่าเฉลี่ยรวมของระดับของ B เหมือนกันหรือใกล้เคียงกันแล้ว เราจะ

สรุปว่าไม่มีอิทธิพลหลักสำหรับองค์ประกอบ B ผลของปฏิสัมพันธ์จะถูกกำหนดโดยค่าเฉลี่ยของเซลสำหรับระดับของ A ที่มีการทำให้ระดับของ B คงที่ (หรือในทางกลับกัน) ซึ่งจะแสดงง่าย ๆ เป็น profile plot โดยการตรวจสอบเห็นการไม่เป็นเส้นตรงที่คู่ขนานกัน เส้นตรงที่คู่ขนานกันจะบ่งชี้ถึงการไม่มีปฏิสัมพันธ์ เมื่อเส้นที่ไม่คู่ขนานกันจะแสดงถึงการมีปฏิสัมพันธ์ แน่นอนว่ามีการทดสอบนัยสำคัญทางสถิติสำหรับอิทธิพลหลักและผลปฏิสัมพันธ์โดยการใช้อัตราส่วน F สำหรับ profile plots จะแสดงให้เห็นถึงการมีอยู่ของปฏิสัมพันธ์ ตัวอย่าง เส้นตรงที่เกือบคู่ขนานกันเป็นไปได้ที่จะไม่แสดงถึงการมีนัยสำคัญของปฏิสัมพันธ์

แผนภาพแสดงดังภาพประกอบ 1 แสดงถึง 8 แบบแผนที่อาจจะเป็นไปได้ในการออกแบบ 2 องค์ประกอบ มีเพียง 2 ระดับในแต่ละองค์ประกอบ พล็อต a) บ่งชี้ถึงการไม่มีอิทธิพลหลักทั้งองค์ประกอบ A และ B และไม่มีผลปฏิสัมพันธ์ เส้นแกนนนอน (ไม่มีอิทธิพลของ A) ทั้งสองเส้นเข้าใกล้กัน (ไม่มีอิทธิพลของ B) และคู่ขนานกัน (ไม่มีปฏิสัมพันธ์) พล็อต b) แสดงถึงการมีอยู่ของอิทธิพลอันเนื่องมาจาก A โดยเฉพาะ (เส้นไม่เป็นแกนนอนเพราะว่าค่าเฉลี่ยของ A_1 สูงกว่าค่าเฉลี่ยของ A_2) แต่ทั้งสองเส้นเข้าใกล้กัน (ไม่มีอิทธิพลของ B) และคู่ขนานกัน (ไม่มีปฏิสัมพันธ์) ในพล็อต c) จะเป็นการแยกส่วนระหว่างเส้นสำหรับระดับของ B (B_1 สูงกว่า B_2) ดังนั้นมีอิทธิพลหลักของ B แต่เป็นเส้นแนวนอน (ไม่มีอิทธิพลของ A) และคู่ขนานกัน (ไม่มีปฏิสัมพันธ์)

สำหรับพล็อต d) ไม่มีอิทธิพลหลัก (ค่าเฉลี่ยสำหรับระดับของ A เท่ากัน และค่าเฉลี่ยสำหรับระดับของ B เท่ากัน) แต่มีปฏิสัมพันธ์โดยทั้งสองเส้นไม่คู่ขนานกัน พล็อต e) แสดงอิทธิพลหลักของทั้งสององค์ประกอบ โดยมีค่าเฉลี่ยแตกต่างกัน (A_1 ต่ำกว่า A_2 และ B_1 สูงกว่า B_2) แต่ไม่มีปฏิสัมพันธ์กัน (เส้นคู่ขนานกัน) ในพล็อต f) จะเห็นอิทธิพลหลักของ A (A_1 ต่ำกว่า A_2) และปฏิสัมพันธ์ยังมีอยู่แต่ไม่มีอิทธิพลหลักของ B (ความแตกต่างของเส้นมีน้อย) สำหรับพล็อต g) จะปรากฏอิทธิพลหลักของ B (B_1 สูงกว่า B_2) และปฏิสัมพันธ์ แต่ไม่มีอิทธิพลหลักของ A สุดท้าย พล็อต h) จะเห็นอิทธิพลหลักของทั้งองค์ประกอบ A และ B (A_1 ต่ำกว่า A_2 และ B_1 สูงกว่า B_2) และมีปฏิสัมพันธ์ แม้ว่าจะชัดเจนเฉพาะผลที่เป็นไปได้ทั้งหมดจากกรณีเป็นรูปแบบ 2 องค์ประกอบ แบบแผนที่แตกต่างกันขึ้นอยู่กับค่าเฉลี่ยของเซลล์ ในอีกกรณีหนึ่ง ถ้าการศึกษาของคุณ มีนัยสำคัญเฉพาะอิทธิพลของ A profile plot ไม่จำเป็นว่าจะต้องเป็นเหมือน พล็อต b) แต่จะอาจจะเหมือนกับแบบแผนทั่ว ๆ ไปและแปลความหมาย



ภาพประกอบ 1 แสดงความเป็นไปได้ของอิทธิพลใน Two-way ANOVA

ในหนังสือสถิติทั่วไป ผู้แต่งอาจจะแสดงแนวคิดเกี่ยวกับชนิดของปฏิสัมพันธ์แสดงใน profile plot ซึ่งอาจจะทำให้เห็นความแตกต่างระหว่างปฏิสัมพันธ์ที่มีลำดับ (ordinal interaction) และปฏิสัมพันธ์ที่ไม่มีลำดับ (disordinal interaction) ปฏิสัมพันธ์ที่มีลำดับก็คือเส้นที่ยังอาจมีปฏิสัมพันธ์อยู่เมื่อเส้นตรงไม่คู่ขนานแต่ไม่ตัดกัน ลำดับในที่นี้หมายถึงความสัมพันธ์ที่เหมือนกันในลำดับของค่าเฉลี่ยในเซลล์ที่มีอยู่ข้ามระดับของอีกองค์ประกอบหนึ่ง ตัวอย่างเช่น ค่าเฉลี่ยสำหรับระดับ 1 ขององค์ประกอบ B มีมากกว่าค่าเฉลี่ยระดับ 2 ของ B ในทุกระดับของ A สำหรับปฏิสัมพันธ์ที่ไม่มีลำดับ ก็คือเส้นที่ไม่คู่ขนานและตัดกัน ตัวอย่างเช่น ค่าเฉลี่ยของ B_1 สูงกว่าค่าเฉลี่ยของ B_2 ที่ระดับ A_1 แต่ค่าเฉลี่ยของ B_1 ต่ำกว่าค่าเฉลี่ยของ B_2 ที่ระดับ A_2 ที่กล่าวมานี้เป็นความแตกต่างระหว่าง 2 ชนิดของปฏิสัมพันธ์ แต่ไม่แน่นอน ขึ้นอยู่กับองค์ประกอบที่ถูกพล็อตลงบนแกน X นั่นคือเมื่อ องค์ประกอบ A ถูกพล็อตลงบนแกน X อาจจะแสดงปฏิสัมพันธ์แบบไม่มีลำดับ และเมื่อองค์ประกอบ B ถูกพล็อตลงบนแกน X อาจจะแสดงปฏิสัมพันธ์แบบมีลำดับ จุดมุ่งหมายของ profile plot คือจะให้ง่ายในการเห็นผลของปฏิสัมพันธ์ ความกังวลเกี่ยวกับชนิดของปฏิสัมพันธ์ที่วาดขึ้นอาจจะทำให้เกิดความสับสนในการแปลความหมาย

พิจารณาใน 2 กรณีที่เป็นไปได้คือ เมื่อผลปฏิสัมพันธ์มีนัยสำคัญหรือไม่มีนัยสำคัญ ถ้าไม่มีนัยสำคัญของผลปฏิสัมพันธ์แล้ว จะพบอิทธิพลหลักสามารถสรุปอ้างอิงในช่วงความเชื่อมั่นที่กว้างกว่า ในกรณีนี้ อิทธิพลหลักจะเป็นการเพิ่มอิทธิพล (additive effects) และเพิ่มโมเดลเชิงเส้นกับการไม่มีปฏิสัมพันธ์ในเทอมทีควรจะใช้อธิบายข้อมูล ตัวอย่างเช่น ผลอาจจะเป็นไปได้ว่าในองค์ประกอบ A ระดับที่ 1 มีค่าเฉลี่ยสูงกว่าระดับที่ 2 เสมอคือ 10 หน่วย ในทุกระดับของ B ดังนั้นสามารถจะให้รายละเอียดเกี่ยวกับการเพิ่มขึ้นอย่างคงที่ของระดับที่ 1 และ 2 ขององค์ประกอบ A โดยไม่สนใจองค์ประกอบ B นอกจากนี้สำหรับกรณีไม่มีปฏิสัมพันธ์ อิทธิพลหลักในทางสถิติจะเป็นอิสระจากองค์ประกอบอื่น นั่นคือ อิทธิพลหลักแต่ละตัวจะเป็นตัวแปรทำนายที่เป็นอิสระจากกัน

ถ้าผลปฏิสัมพันธ์มีนัยสำคัญแล้วจะพบอิทธิพลหลักไม่สามารถสรุปอ้างอิงไปได้ ในกรณีนี้อิทธิพลหลักไม่เพิ่มและมีปฏิสัมพันธ์ในเทอมที่ต้องรวมอยู่ในโมเดลเชิงเส้น ตัวอย่างเช่น ผลอาจจะมีว่า 1) ค่าเฉลี่ยสำหรับระดับ 1 ขององค์ประกอบ A มากกว่าระดับที่ 2 เมื่อพิจารณาที่ระดับที่ 1 ขององค์ประกอบ B แต่ 2) ค่าเฉลี่ยของระดับที่ 1 องค์ประกอบ A น้อยกว่าระดับที่ 2 เมื่อพิจารณาที่ระดับที่ 2 ขององค์ประกอบ B ดังนั้นเราไม่สามารถให้รายละเอียดเกี่ยวกับความคงที่ที่เพิ่มขึ้นของระดับที่ 1 และ 2 ขององค์ประกอบ A เพิ่มว่าขึ้นอยู่กับระดับขององค์ประกอบ B นอกจากนี้ สำหรับกรณีมีปฏิสัมพันธ์ อิทธิพลหลักในทางสถิติจะไม่เป็นอิสระจากองค์ประกอบอื่น นั่นคือ ในอิทธิพลหลักแต่ละตัวจะไม่เป็นตัวแปรที่เป็นอิสระจากกัน ดังนั้นในการมีอยู่ของนัยสำคัญของปฏิสัมพันธ์ การสรุปอ้างอิงเกี่ยวกับอิทธิพลหลักต้องเป็นเชิงปริมาณ profile plot ควรจะถูกตรวจสอบและแปลความหมายด้วยแผนภาพของปฏิสัมพันธ์และอิทธิพลหลักที่มีอยู่ การมีนัยสำคัญของปฏิสัมพันธ์เป็นการเตือนไม่สามารถสรุปอ้างอิงไปยังอิทธิพลหลักของ A ในทุก

ๆ ระดับของ B ได้ ถ้าคุณยังอ้างอิงถึงการมีนัยสำคัญของปฏิสัมพันธ์ นั่นคือมีความสำคัญในผลที่ได้ จะต้องไม่เพิกเฉยและนำไปสู่การแปลความหมายอิทธิพลหลัก

ข้อตกลงเบื้องต้นและการละเมิดของข้อตกลงเบื้องต้น

จากข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนแบบองค์ประกอบเดียว ก็จะทำนองเดียวกับการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ ข้อตกลงเบื้องต้นที่แน่นอนคือการแจกแจงของความคลาดเคลื่อน ในกรณีสององค์ประกอบนี้ เราจะเกี่ยวข้องกับความคลาดเคลื่อนในระดับของเซลล์ นั่นคือความคลาดเคลื่อนที่อยู่ภายในแต่ละเซลล์

ข้อตกลงเบื้องต้นของการแจกแจงของความคลาดเคลื่อนในทางปฏิบัติจะเป็นชุดของข้อตกลงเบื้องต้น 3 ประการเกี่ยวกับความคลาดเคลื่อน ε_{ijk} คือ ประการแรก ความคลาดเคลื่อนจะสมมติว่าเป็นไปอย่างสุ่มและเป็นอิสระจากกัน นั่นคือจะไม่มีความคลาดเคลื่อนอย่างเป็นระบบ และความคลาดเคลื่อนแต่ละค่าจะต้องเป็นอิสระจากกัน

ในการใช้กลุ่มตัวอย่างที่สุ่มมาอย่างอิสระจากกันสำคัญในการวิเคราะห์ความแปรปรวน อัตราส่วน F จะอ่อนไหวมากในการละเมิดข้อตกลงเบื้องต้นของความเป็นอิสระนี้จะช่วยเพิ่มความคลาดเคลื่อนแบบที่ I และ II ให้สูงขึ้น การละเมิดข้อตกลงเบื้องต้นของความเป็นอิสระอาจจะมีอิทธิพลต่อความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยและมีอิทธิพลส่งไปจนถึงค่าเฉลี่ยรวม ในจุดประสงค์หนึ่งของการสุ่มกลุ่มตัวอย่างเข้ากลุ่มแต่ละกลุ่มควรจะต้องให้เป็นอิสระจากกัน ถ้าแต่ละหน่วยของตัวอย่างถูกสุ่มเข้าในแต่ละเซลล์อย่างอิสระจากกันแล้ว ข้อตกลงเบื้องต้นของความเป็นอิสระของความคลาดเคลื่อนนี้จะเป็นจริง

กระบวนการอย่างง่ายสำหรับการประเมินความเป็นอิสระของความคลาดเคลื่อนโดยการนำความคลาดเคลื่อนแต่ละเซลล์มาพล็อตกราฟ ถ้าข้อตกลงเบื้องต้นของความเป็นอิสระของความคลาดเคลื่อนเป็นจริงแล้ว ความคลาดเคลื่อนจะตกอยู่อย่างสุ่ม ถ้าข้อตกลงเบื้องต้นนี้ถูกละเมิดแล้วความคลาดเคลื่อนจะตกอยู่เป็นวงล้อมกัน สามารถใช้สถิติ Durbin-Watson ในการตรวจสอบ autocorrelation การละเมิดข้อตกลงเบื้องต้นของความเป็นอิสระของความคลาดเคลื่อนมีโอกาสเกิดขึ้นได้ 3 กรณีคือ ข้อมูลเป็นแบบอนุกรมเวลา (time series data) ค่าสังเกตถูกจัดเป็นบล็อกและมีการวัดซ้ำ สำหรับการแก้ไขปัญหาการละเมิดข้อตกลงเบื้องต้นนี้อาจจะใช้การแปลงรูปข้อมูลหรือใช้การทดสอบแบบไร้พารามิเตอร์ (nonparametric tests) ตัวอย่างการละเมิดข้อตกลงของข้อมูลที่เป็นอนุกรมเวลา autocorrelation ระหว่างแถวอาจจะมีอิทธิพลในการทดสอบของสดมภ์ และในทางกลับกัน ถ้าการละเมิดเนื่องมาจากองค์ประกอบแทรกซ้อนภายนอก ในกรณีควรจะรวมองค์ประกอบแทรกซ้อนเข้าไว้ในการวิเคราะห์เป็นตัวแปรอิสระอีกตัวหนึ่งในโมเดลด้วย

ส่วนที่ 2 ของข้อตกลงเบื้องต้นคือการแจกแจงของความคลาดเคลื่อนสำหรับแต่ละเซลล์จะมีความแปรปรวนคงที่ (σ^2_{res}) มีเสมอที่ถูกอ้างอิงว่าเป็นข้อตกลงเบื้องต้นของความเป็นเอกพันธ์ของความแปรปรวน (homogeneity of variance or homoscedasticity) ในอีกกรณีหนึ่ง

สำหรับทุก ๆ เซลล์แล้ว การแจกแจงแบบมีเงื่อนไขของความคลาดเคลื่อนจะมีความแปรปรวนเท่ากัน ถ้าข้อตกลงเบื้องต้นเป็นจริงแล้ว s^2_{res} จะเป็นตัวประมาณค่าที่ไม่ลำเอียงของ σ^2_{res} สำหรับแต่ละเซลล์

การละเมิดข้อตกลงเบื้องต้นของความเป็นเอกพันธ์ของความแปรปรวนอาจจะนำไปสู่ความลำเอียงใน SS_{with} ซึ่งจะเป็นการเพิ่มความคลาดเคลื่อนแบบที่ I และเป็นไปได้ที่จะเพิ่มความคลาดเคลื่อนแบบที่ II ด้วย อิทธิพลของการละเมิดนี้ดูเหมือนจะเล็กน้อยกับการออกแบบที่แต่ละกลุ่มมีตัวอย่างเท่ากันหรือใกล้เคียงกัน (balanced or nearly balanced designs) นอกจากนี้กรณีที่มีการเพิ่ม n ให้มากขึ้น ผลที่เกิดจากการละเมิดข้อตกลงเบื้องต้นนี้จะลดลง ดังนั้นจะมีปัญหาน้อยหากใช้การออกแบบที่กลุ่มตัวอย่างแต่ละกลุ่มเท่ากันหรือใกล้เคียงกัน

ส่วนที่ 3 ของข้อตกลงเบื้องต้นก็คือ การแจกแจงของความคลาดเคลื่อนจะเป็นโค้งปกติ นั่นคือสำหรับทุก ๆ เซลล์แล้ว ความคลาดเคลื่อนจะมีการแจกแจงเป็นโค้งปกติ ผลของการละเมิดข้อตกลงเบื้องต้นเหมือนกับข้อตกลงเบื้องของความเป็นเอกพันธ์ของความแปรปรวน การค้นหาการละเมิดข้อตกลงเบื้องต้นของการแจกแจงปกติสามารถใช้กราฟพิจารณาการแจกแจงของความคลาดเคลื่อนของแต่ละเซลล์

ตารางสรุปข้อตกลงเบื้องต้นและผลของการละเมิดข้อตกลงเบื้องต้นแสดงในตาราง 2

ตาราง 2 ข้อตกลงเบื้องต้นและอิทธิพลของการละเมิดข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ

ข้อตกลงเบื้องต้น	ผลของการละเมิดข้อตกลงเบื้องต้น
1. ความเป็นอิสระของความคลาดเคลื่อน	การเพิ่มขึ้นของ Type I และ Type II error ในการทดสอบ F และอิทธิพลต่อความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย อิทธิพลต่อค่าเฉลี่ยรวม
2. ความเป็นเอกพันธ์ของความคลาดเคลื่อน	มีความลำเอียงใน SS_{with} , การเพิ่มขึ้นของ Type I และ Type II error, มีอิทธิพลน้อยต่อการออกแบบที่กลุ่มตัวอย่างแต่ละเซลล์เท่ากันหรือใกล้เคียงกัน, และมีอิทธิพลน้อยเมื่อเพิ่ม n ให้มากขึ้น
3. การแจกแจงปกติของความคลาดเคลื่อน	มีอิทธิพลน้อยต่อการออกแบบที่กลุ่มตัวอย่างแต่ละเซลล์เท่ากันหรือใกล้เคียงกัน, มีอิทธิพลน้อยเมื่อเพิ่ม n ให้มากขึ้น

การคำนวณผลรวมกำลังสอง (Sums of Squares)

การแบ่งส่วนผลรวมกำลังสองเป็นแนวคิดสำคัญทั้งในการวิเคราะห์การถดถอยและการวิเคราะห์ความแปรปรวน เริ่มต้นจากผลรวมกำลังสองของรวมทั้งหมดในตัวแปร Y ใช้สัญลักษณ์ว่า SS_{total} เทอม SS_{total} จะแสดงปริมาณของความแปรปรวนรวมทั้งหมดของค่าสังเกต ในขั้นถัดมาจะแบ่งส่วนความแปรปรวนร่วมเป็นความแปรปรวนระหว่างระดับขององค์ประกอบ A (สัญลักษณ์ SS_A) ความแปรปรวนระหว่างระดับขององค์ประกอบ B (สัญลักษณ์ SS_B) ความแปรปรวนเนื่องจากปฏิสัมพันธ์ของระดับในองค์ประกอบ A และ B (สัญลักษณ์ SS_{AB}) และความแปรปรวนภายในเซล (สัญลักษณ์ SS_{with}) ในการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบสามารถแบ่งส่วน SS_{total} ได้ดังนี้

$$SS_{total} = SS_A + SS_B + SS_{AB} + SS_{with}$$

หรือเขียนในเทอมของพารามิเตอร์ของประชากรได้ดังนี้

$$\sum_i^n\sum_j^J\sum_k^K(Y_{ijk}-\mu_{...})^2 = nK\sum_j^J\alpha_j^2 + nJ\sum_k^K\beta_k^2 + n\sum_j^J\sum_k^K(\alpha\beta)_{jk}^2 + \sum_i^n\sum_j^J\sum_k^K\varepsilon_{ijk}^2$$

อาจอ้างอิงสูตรของการแบ่งส่วนผลรวมกำลังสองเรียกว่า definitional formula เพราะว่าในแต่ละเทอมจะเป็นนิยามรูปแบบของความแปรปรวน

เนื่องจากการคำนวณที่ซับซ้อน จึงจะนำเสนอสูตรที่ใช้กับข้อมูลจริง โดยสามารถแทนที่ค่าต่าง ๆ ลงสูตรและคำนวณด้วยมือได้ ซึ่งมีสูตรคำนวณดังนี้

$$SS_{total} = \sum_i^n\sum_j^J\sum_k^K Y_{ijk}^2 - \left(\sum_i^n\sum_j^J\sum_k^K Y_{ijk}\right)^2 \Big/ N$$

$$SS_A = \sum_j^J\left[\left(\sum_i^n\sum_k^K Y_{ijk}\right)^2 \Big/ nK\right] - \left(\sum_i^n\sum_j^J\sum_k^K Y_{ijk}\right)^2 \Big/ N$$

$$SS_B = \sum_k^K\left[\left(\sum_i^n\sum_j^J Y_{ijk}\right)^2 \Big/ nJ\right] - \left(\sum_i^n\sum_j^J\sum_k^K Y_{ijk}\right)^2 \Big/ N$$

$$SS_{AB} = \sum_j^J\sum_k^K\left[\left(\sum_i^n Y_{ijk}\right)^2 \Big/ n\right] - \sum_j^J\left[\left(\sum_i^n\sum_k^K Y_{ijk}\right)^2 \Big/ nK\right] - \sum_k^K\left[\left(\sum_i^n\sum_j^J Y_{ijk}\right)^2 \Big/ nJ\right] + \left(\sum_i^n\sum_j^J\sum_k^K Y_{ijk}\right)^2 \Big/ N$$

$$SS_{with} = \sum_i^n\sum_j^J\sum_k^K Y_{ijk}^2 - \sum_j^J\sum_k^K\left[\left(\sum_i^n Y_{ijk}\right)^2 \Big/ n\right]$$

ในกรณีที่มีความซับซ้อนมากกว่าหนึ่งองค์ประกอบ ซึ่งจะเห็นได้จากชุดของสมการที่มีจำนวนมากและเทอมที่ต้องคำนวณแตกต่างกันมากถึง 5 เทอม

ตารางสรุป ANOVA

เมื่อสามารถคำนวณหาผลรวมกำลังสองมาได้หมดแล้ว ขั้นถัดมาคือการลงตารางสรุป ANOVA โดยมีจุดประสงค์เพื่อจะสรุปผลการวิเคราะห์ความแปรปรวนในรูปที่ง่ายที่สุด รูปแบบทั่วไปของตารางสรุปสำหรับกรณีสององค์ประกอบปรากฏดังตาราง 3 สดมภ์แรกจะเป็นแหล่งของความแปรปรวนในโมเดล สังเกตว่าความแปรปรวนร่วมจะถูกแบ่งออกเป็นความแปรปรวนภายในกลุ่ม และระหว่างกลุ่ม ซึ่งความแปรปรวนระหว่างกลุ่มจะแบ่งย่อยออกเป็นความแปรปรวนเนื่องมาจากองค์ประกอบ A, B และปฏิสัมพันธ์ระหว่าง A และ B ทำนองเดียวกับกรณีหนึ่งองค์ประกอบที่แบ่งออกเป็นความแปรปรวนภายในกลุ่มและความแปรปรวนระหว่างกลุ่ม สดมภ์ที่สองจะเป็นผลรวมกำลังสองที่คำนวณมาจากสูตรทั้งหมดในหัวข้อข้างต้น

สดมภ์ที่สามจะเป็นองศาแห่งความเป็นอิสระ (degrees of freedon) สำหรับแต่ละแหล่ง โดยปกติองศาแห่งความเป็นอิสระจำเป็นจำนวนของค่าสังเกตที่เป็นอิสระ ถ้าองค์ประกอบมี J ระดับแล้ว จำนวนองศาแห่งความเป็นอิสระของ A จะเท่ากับ J - 1 ทำไม เมื่อมีค่าเฉลี่ย J ค่า และเรารู้ค่าเฉลี่ยรวมแล้ว จะมีค่าเฉลี่ยเพียง J - 1 ค่าเท่านั้นที่สามารถแปรค่าได้อย่างอิสระ สำหรับองค์ประกอบ B มีจำนวน K ระดับ องศาแห่งความเป็นอิสระจะเท่ากับ K - 1 สำหรับปฏิสัมพันธ์ระหว่าง A และ B ก็จะมีองศาแห่งความเป็นอิสระทั้งของ A และ B หรือเท่ากับ (J - 1)(K - 1) ส่วนความแปรปรวนภายในกลุ่มจะเท่ากับจำนวนค่าสังเกตรวมทั้งหมดลบกับจำนวนเซลล์ในโมเดลหรือคือ N - JK สุดท้ายองศาแห่งความเป็นอิสระของรวมจะเท่ากับ N - 1

ถัดมาการนำผลรวมกำลังสองมาถ่วงน้ำหนักด้วยองศาแห่งความเป็นอิสระจะเรียกว่าค่าเฉลี่ยกำลังสอง ดังนั้น $MS_A = SS_A/df_A$ และสดมภ์สุดท้ายของตารางสรุป ANOVA จะเป็นค่า F ซึ่งจะเป็นค่าสถิติสรุปสำหรับการวิเคราะห์ความแปรปรวน โดยมีสมมติฐาน 3 ข้อที่ต้องสนใจทดสอบ สถิติ F-test ทั้งสามนี้ จะทดสอบอิทธิพลหลัก 2 ค่าและทดสอบผลปฏิสัมพันธ์ 1 ค่า กรณีเป็นการวิเคราะห์ความแปรปรวนแบบอิทธิพลกำหนด (fixed-effects) ค่าสถิติ F นี้สามารถคำนวณได้โดยการนำ MS นั้น ๆ มาหารด้วย MS_{with} ดังนั้นแต่ละสมมติฐานที่ถูกทดสอบจะมีความคลาดเคลื่อนเทอมเดียวกัน

สถิติ F แต่ละค่าจะถูกเปรียบเทียบกับค่าวิกฤติที่เปิดจากตาราง ค่าวิกฤติของการทดสอบองค์ประกอบ A จากตาราง ${}_{(1-\alpha)}F_{(J-1,N-JK)}$ ค่าวิกฤติสำหรับการทดสอบองค์ประกอบ B จากตาราง ${}_{(1-\alpha)}F_{(K-1,N-JK)}$ และค่าวิกฤติสำหรับการทดสอบผลปฏิสัมพันธ์ AB จากตาราง ${}_{(1-\alpha)}F_{[(J-1)(K-1),N-JK]}$ แต่ละการทดสอบจะเป็นแบบทิศทางเดียวซึ่งสอดคล้องกับสมมติฐานอื่นที่กำหนดไว้ สมมติฐานศูนย์จะถูกปฏิเสธถ้าค่า F ที่คำนวณได้มีค่ามากกว่าค่า F วิกฤติจากตาราง

ถ้าสถิติทดสอบ F มีค่ามากกว่าค่าวิกฤติจากตาราง แล้วมีความชัดเจนว่าสมมติฐานศูนย์จะถูกปฏิเสธ ในกรณีนี้อาจจะต้องมีการเปรียบเทียบพหุคูณเพื่อค้นหาความแตกต่างของค่าเฉลี่ยระหว่างแต่ละกลุ่ม ในหัวข้อถัดไป

ตาราง 3 ตารางสรุปผลการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม :				
A	SS_A	J - 1	MS_A	MS_A / MS_{with}
B	SS_B	K - 1	MS_B	MS_B / MS_{with}
AB	SS_{AB}	(J - 1)(K - 1)	MS_{AB}	MS_{AB} / MS_{with}
ภายในกลุ่ม	SS_{with}	N - JK	MS_{with}	
รวม	SS_{total}	N - 1		

วิธีการเปรียบเทียบพหุคูณ (Multiple Comparison Procedures)

ในหัวข้อนี้เราจะขยายแนวคิดที่เกี่ยวข้องกับวิธีการเปรียบเทียบพหุคูณ (MCPs) ที่เคยเสนอไว้ในเรื่องการวิเคราะห์ความแปรปรวนแบบหนึ่งองค์ประกอบ ซึ่งการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบจะรวมถึงอิทธิพลหลักและผลปฏิสัมพันธ์ ซึ่งอาจจะต้องเปรียบเทียบทั้งคู่ของอิทธิพลหลักและผลปฏิสัมพันธ์ ซึ่งวิธีโดยทั่วไปได้อธิบายไปแล้วในเรื่องการวิเคราะห์ความแปรปรวนแบบหนึ่งองค์ประกอบที่สามารถนำมาประยุกต์ใช้กับกรณีสององค์ประกอบได้ นั่นคือเราจะมีการเปรียบเทียบค่าเฉลี่ยทั้งในแนวแถวและสดมภ์ และค่าเฉลี่ยภายในเซลล์ ดังนั้นต้องมีความระมัดระวังและชัดเจนเกี่ยวกับค่าเฉลี่ยที่จะนำมาเปรียบเทียบ

เริ่มต้นดำเนินการเปรียบเทียบอิทธิพลหลัก ถ้าอิทธิพลสำหรับองค์ประกอบ A มีนัยสำคัญและมีมากกว่า 2 ระดับในองค์ประกอบ A เราจะต้องเปรียบเทียบค่าเฉลี่ยสำหรับแต่ละระดับขององค์ประกอบ A โดยใช้ค่าเฉลี่ยในแต่ละระดับขององค์ประกอบ A นอกจากนี้ n ควรจะมีอิทธิพลเป็นจำนวนของค่าเฉลี่ยที่ต้องนำมาเปรียบเทียบ และจำนวนของค่าสังเกตในแต่ละเซลล์ ดังนั้นการเปรียบเทียบที่เกี่ยวข้องกับองค์ประกอบ A จำนวนของค่าสังเกตจะเท่ากับ nK (สัญลักษณ์คือ n_j) และสำหรับการเปรียบเทียบขององค์ประกอบ B จำนวนของค่าสังเกตคือ nJ (สัญลักษณ์คือ n_k) พิจารณาแต่ละองค์ประกอบแยกกัน มีสถิติบางตัวที่จะแนะนำสำหรับการเปรียบเทียบกรณีเป็นโมเดลองค์ประกอบเดียวกับมีระดับ JK ระดับเมื่อใช้ MCPs ในการตรวจสอบอิทธิพลหลัก ซึ่งจะมีสอดคล้องกับการออกแบบและมีแนวโน้มจะต้องแยกวิเคราะห์แต่ละอิทธิพล และไม่แนะนำ

สำหรับการเปรียบเทียบที่เกี่ยวข้องกับปฏิสัมพันธ์จะมีความซับซ้อน โดยเริ่มต้นการเปรียบเทียบปฏิสัมพันธ์ ถ้าในโมเดลมีมากกว่า 4 เซลล์ ดังนั้นรูปแบบ 4 x 4 จะถูกออกแบบสอดคล้องกับระดับองค์ประกอบ A มี 4 ระดับ และระดับองค์ประกอบ B อีก 4 ระดับ หรืออาจจะ

เปรียบเทียบรูปแบบของ 4 x 2 หรือวิธีการสอน 4 ระดับและผู้สอน 2 ระดับ ตัวอย่างการเปรียบเทียบ

$$\Psi' = \frac{1}{4}(\overline{Y}_{.11} + \overline{Y}_{.21} + \overline{Y}_{.31} + \overline{Y}_{.41}) - \frac{1}{4}(\overline{Y}_{.12} + \overline{Y}_{.22} + \overline{Y}_{.32} + \overline{Y}_{.42})$$

และความคลาดเคลื่อนมาตรฐาน

$$SE(\Psi') = \sqrt{[MS_{error}(\sum_j \sum_k \frac{c_{jk}^2}{n_{jk}})]}$$

เมื่อ n_{jk} คือจำนวนของค่าสังเกตในเซลล์ jk ซึ่งการเปรียบเทียบนี้ควรจะตรวจสอบปฏิสัมพันธ์ระหว่างวิธีการทั้ง 4 ของวิธีการสอนและอีก 2 ระดับของผู้สอน การเปรียบเทียบปฏิสัมพันธ์ที่ซับซ้อนควรจะพิจารณาปฏิสัมพันธ์ระหว่างวีธีการสอนทั้ง 4 และผู้สอนอีก 2 ระดับที่เหลือ

ถ้าการเปรียบเทียบปฏิสัมพันธ์ที่ซับซ้อนมีนัยสำคัญแล้ว การเปรียบเทียบปฏิสัมพันธ์อย่างง่ายจะเกี่ยวข้องกับค่าเฉลี่ยที่มีเพียงแค่ 4 เซลล์ จะเกี่ยวข้องกับองค์ประกอบ A และ B ที่มีเพียงองค์ประกอบละ 2 ระดับเท่านั้น สามารถเปรียบเทียบได้ดังนี้

$$\Psi' = (\overline{Y}_{.11} - \overline{Y}_{.21}) - (\overline{Y}_{.12} - \overline{Y}_{.22})$$

กับเทอมของความคลาดเคลื่อนมาตรฐานเดียวกัน ใช้ตัวอย่างนี้สำหรับการเปรียบเทียบระหว่างวิธีเดียวจาก 2 วิธีการสอนและระดับเพียงจาก 2 ระดับของผู้สอน

สถิติอื่น ๆ ที่แนะนำในการใช้สำหรับอิทธิพลหลักในการทดสอบที่ปฏิสัมพันธ์มีนัยสำคัญ จะเกี่ยวข้องกับเปรียบเทียบระดับขององค์ประกอบ A ในแต่ละระดับขององค์ประกอบ B และโดยทั่วไปจะใช้การแบ่งส่วนผลรวมกำลังสอง อย่างไรก็ตามอิทธิพลง่าย ๆ ในผลรวมกำลังสองจะแสดงสัดส่วนของอิทธิพลหลักรวมกับผลปฏิสัมพันธ์

MCPs โดยมากที่ได้อธิบายไปแล้วในเรื่องการวิเคราะห์ความแปรปรวนแบบองค์ประกอบเดียวสามารถใช้ทดสอบอิทธิพลหลักและผลปฏิสัมพันธ์ได้ซึ่งคุณต้องแน่ใจว่าสามารถเปรียบเทียบเซลล์ของค่าเฉลี่ยได้เหมาะสมกับจำนวนของค่าสังเกตที่ถูกต้องและองศาแห่งความเป็นอิสระที่ถูกต้อง

การวัดความสัมพันธ์ (Measures of Association)

มีการวัดหลากหลายที่เกี่ยวข้องกับความสัมพันธ์ระหว่างอิทธิพลและ Y ซึ่งจะมีการวัดอยู่ 2 อย่าง สมมติว่าแต่ละเซลล์มีความแปรปรวนเท่ากันแล้ว การวัดแรกคือ η^2 ซึ่งคืออัตราส่วนความสัมพันธ์ ซึ่งจะแสดงสัดส่วนของความแปรปรวนใน Y ที่ถูกอธิบายได้ด้วยอิทธิพลที่สนใจศึกษา (เช่น องค์ประกอบ A หรือองค์ประกอบ B หรือปฏิสัมพันธ์ AB) สามารถคำนวณ η^2 ได้ดังนี้

$$\eta_A^2 = SS_A / SS_{total}$$

$$\eta_B^2 = SS_B / SS_{total}$$
$$\eta_{AB}^2 = SS_{AB} / SS_{total}$$

สถิตินี้มีแนวคิดทำนองเดียวกับสถิติ R^2 ในการวิเคราะห์การถดถอย

อีกการวัดหนึ่งที่เป็นความสัมพันธ์ระหว่างอิทธิพลและ Y คือสถิติ ω^2 สามารถคำนวณ ω^2 ได้ดังนี้

$$\omega_A^2 = [SS_A - (J-1)MS_{with}]/[SS_{total} + MS_{with}]$$
$$\omega_B^2 = [SS_B - (K-1)MS_{with}]/[SS_{total} + MS_{with}]$$
$$\omega_{AB}^2 = [SS_{AB} - (J-1)(K-1)MS_{with}]/[SS_{total} + MS_{with}]$$

ขึ้นอยู่กับอิทธิพลที่สนใจศึกษา

ในแต่ละการวัดนี้จะมีข้อจำกัดคือแต่ละกลุ่มควรมีความแปรปรวนเท่ากัน หรือมี n เท่ากัน ดังนั้นเป็นความลำบากที่จะแนะนำการใช้การวัดนี้ นอกจากนี้ยังไม่มีกฎใด ๆ สำหรับการแปลความหมายขนาดของสถิติที่ได้ ซึ่งโดยปกติค่าการวัดความสัมพันธ์มีค่าตั้งแต่ 0 (ไม่มีความสัมพันธ์) จนถึง 1 (สัมพันธ์กันอย่างสมบูรณ์) ขึ้นอยู่กับผู้วิจัยที่จะแปลความหมายขนาดของความสัมพันธ์ที่ได้อาจเปรียบเทียบค่าที่ได้กับงานวิจัยอื่น ๆ ที่วิจัยทำนองเดียวกันนี้

ตัวอย่างคำนวณ

พิจารณาตัวอย่างการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ เราจะขยายเพิ่มจากเนื้อหาการวิเคราะห์ความแปรปรวนแบบหนึ่งองค์ประกอบ โดยเพิ่มองค์ประกอบที่สองเข้าไปในโมเดล ตัวแปรตามก็จะเป็นระยะเวลาในการเข้าห้องแล็บในภาคเรียนหนึ่ง เมื่อองค์ประกอบ A คือการใช้เวลาสอนในห้องแล็บ (ซึ่งผู้สอนแต่ละคนมีเพศและองค์ประกอบเท่าเทียมกัน) และองค์ประกอบ B คือช่วงเวลาที่ใช้ห้องแล็บ ดังนั้นผู้วิจัยสนใจที่จะศึกษาอิทธิพลของระยะเวลาการสอนในห้องแล็บ ช่วงเวลาที่ใช้ห้องแล็บที่มีอิทธิพลต่อนระยะเวลาการเข้าใช้ห้องแล็บของนักเรียนในภาคเรียนหนึ่ง โดยระดับของเวลาสอนในห้องแล็บมี 4 ระดับคือ 1) ไม่ใช้ห้องแล็บในการสอน 2) ใช้ห้องแล็บในการสอนเพียงเล็กน้อย 3) ใช้ห้องแล็บในการสอนปานกลาง และ 4) ใช้ห้องแล็บในการสอนมาก ส่วนช่วงเวลาที่ใช้ห้องแล็บแบ่งออกเป็น 2 ระดับคือ 1) ใช้แล็บในเวลาบ่าย และ 2) ใช้แล็บในเวลาเย็น นักเรียนแต่ละคนถูกสุ่มเข้ารับผู้สอนแล็บแต่ละคนและช่วงเวลาที่ใช้แล็บ โดยมีนักเรียน 4 คน ในแต่ละเซลล์และรวมทั้งหมดมี 8 เซลล์ รวมนักเรียนทั้งหมด 32 คน นักเรียนใช้เวลาเรียนในรายวิชานี้ทั้งหมด 30 ชั่วโมง ตาราง 4 จะแสดงข้อมูลและค่าเฉลี่ยของกลุ่มตัวอย่างสำหรับเซลล์ทั้งหมดทั้งในแนวแถว สดมภ์และรวม

ตาราง 4 ข้อมูลตัวอย่างการใช้ห้องแล็บ

	ช่วงบ่าย	ช่วงเย็น	ค่าเฉลี่ย (แถว)
ไม่ใช้ห้องแล็บในการสอน	15	10	11.13
	12	8	
	21	7	
	13	3	
	15.25	7.00	
ใช้ห้องแล็บสอนเล็กน้อย	20	13	17.88
	22	9	
	24	18	
	25	12	
	22.75	13.00	
ใช้ห้องแล็บสอนปานกลาง	24	10	20.25
	29	12	
	27	21	
	25	14	
	26.25	14.25	
ใช้ห้องแล็บสอนมาก	30	22	24.38
	26	20	
	29	25	
	28	15	
	28.25	20.50	
ค่าเฉลี่ย (สดมภ์)	23.13	13.69	18.41 (ค่าเฉลี่ยรวม)

เริ่มวิเคราะห์ความแปรปรวน โดยประการแรกเราจะต้องคำนวณผลรวมกำลังสอง

$$SS_{total} = \sum_{i}^{n}\sum_{j}^{J}\sum_{k}^{K} Y_{ijk}^2 - \left(\sum_{i}^{n}\sum_{j}^{J}\sum_{k}^{K} Y_{ijk}\right)^2 \Big/ N$$

$$= 12591 - (346921/32)$$

$$= 1749.72$$

$$SS_A = \sum_j^J \left[\left(\sum_i^n \sum_k^K Y_{ijk} \right)^2 \Big/ nK \right] - \left(\sum_i^n \sum_j^J \sum_k^K Y_{ijk} \right)^2 \Big/ N$$

$$= 11579.88 - (346921/32)$$

$$= 738.60$$

$$SS_B = \sum_k^K \left[\left(\sum_i^n \sum_j^J Y_{ijk} \right)^2 \Big/ nJ \right] - \left(\sum_i^n \sum_j^J \sum_k^K Y_{ijk} \right)^2 \Big/ N$$

$$= 11553.81 - (346921/32)$$

$$= 712.53$$

$$SS_{AB} = \sum_j^J \sum_k^K \left[\left(\sum_i^n Y_{ijk} \right)^2 \Big/ n \right] - \sum_j^J \left[\left(\sum_i^n \sum_k^K Y_{ijk} \right)^2 \Big/ nK \right]$$

$$- \sum_k^K \left[\left(\sum_i^n \sum_j^J Y_{ijk} \right)^2 \Big/ nJ \right] + \left(\sum_i^n \sum_j^J \sum_k^K Y_{ijk} \right)^2 \Big/ N$$

$$= 12314.25 - 11579.88$$

$$- 11553.81 + (346921/32)$$

$$= 21.84$$

$$SS_{with} = \sum_i^n \sum_j^J \sum_k^K Y_{ijk}^2 - \sum_j^J \sum_k^K \left[\left(\sum_i^n Y_{ijk} \right)^2 \Big/ n \right]$$

$$= 12591 - 12314.25$$

$$= 276.75$$

ถัดมาคำนวณค่าเฉลี่ยกำลังสอง

$$MS_A = SS_A/df_A = 738.60/3 = 246.20$$

$$MS_B = SS_B/df_B = 712.53/1 = 712.53$$

$$MS_{AB} = SS_{AB}/df_{AB} = 21.84/3 = 7.28$$

$$MS_{with} = SS_{with}/df_{with} = 276.75/2 = 11.53$$

สุดท้ายคำนวณหาสถิติทดสอบ F ได้ค่า

$$F_A = MS_A/MS_{with} = 246.20/11.53 = 21.35$$

$$F_B = MS_B/MS_{with} = 712.53/11.53 = 61.79$$

$$F_{AB} = MS_{AB}/MS_{with} = 7.28/11.53 = 0.63$$

สถิติทดสอบจะนำไปเปรียบเทียบค่าวิกฤติ ${}_{.95}F_{3,24}$ = 3.01 สำหรับอิทธิพลของ A และอิทธิพลของ AB และ ${}_{.95}F_{1,24}$ = 4.26 สำหรับอิทธิพลของ B สามารถเปิดได้จากตาราง ที่ระดับนัยสำคัญ 0.05 สถิติทดสอบมากกว่าค่าวิกฤติสำหรับอิทธิพลของ A และ B เท่านั้น ดังนั้นเราจะ

ปฏิเสธ H_0 และสรุปว่าทั้งคู่ของการใช้เวลาในการสอนแล็บและช่วงเวลาสอนแล็บมีความสัมพันธ์ทำให้เกิดความแตกต่างระหว่างค่าเฉลี่ยของระยะเวลาในการใช้แล็บของผู้เรียน ผลปฏิสัมพันธ์ไม่มีนัยสำคัญ นำค่าต่าง ๆ ที่คำนวณได้มาสรุปลงตาราง ANOVA ดังตาราง 5

ตาราง 5 ตารางสรุปผลการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม :				
A	738.60	3	246.20	21.35[a]
B	712.53	1	712.53	61.79[b]
AB	21.84	3	7.28	0.63[a]
ภายในกลุ่ม	276.75	24		
รวม	1749.72	31		

[a] ${}_{.95}F_{3,24} = 3.01$

[b] ${}_{.95}F_{1,24} = 4.26$

ถัดมาประมาณค่าอิทธิพลหลักและผลปฏิสัมพันธ์ อิทธิพลหลักสำหรับแต่ละระดับของ A ประมาณค่าได้ดังนี้

$$a_1 = \bar{Y}_{.1.} - \bar{Y}_{...} = 11.13 - 18.41 = -7.28$$

$$a_2 = \bar{Y}_{.2.} - \bar{Y}_{...} = 17.88 - 18.41 = -0.53$$

$$a_3 = \bar{Y}_{.3.} - \bar{Y}_{...} = 20.25 - 18.41 = +1.84$$

$$a_4 = \bar{Y}_{.4.} - \bar{Y}_{...} = 24.38 - 18.41 = +5.97$$

อิทธิพลหลักสำหรับระดับของ B ประมาณค่าได้ดังนี้

$$b_1 = \bar{Y}_{..1} - \bar{Y}_{...} = 23.13 - 18.41 = +4.72$$

$$b_2 = \bar{Y}_{..2} - \bar{Y}_{...} = 13.69 - 18.41 = -4.72$$

สุดท้ายผลปฏิสัมพันธ์สำหรับการรวมกันของแต่ละระดับขององค์ประกอบ A และ B ประมาณค่าได้ดังนี้

$$ab_{11} = \bar{Y}_{.11} - (\bar{Y}_{.1.} + \bar{Y}_{..1} - \bar{Y}_{...})$$

$$= 15.25 - (11.13 + 23.13 - 18.41) = -0.59$$

$$ab_{12} = \bar{Y}_{.12} - (\bar{Y}_{.1.} + \bar{Y}_{..2} - \bar{Y}_{...})$$

$$= 7.00 - (11.13 + 13.69 - 18.41) = +0.59$$

$$ab_{21} = \bar{Y}_{.21} - (\bar{Y}_{.2.} + \bar{Y}_{..1} - \bar{Y}_{...})$$

$= 22.75 - (17.88 + 23.13 - 18.41) = +0.16$

$ab_{22} = \bar{Y}_{.22} - (\bar{Y}_{.2.} + \bar{Y}_{..2} - \bar{Y}_{...})$

$= 13.00 - (17.88 + 13.69 - 18.41) = -0.16$

$ab_{31} = \bar{Y}_{.31} - (\bar{Y}_{.3.} + \bar{Y}_{..1} - \bar{Y}_{...})$

$= 26.25 - (20.25 + 23.13 - 18.41) = +1.28$

$ab_{32} = \bar{Y}_{.32} - (\bar{Y}_{.3.} + \bar{Y}_{..2} - \bar{Y}_{...})$

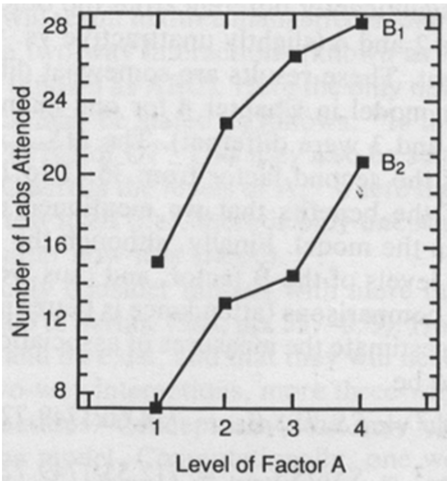
$= 14.25 - (20.25 + 13.69 - 18.41) = -1.28$

$ab_{41} = \bar{Y}_{.41} - (\bar{Y}_{.4.} + \bar{Y}_{..1} - \bar{Y}_{...})$

$= 28.25 - (24.38 + 23.13 - 18.41) = -0.84$

$ab_{42} = \bar{Y}_{.42} - (\bar{Y}_{.4.} + \bar{Y}_{..2} - \bar{Y}_{...})$

$= 20.50 - (24.38 + 13.69 - 18.41) = +0.84$



ภาพประกอบ 2 Profile Plot ของข้อมูลที่เป็นตัวอย่าง

ทำ profile plot แสดงในภาพประกอบ 2 อิทธิพลขององค์ประกอบ A มีนัยสำคัญและมีมากกว่า 2 ระดับ ต่อไปจะพิจารณาถึงวิธีการเปรียบเทียบพหุคูณด้วยการทดสอบ Tukey's HSD test ซึ่งเป็นการทดสอบแบบ family-wise ที่มีความเหมาะสมมากสำหรับการเปรียบเทียบเป็นรายคู่ที่มีรูปแบบสมดุล (มีจำนวนข้อมูลเท่ากันในแต่ละเซลล์) สามารถคำนวณได้ดังนี้

ค่าวิกฤติ (อ้างอิงจากตาราง)

$$_{1-\alpha}q_{df(errof,J)} = {}_{.95}q_{24,4} = 3.90$$

ความคลาดเคลื่อนมาตรฐาน

$$s_{\psi'} = \sqrt{[MS_{(error)}/n_j]}$$

$$= \sqrt{(11.53/8)}$$

$$= 1.20$$

สถิติทดสอบ

$$q_1 = (\bar{Y}_{.4.} - \bar{Y}_{.1.}) / s_{\psi'}$$
$$= (24.38 - 11.13)/1.20$$
$$= 11.04 \text{ (มีนัยสำคัญ)}$$
$$q_2 = (\bar{Y}_{.4.} - \bar{Y}_{.2.}) / s_{\psi'}$$
$$= (24.38 - 17.88)/1.20$$
$$= 5.41 \text{ (มีนัยสำคัญ)}$$
$$q_3 = (\bar{Y}_{.4.} - \bar{Y}_{.3.}) / s_{\psi'}$$
$$= (24.38 - 20.25)/1.20$$
$$= 3.44 \text{ (ไม่มีนัยสำคัญ)}$$
$$q_4 = (\bar{Y}_{.3.} - \bar{Y}_{.1.}) / s_{\psi'}$$
$$= (20.25 - 11.13)/1.20$$
$$= 4.60 \text{ (มีนัยสำคัญ)}$$
$$q_5 = (\bar{Y}_{.3.} - \bar{Y}_{.2.}) / s_{\psi'}$$
$$= (20.25 - 17.88)/1.20$$
$$= 1.98 \text{ (ไม่มีนัยสำคัญ)}$$
$$q_6 = (\bar{Y}_{.2.} - \bar{Y}_{.1.}) / s_{\psi'}$$
$$= (17.88 - 11.13)/1.20$$
$$= 5.62 \text{ (มีนัยสำคัญ)}$$

ผลบ่งชี้ว่าค่าเฉลี่ยสำหรับระดับขององค์ประกอบ A มีนัยสำคัญแตกต่างกันระหว่างระดับ 1 และ 4, 2 และ 4, 1 และ 3 และ 1 และ 2 ดังนั้นระดับที่ 1 ไม่ใช้ห้องแล็บในการสอนมีนัยสำคัญแตกต่างจากระดับอื่น ๆ อีก 3 ระดับ และระดับที่ 2 และ 4 (ใช้ห้องแล็บในการสอนกับใช้ห้องแล็บในการสอนมาก) มีนัยสำคัญแตกต่างกัน ผลของความแตกต่างที่พบจะแปลความหมายทำนองเดียวกับในเรื่องของการวิเคราะห์ความแปรปรวนแบบหนึ่งองค์ประกอบ และสุดท้ายอิทธิพลขององค์ประกอบ B มีนัยสำคัญ ซึ่ง B มีเพียงแค่ 2 ระดับเท่านั้นจึงไม่จำเป็นต้องเปรียบเทียบเป็นรายคู่

สุดท้ายจะประมาณค่าการวัดความสัมพันธ์ อัตราส่วนความสัมพันธ์ η^2 คำนวณได้ว่า

$$\eta_A^2 = SS_A / SS_{total} = 738.60/1749.72 = 0.42$$
$$\eta_B^2 = SS_B / SS_{total} = 712.53/1749.72 = 0.41$$
$$\eta_{AB}^2 = SS_{AB} / SS_{total} = 21.84/1749.72 = 0.01$$

คำนวณค่าสถิติ ω^2 ได้ดังนี้

$$\omega_A^2 = [SS_A - (J-1)MS_{with}]/[SS_{total} + MS_{with}]$$

= [738.60 - (3)11.53]/[1749.72 + 11.53] = 0.40

$$\omega_B^2 = [SS_B - (K-1)MS_{with}]/[SS_{total} + MS_{with}]$$

= [712.53 - (1)11.53]/[1749.72 + 11.53] = 0.40

$$\omega_{AB}^2 = [SS_{AB} - (J-1)(K-1)MS_{with}]/[SS_{total} + MS_{with}]$$

= [21.84 - (3)(1)11.53]/[1749.72 + 11.53] = -0.01

พื้นฐานของการวัดความสัมพันธ์และปราศจากความรู้ในงานวิจัยอื่น ๆ ที่ศึกษาทำนองเดียวกันนี้ อาจจะสรุปผลการวัดความสัมพันธ์ระหว่างปริมาณการใช้เวลาสอนแล็บของผู้สอนแต่ละคนสัมพันธ์กับระยะเวลาการใช้ห้องแล็บ และช่วงเวลาที่ใช้สอนแล็บมีความสัมพันธ์กับระยะเวลาการใช้ห้องแล็บ แต่ไม่มีความสัมพันธ์ระหว่างปฏิสัมพันธ์ของปริมาณการใช้เวลาสอนแล็บและช่วงเวลาการใช้ห้องแล็บกับระยะเวลาการใช้ห้องแล็บ

การวิเคราะห์ความแปรปรวนกรณีที่ n ไม่เท่ากัน

ในหัวข้อนี้จะกล่าวถึงกรณีที่ n แต่ละเซลล์ไม่เท่ากัน (รูปแบบไม่สมดุลย์) ซึ่งจะมีสูตรและสมการคำนวณค่าต่าง ๆ ที่แตกต่างกันออกไป

ประการแรกจะเป็นความแตกต่างของรูปแบบคือรูปแบบกรณีที่ n เท่ากันกับรูปแบบกรณีที่ใช้สัดส่วนของ n กรณีที่ n เท่ากันนั้นดังได้กล่าวแล้ว สำหรับกรณีใช้สัดส่วนของ n ในเซลล์ที่แต่ละเซลล์คือ

$$n_{jk} = n_j n_k / N$$

สำหรับทุก j และ k เมื่อ n_{jk} คือจำนวนของค่าสังเกตในเซลล์ jk แล้ว n_j คือจำนวนของค่าสังเกตในระดับของ j ขององค์ประกอบ A และ n_k คือจำนวนของค่าสังเกตในระดับ k ขององค์ประกอบ B และ N คือจำนวนค่าสังเกตรวมทั้งหมด ซึ่งตัวอย่างข้อมูลกรณี n เป็นสัดส่วนกันแสดงในตาราง 7 การคำนวณผลรวมกำลังสองสามารถทำได้ดังนี้

$$SS_{total} = \sum_i^n \sum_j^J \sum_k^K Y_{ijk}^2 - \left(\sum_i^n \sum_j^J \sum_k^K Y_{ijk} \right)^2 \Big/ N$$

$$SS_A = \sum_j^J \left[\left(\sum_i^n \sum_k^K Y_{ijk} \right)^2 \Big/ n_j \right] - \left(\sum_i^n \sum_j^J \sum_k^K Y_{ijk} \right)^2 \Big/ N$$

$$SS_B = \sum_k^K \left[\left(\sum_i^n \sum_j^J Y_{ijk} \right)^2 \Big/ n_k \right] - \left(\sum_i^n \sum_j^J \sum_k^K Y_{ijk} \right)^2 \Big/ N$$

$$SS_{AB} = \sum_j^J \sum_k^K \left[\left(\sum_i^n Y_{ijk} \right)^2 \Big/ n_{jk} \right] - \sum_j^J \left[\left(\sum_i^n \sum_k^K Y_{ijk} \right)^2 \Big/ n_j \right]$$

$$= -\sum_{k}^{K}\left[\left(\sum_{i}^{n}\sum_{j}^{J}Y_{ijk}\right)^2 \Big/ n_k\right] - \left(\sum_{i}^{n}\sum_{j}^{J}\sum_{k}^{K}Y_{ijk}\right)^2 \Big/ N$$

$$SS_{with} = \sum_{i}^{n}\sum_{j}^{J}\sum_{k}^{K}Y_{ijk}^2 - \sum_{j}^{J}\sum_{k}^{K}\left[\left(\sum_{i}^{n}Y_{ijk}\right)^2 \Big/ n_{jk}\right]$$

เมื่อผลรวมของ n มีค่าตั้งแต่ i = 1, …, n_{jk}

ตาราง 7 ตัวอย่างข้อมูลที่ n เป็นสัดส่วนกัน

ตัวอย่างที่ 1					
		b_1	b_2	b_3	
	a_1	$n_{11} = 5$	$n_{12} = 5$	$n_{13} = 5$	$n_{1.} = 15$
	a_2	$n_{21} = 7$	$n_{22} = 7$	$n_{23} = 7$	$n_{2.} = 21$
		$n_{.1} = 12$	$n_{.2} = 12$	$n_{.3} = 12$	$N = 36$
ตัวอย่างที่ 2					
		b_1	b_2	b_3	
	a_1	$n_{11} = 4$	$n_{12} = 6$	$n_{13} = 8$	$n_{1.} = 18$
	a_2	$n_{21} = 4$	$n_{22} = 6$	$n_{23} = 8$	$n_{2.} = 18$
		$n_{.1} = 8$	$n_{.2} = 12$	$n_{.3} = 16$	$N = 36$
ตัวอย่างที่ 3					
		b_1	b_2	b_3	
	a_1	$n_{11} = 2$	$n_{12} = 4$	$n_{13} = 6$	$n_{1.} = 12$
	a_2	$n_{21} = 4$	$n_{22} = 8$	$n_{23} = 12$	$n_{2.} = 24$
		$n_{.1} = 6$	$n_{.2} = 12$	$n_{.3} = 18$	$N = 36$

กรณีที่ n ไม่เป็นสัดส่วนกันมีโอกาสเกิดขึ้นได้เมื่อ n ไม่เท่ากันในแต่ละกลุ่มและไม่สามารถกำหนดเป็นสัดส่วนได้ จึงอาจเป็นเหตุให้อิทธิพลหลักและผลปฏิสัมพันธ์ไม่เป็นอิสระจากกัน ในอีกกรณีหนึ่ง ผลรวมกำลังสองไม่สามารถจะแบ่งส่วนอิทธิพลได้อย่างอิสระจากกัน นำไปสู่ความพยายามใช้วิธีการที่หลากหลายในการแบ่งส่วนซึ่งไม่สามารถแบ่งส่วนได้ แต่มีวิธีอยู่ 3 วิธีที่ช่วยแก้ไขปัญหากรณี n ไม่เป็นสัดส่วนกัน ในแต่ละวิธีจะให้การทดสอบสมมติฐานที่แตกต่างกันออกไป วิธีโดยสรุปมีดังนี้ 1) the sequential approach หรือ the hierarchical sums of squares approach 2) the partially sequential approach หรือ the partially heirarchical หรือ

experimental design หรือ method of fitting constants approach และ 3) the regression approach หรือ the marginal means approach

สรุป

ในเนื้อหานี้จะเป็นวิธีเกี่ยวข้องกับการวิเคราะห์ความแปรปรวนแบบสององค์ประกอบ ซึ่งจะกล่าวถึงลักษณะสำคัญของการวิเคราะห์ รวมถึง 1) มีตัวแปรอิสระ 2 ตัวหรือมากกว่า และแต่ละตัวแปรต้องเป็นแบบกำหนด 2) กลุ่มตัวอย่างถูกสุ่มเข้าในแต่ละระดับของตัวแปรอิสระ 3) องค์ประกอบจะไขว้กันอย่างสมบูรณ์ในทุกระดับขององค์ประกอบิสระ และ 4) ตัวแปรตามถูกวัดอย่างน้อยในระดับช่วง (interval scale) โมเดล ANOVA ถูกตรวจสอบอิทธิพลหลัก และผลปฏิสัมพันธ์ ข้อตกลงเบื้องต้นของ ANOVA ผลของการละเมิดข้อตกลงเบื้องต้น ตารางสรุป ANOVA สูตรสำหรับคำนวณผลรวมกำลังสอง ค่าเฉลี่ยกำลังสอง และการทดสอบนัยสำคัญด้วยสถิติ F

แปลและเรียบเรียงจาก

Lomax, Richard G. (1992). **Statistical Concepts : A Second Course for Education and the Behavioral Sciences**. London : Lawrence Erlbaum Associates, Inc.