บทที่ 13
การวิเคราะห์ความแปรปรวนร่วม
สำหรับการออกแบบการทดลองแบบสุ่มสมบูรณ์

1. แนวคิดเกี่ยวกับการวิเคราะห์ความแปรปรวนร่วม

การวิเคราะห์ความแปรปรวนร่วม คือ การผสมผสานของการวิเคราะห์การถดถอยกับการวิเคราะห์ความแปรปรวน เราใช้ความแปรปรวนร่วมเมื่อตัวแปรตาม Y นอกจากได้รับอิทธิพลจากทรีทเมนต์แล้วยังมีความสัมพันธ์เชิงเส้นกับตัวแปรร่วม (covariates) ตัวอื่น ๆ ด้วยในบทนี้มีขอบเขตเฉพาะการถดถอยเชิงเส้นอย่างง่าย คือ ตัวแปรตาม Y ถดถอยบนตัวแปรอธิบายตัวเดียว และ Y ยังมีความสัมพันธ์เชิงเส้นกับตัวแปรร่วมอื่นอีก 1 ตัวแปร ซึ่งต้องอาศัยการวิเคราะห์ความแปรปรวนแบบทางเดียว

การวิเคราะห์ความแปรปรวนร่วมมีวัตถุประสงค์ที่สำคัญคือ

1. เพื่อเพิ่มความถูกต้องแม่นยำในการทดลอง
2. เพื่อควบคุมตัวแปรร่วมอื่น ๆ ในการทดลอง
3. เพื่อเปรียบเทียบการถดถอยภายในกลุ่มต่าง ๆ

1.1 การเพิ่มความถูกต้องแม่นยำในการทดลอง

ตัวอย่างเช่นการศึกษาเปรียบเทียบวิธีการลดน้ำหนัก 3 วิธี การใช้การวิเคราะห์ความแปรปรวนอาจไม่สามารถตัดสินใจได้ว่ามีความแตกต่างของวิธีลดน้ำหนักทั้ง 3 วิธีนี้ เพราะความแปรปรวนภายในแต่ละวิธีการลดน้ำหนักมีขนาดใหญ่มาก ซึ่งอาจเนื่องมาจากน้ำหนักที่ลดลง (Y) มีความสัมพันธ์เชิงเส้นกับตัวแปรร่วมตัวอื่นนอกจากวิธีการลดน้ำหนัก เช่น น้ำหนักตอนเริ่มการทดลองของตัวอย่าง กรณีตัวอย่างนี้เราสามารถลดความแปรปรวนภายในแต่ละวิธีการลดน้ำหนักได้โดยการวิเคราะห์ความแปรปรวนร่วมซึ่งเป็นการผสมผสานของการวิเคราะห์การถดถอยของ Y

บน X กับการวิเคราะห์ความแปรปรวนบน Y ในกรณีนี้อาจใช้เทคนิคการบล็อกน้ำหนักตอนเริ่มต้นทดลองก็ได้

1.2 การควบคุมตัวแปรร่วมอื่น ๆ ในการทดลอง

ตัวอย่างเช่น การศึกษาเงินเดือนครู (Y) ในโรงเรียนต่าง ๆ ที่มีระบบแตกต่างกัน 3 ระบบ เป็นทรีทเมนต์ โดยมีระดับวุฒิการศึกษาของครูเป็นตัวแปรร่วม (X) ถ้า Y มีความสัมพันธ์เชิงเส้นกับ X เราสามารถใช้การวิเคราะห์ความแปรปรวนร่วม ซึ่งเป็นการแปลงข้อมูลแต่ละตัว (x_{ij} , y_{ij}) ด้วยระดับการศึกษาเฉลี่ยของครูทุกคนในการทดลอง ได้เป็นเงินเดือนที่ปรับแล้วของครูคนที่ j ที่มีระดับการศึกษาเฉลี่ย ($\bar{x}..$, y'_{ij}) และจะทำให้ค่าเฉลี่ยของกลุ่มทั้ง 3 กลุ่มถูกแปลงด้วย

2. เกี่ยวกับทฤษฎีการวิเคราะห์ความแปรปรวนร่วม

ข้อกำหนดในการวิเคราะห์ความแปรปรวนร่วมคือ

(1) ตัวแปรร่วม (covariate) X เป็นค่าที่สังเกตได้ไม่ใช่อิทธิพลของทรีทเมนต์ ส่วนตัวแปรตาม Y เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติ

(2) ความคลาดเคลื่อนของการทดลอง (e_{ij}) เกิดขึ้นโดยสุ่มมีความเป็นอิสระกันและมีการแจกแจงแบบปกติ โดยมีค่าเฉลี่ยเป็น 0 และมีค่าความแปรปรวนร่วมกัน

(3) การถดถอยของ Y บน X หลังจากหักอิทธิพลอื่น ๆ ในการวิเคราะห์ความแปร ปรวนออกแล้วต้องเป็นเส้นตรง

(4) เส้นการถดถอยทุกกลุ่มของทรีทเมนต์มีความชันเท่ากัน

ตัวแบบสถิติของการวิเคราะห์ความแปรปรวนร่วมสำหรับการออกแบบการทดลองแบบสุ่มสมบูรณ์คือ

$$y_{ij} = \mu + \alpha_i + \beta(x_{ij} - \bar{x}..) + e_{ij} \qquad (1)$$

$i = 1, 2, \ldots, a \;\; ; \;\; j = 1, 2, \ldots, n_i$

เมื่อ μ คือ ค่าเฉลี่ยทั้งหมด

α_i คือ อิทธิพลของทรีทเมนต์ที่ i

β คือ ความชันของเส้นถดถอย

$\bar{x}..$ คือ ค่าเฉลี่ยทั้งหมดของตัวแปรร่วมหรือตัวแปรภายนอก

e_{ij} คือ ความคลาดเคลื่อนสุ่มของสมาชิกตัวที่ j ในทรีทเมนต์ที่ i

ซึ่งอาจเขียน (1) ได้ใหม่ คือ

$$y_{ij} - \beta(x_{ij} - \bar{x}..) = \mu + \alpha_i + e_{ij} \quad (2)$$

หรือ $$y_{ij} - \alpha_i = \mu + \beta(x_{ij} - \bar{x}..) + e_{ij} \quad (3)$$

ถ้าเขียนตัวแบบสถิติในรูป (2) หมายความว่า ผู้วิจัยต้องการวิเคราะห์ความแปรปรวนของ Y ที่ได้รับการปรับเอาอิทธิพลของตัวแปรร่วม X ออกแล้ว

ถ้าเขียนตัวแบบสถิติในรูป (3) หมายความว่า ผู้วิจัยต้องการวิเคราะห์การถดถอยโดยหาการวิเคราะห์ การถดถอยเชิงเส้นของ Y บน X เมื่อหักอิทธิพลของทรีทเมนต์ออกแล้ว

3. การตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเท่ากันของความชันของเส้นถดถอย (test of homogeneity of slopes)

3.1 การตรวจสอบเกี่ยวกับความเท่ากันของความชันของเส้นถดถอยทุกกลุ่มของทรีทเมนต์

จากตัวแบบสถิติของการวิเคราะห์ความแปรปรวนร่วมสำหรับการออกแบบการทดลองแบบสุ่มสมบูรณ์คือ

$$y_{ij} = \mu + \alpha_i + \beta(x_{ij} - \bar{x}..) + e_{ij}$$

อาจเขียนใหม่เรียกว่า full model สมมติให้เป็น model 1 คือ

$$y_{ij} = \mu_i + \beta(x_{ij} - \bar{x}..) + e_{ij}$$

เมื่อ μ_i คือ ค่าเฉลี่ยของทรีทเมนต์กลุ่ม i

β คือ ความชันของเส้นถดถอยของทรีทเมนต์ทุกกลุ่ม

และให้ตัวแบบการถดถอยเชิงเส้นตรงของทรีทเมนต์แต่ละกลุ่มเป็น model 2 คือ

$$y_{ij} = \mu_i + \beta_i(x_{ij} - \bar{x}_{i.}) + e_{ij}$$

; $i = 1, 2, \ldots, a$; $j = 1, 2, \ldots, n$

เมื่อ β_i คือ สัมประสิทธิ์การถดถอยของทรีทเมนต์กลุ่ม i สมมติฐานทางสถิติที่ต้องการทดสอบ คือ $H_0 : \beta_1 = \beta_2 = \ldots = \beta_a$ คู่กับ H_1 : มี β_i อย่างน้อย 1 ตัวที่ไม่เท่ากับตัวอื่น ๆ การคำนวณผลบวกกำลังสอง (sum of square) ของการทดสอบเกี่ยวกับความเท่ากันของความชันของเส้นถดถอยทุกกลุ่มของทรีทเมนต์มีสูตรคือ

$$SS_{homogeneity} = SSE_1 - SSE_2$$

ที่มี $df = (N - a - 1) - (N - 2a) = (a - 1)$

เมื่อ SSE_1 คือ sum of square ของความคลาดเคลื่อนจาก model 1 ที่มี $df = (N - a - 1)$

SSE_2 คือ sum of square ของความคลาดเคลื่อนจาก model 2 ที่มี $df = (N - 2a)$

การคำนวณ SS homogeneity โดยใช้โปรแกรม SPSS ช่วยในการคำนวณ จะอยู่ในเทอมของอิทธิพลร่วมระหว่างทรีทเมนต์กับตัวแปรร่วม

สถิติทดสอบคือ

$$F_0 = \frac{MS\ homogeneity}{MSE_2}$$

ซึ่งมีการแจกแจงแบบ F ที่มีจำนวนชั้นอิสระ $(a - 1)$ และ $(N - 2a)$

ตัวอย่างเช่น การศึกษาเปรียบเทียบสื่อโฆษณา 3 แบบ ของร้านอาหารจานด่วนคือ สื่อทางวิทยุ สื่อทางหนังสือพิมพ์ และสื่อทางโทรทัศน์ กลุ่มตัวอย่างคือ ร้านอาหารที่มีเครื่องหมายการค้าเดียวกัน 15 ร้าน ตั้งอยู่ในที่ต่าง ๆ กัน แต่ตั้งอยู่ในเมืองขนาดเดียวกัน ออกแบบการทดลองแบบสุ่มสมบูรณ์โดยสุ่มร้านอาหารให้ได้รับทรีทเมนต์คือสื่อโฆษณาแบบที่หนึ่ง 5 ร้าน แบบที่สอง 5 ร้าน และแบบที่สาม 5 ร้าน กำหนดช่วงเวลาการโฆษณาด้วยสื่อทั้ง 3 แบบ ในช่วงเวลาเดียวกัน แล้วเก็บข้อมูลเป็นรายได้ (y) หน่วยเป็น 50,000 บาท ถึงแม้ว่าร้านอาจจะมีขนาดเท่ากัน แต่มีจำนวนคนงานไม่เท่ากัน ซึ่งจำนวนคนงานอาจมีความสัมพันธ์กับรายได้ จึงให้จำนวนคนงานเป็นตัวแปรร่วม (x)

ตารางที่ 13.1 ข้อมูลรายได้ของร้านอาหารที่ใช้สื่อโฆษณาแตกต่างกัน 3 แบบ

สื่อโฆษณา					
วิทยุ		หนังสือพิมพ์		โทรทัศน์	
x	y	x	y	x	y
10	30	21	24	34	17
14	18	26	20	39	11
19	13	31	7	43	3
25	6	36	4	47	-6
27	3	41	-5	52	-10

กำหนดให้ตัวแปร treat และ group แทนทรีทเมนต์คือ สื่อโฆษณา 3 แบบ (1 = วิทยุ , 2 = หนังสือพิมพ์ , 3 = โทรทัศน์) ให้ตัวแปร employee แทนจำนวนคนงานเป็นตัวแปรร่วม และตัวแปร money แทนรายได้เป็นตัวแปรตาม

สมการเส้นถดถอยของรายได้ของร้านอาหารที่ถดถอยบนสื่อโฆษณาคือ

$$(\text{รายได้})_{ij} = \mu_i + \beta_i(x_{ij} - \bar{x}_{i.}) + e_{ij}$$

i = สื่อวิทยุ , สื่อหนังสือพิมพ์ , สื่อโทรทัศน์ ; j = ร้านที่ 1 , ร้านที่ 2 , ร้านที่ 3 , ร้านที่ 4 , ร้านที่ 5

ต้องการตรวจสอบว่าความชันของเส้นถดถอยทั้ง 3 สื่อโฆษณาเท่ากันหรือไม่ สมมติฐานทางสถิติที่ต้องการทดสอบคือ $H_0 : \beta_{\text{วิทยุ}} = \beta_{\text{หนังสือพิมพ์}} = \beta_{\text{โทรทัศน์}}$ คู่กับ H_1 : มีอย่างน้อย 1 ค่าที่ไม่เท่ากับค่าอื่น ๆ

ถ้าตัวแปรร่วมคือจำนวนคนงานในร้านอาหารมีอิทธิพลต่อรายได้ของร้านอาหารที่ใช้สื่อโฆษณา 3 แบบ แตกต่างกันแม้แต่เพียงแบบเดียวจะทำให้ตัวแบบการถดถอยของข้อมูลทั้งหมดอาจมีความผิดพลาดเพราะตัวแบบการถดถอยไม่สามารถอธิบายรูปแบบของความสัมพันธ์ของทุกกลุ่มทรีทเมนต์คือ สื่อโฆษณาทั้ง 3 แบบได้ ข้อตกลงเบื้องต้นข้อนี้มีความสำคัญมาก เรียกว่า ข้อตกลงเบื้องต้นเกี่ยวกับความเท่ากันของความชันของเส้นถดถอย (homogeneity of regression slopes)

วิธีที่สามารถทำให้เห็นเกี่ยวกับข้อตกลงเบื้องต้นนี้ได้คือ การพล็อตกราฟการกระจายของข้อมูลที่มีแกนหนึ่งคือ ตัวแปรร่วมได้แก่ จำนวนคนงานในร้านอาหาร และอีกแกนหนึ่งคือ ตัวแปรตามได้แก่ รายได้ของร้านอาหาร แล้วคำนวณสมการถดถอยของแต่ละทรีทเมนต์ ได้แก่ สื่อโฆษณา 3 แบบ ที่เหมาะสมกับข้อมูลที่มีการกระจายเหล่านี้ ซึ่งจะทำให้เรามองเห็นได้ว่าเส้นถดถอยทั้ง 3 เส้น เหมือนกันมากน้อยเพียงใด นั่นคือค่าของ β ของแต่ละกลุ่มตัวอย่างควรจะเท่ากัน เราสามารถใช้โปรแกรม SPSS ช่วยในการพล็อตกราฟได้โดยใช้คำสั่ง **Graphs , Scatter…**

3.2 การใช้คำสั่ง Graphs, Scatter….

เพื่อดูรูปแบบความสัมพันธ์ระหว่างตัวแปรร่วมกับตัวแปรตามของทรีทเมนต์แต่ละกลุ่ม

จากตัวอย่างการศึกษาเปรียบเทียบสื่อโฆษณา 3 แบบของร้านอาหารทำการบันทึกข้อมูลลงในแฟ้มข้อมูล ancova1.sav มีรูปแบบดังนี้

ตารางที่ 13.2 รูปแบบการบันทึกข้อมูลรายได้ของร้านอาหารที่ใช้สื่อโฆษณาแตกต่างกัน 3 แบบ

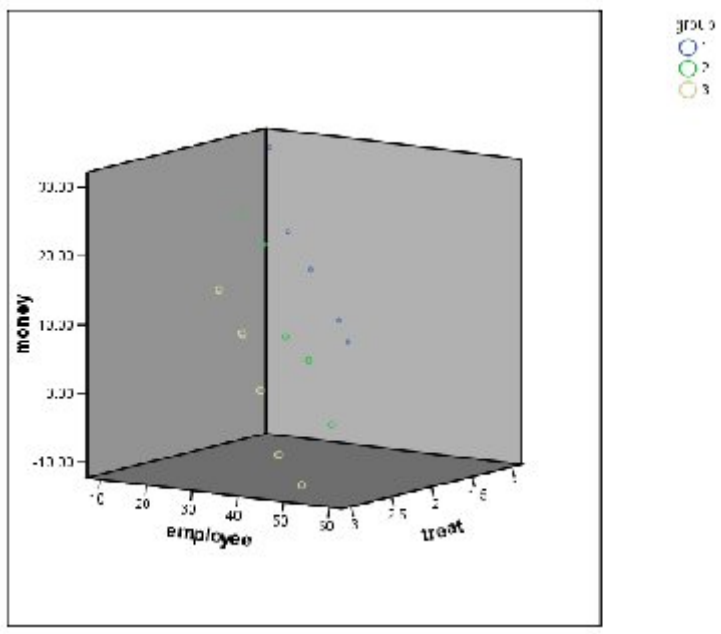
treat	group	employee	money	treat	group	employee	money
1	1	10	30	2	2	36	4
1	1	14	18	2	2	41	-5
1	1	19	13	3	3	34	17
1	1	25	6	3	3	39	11
1	1	27	3	3	3	43	3
2	2	21	24	3	3	47	-6
2	2	26	20	3	3	52	-10
2	2	31	7				

ขั้นตอนการใช้คำสั่งคือ

1. ไปที่เมนูบาร์ คลิกที่ Graphs, Scatter… จะได้หน้าต่าง Scatterplot
2. ในหน้าต่าง Scaterplot
   คลิกที่คำสั่ง 3-D แล้วกดปุ่ม Define จะได้หน้าต่าง 3-D Scatterplot
3. ในหน้าต่าง 3-D Scatterplot

คลิกที่ตัวแปร money ให้ย้ายเข้าไปอยู่ในช่อง Y Axis :
คลิกที่ตัวแปร employee ให้ย้ายเข้าไปอยู่ในช่อง X Axis:
คลิกที่ตัวแปร treat ให้ย้ายเข้าไปอยู่ในช่อง Z Axis:
คลิกที่ตัวแปร group ให้ย้ายเข้าไปอยู่ในช่อง Set Markers by:
แล้วคลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.1

Graph



ภาพที่ 13.1

ภาพที่ 13.1 แสดงภาพการกระจายของข้อมูลแบบ 3 มิติ ที่แสดงรูปแบบของความสัมพันธ์ระหว่างจำนวนคนงานในร้านอาหาร (ตัวแปรร่วม) กับรายได้ของร้านอาหาร (ตัวแปรตาม) ของกลุ่มตัวอย่าง 3 กลุ่ม โดยที่กลุ่ม 1 คือร้านอาหารที่ใช้สื่อโฆษณาวิทยุ กลุ่ม 2 คือร้านอาหารที่ใช้สื่อโฆษณาหนังสือพิมพ์ และกลุ่ม 3 คือ ร้านอาหารที่ใช้สื่อโฆษณาโทรทัศน์ แต่ละกลุ่มมี 5 จุด ในภาพไม่ได้แสดงเส้นถดถอยของทั้ง 3 กลุ่ม แต่เราก็สามารถจินตนาการได้ง่าย ๆ

ทำให้เห็นได้ว่าเส้นถดถอยของทั้ง 3 กลุ่มมีความชันเป็นบวก และเส้นถดถอยทั้ง 3 เส้นนี้มีความชันพอ ๆ กัน แต่ก็ยังไม่สามารถบอกได้ว่าเส้นถดถอยทั้ง 3 เส้นนี้มีความชันเท่ากันหรือไม่

3.3 การใช้คำสั่ง General Linear Model แบบ Univariate

เพื่อทดสอบข้อตกลงเบื้องต้นเกี่ยวกับความเท่ากันของความชันของเส้นถดถอย

1. ไปที่เมนูบาร์ คลิกที่ Analyze, General Linear Model, Univariate … จะได้หน้าต่าง Univariate

2. ในหน้าต่าง Univariate คลิกที่ตัวแปรตาม money ให้ย้ายเข้าไปอยู่ในช่อง Dependent Variable : และคลิกที่ตัวแปร treat ให้ย้ายเข้าไปอยู่ในช่อง Fixed Factor(s) : และคลิกที่ตัวแปรร่วม employee ให้ย้ายเข้าไปอยู่ในช่อง Covariate(s) แล้วคลิกที่ปุ่ม Model … จะได้หน้าต่าง Univariate Model

3. ในหน้าต่าง Univariate Model

ในกรอบ Specify Model เลือก O Custom

ในกรอบ Factors & Covariates : เลือกตัวแปร treat ในกรอบ Build Term(s) เลือก Main effects แล้วคลิกที่ปุ่มลูกศร ➢ให้ตัวแปร treat ย้ายเข้าไปในกรอบของ Model ทำเช่นเดียวกันกับตัวแปร employee เพื่อให้ทั้งตัวแปร treat และ employee เป็นอิทธิพลหลักในตัวแบบสถิติ แล้วเลือกตัวแปร treat และ employee พร้อมกัน ในกรอบ Build Term(s) เลือก Interaction แล้วคลิกที่ปุ่มลูกศร ➢ จะได้อิทธิพลร่วมของ treat * employee ย้ายเข้าไปอยู่ในกรอบของ Model

แล้วคลิกที่ปุ่ม Continue หน้าต่างนี้จะถูกปิดไป

4. ในหน้าต่าง Univariate คลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.2

Univariate Analysis of Variance

Between-Subjects Factors

		N
treat	1	5
	2	5
	3	5

Tests of Between-Subjects Effects

Dependent Variable: money

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1793.055[a]	5	358.611	63.353	.000
Intercept	1826.401	1	1826.401	322.654	.000
treat	93.164	2	46.582	8.229	.009
employee	1470.924	1	1470.924	259.855	.000
treat * employee	2.500	2	1.250	.221	.806
Error	50.945	9	5.661		
Total	3059.000	15			
Corrected Total	1844.000	14			

a. R Squared = .972 (Adjusted R Squared = .957)

ภาพที่ 13.2 ผลลัพธ์การตรวจสอบเกี่ยวกับความเท่ากันของความชันของเส้นถดถอยทุกกลุ่มของทรีทเมนต์

ผลลัพธ์ที่ได้คือ ตาราง Tests of Between Subjects Effects ดูที่แถวของเทอมปฏิสัมพันธ์ treat*employee ได้ค่าสถิติทดสอบ F เท่ากับ .221 และค่า Sig. เท่ากับ .806 ซึ่งมากกว่าระดับนัยสำคัญที่กำหนด .05 ดังนั้นจึงสรุปว่าไม่สามารถปฏิเสธ $H_0 : \beta_{วิทยุ} = \beta_{หนังสือพิมพ์} = \beta_{โทรทัศน์}$ นั่นคือ ความชันของเส้นแสดงรูปแบบความสัมพันธ์ระหว่างตัวแปรตาม money กับตัวแปรร่วม employee ไม่แตกต่างกัน เมื่อมีระดับของตัวแปรอิสระ treat แตกต่างกันหรือเส้นตรงที่แสดงความสัมพันธ์ระหว่าง money กับ employeeขนานกันเมื่อใช้สื่อโฆษณาแตกต่างกัน หมายความว่าความสัมพันธ์ระหว่างรายได้กับจำนวนคนงานของร้านอาหารที่ใช้สื่อโฆษณาแตกต่างกัน มีความสัมพันธ์แตกต่างกันอย่างไม่มีนัยสำคัญคือ ไม่แตกต่างกัน จึงทำการวิเคราะห์ความแปรปรวนร่วมต่อไป

4. การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของทรีทเมนต์ที่ได้ปรับอิทธิพลของตัวแปรร่วม X ออกไปแล้ว

การวิเคราะห์ความแปรปรวนร่วมในการออกแบบการทดลองแบบสุ่มสมบูรณ์ หรือการวิเคราะห์ความแปรปรวนร่วมทางเดียว (one way analysis of covariance) ต้องการทดสอบความแตกต่างระหว่างอิทธิพลของทรีทเมนต์ต่าง ๆ ซึ่งเป็นการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยของทรีทเมนต์ที่ได้ปรับอิทธิพลของตัวแปรร่วม X ออกไปแล้ว ซึ่งเรียกว่า adjusted treatment means

จากตัวแบบสถิติ $y_{ij} = \mu_i + \beta(x_{ij} - \bar{x}..) + e_{ij}$

สมมติฐานทางสถิติที่ต้องการทดสอบคือ

(1) ทดสอบนัยสำคัญของอิทธิพลของตัวแปรร่วม employee ที่มีต่อตัวแปรตาม money สมมติฐานทางสถิติที่ต้องการทดสอบคือ

$$H_0 : \beta = 0 \quad \text{คู่กับ} \quad H_1 : \beta \neq 0$$

สถิติทดสอบคือ

$$F_0 = \frac{\text{MS employee}}{\text{MSE}}$$

ซึ่งมีการแจกแจงแบบ F ที่มีจำนวนชั้นอิสระ $(a-1)$ และ $N\text{-}a\text{-}1$ เขตวิกฤตคือ $F_{\alpha,\,(a-1),\,(N\text{-}a\text{-}1)}$

(2) ทดสอบนัยสำคัญของอิทธิพลของทรีทเมนต์ภายหลังการปรับเอาอิทธิพลของตัวแปรร่วม employee ออกแล้ว สมมติฐานทางสถิติที่ต้องการทดสอบคือ

$H_0 : \mu_1 = \mu_2 = \ldots = \mu_a$ คู่กับ H_1 : มีอย่างน้อย 1 ค่าที่ไม่เท่ากับค่าอื่น

สถิติทดสอบคือ

$$F_0 = \frac{\text{MS treat (adjusted)}}{\text{MSE}}$$

ซึ่งมีการแจกแจงแบบ F ที่มีจำนวนชั้นอิสระ $(a-1)$ และ $N\text{-}a\text{-}1$ เขตวิกฤตคือ $F_{\alpha,\,(a-1),\,(N\text{-}a\text{-}1)}$

จากตัวอย่างการศึกษาเปรียบเทียบสื่อโฆษณา 3 แบบ ของร้านอาหารจานด่วน สมมติฐานทางสถิติที่ต้องการทดสอบคือ H_0 : $\mu_{วิทยุ} = \mu_{หนังสือพิมพ์} = \mu_{โทรทัศน์}$ คู่กับ H_1 : มีอย่างน้อย 1 ค่า ที่ไม่เท่ากับค่าอื่น

ผู้วิจัยต้องการทดสอบสมมติฐานเกี่ยวกับอิทธิพลของตัวแปรร่วม employee ที่มีต่อตัวแปรตาม money และทดสอบสมมติฐานเกี่ยวกับอิทธิพลของทรีทเมนต์ภายหลังการปรับเอาอิทธิพลของตัวแปรร่วม employee ออกแล้ว มีขั้นตอนดังนี้

ขั้นตอนการวิเคราะห์ความแปรปรวนร่วมโดยใช้คำสั่ง **General Linear Model** แบบ **Univariate** คือ

1. ไปที่เมนูบาร์ คลิกที่ Analyze , General Linear Model , Univariate… จะได้หน้าต่าง Univariate

2. ในหน้าต่าง Univariate คลิกที่ตัวแปรตาม money ให้ย้ายเข้าไปอยู่ในช่อง Dependent Variable : และคลิกที่ตัวแปร treat ให้ย้ายเข้าไปอยู่ในช่อง Fixed Factor(s) : และคลิกที่ตัวแปรร่วม employee ให้ย้ายเข้าไปอยู่ในช่อง Covariate(s)

   คลิกปุ่ม Model … จะได้หน้าต่าง Univariate Model

   คลิกปุ่ม Option… จะได้หน้าต่าง Univariate : Options

3. ในหน้าต่าง Univariate Model

   ในกรอบ Specify Model เลือก ○ Full factorial

   แล้วคลิกปุ่ม Continue หน้าต่างนี้จะถูกกปิดไป

4. ในหน้าต่าง Univariate Options

เพื่อทำ post hoc test เป็นการทดสอบความแตกต่างของค่าเฉลี่ยของประชากรเป็นรายคู่ ตามวิธีของฟิชเชอร์ (LSD) , Bonferroni , และ Sidak การเลือกใช้วิธีใดเราพิจารณาจาก 3 สิ่งคือ (1) สามารถควบคุมความคลาดเคลื่อนประเภท I ให้เป็นไปตามที่กำหนด (2) สามารถควบคุมความคลาดเคลื่อนประเภท II ให้เป็นไปตามที่กำหนด และ (3) มีความเที่ยง (reliable) เมื่อไม่เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวน

สำหรับวิธี LSD เป็นการเปรียบเทียบความแตกต่างของทุกกลุ่มโดยเลือกมาทีละคู่แล้วใช้การทดสอบ t-test ซึ่งมักเกิดปัญหาเกี่ยวกับความคลาดเคลื่อนประเภท I ไม่เป็นไปตามที่กำหนด จึงมักนิยมใช้วิธี Bonferroni ที่สามารถควบคุมความคลาดเคลื่อนประเภท I ได้โดยการแบ่งความคลาดเคลื่อนประเภท I ออกเป็นส่วนเท่า ๆ กัน ตามจำนวนการเปรียบเทียบความ

แตกต่างของกลุ่มตัวอย่างเป็นรายคู่ ตัวอย่างเช่น กำหนดความคลาดเคลื่อนประเภท I เท่ากับ 0.05 ถ้ามีการทดสอบ 10 คู่ เราจะให้แต่ละการทดสอบมีความคลาดเคลื่อนประเภท I เท่ากับ 0.05/10 = 0.005 สำหรับวิธี Sidak คล้ายกับวิธี Bonferroni เราจะเลือกวิธีนี้ถ้าคำนึงถึงความคลาดเคลื่อนประเภท II ซึ่ง Bonferroni ไม่สามารถควบคุมความคลาดเคลื่อนประเภท II นี้ได้ (Field, A., 2000)

ขั้นตอนการใช้คำสั่งคือ

ในกรอบ Estimated Marginal Means

- คลิกที่ตัวแปรอิสระ treat ในช่อง Factor(s) and Factor Interactions: ให้ย้ายเข้าไปอยู่ในช่อง Display Means for :

- คลิกเลือก ☐ Compare main effects จะทำให้คำสั่ง Confidence interval adjustment: ทำงานได้ (active) ซึ่งภายใต้คำสั่งนี้จะมีวิธีของ post hoc test ให้เลือก 3 วิธี คือ LSD , Bonferroni , และ Sidak

สำหรับตัวอย่างนี้เลือกวิธี Sidak

ในกรอบ Display

- คลิกเลือก ☐ Descriptive statistics จะได้ผลลัพธ์เป็นค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างแต่ละกลุ่ม

- คลิกเลือก ☐ Parameter estimates จะได้ผลลัพธ์เป็นค่าสัมประสิทธิ์การถดถอย และ การทดสอบนัยสำคัญของตัวแปรที่อยู่ในตัวแบบการถดถอย

- คลิกเลือก ☐ Homogeneity tests จะได้ผลลัพธ์เป็น Levene's test ซึ่งเป็นการตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ของความแปรปรวนของประชากรทุกกลุ่ม

คลิกที่ปุ่ม Continue หน้าต่างนี้จะถูกปิดไป

5. ในหน้าต่าง Univariate

คลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.3

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
treat	1	radio	5
	2	newspaper	5
	3	tv	5

Descriptive Statistics

Dependent Variable: money

treat	Mean	Std. Deviation	N
radio	14.0000	10.70047	5
newspaper	10.0000	11.89538	5
tv	3.0000	11.29159	5
Total	9.0000	11.47668	15

Levene's Test of Equality of Error Variances [a]

Dependent Variable: money

F	df1	df2	Sig.
.387	2	12	.687

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+employee+treat

Tests of Between-Subjects Effects

Dependent Variable: money

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1790.555[a]	3	596.852	122.844	.000
Intercept	2014.134	1	2014.134	414.550	.000
employee	1480.555	1	1480.555	304.729	.000
treat	502.096	2	251.048	51.671	.000
Error	53.445	11	4.859		
Total	3059.000	15			
Corrected Total	1844.000	14			

a. R Squared = .971 (Adjusted R Squared = .963)

Parameter Estimates

Dependent Variable: money

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	67.897	3.846	17.65	.000	59.432	76.362
employee	-1.509	.086	-17.5	.000	-1.700	-1.319
[treat=1]	-25.222	2.500	-10.1	.000	-30.724	-19.720
[treat=2]	-11.111	1.738	-6.394	.000	-14.936	-7.286
[treat=3]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Estimated Marginal Means
treat

Estimates

Dependent Variable: money

treat	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
radio	-4.111[a]	1.431	-7.261	-.961
newspaper	10.000[a]	.986	7.830	12.170
tv	21.111[a]	1.431	17.961	24.261

a. Covariates appearing in the model are evaluated at the following values: employee = 31.00.

Pairwise Comparisons

Dependent Variable: money

(I) treat	(J) treat	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
radio	newspaper	-14.111*	1.738	.000	-18.995	-9.227
	tv	-25.222*	2.500	.000	-32.247	-18.196
newspaper	radio	14.111*	1.738	.000	9.227	18.995
	tv	-11.111*	1.738	.000	-15.995	-6.227
tv	radio	25.222*	2.500	.000	18.196	32.247
	newspaper	11.111*	1.738	.000	6.227	15.995

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Sidak.

Univariate Tests

Dependent Variable: money

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	502.096	2	251.048	51.671	.000
Error	53.445	11	4.859		

The F tests the effect of treat. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

ภาพที่ 13.3

ผลลัพธ์ที่ได้คือ

1. ในตาราง Between-Subjects Factors มีกลุ่มตัวอย่าง 3 กลุ่ม ขนาดกลุ่มตัวอย่างเท่ากับ 5 ทุกกลุ่ม

2. ผลการทดสอบความเป็นเอกพันธ์ของความแปรปรวนคือ ความแปรปรวนของประชากรแต่ละกลุ่มเท่ากัน เป็นการทดสอบความแปรปรวนของข้อมูล money เปรียบเทียบระหว่างระดับของตัวแปรอิสระ treat อยู่ในตาราง Levene's Test of Equality of Error Variances โดยใช้สถิติทดสอบ Levene statistic ได้ค่า F = 0.387 และค่า Sig. = 0.687 ซึ่งมากกว่าระดับนัยสำคัญที่กำหนด (.05) ดังนั้นจึงยอมรับ H_0 : $\sigma^2_{\text{วิทยุ}} = \sigma^2_{\text{หนังสือพิมพ์}} = \sigma^2_{\text{โทรทัศน์}}$ นั่นคือ รายได้ของร้านอาหารที่ใช้สื่อโฆษณาต่างกัน 3 แบบ มีความแปรปรวนเท่ากัน

3. ผลการทดสอบสมมติฐานเกี่ยวกับอิทธิพลของตัวแปรร่วม employee ที่มีต่อตัวแปรตาม money ดูที่ตาราง Tests of Between-Subjects Effects ในคอลัมน์ Source ที่บรรทัด EMPLOYEE ได้ค่าสถิติทดสอบ F เท่ากับ 304.729 และค่า Sig. เท่ากับ .000 ซึ่งน้อยกว่าระดับนัยสำคัญที่กำหนด (.05) ดังนั้นจึงสรุปว่าปฏิเสธ H_0: $\beta = 0$ นั่นคือ จำนวนคนงานมีอิทธิพลต่อรายได้อย่างมีนัยสำคัญทางสถิติ และสรุปได้ว่าการปรับเอาอิทธิพลของจำนวนคนงานออกแล้วช่วยลดความแปรปรวนของการทดลองได้อย่างมีนัยสำคัญทางสถิติ

ผลการทดสอบสมมติฐานเกี่ยวกับอิทธิพลของทรีทเมนต์ภายหลังการปรับเอาอิทธิพลของตัวแปรร่วม employee ออกแล้ว ดูที่บรรทัด TREAT ได้ค่าสถิติทดสอบ F เท่ากับ 51.671 และค่า Sig. เท่ากับ .000 ดังนั้นจึงสรุปว่าปฏิเสธ H_0 นั่นคือ การใช้สื่อโฆษณาทั้ง 3 แบบ ทำให้รายได้ของร้านอาหารจานด่วนแตกต่างกันอย่างมีนัยสำคัญทางสถิติภายหลังการปรับอิทธิพลของจำนวนคนงานออกแล้ว

โดยมีค่าสัมประสิทธิ์การตัดสินใจ R^2 = 0.971 หมายความว่า จำนวนคนงานและสื่อโฆษณา สามารถอธิบายความผันแปรของรายได้ของร้านอาหารได้ 97.1%

4. ในตาราง Parameter Estrmates การประมาณค่าพารามิเตอร์ในการวิเคราะห์การถดถอยที่มีตัวแปรอิสระ treat แบ่งเป็น 3 กลุ่ม เป็นตัวแปรเชิงคุณภาพ การคำนวณในโปรแกรม SPSS จะกำหนดใช้กลุ่มสุดท้ายของตัวแปร treat คือสื่อโทรทัศน์เป็นกลุ่มอ้างอิง (reference group) คือ กลุ่ม treat = 3 สำหรับบรรทัด treat = 2 แทนความแตกต่างระหว่างกลุ่ม 2 คือ สื่อหนังสือพิมพ์กับกลุ่มอ้างอิง และบรรทัด treat = 1 แทนความแตกต่างระหว่างกลุ่ม 1 คือ สื่อวิทยุกับกลุ่มอ้างอิง

ค่า B คือความแตกต่างระหว่างค่าเฉลี่ยของรายได้ของกลุ่มที่อ้างถึง เช่น treat = 1 คือ -4.111 - 21.111 = -25.222 แล้วทดสอบความแตกต่างด้วย t-test ดูที่ค่า Sig. = .000 ซึ่งน้อยกว่าระดับนัยสำคัญที่กำหนด (0.05) สรุปผลการทดสอบได้ว่า ปฏิเสธ H_0 นั่นคือ มีความแตกต่างระหว่างค่าเฉลี่ยของรายได้ของกลุ่ม 1 กับกลุ่ม 3 และมีความแตกต่างระหว่างค่าเฉลี่ยของรายได้ของกลุ่ม 2 กับกลุ่ม 3

ค่า B ของตัวแปรร่วม employee = -1.509 หมายความว่า ถ้าค่าของตัวแปรร่วม employee เพิ่มขึ้น 1 คน จะทำให้ตัวแปรตาม money ลดลง 1.509 หน่วย (50,000 บาท)

แบบฝึกหัดบทที่ 13

1. ในการศึกษาผลกระทบของการดูจากภาพถ่ายและการได้เห็นตัวจริงของผู้ทดลองต่อการประเมินความสำเร็จในชีวิตของคนคนนั้นโดยให้เป็นคะแนนจาก -10 ถึง +10 คะแนน -10 หมายถึง ประสบความล้มเหลวที่สุด และ +10 หมายถึง ประสบความสำเร็จที่สุด กลุ่มตัวอย่างคือ เจ้าหน้าที่ผู้ชาย 10 คน แบ่งออกเป็น 2 กลุ่มอย่างสุ่ม ดำเนินการทดลองโดยให้ผู้ทดลองได้ดูภาพถ่ายปกติ และดูตัวจริง เจ้าหน้าที่กลุ่มหนึ่งบ่อยเท่าที่ต้องการในระหว่างการให้คะแนน สำหรับเจ้าหน้าที่อีกกลุ่มหนึ่งผู้ทดลองได้ดูเฉพาะภาพถ่ายปกติของเจ้าหน้าที่เท่านั้น ได้ผลการให้คะแนนและอายุของเจ้าหน้าที่ดังตาราง

ตาราง อายุและคะแนนความสำเร็จของเจ้าหน้าที่ซึ่งถูกประเมินจากการดูตัวจริงและดูภาพถ่ายเท่านั้น

เจ้าหน้าที่	ดูตัวจริง		ดูภาพถ่าย	
	อายุ (x)	คะแนน (y)	อายุ (x)	คะแนน (y)
1	32	+1	58	-3
2	49	-3	27	+2
3	29	+2	43	-4
4	53	-4	30	-1
5	36	+1	31	0

แหล่งที่มา : ปรับจากข้อมูลของ Neter, J. et al (1974)

ก. จงเขียนตัวแบบสถิติของการวิเคราะห์ความแปรปรวนร่วมของการทดลองนี้ โดยกำหนดให้อายุเป็นตัวแปรร่วม

ข. จงตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเท่ากันของความชันของเส้นถดถอยทุกกลุ่มของทรีทเมนต์ที่ระดับนัยสำคัญ $\alpha = .01$ และสรุปผลการวิเคราะห์

ค. จงทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของทรีทเมนต์ที่ได้ปรับอิทธิพลของตัวแปรร่วม x ออกไปแล้ว

2. ในการศึกษาอิทธิพลของวิธีการแสดงสินค้าที่มีผลต่อการขายสินค้านั้น กลุ่มตัวอย่างคือ ร้านค้าที่มีคุณลักษณะต่าง ๆ ของร้านเหมือนกัน 15 ร้าน แบ่งออกเป็น 3 กลุ่มอย่างสุ่ม แต่ละกลุ่มจะได้รับทรีทเมนต์หนึ่งเป็นวิธีการแสดงสินค้าแบบหนึ่ง ผู้วิจัยสนใจศึกษาวิธีการแสดงสินค้า 3 แบบคือ (1) จัดเคาน์เตอร์แสดงสินค้าที่เดิม (2) จัดแสดงสินค้าที่ด้านหน้าของร้าน และ (3) จัด ทั้ง 2 แห่ง คือ ที่เคาน์เตอร์แสดงสินค้าที่เดิม และที่หน้าร้านด้วย เก็บข้อมูลการขายในช่วง 3 สัปดาห์แรก แล้วจัดแสดงสินค้าที่ออกใหม่อีก เก็บข้อมูลการขายในช่วง 3 สัปดาห์ที่สอง ได้ข้อมูลดังตาราง

ตาราง ข้อมูลการขายสินค้าในร้านที่จัดแสดงสินค้าใหม่ ในรูปแบบแตกต่างกัน 3 แบบ

การจัดสินค้า	3 สัปดาห์	ร้าน				
		1	2	3	4	5
แบบที่ 1	1 (x)	67	74	52	68	92
	2 (y)	43	58	38	57	70
แบบที่ 2	1 (x)	81	69	73	77	80
	2 (y)	75	73	78	74	82
แบบที่ 3	1 (x)	42	65	81	73	69
	2 (y)	49	63	84	75	76

แหล่งที่มา : Neter, J. et al (1974)

ก. จงเขียนตัวแบบสถิติของการวิเคราะห์ความแปรปรวนร่วมของการทดลองนี้ โดยกำหนดให้การขายใน 3 สัปดาห์แรกเป็นตัวแปรร่วม

ข. จงตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเท่ากันของความชันของเส้นถดถอยทุกกลุ่มของทรีทเมนต์ที่ระดับนัยสำคัญ $\alpha = .01$ และสรุปผลการวิเคราะห์

ค. จงทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของทรีทเมนต์ที่ได้ปรับอิทธิพลของตัวแปรร่วม x ออกไปแล้ว

3. ในการศึกษาอิทธิพลของการออกกำลังกายที่มีต่อการระบายออกซิเจน (oxygen ventilation) ทรีทเมนต์คือ การออกกำลังกาย 2 แบบคือ แอโรบิก และการวิ่ง กลุ่มตัวอย่างคือ ผู้ชายสุขภาพดี 12 คน แบ่งออกเป็น 2 กลุ่มอย่างสุ่ม วัดระบบการหายใจของกลุ่มตัวอย่างก่อนการทดลองได้ผลไม่แตกต่างกันทั้ง 2 กลุ่ม ดำเนินการทดลองโดยให้กลุ่มหนึ่งออกกำลังกายแบบแอโรบิก อีกกลุ่มหนึ่งออกกำลังกายแบบวิ่งบนพื้นราบ ทำการทดลองเป็นเวลา 12 สัปดาห์ เมื่อสิ้นสุดการทดลองวัดความเปลี่ยนแปลงของการระบายอากาศที่มากที่สุด (maximal ventilation change) ของออกซิเจนเป็น liters/minutes ได้ข้อมูลดังตาราง

ตาราง ข้อมูลความเปลี่ยนแปลง (maximal ventilation change) วัดเป็น liters/minute หลังจากได้รับโปรแกรมการออกกำลังกายเป็นเวลา 12 สัปดาห์

คนที่	แอโรบิก		การวิ่ง	
	อายุ (x)	ความเปลี่ยนแปลง (y)	อายุ (x)	ความเปลี่ยนแปลง (y)
1	31	17.05	23	- 0.87
2	23	4.96	22	- 10.74
3	27	10.40	22	- 3.27
4	28	11.05	25	- 1.97
5	22	0.26	27	7.50
6	24	2.51	20	- 7.25

แหล่งที่มา : Allen, D. Exercise Physiology, University of Arisona อ้างถึงใน Kuehl, O.R. 1994.

ก. จงพล็อตกราฟการกระจายระหว่าง y และ x ในการออกกำลังกายทั้งแบบแอโรบิกและการวิ่ง แล้วพิจารณารูปแบบของความสัมพันธ์ระหว่าง y และ x ของการออกกำลังกายทั้ง 2 แบบเหมือนกันหรือไม่

ข. จงเขียนตัวแบบสถิติของการวิเคราะห์ความแปรปรวนร่วม ที่มี y เป็นตัวแปรตาม และ x เป็นตัวแปรร่วม และสมมติให้ y และ x มีความสัมพันธ์กันแบบเชิงเส้น

ค. จงตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเหมือนกันของความชันของเส้นถดถอยสำหรับความเปลี่ยนแปลง (y) ที่ถดถอยบนอายุ (x) ของกลุ่มที่ออกกำลังกายแบบแอโรบิก และกลุ่มที่ออกกำลังกายแบบวิ่ง

ง. จงทดสอบนัยสำคัญของอิทธิพลของตัวแปรร่วม (x) ที่มีต่อตัวแปรตาม (y) และสรุปผลการทดสอบ

จ. จงทดสอบนัยสำคัญของอิทธิพลของทรีทเมนต์ภายหลังการปรับเอาอิทธิพลของตัวแปรร่วม (x) ออกแล้ว และสรุปผลการทดสอบ