บทที่ 10
การวิเคราะห์ความแปรปรวนร่วม
(Analysis of Coveriance)

ความจำเป็นของการใช้การวิเคราะห์ความแปรปรวนร่วม

ลองพิจารณาตัวอย่างงานทดลองต่อไปนี้

งานทดลองที่ 1 – เป็นการเปรียบเทียบสูตรอาหารที่ใช้เลี้ยงหมู สูตรที่ดีคือสูตรที่ให้น้ำหนักเพิ่มมาก ๆ หลังการทดลองระยะเวลาหนึ่ง สูตรอาหารเหล่านั้นอาจให้ผลการทดสอบว่าความแตกต่างมีนัยสำคัญหรืออาจไม่มีนัยสำคัญ ก็ได้ อย่างไรก็ตาม ผู้ทดลองอาจฉุกคิดหรือตั้งปัญหาในใจได้ว่า หากเขาใช้หมูที่มีน้ำหนักต่างกันตอนเริ่มทำการทดลอง เขาจะทราบได้อย่างไรว่าความแตกต่างของวิธีการไม่ได้รวมความแตกต่างของน้ำหนักก่อนเริ่มทำการทดลองไว้ด้วย หากกำหนดให้น้ำหนักของหมูแต่ละตัวก่อนเริ่มการทดลองเป็น X ส่วน Y คือน้ำหนักเมื่อได้รับวิธีการแล้ว จะทำอย่างไรจึงจะปรับปรุงค่า Y สำหรับค่า X ที่ต่างกันได้ คำตอบคือต้องใช้วิธีการวิเคราะห์ความแปรปรวนร่วม

งานทดลองที่ 2 – เป็นการเปรียบเทียบวิธีการต่าง ๆ ในการสอนวิชาสถิติ โดยวัดคะแนน Y จากนักเรียนแต่ละคนภายใต้วิธีการต่าง ๆ แต่ให้สอบข้อสอบชุดเดียวกัน จะแน่ใจได้อย่างไรว่าความแตกต่างของคะแนนสอบภายใต้วิธีการต่าง ๆ ไม่ได้รวมความแตกต่างของระดับ I.Q. (X) ของเด็กนักเรียนแต่ละคนเข้าไปด้วย คำตอบคือต้องปรับปรุงค่า Y สำหรับค่า X โดยวิธีการวิเคราะห์ความแปรปรวนร่วมกัน

ข้อสมมุติ และ แบบจำลอง

จะใช้ข้อสมมุติที่จำเป็นสำหรับการวิเคราะห์ความแปรปรวนและความถดถอย กล่าวคือ มีการแจกแจงแบบอิสระกัน แบบปกติ และมีความแปรปรวนไม่ต่างกัน ส่วน X เป็นตัวแปรที่สามารถกำหนดค่าล่วงหน้าได้ และมีแบบจำลอง ดังนี้

แบบแผนงานทดลองแบบสุ่มสมบูรณ์

$$Y_{ij} = \xi + \tau_i + \beta X_{ij} + \varepsilon_{ij} \quad ; \quad \begin{matrix} i = 1,2,\ldots,t \\ j = 1,2,\ldots,r \end{matrix}$$

แผนงานทดลองแบบบล็อคสมบูรณ์

$$Y_{ij} = \xi + \tau_i + \rho_j + \beta X_{ij} + \varepsilon_{ij} \quad ; \quad \begin{matrix} i = 1,2,\ldots,t \\ j = 1,2,\ldots,r \end{matrix}$$

แผนงานทดลองแบบจัตุรัสลาติน

$$Y_{ij} = \xi + \rho_i + \gamma_j + \tau_k + \beta X_{ij(k)} + \varepsilon_{ij(k)} \quad ; \begin{matrix} i = 1,2,\ldots,r \\ j = 1,2,\ldots,r \\ k = 1,2,\ldots,r \end{matrix}$$

แผนงานทดลองแบบแฟคทอเรียลที่มี 2 แฟคเตอร์ ในบล็อคสมบูรณ์

$$Y_{ij} = \xi + \rho_i + \alpha_j + \gamma_k + (\alpha\gamma)_{jk} + \beta X_{ijk} + \varepsilon_{ijk} \quad ; \begin{matrix} i = 1,2,\ldots,r \\ j = 1,2,\ldots,a \\ k = 1,2,\ldots,c \end{matrix}$$

แต่ในทางปฏิบัติ มักจะใช้แบบจำลองในรูปส่วนเบี่ยงเบนของ X จากค่าเฉลี่ยดังนั้น จะได้แบบจำลองในรูปส่วนเบี่ยงเบนจากค่าเฉลี่ยของ X ดังนี้

แบบสุ่มสมบูรณ์

$$Y_{ij} = \mu + \tau_i + \beta(X_{ij} - \bar{X}) + \varepsilon_{ij} \quad ; \begin{matrix} i = 1,2,\ldots,t \\ j = 1,2,\ldots,r \end{matrix}$$

แบบบล็อคสมบูรณ์

$$Y_{ij} = \mu + \tau_i + \rho_j + \beta(X_{ij} - \bar{X}) + \varepsilon_{ij} \quad \begin{matrix} i = 1,2,\ldots,t \\ j = 1,2,\ldots,r \end{matrix}$$

แบบจัตุรัสลาติน

$$Y_{ij} = \mu + \rho_i + Y_j + \tau_k + \beta(X_{ij(k)} - \bar{X}) + \varepsilon_{ij(k)} \quad ; \begin{matrix} i = 1,2,\ldots,r \\ j = 1,2,\ldots,r \\ k = 1,2,\ldots,r \end{matrix}$$

แบบแฟคทอเรียลที่มี 2 แฟคเตอร์ในบล็อคสมบูรณ์

$$Y_{ijk} = \mu + \rho_i + \alpha_j + \gamma_k + (\alpha\gamma)_{jk} + \beta(X_{ijk} - \bar{X}) + \varepsilon_{ijk}$$

$i = 1,\ldots,r \; ; \; j = 1,\ldots,a \; ; \; k = 1,\ldots,c$

โดยที่ $\mu = \xi + \beta\bar{X}$

แผนงานทดลองแบบสุ่มสมบูรณ์

ต่อไปจะใช้สัญญลักษณ์ต่าง ๆ ต่อไปนี้

Σx^2 = corrected total sum of squares for X

$$= \sum_i \sum_j X_{ij}^2 - (\sum_i \sum_j X_{ij})^2/N$$

Σxy = corrected total sum of products for X and Y

$$= \sum_i \sum_j X_{ij} Y_{ij} - (\sum_i \sum_j X_{ij})(\sum\sum_{ij} Y_{ij})$$

Σy^2 = corrected total sum of squares for Y

$$= \sum_i \sum_j Y_{ij}^2 - (\sum_i \sum_j Y_{ij})^2/N$$

T_{xx} = treatment sum of squares for X

$$= \sum_i \frac{(\sum_j X_{ij})^2}{r_i} - (\sum_i \sum_j X_{ij})^2/N$$

T_{xy} = treatment sum of products for X and Y

$$= \sum_i \frac{(\sum_j X_{ij})(\sum_j Y_{ij})}{r_i} - \frac{(\sum_i \sum_j X_{ij})(\sum_i \sum_j Y_{ij})}{N}$$

T_{yy} = treatment sum of squares for Y

$$= \sum_i \frac{(\sum_j Y_{ij})^2}{r_i} - \frac{(\sum_i \sum_j Y_{ij})^2}{N}$$

E_{xx} = experimental error sum of squares for X

$$= \Sigma X^2 - T_{xx}$$

E_{xy} = experimental error sum or products for X and Y

$$= \Sigma xy - T_{xy}$$

E_{yy} = experimental error sum of squares for Y

$$= \Sigma y^2 - T_{yy}$$

ตัวสถิติที่ใช้ทดสอบความแตกต่างของวิธีการภายหลังจากการปรับปรุงอิทธิพลของ X คือ

$$F = \frac{(S_{T+E} - S_E)/(t-1)}{S_E^2} ; \quad \nu_1 = t - 1, \quad \nu_2 = N - t - 1$$

ส่วนค่าเฉลี่ยของวิธีการที่ได้ปรับปรุงแล้ว คือ

$$\text{adj. } \bar{Y}_i = \bar{Y}_i - b(\bar{X}_i - \bar{X}) ; \quad i = 1,\ldots,t$$

b = regression coefficient = E_{xy}/E_{xx}

$$\hat{V}\ \text{adj.}\ \bar{Y}_i\ ) \;=\; S_E^2\ \ (1/r_i + \ (\bar{X}_i - \bar{X}^2)/E_{xx})$$

ทั้งนี้เนื่องจากในแบบจำลอง ได้มีข้อสมมุติว่า สัมประสิทธิ์ความถดถอย คือ $\beta = 0$ เพราะถ้า $\beta = 0$ การใช้ X มาปรับปรุงอิทธิพลของวิธีการจะไม่มีประโยชน์อันใดเลย ดังนั้น หากต้องการตรวจสอบว่าข้อสมมุตินี้เป็นจริงหรือไม่ ทำได้ดังนี้

$$Ho\ :\ \beta = 0\ ,\quad Ha\ :\ \beta \neq 0$$

$$F \;=\; \frac{E_{xy}^2\ /\ E_{xx}}{S_E\ /\ (\ N - t - 1)} \;=\; \frac{E_{xy}^2\ /\ E_{xx}}{S_E^2}$$

$$\nu_1 = 1\ ,\quad \nu_2 = N - t - 1$$

General Analysis of Covariance for data from a CRD

Source of Variation	df.	Sum of Squares and Products			Deviations about Regression		
		Σx^2	Σxy	Σy^2	$\Sigma y^2 - (\Sigma xy)^2/\Sigma x^2$	df	Mean Square
Among treatments	t - 1	T_{xx}	T_{xy}	T_{yy}	.....	.....	.....
Within treatments	N - t	E_{xx}	E_{xy}	E_{yy}	$S_E = E_{yy} - E_{xy}^2/E_{xx}$	N-t-1	$S_E^2 = \frac{S_E}{N-t-1}$
Among treatments + Within treatments (= total)	N - 1	$S_{xx} = T_{xx}+E_{xx}$	$S_{xy} = T_{xy}+E_{xy}$	$S_{yy} = T_{yy}+E_{yy}$	$S_{T+E} = S_{yy} - S_{xy}^2/S_{xx}$	N - 2	
Difference for testing among adjusted treatment means					$S_{T+E} - S_E = T_{yy} - \frac{S_{xy}^2}{S_{xx}} + E_{xy}^2/E_{xx}$	t - 1	$\frac{(S_{T+E} - S_E)}{(t - 1)}$

ตัวอย่าง

น้ำหนักที่เพิ่มขึ้น (Y) และน้ำหนักสุดท้ายขณะเริ่มการทดลอง (X) ของหมูจากการทดลองเรื่องอาหาร

	วิธีการ							
	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
	30	165	24	180	34	156	41	201
	27	170	31	169	32	189	32	173
	20	130	20	171	35	138	30	200
	21	156	26	161	35	190	35	193
	33	167	20	180	30	160	28	142
	29	151	25	170	29	172	36	189
total	160	939	146	1031	195	1005	202	1098

Analysis of Covariance

Source of Variation	df.	Sum of squares and products			Deviations about Regression		
		Σx^2	Σxy	Σy^2	$\Sigma y^2 - \frac{(\Sigma xy)^2}{\Sigma x^2}$	df.	Mean Square
Among treatments	3	365.46	451.21	2163.13	.....	.....	.....
Within treatments	20	361.50	496.83	5937.83	5255.01	19	276.58
Total	23	726.96	948.04	8100.96	6864.61	22	.....
Difference for testing among adjusted treatment means					1609.60	3	536.53

$\Sigma X^2 = (30^2 + ... + 36^2) - (703)^2/24 = 726.96$

$\Sigma xy = (30)(165) + ... + (36)(189) - (703)(4073)/24 = 948.04$

$\Sigma y^2 = (165^2 + ... + 189^2) - (4073)^2/24 = 3100.96$

$T_{xx} = (160^2 + 146^2 + 195^2 + 202^2)/6 - (703)^2/24 = 365.46$

$T_{xy} = 1/6\{(160)(939) + (146)(1031) + (195)(1005) + (202)(1098)\}$

$- (703)(4073)/24 = 451.21$

$$T_{yy} = 1/6(939^2 + 1031^2 + 1005^2 + 1098^2) - (4073)^2/24 = 2163.13$$

$$E_{xx} = S_{xx} - T_{xx} = 361.50$$

$$E_{xy} = S_{xy} - T_{xy} = 496.83$$

$$E_{yy} = S_{yy} - T_{yy} = 5937.83$$

ตัวสถิติที่ใช้ทดสอบความแตกต่างของวิธีการที่ปรับปรุงแล้ว คือ

$$F = 536.53/276.58 = 1.94 \; ; \quad \nu_1 = 3 \; , \; \nu_2 = 19$$

สรุปว่าอิทธิพลของวิธีการทั้ง 4 วิธีการไม่ต่างกันอย่างมีนัยสำคัญ

สำหรับตัวอย่างนี้ ปรากฏว่าผลการทดสอบของวิธีการภายหลังการปรับปรุง และ ก่อนการปรับปรุงให้ผลสรุปแบบเดียวกัน แต่มีหลายกรณีที่ภายหลังจากการใช้เทคนิคของการวิเคราะห์ความแปรปรวนร่วมกันแล้วจะให้ผลสรุปที่แตกต่างกัน

แสดงการคำนวณค่าเฉลี่ยวิธีการที่ปรับปรุงแล้ว

($\overline{X}$ = 29.29, $\overline{Y}$ = 169.71, b = 496.83/361.50 = 1.374)

วิธีการ	1	2	3	4
$\overline{X}_i$	26.67	24.33	32.50	33.67
$\overline{X}_i - \overline{X}$	− 2.62	− 4.96	3.21	4.38
b $(\overline{X}_i - \overline{X})$	− 3.60	− 6.82	4.41	6.02
Y_i	156.50	171.83	167.50	183.00
adj. $\overline{Y}_i$	160.10	178.65	163.09	176.98
Standard error of adjusted $\overline{Y}_i$	7.17	8.06	7.35	7.80

การวิเคราะห์ความแปรปรวนร่วม เมื่อใช้แผนงานทดลองแบบบล็อคสมบูรณ์

ตัวอย่าง ข้อมูลในตารางข้างล่างคือผลผลิตของพืช 3 สายพันธุ์ ในการทดลองแบบบล็อคสมบูรณ์ที่มี 4 บล็อค โดยที่ X คือผลผลิตจากแปลงทดลองก่อนเริ่มทำการทดลอง และ Y คือ ผลผลิตจากแปลงเดิมในระหว่างการทดลองที่ใช้พืช 3 สายพันธุ์

Block		สายพันธุ์ A	B	C	total
1	X	54	51	57	162
	Y	64	65	72	201
2	X	62	64	60	186
	Y	68	69	70	207
3	X	51	47	46	144
	Y	54	60	57	171
4	X	53	50	41	144
	Y	62	66	61	189
Total	X	220	212	204	636
	Y	248	260	260	768

$\Sigma X^2 = (54^2 + \ldots + 41^2) - (636)^2/12 = 514$

$R_{XX} = (162^2 + 186^2 + 144^2 + 144^2)/3 - (636)^2/12 = 396$

$T_{xx} = (220^2 + 212^2 + 204^2)/4 - (636)^2/12 = 32$

$E_{xx} = 514 - 396 - 32 = 86$

$\Sigma y^2 = (64^2 + \ldots + 61^2) - (768)^2/12 = 324$

$R_{yy} = (201^2 + 207^2 + 171^2 + 189^2)/3 - (768)^2/12 = 252$

$T_{yy} = (248^2 + 260^2 + 260^2)/4 - (768)^2/12 = 24$

$E_{yy} = 324 - 252 - 24 = 48$

$\Sigma xy = \{(54)(64) + \ldots + (41)(61)\} - (636)(768)/12 = 286$

$R_{xy} = \{(162)(201) + (186)(207) + \ldots (144)(189)\}/3 - (636)(768)/12$
$= 264$

$T_{xy} = \{(220)(248) + (212)(260) + (204)(260)\}/3 = (636)(768)/12$
$= -24$

$E_{xy} = 286 - 264 - (-24) = 46$

Source of Variation	df.	Sum of squares and products			Deviation about Regression		
		Σx^2	Σxy	Σy^2	$\Sigma y^2 - (\Sigma xy)^2/\Sigma x^2$	df.	Mean Square
Replicates (blocks)	3	396	264	252	.....	.....	.....
Treatments	2	32	−24	24	.....	.....	.....
Experimental error	6	86	46	48	23.4	5	4.68
Treatment + error	8	118	22	72	67.9	7	.....
Diffence for testing adj. treatment means					44.5	2	22.25

- เพื่อจะทดสอบว่าควรใช้เทคนิคของการวิเคราะห์ความแปรปรวนร่วมหรือไม่ มีวิธีการ ดังนี้

$$Ho : \beta = 0 \quad , \quad Ha : \beta \neq 0$$

$$F = \frac{(46)^2/86}{4.68} = 5.26 \quad ; \quad f_{1,5(.05)} = 6.61$$

สรุปว่า X ไม่มีอิทธิพลในการกำหนดค่า Y ดังนั้นจึงยังไม่ควรใช้การวิเคราะห์ความแปรปรวนร่วม.

- เพื่อจะทดสอบว่าก่อนการใช้การวิเคราะห์ความแปรปรวนร่วม อิทธิพลของวิธีการเป็นอย่างไร ?

$$F = (24/2) / (48/6) = 1.5 \; ; \; f_{2,6\,(.05)} = 5.14$$

สรุปว่าความแตกต่างของวิธีการไม่มีนัยสำคัญ

- เพื่อจะทดสอบว่าอิทธิพลของวิธีการภายหลังที่ได้ปรับปรุงแล้วมีความแตกต่างกันหรือไม่ ?

$$F = 22.25/4.68 = 4.75 \; ; \; f_{2,5(.10)} = 3.78 ; f_{2,5(.05)} = 5.79$$

ถ้าใช้ระดับนัยสำคัญ .05 จะยังปฏิเสธ Ho ไม่ได้เช่นเดียวกัน แต่ค่าสถิติที่คำนวณได้เกือบจะปฏิเสธ Ho ได้ ดังนั้น หากเพิ่มจำนวนซ้ำให้มากกว่าเดิม อาจจะปฏิเสธ Ho ได้ก็ได้

การใช้ความแปรปรวนร่วมในแผนงานทดลองแบบจัตุรัสลาติน

มีหลักการคล้ายกับการวิเคราะห์ความแปรปรวนร่วมในการวางแผนแบบสุ่มสมบูรณ์และแบบบล็อคสมบูรณ์ และตารางวิเคราะห์ความแปรปรวนร่วม มีดังนี้

Source of Variation	Sum of squares and products					Deviations about Regression	
	df.	Σx^2	Σxy	Σy^2	$\Sigma y^2 - \frac{(\Sigma xy)^2}{\Sigma x^2}$	df	Mean square
Rows	r - 1	R_{xx}	R_{xy}	R_{yy}	.....	.....	.....
Columns	r - 1	C_{xx}	C_{xy}	C_{yy}	.....	.....	.....
Treatments	r - 1	T_{xx}	T_{xy}	T_{yy}	.....	.....	.....
Experimental error	(r-1)(r-2)	E_{xx}	E_{xy}	E_{yy}	$S_E = E_{yy} - \frac{E_{xy}^2}{E_{xx}}$	(r-1)(r-2)	$S_E^2 = \frac{S_E}{(r-1)(r-2)-1}$
Treatment + Error	$(r-1)^2$	$S_{xx} = T_{xx} + E_{xx}$	$S_{xy} = T_{xy} + E_{xy}$	$S_{yy} = T_{yy} + E_{yy}$	$S_{T+E} = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$	$(r-1)^2-1$	
Difference for testing adj.trt. means					$S_{T+E} - S_E$	(r-1)	$(S_{T+E} - S_E)/(r-1)$

$$Ho : \beta = 0 \quad ; \quad Ha : \beta \neq 0$$

$$F = \frac{E_{xy}^2/E_{xx}}{S_E^2} \qquad \nu_1 = 1 \; ; \; \nu_2 = (r-1)(r-2)-1$$

Test for adjusted treatment means :

$$F = \frac{(S_{T+E} - S_E)/(r-1)}{S_E^2} \qquad \nu_1 = r-1$$

$$\nu_2 = (r-1)(r-2)-1$$

การวิเคราะห์ความแปรปรวนร่วมของ 2 แฟคเตอร์ในบล็อคสมบูรณ์

ตัวอย่าง เป็นข้อมูลจากการเลี้ยงหมูด้วยสูตรอาหาร 3 สูตร และ 2 เพศ คือ หมูเพศผู้และเมียโดยเลี้ยงในคอกที่แบ่งแยกต่างหากจากกัน 5 คอก (เรพปริเคท) X คือน้ำหนักลูกหมูตอนเริ่มทำการทดลอง Y คือน้ำหนักที่เพิ่มภายหลังได้รับวิธีการต่าง ๆ แล้ว

คอกหมู (เรพปริเคท)		อาหาร						Totals
		A		B		C		
		ผู้	เมีย	ผู้	เมีย	ผู้	เมีย	
1	X	38	48	39	48	48	48	269
	Y	9.52	9.94	8.51	10.00	9.11	9.75	56.83
2	X	35	32	38	32	37	28	202
	Y	8.21	9.48	9.95	9.24	8.50	8.66	54.04
3	X	41	35	46	41	42	33	238
	Y	9.32	9.32	8.43	9.34	8.90	7.63	52.94
4	X	48	46	40	46	42	50	272
	Y	10.56	10.90	8.86	9.68	9.51	10.37	59.88
5	X	43	32	40	37	40	30	222
	Y	10.42	8.82	9.20	9.67	8.76	8.57	55.44
Totals	X	205	193	203	204	209	189	1203
	Y	48.03	48.46	44.95	47.93	44.78	44.98	279.13

Source of Variation	df.	Σx^2	Σxy	Σy^2	$\Sigma y^2 - \frac{(\Sigma xy)^2}{\Sigma x^2}$	df.	Mean Square
Replicates	4	605.87	39.905	4.8518	.....	.....	.....
Treatments							
A : อาหาร	2	5.40	−0.147	2.2686	.....	.....	.....
B : เพศ	1	32.03	−3.730	0.4344	.....	.....	.....
A × B	2	22.47	3.112	0.4761	.....	.....	.....
Error	20	442.93	39.367	8.3144	4.8155	19	0.2534
A + Error	22	448.33	39.220	10.5830	7.1520	21	
Difference for testing among adj. food means..					2.3365	2	1.16825
B + Error	21	474.96	35.637	8.7488	6.0749	20	.....
Difference for testing among adj. sex means..					1.2594	1	1.2594
AB + Error	22	465.40	42.479	8.7905	4.9133	21	.....
Difference for testing among food × sex effects.					.0978	2	0.0489

$$Ho : \beta - 0 \quad ; \quad Ha : \beta \neq 0$$

$$F = \frac{(39.367)^2 / 442.93}{0.2534} = 13.81^{**}$$

$$f_{1,19(.01)} = 8.1$$

สรุปว่า X มีอิทธิพลต่อการเปลี่ยนแปลงของ Y

การทดสอบความแตกต่างของอิทธิพลต่าง ๆ ของแฟคเตอร์ มีดังนี้.

ทดสอบอิทธิพลของอาหาร (ที่ปรับปรุงแล้ว) :

$$F = 1.16825/0.2534 = 4.61$$

ทดสอบอิทธิพลของเพศ (ที่ปรับปรุงแล้ว) :

$$F = 1.2594/0.2534 = 4.97$$

ทดสอบอิทธิพลร่วมกันของอาหารและเพศ (ซึ่งได้ปรับปรุงแล้ว) :

$$F = 0.0489/0.2534 = 0.19$$

cccovariancecccovariancecccovariancecc

แบบฝึกหัด

กำหนดให้ข้อมูลจากงานทดลองหนึ่งมีดังนี้

1. กำหนดให้ข้อมูลจากงานทดลองหนึ่งมีดังนี้

Source of Variation	**df.**	Σx^2	Σxy	Σy^2
Replicates	5	200	600	4000
Treatments	5	100	200	2500
Experimental error	25	300	1200	7500

ก. ความถดถอยของ Y บน X มีนัยสำคัญหรือไม่ ที่ระดับนัยสำคัญ 5%

ข. ความแตกต่างของค่าเฉลี่ยวิธีการ Y ซึ่งได้ปรับปรุงแล้ว มีนัยสำคัญหรือไม่ที่ระดับ 5%

2. ทดลองปลูกถั่วเหลือง 10 พันธุ์ ในแผนงานทดลองแบบบล็อคสมบูรณ์ที่มี 4 เรพปริเคท และได้ทดสอบความแตกต่างของวิธีการ พบว่าไม่มีนัยสำคัญ มีผู้ให้ข้อสังเกตว่าความทนทานต่อศัตรูพืช (X) ของถั่วแต่ละพันธุ์ไม่เท่ากัน อาจมีผลต่อผลผลิต (Y) ได้ และมีข้อมูลเพิ่มเติมดังนี้

Source of Variation	**df.**	Σx^2	Σxy	Σy^2
พันธุ์ถั่วเหลือง	9	4684	−532	112
Error	27	3317	−650	216

ก. จงแสดงการทดสอบความแตกต่างของผลผลิตภายหลังการปรับปรุงสำหรับแรงต้านทานโรค

ข. จงหาส่วนของความผันแปรของผลผลิตที่ไม่สามารถอธิบายด้วยความถดถอย

3. งานทดลองเกี่ยวกับการปรับสภาพดิน 9 วิธีการ เพื่อใช้ปลูกพืชชนิดหนึ่ง ให้ X คือน้ำหนักของเมล็ดพันธุ์พืช Y คือน้ำหนักของผลผลิตจากเมล็ดพันธุ์นั้นภายหลังปลูก 3 สัปดาห์ ได้ข้อมูลสรุป ดังนี้

Source of Variation	**df.**	Σx^2	Σxy	Σy^2
Replicates	5	4	16	96
Treatments	8	16	32	80
Error	40	20	40	160

จงเติมตารางให้สมบูรณ์ และทำการทดสอบตามที่เห็นสมควร

4. ตารางวิเคราะห์ข้างล่างได้จากการศึกษารายได้จากผลผลิตของชาวไร่ ชาวไร่จำแนกเป็น 2 ประเภท คือ เป็นเจ้าของที่ดิน และ เช่าจากผู้อื่น และจำแนกตามขนาดของพื้นที่กสิกรรม ให้ X เป็นขนาดของพื้นที่กสิกรรม และ Y เป็นผลผลิตจากพื้นที่กสิกรรม

ก. ผลสรุปการทดสอบความแตกต่างของวิธีการก่อนการปรับปรุง และภายหลังการปรับปรุงสำหรับขนาดของพื้นที่ มีความแตกต่างกันหรือไม่

ข. ขนาดของพื้นที่มีอิทธิพลต่อความผันแปรของผลผลิตหรือไม่

Source of Variation	**df.**	Σx^2	Σxy	Σy^2
Replicates (พื้นที่)	4	20000	14010	13600
ชาวไร่	1	61000	13260	4200
Interaction	4	4000	−13270	6100
ภายในกลุ่มย่อย	50	40000	19000	12700
Total	59	125000	33000	36600

5. จงวิเคราะห์ข้อมูลข้างล่าง โดยที่ X คือขนาดของฟาร์ม และ Y คือผลผลิตจากฟาร์ม ฟาร์มแบ่งเป็น 2 แบบคือ เป็นเจ้าของ และเช่าจากผู้อื่น

ก. จงทดสอบ $H_o : \beta = 0 \; ; \; H_a : \beta \neq 0$

ข. จงทดสอบความแตกต่างของวิธีการที่ได้ปรับปรุงแล้ว

Block 1			Block 2			Block 3		
ฟาร์มเลขที่	Y เป็นเจ้าของ	X	ฟาร์มเลขที่	Y เป็นเจ้าของ	X	ฟาร์มเลขที่	Y เป็นเจ้าของ	X
22	6309	160	27	2490	90	17	4489	120
13	8456	320	24	5349	154	25	10026	245
20	8453	200	11	5518	160	1	5659	160
8	4891	160	34	10417	234	26	5475	160
21	3491	120	38	4278	120	4	11382	320
	เป็นผู้เช่า			เป็นผู้เช่า			เป็นผู้เช่า	
31	6944	160	13	4936	160	20	5731	160
30	6971	160	1	7376	200	15	6787	173
11	4053	120	19	6216	160	7	5814	134
6	8767	280	32	10313	240	5	9607	239
16	6765	160	28	5124	120	25	9817	320